Applied Analysis

Monika Dörfler

Winter term 2025/26, We 13:00 – 14:30 pm, Fr 09:45 - 11.15 pm,

January 25, 2026

Contents

List of Important Theorems
The Proofs of these theorems must be prepared for the exam

LIST OF IMPORTANT THEOREMSTHE PROOFS OF THESE THEOREMS MUST BE PREPARED

TIMELINE:

1. 8.10. Introduction: (Small) waves and signal representations
2. 10.10. Fourier Series 1
3. 15.10. Fourier Series 2
4. 17.10. Fourier Transform 1
5. 22.10. Fourier Transform 2
6. 24.10. Fourier Transform 3 / Tempered Distributions
7. 29.10. Sampling Theorem 1
8. 31.10. Sampling Theorem 2
9. 5.11. Time-frequency and Uncertainty
10. 7.11. Short-time Fourier Transform
11. 12.11. Frames and Decomposition
12. 14.11. Gabor Frames and Audio Analysis
13. 19.11. Wavelets and Image Analysis
14. 21.11. Loose Ends / Wrap Up

Chapter 1

Introduction

The fundamental concept in (Applied) Harmonic Analysis is the decomposition of data/ signals/ functions using representation systems with prescribed properties for a given class of mathematical objects (data/signals/functions). This leads to useful ways to decompose and modify or transform various such classes. Given (a subset of) a Hilbert space $\mathcal{H}$, the basic idea consists in the construction of a representation system $\Phi = \{\varphi_j\}_{j\in J}$ – so that there exists some mapping $\mathbf{C} : \mathcal{H} \to \ell^\infty$, such that

$$f = \sum_{j=1}^{N} \mathbf{C}(f)(j)\varphi_j.$$

1.1 Motivation: Sines and Cosines

A large part of this lecture will be concerned with Fourier Analysis in its different versions, that is, with Fourier series, and discrete and continuous Fourier transforms. At the base of all this are the trigonometric functions, which may be seen as the simplest periodic functions. We will now see, why.
"Sinusoids describe many natural, periodic processes"

1.1.1 Harmonic Oscillator and the Spring Equation

We will begin our study of wave phenomena by reviewing the harmonic oscillator. Consider a block with mass, m, free to slide on a frictionless air-track, but attached to a Hooke's law spring with its other end attached to a fixed wall.

Light here means that the mass of the spring is small enough to be ignored in the analysis of the motion of the block. This system has only one relevant degree of freedom. In general, the number of degrees of freedom of a system is the number of coordinates that must be specified in order to determine the configuration completely. In this case, because the spring is light, we can assume that it is uniformly stretched from the fixed wall to the block. Then the only important coordinate is the position of the block. In this situation, gravity plays no role in the motion of the block.

Recall Newton's second law: The acceleration a of a body is parallel and directly proportional to the net force F and inversely proportional to the mass m, i.e., $F = ma = m \cdot x''$.
As a second instance of a harmonic oscillator, consider a tuning fork, that is struck and thus produces a sound. What happens as the tuning fork is struck? It will be deformed and a restoring force F strives to restore the equilibrium, the fork overshoots, etc. This motion produces air pressure waves that are picked up by our ears.

How can the motion in these situations be modeled?

1. F is proportional to the displacement $x(t)$ from equilibrium:

$$F = -kx, \tag{1.1}$$

   where k is an elasticity constant (whose unit would be $\frac{N}{m}$.)

2. F produces acceleration proportional to itself:

$$F = ma, \tag{1.2}$$

   where m is the mass of the block (or the tine of the tuning fork).

3. Now recall that $a = d^2x/dt^2$, i.e. acceleration is the second derivative of displacement x with respect to time t. We thus obtain the ordinary differential equation for the harmonic oscillator:

$$\frac{d^2x}{dt^2} = -\frac{k}{m}x(t) \tag{1.3}$$

4. In order to understand the motion of the struck tine, we therefore have to find functions $x(t)$ that are proportional to their second derivative by a negative number $c = -\frac{k}{m}$:

$$\frac{d^2}{dt^2}\sin(wt) = -w^2\sin(wt) \tag{1.4}$$

   and the $\cos(wt)$ fulfills (1.3) analogously. In both cases $w = \sqrt{\frac{k}{m}}$ and the period of the oscillation is then given by $T = 2\pi\sqrt{\frac{m}{k}} = \frac{2\pi}{w} = \frac{1}{f}$, where f is the usual frequency in Hertz (whereas w is measured in radians per seconds, thus $w = 2\pi f$).

1.1.2 The Vibrating String and the Wave Equation

A vibrating string is a simple but fundamental model of wave motion. When a string (as in a guitar, violin, or piano) vibrates, the oscillations propagate along its length and produce sound. The dominant frequency of vibration determines the pitch of the note, which is usually perceived as constant over time. Vibrating strings are thus the basis of any string instrument such as the guitar, cello, or piano.

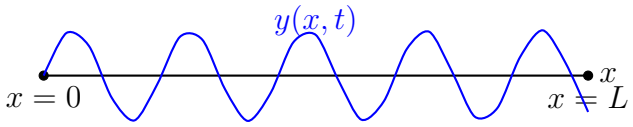


Figure 1.1: A vibrating string fixed at both ends.

Let $y(x,t)$ denote the vertical displacement of the string at position $x \in [0, L]$ and time t. Assuming the string is perfectly flexible, under uniform tension, and fixed at both ends, its dynamics are governed by the *wave equation*:

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \tag{1.5}$$

where $c = \sqrt{T/\mu}$ is the wave speed, with T the string tension and μ its linear mass density. This equation balances acceleration in time with curvature in space.

General Structure of Solutions

The wave equation is linear, and its most general solution can be written as

$$y(x,t) = f(x - ct) + g(x + ct),$$

where f and g are arbitrary twice-differentiable functions. Here: - $f(x - ct)$ represents a wave travelling to the right with speed c, - $g(x + ct)$ represents a wave travelling to the left with speed c.

Thus, the equation admits *any wave shape* provided it propagates undistorted at speed c. This general result is known as *d'Alembert's solution* of the one-dimensional wave equation.

Why Sinusoids Solve the Wave Equation

Although any shape $f(x - ct)$ or $g(x + ct)$ is allowed, sinusoidal waves play a special role. Consider the (one-dimensional) wave equation again:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}.$$

A sinusoidal wave of the form $u(x,t) = \sin(kx - \omega t)$ satisfies:

$$\frac{\partial^2 u}{\partial x^2} = -k^2 \sin(kx - \omega t), \quad \frac{\partial^2 u}{\partial t^2} = -\omega^2 \sin(kx - \omega t).$$

Substitution into the wave equation gives

$$-\omega^2 \sin(kx - \omega t) = -c^2 k^2 \sin(kx - \omega t),$$

which implies the *dispersion relation* $\omega = ck$.

Sinusoids are therefore natural eigenfunctions of the second derivative and automatically satisfy the wave equation. Moreover, because the equation is linear, any superposition of sinusoidal solutions is also a valid solution.

Boundary Conditions and Periodicity

For a string fixed at both ends, we impose

$$y(0,t) = y(L,t) = 0 \qquad \text{for all } t.$$

These boundary conditions restrict the general solution to oscillations that are periodic in space and time. We can demonstrate this explicitly.

From $y(0,t) = 0$ we have $f(-ct) + g(ct) = 0 \Rightarrow g(ct) = -f(-ct)$. Applying $y(L,t) = 0$ gives

$$f(L - ct) + g(L + ct) = 0.$$

Substituting the first relation yields

$$f(L - ct) = f(-L - ct),$$

implying that f (and hence y) is periodic with period $2L$ in its argument, corresponding to a temporal period of $\frac{2L}{c}$.

Separation of Variables Approach

An alternative way to solve (1.5) is via *separation of variables*. Assume $y(x,t) = T(t)X(x)$. Substituting into (1.5) gives

$$\frac{T''(t)}{c^2 T(t)} = \frac{X''(x)}{X(x)} = -\lambda,$$

for some constant $\lambda \geq 0$.

- The spatial equation

$$X''(x) + \lambda X(x) = 0, \quad X(0) = X(L) = 0$$

  has nontrivial solutions

$$X_n(x) = \sin\Big(\frac{n\pi x}{L}\Big), \qquad \lambda_n = \Big(\frac{n\pi}{L}\Big)^2, \quad n \in \mathbb{N}.$$

- The temporal equation

$$T''(t) + c^2 \lambda_n T(t) = 0$$

  has general solutions

$$T_n(t) = A_n \cos\Big(\frac{n\pi c}{L} t\Big) + B_n \sin\Big(\frac{n\pi c}{L} t\Big).$$

Standing Waves and Harmonics

Combining both parts, the normal modes of vibration are

$$y_n(x,t) = \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(\frac{n\pi c}{L}t\right) + B_n \sin\left(\frac{n\pi c}{L}t\right)\right], \quad n = 1, 2, \ldots$$

These are *standing waves* with $n-1$ internal nodes. The corresponding frequencies are

$$f_n = \frac{nc}{2L},$$

integer multiples of the fundamental frequency $f_1 = \frac{c}{2L}$. This explains the harmonic structure of sounds produced by string instruments: the vibration is a superposition of sinusoidal modes, which Fourier analysis precisely captures.

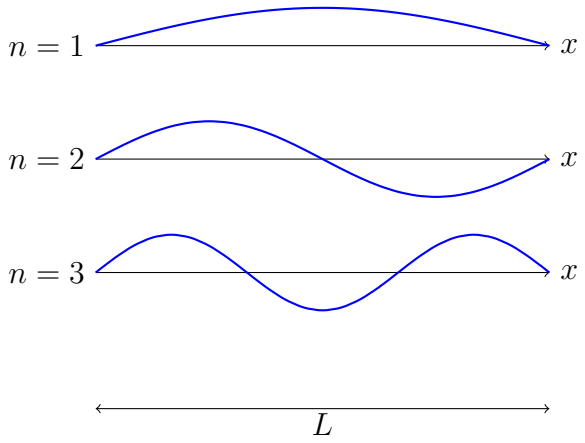


Figure 1.2: First three standing-wave modes on a string of length L. Each mode has $n-1$ internal nodes.

Periodicity and Sinusoidal Basis

A function f on $\mathbb{R}$ is *periodic* with period $p > 0$ if $f(x+p) = f(x)$ for all $x \in \mathbb{R}$. For example, $e^{2\pi i n \frac{c}{2L}t}$ is $\frac{2L}{c}$-periodic for all integers n, since

$$e^{2\pi i n \frac{c}{2L}(t+\frac{2L}{c})} = e^{2\pi i n} e^{2\pi i n \frac{c}{2L}t} = e^{2\pi i n \frac{c}{2L}t}.$$

Therefore, linear combinations of sinusoids of frequencies $n\frac{c}{2L}$ form the general periodic solution for a fixed string.

Summary. The wave equation admits a wide class of traveling-wave solutions $f(x-ct) + g(x+ct)$. Boundary conditions on a string select discrete, standing sinusoidal modes. These modes explain the harmonic overtones of musical instruments and demonstrate why *sinusoids are the natural building blocks* of solutions to the wave equation.

Chapter 2

Discrete Fourier analysis

2.1 Definition of Fourier series and examples

Definition 2.1.1 ($\mathbb{Z}^d$-periodic functions)**.** *A function $f : \mathbb{R}^d \to \mathbb{C}$ is called* $\mathbb{Z}^d$-periodic *if*

$$f(x+k) = f(x) \quad \text{for all } x \in \mathbb{R}^d \text{ and all } k \in \mathbb{Z}^d.$$

That is, f repeats its values whenever its argument is shifted by an integer vector. Equivalently, f has period 1 in each coordinate direction:

$$f(x_1, \ldots, x_j + 1, \ldots, x_d) = f(x_1, \ldots, x_j, \ldots, x_d), \qquad j = 1, \ldots, d.$$

Examples.

- *For $d = 1$, $f(x) = e^{2\pi i n x}$ is $\mathbb{Z}$-periodic for any integer n.*
- *For $d = 2$, $f(x, y) = \sin(2\pi x)\cos(2\pi y)$ is $\mathbb{Z}^2$-periodic.*

Definition 2.1.2 (d-torus)**.** *The* d-dimensional torus *is defined as the quotient space*

$$\mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d.$$

This means that points in $\mathbb{R}^d$ that differ by an integer vector are identified:

$$x \sim y \quad \text{if and only if} \quad x - y \in \mathbb{Z}^d.$$

Each coordinate of $\mathbb{T}^d$ can thus be viewed as lying in the interval $[0, 1)$ with endpoints identified. For instance:

$$\mathbb{T}^1 \cong [0, 1) \text{ with } 0 \sim 1, \qquad \mathbb{T}^2 \cong [0, 1)^2 \text{ with opposite edges identified.}$$

Remark. *Any $\mathbb{Z}^d$-periodic function $f : \mathbb{R}^d \to \mathbb{C}$ naturally defines a function on the torus:*

$$f : \mathbb{T}^d \to \mathbb{C}, \qquad f([x]) := f(x),$$

where $[x]$ denotes the equivalence class of x in $\mathbb{T}^d$. This is well-defined because $f(x + k) = f(x)$ for all $k \in \mathbb{Z}^d$.

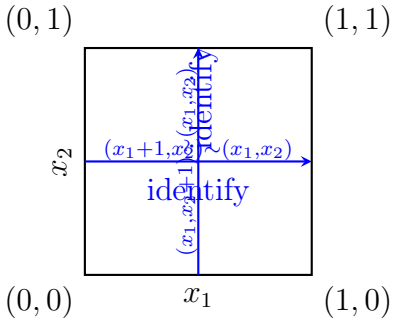


Figure 2.1: The 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Opposite edges of the unit square are identified, forming a surface topologically equivalent to a donut.

Now we dive into the question of how periodic functions may be described. We start with the definition of the real Fourier series.

Definition 2.1.3. *For a p-periodic function $f(x)$ that is integrable on $[-\frac{p}{2}, \frac{p}{2}]$, the numbers*

$$a_n = \frac{2}{p}\int_{-p/2}^{p/2} f(x)\cos(\frac{2\pi nx}{p})\,dx, \quad n \geq 0 \tag{2.1}$$

and

$$b_n = \frac{2}{p}\int_{-p/2}^{p/2} f(x)\sin(\frac{2\pi nx}{p})\,dx, \quad n \geq 1 \tag{2.2}$$

are called the Fourier coefficients of f. The expression

$$(S_N f)(x) = \frac{a_0}{2} + \sum_{n=1}^{N}[a_n\cos(\frac{2\pi nx}{p}) + b_n\sin(\frac{2\pi nx}{p})], \quad N \geq 0. \tag{2.3}$$

is called trigonometric polynomial of degree N. The infinite sum

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}[a_n\cos(\frac{2\pi nx}{p}) + b_n\sin(\frac{2\pi nx}{p})] \tag{2.4}$$

is called the Fourier series of f.

Example 2.1.4. *The Fourier series of a square wave Consider the $[0,1]$-periodic function*

$$f(x) := \begin{cases} 1 & \text{for } 0 \leq x < \frac{1}{2} \\ -1 & \text{for } \frac{1}{2} \leq x < 1 \end{cases}$$

Then its Fourier series is given by

$$f(x) = \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{1}{(2k-1)}\sin(2\pi(2k-1)x)$$

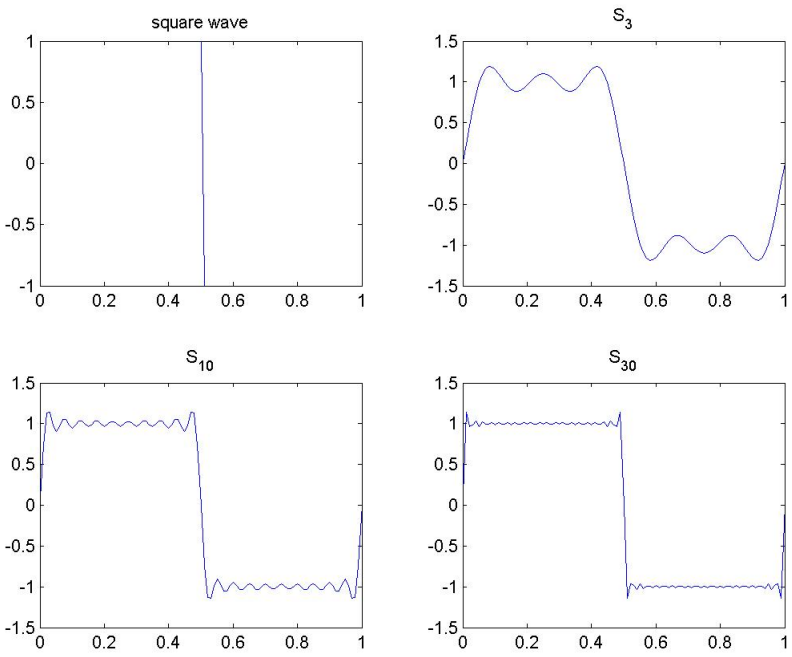


Figure 2.2: Fourier Series Square Wave

Remark 2.1.5. *For a bounded, piecewise continuous function f, the Fourier coefficients* (2.1) *and* (2.3) *yield the best approximation with a trigonometric polynomial of degree N. Furthermore, if f is piecewise smooth with finitely many discontinuities, its Fourier series converges pointwise.*

Remark 2.1.6. *Note that best approximation means, that the error which occurs, when approximating a given, p-periodic functions by a trigonometric polynomial of degree N, as in* (2.3)*, is minimal, if the coefficients a_n, b_n are the Fourier coefficients. This is an immediate consequence of the fact, that the sinusoids form an orthonormal basis for "all" periodic functions (which are sufficiently nice). We will consider this property in Proposition 2.1.8.*

The complex version of Fourier series: We can use Euler's formula, $e^{2\pi i\frac{n}{p}x} = \cos(2\pi\frac{n}{p}x) + i\sin(2\pi\frac{n}{p}x)$ where i is the imaginary unit, to give a more concise formula of the Fourier series of a function f:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i\frac{n}{p}x}, \tag{2.5}$$

with the Fourier coefficients[1] given by:

$$\hat{f}[n] = c_n = \frac{1}{p}\int_{-p/2}^{p/2} f(x)e^{-2\pi i\frac{n}{p}x}\,dx. \tag{2.6}$$

If we assume here $p = 1$, the above formulas simplify further and we can use Euler's formula, $e^{2\pi inx} = \cos(2\pi nx) + i\sin(2\pi nx)$ where i is the imaginary unit, to give a more concise formula of the Fourier series of a function f:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi inx}, \tag{2.7}$$

and

$$\hat{f}[n] = c_n = \int_{-1/2}^{1/2} f(x)e^{-2\pi inx}\,dx. \tag{2.8}$$

Example 2.1.7. The Fourier series of a sum of sinusoids
We consider the functions $f_1(t) = \sin 2\pi\omega_0 t$ and $f_2(t) = \cos 2\pi 3\omega_0 t$, for arbitrary $\omega_0 \in \mathbb{N}$ and $h(t) = f_1(t) + \frac{1}{2}\cdot f_2(t)$. We want to compute and interpret the Fourier series of these three functions. Obviously, f_1 and f_2 are pure sinusoids with frequencies ω_0 and $3\omega_0$, respectively, hence, with periods $p_1 = \frac{1}{\omega_0}$ and $p_2 = \frac{1}{3\omega_0}$. It is clear, that f_2 is also periodic with the longer period $\frac{1}{\omega_0}$. (Check this by invoking the definition of periodic functions!)
We first consider the Fourier coefficients of f_1. Since its period is $\frac{1}{\omega_0}$, we are looking for the coefficients a_n, b_n in the expansion

$$f_1(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}[a_n\cos(2\pi n\omega_0 t) + b_n\sin(2\pi n\omega_0 t)], \quad n \geq 0. \tag{2.9}$$

We can now compute the coefficients (do it!!), according to the formulas given in Definition 2.1.3, to find, using the orthogonality relations, that $a_n = 0\,\forall n$ and $b_1 = 1$, $b_n = 0$ for $n \neq 1$. Alternatively we simply look at the given form of f_1 and argue that, since the expansion in an orthogonal system is unique, the coefficients have to be of the very same form! As a third version, compute the Fourier coefficients according to (2.6).

We immediately derive (do it!!) that $c_1 = \hat{f}[1] = \frac{1}{2i}$ and $c_{-1} = \hat{f}[-1] = -\frac{1}{2i}$, which leads us directly to the expression of the sine-function via Euler's formula:

$$f_1(t) = \sin 2\pi\omega_0 t = \frac{e^{2\pi i\omega_0 t} - e^{-2\pi i\omega_0 t}}{2i}!$$

Let us first interpret these findings: obviously, the coefficients a_n, b_n in the Fourier series express the "contribution" or energy of the cosine (or sine) function to the periodic signal we wish to express. If we use the complex form, we split the energy contained in one

[1]The Fourier coefficients c_n are often denoted by $\hat{f}[n]$, since $\hat{f}$ is the most common notation for the Fourier transform of f.

sinusoid into a positive and a negative part of equal absolute value (in the case of real functions). If we use the real part, the contributions to "one frequency component" may be split in cosine and sine parts. Since this is usually more complicated, the complex form is usually preferred.
We now turn to f_2, periodic with period p_2, hence, if we consider the orthonormal basis $\{\cos(2\pi n3\omega_0 t), \quad n \geq 0\} \cup \{\sin(2\pi n3\omega_0 t), \quad n \geq 1\}$ in complete analogy to before, we compute, or derive from the properties of our orthonormal basis, that $a_n = 0\,\forall n \neq 1$ and $a_1 = 1$, $b_n = 0\forall n$. On the other hand, if we consider the basis $\{\cos(2\pi n\omega_0 t), \quad n \geq 0\} \cup \{\sin(2\pi n\omega_0 t), \quad n \geq 1\}$, which we will also have to use for $h(t)$, we find that

$$f_2(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(2\pi n\omega_0 t) + b_n \sin(2\pi n\omega_0 t)], \quad n \geq 0. \tag{2.10}$$

with $a_n = 0\,\forall n \neq 3$ and $a_3 = 1$, $b_n = 0\forall n$! We can also derive the coefficients of the complex form:
Combining all the above considerations, we now derive the Fourier coefficients of $h(t)$ according to (2.6): $c_1 = \hat{f}[1] = \frac{1}{2i}$, $c_{-1} = \hat{f}[-1] = -\frac{1}{2i}, c_3 = \hat{f}[3] = \frac{1}{2}$, $c_{-3} = \hat{f}[-3] = \frac{1}{2}$. The absolute values of these Fourier coefficients as well as the functions $h(t)$ are shown in Figure 2.3for $\omega_0 = 10$. Please also write out the real form of the Fourier series and verify that the two forms are identical. In the figure, note that the x-axis is labeled with the frequencies in Hertz. Of course, this is an interpretation of our observation that the coefficients in the Fourier series correspond to the pure frequencies in the function (signal) of interest: c_0 corresponds to $0 \cdot \omega_0 Hz$, c_1 corresponds to $1 \cdot \omega_0 Hz$, etc.

We will often denote the Fourier coefficients of a function f by $F[n]$ or $\hat{f}(n)$. More precise explanations on scalar product, (norm and minimal error), orthogonality, and the correspondence to ONBs:

Proposition 2.1.8 (ONB of exponential functions)**.** *The family of functions $\{\frac{1}{\sqrt{p}} e^{2\pi i \frac{k}{p} x}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2([-\frac{p}{2}, \frac{p}{2}])$.*

The proof consists of an easy part, namely showing the orthonormality of the functions e_k, with respect to the inner product on $L^2([-\frac{p}{2}, \frac{p}{2}])$, and of a more involved part, which aims at showing completeness. This part may be accomplished in different ways. We may, e.g., use Lemma 2.2.2, which claims that $\hat{f}(m) = \hat{g}(m)$ for all m implies $f == g$, *a.e.* for all periodic L^1 functions. Then completeness follows from general properties of an orthonormal system, since $\langle f, e_m \rangle = \hat{f}(m)$ and hence $\hat{f}(m) = \langle f, e_m \rangle = 0$ for all m implies $f == 0$, cf. Theorem 2.1.9 below.

Alternatively, we state for the real sinusoids: The sines and cosines form an orthogonal set: (note that the constant function is $\cos(2\pi \frac{m}{p} x)$ for $m = 0$).

$$\int_{-\frac{p}{2}}^{\frac{p}{2}} \cos(2\pi \frac{m}{p} x)\, \cos(2\pi \frac{n}{p} x)\, dx = \delta_{mn}, \quad m \geq 0, n \geq 1 \tag{2.11}$$

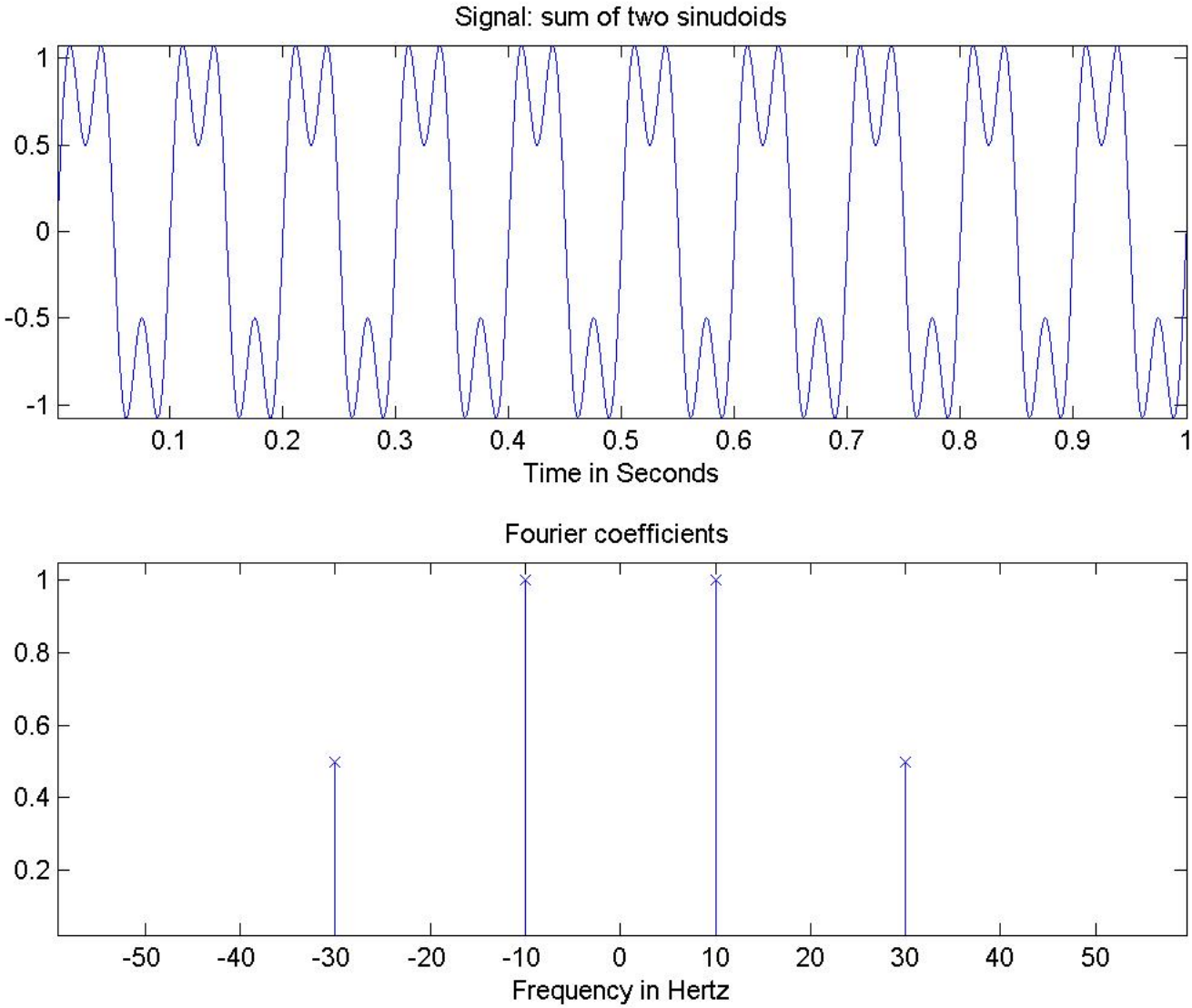


Figure 2.3: Fourier coefficients of the sum of two sinusoids

$$\int_{-\frac{p}{2}}^{\frac{p}{2}} \sin(2\pi\frac{m}{p}x)\,\sin(2\pi\frac{n}{p}x)\,dx = \delta_{mn}, \quad m,n \geq 1 \tag{2.12}$$

(here δ_{mn} is the Kronecker delta), and

$$\int_{-\frac{p}{2}}^{\frac{p}{2}} \cos(2\pi\frac{m}{p}x)\,\sin(2\pi\frac{n}{p}x)\,dx = 0, \quad m \geq 0, n \geq 1 \tag{2.13}$$

The span of these sines is dense in $L^2([-\frac{p}{2},\frac{p}{2}])$, hence they form an orthonormal basis of this vector space.

Theorem 2.1.9 (Properties ONB)**.** *Suppose that $\{\varphi_n\}_{n=1}^{\infty}$ is an orthonormal sequence in a Hilbert space H. Let*

$$V_N = \operatorname{span}\{\varphi_1,\varphi_2,\ldots,\varphi_N\}, \quad V = \bigcup_{N=1}^{\infty} V_N.$$

Then the following are equivalent:

(a) V is dense in H (with respect to the distance $d(f,g) = \|f-g\|$),

(b) If $f \in H$ and $\langle f,\varphi_n\rangle = 0$ for all n, then $f = 0$,

(c) If $f \in H$ and $s_N = \sum_{n=1}^{N}\langle f,\varphi_n\rangle\varphi_n$, then $\|s_N - f\| \to 0$ as $N \to \infty$,

(d) If $f \in H$, then

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle f,\varphi_n\rangle|^2.$$

If these properties hold, $\{\varphi_n\}_{n=1}^{\infty}$ is called an orthonormal basis or a complete orthonormal system for H.

Proof. **(a) $\implies$ (b):** Suppose $f \in H$ satisfies $\langle f,\varphi_n\rangle = 0$ for all n. Then, for any $v \in V$ (finite linear combinations of $\{\varphi_n\}$), it follows that $\langle f, v\rangle = 0$.

Since V is dense in H, there exists a sequence $\{v_j\} \subset V$ such that $\|v_j - f\| \to 0$ as $j \to \infty$. By continuity of the inner product, we have

$$\langle f, v_j\rangle \to \langle f, f\rangle.$$

However, since $\langle f, v_j\rangle = 0$ for all j, it follows that $\langle f, f\rangle = 0$, which implies $\|f\|^2 = 0$. Thus, $f = 0$.

(b) $\implies$ (c): Let $f \in H$ and denote $c_n = \langle f,\varphi_n\rangle$ and $s_N = \sum_{n=1}^{N} c_n\varphi_n$. By Bessel's inequality,

$$\sum_{n=1}^{\infty} |c_n|^2 \leq \|f\|^2 < \infty.$$

For $M < N$, the orthonormality of $\{\varphi_n\}$ gives

$$\|s_N - s_M\|^2 = \left\| \sum_{n=M+1}^{N} c_n \varphi_n \right\|^2 = \sum_{n=M+1}^{N} |c_n|^2 \to 0 \quad \text{as } M, N \to \infty.$$

Thus, $\{s_N\}$ is a Cauchy sequence in H, and by the completeness of H, there exists $u \in H$ such that $\|s_N - u\| \to 0$.

Moreover, for all n,

$$\langle f - s_N, \varphi_n \rangle = \langle f, \varphi_n \rangle - \langle s_N, \varphi_n \rangle = 0 \quad \text{for } N \geq n.$$

Taking the limit as $N \to \infty$ with n fixed, we have $\langle f - u, \varphi_n \rangle = 0$ for all n. By (b), $f - u = 0$, so $\|s_N - f\| \to 0$.

(c) $\implies$ (d): Using the Pythagorean decomposition of the norm, we have

$$\|f\|^2 = \|f - s_N\|^2 + \|s_N\|^2,$$

where $\|s_N\|^2 = \sum_{n=1}^{N} |c_n|^2$ by orthonormality. Taking the limit as $N \to \infty$, and using (c) to conclude $\|f - s_N\|^2 \to 0$, we obtain

$$\|f\|^2 = \sum_{n=1}^{\infty} |c_n|^2.$$

(d) $\implies$ (a): Using the Pythagorean decomposition again:

$$\|f\|^2 = \|f - s_N\|^2 + \sum_{n=1}^{N} |c_n|^2.$$

Taking the limit as $N \to \infty$, the rightmost term tends to $\|f\|^2$ by (d), so $\|f - s_N\|^2 \to 0$. Thus, $s_N \to f$ in norm, and since $s_N \in V_N \subset V$, it follows that V is dense in H.

□

Exercise 2.1.10. *Formulate the statements of Theorem 2.1.9 for the orthonormal basis of complex exponentials given in Proposition 2.1.8.*

Corollary 2.1.11. *Consider the subspace which is spanned by the first $2N + 1$ functions of the ONB $\{e^{2\pi ikt}\}_{k\in\mathbb{Z}}$, i.e. by $\{e^{2\pi ikt} : k = -N, \ldots, N\}$. Then, the best approximation of $f \in C([-\frac{1}{2}, \frac{1}{2}], \mathbb{C})$ by an arbitrary linear combination in $M_N(t) = \sum_{k=-N}^{N} c_k e^{2\pi ikt}$ is given by $S_N(t) = \sum_{k=-N}^{N} \hat{f}[k] e^{2\pi ikt}$.*

Proof. Let us assume that, for some $c_k \neq \hat{f}[k] = \langle f, e^{2\pi ikt} \rangle$, we can achieve a better approximation than with S_N:

$$\|f - M_N\|_2^2 < \|f - S_N\|_2^2 \tag{2.14}$$

We now show that this leads to a contradiction. We compute:

$$\begin{aligned}
\|f - M_N\|_2^2 =& \langle f - M_N, f - M_N\rangle \\
=& \langle f, f\rangle + \langle M_N, M_N\rangle - 2\mathcal{R}e\langle f, M_n\rangle \\
=& \|f\|_2^2 + \int_{-\frac{1}{2}}^{\frac{1}{2}} M_N(t)\overline{M_N}(t)dt - 2\mathcal{R}e[\langle f, \sum_{k=-N}^{N} c_k e^{2\pi ikt}\rangle] \\
=& \|f\|_2^2 + \sum_k \sum_{k'} c_k \overline{c_{k'}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi ikt} e^{-2\pi ik't} dt - 2\mathcal{R}e[\sum_{k=-N}^{N} c_k \langle f, e^{2\pi ikt}\rangle] \\
=& \|f\|_2^2 + \sum_k |c_k|^2 - 2\mathcal{R}e[\sum_{k=-N}^{N} \overline{c_k}\langle f, e^{2\pi ikt}\rangle],
\end{aligned}$$

where the last step follows from the orthogonality of the basis functions $\{e^{2\pi ikt}\}$. We carry out the same steps for S_n and obtain:

$$\begin{aligned}
\|f - S_N\|_2^2 =& \|f\|_2^2 + \sum_k |\hat{f}[k]|^2 - 2\mathcal{R}e[\sum_{k=-N}^{N} \hat{f}[k]\langle f, e^{2\pi ikt}\rangle] \\
=& \|f\|_2^2 + \sum_k |\hat{f}[k]|^2 - 2\sum_{k=-N}^{N} |\hat{f}[k]|^2 = \|f\|_2^2 - \sum_k |\hat{f}[k]|^2
\end{aligned}$$

Hence, our assumption (2.14) is equivalent to assuming

$$\sum_k |c_k|^2 - 2\mathcal{R}e[\sum_{k=-N}^{N} \overline{c_k}\langle f, e^{2\pi ikt}\rangle] < -\sum_k |\hat{f}[k]|^2$$

for some $c_k, k = -N, \ldots, N$. We rewrite this as

$$\sum_k |c_k|^2 - 2\mathcal{R}e[\sum_{k=-N}^{N} \overline{c_k}\hat{f}[k]\rangle] + \sum_k |\hat{f}[k]|^2 < 0$$

hence

$$\sum_k \left[|c_k|^2 - 2\mathcal{R}e[\overline{c_k}\hat{f}[k]] + |\hat{f}[k]|^2\right] = \sum_k |c_k - \hat{f}[k]|^2 < 0$$

and obviously the sum of positive values can never be negative. This contradiction concludes the proof. □

From general properties of ONBs we can now easily deduce the following properties of Fourier series:

Proposition 2.1.12 (Parseval Identity)**.**

$$\langle f, g\rangle_{L^2([-\frac{p}{2},\frac{p}{2}])} = p\sum_{k\in\mathbb{Z}} \hat{f}(k)\overline{\hat{g}(k)} =: p\langle \hat{f}, \hat{g}\rangle_{\ell^2} \tag{2.15}$$

In particular, setting $f = g$, it follows, that

$$\|f\|^2_{L^2([-\frac{p}{2},\frac{p}{2}])} = \langle f, f\rangle_{L^2([-\frac{p}{2},\frac{p}{2}])} = \int_{-p/2}^{p/2} |f(x)|^2 dx = p\sum_{k\in\mathbb{Z}} |\hat{f}(k)|^2.$$

Proof. A direct proof, assuming that the interchange of sum and integral is justified:

$$\begin{aligned}
\langle f, g\rangle_{L^2([-\frac{p}{2},\frac{p}{2}])} &= \int_{-p/2}^{p/2} f(x)\overline{g(x)}dx \\
&= \int_{-p/2}^{p/2} f(x)\overline{\sum_{n\in\mathbb{Z}} G[n]e^{2\pi i\frac{n}{p}x}}dx \\
&= \sum_{n\in\mathbb{Z}} \overline{G[n]} \int_{-p/2}^{p/2} f(x)e^{-2\pi i\frac{n}{p}x}dx = p\sum_{k\in\mathbb{Z}} F[k]\overline{G[k]} \\
&= p\sum_{k\in\mathbb{Z}} \hat{f}(k)\overline{\hat{g}(k)} =: p\langle \hat{f}, \hat{g}\rangle_{\ell^2}
\end{aligned}$$

□

2.1.1 Computing the truncation error

$$f(t) = S_N(t) + \sum_{|k|>N} \hat{f}(k)e^{2\pi ikt}$$

How big is $E_N(t) = f(t) - S_N(t)$? And how do we measure?

$$\|f - S_N\|_2^2 = \int_{1/2}^{1/2} |E_N(t)|^2 dt = \sum_{|k|>N} |\hat{f}(k)|^2$$

Because of the isometry of the Fourier transform, the error term can be computed by

$$\sum_{|k|>N} |\hat{f}(k)|^2 = \|f\|_2^2 - \sum_{|k|\le N} |\hat{f}(k)|^2$$

Example 2.1.13. *Recall the square wave*

$$f(x) = \begin{cases} 1, & 0 \le x < \frac{1}{2}, \\ -1, & \frac{1}{2} \le x < 1, \end{cases}$$

with Fourier series

$$f(x) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{1}{2k-1} \sin\big(2\pi(2k-1)x\big),$$

hence

$$b_n = \frac{4}{\pi n} \quad \textit{(nonzero only for odd } n\textit{)}.$$

If we keep only the first 5 *nonzero terms (i.e.* $k = 1, \dots, 5$*, corresponding to* $n = 1, 3, 5, 7, 9$*), the residual (mean-square) error over one period is the energy in the omitted coefficients. Using Parseval's identity (for period* 1*) one has*

$$\int_0^1 f(x)^2 \, dx = \frac{1}{2} \sum_{n \geq 1} b_n^2.$$

Since $f^2 \equiv 1$ *we get* $\int_0^1 f^2 \, dx = 1$ *and consequently*

$$1 = \frac{1}{2} \sum_{n \geq 1} b_n^2.$$

Therefore the mean-square error (residual energy) after truncating to the first 5 *nonzero terms is*

$$E_{\text{ms}} \;=\; \frac{1}{2} \sum_{n \in \{odd\} \setminus \{1,3,\dots,9\}} b_n^2 \;=\; \frac{1}{2} \sum_{k=6}^{\infty} \Big(\frac{4}{\pi(2k-1)} \Big)^2 \;=\; \frac{8}{\pi^2} \sum_{k=6}^{\infty} \frac{1}{(2k-1)^2}.$$

Equivalently (using the total energy $1 = \frac{1}{2} \sum_{k \geq 1} b_{2k-1}^2$*) one can write*

$$E_{\text{ms}} \;=\; 1 - \frac{1}{2} \sum_{k=1}^{5} \Big(\frac{4}{\pi(2k-1)} \Big)^2 \;=\; 1 - \frac{8}{\pi^2} \sum_{k=1}^{5} \frac{1}{(2k-1)^2}.$$

Evaluating numerically,

$$\sum_{k=6}^{\infty} \frac{1}{(2k-1)^2} \approx 0.0498355967474,$$

hence

$$E_{\text{ms}} \approx \frac{8}{\pi^2} \cdot 0.0498355967474 \approx 0.0403952132.$$

The L^2 *(RMS) error is the square root of the mean-square error:*

$$E_{L^2} = \sqrt{E_{\text{ms}}} \approx \sqrt{0.0403952132} \approx 0.2009856.$$

Thus:

$$\boxed{\textit{Mean-square error} \approx 0.0403952, \qquad \textit{RMS error} \;\approx 0.2009856.}$$

For interpretation: the first 5 nonzero Fourier terms capture about $1 - E_{\text{ms}} \approx 0.959605$ *or* $\approx 95.96\%$ *of the signal energy.*

Example 2.1.14. *In Figure 2.4, you can see a harmonic signal with added noise, i.e. which is of the form*

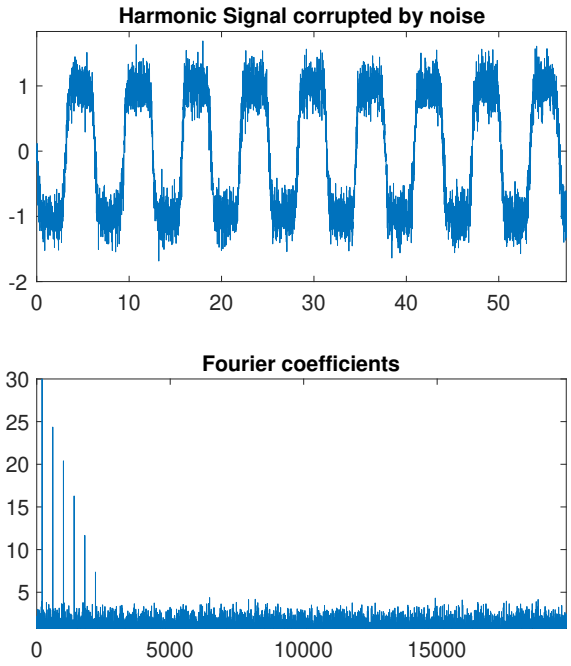


Figure 2.4: A harmonic signal corrupted by noise (upper plot) and its Fourier coefficients (lower plot).

Historical Development of Convergence of Fourier Series

- **Fourier (1822)** Jean-Baptiste Joseph Fourier introduced the concept of representing functions as infinite trigonometric series in his work on heat conduction:

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx.$$

  Fourier conjectured that such series could represent a wide class of functions. However, his work lacked rigorous justification, particularly regarding the conditions under which the series converges.

- **Dirichlet and Pointwise Convergence (1829)** Peter Gustav Lejeune Dirichlet provided the first rigorous proof of the pointwise convergence of Fourier series. He showed that if $f(x)$ is piecewise continuous and has bounded variation on $[-\pi, \pi]$, then its Fourier series converges pointwise to $f(x)$ at points of continuity and to the average of left and right limits at discontinuities.

- **Riemann and Conditions for Convergence (1854)** Bernhard Riemann extended the theory of Fourier series by studying conditions for convergence using the concept

of Riemann integration. He demonstrated that boundedness of a function does not guarantee convergence of its Fourier series.

- **Fejér and Cesàro Summability (1900)** Lipót Fejér introduced Cesàro summability, showing that the arithmetic means of the partial sums of a Fourier series always converge uniformly to a continuous function. Fejér's theorem provided a significant advancement in understanding the norm convergence of Fourier series:

$$\sigma_N(x) = \frac{1}{N+1} \sum_{k=0}^{N} S_k(x),$$

where $\sigma_N(x)$ denotes the Cesàro mean of the partial sums $S_k(x)$.

- **Kolmogorov and Divergence Examples (1923)** Andrey Kolmogorov constructed an example of an L^1 function whose Fourier series diverges almost everywhere. This result showed that the convergence of Fourier series is not guaranteed for all integrable functions and highlighted the limitations of Fourier's original conjectures.

- **Carleson and Almost Everywhere Convergence (1966)** Lennart Carleson achieved a major breakthrough by proving that the Fourier series of any square-integrable function ($f \in L^2$) converges almost everywhere to $f(x)$. His result resolved a long-standing open problem and was further refined by Richard Hunt (1968) to apply to functions in L^p for $p > 1$.

- **Summary of Key Results**
  Pointwise convergence: Established for piecewise continuous and bounded variation functions (Dirichlet).
  Norm convergence: Cesàro summability guarantees convergence for continuous functions (Fejér).
  Divergence: Constructed for certain L^1 functions (Kolmogorov).
  Almost everywhere convergence: Proven for L^2 functions (Carleson), extended to L^p for $p > 1$ (Hunt).

2.2 Pointwise convergence of Fourier series

Pointwise convergence of Fourier series is crucial in applications where local behavior of functions or signals is significant. Key practical aspects include:

1. **Signal Reconstruction:** Fourier series are widely used in audio, image, and signal processing to approximate or reconstruct signals. Pointwise convergence ensures the approximation closely matches the original signal at specific points, preserving critical details.

2. **Analysis of Discontinuities:** Pointwise convergence highlights phenomena such as the Gibbs phenomenon, providing insights into the limitations of Fourier series in approximating functions with jumps or sharp changes.

3. **Boundary Conditions in Physics:** Many physical problems (e.g., heat and wave equations) rely on Fourier series to satisfy boundary conditions. Pointwise convergence ensures the solution matches physical constraints at key points.

4. **Numerical and Computational Applications:** Practical computations often involve pointwise evaluations of Fourier series. Understanding pointwise convergence aids in error estimation and improving numerical algorithms for approximating functions.

5. **Piecewise Smooth Functions:** Many real-world functions, such as signals, are piecewise smooth rather than globally smooth. Pointwise convergence ensures the Fourier series represents these functions accurately at most points.

While other forms of convergence (e.g., uniform or L^2) are also important, pointwise convergence provides a critical perspective on how Fourier series approximate functions at specific locations, making it indispensable in both theoretical and practical contexts.

2.2.1 Version I: Fourier series in $\mathbb{R}$: Fejer

We first want to establish the following statement, which answers a question raised in reaction to du Bois-Reymond 's construction of a continuous function whose Fourier series is not convergent: is a continuous function uniquely determined by its Fourier coefficients? Fejer proved that the answer is yes and deduced it from the following statement.

Theorem 2.2.1 (Fejer's Theorem)**.** *Let $f : \mathbb{T} \to \mathbb{C}$ be continuous then*

$$\sigma_N(f,t) := \sum_{n=-N}^{N} \frac{N+1-|n|}{N+1} \hat{f}(n) e^{2\pi i n t} \longrightarrow_{N\to\infty} f(t) \tag{2.16}$$

Corollary 2.2.2. *Let $f, g : \mathbb{T} \to \mathbb{C}$ be continuous and $\hat{f}(m) = \hat{g}(m)\ \forall\, m \in \mathbb{Z}$, then $f = g$.*

Proof. Since all Fourier coefficients of f and g coincide, we see immediately that

$$0 = \sigma_N(f,t) - \sigma_N(g,t) \longrightarrow_{N\to\infty} f(t) - g(t).$$

□

Note that $\sigma_N(f,t) = K_N * f(t)$, where $K_N(t) = \sum_{n=-N}^{N} \frac{N+1-|n|}{N+1} e^{2\pi i n t}$ are the Fejer kernels, basically the Cesaro sum of the Dirichlet kernels, which are defined in the appendix. For the Fejer kernels, we can observe the following properties, which will enlighten the strategy of the proof of Proposition 2.2.1.

Lemma 2.2.3. *1.* $K_N(0) = N+1$ *and* $K_N(s) = \frac{1}{N+1}\left[\frac{\sin(N+1)\pi s}{\sin(\pi s)}\right]^2$

2. $K_N \geq 0$ *on* $\mathbb{T}$.

3. $K_N(s) \to 0$ *uniformly outside of* $[-\delta, \delta]$ *for all* $\delta > 0$.

4. $\int_{\mathbb{T}} K_N(s)ds = 1$ *for all* N.

Proof. (1) We begin by evaluating $K_N(s)$ explicitly:

$$K_N(s) = \frac{1}{N+1}\sum_{n=-N}^{N}(N+1-|n|)\,e^{2\pi ins}.$$

(a) $K_N(0) = N+1$**:** For $s = 0$, $e^{2\pi ins} = 1$ for all n. Thus:

$$K_N(0) = \frac{1}{N+1}\sum_{n=-N}^{N}(N+1-|n|)\,.$$

The summation is symmetric around $n = 0$. Breaking it into terms for $n \geq 0$ and $n < 0$, we compute:

$$\sum_{n=-N}^{N}(N+1-|n|) = 2\sum_{n=0}^{N}(N+1-n) - (N+1)\,.$$

The sum $\sum_{n=0}^{N}(N+1-n)$ is a simple arithmetic progression:

$$\sum_{n=0}^{N}(N+1-n) = (N+1) + N + \cdots + 1 = \frac{(N+1)(N+2)}{2}.$$

Thus:

$$K_N(0) = \frac{1}{N+1}\cdot\left[\frac{(N+1)(N+2)}{2} + \frac{(N+1)(N+2)}{2} - (N+1)\right] = N+1.$$

(b) Trigonometric Expression for $K_N(s)$**:** The expression for $K_N(s)$ can be derived using the Fejér summation (see below) formula:

$$\sum_{n=-N}^{N}(N+1-|n|)\,e^{2\pi ins} = \frac{\sin^2((N+1)\pi s)}{\sin^2(\pi s)}.$$

Thus:

$$K_N(s) = \frac{1}{N+1}\left[\frac{\sin((N+1)\pi s)}{\sin(\pi s)}\right]^2.$$

(2) obvious.
(3) For $|s| \geq \delta > 0$, the denominator $\sin(\pi s)$ is bounded away from zero:

$$|\sin(\pi s)| \geq \sin(\pi\delta) > 0.$$

In the numerator, $|\sin((N+1)\pi s)|$ oscillates but remains bounded by 1. Thus, outside $[-\delta, \delta]$, we have:

$$K_N(s) = \frac{1}{N+1}\left[\frac{\sin((N+1)\pi s)}{\sin(\pi s)}\right]^2 \leq \frac{1}{N+1}\frac{1}{\sin^2(\pi\delta)}.$$

As $N \to \infty$, the factor $\frac{1}{N+1} \to 0$, implying that $K_N(s) \to 0$ uniformly for $s \notin [-\delta, \delta]$.
(4) Going back to the original definition of K_N, the integral of $K_N(s)$ over one period is:

$$\int_{\mathbb{T}} K_N(s)\, ds = \int_{\mathbb{T}} K_N(t) \sum_{n=-N}^{N} \frac{N+1-|n|}{N+1} e^{2\pi i n t} ds,$$

which is equal to 0 for all $n \neq 0$. For $n = 0$: $\int_{\mathbb{T}} 1\, ds = 1$. □

Proof of Fejer's theorem. using the properties in the previous lemma, the proof is conducted as follows. Choose some sufficiently small $\delta > 0$ and sufficiently big N, then

$$\sigma_N(f,t) = K_N * f(t) = K_N * f(t) \approx \int_{-\delta}^{\delta} K_N(s) f(t-s) ds,$$

since by the above Lemma, (3), K_N is negligible outside $[-\delta, \delta]$. Furthermore, since f is continuous at t, it is approximately constant on $[t-\delta, t+\delta]$, hence $f(t-s) \approx f(t)$ for $s \in [-\delta, \delta]$ and we can write

$$\sigma_N(f,t) =\approx \int_{-\delta}^{\delta} K_N(s) f(t) ds \approx \int_{-1/2}^{1/2} K_N(s) f(t) ds,$$

where in the last step we argue again with negligibility of K_N outside $[-\delta, \delta]$. Hence, due to (4) in the previous Lemma, $\sigma_N(f,t) \approx f(t) \int_{-1/2}^{1/2} K_N(s) ds = f(t)$.
We now provide the technical steps:.
We aim to show that the Fejér summation operator

$$\sigma_N(f,t) := \sum_{n=-N}^{N} \frac{N+1-|n|}{N+1} \hat{f}(n) e^{2\pi i n t}$$

converges uniformly to $f(t)$ as $N \to \infty$, where $\hat{f}(n)$ are the Fourier coefficients of f, given by

$$\hat{f}(n) = \int_0^1 f(x) e^{-2\pi i n x}\, dx.$$

The Fejér summation can be expressed as a convolution:

$$\sigma_N(f,t) = \int_0^1 K_N(t-x)f(x)\,dx,$$

where $t-x$ is interpreted modulo 1.

Since $K_N(s)$ is a positive kernel that integrates to 1, it acts as an approximation to the identity.

Using the uniform continuity of f (which is implied if f is continuous on the compact interval $[0,1]$), we can write

$$|\sigma_N(f,t) - f(t)| = \left| \int_0^1 K_N(t-x)(f(x)-f(t))\,dx \right|.$$

By the triangle inequality and the boundedness of $K_N(s)$, we have

$$|\sigma_N(f,t) - f(t)| \leq \int_0^1 K_N(t-x)\,|f(x)-f(t)|\,dx.$$

For any $\epsilon > 0$, by the uniform continuity of f, there exists a $\delta > 0$ such that $|f(x) - f(t)| < \epsilon$ whenever $|x-t| < \delta$. Split the integral as:

$$\int_0^1 K_N(t-x)\,|f(x)-f(t)|\,dx = \int_{|x-t|<\delta} K_N(t-x)\,|f(x)-f(t)|\,dx + \int_{|x-t|\geq\delta} K_N(t-x)\,|f(x)-f(t)|\,dx.$$

1. For $|x-t| < \delta$, $|f(x)-f(t)| < \epsilon$, and since $\int_{|x-t|<\delta} K_N(t-x)\,dx \leq 1$, this term is bounded by ϵ. 2. For $|x-t| \geq \delta$, $K_N(t-x) \to 0$ as $N \to \infty$, and since f is bounded on $[0,1]$, the contribution of this term vanishes.

Thus, for sufficiently large N, we have

$$|\sigma_N(f,t) - f(t)| < \epsilon.$$

□

Theorem 2.2.4 (Pointwise Convergence I: continuous functions with ℓ^1-Fourier coefficients)**.** *If $f : \mathbb{T} \longrightarrow \mathbb{C}$ is continuous and $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty$, then $S_N f \xrightarrow{N\to\infty} f$ uniformly on $\mathbb{T}$.*

Proof. According to the ℓ^1 assumption on the coefficients $\hat{f}(n)$, we obtain, via the usual Cauchy criterion, that one can find an ϵ-dependent N_0, such that for all $n \leq m < \epsilon$ we have

$$|\sum_{n\leq|k|\leq m} \hat{f}(k)e^{2\pi ikt}| \leq \sum_{n\leq|k|\leq m} |\hat{f}(k)e^{2\pi ikt}| \leq \sum_{n\leq|k|\leq m} |\hat{f}(k)| \leq \epsilon,$$

such that $S_N(t)$ tends uniformly to some continuous $g(t)$ on $[0,1]$ for $N \to \infty$. We still need to show, that $f = g$. Since $\int_0^1 e^{2\pi ikt}\,dt = 1$ if and only if $k = 0$ and the integral is zero for all other integers $k \in \mathbb{Z}$, we observe:

$$\begin{aligned}\hat{f}(k) &= \sum_{n=-N}^{N} \hat{f}(n) \int_0^1 e^{2\pi i(n-k)t}\,dt = \int_0^1 \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i(n-k)t}\,dt \\ &= \int_0^1 \sum_{n=-N}^{N} S_N f(t) e^{-2\pi ikt}\,dt = \xrightarrow{N\to\infty} \int_0^1 \sum_{n=-N}^{N} g(t) e^{-2\pi ikt}\,dt = \hat{g}(k)\end{aligned}$$

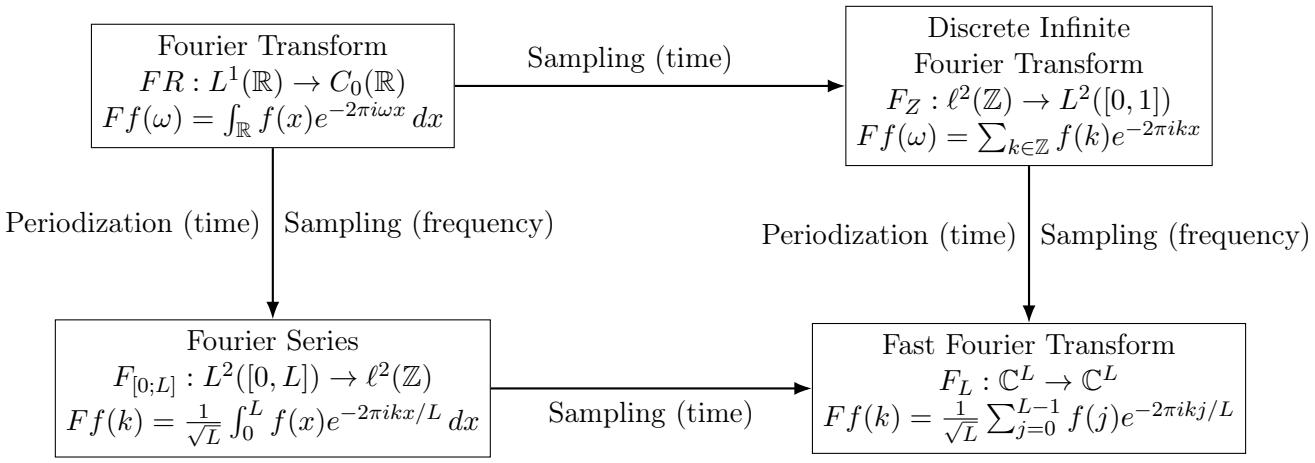
and it follows by Corollary 2.2.2 that $f = g$. □

Chapter 3

Fourier transform of functions on $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}^N$

3.1 The transition to other domains

We first introduced the Fourier series, since they are, in a certain sense, the most natural instance of Fourier transforms. The basic idea should be clear by now: a (periodic, so far) function can be represented by a sum of weighted sinusoids, and the sinusoids can be interpreted as the frequencies present in the function (signal). We will now push this concept a bit further. First, we will simply turn around the interpretation of the two variables involved in the definition of Fourier series, namely time t and frequency, which has so far been chosen to live on a discrete subset of $\mathbb{R}$ and labeled by the integers. The next step is of vital importance for understanding the world of digital signal processing, which is, in fact behind almost any modern technical tools we use. Indeed, if we assume, that the frequency information contained in a signal is contained in an interval of finite length, in other words, if the signal (function) is *band-limited*, we may - mutatis mutandis - expand its frequency-information in a "Fourier series", and the corresponding coefficients will then contain the time-information. However, as we have seen so far, this is discrete information, indexed by $k \in \mathbb{Z}$ - in Section 3.1.1 we thus arrive naturally at a concept of Fourier transform for discrete signals - as a dual concept of the Fourier series.
On the other hand, and complementary to the approach just described, we may think of a periodic time-signal for gradually growing period, which means that, for $p \to \infty$, the size of $\frac{k}{p}$ in the definition of the sinusoids $\{e^{2\pi i \frac{k}{p} x}\}_{k\in\mathbb{Z}}$ becomes infinitely small. This idea leads to Fourier transforms on $\mathbb{R}$, introduced in Section 3.1.2, with an integral replacing the sum in the representation. The entire landscape of these different Fourier transforms can be summarized as follows:



3.1.1 The discrete Fourier transform

Let $f : \mathbb{Z} \mapsto \mathbb{C}$ be a function defined on the integers. We consider the complex exponentials $e^{2\pi isn}$, $n \in \mathbb{Z}$, $s \in \mathbb{R}$ and observe immediately, that

$$e^{2\pi i(s+m)n} = e^{2\pi isn} \text{ for all } n, m \in \mathbb{Z}.$$

In other words, the exponentials $e^{2\pi isn}, e^{2\pi i(s\pm 1)n}, e^{2\pi i(s\pm 2)n}, \ldots$ cannot be distinguished if n are integer values. Hence, in order to avoid ambiguity, we will synthesize f from $e^{2\pi isn/p}$, for $0 \le s < p$ and $n \in \mathbb{Z}$.

Definition 3.1.1 (DFT)**.** *The discrete Fourier transform of a (suitably regular) function f on $\mathbb{Z}$ is defined as*

$$\hat{f}(s) = F(s) = \frac{1}{p} \sum_{n=-\infty}^{\infty} f[n]e^{-2\pi isn/p} \quad \textit{for } 0 \le s < p \tag{3.1}$$

f can then be written as

$$f[n] = \int_{s=0}^{p} F(s)e^{2\pi isn/p} ds. \tag{3.2}$$

Note that the Fourier transform $\hat{f} = F$ of a discrete-valued function is a function on the circle with diameter of length p.

Remark 3.1.2. *Note that normally p is set to 1, so that the frequencies that occur in a signal are normalized, on the unit circle. We will see later, when we discuss sampling (in fact, any signal on $\mathbb{Z}$ is a digital, hence sampled signal, unless it stems from a, inherently discrete process, e.g. a time-series of stock exchange values), that $\frac{1}{2}$ corresponds to the highest frequency that occurs in a real signal. The frequencies in the interval $]\frac{1}{2}, 1[$ are then the negative frequencies.*
We introduced p in the above definition to guarantee generality and to emphasize the parallelism with the Fourier series.

We now meet the concept of Dirac Impulse for the first time, see Appendix A.2.

Example 3.1.3 (Dirac Impulse)**.** *Consider the function δ defined on $\mathbb{Z}$, that is equal to 0 everywhere, except for $\delta[0] = 1$. It is then easy to see, that the DFT of δ is given by $\hat{\delta}(s) = 1/p$ for all $s \in [0, p]$.*

Example 3.1.4 (Sinusoid)**.** *Here, we are faced with the opposite situation: what is a pure sinusoid f's Fourier transform? $\hat{f}$ has one positive entry only, at frequency ω_0, and should be equal to 0 everywhere else. So, using (3.2):*

$$f[n] = \int_{s=0}^{1} \delta(\omega - \omega_0) e^{2\pi i \omega n} ds = e^{2\pi i \omega_0 n}$$

3.1.2 The Fourier transform of functions on $\mathbb{R}$

We now consider integrable functions on $\mathbb{R}$.
Remark: The Fourier transform provides a powerful tool for analyzing the frequency components of functions. In this section, we focus on functions that are integrable on $\mathbb{R}$, meaning they belong to the space $L^1(\mathbb{R})$, as well as square-integrable functions in $L^2(\mathbb{R})$. The distinction between these spaces is important because the Fourier transform is defined differently for each: for $L^1(\mathbb{R})$-functions, it is defined by the integral in Definition 3.1.5, while for $L^2(\mathbb{R})$-functions, the Fourier transform is often extended using the theory of distributions and the concept of density. The interplay between these spaces and the Fourier transform allows for a comprehensive analysis of functions in various settings.

Definition 3.1.5. *The Fourier transform of a function $f \in L^1(\mathbb{R})$ is defined by*

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt. \tag{3.3}$$

Note that, if f is integrable, the integral in (3.19) converges and

$$|\hat{f}(\omega)| \leq \int_{-\infty}^{\infty} |f(t)| dt < \infty.$$

Proposition 3.1.6 (Inverse Fourier transform)**.** *If $\hat{f} \in L^1(\mathbb{R})$, then f is given by the inverse Fourier transform of $\hat{f}$:*

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\omega) e^{2\pi i \omega t} d\omega \tag{3.4}$$

Remark 3.1.7. *Note that the Fourier transform is usually extended to all functions in $L^2(\mathbb{R})$ by using a density argument, similar to our approach in the proof of Proposition 2.1.8. Then, as before, an inner product can be defined on $L^2(\mathbb{R})$, and most arguments work similar to the case of periodic functions.*

Example 3.1.8 (Fourier transform of the box function)**.** *Consider the function*

$$\Pi(x) := \begin{cases} 1 & for\ -\frac{1}{2} < x < \frac{1}{2} \\ 0 & else \end{cases}$$

To compute the Fourier transform, first note that Π *is even, so that we can omit the sine-part (generally we can observe, that the Fourier transform of even (symmetric) functions is always real! On the other hand, the Fourier transform of real functions is symmetric and we only have to consider the positive frequencies. This property is heavily exploited in the processing of speech and music signals, which are always real.) We therefore have*

$$\begin{aligned} \hat{\Pi}(\omega) &= \int_{\mathbb{R}} \Pi(x) e^{-2\pi i\omega x} dx = \int_{\mathbb{R}} \Pi(x) \cos(2\pi\omega x) dx \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos(2\pi\omega x) dx = \frac{\sin(2\pi\omega x)}{2\pi\omega}\Big|_{x=-\frac{1}{2}}^{\frac{1}{2}} \\ &= \frac{\sin(\pi\omega)}{2\pi\omega} - \frac{\sin(-\pi\omega)}{2\pi\omega} = \frac{\sin(\pi\omega)}{\pi\omega} =: sinc(\omega) \end{aligned}$$

Note that $sinc(x)$ *can be defined as* $sinc = \frac{\sin(\pi x)}{\pi x}$ *only for* $x \neq 0$*. However according to L'Hôpital's rule, we have that* $\lim_{x\to 0} sinc(x) = 1$*, since, for any open interval* I *around* 0*, we have* $h'(x) = 1 \neq 0$ *for* $h(x) = x$ *and, since* $\frac{d}{dx}\sin x = cosx$ *and* $\lim_{x\to 0}\cos(x) = 1$*, we have* $\lim_{x\to 0}\frac{\sin' x}{g'(x)} = 1$*, and so* $\lim_{x\to 0}\frac{\sin x}{x} = 1$*. More directly, we may consider the Taylor series* $\sin x = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n+1}$*, such that* $sinc(x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n}$ *for* $x \neq 0$*, and convergence to* 1 *is obvious.*

We will next address two very basic operators that can act on a function or signal, namely translation, or *time-shift* and modulation, or *frequency-shift*.

Example 3.1.9 (Translation and Modulation)**.** *For any real number* x_0*, if* $g(x) = T_{x_0}f(x) := f(x - x_0)$*, then* $\hat{g}(\omega) = e^{-2\pi i x_0\omega}\hat{f}(\omega)$*.*
For any real number ω_0*, if* $g(x) = M_{\omega_0} := e^{2\pi i x\omega_0}f(x)$*, then* $\hat{g}(\omega) = \hat{f}(\omega - \omega_0)$*.*

Example 3.1.10 (Dilation)**.** *Let* $a \neq 0 \in \mathbb{R}$*. Set* $g(x) = D_a f(x) := \sqrt{a}f(ax)$*. Let* $\hat{f}$ *be the Fourier transform of* f*. Then, the Fourier transform of* g *is given by*

$$\hat{g}(\omega) = \frac{1}{a}\hat{f}(\frac{\omega}{a}).$$

Example 3.1.11. *Consider the Fourier transform of a damped oscillator:*

$$f(t) = e^{-\gamma t}\cos(\omega_0 t)\theta(t), \tag{3.5}$$

where the unit-step function $\theta(t)$ *is defined as:*

$$\theta(t) = \begin{cases} 1, & t > 0, \\ 0, & t \leq 0. \end{cases} \tag{3.6}$$

The step function ensures the oscillator starts at $t = 0$. Without it, the amplitude would blow up as $t \to -\infty$.

Rewriting $f(t)$:

$$f(t) = \frac{1}{2}e^{-\gamma t}e^{2\pi i\omega_0 t}\theta(t) + \frac{1}{2}e^{-\gamma t}e^{-2\pi i\omega_0 t}\theta(t). \tag{3.7}$$

Starting with the first term:

$$\tilde{f}_{+\omega_0}(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-\gamma t}e^{-2\pi i(\omega-\omega_0)t}\theta(t)\,dt \tag{3.8}$$

$$= \frac{1}{2}\int_{0}^{\infty} e^{(-\gamma-2\pi i\omega+2\pi i\omega_0)t}\,dt. \tag{3.9}$$

Evaluating the integral:

$$\tilde{f}_{+\omega_0}(\omega) = \frac{1}{2}\left[\frac{1}{-\gamma-2\pi i(\omega-\omega_0)}e^{(-\gamma-2\pi i\omega+2\pi i\omega_0)t}\right]_0^{\infty} \tag{3.10}$$

$$= \frac{1}{2}\frac{1}{\gamma+2\pi i(\omega-\omega_0)}. \tag{3.11}$$

For the second term, the result is the same with $\omega_0 \to -\omega_0$:

$$\tilde{f}_{-\omega_0}(\omega) = \frac{1}{2}\frac{1}{\gamma+2\pi i(\omega+\omega_0)}. \tag{3.12}$$

The full Fourier transform is:

$$\tilde{f}(\omega) = \tilde{f}_{+\omega_0}(\omega) + \tilde{f}_{-\omega_0}(\omega) \tag{3.13}$$

$$= \frac{1}{2}\left[\frac{1}{\gamma+2\pi i(\omega-\omega_0)} + \frac{1}{\gamma+2\pi i(\omega+\omega_0)}\right] \tag{3.14}$$

3.2 Filters and convolution

We all know filters, since they are all around us. Every room is a filter, our own mouth is a filter, and of course filters are part of any modern audio equipment. Light is filtered by the air etc.

If we think about the characteristics of filters, then one of the most striking one is the fact that it shouldn't matter whether a signal is filter at an earlier time or later on. In other words, a filter is a time-invariant system. Let us denote our filter by L, and we assume that any input signal f is then mapped to an output Lf. We will hope to work with linear filters, so that we arrive at the class of linear, time-invariant systems.

Definition 3.2.1 (Linear, time-invariant (LTI) systems)**.** *A linear operator L that maps functions $f \in V$ to $Lf \in V$, where V is a vector space, is called time-invariant, if*

$$L(f(t-u)) = L(f)(t-u), \text{ equivalently: } L(T_u f) = T_u(Lf).$$

Let us now look at a very fundamental concept, the impulse response. Any LTI-system is completely characterized by its impulse response. That is, for any input function, the output function can be calculated in terms of the input and the impulse response. The impulse response of a linear transformation is the image of Dirac's delta function under the transformation. [1]
We now consider the mathematical derivation of impulse response of an LTI-system L . Note that

$$f(t) = \int_u f(u)\delta_u(t)du = \int_u f(u)\delta(t-u)du \tag{3.15}$$

hence, because L is linear

$$Lf(t) = \int_u f(u)L\delta_u(t)du \tag{3.16}$$

Finally, we use the last property of L, namely time-invariance, to see that $L\delta_u(t) = L(\delta(t-u)) = (L\delta)(t-u)$, hence

$$Lf(t) = \int_u f(u)L\delta_u(t)du = \int_u f(u)(L\delta)(t-u)du \tag{3.17}$$

Setting $h(t) := (L\delta)(t)$, we achieve

$$Lf(t) = \int_u f(u)h(t-u)du =: h * f. \tag{3.18}$$

As we see from (3.34), an LTI-system is completely characterized by its impulse response.

Definition 3.2.2 (Impulse Response)**.** *Let L be an LIT-system. Its impulse response is defined as $h(t) = L\delta(t)$.*

Example 3.2.3 (Discrete Impulse response)**.** *The **Discrete Impulse Response** refers to the output of a discrete-time system when the input is a discrete-time impulse signal, often represented as $\delta[n]$, where $\delta[n]$ is defined as:*

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

*The **Impulse Response** $h[n]$, characterizes the behavior of a (LTI) system. For an digital LTI system, the output for any arbitrary input $x[n]$ can be calculated using the **discrete convolution**:*

[1] In practical situations, it is not possible to produce a true impulse used for testing. Therefore, some other brief, explosive sound is sometimes used as an approximation of the impulse. In acoustic and audio applications, impulse responses enable the acoustic characteristics of a location, such as a concert hall, to be captured. These impulse responses can be used in applications to mimic the acoustic characteristics of a particular location.

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k].$$

Thus, knowing the impulse response $h[n]$ completely determines the system's response to any input. Some properties:

1. ***Causality**: If the system is causal, $h[n] = 0$ for $n < 0$.*
2. ***Stability**: For the system to be stable, the impulse response must satisfy:*

$$\sum_{n=-\infty}^{\infty} |h[n]| < \infty.$$

3. ***FIR (Finite Impulse Response) Systems**: The impulse response $h[n]$ is nonzero only for a finite range of n.*
4. ***IIR (Infinite Impulse Response) Systems**: The impulse response $h[n]$ is nonzero for an infinite range of n, typically decaying over time.*

Definition 3.2.4. *The Fourier transform of a function $f \in L^1(\mathbb{R})$ is defined by*

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t}dt. \tag{3.19}$$

Note that, if f is integrable, the integral in (3.19) converges and

$$|\hat{f}(\omega)| \leq \int_{-\infty}^{\infty} |f(t)|dt < \infty.$$

Proposition 3.2.5 (Inverse Fourier transform)**.** *If $\hat{f} \in L^1(\mathbb{R})$, then f is given by the inverse Fourier transform of $\hat{f}$:*

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\omega)e^{2\pi i\omega t}d\omega \tag{3.20}$$

Remark 3.2.6. *Note that the Fourier transform is usually extended to all functions in $L^2(\mathbb{R})$ by using a density argument, similar to our approach in the proof of Proposition 2.1.8. Then, as before, an inner product can be defined on $L^2(\mathbb{R})$, and most arguments work similar to the case of periodic functions.*

Example 3.2.7 (Fourier transform of the box function)**.** *Consider the function*

$$\Pi(x) := \begin{cases} 1 & for\ -\frac{1}{2} < x < \frac{1}{2} \\ 0 & else \end{cases}$$

To compute the Fourier transform, first note that Π *is even, so that we can omit the sine-part (generally we can observe, that the Fourier transform of even (symmetric) functions is always real! On the other hand, the Fourier transform of real functions is symmetric and we only have to consider the positive frequencies. This property is heavily exploited in the processing of speech and music signals, which are always real.) We therefore have*

$$\begin{aligned}\hat{\Pi}(\omega) &= \int_{\mathbb{R}} \Pi(x) e^{-2\pi i\omega x} dx = \int_{\mathbb{R}} \Pi(x)\cos(2\pi\omega x) dx \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos(2\pi\omega x) dx = \frac{\sin(2\pi\omega x)}{2\pi\omega}\Big|_{x=-\frac{1}{2}}^{\frac{1}{2}} \\ &= \frac{\sin(\pi\omega)}{2\pi\omega} - \frac{\sin(-\pi\omega)}{2\pi\omega} = \frac{\sin(\pi\omega)}{\pi\omega} =: sinc(\omega)\end{aligned}$$

Note that $sinc(x)$ *can be defined as* $sinc = \frac{\sin(\pi x)}{\pi x}$ *only for* $x \neq 0$*. However according to L'Hôpital's rule, we have that* $\lim_{x\to 0} sinc(x) = 1$*, since, for any open interval* I *around* 0*, we have* $h'(x) = 1 \neq 0$ *for* $h(x) = x$ *and, since* $\frac{d}{dx}\sin x = cosx$ *and* $\lim_{x\to 0}\cos(x) = 1$*, we have* $\lim_{x\to 0}\frac{\sin' x}{g'(x)} = 1$*, and so* $\lim_{x\to 0}\frac{\sin x}{x} = 1$*. More directly, we may consider the Taylor series* $\sin x = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n+1}$*, such that* $sinc(x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n}$ *for* $x \neq 0$*, and convergence to* 1 *is obvious.*

We will next address two very basic operators that can act on a function or signal, namely translation, or *time-shift* and modulation, or *frequency-shift.*

Example 3.2.8 (Translation and Modulation)**.** *For any real number* x_0*, if* $g(x) = T_{x_0}f(x) := f(x - x_0)$*, then* $\hat{g}(\omega) = e^{-2\pi i x_0\omega}\hat{f}(\omega)$*.*
For any real number ω_0*, if* $g(x) = M_{\omega_0} := e^{2\pi i x\omega_0}f(x)$*, then* $\hat{g}(\omega) = \hat{f}(\omega - \omega_0)$*.*

Example 3.2.9 (Dilation)**.** *Let* $a \neq 0 \in \mathbb{R}$*. Set* $g(x) = D_a f(x) := \sqrt{a}f(ax)$*. Let* $\hat{f}$ *be the Fourier transform of* f*. Then, the Fourier transform of* g *is given by*

$$\hat{g}(\omega) = \frac{1}{a}\hat{f}(\frac{\omega}{a}).$$

Example 3.2.10. *Consider the Fourier transform of a damped oscillator:*

$$f(t) = e^{-\gamma t}\cos(\omega_0 t)\theta(t), \tag{3.21}$$

where the unit-step function $\theta(t)$ *is defined as:*

$$\theta(t) = \begin{cases} 1, & t > 0, \\ 0, & t \leq 0. \end{cases} \tag{3.22}$$

The step function ensures the oscillator starts at $t = 0$*. Without it, the amplitude would blow up as* $t \to -\infty$*.*

Rewriting $f(t)$:

$$f(t) = \frac{1}{2}e^{-\gamma t}e^{2\pi i\omega_0 t}\theta(t) + \frac{1}{2}e^{-\gamma t}e^{-2\pi i\omega_0 t}\theta(t). \tag{3.23}$$

Starting with the first term:

$$\tilde{f}_{+\omega_0}(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-\gamma t}e^{-2\pi i(\omega-\omega_0)t}\theta(t)\,dt \tag{3.24}$$

$$= \frac{1}{2}\int_{0}^{\infty} e^{(-\gamma-2\pi i\omega+2\pi i\omega_0)t}\,dt. \tag{3.25}$$

Evaluating the integral:

$$\tilde{f}_{+\omega_0}(\omega) = \frac{1}{2}\left[\frac{1}{-\gamma-2\pi i(\omega-\omega_0)}e^{(-\gamma-2\pi i\omega+2\pi i\omega_0)t}\right]_0^{\infty} \tag{3.26}$$

$$= \frac{1}{2}\frac{1}{\gamma+2\pi i(\omega-\omega_0)}. \tag{3.27}$$

For the second term, the result is the same with $\omega_0 \to -\omega_0$:

$$\tilde{f}_{-\omega_0}(\omega) = \frac{1}{2}\frac{1}{\gamma+2\pi i(\omega+\omega_0)}. \tag{3.28}$$

The full Fourier transform is:

$$\tilde{f}(\omega) = \tilde{f}_{+\omega_0}(\omega) + \tilde{f}_{-\omega_0}(\omega) \tag{3.29}$$

$$= \frac{1}{2}\left[\frac{1}{\gamma+2\pi i(\omega-\omega_0)} + \frac{1}{\gamma+2\pi i(\omega+\omega_0)}\right] \tag{3.30}$$

3.3 Filters and convolution

We all know filters, since they are all around us. Every room is a filter, our own mouth is a filter, and of course filters are part of any modern audio equipment. Light is filtered by the air etc.
If we think about the characteristics of filters, then one of the most striking one is the fact that it shouldn't matter whether a signal is filter at an earlier time or later on. In other words, a filter is a time-invariant system. Let us denote our filter by L, and we assume that any input signal f is then mapped to an output Lf. We will hope to work with linear filters, so that we arrive at the class of linear, time-invariant systems.

Definition 3.3.1 (Linear, time-invariant (LTI) systems)**.** *A linear operator L that maps functions $f \in V$ to $Lf \in V$, where V is a vector space, is called time-invariant, if*

$$L(f(t-u)) = L(f)(t-u), \text{ equivalently: } L(T_u f) = T_u(Lf).$$

Let us now look at a very fundamental concept, the impulse response. Any LTI-system is completely characterized by its impulse response. That is, for any input function, the output function can be calculated in terms of the input and the impulse response. The impulse response of a linear transformation is the image of Dirac's delta function under the transformation. [2]
We now consider the mathematical derivation of impulse response of an LTI-system L . Note that

$$f(t) = \int_u f(u)\delta_u(t)du = \int_u f(u)\delta(t-u)du \tag{3.31}$$

hence, because L is linear

$$Lf(t) = \int_u f(u)L\delta_u(t)du \tag{3.32}$$

Finally, we use the last property of L, namely time-invariance, to see that $L\delta_u(t) = L(\delta(t-u)) = (L\delta)(t-u)$, hence

$$Lf(t) = \int_u f(u)L\delta_u(t)du = \int_u f(u)(L\delta)(t-u)du \tag{3.33}$$

Setting $h(t) := (L\delta)(t)$, we achieve

$$Lf(t) = \int_u f(u)h(t-u)du =: h * f. \tag{3.34}$$

As we see from (3.34), an LTI-system is completely characterized by its impulse response.

Definition 3.3.2 (Impulse Response)**.** *Let L be an LIT-system. Its impulse response is defined as $h(t) = L\delta(t)$.*

Example 3.3.3 (Discrete Impulse response)**.** *The **Discrete Impulse Response** refers to the output of a discrete-time system when the input is a discrete-time impulse signal, often represented as $\delta[n]$, where $\delta[n]$ is defined as:*

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

*The **Impulse Response** $h[n]$, characterizes the behavior of a (LTI) system. For an digital LTI system, the output for any arbitrary input $x[n]$ can be calculated using the **discrete convolution**:*

[2] In practical situations, it is not possible to produce a true impulse used for testing. Therefore, some other brief, explosive sound is sometimes used as an approximation of the impulse. In acoustic and audio applications, impulse responses enable the acoustic characteristics of a location, such as a concert hall, to be captured. These impulse responses can be used in applications to mimic the acoustic characteristics of a particular location.

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k].$$

Thus, knowing the impulse response $h[n]$ completely determines the system's response to any input. Some properties:

1. ***Causality**: If the system is causal, $h[n] = 0$ for $n < 0$.*

2. ***Stability**: For the system to be stable, the impulse response must satisfy:*

$$\sum_{n=-\infty}^{\infty} |h[n]| < \infty.$$

3. ***FIR (Finite Impulse Response) Systems**: The impulse response $h[n]$ is nonzero only for a finite range of n.*

4. ***IIR (Infinite Impulse Response) Systems**: The impulse response $h[n]$ is nonzero for an infinite range of n, typically decaying over time.*

Here are the general definitions of convolution:

Definition 3.3.4 (Convolution)**.** *The convolution of two functions $f, g \in L^1(\mathbb{R})$ is defined by*

$$(f * g)(t) = \int_u f(u)g(t-u)du = \int_u g(u)f(t-u)du = (g * f)(t) \tag{3.35}$$

For functions f, g on $\mathbb{Z}$, we define

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m] = \sum_{m=-\infty}^{\infty} g[m]f[n-m] = (g * f)[n] \tag{3.36}$$

For functions f, g on $\mathbb{Z}_N$, which are simply vectors in $\mathbb{C}^N$, periodically extended, i.e., we let $f[n] = f[n']$ and $g[n] = g[n']$ if $n \equiv n'(mod N)$ and define

$$(f * g)[n] = \sum_{m=0}^{N-1} f[m]g[n-m]du = \sum_{m=0}^{N-1} g[m]f[n-m]du = (g * f)[n]. \tag{3.37}$$

The following examples help understand the action of convolution.

Example 3.3.5. *Linear averaging over $[-T, T]$:*

$$Lf(t) = \frac{1}{2T}\int_{t-T}^{t+T} f(u)du = \frac{1}{2T}\int_t 1_{[-T,T]}(t-u)f(u)du = (1_{[-T,T]} * f)(t)$$

Excursus: eigenfunctions

Eigenfunctions of a linear operator L defined on some function space are non-zero functions h such that

$$Lh = \lambda h$$

for some scalar λ, the corresponding eigenvalue. In the theory of signals and systems, the eigenfunction of a system is a signal h which produces a scalar multiple (possibly complex) of itself as a response to the system:

$$Lh = \lambda h, \qquad \lambda \in \mathbb{C}.$$

Now assume that there exists an orthonormal basis (ONB) $\{\varphi_k\}_{k\in\mathbb{Z}}$ of V consisting of eigenfunctions of a mapping (system, operator) L, i.e.

$$f = \sum_k c_k \varphi_k, \quad \text{for all } f \in V.$$

Then:

$$Lf = \sum_k c_k L\varphi_k = \sum_k \lambda_k c_k \varphi_k.$$

In other words, L acts as a **multiplication operator** on the coefficients of the function's expansion.

End of excursus.

Since we have already seen that complex exponentials provide natural signal expansions with clear interpretations, let us next consider what happens when an LTI system acts on them. Intuitively, since LTI systems are filters, we should expect that complex exponentials corresponding to a particular frequency are merely amplified, damped, or phase-shifted.

Now, for an LTI system L with impulse response h, we have:

$$Le^{2\pi i\omega t} = \int_{\mathbb{R}} e^{2\pi i\omega u} h(t-u)\, du \tag{3.38}$$

$$= \int_{\mathbb{R}} e^{2\pi i\omega (t-u)} h(u)\, du \tag{3.39}$$

$$= e^{2\pi i\omega t} \int_{\mathbb{R}} e^{-2\pi i\omega u} h(u)\, du = e^{2\pi i\omega t}\, \hat{h}(\omega). \tag{3.40}$$

Thus, $e^{2\pi i\omega t}$ is an eigenfunction of any LTI system, with eigenvalue $\hat{h}(\omega)$.

Proposition 3.3.6 (Convolution relation)**.** *For $f, g \in L^1(\mathbb{R}^d)$, we have*

$$\widehat{f * g} = \hat{f}\, \hat{g}.$$

Proof. Theorem A.1.3 ensures $f * g \in L^1(\mathbb{R}^d)$. Fubini's theorem gives

$$\int_{\mathbb{R}} |(f*g)(t)|\,dt = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(t-s)g(s)\,ds \right| dt \leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(t-s)||g(s)|\,ds\,dt = \|f\|_1 \cdot \|g\|_1.$$

Now compute:

$$\begin{aligned}
\widehat{f*g}(\xi) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y)g(x-y)e^{-2\pi i\langle x,\xi\rangle}\,dy\,dx \\
&= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x-y)e^{-2\pi i\langle x-y,\xi\rangle}\,dx \right) e^{-2\pi i\langle y,\xi\rangle}\,dy \\
&= \int_{\mathbb{R}^d} f(y)\,\hat{g}(\xi)\,e^{-2\pi i\langle y,\xi\rangle}\,dy = \hat{f}(\xi)\,\hat{g}(\xi).
\end{aligned}$$

□

Example 3.3.7 (Convolution of box functions)**.** *Define the triangle function:*

$$\Lambda(x) = \begin{cases} 1-|x|, & |x| \leq 1, \\ 0, & \textit{else.} \end{cases}$$

Then for $I = [-\frac{1}{2}, \frac{1}{2}]$, we have

$$(\Pi_I * \Pi_I)(x) = \Lambda(x),$$

since

$$\begin{aligned}
(\Pi_I * \Pi_I)(y) &= \int_{-\infty}^{\infty} \Pi_I(x)\Pi_I(y-x)\,dx \\
&= \begin{cases} \int_{-\frac{1}{2}}^{y+\frac{1}{2}} 1\,dx, & -1 < y \leq 0, \\ \int_{y-\frac{1}{2}}^{\frac{1}{2}} 1\,dx, & 0 < y < 1. \end{cases} \\
&= \begin{cases} y+1, & -1 < y \leq 0, \\ 1-y, & 0 < y < 1, \\ 0, & \textit{else.} \end{cases}
\end{aligned}$$

See also: `https://en.wikipedia.org/wiki/File:Convolution_of_box_signal_with_itself2.gif`

Example 3.3.8 (Convolution with sinc function via Fourier transforms)**.** *The sinc function is defined as:*

$$sinc(t) = \frac{\sin(\pi t)}{\pi t}.$$

The convolution of two functions $f(t)$ and $g(t)$ is given by:

$$(f*g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\,d\tau.$$

By the convolution theorem,

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\}.$$

Since the Fourier transform of sinc(t) is the box function:

$$\mathcal{F}\{sinc(t)\} = \Pi(\omega),$$

we have:

$$\mathcal{F}\{f * sinc\}(\omega) = F(\omega)\,\Pi(\omega).$$

Hence,

$$(f * sinc)(t) = \mathcal{F}^{-1}\{F(\omega)\Pi(\omega)\},$$

*which shows that convolution with a sinc function acts as a **low-pass filter**.*

Proposition 3.3.9 (Fourier invariance of the Gaussian)**.** *For $\varphi : \mathbb{R}^d \to \mathbb{R}$, $\varphi(x) = e^{-\pi\|x\|^2}$, we have*

$$\widehat{\varphi}(\xi) = \varphi(\xi),$$

i.e., the Gaussian is invariant under the Fourier transform.

Proof. We start in one dimension with $f(x) = e^{-\pi x^2}$.

The Fourier transform is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx.$$

We first find a differential equation satisfied by $\hat{f}$. Differentiate $\hat{f}(\xi)$ with respect to ξ:

$$\hat{f}'(\xi) = \int_{\mathbb{R}} e^{-\pi x^2} \frac{d}{d\xi}(e^{-2\pi i x\xi})\, dx = -2\pi i \int_{\mathbb{R}} x e^{-\pi x^2} e^{-2\pi i x\xi}\, dx.$$

Notice that the right-hand side can be expressed in terms of the Fourier transform of the derivative of f. Since

$$f'(x) = -2\pi x e^{-\pi x^2},$$

we have

$$\mathcal{F}\{f'\}(\xi) = 2\pi i \xi \hat{f}(\xi).$$

Combining both expressions gives the differential equation:

$$\frac{d}{d\xi}\hat{f}(\xi) = -2\pi\xi \hat{f}(\xi).$$

This is a separable ODE:

$$\frac{\hat{f}'(\xi)}{\hat{f}(\xi)} = -2\pi\xi.$$

Integrating both sides yields:

$$\ln \hat{f}(\xi) = -\pi\xi^2 + C \quad \Rightarrow \quad \hat{f}(\xi) = Ce^{-\pi\xi^2}.$$

To find C, evaluate at $\xi = 0$:

$$\hat{f}(0) = \int_{\mathbb{R}} e^{-\pi x^2}\, dx = 1,$$

since $\int_{\mathbb{R}} e^{-\pi x^2} dx = 1$. Thus $C = 1$, and

$$\hat{f}(\xi) = e^{-\pi\xi^2}.$$

In d dimensions, separability implies

$$\widehat{e^{-\pi\|x\|^2}}(\xi) = e^{-\pi\|\xi\|^2},$$

since the multidimensional Gaussian factorizes into products of one-dimensional Gaussians. Hence the Gaussian is invariant under the Fourier transform. □

As an exercise, prove the following generalization: The dilated Gaussian $\varphi_a(t) = e^{-\pi t^2/a}$ has the Fourier transform $\hat{\varphi_a}(\omega) = \sqrt{a}e^{-\pi a\omega^2}$.

3.3.1 The Fourier transform on L^2

Definition 3.3.10 (Fourier transform and inverse Fourier transform)**.** *Let $\langle \xi, x\rangle$ denote the standard inner product in $\mathbb{R}^d$. For $f \in L^1(\mathbb{R}^d)$, we call*

$$(\mathcal{F}f)(\xi) := \hat{f}(\xi) := \int_{\mathbb{R}^d} f(x)\mathrm{e}^{-2\pi\mathrm{i}\langle \xi, x\rangle}\mathrm{d}x \tag{3.41}$$

the Fourier transform *of f.*
For $f \in L^1(\mathbb{R}^d)$, we define the inverse Fourier transform $\check{f} : \mathbb{R}^d \to \mathbb{C}$ by

$$\check{f}(\xi) := \hat{f}(-\xi) = \int_{\mathbb{R}^d} f(x)\mathrm{e}^{2\pi i\langle x, \xi\rangle}\mathrm{d}x.$$

Lemma 3.3.11. *For $f \in L^1(\mathbb{R}^d)$, we have $\hat{f} \in \mathcal{C}(\mathbb{R}^d)$ and $\|\hat{f}\|_\infty \le \|f\|_{L^1}$.*

Proof. We estimate

$$\left|\hat{f}(\xi)\right| = \left|\int_{\mathbb{R}^d} f(x)\mathrm{e}^{-2\pi\mathrm{i}\langle x, \xi\rangle}\mathrm{d}x\right| \le \int_{\mathbb{R}^d} |f(x)|\mathrm{d}x.$$

If $\xi_n \to \xi$, then $f(x)\mathrm{e}^{-2\pi\mathrm{i}\langle \xi_n, x\rangle} \to f(x)\mathrm{e}^{-2\pi\mathrm{i}\langle \xi, x\rangle}$, for all $x \in \mathbb{R}^d$. Since $\left|f(x)\mathrm{e}^{-2\pi\mathrm{i}\langle \xi_n, x\rangle}\right| \le |f(x)|$, the dominated convergence Theorem A.1.2 implies $\hat{f}(\xi_n) \to \hat{f}(\xi)$. □

For $f \in L^1(\mathbb{R}^d)$, we would like to apply the "inverse" Fourier transform to $\hat{f}$ but Lemma 3.3.11 suggests that we could be faced with $\hat{f} \notin L^1(\mathbb{R}^d)$. Therefore, we replace $L^1(\mathbb{R}^d)$ with a "nicer" function space that better fits to the Fourier transform. For $f \in \mathcal{C}^\infty(\mathbb{R}^d)$ and $\alpha, \beta \in \mathbb{N}^d$, define

$$p_{\alpha,\beta}(f) := \sup_{x \in \mathbb{R}^d} |x^\alpha \partial^\beta f(x)|$$

Definition 3.3.12 (Schwartz space)**.** *The* Schwartz space *is*

$$\mathcal{S}(\mathbb{R}^d) := \{f \in \mathcal{C}^\infty(\mathbb{R}^d) : \forall \alpha, \beta \in \mathbb{N}^d \; p_{\alpha,\beta}(f) < \infty\},$$

where $f_n \to f$ if and only if $p_{\alpha,\beta}(f_n - f) \to 0 \; \forall \alpha, \beta \in \mathbb{N}^d$.

Proposition 3.3.13. *For $f, g \in \mathcal{S}(\mathbb{R}^d)$, we have*

$$\forall \alpha \in \mathbb{N}^d \qquad \partial^\alpha f, \quad f \cdot g, \quad f * g \quad \in \; \mathcal{S}(\mathbb{R}^d),$$

and $\partial^\alpha (f * g) = (\partial^\alpha f) * g$.

Proof. [1]. □

Although

$$\Delta : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \tag{3.42}$$

there are other issues...

Lemma 3.3.14. *For $f \in \mathcal{S}(\mathbb{R}^d)$ and $\alpha \in \mathbb{N}^d$, we have*

$$\widehat{\partial^\alpha f}(\xi) = (2\pi \mathrm{i}\xi)^\alpha \hat{f}(\xi), \qquad \partial^\alpha \hat{f} = \widehat{((-2\pi \mathrm{i}\cdot)^\alpha f)}.$$

Proof. Integration by parts with $f(\pm\infty) = 0$ yields

$$\begin{aligned}
\widehat{\frac{\partial}{\partial x_k} f}(\xi) &= \int_{\mathbb{R}^d} \frac{\partial}{\partial x_k} f(x) \mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} \mathrm{d}x \\
&= -\int_{\mathbb{R}^d} f(x) \frac{\partial}{\partial x_k} \mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} \mathrm{d}x \\
&= -\int_{\mathbb{R}^d} f(x)(-2\pi \mathrm{i}\xi_k) \mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} \mathrm{d}x = 2\pi \mathrm{i}\xi_k \hat{f}(\xi).
\end{aligned}$$

Interchange of differentiation and integration yields

$$\begin{aligned}
(\frac{\partial}{\partial_k \xi} \hat{f})(\xi) &= \frac{\partial}{\partial_k \xi} \int_{\mathbb{R}^d} f(x) \mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} \mathrm{d}x \\
&= \int_{\mathbb{R}^d} \frac{\partial}{\partial_k \xi} \left(f(x) \mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} \right) \mathrm{d}x \\
&= -\int_{\mathbb{R}^d} 2\pi \mathrm{i} x_k f(x) \mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} \mathrm{d}x.
\end{aligned}$$

□

Lemma 3.3.14 implies that (3.42) is equivalent to

$$4\pi^2\|\xi\|^2\hat{u}(\xi) = \hat{f}(\xi), \qquad \xi \in \mathbb{R}^d. \tag{3.43}$$

Definition 3.3.15 (Approximate identity)**.** *An approximate identity is a family $(u_\epsilon)_{\epsilon>0} \subset L^1(\mathbb{R}^d)$ such that it holds:*

(a) $\exists c > 0$ *such that* $\|u_\epsilon\|_{L^1} \leq c$, $\forall \epsilon > 0$,

(b) $\int u_\epsilon = 1$, $\forall \epsilon > 0$,

(c) for any neighborhood U of 0,

$$\int_{X\setminus U} |u_\epsilon| \xrightarrow{\epsilon\to 0} 0.$$

Proposition 3.3.16. *Let $\mathcal{U} = (u_\epsilon)_\varepsilon$ be an approximate identity. Then for $f \in L^1(\mathbb{R}^d)$, it holds that*

$$\lim_{\varepsilon\to 0} \|f - f * u_\epsilon\|_1 = 0$$

Furthermore, there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$, $\varepsilon_k > 0$ such that

$$\lim_{k\to\infty} f * u_\epsilon(t) = f(t) \qquad a.e.$$

Moreover, if f is continuous (or uniformly continuous and bounded), then $\lim_{\epsilon\to 0}(f * u_\epsilon)(t) = f(t)$ pointwise (or uniformly).

Proof. **Step 1: Approximation for Simple Functions**

First, consider $f = \chi_Q$, where $Q \subset \mathbb{R}$ is a measurable subset. We aim to show:

Using the definition of convolution:

Applying Fubini's theorem to interchange the order of integration:

Define $q_s(t) = \left|\chi_Q(t) - \chi_Q(t-s)\right|$. Then:

Step 2: Estimating $q_s(t)$

For fixed s, note that $q_s(t) = \chi_{Q\cup(s+Q)}(t) - \chi_{Q\cap(s+Q)}(t)$. The measure of the symmetric difference $(Q \cup (s+Q)) \setminus (Q \cap (s+Q))$ can be made arbitrarily small for small $|s|$:

Thus, for $s \in [-\delta, \delta]$:

Step 3: Completing the Argument

Using the properties of g_ϵ:

Split the integral into two parts:

For $s \notin [-\delta, \delta]$, the integral $\int_{\mathbb{R}} q_s(t)dt$ is bounded by $2|\chi_Q|1$, and $\int \mathbb{R} \setminus [-\delta, \delta] g_\epsilon(s)ds \to 0$ as $\epsilon \to 0$. For $s \in [-\delta, \delta]$, $\int_{\mathbb{R}} q_s(t)dt < \tau$:

Since $\tau > 0$ is arbitrary:

Step 4: Extending to General $f \in L^1(\mathbb{R})$

For general $f \in L^1(\mathbb{R})$, given $\tau > 0$, choose a step function h such that $|f - h|_1 < \tau$. Using the triangle inequality:

The first term is less than τ by construction, and the second term converges to 0 as $\epsilon \to 0$ because h is a step function. The third term is bounded by $|h - f|1 \cdot |g\epsilon|_1 < \tau$. Thus:

Step 5: Almost Everywhere Pointwise Convergence

The almost everywhere convergence follows from the Riesz-Fischer theorem, which states that L^1 convergence implies the existence of a subsequence ϵ_k with $\lim_{k\to\infty} \epsilon_k = 0$ such that $f * g_{\epsilon_k}(t) \to f(t)$ almost everywhere. □

Example 3.3.17. *The family $\mathcal{U} = (u_\epsilon)_\varepsilon$, defined as*

$$g_\varepsilon(t) = e^{\frac{-\pi x^2}{\varepsilon}}$$

is an approximate identity.

Theorem 3.3.18 (Inversion of Fourier transform on L^1)**.** *Let $f \in L^1(\mathbb{R})$ with $\hat{f} \in L^1(\mathbb{R})$.*

Then the function f can be reconstructed (almost everywhere) from $\hat{f}$ using the formula:

$$f(t) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i \omega t} d\omega.$$

In particular, f can be modified on a set of measure zero to yield a continuous function.

Proof. Formally:

$$\int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i \omega t}\, d\omega = \int_{\mathbb{R}} \int_{\mathbb{R}} f(s) e^{-2\pi i \omega s} e^{2\pi i \omega t}\, ds\, d\omega.$$

Interchange integration (Fubini's theorem):

$$= \int_{\mathbb{R}} f(s) \left(\int_{\mathbb{R}} e^{2\pi i \omega (t-s)}\, d\omega \right) ds.$$

Using the property of the Dirac delta:

$$\int_{\mathbb{R}} e^{2\pi i \omega x}\, d\omega = \delta(x),$$

it follows:

$$f(t) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i \omega t}\, d\omega.$$

Technically: Let us write out the integral we aim to study:

$$\int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i \omega t} d\omega = \int_{\mathbb{R}} \int_{\mathbb{R}} f(s) e^{-2\pi i \omega s} e^{2\pi i \omega t} ds d\omega.$$

We multiply the integrand by a Gaussian weight function to ensure the integral converges more rapidly. Define:

$$I_\epsilon(t) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i\omega t} e^{-\epsilon^2\omega^2/2} d\omega,$$

where $\epsilon > 0$ is arbitrary. Since $\hat{f} \in L^1(\mathbb{R})$, the dominated convergence theorem implies:

$$\lim_{\epsilon\to 0} I_\epsilon(t) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i\omega t} d\omega, \quad \text{for all } t \in \mathbb{R}.$$

Next, we analyze $I_\epsilon(t)$ by interchanging the order of integration, which is valid by the Fubini-Tonelli theorem due to the Gaussian factor. This gives:

$$I_\epsilon(t) = \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} e^{-2\pi i\omega s} e^{2\pi i\omega t} e^{-\epsilon^2\omega^2/2} d\omega ds.$$

The inner integral is a Gaussian-weighted Fourier transform:

$$\int_{\mathbb{R}} e^{2\pi i\omega(t-s)} e^{-\epsilon^2\omega^2/2} d\omega = \sqrt{\frac{2\pi}{\epsilon^2}} e^{-\frac{(t-s)^2}{2\epsilon^2}}.$$

Substituting this result, we obtain:

$$I_\epsilon(t) = \int_{\mathbb{R}} f(s) \sqrt{\frac{2\pi}{\epsilon^2}} e^{-\frac{(t-s)^2}{2\epsilon^2}} ds = f * u_\epsilon(t),$$

where $u_\epsilon(t) = \sqrt{\frac{2\pi}{\epsilon^2}} e^{-\frac{t^2}{2\epsilon^2}}$ is a Gaussian kernel.

By Proposition 3.3.16 on approximate identities, there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$, $\varepsilon_k > 0$ such that: $\lim_{k\to\infty} I_{\epsilon_k}(t) = f(t)$, almost everywhere, hence

$$\int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i\omega t} d\omega = \lim_{\epsilon\to 0} I_\epsilon(t) = f(t), \quad \text{almost everywhere}$$

Since the function $\int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i\omega t} d\omega$ is continuous when $\hat{f} \in L^1(\mathbb{R})$, f can be modified in a set of zero measures to become continuous.

□

Corollary 3.3.19. *Suppose that $f \in L^1(\mathbb{R})$ is continuous in 0 and that $\hat{f}(\omega) \geq 0$ for all $\omega \in \mathbb{R}$. Then*

$$f(0) = \int_{\mathbb{R}} \hat{f}(\omega) \mathrm{d}\omega.$$

Using the previous result we can show the following.

Theorem 3.3.20 (Isometry of Fourier transform)**.** *For $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, we have*

$$\|f\|_{L^2}^2 = \|\hat{f}\|_{L^2}^2. \tag{3.44}$$

Furthermore, we have, by polarization [3]*, for* $f, g \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$,

$$\langle f, g\rangle_{L^2} = \langle \hat{f}, \hat{g}\rangle_{L^2}. \tag{3.45}$$

Proof. Consider $f' \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ defined as $f'(t) = \overline{f(-t)}$. It is easy to see that the Fourier transform of f' is $\hat{f}'(\omega) = \overline{\hat{f}(\omega)}$. Define

$$F = f * f' \quad \text{(the convolution of } f \text{ and } f').$$

As an exercise, show that F is continuous. Moreover, the Fourier transform of F is given by:

$$\hat{F}(\omega) = |\hat{f}(\omega)|^2 \geq 0.$$

By Corollary 3.3.19, it follows that

$$\|f\|_2^2 = F(0) = \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 d\omega.$$

Using polarization, we get the fundamental result related to the preservation of the L^2-inner product. □

We now extend $\mathcal{F}$ to $L^2(\mathbb{R}^d)$: for $f \in L^2(\mathbb{R}^d, \mathbb{C})$,

$$\int_{\mathbb{R}^d} \underbrace{f(x)\mathrm{e}^{-2\pi \mathrm{i}\langle x,\xi\rangle}}_{\notin L^1} \mathrm{d}x$$

is not well-defined.

Theorem 3.3.21 (Unitarity of Fourier transform on L^2)**.** *$\mathcal{F}$ can be uniquely and continuously extended from $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ to an isometric isomorphism $\mathcal{F} : L^2(\mathbb{R}^d) \mapsto L^2(\mathbb{R}^d)$, that is, for all $f, g \in L^2(\mathbb{R}^d)$ we have*

$$\langle f, g\rangle_{L^2} = \langle \hat{f}, \hat{g}\rangle_{L^2}. \tag{3.46}$$

and

$$f = \mathcal{F}^*\mathcal{F}f,$$

where $\mathcal{F}^*$ is the adjoint of $\mathcal{F}$, as defined in Definition 3.3.10:

$$\mathcal{F}^* f(\xi) = \int_{\mathbb{R}^d} f(x)\mathrm{e}^{2\pi i\langle x,\xi\rangle}\mathrm{d}x.$$

[3] For a complex inner product space V, the inner product is generally conjugate symmetric, i.e. $\langle x, y\rangle = \overline{\langle y, x\rangle}$. The polarization identity for complex spaces is:

$$\langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big).$$

Proof. This is a standard density proof. Let $f \in L^2(\mathbb{R})$. For each $n \in \mathbb{N}$, define

$$f_n := f\,\chi_{[-n,n]}.$$

Since f_n has compact support and $f \in L^2(\mathbb{R})$, it follows that

$$f_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \quad \text{for all } n.$$

Moreover, because $\chi_{[-n,n]} \to 1$ pointwise and monotonically, we have

$$\|f - f_n\|_{L^2}^2 = \int_{\mathbb{R}} |f(x)|^2\, \chi_{\mathbb{R}\setminus[-n,n]}(x)\, dx \longrightarrow 0 \quad \text{as } n \to \infty,$$

by the monotone convergence theorem. Hence, (f_n) converges to f in $L^2(\mathbb{R})$ and is therefore a Cauchy sequence in $L^2(\mathbb{R})$.

By Plancherel's theorem, the Fourier transform is an isometry on $L^2(\mathbb{R})$, so

$$\|\widehat{f_n} - \widehat{f_m}\|_{L^2} = \|f_n - f_m\|_{L^2} \quad \text{for all } m, n \in \mathbb{N}.$$

Thus, $(\widehat{f_n})$ is a Cauchy sequence in $L^2(\mathbb{R})$ and hence converges to a limit in $L^2(\mathbb{R})$. Furthermore, this limit does not depend on the particular approximating sequence (f_n).

We therefore define the Fourier transform of f by

$$\widehat{f} := \lim_{n\to\infty} \widehat{f_n} \quad \text{in } L^2(\mathbb{R}).$$

This defines a unique linear operator

$$\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$$

which is an isometric extension of the Fourier transform on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. In particular,

$$\|\widehat{f}\|_{L^2} = \|f\|_{L^2} \quad \text{for all } f \in L^2(\mathbb{R}),$$

and all standard identities of the Fourier transform extend to $L^2(\mathbb{R})$ by density. □

3.4 The finite discrete Fourier transform

So far, we have been dealing with functions/signals that are

1. continuous in time and have infinite duration; they have a Fourier transform that is continuous in time and has infinite bandwidth ("duration in frequency").

2. discrete in time and have infinite duration; they have a Fourier transform that is continuous in frequency and has finite bandwidth.

3. continuous in time and have finite duration (in other words; they are periodic); they have a Fourier transform that is discrete in frequency and has infinite duration.

We notice that finite duration (or, ly, periodicity) in one domain (i.e. in time or frequency) leads to discreteness in the other domain. We now address the fourth case, which is the case of signals, that are both finite in duration *and* discrete, or, discrete and periodic.

Remark 3.4.1. *We want to point out that finite duration is equivalent to periodicity only in the sense, that the entire information that is contained in the signal is actually contained in an interval of finite length. in this sense, a periodic function can be identified with a function supported on an interval of finite length or on the torus* $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. *As we will see later, in the context of sampling, we may have to distinguish meticulously between periodic signals and signals that are supported in an interval of finite duration and are zero elsewhere.*

Definition 3.4.2 (Finite discrete Fourier transform)**.** *The finite discrete Fourier transform of* $f \in \mathbb{C}^N$, *i.e. of a vector of* N *complex numbers is given by*

$$\mathcal{F}'f[k] = \hat{f}[k] = F[k] = \frac{1}{N}\sum_{n=0}^{N-1} f[n] \cdot e^{-i2\pi\frac{k}{N}n}. \tag{3.47}$$

The inverse transform yields the expansion of f *as*

$$\mathcal{F}\mathcal{F}'f[n] = \mathcal{F}\hat{f}[n] = f[n] = \sum_{k=0}^{N-1} F[k] \cdot e^{i2\pi\frac{n}{N}k}. \tag{3.48}$$

Remark 3.4.3. *The Fourier transforms we discussed so far, gave us information about the amount of any pure frequency, i.e. complex exponentials, present in a given signal. Obviously, the sinusoid must have the same basic properties as the underlying signal under consideration: for periodic signals with a certain period* p, *we only considered sinusoids with the same period, for discrete-time signals we only considered discrete-time sinusoids, whereas, for continuous time signals of infinite duration, any complex exponential is a candidate in the expansion* (3.19).
For the finite, discrete signals, which are in fact vectors in $\mathbb{C}^N$, *we may ask, how many complex exponentials are eligible for the definition of a corresponding Fourier transform. We have two criteria:*
(a) they should be periodic with length N, *that is, we require that*

$$e^{2\pi is(n+N)} = e^{2\pi isn}, \ \textit{for all } n,$$

which means that $s = \frac{k}{N}$.
(b) we observe that for all $m \in \mathbb{Z}$:

$$e^{2\pi ik\frac{n}{N}} = e^{2\pi i(k+mN)\frac{n}{N}}, \ \textit{for all } n,$$

which is a similar phenomenon as observed for discrete-time sinusoids before. This means, that, since $s = \frac{k}{N}, s = \frac{k\pm N}{N}, s = \frac{k\pm 2N}{N}, \ldots$ *all give the same signals, we have only* N

distinct sinusoid adequate for analyzing our N-finite, discrete signals, namely $e^{2\pi ik\frac{n}{N}}$*, for* $k = 0, \dots, N-1$.
Of course, this is exactly what you should have expected: since the complex exponentials have provided ONBs so far, they should provide an orthonormal basis for $\mathbb{C}^N$ *as well. Obviously, this means, that there should be* N *of them.*

Example 3.4.4. *Note that for* $k = 0$, $e^{2\pi ik\frac{n}{N}}$ *is constantly equal to* 1*, then, the rotation of the vector* $e^{2\pi ik\frac{n}{N}}$*, that rotates, as* n *goes from* 0 *to* $N-1$*, accelerates with growing* k*:* $k = 1$ *corresponds to a single rotation,* $k = 2$ *to* 2 *rotations, etc., up to* $N/2$*, from where the frequencies decrease, since they become negative.*

We next show, that the vectors $e^{2\pi ik\frac{n}{N}}$ in fact form an orthogonal basis.

Proposition 3.4.5 (Discrete Fourier basis)**.** *The vectors* s_k*,* $k = 0, \dots, N-1$*, with entries* $s_k[n] = e^{2\pi ik\frac{n}{N}}$ *are orthogonal in* $\mathbb{C}^N$*. The set* $\{\frac{1}{\sqrt{N}}s_k, k = 0, \dots, N-1\}$ *is an ONB.*

Proof.

$$\begin{aligned}\langle s_k, s_l\rangle &= \sum_{n=0}^{N-1} s_k[n]\overline{s_l[n]} \\ &= \sum_{n=0}^{N-1} e^{2\pi ik\frac{n}{N}} e^{-2\pi il\frac{n}{N}} \\ &= \sum_{n=0}^{N-1} e^{2\pi i(k-l)\frac{n}{N}} = \frac{1-e^{2\pi i(k-l)}}{1-e^{2\pi i(k-l)/N}}\end{aligned} \tag{3.49}$$

where the last step follows from the well-known formula for geometric series: $\sum_{n=0}^{N-1} z^n = \frac{1-z^N}{1-z}$. Now, (3.49) is zero, if $k \neq l$, and for $k = l$, we evaluate the sum as $\sum_{n=0}^{N-1} e^{2\pi i(k-l)\frac{n}{N}} = \sum_{n=0}^{N-1} 1 = N$, therefore, the normalization $\frac{1}{\sqrt{N}}s_k$ leads to $\langle \frac{1}{\sqrt{N}}s_k, \frac{1}{\sqrt{N}}s_k\rangle = \frac{1}{N}\|s_k\|_2^2 = 1$. □

Example 3.4.6.
- *The Delta Function*
- *The Constant Function*
- *The Delta train or Dirac comb on* $\mathbb{C}^N$
  When $m = 1, 2, \dots$ *divides* N *, we define the Dirac comb as*

$$\text{Ш}_m[n] = \begin{cases} 1 & \text{if } n = 0, \pm m, \pm 2m, \dots \\ 0 & \text{otherwise} \end{cases} \tag{3.50}$$

  Here, m *specifies the spacing between the "teeth" hence* $m' := N/m$ *is the number of teeth. We can easily verify that* Ш_m *has the Fourier transform*

$$\widehat{\text{Ш}}_m[k] = \frac{1}{m}\text{Ш}_{N/m}[k]. \tag{3.51}$$

- *Periodicity on $\mathbb{C}^N$:*
  Let $N = m \cdot m'$, $m, m' \in \mathbb{N}^+$ and assume that f is m-periodic on $\mathbb{C}^N$. We show that $\hat{f}[k] = 0$ if k is not a multiple of m':
  Since f is m-periodic, we have $f[n+m] - f[n] = 0$. Now, since, for $n_0 \in \mathbb{Z}$

  $$g[n] = f[n - n_0] \text{ has the Fourier transform } \hat{g}[k] = e^{-2\pi i k n_0/N} \hat{f}[k],$$

  we may write

  $$(e^{2\pi i k m/N} - 1)\hat{f}[k] = (e^{2\pi i k/m'} - 1)\hat{f}[k] = 0$$

  hence $\hat{f}[k] = 0$ if $m' \neq k$.

Now this is all really nice, but maybe also sometimes confusing; it seems that different versions of the Fourier transform are appropriate for different situations and have various advantages and disadvantages. Isn't there something like a unifying approach. Indeed there is! And since it is based on a smart and important mathematical idea, we introduce its basics in the next section.

3.5 Tempered distributions

Recall the definition of Schwartz space $\mathcal{S}$ (Definition 3.3.12). Since neither $\mathcal{S}(\mathbb{R}^d)$ nor $L^2(\mathbb{R}^d)$ are suitable to deal with certain derivatives [4], we need to introduce a new space.

Definition 3.5.1. *The* tempered distributions *are the elements of*

$$\mathcal{S}'(\mathbb{R}^d) = \{L : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C} \mid L \text{ is linear and continuous}\}.$$

We endow $\mathcal{S}'(\mathbb{R}^d)$ with the weak*-topology, so that, for $L_n, L \in \mathcal{S}'(\mathbb{R}^d)$,

$$L_n \to L \qquad \Leftrightarrow \qquad \forall \eta \in \mathcal{S}(\mathbb{R}^d): \quad L_n(\eta) \to L(\eta).$$

For $f : \mathbb{R}^d \to \mathbb{C}$ measurable, we write

$$L_f(\eta) := \int_{\mathbb{R}^d} f(x)\eta(x)\mathrm{d}x, \qquad \eta \in \mathcal{S}(\mathbb{R}^d).$$

If $L_f \in \mathcal{S}'(\mathbb{R}^d)$, we simply write $f \in \mathcal{S}'(\mathbb{R}^d)$. We also define

$$L^1_{loc}(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{C} \mid f \text{ measurable}, \ 1_K \cdot f \in L^1(\mathbb{R}^d) \ \forall K \subset \mathbb{R}^d \text{ compact}\}$$

Proposition 3.5.2. *If $f \in L^1_{loc}(\mathbb{R}^d)$ and $\exists N \in \mathbb{N}$: $\lim_{\|x\|\to\infty} \frac{|f(x)|}{\|x\|^N} \to 0$, then $f \in \mathcal{S}'(\mathbb{R}^d)$.*

Proof. [1]. □

[4] Note that, according to Lemma 3.3.14, given $f \in \mathcal{S}(\mathbb{R}^d)$, it could be that $\frac{1}{4\pi^2\|\cdot\|^2}\hat{f}(\cdot) \notin \mathcal{S}(\mathbb{R}^d)$ (and $\notin L^2(\mathbb{R}^d)$).

Example 3.5.3. - $1 \in \mathcal{S}'(\mathbb{R}^d)$, $1_{[0,\infty)} \in \mathcal{S}'(\mathbb{R})$.

- $\mathcal{S}(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$ *by Theorem 3.5.2.*

- *For* $x_0 \in \mathbb{R}^d$, *we have* $\delta_{x_0} \in \mathcal{S}'(\mathbb{R}^d)$, *where* $\delta_{x_0}(\eta) := \eta(x_0)$, *for* $\eta \in \mathcal{S}'(\mathbb{R}^d)$.

Definition 3.5.4. *For* $L \in \mathcal{S}'(\mathbb{R}^d)$ *and* $g \in \mathcal{S}(\mathbb{R}^d)$, *we define* $L \cdot g \in \mathcal{S}'(\mathbb{R}^d)$ *by*

$$L \cdot g(\eta) := L(g \cdot \eta), \qquad \eta \in \mathcal{S}(\mathbb{R}^d).$$

For $f, g \in \mathcal{S}(\mathbb{R}^d)$, we observe $L_f \cdot g = L_{f \cdot g}$.

Definition 3.5.5. *For* $L \in \mathcal{S}'(\mathbb{R}^d)$ *and* $g \in \mathcal{S}(\mathbb{R}^d)$,

$$L * g(\eta) := L(g^- * \eta), \qquad \eta \in \mathcal{S}(\mathbb{R}^d),$$

yields $L * g \in \mathcal{S}'(\mathbb{R}^d)$, *where* $g^-(x) := g(-x)$, *for* $x \in \mathbb{R}^d$.

This extends the convolution of functions:

Lemma 3.5.6. *For* $f, g \in \mathcal{S}(\mathbb{R}^d)$, *we have* $L_f * g = L_{f*g}$.

Proof. Exercise (Fubini, substitution) □

Lemma 3.5.7. *For* $g \in \mathcal{S}(\mathbb{R}^d)$, *we have* $\delta_0 * g = L_g$.

Hence, $\delta_0 * g = g$ in $\mathcal{S}'(\mathbb{R}^d)$.

Proof. For $\eta \in \mathcal{S}(\mathbb{R}^d)$, we compute

$$\delta_0 * g(\eta) = \delta_0(g^- * \eta) = \int_{\mathbb{R}^d} g(-y)\eta(0-y)\mathrm{d}y = \int_{\mathbb{R}^d} g(y)\eta(y)\mathrm{d}y = L_g(\eta). \qquad \square$$

Definition 3.5.8. *For* $L \in \mathcal{S}'(\mathbb{R}^d)$ *and* $\alpha \in \mathbb{N}^d$, *we define* $\partial^\alpha L \in \mathcal{S}'(\mathbb{R}^d)$ *by*

$$\partial^\alpha L(\eta) := (-1)^{|\alpha|} L(\partial^\alpha \eta), \qquad \eta \in \mathcal{S}(\mathbb{R}^d).$$

Lemma 3.5.9. *For* $f \in \mathcal{S}(\mathbb{R}^d)$, *we have* $L_{\partial^\alpha f} = \partial^\alpha L_f$.

Proof. For $d = \alpha = 1$, integration by parts yields

$$L_{f'}(\eta) = \int_{-\infty}^{\infty} f'(x)\eta(x)\mathrm{d}x = \underbrace{[f(x)\eta(x)]_{-\infty}^{\infty}}_{0} - \int_{-\infty}^{\infty} f(x)g'(x)\mathrm{d}x.$$

By iteration, we obtain the general formula. □

Example 3.5.10. *We have* $\partial^1 1_{[0,\infty)} = \delta_0$ *since*

$$\partial^1 1_{[0,\infty)}(\eta) = -\int_{\mathbb{R}} 1_{[0,\infty)}(x)\partial^1\eta(x)\mathrm{d}x = -\int_0^\infty \partial^1\eta(x)\mathrm{d}x = \eta(0) = \delta_0(\eta).$$

Definition 3.5.11. *For $L \in \mathcal{S}'(\mathbb{R}^d)$, we define $\hat{L}, \check{L} \in \mathcal{S}'(\mathbb{R}^d)$ by*

$$\hat{L}(\eta) := L(\hat{\eta}), \qquad \check{L}(\eta) := L(\check{\eta}) \qquad \eta \in \mathcal{S}(\mathbb{R}^d).$$

Lemma 3.5.12. *For $f \in \mathcal{S}(\mathbb{R}^d)$, we have $\widehat{L_f} = L_{\hat{f}}$ and $\check{L}_f = L_{\check{f}}$.*

Proof. Fubini leads to

$$\begin{aligned}\widehat{L_f}(\eta) = L_f(\hat{\eta}) &= \int_{\mathbb{R}^d} f(x)\hat{\eta}(x)\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f(x) \int_{\mathbb{R}^d} \eta(y)\mathrm{e}^{-2\pi i\langle y,x\rangle}\mathrm{d}y\mathrm{d}x \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x)\mathrm{e}^{-2\pi i\langle y,x\rangle}\mathrm{d}x\eta(y)\mathrm{d}y \\ &= \int_{\mathbb{R}^d} \hat{f}(y)\eta(y)\mathrm{d}y = L_{\hat{f}}.\end{aligned}$$

The proof for $\check{}$ is analogous. □

Example 3.5.13. *We have $\hat{1} = \delta_0$ since, for $\eta \in \mathcal{S}(\mathbb{R}^d)$,*

$$\hat{1}(\eta) = \int_{\mathbb{R}^d} 1 \cdot \hat{\eta}(x)\mathrm{d}x = \int_{\mathbb{R}^d} \hat{\eta}(x)\mathrm{e}^{2\pi\mathrm{i}\langle 0,x\rangle}\mathrm{d}x = \check{\hat{\eta}}(0) = \eta(0) = \delta_0(\eta).$$

Proposition 3.5.14 (Convolution relation distributions)**.** *For $L \in \mathcal{S}'(\mathbb{R}^d)$ and $g \in \mathcal{S}(\mathbb{R}^d)$, we have*

$$\widehat{L * g} = \hat{L} \cdot \hat{g}.$$

Proof. For $\eta \in \mathcal{S}(\mathbb{R}^d)$, we compute

$$\begin{aligned}\widehat{L * g}(\eta) &= L * g(\hat{\eta}) = L(g^- * \hat{\eta}) \\ &= \hat{L}(\mathcal{F}^{-1}(g^- * \hat{\eta})) \\ &= \hat{L}(\hat{g} \cdot \eta) = (\hat{L} \cdot \hat{g})(\eta).\end{aligned}$$

□

Proposition 3.5.15 (Fourier transform for tempered distributions)**.** *The Fourier transform $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$, $L \mapsto \hat{L}$ is a topological isomorphism with inverse $\mathcal{F}^{-1} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$, $L \mapsto \check{L}$.*

Proof. [1]. □

Example 3.5.16 (The Sha distribution)**.** *The Sha distribution, also known as the Dirac comb, is denoted by Sha.*[5] *It is important in Fourier analysis because it relates Fourier*

[5]This letter was chosen because it resembles the way people visualize the function, a long series of vertical spikes. The function is called the *Dirac comb* for the same reason.

series and Fourier transforms. It connects sampling and periodization. It is its own Fourier transform and, with a few qualifiers discussed later, the only such function.

The Ш *distribution is defined as*

$$Ш(x) = \sum_{n=-\infty}^{\infty} \delta(x-n) = \sum_{n=-\infty}^{\infty} T_n\delta(x),$$

where $\delta(x-n)$ is the Dirac delta distribution centered at n. The action of $\delta(x-n)$ on a test function is to evaluate that function at n. You can envision Ш *as an infinite sequence of spikes, one at each integer. The action of* Ш *on a test function is to add up its values at every integer.*

The product of Ш *with a function f is a new distribution whose action on a test function φ is the sum of $f\varphi$ over all integers. Alternatively, you could think of the distribution as a sort of clothesline on which to hang the sampled values of f, much like how a generating function works.*

Now consider a function f defined on $[0,1]$, i.e., zero everywhere outside the unit interval. The convolution of f with $\delta(x-n)$ is $f(x-n)$, i.e., a copy of f shifted to the interval $[n, n+1]$. By taking the convolution with Ш*, we create copies of f over the entire real line, effectively turning f into a periodic function. Instead of saying* the function f extended to create a periodic function*, you can simply write $f * Ш$.*

What is the Fourier transform of Ш*? The Fourier transform of $\delta(x)$ is 1, i.e., a constant function.*[6]

If you shift a function by n, you modulate its Fourier transform by $\exp(-2\pi i n\omega)$, hence:

$$\mathcal{F}(Ш)(\omega) = \widehat{Ш}(\omega) = \sum_{n=-\infty}^{\infty} \exp(-2\pi i n\omega).$$

This equation only makes sense in terms of distributions: the right-hand side does not converge in any classical sense. However, it turns out that the right-hand side is also Ш*!*

To show that the exponential sum equals Ш*, i.e., that* Ш *is its own Fourier transform, we resort to the definition of the Fourier transform of a distribution As a distribution, $\exp(-2\pi i n\omega)$ acts on a test function φ by integrating against it. This results in the Fourier transform of φ evaluated at n. Thus, the Fourier transform of* Ш *acts on φ by summing the values of φ's Fourier transform over all integers. By the Poisson summation formula, see Section 4.2.1, this is equivalent to summing the values of φ itself over all integers, which is the action of* Ш*. Hence,* Ш *is its own Fourier transform.*

The Ш *distribution is essentially unique. Any tempered distribution with period 1 that equals its own Fourier transform must be a multiple of* Ш*.*

Take a look at `https://dspillustrations.com/pages/posts/misc/the-dirac-comb-and-its-html`

[6] Remember: the more concentrated a function is, the more spread out its Fourier transform.

Chapter 4

Sampling

4.1 How does the Music end up on a CD? Sampling and Filtering

In the previous chapter we introduced continuous and discrete-time signals in a somewhat unrelated manner. This chapter deals with the core of modern digital signal processing: the idea and the basic theory of sampled signals. The principal idea is the following: which conditions of a continuous signal guarantee perfect reconstructions from discrete signal samples?
In order to understand the principal idea, let us first look at what happens to the Fourier transform, if we sample a signal as to obtain f_d from f. In Figure 4.1, you see the plot of the excerpt of a (pseudo-)continuous signal (a piano sound), with its Fourier transform. In the lower plots, a rather coarsely sampled signal (Sampling rate 11025 samples per second) and its Fourier transform are shown. It should be immediately obvious, what happens to the Fourier transform, if we sample f: the Fourier transform $\hat{f}$ of f is periodized!
So, the answer to the next question, namely, how to obtain the original signal from the sampled version, should be really easy: since the sampling process leads to repeated copies of the (hopefully bandlimited) spectrum, all we need to do is multiply with a lowpass filter in order to get rid of the unwanted copies: $\hat{f} = \hat{f}_d \cdot \Pi$, hence $f = f_d * \Pi$. Here we are intentionally sloppy and don't specify any of the involved parameters, since we only want to get across the basic idea - and this seems almost perfect!
However, if we look a bit closer at the spectrum of f, namely, if we apply a logarithmic scale (which actually corresponds to our perception of audio), we can see, that the spectrum of f has not actually dropped to anything close to 0, see Figure 4.2 so, what will happen to the frequencies above the cut-off? In fact, if we don't suppress them by highpass-filtering before the sampling process, those samples will show up as - usually unwanted - aliases in the lower frequency bands.

Aliasing is an effect which is one of the limitations of discrete-time sampling. An example of aliasing can be seen in old movies, e.g. when watching wagon but also car wheels: the wheels appear to go in reverse. This phenomenon can be observed if the

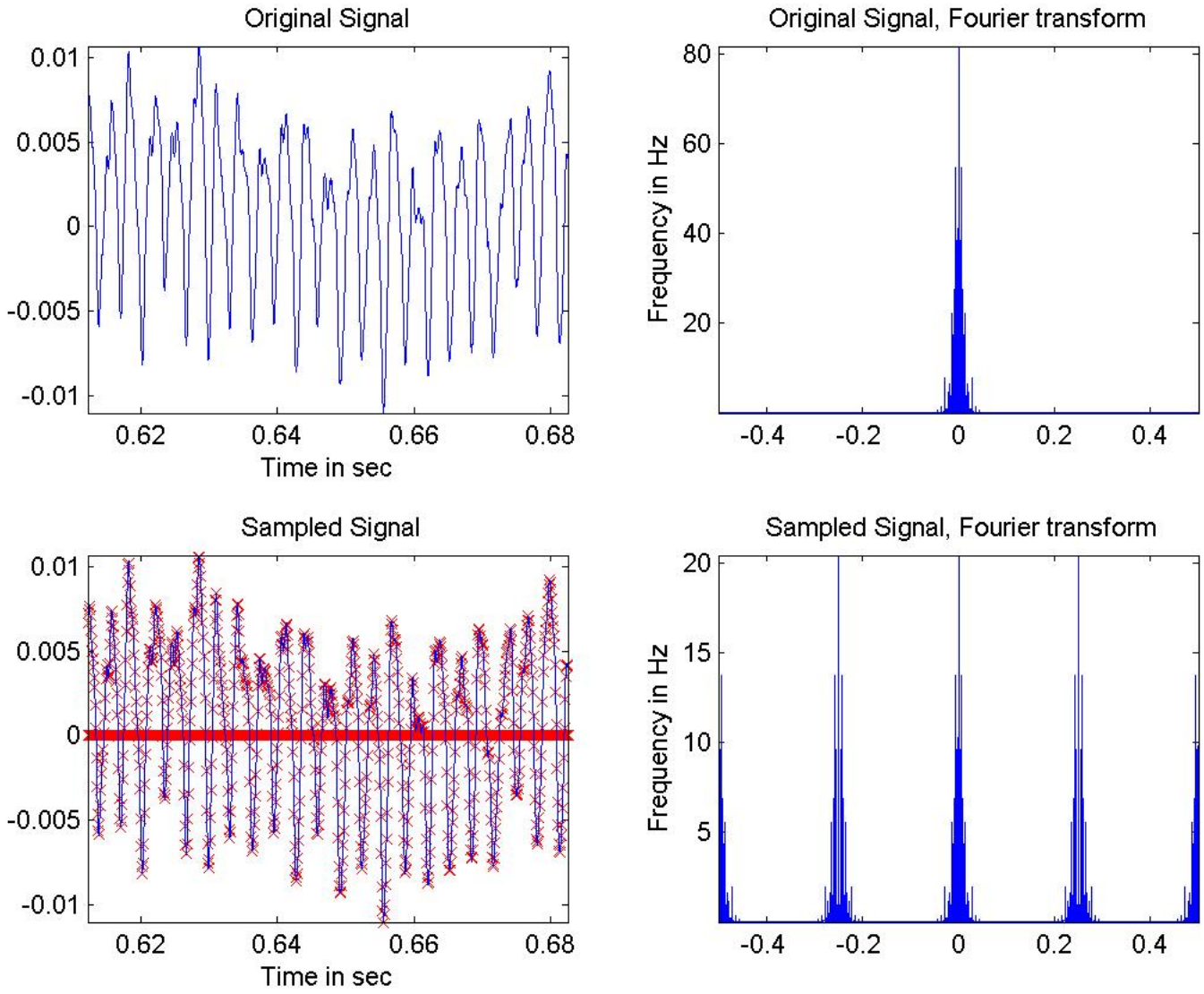


Figure 4.1: Subsampling and resulting Fourier transform

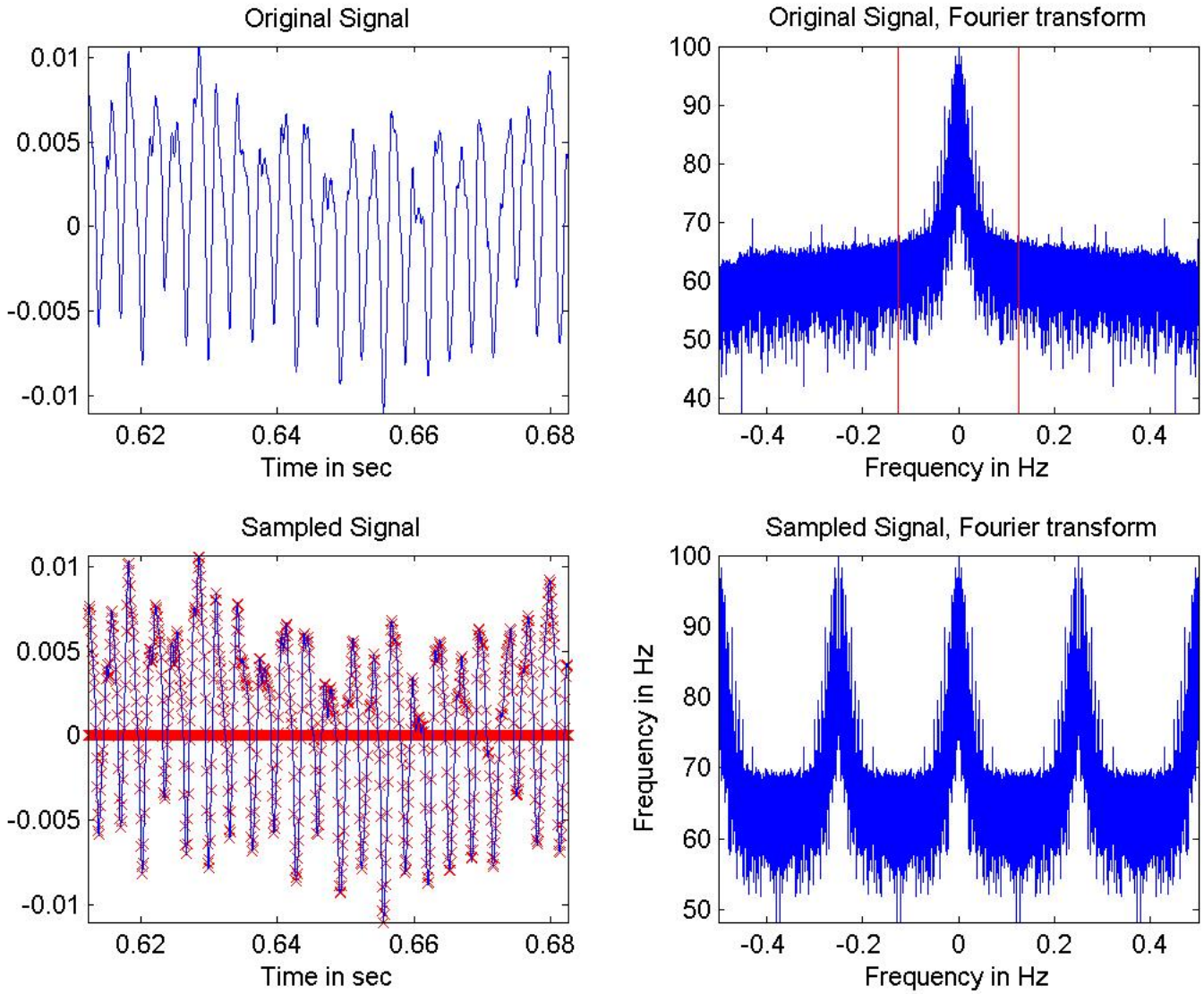


Figure 4.2: Subsampling and resulting Fourier transform

rate of the wagon wheel's spokes spinning approaches the rate of the sampler (the camera operating at about 30 frames per second) [1].

The same thing happens in data acquisition between the sampler and the signal we are sampling. For an example, have a look at Figure 4.3. Here, the effect of undersampling is immediately obvious: the sinusoid of 330Hz appears as a sinudoid with much lower frequency, namely 30Hz. In the lower plot, the wrong sampling rate of 320Hz maps the frequency 330Hz to $10Hz$. We will now study a simple case of this phenomenon mathematically.

Example 4.1.1. *Consider a complex exponential (a phasor) with frequency ω_0, i.e. $\varphi(t) = e^{2\pi i\omega_0 t}$. Now, assume that we sub-sample this phasor to obtain*

$$\varphi_d(n) = e^{2\pi i\omega_0(nT)},$$

i.e., T is the sampling interval. We have seen many times by now, that adding $2\pi ink$ to exponent doesn't change this (discrete) function:

$$\varphi_d(n) = e^{2\pi i\omega_0(nT)+2\pi ink} = e^{2\pi iTn(\omega_0+k/T)}, \text{ for all } k \in \mathbb{Z}.$$

This equation tells us that, after sampling, a sinusoid with frequency ω_0 cannot be distinguished from a sinusoid with frequency $\omega_0 + k/T$, $k \in \mathbb{Z}$. Note that $F_s = 1/T$ is the sampling rate.
If we sample real-valued signals, however, we always have to consider positive and negative frequencies, so, to a real sinusoid (sine or cosine) with frequency ω_0, we have in fact aliases at $\pm + \omega_0 k/T$.

Example 4.1.2. *Recall now the Fourier series of a square wave as defined in Example 2.1.4:*

$$f(x) = \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{1}{(2k-1)}\sin(2\pi(2k-1)x)$$

Obviously, this periodic function does NOT have a finite number of frequencies in it: its spectrum, i.e. the frequencies contained in the square wave decay like $1/n$ - and this really slow! We note that, due to this infinite bandwidth, the square-wave cannot be sampled properly: sampling, no matter how densely must always lead to aliasing, as we shall see next.
We assume a sampling rate of $F_s = 44100Hz$ and consider a square wave with fundamental frequency $F = 700Hz$. Then, the Fourier series of this function is simply

$$f(x) = \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{1}{(2k-1)}\sin(2\pi * 700 * (2k-1)x)$$

[1] This effect is even called "wagon-wheel effect".

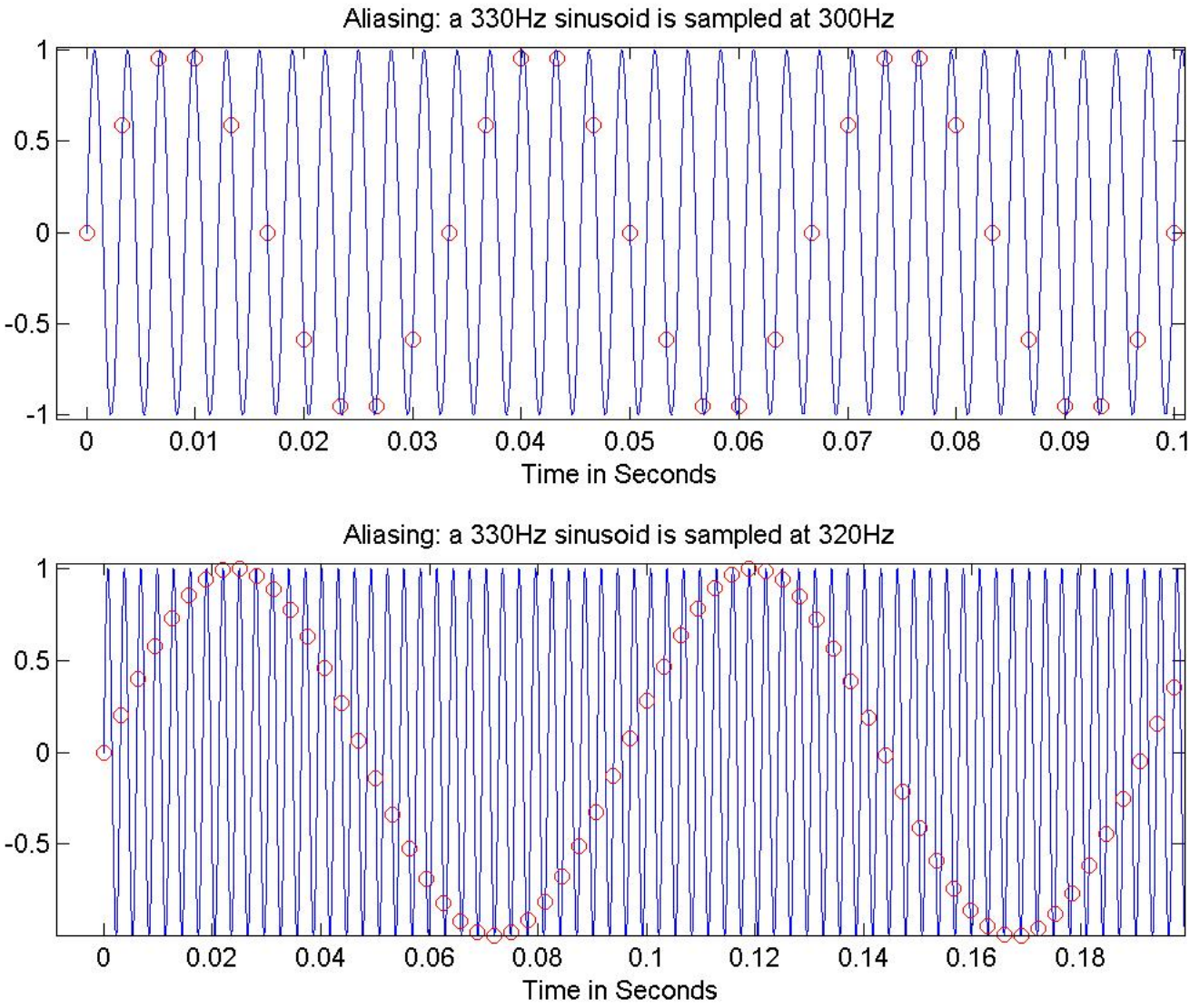


Figure 4.3: Aliasing by subsampling

since now we have 700 *oscillations per second. In this case, the highest frequency that is still below the Nyquist frequency of* $22050Hz$ *is the* $31st$ *harmonic which belongs to the frequency* $31 \cdot 700 = 21700$. *The next frequency contained in the signal, with index* 33 *is* $23100Hz$ *and is above Nyquist. It will therefore show up as an alias at* $(23100 - 44100)Hz = -21000Hz$. *This effect continues for all higher frequencies, and of course, all the negative frequencies turn into positive aliases accordingly. The phenomenon is shown on Figure 4.4.*

Note that in the current case, the fundamental frequency divides the Sampling rate, and the aliases become quasi-harmonics. In contrast, if we choose $F = 800Hz$, *the aliases will occur in frequencies that are not related not the fundamental frequencies, see Figure 4.5. While aliases should be avoided in usual sampling procedure, intentional aliasing can lead to interesting sound effects.*

We now turn to a more technical approach to sampling.

4.2 Formal Sampling

4.2.1 Poisson summation formula

We start by connecting the Fourier transform with Fourier coefficients of a periodic function-

Lemma 4.2.1. *If* $f \in L^1(\mathbb{T}^d)$ *and* $\left(\hat{f}_k\right)_{k\in\mathbb{Z}^d} \in \ell^1(\mathbb{Z}^d)$, *then* $f \in \mathcal{C}(\mathbb{T}^d)$ *with*

$$f = \sum_{k\in\mathbb{Z}^d} \hat{f}_k \mathrm{e}^{2\pi \mathrm{i}\langle k,\cdot\rangle} \tag{4.1}$$

converging uniformly and in $L^1(\mathbb{T}^d)$ *and* $L^2(\mathbb{T}^d)$.

Proof. The condition $\left(\hat{f}_k\right)_{k\in\mathbb{Z}^d} \in \ell^1(\mathbb{Z}^d)$ implies uniform convergence, so that

$$g := \sum_{k\in\mathbb{Z}^d} \hat{f}_k \mathrm{e}^{2\pi \mathrm{i}\langle k,\cdot\rangle} \in \mathcal{C}(\mathbb{T}^d) \subset L^2(\mathbb{T}^d) \subset L^1(\mathbb{T}^d).$$

Since $\ell^1(\mathbb{Z}^d) \subset \ell^2(\mathbb{Z}^d)$, we have $f \in L^2(\mathbb{T}^d)$. Obviously, $\hat{g}_k = \hat{f}_k$, so that we deduce $g = f$ in $L^2(\mathbb{T}^d)$, hence, $f = g$ almost everywhere. □

Proposition 4.2.2. *If* $f \in L^1(\mathbb{R}^d)$, *then*

$$\overline{\omega} f := \sum_{k\in\mathbb{Z}^d} f(\cdot + k)$$

converges pointwise almost everywhere and $\overline{\omega} f \in L^1(\mathbb{T}^d)$ *with*

$$\|\overline{\omega} f\|_{L^1(\mathbb{T}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}.$$

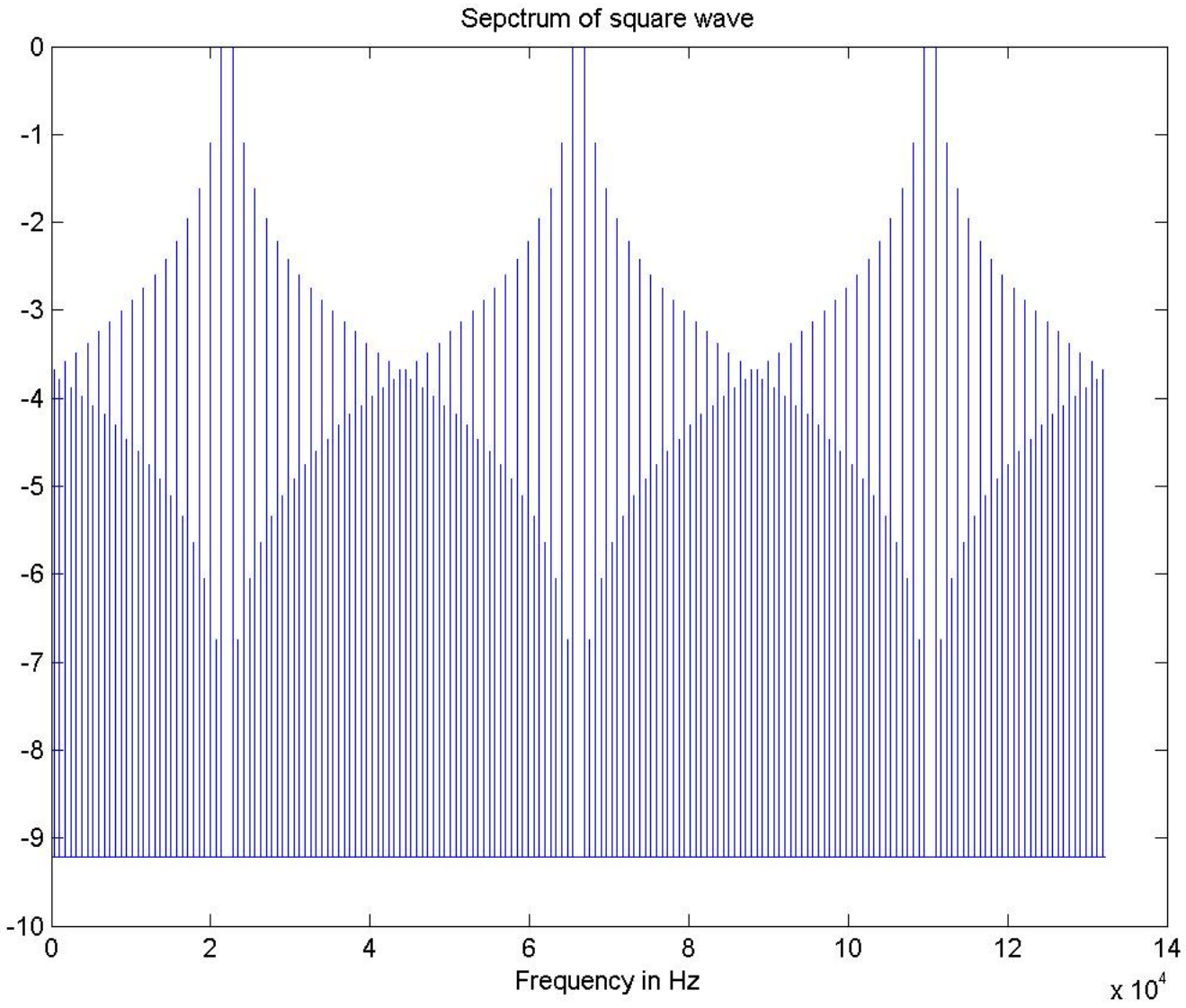


Figure 4.4: Aliasing by discretization of Square Wave

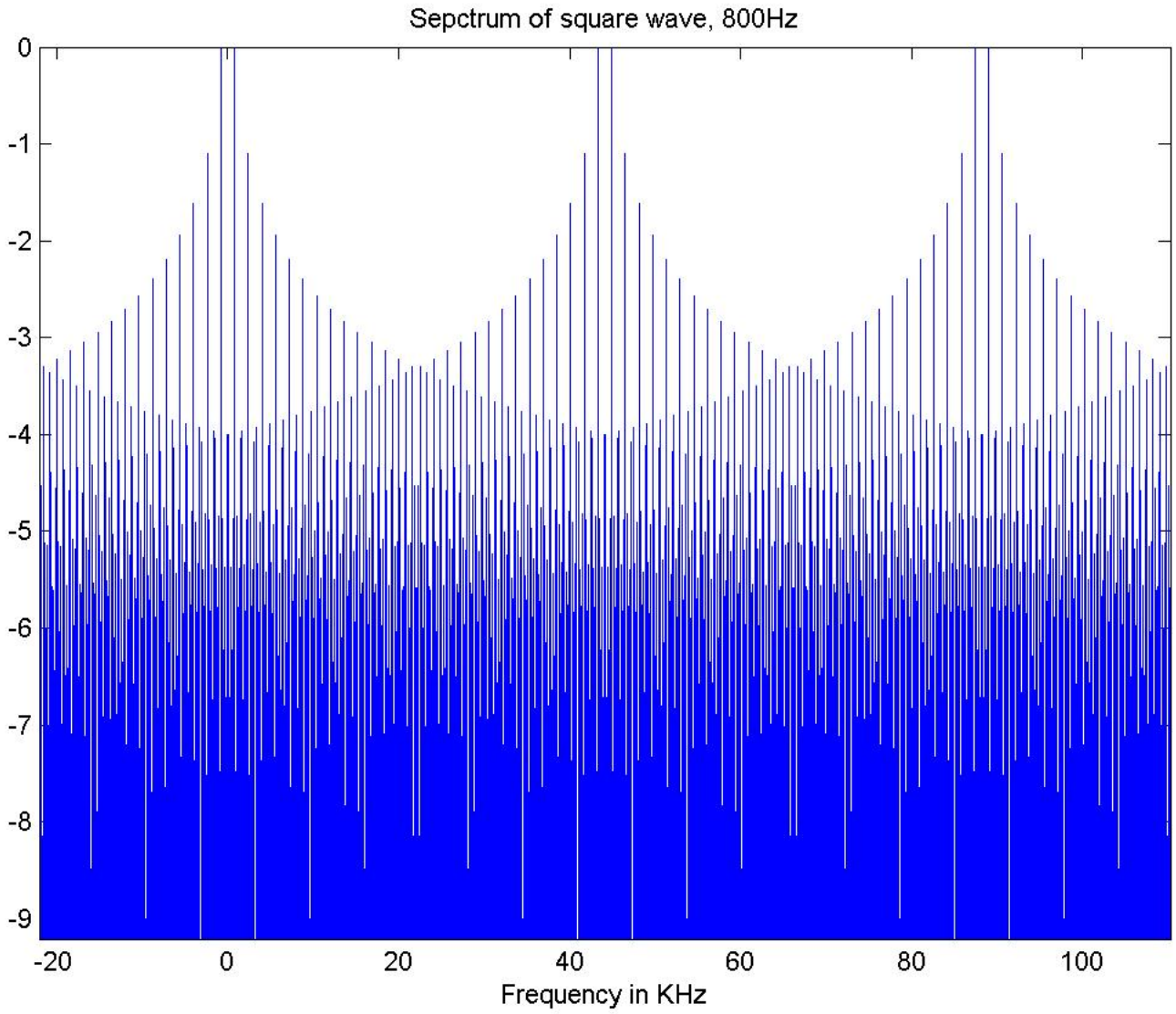


Figure 4.5: Aliasing by discretization of Square Wave

Proof. According to Theorem A.1.2 and the monoton convergence theorem,

$$\begin{aligned}\|f\|_{L^1(\mathbb{R}^d)} &= \sum_{k\in\mathbb{Z}^d} \int_{I^d+k} |f| \\ &= \sum_{k\in\mathbb{Z}^d} \int_{I^d} |f(\cdot+k)| \\ &= \int_{I^d} \sum_{k\in\mathbb{Z}^d} |f(\cdot+k)|,\end{aligned}$$

where $\sum_{k\in\mathbb{Z}^d} |f(\cdot+k)|$ converges almost everywhere (and in $L^1(\mathbb{R}^d)$. Hence, $\sum_{k\in\mathbb{Z}^d} f(\cdot+k)$ converges almost everywhere with upper bound $\overline{\omega}|f|$ and $\overline{\omega}f$ is $\mathbb{Z}^d$-periodic, and we have

$$\|\overline{\omega}f\|_{L^1(\mathbb{T}^d)} \le \|\overline{\omega}|f|\|_{L^1(\mathbb{T}^d)} = \|f\|_{L^1(\mathbb{R}^d)}.$$ □

For $f \in L^1(\mathbb{R}^d)$, how do $\hat{f}(k)$ and $(\widehat{\overline{\omega}f})_k$ relate to each other?

Theorem 4.2.3 (Poisson formula)**.** *Let $f \in L^1(\mathbb{R}^d)$.*

a) For $k \in \mathbb{Z}^d$,

$$(\widehat{\overline{\omega}f})_k = \hat{f}(k), \qquad k \in \mathbb{Z}^d\,.$$

b) If $\left((\widehat{\overline{\omega}f})_k\right)_{k\in\mathbb{Z}^d} \in \ell^1(\mathbb{Z}^d)$ and $\overline{\omega}f$ is continuous, then

$$(\overline{\omega}f)(x) = \sum_{k\in\mathbb{Z}^d} \hat{f}(k)\mathrm{e}^{2\pi\mathrm{i}\langle k,x\rangle}, \qquad x \in \mathbb{T}^d.$$

c) In particular, for $x = 0$, and $d = 1$ we obtain

$$\sum_{k\in\mathbb{Z}} f(k) = \sum_{k\in\mathbb{Z}} \hat{f}(k).$$

Proof. By Proposition 4.2.2 and Theorem A.1.2, we have

$$\begin{aligned}(\widehat{\overline{\omega}f})_k &= \int_{I^d} (\overline{\omega}f)(x)\ \mathrm{e}^{-2\pi\mathrm{i}\langle k,x\rangle}\mathrm{d}x \\ &= \sum_{l\in\mathbb{Z}^d} \int_{I^d} f(x+l)\mathrm{e}^{-2\pi\mathrm{i}\langle k,x\rangle}\mathrm{d}x \\ &= \sum_{l\in\mathbb{Z}^d} \int_{I^d+l} f(x)\ \mathrm{e}^{-2\pi\mathrm{i}\langle k,x\rangle}\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f(x)\mathrm{e}^{-2\pi\mathrm{i}\langle k,x\rangle}\mathrm{d}x = \hat{f}(k).\end{aligned}$$

Part b) follows from (4.1) since both sides are continuous. □

Figure 4.6: $\mathrm{sinc}(\xi) = \frac{\sin(\pi\xi)}{\pi\xi}$, for $|x| \leq 10$.

Corollary 4.2.4. *For $f \in \mathcal{S}(\mathbb{R}^d)$, we have*

$$(\overline{\omega} f)(x) = \sum_{k \in \mathbb{Z}^d} \hat{f}(k) \mathrm{e}^{2\pi \mathrm{i}\langle k, x\rangle}, \qquad x \in \mathbb{T}^d.$$

Proof. Since $\overline{\omega} f$ converges uniformly, Theorem 4.2.3 implies the claim. □

4.2.2 The Shannon Sampling Theorem

We define

$$\mathrm{sinc}(\xi) := \frac{\sin(\pi\xi)}{\pi\xi}, \quad \xi \in \mathbb{R},$$

see Figure A.2.

Lemma 4.2.5 (Fourier transform Box)**.** *For $d = 1, 2, \ldots$, we have*

$$\mathcal{F}^{\pm 1}\left(1_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d}\right)(\xi) = \prod_{i=1}^{d} \mathrm{sinc}(\xi_i), \qquad \xi \in \mathbb{R}^d. \tag{4.2}$$

Proof. For $d = 1$, we compute

$$\begin{aligned}
\mathcal{F}\left(1_{\left[-\frac{1}{2}, \frac{1}{2}\right]}\right)(\xi) &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{e}^{-2\pi \mathrm{i} x \xi} \mathrm{d}x \\
&= \frac{1}{-2\pi \mathrm{i}\xi} \left[\mathrm{e}^{-2\pi \mathrm{i} x \xi}\right]_{-\frac{1}{2}}^{\frac{1}{2}} \\
&= \frac{\mathrm{e}^{-\pi \mathrm{i}\xi} - \mathrm{e}^{\pi \mathrm{i}\xi}}{-2\pi \mathrm{i}\xi} \\
&= \frac{\cos(-\pi\xi) + \mathrm{i}\sin(-\pi\xi) - \cos(\pi\xi) - \mathrm{i}\sin(\pi\xi)}{-2\pi\xi} \\
&= \frac{\sin(\pi\xi)}{\pi\xi}.
\end{aligned}$$

For $d \geq 2$, we observe

$$\begin{aligned}
1_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d}(x) &= 1_{\left[-\frac{1}{2}, \frac{1}{2}\right]}(x_1) \cdots 1_{\left[-\frac{1}{2}, \frac{1}{2}\right]}(x_d), \\
\mathrm{e}^{-2\pi \mathrm{i}\langle x, \xi\rangle} &= \mathrm{e}^{-2\pi \mathrm{i} x_1 \xi_1} \cdots \mathrm{e}^{-2\pi \mathrm{i} x_d \xi_d},
\end{aligned}$$

so that we derive

$$\mathcal{F}\left(1_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d}\right)(\xi) = \int_{\mathbb{R}} 1_{\left[-\frac{1}{2}, \frac{1}{2}\right]}(x_1) \mathrm{e}^{-2\pi \mathrm{i} x_1 \xi_1} \mathrm{d}x_1 \cdots \int_{\mathbb{R}} 1_{\left[-\frac{1}{2}, \frac{1}{2}\right]}(x_d) \mathrm{e}^{-2\pi \mathrm{i} x_d \xi_d} \mathrm{d}x_d.$$

Since $\mathrm{sinc}(-\xi_i) = \mathrm{sinc}(\xi_i)$ and $\mathcal{F}^{-1} f(\xi) = \mathcal{F} f(-\xi)$, we conclude the proof. □

Definition 4.2.6. *For $t > 0$, the Paley-Wiener space* $\mathrm{PW}(t)$ *is*

$$\mathrm{PW}(t) = \left\{ f \in L^2(\mathbb{R}^d) : \operatorname{supp}(\hat{f}) \subset [-t,t]^d \right\}.$$

Theorem 4.2.7 (Shannon's sampling theorem)**.** *Let $\varphi(\xi) := \prod_{i=1}^d \operatorname{sinc}(\xi_i)$. If $f \in \mathrm{PW}\left(\frac{1}{2}\right)$, then*

$$f = \sum_{k\in\mathbb{Z}^d} f(k)\varphi(\cdot - k) \tag{4.3}$$

holds in $L^2(\mathbb{R}^d)$ and uniformly in $\mathbb{R}^d$.

Proof. Since $f \in \mathrm{PW}\left(\frac{1}{2}\right)$ implies $\hat{f} \in L_1(\mathbb{R}^d)$, Poisson's formula yields

$$f(-k) = (\mathcal{F}\hat{f})(k) = (\widehat{\overline{\omega}\hat{f}})_k. \tag{4.4}$$

Due to $\operatorname{supp}(\hat{f}) \subset I^d$, we have

$$1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d}\, \overline{\omega}\hat{f} = \hat{f},$$

so that $\overline{\omega}\hat{f} \in L^2(\mathbb{T}^d)$ because

$$\|\overline{\omega}\hat{f}\|^2_{L^2(\mathbb{T}^d)} = \int_{I^d} \left|(\overline{\omega}\hat{f})(x)\right|^2 \mathrm{d}x = \int_{I^d} \left|\hat{f}(x)\right|^2 \mathrm{d}x = \|f\|^2_{L^2(\mathbb{R}^d)} < \infty.$$

Thus, (4.4) implies $(f(k))_{k\in\mathbb{Z}^d} \in \ell^2(\mathbb{Z}^d)$. We further derive in $L^2(\mathbb{R}^d)$

$$\begin{aligned}
\hat{f} = 1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d}\, \overline{\omega}\hat{f} &= 1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d} \sum_{k\in\mathbb{Z}^d} f(-k)\mathrm{e}_k \\
&= \sum_{k\in\mathbb{Z}^d} f(k) 1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d}\, \mathrm{e}_{-k} \\
&= \sum_{k\in\mathbb{Z}^d} f(k)\mathcal{F}\left(\varphi(\cdot - k)\right),
\end{aligned}$$

where we have used Lemma 4.2.5 for $\mathcal{F}^{-1}$. Applying $\mathcal{F}^{-1}$ to both sides implies (4.3) in $L^2(\mathbb{R}^d)$.

To verify uniform convergence, recall $|\varphi|^2 \in L^1(\mathbb{R}^d)$ with

$$\left(\widehat{\overline{\omega}|\varphi|^2}\right)_k = \mathcal{F}\left(|\varphi|^2\right)(k), \qquad k \in \mathbb{Z}^d. \tag{4.5}$$

Since $\mathcal{F}\left(|\varphi|^2\right) = \mathcal{F}\left(\varphi \cdot \overline{\varphi}\right) = \widehat{\varphi} * \widehat{\overline{\varphi}} = \widehat{\varphi} * \overline{\widehat{\varphi}(-\cdot)}$, we obtain

$$\operatorname{supp}\mathcal{F}\left(|\varphi|^2\right) \subset [-1,1]^d.$$

The function $\mathcal{F}(|\varphi|^2)$ is continuous, so that (4.5) yields $\left(\widehat{\overline{\omega}|\varphi|^2}\right)_k = 0$ for all $0 \neq k \in \mathbb{Z}^d$. We deduce $\overline{\omega}|\varphi|^2 = c \in \mathbb{C}$ is constant. Cauchy-Schwartz leads to

$$\left| \sum_{\|k\| \geq m} f(k)\varphi(\cdot - k) \right| \leq \left(\sum_{\|k\| \geq m} |f(k)|^2 \right)^{1/2} c \quad \xrightarrow{m \to \infty} \quad 0$$

uniformly since $(f(k))_{k \in \mathbb{Z}^d} \in \ell^2(\mathbb{Z}^d)$. □

We now allow for $f \in \mathrm{PW}(t)$.

Corollary 4.2.8. *Suppose that* $0 < a \leq \frac{1}{2t}$ *and* $\varphi \in \mathrm{PW}\left(\frac{1}{2}\right)$ *satisfies*

$$\widehat{\varphi}(\xi) = 1, \quad \xi \in [-at, at]^d. \tag{4.6}$$

If $f \in \mathrm{PW}(t)$, *then*

$$f = \sum_{k \in \mathbb{Z}^d} f(ka)\varphi\left(\frac{\cdot}{a} - k\right)$$

holds in $L^2(\mathbb{R}^d)$ *and uniformly in* $\mathbb{R}^d$.

Note that φ can be chosen as in (4.2), but it is not necessary if $0 < a < \frac{1}{2t}$. The number $\frac{1}{2t}$ is called *Shannon's sampling rate* or Nyquist rate.

Proof. Put $g = f(\cdot a)$, so that $g \in \mathrm{PW}(at) \subset \mathrm{PW}\left(\frac{1}{2}\right)$. EXERCISE: Go through the proof of Theorem 4.2.7 and observe that the condition (4.6) with $0 < a \leq \frac{1}{2t}$ is sufficient. □

4.2.3 An alternative view: sampling is periodization in the Fourier domain

In Section 3.4, we looked at the finite DFT of a Dirac comb and noted that it exhibits periodicity. Now, we will prove the continuous analog of this observation: if a continuous function is sampled, its Fourier transform is periodic.

Theorem 4.2.9 (Sampling is periodization in the Fourier domain)**.** *The Fourier transform of the discrete signal* f_d *obtained by sampling a continuous signal* f *at a sampling interval* T *is*

$$\hat{f}_d(\omega) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \hat{f}(\omega - \frac{k}{T}). \tag{4.7}$$

Proof. First note that, considered as a distribution, f_d has the form

$$f_d = f \cdot \text{Ш}_T.$$

Now we invoke the convolution theorem for the Fourier transform: if $h(t) = f(t) \cdot g(t)$, then $\hat{h}(\omega) = [\hat{f} * \hat{g}](\omega)$, hence

$$\hat{f}_d(\omega) = (\hat{f} * \widehat{\text{Ш}_T})(\omega) = \frac{1}{T}(\hat{f} * \text{Ш}_{\frac{1}{T}})(\omega) = \frac{1}{T}\sum_{k\in\mathbb{Z}} \hat{f}(\omega - \frac{k}{T})$$

□

Let us now recollect, what we have observed so far: if we are given a continuous signal f, and we sample it with a sampling rate of $1/T$ samples per second: $f_d[n] = f(Tn)$. Then, the corresponding Fourier transform $\hat{f}_d$ is a periodized version of the original $\hat{f}$, with period equal to the sampling rate, i.e. $1/T$. Now, under the assumption, that the original signal was effectively bandlimited to the interval $[-\frac{1}{2T}, \frac{1}{2T}]$ it is now quite obvious what needs to be done to recover the continuous waveform: we have to cut off all the unnecessary copies of the spectrum. The elimination of all frequencies above $|\frac{1}{2T}|$ leads to interpolation, in other words, to the reconstruction of the original function from its samples.

Theorem 4.2.10 (Shannon Sampling Theorem II)**.** *Assume that the Fourier transform of a continuous signal f is contained in the interval $[-\frac{1}{2T}, \frac{1}{2T}]$, then f can be perfectly reconstructed from its samples at nT, i.e. from $f_d = f \cdot \text{Ш}_T$ as follows*

$$f(t) = \sum_{n\in\mathbb{Z}} f(nT)sinc(\frac{t - nT}{T}) \tag{4.8}$$

Proof. Note that for $n \neq 0$ the support of $\hat{f}(\omega - \frac{n}{T})$ does not intersect with the support of $\hat{f}(w)$, since $\hat{f}(\omega) = 0$ for $|\omega| \geq \frac{1}{2T}$. Therefore

$$\hat{f}_d(\omega) = \frac{1}{T}\hat{f}(\omega) \text{ for } |\omega| \leq \frac{1}{2T}$$

We then have $\hat{f}(\omega) = [T \cdot \Pi_T \cdot \hat{f}_d](\omega)$ and thus

$$f(t) = \mathcal{F}^{-1}(T \cdot \Pi_T \cdot \hat{f}_d)(t)$$

and $\mathcal{F}^{-1}\Pi_T = \widehat{\Pi_T}$, since Π_T is symmetric, hence

$$\begin{aligned} f(t) =& (T\widehat{\Pi_T} * f_d)(t) = T\widehat{\Pi_T} * \sum_{n\in\mathbb{Z}} f(nT)\delta(t - nT) \\ =& \sum_{n\in\mathbb{Z}} f(nT)T\widehat{\Pi_T}(t - nT) \end{aligned}$$

Recall that $\hat{\Pi}(x) = sinc(x)$ and, since $\widehat{\Pi_T}(x) = \widehat{D_{\frac{1}{T}}\Pi}(x) = \frac{1}{T}D_T\hat{\Pi}(x)$, we find

$$f(t) = \sum_{n\in\mathbb{Z}} f(nT)sinc(\frac{t}{T} - n).$$

□

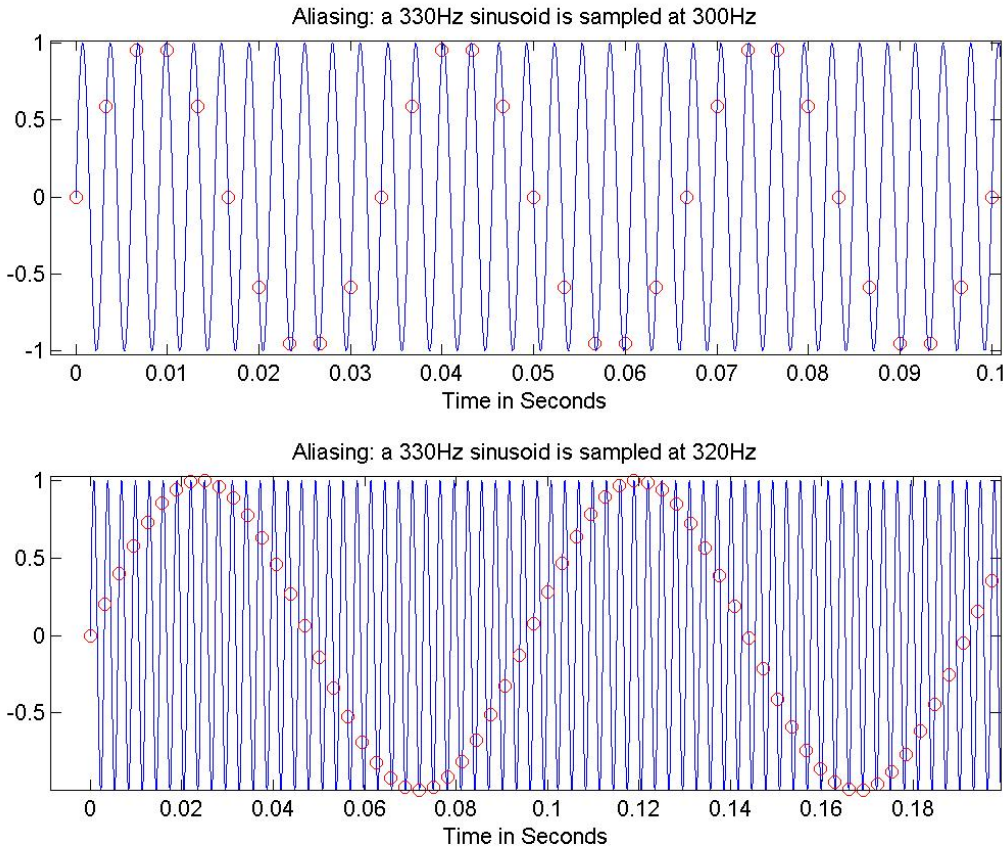


Figure 4.7: Sampling too sparsely.

Undesired aliases can be suppressed by applying prefiltering to analog signals before sampling them.[2]

4.2.4 Aliasing

If we do not sample sufficiently dense, then we cannot expect to reconstruct the correct function, see Figure 4.7.

If $a > \frac{1}{2t}$, then *aliasing* occurs, i.e., higher frequency components are falsely taken as lower frequency contributions. More specifically, write $f \in \mathrm{PW}(t)$ as the finite sum

$$f = \sum_m f_m, \qquad \text{with} \qquad \mathrm{supp}(\hat{f}_m) \subset \left[-\frac{1}{2}, \frac{1}{2}\right]^d + m, \qquad m \in \mathbb{Z}^d,$$

[2]Anti-alias Filter: The pre-filtering of an analog signal, before it is digitized or sampled, to remove or substantially attenuate the undesired aliasing components, i.e. those components that have higher frequency than half the sampling rate.

which is derived from $f_m := \mathcal{F}^{-1}\left(\hat{f} \cdot 1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d+m}\right)$. For the sampling rate $a = 1 > \frac{1}{2t}$, we consider

$$F(x) := \sum_{k\in\mathbb{Z}^d} f(k) sinc(x-k).$$

The periodization of $\hat{f}$ satisfies

$$\overline{\omega}\hat{f} = \sum_{l\in\mathbb{Z}^d}\sum_m \hat{f}_m(\cdot - l) = \sum_m \hat{f}_m(\cdot + m).$$

Therefore, we derive

$$\hat{F}(\xi) = 1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d}\,\overline{\omega}\hat{f} = 1_{\left[-\frac{1}{2},\frac{1}{2}\right]^d}\sum_m \hat{f}_m(\cdot + m),$$

so that F does not contain any frequencies beyond $\left[-\frac{1}{2},\frac{1}{2}\right]^d$. The Fourier reconstruction leads to

$$F(x) = \int_{\left[-\frac{1}{2},\frac{1}{2}\right]^d} \sum_m \hat{f}_m(\xi+m)\mathrm{e}^{2\pi\mathrm{i}\langle x,\xi\rangle}\mathrm{d}\xi.$$

The lower frequency components with $\xi \in \left[-\frac{1}{2},\frac{1}{2}\right]^d$ should be $\hat{f}_0(\xi)$. However, there are the additional contributions $\sum_{m\neq 0}\hat{f}_m(\xi+m)$ that are induced from higher frequencies. Further, we obtain

$$\begin{aligned} F(x) &= \sum_m \int_{\left[-\frac{1}{2},\frac{1}{2}\right]^d} \hat{f}_m(\xi+m)\mathrm{e}^{2\pi\mathrm{i}\langle x,\xi\rangle}\mathrm{d}\xi \\ &= \sum_m \int_{\left[-\frac{1}{2},\frac{1}{2}\right]^d+m} \hat{f}_m(\xi)\mathrm{e}^{2\pi\mathrm{i}\langle x,\xi\rangle}\mathrm{e}^{-2\pi\mathrm{i}\langle x,m\rangle}\mathrm{d}\xi \\ &= \sum_m f_m(x)\mathrm{e}^{-2\pi\mathrm{i}\langle x,m\rangle}. \end{aligned}$$

Thus, instead of $f = f_0 + \sum_{m\neq 0} f_m$, we obtain $F = f_0 + \sum_{m\neq 0} f_m(x)\mathrm{e}^{-2\pi\mathrm{i}\langle x,m\rangle}$.

Exercises: Understanding Aliasing and the Nyquist Rate

Exercise 1: Nyquist Rate and Bandwidth For each of the following signals, calculate the Nyquist rate:

1. $f_1(t) = \sin(2\pi \cdot 10t) + \sin(2\pi \cdot 20t)$
2. $f_2(t) = \cos(2\pi \cdot 50t)$
3. $f_3(t) = e^{-t^2}\sin(2\pi \cdot 5t)$

Suggest a sampling rate for each signal that avoids aliasing.

Exercise 2: Fourier Transform and Aliasing Given the signal $f(t) = \sin(2\pi \cdot 5t)$ sampled at rate $a = 8$ Hz:

1. Compute the Fourier Transform of the sampled signal.
2. Show how the higher frequency components fold back into the interval $[-\frac{1}{2}, \frac{1}{2}]$.

Exercise 3: Reconstruction and Error Analysis Consider the signal $f(t) = \cos(2\pi \cdot 3t)$ sampled at rate $a = 4$ Hz.

1. Reconstruct the signal using the given sampling points.
2. Compare the reconstructed signal to the original by plotting both on the same graph.
3. Quantify the error between the original and reconstructed signals.

Exercise 4: Simulating Aliasing Write a Python script to sample the continuous-time signal $f(t) = \sin(2\pi \cdot 10t)$ at various sampling rates $a = 5, 10, 15$ Hz. For each case:

1. Plot the original and reconstructed signals.
2. Observe and describe any aliasing effects.

Exercise 6: Frequency Domain Representation Using the following skeleton code, visualize the spectrum of the sampled signal $f(t) = \cos(2\pi \cdot 25t)$:

```
import numpy as np
import matplotlib.pyplot as plt

# Parameters
fs = 50  # Sampling frequency
T = 1/fs # Sampling period
t = np.arange(0, 1, T)

# Signal
t_signal = np.linspace(0, 1, 1000)
f = 25 # Frequency in Hz
signal = np.cos(2 * np.pi * f * t_signal)
sampled_signal = np.cos(2 * np.pi * f * t)

# Compute FFT
fft_sampled = np.fft.fft(sampled_signal)
freqs = np.fft.fftfreq(len(fft_sampled), T)

# Plot
```

```
plt.figure()
plt.plot(freqs, np.abs(fft_sampled))
plt.title("Frequency Spectrum")
plt.show()
```

Complete the code to:

- Plot the signal in both time and frequency domains.
- Identify aliasing in the frequency domain.
- Suggest ways to mitigate aliasing (e.g., filtering).

Chapter 5

More flexibel transformations

5.1 Introduction - Uncertainty principle and time-frequency molecules

So far, we have been looking at a signal as being described either in time or in frequency - with the Fourier transform as a means of transformation from one domain to the other. Many signals of interest, however - above all: music and speech - are not stationary over time, they are time-variant, and we will be more interested in finding the frequency-content at a particular point in time, rather than knowing which frequencies comprise the entire signal. With this question we enter the important realm of *time-frequency (TF) analysis* (TF-Analysis). The desire to have complete insight in the local time-frequency structure, however, is impaired by a central fact of both TF-analysis and - more famously - quantum mechanics, called *uncertainty principle*. Loosely speaking, it states that, the more concentrated $f(t)$ is, the wider its Fourier transform $\hat{f}(\omega)$ must be. In other words, the scaling property of the Fourier transform may be seen as saying: if we "squeeze" a function in t (time), its Fourier transform "stretches out" in ω (frequency). It is not possible to arbitrarily concentrate both a function and its Fourier transform.

Exercise 5.1.1 (Gaussian Window)**.** *The Fourier transform of the normalized Gaussian window $\varphi_0(t) = e^{-\pi t^2}$ is given by $\hat{\varphi}_0(\omega) = e^{-\pi\omega^2}$, in other words: the Gaussian is invariant under Fourier transform. Then, the dilated Gaussian $\varphi_a(t) = e^{-\pi t^2/a}$ has the Fourier transform $\hat{\varphi}_a(\omega) = \sqrt{a}e^{-\pi a\omega^2}$.*

We can regard a function either as represented as $f(t)$, which measures the correlation of f with $\delta(t)$, e.g., what 'happens at time t'. Alternatively, we may represent a function as its Fourier transform $\hat{f}(\omega)$, which measures the correlation of f with an exponential $e^{2\pi i\omega t}$, e.g., how much of the pure frequency ω is contained in the function.

Can we measure how much of a specific frequency is contained at a specific time? In other words, can we measure simultaneously time and frequency? To achieve this, we would represent a function by correlation of f with functions (or distributions) which are concentrated in both time and frequency. The Heisenberg uncertainty principle yields a

fundamental lower bound for such a concentration: time and frequency cannot be simultaneously measured up to arbitrary precision!

To make this precise, we define for $f \in L^2(\mathbb{R})$ its spread via

$$\sigma(f)^2 := \inf_{t_0 \in \mathbb{R}} \int_{\mathbb{R}} (t - t_0)^2 |f(t)|^2 \, dt \Big/ \|f\|^2.$$

Clearly, $\sigma(f)^2$ can be interpreted as the standard deviation of a random variable with density $\frac{|f(t)|^2}{\|f\|^2}$.

We are interested in the construction of a function f which is optimally concentrated in both time and frequency.

Our goal is the following theorem.

Theorem 5.1.2 (Heisenberg Uncertainty)**.** *For any nonzero $f \in L^2(\mathbb{R})$, it holds that*

$$\sigma(f) \cdot \sigma(\hat{f}) \geq \frac{1}{4\pi}.$$

Equality holds if and only if f is a translation and modulation of a Gaussian.

Before we prove this, we show an auxiliary result.

Definition 5.1.3. *Let H be a Hilbert space and A, B be (possibly unbounded) linear operators on H. Then we define the Poisson bracket*

$$[A, B] := AB - BA.$$

Lemma 5.1.4. *Suppose that A, B are self-adjoint operators on a Hilbert space H. Then for all $a, b \in \mathbb{R}$ and $f \in H$, it holds that*

$$\|(A - a)f\|_H \cdot \|(B - b)f\|_H \geq \frac{1}{2} \cdot |\langle [A, B]f, f \rangle_H|.$$

Equality holds if and only if

$$(A - a)f = ic(B - b)f$$

for some constant $c \in \mathbb{R}$.

Proof. Note that $[A, B] = [A - a, B - b]$. Using the self-adjointness of A and B, we write

$$\langle [A, B]f, f \rangle = \langle [A-a, B-b]f, f \rangle = \langle (B-b)f, (A-a)f \rangle - \langle (A-a)f, (B-b)f \rangle = 2i \, \Im \langle (B-b)f, (A-a)f \rangle.$$

It follows that

$$|\langle [A, B]f, f \rangle| \leq 2 \, |\langle (B - b)f, (A - a)f \rangle| \leq 2 \, \|(B - b)f\| \cdot \|(A - a)f\|,$$

by the Cauchy-Schwarz inequality. This proves the first part. The part about equality follows from the fact that it only holds if $\langle (B - b)f, (A - a)f \rangle$ is purely imaginary AND if $(B - b)f$ is a scalar multiple of $(A - a)f$. □

Now we are ready for the proof of Theorem 5.1.2. Without loss of generality we assume $\|f\|^2 = 1$ and

$$\sigma^2(f) = \int_{\mathbb{R}} t^2 |f(t)|^2 \, dt \quad \text{and} \quad \sigma^2(\hat{f}) = \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 \, d\omega.$$

Let $H = L^2(\mathbb{R})$. Put $Af(t) := t \cdot f(t)$ and $Bf(t) := \frac{1}{2\pi i} f'(t)$. Then

$$\sigma^2(f) \cdot \sigma^2(\hat{f}) = \|Af\|^2 \cdot \|Bf\|^2.$$

We may now use Lemma 5.1.4 to deduce that

$$\sigma(f) \cdot \sigma(\hat{f}) \geq \frac{1}{2} \cdot |\langle [A, B]f, f \rangle|.$$

So let's compute the Poisson bracket of A and B:

$$ABf(t) = \frac{1}{2\pi i} t \cdot f'(t), \quad BAf(t) = \frac{1}{2\pi i} (t \cdot f(t))',$$

which implies that $[A, B]f = -\frac{1}{2\pi i} f$ and hence the Heisenberg inequality. The case of equality follows from the corresponding statement in Lemma 5.1.4 by considering the differential equation resulting from

$$(B - b)f = ic(A - a)f$$

´namely

$$f' - b2\pi i f = -2\pi c(x - a)f.$$

Setting $a = b = 0$, we immediately deduce $f(x)\varphi_c(x) = e^{-\pi x^2}$ and the general case follows accordingly.

Remark 5.1.5. *In quantum mechanics, time and frequency are replaced by momentum and position and the inequality above is the famous statement of the Heisenberg uncertainty principle.*

5.2 The Short-time Fourier transform and the Spectrogram

Among the many existing and partially quite sophisticated signal representations which conceptually go beyond pure Fourier transforms, the short-time Fourier transform (STFT) is both the best-known and most straight-forward to derive from the principles of Fourier analysis: is function is localized by multiplication with a window function and then, a Fourier transform is applied. This a priori simple and beautiful idea entails a lot of interesting mathematical questions, even before going into problems such as discretization, which we will look at in Section 5.3.

5.2.1 Analysis of a time-variant signal: Short-time Fourier transform

The STFT is used to determine the frequency content of local sections of a signal that changes over time. The function f of interest is multiplied by a window function which is nonzero for only a short period of time and the Fourier transform of the resulting, localized signal is computed.

Definition 5.2.1 (STFT)**.** *Let $\varphi \in L^2(\mathbb{R})$ be the window function and $f \in L^2(\mathbb{R})$ the signal to be analysed.*

$$\mathbf{STFT}\,\{f\} = \mathcal{S}_\varphi f(\tau,\omega) = \int_{-\infty}^{\infty} f(t)\overline{\varphi}(t-\tau)e^{-2\pi \mathrm{i}\omega t}\,dt = \int_{-\infty}^{\infty} f(t)\overline{M_\omega T_\tau \varphi(t)}\,dt,$$

Note that $\mathcal{S}_\varphi f$ is essentially the Fourier Transform of $f \cdot \varphi$, a complex function representing the phase and magnitude of the signal over time and frequency.

Instantaneous Frequencies and some examples

To gain some intuitive understanding of the kind of signal representation the STFT provides, let us consider the concept of "instantaneous frequency", which may be defined for signals with *slowly* varying frequency. Assume that a signal is given by $f(t) = \sin(\theta(t))$, with $\theta(t)$ smooth, so that it can be approximated by its Taylor polynomial in the vicinity of time t:

$$\theta(t+\tau) \approx \theta(t) + \theta'(t)\cdot\tau,$$

then, for small τ

$$\sin(\theta(t+\tau)) \approx \sin(\theta(t) + 2\pi\cdot\frac{\theta'(t)}{2\pi}\cdot\tau),$$

and $\nu^{loc} = \frac{\theta'(t)}{2\pi}$ can be interpreted as "local frequency" at time $t+\tau$.

Example 5.2.2. *Consider the three signals*

$$\begin{aligned} f_1 &= \sin(2\pi\cdot 500\cdot t)\\ f_2 &= \sin(2\pi\cdot(500\cdot t + (50/\pi)\sin 2\pi t))\\ f_3 &= \sin(2\pi\cdot(500\cdot t + (125/2)t^2)), \end{aligned}$$

then $\nu_1^{loc} = 500Hz$, $\nu_2^{loc} = (500 + 100\cos 2\pi t)Hz$ and $\nu_3^{loc} = (500 + 125t)Hz$. Compare these findings with the spectrograms shown in Figure 5.1.

We next answer the obvious question whether a function f can be reconstructed from its STFT and what are the necessary conditions on the window φ. We will need the following result, from which inversion of the STFT follows as a corollary.

Theorem 5.2.3 (Orthogonality relations for the STFT)**.** *Let $f_1, f_2, \varphi_1, \varphi_2 \in L^2(\mathbb{R})$. Then $\mathcal{S}_{\varphi_j} f_j \in L^2(\mathbb{R}^2)$ and*

$$\langle \mathcal{S}_{\varphi_1} f_1, \mathcal{S}_{\varphi_2} f_2\rangle_{L^2(\mathbb{R}^2)} = \langle f_1, f_2\rangle\overline{\langle \varphi_1, \varphi_2\rangle}. \tag{5.1}$$

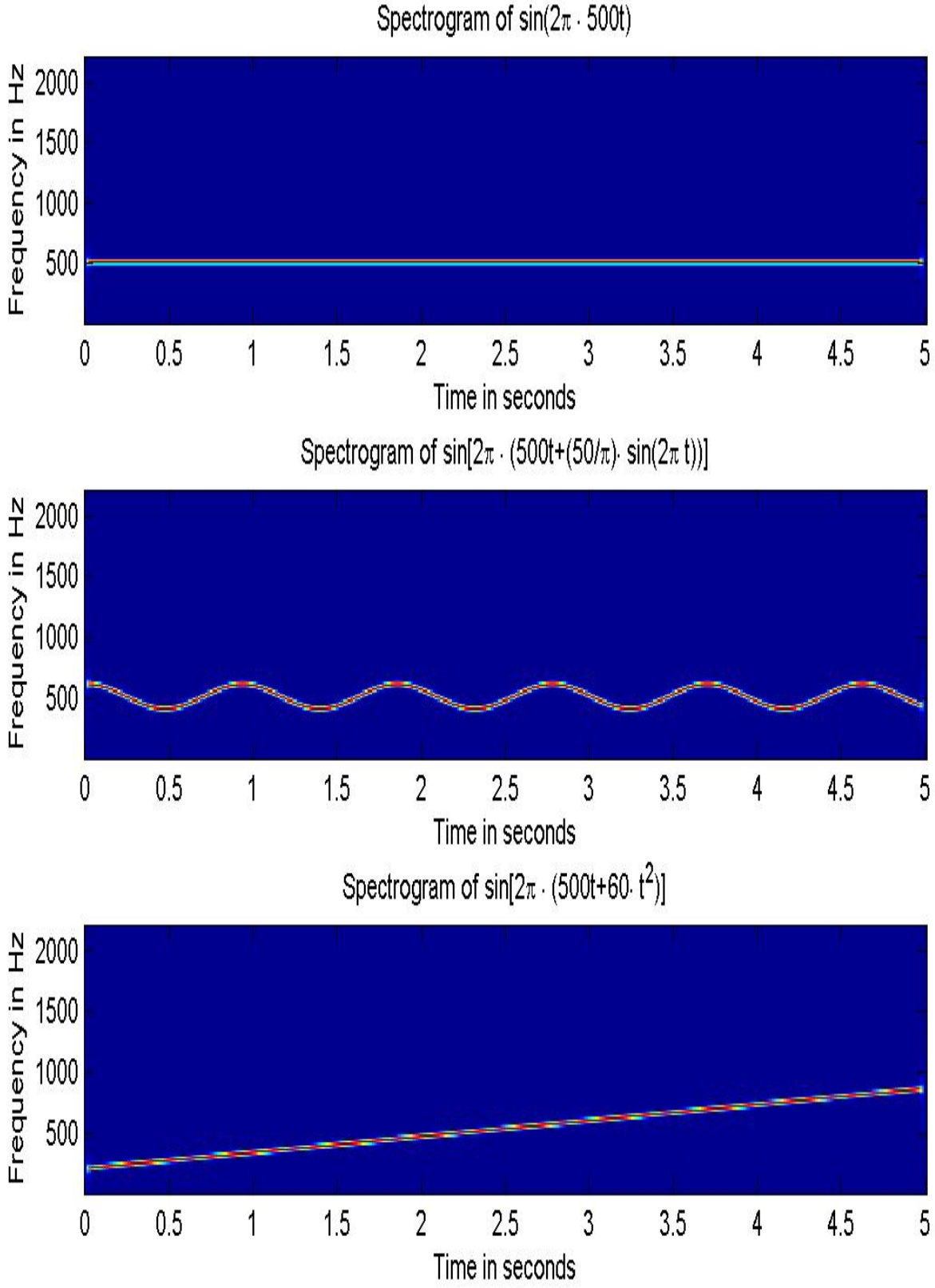


Figure 5.1: Spectrograms of slowly time-variant signals. The instantaneous frequency, as computed in Example 5.2.2 is visible.

Proof.

$$\langle \mathcal{S}_{\varphi_1} f_1, \mathcal{S}_{\varphi_2} f_2 \rangle_{L^2(\mathbb{R}^2)} = \int_\tau \int_\xi \mathcal{S}_{\varphi_1} f_1(\tau,\xi) \overline{\mathcal{S}_{\varphi_2} f_2(\tau,\xi)} d\tau d\xi \tag{5.2}$$

$$= \int_\tau (\int_\xi \mathcal{F}(T_\tau \bar{\varphi_1} \cdot f_1)(\xi) \overline{\mathcal{F}(T_\tau \overline{\varphi_2} \cdot f_2)(\xi)}) d\tau) d\xi \tag{5.3}$$

$$= \int_\tau (\int_t (T_\tau \bar{\varphi_1} \cdot f_1)(t) \overline{(T_\tau \overline{\varphi_2} \cdot f_2)(t)}) d\tau) dt, \tag{5.4}$$

where the last equality holds because of the unitarity of the Fourier transform. Hence we have

$$\langle \mathcal{S}_{\varphi_1} f_1, \mathcal{S}_{\varphi_2} f_2 \rangle_{L^2(\mathbb{R}^2)} = \int_\tau \int_t \overline{\varphi_1}(t-\tau) f_1(t) \overline{f_2(t)} \varphi_2(t-\tau)) d\tau) dt = \tag{5.5}$$

$$= \langle f_1, f_2 \rangle \int_\tau \bar{\varphi_1}(\tau) \varphi_2(\tau) d\tau = \langle f_1, f_2 \rangle \overline{\langle \varphi_1, \varphi_2 \rangle}. \tag{5.6}$$

□

Corollary 5.2.4. *Let $f\varphi_1, \varphi_2 \in L^2(\mathbb{R})$ with $\langle \varphi_1, \varphi_2 \rangle \neq 0$. Then*

1. *Isometry of the STFT:* $\|\mathcal{S}_{\varphi_1} f\|_2 = \|f\|_2 \|\varphi_1\|_2$
2. *Inversion formula for STFT:*

$$f = \frac{1}{\langle \varphi_1, \varphi_2 \rangle} \int_x \int_\omega \mathcal{S}_{\varphi_1} f(x,\omega) M_\omega T_x \varphi_2 d\omega \, dx, \tag{5.7}$$

where the vector-valued integral must be understood in a weak sense.

Proof. This follows immediately from the previous Theorem, since

$$\langle \frac{1}{\langle \varphi_1, \varphi_2 \rangle} \int_x \int_\omega \mathcal{S}_{\varphi_1} f(x,\omega) M_\omega T_x \varphi_2 d\omega \, dx, h \rangle = \frac{1}{\langle \varphi_1, \varphi_2 \rangle} \int_x \int_\omega \mathcal{S}_{\varphi_1} f(x,\omega) \cdot \overline{\mathcal{S}_{\varphi_2} h(x,\omega)} dx d\omega$$

$$\frac{1}{\langle \varphi_1, \varphi_2 \rangle} \langle \mathcal{S}_{\varphi_1} f, \mathcal{S}_{\varphi_2} h \rangle = \langle f, h \rangle$$

where the last step follows from the Orthogonality relations for the STFT. □

Example 5.2.5 (STFT of Gaussian)**.** *For the Gaussian $g(t) = e^{-\pi t^2}$, note that:*

1. *The Gaussian is real and symmetric, so $\overline{g(t-\tau)} = g(t-\tau)$.*
2. *The Fourier transform of a Gaussian $e^{-\pi t^2}$ is also a Gaussian $e^{-\pi \omega^2}$.*

Substituting $g(t)$ into the STFT:

$$S_g g(\tau, \omega) = \int_{-\infty}^{\infty} e^{-\pi t^2} e^{-\pi(t-\tau)^2} e^{-2\pi i \omega t}\, dt.$$

Expanding the terms:

$$e^{-\pi(t-\tau)^2} = e^{-\pi(t^2 - 2t\tau + \tau^2)} = e^{-\pi t^2} e^{2\pi t\tau} e^{-\pi\tau^2}.$$

Substituting this back into the integral:

$$S_g g(\tau, \omega) = e^{-\pi\tau^2} \int_{-\infty}^{\infty} e^{-2\pi t^2} e^{2\pi t(\tau - i\omega)}\, dt.$$

This integral has the form of a Gaussian integral. Completing the square in the exponent:

$$-2\pi t^2 + 2\pi t(\tau - i\omega) = -2\pi \left(t^2 - t(\tau - i\omega)\right).$$

Completing the square:

$$t^2 - t(\tau - i\omega) = \left(t - \frac{\tau - i\omega}{2}\right)^2 - \frac{(\tau - i\omega)^2}{4}.$$

Thus, the exponent becomes:

$$-2\pi t^2 + 2\pi t(\tau - i\omega) = -2\pi \left[\left(t - \frac{\tau - i\omega}{2}\right)^2 - \frac{(\tau - i\omega)^2}{4}\right].$$

Factoring this into:

$$e^{-2\pi t^2 + 2\pi t(\tau - i\omega)} = e^{-\frac{\pi}{2}(\tau - i\omega)^2} e^{-2\pi\left(t - \frac{\tau - i\omega}{2}\right)^2}.$$

Now the integral becomes:

$$S_g g(\tau, \omega) = e^{-\pi\tau^2} e^{-\frac{\pi}{2}(\tau - i\omega)^2} \int_{-\infty}^{\infty} e^{-2\pi\left(t - \frac{\tau - i\omega}{2}\right)^2} dt.$$

The integral of a Gaussian is well known:

$$\int_{-\infty}^{\infty} e^{-a(t-b)^2}\, dt = \sqrt{\frac{\pi}{a}}, \quad \textit{for } a > 0.$$

Here, $a = 2\pi$, so the integral evaluates to:

$$\int_{-\infty}^{\infty} e^{-2\pi\left(t - \frac{\tau - i\omega}{2}\right)^2} dt = \sqrt{\frac{\pi}{2\pi}} = \frac{1}{\sqrt{2}}.$$

Thus:

$$S_g g(\tau,\omega) = e^{-\pi\tau^2} e^{-\frac{\pi}{2}(\tau - i\omega)^2} \cdot \frac{1}{\sqrt{2}}.$$

Simplifying the exponentials:

$$-\pi\tau^2 - \frac{\pi}{2}(\tau - i\omega)^2 = -\pi\tau^2 - \frac{\pi}{2}(\tau^2 - 2i\omega\tau - \omega^2).$$

$$= -\pi\tau^2 - \frac{\pi}{2}\tau^2 + \pi i\omega\tau - \frac{\pi}{2}\omega^2.$$

Combining terms:

$$S_g g(\tau,\omega) = \frac{1}{\sqrt{2}} e^{-\frac{3\pi}{2}\tau^2 - \frac{\pi}{2}\omega^2 + \pi i\omega\tau}.$$

Thus, the STFT of $g(t) = e^{-\pi t^2}$ with respect to itself is:

$$S_g g(\tau,\omega) = \frac{1}{\sqrt{2}} e^{-\frac{3\pi}{2}\tau^2 - \frac{\pi}{2}\omega^2 + \pi i\omega\tau}.$$

5.2.2 The spectrogram as energy density

The spectrogram is modulus squared of the STFT:

Definition 5.2.6. *Let $\varphi \in L^2(\mathbb{R})$ with $\|\varphi\|_2 = 1$, then the spectrogram of f with respect to φ is defined as*

$$Spec_\varphi f(x,\omega) := |\mathcal{S}_\varphi f(x,\omega)|^2.$$

Exercise 5.2.7. *Derive the following properties from the corresponding statements about the STFT:*

- *$Spec_\varphi f(x,\omega) \geq 0$ for all $x, \omega \in \mathbb{R}$.*
- *$Spec_\varphi(T_u M_\eta f)(x,\omega) = Spec_\varphi f(x-u, \omega-\eta)$.*
- *$\int_x \int_\omega Spec_\varphi f(x,\omega) dx d\omega = \|f\|_2^2$.*

Given these properties, it makes sense to interpret the spectrogram as an energy density in the TF-plane. It is also used in all the visualizations of the STFT, such as those that we have seen so far, as well as for further processing in many applications. An important question is then, whether and in which sense f may be reconstructed from a corresponding spectrogram. The following statement gives an answer and is a prototype of results from the active research area of phase retrieval, cf. e.g. [2][1] The following statement is prototypical and given without a proof.

Proposition 5.2.8. *Let $f, h, \varphi \in L^2(\mathbb{R})$ and let the (two-dimensional) Fourier transform of $Spec_\varphi \varphi$ be nonzero almost everywhere. If $Spec_\varphi f = Spec_\varphi h$, then $h = e^{i\alpha}$ for some $\alpha \in \mathbb{R}$.*

[1] Master thesis?

5.3 Frames

Just as the inversion of non-singular square matrices is not the end of the story of matrix inversion, bases, let alone orthogonal bases, are not the end of the story of the expansion of functions, or signals. In particular, when we are interested in the analysis of speech or music signals, we face insurmountable difficulties if we stick with the idea of analyzing functions or signals by using (orthonormal) bases. In this section, we will encounter an important generalization of bases, the so-called concept of *frames*. We will motivate this new idea with the most important analysis method used in audio signal processing.

Motivation: the deficiencies of bases in audio analysis

It is quite obvious that a reasonable analysis of time-variant signals such as music (or speech) requires a transformation that is local in both time and frequency. In other words, it doesn't help a lot to have the frequency information for an entire music piece: we would like to know *which* frequency sounds at *which* time. However, why can't we just cut our - probably sampled and thus discrete - signal into pieces of finite length and analyze each of them separately?
The answer to this question can be seen in Figure 5.2 and Figure 5.4 and is related to the behavior of the Fourier transform of the box function. Indeed, "cutting the signal into pieces" is equivalent to multiplying it with a number of shifted box functions, one piece would then be given by $f_{loc} = f \cdot \Pi$, hence $\widehat{f_{loc}} = \hat{f} * \hat{\Pi} = \hat{f} * sinc$. Then, each piece can be easily recovered by means of the inverse Fourier transform and in fact, it is even straightforward to see that we obtain an ONB if the shifted box-functions are chosen accordingly.
So far so good, but look at Figure 5.4, where the spectrogram (the magnitude squared of the STFT) of such an analysis of a very simple signal is shown. The problem with the cutting approach is the fact that *sinc* decays slowly, and that means that, as a result of $\widehat{f_{loc}} = \hat{f} * sinc$, the frequency information appears totally smeared. Note that the signal shown is extremely simple, just two sinusoids multiplied by an envelope, but even for this signal it would be hard to separate the two components.
On the other hand, have a look at the other proposed window, a smoother window, called Hanning window, which is very popular in audio signal processing. This window's Fourier transform has a much better decay, look at Figure 5.2, and as a result, the frequency resolution in the STFT obtained by application of this window is a lot better. However, since the Hanning window decays towards 0 smoothly, in order not to lose information, we have to assure some overlap in the translation of the windows. In that way, we are not able to obtain orthogonal bases.

5.3.1 Frames

Redundancy and Robust Reconstruction in Frames

Frames versus Bases

A *frame* for a Hilbert space $\mathcal{H}$ is a (typically redundant) collection of vectors $\{\varphi_k\}_{k=1}^{M}$

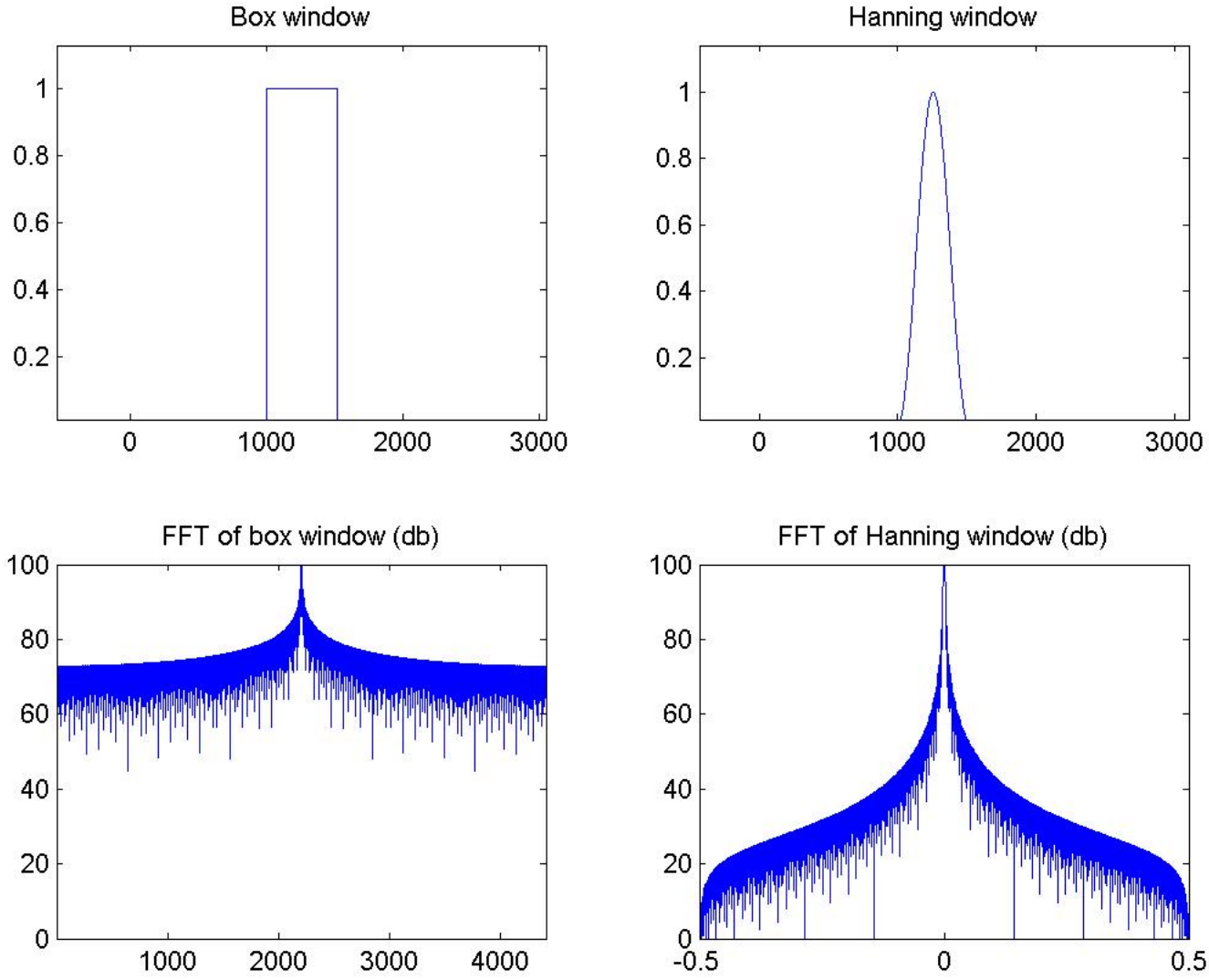


Figure 5.2: A box function as a window and a Hanning window. The lower plots show the respective Fourier transforms.

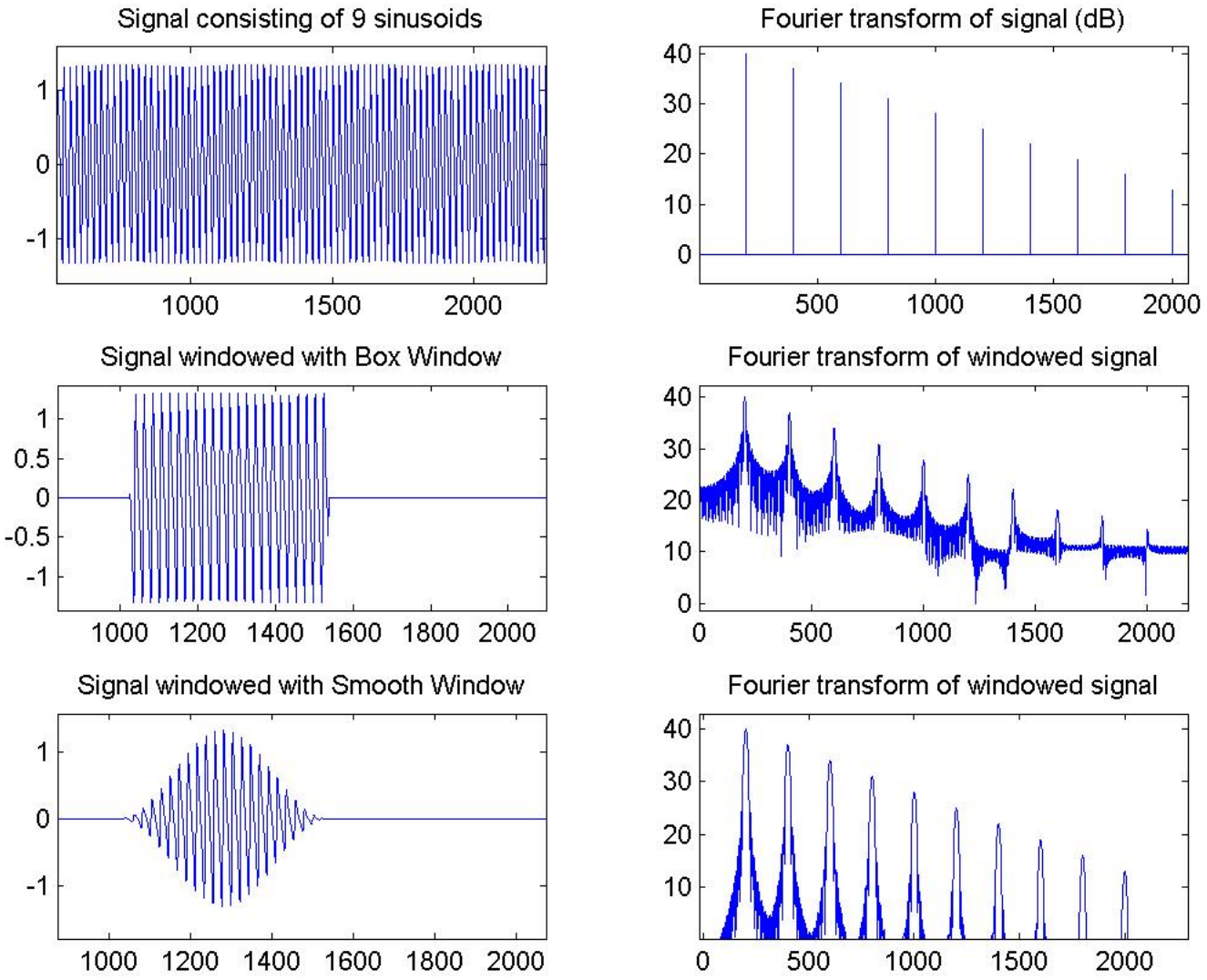


Figure 5.3: Comparing the respective Fourier transforms of a simple signal and its windowed versions: once with a box function, once with a hanning window.

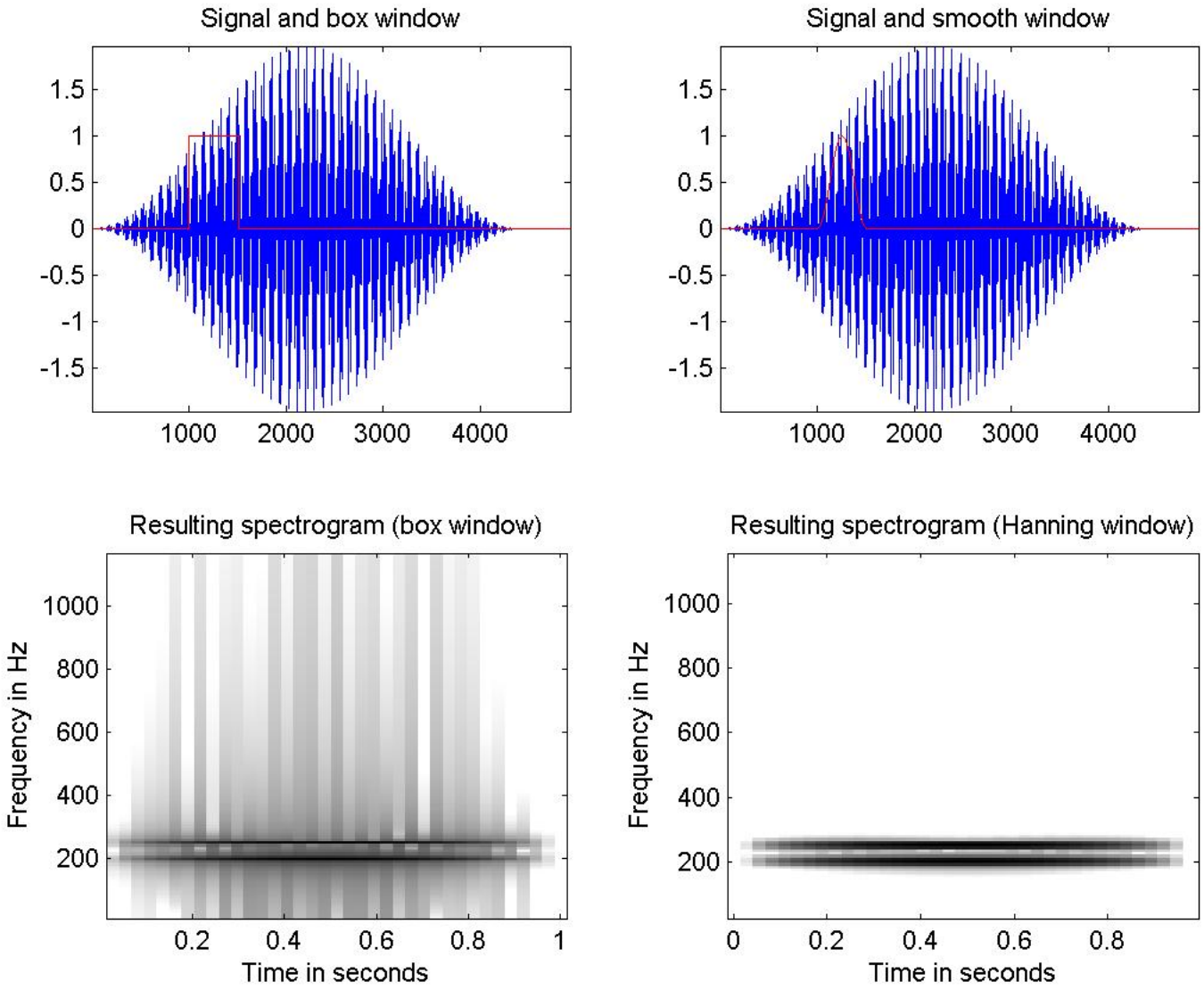


Figure 5.4: Spectrograms resulting from the usage of different windows.

satisfying

$$A\|f\|^2 \leq \sum_{k=1}^{M} |\langle f, \varphi_k \rangle|^2 \leq B\|f\|^2, \qquad \forall f \in \mathcal{H},$$

with frame bounds $0 < A \leq B < \infty$. Unlike an orthonormal basis, frames may contain more vectors than the dimension of the space, which provides *redundancy*. This redundancy improves numerical stability, offers flexibility in representation, and allows partial reconstruction even when some coefficients are missing or corrupted.

If Φ denotes the analysis operator (matrix of size $M \times N$ with $M \geq N$), the frame operator is $S = \Phi^*\Phi$. For perfect, uncorrupted coefficients $c = if$, the canonical dual reconstruction is given by

$$f = S^{-1}\Phi^* c.$$

This formula guarantees exact recovery when all coefficients are available.

Effect of Coefficient Corruption

In practical scenarios, coefficients may be *missing* or *corrupted*. Let M denote a diagonal masking operator with entries $m_k \in \{0, 1\}$ indicating whether the coefficient c_k is available. The observed data are then

$$c_{\text{obs}} = Mc = M\Phi f.$$

If one naively applies the canonical dual reconstruction

$$\hat{f}_{\text{can}} = S^{-1}\Phi^* c_{\text{obs}},$$

the missing coefficients are implicitly treated as zeros. This introduces bias and may significantly degrade reconstruction quality, especially when many coefficients are lost.

Masked and Regularized Reconstruction

A more principled approach is to explicitly incorporate the mask M into the reconstruction problem by solving a regularized least-squares system:

$$\hat{f}_\lambda = \arg\min_f \|M(\Phi f) - c_{\text{obs}}\|_2^2 + \lambda\|f\|_2^2,$$

which leads to the *Tikhonov-regularized normal equations*

$$(\Phi^* M\Phi + \lambda I)\hat{f}_\lambda = \Phi^* M c_{\text{obs}}.$$

Here $\lambda > 0$ controls the trade-off between fidelity and stability. Typical values are chosen as $\lambda \approx 10^{-3}$–$10^{-2}\,\lambda_{\max}(S)$. For small- to medium-size problems, this system can be solved directly; for large-scale frames, iterative solvers such as conjugate gradients (PCG) or `lsqr` are preferred.

Alternative Formulations

- **Submatrix reconstruction:** Extract the rows of Φ corresponding to observed coefficients, Φ_K, and solve
$$(\Phi_K^* \Phi_K + \lambda I)\hat{f} = \Phi_K^* c_K.$$
This is equivalent to the masked formulation and efficient when many coefficients are missing.

- **Iterative reconstruction (PCG):** Define the operator $A(f) = \Phi^*(M(\Phi f)) + \lambda f$ and solve $A(f) = \Phi^* M c_{\text{obs}}$ iteratively using PCG.

- **Sparsity-promoting reconstruction:** If the signal is known to be sparse in some domain, one may use an ℓ_1-regularized problem
$$\hat{f} = \arg\min_f \|M(\Phi f) - c_{\text{obs}}\|_2^2 + \tau \|f\|_1,$$
which can be solved using LASSO, SPGL1, or proximal gradient methods (ISTA/-FISTA).

Practical Insights

1. Redundant frames ($M > N$) provide resilience against coefficient loss.
2. The canonical dual reconstruction is optimal only when all coefficients are available.
3. Incorporating the mask M leads to significantly improved robustness.
4. Regularization (λI) prevents amplification of noise and numerical instability.
5. In high-dimensional cases, iterative solvers offer computational efficiency.

MATLAB Implementation Summary
The following code fragment illustrates masked Tikhonov reconstruction:

```
Mdiag = spdiags(double(mask), 0, M, M);
lambda = 1e-3 * max(eig(Phi'*Phi));
A = Phi' * Mdiag * Phi + lambda * speye(N);
b = Phi' * (Mdiag * c_obs);
f_rec = real(A \ b);
```

This formulation correctly accounts for missing coefficients, yielding more accurate and stable results than the naive canonical dual reconstruction.

In summary, *redundant frames transform signal analysis into an overdetermined, and hence robust, representation system.* Their redundancy enables meaningful reconstruction even under partial data loss, provided the reconstruction step explicitly incorporates the structure of missing data.

Direct Motivation: sample the STFT

Replace the integrals in the inversion formula for the STFT, see (5.7), by discrete sums to obtain an approximation by Riemann sums. In other words, hope to obtain a reconstruction of f from samples of $\mathcal{S}_\varphi f$ as:

$$f = \sum_{k\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} \mathcal{S}_{\varphi_1} f(\alpha k, \beta n) M_{\beta n} T_{\alpha k} \varphi_2. \tag{5.8}$$

So, what do we need to know about the collection of functions $\{M_{\beta n} T_{\alpha k}\varphi\}$ in order to obtain such an equality?

General Frames

A set $\Phi = \{\varphi_j\}_{j\in J}$ in a Hilbert space $\mathcal{H}$ is complete if every element in $\mathcal{H}$ can be approximated arbitrarily well (in norm) by finite linear combinations of elements in Φ. For a finite-dimensional vector space such as $\mathbb{R}^n$ or $\mathbb{C}^n$, this simply means that any vector in $\mathcal{H}$ can be written as a linear combination of the $\{\varphi_j\}$, in other words, that Φ spans $\mathcal{H}$.
A complete set is overcomplete, if removal of one element of the set still results in a complete system. In signal processing, overcompleteness can help to achieve a more stable, more robust, or more compact decomposition than the usage of a basis which implies uniqueness. *Frames* are an interesting generalization of bases and are widely used in mathematics, computer science, engineering, and statistics.

Definition 5.3.1 (Frames)**.** *Let $\mathcal{H}$ be a Hilbert spaces (a finite-dimensional vector space with inner product. A frame is defined to be a countable family of non-zero vectors $\{\varphi_j\}_{j\in J}$ in $\mathcal{H}$, such that for arbitrary $f \in \mathcal{H}$,*

$$C_l \|f\|^2 \leq \sum_{j\in J} |\langle f, \varphi_j\rangle|^2 \leq C_u \|f\|^2$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product, C_l and C_u are positive constants called lower and upper frame bound, respectively. If C_l, C_u can be chosen such that $C_l = C_u$, the frame is called a tight frame.

Note that the above inequality can be understood as an "approximate Plancherel formula" and in that sense, an ONB is a special case of a (tight) frame. Recall, that any ONB $\{\psi_j\}_{j\in J}$ implies a convenient signal representation for all $f \in \mathcal{H}$, given by

$$f = \sum_{j\in J} \langle f, \psi_j\rangle \psi_j \quad \text{with} \quad \sum_{j\in J} |\langle f, \psi_j\rangle|^2 = \|f\|_2^2.$$

For frames, the situation is, obviously, slightly more complicated. In general, we can not expect the mapping $S : f \to \sum_{j\in J} \langle f, \varphi_j\rangle \varphi_j$ to be identity, but at least, this mapping will be invertible, a property that opens the door to the virtues of frames.

Definition 5.3.2 (Frame operator)**.** *Consider a vector space $\mathcal{V}$ and a family of elements $\{\varphi_j\}_{j\in J}$ in $\mathcal{V}$.*

$$Sf = \sum_{j\in J} \langle f, \varphi_j\rangle \varphi_j \tag{5.9}$$

Note that

$$\langle Sf, f\rangle = \langle \sum_{j\in J} \langle f, \varphi_j\rangle \varphi_j, f\rangle = \sum_{j\in J} |\langle f, \varphi_j\rangle|^2.$$

Also note that the frame operator (5.9) can be written as the composition of an *analysis operator* $C : \mathcal{V} \to \mathbb{C}^k$, given by $C : v \to \{\langle v, \varphi_j\rangle\}_{j\in J}$ and the synthesis operator $D : \mathbb{C}^k \to \mathcal{V}$, given by $D : \mathbf{c} \to \sum_{j\in J} c_j\varphi_j$. In fact, D is the adjoint operator, i.e. the transposed, complex conjugate of C: $D = C' = \overline{C^T}$ and thus $S' = (DC)' = (C'C)' = C'C = S$ and S is self-adjoint.

Proposition 5.3.3 (Properties of the frame operator)**.** *Let $\mathcal{V}$ be a finite-dimensional vector space and Φ a frame for $\mathcal{H}$.*
(i) S is invertible and self-adjoint.
(ii) There exist dual frames $\tilde{\varphi}_k$, allowing an expansion of f as:

$$f = \sum_j \langle f, \varphi_j\rangle \tilde{\varphi}_j = \sum_j \langle f, \tilde{\varphi}_j\rangle \varphi_j \tag{5.10}$$

In particular, the canonical dual frame is given by $\tilde{\varphi}_k = S^{-1}\varphi_k$.
(iii) If the frame is not a basis, then the coefficients $c_j = \langle f, \tilde{\varphi}_j\rangle$ are not unique, but optimal in the sense of minimizing $\sum_j |c_j|^2$.

Remark 5.3.4. *Completely analogous statements hold for infinite dimensional Hilbert spaces, but the proof is beyond our scope.*

Proof. (i) We show that S is injective: assume that $Sf = 0$ for $f \in \mathcal{V}$, then

$$0 = \langle Sf, f\rangle = \sum_{j=1}^{k} |\langle f, \varphi_j\rangle|^2 \geq C_l \|f\|^2 \Rightarrow \|f\| = 0 \Rightarrow f = 0.$$

(ii) Since S is invertible, we can write

$$f = S^{-1}Sf = S\left(\sum_j \langle f, \varphi_j\rangle \varphi_j\right) = \sum_j \langle f, \varphi_j\rangle S^{-1}(\varphi_j)$$

and setting $\tilde{\varphi}_j = S^{-1}(\varphi_j)$ proves the first equality. Changing order: $f = SS^{-1}f$ and noting that self-adjointness of S leads to self-adjointness of S^{-1}, we obtain

$$f = \sum_j \langle S^{-1}f, \varphi_j\rangle \varphi_j = \sum_j \langle f, S^{-1}\varphi_j\rangle \varphi_j.$$

(iii) The proof of this statement is similar to the proof of Proposition A.4.18. In fact, suppose that the coefficient vector $\mathbf{c} \in \mathbb{C}^k$ fulfills $f = \sum_{j=1}^k c_j \varphi_j$, hence, since $c_j = c_j - \langle f, S^{-1}\varphi_j\rangle + \langle f, S^{-1}\varphi_j\rangle$ and $f = \sum_{j=1}^k \langle f, S^{-1}\varphi_j\rangle \varphi_j$, we have

$$\sum_{j=1}^k (c_j - \langle f, S^{-1}\varphi_j\rangle)\varphi_j = 0.$$

Setting $d_j = c_j - \langle f, S^{-1}\varphi_j\rangle$, this can be written as $D\mathbf{d} = 0$ and thus the sequence $\mathbf{d}$, given by $d_j = c_j - \langle f, S^{-1}\varphi_j\rangle$ is in the kernel of $D = C'$, which is orthogonal to the range of C. On the other hand, the sequence $\{\langle f, S^{-1}\varphi_j\rangle\}_{j=1}^k$ is in the range of C, because

$$\{\langle f, S^{-1}\varphi_j\rangle\}_{j=1}^k = \{\langle S^{-1}f, \varphi_j\rangle\}_{j=1}^k = C(S^{-1}f).$$

Therefore, as in the proof of Proposition A.4.18:

$$\begin{aligned}\sum_{j=1}^k |c_j|^2 &= \sum_{j=1}^k |c_j - \langle f, S^{-1}\varphi_j\rangle + \langle f, S^{-1}\varphi_j\rangle|^2 \\ &= \sum_{j=1}^k |c_j - \langle f, S^{-1}\varphi_j\rangle|^2 + \sum_{j=1}^k |\langle f, S^{-1}\varphi_j\rangle|^2 + \\ &+ \sum_{j=1}^k \langle c_j - \langle f, S^{-1}\varphi_j\rangle, \langle f, S^{-1}\varphi_j\rangle\rangle,\end{aligned}$$

and since the last sum is 0 due the orthogonality of the range of C and the kernel of D, we have

$$\sum_{j=1}^k |c_j|^2 = \sum_{j=1}^k |c_j - \langle f, S^{-1}\varphi_j\rangle|^2 + \sum_{j=1}^k |\langle f, S^{-1}\varphi_j\rangle|^2 \geq \sum_{j=1}^k |\langle f, S^{-1}\varphi_j\rangle|^2.$$

□

5.3.2 Gabor Frames: Structure and Existence

We now look at the special kind of frames that are associated with the STFT: Gabor frames.

For a given function $g \in L^2(\mathbb{R})$, and two non-negative constants a, b, the family

$$\mathcal{G}_{g,a,b} = \{g_{m,n} := M_{mb}T_{na}g, m, n \in \mathbb{Z}\}$$

for $m, n \in \mathbb{Z}$, is called a Gabor system with analysis window g, time-shift parameter a and frequency-shift parameter b. The associated frame operator is then given by

$$S_{g,a,b}f := \sum \langle f, g_{m,n}\rangle g_{m,n}. \tag{5.11}$$

While the dual frame of a given frame (whose frame inequalities are also not always trivial to derive) is in theory simply given by the inversion of the frame operator, the latter is infeasible in higher dimensions. Thus, we must take a closer look at the structure of the Gabor frame operator in order to get some easier solutions. The following three statements provide the necessary insights.

Proposition 5.3.5 (Gabor frame operator)**.** *Let a Gabor system $\mathcal{G}_{g,a,b}$ be given. Then $S_{g,a,b}$ commutes with the time-frequency shifts $M_{mb}T_{na}$ and hence there exists a dual window γ, such that $\mathcal{G}_{\gamma,a,b}$ is a dual frame for $\mathcal{G}_{g,a,b}$.*

Proof. 1. **Frame Operator Definition:** The Gabor frame operator $S_{g,a,b}$ is given by:

$$S_{g,a,b}f = \sum_{m,n\in\mathbb{Z}} \langle f, M_{mb}T_{na}g\rangle M_{mb}T_{na}g,$$

where the operators T_{na} (time shift) and M_{mb} (frequency shift) are defined as:

$$T_{na}g(t) = g(t-na), \quad M_{mb}g(t) = e^{2\pi imbt}g(t).$$

2. **Commutativity with Time-Frequency Shifts:** Let $S_{g,a,b}$ act on a time-frequency shift $M_{m'b}T_{n'a}f$, where f is an arbitrary function. We aim to show:

$$S_{g,a,b}(M_{m'b}T_{n'a}f) = M_{m'b}T_{n'a}(S_{g,a,b}f).$$

- Compute $S_{g,a,b}(M_{m'b}T_{n'a}f)$:

$$S_{g,a,b}(M_{m'b}T_{n'a}f) = \sum_{m,n\in\mathbb{Z}} \langle M_{m'b}T_{n'a}f, M_{mb}T_{na}g\rangle M_{mb}T_{na}g.$$

Using the orthogonality and commutation properties of time-frequency shifts, we have:

$$\langle M_{m'b}T_{n'a}f, M_{mb}T_{na}g\rangle = \langle f, M_{(m-m')b}T_{(n-n')a}g\rangle.$$

Substituting this back, we get:

$$S_{g,a,b}(M_{m'b}T_{n'a}f) = \sum_{m,n\in\mathbb{Z}} \langle f, M_{(m-m')b}T_{(n-n')a}g\rangle M_{mb}T_{na}g.$$

- Rewrite the summation: Let $m' = m - m'$ and $n' = n - n'$. Then:

$$S_{g,a,b}(M_{m'b}T_{n'a}f) = \sum_{m',n'\in\mathbb{Z}} \langle f, M_{m'b}T_{n'a}g\rangle M_{(m'+m')b}T_{(n'+n')a}g.$$

Noting that $M_{(m'+m')b}T_{(n'+n')a} = M_{m'b}T_{n'a}M_{mb}T_{na}$, we can factor out $M_{m'b}T_{n'a}$:

$$S_{g,a,b}(M_{m'b}T_{n'a}f) = M_{m'b}T_{n'a}\sum_{m,n\in\mathbb{Z}} \langle f, M_{mb}T_{na}g\rangle M_{mb}T_{na}g.$$

- Recognize the summation as $S_{g,a,b}f$:

$$S_{g,a,b}(M_{m'b}T_{n'a}f) = M_{m'b}T_{n'a}(S_{g,a,b}f).$$

Thus, $S_{g,a,b}$ commutes with $M_{m'b}T_{n'a}$.

3. **Existence of a Dual Frame:** Since $S_{g,a,b}$ is a positive, self-adjoint, and invertible operator (as the Gabor system is a frame), it admits a unique inverse $S_{g,a,b}^{-1}$. Define the dual window γ as:

$$\gamma = S_{g,a,b}^{-1}g.$$

Then the Gabor system $\mathcal{G}(\gamma, a, b)$ forms a dual frame for $\mathcal{G}(g, a, b)$. That is:

$$\langle f, M_{mb}T_{na}\gamma\rangle = c_{m,n},$$

where $c_{m,n}$ reconstructs f in the dual frame decomposition. □

Theorem 5.3.6 (Walnut representation)**.** *Assume that the window $g \in L^2(\mathbb{R})$ is piece-wise continuous and decays sufficiently fast. Then the frame operator $S_{g,a,b}$ can be written as*

$$S_{g,a,b} = b^{-1}\sum_{n\in\mathbb{Z}} G_n \cdot T_{\frac{n}{b}}, \tag{5.12}$$

where $G_n(x) = \sum_{k\in\mathbb{Z}} \overline{g}(x - \frac{n}{b} - ak)g(x - ak)$.

Remark 5.3.7. *The technical assumption is, that g lies in the Wiener space: $g \in W(L^\infty, \ell^1)$, in detail: : The Wiener amalgam space $W(L^\infty, \ell^1)$ consists of functions $f \in L^\infty_{loc}(\mathbb{R})$ such that $\|f\|_{W(L^\infty,\ell^1)} = \sum_{k\in\mathbb{Z}} \|f \cdot \chi_{[k,k+1]}\|_{L^\infty} < \infty$, where $\chi_{[k,k+1]}$ is the indicator function for the interval $[k, k + 1]$. Intuitively, this norm measures the local L^∞ norm of f over each unit interval and sums these values globally.*

Proof. Fix $f \in C_c(\mathbb{R})$ and $k \in \mathbb{Z}$. Then $f \cdot T_{ak}g$ is bounded and compactly supported. The Gabor frame operator $S_{g,a,b}$ is given by:

$$S_{g,a,b}f = \sum_{n\in\mathbb{Z}}(\sum_{k\in\mathbb{Z}}\langle f, M_{kb}T_{na}g\rangle M_{kb})T_{na}g,$$

which, since the sequence $(\langle f, M_{kb}T_{na}g\rangle)_k$ is in ℓ^2, may be rewritten as

$$S_{g,a,b}f(x) = \frac{1}{b}\sum_{n\in\mathbb{Z}}(\sum_{k\in\mathbb{Z}}(f\cdot T_{an}\overline{g})(x-\frac{k}{b}))T_{na}g(x) = \frac{1}{b}\sum_{k\in\mathbb{Z}}(\sum_{n\in\mathbb{Z}}\overline{g}(x-\frac{k}{b}-na)g(x-na))f(x-\frac{n}{b}),$$

where exchanging the order of the sums is allowed due to the compactness assumption. Note that G_n is the a-periodization of $g \cdot T_{k/b}\overline{g}$ and that, by assumption, we have $G_n \in L^\infty(\mathbb{R})$. We can hence extend the new form of the operator $S_{g,a,b}$ to all of L^2 and obtain

$$S_{g,a,b}f(x) = \frac{1}{b}\sum_{n\in\mathbb{Z}} G_n \cdot T_{\frac{n}{b}}f(x), \tag{5.13}$$

as claimed. □

Corollary 5.3.8 (Painless Non-Orthogonal Expansions)**.** *Suppose that $g \in L^\infty(\mathbb{R})$ is supported in the interval $[0, L]$. If $a \leq L$ and $b \leq \frac{1}{L}$, then the frame operator $S_{g,a,b}$ is the multiplication operator*

$$S_{g,a,b}f(x) = (\frac{1}{b}\sum_{k\in\mathbb{Z}} |g(x-ak)|^2)f(x). \tag{5.14}$$

Proof: exercise.

Corollary 5.3.9. *If $g \in W(L^\infty, \ell^1)$, then $G(g, a, b)$ is a frame for all sufficiently small a, b.*

Theorem 5.3.10 (Balian–Low, Hilbert space version)**.** *If $\mathcal{G}(g, \mathbb{Z}^2)$ is an orthonormal basis for $L^2(\mathbb{R})$, then either $xg(x) \notin L^2(\mathbb{R})$ or $g'(x) \notin L^2(\mathbb{R})$.*

Proof. The proof relies on the commutation relations for the position and momentum operators, which are defined as

$$(Xg)(x) = xg(x) \quad \text{and} \quad (Pg)(x) = \frac{1}{2\pi i}g'(x).$$

Recall that X and P are self-adjoint and that

$$(PX - XP)g = \frac{1}{2\pi i}g,$$

for $g \in \mathrm{dom}(XP) \cap \mathrm{dom}(PX)$. Furthermore, we have $\mathcal{F}(Pg) = X\mathcal{F}g$.

We prove the theorem by contradiction. Suppose $\mathcal{G}(g, \mathbb{Z} \times \mathbb{Z})$ is an orthonormal basis of $L^2(\mathbb{R})$, in particular $g \neq 0$, and suppose that $Xg \in L^2(\mathbb{R})$ and $Pg \in L^2(\mathbb{R})$. Then, using the orthonormal expansion of Xg, we have

$$\langle Xg, Pg\rangle = \sum_{k,l\in\mathbb{Z}} \langle Xg, M_lT_kg\rangle \langle M_lT_kg, Pg\rangle.$$

We now re-write the inner products in the series expansion. Since

$$XM_lT_kg(x) = (x)e^{2\pi ilx}g(x-k) = kM_lT_kg(x) + M_lT_kXg(x),$$

we have

$$\begin{aligned}\langle Xg, M_lT_kg\rangle &= \langle g, XM_lT_kg\rangle \\ &= k\langle g, M_lT_kg\rangle + \langle g, M_lT_kXg\rangle \\ &= 0 + \langle T_{-k}M_{-l}g, Xg\rangle,\end{aligned}$$

where we used the fact that $\langle g, M_lT_kg\rangle = 0$ because we assume $\mathcal{G}(g, \mathbb{Z} \times \mathbb{Z})$ to be an orthonormal basis.

Similarly,

$$\begin{aligned}\langle PM_lT_kg, g\rangle &= \langle X\widehat{T_lM_{-k}}g, \widehat{g}\rangle \\ &= l\langle \widehat{T_lM_{-k}}g, \widehat{g}\rangle + \langle T_l\widehat{M_{-k}X}g, \widehat{g}\rangle \\ &= \langle M_lT_kPg, g\rangle.\end{aligned}$$

Combining the above results, we obtain

$$\langle Xg, Pg\rangle = \sum_{k,l\in\mathbb{Z}} \langle Pg, T_{-k}M_{-l}g\rangle \langle T_{-k}M_{-l}g, Xg\rangle = \langle Pg, Xg\rangle.$$

If we knew that $g \in \operatorname{dom}(PX) \cap \operatorname{dom}(XP)$, then we could rewrite this identity as

$$0 = \langle (PX - XP)g, g\rangle = \frac{1}{2\pi i}\|g\|_2^2.$$

We choose a sequence $(g_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R})$, such that $\|g_n - g\|_2 \to 0$, $\|Xg_n - Xg\|_2 \to 0$, and $\|Pg_n - Pg\|_2 \to 0$ (such a sequence exists). Then

$$\lim_{n\to\infty} \big(\langle Xg_n, Pg_n\rangle - \langle Pg_n, Xg_n\rangle\big) = \langle Xg, Pg\rangle - \langle Pg, Xg\rangle = 0.$$

On the other hand, since $g_n \in \mathcal{S}(\mathbb{R}) \subset \operatorname{dom}(PX) \cap \operatorname{dom}(XP)$, this limit is also

$$\lim_{n\to\infty} \langle (PX - XP)g_n, g_n\rangle = \frac{1}{2\pi i} \lim_{n\to\infty} \|g_n\|_2^2 = \frac{1}{2\pi i}\|g\|_2^2.$$

Thus, $g = 0$, contradicting the assumption that $\mathcal{G}(g, \mathbb{Z}\times\mathbb{Z})$ is an orthonormal basis. □

Remark 5.3.11. *While the classical uncertainty principle provides a lower bound on the deviations in time and frequency, i.e.,*

$$\|Xg\|_2 \, \|Pg\|_2 = \|Xg\|_2 \, \|X\widehat{g}\|_2 \geq \frac{1}{4\pi}\|g\|_2^2,$$

the Balian–Low theorem (BLT) implies that a window of an orthonormal Gabor basis possesses the maximal uncertainty,

$$\|Xg\|_2 \, \|Pg\|_2 = \infty.$$

For the Hilbert space $L^2(\mathbb{R}^d)$, this result holds for the conjugate variables (x_k, ω_k), $k = 1, \ldots, d$. There are also more general versions for (symplectic) lattices and more general index sets. For more details on this we refer, e.g., to [?]. ◇

We will now state the Wiener amalgam version of the BLT.

Proposition 5.3.12 (Balian–Low theorem, Wiener amalgam version)**.** *If $\mathcal{G}(g, \mathbb{Z}^{2d})$ is a frame for $L^2(\mathbb{R}^d)$, then both*

$$g \notin W_0(\mathbb{R}^d) \quad \textit{and} \quad \widehat{g} \notin W_0(\mathbb{R}^d).$$

To prove the amalgam version of the BLT, we show a special property of the Zak transform.

Definition 5.3.13 (Zak transform)**.** *For $f \in L^2(\mathbb{R}^d)$, the* Zak transform Zf *is defined by*

$$(Zf)(x,\omega) = \sum_{k\in\mathbb{Z}^d} f(x-k)\, e^{2\pi i k\cdot\omega}, \quad (x,\omega) \in \mathbb{R}^{2d}.$$

The Zak transform is $\mathbb{Z}^{2d}$–periodic in (x,ω) and unitary from $L^2(\mathbb{R}^d)$ onto $L^2(Q\times Q)$, where $Q = [0,1)^d$.

Corollary 5.3.14. *Let $g \in L^2(\mathbb{R}^d)$ and consider the Gabor system $\mathcal{G}(g,\mathbb{Z}^{2d})$.*

1. *The frame operator S_g is bounded on $L^2(\mathbb{R}^d)$ if and only if $|Zg|^2 \in L^\infty(\mathbb{R}^{2d})$.*
2. *$\mathcal{G}(g,\mathbb{Z}^{2d})$ is a frame for $L^2(\mathbb{R}^d)$ if and only if there exist constants $0 < a \le b < \infty$ such that*

$$0 < a \le |Zg(x,\omega)| \le b < \infty \quad \textit{for almost all } (x,\omega) \in \mathbb{R}^{2d}.$$

In this case, the optimal frame bounds are

$$A = \operatorname*{ess\,inf}_{(x,\omega)\in Q\times Q} |Zg(x,\omega)|^2, \qquad B = \operatorname*{ess\,sup}_{(x,\omega)\in Q\times Q} |Zg(x,\omega)|^2.$$

Proof. Conjugation by the Zak transform diagonalizes the frame operator:

$$(ZS_gZ^{-1}F)(x,\omega) = |Zg(x,\omega)|^2 F(x,\omega), \qquad F \in L^2(Q\times Q).$$

Hence S_g is bounded on L^2 if and only if multiplication by $|Zg|^2$ is bounded, which is equivalent to $|Zg|^2 \in L^\infty(\mathbb{R}^{2d})$, yielding (a).

Similarly, S_g is invertible if and only if both $|Zg|^2$ and its reciprocal are essentially bounded, giving (b) and the explicit expressions for the optimal frame bounds.

For (c), if $\mathcal{G}(g,\mathbb{Z}^{2d})$ is an orthonormal basis, then $S_g = I_{L^2}$, so ZS_gZ^{-1} acts as multiplication by $|Zg|^2 = 1$. Conversely, if $|Zg|^2 = 1$ a.e., then by (b) $\mathcal{G}(g,\mathbb{Z}^{2d})$ is a tight frame with bound 1. Since $\|M_lT_kg\|_2 = \|g\|_2 = 1$, it follows from Lemma 6.16 that $\mathcal{G}(g,\mathbb{Z}^{2d})$ is an orthonormal basis. □

Lemma 5.3.15. *If Zf is continuous on $\mathbb{R}^{2d}$, then Zf has a zero in $Q\times Q$.*

Sketch of proof of the Zak zero lemma. Suppose Zf is continuous and nonvanishing on the torus $Q\times Q$. Then $(Zf)^{-1}$ is continuous and $\mathcal{Z}^{-1}((Zf)^{-1})$ is a bounded (continuous) function on $\mathbb{R}^d$, which can be used to construct a dual window in the Wiener amalgam class; this contradicts the frame assumption (more concretely: a continuous nonvanishing Zak transform leads to a well-behaved dual Gabor atom, impossible in the critical density case). A detailed version of this argument (or alternative topological proofs using the degree/winding-number of the phase of Zf) can be found in [**?**, **?**]. □

Proof of Theorem 5.3.12. Suppose, for contradiction, that the window $g \in W_0(\mathbb{R}^d)$. By Lemma 5.3.15, the Zak transform Zg is continuous on $\mathbb{R}^{2d}$. Hence, by the preceding lemma, Zg has a zero in $Q \times Q$. According to Corollary 9.7, this implies that the lower frame bound of $\mathcal{G}(g, \mathbb{Z}^{2d})$ is zero, so $\mathcal{G}(g, \mathbb{Z}^{2d})$ cannot be a frame. Therefore, if $\mathcal{G}(g, \mathbb{Z}^{2d})$ *is* a frame, then necessarily $g \notin W_0(\mathbb{R}^d)$.

Finally, note that $\mathcal{G}(g, \mathbb{Z}^{2d})$ is a frame if and only if $\mathcal{G}(\widehat{g}, \mathbb{Z}^{2d})$ is a frame. Applying the same argument to $\widehat{g}$ yields $\widehat{g} \notin W_0(\mathbb{R}^d)$. □

5.3.3 Frames in $\mathbb{C}^n$, matrices and PINV

Excursus: Generalized Inverses and the Moore–Penrose Inverse

In linear algebra, the ordinary inverse of a matrix exists only for square matrices whose columns (equivalently, rows) are linearly independent. However, many applications—such as solving inconsistent or underdetermined systems of linear equations, performing least-squares fitting, or analyzing non-square linear transformations—require a more flexible concept of "inversion." This leads to the notion of a *generalized inverse*.

Unfortunately, the generalization of matrix inversion to singular or rectangular matrices is not unique. Different authors adopt different axioms depending on the intended application. The approach followed here is classical and motivated by Ben-Israel and Greville [**?**].

Axioms for a Generalized Inverse

Following Ben-Israel [**?**], a reasonable definition of a generalized inverse $A^\dagger$ of a matrix A should satisfy at least:

1. If A is invertible, then $A^\dagger$ coincides with the ordinary inverse A^{-1}.
2. Some singular matrices should also possess a generalized inverse.
3. The generalized inverse should share certain algebraic properties with the ordinary inverse.

A minimal condition often required is:

$$ABA = A.$$

Any matrix B satisfying this property is called a *reflexive generalized inverse*. Koecher [**?**] strengthens this by adding the dual condition

$$BAB = B,$$

which eliminates degeneracies and ensures that B behaves like an inverse on the relevant subspaces.

The Moore–Penrose Inverse

Among all generalized inverses, one stands out as the most important: the *Moore–Penrose inverse*. For a matrix $A \in \mathbb{C}^{m\times n}$, the Moore–Penrose inverse $A^+ \in \mathbb{C}^{n\times m}$ is uniquely defined by the four *Moore–Penrose conditions*:

$$AA^+A = A, \tag{5.15}$$
$$A^+AA^+ = A^+, \tag{5.16}$$
$$(AA^+)^* = AA^+, \tag{5.17}$$
$$(A^+A)^* = A^+A. \tag{5.18}$$

Conditions (5.17)–(5.18) require that AA^+ and A^+A be orthogonal projections. This makes the Moore–Penrose inverse the unique generalized inverse that is compatible with orthogonality, least-squares problems, and singular value decomposition.

Special Cases and Explicit Formulas

The Moore–Penrose inverse is easy to compute in several important situations.

Full column rank. If the columns of A are linearly independent, then A^*A is invertible, and

$$A^+ = (A^*A)^{-1}A^*.$$

In this case,

$$A^+A = I_n,$$

so A^+ acts as a left inverse.

Full row rank. If the rows of A are linearly independent, then AA^* is invertible, and

$$A^+ = A^*(AA^*)^{-1},$$

which satisfies

$$AA^+ = I_m.$$

Full rank square matrices. If A is invertible, then the Moore–Penrose inverse reduces to the ordinary inverse:

$$A^+ = A^{-1}.$$

Products involving a unitary matrix. If AB is defined and one of the matrices is unitary, then

$$(AB)^+ = B^+A^+.$$

Scalars and vectors. Treating scalars and vectors as 1×1 and $n \times 1$ matrices, respectively,

$$x^+ = \begin{cases} 0, & x = 0, \\ x^{-1}, & x \neq 0, \end{cases}$$

and for a nonzero vector x,

$$x^+ = \frac{x^*}{x^* x}.$$

Hermitian matrices. If A is Hermitian, its eigen-decomposition $A = UDU^*$ leads to

$$A^+ = UD^+U^*,$$

where D^+ is obtained by inverting nonzero diagonal entries and leaving zeros intact.

Computing the Moore–Penrose Inverse

Several computational formulas are available, depending on rank information.

Rank factorization. If A has rank k, we may write $A = BC$ with $B \in \mathbb{C}^{m \times k}$ and $C \in \mathbb{C}^{k \times n}$. Then

$$A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*.$$

Singular value decomposition (SVD). This is the most robust and conceptually illuminating method. If

$$A = U\Sigma V^*$$

is the SVD, then

$$A^+ = V\Sigma^+ U^*,$$

where Σ^+ is formed by inverting all nonzero singular values and transposing the diagonal matrix.

Numerical considerations. Although the formula $(A^*A)^{-1}A^*$ is easy to implement when A has full column rank, it is numerically unstable because the condition number of A^*A is the square of that of A. More stable methods include:

- QR-based methods (efficient and reasonably stable),
- SVD-based methods (most accurate but computationally more expensive),
- Greville's algorithm [**?**] (incremental column-by-column computation).

Applications

The Moore–Penrose inverse is the natural tool for solving inconsistent or underdetermined systems.

Least-squares solutions. If $Ax = b$ has no exact solution, the vector

$$\bar{x} = A^{+}b$$

minimizes the Euclidean norm $\|Ax - b\|_2$ and is the solution of smallest norm.

General solution of consistent systems. If $Ax = b$ has infinitely many solutions, then every solution can be written as

$$x = A^{+}b + (I_n - A^{+}A)\, y, \qquad y \in \mathbb{C}^n.$$

The term $(I_n - A^{+}A)y$ describes the freedom in the nullspace of A, while $A^{+}b$ is the unique solution of minimal norm.
References:

- *BenIsraelGreville A. Ben-Israel and T. N. E. Greville,* Generalized Inverses: Theory and Applications*, Springer, 2nd ed., 2003.*
- *Koecher M. Koecher,* Lineare Algebra und analytische Geometrie*, Springer, 1986.*
- *Greville T. N. E. Greville, "Some applications of the pseudoinverse of a matrix,"* SIAM Review*, 2(1), 1960.*
- *GolubVanLoan G. H. Golub and C. F. Van Loan,* Matrix Computations*, Johns Hopkins University Press, 4th ed., 2013.*

End of Excursus.

We now turn to finite-dimensional frame theory and make explicit the connection to linear algebra. Throughout, we work in $\mathbb{C}^n$ with the standard inner product. Let $\{\varphi_j\}_{j=1}^{k}$ be a frame for $\mathbb{C}^n$.

Recall that the *analysis operator*

$$C : \mathbb{C}^n \to \mathbb{C}^k, \qquad Cv = (\langle v, \varphi_j \rangle)_{j=1}^{k},$$

and the *synthesis operator*

$$D : \mathbb{C}^k \to \mathbb{C}^n, \qquad Dc = \sum_{j=1}^{k} c_j \varphi_j,$$

satisfy $D = C^*$ (conjugate transpose). Writing these as matrices:

- C is a $k \times n$ matrix whose jth row is $\overline{\varphi_j}^{\top}$,
- $D = C^*$ is an $n \times k$ matrix whose jth column is φ_j,

- the frame operator is $S = C^*C = DC$.

Since S is of the form C^*C, it is always Hermitian and positive-definite whenever $\{\varphi_j\}$ is a frame.

Proposition 5.3.16. *Let A be a $k \times n$ matrix and $\nu \in \mathbb{C}^n$. The following are equivalent:*

[(i)]There exists $C_u > 0$ such that

$$C_u\|\nu\|_2^2 \le \|A\nu\|_2^2 \qquad \forall \nu \in \mathbb{C}^n. \tag{5.19}$$

The n columns of A are linearly independent. The k rows of A form a frame for $\mathbb{C}^n$.

Proof. Let ψ_l denote the lth column of A. Then

$$A\nu = \sum_{l=1}^{n} v_l \psi_l, \qquad \nu = (v_l)_{l=1}^n.$$

(i) $\Rightarrow$ (ii): If the columns were dependent, some nonzero ν would satisfy $A\nu = 0$, contradicting (5.20).

(i) $\Rightarrow$ (iii): Writing the rows as $\{\varphi_j\}_{j=1}^k$,

$$\|A\nu\|_2^2 = \sum_{j=1}^{k} |\langle \nu, \varphi_j \rangle|^2,$$

so the lower frame bound holds; the upper bound is automatic in finite dimensions.

(iii) $\Rightarrow$ (ii): If the rows form a frame, then C is injective, hence C^*C is invertible, which implies independence of the columns. □

Example 5.3.17. *Let*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

The rows form a frame for $\mathbb{R}^2$, while the columns span a two-dimensional subspace of $\mathbb{R}^3$ and are linearly independent.

Proposition 5.3.18. *Let $\{\varphi_j\}_{j=1}^k$ be a frame in $\mathbb{C}^n$ with analysis operator C, synthesis operator $D = C^*$, and frame operator $S = C$. Then*

$$D^+ v = (\langle v, S^{-1}\varphi_j \rangle)_{j=1}^k, \qquad v \in \mathbb{C}^n.$$

Thus, the canonical dual frame is given by the columns of $C^+ = (C^*)^+$.

Proof. The canonical reconstruction formula is

$$v = \sum_{j=1}^{k} \langle v, S^{-1}\varphi_j \rangle \varphi_j = Dc, \qquad c_j = \langle v, S^{-1}\varphi_j \rangle.$$

By the Moore–Penrose theory, $c = D^+ v$ is the minimal-norm solution to $Dc = v$. This matches the coefficients above. □

Example 5.3.20 continued. The pseudoinverse of A is

$$A^+ = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}.$$

The frame operator is $A^= \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, and the dual frame consists of the columns of $(A^)^{-1}A^* = A^+$.

Frames and Stability

Suppose $f \in \mathbb{C}^n$ is transmitted using

- a basis representation: $c_B = Bf$ for an invertible $n \times n$ matrix B,
- a frame representation: $c_F = Af$ for an $n \times k$ matrix A with full rank.

If noise is added,

$$\tilde{c}_B = c_B + n_B, \qquad \tilde{c}_F = c_F + n_F,$$

then in the basis case (especially if B is unitary), the noise is fully preserved. In contrast, in the frame case one reconstructs via

$$\tilde{f} = A^+\tilde{c}_F,$$

and A^+ annihilates components in $\ker(A^*)$, thus reducing noise in directions orthogonal to the range of A. This illustrates the robustness provided by redundancy.

Time–Frequency Shifts in $\mathbb{C}^L$

Let $f \in \mathbb{C}^L$, extended periodically by $f(t+L) = f(t)$. Define:

$$T_k f(t) = f(t-k), \qquad M_\ell f(t) = e^{\frac{2\pi i \ell t}{L}} f(t), \qquad \ell \in \mathbb{Z}.$$

The operators $M_\ell T_k$ are called *time–frequency shifts*.

Sampling the time–frequency plane with steps a (time) and b (frequency) gives the lattice

$$\Lambda = a\mathbb{Z} \times b\mathbb{Z},$$

with redundancy $L/(ab)$. For $Na = Mb = L$, define the discrete Gabor system

$$g_{m,n} := M_{mb}T_{na}g, \qquad m = 0, \ldots, M-1, \;\; n = 0, \ldots, N-1.$$

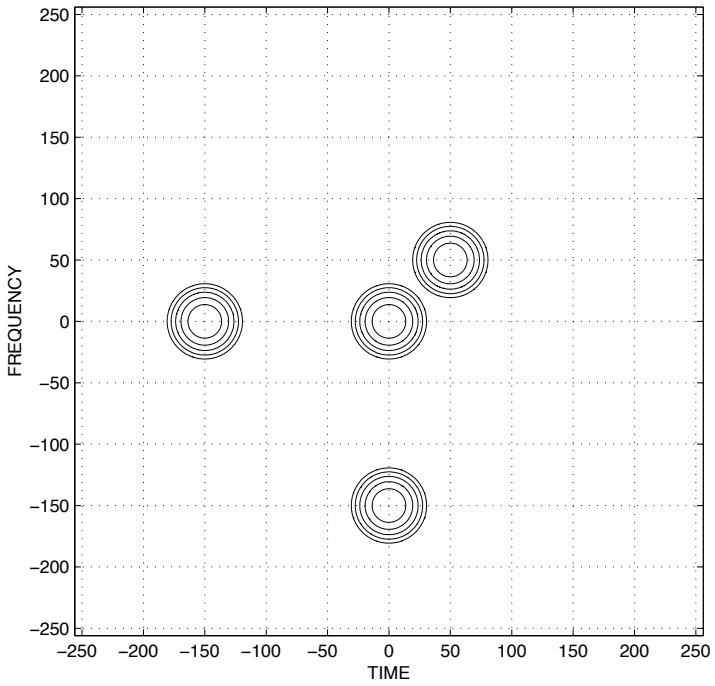


Figure 5.5: Time–frequency shifted versions of a Gaussian window.

Frame Bounds

Let $\{g_{m,n}\}_{m,n}$ be the $k = MN$ Gabor atoms in $\mathbb{C}^L$. Form the *Gabor analysis matrix*

$$G \in \mathbb{C}^{k\times L}, \qquad \mathrm{row}_{m+nM}(G) = \overline{g_{m,n}}^{\top}.$$

The system is a frame iff $k \geq L$ and $\mathrm{rank}(G) = L$. Its frame operator is $S = G^*G$, and the frame bounds are the smallest and largest eigenvalues of S.

The *canonical dual frame* is given by

$$\tilde{g}_{m,n} = S^{-1}g_{m,n},$$

and reconstruction reads

$$f = \sum_{m,n} \langle f, g_{m,n}\rangle\, \tilde{g}_{m,n}.$$

For a tight frame, $S = AI$ and hence $\tilde{g}_{m,n} = \frac{1}{A}g_{m,n}$.

We will now look more closely at the links to linear algebra.
Let $\{\varphi_j\}_{j=1}^k$ be a frame for $\mathbb{C}^n$. First observe, that the frame operator (5.9) can be written as the composition of the analysis operator $C : \mathbb{C}^n \to \mathbb{C}^k$, given by $C : v \to \{\langle v, \varphi_j\rangle\}_{j=1}^k$ and the synthesis operator $D : \mathbb{C}^k \to \mathbb{C}^n$, given by $D : \mathbf{c} \to \sum_{j=1}^k c_j\varphi_j$. In fact, D is the adjoint operator, i.e. the transposed, complex conjugate of C: $D = C'$.
We can now directly write out the matrices corresponding to these "operators" (linear maps). C, the analysis operator, is then a $k \times n$ matrix with $\overline{\varphi_{j_0}}$ in its j_0-th row. It follows immediately, from $D = C'$, that D is the $n \times k$ matrix with the vector φ_{j_0} in its j_0-th column, and $S = C' \cdot C = D \cdot C$ is then a selfadjoint map, hence a symmetric matrix, since $S' = (DC)' = (C'C)' = C'C = DC = S$.

Proposition 5.3.19. *Let A be a $k \times n$ matrix and $\nu \in \mathbb{R}^n$ a vector given by its entries v_l, $l = 1, \ldots, n$. Then the following are equivalent:*
(i) There exists a constant C_u:

$$C_u \sum_{l=1}^{n} |v_l|^2 \leq \|A\nu\|_2^2,\ \forall \nu \in \mathbb{C}^n. \tag{5.20}$$

(ii) The columns of A form a basis for their span in $\mathbb{C}^k$.
(iii) The rows of A form a frame for $\mathbb{C}^n$.

Proof. Recall that the ℓ^2-norm of $A\nu$ is given by $\|A\nu\|_2^2 = \sum_{j=1}^k |(A\nu)_l|^2$. Let $\varphi_{j_0}[l]$, $l = 1, \ldots, n$ denote the j_0th column of A', then the columns of A are given by the n vectors

$$\psi_l = \begin{pmatrix} \overline{\varphi_1[l]} \\ \overline{\varphi_2[l]} \\ \vdots \\ \overline{\varphi_k[l]} \end{pmatrix}.$$

(i)⇒(ii): (5.20) means that $C_u \sum_{l=1}^n |v_l|^2 \leq \|\sum_{l=1}^n v_l\psi_l\|^2\ \forall \nu \in \mathbb{C}^n$. Now assume that

the ψ_l, $l = 1, \ldots, n$ do not form a basis for their linear span, which means that they are linearly dependent, i.e. there exists a non-zero vector $\nu \in \mathbb{R}^n$, such that $0 = \sum_{l=1}^n v_l \psi_l$, which contradicts (5.20).
(i)⇒(iii): (5.20) can also be written as

$$C_u \sum_{l=1}^n |v_l|^2 \leq \sum_{l=1}^n |\langle \nu, \varphi_l \rangle|^2, \ \forall \nu \in \mathbb{C}^n,$$

from which the frame property of the columns follows, since the upper frame bound is automatically satisfied. (iii)⇒(ii) holds by the definition of a frame. □

Example 5.3.20. *Consider the matrix* $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}$. *Obviously, the rows constitute a frame for* $\mathbb{R}^3$, *while the columns span a two-dimensional subspace of* $\mathbb{R}^3$.

Proposition 5.3.21. *Let* $\{\varphi_j\}_{j=1}^k$ *be a frame for* $\mathbb{C}^n$ *with analysis operator* $C : \mathbb{C}^n \to \mathbb{C}^k$ *synthesis operator* $D : \mathbb{C}^k \to \mathbb{C}^n$ *and frame operator* S. *Then*

$$D^+ v = \{\langle v, S^{-1}\varphi_j \rangle\}_{j=1}^k, \ \forall v \in \mathbb{C}^n.$$

In other words, the canonical dual frame is given by the columns of the PINV of C.

Proof. We know that $v = \sum_{j=1}^k \{\langle v, S^{-1}\varphi_j \rangle \varphi_j$ and that $v = \sum_{j=1}^k c_j \varphi_j = D\mathbf{c}$ for the vector $\mathbf{c} \in \mathbb{R}^k$. From Proposition A.4.18 we have that $\mathbf{c} = D^+ v$ is the least-squares solution of this problem with the smallest norm and in Proposition 5.3.3(iii) it was shown that this solution is given by the coefficients $\{\langle v, S^{-1}\varphi_j \rangle\}_{j=1}^k = v \cdot C \cdot (C' * C)^{-1}$. □

Example 29 continued: The pseudoinverse of A in Example 5.3.20 is given by

$$A^+ = \frac{1}{3} \cdot \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}$$

. Consider the frame consisting of the rows of A, its frame operator is given by $A' * A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ and the dual frame consists of the columns of $(A' * A)^{-1} \cdot A' = A^+$.

Frames and stability

Assume that data are being transmitted using (a) a basis representation (b) a frame. That means, a sender wants to transmit $f \in \mathbb{C}^n$ by using (a) an invertible $n \times n$ matrix B and sends $c_b = B \cdot f$. Another sender uses an $n \times k$ matrix A with full rank and transmits $c_f = A \cdot f$. Now assume that the data are corrupted by some noise during transmission, i.e. $\tilde{c}_b = c_b + n_b$ and $\tilde{c}_f = c_f + n_f$, respectively. In case (a), the mapping B is unitary, so, any noise added in the transmission will be completely picked up by the receiver. On the other hand, if the frame coefficients c_f are corrupted by noise and $\tilde{c}_b$ is received then there

is justified hope, that an essential part of the noise is from the orthogonal complement of the range of A, i.e., in $N(A')$ and will, therefore, be set to 0 in the reconstruction.

Now we will introduce the finite discrete version of Gabor frames, that has become the most important one for audio signal processing and is tightly linked with the STFT.

Definition 5.3.22 (Time-frequency shifts in $\mathbb{C}^L$)**.** *Let $f \in \mathbb{C}^L$ and consider this vector extended to its L-periodic version by $f(k + lL) = f(k)$ whereever necessary.*
$T_k f(t) := f(t - k)$ is called translation operator or time shift.
$M_l f(t) := e^{\frac{2\pi i l t}{L}} f(t), l \in \mathbb{Z}$ is called modulation operator or frequency shift.
The composition of these operators, $M_l T_k$, is a time-frequency shift operator.

Generally, we are not interested in calculating $STFT(\lambda)$ in every point $\lambda \in \mathbb{Z} \times \mathbb{Z}$.[2] This would yield a redundancy of L, the length of the given signal. We down-sample in time by a and in frequency by b, so that the redundancy reduces to $\frac{L}{ab}$, and the sampling lattice is $\Lambda = a\,\mathbb{Z} \times b\,\mathbb{Z}$. a and b are referred to time- and frequency-shift parameters. The family

$$g_{m,n} := M_{mb} T_{na} g$$

for $m = 0, \ldots, M - 1$ and $n = 0, \ldots, N - 1$, where $Na = Mb = L$, is called the set of *Gabor analysis functions.*
Let us assume the $g_{m,n}$ were an orthogonal basis for a moment. In this case, the inner products $\langle f, g_{m,n} \rangle$ uniquely determine f, each representing a single and unique coefficient in the expansion

$$f = \sum_m \sum_n \langle f, g_{m,n} \rangle g_{m,n}$$

Together with Plancherel's formula $\|f\|_2^2 = \sum_m \sum_n |\langle f, g_{m,n} \rangle|^2$ this gives a beautiful split of f in pieces, preserving the signal's energy in the coefficients.

Framebounds

In the finite discrete case of $f \in \mathbb{C}^L$ a collection $\{g_{m,n}\} \in \mathbb{C}^L$ with $k = NM$ can only be a frame, if $L \leq k$ and if the matrix G, defined as the $k \times L$ matrix having $\overline{g_{m,n}}$ as its $(m + nM) - th$ row, has full rank. In this case the frame bounds are the maximum and minimum eigenvalues of S, respectively. They yield information about numerical stability. The closer the frame-bounds are, the closer the frame operator will be be to a diagonal matrix. If the frame bounds differ too much, the inversion of the frame operator is numerically unstable. The inversion of the frame operator provides reconstruction, as we saw in Proposition 5.3.3. The canonical *dual frame* $\tilde{g}_{m,n}$, which yields reconstruction as in (5.10), is given by

$$\tilde{g}_{m,n} = S^{-1} g_{m,n}$$

[2] Calculating the inner product in every point of the time-frequency lattice would yield the *full short-time Fourier transform*, a representation of redundancy L. The term "short-time Fourier transform" is often used for sampled short-time Fourier transforms as well. The spectrogram, its modulus squared, is one of the most popular time-frequency representations.

as

$$f = S^{-1}Sf = \sum \langle f, g_{m,n}\rangle S^{-1} g_{m,n} = \sum \langle f, g_{m,n}\rangle \tilde{g}_{m,n}$$

In the case of a tight frame, we have $S = C_u I$. (I denotes the identity operator) and therefore $S^{-1} = \frac{1}{C_u} I$.
The next section introduces the special case arising from applications in audio signal processing. We will see that in this case the frame operator takes a simple form.

Gabor frames for audio

Let us from now on assume that we are given a signal $f \in \mathbb{C}^L$. This signal represents a piece of music or a spoken sentence etc., which we are interested to investigate and/or modify. Modifications might aim to achieve noise reduction in old or degraded recordings. Another issue might be the extraction of certain features of the signal, for example single instrument components. Let us further assume that an engineer approaches the problem by using a Fourier transform of length l in a first step. This implies that the window used for cutting out the part of interest must have this length. Looking at the definition of the Gabor coefficients:

$$c_{m,n} = \langle f, g_{m,n}\rangle = \sum_{j=0}^{L-1} f(j)\overline{g_{m,n}(j)}$$

as an inner product, which can be interpreted as correlation between the window and the respective part of the signal, we can see that the signals f and $g_{m,n}$ must have the same length, at least theoretically. Practically, of course, as $l \ll L$, most of the "theoretical" g would be zero. As we don't tend to waste computation time on multiplying with 0, only the *effective* length of g, here l, is multiplied with the part of interest of f. This procedure implicitly introduces a frequency lattice constant $b = \frac{L}{l}$. The time constant a is related to what is often called *overlap*. If $a = \frac{l}{2}$ or $a = \frac{l}{4}$, the overlap is $\frac{l}{2}$ and $\frac{3l}{4}$, respectively. The redundancy of the representation is thus given by $red = \frac{l}{a}$, e.g., if the overlap is half the window length, we get twice as many data points as in the original signal. This is in accordance with the general case where

$$red = \frac{\frac{L}{a}\frac{L}{b}}{L} = \frac{L}{ab} = \frac{L}{a\frac{L}{l}} = \frac{l}{a}$$

Remark:

The reduction of redundancy from L in the case of the full short-time Fourier transform to a reasonable amount of redundancy in the Gabor setting ensures a balance between computability on the one hand and sufficient localisation on the other hand. The choice of a reasonable window-length and overlap common in applications corresponds roughly to such a rather balanced situation in the Gabor setting. Gabor theory, though, allows for more general choices of lattices, concerning the redundancy as well as the distribution of the lattice-points. It also provides detailed knowledge about the dependence of results on the choice of analysis parameters. This is especially decisive in the case of modification of the synthesis coefficients, which are non-unique due to the

Let us now look at the calculation of the inner products $c_{m,n} = \langle f, g_{m,n} \rangle$ more closely. They can also be written as

$$(c_{m,n})_{m=1,\ldots,M;n=1,\ldots,N} = G \cdot f$$

where G is the operator (matrix) introduced in Section 5.3.3. G consists of blocks

$$G_n, \ n = 0, \ldots, N-1$$

each corresponding to one time-position of the window g. If we define g^l as the restriction of $g \in \mathbb{C}^L$ to its non-zero part of length l, we get the following. The block G_n acts on the samples $f(na+1), \ldots, f(na+l) =: f_{na}(t)$ by taking inner products of this slice f_{na} of the signal with each of the l modulated windows

$$\begin{aligned} M_{mb} g^l(t) &= e^{\frac{-2\pi imbt}{L}} g^l(t) \\ = e^{\frac{-2\pi im\frac{L}{l}t}{L}} g^l(t) &= e^{\frac{-2\pi imt}{l}} g^l(t) \\ m = 0, \ldots, M-1 \quad &\text{and} \quad t = 0, \ldots, l-1 \end{aligned}$$

The coefficients $e^{\frac{-2\pi imt}{l}}$ are exactly the entries of the Fourier matrix $\mathcal{F}_l$ of the FFT of length l with $\hat{f} = \mathcal{F}_l f$. Therefore

$$\begin{aligned} G_n f_{na}(t) &= \mathcal{F}_l(f_{na} \cdot g^l)(t) \\ t = 0, \ldots, l-1 \quad &\text{and} \quad n = 0, \ldots, N-1 \end{aligned}$$

and the action of G_n on f_{na} corresponds to multiplying f with g, skipping zero entries and taking the Fourier transform of the non-zero part.

Remarks:

1. Although for implementation in real-life situations, the FFT-approach is always preferred, it is useful to look at the expansion from an operator point of view. Many important theoretical issues, yielding better understanding also for the applications, can be investigated more easily.

2. As mentioned before, all operators in Gabor theory generally act on the whole signal length L. In the definition of the building blocks $g_{m,n}$, the modulation operator is therefore defined as

   $$M_{mb} g(t) = e^{\frac{-2\pi imbt}{L}} g(t) \text{ for } m = 0, \ldots, N-1 \text{ and } t = 0, \ldots, L-1$$

   The blocks G_n, as opposed to the situation arising from implementation as discussed above, will not have identical entries, as the zero entries are in different positions.
   Example:
   Let $g \in \mathbb{C}^{32}$ with

   $$g(t) \begin{cases} \neq 0 & \text{for } t = 0, \ldots, 7 \\ = 0 & \text{else} \end{cases}$$

Then (by assumption $b = \frac{L}{l}$, so that $e^{\frac{-2\pi imbt}{L}} = e^{\frac{-2\pi imt}{l}}$)

$$M_{mb}g(t) = (g(0), e^{\frac{-2\pi im}{l}}g(1), e^{\frac{-2\pi i2m}{l}}g(2), \ldots, e^{\frac{-2\pi i7m}{l}}g(7), 0, \ldots, 0)$$

whereas

$$M_{mb}T_a g(t) = (0, \ldots, 0, e^{\frac{-2\pi ima}{l}}g(a), e^{\frac{-2\pi im(a+1)}{l}}g(a+1), \ldots, e^{\frac{-2\pi im(a+7)}{l}}g(a+7), 0, \ldots, 0)$$

$e^{\frac{-2\pi ima}{l}}$ is not necessarily 1, so that the blocks will differ by a phase factor.

3. The restriction that a be a divisor of l is also due to the usual choice of parameters in applications. Two common cases would be $a = \frac{l}{2}$ and $a = \frac{l}{4}$, in which cases the number of *different* kinds of blocks reduce to 2 and 4, respectively.
The difference only concerns the phase spectrum, which is usually not considered in further processing, except for reconstruction. The dual window does not depend on the phase factor in the case discussed in the theorem as will be seen below.

Mastering the frame operator - the Walnut representation revisited

Let us now come back to the central question of how to find a set of windows $\tilde{g}_{m,n}$ for reconstruction as in (5.10). If it is possible to find a window $\tilde{g}$ which is smooth and similar to the original window g especially in decaying to zero and if the rest of the dual family can be deduced in analogy to the Gabor analysis function set by time-frequency shifts, this will make reconstruction in a kind of overlap-add process easier. Infact, all the above conditions can be fulfilled. Generally, the elements of the dual frame $(\tilde{g}_{m,n})$ are generated from a single function (the dual window $\tilde{g}$), just as the original family. This follows from the fact that S and S^{-1} (the frame operator and its inverse) commute with the modulation and translation operators M_{nb} and T_{ma}, for $m = 1, \ldots, M$ and $n = 1, \ldots, N$.
The higher redundancy, the closer the shape of the dual window gets to the original window's shape. As in applications redundancy of 2, 4 or even higher are common, well localised dual windows can be found. Even more is true. The special situation in which the effective length of the window g equals or is shorter[3] than the FFT-length, the frame operator takes a surprisingly simple form.
From the definition of the frame operator

$$Sf = \sum_{m,n} \langle f, g_{m,n} \rangle g_{m,n}$$

we deduce that the single entries of S are given by

$$S_{j,k} = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} M_{mb}T_{na}g(j)\overline{M_{mb}T_{na}g(k)}$$

[3]E.g. in the case of zero padding.

Looking at the inner sum, note that $\sum_{m=0}^{M-1} e^{\frac{2\pi imb(j-k)}{L}} = 0$ if $(j-k)$ is not equal to 0 or a multiple of M. In these cases

$$\sum_{m=0}^{M-1} e^{\frac{2\pi imb(j-k)}{L}} = \sum_{m=0}^{M-1} e^{\frac{2\pi imbM}{L}} \overset{bM=L}{=} M$$

This leads to the *Walnut representation* of the frame operator for the discrete case:

$$S_{jk} = \begin{cases} M \sum_{n=0}^{N-1} T_{na} g(j) \overline{T_{na} g(k)} & \text{if } |j-k| \bmod M = 0 \\ 0 & \text{else} \end{cases} \tag{5.21}$$

There will obviously be non-zero entries in the diagonal, $j = k$, but as $M = l$, i.e. the window-length, $j = k$ is infact the only case for which $|j-k| \bmod M = 0$ holds and $g(j)$ and $g(k)$ both have non-zeros values. Therefore, the frame operator is diagonal and the dual window $\tilde{g}$ is calculated as

$$\tilde{g}(t) = g/(M \sum_{n=0}^{N-1} T_{na} |g(t)|^2)$$

5.3.4 Wavelet Bases, Frames, and Multiresolution

In the previous sections we studied time–frequency representations based on the Short-Time Fourier Transform and Gabor frames. These systems describe signals by translating and modulating a fixed window function. However, such representations have a *uniform resolution* in time and frequency, which limits their ability to capture phenomena whose characteristic frequencies change over time. Wavelet analysis overcomes this restriction by replacing fixed-frequency modulation with *scaling*: instead of sliding a window of constant width, we stretch and compress a prototype function (the *mother wavelet*) to achieve fine temporal resolution at high frequencies and coarse resolution at low frequencies. This approach leads naturally to the concept of **multiresolution analysis** (MRA), which provides the mathematical foundation for constructing orthogonal and biorthogonal wavelet bases and filter-bank implementations.

Wavelets and Multiresolution Analysis

Wavelets provide localized time–frequency representations of signals. In contrast to the Fourier basis, where all functions are globally supported, wavelets are well localized in both time and frequency. A *mother wavelet* $\psi \in L^2(\mathbb{R})$ generates a family of functions by translations and dyadic dilations:

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), \qquad j, k \in \mathbb{Z}.$$

- The index j controls the *scale* (or resolution): large j corresponds to fine detail.
- The index k controls the *position* of the wavelet.

Continuous Wavelet Transform. Given $f \in L^2(\mathbb{R})$, its continuous wavelet transform (CWT) is

$$W_f(a,b) = \int_{\mathbb{R}} f(t)\, \psi^*_{a,b}(t)\, dt, \qquad \psi_{a,b}(t) = \frac{1}{\sqrt{a}}\, \psi\!\left(\frac{t-b}{a}\right),$$

which measures the similarity of f to dilated (a) and translated (b) copies of ψ. The CWT provides a redundant but highly interpretable time–scale representation. For practical computation, we use its discrete version, obtained from the theory of *multiresolution analysis* (MRA).

Multiresolution Analysis (MRA)

An MRA is a sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ satisfying:

1. $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$,
2. $\bigcup_{j\in\mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R})$,
3. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,
4. $f(x) \in V_j \Rightarrow f(2x) \in V_{j+1}$,
5. There exists a *scaling function* $\varphi \in L^2(\mathbb{R})$ such that $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_0.

Each space V_j represents the approximation of signals at resolution 2^j and is spanned by $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$. The refinement relation

$$\varphi(x) = \sum_k h_k\, \varphi(2x-k),$$

introduces the *scaling coefficients* h_k, which play the role of a discrete low–pass filter.

The details between successive spaces are captured by complementary subspaces W_j defined via

$$V_{j+1} = V_j \oplus W_j.$$

Each W_j is spanned by wavelets $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, where the wavelet satisfies

$$\psi(x) = \sum_k g_k\, \varphi(2x-k), \qquad g_k = (-1)^k h_{1-k}.$$

The sequence (g_k) acts as a discrete high–pass filter.

Remark 5.3.23 (Filter Bank Interpretation)**.** *The relations above form the basis of the* two–channel filter bank*: a signal is decomposed by convolution with (h_k) (approximation) and (g_k) (detail), followed by downsampling by 2. Reconstruction uses the inverse filters and upsampling. This discrete implementation is known as the* ***Discrete Wavelet Transform (DWT)****.*

Multiresolution Approximation of a Signal

For $f \in L^2(\mathbb{R})$, its approximation at level j is the orthogonal projection onto V_j:

$$P_j f(x) = \sum_k c_k^j \, \varphi_{j,k}(x), \qquad c_k^j = \langle f, \varphi_{j,k} \rangle.$$

The detail at level j is the projection onto W_j:

$$D_j f(x) = \sum_k d_k^j \, \psi_{j,k}(x), \qquad d_k^j = \langle f, \psi_{j,k} \rangle.$$

Thus

$$f(x) = P_J f(x) + \sum_{j \geq J} D_j f(x),$$

with coarse approximation $P_J f$ and progressively finer details $D_j f$.

Remark 5.3.24 (Discrete Algorithm)**.** *In practice, the coefficients $\{c_k^j, d_k^j\}$ are computed by convolution and subsampling:*

$$c_k^{j+1} = \sum_n h_n \, c_{2k-n}^j, \qquad d_k^{j+1} = \sum_n g_n \, c_{2k-n}^j.$$

Iterating these equations yields a multilevel decomposition of f.

Two-Dimensional Wavelet Transform

For image processing, separable 2-D wavelets are constructed from tensor products of 1-D scaling and wavelet functions:

$$\begin{aligned} \psi^{(1)}(x,y) &= \varphi(x)\psi(y), \\ \psi^{(2)}(x,y) &= \psi(x)\varphi(y), \\ \psi^{(3)}(x,y) &= \psi(x)\psi(y). \end{aligned}$$

A single decomposition step produces four subbands:

LL (approximation), LH (horizontal detail), HL (vertical detail), HH (diagonal detail).

Repeated application to the LL band yields the familiar wavelet pyramid used in image analysis and compression (e.g. JPEG 2000).

Example: Decomposition and Reconstruction

Let $f(x)$ be a piecewise smooth signal. Its decomposition at level J consists of approximation and detail coefficients:

$$f(x) = \sum_k c_k^J \, \varphi_{J,k}(x) + \sum_{j=J}^{J_{\max}} \sum_k d_k^j \, \psi_{j,k}(x).$$

Perfect reconstruction is achieved by inverting the filter bank: upsampling and convolution with synthesis filters $\tilde{h}_k, \tilde{g}_k$ followed by summation of the approximation and detail branches.

Remark 5.3.25 (Applications of Multiresolution Analysis).

- ***Signal Denoising:*** *Remove noise from signals by thresholding detail coefficients at fine scales.*
- ***Image Compression:*** *Efficiently represent images by retaining significant coefficients in the wavelet domain.*
- ***Feature Extraction:*** *Identify important features in data by analyzing variations at different scales.*

Appendix A

Appendix:

A.1 A primer on Lebesgue spaces

We refer to [1] for proofs of this section. Let $(X, \Sigma(X), \mu)$ be a measure space. For $1 \leq p < \infty$, define

$$\mathcal{L}^p(X,\mu) := \{f : X \to \mathbb{C} \mid f \text{ measurable}, \int_X |f(x)|^p \mathrm{d}\mu(x) < \infty\}.$$

For $f \in \mathcal{L}_p(X,\mu)$, let

$$\|f\|_{L^p} := \left(\int_X |f(x)|^p \mathrm{d}\mu(x)\right)^{\frac{1}{p}}.$$

The equivalence relation $f \sim g$ if and only if $\|f-g\|_{L^p} = 0$ leads to the Banach space

$$L^p(X,\mu) := \mathcal{L}^p(X,\mu)/\sim .$$

Exercise: Define $L^\infty(X,\mu)$.
If the choice of μ is clear from the context, we simply write $L^p(X)$.

Example A.1.1. *The sequences of functions*

$$f_n = 1_{[n,n+1]}, \qquad g_n = \frac{1}{n}1_{[0,n]}, \qquad h_n = n1_{[\frac{1}{n},\frac{2}{n}]}$$

converge pointwise to 0 *but*

$$1 = \|f_n\|_{L^1} = \|g_n\|_{L^1} = \|h_n\|_{L^1},$$

so that they do not converge to 0 *in* $L^1(\mathbb{R})$. *Note that* $g_n \to 0$ *even uniformly.*

Proposition A.1.2 (Dominated convergence theorem)**.** *Let* $L^1(X,\mu) \ni f_k \to f : X \to \mathbb{C}$ *pointwise a.e.*

$$\exists g \in L^1(X,\mu) : |f_k| \leq g \text{ a.e.} \qquad \Rightarrow \qquad f \in L^1(X,\mu), \quad \int f_k \to \int f.$$

Figure A.1: $u_\epsilon(x) = \frac{1}{\epsilon}\mathrm{e}^{-\pi\left(\frac{x}{\epsilon}\right)^2}$, for $|x| \leq 2$ and $\epsilon = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$.

Proposition A.1.3 (Young's convolutional inequality)**.** *Let $1 \leq p < \infty$. For $f \in L^1(\mathbb{R}^d)$ and $g \in L_p(\mathbb{R}^d)$,*

$$(f * g)(\cdot) := \int f(y)g(\cdot - y)\mathrm{d}y \in L^p(\mathbb{R}^d)$$

is well-defined a.e. and it holds

$$\|f * g\|_{L^p} \leq \|f\|_{L^1}\|g\|_{L^p}. \tag{A.1}$$

Definition A.1.4 (Approximate identity)**.** *An approximate identity is a family $(u_\epsilon)_{\epsilon>0} \subset L^1(\mathbb{R}^d)$ such that it holds:*

(a) $\exists c > 0$ *such that* $\|u_\epsilon\|_{L^1} \leq c$, $\forall \epsilon > 0$,

(b) $\int u_\epsilon = 1$, $\forall \epsilon > 0$,

(c) for any neighborhood U of 0,

$$\int_{X\setminus U} |u_\epsilon| \xrightarrow{\epsilon\to 0} 0.$$

Example A.1.5. *If $u \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} u(x)\mathrm{d}x = 1$, then*

$$u_\epsilon(x) := \epsilon^{-d}u(\epsilon^{-1}x) \tag{A.2}$$

is an approximate identity, cf. Figure A.1 for $d = 1$ and $u(x) = \mathrm{e}^{-\pi x^2}$. Parts (a,b) are transformation identities and, Part (c) is an exercise.

Proposition A.1.6 (Approximate Identity)**.** *Let $1 \leq p < \infty$ and $f \in L^p(\mathbb{R}^d)$. If $(u_\epsilon)_{\epsilon>0}$ is an approximate identity, then*

$$u_\epsilon * f \xrightarrow{\epsilon\to 0} f \quad \text{in } L^p(\mathbb{R}^d).$$

Proof. Let $p = 1$. We have

$$\begin{aligned}
\|f - u_\epsilon * f\|_{L^1} &= \int |f(x) - u_\epsilon * f(x)|\mathrm{d}\lambda(x) \\
&= \int \left| f(x)\int u_\epsilon(y)\mathrm{d}\lambda(y) - \int u_\epsilon(y)f(y^{-1}x)\mathrm{d}\lambda(y)\right| \mathrm{d}\lambda(x) \\
&\leq \int\int \left|f(x) - f(y^{-1}x)\right| |u_\epsilon(y)|\,\mathrm{d}\lambda(x)\mathrm{d}\lambda(y).
\end{aligned}$$

Splitting $X = U \cup (X \setminus U)$ with U as in Lemma **??** leads to

$$\|f - u_\epsilon * f\|_{L^1} \leq \int_U \int \left|f(x) - f(y^{-1}x)\right| \mathrm{d}\lambda(x)\,|u_\epsilon(y)|\,\mathrm{d}\lambda(y) + \int_{X\setminus U} \int \left|f(x) - f(y^{-1}x)\right| \mathrm{d}\lambda(x)\,|u_\epsilon(y)|\,\mathrm{d}\lambda$$

The case $1 < p < \infty$ is analogous. □

Figure A.2: $f * u_\epsilon$ for $f = -1_{[-1,0)} + 1_{[0.1]}$ and $u_\epsilon(x) = \frac{1}{\epsilon} \mathrm{e}^{-\pi\left(\frac{x}{\epsilon}\right)^2}$, for $|x| \leq 2$ and $\epsilon = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$.

We know that $\mathcal{C}_c(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$, for $1 \leq p < \infty$.

Corollary A.1.7. *The space $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$, for $1 \leq p < \infty$.*

A.2 Dirac Impulse

The **Dirac impulse**, also known as the **Dirac delta function**, is a generalized function used in signal processing, physics, and engineering. It represents an infinitely narrow and high peak at a single point, with an integral of 1.

Continuous Version (Dirac Delta Function)

In the continuous domain, the **Dirac delta function** $\delta(t)$ is defined by the following properties:

1. **It is zero everywhere except at $t = 0$:**

$$\delta(t) = 0 \quad \text{for} \quad t \neq 0$$

2. **Its integral over the entire real line is equal to 1:**

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1$$

3. **Sifting property (for any test function $f(t)$):**

$$\int_{-\infty}^{\infty} f(t)\delta(t - t_0)\, dt = f(t_0)$$

This property essentially "picks out" the value of the function $f(t)$ at $t = t_0$.

Mathematically, the Dirac delta function can also be approximated as:

$$\delta(t) = \lim_{\epsilon \to 0} \frac{1}{\sqrt{\pi\epsilon}} e^{-t^2/\epsilon}$$

This expression represents the Dirac delta as a Gaussian function that becomes narrower and higher as $\epsilon \to 0$.

Discrete Version (Dirac Delta Sequence)

In the discrete case, the Dirac impulse is represented as a sequence of Kronecker delta functions $\delta[n]$, defined as:

$$\delta[n] = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0 \end{cases}$$

The discrete delta function has the following properties:
1. **It is zero for all values except at $n = 0$:**

$$\delta[n] = 0 \quad \text{for} \quad n \neq 0$$

2. **Its summation over all integers is equal to 1:**

$$\sum_{n=-\infty}^{\infty} \delta[n] = 1$$

3. **Sifting property (for any sequence $x[n]$):**

$$\sum_{n=-\infty}^{\infty} x[n]\delta[n - n_0] = x[n_0]$$

This property "picks out" the value of $x[n]$ at $n = n_0$.

A.3 Version II of pointwise convergence of Fourier series in $\mathbb{R}^d$: Dirichlet

Let $I := \left[-\frac{1}{2}, \frac{1}{2}\right]^d$, so that the inner product in $L^2(\mathbb{T}^d)$ is

$$\langle f, g\rangle_{L^2(\mathbb{T}^d)} = \int_{I^d} f(x)\overline{g(x)}\mathrm{d}x.$$

For $k \in \mathbb{Z}^d$, we consider

$$\mathrm{e}_k(x) := \mathrm{e}^{2\pi \mathrm{i}\langle k,x\rangle}.$$

It is clear that $\{\mathrm{e}_k : k \in \mathbb{Z}^d\}$ is an orthonormal basis for $L^2(\mathbb{T}^d)$. Let

$$\Pi_N := \operatorname{span}\{\mathrm{e}_k : \|k\|_\infty \leq N\}$$

denote the trigonometric polynomials of degree $N \in \mathbb{N}$. The orthogonal projection onto Π_N is

$$S_N : L^2(\mathbb{T}^d) \to \Pi_t, \qquad f \mapsto \sum_{\substack{k\in\mathbb{Z}^d \\ \|k\|_\infty \leq N}} \langle f, \mathrm{e}_k\rangle_{L^2(\mathbb{T}^d)}\mathrm{e}_k, \tag{A.3}$$

so that $S_N f \xrightarrow{N\to\infty} f$ in $L^2(\mathbb{T}^d)$.

The *Fourier coefficients* of $f \in L^1(\mathbb{T}^d)$ are

$$\hat{f}(k) := \int_{I^d} f(x)\mathrm{e}^{-2\pi \mathrm{i}\langle k,x\rangle}\mathrm{d}x, \quad k \in \mathbb{Z}^d\,.$$

Lemma A.3.1 (Riemann-Lebesgue for $\mathbb{T}^d$)**.** *If $f \in L^1(\mathbb{T}^d)$, then $\hat{f}_k \xrightarrow{\|k\|\to\infty} 0$.*

Proof. Since $\mathbb{T}^d$ is compact, we have $\mathcal{C}(\mathbb{T}^d) \subset L^2(\mathbb{T}^d) \subset L^1(\mathbb{T}^d)$ are dense. For $f \in L^2(\mathbb{T}^d)$, we have $\left(\hat{f}_k\right)_{k\in\mathbb{Z}^d} \in \ell^2(\mathbb{Z}^d)$, so that $\hat{f}_k \xrightarrow{\|k\|\to\infty} 0$.

Exercise: complete the proof for $f \in L^1(\mathbb{T}^d)$. □

We have already used that $f \in L^2(\mathbb{T}^d)$ leads to $\hat{f}_k = \langle f, \mathrm{e}_k\rangle_{L^2(\mathbb{T}^d)}$. Note that S_N in (A.3) can be extended to

$$S_N : L^1(\mathbb{T}^d) \to \Pi_t, \qquad f \mapsto \sum_{\substack{k\in\mathbb{Z}^d \\ \|k\|_\infty \le t}} \hat{f}_k \mathrm{e}_k.$$

The Dirichlet kernel allows to write the partial Fourier sums as convolutions in a similar manner as the Fejer kernel.

Definition A.3.2 (Dirichlet kernel)**.** *The function*

$$D_t(z) := \sum_{\|k\|_\infty \le t} \mathrm{e}_k(z), \qquad z \in \mathbb{R}^d,$$

is called the Dirichlet kernel*.*

Proposition A.3.3. *For $0 < t \in \mathbb{N}$, we have*

$$(S_N f)(x) = \int_{I^d} f(y) D_t(x-y)\mathrm{d}y$$

and the Dirichlet kernel satisfies

$$D_t(z) = \prod_{i=1}^{d} \left(\frac{\mathrm{e}_{t+1}(z_i) - \mathrm{e}_t(z_i)}{\mathrm{e}_1(z_i) - 1} \right) \tag{A.4}$$

Proof. Exercise. (hint: prove (A.4) first for $d = 1$ via geometric progression and observe the tensor structure) □

Proposition A.3.4 (Pointwise Convergence II: continuously differentiable functions.)**.** *If $f \in \mathcal{C}^1(\mathbb{T})$, then $S_N f \xrightarrow{N\to\infty} f$ pointwise.*

Proof. Let $x \in \mathbb{R}$ be fixed. Since $\mathrm{e}_1 = 1$, the orthogonality implies $\int_I \overline{D_t(z)}\mathrm{d}z = 1$, so that we derive with $\overline{D_t(z)} = D_t(-z)$,

$$S_N f(x) - f(x) = \int_I f(y) D_t(x-y)\mathrm{d}y - \int_I f(x) D_t(-z)\mathrm{d}z$$

Substitution and periodicity lead to

$$\begin{aligned}
S_N f(x) - f(x) &= -\int_{x+\frac{1}{2}}^{x-\frac{1}{2}} f(x-z) D_t(z) \mathrm{d}z - \int_I f(x) D_t(-z) \mathrm{d}z \\
&= \int_{-x-\frac{1}{2}}^{-x+\frac{1}{2}} f(x+z) D_t(-z) \mathrm{d}z - \int_I f(x) D_t(-z) \mathrm{d}z \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x+z) D_t(-z) \mathrm{d}z - \int_I f(x) D_t(-z) \mathrm{d}z \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(f(x+z) - f(x)\right) D_t(-z) \mathrm{d}z \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} g(z) \left(\mathrm{e}_{-(t+1)}(z) - \mathrm{e}_{-t}(z)\right) \mathrm{d}z,
\end{aligned}$$

where we have used (A.4) and

$$g(z) := \frac{f(x+z) - f(x)}{\mathrm{e}_1(z) - 1}.$$

We have $g \in L^1(\mathbb{T})$ because l'Hospital's rule yields

$$\lim_{0 \neq z \to 0} g(z) = \lim_{0 \neq z \to 0} \frac{f(x+z) - f(x)}{z} \cdot \underbrace{\frac{z}{\mathrm{e}_1(z) - 1}}_{\to \frac{1}{2\pi i}} = \frac{f'(x)}{2\pi i}.$$

The Riemann-Lebesgue Lemma A.3.1 leads to

$$S_N f(x) - f(x) = \hat{g}_{t+1} - \hat{g}_t \to 0. \qquad \square$$

A.3.1 Approximation of Sobolev functions

Definition A.3.5 (Sobolev space)**.** *For $s > 0$, the Sobolev space $H^s(\mathbb{T}^d)$ is*

$$H^s(\mathbb{T}^d) := \{f \in L^2(\mathbb{T}^d) : \sum_{k \in \mathbb{Z}^d} \left(1 + \|k\|^2\right)^s \left|\hat{f}_k\right|^2 < \infty\}$$

with inner product

$$\langle f, g \rangle_{H^s} := \sum_{k \in \mathbb{Z}^d} \left(1 + \|k\|^2\right)^s \hat{f}_k \overline{\hat{g}_k}.$$

Proposition A.3.6. *For $s, t > 0$ and $f \in H^s(\mathbb{T}^d)$, we have*

$$\|S_N f - f\|_{L^2} \leq t^{-s} \|f\|_{H^s}.$$

Proof. Eventually, the Hölder inequality yields

$$\begin{aligned}\|S_N f - f\|_{L^2}^2 &= \left\| \sum_{\|k\|_\infty > t} \hat{f}_k \mathrm{e}_k \right\|_{L^2}^2 = \sum_{\|k\|_\infty > t} \left| \hat{f}_k \right|^2 \\ &= \sum_{\|k\|_\infty > t} \left(1 + \|k\|^2\right)^{-s} \left(1 + \|k\|^2\right)^s \left| \hat{f}_k \right|^2 \\ &\leq (1+t^2)^{-s} \|f\|_{H^s}^2. \end{aligned}$$

□

Lemma A.3.7. *For $\alpha \in \mathbb{N}^d$ and $f \in \mathcal{C}^{|\alpha|}(\mathbb{T}^d)$, we have $\widehat{(\partial^\alpha f)}_k = (2\pi \mathrm{i} k)^\alpha \, \hat{f}_k$.*

Proof. For $d = 1 = \alpha$, integration by parts yields

$$\begin{aligned}\widehat{\partial f}_k &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \partial f(x) \mathrm{e}^{-2\pi \mathrm{i} k x} \mathrm{d}x \\ &= \underbrace{\left[f(x) \mathrm{e}^{-2\pi \mathrm{i} k x} \right]_{-\frac{1}{2}}^{\frac{1}{2}}}_{0} - \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x) (-2\pi \mathrm{i} k) \mathrm{e}^{-2\pi \mathrm{i} k x} \mathrm{d}x \\ &= 2\pi \mathrm{i} k \hat{f}_k. \end{aligned}$$

□

Lemma A.3.7 enables us to define $\partial^\alpha f$ via its Fourier coefficients. For $\alpha \in \mathbb{N}^d$ with $s \geq |\alpha| \geq 1$ and $f \in H^s(\mathbb{T}^d)$, we observe

$$\begin{aligned}\|\partial^\alpha f\|_{H^{s-|\alpha|}} &= \sum_{k \in \mathbb{Z}^d} (1 + \|k\|^2)^{s-|\alpha|} |2\pi k|^{2\alpha} |\hat{f}_k|^2 \\ &= (2\pi)^{|\alpha|} \sum_{k \in \mathbb{Z}^d \setminus \{0\}} (1 + \|k\|^2)^s |\hat{f}_k|^2 \frac{k^{2\alpha}}{(1 + \|k\|^2)^{|\alpha|}} \\ &\leq (2\pi)^{|\alpha|} \sum_{k \in \mathbb{Z}^d \setminus \{0\}} (1 + \|k\|^2)^s |\hat{f}_k|^2 \underbrace{\frac{\|k\|_\infty^{2|\alpha|}}{\|k\|^{2|\alpha|}}}_{\leq 1} \\ &\leq (2\pi)^{|\alpha|} \|f\|_{H^s}^2. \end{aligned}$$

Let us summarize our computations:

Proposition A.3.8. *For $\alpha \in \mathbb{N}^d$ with $s \geq |\alpha| \geq 1$, we have $\partial^\alpha : H^s(\mathbb{T}^d) \to H^{s-|\alpha|}(\mathbb{T}^d)$, where*

$$\widehat{(\partial^\alpha f)}_k = (2\pi \mathrm{i} k)^\alpha \, \hat{f}_k.$$

For $s \geq 2$ and $f \in H^{s-2}(\mathbb{T}^d)$, consider Poisson's equation

$$-\Delta u = f \tag{A.5}$$

in $L^2(\mathbb{T}^d)$. By applying $\Delta = \sum_{i=1}^d \partial_i^2$, the Fourier coefficients must satisfy

$$4\pi^2\|k\|^2\hat{u}_k = \hat{f}_k, \quad k \in \mathbb{Z}^d\,.$$

Hence, we may define u by

$$\hat{u}_k = \begin{cases} \frac{1}{4\pi^2\|k\|^2}\hat{f}_k, & k \in \mathbb{Z}^d \setminus \{0\}, \\ c, & k = 0, \end{cases} \tag{A.6}$$

where $c \in \mathbb{C}$ is an arbitrary constant. It yields $\int_{I^d} u(x)\mathrm{d}x = c$ and we indeed observe $u \in H^s(\mathbb{T}^d)$.

Corollary A.3.9. *For $s \geq 2$ and $f \in H^{s-2}(\mathbb{T}^d)$, we may define $u^t \in \Pi_t$ by*

$$(\widehat{u^t})_k := \begin{cases} \frac{1}{4\pi^2\|k\|^2}\hat{f}_k, & 1 \leq \|k\|_\infty \leq t, \\ 0, & \textit{otherwise.} \end{cases}$$

If $u \in H^s(\mathbb{T}^d)$ solves Poisson's equation (A.5) *with $c = 0$ in* (A.6)*, then we have*

$$\left\|u^t - u\right\|_{L^2} \leq t^{-s}\frac{\|f\|_{H^{s-2}}}{2\pi^2}.$$

Proof. Theorem A.3.6 implies

$$\left\|u^t - u\right\|_{L^2} \leq t^{-s}\|u\|_{H^s}.$$

We further estimate

$$\begin{aligned}
\|u\|_{H^s}^2 &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} (1+\|k\|^2)^s \left|\hat{u}_k\right|^2 \\
&= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} (1+\|k\|^2)^{s-2} \left|\hat{f}_k\right|^2 \frac{(1+\|k\|^2)^2}{(4\pi^2\|k\|^2)^2} \\
&= \frac{1}{16\pi^4}\sum_{k\in\mathbb{Z}^d\setminus\{0\}} (1+\|k\|^2)^{s-2} \left|\hat{f}_k\right|^2 \left(\frac{1+\|k\|^2}{\|k\|^2}\right)^2 \\
&\leq \frac{1}{4\pi^4}\sum_{k\in\mathbb{Z}^d\setminus\{0\}} (1+\|k\|^2)^{s-2} \left|\hat{f}_k\right|^2 \underbrace{\left(\frac{1+\|k\|^2}{2\|k\|^2}\right)^2}_{\leq 1} \\
&\leq \frac{1}{4\pi^4}\|f\|_{H^{s-2}}^2.
\end{aligned}$$

□

A.4 Approximation

This appendix recalls or summarizes some results from linear algebra which can be helpful to understand concepts in frame theory.

A.4.1 Least squares method

The method of least squares is a standard approach to the approximate solution of overdetermined systems, i.e., sets of equations in which there are more equations than unknowns. "Least squares" means that the overall solution minimizes the sum of the squares of the errors made in solving every single equation.

We first recall the following system of linear equations:

$$Ax = b, \tag{A.7}$$

where A is a $k \times n$ matrix, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^k$. A is the matrix of a linear map from $\mathbb{R}^n$ to $\mathbb{R}^k$ with respect to some basis. The system of linear equations (A.7) can be solved with a unique solution, if and only if $k = n$ and A is invertible. In this case, the rank of the matrix is n and so is the dimension of its range (column space $C(A)$), while the dimension of its kernel (null-space $N(A)$) is 0.[1] This is the almost trivial, and in this section, we will be more interested in the cases when $k \neq n$ and in particular, when $k > n$, in which case the problem (A.7) is over-determined and we will usually only be able to find approximate solutions. Talking in the context of the subspaces mentioned so far: the range of our matrix A is a proper subspace of $\mathbb{R}^k$.
On the other hand, if $k < n$, we have less equations than parameters and usually (A.7) is thus under determined. In this case, the rank of the matrix, and hence the dimension of its range, is at most k, so that there must be a non-trivial null-space $N(A)$. i.e. there are $y \neq 0 \in \mathbb{R}^n$, such that $Ay = 0$. Assume now, that x solves (A.7), then, for any scalar c and $x + cy$ we have

$$A(x + cy) = Ax + cAy = Ax = b$$

and hence there are always infinitely many solutions. Use Figure A.3 to get a good orientation in the four subspaces involved in linear maps[2].

Projection onto a one-dimensional subspace

Let us assume that we want to project a vector $b \in \mathbb{R}^n$ onto a vector $a \in \mathbb{R}^n$. The corresponding $n \times n$- matrix P_a is then given by $a \cdot a^T$, since

$$Pb = a \cdot \frac{\langle b, a, \rangle}{\|a\|^2} = a \cdot \frac{a^T \cdot b}{\|a\|^2}.$$

Example A.4.1. *We want to project the vector*

$$b = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 2 \\ -3 \end{pmatrix} \text{ onto (a) } a_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \text{ (b) } a_2 = \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

[1] We will encounter two more of what G. Strang calls "Four fundamental Subspaces"', cf. *http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/video-lectures/lecture-10-the-four-fundamental-subspaces/*.

[2] In Figure A.3, m is used instead of k, since this plot is taken from a book of G.Strang.

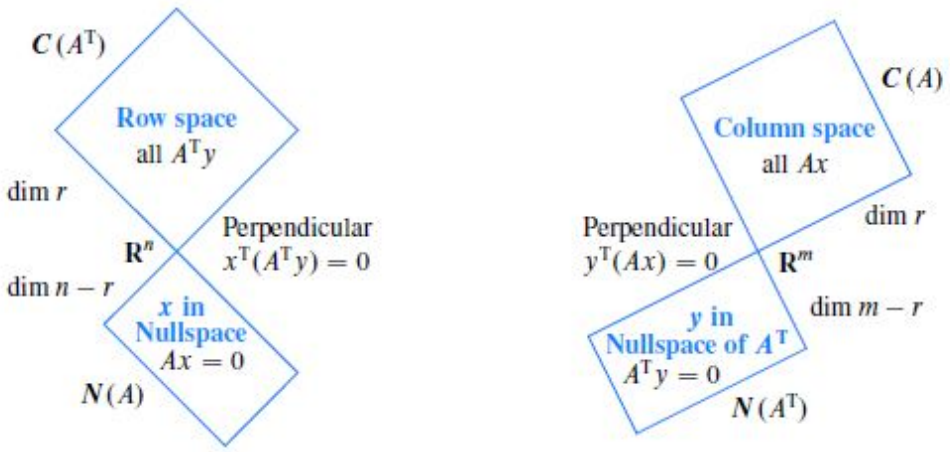


Figure A.3: Dimensions and orthogonality for any m by n matrix A of rank r.

In the first case, $P_{a_1} = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ *and the resulting projection is the* 0*-vector. This is the obvious result, if we try to approximate a vector (here* b*) by a vector that is orthogonal (perpendicular) to* b*. On the other hand,* $P_{a_2} = \frac{1}{4}\begin{pmatrix} 1 & i & -1 & -i \\ -i & 1 & i & -1 \\ -1 & -i & 1 & i \\ i & -1 & -i & 1 \end{pmatrix}$ *and* $P_{a_2} \cdot b = \begin{pmatrix} -i \\ -1 \\ i \\ 1 \end{pmatrix}$.

Note that the projection onto a vector b (a one-dimensional subspace) is equivalent to the approximation by b. We next generalize this to the approximation by several vectors.

Projection onto a subspace of higher dimension

We now assume that we are given n vectors $a_j \in \mathbb{R}^k$, and we consider the projection of a vector $b \in \mathbb{R}^k$ onto the subspace $\mathcal{A} = span(a_1, \ldots, a_n)$, in other words, we want to approximate b by an arbitrary linear combination of the given vectors. Since we assume $k > n$, we can also assume that the n vectors are linearly independent. If k is significantly larger than n we cannot expect to find an exact solution, so we will try to minimize the following error:

$$e(\hat{x}) = \|\hat{x}_1 a_1 + \ldots + \hat{x}_n a_n - b\|_2^2,$$

so we are looking for the coefficient vector $\hat{x} \in \mathbb{R}^n$ such that $e(\hat{x}) = min_{\hat{x}\in\mathbb{R}^n}\|\hat{x}_1 a_1 + \ldots + \hat{x}_n a_n - b\|_2^2$, or $\hat{x} = \operatorname{argmin} e(x)$.

Since the error $e = b - A\hat{x}$, where A is the $k \times n$ matrix with the a_j as n columns, must be orthogonal to $a_1, \ldots, a_n$, we obtain the new set of linear equations

$$A^T(b - A\hat{x}) = 0 \Longleftrightarrow A^T \cdot A\hat{x} = A^T b$$

and since the vectors a_j where supposed to be linearly independent, the $n \times n$ matrix $A^T \cdot A$ is invertible.
Hence, the coefficients of the best approximation are given by

$$\hat{x} = (A^T \cdot A)^{-1} \cdot A^T b$$

and the best approximation, or orthogonal projection onto $\mathcal{A}$ is then

$$p = P_{\mathcal{A}} b = \underbrace{A \cdot (A^T \cdot A)^{-1} \cdot A^T}_{P_{\mathcal{A}}} b.$$

Example A.4.2. *Determine the projection of* $b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$ *onto* $= span\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right)$.

Result: $\hat{x} = (5, -3)$, $p = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}$.

Proposition A.4.3. *A matrix $A^T A$ is invertible if and only if the columns of A are linearly independent.*

Proof. We show that $ker(A^T A) = ker(A)$, which is equivalent to the statement of the theorem (argue, why!)
(I) Let $x \in ker(A)$, then $Ax = 0$, hence, by linearity, $A^T Ax = A^T 0 = 0$.
(II) Let $x \in ker(A^T A)$, i.e. $A^T Ax = 0$, then $x^T A^T Ax = 0$, hence $(Ax)^T(Ax) = ||Ax||^2 = 0$ and therefore $Ax = 0$ and x is also in the kernel of A. □

Corollary A.4.4. *Let A be an $k \times n$ matrix. If $n > k$, then $A^T A$ cannot be invertible.*

This follows immediately from the proposition above. Argue, why!

An important application: data fitting with polynomials

In engineering and many other applications, it is often necessary to fit a line to a set of data. A line is a a first degree polynomial:

$$s = ct + d$$

in other words, a line with slope c. Our goal is to determine the coefficients c and d of the polynomial that lead to the "best fit" of a line to the data. Now assume that we are

given data $(t_i, s_i), i = 1, \ldots, k$, i.e. k measurements s_i at time points t_i. We then want to minimize the error

$$e(c, d,) = \sum_{i=1}^{k} |s_i - (ct_i + d)|^2 ,$$

and, setting

$$\hat{x} = \begin{pmatrix} d \\ c \end{pmatrix} ; A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_k \end{pmatrix} ; s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{pmatrix}$$

we can obtain the solution by solving the overdetermined linear system $A\hat{x} = s$.
Why is the best solution to this system of linear equations also the Žminimizer of $e(c, d)$?

Example A.4.5. *Find the line that best approximates the three measurements* $(0, 6); (1, 0); (2, 0)$. *Result:* $s = -3t + 5$.

The fitting process can be generalized to determine the coefficients of the Nth-order polynomial that best fits $N + 1$ or more data points. The determination of the coefficients can be done in MATLAB by the function polyfit.
For example, for $N = 3$:

$$s = c_0 + c_1 t + c_2 t^2$$

.

This will exactly fit a simple curve to three points, as we see in the following example.

Example A.4.6. *Fit a polynomial of degree* 2 *to the data points from Example A.4.5. Now we use*

$$\hat{x} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} ; A = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \end{pmatrix} ; s = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}$$

and obtain the result $s = 6 - 9t + 3t^2$, *i.e., the coefficient-vector is given by* $\hat{x} = \begin{pmatrix} 6 \\ -9 \\ 3 \end{pmatrix}$.
In this case, the "projection"-matrix onto the range of A (CS(A), SR(A)) is the identity, since the columns of A span all of $\mathbb{R}^3$

Remark A.4.7. *In the lecture we discussed that there are three ways to interpret the method of least squares: geometrically (looking for the point in the hyperplane spanned by the columns of A, that is closest to b), algebraically, by removing the part of b, that is orthogonal to the range of A (i.e. to all the columns of A), and solve* $Ax = p$ *instead of* $Ax = b = p + e$. *The removed part e is the - inevitable - error due to the non-empty null-space of* A^T. *Lastly, there is the analytic interpretation obtained by taking partial derivatives with respect to the unknown coefficients in* $\hat{x}$ *and deriving the minimizer thereof.*

A.4.2 Eigenvalues and singular values

We first recall some facts about matrices and diagonalization.

A square matrix A is called diagonalizable if it is similar to a diagonal matrix, i.e., if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

The finite-dimensional spectral theorem says that any symmetric matrix whose entries are real can be diagonalized by an *orthogonal matrix*. More explicitly: For every symmetric real matrix A there exists a real orthogonal matrix Q such that $D = Q'AQ$ is a diagonal matrix.
Here, Q' is the transpose of Q for real matrices and the conjugate transpose (Hermitian transpose or adjoint matrix) for a complex matrix: $\mathbf{A}' = (\overline{\mathbf{A}})^{\mathrm{T}} = \overline{\mathbf{A}^{\mathrm{T}}}$.
Every real symmetric matrix is Hermitian, and therefore all its eigenvalues are real. As a consequence, since $Q^{-1} = Q'$ for unitary (or orthogonal) matrices[3], inversion of symmetric real matrices is straight-forward once its decomposition $QDQ' = A$ is known: $A^{-1} = (QDQ')^{-1} = QD^{-1}Q'$.

Example A.4.8. *The matrix* $M_C = \begin{pmatrix} 0.1 & 0 & 0.1 & 1 \\ 1 & 0.1 & 0 & 0.1 \\ 0.1 & 1 & 0.1 & 0 \\ 0 & 0.1 & 1 & 0.1 \end{pmatrix}$ *is the matrix corresponding to finite discrete convolution with the vector* $k_C = (.1, 1, 0.1, 0)$. *Check that* $M_C \cdot \delta = k_C$. *The eigenvectors of this matrix are therefore given by the vectors*

$$s_l[n] = e^{2\pi i l \frac{n}{4}},\ l = 0, \ldots, 3.$$

We can easily compute the eigenvalues - and therefore the inverse of M_C*: since the eigenvectors* s_l *constitute exactly the matrix of the finite discrete Fourier transform, we have, for any vector* $v \in \mathbb{R}^4$*:*

$$M_C v = \mathcal{F} D \mathcal{F}' v$$

and this is, once more, the convolution relation for the Fourier transform: instead of convolving two vectors, we can take their Fourier transforms and apply pointwise multiplication. The action of pointwise multiplication written by means of a matrix is the multiplication with a diagonal matrix:

$$\mathcal{F}' \left(M_C v \right) = D \left(\mathcal{F}' v \right).$$

The entries of the diagonal matrix D *are, of course, the eigenvalues of the convolution, and are therefore given by the Fourier transform of the convolution kernel* k_C*:*

$$\mathcal{F}'(v * k_C) = \hat{v} \cdot \widehat{k_C} = D \left(\mathcal{F}' v \right).$$

In our concrete example, we have $\widehat{k_C} = (1.2, -\mathrm{i}, -0.8, \mathrm{i})$ *and therefore*

$$M_C = \mathcal{F} D \mathcal{F}' = \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5\mathrm{i} & -0.5 & -0.5\mathrm{i} \\ 0.5 & -0.5 & 0.5 & -0.5 \\ 0.5 & -0.5\mathrm{i} & -0.5 & 0.5\mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 1.2 & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ 0 & 0 & -0.8 & 0 \\ 0 & 0 & 0 & \mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & -0.5\mathrm{i} & -0.5 & 0.5\mathrm{i} \\ 0.5 & -0.5 & 0.5 & -0.5 \\ 0.5 & 0.5\mathrm{i} & -0.5 & -0.5\mathrm{i} \end{pmatrix}.$$

[3] A unitary matrix in which all entries are real is an orthogonal matrix.

The inverse of M_C is then

$$M_C^{-1} = \mathcal{F}D^{-1}\mathcal{F}' = \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5\mathrm{i} & -0.5 & -0.5\mathrm{i} \\ 0.5 & -0.5 & 0.5 & -0.5 \\ 0.5 & -0.5\mathrm{i} & -0.5 & 0.5\mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 5/6 & 0 & 0 & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & -5/4 & 0 \\ 0 & 0 & 0 & -\mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & -0.5\mathrm{i} & -0.5 & 0.5\mathrm{i} \\ 0.5 & -0.5 & 0.5 & -0.5 \\ 0.5 & 0.5\mathrm{i} & -0.5 & -0.5\mathrm{i} \end{pmatrix}.$$

Not all square matrices are invertible, let alone any rectangular matrices. We saw in the section on least squares approximation that we might still be interested in an inversion of the action of A, "where it is possible", in other words, on the range of A. The pseudoinverse does exactly that: it inverts the action of A mapping the row-space to the column-space. Before we look at this new inversion, we have to introduce a generalization of the eigenvalues, which are the singular values. Again, *any* matrix has a singular value decomposition (SVD)!

Singular value decomposition

The main feature of diagonalization of symmetric matrices is the fact, that the action of A can be written as a diagonal matrix by means of a change of basis. As we saw above, even inversion is then easily realized.
We are seeking a similar representation for all matrices, in particular for $k \times n$-matrices, when $n \neq k$. In the general setting, however, we will have to work with two ONBs: one for the domain space ($\mathbb{R}^n$), one for the range space ($\mathbb{R}^k$).
The SVD is a factorization of a real or complex $k \times n$ matrix A with rank r of the form $A = U\Sigma V'$ where U is a $k \times k$ unitary matrix , Σ is a $k \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and V is an $n \times n$ unitary matrix and V' denotes the complex transposition (adjoint) of A, or just transposition in the real case. The diagonal entries Σ are the singular values of A. The columns of U are the left singular vectors and form an ONB of $\mathbb{R}^k$ and the columns of V are the right singular vectors of A and form an ONB of $\mathbb{R}^n$. The SVD has many useful applications in signal processing and statistics.

The singular value decomposition and the eigen-decomposition are closely related, since

- the left singular vectors of A are eigenvectors of AA' .
- the right singular vectors of A are eigenvectors of $A'A$.
- the non-zero singular values of A are the square roots of the non-zero eigenvalues of AA' or AA'.

Proposition A.4.9 (SVD)**.** *Let $A : \mathbb{R}^n \mapsto \mathbb{R}^k$ with rank r. There exist an ONB $\mathcal{V} = \{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ and an ONB $\mathcal{U} = \{u_1, \ldots, u_k\}$ of $\mathbb{R}^k$ such that*

$$Av_i = s_i u_i;\ A'u_i = s_i v_i \ \text{ and } A'Av_i = s_i^2 u_i, \ \text{ with } s_i > 0 \text{ for } i \leq r.$$

Proof. (1) $A'A$ is symmetric and can therefore by diagonalized by an ONB $\mathcal{V} = \{v_1, \ldots, v_n\}$. Recall that $rank(A'A) = rank(A) = r$, hence, we may order the eigenvalues such that

$$\lambda_1 \geq \ldots \geq \lambda_r >= 0 = \lambda_{r+1} \ldots \lambda_n$$

and the vectors $v_1, \ldots, v_r$ form an ONB for the range of A' (row-space), while $v_{r+1}, \ldots, v_n$ form an ONB for the kernel of A.
(2) Set $s_i = \sqrt{\lambda_i}$ and $u_i = \frac{1}{\sqrt{\lambda_i}} A v_i$ for $i = 1, \ldots, r$. Then, the $\{u_i, i = 1, \ldots, r\}$ span the range of A and they form an ONB, since

$$\begin{aligned}\langle u_i, u_j \rangle =& \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}} \langle Av_i, Av_j \rangle = \frac{1}{\sqrt{\lambda_i \lambda_j}} \langle A'Av_i, v_j \rangle \\ =& \frac{\lambda_i}{\sqrt{\lambda_i \lambda_j}} \langle v_i, v_j \rangle = \frac{1}{\sqrt{\lambda_j}} \langle v_i, v_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else,} \end{cases}\end{aligned}$$

since $\mathcal{V}$ is an ONB. $\{u_i, i = 1, \ldots, r\}$ is an ONB for the range (column space) of A and, since the kernel of A' is orthogonal to the range of A, can be extended to an ONB $\mathcal{U}$ of all $\mathbb{R}^k$ by adding an ONB of the kernel of A'. □

Corollary A.4.10 (SVD)**.** *Every $k \times n$ matrix A with rank r can be decomposed by $A = U\Sigma V'$, with U, V unitary (orthogonal) $k \times k$ and $n \times n$ matrices, respectively, and*

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

is a $k \times n$ matrix in which D is an $r \times r$ diagonal matrix with the positive singular values $s_i, i = 1, \ldots, r$ of A in the diagonal.

Proof. This decomposition follows directly from the construction of Theorem A.4.9, by letting V be the unitary $n \times n$ matrix with the vectors of the ONB $\mathcal{V}$ as its columns and U the $k \times k$ matrix with the vectors of the ONB $\mathcal{U}$ as its columns. Then

$$AV = (s_1 u_1 \cdots s_r u_r \, 0 \cdots 0) = U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

and multiplication with A' from the right completes the proof. □

Example A.4.11. *The SVD of*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}$$

is given by

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Example A.4.12. *The SVD of*

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

is given by

$$A = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{pmatrix} \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}$$

Example A.4.13. *The matrix*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

cannot be diagonalized by means of an orthonormal basis! However, we can find its SVD, given by:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{10} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 1 & 0 & 0 \\ 0 & 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}$$

A.4.3 Pseudoinverse: a generalization of matrix inversion

The singular value decomposition can be used for computing the *pseudoinverse* (PINV) of a matrix. PINV is a generalized inverse and is constructed according to the following ideas:
Given a $k \times n$ matrix A, $\mathbb{R}^n \to \mathbb{R}^k$, we consider the injective mapping given by restricting A to the orthogonal complement of its kernel $N(A)$, which is the row-space of A, or column space of A': $C(A')$, cp. Figure A.3. So, we consider the injective mapping

$$\tilde{A} : C(A') \to \mathbb{R}^k.$$

Now, A and $\tilde{A}$ have the same range, which is $C(A)$ and $\tilde{A}$ considered as a mapping $C(A') \to C(A)$ has an inverse:

$$\tilde{A}^{-1} : C(A) \to C(A'),$$

and we expect that, since $Av_i = s_i u_i$ for the members of the ONBs of $C(A')$ and $C(A)$, $i = 1, \ldots, r$, that

$$\tilde{A}^{-1} u_i = \frac{v_i}{s_i}.$$

The mapping $\tilde{A}^{-1}$ can be extended to an operator $A^+ : \mathbb{R}^k \to \mathbb{R}^n$ by defining

$$A^+(u_1 + u_3) = \tilde{A}^{-1}(u_1) \text{ for } u_1 \in C(A)\, u_3 \in N(A') = C(A)^\perp.$$

In other words, the part of the vector $u = u_1 + u_3$ that is orthogonal to the range of A and is thus in the kernel of A', is set to 0 by the pseudoinverse. Have a look at Figure A.4.

Pseudoinverse $A^+ u_i = \dfrac{v_i}{\sigma_i}$ for $i = 1, \ldots, r$.

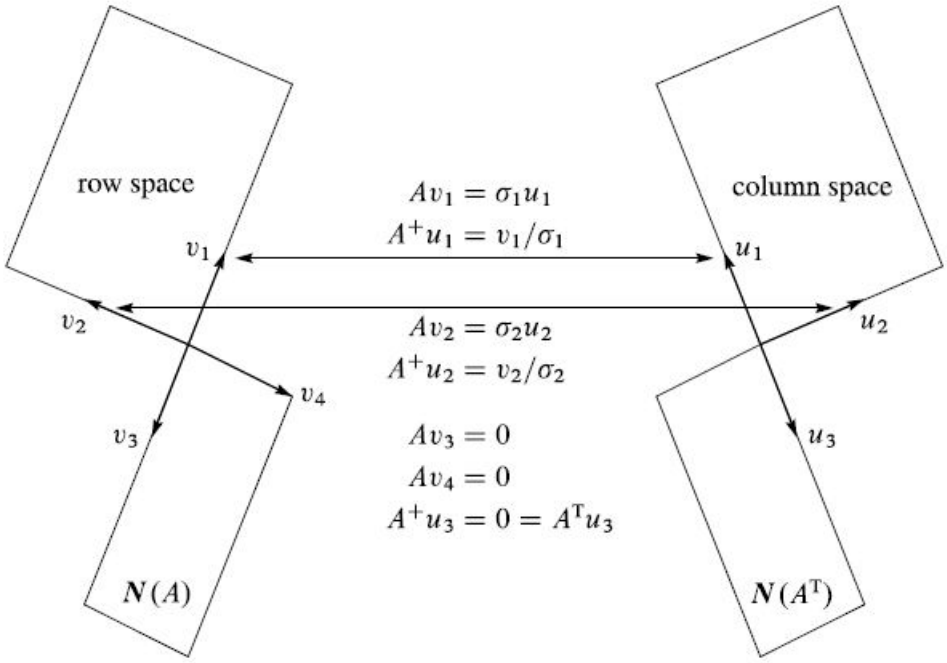


Figure A.4: Orthonormal bases that diagonalize A (3 by 4) and A^+ (4 by 3), matrices with rank 2.

It is immediately clear, that A^+ fulfills the desired property $AA^+u = u$ if $u \in C(A)$, i.e. the product is the identity on the range of A, compare this to the corresponding property of an invertible matrix!

Now we will see that the work we did in the previous section was not in vain! Indeed, the pseudoinverse of the matrix A with singular value decomposition $A = U\Sigma V'$ is

$$A^+ = V\Sigma^+U', \tag{A.8}$$

where Σ^+ is the pseudoinverse of Σ, and Σ^+ is formed by replacing every nonzero diagonal entry by its reciprocal and transposing the resulting matrix. In other words, using the notation of Corollary A.4.10, we have

$$\Sigma^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Example A.4.14. *The PINV of A from Example A.4.11 is given by*

$$A^+ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Example A.4.15. *The PINV of A from Example A.4.12 is given by*

$$A^+ = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}$$

Example A.4.16. *The PINV of A from Example A.4.13 is given by:*

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1/\sqrt{5} & 0 & 2/\sqrt{5} \\ 2/\sqrt{5} & 0 & -1/\sqrt{5} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{10}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

Proposition A.4.17. *Let A be a $k \times n$ matrix and let $A^+ = V\Sigma^+U'$ be its PINV as defined in* (A.8)*.*
(i)A^+ maps the range of A, i.e. $C(A)$ onto the row space of A, which is $C(A')$. The kernel of A', i.e., the orthogonal complement of the range of A is mapped to 0*.*
(ii) A^+ is the unique $n \times k$ matrix for which AA^+ is the orthogonal projection onto the range of A ($C(A)$) and A^+A is the orthogonal projection onto the range of A' ($C(A')$).

Proof. (i) Recall that the columns of U are the members of an ONB $\mathcal{U}$ for $\mathbb{R}^k$, in which the first r vectors $u_1, \ldots, u_r$ span the range of A, $C(A)$. Hence, the range of A consists of all linear combinations $\sum_{j=1}^r c_j u_j =: v$. Then

$$U' \cdot v = \begin{pmatrix} \langle \sum_{j=1}^r c_j u_j, u_1 \rangle \\ \langle \sum_{j=1}^r c_j u_j, u_2 \rangle \\ \vdots \\ \langle \sum_{j=1}^r c_j u_j, u_r \rangle \\ \langle \sum_{j=1}^r c_j u_j, u_{r+1} \rangle \\ \vdots \\ \langle \sum_{j=1}^r c_j u_j, u_k \rangle \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

by orthonormality of $\mathcal{U}$. Consequently, the $n \times 1$-vector $\Sigma^+ \cdot U' \cdot \upsilon$ is given by

$$\Sigma^+ \cdot U' \cdot \upsilon = \begin{pmatrix} \frac{c_1}{\sigma_1} \\ \vdots \\ \frac{c_r}{\sigma_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and therefore

$$V\Sigma^+ \cdot U' \cdot \upsilon = \sum_{j=1}^{r} \frac{c_j}{\sigma_j} v_j,$$

which is in the row space $C(A')$ of A, since $v_1, \ldots, v_r$ span $C(A')$. From the same derivation it is apparent that the kernel of A', for which $u_{r+1}, \ldots, u_k$ form an ONB , i.e., the orthogonal complement of the range of A, is mapped to 0.
(ii) To prove the second statement, first note that AA^+ is an orthogonal projection, since, by unitarity of V

$$AA^+ = U \cdot \Sigma \cdot V' \cdot V \cdot \Sigma' \cdot U' = U \cdot \Sigma \cdot \Sigma' \cdot U'$$

and

$$\Sigma \cdot \Sigma' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

such that $(AA^+)^2 = AA^+$ and also $(AA^+)' = AA^+$.
From $AA^+ = U \cdot \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \cdot U'$ it is easy to see that, for $\nu = \sum_{j=1}^{k} c_j u_j \in \mathbb{R}^k$, we have

$$AA^+\nu = \sum_{j=1}^{r} \langle \nu, u_j \rangle u_j,$$

because all coefficients corresponding to indices above r are set to zero by the diagonal operator $\Sigma \cdot \Sigma'$. Thus, AA^+ is the orthogonal projection onto the range of A. The statement for A^+A is shown completely analogously. □

A.4.4 Pseudoinverse and least squares

The pseudoinverse is one way to solve linear least squares problems.

Proposition A.4.18. *Let A be a $k \times n$ matrix, Then $x^+ = A^+b$ is a least squares solution of $Ax = b$ and for any other least squares solution $\tilde{x}$, we have $\|\hat{x}\|_2 \leq \|\tilde{x}\|_2$, so $\hat{x}$ has minimal norm.*

Proof. Recall that $\hat{x}$ of a least-squares solution of $Ax = b$ is and only if it solves $A'Ax = A'b$. First note that x^+ is a least squares solution by observing that $Ax^+ - b = AA^+b - b$ and since AA^+b is the orthogonal projection of b onto the range of A, $AA^+b - b$ is in the orthogonal complement of the range, hence in th kernel of A' and therefore $A'(AA^+b - b) = 0$.
We still have to show that for any other least squares solution $\hat{x}$ of $Ax = b$, we have $\|\hat{x}\| \leq \|x^+\|$. To see this, note that

$$A'(A\hat{x} - b) = A'(Ax^+ - b) = 0 \Rightarrow A'A\hat{x} = A'Ax^+$$

and thus $\hat{x} - x^+ \in N(A'A)$ and also $\hat{x} - x^+ \in N(A)$ (by Proposition 8). However, $x^+ \perp N(A)$ and we can estimate:

$$\begin{aligned}\|\hat{x}\|^2 =& \|x^+ - x^+ + \hat{x}\|^2 \\ =& \|x^+\|^2 + \|\hat{x} - x^+\|^2 + 2Re(\langle x^+, \hat{x} - x^+\rangle) \\ =& \|x^+\|^2 + \|\hat{x} - x^+\|^2 \geq \|x^+\|^2\end{aligned}$$

□

A.5 An Application: Distributional Poisson's equation

For $s > 2$ and $f \in \mathcal{S}(\mathbb{R}^d)$, consider

$$-\Delta u = f \qquad \text{in } \mathcal{S}'(\mathbb{R}^d). \tag{A.9}$$

Lemma A.5.1. *For $d \geq 3$, $f : \mathbb{R}^d \to \mathbb{C}$, $x \mapsto \frac{1}{\|x\|^2}$ satisfies $f \in \mathcal{S}'(\mathbb{R}^d)$.*

Proof. Let $B \subset \mathbb{R}^d$ denote the unit ball. By spherical coordinates, we observe

$$\int_B \frac{1}{\|x\|^2}\mathrm{d}x = \mathrm{vol}(\mathbb{S}^{d-1}) \int_0^1 \frac{1}{r^2} r^{d-1}\mathrm{d}r = \int_0^1 r^{d-3}\mathrm{d}r < \infty,$$

so that $f \in L^1_{loc}(\mathbb{R}^d)$. Theorem 3.5.2 concludes the proof. □

According to Lemma A.5.1, there is $E \in \mathcal{S}'(\mathbb{R}^d)$ such that $\hat{E} = \frac{1}{4\pi^2\|\cdot\|^2}$. For $f \in \mathcal{S}(\mathbb{R}^d)$, we solve (A.9) by putting

$$u := E * f$$

since then $\hat{u} = \hat{E}\hat{f} = \frac{1}{4\pi^2\|\xi\|^2}\hat{f}$, so that (3.43) is satisfied.

Remark A.5.2. *In a course on PDE one observes that $E(x) = \frac{c_d}{\|x\|^{d-2}}$, for some suitable constant $c_d \in \mathbb{R}$.*

Definition A.5.3. *For $s \in \mathbb{R}$, the Sobolev space is*

$$H^s(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : (1 + \|\cdot\|^2)^{s/2}\hat{f} \in L^2(\mathbb{R}^d)\}$$

with inner product

$$\langle f, g\rangle_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} (1 + \|\xi\|^2)^s \hat{f}(\xi)\overline{\hat{g}(\xi)}\mathrm{d}\xi.$$

Lemma A.5.4. *The Sobolev space $H^s(\mathbb{R}^d)$ is a Hilbert space.*

Proof. Due to Cauchy-Schwartz, the inner product is well-defined. To verify completeness, let $(f_n)_{n\in\mathbb{N}} \subset H^s(\mathbb{R}^d)$ be Cauchy. We observe that $(1+\|\xi\|^2)^{s/2}\hat{f}_n(\xi)$ is Cauchy in $L^2(\mathbb{R}^d)$, hence, converges towards $g \in L^2(\mathbb{R}^d)$. Define

$$f := \mathcal{F}^{-1}(g(\xi)(1+\|\xi\|^2)^{-s/2}),$$

so that $\hat{f}(\xi)(1+\|\xi\|^2)^{s/2} \in L^2(\mathbb{R}^d)$, hence, $f \in H^s(\mathbb{R}^d)$. We estimate

$$\begin{aligned}\|f_n - f\|^2_{H^s(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} (1+\|\xi\|^2)^s |\hat{f}_n(\xi) - \hat{f}(\xi)|^2 \mathrm{d}(\xi) \\ &= \int_{\mathbb{R}^d} |(1+\|\xi\|^2)^{s/2}\hat{f}_n(\xi) - g(\xi)|^2 \mathrm{d}(\xi) \to 0. \end{aligned}$$ □

Corollary A.5.5. *For $s > 2$, we have $\Delta : H^s(\mathbb{R}^d) \to H^{s-2}(\mathbb{R}^d)$ continuously.*

Proof. For $u \in H^s(\mathbb{R}^d)$, we observe

$$\begin{aligned}\|\Delta u\|^2_{H^{s-2}(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} (1+\|\xi\|^2)^{s-2} \left|\widehat{\Delta u}(\xi)\right|^2 \mathrm{d}\xi \\ &= 16\pi^4 \int_{\mathbb{R}^d} (1+\|\xi\|^2)^{s-2} \underbrace{\|\xi\|^4}_{\leq (1+\|\xi\|^2)^2} |\hat{u}(\xi)|^2 \,\mathrm{d}\xi \\ &\leq 16\pi^4 \int_{\mathbb{R}^d} (1+\|\xi\|^2)^s \,|\hat{u}(\xi)|^2 \,\mathrm{d}\xi \\ &= 16\pi^4 \|u\|^2_{H^s(\mathbb{R}^d)}. \end{aligned}$$ □

For $f \in H^s(\mathbb{R}^d)$, consider

$$(I - \Delta)u = f.$$

A solution is given by

$$u := \mathcal{F}^{-1}\left(\frac{1}{1+4\pi^2\|\cdot\|^2}\hat{f}\right)$$

and the obvious inequality $\frac{1+\|\xi\|^2}{1+4\pi^2\|\xi\|^2} \leq 1$ leads to $u \in H^{s+2}(\mathbb{R}^d)$.

Bibliography

[1] L. Grafakos, *Classical fourier analysis*, Graduate Texts in Mathematics, 2014.

[2] Philipp Grohs and Lukas Liehr, *Injectivity of gabor phase retrieval from lattice measurements*, Applied and Computational Harmonic Analysis **62** (2023), 173–193 (English), Publisher Copyright: © 2022 The Author(s).