Introduction to Signal Processing

Guillaume Tochon and Jim Newton

September 22, 2026

Contents

Chapter 0

Introduction

The content in this volume is a result primarily of a French to English translation of hand-written lecture notes prepared by Guillaume Tochon. Many sections however, have been augmented and clarified. Some proofs have been added and some background information has been supplied.

0.1 Generalities

Welcome to the marvelous world of the mathematics of signal processing!

We start with a very general, informal, definition of a signal.

Definition 0.1.1 (*signal*)

A *signal* is the observation or the measurement of a physical or chemical phenomenon which varies in time or space and conveys information.

Table 1 shows some non-exhaustive examples of different types of signals. In short, a *signal* is a sensor measurement representing some real physical phenomenon.

Definition 0.1.2 (*signal processing*)

A discipline at the interface of mathematics, electronics (analog and digital), and computer science (informatics) with the purpose of designing algorithms (methods, techniques) for transforming (manipulating, studying) signals in order to solve problems.

Signals may therefore have natural or artificial origins, spanning a variety of physical phenomena, and may be exploited for very diverse applications

- Red-eye correction in a digital photograph
- Avoid spectral folding prior to sampling an audio signal.
- ...

Signal processing provides a unified framework of mathematical operations, vocabulary, and algorithms for all of this.

As the possible inputs and outputs are diverse, it is worth classifying all this into different categories.

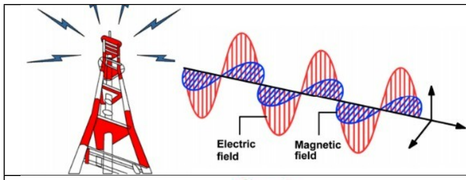
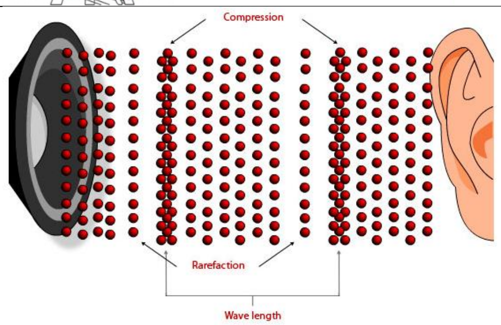
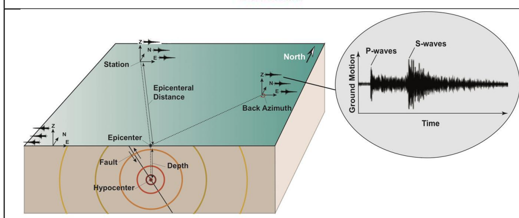
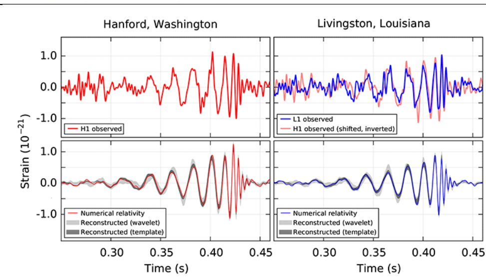
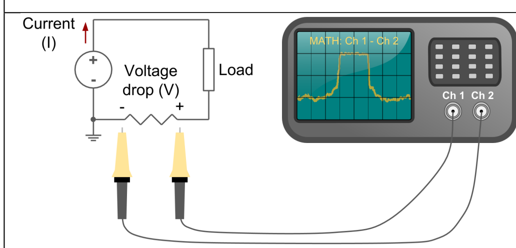
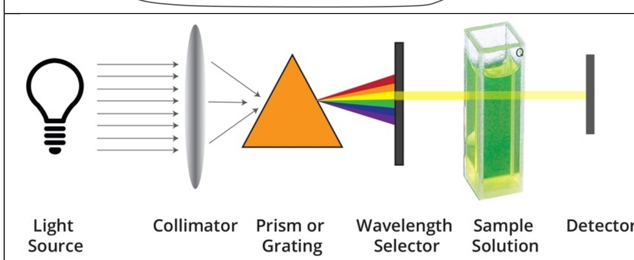
<table>
<tr><td>
</td><td>Light</td></tr>
<tr><td>
</td><td>Sound</td></tr>
<tr><td>
</td><td>Seismic</td></tr>
<tr><td>
</td><td>Gravitational</td></tr>
<tr><td>
</td><td>AC Current or Voltage</td></tr>
<tr><td>
</td><td>Chemical</td></tr>
</table>

Table 1: Examples of signals

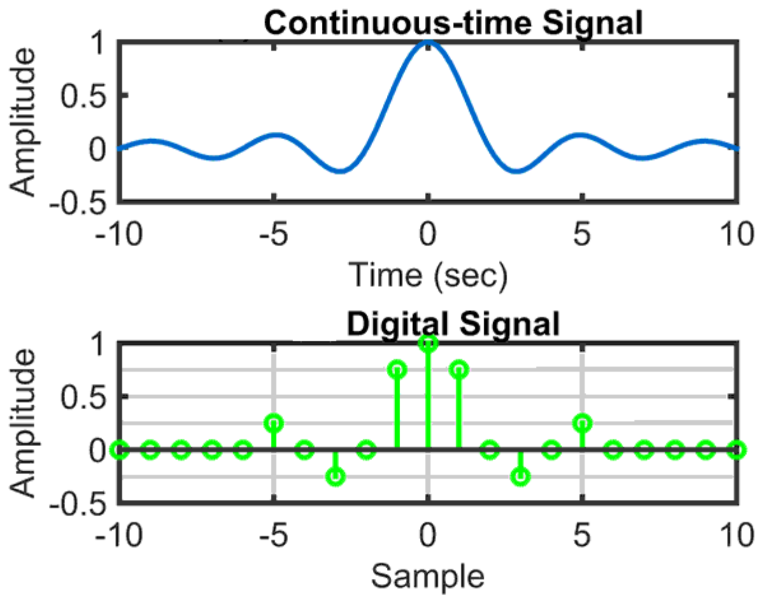


Figure 1: Analog vs Digital signal

0.2 Rough classification of signals

To be well-positioned for the scope of this course we begin with some tentative rough categories.

As shown in Figure 1, a signal may be analog or digital.

Definition 0.2.1 (***analog signal***)

An *analog* signal varies in a continuous manner with respect to its input variables and in its output values — all measurable physical quantities.

Definition 0.2.2 (***digital signal***)

A *digital* signal varies in a discrete manner with respect to its input and output values. A discrete signal may be stored in finite memory to be manipulated by a computer.

As shown in Figure 2, a signal may be deterministic or stochastic.

Definition 0.2.3 (***deterministic signal***)

A *deterministic* signal is perfectly clean, without external perturbations and without measurement noise.

Definition 0.2.4 (***stochastic signal***)

A *stochastic* signal involves one or more random processes (noise), in addition to its normal behavior. Welcome to the real world.

In this course, we find ourselves most of the time in the ideal, theoretical world of analog, deterministic signals.

- We will, however, talk a bit about digital signals in the TPs.
- We will talk about noise from time to time, but without introducing the mathematical tools for rigorous treatment of stochastic signals.

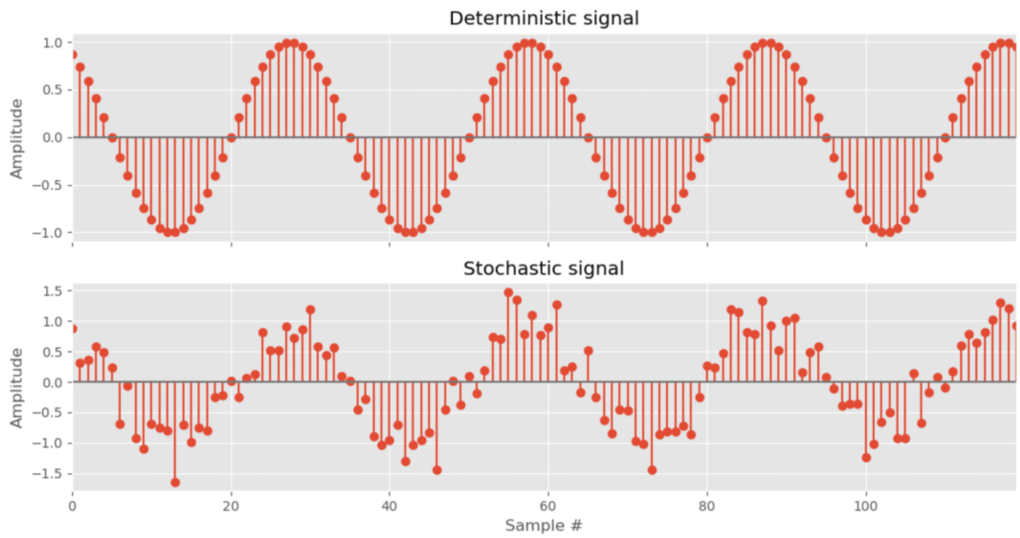


Figure 2: Deterministic vs Stochastic signal

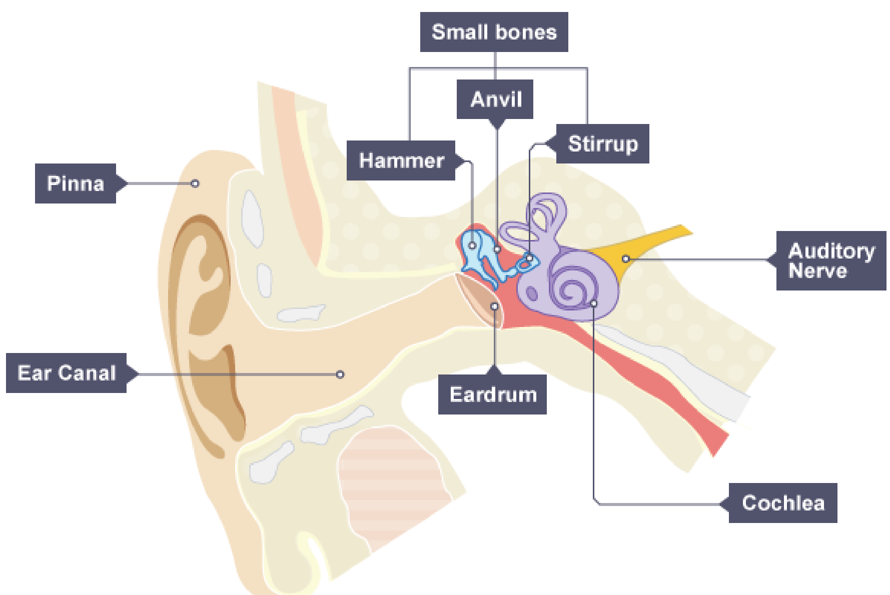


Figure 3: External and internal (inner) anatomy of the human ear

0.3 Example: Function of the human ear

Figure 3 shows an image of the human ear.

1. The acoustic wave is amplified by the external ear canal,
2. causing the eardrum to vibrate, transmitting the wave to the inner ear, via the hammer, anvil, and stirrup.
3. This causes a vibration within the endolymph fluid in the cochlea, which vibrates the hair cell filaments lining the cochlea.
4. These tiny hair cells convert the mechanical movement into electrical impulses to the brain via the auditory nerve. Fast (high-frequency) vibrations attenuate more quickly than slow (low-frequency) vibrations as they travel through the fluid, so they are detected by the hair cells near the beginning of the cochlea. The further into the cochlea a vibration travels before being detected, the slower (lower-frequency) it is.

The ear therefore decomposes the sound wave into components consisting of high, medium, and low frequencies.

> The ear is a Fourier transform.

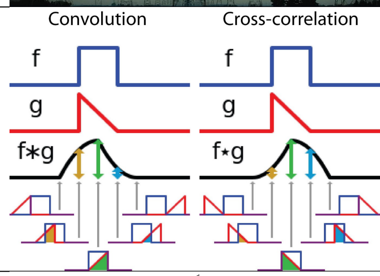
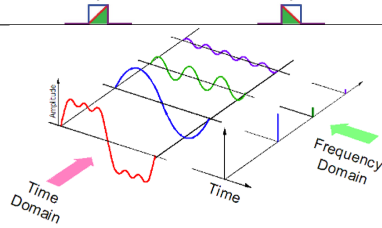
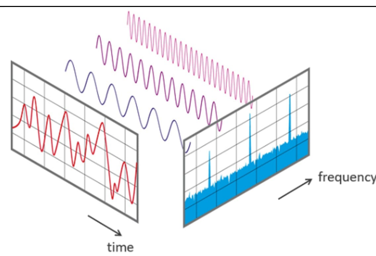
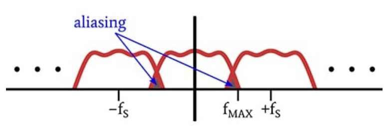
0.4 Objectives and syllabus

0.4.1 Global objective

To answer the question: why do we sample sound at 44.1 kHz?

The human ear detects sounds up to ≈ 20 kHz, and the Shannon Sampling Theorem stipulates that we must sample at twice the maximum frequency. But what are *sampling* and *frequency*?

0.4.2 Syllabus

<table>
<tr><th>Lectures</th><th>Chapter</th><th></th><th>Topic</th></tr>
<tr><td>1</td><td>1</td><td></td><td>Intro, Energy, and Power</td></tr>
<tr><td>2</td><td>2</td><td>
</td><td>Correlation and Convolution</td></tr>
<tr><td>3</td><td>3</td><td>
</td><td>Fourier Series</td></tr>
<tr><td>4,5</td><td>4</td><td>
</td><td>Fourier Transform</td></tr>
<tr><td>6,7</td><td>5</td><td>
</td><td>A to D Conversion
Sampling and the Shannon Sampling Theorem
Quantization</td></tr>
</table>

0.5 How to Read This Document

Throughout this document, margin annotations and colored boxes serve as navigation aids.

Margin annotations. Icons appearing in the margin signal the nature of the adjacent content.

- **Essential** — a core concept you must understand and be able to apply.
- **Recall** — a result or technique introduced elsewhere; look it up if needed.
- **Warning** — a common mistake or a subtle point requiring care.
- **Enrichment** — optional deeper material, or a building block used only to support another proof; not essential to memorize.

Colored boxes. Different types of mathematical content are distinguished by background color.

Chapter Objectives — the skills and understanding you should have by the end of the chapter.

Theorem — a mathematical statement that has been proven true.

Definition — a precise meaning assigned to a mathematical term.

Proposition — a useful mathematical claim, with short proof.

Lemma — intermediate result used in the proof of a larger theorem.

Corollary — a result following directly from a preceding theorem.

Example — a worked illustration of a concept or technique.

Proof — a logical argument establishing a claim.

Remark — a side comment clarifying notation or interpretation.

Reminder — a recap of material from an earlier chapter.

Exercise — a problem left for the reader to work out.

0.6 Review

1. Vague definition of signal and signal processing
2. Examples of signals
3. Classification of Signals:
    - analog vs digital
    - deterministic vs stochastic
4. Example of human ear
5. Purpose of the course: Shannon Sampling Theorem
6. Syllabus

Chapter 1

Energy and Power of a Signal

Chapter Objectives

- Define a signal precisely (bounded, piecewise-continuous, complex-valued) and build the vector-space structure (inner product, norm, distance) that the space of signals inherits.
- Compute the energy of a signal, $E_x = \langle x, x \rangle$, and determine whether a given signal has finite energy.
- Recognize that periodic signals have infinite energy, and compute their average power, P_x, instead.
- Exploit periodicity to reduce the computation of average power to a single period.

We would like to define *signal* for this course, but we must first define piecewise continuous.

Definition 1.0.1 (*piecewise continuous*)

Let $x : [a, b] \to \mathbb{C}$. We say that x is *piecewise continuous, PWC*, on $[a, b]$ if

1. $\lim\limits_{t \to \tau^+} x(t)$ exists for $a \le \tau < b$
2. $\lim\limits_{t \to \tau^-} x(t)$ exists for $a < \tau \le b$,
3. x is continuous on $[a, b]$ except at finitely many points.

If $x : \mathbb{R} \to \mathbb{C}$, then x is said to be PWC (on $\mathbb{R}$) if x is PWC on every closed interval $[a, b] \subset \mathbb{R}$.

Figure 1.1 [Left] shows an example of a function

$$f(t) = \begin{cases} 0 & ;\ \text{if } t = 0 \\ \sin(1/t) & ;\ \text{if } \textit{otherwise}\,, \end{cases}$$

which is not PWC, because the right and left limits at zero do not exist. Figure 1.1 [Right] shows an example of a function

$$f(t) = \begin{cases} \operatorname{mod}(t \left\lfloor \frac{1}{t} \right\rfloor, 3) & ;\ \text{if } t \neq 0 \\ 0 & ;\ \text{if } t = 0 \end{cases}$$

which is not PWC, because every interval around zero contains infinitely many discontinuities.

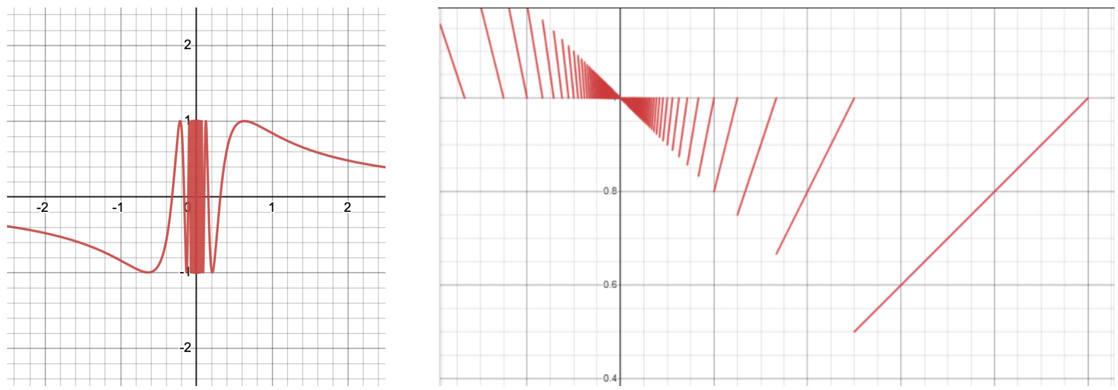


Figure 1.1: Examples of functions which are not PWC

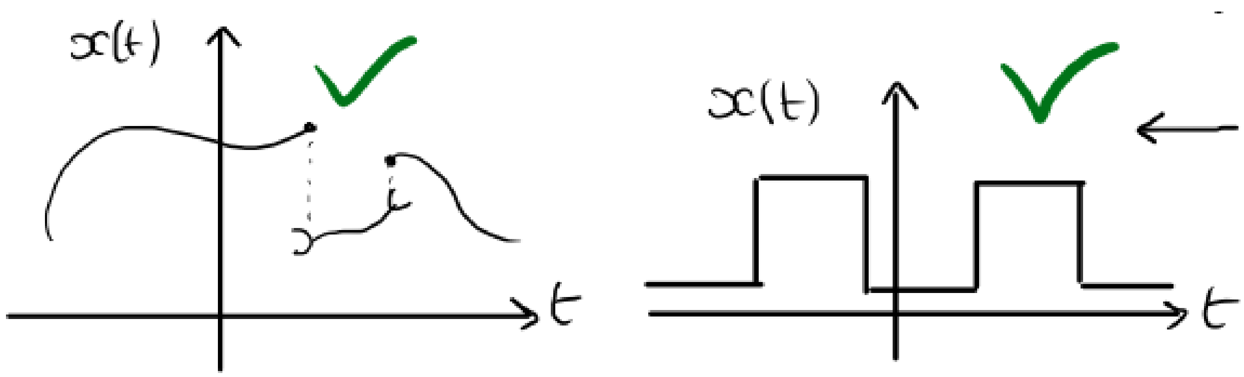


Figure 1.2: Example functions which are signals

Definition 1.0.2 (***complex valued signal***)

Let $I \subseteq \mathbb{R}$ be a closed interval or $I = \mathbb{R}$. A *signal*, x, is a complex valued function of one variable (this is already complicated enough)

$$x : I \to \mathbb{C}\,,$$

for which

- x is bounded: $\exists M > 0$ such that $|x(t)| < M \ \forall t \in I$
- x is PWC on I.

Normally, t represents time, but we also allow $t < 0$.

While we may consider complex valued, $\mathbb{C}$, signals, as is indicated in Definition 1.0.2, the majority of signals we will discuss are real valued, $\mathbb{R}$.

We will also consider the graph $(t, x(t))$ as the temporal representation of signal x.

Figure 1.2 illustrates some functions which are signals, while Figure 1.3 illustrates some functions which are not signals.

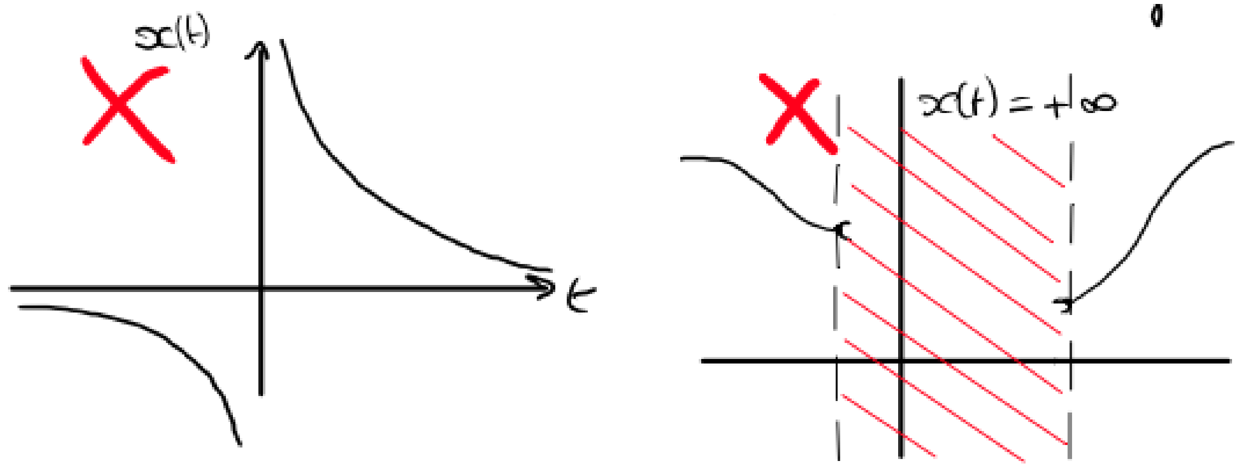


Figure 1.3: Example functions which are not signals

1.1 Groups and Fields

Definition 1.1.1 (*Abelian group*)

A set G, together with a binary operation, $\circ$, is called a *group*, if the following hold, for $a, b, c \in G$,

$$a \circ b \in G \qquad \text{; closed} \tag{1.1}$$

$$a \circ (b \circ c) = (a \circ b) \circ c \qquad \text{; associative} \tag{1.2}$$

$$\exists e \in G \mid a \in G \implies e \circ a = a \circ e = a \qquad \text{; identity} \tag{1.3}$$

$$\exists a^{-1} \in G \mid (a \circ a^{-1} = e) \wedge (a^{-1} \circ a = e) \qquad \text{; inverse} \tag{1.4}$$

If we say that $(G, \circ)$ is a group, we mean that G is a group with binary operation $\circ$.

If $(G, \circ)$ is a group and $\forall a, b \in G$, we have $a \circ b = b \circ a$, then we say $(G, \circ)$ is an *Abelian* group.

Definition 1.1.2 (*field*)

A set K, together with two binary operators, $+$ and $\cdot$, is called a *field* if:

$$(K, +) \text{ is an Abelian group, whose identity is denoted } 0 \tag{1.5}$$

$$(K \setminus \{0\}, \cdot) \text{ is an Abelian group, whose identity is denoted } 1 \tag{1.6}$$

$$a, b, c \in K \implies a \cdot (b + c) = (a \cdot b) + (a \cdot c) \tag{1.7}$$

1.2 Vector Space

We assume the reader is already familiar with vector spaces—at least finite dimensional vector spaces. We would like to view the set of signals as a vector space. As a reminder, we state here the definition of *vector space*. We state the definition in terms of *group* (Definition 1.1.1) and *field* (Definition 1.1.2).

Definition 1.2.1 (*vector space*)

A *vector space* is an (additive) Abelian group V together with a field K and a scalar multiplication of elements of K with elements of V such that if $u, v \in V$ and $a, b \in K$, the following hold:

$$av \in V \qquad ; \text{closure of scalar multiplication} \tag{1.8}$$
$$a(bv) = (ab)v \qquad ; \text{association of scalar multiplication} \tag{1.9}$$
$$a(u+v) = au + av \qquad ; \text{distribution law 1} \tag{1.10}$$
$$(a+b)v = av + bv \qquad ; \text{distribution law 2} \tag{1.11}$$
$$1v = v \qquad ; \text{scalar identity interaction} \tag{1.12}$$

Vector spaces come in various forms, and many of these particular forms have names. Some vector spaces are normed, if every vector has a *length* (or norm) which generates a distance. A complete normed vector space is called a Banach space, meaning that every Cauchy sequence of vectors *converges* to a vector within the space. Some vector spaces are inner product spaces, in that an *inner product* is defined which fulfills Definition 1.4.1. If the norm of a Banach space is defined by an inner product, *i.e.* if $\|v\|^2 = \langle v, v\rangle$, then the space is called a *Hilbert* space.

In this course we are interested in equivalence classes of signals. Two signals are considered to be in the same equivalence class if they are pointwise equal except perhaps at their points of discontinuity. With the help of definitions 1.3.1, 1.3.2, and Equation (1.32), can say that two signals, x and y, are in the same equivalence class if

$$d(x,y) = \|x-y\| = \sqrt{\int_{\mathbb{R}} |x(t)-y(t)|^2 \, \mathrm{d}t} = 0\,,$$

or equivalently,

$$\int_{\mathbb{R}} |x(t)-y(t)|^2 \, \mathrm{d}t = 0\,.$$

This set of equivalence classes forms an additive group of *signals*. This set of functions (or more correctly, equivalence classes of functions) forms a vector space where we will examine Fourier sequences of functions and consider the signals those sequences converge to. The question under which conditions this vector space is a Hilbert space is studied in *functional analysis* and is out of the scope of this course.

Theorem 1.2.2
The set of signals as in Definition 1.0.2 forms a vector space.

- If x_1, x_2 are signals, and $\alpha \in \mathbb{R}$ (or $\mathbb{C}$) then $x_1 + \alpha x_2$ is a signal.
- We may define a basis (but not the basis you know from Linear Algebra).
- We may define an inner product (scalar product).
- We may define a norm, and therefore a distance.

1.3 Norm

As mentioned above, we'd like to measure the *distance* between two elements of the vector space of signals. We will define *distance* and *norm* and Theorem 1.3.3 asserts that we can define a distance in terms of a norm.

Definition 1.3.1 (*distance*)

If V is a vector space, a function $d : V \times V \to \mathbb{R}$, is called a *distance* function, if it satisfies the following, for $u, v, w \in V$:

$$d(u, w) = d(w, u) \tag{1.13}$$
$$d(u, w) \geq 0 \tag{1.14}$$
$$d(u, w) = 0 \iff u = w \tag{1.15}$$
$$d(u, w) \leq d(u, v) + d(v, w) \tag{1.16}$$

Definition 1.3.2 (*norm*)

If V is a vector space over $\mathbb{C}$ and there exists a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties, then the function is called a *norm* for the vector space. For $u, v \in V$, and $\alpha \in \mathbb{C}$:

$$\|\alpha u\| = |\alpha|\|u\| \tag{1.17}$$
$$\|u\| \geq 0 \tag{1.18}$$
$$\|u\| = 0 \iff u = 0 \tag{1.19}$$
$$\|u + v\| \leq \|u\| + \|v\| \tag{1.20}$$

Theorem 1.3.3

Equation (1.21) defines a distance.

$$d(u, w) = \|u - w\| \tag{1.21}$$

Proof. Exercise for student

1.4 Inner Product

In a vector space, we may compute the similarity between two vectors, x, y, using the inner product $\langle x, y \rangle$. Definition 1.4.1 introduces the *inner product*. Some texts refer to this as a *scalar product* because it is a binary function which maps pairs of vectors from the vector space to a scalar in the underlying field.

Definition 1.4.1 (*inner product, scalar product*)

In a vector space, V over $\mathbb{C}$, a function, $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$, is a called an *inner product* if it satisfies the following axioms: For $u, v, w \in V$ and $\alpha \in \mathbb{C}$:

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \tag{1.22}$$
$$\langle \alpha v, w \rangle = \alpha \langle v, w \rangle \tag{1.23}$$
$$\langle v, w \rangle = \overline{\langle w, v \rangle} \qquad ; \overline{v} \text{ denotes complex conjugate} \tag{1.24}$$
$$\langle v, v \rangle \geq 0 \tag{1.25}$$
$$\langle v, v \rangle = 0 \iff v = 0 \tag{1.26}$$

Note about Equation (1.25), $\langle v, v \rangle \geq 0$, this inequality only makes sense if $\langle v, v \rangle$ is real; however, we stated in the assumptions of Definition 1.4.1 that $\langle \cdot, \cdot \rangle \in \mathbb{C}$. How do we know that $\langle v, v \rangle$ is real? To make sense of (1.25) we note that $\langle v, v \rangle$ is real, as shown in Corollary 1.4.2.

Corollary 1.4.2

$$\langle v, v \rangle \in \mathbb{R} \tag{1.27}$$

Proof. Immediate consequence of (1.24).

Corollary 1.4.3

$$\langle 0, u\rangle = \langle u, 0\rangle = 0 \tag{1.28}$$

Proof.

$$\begin{aligned}\langle 0, u\rangle &= \langle 0u, u\rangle && ;\text{ by } (1.23)\\ &= 0\,\langle u, u\rangle \\ &= 0 && (1.29)\end{aligned}$$

Furthermore,

$$\begin{aligned}\langle u, 0\rangle &= \overline{\langle 0, u\rangle} && ;\text{ by } (1.24)\\ &= \overline{0} && ;\text{ by } (1.29)\\ &= 0\end{aligned}$$

From Definition 1.4.1 we easily derive Equations (1.30) and (1.31) :

Corollary 1.4.4

$$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle \tag{1.30}$$

Proof.

$$\begin{aligned}\langle u, v+w\rangle &= \overline{\langle v+w, u\rangle} && ;\text{ by } (1.24)\\ &= \overline{\langle v, u\rangle + \langle w, u\rangle} && ;\text{ by } (1.22)\\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} \\ &= \langle u, v\rangle + \langle u, w\rangle && ;\text{ by } (1.24)\end{aligned}$$

Corollary 1.4.5

$$\langle u, \alpha v\rangle = \overline{\alpha}\,\langle u, v\rangle \tag{1.31}$$

Proof.

$$\begin{aligned}\langle u, \alpha v\rangle &= \overline{\langle \alpha v, u\rangle} && ;\text{ by } (1.24)\\ &= \overline{\alpha\,\langle v, u\rangle} && ;\text{ by } (1.23)\\ &= \overline{\alpha}\,\overline{\langle v, u\rangle} \\ &= \overline{\alpha}\,\langle u, v\rangle && ;\text{ by } (1.24)\end{aligned}$$

One intuition of the meaning of a real-valued inner product is that the more *similar* two vectors are, the closer the numerical value of inner product is to the product of the *lengths* of the two vectors. More precisely,

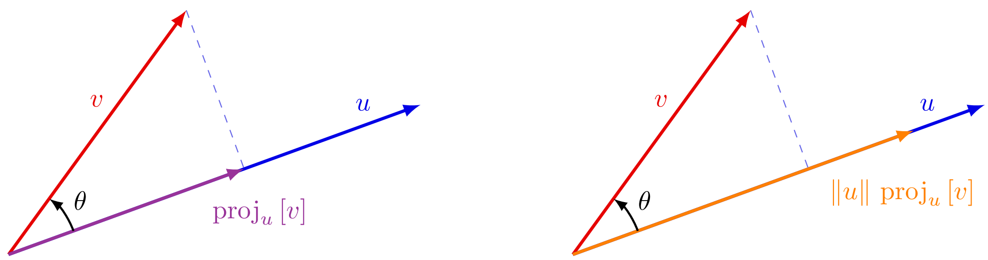


Figure 1.4: (Left)Illustration of $\text{proj}_u[v]$; (Right) $\|u\|\ \text{proj}_u[v]$

if u and w are vectors with lengths $\|u\|$ and $\|w\|$ respectively, then $\langle u, w\rangle$ is a value between $-\|u\|\|w\|$ and $\|u\|\|w\|$, such that if u and w point in the same direction then $\langle u, w\rangle = \|u\|\|w\|$.

Since

$$-\|u\|\|w\| \leq \langle u, w\rangle \leq \|u\|\|w\|,$$

which is proven as Theorem 1.10.2, then there necessarily exists an angle θ such that

$$\langle u, w\rangle = \|u\|\|w\| \cos\theta .$$

Having a concept of *angle between vectors*, we can extend our intuition to say that two vectors are *close* if their lengths are *close* to each other, and the angle between them is close to 0.

However, this intuition supposes that we have a way to measure length. Typically, in a vector space, we define length in terms of the inner product, so the intuition is meta-circular.

For $x, y \in \mathbb{C}^n$, the inner product is typically defined as:

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i},$$

where $\overline{y_i}$ denotes the complex conjugate of y_i. However, if $x, y \in \mathbb{R}^n$, this reduces to

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i .$$

1.5 Motivation for Inner Product

Clearly the inner product is a generalization of the 2-dimensional or 3-dimensional dot product.

If we define $\text{proj}_u[v]$, the projection of vector v onto vector u as shown in Figure 1.4 (left), then we can think of the inner product (in the case of vectors in $\mathbb{R}^n$) as being the positive length of this projection vector scaled by the length of vector u if the angle is acute, $0 \leq \theta \leq \frac{\pi}{2}$, as shown in Figure 1.4 (right), and the negative of this length if the angle is obtuse, $\frac{\pi}{2} \leq \theta \leq \pi$ (Figure 1.5).

$$\langle u, v\rangle = \begin{cases} \|u\|\|\text{proj}_u[v]\| & \text{;if angle acute} \\ -\|u\|\|\text{proj}_u[v]\| & \text{; if angle obtuse} \end{cases}$$

We can use Figure 1.4 (left) and a bit of trigonometry to derive the length of $\text{proj}_u[v]$.

$$\cos\theta = \frac{\|\text{proj}_u[v]\|}{\|v\|}$$
$$\|\text{proj}_u[v]\| = \|v\| \cos\theta$$

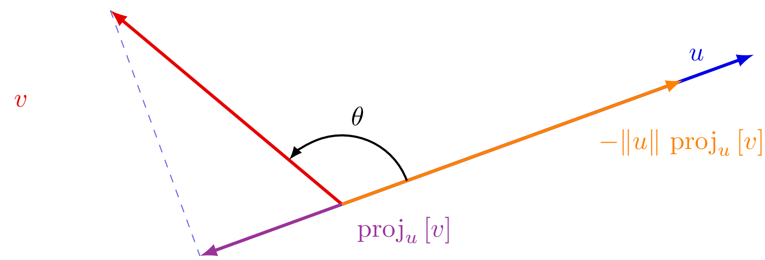


Figure 1.5: Illustration of $\mathrm{proj}_u[v]$ when $\cos\theta < 0$

As shown in Figure 1.4 (right), if we extend $\mathrm{proj}_u[v]$ by the length of u, we have the vector $\|u\|\ \mathrm{proj}_u[v]$. Now when we compute the length of $\|u\|\ \mathrm{proj}_u[v]$. We call this length $|\langle u, v\rangle|$.

$$\begin{aligned}\left\|\overbrace{\|u\|\ \mathrm{proj}_u[v]}^{\text{scaled projection}}\right\| &= \|u\|\ \|\mathrm{proj}_u[v]\| \\ &= \|u\|\ \|v\| |\cos\theta| \\ &= |\langle u, v\rangle|\end{aligned}$$

If $\cos\theta \geq 0$, *i.e.*, if $0° \leq \theta \leq 90°$, then we have $|\cos\theta| = \cos\theta$ and

$$\langle u, v\rangle = \|u\|\ \|v\|\ \cos\theta \geq 0\,.$$

However, if $\cos\theta \leq 0$ then we have $|\cos\theta| = -\cos\theta$ and

$$\langle u, v\rangle = \|u\|\ \|v\|\ \cos\theta \leq 0\,.$$

So in either case

$$\langle u, v\rangle = \|u\|\ \|v\|\ \cos\theta\,.$$

The case where $\cos\theta < 0$ is illustrated in Figure 1.5. We see that $\mathrm{proj}_u[v]$ points in the opposite direction of u. So $-\mathrm{proj}_u[v]$ is parallel to u.

Finally, if $\theta = 90°$ then u and v are perpendicular, and $\cos\theta = 0$ so we see that the inner product of orthogonal vectors is 0.

$$\begin{aligned}\langle u, v\rangle &= \|u\|\ \|v\|\ \cos\theta \\ &= \|u\|\ \|v\| \times 0 \\ &= 0\end{aligned}$$

The equality $\langle u, v\rangle = \|u\|\ \|v\|\ \cos\theta$ presupposes that $\langle u, v\rangle \in \mathbb{R}$, in other words that the field underlying the vector space is $\mathbb{R}$. However, in the case that the underlying field is $\mathbb{C}$, we have to slightly adjust the equation because $\langle u, v\rangle \in \mathbb{C}$. Thus, for a complex vector space,

$$\Re\{\langle u, v\rangle\} = \|u\|\ \|v\|\ \cos\theta\,,$$

where $\Re\{z\}$ denotes the real part of the complex number z.

1.6 Inner Product for Signals

We need an inner product for signals. To extend the inner product to the vector space of functions, we would like to replace $\sum$ by $\int$, to achieve

$$\langle x, y\rangle = \int_{\mathbb{R}} x(t)\,\overline{y(t)}\;\mathrm{d}t\,. \tag{1.32}$$

However, we have to take a few precautions. The definition of inner product does not imply uniqueness. In fact there are multiple ways of defining an inner product for a given vector space. We will define two different inner products for the vector space of signals in Section 1.9 and another one in Section 1.12.

Definition 1.6.1 (*integrable on an interval*)

Let $F \in \{\mathbb{R}, \mathbb{C}\}$, and let $I \subseteq \mathbb{R}$ be a closed interval. We say a function, $f : I \to F$, is *integrable on* I, if

$$\int_I |f(t)|\,\mathrm{d}t < +\infty\,.$$

Likewise, if $p \in \mathbb{N}$, we say f is *p-integrable on* I, if

$$\int_I |f(t)|^p\,\mathrm{d}t < +\infty\,.$$

We may also define *integrable on a non-necessarily closed interval*, simply by redefining the function at the end-points of the interval. Two functions which differ in value only at a finite number of points, have the same integral.

Note that even when $F = \mathbb{C}$, we are still considering functions whose domain is $\mathbb{R}$ or an interval in $\mathbb{R}$. In this course we do not consider signals whose domain is multi-dimensional, $\mathbb{C}$ or otherwise.

Definition 1.6.2 ($\mathcal{L}^p$)

$\mathcal{L}^p(I)$ denotes the set of p-integrable functions on the interval I.

Proposition 1.6.3

If $p \in \mathbb{N}, p > 0$, and I is an interval in $\mathbb{R}$, then $\mathcal{L}^p(I)$ is a vector space.

In particular $\mathcal{L}^2(I)$ is the vector space of square integrable functions over interval I:

$$\int_I |x(t)|^2 < +\infty\,.$$

$\mathcal{L}^1(I)$ is a vector space of *absolutely* integrable functions over inteval I.

1.7 Relation of $\mathcal{L}^1$ and $\mathcal{L}^2$

One might wonder whether $\mathcal{L}^1$ and $\mathcal{L}^2$ are the same or whether one is a subset of the other. To answer these questions it suffices to look at two specially chosen functions, (1.33) and (1.34). Thanks to Jonas Meyer for these example functions.[1] As is seen in Theorem 1.7.1 since neither $\mathcal{L}^1$ nor $\mathcal{L}^2$ is a subset of the other, they are certainly not equal.

Theorem 1.7.1

$\mathcal{L}^1 \not\subset \mathcal{L}^2$ and $\mathcal{L}^2 \not\subset \mathcal{L}^1$.

[1] Jonas Meyer, `https://math.stackexchange.com/users/1424/jonas-meyer`. See the discussion on Stack Exchange: `https://math.stackexchange.com/a/18399`.

Proof.

$$f(t) = \begin{cases} \frac{1}{t} & \text{if } t \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{1.33}$$

$$g(t) = \begin{cases} \frac{1}{\sqrt{t}} & \text{if } 0 < t \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{1.34}$$

Goal 1: $\mathcal{L}^2(\mathbb{R}) \not\subset \mathcal{L}^1(\mathbb{R})$.

$$\begin{aligned} \int_{\mathbb{R}} |f(t)| \,\mathrm{d}t &= \int_1^\infty \left|\frac{1}{t}\right| \mathrm{d}t = \int_1^\infty \frac{1}{t} \,\mathrm{d}t = \ln(t)|_1^\infty = \infty \\ &\implies f \notin \mathcal{L}^1(\mathbb{R}) \end{aligned}$$

$$\begin{aligned} \int_{\mathbb{R}} |f(t)|^2 \,\mathrm{d}t &= \int_1^\infty \left|\left(\frac{1}{t}\right)^2\right| \mathrm{d}t = \int_1^\infty \frac{1}{t^2} \,\mathrm{d}t = -\frac{1}{t}\Big|_1^\infty = 0 - -1 < \infty \\ &\implies f \in \mathcal{L}^2(\mathbb{R}) \end{aligned}$$

We see that $f \in \mathcal{L}^2(\mathbb{R}) \setminus \mathcal{L}^1(\mathbb{R})$, so $\mathcal{L}^2(\mathbb{R}) \not\subset \mathcal{L}^1(\mathbb{R})$.

Goal 2: $\mathcal{L}^1(\mathbb{R}) \not\subset \mathcal{L}^2(\mathbb{R})$.

$$\begin{aligned} \int_{\mathbb{R}} |g(t)| \,\mathrm{d}t &= \int_0^1 \left|\frac{1}{\sqrt{t}}\right| \mathrm{d}t = \int_0^1 \frac{2}{\sqrt{t}} \,\mathrm{d}t = 2\sqrt{t}\Big|_0^1 = 2 - 0 < \infty \\ &\implies g \in \mathcal{L}^1(\mathbb{R}) \\ \int_{\mathbb{R}} |g(t)|^2 \,\mathrm{d}t &= \int_0^1 \left|\frac{1}{\sqrt{t}}\right|^2 \mathrm{d}t = \int_0^1 \frac{1}{t} \,\mathrm{d}t = \ln t \,|_0^1 = 0 - -\infty = \infty \\ &\implies g \notin \mathcal{L}^2(\mathbb{R}) \end{aligned}$$

We see that $g \in \mathcal{L}^1(\mathbb{R}) \setminus \mathcal{L}^2(\mathbb{R})$, so $\mathcal{L}^1(\mathbb{R}) \not\subset \mathcal{L}^2(\mathbb{R})$.

If we consider only bounded domains $I = [a, b] \subset \mathbb{R}$, then $\mathcal{L}^2(I) \subset \mathcal{L}^1(I)$. If we consider only the set of bounded functions on $\mathbb{R}$, call it $\mathfrak{B}$. In this case $\mathcal{L}^1(\mathbb{R}) \cap \mathfrak{B} \subset \mathcal{L}^2(\mathbb{R}) \cap \mathfrak{B}$.

1.8 Concrete Example: Radar

As shown in Figure 1.6, we send a wave (reference signal), x_{ref}. The signal propagates toward the target, reflects on the target, and the echo returns to the emitter/receptor which registers x'_{echo}. Both signals are illustrated in Figure 1.7. We must measure the time delay of x'_{echo} with regard to x_{ref}, and from that time-delay we can calculate the distance to the target.

How can we most reliably compute the time-delay of the echo?

The idea is that if we could measure the similarity between the signals x_{ref} and x'_{echo}, then it would suffice to translate x'_{echo} and superimpose it over x_{ref} at a position which maximizes the similarity between the two signals. The optimal translation amount would give us the delay we seek.

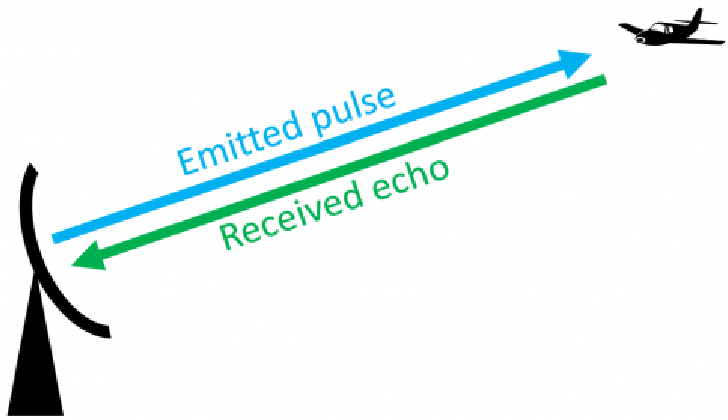
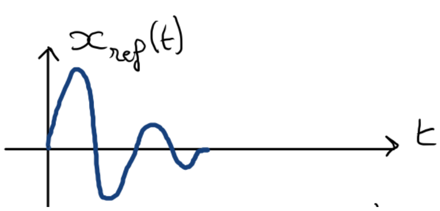
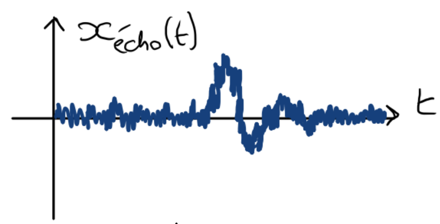


Figure 1.6: Emitted and received signal from radar



Reference signal, x_{ref}, sent by radar device



Echo signal, x'_{echo}, returning to the radar sensor.

Figure 1.7: Radar signal

1.9 Finite Energy Signals

Definition 1.9.1 (*energy of a signal*)

In signal processing we call the quantity

$$E_x = \int_I |x(t)|^2 dt$$

the *energy of signal* x.

According to Definitions 1.6.2 and 1.9.1, we may say that $\mathcal{L}^2(I)$ is the set of finite energy signals on interval I.

We will work with $\mathcal{L}^2(\mathbb{R})$, the set of finite energy signals on $\mathbb{R}$.

Definition 1.9.2 (*inner product in* $\mathcal{L}^2(\mathbb{R})$)

In $\mathcal{L}^2(\mathbb{R})$, we may define the *inner product*,

$$\langle\ ,\ \rangle : \mathcal{L}^2(\mathbb{R}) \times \mathcal{L}^2(\mathbb{R}) \to \mathbb{R}\,,$$

as follows:

$$\langle x, y\rangle = \int_{\mathbb{R}} x(t)\,\overline{y(t)}\;dt\,. \tag{1.35}$$

Theorem 1.9.3

Definition 1.9.2 defines an inner product which satisfies the axioms of Definition 1.4.1.

Proof. We must show the inner product from Definition 1.9.2 fulfills all the axioms of Definition 1.4.1.

In each case we will assume that the integrals exist.

- $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (1.22)

$$\begin{aligned}\langle u+v, w\rangle &= \int_{\mathbb{R}} (u(t)+v(t))\,\overline{w(t)}\;dt\\ &= \int_{\mathbb{R}} \Big(u(t)\,\overline{w(t)} + v(t)\,\overline{w(t)}\Big)\,dt\\ &= \int_{\mathbb{R}} u(t)\,\overline{w(t)}\;dt + \int_{\mathbb{R}} v(t)\,\overline{w(t)}\;dt\\ &= \langle u, w\rangle + \langle v, w\rangle\end{aligned}$$

- $\langle \alpha v, w\rangle = \alpha\,\langle v, w\rangle$ (1.23)

$$\begin{aligned}\langle \alpha v, w\rangle &= \int_{\mathbb{R}} \alpha v(t)\,\overline{w(t)}\;dt\\ &= \alpha \int_{\mathbb{R}} v(t)\,\overline{w(t)}\;dt\\ &= \alpha\,\langle v, w\rangle\end{aligned}$$

- $\langle v, w\rangle = \overline{\langle w, v\rangle}$ (1.24)

  We will show that $\langle v, w\rangle - \overline{\langle w, v\rangle}$.

$$
\begin{aligned}
\langle v, w\rangle &= \int_{\mathbb{R}} v(t)\,\overline{w(t)}\,dt \\
&= \int_{\mathbb{R}} (\Re\{v(t)\} + i\Im\{v(t)\})\,\overline{(\Re\{w(t)\} + i\Im\{w(t)\})}\,dt \\
&= \int_{\mathbb{R}} (\Re\{v(t)\} + i\Im\{v(t)\})(\Re\{w(t)\} - i\Im\{w(t)\})\,dt \\
&= \int_{\mathbb{R}} (\Re\{v(t)\}\,\Re\{w(t)\} - \Im\{v(t)\}\,\Im\{w(t)\} \\
&= \qquad +i\Im\{v(t)\}\,\Re\{w(t)\} - i\Re\{v(t)\}\,\Im\{w(t)\})\,dt
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
\langle w, v\rangle &= \int_{\mathbb{R}} w(t)\,\overline{v(t)}\,dt \\
&= \int_{\mathbb{R}} (\Re\{w(t)\}\,\Re\{v(t)\} - \Im\{w(t)\}\,\Im\{v(t)\} \\
&= \qquad +i\Im\{w(t)\}\,\Re\{v(t)\} - i\Re\{w(t)\}\,\Im\{v(t)\})\,dt \\
\overline{\langle w, v\rangle} &= \int_{\mathbb{R}} (\Re\{w(t)\}\,\Re\{v(t)\} - \Im\{w(t)\}\,\Im\{v(t)\} \\
&= \qquad -i\Im\{w(t)\}\,\Re\{v(t)\} + i\Re\{w(t)\}\,\Im\{v(t)\})\,dt
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\langle v, w\rangle - \overline{\langle w, v\rangle} &= i\int_{\mathbb{R}} (\Im\{v(t)\}\,\Re\{w(t)\} - \Re\{v(t)\}\,\Im\{w(t)\} \\
&\qquad + \Im\{w(t)\}\,\Re\{v(t)\} - \Re\{w(t)\}\,\Im\{v(t)\})\,dt \\
&= \int_{\mathbb{R}} 0\,dt = 0
\end{aligned}
$$

- $\langle v, v\rangle \geq 0$ (1.25)

$$
\begin{aligned}
\langle v, v\rangle &= \int_{\mathbb{R}} v(t)\,\overline{v(t)}\,dt \\
&= \int_{\mathbb{R}} |v(t)|^2\,dt \\
&\geq 0 && \text{because } |v(t)|^2 \geq 0 \text{ for all } t
\end{aligned}
$$

- $v(t) = 0 \implies \langle v, v\rangle = 0$ (1.26)
  Let $v(t) = 0$

$$
\begin{aligned}
\langle v, v\rangle &= \int_{\mathbb{R}} v(t)\,\overline{v(t)}\,dt \\
&= \int_{\mathbb{R}} (0)(\overline{0})\,dt \\
&= \int_{\mathbb{R}} 0\,dt = 0
\end{aligned}
$$

$x, y \in \mathbb{R}^n$	$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$
$x(t), y(t) \in \mathbb{R}$	$\langle x, y\rangle = \int_{\mathbb{R}} x(t)y(t)\, dt$
$x, y \in \mathbb{C}^n$	$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$
$x(t), y(t) \in \mathbb{C}$	$\langle x, y\rangle = \int_{\mathbb{R}} x(t)\overline{y(t)}\, dt$

Figure 1.8: Inner product definitions

- $\langle v, v\rangle = 0 \implies v(t) = 0$ (1.26)

$$\begin{aligned} 0 &= \langle v, v\rangle \\ &= \int_{\mathbb{R}} v(t)\overline{v(t)}\, dt \\ &= \int_{\mathbb{R}} |v(t)|^2\, dt \end{aligned}$$

  Unfortunately,

$$\int_{\mathbb{R}} |v(t)|^2\, dt = 0 \;\not\Longrightarrow\; v(t) = 0\,.$$

  However, since we are only concerned with functions which are PWC, Definition 1.0.1, then we shall consider two signals (PWC functions) to be in the same *equivalence class*, whenever they differ only at their points of discontinuity, thus which differ only at countably many points. This means $v(t) = 0$ except at possibly countably many point—not effecting the value of the integral.

If we only consider real valued functions, $t \in \mathbb{R} \implies x(t) \in \mathbb{R}$, then the inner product simplifies to:

$$\langle x, y\rangle = \int_{\mathbb{R}} x(t)y(t)\, dt\,.$$

If we consider the inner product of a vector (signal) with itself, then we observe that the value is real ($\in \mathbb{R}$) and non-negative, even if the signal is complex valued.

$$\begin{aligned} \langle x, x\rangle &= \int_{\mathbb{R}} x(t)\overline{x(t)}\, dt \\ &= \int_{\mathbb{R}} |x(t)|^2\, dt \qquad (1.36) \\ &= E_x \end{aligned}$$

Figure 1.8 summarizes the resemblance between the various definitions of inner product.

Remark 1.9.4
Even if it is an abuse of notation, we sometimes (often) use $\langle x(t), y(t)\rangle$, to mean $\langle x, y\rangle$ in the case that x and y are signals. *I.e.*, sometimes we use the notation, $x(t)$, to mean a scalar value such as $x(t) \in \mathbb{R}$, and sometimes we use the notation, $x(t)$, to mean a function such as $x(t) \in \mathcal{L}^2(\mathbb{R})$.

1.10 Norms and Inner Products

As we mentioned earlier, once an inner product exists for a vector space, we can derive a norm from it.

Theorem 1.10.1

If a vector space has an inner product, $\langle\ ,\ \rangle$, then (1.37) defines a norm.

$$\|\cdot\| : x \mapsto \|x\| = \sqrt{\langle x, x\rangle}\,. \tag{1.37}$$

Proof. Exercise for student

With Definition 1.3.2, we can verify that Equation (1.21) indeed defines a distance function.

Theorem 1.10.2 (*Cauchy-Schwarz Inequality*)

If $\|\cdot\|$ is induced from the inner product $\langle\cdot,\cdot\rangle$, then

$$|\langle u, v\rangle| \le \|u\|\|v\|\,.$$

This proof is inspired by stackexchange post `https://math.stackexchange.com/q/478404`, thanks to `walcher`, `https://math.stackexchange.com/users/89844/walcher`.

Proof. Let $a = u$ and $b = \frac{\langle u,v\rangle}{\|v\|^2} v$,

$$\begin{aligned}
0 &\le \|a-b\|^2 && \text{by (1.18)}\\
&= \langle a-b, a-b\rangle \\
&= \langle a, a-b\rangle - \langle b, a-b\rangle && \text{by (1.22)}\\
&= \langle a,a\rangle - \langle a,b\rangle - \langle b,a\rangle + \langle b,b\rangle && \text{by (1.30)}\\
&= \langle a,a\rangle - \langle a,b\rangle - \overline{\langle a,b\rangle} + \langle b,b\rangle && \text{by (1.24)}\\
&= \|a\|^2 - 2\mathfrak{Re}\,\langle a,b\rangle + \|b\|^2 && \text{by Theorem 1.10.1}\\
&= \|u\|^2 - 2\mathfrak{Re}\left\langle u, \frac{\langle u,v\rangle}{\|v\|^2} v\right\rangle + \left\|\frac{\langle u,v\rangle}{\|v\|^2} v\right\|^2 \\
&= \|u\|^2 - \frac{2}{\|v\|^2}\mathfrak{Re}\,\langle u, \langle u,v\rangle v\rangle + |\langle u,v\rangle|^2 \frac{\|v\|^2}{\|v\|^4} && \text{by (1.24)}\\
&= \|u\|^2 - \frac{2}{\|v\|^2}\mathfrak{Re}\Big(\overline{\langle u,v\rangle}\,\langle u,v\rangle\Big) + \frac{|\langle u,v\rangle|^2}{\|v\|^2} && \text{by (1.24)}\\
&= \|u\|^2 - \frac{2}{\|v\|^2}\mathfrak{Re}\Big(|\langle u,v\rangle|^2\Big) + \frac{|\langle u,v\rangle|^2}{\|v\|^2} && \text{by (1.24)}\\
&= \|u\|^2 - \frac{2|\langle u,v\rangle|^2}{\|v\|^2} + \frac{|\langle u,v\rangle|^2}{\|v\|^2} && |\langle u,v\rangle|^2 \ \text{ is real}\\
&= \|u\|^2 - \frac{|\langle u,v\rangle|^2}{\|v\|^2}
\end{aligned}$$

$$\begin{aligned}
0 \le \|u\|^2\|v\|^2 - |\langle u,v\rangle|^2 &\implies |\langle u,v\rangle|^2 \le \|u\|^2\|v\|^2 \\
&\implies |\langle u,v\rangle| \le \|u\|\|v\|
\end{aligned}$$

So which inner product induces a norm on the vector space of finite energy signals?

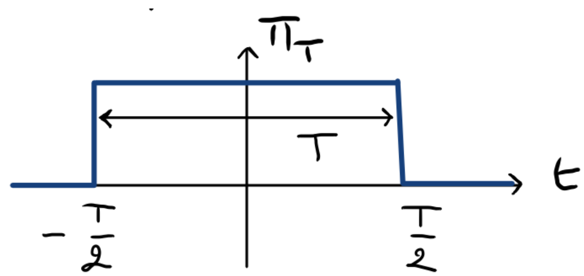


Figure 1.9: Window (pulse) of size T

Let $x \in \mathcal{L}^2(\mathbb{R})$,

$$\begin{aligned}
\|x\| &= \sqrt{\langle x, x\rangle} \\
\|x\|^2 &= \langle x, x\rangle = \int_{\mathbb{R}} x(t)\,\overline{x(t)}\,dt \\
&= \int_{\mathbb{R}} |x(t)|^2\,dt && ;z \in \mathbb{C} \implies z\,\overline{z} = |z|^2 \\
&= E_x < +\infty && ;\texttt{because } x \in \mathcal{L}^2(\mathbb{R})
\end{aligned}$$

Thus we have

$$E_x = \|x\|^2,$$

which means the energy of a signal is the square of the norm of the signal. If x is a finite-energy signal, then its norm is also finite.

Definition 1.10.3 (*window, gate, pulse, rectangle, boxcar*)

Let $T > 0$, then $\Pi_T : \mathbb{R} \to \mathbb{R}$ is defined as follows:

$$\Pi_T(t) = \begin{cases} 1 & ;\text{if } -\frac{T}{2} \leq x \leq \frac{T}{2} \\ 0 & ;\text{otherwise} \end{cases},$$

as shown in Figure 1.9.

Example 1.10.4

Given $A > 0$, we will compute the energy of x defined as follows:

$$x(t) = A\Pi_T(t)$$

$$
\begin{aligned}
E_x &= \|x\|^2 \\
&= \langle x, x \rangle \\
&= \int_R |x(t)|^2 dt \\
&= \int_{-\infty}^{-\frac{T}{2}} |x(t)|^2 dt + \int_{-T/2}^{T/2} |x(t)|^2 dt + \int_{-T/2}^{T/2} |x(t)|^2 dt \\
&= \int_{-\infty}^{-\frac{T}{2}} 0 dt + \int_{-T/2}^{T/2} A^2 dt + \int_{\frac{T}{2}}^{\infty} 0 dt \\
&= 0 + A^2 t \Big|_{-\frac{T}{2}}^{\frac{T}{2}} + 0 \\
&= A^2 T \quad < +\infty
\end{aligned}
$$

Therefore, $x = A\Pi_T$ is a finite energy signal, $x \in \mathcal{L}^2(\mathbb{R})$.

Definition 1.10.5 (*support*)

Let $f : \mathbb{R} \to \mathbb{C}$ be a signal. The *support* of f is the set:

$$\texttt{supp}(f) = \{t \in \mathbb{R} \mid f(t) \neq 0\}$$

Definition 1.10.6 (*bounded support*)

Let $f : \mathbb{R} \to \mathbb{C}$. f is said to have *bounded support* if $\texttt{supp}(f)$ is a bounded set; *i.e.*, if $\exists\ M > 0$ such that $|t| < M\ \forall t \in \texttt{supp}(f)$.

Theorem 1.10.7

Every signal with bounded support is a finite energy signal.

Proof. Let $f : \mathbb{R} \to \mathbb{C}$ be a signal with bounded support. Since f has bounded support, let L be a bound of the support. Since f is bounded (by Definition 1.0.2), choose M such that $|f(t)| < M\ \forall t \in \mathbb{R}$.

$$
\begin{aligned}
E_x &= \int_{\mathbb{R}} |f(t)|^2 dt \\
&= \underbrace{\int_{-\infty}^{-L} |f(t)|^2 dt}_{=0} + \int_{-L}^{L} |f(t)|^2 dt + \underbrace{\int_{L}^{\infty} |f(t)|^2 dt}_{=0} \\
&= \int_{-L}^{L} |f(t)|^2 dt \\
&\leq \int_{-L}^{L} M^2 dt \\
&= 2LM^2 < \infty
\end{aligned}
$$

1.11 Infinite Energy Signals

So what is an example of a signal without bounded support?

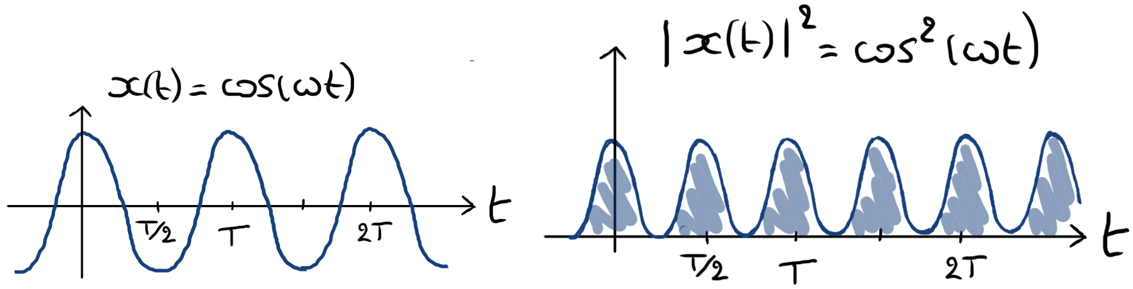


Figure 1.10: The cos and $\cos^2$ signals

Example 1.11.1 (***Function without bounded support***)

Let x be the cos signal as shown in Figure 1.10

$$x(t) = \cos(\omega t)\,,$$

where

$$\omega = 2\pi f = \frac{2\pi}{T}\,,$$

with f denoting the frequency, ω denoting the pulsation, and T denoting the period.

Theorem 1.11.2

$\cos(\omega t)$ is an infinite energy signal.

The following proof is perhaps an overly complicated way of proving the obvious—that the area under the curve of $cos^2(\omega t)$ in Figure 1.10 is infinite.

Proof. When we look at Figure 1.10 we can note by inspection that $\cos^2$ is integrable on the closed interval $[0, \frac{T}{2}]$. So denote

$$M = \int_0^{\frac{T}{2}} \cos^2(\omega t) dt\,.$$

We note that $0 < M < \infty$.

$$
\begin{aligned}
E_x &= \int_{\mathbb{R}} |x(t)|^2 dt \\
&= \int_{-\infty}^{\infty} \cos^2(\omega t) dt \\
&= \lim_{n\to\infty} \int_{-nT}^{nT} \cos^2(\omega t) dt \\
&= 2 \times \lim_{n\to\infty} \int_{0}^{nT} \cos^2(\omega t) dt \\
&= 2 \times \lim_{n\to\infty} \sum_{i=0}^{n-1} \int_{iT}^{(i+1)T} \cos^2(\omega t) dt \\
&= 2 \times \lim_{n\to\infty} \sum_{i=0}^{n-1} \int_{0}^{T} \cos^2(\omega t) dt && \text{; because } \cos^2 \text{ is } \frac{T}{2}\text{-periodic} \\
&= 2 \times \lim_{n\to\infty} \sum_{i=0}^{n-1} 2M \\
&= 2 \times \lim_{n\to\infty} 2Mn \\
&= +\infty
\end{aligned}
$$

The idea in Example 1.11.1 is that a sinusoidal signal is infinite power. The same idea applies to any combination of sin/cos signals, no matter their amplitudes or frequencies (periods/pulsations).

Theorem 1.11.3
If $0 < T < \infty$, then every T-periodic signal is non-energy-finite.

Proof. Exercise for student. Follow the pattern of Theorem 1.11.2.

The fact that periodic signals are not finite energy signals somewhat troublesome, because periodic signals are very important to the study of signal processing.

1.12 Average power signals

The vector space of finite energy signals, $\mathcal{L}^2(\mathbb{R})$, is insufficiently exhaustive to make possible a general description of signals we regularly encounter in signal processing. Therefore, we will need something more general.

Definition 1.12.1 (*average power*)
Let $x : \mathbb{R} \to \mathbb{C}$ be a signal. The quantity

$$P_x = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt$$

will be called the *average power* of signal x over $\mathbb{R}$.

We recall from our physics courses that normally a power (in Watts) is expressed as energy (in Joules) per time (in seconds):

$$1W = 1J/s.$$

Here we find the same idea. $\int\limits_{-T/2}^{T/2} |x(t)|^2 dt$ corresponds to the energy of x over one period, $[-\frac{T}{2}, \frac{T}{2}]$, as opposed to over all of $\mathbb{R}$. So $\frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt$ is analogous to power.

Definition 1.12.2 (*finite average power signal*)

Similar to finite-energy signals, we say that a signal x is a *finite average power* signal, with average power, P_x, if

$$P_x = \lim_{T \to +\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt < +\infty$$

Definition 1.12.3 (*inner product in* $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$)

In $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, we may define the *inner product* as follows:

$$\langle\ ,\ \rangle : (x, y) \mapsto \langle x, y\rangle = \lim_{T \to +\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)\, \overline{y(t)}\, dt\,. \tag{1.38}$$

Theorem 1.12.4

Equation (1.38) in Definition 1.12.3 defines an inner product.

Proof. Exercise for student

Definitions 1.12.1 and 1.12.2 deserve several remarks.

- As opposed to $\mathcal{L}^2(\mathbb{R})$, there is no standard notation for the space of finite average power signals. We will denote this space $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$.
- Like $\mathcal{L}^2(\mathbb{R})$, $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ is a vector space; for

$$x, y \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R}), \lambda \in \mathbb{C} \implies x + \lambda y \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})\,.$$

- Derived from the inner product on $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, we recover the equality

$$\langle x, x\rangle = \|x\|^2 = P_x\,.$$

- The average power of a T-periodic signal is given by

$$P_x = \frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt\,.$$

  See Theorem 1.13.3.

1.13 Power of Periodic Signals

We would like to present Theorem 1.13.3. To prove Theorem 1.13.3 we need Lemma 1.13.2, which we prove using Lemma 1.13.1.

Lemma 1.13.1

If signal x is periodic with period T, then

$$\int_a^b x(t)dt = \int_{a+nT}^{b+nT} x(t)dt\,.$$

Proof. Note that x periodic means that $\forall n \in \mathbb{Z}, x(t) = x(t+nT)$. Consequently

$$\begin{aligned} I &= \int_a^b x(t)dt \\ &= \int_a^b x(t+nT)dt \end{aligned}$$

Let $u = t + nT$, so $dt = du$; $t = a \implies u = a + nT$; $t = b \implies u = b + nT$.

$$I = \int_{a+nT}^{b+nT} x(u)du = \int_{a+nT}^{b+nT} x(t)dt$$

Lemma 1.13.2

If x is a periodic signal with period $T > 0$, then

$$\int_0^T x(t)dt = \int_\alpha^{\alpha+T} x(t)dt\,.$$

I.e., we can integrate over any period of length T starting at any value $t = \alpha$.

Proof. If $0 \leq \alpha \leq T$, then

$$\begin{aligned} \int_0^T x(t)dt &= \int_0^\alpha x(t)dt + \int_\alpha^T x(t)dt \\ &= \int_T^{T+\alpha} x(t)dt + \int_\alpha^T x(t)dt && \text{by Lemma 1.13.1} \\ &= \int_\alpha^{\alpha+T} x(t)dt && (1.39) \end{aligned}$$

If $\alpha < 0$ or $\alpha \geq T$, then choose m such that $0 \leq \alpha - mT \leq T$.

$$\int_0^T x(t)dt = \int_{\alpha-mT}^{\alpha-mT+T} x(t)dt \qquad \text{by (1.39)}$$
$$= \int_{\alpha}^{\alpha+T} x(t)dt$$

Theorem 1.13.3 shows that to compute the power of a periodic signal, we need only integrate over one period.

Theorem 1.13.3
If x is a periodic signal with period $T > 0$, then

$$P_x = \frac{1}{T}\int_{-T/2}^{T/2} |x(t)|^2 dt\,.$$

We demonstrate two proofs of Theorem 1.13.3.

First Proof. By Definition 1.12.1

$$P_x = \lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} |x(t)|^2 dt$$

To verify that the average power does not depend on the number of periods, we show (by induction) that it is the same for any finite number of periods.

First, we compute the power for two periods.

$$\frac{1}{2T}\int_{-T}^{T} |x(t)|^2 dt = \frac{1}{2T}\left(\int_{-T}^{0} |x(t)|^2 dt + \int_{0}^{T} |x(t)|^2 dt\right)$$
$$= \frac{1}{2T}\underbrace{\left(\int_{-T/2}^{T/2} |x(t)|^2 dt + \int_{-T/2}^{T/2} |x(t)|^2 dt\right)}_{\text{by Lemma 1.13.2}}$$
$$= \frac{1}{2T}\left(2\int_{-T/2}^{T/2} |x(t)|^2 dt\right)$$
$$= \frac{1}{T}\int_{-T/2}^{T/2} |x(t)|^2 dt$$

For the inductive step we suppose that for $2 \le k \le n$ the average power over k periods equals P_x, and we compute the average power over $n+1$ periods.

$$\begin{aligned} P_{n+1} &= \frac{1}{(n+1)T}\int_0^{(n+1)T} |x(t)|^2 dt \\ &= \frac{1}{(n+1)T}\left(\underbrace{\int_0^{(n-1)T} |x(t)|^2 dt}_{\text{energy over } n-1 \text{ periods}} + \underbrace{\int_{(n-1)T}^{(n+1)T} |x(t)|^2 dt}_{\text{energy over two periods}} \right) \end{aligned}$$

By inductive hypothesis, the average power over $n-1$ periods is P_x, and also the average power over 2 periods is P_x. So the energy over $n-1$ periods is $(n-1)TP_x$, and the energy over 2 periods is $2TP_x$.

$$\begin{aligned} P_{n+1} &= \frac{1}{(n+1)T}((n-1)TP_x + 2TP_x) \\ &= \frac{1}{(n+1)T}(n-1+2)TP_x \\ &= \frac{1}{(n+1)T}(n+1)TP_x \\ &= P_x \end{aligned}$$

We have shown that the average power of x over n periods is P_x for all $n \in \mathbb{N}$. So if we view the sequence, $P_n = \left(\frac{1}{nT} \int_{-\frac{nT}{2}}^{\frac{nT}{2}} |x(t)|^2 dt \right)_{n=1}^{\infty}$, and look at the limit as $n \to \infty$, we see:

$$\begin{aligned} \lim_{n\to\infty} P_n &= \lim_{n\to\infty} \frac{1}{nT} \int_{-\frac{nT}{2}}^{\frac{nT}{2}} |x(t)|^2 dt \\ &= \lim_{n\to\infty} P_x \\ &= P_x \end{aligned}$$

$P_n = \frac{1}{nT} \int_{-\frac{nT}{2}}^{\frac{nT}{2}} |x(t)|^2 dt$ is a constant sequence, thus converges. If $P(\tau) = \frac{1}{\tau T} \int_{-\frac{\tau T}{2}}^{\frac{\tau T}{2}} |x(t)|^2 dt$, then $P_n = P(\tau)$ whenever $n = \tau$. Therefore if, $\lim_{\tau\to\infty} P(\tau)$ exists,

$$\lim_{\tau\to\infty} P(\tau) = \lim_{n\to\infty} P_n = P_x\,.$$

We now provide a second proof of Theorem 1.13.3.

Second Proof. Let $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, and $T \in \mathbb{R}$ with $T > 0$, such that x is T-periodic. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(y) = \begin{cases} 0 & ;\text{ if } y = 0 \\ \frac{1}{y} \int_{-y/2}^{y/2} |x(t)|^2 dt & ;\text{ otherwise} \end{cases}$$

Now,

$$x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R}) \implies \lim_{y\to\infty} f(y) = P_x \in \mathbb{R}\,.$$

Now define $U_n = nT$, for $n \in \mathbb{N}$. And we have

$$\begin{aligned}
\lim_{y\to+\infty} f(y) &= \lim_{n\to+\infty} f(U_n) \\
&= \lim_{n\to+\infty} \frac{1}{nT} \int\limits_{-nT/2}^{nT/2} |x(t)|^2 dt \\
&= \lim_{n\to+\infty} \frac{n}{nT} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt \\
&= \lim_{n\to+\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt \\
&= P_x
\end{aligned}$$

1.14 Exercises: TD 1

1. What is the link between $\mathcal{L}^2(\mathbb{R})$ and $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$? That is to say: what can be said about the energy of a finite average power signal, and about the average power of a finite energy signal?

2. Compute the average power of
$$x(t) = A\cos(\frac{2\pi}{T}\,t)$$
with $T > 0$ being the period of x. Suggestion, start with $T = 1$ or $T = 2\pi$. Reminders:
$$\cos^2(a) = \frac{1}{2}\left(1 + \cos(2a)\right) \tag{1.40}$$
$$\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b) \tag{1.41}$$

3. Does the previous result change if a phase ϕ is imposed as follows
$$x(t) = A\cos\left(\frac{2\pi}{T}t + \phi\right)?$$

4. What is the average power P_{x+y} of the sum of two signals. *I.e.*, if $x, y \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, what is the average power of $x + y$? Same question for the energy of the sum of two finite energy signals.

1.15 Review

1. What is a signal: a bounded, PWC, $\mathbb{C}$-valued function
2. Vector space, norm, inner product, induced norm, angle.
3. Vector space of signals: linear combinations, inner product, norm, basis.
4. Integrable functions: $\mathcal{L}^p$, $\mathcal{L}^1$, $\mathcal{L}^2$
5. Inner product of $\mathcal{L}^2$: $\langle x, y \rangle = \int\limits_{\mathbb{R}} x(t)\,\overline{y(t)}\ dt$
6. Example of inner product with radar signals.
7. Signal Energy: $E_x = \langle x, x \rangle = \int\limits_{\mathbb{R}} |x(t)|^2\ dt$
8. Infinite energy: cos, every non-zero, periodic function
9. Every signal of bounded support has finite energy: window function, `supp`
10. Average power signals
11. $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, inner product $\langle x, y \rangle = \lim\limits_{T \to +\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)\,\overline{y(t)}\, dt$
12. $P_x = \langle x, x \rangle = \lim\limits_{T \to +\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt$
13. If x is T-periodic: $P_x = \lim\limits_{T \to +\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt = \frac{1}{T} \int\limits_{-T/2}^{T/2} |x(t)|^2 dt$

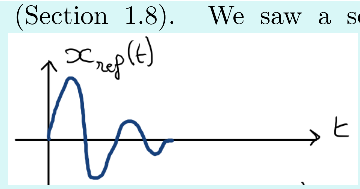
Chapter 2

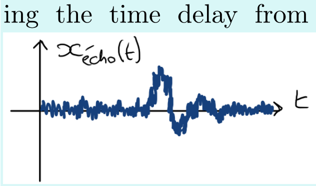
Correlation and Convolution

Chapter Objectives

- Define autocorrelation and cross-correlation for finite-energy and finite-average-power signals, and compute them for concrete signals.
- Apply the key properties of autocorrelation — Hermitian symmetry, maximum at zero, connection to energy/power — to determine the time delay of a noisy echo.
- Define and compute the convolution product of two signals, both by direct integration and by geometric/visual reasoning.
- Relate convolution, cross-correlation, and the inner product to one another, and translate between the three representations.
- Correctly interpret time-shifted signals $x(t-\tau)$ (delay vs. advance) when manipulating correlation and convolution expressions.

Reminder 2.0.1

Recall the example of radar in the previous chapter (Section 1.8). We saw a solution for measuring the time delay from the reference signal, x_{ref}, 

, to its echo x'_{echo}, 

. The solution was to compare $t \mapsto x_{ref}(t)$ against all versions of $t \mapsto x_{echo}(t-\tau)$, *i.e.*, for all values of τ. We defined the time-delay as the value of τ which maximizes the resemblance (the inner product) from $t \mapsto x_{ref}(t)$ to $t \mapsto x_{echo}(t-\tau)$.

Thus mathematically, we would expect to manipulate a function $\tau \mapsto \langle x_{ref}(t), x'_{echo}(t-\tau)\rangle$. We call this function the *cross-correlation* of x_{ref} to x'_{echo}, and will be defined more formally in Section 2.4.

2.1 Remark about notation

We must make a remark about the notation here. The notation,

$$\langle x(t), y(t-\tau) \rangle \ ,$$

is an abuse of notation, because the inner product is an operation between elements of a vector space, signals x and y, rather than the values of the signals at time t, $x(t)$ and $y(t-\tau)$. To be completely consistent, we would need to define a translated signal (a function of t)

$$y_\tau(t) = y(t-\tau)\,.$$

y_τ is a function of t which translates y by τ time units. Now $\tau \mapsto \langle x, y_\tau \rangle$ is indeed a function of τ.

Even though such a notation would be more correct, it is also awkward and rarely used, except in very detailed proofs. Instead we allow ourselves to write $\langle x(t), y(t-\tau)\rangle$, and just remember that this is shorthand for $\tau \mapsto \langle x, y_\tau \rangle$. We keep in mind that the inner product applies to signals living in a vector space of functions (signals), and not on particular real or complex values of those functions at a point in time.

2.2 Autocorrelation

Before studying the cross-correlation function, we simplify the problem a bit and consider the case that $y = x_{ref}$.

Definition 2.2.1 (*autocorrelation in* $\mathcal{L}^2(\mathbb{R})$)

If $x \in \mathcal{L}^2(\mathbb{R})$, we define the *autocorrelation* of x as follows:

$$\Gamma_{xx}(\tau) = \langle x(t), x(t-\tau) \rangle = \int_{\mathbb{R}} x(t)\,\overline{x(t-\tau)}\,\mathrm{d}t$$

Corollary 2.2.2 (*autocorrelation of a real signal*)

If $x(t) \in \mathbb{R}$

$$\Gamma_{xx}(\tau) = \int_{\mathbb{R}} x(t)x(t-\tau)\,\mathrm{d}t \tag{2.1}$$

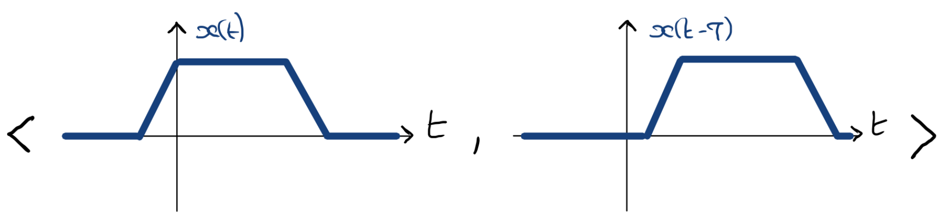
Proof. If $x(t) \in \mathbb{R}$ then $\overline{x(t)} = x(t)$.

Reminder 2.2.3

Recall the discussion from Section 2.1 that $\langle x(t), x(t-\tau)\rangle$ in Definition 2.2.1, really means $\langle x, y_\tau\rangle$, where $y_\tau(t) = x(t-\tau)$.

$\Gamma_{xx}(\tau)$ is the similarity (inner product) between x and x translated by τ time units. This translation is a delay (pushed into the future) if $\tau > 0$, or advanced (pushed into the past) if $\tau < 0$. See Section 2.11.

$\Gamma_{xx}(\tau) =$



Several properties follow from Definition 2.2.1.

Theorem 2.2.4 (*autocorrelation is finite*)

If $x \in \mathcal{L}^2(\mathbb{R})$ then

$$|\Gamma_{xx}(\tau)| < +\infty \quad \forall \tau \in \mathbb{R}\,.$$

The proof should be evident from the properties which follow. We prove Theorem 2.2.4 on page 41.

Corollary 2.2.5 (*autocorrelation at zero equals energy*)

If $x \in \mathcal{L}^2(\mathbb{R})$, then

$$\Gamma_{xx}(0) = E_x\,.$$

Proof.

$$\begin{aligned}\Gamma_{xx}(0) &= \langle x(t), x(t)\rangle \\ &= \int_{\mathbb{R}} |x(t)|^2\,\mathrm{d}t \\ &= \|x\|^2 \\ &= E_x\,.\end{aligned}$$

We know $E_x < +\infty$, because $x \in \mathcal{L}^2(\mathbb{R})$.

Corollary 2.2.6 (***Hermitian symmetry of autocorrelation***)

If $x \in \mathcal{L}^2(\mathbb{R})$, then Γ_{xx} is symmetric Hermitian.

$$\Gamma_{xx}(-\tau) = \overline{\Gamma_{xx}(\tau)}\,.$$

Proof.

$$\Gamma_{xx}(-\tau) = \int_{\mathbb{R}} x(t)\,\overline{x(t+\tau)}\,\mathrm{d}t$$

Let $u = t + \tau$; $t = u - \tau$; $\mathrm{d}u = \mathrm{d}t$.

$$\begin{aligned}\Gamma_{xx}(-\tau) &= \int_{\mathbb{R}} x(u-\tau)\,\overline{x(u)}\,\mathrm{d}u \\ &= \overline{\int_{\mathbb{R}} x(u)\,\overline{x(u-\tau)}\,\mathrm{d}u} \\ &= \overline{\Gamma_{xx}(\tau)}\end{aligned}$$

Corollary 2.2.7 (*autocorrelation of a real signal is even*)

If $x \in \mathcal{L}^2(\mathbb{R})$, and x is real valued, $x(t) \in \mathbb{R} \quad \forall t$, then Γ_{xx} is an even function.

Proof. $x(t)$ is real, so $x(t) = \overline{x(t)}$.

$$\begin{aligned}
\Gamma_{xx}(-\tau) &= \overline{\Gamma_{xx}(\tau)} && \text{; by Corollary 2.2.6} \\
&= \overline{\int_{\mathbb{R}} x(t)\,\overline{x(t-\tau)}\,\mathrm{d}t} && \text{; definition of } \Gamma_{xx} \\
&= \overline{\int_{\mathbb{R}} x(t)x(t-\tau)\,\mathrm{d}t} && \text{; } x(t) \in \mathbb{R} \forall t \in \mathbb{R} \\
&= \int_{\mathbb{R}} x(t)x(t-\tau)\,\mathrm{d}t && \text{; integral of real valued function} \\
&= \Gamma_{xx}(\tau) && \text{; by Corollary 2.2.2}
\end{aligned}$$

Theorem 2.2.8 (*time-shifting preserves energy*)

If $x \in \mathcal{L}^2(\mathbb{R})$, and if the energy of signal x is E_x, then the energy of signal x_τ is also E_x, where

$$x_\tau(t) = x(t-\tau)\,.$$

Proof. From the hypothesis we have:

$$\begin{aligned}
E_x &= \|x\|^2 \\
&= \int_{\mathbb{R}} |x(t)|^2\,\mathrm{d}t\,.
\end{aligned}$$

So we compute the energy of x_τ.

$$\begin{aligned}
E_{x_\tau} &= \|x_\tau\|^2 \\
&= \int_{\mathbb{R}} |x_\tau(t)|^2\,\mathrm{d}t \\
&= \int_{\mathbb{R}} |x(t-\tau)|^2\,\mathrm{d}t
\end{aligned}$$

Let $u = t - \tau$; $\mathrm{d}u = \mathrm{d}t$.

$$\begin{aligned}
E_{x_\tau} &= \int_{\mathbb{R}} |x(u)|^2\,\mathrm{d}u \\
&= E_x
\end{aligned}$$

Theorem 2.2.9 (*autocorrelation is maximal at zero*)

Let $x \in \mathcal{L}^2(\mathbb{R})$, and let $\tau \in \mathbb{R}$.

$$|\Gamma_{xx}(\tau)| \le \Gamma_{xx}(0)\,.$$

For the proof, we recall the Cauchy-Schwarz inequality (Theorem 1.10.2): for x, y in a vector space,

$$|\langle x, y\rangle| \le \|x\| \cdot \|y\|\,.$$

We will also abuse the notation and use $\|x(t)\|$ to mean $\|x\|$ and $\|x(t-\tau)\|$ to mean $\|y_\tau\|$ where $y_\tau(t) =$

$x(t-\tau)$.

Proof.

$$\begin{aligned}|\Gamma_{xx}(\tau)| &= |\langle x(t), x(t-\tau)\rangle| \\ &\leq \|x(t)\| \cdot \|x(t-\tau)\|\end{aligned}$$

We know that $E_x = \|x\|^2$, so $\|x\| = \sqrt{E_x}$. Similarly, since the energy of a signal does not change when the signal is shifted in time (Theorem 2.2.8), $\|x_\tau\| = \|x(t-\tau)\| = \sqrt{E_x}$. Therefore,

$$|\Gamma_{xx}(\tau)| \leq \sqrt{E_x}\sqrt{E_x} = E_x = \Gamma_{xx}(0)\,.$$

If $x \in \mathcal{L}^2(\mathbb{R})$, then the autocorrelation obtains its maximum at $\tau = 0$ (Theorem 2.2.9), meaning that the signal that maximizes its resemblance to a given signal is the signal itself, rather than the signal translated some distance into the past or future.

Proof of Theorem 2.2.4. We wish to prove that $|\Gamma_{xx}(\tau)| < +\infty$

$$\begin{aligned}|\Gamma_{xx}(\tau)| &\leq \Gamma_{xx}(0) && \text{; by Theorem 2.2.9} \\ &= E_x && \text{; by Corollary 2.2.5} \\ &< +\infty && x \in \mathcal{L}^2(\mathbb{R})\end{aligned}$$

2.3 Autocorrelation and periodic signals

And what about periodic signals? In particular, if x is T-periodic, then $x(t-T) = x(t) \quad \forall t \in \mathbb{R}$; thus

$$\begin{aligned}\Gamma_{xx}(T) &= \langle x(t), x(t-T)\rangle \\ &= \langle x(t), x(t)\rangle \\ &= \Gamma_{xx}(0) \\ &= \Gamma_{xx}(nT) \quad n \in \mathbb{Z}\,.\end{aligned}$$

From this it would seem that the autocorrelation of a periodic function is also periodic. Can this be true?

Well, not exactly.

Reminder 2.3.1

Recall that in Definition 2.2.1, x was assumed to be a finite energy signal ($x \in \mathcal{L}^2(R)$); and we have already seen that all non-zero periodic signals fail to be finite energy signals ($E_x = +\infty$).

However, periodic signals are finite average power signals.

Finally the only thing that is missing to define autocorrelation for periodic signals is an inner product. That's fortunate, because we have already defined (Definition 1.12.3) an inner product for $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$.

$$\langle x, y\rangle = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\,\overline{y(t)}\,dt\,.$$

At this point is suffices to replace inner product formula in the autocorrelation definition.

Definition 2.3.2 (***autocorrelation in*** $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$)

Let $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, the autocorrelation of x is defined as

$$\begin{aligned}\Gamma_{xx}(\tau) &= \langle x(t), x(t-\tau)\rangle \\ &= \lim_{T\to+\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)\,\overline{x(t-\tau)}\,dt\end{aligned}$$

Properties similar to what we know for $\mathcal{L}^2(\mathbb{R})$ still apply, sometimes replacing E_x with P_x.

Corollary 2.3.3 (***autocorrelation at zero equals average power***)

If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, then

$$\Gamma_{xx}(0) = \langle x(t), x(t)\rangle = P_x\,.$$

Corollary 2.3.4 (***Hermitian symmetry of autocorrelation (average power)***)

If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, then Γ_{xx} is symmetric Hermitian:

$$\Gamma_{xx}(-\tau) = \overline{\Gamma_{xx}(\tau)}\,.$$

Corollary 2.3.5 (***autocorrelation of a real signal is even (average power)***)

If $x(t) \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ is real valued, then Γ_{xx} is an even function:

$$\Gamma_{xx}(-\tau) = \Gamma_{xx}(\tau)\,.$$

Corollary 2.3.6 (***autocorrelation is maximal at zero (average power)***)

If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, and $\tau \in \mathbb{R}$, then $|\Gamma_{xx}(\tau)| \leq \Gamma_{xx}(0) = P_x$.

Corollary 2.3.7 (***autocorrelation of a periodic signal is periodic***)

If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, and if, additionally, x is T-periodic, then: Γ_{xx} is also T-periodic.

Proof. This is because if you translate x in time a distance of T, then x superimposes onto itself.

The definition of Γ_{xx} can be simplified as follows, by following the same reasoning as we used (in Theorem 1.13.3) when showing the average power of a T-periodic signal is simply the average power over one period.

$$\begin{aligned}\Gamma_{xx}(\tau) &= \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)\,\overline{x(t-\tau)}\,dt \\ &= \frac{1}{T}\int_0^T x(t)\,\overline{x(t-\tau)}\,\mathrm{d}t\end{aligned}$$

Table 2.1 summarizes. We can be happy to remember that

$$\Gamma_{xx}(\tau) = \langle x(t), x(t-\tau)\rangle\ ,$$

and not have to care about whether $x \in \mathcal{L}^2(\mathbb{R})$, or $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$.

$x \in \mathcal{L}^2(\mathbb{R})$	$\Gamma_{xx}(\tau) = \int_{\mathbb{R}} x(t)\,\overline{x(t-\tau)}\,\mathrm{d}t$ $\Gamma_{xx}(0) = E_x$
$x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$	$\Gamma_{xx}(\tau) = \lim_{T\to+\infty} \frac{1}{T} \int_0^T x(t)\,\overline{x(t-\tau)}\,\mathrm{d}t$ $\Gamma_{xx}(0) = P_x$
If x is T-periodic, and $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$	$\Gamma_{xx}(\tau) = \frac{1}{T} \int_0^T x(t)\,\overline{x(t-\tau)}\,\mathrm{d}t$ $\Gamma_{xx}(0) = \Gamma_{xx}(nT) = P_x$ Γ_{xx} is T-periodic

Table 2.1: Summary of computation of Γ_{xx}

2.4 Cross-Correlation

In Section 2.2, we looked at the special case of cross-correlation of a signal with itself, $x = y$. Now we can attack the case of $x \neq y$.

Definition 2.4.1 (*cross-correlation*)

Let x, y be signals. We call the function $\Gamma_{xy} : \mathbb{R} \to \mathbb{R}$ the cross-correlation between the signals x and y. $\Gamma_{xy}(\tau)$ measures the resemblance between x and y delayed τ times units into the future.

Γ_{xy} is defined as:

$$\Gamma_{xy}(\tau) = \langle x(t), y(t-\tau) \rangle \ .$$

As we have seen, this inner product is different in different vector spaces. If $x, y \in \mathcal{L}^2(\mathbb{R})$,

$$\Gamma_{xy}(\tau) = \langle x(t), y(t-\tau) \rangle = \int_{\mathbb{R}} x(t)\, \overline{y(t-\tau)}\, \mathrm{d}t \, .$$

If $x, y \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$,

$$\Gamma_{xy}(\tau) = \langle x(t), y(t-\tau) \rangle = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, \overline{y(t-\tau)}\, dt \, .$$

And if $x, y \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, both T-periodic, (*i.e.*, if both x and y have the *same* period)

$$\Gamma_{xy}(\tau) = \langle x(t), y(t-\tau) \rangle = \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, \overline{y(t-\tau)}\, dt \, .$$

In Section 2.2, we saw several notable properties of Γ_{xx}. At this point there are no analogous properties of Γ_{xy} to discuss. We will see some such properties later when we discuss the Fourier transform.

Definition 2.4.2 (*reflection*)

Let $y : \mathbb{R} \to \mathbb{C}$ be a signal. To ease with notation we define the function y^- as follows:

$$y^-(t) = y(-t) \, .$$

We call y^-, the reflection of y about 0.

Lemma 2.4.3 (*correlation of reflected signals*)

$$\Gamma_{x^- y^-}(\tau) = \Gamma_{xy}(-\tau)$$

Proof.

$$
\begin{aligned}
\Gamma_{x^-y^-}(\tau) &= \int_{\mathbb{R}} x^-(t)\,\overline{y^-(t-\tau)}\,\mathrm{d}t \\
&= \int_{\mathbb{R}} x(-t)\,\overline{y(\tau-t)}\,\mathrm{d}t \\
&= \int_{\infty}^{-\infty} x(u)\,\overline{y(\tau+u)}\,(-\,\mathrm{d}u) && \text{letting } u=-t, \mathrm{d}u=-\,\mathrm{d}t \\
&= \int_{\mathbb{R}} x(u)\,\overline{y(\tau+u)}\,\mathrm{d}u \\
&= \int_{\mathbb{R}} x(u)\,\overline{y(u-(-\tau))}\,\mathrm{d}u \\
&= \Gamma_{xy}(-\tau)
\end{aligned}
$$

Theorem 2.4.4 (***cross-correlation non-commutative***)

$$\Gamma_{xy}(\tau) = \overline{\Gamma_{yx}(-\tau)}\ .$$

Proof.

$$
\begin{aligned}
\Gamma_{xy}(\tau) &= \int_{\mathbb{R}} x(t)\,\overline{y(t-\tau)}\,\mathrm{d}t \\
&= \int_{-\infty}^{\infty} x(t)\,\overline{y(t-\tau)}\,\mathrm{d}t
\end{aligned}
$$

Change of variables. Let $u = \tau - t$, $\mathrm{d}u = -\,\mathrm{d}t$, and $t = \tau - u$.

$$
\begin{aligned}
t = -\infty &\implies u = \infty && \text{; computing lower limit of integration} \\
t = \infty &\implies u = -\infty && \text{; computing upper limit of integration}
\end{aligned}
$$

$$
\begin{aligned}
\Gamma_{xy}(\tau) &= \int_{\infty}^{-\infty} x(\tau-u)\,\overline{y(-u)}\,(-\,\mathrm{d}u) \\
&= -\int_{\infty}^{-\infty} x(\tau-u)\,\overline{y(-u)}\,\mathrm{d}u \\
&= \int_{-\infty}^{\infty} x(\tau-u)\,\overline{y(-u)}\,\mathrm{d}u \\
&= \int_{\mathbb{R}} \overline{y(-u)}\,x(-(u-\tau))\,\mathrm{d}u \\
&= \int_{\mathbb{R}} \overline{y^-(u)}\,x^-(u-\tau)\,\mathrm{d}u \\
&= \overline{\int_{\mathbb{R}} y^-(u)\,\overline{x^-(u-\tau)}\,\mathrm{d}u} \\
&= \overline{\Gamma_{y^-x^-}(\tau)} \\
&= \overline{\Gamma_{yx}(-\tau)} && \text{by Lemma 2.4.3}
\end{aligned}
$$

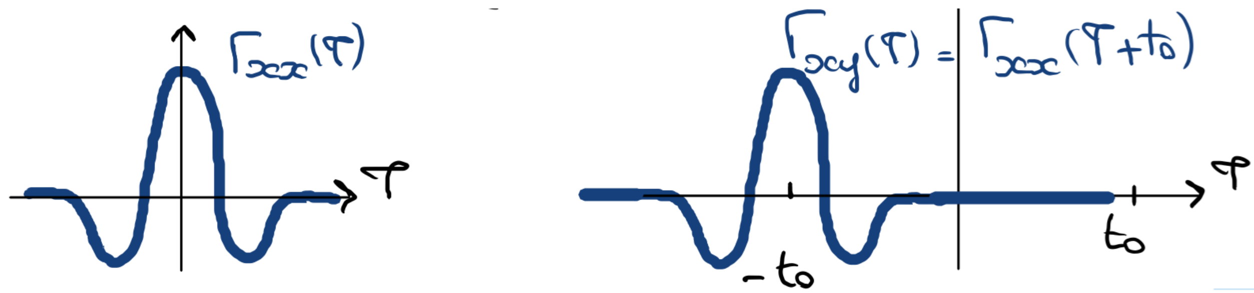


Figure 2.1: Autocorrelation delayed compared to cross-correlation

2.5 Back to the Radar Example

With the tools we have developed thus far, we may now correctly treat the example of the radar signal from start to finish.

2.5.1 Echo without noise

In the case of a noise-free echo:

$$x'_{echo}(t) = x_{ref}(t - t_o)\,,$$

with $t_0 > 0$ being an unknown time delay which we would like to estimate. To simplify the notation let

$$\begin{aligned} x(t) &= x_{ref}(t) \\ y(t) &= x'_{echo}(t) \\ &= x(t - t_0) \end{aligned}$$

What is the value of Γ_{xy}?

$$\begin{aligned} \Gamma_{xy}(\tau) &= \langle x(t),\ y(t-\tau)\rangle \\ &= \langle x(t),\ x((t-t_0)-\tau)\rangle \\ &= \langle x(t),\ x(t-(t_0+\tau))\rangle \\ &= \Gamma_{xx}(t_0+\tau) \end{aligned}$$

We see that Γ_{xy} is a delayed version of Γ_{xx} — pushed into the future by exactly t_0. We know that Γ_{xx} obtains its maximum value at 0, so $\tau \mapsto \Gamma_{xy}(\tau)$ is maximal when $t_0 + \tau = 0$, or $\tau = -t_0$. See Figure 2.1.

To determine the unknown delay, from t_0, to x and y (with $y(t) = x(t - t_0)$), it suffices to follow these steps:

1. Calculate the cross-correlation: Γ_{xy} of x to y.
2. Find the time point $\tau_{\max}$ where Γ_{xy} is maximized:

$$\tau_{\max} = \texttt{argmax}\ \Gamma_{xy}(\tau)\,.$$

   The `argmax` function returns the value $\tau \in \mathbb{R}$ which maximizes $\Gamma_{xy}(\tau)$.

3. $t_0 + \tau = t_0 + \tau_{\max} = 0 \implies t_0 = -\tau_{\max}$. Since the echo is delayed into the future, we know $t_0 > 0$, which means $\tau_{\max} < 0$.

This makes sense logically; we must translate y to the left (into the past) to make the resemblance maximal: $\tau_{\max} = -t_0 < 0$.

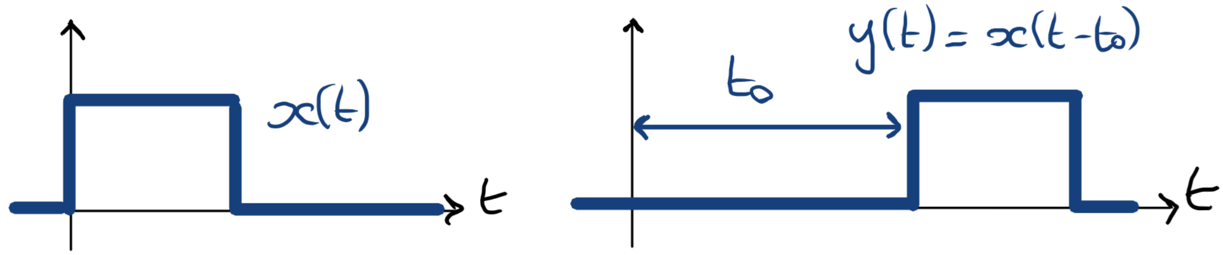


Figure 2.2: Translate y to maximize resemblance.

2.5.2 Echo with noise

In the case of echo with noise:

$$x'_{echo}(t) = x_{ref}(t - t_0) + \eta(t)\,,$$

with η representing a noise independent of the signal x_{ref}. For η, we consider white Gaussian noise, $\eta(t) \approx \mathcal{N}(0, \sigma^2)$.

Similar to Section 2.5.1, we denote:

$$\begin{aligned} x(t) &= x_{ref}(t) \\ y(t) &= x'_{echo}(t) \\ &= x(t - t_0) + \eta(t) \end{aligned}$$

To rigorously treat this type of signal, we would need to extend the notation of inner product to accommodate random signals (call stochastic processes) but this is beyond the scope of this course. Let's see what happens when we just use the definitions we already have.

$$\begin{aligned} \Gamma_{xy}(\tau) &= \langle x(t),\ y(t - \tau)\rangle \\ &= \langle x(t),\ x(t - t_0 - \tau) + \eta(t - \tau)\rangle \\ &= \langle x(t),\ x(t - t_0 - \tau)\rangle + \langle x(t),\ \eta(t - \tau)\rangle && \text{; by (1.30)} \\ &= \langle x(t),\ x(t - (t_0 + \tau))\rangle + \langle x(t),\ \eta(t - \tau)\rangle \\ &= \Gamma_{xx}(t_0 + \tau) + \Gamma_{x\eta}(\tau) \end{aligned}$$

Since we are assuming the noise, η, is independent of the signal, x, we can conclude that $\Gamma_{x\eta}(\tau) = 0 \quad \forall \tau \in \mathbb{R}$. So we have,

$$\Gamma_{xy}(\tau) = \Gamma_{xx}(t_0 + \tau)\,,$$

which is exactly the equivalence we saw in the noise-free case in Section 2.5.1.

Noise has no impact (in theory) on the estimation of the time delay in cross-correlation.

For general purposes, the cross-correlation function proves very useful in signal processing in shape recognition, as illustrated in Figure 2.3. In such a case, one of the two signals plays the role of template, and all the maxima of the cross-correlation function identify an occurrence of the template within the signal being analyzed.

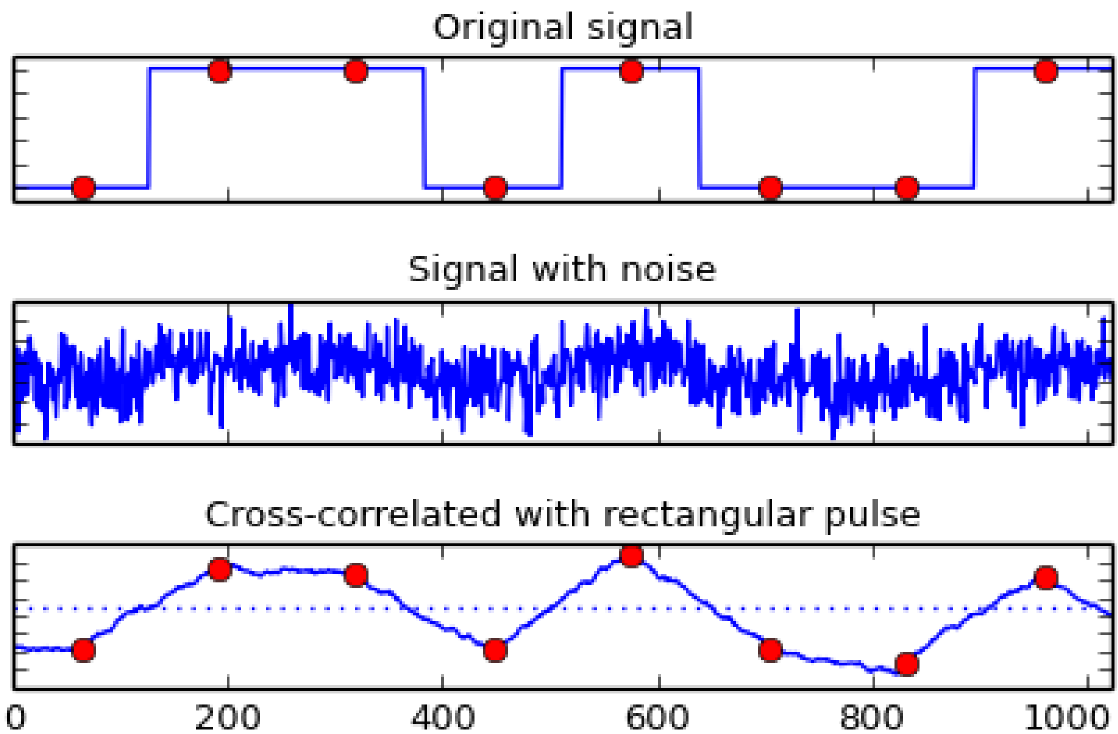


Figure 2.3: Using cross-correlation for pattern recognition

2.6 The Convolution Product

Everyone today has certainly heard talk about the *convolution product* (also called simply *convolution*), because this operation is at the core of convolutional neural networks (CNNs). But this bizarre operation between two tables of values is only the 2D discrete representation (Figure 2.4) of a mathematical operation defined at the lowest level on functions. More precisely (and this is anyway the role played by CNNs), the product of convolution is the mathematical representation of the notion of linear filter. As shown in Figure 2.5, it is the operation of convolution which makes model filtering possible.

Notably, thanks to the theory of Plancherel, which we will see later, when we talk about the Fourier transform, the convolution is unavoidable.

Definition 2.6.1 (*convolution*)

Let x and y be signals. The *convolution product*, $(x * y) : \mathbb{R} \to \mathbb{C}$ of x and y is defined as:

$$(x * y)(\tau) = \int_{\mathbb{R}} x(t)y(\tau - t)\,\mathrm{d}t$$

In other words:

- We start with two signals x and y.
- We transform (we reflect) y into y^{-}.
- As shown in Figure 2.6, we translate y^{-} to the left by the amount τ: $y(-t + \tau) = y(\tau - t)$.
- For that fixed value of τ, we compute the area of the point-wise product of x with y reflected and

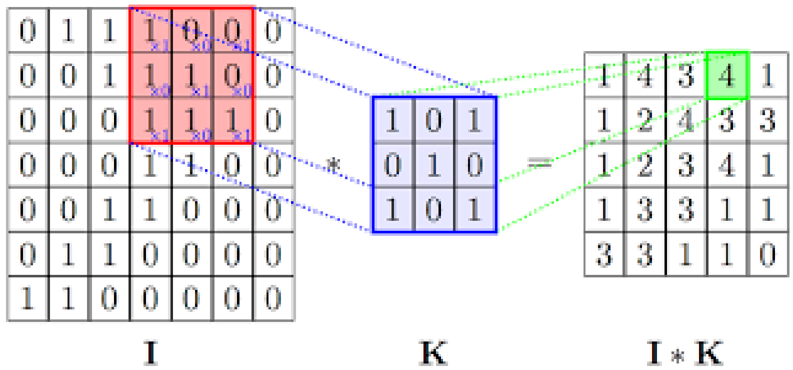


Figure 2.4: 2D discrete convolution

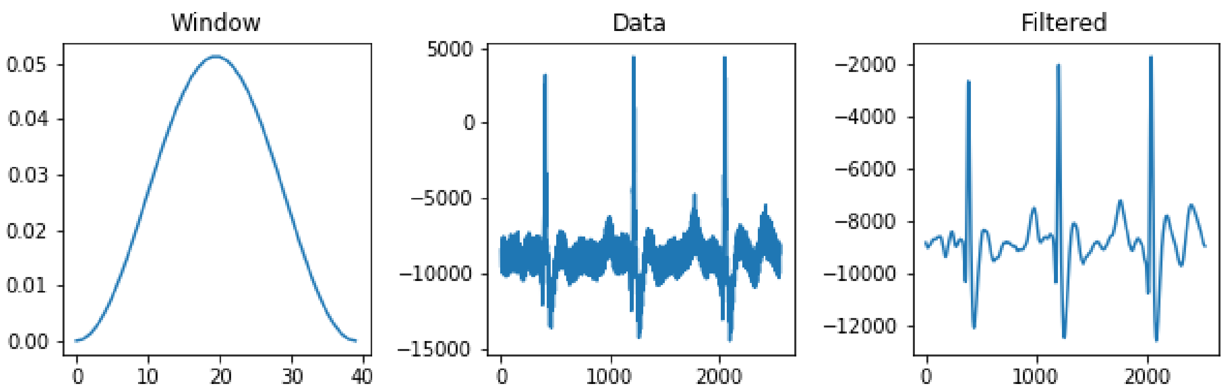


Figure 2.5: Convolution makes it possible to model filtering

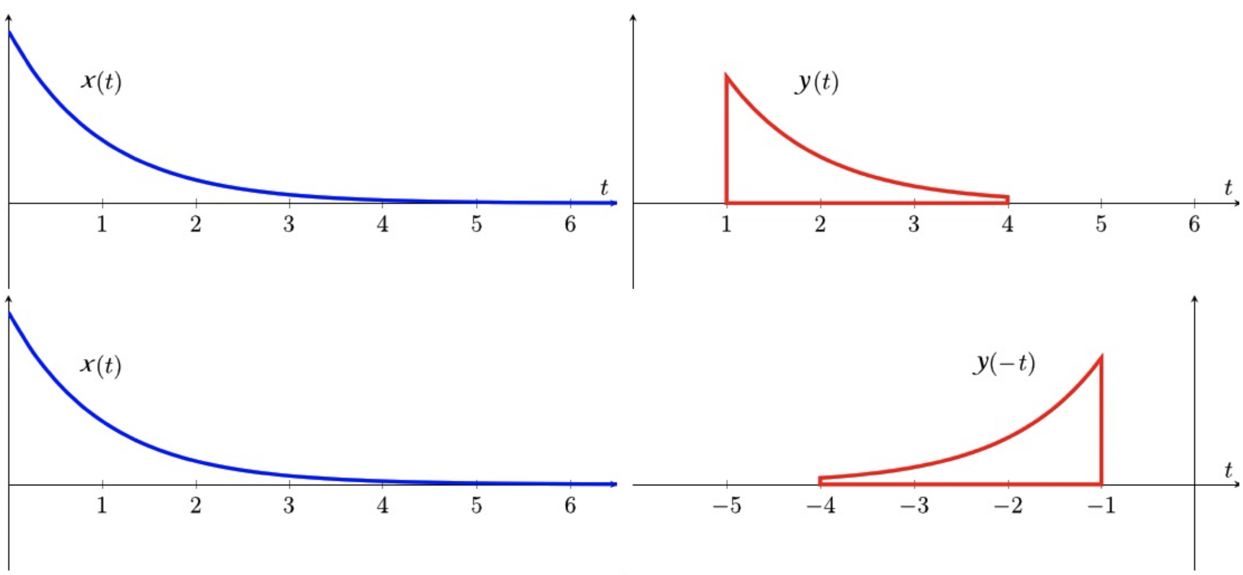


Figure 2.6: Signal reflected about the y-axis

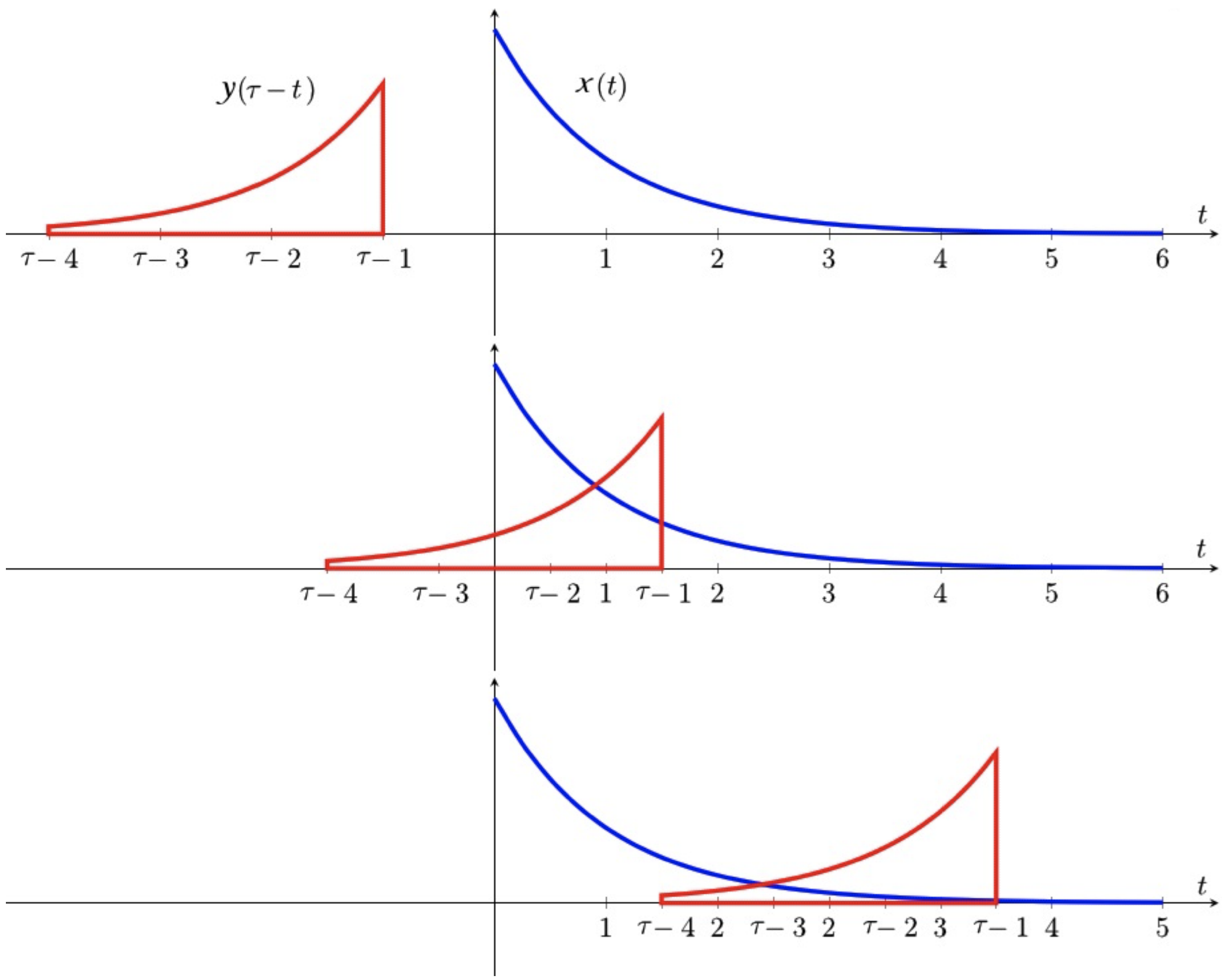


Figure 2.7: To compute the convolution, y^- is translated by τ, and the integral of the produce is computed for each τ.

translated:

$$\int_{\mathbb{R}} x(t)y(\tau - t)\,\mathrm{d}t$$

- This gives us the value of the convolution for the fixed value of τ:

$$(x * y)(\tau) = \int_{\mathbb{R}} x(t)y(\tau - t)\,\mathrm{d}t\,.$$

- We compute this integral as a function of τ. As shown in Figure 2.7, as τ varies, $y(\tau - t)$ *slides* along the x axis, and the value $\int_{\mathbb{R}} x(t)y(\tau - t)\,\mathrm{d}t$ is effectively computed for each value of $\tau \in \mathbb{R}$.

The signal y plays the role of *sliding window* through which we observe x. The convolution may be seen as the generalization of the running (sliding) average for functions.

2.7 Properties of Convolution

Theorem 2.7.1 (***Fubini's Theorem***)

If $\int_{\mathbb{R}\times\mathbb{R}} |f(x,y)|\, d(x,y) < \infty$, then

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y) dx \right) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y) dy \right) dx$$

Fubini's theorem gives us a sufficient condition from commuting the integral signs of a double integral.

Example 2.7.2 (***Fubini's theorem can fail***)

$$\int_{x=0}^{1} \left(\int_{y=0}^{1} \frac{x^2 - y^2}{(x^2+y^2)^2} dy \right) dx = \frac{\pi}{4}$$
$$\int_{y=0}^{1} \left(\int_{x=0}^{1} \frac{x^2 - y^2}{(x^2+y^2)^2} dx \right) dy = -\frac{\pi}{4}$$

These two integrals give different results because the conditions of Fubini's theorem do not hold.

$$\int_0^1 \int_0^1 \left| \frac{x^2 - y^2}{(x^2+y^2)^2} \right| dx\, dy = \infty$$

There are several important properties of the convolution:

Existence If $x, y \in \mathcal{L}^1(\mathbb{R})$, then $(x * y)$ exists and $(x * y) \in \mathcal{L}^1(\mathbb{R})$. This is a consequence of the Theorem of Fubini (Theorem 2.7.1).

Commutativity The convolution product is commutative:

$$x * y = y * x \tag{2.2}$$
$$\int_{\mathbb{R}} x(t) y(\tau - t)\, \mathrm{d}t = \int_{\mathbb{R}} x(t - \tau) y(t)\, \mathrm{d}t \quad \forall \tau \in \mathbb{R} \tag{2.3}$$
$$\tag{2.4}$$

Bilinearity

$$x * (y + \lambda z) = x * y + \lambda (x * z) \tag{2.5}$$

Associativity

$$x * y * z = (x * y) * z = x * (y * z)$$

Identity There is no function, $n \in \mathcal{L}^1(\mathbb{R})$, which satisfies the identity $n * x = x$ nor $x * n = x$. However, it is useful (and powerful) to create a notation which we will see later, to make believe that such an identity exists.

Bias If $y \in \mathcal{L}^1(\mathbb{R})$, $\int_{\mathbb{R}} x(t)\, \mathrm{d}t < \infty$, and C is a constant,

$$x * (y + C) = (x * y) + C \int_{\mathbb{R}} x(t)\, \mathrm{d}t\,.$$

Symmetries If x and y are signals, even or odd,

- $x * y$ is even, if x and y are both even or both odd.
- $x * y$ is odd, if x and y have different parity, *i.e.*, if x is even and y is odd, or if x is odd and y is even.

Differentiation If x and y are differentiable, and $x, x', y, y' \in \mathcal{L}^1(\mathbb{R})$, then $x * y$ is differentiable and:

$$(x * y)' = (x' * y) = (x * y')\,.$$

Integration If $X(\tau) = \int_{-\infty}^{\tau} x(t)\,\mathrm{d}t$ and $Y(\tau) = \int_{-\infty}^{\tau} y(t)\,\mathrm{d}t$,

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

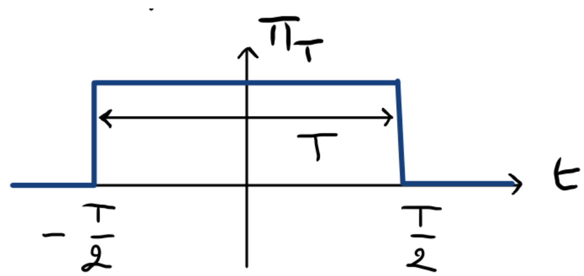


Figure 2.8: Window (pulse) of size T

2.8 Example Computation of Convolution

In this example we *convolve*[1] the window function of size T, Π_T with itself. The window function of size T is illustrated in Figure 2.8

$$\Pi_T = \begin{cases} 1 & \text{if } -\frac{T}{2} <= t \leq \frac{T}{2} \\ 0 & \text{otherwise} \end{cases}$$

We wish to compute

$$\left(\Pi_T * \Pi_T\right)(\tau) = \int_{\mathbb{R}} \Pi_T(t)\Pi_T(\tau - t)\,\mathrm{d}t\,.$$

We will compute the convolution integral in two different ways shown in Sections 2.8.1 and 2.8.2. Both approaches give the same result—the function whose graph is shown in Figure 2.9. The solution in Section 2.8.1 is provided for completeness, but the student will probably find the solution in Section 2.8.2 more intuitive.

2.8.1 Computing convolution analytically

$$\begin{aligned}
I(\tau) &= (\Pi_T * \Pi_t)(\tau) \\
&= \int_{\mathbb{R}} \Pi_T(t)\Pi_T(\tau - t)\,\mathrm{d}t \\
&= \int_{-\infty}^{-T/2} \underbrace{\Pi_T(t)\Pi_T(\tau - t)\,\mathrm{d}t}_{\Pi_T \text{ is 0, if } t<-T/2} \\
&\qquad + \int_{-T/2}^{T/2} \underbrace{\Pi_T(t)\Pi_T(\tau - t)\,\mathrm{d}t}_{\Pi_T \text{ is 1, if } T/2<t<T/2} \; dt \\
&\qquad + \int_{T/2}^{\infty} \underbrace{\Pi_T(t)\Pi_T(\tau - t)\,\mathrm{d}t}_{\Pi_T \text{ is 0, if } T/2<t} \\
&= \int_{-T/2}^{T/2} \Pi_T(\tau - t)dt
\end{aligned}$$

[1]The verb (in English) corresponding to the noun, *convolution* is *convolve*, not *convolute*. To *convolute* means to make something confusing. Even though the convolution product may be confusing at times, it is correct to say: we compute the convolution of two signals by *convolving* them.

Now we employ a simple change of variables:

$$\begin{aligned} u &= \tau - t \\ \mathrm{d}u &= -\,\mathrm{d}t \\ -\,\mathrm{d}u &= \mathrm{d}t \end{aligned}$$

So we can rewrite the integral above as:

$$\begin{aligned} I(\tau) &= \int_{\tau+T/2}^{\tau-T/2} -\Pi_T(u)\,\mathrm{d}u \\ &= \int_{\tau-T/2}^{\tau+T/2} \Pi_T(u)\,\mathrm{d}u \end{aligned}$$

To compute I, we must integrate a window function over a range. The gate, Π_T has value 1 between $-T/2$ and $T/2$ and 0 otherwise; while the range is $\tau - T/2$ to $\tau + T/2$. Since $T > 0$, we have $-T/2 < T/2$ and $\tau - T/2 < \tau + T/2$. This means 4 cases are possible.

Case 1: $\tau - T/2 < \tau + T/2 \leq -T/2 < T/2$

$$I(\tau) = \int_{\tau-T/2}^{\tau+T/2} 0\,\mathrm{d}u = 0$$

In this case

$$\tau + T/2 \leq -T/2 \implies \tau < -T$$

So if $\tau < -T$, we have $I(\tau) = 0$.

Case 2: $\tau - T/2 \leq -T/2 \leq \tau + T/2 \leq T/2$

$$\begin{aligned} I(\tau) &= \int_{\tau-T/2}^{-T/2} 0\,\mathrm{d}u + \int_{-T/2}^{\tau+T/2} 1\,\mathrm{d}u \\ &= 0 + u\Big|_{-T/2}^{\tau+T/2} \\ &= \tau + T/2 + T/2 \\ &= \tau + T \end{aligned}$$

In this case

$$\begin{aligned} -T/2 \leq \tau + T/2 \leq T/2 &\implies -T/2 - T/2 \leq \tau \leq T/2 - T/2 \\ &\implies -T \leq \tau \leq 0 \end{aligned}$$

So if $-T \leq \tau \leq 0$, we have $I(\tau) = \tau + T$. The graph (Figure 2.9) consists of a line of slope of 1, y-intercept of T, and x-intercept of $-T$.

Case 3: $-T/2 \leq \tau - T/2 \leq T/2 \leq \tau + T/2$

$$\begin{aligned} I(\tau) &= \int_{\tau-T/2}^{T/2} 1\,\mathrm{d}u + \int_{T/2}^{\tau+T/2} 0\,\mathrm{d}u \\ &= u\Big|_{\tau-T/2}^{T/2} + 0 \\ &= T/2 - (\tau - T/2) \\ &= T - \tau \end{aligned}$$

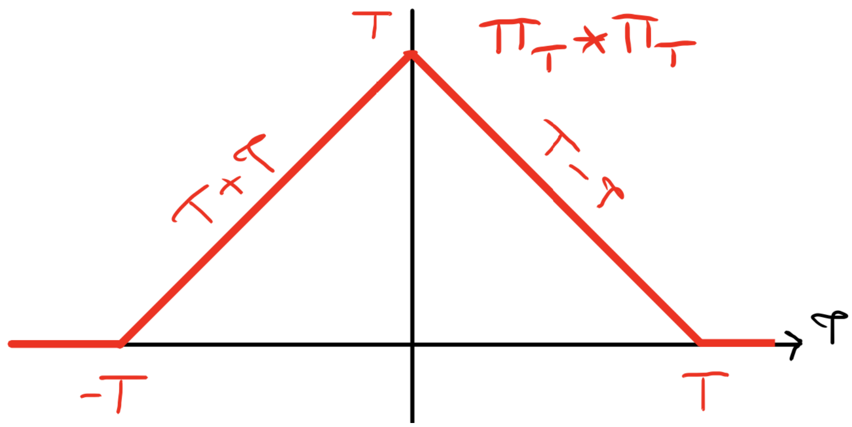


Figure 2.9: Graph of $I(\tau) = (\Pi_T * \Pi_T)(\tau)$

In this case

$$\begin{aligned} -T/2 \leq \tau - T/2 \leq T/2 &\implies -T/2 + T/2 \leq \tau \leq T/2 + T/2 \\ &\implies 0 \leq \tau \leq T \end{aligned}$$

So if $0 \leq \tau \leq T$, we have $I(\tau) = T - \tau$. The graph (Figure 2.9) consists of a line of slope -1, y-intercept of T, and x-intercept of T.

Case 4: $-T/2 < T/2 \leq \tau - T/2 < \tau + T/2$

$$I(\tau) = \int_{\tau - T/2}^{\tau + T/2} 0 \,\mathrm{d}u = 0$$

In this case

$$\begin{aligned} T/2 \leq \tau - T/2 &\implies T/2 + T/2 \leq \tau \\ &\implies \tau \geq T \end{aligned}$$

So if $\tau \geq T$, we have $I(\tau) = 0$.

From these 4 cases, we can construct the graph shown in Figure 2.9 as

$$I(\tau) = \begin{cases} 0 & \text{if } \tau \leq -T \\ T + \tau & \text{if } -T \leq \tau \leq 0 \\ T - \tau & \text{if } 0 \leq \tau \leq T \\ 0 & \text{if } T < \tau \end{cases}$$

2.8.2 Computing convolution visually

The second solution we will look at for computing $I(\tau) = \Pi_T * \Pi_T(\tau)$ is also based on 4 cases, but these cases are more visual than those of the solution in Section 2.8.1.

In Figure 2.10 we see two cases when the two window pulses do not overlap. These are numbered cases 1 and 4. In Figure 2.11 we see two cases of the two window pulses overlapping. These are numbered cases 2 and 3.

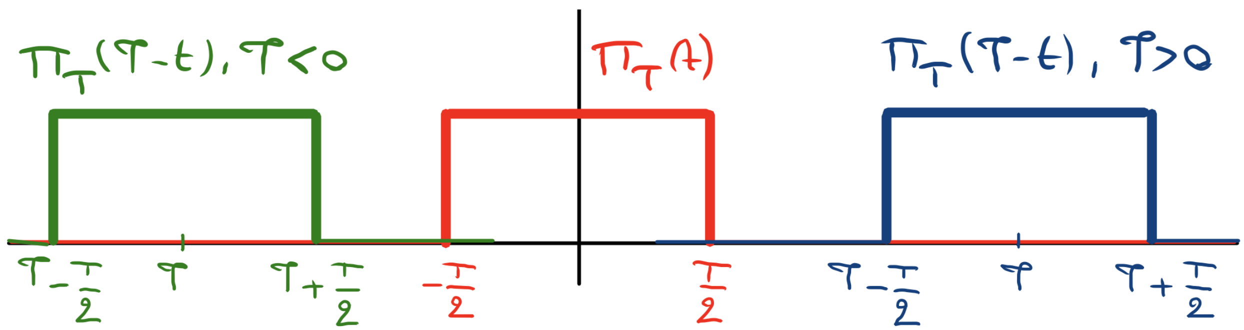


Figure 2.10: $\Pi_T(\tau)$ not overlapping $\Pi_T(\tau - T)$

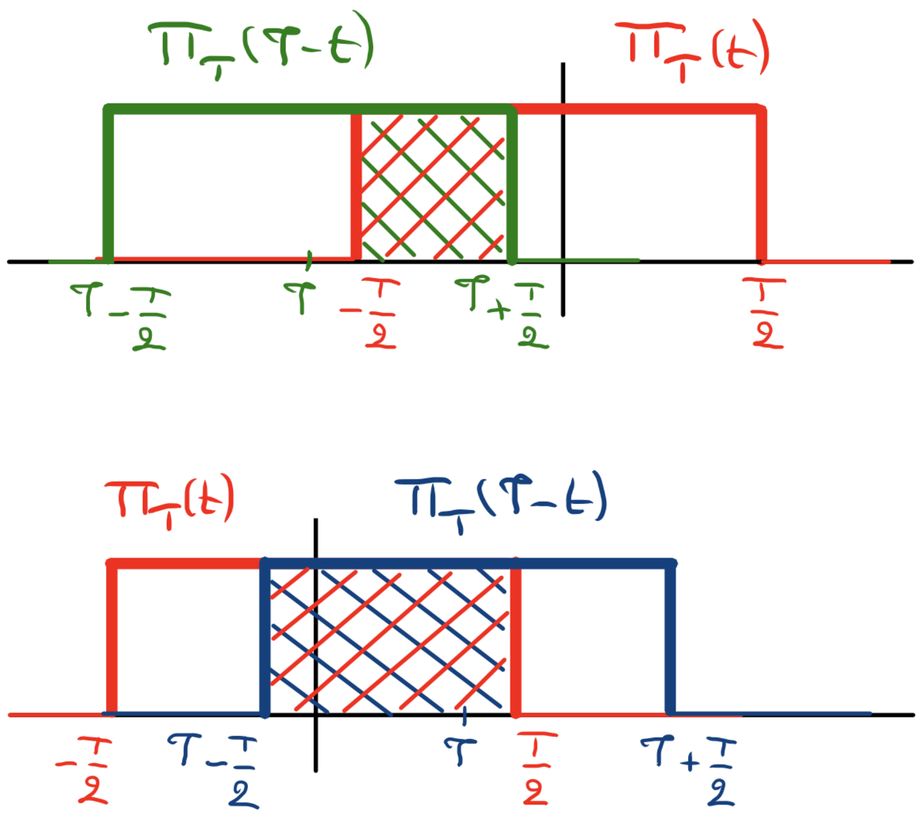


Figure 2.11: $\Pi_T(\tau)$ overlapping $\Pi_T(\tau - T)$

Case 1: Non-overlapping: $\tau + T/2 \leq -T/2 \implies \tau \leq -T$
From Figure 2.10 we can see in this range the product $\Pi_T(t) \times \Pi_T(\tau - t) = 0$, because the window pulses do not overlap, so

$$I(\tau) = 0 \text{ ; if } \tau \leq -T$$

Case 2: Overlapping: $-T/2 \leq \tau + T/2 \leq T/2 \implies -T \leq \tau \leq 0$
From Figure2.11 (top) we can see the area of the overlapping (cross-hatched) region is

$$\begin{aligned} I(\tau) &= \texttt{width} \times \texttt{height} \\ &= (\tau + T/2 - (-T/2)) \times 1 \\ &= \tau + T \qquad\qquad\qquad\qquad ; \text{if } -T \leq \tau \leq 0 \end{aligned}$$

In this range, the graph has a slope of 1, y-intercept of T, and x-intercept of $-T$.

Case 3: Overlapping: $-T/2 \leq \tau - T/2 \leq T/2 \implies 0 \leq \tau \leq T$
From Figure2.11 (bottom) we can see the area of the overlapping (cross-hatched) region is

$$\begin{aligned} I(\tau) &= \texttt{width} \times \texttt{height} \\ &= (T/2 - (\tau - T/2)) \times 1 \\ &= T - \tau \qquad\qquad\qquad\qquad ; \text{if } 0 \leq \tau \leq T \end{aligned}$$

In this range, the graph has a slope of -1, y-intercept of T, and x-intercept of T.

Case 4: Non-overlapping, $T/2 \leq \tau - T/2 \implies \tau \geq T$
From Figure 2.10 we can see in this range the product $\Pi_T(t) \times \Pi_T(\tau - t) = 0$, because the window pulses do not overlap, so

$$I(\tau) = 0 \text{ ; if } \tau \geq T$$

From these 4 cases, we can construct the graph shown in Figure 2.9 as

$$I(\tau) = \begin{cases} 0 & \text{if } \tau \leq -T \\ T + \tau & \text{if } -T \leq \tau \leq 0 \\ T - \tau & \text{if } 0 \leq \tau \leq T \\ 0 & \text{if } T < \tau \end{cases},$$

which is the exact equation we found in Section 2.8.1.

We also observe from Figure 2.9 that

$$(\Pi_T * \Pi_T)(0) = T\,,$$

and this is the maximum value of $(\Pi_T * \Pi_T)(\tau)$.

Finally, we notice that it is the case (and general) that the support (see Definition 1.10.5) of the convolution in Figure 2.9 is larger than that of either of the functions being convolved.

2.9 Convolution and Big-Integer Multiplication

Let's look at an example of executing algorithm you leaned as a child to compute integer multiplication; although you were probably not told that this is an algorithm. We will multiply

$$6,543,210 \times 54,321 = 355,433,710,410\,.$$

					6	5	4	3	2	1	0	
$\times$							5	4	3	2	1	
					6	5	4	3	2	1	0	shifting rows
				12	10	8	6	4	2	0		
			18	15	12	9	6	3	0			
		24	20	16	12	8	4	0				
	30	25	20	15	10	5	0					
	30	49	58	58	50	35	20	10	4	1	0	column sums

Once the column sums have been computed, we may add them with carry propagation to obtain the final sequence of digits.

3	4	5	5	5	3	2	1					carry propagation
	0	9	8	8	0	5	0	0	4	1	0	
3	5	5	4	3	3	7	1	0	4	1	0	

This multiplication algorithm has quadratic complexity, related to $m \times n$, if we consider that 6543210 has n digits and 54321 has m digits. The carry propagation has linear complexity related to $m + n$, which is negligible.

Somewhat surprisingly we may compute the column sums directly using convolution. We consider the digits as a signal $x = (6, 5, 4, 3, 2, 1, 0)$ and $y = (5, 4, 3, 2, 1)$, then compute $x * y$ as follows: $y^- = (1, 2, 3, 4, 5)$. *E.g.*, we obtain the 5th column , 20, (counting from the right) by computing

$$\begin{aligned}\langle(4,3,2,1,0),(1,2,3,4,5)\rangle &= 4\times 1 + 3\times 2 + 2\times 3 + 1\times 4 + 0\times 5\\ &= 4 + 6 + 6 + 4 + 0\\ &= 20\,.\end{aligned}$$

The other column sums can be computed in Table 2.2. A justification why a column sum is computed as an inner product is explained in Section 2.10.2.

In elementary school you learned the algorithm to compute the shifting rows, them add them up to get the column sums. However, using the convolution product, you compute the column sums directly, each as an inner product of one factor with the reverse of the other factor.

To compute each of the column sums, we compute an inner product of two vectors, each having a maximum length of m. So at first glance it would seem that the complexity is roughly the same as before $O((n{+}m){\times}m)$. However, since the sequence of column sums is simply the convolution of x with y, it can be computed with complexity $O(n \log n)$, using FFT techniques.

2.10 Connection Between Convolution and Cross-Correlation

In view of the observation that the two formulas

$$\begin{aligned}\Gamma_{xy}(\tau) &= \int_{\mathbb{R}} x(t)\,\overline{y(t-\tau)}\,\mathrm{d}t\\ (x * y)(\tau) &= \int_{\mathbb{R}} x(t) y(\tau - t)\,\mathrm{d}t\end{aligned}$$

are similar, we should most certainly be able to find a connection between these operations.

column	u	v	**sum** $= \langle u, v\rangle$
1	(0)	(1)	0
2	(1,0)	(1,2)	1
3	(2,1,0)	(1,2,3)	4
4	(3,2,1,0)	(1,2,3,4)	10
5	(4,3,2,1,0)	(1,2,3,4,5)	20
6	(5,4,3,2,1)	(1,2,3,4,5)	35
7	(6,5,4,3,2)	(1,2,3,4,5)	50
8	(6,5,4,3)	(2,3,4,5)	58
9	(6,5,4)	(3,4,5)	58
10	(6,5)	(4,5)	49
11	(6)	(5)	30

Table 2.2: Column sums expressed as inner products

2.10.1 Convolution as Cross-Correlation

Let x, y be finite energy signals such that

$$x, y \in \mathcal{L}^2(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})\,.$$

We require $x, y \in \mathcal{L}^2(\mathbb{R})$ to assure that the cross-correlation exist, and we require $x, y \in \mathcal{L}^1(\mathbb{R})$ to assure that the convolution be defined.

We start with $x * y$ to try to recover Γ_{xy}:

$$(x * y)(\tau) = \int_{\mathbb{R}} x(t)y(\tau - t)\,\mathrm{d}t\,.$$

We note that (Definition 2.4.2) $y^- : t \mapsto y(-t)$ (the symmetric of y with respect to the ordinate axis (y-axis):

$$(x * y^-)(\tau) = \int_{\mathbb{R}} x(t)y^-(\tau - t)\,\mathrm{d}t = \int_{\mathbb{R}} x(t)y(t - \tau)\,\mathrm{d}t\,.$$

Now we simply have to introduce the complex conjugate:

$$(x * \overline{y^-})(\tau) = \int_{\mathbb{R}} x(t)\overline{y(t - \tau)}\,\mathrm{d}t = \Gamma_{xy}(\tau)\,.$$

So we have the result

$$\Gamma_{xy} = (x * \overline{y^-})\,.$$

In the case that the signal, y, is real valued, $y(t) \in \mathbb{R}$, then we know that $\overline{y(t)} = y(t)$, so

$$\Gamma_{xy} = (x * y^-)\,.$$

The cross-correlation of x and y is identical to the convolution of x with y reflected (y^-).

2.10.2 Convolution as Inner Product

If we recall the notation:

$$\begin{aligned} y^-(t) &= y(-t) \\ y_\tau &= y(t - \tau) \\ y^-_\tau &= y(\tau - t)\,, \end{aligned}$$

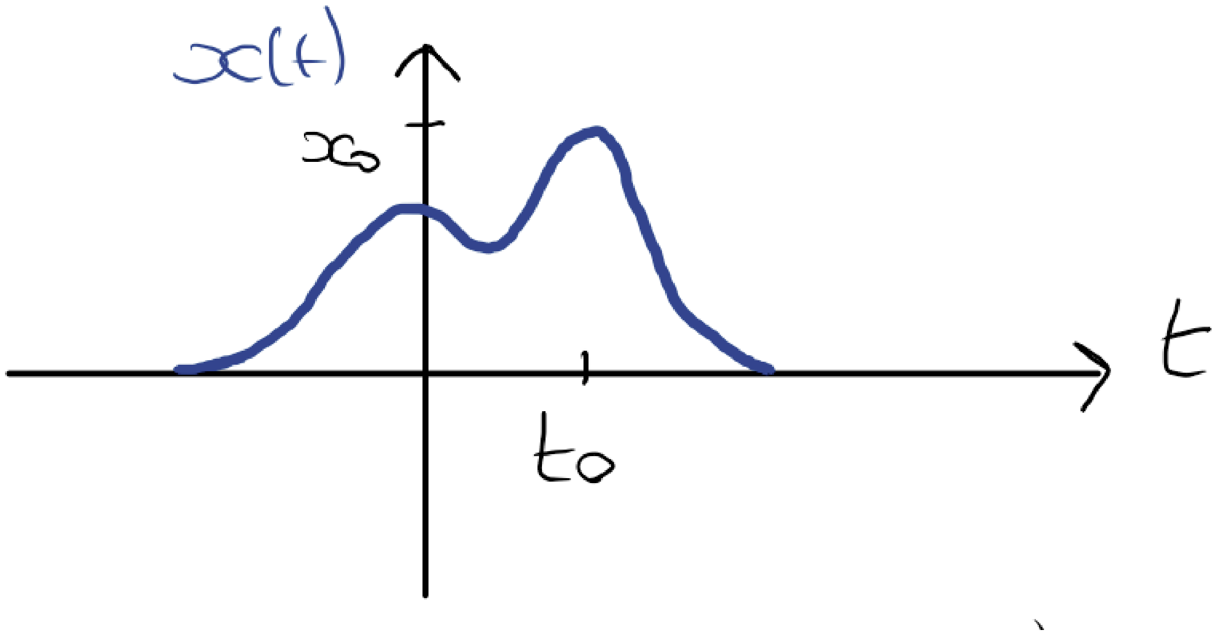


Figure 2.12: Signal of reference, $x(t)$; we'd like to derive $x(t-\tau)$ from $x(t)$.

We can rewrite the convolution integral at τ as follows:

$$\begin{aligned}(x * y)(\tau) &= \int_{\mathbb{R}} x(t) y(\tau - t)\,\mathrm{d}t \\ &= \int_{\mathbb{R}} x(t) y_\tau^-(t)\,\mathrm{d}t \\ &= \left\langle x(t),\, \overline{y_\tau^-(t)} \right\rangle \end{aligned} \tag{2.6}$$

The meaning of (2.6) is that to compute the value of the convolution at τ, we must compute the inner product $\left\langle x(t),\, \overline{y_\tau^-(t)} \right\rangle$, or simply $\langle x(t), y_\tau^-(t)\rangle$ if x and y are real valued. This is also illustrated in Table 2.2.

2.11 Translating a Function into the Past or Future

In this course we often have to manipulate expressions of the form $x(t-\tau)$ where $\tau \in \mathbb{R}$. It is necessary to convince ourselves that $x(t-\tau)$ is a copy of $x(t)$ but pushed τ distance *into the future* (to the right) if $\tau > 0$ and into the past (to the left) if $\tau < 0$. Even if (especially if) this is non-intuitive at first, the signal is indeed delayed (into the future) if $\tau > 0$ and it is advanced (into the past) if $\tau < 0$.

Consider the signal in Figure 2.12, the blue curve. The signal obtains a value of x_0 at $t = t_0$; *i.e.*,

$$x(t_0) = x_0 \,.$$

We'd like to construct the red curve in Figure 2.13 by *reading* the values from the blue curve. We consider the question: when does the red curve, $t \mapsto x(t-\tau)$ obtain the value x_0, given that $t \mapsto x(t)$ obtains the value t_0 at $t = 0$.

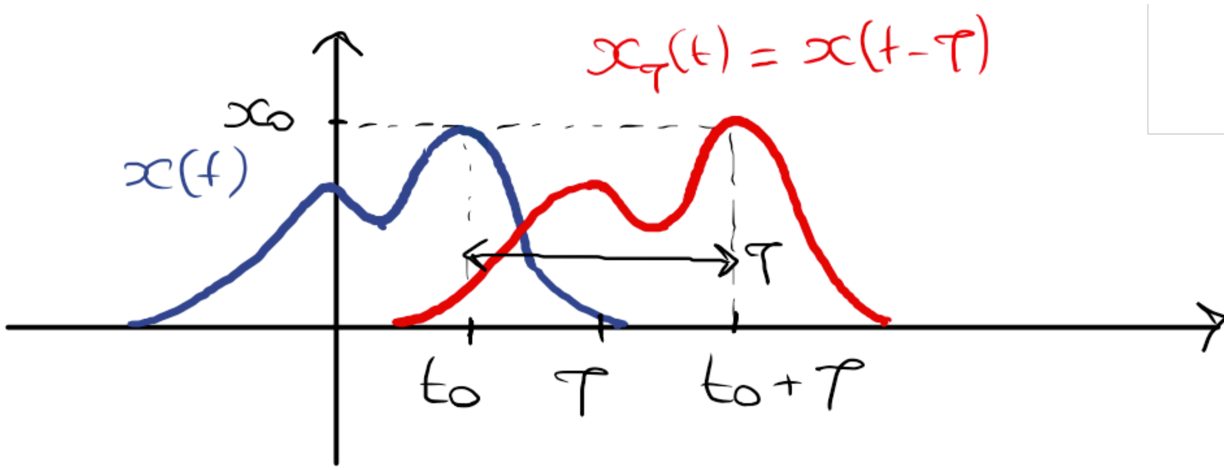


Figure 2.13: Signal $x(t-\tau)$ is $x(t)$ delayed into the future by τ time units.

$$
\begin{aligned}
x(t-\tau) = x_0 &\iff x-\tau = t_0 && \text{because } x(t_0) = x_0 \\
&\iff t = t_0 + \tau
\end{aligned}
$$

The function $t \mapsto x(t-\tau)$ obtains x_0 when $t = t_0 + \tau$, thus t_0 delayed τ into the future.

Another way to think of the problem: Which value does the red curve obtain at $t = t_0 + \tau$? Well, plug it in:

$$
\begin{aligned}
x(t-\tau) &= x((t_0+\tau) - tau) \\
&= x(t_0 + (\tau - \tau)) \\
&= x(t_0) \\
&= x_0.
\end{aligned}
$$

So $x(t-\tau)$ obtains a value of x_0 by looking *into the past* τ units of time, not into the future.

Similarly, $x(t+\tau)$ for $\tau > 0$ will translate (advance) the signal to the left, into the past.

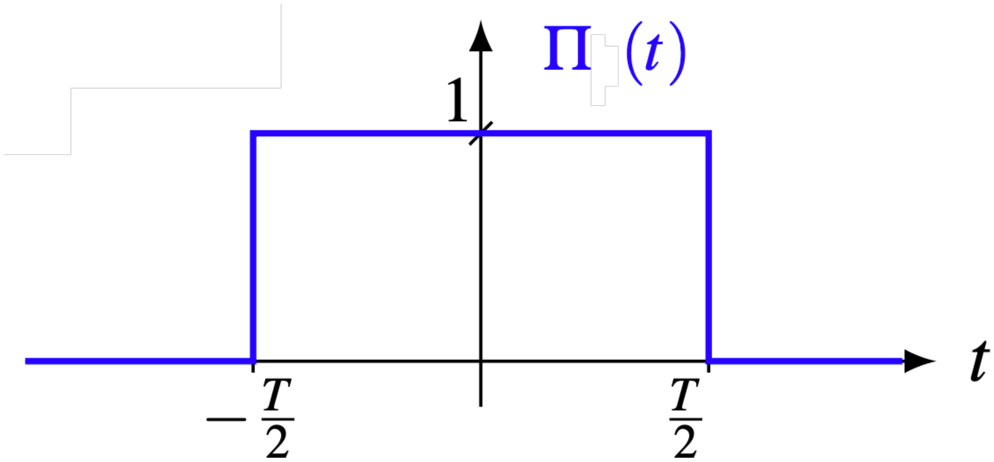


Figure 2.14: $\Pi(t)$ function from Equation (2.7)

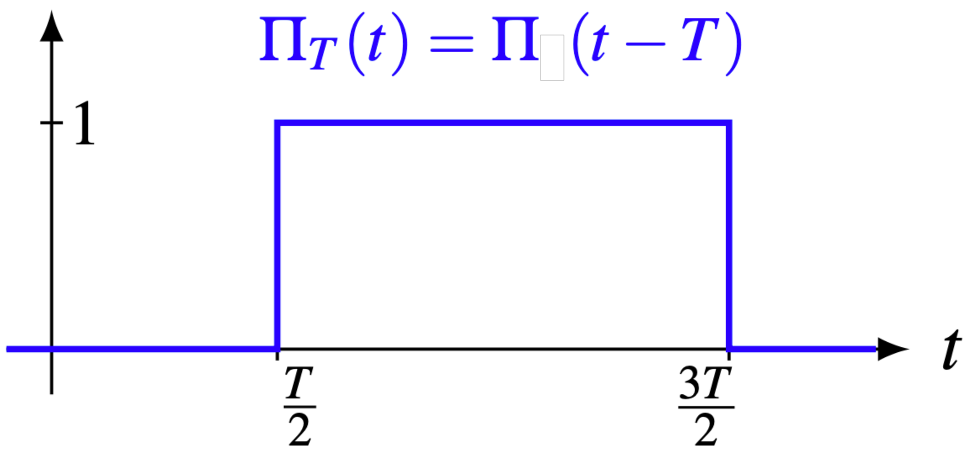


Figure 2.15: $\Pi_T(t)$ function from Equation (2.8)

2.12 Exercises: TD 2

We will consider two window functions:

$$\Pi(t) = \begin{cases} 1 & \text{if } -\frac{T}{2} \leq t \leq \frac{T}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

$$\Pi_T(t) = \begin{cases} 1 & \text{if } \frac{T}{2} \leq t \leq \frac{3T}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2.8}$$

I.e., Π_T is a window function (shown in Figure 2.14) centered at T. $\Pi_T(\tau)$ is a version of $\Pi(\tau)$ delayed T units of time into the future, (shown in Figure 2.15).

1. Compute and draw the graph of the autocorrelation of Π with itself: $\Gamma_{\Pi\,\Pi}(\tau)$. How is the result different from the convolution: $(\Pi * \Pi)(\tau)$?

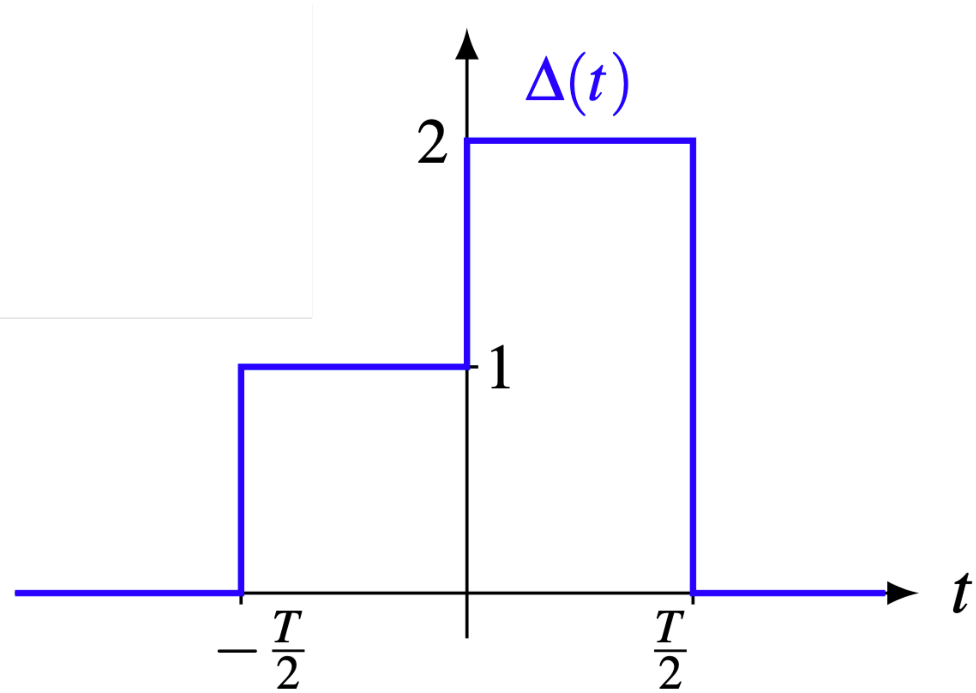


Figure 2.16: $\Delta(t)$ function from Equation (2.9)

2. Compute the intercorrelation $\Gamma_{\Pi\,\Pi_T}(\tau)$ and the convolution $(\Pi * \Pi_T)(\tau)$. Where (at which time) do the maximums of the autocorrelation and the convolution occur?

3. Now we consider the double window function, shown in Figure 2.16:

$$\Delta(t) = \begin{cases} 1 & \text{if } -\frac{T}{2} \leq t < 0 \\ 2 & \text{if } 0 \leq t \leq \frac{T}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

Compute and draw the graphs of the intercorrelation $\Gamma_{\Pi\,\Delta}(\tau)$ and the convolution $(\Pi * \Delta)(\tau)$. Where (at which time) do the maximums of the autocorrelation and the convolution occur?

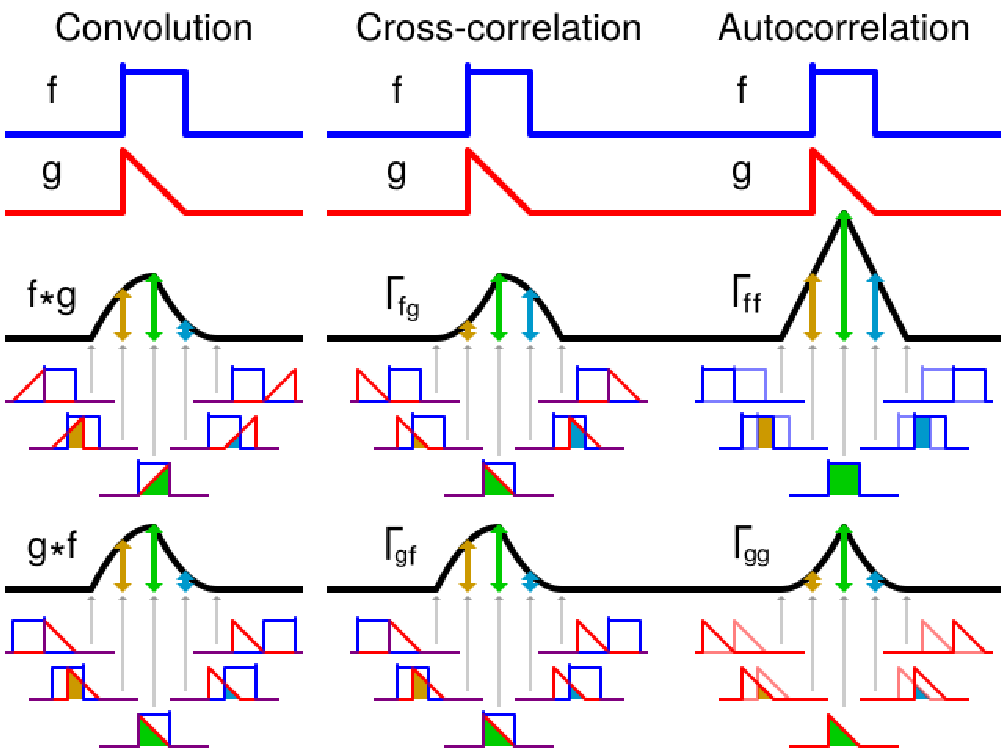


Figure 2.17: Comparison of convolution, cross-correlation, and autocorrelation

2.13 Review

1. Inner product

   - If $x \in \mathcal{L}^2(\mathbb{R})$, then

$$\langle x(t), y(t)\rangle = \int_{\mathbb{R}} x(t)\,\overline{y(t)}\ \mathrm{d}t$$

   - If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$, then

$$\begin{aligned}\langle x(t), y(t)\rangle &= \lim_{T\to+\infty} \frac{1}{T}\int_0^T x(t)\,\overline{y(t)}\ \mathrm{d}t\\ &= \lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)\,\overline{y(t)}\,dt\end{aligned}$$

   - If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ and x periodic with period T, then

$$\begin{aligned}\langle x(t), y(t)\rangle &= \frac{1}{T}\int_0^T x(t)\,\overline{y(t)}\ \mathrm{d}t\\ &= \frac{1}{T}\int_{-T/2}^{T/2} x(t)\,\overline{y(t)}\,dt\end{aligned}$$

2. Autocorrelation

$$\Gamma_{xx}(\tau) = \langle x(t), x(t-\tau) \rangle$$

- $x \in \mathcal{L}^2(\mathbb{R})$
- $\Gamma_{xx}(0) = E_x$.
- Γ_{xx} is an even function.
- Γ_{xx} measures the resemblance of x to a translated version of itself.
- If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ and x is T-periodic, so is Γ_{xx}. However, be careful to use the $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ inner product.

3. Cross-correlation

$$\Gamma_{xy}(\tau) = \langle x(t), y(t-\tau) \rangle$$

- $x \in \mathcal{L}^2(\mathbb{R})$.
- Γ_{xy} measures the resemblance of x to y.
- Not commutative: $\Gamma_{xy}(\tau) = \overline{\Gamma_{yx}(-\tau)}$
- If $x \in \mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ and x is T-periodic, so is Γ_{xx}. However, be careful to use the $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$ inner product.

4. Convolution

$$\begin{aligned}(x * y)(\tau) &= \left\langle x(t), \overline{y(\tau - t)} \right\rangle \\ &= \int_{\mathbb{R}} x(t) y(\tau - t)\,\mathrm{d}t\end{aligned}$$

- $x \in \mathcal{L}^1(\mathbb{R})$.
- Commutative: $(x * y) = (y * x)$
- Practical applications for operations of filtering.
- Weighted sliding average of a function seen through another
- Example computation, convolution of window function with itself.
- Connected with the cross-correlation:

$$\Gamma_{xy} = (x * \overline{y^{-}})$$

Chapter 3

Harmonic Analysis: Act I, Fourier Series

Chapter Objectives

- Approximate a periodic signal by a finite sum of trigonometric harmonics, and state Dirichlet's theorem giving conditions under which the series converges to the signal itself.
- Compute the real Fourier coefficients (a_0, a_n, b_n) and the complex coefficients (c_n) of a periodic signal, and convert between the two representations.
- Recognize the Gibbs phenomenon and explain why it does not contradict Dirichlet's theorem.
- Interpret the Fourier series geometrically as the decomposition of a signal onto an orthonormal basis of the vector space of periodic signals.
- Apply Parseval's equality to relate the energy of a signal in the time domain to the energy carried by its harmonics.

A brief history:

- ≈ 1750 : work on vibrating strings by d'Alembert, Euler, and Bernoulli
- between 1807 and 1822: work of Fourier on heat diffusion, the solution of heat diffusion by decomposing into trigonometric series.
- 1829: first formal result on the convergence of the Fourier series by Dirichlet (Dirichlet's theorem).

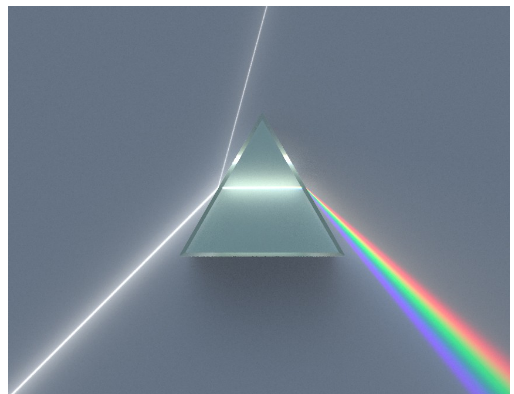
Starting point for Fourier: the principle of superposition:
In physics, a phenomenon of complex vibration can be solved by decomposing it into a linear superposition of phenomena, each a much simpler element, as illustrated in Figure 3.1.

3.1 Fourier series as finite approximation

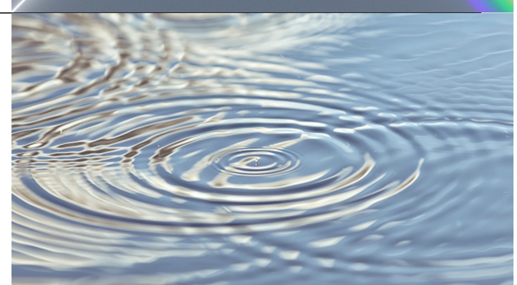
The general idea of Fourier was to approximate a wave (a T-periodic signal) $x : t \mapsto x(t)$ by a superposition of trigonometric waves of the form

$$A_n \sin\left(\frac{2\pi nt}{T} + \phi_n\right) \quad , n = 1, 2, \ldots, N\,.$$

Decomposition of white light into a superposition of monochromatic colors



Interference of wavelets on the surface of water.



Vibration of a guitar string

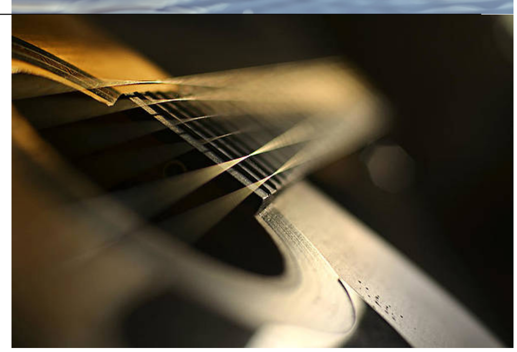
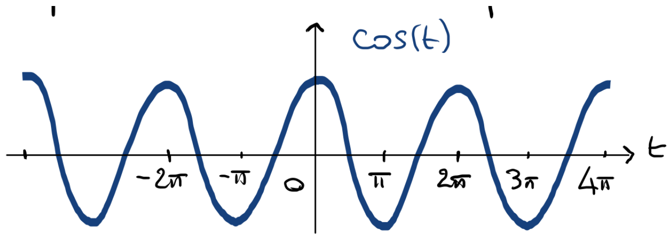


Figure 3.1: Superposition of waves



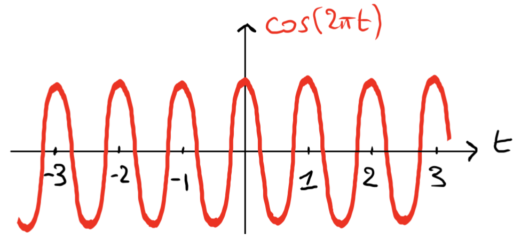


Figure 3.2: Periods of cos with t and $2\pi t$ as argument

Some vocabulary: For a T-periodic signal, we denote:

T	the period
$f = \frac{1}{T}$	the fundamental frequency
$\frac{n}{T} = nf$	the harmonic of order n

The illustration in Figure 3.2, hints at the following general facts.

$$\begin{array}{rcll} t \mapsto \sin(t+\phi) & \text{and} & t \mapsto \cos(t+\phi) & \text{are } 2\pi\text{-periodic.} \\ t \mapsto \sin(2\pi t+\phi) & \text{and} & t \mapsto \cos(2\pi t+\phi) & \text{are 1-periodic.} \\ t \mapsto \sin\left(\frac{2\pi}{T}t+\phi\right) & \text{and} & t \mapsto \cos\left(\frac{2\pi}{T}t+\phi\right) & \text{are T-periodic.} \end{array}$$

In each case, the period is independent of the phase, ϕ.

Returning to the idea of Fourier, we would like to approximate $x : t \mapsto x(t)$ by a superposition of N functions of the form

$$A_n \sin\left(\frac{2\pi n t}{T} + \phi_n\right) \quad , n = 1, 2, \ldots, N\,.$$

$$\begin{aligned} x(t) &\approx \sum_{n=1}^{N} A_n \sin\left(\frac{2\pi n}{T}\, t + \phi_n\right) && (3.1) \\ &= \sum_{n=1}^{N} A_n \left(\sin\left(\frac{2\pi n}{T}\, t\right)\cos(\phi_n) + \cos\left(\frac{2\pi n}{T}\, t\right)\sin(\phi_n)\right) && (3.2) \end{aligned}$$

If we denote:

$$\begin{aligned} a_n &= A_n \sin(\phi_n) \\ b_n &= A_n \cos(\phi_n)\,, \end{aligned}$$

then we can rewrite the approximation of $x(t)$ in Equation (3.2) as:

$$x(t) \approx \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi n}{T}\, t\right) + b_n \sin\left(\frac{2\pi n}{T}\, t\right)\right)\,. \tag{3.3}$$

The above treatment, including Equation (3.3), assumes the signal is biased at zero, as shown in Figure 3.2. What can we do if the average is non-zero for a period? In this case, we simply include an offset, which we will denote a_0:

$$x(t) \approx a_0 + \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi n}{T}\, t \right) + b_n \sin\left(\frac{2\pi n}{T}\, t \right) \right) .$$

Finally, to make the notation less verbose we can define

$$x_n(t) = a_n \cos\left(\frac{2\pi n}{T}\, t \right) + b_n \sin\left(\frac{2\pi n}{T}\, t \right) , \tag{3.4}$$

which simplifies the formula to

$$x(t) \approx a_0 + \sum_{n=1}^{N} x_n(t) . \tag{3.5}$$

Equation (3.5) is an order-N approximation of x; *i.e.*, it is the partial sum $S_N(t)$ of an infinite series of functions, and $x(t) - S_N(t)$ is the error of approximation of order N.

3.2 Fourier series as infinite series

How can we add a new term

$$x_{N+1} : t \mapsto a_{N+1} \cos\left(2\pi \frac{N+1}{T}\, t \right) + b_{N+1} \sin\left(2\pi \frac{N+1}{T}\, t \right)$$

onto $S_N(t)$? To do this we must determine the coefficients a_{N+1}, b_{N+1} such that x_{N+1} which best corrects the error of approximation.

In practice:

- In general we would need infinitely many terms in the series for the error of approximation to become zero, except in the special case that the signal is already a superposition of finitely many sinusoids.
- We have explicit formulas for explicitly computing the coefficients $\{a_n\}_{n\geq 0}$ and $\{b_n\}_{n\geq 1}$.
- We must, however, pay attention to the regularity of the signal, x, to assure it can be written as a Fourier series decomposition.
- All these considerations are captured in Dirichlet's theorem (Theorem 3.3.1).

A few words about the regularity.

3.2.1 Piecewise continuity

Reminder 3.2.1 (*piecewise continuity*)

Recall Definition 1.0.1 (page 12), *piecewise continuous, PWC* on an interval, $[a, b]$. A PWC function is allowed to be discontinuous at a finite number of points, but if x is discontinuous at $a < t_0 < b$ then we require that

$$\lim_{\substack{t \to t_0 \\ a < t < t_0 \leq b}} x(t) = x\left(t_0^-\right) ,$$

and

$$\lim_{\substack{t \to t_0 \\ a \leq t_0 < t < b}} x(t) = x\left(t_0^+\right) ,$$

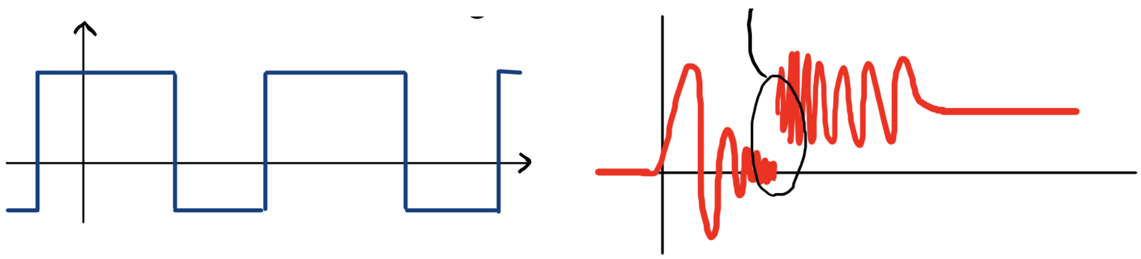


Figure 3.3: Illustrations of piecewise differentiability

but potentially

$$x(t_0^-) \neq x(t_0^+) \neq x(t_0)\,.$$

Definition 3.2.2 (*piecewise continuous signals*)

The set of piecewise continuous signals on $[a, b]$ will be denoted

$$\mathcal{C}^0_{pm}\left([a, b]\right)\,.$$

3.2.2 Piecewise differentiability

The notion of piecewise continuous extends to piecewise differentiable, which is exactly what we need for Dirichlet's theorem.

Definition 3.2.3 (*piecewise differentiable*)

A signal, $x : t \mapsto x(t)$ is said to be *piecewise differentiable* on the closed interval $[a, b]$, if it is piecewise continuous on $[a, b]$, and differentiable except possibly at finitely many points for which x has limits from the left and right.

Definition 3.2.4 (*piecewise differentiable signals*)

The set of piecewise differentiable signals on $[a, b]$ will be denoted

$$\mathcal{C}^1_{pm}\left([a, b]\right)\,.$$

To remember: piecewise differentiable signal

- may be discontinuous
- but despite the discontinuities is sufficiently well-behaved near those points of discontinuity that the derivative has left and right limits (potentially different limits from the left and right).

As examples, consider the functions illustrated in Figure 3.3. The square wave function (left) is piecewise differentiable. However, the function on the right (in Figure 3.3) is not piecewise differentiable. The illustration is intended to denote a function which oscillates more and more rapidly on both sides of the point of discontinuity, as if having some component such as $t \mapsto \sin(1/(t - t_0))$. Luckily we rarely encounter such badly behaved functions, and never in this course.

3.3 The Theorem of Dirichlet

Theorem 3.3.1 (*Dirichlet's theorem*)

Let $x : t \mapsto x(t)$ be a T-periodic signal in $\mathcal{C}^1_{pm}([0, T])$. At every point of continuity of x, the Fourier series converges to $x(t)$. If x is discontinuous at t, then the series converges to $\frac{1}{2}\left(x(t^-)+x(t^+)\right)$.

$$x(t)=a_0+\sum_{n=1}^{\infty}\left(a_n\cos\left(\frac{2\pi n}{T}\,t\right)+b_n\sin\left(\frac{2\pi n}{T}\,t\right)\right). \tag{3.6}$$

Moreover, the coefficients are given as follows:

$$a_0=\frac{1}{T}\int_0^T x(t)dt \tag{3.7}$$

$$a_n=\frac{2}{T}\int_0^T x(t)\cos\left(\frac{2\pi n}{T}\,t\right)dt \qquad \text{if } n>0 \tag{3.8}$$

$$b_n=\frac{2}{T}\int_0^T x(t)\sin\left(\frac{2\pi n}{T}\,t\right)dt \qquad \text{if } n>0 \tag{3.9}$$

Dirichlet's theorem may be formulated a bit more rigorously in saying that the sequence of partial sums,

$$S_N : t\mapsto a_0+\sum_{n=1}^{N}\left(a_n\cos\left(\frac{2\pi n}{T}\,t\right)+b_n\sin\left(\frac{2\pi n}{T}\,t\right)\right), \tag{3.10}$$

converges (simply as opposed to uniformly) to

$$\frac{1}{2}\left(x(t^-)+x(t^+)\right)$$

at every point, because

$$\frac{1}{2}\left(x(t^-)+x(t^+)\right)=x(t)$$

at every point of continuity.

Additionally, if $x\in\mathcal{C}^1_{pm}([0,T])$ is actually continuous on $[0,T]$, then the sequence of partial sums converges uniformly.

Anyway, Dirichlet's theorem is a major result because it gives (sufficient) conditions under which we effectively have equality between the signal and its Fourier series decomposition, and also the identity of the coefficients, $a_0, \{a_n, b_n\}_{n>0}$, which make it possible to realize this equality.

3.4 The Gibbs Phenomenon

But the theorem does not say what happens near the points of discontinuity. In effect, Equation (3.10) is continuous because it is a finite sum of continuous functions. But at a point of discontinuity of x, $S_N(t)$ is forced to be discontinuous in the limit (as $N\to\infty$). To manifest this behavior, there is a peculiar kind of oscillation near the point of discontinuity as we observe values of $S_N(t)$ for larger values of N. Figure 3.4 illustrates this behavior. This peculiar behavior is described by the Gibbs phenomenon.

Observation 1:
The amplitude of the overshoot converges to a (constant) value which depends on the height of the discontinuity. If x is discontinuous at t_0 and $\Delta x(t_0)=\left|x\left(t_0^-\right)-x\left(t_0^+\right)\right|$ is the height of the discontinuity, then the height of the overshoot of S_N near t_0 converges to $0.09\times\Delta x(t_0)$.

Example 3.4.1 (*Gibbs overshoot for a unit jump*)

For a signal discontinuity gap from 0 (left hand size) to 1 (right hand side), the amplitude of the discontinuity is 1. So the Gibbs phenomenon predicts a 1.09 in overshoot at the right hand side of the discontinuity and 0.09 as overshoot (in the negative direction) at the 0 left hand side.

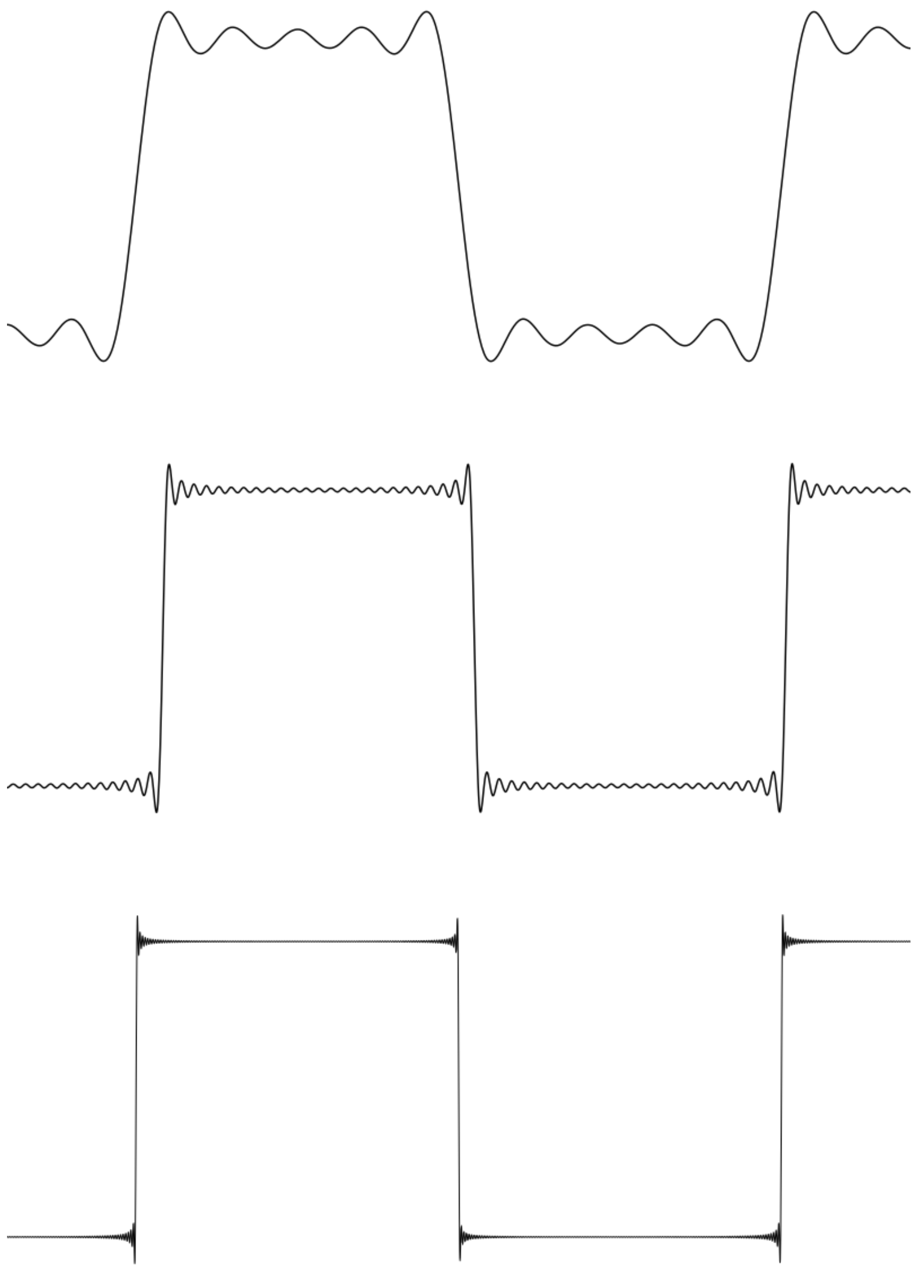


Figure 3.4: The Gibbs phenomenon. With higher order approximation, the overshoot moves horizontally closer to the point of discontinuity.

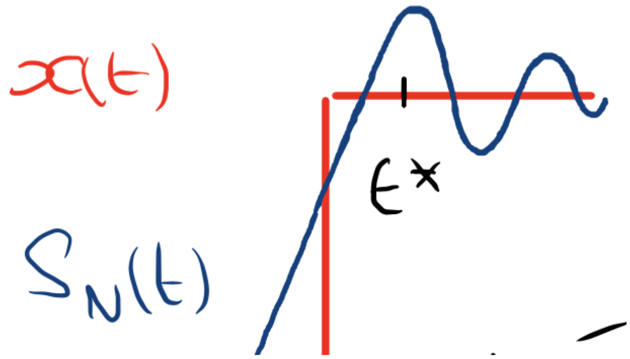


Figure 3.5: The time point of maximum overshoot moves toward the discontinuity as $N \to \infty$

The number 0.09 is actually called Gibbs constant, and its precise value is

$$G = \frac{1}{\pi} \lim_{a \to 0} \int_a^\pi \frac{\sin(t)}{t} dt - \frac{1}{2}$$

Observation 2:
How can the simple convergence of $S_N(t) \to \frac{1}{2}\left(x(t^-) + x(t^+)\right)$ be compatible with the Gibbs phenomenon?

In Figure 3.5, the time point of maximum overshoot is labeled t^*. Basically, as N grows larger, t^* moves closer to t_0. So for any $t \neq t_0$ close to the point of discontinuity, there is a sufficiently large N such that $t_0 < t^* < t$. This convergence to t_0 of the sequence of successive values of t^*, ensures that $S_N(t) \xrightarrow[N\to\infty]{simple} x(t)$.

The equality between $S_N(t)$ and $x(t)$ predicted by Dirichlet's theorem is only true in the limit, and not for any finite N.

3.5 Symmetry of the Fourier Series

If $x : t \mapsto x(t)$ is an odd function, then $a_n = 0 \quad \forall n > 0$. If $x : t \mapsto x(t)$ is an even function, then $b_n = 0 \quad \forall n > 0$. If x is odd, then in fact $a_0 = 0$ as well (it is the coefficient of the even function $\cos(0) = 1$). Even-ness of x, however, does not force $a_0 = 0$.

This fact is not hard to understand, as shown in Figure 3.6. The sequence, (a_n), are coefficients of even functions, $t \mapsto \cos\left(\frac{2\pi n}{T} t\right)$, thus their contribution is zero for reconstructing a signal which is an odd function.

Vice versa if x is an even function; $b_n = 0 \quad \forall n > 0$.

3.6 Complex Decomposition of the Fourier Series

The coefficients $a_0, (a_n)_{n=1}^\infty, (b_n)_{n=1}^\infty$ are called the coefficients of the real decomposition, because there also exists a complex decomposition.

Consider the Euler formulas for $\theta \in \mathbb{R}$ as illustrated in Figure 3.7:

$$e^{i\theta} = \cos\theta + i \sin\theta \tag{3.11}$$

$$\cos\theta = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right) \tag{3.12}$$

$$\sin\theta = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right) \tag{3.13}$$

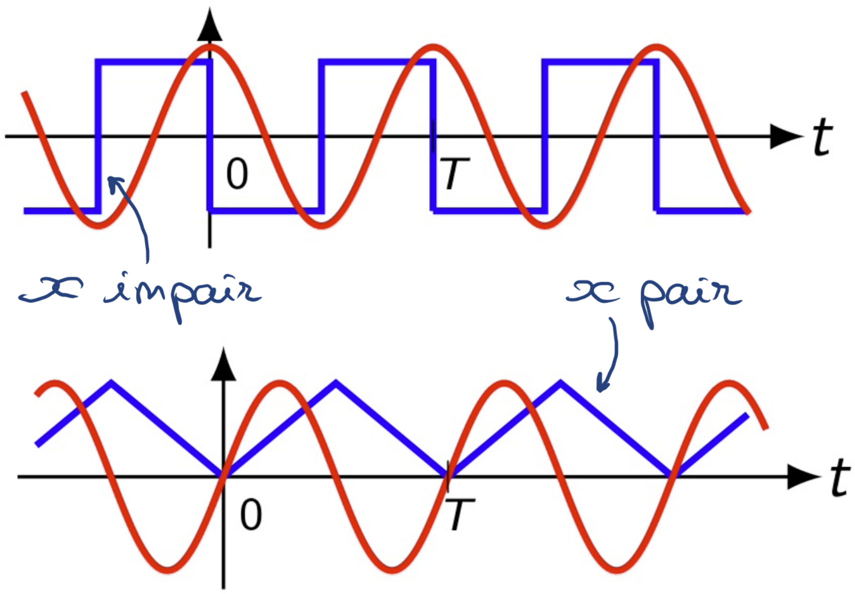


Figure 3.6: Odd and even contributions to the Fourier Series

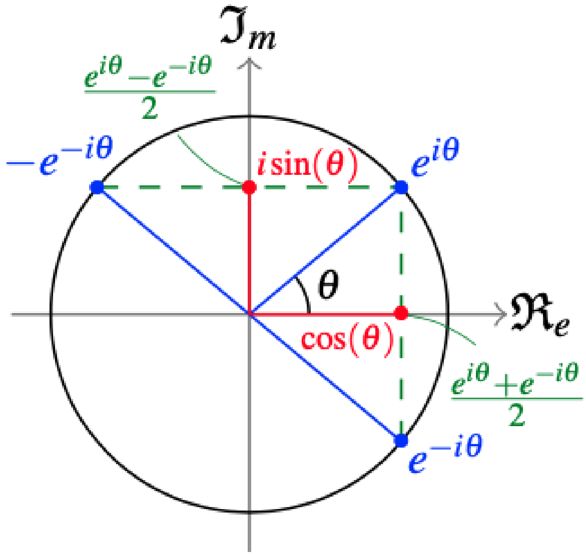


Figure 3.7: Euler identities for exponential, cos, and sin

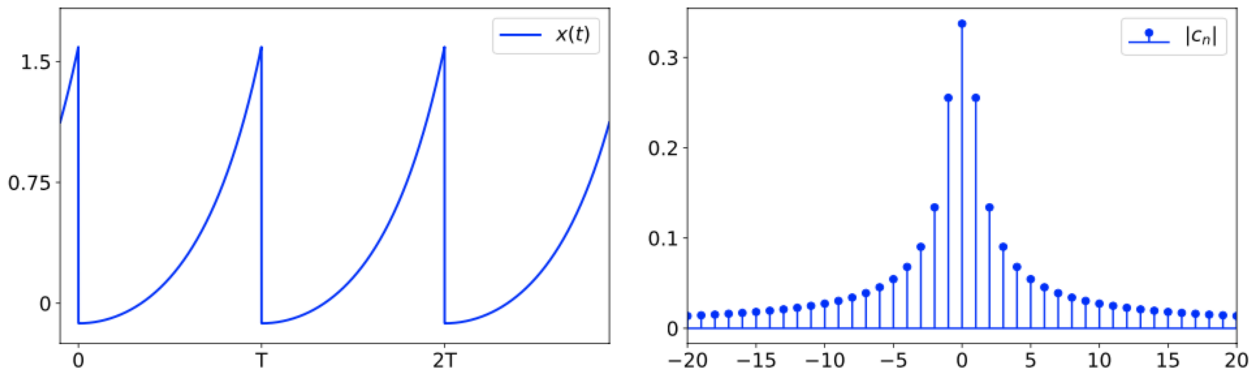


Figure 3.8: Signal and its spectrum

Using Equations (3.11), (3.12), and (3.13), in Dirichlet's theorem Equation 3.6, we can replace $\cos\left(\frac{2\pi n}{T}\, t\right)$ by $\frac{1}{2}\left(e^{i\theta}+e^{-i\theta}\right)$, and $\sin\left(\frac{2\pi n}{T}\, t\right)$ by $\frac{1}{2i}\left(e^{i\theta}-e^{-i\theta}\right)$, and a bit of algebra, we can arrive at the following complex decomposition of the Fourier series:

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{+\infty} c_n e^{\frac{2\pi i n}{T}\, t} && (3.14)\\ &= \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i n}{T}\, t\right) \\ c_n &= \frac{1}{T}\int_0^T x(t) e^{-\frac{2\pi i n}{T}\, t} dt && (3.15)\\ &= \frac{1}{T}\int_0^T x(t)\exp\left(-\frac{2\pi i n}{T}\, t\right) dt\,. \end{aligned}$$

The two decompositions (real and complex) are strictly equivalent. The hypotheses and conclusion of Dirichlet's theorem remain the same. So why bother with the complex notation in such a case?

The complex formulation (decomposition) ...

- ...gives us one single formula for the coefficients, $c_n \quad n \in \mathbb{Z}$, as opposed to three formula for a_0, a_n, and b_n. (We can easily verify that $c_0 = a_0$.)
- ...allows for a geometric interpretation of the Fourier series decomposition. See Section 3.8.
- ...prepares the ground for the Fourier transform.

3.7 Relations between Real and Complex Fourier Coefficients

$$\forall n \in \mathbb{N} \cup \{0\} \begin{cases} a_n &= c_n + c_{-n} \\ b_n &= i\left(c_n - c_{-n}\right) \end{cases} \iff \begin{cases} c_n &= \frac{1}{2}\left(a_n - i b_n\right) \\ c_{-n} &= \frac{1}{2}\left(a_n + i b_n\right) \end{cases} \tag{3.16}$$

If $x : t \mapsto x(t)$ is a real valued signal, then a_n and b_n are real, and consequently

$$\underbrace{c_n = \overline{c_{-n}}}_{\text{Hermitian symmetry}} \qquad \forall n > 0\,, \tag{3.17}$$

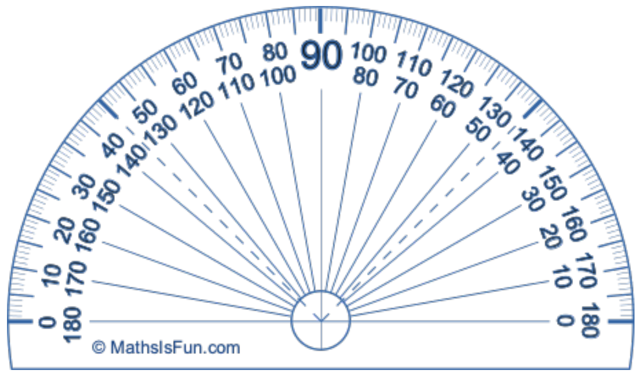


Figure 3.9: An inner product on a vector space makes it possible to measure angles.

in which case $\left(|c_n|_{n\in\mathbb{Z}}\right)$ is an even sequence:

$$|c_n| = |c_{-n}| \quad \forall n > 0\,.$$

Definition 3.7.1 (*spectrum*)

If $(c_n)_{n\in\mathbb{Z}}$ are the complex Fourier series coefficients of signal, x, then $(|c_n|)_{n\in\mathbb{Z}}$ is called the *spectrum* of signal x.

3.8 Geometric Interpretation

Just as in a finite dimensional vector space, any vector can be written uniquely as a (finite) linear combination of the basis elements, so can a periodic function (with certain reasonable restrictions) be written as an infinite series (infinite sum) of elements of a countably infinite set of normalized linearly independent signals. Of course, care must be taken to assure that the infinite series converges.

Definition 3.8.1 (*orthonormal*)

In a vector space, V, with inner product, $\langle \cdot, \cdot \rangle$, a set of vectors $\mathfrak{U} \subseteq V$ is said to be *orthonormal* if

- $0 \notin \mathfrak{U}$
- $u \in \mathfrak{U} \implies \|u\| = 1$ (the normal part)
- $\langle u, v \rangle = 0 \quad \forall u, v \in \mathfrak{U}, u \neq v$ (the *ortho* part).

Reminder 3.8.2 (*norm and angle from an inner product*)

In Definition 3.8.1 we mentioned $\|u\| = 1$ and $\langle u, v \rangle = 0$. Recall that given an inner product, we may define the norm of a vector and the angle between two vectors as:

$$\|u\| = \sqrt{\langle u, u \rangle}$$
$$\cos\theta = \frac{\langle u, v \rangle}{\|u\|\|v\|}$$

In $\mathbb{R}^n$ this norm and angle have an intuitive interpretation of an actual length and angle which can be measured with a ruler and protractor. However, in more abstract vector spaces they may not have an obvious interpretation. Nevertheless, every vector space equipped with an inner product (even an infinite dimensional vector space), necessarily has such a built-in *ruler* and *protractor*; thus we can speak about a vector being *normal* (unit length) and about two vectors being orthogonal, having a 90° angle between them. When

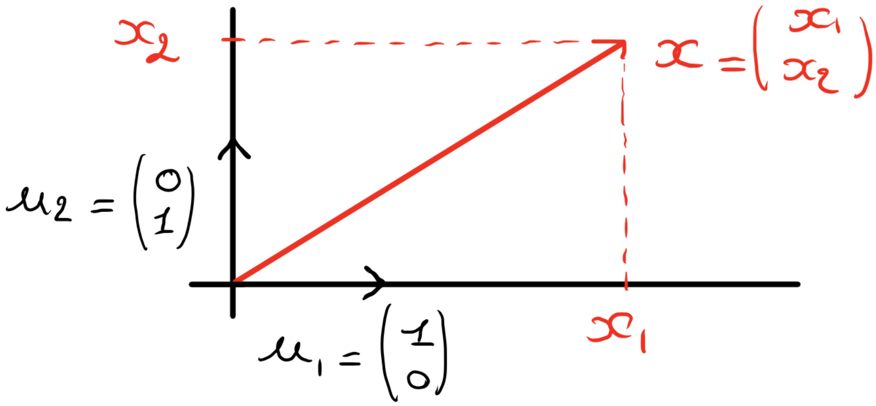


Figure 3.10: Vector decomposition with respect to 2D basis

constructing a vector space, we have a choice of different inner products to use, as long as the inner product satisfies Definition 1.4.1 (page 16). A different choice of inner product will lead to different length and angle measurements.

Definition 3.8.3 (*span*)

In a vector space, V, over field K, let $V' \subseteq V$. The *span* of V', *span*(V'), is the set of exactly all finite linear combinations of finite subsets of V'. *I.e.*, if $B = \{v_1, v_2, \ldots, v_m\} \subseteq V'$, and $A = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in K^m$ (*i.e.*, $\alpha_i \in K$ for $i = 1, 2, \ldots, m$), then $\sum_{i=1}^{m} \alpha_i v_i \in$ *span*(V').

Definition 3.8.4 (*basis*)

In a vector space, V, a subset B is called a *basis*, if $V =$ *span*(B), and every finite subset of $V' = \{v_1, v_2, \ldots, v_m\} \subset B$ is linearly independent,

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = 0 \implies \alpha_i = 0 \quad \forall i = 1 \ldots m\,.$$

The complex decomposition of a Fourier series, Equation (3.14), can be interpreted geometrically (somewhat abstractly) as a decomposition of a vector into coordinates with respect to an orthonormal basis. In this interpretation the coefficients, Equation (3.15), correspond to the coordinates of an infinite dimensional vector, x.

3.8.1 Decomposing a Vector in $\mathbb{R}^2$

As in Figure 3.10, let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$, and let

$$\mathcal{U} = (u_1, u_2) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

be an orthonormal basis of $\mathbb{R}^2$.

x may be written as a linear combination of u_1, u_2 as

$$\begin{aligned} x &= x_1 u_1 + x_2 u_2 \\ &= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

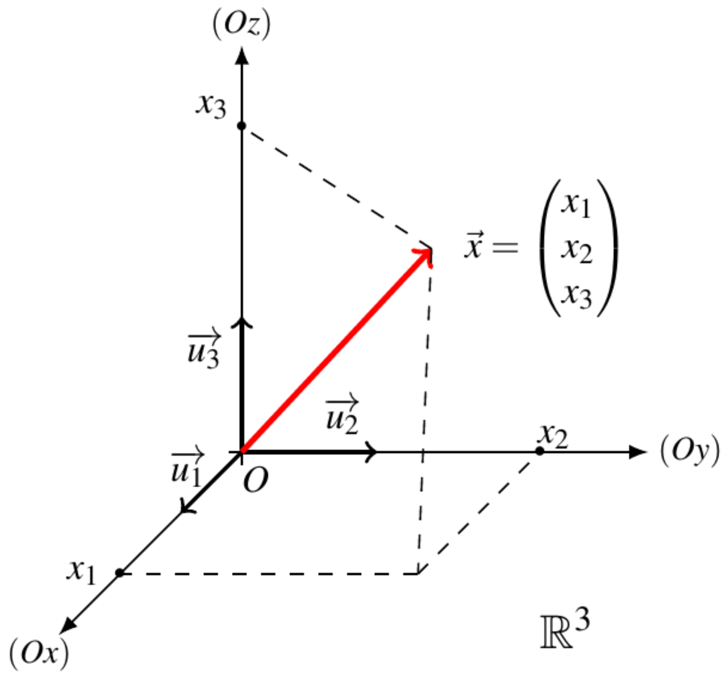


Figure 3.11: Euclidean Representation of Vector in $\mathbb{R}^3$

The coordinate, x_1, is the projection of vector, x, onto the line passing through u_1. Similar for x_2.

$$\begin{aligned} x_1 &= \langle x, u_1 \rangle \\ &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \\ x_2 &= \langle x, u_2 \rangle \\ &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle \end{aligned}$$

Finally, $x = x_1u_1 + x_2u_2$ may be written as $x = \langle x, u_1 \rangle u_1 + \langle x, u_2 \rangle u_2$.

The analogous process works for higher dimensions. Figure 3.11 illustrates this point for $\mathbb{R}^3$, but it will also work for $\mathbb{R}^n$.

Let $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, and let $\mathcal{U} = (u_1, u_2, \ldots, u_n)$ be an orthonormal basis for $\mathbb{R}^n$.

$$x = x_1u_1 + x_2u_2 + \cdots + x_nu_n \tag{3.18}$$

$$= \sum_{i=1}^{n} x_iu_i \tag{3.19}$$

$$= \sum_{i=1}^{n} \langle x, u_i \rangle u_i \,. \tag{3.20}$$

We would like to generalize Equation (3.20) to work with infinite dimensional vector spaces, but there are some obstacles. Definition 3.8.4 defines *basis* only for finite dimensional vector spaces. In Section 3.8.2, we will talk more about this problem.

3.8.2 Vector decomposition in an infinite dimensional vector space

If x is an element of an infinite dimensional vector space, we would like to write

$$x = \sum_{i=1}^{+\infty} \langle x, u_i \rangle \, u_i \,, \tag{3.21}$$

where $\mathcal{U} = (u_1, u_2, \ldots)$, and $\mathcal{U}$ is an orthonormal basis. However, Definition 3.8.4 defines *basis* only for finite dimensional vector spaces.

3.8.3 Convergence

Extending the definition of basis to infinite dimensional vector spaces is subtle. If we simply extend the Definition 3.8.3 to allow an infinite sum (as one might be tempted to do), then if we take a subset of V, say $\{v_i\}_{i\in\mathbb{Z}}$ and a sequence in $\mathbb{C}$, $(\alpha_i)_{i\in\mathbb{Z}}$, there is no guarantee that $\sum\limits_{i\in\mathbb{Z}} \alpha_i v_i$ even converges; in general it does not. As a simple example take $1 = \alpha_i, \forall i \in \mathbb{Z}$, in which case $\sum\limits_{i\in\mathbb{Z}} \alpha_i u_i$ definitely does not converge.

There are certainly sequences, $(\alpha_i)_{i\in\mathbb{Z}}$, for which $\sum\limits_{i\in\mathbb{Z}} \alpha_i u_i$ converges. We call the set of all such sequences, $\ell^2(\mathbb{C})$. As a trivial case, $\alpha_i = 0, \forall i \in \mathbb{Z}$, in which case $\sum\limits_{i\in\mathbb{Z}} \alpha_i u_i = \sum\limits_{i\in\mathbb{Z}} 0 u_i = 0$

In the coming sections, we wish to find a set, $\mathfrak{U}$, and show that every $x \in \mathcal{L}^2([0,T])$ corresponds to a sequence of complex numbers $(\alpha_i)_{i\in\mathbb{Z}} \in \ell^2(\mathbb{C})$ such that $x = \sum_{i\in\mathbb{Z}} \alpha_i u_i$. We would like to compute α_i.

The fact that given $(\alpha_i)_{i\in\mathbb{Z}} \in \ell^2(\mathbb{C})$, we have $\sum\limits_{i\in\mathbb{Z}} \alpha_i u_i$ converging to some $x \in \mathcal{L}^2([0,T])$ means there is a map $\ell^2(\mathbb{C}) \to \mathcal{L}^2([0,T])$. However, the mapping is neither injective nor surjective. The fact that given $x \in \mathcal{L}^2([0,T])$ we can compute $(\alpha_i)_{i\in\mathbb{Z}} \in \ell^2(\mathbb{C})$ (its Fourier series) does not mean the mapping is surjective, because two different functions $x_1, x_2 \in \mathcal{L}^2([0,T])$ may correspond to the same sequence from $\ell^2(\mathbb{C})$ if x_1 and x_2 differ in value (for example) at finitely many points. This means that $\ell^2(\mathbb{C})$ and $\mathcal{L}^2([0,T])$ fail to be isomorphic.

It is possible, albeit beyond the scope of this course, to partition $\mathcal{L}^2([0,T])$ into equivalence classes, and demonstrate that $\ell^2(\mathbb{C})$ is isomorphic to that quotient space of equivalence classes.

3.8.4 Constructing a Basis for $\mathcal{L}^2([0,T])$

We have already seen that non-zero periodic signals are not finite energy signals, so we cannot work in $\mathcal{L}^2(\mathbb{R})$ and with its inner product

$$\langle x, y \rangle = \int_{\mathbb{R}} x(t)\,\overline{y(t)}\,dt\,.$$

On the contrary, periodic signals are finite average power signals, whose average power is exactly the power over one period. We can, therefore, work in $\mathcal{L}^{\texttt{pm}}(\mathbb{R})$, and we may use its inner product

$$\begin{aligned}\langle x, y \rangle &= \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\,\overline{y(t)}\,dt \\ &= \frac{1}{T}\int_0^T x(t)\,\overline{y(t)}\,dt\end{aligned}$$

Notice that this puts us back in $\mathcal{L}^2([0,T])$ — the space of signals having finite energy over one period. We now have an inner product, we now need a basis, or something similar to a basis.

Definition 3.8.5 (***the complex exponential harmonics***)

Let $\mathfrak{U}([0,T])$ be defined as the following ordered infinite sequence of functions.

$$\mathfrak{U}([0,T]) = \left\{ u_n : t \mapsto u_n(t) = \exp\left(\frac{2\pi i n}{T}\, t\right) \right\}_{n\in\mathbb{Z}}$$

Lemma 3.8.6 (***natural-number powers of the exponential***)

If $t \in \mathbb{C}$ and $k \in \mathbb{N}$ then $e^{kt} = (e^t)^k$.

Notice that Lemma 3.8.6 is not true if $k \notin \mathbb{N}$. A counter example suffices:

$$e^{2i\pi t} \neq ((e^{i\pi})^2)^t = ((-1)^2)^t = 1^t = 1$$

Proof. By Induction on k. We will use without proof $e^{x+y} = e^x e^y$ and $e^x = (e^x)^1$.

Base case: $k = 0$

$$e^{0t} = e^0 = 1 = 1^t = (e^0)^t$$

Inductive case: $k > 0$ Suppose $e^{t(k-1)} = (e^t)^{k-1}$.

$$\begin{aligned} e^{tk} &= e^{t(k-1)} e^t \\ &= (e^t)^{k-1} e^t && \text{; by inductive hypothesis} \\ &= (e^t)^{k-1} (e^t)^1 \\ &= (e^t)^{k-1+1} \\ &= (e^t)^k \end{aligned}$$

Proposition 3.8.7 (***the complex exponentials are orthonormal***)

$\mathfrak{U}([0,T])$ is orthonormal in $\mathcal{L}^2([0,T])$.

Proof. Let u, v be elements of $\mathfrak{U}([0,T])$. Choose n, m such that

$$u = \exp\left(\frac{2\pi i n}{T}\, t\right)$$
$$v = \exp\left(\frac{2\pi i m}{T}\, t\right)$$

To satisfy Definition 3.8.1 we must show three conditions are satisfied.

First, we compute $\langle u, v\rangle$

$$
\begin{aligned}
\langle u, v\rangle &= \frac{1}{T}\int_0^T u(t)\,\overline{v(t)}\,dt \\
&= \frac{1}{T}\int_0^T \exp\left(\frac{2\pi i n}{T}\,t\right)\overline{\exp\left(\frac{2\pi i m}{T}\,t\right)}\,dt \\
&= \frac{1}{T}\int_0^T \exp\left(\frac{2\pi i n}{T}\,t\right)\exp\left(-\frac{2\pi i m}{T}\,t\right)dt \\
&= \frac{1}{T}\int_0^T \exp\left(\frac{2\pi i n}{T}\,t - \frac{2\pi i m}{T}\,t\right)dt \\
&= \frac{1}{T}\int_0^T \exp\left(\frac{2\pi i(n-m)}{T}\,t\right)dt
\end{aligned} \tag{3.22}
$$

Show $0 \notin \mathfrak{U}([0,T])$

The exponential function is never zero.

Show $u \in \mathfrak{U}([0,T]) \implies \|u\| = 1$

$$
\begin{aligned}
\|u\|^2 &= \langle u, u\rangle \\
&= \frac{1}{T}\int_0^T \exp\left(\frac{2\pi i(n-n)}{T}\,t\right)dt \quad ; \text{by (3.22)} \\
&= \frac{1}{T}\int_0^T \exp(0)dt \\
&= \frac{1}{T}\int_0^T 1dt = \frac{1}{T}T = 1\,.
\end{aligned}
$$

Thus $\|u\| = 1$.

Show $u \neq v \implies \langle u, v\rangle = 0$

Suppose $u \neq v$, so $n \neq m$. Since $\langle v, u\rangle = \overline{\langle u, v\rangle}$, proving $\langle u, v\rangle = 0$ in the case $n > m$ also settles the case $n < m$ (by swapping the roles of u and v), so without loss of generality assume $n > m$, so that $2(n-m) \in \mathbb{N}$ and Lemma 3.8.6 applies directly below.

$$
\begin{aligned}
\langle u, v\rangle &= \frac{1}{T}\int_0^T \exp\left(\frac{2\pi i(n-m)}{T}\,t\right)dt \quad ; \text{by (3.22)} \\
&= \frac{1}{T}\left[\frac{T}{2\pi i(n-m)}\exp\left(\frac{2\pi i(n-m)}{T}\,t\right)\right]_0^T \\
&= \frac{1}{2\pi i(n-m)}\left[\exp\left(\frac{2\pi i(n-m)T}{T}\right) - \exp\left(\frac{2\pi i(n-m)\cdot 0}{T}\right)\right] \\
&= \frac{e^{2\pi i(n-m)} - e^0}{2\pi i(n-m)} = \frac{(e^{i\pi})^{2(n-m)} - 1}{2\pi i(n-m)} \quad \text{by Lemma 3.8.6} \\
&= \frac{(-1)^{2(n-m)} - 1}{2\pi i(n-m)} = \frac{1-1}{2\pi i(n-m)} = \frac{0}{2\pi i(n-m)} = 0
\end{aligned}
$$

The three axioms of Definition 3.8.1 are satisfied, so we can conclude that $\mathfrak{U}([0,T])$ is orthonormal.

In resume, for $\mathcal{L}^2([0,T])$ we now have

- an inner product: $\langle x, y\rangle = \frac{1}{T}\int_0^T x(t)\,\overline{y(t)}\,dt$
- an orthonormal set:

$$\mathfrak{U}([0,T]) = \left\{u_n : t \mapsto u_n(t) = \exp\left(\frac{2\pi i n}{T}\,t\right)\right\}_{n\in\mathbb{Z}} .$$

We want to ask now whether we can extend Equation (3.20) (page 80), where $x \in \mathbb{R}^n$:

$$x = \sum_{i=1}^{n} \langle x, u_i\rangle\, u_i ,$$

to $\mathcal{L}^2([0,T])$ as

$$x(t) = \sum_{n\in\mathbb{Z}} \langle x, u_n\rangle\, u_n(t) .$$

We will assert Proposition 3.8.8 without proof.

Proposition 3.8.8 ($\mathfrak{U}([0,T])$ *is a basis for* $\mathcal{L}^2([0,T])$)

Let $x \in \mathcal{L}^2([0,T])$,

$$\begin{aligned} x &= \sum_{u\in\,\mathfrak{U}([0,T])} \langle x, u\rangle\, u \\ &= \lim_{n\to+\infty} \sum_{j=-n}^{n} \langle x, u_j\rangle\, u_j \end{aligned}$$

The proof of Proposition 3.8.8 would involve showing that $\mathcal{L}^2([0,T])$ is a *separable Hilbert space* for which $\mathfrak{U}([0,T])$ is an orthonormal Schauder basis. Obviously such a proof would be well beyond the scope of this course, and is better treated in a course in Functional analysis. We refer the reader to the lecture notes of Frederic Schuller *Lecture 03 - Separable Hilbert spaces (Schuller's Lectures on Quantum Theory)*[1].

If we accept Proposition 3.8.8 as true, then we can express $x(t)$ as follows:

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{+\infty} \langle x, u_n\rangle\, u_n(t) && \text{with } u_n : t \mapsto \exp\left(\frac{2\pi i n}{T}\,t\right) \\ &= \sum_{n=-\infty}^{+\infty} \langle x, u_n\rangle \exp\left(\frac{2\pi i n}{T}\,t\right) . && \end{aligned} \tag{3.23}$$

Examining the inner product:

$$\begin{aligned} \langle x, u_n\rangle &= \frac{1}{T}\int_0^T x(t)\,\overline{u_n(t)}\,dt \\ &= \frac{1}{T}\int_0^T x(t)\exp\left(-\frac{2\pi i n}{T}\,t\right) \\ &= c_n && \text{from Equation (3.15)} \end{aligned} \tag{3.24}$$

We can now substitute Equation (3.24) into Equation 3.23 to obtain (recover) Dirichlet's Theorem:

$$x(t) = \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i n}{T}\,t\right) . \tag{3.25}$$

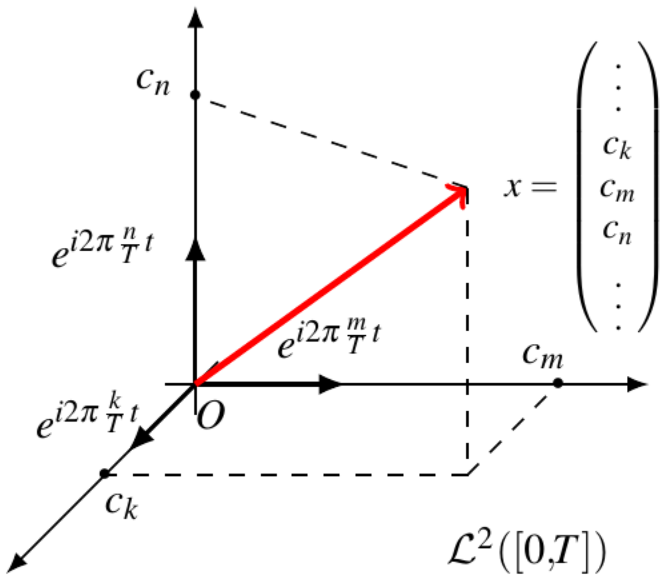


Figure 3.12: Euclidean Representation of Fourier Series

So $c_n = \frac{1}{T}\int_0^T x(t)\exp\bigl(-\frac{2\pi i n}{T}\,t\bigr)$ is the projection of signal, x, onto the *line* through $u_n : t \mapsto \exp\bigl(\frac{2\pi i n}{T}\,t\bigr)$. *I.e.*, c_n is the n^{th} coordinate of x with respect to $\mathfrak{U}([0,T])$.

Figure 3.12 illustrates that the complex Fourier coefficients $(c_n)_{n\in\mathbb{Z}}$ of a signal x can be interpreted geometrically as the infinite dimensional coordinates of a vector relative to a (Schauder) basis $\mathfrak{U}([0,T])$.

3.9 Parseval's Equality

For the signals $\mathcal{L}^2(\mathbb{R})$, we defined the energy of a signal, x as

$$E_x = \langle x, x\rangle = \|x\|^2 = \int_{\mathbb{R}} |x(t)|^2 dt\,.$$

In like manner, we can define the energy of a T-periodic signal $x \in \mathcal{L}^2([0,T])$ with the adapted inner product.

$$E_x = \|x\|^2 = \frac{1}{T}\int_0^T |x(t)|^2 dt \tag{3.26}$$

Equation (3.26) identifies the energy of a signal in the *time domain*. Equation (3.26) happens to equal the average power of x as calculated by the inner product of $\mathcal{L}^{\mathtt{pm}}(\mathbb{R})$.

If we apply Dirichlet's theorem

$$x(t) = \sum_{n=-\infty}^{+\infty} c_n \exp\biggl(\frac{2\pi i n}{T}\,t\biggr) \quad \forall n \in \mathbb{Z}\,,$$

we get:

[1] `https://mathswithphysics.blogspot.com/2016/07/frederic-schullers-lectures-on-quantum.html`

- x is the superposition of infinitely many harmonics:

$$x_n : t \mapsto c_n \exp\left(\frac{2\pi i n}{T} t\right) \quad \forall n \in \mathbb{Z} \tag{3.27}$$

- The time domain energy of the n^{th} harmonic is therefore

$$\begin{aligned}
E_{x_n} &= \frac{1}{T}\int_0^T |x_n(t)|^2 dt \\
&= \frac{1}{T}\int_0^T \left|c_n \exp\left(\frac{2\pi i n}{T} t\right)\right|^2 dt && \text{by Eq. (3.27)} \\
&= |c_n|^2 \frac{1}{T}\int_0^T \underbrace{\left|\exp\left(\frac{2\pi i n}{T} t\right)\right|^2}_{=1} dt \\
&= |c_n|^2 \frac{1}{T}\int_0^T 1^2 dt \\
&= |c_n|^2 \frac{1}{T}\,[t]_0^T = |c_n|^2 \frac{T-0}{T} \\
E_{x_n} &= |c_n|^2
\end{aligned}$$

- The energy for all the harmonics combined is thus $\sum_{n=-\infty}^{+\infty} |c_n|^2$.

Theorem 3.9.1 (*Parseval's equality*)

The energy of a signal in the time domain is equal to its energy in the frequency domain.

$$\begin{aligned}
E_x &= \frac{1}{T}\int_0^T |x(t)|^2 dt \\
&= \sum_{n=-\infty}^{+\infty} |c_n|^2 \\
&= |a_0|^2 + \frac{1}{2}\sum_{n=1}^{+\infty}\left(|a_n|^2 + |b_n|^2\right)
\end{aligned}$$

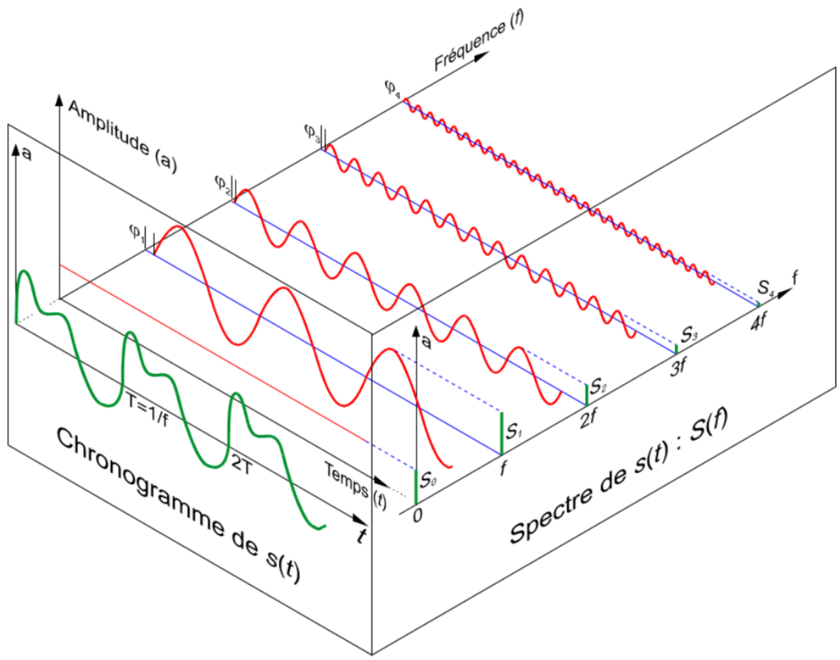


Figure 3.13: Time and Frequency Domains of Fourier Series

Proof.

$$
\begin{aligned}
E_x &= \frac{1}{T}\int_0^T |x(t)|^2 dt \\
&= \frac{1}{T}\int_0^T \left|\sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i n}{T}\,t\right)\right|^2 dt \\
&= \frac{1}{T}\int_0^T \sum_{n=-\infty}^{+\infty} \left|c_n \exp\left(\frac{2\pi i n}{T}\,t\right)\right|^2 dt && \text{; by Pythagoras} \\
&= \frac{1}{T}\int_0^T \sum_{n=-\infty}^{+\infty} |c_n|^2 \underbrace{\left|\exp\left(\frac{2\pi i n}{T}\,t\right)\right|^2}_{=1} dt \\
&= \frac{1}{T}\int_0^T \sum_{n=-\infty}^{+\infty} |c_n|^2 dt \\
&= \sum_{n=-\infty}^{+\infty} |c_n|^2 \frac{1}{T} \underbrace{\int_0^T dt}_{=T} \\
&= \sum_{n=-\infty}^{+\infty} |c_n|^2
\end{aligned}
$$

The interpretation of Parseval's equality (Theorem 3.9.1) is that as far as the energy of a signal is concerned, the temporal representation and the spectral representation are equivalent. In this sense the graph (time domain) and the spectrum (frequency domain) represent the same information, *i.e.*, the signal. This fact is alluded to in Figure 3.13.

3.10 Asymptotic Behavior

We have seen that the smoother (more regular) a signal is (example continuous rather than piecewise continuous), the faster the sequence of coefficients shrink (in absolute value). We now mathematically formalize this observation.

- For a signal in $\mathcal{C}^1_{pm}$ (so that Dirichlet's theorem applies) but which has discontinuities (*e.g.*, square wave), we have
$$|c_n| \to \frac{1}{n} \quad \text{as } n \to +\infty$$
This is the minimal asymptotic decrease possible, and is the same for all discontinuous signals.

- More generally: if x is at least k times differentiable, and if the k^{th} derivative $x^{[k]}$ is the first to have discontinuities, then
$$|c_n| \to \frac{1}{n^{k+1}} \quad \text{as } n \to +\infty$$

Example 3.10.1 (*Fourier coefficient decay for a triangle wave*)

If x is a continuous signal but has *angular* points such as a triangle wave, then the first derivative x' is discontinuous at those points. Since $k = 1$, we have

$$|c_n| \to \frac{1}{n^2} \quad \text{as } n \to +\infty$$

For a Fourier series, the smoother and more regular the signal is, the less energy is contained in high order harmonics; *i.e.*, the smaller is $|c_n|$ for large n.

3.11 Review

- The Fourier series makes it possible to decompose a complex periodic signal into a sum of *simple* trigonometric functions.

- Dirichlet's theorem
$$\begin{aligned} x(t) &= a_0 + \sum_{n=1}^{\infty}\left(a_n \cos\left(\frac{2\pi n}{T}\,t\right) + b_n \sin\left(\frac{2\pi n}{T}\,t\right)\right) \\ &= \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i n}{T}\,t\right) \end{aligned}$$

The coefficients are as follows:

$$\begin{aligned}
a_0 &= \frac{1}{T}\int_0^T x(t)dt && \text{average value for a period}\\
a_n &= \frac{2}{T}\int_0^T x(t)\cos\left(\frac{2\pi n}{T}\,t\right)dt && \text{if } n>0\\
b_n &= \frac{2}{T}\int_0^T x(t)\sin\left(\frac{2\pi n}{T}\,t\right)dt && \text{if } n>0\\
c_n &= \frac{1}{T}\int_0^T x(t)\exp\left(-\frac{2\pi i n}{T}\,t\right)dt && n\in\mathbb{Z}
\end{aligned}$$

- If x is discontinuous at t
  - the series for $x(t)$ converges to
  $$\frac{1}{2}\left(x(t^-)+x(t^+)\right)$$
  - Gibbs phenomenon
  - $|c_n| \to \frac{1}{n}$ as $n\to\infty$, if x is discontinuous.
  - $|c_n| \to \frac{1}{n^{k+1}}$ as $n\to\infty$, if x is k times differentiable and $x^{[k]} = \frac{d^k}{dt^k}$ is discontinuous.
- If x is an even function, then $b_n = 0 \quad \forall n>0$
- If x is an odd function, then $a_n = 0 \quad \forall n>0$
- If x is real valued, then $c_{-n} = \overline{c_n}$, and $|c_n| = |c_{-n}|$
- Geometrical interpretation

$$\begin{aligned}
x(t) &= \sum_{n\in\mathbb{Z}} c_n \exp\left(\frac{2\pi i n}{T}\,t\right)\\
&= \sum_{n\in\mathbb{Z}} \langle x, u_n\rangle\, u_n(t)
\end{aligned}$$

c_n is the n^{th} coordinates of x in the basis

$$\begin{aligned}
\mathfrak{U}([0,T]) &= \left\{u_n : t \mapsto \exp\left(\frac{2\pi i n}{T}\,t\right)\right\}_{n\in\mathbb{Z}}\\
\langle x, u_n\rangle &= \frac{1}{T}\int_0^T x(t)\overline{u_n}\,dt\\
&= \frac{1}{T}\int_0^T x(t)\exp\left(-\frac{2\pi i n}{T}\,t\right)dt\\
&= c_n
\end{aligned}$$

- Parseval's equality

$$\begin{aligned}
E_x &= \frac{1}{T}\int_0^T |x(t)|^2 dt\\
&= \sum_{n=-\infty}^{+\infty} |c_n|^2\\
&= a_0^2 + \frac{1}{2}\sum_{n=1}^{+\infty}\left(|a_n|^2 + |b_n|^2\right)
\end{aligned}$$

The time domain and frequency domain representations are equivalent.

Chapter 4

Harmonic Analysis: Act II, Fourier Transform

Chapter Objectives

- Define the Fourier transform and its inverse, and state a sufficient condition ($x \in \mathcal{L}^1(\mathbb{R})$) for the transform to exist.
- Compute Fourier transforms of standard signals (window/gate, sinc, Dirac delta, $\cos$, $\sin$) using the definition together with linearity, time-shifting, and symmetry properties.
- Use the Dirac delta function and the Sifting theorem to simplify integrals and to compute Fourier transforms indirectly (e.g., $\mathcal{F}(1)$).
- Apply Plancherel's theorem to relate convolution and multiplication of signals in the time and frequency domains, and use it to connect the Fourier transform to cross-correlation.
- Apply Parseval's identity to relate the energy of a signal in the time and frequency domains.
- State Bernstein's theorem and interpret how the bandwidth of a signal bounds its rate of change.

The Fourier transform is undoubtedly one of the most powerful tools in signal processing (filtering, noise cleaning, voice recognition, digital transmission, etc.), but equally in other domains which are not obviously connected (optics, quantum theory, control theory, arbitrary precision integer arithmetic).

4.1 From Fourier Series to Fourier Transform

The Fourier series permits the decomposition of a large class of periodic signals into a superposition of trigonometric functions of quantified frequencies, $\left\{\frac{n}{T}\right\}_{n\in\mathbb{Z}}$ multiples of the fundamental frequency, $\frac{1}{T}$, of the given signal. The amplitude of the harmonics expressed in terms of coefficients a_n and b_n (if the decomposition is in terms of sin and cos) or in terms of coefficients c_n (if the decomposition is in terms of exponentials), gives a *weight* to each harmonic in the reconstruction of the signal, and makes up its spectrum.

What can we do for a signal which is not periodic? Can't we simply proceed with its harmonic analysis?

The Fourier transform generalizes periodic signals to non-periodic signals, by generalizing the tools introduced in the scope of the Fourier series. The two notions are intimately connected.

Figure 4.1: Joseph Fourier

To see the connection, we look again at the expression given by Dirichlet's theorem:

$$x(t) = \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i\, n}{T} t\right).$$

- The coefficient, c_n of the n^{th} harmonic is associated with the discrete frequency $\nu_n = \frac{n}{T}$
- The coefficient, c_{n+1}, (the successive harmonic) is thus associated with the frequency $\nu_{n+1} = \frac{n+1}{T}$
- The *gap*, in radians, between two consecutive harmonics is thus

$$\Delta\nu = \nu_{n+1} - \nu_n = \frac{n+1}{T} - \frac{n}{T} = \frac{1}{T}.$$

- $\Delta\nu \to 0$ as $T \to +\infty$. This means the gap between successive harmonic frequencies decreases as the period of the signal increases. We can imagine that the difference between the frequencies becomes zero at $T = +\infty$. *I.e.*, to have a continuum of frequencies (without gaps) we need an infinite period as alluded to in Figure 4.2.

An aperiodic signal can be interpreted as a signal with infinite period, $T = +\infty$, which is never repeated. The Fourier series decomposition for such a signal gives harmonics whose frequencies vary in a continuous manner. This is the Fourier transform.

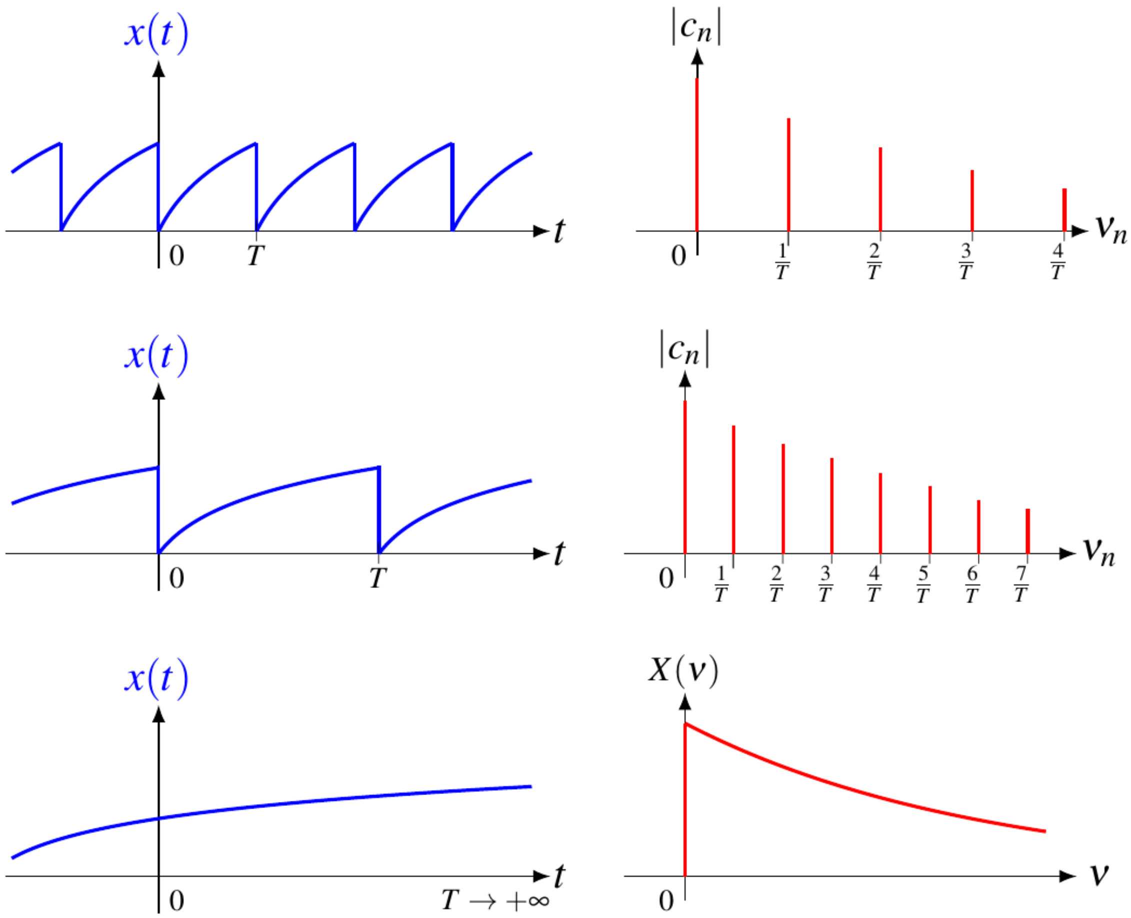


Figure 4.2: Continuum of Frequencies

4.2 The Fourier Transform

Definition 4.2.1 (*Fourier transform*)

The *Fourier transform* of a signal $x : t \mapsto x(t)$ is defined (in case it exists) by the function $X : \mathbb{R} \to \mathbb{C}$ such that

$$X(\nu) = \int_{\mathbb{R}} x(t) \exp(-2\pi i \nu t) dt$$

A comment about notation. The Fourier transform of a signal of one variable is a function of one variable. However, the Fourier transform operator itself, $\mathcal{F}$, is a function which operates on functions, mapping them to functions. We may say $\mathcal{F}(x) = X$, meaning that $\mathcal{F}$ transforms the signal, x, into the signal X, where x and X are both functions: $x : \mathbb{R} \to \mathbb{C}$ and $X : \mathbb{R} \to \mathbb{C}$. *I.e.*, x and X sit in the same function space. However, it is idiomatic to use different input variable names for the two functions. In this treatment we'll specify them as $x(t)$ and $X(\nu)$. We will also denote $\mathcal{F}(x(t)) = X(\nu)$.

Remember that the Fourier transform of a signal is a complex ($\mathbb{C}$) valued function of a real ($\mathbb{R}$) variable, even if the original signal is real valued.

In Definition 4.2.1, ν, the frequency, is analogous to the discrete frequency, $\nu_n = \frac{n}{T}$, of the n^{th} harmonic of the Fourier series, while $X(\nu)$ corresponds to the amplitude of the harmonic at frequency ν, and is analogous to c_n.

There is a bit of familiarity:

$$c_n = \frac{1}{T} \int_0^T x(t) \exp\Big(-2\pi i \frac{n}{T} t\Big) dt \tag{4.1}$$

$$= \frac{1}{T} \int_{-T/2}^{T/2} x(t) \exp\Big(-2\pi i \frac{n}{T} t\Big) dt \tag{4.2}$$

$$X(\nu) = \int_{-\infty}^{\infty} x(t) \exp(-2\pi i \nu t) dt$$

As shown in Lemma 1.13.2, we can pass from (4.1) to (4.2) because the actual limits of integration are not important as long as they differ by T.

The Fourier transform corresponds to the Fourier series decomposition where the harmonics vary in a continuous manner.

In engineer's notation we say:

$$x(t) \overset{\mathcal{F}/\mathcal{F}^{-1}}{\longleftrightarrow} X(\nu)\,.$$

However, mathematicians usually prefer a different notation,

$$x(t) \overset{\mathcal{F}/\mathcal{F}^{-1}}{\longleftrightarrow} \hat{x}(f)\,.$$

In either case, we use $\mathcal{F}$ to denote the Fourier transform operator, $\mathcal{F}(x)$, and $\mathcal{F}^{-1}(X)$ for its inverse.

Definition 4.2.2 (*kernel*)

If the Fourier transform of x is expressed as

$$X(\nu) = \int_{\mathbb{R}} x(t) \exp(-2\pi i \nu t) dt$$

then $\exp(-2\pi i \nu t)$ is called the *kernel* of the transform defined in frequency.

The kernel could also be expressed in pulsation as $e^{-i\omega t}$, but this would impact the definition of the inverse transform.

4.3 Sufficient Condition of Existence

Under which conditions does the Fourier transform exist? Theorem 4.3.1 argues a sufficient but not necessary condition for the existence of the Fourier transform.

Theorem 4.3.1 (*Sufficient condition for existence*)
If signal x is integrable, then the Fourier transform exists.

$$x \in \mathcal{L}^1(\mathbb{R}) \implies X = \mathcal{F}(x) \text{ exists}$$

We prove Theorem 4.3.1 by showing that $|X(\nu)| < +\infty$.

Proof. Let $x \in \mathcal{L}^1(\mathbb{R})$, and let $\int_{\mathbb{R}} |x(t)|dt = M < +\infty$.

$$\begin{aligned}
|X(\nu)| &= \left| \int_{\mathbb{R}} x(t) \exp(-2\pi i \nu t) dt \right| \\
&\leq \int_{\mathbb{R}} |x(t) \exp(-2\pi i \nu t)| dt \\
&= \int_{\mathbb{R}} |x(t)| \underbrace{|\exp(-2\pi i \nu t)|}_{=1} dt \\
&= \int_{\mathbb{R}} |x(t)| dt \\
&= M < +\infty
\end{aligned}$$

However, the converse of Theorem 4.3.1 is not true. There are functions outside of $\mathcal{L}^1(\mathbb{R})$ which have a Fourier transform. An example of such a function is the sinc function discussed in Section 4.5. sinc $\notin \mathcal{L}^1(\mathbb{R})$; nevertheless $\mathcal{F}(\text{sinc}) \in \mathcal{L}^1(\mathbb{R})$. For discussion of the Fourier transform of the sinc function and its inverse transform, see Section 4.6 and Figures 4.11 (page 112) and 4.14 (page 134)

4.4 Inverse Fourier Transform

Definition 4.4.1 (*Inverse Fourier transform*)
If x has a Fourier transform

$$X : \nu \mapsto \int_{\mathbb{R}} x(t) e^{-2\pi i \nu t} dt$$

and if $X \in \mathcal{L}^1(\mathbb{R})$, then

$$x(t) = \int_{\mathbb{R}} X(\nu) e^{2\pi i \nu t} d\nu$$

For the inverse Fourier transform, we integrate with respect to ν rather than t, and with a $+$ on the kernel $e^{2\pi i \nu t}$.

$$x \in \mathcal{L}^1(\mathbb{R}) \implies X \text{ exists}$$

but

$$\not\Longrightarrow X \in \mathcal{L}^1(\mathbb{R})\,!$$

Again we see some similarity to Dirichlet's theorem:

$$x(t) = \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i\, n}{T} t\right).$$

We see that in fact, x **is a continuous combination of harmonics.**

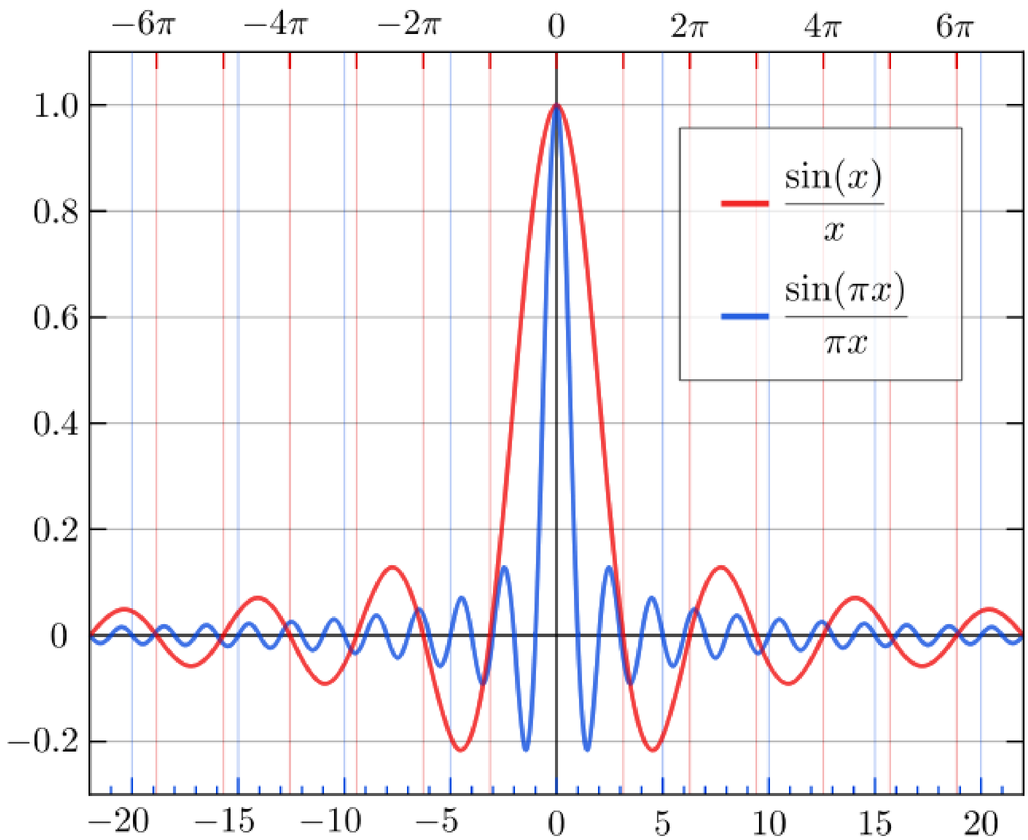


Figure 4.3: The sinc function.

4.5 The sinc function

From Section 4.6 we will need the definition of the sinc function.

Definition 4.5.1 (sinc)

We define the sinc : $\mathbb{R} \to \mathbb{R}$ (cardinal sin) function as follows:

$$\mathrm{sinc}(t) = \begin{cases} \frac{\sin(t)}{t} & ; \text{if } t \neq 0 \\ 1 & ; \text{if } t = 0 \end{cases}$$

Figure 4.3 illustrates the sinc function.

Theorem 4.5.2 (***Continuity of*** sinc ***at zero***)

The *sinc* function is continuous at 0.

Proof.

$$\begin{aligned}
\lim_{t\to 0}\operatorname{sinc}(t) &= \lim_{t\to 0}\frac{\sin(t)}{t} \\
&= \lim_{t\to 0}\frac{\frac{d}{dt}\sin(t)}{\frac{d}{dt}t} && \text{; L'hôpital's rule} \\
&= \lim_{t\to 0}\frac{\cos(t)}{1} \\
&= 1 \\
&= \operatorname{sinc}(0)
\end{aligned}$$

Theorem 4.5.3 (*Finite energy of* sinc)

The *sinc* function is a finite energy signal.

We take the proof from `https://planetmath.org/sincisl2`

Proof. We simply need to show $\operatorname{sinc} \in \mathcal{L}^2(\mathbb{R})$. We note that *sinc* is an even function so it suffices to show $\operatorname{sinc} \in \mathcal{L}^2([0,\infty])$. Furthermore, we note that $\int_0^\pi |\operatorname{sinc}(t)|^2 dt < \infty$, so it further suffices to show $\operatorname{sinc} \in \mathcal{L}^2([\pi,\infty])$.

$$\begin{aligned}
\int_\pi^\infty \left|\operatorname{sinc}^2(t)\right| dt &= \int_\pi^\infty \operatorname{sinc}^2(t)dt \\
&= \lim_{n\to\infty}\sum_{k=1}^{n}\int_{k\pi}^{(k+1)\pi}\frac{\sin^2(t)}{t^2}dt \\
&\le \lim_{n\to\infty}\sum_{k=1}^{n}\int_{k\pi}^{(k+1)\pi}\frac{\sin^2(t)}{(k\pi)^2}dt \\
&= \sum_{k=1}^{\infty}\frac{1}{2k^2\pi} \\
&\le \infty
\end{aligned}$$

4.6 Computing a Fourier Transform

We rarely calculate the Fourier transform by hand. Instead we rely heavily on predefined tables, *e.g.*, Figure 4.14. Nevertheless, we will compute the Fourier transform of Π_T as this is an equality which we will often return to in this course.

Reminder 4.6.1 (*the gate function* Π_T)

E.g., see Figure 2.14 on page 64 for a reminder of the Π function.

Recall

$$\Pi_T : t \mapsto \begin{cases} 1 & ; \text{if } -T/2 \le t \le T/2 \\ 0 & ; \text{otherwise} \end{cases}$$

We wish to compute the Fourier transform of Π_T:

Lemma 4.6.2 (*Euler's formulas for* $\sin$ *and* $\cos$)

$$\sin\theta = \frac{\exp(i\theta) - \exp(-i\theta)}{2i}$$

and

$$\cos\theta = \frac{\exp(i\theta) + \exp(-i\theta)}{2}$$

Proof.

$$\begin{aligned} \exp(i\theta) &= \cos(\theta) + i\sin\theta && (4.3) \\ \exp(-i\theta) &= \cos(-\theta) + i\sin(-\theta) \\ &= \cos(\theta) - i\sin(\theta) && (4.4) \end{aligned}$$

Adding (4.3) and (4.4), we obtain:

$$\begin{aligned} \exp(i\theta) + \exp(-i\theta) &= \cos(\theta) + i\sin\theta + \cos(\theta) - i\sin(\theta) \\ &= 2\cos(\theta) \\ \frac{\exp(i\theta) + \exp(-i\theta)}{2} &= \cos\theta \end{aligned}$$

Subtracting (4.3) and (4.3), we obtain:

$$\begin{aligned} \exp(i\theta) - \exp(-i\theta) &= (\cos(\theta) + i\sin(\theta)) - (\cos(\theta) - i\sin(\theta)) \\ &= 2i\sin(\theta) \\ \frac{\exp(i\theta) - \exp(-i\theta)}{2i} &= \sin\theta \end{aligned}$$

Theorem 4.6.3 (*Fourier transform of the gate function*)

$$\mathcal{F}(\Pi_T(t)) = T\,\mathrm{sinc}(\pi\nu T)$$

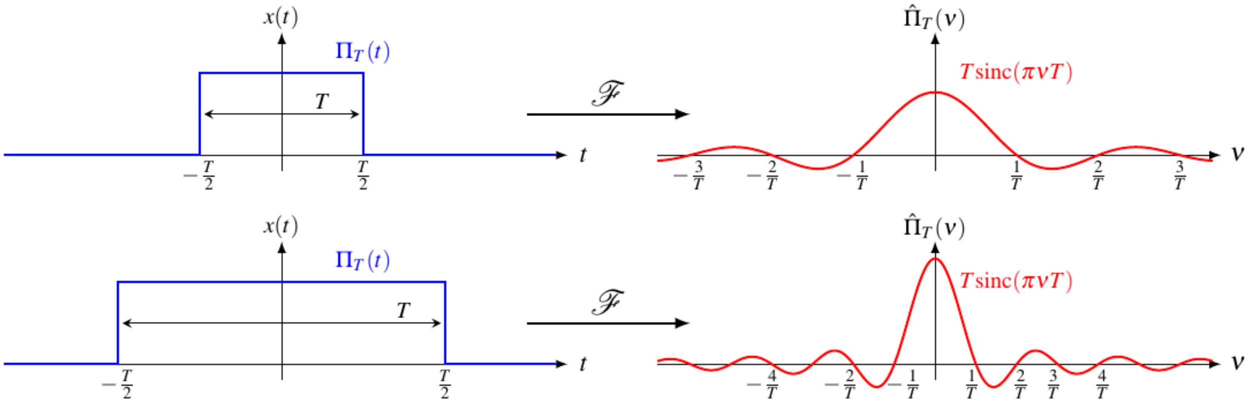


Figure 4.4: [Top] Fourier transform of window function with a narrow pulse. [Bottom] Fourier transform of window function with a wide pulse.

Proof.

$$
\begin{aligned}
\mathcal{F}(\Pi_T(t)) &= \int_{-\infty}^{+\infty} \Pi_T(t) e^{-2\pi i\nu t} dt \\
&= \int_{-T/2}^{T/2} e^{-2\pi i\nu t} dt \\
&= \big[\frac{1}{-2\pi i\nu} e^{-2\pi i\nu t}\big]_{-T/2}^{T/2} \\
&= \frac{1}{-2\pi i\nu}\big(e^{-i\pi\nu T} - e^{i\pi\nu T}\big) \\
&= \frac{e^{i\pi\nu T} - e^{-i\pi\nu T}}{2\pi i\nu} \\
&= \frac{1}{\pi\nu} \cdot \frac{e^{i\pi\nu T} - e^{-i\pi\nu T}}{2i} \\
&= \frac{\sin(\pi\nu T)}{\pi\nu} && \text{by Lemma 4.6.2} \\
&= T\frac{\sin(\pi\nu T)}{\pi\nu T} \\
&= T\,\mathrm{sinc}(\pi\nu T)
\end{aligned}
$$

So we have

$$
\mathcal{F}(\Pi_T(t)) = \widehat{\Pi_T}(\nu) = T\,\mathrm{sinc}(\pi\nu T)\,. \tag{4.5}
$$

If T is small, then Π_T is a narrow window pulse. Consequently, $\widehat{\Pi_T}$ has a wide and small in height principal lobe, as shown in Figure 4.4 [Top]. On the contrary, if T is large, then Π_T is a wide window pulse. Consequently, $\widehat{\Pi_T}$ has a narrow and tall principal lobe, as shown in Figure 4.4 [Bottom].

One might ask what is the limit as $T \to 0$ and as $T \to +\infty$ of $\mathcal{F}(\Pi_T)$? We discuss this question in the following sections.

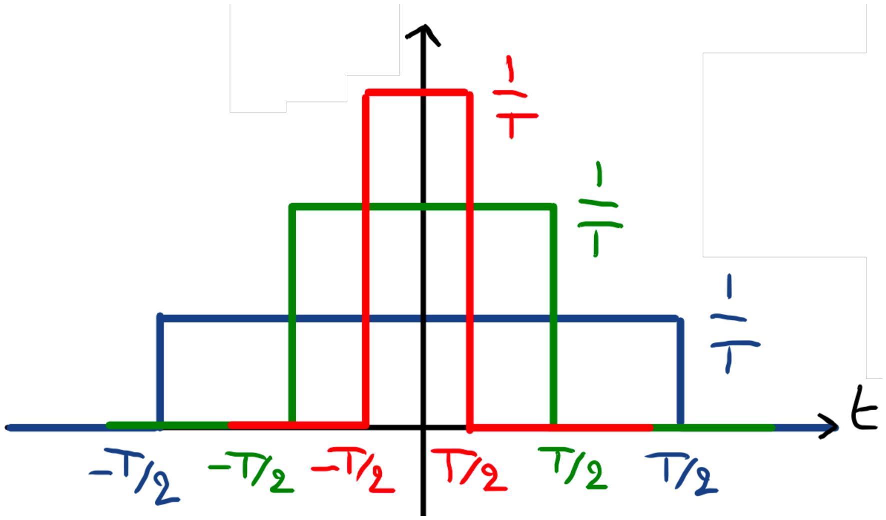


Figure 4.5: Thought experiment to justify the integration property (Eq (4.8)) of the Dirac δ-function.

4.7 Dirac Delta

We now introduce a notation which is intended to aid in computation. We will use the term *delta function*, but it is not at all a function, and we will use the symbology

$$\int_{\mathbb{R}} x(t)\delta(t)dt\,.$$

However, this is certainly not the integral (Riemann or otherwise) you've studied from calculus.

Definition 4.7.1 (*Dirac delta function*)

We *define* the *Dirac delta function* as having the following two otherwise contradictory properties.

$$\delta(t) = \begin{cases} +\infty & ;\text{ if } t = 0 \\ 0 & ;\text{ otherwise} \end{cases} \tag{4.6}$$

$$\int_{\mathbb{R}} \delta(t)dt = 1 \tag{4.7}$$

Equation (4.7) is referred to as the *mass* of the Dirac function.

The observant student has every right to be troubled by Definition 4.7.1. There are ways to formalize the mathematics, but we will not do so in this course. Some literature refers to δ as a *generalized function*, while other literature refers to it as a *distribution*. However, we will not attempt to define either of these terms.

To ease your mind you may think of $f : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, then

$$\int_{\mathbb{R}} f(\delta(t), x(t))dt$$

as a sort of shorthand for

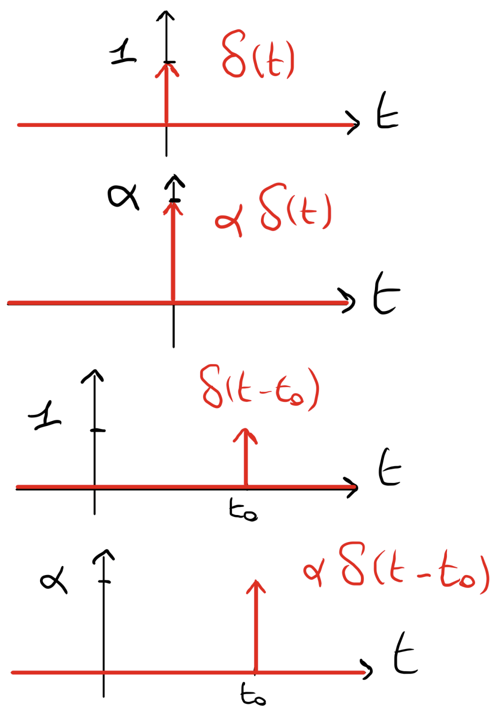
$$\int_{\mathbb{R}} f(\delta(t), x(t))dt = \lim_{T\to 0} \int_{\mathbb{R}} f\left(\frac{\Pi_T(t)}{T}, x(t)\right)dt$$

as alluded to in Figure 4.5.

In particular if $f(u, v) = u \times v$, then

$$\begin{aligned}\int_{\mathbb{R}} \delta(t)x(t)dt &= \lim_{T\to 0} \int_{\mathbb{R}} \frac{\Pi_T(t)}{T}\, x(t)dt \\ &= \lim_{T\to 0} \frac{1}{T} \int_{\mathbb{R}} \Pi_T(t)x(t)dt \end{aligned} \tag{4.8}$$

However, that analogy does not always work, because Π_T is not continuous and not infinitely differentiable. You might, however, imagine a cousin function of Π_T which is smooth.



$$\int_{\mathbb{R}} \delta(t)dt = 1 \quad \text{Eq (4.7)}$$

$$\int_{\mathbb{R}} \alpha\delta(t)dt = \alpha \tag{4.9}$$

$$\int_{\mathbb{R}} \delta(t-t_0)dt = 1 \tag{4.10}$$

$$\int_{\mathbb{R}} \alpha\delta(t-t_0)dt = \alpha \tag{4.11}$$

Figure 4.6: Variants of the δ function

Graphically, we represent the δ-function as an arrow of height equal to its mass, *i.e.*, the numerical value of its integral, as shown in Figure 4.6.

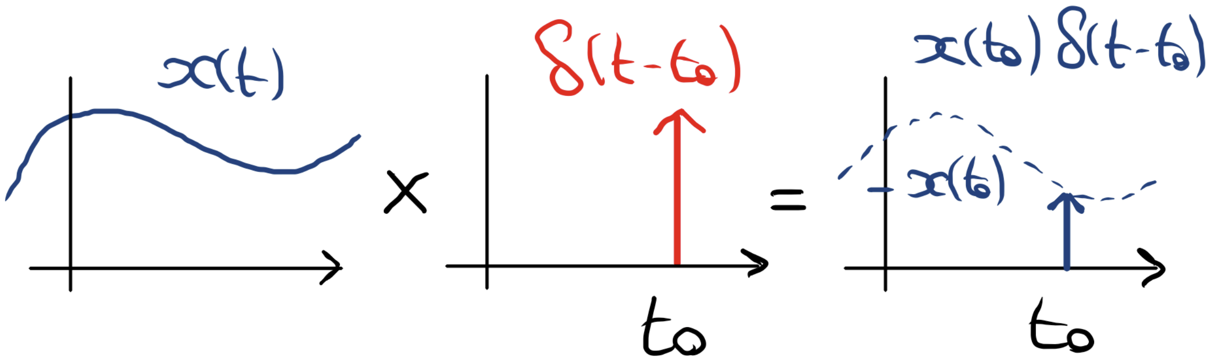


Figure 4.7: Sampling using the δ-function

4.7.1 Properties of the δ-function

There are several properties which are useful when manipulating integrals using the δ-function.

$$x(t)\delta(t-t_0) = x(t_0)\delta(t-t_0) \tag{4.12}$$

The δ function is even.

$$\delta(-x) = \delta(x) \tag{4.13}$$

This means we can also represent Equation (4.12) as (4.14).

$$x(t)\delta(t_0-t) = x(t_0)\delta(t_0-t) \tag{4.14}$$

Equation (4.12) is illustrated in Figure 4.7, and allows the following simplification:

Theorem 4.7.2 (***The Sifting Theorem***)

$$\int_{\mathbb{R}} x(t)\delta(t-t_0)dt = x(t_0)$$

Proof.

$$\begin{aligned}
\int_{\mathbb{R}} x(t)\delta(t-t_0)dt &= \int_{\mathbb{R}} x(t_0)\delta(t-t_0)dt && ;\text{ by } (4.12)\\
&= x(t_0)\underbrace{\int_{\mathbb{R}} \delta(t-t_0)dt}_{=1} \\
&= x(t_0) && ;\text{ by } (4.10)
\end{aligned}$$

We have recovered the value of x at t_0; *i.e.*, we have *sampled* x at t_0.

4.7.2 Convolution of δ

The second important property is:

Theorem 4.7.3 (*δ is the convolution identity*)

$$\delta * x = x * \delta = x$$

This means that δ acts as the identity element for the convolution operation. As this is not really a function, it is not really claiming that convolution has an identity. We will prove Theorem 4.7.3 in two different ways. First using the limit as in Equation (4.8), and second using Theorem 4.7.2.

Note that there are several different relations to prove here. We know that convolution is commutative on functions from $\mathcal{L}^1(\mathbb{R})$, but since $\delta \notin \mathcal{L}^1(\mathbb{R})$, we need to show that $x * \delta = \delta * x$; thereafter we can show that $x * \delta = x$.

No. 1 of Theorem 4.7.3.

First we show that $\underbrace{x * \delta}_{I} = \underbrace{\delta * x}_{J}$.

$$\begin{aligned} I(\tau) &= (x * \delta)(\tau) \\ &= \int_{\mathbb{R}} x(t)\delta(\tau - t)dt \\ &= \lim_{T\to 0} \frac{1}{T} \int_{\mathbb{R}} x(t)\Pi_T(\tau - t)dt && \text{; by (4.8)} \\ &= \lim_{T\to 0} \frac{1}{T} \int_{\tau - T/2}^{\tau + T/2} x(t)dt \end{aligned} \tag{4.15}$$

$$\begin{aligned} J(\tau) &= (\delta * x)(\tau) \\ &= \int_{\mathbb{R}} \delta(t)x(\tau - t)dt \\ &= \lim_{T\to 0} \frac{1}{T} \int_{\mathbb{R}} x(\tau - t)\Pi_T(t)dt && \text{; by (4.8)} \\ &= \lim_{T\to 0} \frac{1}{T} \int_{-T/2}^{T/2} x(\tau - t)dt \end{aligned}$$

Let $u = \tau - t$ and $du = -dt$, so that

$$\begin{aligned} J(\tau) &= \lim_{T\to 0} \frac{1}{T} \int_{\tau + T/2}^{\tau - T/2} -x(u)du \\ &= \lim_{T\to 0} \frac{1}{T} \int_{\tau - T/2}^{\tau + T/2} x(u)du \\ &= I(\tau) && \text{by (4.15)} \end{aligned}$$

We have now established that $I(\tau) = J(\tau)$, so

$$x * \delta = \delta * x\,.$$

Now we will show that $x * \delta = x$.

Let P be a primitive of x so that

$$\frac{d}{dt}P(t) = x(t)\,. \tag{4.16}$$

$$\begin{aligned}
x * \delta &= \lim_{T\to 0} \frac{1}{T} \int_{\tau - T/2}^{\tau + T/2} x(t)dt && ; \text{by } (4.15) \\
&= \lim_{T\to 0} \frac{1}{T} \big[P(t)\big]_{\tau-T/2}^{\tau+T/2} \\
&= \lim_{T\to 0} \frac{P(\tau + T/2) - P(\tau - T/2)}{T} \\
&= \lim_{T\to 0} \frac{P(\tau + T/2) - P(\tau) + P(\tau) - P(\tau - T/2)}{T} \\
&= \lim_{T\to 0} \frac{P(\tau + T/2) - P(\tau)}{T} + \lim_{T\to 0} \frac{P(\tau) - P(\tau - T/2)}{T}
\end{aligned}$$

Now let $T_1 = T/2$, and $T_2 = -T/2$. So $T = 2T_1 = -2T_2$

$$\begin{aligned}
x * \delta &= \lim_{T_1\to 0} \frac{P(\tau + T_1) - P(\tau)}{2T_1} + \lim_{T_2\to 0} \frac{P(\tau) - P(\tau + T_2)}{-2T_2} \\
&= \lim_{T_1\to 0} \frac{P(\tau + T_1) - P(\tau)}{2T_1} + \lim_{T_2\to 0} \frac{P(\tau + T_2) - P(\tau)}{2T_2} \\
&= \frac{1}{2}\frac{d}{dt}P(\tau) + \frac{1}{2}\frac{d}{dt}P(\tau) \\
&= \frac{d}{dt}P(\tau) = x(\tau) && ; \text{by } (4.16)
\end{aligned}$$

No. 2 of Theorem 4.7.3.

$$\begin{aligned}
(x * \delta)(\tau) &= \int_{\mathbb{R}} x(t)\delta(\tau - t)dt \\
&= \int_{\mathbb{R}} x(\tau)\delta(\tau - t)dt && ; \text{by } (4.14) \\
&= x(\tau) \underbrace{\int_{\mathbb{R}} \delta(\tau - t)dt}_{=1\ ;\ \text{by } (4.7)} \\
&= x(\tau)
\end{aligned}$$

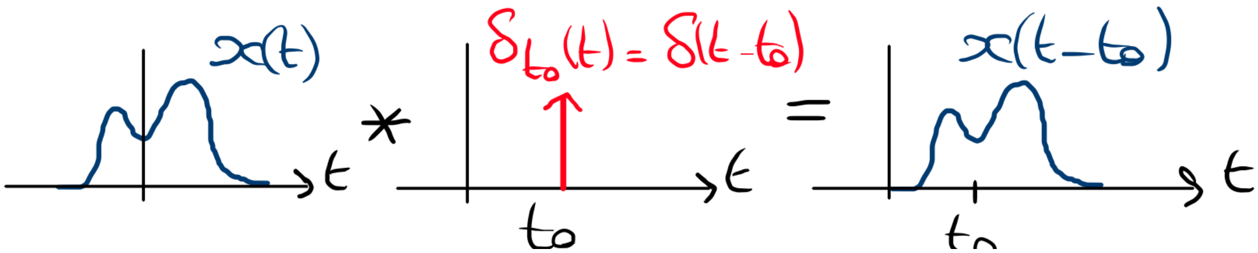


Figure 4.8: Convolution of a signal with shifted δ-function.

Also,

$$\begin{aligned}
(\delta * x)(\tau) &= \int_{-\infty}^{+\infty} \delta(t)x(\tau - t)dt \\
&= \int_{+\infty}^{-\infty} -\delta(\tau - u)x(u)du && \text{; Let } u = \tau - t \text{ and } du = -dt \\
&= \int_{-\infty}^{\infty} \delta(\tau - u)x(u)du \\
&= \int_{-\infty}^{\infty} \delta(\tau - u)x(\tau)du && \text{; by (4.14)} \\
&= x(\tau) \underbrace{\int_{-\infty}^{+\infty} \delta(\tau - u)du}_{=1 \text{ ; by (4.7)}} \\
&= x(\tau)
\end{aligned}$$

It is easy to see from the second proof of Theorem 4.7.3 how well adapted the δ notation is for performing signal related computations.

Definition 4.7.4 (***the shifted delta notation*** δ_{t_0})

We use the symbol δ_{t_0} to denote $\delta(t - t_0)$, *i.e.*, the Dirac δ-function centered at t_0.

Theorem 4.7.5 is illustrated in Figure 4.8.

Theorem 4.7.5 (***Convolution with a shifted*** δ)

Convolving a signal with a delayed δ is equivalent to delaying the signal by that delay.

$$(x * \delta_{t_0})(\tau) = x(\tau - t_0)$$

Proof. Recall $x(t) * y(t) = \int_{-\infty}^{\infty} x(t)y(\tau - t)dt$. So let

$$y(t) = \delta(t - t_0)$$
$$y(\tau - t) = \delta(\tau - t - t_0)$$

Since the argument of δ is 0 when $0 = \tau - t - t_0 \implies t = \tau - t_0$, by the Sifting Theorem 4.7.2, we will need to evaluate x at $t = \tau - t_0$

Now, let's evaluate the convolution.

$$
\begin{aligned}
x(t) * \delta(t - t_0) &= x(t) * y(t) \\
&= \int_{-\infty}^{\infty} x(t) y(\tau - t) dt \\
&= \int_{-\infty}^{\infty} x(t) \delta(\tau - t - t_0) dt \\
&= \int_{-\infty}^{\infty} x(\tau - t_0) \delta(\tau - t - t_0) dt && \text{By Theorem 4.7.2} \\
&= x(\tau - t_0) \int_{-\infty}^{\infty} \delta(\tau - t - t_0) dt \\
&= x(\tau - t_0) \times 1 && \text{By (4.10)} \\
&= x(\tau - t_0)
\end{aligned}
$$

4.7.3 Other Integrals involving Dirac delta

Theorem 4.7.6 (***Vanishing tails of*** $\int x\delta$)

If x is a signal and $a > 0$, then

$$\int_{-\infty}^{t_0 - a} x(t) \delta(t - t_0) dt = 0 \,,$$

and

$$\int_{t_0 + a}^{\infty} x(t) \delta(t - t_0) dt = 0 \,.$$

Proof. Let $a > 0$, then by (4.6),

$$
\begin{aligned}
t < t_0 - a &\implies t - t_0 < -a < 0 \\
&\implies \delta(t - t_0) = 0 \\
&\implies \int_{-\infty}^{t_0 - a} x(t) \delta(t - t_0) dt = \int_{-\infty}^{t_0 - a} x(t) \; 0 \; dt = 0
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
t > t_0 + a &\implies t - t_0 > a > 0 \\
&\implies \delta(t - t_0) = 0 \\
&\implies \int_{t_0 + a}^{\infty} x(t) \delta(t - t_0) dt = \int_{t_0 + a}^{\infty} x(t) \; 0 \; dt = 0
\end{aligned}
$$

Corollary 4.7.7 (δ ***integral vanishing away from*** t_0)

If x is a signal and $a > b > 0$, then

$$\int_{t_0 - b}^{t_0 - a} x(t) \delta(t - t_0) dt = 0$$

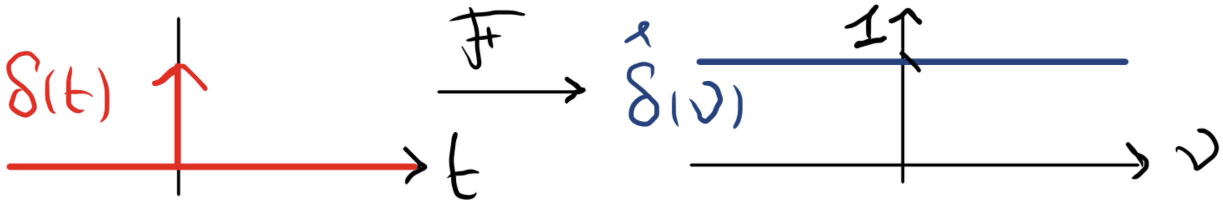


Figure 4.9: Fourier transform of the δ function

Proof. Let $a > b > 0$.

$$\begin{aligned} I &= \int_{t_0-b}^{t_0-a} x(t)\delta(t-t_0)dt \\ &= \int_{-\infty}^{t_0-a} x(t)\delta(t-t_0)dt - \int_{-\infty}^{t_0-b} x(t)\delta(t-t_0)dt \\ &= 0-0 && \text{by Theorem 4.7.6} \\ &= 0 \end{aligned}$$

Corollary 4.7.8 (*δ integral over an interval avoiding 0*)
If $[a, b]$ is an interval that does not contain 0, then

$$\int_a^b x(t)\delta(t)dt = 0\,.$$

Proof. If $0 < a \leq b$, then, taking $t_0 = 0$ in the right tail of Theorem 4.7.6, $\int_a^\infty x(t)\delta(t)dt = 0$ and $\int_b^\infty x(t)\delta(t)dt = 0$, so

$$\begin{aligned} \int_a^b x(t)\delta(t)dt &= \int_a^\infty x(t)\delta(t)dt - \int_b^\infty x(t)\delta(t)dt \\ &= 0-0=0\,. \end{aligned}$$

If $a \leq b < 0$, then, taking $t_0 = 0$ in the left tail of Theorem 4.7.6, $\int_{-\infty}^a x(t)\delta(t)dt = 0$ and $\int_{-\infty}^b x(t)\delta(t)dt = 0$, so

$$\begin{aligned} \int_{t=a}^{t=b} x(t)\delta(t)dt &= \int_{-\infty}^b x(t)\delta(t)dt - \int_{-\infty}^a x(t)\delta(t)dt \\ &= 0-0=0 \end{aligned}$$

4.7.4 Fourier transform and inverse transform of δ

Since we are allowing ourselves to compute integrals involving the δ-function, we might ask whether the δ-function has a Fourier transform. The Fourier transform is illustrated in Figure 4.9.

Theorem 4.7.9 (*Fourier transform of* δ)

$$\mathcal{F}(\delta(t)) = 1$$

Proof.

$$\begin{aligned}
\mathcal{F}(\delta(t)) &= \hat{\delta}(t) \\
&= \int_{\mathbb{R}} \delta(t) e^{-2\pi i \nu t} dt \\
&= \int_{\mathbb{R}} \delta(t) e^{-2\pi i \nu \times 0} dt && ; \text{by } (4.12) \\
&= \int_{\mathbb{R}} \delta(t) \times 1 dt \\
&= \int_{\mathbb{R}} \delta(t) dt = 1 && ; \text{by } (4.7)
\end{aligned}$$

From the frequency perspective, the Dirac δ-function is a superposition of all frequencies having the same amplitude.

Can we compute $\mathcal{F}(1)$? If we try, we find that we need to compute the integral,

$$\int_{+\infty}^{-\infty} \exp(2\pi i \nu t) dt = \frac{1}{2\pi i \nu} \left(\exp(2\pi i \nu \times +\infty) - \exp(2\pi i \nu \times -\infty)\right),$$

which we can't solve so easily.

However, we can compute $\mathcal{F}(1)$ in an indirect way.

Theorem 4.7.10 (*Fourier transform of the constant signal 1*)

$$\mathcal{F}(1) = \delta(\nu)$$

We will show $\mathcal{F}^{-1}(\delta(\nu)) = 1$, which implies $\mathcal{F}(1) = \delta(\nu)$.

Proof.

$$\begin{aligned}
\mathcal{F}^{-1}(\delta(\nu)) &= \int_{\mathbb{R}} \delta(\nu) e^{2\pi i \nu t} d\nu \\
&= \int_{\mathbb{R}} \delta(\nu) e^{2\pi i t \times 0} d\nu && ; \text{by } (4.12) \\
&= \int_{\mathbb{R}} \delta(\nu) d\nu = 1 && ; \text{by } (4.7)
\end{aligned}$$

4.8 Properties of the Fourier Transform

In the discussion which follows, we will take many liberties which may not be strictly mathematically correct. A rigorous course in functional analysis would fill in the gaps. That amount of rigor is beyond the scope of this course in introduction to signal processing.

4.8.1 Linearity

The Fourier transform is a linear operator.

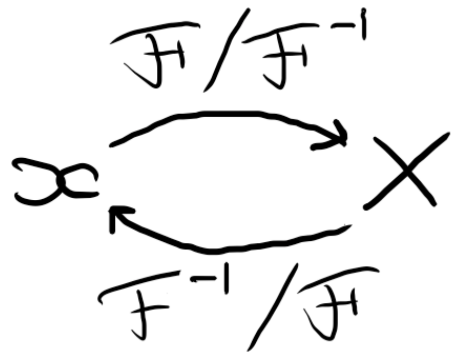


Figure 4.10: Even functions are involutive.

Theorem 4.8.1 (*Linearity of the Fourier Transform*)

If $\alpha \in \mathbb{C}$ and $\mathcal{F}(x)$ and $\mathcal{F}(y)$ exist, then

$$\mathcal{F}(x(t) + y(t)) = \mathcal{F}(x(t)) + \mathcal{F}(y(t)) \tag{4.17}$$
$$\mathcal{F}(\alpha x(t)) = \alpha \mathcal{F}(x(t)) \,. \tag{4.18}$$

Proof.

$$\begin{aligned}
\mathcal{F}(\alpha x(t) + y(t)) &= \int_{\mathbb{R}} (\alpha x(t) + y(t)) \exp(-2\pi i \nu t) dt \\
&= \int_{\mathbb{R}} (\alpha x(t)) \exp(-2\pi i \nu t) dt + \int_{\mathbb{R}} y(t) \exp(-2\pi i \nu t) dt \\
&= \alpha \int_{\mathbb{R}} (x(t)) \exp(-2\pi i \nu t) dt + \int_{\mathbb{R}} y(t) \exp(-2\pi i \nu t) dt \\
&= \alpha\, \mathcal{F}(x(t)) + \mathcal{F}(y(t))
\end{aligned} \tag{4.19}$$

From (4.19), let $\alpha = 1$, to obtain (4.17). From (4.19), let $\alpha = -1$ and let $x = y$, to obtain:

$$\mathcal{F}(0) = 0 \,. \tag{4.20}$$

From (4.19) and (4.20), let $y = 0$ to obtain (4.18).

4.8.2 Involutive

Reminder 4.8.2 (*the reflection notation* x^-)

For Theorem 4.8.3, recall the notation: $x^-(t) = x(-t)$.

Theorem 4.8.3 (*Double Fourier transform*)

If x is a signal which has a Fourier transform,

$$\mathcal{F}(\mathcal{F}(x)) = x^- \,.$$

Proof. Suppose $X(\nu) = \mathcal{F}(x(t))$.

$$\begin{aligned}
\mathcal{F}(x(t)) &= X(\nu) \\
\mathcal{F}^{-1}(\mathcal{F}(x(t))) &= \mathcal{F}^{-1}(X(\nu)) \\
x(t) &= \int_{\mathbb{R}} X(\nu) e^{2\pi i\nu t} d\nu \\
x(-t) &= \int_{\mathbb{R}} X(\nu) e^{2\pi i\nu(-t)} d\nu \\
x(-t) &= \int_{\mathbb{R}} X(\nu) e^{-2\pi i\nu t} d\nu \qquad (4.21)
\end{aligned}$$

Notice that (4.21) is simply the Fourier transform of $X(\nu)$ but with t and ν reversed from their conventional usage. Recall that the signal x and its Fourier transform X are both complex valued signals of a real variable: $x : \mathbb{R} \to \mathbb{C}$ and also $X : \mathbb{R} \to \mathbb{C}$. As it is written, Eq (4.21) is a function of t, so we have:

$$\begin{aligned}
x^{-} &= t \mapsto \int_{\mathbb{R}} X(\xi) e^{-2\pi i\xi t} d\xi \\
&= \mathcal{F}(X) \\
&= \mathcal{F}(\mathcal{F}(x))
\end{aligned}$$

Corollary 4.8.4 (*Involution of the Fourier transform on even signals*)

The Fourier transform is involutive for all even functions.

$$\mathcal{F}(\mathcal{F}(x)) = x$$

or

$$\mathcal{F}(x) = \mathcal{F}^{-1}(x)$$

Proof. Since x is an even function, $x = x^{-}$, so apply Theorem 4.8.3.

In particular, since Π_T, sinc, δ, and 1 are even functions, and we've already shown that

$$\begin{aligned}
\mathcal{F}(\Pi_T(t)) &= T\,\mathrm{sinc}(\pi\nu T) \\
\mathcal{F}(\delta) &= 1\,,
\end{aligned}$$

then, as shown in Figure 4.11 we have automatically (thanks to Corollary 4.8.4) we also have

$$\begin{aligned}
\mathcal{F}(T\,\mathrm{sinc}(\pi\nu T)) &= \Pi_T(t) \qquad (4.22) \\
\mathcal{F}(1) &= \delta\,. \qquad (4.23)
\end{aligned}$$

4.8.3 Negative Frequencies

Theorem 4.8.5 (*Hermitian symmetry of the Fourier transform*)

If x is a real valued signal: $x(t) \in \mathbb{R}$, then X is a symmetric Hermitian

$$\overline{X(\nu)} = X(-\nu)$$

Proof. Let x be a real valued signal.

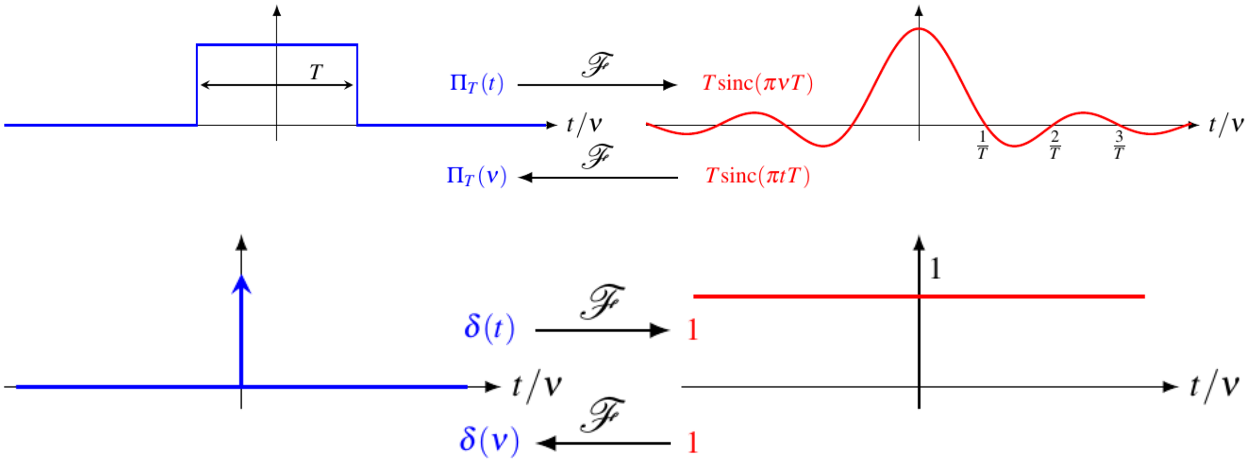


Figure 4.11: Some Involutive Fourier Transforms

$$
\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt \\
\overline{X(\nu)} &= \overline{\int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt} \\
&= \int_{\mathbb{R}} x(t)\,\overline{e^{-2\pi i\nu t}}\,dt \\
&= \int_{\mathbb{R}} x(t)e^{2\pi i\nu t}dt \\
&= \int_{\mathbb{R}} x(t)e^{-2\pi i(-\nu)t}dt \\
&= X(-\nu)
\end{aligned}
$$

Theorem 4.8.5 should remind you that the Fourier series of a real valued signal, the coefficients:

$$c_{-n} = \overline{c_n}\,.$$

Theorem 4.8.6 (*Symmetry of the spectrum*)
The spectrum of a real valued signal is symmetric.

$$|X(-\nu)| = |X(\nu)|$$

Proof. Suppose x is a real valued signal. From Theorem 4.8.5, we have $\overline{X(\nu)} = X(-\nu)$. And we already know $\left|\overline{X(\nu)}\right| = |X(\nu)|$. So it follows that

$$\left|\overline{X(\nu)}\right| = |X(\nu)| = |X(-\nu)|\,.$$

Similar to Theorem 4.8.5, Lemma 4.8.7 is an identity which will be useful later in the proof of Theorem 4.12.2.

Lemma 4.8.7 (*Fourier transform of $\overline{y^-}$*)

$$\mathcal{F}\left(\overline{y^-}\right) = \overline{\mathcal{F}(y(\nu))}$$

Proof.

$$\begin{aligned}
\mathcal{F}\left(\overline{y^-}\right) &= \int_{\mathbb{R}} \overline{y^-}\, e^{-2\pi i\nu t} dt = \int_{\mathbb{R}} \overline{y(-t)}\, e^{-2\pi i\nu t} dt \\
&= \int_{\mathbb{R}} \overline{y(-t)}\ \overline{e^{2\pi i\nu t}}\, dt = \int_{\mathbb{R}} \overline{y(-t)\ e^{2\pi i\nu t}}\, dt \\
&= \overline{\int_{t=-\infty}^{t=+\infty} y(-t) e^{2\pi i\nu t} dt} \\
&= \overline{\int_{u=+\infty}^{u=-\infty} \underbrace{-y(u) e^{-2\pi i\nu u} du}_{;\ \text{letting}\ u=-t,\ du=-dt}} \\
&= \overline{\int_{u=-\infty}^{u=+\infty} y(u) e^{-2\pi i\nu u} du} \\
&= \overline{\mathcal{F}(y(\nu))}
\end{aligned}$$

4.8.4 Parity

Theorem 4.8.8 (*Fourier transform of a real even signal*)

If x is real and even (an even signal), then X is real and even.

First approach.

$$\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t) e^{-2\pi i\nu t} dt \\
&= \int_{\mathbb{R}} x(t)(\cos(2\pi\nu t) + i \sin(2\pi\nu t)) dt \\
&= \int_{\mathbb{R}} \underbrace{x(t)\ \cos(2\pi\nu t)}_{\in\mathbb{R}} dt + i \lim_{T\to\infty} \underbrace{\int_{-T}^{T} \overbrace{x(t)\sin(2\pi\nu t)}^{\text{odd function}} dt}_{=0} \\
&= \int_{\mathbb{R}} x(t)\cos(2\pi\nu t) dt
\end{aligned}$$

Now if we look at $X(-\nu)$,

$$\begin{aligned}
X(-\nu) &= \int_{\mathbb{R}} x(t) \underbrace{\cos(2\pi(-\nu)t)}_{= \cos(2\pi\nu t)} dt \\
&= \int_{\mathbb{R}} x(t)\cos(2\pi\nu t)dt \\
&= X(\nu)
\end{aligned}$$

Thus X is real and even.

Second approach. Let x be an even function so that $x(t) = x(-t)$.

$$\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt \\
&= \underbrace{\int_{t=-\infty}^{t=0} x(t)e^{-2\pi i\nu t}dt}_{\text{Let } u=-t, du=-dt} + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= \int_{u=\infty}^{u=0} \underbrace{x(-u)}_{\text{even}} e^{-2\pi i\nu(-u)}(-du) + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= -\int_{u=\infty}^{u=0} x(u)e^{2\pi i\nu u}du + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= \int_{u=0}^{u=\infty} x(u)e^{2\pi i\nu u}du + \int_0^\infty x(t)e^{-2\pi i\nu t}dt
\end{aligned}$$

Now change variables $t = u$,

$$\begin{aligned}
X(\nu) &= \int_0^\infty x(t)e^{2\pi i\nu t}dt + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= \int_0^\infty \left(x(t)e^{2\pi i\nu t} + x(t)e^{-2\pi i\nu t}\right)dt \\
&= \int_0^\infty x(t) \underbrace{\left(e^{2\pi i\nu t} + e^{-2\pi i\nu t}\right)}_{= 2\cos(2\pi\nu t)} dt \\
&= 2\int_0^\infty x(t)\cos(2\pi\nu t)dt
\end{aligned}$$

Now, look at $X(-\nu)$.

$$\begin{aligned}
X(-\nu) &= 2\int_0^\infty x(t) \underbrace{\cos(2\pi(-\nu)t)}_{= \cos(2\pi\nu t)} dt \\
&= 2\int_0^\infty x(t)\cos(2\pi\nu t)dt \\
&= X(\nu)
\end{aligned}$$

Thus $X(\nu)$ is real and even.

Theorem 4.8.9 (*Fourier transform of a real odd signal*)
If x is a real valued odd function, then X is odd and purely imaginary.

First approach.

$$\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt \\
&= \underbrace{\lim_{T\to\infty}\int_{-T}^{T} \underbrace{\overbrace{x(t)}^{\text{odd}}\,\overbrace{\cos(2\pi\nu t)}^{\text{even}}}_{\text{odd}}\,dt}_{=0} + i\underbrace{\int_{\mathbb{R}} x(t)\sin(2\pi\nu t)dt}_{\in\mathbb{R}} \\
&= i\int_{\mathbb{R}} x(t)\sin(2\pi\nu t)dt
\end{aligned}$$

Now, we look at $X(-\nu)$.

$$\begin{aligned}
X(-\nu) &= i\int_{\mathbb{R}} x(t)\underbrace{\sin(2\pi(-\nu)t)}_{=\,-\sin(2\pi\nu t)}\,dt \\
&= -i\int_{\mathbb{R}} x(t)\sin(2\pi\nu t)dt
\end{aligned}$$

So $X(\nu)$ is pure imaginary and odd.

Second approach. Let x be an odd function so that $x(t) = -x(-t)$.

$$\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt \\
&= \underbrace{\int_{t=-\infty}^{t=0} x(t)e^{-2\pi i\nu t}dt}_{\text{Let } u=-t, du=-dt} + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= \int_{u=\infty}^{u=0} \underbrace{x(-u)}_{\text{odd}}\, e^{-2\pi i\nu(-u)}(-du) + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= -\int_\infty^0 -x(u)e^{2\pi i\nu u}du + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= \int_\infty^0 x(u)e^{2\pi i\nu u}du + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= -\int_0^\infty x(u)e^{2\pi i\nu u}du + \int_0^\infty x(t)e^{-2\pi i\nu t}dt
\end{aligned}$$

Now change variables $t = u$,

$$\begin{aligned}
X(\nu) &= -\int_0^\infty x(t)e^{2\pi i\nu t}dt + \int_0^\infty x(t)e^{-2\pi i\nu t}dt \\
&= \int_0^\infty \left(x(t)e^{2\pi i\nu t} - x(t)e^{-2\pi i\nu t}\right)dt \\
&= \int_0^\infty x(t)\, \underbrace{\left(e^{2\pi i\nu t} - e^{-2\pi i\nu t}\right)}_{=\ 2i\sin(2\pi\nu t)} dt \\
&= 2i\int_0^\infty x(t)\sin(2\pi\nu t)dt
\end{aligned}$$

Now, look at $X(-\nu)$.

$$\begin{aligned}
X(-\nu) = 2i\int_0^\infty x(t)\, &\underbrace{\sin(2\pi(-\nu)t)}_{=\ -\sin(2\pi\nu t)}\, dt \\
&= -2i\int_0^\infty x(t)\sin(2\pi\nu t)dt \\
&= -X(\nu)
\end{aligned}$$

Thus $X(\nu)$ is purely imaginary and odd.

Exercise 4.8.10 (*Fourier transform of purely imaginary signals*)

What can you say about the Fourier transform of a purely imaginary odd signal? What about a purely imaginary even signal?

4.8.5 Fourier Transform and Derivatives

In this section we sweep a few things under the rug: existence and integrability of the derivative of $x(t)$, existence of the derivative of $X(\nu)$, and integrability of $tx(t)$.

Theorem 4.8.11 (*Fourier Transform of Derivative*)

If $\mathcal{F}(x(t)) = X(\nu)$, then

$$\mathcal{F}\left(\frac{d}{dt}\, x(t)\right) = 2\pi i\ \nu\ X(\nu)\,.$$

For this proof we'll make use of the Leibniz integral rule. Many proofs are available for this rule, for example Wikipedia.

$$\frac{d}{dx}\int_a^b f(x,t)dt = \int_a^b \frac{\partial}{\partial x}f(x,t)dt \tag{4.24}$$

Proof. Suppose $\mathcal{F}(x(t)) = X(\nu)$.

$$\begin{aligned}
x(t) &= \int_{\mathbb{R}} X(\nu)\exp(2\pi i\nu t)dt \\
\frac{d}{dt}x(t) &= \frac{d}{dt}\int_{\mathbb{R}} X(\nu)\exp(2\pi i\nu t)dt \\
&= \int_{\mathbb{R}} \frac{\partial}{\partial t}X(\nu)\exp(2\pi i\nu t)dt && \text{;by (4.24)} \\
&= \int_{\mathbb{R}} 2\pi i\nu X(\nu)\exp(2\pi i\nu t)dt \\
&= 2\pi i\int_{\mathbb{R}} \nu X(\nu)\exp(2\pi i\nu t)dt \\
&= 2\pi i\ \mathcal{F}^{-1}(\nu X(\nu))
\end{aligned}$$

Now we take the Fourier transform of both sides to obtain:

$$\begin{aligned}
\mathcal{F}\left(\frac{d}{dt}x(t)\right) &= \mathcal{F}\big(2\pi i\ \mathcal{F}^{-1}(\nu X(\nu))\big) \\
&= 2\pi i\ \mathcal{F}\big(\mathcal{F}^{-1}(\nu X(\nu))\big) \\
&= 2\pi i\ \nu X(\nu)
\end{aligned}$$

Theorem 4.8.12 (*Inverse Fourier Transform of Derivative*)

If $\mathcal{F}(x(t)) = X(\nu)$, then

$$\mathcal{F}^{-1}\left(\frac{d}{d\nu}\ X(\nu)\right) = -2\pi i\ t\ x(t)\,; \tag{4.25}$$

equivalently,

$$\mathcal{F}(t\ x(t)) = \frac{i}{2\pi}\frac{d}{d\nu}X(\nu)\,. \tag{4.26}$$

Proof. Suppose $\mathcal{F}(x(t)) = X(\nu)$.

$$\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t)\exp(-2\pi i\nu t)dt \\
\frac{d}{d\nu}X(\nu) &= \frac{d}{d\nu}\int_{\mathbb{R}} x(t)\exp(-2\pi i\nu t)dt \\
&= \int_{\mathbb{R}} \frac{\partial}{\partial\nu}x(t)\exp(-2\pi i\nu t)dt && \text{;by (4.24)} \\
&= -2\pi i\int_{\mathbb{R}} t\ x(t)\exp(-2\pi i\nu t)dt \\
&= -2\pi i\ \mathcal{F}(t\ x(t)) \\
\mathcal{F}(t\ x(t)) &= \frac{1}{-2\pi i}\frac{d}{d\nu}X(\nu) \\
&= \frac{i}{2\pi}\frac{d}{d\nu}X(\nu)
\end{aligned}$$

This equality confirms (4.26). Now we can take the inverse Fourier transform of both sides.

$$\begin{aligned}
\mathcal{F}^{-1}\left(\frac{1}{-2\pi i}\frac{d}{d\nu}X(\nu)\right) &= \mathcal{F}^{-1}(\mathcal{F}(t\ x(t))) \\
\frac{1}{-2\pi i}\ \mathcal{F}^{-1}\left(\frac{d}{d\nu}X(\nu)\right) &= t\ x(t) \\
\mathcal{F}^{-1}\left(\frac{d}{d\nu}X(\nu)\right) &= -2\pi i\ t\ x(t)
\end{aligned}$$

This equality confirms (4.25).

In summary, multiplying by t in the time domain corresponds to taking the derivative in the frequency domain; and taking the derivative in the time domain corresponds to multiplying by ν in the frequency domain—modulo a constant multiple.

4.8.6 Shifting in Time

- The modulus of the Fourier transform gives the distribution of the amplitude of frequencies in the signal.
- The phase of the Fourier transform gives the localization (in time) of the frequencies in the signal.

In general

$$\begin{aligned}
X(\nu) \in \mathbb{C} \implies X(\nu) &= \mathfrak{Re}\{X(\nu)\} + i\mathfrak{Im}\{X(\nu)\} \\
&= |X(\nu)|e^{i\phi(X(\nu))}
\end{aligned} \tag{4.27}$$

Definition 4.8.13 (*spectrum, phase*)

In the formulation in Equation (4.27), we call $|X(\nu)|$ the *spectrum*, and we call $\phi(X(\nu))$ the *phase* of the Fourier transform.

The spectrum represents the amplitude of the frequencies of signal, x.

- Can we analyze the information contained in the phase $\phi(X(\nu))$?
- What happens if we compute $\mathcal{F}(x_{t_0}) = \mathcal{F}(x(t - t_0))$

Theorem 4.8.14 (*Time Shifting*)

Time-shifting a signal, x, by t_0 corresponds to multiplying the Fourier transform by $e^{-2\pi i\nu t_0}$.

$$\mathcal{F}(x(t - t_0)) = X(\nu)\exp(-2\pi i\nu t_0)$$

Proof.

$$\begin{aligned}
\mathcal{F}(x_{t_0}(t)) &= \int_{\mathbb{R}} x_{t_0}(t)e^{-2\pi i\nu t}dt \\
&= \int_{\mathbb{R}} x(t-t_0)e^{-2\pi i\nu t}dt \\
&= \int_{\mathbb{R}} x(u)e^{-2\pi i\nu(u+t_0)}du && \text{Letting } u = t - t_0 \\
&= \int_{\mathbb{R}} x(u)e^{-2\pi i\nu u}e^{-2\pi i\nu t_0}du \\
&= e^{-2\pi i\nu t_0}\underbrace{\int_{\mathbb{R}} x(u)e^{-2\pi i\nu u}du}_{\mathcal{F}(x(t))} \\
&= X(\nu)e^{-2\pi i\nu t_0}
\end{aligned}$$

Theorem 4.8.15 (*Effect of time-shifting on modulus and phase*)
Shifting signal, x in time does not change the modulus of the Fourier transform, but it does change the phase.

Proof.

$$\begin{aligned}
|X_{t_0}(\nu)| &= |\mathcal{F}(x(t-t_0))| \\
&= \left|X(\nu)e^{-2\pi i\nu t_0}\right| \\
&= |X(\nu)| \cdot \left|e^{-2\pi i\nu t_0}\right| \\
&= |X(\nu)|
\end{aligned}$$

$$\begin{aligned}
\phi(X_{t_0}(\nu)) &= \phi(X(\nu)e^{-2\pi i\nu t_0}) \\
&= \phi(X(\nu)) - 2\pi\nu t_0 \\
&\neq \phi(X(\nu))
\end{aligned}$$

4.8.7 Interpretation

- The modulus/spectrum of the Fourier transform gives the distribution of amplitude of the frequencies in the signal.
- The phase of the Fourier transform gives the temporal location of the frequencies in the signal.

We have just seen the property:

$$\text{time shift} \iff \text{phase shift}$$
$$\mathcal{F}(x(t-t_0)) = X(\nu)e^{-2\pi i\nu t_0}$$

There is also a dual version of this property. Consider a signal, x, with Fourier transform, $X = \mathcal{F}(x)$. What is the impact of multiplying x by a complex exponential?

Theorem 4.8.16 (*Frequency shifting (modulation)*)

If $\mathcal{F}(x) = X$, then
$$\mathcal{F}(x(t)\exp(2\pi i \nu_0 t)) = X(\nu - \nu_0)\,.$$

Proof.
$$\begin{aligned}\mathcal{F}\big(x(t)e^{2\pi i \nu_0 t}\big) &= \int_{\mathbb{R}} x(t)e^{2\pi i \nu_0 t}e^{-2\pi i \nu t}dt \\ &= \int_{\mathbb{R}} x(t)e^{-2\pi i(\nu-\nu_0)t}dt \\ &= X(\nu - \nu_0)\end{aligned}$$

Multiplying a signal by a complex exponential corresponds to shifting the frequencies in its Fourier transform.

Using Theorem 4.8.16, we can compute the Fourier transform of $x(t) = e^{2\pi i \nu t}$ knowing from Eq (4.23) that $\mathcal{F}(1) = \delta$.

$$\begin{aligned}\mathcal{F}\big(e^{2\pi i \nu_0 t}\big) &= \mathcal{F}\big(1 \times e^{2\pi i \nu_0 t}\big) \\ &= \delta(\nu - \nu_0) \qquad (4.28)\end{aligned}$$

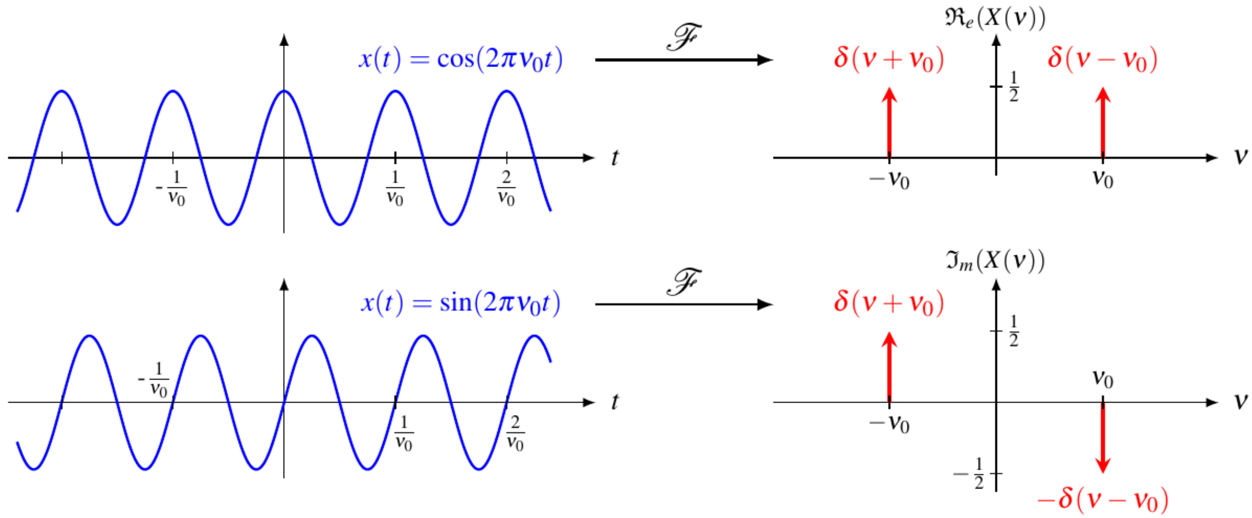


Figure 4.12: [Top] Fourier Transform of cos. [Bottom] Fourier Transform of sin

4.9 Fourier transforms of cos and sin

We'd like to compute the Fourier transforms of $t \mapsto \cos(2\pi\nu_0 t)$ and $t \mapsto \sin(2\pi\nu_0 t)$, with $\nu_0 = \frac{1}{T_0}$, the fundamental frequency.

Note that $t \mapsto \cos(2\pi\nu_0 t)$ and $t \mapsto \sin(2\pi\nu_0 t) \notin \mathcal{L}^1(\mathbb{R})$; nevertheless, we can compute their Fourier transforms in Theorems 4.9.1 and 4.9.2.

Theorem 4.9.1 (*Fourier transform of* cos)
As shown in Figure 4.12 [Top],

$$\mathcal{F}(\cos(2\pi\nu_0 t)) = \frac{1}{2}(\delta(\nu+\nu_0) + \delta(\nu-\nu_0))$$

Proof.

$$\begin{aligned}
\cos(2\pi\nu_0 t) &= \frac{1}{2}\left(e^{-2\pi i\nu_0 t} + e^{2\pi i\nu_0 t}\right) && \text{Euler's formula} \\
\mathcal{F}(\cos(2\pi\nu_0 t)) &= \mathcal{F}\left(\frac{1}{2}\left(e^{-2\pi i\nu_0 t} + e^{2\pi i\nu_0 t}\right)\right) \\
&= \frac{1}{2}\mathcal{F}\left(e^{-2\pi i\nu_0 t}\right) + \frac{1}{2}\mathcal{F}\left(e^{2\pi i\nu_0 t}\right) && \text{; by Theorem 4.8.1} \\
&= \frac{1}{2}(\delta(\nu+\nu_0) + \delta(\nu-\nu_0)) && \text{; by (4.28)}
\end{aligned}$$

The Fourier transform of the sin function is similar to that of the cos function.

Theorem 4.9.2 (*Fourier transform of* sin)
As shown in Figure 4.12 [Bottom],

$$\mathcal{F}(\sin(2\pi\nu_0 t)) = \frac{i}{2}(\delta(\nu+\nu_0) - \delta(\nu-\nu_0))$$

Proof.

$$\begin{aligned}
\sin(2\pi\nu_0 t) &= \frac{i}{2}\left(e^{-2\pi i\nu_0 t} - e^{2\pi i\nu_0 t}\right) && \text{Euler's formula}\\
\mathcal{F}(\sin(2\pi\nu_0 t)) &= \mathcal{F}\left(\frac{i}{2}\left(e^{-2\pi i\nu_0 t} - e^{2\pi i\nu_0 t}\right)\right)\\
&= \frac{i}{2}\mathcal{F}\left(e^{-2\pi i\nu_0 t}\right) - \frac{i}{2}\mathcal{F}\left(e^{2\pi i\nu_0 t}\right) && \text{; by Theorem 4.8.1}\\
&= \frac{i}{2}(\delta(\nu+\nu_0) - \delta(\nu-\nu_0)) && \text{; by (4.28)}
\end{aligned}$$

When computing the Fourier transforms of cos and sin (Theorems 4.9.1 and 4.9.2, we find again that if x is real and even, cos, then X is real and even; and if x is real and odd, then X is purely imaginary and odd.

In Examples 4.9.3 and 4.9.4 we develop $\mathcal{F}(\sin)$ two different ways.

Example 4.9.3 (*Alternative derivation of* $\mathcal{F}(\sin(2\pi\nu_0 t))$)

Show

$$\mathcal{F}(\sin(2\pi\nu_0 t)) = \frac{1}{2i}[\delta(\nu-\nu_0) - \delta(\nu+\nu_0)],$$

using:

$$\begin{aligned}
\sin(t) &= -\frac{d}{dt}\cos(t)\\
\mathcal{F}\left(\frac{d}{dt}x(t)\right) &= 2\pi i\nu X(\nu) && \text{Theorem 4.8.11}\\
\delta(t_0 - t)x(t) &= \delta(t_0 - t)x(t_0) && \text{Sifting Theorem (4.14) page 103}
\end{aligned}$$

Proof.

$$\begin{aligned}
\mathcal{F}(\sin(2\pi\nu_0 t)) &= \mathcal{F}\left(-\frac{d}{dt}\left(\frac{1}{2\pi\nu_0}\cos(2\pi\nu_0 t)\right)\right)\\
&= \frac{-1}{2\pi\nu_0}\mathcal{F}\left(\frac{d}{dt}(\cos(2\pi\nu_0 t))\right)\\
&= \left(\frac{-1}{2\pi\nu_0}\right)\underbrace{(2\pi i\nu)\ \mathcal{F}(\cos(2\pi\nu_0 t))}_{\text{By Theorem 4.8.11}}\\
&= \left(\frac{-i\nu}{\nu_0}\right)\left(\frac{1}{2}\right)\underbrace{[\delta(\nu-\nu_0) + \delta(\nu+\nu_0)]}_{\text{By Theorem 4.9.1}}\\
&= \frac{-i}{2}\left[\delta(\nu-\nu_0)\left(\frac{\nu}{\nu_0}\right) + \delta(\nu+\nu_0)\left(\frac{\nu}{\nu_0}\right)\right]\\
&= \frac{-i}{2}\left[\delta(\nu-\nu_0)\left(\frac{\nu_0}{\nu_0}\right) + \underbrace{\delta(\nu+\nu_0)\left(\frac{-\nu_0}{\nu_0}\right)}_{\text{Sifting Theorem}}\right]\\
&= \frac{1}{2i}\left[\delta(\nu-\nu_0) - \delta(\nu+\nu_0)\right]
\end{aligned}$$

Example 4.9.4 (***Alternative derivation of*** $\mathcal{F}(sin(2\pi\nu_0 t))$)

Show

$$\mathcal{F}(\sin(2\pi\nu_0 t)) = \frac{1}{2i}[\delta(\nu-\nu_0) - \delta(\nu+\nu_0)]\,,$$

using

$$\begin{aligned}
\sin(t) &= \cos\Big(t - \frac{\pi}{2}\Big) && \\
\mathcal{F}(x(t-t_0)) &= X(\nu)e^{-2\pi i\nu t_0} && \text{Theorem 4.8.14} \\
\delta(t_0-t)x(t) &= \delta(t_0-t)x(t_0) && \text{Sifting Theorem (4.14) page 103}
\end{aligned}$$

Proof.

$$\begin{aligned}
\mathcal{F}(\sin(2\pi\nu_0 t)) &= \mathcal{F}\Big(\cos\Big(2\pi\nu_0 t - \frac{\pi}{2}\Big)\Big) \\
&= \mathcal{F}\left(\cos\left(2\pi\nu_0 t - \left(\frac{2\pi\nu_0}{2\pi\nu_0}\right)\Big(\frac{\pi}{2}\Big)\right)\right) \\
&= \mathcal{F}\left(\cos\left(2\pi\nu_0\left(t - \frac{1}{4\nu_0}\right)\right)\right) \\
&= \underbrace{\mathcal{F}(\cos(2\pi\nu_0 t))e^{-2\pi i\nu\times\frac{1}{4\nu_0}}}_{\text{By Theorem 4.8.14}} \\
&= \underbrace{\frac{1}{2}\left[\delta(\nu-\nu_0) + \delta(\nu+\nu_0)\right]}_{\text{By Theorem 4.9.1}} e^{-\frac{\pi i\nu}{2\nu_0}} \\
&= \frac{1}{2}\left[\delta(\nu-\nu_0)e^{-\frac{\pi i\nu}{2\nu_0}} + \delta(\nu+\nu_0)e^{-\frac{\pi i\nu}{2\nu_0}}\right] \\
&= \frac{1}{2}\left[\delta(\nu-\nu_0)\underbrace{e^{-\frac{\pi i\nu_0}{2\nu_0}}}_{\text{Sifting Theorem}} + \delta(\nu+\nu_0)\underbrace{e^{+\frac{\pi i\nu_0}{2\nu_0}}}_{\text{Sifting Theorem}}\right] \\
&= \frac{1}{2}\left[\delta(\nu-\nu_0)\underbrace{e^{-\frac{\pi i}{2}}}_{=-i} + \delta(\nu+\nu_0)\underbrace{e^{+\frac{\pi i}{2}}}_{=i}\right] \\
&= \frac{1}{2}\left[-i\delta(\nu-\nu_0) + i\delta(\nu+\nu_0)\right] \\
&= \frac{i}{2}\left[-\delta(\nu-\nu_0) + \delta(\nu+\nu_0)\right] \\
&= \left(\frac{i}{i}\right)\left(\frac{i}{2}\right)\left[-\delta(\nu-\nu_0) + \delta(\nu+\nu_0)\right] \\
&= \frac{-1}{2i}\left[-\delta(\nu-\nu_0) + \delta(\nu+\nu_0)\right] \\
&= \frac{1}{2i}\left[\delta(\nu-\nu_0) - \delta(\nu+\nu_0)\right]
\end{aligned}$$

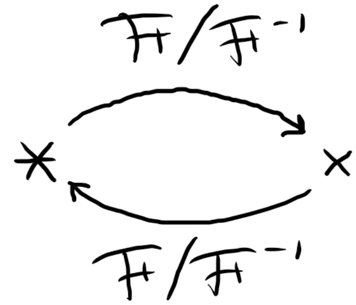


Figure 4.13: Algebraic product and convolution product are duals with respect to the Fourier Transform.

4.10 Plancherel's Theorem

Theorem 4.10.1 (*Plancherel*)

$$\mathcal{F}(x * y) = \mathcal{F}(x)\mathcal{F}(y) \tag{4.29}$$
$$\mathcal{F}(xy) = \mathcal{F}(x) * \mathcal{F}(y) \tag{4.30}$$
$$\mathcal{F}^{-1}(X * Y) = \mathcal{F}^{-1}(X)\mathcal{F}^{-1}(Y) \tag{4.31}$$
$$\mathcal{F}^{-1}(XY) = \mathcal{F}^{-1}(X) * \mathcal{F}^{-1}(Y) \tag{4.32}$$

The tricky part of this proof is that we have an integral inside another integral and we need two swap the order of integration.

Proof of (4.29). We wish to show

$$\mathcal{F}(x * y) = \mathcal{F}(x)\mathcal{F}(y)\,.$$

Since we will need to integrate a function of t, we need to use different variables to avoid conflating the same variable for two different entities. Let's use this definition of convolution

$$(x * y)(t) = \int_{\mathbb{R}} x(s)y(t-s)ds\,,$$

and this definition for Fourier Transform.

$$(\mathcal{F}(x))(\nu) = \int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt\,.$$

$$\begin{aligned}\mathcal{F}(x * y)(\nu) &= \int_{\mathbb{R}} (x * y)(t)\; e^{-2\pi i\nu t}dt \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} x(s)y(t-s)ds\right)e^{-2\pi i\nu t}dt\end{aligned}$$

The Fubini/Tonelli theorem allows us to interchange the integrals, because $x, y \in \mathcal{L}^2$ and

$$\int_{\mathbb{R}}\int_{\mathbb{R}} |x(s)||y(t-s)|ds\; dt = \|x\|_{\mathcal{L}^1}\|y\|_{\mathcal{L}^1} < \infty\,,$$

where

$$\|x\|_{\mathcal{L}^1} = \int_{\mathbb{R}} |x(t)|dt\,.$$

So we have

$$\begin{aligned}
\mathcal{F}(x*y)(\nu) &= \int_{\mathbb{R}} x(s) \underbrace{\left(\int_{\mathbb{R}} y(t-s)e^{-2\pi i\nu t}dt\right)}_{\mathcal{F}(y(t-s))} ds \\
&= \int_{\mathbb{R}} x(s)\mathcal{F}(y(t))e^{-2\pi i\nu s}ds \\
&= \underbrace{\left(\int_{\mathbb{R}} x(s)e^{-2\pi i\nu s}ds\right)}_{\mathcal{F}(x(s))} Y(\nu) \\
&= X(\nu)Y(\nu)
\end{aligned}$$

Reminder 4.10.2 (*closure of $\mathcal{L}^1$ under convolution*)

Recall that if $x, y \in \mathcal{L}^1(\mathbb{R})$ then $z = x * y \in \mathcal{L}^1(\mathbb{R})$.

If we denote $Z = \mathcal{F}(z)$, $X = \mathcal{F}(x)$, and $Y = \mathcal{F}(y)$, then

$$\begin{aligned}
\mathcal{F}(x*y) &= \mathcal{F}(x)\mathcal{F}(y) \\
Z(\nu) &= X(\nu)Y(\nu)\,.
\end{aligned}$$

The Fourier transform transforms a convolution product into an algebraic product, but this works also with a product of signals, as alluded to in Figure 4.13. From a previous result $x, y \in \mathcal{L}^1(\mathbb{R}) \implies xy \in \mathcal{L}^1(\mathbb{R})$. If we denote $z(t) = x(t)y(t)$, then

$$\begin{aligned}
\mathcal{F}(xy) &= \mathcal{F}(x) * \mathcal{F}(y) \\
Z(\nu) &= (X * Y)(\nu)
\end{aligned}$$

This fact is a major result which in and of itself justifies the practical interest in the Fourier transform. For example, in digital signal processing, we represent x and y as follows.

$$\begin{aligned}
x &= [x(0), x(1), \ldots, x(N)] \\
\text{and } y &= [y(0), y(1), \ldots, y(N)]\,.
\end{aligned}$$

And we represent the discrete convolution as

$$(x*y)(n) = \sum_{k=0}^{N} x(k)y(n-k)\,,$$

which is the discrete version of

$$(x*y)(t) = \int_{\mathbb{R}} x(\tau)y(t-\tau)dt\,.$$

The discrete convolution would seem to have

- complexity $O(N)$ for computing a single value of the convolution, and thus
- complexity $O(N^2)$ for computing all the values.

However, $(x*y) = \mathcal{F}^{-1}(\mathcal{F}(x)\mathcal{F}(y))$. Since, thanks to the Fast Fourier Transform (FFT), we can compute the Fourier transform, $X = \mathcal{F}(x)$ with complexity $O(N \log N)$

The computation $(x*y)$ can be done with $O(N \log N)$ complexity by computing the Fourier transform and applying Plancherel's theorem.

4.11 Parseval's Identity

Theorem 4.11.1 (*Parseval's Identity*)

If $x, y \in \mathcal{L}^2(\mathbb{R})$ and $X = \mathcal{F}(x) \in \mathcal{L}^2(\mathbb{R})$, $Y = \mathcal{F}(y) \in \mathcal{L}^2(\mathbb{R})$, then

$$\int_{\mathbb{R}} x(t)\,\overline{y(t)}\,dt = \int_{\mathbb{R}} X(\nu)\,\overline{Y(\nu)}\,d\nu\,.$$

In particular, if $x = y$, then

$$\int_{\mathbb{R}} |x(t)|^2 dt = \int_{\mathbb{R}} |X(\nu)|^2 d\nu\,.$$

Theorem 4.11.1 is proven in Section 4.12.

This is the Fourier transform analog of the result we have already obtained for the Fourier series:

$$\frac{1}{T}\int_0^T |x(t)|^2 dt = \sum_{n\in\mathbb{Z}} |c_n|^2\,.$$

Thus the energy is conserved when we pass from the time domain to the frequency domain.

$$\int_{\mathbb{R}} |x(t)|^2 dt = E_x \tag{4.33}$$

$$= \int_{\mathbb{R}} |X(\nu)|^2 d\nu \tag{4.34}$$

Equation (4.33) is the computation of signal energy which we've known from the beginning of the course. Equation (4.34) is the energy in the frequency domain.

Definition 4.11.2 (*spectral energy density*)

The function, $\nu \mapsto |X(\nu)|^2$, is called the *spectral energy density*.

Reminder 4.11.3 (*modulus squared and Hermitian symmetry*)

Recall that $|X(\nu)|^2 = X(\nu)\,\overline{X(\nu)}$, and that if x is a real valued signal, then X is Hermitian symmetric, $X(-\nu) = \overline{X(\nu)}$.

Therefore, for a real valued signal:

$$|X(\nu)|^2 = X(\nu)\,\overline{X(\nu)} = X(\nu)X(-\nu) = |X(-\nu)|^2\,,$$

the spectral energy density is also an even function.

4.12 Connection with Correlation

Recall that

$$\Gamma_{xy}(\tau) = \langle x(t), y(t-\tau)\rangle = \int_{\mathbb{R}} x(t)\,\overline{y(t-\tau)}\,dt\,.$$

What is the Fourier transform of Γ_{xy}?

$$\begin{aligned}\mathcal{F}(\Gamma_{xy}(\tau)) &= \int_{\mathbb{R}} \Gamma_{xy}(\tau)e^{-2\pi i\nu\tau}d\tau \\ &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} x(t)\,\overline{y(t-\tau)}\,dt\right)e^{-2\pi i\nu\tau}d\tau \qquad (4.35)\end{aligned}$$

Sadly, Equation (4.35) is not very enlightening. However, Plancherel's theorem gives us a different way of computing $\mathcal{F}(\Gamma_{xy}(\tau))$.

Reminder 4.12.1 (*cross-correlation as a convolution*)

Recall the link between cross-correlation and convolution. Recall also, $y^-(t) = y(-t)$:

$$\Gamma_{xy}(\tau) = (x * \overline{y^-})(\tau) \qquad (4.36)$$

This equality, (4.36), leads to Theorem 4.12.2. Lemma 4.8.7 will be useful to prove Theorem 4.12.2.

Theorem 4.12.2 (*Fourier transform of cross-correlation*)

$$\mathcal{F}(\Gamma_{xy}(\tau)) = X(\nu)\overline{Y(\nu)}\,.$$

Proof.

$$\begin{aligned}\mathcal{F}(\Gamma_{xy}) &= \mathcal{F}\Big(x * \overline{y^-}\Big) && \text{by (4.36)} \\ &= \mathcal{F}(x)\ \mathcal{F}\Big(\overline{y^-}\Big) && \text{by Plancherel} \\ &= X(\nu)\ \overline{Y(\nu)} && \text{by Lemma 4.8.7}\end{aligned}$$

Definition 4.12.3 (*cross spectral energy density*)

We define γ as follows

$$\gamma_{xy}(\nu) = \mathcal{F}(\Gamma_{xy}(\tau))$$

We call the function $\nu \mapsto \gamma_{xy}(\nu)$ the *cross spectral energy density*

Corollary 4.12.4 (*Fourier transform of the autocorrelation (Wiener–Khinchin)*)

$$\mathcal{F}(\Gamma_{xx}) = \gamma_{xx}(\nu) = |X(\nu)|^2$$

Proof. Let $x = y$ and apply Theorem 4.12.2.

$$\begin{aligned}\mathcal{F}(\Gamma_{xx}) &= \gamma_{xx}(\nu) \\ &= X(\nu)\overline{X(\nu)} \\ &= |X(\nu)|^2\end{aligned}$$

Thus, the Fourier transform of the autocorrelation function is the spectral energy density.

It is worth emphasizing that in order to compute Γ_{xx} we may proceed in two ways.

- We may directly compute

$$\Gamma_{xx}(\tau) = \int_{\mathbb{R}} x(t)\,\overline{x(t-\tau)}\,dt\,.$$

- We may compute the inverse Fourier transform of the spectral energy density

$$\Gamma_{xx}(\tau) = \mathcal{F}^{-1}(\gamma(\nu))\,.$$

Computing Γ_{xx} via the inverse Fourier transform offers the same improvement in complexity ($O(N \log N)$ vs $O(N^2)$) as for computing the convolution of digital signals.

We get one more somewhat surprising result from the computation of the autocorrelation.

Reminder 4.12.5 (*energy from autocorrelation at zero*)

Recall:

$$\Gamma_{xx}(0) = E_x = \int_{\mathbb{R}} |x(t)|^2 dt$$

$$\Gamma_{xx}(t) = \mathcal{F}^{-1}(\gamma_{xx}(\nu)) = \int_{\mathbb{R}} |X(\nu)|^2 e^{2\pi i \nu t} d\nu$$

Now we can prove Parseval's Identity, Theorem 4.11.1.

Proof of Theorem 4.11.1.

$$\begin{aligned}
\int_{\mathbb{R}} |x(t)|^2 dt &= \Gamma_{xx}(0) \\
&= \Gamma_{xx}(t)\Big]_{t=0} \\
&= \int_{\mathbb{R}} |X(\nu)|^2 e^{2\pi i \nu t} d\nu\Big]_{t=0} \\
&= \int_{\mathbb{R}} |X(\nu)|^2 e^{2\pi i \nu \times 0} d\nu \\
&= \int_{\mathbb{R}} |X(\nu)|^2\,.
\end{aligned}$$

4.13 Bernstein's Theorem

Definition 4.13.1 (***band limited signal***)

A signal, x with Fourier transform $X = \mathcal{F}(x)$, is said to be *band limited* if there exists $B \in \mathbb{R}$, $B \geq 0$ such that

$$|X(\nu)| = 0 \quad \forall |\nu| \geq B\,.$$

A band limited signal is therefore a signal whose Fourier transform has bounded support.

Theorem 4.13.2 (***Finite energy of band-limited signals***)

Suppose x is a band limited signal, with band B, and that the Fourier transform of x is bounded, then x is a finite energy signal.

Proof. Let $M < \infty$ be a bound of X.

$$\begin{aligned}
E_x &= \int_{\mathbb{R}} |x(t)|^2 dt \\
&= \int_{\mathbb{R}} \underbrace{|X(\nu)|^2 d\nu}_{\text{by (4.34)}} \\
&= \int_{-\infty}^{-B} \underbrace{|X(\nu)|^2}_{=0 \text{ by Def 4.13.1}} d\nu + \int_{-B}^{B} |X(\nu)|^2 d\nu + \int_{B}^{\infty} \underbrace{|X(\nu)|^2}_{=0 \text{ by Def 4.13.1}} d\nu \\
&= \int_{-B}^{B} \underbrace{|X(\nu)|^2}_{<M^2} d\nu \\
&\leq 2BM^2 < +\infty
\end{aligned}$$

Lemma 4.13.3 (***Derivatives of band-limited signals***)

Suppose x is a band limited signal, with band B, and that the Fourier transform of x is bounded, then, for every integer $n \geq 1$,

$$\frac{d^n}{dt^n} x(t) = (2\pi i)^n \int_{-B}^{B} \nu^n X(\nu) e^{2\pi i \nu t} d\nu$$

For the proof of Lemma 4.13.3 we'll make use of the Leibniz integral rule. Many proofs thereof are available for this rule, for example Wikipedia.

$$\frac{d}{dx} \int_a^b f(x,t) dt = \int_a^b \frac{\partial}{\partial x} f(x,t) dt \tag{4.37}$$

Proof by induction.

Base case: $n = 1$

$$
\begin{aligned}
\frac{d}{dt}x(t) &= \frac{d}{dt}\left(\int_{-B}^{B} X(\nu)e^{2\pi i\nu t}d\nu\right) \\
&= \int_{-B}^{B} \frac{\partial}{\partial t} \underbrace{X(\nu)}_{\text{constant w.r.t } t} e^{2\pi i\nu t}d\nu && \text{; by (4.37)} \\
&= \int_{-B}^{B} X(\nu) \underbrace{\frac{\partial}{\partial t}e^{2\pi i\nu t}}_{2\pi i\nu\, e^{2\pi i\nu t}} d\nu \\
&= 2\pi i \int_{-B}^{B} \nu X(\nu)e^{2\pi i\nu t}d\nu && (4.38)
\end{aligned}
$$

Inductive case: Suppose for some $n \geq 1$ we have

$$
\frac{d^n}{dt^n}x(t) = (2\pi i)^n \int_{-B}^{B} \nu^n X(\nu)e^{2\pi i\nu t}d\nu\,.
$$

We can examine $\frac{d^{n+1}}{dt^{n+1}}x(t)$.

$$
\begin{aligned}
\frac{d^{n+1}}{dt^{n+1}}x(t) &= \frac{d}{dt}\left(\frac{d^n}{dt^n}x(t)\right) \\
&= \frac{d}{dt}\left((2\pi i)^n \int_{-B}^{B} \nu^n X(\nu)e^{2\pi i\nu t}d\nu\right) && \text{by inductive hypothesis} \\
&= (2\pi i)^n \int_{-B}^{B} \frac{\partial}{\partial t}\nu^n X(\nu)e^{2\pi i\nu t}d\nu && \text{by (4.37)} \\
&= (2\pi i)^n \int_{-B}^{B} \nu^n X(\nu)\frac{\partial}{\partial t}e^{2\pi i\nu t}d\nu \\
&= (2\pi i)^n \int_{-B}^{B} \nu^n X(\nu)\ (2\pi i\nu)e^{2\pi i\nu t}d\nu \\
&= (2\pi i)^{n+1} \int_{-B}^{B} \nu^{n+1} X(\nu)e^{2\pi i\nu t}d\nu
\end{aligned}
$$

Theorem 4.13.4 (*Bernstein's theorem (first derivative)*)

Let x be a differentiable, band limited signal with band, B. Let X denote Fourier transform of x.

$$
\left|\frac{d}{dt}x(t)\right| \leq 2\pi B \int_{-B}^{B} |X(\nu)|d\nu
$$

Proof. Since x is band limited,

$$
x(t) = \int_{-B}^{B} X(\nu)e^{2\pi i\nu t}d\nu \tag{4.39}
$$

The derivative of x at t, $\frac{d}{dt}x(t)$ gives the instantaneous rate of change of x at time, t.

$$\begin{aligned}
\left|\frac{d}{dt}x(t)\right| &= \left|2\pi i\int_{-B}^{B}\nu X(\nu)e^{2\pi i\nu t}d\nu\right| && ;\text{ by (4.38)}\\
&= 2\pi\left|\int_{-B}^{B}\nu X(\nu)e^{2\pi i\nu t}d\nu\right|\\
&\leq 2\pi\int_{-B}^{B}\left|\nu X(\nu)e^{2\pi i\nu t}\right|d\nu\\
&\leq 2\pi\int_{-B}^{B}|\nu|\cdot|X(\nu)|\cdot\underbrace{\left|e^{2\pi i\nu t}\right|}_{=1}d\nu\\
&\leq 2\pi B\int_{-B}^{B}|X(\nu)|d\nu && ;\ |\nu|\leq B\quad\forall\nu\in[-B,B]
\end{aligned}$$

This result in Theorem 4.13.4 generalizes to $\frac{d^n}{dt^n}x(t)$ as shown in Theorem 4.13.5. To prove Theorem 4.13.5, we will make use of Lemma 4.13.3

Theorem 4.13.5 (*Bernstein's theorem (n^{th} derivative)*)
Let $n \geq 1$ be an integer, and let x be an n-times differentiable, band limited signal with band, B. Let X denote Fourier transform of x.

$$\left|\frac{d^n}{dt^n}x(t)\right| \leq (2\pi B)^n\int_{-B}^{B}|X(\nu)|d\nu\,.$$

Proof. Since x is band limited,

$$x(t) = \int_{-B}^{B}X(\nu)e^{2\pi i\nu t}d\nu$$

$$\begin{aligned}
\frac{d^n}{dt^n}x(t) &= (2\pi i)^n\int_{-B}^{B}\nu^n X(\nu)e^{2\pi i\nu t}d\nu && \text{by Lemma 4.13.3}\\
\left|\frac{d^n}{dt^n}x(t)\right| &= \left|(2\pi i)^n\int_{-B}^{B}\nu^n X(\nu)e^{2\pi i\nu t}d\nu\right|\\
&\leq |(2\pi)^n|\ |i^n|\ \int_{-B}^{B}|\nu^n|\ |X(\nu)|\ \left|e^{2\pi i\nu t}\right|d\nu\\
&\leq (2\pi)^n\ \int_{-B}^{B}|B^n|\ |X(\nu)|d\nu\\
&\leq (2\pi B)^n\ \int_{-B}^{B}|X(\nu)|d\nu
\end{aligned}$$

To understand Bernstein's theorem: the maximum rate of change of a band limited signal is bounded by a quantity which depends on the value of the band, B.

- If B is small, *i.e.*, narrow support, then the maximum rate of change of x is weak. This means x only contains low frequencies.
- If B is large, *i.e.*, wide support, then the maximum rate of change of x is large. This means x contains high frequencies.

4.14 Review

- The Fourier transform generalizes the Fourier series for non-periodic signals—representing as a superposition of a continuum of harmonics.
- If $x \in \mathcal{L}^1(\mathbb{R})$, then the Fourier transform of x is

$$X(\nu) = \int_{\mathbb{R}} x(t)e^{-2\pi i\nu t}dt\,.$$

  If $X \in \mathcal{L}^1(\mathbb{R})$, then the inverse Fourier transform of X is

$$x(t) = \int_{\mathbb{R}} X(\nu)e^{2\pi i\nu t}d\nu\,.$$

- Some useful Fourier transforms are given in Figure 4.14.
- Symmetric properties
  - x real $\implies$ X is symmetric, Hermitian:

$$X(-\nu) = \overline{X(\nu)}$$

  - x even $\implies$ $\mathcal{F}$ is involutive: $\mathcal{F} = \mathcal{F}^{-1}$

$$x = \mathcal{F}(\mathcal{F}(x))$$

  - x real and even $\implies$ X is real and even.
  - x real and odd $\implies$ X is purely imaginary and odd
- Duality time-shifting, phase-shifting.

$$\mathcal{F}(x(t-t_0)) = X(\nu)e^{-2\pi i\nu t_0}$$
$$\mathcal{F}\big(x(t)e^{2\pi i\nu_0 t}\big) = X(\nu-\nu_0)$$

- Interpretation of modulus and phase: $X(\nu) \in \mathbb{C}$ in general

$$X(\nu) = |X(\nu)|e^{i\phi(X(\nu))}$$

  - $|X(\nu)|$ is the spectrum, distribution of the amplitude of the frequencies.
  - $\phi(X(\nu))$ is the phase, spatial localization of the frequencies.
- Duality of algebraic product and convolution product: Theorem of Plancherel.

$$\mathcal{F}((x*y)(t)) = X(\nu)Y(\nu)$$
$$\mathcal{F}(x(t)y(t)) = (X*Y)(\nu)$$

- Parseval's identity: conservation of energy between frequency and time domains.

$$\int_{\mathbb{R}} |x(t)|^2dt = \int_{\mathbb{R}} |X(\nu)|^2d\nu$$

- Link between spectral energy density γ_{xx} and autocorrelation, Γ_{xx}:

$$\gamma_{xx} = |X(\nu)|^2 = \mathcal{F}(\Gamma_{xx}(\tau))$$

- Bernstein's Theorem:

$$\left|\frac{d}{dt}x(t)\right| \le 2\pi B\int_{-B}^{B}|X(\nu)|d\nu$$

Signal $x(t)$	**Fourier Transform** $X(\nu)$
$\Pi_T(t)$	$T\,\mathrm{sinc}(\pi T\nu)$
$T\,\mathrm{sinc}(\pi Tt)$	$\Pi_T(\nu)$
$\delta(t-t_0)$	$\exp(-2\pi i\nu t_0)$
$\delta(t)$	1
1	$\delta(\nu)$
$\exp(2\pi i\nu_0 t)$	$\delta(\nu-\nu_0)$
$\cos(2\pi\nu_0 t)$	$\frac{1}{2}(\delta(\nu+\nu_0)+\delta(\nu-\nu_0))$
$\sin(2\pi\nu_0 t)$	$\frac{i}{2}(\delta(\nu+\nu_0)-\delta(\nu-\nu_0))$
$\Gamma_{xy}(\tau)$	$X(\nu)\overline{Y(\nu)}$
$\Gamma_{xx}(\tau)$	$\lvert X(\nu)\rvert^2$
$\alpha x(t)+y(t)$	$\alpha X(\nu)+Y(\nu)$
$x(t)\exp(2\pi i\nu_0 t)$	$X(\nu-\nu_0)$
$x(t-t_0)$	$X(\nu)\exp(-2\pi i\nu t_0)$
$\frac{d}{dt}x(t)$	$2\pi i\nu X(\nu)$
$t\ x(t)$	$\frac{i}{2\pi}\frac{d}{d\nu}X(\nu)$

Figure 4.14: Useful Fourier Transforms

Chapter 5

Analog-Digital Conversion

Chapter Objectives

- Model the sampling process using the Dirac comb, and compute the Fourier transform of a sampled signal.
- State and apply the Shannon-Nyquist sampling theorem to determine an adequate sampling frequency for a band-limited signal.
- Recognize aliasing (spectral folding) when the Nyquist condition is violated, and reconstruct a signal from its samples via the Shannon-Whittaker interpolation formula when it is respected.
- Quantize a sampled signal (mid-riser and mid-tread strategies) and compute the resulting quantization step and rounding error.
- Compute the signal-to-noise ratio (SNR) of a quantized signal as a function of the number of encoding bits.

5.1 Review of Mathematical Relations

There are several relations which will be useful for the development in this chapter. The following is a brief review of these properties.

Reminder 5.1.1 (*Fourier Transform and Series*)

$$
\begin{aligned}
X(\nu) &= \int_{\mathbb{R}} x(t)\exp(-2\pi i\nu t)\,\mathrm{d}t \\
x(t) &= \int_{\mathbb{R}} X(\nu)\exp(2\pi i\nu t)\,\mathrm{d}\nu \\
x(t) &= \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi i n}{T}\,t\right) \\
c_n &= \frac{1}{T}\int_0^T x(t)\exp\left(-\frac{2\pi i n}{T}\,t\right)\mathrm{d}t\,.
\end{aligned}
$$

Reminder 5.1.2 (***Properties of convolution***)

$$(x * y)(\tau) = \int_{\mathbb{R}} x(t)y(\tau - t)\,\mathrm{d}t$$

$$x * y = y * x \qquad \text{Eq (2.2)}$$

$$x * (y + \lambda z) = x * y + \lambda(x * z) \qquad \text{Eq (2.5)}$$

Reminder 5.1.3 (***Properties of Fourier Transform***)

$$\mathcal{F}(\alpha x(t) + \beta y(t)) = \alpha\mathcal{F}(x(t)) + \beta\mathcal{F}(y(t)) \qquad \text{Theorem 4.8.1}$$

$$\mathcal{F}^{-1}(\alpha x(t) + \beta y(t)) = \alpha\mathcal{F}^{-1}(x(t)) + \beta\mathcal{F}^{-1}(y(t))$$

Reminder 5.1.4 (***Plancherel's Theorem 4.10.1***)

$$\mathcal{F}(x * y) = \mathcal{F}(x)\mathcal{F}(y) \qquad \text{Eq (4.29)}$$

$$\mathcal{F}(xy) = \mathcal{F}(x) * \mathcal{F}(y) \qquad \text{Eq (4.30)}$$

$$\mathcal{F}^{-1}(X * Y) = \mathcal{F}^{-1}(X)\mathcal{F}^{-1}(Y) \qquad \text{Eq (4.31)}$$

$$\mathcal{F}^{-1}(XY) = \mathcal{F}^{-1}(X) * \mathcal{F}^{-1}(Y) \qquad \text{Eq (4.32)}$$

Reminder 5.1.5 (***Cardinal Sine***)

$$\operatorname{sinc}(t) = \begin{cases} \frac{\sin(t)}{t} & ; \text{if } t \neq 0 \\ 1 & ; \text{if } t = 0 \end{cases}$$

$$\Pi_T(t) = \begin{cases} 1 & ; \text{ if } -T/2 \leq t \leq T/2 \\ 0 & ; \text{ otherwise} \end{cases}$$

$$\mathcal{F}(\Pi_T(t)) = T\operatorname{sinc}(\pi T\nu)$$

$$\mathcal{F}(T\operatorname{sinc}(\pi T t)) = \Pi_T(\nu)$$

Reminder 5.1.6 (***Properties of Dirac Delta***)

Definition 4.7.1

$$\delta(t) = \begin{cases} +\infty & ; \text{ if } t = 0 \\ 0 & ; \text{ otherwise} \end{cases} \qquad \textit{Eq (4.6)}$$

$$\int_{\mathbb{R}} \delta(t)\,\mathrm{d}t = 1 \qquad \textit{Eq (4.7)}$$

$$\int_a^b \delta(t - t_0)\,\mathrm{d}t = \begin{cases} 0 & \text{if } t_0 < a < b \\ 1 & \text{if } a < t_0 < b \\ 0 & \text{if } a < b < t_0 \end{cases} \qquad \text{Theorem 4.7.6}$$

Reminder 5.1.7 (***Sampling properties of Dirac Delta***)

$$x(t)\delta(t_0 - t) = x(t_0)\delta(t_0 - t) \qquad \text{Eq (4.14)}$$

$$x(t) * \delta(t - t_0) = x(t - t_0) \qquad \text{Theorem 4.7.5}$$

$$\int_{\mathbb{R}} \alpha\delta(t - t_0)dt = \alpha \qquad \text{Eq (4.11)}$$

$$\int_{\mathbb{R}} x(t)\delta(t - t_0)dt = x(t_0) \qquad \text{Theorem 4.7.2}$$

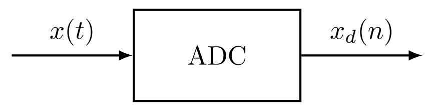


Figure 5.1: Analog-Digital converter ADC (CAN in French)

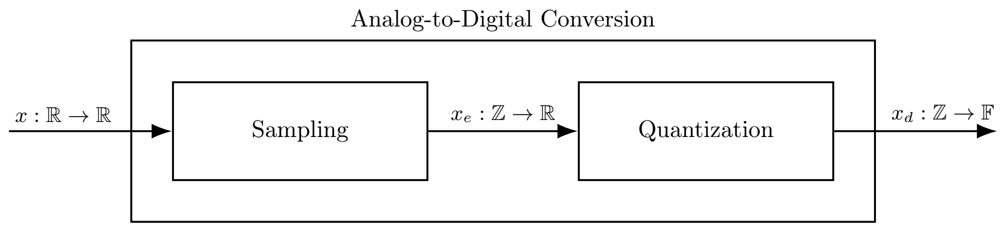


Figure 5.2: Analog-Digital conversion made up of sampling and quantization

5.2 Sampling vs Quantizing

Signals naturally emitted from measurable physical phenomena (pressure, current, voltage, light etc) are continuous with respect to their variables (time and position) and the set of (admissible) values those signals may obtain.

- It is necessary to discretize (render them finite) for storage in memory and for digital manipulation in finite memory.
- This is the role of an ADC analog-digital converter (CAN in French)

In Figure 5.1, $x : \mathbb{R} \to \mathbb{R}$ is a natural analog signal, continuous with respect to its input variables and output value, $x_d : \mathbb{Z} \to \mathbb{F}$ (the output of the ADC) is a discrete digital signal. The signal is sampled on an infinite sequence of input values ($\mathbb{Z}$), and $\mathbb{F}$ denotes a finite dictionary of admissible values.

As shown in Figure 5.2, the conversion takes place in two steps:

Sampling : discretizing with respect to the input variable:

$$x : \mathbb{R} \to \mathbb{R} \xrightarrow{sampling} x_e : \mathbb{Z} \to \mathbb{R}$$

Quantizing : discretizing with respect to the signal values:

$$x_e : \mathbb{Z} \to \mathbb{R} \xrightarrow{quantizing} x_d : \mathbb{Z} \to \mathbb{F}$$

5.3 Sampling

Sampling a continuous analog signal $x : \mathbb{R} \to \mathbb{R}$, consists of lifting the values of x at certain points of measurement $(t_n)_{n\in\mathbb{Z}}$ to produce a sequence of discrete values $(x(t_n))_{n\in\mathbb{Z}}$, called *samples*. Each measurement point $t_n, \forall n \in \mathbb{Z}$ corresponds to one sample $x(t_n)$, which may also be denoted $x(n)$.

Question: How can we discretize the variable $t \to (t_n)_{n\in\mathbb{Z}}$ without losing information or losing the least amount possible?

This question (falsely) implies that we already know how to define *information* and we already know how to quantify its loss.

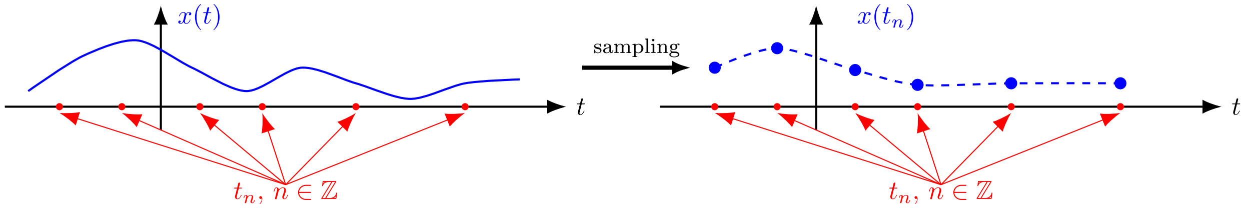


Figure 5.3: Using δ function for non-uniform sampling

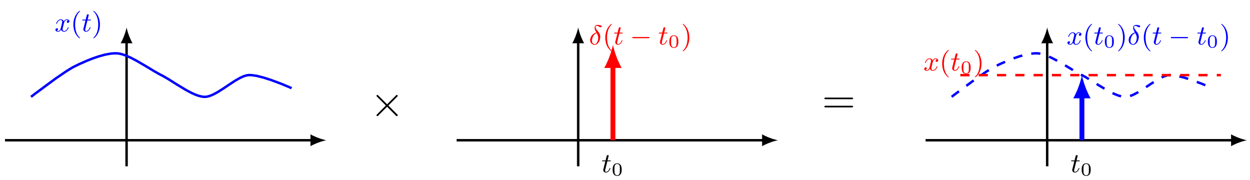


Figure 5.4: Using $\delta(t-t_0)$ to sample $x(t)$ at $t=t_0$

We will begin by simplifying our life a bit and take samples only at regularly spaced intervals.

$$
\begin{aligned}
&t_n = nT_e, \forall n \in \mathbb{Z} && \\
&T_e > 0 && \text{sampling period} \\
&\nu_e = \frac{1}{T_e} && \text{sampling frequency}
\end{aligned}
$$

The question above becomes: how to optimally choose T_e and thus ν_e.

To do this we must mathematically model the sampling operations. For this we will use the Dirac delta function.

Reminder 5.3.1 (*the Dirac delta function, for modeling sampling*)

Recall, $\delta : \mathbb{R} \to \mathbb{R}$ is defined to have the following properties

$$
t \mapsto \delta(t) = \begin{cases} +\infty & \text{if } t = 0 \\ 0 & \text{otherwise} \end{cases}
$$

$$
\int_{\mathbb{R}} \delta(t)dt = 1
$$

$$
x(t)\delta(t-t_0) = x(t_0)\delta(t-t_0) \tag{5.1}
$$

$$
\int_{\mathbb{R}} x(t)\delta(t-t_0)dt = x(t_0) \tag{5.2}
$$

Equation (5.1) means that we transfer the value of x at t_0 to the mass of δ. Equation (5.2) means that we effectively *read* the value of x at t_0 as shown in Figure 5.4.

And if we wish to lift the value of x at two instants, t_0 and t_1?

$$
\begin{aligned}
\underbrace{x(t)\delta(t-t_0)}_{\text{sample at } t_0} + \underbrace{x(t)\delta(t-t_1)}_{\text{sample at } t_1} &= x(t_0)\delta(t-t_0) + x(t_1)\delta(t-t_1) \\
&= x(t)(\delta(t-t_0) + \delta(t-t_1))
\end{aligned}
$$

And if we wish to lift a sample for $t = nT_e, \forall n \in \mathbb{Z}$? In a similar manner, we use a δ function at every

$t = nT_e \mapsto \delta(t - nT_e)$,

$$x(t) \left(\underbrace{\cdots + \underbrace{\delta(t+T_e)}_{\text{sample at } t=-T_e} + \underbrace{\delta(t)}_{t=0} + \underbrace{\delta(t-T_e)}_{t=T_e} + \underbrace{\delta(t-2T_e)}_{t=2T_e} + \cdots}_{\sum_{n=-\infty}^{+\infty} \delta(t-nT_e)} \right)$$

So the sampled signal can be represented as

$$x_e(t) = x(t) \sum_{n=-\infty}^{+\infty} \delta(t - nT_e)\,.$$

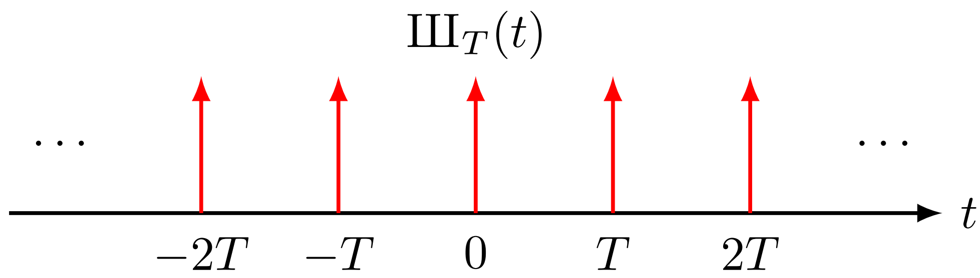


Figure 5.5: Visualization of Dirac comb: $Ш_T(t)$

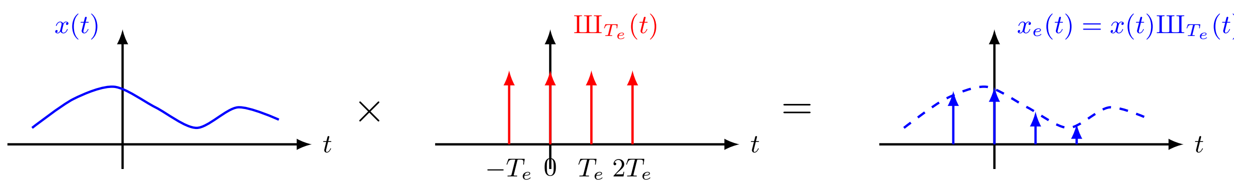


Figure 5.6: Sampling a signal, $x(t)$, using the Dirac comb, to generate $x_e(t) = x(t)Ш_{T_e}(t)$.

5.4 Dirac Comb

Definition 5.4.1 (*Dirac Comb*)

We define the *Dirac comb* as follows

$$Ш_T(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)\,.$$

We will often think of Ш as a function as shown in Figure 5.5, but we keep in the back of our minds that it is not a function, but rather a notation which is shorthand for a limiting process.

As illustrated in Figure 5.6, we can write a sampled signal as Definition 5.4.2.

Definition 5.4.2 (*sampled signal*)

$$\begin{aligned} x_e(t) &= x(t)Ш_{T_e}(t) \\ &= x(t) \sum_{n=-\infty}^{+\infty} \delta(t - nT_e) && \text{by Definition 5.4.1} \\ &= \sum_{n=-\infty}^{+\infty} x(nT_e)\delta(t - nT_e) && (5.3) \end{aligned}$$

But this still fails to answer the question of how to optimally choose T_e (and thus ν_e) for sampling the signal x.

If the sampling frequency is to small (if the sampling period is too large), Figure 5.7 (top), with respect to the variations in the signal, then all the variation in the signal will be lost. Under-sampling implies lost of information.

If the sampling frequency is too large (if the sampling period is too small), Figure 5.7 (bottom), then excellent resolution of the signal is retained. However, is it really necessary to retain so much information? Oversampling incurs wasted memory.

The sampling frequency/period must be linked to the maximal rate of change of the signal being sampled.

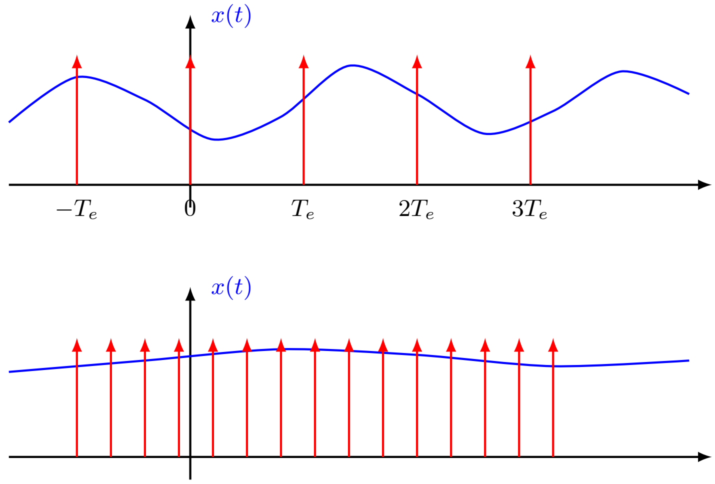


Figure 5.7: Top: too small a sampling frequency incurs lost information. Bottom: too large a sampling frequency wastes memory.

5.5 Fourier Transform of Dirac Comb

We now return to the Fourier transform in order to answer the optimal sampling frequency question.

There are plenty of reasons to be skeptical that x_e has a Fourier transform, or that the Fourier transform is anything other than the Fourier transform of the constant 0 signal.

One reason to think the Fourier transform might exist is that a sufficient condition for its existence is that

$$\sum_{n=-\infty}^{+\infty} |x(nT_e)| < +\infty$$

which seems plausible if $x \in \mathcal{L}^1(\mathbb{R})$ which means $\int_{\mathbb{R}} |x(t)|dt < +\infty$.

However, if we accept that the sampled signal x_e has a Fourier transform, what does it look like?

$$\begin{aligned}
x_e(t) &= x(t)Ш_{T_e} \\
&= \sum_{n=-\infty}^{+\infty} x(nT_e)\delta(t - nT_e) \\
X_e(\nu) &= \mathcal{F}(x_e(t)) \\
&= \mathcal{F}(x(t)Ш_{T_e}) \\
&= \mathcal{F}(x(t)) * \mathcal{F}(Ш_{T_e}) && \text{By Plancherel's theorem} \\
&= X(\nu) * \underbrace{\mathcal{F}(Ш_{T_e})}_{\text{But what is this?}} && (5.4)
\end{aligned}$$

In Theorem 5.5.4 we develop the Fourier transform of $Ш_{T_e}$. Before that, in Example 5.5.1 we show a partial attempt to compute $\mathcal{F}(Ш_{T_e})$, which leads to a curious obstacle.

Example 5.5.1 (*Partial attempt*)

$$\begin{aligned}
\mathcal{F}(\text{Ш}_{T_e}) &= \int_{\mathbb{R}} \text{Ш}_{T_e}(t) e^{-2\pi i\nu t}\,\mathrm{d}t \\
&= \int_{\mathbb{R}} \left(\sum_{n=-\infty}^{+\infty} \delta(t-nT_e)\right) e^{-2\pi i\nu t}\,\mathrm{d}t \\
&= \sum_{n=-\infty}^{+\infty} \int_{\mathbb{R}} \delta(t-nT_e) e^{-2\pi i\nu t}\,\mathrm{d}t \\
&= \sum_{n=-\infty}^{+\infty} e^{-2\pi i\nu nT_e} && \text{by (5.2)}
\end{aligned}$$

It is not clear how to proceed from here, so we push this aside for the moment. We will revisit this intermediate result in Theorem 5.5.4.

Before attacking Theorem 5.5.4, we observe that we can compute the Fourier series coefficients of Ш_{T_e}. Since this computation is possible, we will suppose (without proof) that Ш_{T_e} has a Fourier series expansion.

Lemma 5.5.2 (*Fourier Series of* Ш_{T_e})

$$\text{Ш}_{T_e}(t) = c_n \sum_{n=-\infty}^{+\infty} \exp\left(\frac{2\pi in}{T_e}t\right) = \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \exp\left(\frac{2\pi in}{T_e}t\right)$$

$$c_n = \frac{1}{T_e}$$

We will need the fact that if $\frac{-T_e}{2} \leq t \leq \frac{T_e}{2}$, then

$$\text{Ш}_{T_e}(t) = \delta(t)\,. \tag{5.5}$$

Proof. Since Ш_{T_e} is T_e-periodic, we assert (without proof) that Ш_{T_e} is decomposable as its Fourier series. Thus there exists a sequence, $(c_n)_{n\in\mathbb{Z}}$, for which

$$\text{Ш}_{T_e}(t) = \sum_{n=-\infty}^{+\infty} c_n \exp\left(\frac{2\pi in}{T_e}t\right). \tag{5.6}$$

Moreover, we can compute c_n as:

$$\begin{aligned}
c_n &= \frac{1}{T_e} \int\limits_{-T_e/2}^{T_e/2} Ш_{T_e}(t) \exp\left(\frac{-2\pi i n}{T_e} t\right) \mathrm{d}t \\
&= \frac{1}{T_e} \int\limits_{-T_e/2}^{T_e/2} \delta(t) \exp\left(\frac{-2\pi i n}{T_e} t\right) \mathrm{d}t && \text{; by (5.5)} \\
&= \frac{1}{T_e} \underbrace{\exp\left(\frac{-2\pi i n}{T_e} \times 0\right)}_{=1} && \text{; by Theorem 4.7.2} \\
c_n &= \frac{1}{T_e} && (5.7)
\end{aligned}$$

Now, we can substitute (5.7) into (5.6).

$$Ш_{T_e}(t) = \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \exp\left(\frac{2\pi i n}{T_e} t\right)$$

Notice in the proof of Lemma 5.5.2, that c_n, the Fourier coefficient is not a function of n, rather a constant $c_n = \frac{1}{T_e}$ for all $n \in \mathbb{Z}$.

Lemma 5.5.2 may also be thought of as an alternate (equivalent) definition of the Dirac comb.

Corollary 5.5.3 (*Fourier series form of the reciprocal Dirac comb*)

$$Ш_{1/T_e}(\nu) = T_e \sum_{n=-\infty}^{+\infty} e^{-2\pi i n T_e \nu}$$

Proof. We compute $Ш_{1/T_e}$ by substituting $\frac{1}{T_e}$ for T_e and ν for t into Lemma 5.5.2.

$$\begin{aligned}
Ш_{1/T_e}(\nu) &= T_e \sum_{n=-\infty}^{+\infty} e^{2\pi i n T_e \nu} \\
&= T_e \sum_{(-n)=-\infty}^{\infty} e^{2\pi i (-n) T_e \nu} \\
&= T_e \sum_{n=-\infty}^{+\infty} e^{-2\pi i n T_e \nu} && (5.8)
\end{aligned}$$

Theorem 5.5.4 (*Fourier transform of Dirac Comb*)

The Dirac comb, $Ш_{T_e}$, with period T_e has a Fourier transform which is equal to a scaled Dirac comb with period $\frac{1}{T_e}$.

$$\mathcal{F}(Ш_{T_e}(t)) = \frac{1}{T_e} Ш_{1/T_e}(\nu)\,.$$

First proof of Theorem 5.5.4. First, for concision, we repeat the text in Example 5.5.1.

$$\begin{aligned}
\mathcal{F}(\text{Ш}_{T_e})(\nu) &= \int_{\mathbb{R}} \text{Ш}_{T_e}(t) e^{-2\pi i \nu t} \, \mathrm{d}t \\
&= \int_{\mathbb{R}} \left(\sum_{n=-\infty}^{+\infty} \delta(t - nT_e) \right) e^{-2\pi i \nu t} \, \mathrm{d}t \\
&= \sum_{n=-\infty}^{+\infty} \int_{\mathbb{R}} \delta(t - nT_e) e^{-2\pi i \nu t} \, \mathrm{d}t \\
&= \sum_{n=-\infty}^{+\infty} e^{-2\pi i \nu n T_e} && \text{by (5.2)} \\
&= \frac{1}{T_e} \left(T_e \sum_{n=-\infty}^{+\infty} e^{-2\pi i \nu n T_e} \right) \\
&= \frac{1}{T_e} \text{Ш}_{1/T_e}(\nu) && \text{by Corollary 5.5.3}
\end{aligned}$$

Second proof of Theorem 5.5.4. Recall from Def 5.4.1 that $\text{Ш}_{T_e}(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT_e)$. We compute the Fourier transform of Ш_{T_e}.

$$\begin{aligned}
\mathcal{F}(\text{Ш}_{T_e})(\nu) &= \mathcal{F}\left(\frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \exp\left(\frac{2\pi i n}{T_e} t \right) \right) && \text{; by Lemma 5.5.2} \\
&= \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \mathcal{F}\left(\exp\left(\frac{2\pi i n}{T_e} t \right) \right) \\
&= \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \mathcal{F}\left(\exp\left(2\pi i \left(\frac{n}{T_e} \right) t \right) \right) \\
&= \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \delta\left(\nu - \frac{n}{T_e} \right) && \text{; by (4.28)} \\
&= \frac{1}{T_e} \text{Ш}_{1/T_e}(\nu) && \text{; by Def 5.4.1}
\end{aligned}$$

Theorem 5.5.4 demonstrates one function which is its proportional to its own Fourier transform, *i.e.*, an *eigenfunction* of $\mathcal{F}$. See Section A.2 on page 164 for more such self-transforming functions.

5.6 Fourier Transform of Sampled Signal

In light of Theorem 5.5.4, we may continue the derivation of $\mathcal{F}(x_e(t))$ which we attempted earlier in Equation (5.4).

$$X_e(\nu) = \nu_e(\ldots + X(\nu+\nu_e) + X(\nu) + X(\nu+\nu_e) + \ldots)$$

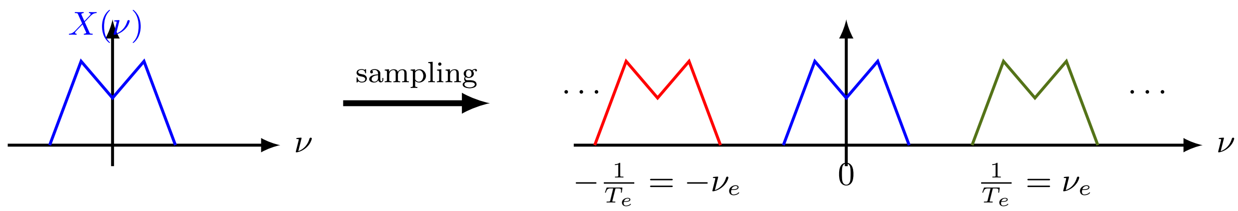


Figure 5.8: $X_e(\nu)$ expressed as periodized copies of $X(\nu)$. Sampling and periodizing are dual operations with respect to the Fourier transform.

$$\begin{aligned}
X_e(\nu) &= \mathcal{F}(x_e(t)) \\
&= \mathcal{F}(x(t)\text{Ш}_{T_e}) \\
&= \mathcal{F}(x(t)) * \mathcal{F}(\text{Ш}_{T_e}) \\
&= X(\nu) * \underbrace{\mathcal{F}(\text{Ш}_{T_e})}_{\text{apply Theorem 5.5.4}} && \text{Same as Eq(5.4)} \\
&= X(\nu) * \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} \delta\left(\nu - \frac{n}{T_e}\right) \\
&= \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} X(\nu) * \delta\left(\nu - \frac{n}{T_e}\right) && \text{By Eq (2.5)}
\end{aligned}$$

Reminder 5.6.1 (*delay via convolution with* δ)

Recall that convolving a signal with $\delta(t - t_0)$ simply delays the signal by that amount of time:

$$x(t) * \delta(t - t_0) = x(t - t_0)\,.$$

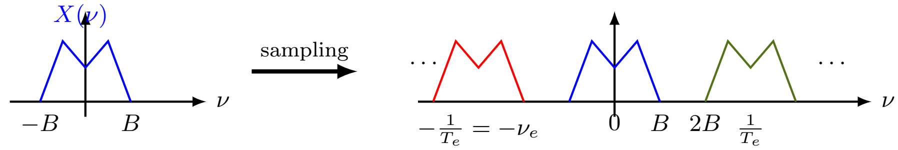


Figure 5.9: Relabeled version of Figure 5.8

So, as in Figure 5.8 we have

$$\begin{aligned}
X_e(\nu) &= \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} X\left(\nu - \frac{n}{T_e}\right) \\
&= \nu_e \sum_{n=-\infty}^{+\infty} X(\nu - n\nu_e) \\
&= \nu_e \Bigg(\cdots \quad + \quad \underbrace{X(\nu+\nu_e)}_{\text{spectrum of } x \text{ advanced by } \nu_e = \frac{1}{T_e}} \\
&\quad + \quad \underbrace{X(\nu)}_{\text{spectrum of } x} \\
&\quad + \quad \underbrace{X(\nu-\nu_e)}_{\text{spectrum of } x \text{ delayed by } \nu_e = \frac{1}{T_e}} \\
&\quad + \quad \underbrace{X(\nu-2\nu_e)}_{\text{delayed by } 2\nu_e = \frac{2}{T_e}} \quad + \cdots \Bigg)
\end{aligned}$$

The spectrum of x_e is obtained by *periodizing* the spectrum of x with a period of $\frac{1}{T_e}$ in the frequency domain.

Sampling in the time domain is equivalent to periodizing in the frequency domain.

One may also show that sampling the Fourier transform is equivalent to periodizing in the time domain. Sampling and periodizing are dual operations with respect to the Fourier transform. See Figure 5.8 (Bottom).

5.7 Shannon's Sampling Theorem

We can relabel the plot in Figure 5.8 as Figure 5.9. Two cases are possible.

1. The signal x is band limited with $X(\nu) = 0, \forall |\nu| \geq B$. In this case if $\nu_e = \frac{1}{T_e} \geq 2B$, there is no overlap between the different copies of the spectrum. However, if $\nu_e = \frac{1}{T_e} < 2B$, then there is overlap between the copies of the spectrum—spectral folding or *aliasing*.

2. The signal, x, is not band limited, in which case we have spectral folding (aliasing) regardless of the sample frequency, ν_e.

If we rename $B = \nu_{max}$, as the maximal frequency of the band limited signal, then we can state Shannon's sampling theorem.

Theorem 5.7.1 (*Shannon-Nyquist sampling theorem*)

If x is a band limited signal with maximum frequency, ν_{max}, then we may sample x at sampling frequency

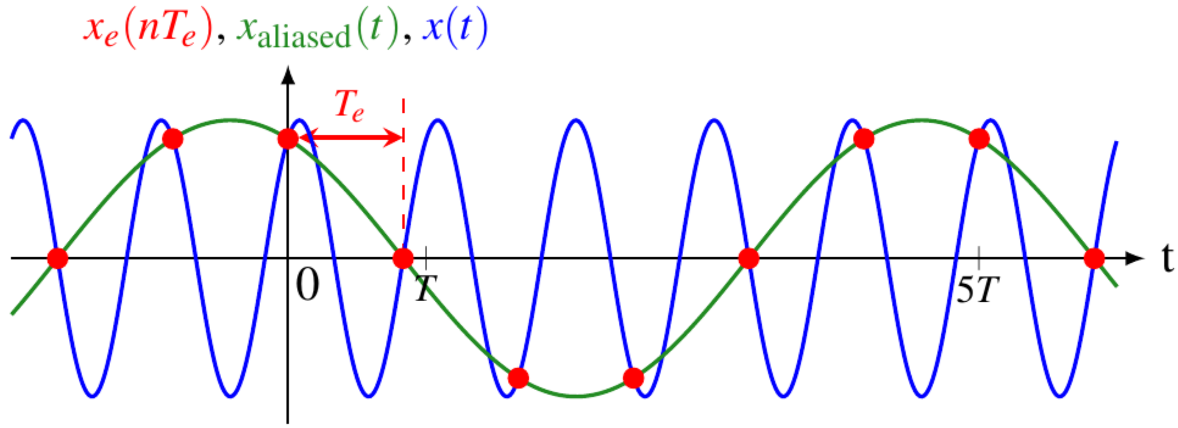


Figure 5.10: If sampling frequency is to small, it becomes impossible to reconstruct the original signal.

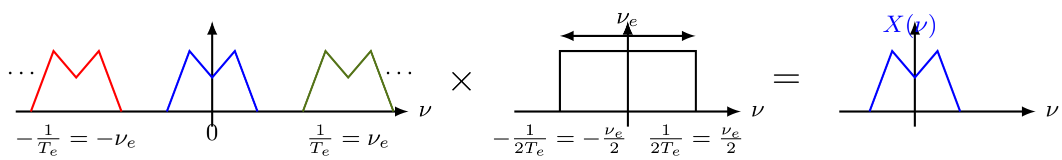


Figure 5.11: Recuperating the center copy of the duplicated spectrum.

ν_e without loss of information if

$$\underbrace{\nu_e \geq 2\nu_{max}}_{\text{called the Nyquist condition}} \quad .$$

Shannon's theorem explains the sampling rate $44.1kHz$ of the human ear which hears up to $\nu_{max} = 20kHz$. However, sounds in nature may contain frequencies greater than $20kHz$. We must filter these frequencies (with an analog filter) before the sampling step. The margin of $4.1kHz$ takes into account the attenuation of the analog low pass filter.

5.8 Reconstruction and Aliasing

If $\nu_e < 2\nu_{max}$, the Nyquist condition is not respected and there is spectral folding; consequently the signal reconstructed with different samples $x(nT_e)$ is not the original signal, x. Rather the reconstructed signal is a signal such that its sampled signal at sampling frequency ν_e is the same as that of x, as illustrated in Figure 5.10.

On the other hand, if the Nyquist condition is respected, then we can, in theory, perfectly reconstruct x given only the samples $x(nT_e), n \in \mathbb{Z}$. We start with the expression

$$X_e(\nu) = \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} X(\nu - n\nu_e) \iff T_e X_e(\nu) = \sum_{n=-\infty}^{+\infty} X(\nu - n\nu_e) \tag{5.9}$$

The idea is this: we filter $\sum_{n=-\infty}^{+\infty} X(\nu - n\nu_e)$ in a way that keeps only the central copy, $X(\nu)$ for $n = 0$,

(Figure 5.11). Thereafter, we can invert the Fourier transform to recuperate x.

$$T_e X_e(\nu) = \underbrace{\sum_{n=-\infty}^{+\infty} X(\nu - n\nu_e)}_{\text{by (5.9)}} \tag{5.10}$$

$$= [\cdots + X(\nu+\nu_e) + X(\nu) + X(\nu-\nu_e) + \cdots] \tag{5.11}$$

$$T_e X_e(\nu)\Pi_{\nu_e} = [\cdots + X(\nu+\nu_e) + X(\nu) + X(\nu-\nu_e) + \cdots]\,\Pi_{\nu_e}(\nu)$$

$$= X(\nu) \tag{5.12}$$

To obtain $X(\nu)$ given $T_e X_e(\nu)$, we filter (algebraic multiplication in the frequency domain) by a window function (template of an ideal low pass filter) of size ν_e, thus keeping all frequencies $|\nu| \leq \frac{\nu_e}{2}$ and discarding all others.

Theorem 5.8.1 (*The Shannon-Whittaker interpolation formula*)

The signal, x, is perfectly reconstructed as an infinite sum of sinc functions centered at $nT_e, n \in \mathbb{Z}$, weighted by the samples $x(nT_e)$:

$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT_e)\,\mathrm{sinc}\left(\frac{\pi}{T_e}(t - nT_e)\right)$$

Proof.

$$\begin{aligned}
X(\nu) &= (T_e X_e(\nu)) \times \Pi_{\nu_e}(\nu) && \text{by (5.12)}\\
x(t) &= \mathcal{F}^{-1}(X(\nu))\\
&= \mathcal{F}^{-1}((T_e X_e(\nu)) \times \Pi_{\nu_e}(\nu))\\
&= T_e\ \mathcal{F}^{-1}(X_e(\nu)) * \mathcal{F}^{-1}(\Pi_{\nu_e}(\nu)) && \text{by Plancherel}\\
&= T_e\ x_e(t) * \mathcal{F}^{-1}(\Pi_{\nu_e}(\nu))\\
&= T_e\ x_e(t) * \nu_e\,\mathrm{sinc}(\pi\nu_e t) && \text{by Figure 4.14}\\
&= T_e\ x_e(t) * \frac{1}{T_e}\,\mathrm{sinc}\left(\frac{\pi}{T_e}t\right)\\
&= x_e(t) * \mathrm{sinc}\left(\frac{\pi}{T_e}t\right) && \text{canceling } T_e\\
&= \left(\sum_{n=-\infty}^{+\infty} x(nT_e)\delta(t - nT_e)\right) * \mathrm{sinc}\left(\frac{\pi}{T_e}t\right) && \text{by (5.3) p. 140}\\
&= \sum_{n=-\infty}^{+\infty} \left(x(nT_e)\,\mathrm{sinc}\left(\frac{\pi}{T_e}t\right) * \delta(t - nT_e)\right)\\
&= \sum_{n=-\infty}^{+\infty} x(nT_e)\,\mathrm{sinc}\left(\frac{\pi}{T_e}(t - nT_e)\right) && \text{by Th 4.7.5 p. 106}
\end{aligned}$$

Even though x can in theory be reconstructed, the reconstruction of x at instant t requires knowledge of all the samples, $x(nT_e)$, past and future. This means the reconstruction is not causal (a consequence precedes its cause). Therefore, Theorem 5.8.1 is not usable in practice.

5.9 Quantization

In this section we look at the final step of converting the signal $x(t)$ where $x : \mathbb{R} \to \mathbb{R}$ to the sequence $(x_e(n))_{n\in\mathbb{Z}}$ where $x_e : \mathbb{Z} \to \mathbb{F}$.

5.9.1 A Sequence of Samples

Reminder 5.9.1 (*the sampled signal* $x_e(t)$)

Recall the *sampled sequence*, $x_e(t)$, from Definition 5.4.2. After the sampling step,

$$x_e(t) = \sum_{n=-\infty}^{+\infty} x(nT_e)\delta(t - nT_e)\,, \qquad \text{(5.3) from Definition 5.4.2}$$

which is still defined as a signal[1] of a continuous variable, even though it is supposed to be discrete at the entry point of the quantization step.

To solve this problem we make an abuse of notation. We will define a sequence $(x_e(n))_{n\in\mathbb{Z}}$ whose values are given in Definition 5.9.2. Why is this notation abusive? Because we also use the symbol x_e to represent the function in Equation 5.3, where $x_e : \mathbb{R} \to \mathbb{C}$.

Definition 5.9.2 (*sequence of samples*)

Let

$$\Delta n = [nT_e - \frac{T_e}{2}, nT_e + \frac{T_e}{2}]\,. \tag{5.13}$$

We define the *sequence of samples*, $x_e : \mathbb{Z} \to \mathbb{C}$, as follows Let

$$x_e(n) = \int_{\Delta n} x_e(t)\,\mathrm{d}t \tag{5.14}$$

In Theorem 5.9.4 we claim explicitly that the values of the sequence $(x_e(n))_{n\in\mathbb{Z}}$, from Definition 5.9.2, are in agreement with the values of the signal $x(t)$ when $t = nT_e$.

Before presenting Theorem 5.9.4, we present Lemma 5.9.3 which makes the proof simpler. The Lemma defines a useful sequence, $(A_{k,n})_{k,n\in\mathbb{Z}}$.

Lemma 5.9.3 (*The* $A_{k,n}$ *sequence*)

Let $T_e > 0$ and

$$A_{k,n} = \int_{nT_e - \frac{T_e}{2}}^{nT_e + \frac{T_e}{2}} x(kT_e)\delta(t - kT_e)\,\mathrm{d}t\,.$$

The three following hold.

$$k < n \implies A_{k,n} = 0 \tag{5.15}$$
$$k > n \implies A_{k,n} = 0 \tag{5.16}$$
$$k = n \implies A_{k,n} = x(nT_e) \tag{5.17}$$

For the following proofs we employ a change of variables: $u = t - kT_e$, so $du = dt$. This gives us

$$A_{k,n} = \int_{nT_e - kT_e - \frac{T_e}{2}}^{nT_e - kT_e + \frac{T_e}{2}} x(kT_e)\delta(u)du \tag{5.18}$$

of (5.15). Fix k such that $k < n$. Since $k, n \in \mathbb{Z}$, then $k \le n - 1$ so $n - k \ge 1$.

[1] The careful reader will notice that this definition of x_e actually violates Definition 1.0.2 of signal in that it is not bounded.

$$
\begin{aligned}
nT_e - kT_e + \frac{T_e}{2} &> nT_e - kT_e - \frac{T_e}{2} \\
&= T_e(n-k-1/2) && ; n-k \geq 1 \\
&\geq T_e(1-1/2) \\
&\geq \frac{T_e}{2} > 0
\end{aligned}
$$

So by Corollary 4.7.8 (page 108),

$$
A_{k,n} = \int_{nT_e - kT_e - \frac{T_e}{2}}^{nT_e - kT_e + \frac{T_e}{2}} x(kT_e)\delta(u)du = 0
$$

of (5.16). Fix k such that $k > n$. Since $k, n \in \mathbb{Z}$, then $k \geq n+1$ so $n-k \leq -1$. Let $u = t - kT_e$, so $du = dt$.

$$
\begin{aligned}
nT_e - kT_e - \frac{T_e}{2} &< nT_e - kT_e + \frac{T_e}{2} \\
&= T_e(n-k+1/2) \\
&\leq T_e(-1+1/2) && n-k \leq -1 \\
&< 0
\end{aligned}
$$

So by Corollary 4.7.8 (page 108),

$$
A_{k,n} = \int_{nT_e - kT_e - \frac{T_e}{2}}^{nT_e - kT_e + \frac{T_e}{2}} x(kT_e)\delta(u)du = 0
$$

of (5.17). Let $n = k$. First we show:

$$
\begin{aligned}
\underbrace{\int_{-\infty}^{\infty} \delta(t)}_{=1 \text{ by } (4.7)} &= \underbrace{\int_{-\infty}^{-\frac{T_e}{2}} \delta(t)}_{=0 \text{ by } (4.7.6)} + \int_{-\frac{T_e}{2}}^{\frac{T_e}{2}} \delta(t) + \underbrace{\int_{\frac{T_e}{2}}^{\infty} \delta(t)}_{=0 \text{ by } (4.7.6)} \\
1 &= 0 + \int_{-\frac{T_e}{2}}^{\frac{T_e}{2}} \delta(t) + 0 \\
1 &= \int_{-\frac{T_e}{2}}^{\frac{T_e}{2}} \delta(t) && (5.19)
\end{aligned}
$$

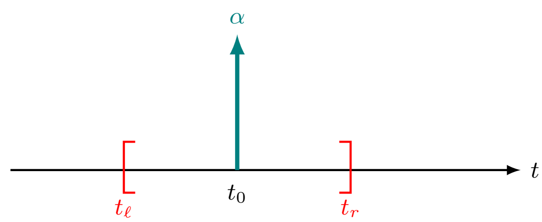


Figure 5.12: Integration of Dirac delta over bounded interval

$$
\begin{aligned}
A_{k,n} &= A_{n,n} \\
&= \int_{nT_e - nT_e - \frac{T_e}{2}}^{nT_e - nT_e + \frac{T_e}{2}} x(nT_e)\delta(u)du \\
&= \int_{-\frac{T_e}{2}}^{\frac{T_e}{2}} x(nT_e)\delta(u)du \\
&= x(nT_e) \underbrace{\int_{-\frac{T_e}{2}}^{\frac{T_e}{2}} \delta(u)du}_{=1 \text{ by } (5.19)} \\
&= x(nT_e)
\end{aligned}
$$

The conclusion of Lemma 5.9.3 is not difficult to understand, even if the notation is tedious. We already know from (4.11), that $\int_{-\infty}^{\infty} \alpha\delta(t - t_0) = \alpha$. In Lemma 5.9.3, we simply assert that the limits if integration need not be $[-\infty, \infty]$; but rather any bounded, non-zero interval suffices as long as it includes t_0 in its interior. In Figure 5.12, $\int\limits_{t_\ell}^{t_r} \alpha\delta(t - t_0) = \alpha$. In particular, in Lemma 5.9.3: $t_\ell = nT_e - \frac{T_e}{2}$, $t_r = nT_e + \frac{T_e}{2}$, $t_0 = kT_e$, and $\alpha = x(kT_e)$.

With Lemma 5.9.3 under our belt, we can state Theorem 5.9.4.

Theorem 5.9.4 (*The sample sequence agrees with* x *at* $t = nT_e$)

The sequence, $(x_e(n))_{n\in\mathbb{Z}}$, agrees with the original function $x(t)$ when $t = nT_e$.

$$x_e(n) = x(nT_e)\,.$$

Proof. For Theorem 5.9.4, we *read* the weight of each Dirac-δ:

$$
\begin{aligned}
x_e(n) &= \int_{\Delta n} x_e(t)\,\mathrm{d}t \\
&= \int_{\Delta n} \sum_{k=-\infty}^{\infty} x(kT_e)\delta(t-kT_e)\,\mathrm{d}t && \text{by (5.3)} \\
&= \int\limits_{nT_e-\frac{T_e}{2}}^{nT_e+\frac{T_e}{2}} \sum_{k=-\infty}^{\infty} x(kT_e)\delta(t-kT_e)\,\mathrm{d}t && \text{by (5.13)} \\
&= \sum_{k=-\infty}^{\infty} \int_{nT_e-\frac{T_e}{2}}^{nT_e+\frac{T_e}{2}} x(kT_e)\delta(t-kT_e)\,\mathrm{d}t && (5.20)
\end{aligned}
$$

Notice that (5.20) is exactly the form of Lemma 5.9.3.

$$
\begin{aligned}
x_e(n) &= \sum_{k=-\infty}^{\infty} A_{k,n} && \text{by Lemma 5.9.3} \\
&= \sum_{k<n} A_{k,n} + A_{n,n} + \sum_{k>n} A_{k,n} \\
&= \sum_{k=-\infty}^{n-1} 0 + x(nT_e) + \sum_{k=n+1}^{\infty} 0 && \text{Eq (5.15), (5.16), (5.17)} \\
&= x(nT_e)
\end{aligned}
$$

This operation of integration may be realized in practice by building an analog integrator (RC filter).

5.9.2 Partitioning Range into Intervals

There is a problem with the sequence $(x_e(n))_{n\in\mathbb{Z}}$ defined in Definition 5.9.2. The values of $(x_e(n))_{n\in\mathbb{Z}}$ obtained as in the definition are real values, $x_e(n) \in \mathbb{R}$, which is not compatible with digital storage. We must round the values. We call this process of rounding, quantization. In the coming sections we will discuss different strategies to quantization.

Definition 5.9.5 (*quantization*)

We define *quantization* as rounding with respect to some finite dictionary of values, $\mathbb{F}$.

In practice, the signal, x, and thus its sampled sequence, x_e, is bounded; *i.e.*, there exists an $A \in \mathbb{R}$ such that $\forall t \in \mathbb{R}, n \in \mathbb{Z}$, we have $|x(t)| \leq A$, and thus $|x_e(n)| \leq A$. Figure 5.13 shows the general idea:

- Divide the interval $[-A, A]$ into L-many subintervals $[a_k, a_{k+1}]\,, k = 0, \cdots, L-1$, such that

$$[-A, A] \in \bigcup_{i=0}^{L-1} [a_i, a_{i+1}]\ .$$

- Define a quantization level, α_k for each $[a_k, a_{k+1}]$.
- Round to α_k all the values

$$x_e(n) \in [a_k, a_{k+1}] : Q(x_e(n) \in [a_k, a_{k+1}]) = \alpha_k\ .$$

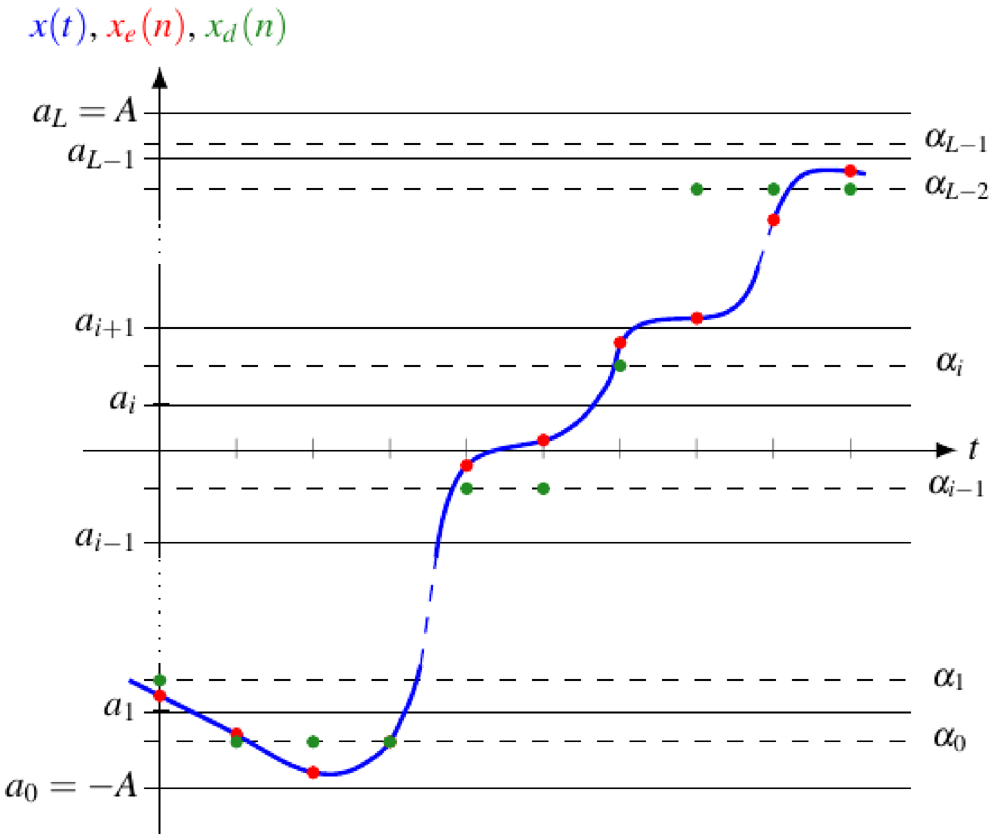


Figure 5.13: Division of the range $[-A, A]$ into intervals.

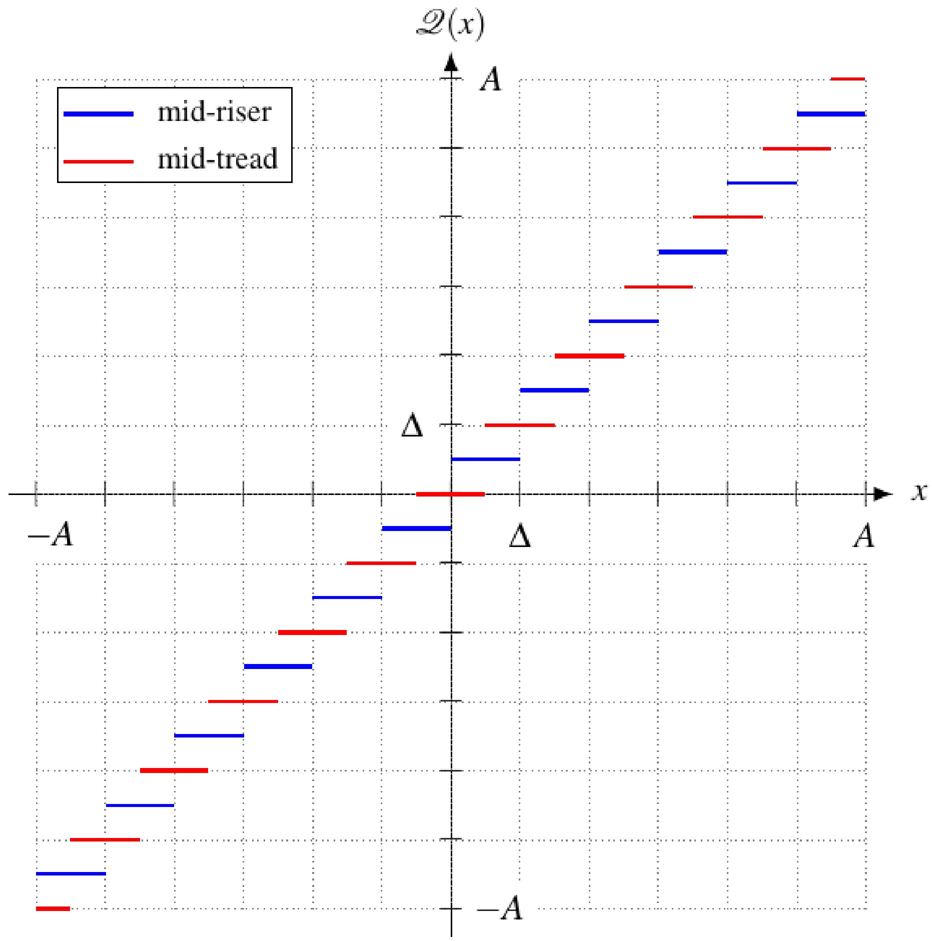


Figure 5.14: Quantization Strategies

We call $\Delta k = \alpha_{k+1} - \alpha_k$ the quantization step.

During the quantization step, all the values which fall into the same interval are rounded to the same level. Thus quantization incurs irreversible loss of information.

In practice, the values after quantization are encoded into b-bits. Let $L = 2^b$ be the quantization level and let the quantization step Δk be uniform and constant.

$$\begin{aligned}
\Delta k &= \Delta \\
&= 2\frac{A}{2^b} && \text{because the range is } [-A, A] \\
&= \frac{A}{2^{b-1}}
\end{aligned}$$

There are several possible strategies for defining the quantization levels, illustrated in Figure 5.14 and explained in Sections 5.9.3 and 5.9.4

5.9.3 Mid-riser quantization

The level, α_k is defined by the midpoint of the interval $[a_k, a_{k+1}]$,

$$\alpha_k = \frac{a_k + a_{k+1}}{2}$$

For example, the interval $[0, \Delta]$ is rounded to $\frac{\Delta}{2}$, while the next interval $[\Delta, 2\Delta]$ is rounded to $\frac{3\Delta}{2}$, etc.

Mathematically, we denote the rounding of the value $x(n)$ as

$$Q(x(n)) = \Delta \times \left(\left\lfloor \frac{x(n)}{\Delta} \right\rfloor + \frac{1}{2} \right). \tag{5.21}$$

advantage Divides the interval $[-A, A]$ into an even number of quantization levels, so it can be effectively encoded into b bits.

disadvantage $[-\Delta, 0]$ is rounded to $\frac{-\Delta}{2}$, and $[0, \Delta]$ is rounded to $\frac{\Delta}{2}$. This means that a 0-valued signal will not be quantified to 0, but rather to either $\frac{-\Delta}{2}$ or $\frac{\Delta}{2}$ — wasted energy.

5.9.4 Mid-tread quantization

Every value which falls into the interval $[\alpha_k - \frac{\Delta}{2}, \alpha_k + \frac{\Delta}{2}]$ is rounded to α_k. For example, $[\frac{-\Delta}{2}, \frac{\Delta}{2}]$ is rounded to 0, while $[\frac{\Delta}{2}, \frac{3\Delta}{2}]$ is rounded to Δ, etc.

Mathematically, we denote the rounding of the value $x(n)$ as

$$Q(x(n)) = \Delta \times \left\lfloor \frac{x(n)}{\Delta} + \frac{1}{2} \right\rfloor . \tag{5.22}$$

advantage Solves the problem of mid-riser. A 0-valued signal quantifies as a 0-valued signal.

disadvantage The interval $[-A, -A + \frac{\Delta}{2}]$ is rounded to $-A$, and the interval $[A - \frac{\Delta}{2}, A]$ is rounded to A. This is because the actual intervals are $[-A - \frac{\Delta}{2}, -A + \frac{\Delta}{2}]$ and $[A - \frac{\Delta}{2}, A + \frac{\Delta}{2}]$, half of which are unused, as the signal is bounded by A.

- The interval $[-A, A]$ is divided into an odd number of quantization levels, thus not fully taking advantage of b bits of encoding.
- We must abandon one of the intervals, either $[-A, -A + \frac{\Delta}{2}]$ or $[A - \frac{\Delta}{2}, A]$, in order to have an even number of quantization levels, and encodable into b bits.

5.10 Evaluating the Quantization Error

Because the quantization operation is a rounding operation, it incurs irreversible rounding errors.

$$x_e(n) = x_d(n) + e(n)$$

with $x_d(n) = Q(x_e(n))$ and $e(n)$ being the rounding error for the quantization of $x_e(n)$ to $x_d(n)$.

By construction, the rounding error cannot exceed $\frac{\Delta}{2}$ so $|e(n)| \leq \frac{\Delta}{2}$.

$$e(n) \in [\frac{-\Delta}{2}, \frac{\Delta}{2}] .$$

As illustrated in Figure 5.15, if the quantization step, Δ is sufficiently small, we can approximate the round-off error by a uniform random variable on $[\frac{-\Delta}{2}, \frac{\Delta}{2}]$; $e(n) \approx \mathcal{U}\big([\frac{-\Delta}{2}, \frac{\Delta}{2}]\big)$.

How can we numerically quantify the amount of information lost during the quantization step? This is the role of the signal-to-noise ratio (SNR, or RSB in French, *rapport signal sur bruit*) which in signal processing, measures the ratio of the strength of the signal (or its energy) to the strength of the noise. The higher the SNR, the more the signal dominates the noise, and the better the signal quality.

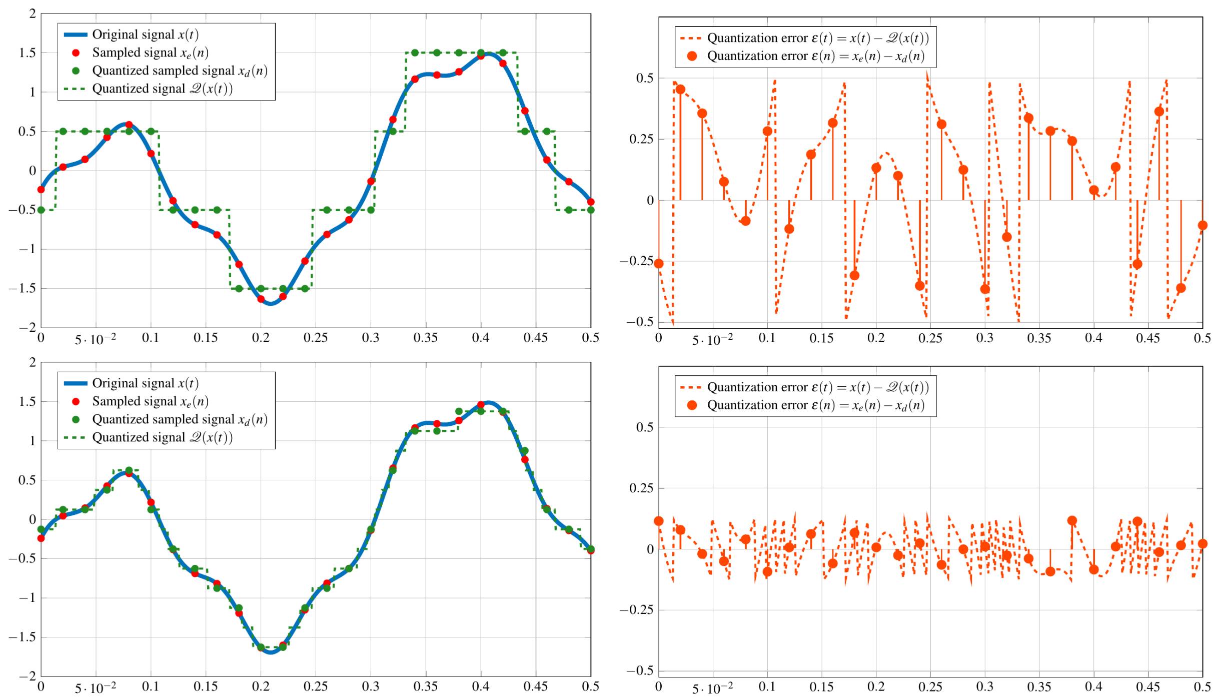


Figure 5.15: Top-left: $b = 2$ bits, 4 quantization levels. Top-right: large error w.r.t. original signal. Bottom-left: $b = 4$ bits, 16 quantization levels. Bottom-right: increasing the number of quantization levels, decreases the error.

Definition 5.10.1 (*signal to noise ratio,* SNR)

The *signal to noise ration* (SNR or sometimes S/N) of signal x to noise e is defined as

$$\mathrm{SNR}_{dB} = 10\log_{10}\frac{P_x}{P_e}$$

with P_x being the power of signal x and P_e being the power of the noise.

As are all log relations relating to power, the SNR is expressed in decibels (dB).

As the noise is neither a finite energy signal nor periodic, we have thus far defined no way to compute its power.

Here the noise of quantization is modeled by a random uniform variable $\mathcal{U}\left([\frac{-\Delta}{2}, \frac{\Delta}{2}]\right)$, for a given quantization step, Δ.

Remark 5.10.2 (*power is proportional to amplitude squared*)

We will assume that the power of a signal is proportional to the square of its amplitude.

$$P_x \propto A_x^2$$

We will not prove the equation in Remark 5.10.2, but we can give some motivation, by computing the ratio of the average power of a scaled signal to the power of the un-scaled signal.

$$\begin{aligned}
P_x &= \lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T}|x(t)|^2 dt\\
P_{Ax} &= \lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T}|Ax(t)|^2 dt\\
&= A^2\lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T}|x(t)|^2 dt\\
&= A^2 P_x\\
\frac{P_{Ax}}{P_x} &= \frac{A^2 P_x}{P_x} = A^2
\end{aligned}$$

If we assume that the sampled signal is quantized into b bits, then by Remark 5.10.2 we have $A_x = \frac{2^b}{2}\Delta$, where b is the number of bits used for quantization, assuming each quantization interval is the same size. The relation $\frac{2^b}{2}\Delta$ comes from the fact that the excursion of the signal is from $-A_x$ to A_x which is a total of $2A_x$, and $2A_x$ is divided equally into 2^b many subintervals of Δ each.

$$\begin{aligned}
2A_x &= \Delta\times 2^b\\
A_x &= 2^b\frac{1}{2}\Delta\\
&= 2^{b-1}\Delta
\end{aligned}$$

$A_e = \frac{\Delta}{2}$, because x has b bits of dynamic range and the error signal has maximally 1 bit of dynamic range. The $\frac{\Delta}{2}$ factor comes from the fact that the peak-to-peak amplitude fills the dynamic range, so the amplitudes proper (half the peak-to-peak amplitude), is half of $2^b\Delta$ and half of Δ respectively.

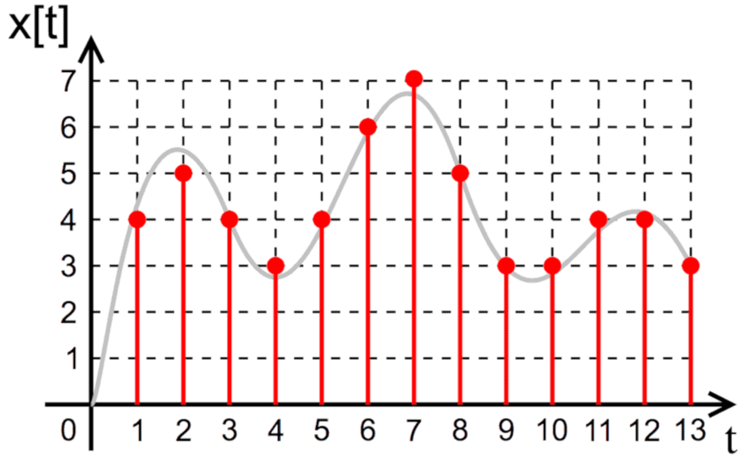


Figure 5.16: Shannon-Whittaker interpolation

Now, we can compute the SNR_{dB}.

$$\begin{aligned}
\text{SNR}_{dB} &= 10\log_{10}\frac{P_x}{P_e} && \text{by Def 5.10.1} \\
&= 10\log_{10}\left(\frac{A_x}{A_e}\right)^2 \\
&= 10\log_{10}\left(\frac{2^{b-1}\Delta}{\frac{\Delta}{2}}\right)^2 \\
&= 10\log_{10}\left(2^b\right)^2 = 10\log_{10}\left(2^{2b}\right) \\
&= 2b \times 10\log_{10} 2 = 20b\log_{10} 2 \approx 6.02\ b
\end{aligned}$$

The SNR of quantization increases $6.02dB$ each time the signal is quantified with one additional bit.

Example 5.10.3 (*SNR for audio and image encoding*)

An encoded audio signal in 16 bits ($b = 16$): $\text{SNR}_{dB} \approx 96dB$.
Classic image encoded into 8 bits ($b = 8$): $\text{SNR}_{dB} \approx 48dB$.

5.11 Review

- Analog-Digital conversion transforms a continuous signal into a discrete signal in two steps: sampling (discretizing the input variable), and quantization (discretizing the signal value).

- Sampling with a period of T_e

$$\begin{aligned}
x_e(t) &= x(t)\text{Ш}_{T_e} \\
&= \sum_{n=-\infty}^{+\infty} x(nT_e)\delta(t - nT_e)
\end{aligned}$$

- Sampling in time corresponds to periodizing in frequency

$$
\begin{aligned}
X_e(\nu) &= \frac{1}{T_e} \sum_{n=-\infty}^{+\infty} X(\nu - \frac{n}{T_e}) \\
&= \nu_e \sum_{n=-\infty}^{+\infty} X(\nu - n\nu_e)
\end{aligned}
$$

- Shannon sampling theorem: no loss of information if $\nu_e \geq 2\nu_{max}$
- If the Nyquist condition is met, (Shannon-Whittaker interpolation, Figure 5.16)

$$
x(t) = \sum_{n=-\infty}^{+\infty} x(nT_e) \, \mathrm{sinc}\left(\frac{\pi}{T_e}(t - nT_e) \right).
$$

  if the Nyquist condition is not met, spectral aliasing.

- Quantization involves rounding, and round-off error which is irreversible loss of information, which is quantified by $\mathrm{SNR}_{dB} \propto 6.02b$, for b-bit encoding.

Appendix A

Supplementary Material

Chapter Objectives

- Compute the Fresnel integral $\int_{\mathbb{R}} \cos(t^2)dt = \int_{\mathbb{R}} \sin(t^2)dt = \sqrt{\pi/2}$ using differentiation under the integral sign (the Leibniz integral rule).
- Recognize which signals are their own Fourier transform (self-transforming signals, i.e., eigenfunctions of $\mathcal{F}$), and verify this property for the Gaussian and for $1/\sqrt{|t|}$.
- Derive the Fourier transform of a Gaussian by solving the differential equation it satisfies.

A.1 The Fresnel Integral

In this section we compute

$$\int_{\mathbb{R}} \cos(t^2)dt$$

which is used in Section A.2.2 in the proof of Theorem A.2.2. A plot of this function and its sin counterpart can be seen in Figure A.1.

Definition A.1.1 (*Fresnel integrals*)

The *Fresnel integrals*, $S, C : \mathbb{R} \to \mathbb{R}$, are defined as

$$S(x) = \int_0^x \sin(t^2)\,\mathrm{d}t\,, \qquad C(x) = \int_0^x \cos(t^2)\,\mathrm{d}t\,.$$

By the Fundamental Theorem of Calculus, S and C are respectively *the* antiderivatives of $\sin(t^2)$ and $\cos(t^2)$ satisfying $S(0) = C(0) = 0$. Neither $\sin(t^2)$ nor $\cos(t^2)$ has an antiderivative expressible in elementary closed form, so S and C are defined directly as these integrals, the same way erf is defined from e^{-t^2}.

Theorem A.1.2

$$\int_{\mathbb{R}} \cos(t^2)dt = \int_{\mathbb{R}} \sin(t^2)dt = \sqrt{\frac{\pi}{2}} \tag{A.1}$$

The following proof follows very closely to the YouTube Flammable Maths video `https://www.youtube.com/watch?v=iCwiDPL3EMI` where `@papaflammy` performs the proof.

The proof uses a general form of the Leibniz integration rule.

$$\frac{d}{d\xi}\left(\int_{a(\xi)}^{b(\xi)} f(\xi,t)dt\right) = f(\xi,b(\xi))\frac{d}{d\xi}b(\xi) - f(\xi,a(\xi))\frac{d}{d\xi}a(\xi) + \int_{a(\xi)}^{b(\xi)} \frac{\partial}{\partial\xi} f(\xi,t)dt \tag{A.2}$$

If $a(\xi) = \alpha$ and $b(\xi) = \beta$ are constants with respect to ξ, then (A.2) reduces to (A.3).

$$\frac{d}{d\xi}\left(\int_{\alpha}^{\beta} f(\xi,t)dt\right) = \int_{\alpha}^{\beta} \frac{\partial}{\partial\xi} f(\xi,t)dt \tag{A.3}$$

Proof. We define J, a constant:

$$\begin{aligned} J &= \int_{\mathbb{R}} \exp\left(-it^2\right)dt \\ &= \int_{\mathbb{R}} \cos\left(t^2\right)dt - i \int_{\mathbb{R}} \sin\left(t^2\right)dt \end{aligned} \tag{A.4}$$

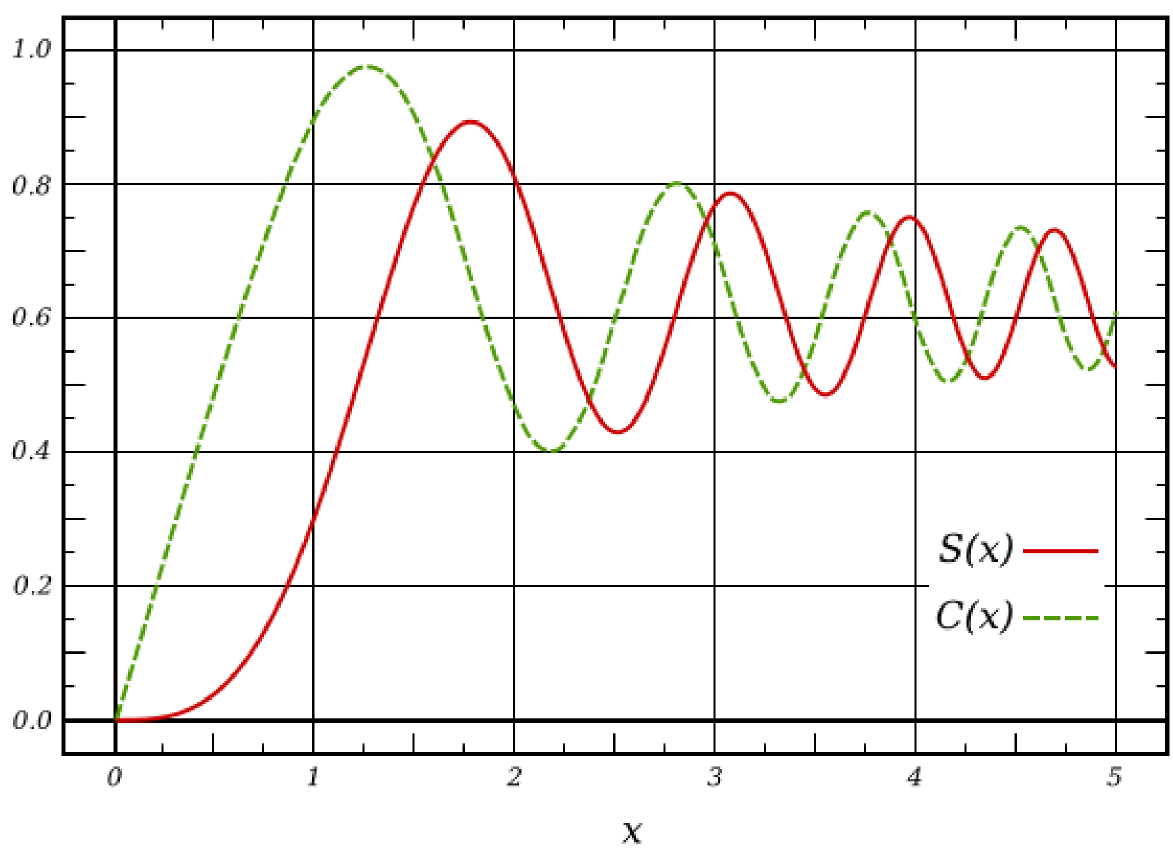


Figure A.1: The Fresnel functions $S(x)$ and $C(x)$ (Definition A.1.1).

We will prove a slightly more general result, namely that

$$\begin{aligned} J = \int_{\mathbb{R}} \exp\left(-it^2\right) dt &= \int_{\mathbb{R}} \cos\left(t^2\right) dt - i \int_{\mathbb{R}} \sin\left(t^2\right) dt \\ &= \sqrt{\frac{\pi}{2}} - i \sqrt{\frac{\pi}{2}} \end{aligned}$$

Notice that $\exp\left(-it^2\right)$ is an even function, so

$$\int_{\mathbb{R}} \exp\left(-it^2\right) dt = 2 \int_0^{\infty} \exp\left(-it^2\right) dt\,. \tag{A.5}$$

Let's define a new function, $I : \mathbb{R} \to \mathbb{C}$.

$$I(\xi) = \left(\int_0^{\xi} e^{-it^2} dt \right)^2 \tag{A.6}$$

$$J = 2 \int_0^{\infty} \exp\left(-it^2\right) dt \tag{A.7}$$

$$= \lim_{\xi \to \infty} 2\sqrt{I(\xi)} \tag{A.8}$$

Now, we consider the derivative of $I(\xi)$:

$$\begin{aligned} \frac{d}{d\xi} I(\xi) &= \frac{d}{d\xi} \left(\int_0^{\xi} \exp\left(-it^2\right) dt \right)^2 \\ &= 2 \left(\int_0^{\xi} \exp\left(-it^2\right) dt \right) \left(\frac{d}{d\xi} \int_0^{\xi} \exp\left(-it^2\right) dt \right) \\ &= 2 \left(\int_0^{\xi} e^{-it^2} dt \right) \left(\int_0^{\xi} \underbrace{\frac{\partial}{\partial \xi} e^{-it^2}}_{=0} dt + e^{-i\xi^2} \underbrace{\frac{d}{d\xi}\xi}_{=1} - \underbrace{e^{-i0^2}}_{=1} \underbrace{\frac{d}{d\xi}0}_{=0} \right) &&\text{; by (A.2)} \\ &= 2 \left(\int_0^{\xi} e^{-it^2} dt \right) \left(e^{-i\xi^2} \right) \\ &= 2 \int_0^{\xi} e^{-it^2} \left(e^{-i\xi^2} \right) dt \\ &= 2 \int_0^{\xi} \exp\left(-i\left(t^2 + \xi^2\right)\right) dt \end{aligned} \tag{A.9}$$

Now we define:

$$\begin{aligned} u &= \frac{t}{\xi} \\ du &= \frac{1}{\xi} dt \\ dt &= \xi du \\ t = 0 &\implies u = 0 \\ t = \xi &\implies u = 1 \end{aligned}$$

Substitute u into (A.9).

$$\begin{aligned}\frac{d}{d\xi}I(\xi) &= 2\int_{t=0}^{t=\xi}\exp\left(-i\xi^2\left(1+\frac{t}{\xi}^2\right)\right)dt\\ &= 2\int_{u=0}^{u=1}\exp\big(-i\xi^2\big(1+u^2\big)\big)\xi\ du\\ &= \int_0^1 2\xi\ \exp\big(-i\xi^2\big(1+u^2\big)\big)\ du\,. \end{aligned} \tag{A.10}$$

Observe:

$$\begin{aligned}2\xi\exp\big(-i\xi^2\big(1+u^2\big)\big) &= \frac{\partial}{\partial\xi}\left(\frac{\exp\big(-i\xi^2\big(1+u^2\big)\big)}{-i(1+u^2)}\right)\\ &= \frac{\partial}{\partial\xi}\left(\frac{i\ \exp\big(-i\xi^2\big(1+u^2\big)\big)}{(1+u^2)}\right). \end{aligned} \tag{A.11}$$

Now, substitute (A.11) into (A.10) and apply the Leibniz integral rule in reverse:

$$\begin{aligned}\frac{d}{d\xi}I(\xi) &= \int_0^1\frac{\partial}{\partial\xi}\left(\frac{i\ \exp\big(-i\xi^2\big(1+u^2\big)\big)}{(1+u^2)}\right)\ du\\ &= \frac{d}{d\xi}\int_0^1\left(\frac{i\ \exp\big(-i\xi^2\big(1+u^2\big)\big)}{(1+u^2)}\right)\ du \qquad ;\text{ by (A.3)}\end{aligned}$$

If two functions have the same derivative, then they differ by a constant, so we have:

$$I(\xi) = \int_0^1\left(\frac{i\ \exp\big(-i\xi^2\big(1+u^2\big)\big)}{(1+u^2)}\right)\ du + C\,. \tag{A.12}$$

We must now determine the value of C, so we substitute (A.6) into (A.12), and evaluate at $\xi = 0$.

$$\begin{aligned}\left(\int_0^\xi e^{-it^2}dt\right)^2\Bigg|_{\xi=0} &= \int_0^1\left(\frac{i\ \exp\big(-i\xi^2\big(1+u^2\big)\big)}{(1+u^2)}\right)\ du\Bigg|_{\xi=0} + C\\ \Bigg(\underbrace{\int_0^0 e^{-it^2}dt}_{=0}\Bigg)^2 &= \int_0^1\left(\frac{i\ \overbrace{\exp\big(-i\times 0^2\big(1+u^2\big)\big)}^{=1}}{(1+u^2)}\right)\ du + C\\ 0 &= \int_1^1\frac{i}{1+u^2}du + C\end{aligned}$$

Solving for C gives:

$$\begin{aligned}C &= -i\int_0^1\frac{du}{1+u^2}\\ &= -i\ \tan^{-1}(u)\Big|_{u=0}^{u=1}\\ &= -i\Big(\frac{\pi}{4}-0\Big)\\ &= -i\frac{\pi}{4}\end{aligned} \tag{A.13}$$

Now we can evaluate J by substituting (A.12) and (A.13) into (A.8).

$$\begin{aligned}
J &= \lim_{\xi\to\infty} 2\sqrt{I(\xi)} && ;\text{ by (A.8)}\\
&= \lim_{\xi\to\infty} 2\sqrt{\int_0^1 \left(\frac{i\ \exp(-i\ \xi^2(1+u^2))}{(1+u^2)}\right) du + C} && ;\text{ by (A.12)}\\
&= \lim_{\xi\to\infty} 2\sqrt{\int_0^1 \left(\frac{i\ \exp(-i\ \xi^2(1+u^2))}{(1+u^2)}\right) du - i\frac{\pi}{4}} && ;\text{ by (A.13)}
\end{aligned}$$

Moving the limit inside the integral:

$$\begin{aligned}
J &= 2\sqrt{\int_0^1 \left(\lim_{\xi\to\infty} \frac{i\ \overbrace{e^{-i\ \xi^2(1+u^2)}}^{\to\ 0 \text{ as } \xi\to\infty}}{(1+u^2)}\right) du - i\frac{\pi}{4}}\\
&= 2\sqrt{\int_0^1 0\ du - i\frac{\pi}{4}}\\
&= 2\sqrt{-i\frac{\pi}{4}}\\
&= \sqrt{\pi}\sqrt{-i}\\
&= \sqrt{\pi}\left(\frac{1}{\sqrt{2}} - i\ \frac{1}{\sqrt{2}}\right)\\
&= \sqrt{\frac{\pi}{2}} - i\ \sqrt{\frac{\pi}{2}} && && \text{(A.14)}\\
&= \int_{\mathbb{R}} \cos(t^2)dt - i\ \int_{\mathbb{R}} \sin(t^2)dt && ;\text{by (A.4)} && \text{(A.15)}
\end{aligned}$$

By (A.14) and (A.15) we conclude that

$$\begin{aligned}
\int_{\mathbb{R}} \cos(t^2)dt &= \sqrt{\frac{\pi}{2}}\\
\int_{\mathbb{R}} \sin(t^2)dt &= \sqrt{\frac{\pi}{2}}
\end{aligned}$$

A.2 Self-transforming signals

In this section we talk about signals which are their own Fourier Transform.

As seen in Theorem 5.5.4, $\mathcal{F}(\text{Ш}_T(t)) = \frac{1}{T}\text{Ш}_{1/T}(t)$. So setting $T = 1$, we have

$$\mathcal{F}(\text{Ш}_1(t)) = \text{Ш}_1(t)\,. \tag{A.16}$$

You might also ask which other signals are their own Fourier transform. Clearly if $\mathcal{F}(x_1(t) = x_1(t))$ and $\mathcal{F}(x_2(t) = x_2(t))$, then by Theorem 4.8.1, (linearity),

$$\mathcal{F}(\alpha x_1(t) + \beta x_2(t)) = \alpha x_1(t) + \beta x_2(t)\,.$$

In particular if $x_1 = x_2$ and $\alpha = \beta = 0$, we have

$$\mathcal{F}(0) = 0\,, \tag{A.17}$$

Which can also independently be verified by the definition:

$$\begin{aligned}
\mathcal{F}(0) &= \int_{\mathbb{R}} 0 \exp(-2\pi i\nu t)dt \\
&= 0 \int_{\mathbb{R}} \exp(-2\pi i\nu t)dt \\
&= 0
\end{aligned}$$

In Sections A.2.1 and A.2.2 we present two other such functions.

A.2.1 Fourier Transform of Gaussian

A well known property of the Fourier transform is that the Fourier transform of a Gaussian is itself a Gaussian. By Gaussian, we mean a bell-shaped curve of the form: $x(t) = e^{-at^2}$ in general and not necessarily the normal distribution curve: $x(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t}{2\sigma^2}\right)$.

Theorem A.2.1

$$\mathcal{F}\left(\exp\left(-\pi t^2\right)\right) = \exp\left(-\pi\nu^2\right).$$

We will actually prove a more general result:

$$\mathcal{F}\left(\exp\left(-at^2\right)\right) = \sqrt{\frac{\pi}{a}} \exp\left(-\frac{\pi^2\nu^2}{a}\right) \tag{A.18}$$

And as a simple corollary, if $a = \pi$, then we have

$$\begin{aligned}
\mathcal{F}\left(\exp\left(-\pi t^2\right)\right) &= \sqrt{\frac{\pi}{\pi}} \exp\left(-\frac{\pi^2\nu^2}{\pi}\right) \\
&= \exp\left(-\pi\nu^2\right)
\end{aligned} \tag{A.19}$$

Proof. We consider the signal

$$x(t) = \exp\left(-at^2\right). \tag{A.20}$$

Recall

$$\begin{aligned}
\int_{\mathbb{R}} x(t)dt &= \int_{\mathbb{R}} \exp\left(-at^2\right)dt \\
&= \sqrt{\frac{\pi}{a}}
\end{aligned} \tag{A.21}$$

Moreover, if $X(\nu) = \mathcal{F}(x(t)) = \mathcal{F}\big(\exp\big(-at^2\big)\big)$, then

$$\begin{aligned}
X(0) &= X(\nu)\Big|_{\nu=0} \\
&= \int_{\mathbb{R}} x(t)\exp(-2\pi i\nu t)dt\Big|_{\nu=0} \\
&= \int_{\mathbb{R}} x(t)\underbrace{\exp(-2\pi i\times 0\times t)}_{=1}dt \\
&= \int_{\mathbb{R}} x(t)dt \\
&= \sqrt{\frac{\pi}{a}} && \text{; by (A.21)} \qquad \text{(A.22)}
\end{aligned}$$

Also notice that x satisfies the following differential equation, which we can solve using results shown in Section 4.8.5: Theorems 4.8.11 and 4.8.12.

$$\begin{aligned}
\frac{d}{dt}x(t) &= \frac{d}{dt}\big(\exp\big(-at^2\big)\big) \\
&= -2at\underbrace{\exp\big(-at^2\big)}_{=x(t)} \\
&= -2at\ x(t)
\end{aligned}$$

We make take the Fourier transform of both sides of this differential equation.

$$\begin{aligned}
\mathcal{F}\left(\frac{d}{dt}x(t)\right) &= \mathcal{F}(-2at\ x(t)) \\
&= -2a\mathcal{F}(t\ x(t)) \\
2\pi i\ \nu X(\nu) &= -2a\frac{i}{2\pi}\frac{d}{d\nu}X(\nu) && \text{;by Th 4.8.11 and 4.8.12}
\end{aligned}$$

So we transform the differential equation into

$$\frac{d}{d\nu}X(\nu) = -\frac{2\pi^2}{a}\nu X(\nu) \tag{A.23}$$

$$\frac{\frac{d}{d\nu}X(\nu)}{X(\nu)} = -\frac{2\pi^2}{a}\nu \tag{A.24}$$

We can solve (A.24) by integrating both sides from 0 to ν.

$$\int_0^\nu \frac{\frac{d}{d\xi}X(\xi)}{X(\xi)}d\xi = \int_0^\nu -\frac{2\pi^2}{a}\xi d\xi \tag{A.25}$$

First, the right-hand side of (A.25)

$$\begin{aligned}
\int_0^\nu -\frac{2\pi^2}{a}\xi\ d\xi &= -\frac{2\pi^2}{a}\left[\frac{\xi^2}{2}\right]_{\xi=0}^{\xi=\nu} \\
&= -\frac{\pi^2\nu^2}{a} \qquad \text{(A.26)}
\end{aligned}$$

Now, the left-hand side of (A.25)

$$\int_0^\nu \frac{\frac{d}{d\xi}X(\xi)}{X(\xi)}d\xi = [\ln(X(\xi))]_{\xi=0}^{\xi=\nu} \tag{A.27}$$

$$= \ln(X(\nu)) - \ln(X(0))$$

$$= \ln(X(\nu)) - \ln\left(\sqrt{\frac{\pi}{a}}\right) \qquad ; \text{by (A.22)} \tag{A.28}$$

Finally, we can combine the left-hand side and left-hand side of equality (A.25), i.e., equality between (A.26) and (A.28).

$$-\frac{\pi^2\nu^2}{a} = \ln(X(\nu)) - \ln\left(\sqrt{\frac{\pi}{a}}\right)$$

$$\ln(X(\nu)) = \ln\left(\sqrt{\frac{\pi}{a}}\right) - \frac{\pi^2\nu^2}{a}$$

$$\exp(\ln(X(\nu))) = \exp\left(\ln\left(\sqrt{\frac{\pi}{a}}\right) - \frac{\pi^2\nu^2}{a}\right)$$

$$X(\nu) = \exp\left(\ln\left(\sqrt{\frac{\pi}{a}}\right)\right)\exp\left(-\frac{\pi^2\nu^2}{a}\right)$$

$$= \sqrt{\frac{\pi}{a}}\exp\left(-\frac{\pi^2\nu^2}{a}\right)$$

Thus we have that the Fourier transform of a Gaussian is also a Gaussian.

A.2.2 Inverse Square Root

Theorem A.2.2

$$\mathcal{F}\left(\frac{1}{\sqrt{|t|}}\right) = \frac{1}{\sqrt{|\nu|}} \tag{A.29}$$

The proof is taken from `stackexchange`, thanks to Job Bouwman, `https://math.stackexchange.com/users/274003/job-bouwman`.

`https://math.stackexchange.com/questions/728670/functions-that-are-their-own-fourier-transform`

Proof.

$$\mathcal{F}\left(\frac{1}{\sqrt{|t|}}\right) = \int_{\mathbb{R}} \frac{1}{\sqrt{|t|}}dt$$

We denote $f(t) = \frac{1}{\sqrt{|t|}}$. Since $f(t)$ is even and real, $F(\nu)$ must be even and real. So to avoid problems with the absolute value, we'll work with $x \geq 0$ and $\nu \geq 0$.

$$
\begin{aligned}
F(\nu) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{|t|}} \exp(2\pi i \nu t) dt \\
&= 2\int_{0}^{\infty} \frac{1}{\sqrt{|t|}} (\cos(2\pi\nu t) + i\sin(2\pi\nu t)) dt \\
&= 2\int_{0}^{\infty} \frac{1}{\sqrt{t}} \cos 2\pi t\nu dt && ;\ f \text{ is real and even}
\end{aligned}
$$

1st change of variables: let

$$
\begin{aligned}
u &= 2\pi t\nu \\
\sqrt{t} &= \sqrt{\frac{u}{2\pi\nu}} \\
dt &= \frac{du}{2\pi\nu}
\end{aligned}
$$

This leads to the following:

$$
\begin{aligned}
F(\nu) &= 2\int_{0}^{\infty} \sqrt{\frac{2\pi\nu}{u}} \cos(u) \frac{du}{2\pi\nu} \\
&= \frac{2}{\sqrt{2\pi\nu}} \int_{0}^{\infty} \frac{1}{\sqrt{u}} \cos(u) du
\end{aligned}
$$

2nd change of variables: let

$$
\begin{aligned}
u &= t^2 \\
\sqrt{u} &= t \\
du &= 2t dt
\end{aligned}
$$

This leads to the following, which makes use of a Fresnel integral. See Section A.1 for details.

$$
\begin{aligned}
F(\nu) &= \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\nu}} \int_{0}^{\infty} \frac{1}{t} \cos\left(t^2\right)\ 2t\ dt \\
&= 2\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\nu}} \qquad \underbrace{\int_{0}^{\infty} \cos\left(t^2\right)\ dt}_{=\sqrt{\frac{\pi}{8}},\ \text{the Fresnel integral}} \\
&= \sqrt{\frac{8}{\pi}} \frac{1}{\sqrt{\nu}} \sqrt{\frac{\pi}{8}} \\
&= \frac{1}{\sqrt{\nu}}
\end{aligned}
$$