Elements of Digital Signal Processing in 1D (2024/25)

Alessandro Verri

Abstract

These notes serve the purpose of giving a succinct account of the topics covered in the first part of the *Digital Signal and Image Processing* class. For whoever is interested in digging deeper references are provided in each section. We start with *Fourier Series* and *Fourier Transform*, spending a considerable amount of time in the issues arising in sampling a signal. After a brief interlude on *Linear Time-Invariant Systems*, we introduce *Kalman Filtering* as an example of what can be done if causality must be taken into account. Finally, we discuss *Discrete Haar Wavelets*, a signal transformation that is local in both time and frequency domain, and *Frames*, an example of redundant signal representation.

Contents

1 Fourier series **3**
1.1 A linear space of functions 3
1.2 Bases 4
1.3 Convergence and generalisation 6
1.4 A simple example and some useful tips 7
1.5 Things you need to know 8

2 Fourier Transform **9**
2.1 Singular functions 9
2.2 A reversible transformation 12
2.3 Main properties 13
2.4 Fourier Pairs 15
2.5 Sampling 17
2.6 Things you need to know 20

3 Linear Time-Invariant Systems **21**
3.1 Definition 21
3.2 Characterisation of an LTI system 21
3.3 Properties 22
3.4 RC circuit 24
3.5 Things you need to know 25

4 Kalman Filtering **26**
4.1 Computing the average revisited 26
4.2 General case 28
4.3 Linear Kalman Filter 30
4.4 Computing the derivative of a function affected by noise 31
4.5 Things you need to know 33

5 A Brief Intro to Wavelets **34**
5.1 A different representation for a 4-pixel image . . . 34
5.2 Haar wavelet transform . . . 34
5.3 Multiresolution analysis . . . 38
5.4 Things you need to know . . . 40

6 Frames **41**
6.1 Beyond orthogonality . . . 41
6.2 Analysis and synthesis filters . . . 42
6.3 Frame operator . . . 44
6.4 Concluding remarks . . . 46
6.5 Things you need to know . . . 48

References

[1] P.G. Casazza, G. Kutyniok & F. Philipp. *Introduction to Finite Frame Theory*. In Finite Frames, pp 1-53, Birkhäuser, 2013.

[2] I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.

[3] A.N. Kolmogorov & S.V. Fomin. *Elements of the Theory of Functions and Functional Analysis*. Martino Fine Books, 2012.

[4] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 2009.

[5] P.S. Maybeck. *Stochastic Models, Estimation, and Control*. Academic Press, 1979.

[6] A. Papoulis. *The Fourier Integral and its Applications*. Mc Graw-Hill, 1962.

[7] N.A. Thacker & J.C. Lacey. *Tutorial: The Kalman Filter*. MIT Tina Memo, 1998.

[8] E.J. Stollnitz, T.D. De Rose & D.H. Salesin. *Wavelet for Computer Graphics: a Primer*. University of Washington Technical Report 94-9-11, 1994.

1 Fourier series

Any textbook on Fourier Series can be used as a reference. A rigorous treatment can be found in [3].

1.1 A linear space of functions

We know that a vector $\mathbf{v}$ in a Euclidean linear space V of dimension N can be represented by N real numbers $v_1, \ldots, v_N$, projections of $\mathbf{v}$ on each of the N mutually orthogonal unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_N$ which form an orthonormal basis for V. The vector $\mathbf{v}$ can be written as

$$\mathbf{v} = \sum_{i=1}^{N} v_n \mathbf{e}_n$$

We now consider a Euclidean linear space V in which a vector is a piece-wise continuous function over a fixed interval $[a, b]$. We will learn that a basis for this space can be constructed starting from an infinite number of mutually orthogonal periodic functions of period equal to $(b-a)/N$ for $N = 1, \ldots, +\infty$. Not surprisingly, each vector in V can be written as a linear combination of the basis elements. Neglecting the technical difficulties posed by the fact that the basis has infinite cardinality, we obtain a remarkable result: **any piece-wise continuous function can be written as a linear combination of a fixed and countable set of functions and can thus be represented as a discrete sequence of real numbers, components of the function in some fixed basis.** Like in the finite dimensional case, we will be able to compute the components by simply taking the scalar product between the function and each basis element.

The fact that a and b are finite has an important consequence: **the expansion we find of a function in the interval $[a, b]$ is a periodic function of period $b - a$ over the real line**.

Definition 1.1. *Piece-wise continuitity*
A *piece-wise continuous* function $f : \mathbb{R} \to [a, b]$ has at most a finite number of points between a and b in which it is not continuous. □

Let V be the set of all real valued *piece-wise continuous* functions over $[a, b]$ for $a < b \in \mathbb{R}$.

Fact 1.1. *V is a linear space*
Since any linear combination of piece-wise continuous functions over $[a, b]$ is piece-wise continuous, it is straightforward to realise that V is a linear space. For notational consistency we denote with $\mathbf{f}$ a piece-wise continuous function $f : [a, b] \to \mathbb{R}$ thought of as en element of V.

Definition 1.2. *A scalar product between functions*
We define the scalar product between two functions $\mathbf{f}$ and $\mathbf{g} \in V$ as

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b f(t) g(t) \mathrm{d}t \tag{1}$$

We note that the integral is always well defined, since the product of piece-wise continuous functions is piece-wise continuous and, hence, Riemann integrable.

Observation 1.1. *Norm of a function*
We can measure the norm of a function $\mathbf{f}$ as

$$\|\mathbf{f}\| = \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle} = \sqrt{\int_a^b f^2(t) \mathrm{d}t}$$

Observation 1.2. *Orthogonality between functions*
Two functions $\mathbf{f}$ and $\mathbf{g}$ are *orthogonal* if

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b f(t)g(t)\mathrm{d}t = 0$$

Observation 1.3. *Complex valued case*
If $\mathbf{f}$ and $\mathbf{g}$ are complex valued we define the scalar product as

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b f(t)g^*(t)\mathrm{d}t$$

where $g^*(t)$ is the complex conjugate of $g(t)$ and $\langle \mathbf{f}, \mathbf{g} \rangle = \langle \mathbf{g}, \mathbf{f} \rangle^*$.

Observation 1.4. *The intuition behind*
Let us assume, for simplicity, that the two functions $\mathbf{f}$ and $\mathbf{g}$ in Equation (1) are continuous. We divide the interval $[a, b]$ in M equal non-overlapping intervals of width $(b-a)/M$ and consider the M midpoints t_m with $m = 1, \ldots, M$. The integral in Equation (1) can then be rewritten as

$$\langle \mathbf{f}, \mathbf{g} \rangle = \lim_{M \to +\infty} \frac{b-a}{M} \sum_{m=1}^{N} f(t_m)g(t_m)$$

Aside from the scaling factor $(b-a)/M$ and the limit, the analogy with the scalar product of two vectors $\mathbf{f}$ and $\mathbf{g}$ in a Euclidean vector space of dimension M

$$\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{m=1}^{M} f_m g_m$$

is immediate.

1.2 Bases

Definition 1.3. *Orthogonal and orthonormal system of functions*
If $\langle \boldsymbol{\phi}_n, \boldsymbol{\phi}_m \rangle = 0 \ \forall n \neq m \in \mathbb{N}$, then $S_\phi = \{\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \ldots\}$ is an *orthogonal system*. Furthermore, if $\langle \boldsymbol{\phi}_n, \boldsymbol{\phi}_n \rangle = 1 \ \forall n \in \mathbb{N}$, S_ϕ is *orthonormal*.

Example 1.1. *Orthonormality for* $[a, b] = [-\pi, \pi]$
The system $S_\phi = \{\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \ldots\}$ with

$$\boldsymbol{\phi}_0 = \phi_0(t) = \frac{1}{\sqrt{2\pi}}, \ \ \boldsymbol{\phi}_{2n-1} = \phi_{2n-1}(t) = \frac{\sin nt}{\sqrt{\pi}}, \ \ \boldsymbol{\phi}_{2n} = \phi_{2n}(t) = \frac{\cos nt}{\sqrt{\pi}}$$

with $n \in \mathbb{N}$ is orthonormal.

Example 1.2. *Easier with complex exponentials*
The system $S_\psi = \{\boldsymbol{\psi}_0, \boldsymbol{\psi}_1, \boldsymbol{\psi}_{-1}, \ldots\}$ with

$$\boldsymbol{\psi}_n = \psi_n(t) = \frac{e^{int}}{\sqrt{2\pi}} \ \text{ with } \ n \in \mathbb{Z} \tag{2}$$

is also orthonormal.

The orthonormality of the system S_ψ can be easily checked. Indeed using the definition of scalar product with the complex conjugate, for all $n \in \mathbb{Z}$ we have

$$\langle \boldsymbol{\psi}_n, \boldsymbol{\psi}_n \rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{int}e^{-int}\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}t = \frac{2\pi}{2\pi} = 1$$

and for $n \neq m \in \mathbb{Z}$

$$\langle \boldsymbol{\psi}_n, \boldsymbol{\psi}_m \rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)t}\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi}\cos{(n-m)t}\mathrm{d}t + \frac{i}{2\pi}\int_0^{2\pi}\sin{(n-m)t}\mathrm{d}t = 0$$

because both the real and the imaginary part are integral of sinusoidal functions of period $2\pi/|n-m|$ over an interval of width 2π.

We now extend the notion of linear independence to infinite sets. We recall that if a finite set of N vectors $\{\boldsymbol{\phi}_1, \ldots \boldsymbol{\phi}_N\}$ are linearly independent, then

$$\sum_{n=1}^{N} d_n\phi_n(t) = 0 \;\; \forall t \in [a,b] \text{ and } d_n \in \mathbb{R}$$

implies

$$d_n = 0 \quad \text{for } n = 1, \ldots, N$$

Observation 1.5. *Infinite case*
If S is infinite, the vectors in S are linearly independent if all finite subsets of S consist of linearly independent vectors.

Observation 1.6. *Mutual orthogonality implies linearly independence*
The vectors of an orthogonal system S are linearly independent. □

A natural question to ask is whether an infinite set of orthonormal functions, like S_ϕ or S_ψ, is a basis for a certain vector space. For example, let us consider with the $\boldsymbol{\psi}_n$ as in Equation (2) the *Fourier Series*

$$\sum_{n=0}^{+\infty} f_n\psi_n(t) \quad \text{with} \quad f_n = \langle \mathbf{f}, \boldsymbol{\psi} \rangle = \int_0^{2\pi} f(s)\psi_n^*(s)\mathrm{d}s$$

For which values of n the series converge? And if it does, does it converge to $f(t)$?

We first answer a simpler question. We let

$$\mathbf{f}_N = \sum_{n=1}^{N} f_n\boldsymbol{\phi}_n \;\text{ and }\; \mathbf{d}_N = \sum_{n=1}^{N} d_n\boldsymbol{\phi}_n$$

While the coefficients f_n are computed by projecting the function $\mathbf{f}$ onto each of the $N < +\infty$ basis vectors considered, the d_n are arbitrary real numbers.

Theorem 1.1. *Best approximation property*

$$\forall N > 0 \;\; \|\mathbf{f} - \mathbf{d}_N\|^2 \geq \|\mathbf{f} - \mathbf{f}_N\|^2$$

Proof
Since

$$\|\mathbf{f}_N\|^2 = \sum_{n=1}^{N} f_n^2 \;\text{ and }\; \langle \mathbf{f}, \mathbf{f}_N \rangle = \sum_{n=1}^{N} f_n^2$$

we have

$$\|\mathbf{f}-\mathbf{f}_N\|^2 = \|\mathbf{f}\|^2 - \sum_{n=1}^{N} f_n^2 \geq 0 \tag{3}$$

Furthermore, since

$$\|\mathbf{d}_N\|^2 = \sum_{n=1}^{N} d_n^2 \ \text{ and } \ \langle \mathbf{f}, \mathbf{d}_N\rangle = \sum_{n=1}^{N} d_n f_n$$

adding and subtracting $\sum f_n^2$ and using inequality (3) we obtain

$$\begin{aligned}\|\mathbf{f}-\mathbf{d}_N\|^2 &= \|\mathbf{f}\|^2 + \|\mathbf{d}_N\|^2 - 2\langle \mathbf{f}, \mathbf{d}_N\rangle = \|\mathbf{f}\|^2 + \sum_{n=1}^{N} d_n^2 + \sum_{n=1}^{N} f_n^2 - \sum_{n=1}^{N} f_n^2 - 2\sum_{n=1}^{N} d_n f_n \\ &= \|\mathbf{f}\|^2 + \sum_{n=1}^{N}(d_n - f_n)^2 - \sum_{n=1}^{N} f_n^2 = \|\mathbf{f}-\mathbf{f}_N\|^2 + \sum_{n=1}^{N}(d_n-f_n)^2 \geq \|\mathbf{f}-\mathbf{f}_N\|^2\end{aligned}$$

■

Observation 1.7. *A general result*
The result we proved holds true for any Euclidean space: no other linear combination leads to a better approximation of a vector $\mathbf{f}$ in a certain subspace than the one obtained by projecting $\mathbf{f}$ along the corresponding basis vectors of that subspace.

Fact 1.2. *Parseval equality (Pythagora, once again)*
From Equation (3) we have that $\sum f_n^2 \leq \|\mathbf{f}\|^2$ and thus

$$\sum f_n^2 = \|\mathbf{f}\|^2 \Longleftrightarrow \text{for } \ N \to \infty \ \ \|\mathbf{f}-\mathbf{f}_N\| \to 0$$

Observation 1.8. *Sanity check*
Parseval formula can thus be used to control whether the representation of a vector in terms of linear combination of certain basis elements is complete. If the square of the vector norm equals the sum of the square of the coefficients we are done. If it is strictly larger, then something is missing. Furthermore, if the Parseval formula holds true for all $\mathbf{f}$, the orthonormal system S is a basis.

Fact 1.3. *Completeness*
The orthonormal systems of **Example 1.1.** and **1.2.** are complete.

1.3 Convergence and generalisation

Fact 1.4. *Convergence*
Let $\mathbf{f}$ be piece-wise continuous and t a point of continuity for $\mathbf{f}$. For all $\epsilon > 0$ there exists N^* such that for all $N \geq N^*$

$$\left|\frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(nt) + b_n \sin(nt) - f(t)\right| < \epsilon$$

Fact 1.5. *Gibbs phenomenon*
The Fourier Series converges also at a point t^* of discontinuity for $\mathbf{f}$, but it converges to the average of the limit of $f(t)$ for $t \to t^*_+$ and $t \to t^*_-$. For finite values of n the truncated series at points nearby t^* oscillates somewhat wildly, see Figure (1). This behaviour is known as Gibbs phenomenon.

Observation 1.9. *Arbitrary period*
There is nothing special in the period 2π. If the period is $T = 1/f$ the orthogonal systems vary accordingly. For example, in the case of complex exponentials, we have

$$\psi_n(t) = \frac{e^{i2\pi fnt}}{\sqrt{T}}$$

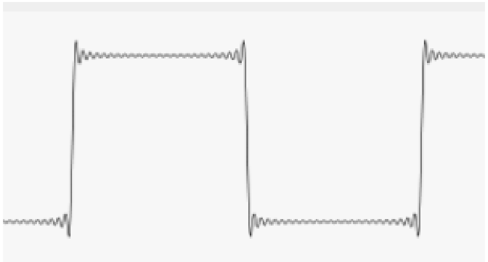

Figure 1: Gibbs phenomenon for a square wave.

1.4 A simple example and some useful tips

Once in a lifetime let us compute by hand the Fourier coefficients in a super simple example. We want to compute the Fourier series associated with the rectangular function $p_{\pi/2}(\cdot)$ where

$$p_{\pi/2}(t) = \begin{cases} 1 & -\pi/2 \leq t \leq \pi/2 \\ 0 & \text{otherwise} \end{cases}$$

We first rewrite the Fourier series the way it is typically found in textbooks. Using the $\phi_n(t)$ defined by Equation (2) we have

$$\frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos(nt) + b_n \sin(nt) \tag{4}$$

Observation 1.10. *Taking care of normalisation*
The coefficients of the linear expansion of Equation (4) are proportional but not equal to the projection of the original function on each of the basis element. There is a missing factor $1/\sqrt{\pi}$ in the projection and another $1/\sqrt{\pi}$ when the basis element is used in the expansion. When computing the square of the norm of a functions using the coefficients of Equation (4) this has to be taken into account. In addition, we write $a/2$ in place of a_0 so that we can compute a_0 using the same normalisation factor, $1/\sqrt{\pi}$, in use for a_n with $n > 0$. □

Since the rectangular function is even we only need to consider the projection of $p_{\pi/2}(\cdot)$ onto the cosine harmonics (including the constant). We thus find

$$\frac{a_0}{2} = \frac{1}{2}\frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} p_{\pi/2}(s)\mathrm{d}s = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2}\mathrm{d}s = \frac{1}{2}$$

and

$$a_n = \frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} p_{\pi/2}(s)\cos(ns)\mathrm{d}s = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\cos(ns)\mathrm{d}s = \frac{2\sin(n\pi/2)}{\pi n} = \frac{2}{\pi n}(-1)^{(n-1)/2}$$

We see that $a_n > 0$ for $n = 1, 5, \ldots$ and negative otherwise. Therefore if we write $n = 2k - 1$ for $k = 1, \ldots$ and replace k with n we obtain

$$p_{\pi/2}(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{\cos((2n-1)t)}{2n-1} = \frac{1}{2} + \frac{2}{\pi}\left(\cos t - \frac{\cos 3t}{3} + \frac{\cos 5t}{5} - \ldots\right) \tag{5}$$

Observation 1.11. *Beware of the equality sign*
The equality sign holds true at all points t for which $p_{\pi/2}$ is continuous. At the points $-\pi/2$ and $\pi/2$ the series converges to $1/2$.

Observation 1.12. *Hard to obtain with a computer!*
For $t = 0$, $p_{\pi/2}(0) = 1$ and Equation (5) gives

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

Observation 1.13. *Constants matter*
Normalisation factors are not the only constants to be checked. An additional source of confusion is related to M, the number of points used for sampling the considered interval. To be on the safe side M needs to be at least an order of magnitude greater than the highest harmonic. Furthermore, in the Parseval equality each coefficient is computed by taking the scalar product between the function and an element of an orthonormal basis. The sum of the square of the coefficients, therefore, equals the square of the norm of the function divided by the square of the normalisation factor (each basis element appears twice: in the scalar product computation and as one of the vectors in the linear combination).

1.5 Things you need to know

1. Any piece-wise continuous real valued function $\mathbf{f} : [a, b] \to \mathbb{R}$ can be written as a linear combination of properly normalised harmonic functions of period $T = b - a$

$$\frac{1}{\sqrt{T/2}} \sin\left(\frac{2\pi nt}{T}\right) \quad \text{and} \quad \frac{1}{\sqrt{T/2}} \cos\left(\frac{2\pi nt}{T}\right) \quad \text{for } n = 1, \ldots$$

   including the unit norm constant function $1/\sqrt{T}$.

2. Each coefficient in the linear combination is the scalar product between $\mathbf{f}$ and the corresponding harmonic, or the projection of $\mathbf{f}$ along the harmonic.

3. The larger the square of the coefficient, the higher the weight of the corresponding harmonic in the approximation. Eventually, the coefficients of higher order harmonics tend to 0.

4. Since the reconstructed function is periodic of period $T = b - a$ over the entire real line, the original function can also be viewed as periodic of period T and reconstructed accordingly.

5. Since the normalised harmonics form an orthonormal basis, the sum of the square of the coefficients is no greater than $||\mathbf{f}||^2$ (actually equal if we take all the coefficients). The smaller the difference between the two, the better the obtained approximation.

6. The series converges to $f(t)$ point-wise if t is a point in which f is continuous, or to the average of the left and right limit of $f(t)$

$$\frac{f(t_+) + f(t_-)}{2}$$

   if t is a point of discontinuity.

2 Fourier Transform

Most of the material of this section is taken from [6]. If you are after mathematical rigor check out [3]. The Fourier Transform can be viewed as a generalisation of the Fourier Series for functions defined over the entire real line. Instead of an infinite set of discrete coefficients, **the Fourier Transform is a complex valued function of the frequency**. While the Fourier Series of a signal in the interval $[a, b]$ represents and reconstructs the signal as a periodic function of period $b - a$, the Fourier Transform represents and reconstructs the signal for what it is.

2.1 Singular functions

A first model for the δ function Let $T > 0$ and consider the function

$$r_T(t) = \frac{1}{2T} p_T(t)$$

where, as usual,

$$p_T(t) = \begin{cases} 1 & -T \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$

For all values of T, $r_T(\cdot)$ defines a *rectangle* of unit area since

$$\int_{-\infty}^{+\infty} r_T(t)\mathrm{d}t = \frac{1}{2T}\int_{-T}^{T} \mathrm{d}t = \frac{2T}{2T} = 1$$

If $\phi : \mathbb{R} \to \mathbb{R}$ is any function continuous in $t = 0$, for small values of T we have

$$\int_{-\infty}^{+\infty} r_T(t)\phi(t)\mathrm{d}t = \frac{1}{2T}\int_{-T}^{T} \phi(t)\mathrm{d}t \approx \frac{1}{2T}2T\phi(0) = \phi(0)$$

If we take the limit for $T \to 0$ we thus have

$$\lim_{T\to 0}\int_{-\infty}^{+\infty} r_T(t)\phi(t)\mathrm{d}t = \phi(0) \tag{6}$$

Exchanging the *limit* with the *integral* in Equation (6), we could write $\lim_{T\to 0} r_T(t) = \delta(t)$ with

$$\delta(t) = \begin{cases} 0 & t \neq 0 \\ +\infty & t = 0 \end{cases}$$

and

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

Clearly, no ordinary function behaves like $\delta(t)$ but

$$\int_{-\infty}^{+\infty} \delta(t)\phi(t)\mathrm{d}t = \phi(0) \tag{7}$$

is a suggestive way to represent the *distribution* N_δ which associates to a function $\phi(\cdot) : \mathbb{R} \to \mathbb{R}$, continuous in $t = 0$, its value $\phi(0)$, or

$$N_\delta[\phi] = \phi(0) \tag{8}$$

Observation 2.1. *Why bother*
Let alone the fact that integrals with a δ function inside are very easy to evaluate, the rationale for introducing $\delta(t)$, example of a *singularity* function or *generalised* function, is that Equation (7) makes it apparent that the distribution N_δ in Equation (8) enjoys all the properties applicable to an integral.

Observation 2.2. *Sampling*
In this course, the distribution N_δ is of paramount importance. It describes the process through which a 1D time-varying signal is sampled at the specific time stamp, or a 2D image at a specific pixel!

Observation 2.3. *Distributions*
We have introduced the concept of distribution starting with the somewhat awkward example of a singularity function. Here are three examples you already came across in terms of ordinary functions. We invariably assume that for all the functions we use the integrals we write exist.

1. For all $T \neq 0$ we write

$$N_{p_T}[\phi] = \int_{-\infty}^{+\infty} p_T(t)\phi(t)\mathrm{d}t = \int_{-T}^{T} \phi(t)\mathrm{d}t$$

   The distribution $N_{p_T}[\phi]$ associates to ϕ the integral of ϕ between $-T$ and T.

2. If

$$U(t) = \begin{cases} 1 & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

   is the *unit step* function, we have

$$N_U[\phi] = \int_{-\infty}^{+\infty} U(t)\phi(t)\mathrm{d}t = \int_{0}^{+\infty} \phi(t)\mathrm{d}t$$

   The distribution $N_U[\phi]$ associates to ϕ the integral of ϕ between 0 and $+\infty$.

3. If X is a random variable taking on values in the interval $[-\infty, \infty]$, $\phi(X)$ is function of X, and $f(x) \geq 0$ the probability density function of X with

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$$

   the *expected value* of $\phi(X)$, $\mathbb{E}[\phi(X)]$, is the distribution

$$N_f[\phi] = \int_{-\infty}^{+\infty} \phi(x)f(x)\mathrm{d}x$$

   Notice that the notation $N_f[\phi]$, unlike $\mathbb{E}[\phi(X)]$, makes it explicit the dependence of the expected value on the probability density function.

Properties In what follows we assume that any function ϕ can be differentiated as many times as we need and $\phi(t) \to 0$ faster than any polynomials for $t \to \pm\infty$.

Linearity For any distribution $g(t)$, as an immediate consequence of the integral notation, we have

$$\int_{-\infty}^{+\infty} g(t)\left(\alpha\phi_1(t) + \beta\phi_2(t)\right)\mathrm{d}t = \alpha \int_{-\infty}^{+\infty} g(t)\phi_1(t)\mathrm{d}t + \beta \int_{-\infty}^{+\infty} g(t)\phi_2(t)\mathrm{d}t$$

Sifting Substituting $t - t_0 \to t$ gives for any distribution $g(t)$

$$\int_{-\infty}^{+\infty} g(t-t_0)\phi(t)\mathrm{d}t = \int_{-\infty}^{+\infty} g(t)\phi(t+t_0)\mathrm{d}t$$

For the δ function this reduces to

$$\int_{-\infty}^{+\infty} \delta(t-t_0)\phi(t)\mathrm{d}t = \phi(t_0)$$

Scale For all $a \neq 0$, substituting $at \to t$ gives that for any distribution $g(t)$

$$\int_{-\infty}^{+\infty} g(at)\phi(t)\mathrm{d}t = \frac{1}{|a|}\int_{-\infty}^{+\infty} g(t)\phi\left(\frac{t}{a}\right)\mathrm{d}t$$

The absolute value is due to the need of interchanging the limits, when $a < 0$, to restore integration from $-\infty$ to $+\infty$.

Derivative Integrating by parts (since $\phi(t) \to 0$ for $t \to \pm\infty$) we have that for the derivative of a distribution $g(t)$ we have

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}g(t)}{\mathrm{d}t}\phi(t)\mathrm{d}t = -\int_{-\infty}^{+\infty} g(t)\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}\mathrm{d}t$$

Since

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}U(t)}{\mathrm{d}t}\phi(t)\mathrm{d}t = -\int_{-\infty}^{+\infty} U(t)\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}\mathrm{d}t = -\int_{0}^{+\infty} \frac{\mathrm{d}\phi(t)}{\mathrm{d}t}\mathrm{d}t = \phi(0) - \phi(+\infty) = \phi(0)$$

we see that the $\delta(t)$ can be seen as the derivative of the unit step $U(t)$. This result is consistent with the intuition that the derivative of the unit step is 0 for all $t \neq 0$ and diverges in 0 because of the discontinuity.

Integral of the *sinc* We start by observing that

$$\begin{aligned}\int_{0}^{+\infty} e^{-xt}\sin x\mathrm{d}x &= \int_{0}^{+\infty} e^{-xt}\frac{e^{ix}-e^{-ix}}{2i}\mathrm{d}x = \frac{1}{2i}\left(\int_{0}^{+\infty} e^{x(i-t)}\mathrm{d}x - \int_{0}^{+\infty} e^{-x(i+t)}\mathrm{d}x\right) \\ &= \frac{1}{2i}\left.\frac{e^{x(i-t)}}{i-t}\right|_{0}^{+\infty} - \frac{1}{2i}\left.\frac{e^{-x(i+t)}}{-i-t}\right|_{0}^{+\infty} = \frac{1}{2i}\left(\frac{1}{t-i} - \frac{1}{t+i}\right) = \frac{1}{1+t^2} \end{aligned} \tag{9}$$

Since

$$\int_{0}^{+\infty} e^{-xt}\mathrm{d}t = -\left.\frac{e^{-xt}}{x}\right|_{0}^{+\infty} = \frac{1}{x}$$

we have

$$\int_{-\infty}^{+\infty} \frac{\sin x}{\pi x}\,\mathrm{d}x = \frac{2}{\pi}\int_{0}^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x = \frac{2}{\pi}\int_{-\infty}^{+\infty} \sin x\left(\int_{0}^{+\infty} e^{-xt}\,\mathrm{d}t\right)\mathrm{d}x \tag{10}$$

and if we swap the order of integration in Equation (10) and use Equation (9) we finally obtain

$$\int_{-\infty}^{+\infty}\frac{\sin x}{\pi x}\,\mathrm{d}x = \frac{2}{\pi}\int_{-\infty}^{+\infty}\left(\int_{0}^{+\infty} e^{-xt}\sin x\,\mathrm{d}x\right)\mathrm{d}t = \frac{2}{\pi}\int_{-\infty}^{+\infty}\frac{1}{1+t^2}\,\mathrm{d}t = \frac{2}{\pi}\arctan t\Big|_{0}^{+\infty} = 1 \tag{11}$$

Observation 2.4. *Can it always be done?*
The conditions under which it is safe to swap the order of integration in a multiple integral are rather mild. In essence, we can always do it as far as the integral, as a whole, converges.

The Riemann-Lebesgue lemma In the case of Fourier Series we know that for a convergent series the coefficients f_k must approach 0 for $k \to \pm\infty$. The Riemann-Lebesgue lemma ensures that a similar result holds true in the continuous case or that

$$\lim_{\omega\to+\infty}\int_{a}^{b} f(t)\cos(\omega t)\mathrm{d}t = 0 \text{ and } \lim_{\omega\to+\infty}\int_{a}^{b} f(t)\sin(\omega t)\mathrm{d}t = 0$$

for all finite or infinite a and b. The proof is beyond the scope of this class. A hand-waving explanation is that for large values of ω, over each small interval around a point t, the functions $f(t)\cos(t)$ and $f(t)\sin(t)$ oscillate infinitely many times between $f(t)$ and $-f(t)$ thereby zeroing both integral values in the interval.

Another model for $\delta(t)$ We are now ready to show that

$$\lim_{\omega\to+\infty}\frac{\sin(\omega t)}{\pi t}=\delta(t)$$

Proof: We start by writing for some small $\epsilon>0$

$$\lim_{\omega\to+\infty}\int_{-\infty}^{+\infty}\frac{\sin(\omega t)}{\pi t}\phi(t)\mathrm{d}t=\lim_{\omega\to+\infty}\left(\int_{-\infty}^{-\epsilon}\frac{\sin(\omega t)}{\pi t}\phi(t)\mathrm{d}t+\int_{-\epsilon}^{+\epsilon}\frac{\sin(\omega t)}{\pi t}\phi(t)\mathrm{d}t+\int_{+\epsilon}^{+\infty}\frac{\sin(\omega t)}{\pi t}\phi(t)\mathrm{d}t\right)$$

In virtue of the *Riemann-Lebesgue* lemma the first and the third integral in the limit vanish. If we set $x=\omega t$ and take into account Equation (11), the middle integral writes

$$\lim_{\omega\to+\infty}\int_{-\epsilon}^{+\epsilon}\frac{\sin(\omega t)}{\pi t}\phi(t)\mathrm{d}t=\phi(0)\lim_{\omega\to+\infty}\int_{-\epsilon\omega}^{+\epsilon\omega}\frac{\sin x}{\pi x}\mathrm{d}x=\phi(0)\int_{-\infty}^{+\infty}\frac{\sin x}{\pi x}\mathrm{d}x=\phi(0)$$

■

In summary, the functions $r_T(t)$ for $T\to 0$ and $\sin(\omega t)/(\pi t)$ for $\omega\to\infty$ behave very similarly: they both approximate an impulse of unit area at the origin.

And yet another (very important) model We are not finished. We also have another way to obtain the δ function. Namely

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\cos(\omega t)\mathrm{d}\omega=\lim_{\Omega\to\infty}\int_{-\Omega}^{+\Omega}\frac{\cos(\omega t)}{2\pi}\mathrm{d}\omega=\lim_{\Omega\to\infty}\left.\frac{\sin(\omega t)}{2\pi t}\right|_{-\Omega}^{+\Omega}=\lim_{\Omega\to\infty}\frac{\sin(\Omega t)}{\pi t}=\delta(t)$$

Notice that the above integral exists only as a distribution.

Observation 2.5. *One more step*
The parity of the *cosine* and *sine* functions tells us that we also have

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{i\omega t}\mathrm{d}\omega=\frac{1}{2\pi}\left(\int_{-\infty}^{+\infty}\cos\omega t\mathrm{d}\omega+i\int_{-\infty}^{+\infty}\sin\omega t\mathrm{d}\omega\right)=\delta(t)+0=\delta(t)\tag{12}$$

Here again, the second integral exists and is equal to 0 only as a distribution.

2.2 A reversible transformation

We are now fully equipped to introduce the Fourier Transform.

Back and forth We define the Fourier Transform $F(\omega)$ of a function $f(t)$ as

$$F(\omega)=\int_{-\infty}^{+\infty}f(t)e^{-i\omega t}\mathrm{d}t.$$

The representation of $f(t)$ provided by $F(\omega)$ is *reversible* in the sense that we also have

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}F(\omega)e^{i\omega t}\mathrm{d}\omega$$

Indeed we have

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}F(\omega)e^{i\omega t}\mathrm{d}\omega=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty}f(t')e^{-i\omega t'}\mathrm{d}t'\right)e^{i\omega t}\mathrm{d}\omega\tag{13}$$

If we change the order of integration in Equation (13), and use Equation (12) and the sifting property, we finally obtain

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}F(\omega)e^{i\omega t}\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty}e^{i\omega(t-t')}\mathrm{d}\omega\right)f(t')\mathrm{d}t' = \int_{-\infty}^{+\infty}\delta(t-t')f(t')\mathrm{d}t' = f(t)$$

In this definition we used the angular frequency $\omega = 2\pi\nu$ (where ν is the true frequency).

Observation 2.6. *Existence*
If the function $f(t)$ is such that

$$\int_{-\infty}^{\infty}|f(t)|\mathrm{d}t < +\infty$$

then $F(\omega)$ exists as an ordinary function. Similarly to the Fourier Series case, the inversion formula converges point-wise at the points where f is continuous and to the average at the points of discontinuity. We will often consider cases in which the Fourier Transform and its inverse exist as distributions.

Observation 2.7. *Analogy with Fourier Series*
The analogy with the Fourier Series with complex exponentials is apparent. The Fourier Transform $F(\omega)$ corresponds to the coefficients c_n, while the Inverse Fourier Transform to the reconstruction of the signal as linear combination of the basis elements each multiplied by the corresponding coefficient. The sign change in the complex exponential between the Transform and the Inverse Transform completes the analogy.

2.3 Main properties

Linearity If $f(t) = ag(t) + bh(t)$, then $F(\omega) = aG(\omega) + bH(\omega)$. Immediate consequence of the definition of Fourier Transform as an integral

Conjugation If $f(t) \Leftrightarrow F(\omega)$, then $f^*(t) \Leftrightarrow F^*(-\omega)$. Since $zw^* = (z^*w)^*$, we have

$$\int_{-\infty}^{+\infty}f^*(t)e^{-i\omega t}\mathrm{d}t = \left(\int_{-\infty}^{+\infty}f(t)e^{i\omega t}\mathrm{d}t\right)^* = F^*(-\omega)$$

Duality If $f(t) \Leftrightarrow F(\omega)$, then $F(t) \Leftrightarrow 2\pi f(-\omega)$. Indeed viewing $F(t)$ has a Fourier Transform by setting $t = \Omega$ we obtain

$$\int_{-\infty}^{+\infty}F(t)e^{-i\omega t}\mathrm{d}t = \frac{2\pi}{2\pi}\int_{-\infty}^{+\infty}F(\Omega)e^{-i\omega\Omega}\mathrm{d}\Omega = 2\pi f(-\omega)$$

Time shift If $f(t) \Leftrightarrow F(\omega)$, then $f(t-t_0) \Leftrightarrow e^{-i\omega t_0}F(\omega)$. The substitution $t - t_0 \to t$ yields

$$\int_{-\infty}^{+\infty}f(t-t_0)e^{-i\omega t}\mathrm{d}t = \int_{-\infty}^{+\infty}f(t)e^{-i\omega(t+t_0)}\mathrm{d}t = e^{-i\omega t_0}\int_{-\infty}^{+\infty}f(t)e^{-i\omega t}\mathrm{d}t = e^{-i\omega t_0}F(\omega)$$

Frequency shift If $f(t) \Leftrightarrow F(\omega)$, then $e^{i\omega_0 t}f(t) \Leftrightarrow F(\omega-\omega_0)$. The substitution $\omega - \omega_0 \to \omega$ yields

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}F(\omega-\omega_0)e^{i\omega t}\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty}F(\omega)e^{i(\omega+\omega_0)t}\mathrm{d}\omega = e^{i\omega_0 t}\frac{1}{2\pi}\int_{-\infty}^{+\infty}F(\omega)e^{i\omega t}\mathrm{d}\omega = e^{i\omega_0 t}f(t)$$

Derivative If $f(t) \Leftrightarrow F(\omega)$, then $\mathrm{d}f/\mathrm{d}t \Leftrightarrow i\omega F(\omega)$. If F', the Fourier Transform of the derivative, exists we have

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} = \frac{1}{2\pi}\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{-\infty}^{+\infty}F(\omega)e^{i\omega t}\mathrm{d}\omega\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}i\omega F(\omega)e^{i\omega t}\mathrm{d}\omega$$

and thus $F'(\omega) = i\omega F(\omega)$.

Integral If $f(t) \Leftrightarrow F(\omega)$ and $F(0) = 0$, then $\int^t f(\tau)\,\mathrm{d}\tau \Leftrightarrow F(\omega)/(i\omega)$. Indeed, interchanging the order of integration we find

$$\int_{-\infty}^{t} f(\tau)\mathrm{d}\tau = \frac{1}{2\pi}\int_{-\infty}^{t}\int_{-\infty}^{+\infty} F(\omega)e^{i\omega\tau}\mathrm{d}\omega\mathrm{d}\tau = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)\int_{-\infty}^{t} e^{i\omega\tau}\mathrm{d}\tau\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{F(\omega)}{i\omega}e^{i\omega t}\mathrm{d}\omega.$$

Convolution If $h(t) = f * g$, then $H(\omega) = F(\omega)G(\omega)$. Since

$$h(t) = f * g = \int_{-\infty}^{+\infty} f(x)g(t-x)\mathrm{d}x,$$

interchanging the order of integration and using the time shift property we obtain

$$\begin{aligned} H(\omega) &= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} f(x)g(t-x)\mathrm{d}x\right)e^{-i\omega t}\,\mathrm{d}t = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} e^{-i\omega t}g(t-x)\,\mathrm{d}t\right)f(x)\,\mathrm{d}x \\ &= G(\omega)\int_{-\infty}^{+\infty} e^{-i\omega x}f(x)\,\mathrm{d}x = G(\omega)F(\omega) \end{aligned}$$

We also have that if $h(t) = f(t)g(t)$, then $2\pi H(\omega) = F * G$.

Parseval formula and Plancherel Theorem

$$\int_{-\infty}^{+\infty} f(t)g^*(t)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)G^*(\omega)\mathrm{d}\omega$$

Interchanging twice the order of integration, and using Equation (12) and the sifting property we find

$$\begin{aligned} \int_{-\infty}^{+\infty} f(t)g^*(t)\mathrm{d}t &= \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} F(\omega)e^{i\omega t}\mathrm{d}\omega\right)\left(\int_{-\infty}^{+\infty} G^*(-\omega')e^{i\omega' t}\mathrm{d}\omega'\right)\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(\omega)G^*(-\omega')\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i(\omega+\omega')t}\mathrm{d}t\right)\,\mathrm{d}\omega\,\mathrm{d}\omega' \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(\omega)G^*(-\omega')\delta(\omega+\omega')\,\mathrm{d}\omega\,\mathrm{d}\omega' \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} G^*(-\omega')\delta(\omega+\omega')\,\mathrm{d}\omega'\right)F(\omega)\,\mathrm{d}\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)G^*(\omega)\mathrm{d}\omega \end{aligned}$$

Setting $g(t) = f^*(t)$ Parseval formula becomes Plancherel theorem (aka Pythagora theorem), or

$$\int_{-\infty}^{+\infty} |f(t)|^2\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |F(\omega)|^2\mathrm{d}\omega$$

□

An important difference between Fourier series and Fourier Transform is that while the Fourier series of a real valued function is always real valued, the Fourier Transform of a real valued function f is, in general, complex. This can be readily seen applying the definition since we have $F(\omega) = R(\omega) - iX(\omega)$ with

$$R(\omega) = \int_{-\infty}^{+\infty} f(t)\cos(\omega t)\mathrm{d}t \quad \text{and} \quad X(\omega) = \int_{-\infty}^{+\infty} f(t)\sin(\omega t)\mathrm{d}t.$$

Exercise 2.1. *Real valued and pure imaginary Fourier Transform*
Which functions have real and purely imaginary Fourier Transform?

Observation 2.8. *Module and phase*
Module and phase of a Fourier Transform follow the usual definition. We thus have

$$F(\omega) = R(\omega) + iX(\omega) = A(\omega)e^{i\Phi(\omega)}$$

The module $A(\omega)$ tells you how much of the signal is captured at the angular frequency ω, the phase $\Phi(\omega)$ how the frequencies are lined up for reconstructing the signal.

Observation 2.9. *Dependencies*
Unless some prior information on the signal (like causality, for example) is available, the real and the imaginary part, or the module and phase of a Fourier Transform are independent.

Fact 2.1. *You cannot have both*
From the pairs we have seen, it appears that a function and its Fourier Transform exhibit a dual property. A function of finite support has a Fourier Transform with infinite support and the other way around.

2.4 Fourier Pairs

1. From
$$p_T(t) \Leftrightarrow 2\frac{\sin(\omega T)}{\omega}$$
and by interchanging the dependency on t and ω, we have
$$\frac{\sin(\omega_b t)}{\pi t} \Leftrightarrow p_{\omega_b}(\omega)$$
This pair yields the first example of a Fourier Transform with *limited* support in the frequency domain. Functions with Fourier Transform of finite support are called *band limited*. We will see that the band limited functions can always be sampled safely.

   Then, we consider a translation in time: $p_T(t) \to p_T(t - t_0)$. Using the translation property we immediately have
$$p_T(t - t_0) \Leftrightarrow \frac{2\sin(\omega T)}{\omega}e^{-i\omega t_0}$$
From this pair we can appreciate the meaning of module and phase. The module is left unchanged while the phase change, not surprisingly, is determined by t_0.

   Furthermore, since
$$p_T(t + 2T) = \frac{2\sin(\omega T)}{\omega}e^{i2\omega T} \quad \text{and} \quad p_T(t - 2T) = \frac{2\sin(\omega T)}{\omega}e^{-i2\omega T}$$
we have
$$p_T(t + 2T) + p_T(t - 2T) \Leftrightarrow \frac{4\sin(\omega T)}{\omega}\cos(2\omega T)$$
Since
$$p_{T/2}(t + T/2) = \frac{2\sin(\omega T/2)}{\omega}e^{i\omega T/2} \quad \text{and} \quad p_{T/2}(t - T/2) = \frac{2\sin(\omega T/2)}{\omega}e^{-i\omega T/2}$$
we also have
$$p_{T/2}(t + T/2) - p_{T/2}(t - T/2) \Leftrightarrow \frac{4i\sin^2(\omega T/2)}{\omega}$$

Finally, since

$$p_T(t)\cos(\omega_0 t) = p_T(t)\frac{e^{i\omega_0 t} + e^{-i\omega_0 t}}{2}$$

using the frequency shift property we find

$$p_T(t)\cos(\omega_0 t) \Leftrightarrow \frac{\sin[(\omega-\omega_0)t]}{(\omega-\omega_0)t} + \frac{\sin[(\omega+\omega_0)t]}{(\omega+\omega_0)t}$$

2. We now prove that

$$q_T(t) \Leftrightarrow 4\frac{\sin^2(\omega T/2)}{T\omega^2}$$

We consider the triangle function $q_T(t)$ defined as

$$q_T(t) = \begin{cases} 1 + t/T & -T \leq t \leq 0 \\ 1 - t/T & 0 < t \leq T \\ 0 & \text{otherwise.} \end{cases}$$

Since

$$q_T(t) = \frac{1}{T}\int_{-\infty}^{t} \left(p_{T/2}(\tau + T/2) - p_{T/2}(\tau - T/2)\right) \mathrm{d}\tau$$

the pair $q_T(t) \Leftrightarrow 4\sin^2(\omega T/2)/(T\omega^2)$ follows from the integral property.

Applying the symmetry property $F(t)/(2\pi) \Leftrightarrow f(-\omega)$ to this pair with $T = 2a$ we obtain

$$\frac{\sin^2(at)}{\pi a t^2} \Leftrightarrow q_{2a}(\omega)$$

3. Furthermore, since

$$\int_{-\infty}^{+\infty} e^{-\alpha t}U(t)e^{-i\omega t}\mathrm{d}t = \int_{0}^{+\infty} e^{-(\alpha+i\omega)t}\mathrm{d}t = \left.\frac{-e^{-(\alpha+i\omega)t}}{\alpha+i\omega}\right|_{0}^{+\infty} = \frac{1}{\alpha+i\omega}$$

we have that

$$e^{-\alpha t}U(t) \Leftrightarrow \frac{1}{\alpha+i\omega}$$

For the even part of $e^{-\alpha t}U(t)$ we have $f_e(t) = e^{-\alpha|t|}/2$. Since $f_e(t) \Leftrightarrow R(\omega)$, the Fourier Transform of $e^{-\alpha|t|}$ is simply twice the real part of $1/(\alpha + i\omega)$, or

$$2\mathrm{Re}\left(\frac{1}{\alpha+i\omega}\right) = \frac{2\alpha}{\alpha^2+\omega^2}$$

4. Finally, let us show that

$$e^{-t^2/2\sigma^2} \Leftrightarrow \sqrt{2\pi}\sigma e^{-\omega^2\sigma^2/2}$$

Let $g(t) = e^{-t^2/2\sigma^2}$. For the Fourier Transform we have

$$G(\omega) = \int_{-\infty}^{+\infty} e^{-t^2/2\sigma^2}e^{-i\omega t}\mathrm{d}t$$

taking the derivative of $G(\omega)$ with respect to ω we obtain

$$\frac{\mathrm{d}G(\omega)}{\mathrm{d}\,\omega} = \int_{-\infty}^{+\infty} (-it)e^{-t^2/2\sigma^2}e^{-i\omega t}\mathrm{d}t$$

If we integrate by parts the integral in the right hand side with

$$u(t) = e^{-i\omega t} \quad \text{and} \quad v(t) = ie^{-t^2/(2\sigma^2)}$$

since

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = -i\omega e^{-i\omega t} \quad \text{and} \quad \frac{\mathrm{d}v(t)}{\mathrm{d}t} = \frac{-it}{\sigma^2} e^{-t^2/(2\sigma^2)}$$

and $v(t) \to 0$ for $t \to \pm\infty$ with $u(t)$ in $[-1, 1]$, we find

$$\frac{\mathrm{d}G(\omega)}{\mathrm{d}\,\omega} = \sigma^2 \int_{-\infty}^{+\infty} \frac{\mathrm{d}v}{\mathrm{d}t} u \mathrm{d}t = uv\Big|_{-\infty}^{+\infty} - \sigma^2 \int_{-\infty}^{+\infty} \frac{\mathrm{d}u}{\mathrm{d}t} v \mathrm{d}t = -\omega\sigma^2 \int_{-\infty}^{+\infty} e^{-t^2/2\sigma^2} e^{-i\omega t} \mathrm{d}t = -\omega\sigma^2 G(\omega)$$

The general solution to this differential equation, for some constant $K \in \mathbb{R}$, is

$$G(\omega) = Ke^{-\omega^2\sigma^2/2}$$

But we know that

$$G(0) = \int_{-\infty}^{+\infty} e^{-t^2/(2\sigma^2)} \mathrm{d}t = \sigma\sqrt{2\pi}$$

therefore $K = G(0) = \sigma\sqrt{2\pi}$ and we finally have

$$G(\omega) = \sigma\sqrt{2\pi} e^{-\omega^2\sigma^2/2}$$

Observation 2.10. *The Gaussian is an eigenfunction of the Fourier operator*
The Fourier Transform of a Gaussian is a Gaussian! If σ^2 is the variance in time, $1/\sigma^2$ is the variance in frequency. Therefore the Fourier Transform of a Gaussian peaked around the origin in space is a shallow Gaussian in the frequency domain. *Viceversa*, the Fourier Transform of a shallow Gaussian in the spatial domain is a Gaussian peaked in the low frequency range.

2.5 Sampling

Infinite Impulse Train An infinite impulse train with period T_s can be written as

$$i(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT_s) \tag{14}$$

Setting $\omega_s = 2\pi/T_s$ we can expand $i(t)$ in its Fourier Series in the interval $[-T_s/2, T_s/2]$ and obtain

$$i(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ikt\omega_s}$$

Using Equation (14) and observing that the only relevant δ in $[-T_s/2, T_s/2]$ is obtained for $n = 0$, for the coefficient c_k we find

$$c_k = \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} i(t) e^{-ikt\omega_s} \mathrm{d}t = \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} \delta(t) e^{-ikt\omega_s} \mathrm{d}t = \frac{1}{T_s}$$

and thus we have

$$i(t) = \frac{1}{T_s} \sum_{k=-\infty}^{+\infty} e^{ikt\omega_s}$$

If we now take the Fourier Transform of $i(t)$ and interchange the integral with the series we obtain

$$I(\omega) = \frac{1}{T_s} \int_{-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} e^{ikt\omega_s} e^{-i\omega t} \mathrm{d}t = \frac{2\pi}{T_s} \sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_s) = \omega_s \sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_s)$$

Observation 2.11. *Duality*
The Fourier Transform of a periodic impulse train is a periodic impulse train. Interestingly, if the impulses are T_s far apart in time, the transformed impulses are $2\pi/T_s$ far apart in frequency.

What really happens We are now ready to analyse what happens when a signal is sampled over time. Let $s_{smp}(t)$ be the sampling of $s(t)$ at a discrete infinite sequence of equidistant times t_n with $n \in (-\infty, \infty)$. By definition of sampling we have

$$s_{smp}(t) = \begin{cases} s(t) & t = t_n \\ 0 & \text{otherwise} \end{cases}$$

We set $t_n = nT_s$ for some fixed time interval T_s and we model the sampled signal s_{smp} through a train of δ functions

$$s_{smp}(t) = s(t)i(t) = \sum_{n=-\infty}^{+\infty} s(t)\delta(t - nT_s) = \sum_{n=-\infty}^{+\infty} s(nT_s)\delta(t - nT_s).$$

Via the convolution property we have

$$S_{smp}(\omega) = \frac{1}{2\pi} S * I = \frac{1}{2\pi}\omega_s \sum_{k=-\infty}^{+\infty} \int_{-\infty}^{+\infty} S(\omega')\delta(\omega - k\omega_s - \omega')\mathrm{d}\omega' = \frac{1}{T_s}\sum_{k=-\infty}^{+\infty} S(\omega - k\omega_s) \quad (15)$$

Observation 2.12. *Aliasing*
From Equation (15) we see that $S_{smp}(\omega)$ is the sum of equidistant copies of $S(\omega)$, ω_s far apart. Consequently, if $\mathcal{L}(S)$ - length of the smallest interval over which $S(\omega) \neq 0$ - is larger than ω_s, $S_{smp}(\omega)$ and $S(\omega)$ over the interval $[-\omega_s/2, \omega_s/2]$ are different. This phenomenon is known as *aliasing* because higher frequencies of $S(\omega)$ are disguised as lower frequencies of $S_{smp}(\omega)$, see Figure (2). $\square$

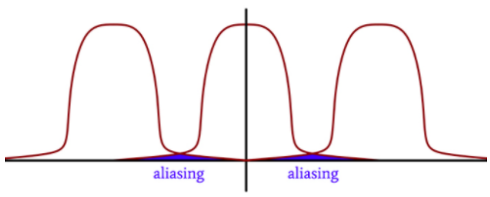

Figure 2: See text.

We are now ready to state and prove the Shannon Sampling Theorem, the cornerstone of digital signal processing. The Theorem establishes the condition under which the sampling of a signal does not induce aliasing.

Theorem 2.1. *Sampling theorem*
Let $S(\omega)$ be the Fourier Transform of a signal $s(t)$ with $S(\omega) = 0$ for $|\omega| \geq \omega_b$ (we thus know that $\mathcal{L}(S) \leq 2\omega_b$). If $t_n = n\pi/\omega_b$ with $n \in (-\infty, \infty)$ are equidistant times and $s_{smp}(t)$ the sampling of $s(t)$ at the times t_n, then $s(t)$ can be reconstructed *exactly* for all t from $s_{smp}(t)$ through the formula

$$s(t) = \sum_{n=-\infty}^{+\infty} s_{smp}(t_n)\frac{\sin(\omega_b(t - t_n))}{\omega_b(t - t_n)}.$$

As usual whenever we write $(\sin x)/x$ we assume that for $x = 0$ $(\sin x)/x = 1$.
Proof
From the hypothesis we have that for the Inverse Fourier Transform of $S(\omega)$ we have

$$s(t) = \frac{1}{2\pi}\int_{-\omega_b}^{\omega_b} S(\omega)e^{i\omega t}\mathrm{d}\omega \tag{16}$$

For $t_n = n\pi/\omega_b$, therefore, Equation (16) writes

$$s_{smp}(t_n) = s(t_n) = s\left(n\frac{\pi}{\omega_b}\right) = \frac{1}{2\pi}\int_{-\omega_b}^{\omega_b} S(\omega)e^{in\pi\omega/\omega_b}\mathrm{d}\omega \tag{17}$$

On the other hand, the Fourier Series expansion of $S(\omega)$ in the interval $[-\omega_b, \omega_b]$ is

$$S(\omega) = \sum_{n=-\infty}^{+\infty} c_n e^{in\pi\omega/\omega_b}$$

with

$$c_n = \frac{1}{2\omega_b}\int_{-\omega_b}^{\omega_b} S(\omega)e^{-in\pi\omega/\omega_b}\mathrm{d}\omega \tag{18}$$

Comparing Equation (17) and (18) we find

$$c_n = \frac{\pi}{\omega_b}s\left(-n\frac{\pi}{\omega_b}\right) = \frac{\pi}{\omega_b}s(t_{-n}) = \frac{\pi}{\omega_b}s_{smp}(t_{-n})$$

Absorbing the change of sign we write

$$S_{smp}(\omega) = \sum_{n=-\infty}^{\infty}\frac{\pi}{\omega_b}s(t_{-n})e^{in\pi\omega/\omega_b} = \frac{\pi}{\omega_b}\sum_{n=-\infty}^{\infty} s_{smp}(t_n)e^{-in\pi\omega/\omega_b}$$

The function $S_{smp}(\omega)$ is periodic of period $2\omega_b$ with $S_{smp}(\omega) = S(\omega)$ for $\omega \in [-\omega_b, \omega_b]$. Therefore, we can write

$$S(\omega) = p_{\omega_b}S_{smp}(\omega) = \frac{\pi}{\omega_b}\sum_{n=-\infty}^{\infty} s_{smp}(t_n)p_{\omega_b}(\omega)e^{-in\pi\omega/\omega_b}.$$

Finally, from the pair

$$\frac{\omega_b}{\pi}\frac{\sin(\omega_b t - n\pi)}{\omega_b t - n\pi} \Leftrightarrow p_{\omega_b}(\omega)e^{-in\pi\omega/\omega_b}$$

and the convolution property we find

$$s(t) = \sum_{n=-\infty}^{+\infty} s_{smp}(t_n)\frac{\sin(\omega_b(t - t_n))}{\omega_b(t - t_n)} \tag{19}$$

■

Observation 2.13. *Nyquist frequency (beware of a factor 2...)*
The frequency ω_b/π is the celebrated Nyquist frequency. Its inverse, π/ω_b is time lag between consecutive samples. Sampling at the inverse of the Nyquist frequency or higher allows for perfect signal reconstruction. Even if you forget the proof of Sampling Theorem, all what you need to remember is the duality between an infinite impulse train in time with time lag equal to T_s and its Fourier Transform, an infinite impulse train in frequency with frequency lag equal to $2\pi/T_s$. Aliasing is avoided if

$$T_s \leq \frac{\pi}{\omega_b} = \frac{2\pi}{\mathcal{L}(S)}$$

Notice that $\omega_b = \mathcal{L}(S)/2$!

2.6 Things you need to know

1. The Fourier Transform $F(\omega)$ of a real valued function $f : (-\infty, +\infty) \to \mathbb{R}$, the absolute value of which is integrable, is
$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\mathrm{d}t$$

2. The original signal $f(t)$ can be reconstructed from the Fourier Transform $F(\omega)$ as
$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\mathrm{d}\omega$$

3. How to obtain the real part, the imaginary part, the module and the phase of a Fourier Transform.

4. How to derive the properties of linearity, conjugation, duality, time shift, frequency shift, and convolution and the Plancherel theorem

5. That the Fourier Transform of a train of impulses is a train of impulses, statement of the sampling theorem, the Nyquist sampling frequency and being able to explain the aliasing effect.

3 Linear Time-Invariant Systems

Here again, we took inspiration from [6].

3.1 Definition

A physical system $\mathcal{L}$ is a transducer which, for a signal $f_i(t)$ in input, produces a signal $f_o(t)$ in output. The output signal $f_o(t)$, also called *effect*, is entirely determined by the input signal $f_i(t)$, also called *cause*. The system $\mathcal{L}$ can be thought of as a transformation with

$$\mathcal{L}\{f_i(t)\} = f_o(t)$$

We immediately restrict our attention to systems which satisfy two constraints.

Linearity A system $\mathcal{L}$ is *linear* if for all possible input signals $f_i(t)$ and $g_i(t)$ from

$$\mathcal{L}\{f_i(t)\} = f_o(t) \quad \text{and} \quad \mathcal{L}\{g_i(t)\} = g_o(t)$$

it follows that

$$\mathcal{L}\{af_i(t) + bg_i(t)\} = af_o(t) + bg_o(t)$$

.

Time-invariance A linear system $\mathcal{L}$ is *time-invariant* if $\forall s \in \mathbb{R}$ and input signal $f_i(t)$

$$\mathcal{L}\{f_i(t)\} = f_o(t) \quad \Rightarrow \quad \mathcal{L}\{f_i(t-s)\} = f_o(t-s).$$

In words, the parameters of a linear, time-invariant system do not change over time.

3.2 Characterisation of an LTI system

Impulse response Let $h(t)$ be the response of a *linear time-invariant* (LTI) system to an impulse $\delta(t)$. We have

$$\mathcal{L}\{\delta(t-s)\} = h(t-s).$$

The power of convolution If we write a signal $f_i(t)$ as an infinite sum of impulses,

$$f_i(t) = \int_{-\infty}^{\infty} f_i(s)\delta(t-s)\mathrm{d}s,$$

and use the linearity of the integral and the time-invariance property we have

$$f_o(t) = \mathcal{L}\{f_i(t)\} = \int_{-\infty}^{\infty} f_i(s)\mathcal{L}\{\delta(t-s)\}\mathrm{d}s = \int_{-\infty}^{\infty} f_i(s)h(t-s)\mathrm{d}s \tag{20}$$

Equation (20) provides a powerful characterisation of an LTI system. The output of an LTI system can be obtained as the *convolution of the input with the impulse response*. Remarkably, all what is needed to compute the output $f_o(t)$ of a linear time-invariant system for a given input $f_i(t)$ is the knowledge of just one function of time: the impulse response $h(t)$.

3.3 Properties

Stability A system is *stable* if its response to a bounded input is bounded, or if

$$|f_i(t)| < M \quad \Rightarrow \quad |f_o(t)| < MI$$

where I is a constant independent of the input.

Fact 3.1. *The integrability of the absolute value of the impulse response implies stability*

$$I = \int_{-\infty}^{\infty} |h(t)|dt < \infty \quad \Rightarrow \quad |f_o(t)| \leq \int_{-\infty}^{\infty} |f_i(t-s)h(s)|ds < M \int_{-\infty}^{\infty} |h(s)|ds = MI$$

■

Causality A system is *causal* if when $f_i(t) = 0$ for $t < t_0$ we also have that $f_o(t) = 0$ for $t < t_0$. For a causal LTI system $\mathcal{L}$ we thus find

$$f_o(t) = \mathcal{L}\{f_i(t)\} = 0 \quad \text{for } \ t < t_0.$$

Fact 3.2. *The impulse response of a causal LTI system is causal and viceversa.*
In the Fourier class we saw that a function $f(t)$ is *causal* if $f(t) = 0$ for $t < 0$.

$\rightarrow$ We know that if $f_i(t) = 0$ for $t < t_0$, then $f_o(t) = 0$ for $t < t_0$. Now, if we rewrite the convolution of Equation (20) as

$$f_o(t) = \mathcal{L}\{f_i(t)\} = \int_{-\infty}^{\infty} f_i(s)h(t-s)\mathrm{d}s$$

we see that in order to have $f_o(t) = 0$ for $t < t_0$ we need $h(t_0 - s) = 0$ for $s > t_0$, or that $h(t) = 0$ for $t < 0$.

$\leftarrow$ We know that $h(t) = 0$ for $t < 0$. Therefore, we can write

$$f_o(t) = \int_{-\infty}^{\infty} f_i(s)h(t-s)\mathrm{d}s = \int_{-\infty}^{t} f_i(s)h(t-s)\mathrm{d}s.$$

If for $t < t_0$ $f_i(t) = 0$ we immediately see that for $t < t_0$, since $s < t$, the integrand vanishes identically in the integration interval and $f_o(t) = 0$. ■

Eigenfunctions and eigenvalues Let us consider the output of LTI system $\mathcal{L}$ for an exponential input $f_i(t) = e^{i\omega_0 t}$. We have

$$\mathcal{L}\{e^{i\omega_0 t}\} = f_o(t)$$

for some function $f_o(t)$. For the time-invariance we can write

$$\mathcal{L}\{e^{i\omega_0 (t+s)}\} = f_o(t+s).$$

For the linearity and the fact that $e^{i\omega_0(t+s)} = e^{i\omega_0 t}e^{i\omega_0 s}$ we have

$$\mathcal{L}\{e^{i\omega_0(t+s)}\} = \mathcal{L}\{e^{i\omega_0 t}e^{i\omega_0 s}\} = e^{i\omega_0 s}\mathcal{L}\{e^{i\omega_0 t}\} = e^{i\omega_0 s}f_o(t).$$

Therefore we conclude that

$$f_o(t+s) = e^{i\omega_0 s}f_o(t)$$

If we set $t = 0$, $f_o(0) = k$, and observe that s is arbitrary we obtain (with $s \rightarrow t$)

$$f_o(t) = ke^{i\omega_0 t}. \tag{21}$$

We thus find that the complex exponential is an eigenfunction of $\mathcal{L}$, since the output to a complex exponential input is the same complex exponential. We now express the eigenvalue k in terms of the Fourier Transform of the impulse response.

System function Through the convolution property of the Fourier Transform we have that taking the Fourier Transform of Equation (20) we find

$$F_o(\omega) = F_i(\omega)H(\omega).$$

The Fourier Transform $H(\omega)$ of the impulse response $h(t)$ is known as *system function*.

Fact 3.3. *Eigenvalue characterisation*
Let $f_i(t) = e^{i\omega_0 t}$. Since the Fourier Transform of $f_i(t) = e^{i\omega_0 t}$ is $F_i(\omega) = 2\pi\delta(\omega - \omega_0)$ we find

$$F_o(\omega) = 2\pi\delta(\omega - \omega_0)H(\omega) = 2\pi\delta(\omega - \omega_0)H(\omega_0).$$

Therefore, for the Inverse Fourier Transform we find

$$f_o(t) = \frac{1}{2\pi}H(\omega_0)\int_{-\infty}^{\infty} 2\pi\delta(\omega - \omega_0)e^{i\omega t}\mathrm{d}\omega = H(\omega_0)e^{i\omega_0 t}. \tag{22}$$

Hence from Equation (21) and (22) we find

$$k = H(\omega_0). \tag{23}$$

Observation 3.1. *Consequences*
We already knew that the output of a sinusoidal input is a sinusoid with the same angular frequency ω_0. From Equation (21) we can also compute the attenuation factor $A(\omega_0) = |H(\omega_0)|$ and the phase Φ as in $H(\omega_0) = A(\omega_0)e^{i\Phi(\omega_0)}$. For $f_i(t) = \cos(\omega_0 t)$, for example, we have

$$f_o(t) = \mathcal{L}\{\cos(\omega_0 t)\} = A(\omega_0)\cos(\omega_0 t + \Phi(\omega_0))$$

All input sinusoids maintain the same frequency on output but are *attenuated* by a frequency dependent factor and *delayed* by a frequency dependent phase!

Observation 3.2. *Constant input*
If $\omega_0 = 0$, we find that the output signal of a constant input signal, $f_i(t) = \bar{f}_i$ is constant with

$$f_o(t) = \bar{f}_o = H(0)\bar{f}_i \tag{24}$$

A second characterisation Since $U(t) = \int_{-\infty}^{t}\delta(s)\mathrm{d}s$ and $h(t) = 0$ for $t < 0$, we have that

$$a(t) = \int_0^t h(s)\mathrm{d}s.$$

For a causal system, it is sometimes easier to compute the response $a(t)$ to the unit step $U(t)$,

$$\mathcal{L}\{U(t)\} = a(t).$$

We start by writing the input signal $f_i(t)$ as

$$f_i(t) = f_i(-\infty) + \int_{-\infty}^{t}\frac{\mathrm{d}f_i(s)}{\mathrm{d}s}\mathrm{d}s = f_i(-\infty) + \int_{-\infty}^{\infty}\frac{\mathrm{d}f_i(s)}{\mathrm{d}s}U(t-s)\mathrm{d}s$$

Thus, by the linearity of $\mathcal{L}$ and using Equation (24), we obtain

$$\begin{aligned} f_o(t) = \mathcal{L}\{f_i(t)\} &= \mathcal{L}\{f_i(-\infty)\} + \int_{-\infty}^{\infty}\frac{\mathrm{d}f_i(s)}{\mathrm{d}s}\mathcal{L}\{U(t-s)\}\mathrm{d}s \\ &= H(0)f_i(-\infty) + \int_{-\infty}^{\infty}\frac{\mathrm{d}f_i(s)}{\mathrm{d}s}a(t-s)\mathrm{d}s \\ &= H(0)f_i(-\infty) + \int_{-\infty}^{t}\frac{\mathrm{d}f_i(s)}{\mathrm{d}s}a(t-s)\mathrm{d}s \end{aligned}$$

since, clearly, $a(t)$ is a causal function.

Let us know discuss an example of a stable and causal LTI system.

3.4 RC circuit

We consider the circuit in Figure (3). All what you are supposed to remember about the physics of a circuit are two facts: *(a)* when you apply a voltage V_{in}, a current will start to flow across the circuit, and *(b)* the same current, with no leakage, flows across the entire circuit.

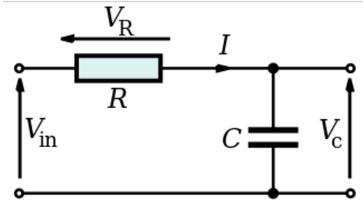

Figure 3: See text.

An equation involving the derivative of a function We have that

$$V_R + V_C = V_{in}$$

From Ohm's law, we know that for the voltage drop between the resistor ends we have $V_R = R\,\mathrm{d}q/\mathrm{d}t$, while the voltage at the plates of a capacitor is inversely proportional to the capacitance $V_C = q/C$. We thus have

$$R\,\frac{\mathrm{d}q(t)}{\mathrm{d}t} + \frac{1}{C}\,q(t) = V_{in}(t)$$

If we set $V_C = q(t)/C = V_{out}(t)$ we find

$$\frac{\mathrm{d}V_{out}(t)}{\mathrm{d}t} + \frac{1}{\tau}\,V_{out}(t) = \frac{1}{\tau}\,V_{in}(t) \tag{25}$$

where we introduce the time constant $\tau = RC$. In order to write $V_{out}(t)$ in terms of $V_{in}(t)$ we need to solve the differential Equation (25) for $V_{out}(t)$.

The trick is to multiply each term by a differentiable function $e^{t/\tau}$

$$\frac{\mathrm{d}V_{out}(t)}{\mathrm{d}t}e^{t/\tau} + \frac{1}{\tau}V_{out}(t)e^{t/\tau} = \frac{1}{\tau}V_{in}(t)e^{t/\tau} \tag{26}$$

and observe that the left hand side of Equation (26) can be written as a derivative. Indeed, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(V_{out}(t)e^{t/\tau}\right) = \frac{\mathrm{d}V_{out}(t)}{\mathrm{d}t}e^{t/\tau} + \frac{1}{\tau}V_{out}(t)e^{t/\tau}.$$

A causal, stable LTI system Therefore, integrating both sides of Equation (26) we find

$$V_{out}(t)e^{t/\tau} = \frac{1}{\tau}\int_{-\infty}^{t} e^{s/\tau}V_{in}(s)\mathrm{d}s \quad \rightarrow \quad V_{out}(t) = \frac{1}{\tau}\int_{-\infty}^{t} e^{-(t-s)/\tau}V_{in}(s)\mathrm{d}s \quad \text{for } t \geq s$$

We thus obtain for the impulse response

$$h(t) = \frac{1}{\tau}e^{-t/\tau}U(t)$$

where $U(t)$ ensures that $h(t) = 0$ for $t < 0$. Intuitively, the capacitor discharges at an exponential rate depending on the characteristic time $\tau = RC$.

The system is clearly linear, time-invariant, and causal. Stability is guaranteed by its physical nature.

Unit step response We have

$$a(t) = \frac{e^{-t/\tau}}{\tau} \int_0^t e^{s/\tau} \mathrm{d}s = \frac{e^{-t/\tau}}{\tau} \tau e^{s/\tau} \Bigg|_{s=0}^{s=t} = \left(1 - e^{-t/\tau}\right) \quad \text{for } t \geq 0$$

The constant τ is the characteristic time needed by the system, see Figure (4), to respond to the unit step input.

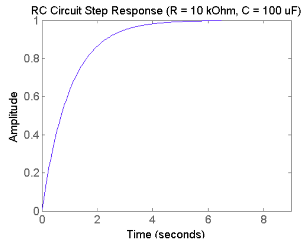

Figure 4: See text.

Impulse response The impulse response can also be found by taking the derivative of the unit step response. Indeed we have

$$h(t) = \frac{\mathrm{d}}{\mathrm{d}t} \left(1 - e^{-t/\tau}\right) = \frac{1}{\tau} e^{-t/\tau} U(t)$$

System function For the Fourier transform of $h(t)$ we have

$$H(\omega) = \frac{1}{\tau} \frac{1}{1/\tau + i\omega} = \frac{1}{1 + i\tau\omega} = A(\omega) e^{i\Phi(\omega)} = \frac{1}{\sqrt{1 + \tau^2\omega^2}} e^{i\tau\omega}$$

This is a first example of filter. The shape of both the unit step response and the impulse response can be understood in terms of the action of the system function on the Fourier Transform of $U(t)$ and $\delta(t)$. The higher the frequency, the stronger the attenuation in the module $A(\omega)$. Here again, the constant τ is the characteristic time. The larger τ the more visible is the filter effect (for the same frequency). The phase $\Phi(\omega)$, instead, introduces distortion in the reconstruction.

3.5 Things you need to know

1. Make sure you are able to write down and discuss the linearity and the time-invariance property of a system

2. Make sure you know how to restate the convolution property which characterises a linear time-invariant system in the frequency domain

3. What is the response of a linear time-invariant system to a sinusoidal input

4. What is a linear filter, and low-pass and high-pass filter in particular.

4 Kalman Filtering

The filters we discussed so far do not take into account the fact that, in time-varying signal, future samples are often correlated with the current sample. We now present a scheme, Kalman filtering, in which this correlation is exploited to design optimal filters which can also be implemented efficiently. As usual the topic is vast and in class we only cover the fundamentals. We invariably assume the discrete time setting. We are partially indebted to [7]. For a thorough reference on the subject see [5].

4.1 Computing the average revisited

The scenario Let us assume we want to estimate the height of a building. We refer to s as the unknown quantity to be estimated, the true height, and view it as the *hidden state* of a certain system, the building (in the far-fetched assumption that all what we are interested in is the building height). We do not have direct access to the state but at each time $t = 1, \dots$ we acquire a measurement, m_t. The problem is to obtain the optimal estimate $\hat{s}_t$ of s. The $\hat{}$ sign tells us that $\hat{s}_t$ is an *estimate* of s, while the subscript $_t$ that the t-th measurement has been acquired. Here the hidden state s does not change over time.

Known answer We already know the solution we should find: the best we can do is to compute the empirical average of the measurements acquired up to time t,

$$\hat{s}_t = \frac{1}{t}\sum_{i=1}^{t} m_i$$

Problem setting Let us derive this result in the *Kalman filtering* setting. We write

$$m_t = s + r_t \tag{27}$$

where r_t is the noise affecting the measurement m_t. We assume that the noise is zero mean, $\mathbb{E}[r_t] = 0$, with variance $\mathbb{E}[r_t^2] = \mathrm{R} > 0$.

If we let $\hat{s}_t^-$ denote the hidden state estimate we predict *before* the acquisition of the t-th measurement, we write the **state update** equation as

$$\hat{s}_t = \hat{s}_t^- + \mathrm{K}_t(m_t - \hat{s}_t^-) \tag{28}$$

and aim at finding K_t, the *Kalman gain*, which allows us to recursively obtain the optimal estimate $\hat{s}_t$ as the weighted sum of $\hat{s}_t^-$, and the difference between the new measurement m_t and $\hat{s}_t^-$, called *innovation* or *measurement residual*. If $e_t = s - \hat{s}_t$ is the *error*, we look for the Kalman gain by minimising the *mean squared error*, or variance of the state estimate *after* the t-th measurement is acquired,

$$\mathbb{E}[e_t^2] = \mathbb{E}[(s - \hat{s}_t)^2]$$

Solution Using Equation (27) and (28) we have

$$e_t = s - \hat{s}_t = s - \hat{s}_t^- - \mathrm{K}_t h - \mathrm{K}_t r_t + \mathrm{K}_t \hat{s}_t^- = (1 - \mathrm{K}_t)(s - \hat{s}_t^-) - \mathrm{K}_t r_t$$

and thus

$$e_t^2 = (1 - \mathrm{K}_t)^2 (s - \hat{s}_t^-)^2 + \mathrm{K}_t^2 r_t^2 - 2\mathrm{K}_t(1 - \mathrm{K}_t) r_t (s - \hat{s}_t^-)$$

We now consider what happens in expectation. Taking into account that the error in the state estimate and the noise measurement are uncorrelated, the expected value of the last term in the right hand side of the previous equality vanishes and we are left with

$$\mathbb{E}[e_t^2] = (1 - \mathrm{K}_t)^2 \mathbb{E}[(s - \hat{s}_t^-)^2] + \mathrm{K}_t^2 \mathbb{E}[r_t^2] \tag{29}$$

If we denote with P_t the variance $\mathbb{E}[e_t^2] = \mathbb{E}[(s-\hat{s}_t)^2]$ and let $\mathrm{P}_t^- = \mathbb{E}[(s-\hat{s}_t^-)^2]$ for the variance of the state estimate predicted *before* the t-th measurement has been acquired, since $\mathbb{E}[r_t^2] = \mathrm{R}$, we rewrite Equation (29) as

$$\mathrm{P}_t = \mathrm{P}_t^- + \mathrm{K}_t^2(\mathrm{P}_t^- + \mathrm{R}) - 2\mathrm{K}_t\mathrm{P}_t^- \tag{30}$$

Setting the derivative of P_t in Equation (30) with respect to K_t equal to 0 we obtain the **Kalman gain** minimising the mean squared error P_t as

$$\frac{\mathrm{dP}_t}{\mathrm{dK}_t} = (\mathrm{K}_t - 1)\mathrm{P}_t^- + \mathrm{K}_t\mathrm{R} = 0 \qquad \rightarrow \qquad \mathrm{K}_t = \frac{\mathrm{P}_t^-}{\mathrm{P}_t^- + \mathrm{R}}$$

By substituting this expression for K_t in Equation (28) we are now in a position to compute the state update following the acquisition of the t-th measurement. In order to complete the derivation we still need to update the variance, and project the state and variance estimates from t to $t+1$. For the **variance update**, by plugging the expression for K_t into Equation (30), we obtain

$$\mathrm{P}_t = \mathrm{P}_t^- + \frac{(\mathrm{P}_t^-)^2}{\mathrm{P}_t^- + \mathrm{R}} - 2\frac{(\mathrm{P}_t^-)^2}{\mathrm{P}_t^- + \mathrm{R}} = \mathrm{P}_t^-(1 - \mathrm{K}_t)$$

For the **state estimate projection**, since we know that s is not changing over time, we simply have

$$\hat{s}_{t+1}^- = \hat{s}_t \tag{31}$$

For the **variance estimate projection**, using Equations (31), we obtain

$$\mathrm{P}_{t+1}^- = \mathbb{E}[(s - \hat{s}_{t+1}^-)^2] = \mathbb{E}[(s-\hat{s}_t)^2] = \mathrm{P}_t$$

In summary, the solution to our problem - relying on the prior knowledge of the measurement noise variance R - consists of 7 steps.

Algorithm 4.1. *FixedStateKalman*
Acquire the measurement m_1 and set

$$\hat{s}_2^- = \hat{s}_1 = m_1, \quad \text{and} \quad P_2^- = P_1 = \mathrm{R}$$

Set $t = 2$ and repeat

1.	Acquire the t-th measurement:	m_t
2.	Set the Kalman gain:	$\mathrm{K}_t = \mathrm{P}_t^-/(\mathrm{P}_t^- + \mathrm{R})$
3.	State update:	$\hat{s}_t = \hat{s}_t^- + \mathrm{K}_t(m_t - \hat{s}_t^-)$
4.	Variance update:	$\mathrm{P}_t = \mathrm{P}_t^-(1 - \mathrm{K}_t)$
5.	State estimate projection:	$\hat{s}_{t+1}^- = \hat{s}_t$
6.	Variance estimate projection:	$\mathrm{P}_{t+1}^- = \mathrm{P}_t$
7.	$t \leftarrow t+1$	

Exercise 4.1. *Checking the obtained result*
Go through a few iterations of the *FixedStateKalman* algorithm and verify that

- for the Kalman gain we get $\mathrm{K}_t = 1/t$
- the state update is the recursive computation of the empirical average, and
- the variance update is $\mathrm{P}_t = \mathrm{R}/t$

4.2 General case

We now generalise what we have seen so far in several directions. First, we consider a system described by an N-dimensional hidden state vector $\mathbf{s}_t$ where each component changes, or might change, with t. As before, we do have some knowledge about the way the state evolves over time.

Definition 4.1. *Linear (stationary) process model*
A *linear process model* expresses $\mathbf{s}_{t+1}$ in terms of $\mathbf{s}_t$ through an $N \times N$ *transition* matrix $\mathbf{\Phi}$

$$\mathbf{s}_{t+1} = \mathbf{\Phi}\mathbf{s}_t + \mathbf{q}_t \tag{32}$$

where the N-dimensional vector $\mathbf{q}_t$ is the *process noise*.

Definition 4.2. *Process noise expected value and covariance*
We assume that $\mathbb{E}[\mathbf{q}_t] = 0$ and denote the $N \times N$ *process noise covariance* matrix with

$$\mathbf{Q} = \mathbb{E}[\mathbf{q}_t\mathbf{q}_t^\top]$$

Observation 4.1. *Back to the average*
In our first, very simple example the state $\mathbf{s}_t$, the process noise $\mathbf{q}_t$, and the covariance matrix $\mathbf{Q}$ were scalar with no dependence on t. Since Equation (32) reads $s_{t+1} = s_t$, we have $\mathbf{\Phi} = 1$, $q_t = 0$ for all t, and $\mathbf{Q} = 0$. That process model was *exact*.

Definition 4.3. *Linear observation model*
Let $\mathbf{m}_t$ be the M-dimensional *measurement* vector and $\mathbf{r}_t$ the M-dimensional *measurement noise* vector. If $\mathbf{H}$ is an $M \times N$ *measurement* matrix, the *observation model* expresses $\mathbf{m}_t$ in terms of $\mathbf{s}_t$ and $\mathbf{r}_t$, or

$$\mathbf{m}_t = \mathbf{H}\mathbf{s}_t + \mathbf{r}_t \tag{33}$$

Definition 4.4. *Measurement noise expected value and covariance*
We assume that $\mathbb{E}[\mathbf{r}_t] = 0$ and denote the $M \times M$ *measurement noise covariance* matrix with

$$\mathbf{R} = \mathbb{E}[\mathbf{r}_t\mathbf{r}_t^\top]$$

Observation 4.2. *Back to averaging, once again*
In our first example both $\mathbf{m}_t$ and $\mathbf{r}_t$ were scalar. Since $m_t = s + r_t$, we have $\mathbf{H} = 1$ and the covariance matrix $\mathbf{R}$ reduces to the noise variance R.

Definition 4.5. *Error*
Let $\hat{\mathbf{s}}_t$ be the *estimate* of $\mathbf{s}_t$ obtained *after* the acquisition of the measurement vector $\mathbf{m}_t$. The difference between the true hidden state and the *posterior state estimate* is expressed by the N-dimensional *error* vector

$$\mathbf{e}_t = \mathbf{s} - \hat{\mathbf{s}}_t$$

Definition 4.6. *Posterior covariance*
This time $\mathbf{P}_t$ is an $N \times N$ matrix, the *posterior covariance*, or

$$\mathbf{P}_t = \mathbb{E}[\mathbf{e}_t\mathbf{e}_t^\top] \tag{34}$$

Same objective and plan As before, we use the "$^-$" superscript to denote the projection of a given quantity from time t to time $t+1$ *before* the acquisition of the measurement at time $t+1$. However, this time the *Kalman gain* is the $N \times M$ matrix $\mathbf{K}_t$ which, after the acquisition of the M measurement $\mathbf{m}_t$ gives the N-dimensional **state update** equation at time t as

$$\hat{\mathbf{s}}_t = \hat{\mathbf{s}}_t^- + \mathbf{K}_t(\mathbf{m}_t - \mathbf{H}\hat{\mathbf{s}}_t^-) \tag{35}$$

The state estimate $\hat{\mathbf{s}}_t$, again, is a weighted average between $\hat{\mathbf{s}}_t^-$, the state estimate at time t we predict to obtain before the t-th measurement is acquired, and the innovation $\mathbf{m}_t - \mathbf{H}\hat{\mathbf{s}}_t^-$. The optimal weight, the Kalman gain matrix, is obtained by minimising the trace of the covariance matrix. There is only one difference with respect to the starting example: we have to deal with matrices every step of the way.

Example 4.1. *Tracking*
Assume you want to track a ball moving with constant velocity on the billiard table. The hidden state is a 4-dimensional state vector $\mathbf{s}_t = [x_t \ u_t \ y_t \ v_t]^\top$: x_t and y_t are the two components for the ball position, and u_t and v_t the two components for its constant velocity along the X and Y direction at time t. According to the linear model of Equation (32) the velocity does not depend on t, and thus

$$u_{t+1} = u_t \quad \text{and} \quad v_{t+1} = v_t$$

while the ball position at time $t+1$ can be computed from the ball position and velocity at time t as

$$x_{t+1} = x_t + u_t\Delta t \quad \text{and} \quad y_{t+1} = y_t + v_t\Delta t$$

Since $\Delta t = (t+1) - t = 1$, if we set

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Equation (32) reads

$$\begin{bmatrix} x_{t+1} \\ u_{t+1} \\ y_{t+1} \\ v_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ u_t \\ y_t \\ v_t \end{bmatrix} + \begin{bmatrix} q_{x_t} \\ q_{u_t} \\ q_{y_t} \\ q_{v_t} \end{bmatrix}$$

with the process noise vector $\mathbf{q}_t = [q_{x_t} \ q_{u_t} \ q_{y_t} \ q_{v_t}]^\top = 0$ for all t. The role of the process noise is to allow some variability in the model. In this case, for example, $\mathbf{q}_t \neq 0$ models a process in which the ball velocity is only piece-wise constant and includes the tracking of a ball rebounding from the cushion.

At each time t through some computer vision algorithm we estimate the ball position $\mathbf{m}_t = [m_{x_t} \ m_{y_t}]^\top$ through an image captured by a camera calibrated with respect to the billiard table. We thus have

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and Equation (33) reads

$$\begin{bmatrix} m_{x_t} \\ m_{y_t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ u_t \\ y_t \\ v_t \end{bmatrix} + \begin{bmatrix} r_{x_t} \\ r_{y_t} \end{bmatrix}$$

where $\mathbf{r}_t = [r_{x_t} \ r_{y_t}]$ is the measurement noise vector. $\square$

Observation 4.3. *To the moon (and back) in 1969!*
This scenario is much richer than before. Starting from a linear model of how a state evolves over time and a linear model of how measurements are obtained, for each t our goal is to obtain the Kalman gain, the state and covariance estimate updates, and project our estimates at time $t+1$. The recursive nature of the solution, combined with the predictive property, ensures that this program can be carried out effectively and with limited computational resources. Since many decades Kalman filtering has been, and still is, the basis for tracking and control tasks in many different application domains.

4.3 Linear Kalman Filter

As before, the program is to find the *Kalman gain* matrix $\mathbf{K}_t$. By plugging Equation (32) into Equation (35) we rewrite Equation (35) as

$$\hat{\mathbf{s}}_t = \hat{\mathbf{s}}_t^- + \mathbf{K}_t(\mathbf{H}\mathbf{s}_t + \mathbf{r}_t) - \mathbf{K}_t\mathbf{H}\hat{\mathbf{s}}_t^- \tag{36}$$

Using Equation (36), Equation (34) becomes

$$\mathbf{P}_t = \mathbb{E}\left[[(\mathbf{I} - \mathbf{K}_t\mathbf{H})(\mathbf{s}_t - \hat{\mathbf{s}}_t^-) - \mathbf{K}_t\mathbf{r}_t][(\mathbf{I} - \mathbf{K}_t\mathbf{H})(\mathbf{s}_t - \hat{\mathbf{s}}_t^-) - \mathbf{K}_t\mathbf{r}_t]^\top\right]$$

which, since the error of the prior state estimate and the measurement noise are uncorrelated, can be written

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbb{E}[(\mathbf{s}_t - \hat{\mathbf{s}}_t^-)(\mathbf{s}_t - \hat{\mathbf{s}}_t^-)^\top](\mathbf{I} - \mathbf{K}_t\mathbf{H})^\top + \mathbf{K}_t\mathbb{E}[\mathbf{r}_t\mathbf{r}_t^\top]\mathbf{K}_t^\top \tag{37}$$

If we set $\mathbf{P}_t^- = \mathbb{E}[(\mathbf{s}_t - \hat{\mathbf{s}}_t^-)(\mathbf{s}_t - \hat{\mathbf{s}}_t^-)^\top]$, since $\mathbb{E}[\mathbf{r}_t\mathbf{r}_t^\top] = \mathbf{R}$, we rewrite Equation (37) as

$$\begin{aligned}\mathbf{P}_t &= (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{P}_t^-(\mathbf{I} - \mathbf{K}_t\mathbf{H})^\top + \mathbf{K}_t\mathbf{R}\mathbf{K}_t^\top \\ &= \mathbf{P}_t^- - \mathbf{K}_t\mathbf{H}\mathbf{P}_t^- - \mathbf{P}_t^-\mathbf{H}^\top\mathbf{K}_t^\top + \mathbf{K}_t(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})\mathbf{K}_t^\top \end{aligned} \tag{38}$$

We know that the trace of $\mathbf{P}_t$, the scalar $\mathrm{Tr}[\mathbf{P}_t]$, is the sum of the mean squared errors. To minimise this sum we have to solve the $N \times M$ equations

$$\frac{\mathrm{d}\mathrm{Tr}[\mathbf{P}_t]}{\mathrm{d}\mathbf{K}_t} = 0 \tag{39}$$

With the derivative with respect to the matrix $\mathbf{K}_t$, we actually mean the derivative of $\mathrm{Tr}[\mathbf{P}_t]$ with respect to each of the $N \times M$ entries of $\mathbf{K}_t$.

For the trace we have

$$\mathrm{Tr}[\mathbf{P}_t] = \mathrm{Tr}[\mathbf{P}_t^-] - 2\mathrm{Tr}[\mathbf{K}_t\mathbf{H}\mathbf{P}_t^-] + \mathrm{Tr}[\mathbf{K}_t(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})\mathbf{K}_t^\top]$$

Notice that the matrix $\mathbf{K}_t(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})\mathbf{K}_t^\top$ is symmetric. Computing each derivative entry-wise, see the last section for details, in matrix notation Equation (39) becomes

$$0 = -2(\mathbf{H}\mathbf{P}_t^-)^\top + 2\mathbf{K}_t(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})$$

which can be solved for $\mathbf{K}_t$ to give the **Kalman gain**

$$\mathbf{K}_t = \mathbf{P}_t^-\mathbf{H}^\top(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})^{-1} \tag{40}$$

From Equation (40) we have that

$$\hat{\mathbf{K}}_t(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R}) = \mathbf{P}_t^-\mathbf{H}^\top(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})^{-1}(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R}) = \mathbf{P}_t^-\mathbf{H}^\top$$

Therefore, the third and fourth term in the sum of the right hand side of Equation (38) cancel out and we finally obtain the **update covariance** equation

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{P}_t^-$$

To complete the program, as before, we use the process model to obtain the **state estimate projection** equation, or the estimate of the state $\mathbf{s}_t$ *before* the acquisition of the measurement vector $\mathbf{m}_t$

$$\hat{\mathbf{s}}_{t+1}^- = \mathbf{\Phi}\hat{\mathbf{s}}_t \tag{41}$$

Using Equation (32) and (41) we obtain the **covariance estimate projection** equation

$$\begin{aligned}
\mathbf{P}_{t+1}^- &= \mathbb{E}[\mathbf{e}_{t+1}^-\mathbf{e}_{t+1}^{-\top}] = \mathbb{E}[(\mathbf{s}_{t+1} - \hat{\mathbf{s}}_{t+1}^-)(\mathbf{s}_{t+1} - \hat{\mathbf{s}}_{t+1}^-)^\top] \\
&= \mathbb{E}[(\mathbf{\Phi}\mathbf{s}_t + \mathbf{q}_t - \mathbf{\Phi}\hat{\mathbf{s}}_t)(\mathbf{\Phi}\mathbf{s}_t + \mathbf{q}_t - \mathbf{\Phi}\hat{\mathbf{s}}_t)^\top] = \mathbb{E}[(\mathbf{\Phi}\mathbf{e}_t + \mathbf{q}_t)(\mathbf{\Phi}\mathbf{e}_t + \mathbf{q}_t)^\top] \\
&= \mathbb{E}[\mathbf{\Phi}\mathbf{e}_t\mathbf{\Phi}\mathbf{e}_t^\top] + \mathbb{E}[\mathbf{q}_t\mathbf{q}_t^\top] = \mathbf{\Phi}\mathbf{P}_t\mathbf{\Phi}^\top + \mathbf{Q}
\end{aligned}$$

Similarly to the static case, the solution to our problem - this time relying on the prior knowledge of the model noise and measurement noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$ - consists of 7 steps.

Algorithm 4.2. *LinearKalmanFilter*
Guess the initial estimate of the hidden state $\mathbf{s}_1$ and the matrix $\mathbf{P}_1^-$ (random initialisation is an option). Set $t = 1$ and repeat

1.	Acquire the t-th measurement:	$\mathbf{m}_t$
2.	Set the Kalman gain:	$\mathbf{K}_t = \mathbf{P}_t^-\mathbf{H}^\top(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^\top + \mathbf{R})^{-1}$
3.	State update:	$\hat{\mathbf{s}}_t = \hat{\mathbf{s}}_t^- + \mathbf{K}_t(\mathbf{m}_t - \mathbf{H}\hat{\mathbf{s}}_t^-)$
4.	Covariance update:	$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{P}_t^-$
5.	State estimate projection:	$\hat{\mathbf{s}}_{t+1}^- = \mathbf{\Phi}\hat{\mathbf{s}}_t$
6.	Covariance estimate projection:	$\mathbf{P}_{t+1}^- = \mathbf{\Phi}\mathbf{P}_t\mathbf{\Phi}^\top + \mathbf{Q}$
7.	$t \leftarrow t + 1$	

4.4 Computing the derivative of a function affected by noise

We conclude with a non-trivial example which can be implemented without matrix inversion. We assume to consider an unspecified system described by a time-varying function f and its derivative f' with respect to time. For the true hidden state we thus write $\mathbf{s}_t = [f_t \; f_t']^\top$ where f_t and f_t' are the true value of the function f and of its temporal derivative both evaluated at time t. We do not have access to the true state $\mathbf{s}_t$ but at each time t we obtain a sample m_t and we want to use Kalman filtering to obtain the optimal state estimate $\hat{\mathbf{s}}_t = [\hat{f}_t \; \hat{f}_t']^\top$. Let us proceed by writing down what we have in this case.

As for $\mathbf{\Phi}$ we could use

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Assuming again a unit time interval $\Delta t = 1$ between samples, we model f_{t+1} as

$$f_{t+1} = f_t + f_t' \, \Delta t = f_t + f_t'$$

and the temporal derivative as a constant. We let $\mathbf{Q}$ be the 2×2 diagonal matrix

$$\mathbf{Q} = \begin{bmatrix} q_1^2 & 0 \\ 0 & q_2^2 \end{bmatrix}$$

with q_1^2 and q_2^2 the variances of the noise affecting the piece-wise linear model we are adopting and thus allowing for changes in the derivative value.

The measurement model is immediate. We have $\mathbf{H} = [1\ 0]$ and

$$m_t = [1\ 0] \begin{bmatrix} f_t \\ f'_t \end{bmatrix} + r_t = f_t + r_t$$

with R, as usual, for the variance of the noise r_t. If

$$\mathbf{P}_t = \begin{bmatrix} p_{11,t} & p_{12,t} \\ p_{21,t} & p_{22,t} \end{bmatrix} \quad \text{and} \quad \mathbf{P}_t^- = \begin{bmatrix} p_{11,t}^- & p_{12,t}^- \\ p_{21,t}^- & p_{22,t}^- \end{bmatrix}$$

since

$$\mathbf{P}_t^- \mathbf{H}^\top = \begin{bmatrix} p_{11,t}^- & p_{12,t}^- \\ p_{21,t}^- & p_{22,t}^- \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} p_{11,t}^- \\ p_{21,t}^- \end{bmatrix}$$

and

$$\mathbf{H}\mathbf{P}_t^- \mathbf{H}^\top = [1\ 0] \begin{bmatrix} p_{11,t}^- & p_{12,t}^- \\ p_{21,t}^- & p_{22,t}^- \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = [1\ 0] \begin{bmatrix} p_{11,t}^- \\ p_{21,t}^- \end{bmatrix} = p_{11,t}^-$$

we have that

$$(\mathbf{H}\mathbf{P}_t^- \mathbf{H}^\top + \mathrm{R})^{-1} = \frac{1}{p_{11,t}^- + \mathrm{R}}$$

so that the **Kalman gain**, $\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}^\top (\mathbf{H}\mathbf{P}_t^- \mathbf{H}^\top + \mathrm{R})^{-1}$, is given by

$$\mathbf{K}_t = \frac{1}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} p_{11,t}^- \\ p_{21,t}^- \end{bmatrix}$$

For the **state update**, $\hat{\mathbf{h}}_t = \hat{\mathbf{h}}_t^- + \mathbf{K}_t(\mathbf{m}_t - \mathbf{H}\hat{\mathbf{h}}_t^-)$, we thus have

$$\begin{bmatrix} \hat{f}_t \\ \hat{f}'_t \end{bmatrix} = \begin{bmatrix} \hat{f}_t^- \\ \hat{f}_t'^- \end{bmatrix} + \frac{m_t - \hat{f}_t^-}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} p_{11,t}^- \\ p_{21,t}^- \end{bmatrix}$$

Since

$$\mathbf{I} - \mathbf{K}_t\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} p_{11,t}^- \\ p_{21,t}^- \end{bmatrix} [1\ 0] = \frac{1}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} \mathrm{R} & 0 \\ -p_{21,t}^- & p_{11,t}^- + \mathrm{R} \end{bmatrix}$$

the **covariance update** reads

$$\begin{aligned}
\begin{bmatrix} p_{11,t} & p_{12,t} \\ p_{21,t} & p_{22,t} \end{bmatrix} &= \frac{1}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} \mathrm{R} & 0 \\ -p_{21,t}^- & p_{11,t}^- + \mathrm{R} \end{bmatrix} \begin{bmatrix} p_{11,t}^- & p_{12,t}^- \\ p_{21,t}^- & p_{22,t}^- \end{bmatrix} \\
&= \frac{1}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} \mathrm{R}p_{11,t}^- & \mathrm{R}p_{12,t}^- \\ \mathrm{R}p_{21,t}^- & -p_{21,t}^- p_{12,t}^- + p_{11,t}^- p_{22,t}^- + p_{22,t}^- \mathrm{R} \end{bmatrix} \\
&= \frac{\mathrm{R}}{p_{11,t}^- + \mathrm{R}} \begin{bmatrix} p_{11,t}^- & p_{12,t}^- \\ p_{21,t}^- & p_{22,t}^- + |\mathbf{P}_t^-|/\mathrm{R} \end{bmatrix}
\end{aligned}$$

where $|\mathbf{P}_t^-|$ denotes the determinant of $\mathbf{P}_t^-$. For the **state estimate projection**, instead, we have

$$\begin{bmatrix} \hat{f}_{t+1}^- \\ \hat{f}_{t+1}'^- \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{f}_t \\ \hat{f}'_t \end{bmatrix} = \begin{bmatrix} \hat{f}_t + \hat{f}'_t \\ \hat{f}'_t \end{bmatrix}$$

Finally, since

$$
\begin{aligned}
\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} p_{11,t} & p_{12,t} \\ p_{21,t} & p_{22,t} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} p_{11,t}+p_{12,t} & p_{12,t} \\ p_{21,t}+p_{22,t} & p_{22,t} \end{bmatrix} \\
&= \begin{bmatrix} p_{11,t}+p_{12,t}+p_{21,t}+p_{22,t} & p_{21,t}+p_{22,t} \\ p_{21,t}+p_{22,t} & p_{22,t} \end{bmatrix}
\end{aligned}
$$

we write the **covariance estimate projection** as

$$
\begin{bmatrix} p^-_{11,t} & p^-_{12,t} \\ p^-_{21,t} & p^-_{22,t} \end{bmatrix} = \begin{bmatrix} p_{11,t}+p_{12,t}+p_{21,t}+p_{22,t} & p_{12,t}+p_{22,t} \\ p_{21,t}+p_{22,t} & p_{22,t} \end{bmatrix} + \begin{bmatrix} q_1^2 & 0 \\ 0 & q_2^2 \end{bmatrix}
$$

The derivative of the trace with respect to a matrix

For the sake of simplicity we consider 2×2 matrices, leaving the generalisation as an exercise. $\mathbf{A}$ and $\mathbf{B}$ are arbitrary matrices while $\mathbf{C}$ is symmetric. We thus have

$$
\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}
$$

We first compute $\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top]$ and obtain

$$
\begin{aligned}
\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top] = & -2(a_{11}b_{11} + a_{12}b_{21} + a_{21}b_{12} + a_{22}b_{22}) \\
& + a_{11}^2 c_{11} + 2a_{11}a_{12}c_{12} + a_{12}^2 c_{22} + a_{21}^2 c_{11} + 2a_{21}a_{22}c_{12} + a_{22}^2 c_{22}
\end{aligned}
$$

For the derivative of $\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top]$ with respect to the entries of $\mathbf{A}$ we obtain

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}a_{11}}\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top] &= -2b_{11} + 2a_{11}c_{11} + 2a_{12}c_{12} \\
\frac{\mathrm{d}}{\mathrm{d}a_{12}}\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top] &= -2b_{21} + 2a_{11}c_{12} + 2a_{12}c_{22} \\
\frac{\mathrm{d}}{\mathrm{d}a_{21}}\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top] &= -2b_{12} + 2a_{21}c_{11} + 2a_{22}c_{12} \\
\frac{\mathrm{d}}{\mathrm{d}a_{22}}\mathrm{Tr}[-2\mathbf{AB} + \mathbf{ACA}^\top] &= -2b_{22} + 2a_{21}c_{12} + 2a_{22}c_{22}
\end{aligned}
$$

which in matrix form writes

$$
-2\mathbf{B}^\top + 2\mathbf{AC}
$$

This is precisely the result we got for $\mathrm{dTr}[\mathbf{P}_t]/\mathrm{d}\mathbf{K}_t$ with $\mathbf{K}_t = \mathbf{A}$, $\mathbf{HP}_t^- = \mathbf{B}$, and $\mathbf{HP}_t^-\mathbf{H}^\top + R = \mathbf{C}$.

4.5 Things you need to know

1. The notions of model and measurement model, the role played by noise in the two cases.
2. How to build simple models for tracking objects or estimate a function and its derivative.
3. The notion of Kalman gain and the four recursive equations: update state and covariance and state and projection estimate covariance. No need to remember the equations by heart but you should be able to recognise each of them and parse their meaning.
4. The role played by the covariance of the process noise and model noise in solving real life problems.

5 A Brief Intro to Wavelets

For the most part, this section is adapted from [8]. For an authoritative account on wavelets see [4].

5.1 A different representation for a 4-pixel image

We start off by discussing a very simple example. We are given a *4-pixel* image

$$\mathcal{I}_4 = [\,8\;4\;1\;3\,]$$

and we want to construct a multi-resolution representation for $\mathcal{I}$. If we average the first and the second pixel, and the third and the fourth, we obtain a lower resolution *2-pixel* image

$$\mathcal{I}_2 = \left[\,\frac{8+4}{2}\;\frac{1+3}{2}\,\right] = [\,6\;2\,]$$

The information lost in the averaging can be recovered if along with the averages, 6 and 2, we also save the *detailed coefficients*: **2** for the first average and **-1** for the second. Combining the averages with the detailed coefficients we obtain back the four original pixel values of $\mathcal{I}_4$,

$$6+\mathbf{2}=8,\; 6-\mathbf{2}=4,\; 2+(-\mathbf{1})=1,\; \text{and } 2-(-\mathbf{1})=3.$$

If we repeat the process by averaging the two pixels of the lower resolution *2-pixel* image, we obtain the *1-pixel* image

$$\mathcal{I}_1 = \left[\,\frac{6+2}{2}\,\right] = [\,4\,]$$

and the detail coefficient **2** needed to recover the two pixel values of $\mathcal{I}_2$. The full process is summarised in Figure 5.

Resolution	Averages	Detail coefficients
4	$[\,8\;\;4\;\;1\;\;3\,]$	
2	$[\,6\;\;2\,]$	$[\,2\;\;-1\,]$
1	$[\,4\,]$	$[\,2\,]$

Figure 5: From [8].

Nothing is gained and nothing is lost if, instead of saving $\mathcal{I}_4$, we save $\mathcal{I}_1$ and *three* detail coefficients: *one* for the reconstruction of $\mathcal{I}_2$ from $\mathcal{I}_1$, and *two* for the reconstruction of $\mathcal{I}_4$ from $\mathcal{I}_2$. It is conceivable that in this new representation the truncation of small detail coefficients at high resolution could lead to a lossy image compression of higher quality than straightforward downsampling.

The representation of images like $\mathcal{I}_4$ in terms of a single average value and three detailed coefficients is a very simple example of wavelet transform, known as *Haar wavelet transform*.

Let us go through the same example this time using wavelet notation and concepts along the way,

5.2 Haar wavelet transform

The set V^0 of all constant functions over the interval $[0, 1)$ allows us to describe *1-pixel* images. We increase the resolution by introducing the set V^1 of all the functions made by two constant pieces: the first over the interval $[0, 1/2)$ and the second over the interval $[1/2, 1)$. V^1 is what we need to describe *2-pixel* images. If we keep going, the set V^j consists of all piece-wise constant functions made by 2^j pieces (or $2^j - pixel$ images), with the interval $[0, 1)$ divided in 2^j disjoint $1/2^j$ wide intervals. As shown in Figure (6) the accuracy of the approximation of a given function increases with j.

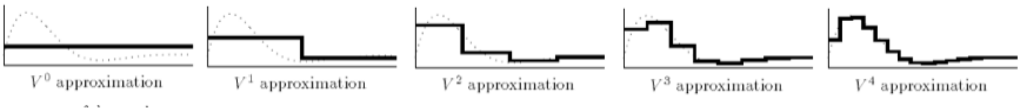

Figure 6: From [8].

Fact 5.1. *Multiresolution analysis*
For all j, the set V^j of all 2^j*-pixel* images is clearly a vector space. In addition, any 2^j*-pixel* image in V^j is also a 2^{j+1}*-pixel* image in V^{j+1}.

Exercise 5.1. *Nested structure of vector spaces*
Prove that

$$V^0 \subset V^1 \subset V^2 \subset \ldots$$

□

As basis functions for each space V^j we introduce the set of *scaled* and *translated* functions

$$\phi_i^j = \phi(2^j x - i), \quad \text{with } i = 0, \ldots, 2^j - 1$$

where the **box** function

$$\phi(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

is the *scaling* function.

Observation 5.1. *Scale and translation*
The 2^j functions ϕ_i^j of V^j for $i = 0, \ldots, 2^j - 1$ are non-overlapping translated copies of a scaled box function of width $1/2^j$ in the interval $[0, 1)$. Figure (7) shows the four functions of V^2.

Figure 7: From [8].

Let us adopt the standard definition for computing the scalar product between two vectors in V^j. If $\mathbf{u} = u(x)$ and $\mathbf{v} = v(x)$ are two vectors in V^j, or two 2^j*-pixel* images we have

$$\langle \mathbf{u}, \mathbf{v} \rangle = \int_0^1 u(x)v(x)\mathrm{d}x.$$

Fact 5.2. *An orthogonal basis for* V^j
The 2^j vectors ϕ_i^j for $i = 0, \ldots, 2^j - 1$ are mutually orthogonal and thus a basis for the 2^j dimensional

vector space V^j. Indeed no two functions ϕ_i^j and ϕ_k^j for $i \neq k$ are both different from zero on each of the 2^j subinterval (see again Figure (7) for $j = 2$). Therefore, we have

$$
\begin{aligned}
\langle \phi_i^j, \phi_k^j \rangle &= \int_0^1 \phi(2^j x - i)\phi(2^j x - k)\mathrm{d}x \\
&= \int_0^{1/2^j} \phi(2^j x - i)\phi(2^j x - k)\mathrm{d}x + \cdots + \int_{(2^j-1)/2^j}^1 \phi(2^j x - i)\phi(2^j x - k)\mathrm{d}x \\
&= \underbrace{0 + \cdots + 0}_{2^k} = 0
\end{aligned}
$$

Observation 5.2. *Easy step: from orthogonal to orthonormal*
The ϕ_i^j are not orthonormal since for $i = j$ we have

$$
\langle \phi_k^j, \phi_k^j \rangle = \int_0^1 \phi(2^j x - k)\phi(2^j x - k)\mathrm{d}x = \int_{k/2^j}^{(k+1)/2^j} \mathrm{d}x = \frac{1}{2^j}.
$$

To obtain an orthonormal basis we multiply each ϕ_i^j by the square root of 2^j and define $\hat{\phi}_i^j$ as

$$
\hat{\phi}_i^j = 2^{j/2}\phi(2^j x - i), \quad \text{with } i = 0, \ldots, 2^j - 1. \square
$$

Observation 5.3. *The squared norm of $\mathcal{I}_4$*
Rewriting $\mathcal{I}_4$ and using the orthonormality for the basis of V^2 we obtain

$$
\begin{aligned}
\|\mathcal{I}_4\|^2 &= \left\langle \frac{8}{2}\hat{\phi}_0^2 + \frac{4}{2}\hat{\phi}_1^2 + \frac{1}{2}\hat{\phi}_2^2 + \frac{3}{2}\hat{\phi}_3^2, \frac{8}{2}\hat{\phi}_0^2 + \frac{4}{2}\hat{\phi}_1^2 + \frac{1}{2}\hat{\phi}_2^2 + \frac{3}{2}\hat{\phi}_3^2 \right\rangle \\
&= \left\langle 4\hat{\phi}_0^2 + 2\hat{\phi}_1^2 + \frac{1}{2}\hat{\phi}_2^2 + \frac{3}{2}\hat{\phi}_3^2, 4\hat{\phi}_0^2 + 2\hat{\phi}_1^2 + \frac{1}{2}\hat{\phi}_2^2 + \frac{3}{2}\hat{\phi}_3^2 \right\rangle \\
&= 4 \cdot 4 + 2 \cdot 2 + \frac{1}{2} \cdot \frac{1}{2} + \frac{3}{2} \cdot \frac{3}{2} = 22.5
\end{aligned}
$$

$\square$

Before completing our first wavelet construction, we review the notion of orthogonal complement of a vector subspace.

Fact 5.3. *Orthogonal complement as a subspace*
Let V be a vector space. The orthogonal complement $W^\perp$ of a vector subspace U in V, or the set of all vectors $w \in V$ orthogonal to all vectors in U, is a vector subspace.

Example 5.1. *The XY-plane, the Z-axis, and the XYZ-space*
The set U of vectors which can be written as $(s_1, s_2, 0)$ for all s_1 and $s_2 \in \mathbb{R}$, the points in the *XY-plane*, is a subspace of $V = \mathbb{R}^3$, the *XYZ-space*. The orthogonal complement of U is the set $W^\perp$ of all vectors which can be written as $(0, 0, s_3)$ for all $s_3 \in \mathbb{R}$, the points in the *Z-axis*. It is trivial to verify that $W^\perp$ is a subspace and that the scalar product between any vector in U with any vector in $W^\perp$ is always 0.

Fact 5.4. *Sum of the dimension of a subspace and of its orthogonal complement*
For finite dimensional spaces it is always true that

$$
dim(V) = dim(U) + dim(W^\perp).
$$

Moreover a basis for V can be obtained as the union of a basis for U with a basis for $W^\perp$.

Observation 5.4. *Orthogonal complement of V^j in V^{j+1}*
An orthogonal basis for V^{j+1} can be obtained by adding to the 2^j mutually orthogonal ϕ_i^j, basis of V^j, the 2^j mutually orthogonal vectors ϕ_i^j basis for the vector space W^j, the orthogonal complement of V^j in V^{j+1}. $\square$

The main idea leading to the wavelet construction is that the basis for the orthogonal complement of V^j are the basis functions needed to recover the details lost in the averaging from V^{j+1} to V^j.

In the case of the *box* scaling function the basis functions are defined as

$$\psi_i^j = \psi(2^j x - i), \quad \text{with } i = 0, \ldots, 2^j - 1$$

where

$$\psi(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1/2 \\ -1 & \text{for } 1/2 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

is the *wavelet* function.

Figure (7) shows ψ_0^1 and ψ_1^1, the two wavelets of W^1.

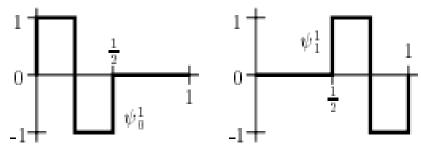

Figure 8: From [8].

Exercise 5.2. *Constructing an orthogonal basis for V^{j+1}*
Prove that the ϕ_i^j and the ψ_i^j for $i = 0, \ldots, 2^j - 1$ are a basis for V^{j+1}.

Observation 5.5. *Orthonormality*
Like in the case of the scaling functions ϕ_i^j, to obtain unit norm vectors we need to set $\hat{\psi}_i^j = 2^{j/2}\psi_i^j$.

In V^2 we write

$$\mathcal{I}_4(x) = c_0^2\phi_0^2(x) + c_1^2\phi_1^2(x) + c_2^2\phi_2^2(x) + c_3^2\phi_3^2(x) = 8\phi_0^2(x) + 4\phi_1^2(x) + 1\phi_2^2(x) + 3\phi_3^2(x)$$

The weight of each of the four scaled and translated boxes, ϕ_i^2 $i = 1, \ldots, 4$, is simply the value of the corresponding pixel i.

In V^1 and W^1, instead,

$$\mathcal{I}_4(x) = c_0^1\phi_0^1(x) + c_1^1\phi_1^1(x) + d_0^1\psi_0^1(x) + d_1^1\psi_1^1(x) = 6\phi_0^1(x) + 2\phi_1^1(x) + 2\psi_0^1(x) - 1\psi_1^1(x)$$

The weights of the two scaled and translated boxes ϕ_0^1 and ϕ_1^1 are the two averages, 6 and 2, while the weights of two wavelets ψ_0^1 and ψ_1^1 the corresponding detailed coefficients, 2 and -1.

Finally, in V^0, W^0, and W^1, we have

$$\mathcal{I}_4(x) = c_0^0\phi_0^0(x) + d_0^0\psi_0^0(x) + d_0^1\psi_0^1(x) + d_1^1\psi_1^1(x) = 4\phi_0^0(x) + 2\psi_0^0(x) + 2\psi_0^1(x) - 1\psi_1^1(x)$$

The weights of the only box at the lowest resolution, ϕ_0^0, is the average, 4, while the rest of the basis functions are wavelets at different resolution each weighted by the corresponding detailed coefficient.

Observation 5.6. *Pythagora, once again!*
If we rewrite the last two expansions using unit norm orthogonal vectors we find

$$\begin{aligned}\mathcal{I}_4(x) &= \frac{6}{\sqrt{2}}\hat{\phi}_0^1(x) + \frac{2}{\sqrt{2}}\hat{\phi}_1^1(x) + \frac{2}{\sqrt{2}}\hat{\psi}_0^1(x) - \frac{1}{\sqrt{2}}\hat{\psi}_1^1(x) \\ &= 4\hat{\phi}_0^0(x) + 2\hat{\psi}_0^0(x) + \frac{2}{\sqrt{2}}\hat{\psi}_0^1(x) - \frac{1}{\sqrt{2}}\hat{\psi}_1^1(x)\end{aligned}$$

Here again, the sum of the square of the coefficients returns $\|\mathcal{I}_4\|^2$, since

$$\frac{6}{\sqrt{2}}\cdot\frac{6}{\sqrt{2}}+\frac{2}{\sqrt{2}}\cdot\frac{2}{\sqrt{2}}+\frac{2}{\sqrt{2}}\cdot\frac{2}{\sqrt{2}}+\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} = 4\cdot 4+2\cdot 2+\frac{2}{\sqrt{2}}\cdot\frac{2}{\sqrt{2}}+\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} = 22.5$$

5.3 Multiresolution analysis

We now go through the same steps once more and show that both the *analysis*, in which average and detail coefficients are computed at all scales, and *synthesis*, in which the original function is reconstructed from the computed coefficients, can be described in terms of linear filters and, therefore, matrices. Our analysis is limited to Haar wavelets. Wherever appropriate we mention what happens in the general case. For simplicity we stick to the *power of 2* scenario.

Synthesis fitlers

If we introduce the 2^j*-D row* vector

$$\mathbf{\Phi}^j(x) = \left[\phi_0^j \ \dots \ \phi_{2^j-1}^j\right]$$

and the $2^j \times 2^{j-1}$ constant matrix

$$\mathbf{P}^j = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ & & & & \dots & & \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

we can obtain the 2^{j-1}*-D row* vector $\mathbf{\Phi}^{j-1}(x)$ of scaling functions at the coarser level as

$$\mathbf{\Phi}^{j-1}(x) = \mathbf{\Phi}^j(x)\mathbf{P}^j.$$

Therefore, each scaling function at the level $j-1$ can be written as a linear combination of the scaling functions at the finer scale j.

Similarly, we introduce the 2^j*-D row* vector

$$\mathbf{\Psi}^j(x) = \left[\psi_0^j \ \dots \ \psi_{2^j-1}^j\right]$$

and the $2^j \times 2^{j-1}$ constant matrix

$$\mathbf{Q}^j = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 0 & 0 & \dots & 0 & 0 \\ & & & & \dots & & \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 \end{bmatrix} \tag{42}$$

and write

$$\mathbf{\Psi}^{j-1}(x) = \mathbf{\Phi}^j(x)\mathbf{Q}^j.$$

Here again, each wavelet function at the level $j-1$ can be written as a linear combination of the scaling functions at the finer scale j.

Example 5.2. *Back to $j = 2$*
Since

$$\mathbf{P}^2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } \mathbf{Q}^2 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}$$

we have

$$\begin{bmatrix}\phi_0^1(x) & \phi_1^1(x)\end{bmatrix} = \begin{bmatrix}\phi_0^2(x) & \phi_1^2(x) & \phi_2^2(x) & \phi_3^2(x)\end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix}\psi_0^1(x) & \psi_1^1(x)\end{bmatrix} = \begin{bmatrix}\phi_0^2(x) & \phi_1^2(x) & \phi_2^2(x) & \phi_3^2(x)\end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}$$

Observation 5.7. *Is $\mathbf{Q}^j$ uniquely determined by the choice of the ϕ_i^j and $\mathbf{P}^j$?*
The answer is in the negative. The columns of $\mathbf{Q}^j$ must form a basis for the orthogonal complement of V^{j-1} in V^j. Of the many possible choices available once we choose the scaling functions Φ^j and the synthesis filters $\mathbf{P}^j$, the $\mathbf{Q}^j$ matrix of the Haar wavelet in Equation (42) is uniquely determined by the additional constraint of having the least number of non-zero entries in each column. $\square$

We now introduce the 2^j*-D column* vectors $\mathbf{C}^j$ and $\mathbf{D}^j$ with

$$\mathbf{C}^j = \begin{bmatrix}c_0^j & \dots & c_{2^j-1}^j\end{bmatrix}^\top \text{ and } \mathbf{D}^j = \begin{bmatrix}d_0^j & \dots & d_{2^j-1}^j\end{bmatrix}^\top$$

where the c_i^j are the coefficients of the scaling functions at the level j needed to approximate a given function $f(x)$, and d_i^j, the detail coefficients for $i = 0, \dots, 2^j - 1$ needed to reconstruct the function $f(x)$ at the finer scale $j+1$.

The $\mathbf{P}^j$ and $\mathbf{Q}^j$ are called *synthesis* filters since we have

$$\mathbf{C}^j = \mathbf{P}^j\mathbf{C}^{j-1} + \mathbf{Q}^j\mathbf{D}^{j-1} = \begin{bmatrix}\mathbf{P}^j | \mathbf{Q}^j\end{bmatrix} \begin{bmatrix}\mathbf{C}^{j-1} \\ \hline \mathbf{D}^{j-1}\end{bmatrix}. \tag{43}$$

Exercise 5.3. *Sanity check*
Verify that the dimension of all the vectors and matrices in Equation (43) are consistent.

Example 5.3. *Again the case of $j = 2$*
If

$$\mathbf{C}^2 = [\,8\,4\,1\,3\,]^\top \quad \mathbf{C}^1 = [\,6\,2\,]^\top \quad \mathbf{D}^1 = [\,2\ \ -1\,]^\top \quad \mathbf{C}^0 = [\,4\,] \text{ and } \mathbf{D}^0 = [\,2\,],$$

consistently with what we just derived with

$$\mathbf{P}^2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \mathbf{Q}^2 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \quad \mathbf{P}^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{Q}^1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

we have

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 1 \\ 3 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} [\,4\,] + \begin{bmatrix} 1 \\ -1 \end{bmatrix} [\,2\,] = \begin{bmatrix} 4 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}.$$

Analysis filters

We now want to find the $2^{j-1} \times 2^j$ matrices $\mathbf{A}^j$ and $\mathbf{B}^j$ needed to obtain the lower resolution coefficients

$$\mathbf{C}^{j-1} = \mathbf{A}^j \mathbf{C}^j$$

$$\mathbf{D}^{j-1} = \mathbf{B}^j \mathbf{C}^j$$

For the unnormalised Haar basis we have

$$\mathbf{A}^j = \frac{1}{2}(\mathbf{P}^j)^\top \text{ and } \mathbf{B}^j = \frac{1}{2}(\mathbf{Q}^j)^\top. \tag{44}$$

Exercise 5.4. *Boring check*
Verify Equation (44) in the case of **Example 5.3**.

Observation 5.8. *Analysis filters*
The matrices $\mathbf{A}^j$ and $\mathbf{B}^j$ are *analysis* filters since they are used to obtain the coefficients at level $j-1$ from the finer resolution at level j.

Observation 5.9. *Relationship between analysis and synthesis filters*
In the general case the analysis filters and the synthesis filters are not the transpose of one another though it is always true that, stacking $\mathbf{A}^j$ and $\mathbf{B}^j$ one on top of the other and $\mathbf{P}^j$ and $\mathbf{Q}^j$ next to each other, it holds

$$\left[\frac{\mathbf{A}^j}{\mathbf{B}^j}\right] = \left[\mathbf{P}^j | \mathbf{Q}^j\right]^{-1}$$

5.4 Things you need to know

1. How to obtain the scaled and translated version of the pair scaling and wavelet functions ϕ and ψ in the case of the Haar Wavelet Transform in the interval $[0, 1]$.
2. Orthogonality between the subspaces V^j and W^j and the fact that V^{j+1} can be viewed as the direct sum of V^j and W^j
3. Analysis and synthesis filters for the Haar Wavelet.
4. Be able to work out some wavelet analysis and synthesis on simple examples.

6 Frames

We now touch upon the notion of redundant representations. We restrict our attention to the particular case of finite dimensional Euclidean vector spaces. If you want to learn more you may start from [1] and then move on to [2].

6.1 Beyond orthogonality

Signal representations built on orthonormal bases, like Fourier Series, Fourier Transform, and the Haar Wavelet Transform we came across so far, share two desirable properties: (1) the projection on a unit norm vector tells how much of that vector is needed to reconstruct the signal, and (2) orthogonality ensures that no other projection carries information on the signal in that vector direction. The use of an orthonormal basis, however, presents two main drawbacks. First, signal and noise can be hardly distinguished: a missing coefficient, for example, cannot be recovered or noise effect cannot be mitigated. Second, since many coefficients are different from zero, sparse representations, which might be preferable in sampling or compression applications, cannot be obtained. A way to overcome both limitations is provided by frames which yield an over-complete representation.

Definition 6.1. *Frames in* $\mathbb{R}^n$
A *frame* is a *sequence* of vectors $\{\mathbf{f}_i\}_{i=1}^m$ in $\mathbb{R}^n$, with $m \geq n$, for which there exist constants A and B with $0 < A \leq B < +\infty$ such that for all $\mathbf{v} \in \mathbb{R}^n$

$$A\|\mathbf{v}\|^2 \leq \sum_{i=1}^{m} \langle \mathbf{f}_i, \mathbf{v} \rangle^2 \leq B\|\mathbf{v}\|^2 \tag{45}$$

If $A = B$ the frame is *tight*. If $\|\mathbf{f}_i\| = 1$ for $i = 1, \ldots, m$, the frame is *unit-norm*. $\square$

Fact 6.1. *In finite dimension frames are equivalent to spanning sets*
A sequence of vectors $\{\mathbf{f}_i\}_{i=1}^m \in \mathbb{R}^n$ is a frame if and only if the $\{\mathbf{f}_i\}_{i=1}^m$ span $\mathbb{R}^n$.
Proof

$\rightarrow$: We proceed by contradiction. We assume that the sequence $\{\mathbf{f}_i\}_{i=1}^m$ is a frame and that it does not span $\mathbb{R}^n$. We can thus find a vector $\mathbf{v}' \neq 0$ for which $\langle \mathbf{f}_i, \mathbf{v}' \rangle = 0$ for $i = 1, \ldots, m$. But since $A > 0$ this violates the leftmost inequality in (45).

$\leftarrow$: We need to show that when the vectors $\{\mathbf{f}_i\}_{i=1}^m$ span $\mathbb{R}^n$, $\exists\, A, B$ with $0 < A \leq B < +\infty$ such that for all $\mathbf{v} \in \mathbb{R}^n$ both the inequalities in (45) hold true. We first show that $\sum_{i=1}^m \|\mathbf{f}_i\|^2$ is a possible choice for the constant B. Indeed, from the inequality $\langle \mathbf{f}_i, \mathbf{v} \rangle^2 \leq \|\mathbf{f}_i\|^2 \|\mathbf{v}\|^2$, which holds true for all $\mathbf{v} \in \mathbb{R}^n$ and $i = 1, \ldots, m$, we find

$$\sum_{i=1}^{m} \langle \mathbf{f}_i, \mathbf{v} \rangle^2 \leq (\sum_{i=1}^{m} \|\mathbf{f}_i\|^2)\|\mathbf{v}\|^2$$

We now let

$$A = \min_{\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \sum_{i=1}^{m} \left\langle \mathbf{f}_i, \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\rangle^2$$

and notice that since the set of unit vectors in $\mathbb{R}^n$ is closed and bounded and the $\{\mathbf{f}_i\}_{i=1}^m$ span $\mathbb{R}^n$, A exists and is strictly positive. Therefore, for all $\mathbf{v} \in \mathbb{R}^n$, we obtain

$$A\|\mathbf{v}\|^2 \leq \sum_{i=1}^{m} \langle \mathbf{f}_i, \mathbf{v} \rangle^2$$

$\blacksquare$

Four observations are in order. First, being a sequence and not a set of vectors, a frame may contain the same element more than once. Second, an orthonormal basis is a tight unit-norm frame with $A = B = 1$. Third, if $m > n$, a frame can be viewed as a *redundant* basis. Lastly, if $m < n$ the sequence of m vectors cannot span $\mathbb{R}^n$ and, thus, cannot be a frame.

Example 6.1. *A first redundant basis*
The six unit vectors of $\mathbb{R}^2$ in Figure 9

$$\mathbf{f}_i = \left[\cos\left(\frac{2(i-1)\pi}{6}\right) \quad \sin\left(\frac{2(i-1)\pi}{6}\right)\right]^\top \quad \text{with } i = 1, \ldots, 6$$

span $\mathbb{R}^2$ and thus define a unit-norm frame. Since $\mathbf{f}_1 = -\mathbf{f}_4$, $\mathbf{f}_2 = -\mathbf{f}_5$ and $\mathbf{f}_3 = -\mathbf{f}_6$ this frame actually consists of three vectors, repeated twice, which span $\mathbb{R}^2$.

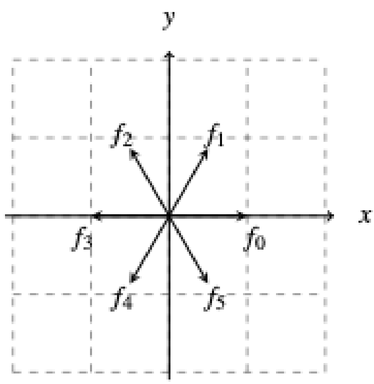

Figure 9: See text.

Example 6.2. *A second redundant basis*
The three vectors of $\mathbb{R}^2$

$$\mathbf{f}'_1 = [1 \quad 0]^\top \quad \mathbf{f}'_2 = [0 \quad 1]^\top \quad \text{and} \quad \mathbf{f}'_3 = [-1 \quad -1]^\top$$

span $\mathbb{R}^2$ and thus define a frame. Since $\|\mathbf{f}'_3\| = \sqrt{2}$ this frame is not unit-norm.

6.2 Analysis and synthesis filters

Analysis filter

The linear map $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^m$ defined as

$$\mathcal{F}(\mathbf{v}) = [\langle \mathbf{f}_1, \mathbf{v}\rangle \quad \ldots \quad \langle \mathbf{f}_m, \mathbf{v}\rangle]^\top = \mathbf{F}\mathbf{v}$$

maps a vector $\mathbf{v} \in \mathbb{R}^n$ into its m projections on $\mathbf{f}_1, \ldots \mathbf{f}_m$. If we adopt the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$, $\mathcal{F}$ can be represented by the $m \times n$ matrix $\mathbf{F}$, the i-th row of which is $\mathbf{f}_i^\top$. The matrix $\mathbf{F}$ plays the role of an **analysis filter**.

Exercise 6.1. *Analysis filter for a tight frame*
Find the linear map $\mathcal{F}$ and the matrix $\mathbf{F}$ for the frame $\{\mathbf{f}_i\}_{i=1}^6$ of **Example 6.1**. Then, show that the

frame is tight with $A = B = 3$.
Solution
For the six projections of a vector $\mathbf{v} = [x, y]^\top \in \mathbb{R}^2$ on the frame $\{\mathbf{f}_i\}_{i=1}^6$ the matrix $\mathbf{F}$ is 6×2 and we find

$$\mathcal{F}(\mathbf{v}) = \mathbf{F}\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 1/2 & \sqrt{3}/2 \\ -1/2 & \sqrt{3}/2 \\ -1 & 0 \\ -1/2 & -\sqrt{3}/2 \\ 1/2 & -\sqrt{3}/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ x/2 + \sqrt{3}y/2 \\ -x/2 + \sqrt{3}y/2 \\ -x \\ -x/2 - \sqrt{3}y/2 \\ x/2 - \sqrt{3}y/2 \end{bmatrix}$$

This frame is tight with $A = B = 3$ since for any $\mathbf{v} = [x \;\; y]^\top \in \mathbb{R}^n$ we have

$$\begin{aligned} \sum_{i=1}^{6} \langle \mathbf{f}_i, \mathbf{v} \rangle^2 = & \; x^2 + x^2/4 + y^2 3/4 + 2xy\sqrt{3}/4 + x^2/4 + y^2 3/4 - 2xy\sqrt{3}/4 + \\ & x^2 + x^2/4 + y^2 3/4 + 2xy\sqrt{3}/4 + x^2/4 + y^2 3/4 - 2xy\sqrt{3}/4 = 3(x^2 + y^2) = 3\|\mathbf{v}\|^2 \end{aligned}$$

Exercise 6.2. *Analysis filter for a non unit-norm frame*
Find the linear map $\mathcal{F}'$ and the matrix $\mathbf{F}'$ for the frame $\{\mathbf{f}'\}_{i=1}^3$ of **Example 6.2**. Then show that the optimal values for A and B for this frame are $A = 1$ and $B = 3$.
Solution

$$\mathcal{F}'(\mathbf{v}) = \mathbf{F}'\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \\ -x - y \end{bmatrix}$$

Furthermore, since

$$\sum_{i=1}^{3} \langle \mathbf{f}'_i, \mathbf{v} \rangle^2 = x^2 + y^2 + (x + y)^2$$

the minimum is reached for $\mathbf{v} = [x \quad -y]^\top$ which gives $\sum_{i=1}^{3} \langle \mathbf{f}'_i, \mathbf{v} \rangle^2 = x^2 + y^2 = \|\mathbf{v}\|^2$ and thus $A = 1$. In order to show that the optimal value for B is 3, without loss of generality, we assume $xy > 0$ and write $y = x + \epsilon$. We have

$$\begin{aligned} 0 &\leq \epsilon^2 \\ 2xy &\leq \epsilon^2 + 2xy = \epsilon^2 + 2x(x + \epsilon) = x^2 + x^2 + 2x\epsilon + \epsilon^2 = x^2 + y^2 \\ x^2 + y^2 + 2xy &\leq 2(x^2 + y^2) \\ (x + y)^2 &\leq 2\|\mathbf{v}\|^2 \end{aligned}$$

and thus $\sum_{i=1}^{3} \langle \mathbf{f}'_i, \mathbf{v} \rangle^2 \leq 3\|\mathbf{v}\|^2$ with the equal sign reached for $\epsilon = 0$.

Synthesis filter

The linear map $\mathcal{F}^* : \mathbb{R}^m \to \mathbb{R}^n$ defined as

$$\mathcal{F}^*(\mathbf{c}) = \sum_{i=1}^{m} c_i \mathbf{f}_i$$

maps a vector $\mathbf{c} = [c_1 \;\; \ldots \;\; c_m]^\top \in \mathbb{R}^m$ into the vector $\sum_{i=1}^{m} c_i \mathbf{f}_i \in \mathbb{R}^n$. If we adopt again the standard basis of $\mathbb{R}^m$ and $\mathbb{R}^n$, $\mathcal{F}^*$ can be represented by the transpose matrix $\mathbf{F}^\top$. The matrix $\mathbf{F}^\top$ plays the role of a **synthesis filter**.

In the case of **Example 6.1** $\mathbf{F}^\top$ is 2×6 and we have

$$\begin{aligned}\mathcal{F}^*(\mathbf{c}) &= \mathbf{F}^\top \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & -1/2 & -1 & -1/2 & 1/2 \\ 0 & \sqrt{3}/2 & \sqrt{3}/2 & 0 & -\sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} \\ &= \begin{bmatrix} c_1 + c_2/2 - c_3/2 - c_4 - c_5/2 + c_6/2 \\ c_2\sqrt{3}/2 + c_3\sqrt{3}/2 - c_5\sqrt{3}/2 - c_6\sqrt{3}/2 \end{bmatrix}\end{aligned}$$

In the case of **Example 6.2**, instead, $\mathbf{F'}^\top$ is 2×3 and we have

$$\mathcal{F}'^*(\mathbf{c}) = \mathbf{F'}^\top \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_1 - c_3 \\ c_2 - c_3 \end{bmatrix}$$

6.3 Frame operator

We first prove three simple propositions.

Proposition 6.1. *Injectivity*
The linear map $\mathcal{F}$ is injective.
Proof: if for some $\mathbf{v}$ and $\mathbf{w} \in \mathbb{R}^n$ we have that $\mathcal{F}(\mathbf{v}) = \mathcal{F}(\mathbf{w})$, from the linearity of $\mathcal{F}$ it follows that

$$0 = \mathcal{F}(\mathbf{v}) - \mathcal{F}(\mathbf{w}) = \langle (\mathbf{v} - \mathbf{w}), \mathbf{f}_i \rangle = 0 \quad \text{for} \quad i = 1, \dots, m$$

But since the m vectors $\mathbf{f}_1, \dots, \mathbf{f}_m$ span $\mathbb{R}^n$ we necessarily have $\mathbf{v} - \mathbf{w} = \mathbf{0}$. ■

Proposition 6.2. *The standard basis in* $\mathbb{R}^m$
If $\mathbf{e}_1, \dots, \mathbf{e}_m$ is the standard basis of $\mathbb{R}^m$ then

$$F^\top \mathbf{e}_i = \mathbf{f}_i \quad \text{with } i = 1, \dots, m$$

Proof: it follows from the fact that for $i = 1, \dots, m$

$$\langle \mathbf{v}, \mathbf{F}^\top \mathbf{e}_i \rangle = \langle \mathbf{F}\mathbf{v}, \mathbf{e}_i \rangle = \langle \mathbf{v}, \mathbf{f}_i \rangle$$

where the first equality is obtained by transposing $\mathbf{F}$ inside the scalar product. ■

Proposition 6.3. *Surjectivity*
The linear map $\mathcal{F}^*$ is surjective.
Proof: $\forall \mathbf{v} \in \mathbb{R}^n$ since the m vectors $\mathbf{f}_i$ span $\mathbb{R}^n$ there exist $\alpha_1, \dots \alpha_m \in \mathbb{R}$ such that

$$\mathbf{v} = \alpha_1 \mathbf{f}_1 + \cdots + \alpha_m \mathbf{f}_m = \alpha_1 \mathbf{F}^\top \mathbf{e}_1 + \cdots + \alpha_m \mathbf{F}^\top \mathbf{e}_m = \mathbf{F}^\top (\alpha_1 \mathbf{e}_1 + \cdots + \alpha_m \mathbf{e}_m)$$

■

Definition 6.2. *Frame operator*
Given a frame $\{\mathbf{f}_i\}_{i=1}^m$, the *frame operator* is the linear invertible map $\mathcal{S} : \mathbb{R}^n \to \mathbb{R}^n$ that, in the standard basis, reads

$$\mathcal{S}(\mathbf{v}) = \mathbf{S}\mathbf{v} = \mathbf{F}^\top \mathbf{F}\mathbf{v} = \sum_{i=1}^m \langle \mathbf{v}, \mathbf{f}_i \rangle \mathbf{f}_i$$

□

We thus have

$$\langle \mathbf{S}\mathbf{v}, \mathbf{v}\rangle = \sum_{i=1}^{m} \langle \mathbf{v}, \mathbf{f}_i\rangle^2$$

Fact 6.2. *Frame operator properties*
The matrix $\mathbf{S}$ is symmetric, invertible, and positive definite. □

Therefore, we have

$$\mathbf{v} = \mathbf{S}\mathbf{S}^{-1}\mathbf{v} = \mathbf{S}\sum_{i=1}^{m}\langle \mathbf{S}^{-1}\mathbf{v}, \mathbf{f}_i\rangle \mathbf{f}_i = \sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{S}^{-1}\mathbf{f}_i\rangle \mathbf{f}_i = \sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{g}_i\rangle \mathbf{f}_i \tag{46}$$

and

$$\mathbf{v} = \mathbf{S}^{-1}\mathbf{S}\mathbf{v} = \mathbf{S}^{-1}\sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{f}_i\rangle \mathbf{f}_i = \sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{f}_i\rangle \mathbf{S}^{-1}\mathbf{f}_i = \sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{f}_i\rangle \mathbf{g}_i \tag{47}$$

The frame $\{\mathbf{g}_i\}_{i=1}^m = \{\mathbf{S}^{-1}\mathbf{f}_i\}_{i=1}^m$, which is also a frame because $\mathcal{S}$ is invertible, is the *canonical* dual frame of $\{\mathbf{f}_i\}_{i=1}^m$. This result is a **generalisation of the expansion obtained with an orthonormal basis**: when we expand $\mathbf{v}$ on a frame $\{\mathbf{f}_i\}_{i=1}^m$ we compute its projections on the dual frame $\{\mathbf{g}_i\}_{i=1}^m$, see Equation (46). Alternatively, when we expand the same vector $\mathbf{v}$ on the dual frame $\{\mathbf{g}_i\}_{i=1}^m$ we compute its projection on the original frame $\{\mathbf{f}_i\}_{i=1}^m$, see Equation (47).

Exercise 6.3. *Frame operator, its inverse, and dual frame for a tight frame*
Find the frame operator, its inverse, and the dual frame for the frame of **Example 6.1**
Solution
For the frame operator we find

$$\mathbf{S} = \begin{bmatrix} 1 & 1/2 & -1/2 & -1 & -1/2 & 1/2 \\ 0 & \sqrt{3}/2 & \sqrt{3}/2 & 0 & -\sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1/2 & \sqrt{3}/2 \\ -1/2 & \sqrt{3}/2 \\ -1 & 0 \\ -1/2 & -\sqrt{3}/2 \\ 1/2 & -\sqrt{3}/2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = 3\mathbf{I}$$

Therefore we immediately find that

$$\mathbf{S}^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix}$$

with $\{\mathbf{g}_i\}_{i=1}^6 = \{\mathbf{f}_i/3\}_{i=1}^6$. □

Since the frame operator for a tight frame is a multiple of the identity, we can rewrite the expansion without explicitly using the dual frame and obtain

$$\mathbf{v} = \frac{1}{A}\sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{f}_i\rangle \mathbf{f}_i \tag{48}$$

Observation 6.1. *Parseval, once again*
From Equation (48) we obtain

$$\|\mathbf{v}\|^2 = \frac{1}{A}\sum_{i=1}^{m}\langle \mathbf{v}, \mathbf{f}_i\rangle^2$$

which can be viewed as a generalisation of Parseval identity. The constant A balances the redundancy effect of not using an orthonormal basis.

Exercise 6.4. *Frame operator, its inverse, and dual frame for a non unit-norm frame*
Find the frame operator, its inverse, and the dual frame for the frame of **Example 6.2**
Solution
For the frame operator we find

$$\mathbf{S}' = \mathbf{F}'^{\top}\mathbf{F}' = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

In this case for $\mathbf{S}'^{-1}$ we find

$$\mathbf{S}'^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix}$$

while for the dual frame we obtain

$$\mathbf{g}'_1 = \mathbf{S}'^{-1}\mathbf{f}'_1 = \begin{bmatrix} 2/3 \\ -1/3 \end{bmatrix}, \ \mathbf{g}'_2 = \mathbf{S}'^{-1}\mathbf{f}'_2 = \begin{bmatrix} -1/3 \\ 2/3 \end{bmatrix}, \ \text{and} \ \mathbf{g}'_3 = \mathbf{S}'^{-1}\mathbf{f}'_3 = \begin{bmatrix} -1/3 \\ -1/3 \end{bmatrix} \tag{49}$$

We conclude this brief introduction on frames with three important remarks.

6.4 Concluding remarks

Spectral properties of the frame operator

We prove two simple results involving the eigenvalues of the frame matrix S.

Fact 6.3. *Optimal values*
The optimal lower and upper frame bound for a finite dimensional frame $\{\mathbf{f}_i\}_{i=1}^m$ are the smallest and the largest eigenvalue of the matrix S.
Proof
Since the matrix S is positive definite there exists an orthonormal basis of n eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ with n corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ such that every vector $\mathbf{v} \in \mathbb{R}^n$ can be written as

$$\mathbf{v} = \sum_{j=1}^{n} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$$

Applying $\mathbf{S}$ on both sides, since $\mathbf{S}\mathbf{u}_j = \lambda_j \mathbf{u}_j$ we obtain

$$\mathbf{S}\mathbf{v} = \sum_{j=1}^{n} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{S}\mathbf{u}_j = \sum_{j=1}^{n} \lambda_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$$

and thus

$$\langle \mathbf{S}\mathbf{v}, \mathbf{v} \rangle = \sum_{j=1}^{n} \lambda_j \langle \mathbf{v}, \mathbf{u}_j \rangle^2$$

But since $\langle \mathbf{S}\mathbf{v}, \mathbf{v} \rangle = \sum_{i=1}^{m} \langle \mathbf{v}, \mathbf{f}_i \rangle^2$ and $\sum_{j=1}^{n} \langle \mathbf{v}, \mathbf{u}_j \rangle^2 = \|\mathbf{v}\|^2$, we finally have

$$\lambda_{min}\|\mathbf{v}\|^2 \leq \sum_{i=1}^{m} \langle \mathbf{f}_i, \mathbf{v} \rangle^2 \leq \lambda_{max}\|\mathbf{v}\|^2$$

Taking $\mathbf{u}_{min}$ and $\mathbf{u}_{max}$ to be eigenvectors corresponding to the eigenvalue λ_{min} and λ_{max} respectively proves that λ_{min} and λ_{max} are the optimal bounds.

Fact 6.4. $\sum_{j=1}^{n} \lambda_j = \sum_{i=1}^{m} \|\mathbf{f}_i\|^2$
Proof
Since for all $j = 1, \dots, n$ $\langle \mathbf{S}\mathbf{u}_j, \mathbf{u}_j \rangle = \sum_{i=1}^{m} \langle \mathbf{u}_j, \mathbf{f}_i \rangle^2$ we have

$$\sum_{j=1}^{n} \lambda_j = \sum_{j=1}^{n} \lambda_j \langle \mathbf{u}_j, \mathbf{u}_j \rangle = \sum_{j=1}^{n} \langle \mathbf{S}\mathbf{u}_j, \mathbf{u}_j \rangle = \sum_{j=1}^{n} \sum_{i=1}^{m} \langle \mathbf{u}_j, \mathbf{f}_i \rangle^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \langle \mathbf{u}_j, \mathbf{f}_i \rangle^2 = \sum_{i=1}^{m} \|\mathbf{f}_i\|^2$$

When the frame is tight and unit-norm

$$\sum_{j=1}^{n} \lambda_j = nA \quad \text{and} \quad \sum_{i=1}^{m} \|\mathbf{f}_i\|^2 = m$$

and thus we find $A = m/n$ known as the **frame redundancy ratio**.

What is left of uniqueness

Since the frame is a redundant basis, the coefficients of the expansion are not unique. Dual frames alternative to the canonical dual frame can also be found. Digging deeper in this direction is beyond the scope of this slim introduction to frames but we are in a position to characterise the coefficients obtained by projecting a vector on the canonical dual frame in simple terms.

Fact 6.5. *Minimal norm*
Let $\{\mathbf{f}_i\}_{i=1}^{m}$ be a frame and $\{\mathbf{g}_i = \mathbf{S}^{-1}\mathbf{f}_i\}_{i=1}^{m}$ its canonical dual frame. We write $\mathbf{v} = \sum_{i=1}^{m} \langle \mathbf{v}, \mathbf{g}_i \rangle \mathbf{f}_i$ and $\mathbf{v} = \sum_{i=1}^{m} a_i \mathbf{f}_i$ with some $a_i \neq \langle \mathbf{v}, \mathbf{g}_i \rangle$ since the frame is redundant ($m > n$). Then

$$\sum_{i=1}^{m} a_i^2 = \sum_{i=1}^{m} \langle \mathbf{v}, \mathbf{g}_i \rangle^2 + \sum_{i=1}^{m} (a_i - \langle \mathbf{v}, \mathbf{g}_i \rangle)^2 \tag{50}$$

Therefore, the coefficients obtained by projecting a vector over the canonical dual frame have the **smallest quadratic norm**.
Proof
By exploiting the symmetry of $\mathbf{S}^{-1}$ we find that the vector

$$[\langle \mathbf{v}, \mathbf{g}_1 \rangle \quad \dots \quad \langle \mathbf{v}, \mathbf{g}_m \rangle]^\top = [\langle \mathbf{v}, \mathbf{S}^{-1}\mathbf{f}_1 \rangle \quad \dots \quad \langle \mathbf{v}, \mathbf{S}^{-1}\mathbf{f}_m \rangle]^\top = [\langle \mathbf{S}^{-1}\mathbf{v}, \mathbf{f}_1 \rangle \quad \dots \quad \langle \mathbf{S}^{-1}\mathbf{v}, \mathbf{f}_m \rangle]^\top$$

belongs to the **range of the analysis filter**, $ran_{\mathbf{F}}$. At the same time since $\sum_{i=1}^{m} (a_i - \langle \mathbf{v}, \mathbf{g}_i \rangle) \mathbf{f}_i = 0$ we have that the vector

$$[a_1 - \langle \mathbf{v}, \mathbf{g}_1 \rangle \quad \dots \quad a_m - \langle \mathbf{v}, \mathbf{g}_m \rangle]^\top$$

belongs to the **kernel of the synthesis filter**, $ker_{\mathbf{F}^\top}$. Since for $i = 1, \dots, m$

$$a_i = \langle \mathbf{v}, \mathbf{g}_m \rangle + (a_i - \langle \mathbf{v}, \mathbf{g}_m \rangle)$$

Equation (50) follows from the fact that $ker_{\mathbf{F}^\top} = (ran_{\mathbf{F}})^\perp$, the orthogonal complement of $ran_{\mathbf{F}}$, with $\mathbb{R}^n = ran_{\mathbf{F}} \oplus (ran_{\mathbf{F}})^\perp$.

Exercise 6.5. *Just in case you don't trust theorems*
Verify that the expansion of the vector $\mathbf{v} = [9 \quad 12]^\top$ with the frame $\{\mathbf{f}'_i\}_{i=1}^{3}$ of **Example 6.2** is

$$\sum_{i=1}^{3} \langle \mathbf{v}, \mathbf{g}'_i \rangle \mathbf{f}'_i = 2\mathbf{f}'_1 + 5\mathbf{f}'_2 - 7\mathbf{f}'_3$$

Then, through the use of Equation (50), verify that the coefficients obtained by using the dual frame $\{\mathbf{g}'_i\}_{i=1}^3$ have smaller norm than those obtained in the standard basis.
Solution
Using the dual frame $\mathbf{g}'_1 = [2/3 \quad -1/3]^\top$, $\mathbf{g}'_2 = [-1/3 \quad 2/3]^\top$, and $\mathbf{g}'_3 = [-1/3 \quad -1/3]^\top$ we find

$$\langle \mathbf{v}, \mathbf{g}'_1 \rangle = 6 - 4 = 2, \quad \langle \mathbf{v}, \mathbf{g}'_2 \rangle = -3 + 8 = 5, \quad \text{and} \quad \langle \mathbf{v}, \mathbf{g}'_3 \rangle = -3 - 4 = -7$$

In the standard basis we have $a_1 = 9$, $a_2 = 12$, and $a_3 = 0$. Therefore, Equation (50) reads

$$\begin{aligned} a_1^2 + a_2^2 + a_3^2 &= 2 \times 2 + 5 \times 5 + 7 \times 7 + ((a_1 - 2)^2 + (a_2 - 5)^2 + (a_3 - 7)^2) \\ 9 \times 9 + 12 \times 12 &= 4 + 25 + 49 + (7 \times 7 + 7 \times 7 + 7 \times 7) \\ 81 + 144 &= 78 + (49 + 49 + 49) \\ 225 &= 225 \end{aligned}$$

or $\sum_{i=1}^3 a_i^2 = 225$ and $\sum_{i=1}^3 \langle \mathbf{v}, \mathbf{g}'_i \rangle^2 = 78$.

Building a frame

Even in the finite dimensional case frame design is not an easy task. A starting point is given by the following elementary result which, nevertheless, we state without proof. Let $m \geq n$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ a finite sequence of vectors in $\mathbb{R}^n$.

Definition 6.3. *Basis extension*
The sequence $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ has an *extension to a basis* for $\mathbb{R}^m$ if there exist vectors $\{\mathbf{h}_1, \ldots, \mathbf{h}_m\}$ in $\mathbb{R}^{m-n}$ such that

$$\left\{ \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{h}_1 \end{bmatrix} \cdots \begin{bmatrix} \mathbf{f}_m \\ \mathbf{h}_m \end{bmatrix} \right\}$$

is a basis for $\mathbb{R}^m$.

Fact 6.6. *From a frame to a basis*
If the sequence $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ is a **frame** in $\mathbb{R}^n$ then there exists an **extension** $\{\mathbf{h}_1, \ldots, \mathbf{h}_m\}$ to a basis for $\mathbb{R}^m$.

A simple method to construct a frame, therefore, can be obtained by projecting a basis of $\mathbb{R}^m$ onto $\mathbb{R}^n$. The design of a projection leading to a frame with some desired properties, however, is still subject of research.

6.5 Things you need to know

1. Analysis and synthesis filter for a frame in $\mathbb{R}^n$
2. Frame operator, canonical dual frame
3. Tight frames
4. Frame expansion and minimal norm property of the computed projections