Fall 2025

Data Science

or rather introduction to signals and systems

Ferréol Soulez
ferreol.soulez@univ-lyon1.fr

This document is written in Typst. The source files are available on GitHub: FerreolS/Signals-and-systems-m1-tpe. Please file an issue or submit a pull request for any typos or misunderstandings.

Avant propos

This lecture was initially supposed to be about Data Science. Data science is “a concept to unify statistics, data analysis, informatics, and their related methods” to “understand and analyze actual phenomena with data” (Hayashi, 1998). To learn about “data science”, it became clear to me that knowledge of basic signal processing concepts is essential. This is why, ultimately, instead of a Data Science course as indicated in your schedule, I propose you an introductory course on signals and systems that will present the essential concepts for any physicist who works with data. Compared to signals and systems courses in engineering curricula, this course will skip some traditionally taught concepts (z-transform, Laplace transform, feedback and control…). To go further, there are many resources such as (Oppenheim et al., 1997) or MIT video lectures (Freeman, 2011).

Contents

1 Signals & systems

1.1 Signals

A signal represents the variation of a quantity with time or another independent variable, such as space. Signals may describe a wide variety of physical phenomena: *e.g.*, an electrical voltage across a circuit as a function of time $v(t)$, or the light intensity in the plane of a detector $I(x, y)$. Signals are represented mathematically as functions of one or more independent variables.

In this course, for simplicity, we will generally represent a signal as a function of a single variable: time. However, this can easily be generalized to functions of several variables of any physical dimension. For example, in geophysics, a signal can represent variations with latitude, longitude, and depth of quantities such as density, resistivity, *etc.*

Continuous signals: A continuous signal, written $x(t)$ and also called an analog signal, is a function defined for all values of time (possibly in an interval). For example, the fluctuations in the current produced by a coil in an electromagnetic microphone form a continuous signal. Such signals can take any value in a continuous range. Figure 1.1 shows an example.

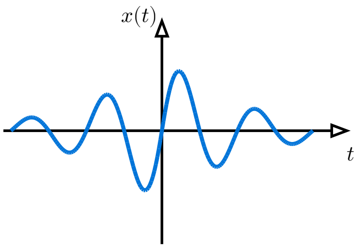

Figure 1.1: Continuous signal

Discrete signals: A discrete signal written $x[n]$ is defined only at discrete time indices n (integers). It is sometimes referred to as a discrete-time sequence. For example, the daily average temperature measured at a weather station is a discrete signal, as it is only known at specific times (once per day in this case). Discrete signals often arise from sampling continuous signals at regular intervals. Figure 1.2 shows a discrete signal obtained by sampling the continuous signal presented in Figure 1.1.

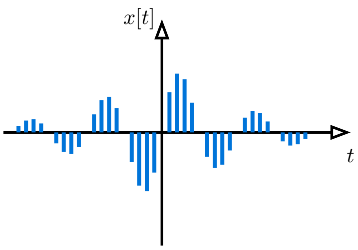

Figure 1.2: Discrete signal

A discrete signal (or discrete time signal) should not be confused with discrete valued signal (which can be either continuous or discrete time). A discrete valued signal also called quantized signal can take only a finite or countable number of values. An example of a discrete valued signal is a digital signal used in digital electronics, which can take only two values (0 and 1).

Periodic signals: A periodic signal is a signal that repeats itself at regular intervals over time. A continuous signal $x(t)$ is periodic if there exists a positive T (the period) that satisfies the condition:

$$x(t) = x(t + T), \quad \forall t \in \mathbb{R} \tag{1.2}$$

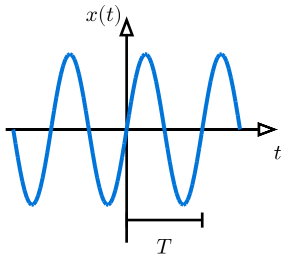

Figure 1.3: Periodic signal with period T

In other words, a periodic signal is unchanged by a time shift of T. In this case, we say that $x(t)$ is periodic with period T. If a signal is periodic of period T, any integer multiple nT (for a positive integer n) is also a period. The least positive period is called the fundamental period. Often, "the" period of a signal is used to refer to its fundamental period.

Discrete periodic signals are defined analogously. A discrete signal $x[n]$ is periodic if there exists a positive period N where:

$$x[n] = x[n+N], \quad \forall n \in \mathbb{N} \tag{1.3}$$

Energy: The energy E of a continuous signal $x(t)$ is defined as:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 \, \mathrm{d}t \tag{1.4}$$

The unit of energy is the square of the unit of $x(t)$. In this context, this energy is not, strictly speaking, the same as the conventional notion of energy in physics (usually in joules).

For some signals the integral in Eq. (1.4) might not converge: *e.g.* if $x(t)$ or $x[n]$ is periodic. Such signals have infinite energy: $x(t)$ is not a square-integrable function (*i.e.* does not belong to the L^2 space). Signals of finite energy (*i.e.* $E < \infty$) are often called energy signals.

Power: Power P of the signal $x(t)$ is defined as the amount of energy per unit time:

$$P = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 \, \mathrm{d}t \tag{1.5}$$

This quantity is useful to work with infinite energy signals. By construction, $P = 0$ for energy signals (*i.e.* $E < \infty$).

For periodic signals, this amounts to calculating the average power over a single period:

$$P = \frac{1}{T} \int_{t_0}^{t_0+T} |x(t)|^2 \, \mathrm{d}t \,. \tag{1.6}$$

Signals of non-zero but finite power (*i.e.* $0 < P < \infty$) are often called power signals. Periodic or constant signals are examples of power signals. There are signals, like $x(t) = t$, with infinite power that are neither energy nor power signals.

These quantities are also defined for discrete signals $x[n]$:

$$E = \sum_{n=-\infty}^{\infty} |x[n]|^2 \tag{1.7}$$

$$P = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x[n]|^2 \tag{1.8}$$

It is important to remember that, in this lecture, the terms “power” and “energy” are used independently of whether these quantities are actually related to physical energy.

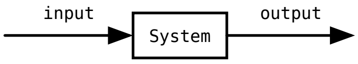

Figure 1.4: Block diagram of a system

1.2 Systems

A system is a powerful conceptual tool used across a wide range of scientific fields, particularly in physics. In this abstraction, described as a block diagram in Figure 1.4, a system transforms an input signal into an output signal.

$$y(t) = H\{x(t)\} \tag{1.9}$$

A seismometer is a good example of a system: the physical ground motion $x(t)$ is the input and the seismometer transforms it into an electrical voltage $y(t)$. The **description** of a system is **arbitrary**, and its inputs/outputs can be defined to facilitate calculations or the understanding of the system. For example, in the case of a seismometer the input can be ground displacement, its velocity, its acceleration, *etc.*.

Properties of systems:

BIBO Stability A system is said to be bounded input bounded output (BIBO) stable if the output is bounded for every bounded input to the system.

Causality A system is causal if the output at any time depends only on values of the input at the present time and in the past. If any value of the output signal depends on a future value of the input signal, then the system is non-causal.

Linearity A system is said to be linear if it satisfies the **principle of superposition** (additivity and homogeneity) where for any $(a_1, a_2) \in \mathbb{C}^2$:

$$H\{a_1x_1(t) + a_2x_2(t)\} = H\{a_1x_1(t)\} + H\{a_2x_2(t)\} \quad \textbf{additivity} \tag{1.10}$$

$$= a_1H\{x_1(t)\} + a_2H\{x_2(t)\} \quad \textbf{homogeneity} \tag{1.11}$$

Time invariance A system is said to be time-invariant if its behavior does not change over time. This means delaying the input by some amount simply delays the output by the same amount:

$$y(t+\tau) = H\{x(t+\tau)\} \tag{1.12}$$

Modularity: One of the main advantages of this description is that a system is **modular**: it can be decomposed into smaller elementary systems (*e.g.* mass/spring system + transducer + analog to digital converter + *etc.*) or be included in a larger system (*e.g.* seismic source + propagation medium + array of seismometers + *etc.*) as illustrated in Figure 1.5.

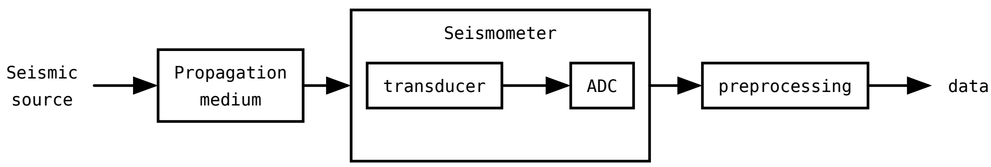

Figure 1.5: Seismometer as a modular system

Why studying systems?: There are many reasons to describe a problem as a system. Depending on the final goal, one can use this formalism to study:

- **output**: predicting the output from the input given a known system, *e.g.* the *Deep Thought* supercomputer answering *42* to the *Ultimate Question of Life, the Universe, and Everything.*
- **system**: characterizing the effect of the system (distortion, attenuation, *etc.*) on measurements, *e.g.* understanding how the transmission medium is changing communication signals.
- **input**: inferring the input from the output. This is often tackled within an "inverse problem" framework, *e.g.* estimating the position and energy of an earthquake from seismograms, or estimating the question to which the answer is *42*.

1.3 Representation of signals

Unit impulse: In discrete time, the unit impulse, also known as the delta function, is the simplest discrete signal.

$$\delta[n] = \begin{cases} 1 & \text{if } n = 0\,, \\ 0 & \text{otherwise.} \end{cases} \tag{1.13}$$

Representation of discrete time signal: Any discrete-time signal can be viewed as a sequence of scaled individual unit pulses:

$$x[n] = \sum_{k=-\infty}^{+\infty} x[k]\,\delta[n-k] \tag{1.14}$$

This can be used to represent any arbitrary sequence as a linear combination of shifted unit impulses $\delta[n-k]$, where the weights are $x[k]$. This is sometimes called the *sifting property* of the discrete-time unit impulse, where the impulse acts as a selector preserving only the value corresponding to $k = n$. The impulse functions form a **complete basis set** for discrete-time signals. The coefficient for each basis function is simply $x[k]$.

Dirac delta function: In continuous time we do not have a discrete sequence of values, but we can think of a continuous unit impulse function $\delta_{\{\Delta\}}$ as a pulse of width Δ:

$$\delta_\Delta(t) = \begin{cases} \frac{1}{\Delta} & \text{if } |t| < \frac{\Delta}{2}\,, \\ 0 & \text{otherwise.} \end{cases} \tag{1.15}$$

As $\Delta \to 0$, δ_Δ approaches the Dirac delta distribution. The Dirac delta function (or distribution) is a generalized function on the real numbers, whose value is zero everywhere except at zero, and whose integral over the entire real line is equal to one:

$$\delta(t) = \begin{cases} +\infty & \text{if } t = 0\,, \\ 0 & \text{otherwise.} \end{cases} \tag{1.16}$$

such that

$$\int_{-\infty}^{+\infty} \delta(t)\,\mathrm{d}t = 1 \tag{1.17}$$

It is a generalized function that appears only under an integral. It is often defined as the limit of a sequence of functions, such as the sequence defined in Eq. (1.15) with decreasing Δ, or a sequence of Gaussian distributions centered at the origin with variance tending to zero.

With this Dirac delta distribution, similarly to the discrete case Eq. (1.14), we can represent any continuous signal as:

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\,\delta(t-\tau)\ \mathrm{d}\tau \tag{1.18}$$

Again this can be viewed as a "weighted sum" of shifted impulses where the weight on the impulse $\delta(t-\tau)$ is $x(\tau)$. Contrary to the discrete case, the Dirac delta distribution does not, strictly speaking, define a basis of the space of continuous signals as δ itself does not belong to this space.

Properties of the Dirac delta distribution:

Translation

$$\int_{-\infty}^{+\infty} f(t)\,\delta(t-\tau)\,\mathrm{d}t = f(\tau) \tag{1.19}$$

Scaling

$$\delta(a\,t) = \frac{1}{|a|}\delta(t) \tag{1.20}$$

Symmetry δ is even

$$\delta(t) = \delta(-t) \tag{1.21}$$

Heavyside step function:

The Heaviside function or unit step function $u(t)$ is a step function defined as:

$$u(t) = \begin{cases} 0 & \text{if } t < 0\,, \\ 1 & \text{if } t \geq 0. \end{cases} \tag{1.22}$$

It is the indicator function of $\mathbb{R}^+$. Two conventions exist: either $u(0) = 1$ or $u(0) = \frac{1}{2}$. It is related to the Dirac delta function by:

$$\frac{\mathrm{d}}{\mathrm{d}t}u(t) = \delta(t) \tag{1.23}$$

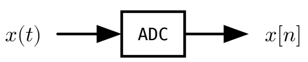

Figure 1.6: Sampling system

1.4 Sampling

The digitization is the process of converting an analog signal $x(t)$ into a discrete signal $x[n]$, usually done in practice by an analog to digital converter (ADC). The digitization itself is composed of two operations:

- **sampling** that goes from continuous time to discrete time
- **quantization** that goes from continuous value to a finite number of levels

Quantization, being a non-linear operation, will not be treated in this lecture.

The sampling of the function $f(t)$ with the sampling period Δ is given by the equation:

$$f[n] = \int_{-\infty}^{+\infty} f(t)\,\delta(t - n\,\Delta)\,\mathrm{d}t\,, \tag{1.24}$$

The condition to perfectly reproduce $f(t)$ from $f[n]$ will be treated in Section 5.5.

2 Linear Time-Invariant Systems

Among the properties of a system given in Section 1.2, linearity and time invariance play a fundamental role in signal and system analysis. First, the linearity and time invariance properties are fortunately shared by numerous physical phenomena. In addition, signal and system analysis provides powerful tools to analyze LTI systems in great detail, going deep into their properties.

The main reason LTI systems can be deeply analyzed is a consequence of the superposition property given in Section 1.2.a: if we can represent the input to an LTI system in terms of a linear combination of a set of basic signals, we can compute its output as the combination of its responses to these basic signals.

2.1 Discrete time LTI systems

As shown in Section 1.3.b, a discrete signal $x[n]$ can be represented as a sum of impulses: $x[n] = \sum_k x[k]\,\delta[n-k]$. In this representation, an arbitrary sequence is a linear combination of shifted unit impulses $\delta[n-k]$, where the weights are $x[k]$.

Discrete time linear system: If H is a **linear system** then its output $y[n]$ is simply the weighted linear combination of shifted unit impulse responses:

$$y[n] = H(x[n]), \tag{2.2}$$

$$= H\left(\sum_k x[k]\,\delta[n-k]\right), \tag{2.3}$$

$$= \sum_k x[k]\,H(\delta[n-k]), \tag{2.4}$$

$$= \sum_k x[k]\,h_k[n-k] \tag{2.5}$$

where $h_k[n] = H(\delta[n-k])$ denotes the response of the linear system H to the shifted unit impulse $\delta[n-k]$.

In matrix–vector notation we define the input vector $\mathbf{x} = \left[x_1, x_2, ..., x_{\{N-1\}}\right]^T$, the output $\mathbf{y} = \left[y_0, y_1, ..., y_{\{M-1\}}\right]^T$; the linear system is described by the linear operator (matrix) M_H:

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{x} \tag{2.6}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{M-1} \end{pmatrix} = x_1 \begin{pmatrix} h_0[0] \\ h_0[1] \\ \vdots \\ h_0[N-1] \end{pmatrix} + x_2 \begin{pmatrix} h_1[-1] \\ h_1[0] \\ \vdots \\ h_1[N-2] \end{pmatrix} + ... \tag{2.7}$$

$$= \begin{pmatrix} h_0[0] & h_1[-1] & ... & h_N[1-N] \\ h_0[1] & h_1[0] & ... & h_N[2-N] \\ \vdots & \vdots & \ddots & \vdots \\ h_0[M-1] & h_1[M-2] & ... & h_N[M-N] \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{pmatrix} \tag{2.8}$$

Discrete-Time Linear Time-Invariant System: If the linear system H is also **time-invariant** all $h_k[n]$ are time-shifted versions of a single sequence $h_0[n]$:

$$h_k[n] = h_0[n-k] \tag{2.9}$$

Dropping the subscript 0 for convenience, we define the **impulse response** h of the system:

$$h[n] = h_0[n] \tag{2.10}$$

$h[n]$ being the output of the LTI system H to the unit impulse $\delta[n]$.

Then for any LTI systems, its output depends only on the input and its impulse response:

$$y[n] = \sum_{k=-\infty}^{+\infty} x[k]\,h[n-k] \tag{2.11}$$

This is a **convolution sum** or superposition sum, and the operation on the right-hand side is known as the convolution of the sequences x[n] and h[n]. This convolution is denoted with the symbol $*$ as in :

$$y[n] = (x * h)[n]. \tag{2.12}$$

The convolution matrix is then a band-diagonal matrix:

$$\mathbf{H} = \begin{pmatrix} h[0] & h[-1] & \dots & h[1-N] \\ h[1] & h[0] & \dots & h[2-N] \\ \vdots & \vdots & \ddots & \vdots \\ h[M-1] & h[M-2] & \dots & h[M-N] \end{pmatrix} \tag{2.13}$$

2.2 Continuous time LTI system

Continuous time linear system: Let us define $h(t, \tau)$ as the response at time t to a unit impulse $\delta(t-\tau)$ applied at time τ. Similarly to the discrete case, the output of the system is:

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)\, h(t,\tau)\, \mathrm{d}\tau \tag{2.14}$$

If we intuitively think of $x(t)$ as a "sum" of weighted shifted impulses (where the weight on the impulse $\delta(t-\tau)$ is $x(\tau)$) this Eq. (2.14) represents the superposition of the responses to each of these inputs, and by linearity, the weight on the response $h(t,\tau)$ to the shifted impulse $\delta(t-\tau)$ is also $x(\tau)$.

Continuous-Time Linear Time-Invariant System: If, in addition to being linear, the system is time-invariant, its response no longer depends on the instant τ where the impulse was applied but only on the time difference $t-\tau$:

$$h(t,\tau) = h(t-\tau)\,. \tag{2.15}$$

With this definition, $h(t)$ is the response of the system to the impulse $\delta(t)$, that is the **impulse response** of the system.

For Linear Time-Invariant systems, Eq. (2.14) becomes the **convolution integral**:

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)\, h(t-\tau)\, \mathrm{d}\tau \tag{2.16}$$

2.3 Convolution

The convolution of two functions f and g, written $f * g$, is defined as the integral of the product of the two functions after one is reflected and shifted:

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)\, g(t-\tau)\, \mathrm{d}\tau \tag{2.17}$$

This operation is well defined only if f and g decay sufficiently rapidly at infinity so that the integral exists. Existence conditions include:

- f and g have compact support (then $f * g$ exists and also has compact support),

- $finL^p$ and $ginL^q$ with $1le\frac{1}{p}+\frac{1}{q}le2$, ensuring integrability (Young's convolution inequality).

In physics, signals typically have finite energy or power, so these conditions are usually satisfied in practice.

Properties of the convolution operation:

Commutativity

$$f * g = g * f \tag{2.18}$$

Associativity

$$f * (g * h) = (f * g) * h \tag{2.19}$$

Distributivity

$$f * (g + h) = f * g + f * h \tag{2.20}$$

Identity element

$$f * \delta = f \tag{2.21}$$

Time reversal

$$(f * g)(-t) = f(-t) * g(-t) \tag{2.22}$$

Conjugation

$$\overline{f * g} = \overline{f} * \overline{g} \tag{2.23}$$

Differentiation

$$\frac{\mathrm{d}}{\mathrm{d}t}(f * g) = \frac{\mathrm{d}f}{\mathrm{d}t} * g = f * \frac{\mathrm{d}g}{\mathrm{d}t} \tag{2.24}$$

Integration

$$\int_{-\infty}^{+\infty} (f * g)(t)\,\mathrm{d}t = \left(\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t\right)\left(\int_{-\infty}^{+\infty} g(t)\,\mathrm{d}t\right) \tag{2.25}$$

Green's function G of a system with impulse response h

$$h * G = \delta \quad \text{by definition} \tag{2.26}$$

Eigenfunctions of an LTI system: An eigenfunction is a function f for which the output of the operator is the same function scaled by some constant:

$$H\{f\} = \lambda f\,, \tag{2.27}$$

where λ is the eigenvalue (a constant).

Complex exponential functions are eigenfunctions of any LTI system. That is, when a complex exponential is applied as input to an LTI system, the output is simply a scaled version of the input:

$$h * e^{j\omega t} = \lambda\, e^{j\omega t} \tag{2.28}$$

Demonstration

$$\begin{aligned} h * e^{j\omega t} &= \int_{\mathbb{R}} h(\tau)\, e^{j\omega(t-\tau)}\, \mathrm{d}\tau \\ &= \int_{\mathbb{R}} h(\tau)\, e^{j\omega t}\, e^{-j\omega\tau}\, \mathrm{d}\tau \\ &= \underbrace{e^{j\omega t}}_{\substack{\text{eigen}\\ \text{function}}} \underbrace{\int_{\mathbb{R}} h(\tau)\, e^{-j\omega\tau}\, \mathrm{d}\tau}_{\lambda \text{ (scalar)}} \end{aligned} \tag{2.29}$$

This property is extremely important because the effect of an LTI system on a linear combination of complex exponentials is, thanks to the superposition property, the same complex exponentials weighted by the eigenvalues λ that can be computed independently once and for all.

If we can decompose any input signal as a sum of complex exponentials, then computing the output of any LTI system is trivial:

$$\begin{aligned} h * \left(a_1 e^{j\omega_1 t} + a_2 e^{j\omega_2 t}\right) &= a_1\, h * e^{j\omega_1 t} + a_2\, h * e^{j\omega_2 t} \\ &= a_1\, \lambda_1\, e^{j\omega_1 t} + a_2\, \lambda_2\, e^{j\omega_2 t}. \end{aligned}$$

with

$$\begin{aligned} \lambda_1 &= \int_{\mathbb{R}} h(\tau)\, e^{-j\omega_1\tau}\, \mathrm{d}\tau \\ \lambda_2 &= \int_{\mathbb{R}} h(\tau)\, e^{-j\omega_2\tau}\, \mathrm{d}\tau \end{aligned}$$

3 Fourier Series

To use the fact that complex exponentials are eigenfunctions of LTI systems, one has to decompose the input signal into complex exponentials. First, we will see that any periodic signal can be decomposed as a sum of complex exponentials via Fourier series; then we will extend to aperiodic signals (using the Fourier transform).

3.1 Harmonic Signals

As stated in Section 1.1.c, a signal is periodic if for some $T > 0$:

$$x(t) = x(t+T), \quad \forall t \in \mathbb{R}. \tag{3.2}$$

A complex exponential $e^{j\omega t}$ is periodic of period T if and only if

$$\begin{aligned} e^{j\omega t} &= e^{j\omega(t+T)} \\ &= e^{j\omega t} e^{j\omega T} \\ \Leftrightarrow e^{j\omega T} &= 1 \\ \Leftrightarrow \omega &= n\,\frac{2\pi}{T}, \quad \forall n \in \mathbb{Z} \end{aligned} \tag{3.3}$$

Every signal $x(t)$ that is linear combination of complex exponentials, periodic of period T can be expressed as:

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\,\omega_0}, \tag{3.4}$$

where $\omega_0 = \frac{2\pi}{T}$ is the fundamental frequency (or pulsation). This Eq. (3.4) is referred to as the synthesis equation. The signal x is a **harmonic signal** with frequencies that are integer multiples of the fundamental frequency ω_0 (the harmonics). The coefficients $a_k = \rho_k\, e^{j\,\varphi_k}$ are complex and can also be expressed in terms of phase and amplitude. The component a_0 corresponding to the mean of x is sometimes called the DC component.

Real harmonic signal:

Euler formulae

$$\cos(\theta) = \frac{e^{j\,\theta} + e^{-j\,\theta}}{2}$$

$$\sin(\theta) = \frac{e^{j\,\theta} - e^{-j\,\theta}}{2j}$$

Using Euler formula and defining $a_k = b_k + jc_k$, we can rewrite Eq. (3.4):

$$\begin{aligned} x(t) = a_0 &+ \sum_{n=1}^{+\infty} (a_n e^{jn\,\omega_0} + a_{-n} e^{-jn\,\omega_0}) \\ = a_0 &+ \sum_{n=1}^{+\infty} \left(\frac{a_n + a_{-n}}{2} \cos(n\,\omega_0) - j \frac{a_n - a_{-n}}{2} \sin(n\,\omega_0) \right) \\ = a_0 &+ \sum_{n=1}^{+\infty} \left(\frac{\Re(a_n) + \Re(a_{-n})}{2} \cos(n\,\omega_0) + \frac{\Im(a_n) - \Im(a_{-n})}{2} \sin(n\,\omega_0) \right) \\ &+ j \sum_{n=1}^{+\infty} \left(\frac{\Im(c_n) + \Im(c_{-n})}{2} \cos(n\,\omega_0) - \frac{\Re(c_n) - \Re(c_{-n})}{2} \sin(n\,\omega_0) \right) \end{aligned}$$

Real harmonic signals are complex harmonic signals with zero imaginary part:

$$\begin{aligned} x(t) &= \sum_{k=-\infty}^{+\infty} \Re\big(a_k e^{jk\,\omega_0}\big), \\ &= a_0 + \sum_{n=1}^{+\infty} \left(\frac{\Re(a_n) + \Re(a_{-n})}{2} \cos(n\,\omega_0) + \frac{\Im(a_n) - \Im(a_{-n})}{2} \sin(n\,\omega_0) \right) \\ &= a_0 + \sum_{n=1}^{+\infty} b_n \cos(n\,\omega_0) + c_n \sin(n\,\omega_0)) \\ &= a_0 + \sum_{n=1}^{+\infty} \rho_n \cos(n\,\omega_0 + \varphi_n) \end{aligned} \tag{3.5}$$

where we define:

$$b_n = \frac{\Re(a_n) + \Re(a_{-n})}{2} \tag{3.6}$$

$$c_n = \frac{\Im(a_n) - \Im(a_{-n})}{2} \tag{3.7}$$

$$\rho_n = \sqrt{b_n^2 + c_n^2} \qquad \textbf{cartesian representation} \tag{3.8}$$

$$\varphi_n = \arctan\left(\frac{c_n}{b_n}\right) \qquad \textbf{polar representation} \tag{3.9}$$

$$a_n = \begin{cases} \frac{1}{2}(b_n - j\,c_n) & \text{if } n < 0\,, \\ a_n & \text{if } n = 0\,, \\ \frac{1}{2}(b_n + j\,c_n) & \text{if } n > 0\,. \end{cases} \tag{3.10}$$

b_n and c_n will describe the even and odd components of x respectively.

3.2 Fourier Series Representation of Continuous Periodic Signals

The idea of decomposing any periodic function into the sum of simple oscillating functions was initially proposed by Fourier in 1807. He stated that any periodic function $x(t)$ of period T can be represented as a sum of complex exponentials of frequencies that are integer multiples of the fundamental frequency $\omega_0 = \frac{2\pi}{T}$ as in the synthesis Eq. (3.4).

To determine Fourier coefficients a_k from any periodic function $x(t)$ of period T we will use two properties of periodic signal:

- the integration of a periodic signal x over any interval of length equals to its period T is:

$$\int_T x(t)\ \mathrm{d}t = \int_{t_0}^{t_0+T} x(t)\ \mathrm{d}t, \quad \forall t \in \mathbb{R}. \tag{3.11}$$

- the integral of a complex exponential over a period T is zero excepted for $k = 0$:

$$\frac{1}{T}\int_T e^{-jk\omega_0 t}\ \mathrm{d}t = \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases} \\ = \delta[k] \tag{3.12}$$

If any periodic signal $x(t)$ can be expressed as a weighted sum of complex exponentials thanks to the synthesis Eq. (3.4), then we can compute its correlation with a complex exponential of frequency $k\omega_0$ for any $k \in \mathbb{Z}$ as follows:

$$
\begin{aligned}
\frac{1}{T}\int_T x(t)\, e^{-jk\,\omega_0 t}\ \mathrm{d}t &= \frac{1}{T}\int_T \sum_{\ell=-\infty}^{+\infty} a_\ell\, e^{j\ell\,\omega_0 t}\, e^{-jk\,\omega_0 t}\ \mathrm{d}t,\\
&= \frac{1}{T}\int_T \sum_{\ell=-\infty}^{+\infty} a_\ell\, e^{j(\ell-k)\,\omega_0 t}\ \mathrm{d}t,\\
&= \frac{1}{T}\sum_{\ell=-\infty}^{+\infty}\int_T a_\ell\, e^{j(\ell-k)\,\omega_0 t}\ \mathrm{d}t,\\
&= a_\ell\, \delta[\ell-k],\\
&= a_k
\end{aligned}
$$

This defines the *analysis–synthesis* set of equations of the Fourier series:

$$\hat{x}_k = \frac{1}{T}\int_{t_0}^{t_0+T} x(t)\, e^{-jk\,\omega_0 t}\ \mathrm{d}t, \qquad \textbf{analysis} \qquad (3.13)$$

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\,\omega_0}. \qquad \textbf{synthesis} \qquad (3.14)$$

Depending on the field, the Fourier series coefficients $\hat{x}_k$ can also be denoted $\hat{x}[k]$ or $X[k]$.

3.3 Convergence of Fourier Series

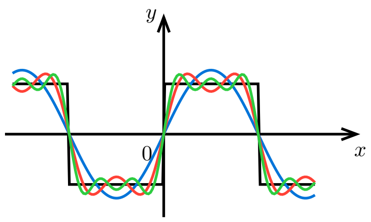

Figure 3.7: eq-fourier-transform-synthesis ($N = 1, 3, 5$) orders of the Fourier series of a square wave

The question of the convergence of Fourier series, *i.e.* does all periodic function can be represented by its Fourier series?, was only solved by Dirichlet in 1829. He showed that the Fourier series of a periodic function $x(t)$ converges to $x(t)$ at all points where x is continuous and to the average of the left-hand and right-hand limits at points of discontinuity, provided that:

- $x(t)$ is absolutely integrable over a period, *i.e.* $\int_T |x(t)|\ \mathrm{d}t < \infty$
- $x(t)$ has a finite number of maxima and minima in any given period,
- $x(t)$ has a finite number of discontinuities in any given period.

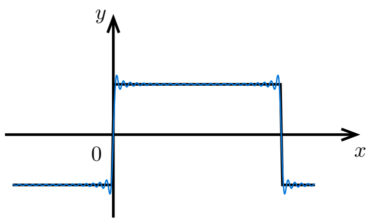

Figure 3.8: Gibbs phenomena on the Fourier series (order $N = 50$) of a square wave

The point-wise convergence is only *almost everywhere*, meaning that the Fourier series may not converge to $x(t)$ for some points, *i.e.* at discontinuities. Indeed, a truncated Fourier series approximation of a discontinuous signal will in general exhibit high-frequency ripples and overshoot x(t) near the discontinuities. These ripples, known as *Gibbs phenomena*, are present no matter how large the approximation order, as seen in Figure 3.8 (at least 9% overshoot for a unit square wave). However, large enough approximation order can always be chosen so as to guarantee that the total energy in these ripples is insignificant.

$$\lim_{N\to+\infty}\int_T \left| x(t) - \sum_{k=-N}^{k=+N} \hat{x}_k\, e^{j\,k\,\omega_0\,t} \right|^2 \mathrm{d}t = 0 \qquad (3.15)$$

3.4 Orthonormal basis of harmonic signals space

The space of square-integrable periodic functions on the period $[0, T]$ forms the Hilbert space $L^2([0, T])$. This space is equipped with the inner product:

$$\langle f, g \rangle = \frac{1}{T} \int_0^T f(t)\, \overline{g(t)}\ \mathrm{d}t, \tag{3.16}$$

where $\overline{g(t)}$ is the complex conjugate of $g(t)$. This inner product induces the norm:

$$\|f\| = \sqrt{\langle f, f \rangle} \tag{3.17}$$

The scalar product between two complex exponentials is:

$$\begin{aligned} < e^{jk\,\omega_0}, e^{j\ell\,\omega_0} >&= \frac{1}{T} \int_0^T e^{jk\,\omega_0}\, e^{-j\ell\,\omega_0}\ \mathrm{d}t, \\ &= \frac{1}{T} \int_0^T e^{j(k-\ell)\,\omega_0}\ \mathrm{d}t \\ &= \delta[k-\ell] \end{aligned} \tag{3.18}$$

That means that $\{e^{jk\,\omega_0} : k \in \mathbb{Z}\}$ forms an **orthonormal basis** of $L^2([0, T])$. In other words, every square-integrable periodic function can be represented as a Fourier series as defined by the analysis-synthesis equations above.

3.5 Parseval theorem

The Parseval theorem states that the energy of a signal $x(t)$ over a period T is equal to the sum of the squared magnitudes of its Fourier series coefficients:

$$\frac{1}{T} \int_0^T |x(t)|^2\ \mathrm{d}t = \sum_{k=-\infty}^{k=+\infty} |\hat{x}_k|^2 \tag{3.19}$$

Demonstration:

$$
\begin{aligned}
\frac{1}{T}\int_0^T |x(t)|^2 \,\mathrm{d}t &= \frac{1}{T}\int_0^T x(t)\,\overline{x}(t)\,\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \sum_{k=-\infty}^{k=+\infty} \hat{x}_k\, e^{j\,k\,\omega_0\,t}\, \overline{\sum_{k=-\infty}^{k=+\infty} \hat{x}_k\, e^{j\,k\,\omega_0\,t}}\,\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \sum_{k=-\infty}^{k=+\infty} \hat{x}_k\, e^{j\,k\,\omega_0\,t} \sum_{k=-\infty}^{k=+\infty} \overline{\hat{x}_k}\, e^{-j\,k\,\omega_0\,t}\,\mathrm{d}t \\
&= \frac{1}{T}\sum_{k=-\infty}^{k=+\infty}\sum_{k'=-\infty}^{k'=+\infty}\int_0^T \hat{x}_k\overline{\hat{x}_{k'}}\, e^{j\,(k-k')\,\omega_0\,t}\,\mathrm{d}t \\
&= \frac{1}{T}\sum_{k=-\infty}^{k=+\infty}\sum_{k'=-\infty}^{k'=+\infty} \hat{x}_k\overline{\hat{x}_{k'}}\, T\,\delta[k-k'] \\
&= \sum_{k=-\infty}^{k=+\infty} |\hat{x}_k|^2
\end{aligned}
$$

3.6 Properties of Fourier Series

Linearity

$$z(t) = a\,x(t) + b\,y(t) \Leftrightarrow \hat{z}_k = a\,\hat{x}_k + b\,\hat{z}_k \tag{3.20}$$

Time Shifting

$$y(t) = x(t - t_0) \Leftrightarrow \hat{y}_k = e^{j\,k\,\omega_0\,t_0}\hat{x}_k \tag{3.21}$$

Time reversal

$$y(t) = x(-t) \Leftrightarrow \hat{y}_k = \hat{x}_{-k} \tag{3.22}$$

Frequency Shifting

$$y(t) = e^{j\,k_0\,\omega_0\,t}\,x(t) \Leftrightarrow \hat{y}_k = \hat{x}_{k-k_0} \tag{3.23}$$

Scaling

$$y(t) = x(a\,t) \Leftrightarrow \hat{y}_k = \frac{1}{|a|}\hat{x}\left[\frac{k}{a}\right] \tag{3.24}$$

Multiplication

$$z(t) = x(t)\,y(t) \Leftrightarrow \hat{z}_k = \sum_{\ell=-\infty}^{+\infty} \hat{x}_\ell\,\hat{y}_{k-\ell} \tag{3.25}$$

Conjugation

$$y(t) = \overline{x(t)} \Leftrightarrow \hat{y}_k = \overline{\hat{x}_{-k}} \tag{3.26}$$

Differentiation

$$y(t) = \frac{\mathrm{d}x(t)}{\mathrm{d}t} \Leftrightarrow \hat{y}_k = j\,k\,\omega_0\,\hat{x}_k \tag{3.27}$$

Symmetry for real signals

$$x(t) \in \mathbb{R} \Leftrightarrow \hat{x}_{-k} = \overline{\hat{x}_k} \tag{3.28}$$

Even real signals

$$x(t) = x(-t) \in \mathbb{R} \Leftrightarrow \hat{x}_k \in \mathbb{R} \tag{3.29}$$

Odd real signals

$$x(t) = -x(-t) \in \mathbb{R} \Leftrightarrow \hat{x}_k \in j\,\mathbb{R} \tag{3.30}$$

3.7 Fourier Series Representation of Discrete Periodic Signals

From Eq. (1.3), a discrete time signal is periodic with period N if:

$$x[n] = x[n+N], \quad \forall n \in \mathbb{N} \tag{3.31}$$

The fundamental frequency is $\omega_0 = \frac{2\pi}{N}$ is defined from the fundamental period N, the smallest integer for which the Eq. (3.31) holds.

The set of all discrete-time complex exponentials that are periodic with period N is finite and given by:

$$\left\{e^{j\,\omega_0\,k\,n} : k = 0, 1, ..., N-1\right\} \tag{3.32}$$

since for any $k \geq N$ or $k < 0$:

$$e^{j\,\omega_0\,k\,n} = e^{j\,\omega_0 (k \bmod N)\,n} \tag{3.33}$$

Analysis-Synthesis Equations: As for continous-time signal described Section 3.2, any discrete periodic signal $x[n]$ of period N can be expressed as a weighted sum of these complex exponentials. The analysis and synthesis equations for discrete-time Fourier series are:

$$\hat{x}[k] = \frac{1}{N} \sum_{n=1}^{N} x[n]\, e^{-j\,\omega_0\,k\,n}, \qquad \textbf{analysis} \tag{3.34}$$

$$x[n] = \sum_{k=1}^{N} \hat{x}[k]\, e^{j\,\omega_0\,k\,n}. \qquad \textbf{synthesis} \tag{3.35}$$

In these equations, the limits of the summation can be any contiguous range in $\mathbb{N}$ (*e.g.* $k = 0, 1, ..., N-1$ or $k = 1, 2, ..., N$).

Parseval Theorem: The Parseval theorem holds equivalently in discrete-time:

$$\sum_{n=1}^{n=N} |x[n]|^2 = \sum_{k=1}^{k=N} |\hat{x}[k]|^2 \tag{3.36}$$

Properties of Discrete-Time Fourier Series: The properties of Fourier series decomposition for discrete-time signals are similar to the continuous signal ones described in Section 3.6.

Property	Time Domain	Frequency Domain
Linearity	$z[n] = a\,x[n] + b\,y[n]$	$\hat{z}[k] = a\,\hat{x}[k] + b\,\hat{y}[k]$
Time Shifting	$y[n] = x[n - n_0]$	$\hat{y}[k] = e^{-j\,\omega_0\,k\,n_0}\hat{x}[k]$
Time Reversal	$y[n] = x[-n]$	$\hat{y}[k] = \hat{x}[-k]$
Frequency Shifting	$y[n] = e^{j\,\omega_0\,k_0\,n}\,x[n]$	$\hat{y}[k] = \hat{x}[k - k_0]$
Multiplication	$z[n] = x[n]\,y[n]$	$\hat{z}[k] = \sum_{\ell=0}^{N-1} \hat{x}[\ell]\,\hat{y}[k-\ell]$
Conjugation	$y[n] = \overline{x[n]}$	$\hat{y}[k] = \overline{\hat{x}[-k]}$
First Difference	$y[n] = x[n] - x[n-1]$	$\hat{y}[k] = \left(1 - e^{-j\,\omega_0\,k}\right)\hat{x}[k]$
Symmetry	$x[n] \in \mathbb{R}$	$\hat{x}[-k] = \overline{\hat{x}[k]}$
Even Real Signals	$x[n] = x[-n] \in \mathbb{R}$	$\hat{x}[k] \in \mathbb{R}$
Odd Real Signals	$x[n] = -x[-n] \in \mathbb{R}$	$\hat{x}[k] \in j\,\mathbb{R}$

Table 3.1: Properties of Discrete Time Fourier Series

4 Fourier transform

As complex exponentials are eigenfunctions of LTI systems (see Eq. (2.28)), the Fourier series decomposition of the output signal y of an LTI system with impulse response h can be easily computed from the Fourier series representation of the input signal x:

$$\begin{aligned} y(t) &= h * x \\ &= h * \left(\sum_{k=-\infty}^{+\infty} \hat{x}_k\, e^{j\,\omega_0\,k\,t} \right), \\ &= \sum_{k=-\infty}^{+\infty} \hat{x}_k \left(h * e^{j\,\omega_0\,k\,t} \right), \\ &= \sum_{k=-\infty}^{+\infty} \hat{x}_k\, \lambda_k\, e^{j\,\omega_0\,k\,t}, \\ \hat{y}_k &= \hat{x}_k \lambda_k \end{aligned} \tag{4.2}$$

with

$$\lambda_k = \int_{\mathbb{R}} h(\tau)\, e^{-j\,\omega_0\,k\,\tau}\, \mathrm{d}\tau \tag{4.3}$$

The operation on the right-hand side of Eq. (4.3) is the Fourier transform of the impulse response h taken at frequency $\omega = k\,\omega_0$.

4.1 Fourier transform of Continuous Signals

The Fourier transform can be derived from the Fourier series representation of periodic signals by considering the limit when the period T

tends to infinity. In this case, the fundamental frequency $\omega_0 = \frac{2\pi}{T}$ tends to zero and the frequencies $k\,\omega_0$ become continuous over the real line $\mathbb{R}$.

This defines the Fourier transform pair in angular frequency ω:

$$\hat{x}(\omega) = \frac{1}{2\pi}\int_{\mathbb{R}} x(t)\,e^{-j\,\omega\,t}\;\mathrm{d}t, \qquad \textbf{Forward transform} \tag{4.4}$$

$$x(t) = \int_{\mathbb{R}} \hat{x}(\omega)\,e^{j\,\omega\,t}\;\mathrm{d}\omega. \qquad \textbf{Inverse transform} \tag{4.5}$$

or equivalently in ordinary frequency ν:

$$\hat{x}(\nu) = \int_{\mathbb{R}} x(t)\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}t, \qquad \textbf{Forward transform} \tag{4.6}$$

$$x(t) = \int_{\mathbb{R}} \hat{x}(\nu)\,e^{j\,2\pi\,\nu\,t}\;\mathrm{d}\omega. \qquad \textbf{Inverse transform} \tag{4.7}$$

The angular frequency Fourier transform can be made unitary as:

$$\hat{x}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x(t)\,e^{-j\,\omega\,t}\;\mathrm{d}t, \qquad \textbf{Forward transform} \tag{4.8}$$

$$x(t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{x}(\omega)\,e^{j\,\omega\,t}\;\mathrm{d}\omega. \qquad \textbf{Inverse transform} \tag{4.9}$$

For periodic functions of period T, the Fourier transform is non-zero only at discrete frequencies $\nu = \frac{k}{T}$ with $k \in \mathbb{Z}$ and is related to the Fourier series coefficients as:

$$\hat{x}(\nu) = \sum_{k=-\infty}^{+\infty} \hat{x}_k\,\delta\left(\nu - \frac{k}{T}\right). \tag{4.10}$$

4.2 Convergence of Fourier transform

If $x(t)$ is square integrable over $\mathbb{R}$, *i.e.* $\int_{\mathbb{R}} |x(t)|^2\;\mathrm{d}t < \infty$, then its Fourier transform exists and $\hat{x}(\omega)$ is also square integrable over $\mathbb{R}$. Since many signals in physics are of finite energy, this condition holds in many applications involving physical quantities.

For periodic signals, the Fourier transform of a signal $x(t)$ exists under the same Dirichlet conditions stated in Section 3.3, which require that:

- $x(t)$ is absolutely integrable over $\mathbb{R}$, *i.e.* $\int_{\mathbb{R}}|x(t)|\;\mathrm{d}t < \infty$.
- $x(t)$ is of bounded variation (*i.e.* there is a finite number of maxima and minima within any finite interval)
- $x(t)$ has a finite number of discontinuities within any finite interval.

4.3 Plancherel-Parseval theorem

The Plancherel-Parseval theorem states that the total energy of a signal $x(t)$ is equal to the total energy of its Fourier transform $\hat{x}(\nu)$:

$$\int_{\mathbb{R}} x(t)\,\overline{y(t)}\,\mathrm{d}t = \int_{\mathbb{R}} \hat{x}(\nu)\,\overline{\hat{y}(\nu)}\,\mathrm{d}\nu \qquad \textbf{inner product} \quad (4.11)$$

$$\int_{\mathbb{R}} |x(t)|^2\,\mathrm{d}t = \int_{\mathbb{R}} |\hat{x}(\nu)|^2\,\mathrm{d}\nu \qquad \textbf{energy preservation} \quad (4.12)$$

Demonstration:

$$\begin{aligned}
\int_{\mathbb{R}} |x(t)|^2\,\mathrm{d}t &= \int_{\mathbb{R}} x(t)\,\overline{x(t)}\;\mathrm{d}t \\
&= \int_{\mathbb{R}} x(t)\,\overline{\int_{\mathbb{R}} \hat{x}(\nu)\,e^{j\,2\pi\,\nu\,t}\;\mathrm{d}\nu}\;\mathrm{d}t \\
&= \int_{\mathbb{R}} x(t) \int_{\mathbb{R}} \overline{\hat{x}(\nu)}\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}\nu\;\mathrm{d}t \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} x(t)\,\overline{\hat{x}(\nu)}\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}t\;\mathrm{d}\nu \\
&= \int_{\mathbb{R}} \overline{\hat{x}(\nu)} \left(\int_{\mathbb{R}} x(t)\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}t \right) \mathrm{d}\nu \\
&= \int_{\mathbb{R}} |\hat{x}(\nu)|^2\;\mathrm{d}\nu
\end{aligned}$$

4.4 Convolution theorem

The convolution theorem states that the Fourier transform of the convolution of two signals is equal to the product of their Fourier transforms:

$$z(t) = x(t) * y(t) \Leftrightarrow \hat{z}(\nu) = \hat{x}(\nu)\,\hat{y}(\nu) \qquad (4.13)$$

Demonstration:

$$\begin{aligned}
\hat{z}(\nu) &= \int_{\mathbb{R}} z(t)\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}t \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} x(\tau)\,y(t-\tau)\;\mathrm{d}\tau \right) e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}t \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} x(\tau)\,y(t-\tau)\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}\tau\;\mathrm{d}t \\
&= \int_{\mathbb{R}} x(\tau) \left(\int_{\mathbb{R}} y(t-\tau)\,e^{-j\,2\pi\,\nu\,t}\;\mathrm{d}t \right) \mathrm{d}\tau \\
&= \int_{\mathbb{R}} x(\tau) \left(\int_{\mathbb{R}} y(u)\,e^{-j\,2\pi\,\nu\,(u+\tau)}\;\mathrm{d}u \right) \mathrm{d}\tau \\
&= \int_{\mathbb{R}} x(\tau)\,e^{-j\,2\pi\,\nu\,\tau} \left(\int_{\mathbb{R}} y(u)\,e^{-j\,2\pi\,\nu\,u}\;\mathrm{d}u \right) \mathrm{d}\tau \\
&= \hat{x}(\nu)\,\hat{y}(\nu)
\end{aligned}$$

4.5 Properties of the continuous Fourier transform

Property	Time Domain	Frequency Domain
Linearity	$z(t) = a\,x(t) + b\,y(t)$	$\hat{z}(\nu) = a\,\hat{x}(\nu) + b\,\hat{y}(\nu)$
Time Shifting	$y(t) = x(t - t_0)$	$\hat{y}(\nu) = e^{-j\,2\pi\,\nu\,t_0}\hat{x}(\nu)$
Time Reversal	$y(t) = x(-t)$	$\hat{y}(\nu) = \hat{x}(-\nu)$
Frequency Shifting	$y(t) = e^{j\,2\pi\,\nu_0\,t}\,x(t)$	$\hat{y}(\nu) = \hat{x}(\nu - \nu_0)$
Time Scaling	$y(t) = x(a\,t)$	$\hat{y}(\nu) = \frac{1}{\lvert a\rvert}\hat{x}\left(\frac{\nu}{a}\right)$
Convolution	$z(t) = x(t) * y(t)$	$\hat{z}(\nu) = \hat{x}(\nu)\,\hat{y}(\nu)$
Multiplication	$z(t) = x(t)\,y(t)$	$\hat{z}(\nu) = \hat{x}(\nu) * \hat{y}(\nu)$
Conjugation	$y(t) = \overline{x(t)}$	$\hat{y}(\nu) = \overline{\hat{x}(-\nu)}$
Differentiation	$y(t) = \frac{\mathrm{d}}{\mathrm{d}t}x(t)$	$\hat{y}(\nu) = j\,2\pi\,\nu\,\hat{x}(\nu)$
Integration	$y(t) = \int_{-\infty}^{t} x(\tau)\,\mathrm{d}\tau$	$\hat{y}(\nu) = \frac{\hat{x}(\nu)}{j\,2\pi\,\nu} + \hat{x}(0)\,\delta(\nu)$
Symmetry	$x(t) \in \mathbb{R}$	$\hat{x}(-\nu) = \overline{\hat{x}(\nu)}$
Even Real Signals	$x(t) = x(-t) \in \mathbb{R}$	$\hat{x}(\nu) \in \mathbb{R}$
Odd Real Signals	$x(t) = -x(-t) \in \mathbb{R}$	$\hat{x}(\nu) \in j\,\mathbb{R}$

Table 4.2: Properties of Continuous Fourier Transform

4.6 Symmetry

When the real and imaginary parts of a complex function are decomposed into their even and odd parts, there are four components, each with a specific symmetry property:

- The real part of the Fourier transform of a real signal is an even function.
- The imaginary part of the Fourier transform of a real signal is an odd function.
- The real part of the Fourier transform of an imaginary signal is an odd function.
- The imaginary part of the Fourier transform of an imaginary signal is an even function.

The transform of a real-valued function thus exhibits conjugate symmetry. Conversely, if a function's Fourier transform has conjugate symmetry, the original function is real-valued.

4.7 Incertitude Principle

The Fourier transform uncertainty principle states that a signal cannot be simultaneously localized in time and frequency. More precisely, considering a centered signal ($\int_{\mathbb{R}} t\,x(t)\,\mathrm{d}t = 0$) of unit energy ($\int_{\mathbb{R}} \lvert x(t)\rvert^2\,\mathrm{d}t = 1$) for simplicity, if we define the time spread (standard deviation) Δt and the frequency spread $\Delta\nu$ of a signal $x(t)$ as:

$$\Delta t = \sqrt{\int_{\mathbb{R}} t^2 |x(t)|^2 \, \mathrm{d}t} \tag{4.14}$$

$$\Delta \nu = \sqrt{\int_{\mathbb{R}} \nu^2 |\hat{x}(\nu)|^2 \, \mathrm{d}\nu}\,, \tag{4.15}$$

then the uncertainty principle states that:

$$\Delta t \, \Delta \nu \geq \frac{1}{4\pi}\,. \tag{4.16}$$

In quantum mechanics, since the momentum and position wave functions are related by a Fourier transform (up to Planck's constant), this inequality becomes the Heisenberg uncertainty principle.

4.8 Notable Fourier transforms

Signal	Time Domain	Frequency Domain
Rectangle	$\mathrm{rect}(t) = \begin{cases} 1 & \text{if } \lvert t\rvert < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$	$\widehat{\mathrm{rect}}(\nu) = \mathrm{sinc}(\nu) = \frac{\sin(\pi\nu)}{\pi\nu}$
Sinc	$x(t) = \mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$	$\hat{x}(\nu) = \mathrm{rect}(\nu)$
Gaussian	$x(t) = e^{-\pi t^2}$	$\hat{x}(\nu) = e^{-\pi \nu^2}$
Exponential decay	$x(t) = e^{-at} u(t), \quad a > 0$	$\hat{x}(\nu) = \frac{1}{a + j2\pi\nu}$
Two-sided exponential	$x(t) = e^{-a\lvert t\rvert}, \quad a > 0$	$\hat{x}(\nu) = \frac{2a}{a^2 + (2\pi\nu)^2}$
Dirac delta	$x(t) = \delta(t)$	$\hat{x}(\nu) = 1$
Constant	$x(t) = 1$	$\hat{x}(\nu) = \delta(\nu)$
Complex exponential	$x(t) = e^{j2\pi\nu_0 t}$	$\hat{x}(\nu) = \delta(\nu - \nu_0)$
Cosine	$x(t) = \cos(2\pi\nu_0 t)$	$\hat{x}(\nu) = \frac{1}{2}[\delta(\nu - \nu_0) + \delta(\nu + \nu_0)]$
Sine	$x(t) = \sin(2\pi\nu_0 t)$	$\hat{x}(\nu) = \frac{1}{2j}[\delta(\nu - \nu_0) - \delta(\nu + \nu_0)]$
Sign	$x(t) = \mathrm{sgn}(t) = \begin{cases} 1 & t > 0 \\ -1 & t < 0 \end{cases}$	$\hat{x}(\nu) = \frac{1}{j\pi\nu}$
Unit step	$x(t) = u(t)$	$\hat{x}(\nu) = \frac{1}{j2\pi\nu} + \frac{1}{2}\delta(\nu)$

Table 4.3: Notable Fourier Transform Pairs

4.9 Filtering

Thanks to the convolution theorem of the Fourier transform, the action of an LTI system with impulse response h on a signal x is:

$$y(t) = (h * x)(t) \tag{4.17}$$

$$\hat{y} = \hat{h}\,\hat{x}\,, \tag{4.18}$$

where $\hat{h}$, the Fourier transform of h, is called the transfer function.

We call the action of these LTI systems *filtering* of the input signal. It is important to understand that filters act on each frequency independently. We can study the transfer function of filters either to understand their effect on signals or to design filters with specific frequency response. This is widely used in signal processing, communications, control systems, and many other fields. This study generally amounts to studying the magnitude and phase of the transfer function as a function of frequency:

$$h(\nu) = \left|\hat{h}(\nu)\right| e^{j\,\varphi(\nu)}\,, \tag{4.19}$$

where $\varphi(\nu)$ is the phase response and $\left|\hat{h}(\nu)\right|$ is the magnitude response of the filter.

Several categories of filters can be described:

- Low-pass filters: These filters allow low-frequency components to pass through while attenuating high-frequency components. They are commonly used to remove high-frequency noise from signals.
- High-pass filters: These filters allow high-frequency components to pass through while attenuating low-frequency components. They are used to eliminate low-frequency noise or drift.
- Band-pass filters: These filters allow a specific range of frequencies to pass through while attenuating frequencies outside this range. They are used in applications such as audio processing and communications.
- Band-stop filters: These filters attenuate a specific range of frequencies while allowing frequencies outside this range to pass through. They are used to eliminate unwanted frequency components, such as interference.
- All-pass filters: These filters allow all frequencies to pass through but alter the phase of the signal.

Filters can be easily combined by multiplying their transfer functions. This property is particularly useful in designing complex filtering systems by cascading simpler filters.

5 Discrete Fourier Transform

5.1 Discrete-Time Fourier Transform

$x(t)$ is a continuous signal and its Fourier transform is:

$$\hat{x}(\nu) = \int_{\mathbb{R}} x(t)\, e^{-j\,2\pi\,\nu\,t}\ \mathrm{d}t \tag{5.2}$$

We define x_T the signal x sampled at interval of T seconds, it becomes:

$$\widehat{x_T}(\nu) = \sum_{n=-\infty}^{+\infty} T\, x(nT) e^{-j\,2\pi\,\nu\,T\,n} \tag{5.3}$$

in angular frequencies, taking $\omega = 2\pi\,\nu T$, the function $\hat{x}_{2\pi}(\omega)$ became periodic of period 2π. $\widehat{x_{2\pi}}(\nu)$ is called the Discrete-Time Fourier Transform (DTFT) of the discrete-time signal $x[n] = x(nT)$:

$$\widehat{x_T}(\omega) = \sum_{n=-\infty}^{+\infty} x[n]\, e^{-j\,\omega\, n}, \qquad \textbf{Forward DTFT} \tag{5.4}$$

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{x_T}(\omega)\, e^{j\,\omega\, n}\ \mathrm{d}\omega. \qquad \textbf{Inverse DTFT} \tag{5.5}$$

5.2 Discrete Fourier Transform

The Discrete Fourier Transform (DFT) is a sampled version of the DTFT. It is defined for a finite-length discrete-time signal $x[n]$ of length N as:

$$\hat{x}[k] = \sum_{n=0}^{N-1} x[n]\, e^{-j\,2\pi\, k\, \frac{n}{N}}, \qquad \textbf{Forward DFT} \tag{5.6}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{x}[k]\, e^{j\,2\pi\, k\, \frac{n}{N}}. \qquad \textbf{Inverse DFT} \tag{5.7}$$

where $\omega_N = e^{j\,\frac{2\pi}{N}}$ is the Nth-root of unity.

We can define the DFT matrix of size N as:

$$F_N[k,n] = e^{-j\,2\pi \frac{k\,n}{N}}, \quad k, n = 0, 1, ..., N-1 \tag{5.8}$$

$$\boldsymbol{F}_N = \begin{pmatrix} 1 & 1 & 1 & ... & 1 \\ 1 & \omega_N & \omega_N^2 & ... & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & ... & \omega_N^{2(N-1)} \\ ... & ... & ... & ... & ... \\ 1 & \omega_N^{N-1} & \omega_N^{2(N-1)} & ... & \omega_N^{(N-1)(N-1)} \end{pmatrix} \tag{5.9}$$

the discrete Fourier transform becomes:

$$\hat{\boldsymbol{x}} = \boldsymbol{F}_N \cdot \boldsymbol{x}, \qquad \textbf{Forward DFT} \tag{5.10}$$

$$\boldsymbol{x} = \frac{1}{N} \boldsymbol{F}_N^H \cdot \hat{\boldsymbol{x}} \qquad \textbf{Inverse DFT} \tag{5.11}$$

$$\boldsymbol{x} = \boldsymbol{F}_N^{-1} \cdot \hat{\boldsymbol{x}} \tag{5.12}$$

where $\boldsymbol{F}_N^H$ is the Hermitian transpose or conjugate transpose:

$$F_N^H[i,j] = \overline{F_N[j,i]} \tag{5.13}$$

.

As $\boldsymbol{F}_N^H \cdot \boldsymbol{F}_N = \frac{1}{N}\, \mathbf{I} \neq \mathbf{I}$, this DFT matrix is not unitary. This leads to the unitary definition of the DFT matrix:

$$\boldsymbol{U}_N = \frac{1}{\sqrt{N}} \boldsymbol{F}_N \tag{5.14}$$

and $\boldsymbol{U}_N^H \cdot \boldsymbol{U}_N = \mathbf{I}$.

In this lecture, for the sake of simplicity we will use the unitary matrix as the DFT matrix $\boldsymbol{F}_N = \boldsymbol{U}_N$ such that $\boldsymbol{F}_N^{-1} = \boldsymbol{U}_N^H$

The DFT can be computed efficiently using the Fast Fourier Transform (FFT) algorithm that applies the DFT matrix $\boldsymbol{F}_N$ in $O(N \log(N))$ operations.

5.3 Properties

The properties of the Discrete Fourier Transform are similar to those of the continuous Fourier transform described in Section 4.5. They can be derived from the properties of the DTFT.

Property	Time Domain	Frequency Domain
Linearity	$z[n] = a\,x[n] + b\,y[n]$	$\hat{z}[k] = a\,\hat{x}[k] + b\,\hat{y}[k]$
Time Shifting	$y[n] = x[n - n_0]$	$\hat{y}[k] = e^{-j\,2\pi\,k\,\frac{n_0}{N}}\,\hat{x}[k]$
Frequency Shifting	$y[n] = e^{j\,2\pi\,k_0\,\frac{n}{N}}\,x[n]$	$\hat{y}[k] = \hat{x}[(k - k_0) \bmod N]$
Time Reversal	$y[n] = x[-n \bmod N]$	$\hat{y}[k] = \hat{x}[-k \bmod N]$
Convolution	$z[n] = (x * y)[n] = \sum_{m=0}^{N-1} x[m]\,y[(n - m) \bmod N]$	$\hat{z}[k] = \hat{x}[k]\,\hat{y}[k]$
Multiplication	$z[n] = x[n]\,y[n]$	$\hat{z}[k] = \frac{1}{N} \sum_{\ell=0}^{N-1} \hat{x}[\ell]\,\hat{y}[(k - \ell) \bmod N]$
Conjugation	$y[n] = \overline{x[n]}$	$\hat{y}[k] = \overline{\hat{x}[-k \bmod N]}$
First Difference	$y[n] = x[n] - x[n - 1 \bmod N]$	$\hat{y}[k] = \left(1 - e^{-j\,2\pi\,\frac{k}{N}}\right)\hat{x}[k]$
Symmetry	$x[n] \in \mathbb{R}$	$\hat{x}[-k \bmod N] = \overline{\hat{x}[k]}$
Even Real Signals	$x[n] = x[-n \bmod N] \in \mathbb{R}$	$\hat{x}[k] \in \mathbb{R}$
Odd Real Signals	$x[n] = -x[-n \bmod N] \in \mathbb{R}$	$\hat{x}[k] \in j\,\mathbb{R}$

Table 5.4: Properties of Discrete Fourier Transform

5.4 Discrete convolution matrix

Given a finite-length discrete impulse response $\boldsymbol{h}$ of length M, the convolution matrix $\boldsymbol{H}$ generated by $\boldsymbol{h}$ for an input signal $\boldsymbol{x}$ of size N is defined as:

$$\boldsymbol{y} = \boldsymbol{H} \cdot \boldsymbol{x} \tag{5.15}$$

$$\boldsymbol{H} = \begin{pmatrix} h[0] & 0 & 0 & \dots & 0 \\ h[1] & h[0] & 0 & \dots & 0 \\ h[2] & h[1] & h[0] & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ h[M-1] & h[M-2] & h[M-3] & \dots & h[0] \\ 0 & h[M-1] & h[M-2] & \dots & h[1] \\ 0 & 0 & h[M-1] & \dots & h[2] \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & h[M-1] \end{pmatrix} \tag{5.16}$$

This matrix is of size $(N + M - 1) \times N$ and the output signal $\boldsymbol{y}$ is then of size $N + M - 1$. This formulation assumes that the vector $\boldsymbol{x}$ is zero outside its defined range.

Circulant Matrices: If we assume the input vector is periodic with period N, the matrix $\boldsymbol{H}$ becomes a Toeplitz circulant matrix of size $N \times N$:

$$\boldsymbol{H} = \begin{pmatrix} h[0] & h[N-1] & h[N-2] & ... & h[1] \\ h[1] & h[0] & h[N-1] & ... & h[2] \\ h[2] & h[1] & h[0] & ... & h[3] \\ ... & ... & ... & ... & ... \\ h[N-1] & h[N-2] & h[N-3] & ... & h[0] \end{pmatrix} \tag{5.17}$$

This matrix is Toeplitz as each descending diagonal from left to right is constant.

Diagonalization of convolution matrix: The circulant convolution matrix (also known as Toeplitz matrix) $\boldsymbol{H}$ of size N can be diagonalized using the DFT matrix $\boldsymbol{F}_N$ as:

$$\boldsymbol{H} = \boldsymbol{F}_N^{-1} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{F}_N \tag{5.18}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{H}$ given by the DFT of its first column:

$$\Lambda_{k,k} = \hat{h}_k \tag{5.19}$$

In other words,

$$\boldsymbol{H} = \boldsymbol{F}_N^{-1} \cdot \mathrm{diag}\left(\hat{h}\right) \cdot \boldsymbol{F}_N \tag{5.20}$$

where $\mathrm{diag}\left(\hat{h}\right)$ is the diagonal matrix containing the DFT of the impulse response $\boldsymbol{h}$.

Periodic Border Conditions:

When the convolution matrix $\boldsymbol{H}$ is circulant, it implies periodic border conditions for the input signal $\boldsymbol{x}$. This means that the signal is assumed to repeat itself outside its defined range. Specifically, for an input signal of length N, we have:

$$x[n + N] = x[n], \quad \forall n \in \mathbb{Z}\,. \tag{5.21}$$

This assumption may induce artifacts at the borders of the output signal, especially if the input signal has significant discontinuities at its boundaries.

5.5 Sampling

When a continuous-time signal $x(t)$ is sampled at regular intervals of period T, the resulting discrete-time signal is given by:

$$x[n] = x(nT)\,, \quad n \in \mathbb{Z}\,. \tag{5.22}$$

The sampling frequency is defined as $f_s = \frac{1}{T}$.

Shannon-Nyquist sampling theorem: When the continuous-time signal $x(t)$ is band-limited with a maximum frequency $f_{\max}$ (i.e., its Fourier transform $\hat{x}(\nu) = 0$ for $|\nu| > f_{\max}$), the **Shannon-Nyquist sampling theorem** states:

A continuous-time signal can be perfectly reconstructed from its samples if the sampling frequency satisfies:

$$f_s > 2\, f_{\max}\,. \tag{5.23}$$

If the sampling frequency is lower than twice the maximum frequency of the signal, **aliasing** occurs, leading to distortion in the reconstructed signal. Aliasing occurs when higher frequency components of the signal are misrepresented as lower frequency components in the sampled signal. Moiré patterns in images are a common example of aliasing.

Border artifacts: When sampling a continuous-time signal, border artifacts can occur due to the finite duration of the sampled signal. These artifacts are particularly noticeable when the original signal has discontinuities or rapid changes at the boundaries of the sampling interval. To mitigate border artifacts, several techniques can be employed:
- Removing the slope: The mean slope of the signal before sampling can help reduce discontinuities at the borders.
- Windowing: Applying a window function (*e.g.* Gaussian, Hanning, …) to the signal before sampling.
- Padding: Extending the signal with additional samples (*e.g.* zeros, mean, or mirrored values) can help minimize edge effects during processing.

6 Random signals

6.1 Definitions

Discrete random variable: A discrete random variable X is a variable that can take on a countable number of distinct values, each associated with a specific probability. The probability mass function (PMF) $p_{X(x)}$ defines the probability that the random variable takes on the value x:

$$p_X(x) = P(X = x), \quad \forall x \in S_X, \tag{6.2}$$

where S_X is the set of possible values of X. The PMF must satisfy the following properties:
- Non-negativity: $p_{X(x)} \geq 0, \quad \forall x \in S_X$
- Normalization: $\sum_{x \in S_X} p_{X(x)} = 1$.

Continuous random variable: A continuous random variable X is a variable that can take on an infinite number of values within a given range. The probability density function (PDF) $f_{X(x)}$ defines the likelihood that the random variable falls within an interval $[a, b]$ as:

$$P(a \leq X \leq b) = \int_a^b f_X(x)\ \mathrm{d}x \tag{6.3}$$

It can also be defined by its cumulative distribution function $F_X(x)$:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,\mathrm{d}u \tag{6.4}$$

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} P(X \leq x), \quad \forall x \in \mathbb{R}, \tag{6.5}$$

The PDF must satisfy the following properties:

- Non-negativity: $f_{X(x)} \geq 0, \quad \forall x \in \mathbb{R}$
- Normalization: $\int_{\mathbb{R}} f_{X(x)}\ \mathrm{d}x = 1$

Moments:

The n^{th} moment of a random variable X is defined as the expected value of its n^{th} power:

$$\mathbb{E}[X^n] = \begin{cases} \sum_{x \in S_X} x^n\, p_X(x) & \text{discrete random variables} \\ \int_{\mathbb{R}} x^n\, f_X(x)\ \mathrm{d}x & \text{continuous random variables} \end{cases} \tag{6.6}$$

- The first moment $(n = 1)$ is the mean or expected value $\mu_X = \mathbb{E}[X]$.

The n^{th} central moment of a random variable X is defined as the expected value of the n^{th} power of the deviation of X from its mean μ_X:

$$\mathbb{E}[(X - \mu_X)^n] = \begin{cases} \sum_{x \in S_X} (x - \mu_X)^n\, p_X(x) & \text{discrete} \\ \int_{\mathbb{R}} (x - \mu_X)^n\, f_X(x)\ \mathrm{d}x & \text{continuous} \end{cases} \tag{6.7}$$

- The second central moment $(n = 2)$ is the variance σ_X^2, which measures the spread of the random variable around its mean:

$$\sigma_X^2 = \mathbb{E}\left[(X - \mu_X)^2\right] \tag{6.8}$$

$$= \mathbb{E}[X^2] - \mu_X^2 \tag{6.9}$$

,

- The standard deviation σ_X is the square root of the variance:

$$\sigma_X = \sqrt{\sigma_X^2} \tag{6.10}$$

.

6.2 Random signals

A random signal $X(t, s)$, also known as a stochastic process, is a function of time (or another variable) whose amplitude at any given time t is a

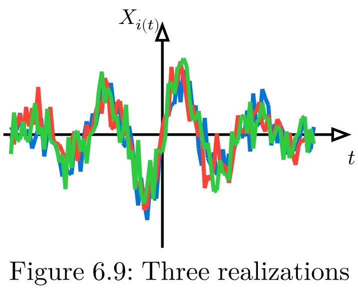

Figure 6.9: Three realizations of a non-stationary random signal

random variable. It is a set of functions of t, the set being indexed by s as illustrated in Figure 6.9. A random signal is thus a bivariate quantity. When $s = s_i$ is fixed, we get a realization of the random process, denoted $X(t, s_i)$ or, more simply, $X_i(t)$. When t is fixed, the random process reduces to a simple random variable. A random signal can be either continuous or discrete in time or value.

Correlation: The correlation function between two random signals $X(t)$ and $Y(t)$ is defined as:

$$\Gamma_{XY}(t_1, t_2) = \mathbb{E}\Big[X(t_1)\,\overline{Y(t_2)}\Big], \quad \forall t_1, t_2 \in \mathbb{R}, \tag{6.11}$$

The correlation function measures the statistical dependence between the values of both random signals at different times. It presents an Hermitian symmetry:

$$\Gamma_{XY}(t_1, t_2) = \overline{\Gamma_{YX}(t_2, t_1)} \tag{6.12}$$

.

Autocorrelation: When $X(t) = Y(t)$, the correlation function reduces to the **autocorrelation function**:

$$\Gamma_{X}(t_1, t_2) = \mathbb{E}\Big[X(t_1)\,\overline{X(t_2)}\Big], \quad \forall t_1, t_2 \in \mathbb{R}, \tag{6.13}$$

The autocorrelation function measures the statistical dependence of the random signal with itself at different times. It is a Hermitian function, real-valued along the diagonal, and $\Gamma_X(t_1, t_2) = \overline{\Gamma_X(t_2, t_1)}$.

AutoCovariance: The **autocovariance function** of a random signal $X(t)$ is defined as:

$$C_X(t_1, t_2) = \mathbb{E}\Big[(X(t_1) - \mu_X(t_1))\,\overline{X(t_2) - \mu_X(t_2)}\Big], \quad \forall t_1, t_2 \in \mathbb{R}, \tag{6.14}$$

$$= \Gamma_X(t_1, t_2) - \mu_X(t_1)\,\overline{\mu_X(t_2)} \tag{6.15}$$

The autocovariance function is also a Hermitian function with $C_X(t_1, t_2) = \overline{C_X(t_2, t_1)}$.

Stationarity: A random signal is **wide-sense stationary** if its mean and autocovariance are finite and independent of the choice of the origin of time:

$$\mathbb{E}\big[X^2(t)\big] < \infty, \quad \forall t \in \mathbb{R} \tag{6.16}$$

$$\mathbb{E}[X(t)] = \mu, \quad \forall t \in \mathbb{R} \tag{6.17}$$

$$\mathbb{E}[X(t)\, X^*(t + \tau)] = \gamma_X(\tau) \tag{6.18}$$

$\gamma_X(\tau)$ is the correlation function:

$$\gamma_X(\tau) = \Gamma(t, t+\tau) \tag{6.19}$$

Ergodicity:

A random signal is **ergodic** if its time averages are equal to its ensemble averages. This means that statistical properties of the random signal can be estimated from a single realization over time, rather than needing multiple realizations. For an ergodic random signal, the following holds:

- The time average of the signal equals its expected value:

$$\lim_{T\to\infty}\left(\frac{1}{2T}\right)\int_{-T}^{T} X(t)\ \mathrm{d}t = \mathbb{E}[X(t)]\,, \quad \forall t \in \mathbb{R}. \tag{6.20}$$

- The time average of the autocorrelation equals its ensemble autocorrelation:

$$\lim_{T\to\infty}\left(\frac{1}{2T}\right)\int_{-T}^{T} X(t)\,\overline{X(t+\tau)}\ \mathrm{d}t = \Gamma_X(\tau), \quad \forall \tau \in \mathbb{R}. \tag{6.21}$$

Discrete random signal: In the discrete case , the definitions are similar with $t \in \mathbb{Z}$.

- For a discrete random signal X and Y of size N and M respectively, the **correlation matrix** is a $N \times M$ matrix defined as:

$$\Gamma_{XY}[m,n] = \mathbb{E}\left[X[m]\,\overline{Y[n]}\right], \quad \forall m,n \in \mathbb{Z}, \tag{6.22}$$

$$\boldsymbol{\Gamma}_{XY} = \mathbb{E}[\boldsymbol{X} \cdot \boldsymbol{Y}^H] \tag{6.23}$$

- When $Y = X$, the **autocorrelation matrix** is :

$$\boldsymbol{\Gamma}_X = \mathbb{E}[X \cdot X^H] \tag{6.24}$$

- the **autocovariance matrix** is:

$$\boldsymbol{C}_X = \mathbb{E}[X \cdot X^H] - E[X] \cdot E[X]^H \tag{6.25}$$

By construction this autocovariance matrix is an Hermitian matrix and its diagonal is real valued and equal to the variance of the signal:

$$\sigma_X^2[n] = C_X[n,n] \tag{6.26}$$

In the case of a wide-sense stationary discrete random signal, the autocovariance matrix is a Toeplitz matrix:

$$C_X[m,n] = \gamma_X[m-n], \quad \forall m,n \in \mathbb{Z}\,. \tag{6.27}$$

As a consequence, the autocovariance matrix can be diagonalized using the DFT matrix:

$$\boldsymbol{C}_X = \boldsymbol{F}_N^{-1} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{F}_N \tag{6.28}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of C_X given by the DFT of its first column:

$$\Lambda_{k,k} = \widehat{\gamma_X}_k \tag{6.29}$$

In other word

$$\boldsymbol{C}_X = \boldsymbol{F}_N^{-1} \cdot \operatorname{diag}(\widehat{\gamma_X}) \cdot \boldsymbol{F}_N \tag{6.30}$$

where $\operatorname{diag}(\widehat{\gamma_X})$ is the diagonal matrix containing the DFT of the autocovariance sequence γ_X.

If it exists, the inverse of the covariance matrix $\boldsymbol{W} = \boldsymbol{C}^{-1}$ is called the **precision matrix**.

6.3 Fourier transform of random signals

The Fourier transform of a random signal is itself a random signal in the frequency domain. The statistical properties of the Fourier transform of a random signal can be analyzed similarly to those of the original signal.

Power Spectral Density:

For finite time random signal $X(t)$, we define the **Energy Spectral Density** is given by:

$$S_X(\nu) = \left|\hat{X}(\nu)\right|^2 \tag{6.31}$$

For a (wide-sense) stationary random signal $X(t)$, we define the **Power Spectral Density** (PSD) $S_{X(\nu)}$ as:

$$S_X(\nu) = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\left|\widehat{X_T}(\nu)\right|^2\right], \quad \forall \nu \in \mathbb{R}, \tag{6.32}$$

where $\widehat{X_T}(\nu)$ is the Fourier transform of the truncated signal $X_T(t) = X(t)\operatorname{rect}(\frac{t}{T})$. The power spectral density describes how the power of the random signal is distributed across different frequency components.

6.4 Wiener-Khinchin theorem

The Wiener-Khinchin theorem states that the power spectral density $S_{X(\nu)}$ of a wide-sense stationary random signal $X(t)$ is the Fourier transform of its autocorrelation function $\gamma_X(\tau)$:

$$\begin{aligned} S_X(\nu) &= \int_{\mathbb{R}} \gamma_X(\tau)\, e^{-j\,2\pi\,\nu\,\tau}\ \mathrm{d}\tau \\ \gamma_X(\tau) &= \int_{\mathbb{R}} S_X(\nu)\, e^{j\,2\pi\,\nu\,\tau}\ \mathrm{d}\nu\,. \end{aligned} \tag{6.33}$$

Demonstration:

Consider the truncated signal $X_T(t) = X(t)\operatorname{rect}(\frac{t}{T})$. Its Fourier transform is:

$$\begin{aligned}
\widehat{X_T}(\nu) &= \int_{\mathbb{R}} X_T(t)\, e^{-j\,2\pi\,\nu\,t}\ \mathrm{d}t \\
&= \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t)\, e^{-j\,2\pi\,\nu\,t}\ \mathrm{d}t
\end{aligned}$$

The expected value of the squared magnitude of the Fourier transform is:

$$\begin{aligned}
\frac{1}{T}\mathbb{E}\Big[\big|\widehat{X_T}(\nu)\big|^2\Big] &= \frac{1}{T}\mathbb{E}\Big[\widehat{X_T}(\nu)\,\overline{\widehat{X_T}(\nu)}\Big] \\
&= \frac{1}{T}\mathbb{E}\left[\int_{-\frac{T}{2}}^{\frac{T}{2}} X(t_1)\, e^{-j\,2\pi\,\nu\,t_1}\ \mathrm{d}t_1 \int_{-\frac{T}{2}}^{\frac{T}{2}} \overline{X(t_2)}\, e^{j\,2\pi\,\nu\,t_2}\ \mathrm{d}t_2\right] \\
&= \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbb{E}\Big[X(t_1)\,\overline{X(t_2)}\Big]\, e^{-j\,2\pi\,\nu\,(t_1-t_2)}\ \mathrm{d}t_1\ \mathrm{d}t_2 \\
&= \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}\int_{-\frac{T}{2}}^{\frac{T}{2}} \gamma_X(t_1-t_2)\, e^{-j\,2\pi\,\nu\,(t_1-t_2)}\ \mathrm{d}t_1\ \mathrm{d}t_2 \\
&= \frac{1}{T}\left(\int_{-\frac{T}{2}}^{\frac{T}{2}} \gamma_X(\tau)\, e^{-j\,2\pi\,\nu\,\tau}\ \mathrm{d}\tau\right)\left(\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}t_2\right) \\
&= \int_{-\frac{T}{2}}^{\frac{T}{2}} \gamma_X(\tau)\, e^{-j\,2\pi\,\nu\,\tau}\ \mathrm{d}\tau
\end{aligned}$$

Taking the limit when T goes to infinity, we obtain the Wiener-Khinchin theorem.

$$\begin{aligned}
S_X(\nu) &= \lim_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\big|\widehat{X_T}(\nu)\big|^2\Big] \\
&= \widehat{\gamma_X}(\nu)
\end{aligned}$$

6.5 White noise

A random signal $X(t)$ is called **white noise** if its power spectral density is constant across all frequencies:

$$S_X(\nu) = \sigma_X^2, \quad \forall \nu \in \mathbb{R}, \tag{6.34}$$

where σ_X^2 is the variance of the noise. This implies that the autocorrelation function of white noise is a Dirac delta function:

$$\gamma_X(\tau) = \sigma_X^2\,\delta(\tau), \quad \forall \tau \in \mathbb{R}. \tag{6.35}$$

This means that white noise has no correlation between its values at different times and its covariance matrix is proportional to the identity matrix.

Signal to noise ratio: For a signal $y(t) = x(t) + n(t)$ that is the sum of a signal of interest x and an additive noise term n, the Signal-to-Noise Ratio (SNR) is a measure used to quantify the level of the desired signal relative to the level of background noise. It is defined as the ratio of the power of the signal to the power of the noise:

$$\begin{aligned}\mathrm{SNR} &= 10\log_{10}\left(\frac{P_x}{P_n}\right)\,\mathrm{dB} \\ &= 10\log_{10}\frac{\int_{\mathbb{R}}\mathbb{E}(|x(t)|^2)\,\mathrm{d}t}{\int_{\mathbb{R}}\mathbb{E}(|n(t)|^2)\,\mathrm{d}t} \\ &= 10\log_{10}\frac{\int_{\mathbb{R}}\mathbb{E}(|\hat{x}(\nu)|^2)\,\mathrm{d}\nu}{\int_{\mathbb{R}}\mathbb{E}(|\hat{n}(\nu)|^2)\,\mathrm{d}\nu} \qquad (6.36)\end{aligned}$$

If the noise is white: $P_n = \sigma_n^2$ and the SNR becomes:

$$\begin{aligned}\mathrm{SNR} &= 10\log_{10}\left(\frac{\sigma_x^2}{\sigma_n^2}\right) \\ &= 20\log_{10}\left(\frac{\sigma_x}{\sigma_x}\right) \qquad (6.37)\end{aligned}$$

where σ_x is the standard deviation of the signal

6.6 Filtering random signal

When a wide-sense stationary random signal $X(t)$ is passed through a linear time-invariant (LTI) system with impulse response $h(t)$, the output random signal $Y(t)$ is also wide-sense stationary. The power spectral density of the output signal $Y(t)$ can be determined using the transfer function of the LTI system $\hat{h}$. The power spectral density of the output signal $Y(t)$ is given by:

$$S_Y(\nu) = \left|\widehat{h(\nu)}\right|^2 S_X(\nu), \quad \forall\nu\in\mathbb{R}. \qquad (6.38)$$

Whitening: Whitening is the process of transforming a random signal so that its power spectral density becomes equal to one across all frequencies, effectively converting it into white noise. The Whitening filter w is defined such that:

$$\begin{aligned}S_Y(\nu) &= \left|\widehat{w(\nu)}\right|^2 S_X(\nu), \quad \forall\nu\in\mathbb{R} \\ w(\nu) &= (S_X(\nu))^{-\frac{1}{2}} \qquad (6.39)\end{aligned}$$

For discrete signal the whitening matrix $\boldsymbol{B}$ is the square-root of the precision matrix:

$$
\begin{aligned}
\boldsymbol{B} &= \boldsymbol{C}^{-\frac{1}{2}} \\
&= \boldsymbol{W}^{\frac{1}{2}} \\
&= \boldsymbol{F}^{-1} \cdot \operatorname{diag}\left(\frac{1}{\sqrt{\widehat{\gamma_X}}}\right) \cdot \boldsymbol{F}
\end{aligned} \tag{6.40}
$$

Gaussian process generation: A Gaussian process is a random signal where any finite collection of samples follows a multivariate normal distribution. To generate such a Gaussian process with a specified autocovariance function $\gamma_X(\tau)$, we just need to filter white Gaussian noise signal $n(t)$ with zero mean and unit variance by the inverse of the whitening filter.

In the discrete case, with $\boldsymbol{n}$ a white Gaussian noise signal with zero mean and unit variance, the generated signal is:

$$
\boldsymbol{x} = \boldsymbol{B}^{-1} \cdot \boldsymbol{n} \tag{6.41}
$$

And its covariance matrix is:

$$
\begin{aligned}
\boldsymbol{C}_X &= \mathbb{E}(\boldsymbol{x} \cdot \boldsymbol{x}^H) \\
&= \mathbb{E}\left(\boldsymbol{B}^{-1} \cdot \boldsymbol{n} \cdot (\boldsymbol{B}^{-1} \cdot \boldsymbol{n})^H\right) \\
&= \mathbb{E}(\boldsymbol{B}^{-1} \cdot \boldsymbol{n} \cdot \boldsymbol{n}^H \cdot \boldsymbol{B}^{-H}) \\
&= \boldsymbol{B}^{-1} \cdot \mathbb{E}(\boldsymbol{n} \cdot \boldsymbol{n}^H) \cdot \boldsymbol{B}^{-H} \\
&= \boldsymbol{B}^{-1} \cdot \boldsymbol{B}^{-H} \\
&= \boldsymbol{F}^{-1} \cdot \operatorname{diag}\left(\sqrt{\widehat{\gamma_X}}\right) \cdot \boldsymbol{F} \cdot \boldsymbol{F}^{-1} \cdot \operatorname{diag}\left(\sqrt{\widehat{\gamma_X}}\right) \cdot \boldsymbol{F} \\
&= \boldsymbol{F}^{-1} \cdot \operatorname{diag}\left(\widehat{\gamma_X}\right)) \cdot \boldsymbol{F}
\end{aligned}
$$

This process ensures that the generated Gaussian process $X(t)$ has the desired autocovariance function

Connection with Principal Component Analysis:

Principal Component Analysis (PCA) is closely related to the whitening process and provides a complementary perspective on signal decorrelation. Indeed, PCA is based on the diagonalization of the covariance matrix $\boldsymbol{C}$ and for stationary signal this matrix is naturaly diagonalized by the DTF:

$$
\boldsymbol{C}_X = \boldsymbol{F}^{-1} \operatorname{diag}(\widehat{\gamma_{\boldsymbol{X}}})\boldsymbol{F} \tag{6.42}
$$

Then the principal components are the row of the DFT matrix (some complex exponentials) corresponding to the highest values of $|\widehat{\gamma_{\boldsymbol{X}}}|^2$.

Bibliography

Freeman, D. (2011). 6.003: Signals And Systems. Retrieved from https://youtube.com/playlist?list=PLUl4u3cNGP61kdPAOC7CzFjJZ8f1eMUxs&si=EIZ02tZiU6OtpFWd

Hayashi, C. (1998). What is data science? Fundamental concepts and a heuristic example. In *Data Science, Classification, and Related Methods: Proceedings of the Fifth Conference of the International Federation of Classification Societies (IFCS-96), Kobe, Japan, March 27–30, 1996* (pp. 40–51).

Oppenheim, A. V., Willsky, A. S., & Nawab, S. H. (1997). *Signals & systems*. Pearson Educación.