Master of Turbulence

EC Lille & ENSMA Poitiers

& ENSI Poitiers

Master 2 - Year 2022-2023

Course of Signal Processing

D. Marx, Institut PPRIME

Download the course from:
https://pprime.fr/marx-david/

1 Introduction

This course presents some signal processing techniques that can be of interest for processing data in fluid mechanics and acoustics.

Several types of signals

The question of what a signal is should be answered first. Generally speaking a signal is a random process $x(t, \xi)$. This definition involves two variables. The variable t represents time and the variable ξ represents an outcome number which is made necessary by the fact that time traces obtained for different outcomes are not the same for a random process. If the signal x is a white noise, both variables t and ξ are necessary. If the signal is not random (it is then said to be deterministic) the variable ξ is not necessary anymore and the signal becomes a usual function $x(t)$ depending solely on time t. Hence, the first thing to know about a signal is probably whether it is random or not. A classification of signals is presented in Fig. 1.1; it follows that given in [3, chapter 1]. For deterministic signals a further distinction may be made between periodic and

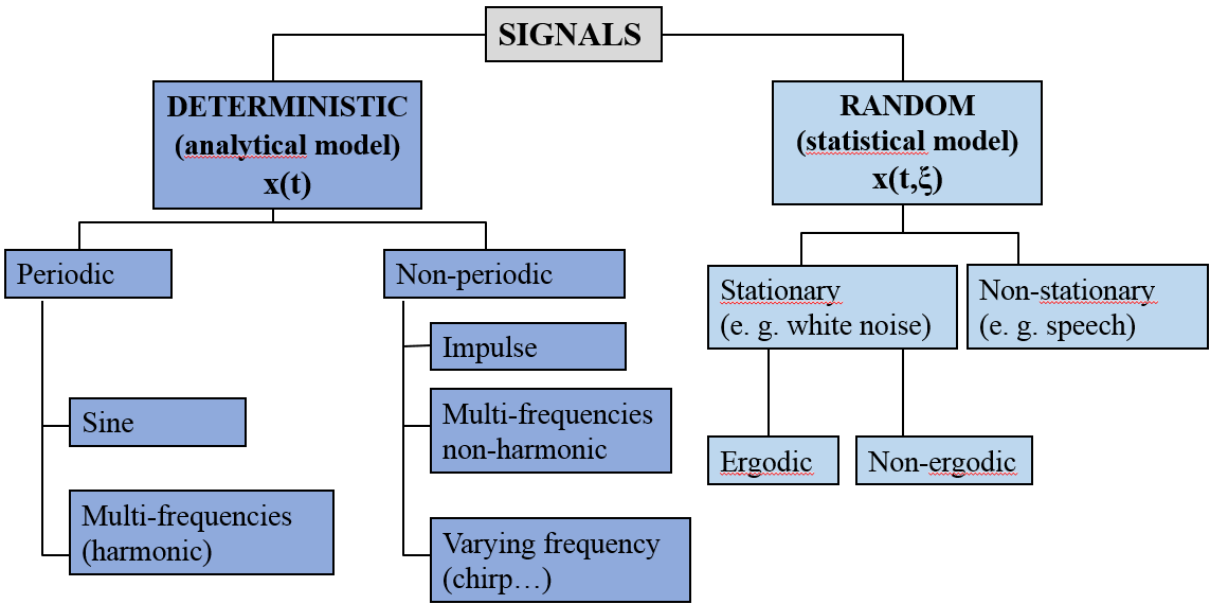


Figure 1.1: Several types of signals.

non-periodic signals. For random signals a similar distinction is made between stationary and non-stationary signals. The goal of signal processing is to extract useful information from these signals. Ideally, one should use a method that is well suited to the particular signal considered.

Signal representation
In this course signal processing will mainly consist of breaking down the signal into pieces. A signal $x(t)$ will be represented as a sum of other signals $x_n(t)$ according to an expansion such as:

$$x(t) = \sum_n a_n x_n(t) \tag{1.1}$$

Of course, such an expansion (or decomposition) is useful only if the pieces $x_n(t)$ have some useful properties. This encompasses nice mathematical properties (for example, the set $\{x_n(t)\}$ can be an orthogonal basis of the considered functional space) or physical ones (nice physical interpretation).

In producing the above expansion there are two underlying issues. The first is to find the set $\{x_n(t)\}$ (or basis, or family, or dictionary) that is relevant to a given type of signal. Some bases are nice for periodic or stationary signals (Fourier), other bases are good for transient signals (wavelets). Some bases are nice to compress a signal (cosines in jpeg image compression, wavelets). Some bases have a good connection with the underlying physics (POD). Some bases are fixed in advance (Fourier), others depend on the signal itself (POD). Some bases are orthogonal, some other are not (redundancy in the continuous wavelet transform).

The second issue, once a basis has been chosen, is to obtain a recipe for expanding the original function into this basis, that is, a recipe for obtaining the coefficients a_n. Finding the coefficients is often equivalent to performing some transform adapted to the considered basis: the Fourier transform, the Proper Orthogonal Decomposition, the wavelet transform, etc. Very often these transforms are or resemble a scalar product between the signal and the basis functions.

When do two signals resemble each other ?
A related question that arises frequently is whether a signal resembles another signal (or also whether a part of a signal resembles a member of the basis). This is so because a nice basis is usually a basis that resembles the signal in some sense. An answer to that question can be given provided the signal is in the right mathematical space. In particular, when the considered space is an Hilbert space an inner product (or scalar product) is defined. For example, the space of square integrable functions is an Hilbert space (this space is particularly useful because any physical signal measured over a finite duration belongs to that space) and the scalar product is defined by:

$$< x, y >= \int_{-\infty}^{\infty} x(t) y^*(t) \mathrm{d}t \tag{1.2}$$

where y^* stands for the complex conjugate of y. If $< x, y >= 0$ the functions are orthogonal, they do not resemble to each other (actually this may not be completely true, consider for example the cos and sin functions). If $< x, y >$ is large, then the two functions resemble to each other. In particular, when one is positive (resp. negative) the other tends to be positive (resp. negative).

	Lecture/Lab	Content	Material to read
Part 1 Harmonic Analysis	Lecture 1 (2h)	Fourier Transform	Chaps. 2-3, Appendix A
	Lecture 2 (2h)	Discrete Fourier Transform	Chap. 4, Appendix B
	Lecture 3 (2h)	Random processes	Chap. 5, Appendix C
	Lab 1 (Matlab) (4h)	Spectral estimation	Text in Appendix E
Part 2 Time-Frequency Analysis	Lecture 4 (2h)	Time Frequency 1: introduction	Chap. 6
	Lecture 5 (2h)	Time Frequency 2: distributions	Chaps. 7-8, Appendix C
	Lecture 6 (2h)	Time Frequency 3: wavelets	Chap. 9
	Lab 2 (Matlab) (4h)	Wavelet transform	Text in Appendix E
Part 3 POD	Lecture 7 (2h)	Proper Orthogonal Decomposition	Chap. 10, Appendix D
	Lab 3 (Matlab) (4h)	Proper Orthogonal Decomposition	Text in Appendix E

Table 1.1: List of lectures and labs

Correlation
This is also linked to the concept of correlation. The autocorrelation of a complex signal $x(t)$ is given by

$$C_{xx}(\tau) =< x(t)x(t+\tau) >= \int_{-\infty}^{\infty} x(t)x^*(t+\tau)\mathrm{d}t \qquad \text{(Autocorrelation)} \tag{1.3}$$

It tells whether the signals $x(t)$ and $x(t+\tau)$ (the signal translated by τ) resemble each other. The intercorrelation between two complex signals $x(t)$ and $y(t)$ is given by:

$$C_{xy}(\tau) =< x(t)y(t+\tau) >= \int_{-\infty}^{\infty} x(t)y^*(t+\tau)\mathrm{d}t \qquad \text{(Intercorrelation)} \tag{1.4}$$

It tells wether the signals $x(t)$ and $y(t+\tau)$ (the signal translated by τ) resemble each other.

Note: most of the time we will consider real signals, for which the above definitions reduce to (no complex conjugation needed):

$$C_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)\mathrm{d}t \qquad \text{(Autocorrelation, real signal)} \tag{1.5}$$

$$C_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y(t+\tau)\mathrm{d}t \qquad \text{(Intercorrelation, real signals)} \tag{1.6}$$

Overview of the course
The course has 3 parts summarized in Table 1.1. Each part is concluded by a lab based on Matlab programming.

The first part deals with harmonic analysis for deterministic or random signals. Harmonic analysis is probably the first tool to be used for post-processing measured signals, especially when these signals (or sub-parts of them) are periodic or stationary. Two chapters review the theory of Fourier series and Fourier transform. A third chapter is dedicated to the discrete

Fourier transform, which is a tool that needs to be mastered for performing digital signal processing on a computer. The last chapter in this part gives an introduction to random processes. Many signals, starting with those measured in turbulent flows, are random and require some statistical tools that are presented. The lab for this part is about spectral analysis. Tools presented during the course are programmed and used on typical hot wire measurements.

While it is useful, harmonic analysis has some limitations. The Fourier transform represents the signal as a sum of harmonic waves characterized by a frequency, and moves the signal from the time domain to the frequency domain. Sometimes, one would like to consider both domains simultaneously, to observe for example how the amplitude of a given frequency component changes with time. More generally, some signals, such as transient signals do not resemble harmonic waves, and as we have said, a nice basis should usually resemble the signal. The second part gives an introduction to time-frequency analysis that tries of solve some of these issues. A chapter presents general concepts. Another chapter presents time-frequency energy distributions. Finally, atomic decompositions, such as the short time Fourier transform and the wavelet transform, are presented. Presently, the latter is the most widely spread time-frequency technique used for the analysis of turbulent flows. The lab for this part will consist in handling the continuous wavelet transform, and an example of application for a fluid flow will be presented.

The last part contains only one chapter on proper orthogonal decomposition (POD) and a related lab. POD is routinely used when analysing flow measurements to extract coherent structures (most energetic modes). Compared to the other methods presented in the course, POD devises a basis of POD modes from the measurements, ie the basis is built so as to resemble the data.

2 Fourier Series

In this chapter the Fourier series is introduced for periodic signals. The extension to signals defined on infinite intervals (with infinite period) is given by the Fourier Transform which is presented in the next chapter. This chapter introduces the ingredients that are found in the Fourier transform and in many transforms in this course:

- Expansion of a signal using a family of functions
- Analysis and reconstruction steps (transform and inverse transform)
- Link with bases and scalar product.

2.1 Fourier series expansion

The signal $x(t)$ that is considered is given on a *finite* interval $(0, T_0)$. It can be extended to all t in $(-\infty, \infty)$ by assuming that it repeats with a period T_0. Thus, a signal defined on $(0, T_0)$ and its periodic extension with period T_0 are equivalent, and both of them will be called periodic signals in this chapter. The **periodic** signals are precisely those for which the Fourier series should be used.

The **Fourier series (FS)** expansion of a T_0-periodic signal $x(t)$ is defined by:

$$x(t) = \sum_{n=-\infty}^{\infty} X_n e^{\mathrm{j}2\pi n f_0 t} \qquad (\mathbf{FS}) \left(\sim \mathbf{FST}^{-1}\right) \tag{2.1}$$

The notation FST^{-1} is unusual and specific to the present course, it stands for "Fourier Series Transform Inverse". This allows drawing a parallel with the inverse Fourier Transform to be introduced in the next chapter. Usually, an *inverse transform* corresponds simply to the *expansion* of a signal $x(t)$ using some expansion family. Presently, Eq. (2.1) means that the signal is a sum of harmonic functions $b_n(t)$ that belong to the set $\{b_n(t) = e^{\mathrm{j}2n\pi f_0 t}; n = -\infty...\infty\}$. In the sum (2.1) each of the functions $b_n(t)$ is weighted by a complex amplitude $X_n = |X_n|e^{\mathrm{j}\phi_n}$ that is also called a **Fourier coefficient**, or an expansion coefficient. Note that the frequency of the functions $b_n(t)$ is not arbitrary and has to be an integer multiple $n \times f_0$ (with $n{=}-\infty...\infty$) of the fundamental frequency $f_0 = 1/T_0$ that is the reciprocal of the period T_0. The part of the signal associated with the frequency nf_0 is called the $\mathbf{n}^{th}$ **harmonic**.

$$H_n(t) = X_{-n}e^{-\mathrm{j}2\pi n f_0 t} + X_n e^{\mathrm{j}2\pi n f_0 t} \quad \forall n = 0...\infty \qquad (n^{th} \text{ harmonic})$$

The component corresponding to $n = \pm 1$ is called the first harmonic or also the fundamental component, the one with $n = \pm 2$ is called the second harmonic, the one with $n = \pm 3$ is called the third harmonic, etc. The component with $n = 0$ is constant and represents the time average of the signal. Hence, one may write the Fourier series as a sum of harmonics:

$$x(t) = \underbrace{X_0}_{\text{mean}} + \underbrace{X_1 e^{\mathrm{j}2\pi f_0 t} + X_{-1} e^{-\mathrm{j}2\pi f_0 t}}_{\text{fundamental (1st harmonic)}} + \underbrace{X_2 e^{2\mathrm{j}2\pi f_0 t} + X_{-2} e^{-2\mathrm{j}2\pi f_0 t}}_{\text{2nd harmonic}} + \underbrace{X_3 e^{3\mathrm{j}2\pi f_0 t} + X_{-3} e^{-3\mathrm{j}2\pi f_0 t}}_{\text{3rd harmonic}} + ... \tag{2.2}$$

Note also that even if the original signal $x(t)$ is initially defined only on $(0, T_0)$, the series in Eq. (2.1) is defined over R. It equals the original signal within $(0, T_0)$ and is also defined outside this interval as it is periodic of period T_0. This is the reason why one had better consider $x(t)$ to be periodic from the start.

The Fourier coefficient X_n is calculated from:

$$X_n = \frac{1}{T_0} \int_0^{T_0} x(t) e^{-\mathrm{j}2\pi n f_0 t} dt \qquad (\sim \mathbf{FST}) \tag{2.3}$$

We have here introduced the unusual notation FST (for Fourier Series Transform) to draw a parallel with the inverse Fourier Transform defined in the next chapter. The transform allows calculating the expansion coefficients needed to calculate the expansion in the inverse transform. It is recalled that the signal $x(t)$ is periodic of period T_0, meaning the Fourier coefficient can be calculated over any interval of length T_0. On considering the interval $(-T_0/2, T_0/2)$, one may write as well:

$$X_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}2\pi n f_0 t} dt \qquad (\sim \mathbf{FST}) \tag{2.4}$$

Note: using Eq. (2.3) we have in particular:

$$X_0 = \frac{1}{T_0} \int_0^{T_0} x(t) dt \tag{2.5}$$

which confirms that X_0 is the time-average of the signal.

2.2 Convergence of the Fourier series

The Fourier series, Eq. (2.1), contains an infinite sum of functions and one should worry about the convergence of this series. The **partial sum** is defined as:

$$x_K(t) = \sum_{n=-K}^{K} X_n e^{\mathrm{j}2\pi n f_0 t} \qquad \text{(partial sum)} \tag{2.6}$$

The question is then to know whether $x_K \to x$ when $K \to \infty$, and in which sense the convergence is defined. There exist many theorems giving the conditions for a Fourier series to converge toward a function, and these are beyond the scope of the course. Let us give one theorem for pointwise convergence:

Theorem 2.2.1 (Dirichlet)**.** Let $x(t)$ be a periodic function with period T_0 and of class C^1 piecewise (x: $\mathbf{R} \to \mathbf{C}$). Then the Fourier series, Eq. (2.1), converges to $\dfrac{x(t^-) + x(t^+)}{2}$.

Note that the function should be piecewise continuous, and so should be its derivative. You may find slightly different forms of this theorem in the literature, with conditions on the total variations of $x(t)$ replacing the piecewise continuous character of the derivative. You are referred to Boyce and DiPrima [7, chapter 10] or to Whittaker and Watson [42, chapter IX], for example, to get more theorems and proofs. See also the book by Lanczos. An important feature of this theorem is that if a function x is continuous at point t (meaning its left and right limit satisfy $x(t^-)$=$x(t^+)$=$x(t)$) then its Fourier series converge to $x(t)$ (if the derivative satisfies the requirements of the theorem).
Note: a piecewise continuous function has a *finite* number of discontinuities, and at any discontinuity it should posess a finite limit on the left as well as a finite limit on the right. Hence, the theorem does not apply to functions having infinitly many discontinuities, or to functions having infinite discontinuities, such as $1/t^2$ at t=0.

2.3 Cosine/Sine form of the Fourier Series

Using the Euler formula Eq. (A.2), it is possible to write the Fourier series in terms of cos and sin functions in place of the complex exponentials:

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + b_n \sin(2\pi n f_0 t) \tag{2.7}$$

where the coefficients are given by

$$a_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cos(2\pi n f_0 t) \mathrm{d}t \tag{2.8}$$

$$b_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \sin(2\pi n f_0 t) \mathrm{d}t \tag{2.9}$$

One advantage of this expansion is that the coefficients a_n and b_n remain real when the function $x(t)$ is real.
Note: a_0 is twice the mean value of $x(t)$.

The coefficients a_n and b_n are related to those in Eq. (2.3) by:

$$\begin{cases} X_n = \frac{1}{2}(a_n - \mathrm{j} b_n) \\ X_{-n} = \frac{1}{2}(a_n + \mathrm{j} b_n) \end{cases} \tag{2.10}$$

or equivalently by:

$$\begin{cases} a_n = X_n + X_{-n} \\ b_n = \mathrm{j}(X_n - X_{-n}) \end{cases} \tag{2.11}$$

Even function
For an *even* function $x(t)$, that is, such that $x(-t) = x(t)$, we have:

$$b_n = 0 \tag{2.12}$$

$$a_n = \frac{4}{T_0}\int_0^{T_0/2} x(t)\cos(2\pi n f_0 t)\mathrm{d}t \tag{2.13}$$

Equation (2.7) then reduces to a cosine Fourier series:

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) \qquad \left(\begin{array}{l}\text{cosine Fourier series}\\ \text{for even function}\end{array}\right) \tag{2.14}$$

Odd function
For an *odd* function:

$$a_n = 0 \tag{2.15}$$

$$b_n = \frac{4}{T_0}\int_0^{T_0/2} x(t)\sin(2\pi n f_0 t)\mathrm{d}t \tag{2.16}$$

Equation (2.7) then reduces to a sine Fourier series:

$$x(t) = \sum_{n=1}^{\infty} b_n \sin(2\pi n f_0 t) \qquad \left(\begin{array}{l}\text{sine Fourier series}\\ \text{for odd function}\end{array}\right) \tag{2.17}$$

2.4 Link with bases and the scalar product

The most natural space for Fourier series is the space $\mathrm{L}^2(0,T_0)$, the space of square integrable functions over the interval $(0,T_0)$. In that space the family of functions $\{b_n(t) = e^{\mathrm{j}2n\pi f_0 t}; n = -\infty...\infty\}$, is an orthogonal basis (an Hilbert basis). Moreover, X_n in Eq. (2.3) is calculated by the scalar product of the signal $x(t)$ and the basis function $b_n(t)$, that is:

$$X_n = \langle x, b_n\rangle = \left\langle x, e^{\mathrm{j}2\pi n f_0 t}\right\rangle = \frac{1}{T_0}\int_0^{T_0/2} x(t)\left(e^{\mathrm{j}2\pi n f_0 t}\right)^* dt \qquad (\sim \mathbf{FST}) \tag{2.18}$$

Equation (2.1) is then nothing but the expansion of the signal $x(t)$ in the basis $\{b_n; n = -\infty...\infty\}$. One may write:

$$x(t) = \sum_{n=-\infty}^{\infty} \langle x, b_n\rangle\, b_n(t) \qquad (\mathbf{FS})\left(\sim \mathbf{FST}^{-1}\right) \tag{2.19}$$

Figure 2.1 summarizes these results. One may recognize two different steps:
• an analysis step (FST): the signal is projected onto the basis to get the Fourier coefficients (Eq. (2.3));
• a synthesis step (FS$\sim$FST^{-1}), or reconstruction step: the signal is reconstructed (Eq. (2.1)) from the knowledge of the basis and the Fourier coefficients.
The process of going from the function (signal) space on the top of the figure to the coefficient space in the bottom by using a transform and an inverse transform will be found in most of the

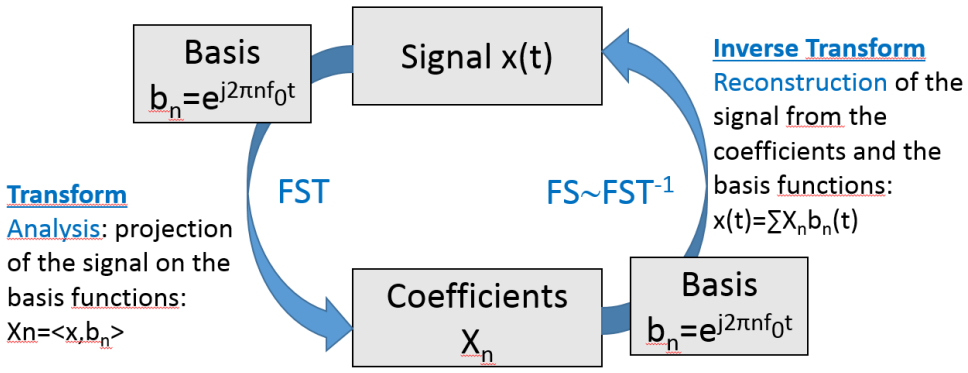


Figure 2.1: Analysis and reconstruction of a signal in the case of a Fourier series. FS denotes the Fourier Series, which in this course is also denoted FST^{-1} (this is the inverse transform). FST is also a notation found only in this course and represents the forward transform. The process of going from the function (signal) space on the top to the coefficient space in the bottom by using a transform and an inverse transform will be found in most of the transforms presented in this course.

transforms presented in this course.

For functions that do not belong to $\mathrm{L}^2(0,T_0)$ (for example functions in $\mathrm{L}^1(0,T_0)$) the scalar product is not defined and the interpretation above is not strictly valid, but everything happens "as if" it was the case.

2.5 About negative frequencies

In formula such as Eq. (2.1), the quantity nf_0 is the frequency associated with the component number n, and n is an integer varying from $-\infty$ to ∞, meaning one has both negative and positive frequencies. One should not worry about these negative frequencies. Positive frequencies are indeed not more meaningfull than negative ones. Physically, a phenomenon $x(t)$ has frequency $\pm f_0$ if there are $|f_0|$ cycles of this phenomenon per second. This may be written mathematically in several ways:

$$x(t) = a\cos(2\pi f_0 t + \phi) \tag{2.20}$$

or

$$x(t) = b\sin(2\pi f_0 t + \phi') \tag{2.21}$$

or

$$x(t) = c\cos(2\pi f_0 t) + d\sin(2\pi f_0 t) \tag{2.22}$$

In these equations, one may use arbitrarily $f_0 > 0$ or $f_0 < 0$ as this would only change the values of the amplitudes (a,b,c,d) or of the phase (ϕ,ϕ').

Let us consider for example the signal $x(t) = A\cos(2\pi f_0 t)$ with $f_0 > 0$. This signal may be written:

$$x(t) = \frac{A}{2}\left(e^{\mathrm{j}2\pi f_0 t} + e^{-\mathrm{j}2\pi f_0 t}\right) \tag{2.23}$$

This latter expression is nothing but the Fourier series for the cosine function, and it contains both positive and negative frequencies, f_0 and $-f_0$, and both of them are present in that case to express that the signal has $|f_0|$ cycles per second.

2.6 Asymptotic decay rate of X_n

The rate at which X_n decreases as $n \to \pm\infty$ is very important. Why is that so ? A periodic function $x(t)$ may be calculated exactly with the Fourier series in Eq. (2.1). In practice, one often wants to *approximate* the function by *truncating* the series and keeping only a finite number of terms. By keeping only the first K terms, one obtains the partial sum x_K already defined in Eq. (2.6) and given here again:

$$x_K(t) = \sum_{n=-K}^{K} X_n e^{\mathrm{j}2\pi n f_0 t} \qquad \text{(partial sum, or truncated } \mathbf{FS}\text{)} \tag{2.24}$$

Formally, one has:

$$x(t) = x_K(t) + \epsilon_K(t)$$

For the approximation to be good, the **truncation error**

$$\epsilon_K(t) = \sum_{n=K+1}^{\infty} X_n e^{\mathrm{j}2\pi n f_0 t} + \sum_{n=-\infty}^{-K-1} X_n e^{\mathrm{j}2\pi n f_0 t} \tag{2.25}$$

should be as small as possible, and this is the case if X_{K+1} and X_{-K-1} are small. Hence, it is important to know how small X_n becomes as n becomes large.

First, note that

$$X_n \to 0 \text{ as } n \to \pm\infty \quad .$$

The question then is to know *how fast* the 0 limit is reached. It can be shown that:

$$x(t) \in C^\infty \Leftrightarrow X_n = o\left(\frac{1}{|n|^p}\right) \text{ when } n \to \pm\infty, \qquad \forall p \in N \tag{2.26}$$

Hence, the rate at which X_n decreases as $n \to \pm\infty$ is directly related to the **regularity of the function**. One should remember:

The smoother the function is, the more rapidly X_n decreases.

If the function is only of class C^{k-1}, and is piecewise C^k then:

$$X_n = O\left(\frac{1}{|n|^k}\right) \tag{2.27}$$

In the worst possible case, that is for a discontinuous function that does not belong to C^0, X_n will decrease according to:

$$X_n \underset{n\to\infty}{\sim} O\left(\frac{1}{n}\right) \tag{2.28}$$

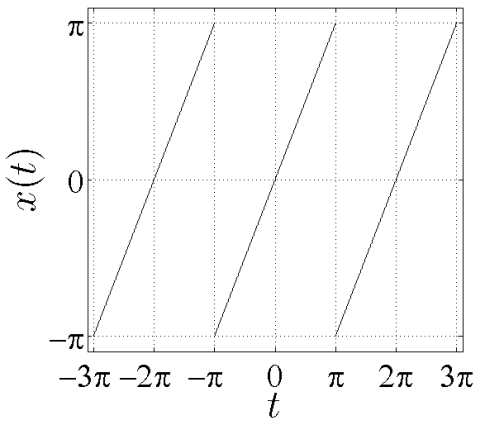


Figure 2.2: Sawtooth signal.

which is the worst possible decay rate. That the decay rate is small makes sense: a discontinuous function has very rapid variations at the point of discontinuity, and rapid variations correspond to high frequencies, that is, to large values of n. Thus, discontinuities induce a high frequency spectrum, and large values of X_n for large n.

Conversely, the regularity of the function may be determined from the asymptotic decay rate of X_n:

$$\text{If } X_n = o\left(\frac{1}{|n|^{k+2}}\right) \text{ then } x(t) \in C^k$$

2.7 Examples

2.7.1 Example 1: sawtooth wave

Consider the sawtooth signal $x(t)$=t over the interval (-π,π). This is represented in Fig. 2.2. In that case, T_0=2π (that is, $2\pi f_0$=1). Using Eq. (2.4), the Fourier coefficients are:

$$X_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} te^{-\mathrm{j}nt}\mathrm{d}t$$

which gives (after integration by parts):

$$X_n = \frac{\mathrm{j}(-1)^n}{n}$$

Note in particular that:

$$X_n \underset{n\to\infty}{\sim} O\left(\frac{1}{n}\right)$$

This is expected for a discontinuous signal (see Eq. (2.28)). This is the worst possible decay rate, meaning the convergence will be slow. The Fourier series for the signal is:

$$x(t) = \sum_{n=-\infty}^{\infty} \frac{\mathrm{j}(-1)^n}{n} e^{\mathrm{j}nt} \tag{2.29}$$

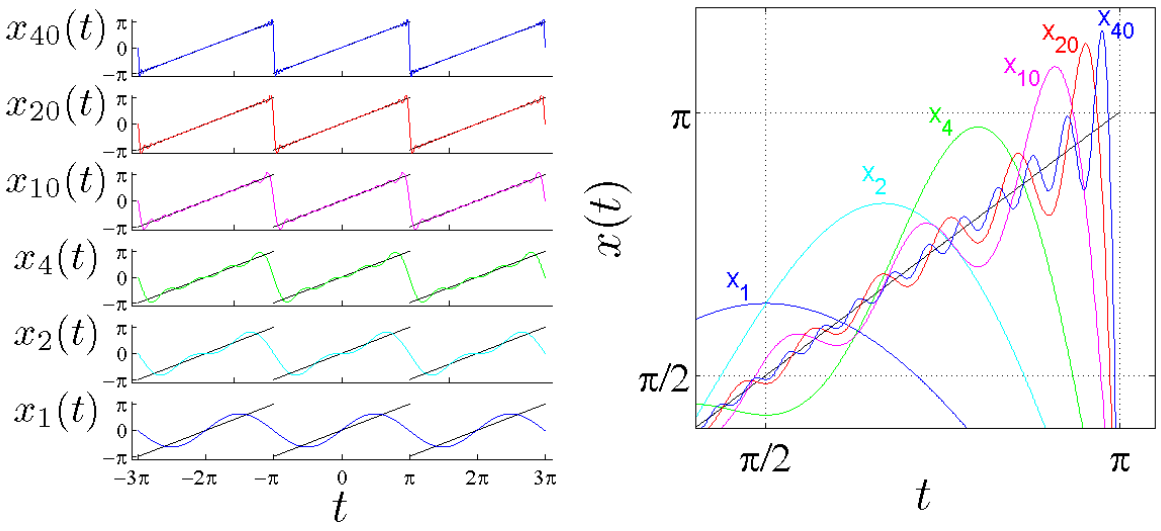


Figure 2.3: Left: Partial sums $x_K(t)$ for the sawtooth signal for K=1; 2; 4; 10; 20; 40 (color: partial sums; black: original signal). Right: zoom in on the discontinuity at x=π.

You may have noticed that the signal is odd and rightly deduced that the signal may be written as a Fourier sine series. Using Eq. (2.11) we find

$$b_n = \mathrm{j}(X_n - X_{-n}) = (-1)^{n+1}\frac{2}{n}$$

which is checked to be real (since $x(t)$ is real). Of course, it is also $O(1/n)$ as $n \to \infty$. The Fourier sine series for the signal is:

$$x(t) = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{2}{n}\sin(nt) \tag{2.30}$$

The partial sums are given by:

$$x_K(t) = \sum_{n=1}^{K}(-1)^{n+1}\frac{2}{n}\sin(nt)$$

These are plotted for some values of K in Fig. 2.3 together with the original signal. For a given time t, the difference $|x_K(t) - \frac{x(t^-)+x(t^+)}{2}|$ tends to zero as K increases. That is, at points t at which $x(t)$ is continuous $|x_K(t) - x(t)| \to 0$ as K increases. At the points t_k at which $x(t)$ is discontinuous (t_k=$\pi + 2k\pi, k \in \mathbf{N}$), the partial sum converges to 0=$\frac{-\pi+\pi}{2}$, which is half the sum of the left and right limits of $x(t)$ at those points.This is expected from theorem 2.2.1 on page 9.

Gibb's phenomenon

This is the right place to introduce Gibb's phenomenon that reflects the fact that the partial sum of a Fourier series is oscillatory at the **discontinuities** of the function, as seen in the right part of Fig. 2.3. Mathematically, there is simple convergence at discontinuities (this is the Dirichlet theorem 2.2.1) but there is not uniform convergence, which means:

$$||x_K - x||_\infty = \max_{\text{over } t}(|x_K(t) - x(t)|) \underset{K\to\infty}{\not\to} 0$$

The lack of uniform convergence is visible in the overshoot at $t = \pi$. At K increases, the position of the overshoot moves closer to π, but the amplitude of the overshoot does not decrease (observe that the overshoot of x_{40} is about the same as that of x_{20}).

2.7.2 Example 2

Let $x(t)$ be a periodic function with period $T_0 = 2\pi$. Over one period it is defined by:

$$x(t) = \begin{cases} \sin(t) & 0 < t < \pi \\ 0 & \pi < t < 2\pi \end{cases} \tag{2.31}$$

Contrary to the sawtooth function, this function is continuous. However, its derivative is not continuous at t=π. The cos/sin form of the Fourier series is used, the Fourier coefficients are given by:

$$a_n = \frac{1}{\pi}\int_0^{\pi} \sin(t)\cos(nt)\mathrm{d}t$$

$$b_n = \frac{1}{\pi}\int_0^{\pi} \sin(t)\sin(nt)\mathrm{d}t$$

On using trigonometric formula, one obtains:

$$a_n = \frac{1}{2\pi}\frac{2}{n^2-1}\left((-1)^{n+1}-1\right)$$

$$b_1 = \frac{1}{2} \quad ; \quad b_n = 0 \text{ for } n > 1$$

Note that a_n=0 for n odd. For n even, one has now:

$$a_n \underset{n\to\infty}{\sim} O\left(\frac{1}{n^2}\right)$$

This represents a faster convergence than for the sawtooth wave (for which the coefficients scale as $O(1/n)$), because the function is smoother than the sawtooth signal. Indeed, the function is C^0 and C^1 piecewise, so that Eq. (2.27) used with k=1 predicts the scaling $a_n = 0(1/n^2)$.

2.8 Conclusion

Periodic signals (signals on a finite interval) of period $T_0 = 1/f_0$ can be represented as Fourier series, that is, as an infinite sum of harmonic functions whose frequencies are multiple of the fundamental frequency f_0. Depending on the space considered, these functions form an orthogonal *basis*. While Fourier Series work for any periodic signal, they work better (converge more rapidly) for *regular* (continuous, differentiable) signals.
The Fourier series establishes a link between a *signal space* and the *Fourier coefficients space*, and one travels from one space to the other by using a transform (that resembles a scalar product) and an inverse tranform (which is the expansion itself, presently, the Fourier series expansion). This principle will be used repeatedly in most of the following chapters.

3 Fourier Transform

The Fourier series has been introduced in the previous chapter for periodic signals defined on finite intervals $(0, T_0)$. The extension to non-periodic signals defined on infinite intervals $(-\infty, \infty)$ is precisely given by the Fourier Transform introduced in the present chapter.
The Fourier transform, its inverse, and its properties are first adressed. The Heisenberg-Gabor principle relating the length of a signal in the time domain to the length of its Fourier transform in the frequency domain is then presented. The theory of linear time-invariant systems is not connected to Fourier transform. However, it is reviewed here (including the impulse response and the convolution product), so that the Fourier transform of the output of such systems can be investigated. The important convolution theorem is then given. Energy relations are presented next. Concepts such as correlation and power spectral density are introduced. The important Wiener-Khintchine theorem is given: it says that the power spectral density is the Fourier transform of the autocorrelation. Finally, windowing is briefly presented.
Only determistic signals are considered in the present chapter and extensions to random signals will be presented in a subsequent chapter.
Classical books on Fourier Transform: Bracewell [8].

3.1 Definition

Compared with the Fourier series, the signal is now defined on an infinite interval $(-\infty, \infty)$. The Fourier Transform expresses the fact that a function $x(t)$, under some conditions to be specified in section 3.2, can be expanded using the family $\{e^{\mathrm{j}2\pi ft}; f \in \mathbf{R}\}$. In that case the function $x(t)$ has the following expansion:

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{\mathrm{j}2\pi ft} \mathrm{d}f \tag{3.1}$$

In words, $x(t)$ is an integral sum of harmonic functions $e^{\mathrm{j}2\pi ft}$ in which each of these functions is weighted by a complex amplitude $X(f) = |X(f)|e^{\mathrm{j}\phi(f)}$. Note the difference with the Fourier series expansion (2.1) that contains a series (an infinite sum) rather than an integral.

The **Fourier Transform** (FT) is precisely a way of calculating the complex amplitude $X(f)$.

The Fourier tranform of a signal $x(t)$ (FT) reads:

$$X(f) = \mathrm{FT}[x(t)](f) = \int_{-\infty}^{\infty} x(t)e^{-\mathrm{j}2\pi ft}\mathrm{d}t \qquad (\mathbf{FT}) \tag{3.2}$$

Here, it is assumed of course that this quantity can be calculated.

The inverse transform is nothing but the expansion (3.1) given in the beginning of the section. The **inverse Fourier transform** (IFT or FT^{-1}) reads:

$$x(t) = \mathrm{FT}^{-1}[X(f)](t) = \int_{-\infty}^{\infty} X(f)e^{\mathrm{j}2\pi ft}\mathrm{d}f \qquad \left(\mathbf{FT}^{-1}\right) \tag{3.3}$$

One may also view this equation as a reconstruction formula, whereby a signal is recontructed from its (harmonic) pieces. Within the signal $x(t)$, the wave $e^{\mathrm{j}2\pi ft}$ with frequency f has an amplitude $|X(f)|$ and a phase $\phi(f) = arg(X(f))$.

3.2 Conditions for being able to compute the FT and the IFT

There exists many theorems giving the conditions for the FT and the IFT to be computable quantities and these are beyond the scope of this course. Only some basics are recalled here.

Functions in L^1:
L^1 is the most natural space here because integrable functions admit a Fourier transform. However, this transform does not necessarily belong to L^1 itself, meaning that the inverse tranform (of the transform) is not always defined for functions in L^1. That is, there is not necessarily a way back from the frequency domain, and the possibility of reconstructing a signal from its pieces is compromised.

Functions in L^2:
The Fourier Transform *and* its inverse are both defined for signals $x(t)$ in L^2, the space of square integrable signals. By definition, these signals verify:

$$\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty \qquad (x(t) \in L^2) \tag{3.4}$$

This integral is defined as the energy of the signal (see Section 3.10.1), and the signals belonging to L^2 are thus the signals having finite energy. Example: the Gaussian function $x(t) = e^{-\alpha^2 t^2}$ belongs to L^2, its FT is defined and is $X(f) = TF[x(t)] = \frac{\sqrt{\pi}}{\alpha} e^{-\frac{\pi^2 f^2}{\alpha^2}}$, and in addition $FT^{-1}[X(f)] = x(t)$. The class of signals with finite energy is especially important because signals observed experimentally are always windowed somehow and fall into this category. Hence, L^2 is the most natural space for measured signals. Measured signals will have a FT and a FT^{-1}.

Other functions:
While measured signals have finite energy because we have made them so by windowing, many

other signals do not have this property. This is the case for stationary random processes in general, or more simply for cosines, sines, or complex exponentials that are very useful in signal processing. For some signals that are not in L^2 or even in L^1 it is still possible to define a FT when one is considering distributions (also called generalized functions). For example the complex exponential $e^{\mathrm{j}2\pi f_0 t}$ does not belong to L^2 (it has infinite energy) or in L^1, but is has a Fourier transform:

$$TF[e^{\mathrm{j}2\pi f_0 t}] = \delta(f - f_0)$$

For a cosine, the Fourier transform is given by:

$$TF[\cos(2\pi f_0 t)] = \frac{1}{2}[\delta(f - f_0) + \delta(f - f_0)]$$

These transforms contain a Dirac delta function, which is not a regular fonction but a distribution. Some simple properties of the Dirac delta function are recalled in the appendix A.

3.3 Link with bases and the scalar product

As we have seen, the space L^2 is particularly interesting because measured signals (those with finite energy) belong to that space. L^2 is an Hilbert space (a vector space that has an inner/scalar product and that is complete for the norm associated with this scalar product). The **scalar product** between two functions $x(t)$ and $y(t)$ in L^2 is:

$$\langle x, y \rangle = \int_{-\infty}^{\infty} x(t) y^*(t) dt \tag{3.5}$$

Unfortunately, the set $\{b_f(t) = e^{\mathrm{j}2\pi ft}; f \in R\}$ is not a basis for that space (these functions do not belong to L^2, and the set is not countable either). Nevertheless, everythings happens nearly *as if* this set worked as a basis.

Two members of this set are **orthogonal**:

$$\left\langle e^{\mathrm{j}2\pi ft}, e^{\mathrm{j}2\pi f't} \right\rangle = \int_{-\infty}^{\infty} e^{\mathrm{j}2\pi(f-f')t} dt = \delta(f - f') \tag{3.6}$$

Any function can be expanded using this set, and this expansion is exactly what the FT^{-1} is:

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{\mathrm{j}2\pi ft} \mathrm{d}f = \int_{-\infty}^{\infty} X(f) b_f(t) \mathrm{d}f \tag{3.7}$$

The $X(f)$ are the complex coefficients of the expansion, and they may be found by projecting the function onto the set using the scalar product. This projection is exactly what the Fourier transform does:

$$X(f) = \mathrm{FT}[x(t)](f) = \int_{-\infty}^{\infty} x(t) e^{-\mathrm{j}2\pi ft} \mathrm{d}t = \left\langle x(t), e^{\mathrm{j}2\pi ft} \right\rangle \tag{3.8}$$

Synthetically, one can write:

$$x(t) = \int_{-\infty}^{\infty} \left\langle x(t), e^{\mathrm{j}2\pi ft} \right\rangle e^{\mathrm{j}2\pi ft} \mathrm{d}f \tag{3.9}$$

This relation is the equivalent for an uncoutable set to Eq. (2.19) that was given earlier for the Fourier series.

3.4 Properties of the Fourier Transform

Some classical properties of the FT are given in table 3.1.

Dilatation/contraction property:
This property states that:

$$g(t) = h(kt) \rightleftharpoons G(f) = \frac{1}{|k|} H(f/k) \tag{3.10}$$

Hence, if the function gets contracted in the time domain ($k > 1$) then its Fourier transform gets expanded in the frequency domain. The physical reason behind this is the following: when a function is contracted in the time domain, its variations are enhanced, which leads to more rapid variations. Hence, its high-frequency content increases, that is, its Fourier transform is shifted toward higher frequencies. See also the Heisenberg-Gabor principle below that expresses the same idea.

	Signal	FT of the signal
Linearity	$g(t) + h(t)$	$G(f) + H(f)$
Translation in time	$g(t + \tau)$	$G(f)e^{j2\pi f\tau}$
Modulation (translation in frequency)	$g(t)e^{j2\pi f_0 t}$	$G(f - f_0)$
Dilatation ($k < 1$) -contraction ($k > 1$)	$g(kt)$	$\frac{1}{\|k\|}G(f/k)$
Time reversal	$g(-t)$ with g real	$G^*(f)$
Real signal	$g(t)$ real	$G(-f) = G^*(f)$
Time derivative	$g(t) = \frac{\mathrm{d}h}{\mathrm{d}t}$	$G(f) = \mathrm{j}2\pi f H(f)$
nth order time derivative	$g(t) = \frac{\mathrm{d}^n h}{\mathrm{d}t^n}$	$G(f) = (\mathrm{j}2\pi f)^n H(f)$

Table 3.1: Some properties of the Fourier Transform.

3.5 Fourier Transform of classical signals

Some classical signals together with their Fourier transform is given in table 3.2.

The constant signal is defined by $x(t) = 1 \quad , \forall t$. This signal does not vary in time. Hence, it contains no time varying components. Its Fourier transform is $X(f) = \delta(f)$, meaning that there is only a $f = 0$ component in the spectrum (the f=0 component corresponds to the mean of the signal).

The Dirac impulse is defined by $x(t) = \delta(t)$. This signal is varying extremely fast at $t = 0$. To be able to describe this signal as a sum of sinusoidal waves, one has to add waves with all possible frequencies, all waves having an equal amplitude. This is expressed mathematically by: $X(f) = 1 \quad , \forall f$. Physically, one excites acoustic waves with all possible frequencies by clapping one's hands or by firing a gun shot.

That the Fourier transform of $x(t) = \delta(t - t_0)$ is given by $X(f) = e^{-j2\pi f t_0}$ is simply a result of the translation property explained in the previous section.

	Signal	FT of the signal
Constant	1	$\delta(f)$
Dirac delta function	$\delta(t)$	1
Translated Dirac delta function	$\delta(t - t_0)$	$e^{-j2\pi f t_0}$
Complex exponential	$e^{j2\pi f_0 t}$	$\delta(f - f_0)$
Cosine	$\cos(2\pi f_0 t)$	$\frac{1}{2}[\delta(f - f_0) + \delta(f + f_0)]$
Sine	$\sin(2\pi f_0 t)$	$\frac{1}{2j}[\delta(f - f_0) - \delta(f + f_0)]$
Dirac comb	$\sum_{k=-\infty}^{\infty} \delta(t - kT)$	$\frac{1}{T}\sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T}\right)$
Rectangular Window	$\Pi_T(\mathrm{t})$	$T\mathrm{sinc}(\pi f T)$=$T\frac{sin(\pi f T)}{\pi f T}$
Gaussian function	$e^{-\alpha^2 t^2}$	$\frac{\sqrt{\pi}}{\alpha} e^{-\frac{\pi^2 f^2}{\alpha^2}}$

Table 3.2: Some usual signals and their Fourier Transform.

Exercise: *show that the Fourier transform of the rectangular function* $\Pi_T(t)$ *is a* sinc *function as given in the Table. Is* $\Pi_T(t)$ *continous? Is its FT continuous? When* $f \to \infty$ *how decreases the* sinc *function ?*

3.6 Relationship between the Fourier Transform and the Fourier Series

The Fourier series applies to a signal with finite time support $(0, T_0)$. The Fourier transform is used for signals defined on $(-\infty, \infty)$. For a signal that is defined on $(-\infty, \infty)$ but is periodic with period T_0, both the Fourier transform and the Fourier series can be used, and both transforms in that case are linked indeed, as is now shown.

Consider a T_0-periodic signal. It has a Fourier series:

$$x(t) = \sum_{n=-\infty}^{\infty} X_n e^{\mathrm{j}2\pi n f_0 t} \tag{3.11}$$

Let's take the FT of this signal:

$$\begin{aligned} X(f) = \mathrm{FT}\left[x(t)\right] &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} X_n e^{\mathrm{j}2\pi n f_0 t} e^{-\mathrm{j}2\pi f t} \mathrm{d}t & (3.12)\\ &= \sum_{n=-\infty}^{\infty} X_n \int_{-\infty}^{\infty} e^{\mathrm{j}2\pi n f_0 t} e^{-\mathrm{j}2\pi f t} \mathrm{d}t & (3.13)\\ &= \sum_{n=-\infty}^{\infty} X_n \mathrm{FT}\left[e^{\mathrm{j}2\pi n f_0 t}\right] & (3.14)\\ &= \sum_{n=-\infty}^{\infty} X_n \delta(f - n f_0) & (3.15) \end{aligned}$$

The Fourier transform of a periodic signal is thus a Dirac comb with period $f_0 = 1/T_0$, and each of the Dirac $\delta(f - nf_0)$ in this comb is weighted by the Fourier coefficient X_n. See Fig. 3.1. For a periodic signal, the FS and FT are equivalent.

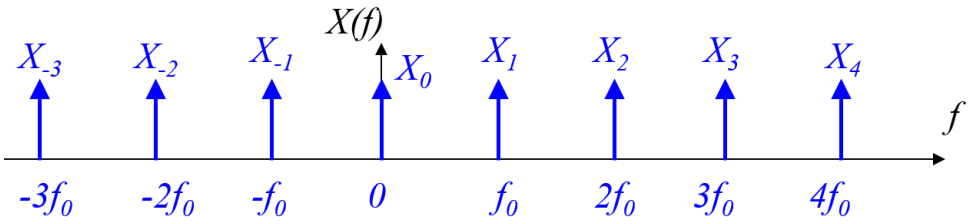


Figure 3.1: Fourier transform of a periodic signal.

3.7 Heisenberg-Gabor Principle

The Heisenberg-Gabor principle relates the extent of a function in the time domain and the extent of its Fourier transform in the frequency domain. Such a relation was already expressed in the dilatation property of the FT. The principle is simple yet very important and has important consequences for the Fourier transform as well as for time-frequency methods that are presented in subsequent chapters.

Before presenting this principle, some *gross characterizations* of the signal are introduced. For a signal $x(t)$, the **mean time** is given by:

$$t_m = \frac{1}{E_x} \int_{-\infty}^{\infty} t|x(t)|^2 \mathrm{d}t \tag{3.16}$$

where the energy E_x of the signal is defined by:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \mathrm{d}t \tag{3.17}$$

Hence, the mean time can be calculated for a signal having finite energy. The mean time is the time around which most of the energy of the signal occurs. The **effective duration**, T_e, tells you how the signal is spread arount that mean time. It is defined by:

$$T_e^2 = \frac{1}{E_x} \int_{-\infty}^{\infty} (t - t_m)^2 |x(t)|^2 \mathrm{d}t \tag{3.18}$$

It gives the typical extent of $x(t)$ in the time domain, most of the energy being contained in the time interval $(t_m - T_e, t_m + T_e)$. Equivalent characterizations are available in the frequency domain. The **mean frequency** is defined by

$$f_m = \frac{1}{E_x} \int_{-\infty}^{\infty} f|X(f)|^2 \mathrm{d}f \tag{3.19}$$

where $X(f)$ is the Fourier transform of $x(t)$. The mean frequency is the frequency at which most of the energy spectral density is located (this quantity is defined more precisely later in

this course). For a real signal we have $|X(-f)|=|X(f)|$, which implies that $f_m=0$. The **effective bandwidth** is defined by

$$B_e^2 = \frac{1}{E_x}\int_{-\infty}^{\infty}(f-f_m)^2|X(f)|^2\mathrm{d}f \tag{3.20}$$

It gives the typical extent of $X(f)$ in the frequency domain. These gross characterizations of the signal in the time and frequency domains are summarized in Fig. 3.2.

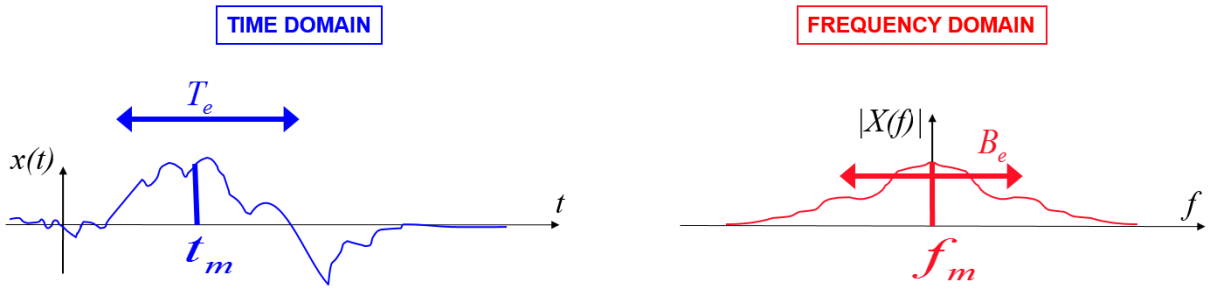


Figure 3.2: Figure showing t_m and T_e for the signal $x(t)$ as well as f_m and B_e for its Fourier Transform $X(f)$. Since $X(f)$=TF$[x(t)]$, the quantities T_e and B_e are not independent and are related by the Heisenberg-Gabor principle.

T_e depends on $x(t)$ and B_e depends on $X(f)$. Moreover, $x(t)$ and $X(f)$ are not independent from each other since the latter is the Fourier transform of the former. Hence, the extent of a function in the time domain and the extent of its Fourier transform in the frequency domain are not independent. Indeed, the **Heisenberg-Gabor Principle** states that they are inverse proportional. For any signal (with finite energy), the effective duration and effective bandwidth verify:

$$T_e B_e \geq \frac{1}{4\pi} \qquad \text{(Heisenberg-Gabor principle)} \tag{3.21}$$

In words, B_e and T_e can not be made arbitrarily small simultaneously. A short signal, one with a small T_e, will necessarily have a large B_e, that is, it will have a large frequency range. On the other hand, suppose that you want to measure the frequency of a harmonic signal with a high precision. A high precision in the frequency space requires B_e small, and the above principle then states that T_e should be large. Thus, it is possible to know the frequency of a harmonic signal with a high precision only if this signal is observed over a long enough duration (see also Section 3.11 on windowing). One should remember:

$$\boxed{\text{High precision on } f \rightleftharpoons \text{long observation time } (T_{obs} >> 1/f)}$$

Exercise: *show that the equality in Eq. (3.21) is met for a Gaussian signal.*
Hint: use the FT of a Gaussian signal (see tables), use the definition of the energy, E_x. Due to parity, show that t_m=0 and f_m=0. Use integration by parts and the Gauss integral:

$$\int_{-\infty}^{\infty} e^{-u^2}\mathrm{d}u = \sqrt{\pi}$$

Figure 3.3 gives a qualitative illustration of the Heisenberg-Gabor principle by considering 4 functions in the top row, and their corrresponding Fourier transforms in the bottom row (note that the cases (a)/(a') and (d)/(d') do not correspond to a finite energy either in the time or frequency domain but are used nevertheless for this illustration):
Pair (a)/(a') : the time function in (a) is a Dirac delta function, which may be understood as the limit of a rectangular function whose time-support goes to zero (while its integral is kept constant). It is extremely compact since it is non-zero only at t=0. Its effective duration is "T_e=0". Its Fourier transform in (a') has an infinite effective bandwidth: B_e=∞.
Pair (b)/(b') : the time function in (b) is a rectangular function with T_e=$T_{e,1}$, corresponding to some bandwidth $B_{e,1}$. Compared to the case (a)/(a') the time extent has increased, and as a result of the Heisenberg-Gabor principle the bandwidth has decreased: $B_{e,1} < \infty$.
Pair (c)/(c') : the time function in (c) is now a rectangular function having $T_{e,2} > T_{e,1}$. As a result the corresponding bandwidth in (c') is smaller that than in (b'): $B_{e,2} < B_{e,1}$.
Pair (d)/(d') : this case is the reverse of case (a)/(a'). The signal in (d) is now the constant signal, which is also a rectangular function whose support extends to infinity. Thus, we have "T_e=∞". As a result, its FT (the Dirac delta function) satifies: B_e=0.
Overall, when moving from (a)/(a') to (d)/(d') the effective duration increases and at the same time the bandwidth decreases so as to satisfy the Heisenberg-Gabor principle.

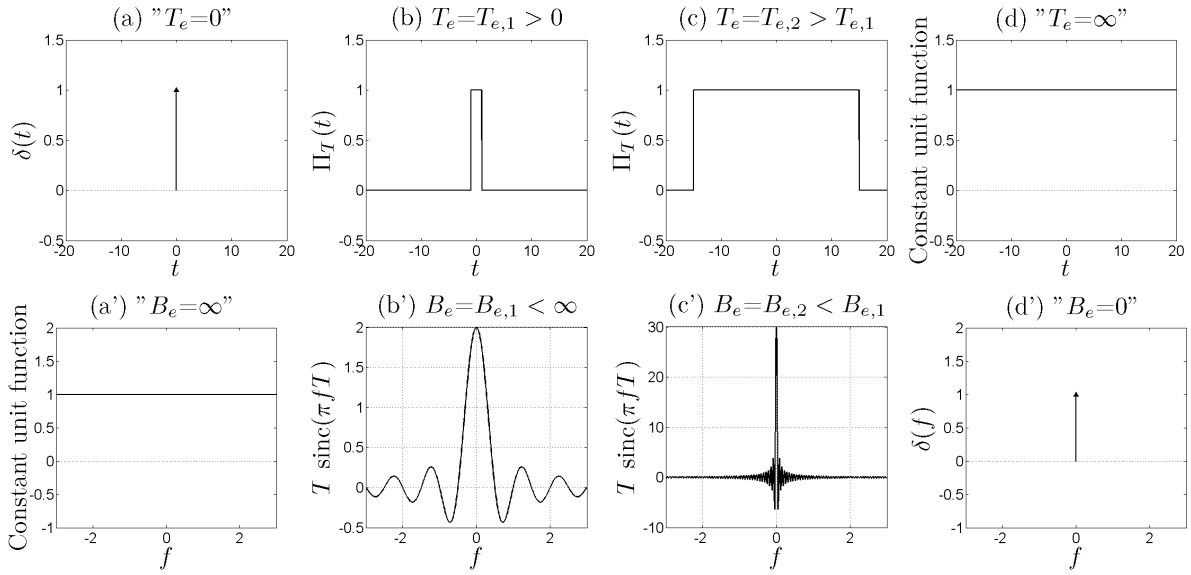


Figure 3.3: This figure illustrates that the more compact a function is in the time domain the more extended its Fourier transform is in the frequency domain. The top row represents four functions in the time domain: (a) the Dirac delta function (the more compact within the four); (b) a narrow rectangular window; (c) a wide rectangular function; (d) the constant function (the less compact within the four). The bottom row represents the Fourier transforms of the functions in the top row.

3.8 Regularity of $x(t)$ and asymptotic decreasing of $X(f)$

When studying the Fourier series, it has been seen in Section 2.6 that the decreasing of the Fourier coefficients X_n when $n \to \infty$ is related to the regularity of the function. The more regular the function (ie, the more derivatives can be calculated, or also, the smoother it is), the more rapid the decreasing of its Fourier coefficients. A similar property holds for Fourier transform pairs:
• if the function $x(t)$ is C^K, and if each of its derivatives $x^{(k)}(t)$, k=0...K are integrable, then the Fourier transform $X(f)$ verifies:

$$X(f) \leq O\left(\frac{1}{f^K}\right) \quad \text{when} \quad f \to \infty$$

Thus, **the more regular $x(t)$, the faster $X(f)$ decreases as $f \to \infty$.**

Conversely:
• if the function $x(t)$ decreases rapidly with t in the following sense: $t^k x(t)$ can be integrated $\forall k = 0...K$, then its Fourier transform $X(f)$ can be derivated K-times. Hence, **the more rapidly $x(t)$ decreases as $t \to \infty$, the more regular its Fourier transform $X(f)$ is.**

One will remember:

The regularity (smoothness) in one domain is equivalent to the asymptotic decreasing rate in the dual domain.

Example: the FT of the rectangular window $\Pi_T(t)$ is $T\text{sinc}(\pi fT)$=$T\frac{sin(\pi fT)}{\pi fT}$ (see Table 3.2 on page 21). The function $\Pi(t)$ is discontinuous (that is, not smooth, or not regular), and its Fourier transform decays only as $1/f$ as $f \to \infty$. Conversely, when $X(f) \sim 1/f$ as $f \to \infty$, one can infer that the signal $x(t)$ is discontinuous. Just by looking at $T\text{sinc}(\pi fT)$ one knows it is the Fourier transform of a discontinuous signal.

Example 2: the Fourier transform of a Gaussian is a Gaussian (again, see Table 3.2 on page 21). The Gaussian signal $x(t)$ is very regular since it belongs to C^∞. And one checks that its FT is fastly decreasing as $f \to \infty$, since it is a Gaussian that decreases faster than any power of t. Being a Gaussian, the FT is also C^∞. As a result, the corresponding signal is fastly decreasing. Thus, for the Gaussian, both the signal and its FT are both fastly decreasing and highly regular, one of these properties in one domain being linked to the other property in the dual domain.

3.9 Linear time invariant systems, Convolution, Impulse Response

Frequently, one would like to establish a connection between two signals. These two signals can result, for example, of the measurements made with two different sensors. It frequently happens, but not always, that a linear relationship exists between the two signals. What interests us here is what this relationship is in the frequency domain: in short, what is the relationship between one frequency component of the first signal and the corresponding frequency component of the

second signal? Before adressing this question, the theory of linear time invariant systems needs to be reviewed in the time domain.

3.9.1 Linear time-invariant systems

Two signals $x(t)$ and $y(t)$ may be linked by a linear time-invariant (LTI) relationship. This means that there exists some linear differential equation relating $x(t)$ and $y(t)$:

$$F\left(y, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}^2y}{\mathrm{d}t^2}, ..., \frac{\mathrm{d}^{N-1}y}{\mathrm{d}t^{N-1}}, x, \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}^2x}{\mathrm{d}t^2}, ..., \frac{\mathrm{d}^{M-1}x}{\mathrm{d}t^{M-1}}\right) = 0 \quad . \tag{3.22}$$

The coefficients appearing in front of the derivatives in this equation have to be *time invariant*. Equivalently, $x(t)$ may be seen as the input of a system, and $y(t)$ as its output. Such a system is generally represented by a box as in Fig. 3.4. In electric engineering a resistor is such a system

Figure 3.4: Linear Time Invariant (LTI) system with one input and one output.

(if its resistance does not increase with time due to heating, in which case it would not be time invariant). However, one should be aware that more complex situations may be represented by a LTI system. For example, Fig. 3.5 represents a forward step flow. One may wonder whether the relationship between the velocities measured at two different points can be modeled by a LTI. The dynamics of the flow is governed by the Navier-Stokes equations which are nonlinear. As a result, strictly speaking, the answer to the question is 'no'. However, if the two points are close enough, a LTI can approximately be observed between the two points (see chapter 5).

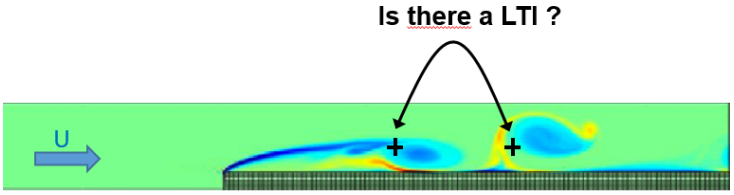


Figure 3.5: Forward step flow (vorticity).

In general, LTI system are useful but cannot describe all possible situations. In particular they do not describe nonlinear relationships. For example, the relationship

$$y(t) = x(t)\frac{\mathrm{d}x}{dt}$$

cannot be represented by a LTI system having $x(t)$ as input and $y(t)$ as output because $y(t)$ depends nonlinearly on $x(t)$. The Navier-Stokes equations are nonlinear. Hence, in general, there is no LTI linking two points in a flow. However, the linearized Navier-Stokes equations used in stability are linear by definition, as well as those in linear acoustics. LTI are much used in acoustics as a result.

3.9.2 Characterization of a LTI system - Impulse response

From the theory of differential equations [7], you should know that the solution $y(t)$ of the linear inhomogenous differential equation $Ly(t) = x(t)$ (L is a *linear* differential operator) is given by $y = G * x$, where $G(t)$ is the Green function and $*$ is the time convolution. The Green function is solution to $LG = \delta$ where $\delta(t)$ is the Dirac distribution. This means that the solution of the differential equation having as source a Dirac impulse is of special importance. This is no different for a LTI system since such a system is underpinned by a *linear* differential equation.

The **impulse response (IR)** of a LTI system is the response $h(t)$ when the input is the Dirac impulse $\delta(t)$, as schematically shown in Fig. 3.6. Note that the impulse response is nothing but the Green function in the language of the theory of differential equations. For satisfying

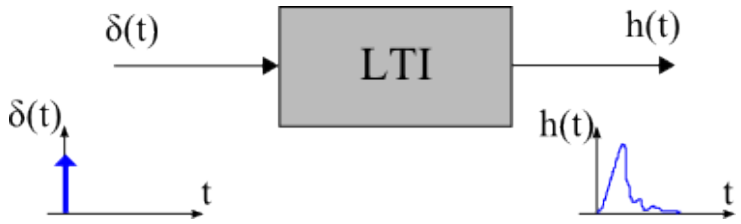


Figure 3.6: Impulse response of a LTI.

causality for a physical system, one has necessarily:

$$h(t) = 0 \qquad \text{for} \qquad t < 0 \qquad \text{(causality)} \tag{3.23}$$

In addition, the system has to be stable: BIBO stability (for Bounded Input Bounded Output) states that the output of the system is bounded when the input is bounded. This requires:

$$\int_{-\infty}^{\infty} |h(t)|dt < \infty \qquad \text{(BIBO stability)} \tag{3.24}$$

In particular this implies that:

$$h(t) \underset{t\to\infty}{\to} 0 \tag{3.25}$$

meaning the impulse response vanishes for large times.

The impulse response is very important because it fully characterizes the LTI system. Once it is known, the response to an arbitray input can be determined by using the convolution product. This is the reason why this product is now introduced.

The **convolution product**, usually represented by a $*$ symbol, between two signals $a(t)$ and $b(t)$ both depending on the variable t is a signal depending on t as well and defined by:

$$a(t) * b(t) = \int_{-\infty}^{\infty} a(\tau)b(t-\tau)\mathrm{d}\tau = \int_{-\infty}^{\infty} b(\tau)a(t-\tau)\mathrm{d}\tau \qquad \text{(convolution product)} \tag{3.26}$$

The second equality stems from the convolution product being commutative, ie $a * b = b * a$. From the definition (see Eq. (A.7)) of the Dirac delta function, one may easily check that the Dirac delta function is the neutral element for the convolution, that is:

$$a(t) * \delta(t) = \delta(t) * a(t) = a(t) \qquad \forall a(t) \tag{3.27}$$

The convolution also allows translating a signal using a translated Dirac delta function:

$$a(t) * \delta(t - t_0) = a(t - t_0) \qquad \forall a(t) \tag{3.28}$$

This translation property can be used to repeat a pattern with a given period, see Fig. 3.7. This property underpins the theory of sampling (see chapter 4).

The translation property of the convolution can be used to periodize some given pattern. Suppose that the pattern is the rectangular window of width $T/2$ and that is needs to be periodized with a period T. One would calculate:

$$z(t) = \Pi_{T/2}(t) * \underbrace{\sum_{k=-\infty}^{\infty} \delta(t - kT)}_{\text{Dirac comb}} = \sum_{k=-\infty}^{\infty} \Pi_{T/2}(t - kT)$$

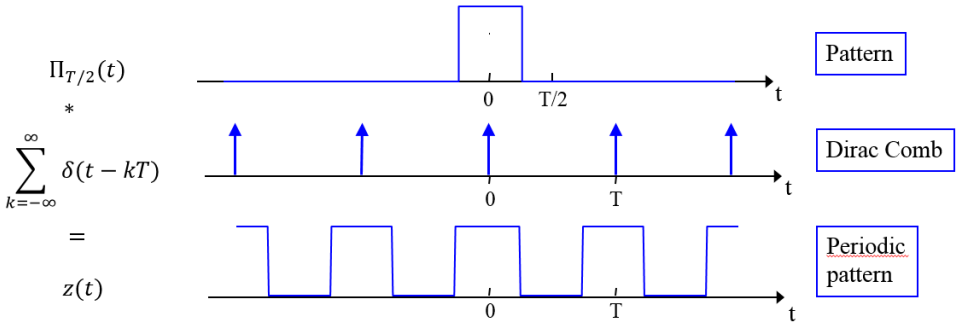


Figure 3.7: Periodization of a pattern using the convolution product of the pattern with a Dirac comb.

Knowing the impulse response $h(t)$ of a LTI system, the response $y(t)$ to an arbitrary input signal $x(t)$ is given by:

$$y(t) = h(t) * x(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)\mathrm{d}\tau = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)\mathrm{d}\tau$$

That is, the output is the convolution product between the input and the impulse response of the LTI system. When $h(t)$ is causal, as it should be the case for the system to be physically relevant, one has:

$$y(t) = h(t) * x(t) = \int_{0}^{\infty} h(\tau)x(t - \tau)\mathrm{d}\tau \qquad \text{(output of a causal LTI system)}$$

The meaning of the equation is the following: to calculate the output $y(t)$ at times t one needs to integrate the input values $x(t - \tau)$ at the past times $t - \tau < t$, and these values are weighted by a weighting factor $h(\tau)$. The larger $h(\tau)$ the larger the importance of past values $x(t - \tau)$ in the calculation of $y(t)$. Recall that stability (Eq. (3.25)) implies that $h(t) \to 0$ as $t \to \infty$, which means that values of the input in the far past have low effect on the value of the output at the present time.

3.9.3 LTI and Fourier Transform

Up to now the LTI system has been considered in the time domain. It is also possible to consider the frequency domain. We have 3 signals: the input $x(t)$, the output $y(t)$, and the impulse response $h(t)$. Assuming they can be Fourier-transformed, we have the 3 Fourier transforms:

$$X(f) = \mathrm{FT}\left[x(t)\right] \tag{3.29}$$
$$Y(f) = \mathrm{FT}\left[y(t)\right] \tag{3.30}$$
$$H(f) = \mathrm{FT}\left[h(t)\right] \qquad \text{(frequency response)} \tag{3.31}$$

$H(f) = \mathrm{FT}\left[h(t)\right]$ is called the **frequency response** of the LTI. It is a complex quantity, which has a norm and a phase: $H(f) = |H(f)|e^{\mathrm{j}\phi_H(f)}$. Since $h(t)$ can be obtained from $H(f)$ by the inverse Fourier transform, the system can be characterized equivalently by $h(t)$ or by $H(f)$.

The relationship that holds between the input and output in the frequency domain is the frequency domain equivalent of the convolution product $y(t) = x(t) * h(t)$. The connection is given by the **convolution theorem**:

$$\mathrm{FT}\left[a(t) * b(t)\right] = \mathrm{FT}\left[a(t)\right] \cdot \mathrm{FT}\left[b(t)\right] \qquad \text{(Convolution theorem)} \tag{3.32}$$

that is, the Fourier transform turns the convolution product into a regular product. Note that the converse is true as well:

$$\mathrm{FT}\left[a(t) \cdot b(t)\right] = \mathrm{FT}\left[a(t)\right] * \mathrm{FT}\left[b(t)\right] \qquad \text{(Convolution theorem 2)} \tag{3.33}$$

The Fourier transform turns a regular product into a convolution product (this is used in Section 3.11 on windowing).

Thus, by taking the Fourier transform of $y(t) = x(t) * h(t)$ and using the convolution theorem, one obtains:

$$Y(f) = X(f) \cdot H(f) \tag{3.34}$$

In the frequency domain, the output is the regular product of the input and the frequency response. In particular, the output component at frequency f depends only on the component with the *same* frequency in the input. This would not be the case for nonlinear systems: for example, the turbulent cascade from large to small scale in turbulent flows is a nonlinear effect, where different scales can interact. In linear systems considered here, different scales/frequencies do not interact.

Note: in the simple case when the input is a wave $x(t) = e^{j2\pi f_0 t}$, its FT is given by $X(f) = \delta(f - f_0)$. The FT of the output is: $Y(f) = H(f)X(f) = H(f_0)\delta(f - f_0)$. The output signal is then

$$y(t) = H(f_0)e^{j2\pi f_0 t} = |H(f_0)|e^{j(2\pi f_0 t + \phi_H(f_0))}$$

This confirms that the effect of the LTI on a wave at frequency f_0 is simply to modify its amplitude and its phase according to the complex frequency response $H(f_0)$ at this frequency.

Figure 3.8 summarizes the link between the input and the output for a LTI system. Usually, *filtering* a signal refers to a multiplication by $H(f)$ in the frequency domain, and to a convolution with $h(t)$ in the time domain.

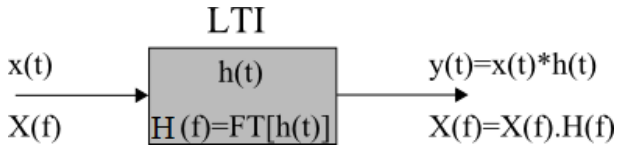


Figure 3.8: Link between the input and the output in the time and frequency domains for a LTI system.

Exercise: *A LTI system has the following impulse response:*

$$h(t) = \begin{cases} 1 & \text{for } |t| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.35}$$

What is the fundamental problem with this response is one was to built practically the system? Calculate the output y(t) for the following input:

$$x(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t \geq 0 \end{cases}$$

3.10 Signal energy, Autocorrelation (deterministic case)

Note: in the present chapter, only determistic signals are considered and extensions to random signals will be presented in a subsequent chapter.

3.10.1 Signals having finite energy

Signals with finite energy are those in the functionnal space L^2. For these signals, the **energy** E_x is defined by:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \mathrm{d}t \qquad < \infty \qquad \text{(for finite-energy signals)} \tag{3.36}$$

Remark: this is a mathematical definition. The physical energy is usually obtained after multiplying by some constant having some physical unit.

The quantity $|x(t)|^2$ is a time density of energy, that is, it is an energy per time unit (density refers to the "*per* something"). To obtain an energy, one needs to integrate this density with respect to time. A similar density is defined in the frequency domain. The **Energy Spectral Density** (ESD), $S_{xx}(f)$, is a real positive quantity defined by:

$$S_{xx}(f) = |X(f)|^2 \qquad \text{(Energy spectral density, for finite-energy signals)} \tag{3.37}$$

This is an energy per frequency band and it needs to be integrated with respect to frequency for an energy to be obtained.

The **Parseval relation** states that the energy of a signal may be calculated either in the time or frequency domain:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty} |X(f)|^2 \mathrm{d}f = \int_{-\infty}^{\infty} S_{xx}(f) \mathrm{d}f \tag{3.38}$$

This is illustrated in Fig. 3.9.

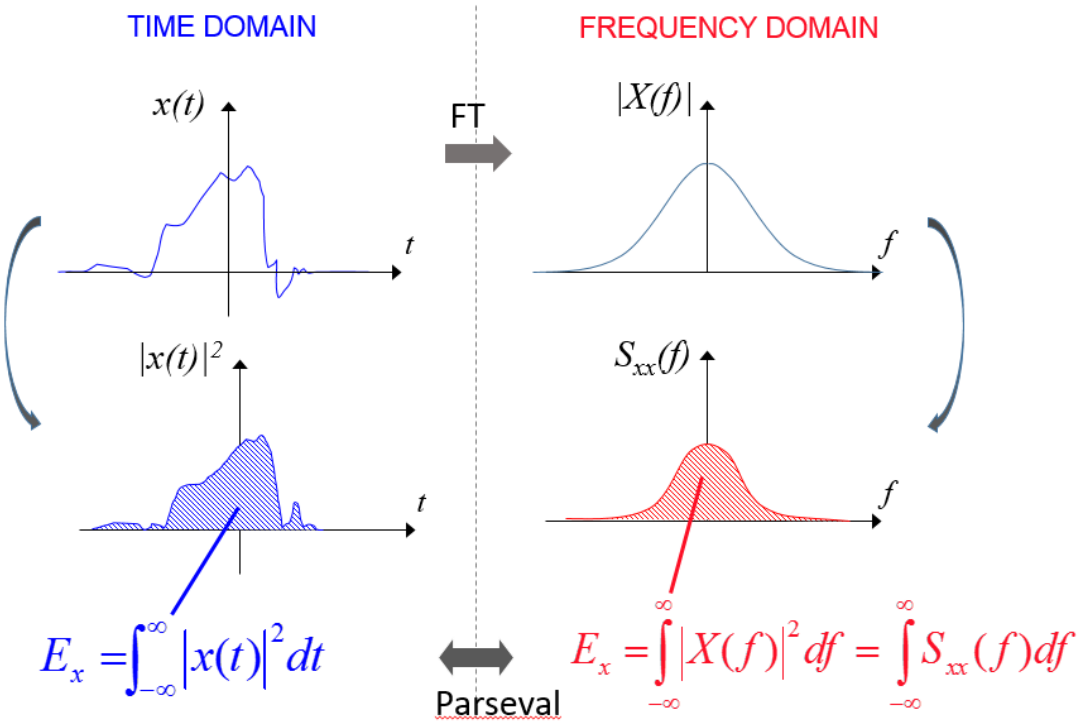


Figure 3.9: Illustration of the Parseval theorem. The energy of the signal can be calculated either directly in the time domain (left) or in the frequency domain using the energy spectral density (right). The link between the time and frequency domains is provided by the Fourier transform.

In the formula above, the integration in the frequency domain is performed for all frequencies ranging from $-\infty$ to ∞. It is also possible to calculate the energy contained in some specified frequency band. Suppose the signal is band-limited with maximal frequency 600Hz. The energy contained in the bands (200Hz;400Hz) and (-400Hz:-200Hz) is obtained by integrating the energy spectral density in the corresponding bands (see Fig. 3.10a):

$$E_{[200-400]} = \int_{-400}^{-200} S_{xx}(f)\mathrm{d}f + \int_{200}^{400} S_{xx}(f)\mathrm{d}f$$

The energy so calculated corresponds to components having between 200 and 400 cycles per second, including positive and negative frequencies (see section 2.5), both of which are included in the calculation. Note that the signal $x(t)$ is here assumed to be real. As a result, its Fourier transform has hermitian symmetry ($X(-f)=X^*(f)$, see Table 3.1), meaning the energy spectral density is even ($S_{xx}(-f)=S_{xx}(f)$). For this reason, a **one-sided energy spectral density** defined only for positive frequencies is sometines defined for real signals. This spectral density, defined only for $f \geq 0$, is given by:

$$S_{xx}^{1sd}(f) = 2S_{xx}(f) \qquad \text{one-sided energy spectral density, defined for } f \geq 0$$

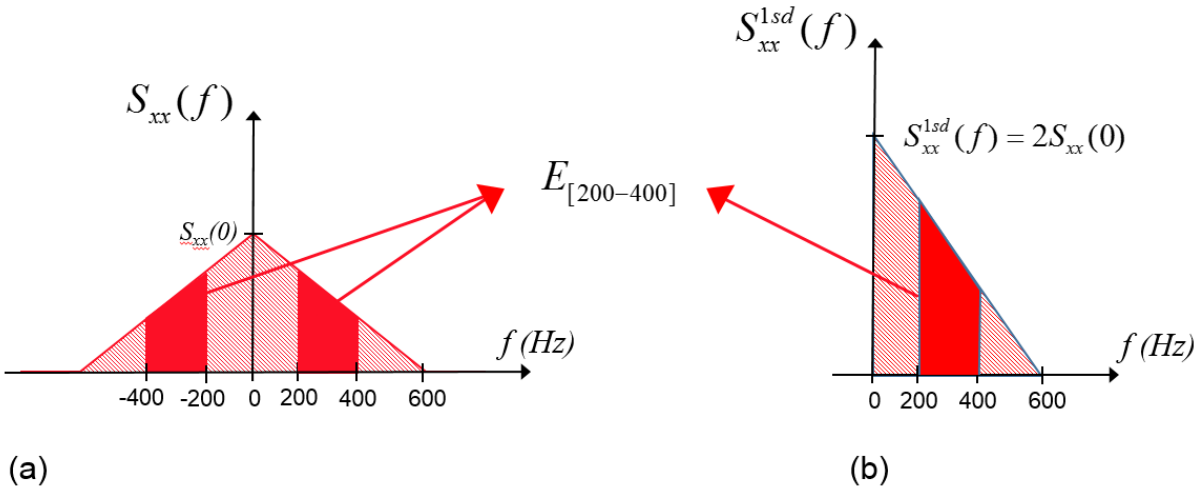


Figure 3.10: Calculation of the energy contained in the signal Fourier modes having between 200 and 400 cycles per second: (a) using the two-sided energy spectral density; (b) using the one-sided energy spectral density.

By contrast, S_{xx} is called a **two-sided energy spectral density**, defined both for negative and positive frequencies. When using the one-sided energy spectral density, one calculates the energy as:

$$E_x = \int_0^\infty S_{xx}^{1sd}(f)\mathrm{d}f$$

In that case, the energy contained by components having between 200 and 400 cyles per second is (see Fig. 3.10b):

$$E_{[200-400]} = \int_{200}^{400} S_{xx}^{1sd}(f)\mathrm{d}f$$

3.10.2 Autocorrelation of signals having finite energy

For a *real* signal $x(t)$ having finite energy, the **autocorrelation function** is defined by:

$$C_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)\mathrm{d}t \qquad \text{(autocorrelation of a real signal having finite energy)} \quad (3.39)$$

The autocorrelation depends on a time τ called the lag. It is computed by integrating the product of the signal and the signal translated by the lag. The autocorrelation is even. It is maximal at lag τ=0 with a value equal to the energy of the signal:

$$C_{xx}(\tau = 0) = E_x \quad (3.40)$$

As recalled in the general introduction to the course, the correlation between two signals is a means to assess whether they resemble each other. For the autocorrelation, one wants to know whether the signal resembles a translated version of itself. It thus takes a large value if τ is such that $x(t)$ and $x(t+\tau)$ resemble each other and a low value otherwise. This may be useful to detect the echoes in a signal for example (see Fig. 3.11).

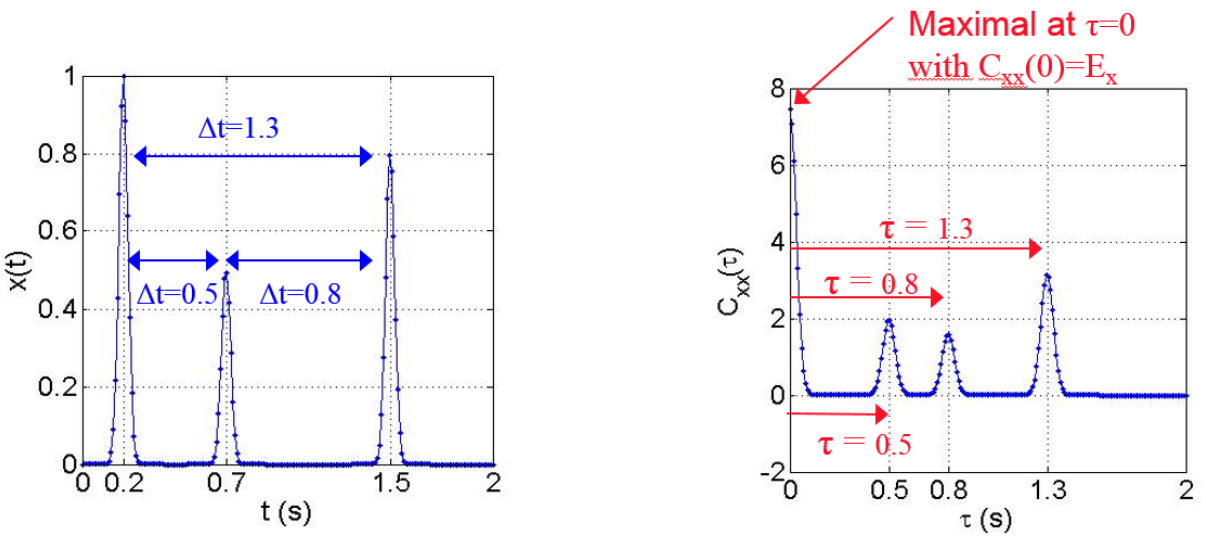


Figure 3.11: Left: a pressure signal with a main peak and some echoes. Right: the autocorrelation of the signal in the left part. The autocorrelation is maximal at τ=0, and has local extrema at lags at which there is an echo in the signal.

In addition, there exists a relation between the autocorrelation and the spectral energy density which makes the autocorrelation attractive. The **Wiener-Khintchine** theorem (to be remembered) states that the energy spectral density is the Fourier transform of the autocorrelation:

$$S_{xx}(f) = \mathrm{FT}\left[C_{xx}(\tau)\right] = X^*(f)X(f) = |X(f)|^2 \qquad \left(\begin{array}{l}\text{Wiener-Khintchine theorem}\\ \text{Signals with finite energy}\end{array}\right) \tag{3.41}$$

Note: this theorem is usually known under this name for random processes and will be mentionned again in the chapter on random processes. It remains true and much more easy to demonstrate for the deterministic signals considered in this chapter.

> ***Exercise:*** *For a signal $x(t)$, what is the definition of the autocorrelation? Show that it is an even function. Consider the signal $x(t)$=$\Pi_T(t)$. Calculate the autocorrelation and its Fourier transform. What is the Fourier transform of $x(t)$? Is the Wiener-Khintchine theorem verified in this case?*

3.10.3 Signals having finite power

Many signals we are going to work with have not a finite energy (a sine for example). However, in many instances, a finite **power** can be defined by:

$$P_x = \lim_{T\to\infty}\frac{1}{T}\int_0^T |x(t)|^2 \mathrm{d}t \qquad < \infty \quad \text{(signals with finite power)} \tag{3.42}$$

eventhough E_x=∞.

For signals with finite power, the definition of the autocorrelation needs to be modified:

$$C_{xx}(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)x(t+\tau)\mathrm{d}t \qquad \left(\begin{array}{l}\text{Autocorrelation for real signals}\\ \text{having finite power}\end{array}\right) \tag{3.43}$$

The autocorrelation at lag τ=0 is now the power of the signal: $C_{xx}(0) = P_x$

The energy density in the frequency domain is now the **Power Spectral Density** (PSD):

$$S_{xx}(f) = \lim_{T\to\infty} \frac{1}{T} X_T^*(f) X_T(f) = \lim_{T\to\infty} \frac{1}{T} |X_T(f)|^2 \qquad \left(\begin{array}{l} \text{Power Spectral Density} \\ \text{Signals with finite power} \end{array} \right) \tag{3.44}$$

where X_T is the Fourier transform of the signal after windowing on the time interval $(0, T)$:

$$X_T(f) = \mathrm{FT}\left[x(t)\Pi_T(t)\right]$$

Compared with the signals with finite energy, what remains unchanged is the Wiener-Khintchine theorem: the power spectral density is the Fourier transform of the autocorrelation:

$$S_{xx}(f) = \mathrm{FT}\left[C_{xx}(\tau)\right]$$

The Parseval relation now tells that the power of the signal may be calculated by integrating the PSD with respect to frequency:

$$P_x = \lim_{T\to\infty} \frac{1}{T} \int_0^T |x(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty} S_{xx}(f) \mathrm{d}f$$

In the same fashion as for finite-energy signals, a one-sided PSD can be defined.

For signals with finite power, all quantities are thus defined by considering finite-duration intervals, and by taking the limit when the interval duration tends to ∞. Note that for an interval of finite duration, the energy is defined, and the power is obtained by dividing this energy by the interval duration.

3.10.4 Summary

Figure 3.12 provides a summary of the energy relations.

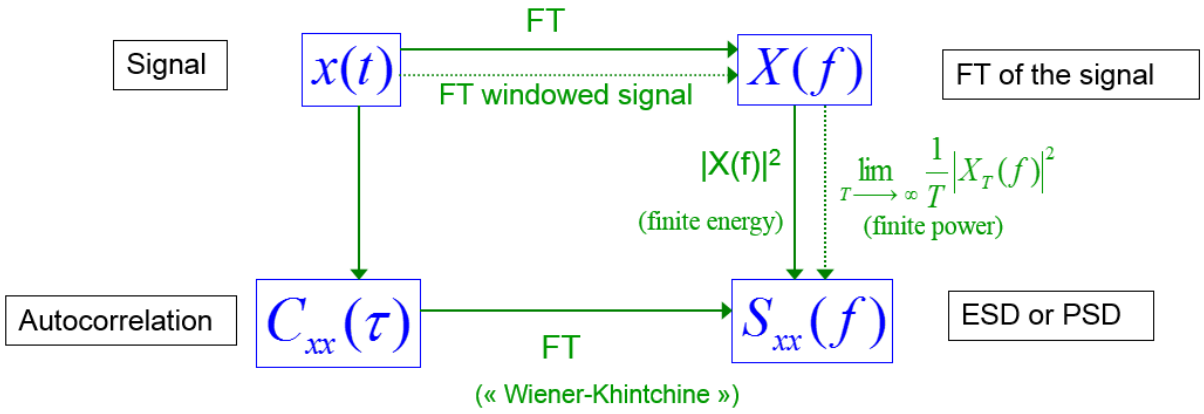


Figure 3.12: Two paths for calculating the energy or power spectral density of a signal. On the upper path, the dotted arrows correspond to signal having finite power while the plain arrows correspond to signals having finite energy.

3.11 Windowing

A signal $x(t)$ is always measured over some finite time interval T. This is equivalent to multiplying the signal by the rectangular function $\Pi_T(t)$. What we get is the windowed signal $x_w(t)$ given by:

$$x_w(t) = x(t) \cdot \Pi_T(t) \tag{3.45}$$

More generally, an arbitrary window $w(t)$ can be used as well, with:

$$x_w(t) = x(t) \cdot w(t) \tag{3.46}$$

The spectrum is obtained by taking the Fourier transform of the latter equation and using the convolution theorem:

$$X_w(f) = X * W(f) \tag{3.47}$$

The FT of the windowed signal is thus the convolution product of the signal FT with the window FT. The convolution with W is equivalent to blurring and the spectrum so obtained contains information on the (blurring) window. Note: to have $X_w = X$ (which is one would like to get ideally), ones needs $W(f) = \delta(f)$. This corresponds to $w(t) = 1\ \forall t$. This means that no window is used or equivalently that an infinite observation time is possible.
The effect of windowing is illustrated in Fig. 3.13. In the top row, an everlasting harmonic signal has a Dirac in its spectrum. The middle row shows the same signal after windowing using a rectangular window, and its spectrum. In the time domain, windowing creates discontinuities (either in the signal or in its derivatives), meaning the windowed signal is less smooth than the original one. The dual observation in the frequency domain is that the signal FT contains more frequency components. The FT of the windowed signal in the frequency domain is now only an approximation of the original Dirac. For a rectangular window the FT of the windowed signal is a sinc function, and it contains, in addition to the main lobe, some secondary lobes whose amplitude is significant. One expects reducing the discontinuities introduced by the window by taking a smoother window. Classical smooth windows are the Hamming or Hanning windows. The bottom row in Fig. 3.13 is for a Hanning window. Compared with the rectangular window, this window falls off more gently to zero. A a result, the secondary lobes in the spectrum of the windowed signal have now a far lesser amplitude than the main lobe. However, the main lobe is wider compared to that due to the rectangular window. This is something quite general: the more the window $w(t)$ tapers off the signal in the time domain, the less oscillations $W(f)$ has in the frequency domain, but the widest its main lobe will be. There is indeed only one thing to remember about windowing:

> There is a trade-off in the frequency domain between the width of the main lobe of $W(f)$ and the amplitudes of its side lobes.
> Abrupt window $w(t)$ (such as rectangular) $\rightleftharpoons$ Fine main lobe for $W(f)$, high-amplitude secondary lobes.
> Smooth window $w(t)$ (such as Hanning) $\rightleftharpoons$ Broad main lobe for $W(f)$, low-amplitude secondary lobes.

It leaves the question as to whether one would prefer a large main lobe with no side lobes or a fine main lobe with many side lobes. A lobe (hopefully) corresponds to the detection of an harmonic component. The width of a lobe is related to the precision on the frequency, a good

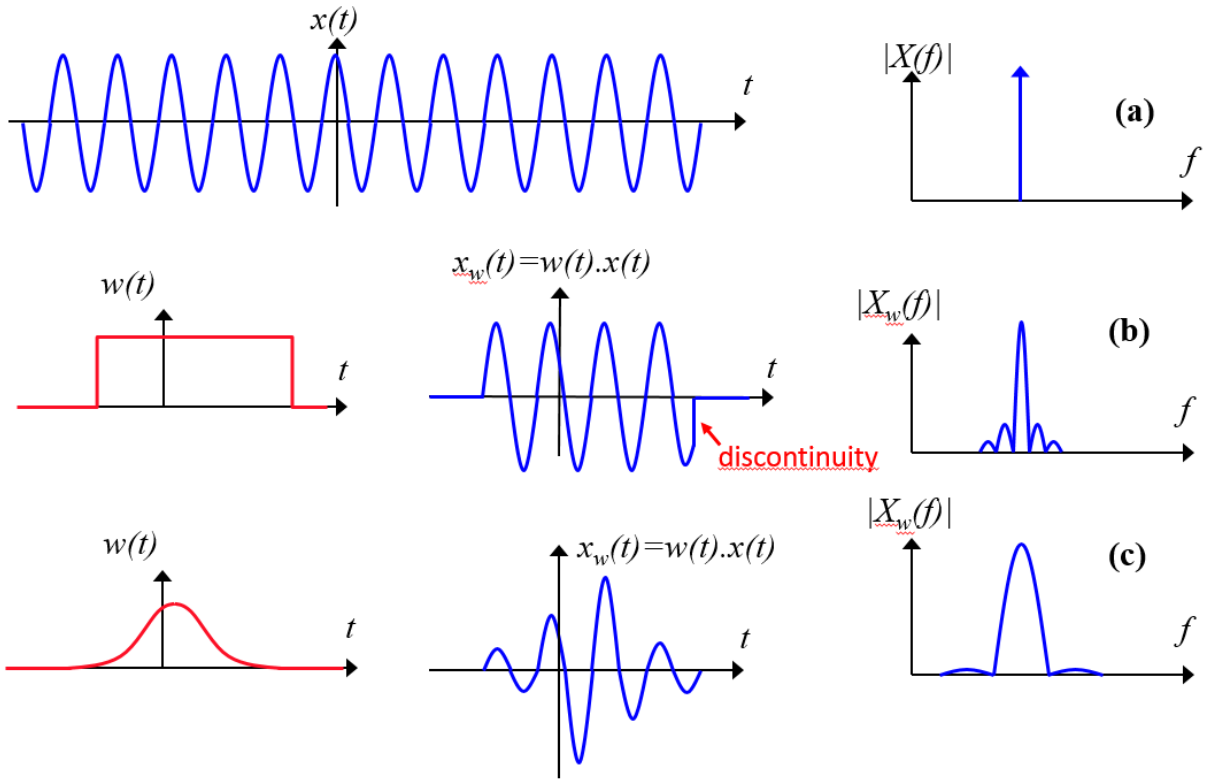


Figure 3.13: Top row: an everlasting harmonic signal and its spectrum consisting of a Dirac. Middle row: a rectangular window, the signal after windowing with the window, and its spectrum. Bottom row: a Hanning window, the signal after windowing, and its spectrum. Note: the spectrum of the original signal is given in (a). The one that is measured is given in (b) for a rectangular window, and in (c) for a Hanning window. In the real world, you do not know the spectrum in (a) (of course) since this is the one you would like to measure ! You need to infer it from the spectra in (b) or (c), the only ones that are available to you because the measurement is done over a finite time interval.

precision being obtained only for a fine lobe.

Hence, a rectangular window allows a good precision in the frequency domain (the fine main lobe in the frequency response allows detecting two close peaks), but there are many side lobes in the frequency response, increasing the risk of spurious frequency peak detection.

A smooth window such as the Hanning window allows only a low precision in the frequency domain (the wide main lobe in the frequency response increases the chance of missing peaks close to one another), but there are only weak side lobes in the frequency response, decreasing the risk of spurious frequency peak detection. For example, one would use the Hanning window when the signal FT is known to have two peaks with different amplitudes and at frequencies that are far apart in the spectrum. In that case, the peak with a small amplitude will neither be masked by the main lobe of the other peak (because it is far from it), nor will it be masked by the side lobes of the other peak (that are weak).

In general, it is advised to "try" several windows. There is no all-proof window, though. By observing a signal on a finite time interval, one looses information about this signal, and no window is going to change this. For a given window the width of the main lobe (the accuracy of frequency detection) can be reduced only by increasing the duration of the window in the time domain. This is related to a statement made earlier (when discussing the Heisenberg-Gabor principle, on page 23): a high precision on frequency requires a long observation time. So the rule is: **make as long measurements as possible** ! (This is even truer for random signals for which statistical averaging is necessary, see chapter on random signals).

3.12 Conclusion

The theory of Fourier transform tells us that any signal defined over the real line (assuming that the FT can be calculated) is made of a superposition of harmonic waves. More precisely, the FT is a "kind of projection" of the signal on harmonic waves that allows computing the complex amplitudes of these waves within the signal. As for all other transforms in this course, it is good to have an idea as to which signals it should be used for. One case when it is particularly useful is when the signal comes from a phycical system having some resonances: for measuring pure tones, the FT is very good. There are of course many more situations when the FT is useful (stationary random signals in Chapter 5 are other good candidates). Indeed, the danger is that the FT can be virtually used for any signal. However, it should be said that this decomposition into harmonics is sometimes more a mathematical fact than a physical one. Take the example of a function with a compact support (this function is zero outside a bounded interval): a fluid example would be a compactly supported vortex (in this case the spatial FT would bring the vortex from the physical space (x, y, z) to the wavenumber space). The FT says that this function is a sum of harmonic waves, each of which is not compactly supported and lasts forever. Hence, while the signal does not start before some time (some position for the vortex), it is made of waves that all start long before that time. Of course, these waves are clever and cancel each other until the time the signal becomes non-null, so that there is no inconsistency. Hence, mathematically, the FT can be used for any signal for which it can be computed. However, physically, one would say that some transient signals (with compact support) are not the best field of application for the FT. This has just to be kept in mind. This kind of argument will be given again when introducing time-frequency analysis and proposing other possible expansions.

The FT allows going from the time domain to the frequency domain, and conversely when possible. What happens in one domain (time/frequency) thus reflects what happens in the dual domain (frequency/time):
• The effective duration T_e and effective bandwidth B_e are roughly inverse proportional (Heisenberg-Gabor principle).
• The regularity (continuity, derivability) in one domain reflects the rate of decreasing at infinity in the dual domain.

The theory of time-invariant linear systems has been quickly reviewed. For these systems, the output is the convolution of the input with the impulse response of the system. In the frequency domain, due to the convolution theorem, a frequency component in the input is mul-

tiplied (regular product) by the frequency response of the system (the frequency response is the TF of the impulse response) to give the corresponding frequency component (that is, at the same frequency) in the output. For linear systems, there is no coupling between the different frequencies when the signal goes through the system.

Some energy relations have been presented. The energy in the signal can be calculated either in the time-domain or in the frequency domain, by application of the Parseval theorem. The relevant quantity in the frequency domain is the energy or power spectral density, that reveals how the energy is split into the different waves. This is usually the engineering quantity that is presented in many scientific reports. The Wiener-Khintchine theorem (that will be encountered again for random signals) says that the energy or power spectral density is the Fourier transform of the autocorrelation. There are thus two ways of calculating the spectral density: 1) calculate the Fourier transform, take the modulus, and square; or 2) calculate the autocorrelation and take the Fourier transform of it.

Finally, windowing has been presented. It should never be forgotten that as a result of measuring the signal over a finite time interval, some information is lost. The FT of the signal (or its spectral density) will in general be a blurred version of the original one. By selecting a window that more or less tapers off the signal at the limits of the measured time interval, the blurring can be modified so as to privilege frequency resolution or on the contrary reduce the possibility of false spike detection.

4 Sampling Discrete Fourier Transform

Signal processing is mainly performed in the digital world: discrete data are obtained either from analog experimental data (after sampling), or from numerical data that are discrete from the very beginning. As a result, in practice, one will mostly use the discrete Fourier transform (DFT), the digital version of the Fourier Transform. In the present chapter, fundamentals of sampling theory are reviewed (including the famous Shannon theorem). In that theory, a sampled signal is a signal multiplied by a Dirac comb, and this is not exactly a discrete signal. Thus, discrete signals are introduced next; these are nothing but sequences of numbers. For these signals, the discrete Fourier Transform and its properties are presented. The DFT expresses the fact that a discrete signal can be expanded as a sum of orthogonal discrete harmonic functions. The objective of this chapter is to provide the tools necessary to process a signal with a computer. Some examples are produced with the program Matlab. You are encouraged to reproduce these examples. Essential Matlab commands are given in Appendix B.

Classical books on the subject are those by Oppenheim and Schafer [25] and Brigham [9]. A very clear and progressive introduction is provided by Van den Enden and Verhoeckx [40]. Chapters 12-13 on FFT and its applications in the classical book by Press et al [28] are also useful.

4.1 Sampling

Sampling is a way to pick up signal values at some discrete times. How to pick up these values in practice (Analog to Digital (AD) converters) is not an easy matter. We are solely concerned by the theoretical aspect in the following. It is possible to sample a signal $x(t)$ at time $t = t_0$ by taking the product with a Dirac centered at $t = t_0$:

$$y(t) = x(t) \cdot \delta(t - t_0) = x(t_0) \cdot \delta(t - t_0) \qquad \text{(note: } \neq x(t_0)\text{!)} \tag{4.1}$$

The result, $y(t)$, is a signal which is nul almost everywhere except at $t = t_0$ where it is made up of a spike (a Dirac) with amplitude $x(t_0)$. Note that $y(t)$ is an analog signal as it still depends on the continuous variable t (completely discrete signals will be considered in section 4.2 and in all the subsequent sections). Equation (4.1) is a *model* for sampling that turns out to be fruitful

for analysis.

Usually, one does not sample a signal at a single time but rather at a sequence of equi-spaced times, with a sampling period noted T_s. The corresponding sampling frequency is $f_s = 1/T_s$ (this is also called the sampling rate). To perform sampling at all times $t_n = n \cdot T_s, n = -\infty...\infty$, one uses the sum of equi-spaced Dirac spikes called a Dirac comb, $p(t)$, defined by:

$$p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \qquad \text{(Dirac comb)} \tag{4.2}$$

The sampled signal $x_s(t)$ is obtained by multiplying the signal $x(t)$ by the Dirac comb:

$$x_s(t) = x(t) \cdot p(t) \qquad \text{(Sampled signal)} \tag{4.3}$$

This is also:

$$x_s(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) = \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s) \tag{4.4}$$

Once the sampled signal is obtained, one would like to know whether it contains all the information in $x(t)$. Equivalently, one wants to know whether it is possible to reconstruct the signal $x(t)$ from its samples $x(nT_s)$, $n = -\infty...\infty$. The answer is given by the following theorem:

Theorem 4.1.1 (Shannon, 1948)**.** If a signal $x(t)$ is such that its Fourier Transform $X(f)$ has a compact support (ie $X(f)$=0 for $|f| > B$) then it is possible to sample the signal without any loss of information. It is indeed sufficient for the sampling frequency (also called sampling rate) to satisfy: $f_s > 2B$ (if the latter condition is verified, the signal $x(t)$ can be obtained from its samples $x(nT_s)$, $n = -\infty...\infty$). B is the maximal frequency contained in the signal, it is also called the bandwidth.

Definition: the frequency $f_s/2$, half the sampling rate, will appear frequently, it is called the **Nyquist frequency**.

This theorem can be proved in the Fourier domain. In that domain, the question is turned into the following: is it possible to know $X(f)$ when one knows $X_s(f)$? From Table 3.2, we know that the Fourier transform of a Dirac comb in the time domain is a Dirac comb in the frequency domain:

$$P(f) = \mathrm{TF}[p(t)] = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_s}\right) = f_s \sum_{n=-\infty}^{\infty} \delta\left(f - nf_s\right)$$

Taking the Fourier transform of Eq. (4.3) and accounting for the convolution theorem (Eq. (3.32)) yields:

$$X_s(f) = X(f) * P(f) \tag{4.5}$$

where $X_s = \mathrm{TF}[x_s(t)]$. Hence, the Fourier transform of the sampled signal is given by the convolution product of the Fourier transform of the original signal, $X(f)$, with the Dirac comb, $P(f)$. Remember that the effect of such a product is to periodize the pattern $X(f)$ with the period f_s (the period in the frequency domain) of the comb (see Eq. (3.7) and Fig. 3.7). This is indeed a rule to be remembered:

Sampling a signal in the time domain is equivalent to periodizing its Fourier transform in the frequency domain. The converse is true: sampling in the frequency domain is equivalent to periodizing in the time domain (can you explain why it is so ?).

Hence, the question is now to know whether it is possible to reconstruct $X(f)$ once it has been periodized due to sampling. This depends on whether the periodized bits overlap, and this in turn depends on respective values of the maximal frequency contained in the signal and the sampling frequency, which is explained now. There are two possible cases:

Case 1 - sufficient sampling rate : $f_s > 2B$
This case is considered in Fig. 4.1. The FT of the original signal is shown in Fig. 4.1(a). The

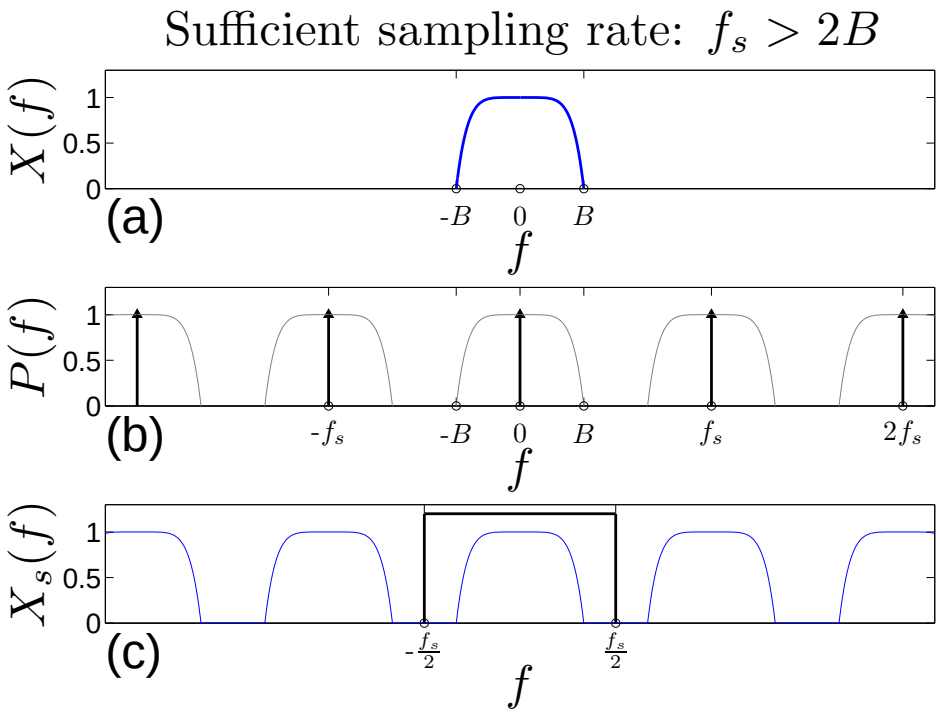


Figure 4.1: Signal sampled with a sufficient sampling rate: (a) FT of the original signal; (b) FT of the sampling comb (in light gray are indicated the different signal FTs sticked on each of the Diracs of the comb); (c) FT of the sampled signal. The thick black rectangular box in (c) represents a low-pass filter $H(f)$.

signal is bandlimited with maximal frequency B. The FT of the sampling comb is shown in Fig. 4.1(b). The spacing between the Dirac spikes in the frequency domain is f_s. The effect of sampling is to copy $X(f)$ on each of the Dirac positions. The copies are represented in a light gray shade. The spectrum of the sampled signal, obtained by adding all these copies (by Eq. (4.5)), is shown in Fig. 4.1(c). Since $f_s > 2B$, the different copies do not overlap, and it is indeed possible to obtain $X(f)$ back from $X_s(f)$. One has simply to apply a low-pass filter $H(f)$ to $X_s(f)$:

$$X(s) = H(f) \cdot X_s(f)$$

where

$$H(f) = \begin{cases} 1 & \text{if } |f| < f_s/2 \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

(actually, one may choose the filter cut-off frequency either as $f_s/2$ or as B, since $X_s(f)$=0 in between anyway.) The knowledge of $X(f)$ allows computing $x(t)$ by FT^{-1}. Hence, it is possible to reconstruct the original signal from the sampled signal. The theorem is therefore proved.

Case 2 - insufficient sampling rate : $f_s < 2B$
Let's see what happen when the conditions of the theorem are not met, which is considered in Fig. 4.2. The signal is unchanged, ie it is bandlimited with a maximal frequency B. However,

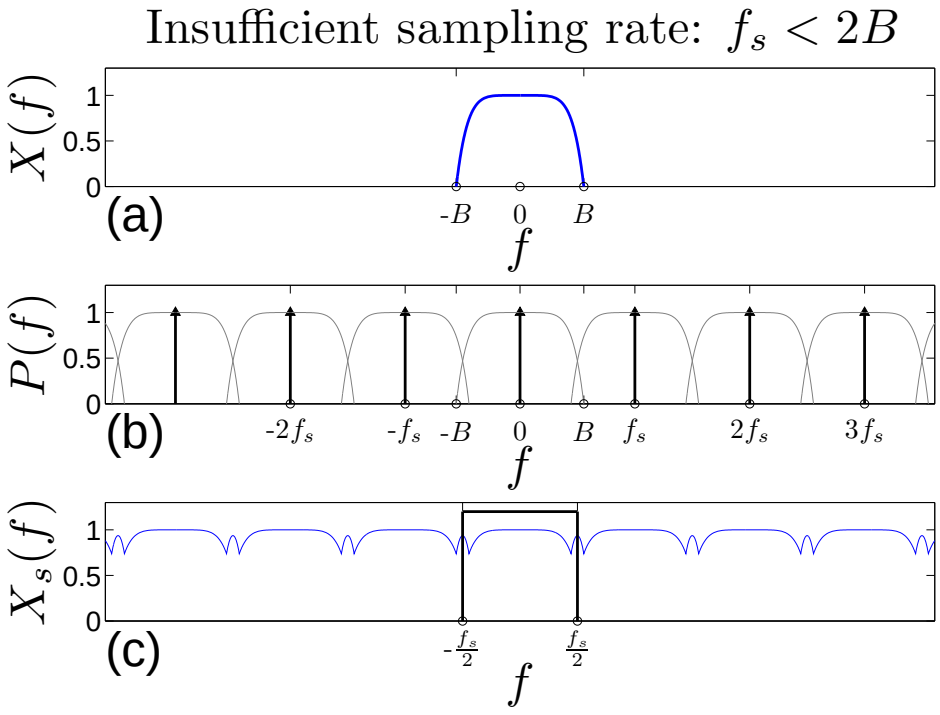


Figure 4.2: Signal sampled with an insufficient sampling rate. Same caption as in Fig. 4.1.

the sampling frequency f_s is not large enough and does not satisfy the requirement of Shannon's theorem. Compared with the previous case, the Dirac spikes are close to one another, and as a result the copies of $X(f)$ do overlap. This overlap is called *aliasing*. A low-pass filtering of the sampled spectrum in the frequency band $(-f_s/2, f_s/2)$ does not yield the original $X(f)$ anymore. It is thus in general not possible to reconstruct the original signal from its samples: information about the signal has been lost.

To be more specific, consider $X_s(f)$ in the frequency band $(f_s - B, f_s/2)$ (Fig. 4.2(b) and (c) and Fig. 4.3). It contains a contribution of $X(f)$ at these frequencies (as it should be) due to the Dirac $\delta(f)$ in $P(f)$. However, due to the overlap with the spectrum $X(f)$ that is sticked onto the Dirac $\delta(f - f_s)$, $X_s(f)$ also contains a contribution of $X(f)$ for frequencies in the band $(-B, -f_s/2)$. The latter contribution is shown in red color in Fig. 4.3. Hence, all the frequencies

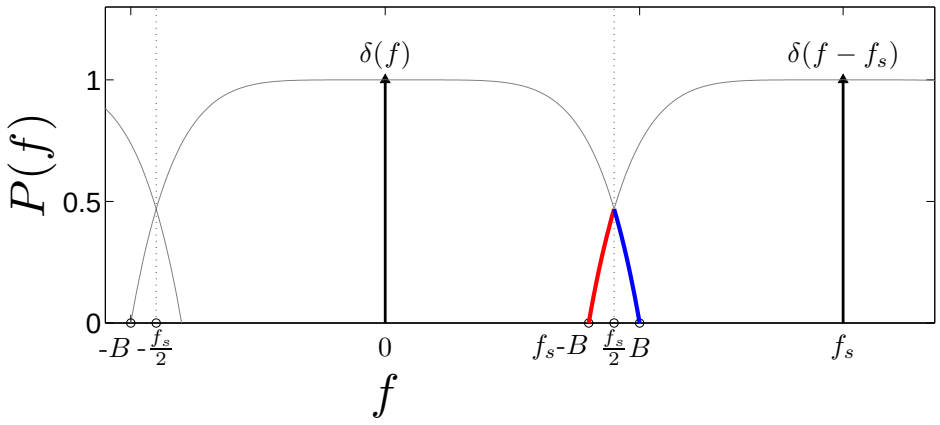


Figure 4.3: Zoom of Fig. 4.2(b).

in $X(f)$ that are beyond $|f_s/2|$ are *aliased* in the band $(-f_s/2,f_s/2)$. When using a low-pass filter in the frequency band $(-f_s/2, f_s/2)$, not only one misses the values of $X(f)$ for frequencies that are above $f_s/2$, but these frequencies spoil the result for frequencies below $f_s/2$.

Effect of aliasing
Aliasing is responsible for the appearance of *spurious* frequencies in the sampled signal. From what has been said above, a component with the original frequency f in the signal will appear as a component with the aliased frequency f_a defined by:

$$f_a = f \pm m f_s$$

where the integer m is such that: $f_a \in (-f_s/2, f_s/2)$ (if the sampling frequency satisfies the requirement of the Shannon theorem, then $f < B < f_s/2$, m=0, and f_a=f.)

Example:
One has to sample a signal $x(t) = \cos(2\pi 50t)$. The sampling rate does not satisfy the Shannon criterion: f_s=56 Hz $< 2B$=100Hz. The sampled signal contains spurious components at $\pm$ 6 Hz (see Fig. 4.4), due to $m = \pm 1$. After low-pass filtering in the frequency band (-28Hz,28Hz)

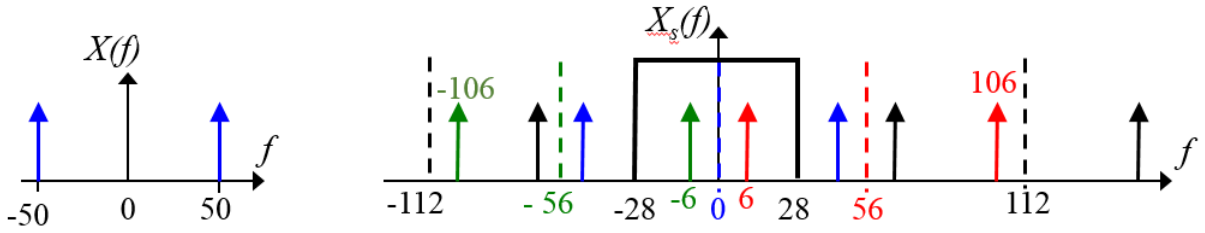


Figure 4.4: Left: spectrum $X(f)$ of the original signal; Right: spectrum $X_s(f)$ of the sampled signal with a sampling rate f_s=56 Hz.

this is the only frequency that is retained. This change of frequency is illustrated in Fig. 4.5. The original signal $x(t)$ and the samples are shown in Fig. 4.5(a). The sampled signal (with

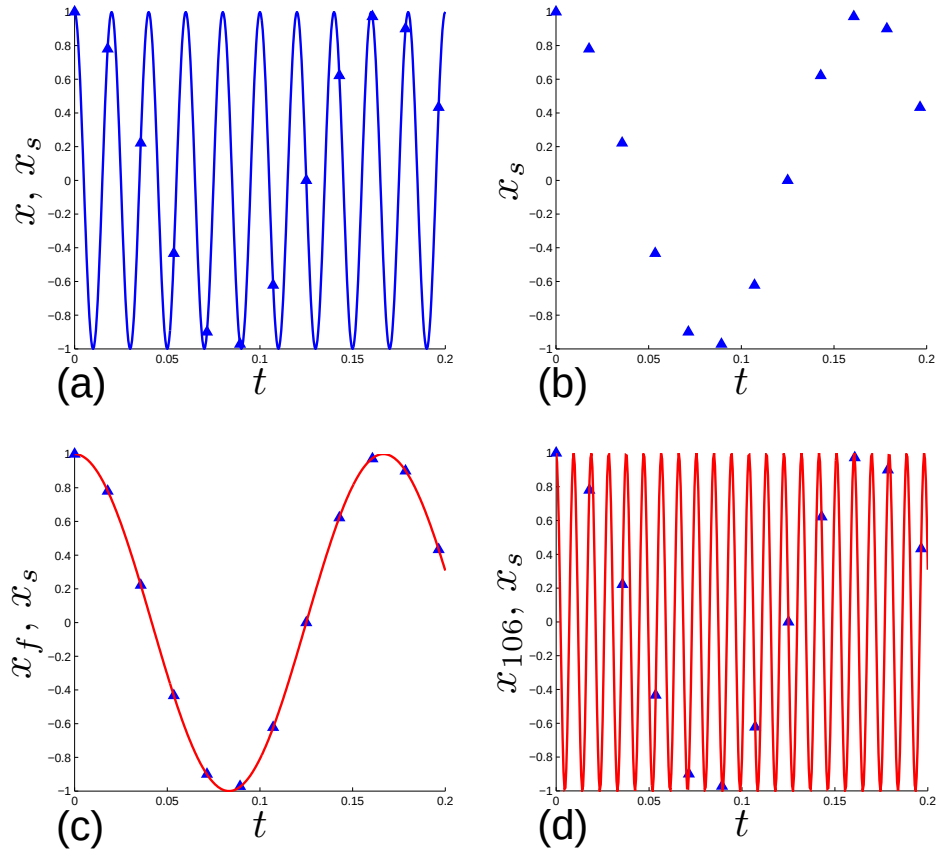


Figure 4.5: Illustration of the effect of subsampling on an harmonic signal: (a) Original signal at 50 Hz; (b) samples for a too small sampling frequency (f_s=56 Hz); (c) resulting reconstructed signal at 6 Hz (the signal in (a) is not recovered !); (d) example of another cosine at 106 Hz that could be passed within the samples. *Note that for this figure, the sampled signal is actually made of the samples indicated by a symbol multiplied by a Dirac, see Eq. (4.4).*

f_s=56 Hz) alone is shown in Fig. 4.5(b) and our problem is to recover a continuous signal from the samples. It should be noted that infinitly many cosines could be passed within the samples, since

$$\cos(2\pi f t_j) = \cos(2\pi(f + m f_s)t_j) \quad \forall j \in Z,\ \forall m \in Z$$

where $t_j = jT_s = j/f_s$ are the sampling times. These different cosines correspond to spike pairs in the right part of Fig. 4.4. For $m = 0$ the original components at $\pm$50 Hz are retrieved. For $m = \pm 1$ components at $\pm$6 Hz can be passed through the samples, as shown in Fig. 4.5 (c). For $m = \pm 2$ components at $\pm$ 106 Hz can be passed through the samples, as shown in Fig. 4.5(d). After low-pass filtering in the band (-28Hz,28Hz), the retained continuous signal has frequency 6 Hz (Fig. 4.5 (c)). Hence, we end up with a signal at 6 Hz while the original signal was at 50 Hz, as a result of under-sampling.

Exercise

Do the homework in Fig. 4.6.

Homework with Matlab: spectrogram of a chirp and aliasing

Run the following commands with Matlab:

```
fs=10000;Ts=1/fs;                % sampling frequency 10kHz
t=0:Ts:4;                        % time running from 0 to 4s with steps Ts
x=0.5*cos(2*pi.*(2000*t).*t);    % signal
sound(x,fs);                     % plays signal on loudspeaker at freq.  fs
figure
specgram(x,[],fs);               % time-frequency plot of x
```

The spectrogram plots the energy of the signal in the time-frequency plane and tells you which frequencies are present at time t. Change the sampling frequency to f_s=32kHz and observe the difference. Explain the result knowing that for a signal $x(t) = \cos(\Phi(t))$ the instantaneous frequency is given by:

$$f_{inst}(t) = \frac{1}{2\pi}\frac{\mathrm{d}\Phi}{\mathrm{d}t}$$

(plot it!).

The concept of instantaneous frequency will be considered in depth in the chapters on time-frequency analysis.

Figure 4.6: Matlab Homework: illustration of aliasing.

Are some given measurements suffering from aliasing ?

After sampling a signal, one would like to know whether this signal is spoiled by aliasing. The answer depends on the bandwidth of the original signal and the sampling frequency, and you do not know the signal (the only thing you know is the sampled signal !). Hence, "after" is too late. *After* sampling, the best you can have is a rough assessment as to whether the sampled signal is corrupted by aliasing. Figures 4.7(a) and (b) show two spectra of sampled signals. The spectrum in (a) has no aliased components, whereas the spectrum in (b) does have some. In (a) the spectrum falls off before the Nyquist frequency, while that in (b) has a small increases just before that frequency. In general, a spectrum that is not zero at the Nyquist frequency and increases near that frequency is an indication that some aliasing is present.

Aliasing can be avoided in some instances *prior to* sampling. If a signal depends on time for example and one uses a sampling frequency f_s, then one could use an analog filter (anti-aliasing filter) to suppress components having a frequency higher than f_s so that these frequencies are not aliased in the final frequency band. This filter should be used before sampling the signal, of course. Sometimes aliasing cannot be avoided, in particular when working with a spatial coordinate. For example, in a round jet, when using a ring of microphones for computing azimuthal spectra of the pressure, nothing can be done to avoid spatial aliasing (time aliasing is suppressed by anti-aliasing filters). One would just use as many microphones as possible to increase the azimuthal Nyquist frequency and assess a posteriori whether aliasing is likely or

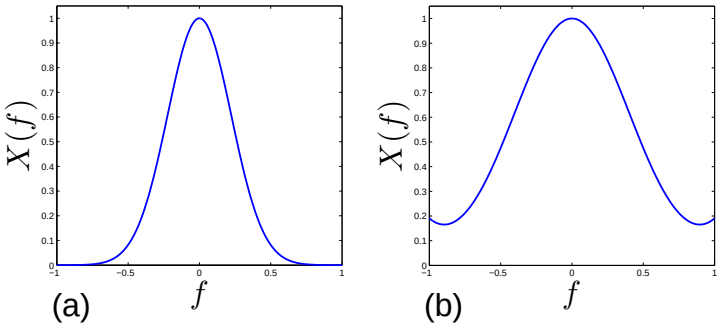


Figure 4.7: (a) a measured spectrum not affected by aliasing; (b) a measured spectrum affected by aliasing.

not.
It is worth pointing out that the Shannon theorem is valid for bandlimited signals, and the signal one needs to measure may not satisfy this requirement. In that case, aliasing is present but may be unimportant if the original spectrum decays rapidly as a function of frequency.

4.2 Discrete-time signals

4.2.1 Introduction

The signal $x_s(t)$ in the previous section is still a continuous (analog) signal for which the tools studied in the previous chapter are applicable (FT, PSD,...). From now on, we are going to work with a truly discrete-time signal $x_s[n]$, that is, a sequence of numbers indexed by an integer n. There are indeed two different things not be confused:
- the signal is given at discrete times: this is a *discrete-time* signal;
- in addition the amplitude of a signal at these discrete times could take discrete values, that is, could have a discrete amplitude (as is the case if the amplitude needs to be represented with 8 bytes for example): in that case the signal is a *digital* signal. Only digital signals can be handled and stored on computers.
In the following, we do not consider the effect of amplitude discretization: we solely consider discrete-time signals (even if sometimes the word digital is used abusively). The square bracket in $x[]$ is used to emphasize that the signal is now a discrete-time sequence. $x_s[n]$ can be imagined to be the value of the continuous signal x at time $t_n = nT_s$ where T_s is the sampling period. Hence, $x_s[n] = x(nT_s)$. In the following the subscript s that indicates that the discrete signal results from some form of sampling is removed unless it is necessary. Examples of discrete signals are:
• The discrete impulse (discrete form of the Dirac) $\delta[n]$ defined by (see Fig. 4.8(a)):

$$\delta[n] = \begin{cases} 1 & \text{if } n{=}0 \\ 0 & \text{otherwise} \end{cases} \qquad \forall n \in Z \tag{4.7}$$

Compared with the continuous-time case, the impulse is here defined without any difficulty (no distribution needed).

• The discrete Heaviside function:

$$\mathrm{he}[n] = \begin{cases} 1 & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases} \qquad \forall n \in Z \tag{4.8}$$

• The discrete exponential: $x[n]{=}e^{\alpha n}$, where $n{\in}\mathrm{Z}$, and $\alpha{\in}\mathrm{C}$. If $\alpha = \mathrm{j}\omega$ this is a complex exponential (ie a wave). Figure 4.8(b) represents the signal $x[n] = \mathrm{he}[n]e^{-n/2}$. Figure 4.8(c) represents the real part of the signal $x[n] = e^{\mathrm{j}2\pi n/7}$ which is a discrete-time wave with 7 points per period.

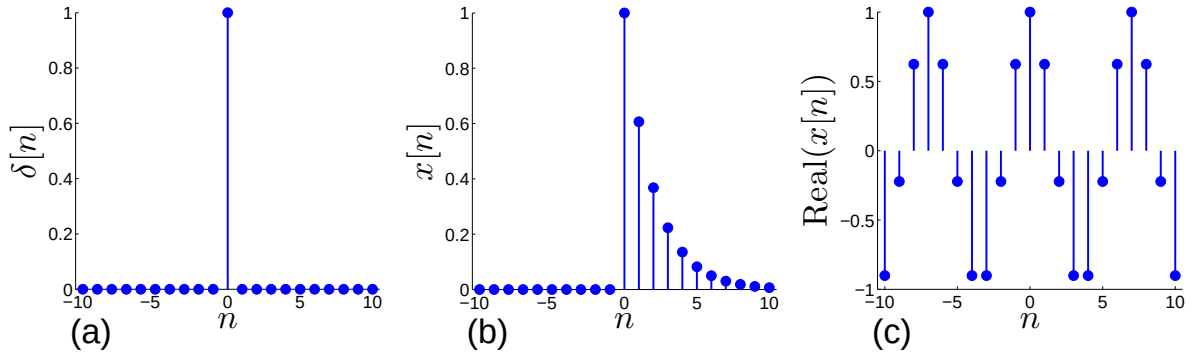


Figure 4.8: . (a) Discrete impulse $\delta[n]$; (b) discrete exponential multiplied by a Heaviside function, $x[n] = \mathrm{he}[n]e^{-n/2}$; (c) real part of a discrete wave, $\mathrm{Re}\,(\mathrm{x[n]}) = \mathrm{Re}\left(\mathrm{e}^{\mathrm{j}2\pi\mathrm{n}/7}\right) = \cos(2\pi\mathrm{n}/7)$, having 7 points per period.

4.2.2 Spectrum of discrete-time signals (DTFT)

For discrete-time signals, it is possible to obtain a continuous and periodic spectrum. This is achieved by the discrete-time Fourier transform (DTFT) that provides a *continuous* spectrum out of a *discrete* signal. This is given by:

$$X(f) = \mathrm{DTFT}[x[n]] = \sum_{n=-\infty}^{\infty} x[n]e^{-j2\pi f n T_s} \tag{4.9}$$

One may show that this is periodic in f of period f_s, and the useful interval is of course $(-f_s/2; f_s/2)$. There is also an inverse, IDTFT, that gives the discrete signals back from a continuous periodic spectrum. No more is said here, the DTFT will be used below only when explaining zero-padding.

4.3 Discrete Fourier Transform

4.3.1 Definition of the DFT and IDFT

In the examples above, the integer n can take any value from $-\infty$ to ∞. For a sequence to be stored on a computer, n has to belong to a finite set. Hence, we will now consider sequences with N points, that is, sequences $x[n]$, $n = 0...N-1$. An example of such a sequence is shown in Fig. 4.9. This is a discrete impulse with $N{=}4$ points; observe the difference with Fig. 4.8(a)

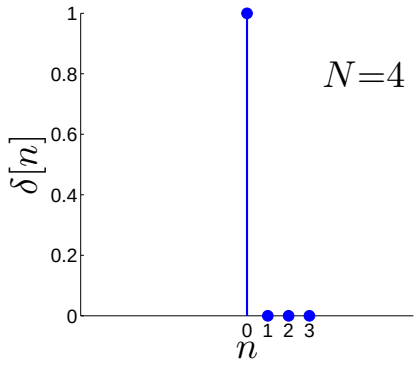


Figure 4.9: Discrete impulse $\delta[n]$ for a N=4.

where the impulse is defined over $n \in Z$. The discrete Fourier transform $X[k]$ of the signal $x[n]$ needs also to be a discrete signal (as it also needs to be handled by a computer), and as we are going to see, it is also of length N. Hence, the discrete Fourier transform (DFT) and its inverse (IDFT or DFT^{-1}) establish a link between two discrete signals with N points:

$$x[n], n = 0...N-1 \quad \underset{\text{IDFT}}{\overset{\text{DFT}}{\rightleftharpoons}} \quad X[k], k = 0...N-1$$

$x[n]$ is the signal in the time domain, and its transform $X[k]$ belongs to the frequency domain. In the following, we will generally use an index n when the signal belongs to the time domain, and an index k when it belongs to the frequency domain.

We start with the definition of the inverse discrete Fourier transform (IDFT or DFT^{-1}):

$$\boxed{x[n] = \text{IDFT}\left[X[k]\right] = \frac{1}{N}\sum_{k=0}^{N-1} X(k)e^{\mathrm{j}2\pi kn/N} \qquad \text{(IDFT)}} \tag{4.10}$$

It says that any signal $x[n]$ with N points is a sum of N discrete harmonic functions $b_k[n]$ defined by:

$$b_k[n] \overset{\text{def}}{=} e^{\mathrm{j}2\pi kn/N} \qquad k = 0...N-1 \tag{4.11}$$

Said otherwise, the family $b_k[n]$, $k = 0...N-1$, is a basis for the space of signals with N points. In addition, this basis is orthogonal, meaning the scalar product between two different functions is null. Specifically, defining the scalar product between two N-point sequences by:

$$\langle x[n], y[n]\rangle = \sum_{n=0}^{N-1} x[n]y^*[n] \qquad \text{(scalar product)} \tag{4.12}$$

where $*$ indicates the complex conjugate, we have:

$$\langle b_k[n], b_l[n]\rangle = \sum_{n=0}^{N-1} e^{\mathrm{j}2\pi(k-l)n/N} = N\delta_{kl} \qquad \forall k = 0...N-1, l = 0...N-1 \tag{4.13}$$

This comes from the more general relation:

$$\boxed{\sum_{n=0}^{N-1} e^{\mathrm{j}2\pi(k-l)n/N} = \begin{cases} N \text{ if } k = l + qN,\ q \text{ integer} \\ 0 \text{ otherwise} \end{cases}} \tag{4.14}$$

which is true for any integer k or l (not necessarily in the range 0...N-1).

In expansion (4.10), each of the basis functions has a complex amplitude $X[k]$. The complex amplitude $X[k]$ is calculated by computing the DFT:

$$\boxed{X[k] = \mathrm{DFT}\left[x[n]\right] = \sum_{n=0}^{N-1} x(n) e^{-\mathrm{j}2\pi kn/N} \qquad \text{(DFT)}} \tag{4.15}$$

This amounts to projecting the signal onto the members of the basis using the scalar product. One has indeed:

$$X[k] = \langle x[n], b_k[n] \rangle$$

The complex conjugate in the scalar product definition, Eq. (4.12), is responsible for the minus sign in the exponentials in Eq. (4.15). The situation is formally the same as for the Fourier series and the Fourier transform.

Figure 4.10 represents the functions $b_k[n] = e^{\mathrm{j}2\pi kn/N}$ as functions of n in the particular case N=10. Any 10-point signal is a weighed sum, Eq. (4.10), of these 10 functions. A few remarks may be done. One has $b_0[n] = 1\ \forall n$. All other functions are sine-like and of zero mean. This results from the property:

$$\frac{1}{N} \sum_{n=0}^{N-1} b_k[n] = \frac{1}{N} \sum_{n=0}^{N-1} e^{\mathrm{j}2\pi nk/N} = \delta_{k0} \tag{4.16}$$

The function $b_1[n]$ is a low frequency wave, $b_2[n]$ has a higher frequency, and so on until $b_5[n]$=$b_{N/2}[n]$. The function $b_5[n]$ has alternating -1 and +1 values, which corresponds to what it referred to as grid-to-grid oscillations. This is the wave with the highest possible frequency that can be represented on a discrete grid (grid-to-grid oscillations correspond of course to the 2 points per wavelength criterion in the Shannon theorem, that is, also, to the Nyquist frequency). From $b_6[n]$ until $b_9[n]$, the frequency is decreasing again. Indeed the frequency corresponding to b_6 is the same as for b_4, the frequency corresponding to b_7 is the same as for b_3 and so on. More generally, b_k and b_{N-k} correspond to the same frequency. This can be simply understood: one has

$$b_{N-k}[n] = e^{\mathrm{j}2\pi n(N-k)/N} = e^{\mathrm{j}2\pi n(-k)/N}$$

which means that $b_k[n]$ has frequency k/N whereas $b_{N-k}[n]$ has the opposite frequency $-k/N$, which corresponds to the same rate (we will simply say they have the same frequency in the following). From the plots, it would appear that b_6 and b_4 are the same, but only the real part is shown in the figure and the imaginary part is different, so that b_6 and b_4 are different. From what precedes, in the decomposition (4.10) the coefficient $X[0]$ will bear the mean value (or, which is nearly the same, the sum) of the signal. The coefficients $X[1]$ and $X[9]$ correspond to

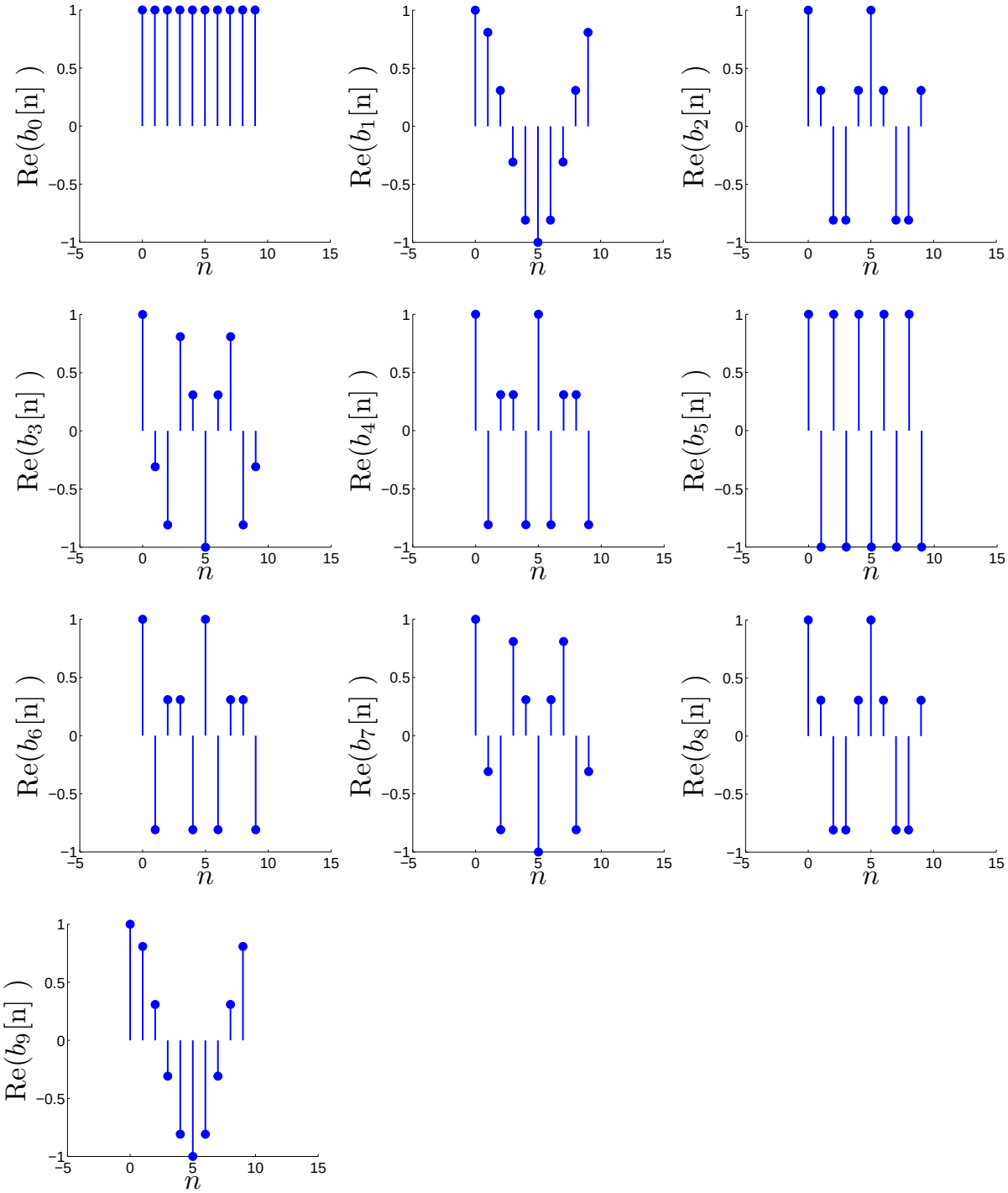


Figure 4.10: Basis for the DFT representation of a signal , in the case N=10. For every k=0 to 9, the real part of the k-th discrete wave $b_k[n] = e^{\mathrm{j}2\pi nk/N}$ is represented.

the low frequency part of the signal, $X[2]$ and $X[8]$ will correspond to a higher frequency part, and so on. $X[5]$ will give the complex amplitude of the highest frequency part in the signal. The meaning of Eq. (4.10) is that any signal having N=10 points is a linear combination of the functions shown in Fig. 4.10.

Exercise

What is the DFT of the following N-point discrete signals?
a) $x[n] = \delta[n]$; (b) $x[n] = \delta[n - n_0]$ (with $n_0 < N$); (c) $x[n] = a^n$; (d) $x[n] = e^{-n\omega_0}$;
Choose a value for N, n_0, ω_0 and check your results with Matlab.

Exercise

Use the definitions in Eq. (4.15) and Eq. (4.10) and show that $X[k]$ and $x[n]$ are both N-periodic (in this exercise n and k can take any integer value and are not to be restricted to the (0; N-1) interval).

4.3.2 Periodicity of $x[n]$ and $X[k]$

The sequences $x[n]$ and $X[k]$ have been initially defined for n=0...N-1 and k=0...N-1. However, they can be defined for any $n \in Z$ and any $k \in Z$. Indeed, using Eq. (4.10), it is easy to prove that $x[n + jN] = x[n]$, $\forall j \in Z$ and n=0...$N - 1$. Similarly, by using Eq. (4.15) one can prove that $X[k + jN] = X[k]$, $\forall j \in Z$ and k=0...$N - 1$. Hence, $x[n]$ and $X[k]$ are both N-periodic and defined for any $n \in Z$ and $k \in Z$. One period is given over N points. It should be remembered from section 4.1 that sampling in one domain (time/frequency) is equivalent to periodizing in the dual domain (frequency/time). As a result, since in the digital world both the signal and its transform are discrete, both of them also need to be periodic. This is illustrated in Fig. 4.11. One should remember that Eqs. (4.15-4.10) can be used for any integer value of n and k, respectively.

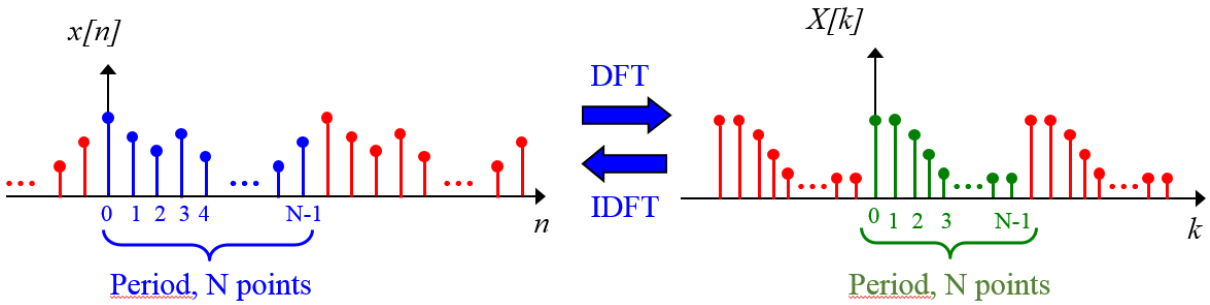


Figure 4.11: Both $x[n]$ and $X[k]$ are N-periodic, due to $X[k]$ and $x[n]$ being discrete signals, respectively.

4.3.3 Time and frequency scales

For any discrete-time sequence, we may write $x[n] = x(nT_s)$ where T_s is the sampling period. Then, we have the correspondence:

$$\text{index } n \rightleftharpoons \text{time } t_n = nT_s \quad \forall n = 0...N-1$$

whereby the index n corresponds to a time $t_n = nT_s$.
In a similar fashion, the index k corresponds to a frequency f_k and this should belong to the frequency band (-f_s/2,f_s/2) where $f_s = 1/T_s$ is the sampling frequency. Specifically, we have:
for N even:

$$f_k = \begin{cases} \dfrac{k}{NT_s} & \text{for} \quad k = 0...\frac{N}{2}-1 \\ -f_{N-k} = -\dfrac{N-k}{NT_s} & \text{for} \quad k = \frac{N}{2}...N-1 \end{cases} \tag{4.17}$$

For N odd:

$$f_k = \begin{cases} \dfrac{k}{NT_s} & \text{for} \quad k = 0...\frac{N-1}{2} \\ -f_{N-k} = -\dfrac{N-k}{NT_s} & \text{for} \quad k = \frac{N-1}{2}+1...N-1 \end{cases} \tag{4.18}$$

One should be careful about the order of the frequencies. A computer will return a DFT $X(k)$ with k=0...N-1. According the previous formula (limiting the discussion to N even), the first $N/2$ samples of $X(k)$ correspond to positive frequencies ($f_k > 0$), while the next $N/2$ samples correspond to negative frequencies ($f_k < 0$). The most positive frequency corresponds to k=$N/2$-1 and its value is:

$$f_{k=N/2-1} = \frac{N/2-1}{NT_s} = \frac{N/2-1}{N} f_s \sim \frac{f_s}{2} \qquad \text{(most positive frequency)}$$

The most negative frequency correspond to k=$N/2$ (it comes just next to the most positive one) and its value is:

$$f_{k=N/2} = -\frac{N-N/2}{NT_s} = -\frac{N/2}{N} f_s = -\frac{f_s}{2} \qquad \text{(most negative frequency)}$$

The frequency therefore lies in the frequency band (-f_s/2,f_s/2), as expected.
Figure 4.12 summarizes how one should re-order the DFT output $X[k]$ in order to have the positive and negative frequencies in the right order (most progammation languages have a command called "fftshift" or something similar to perform that re-ordering). An example is now given:
Example
Consider the sequence $x[\cdot]$=[1 7 5 6 4 3] with N=6, and a sampling frequency f_s=60Hz. The DFT returns: $X[\cdot] \sim$ [26; -4.5-4.3i; -2.5 - 2.6i; -6 ;-2.5 + 2.6i; -4.5 + 4.3i]. The signal $x[n]$ is shown in Fig. 4.13(a) as a function of n. The same signal plotted versus t_n is shown in Fig. 4.13(c). The DFT $X[k]$ is shown in Fig. 4.13(b) as a function of k. The 3 samples indicated by markers (square, circle, diamond) correspond to negative frequencies. The frequencies f_k as defined in Eq. (4.17) are listed in the following table:

$k =$	0	1	2 (=N/2-1)	3 (=N/2)	4	5
$f_k =$	0	10	20	-30	-20	-10

When X is plotted versus frequency f_k, these samples need to be shifted to the left, as shown in Fig. 4.13(d). The Matlab code for producing these figures is the following:

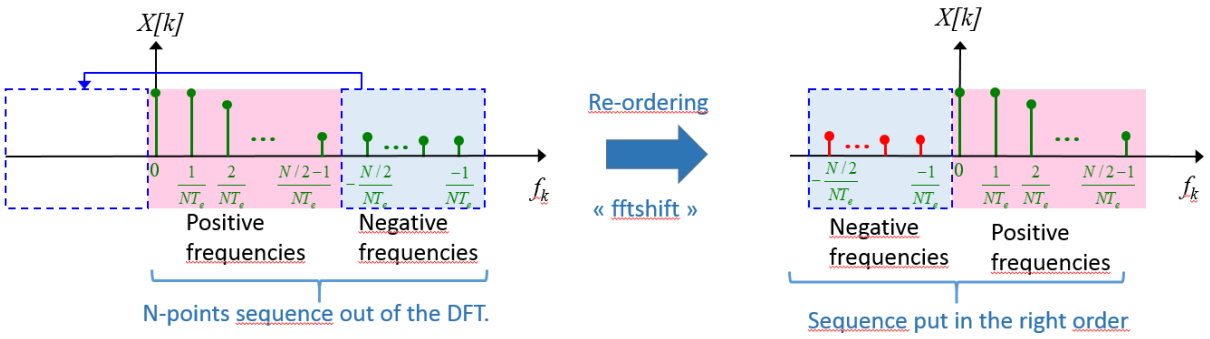


Figure 4.12: In the DFT $X[k]$, the first half of the samples correspond to positive frequencies and the second half to negative frequencies. The samples need to be re-ordered before plotting X as a function of frequency f_k.

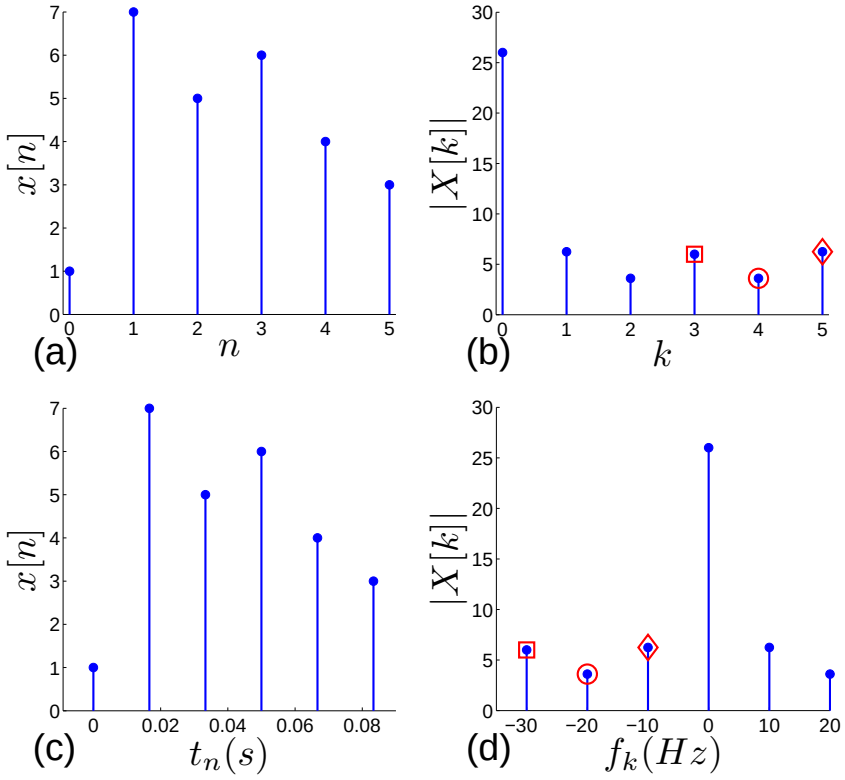


Figure 4.13: (a) A discrete signal $x[n]$ plotted versus n; (b) its DFT $X[k]$ plotted versus k; (c) $x[n]$ plotted versus t_n; (d) $X[k]$ plotted versus f_k.

% Matlab code for producing Fig. 4.13

```
x=[1 7 5 6 4 3];                % The signal.
X=fft(x);                       % Its DFT.
Xshift=fftshift(X);             % Re-ordering of the elements
```

```
                                                % to have negative frequencies first.
N=length(x);
n=(0:1:N-1);
k=(0:1:N-1);
fs=60;
Ts=1/fs;
t=n*Ts;
fk=(-N/2:N/2-1)/(N*Ts);                         % Note:  fk defined directly
                                                % from negative to positive.
                                                % Other possibility:  define fk
                                                % as in Eq. (4.17)
                                                % and use:  fk=fftshift(fk).
figure(1);
plot(n,x,'.-b')
xlabel('n');ylabel('x[n]');
figure(2);
plot(k,abs(X),'.-b')
xlabel('k');ylabel('|X[k]|');
figure(3);
plot(t,x,'.-b')
xlabel('t_n (s)');ylabel('x[t_n]');
figure(4);
plot(fk,abs(Xshift),'.-b')
xlabel('f_k (Hz)');ylabel(|X[f_k]|');
```

Finally, the different time and frequency scales need to be discussed. There are two time scales:
• the sampling period T_s, which is the small scale, and which we will call the time resolution;
• the signal duration $D = (N-1)T_s$ which is the larger time scale.
There are also two frequency scales:
• the frequency resolution Δf between two consecutive frequency samples;
• the total frequency span which is also twice the bandwidth, $f_s = 2B$, where $B = f_s/2$ is the bandwidth.
These scales are shown in Fig. 4.14 and some relations exist between them, the small scale in one domain being related to the large scale in the dual domain. The large scale in the frequency domain, f_s, is inverse proportional to the sampling period, T_s, and this is the small scale in the time domain. As you know, the maximal frequency is $f_s/2 = 1/2/T_s$, due to the Shannon theorem.
The frequency resolution Δf is related to the acquired signal duration (the large scale in the time domain). The longer the signal is observed, the smaller the first non-zero frequency produced by the DFT. However, the fact that the DFT does not display information between two samples in the frequency domain does not mean that the information is not there. It is indeed possible to obtain a value of X for any value of the frequency between $-f_s/2$ and $f_s/2$. Hence, Δf is an apparent resolution that can be improved (see below how this can be achieved using zero-padding). On the contrary, the maximal frequency $f_s/2$ is an absolute barrier that is fixed by the sampling period.

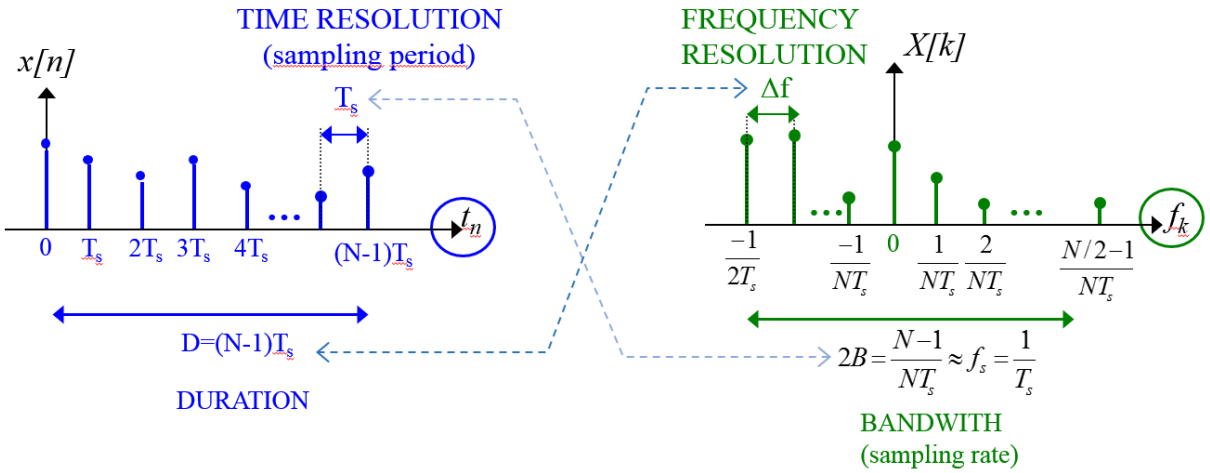


Figure 4.14: Time and frequency scales involved in the DFT.

4.3.4 Link with the Fourier transform

$x[n]$ and $X[k]$ are discrete-time sequences of N points and the DFT and IDFT have an existence in their own, ie they do not systematically result from approximating the continuous (analog) Fourier transform. For example, suppose you perform a numerical simulation of turbulence on a grid with 500^3 points, then the axial velocity along a coordinate is a 500-point sequence. This having been said, it is possible to establish a connection between the analog Fourier transform and the digital Fourier transform. For this, consider the analog signal $x(t)$, and let $X(f)$ be its Fourier transform. $x[n]$ is obtained from sampling the analog signal at times $t_n = nT_s$, that is, $x[n] = x(nT_s)$. It is recalled that T_s is the sampling period and $f_s = 1/T_s$ is the sampling frequency. For a sampled signal, we know that the frequency should take values between $-f_s/2$ and $f_s/2$. Let's take a set of frequencies f_k between 0 and $f_s/2$ by writing:

$$f_k = \frac{kf_s}{N} \qquad k = 0...N/2$$

A more complete definition of f_k was given above (see Eq. (4.17) and Eq. (4.18)). The Fourier transform of the signal, calculated at frequency $f = f_k$ is then:

$$X(f_k) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi f_k t}\mathrm{d}t \tag{4.19}$$

Now suppose the signal has been measured over a time interval $D \sim N\Delta t$ and is assumed to be zero outside this interval. The FT becomes:

$$X(f_k) = \int_{0}^{D} x(t)e^{-j2\pi f_k t}\mathrm{d}t \tag{4.20}$$

Finally, let's approximate the integral by a sum over the samples, so that $x(t)$ is replaced by $x[n]$ and t is replaced by $t_n = nT_s$, while $\mathrm{d}t$ is approximated by T_s. In addition, use is made of $f_k = kf_s/N$, and we obtain:

$$X(f_k) \sim T_s \sum_{n=0}^{N-1} x[n]e^{-j2\pi nT_s kf_s/N} \tag{4.21}$$

Using $f_s T_s = 1$ finally yields:

$$X(f_k) \sim T_s \underbrace{\sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N}}_{X(k)} \tag{4.22}$$

The approximate relation between the continuous FT and the DFT is therefore:

$$X(f_k) \sim T_s X[k] \quad \forall k \tag{4.23}$$

Hence, the FT at frequency f_k is the FFT at index k multiplied by the sampling time. The product $t.f$ in the analog transform is replaced by the product $n.k$ in the digital world.

4.4 Properties of the DFT

Many properties of the Fourier Transform have an equivalent version for the DFT.

4.4.1 Parity of the DFT for a real signal

For a real signal $x[n]$, the real part of the DFT is even and its imaginary part is odd. Only the knowledge of $X[k]$, k=0...$N/2$-1 is therefore necessary. Specifically, we have:

$$X[k] = X[N-k] \qquad \text{for} \quad k = 0...N-1 \qquad \text{(for a real signal } x[n]) \tag{4.24}$$

In terms of modulus and phase, this becomes:

$$|X[k]| = |X[N-k]| \qquad \text{(for a real signal } x[n]) \tag{4.25}$$

$$\phi_X[k] = -\phi_X[N-k] \qquad \text{(for a real signal } x[n]) \tag{4.26}$$

Hence, when plotted as a function of k, $|X[k]|$ is even with respect to $N/2$, and the phase is odd. As a function of f_k, $|X[f_k]|$ is even and $\phi_X(f_k)$ is odd.

4.4.2 Power Spectral Density (PSD) and Parseval's relation

The discrete version of the Parseval relation is:

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2 \qquad \text{(Discrete Parseval relation)} \tag{4.27}$$

This relation states that the energy computed in the time domain equals that computed in the frequency domain. As for the Fourier transform, the discrete Parseval relation leeds to the definition of a **power spectral density**:

$$S_{xx}[k] = \frac{1}{N} |X[k]|^2 \quad k = 0...N-1 \qquad \text{(Discrete power spectral density)} \tag{4.28}$$

Exercise

Prove Eq. (4.27). For this, start from the left hand side and write $|x[n]|^2 = x[n]x^*[n]$. Plug in the IDFT expansion for $x[n]$ (that is, Eq. (4.10)). Finally, use Eq. (4.14).

4.4.3 DFT and circular convolution

For two discrete sequences $a[n]$ and $b[n]$ defined for any $n \in Z$ the discrete convolution product (also called the discrete linear convolution product, as opposed to the circular convolution product to be defined below) is defined as:

$$(a * b)[n] = \sum_{k=-\infty}^{\infty} a[k]b[n-k] \tag{4.29}$$

the product $a[k]b[n-k]$ is the equivalent of the product $a(\tau)b(t-\tau)$ found in the continuous convolution product (see Eq. (3.26)). One may actually arrive at Eq. (4.29) by discretizing Eq. (3.26), to within a multiplicative factor.

With the DFT, sequences of length N are used, and they are implicitly assumed to be periodic. For two sequences $a[n]$ and $b[n]$ of length N the **circular convolution product** is noted:

$$(a \otimes b)[n] = \sum_{k=0}^{N-1} a[k]b[n-k] \qquad \text{(Circular convolution product)} \tag{4.30}$$

The reason for the term "circular" is the following: suppose N=5 and you want to calculate $(a \otimes b)[n=1]$. The term in the sum over k for k=3 is $a[1]b[-2]$. The index -2 is out of range for n=0...N-1. However, remember that the signals with N-points produced by the DFT and IDFT are actually N-periodic. Hence, one may take: $b[-2]$=$b[N-2]$=$b[3]$. In Eq. (4.30) the indexes appearing in the sum are computed modulo N. When an index goes out of range through one extremity, one goes on by taking the values at the other extremity. The term circular thus comes from the N-periodicity of the signals.

The circular convolution product being defined, the discrete (circular) convolution theorems can be given. For two N-points signals $x[n]$ and $y[n]$ with respective DFT $X[k]$ and $Y[k]$, we have:

$$\text{DFT}\left((x \otimes y)[n]\right)[k] = X[k] \cdot Y[k] \tag{4.31}$$

and

$$\text{DFT}\left((x \cdot y)[n]\right)[k] = \frac{1}{N} X[k] \otimes Y[k] \tag{4.32}$$

The DFT transforms the circular convolution product into a regular product and conversely. The first relation is useful when x is the input to a discrete LTI with discrete impulse response y. The second relation is useful when x is a signal and y a discrete window.

4.4.4 Time and Frequency shift

If $X[k]$=DFT$[x[n]]$, then the DFT of the signal translated in time is:

$$\text{DFT}[x[n-n_0]] = X[k]e^{-j2\pi n_0 k/N}$$

When a signal is translated in time, all its wave components are translated as well. This translation corresponds for each of the components to a phase shift. Hence, the DFT of the initial signal, $X[k]$, gets multiplied by a factor $e^{-j2\pi n_0 k/N}$ representing this phase shift (it depends on the frequency, k, of the component).

It is also possible to modulate the signal (ie, to multiply this signal by a harmonic signal) so as to translate the DFT of the signal in the frequency domain. We have:

$$\text{DFT}[x[n]e^{j2\pi n k_0/N}] = X[k-k_0]$$

4.5 Fast Fourier Transform

A naive implementation of the DFT would be the following:

```
do k=0,N-1
   do n=0,N-1
      X(k)=X(k)+x[n] e^{-j2πnk/N}
  end
end
```

The operation count for this implementation scales as $N \times N$, because one has to calculate N frequency components (loop k=0...N-1.) and for each of these components a sum over N samples (loop over n=0...N-1) is performed. The FFT is an algorithmic calculation of the DFT (Cooley and Tukey, 1965) that improves the computational cost: the result produced by the FFT is the DFT, but the loops are modified so that the calculation is faster (see [28], chapter 12). You need to remember:

> The FFT calculates the DFT with an operation count that scales as $N\ln(N)$.
> The number of points has to be a power of 2: $N = 2^m$.

This efficiency means that the FFT is also used to calculate more than DFTs. For example, the FFT is used for a fast calculation of discrete correlations. For example to calculate the correlation between two N-point sequences, on would use the DFT (FFT) and IDFT (IFFT). The "circular" correlation is:

$$C_{xx}[m] = \sum_{n=0}^{N-1} x[n]x[n-m]$$

where the indexes are calculated modulo N when they go out of bounds. Then:

$$C_{xx}[\cdot] = IDFT\left[|X[\cdot]|^2\right]$$

where $X[\cdot]$ is the DFT/FFT of $x[\cdot]$. Hence, to calculate the correlation, you would apply the DFT, square the modulus of the result, and take the IDFT. The resulting sequence $C_{xx}[\cdot]$ has N-points, and like x it can be seen to be N-periodic. The first $N/2$ samples correspond to positive

lags, and the $N/2$ following to negative lags. The autocorrelation needs to be re-ordered (using fftshift) in exactly the same way as $X[k]$. This is called fast correlation. Note that the circular correlation is calculated this way; if one wants to obtain the linear correlation, some zero-padding is required in addition. This is not discussed into more details in this course.

4.6 Windowing

The effect of windowing has already been studied for the continuous Fourier transform in Section 3.11. Little is changed for discrete signals. One thing is changed: for a continuous signal, the signal is not zero within some interval, and zero outside this interval; for discrete-time signals having a DFT, the signal is non-zero within an interval, and implicitly assumed periodic. Hence, for instance: for continuous-time signals, abrupt windowing was creating a jump in the signal, from the signal last-known value to zero. For discrete-time signals, the jump occurs between the signal last-known value ($x[N-1]$) and ... the first one ($x[0]$). This, as it turns out, does not lead to any significant difference of behaviour between the continuous- and discrete-time signals. However, the case of a discrete-time signal is considered rather comprehensively in the following.

Let us recall first that windowing refers to the fact of having to consider the signal over a duration shorter than its true duration. Except for transient signals having compact support, this can't be avoided when using DFT. Mathematically, the signal one is dealing with is not the original signal, but the original signal multiplied by a window:

$$x_w[n] = x[n] \cdot w[n] \tag{4.33}$$

where x_w is the windowed signal. In the case of finite-length sequences used by the DFT, a signal $x_\infty[n]$, initially defined for n=-∞...∞ is truncated into a signal $x[n]$, n=0...N-1, that is implicitly N-periodic. This trunctation is equivalent to windowing using a rectangular window ($w[n]$=1 $\forall n$=0...N-1).

Let us first consider the case of a rectangular window. An example is shown in Fig. 4.15. In this figure, a cosine $x_\infty[n] = \cos(2\pi 1 n T_s) = cos(2\pi n/6)$ is initially defined for all values of n. This signal has a frequency 1 Hz, and the sampling frequency is f_s=6 Hz. The signal x_∞ is shown in Fig. 4.15(a). The signal is restricted to an interval with N points, which gives $x[n]$, n=0...N-1. This signal is implicitly periodized due to our using the DFT. The signal $x[n]$ and its periodic copies are shown in Fig. 4.15(b) for N=6. For this value of N, one period is retained in the windowing and after periodization, the signal is exactly the original one. The spectrum of $x[n]$ is shown in Figs. 4.15(c)-(d), either as a function of k or of f_k. The spectrum contains only one frequency component at ± 1 Hz, that is, at the only frequency of the signal. Note that one frequency f_k corresponds precisely to a frequency contained in the signal, and this is due to the fact that $x[n]$ contains exactly one period of the original signal.

Let us consider now the case for N=8 in Fig. 4.16: a little bit more than one period of the original signal is retained, and after periodization the signal $x[n]$ shown in Fig. 4.16(a) is different from the original cosine signal in Fig. 4.15(a). The spectrum of $x[n]$ is shown in Figs. 4.16(c)-(d), either as a function of k or of f_k. It is seen that the spectrum does not contain a single frequency anymore. One reason is that in Fig. 4.16(d) no value of f_k corresponds to the

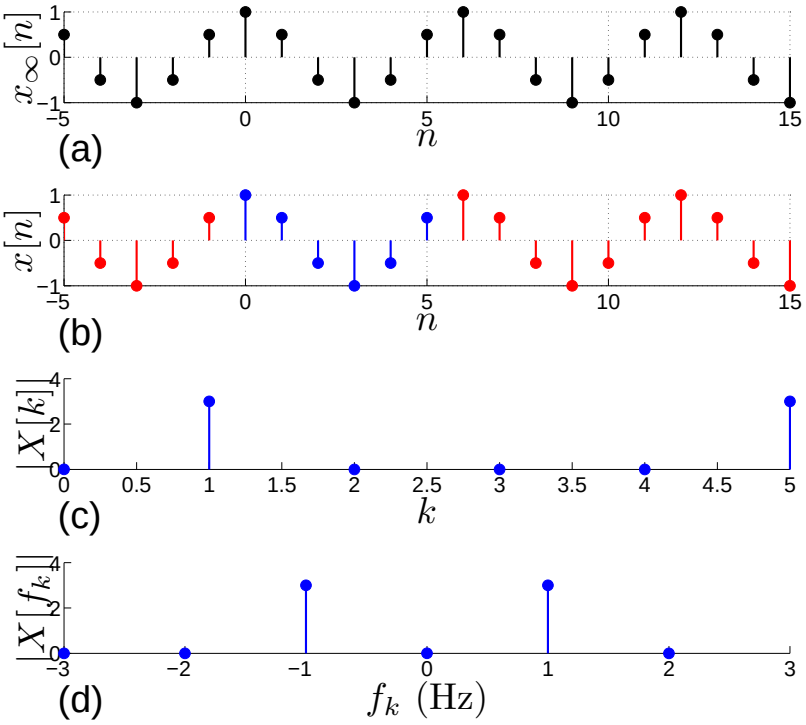


Figure 4.15: (a) A harmonic signal $x_\infty[n] = \cos(2\pi n/6)$ defined for n=-∞...∞; (b) its windowed and periodized version $x[n]$ with N=6; (c) DFT of $x[n]$ versus k; (d) DFT of $x[n]$ versus f_k.

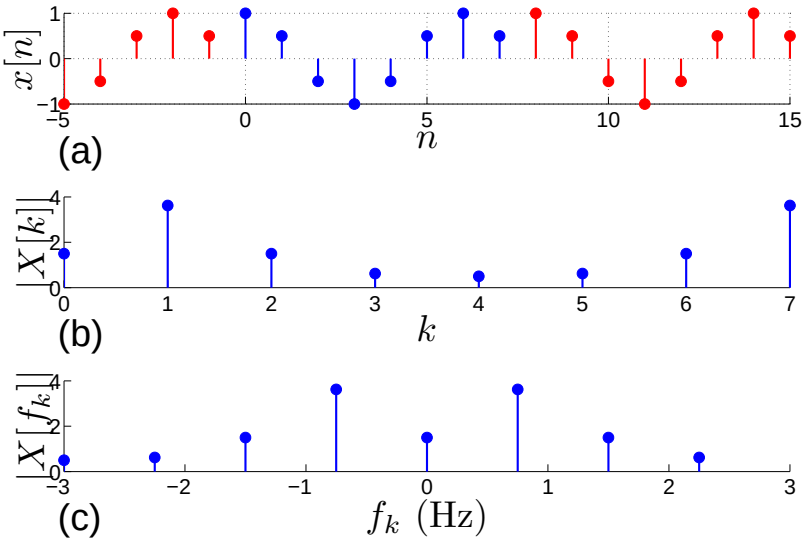


Figure 4.16: (a) The windowed and periodized version $x[n]$ of the signal shown in Fig. 4.15(a), for N=8 ; (b) DFT of $x[n]$ versus k; (c) DFT of $x[n]$ versus f_k.

original 1 Hz of the signal. Hence, this original frequency of the signal has to be spread over neighbouring values. This is called **spectral leakage**.

This gives the following rule:

> When calculating the DFT of a periodic signal, the number of points N should correspond to one period in the signal.

Of course, this is a rather theoretical result: often, you do not know the value of the period, and this is why you perform spectral analysis. However, in cases when you know the period and what you want is to obtain the respective amplitudes of the different harmonics in the signal, then the sampling frequency and the choice of N should satisfy the rule above.

The important point is that it is possible to diminish (or at least modify) spectral leakage by using a window. In what precedes, the signal $x[n]$ results from windowing the signal x_∞ by using a rectangular window. This picks up some values of x_∞ without modifying their amplitudes (as when going from Fig. 4.15(a) to Fig. 4.15(b) or to Fig. 4.16(a)). It is possible to further pre-multiply $x[n]$ by a window, as indicated in Eq. (4.33), prior to calculating the DFT. Taking the DFT of Eq. (4.33) and using the convolution theorem yields:

$$X_w(k) = \frac{1}{N} X(k) \otimes W(k) \tag{4.34}$$

This is going to modify the spectral leakage in the frequency domain. Let's see how it works on an exemple. The action of a Hanning window is considered in Fig. 4.17. The window is shown

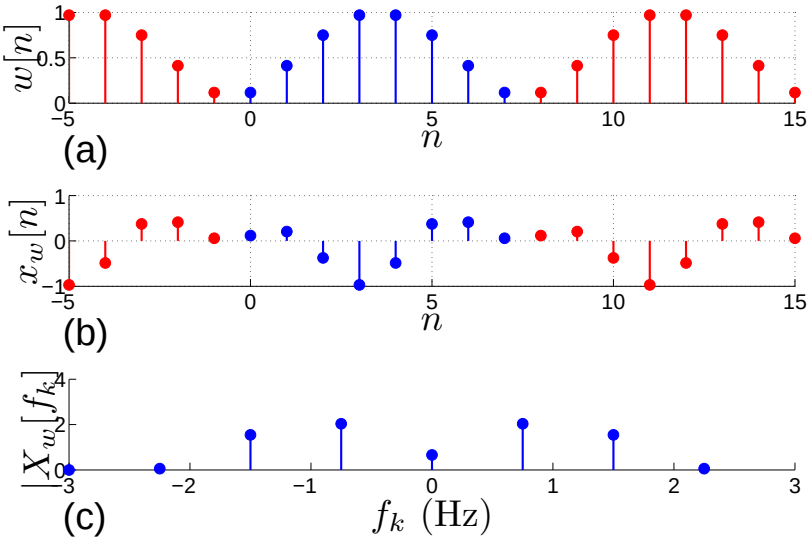


Figure 4.17: (a) The discrete Hanning window with N=8 points (one period is shown in blue color and its periodic copies in red color); (b) signal $x_w[n]$ obtained by multiplying $x[n]$ in Eq. 4.16(a) with the Hanning window; (c) DFT of $x_w[n]$.

in Fig. 4.17(a) for N=8. The resulting windowed signal is shown in Fig. 4.17(b). It is obtained by multiplying the signal in Fig. 4.16(a) by the window. The effect on the signal is to smooth out the jumps introduced due to periodization by tapering off the signal at the stitching points. The effect on the spectrum is shown in Fig. 4.17(c). Compared with the spectrum in Fig. 4.16(c)

for the rectangular window, the leakage has been modified. The frequency extent of the leakage has been reduced, meaning frequency components far from the signal original component at 1 Hz are less contaminated by that component. The samples for $f_k \geq 2$ and $f_k \leq -2$ for example are almost nul in Fig. 4.17(c) while they are not in Fig. 4.16(c). The counterpart is that what was one spike in Fig. 4.16(c) (making it clear there was only a single pure tone) is transformed with the Hanning window into nearly two spikes in Fig. 4.17(c). In that case, one may wonder whether there is one component or two in the vicinity of 1 Hz.

Finally, the effect of using a window on the spectrum is much the same as introduced in Section 3.11 for continuous-time functions. There is a balance in the spectral domain between the sharpness of the main lobe of $W[k]$ and the rate of decrease of $W[k]$ for large k. The sharpness of the main lobe means there is little leakage to immediate neighbouring frequencies. The strong decrease with k means there is little leakage to far frequencies. An abrupt window (eg rectangular) favors the sharpness of the main lobe, while a smooth window (eg Hanning) favors a strong decrease with large k.

There is one last point that needs to be mentioned: using a window decreases the energy of the signal. This is the case because when $x_w[n] = x[n]w[n]$ then $x_w[n]^2 = x[n]^2 w[n]^2 \leq x[n]^2$where it is assumed that $w[n] \leq 1$. As a result, one has:

$$\sum_{n=0}^{N-1} x_w^2[n] < \sum_{n=0}^{N-1} x^2[n]$$

the first being the energy of the windowed signal, and the second the energy of the signal without windowing (or with windowing using a rectangular window). When computing spectral power densities, a corrective factor C_w needs to be applied to recover the energy that has been lost due to windowing (see Lab 1).

Exercise

The aim of this exercise is to plot the frequency response of two different windows: the rectangular window and the Hanning window. The Hanning window is given by:

$$w(n) = 0.5\left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right)$$

The quantity C_w is defined by (see Lab 1):

$$C_w = \frac{N}{\sum_{n=0}^{N-1} |w(n)|^2} \tag{4.35}$$

It is the inverse of the window's power and depends on the window that is used.
a) Create a Hanning window w of size N=128 (use the `hanning` Matlab command). Plot it.
b) Check that $C_w \sim$2.67 for that window. What is it for the rectangular window?
c) We would like to study the first side-lobe attenuation for the Hanning window. To

have a good frequency resolution we append some zeros to the Hanning window (zero-padding). In the following command lines, the window has initially a size 128 and trailing zeros are added to obtain a length 2048. This is done directly in the `fft` command.

```
N=128;
Npadding=2048;
w=hanning(N);
W=fft(w,Npadding);
W=fftshift(W);
f=(-Npadding/2:Npadding/2-1)/Npadding;
figure
plot(f,20*log10(abs(W)./max(abs(W))),'b')
```

Here w is the window, and W is its DFT transform. In the last command, the DFT $W[k]$ is normalised so that its value is 1 at zero frequency. The frequency is normalized with respect to the Nyquist's one, and takes values between -0.5 and 0.5.
Create a new m-file that includes the above commands. Run this program.
What is the effect of the Matlab command `fftshift`?
The gain in dB is defined by $20\log(|W|)$. What is the difference of gain in dB between the main lobe and the first side-lobe?
On the plot, add the result obtained for a rectangular window (in that case, take `w=ones(N,1)`). What trade-off do these curves illustrate?
Hint: for one window the difference is about 31dB, for the other it is about 13dB.

4.7 Zero-padding

Taking as input a signal with N points, the DFT returns N points $X[k]$ in the frequency domain. They correspond to N equi-spaced frequencies in the interval $[-f_s/2\ f_s/2]$. As we have seen above, this leads to an apparent frequency resolution $\Delta f = f_s/N = 1/(NT_s) \sim D$, and the smallest non-zero frequency is Δf. However, this lowest value is more a "technical" limitation than a theoretical one. Indeed, there is no theoretical lower bound on the frequency that can be calculated: we can get a value of $X[f_k]$ for an as-low value of f_k as we may desire. By contrast, there is an upper bound which is the Nyquist frequency, $f_s/2$.

For discrete-time signals, we have seen that it is possible to get a frequency spectrum over the continuous frequency range $(-f_s/2; f_s/2)$ using the DTFT given in Eq. (4.9). The DFT, working with only N frequency points, simply does not not display this information. Actually, the DFT may be shown to be the sampling of the DTFT at N equispaced points. The question is then to know how to display the hidden information using only the DFT. This is where zero-padding comes in. Recall that $\Delta f = f_s/N$. From this relation one can imagine reducing Δf by increasing N. N in theory is fixed, since it is the number of samples we have at our disposal. However, it is possible to increase this number by adding trailing zeros to the signal, which is known as the zero-padding technique.

An example of how this works is shown in Fig. 4.18. Initially, we have a sequence $x[n] =$ [1, 2, 2, 1], with N=4, and f_s=1000 Hz. This sequence is shown in Fig. 4.18(a). Its DFT is plotted versus f_k in Fig. 4.18(b). It contains N=4 samples as well. Also plotted in the figure is

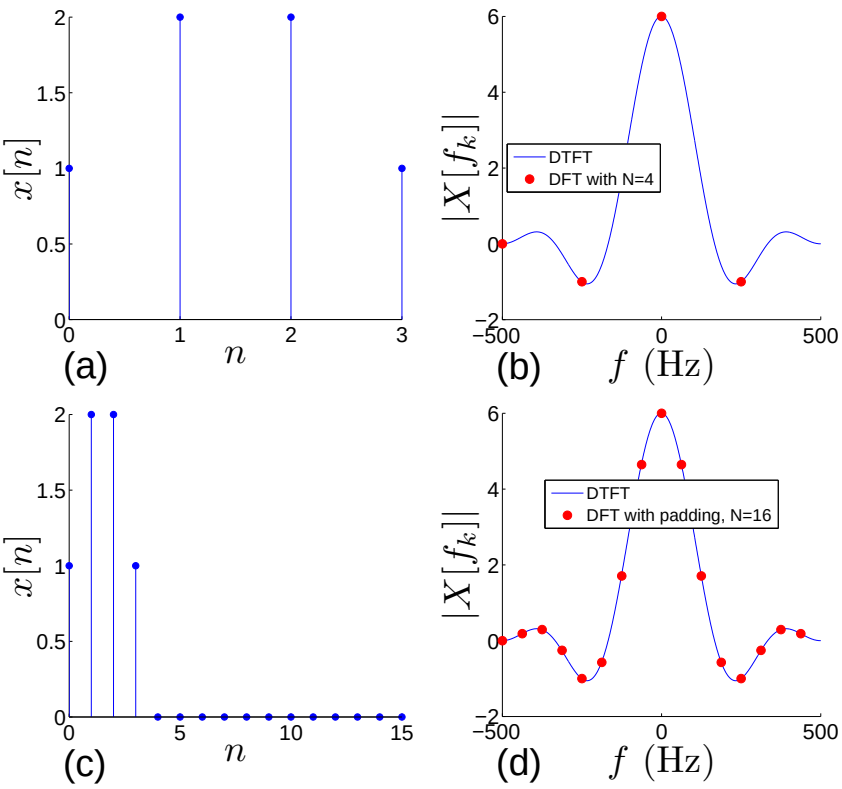


Figure 4.18: (a) A N=4 point signal; (b) its DFT; (c) the same signal padded with 12 zeros, so that altogether there are now N=16 points; (d) the DFT of the zero-padded signal.

the spectrum obtained by the DTFT. In that case the DTFT is computed from:

$$X(f) = \text{DTFT}[x[n]] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi f n T_s}$$

(that is, to calculate the DTFT, the signal is assumed to be null outside the interval 0...N-1. Recall that the DTFT, unlike the DFT, takes a signal defined over n=-∞...∞, and no periodicity is assumed, since the spectrum is continuous). The N=4 samples of the DFT fall on the continuous DTFT, but they do not catch the details of this curve. To improve the apparent resolution of the method, the signal is padded with 12 zeros, so that totally the DFT is performed on N=16 points. Be aware though that over this 16 points, only 4 contains information about the signal! The padded signal is shown in Fig. 4.18(c). Its spectrum is shown in Fig. 4.18(d) and the 16 frequency samples now give a better idea of the continuous spectrum provided by the DTFT. The maximal frequency, $f_s/2$=500 Hz, has not changed.

4.8 Overview of the different steps leading to the DFT process

Figure 4.19 from Brigham [9] gives a nice summary of the different steps involved when taking a signal from the analog world up to the DFT process.

The continuous signal first needs to be sampled and windowed:
(a) analysed signal (finie energy, or finite power)
(a') FT of the signal
(b) Dirac Comb for time sampling
(b') FT of the comb (b) $\rightarrow$ a frequency comb
(c) sampled signal (product of (a) and (b))
(c') FT of the discrete signal (c), obtained by convolution of (a') and (b')
(d) time window
(d') FT of the window
(e) sampled windowed signal, product of (c) and (d)
(e') FT of (e), obtained by convolution of (c') and (d')
For getting a DFT, a frequency sampling is also required
(f') Frequency sampling with a frequential Dirac comb
(f) IFT of the comb (a time comb)
(g') Spectrum sampled (product of (e') and (f'))
(g) Signal sampled and windowed, whose spectrum has been sampled as well (convolution of (e) and (f))
Two losses of information appear in the process: one due to sampling, one due to windowing (that is due to observing the signal over a finite duration time interval). Hence, measure you signal over as long durations as possible (as already mentioned in Section 3.11) with an as high sampling frequency as possible.

4.9 Examples

In the present section we give examples of classical signals together with their DFTs.

Figure 4.20 corresponds to the signal:

$$x[n] = e^{-10(n.T_s)^2} \cos(2\pi.30.nT_s)$$

which is a harmonic signal at 30 Hz damped with a Gaussian envelope. $T_s = 1/f_s$ and f_s=1 kHz. The number of points is N=1024=2^{10}. The signal is shown in Figs. 4.20(a)-(b) versus n and t_n, respectively. The DFT is shown in Figs. 4.20(c)-(d) versus k and f_k, respectively. The location of the spikes correspond to the cosine frequency, $\pm$30 Hz. A Gaussian is sticked on each of theses positions (because the Fourier transform of a Gaussian is a Gaussian).

The second example deals with the following cosine signal at 30 Hz:

$$x[n] = \cos(2\pi.30.nT_s)$$

with, as for the first example: f_s=1 kHz, N=1024=2^{10}. The signal is shown in Fig. 4.21(a). Its spectrum is shown in Fig. 4.21(b) as a function of f_k. And a zoom in on the spectrum is shown in Fig. 4.21(c). While the signal contains only a single frequency component at 30 Hz, the spectrum contains non-zero components for several frequencies around this value. The reason is clear: the N points used do not contain an integer multiple of the signal period, and as a result there is some spectral leakage. Equivalently, in Fig. 4.21(c) it is seen that no frequency f_k exactly corresponds to 30 Hz.

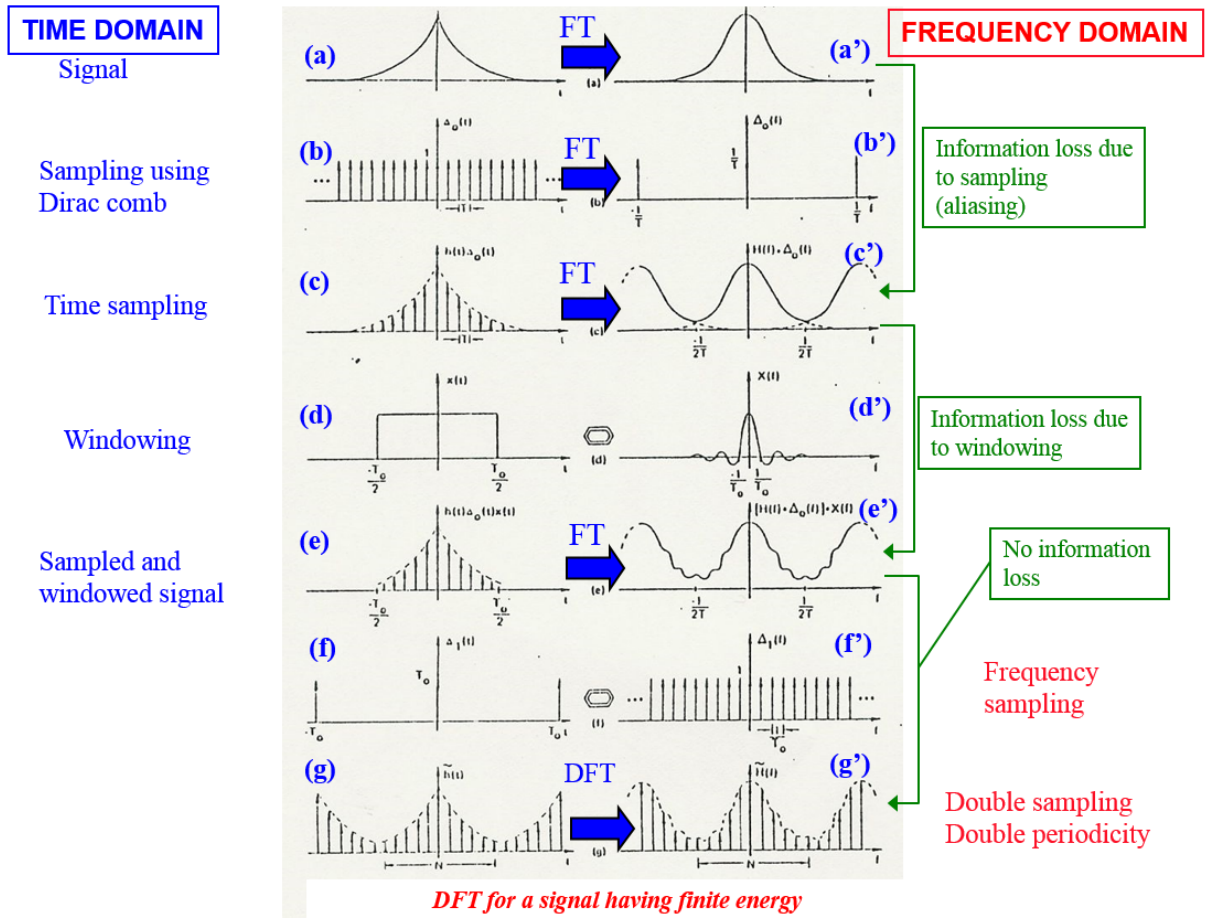


Figure 4.19: Different steps involved in the DFT process (From Brigham [9], Chap. 6).

If one wants a precise estimation of the frequency in that case, is is possible to use zero-padding. Zeros are appended to the signal so that the number of points in the resulting signal is a power of 2 (for using FFT). N_{pad}=7168 zeros are appended, for a total N=8192=2^{13}. The signal is shown in Fig. 4.21(a). Its spectrum is shown in Fig. 4.21(b) as a function of f_k. And a zoom in on the spectrum is shown in Fig. 4.21(c). The frequency resolution has increased, and the DFT is much more continuous than before (what is seen is actually the absolute value of a sinc function). It is now possible to estimate the frequency of the signal with much more precision.

Exercise

N=12. Consider 3 differents DFTs:

$$X_1 = [0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]$$
$$X_2 = [0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0]$$
$$X_3 = [1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]$$

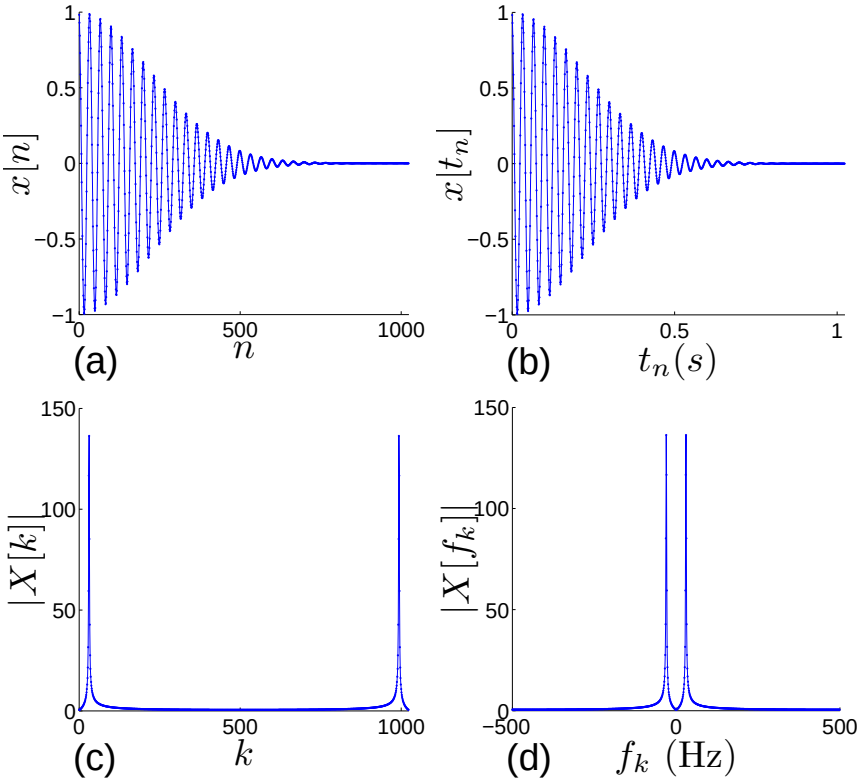


Figure 4.20: (a) Signal vs n; (b) Signal vs t_n; (c) DFT vs k ; (d) DFT vs f_k.

We consider the 3 signals calculated by IDFT:

$$x_1 = IDFT[X_1]$$
$$x_2 = IDFT[X_2]$$
$$x_3 = IDFT[X_3]$$

Which of these signals is real? Which of them is constant? Which of them is grid-to-grid? Which of them has a "moderate" frequency?
The signal x_1 is complex. Why? Suppose you want a signal at the same frequency, but real: which digits in X_1 would you change prior to calculating the IDFT?
Check your results with Matlab (in case the signals are complex, plot their real part for example).
Note: when a signal is sampled with a frequency f_s, a component with $\pm f_s/2$ is said to be grid to grid, it looks like a sawtooth wave.

4.10 Conclusion

To perform digital signal processing, one needs to handle digital (discrete) data. The first question in this chapter was to know if discrete samples contain all the information of a continuous signal. If the signal is bandlimited, and if the sampling rate is at least twice as large as the

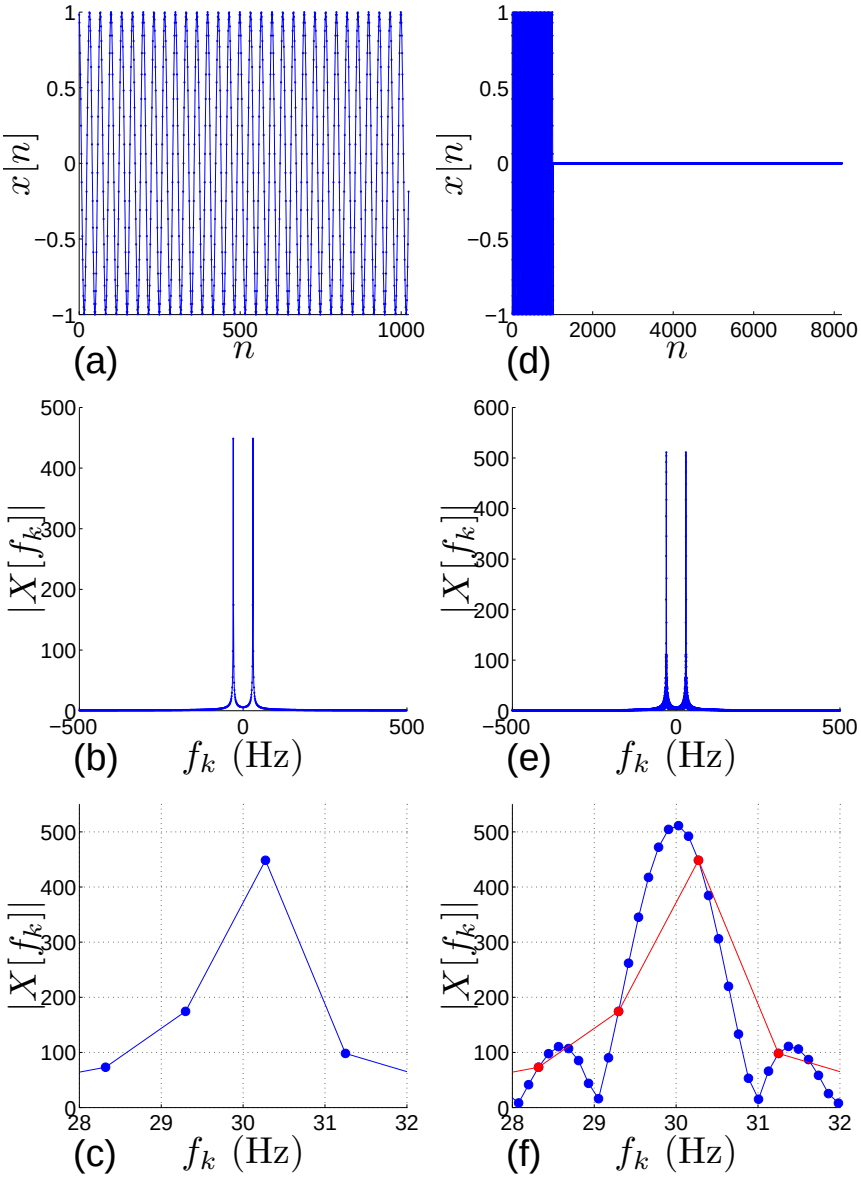


Figure 4.21: (a) Signal vs n, no padding (N=1024); (b) DFT, no padding (N=1024); (c) zoom in on the DFT, no padding (N=1024); (d) Same signal as in (a) after zero padding up to N=8192; (e) DFT of the zero-padded signal (N=8192); (f) zoom in on the DFT, for the padded signal (N=8192). For comparison, the red curve in (f) is a copy of the curve in (c).

maximal frequency contained in the signal, then the signal can be reconstructed from its samples (Shannon's theorem). To prove this theroem, we have used the fact that sampling in one domain (ie multiplying with a Dirac comb) is equivalent to periodizing in the dual domain (ie convoluting with a Dirac comb).

The signals used to prove the Shannon theorem are not exactly discrete, they are a contin-

uous model to a discrete signal. We have thus introduced discrete signals. Using a transform called the Discrete-Time Fourier Transform (DTFT), a spectrum can be calulated from these signals. While the signal is discrete, the spectrum is continuous and periodic.

Since a computer can store only discrete signals, for both time and frequency sequences, both the signal and its transform in the frequency domain need to be discrete and periodic. The discrete Fourier transform and its inverse work with N-points signals that are the periods of periodic signals. The DFT and IDFT are a discrete form of harmonic analysis. Signals are expanded as a sum of discrete harmonic waves (IDFT), and the complex amplitudes of these waves are calculated using the DFT.

Many properties of the FT have an equivalent for the DFT. However, the periodic nature of the sequences need to be accounted for. This is the reason why the circular convolution product has been introduced. With this circular convolution product, there exists a convolution theorem for the DFT. It is also because of periodicity that the first $N/2$ (or $N/2$-1, depending on parity) samples returned by the DFT correspond to positive frequencies, while the last $N/2$ samples correspond to negative frequencies. This has to be accounted for when interpreting the spectra calculated using DFTs.

Windowing has been adressed. When a finite-length sequence $x[n]$, n=0...N-1, is obtained from a longer sequence by picking up its values without any modification of their amplitudes, this corresponds to a rectangular window. This is likely to cause spectral leakage in the frequency domain, ie, one frequency component affects the frequency components whose frequencies are far away in the spectral space. For a periodic signal, this can be avoided if the finite-length sequence exactly corresponds to one period of the signal. Otherwise, spectral leakage (to far away frequencies) can be reduced by using a smooth window that tapers the signal in the time domain.

Subtleties of zero-padding have been introduced. When one needs a better precision in the frequency domain, it is always possible to achieve this by appending zeros to the signal prior to calculating the DFT.

Finally, the FFT is nothing but a fast way of calculating the DFT, in $\sim N\ln N$ operations, where N needs to be a power of 2. FFT is also used to calculate in a fast way correlations or convolutions, which are then called fast correlations and fast convolutions.

5 Introduction to random processes

Until now, only deterministic processes have been considered. We now consider the case of random processes. First, the definition of a random process is given. Statistical and time descriptions of the process are then studied in turn: they rely respectively on an ensemble average operator, and on a time average operator. Important quantities are the mean of the process, its variance, its autocorrelation. For two processes, an important quantity is the cross-correlation. The definition of stationarity and ergodicity are given. Ergodicity means that the ensemble average and time average give similar results. It is often supposed in practice, which allows performing a single experiment instead of a lot of experiments. Then, spectral characteristics of random processes are given. Auto-spectral and cross-spectral power densities are introduced. These can be connected to the quantities introduced earlier, since the Wiener-Khintchine theorem tells us that spectral densities are the Fourier transform of correlation functions. The practical way of computing a power spectral density is then given: the method relies on the averaged periodogram, also called Welch's method. Finally, the transformation (filtering) of a random process by a linear time invariant filter is considered. The coherence function allows determining whether two random processes are linked by a linear time-invariant relationship. Finally, a practical example is given.

A knowledge of random variables is necessary for studying random processes. This is given in Appendix C. Random variables and processes are adressed in Papoulis [26], or in Bendat and Piersol [3]. In connection with turbulence, they are also presented in Chap. 6 of Tennekes and Lumley [37], or in Chap. 3 of Pope [27], amongst others.

5.1 Definition of a random process

Many practical signals are random: electronical noise, background noise in a factory, speech signals, stock exchange fluctuations, industrial vibrations, seismic signals, turbulent flows,... In most cases experimental signals are acually composite: they contain a useful part (that is itself deterministic or random), mixed with noise (random most of the time). An objective of signal processing is often to extract the useful part of the information.

A practical consequence of the signal being random is that performing several times an experiment does not provide the same result. The typical measurement/experiment we will consider in this chapter consists of:
- setting up a turbulent flow in a channel or a jet by switching on, say, a fan;

- measuring in time the velocity (using a hot wire for example) at a given point in space (or at several given points in space), over a certain duration.
Making a new measurement implies: switching off the fan, then restart the procedure. In appendix C, we consider a similar but simpler situation where the measurement is made at a unique given time after switch on: this corresponds to a random variable. By contrast, a measurement in this chapter provides a whole time series (also called a sample function), and this corresponds to a random process. Two consecutive measurements (time series, sample functions) are not the same (it is not possible to super-impose the first time trace with the second one). However, what these two time traces have in common are their statistical properties (the moments of the signal): mean value, rms value, skewness,... Hence, the useful tools for dealing with random signals are statistical tools.

To be more specific, a random process has the following mathematical modelling: $x(t, \xi)$. It is shown in Fig. 5.1. The important thing to note is that it depends on two variables:

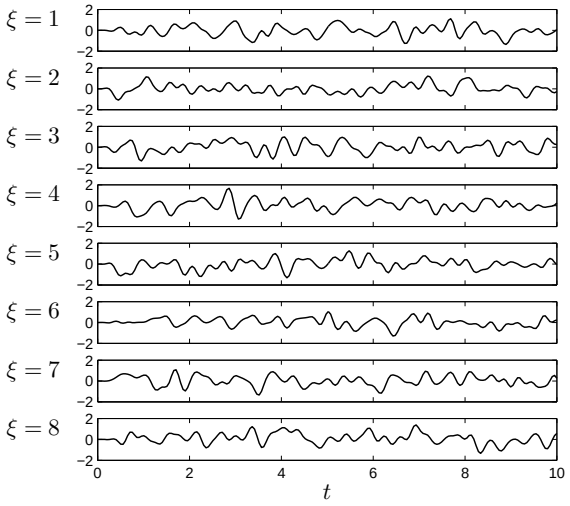


Figure 5.1: Representation of a random process. Eight sample functions ($\xi = 1...8$) of the process are shown, and each sample function is shown only on the time interval (0,10).

- the time t, a real continuous variable;
- the experiment number, ξ, a natural integer (in general, ξ has not to be countable though).

The latter accounts for the random character of the signal. For a given experiment ξ, one obtains a full times series. This is, as we said, the difference between a random process $x(t, \xi)$ and a random variable $x(\xi)$ whose output is a single real number.

There are two ways for understanding a random signal:
1) The temporal description (fixed experiment, horizontal study) is shown in Fig. 5.2(a): the experiment is fixed, $\xi = \xi_0$, which corresponds to a single measurement. The random process reduces to a time function: $x_0(t) = x(t, \xi_0)$ for $t \in (0, \infty)$. This function is a called a **sample function** of the random process.
2) The statistical description (fixed time, vertical study) is shown in Fig. 5.2(b): the time is fixed: $t = t_0$. We look at the result of several measurements (ξ=1,2,3,...∞) at this particular

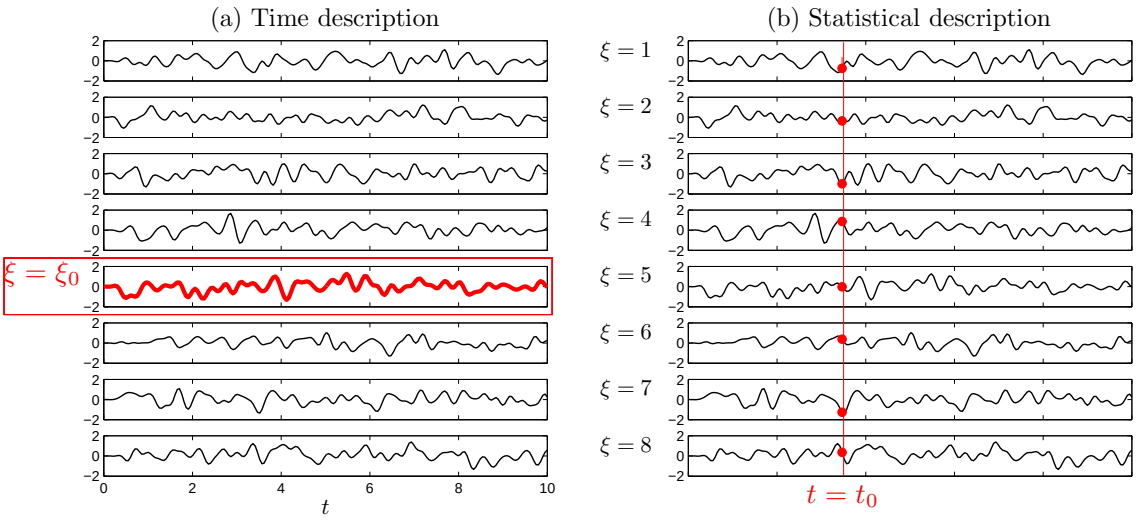


Figure 5.2: Two possible descriptions of a random process.

time. The random process reduces to a random variable, $x_0(\xi) = x(t_0, \xi)$.

When facing a random signal, an experimentalist can ask himself several questions:
1. How to characterize a signal using statistical tools?
2. How to characterize this signal given that it is not possible to realize infinitely many experiments (with t=0...∞ and ξ=1,2,...,∞)? Would it be possible to use a finite number of experiments? In fact, would it be possible to use only a single measurement and perform a single horizontal study?

In the following the random processes take real values (no complex conjugate needs to be used in the correlations).

5.2 Statistical description

5.2.1 Description at one time

In the present section, we got interested in the vertical study in Fig. 5.2(b). We suppose that we are able **to perform an experiments many times**, and we look at the different samples ξ of this experiment at a fixed time $t = t_0$. Then, the random process reduces to a random variable $X(\xi) = x(t_0, \xi)$ and this depends only on the sample number ξ.

Let us recall that a random variable can be studied using the **statistical tools** presented in Appendix C. The random variable is characterized by its **probability density function (PDF)** $p_X(x)$ characterizing the probability $p_X(x)\mathrm{d}x$ that the random variable takes a value between x and $x + dx$. From this, the **moments** of the random variable may be determined. The nth moment is given by:

$$M_n = \int_{-\infty}^{\infty} x^n p_X(x) dx \tag{5.1}$$

Conversely, one needs to know all the moments of a random variable to fully know this random variable. Very often, one uses only the first few moments. These are:

The expected value of $X(\xi)$:
It is also called the ensemble average and is given by:

$$m_X(t) = E\left[x(t,\xi)\right] = \int_{-\infty}^{\infty} x p_X(x,t)\mathrm{d}x \tag{5.2}$$

For the experiment we are considering as an example, X is a random variable that returns a velocity, and the integration variable x represents a velocity as well. m_X is the ensemble averaged velocity of the flow. The PDF $p_X(x,t)$ is the probability that, at time t, the velocity is between x and $x+\mathrm{d}x$. In general, it depends on time. Let us precise why it should be so by considering again our flow measurement where we swicth on the fan at time t=0: during the initial stage at least, the flow will undergo transition to turbulence, and the ensemble averages of the flow before transition and after transition are likely to be different. A practical formula to calculate the expected value is:

$$m_X(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{\xi=1}^{\infty} x(t,\xi)$$

The variance of $X(\xi)$:
The variance is given by:

$$\sigma_X^2(t) = E\left[(x(t,\xi) - m_X)^2\right] = \int_{-\infty}^{\infty} (x - m_X(t))^2 p_X(x,t)\mathrm{d}x \tag{5.3}$$

where the standard deviation, $\sigma_X(t)$, is also a velocity in m/s and represents the typical velocity fluctuation magnitude with respect to the ensemble average. Like the ensemble average, it depends on time. A practical formula to calculate the variance is:

$$\sigma_X^2(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{\xi=1}^{\infty} (x(t,\xi) - m_X(t))^2$$

For a finer description of the random variable at a given time, more moments would be needed (skewness,...). In addition, the random process depends on time, and a characterization should account for the connection between what happens at several different times. Indeed, for any number N of times, the quantities $x(t_1,\xi)$, $x(t_2,\xi)$,..., $x(t_N,\xi)$ represent N random variables whose joint knowledge would require a N-dimensional probability density function, and this is true for any value of N. This task is impossible and in practice only two times are considered, as in the next subsection.

5.2.2 Description at two times

Two times are considered, meaning we are now dealing with two random variables: either $x(t,\xi)$ and $x(t+\tau,\xi)$, or $x(t,\xi)$ and $y(t+\tau,\xi)$.

Autocorrelation of $X(\xi)$:
There is now one signal at two different times, see Fig. 5.3. The **autocorrelation** for the process

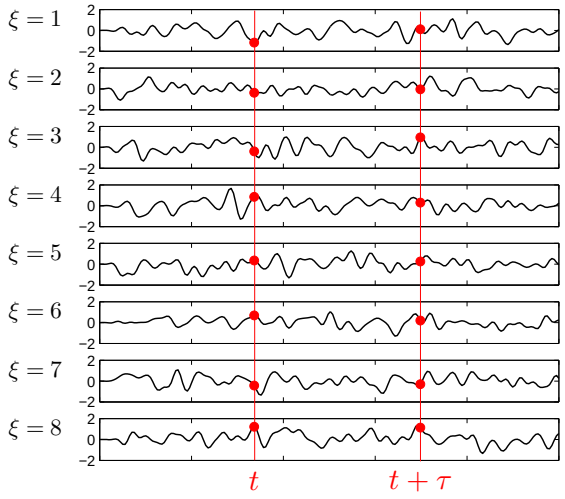


Figure 5.3: One random process is considered at two different times t and $t+\tau$.

x is:

$$R_{xx}(t,\tau) = \mathrm{E}\left[x(t,\xi)x(t+\tau,\xi)\right] \tag{5.4}$$

Cross-correlation of $X(\xi)$ and $Y(\xi)$:
There are now two signals at two different times. This typically occurs when a measurement involving several sensors is performed. The two random processes x and y are measured *simultaneously*. The process x is considered at time t and the process y is considered at time $t+\tau$, as shown in Fig. 5.4. The **cross-correlation** between the process x and y is:

$$R_{xy}(t,\tau) = \mathrm{E}\left[x(t,\xi)y(t+\tau,\xi)\right] \tag{5.5}$$

5.2.3 Stationary random processes

In general, the statistical quantities $m_x(t)$, $\sigma_x(t)$, $R_{xx}(t)$, and $R_{xy}(t)$ depend on time. A **stationary process** is a process for which they do not depend on time. That is, for a stationnary process:

$$m_x(t) = m_x \qquad \text{(independent on time, stationary process)} \tag{5.6}$$
$$\sigma_x^2(t) = \sigma_x^2 \qquad \text{(independent on time, stationary process)} \tag{5.7}$$
$$R_{xx}(t,\tau) = R_{xx}(\tau) \qquad \text{(independent on time, stationary process)} \tag{5.8}$$
$$R_{xy}(t,\tau) = R_{xy}(\tau) \qquad \text{(independent on time, stationary process)} \tag{5.9}$$

The statistical properties of the signal do not depend on the observation time.
The following relations hold:

$$R_{xx}(-\tau) = R_{xx}(\tau) \tag{5.10}$$
$$|R_{xx}(\tau)| \leq R_{xx}(0) \tag{5.11}$$
$$R_{xy}(-\tau) = R_{yx}(\tau) \tag{5.12}$$
$$|R_{xy}(\tau)|^2 \leq R_{xx}(0)R_{yy}(0) \tag{5.13}$$

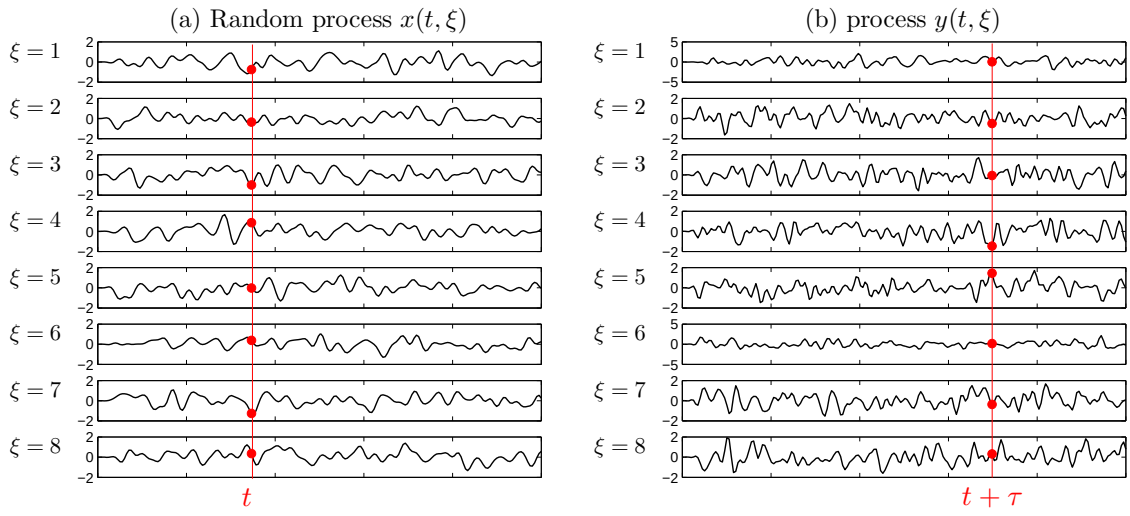


Figure 5.4: Two different random processes. The first one is considered at time t and the second one is considered at time $t+\tau$.

As a result, these quantities are plotted for $\tau > 0$ only. $R_{xx}(0)$ corresponds to the mean (the expected) square value of the signal, which is its power.

Two processes satisfying $R_{xy}(\tau)=0\ \forall\tau$ are said to be **uncorrelated**.

In the following, only stationary random processes are considered.

5.3 Time description

We now fix an experiment ξ_0 and look at the time evolution of the single sample function $x(t,\xi_0)$, as in Fig. 5.2(a). In this section we simply write $x(t)$ for $x(t,\xi_0)$. This sample function is supposed stationary, meaning that quantities such as the average and correlation do not depend on time. The tools defined for deterministic signals in Chapter 3 can be used.
The time average of $x(t)$ is:

$$\bar{x} = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\mathrm{d}t \tag{5.14}$$

Its variance is:

$$s_x^2 = \lim_{T\to\infty} \frac{1}{T} \int_0^T (x(t)-\bar{x})^2\mathrm{d}t \tag{5.15}$$

The autocorrelation is:

$$C_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)\mathrm{d}t \tag{5.16}$$

and $C_{xx}(\tau=0)$ is the power of the sample function.
If in addition to x, we have y, the cross-correlation is:

$$C_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)y(t+\tau)\mathrm{d}t \tag{5.17}$$

These quantities are calculated in the limit $T \to \infty$. In practice, calculations are performed over a finite time interval and what we obtain is an estimate of these quantities.

5.4 Ergodicity

Ergodic processes form a subclass of stationary processes, as shown in Fig. 5.5. A stationary

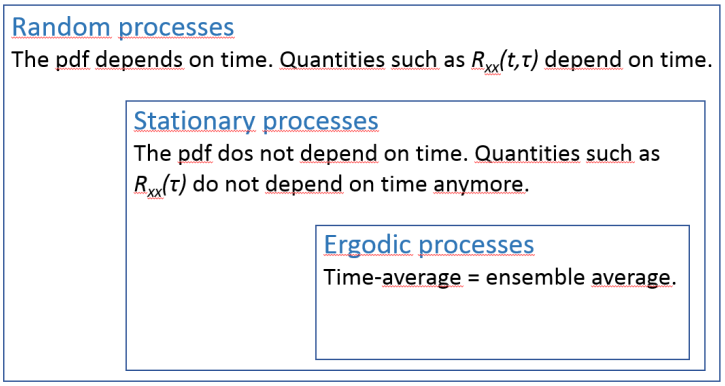


Figure 5.5: Hierarchy of processes. Stationary processes form a subclass of random processes, and ergodic processes a subclass of stationary processes.

random process is ergodic when the statistical properties calculated using ensemble averages (expected values) are the same as that calculated using time averages. In particular for the low order moments:

$$m_x = \bar{x} \tag{5.18}$$
$$\sigma_x^2 = s_x^2 \tag{5.19}$$
$$R_{xx}(\tau) = C_{xx}(\tau) \tag{5.20}$$
$$R_{xy}(\tau) = C_{xy}(\tau) \tag{5.21}$$

Ergodicity, which includes stationarity, thus tells us that the horizontal and vertical studies in Fig. 5.2 are equivalent.
Ergodicity is often supposed in practice: it allows calculating the mean and the variance of a random process based on a single sample function of this process (a single experiment) by using time averages. Experimentally, it is easier to observe a signal over a long time than to repeat an experiment a large number of times. Ideally, the experiment should last infinitely long, with $T \to \infty$. This is not possible and the statistical properties that are measured are only estimates.

For a random process to be ergodic, a sufficient condition is that the processes be stationary, and that the time-averages (mean, correlation,...) calculated using sample functions ξ be independent of ξ. It is indeed obvious that if one wants to calculate the statistical quantities using only one experiment, then any experiment should give the same result. This condition is used in the following exercise.

Exercise : a basic random process

Consider the following random process:

$$x(t,\xi) = A(\xi)\sin(2\pi f_0 t + \phi(\xi))$$

where A and ϕ are two random variables.
Plot what could be the first 3 sample functions of this process.
Calculate the auto-correlation for some particular event ξ_0.
Can this process be ergodic? What if the amplitude A does not depend on ξ?

5.5 Description of the random process in the frequency domain

5.5.1 Power spectral density (PSD)

In the following $x(t,\xi)$ and $y(t,\xi)$ are two stationary random processes. The Fourier transform of the process x is given by:

$$X_T(f,\xi) = \int_0^T x(t,\xi)e^{-\mathrm{j}2\pi ft}\mathrm{d}t \tag{5.22}$$

This equation says that for a given experiment ξ the sample function $x(t,\xi)$ is windowed by a rectangular window of length T and the Fourier transform of the result is calculated. This Fourier transform depends of the experiment ξ and so it is a random variable itself. For the experiment ξ, $X_T(f,\xi)$ gives the complex amplitude of the wave with frequency f contained in $x(t,\xi)$. Similarly, the Fourier transform of the process y is given by:

$$Y_T(f,\xi) = \int_0^T y(t,\xi)e^{-\mathrm{j}2\pi ft}\mathrm{d}t \tag{5.23}$$

Using these 2 windowed Fourier transforms, we define the cross-spectral power density of the windowed signals by:

$$S_{xy,T}(f,\xi) = \frac{1}{T}X_T^*(f,\xi)Y_T(f,\xi) \tag{5.24}$$

where as usual $*$ denotes the complex conjugate. This depends on the length T of the window. This depends also on ξ, so for now $S_{xy,T}(f,\xi)$ is itself a random variables that is a function of the random variables X_T and Y_T. To obtain a power spectral density of the process, we are going to average the spectra obtained for all the experiments. Mathematically, we take the expected value. The **power cross-spectral density** (cross-PSD) is finally:

$$S_{xy}(f) = \lim_{T\to\infty} E\left[S_{xy,T}(f,\xi)\right] \qquad \text{(complex)} \tag{5.25}$$

Taking the expected value has removed the dependence on ξ. The limit on T consisting in taking an infinitely wide window has removed the dependence on T. We are left with a power spectral density that does just depend on f. The meaning of this quantity is the following: its modulus is large when both x and y contain a large amplitude component at frequency f. Its phase is

the phase between the component of y at frequency f and that of x (that is, the phase of y minus the phase of x at frequency f, the minus arising from the complex conjugation in Eq. (5.24)).

By setting y=x, it is possible to calculate the **power auto-spectral density** (PSD) for the signal x:

$$S_{xx}(f) = \lim_{T\to\infty} E\left[S_{xx,T}(f,\xi)\right] \qquad \text{(even, real, positive)} \tag{5.26}$$

It is real (it is not complex since there is no phase lag between x and itself). It is large when x has a large amplitude component at frequency f.

Link with the deterministic case:

A deterministic experiment is one that gives always the same result, independently of ξ. Hence, for a deterministic process: $x(t,\xi) = x(t)\ \forall\xi$. The same is true for the Fourier transform and for the spectral densities, that is: $X_T(f,\xi)$=$X_T(f)\ \forall\xi$, and $S_{xx,T}(f,\xi) = S_{xx,T}(f)\ \forall\xi$. As a result $E[S_{xx,T}(f,\xi)]$=$S_{xx,T}(f)$. The power spectral is then:

$$S_{xx}(f) = \lim_{T\to\infty} E\left[S_{xx,T}(f,\xi)\right] = \lim_{T\to\infty} S_{xx,T}(f) = \lim_{T\to\infty} \frac{1}{T} X_T^*(f) X_T(f)$$

This is exactly the same result as given in Chapter 3 for deterministic signals with finite power (see Eq. (3.44)). All the definitions given for random processes remain valid for deterministic processes.

5.5.2 Wiener-Khintchine theorem

The Wiener-Khintchine theorem states that the auto-spectral and cross-spectral power densities defined above are the Fourier transforms of the auto-correlation and of the cross-correlation. That is:

$$S_{xx}(f) = \mathrm{FT}\left[R_{xx}(\tau)\right] \qquad \text{(Wiener-Khintchine)} \tag{5.27}$$

$$S_{xy}(f) = \mathrm{FT}\left[R_{xy}(\tau)\right] \qquad \text{(Wiener-Khintchine)} \tag{5.28}$$

For a proof, see Chap. 5 of Bendat and Piersol's [3]. Here, the auto and cross correlations are obtained by ensemble average. For ergodic processes they can be calculated by time average. The "Wiener-Khintchine" theorem has already been introduced in Chap. 3 for deterministic signals. It is the present extension for random signals that generally bears the Wiener-Khintchine name.

5.5.3 Power of random processes

The mean power P_x of a stationary random process can be defined by:

$$P_x = R_{xx}(0)$$

(for an ergodic process, this is also $P_x = \lim_{T\to\infty} 1/T \int_0^T x(t)^2 \mathrm{d}t$).

The Wiener-Khincthine theorem gives:

$$R_{xx}(\tau) = \mathrm{FT}^{-1}\left[S_{xx}\right] = \int_{-\infty}^{\infty} S_{xx}(f) e^{\mathrm{j}2\pi f\tau} \mathrm{d}f$$

Thus, at τ=0:

$$P_x = R_{xx}(0) = \int_{-\infty}^{\infty} S_{xx}(f)\mathrm{d}f$$

The power is the integral over frequency of the power spectral density, as for deterministic signals.

5.5.4 Examples of autocorrelation

Two examples of random signals and their correlations and power spectral densities are given in Fig. 5.6. A sample function for a white noise, $x(t)$, is shown in Fig. 5.6(a). White noise refers

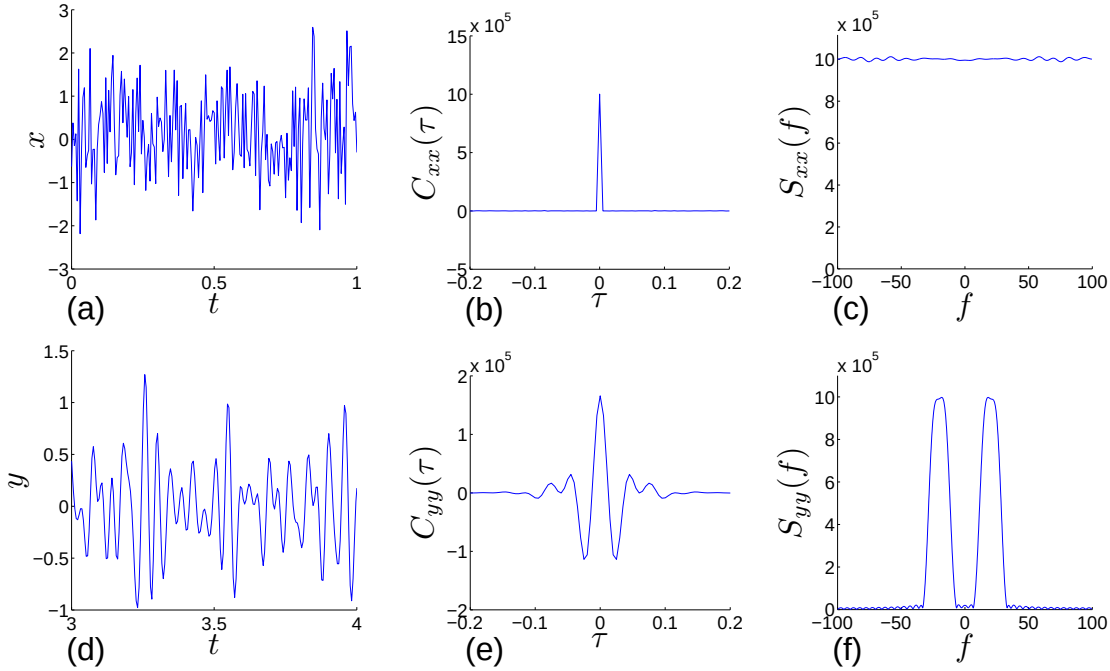


Figure 5.6: (a) White noise; (b) Autocorrelation of white noise; (c) PSD of the white noise obtained by taking the FT of the autocorrelation; (d) Band limited noise; (e) Autocorrelation of band limited noise; (f) PSD of the band limited noise obtained by taking the FT of the autocorrelation.

to the spectrum being flat (no preferred frequency). In terms of amplitudes, the noise is taken to be Gaussian, that is, its PDF is Gaussian. The autocorrelation C_{xx} is shown in Fig. 5.6(b). It is a Dirac. The autocorrelation is defined by $E[x(t)x(t+\tau)]$. It is large for a given τ if $x(t)$ and $x(t+\tau)$ resemble each other *on average*, that is, statistically. For a white noise, the signal is erratic and changes so rapidly that $x(t)$ and $x(t+\tau)$ resemble each other only for τ=0. Finally, the PSD is given in Fig. 5.6(c). It can be calculted either directly from $x(t)$ (this direct estimation using the periodogram method is the subject of a subsequent section), or by Fourier transform of $C_{xx}(\tau)$ by application of the Wiener-Khintchine theorem. Surpriseless, the spectrum is constant since it is the Fourier transform of a Dirac. And this is the characteristic of a white noise. The second signal is a band limited noise for which a sample function, $y(t)$, is shown in Fig. 5.6(d). In practice, this noise is created by bandpass filtering a white noise, and

here $y(t)$ is obtained by filtering $x(t)$. The signal so created has less frequency components: as a result it appears smoother and more harmonic. Its autocorrelation in Fig. 5.6(e) has a longer coherence time (compared with the Dirac corresponding to the white noise) and is actually close to a sinc function. The PSD in Fig. 5.6(e) is the Fourier transform of the autocorrelation. It is non-zero and almost constant over a positive frequency band (the symmetric part for negative frequencies is due to the band pass filter having a real impulse response) and it approximates a rectangular window in the frequency domain. Of course this frequency band corresponds to the pass band filter that has been applied to x to produce y. The inverse Fourier transform of a rectangular window is a sinc function, which explains why the autocorrelation in Fig. 5.6(e) looks like a sinc function (more precisely, the spectrum is made of a translated rectangular function, and the autocorrelation is a sinc function multiplied by a cosine function).

5.5.5 Periodogram

We have introduced the PSD of a random process. This is an important quantity that one will always want to obtain when performing harmonic analysis. It can be calculated:
- either by Fourier transform of the signal, as in Eq. (5.26), this is the periodogram method;
- or by Fourier transform of the correlation (Wiener-Khintchine theorem), this is the correlogram method.

In the following, a "recipe" is provided for estimating the PSD from measurements using the periodogram method, supposing the process is **ergodic**. The PSD is going to be calculated from (see Eq. (5.26)):

$$S_{xx}(f) = \lim_{T\to\infty} E\left[\frac{1}{T} X_T^*(f) X_T(f)\right] \tag{5.29}$$

Similar formula hold for cross-spectra. The process being ergodic, a single measurement of the process is carried out, over a duration T. In practice, T has to be finite and the limit in the above definition is omitted. The question is the following: how to calculate the PSD from this single measurement ?

Raw periodogram (the naive way)

The raw periodogram is a way to estimate the PSD by *omitting the expectation* in Eq. (5.29) (the limit on T is omitted as well). The PSD is estimated by:

$$\tilde{S}_{xx}(f) = \frac{1}{T} X_T^*(f) X_T(f) \tag{5.30}$$

The PSD is then calculated as for a deterministic signal (see Eq. (3.44)). Of course this cannot work well for a random signal. Let's denote by $\sigma\left[\tilde{S}_{xx}(f)\right]$ the standard deviation (the typical random error) of the estimate. This may be understood as the typical error done in the estimation at a given frequency. We have:

$$\frac{\sigma\left[\tilde{S}_{xx}(f)\right]}{S_{xx}(f)} = O(1) \tag{5.31}$$

In words, the typical random error of the estimate has the same magnitude as the quantity that needs to be estimated. Moreover, this is independent of T: increasing T does not decrease the

standard deviation.

Let us see how the PSD is estimated practically using a computer. The measured signal is written $x[n]$, n=0...N-1. Its N-point DFT is:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-\mathrm{j}2\pi nk/N} \qquad \forall k = 0...N-1$$

The DSP is then calculated by:

$$S_{xx}[k] = \frac{1}{N}|X[k]|^2 \qquad \forall k = 0...N-1$$

where the index k corresponds to a frequency f_k, as explained in the chapter on DFT.

A DSP calculated using the raw periodogram is shown in Fig. 5.7(a) and is compared to the known DSP of the random process. The DSP is very noisy, and it is observed that the typical error has the same order of magnitude as the estimated spectrum, as indicated in Eq. (5.31). This would not be improved by increasing N (that is, by increasing the duration T of the measurement). By contrast, the raw periodogram would work in a satisfying manner for a deterministic signal.

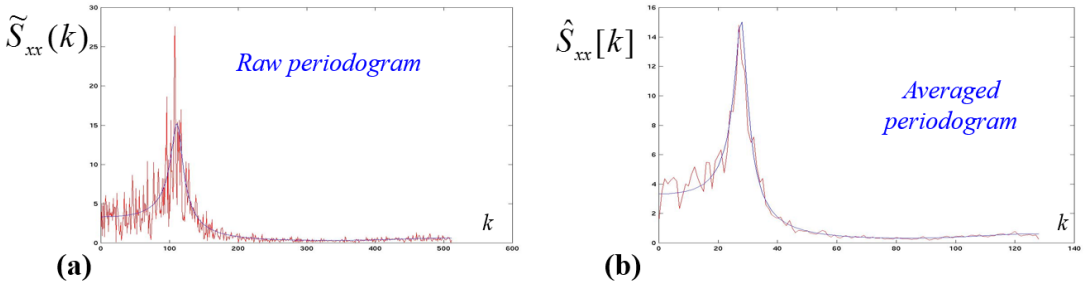


Figure 5.7: (a) DSP estimated using the raw periodogram; (b) DSP estimated using the averaged periodogram.

Averaged periodogram (Welch's method)

The averaged periodogram method, also called **Welch's method**, improves the quality of the PSD estimate. To decrease the typical error of the estimate, the expected value in Eq. (5.29) should be kept somehow. The expected value relies on the fact that several sample functions are available, but we just have one. The solution is to split this unique sample function into M sample functions, or M blocks, of smaller duration. These blocks may possibly overlap each others, and using a 50% overlap is common. A window may possibly be used for each block to smooth out the discontinuity introduced by the splitting. The principle is shown in Fig. 5.8. In this figure, the total duration of the signal is T, the duration of one block is T_b, and there are M blocks that do not overlap, so that T=MT_b. The steps involved for computing the PSD are:
-Step 1: the PSD for each block i=1...M is calculated by:

$$\tilde{S}_{xx}^{(i)}(f) = \frac{1}{T_b} X_T^{(i)*}(f) X_T^{(i)}(f)$$

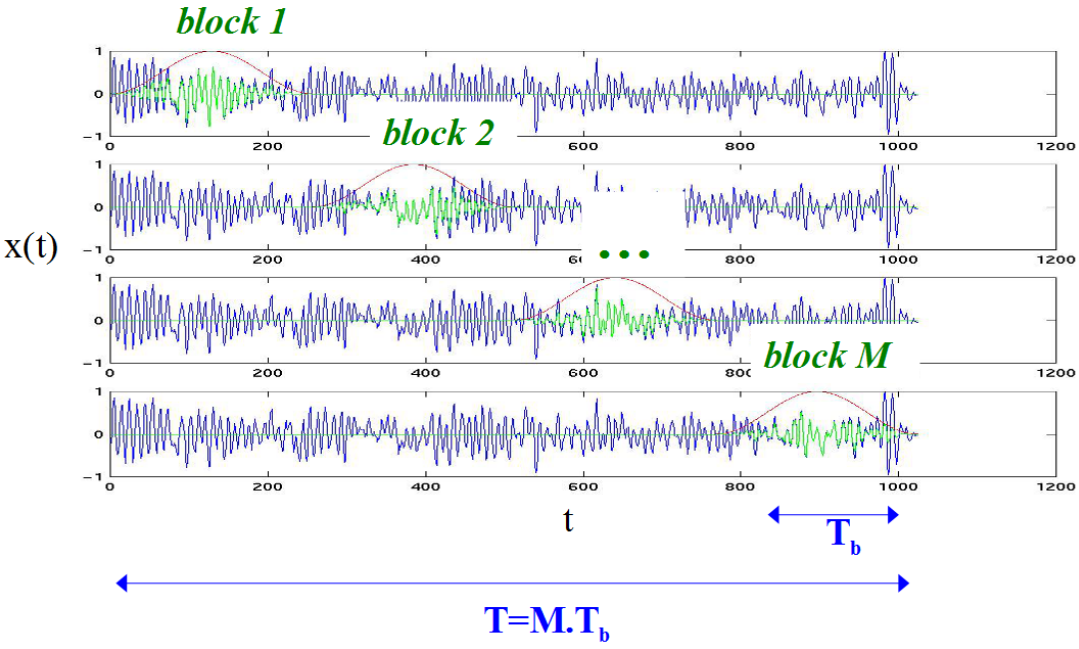


Figure 5.8: Splitting of the signal into blocks for the averaged periodogram (Welch's method). There is only one single sample function (the process is supposed to be ergodic), but four copies are plotted to indicate the positions of the blocks with more clarity.

-Step 2 : The PSD of the process is obtained by averaging the PSD of the blocks.

$$\hat{S}_{xx}(f) = \frac{1}{M}\sum_{i=1}^{M} \tilde{S}_{xx}^{(i)}(f)$$

This is this averaging operation that replaces the original expected value. Finally, the standard deviation of the estimate now satisfies:

$$\frac{\sigma\left[\tilde{S}_{xx}(f)\right]}{S_{xx}(f)} = O\left(\frac{1}{\sqrt{M}}\right) \tag{5.32}$$

The more blocks the less this error. The counterpart is that the frequency resolution is decreased since its value is $1/T_b$ instead of $1/T$ initially.

The numerical implementation of the DSP is simply given by:

$$\hat{S}_{xx}[k] = \frac{1}{M}\sum_{i=0}^{M-1} \frac{1}{N}|X_i[k]|^2 \qquad k = 0...N-1 \tag{5.33}$$

where $X_i[k]$ is the N-point DFT of the signal $x_i[n]$ in the i-th block, N is now the number of samples in one block (not the total number of samples), and M is the number of blocks.
For a cross-DSP between x and y, we would have:

$$\hat{S}_{xy}[k] = \frac{1}{M}\sum_{i=0}^{M-1} \frac{1}{N}X_i^*[k]Y_i[k] \qquad k = 0...N-1 \tag{5.34}$$

The DSP obtained by the averaged periodogram method is shown in Fig. 5.7(b). It is seen that the estimated DSP lies closer to the known one compared with that obtained previously with the raw periodogram.

5.6 Filtering of random processes, coherence function

5.6.1 Input/output relationship

When going through a time-invariant linear (LTI) system, a random process x is transformed into another random process y. The LTI may be a model for a physical system, a model for signal propagation, the response of a sensor,... It has an impulse response h that is deterministic. The system is shown in Fig. 5.9.

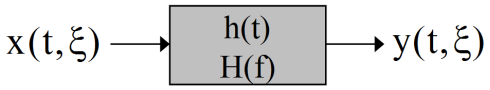


Figure 5.9: Schematic of a LTI system with random intput and output.

It is already known that the input and the output are related by:

$$y(t,\xi) = h(t) * x(t,\xi)$$

In addition, since a random process is characterized by its statistics, one would like to know how statistical quantities are modified by the filter. We have the following input/output relationships:

$$m_y = H(0)m_x \tag{5.35}$$
$$R_{xy}(\tau) = h(\tau) * R_{xx}(\tau) \tag{5.36}$$
$$R_{yy}(\tau) = R_{hh}(\tau) * R_{xx}(\tau) \tag{5.37}$$
$$S_{xy}(f) = H(f)S_{xx}(f) \tag{5.38}$$
$$S_{yx}(f) = \frac{S_{yy}(f)}{H(f)} \tag{5.39}$$
$$S_{yy}(f) = |H(f)|^2 S_{xx}(f) \tag{5.40}$$

These relations hold for a random processes, including deterministic signals. Moreover, the relations in the frequency domain are obtained by Fourier transforming the relations in the time domain and using the Wiener-Khintchine theorem.
The last relation links the DSP of the output to the DSP of the input, the link being the modulus squared of the frequency response of the LTI, and it is obvious that the power at the output is the power at the input multiplied by the gain in power of the LTI, $|H(f)|^2$. In this relation, it is important to note that no complex quantity is involved: there is no phase factor in this relation, because the power (frequency per frequency) is not a matter of phase. On the contrary, the fourth relation, $S_{xy}(f) = H(f)S_{xx}(f)$, involves complex quantities. Let's see what this relation means. It is recalled that the cross-DSP $S_{xy}(f)$ counts the power that x and y have in common: it is large for some frequency f when both x and y have a large amplitude component at this frequency. The relation shows that this is the case when the input contains the component ($S_{xx}(f)$ is large) AND the LTI system lets this component go through ($H(f)$

large). In addition, the phase of $S_{xy}(f)$ is the phase between the components in x and y at frequency f. The knowledge of a phase difference between two signals requires a cross-PSD. The PSD $S_{xx}(f)$ is real, and the phase in S_{xy} is therefore that caused by the LTI, and this is the phase of the complex frequency response, $H(f)$.

5.6.2 Coherence

The coherence function between signals x and y is defined by:

$$\mathrm{Coh}_{xy}(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}} \tag{5.41}$$

Here, the complex version of the coherence function is used[1]. We could also call it: the normalized cross-spectrum (cross-PSD). Coherence is used to test a linear stationary relationship between two random processes. Said otherwise: it is used to test whether the two signals are linked by a LTI system. Coherence is a complex number. Its phase is that of the cross-spectrum. Its modulus verifies: $0 \leq |\mathrm{Coh}_{xy}(f)| \leq 1$. Three cases may be distinguished:
• $|\mathrm{Coh}_{xy}(f)|$=1: x and y are completely coherent at frequency f: there exists a linear and stationary relationship between them.
• $|\mathrm{Coh}_{xy}(f)|$=0: x et y are completely incoherent at frequency f
• $0 < |\mathrm{Coh}_{xy}(f)| < 1$: this is the case "in between" where there are 3 possibilities:
  1. Noise is present.
  2. The relation between $x(t)$ and $y(t)$ is not linear.
  3. The signal $y(t)$ depends on $x(t)$ and on some other signals as well.

Coherence has a meaning only for random processes, and requires an averaging procedure that is present in the DSPs through the expectation operation. Often, a coherence larger than 0.8 or 0.9 is judged significant enough for a linear and stationary relationship between the two signals to exist. However, the causality is not established by the coherence: one does not know whether x causes y, or whether y causes x.

For signals related by a LTI, such as the processes x and y in Fig. 5.9, the modulus of the coherence satisfies:

$$|\mathrm{Coh}_{xy}(f)| = 1$$

A coherence with modulus equal to 1 corresponds to a linear stationary relationship, which is to say that the two signals are respectively the input and the output of a LTI system.

In general, some noise is present (think of a sensor that always picks up some noise). This situation is shown in the model in Fig. 5.10 where some noise $b(t,\xi)$ is supposed to be added at the output of the LTI. In that case, the output one would measure is $s(t,\xi)$, the sum of the useful signal $y(t,\xi)$ and of the noise $b(t,\xi)$. The coherence between the input and the output is given by:

$$\mathrm{Coh}_{xs}(f) = \frac{1}{\sqrt{1 + \frac{S_{bb}(f)}{S_{yy}(f)}}}$$

[1] $\mathrm{Coh}_{xy}(f)$ corresponds to the *complex coherence function* $\gamma_{xy}(f)$ defined in Bendat and Piersol [3], see their Eq. (5.86). They also introduce in their Eq. (5.84) the square of the *modulus* of this quantity, which they denote by $\gamma^2_{xy}(f) = |\gamma_{xy}(f)|^2$ and call it the *coherence function*, a real quantity. Different authors will call coherence either $\mathrm{Coh}_{xy}(f) = \gamma_{xy}(f)$ or $\gamma^2_{xy}(f)$.

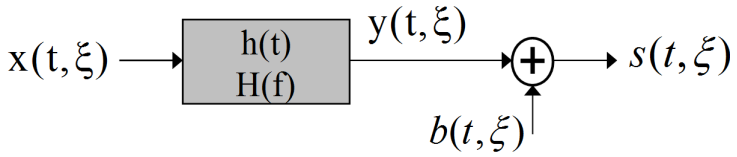


Figure 5.10: Schematic of a LTI system when noise is added to the output.

Overall, the addition of noise decreases the coherence. The larger $S_{bb}(f)$ (this is the power of the noise at frequency f) is compared to the power of the useful signal ($S_{yy}(f)$) the smaller the coherence (this is point 1. in the list above).

Finally, numerical formulas are given to calculate the coherence using the DFT. The coherence function is given by Eq. (5.41) that involves PSDs and cross-PSDs. We know how to calculate these numerically using the averaged periodogram. Hence, calculating the coherence is straightforward. The signals are split into M blocks of N samples each. The coherence is calculated as:

$$\mathrm{Coh}[k] = \frac{\frac{1}{M}\sum_{i=1}^{M}\frac{1}{N}X_i^*[k]Y_i[k]}{\sqrt{\frac{1}{M}\sum_{i=1}^{M}\frac{1}{N}|X_i[k]|^2\frac{1}{M}\sum_{i=1}^{M}\frac{1}{N}|Y_i[k]|^2}} = \frac{\sum_{i=1}^{M}X_i^*[k]Y_i[k]}{\sqrt{\sum_{i=1}^{M}|X_i[k]|^2\sum_{i=1}^{M}|Y_i[k]|^2}} \tag{5.42}$$

where $X_i[k]$ is the N-point DFT of the signal $x_i[n]$ in block i, and likewise for $Y_i[k]$.
Note: if in this expression the averaging process is omitted (only M=1 block is then used), that is, if the raw periodogram is used instead of the averaged periodogram, we necessarily obtain: $|\mathrm{Coh}_{xy}[k]| = 1\ \forall k$. The **averaged periodogram** needs to be used for calculating the **coherence function**!

5.7 Examples in fluid mechanics

The aim of this section is to give an idea of how the preceding theory is used in practice. Spectral estimation is much used as a first tool during post-processing of data in fluid mechanics and turbulence. This allows answering the following questions: are there oustanding frequencies (resonances) in the flow ? Are distant points of the flow related (multi-sensor measurement)? What is the convection velocity of the large structures?

We consider here the example of a forward facing step, which is a well studied flow with a detachment zone. See Fig. 5.11. The salient features in this flow are the vortex at the foot of the step, a detached shear layer arising from the corner of the step, the flapping motion of this shear layer, and its roll-up into vortices creating vortex shedding. The shear layer is re-attaching in average at the position x=x_r, the origin x=0 being counted from the step leading edge. An illustration of vortex shedding is shown in Fig. 5.12. This is obtained at a low Reynolds number, and at higher Re often encountered, the vortex street is not as clear cut. Some Particle Image Velocimetry measurements made at the Lab are shown in Fig. 5.13. The step is 3 cm high and the upstream flow speed U_∞ varies from 30 m/s to 50 m/s. The light sheet is set perpendicularly to the step as shown in the inset in Fig. 5.13(a). Figure 5.13(a) shows the mean axial velocity field and a foot vortex is shown. The shear layer with a jump from low velocities to large ones

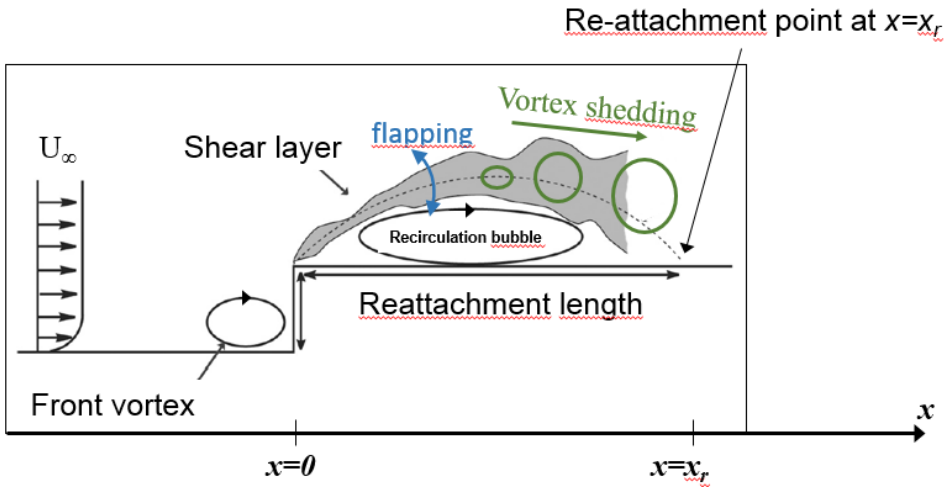


Figure 5.11: Forward step flow and its different features. x=0 is the position of the sharp leading edge of the step. x_r is the reattachment length position.

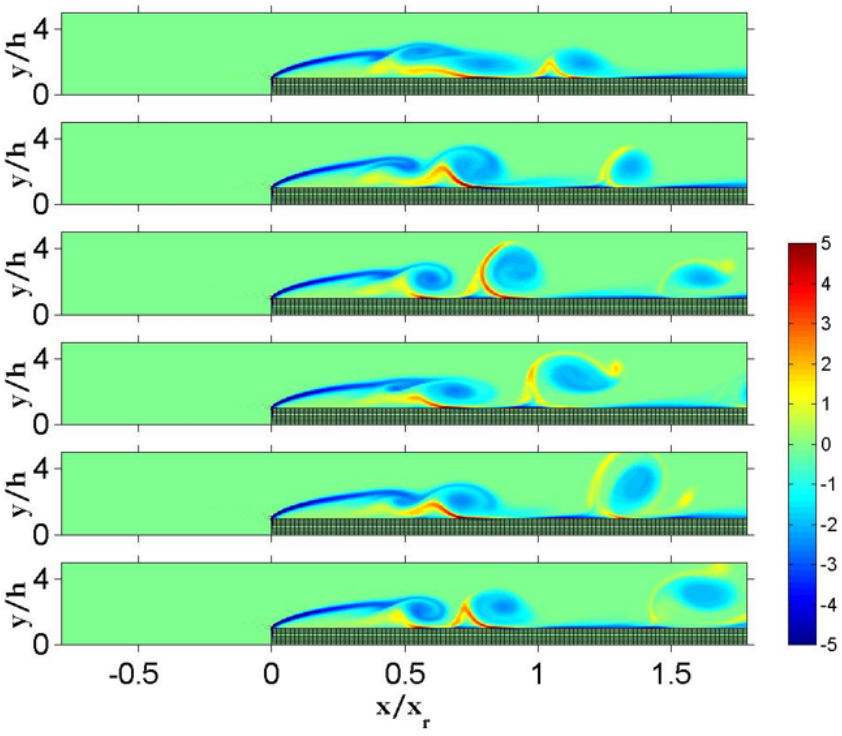


Figure 5.12: . Illustration of shedding at low Re (numerical simulation).

is also seen. Figure 5.13(b) shows the vertical velocity: this is large in the shear layer, which is due to mixing and to the flapping motion. Finally, a zoom in on the flow axial velocity in the region of the recirculation bubble is shown and shows a recirculation length of about 10 cm.
Now the tools we have seen are applied to this step flow. Some pressure sensors are flush mounted on the step, ans their position is indicated in Fig. 5.14. The first microphone axial position is

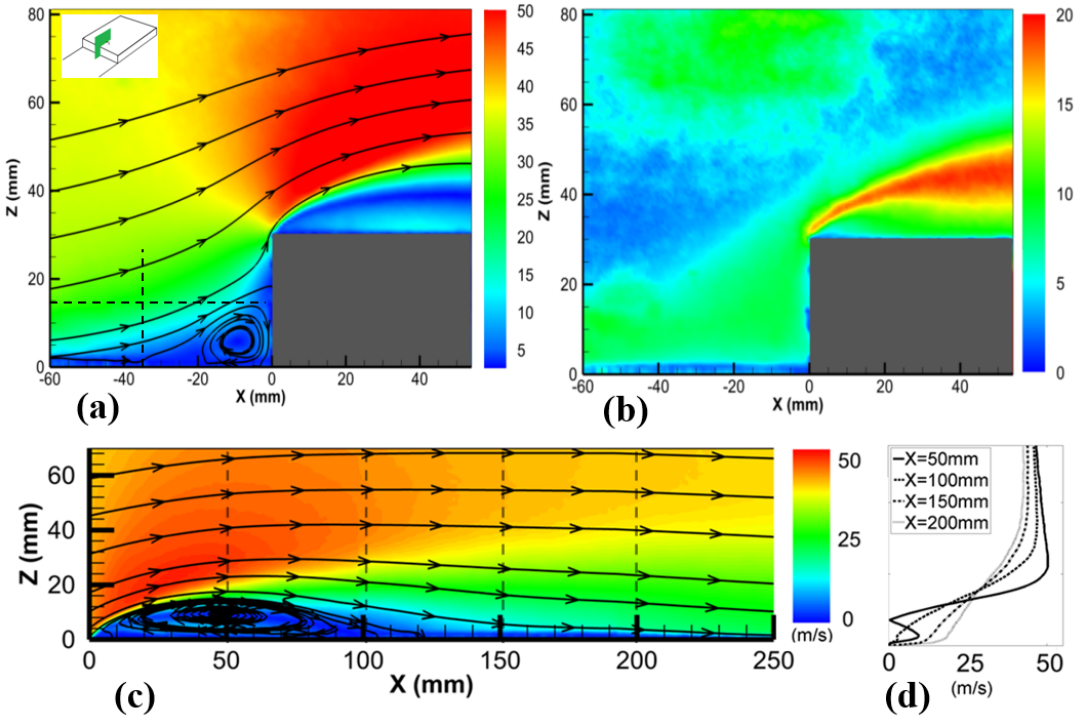


Figure 5.13: Some velocity measurements made by PIV: (a) global axial velocity field with velocity streamlines. The PIV lighting sheet in shown in the inset; (b) Vertical velocity; (c) zoom in on the axial velocity above the step and streamlines; (d) axial velocity profiles at several dowstream positions.

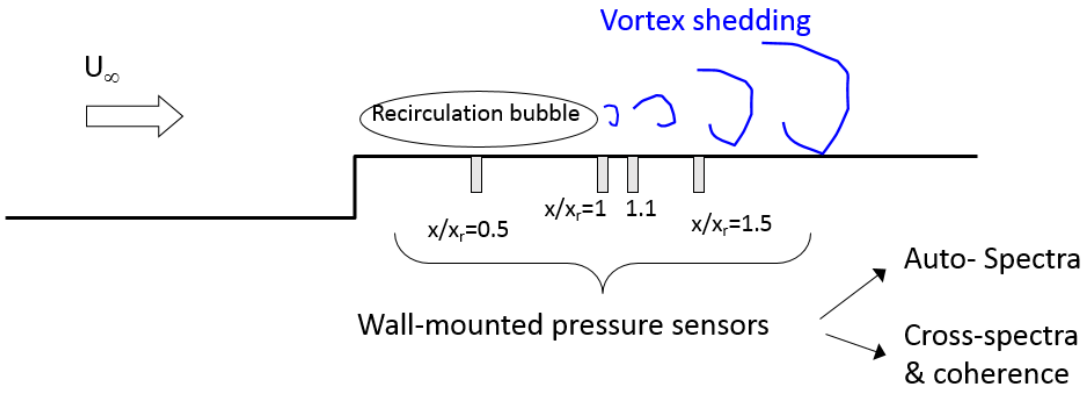


Figure 5.14: Position of the wall mounted microphones.

at the middle of the recirculation bubble (x/x_r=0.5). Other microphones are at, or past, the reattachment point. One may consider these microphones separately or in pairs. An example of auto-PSD is shown in Fig. 5.15(a) for U_∞=30 m/s at several sensor positions. Two humps appear that show the low-frequency flapping motion and the high frequency shedding motion. Figure 5.15(b) shows the auto-PSD at one position but for several flow velocity. The PSDs have an offset, which is typical of a convective character. The PSD in Fig. 5.15 involve only one sensor

measurement, and one could imagine performing these measurements separately during several experiments, with possibly only one microphone whose position would be changed. No phase is needed. Figure 5.16 shows coherence measurements, which is the normalized cross-PSD. In

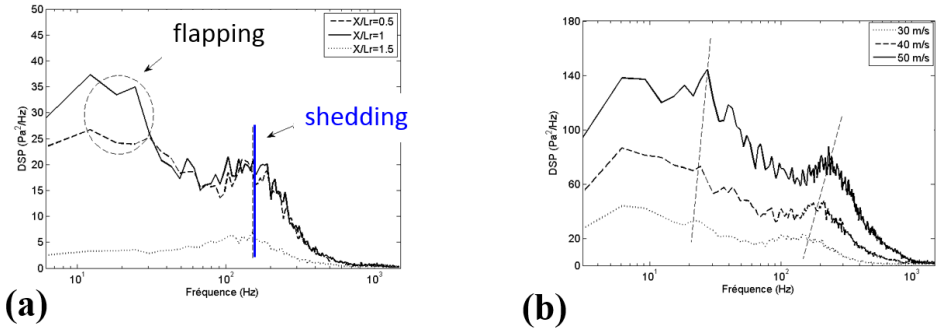


Figure 5.15: (a) PSD for one microphone (U=30m/s); (b) dependence of the PSD with the velocity.

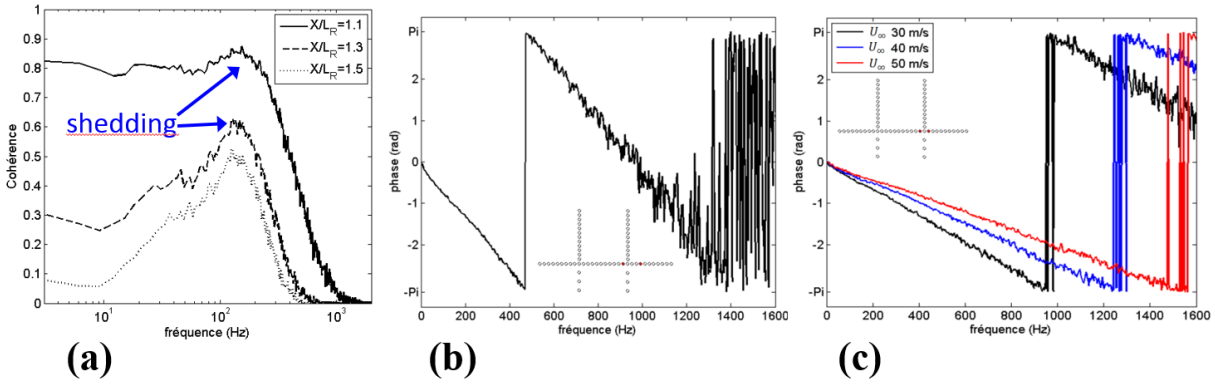


Figure 5.16: (a) Modulus of the coherence between the signal at x=1x_r and the microphone at either x=1.1x_r, x=1.3x_r, or x=1.5x_r ; (b) Phase of the coherence for a given velocity (U=30m/s); (c) Phase of the coherence for different velocities.

particular, the coherence function contains a phase that requires simultaneous recording of the two sensor signals. Here the coherence between a microphone at some value of x/x_r and the microphone at x/x_r=1 is calculated. The modulus of the coherence is shown in Fig. 5.16(a). For x/x_r=1.1 there is a high coherence on the whole low frequency range. This is because the two positions x/x_r=1 and x/x_r=1.1 are close, and the noise (or nonlinearities) had not enough time to spoil the convection motion. This motion is a linear time invariant system (see exercise below) and causes large values of the coherence. For larger values of x/x_r, the coherence decreases, but keeps high at the vortex shedding frequency, even if it is only 0.5. The conclusion is that there is a link (linear time invariant) between the sensors located downstream of the reattachment point at the shedding frequency, even if it is weaker when considering downstream positions. We suspect this link to be due to a convection motion. This can be evidenced by looking at the phase of the coherence in Fig. 5.16(b). The phase is linear with frequency, which is due to a convection motion. The slope in addition depends on the convection speed. It is verified in

Fig. 5.16(c) that the slope changes when the upsteam velocity U_∞ is changed.

This kind of post-processing will be practiced during the Matlab lab on harmonic analysis. To conclude this section, the following exercise explains why convection is considered as a LTI system.

Exercise : convection motion and LTI systems

Consider the following convection equation for pressure, with convection speed c.

$$\frac{\mathrm{d}p}{\mathrm{d}t} + c\frac{\mathrm{d}p}{\mathrm{d}x} = 0$$

Subject to the initial condition $p(x, t = 0) = p_0(x)$ where $p_0(x)$ is given, the solution is:

$$p(x, t) = p_0(x - ct)$$

The solution is obtained by taking the spatial Fourier transform of the equation, solving the equation in t, and coming back in the spatial domain by an inverse Fourier transform. If some losses are present, the equation may be changed slightly, with:

$$\frac{\mathrm{d}p}{\mathrm{d}t} + c\frac{\mathrm{d}p}{\mathrm{d}x} + \sigma p = 0$$

with σ a friction coefficient. The solution in that case is:

$$p(x, t) = p_0(x - ct)e^{-\sigma t}$$

Let's consider two sensors A and B separated by the distance Δx (A is located upstream). Then the pressures in A and B are related by:

$$p_B(t) = p_A(t - \Delta x/c)e^{-\sigma \Delta x/c}$$

Let's note $a(t) = p_A(t)$ and $b(t) = p_B(t)$, so that:

$$b(t) = \kappa a(t - \tau_0) \tag{5.43}$$

where $\tau_0 = \Delta x/c$ and $\kappa = e^{-\sigma \Delta x/c}$. Of course, the two signals are random processes, so that it would be more precise to write $a(t, \xi)$ and $b(t, \xi)$.
Equation (5.43) is a LTI relationship between a and b: b is the output of a filter with input a and impulse response $h(t)$ so that $b(t) = h(t) * a(t)$.
1. What is $h(t)$? (no calculation)
2. Deduce that $R_{ab}(\tau) = \kappa R_{aa}(\tau - \tau_0)$
3. Calculate the Cross-PSD $S_{ab}(f)$ as a function of $S_{aa}(f)$.
Show that the phase of $S_{ab}(f)$ is linear with respect to frequency with a slope to be determined.
Note: The phase of the coherence function being that of the cross-PSD, this shows that the phase of the coherence is linear with respect to frequency for a convection process, as observed

in Fig. 5.16(b).
Deduce that if the spacing Δx between the microphones is known then the convection velocity c can be measured from the slope.
Note: it has been seen in Fig. 5.16(c) that the slope of the phase of the coherence function depends on the convection velocity.

Exercices on harmonic analysis

Exercise 1: A basic random process
Consider the following random process:

$$x(t, \xi) = A(\xi) \sin(2\pi f_0 t + \phi(\xi))$$

where A and ϕ are two random variables.
Plot what could be the first 3 realizations of this process.
Calculate the auto-correlation for some particular event ξ_0.
Can this process be ergodic? What if the amplitude A does not depend on ξ?

Exercise 2: Practical PSD analysis
We want to analyze a signal $x(t)$ made of three pure tones of equal amplitudes with respective frequencies: f_1=1000 Hz, f_2=1010Hz, and f_3=5000 Hz. The analysis is performed using a digital analyzer that uses N-points Discrete Fourier Transforms (DFT). We select N=512, and a rectangular window of length T is used.

1. What is the Fourier Transform of the window. Give the width Δf of its main lobe. We suppose in the following that the resolution capacity of the window is the half-width, $\Delta f/2$.
2. What is the shape of the modulus of the FT of the windowed signal $x(t)$ (before sampling)?
3. What should be the resolution capacity for all the frequencies in $x(t)$ to be distinguished? What minimum analysis duration (T) should then be used?
4. What is the minimum sampling frequency, and what frequency resolution capacity, $\Delta f'$, does this correspond to ? Does it match that of the previous question. In case it does not, what solution would you propose?

Suppose now that our signal contains some extra noise, making it a random process. To deal with it, the Welch's method is used to calculate the power spectral density, by averaging M blocks, with no overlap. The PSD estimation error decreases with the number of blocks as:

$$\epsilon = \frac{1}{\sqrt{M}}$$

(Find which formula in the course corresponds to this).

5. Each block contains N=2048 points, and the signal is sampled at the Shannon frequency. What is the number of averages that needs to be done to obtain a 5% error on the spectrum. What is the duration of the analysis.

6. The measurement time is only 30s, what is the precision?

Exercise 3: Input-output relationships

1. The system in figure 5.17 is a linear time invariant system. Both the input and the output are random stationnary processes (the variable ξ is an experiment number).

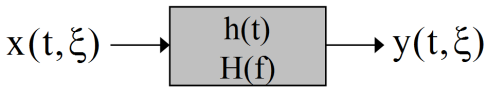


Figure 5.17: Linear Time Invariant System, with random input/output.

a) What is the relation between the input and the output?
Recall the definition of the autocorrelation, $R_{xx}(\tau)$, and that of the cross-correlation, $R_{xy}(\tau)$, for a random stationnary process. Show the following input/output relationship:

$$R_{xy}(\tau) = h(\tau) * R_{xx}(\tau)$$

What is the spectral equivalent?
b) Answer to question a) again, now supposing that the processes are ergodic.

c) The following input-output relationships are recalled:

$$S_{xy}(f) = H(f)S_{xx}(f)$$

$$S_{yy}(f) = |H(f)|^2 S_{xx}(f)$$

Show that the modulus of the coherence between signals x et y is equal to 1.

2) Consider now an evolution of the previous case (see figure 5.18), whereby some noise, $b(t,\xi)$, is added at the output. This noise is not correlated with either x or y.

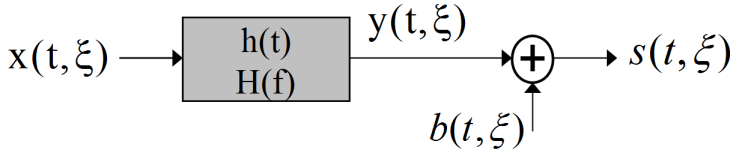


Figure 5.18: Linear Time Invariant system, with random input/output, and some noise on the output.

a) Show that the correlation and spectral density are distributive, that is:

$$C_{u(w+z)} = C_{uw} + C_{uz} \qquad S_{u(w+z)} = S_{uw} + S_{uz}$$

b) What are the values of C_{bx} and S_{bx}?
c) Show that the modulus of the coherence between x and s is:

$$Coh_{xs}(f) = \frac{1}{\sqrt{1 + 1/r(f)}}$$

where $r(f) = S_{yy}(f)/S_{bb}(f)$ is the signal-to-noise ratio.

3) A random white noise $b(t, \xi)$ is used as input for some linear time invariant system, supposed to be a low-pass filter. The output $s(t, \xi)$ is also a random noise (it is filtered though).
What is the meaning of a power spectral density? Sketch the PSDs $S_{bb}(f)$ et $S_{ss}(f)$ versus frequency, f. What does the surface under the curve mean? What is the modulus to the coherence between the signals b et s?

Solutions

Exercise 1: A basic random process

Figure 5.19 shows the first 3 realizations of the process.

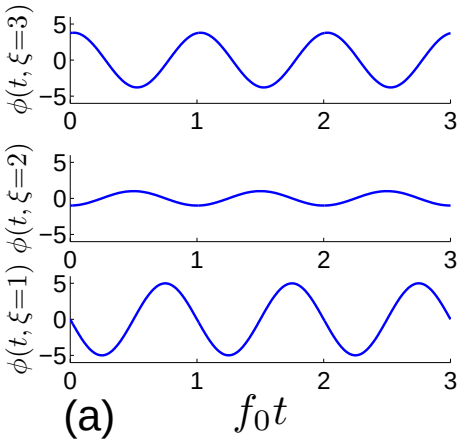


Figure 5.19: **First 3 realizations of the random process.**

For some fixed realization ξ_0, we get a time signal whose correlation is calculated easily:

$$C_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t,\xi_0)x(t+\tau,\xi_0)dt \tag{5.44}$$

$$= \lim_{T\to\infty} \frac{1}{T}\int_0^T A^2(\xi_0)\sin(2\pi f_0 t+\phi(\xi_0))\sin(2\pi f_0(t+\tau)+\phi(\xi_0))dt \tag{5.45}$$

Using relation $\sin a - \sin b = 1/2(\cos(a-b)-\cos(a+b))$, we get:

$$C_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{A^2(\xi_0)}{2}\left[\cos(2\pi f_0\tau)-\cos(2\pi f_0(2t+\tau)+2\phi(\xi_0))\right]dt \tag{5.46}$$

$$= \frac{A^2(\xi_0)}{2}\cos(2\pi f_0\tau) \tag{5.47}$$

A periodic signal has a periodic auto-correlation (with same period).

For a random signal the auto-correlation is given by: $R_{xx}(\tau) = \mathrm{E}\left[x(t,\xi)x(t+\tau,\xi)\right]$. If the signal was ergodic (including stationarity), we would have: R_{xx}=C_{xx}, where C_{xx} could be calculated from *any* realization of the signal. For this to be the case, the auto-correlation should not depend on whatever value of ξ_0 is chosen; this is not true according to Eq. (5.47). Hence, the process is not ergodic.
If the amplitude A does not depend on the realization number ξ, then the process is ergodic.

Exercise 2: Practical PSD analysis

1. The FT of the window is:

$$FT\left[Rect_T(t)\right](f) = \frac{\sin(\pi f T)}{\pi f} = T\text{sinc}(\pi f T)$$

We have $\Delta f = 1/T$.
The resolution capacity is thus $\frac{\Delta f}{2} = \frac{1}{T}$.
Generally speaking, the resolution in the frequency domain, which is the "small interval" in the frequency domain, is the inverse of the "large interval" in the time domain, which is T.

2. See figures 5.20 and 5.21, for T=0.01s and T=0.1s respectively. The sampling frequency is 15kHz, and only the positive frequencies, from 0 to 7.5 kHz, are represented. Depending on T, we will or we won't be able to separate the two frequencies 1000Hz and 1010Hz.

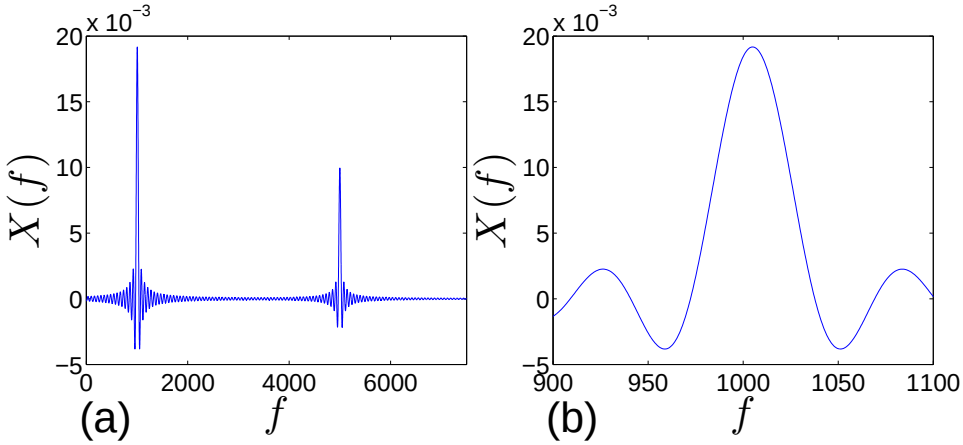


Figure 5.20: **FT of the signal for a small analysis window length, T=0.01s. The plot on the right is a zoom in of the one on the left.**

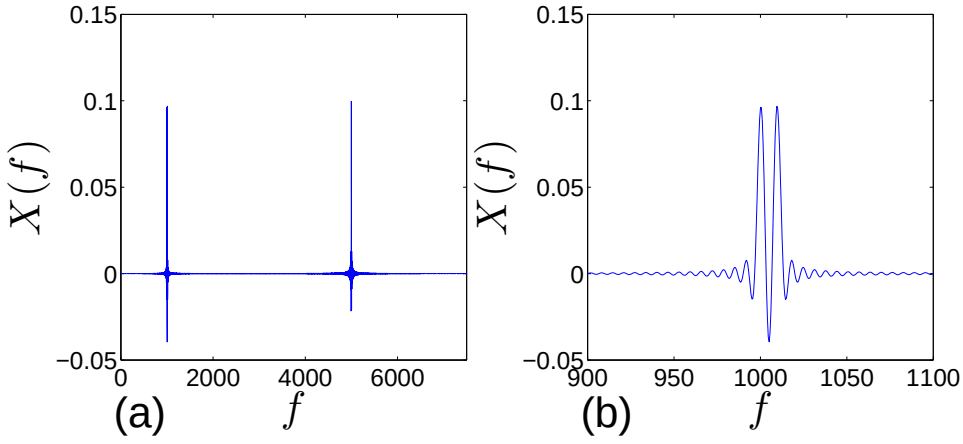


Figure 5.21: **FT of the signal for a large analysis window length, T=0.1s. The plot on the right is a zoom in of the one on the left.**

3. To separate the frequencies 1000Hz and 1010Hz, we need a resolution of 10Hz, that is, T=0.1s.

4. The maximal frequency contained in the signal is 5kHz. Thus, by the Shannon theorem, the sampling frequency should be at least 10kHz. When using the DFT (without zero padding), we deal with discrete frequencies, and the frequency step is $\Delta f' = f_e/N$ where f_e is the sampling frequency. For f_e=10kHz, $\Delta f'$=10000/512=19.5 Hz. This is larger than the value Δf=10Hz which is required. One solution is to increase the number of points, N, which is equivalent to increasing the window length duration (for a given sampling frequency).

5. For ϵ=0.05, we need M=1/0.05^2=400 averages.
The duration of the analysis is then (with no overlap):
$T = \underbrace{M}_{\substack{\text{number} \\ \text{of blocks}}} \quad \underbrace{NT_e}_{\substack{\text{duration of} \\ \text{one block}}} = MN/f_e$= 400*2048/10000 = 81.9s.

6. For T=30s, the number of blocks is $M = T/N * f_e \sim 146$, meaning an error ϵ=8.2%.

Exercise 3: Input-output relationships

1.a) The output is given by the convolution product: $y = h * x = x * h$.
The autocorrelation, which in general is a function of both time and lag, is a function of lag only for a stationary process. It is expressed by:

$$R_{xx}(\tau) = \mathrm{E}\left[x(t,\xi)x(t+\tau,\xi)\right] \qquad \forall t$$

The cross-correlation is:

$$R_{xy}(\tau) = \mathrm{E}\left[x(t,\xi)y(t+\tau,\xi)\right] \qquad \forall t$$

We start from the definition and express the convolution product:

$$y(t,\xi) = h * x = \int_0^\infty h(u)x(t-u,\xi)du$$

If we substitute $t+\tau$ for t:

$$y(t+\tau,\xi) = \int_0^\infty h(u)x(t+\tau-u,\xi)du$$

And thus:

$$x(t,\xi)y(t+\tau,\xi) = \int_0^\infty h(u)x(t,\xi)x(t+\tau-u,\xi)du$$

Now:

$$\begin{aligned} R_{xy}(\tau) &= \mathrm{E}\left[x(t,\xi)y(t+\tau,\xi)\right] && (5.48)\\ &= \int_0^\infty h(u)\mathrm{E}\left[x(t,\xi)x(t+\tau-u,\xi)\right]du && (5.49)\\ &= \int_0^\infty h(u)R_{xx}(\tau-u)du && (5.50)\\ &= h(\tau) * R_{xx}(\tau) && (5.51) \end{aligned}$$

which is the wanted result. To obtain the equivalent in the spectral domain, we take the FT and use the Wiener-Khintchine theorem, which we recall states that $FT[R_{xx}(\tau)] = S_{xx}(f)$. We obtain:

$$S_{xy}(f) = \mathrm{FT}\left[R_{xy}\right] = \mathrm{FT}\left[h * R_{xx}\right] = \mathrm{FT}\left[h\right]\mathrm{FT}\left[R_{xx}\right]$$

which gives:

$$S_{xy}(f) = H(f)S_{xx}(f)$$

1.b) If the process is ergodic, we can calculate directly:

$$C_{xy}(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)y(t+\tau)dt$$

The latter is performed on any realization ξ, that is, on a single measurement. ξ is therefore omitted. As before we have $y = h * x$, which gives:

$$y(t+\tau) = \int_0^\infty h(u)x(t+\tau-u)du$$

and thus:

$$x(t)y(t+\tau) = x(t)\int_0^\infty h(u)x(t+\tau-u)du$$

The cross-correlation is:

$$\begin{aligned}
C_{xy}(\tau) &= \lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)y(t+\tau)dt && (5.52)\\
&= \lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)\int_0^\infty h(u)x(t+\tau-u)dudt && (5.53)\\
&= \int_0^\infty h(u)\underbrace{\lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)x(t+\tau-u)dt}_{=C_{xx}(\tau-u)}du && (5.54)\\
&= \int_0^\infty h(u)C_{xx}(\tau-u)du && (5.55)
\end{aligned}$$

This gives again:

$$C_{xy}(\tau) = h(\tau) * C_{xx}(\tau)$$

1.c) The coherence is a quantity that pertains to the spectral domain. Its modulus is given by:

$$|Coh_{xy}(f)| = \left|\frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}}\right| = \frac{|H(f)|S_{xx}(f)}{\sqrt{S_{xx}(f)|H(f)|^2S_{xx}(f)}} = 1$$

The signal y is the output of a linear time invariant system whose input is x. There is thus a linear time invariant relation between x and y, and we find the result (given during the course) that the modulus of the coherence between two such signals is 1. It means that the frequency content of the signal y at any frequency f is linked to the frequency content of x at the same frequency.

2. a) For a random process: $R_{xy}(\tau) = \mathrm{E}\left[x(t,\xi)y(t+\tau,\xi)\right]$. Then we have:

$$R_{u(w+z)}(\tau) = \mathrm{E}\left[u(t,\xi)w(t+\tau,\xi) + u(t,\xi)z(t+\tau,\xi)\right]$$

that is:

$$R_{u(w+z)}(\tau) = R_{uw}(\tau) + R_{uz}(\tau)$$

Taking the FT and using the Wiener-Khintchine theorem:

$$S_{u(w+z)}(f) = \mathrm{FT}\left[R_{u(w+z)}(\tau)\right] = \mathrm{FT}\left[R_{uw}(\tau) + R_{uz}(\tau)\right] = S_{uw}(f) + S_{uz}(f)$$

Note: if we had supposed that the process is ergodic:

$$\begin{aligned} C_{u(w+z)} &= \lim_{T\to\infty}\frac{1}{T}\int_0^T u(t)\left[w(t+\tau)+z(t+\tau)\right]dt && (5.56)\\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T u(t)w(t+\tau)dt + \lim_{T\to\infty}\frac{1}{T}\int_0^T u(t)z(t+\tau)dt && (5.57)\\ &= C_{uw}(\tau) + C_{uz}(\tau) && (5.58)\end{aligned}$$

The results above are not modified.

3. b) The two signals x and b are not correlated, which means: $R_{xb} = 0$. This also implies $S_{xb} = 0$. For the same reason, $R_{yb} = 0$ and $S_{yb} = 0$.

3.c) The coherence is given by:

$$Coh_{xs}(f) = \frac{S_{xs}(f)}{\sqrt{S_{xx}(f)S_{ss}(f)}}$$

We use the distributivity of the PSD. First we have:

$$S_{xs} = S_{x(y+b)} = S_{xy} + \underbrace{S_{xb}}_{=0} = S_{xy} = H(f)S_{xx}(f) = H(f)\frac{S_{yy}}{|H(f)|^2}$$

and also:

$$S_{xx}(f) = \frac{S_{yy}}{|H(f)|^2}$$

and:

$$S_{ss} = S_{(y+b)(y+b)} = S_{yy} + \underbrace{S_{yb}}_{=0} + \underbrace{S_{by}}_{=0} + S_{bb}$$

By injecting those relations into the Coherence, we obtain:

$$\begin{aligned} Coh_{xs}(f) &= \frac{\left|H(f)\frac{S_{yy}}{|H(f)|^2}\right|}{\sqrt{\frac{S_{yy}}{|H(f)|^2}(S_{yy}+S_{bb})}} && (5.59)\\ &= \frac{1}{\sqrt{1+\frac{S_{bb}}{S_{yy}}}} && (5.60)\\ &= \frac{1}{\sqrt{1+\mathrm{SNR}^{-1}}} && (5.61)\end{aligned}$$

where SNR=S_{yy}/S_{bb} is the Signal to Noise Ratio (it depends on f since there may be more noise at some frequencies than at some others).
In the ideal case when no noise is present: SNR$\rightarrow \infty$, and: $Coh_{xs}(f) = 1 \quad \forall f$, which is the same result as in question 1.c).
In the bad situation in which SNR$\rightarrow \infty$, then $Coh_{xs}(f) = 0 \quad \forall f$. The noise is so important that there is indeed no relation between the input and the output.
For a moderate SNR, the effect of noise is to decrease the coherence between signals x and s. The signal x and y are related by a linear time-invariant relationship, but due to noise the relation between x and $s = y + b$ is only partly linear time-invariant. The less noise, the more important the coherence (or to put it another way, the more linear and time-invariant the relation between x and s).

3) We know that (Wiener-Khintchine): $R_{xx}(\tau) = \mathrm{FT}^{-1}[S_{xx}(f)] = \int_{-\infty}^{\infty} S_{xx}(f) e^{j2\pi f t} df$. In particular: $R_{xx}(0) = \int_{-\infty}^{\infty} S_{xx}(f) df$. And we also know that:

$$R_{xx}(\tau) = \mathrm{E}[x(t,\xi)x(t+\tau,\xi)]$$

which implies

$$R_{xx}(0) = \mathrm{E}\left[x^2(t,\xi)\right]$$

which is the mean square of the signal. By combining the two results, we obtain:

$$\int_{-\infty}^{\infty} S_{xx}(f) = \mathrm{E}\left[x^2(t,\xi)\right] = \text{mean power of the signal}$$

The power spectral density $S_{xx}(f)$ is a power per frequency band df, and by integrating this power density over all possible frequencies, we simply obtain the power of the signal.
Note: for a stationary random process, one has to use the power, since the energy is not finite (that is, the integral defining the energy, $\int_{-\infty}^{\infty} x^2(t,\xi)dt$, is not finite, but that defining the power, $\lim_{T\rightarrow\infty} 1/T \int_0^T x^2(t,\xi)dt$, is finite).

The power spectral density for the white noise $b(t,\xi)$ on the input is by definition: $S_{bb}(f) = 1 \quad \forall f$.
By making use of the input-output relationships, the PSD for the output is:

$$S_{ss}(f) = |H(f)|^2 S_{bb}(f) = |H(f)|^2 \quad \forall f,$$

whose meaning is that by putting a white noise at the input of a system, one obtains the modulus of the frequency response squared at the output. Hence, just by looking at the output, one knows how the system is going to filter any signal. At least one knows how the magnitude is going to be affected at some frequency by the system. We recall that, at a given frequency, a system is characterized by an amplification factor, and by a phase delay. To have access to the latter, a cross spectra between the input and the output is necessary.

The signal b and s are related by a linear time-invariant relationship, hence the modulus of the coherence between these two signals is 1 for all frequencies. The noise (which we repeat is nothing but a particular random signal) has a good coherence with a filtered version of itself ! (at least if the filter is linear time-invariant.)

6 Time-Frequency Analysis: General concepts

The Fourier transform has limitations that are now presented. The concept of an instantaneous frequency is introduced. The instantaneous amplitude and frequency can be found using the Hilbert transform and the analytic signal associated with a real signal. The instantaneous frequency is useful mainly for signals having a single frequency component at any time. For more complex signals, the tools introduced in the following chapters are more appropriate.
An introduction to time-frequency analysis can be found in the first chapters of [6], and in reference [16]. Classical papers on instantaneous frequency are [41] and [5].

6.1 Limitations of the FT and the PSD

6.1.1 Time XOR Frequency

Let us recall the definition of the Fourier transform:

$$X(f) = \mathrm{FT}[x(t)] = \int_{-\infty}^{\infty} x(t) e^{-\mathrm{j}2\pi f t} \mathrm{d}t$$

In this representation, time is integrated out, and the whole signal is needed to calculate the Fourier transform. The inverse Fourier transform is:

$$x(t) = \mathrm{FT}^{-1}[X(f)] = \int_{-\infty}^{\infty} X(f) e^{\mathrm{j}2\pi f t} \mathrm{d}f$$

Here, frequency is integrated out. In harmonic analysis time and frequency are mutually exclusive: it is a time XOR (exclusive) frequency representation. One consequence of this is the following statement:

> The Fourier transform tells us what frequencies are contained in the signal, but is does not tell when a particular frequency appears.

This statement is illustrated by two examples below.

6.1.2 Fourier analysis and singular functions

The previous statement can be extended to singularities. Typically, we can state:

> The Fourier transform tells us if the signal is discontinuous, but does not tell us at which times it is discontinuous.

This conclusion holds for singularities in general. A signal is called singular when it is not continuous, or continuous but not differentiable, at some point. Take for example the rectangular window function, $\Pi_T(t)$: this function is discontinuous at $\pm T/2$. Discontinuities cause the spectrum (the FT modulus) to have a low decrease rate at high frequencies. Namely, is was seen in section 3.8 that the modulus of the FT decreases only as $1/f$ as $f \to \infty$. [1]
Conversely, if one considers the asympotic behaviour of the FT of some function as $f \to \infty$ and observes that this FT decreases as $1/f$, then one knows that the function is discontinuous. However, one does not know *where* the function is discontinuous. And in the particular case of the $\Pi_T(t)$ function, one does not know that the function is actually continuous almost everywhere, except at two points, $\pm T/2$.

6.1.3 Examples

Example 1: a cosine with Gaussian amplitude

Let's consider the signal:

$$x(t) = e^{-500(t-0.5)^2} \cos(2\pi.440.t) = a(t) \cos(2\pi.440.t) \tag{6.1}$$

where $a(t)$ is a time-varying envelope. Figure 6.1(a) shows the signal and Fig. 6.1(b) shows its PSD. The Fourier spectrum consists of a bunch of components with frequencies around 440 Hz. Each of these components is a constant amplitude harmonic wave lasting forever. These harmonic components are delocalized in time. There is actually no way to determine from the PSD *when* the 440 Hz component appears. Of course it is possible to reconstruct the signal from these delocalized components (this is the Inverse Fourier Transform!) by an interference process. The phase of the FT, not included in the DSP, is then very important. However, this interference process is not intuitive. An example of a representation that would be desirable is shown in Fig. 6.2. The spectrogram represented in this figure plots the signal energy density (whose a precise definition will be given later, see Chapter 7) in the time-frequency plane. It tells us that a frequency of about 500 Hz appears between time 0.4 and 0.6 s, with maximal amplitude at 0.5 s. The spectrogram is one of the many time-frequency representations that will be studied later.

Example 2: a linear chirp

A linear chirp is defined by:

$$x(t) = \cos\left(2\pi 40 \frac{t^2}{2}\right) \tag{6.2}$$

[1] To be more specific, the FT of $\Pi_T(t)$ is $T\text{sinc}(\pi f T) = \sin(\pi f T)/(\pi f)$, which exhibits this asymptotic behaviour. Remark: what the asymptotic decrease of the spectrum tells us about the signal may have important consequences in fluid mechanics. Hence, the turbulent energy spectrum $E(k) = k^{-5/3}$, which is not limited by a cutoff wavenumber when the fluid is inviscid, indicates that the turbulent velocity field contains singularities worse than discontinuities.

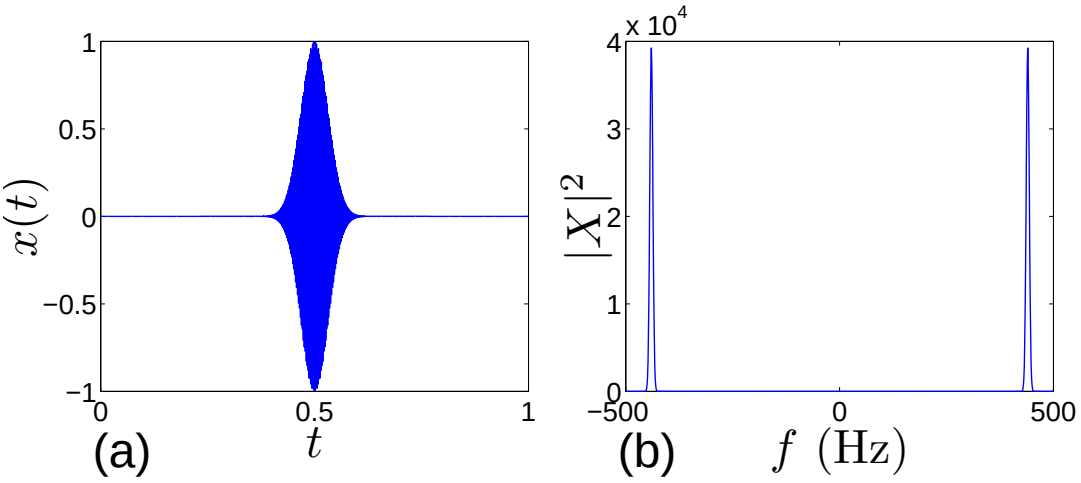


Figure 6.1: (a) Harmonic signal with a Gaussian amplitude (see Eq. (6.1)); (b) The modulus squared of its FT ($\sim$ the PSD).

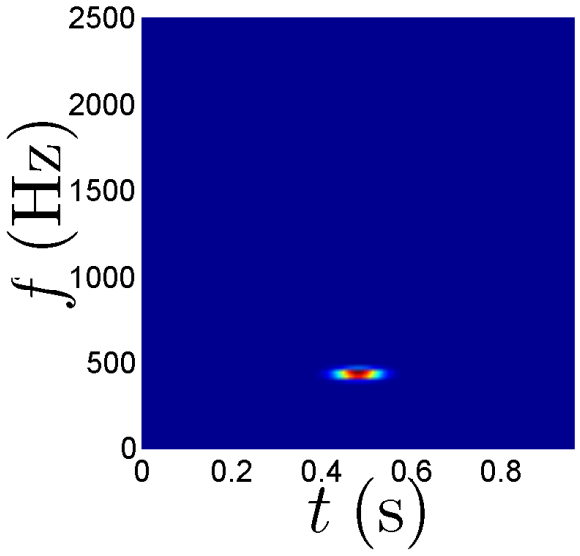


Figure 6.2: Spectrogram of a harmonic signal with a Gaussian amplitude.

We will see later in this lecture that the instantaneous frequency of this signal is $f_i(t)$=40t. Thus, from t=0s to t=1s the instantaneous frequency f_i increases from 0 Hz to 40 Hz. The signal is shown in Fig. 6.3(a) where we see the frequency increasing. The modulus and phase of the Fourier transform of this signal are shown in Fig. 6.3(b) and 6.3(c). Loosely, the modulus of the FT (even) is constant over the frequency band [-40Hz; 40Hz]. The phase is odd (as it should be for a real signal) and gives little intuitive feeling about the signal.

Now let us define $Y(f)$ by: $Y(f) = |X(f)| \; \forall f$. That is to say, the phase of Y is null for all f and its modulus is that of $X(f)$. The modulus and phase of $Y(f)$ are shown in Fig. 6.4(a) and 6.4(b). Now let's calculate the signal y that is the inverse Fourier transform of Y. The

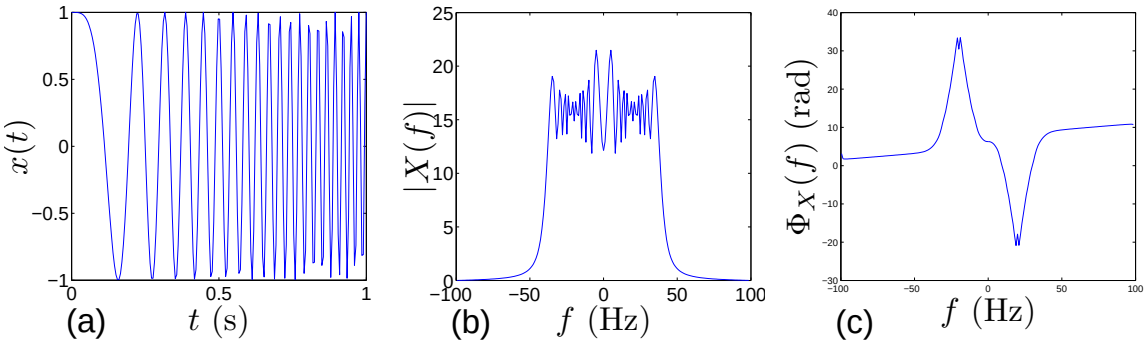


Figure 6.3: (a) Linear chirp in Eq. (6.2); (b) the modulus of its FT; (c) the phase of its FT.

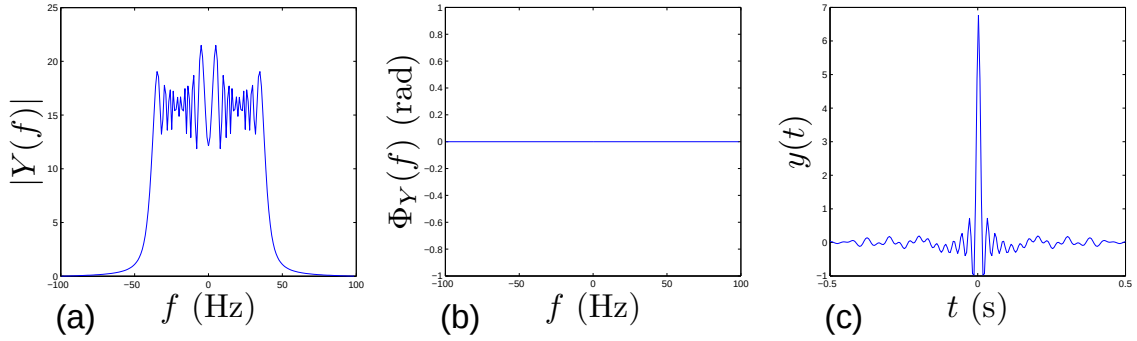


Figure 6.4: (a) Modulus of $Y(f)$; (b) Phase of $Y(f)$; (c) Inverse Fourier transform of $Y(f)$.

signal y is shown in Fig. 6.4(c), it is approximately a sinc function (can you say why?), and is does not resemble x at all. Hence, x and y have the same PSD but are very different. Of course the PSD does not give all the information in the spectral domain since the phase is necessary and the phase is here responsible for x and y being so different. However the PSD is usually the main quantity one plots and from which one would like to get a picture of the signal. A representation that would be desirable is shown in Fig. 6.5 where the spectrograms of x and y are plotted. The spectrogram displays an energy content (as in the PSD), yet allows having some insight about the time characteristics of the signal, as a result of the energy being plotted in the time-frequency domain.

6.1.4 Typical signals in time-frequency analysis

The signal in example 1 of the previous subsection is of the type:

$$a(t)\cos(2\pi f_0 t)$$

It has a constant frequency and a time-varying amplitude. The signal in example 2 is of the type:

$$a\cos(2\pi\alpha(t)t)$$

It has a constant amplitude and a time-varying frequency.
What these examples show is that the basis used in Fourier Transform, which is made of waves

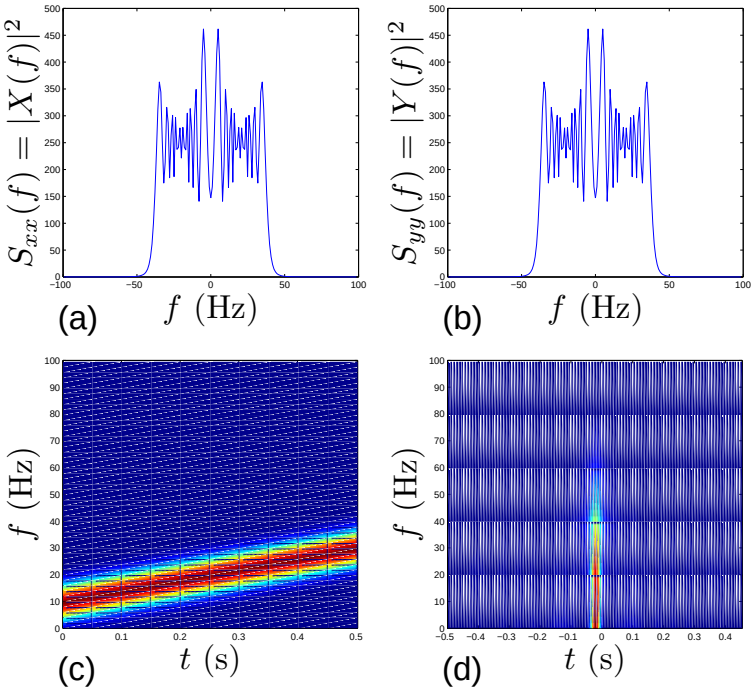


Figure 6.5: (a) PSD of x; (b) PSD of y; (c) Spectrogram of x; (d) Spectrogram of y. This figure shows that x and y, while different, have the same DSP. By contrast, the spectrogam shows a clear difference between x and y.

$A\cos(2\pi f_0 t)$ with A and f_0 constant, does not fit well those signals that have a time-varying amplitude or a time-varying frequency, or both. The later are better represented by a joint time-frequency analysis. A typical signal in time-frequency analysis is the chirp, a signal whose frequency is increasing or decreasing with time. A **linear chirp** is:

$$x(t) = A\cos\left(2\pi f_0 \frac{t^2}{2}\right)$$

or in complex form:

$$x(t) = Ae^{\mathrm{j}\left(2\pi f_0 \frac{t^2}{2}\right)}$$

This signal has a linear frequency modulation, meaning its frequency depends linearly on time (the reason for this is explained in the next section).
A signal that has a sinusoidal frequency modulation is of the form:

$$x(t) = A\cos\left(2\pi f_0 t + \frac{f_d}{f_c}\sin(2\pi f_m t + \Theta) + \Phi\right)$$

Another typical signal is one having a Gaussian amplitude and a linear frequency modulation:

$$x(t) = Ae^{-\beta(t-t_m)^2}\cos\left(2\pi f_0 \frac{t^2}{2}\right)$$

or in complex form:

$$x(t) = Ae^{-\beta(t-t_m)^2}e^{\mathrm{j}2\pi f_0 \frac{t^2}{2}}$$

6.1.5 Domains of applications

There are many practical cases when a time-frequency analysis is useful, in particular in acoustics. A musical score (see Fig. 6.6) is a time-frequency representation. Time-frequency analysis

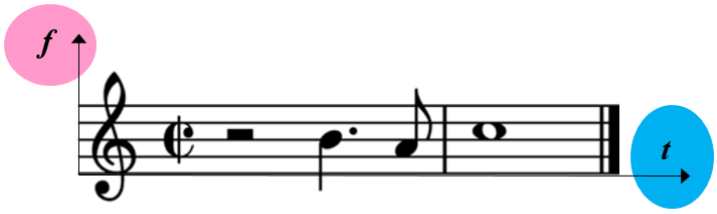


Figure 6.6: A musical scores.

is much used for speech analysis (Fig. 6.7). This can be used for speech recognition, or even for analyzing people's emotion. Sounds emitted by many animals, ranging from whales to bats,

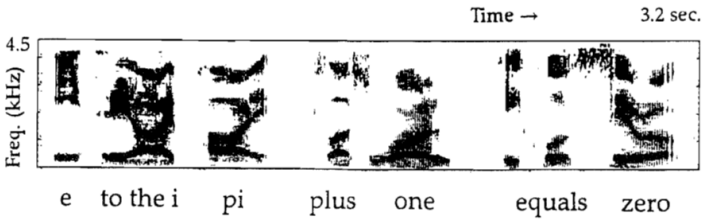


Figure 6.7: Time-frequency plot of a speech signal: $e^{i\pi}+1=0$. From Cohen[14].

are better understood in the time-frequency domain (see Fig. 6.8). It is also used in underwater acoustics, or in damage monitoring. In fluid mechanics, the prevailing technique nowadays is the wavelet transform (which is actually a time-scale analysis) that can be used to detect structures or transient signals (See Chapter 9).

6.2 Hilbert transform, Analytic signal, Instantaneous frequency

6.2.1 Hilbert transform

The Hilbert transform (HT) transforms a time-signal into another time-signal. In the time domain, it consists of a convolution with $1/(\pi t)$:

$$\mathrm{HT}[x(t)](t) = x(t) * \frac{1}{\pi t} = \frac{1}{\pi}\mathtt{vp}\int_{-\infty}^{\infty}\frac{x(\tau)}{(t-\tau)}\mathrm{d}\tau \tag{6.3}$$

In the frequency domain, using the convolution theorem:

$$\mathrm{FT}\left[\mathrm{HT}[x(t)]\right] = \mathrm{FT}[x]\cdot\mathrm{FT}\left[\frac{1}{\pi t}\right]$$

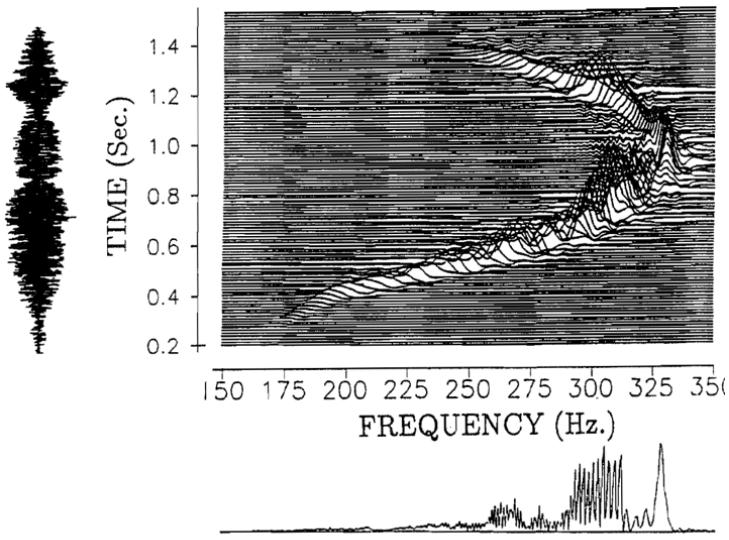


Fig. 5.1 A time-frequency plot of a sound made by a Bowhead whale. The wave form is on the left plot with time increasing upwards. The energy density spectrum is below the main figure and indicates which frequencies existed for the duration. The main figure is a joint time-frequency plot and shows how the frequencies change with time. (From the work of C. Rosenthal and L. Cohen.)

Figure 6.8: Time-frequency analysis of a whale sound. From Cohen[14].

And we have:

$$\mathrm{FT}\left[\frac{1}{\pi t}\right] = -\mathrm{j}\ \mathrm{sgn}(f) \begin{cases} \text{-j} & \text{for } f > 0 \\ 0 & \text{for } f = 0 \\ \mathrm{j} & \text{for } f < 0 \end{cases} \tag{6.4}$$

Letting $X(f) = \mathrm{FT}[x]$, we thus have:

$$\mathrm{FT}\left[\mathrm{HT}[x(t)]\right] = -\mathrm{j}\ \mathrm{sgn}(f)X(f) = \begin{cases} \text{-j}X(f) & \text{for } f > 0 \\ 0 & \text{for } f = 0 \\ \mathrm{j}X(f) & \text{for } f < 0 \end{cases} \tag{6.5}$$

In words, by taking the HT of a signal, the FT of this signal is rotated by -$\pi/2$ for positive frequencies, and by $\pi/2$ for negative frequencies. The interest of this will be seen in the next subsection.

Classical HT are:

$$\mathrm{HT}[\cos(2\pi f_0 t)] = \sin(2\pi f_0 t)$$

$$\mathrm{HT}[\sin(2\pi f_0 t)] = -\cos(2\pi f_0 t)$$

It is also easy to see that:

$$\mathrm{HT}[\mathrm{HT}[x]] = -x$$

meaning that the inverse Hilbert transform (IHT) verifies:

$$\mathrm{IHT} = -\mathrm{HT}$$

6.2.2 Analytic signal

Definition: a complex signal $z(t)$ is said to be **analytic** when:

$$Z(f) = 0 \qquad \text{for } f < 0 \tag{6.6}$$

where

$$Z(f) = \mathrm{FT}[z]$$

Hence, all the components of an analytical signal have a positive frequency.

A real signal $x(t)$ has components for both positive and negative frequencies, since $X(-f) = X^*(f)$. Hence, it cannot be analytic. However, it is possible to associate to this real signal a complex signal that is analytic. For a real signal, we define its **analytic associate** $z_x(t)$ by:

$$z_x(t) = x(t) + \mathrm{j}\mathrm{HT}[x] \tag{6.7}$$

The analytic associate of a real signal is a complex signal whose real part is the real signal and whose imaginary part is the Hilbert transform of the real signal.

Its Fourier transform $Z_x(f) = \mathrm{FT}[z_x]$ is:

$$Z_x(f) = X(f) + \mathrm{j}\left(-\mathrm{j}\ \mathrm{sgn}(f)X(f)\right) = \begin{cases} 2X(f) & \text{for } f > 0 \\ 0 & \text{for } f = 0 \\ 0 & \text{for } f < 0 \end{cases} \tag{6.8}$$

It is thus verified that $Z(f) = 0$ for $f < 0$ and $z_x(t)$ is analytic. $X(f)$ and $Z_x(f)$ are shown in Fig. 6.9. Since we have $X(-f) = X^*(f)$ for a real signal $x(t)$, this signal can be characterized

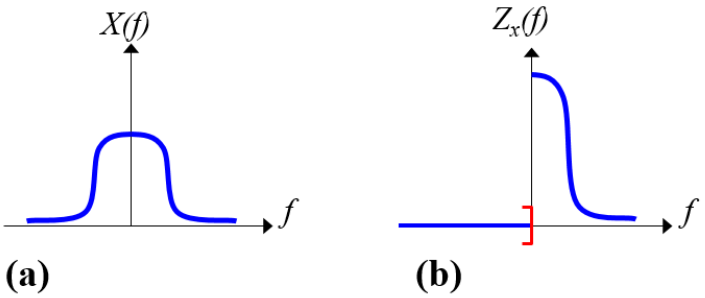


Figure 6.9: (a) Fourier transform of a signal: (b) Fourier transform of its analytic associate.

by the value of $X(f)$ for $f > 0$ only, and this information is contained in $Z_x(f)$. Hence, no information is lost when computing the analytic associate of a signal, and a signal can be recovered from its analytic associate.

We now see that the Hilbert transform is useful to build the analytic associate of a real signal. The advantage of having an analytic associate is explained in the next subsection.

Classical analytic associates are:

$$x(t) = \cos(2\pi f_0 t) \Longrightarrow z_x(t) = e^{\mathrm{j}2\pi f_0 t}$$

$$x(t) = \sin(2\pi f_0 t) \Longrightarrow z_x(t) = -\mathrm{j}e^{\mathrm{j}2\pi f_0 t}$$

6.2.3 Instantaneous frequency

In the present section, we define the instantaneous frequency of a signal. The definition depends on the signal being complex or real.

Complex signal:

Any complex number is defined by a modulus and a phase. Thus, for a complex signal $y(t)$, we have the unique decomposition

$$y(t) = a(t)e^{\mathrm{j}\phi(t)}$$

with $a(t)$ the time-varying amplitude and $\phi(t)$ the time-varying phase.

The **instantaneous frequency** of $y(t) = a(t)e^{\mathrm{j}\phi(t)}$ is defined by:

$$f_i(t) = \frac{1}{2\pi}\frac{\mathrm{d}\phi}{\mathrm{d}t} \tag{6.9}$$

The **instantaneous amplitude** is $a(t)$.

Real signal:

Defining an instantaneous frequency for a real signal is not as straightforward as for a complex signal. A real signal made of a monochromatic wave of constant amplitude has the following representation:

$$x(t) = a\cos(2\pi f_0 t)$$

This may be written as $x(t) = a\cos(\phi(t))$ with $\phi(t) = 2\pi f_0 t$, and in this case the (constant) frequency may be obtained from:

$$f_0 = \frac{1}{2\pi}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}$$

Let's see if this could be extended to a more general signal. Consider the real signal:

$$x(t) = a(t)\cos(\phi(t))$$

From this expression, it is tempting to define again an instantaneous frequency by $\frac{1}{2\pi}\mathrm{d}\phi/\mathrm{d}t$. However, this is not correct because for a real signal, the representation $(a(t);\phi(t))$ is not unique. To see this, write:

$$x(t) = a(t)\frac{1}{b(t)}b(t)\cos(\phi(t)) = c(t)\cos(\psi(t))$$

with $c(t) = a(t)/b(t)$ and $\psi(t) = \mathrm{acos}[b(t)\cos(\phi(t))]$. This is the reason why it is desirable to associate to a real signal a complex signal, for which the amplitude and phase are well (uniquely) determined. This is done using the analytic associate $z_x(t)$ presented in the previous subsection:

$$x(t) \Longrightarrow z_x(t) = x(t) + \mathrm{jHT}[x(t)]$$

This analytic associate being complex, it can be written:

$$z_x(t) = |z_x(t)|e^{\mathrm{j}\phi(t)}$$

where the modulus and phase are uniquely determined. The instantaneous frequency and instantaneous amplitude are then defined by $|z_x(t)|$ and $\frac{1}{2\pi}\frac{\mathrm{d}\phi}{\mathrm{d}t}$.
To summarize: the instantaneous frequency and instantaneous amplitude of a real signal are those of their analytic associate. This is summarized in Fig. 6.10.

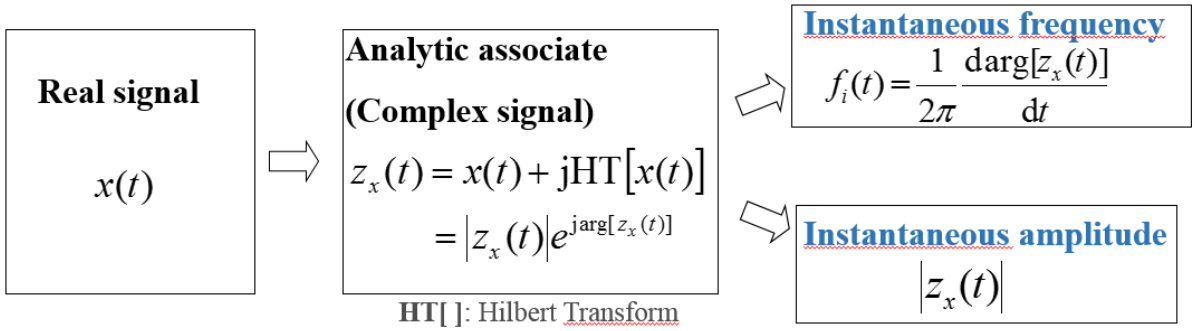


Figure 6.10: Determining the instantaneous frequency and amplitude of a real signal using the analytic associate.

Examples:
For the constant amplitude and constant frequency signal

$$x(t) = a\cos(2\pi f_0 t)$$

the analytic associate is

$$z_x(t) = e^{\mathrm{j}2\pi f_0 t}$$

and the instantaneous frequency is

$$f_i(t) = \frac{1}{2\pi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = f_0$$

Hence, the instantaneous frequency so calculated matches the usual definition of the frequency.
The next two examples are considered numerically, using Matlab:
The first Matlab example is that of a signal with constant frequency and a Gaussian amplitude:

$$x(t) = \cos(2\pi 0.5t)e^{-(t-40)^2/10}$$

The signal is shown in Fig. 6.11(a). The envelope $a(t)$ calculated as the modulus of the analytic associate of $x(t)$ is shown in the same figure. The instantaneous frequency is shown in Fig. 6.11(b). It is 0.5 as expected. It drops off to zero when the signal is actually null to machine accuracy due to the large decrease of its envelope. The Matlab script for plotting the figure is the following:

```
% Matlab code for producing Fig. 6.11
N=2000
t=linspace(0,80,N);
x1=cos(2*pi*0.5*t).*exp(-(t-40).*(t-40)/10);  % frequency 0.5Hz with gaussian envelope.
h1=hilbert(x1);                               % hilbert :  gives the analytic signal
                                              % x1+jH[x1].
a1=abs(h1);                                   % Instantaneous amplitude.
dt=t(2)-t(1);                                 % Time step (for calculating derivative).
fi1=[diff(unwrap(angle(h1))) 0]/2/pi/dt;      % Instantaneous frequency.
                                              % diff(A)/dt is the derivative of A,
```

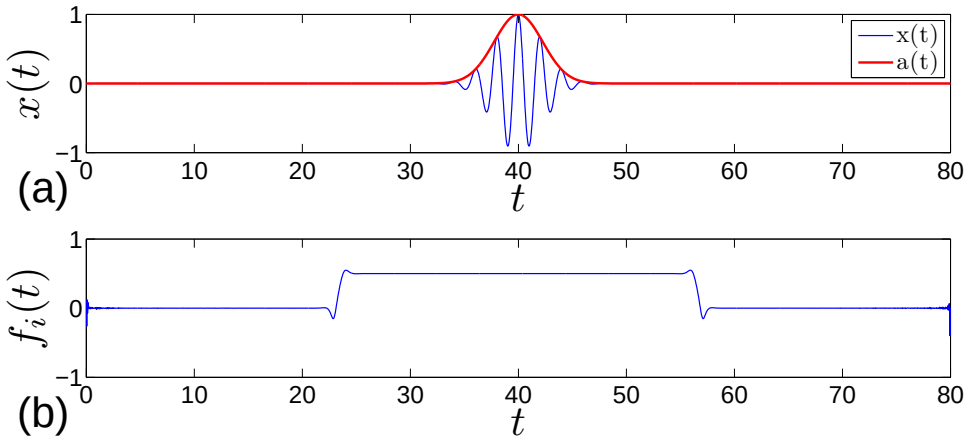


Figure 6.11: (a) Signal and its instantaneous amplitude; (b) instantaneous frequency.

```
% valid for a constant time sampling.
% diff removes one
% sample, so we add one zero at the tail.
% unwrap allows to unwrap the phase so
% that it does contain any jump as is the
% case when it is restricted to the
% [-π; π] interval.
```

Note: the Matlab function `hilbert(x)` does not return the Hilbert transform of x but directly its analytic associate!

The second Matlab example is a linear chirp:

$$x(t) = \cos(2\pi 0.05 t^2/2)$$

This signal is shown in Fig. 6.12(a) with its envelope. The envelope is almost a constant $a(t)$=1 except near the ends of the interval, which is an end effect to due the way the analytical associate is calculated. The instantaneous frequency is shown inf Fig. 6.12(b): as expected, it increases linearly with time, except again at the end of the interval.

6.2.4 Limitations of the instantaneous frequency: mono- and multi-component signal

Three examples "that work" have been given in the end of the previous subsection. Let us consider a simple signal to which it is more difficult to assign an instantaneous frequency:

$$x(t) = \cos(2\pi f_1 t) + \cos(2\pi f_2 t)$$

This is the sum of two waves at two different frequencies that are present at every time. One easily guesses that this reality cannot be described properly by a single instantaneous frequency. It is nevertheless possible to calculate one. Since the signal is real, one has to calculate its analytic associate to obtain the instantaneous frequency and amplitude. One easily gets: $f_i = (f_1 + f_2)/2$. The instantaneous frequency is the average of the two frequencies contained in the signal, and it

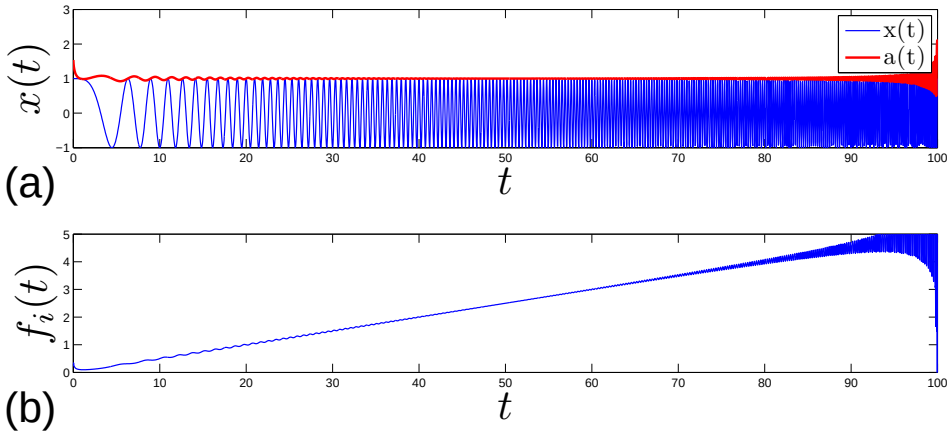


Figure 6.12: (a) Signal and its instantaneous amplitude; (b) instantaneous frequency.

is not a very interesting quantity in this case. The considered signal has several frequencies at a given time and is said to be **multi-components**. By contrast, the signals at the end of the last subsection were **mono-component**: a mono-component signal has only one single frequency at a given time. The difference is shown in Fig. 6.13.

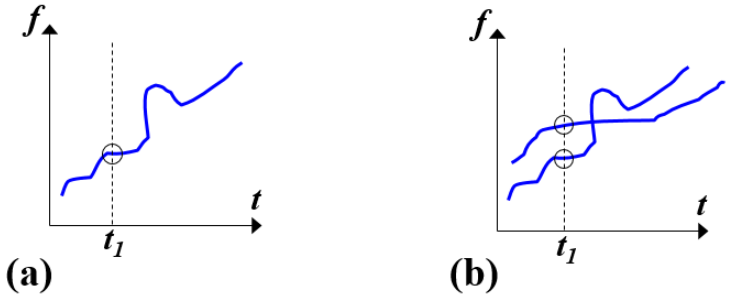


Figure 6.13: (a) A mono-component signal; (b) a multi-component signal.

The **instantaneous frequency** is thus a definition useful mainly for **mono-component** signals. For signals having a rich frequency content (that is, having more than one frequency at a given time), more comprehensive tools are needed: time-frequency energy distributions; atomic decompositions. These provide a full picture of the signal in the t-f plane. They are introduced in the following chapters.

6.2.5 Applications in Fluid mechanics

The Hilbert transform has been used in fluid mechanics by Sreenivasan [34], by Tardu [36], and by Mathis et al [23], amongst other. The latter paper is used as an example of how the method can be applied.

First, let us recall that in wall turbulence, the statistics close to the wall depend mainly on the inner scaling. The inner scaling consists in using u_τ as a velocity scale and z_τ as a length

scale, both of them being computed from the wall shear stress, τ_w, according to:

$$u_\tau = \sqrt{\frac{\tau_w}{\rho}}\,, \qquad z_\tau = \frac{\nu}{u_\tau}$$

In that scaling, the scaled turbulent axial velocity as a function of the scaled distance to the wall, that is,

$$\frac{\sqrt{\overline{u'^2}}}{u_\tau} = f\left(\frac{z}{z_\tau}\right) \quad ,$$

is almost universal and does not depend on the Reynolds number. Noteworthy, the outer scale (the boundary layer thickness δ for a boundary layer flow) is not present in the inner scaling. Hence, while the outer flow determines the friction at the wall, the universal scaling above tends to show that the outer flow does not affect the near wall dynamics. However, at high Reynolds numbers (Retau>2000 at least) this universal scaling does not hold anymore, meaning that the large structures of the outer layer does affect the inner layer statistics [21]. The paper of Mathis et al [23] entitled "Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers" examines the effect of large outer scales on the small inner scales.

Figure 6.14 shows the pre-multiplied energy spectrum of the streamwise velocity fluctuation as a function of the distance to the wall (horizontal axis), and the wavelength (vertical axis), for a boundary layer at Re_τ=7300. Both the normalizations in wall units ($z^+=z/z_\tau$, $\lambda_x^+=\lambda_x/z_\tau$)

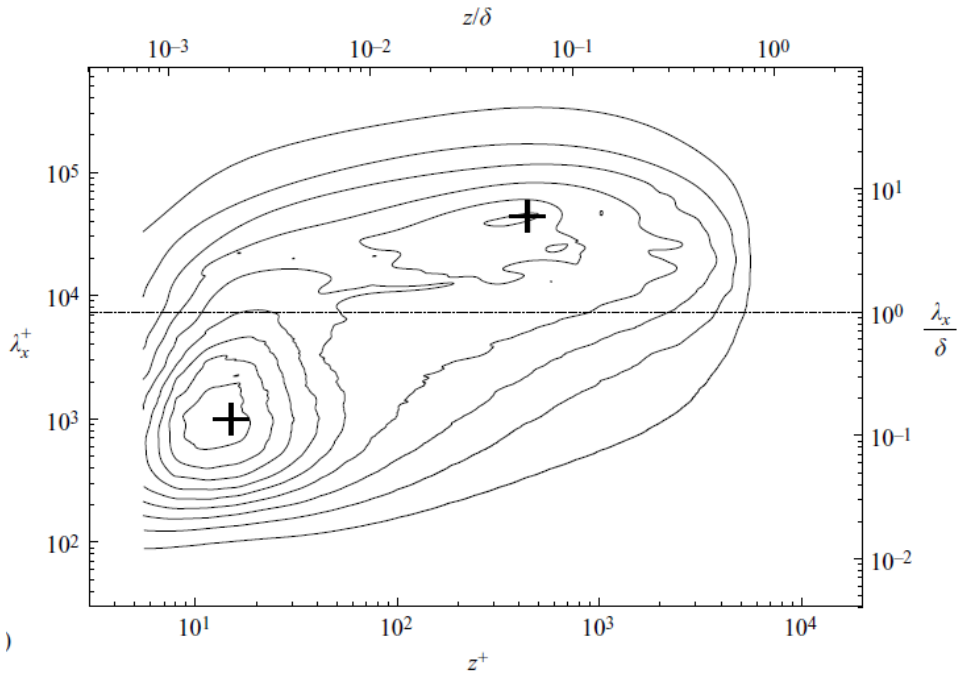


Figure 6.14: From Mathis et (2009) [23]. Contours of the pre-multiplied energy spectra, $k_x \cdot S_{uu}(k_x)$, of the streamwise velocity fluctuation.

and using the boundary layer thickness (z/δ, λ_x/δ) are indicated. The important thing here is that there are two maxima in the spectrum indicated with crosses. One peak scales in wall

units ($z^+ \sim$15, $\lambda_x^+ \sim 1000$) and corresponds to the small scales in the inner layer (the streaks). The other peak scales in outer units ($z/\delta \sim$0.06, $\lambda_x/\delta \sim$6) and corresponds to the large scales in the outer layer (in the log region). It should be pointed out that Fig. 6.14 is obtained using harmonic (Fourier) analysis, that is, using the techniques introduced in the first part of this course. With these techniques, the small scales (corresponding to, say, $\cos[2\pi x/l]$) and the large scales ($\cos[2\pi x/L]$ with $L >> l$) are not correlated (multiplying $\cos[2\pi x/l]$ and $\cos[2\pi x/L]$ and integrating will return zero) and it is difficult to assess the effect of the large scales on the small scales from a sole harmonic analysis. This is the reason why Mathis et al [23] first extract the enveloppe of the small scales using the Hilbert Transform. They then find that this enveloppe correlates well with the large scales (which is possible because the enveloppe itself contains the large scale, $\cos[2\pi x/L]$), which proves that the amplitude of the small scales is modulated by the large scales. Hence, using the Hilbert transform in this context allows the analysis of some kind of nonlinear interaction (that is, an interaction between different scales), which is not possible with the Fourier transform. Below are some more details about their analysis. First, the authors need to separate the large scales from the small scales. For this they use high-pass and low-pass filters. The common cut-off frequency (that is, spatial frequency) of these filters is indicated by the horizontal dashed line in Fig. 6.14 that goes in between the two crosses. A diagram of the subsequent method is shown in Fig. 6.15. In short, the small scales (the signal out of the high-

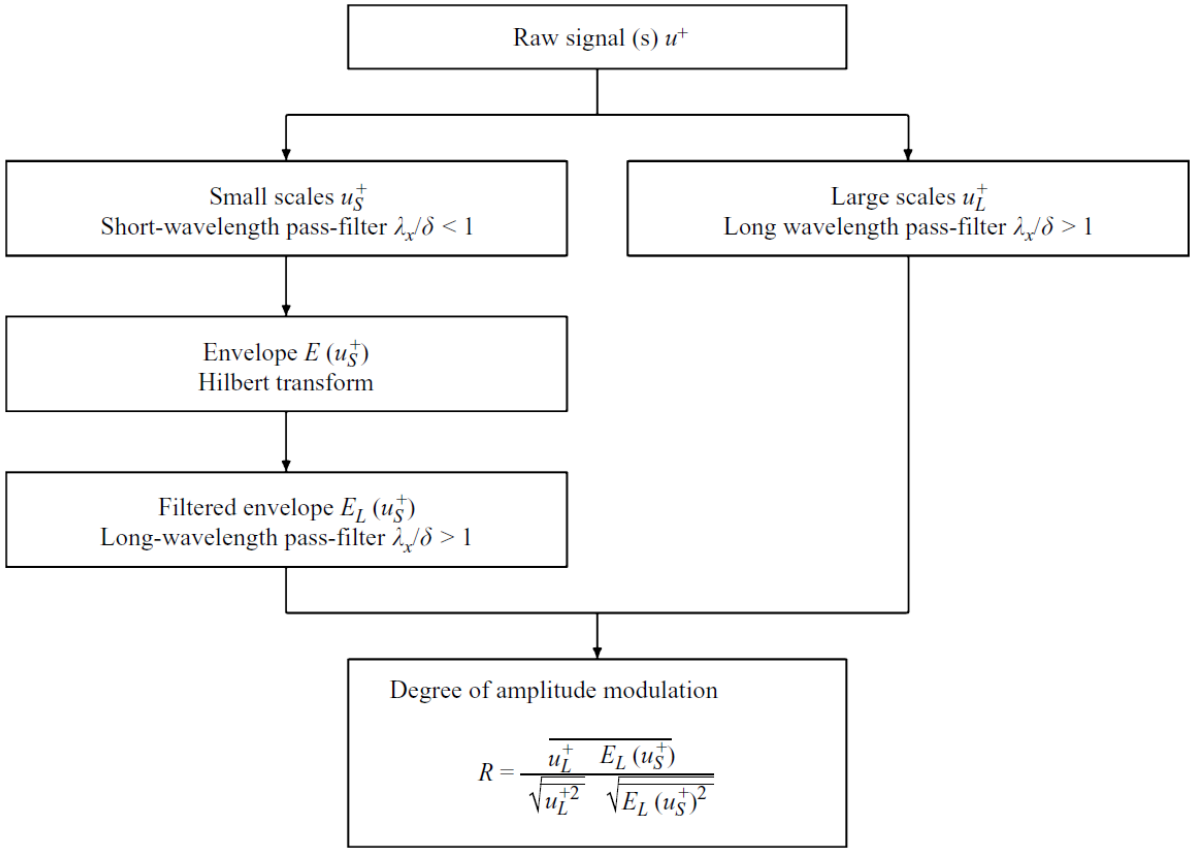
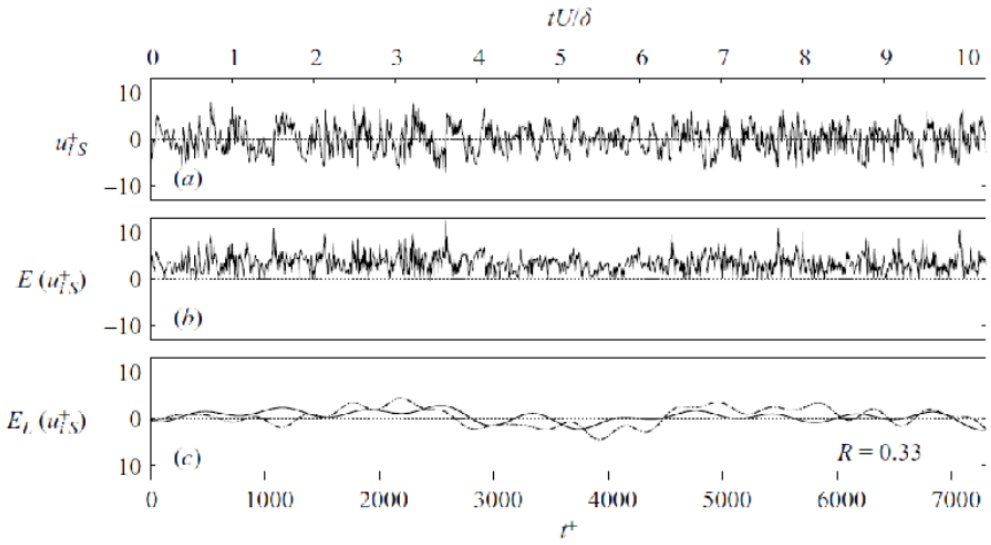


Figure 6.15: From Mathis et (2009) [23]. Diagram of the method to study the modulation of the small scale amplitude by the large scales.

pass filter, on the left) is passed to the Hilbert transform that delivers the enveloppe. Actually, the enveloppe is contaminated by small scales and needs to be cleaned up using a lowpass filter. This cleaned up enveloppe is then correlated with the large scales that are produced by the lowpass filter (on the right), which provides the correlation coefficient R. The latter needs to be large if one wants to prove that the small scales are modulated by the large scales, and this turns out to be the case. Examples of rough signals, enveloppe, and cleaned up enveloppe are given in Fig. 6.16.



Example of small-scale decomposition on fluctuating u velocity signal at $z^+ = 15$: (a) the small-scale signal u_{iS}^+; (b) its envelope; (c) its filtered envelope (solid line) and the large-scale component at $z/\delta = 0.06$ (dot-dashed line). For comparison, the mean of the filtered envelope has been adjusted to zero. Dashed vertical lines show region of negative large-scale u_{oL}^+ fluctuation.

Figure 6.16: From Mathis et (2009) [23]. Example of a processed signal.

Finally, note that Schlatter and Orlu [29] have emitted some doubts about the validity of the present method. Note also that Baars et al [2] have followed the same type of procedure as outlined above, but based on the wavelet transform rather than on the Hilbert transform. The wavelet transform (which will be studied in a subsequent chapter) can also be used to correlate different scales within a signal.

7 Time-Frequency distributions

A signal $x(t)$ has a density of energy per unit time given by $|x(t)|^2$. This density depends on time only. In the frequency domain, the signal has a density of energy per unit frequency given by $S_{xx}(f)$. This density depends on frequency only. In this chapter, time-frequency distributions $D(t, f)$ are presented that depend both on time and frequency. They allow visualizing which frequencies are important at a given time, or at which times a given frequency appears. Their relation to the two densities $|x(t)|^2$ and $S_{xx}(f)$ above is given.

The most famous time-frequency distribution, namely the Wigner-Ville distribution, is introduced. It can be used to construct a whole class of distributions called the Cohen class. Within this class, the Choi-Williams distribution is presented, in which the unwanted cross-terms spoiling the Wigner-Ville distribution are reduced.

The distributions presented in this course are for continuous time and continuous frequency. When processing signals with a computer, discrete-time discrete-frequency distributions are needed. The plots in this chapter are made using the routines in the Matlab package TFTB (for Time Frequency Tool Box) available on the net at: http://tftb.nongnu.org/

Time-frequency distributions are presented in Cohen [13, 14], in Boashash [6], or in Flandrin [16].

7.1 Separate time energy distribution and frequency energy distribution

7.1.1 Energy distribution in the time domain

Assume the signal $y(t)$ has a finite energy given by:

$$E_y = \int_{-\infty}^{\infty} |y(t)|^2 \mathrm{d}t \tag{7.1}$$

Now consider:

$$\rho_y(t) = \frac{|y(t)|^2}{E_y} \tag{7.2}$$

This is positive and satisfies:

$$\int_{-\infty}^{\infty} \rho_y(t)\mathrm{d}t = 1 \tag{7.3}$$

That is, $\rho_y(t)$ satisfies the admissibility condition Eq. (C.1), which is necessary for a function to be a probability density function (also called a distribution function). A time is considered more important (more likely) when it corresponds to a high energy density. A time that corresponds to no energy is unimportant. The quantity $\rho_y(t)\mathrm{d}t$ is the fraction of the signal energy comprised between time t and $t + \mathrm{d}t\mathrm{d}$.

To localize a signal in the time domain, the expected value is calculated by:

$$E[t] = \int_{-\infty}^{\infty} t\rho_y(t)\mathrm{d}t = t_m$$

This is the first order moment of the random variable that has ρ_y for its probability density function, and corresponds exactly to the mean time already defined in Eq. (3.16).

The spread of the signal around this mean time is obtained by calculating the second order moment:

$$E[(t - t_m)^2] = \int_{-\infty}^{\infty} (t - t_m)^2\rho_y(t)\mathrm{d}t = T_e^2 \tag{7.4}$$

which is the squared effective duration given in Eq. (3.18).

7.1.2 Energy distribution in the frequency domain

In the frequency domain, the important quantity for a signal with finite energy is the energy spectral density $S_{yy}(f)$. It is possible to normalize the density to obtain:

$$\sigma_y(f) = \frac{S_{yy}(f)}{E_y} \tag{7.5}$$

so that:

$$\int_{-\infty}^{\infty} \sigma_y(f)\mathrm{d}f = 1 \tag{7.6}$$

Again, this normalized energy distribution $\sigma_y(f)$ may be interpreted as a probability density function. A frequency f is therefore considered to be more important (more likely) when it corresponds to a high energy spectral density.

By calculating the first order moment by

$$E[f] = \int_{-\infty}^{\infty} f\sigma_y(f)\mathrm{d}f = f_m$$

and the second order moment by

$$E[(f - f_m)^2] = \int_{-\infty}^{\infty} (f - f_m)^2\sigma_y(f)\mathrm{d}f = B_e^2 \tag{7.7}$$

one obtains the mean frequency and squared effective bandwidth given previously in Eqs. (3.19-3.20).

7.2 Joint time-frequency energy distribution

The two energy distributions given in the previous section provide a way to characterize the signal in the time and frequency domains separately. However, these distributions do not allow a joint characterization of the signal in the time-frequency plane. For example, from the separate distributions, one may know the mean time of the signal and its mean frequency. However, one would not be able to know the mean frequency given that time is $t{=}t_0$ (we foresee a connection between the mean frequency at a given time and the instantaneous frequency defined in chapter 6). To answer this and similar questions, a **joint energy distribution** $D(t, f)$ is needed. This is equivalent to a joint probability density function, as defined in Appendix C.2. The goal is to find a distribution of energy $D(t, f)$ that gives the energy at a given time t and a given frequency f, per unit time and per unit frequency. There are several requirements, or at least desired properties, for this distribution. First, one would like to have:

$$D(t, f) \in R \tag{7.8}$$

and

$$D(t, f) \geq 0 \tag{7.9}$$

These two relations are desirable because the energy is usually a real positive quantity. One also would like to recover the signal energy by summing over all possible times and frequencies:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(t, f) \mathrm{d}t \mathrm{d}f = E \tag{7.10}$$

where E is the energy of the considered signal. Note that, contrary to the quantities $\rho_y(t)$ and $\sigma_y(f)$ in the previous section that are normalized so that their sum is 1 (see Eqs (7.3) and (7.6)), here $D(t, f)$ is not normalized.

Another requirement is that this distribution has for its **marginals** the energy time density and the energy spectral density (marginals of joint distributions are defined in Appendix C.2). This means that:

$$\int_{-\infty}^{\infty} D(t, f) \mathrm{d}f = |y(t)|^2 \tag{7.11}$$

$$\int_{-\infty}^{\infty} D(t, f) \mathrm{d}t = |Y(f)|^2 = S_{yy}(f) \tag{7.12}$$

Up to a normalization factor, the marginals of $D(t, f)$ are the separate energy distributions defined in the previous section. The effective duration and bandwidth depend on these marginals (see Eqs. (7.4) and (7.7)), and so does the uncertainty principle (Eq. (3.21)). Hence, if the marginals of $D(t, f)$ satisfy the former two relations, then $D(t, f)$ is consistent with the uncertainty principle.

Finally, the distribution has a strong finite support if:

$$D(t, f) = 0 \quad \text{whenever} \quad |y(t)|^2 = 0 \tag{7.13}$$

$$D(t, f) = 0 \quad \text{whenever} \quad |Y(f)|^2 = 0 \tag{7.14}$$

These conditions state that the density should be zero whenever the signal is zero or whenever its spectrum is zero.

Figures. 7.1 and 7.2 illustrate the global picture and the properties on the marginals. The

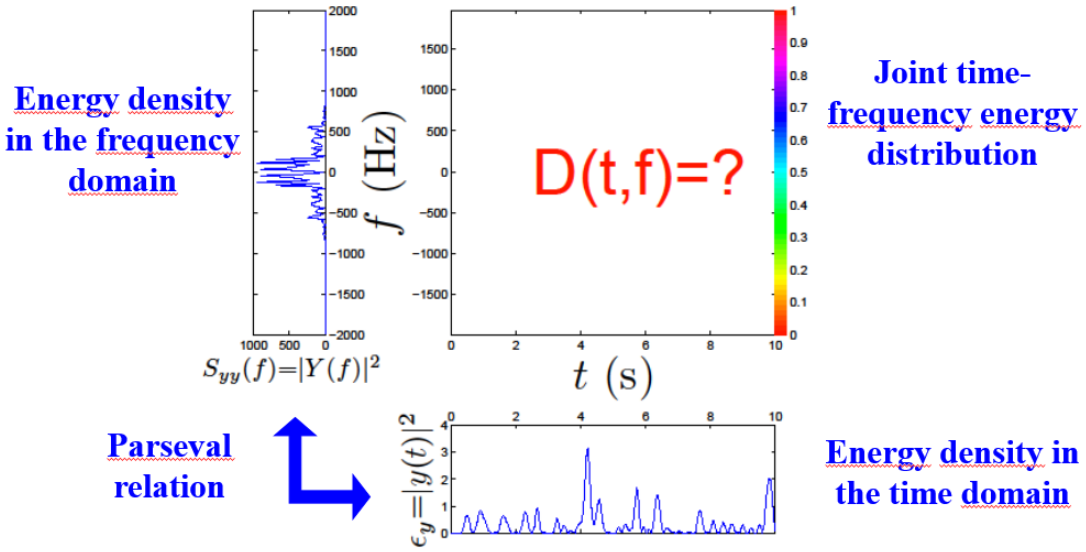


Figure 7.1: Joint time-frequency energy distribution $D(t, f)$ of some signal $y(t)$. Its marginals are the time energy distribution, $|y(t)|^2$, and the frequency energy distribution, $S_{yy}(f)$.

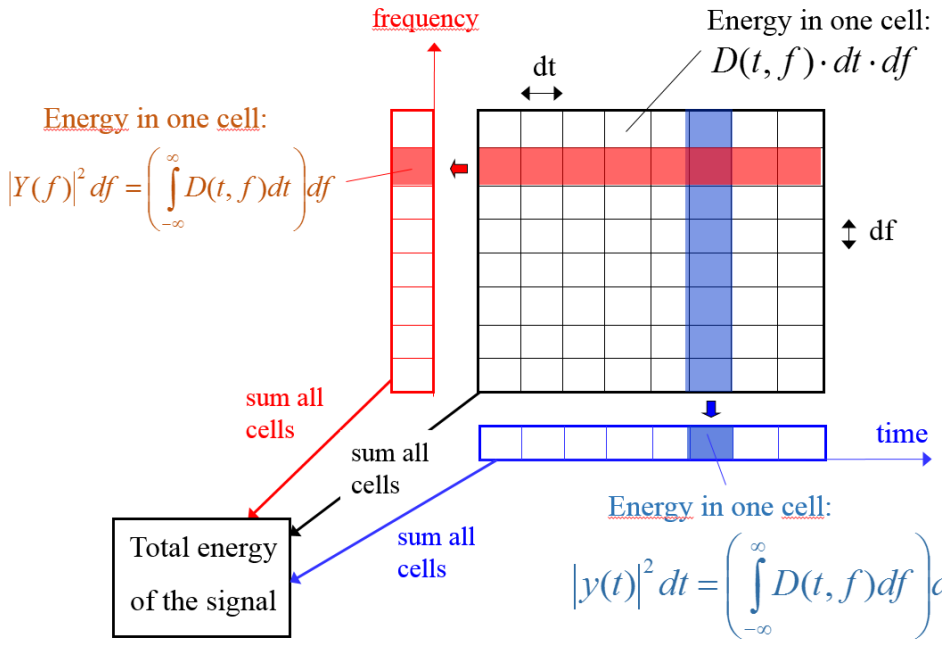


Figure 7.2: Summing over a column in $D(t, f)$ should give $|y(t)|^2$ and summing over a line should give $|Y(f)|^2$ when the marginal properties are satisfied.

question is now to find the distribution $D(t, f)$ of a known signal $y(t)$ that would satisfy some or all of the conditions given above. When this signal is known, so is its Fourier transform. Hence, the targeted marginals are known, and they are linked by the Fourier transform (and as a result by the uncertainty principle). The question is then loosely to devise a two-dimensional joint distribution from its marginals. In general, there are several possibilities, as shown in Fig. 7.3 (see also Fig. C.4 where two different distributions have the same marginals). This is due to the fact that when looking for a distribution, one is actually increasing the dimension of the signal, from a 1D signal to a 2D distribution. Any additional condition that will be put on the distribution will reduce the number of possible candidates. The problem that needs to be solved is summarized by Cohen, a citation of which is given in Fig. 7.4. In the following, several well-known distribution are studied.

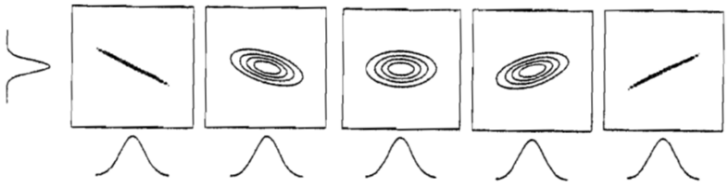


Figure 7.3: Question: how to find one or several joint time-frequency energy distributions $D(t, f)$ when the marginals are known (as is the case for a given signal)? From Cohen [14].

The basic objective of time-frequency analysis is to devise a function that will describe the energy density of a signal simultaneously in time and frequency, and that can be used and manipulated in the same manner as any density. If we had such a distribution we could ask for the fraction of the energy in a certain frequency and time range, we could calculate the density of frequency at a particular time, we could calculate the global and local moments of the distribution such as the mean conditional frequency and its local spread, and so on.

Do there exist joint time-frequency distributions that satisfy our intuitive ideas of a time-varying spectrum? How can they be constructed? Can they be interpreted as true densities? Do they represent the correlations between time and frequency? What reasonable conditions can be imposed to obtain such densities? The hope is that they do exist, but if they don't in the full sense of true densities, what is the best we can do? Are there inherent limitations to such a development? This is the scope of time-frequency analysis.

Figure 7.4: Citation from Cohen [14].

7.3 The Rihaczek distribution

In this section the Rihaczek distribution is introduced. It is not much used but is is presented because it is an example of a distribution that can be derived in a straightforward manner. Let us start from the signal energy:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty} x(t)x^*(t)\mathrm{d}t$$

The conjugated signal is replaced by its Fourier expansion:

$$E_x = \int_{-\infty}^{\infty} x(t) \int_{-\infty}^{\infty} X^*(f) e^{-\mathrm{j}2\pi ft} \mathrm{d}f \mathrm{d}t$$

This may me written:

$$E_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{R}(t,f) \mathrm{d}f \mathrm{d}t \tag{7.15}$$

where:

$$\mathcal{R}(t,f) = x(t) X^*(f) e^{-\mathrm{j}2\pi ft} \tag{7.16}$$

is the Rihaczek distribution. Equation (7.15) expresses the energy as the integral over time and frequency of $\mathcal{R}(t,f)$. However, nothing guarantees that this quantity has all the properties we would like. In particular, this is a complex number and we would prefer a real positive number that would yield itself better to a physical interpretation. However, one can check that the conditions on the marginals are verified, as well as the time and frequency support properties.

7.4 The Wigner-Ville distribution

7.4.1 Definition

We now present the famous Wigner-Ville distribution denoted by $W(t,f)$, and we give an idea of how it can be derived.

Like any other distribution it depends on time and frequency. The frequency dependence can be obtained through a Fourier transform:

$$W(t,f) = \underset{\tau \to f}{\mathrm{FT}} \left[K(t,\tau) \right] \tag{7.17}$$

Remember that the PSD, which is an energy distribution depending on f only, is the Fourier transform of the autocorrelation function (the latter depends on the lag τ only). By analogy, we expect K to be some sort of autocorrelation function that we are going to calculate.

To make some progress, consider a monocomponent signal $z(t) = e^{\mathrm{j}\phi(t)}$. For this signal we would like the Wigner-Ville distribution to have a single component at the instantaneous frequency, that is, we would like to have:

$$W(t,f) = \delta(f - f_i(t)) \quad \text{where} \quad f_i(t) = \frac{1}{2\pi}\phi'(t) \tag{7.18}$$

where $\phi'(t) = \mathrm{d}\phi/\mathrm{d}t$. In that case the autocorrelation function is:

$$K(t,\tau) = \underset{f \to \tau}{\mathrm{FT}}^{-1} \left[W(t,f) \right] = \underset{f \to \tau}{\mathrm{FT}}^{-1} \left[\delta\left(f - \frac{1}{2\pi}\phi'(t) \right) \right] = e^{\mathrm{j}\phi'(t)\tau}$$

The next step is to write an approximation for $\phi'(t)$:

$$\phi'(t) \sim \frac{1}{\tau}(\phi(t + \tau/2) - \phi(t - \tau/2)) \tag{7.19}$$

This approximation is actually exact if $\phi(t)$ is linear or quadratic in t, that is, if the instantaneous frequency of the signal is constant or linear in time. Substituting for this result in the previous equation gives:

$$K(t,\tau) = e^{\mathrm{j}(\phi(t+\tau/2)-\phi(t-\tau/2))} = e^{\mathrm{j}\phi(t+\tau/2)}e^{-\mathrm{j}\phi(t-\tau/2)} = z(t+\tau/2)z^*(t-\tau/2) \tag{7.20}$$

This result is derived in an approximate manner for monocomponent signals. Nevertheless, it is extended to more general signal in the following way. Let's $x(t)$ be the signal to study and let $z_x(t)$ be its **analytic associate** (that is, if $x(t)$ is real, then $z_x(t)$ is its analytic associate, and if $x(t)$ is complex, then $z_x(t)$ is the same as $x(t)$). The **instantaneous autocorrelation function** is defined by:

$$K(t,\tau) = z_x(t+\tau/2)z_x^*(t-\tau/2) \tag{7.21}$$

where t is the analysis time and τ is the lag (this is a time as well).
The Wigner-Ville distribution for the signal x is the Fourier transform of the instantaneous autocorrelation function, whereby τ is transformed into f, as expressed in Eq. (7.17). Namely, this is:

$$W(t,f) = \int_{-\infty}^{\infty} z_x(t+\tau/2)z_x^*(t-\tau/2)e^{-\mathrm{j}2\pi f\tau}\mathrm{d}\tau \qquad \text{(Wigner-Ville distribution)} \tag{7.22}$$

7.4.2 Properties

First, the Wigner-Ville distribution is real. Unfortunately, it has to be negative somewhere in the (t,f) plane, making difficult its interpretation as an energy density.

The WV distribution gives the right marginals:

$$\int_{-\infty}^{\infty} W(t,f)\mathrm{d}f = |x(t)|^2$$

and

$$\int_{-\infty}^{\infty} W(t,f)\mathrm{d}t = S_{xx}(f)$$

This in turn guarantees that:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(t,f)\mathrm{d}t\mathrm{d}f = E_x$$

For any signal $z(t) = a(t)e^{\mathrm{j}\phi(t)}$ (not mono-component), the instantaneous frequency is obtained from:

$$f_i(t) = \frac{1}{2\pi}\phi'(t) = \frac{\int_{-\infty}^{\infty} fW(t,f)\mathrm{d}f}{\int_{-\infty}^{\infty} W(t,f)\mathrm{d}f}$$

The latter quantity is $E[f]$, and the denominator is necessary because $W(t,f)$ is not normalized so that its sum is 1. What the equation tells us is that the WV distribution is concentrated in the neighborhood of the instantaneous frequency. By construction, for the special case of a mono-component signal it is a Dirac at the instantaneous frequency.

The time support of the WV distribution of a signal $z(t)$ is included in the time support of $z(t)$ and its frequency support is included in the frequency support of FT$[z(t)]$. In particular, we have the following properties:

$$z(t) = \delta(t - t_0) \Rightarrow W(t, f) = \delta(t - t_0) \quad \forall f \tag{7.23}$$

and

$$z(t) = e^{\mathrm{j}2\pi f_0 t} \Rightarrow W(t, f) = \delta(f - f_0) \quad \forall t \tag{7.24}$$

The latter relation is a consequence of Eq. (7.18), since for an harmonic signal the instantaneous frequency f_i equals f_0 at all times. The last two equations tell us that the time support of a Dirac impulse (a signal perfectly well localized in time) and the frequency support of a pure harmonic wave (a signal perfectly well localized in frequency) are not spread out by the Wigner-Ville distribution. This is a difference with the short time Fourier transform and the wavelet transform that will be introduced in a later part of the course.

Finally, given two signals $x(t)$ and $y(t)$ and their respective WV distributions $X(t, f)$ and $Y(t, f)$, the Moyal formula is:

$$| < x|y > |^2 = \left| \int_{-\infty}^{\infty} x(t) y^*(t) \mathrm{d}t \right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(t, f) Y(t, f) \mathrm{d}t \mathrm{d}f \tag{7.25}$$

7.4.3 Cross-terms (artifacts)

From its definition, Eq. (7.22), it is clear that the WV distribution is not linear in the signal: the WV distribution of the sum of two signals is not the sum of the WV distributions of the signals. Given two signals $x(t)$ and $y(t)$ (we assume they are complex so that they are already their own analytic associate), the WV distribution of their sum is:

$$W_{x+y}(t, f) = W_x(t, f) + W_y(t, f) + 2\mathrm{Re}[W_{xy}(t, f)] \tag{7.26}$$

where W_{xy} is the **cross Wigner-Ville distribution** of the signals $x(t)$ and $y(t)$ and is given by:

$$W_{xy}(t, f) = \int_{-\infty}^{\infty} x(t + \tau/2) y^*(t - \tau/2) e^{-\mathrm{j}2\pi f \tau} \mathrm{d}\tau$$

This cross WV distribution, that is simply referred to as **cross-terms** or as **artifacts**, is probably the most severe inconvenient of the WV distribution. Consider the simple two-component signal:

$$s(t) = x(t) + y(t) = A_1 e^{\mathrm{j}2\pi f_1 t} + A_2 e^{\mathrm{j}2\pi f_2 t}$$

Intuitively, one would like to see only two important frequencies, f_1 and f_2, in its time-frequency representation. However, from Eq. (7.26) its WV distribution is:

$$W(t, f) = A_1^2 \delta(f - f_1) + A_2^2 \delta(f - f_2) + 2A_1 A_2 \delta(f - (f_1 + f_2)/2) \cos(2\pi(f_2 - f_1)t)$$

Besides the two expected components at frequencies f_1 and f_2, the WV distribution has a component at $(f_1 + f_2)/2$ that corresponds to no physical reality. See also example 2 below.

7.4.4 Examples

The WV distributions in the examples below are computed numerically from a discrete-time signal. The digital Wigner-Ville transform is the one from the TFTB package.

Example 1:
The first signal is a pure tone with a Gaussian envelope:

$$x(t) = \cos(2\pi 50t)e^{-120(t-0.5)^2} \tag{7.27}$$

We have 512 samples of this signal over the time interval $[0, 1]$. The signal is shown in Fig. 7.5(a). This signal is real, and its analytic associate needs to be calculated prior to calculating the Wigner-Ville distribution. The WV distribution is shown in Fig. 7.5(b). There is a good

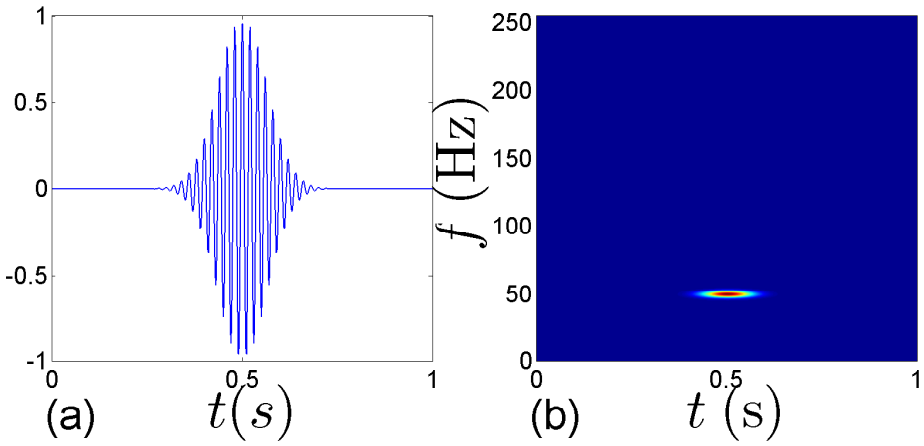


Figure 7.5: (a) Signal; (b) Its Wigner-Ville transform.

simultaneous localization of the signal in the time and frequency domains.

Exercise

How would you plot the energy spectral density $S_{xx}(f)$ of the signal using its WV distribution $W(t, f)$?

Example 2: Cross-terms
In the second example, the signal is made of the sum of two linear chirps:

$$x(t) = \underbrace{e^{\mathrm{j}2\pi(f_{1a}t+(f_{2a}-f_{1a})t^2/2)}}_{\text{component a}} + \underbrace{e^{\mathrm{j}2\pi(f_{1b}t+(f_{2b}-f_{1b})t^2/2)}}_{\text{component b}} \tag{7.28}$$

For the component 'a', the starting frequency is f_{1a}=10 Hz and the ending frequency is f_{2a}=100 Hz. For the component 'b', the starting frequency is f_{1b}=90 Hz and the ending frequency is f_{2b}=200 Hz. The signal is shown in Fig. 7.6(a). Its Wigner-Ville transform is shown in Fig. 7.6(b). The two lines corresponding to the instantaneous frequency $f_{i,a}(t) = f_{1a} + (f_{2a} - f_{1a})t$ for component 'a' and to the instantaneous frequency $f_{i,b}(t) = f_{1b} + (f_{2b} - f_{1b})t$ for component 'b' are observed. These are very fine and the frequency resolution is excellent (see later the corresponding result for other methods). Unfortunately, the result is somewhat spoiled by the cross terms that form a fuzzy area in between the two lines.

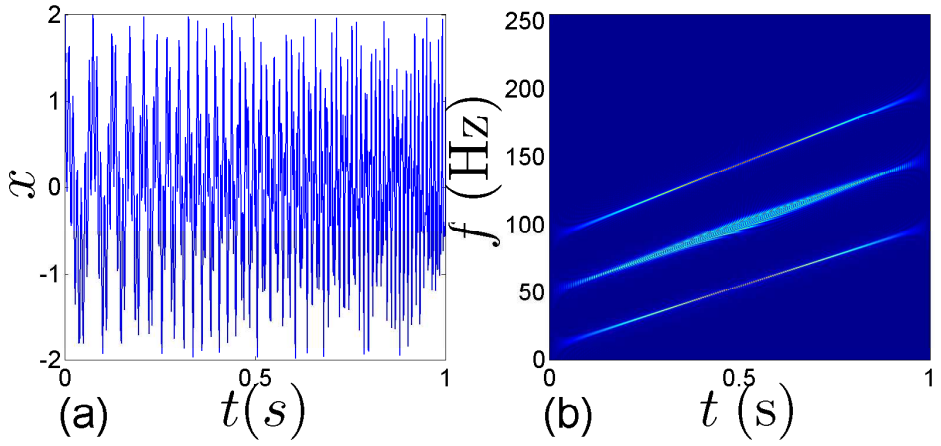


Figure 7.6: (a) Signal; (b) Its Wigner-Ville transform.

7.5 The Cohen class

7.5.1 Definition

The Cohen class is a whole class of energy distributions that can be obtained from the Wigner Ville distribution. The main objective is to get rid of the cross terms.

The method consists in modifying the instantaneous autocorrelation function in Eq. (7.21). Mathematically the modification is realized by convoluting with some **time-lag kernel** $G(t,\tau)$. Hence, the smoothed (blurred, smeared) instantaneous autocorrelation function is:

$$R(t,\tau) = G(t,\tau) \underset{t}{*} K(t,\tau) \qquad \left(\begin{array}{l} \text{Smoothed instantaneous} \\ \text{autocorrelation function} \end{array} \right) \tag{7.29}$$

Once this new autocorrelation function is known, a time-frequency distribution is obtained as before by taking its Fourier transform:

$$D(t,f) = \underset{\tau \to f}{\text{FT}}[R(t,\tau)] = \underset{\tau \to f}{\text{FT}} \left[G(t,\tau) \underset{t}{*} K(t,\tau) \right]$$

Now defining the **time-frequency kernel** by:

$$\gamma(t,f) = \underset{\tau \to f}{\text{FT}}[G(t,\tau)] \qquad \text{(time-frequency kernel)}$$

we have for the energy density:

$$\underbrace{D(t,f)}_{\substack{\text{time-frequency} \\ \text{distribution}}} = \underbrace{W(t,f)}_{\substack{\text{Wigner-Ville} \\ \text{distribution}}} \underbrace{\underset{t\,f}{**}}_{\substack{\text{smoothing} \\ \text{operator}}} \underbrace{\gamma(t,f)}_{\substack{\text{smoothing} \\ \text{function}}} \tag{7.30}$$

This can be spelled out as:

$$D(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t',f') \gamma(t-t', f-f') \mathrm{d}t' \mathrm{d}f' \tag{7.31}$$

Hence, any quadratic time-frequency energy distribution $D(t, f)$ can be obtained from the Wigner-Ville distribution $W(t, f)$ by convoluting the later with a kernel. This is called the kernel method. This is a design method whereby a kernel is designed so that the resulting distribution has nice properties (reduced cross-terms, positivity,...). Usually these properties are conflicting and there is a trade-off to be made.

7.5.2 The Choi-Williams distribution

The Choi-Williams distribution has the following Kernel:

$$G(t, \tau) = \frac{\sqrt{\pi\sigma}}{|\tau|} e^{-\pi^2\sigma t^2/\tau^2}$$

This is a typical example of a **Reduced Interference Distribution** (RID) that aims at reducing the cross terms. This may be the first method to try when computing a time-frequency distribution.

Let's consider again the two-component signal in Eq. (7.28) that was tackled by the WV distribution. The signal and its Choi-Williams distribution are shown in Fig. 7.7. Compared with

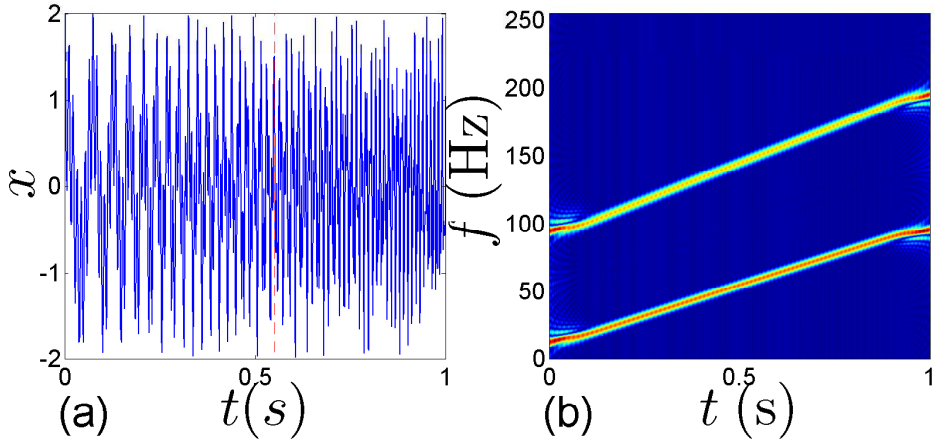


Figure 7.7: (a) Signal; (b) its Choi-Williams distribution.

the Wigner-Ville distribution given in Fig. 7.6, the cross-terms are reduced in the Choi-Williams distribution. The counterpart is that the resolution is reduced, since the lines corresponding to the "physical" components are thicker in Fig. 7.7 than they were in Fig. 7.6. The balance between these two conflicting properties is set by modifying σ.

Another classical signal is the sum of two hyperbolic chirps:

$$x(t) = \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + \cos\left(\frac{\alpha_2}{\beta_2 - t}\right) \tag{7.32}$$

where β_1=0.68 and β_2=0.72. The signal and its Choi-Williams distribution are shown in Fig. 7.8. Note that the instantaneous frequencies of the components are increasing fastly and the time

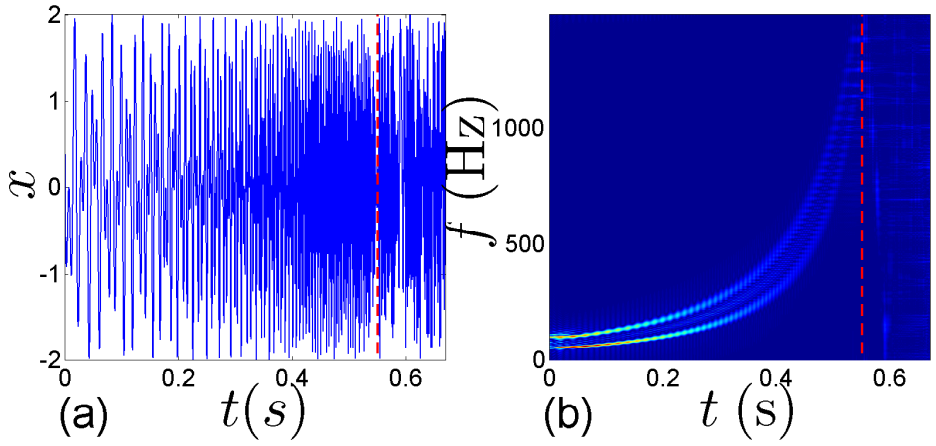


Figure 7.8: (a) Signal (sum of two hyperbolic chirps); (b) its Choi-Williams distribution.

sampling cannot cope with these high frequencies. There is some aliasing present past the red dashed line. This is clearly seen from Fig. 7.8(b) where, past the time indicated by the red dashed line, the lines seem to bounce off the upper frequency limit of the plot. However, until there is aliasing, the Choi-Williams distribution detect the two frequencies. This example will be used again for the short time Fourier transform and for the wavelet transform.

The last example is the signal:

$$x(t) = \cos(2\pi 10t) + \delta(t - 0.5) \tag{7.33}$$

which is the sum of a pure tone and a Dirac impulse. Recall (see Section 7.4.2) that each of the components in this signal is correctly accounted for by using the Wigner-Ville distribution. Unfortunately, the WV distribution of the *sum* is spoiled by cross terms due to the nonlinear interaction between the two components. This is fixed by using the Choi-Williams distribution. The signal and its Choi-Williams distribution are shown in Fig. 7.9. In that case, the signal contains 200 points over the time interval $[0, 1]$. The vertical line in the distribution accounts for the impulse and the horizontal line accounts for the pure tone.

7.6 Conclusion

Quadratic time-frequency distributions were introduced by Ville in 1948. They consider the time-frequency distribution of the energy of the signal. The Wigner-Ville distribution is a bilinear transform from the start, since the product $z_x z_x^*$ is involved in the transform. Bilinearity is responsible for cross-terms that unfortunately spoil the otherwise excellent resolution of the Wigner-Ville distribution. Interferences due to cross terms may be reduced by using a smoothed Wigner-Ville distribution, at the cost of a lower resolution, as in the Choi-Williams distribution for example. Transformations that are linear instead of nonlinear (short time Fourier transform, wavelet transform) and from which an energy distribution can be derived are presented in subsequent chapters.

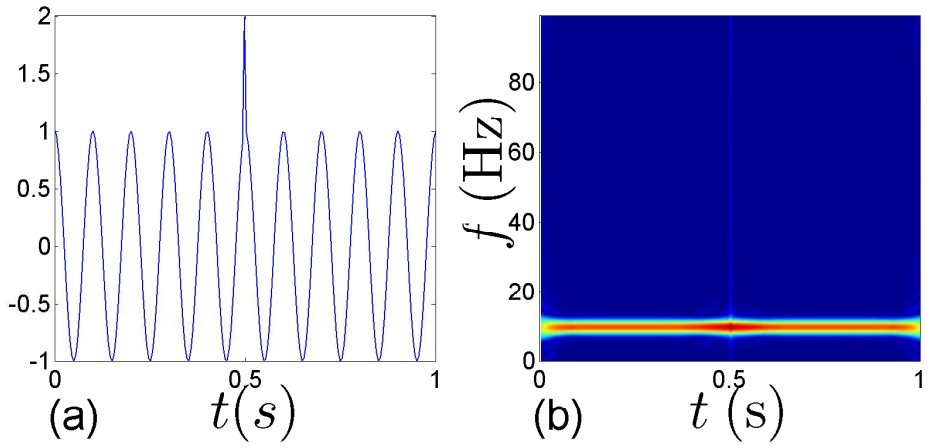


Figure 7.9: (a) Signal (a pure tone + an impulse); (b) its Choi-Williams distribution.

8 Short-Time Fourier Transform

To calculate a time-frequency energy distribution, a slight variation of the Fourier transform may be used: the short-time Fourier transform (STFT). It consists of isolating a small piece of the signal by using a window, and then calculating the Fourier transform of the windowed signal. Squaring the STFT then provides a time-frequency distribution known as the spectrogram. This method has some time and frequency resolution limits that are discussed.

8.1 The STFT

The principle of the short-time Fourier transform is pretty simple. One problem of the Fourier transform it that it uses *all* the signal history. To account for only *a part of* this history, on may multiply the signal by a window $w(t)$ such as that shown in Fig. 8.1. To be more specific,

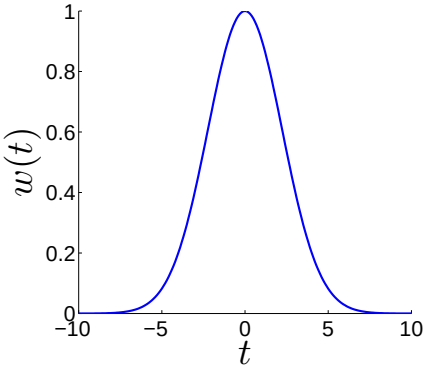


Figure 8.1: Example of an analyzing window for the short-time Fourier transform.

given a signal $x(t)$ defined over the interval $[-\infty\ \infty]$, a windowed signal is defined by:

$$x_u(t) = x(t)w(t-u) \tag{8.1}$$

where the window is translated so that its average position is $t = u$. The signals $x(t)$, $w(t-u)$, and $x_u(t)$ are shown in Fig. 8.2. The windowed signal contains information about the original signal mostly over a time interval around u. Note that the windowed signal $x_u(t)$ depends on two different times:

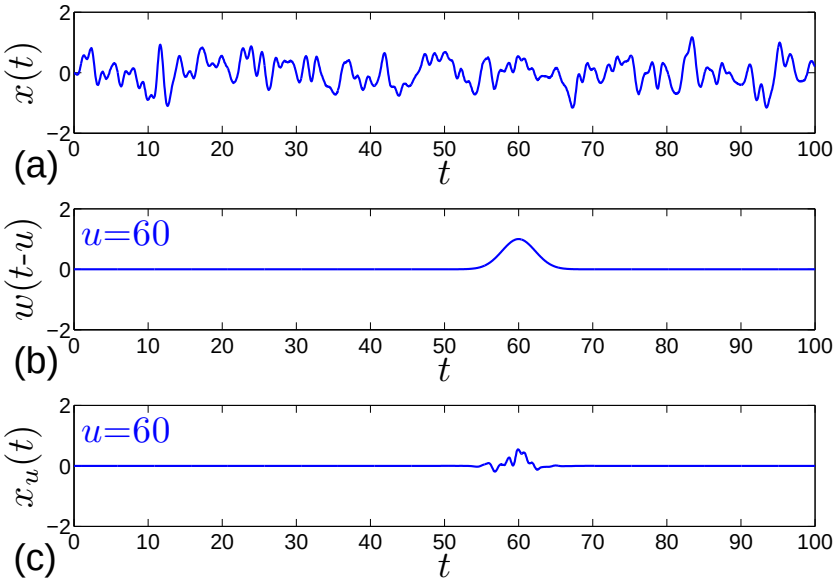


Figure 8.2: (a) Signal $x(t)$; (b) translated window $w(t-u)$ for u=60; (c) $x_u(t) = x(t)w(t-u)$ for u=60.

- the running time, t;
- the time as which the signal is analyzed, u.

The Fourier transform of the windowed signal can then be calculated. This is the **short-time Fourier transform (STFT)**:

$$\text{STFT}[x(t)] = \underset{t\to f}{\text{FT}}[x_u(t)] = X(u,f) = \int_{-\infty}^{\infty} x(t)w(t-u)e^{-\text{j}2\pi ft}\text{d}t \tag{8.2}$$

The STFT $X(u,f)$ gives the amplitudes of waves with frequency f around time u. Some remarks need to be made:

• The STFT amounts to windowing followed by a Fourier transform.

• The frequency f is the dual of the running time t.

• The STFT $X(u,f)$ depends both on the signal *and* on the window. Thus, the STFT entangles the characterictics of the signal with that of the window (this is true as well for the wavelet transform, see Chapter 9).

The signal may be reconstructed from the knowledge of its STFT:

$$x(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X(u,f)w^*(t-u)e^{\text{j}2\pi ft}\text{d}u\text{d}f \tag{8.3}$$

This is the inverse of the STFT.

8.2 The STFT as an atomic decomposition

In Chapter 2 it has been seen that Fourier coefficients in Fourier series for functions in L^2 are obtained by projecting the function onto a basis (see Fig. 2.1). The Fourier series itself is nothing

but the expansion of the function in terms of the basis members. This is not exactly true for the Fourier transform, but everything looks "as if" it were the case. A similar pattern is now shown to be true for the STFT.

Let us define a family of atoms $w_{uf}(t)$ by:

$$w_{uf}(t) = w^*(t-u)e^{\mathrm{j}2\pi ft} \tag{8.4}$$

In addition to the window itself, these atoms are described by two parameters:
- u: the time at which the atom is placed (center of the window);
- f: the frequency.
Both of these parameters can take any real continuous value. Some atoms are shown in Fig. 8.3 for several values of u and f.

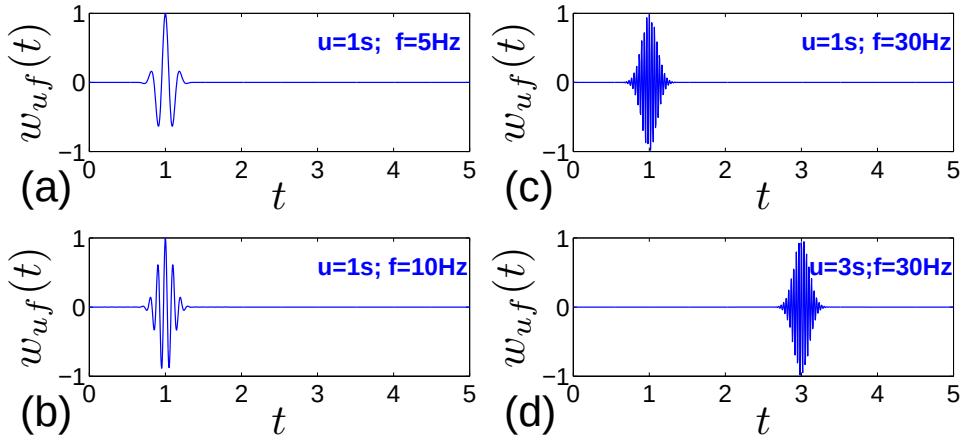


Figure 8.3: Some atoms of the family in Eq. (8.4) with: (a) u=1 s, f=5 Hz; (b) u=1 s, f=10 Hz; (c) u=1 s, f=30 Hz; (d) u=3 s, f=30 Hz.

This set of atoms forms a **basis** onto which the signal is decomposed. The STFT is the projection of the signal onto the atoms:

$$X(u,f) =< x|w_{uf} >= \int_{-\infty}^{\infty} x(t)w_{uf}^*(t)\mathrm{d}t$$

Remember that the scalar product allows comparing the two signals involved in the product: $X(u,f)$ is large if x ressembles w_{uf}, or said otherwise, if x contains w_{uf}. Projecting the signal onto w_{uf} is a way to know the amplitude of the component at frequency f present in the signal around time u. For example, the atoms in Fig. 8.3(a)-(c) allow to make an analysis at the same time, u=1 s, but at 3 different frequencies: 5 Hz, 10 Hz, and 30 Hz.

The inverse STFT is the fact that the signal may be expanded using the family of atoms:

$$x(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X(u,f)w_{uf}(t)\mathrm{d}u\mathrm{d}f = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \langle x|w_{uf}\rangle\, w_{uf}(t)\mathrm{d}u\mathrm{d}f$$

This may be called the **atomic decomposition** of $x(t)$.

8.3 The spectrogram

8.3.1 Definition

A Parseval like relation on the energy can be demonstrated for the STFT. To be more specific, the energy of the signal is given by:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |X(u,f)|^2 \mathrm{d}u\mathrm{d}f \tag{8.5}$$

where it has been assumed that:

$$\int_{-\infty}^{\infty} |w(t)|^2 \mathrm{d}t = 1$$

which means that the window has unit energy. Equation (8.5) prompts us to define a time-frequency energy distribution by:

$$D(u,f) = |X(u,f)|^2 = \left| \int_{-\infty}^{\infty} x(t)w(t-u)e^{-\mathrm{j}2\pi f t}\mathrm{d}t \right|^2 \qquad \text{(Spectrogram)} \tag{8.6}$$

This distribution is called the **spectrogram**. This method was first used in 1946 for analyzing sounds.

8.3.2 Properties

By definition, the energy of the signal is obtained by summing the spectrogram over time and frequency:

$$E_x = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D(u,f)\mathrm{d}u\mathrm{d}f \tag{8.7}$$

and $D(u,f)$ is a real positive quantity.

Marginals:
To see if the conditions on the marginals are satisfied, these are calculated. The marginal on time is obtained by integrating frequency out:

$$M_u(u) = \int_{-\infty}^{\infty} D(u,f)\mathrm{d}f = \int_{-\infty}^{\infty} |X(u,f)|^2 \mathrm{d}f$$

After using the definition of $X(u,f)$ (Eq. 8.2) and exchanging the order of the integrals, one obtains:

$$M_u(u) = \int_{-\infty}^{\infty} |x(t)|^2 |w(u-t)|^2 \mathrm{d}t$$

that is:

$$M_u(u) = (|x|^2 * |w|^2)(u) \neq |x(u)|^2$$

Hence, the marginal on time at time u is not equal to the time energy density, $|x(u)|^2$. It is equal to the convolution of the energy density of the signal with the energy density of the window. As a result, the energy of the signal at some time is spread across the neighbor times. The

condition on the time marginal is not satisfied.
For the marginal on frequency, we have:

$$M_f(f) = \int_{-\infty}^{\infty} D(u,f)\mathrm{d}u = \int_{-\infty}^{\infty} |X(u,f)|^2 \mathrm{d}u$$

On using again the definition of $X(u,f)$, one finds:

$$M_f(f) = \int_{-\infty}^{\infty} |X(f')|^2 |W(f-f')|^2 \mathrm{d}f' \neq |X(f)|^2$$

The condition on the frequency marginal is not satisfied either, the spectral energy of the signal being convoluted with that of the window.
None of the two conditions on the marginals are respected. This is due to the spectrogram scrambling the energy distribution of the signal with that of the window.

Relation with the Wigner-Ville transform:
The spectrogram has be obtained in a way very different from that leading to the Wigner-Ville distribution. However, when studying the Cohen class, it has been shown that any distribution can be obtained by smoothing the Wigner-Ville distribution using some suitable kernel. This is true for the spectrogram, as is now shown.

Writing the STFT as an atomic decomposition, we have:

$$X(u,f) = < x|w_{uf} >$$

Hence, the spectrogram may be written:

$$D(u,f) = |X(u,f)|^2 = | < x|w_{uf} > |^2$$

The Moyal formula, Eq. (7.25), is then used to give:

$$D(u,f) = | < x|w_{uf} > |^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(t',f')W_{w_{uf}}(t',f')\mathrm{d}t'\mathrm{d}f'$$

where W_x is the Wigner-Ville distribution of x and $W_{w_{uf}}$ is the Wigner-Ville transform of w_{uf}. On using $w_{uf}(t) = w^*(t-u)e^{\mathrm{j}2\pi ft}$ and the definition of the Wigner-Ville distribution (Eq. (7.22)), we obtain:

$$D(u,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(t',f')W_w(t'-u,f'-f)\mathrm{d}t'\mathrm{d}f'$$

where W_w is the Wigner-Ville transform of w. Finally, one assumes that the window w is a real even function, that is: $w(-t) = w(t)$. This yields:

$$D(u,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(t',f') \underbrace{W_w(u-t',f-f')}_{\gamma(u-t',f-f')} \mathrm{d}t'\mathrm{d}f'$$

Comparing this expression with Eq. (7.31) yields the following conclusion: for the spectrogram, the smoothing kernel is the Wigner-Ville distribution of the window used in the STFT. This means that, to obtain the spectrogram, one can blur the Wigner-Ville distribution in the $u-f$

plane with a blurring function that depends on the window. The result of this blurring operation is that the resolution of the spectrogram is not as good as that of the Wigner-Ville distribution. However, the level of the cross-terms is decreased in the spectrogram, which is an advantage over the Wigner-Ville distribution.

Spectrogram of a signal localized in time:

Consider the signal $x(t) = \delta(t - t_0)$. This signal contains energy only at time t=t_0. Does the spectrogram satisfy this? Using the properties of the Dirac impulse, the STFT is:

$$X(u, f) = w(t_0 - u)e^{-\mathrm{j}2\pi f t_0}$$

and the spectrogram is:

$$D(u, f) = |X(u, f)|^2 = |w(t_0 - u)|^2 \quad \forall f$$

The energy is then spread over a time interval $[t_0 - T_e/2 \quad t_0 + T_e/2]$ where T_e is the typical effective duration of the window. The resolution of the spectrogram in the time domain is thus about the window duration, which makes sense. This behaviour differs from that of the Wigner-Ville distribution (compare with Eq. (7.23)).

Spectrogram of a signal localized in frequency:

Consider a pure wave: $x(t) = e^{\mathrm{j}2\pi f_0 t}$. The spectrum of this signal should contain energy only at the frequence f_0. Let's check if the spectrogram satisfies this. The STFT is:

$$X(u, f) = \hat{w}(f - f_0)e^{\mathrm{j}2\pi(f - f_0)u}$$

where $\hat{w}(f) = \mathrm{FT}[w]$ is the FT of the window. The spectrogram is:

$$D(u, f) = |X(u, f)|^2 = |\hat{w}(f - f_0)|^2 \quad \forall u$$

The spectral energy density is then spread over a frequency interval $[f_0 - B_e/2 \quad f_0 + B_e/2]$ where B_e is the typical effective bandwidth of the window. The resolution of the spectrogram in the frequency domain is about the window bandwidth. This behaviour differs from that of the Wigner-Ville distribution (compare with Eq. (7.24)).

8.3.3 Trade-off between time and frequency resolutions

In the properties above, the window is seen to have an effect on the result returned by the spectrogram. Another (and related) important consequence of the necessary windowing in the STFT and the spectrogram is now adressed. Namely, the window imposes a trade-off between the time and frequency resolution of the STFT/spectrogram.

According to the uncertainty principle (Heisenberg-Gabor principle), the product of the effective bandwidth and the effective duration of any signal should satisfy:

$$B_e T_e \geq \frac{1}{4\pi}$$

This principle applies to the original signal, $x(t)$, and it also applies to the small piece of signal $x_u(t)$ considered. The window w has some duration, T_e, and one can infer that the effective

duration for $x_u(t)$ is also approximately T_e. Due to the uncertainty principle, the effective bandwidth B_e for $x_u(t)$ has to be large if T_e is small, that is, if the window duration is made small. As a consequence, if we want a good time resolution, we will have a poor frequency resolution, and conversely.

Remark: here, the uncertainty principle applies to the windowed signal, that is, to the signal modified by the window. This a limitation of the technique itself, not one of the original signal. It is indeed the window that fixes T_e and B_e.

The tiling of the time-frequency plane corresponding to the spectrogram is illustrated in Fig. 8.4. For the tiling on the left, the duration T_e of the window is small, meaning there is a good time resolution. The frequency resolution is such that $T_e B_e \geq 1/(4\pi)$. Suppose T_e is decreased to zero (the window is then almost a Dirac impulse): you can see your signal varying in time with a high resolution but you won't be able to know the frequency at any time due to a too large uncertainty. For the tiling on the right, T_e is larger but B_e is smaller, the product $B_e T_e$ being roughly kept constant. In the limit where T_e tends to infinity, the STFT tends to the FT. There is then no time resolution (as you know, in the FT time is integrated out), but a perfect precision on frequency. Note that in the spectrogram $D(u, f)$, both u and f are continuous.

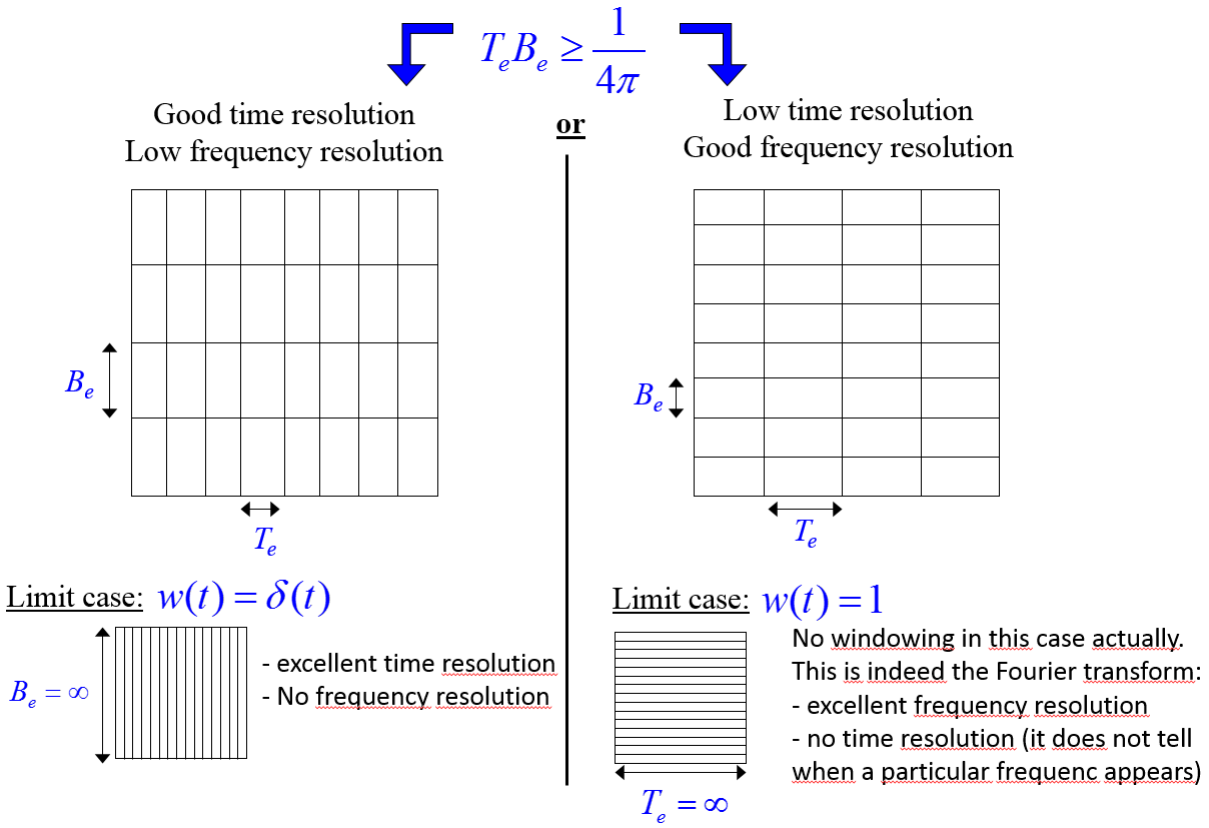


Figure 8.4: Tiling of the time-frequency plane in the spectrogram. The time resolution is T_e and the frequency resolution is B_e, and the product $T_e B_e$ is constant due to the uncertainty principle. If one is small, the other is large, and conversely.

Hence, the tiling in Fig. 8.4 is only a view of the mind. There is actually a small box of width T_e and height B_e around every single point (u, f) in the plane.

8.3.4 Examples

Some examples are finally given. They are the same as that considered with the Wigner-Ville and Choi-Williams distributions.

Example 1:
The first signal is a pure tone with a Gaussian envelope:

$$x(t) = \cos(2\pi 50t)e^{-120(t-0.5)^2} \tag{8.8}$$

This signal is the same as in Eq. (7.27) and is shown in Fig. 7.5(a). A discrete-time signal with N=512 samples is studied with Matlab. Its spectrogram is calculated with the Matlab function `specgram`. This functions uses a DFT for approximating the STFT. For the spectrogram, the duration of the window needs to be selected. For discrete time, this is equivalent to choosing a window size, n. The window is of the Hamming type. It is also possible to choose an overlap between the different positions taken by the window. Here, the overlap is $9n/10$. The spectrogram of the signal is given in Fig. 8.5 for two different values of n. For n=30 in Fig. 8.5(a), the

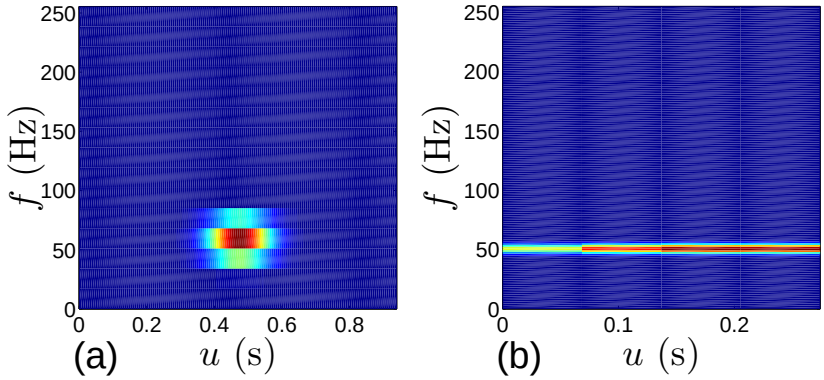


Figure 8.5: Spectrogram of the signal in Eq. (8.8). Two different size of the windows are used: (a) n=30; (b) n=341.

window has a small duration. There is a good time precision, but a poor frequency precision. From the spectrogram, it is indeed difficult to determine the frequency with a good precision. For n=341 in Fig. 8.5(b), the window has a larger duration and the frequency resolution is much better. The frequency is observed to be 50 Hz with a good precision, but the time resolution is poor and it is difficult to say *when* this frequency appears. Compare this result with that given in Fig. 7.5(b) for the Wigner-Ville transform. For this example, there is only one component and the Wigner-Ville transform has no cross-terms. Its resolution is excellent and it is much more precise than the spectrogram.

The Matlab script for plotting the figure is the following:

```
% Matlab code for producing Fig. 8.5(a)
```

```
N=512;                                          % Number of samples of the signal.
t=linspace(0,1,N);                              %
dt=diff(t);dt=dt(1);                            % Sampling period.
fs=1/dt;                                        % Sampling frequency.
x=cos(2*pi*50*t).*exp(-120*(t-0.5).*(t-0.5));   % Signal.
n=30;                                           % Size of the window.
w=hamming(n);                                   % Hamming window.
noverlap=floor(9*n/10);                         % number of samples between
                                                % two times in the spectrogram.
[D,F,T]=specgram(x,n,fs,w,noverlap);            % D:Spectrogram; T: time vector;
                                                % F: frequency vector.
pcolor(T,F,abs(D))                              %
colormap jet;                                   %
shading flat                                    %
```

Example 2:

We now consider the two-component signal made of the sum of two linear chirps:

$$x(t) = \underbrace{e^{\mathrm{j}2\pi(f_{1a}t+(f_{2a}-f_{1a})t^2/2)}}_{\text{component a}} + \underbrace{e^{\mathrm{j}2\pi(f_{1b}t+(f_{2b}-f_{1b})t^2/2)}}_{\text{component b}} \tag{8.9}$$

This is the same signal as in Eq. (7.28). Its Wigner-Ville distribution is given in Fig. 7.6(b) and its Choi-Williams distribution is given in Fig. 7.7(b). The discrete-time signal has N=512 samples and is shown in Fig. 8.6(a). For the spectrogram shown in Fig. 8.6(b), n=50. Compared

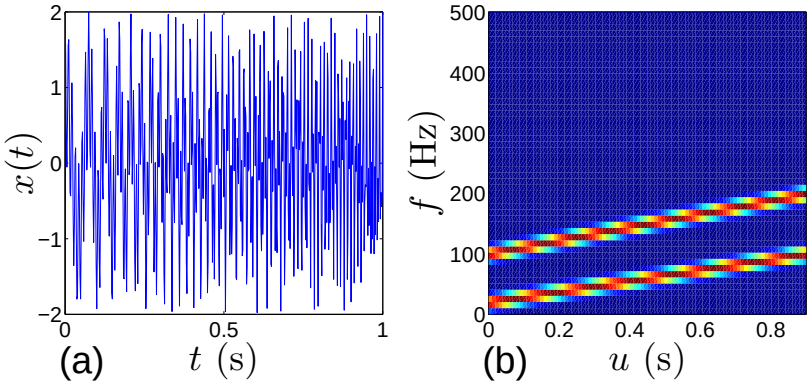


Figure 8.6: (a) Signal made of the sum of two linear chirps (see Eq. (8.9)) (b) its spectrogram with n=50.

with the Wigner-Ville and Choi-Williams distributions, the spectrogam has no cross-terms but has the poorest resolution of the three methods. The Wigner-Ville distribution has an excellent resolution but many cross-terms. It has no parameter. The Choi-Williams is a good compromise between resolution and the number of cross terms.

Example 3:

The signal is a sum of a pure tone and a Dirac impulse:

$$x(t) = \cos(2\pi 10t) + \delta(t - 0.5) \tag{8.10}$$

This is the same as in Eq. (7.33). It is shown in Fig. 7.9(a). There are N=200 samples over the time interval [0 1]. The spectrogram is given for two values n of the window size in Fig. 8.7. The overlap is $9n/10$. For n=2, there is an excellent time precicion. One sees the wave's oscillation.

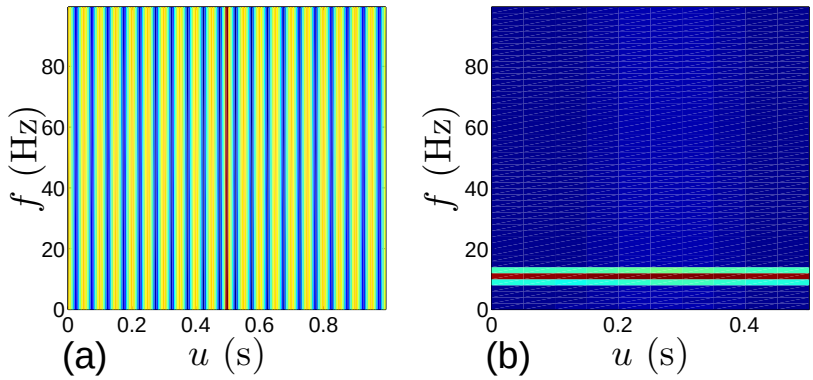


Figure 8.7: Spectrogram of the sum of a Dirac impulse and a pure tone (signal in Eq. (8.10)): (a) n=2; (b) n=100.

One also sees the Dirac impulse. The counterpart is that the frequency of the wave oscillation is not obtained. To obtain the value of this frequency the frequency resolution is increased by increasing the window length. For n=100, the value of the frequency, 10 Hz, may be obtained. However, the time resolution decreases and the Dirac is hardly observed, not to mention its exact appeareance time. The Choi-Williams distribution applied to the same signal is shown in Fig. 7.9(b) and is overall more precise, even if it has more cross terms.

Example 4:
The last example is the sum of hyperbolic chirps:

$$x(t) = \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + \cos\left(\frac{\alpha_2}{\beta_2 - t}\right) \tag{8.11}$$

where β_1=0.68 and β_2=0.72. This is the same signal as in Eq. (7.32) and is plotted in Fig. 7.8(a). There are N=2000 samples and the spectrogram is given in Fig. 8.8 for n=500. The overlap is again $9n/10$. The two frequencies may be separated at the earlier times, but past t=0.3 s they become mixed. The Choi-Williams distribution applied to the same signal is shown in Fig. 7.8(b) and is overall more precise, even if it has more cross terms.

8.4 Conclusion

The spectrogram is a time-frequency distribution obtained by squaring the short-time Fourier transform. The STFT may be seen as an atomic decomposition. Such a decomposition was first introduced by Gabor in 1946. Atomic decompositions, and STFT in particular, consist of taking a linear transform of a signal by projecting the signal onto some basis made of time-frequency atoms. Then, squaring the STFT provides the spectrogram. The method differs from that employed to obtain the Wigner-Ville transform, since the Wigner-Ville transform consists of calculating an instantaneous autocorrelation that is nonlinear in the signal and of taking the

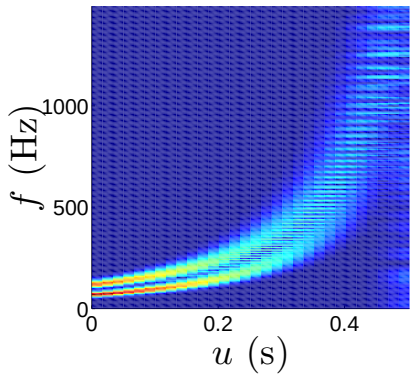


Figure 8.8: Spectrogram of the signal in Eq. (8.11) for n=500.

Fourier transform of this autocorrelation to obtain a time-frequency energy distribution. As a result of the STFT being linear, the spectrogram has no cross-terms. The counterpart is that the precision is not excellent in general. In any case, time and frequency resolutions can not be excellent simultaneously since they have to satisfy the uncertainty principle. Despites the processes leading to the Wigner-Ville transform and to the spectrogram are different, it has been shown that the spectrogram actually belongs to the class of the smoothed Wigner-Ville distributions (the Cohen class) if the correct kernel is employed. Another atomic decomposition is the continuous Wavelet transform that will be studied in Chapter 9.

9 Continuous Wavelet Transform

9.1 Introduction

The wavelet transform is a rather recent tool, since it appeared around 1983 in the works of Morlet and Grossman. Some precursors existed long before that period though (such as the orthogonal Haar basis, 1909).

The term "Wavelet transform" encompasses several kinds of related techniques:
• the continuous wavelet transform;
• the discrete wavelet transform;
• multiresolution algorithms.
A classical and comprehensive reference on wavelets including all these topics is the book by Mallat [20]. This book may be difficult as a first reading, and a much simpler introduction may be found in the book by Addison [1] which considers applications in fluid mechanics and in many other scientific domains. Many wavelet libraries are available on internet, the most commonly encountered is the Wavelab library for Matlab available at the address:
http://www-stat.stanford.edu/~wavelab/
Check also the TFTB library mentioned in the introduction of Chapter 7.

Wavelets are used for:
• Signal analysis, which is performed either with the continuous or the discrete wavelet transform. A discrete-time form of the continuous wavelet transform needs to be used for digital data (and this is different from the discrete wavelet transform).
• Signal compression (sparse representation) and denoising (this involves both an analysis step and a reconstruction step). These are based on the discrete wavelet transform.
• Identification of signal singularities and study of fractal behaviours.

In this course, the continuous wavelet transform is introduced. Like the STFT, with which it shares some properties, it belongs to the class of atomic decompositions. For an atomic decomposition, a family of atoms is needed, and the atoms are wavelets that are scaled and translated versions of a mother wavelet. This is presented in section 9.3. The wavelet transform is the projection of the signal onto the wavelets and depends on two parameters: time and scale. Hence the wavelet transform is called a time-scale analysis. This is very similar to time-frequency analysis, except frequency is replaced by scale, and scale is roughly inverse proportional to frequency.

Squaring the wavelet transform provides the scalogram which is a time-*frequency* energy distribution. This is presented in section 9.4. Some examples of wavelet transform applied to classical signals are presented in section 9.5. Applications in fluid mechanics are considered in section 9.7.

Before these topics are adressed, section 9.2 presents the Haar basis which is the first historical example of a wavelet family. Even if it belongs to the realm of the discrete wavelet transform, this basis is useful to introduce the wavelets in general.

9.2 The Haar basis

In 1909, Haar introduced the following function, called the Haar mother wavelet:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ \text{-1} & 1/2 \leq t < 1 \\ 0 & \text{elsewhere} \end{cases} \qquad \text{(Mother Haar wavelet)} \tag{9.1}$$

This function is plotted in Fig. 9.1. This function can be re-scaled and translated to define a

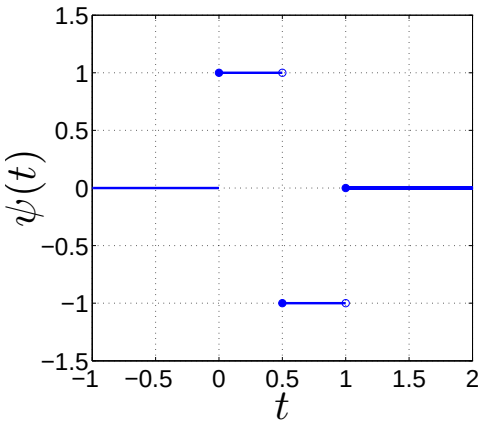


Figure 9.1: Haar wavelet.

whole family of functions that are called the daughter wavelets. The set of functions $\{\psi_{nk}(t), n \in N, 0 \leq k < 2^n\}$ is defined by:

$$\psi_{nk}(t) = 2^{n/2}\psi(2^n t - k) \qquad \text{for} \qquad n \in N,\ 0 \leq k < 2^n \qquad \text{(Haar wavelets)} \tag{9.2}$$

All these functions are null outside the interval [0 1]. The integer n gives the scale: the larger n, the smaller the time support of the function. The integer k gives the position of the function. Some functions are plotted in Fig. 9.2. For n=2, there are 4 possible values of k corresponding to 4 functions. For n=3, there are 8 possible values of k corresponding to 8 functions, and only 4 of them are shown. For n=6, there are 64 possible values of k corresponding to 64 functions, and again only 4 of them are shown. This figure shows how the scale of $\psi_{nk}(t)$ decreases when n increases.

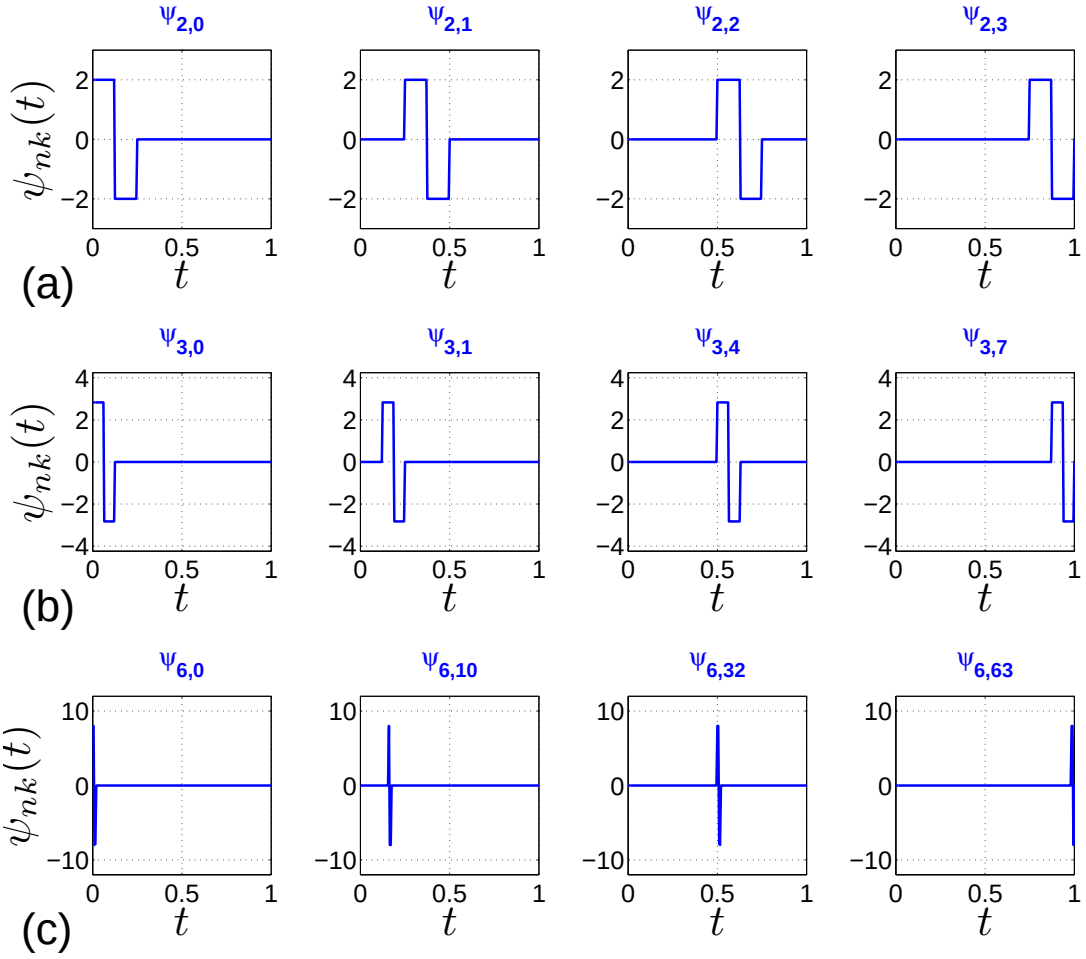


Figure 9.2: Some wavelets derived from the Haar mother wavelet with: (a) n=2, $0 \leq k < 4$; (b) n=3, $0 \leq k < 8$; (c) n=6, $0 \leq k < 64$. Only four wavelets (four values of k) are shown in each case.

What is remarkable about this set of functions is that it forms a complete orthogonal basis for the space of real continuous functions on the interval [0 1]. This means that any function $x(t)$ on this interval may be expanded as:

$$x(t) = \sum_{n=0}^{\infty} \sum_{k=0}^{2^n} C(n,k)\psi_{nk}(t)$$

where C_{nk} is the projection of $x(t)$ on $\psi_{nk}(t)$, that is:

$$C(n,k) = \langle x|\psi_{nk}\rangle = \int_0^1 x(t)\psi_{nk}^*(t)\mathrm{d}t$$

The Haar wavelets have the following property:

$$\int_0^1 \psi_{nk}(t)\mathrm{d}t = 0 \qquad \forall n,\ k$$

Any Haar wavelet has a null average. An equivalent relation will hold for other wavelets. This property allows understanding why wavelets are good to detect signal singularities. Consider first a locally **smooth function**, as in Fig. 9.3(a). The wavelet coefficient is obtained by multiplying

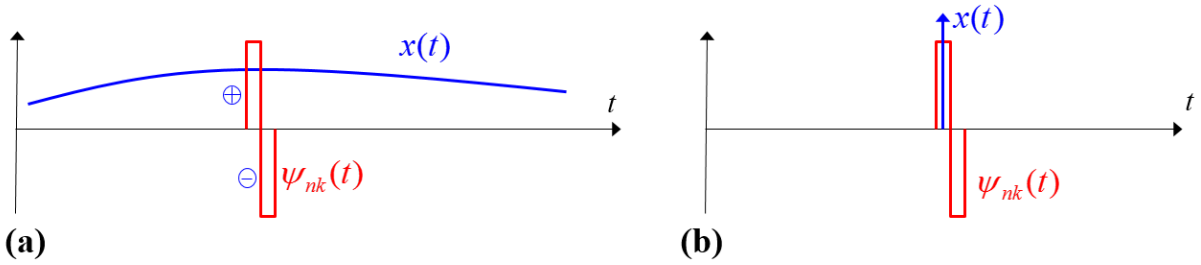


Figure 9.3: (a) smooth function for which $C_{nk} \to 0$ as $n \to \infty$; (b) singular function for which $C_{nk} \neq 0$ as $n \to \infty$.

this function by a wavelet ψ_{nk} and integrating. It is clear that by letting $n \to \infty$, that is to say, by taking a very small scale wavelet, one has:

$$C_{n,k} = \langle x|\psi_{nk}\rangle = 0$$

This is because both the positive and negative parts of the wavelet are multiplied by the same constant (this is the value of the function at the considered point), and the result is then proportional to the average of the wavelet that is null as said above. Let's take a Dirac impulse as our **singular signal** (see Fig. 9.3(b)). The Dirac impulse, if it belongs to the support of the wavelet, falls either in the positive or in the negative part of the wavelet, and in this case:

$$C_{n,k} = \langle x|\psi_{nk}\rangle \neq 0$$

There is always one wavelet ψ_{nk} whose coefficient C_{nk} is different from zero when $n \to \infty$. Hence, by looking at the wavelet transform when $n \to \infty$, one can assess whether the function is smooth or singular, and if it is singular, one knows where the singularity is located.

9.3 The wavelet family

In this section, we define a family of wavelets (atoms) that can be used for the continuous wavelet transform. First a **mother wavelet** is defined, and subsequently a whole family can be obtained from the latter by translation and rescaling.

9.3.1 Mother wavelet and admissibility condition

Not any function can be used as a mother wavelet. To be a mother wavelet, a function $\psi(t)$ has to belong to L^2 (meaning its energy $||\psi||^2 = \int_{-\infty}^{\infty} |\psi(t)|^2 dt$ is finite) and needs to satisfy the following **admissibility condition**:

$$C_\psi < \infty \qquad \text{where} \quad C_\psi = \int_0^\infty \frac{|\hat{\psi}(f)|^2}{|f|} df \qquad \text{(admissibility condition)} \tag{9.3}$$

where $\hat{\psi}(f) = \mathrm{FT}[\psi]$. This condition is a problem at f=0, where $\hat{\psi}$ needs to go down to zero with a sufficiently high rate.
In case ψ is integrable, the condition is equivalent to the simpler form:

$$\hat{\psi}(0) = \int_{-\infty}^{\infty} \psi(t)\mathrm{d}t = 0 \qquad \text{(admissibility condition, form 2)} \tag{9.4}$$

This is the form to be remembered. It states that a function, to be a mother wavelet, should have a **null average**. This was the case already for the Haar wavelet. A mother wavelet has also to be either **real** or **complex analytic**.

In addition, the mother wavelet may sometimes be required to have some further null moments. For instance, to study the n-th derivative of a function, one may want to use a mother wavelet that has null moments up to order n:

$$\int_{-\infty}^{\infty} t^m \psi(t)\mathrm{d}t = 0 \qquad \forall m \leq n$$

9.3.2 Some classical mother wavelets

Some mother wavelets commonly encountered wavelets for the continuous wavelet transform are given in Table 9.1.

	$\psi(t)$	$\hat{\psi}(f)$
Morlet (complex wavelet) ($\omega_0 \geq 5$)	$\frac{1}{\pi^{1/4}} e^{-\frac{t^2}{2}} e^{\mathrm{j}\omega_0 t}$	$\frac{\sqrt{2\pi}}{\pi^{1/4}} e^{-\frac{(\omega-\omega_0)^2}{2}}$
Derivative of Gaussian (DOG) (real wavelet) m: derivative order	$\frac{(-1)^{m+1}}{\sqrt{\Gamma(m+0.5)}} \frac{\mathrm{d}^m}{\mathrm{d}t^m} \left(e^{\frac{-t^2}{2}} \right)$	$\frac{-\sqrt{2\pi}(-\mathrm{j}\omega)^m}{\sqrt{\Gamma(m+0.5)}} e^{\frac{-\omega^2}{2}}$
Mexican hat (DOG for m=2) (real wavelet)	$\frac{-2}{\pi^{1/4}\sqrt{3}}(t^2-1)e^{-\frac{t^2}{2}}$	$\frac{\sqrt{8}\pi^{1/4}\omega^2}{\sqrt{3}} e^{-\frac{\omega^2}{2}}$

Table 9.1: Some classical mother Wavelets. ω stands for $2\pi f$.

The Morlet and Mexican wavelets are plotted in Fig. 9.4. The term wavelet is easily explained from this figure: the functions have oscillations typical of a wave, but these oscillations appear in a limited period of time. In the Morlet wavelet, we find indeed a wave, $e^{\mathrm{j}\omega_0 t}$, multiplied by an envelop $e^{-\frac{t^2}{2}}$.

The wavelets in Table 9.1 satisfy the admissibility condition. They are all normalized so that

$$||\psi|| = 1$$

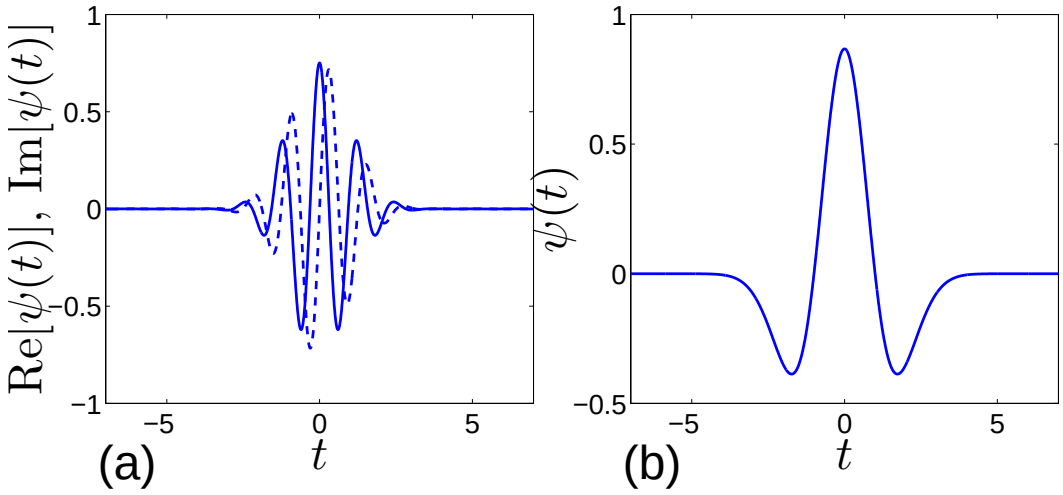


Figure 9.4: (a) The Morlet mother wavelet (for ω_0=5); (b) The Mexican mother wavelet.

. Different normalizations can be used by different authors, and in that case the value of the constant C_ψ will be different. The value of C_ψ for the Morlet (with ω_0=5) and for the Mexican mother wavelets are given in Table 9.2. The Morlet wavelet is complex and analytic. It was

	C_ψ	$f_{m,0}$	$T_{e,0}$
Morlet with ω_0=5	~ 1.2835	~ 0.7958 ($=f_0=\omega_0/(2\pi)$)	~ 0.707
Mexican hat	$\frac{4\sqrt{\pi}}{3}$	~ 0.239	~ 1.08

Table 9.2: Values of some parameters for the Morlet and Mexican mother wavelets: the admissibility constant C_ψ; the mean frequency $f_{m,0}$ (defined in Eq. (9.9)); the mean duration $T_{e,0}$.

introduced by Morlet in 1983. The formula given in the table contains a parameter ω_0 which imposes the number of oscillations under the envelope. In Fig. 9.4(a), we have taken ω_0=5. This corresponds approximately to 5 oscillations under the envelope. For the formula in the table to satisfy the admissibility condition, ω_0 should be chosen larger than 5. Otherwise, the wavelet does not satisfy the admissibility condition (its average is not null), and a more complete formula should be used instead:

$$\psi(t) = \frac{1}{\pi^{1/4}} e^{-\frac{t^2}{2}} (e^{\mathrm{j}\omega_0 t} - e^{-\omega_0^2/2}) \qquad \text{(Morlet wavelet } \forall \omega_0)$$

The Mexican hat wavelet is a real wavelet, and this is a particular case of the DOG wavelet with m=2.

In general, the Morlet wavelet is used to detect waves in signals, while the Mexican wavelet is preferred to detect maxima. The DOG with m=1 is used to detect jumps/discontinuities in signals.

9.3.3 Wavelet family

Once a mother wavelet has been chosen that satisfies the admissibility condition, a whole family of wavelets $\psi_{us}(t)$ can be constructed by translating and rescaling the mother wavelet. The wavelets $\psi_{us}(t)$ are called **daughter wavelets** and are defined by:

$$\psi_{us}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \qquad \forall u \in R,\ \forall s \in R^+ \qquad \text{(daughter wavelets)} \qquad (9.5)$$

The parameter u represents the time at which the wavelet is translated. This time can take *continuous* real values. This is a difference with the Haar wavelets for which the translation parameter k could take integer values only. The parameter s represents the scale (or size) of the wavelet and it can take any *continuous* positive real value. It is positive because a size is positive. The fact that it can take continuous values is a difference with the parameter n in the Haar wavelets. The factor $\frac{1}{\sqrt{s}}$ in Eq. (9.5) warrants that $||\psi_{us}||$=1. Hence, we have a family of wavelets $\psi_{us}(t)$ with u and s taking any real (and positive for s) values. This continuity is one reason for the term continuous wavelet transform to be introduced later. Note that the family $\psi_{us}(t)$ bears some similarities with the atoms $w_{uf}(t)$ in Eq. (8.4) for the STFT. There are also some differences that will be explained below. Finally, note that by taking u=0 and s=1, one recovers the mother wavelet: $\psi_{0,1}(t)$=$\psi(t)$.

Some daughter wavelets are plotted in Fig. 9.5 for several values of s and u, the mother wavelet being the Morlet wavelet (given in Fig. 9.5(a)). This figure may be compared with

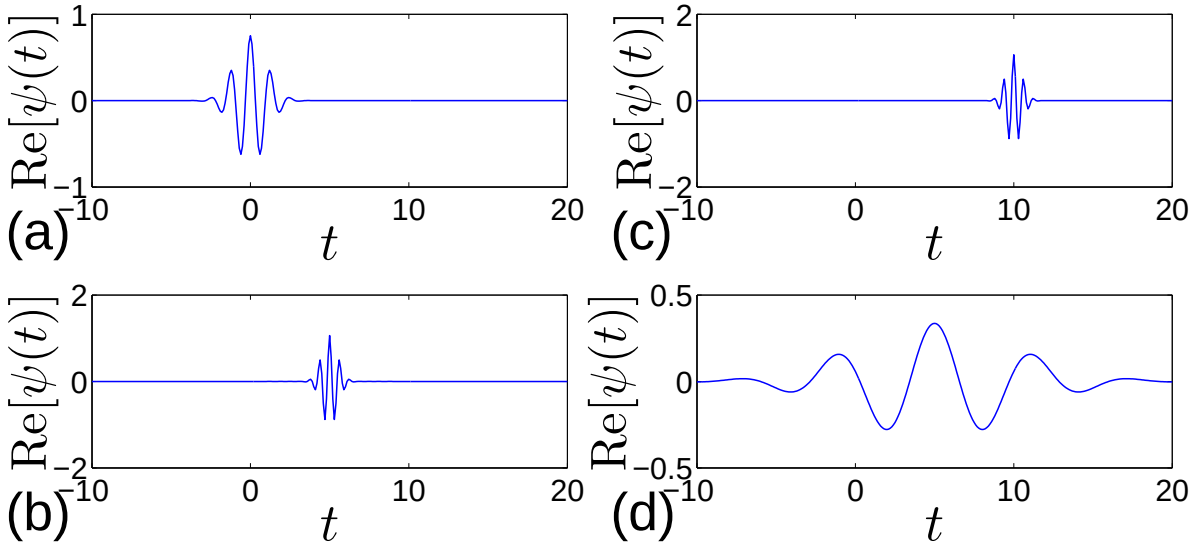


Figure 9.5: Some daughter wavelets based on the Morlet mother wavelet. These are given by Eq. (9.5) with: (a) u=0, s=1 (Morlet mother wavelet); (b) s=0.5, u=5; (c) s=0.5, u=10; (d) s=5, u=5. Note: the Morlet wavelet is complex, and the real part of the wavelets in plotted.

Fig. 8.3 that shows some atoms for the STFT. One important difference should be noted: the envelope for the wavelets is not fixed as it was the case for the STFT. In the STFT, the envelope

is fixed (meaning the window has always the same size), and the number of oscillations present under the envelope changes according to frequency. For the wavelets, the number of oscillations is fixed, and their typical frequency changes according to the scale: when the scale decreases, the oscillations have a higher frequency. One will remember that a frequency may be associated with a scale. We have approximately:

$$f \sim \frac{1}{s} \tag{9.6}$$

Of course, the precise proportionality constant needs to be given and depends on the particular type of wavelets used. One way of specifying this constant is given in the next subsection.

9.3.4 Wavelet Fourier transform

It is fruitful to consider the characteristics of the wavelet family in the frequency domain. Since a wavelet $\psi(t)$ necessarily belongs to L^2 (its energy $\int_{-\infty}^{\infty} |\psi(t)|^2 \mathrm{d}t$ is finite) its Fourier transform can be computed. The expression, Eq. (9.5), for the daughter wavelets is recalled to be:

$$\psi_{us}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right) \tag{9.7}$$

Daughters are obtained from the mother wavelet by translation $(t-u)$ followed by a rescaling $1/s$. The effect of the Fourier Transform on these two operations is well known. Refering to Table 3.1 a translation in the time domain will introduce a phase factor $e^{-j2\pi fu}$ in the frequency domain. The dilation property (see Table 3.1) states that

$$\psi(t) \underset{\mathrm{FT}}{\longrightarrow} \hat{\psi}(f) \quad \Rightarrow \quad \psi(t/s) \underset{\mathrm{FT}}{\longrightarrow} s\hat{\psi}(sf).$$

For $s < 1$ there is a contraction of the wavelet in the time domain and a dilatation of its FT in the frequency domain, and conversely for $s > 1$. All in all, the Fourier Transform of the daughter wavelet is:

$$\hat{\psi}_{us}(f) = \mathrm{FT}[\psi_{us}(t)] = \sqrt{s}\hat{\psi}(sf)e^{-j2\pi fu} \tag{9.8}$$

It is a scaled and phase-shifted version of the Fourier transform of the mother wavelet. The Fourier transform of the mother itself is given in Table 9.1.

Three Morlet wavelets and their FTs are shown in Fig. 9.6. The middle row correrspondsto the mother wavelet ($s = 1$). Since this wavelet consists of the regular product of a wave $e^{\mathrm{j}\omega_0 t}$ with a Gaussian enveloppe $e^{-\frac{t^2}{2}}$, its FT is the convolution product of a Dirac $\delta(f - \omega_0/(2\pi))$ and a Gaussian function, which is a Gaussian function of frequency centered at $f_0 = \omega_0/(2\pi) \sim 1$ (check the precise expression in table 9.1). Hence, the mother wavelet contains a packet of waves with frequencies around $f_0 = \omega_0/(2\pi)$. For the daughter wavelets it is seen in the figure that the center frequency of this packet changes depending on s: when $s < 1$ the oscillations are more rapid, and the packet is shifted to higher frequency values, while for $s > 1$ the contrary occurs. It is also seen that the smaller the scale (the duration) of a wavelet, the larger the scale (the bandwidth) of its FT, a consequence of the dilatation property. Finally note that $\hat{\psi}(f = 0) = 0$, as it should be according to the admissibility condition given in Eq. (9.4). The Morlet wavelet is analytic, and as a result its FT has only components with positive frequencies.

Three Mexican hat wavelets and their FTs are shown in Fig. 9.7. Since the Mexican wavelet is real, its FT includes both positive and negative frequencies, and the modulus of the FT is even. Otherwise, the same general behavior as for the Morlet wavelet is obtained.

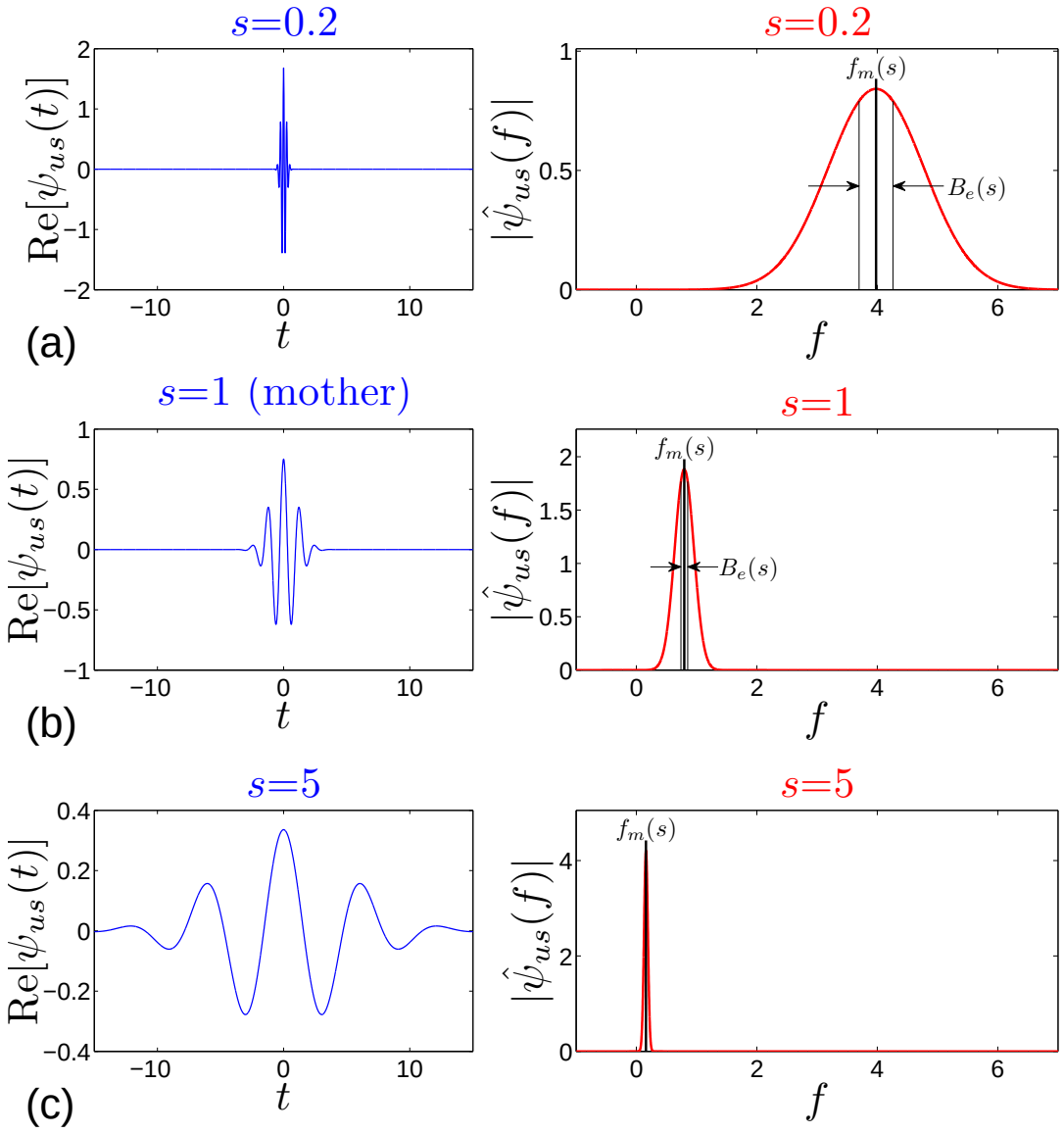


Figure 9.6: Left: real part $\mathrm{Re}[\psi_{us}(t)]$ of three Morlet wavelets having different scales; Right: modulus of their Fourier transform $|\hat{\psi}_{us}(f)|$. (a) s=0.2; (b) s=1 (Mother wavelet); (c) s=5. u=0 in all cases. In the plots on the right: the vertical thick line is located at f=$f_m(s)$; the two vertical thin lines are located at f=$f_m(s) \pm B_e(s)/2$.

The previous observations can be made more quantitative by computing such quantities as the mean time and effective duration of the wavelet in the time domain, and the mean frequency and effective bandwidth in the frequency domain. These will be useful in the following. Let us

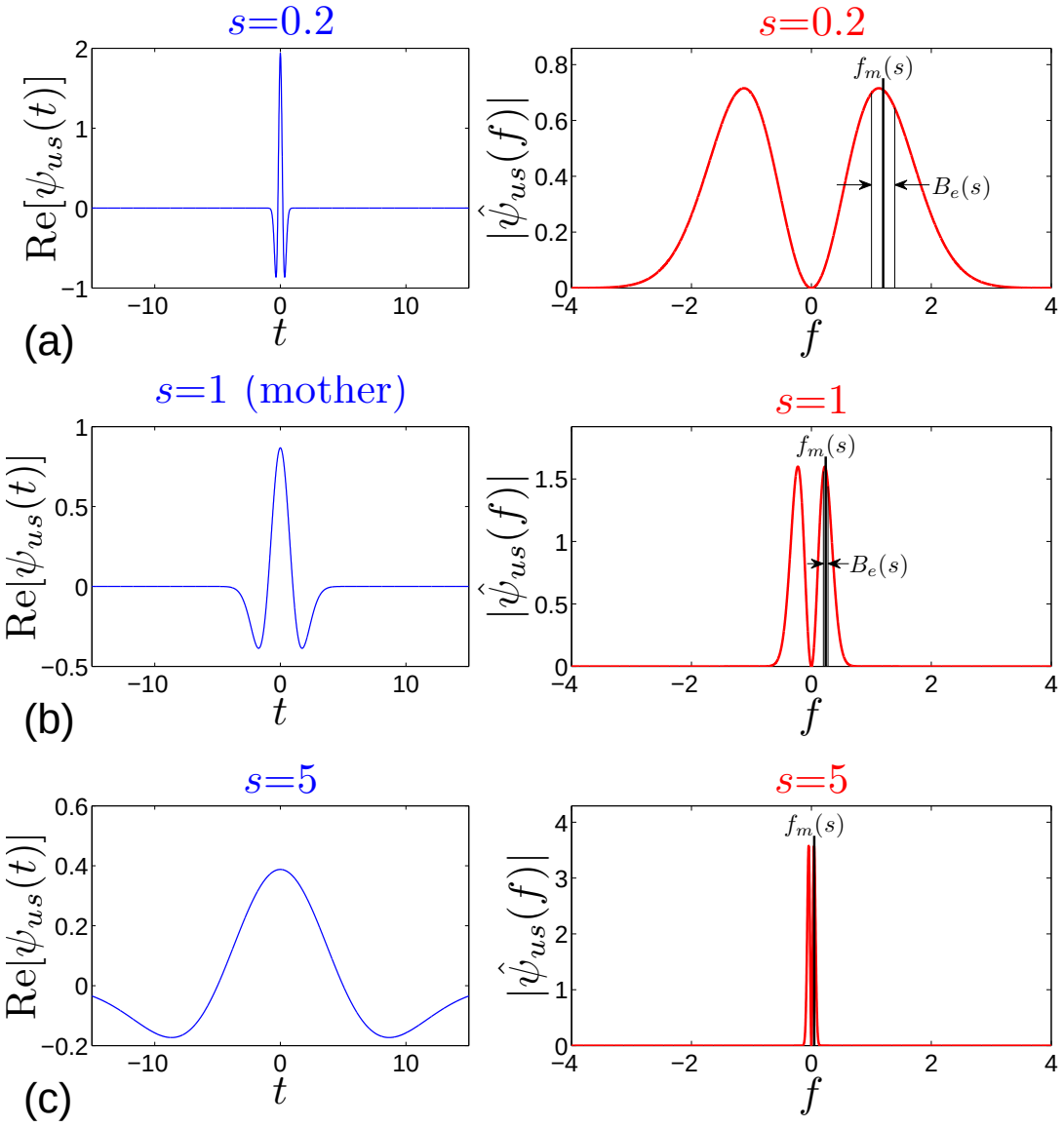


Figure 9.7: Left: three Mexican wavelets $\psi_{us}(t)$ having different scales: Right: the modulus of their Fourier transform $|\hat{\psi}_{us}(f)|$. (a) s=0.2; (b) s=1 (Mother wavelet); (c) s=5. u=0 in all cases. In the plots on the right: the vertical thick line is located at f=$f_m(s)$; the two vertical thin lines are located at f=$f_m(s) \pm B_e(s)/2$.

define:

- $t_{m,0}$ the mean time of the mother wavelet (=0 if wavelet even);
- t_m the mean time of a daughter wavelet;
- $T_{e,0}$ the effective duration of the mother wavelet;
- T_e the effective duration of a daughter wavelet;
- $f_{m,0}$ the mean frequency of the mother wavelet;
- f_m the mean frequency of a daughter wavelet;

- $B_{e,0}$ the effective bandwidth of the mother wavelet;
- B_e the effective bandwidth of a daughter wavelet.

The way f_m and B_e are calculated needs to be discussed in the case of a real wavelet (such as Mexican). The mean frequency we would like to have here is the frequency around which positive frequencies are clustered. This frequency is indeed indicated in the right plots in Figs. 9.6-9.7. However, for a real wavelet, the modulus of the FT is even, and applying Eq. (3.19) to calculate the mean frequency f_m (or $f_{m,0}$) would yield a null mean frequency. The formula for calculating f_m is thus modified in the present chapter so that integration is performed over positive frequencies only:

$$f_m = \frac{\int_0^\infty f|\hat{\psi}_{us}(f)|^2 \mathrm{d}f}{\int_0^\infty |\hat{\psi}_{us}(f)|^2 \mathrm{d}f} \tag{9.9}$$

This is indeed the mean frequency that one would obtain by applying Eq. (3.19) to the analytic associate of the wavelet. For the same reason, once the mean frequency has been determined, the effective bandwidth is calculated by the following relation:

$$B_e^2 = \frac{\int_0^\infty (f-f_m)^2|\hat{\psi}_{us}(f)|^2 \mathrm{d}f}{\int_0^\infty |\hat{\psi}_{us}(f)|^2 \mathrm{d}f} \tag{9.10}$$

The values of B_e are indicated in the top right and middle right plots in Figs 9.6-9.7.

Using the above definitions, the following relations may be derived that connect the characteristics of the mother and daughter wavelets. The relation on mean times is:

$$t_m = t_{m,0} + u \tag{9.11}$$

Equation. (9.11) shows that the mean time for a daughter wavelet placed at time u is simply $t_{m,0} + u$. The meaning is clear: the daughter wavelet is the mother wavelet translated by an amount u; its appearence time is then that of the mother wavelet augmented with u. In most cases the relation is simply $t_m = u$ since $t_{m,0} = 0$, the mother wavelet being symmetric.
The relation on mean durations is:

$$T_e = sT_{e,0} \tag{9.12}$$

It tells us that the typical duration of a daughter wavelet is proportional to its scale, which is also intuitive.
The relation on mean frequencies is:

$$\boxed{f_m = \frac{f_{m,0}}{s}} \tag{9.13}$$

Importantly, Eq. (9.13) confirms what we have announced earlier: a frequency f_m can be associated with the wavelet scale s and is inversely proportional to the latter (this was loosely written $f \sim 1/s$ on page 152). The proportionality constant is now known: it is $f_{m,0}$, the mean frequency associated with the mother wavelet. The specific relation between frequency and scale thus depends on the mother wavelet that has been chosen. The values of $T_{e,0}$ and of $f_{m,0}$ for the Morlet and Mexican mother wavelets are given in Table 9.2. For the Morlet wavelet, $f_{m,0}$ is nothing but $f_0 = \omega_0/(2\pi)$, the value of the frequency for the sinusoid below the Gaussian

envelope ($f_{m,0}$=f_0 results from the Gaussian envelope being even).
Finally the mean bandwidths verify:

$$B_e = \frac{B_{e,0}}{s} \tag{9.14}$$

Exercise:

One considers a daughter wavelet ψ_{us} with u=5 second and s=0.001 second. The mother wavelet is the Morlet wavelet with ω_0=5. What is the value of the mean frequency associated with the daughter wavelet?

The quantities defined above verify:

$$B_e T_e = B_{e,0} T_{e,0} = \text{cst} \geq \frac{1}{4\pi} \tag{9.15}$$

This means that the mother wavelet and the daughter wavelets all satisfy the uncertainty principle. When the scale of the wavelet is modified, its bandwidth is also modified so that this principle remains satisfied. One also has:

$$\frac{B_e}{f_m} = \frac{B_{e,0}}{f_{m,0}} = \text{cst} \tag{9.16}$$

The bandwidth corresponding to any wavelet is therefore proportional to the central (or mean) frequency of this wavelet. This is indeed seen in Figs. 9.6 and 9.7: the higher the frequency corresponding to a wavelet, the larger the bandwidth.

All these properties are useful to understand the tiling of the time-scale plane in the continuous wavelet transform, and to interpret the wavelet transform as a band-pass filter (see the next section).

9.4 The continuous wavelet transform (CWT)

9.4.1 Continuous wavelet transform and Inverse

The objective of the continuous wavelet transform is to tell us which scales s (and hence which frequencies, since scales and frequencies are connected through Eq. (9.13)) are present in a signal x at some time u. It works in the same way as the short time Fourier transform, except its atoms are now the wavelets introduced so far. The **continuous wavelet transform** (CWT) $C(u,s)$ of a signal $x(t)$ is obtained by calculating the scalar product between $x(t)$ and the wavelet $\psi_{us}(t)$:

$$C(u,s) = \langle x(t) | \psi_{us}(t) \rangle = \int_{-\infty}^{\infty} x(t) \psi_{us}^*(t) \mathrm{d}t \qquad \text{(CWT)} \tag{9.17}$$

It depends on the time u at which the wavelet is placed, and on the scale s of the wavelet. The wavelet transform $C(u,s)$ is large when $x(t)$ resembles $\psi_{us}(t)$ (in the approximate time interval $[u - s/2\ u + s/2]$). Said otherwise, the coefficient $C(u,s)$ is large when the signal $x(t)$ contains the scale s at time u. This is illustrated in Fig. 9.8.

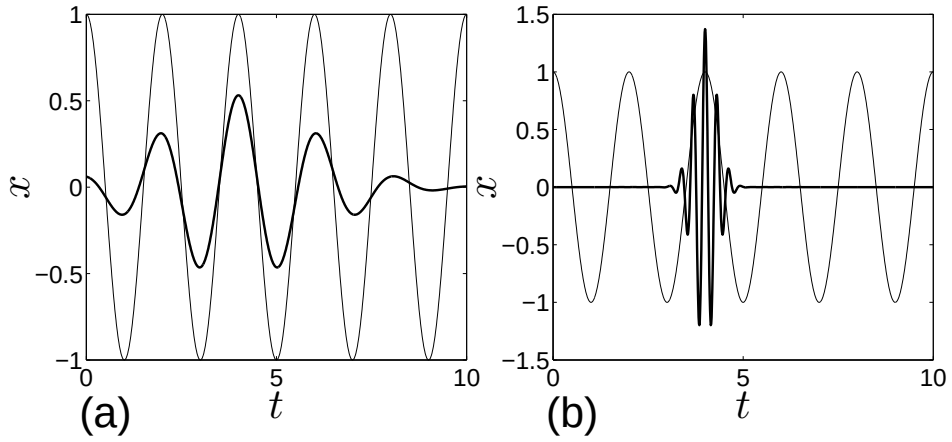


Figure 9.8: Thin line: signal $x(t)$; thick line: wavelet $\psi_{us}(t)$. (a) The signal resembles the wavelet with size s located at time u: $C(u,s) = \langle x|\psi_{us}\rangle \neq 0$; (b) The signal does not resemble the wavelet with size s located at time u: $C(u,s) = \langle x|\psi_{us}\rangle \sim 0$.

9.4.2 Inverse formula and Scalogram

Transforms are useful when they can be inverted and when they conserve energy. This is the case for the CWT. The reconstruction formula is different for real and analytic wavelets.

Real Wavelet
For a real wavelet (e.g. Mexican hat) and a real signal, we have:

$$x(t) = \frac{1}{C_\psi}\int_0^\infty \int_{-\infty}^\infty C(u,s)\psi_{us}(t)du\frac{ds}{s^2} \qquad (\psi \text{ real}) \tag{9.18}$$

Analytic Wavelet
For an analytic wavelet (e. g. Morlet) and a real signal, we have:

$$x(t) = \frac{2}{C_\psi}\mathrm{Re}\left[\int_0^\infty \int_{-\infty}^\infty C(u,s)\psi_{us}(t)du\frac{ds}{s^2}\right] \qquad (\psi \text{ analytic}) \tag{9.19}$$

In both cases, $C_\psi \neq 0$ from the admissibility condition (see Eq. (9.4)), and the signal can be recovered from its wavelet transform by integrating over all times and scales. Note also that the wavelet $\psi_{us}(t)$ appears in Eqs. (9.18)-(9.19) since we now expand the signal $x(t)$ using the atoms $\psi_{us}(t)$. The complex conjugate of $\psi_{us}(t)$ was found in the CWT (see Eq. (9.17)).

There is also a Parseval-like relation expressing the conservation of energy. The energy of the signal is given by:

$$E_x = \int_{-\infty}^\infty |x(t)|^2 \mathrm{d}t = \frac{\kappa}{C_\psi}\int_0^\infty \int_{-\infty}^\infty |C(u,s)|^2 \mathrm{d}u\frac{\mathrm{d}s}{s^2} \tag{9.20}$$

where

$$\kappa = \begin{cases} 1 & \text{for a real } \psi(t) \text{ (and a real } x(t)) \\ 2 & \text{for an analytic } \psi(t) \text{ (and a real } x(t)) \end{cases} \tag{9.21}$$

The square of the modulus of the CWT, $|C(u,s)|^2$ is called the **scalogram**. This represents an energy density, per *time* per *frequency*. Hence, this quantity is of the same kind as the spectrogram obtained by squaring the modulus of the STFT. That the scalogram is a density per time per frequency should be justified. We know that a scale s corresponds to a frequency f by:

$$f \sim \frac{1}{s}$$

Hence:

$$\mathrm{d}f \sim \frac{-\mathrm{d}s}{s^2}$$

and for any function $Q(s)$:

$$\int_0^\infty Q(s)\frac{\mathrm{d}s}{s^2} = \int_0^\infty Q(f)\mathrm{d}f$$

where $Q(f)$ denotes $Q(s = s(f)) = Q(s \sim 1/f)$. Hence,

$$E_x = \frac{\kappa}{C_\psi}\int_0^\infty \int_{-\infty}^\infty |C(u,s)|^2 \mathrm{d}u\frac{\mathrm{d}s}{s^2} = \frac{\kappa}{C_\psi}\int_0^\infty \int_{-\infty}^\infty |C(u,f)|^2 \mathrm{d}u\mathrm{d}f \tag{9.22}$$

where the notation $C(u,f)$=$C(u,s(f))$=$C(u,\sim 1/s)$ is used, which gives the result, since the integration of $|C(u,f)|^2$ is performed over u and f and no other factor depending on u or f is present within the integral.

9.4.3 Tiling of the time-scale plane

The wavelets define a tiling of the time-scale or time-frequency plane. The wavelet transform tell us by how much a signal is projected into each of the boxes of this tiling. If one considers the wavelet transform as a measuring instrument, then the size of the boxes is the precision of this instrument. For instance, if the signal can be represented by a spike in the time-scale plane, the wavelet transform will not return a spike, but a spot having for typical size the size of the box in which the spike is located. The tilings of the time-scale plane and of the time-frequency plane corresponding to the continuous wavelet transform are shown in Fig. 9.9. It should be pointed out that frequency f on the y-axis of the time-frequency plane here represents the *mean frequency* of the wavelet FT (noted f_m above). Elongated boxes in the vertical direction are found on the bottom of the time-scale plane, while they are found in the top of the time-frequency plane. This is due to the approximate relation $f \sim 1/s$ (see Eq.(9.13)). The tiling is not uniform in the plane, which is different from the STFT (see Fig. 8.4). The size of each of the boxes is fixed by the Heisenberg uncertainty principle, so that there is no escaping the fact that when the time resolution is good the frequency resolution is poor, and conversely. However, the trade-off varies in the plane.

Let's see how things work for the wavelet transform: in the wavelet transform, when one considers high frequencies, one accepts a large uncertainty on frequency. This is intuitively reasonable: one prefers to have a 100 Hz resolution on the measurement of a 1 kHz component (producing a 10% error) rather than on a 100 Hz component (producing a 100% error). This is exactly the meaning of Eq. (9.16) which is repeated here:

$$\frac{B_e}{f_m} = \frac{B_{e,0}}{f_{m,0}} = \mathrm{cst}$$

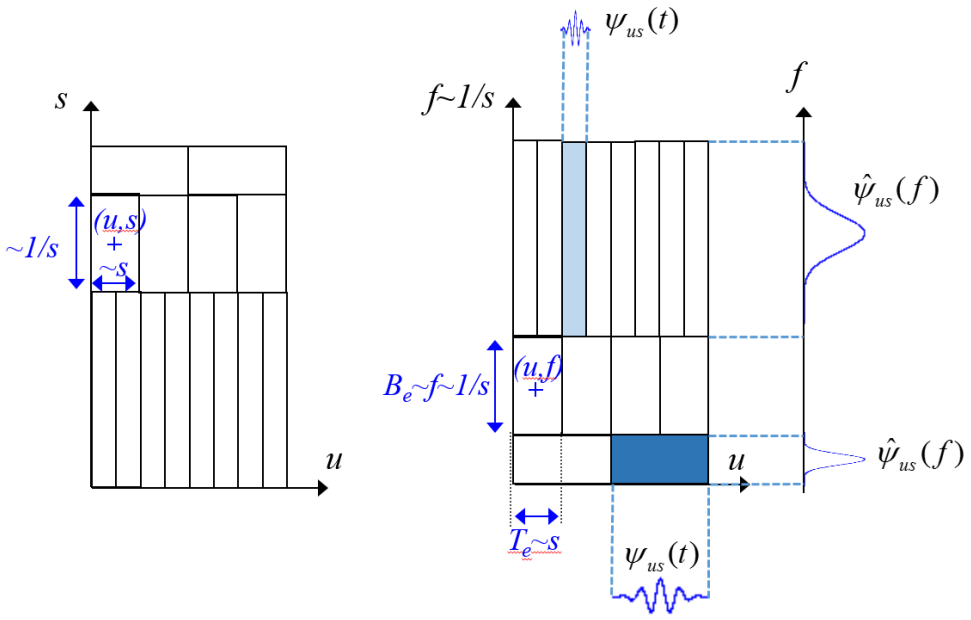


Figure 9.9: Tiling of the time-scale and time-frequency planes in the continuous wavelet transform. In the time-frequency tiling (right part), two Heisenberg boxes are highlighted, and the wavelets and wavelet FTs corresponding to these boxes are shown in the margins. The box colored in a light blue shade corresponds to a short-duration wavelet (small T_e), a large center frequency f (here, f stands for f_m), and a large bandwidth (large B_e); the box colored in a dark blue shade correponds to a wavelet with a longer duration (larger T_e), a smaller center frequency, and a smaller bandwidth (smaller B_e). The two boxes corresponds to the same $T_e B_e$ product.

This is seen in the time-frequency plane in the right part of the figure: the larger f on the y-axis (remember that in the time-frequency plane, f actually stands for f_m), the larger B_e (the poorer the frequency resolution). In the wavelet transform, high frequencies correspond to small scales and to short envelopes. To localize precisely these short events in time, one needs to have a small T_e (a high resolution in time) when s is small, and this is the case since $T_e = sT_{e,0}$ (see Eq. (9.12)). To summarize, for the wavelet transform the precision in time is high and that on frequency is low for the short-duration, high-frequency events. Hence, the wavelet transform allows one to zoom in on an event by considering small scales. This is the reason why the wavelet transform is well suited for discontinuities. This behaviour is different from that of the STFT. It is not better in general, it is better in certain situations.

Finally, note that in Fig. 9.9 separated boxes are represented. However, in reality any continuous position (u, s) of the box is possible for the continuous wavelet transform (the same as for the STFT).

9.4.4 The CWT as a band-pass filter

The CWT may be written in the form of a convolution product. To see this, let's start from the definition:

$$C(u,s) = \int_{-\infty}^{\infty} x(t)\psi_{us}^*(t)\mathrm{d}t = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)\mathrm{d}t \tag{9.23}$$

Let us define $h_s(t)$ by:

$$h_s(t) = \frac{1}{\sqrt{s}}\psi^*\left(\frac{-t}{s}\right) \tag{9.24}$$

It is a rescaled, time-reversed version of the mother wavelet $\psi(t)$. Then, the CWT can be written:

$$C(u,s) = \int_{-\infty}^{\infty} x(t)h_s(u-t)\mathrm{d}t = x * h_s(u) \tag{9.25}$$

The CWT is the convolution product between the signal and the impulse response h_s of a linear filter. This means that if one looks at the CWT $C(u,s)$ on a line in the time-scale plane (s is then fixed and u varies), then the CWT along this line (which is a signal depending solely on time u) is the convolution product of the signal with h_s. $C(u,s)$ as a function of u for a given scale s is related to the time-varying energy contained at scale s (more precisely, at scales about s).

One may look at this in the frequency domain by taking the Fourier transform in time of the CWT. Let's note:

$$\hat{C}(f,s) = \underset{u\to f}{\mathrm{FT}}[C(u,s)]$$

We have:

$$\hat{C}(f,s) = \hat{x}(f)\cdot\hat{h_s}(f) \tag{9.26}$$

with

$$\hat{h_s}(f) = \mathrm{FT}[h_s(u)] = \sqrt{s}\hat{\psi}(sf) \tag{9.27}$$

The quantity $\hat{h_s}(f)$ is the frequency response of the filter having $h_s(u)$ for impulse response. The frequency response tells us what frequency components are allowed to go through the filter (it also tells how these frequency components are phase-shifted but this does not interest us much presently). Equation (9.27) shows that the frequency response of the filter is directly related to that of the wavelet. The latter has been studied in Section 9.3.4. In particular, $\hat{h_s}(f=0)$=0 due to the admissibility condition (Eq. (9.4)). Hence, the low frequencies are cut off by this filter, and one can infer that this filter is a bandpass filter. That this is the case is illustrated in Fig. 9.10 which shows $|\hat{h_s}(f)|$ for 3 different values of s, when ψ is the Morlet wavelet. The plots in this figure are the same as those in the right column of Fig. 9.6. This is the case because from Eq. (9.8) we have $|\hat{h_s}(f)| = |\hat{\psi}_{us}(f)|$ ($\forall u$). As a result, the bandpass centre frequency of the filter is the mean frequency $f_m(s)$ corresponding to the wavelet at scale s (Eq. (9.13)), and its bandwidth is $B_e(s)$ (Eq. (9.14)). We have also seen that the ratio $B_e(s)/f_m(s)$ is constant (see Eq. (9.16)).

To understand the meaning of Eq. (9.26), let's consider again a fixed scale s. The quantity $\hat{C}(f,s)$ tells what frequency components are present at scale s, independent of time. Equation (9.26) shows that these frequency components are those of the signal ($\hat{x}(f)$) if allowed to

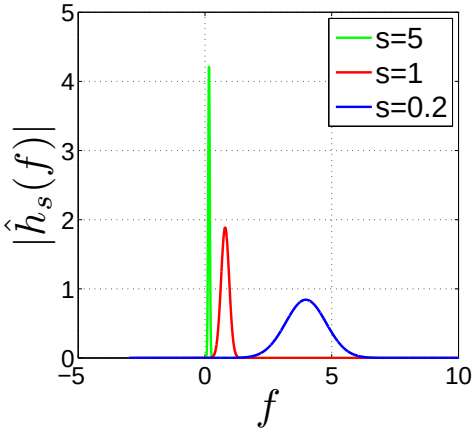


Figure 9.10: Bandpass filter frequency responses correponding to three wavelets having 3 different values of s. The Morlet mother wavelet is used with ω_0=5. The curves are the same as those given in the right column of Fig. 9.6.

pass through the filter $\hat{h_s}(f)$. Each line of the CWT (a line has s fixed) is thus obtained by bandpass filtering the signal with a filter having central frequency $f_m(s)$. It is the wavelet with scale s that plays the role of the filter.

Equations (9.25)-(9.26) have also a practical consequence on the digital computation of the CWT. A line of the CWT could be calculated by using Eq. (9.25). However, it is more efficient numerically to compute Eq. (9.26) and then take the inverse Fourier transform to get a line of the CWT. This *fast convolution* is allowed by the FFT algorithm. See lab 2 to obtain more details on the numerical implementation of the CWT.

9.4.5 Redundancy of the CWT

When calculating the wavelet transform $C(u,s)$ of a signal $x(t)$, we sort of increase the dimension of the signal, from 1D to 2D. It is then intuitively logical that $C(u,s)$ contains some redundant information about the signal. Let u and u_0 represent two different but close positions of two daughter wavelets ψ_{us} and $\psi_{u_0 s_0}$ having two close but different scales s and s_0, as shown in Fig. 9.11. Then, we expect $C(u,s)$ and $C(u_0,s_0)$ to share some information, which is known as redundancy. The **reproducing kernel** is defined by:

$$K(u_0,u,s_0,s) = \langle \psi_{us} | \psi_{u_0 s_0} \rangle = \int_{-\infty}^{\infty} \psi_{us}(t) \psi^*_{u_0 s_0}(t) \mathrm{d}t$$

This is nothing but the inner product between two different wavelets, and it measures the correlation (resemblance) between them. We may say as well that it measures the information shared by the wavelets.

Redundancy is then expressed by the following relation:

$$C(u_0,s_0) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{0}^{\infty} K(u_0,u,s_0,s) C(u,s) \mathrm{d}u \frac{\mathrm{d}s}{s^2} \tag{9.28}$$

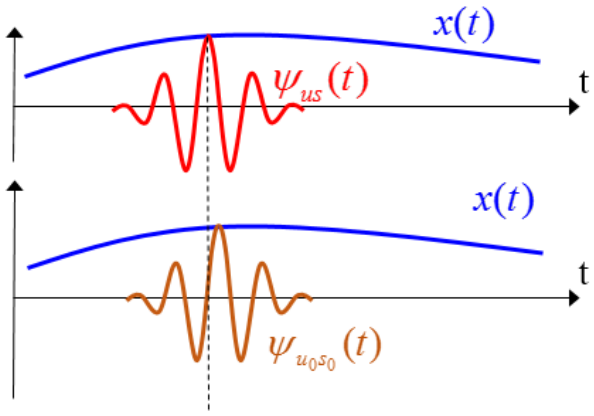


Figure 9.11: Projecting $x(t)$ on ψ_{us} provides $C(u,s)$. Projecting $x(t)$ on $\psi_{u_0 s_0}(t)$ provides $C(u_0,s_0)$. There is some information about the signal shared by $C(u,s)$ and $C(u_0,s_0)$, which is known as redundancy.

The CWT at point (u_0,s_0) depends on the CWT at all neighbors (u,s) for which $K(u_0,u,s_0,s)$ does not vanish. Hence, values of $C(u,s)$ are not "independent", meaning that the CWT is not optimal in terms of signal representation, since it stores the information at several places.

If the wavelets were orthogonal, then the reproducing kernel would be null, except for (u,s)=(u_0,s_0), and there would be no redundancy. The redundancy is then due to the wavelet family not being an orthogonal set (an example of orthogonal set is the Haar set). The wavelet transform based on such orthogonal sets is the discrete wavelet transform. For producing a sparse signal representation, the CWT is thus totally inappropriate and one should consider using the discrete wavelet transform instead. However, the CWT, while redundant, is useful for analysing signals in fluid mechanics, where one wants to understand the signal and not necessarily achieve a sparse representation of the signal. What is important is that the CWT does conserve energy, so that at least one knows the signal information is available. Another representation that is redundant it the STFT/spectrogram.

9.4.6 Smooth functions and functions with singularities

As for the Haar transform, the transform $C(u,s)$ is large for singular signals and negligible for smooth signals as the scale $s \to 0$. The admissibility condition (zero average wavelet) warrants this. The wavelet transform can indeed be used to characterize the regularity of a signal. The regularity can be characterized using the Holder exponent. A function V has an Holder exponent α at point x_0 if

$$|V(x) - V(x_0)| < |x - x_0|^{\alpha(x_0)}$$

with:
• $\alpha > 0$ large $\Rightarrow$ V smooth (or regular)
• $\alpha > 0$ small $\Rightarrow$ V less regular
• $\alpha < 0 \Rightarrow$ V singular
Let n be an integer larger than α, and suppose that the analysing wavelet has at least n vanishing

moments. Then:

$$|C(x_0, s)| < As^{\alpha(x_0)+0.5} \qquad \text{for} \quad s \to 0$$

This means that the wavelet transform can be used to characterize the regularity of a function.

9.5 Examples of continuous wavelet transform

Dirac impulse:
The CWT of $x(t) = \delta(t - t_0)$ is first computed, with here t_0=0.1. In that case, it can be easily obtained by hand calculation:

$$C(u, s) = \frac{1}{\sqrt{s}}\psi^*\left(\frac{t_0 - u}{s}\right)$$

At each scale s, $C(u, s)$ therefore resembles the wavelet translated at the position t_0 of the singularity. The CWT is shown in Fig. 9.12(a) for a Morlet wavelet and in Fig. 9.12(b) for a Mexican wavelet. For the Morlet wavelet, the 5 oscillations are seen (ω_0=5). Note that the vertical axis is $\log_2(s)$, which is commonly used for representing the CWT or the scalogram. Such a scaling "makes" room so that the small scales can be visualized. In both cases, there is a cone

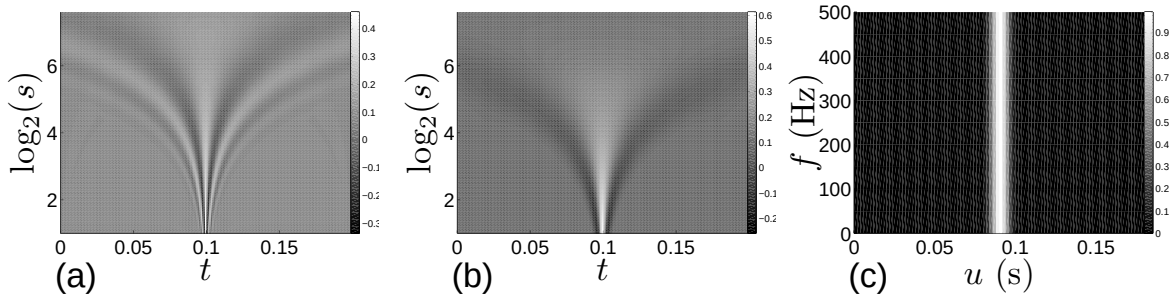


Figure 9.12: CWT of a Dirac impulse: (a) Real part of the CWT calculated with a Morlet wavelet (ω_0=5); (b) CWT calculated using a Mexican Wavelet; (c) Spectrogram.

pointing to the singularity. Hence, one is able to know the position of the singularity by looking at $C(u, s)$ for smaller and smaller values of s. This is a consequence of the shape of the tiling of the time-scale plane (see Fig. 9.9). This zooming capability at small scales is a distinctive feature of the CWT compared with the spectrogram which is shown in Fig. 9.12(c). For the Dirac, the spectrogram displays a vertical strip rather than a cone, and the time resolution is no better at high frequencies than at low frequencies.

White noise:
The CWT of a white noise is shown in Fig. 9.13. This CWT looks as if the signal was highly structured. However, it is not the case. What one can observe is the time-scale tiling.

Sinusoid:
The CWT of the signal $x(t) = \sin(2\pi 10t)$ is shown in Fig. 9.14. For a sinusoid, one is commonly using the Morlet wavelet. As a result, the CWT is complex and has a real part and an imaginary part, or a modulus and a phase. The real and imaginary parts are shown in Fig. 9.14(a) and 9.14(c). They are very similar, except for a time shift of a quarter of a period due to the real

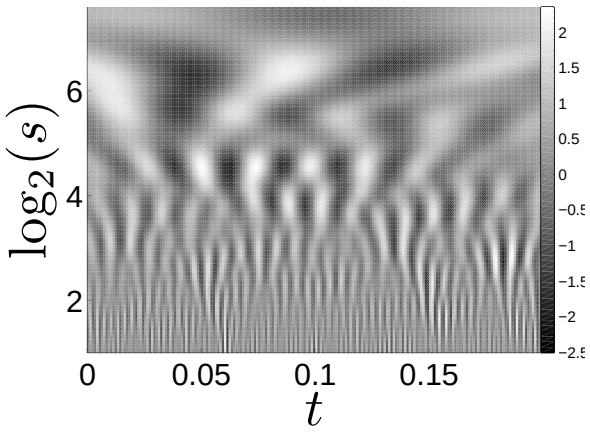


Figure 9.13: Real part of the CWT of a white noise.

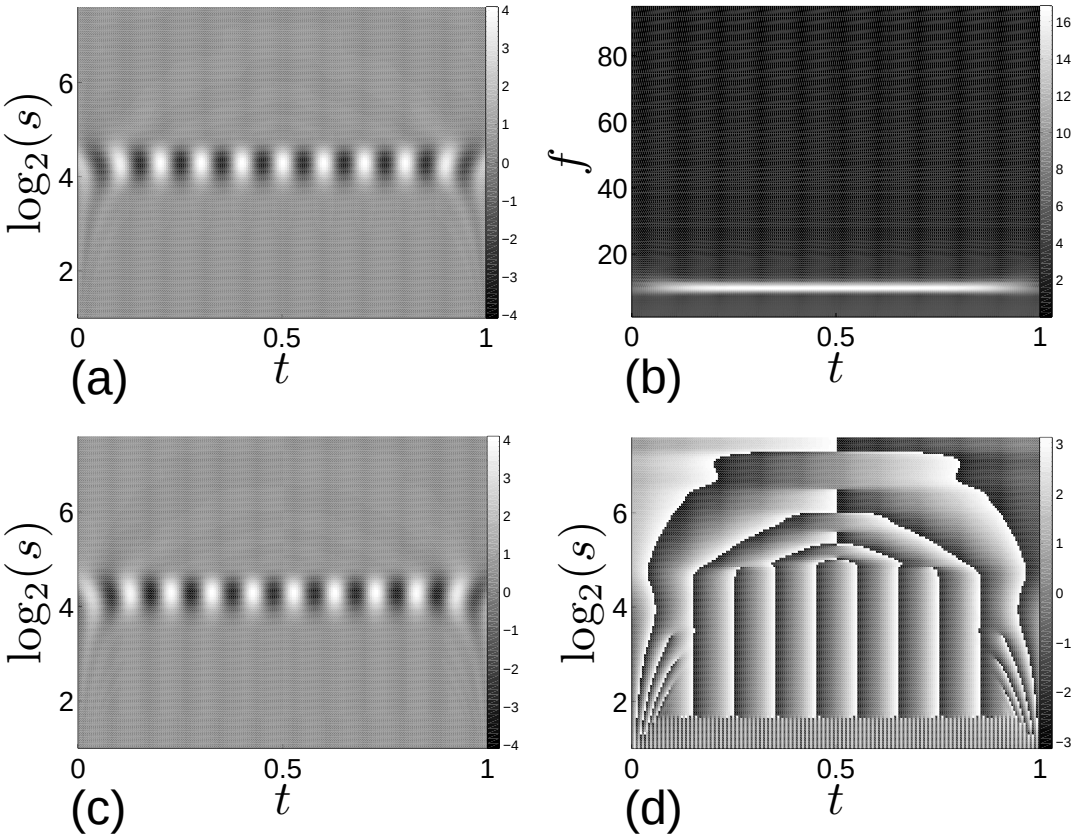


Figure 9.14: CWT of a sinusoid: (a) real part; (b) scalogram, with frequency on the y-axis; (c) imaginary part; (d) phase.

and imaginary part of the Morlet mother wavelet having such a time shift. They both show that there is some privileged scale, and that at this scale, the signal oscillates. The square of the

modulus (the scalogram) is shown in Fig. 9.14(b). The y-axis is linear and represents frequency. According to Eq. (9.13), scale has been converted to frequency by using:

$$f = f_m(s) = \frac{f_{m,0}}{s}$$

where $f_{m,0} = f_0 = 5/(2\pi) \sim 0.796$ (see Table 9.1). A constant energy at frequency 10 Hz (the frequency of the sinusoid) is seen in the scalogram. Finally, the phase of the CWT is shown in Fig. 9.14(d). The phase of the wave can be seen. The phase of the CWT is also used to visualize discontinuities in a signal.

Some artifacts seem to appear in Fig. 9.14(a). These artifacts are dues to edge effects. Because the signal is not available outside the measured interval, there is some discontinuities in the data at the limits of the interval, which shows up in and modify the CWT. At a given scale s, the CWT $C(u, s)$ is obtained by projecting the signal on ψ_{us}. The time support of ψ_{us} is loosely $[u - 3T_e \; u + 3T_e]$, and $C(u, s)$ is corrupted by the boundary when this boundary belongs to that time support. In addition, according to Eq. (9.12), we have $T_e = sT_{e,0}$. Hence, at each scale s, the part of the CWT that is corrupted is within a distance proportional to s (to be specific, a distance $3sT_{e,0}$) from the boundaries. Figure 9.15 plots the **cone of influence** originating from the boundaries which separates the regions affected or not by the boundaries. Below the cone,

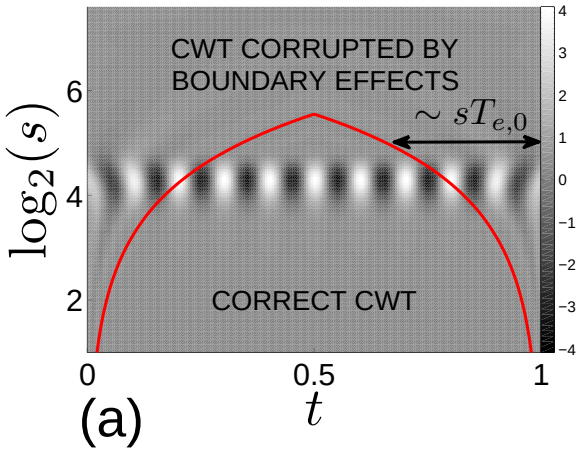


Figure 9.15: Cone of influence.

the CWT is correct. Above this cone, the data are corrupted due to boundary effects.

Sinusoid + Dirac:
The CWT for the sum of a sinusoid and a Dirac is shown in Fig. 9.16. It is the sum of the CWT of the sinusoid and the CWT of the Dirac. In addition, there seems to be a Dirac at each of the boundaries, this is an end effect due to the time interval being bounded.

Sum of two linear chirps:
The signal that was defined in Eq. (7.28) is considered again, and its CWT is shown in Fig. 9.17(a). The scalogram is shown in Fig. 9.17(b) where the y-axis displays the frequency.

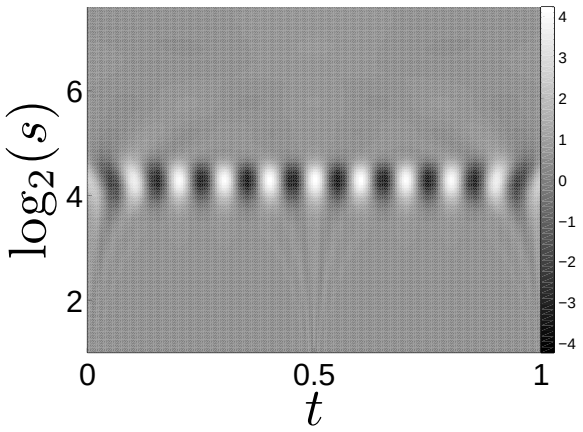


Figure 9.16: Real part of the CWT of the sum of a Dirac impulse and a sine.

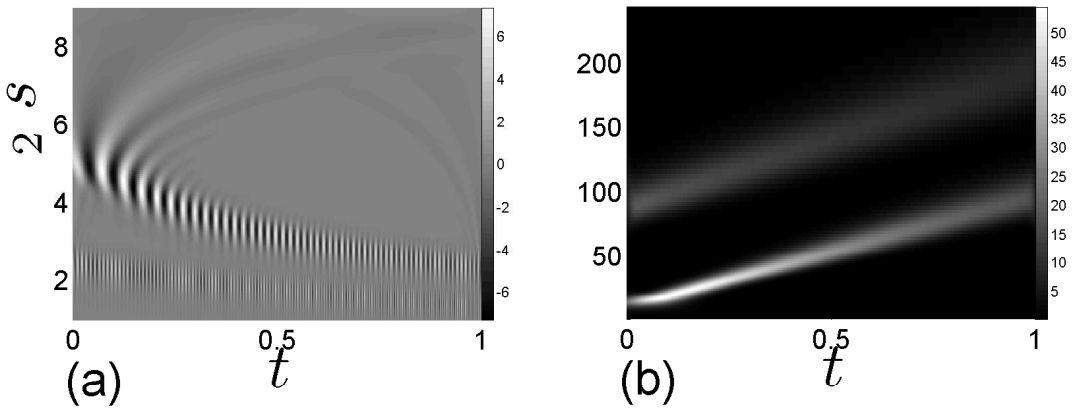


Figure 9.17: CWT of the sum of two linear chirps: (a) real part; (b) scalogram, with frequency on the y-axis.

The two chirps in the signal are seen, and the scalogram works as any time-frequency distribution.

Sum of two hyperbolic chirps:
Finally, the signal in Eqs. (7.32) and (8.11) is analyzed using the CWT. It is made of the sum of two hyperbolic chirps. Here, the CWT performs better than the spectrogram (see Fig. 8.8) and allows to separate the two frequencies over a longer period of time. The reason for this is the difference between the time-scale tilings in the STFT and in the CWT, as illustrated in Fig. 9.19. This evidences once more the high time resolution of the CWT at high frequencies.

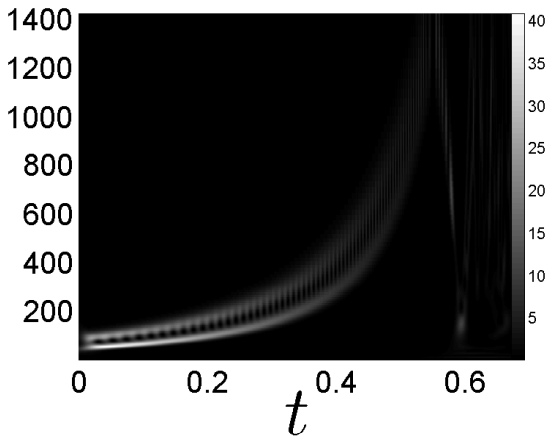


Figure 9.18: Real part of the CWT of the sum of two hyperbolic chirps.

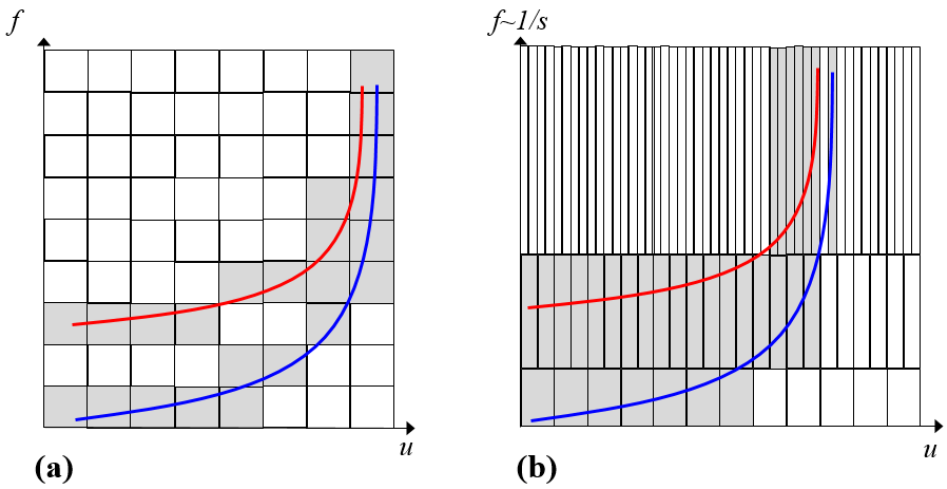


Figure 9.19: Time-scale tilings for (a) the STFT; (b) the CWT. The two hyperbolic chirps are also represented and their frequencies are close at high frequency. At high frequency, they can be separated by the CWT because they fall into two different Heisenberg boxes. They are not separated by the STFT because they fall in the same Heisenberg box.

9.6 The discrete wavelet transform

Let us say just a few words on the discrete wavelet transform (which is not to be confused with a discrete implementation of the continuous wavelet transform).

The discrete wavelet transform works with a family of wavelets which form an **orthogonal set**. For this set, the values taken by u and s (which could both take continuous values in the continuous wavelet transform) are now discrete, and are replaced by integer indexes k and j.

Such a wavelet family may be given by:

$$\psi_{nj}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j n}{2^j}\right) = 2^{-j/2}\psi(2^{-j}t-n) \qquad n \in Z,\ j \in Z \tag{9.29}$$

The wavelets ψ_{nj} are obtained from the mother wavelet by dilatations $s = 2^j$ and translations $u = n2^j$. The Haar family is such a family. The position of the wavelets in the time-scale plane is given by the dyadic grid shown in Fig. 9.20. This grid is sparse and does not cover the full

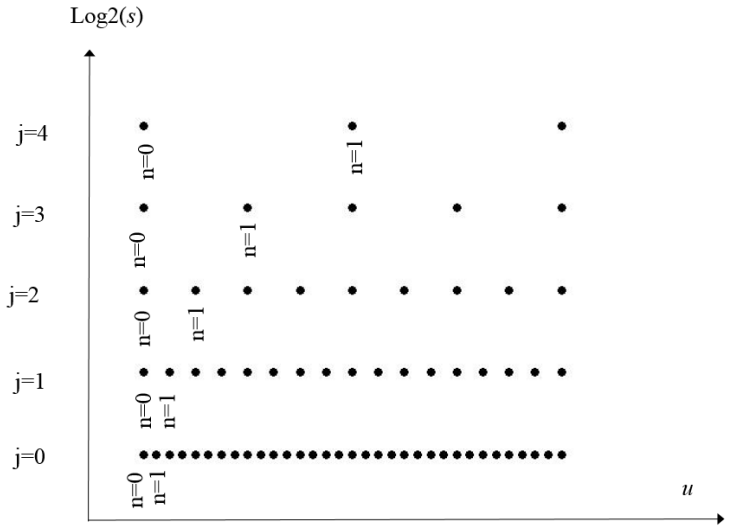


Figure 9.20: Dyadic grid for the discrete wavelet transform.

plane as in the CWT.

The discrete wavelet transform is:

$$C(n,j) = \langle x|\psi_{nj}\rangle$$

It is again the projection of the signal on the wavelets. The reconstruction formula is now given by a sum rather than by an integral:

$$x(t) = \sum_{j=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} C(n,j)\psi_{nj}(t)$$

The energy conservation formula is:

$$\int_{-\infty}^{\infty}|x(t)|^2\mathrm{d}t = \sum_{j=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}|C(n,j)|^2$$

The energy of the signal is given by the sum of the energies held by each of the wavelets, a property resulting from the orthogonality of the wavelets.

The fact that the wavelets are orthogonal means that the information contained in one wavelet is not contained in another wavelet. There is no redundancy of the information, as it

was the case for the CWT.

The discrete wavelet transform is better in terms of signal compression because it is not redundant. For this reason it is more widely spread than the CWT. An inconvenient of the discrete wavelet transform is that it is available only at the points of the dyadic grid, and this grid is coarse. One won't be able to find a frequency with some precision if this frequency does not belong to the dyadic grid. This prompts some researchers to use the CWT. A finer representation is sometimes better for human analysis.

9.7 The continuous wavelet transform in fluid mechanics

Some applications of the wavelet transform in fluid mechanics are now considered. Both the continuous and the discrete wavelet transforms have been used in this field. Two reference papers are those of Meneveau [24] and Farge [15]. See also Chapter 4 of the book by Addison [1] for a wide range of applications in fluid mechanics. Here we focus on the CWT, but things are not much different for the discrete wavelet transform. CWT can be applied to space or time, in one or several dimensions. Sometimes, time measurements are converted into spatial measurements using the Taylor hypothesis of a frozen turbulence, as in Katul et al [18] for example. The wavelet transform is used for several (related) purposes, including:
- statistics of the flow that are scale-dependent;
- intermittent behaviors;
- coherent structures characterization;
- pattern recognition in the time-scale plane.
Some of these applications are reviewed in the following.

The traditional statistical description of turbulence is done in the Fourier space, using a basis made of non-local Fourier modes. As a result, it does not describe properly the coherent structures whose dynamical interaction is essential. Of course, coherent structures can be written as sums of non-local waves whose cancellation/addition gives a local structure, but this relies on the phase of the Fourier transform which is difficult to grasp in a straightforward manner (see Section 6.1). To describe coherent structures, nearly compactly supported functions (in both space and scale) are preferred, and the wavelets are a good candidate.

The Fourier spectrum and the scalogram have some connection since $s \sim 1/f$. Using the CWT, the **global energy spectrum** (also called the wavelet spectrum) is defined by [15]:

$$E(s) = \int_{-\infty}^{\infty} |C(u,s)|^2 \mathrm{d}u \tag{9.30}$$

It is obtained by summing the square of the CWT over time at a given scale. Due to the relation $s \sim 1/f$, one can deduce from $E(s)$ a wavelet spectrum $E(f)$. This spectrum, obtained using the CWT, is not equal to the Fourier spectrum, but is a smoothed version of the Fourier spectrum (this is a result of the passband filter nature of the CWT). This is illustrated in Fig. 9.21 which compares the Fourier spectrum of a signal with wavelet spectra obtained either by continuous or discrete wavelet transform.

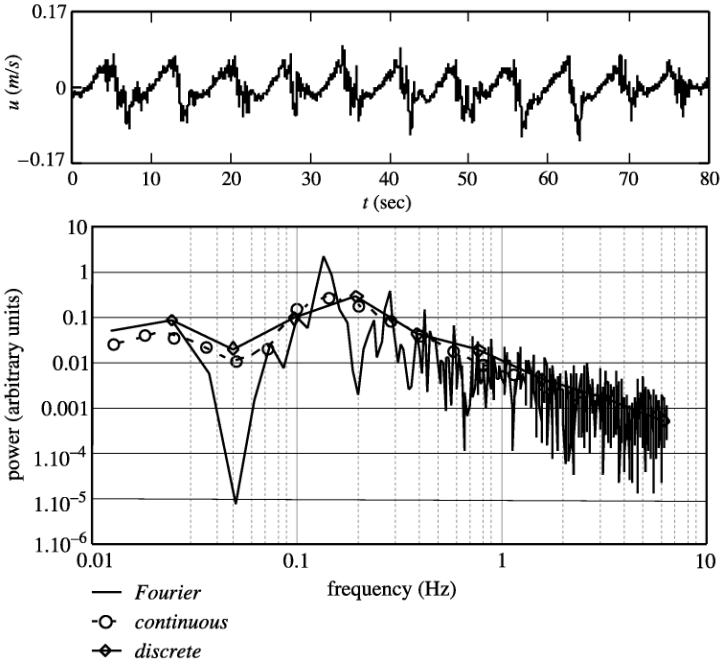


Figure 9.21: Wavelet spectrum compared with the Fourier spectrum. Top: signal downstream of a cylinder in an open channel flow. Bottom: power spectrum (Fourier, continuous wavelet spectrum using the Mexican wavelet, discrete wavelet spectrum). From Fig. 4.1 in Addison [1].

In the Fourier transform, a scale (wavelength) s=$1/f$, if present, is present over the whole space or time. In the wavelet transform, a given scale can be present in some regions of space and not in others. This **non-uniform distribution of the scale** in space or time is called **intermittency**. Intermittency is indeed very important in the understanding of turbulence for the following reason. The Kolmogorov theory (K41) predicts the following scaling:

$$\langle |V(x+r) - V(x)|^p \rangle \sim r^{p/3}$$

where $\langle |V(x+r) - V(x)|^p \rangle$ is the p-order structure function, $< . >$ representing an ensemble average. Many measurements show that this scaling is verified for p=2. However, the scaling becomes less accurate for $p >$2. This is attributed to intermittency in the inertial subrange [18]. Several quantities based on the CWT allows assessing the presence of intermittency.

The **local intermittency measure** (LIM) [15] is defined by:

$$LIM(u,s) = \frac{|C(u,s)|^2}{\langle |C(u,s)|^2 \rangle_u} \tag{9.31}$$

where $< . >_u$ is an average over time u. The LIM represents the energy at some time and scale, normalized by the average energy at that scale. If $LIM(u,s)$=1 $\forall u \forall s$, then any time contains the same energy as any other time, and there is no intermittency. If $LIM(u,s)$=20, say, then the time u contains 20 times more energy on average than the other times at scale s. There is then a strong intermittency at scale s. An example of LIM is shown in Fig. 9.22. The LIM is

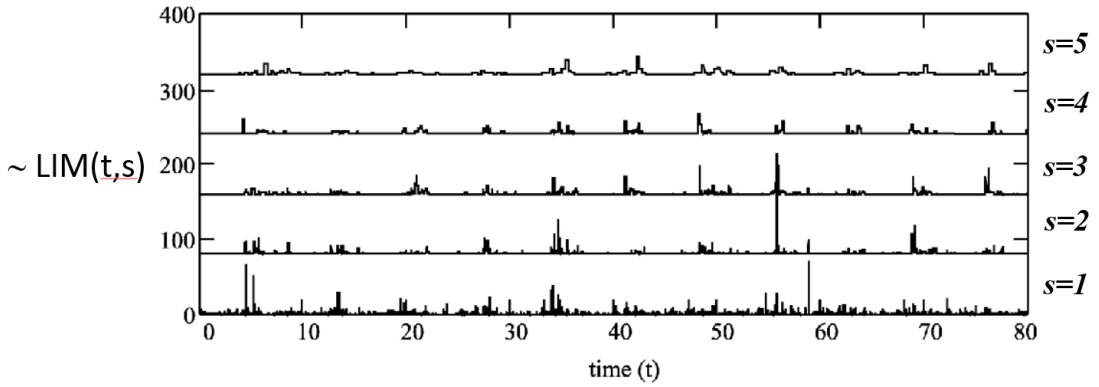


Figure 9.22: Local intermittency measure (LIM) for a signal measured downstream of a cylinder in an open channel flow, for 5 different scales. From Fig. 4.4 in Addison [1].

plotted for several scales. At the smallest scale s=1, there is a lot of spikes, showing a large **intermittency at small scales**. A similar result is obtained for a lot of flows (wakes, boundary layer flows, grid turbulence, jets,...). Hence, the CWT allows calculating the intermittency in time (or space) *at a given scale*. This is out of reach of the Fourier analysis since in the Fourier analysis a given wave at scale $s \sim 1/f$ has the same amplitude, and thus energy, over the whole time or space interval.

Another characteristics of intermittency is that "correlations in small-scale turbulent motions show significant deviations from the Gaussian statistics usually expected for large, randomly interacting systems" [31]. A quantity that is frequently used to assess Gaussianity is the **wavelet flatness factor** (FF) introduced by Meneveau [24] and defined by:

$$FF(s) = \frac{\left\langle |C(u,s)|^4 \right\rangle_u}{\left\langle |C(u,s)|^2 \right\rangle_u^2} \tag{9.32}$$

Compared to the usual flatness factor of a signal, this one depends on scale. A flatness factor FF=3 corresponds to a signal whose values are normally ditributed (that is, follow a Gaussian distribution). As intermittency is related to departure from a Gaussian distribution, a flatness factor which is different from 3 at some scale s indicates intermittency at that scale.

The LIM and FF are used for example by Camussi and Guj [11] to study intermittency in grid turbulence (they use the discrete wavelet transform). Figure 9.23(a) gives the flatness factor for two different Reynolds numbers. For the lower Reynolds number (Re_λ=3) the flatness factor is almost constant with a value of 3 independent of the scale, indicating that there is no intermittency at this Reynolds, for any scale. For the larger Reynods number (Re_λ=12) the flatness factor equals 3 at large scales (these scales are governed by a Gaussian law) but increases at small values of s, indicating intermittency at small scales. The LIM at the smallest resolved scale is given in Fig. 9.23(b). It indicates much more intermittency at Re_λ=12 than at Re_λ=3, which agrees with the results provided by the flatness factor. In addition, the authors characterize the flow structures responsible for intermittency. For doing this, they perform **conditional averaging** based on the LIM at small scales. The procedure is the following: the times at which the LIM exceeds some threshold value are detected, and excerpts of the signal (velocity) in time

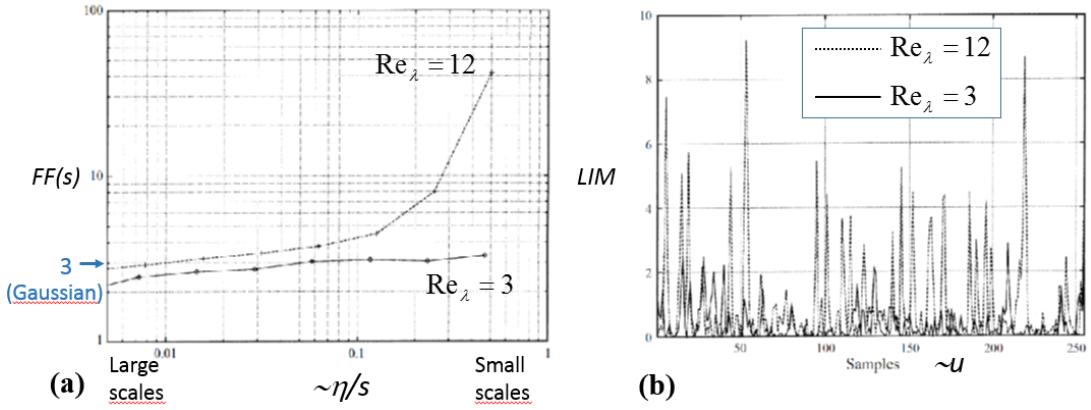


Figure 9.23: Wavelet transform of velocity signals in grid turbulence at two different Reynolds numbers based on the Taylor microscale: Re_λ=3 and Re_λ=12. (a) Flatness factor as a function of the inverse of the scale. η is the Kolmogorov scale. (b) Local Intermittency Measure (LIM) for the smallest resolved scale. From Camussi and Guj [11].

intervals centered on those detection times are isolated. All the excerpts so obtained are then averaged. The result is a structure velocity signature that appears on average when the LIM is large, and which is responsible for intermittency at small scales. This velocity signature is shown in Fig. 9.24. It corresponds to the velocity of a vortex, and the authors conclude that

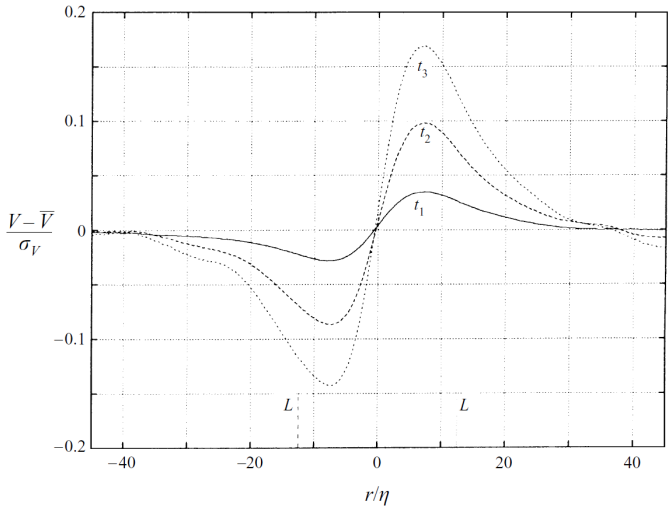


Figure 9.24: Turbulent structure velocity signature responsible for intermittency at small scales at Re_λ=12. This is obtained by conditional averaging, and the velocity depends on the chosen threshold (t_1, t_2, or t_3). From Camussi and Guj [11].

vortex tubes are responsible for small scale intermittency.

Conditional averaging as just seen, or any other type of **thresholding**, are common when performing wavelet transforms. The scalogram $|C(u,s)|^2$ represents an energy density. It is pos-

sible to filter a signal according to the energy contained in its wavelet transform. For example, one would calculate the wavelet transform of a signal, set to zero those values of $C(u,s)$ for which $|C(u,s)|^2$ does not exceed some threshold, and then compute a filtered signal by inverse wavelet transform. Two remarks have to be done:
- Thresholding is interesting if it retains a maximal quantity of energy using a minimal set of wavelet coefficients (see also POD). The threshold should be chosen with this objective in mind.
- The procedure is not fully rigorous for the continuous wavelet transform because the continuous wavelet transform modified that way (ie by cancelling some coefficients) is not necessarily the wavelet transform of a signal (because Eq. (9.28) is then not satisfied anymore). Nevertheless, it is common practice to do this anyway. For the discrete wavelet transform, there is no such inconvenient thanks to the lack of redundancy.
The resulting filtered signal contains only the most energetic features in the time-scale plane. An example of thresholding is shown in Fig. 9.25. The signal is shown in the top row of Fig. 9.25(c).

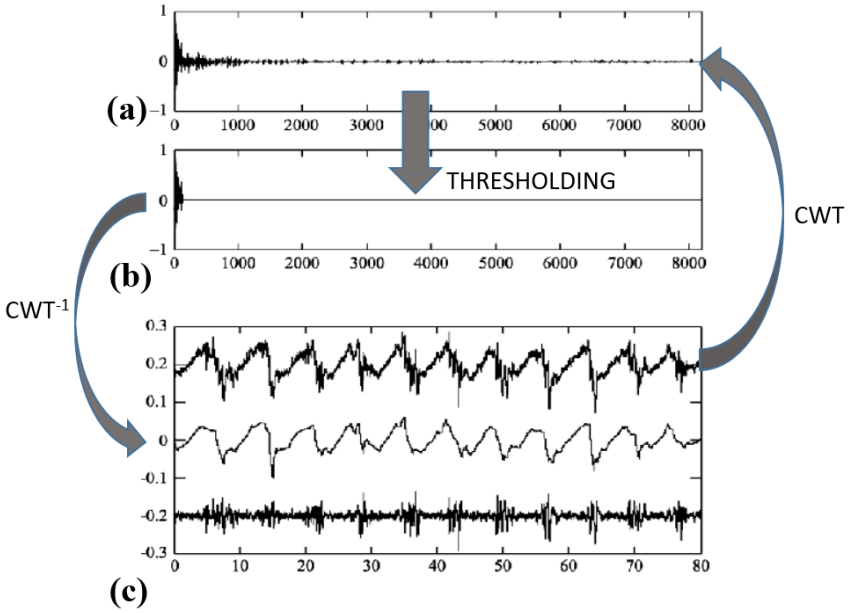


Figure 9.25: Thresholding of the CWT to produce a filtered signal (see text). From Addison [1].

The values of $C(u,s)$ are shown in Fig. 9.25(a) where they are sorted in descending order of their magnitude. By setting to zero the smaller values of $C(u,s)$ (smaller than some given threshold) one obtains the wavelet transform in Fig. 9.25(b). Finally, this filtered wavelet transform is inverted to obtain the signal in the middle row of Fig. 9.25(c). The difference between the original and filtered signal is shown in the bottom row of Fig. 9.25(c). There are many different ways of thresholding a signal using the wavelet transform (for example, the small scales could be suppressed in favor of the large scales).

Another possible use of the CWT consists of characterizing the structures contained in flow velocity fields. For example Camussi et al [10] use CWT to characterize the size of coherent vortices in velocity fields obtained by PIV. They use 1D CWTs (see Lab 2 where an algorithm

similar to theirs is used). Schram et al [30] also detect vortices in PIV-obtained velocity fields by using a 2D wavelet transform. An example of vortex detection performed on an instantaneous flow field is illustrated in Fig. 9.26.

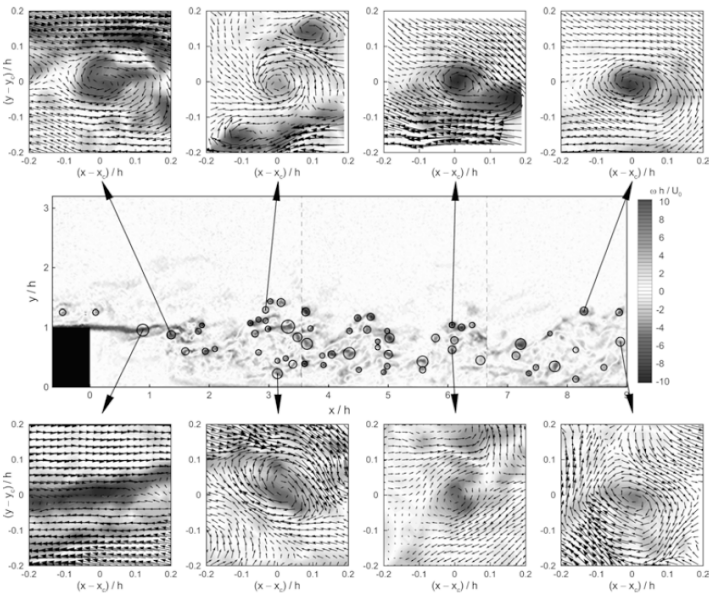


Figure 9.26: Example of vortex detection using 2D CWT and local velocity field in the frame of each vortex. From Schram et al (2004) [30].

Let us conclude with an example in aeroacoustics: Lewalle et al [19] apply the CWT to signals from microphones placed in the far-field of a jet (diameter $D\sim$0.05m, Re$\sim$7 10^5, M$\sim$0.6). The experimental apparatus is shown in Fig. 9.27(a). The three microphones used in the paper are those at 15^o, 30^o, and 45^o. These microphones measure the sound emitted by the jet and the signals they receive have some delay between them because sound does not arrive at the same time at the microphones. Cross-correlations between these microphones are shown in Fig. 9.27(b). Peaks in the cross-correlations are observed indicating that the signals resemble each other when a suitable lag is imposed between them, *on average*. What these cross-correlations do not tell us is whether the signals resemble each other uniformly over the whole signal duration or whether they resemble each other during some bursts of high energy (with these burst dominating the correlations on average). This information can be provided by using the wavelet transform, and Lewalle et al use the CWT based either on the Morlet wavelet or on the Mexican wavelet. Two scalograms are shown in Fig. 9.28(a) and 9.28(b) respectively for the microphones at 15^o and 30^o. Some events can be localized in both scalograms, with a comparable or different magnitude. Three pairs of events appearing at both microphones are indicated by circles. The events appear in a random fashion in the time-scale plane, meaning the sound received at the microphones and the source producing it is intermittent. By cross-correlating the CWT of the different signals at fixed scales, it is possible to isolate events that appear on the three microphones. Figure 9.28(c) give an example of three individual events (one per microphone) that correspond to the same burst. Knowing the lag between them and using triangulation, it is then possible to locate the source position, event per event. This example shows how CWT can be useful to isolate events

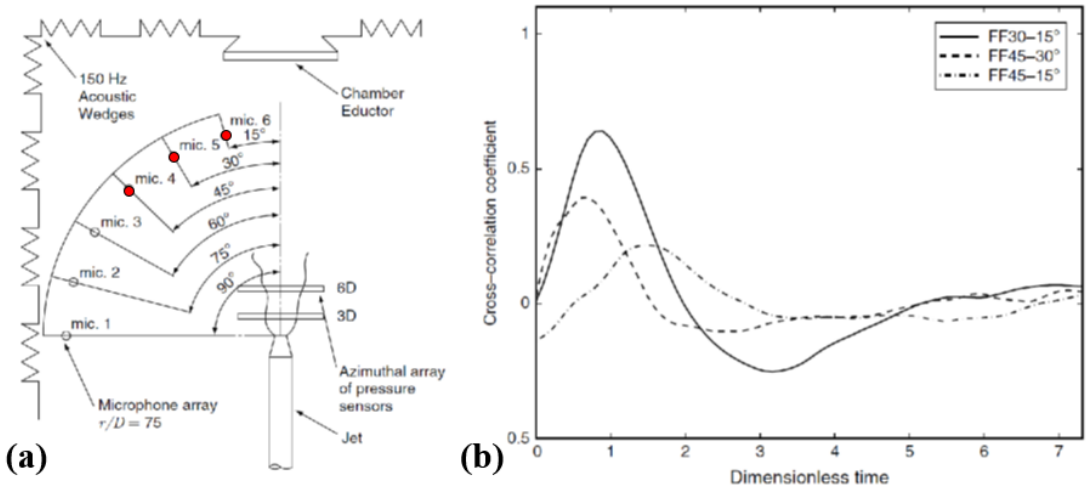


Figure 9.27: (a) Setup of the experiment (the three microphones indicated with red symbols are those used in the study); (b) Some cross-correlations between the microphone signals. From Lewalle et al (2012) [19].

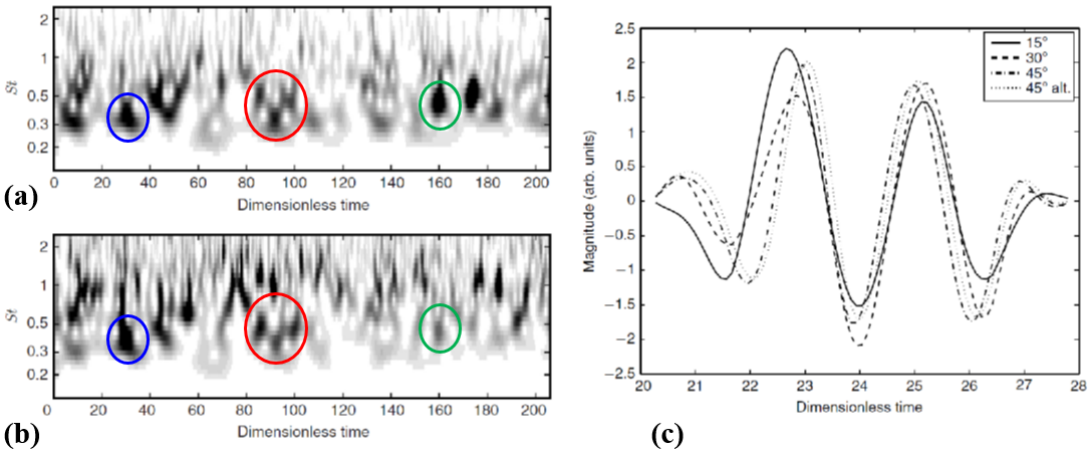


Figure 9.28: (a) Scalogram of an excerpt of the signal at the microphone at 15^o; (b) Scalogram of an excerpt of the signal at the microphone at 30^o; (c) Individual events appearing at the three microphones that can be assigned to a same burst. From Lewalle et al (2012) [19].

in an intermittent signal.

10 Proper Orthogonal Decomposition

10.1 Introduction

Large quantities of data are produced by numerical simulations or by measurements (in particular when using Particle Image Velocimetry (PIV)). There is a need to extract from these high-dimensional data a reduced set of data containing as much information as possible. Specifically, suppose one knows some velocity field $u(x,t)$ over some *bounded* domain. This is written as a sum of modes:

$$u(x,t) = \sum_{i=1}^{\infty} a_i(t)\psi_i(x) \tag{10.1}$$

The modes $\psi_i(x)$ used to expand the velocity can at this stage belong to any basis (the Fourier basis for example). Note that variables x and t are separated. Some data reduction can be obtained if it is possible to *truncate* the series (keeping only the first K terms) yet to have a correct approximation u' for u:

$$u'(x,t) = \sum_{i=1}^{K} a_i(t)\psi_i(x) \tag{10.2}$$

For this purpose, not all $\psi_i(x)$ bases are equivalent. Some bases are better than others.

An issue is to define what a good basis is. In this chapter, a good basis allows maximizing the energy in the truncated series. The energy of the initial field is E (some integral of u^2). The energy of the truncated series is $E' < E$ (E' is some integral of u'^2). An optimal basis will have a maximal retained energy E'. No other basis, for the same number of kept terms (K), will retain as much energy. Proper Orthogonal Decomposition (POD) is precisely a way to find such an **optimal** basis. It is important to note that the basis is *taylored to the data*, it is not known in advance (as it would be the case for a Fourier basis). So POD needs the data to produce its own set of orthogonal eigenfunctions, and these are sometimes called **empirical** eigenfunctions. Different data sets arising from different physical setups will provide different bases. The data can then be expanded using this set of eigenfunctions, and hopefully only a few eigenfunctions will contain most of the energy of the initial data. POD is a way of defining **Coherent Structures** as functions that on average are well correlated with the flow. POD

modes are also useful for reduced order modelling (ROM).

A Classical reference for POD in fluids is the book by Holmes et al [17]. Reviews on POD can also be found in references [12, 4, 32, 33]. There exists also a spectral version of POD, called SPOD, which is connected to other transforms and tools used in fluid mechanics [38]. POD is based on the same principles as the Karhunen-Loève decomposition. Other names related to that matter used in the literature are Singular Value Decomposition (SVD) and Principle Component Analysis (PCA).

This course is organized as follows: POD is introduced in section 10.2. In that section, both time and space variables take continuous values, and therefore we use the name "continuous POD". The definition of POD modes, the equations they satisfy, and their properties are given. The proofs in that section are not detailed because they rely on calculus of variations and on compact self-adjoint operators (see reference [17] for more mathematical details). The continous eigenproblem given in section 10.2, whose solution gives the POD modes, could be discretized in space so as to get the discrete problem that would be solved numerically. However, it is also possible to derive the discrete problem directly, just relying on linear algebra, and especially on the singular value decomposition of a matrix, as shown in Chatterjee [12]. This is the reason why singular value decomposition is reviewed in section 10.3. Then the discrete POD problem is dealt with in section 10.4. When put into the suitable matricial form, discrete POD is nothing but SVD in disguise. The snapshot method [32, 33] is presented; it is the method we will use during Lab 3. Finally, reduced order modelling is shortly adressed in section 10.5 and some applications are considered in section 10.6.

10.2 Continuous POD

10.2.1 Definition of the POD modes

In this section we consider a continuous (in space and time) function $u(x,t)$. The POD decomposition is given by:

$$u(x,t) = \sum_{i=1}^{\infty} a_i(t)\psi_i(x) \tag{10.3}$$

The **yet unknown** POD modes ψ_i need to be determined by using a data set, based on some **criterion**. Let us now specify this criterion.

Consider an ensemble $\{u^{(k)}(x), \quad k = 1, 2, ..., \mathcal{K}\}$ of flow fields over some 1D domain of finite spatial extent, $[0\ L_x]$ (the extension to several dimensions is easy). This ensemble may result from measurements made at some times t_k, in which case $\{u^{(k)}(x)\}$=$\{u(x,t_k)\}$. We are looking for some modes $\psi_i(x)$ that are **a good representation of our ensemble**. As in some of the preceding lectures, a tool for assessing the resemblance between a mode $\psi(x)$ and a function $u(x)$ is required. Whether $\psi(x)$ resembles $u(x)$ is given by the scalar product (u,ψ) of these two functions, with:

- (u,ψ) large $\Rightarrow$ u resembles ψ
- (u,ψ) small $\Rightarrow$ u does not resemble ψ

For a scalar product to be defined, u and ψ have to belong to an Hilbert space, and the space

considered is $L^2([0\ L_x])$ in the 1D case, the space of functions having finite energy. The **scalar product** is then:

$$(u,\psi) = \int_0^{L_x} u(x)\psi^*(x)\mathrm{d}x$$

This is a scalar product over the space variable. The **norm** (squared) of u is then given by:

$$||u||^2 = (u,u) = \int_0^{L_x} u(x)u^*(x)\mathrm{d}x$$

Our criterion to build the POD basis is **optimality**: we want the family $\psi_i(x)$ to resemble the data better than any other family. It means that the truncated expansion

$$u(x,t) = \sum_{i=1}^{K} a_i(t)\psi_i(x) \tag{10.4}$$

should be, on average, closer to the original data ensemble than any other expansion (ie using a different family) of the same size (K terms). A form equivalent to the latter equation applicable to any field of the original data ensemble will also be used in the following:

$$u^{(k)}(x) = \sum_{i=1}^{K} a_i^{(k)}\psi_i(x) \tag{10.5}$$

A mathematical statement of optimatily is that we want to find ψ_1 that reaches:

$$\max_{\psi} \frac{\left\langle |(u,\psi)|^2 \right\rangle}{||\psi||^2}$$

The resemblance is checked *on average*, so that an **ensemble average** $< \cdot >$ over the data ensemble is used. For the ensemble $\{u^{(k)}(x), \quad k = 1,2,...,\mathcal{K}\}$, this is an average over the samples of the set, that is:

$$< \cdot >= \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \tag{10.6}$$

This is an average over k which is equivalent to a time average if the ensemble is obtained by measuring a field at some times. Equation (10.6) defines a single mode ψ_1. On average ψ_1 will resemble the data better than any other function. It is then possible to get more modes by looking in a direction orthogonal to ψ_1. Let define ψ_2 as the function that reaches:

$$\max_{\psi,\psi\perp\psi_1} \frac{\left\langle |(u,\psi)|^2 \right\rangle}{||\psi||^2}$$

In the same fashion, let define ψ_3 as the function that reaches:

$$\max_{\psi,\psi\perp(\psi_1,\psi_2)} \frac{\left\langle |(u,\psi)|^2 \right\rangle}{||\psi||^2}$$

Infinitly many POD modes ψ_i , i=1,2,...∞ can be defined that way. If the meaning of the ψ_i is now clear, one still needs a recipe to compute them. We are not going to give the demonstration,

but a calculation of variations [17] shows that the POD modes $\psi_i(x)$ are solution to the following integral eigenvalue problem:

$$\int_0^{L_x} R(x,x')\psi_i(x')\mathrm{d}x' = \lambda_i\psi_i(x) \qquad \left(\begin{array}{c}\text{Fredholm equation.}\\ \text{Definition of POD modes.}\end{array}\right) \tag{10.7}$$

where

$$R(x,x') = \big\langle u(x)u(x')\big\rangle$$

is the **two-point velocity auto-correlation** calculated using a ensemble average. This is symmetric: $R(x,x') = R(x',x)$. Of course, $R(x,x')$ depends on the data, and so do the POD modes.

The theory of compact self-adjoint operators tells us that problem (10.7) has a countable number of eigenfunctions/POD modes $\psi_i(x)$, i=1,2,3,...∞ [17]. Each of these modes corresponds to an eigenvalue $\lambda_i \geq 0$, i=1,2,3,...∞. In the following, the eigenvalues and the corresponding modes are sorted in descending order of the eigenvalue magnitude: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ....$ The properties of the POD modes are reviewed in the next subsection.

10.2.2 Properties of the POD modes

Orthogonality of the modes:
The POD modes form an orthogonal set. They can be normalized so that their norm is unity:

$$||\psi_i||^2 = (\psi_i,\psi_i) = 1 \qquad \forall i = 1...\infty$$

and in this case, they form an orthonormal set:

$$(\psi_i,\psi_j) = \delta_{ij} \qquad \left|\begin{array}{l}\forall i = 1...\infty\\ \forall j = 1...\infty\end{array}\right. \qquad \left(\begin{array}{c}\text{orthonormality}\\ \text{of the POD modes}\end{array}\right) \tag{10.8}$$

Expansion coefficients:
The modes have been calculated so that they resemble the data. It is possible to express each of the data fields as a sum of these modes:

$$u(x,t) = \sum_{i=1}^{\infty} a_i(t)\psi_i(x) \qquad \text{or} \qquad u^{(k)}(x) = \sum_{i=1}^{\infty} a_i^{(k)}\psi_i(x)$$

The time coefficients in this expansion are obtained by projecting the velocity fields onto the modes. On taking the scalar product of the expansions in the previous equation with mode ψ_j, and on using orthonormality of the modes, one obtains:

$$a_j(t) = (u(x,t),\psi_j(x)) \quad \forall j = 1...\infty \qquad \text{or} \qquad a_j^{(k)} = \Big(u^{(k)}(x),\psi_j(x)\Big) \tag{10.9}$$

Uncorrelatedness of the expansion coefficients:
It can be proved that the expansion coefficients are uncorrelated with respect to the ensemble average:

$$\langle a_i a_j \rangle = \lambda_i \delta_{ij} \tag{10.10}$$

To be more specific, this can be written:

$$\langle a_i a_j \rangle = \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} a_i^{(k)} a_j^{(k)}$$

If the fields result from measuring the velocity at several times, the ensemble average is the time average.

Energy of the modes:
The eigenvalue λ_i, which we recall is positive, is half the energy contained in the projection of the field u on mode ψ_i. To be more specific, defining the averaged energy by:

$$E = \frac{1}{2} \langle (u,u) \rangle = \frac{1}{2} \langle ||u||^2 \rangle = \frac{1}{2} \left\langle \int_0^{L_x} |u(x,t)|^2 \mathrm{d}x \right\rangle \tag{10.11}$$

we have:

$$E = \frac{1}{2} \sum_{i=1}^{\infty} \lambda_i \tag{10.12}$$

This equation says that the averaged energy of the signal is the sum of the energies contained in the modes. The result is easily obtained by injecting the POD decomposition for u in the definition of energy and taking into account the orthogonality in space of the modes $\psi_i(x)$ as well as the uncorrelatedness of the time coefficients $a_i(t)$.

Two-point velocity auto-correlation:
The two-point velocity auto-correlation that has been formed for calculating the modes through Eq. (10.7) can be expressed using the modes and eigenvalues. This is known as the Mercer theorem:

$$R(x,x') = \sum_{i=1}^{\infty} \lambda_i \psi_i(x) \psi_i^*(x') \qquad \text{(Mercer's theorem)} \tag{10.13}$$

Properties shared by the modes and the data fields:
The velocity fields are a linear sum of POD modes; this is actually the POD expansion:

$$u(x,t) = \sum_{i=1}^{\infty} a_i(t) \psi_i(x) \qquad \text{or} \qquad u^{(k)}(x) = \sum_{i=1}^{\infty} a_i^{(k)} \psi_i(x)$$

The converse is true as well: for any mode ψ_i there exists $b_i^{(k)}$, k=1,2,...∞, such that:

$$\psi_i(x) = \sum_{k=1}^{\infty} b_i^{(k)} u^{(k)}(x) \qquad \forall i$$

In words, a mode is a sum of flow fields. As a result, any property shared by the samples $u^{(k)}$ will be shared by the POD modes $\psi_i(x)$. For example, if the $u^{(k)}$ all satisfy some boundary conditions, then the ψ_i satisfy the same boundary conditions. Or, if the $u^{(k)}$ are all incompressible (div($u^{(k)}$)=0) then so are the POD modes (div($\psi_i(x)$)=0).

10.2.3 Snapshot method (discrete times, continuous space variable)

The snapshot method relies on the property mentioned previously that any mode is a sum of flow fields:

$$\psi_i(x) = \sum_{k=1}^{N} b_i^{(k)} u(x, t_k) \qquad \forall i \tag{10.14}$$

where it is supposed that $\mathcal{K}$ measurements of the flow have been performed at times $t_k, k = 1...\mathcal{K}$. By introducing this expansion in the Fredholm equation, Eq. (10.7), and defining the time average as:

$$\langle\rangle = \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}}$$

it is found that the vector $\mathbf{b_i} = [b_i^{(1)} b_i^{(2)} ... b_i^{(\mathcal{K})}]$ of the time coefficients for a given mode should verify:

$$C\mathbf{b} = \lambda \mathbf{b} \tag{10.15}$$

where C is a matrix of size $\mathcal{K}$ x $\mathcal{K}$ with coefficients

$$C_{ij} = (u(x, t_i), u(x, t_j))$$

A matrix eigenvalue problem has then to be solved in order to find $\mathcal{K}$ different eigenvectors $\mathbf{b_i}$, $i = 1...\mathcal{K}$, containing the time coefficients. From these, $\mathcal{K}$ spatial POD modes ψ_i, $i = 1...\mathcal{K}$, may be calculated subsequently using Eq. (10.14). Hence, the snapshot method consists of solving a matrix eigenvalue problem for some time coefficients, rather than the Fredholm equations for the spatial modes.

10.2.4 Case of several components

One may be interested in a multi-component flow field, as when two components of the velocity are needed. In that case, one may write:

$$u(x, t) = \begin{bmatrix} u^{(1)}(x, t) \\ u^{(2)}(x, t) \end{bmatrix} \tag{10.16}$$

The POD modes are then also multi-component, with:

$$\psi_i(x, t) = \begin{bmatrix} \psi_i^{(1)}(x, t) \\ \psi_i^{(2)}(x, t) \end{bmatrix} \tag{10.17}$$

The (spatial) scalar product between $u(x, t)$ and $\psi(x, t)$ is given by:

$$(u, \psi) = \int_0^{L_x} u^{(1)}(x)\psi^{(1)*}(x) + u^{(2)}(x)\psi^{(2)*}(x)\mathrm{d}x \tag{10.18}$$

The correlation is now a tensor obtained using the tensorial product:

$$R(x, x') = \left\langle u(x) \otimes u(x') \right\rangle = \begin{bmatrix} \left\langle u^{(1)}(x)u^{(1)}(x') \right\rangle & \left\langle u^{(1)}(x)u^{(2)}(x') \right\rangle \\ \left\langle u^{(2)}(x)u^{(1)}(x') \right\rangle & \left\langle u^{(2)}(x)u^{(2)}(x') \right\rangle \end{bmatrix} \tag{10.19}$$

As before, the POD modes are solution to a Fredholm problem:

$$\int_0^{L_x} R(x,x')\psi_i(x')\mathrm{d}x' = \lambda_i\psi_i(x)$$

The extension to several space dimensions (ie (x,y,z)) is also straightforward.

10.2.5 Relation to Fourier modes

When turbulence is homogeneous (in the x-direction), the correlation depends on the distance between two points, but not on the particular choice of the points, that is:

$$R(x,x') = R(x-x')$$

For example, the axial and spanwise directions in a channel flow are homogeneous.

For a finite length L_x, $R(x-x')$ can be expanded into a Fourier series:

$$R(x-x') = \sum_{i=1}^{\infty} c_i e^{\mathrm{j}ik(x-x')}$$

where $k = \frac{2\pi}{L_x}$ is the wavenumber and $\mathrm{j} = \sqrt{-1}$. This can also be written:

$$R(x-x') = \sum_{i=1}^{\infty} c_i e^{\mathrm{j}ikx}\left(e^{\mathrm{j}ikx'}\right)^*$$

On comparing this result with that given by the Mercer theorem, see Eq. (10.13), one obtains:

$$\lambda_i = c_i \qquad \psi_i(x) = e^{\mathrm{j}ikx}$$

Hence, for $R(x,x')$ homogenous in one direction, the POD modes are the Fourier modes in that direction. POD may be seen as an extension of the Fourier transform to non-homogeneous directions. POD is useful mainly in non-homogenous cases (otherwise, use Fourier modes).

10.3 Singular Value Decomposition (SVD)

The discrete version of POD will be shown to be equivalent to Singular Value Decomposition (SVD). This is the reason why we introduce the singular value decomposition of a matrix in the present section. The link between SVD and data reduction is also explained.

10.3.1 Definition of SVD

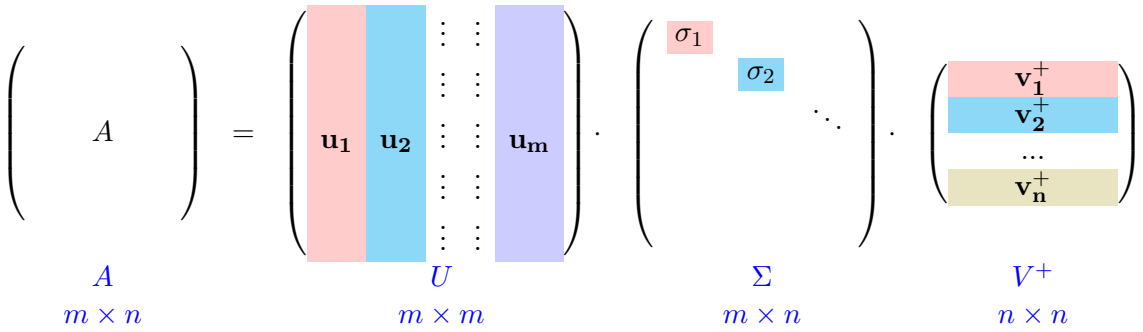
SVD is a matrix factorization technique that can be applied to **any** matrix (including non-square matrices, non-normal matrices, and singular matrices) and is more general than diagonalization (some fundamentals about matrices are summarized in Appendix D). In particular, a matrix may not be diagonalizable and yet have a SVD. Let A be a complex matrix such that:

$$A : C^n \to C^m$$

The size of A is then m x n. The SVD of A reads:

$$A = U\Sigma V^{+} \tag{10.20}$$

where U and V are unitary matrices, and Σ has non-zero entries only in its main diagonal. V^{+} denotes the adjoint (conjugate transpose) of V. More details about these matrices are now given. The shape of the different matrices involved is made clearer by the following expression:



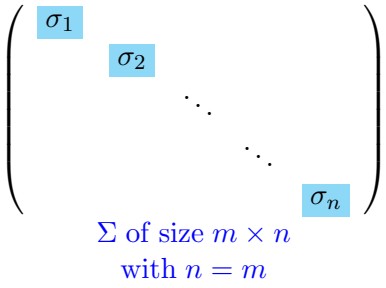
(10.21)

Matrix Σ: Σ is a matrix of size m x n (the same size as A). Its only non-zero entries are real and non-negative numbers $\sigma_i \geq 0$ located on the main diagonal. These are called the **singular values**. The singular values are sorted in descending order of magnitude:

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq ... \geq 0$$

Using row/column permutations it is always possible to arrange U, Σ, and V so that this is the case. The number of non-zero singular values is the **rank** of A and it is at most min(m,n).

A diagonal matrix has to be square and Σ is not necessarily square. It is square when n=m, and in that case it is diagonal:



(10.22)

Its rank is then at most m=n. Some singular values σ_i can be null and the rank is then less than m=n.
For the cases $n < m$ and $n > m$, Σ has the following forms (it cannot be called a diagonal

matrix because it is not square):

$$\underbrace{\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \end{pmatrix}}_{\Sigma \text{ of size } m \times n \text{ with } n < m} \qquad \underbrace{\begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m \end{pmatrix}}_{\Sigma \text{ of size } m \times n \text{ with } n > m} \tag{10.23}$$

For $n < m$, there are n singular values, the rank is at most n, and it is less than n when some singular values are null. For $n > m$, there are m singular values, the rank is at most m, and it is less that m if some singular values are null.

Matrix U: U is a unitary matrix of size m x m. Unitary means: $UU^+ = U^+U = I$. That means that U has m **orthonormal columns**, $\mathbf{u_1}$, $\mathbf{u_2}$, ... $\mathbf{u_m}$, which are called the **left-singular vectors**. This set of vectors forms a basis for the arrival space ($F = C^m$). Note: when A is real, U is real and it is orthonormal (U^+ simply represents the transpose U^t of U in that case).

Matrix V: V is a unitary matrix of size n x n, with again: $VV^+ = V^+V = I$. V has n **orthonormal columns**, $\mathbf{v_1}$, $\mathbf{v_2}$,...$\mathbf{v_n}$, which are called the **right-singular vectors**. These form a basis for the departure space ($E = C^n$). Note: when A is real, V is real and it is orthonormal (V^+ simply represents the transpose V^t of V in that case).

The meaning of the SVD decomposition is better understood when written in the form (obtained by multiplying Eq. (10.20) by V on the right):

$$AV = U\Sigma \tag{10.24}$$

For the case n=m (the departure space and arrival space have the same dimension, that is, the same number of basis vectors $\mathbf{v_i}$ and of basis vectors $\mathbf{u_i}$), this can be written:

$$\begin{pmatrix} A\mathbf{v_1} & A\mathbf{v_2} & \vdots & \vdots & A\mathbf{v_n} \end{pmatrix} = \begin{pmatrix} \sigma_1\mathbf{u_1} & \sigma_2\mathbf{u_2} & \vdots & \vdots & \sigma_n\mathbf{u_n} \end{pmatrix} \tag{10.25}$$

Hence, the orthonormal basis vectors of the departure space (the $\mathbf{v_i}$) and the orthonormal basis vectors of the arrival space (the $\mathbf{u_i}$) satisfy:

$$A\mathbf{v_i} = \sigma_i\mathbf{u_i} \tag{10.26}$$

Under the action of the matrix A, a vector that is aligned with $\mathbf{u_i}$ is thus stretched by σ_i and rotated so that is becomes aligned with $\mathbf{v_i}$. Note: some σ_i may be null; in that case a vector aligned with $\mathbf{v_i}$ becomes the null vector. Indeed, the vectors $\mathbf{v_i}$ with σ_i=0 form a basis of the

null space of the matrix A.

Example: $A : C^3 \to C^3$, and only two singular values are not null (σ_3=0). The rank of A is then 2. We have:
$A\mathbf{v_1} = \sigma_1\mathbf{u_1}$
$A\mathbf{v_2} = \sigma_2\mathbf{u_2}$
$A\mathbf{v_3} = \mathbf{0}$
This is illustrated in Fig. 10.1.

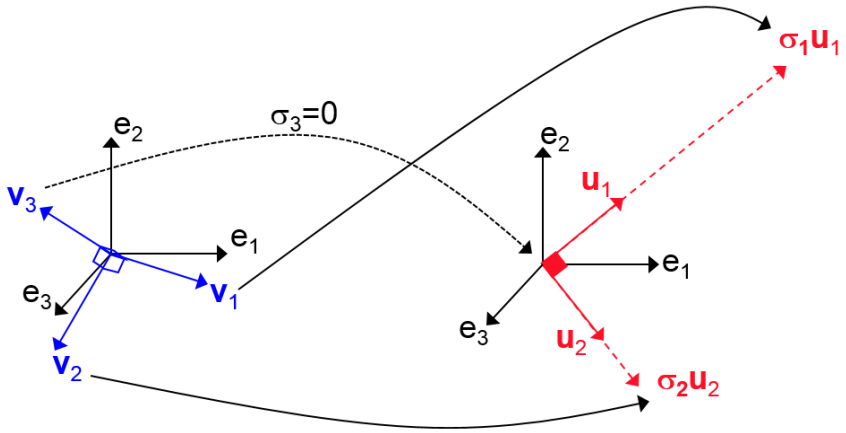


Figure 10.1: Image of the right-singular vectors of the matrix A of size 3x3 and of rank 2 (the third singular value is null: σ_3=0).

We have considered the case n=m, but the cases $n > m$ and $n < m$ are not much different. For example, for the case $n > m$ (the departure space has a higher dimension than the arrival space, that is, there are more vectors $\mathbf{v_i}$ than vectors $\mathbf{u_i}$) we have:

$$\begin{pmatrix} A\mathbf{v_1} & A\mathbf{v_2} & \vdots & \vdots & A\mathbf{v_m} & A\mathbf{v_{m+1}} & \vdots & A\mathbf{v_n} \end{pmatrix} = \begin{pmatrix} \sigma_1\mathbf{u_1} & \sigma_2\mathbf{u_2} & \vdots & \vdots & \sigma_m\mathbf{u_m} & 0 & \vdots & 0 \end{pmatrix} \tag{10.27}$$

The last $n - m$ right-singular vector $\mathbf{v_i}$, i=$m + 1...n$, belong to the null space of A and their images by A are the null vector. Some values of σ_i, $i \leq m$ can be null.

10.3.2 Relation between matrix diagonalization and SVD

In general a matrix A cannot be diagonalized. However, the two square matrices AA^+ and A^+A are normal (a matrix N is normal whenever NN^+=N^+N). The finite-dimensional spectral theorem then warrants that these matrices can be diagonalized (see Appendix D.2.2). This is a good reason for looking at AA^+ and A^+A. Using the SVD of A, we get:

$$A = U\Sigma V^+ \Leftrightarrow A^+ = V^{++}\Sigma^+U^+ = V\Sigma U^+$$

Hence,

$$AA^+ = U\Sigma \underbrace{V^+V}_{=I} \Sigma U^+ = U\Sigma^2 U^+ = U\Sigma^2 U^{-1} \tag{10.28}$$

where $U^+ = U^{-1}$ since U is unitary. Similarly, we have:

$$A^+A = V\Sigma^2 V^+ = V\Sigma^2 V^{-1} \tag{10.29}$$

The former two relations are in the form: unitary x diagonal x unitary inverse. Hence, these relations express nothing but the diagonalization of matrices AA^+ and A^+A (compare with Eq. (D.5)). As a result:
- the eigenvalues of AA^+ or A^+A are nothing but the squares of the singular values of A.
- The left-singular vectors of A (the $\mathbf{u_i}$s) are of the eigenvectors of AA^+.
- The right-singular vectors of A (the $\mathbf{v_i}$s) are the eigenvectors of A^+A.
One should remember:

SVD of A	$\Leftrightarrow$	Diagonalization of AA^+ or A^+A

Note: we will see that calculating the POD modes is equivalent to obtaining U or V, and there will thus be two ways of getting them: SVD of the matrix A containing the data, or diagonalization of the correlation matrix AA^+ (or A^+A) of the data.

10.3.3 Data reduction with SVD

SVD may be used to compute an **approximation** A' of A. The SVD gives: $A = U\Sigma V^+$. Let us consider the diagonal matrix Σ' obtained by setting the smallest singular value of Σ to zero. Σ' contains only the **first** r **largest singular values of** A. Then the matrix

$$A' = U\Sigma' V^+$$

is an approximation of the matrix A with rank r. The Eckart-Young theorem tells us that, among all possible matrices having rank r, A' is the approximation that minimizes the norm $||A - A'||_F$. The norm considered here is the **Frobenius norm**. This is the square root of the sum of the entries squared of a matrix:

$$||A||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{i,j}|^2} = \sqrt{\text{trace}(A^+A)} = \sqrt{\sum_{i=1}^{\min(m,n)} \sigma_i^2} \tag{10.30}$$

To be specific, the norm of the difference between A and its approximation A' (said otherwise, the distance from A' to A) is:

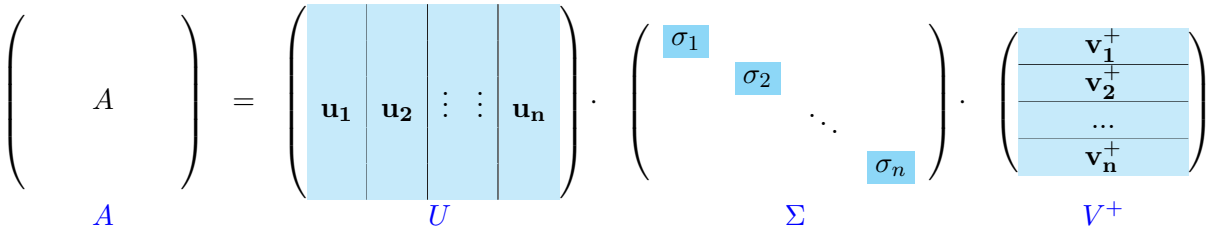
$$||A - A'||_F = \sqrt{\sum_{i=r+1}^{\min(m,n)} \sigma_i^2}$$

Of course, this norm should be as small as possible if A' is to be a good approximation of A. The distance between the matrix and its approximation depends on the values of the neglected

singular values. If the singular values decrease rapidly, one may hope to find a good approximation with a low rank.

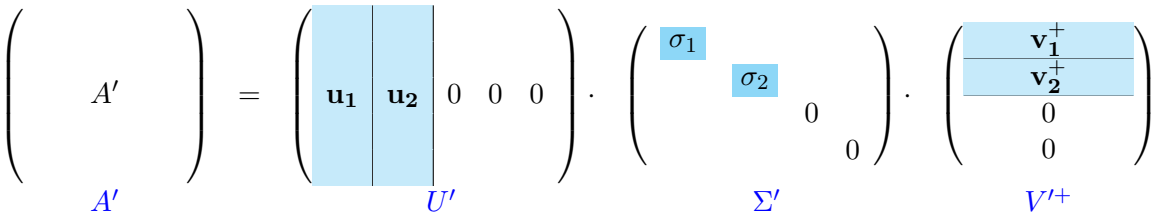
The link between SVD and POD start to show up: in the POD, we want to represent a signal with a truncated expansion (a reduced number of modes), while retaining as much energy as possible. Here, we want to approximate a matrix using only a reduced number of singular values, while keeping the distance between the matrix and its approximation at a low value (which is equivalent to retaining as much as possible the energy in the original matrix, the square of the norm being a measure of energy).

Let us give a graphical illustation. The original matrix A is given by (we take n=m for simplicity):



(10.31)

We would like to have the best rank 2 approximation A' for this matrix. "Best" is in the sense of minimizing the Frobenius norm between A and A'. Knowing the SVD of A, we just have to keep 2 singular values of Σ to build Σ'. It is also possible to keep only the first two left-singular vectors and the first two right-singular vectors to build U' and V'. A' is then obtained by:



(10.32)

Only the data within the remaining blue-colored boxes are used in the calculation of A'. In practice, only these data would be stored on a computer, and this would reduce the quantities of data stored when r is small enough. Calculating an approximation is particularly easy when one has a routine to calculate the SVD. This is only a few lines in a program.

Example: The image in Fig. 10.2(a) is a 300*200 matrix containing integers (levels of gray). The SVD returns 200 singular values (so the initial matrix has at most rank 200, and one can check that it is indeed 200). To be able to produce a good approximation of the matrix by keeping only the first r singular values, a fast enough decay of the singular values needs to be observed. The singular values are plotted in Fig. 10.2(b). They should always be plotted! A flat curve would mean that all singular values are equal and looking for a low-rank approximation would be difficult. On the other hand, if the singular values decay rapidly, then by keeping only the first few, a good approximation is possible. Some approximations with rank r are represented in Fig. 10.3 for several values of r. In this figure, E'/E is the ratio of the energy

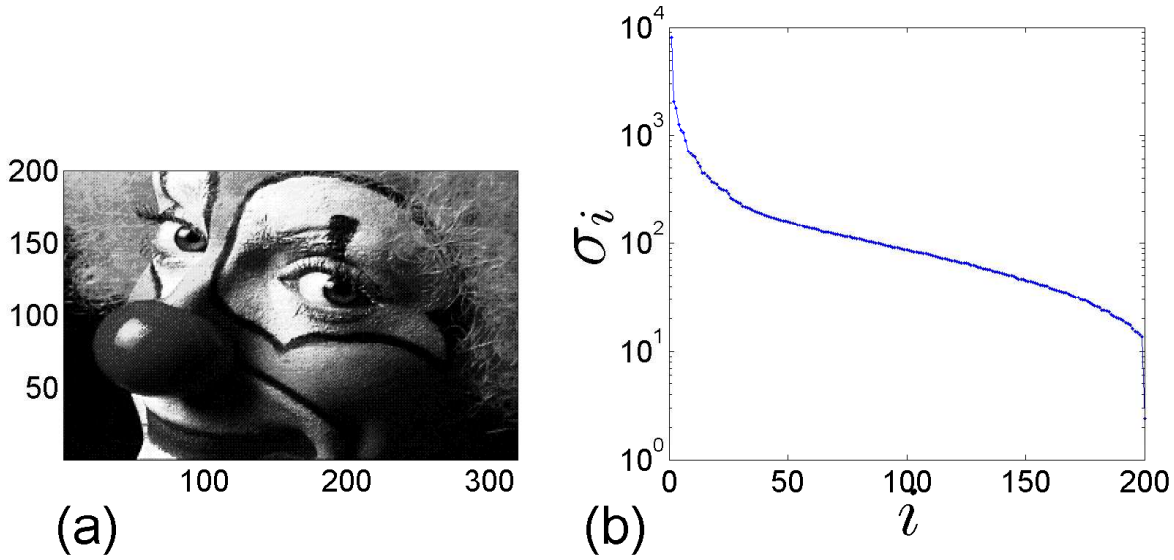


Figure 10.2: (a) Original image; (b) its singular values.

of the approximant to that of the original data, the energy being calculated with the Frobenius norm. The objective of POD is not much different from the one in this figure. Of course, we will apply POD to a fluid velocity or pressure field rather than to the image of a clown. The discrete POD is presented in the next section, and SVD will naturally show up.

Exercise: SVD with Matlab

Define the following matrix in Matlab, and use Matlab to answer the questions. Note: a list of Matlab commands, such as `svd` and `eig`, is available in Appendix B.

$$A = \begin{pmatrix} 5 & 3-\mathrm{j} & 2 \\ 3+\mathrm{j} & 1 & 4+5\mathrm{j} \\ 2 & 4-5\mathrm{j} & 9 \end{pmatrix} \tag{10.33}$$

Is A a normal matrix (compute $AA^+ - A^+A$) ? Did you expect this?
What are the singular values of A? Are they real?
What are the eigenvalues of AA^+? How are they related to the singular values of A?
How can you check that the columns of the left-singular matrix (U) are orthonormal?
What is the first right-singular vector, $\mathbf{v_1}$?
Check that $A\mathbf{v_1} = \sigma_1\mathbf{u_1}$, where $\mathbf{u_1}$ is the first left-singular vector.
What is the rank 2 approximation A' of A? Check the Frobenius norm of $A - A'$ using the `norm` command. To which singular value of A should this be compared to?

10.4 Discrete POD

In this section, discrete POD is considered, that is, the POD of data obtained at discrete spatial positions and discrete instants is performed. We could start directly from the Fredholm

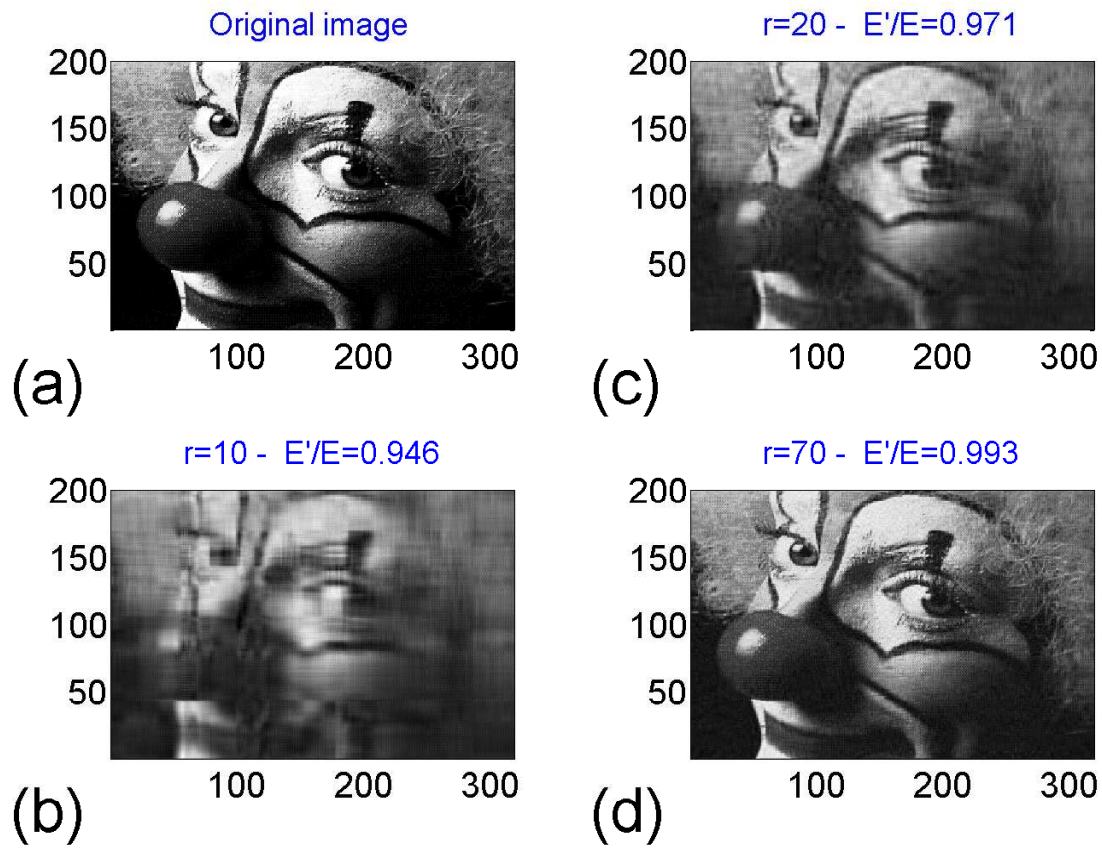


Figure 10.3: (a) original image; (b) rank 10 approximation; (c) rank 20 approximation; (d) rank 70 approximation.

eigenvalue problem in Eq. (10.7), or from the snapshot method in Eq. (10.15), and discretize the equations. However, we prefer to start from the very beginning again, that is, from the modal decomposition given in Eq. (10.4). This modal decomposition, written at discrete positions and times, together with the requirement that the basis be optimal for representing the data, form the discrete POD problem. This problem boils down to performing the SVD of a data matrix, or equivalently, the diagonalisation of the correlation matrix. The so-called snapshot method is presented and will be used during Lab 3. This presentation follows that of Chatterjee [12].

10.4.1 Discrete POD and SVD

We now suppose that we have acquired some data:

- at a finite number M of spatial positions: x_1,x_2,...,x_M
- at a finite number N of instants : t_1,t_2,..., t_N

Note that N is equivalent to $\mathcal{K}$ is section 10.2. The data may typically result from PIV measurements or be the output of some numerical solver.

The data may be stored in a **data matrix** A of size M x N, whose coefficients are given by:

$$A_{mn} = u(x_m, t_n) \quad \left| \begin{array}{l} \forall m = 1...M \\ \forall n = 1...N \end{array} \right. \tag{10.34}$$

The matrix is then:

$$A = \begin{pmatrix} u(x_1,t_1) & u(x_1,t_2) & \cdots & \cdots & u(x_1,t_N) \\ u(x_2,t_1) & u(x_2,t_2) & \cdots & \cdots & u(x_2,t_N) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ u(x_M,t_1) & u(x_M,t_2) & \cdots & \cdots & u(x_M,t_N) \end{pmatrix} \begin{array}{l} \\ \text{given position} \\ \\ \\ \\ \end{array} \tag{10.35}$$

given time

A column of A contains the data for all positions at a given time and corresponds to a snapshot/image. A line of A contains the time series obtained at a given position/sensor.

For discrete times and positions, the proper orthogonal decomposition (10.4) can be written:

$$u(x_m, t_n) = \sum_{i=1}^{\mathrm{N}} a_i(t_n)\psi_i(x_m) \quad \left| \begin{array}{l} \forall m = 1...M \\ \forall n = 1...N \end{array} \right. \tag{10.36}$$

Since $u(x_m, t_n)$ is nothing but the data matrix element A_{mn}, we have:

$$A_{mn} = \sum_{i=1}^{\mathrm{N}} a_i(t_n)\psi_i(x_m) \quad \left| \begin{array}{l} \forall m = 1...M \\ \forall n = 1...N \end{array} \right. \tag{10.37}$$

Let's define the following two matrices:

$$Q = \begin{pmatrix} a_1(t_1) & \cdots & \cdots & a_1(t_N) \\ a_2(t_1) & \cdots & \cdots & a_2(t_N) \\ \vdots & \vdots & \vdots & \vdots \\ a_N(t_1) & \cdots & \cdots & a_N(t_N) \end{pmatrix} \qquad \Psi = \begin{pmatrix} \psi_1(x_1) & \cdots & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \cdots & \cdots & \psi_N(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_1(x_M) & \cdots & \cdots & \psi_N(x_M) \end{pmatrix} \tag{10.38}$$

time coefficient matrix of size N x N

POD modes matrix of size M x N

Q is the N x N matrix whose k-th row contains the time coefficients for the k-th mode; Ψ is the M x N matrix whose k-th column contains the k-th POD mode sampled on the grid x_1, x_2, ...x_M. Equation (10.37) can then be written in the following matricial form:

$$A = \Psi Q \tag{10.39}$$

The data matrix is the product of the modes matrix with the time coefficients matrix. The later equation is nothing but the discrete version of the POD, that is, the discrete version of Eq. (10.4). Note that the time and space variables are separated in the POD, which is seen in the matricial product as well. Now, A is known because the data are known, and the objective

is to find the matrices Ψ and Q. In addition, it is required that Ψ be optimal.

This is where **SVD** is useful. The SVD of the data matrix is:

$$A = U\Sigma V^+$$

where it is assumed that the singular values and singular vectors are arranged so that: $\sigma_1 \geq \sigma_2 \geq \sigma_3$.... By comparing with Eq. (10.39), we get:

$$\left\{ \begin{array}{ll} \Psi = & U \\ Q = & \Sigma V^+ \end{array} \right. \tag{10.40}$$

The (orthonormal) POD modes are thus the left-singular vectors of A, with:

$$\Psi = \left(\begin{array}{c|c|c|c} \psi_1(x_1) & \cdots & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \cdots & \cdots & \psi_N(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_1(x_M) & \cdots & \cdots & \psi_N(x_M) \end{array} \right) = \left(\begin{array}{c|c|c|c} \vdots & & & \vdots \\ \mathbf{u_1} & \cdots & \cdots & \mathbf{u_N} \\ \vdots & & & \vdots \end{array} \right) \tag{10.41}$$

Optimality is automatically satisfied, in the sense that the Frobenius norm of the difference between the data matrix A and some approximation of it obtained by keeping only the first r singular values and singular vectors is smaller than for any other approximation of rank r.

> *Note: you may have noticed that the matrix U was given to be of size $M \times M$, while Ψ is of size $M \times N$. This is not worrying. For $M > N$ you can add $M - N + 1$ orthonormal columns to the matrix Ψ (these are $\mathbf{u_{N+1}}$, $\mathbf{u_{N+2}}$, ...,$\mathbf{u_M}$) so that the augmented matrix has size $M \times M$. However, these additional columns correspond to null singular values and do not account for any energy in the data, so that they are actually useless. For $M < N$, the last $N - M + 1$ columns of Ψ are filled with zero (the vectors $\mathbf{u_{M+1}}$, $\mathbf{u_{M+1}}$,...,$\mathbf{u_N}$ do not even exist and you may imagine they are some null vectors) so that again the useful part of the matrix has size $M \times M$.*

Equation (10.40) means that computing the SVD of a matrix (written in the data matrix form) is equivalent to computing the discrete POD of the data. We have seen that computing the SVD of the matrix A is equivalent to performing the diagonalization of either AA^+ or A^+A. These matrices are called correlation matrices. It is worth remembering the following:

discrete POD of the data	$\Leftrightarrow$	SVD of data matrix A	$\Leftrightarrow$	diagonalization of AA^+ or A^+A

When A has size M x N, AA^+ has size M x M, and A^+A has size N x N. Hence, when $M << N$, one would diagonalize AA^+, while one would diagonalize A^+A when $M >> N$. The later case is known as the snapshot method and is now described.

10.4.2 Discrete snapshot method

It is recalled that M is the number of spatial positions and N is the number of recorded times. The snapshot method is used when $M >> N$. This is typically the case when one measures

a velocity field at N=800 instants and for M=1000 x 1000 pixels using PIV. In that case one performs the diagonalization of A^+A, which is of size N x N. The full steps of the snapshot method are the following (as will be put into practice in Lab 3):

step 1. Store the data in the data matrix A, of size M x N.
step 2. Compute the correlation matrix A^+A of size N x N.
step 3. Diagonalize this correlation matrix. The N eigenvalues λ_i are the squares of the singular values of A: $\lambda_i = \sigma_i^2$. Up to a normalizing factor, the corresponding N eigenvectors are the right-singular vectors of the matrix A (the columns of V, or rows of Q^+). Let's store them in the columns of the matrix V.
step 4. Compute $\Psi = AV$. Since $AV = U\Sigma = \Psi\Sigma$, the columns of Ψ are proportional to the POD modes. That is, the columns of Ψ are orthogonal, but not orthonormal. Hence, normalize each columns of Ψ, by dividing each element in a column by the norm of the column. The resulting matrix Ψ, of size M x N, contains the POD modes in its columns.
step 5. Now compute the time coefficients by $Q = \Psi^+A$ (this is Eq. (10.39)). This is a projection of the data matrix onto the modes to get the time coefficients of the modes. This is equivalent to Eq. (10.9) for the continuous POD.

After completing these steps, one has obtained:
- a M x N POD mode matrix Ψ, containing the spatial POD modes in its columns. Totally, there are N modes in the snapshot method (as many modes as recording times). Each mode is given at M stations (the recording positions, or eventually the number of recording positions multiplied by the number of components that are measured).
- a N x N time coefficients matrix Q containing the time coefficients of the modes in its rows.
- N eigenvalues, one per mode, that give the energy contained in the modes.

Ideally, one hopes that only a few λ_i will take large values, because this warrants that a low-rank approximation for the data can be reached. A low rank approximation of the data matrix is easily produced. Let A' be this low-rank approximation. It is given by:

$$A' = \Psi' Q'$$

where Ψ' is obtained by nullifying the last columns of Ψ, and where Q' is obtained by nullifying the corresponding last rows in Q. Of course, it is assumed that the eigenvalues are sorted from large to small, and the suppressed columns and rows correspond to the smallest values of λ_i.

The snapshot method presented here (for discrete positions and time) is the discrete equivalent of the snapshot method presented in section 10.2.3. In that section, we had to solve $C\mathbf{b} = \lambda\mathbf{b}$ where $\mathbf{b}$ represents a vector of time coefficients that is equivalent to a column of Q^+ in the present approach. The matrix C was a correlation matrix with elements $C_{ij} = (u(x,t_i), u(x,t_j))$ representing the spatial scalar product between two fields obtained at two different times. The equivalent matrix is here A^+A whose elements $(A^+A)_{ij}$ verify:

$$(A^+A)_{ij} = \sum_{m=1}^{M} A^+_{im} A_{mj} = \sum_{m=1}^{M} A^*_{mi} A_{mj} = \sum_{m=1}^{M} u(x_m,t_i)^* u(x_m,t_j) \qquad (10.42)$$

The sum is performed over spatial positions, and the latter equation is nothing but a discrete

approximation for the scalar product (up to a normalizing factor), so that $(A^+A)_{ij}$ is a discrete approximation of C_{ij}.

10.5 Reduced Order Modelling (ROM)

It is possible to project the equations of the flow onto the POD modes, which is known as a Galerkin method. The **partial** differential equations are then transformed into **ordinary** differential equations (ODE) for the time coefficients $a_i(t)$. These allow describing the dynamics of the flow using only a few modes, and a control can also be introduced in the model. Note that the Galerkin projection can be performed for any set of orthogonal modes, but the POD modes are optimal, ie they allow to retain a maximal energy in the reduced order model.

Suppose the governing equation is $L[u(x,t)]$=0. For the Navier-Stokes operator, this is:

$$L[u(x,t)] = \frac{\partial u}{\partial t} + u \cdot \nabla u + \nabla p - \nu \nabla^2 u = 0 \tag{10.43}$$

The projection of $L[u(x,t)] = 0$ onto the first N POD modes is obtained by making use of the scalar product:

$$(L[u(x,t)], \psi_i(x)) = 0 \qquad \forall i = 1...N \tag{10.44}$$

The proper orthogonal decomposition for $u(x,t)$ is:

$$u(x,t) = \sum_{i=1}^{\infty} a_i(t)\psi_i(x) \tag{10.45}$$

By injecting this decomposition into Eq. (10.44), using the expression (10.43) for L, and making use of the orthogonality of the modes ψ_i, one obtains:

$$\frac{\mathrm{d}a_i}{\mathrm{d}t} = \sum_{j,k=0}^{N} q_{ijk} a_j a_k + \nu \sum_{j=0}^{N} n_{ij} a_j + P_i \qquad \forall i = 1...N \tag{10.46}$$

with:

$$\begin{array}{ll} q_{ijk} = (\psi_i, \psi_j \cdot \nabla \psi_k) & \text{(convection term)} \\ n_{ij} = (\psi_i, \nabla^2 \psi_j) & \text{(viscous term)} \\ P_i = (\psi_i, \nabla p) & \text{(pressure term)} \end{array} \tag{10.47}$$

The first two terms may be calculated when the modes are known, since they are integrals over space of the modes and their spatial derivatives. The pressure term is more difficult to deal with; it may be reduced to a boundary term that is sometimes negligible. A set of N coupled nonlinear (quadratic) ordinary differential equations in time is obtained. The spatial dependence has been removed using the scalar product with the modes. Solving this set provides **the dynamics of the most energetic structures of the flow**.

10.6 Applications in fluid mechanics

In this section, a typical application is considered that shows how the most coherent part of a flow can be isolated by using POD.

A typical particle image velocimetry (PIV) experiment is shown in Fig. 10.4. A flow at

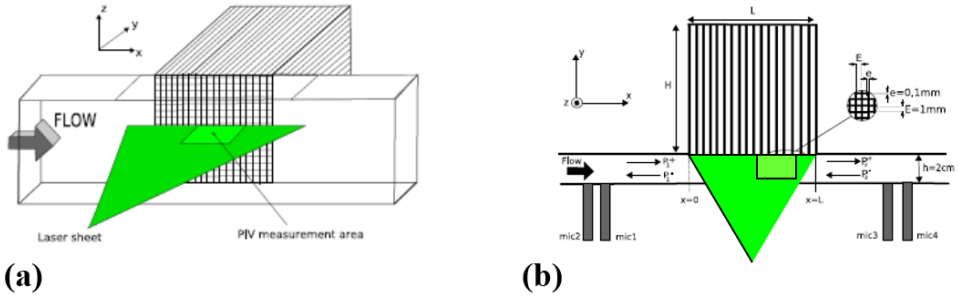


Figure 10.4: Experimental setup: (a) rectangular channel, laser sheet, and material; (b) Side view: the flow velocity field is measured within the small rectangular box. From Marx et al [22].

Mach 0.3 is set up in a rectangular channel (cross-section 2 cm x 8 cm) whose one wall is partly covered with some material (the walls are rigid otherwise). An acoustic wave at frequency 1 kHz is traveling from upstream (where it is excited by a loudspeaker mounted on the channel) and triggers an instability which is due to the material. That some instability is taking place is guessed from the large values of the pressure recorded by the microphones located downstream of the material. The PIV experiment is performed in order to confirm this intuition. A laser sheet is set up perpendicularly to the material, and the PIV technique provides 2D velocity fields of the flow within the light sheet.

The velocity measurements are made at about 10 Hz (a traditional low-rate PIV system is used) and the acquisition is not phased-locked with any sensor (or with the signal exciting the loudspeaker). Hence, the velocity snapshots are measured **at random**. A total of 1000 snapshots are measured but only N=300 are used for the POD. The flow field is acquired at N_x=200 spatial positions in the x-direction and N_y=200 spatial positions in the y-direction. Two components of the velocity, N_c=2, are measured (these are the velocity u in the x-direction and the velocity v in the y-direction). Hence, the effective number of "spatial" measurements at a given time is: M=N_cxN_xxN_y=2x200x200=80000. The data matrix has size M x N. Since $N << M$, the snapshot method is used in this case to calculate the POD modes. It is recalled that the method relies on the diagonalization of the matrix A^+A of size N x N. This provides N POD modes ψ_i (with components $\psi_{u,i}$ and $\psi_{v,i}$), and N eigenvalues λ_i.

First, some instantaneous velocity fields are shown in Fig. 10.5. The top row corresponds to the u component, the bottom row to the v component. These fields correspond to random instants. The ensemble average calculated over all the snapshots has been subtracted to obtain only the fluctuating part. One can guess that some large structures are present within these fields. The POD is then performed. Again, the ensemble-averaged flow is subtracted from all the snapshots so that the POD modes correspond to the fluctuating part of the flow. If this was

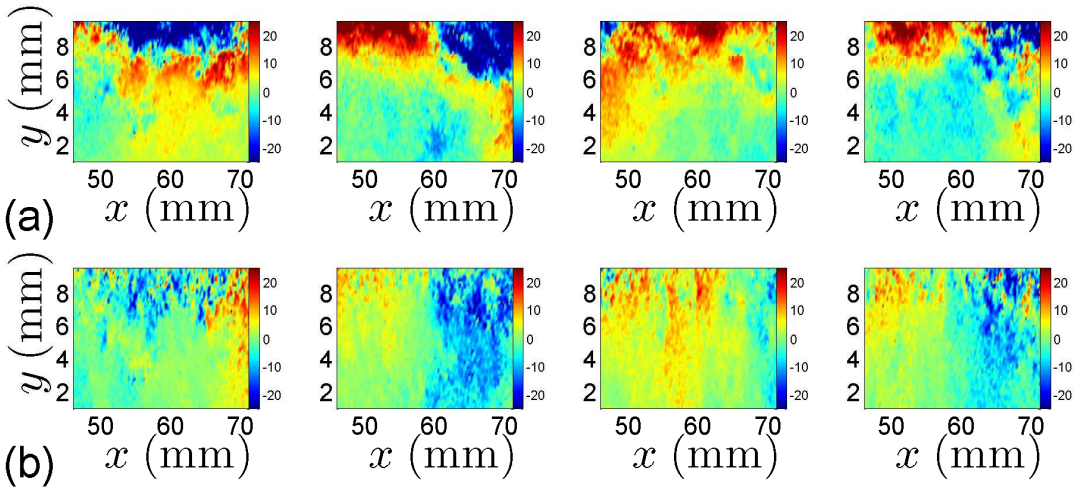


Figure 10.5: Some instantaneous velocity fields: (a) u-component; (b) v-component.

not done, the first POD mode would account for the mean flow and would contain a large part of the (steady) energy. The normalized cumulative sum of the eigenvalues of the POD modes are shown in Fig. 10.6. The normalized cumulative sum is defined by:

$$\sum_{i=1}^{n} \lambda_i / \sum_{i=1}^{N} \lambda_i$$

This is the ratio of the energy contained in the first n modes, divided by the total fluctuating energy, obtained by summing all the eigenvalues (there are N=300 eigenvalues, so that n is at most 300). The term "fluctuating energy" is used since the ensemble average has been removed from the snapshots. From the figure, only the first 5 or 6 eigenvalues stand apart as they cause

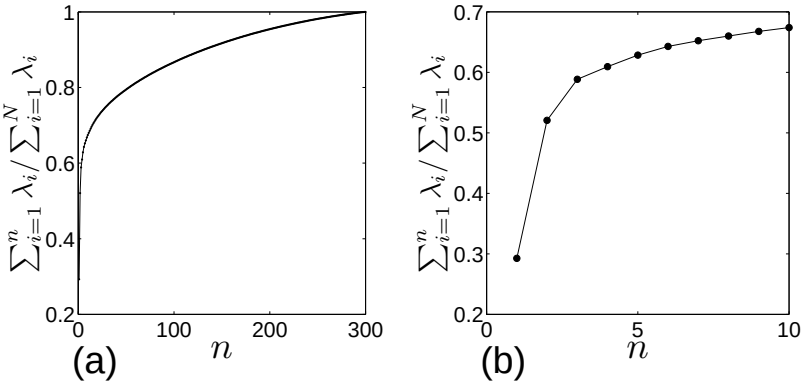


Figure 10.6: Cumulative sum of the eigenvalues: (a) full view; (b) zoom-in.

a jump in the cumulative sum. The other eigenvalues form a continuum and each of them add little energy. More specifically, the first eigenvalue accounts for about 29% of the fluctuating energy, and the first two account for 52% of the fluctuating energy. Clearly, the first two modes

are important for the dynamics of the flow.

The first four POD modes are shown in Fig. 10.7. Each POD mode corresponds to an axial

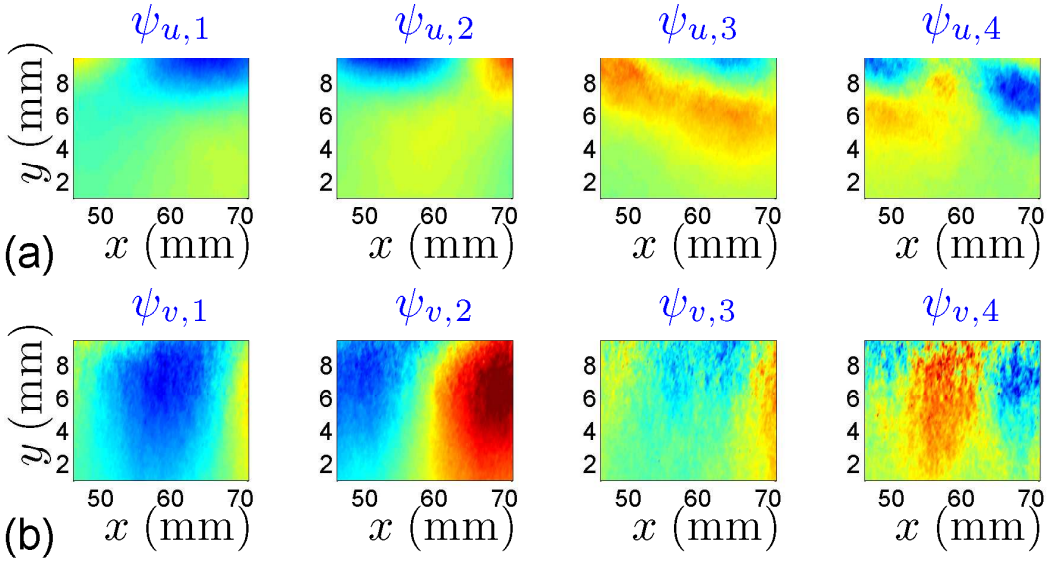


Figure 10.7: First POD modes. (a) u-component; (b) v-component.

component $\psi_{u,i}$ and a transverse component $\psi_{v,i}$. The first two modes contain most of the structures that were guessed in the instantaneous fields in Fig. 10.5. These first two modes resemble each other, the second mode seems to be a translated version (more or less) of the first mode.

Once we know the modes and their energy, it is possible to compute the time coefficients. Of course, the snaphots are taken randomly, and the time coefficients are going to be random as well. However, the relation between them is not random. As a result, it is common to produce some "phase diagram", that is, a plot of a coefficient versus another. The time coefficient for the second mode, a_2, is plotted versus coefficient a_1 in Fig. 10.8(a). Each symbol in this plot corresponds to one of the N snapshots. Interestingly, the symbols seem to collapse approximately on a circle, meaning there is a phase shift of $\pi/2$ between the two time coefficients. A polar angle ϕ is assigned to each measurement in the a_2/a_1 plane. This phase angle allows re-ordering the snapshots by sorting them according to the phase they correspond to. The coefficients a_1 and a_2 are plotted versus the phase angle in Fig. 10.8(b). The phase shift $\pi/2$ is clearly seen. It typically indicates that the first two modes correspond to the convection of a structure. It is further possible to define some angular sectors in the a_1/a_2 plane, and to average either the snapshots or the time coefficients falling within one sector. This allows a kind of phase averaging even though no-phase locking has been used when taking the snapshots. 10 angular sectors have been defined in Fig. 10.8(a), corresponding to 10 phase ϕ_j, j=1,2,...,10. The coefficients $a_{1,avg}(\phi_j)$ and $a_{2,avg}(\phi_j)$ (which are defined for the 10 values of ϕ_j) obtained by averaging a_1 and a_2 in each of the sectors are also shown in Fig. 10.8(b); they represent smoothed versions of a_1 and a_2.

Finally, it is possible to reconstruct the flow field using only two modes and their phase-

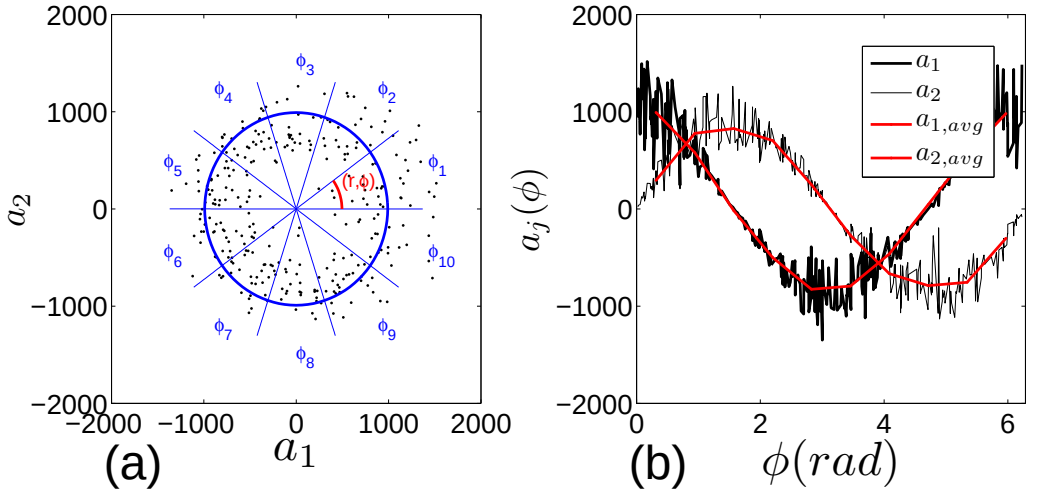


Figure 10.8: (a) Phase diagram: a_2 versus a_1. (b) Coefficients plotted versus the phase (re-ordered coefficients).

averaged time coefficients. For the u component, we have:

$$u_{recons}(x, y, \phi_j) = a_{1,avg}(\phi_j)\psi_{u,1}(x, y) + a_{2,avg}(\phi_j)\psi_{u,2}(x, y) \qquad \forall j = 1, 2, ..., 10$$

This is of course a truncated POD expansion with 2 modes (and optimality means that no other truncation using two modes will contain as much energy as this one). It represents the flow that contains the most energetic (coherent) structures after it has been re-ordered according to the phase (and smoothed as well because $a_{1,avg}$ is an average of coefficients). Figure 10.9 shows this reconstructed flow field at each of the 10 phases. Note: while a phase has been defined, meaning

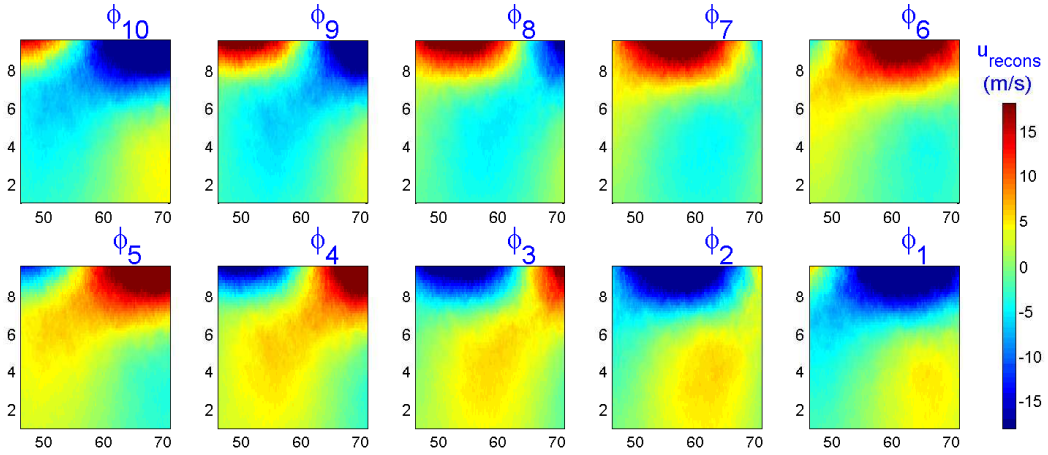


Figure 10.9: Two-mode reconstruction of the flow in the x-direction, $u_{recons}(x, y, \phi_j)$, at each of the 10 phases.

the flow is most probably periodic, we do not know the precise period because the time scale is

missing in that particular case due to the snapshots being captured in a random manner. We do not know either whether we should play the movie corresponding to the images in Fig. 10.9 from phase 1 to phase 10, or from phase 10 to phase 1. However, we have some clues, physically speaking. We know that the flow is from left to right. Hence, the movie should be played from phase 10 to phase 1, so that the structures are convected in the direction of the flow. In addition, we know we are triggering the instability by using an acoustic wave at frequency 1 kHz. Thus, the period of the observed phenenom is likely to be 1 kHz.
Finally, one concludes that the flow is dominated by large scale structures that are convected with the flow and appear periodically (ie, this is close to a wave). It can be checked that the magnitude of the structures increases exponentially in the positive x-direction, which corresponds to a spatial instability.

10.7 Conclusions

The POD provides an orthogonal basis that is optimal in terms of energy.

When used as a signal processing tool, POD is nothing but a filtering operation. The filtering of events is done according to their energy content. For example POD allows filtering out the less energetic features of a flow. This can be compared to other filtering means: using the Fourier transform high frequencies may be filtered out, irrespective of their energy content. Using wavelets, small scales may be filtered out, again irrespective of their energy content.

POD separates low and high energy modes in the signal/flow, irrespective of any physical consideration (for example, acoustic and vorticity "modes" are not going to be separated by using POD). Some physical structures may be made of several POD modes.

POD is also a tool for producing Reduced Order Models (ROM). The governing equations are then projected onto the POD modes using a Galerkin method. This allows studying the dynamics of a flow. In that context, it should be stressed that POD modes are representative of the particular flow samples that have been used to calculate them. For example, modes calculated at Re=4000 may not be efficient at for a ROM at Re=8000. It has also been shown on an example that a ROM model may converge to a wrong limit cycle unless some cautions are taken.

Appendix A | Some functions

A.1 Some functions

Cardinal sine function (or sinc function)
The cardinal sine function is defined by:

$$\operatorname{sinc}(t) = \begin{cases} 1 & \text{if } t{=}0 \\ \frac{\sin(t)}{t} & \text{otherwise} \end{cases} \tag{A.1}$$

Its graph is shown in Fig. A.1.

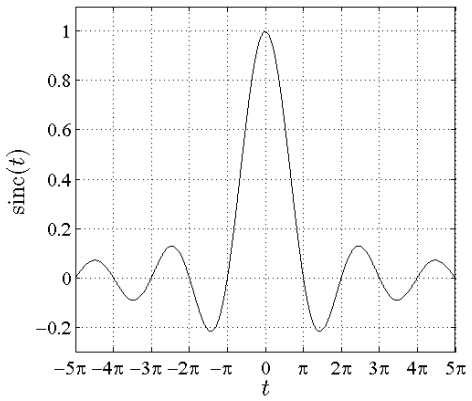


Figure A.1: Graph of the cardinal sine function.

Complex exponential
The Euler formula for the complex exponential is:

$$e^{\mathrm{j}2\pi f_0 t} = \cos(2\pi f_0 t) + \mathrm{j}\sin(2\pi f_0 t) \tag{A.2}$$

The cosine and sine functions can be expressed in terms of exponential functions as well:

$$\cos(2\pi f_0 t) = \frac{1}{2}\left(e^{\mathrm{j}2\pi f_0 t} + e^{-\mathrm{j}2\pi f_0 t}\right) \tag{A.3}$$

$$\sin(2\pi f_0 t) = \frac{1}{2\mathrm{j}}\left(e^{\mathrm{j}2\pi f_0 t} - e^{-\mathrm{j}2\pi f_0 t}\right) \tag{A.4}$$

Rectangular window

The rectangular window is defined by:

$$\Pi_T(t) = \begin{cases} 1 & \text{if } -T/2 \leq t \leq T/2 \\ 0 & \text{otherwise} \end{cases} \tag{A.5}$$

Its graph is given in Fig. A.2.

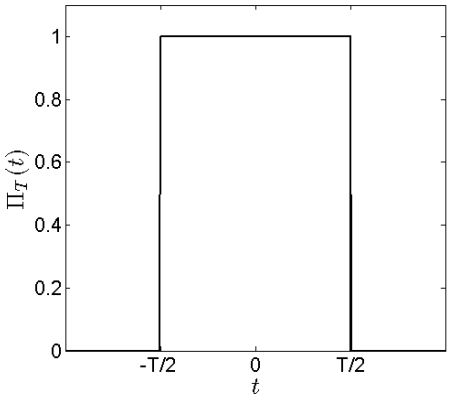


Figure A.2: Graph of the rectangular window function.

Dirac delta function

The Dirac delta function may be seen as the limit of a very narrow rectangular window function when this function becomes narrow while its integral is kept equal to 1. In this limit, the Dirac is a spike with a nul basis, and an infinite height. One may loosely write:

$$\lim_{T\to 0}\frac{1}{T}\Pi_T(x) = \delta(x)$$

The previous definition is not rigorous because the Dirac is not a regular function. It is a distribution, that assigns a real value to a function. The rigorous meaning is that:

$$\lim_{T\to 0}\int_{-\infty}^{\infty}\frac{1}{T}\Pi_T(x)\phi(x)\mathrm{d}x = \phi(0)$$

where $\phi(x)$ has to be a smooth function with compact support.The following notation is generally given as the definition of the Dirac delta function:

$$\int_{-\infty}^{\infty}\phi(x)\delta(x)\mathrm{d}x = \phi(0) \tag{A.6}$$

where a limit is implicit. The Dirac distribution acts on a function $\phi(x)$ and assigns the value $\phi(0)$ to that function.

Many other functions can used to define the Dirac as a limit, as for instance the sinc function. In that case, we have:

$$\lim_{a\to 0}\frac{1}{a}\mathrm{sinc}\left(\frac{x}{a}\right)=\delta(x)$$

the rigorous meaning being:

$$\lim_{a\to 0}\int_{-\infty}^{\infty}\frac{1}{a}\mathrm{sinc}\left(\frac{x}{a}\right)\phi(x)\mathrm{d}x=\phi(0)$$

The following properties are classical:

$$\int_{-\infty}^{\infty}\phi(x)\delta(x)\mathrm{d}x=\phi(0)\quad\text{ or }\int_{-\infty}^{\infty}\phi(x)\delta(x-x_0)\mathrm{d}x=\phi(x_0) \tag{A.7}$$

$$\delta(ax)=\frac{1}{|a|}\delta(x)\quad\text{with in particular}\quad\delta(-x)=\delta(x) \tag{A.8}$$

$$\delta(x-x_0)\cdot f(x)=\delta(x-x_0)f(x_0)\neq f(x_0) \tag{A.9}$$

$$\delta(x)*f(x)=f(x) \tag{A.10}$$

$$\delta(x-x_0)*f(x)=f(x-x_0) \tag{A.11}$$

where $*$ stands for the convolution product. Note that to get rid of a Dirac, an integration is needed. The Dirac is neutral element for the convolution. The convolution by $\delta(x-x_0)$ allows translating a function. Note also that the product of two distributions, such as $\delta(t)^2$ is undefined. Only the product of a distribution and a regular function is allowed.

Appendix B | Matlab commands

Some general commands are given in Table B.1. Commands for signal processing are given in Table B.2. Table B.3 provides some commands in Matlab that are useful for dealing with vectors and matrices. You can obtain help on any command `command` using `help command` or `doc command` in the matlab prompt.

Command	**Description**
`help command`	Gives the help for the command `command`.
`doc command`	Gives the doc for the command `command`.
`plot(t,x,'b')`	Plots x versus t in blue color.
`loglog(t,x,'r')`	Logarithmic Plot of x versus t in red color.
`hold on`	Means that several curves can be plotted in the same figure. If not used, the new plot will replace the old one.
`abs(a)`	Returns the modulus of the complex number a.
`angle(a)`	Returns the argument (phase angle) of the complex number a.

Table B.1: General Matlab commands.

Command	**Description**
`X=fft(x,N)`	Fast Fourier tranform (that is, fast calculation of the DFT) of x using N points (zero padding if size of $x <$ N). N is optional.
`x=ifft(X)`	Fast Fourier tranform Inverse (that is, fast calculation of the IDFT) of X
`X=fftshift(X))`	Re-order the vector X so that the elements are sorted in ascending order of their corresponding frequencies.
`W=hanning(100)`	Creates a discrete hanning window with 100 points.
`y=x.*hanning(length(x))`	y is the vector x windowed by a hanning window.
`Sxx=psd(x)`	Computes the power spectral density of x.
`z=hilbert(x)`	Compute the (complex) analytical associate z of the real signal x. That is, the result is $z = x + \mathrm{j}H(x)$. Warning: the function `hilbert` DOES NOT return the Hilbert transform $H(x)$ of x.
`B=randn(N,1)`	Creates a column vector B of size N containing a white noise.

Table B.2: Some matlab commands for signal processing.

Command	Description
`t=0:0.01:1`	Creates a vector t spannning the interval (0;1) with steps 0.01.
`t=t(:)`	Reshapes t as a column vector (useful when t is row-vector).
`t=linspace(a,b,N)`	Creates a vector t spannning the interval (a;b) with N points.
`length(t)`	Returns the number of elements of a vector.
`A=zeros(N)`	Creates a square matrix of size N initialized with zeros.
`size(A)`	Returns the size (nbrs of columns and rows) of a matrix A. Can be applied to a vector.
`A=zeros(M,N)`	Creates a rectangular matrix of size M x N initialized with zeros.
`A=zeros(N,1)`	Creates a column vector of size N initialized with zeros.
`A=zeros(1,N)`	Creates a row vector of size N initialized with zeros.
`A=randn(1,N)`	Creates a row vector of size N initialized with random numbers normally distributed. This is a white noise (all frequencies equally present) whose amplitude is normally distributed (ie with a Gaussian law).
`A=rand(1,N)`	Creates a row vector of size N initialized with random numbers uniformly distributed.
`A(:,i)`	Returns the i-th column of the matrix A.
`A(i,:)`	Returns the i-th row of the matrix A.
`C=A*B`	The matrix C is the product of A and B.
`C=A.*B`	".*" multiplies vectors of matrices elements by elements. If A and B are matrices, then $C(i,j) = A(i,j)B(i,j)$. If A and B are vectors, then $C(i) = A(i)B(i)$.
`diag(A)`	Applied to a matrix A: returns a vector with the diagonal elements, Applied to a vector A: returns a diagonal matrix whose diagonal is the input vector.
`A'`	Adjoint A^+ of the matrix A (transpose if A real, conjugate transpose if A complex.)
`transpose(A)`	Transpose of the matrix A (for a real or complex matrix).
`rank(A)`	Returns the rank of the matrix A.
`A(:,i)'*A(:,j)`	This is the scalar product of th i-th and j-th columns of A. (=0) if these two columns are orthogonal.
`norm(A,'fro')`	Returns the Frobenius norm of a matrix A.
`eig(A)`	Compute the eigenvalues and eigenvectors of a square matrix. Example: `[V,D]=eig(A)` returns: the diagonal matrix D containing the eigenvalues, and the matrix V containing the eigenvector (columns of V), with: $MV = VD$.
`svd(A)`	Compute the Singular Value Decomposition of a matrix. Example: `[U,S,V]=svd(A)` returns: the matrix U of the left-singular vectors, the diagonal matrix S of the singular values, and the matrix V of the right-singular vectors.

Table B.3: Some useful matlab commands for vectors and matrices.

Appendix C | Probability

Some fundamentals of probability theory are reviewed in the present section. Only the material necessary to address random processes in chapter 5 is covered. This matter in introduced in more depth in Papoulis [26], or in Bendat and Piersol [3]. In connection with turbulence, it is also presented in Chap. 6 of Tennekes and Lumley [37], or in Chap. 3 of Pope [27].

C.1 Random variables, Probability Density Function

C.1.1 Random variable

The Webster English dictionary defines "random" as "lacking a definite plan, purpose, or pattern". This suggests that no intelligible pattern can be found in a given random sequence of symbols. However, eventhough individual events are not certain, there is some regularity in their occurrences. Describing this regularity is the purpose of probability and statistics.

In the following, we perform an experiment whose outcome is random (ie not known in advance). The possible outcomes form a set, and we will call this set the outcome space (often called the sample space). Let the outcomes of the outcome space be indexed by ξ. A **random variable** $X(\xi)$ assigns a real numerical value (generally $\in (-\infty, \infty)$) to each possible outcome ξ of the outcome space. If X takes discrete values, the random variable is said to be discrete, if it takes continuous values, the random variable is said to be continuous.

To be specific, let us consider the typical experiment that will be of interest to us in the course: start a flow a time 0 (for example, by switching on the fan that feeds a channel or a jet), wait until some given time t, and make a velocity measurement at a given point in space. To obtain several measurements, you have to repeat all the steps again. The velocity so measured is random (if the flow is turbulent). Note that while you set the fan speed at the same value from one experiment to the next, and your measuring station is unchanged, the velocity is random because the flow is highly sensitive to initial small perturbations that are not under your control. For this experiment, the continuous random variable is the numerical value of the measured velocity. To simplify the presentation, we say that the first outcome corresponds to ξ=1, the second outcome corresponds to ξ=2, ... (all the way to a very large value). Hence, the random variable assigns a real value (a velocity) to an experiment number ξ. Hence, $X(\xi = 2)$=20 means that the measurement for the second experiment is 20 ms^{-1}.

C.1.2 Probability Density Function (PDF)

What matters is the probability that the random variable takes values within a certain range. This question can be answered when one knows the **probability density function** (PDF) $p_X(u)$ of the random variable X, that fully characterizes the random variable. The requirement for a function to be a PDF is that:

$$p_X(u) \geq 0$$

and that:

$$\int_{-\infty}^{\infty} p_X(u)\mathrm{d}u = 1 \tag{C.1}$$

The factor $p_X(u)\mathrm{d}u$ is the probability that the random variable takes values between u and $u + \mathrm{d}u$. More generally, the probability that the random variable takes values between a and b is:

$$P(a < X \leq b) = \int_a^b p_X(u)\mathrm{d}u \tag{C.2}$$

Note that $\{a < X \leq b\}$ is a set of outcomes. This is the set $\{\xi; X(\xi) > a \text{ and } X(\xi) \leq b\}$, the set of the measurement numbers, such that the velocity is more than a and less than b. Equation (C.1) states that the probability of the velocity taking values between $-\infty$ and ∞ is 1, ie this event is certain.

C.1.3 Moments of the PDF

The PDF can also be characterized by gross quantities that are called the **moments** of the PDF. The n-th order moment is:

$$M_n = \int_{-\infty}^{\infty} u^n p_X(u)\mathrm{d}u \tag{C.3}$$

A PDF is known when all its moments are known. The first order moment is called the **expected value**, $E[X]$, of the random variable, and is given by:

$$m_X = E[X] = \int_{-\infty}^{\infty} u p_X(u)\mathrm{d}u \tag{C.4}$$

In words, a velocity u is multiplied by its probability density and integrated. The integral gives more weight to velocities that are more likely. The result is the expected mean value, and this is also written m_X. The second moment is the mean square value, which is the expected value of X^2:

$$E[X^2] = \int_{-\infty}^{\infty} u^2 p_X(u)\mathrm{d}u \tag{C.5}$$

A related quantity is the variance σ_X^2 given by:

$$\sigma_X^2 = E[(X - m_X)^2] = \int_{-\infty}^{\infty} (u - m_X)^2 p_X(u)\mathrm{d}u \tag{C.6}$$

This is the square of the standard deviation σ_X.
The two moments m_X and $E[X^2]$ (or σ_X^2) are the most widely used gross quantities for a PDF, even if they do not suffice to fully characterize this PDF. They allow to say that, statistically, the random variable will take values mostly in $[m_X - \sigma_X, m_X + \sigma_X]$.

C.1.4 Gaussian PDF

A classical PDF is the Gaussian PDF. It is entirely characterized by its first two moments, that is, it just depends on m_x and σ_X. Other moments depend on the first two. The Gaussian PDF is defined by:

$$p_X(u) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(u-m_X)^2/(2\sigma_X^2)} \tag{C.7}$$

This PDF is shown in Fig. C.1(a) for a mean velocity m_X=20 m/s and a standard deviation σ_X=5 m/s. Most probable velocities are within the range $m_X \pm \sigma_X$. The probability that

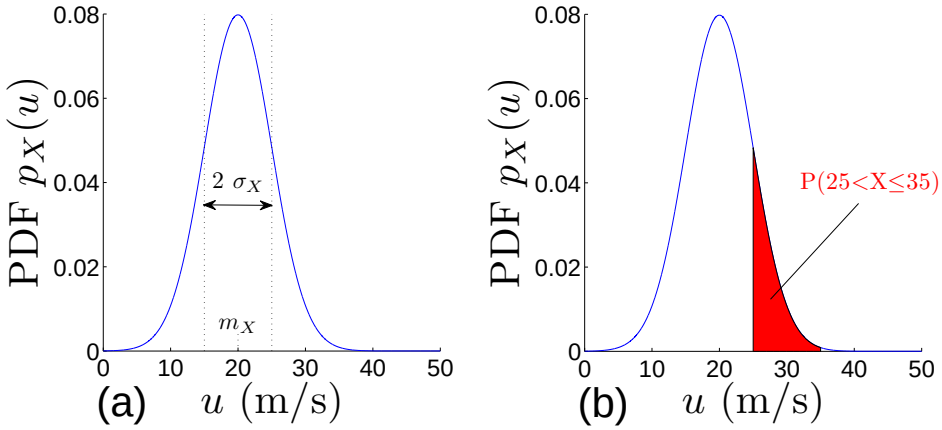


Figure C.1: (a) Gaussian probability function; (b) the probability that the velocity is between 25 and 35 m/s is obtained by integrating the probability density function between these two velocities.

the velocity is between 25 m/s and 35 m/s is obtained by integrating the shaded region in Fig. C.1(b), it is given by:

$$P(25 < X \leq 35) = \int_{25}^{35} p_X(u)\mathrm{d}u \sim 0.157$$

C.1.5 Experimental determination of a PDF

Let us come back to our experiment, and suppose we have made N_{tot}=23 random velocity measurements, ξ=1...23. These are shown in Fig. C.2(a). The question is: how do we find experimentally the PDF for our experiment ? The task is easy and merely a counting operation. First, we define some velocity bins. Here the bins are centered on odd velocities (19,21,23,25,...) and have a width of 2 m/s. So, for example, velocities between 18 m/s and 20 m/s fall in the bin 19 m/s, and velocities between 22 and 24 m/s fall in the bin 23 m/s. Note that by defining these bins, we transform the continuous random variable (that takes any real value) into a discrete

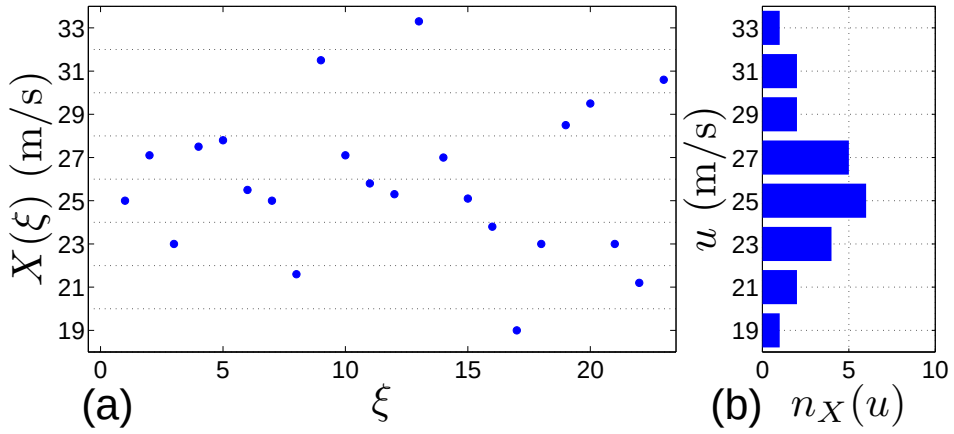


Figure C.2: (a) Samples obtained by measurement of a random variable; (b) Number of samples per velocity plot. A value in a bin/line in (b) is obtained by counting the number of samples in the corresponding line in (a).

one (that takes only discrete values: 19, 21, 23, ...). Second, we count the number of samples per bin u, which we call $n_X(u)$. For example, the number of samples in the bin 23 m/s is $n_X(u = 23)$=4. The plot u versus $n_X(u)$ is plotted that way, see Fig. C.2(b). The probability that a measurement is done in the bin corresponding to velocity u is $p_X(u) = n_X(u)/N_{tot}$, this is the number of samples in the bin u divided by the total number of samples. Finally, the plot of $p_X(u)$ versus u is presented in Fig. C.3 (it is a normalized rotation of the plot u vs n_u). The

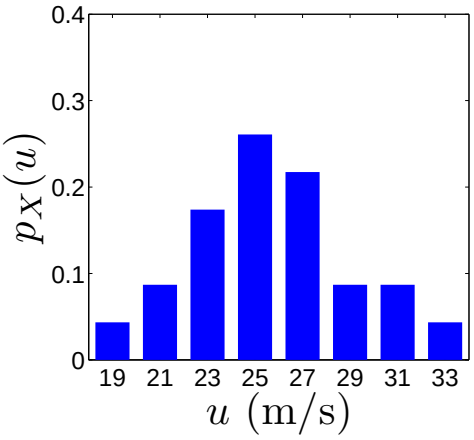


Figure C.3: Experimental PDF.

values taken by the PDF are given in the following table:

bin u (m/s)	19	21	23	25	27	29	31	33
PDF $p_X(u)$	1/23	2/23	4/23	6/23	5/23	2/23	2/23	1/23

Of course we have the discrete equivalent of Eq. (C.1):

$$\sum_{\text{bins } u} p_X(u) = \frac{1}{23} + \frac{2}{23} + \frac{4}{23} + ... = 1$$

To calculate the expected value, we use Eq. (C.4) that we transform into a discrete sum:

$$m_X = \sum_{\text{bins } u} u p_X(u) = 19\frac{1}{23} + 21\frac{2}{23} + 23\frac{4}{23} + ... \sim 25.7 \text{ m/s}$$

The variance is calculated by using the discrete version of Eq. (C.6):

$$\sigma_X^2 = \sum_{\text{bins } u} (u - m_X)^2 p_X(u) = (19 - m_X)^2 \frac{1}{23} + (21 - m_X)^2 \frac{2}{23} + (23 - m_X)^2 \frac{4}{23} + ... \sim 11.3 \text{ m}^2/\text{s}^2$$

And the standard deviation is:

$$\sigma_X \sim 3.37 \text{ m/s}$$

Exercise (using Matlab)

Producing a curve as in Fig. C.2(a) is possible in Matlab for several different PDFs. The two Matlab functions `x=rand(1000,1)` and `x=randn(1000,1)` will create a vector containing 1000 random numbers (or samples). For one of the two Matlab functions (`rand`/`randn`), the amplitudes of the random numbers follow a uniform distribution (distribution = probability density function), and for the other, the amplitudes follow a Gaussian distribution. Use the command `hist(x)` (or `hist(x,Nbin)` to precise a number of bins) to plot a graph similar to that in Fig. C.2(b) and to determine which function corresponds to which PDF. For the Gaussian PDF, judging by eye, what is the mean and the standard deviation of the random variable ? (you may check your estimate by using Matlab commands `mean` and `std`) Using Matlab help, create a vector `x` that has a Gaussian PDF, with an expected value of 11 and a standard deviation of $\sqrt{3}$.

C.2 Two random variables, Joint Probability Density Function

C.2.1 Joint probability density function

Very often, one has to take into account more than just one random variable. This is the case for example when both the axial and transverse velocities are measured during an experiment. In that case, let $X(\xi)$ and $Y(\xi)$ be the two random variables. Let p_X and p_Y be the two corresponding probability density functions. These allow answering questions such as: what are the means of X and Y, what are the standard deviations of X and Y, what is the probability that X exceeds some value (independently of Y), and what is the probability that Y exceeds some value (independently of X). However, these separate PDF do not allow answering questions such as: what is the probability that both X and Y exceeds some value. Answering this question requires the knowledge of a joint probability density function (joint PDF). The **joint density probability function** $p_{XY}(x, y)$ verifies:

$$p_{XY}(x, y) \geq 0 \tag{C.8}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{XY}(x,y)\mathrm{d}x\mathrm{d}y = 1 \tag{C.9}$$

The quantity $p_{XY}(x,y)\mathrm{d}x\mathrm{d}y$ is the probability that the random variable X takes on a value between x and $x+\mathrm{d}x$ and at the same time the random variable Y takes on a value between y and $y+\mathrm{d}y$. More generally, the probability that X be between a and b, and that at the same time Y be between c and d is:

$$P(a < X \leq b, c < Y \leq d) = \int_a^b \int_c^d p_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{C.10}$$

The marginals, that are the separate PDF of X and Y, are obtained from the joint PDF by:

$$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)\mathrm{d}y \tag{C.11}$$

$$p_Y(y) = \int_{-\infty}^{\infty} p_{XY}(x,y)\mathrm{d}x \tag{C.12}$$

Hence, to obtain a PDF for one of the random variables, the joint PDF is integrated with respect to the variable corresponding to the other random variable.

C.2.2 Example: joint normality

An example of joint probability density function is the joint Gaussian PDF:

$$p_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y(1-r_{XY}^2)} e^{-\frac{1}{2(1-r_{XY}^2)}\left(\frac{(x-m_X)^2}{\sigma_X^2} - 2r_{XY}\frac{(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2}\right)} \tag{C.13}$$

where r_{XY} is the correlation coefficient satisfying $|r_{XY}| < 1$. The marginals are also Gaussian, with:

$$p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} \tag{C.14}$$

$$p_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{(y-m_Y)^2}{2\sigma_Y^2}} \tag{C.15}$$

These are as in Eq. (C.7).
Two joint Gaussian PDFs are shown in Fig. C.4. These joint PDFs have the same marginals but the joint PDF in Fig. C.4(a) has a high correlation coefficient, r_{XY}=0.99, while the one in Fig. C.4(b) has a null correlation coefficient, r_{XY}=0. A high correlation coefficient results in the PDF having a straight line shape, while a low coefficient results in the distribution having a round shape.

C.2.3 Expected value of a function $g(X,Y)$

Let $g(X,Y)$ be a function of both X and Y. The expected value of g, $E[g]$, is given by:

$$E[g] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{C.16}$$

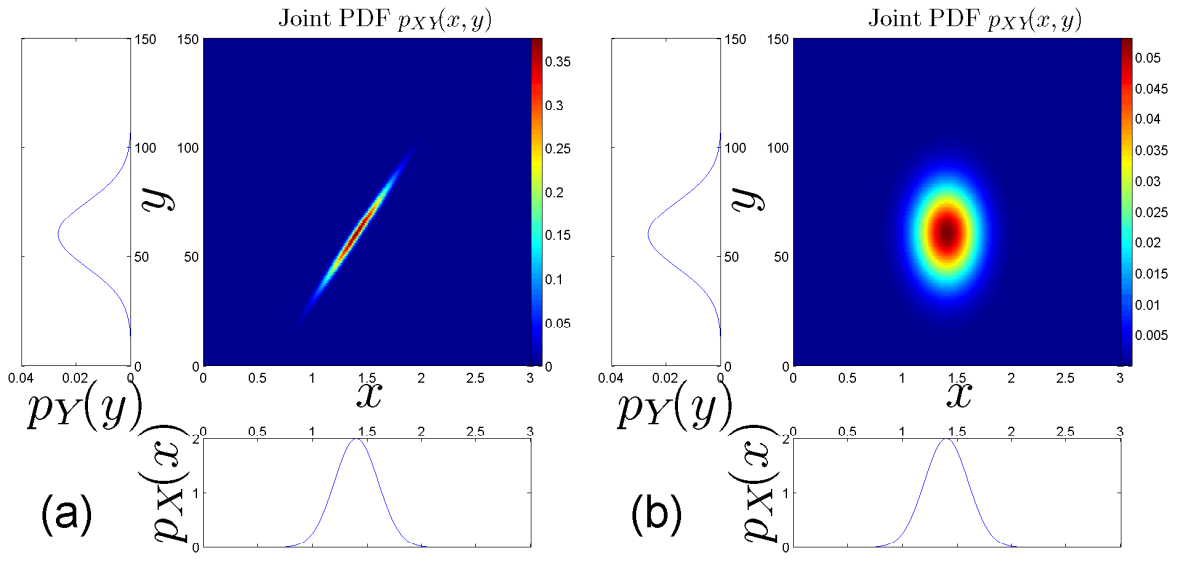


Figure C.4: Two joint Gaussian probability density functions. Both have the same marginals with: m_X=1.4; σ_X=0.2; m_Y=60; σ_Y=15. But their correlation coefficient is not the same, with: (a) r_{XY}=0.99; (b) r_{XY}=0.

Example 1: let $g(X,Y) = X$. In that case we want to calculate the expected value of X, that is, the mean value of X:

$$E[g] = E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x p_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{C.17}$$

$$= \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} p_{XY}(x,y)\mathrm{d}x}_{p_X(x)}\,\mathrm{d}y \tag{C.18}$$

$$= \int_{-\infty}^{\infty} x p_X(x)\mathrm{d}x \tag{C.19}$$

As expected, the result just depends on the marginal p_X (since g does not depend on Y) and is indeed $m_X = E[X]$ as calculated by a monovariate PDF.
Example 2: to calculate the correlation of X and Y, we take $g(X,Y) = XY$, giving:

$$\mathrm{corr}_{XY} = E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy p_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{C.20}$$

Example 3: to calculate the covariance of X and Y, we take $g(X,Y) = (X - m_X)(Y - m_Y)$, giving:

$$\mathrm{cov}_{XY} = E[(X - m_X)(Y - m_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - m_X)(y - m_Y) p_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{C.21}$$

We have the relation: $\mathrm{cov}_{XY} = \mathrm{corr}_{XY} - E[X]E[Y]$ and also: $|\mathrm{cov}_{XY}| \leq \sigma_X\sigma_Y$. Hence, defining the correlation coefficient by:

$$\mathrm{r}_{XY} = \frac{\mathrm{cov}_{XY}}{\sigma_X\sigma_Y}$$

we have: $|r_{XY}| \leq 1$.

C.2.4 Conditional probability

Conditional distributions allow answering questions such as: what is the probability that X takes on values between 20 and 30 ms^{-1} knowing that Y takes on values between 10 and 15 ms^{-1}. The conditional distribution of x assuming y is:

$$p_{XY}(x \mid y) = p_{XY}(x, y)/p_Y(y)$$

Conditional probabilities may be calculated according to:

$$P(a < x \leq b \mid y) = \int_{-\infty}^{\infty} p_{XY}(x \mid y)\mathrm{d}x = \int_{-\infty}^{\infty} \frac{p_{XY}(x, y)}{p_Y(y)}\mathrm{d}x \tag{C.22}$$

Appendix D | Matrices

Some fundamentals on matrices necessary to understand chapter 10 are provided in this appendix. For a course on linear algebra, refer to the classical book by Strang [35]. Another useful reference is Trefethen and Bau [39].

D.1 Generalities on matrices

D.1.1 Matrix and linear operators

A matrix A of size m x n is a representation of a linear operator (also noted A) from a vector space $E = C^n$ of dimension n to a vector space $F = C^m$ of dimension m. Hence:

$$A : C^n \to C^m$$

A basis of $E = C^n$ is the set of vectors $e_j, j = 1...n$. A basis of $F = C^m$ is the set of vectors $f_i, i = 1...m$. The linear application is fully determined when the images $A(e_j) = Ae_j, j = 1...n$, of each the basis vectors of the departure space are known. These images are specified by the elements of the matrix. These elements are noted A_{ij}, $i = 1...m$, $j = 1...n$, and represent the coordinates of the vectors Ae_j in the basis $\{f_i, i = 1...m\}$. To be more specific, we have:

$$Ae_j = \sum_{i=1}^{m} A_{ij} f_i \qquad \forall j = 1...n$$

That is to say, the image Ae_j is the j-th column of A, as represented below:

$$\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \overset{\begin{array}{c} j-\text{th} \\ \text{column} \end{array}}{\begin{pmatrix} \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & Ae_j & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \end{pmatrix}} \tag{D.1}$$

An important quantity is the **rank** of the matrix which is the number of linearly independent columns, which is also the number of linearly independent rows.

Adjoint:
The adjoint A^+ of a *complex* matrix A is the **conjugate transpose** of A.
If the matrix A is *real* then the adjoint of A is simply the **transpose** A^t of A.

Scalar product:
The scalar/inner product (C_1, C_2) between two column vectors is:

$$(C_1, C_2) = C_1^+ C_2$$

In this expression, C_1^+ is a row vector.
In the case when the vectors are *real*, the scalar product reduces to: $(C_1, C_2) = C_1^t C_2$.

Orthogonal and orthonormal vectors:
Two column vectors C_1 and C_2 are **orthogonal** when their scalar product vanishes:

$$(C_1, C_2) = C_1^+ C_2 = 0$$

They are **orthonormal** if they are orthogonal and if they have a unit norm: $C_1^+ C_1$=1; $C_2^+ C_2$=1.
A matrix A has orthonormal columns if: $A^+ A = I$, where I is the identity matrix.

Unitary matrices:
A **unitary matrix** U is a complex square matrix that verifies:

$$UU^+ = U^+ U = I \qquad \text{(unitary matrix)}$$

Hence, a unitary matrix is a matrix whose columns form an orthonormal basis of C^n. The rows also form an orthonormal basis of C^n. Note that U is invertible, with $U^{-1} = U^+$.

D.1.2 Some rules for multiplication

Multiplication of a matrix with vectors C_i in its columns by a matrix A
Let's C be a matrix of size n x p with columns C_1, C_2, ... C_p. These columns represent p n-component vectors belonging to space $E = C^n$. By multiplying this matrix on the left by the matrix A, the result is made of columns AC_i which belong to $F = C^m$. We have:

$$\begin{pmatrix} & A & \end{pmatrix} \cdot \begin{pmatrix} C_1 & \dots & \dots & \dots & C_p \end{pmatrix} = \begin{pmatrix} AC_1 & \dots & \dots & \dots & AC_p \end{pmatrix} \tag{D.2}$$

Left multiplication by a diagonal matrix
On multiplying a matrix by a diagonal matrix on the left, one multiplies the **rows** of this matrix by the entries in the diagonal matrix.

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 r_1 \\ \lambda_2 r_2 \end{pmatrix} \tag{D.3}$$

Right multiplication by a diagonal matrix
On multiplying a matrix by a diagonal matrix on the right, one multiplies the **columns** of this matrix by the entries in the diagonal matrix.

$$\begin{pmatrix} c_1 & c_2 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 c_1 & \lambda_2 c_2 \end{pmatrix} \tag{D.4}$$

D.2 Diagonalisation of a matrix

D.2.1 Diagonalizable matrix

A square matrix A can be diagonalized if there exists a basis of eigenvectors. By putting these eigenvectors in the columns of a matrix P, we may then write:

$$A = PDP^{-1} \tag{D.5}$$

D is a diagonal matrix with the eigenvalues λ_i on the diagonal:

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{D.6}$$

The meaning of Eq. (D.5) is made obvious by multiplying the later relation by P on the right, giving: $AP = PD$. Using Eq. (D.2) and Eq. (D.4), this is:

$$\begin{pmatrix} AP_1 & AP_2 & \dots & \dots & AP_n \end{pmatrix} = \begin{pmatrix} \lambda_1 P_1 & \lambda_2 P_2 & \dots & \dots & \lambda_n P_n \end{pmatrix} \tag{D.7}$$

Hence, $AP_i = \lambda_i P_i$, $\forall i = 1...n$. This is simply a statement that the columns P_i of P are eigenvectors with eigenvalues λ_i.

D.2.2 Finite-dimension spectral theorem (diagonalization with orthonormal eigenvectors)

A **normal matrix** A of size n is a square matrix which commutes with its adjoint:

$$AA^+ = A^+A \qquad \text{(normal matrix)} \tag{D.8}$$

In that case, there exists a unitary matrix U such that:

$$A = UDU^+ \quad \text{or also} \quad A = UDU^{-1} \tag{D.9}$$

where D is a diagonal matrix containing the eigenvalues λ_i. The orthonormal columns $\mathbf{u}_i$ of U are the eigenvectors. Indeed, the former equation is equivalent to $AU = UD$, which means that: $A\mathbf{u}_i = \lambda_i \mathbf{u}_i, \forall i = 1...n$. Said otherwise, a normal matrix is a matrix that can be diagonalized and whose eigenvectors form an orthonormal basis of C^n.

Special case of a Hermitian matrix:
A square matrix A is Hermitian when it is its own conjugate transpose, that is, when $A{=}A^+$. This implies that the matrix A is normal, and as a result A can be diagonalized by a unitary matrix. This is not all: in addition, for an Hermitian matrix, **the eigenvalues are all real**. Hence, for a Hermitian matrix A, one has:

$$A = UDU^+ = UDU^{-1} \quad \text{and} \quad D = \text{diag}(\lambda_1, \lambda_2, ...) \text{ where } \lambda_i \in R \tag{D.10}$$

Matrix A^+A or AA^+:
When the matrix A has nothing special, it is interesting to consider A^+A or AA^+ that are both Hermitian (since $(A^+A)^+ = A^+A^{++} = A^+A$). And the result in the last paragraph indicates that:

$$A^+A = UDU^+ = UDU^{-1} \quad \text{and} \quad D = \text{diag}(\lambda_1, \lambda_2, ...) \text{ where } \lambda_i \in R^+ \tag{D.11}$$

where $U^+U = I = UU^+$. Note that this time, the eigenvalues are real and also positive. This is particularly useful for POD.

Appendix E | Labs

Advanced Signal Processing - LAB 1

Power Spectral Density estimate using Matlab

D. Marx

In this lab, we want to estimate Power Spectral Densities (PSD) of random stationary ergodic signals using Matlab. PSD are important, since they allow to estimate frequency contents of signals, to calculate transfer functions of Linear Time Invariant systems and to estimate relations between signals using coherence.
In part I, a routine to implement Welch's method is constructed and tested.
In part II this routine is used to estimate the properties of signals measured during the study of an air conditioning exhaust noise.

Note:
You need to write a report on this lab. The name of the file containing the report should be:

FAMILYNAME_Firstname_Lab_labnumber.pdf

Part I - Periodogram

1 Background

1.1 Numerical Estimate of the Power Spectral Density using the Raw periodogram

Let $x(n)$ for $n = 0...N-1$ be the N samples of the analysed signal; these samples correspond to instants $t_n = n \cdot T_s$ where T_s is the sampling period. This is the signal whose PSD is ultimately wanted. Let $w(n)$ be the windowing function of size N. The windowed signal is $x_w(n) = x(n) \cdot w(n)$ for $n = 0...N-1$. Its Discrete Fourier Transform (DFT) $X_w(k)$ is:

$$X_w(k) = DFT[x_w(n)] = \sum_{n=0}^{N-1} x_w(n) e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} x(n) w(n) e^{-j2\pi kn/N}$$

for $k = 0...N-1$. This is calculated using function `fft` in Matlab. The index k is used for frequency, and there exists a vector $f(k)$ giving the frequency as a function of k (see lecture 1).

It is possible to estimate the PSD of the *windowed signal* according to the raw periodogram:

$$S_{xx,w}(k) = \frac{|X_w(k)|^2}{N} = \frac{X_w(k) \cdot X_w^*(k)}{N}$$

The power spectral density of the original signal x *before windowing* is, according to the **raw periodogram**:

$$\boxed{S_{xx}(k) = C_w \cdot S_{xx,w}(k) = C_w \frac{|X_w(k)|^2}{N} = C_w \frac{X_w(k) \cdot X_w^*(k)}{N}} \tag{1}$$

This is the quantity we are interested in using Matlab. The corrective factor C_w depends on the window used and is given by:

$$C_w = \frac{N}{\sum_{n=0}^{N-1} |w(n)|^2} \tag{2}$$

Taking into account this factor is necessary for compensating the window effect, which tapers the signal and thus modifies its energy. The corrective factor is by the way the inverse of the energy of the window. It is given in table 1 for some windows used in practice.

Notice that the PSD calculated by Eq. (1) is a perfectly valid estimate if the signal is **deterministic**. This is only a *raw* estimate for a **random** signal. For such a random signal, the typical **error in the estimate**, $\sigma(S_{xx})$, is:

$$\boxed{\frac{\sigma(S_{xx}(k))}{S_{xx}(k)} \sim O(1)} \tag{3}$$

This means that the error is of the order of the calculated quantity, which is not a very good result. For a random signal, the averaged periodogram presented next had better be used.

Window	Expression	C_w (Eq. (2))
Rectangular	$w(n) = 1 \qquad \forall n$	1
Hanning	$w(n) = 0.5\left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right)$	2.67
Blackman	$w(n) = 0.42 - 0.5\cos\left(\frac{2\pi n}{N-1}\right) + 0.08\cos\left(\frac{4\pi n}{N-1}\right)$	3.24

Table 1: Value of C_w for some usual windows (obtained for large window size N).

1.2 Numerical Estimate of the Power Spectral Density using the Averaged periodogram (Welch's method)

When the signal is random, one should use the averaged periodogram to estimate the PSD. This consists in using pieces of the signal that are called blocks. Totally there are M blocks. Each block is of size N and the piece of signal in each block is denoted by $x^{(i)}(n)$ for $n = 0...N-1$, and for $i = 1...M$. Notice that the blocks may overlap (a 50% overlap is used in practice). The PSD of the signal $x(n)$ according to the **averaged periodogram (Welch's method)** is:

$$\boxed{\hat{S}_{xx}(k) = \frac{1}{M}\sum_{i=1}^{M} S_{xx}^{(i)}(k) \qquad \text{for } k = 0...N-1} \tag{4}$$

where $S_{xx}^{(i)}$ is the PSD of the block number i, calculated using Eq. (1). That is:

$$S_{xx}^{(i)}(k) = C_w \cdot \frac{|X_w^{(i)}(k)|^2}{N} = C_w \cdot \frac{X_w^{(i)}(k) \cdot X_w^{(i)*}(k)}{N} \tag{5}$$

where $X_w^{(i)} = DFT[x^{(i)}(n) \cdot w(n)]$ is the Discrete Fourier Transform of the windowed signal in the block number i.

The typical error in the estimate, $\sigma\left(\hat{S}_{xx}\right)$, is:

$$\boxed{\frac{\sigma\left(\hat{S}_{xx}(k)\right)}{\hat{S}_{xx}(k)} \sim O\left(\frac{1}{\sqrt{M}}\right)} \tag{6}$$

The error decreases with the number of blocks. There is a price to pay for this, since at the same time the frequency resolution decreases.

The averaged periodogram can also be used to calculate cross PSDs. The cross PSD of signals $x(n)$ and $y(n)$ is:

$$\hat{S}_{xy}(k) = \frac{1}{M}\sum_{i=1}^{M} S_{xy}^{(i)}(k) \tag{7}$$

with

$$S_{xy}^{(i)}(k) = C_w \frac{1}{N} X_w^{(i)*}(k) Y_w^{(i)}(k) \tag{8}$$

1.3 Signal Power

A stationary signal has finite power. According to Parseval's relation, it may calculated either in the time domain or in the frequency domain:

$$\mathcal{P}_{moy} = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n) \tag{9}$$

$$\mathcal{P}_{moy} = \frac{1}{N}\sum_{k=0}^{N-1} S_{xx}(k) \tag{10}$$

1.4 White noise

For a white noise, $S_{xx}(k)$ does not depend on k. It verifies:

$$S_{xx}(k) = \sigma^2 \qquad \forall k \tag{11}$$

The PSD is thus constant and equal to the noise variance. For the power we have:

$$\mathcal{P}_{moy} = \frac{1}{N}\sum_{k=0}^{N-1} S_{xx}(k) = \sigma^2 \tag{12}$$

A white noise sequence may be generated using Matlab command `randn`.

2 Work to be done for Part I

2.1 Study of a windowing function

Open Matlab.
a) Create a Hanning window w of size 128 (use the `hanning` Matlab command). Plot it.
b) Can you recover the value of the corrective factor C_w given in table 1 for the Hanning window?
c) We would like to study the first side-lobe attenuation for the *continuous* Hanning window. For this we need to append some zeros to the Hanning window created with Matlab. This is called zero-padding. In the following command lines, the window has initially a size 128 and trailing zeros are added to obtain a length 2048. This is done directly in the `fft` command.

```
N=128;
Npadding=2048;
w=hanning(N);
W=fft(w,Npadding);
W=fftshift(W);
f=(-Npadding/2:Npadding/2-1)/Npadding;
figure
plot(f,20*log10(abs(W)./max(abs(W))),'b')
```

Here $w(t)$ is the window, and $W(f)$ is its Fourier transform.
Create a new m-file that includes the above commands. Run this program.
What is the effect of the Matlab command `fftshift`?
The gain in dB is defined by $20\log(|W|)$. What is the difference of gain in dB between the main lobe and the first side-lobe?
Add on the plot the result obtained for a rectangular window (in that case, take `w=ones(N,1)`). What compromise do these curves illustrate?

2.2 Raw periodogram

Now open the main program for part I: `go_part_I.m`. It first defines a signal (that can be a cosine, a white noise, or their sum) and then computes the PSD and calculates the power. You can run this program. At the moment functions called by the program work but return zero-values. In the following you need to complete these functions.

a) When the noise is null (std_noise=0 in the program) the signal is a cosine. This is then a deterministic signal and we can estimate its PSD using the the raw periodogram.
Given the signal definition in the main program, is Shannon condition verified?
Open and complete the function `fct_raw(x,fs,iwindow)`. This should return the periodogram of a signal x according to Eq. (1). It should also return the frequency vector. Both the frequency and the PSD should be ordered so that negative frequencies come first. Once the function works the PSD of the signal is given in figure 2. Do you obtain the expected result? What is the effect of changing the value of N (say N=500 and N=2000)?
b) Open and complete the functions `fct_power_time` and `fct_power_freq`? They should return the power of the signal, calculated according to Eqs. (9) and (10) respectively. Is Parseval relation satisfied?
c) Now we can try to use the raw periodogram for a random signal. Let us use a white noise with variance σ^2. Set `A`=0 and `std_noise`=5 and run the program. From the PSD in figure 2, can you estimate properly the PSD of the noise?
d) Since the PSD of a white noise should be flat and equal to σ^2 for all frequencies (see eq. (11)), there is a simple way to estimate the mean and standard deviation of the PSD estimate: we can estimate these quantities by using only one realization, $S_{xx,raw}$, and considering the values at the different frequencies as different realizations of the PSD estimate. Then use the Matlab command `mean` to calculate the mean of $S_{xx,raw}$, and use the Matlab command `std` to calculate its standard deviation, $\sigma[S_{xx,raw}]$ (this is a single number that does not depend on frequency given the chosen method). How do these compare with Eq. (3)?

2.3 Averaged periodogram

a) Consider again the white noise only. We want to estimate the PSD of this signal using the averaged periodogram. Open and complete the function `fct_welch(x,fs,N,iwindow)`. This should return: the frequency vector going from negative to positive, the PSD calculated according to Eq. (4), and the number of blocks M that have been used. The result will be displayed in figure 2. Use N=300. Is the result more satisfying than with the raw periodogram?
b) Does Parseval relation hold?

2.4 Variance of the PSD estimator

Here we propose to check Eq. (6) in the same way as we have checked Eq. (3) (see 2.2d)). For this we vary the value of N (the size of one block) used as input to the function `fct_welch`, since this will correspond to changing M (the number of blocks). Then, for each value of N belonging to the vector `Nvec`, we want to calculate the mean and standard deviation of $S_{xx,welch}$. To do this, open and complete the function `fct_error(x,fs,Nvec,iwindow)`. The standard deviation is then plotted versus the number of blocks in figure 3, and this is compared with the result of Eq. (6).

Part II - Application to the sound produced by an air conditioning exhaust

3 Presentation of the measurements

3.1 Experimental Setup

The experiment modeling an air conditioning exhaust is presented in figure 1. Air with velocity U is blown through the blowing duct. The flow at the end of the duct is a jet. At the duct lip the shear layer leads to vortex formation. The vortices are convected at the convection velocity U_c. The pressure $z(t)$ at the duct wall is measured using a microphone. Two hot wires are used to measure the velocity. The first hot wire is fixed and provides the measurement of the velocity $x(t)$ in the potential core (where the flow is laminar). The second hot wire is a flying hot-wire that can be displaced in the shear layer along the ξ direction and provides a measurement $y(t)$ of the velocity. Signals are random due to noise and turbulence. The are digitally sampled with a sampling rate f_s=10kHz and the acquisition time is 15s.

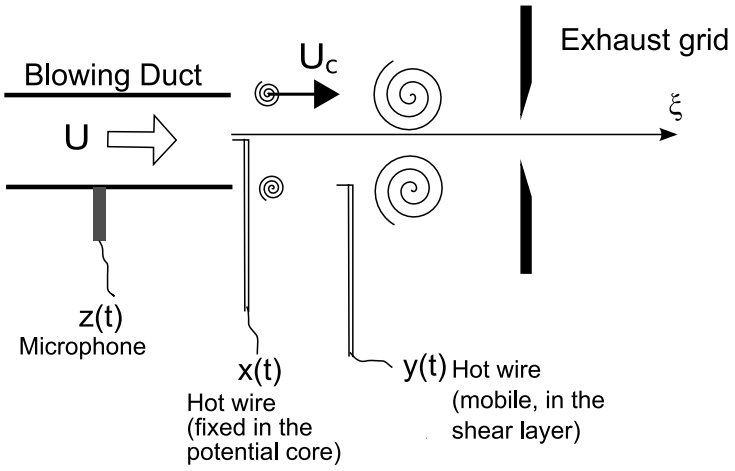


Figure 1: Schematic of the experiments.

3.2 Convection velocity

The structures (vortices) are convected at some convection speed U_c within the shear layer. A classical method to determine this convection speed is to use the cross-PSD of the signals from two hot wires. This is now presented.

First we suppose that the velocity y for the component at frequency f_0 at a point in the shear layer is that of a progressive wave:

$$y(\xi, t) = e^{-j(2\pi f_0 t - k\xi)} \qquad \text{with} \qquad k = \frac{2\pi f_0}{U_c} \tag{13}$$

At a fixed position in space, this represents a pure sine wave at the frequency f_0.
We may also write:

$$y(\xi,t) = e^{-j2\pi f_0 t} e^{j\phi_y(\xi)} \qquad \text{with} \qquad \phi_y(\xi) = \frac{2\pi f_0 \xi}{U_c} + cst \tag{14}$$

with ϕ_y being the phase of y at frequency f_0.
In particular, suppose we know the phase (at frequency f_0) at two positions $\xi = \xi_1$ and $\xi = \xi_2$, with a separation $\Delta\xi = \xi_2 - \xi_1$. Then we have:

$$\Delta\phi_y(f_0) = \phi_y(\xi_2) - \phi_y(\xi_1) = 2\frac{\pi f_0 \Delta\xi}{U_c} \tag{15}$$

This gives the convection speed:

$$\boxed{U_c = 2\frac{\pi f_0 \Delta\xi}{\Delta\phi_y(f_0)}} \tag{16}$$

The positions ξ_1 and ξ_2 correspond to 2 different positions of the flying hot wire, as shown in figure 2. Position ξ_1 corresponds to file `signaux1.lvm`, while position ξ_2 corresponds to file `signaux2.lvm`.

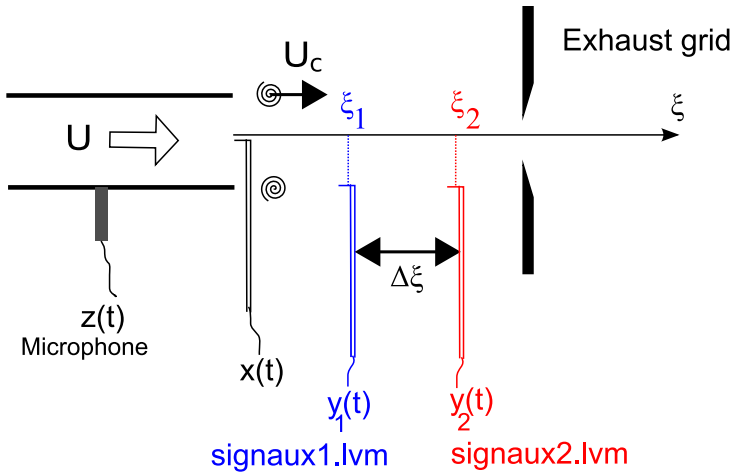


Figure 2: Two different measuring positions.

The geometrical setup is known, with $\Delta\xi$=6mm. This leaves the phase difference $\Delta\phi_y$ to be determined. This is where we need the cross-PSD. The phase difference is indeed given by the argument of the cross-PSD:

$$\boxed{\Delta\phi_y(f_0) = \phi_y(\xi_2) - \phi_y(\xi_1) = arg[S_{y_1x}(f_0)] - arg[S_{y_2x}(f_0)]} \tag{17}$$

Notice that there are two different measurements but that the fixed hot wire, $x(t)$, provides a phase reference that allows making the connection between these two measurements. The quantity $\phi_y(\xi_2) - \phi_y(\xi_1)$ can thus be determined, while the phase ϕ_x does not appear in the final result.

Note: why is the phase of 2 signals linked to the cross-PSD?
Ignoring the random nature of the signals, we may write for the Fourier transform of x:

$$X(f) = |X(f)|e^{j\phi_x(f)}$$

For the signal y:

$$Y(f) = |Y(f)|e^{j\phi_y(f)}$$

and thus

$$Y^*(f) = |Y(f)|e^{-j\phi_y(f)}$$

Now the cross-spectrum is:

$$S_{yx}(f) = Y^*(f)X(f) = |X(f)||Y(f)|e^{j(\phi_x(f)-\phi_y(f))}$$

This gives:

$$\phi_x(f) - \phi_y(f) = arg[S_{yx}(f)]$$

This result was used in Eq. (17).

4 Work to be done for Part II

Now open the main program for part II: `go_part_II.m`.
At the beginning of this program, either the file `signaux1.lvm` or the file `signaux2.lvm` is read. These files contain the data for two different positions, ξ_1 and ξ_2, of the mobile hot wire. At the moment, one can load `signaux1.lvm` and work only with this file.

a) Plot in figure 1 and compare the signals $x(t)$ and $y(t)$ over a time span of 50ms.
b) Compute the mean velocity of signal $x(t)$ and $y(t)$ (the mean of $x(t)$ is actually U). Remove these means from the signals.
c) Compute and display in figure 2 the auto-PSDs S_{xx}, S_{yy}, and S_{zz}. Use Welch's method. Display the result in dB (plot 10.log(S_{xx}) for example).
What is the frequency f_0 that has an high energy content $S_{yy}(f_0)$? This frequency corresponds to vortices passing by in the shear layer (a vortex shedding frequency). This is the frequency appearing in Eq. (13).
d) Compute the cross power spectral density S_{yx} of x and y (open and complete the matlab function `fct_iwelch(u,v,fs,N,iwindow)`. Display the modulus of S_{yx} as a function of frequency in figure 3a and its phase in figure 3b.
e) We would like to estimate the convection velocity of the vortices in the shear layer by using Eqs. (16) and (17). For this you need to work sequentially with files `signaux1.lvm` and `signaux2.lvm`. First open `signaux1.lvm` and compute the cross-PSD S_{y_1x}. Determine its phase, arg$[S_{y_1x}](f_0)$ at the frequency f_0. Then work with file `signaux2.lvm` and determine arg$[S_{y_2x}](f_0)$ at the frequency f_0. From these you can determine the convection speed (remember that $\Delta\xi$=6mm).
f) Determine the coherence $Coh_{yx}(f)$ of x and y, and plot its modulus in figure 3c. Remember that the coherence is nothing but a normalized cross-PSD, and so it is very easily determined using the functions `fct_welch` and `fct_iwelch`. Determine S_{zx}. Plot its modulus and phase in figure 4a and 4b. Plot the modulus of the coherence Coh_{zx} in figure 4c. What information does coherence bring to you?

Advanced Signal Processing - LAB 2

Discrete-Time Continuous Wavelet Transform

Application to intermittent event detection using Wavelets

D. Marx

Note:
You need to write a report on this lab. The name of the file containing the report should be:

FAMILYNAME_Firstname_Lab_labnumber.pdf

1 Exercise 1 - Plotting the wavelets

The objective here is to get familiar with the wavelets. Several functions are available:
- `fct_mother_wavelet(t,mother,param)` : returns the mother wavelet, $\psi(t)$.
- `fct_mother_wavelet_hat(f,mother,param)` : returns the mother wavelet Fourier transform, $\hat{\psi}(f)$.
- `fct_daughter_wavelet(t,u,s,mother,param)` : returns the daughter wavelet, $\psi_{us}(t)$.
- `fct_daughter_wavelet_hat(f,u,s,mother,param)` : returns the daughter wavelet Fourier transform, $\hat{\psi}_{us}(f)$.

In these function, `mother` can be either `'Morlet'` or `'DOG'`. The input `param` is used to pass the value of ω_0 for the Morlet wavelet, or the order for the DOG wavelet. In this lab you need to set param=-1 to use the default value ω_0=6 for the Morlet wavelet, and an order 2 for the DOG wavelet, which in that case is the Mexican hat wavelet.

Create a separate .m file to answer the following questions.

1.a) Use the provided function `fct_daughter_wavelet(t,u,s,mother,param)` to plot the Morlet daughter wavelet in the time interval $[-10\ 10]$ for $u = 0$ and three different values of s (that is, reproduce the examples given during the lecture with, say, s=0.5, s=1, s=2).

1.b) Use the provided function `fct_daughter_wavelet_hat(f,u,s,mother,param)` to plot the Fourier Transform of the Morlet daughter wavelet in the frequency interval $[-5\ 5]$, for the same values of u and s as previously. Is the Morlet wavelet (approximately) analytic?

1.c) The matlab command `quad` allows computing function integrals. For example:
`quad(@(x) sin^2(x),a,b)` calculates the integral of $\sin^2(x)$ over the interval $[a\ b]$.
Using the `quad` command:
- Check that the Morlet wavelet and the Mexican hat wavelet satisfy the admissibility condition:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{1}$$

- For both wavelets, calculate the constant C_ψ given by:

$$C_\psi = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty \tag{2}$$

- Calculate the duration T_e of the mother Morlet and mother Mexican hat wavelets. Since the norm of the wavelets is $||\psi||$=1, and since their mean time is t_m=0, we can compute T_e^2 from:

$$T_e^2 = \int_{-\infty}^{\infty} t^2 |\psi(t)|^2 \mathrm{d}t \tag{3}$$

What about the duration of a daughter having s=2?

2 Exercise 2 - Using the Continuous Wavelet Transform (CWT) and its inverse

Available functions:
- `fct_scale()` : returns a vector of scales (containing the s_j, j=1...N_{scale}).

- `fct_cwt(x,...)` : returns an approximation of the continuous walevet transform (CWT) of `x` (in the notations used in the appendix, this is CWT $W^c(n,j)$), by applying the algorithm in table 1.
- `fct_cwt_fftrecons` : returns a fast reconstruction of the signal out of its CWT.

The main program is:
- `go_wavelet.m`.

2.a) Before going on, you need to complete program `go_wavelet.m` and specify the value `Cpsi` and `Te` for both the Morlet and DOG wavelets. These were calculated in exercise 1. Until you make this change, these values are set to `NaN` and the program will stop before completing.

2.b) Complete the function `fct_scale.m` that returns `SCALES`, the vector of scales, knowing that:
- the scale logarithm are usually equi-spaced, that is:
SCALES(j+1)=SCALES(j) $2^{1/\text{nvoice}}$ where `nvoice` is the number of sub-octaves per octave.
- the minimum scale, `smin`, is an input to the function.
- the approximate maximum scale, `smax`, is an input to the function.
Once this is done you should be able to compute the CWT of a Dirac (`type_signal`=1).

2.c) Use the signal `type_signal`=7. This is taken from [2] and is shown in Fig. 1. Analyze the CWT of this signal using both the Morlet and the Mexican hat wavelet.

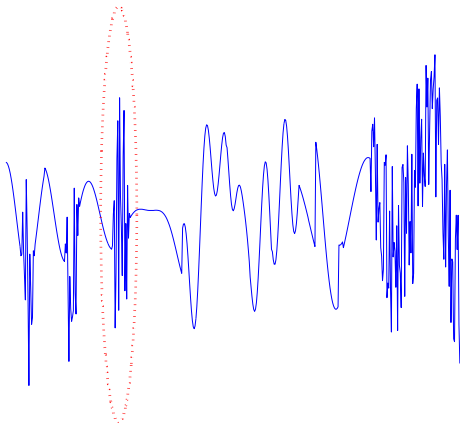


Figure 1: Signal 7. The red ellipse indicates the part of the signal that you need to denoise.

2.d) The function `fct_cwt_fftrecons.m` reconstructs the signal from its discrete-time CWT. This is implemented using Eqs. (31) and (33). The original signal and its reconstruction are compared in figure 4. They should match if SCALES covers most of the important scales in the signal (if necessary, change `smin` and `smax` until this is the case).
For the signal `type_signal`=7, you are asked to produce and plot in figure 4 a modified signal `xls` that is everywhere the same as the original signal except in the ellipse plotted in red in Fig. 1 where the rapid oscillations should be removed (this is a local denoising of the signal).

2.e) Again with `type_signal`=7, reconstruct two different signals:
- a signal x1 that contains only one scale close to s_1=1.7;
- a signal x2 that contains only one scale close to s_2=0.4;
In a single figure, plot the spectrum S_{xx} of the original signal, that $S_{x_1x_1}$ of x_1, and that $S_{x_2x_2}$ of x_2 (use the function `fct_raw` to compute PSDs; use the `semilogy` command for the plot, and represent only the positive frequency range). Did you expect the result?

3 Exercise 3 : A conditional averaging technique based on wavelet analysis

3.1 Introduction

Wavelets can be used to detect intermittent events in a signal. In the signal studied in this part, a few events have been hidden in an otherwise stochastic signal having a Gaussian distribution. At some scales (remember that for a turbulent signal intermittency is observed at the small scales) these events are going to modify the statistics of the signal. More specifically, the flatness of a Gaussian signal is 3, and a departure from this value results from the presence of the events. The technique consists of three steps:
1. Find a scale at which the added events will cause the flatness to depart from the baseline value of 3.
2. At this scale, the presence of events is indicated by large values of the local intermittency measure (LIM).
3. The conditional averaging technique finally provides the typical shape of an event by averaging small portions of the signal located around the times when an event is present.
These steps are typical of the procedure involved in studying the intermittency at small scales in turbulent flows using wavelets (see, for example, reference [3]).

3.2 Work to be done

The main program, in which all the changes are to be done, is `go_eventdetec.m`. The places where changes need to be done in the program are indicated by tags `Qa`, `Qb`, `Qc`, `Qd`, and `Qe`, according to the lettering of the questions below.

a) Check that the events do not modify the overall flatness of the signal (use the function `kurtss` defined in `go_eventdetec.m`). Take some time to observe the signal and its spectrum. Observation of either one does not allow one to say much about events in the signal.

b) Compute the scale-dependent flatness factor FF(s) for all the scales and find at which particular scale s_0 the value departs the most from its baseline value of 3. FF is given by:

$$FF(s) = \frac{\langle |W(t,s)|^4 \rangle_t}{\langle |W(t,s)|^2 \rangle_t^2} \tag{4}$$

c) Compute the local intermittency measure LIM(t,s). The LIM is given by:

$$LIM(t,s) = \frac{|W(t,s)|^2}{\langle |W(t,s)|^2 \rangle_t} \tag{5}$$

d) For the particular scale s_0 find the times at which LIM(t,s_0) exceeds some threshold value (test only those times that correspond to a local maximum of the LIM, which avoids multiple detections to be associated to one single event. Local maxima are found using the function `fct_detect_max`). Fill in the arrays `tdetec` containing the detection times, and the array `idetec` containing the indexes such that `t(idetec)=tdetec`.

e) The objective is now to determine the shape of the event. Educe this shape by performing an average of bits of the signal centered on the times detected in the previous question. The result should be put in the array `xcond`.

Clues:
- the number of events is less than 50, and your detection procedure will miss some of them.
- The educed event shape has a width which is less than 50 time units (a unit time step is used).

APPENDIX

These appendices provide a compact summary of the parts of the course whose knowledge is necessary for this lab. It also provides the details as to how implement a discrete form of the continuous wavelet transform.

A Background

Formulas for the continuous wavelet transform are listed in section A.1 and in particular it is recalled that the continuous wavelet transform can be expressed as a convolution product, see Eq. (12). This is used in section A.2 to compute a fast FFT-based discrete-time version of the transform, see Eq. (23). As recalled in section A.3, a signal can be reconstructed using its CWT. The reconstruction can be written as the integral of a convolution product, see Eq. (27). After discretization, it allows again a fast reconstruction in the discrete-time case, see Eqs. (31) and (33). See also reference [1].

A.1 The Continuous Wavelet Transform (CWT)

The Continuous Wavelet Transform is defined by:

$$W^c(u,s) = \int_{-\infty}^{\infty} x(t)\psi^*_{us}(t)dt \tag{6}$$

with

$$\psi_{us}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \tag{7}$$

where $\psi(t)$ is the mother wavelet, and $\psi_{us}(t)$ will be called the daughter wavelets, with s=0...∞, and u=-∞...∞. The upperscript ^{c} indicates continuous time (as opposed to discrete-time in the next section). The Fourier Transform of the daughter wavelet is:

$$\hat{\psi}_{us}(f) = TF[\psi_{us}(t)] = \sqrt{s}\hat{\psi}(sf)e^{-j2\pi fu} \tag{8}$$

Let us also recall the admissibility condition:

$$C_\psi = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f}df < \infty \tag{9}$$

which implies that:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{10}$$

CWT as a convolution product

We have seen that the CWT can be expressed as a convolution product. Defining $h_s(t)$ by:

$$h_s(t) = \frac{1}{\sqrt{s}}\psi^*\left(\frac{-t}{s}\right) \tag{11}$$

where h_s can be seen as an impulse response at scale s, we have indeed:

$$W^c(u,s) = \int_{-\infty}^{\infty} x(t)h_s(u-t)dt = (x * h_s)(u) \tag{12}$$

This relation will be useful when implementing a "fast" discrete CWT solver.
The Fourier Transform of $h_s(t)$ is:

$$\hat{h}_s(f) = TF[h_s(t)] = \sqrt{s}\hat{\psi}^*(sf) \tag{13}$$

where * stand for complex conjugaison.

CWT of a dirac
Let $x(t) = \delta(t - t_0)$. Its CWT is:

$$W^c(u,s) = \int_{-\infty}^{\infty} \delta(t - t_0)\psi^*_{us}(t)dt \tag{14}$$

Using the definition of a Dirac:

$$W^c(u,s) = \psi^*_{us}(t_0) = \frac{1}{\sqrt{s}}\psi^*\left(\frac{t_0 - u}{s}\right) \tag{15}$$

This relation can be used to check that the result provided by your CWT solver is correct.

A.2 Discrete-Time Continuous Wavelet Transform

Usually one works with a sampled signal $x(n)$, n=0...N-1 and a discrete time wavelet transform needs to be calculated. For this we take advantage of Eq. (12) which represents the wavelet transform as a convolution product. With discrete samples, the convolution product can be written as:

$$W(n,j) = (x * h_{s_j})(n) \quad \forall n = 0...N-1; \quad \forall j = 1...N_{\text{scale}} \tag{16}$$

where $*$ is now the discrete convolution between the two series $x(n)$ and $h_{s_j}(n)$, n=0...N-1. The scales at which the CWT is calculated belong to the discrete set $\{s_j, \quad j = 1...N_{\text{scale}}\}$.
Ideally, the discrete convolution product would be computed from:

$$(x * h_{s_j})(n) = \sum_{m=-\infty}^{\infty} x(m)h_{s_j}(n-m) \tag{17}$$

However, we have at our disposal a finite number of samples and an approximation of this is necessary. A simple (and fast) approximation is to suppose that our N-points series can be extended periodically, that is:

$$x(n \pm N) = x(n) \quad \forall n = 0...N-1$$

and likewise for $h_{s_j}(n)$. Then the convolution product becomes a circular convolution product:

$$W(n,j) = (x \otimes h_{s_j})(n) = \sum_{m=0}^{N-1} x(m)h_{s_j}(n-m) \quad \forall n = 0...N-1; \quad \forall j = 1...N_{\text{scale}} \tag{18}$$

And we have seen that the Fourier transform of a circular convolution product is the regular product of the discrete Fourier Transforms. This means that:

$$W(n,j) = (x \otimes h_{s_j})(n) = \underset{k\to n}{\text{FFT}^{-1}}\left[\underset{n\to k}{\text{FFT}}\left[x(n)\right] \cdot \underset{n\to k}{\text{FFT}}\left[h_{s_j}(n)\right]\right](n) \quad \begin{matrix}\forall n = 0...N-1 \\ \forall j = 1...N_{\text{scale}}\end{matrix} \tag{19}$$

This formula is the basis for a fast computation of the discrete time wavelet transform. Note that often the FT transform $\hat{h}_s(f)$ of $h_s(t)$ is known, since by Eq. (13) it depends of the Fourier Transform of the wavelet. Hence, FFT$[h_{s_j}(n)](k)$, k=0...N-1, can be readily obtained from it by using the relation (36) in appendix:

$$\underset{n\to k}{\text{FFT}}\left[h_{s_j}(n)\right](k) = \frac{1}{\Delta t}\hat{h}_{s_j}(f_k) = \frac{1}{\Delta t}\sqrt{s_j}\hat{\psi}^*(s_j f_k) \tag{20}$$

where f_k, k=0...N-1, is the usual frequency vector associated with the FFT, see Eqs. (34-35) in the appendix. Hence, the FFT of $h_{s_j}(n)$ is actually obtained analytically without resorting to any FFT algorithm. On the contrary, the FFT of $x(n)$ needs to be calculated using the matlab `fft` command.
Once the discrete time wavelet transform $W(n,j)$ is known, one would like to obtain the original continuous-time $W^c(n\cdot\Delta t, s_j)$. We simply have:

$$W^c(n\cdot\Delta t, s_j) = W(n,j)\cdot\Delta t \tag{21}$$

The multiplication by the sampling time comes from Eq. (39) in the appendix, which relates the discrete and continuous convolution products. Combining Eq. (19), Eq. (20), and Eq. (21) we are finally left with ($1/\Delta t$ and Δt cancel each other):

$$W(n,j) = \text{FFT}^{-1}\left[\text{FFT}\left[x\right]\cdot\sqrt{s_j}\hat{\psi}^*(s_j f_k)\right](n) \qquad \begin{array}{l}\forall n = 0...N-1\\ \forall j = 1...N_{\text{scale}}\end{array} \tag{22}$$

The equation is better understood in the following form:

$$\boxed{W(:,j) = \text{FFT}^{-1}\left[\text{FFT}\left[x\right]\cdot\sqrt{s_j}\hat{\psi}^*(s_j f_k)\right] \quad \forall j = 1...N_{\text{scale}}} \tag{23}$$

since it is recalled that the FFT produces a size N vector from a size N vector. It is also clear from this relation that the matrix W is going to be filled in column by column. Finally, to calculate the discrete-time continuous wavelet transform, the algorithm in table 1 is used. The cost of this algorithm is about $N_{\text{scale}}N\ln(N)$. That is: we have N_{scale} columns to fill in, and the cost per column is $N\ln(N)$ using FFT. This is a fast algorithm due to its using the FFT to calculate a convolution product. A direct calculation of the convolution product would require a higher cost: $N_{\text{scale}}N^2$.

Calculate the vector of frequencies fk using `fct_f`.
Calculate the vector xfft(:)=fft(x).
for all j **do**
Calculate the vector hsjfft(:)=$\sqrt{s_j}\hat{\psi}^*(s_j f_k)$.
(Use `fct_mother_wavelet_hat` or `fct_daughter_wavelet_hat`)
Calculate W(:,j)=ifft(xfft.hsjfft).
end for

Table 1: **Algorithm for Fast Discrete-Time Continuous Wavelet Transform.**

A.3 Signal reconstruction

A signal can be recovered from its CWT. The reconstruction formula is (slightly) different for the real and analytic wavelets.

Real Wavelet

For a real wavelet (e.g. Mexican hat) and a real signal, we have:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W^c(u,s)\psi_{us}(t) du \frac{ds}{s^2} \qquad (\psi \text{ real}) \tag{24}$$

that is:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W^c(u,s) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) du \frac{ds}{s^2} \tag{25}$$

By defining: $W_s^c(u) = W^c(u,s)$ and $\psi_s(t) = \psi(t/s)$, we have:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W_s^c(u)\psi_s(t-u) du \frac{ds}{s^{5/2}} \tag{26}$$

We recognize a convolution product:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty (W_s^c * \psi_s)(t) \frac{ds}{s^{5/2}} \tag{27}$$

When working with discrete times, periodicity of the time series is again assumed, and discretizing the integral leads to:

$$x \sim \frac{\Delta t}{C_\psi} \sum_{j=1}^{N_{\text{scale}}} (W_{s_j} \otimes \psi_{s_j})(n) \frac{ds_j}{s_j^{5/2}} \tag{28}$$

where $W_{s_j}(n) = W(n,j)$, the discrete-time CWT at scale index j and time index n. x is of course a size N vector. This formula involves a circular convolution product, which can be calculated efficiently using FFTs and FFT^{-1}s. Indeed we have:

$$W_{s_j} \otimes \psi_{s_j} = \underset{k \to n}{\text{FFT}^{-1}} \left[\text{FFT}[W_{s_j}](k) \cdot \text{FFT}[\psi_{s_j}](k) \right] \tag{29}$$

where k=0...N-1 is the frequency index. We know that:

$$\hat{\psi}_s(f) = FT[\psi(t/s)] = s \cdot FT[\psi](sf) = s\hat{\psi}(sf)$$

By the relation between the FT and FFT, Eq. (36), we thus have:

$$\text{FFT}[\psi_{s_j}](k) = s_j \hat{\psi}(s_j f_k) \frac{1}{\Delta t} \tag{30}$$

meaning that FFT$[\psi_{s_j}](k)$ can be computed from an analytic formula.
Finally, combining Eq. (28), Eq. (29), and Eq. (30), we obtain:

$$\boxed{x \sim \frac{1}{C_\psi} \sum_{j=1}^{N_{\text{scale}}} \underset{k \to n}{\text{FFT}^{-1}} \left[\underset{n \to k}{\text{FFT}}[W_{s_j}](k) \cdot \hat{\psi}(s_j f_k) \right] \frac{ds_j}{s_j^{3/2}}} \tag{31}$$

The factor FFT$[W_{s_j}]$ is the FFT of the jth column of the CWT matrix, the column corresponding to a fixed scale s_j. This factor is calculated using the `fft` matlab command.

Analytic Wavelet

For an analytic wavelet (e. g. Morlet) and a real signal, we have:

$$x(t) = \frac{2}{C_\psi}\mathrm{Re}\left[\int_0^\infty \int_{-\infty}^\infty W^c(u,s)\psi_{us}(t)du\frac{ds}{s^2}\right] \qquad (\psi \text{ analytic}) \tag{32}$$

The discrete reconstruction formula for an analytic wavelet reads:

$$\boxed{x \sim \frac{2}{C_\psi}\mathrm{Re}\left[\sum_{j=1}^{N_{\text{scale}}} \underset{k\to n}{\mathrm{FFT}^{-1}}\left[\underset{n\to k}{\mathrm{FFT}}[W_{s_j}](k)\cdot\hat{\psi}(s_j f_k)\right]\frac{ds_j}{s_j^{3/2}}\right]} \tag{33}$$

B Relation between the DFT (computed using FFT) and the FT of a signal

Let $x_c(t)$ be a signal depending on continuous time t, and $X_c(f)$ be its Fourier Transform. Let $x(n) = x_c(n\Delta t)$ for n=0...N-1 be the N samples of the sampled signal; these samples correspond to instants $t_n = n\cdot\Delta t$ where Δt is the sampling period. Its Discrete Fourier Transform (DFT) $X(k)$ is:

$$X(k) = FFT[x(n)] = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N}$$

for $k = 0...N-1$. The index k is used for frequency, and there exists a vector f_k giving the frequency as a function of k.

For N even:

$$f_k = \begin{cases} \dfrac{k}{N\Delta t} & \text{for} \quad k = 0...\dfrac{N}{2}-1 \\ -f_{N-k} & \text{for} \quad k = \dfrac{N}{2}...N-1 \end{cases} \tag{34}$$

For N odd:

$$f_k = \begin{cases} \dfrac{k}{N\Delta t} & \text{for} \quad k = 0...\dfrac{N-1}{2} \\ -f_{N-k} & \text{for} \quad k = \dfrac{N-1}{2}+1...N-1 \end{cases} \tag{35}$$

Note: in matlab, the integer k takes values between 1 and N (instead of 0 and N-1), meaning that k in the formulas above needs to be replaced by $k-1$. The provided function `fct_f(N,Δt)` returns a vector containing the f_k whatever the parity of N.

There exists a relation between $X(k)$ and $X_c(f)$. This is obtained by writing:

$$X_c(f_k) = \int_{-\infty}^{\infty} x_c(t)e^{-j2\pi f_k t}dt = \int_0^D x_c(t)e^{-j2\pi f_k t}dt \sim \Delta t \underbrace{\sum_{n=0}^{N-1} x(n)e^{-j2\pi nk/N}}_{X(k)}$$

where $D \sim N\Delta t$ is the recorded time. **The relation between the continuous FT and the discrete FFT is therefore:**

$$\boxed{X_c(f_k) \sim \Delta t X(k) \quad \forall k = 0...N-1} \tag{36}$$

The FT at frequency f_k is the FFT at index k multiplied by the sampling time.

C Relation between the continuous convolution and the discrete convolution

The discrete convolution of two series $x(n)$ and $y(n)$ is:

$$(x * y)(n) = \sum_{m=-\infty}^{\infty} x(m) h_s(n-m) \tag{37}$$

The continuous convolution of the continuous signals is:

$$(x * y)(\tau) = \int_{-\infty}^{\infty} x(\tau) y(t-\tau) d\tau \tag{38}$$

and thus, by discretizing the integral:

$$(x * y)(\tau = n\Delta t) \sim \Delta t \sum_{m=-\infty}^{\infty} x(m) h_s(n-m) \sim \Delta t \cdot (x * y)(n)$$

where Δt is the sampling time. The continuous convolution is thus the discrete time convolution multiplied by the sampling time. Since the circular convolution is an approximation of the discrete convolution, the continuous convolution is also equal to the discrete circular convolution multiplied by the sampling time, that is:

$$\boxed{(x * y)(\tau = n\Delta t) \sim \Delta t \cdot (x \otimes y)(n) \quad \forall n = 0...N-1} \tag{39}$$

As for Eq. (36), the multiplication by the sampling time Δt in Eq. (39) results from dicretizing an integral.

REFERENCES

[1] C Torrence, G. Compo, A Practical Guide to Wavelet Analysis, Bulletin of the American Meteorological Society 79, 61-78, 1998. Article and matlab programs available from the net.

[2] H. Nobach, C. Tropea, L. Cordier, J.-P. Bonnet, J. Delville, J. Lewalle, M. Farge, K. Schneider, R. Adrian. Review of some fundamentals of data processing, in *Springer Handbook of experimental fluid mechanics* (chapter 22), Tropea, Yarin, Foss (Eds.), Springer, 2007.

[3] R. Camussi, G. Guj, Orthonormal wavelet decomposition of turbulent flows: intermittency and coherent structures, J. Fluid Mech. 348, 177–199, 1997.

Advanced Signal Processing - LAB 3

Proper Orthogonal Decomposition

—

D. Marx

In this lab, we want to put into practice the SVD and the POD. The SVD is first used to compute low-rank approximations of an image. Then, we consider numerical data for a 2D shear layer and compute the POD modes using the snapshot method. Compared with the material presented during the course (where we had a single component (1C) along a single direction (1D)) we have here a velocity field with two components (2C) in two directions (2D). This just changes the way the data matrix is filled in. This is explained in the "background part" below.

Note:
You need to write a report on this lab. The name of the file containing the report should be:

FAMILYNAME_Firstname_Lab_labnumber.pdf

Background on the snapshot POD method

The POD for a 2D2C (two-dimension, two-component) velocity vector is performed in this lab. This is a straightforward extension to the 1D1C approach presented in the course, and is specified below for more clarity.

Continuous case:
The 2D2C velocity vector:

$$\mathbf{u}(x,y,t) = \begin{bmatrix} u(x,y,t) \\ v(x,y,t) \end{bmatrix} \tag{1}$$

is considered. The number of components is noted N_c=2. For such a vector, the spatial scalar product is:

$$(\mathbf{u_1},\mathbf{u_2}) = \int_0^{L_x}\int_0^{L_y} (u_1(x,y,t)u_2(x,y,t) + v_1(x,y,t)v_2(x,y,t))\,\mathrm{d}x\mathrm{d}y \tag{2}$$

Defining the time average by $< \cdot >= \frac{1}{T}\int_0^T \cdot \mathrm{d}t$ (an ensemble average could be used as well), the time-averaged energy is:

$$E = \frac{1}{2} < (\mathbf{u},\mathbf{u}) >= \frac{1}{2}\frac{1}{T}\int_0^T\int_0^{L_x}\int_0^{L_y} \left(u^2(x,y,t) + v^2(x,y,t)\right)\mathrm{d}x\mathrm{d}y\mathrm{d}t \tag{3}$$

The expansion of $\mathbf{u}$ in POD modes is written:

$$\mathbf{u}(x,y,t) = \sum_{i=1}^{\infty} a_i(t)\boldsymbol{\psi}_i(x,y) \tag{4}$$

where any POD mode $\boldsymbol{\psi}_i$ has the same shape as $\mathbf{u}$, that is, it is a two-component vector:

$$\boldsymbol{\psi}_i(x,y) = \begin{bmatrix} \psi_{u,i}(x,y) \\ \psi_{v,i}(x,y) \end{bmatrix} \tag{5}$$

There is an infinite number of modes for the continuous problem that are obtained by solving the Fredholm eigenvalue problem:

$$\int_0^{L_x}\int_0^{L_y} \mathbf{R}(x,y,x',y')\boldsymbol{\psi}_i(x',y')\mathrm{d}x'\mathrm{d}y' = \lambda_i\boldsymbol{\psi}_i(x,y)$$

where the correlation tensor $\mathbf{R}(x,y,x',y')$ is given by:

$$\mathbf{R}(x,y,x',y') = \left\langle \mathbf{u}(\mathbf{x},\mathbf{y}) \otimes \mathbf{u}(\mathbf{x}',\mathbf{y}') \right\rangle = \begin{bmatrix} \langle u(x,y)u(x',y')\rangle & \langle u(x,y)v(x',y')\rangle \\ \langle v(x,y)u(x',y')\rangle & \langle v(x,y)v(x',y')\rangle \end{bmatrix} \tag{6}$$

The properties of the POD modes are unchanged:
• because the modes are orthonormal, the time coefficients are given by the relation:

$$a_j(t) = \left(\mathbf{u}(x,y,t), \boldsymbol{\psi}_j(x,y)\right) \tag{7}$$

• the time coefficients verify the uncorrelatedness relation:

$$< a_i(t)a_j(t) >= \lambda_i \delta_{ij} \tag{8}$$

where the λ_i are the eigenvalues of the Fredholm equation.
• the energy verifies:

$$E = \sum_{i=1}^{\infty} \lambda_i \tag{9}$$

Discrete case:
N snapshots of the velocity vector are measured at times t_k, $k = 1...N$. For every snapshot, the velocity vector is measured:
• at N_x axial positions x_i, $i = 1...N_x$.
• at N_y transverse positions y_j, $j = 1...N_y$.
In the notations of the course, the number of "spatial positions" is M=$N_c N_x N_y$=$2N_x N_y$. For convenience, let us define M_2=$M/2$=$N_x N_y$. This is a the actual number of spatial positions (the dimension of a snapshot at a given time). It is assumed (as is generally the case for velocity fields obtained by Particle Image Velocimetry or by numerical simulations) that $M \gg N$, so that the **snapshot method** is used. This method relies on the diagonalization of a NxN matrix (rather than a MxM matrix for the regular method) and provides a finite number, N, of modes. There are as many modes as snapshots.

Here are the steps to be followed:

☐ **Step 1.** Arrange the data in the **data matrix** A of size M x N.

$$A = \begin{bmatrix} u(\text{pos}_1, t_1) & u(\text{pos}_1, t_2) & \cdots & u(\text{pos}_1, t_N) \\ u(\text{pos}_2, t_1) & u(\text{pos}_2, t_2) & \cdots & u(\text{pos}_2, t_N) \\ \vdots & \vdots & \vdots & \vdots \\ u(\text{pos}_{M_2}, t_1) & u(\text{pos}_{M_2}, t_2) & \cdots & u(\text{pos}_{M_2}, t_N) \\ v(\text{pos}_1, t_1) & v(\text{pos}_1, t_2) & \cdots & v(\text{pos}_1, t_N) \\ v(\text{pos}_2, t_1) & v(\text{pos}_2, t_2) & \cdots & v(\text{pos}_2, t_N) \\ \vdots & \vdots & \vdots & \vdots \\ v(\text{pos}_{M_2}, t_1) & v(\text{pos}_{M_2}, t_2) & \cdots & v(\text{pos}_{M_2}, t_N) \end{bmatrix} \tag{10}$$

where pos_m, m=$1...N_x N_y$, is one of the $M_2 = N_x N_y$ spatial positions. A 2D2C vector at any given time thus becomes a 1D column in the matrix A. It is necessary to establish a connection between a position in the two-dimensional array ($X(i)$,$Y(j)$) and the corresponding position in the one-dimensional array (pos_m). Namely, the value of $u(X(i), Y(j), t_k), \forall t_k$, is to be found in A at row $\text{pos}_m = (i-1)N_y + j$, $\forall i = 1...N_x, \forall j = 1...N_y$. Equivalently, one can say that the row m in matrix A corresponds to the indexes (i, j) in u with:

$$j = \text{mod}(m, N_y) \text{ and } i = \frac{j-m}{N_y} + 1$$

The matrix form for the POD expansion is:

$$A = \Psi Q \tag{11}$$

with

$$\Psi = \begin{bmatrix} \psi_{u,1}(\text{pos}_1) & \psi_{u,2}(\text{pos}_1) & \cdots & \psi_{u,N}(\text{pos}_1) \\ \psi_{u,1}(\text{pos}_2) & \psi_{u,2}(\text{pos}_2) & \cdots & \psi_{u,N}(\text{pos}_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{u,1}(\text{pos}_{M_2}) & \psi_{u,2}(\text{pos}_{M_2}) & \cdots & \psi_{u,N}(\text{pos}_{M_2}) \\ \psi_{v,1}(\text{pos}_1) & \psi_{v,2}(\text{pos}_1) & \cdots & \psi_{v,N}(\text{pos}_1) \\ \psi_{v,1}(\text{pos}_2) & \psi_{v,2}(\text{pos}_2) & \cdots & \psi_{v,N}(\text{pos}_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{v,1}(\text{pos}_{M_2}) & \psi_{v,2}(\text{pos}_{M_2}) & \cdots & \psi_{v,N}(\text{pos}_{M_2}) \end{bmatrix} \qquad Q = \begin{bmatrix} a_1(t_1) & a_1(t_2) & \cdots & a_1(t_N) \\ a_2(t_1) & a_2(t_2) & \cdots & a_2(t_N) \\ \vdots & \vdots & \vdots & \vdots \\ a_N(t_1) & a_N(t_2) & \cdots & a_N(t_N) \end{bmatrix} \tag{12}$$

where Ψ is the mode matrix, which has the modes in its columns, and Q is the matrix of the time coefficients. The i-th column of Ψ corresponds to the mode $\boldsymbol{\psi}_i$. The i-th row of Q corresponds to the time coefficient for the mode $\boldsymbol{\psi}_i$.
The objective now is to find Ψ and Q using the optimality requirement, which is implicit in the SVD process.

☐ **Step 2.** Compute the correlation matrix:

$$C = \frac{1}{N} A^+ A \tag{13}$$

whose size is N x N. This matrix is hermitian.
Remark: in the course, the matrix to be diagonalized was A^+A (without the extra $1/N$ factor). This is just a matter of how energy is computed. The present form is the standard one in POD. Note that $C_{ij} = \frac{1}{N}(u_i, u_j)$, where $(.,.)$ denotes the spatial scalar product. Writing $C = \frac{1}{N}A^+A$ is indeed a slight approximation to the scalar product. In particular it requires a uniform mesh. In a more correct implementation, C_{ij} should be a numerical approximation to $\frac{1}{N}(u_i, u_j)$ using some standard numerical method of integration for computing the scalar product (e.g. Simpson's rule).

☐ **Step 3a.** Diagonalize the correlation matrix, and obtain the matrix S of size N x N containing the N eigenvalues λ_i as well as the matrix V of the N (column) eigenvectors.

☐ **Step 3b.** Order S and V **in descending order of** λ_i: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ....$

☐ **Step 4a.** Calculate the spatial mode matrix $\Psi = AV$ of size M x N.

☐ **Step 4b.** The columns of Ψ should be orthogonal at this stage, but not orthonormal. This is why the columns should be re-normalized so that the norm of every column is 1. The columns in Ψ are then orthonormal vectors, and these vectors are nothing but the spatial POD modes.

☐ **Step 5.** Compute the matrix of the time coefficients of the modes by calculating:

$$Q = \Psi^+ A \tag{14}$$

A further step can be taken:
☐ **Step 6.** The modes are stored in Ψ as columns, that is, are stored in a 1D format. It may be

convenient to store them in a 2D2C format. The data in Ψ are then unwrapped in the inverse order of step 1.

The POD calculation is now completed. One has calculated:
- a set of N spatial orthonormal vectors $\boldsymbol{\psi}_i$ (the columns of Ψ);
- a set of N eigenvalues λ_i that represent the energy of the velocity field contained in the i-th mode (a rapid decrease in λ_i is necessary for a POD decomposition to be useful ultimately);
- a set of times coefficients a_i for the modes (the rows of Q).

Low-rank approximation

The discrete form of the Proper Orthogonal Decomposition in Eq. (4) is:

$$A = \Psi Q \tag{15}$$

It is possible to produce a low-rank approximation A' of A by computing:

$$A' = \Psi' Q' \tag{16}$$

where Ψ' is obtained by nullifying the last columns in Ψ and Q' is obtained by nullifying the corresponding last rows in Q. Of course, the suppressed columns and rows correspond to the smaller values of λ_i.

Note: in the following, only real data are considered, and the adjoint A^+ of a matrix is the transpose of this matrix.

Work to be done

Exercise 1: SVD

An example of low-rank approximation of a matrix was shown during the course (the clown example). You have to produce this example by yourself, using the `svd` command of matlab. A script `go_clown` is provided that loads the image and set up the figures.
Here are some points you need to consider:
- What are the size and the rank of the original matrix?
- Plot the singular values of this matrix (in figure 1).
- Determine how many singular values are needed to retain 90% of the energy in the signal?
- Plot three approximations of the original matrix (in figure 2): a rank 10 approximation, a rank 20 approximation, a rank 70 approximation. What is the Frobenius norm of the difference between the original matrix and its rank 70 approximation (use the `norm` command in Matlab)? Can you obtain this by considering the singular values of the original matrix?

Exercise 2: POD of 2D shear layer velocity fields

The snapshots obtained during a direct numerical simulation of a 2D low Reynolds shear layer are considered. A typical velocity field may be checked by using `go_plot_data.m`.

The main program for computing the POD of these snapshots is `go_snaphot.m`. It loads a series of data (N=100 snapshots) and calls two functions:
- `fct_snapshot`: this function performs the POD and returns the eigenvalues, the modes, the data matrix, etc.
- `fct_recons`: this function uses the POD modes and the time coefficients to compute a low-rank approximation of the original data. Namely, it makes use of N_{recons} modes to reconstruct the velocity field.

Your main work is to complete the function `fct_snapshot.m` so that it performs the 2D2C snapshot POD according to the steps given in the background part:
☐ Step 1: calculation of the data matrix (2D2C data are stored in 1D columns).
☐ Step 2: calculation of the correlation matrix.
☐ Step 3a/3b: diagonalization of the correlation matrix/re-ordering (in descending order of the eigenvalues).
☐ Step 4a/4b: calculation of the POD modes/normalization of the columns.
☐ Step 5: calculation of the time coefficients.
☐ Step 6: the modes stored in 1D columns are reshaped into the 2D2C form (inverse of step 1).

Steps 1 and 6 are ready to use in the script. The other steps need to be completed.
Once the function `fct_snapshot.m` is completed, go back to the main program, and eventually make some changes there (where indicated) to answer the following questions:

1. Check that the POD modes (columns in Ψ) are orthonormal.

2. Plot the eigenvalues in figure 1 (top). What particularity of the clustering of the eigenvalues do you observe? Check that the sum of the eigenvalues is equal to the energy of the signal (that is, check the discrete equivalent of Eq. (9)).

3. It is interesting to consider the following ratio (normalized cumulative sum):

$$r(n) = \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \tag{17}$$

This is the ratio of the energy contained in the first n modes to the total energy of the signal. Compute and plot this ratio in figure 1 (bottom). How many modes are needed to account for at least 90% of the energy of the signal?

4. The first eight POD modes are shown in figure 2. What observation can you make? (observe modes 1 and 2, and modes 3 and 4,...)

5. Verify the discrete equivalent of Eq. (8). Remember that in the discrete case, we have for the time average: $<>\sim \frac{1}{N}\sum_{i=1}^{N}$.

6. In figure 3 (top), plot the time evolution of $a_1(t)$ and $a_2(t)$. What is the phase difference between them? In figure 3 (bottom), plot the time evolution of $a_3(t)$ and $a_4(t)$. What is the phase difference between them?

7. In figure 4, produce the phase plot a_2/a_1, a_3/a_1, a_5/a_1, and a_4/a_3? What do you observe?

8. Use the function `fct_recons.m` to compute a low rank approximation A_{recons} of the data matrix A. The reconstruction is performed using N_{recons} modes. What should be the value of N_{recons} to retain in A_{recons} 90% of the energy in A? Plot in figure 5 the colormap of the original axial velocity and that of its low rank approximation at some given time. Plot in figure 6 the axial velocity as a function of time at some spatial station for both the original data matrix and its low rank-approximation.

APPENDIX

Table 1 provides some commands in Matlab that are useful for dealing with vectors and matrices. You can obtain help on any command `command` using `help command` of `doc command` in the matlab prompt.

A Matlab commands

Command	Description
`eig(A)`	Compute the eigenvalues and eigenvectors of a square matrix. Example: `[V,D]=eig(A)` returns: the diagonal matrix D containing the eigenvalues, and the matrix V containing the eigenvector (columns of V), with: $MV = VD$.
`svd(A)`	Compute the Singular Value Decomposition of a matrix. Example: `[U,S,V]=svd(A)` returns: the matrix U of the left-singular vectors, the diagonal matrix S of the singular values, and the matrix V of the right-singular vectors.
`diag(A)`	Applied to a matrix A: returns a vector with the diagonal elements, Applied to a vector A: returns a diagonal matrix whose diagonal is the input vector.
`A'` `transpose(A)`	Adjoint of the matrix A (transpose if A real, conjugate transpose if A complex.) Transpose of the matrix A (for a real or complex matrix).

`A=zeros(N)`
`A=zeros(M,N)`
`A=zeros(N,1)`
`A=zeros(1,N)`
	Creates a square matrix of size N initialized with zeros.
	Creates a rectangular matrix of size M x N initialized with zeros.
	Creates a column vector of size N initialized with zeros.
	Creates a row vector of size N initialized with zeros.

Table 1: Some useful matlab commands for vectors and matrices.

REFERENCES

[1] H. Nobach, C. Tropea, L. Cordier, J.-P. Bonnet, J. Delville, J. Lewalle, M. Farge, K. Schneider, R. Adrian. Review of some fundamentals of data processing, in *Springer Handbook of experimental fluid mechanics* (chapter 22), Tropea, Yarin, Foss (Eds.), Springer, 2007.

[2] P. Holmes, J.L. Lumley, G. Berkooz: Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge Monographs on Mechanics, Chapter 3, Cambridge University Press, Cambridge, 1996.

[3] A. Chatterjee, An introduction to the proper orthogonal decomposition, Current Science **78**(7) (2000) 808-817.

[4] T. Smith, J. Moehlis, P. Holmes, Low-Dimensional Modelling of Turbulence Using the Proper Orthogonal Decomposition: A tutorial, Nonlinear Dynamics **41** (2005) 275-307.

Bibliography

[1] Addison, P. S. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance.* CRC PRESS, 2002.

[2] Baars, W. J., Talluru, K. M., Hutchins, N., and Marusic, I. Wavelet analysis of wall turbulence to study large-scale modulation of small scales. *Exp. in Fluids* **56:188** (2015), 1–15.

[3] Bendat, J., and Piersol, A. *Random data: Analysis and Measurement Procedures, 2nd Edition.* Wiley-Interscience, 1986.

[4] Berkooz, G., Holmes, P., and Lumley, J. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics* **25** (1993), 539–575.

[5] Boashash, B. Estimating and interpreting the instantaneous frequency of a signal, part I: Fundamentals. *Proc. IEEE* **80** (1992), 520–538.

[6] Boashash, B. *Time frequency signal analysis and processing - a comprehensive reference.* Elsevier, 2003.

[7] Boyce, W. E., and Prima, R. C. D. *Elementary Differential Equations and Boundary Value Problems, 7th Edition.* John Wiley and Sons, 2001.

[8] Bracewell, R. N. *The Fourier transform and its application, 3rd Edition.* Mac Graw Hill, 2000.

[9] Brigham, E. *The Fast Fourier Transform.* Prentice-Hall, 1974.

[10] Camussi, R. Coherent structures identification from wavelet analysis of particle image velocimetry data. *Exp. in Fluids* **32** (2002), 76–86.

[11] Camussi, R., and Guj, G. Orthonormal wavelet decomposition of turbulent flows: intermittency and coherent structures. *J. Fluid Mech.* **348** (1997), 177–199.

[12] Chatterjee, A. An introduction to the proper orthogonal decomposition. *Current Science* **78** (2000), 808–817.

[13] Cohen, L. Time-frequency distributions - a review. *Proc. IEEE* **77** (1989), 941–981.

[14] Cohen, L. *Time frequency analysis.* Prentice Hall, 1995.

[15] Farge, M. Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics* **24** (1992), 395–457.

[16] Flandrin, P. *Time-frequency / Time-scale analysis.* Academic Press, 1999.

[17] Holmes, P., Lumley, J., and Berkooz, G. *Turbulence, coherent structures, dynamical systems, and symmetry.* Cambridge University Press, 1998.

[18] Katul, G., Parlange, M., and Chu, C. Intermittency, local isotropy, and non-Gaussian statistics in atmospheric surface layer turbulence. *Phys. Fluids* **6***(7)* (1994), 2480–2492.

[19] Lewalle, J., Low, K., and Glauser, M. Properties of the far-field pressure signatures of individual jet noise sources. *J. Aeroacoustics* **11** (2012), 651–674.

[20] Mallat, S. *A wavelet tour of signal processing, 3rd Edition.* Academic Press, 2009.

[21] Marusic, I., McKeon, B. J., Monkewitz, P. A., Nagib, H. M., Smits, A. J., and Sreenivasan, K. R. Wall-bounded turbulent flows at high Reynolds numbers: recent advances and key issues. *Phys. Fluids* **22** (2010), 065103, 1–24.

[22] Marx, D., Aurégan, Y., Bailliet, H., and Valière, J.-C. PIV and LDV evidence of hydrodynamic instability over a liner in a duct with flow. *J. Sound Vib.* **329** (2010), 3798–3812. *Note: POD for this experiment available in: Marx et al, Sound amplification in a lined duct with flow: PIV measurements, proceedings of Acoustics'08, Paris, 2008.*

[23] Mathis, R., Hutchins, N., and Marusic, I. Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers. *J. Fluid Mech.* **628** (2009), 311–337.

[24] Meneveau, C. Analysis of turbulence in the orthonormal wavelet representation. *J. Fluid Mech.* **232** (1991), 469–520.

[25] Oppenheim, A., and Schafer, R. *Digital Signal Processing.* Prentice-Hall Int., 1975.

[26] Papoulis, A., and Pillai, S. *Random variables and stochastic processes.* Mc Graw Hill, 2002.

[27] Pope, S. *Turbulent flows.* Cambridge University Press, 2000.

[28] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. *Numerical recipes in C - the art of scientific programming, 2nd edition.* Cambridge University Press, 1992.

[29] Schlatter, P., and Orlu, R. Quantifying the interaction between large and small scales in wall-bounded turbulent flows: a note of caution. *Phys. Fluids* **22** (2010), 051704, 1–4.

[30] Schram, C., Rambaud, P., and Riethmuller, M. L. Wavelet based eddy structure eduction from a backward facing step flow investigated using particle image velocimetry. *Exp. Fluids* **36** (2004), 233–245.

[31] She, Z.-S., Jackson, E., and Orszag, S. A. Intermittent vortex sctructures in homogeneous isotropic turbulence. *Nature 344* (1990), 226–228.

[32] Sirovich, L. Turbulence and the dynamics of coherent structures, part I-III. *Quaterly of Applied Math.* **XLV** (1987), 561–582.

[33] Smith, T., Moehlis, J., and Holmes, P. Low-dimensional modelling of turbulence using the proper orthogonal decomposition: a tutorial. *Nonlinear Dynamics* **41** (2005), 275–307.

[34] Sreenivasan, K. On the fine-scale intermittency of turbulence. *J. Fluid Mech.* **151** (1985), 81–103.

[35] Strang, G. *Introduction to linear algebra, 5th edition.* Wellesley - Cambridge Press, 2016.

[36] Tardu, S. Stochastic synchronization of the near wall turbulence. *Phys. Fluids* **20** (2008), 045105–1–045105–8.

[37] Tennekes, H., and Lumley, J. *A first course in turbulence.* MIT Press, 1972.

[38] Towne, A., Schmidt, O. T., and Colonius, T. Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.* **847** (2018), 821–867.

[39] Trefethen, L. N., and Bau, D. *Numerical linear algebra.* Society for Industrial and Applied Mathematics, 1997.

[40] Van-Den-Enden, W. M., and Verhoekx, N. A. M. *Discrete-time signal processing : an introduction.* Prentice-hall, New York, 1989.

[41] Ville, J.-M. Théorie et applications de la notion de signal analytique. *Câbles et Transmission* **2A** (1948), 61–74. In French. English translation: I. Selin, Theory and applications of the notion of complex signals, Rand Corporation Report T-92, Santa Monica, 1958.

[42] Whittaker, E. T., and Watson, G. N. *A course of modern analysis, 2nd Edition.* Cambridge University Press, 1915.