A picture of

the Fourier Serie,

the Fourier Transform,

the discrete Fourier Transform

Version 0.1

Bastien Milani

IHU-LIRYC

January 2024, Bordeaux, France

contact : bastien.milani@yahoo.fr

Contents

Introduction: the analysis and the synthesis

The fundamental idea of Fourier analysis is probably much older than Monsieur Fourier and is present in all areas of mathematics, of natural sciences, in medicine, in philosophy, in alchemy and other areas of the human mind. It is the idea of the "Analysis" and the "Synthesis". Or in one word: "Atoms". It is the idea that anything in this word is a combination of fundamental parts, of fundamental ingredients. The decomposition of an object (a body, a piece of a material, a something…) into the fundamental ingredients is the analysis. The result of the analysis tells which quantity of each ingredient is present in the object. That list of quantity is the spectrum. And given a spectrum, the combination of the ingredients with their correct respective quantities in order to form the object is the synthesis. Every time one hears about a spectrum, it means that some decomposition into fundamental ingredients is going on. When Richard Feynman says in the introduction of his "Feynman lectures on Physics" that if the humanity had to save the most essential thing we have learned about nature on a little piece of paper, we should write that the world is made out of atoms, I thought first that it was about the material physical world only. But I am now tempted to think that what Feynman wanted to say (consciously or unconsciously) by "atoms" is that the world is a combination of a few ingredients. The notes of the music. And that the richness of the world arises out of the richness of the combinations. The accords. As if the relations between the notes carry a richness that it not present in the notes taken individually. This idea is so present in sciences, and probably in nature itself if we believe that science describes the nature, that the idea of atoms is one of the most important things human have discovered.

What Joseph Fourier was able to realize, in his work on the series and integral transform of his name, is to identify the atoms of some families of functions, and to find out which quantity of each atom is present in a given function. He found out the magic formulas to decompose a function in fundamental vibrations (or notes), and how to build a functions from its spectrum. The spectrum of a periodic function is made out of its Fourier coefficient, and the synthesis is realized by the Fourier serie. The spectrum of a "sufficiently fast decreasing" function is given by its Fourier transform and the synthesis is realized by the inverse Fourier transform. That being said, it is now clear that we should say "Fourier analysis and synthesis" to be precise but we will keep the title "Fourier analysis" for short.

It is however none of the Fourier coefficients or the Fourier transform that we use in modern digital systems for signal processing, medical image reconstruction, telecommunication…. It is the discrete Fourier transform (DFT), a discrete version of the Fourier transform. Astonishingly, the DFT was not invented together with the use of modern computers. In fact the mathematician Gauss already made use of it by hand in his calculation for astronomical observations and some sources even suggest that antic civilization were using it in some form. But it is only since 1965 that the DFT began to impact our world in such a way like today. What happened in 1965 is that Cooley and Tukey re-discovered the fast-Fourier-transform (FFT) algorithm (which evaluates the DFT in a rapid way) without knowing that it was already invented by Gauss (again) in 1805. But in contrast to Gauss, Cooley and Tukey had computers to take even more advantage of the rapidity of the FFT. While the matrix multiplication that realizes the DFT needs a number of operations that scales with N^2 for a data vector of size N, the FFT algorithm leads to the same answer with a number of operations scaling with $N\ log(N)$. The gain of time is so that it makes some technologies possible, which would be completely impossible without the FFT. The impact on our societies was so that, as Brad Osgood says in his course ("The Fourier Transform and its Applications", Electrical engineering department, Stanford University): "According to some, the modern world began in 1965 when J. Cooley and J. Tukey published their account of an efficient method for numerical computation of the Fourier transform". This should give a feeling about the importance of the DFT.

For the present course, we will divide the Fourier analysis in three variants. I gave here some names to these three variants but those names are not official:

- **Periodic Fourier analysis**: The Fourier coefficients and the associated Fourier serie for periodic functions,
- **Non-periodic Fourier analysis**: The Fourier transform and the inverse Fourier transform for non-periodic functions,
- **Discrete Fourier analysis**: The discrete Fourier transform and the inverse discrete Fourier transform for vectors of finite dimension.

The objects we decompose with the **periodic Fourier analysis** are periodic functions. Not all periodic functions can be decomposed but a large family of them. Moreover, not all periodic function can be synthetized with periodic Fourier analysis and the synthesis is never perfect. How good is the synthesis, is something difficult to quantify and mathematician have invented different kind of convergence-type to qualify how a synthesis approaches the function of interest. But this is out of the scope of this course.

The objects we decompose with **non-periodic Fourier analysis** are functions that converge "sufficiently rapidly" to zero as the argument becomes arbitrary large in norm. What "sufficiently rapidly" means is something subtle and the interested student may take a look at the "Schwartz space" and the space of "quadratic integrable functions". The synthesis is also problematic. For example, a function may accept a

Fourier transform (a spectrum) but the inverse Fourier transform of the spectrum may not converge everywhere. That means that synthetizing back the functions from its spectrum do not always work perfectly. But again, we will not go into that kind of detail in the present course.

The objects we decompose with the **discrete Fourier analysis** (DFT) are vectors of finite dimension and the DFT (as well as its inverse) are linear maps of vector spaces. This is what we need in concrete applications, and fortunately, it is mathematically much simpler than periodic and non-periodic Fourier analysis because we stay in the area of linear algebra. There is therefore no convergence problems in this area. The DFT and its inverse are well defined for any vector and the inverse DFT is exactly what it means: the inverse linear map of the DFT.

After having introduced some needed definitions and notions in part I, we will achieve the two major goals of this course in part II and part III:

- The first is to teach how to approximate the Fourier transform of a function by mean of the discrete Fourier transform via the use of an FFT implementation.
- The second is to highlight what are the precise relations between the discrete Fourier transform, the Fourier transform, and the Fourier serie. We will describe those relations by mean of the convolution product with some appropriate convolution kernels. We show that this can be done without any knowledge about distribution theory.

Part I: Background Notions and Definitions

Functions on the x-space and on the k-space

We assume that the reader is already familiar with many functions and with the use of variables. We will work with functions which accept as argument a 1-dimentional real variable. That mean that the definition domain of the functions we will work with is equal to $\mathbb{R}$. We will briefly generalize some expressions for n-dimensional real variables in one section but all the rest of the text will focus on the 1-dimensional case.

We will use two different definition domains for all our function. These two definition domains are both equal to $\mathbb{R}$, but we will call one the x-space and the other the k-space. They are therefore different because they differ at list by their name. We could say that they are two different copies of $\mathbb{R}$. We will usually write x a variable in the x-space and we will call it a “position” (or a “time”). We will usually write k a variable in the k-space and we will call it a “spatial frequency” (or a “temporal frequency”) or just a “frequency” for short.

We will usually write functions with the “dot-bracket” notation such as $f(\cdot)$ in order to stress the fact that it is a function and not a number. We will sometimes renounce to that notation and just write f instead if the “dot-bracket” becomes unpractical.

To define a real- or complex-valued function $f(\cdot)$ on the x-space we will typically write

$$f: \mathbb{R} \longrightarrow \mathbb{V}$$

$$x \longmapsto f(x)$$

where $\mathbb{V} = \mathbb{C}$ or $\mathbb{V} = \mathbb{R}$. To define a function $g(\cdot)$ on the k-space we will typically write

$$g: \mathbb{R} \longrightarrow \mathbb{V}$$

$$k \longmapsto g(k)$$

Because the author is a physicist, we will consider that any position x has an associated unit which is the meter m (or the second s for a time). Similarly, any frequency k has a unit which is $1/m$ for a spatial frequency (or $1/s$ for a time frequency).

The product $k\ x$ is therefore without unit and we will consider the number

$$\theta = 2\pi\, k\, x$$

to be an angle given in radian. We note that when x is a time, it is usually written t and the time-frequency is written ν. Some authors work in this case sometimes with the angular frequency ω given by

$$\omega = 2\pi\, \nu$$

The corresponding angle is then

$$\theta = \omega\, t = 2\pi\, \nu\, t$$

Periodic functions, non-periodic functions and normalized functions

Le be $\mathbb{V} = \mathbb{R}$ or $\mathbb{V} = \mathbb{C}$ and let

$$f \colon \mathbb{R} \longrightarrow \mathbb{V}$$

$$x \longmapsto f(x)$$

be a real- or complex-valued function. Given a positive length L in meter (or a duration T in second), $f(\cdot)$ is called "***L*-periodic**", or "**of period *L***" if

$$f(x + L) = f(x)$$

for any $x \in \mathbb{R}$. We will say that a function is "**periodic**" if there exist a positive number $L \in \mathbb{R}$ so that the function is L-periodic. That $f(\cdot)$ is "**non-periodic**" obviously means that it is not periodic.

We will say that the L-periodic function $f(\cdot)$ is "**normalized**" if

$$\int_{-L/2}^{+L/2} f(x)\, dx = 1$$

We note that we did not wrote $|f(x)|$ or $|f(x)|^2$ in the integrand. It is really about the functions values themselves.

We will say that a function $f(\cdot)$ "**converges to 0 at infinity**" if

$$\lim_{x \to +\infty} f(x) = \lim_{x \to -\infty} f(x) = 0$$

If $f(\cdot)$ is not periodic, we will say that it is "**normalized**" if

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1$$

provided that the integral converge. It can only converge if $f(\cdot)$ decreases to 0 at infinity "sufficiently rapidly". We will not define what "sufficiently rapidly" means. For us it will just mean "rapidly enough so that the integral converges". The interested reader may explore what is the **Schwartz space**.

Trigonometric functions, standard grids and sampling

We will assume that the reader is already familiar with trigonometric functions. We will give here some relations between them and some indication about how we will use them.

We will write i the square root of -1. The **Euler equation** gives then the relation

$$e^{i\theta} = \cos(\theta) + i\,\sin(\theta)$$

From

$$\theta = 2\pi\, k\, x$$

follows

$$e^{i2\pi\, kx} = \cos(2\pi\, kx) + i\,\sin(2\pi\, kx)$$

$$e^{-i2\pi\, kx} = \cos(2\pi\, kx) - i\,\sin(2\pi\, kx)$$

The expressions $\sin(2\pi\, kx)$ and $\cos(2\pi\, kx)$ can then be written in terms of complex exponential functions as

$$\cos(2\pi\, kx) = \frac{1}{2}(e^{i2\pi\, kx} + e^{-i2\pi\, kx})$$

$$\sin(2\pi\, kx) = \frac{1}{2i}(e^{i2\pi\, kx} - e^{-i2\pi\, kx})$$

The $\cos(\cdot)$ and $\sin(\cdot)$ functions are of period 2π (or 2π-periodic) It follows that the functions

$$\theta \longmapsto e^{i\theta}$$

$$\theta \longmapsto e^{-i\theta}$$

are both 2π-periodic. One can then check that the four functions

$$x \longmapsto \cos(2\pi\, kx)$$

$$x \longmapsto \sin(2\pi\, kx)$$

$$x \longmapsto e^{i2\pi\, kx}$$

$$x \longmapsto e^{-i2\pi\, kx}$$

are periodic with period $1/k$. Alternatively, we can say that they are of frequency k.

We now set a given length L. Because the $\sin(\cdot)$ and $\cos(\cdot)$ functions are 2π-periodic, are the functions

$$x \longmapsto \cos(2\pi\, x/L)$$

$$x \longmapsto \sin(2\pi\, x/L)$$

$$x \longmapsto e^{i2\pi\, x/L}$$

$$x \longmapsto e^{-i2\pi\, x/L}$$

all L-periodic. To the length L (or duration T) we associate the fundamental frequency Δk given by

$$\Delta k = 1/L$$

We define the “discrete frequencies” as the integer multiples of the fundamental frequency given by

$$k_m \coloneqq m\, \Delta k = m/L\,, m \in \mathbb{Z}$$

The positive integer multiples of the fundamental frequency are called the “harmonic frequencies” or just “harmonics” for short. Assuming m to be non-zero, the following two functions are then both of period L/m (i.e. of frequency k_m):

$$x \longmapsto \cos(2\pi\, k_m\, x) = \cos\left(2\pi \frac{m}{L} x\right)$$

$$x \longmapsto \sin(2\pi\, k_m\, x) = \sin\left(2\pi \frac{m}{L} x\right)$$

If $m = 0$ is $\cos(2\pi\, k_m\, x)$ constant and equal to 1, while $\sin(2\pi\, k_m\, x)$ is constant and equal to 0. These two functions can be written in terms of the two complex-valued functions

$$x \longmapsto e^{i2\pi\, k_m\, x} = e^{i2\pi \frac{m}{L} x}$$

$$x \longmapsto e^{-i2\pi\, k_m\, x} = e^{-i2\pi \frac{m}{L} x}$$

which are both of period L/m (i.e. of frequency k_m), except for $m = 0$ in which case they are constant and equal to 1.

Let be N an even positive integer, so that $N/2$ is a positive integer too. We define

$$\Delta x \coloneqq L/N$$

as well as

$$x_n \coloneqq n\, \Delta x$$

We will call the finite set

$$Grid_x = \{x_n | n \in \{-N/2, \dots, N/2 - 1\}\}$$

the "**standard x-grid**". We associate to that grid the interval

$$I_x \coloneqq [-L/2, L/2)$$

which is closed on the left and open on the right. It is an interval of length L centered in 0. We note that

$$Grid_x \subset I_x$$

We also define

$$x_{max} \coloneqq \frac{N}{2} \Delta x = L/2$$

Then is L equal to $2x_{max}$. We not that $-x_{max}$ is in I_x but x_{max} is not.

We have so far define the standard x-grid. We now define similarly the standard k-grid. We have already defined

$$\Delta k \coloneqq 1/L$$

as well as

$$k_m \coloneqq m\, \Delta k$$

We will call the finite set

$$Grid_k = \{k_m | m \in \{-N/2, \dots, N/2 - 1\}\}$$

the “**standard $\boldsymbol{k}$-grid**”. We associate to that grid the interval

$$I_k \coloneqq [-k_{max}, k_{max})$$

where

$$k_{max} \coloneqq \frac{N}{2} \Delta k$$

Again, the interval I_k is closed on the left and open on the right. It is an interval of length $2k_{max}$ centered in 0. We note that

$$Grid_k \subset I_k$$

We note also that $-k_{max}$ is in I_k but k_{max} is not. For convenience, we define

$$W \coloneqq 2k_{max}$$

It is the analogous of L in k-space. We will call L the “**FoV**” (for “field of view”) and we will call W the “**$\boldsymbol{k}$-FoV**”. We have

$$\Delta k = \frac{1}{L} \qquad \text{and} \qquad \Delta x = \frac{1}{W}$$

The standard grids are defined so that

$$\Delta k \, \Delta x = \frac{1}{N}$$

and therefore

$$k_m \, x_n = \frac{m \, n}{N}$$

These two relations are a key for defining the discrete Fourier transform.

Let be $f(\cdot)$ a function (periodic or not) defined on the x-space. We define then the vector $\vec{f}$ to be the column vector with components

$$f_n \coloneqq f(x_n) \qquad \text{with} \qquad x_n \in Grid_x$$

The matrix representation of $\vec{f}$ is

$$\begin{bmatrix} f_{-N/2} \\ \vdots \\ f_{-1} \\ f_0 \\ \vdots \\ f_{N/2-1} \end{bmatrix}$$

We will call vector $\vec{f}$ the sampling of function $f(\cdot)$. Similarly, given a function $g(\cdot)$ on the k-space, we define then the vector $\vec{g}$ to be the column vector with components

$$g_m \coloneqq g(k_m) \quad \text{with} \quad k_m \in Grid_k$$

The matrix representation of $\vec{g}$ is

$$\begin{bmatrix} g_{-N/2} \\ \vdots \\ g_{-1} \\ g_0 \\ \vdots \\ g_{N/2-1} \end{bmatrix}$$

We will call vector $\vec{g}$ the sampling of function $g(\cdot)$.

The standard grids can be summarized as follows.

Given N even and $n, m \in \{-N/2, \dots, N/2-1\}$, then

$$\Delta k = \frac{1}{L} = \frac{1}{2x_{max}} = \frac{W}{N} \qquad \Delta x = \frac{1}{W} = \frac{1}{2k_{max}} = \frac{L}{N}$$

$$k_m = m\,\Delta k \qquad x_n \coloneqq n\,\Delta x$$

$$k_0 = 0 \qquad \text{Position } N/2+1 \qquad x_0 = 0$$

$$k_m\, x_n = \frac{m\,n}{N}$$

$$\Delta k\, \Delta x = \frac{1}{N}$$

The Dirichlet kernels

We define here the "**Dirichlet kernels**" as well as some variants that we will also call "Dirichlet kernels", although it is not an official practice. All these kernels are actually functions. But we call them "kernels" to stress the fact that we will use them in "convolutions" with other functions, as described later in the course where convolutions are defined.

We define the N-th Dirichlet kernel $D_N(\cdot)$ as

$$D_N(x) \coloneqq \sum_{m=-N}^{N} e^{-imx} = \begin{Bmatrix} \sin\left(\left(N+\frac{1}{2}\right)x\right)/\sin\left(\frac{1}{2}x\right) & \text{for x not multiple of } 2\pi \\ 2N+1 & \text{for x multiple of } 2\pi \end{Bmatrix}$$

We note that the sum is symmetric and the definition can alternatively be given by

$$D_N(x) \coloneqq \sum_{m=-N}^{N} e^{imx}$$

It holds in fact

$$D_N(x) = D_N(-x)$$

This is the official definition, but for compatibility with the discrete Fourier transform presented later, we will work with $N/2$ instead of N (where N is considered to be an even number). It holds

$$D_{N/2}(x) = \sum_{m=-N/2}^{N/2} e^{-imx} = \begin{Bmatrix} \sin\left(\left(\frac{N}{2}+\frac{1}{2}\right)x\right)/\sin\left(\frac{1}{2}x\right) & \text{for x not multiple of } 2\pi \\ N+1 & \text{for x multiple of } 2\pi \end{Bmatrix}$$

The Dirichlet kernels are all real-valued and periodic with period 2π. For a later use, we define the "**symmetric Dirichlet kernel of period L**" by

$$\mathcal{S}_{\Delta k}^{N/2}(x) \coloneqq \Delta k \, D_{N/2}(2\pi \, \Delta k \, x) = \Delta k \sum_{m=-N/2}^{N/2} e^{i2\pi \, k_m x}$$

In that definition, the period L is hidden in the fundamental frequency Δk which is equal to $1/L$. The name "symmetric Dirichlet kernel of period L" is not official. We use it here for convenience. We added to adjective "symmetric" to stress the fact that it is defined by a symmetric sum $-N/2$ to $N/2$, since we will define other kernel with an asymmetric sum.

Substituting Δk by Δx in the definition of $\mathcal{S}_{\Delta k}^{N/2}(\cdot)$, and remembering that $2k_{max} = W$ is equal to $1/\Delta x$, we get the "**symmetric Dirichlet kernel of period W**" defined on k-space:

$$\mathcal{S}_{\Delta x}^{N/2}(k) \coloneqq \Delta x \, D_{N/2}(2\pi \, \Delta x \, k) = \Delta x \sum_{n=-N/2}^{N/2} e^{-i2\pi \, x_n k}$$

This is actually just a redefinition of $\mathcal{S}_{\Delta k}^{N/2}(x)$. In fact, substituting k by x (or inversely) in one leads the other. The definitions are chosen so that $\mathcal{S}_{\Delta k}^{N/2}(\cdot)$ and $\mathcal{S}_{\Delta x}^{N/2}(\cdot)$ are normalized:

$$\int_{-L/2}^{L/2} \mathcal{S}_{\Delta k}^{N/2}(x) \, dx = 1$$

$$\int_{-k_{max}}^{k_{max}} \mathcal{S}_{\Delta x}^{N/2}(k) \, dk = 1$$

We get finally the "**assymetric Dirichlet kernel of period L**" by mean of the asymmetric sum

$$\mathcal{A}_{\Delta k}^{N/2}(x) = \Delta k \sum_{m=-N/2}^{N/2-1} e^{-i2\pi \, k_m x} = \mathcal{S}_{\Delta k}^{N/2}(x) - \Delta k \, e^{i2\pi \, k_{max} x}$$

and the similarly "**assymetric Dirichlet kernel of period W**"

$$\mathcal{A}_{\Delta x}^{N/2}(k) = \Delta x \sum_{n=-N/2}^{N/2-1} e^{-i2\pi \, x_n k} = \mathcal{S}_{\Delta x}^{N/2}(k) - \Delta x \, e^{-i2\pi \, x_{max} k}$$

We see that $\mathcal{A}_{\Delta k}^{N/2}(\cdot)$ resp. $\mathcal{A}_{\Delta x}^{N/2}(\cdot)$ are the same like $\mathcal{S}_{\Delta k}^{N/2}(\cdot)$ resp. $\mathcal{S}_{\Delta x}^{N/2}(\cdot)$ up to a high frequency oscillation of small amplitude Δk resp. Δx.

An important fact is that the oscillation $e^{i2\pi \, k_{max} x}$ do not contribute to the "area under the curve" of the complex-valued function $\mathcal{A}_{\Delta k}^{N/2}(\cdot)$ because it holds

$$\int_{-L/2}^{L/2} e^{-i2\pi \, k_{max} x} \, dx = 0$$

The same remark holds for the integration from $-k_{max}$ to k_{max} of $e^{-i2\pi\, x_{max} k}$. The two asymmetric kernels are therefore normalized.

$$\int_{-L/2}^{L/2} \mathcal{A}_{\Delta k}^{N/2}(x)\, dx = 1$$

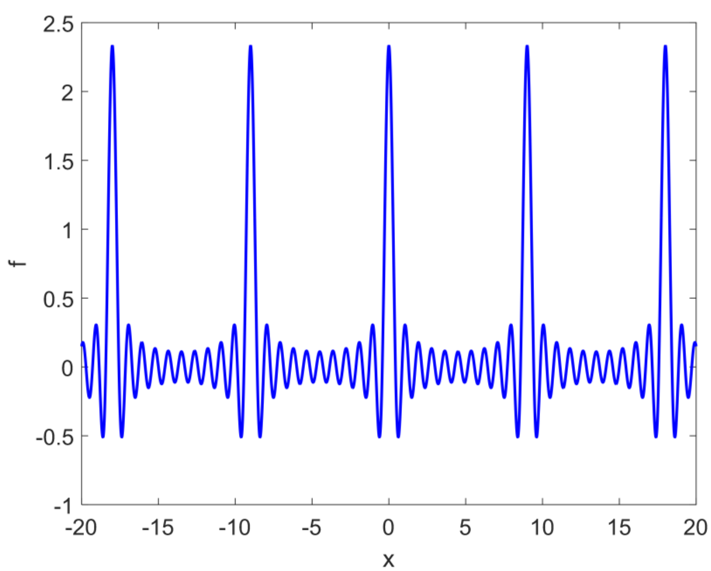
$$\int_{-k_{max}}^{k_{max}} \mathcal{A}_{\Delta x}^{N/2}(k)\, dk = 1$$

In summary, we have:

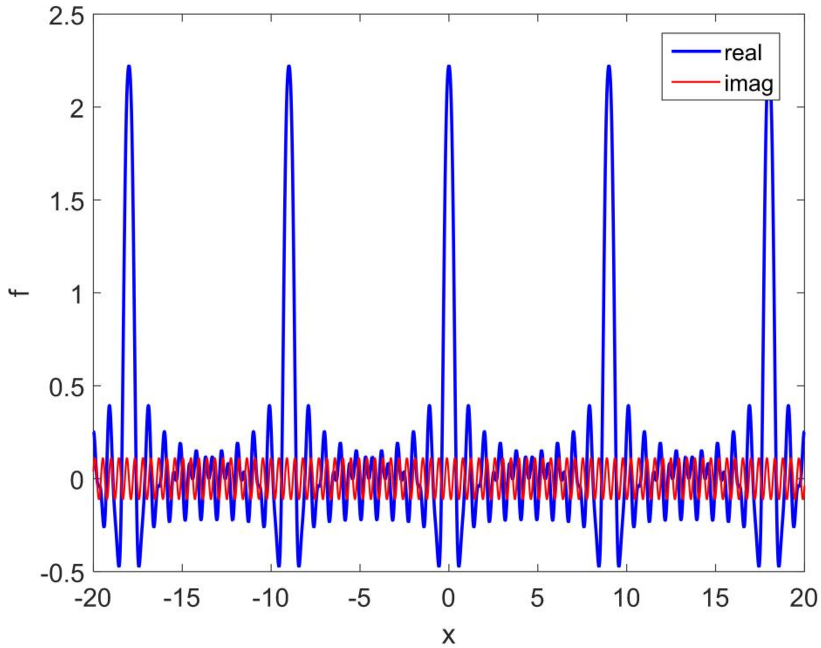
- two symmetric, real-valued, normalized kernels: one of period L on the x-space and one of period W on the k-space,
- two asymmetric, complex-valued, normalized kernels: one of period L on the x-space and one of period W on the k-space.

The complex-valued asymmetric kernels are almost real-valued if we neglect the small complex-valued high frequency oscillation of amplitude Δk resp. Δx.

The following figures displays a symmetric Dirichlet kernel,



while the following displays an asymmetric one.



The door functions

We define the **big-door function** of width L (and centered in 0) as

$$\Pi_L : \mathbb{R} \longrightarrow \mathbb{R}$$

$$x \longmapsto \Pi_L(x) = \begin{Bmatrix} 0 & x < -L/2 \\ 1 & -L/2 \leq x < L/2 \\ 0 & L/2 \leq x \end{Bmatrix}$$

The big-door function of width L centered in a is simply given by

$$x \longmapsto \Pi_L(x - a)$$

We define the little-door function of width L (and centered in 0 as) $\pi_L(\cdot)$ by

$$\pi_L(x) = \frac{1}{L}\Pi_L(x)$$

We observe that the door functions are not exactly symmetric due to the asymmetry in the inequalities in their definitions. In fact is Π_L the characteristic function of interval I_x, which is closed on the left but open on the right.

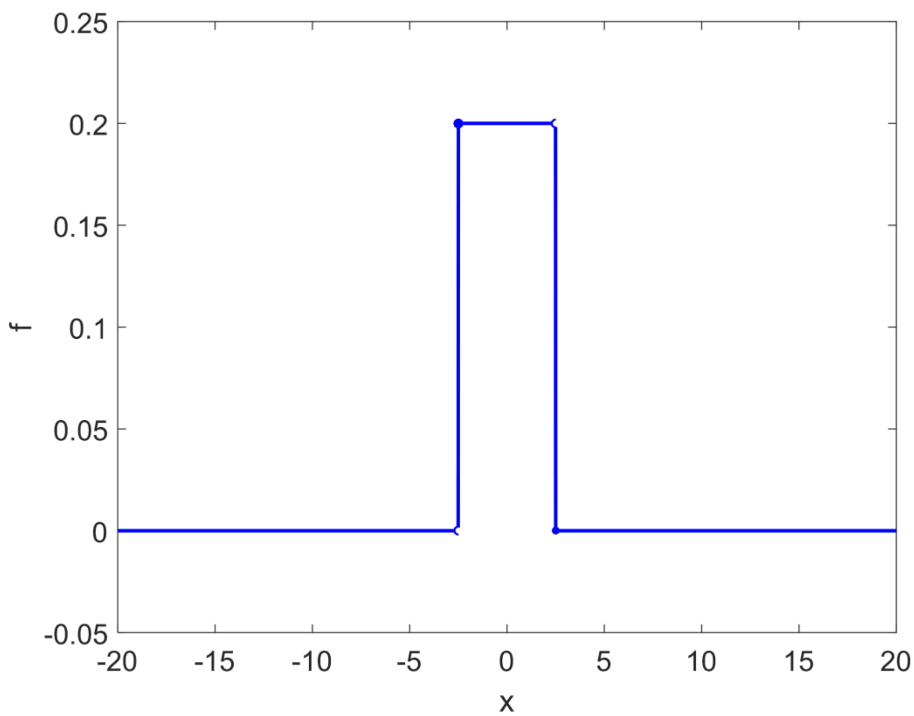
The little-door function is normalized in the sense that

$$\int_{-\infty}^{+\infty} \pi_L(x)dx = 1$$

while the big-door function is not normalized (except if L equals 1) and it holds

$$\int_{-\infty}^{+\infty} \Pi_L(x)dx = L$$

Note finally that for $0 < L < 1$ is the graph of the little-door function is “taller” than the one of the big-door function. “Big” and “little” are just names. The following figure displays de graph of a little-door function



The sinus-cardinal functions

The $sinc(\cdot)$ function (sinus-cardinal) is given by

$$sinc(x) = \begin{cases} \sin(x)/x & x \neq 0 \\ 1 & x = 0 \end{cases}$$

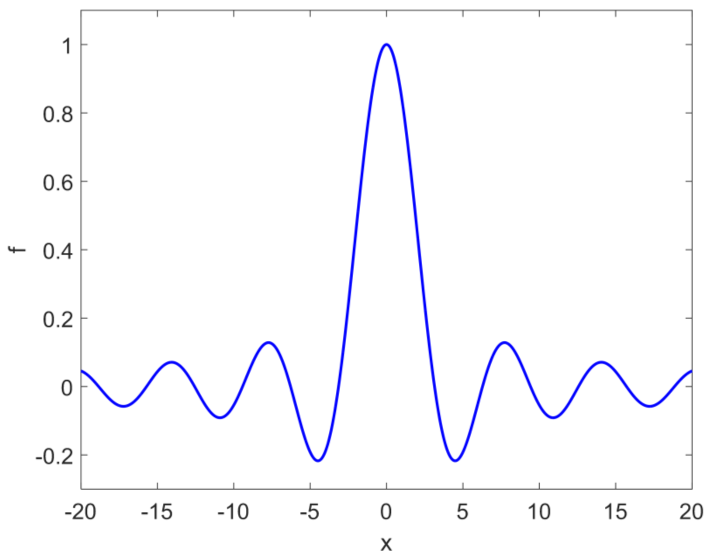
More generally, we will also name “$sinc(\cdot)$-function” any variant given by

$$x \longmapsto A \cdot \mathrm{sinc}(s \cdot (x - a))$$

with real parameters A , a and s. As we will see later, the $sinc(\cdot)$ function

$$k \longmapsto sinc(L\pi\, k)$$

is the Fourier transform of $\pi_L(\cdot)$.The following figure displays the graph of a $sinc(\cdot)$ function



The Gaussian functions

We will name "Gaussian-function" or just "Gaussian" any function of the form

$$x \longmapsto A \cdot \mathrm{e}^{-\frac{1}{2\sigma^2}(s \cdot x - \mu)^2}$$

with real parameters A, σ, μ and s. In the special case where

$$A = \frac{1}{\sqrt{2\pi}\,\sigma} \qquad \text{and} \qquad s = 1$$

is the corresponding Gaussian normalized in the sense that

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma}\, \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2} = 1$$

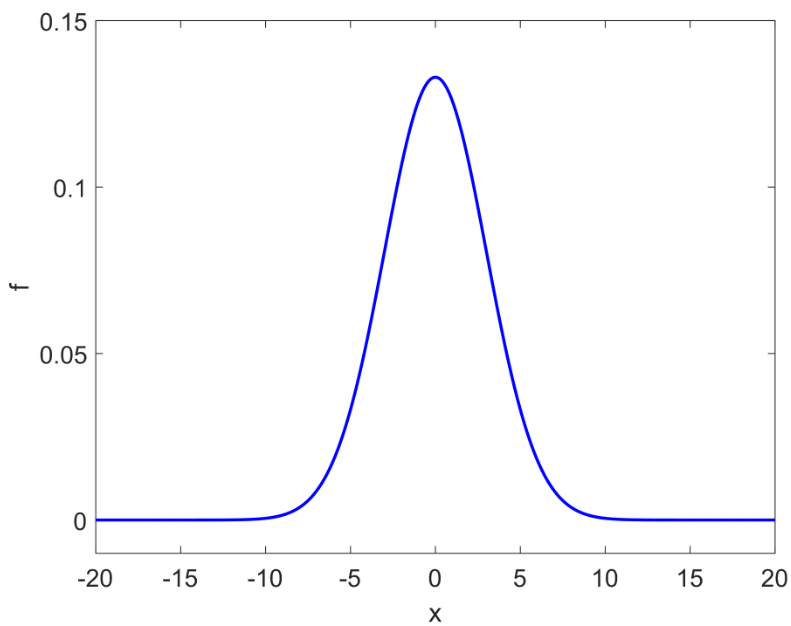
This Gaussian function can then be interpreted as a probability distribution (called "**normal distribution**") with mean μ and variance σ^2. We will work in particular with the Gaussian $\mathcal{G}_\sigma(\cdot)$ define by

$$\mathcal{G}_\sigma(x) \coloneqq \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \mathrm{e}^{-\frac{1}{2\sigma^2}x^2}$$

which is a normal probability distribution with mean 0 and variance σ^2. As we will see, its Fourier transform is the function $\mathcal{F}\mathcal{G}_\sigma(\cdot)$ given by

$$\mathcal{F}\mathcal{G}_\sigma(k) = \mathrm{e}^{-2\pi^2\sigma^2k^2}$$

which is also a Gaussian, but which is not normalized and is therefore not a probability distribution. The following figure displays the graph of a Gaussian function.



Some function operations

Just as numbers, real- and complex-valued functions can be added, subtracted and multiplied (we will not go into the division of a function by another). But unlike numbers, we can perform many more operation on and between them. We give here a few that are useful for our purpose.

Let be two real- or complex-valued functions $f(\cdot)$ and $h(\cdot)$ defined on the line of real numbers. We define the new function $\{f + h\}(\cdot)$ as

$$\{f + h\}(x) \coloneqq f(x) + h(x)$$

We define the new function $\{f \cdot h\}(\cdot)$ as

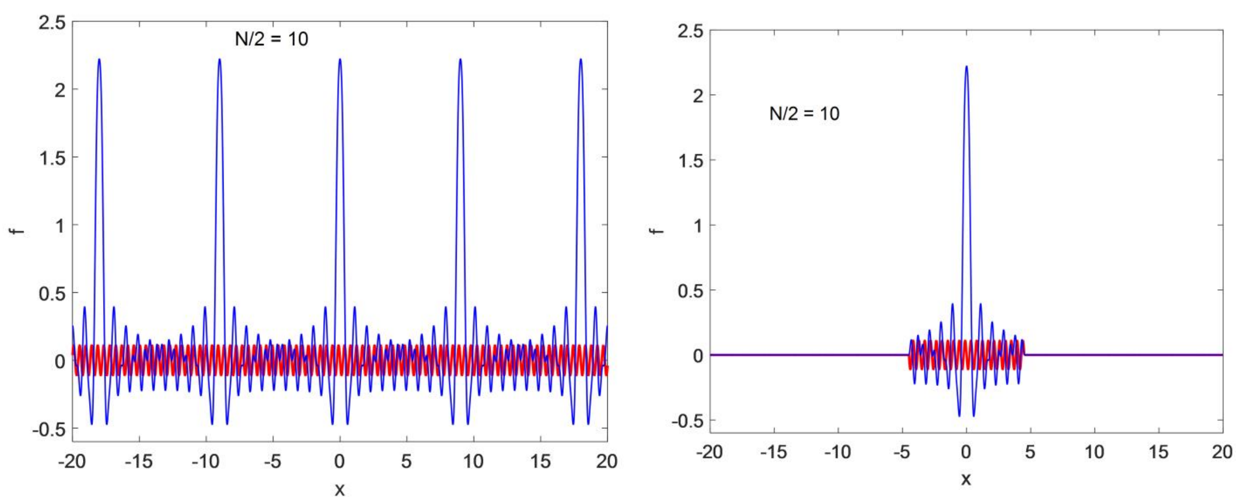
$$\{f \cdot h\}(x) \coloneqq f(x) \cdot h(x)$$

The subtraction can be defined similarly if needed. Given the big-door function $\Pi_L(\cdot)$ of width L, we define the crope of $f(\cdot)$ of width L by the new function $\Pi_L f(\cdot)$ given by

$$\Pi_L f(x) = \Pi_L(x) \cdot f(x) = \begin{Bmatrix} 0 & x < -L/2 \\ f(x) & -L/2 \leq x < L/2 \\ 0 & L/2 \leq x \end{Bmatrix}$$

We observe the subtle asymmetry of the inequalities in the definition. $\Pi_L(\cdot)$ is in fact the indicator function of the interval I_x, which is asymmetric.

The following figure displays the graph of the asymmetric Dirichlet kernel $\mathcal{A}_{\Delta k}^{N/2}(\cdot)$ on the left and of its crope $\Pi_L \mathcal{A}_{\Delta k}^{N/2}(\cdot)$ of width $L = 1/\Delta k$ so that the "remaining" part of $\Pi_L \mathcal{A}_{\Delta k}^{N/2}(\cdot)$ correspond to one period of $\mathcal{A}_{\Delta k}^{N/2}(\cdot)$. The real part is in blue and the imaginary part in red.



Given a real number a, we define the shift of $f(\cdot)$ by a as new function $S_a f(\cdot)$ given by

$$S_a f(x) \coloneqq f(x - a)$$

Depending on the context, we will also write $S_a f(\cdot)$ as $f(\cdot - a)$ if it is more practical.

Given a function $f(\cdot)$, we define its **periodic summation** of period L as the new function $P_L^S f(\cdot)$ given by

$$P_L^S f(\cdot) = \sum_{z \in \mathbb{Z}} f(\cdot - zL) = \sum_{z \in \mathbb{Z}} S_{zL} f(\cdot)$$

The periodic summation of a function is not necessarily well defined. The previous serie must converge for every x for the periodic summation to exist.

For example, if $f(\cdot)$ is identically equal to 1, the serie defining $P_L^S f(\cdot)$ diverges for every x. But the serie will converge if $f(\cdot)$ converges to 0 rapidly enough at infinity on both sides. If the serie converges for every x is then $P_L^S f(\cdot)$ a function of period L.

Given a function $f(\cdot)$, we define its **periodic extension** of period L as the new function $P_L^E f(\cdot)$ given by

$$P_L^E f(\cdot) = P_L^S \Pi_L f(\cdot) = \sum_{z \in \mathbb{Z}} \Pi_L f(\cdot - zL) = \sum_{z \in \mathbb{Z}} S_{zL} \Pi_L f(\cdot)$$

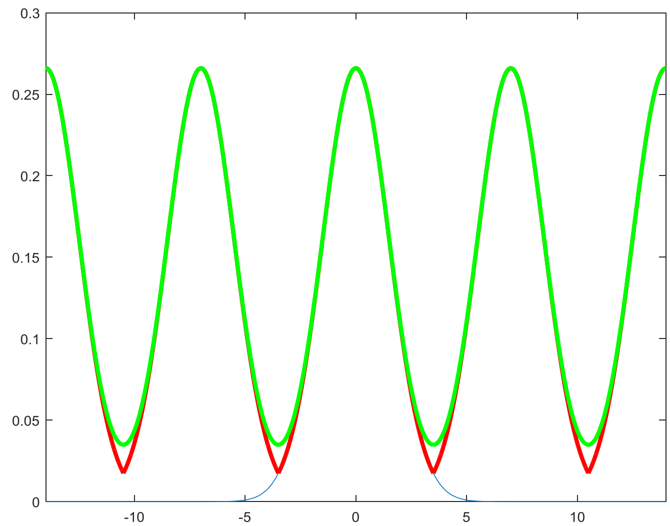
While the periodic summation do not always exists, the periodic extension always exists and is a function of period L. The periodic extension is obtained by cropping the function with width L and then replicating it on both sides of the x-axis with period L.

Given a function that converges to 0 sufficiently rapidly, we have (at least) two ways to fabricate a periodic version of it: the periodic summation and the periodic extension. For any other function, the periodic extension "does the job".

We note finally that a function $f(\cdot)$ is periodic of period L if and only if it is identical to its periodic extension of period L. In that case holds

$$f(\cdot) = P_L^E f(\cdot) = \sum_{z \in \mathbb{Z}} \Pi_L f(\cdot - zL)$$

The following figure shows the graph of a function (in blue) together with the graph of its periodic extension (in red) and its periodic summation (in green):



There is a very important operation called the "convolution" which is related to the Fourier transform in a special way. The convolution of $f(\cdot)$ and $h(\cdot)$ is a new function that we will write $\{f * h\}(\cdot)$ but we will defined it later in the course.

We note at that point that the Fourier transform (resp. its inverse) is also an operation which takes as argument a function $f(\cdot)$ defined on x-space (resp. a function $g(\cdot)$ defined on k-space) and return another function $\mathcal{F}f(\cdot)$ defined on k-space (resp. a function $\mathcal{F}^{-1}g(\cdot)$ defined on x-space).

Part II: Fourier analysis

Periodic Fourier analysis: The Fourier coefficient and the associated Fourier serie

The objects to be decomposed in the context of periodic Fourier analysis are periodic functions. The analysis of periodic functions is done by evaluating its Fourier coefficients. In the present case, the spectrum is the list of Fourier coefficients. The synthesis is achieved by building back the function from its Fourier coefficients. This is done by building the Fourier serie. There is one version of the theory for real-valued functions and another for complex-valued function. We present both.

Let be $f(\cdot)$ a real-valued periodic function of period L defined on the set of real numbers:

$$f: \mathbb{R} \longrightarrow \mathbb{R}$$

$$x \longmapsto f(x)$$

We define the real Fourier coefficients of $f(\cdot)$ as

$$a_0 = \frac{1}{L} \int_{-L/2}^{L/2} f(x)\, dx$$

$$a_m = \frac{2}{L} \int_{-L/2}^{L/2} f(x) \cos(2\pi\, k_m x)\, dx$$

$$b_m = \frac{2}{L} \int_{-L/2}^{L/2} f(x) \sin(2\pi\, k_m x)\, dx$$

for $m = 1, 2, 3, \ldots$. The evaluation of these coefficients constitutes the analysis. The list

$$a_0, a_1, b_1, a_2, b_2, \ldots$$

is the spectrum. For symmetry reasons, one can also add $b_0 = 0$ in the spectrum. Then can the spectrum be seen as a function defined on the set of non-negative discrete frequencies

$$\{k_m |\, k_m = m\, \Delta k, m \in \mathbb{Z}, m \geq 0\}$$

i.e. the fundamental frequency and its harmonics. If seen as a function, the spectrum can be written:

$$k_m \longmapsto \binom{a_m}{b_m}$$

This spectrum can be represented graphically by writing the discrete frequencies on the horizontal axis and the Fourier coefficients on the vertical axis.

We define the function $\mathcal{S}_{real}^{N/2} f(\cdot)$ as

$$\mathcal{S}_{real}^{N/2} f(x) := \; a_0 + \sum_{m=1}^{N/2} a_m \cos(2\pi \, k_m x) + b_m \sin(2\pi \, k_m x)$$

and we call it the "**real Fourier serie with $\boldsymbol{N/2+1}$ terms**" for function $f(\cdot)$. The evaluation of this function constitutes the synthesis. As indicated by its name, it is the **real Fourier serie**. The synthesis is not exactly equal to the original function $f(\cdot)$ but improves as N becomes larger and larger.

Because this serie is partial (we do not consider the limit with N going to infinity) we may also call $\mathcal{S}_{real}^{N/2} f(\cdot)$ a partial Fourier serie. Theoretician also work with the complete serie by including all the terms (an infinite number of terms) and this is what we call the (true) real Fourier serie. But we will only consider partial Fourier series in this course for the sake of applicability.

The real Fourier serie with real coefficient is however limited to the analysis of real periodic functions. The complex version of the analysis allows to handle both real and complex valued periodic functions and we will therefore only work with the complex version in the rest of the text. Here it is.

Let be $f(\cdot)$ a complex-valued periodic function of period L defined on the line of real numbers:

$$f: \mathbb{R} \longrightarrow \mathbb{C}$$

$$x \longmapsto f(x)$$

We define the (complex) Fourier coefficient of $f(\cdot)$ as

$$c_m = \frac{1}{L} \int_{-L/2}^{+L/2} f(x) e^{-i2\pi \, k_m x}$$

and we define the function $\mathcal{S}^{N/2} f(\cdot)$ as

$$\mathcal{S}^{N/2} f(x) := \sum_{m=-N/2}^{N/2} c_m \, e^{i2\pi \, k_m \, x}$$

that we will call the "**Fourier serie of $f(\cdot)$ with $N+1$ terms**". The list of Fourier coefficients constitute the spectrum and it can be seen as a function defined on the set of discrete frequencies

$$\{k_m | \, k_m = m \, \Delta k, m \in \mathbb{Z}\}$$

If seen as a function, the spectrum can be written

$$k_m \longmapsto c_m$$

We note that the definitions are simpler in the complex case than in the real case. It seems that "complex" do not means "complicated".

In general, a complex valued periodic function can only be analyzed in terms of complex Fourier coefficients. But a real valued function can be analyzed both in terms of complex and real Fourier coefficients. For that special case holds

$$a_n = c_n + c_{-n}$$

$$b_n = i(c_n - c_{-n})$$

and

$$c_n = \frac{1}{2}(a_n - i \, b_n)$$

$$c_{-n} = \frac{1}{2}(a_n + i \, b_n) = c_n^*$$

where the star symbol $\cdot^*$ stands for the complex conjugate of a complex value. We note that the sum defining $\mathcal{S}^{N/2} f(\cdot)$ is symmetric in the sense that it starts at index $-N/2$ and ends at index $+N/2$. We will therefore also refer to $\mathcal{S}^{N/2} f(\cdot)$ as the "**symmetric Fourier serie with $N+1$ terms**" for function $f(\cdot)$.

The reason why we introduce the adjective "symmetric" is that we will also need a slightly different serie for the description of the discrete Fourier transform. We define for that purpose the "**asymmetric Fourier serie with N terms**" for function $f(\cdot)$ as

$$\mathcal{A}^{N/2} f(x) \coloneqq \sum_{m=-N/2}^{N/2-1} c_m \, e^{i2\pi \, k_m x}$$

It is the same like $\mathcal{S}^{N/2} f(\cdot)$ but with the last term missing. It is also a synthesis which approaches function $f(\cdot)$ as N becomes larger and larger. It is maybe not as good as $\mathcal{S}^{N/2} f(\cdot)$ because one term is missing. But it is probably as good as $\mathcal{S}^{(N-1)/2} f(\cdot)$. We will come back to that asymmetric serie in the section about the discrete Fourier transform.

Let be $f(\cdot)$ and $g(\cdot)$ two periodic functions of period L and let be $c_m(f)$ the m-th Fourier coefficient of $f(\cdot)$ as well as $c_m(g)$ the m-th Fourier coefficient of $g(\cdot)$. The Parseval-Plancherel identity for periodic functions reads then

$$\frac{1}{L} \int_{-L/2}^{L/2} \overline{f(x)} \, g(x) \, dx = \sum_{m=-\infty}^{+\infty} \overline{c_m(f)} \, c_m(g)$$

We define the bracket $\langle \cdot \mid \cdot \rangle_X$ by

$$\langle f(\cdot) | g(\cdot) \rangle_X \coloneqq \frac{1}{L} \int_{-L/2}^{L/2} \overline{f(x)} \, g(x) \, dx$$

which is an inner product of vector spaces under some conditions. Let be

$$c(f) \coloneqq (\dots, c_{-2}(f), c_{-1}(f), c_0(f), c_1(f), c_2(f), \dots)$$

$$c(g) \coloneqq (\dots, c_{-2}(g), c_{-1}(g), c_0(g), c_1(g), c_2(g), \dots)$$

We define then the bracket $\langle \cdot \mid \cdot \rangle_K$ by

$$\langle c(f) | c(g) \rangle_K \coloneqq \sum_{m=-\infty}^{+\infty} \overline{c_m(f)} \, c_m(g)$$

which is an inner product of vector spaces under some conditions. Then reads the Parseval-Plancherel theorem

$$\langle f(\cdot)|g(\cdot)\rangle_X = \langle c(f)|c(g)\rangle_K$$

The map that send a function $f(\cdot)$ to its list of Fourier coefficients is in that case an isometry, the definitions

$$\|f(\cdot)\|_{X,2}^2 := \langle f(\cdot)|f(\cdot)\rangle_X$$

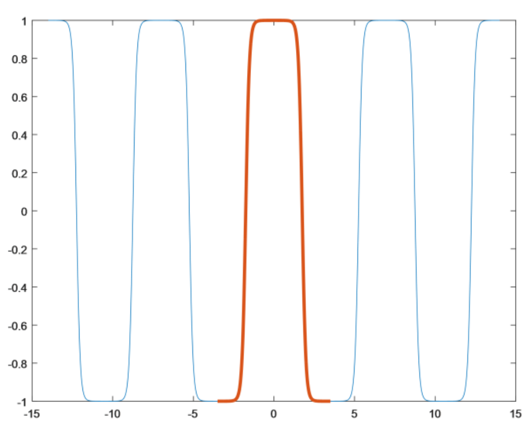
$$\|c(f)\|_{K,2}^2 := \langle c(f)|c(f)\rangle_K$$

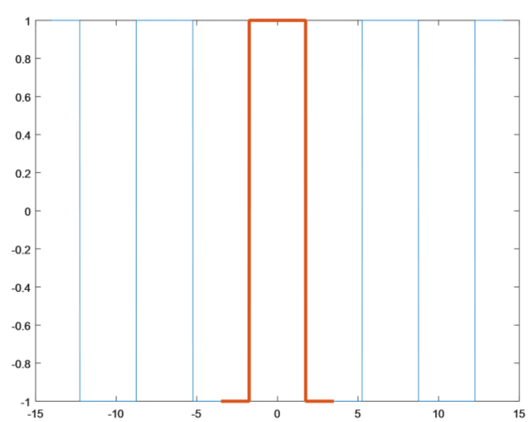
define some 2-norms and the Parseval-Plancherel identity leads to

$$\|f(\cdot)\|_{X,2}^2 = \|c(f)\|_{K,2}^2$$

We terminate this section with two examples. We empirically deduce some general rules out of these examples without proving those rules. But we can be sure they are correct because mathematicians have done the job for us.

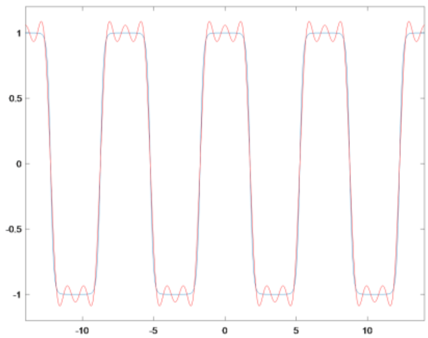
The following figure displays the graph of two periodic functions (one period is in orange). One function is smooth (left) and the other one is not continuous (right). But both have the same period L.



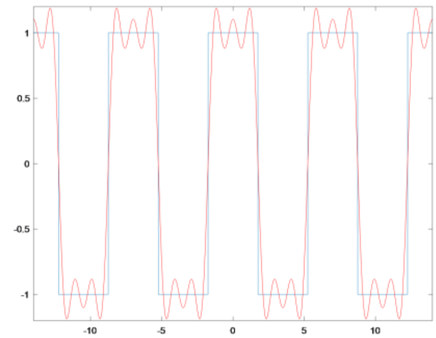


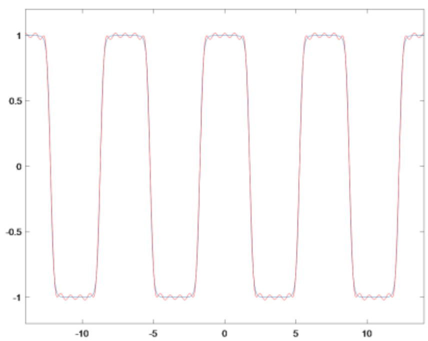
The next figure displays the graphs of some (symmetric) Fourier series (in red) together with the original function (in blue) for an increasing number of terms in the serie. We see that the more terms we add to the Fourier serie, the closer is the Fourier serie to the original function.

We see however that the Fourier serie is closer to the original function in the case of the smooth function (left). For the non-continuous function (right), some oscillations persist and do not improve with an increasing number of terms in the Fourier serie. These oscillations are symptomatic of discontinuities and do not disappear even for a large number of terms. They are called "**Gibb's artefacts**".

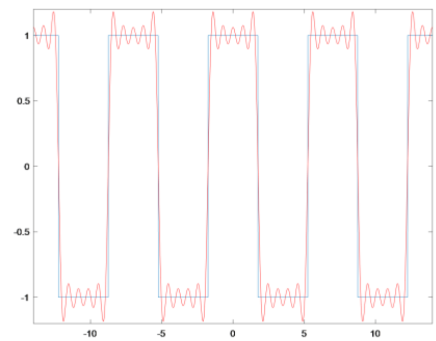


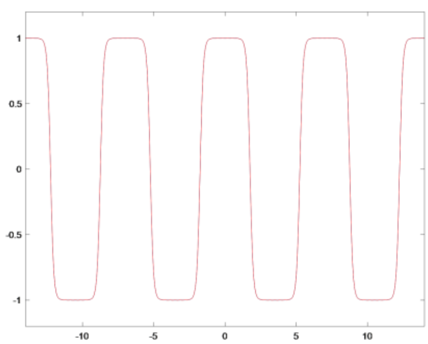
N/2 = 3



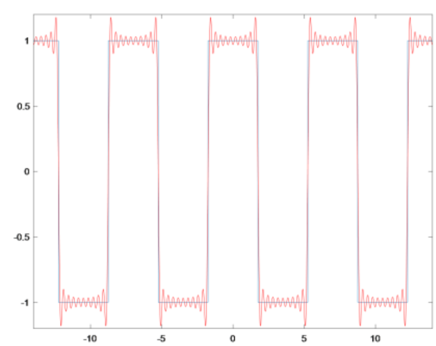


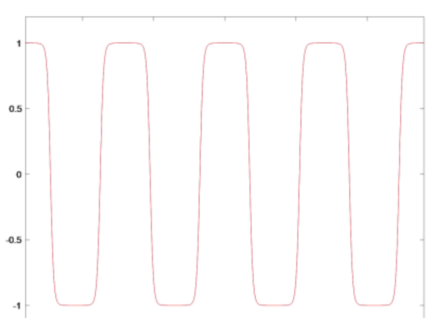
N/2 = 5



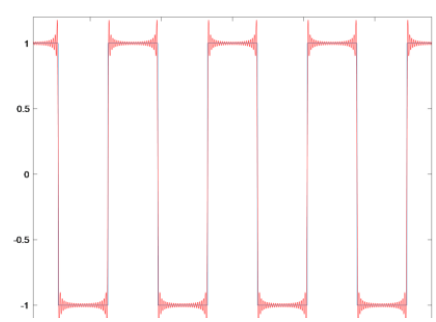


N/2 = 10



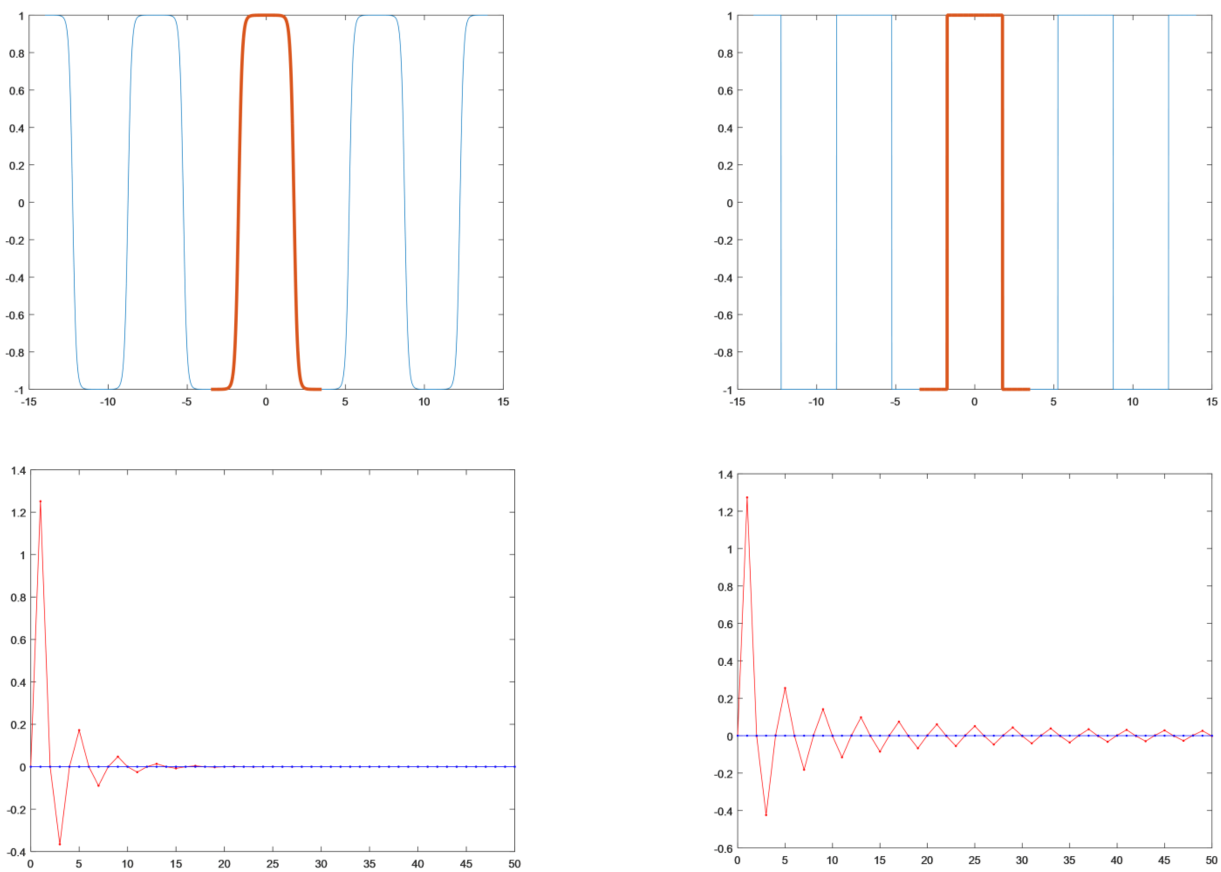


N/2 = 50



The graph of both functions is displayed one more time in the following figure. Their corresponding spectrum, their Fourier coefficients, is displayed below. The cosine-coefficient is in red and sin-coefficient is in blue. The Fourier coefficients are discrete: there is one cosine- and one sin- coefficient for each discrete frequency (the lines between the dots are present only to improve visibility).

We observe that the spectrum of the smooth function (without significant Gibb's artefacts) decays much fast than the spectrum of the non-continuous function. This is true in general: the smoother is a function, the faster decrease its spectrum along increasing frequencies.



Non-periodic Fourier analysis: The Fourier transform and its inverse transform

Let be $f(\cdot)$ a real- or complex-valued function defined on the line of real numbers. We assume that $f(\cdot)$ is sufficiently well behaved (in particular, we assume that it converges to 0 at infinity sufficiently rapidly). We define its Fourier transform (its spectrum) $\mathcal{F}f(\cdot)$ as the new function given by

$$\mathcal{F}f(k) := \int_{-\infty}^{+\infty} f(x)\, e^{-i2\pi\, kx}\, dx$$

Now, if you ask yourself what "sufficiently well behaved" means (or what "converges to 0 at infinity sufficiently rapidly" means) it means any condition that guaranties the convergence of the previous integral.

Assuming a function $g(\cdot)$ to be sufficiently well behaved (or suitable) defined on the k-space (i.e. $g(\cdot)$ is a spectrum), we define its inverse Fourier transform $\mathcal{F}^{-1}g(\cdot)$ as the new function defined by

$$\mathcal{F}^{-1}g(x) := \int_{-\infty}^{+\infty} g(k)\, e^{i2\pi\, kx}\, dk$$

The definition domain of $\mathcal{F}^{-1}g(\cdot)$ is the x-space. In a simple word, the inverse Fourier transform $\mathcal{F}^{-1}$ should be the inverse operation of the Fourier transform $\mathcal{F}$ i.e.

$$\mathcal{F}^{-1}\mathcal{F}f(x) = f(x)$$

and

$$\mathcal{F}\mathcal{F}^{-1}g(k) = g(k)$$

for every x and k. But the word is not simple and although those rules hold for many functions, it breaks down for some other. We will give one example where it works, and another where there is some problem.

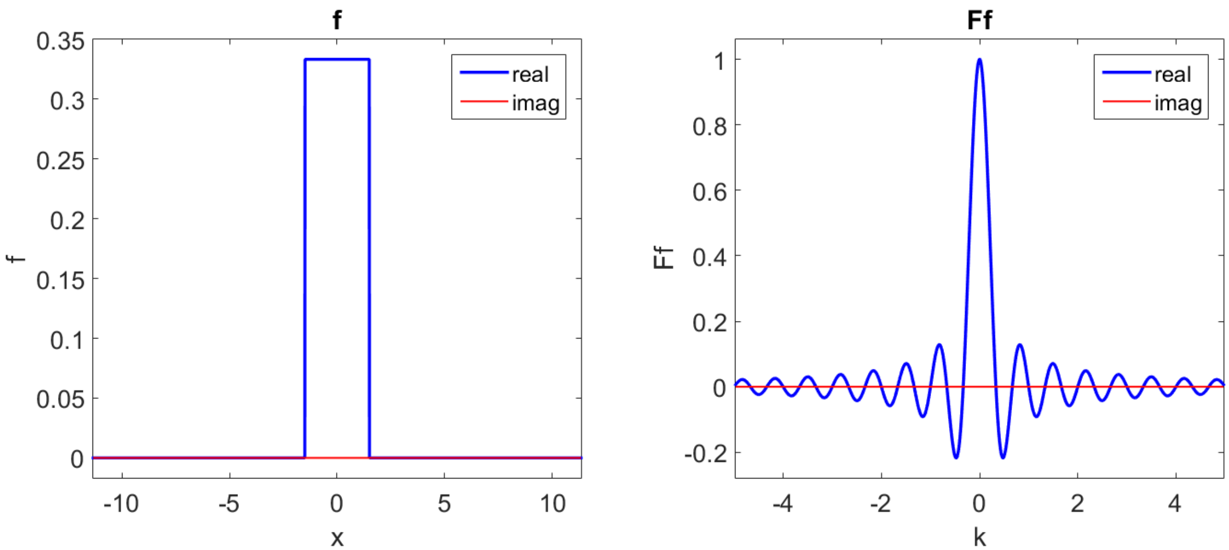
The door functions have a well-defined Fourier transform. The reader can check as an exercise that, by evaluating the Fourier integral, one finds

$$\mathcal{F}\pi_L(k) = sinc(L\pi k)$$

As a consequence of the linearity of the Fourier transform, it holds

$$\mathcal{F}\Pi_L(k) = L\ sinc(L\pi k)$$

The following figure displays little-door function together with its Fourier transform.



Now, the inverse Fourier transform of $\mathcal{F}\pi_L(\cdot)$ does not exist as a function because the Fourier integral

$$\int_{-\infty}^{+\infty} sinc(L\pi k)\ e^{i2\pi\ kx}\ dk$$

does not converge for $x = L/2$ and $x = -L/2$. But this is not surprising because if we would have defined $\pi_L(\cdot)$ with different values at $x = L/2$ and $x = -L/2$, then its Fourier transform would still be exactly $sinc(L\pi\ \cdot)$ (because the integral is insensitive about a change of the integrand on a set of measure 0). Therefore, the door function cannot be retrieved back from its Fourier transform.

Mathematicians have overcome the problem with the theories of distributions, but we will not go into this in the present course. For us, the inverse Fourier transform of $sinc(L\pi\ \cdot)$ just does not exist as a function defined on $\mathbb{R}$. What we may be able to write is

$$\{\mathcal{F}^{-1} sinc(L\pi\ \cdot)\}(x) = \pi_L(x) \quad \text{for } |x| \neq 1$$

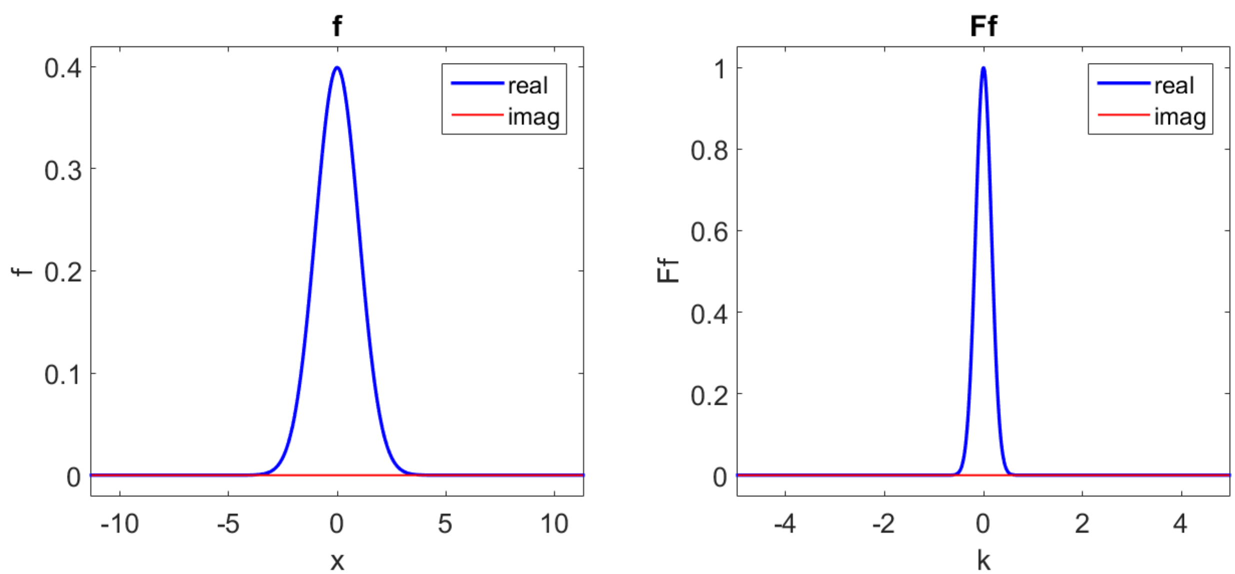
The Gaussian functions have a well define Fourier transform. The Fourier transform of $\mathcal{G}_\sigma(\cdot)$ given by

$$\mathcal{G}_\sigma(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \mathrm{e}^{-\frac{1}{2\sigma^2}x^2}$$

is the function $\mathcal{F}\mathcal{G}_\sigma(\cdot)$ given by

$$\mathcal{F}\mathcal{G}_\sigma(k) = \mathrm{e}^{-\pi^2\sigma^2k^2}$$

and the inverse Fourier transform of the second function is the first one. Everything runs smoothly with Gaussian functions. The following figure displays the graph of both functions.



We terminate this section about the Fourier transform by giving a few identities.

We note that

$$\mathcal{F}f(0) = \int_{-\infty}^{+\infty} f(x)\ dx$$

is the “sum” of the function values.

We also note that

$$\mathcal{F}S_a f(k) = \int_{-\infty}^{+\infty} f(x-a)\, e^{-i2\pi\, kx}\, dx = \int_{-\infty}^{+\infty} f(x')\, e^{-i2\pi\, k(x'+a)}\, dx'$$

$$= \int_{-\infty}^{+\infty} f(x)\, e^{-i2\pi\, kx} e^{-i2\pi\, ka}\, dx = \mathcal{F}f(k) \cdot e^{-i2\pi\, ka}$$

and a similar identity can be shown for the inverse Fourier transform. It follows the Fourier-shift-theorem:

$$\mathcal{F}S_a f(k) = \mathcal{F}f(k) \cdot e^{-i2\pi\, ka}$$

$$\mathcal{F}^{-1}S_a g(x) = \mathcal{F}^{-1}g(k) \cdot e^{i2\pi\, xa}$$

We define the brackets $\langle \cdot \mid \cdot \rangle_X$ and $\langle \cdot \mid \cdot \rangle_K$ by

$$\langle f_1(\cdot) | f_2(\cdot) \rangle_X \coloneqq \int_{-\infty}^{+\infty} \overline{f_1(x)}\, f_2(x)\ dx$$

$$\langle g_1(\cdot) | g_2(\cdot) \rangle_K \coloneqq \int_{-\infty}^{+\infty} \overline{g_1(k)}\, g_2(k)\ dk$$

which define some inner products of vector spaces under some conditions. Each one takes two functions as arguments and returns a complex value. The Parseval-Plancherel identities reads then

$$\langle f_1(\cdot) | f_2(\cdot) \rangle_X = \langle \mathcal{F}f_1(\cdot) | \mathcal{F}f_2(\cdot) \rangle_K$$

$$\langle g_1(\cdot) | g_2(\cdot) \rangle_K = \langle \mathcal{F}^{-1}g_1(\cdot) | \mathcal{F}^{-1}g_2(\cdot) \rangle_X$$

If the brackets defined above are inner products of vector spaces, then

$$\| f(\cdot) \|_{X,2}^2 \coloneqq \langle f(\cdot) | f(\cdot) \rangle_X$$

$$\| g\,(\cdot) \|_{K,2}^2 \coloneqq \langle g(\cdot) | g\,(\cdot) \rangle_K$$

define some 2-norms and the Parseval-Plancherel identities lead to

$$\| f(\cdot) \|_{X,2}^2 = \| \mathcal{F}f(\cdot) \|_{K,2}^2$$

$$\| \mathcal{F}^{-1}g(\cdot) \|_{X,2}^2 = \| g(\cdot) \|_{K,2}^2$$

Discrete Fourier Analysis: the discrete Fourier transform and its inverse Transform

Given a real- or complex-valued column vector $\vec{f}$ of length N and given a positive length L, then can $\vec{f}$ be seen as the sampling on the standard x-grid of step size Δx (equal to L/N) of a function $f(\cdot)$ defined on the x-space. The choice of the function $f(\cdot)$ is obviously not unique, but there exist at least one.

Alternatively, given a function $f(\cdot)$ defined on the x-space and given the standard x-grid associated to the interval I_x (closed on the left and open on the right) of length L and centered in 0, then is the sampling of $f(\cdot)$ a column vector $\overrightarrow{f}$ of length N.

Note that in both cases, function $f(\cdot)$ can be periodic or not. It does not matter. And in both cases we have a vector $\vec{f}$ of length N and a standard x-grid.

We want now do define the discrete Fourier transform of $\vec{f}$. We will write the discrete Fourier transform (DFT) as F. It is a linear and invertible map i.e. an isomorphism of vector spaces. As a consequence, F can be interpreted as a matrix. We define the matrix of F entry wise as follows:

$$F_{mn} \coloneqq \Delta x \, e^{-i2\pi \, k_m x_n} \qquad \text{with} \qquad n, m \in \{-N/2, \dots, N/2-1\}$$

where

$$\Delta x = \frac{L}{N}$$

as defined earlier. From the relation

$$k_m \, x_n = \frac{n\, m}{N}$$

follows

$$F_{mn} \coloneqq \Delta x \, e^{-i2\pi \frac{n\, m}{N}} \qquad \text{with} \qquad n, m \in \{-N/2, \dots, N/2-1\}$$

The way we define the DFT do not only depends on the vector size N as it is usually the case. In our definition, it also depends on the step size Δx of the standard x-grid, what some reader may not accept. We will justify it from a mathematical point of view and from a physical point of view (for consistence of units).

But before that, we will write the matrix representation of F and we will describe the connection with the matrix Ω, which is involved in the fast Fourier transform (FFT). We define the complex number

$$\phi \coloneqq e^{-i2\pi/N}$$

So that

$$F_{mn} \coloneqq \Delta x \, \phi^{n\,m}$$

The matrix of F is then

$$F = \Delta x \begin{bmatrix} \phi^{(-N/2)(-N/2)} & \dots & 1 & \dots & \phi^{(-N/2)(N/2-1)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & \dots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi^{(N/2-1)(-N/2)} & \dots & 1 & \dots & \phi^{(N/2-1)\,(N/2-1)} \end{bmatrix}$$

The column of 1's correspond to $m = 0$ and the raw of 1's corresponds to $n = 0$.

The fast-Fourier-transform algorithm perform a matrix multiplication by the matrix Ω given by

$$\Omega = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \phi^{1\cdot 1} & \phi^{1\cdot 2} & \dots & \phi^{1\cdot(N-1)} \\ 1 & \phi^{2\cdot 1} & \phi^{2\cdot 2} & \dots & \phi^{2\cdot(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \phi^{1\cdot(N-1)} & \phi^{(N-1)\cdot 2} & \dots & \phi^{(N-1)\cdot(N-1)} \end{bmatrix}$$

Component wise, it holds

$$\Omega_{mn} = e^{-i2\pi\frac{n\,m}{N}} \qquad \text{with} \qquad n, m \in \{0, \dots, N-1\}$$

It can be shown that the matrix Ω is equal to the matrix $F/\Delta x$ shifted circularly toward the top by $N/2$ position, and toward the left also by $N/2$ positions. These are the two circular shifts needed to align the raw and column of 1's.

We will call Σ the circular shift toward top by $N/2$ positions of a vector of N entries. Since Σ is a permutation of the entries of a vector, it is a linear map and it can be written as a matrix. Writing $\vec{f}$ as

$$\vec{f} = \begin{bmatrix} f_{-N/2} \\ \vdots \\ f_{-1} \\ f_0 \\ \vdots \\ f_{N/2-1} \end{bmatrix}$$

it holds then

$$\Sigma \, \overrightarrow{f} = \begin{bmatrix} f_0 \\ \vdots \\ f_{N/2-1} \\ f_{-N/2} \\ \vdots \\ f_{-1} \end{bmatrix}$$

As a permutation, Σ has an inverse map which we will write Σ^{-1} and which is also a linear map with its own matrix. It is the circular shift of $N/2$ entries toward the bottom. It holds

$$\Sigma^{-1} \begin{bmatrix} f_0 \\ \vdots \\ f_{N/2-1} \\ f_{-N/2} \\ \vdots \\ f_{-1} \end{bmatrix} = \begin{bmatrix} f_{-N/2} \\ \vdots \\ f_{-1} \\ f_0 \\ \vdots \\ f_{N/2-1} \end{bmatrix} = \overrightarrow{f}$$

Without demonstration, we give the equality

$$\Omega = \Sigma \cdot (F/\Delta x) \cdot \Sigma^{-1}$$

which is equivalent to

$$F = \Delta x \, \Sigma^{-1} \Omega \, \Sigma$$

It follows

$$F\vec{f} = \Delta x \, \Sigma^{-1} \Omega \, \Sigma \, \vec{f}$$

This gives the receipt how to implement the DFT:

- Apply the circular shift to $\vec{f}$ (i.e. multiplication by Σ),
- Perform an FFT (i.e. multiplication by Ω),

- Apply the inverse circular shift (i.e. multiplication by Σ^{-1}),
- Don't forget the scaling factor Δx.

In Matlab for example, for a column vector, it reads something like

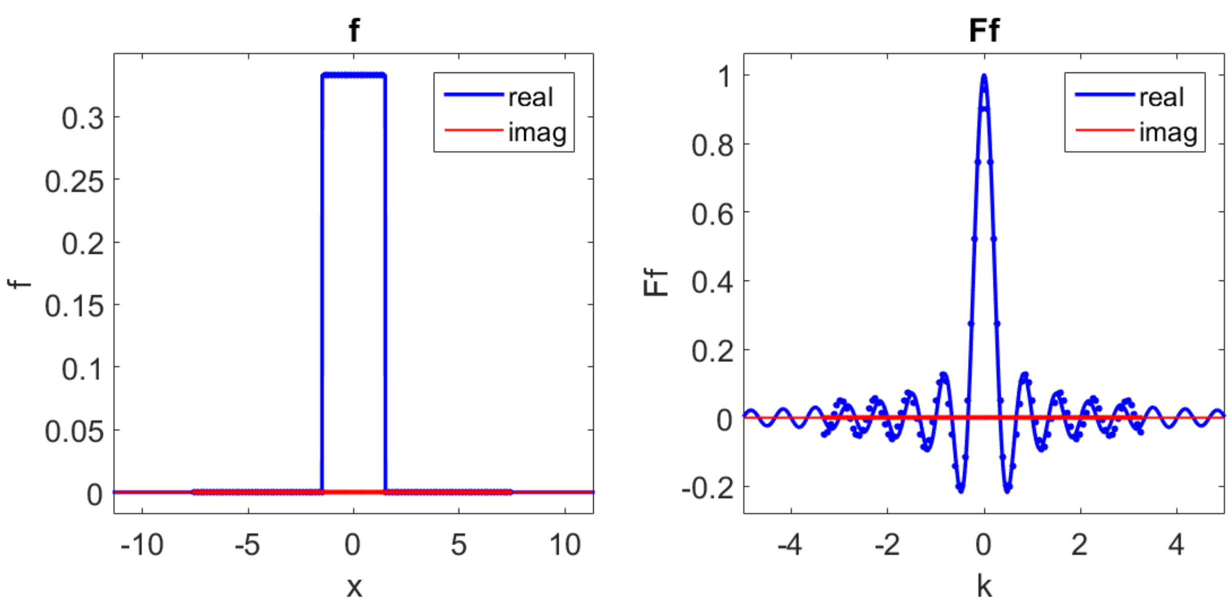
```
Ff = delta_x*fftshift(fft(ifftshift(f, 1), [], 1), 1);
```

But if the vector is a raw vector, don't forget to apply the transforms in the second dimension:

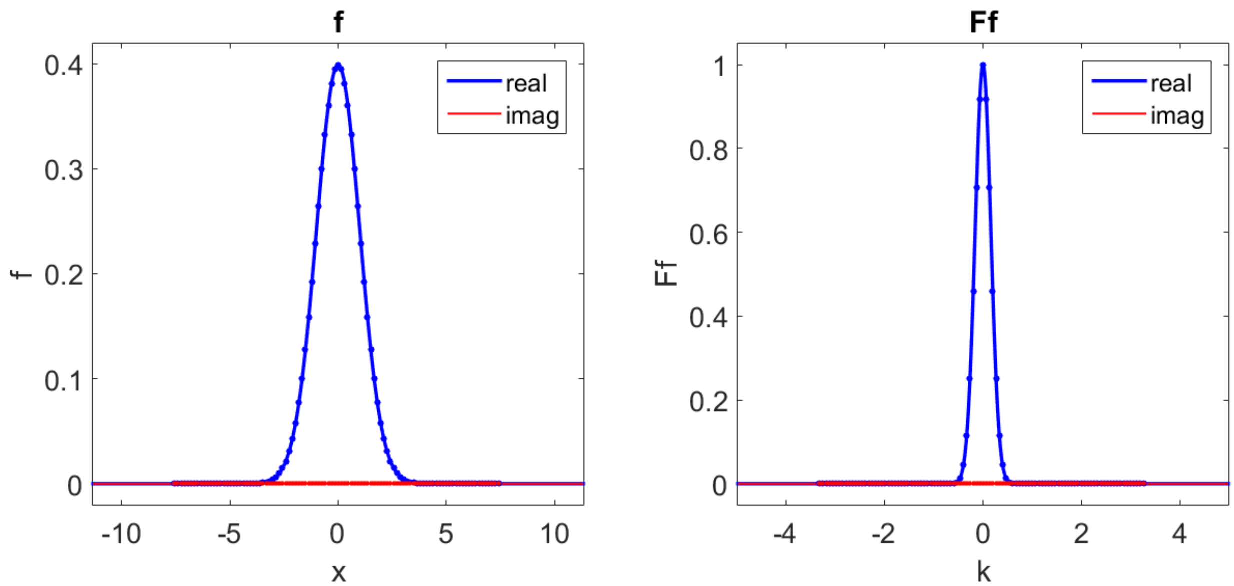
```
Ff = delta_x*fftshift(fft(ifftshift(f, 2), [], 2), 2);
```

The function `ifftshift` is in fact the Matlab implementation of the multiplication by Σ and the function `fftshift` is the one for Σ^{-1}. The Matlab implementation of the FFT is the function `fft`.

The following figure displays on the left the graph of the little-door function together with its sampling on the standard x-grid. On the right is its Fourier transform displayed together if the DFT of the function sampled on the left.



The following figure displays the same for the Gaussian function.



We now justify our original choice for the definition of the DFT. We start from the Fourier integral. Then we assume that the integrand in small outside the interval I_x of length L centered in 0 and we approximate the integral with finite differences:

$$\mathcal{F}f(k) = \int_{-\infty}^{+\infty} f(x)e^{-i2\pi\, kx}\, dx \approx \int_{-\frac{L}{2}}^{+\frac{L}{2}} f(x)e^{-i2\pi\, kx} \approx \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f(x_n)\, e^{-i2\pi\, k\, x_n}\, \Delta x$$

$$= \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f_n\, e^{-i2\pi\, k\, x_n}\, \Delta x$$

We now evaluate this approximation in $k = k_m$:

$$\mathcal{F}f(k_m) \approx \sum_{n=-N/2}^{N/2-1} f_n\, e^{-i2\pi\, k_m\, x_n}\, \Delta x = \sum_{n=-N/2}^{N/2-1} \Delta x\, e^{-i2\pi\, k_m\, x_n}\, f_n = \left(F\vec{f}\right)_m$$

Our sampling grids and our definitions are chosen so that $F\vec{f}$ is an approximation of $\mathcal{F}f(\cdot)$ on the standard k-grid. This can be written compactly as

$$\overrightarrow{\mathcal{F}f} \approx F\vec{f}$$

The factor Δx also makes sense in the definition of F from the view point of units because the unit of $\mathcal{F}f(\cdot)$ is the unit of $f(\cdot)$ multiplied by the unit of x, as given by the Fourier integral.

Now if you ask how good the approximation is, the answer is given by a convolution product that we define in the next part of the course. But we can already show how this answer looks like. We use

$$f_n = f(x_n) = \int_{-\infty}^{+\infty} \mathcal{F}f(k)\, e^{i2\pi\, k x_n}\, dk$$

Then

$$\left(F\vec{f}\right)_m = \sum_{n=-N/2}^{N/2-1} \Delta x\, e^{-i2\pi\, k_m\, x_n}\, f_n = \sum_{n=-N/2}^{N/2-1} \Delta x\, e^{-i2\pi\, k_m\, x_n} \int_{-\infty}^{+\infty} \mathcal{F}f(k)\, e^{i2\pi\, k x_n}\, dk$$

$$= \int_{-\infty}^{+\infty} dk\, \mathcal{F}f(k)\, \Delta x \sum_{n=-N/2}^{N/2-1} e^{-i2\pi\, k_m\, x_n}\, e^{i2\pi\, k x_n} = \int_{-\infty}^{+\infty} dk\, \mathcal{F}f(k)\, \Delta x \sum_{n=-N/2}^{N/2-1} e^{-i2\pi\, (k_m - k)\, x_n}$$

$$= \int_{-\infty}^{+\infty} \mathcal{F}f(k)\, \mathcal{A}_{\Delta x}^{N/2}(k_m - k)\; dk$$

We have shown:

$$\left(F\vec{f}\right)_m = \int_{-\infty}^{+\infty} \mathcal{F}f(k)\, \mathcal{A}_{\Delta x}^{N/2}(k_m - k)\; dk$$

The right hand side is a convolution with $\mathcal{A}_{\Delta x}^{N/2}(\cdot)$, the asymmetric Dirichlet kernel of period W. We will come back to that more in details in part III.

We finally describe the inverse discrete Fourier transform (inverse DFT). It is formally given by the inverse matrix of F i.e. the matrix F^{-1}. Entry wise, it is given by

$$(F^{-1})_{mn} \coloneqq \Delta k\, e^{i2\pi\, k_m x_n}, \qquad \text{with} \qquad n, m \in \{-N/2, \dots, N/2-1\}$$

or alternatively

$$(F^{-1})_{mn} \coloneqq \Delta k\, e^{i2\pi \frac{n\,m}{N}}, \qquad \text{with} \qquad n, m \in \{-N/2, \dots, N/2-1\}$$

We leave as an exercise to show that

$$F\,F^{-1} = F^{-1}\,F = id$$

In order the achieve it, the important following identity is useful:

$$\sum_{n=-N/2}^{N/2-1} e^{-i2\pi\frac{n}{N}(m'-m)} = \begin{Bmatrix} 0 & : & (m'-m) \notin N\cdot\mathbb{Z} \\ N & : & (m'-m) \in N\cdot\mathbb{Z} \end{Bmatrix}$$

where $N \cdot \mathbb{Z}$ is the set of integer multiples of N. Note that the sum is asymmetric. A symmetric sum would not lead to the expression on the right-hand side. It seems that asymmetry has its place in this world.

In the following, we will use the superscript $\cdot^*$ (the "star" symbol) to note the complex conjugate-transpose of a vector or a matrix. Note that this notion is different from the "adjoint" that we will write later with super script $\cdot^\dagger$ (the "dagger" symbol).

It follows from the definitions that

$$(F/\Delta x)^* = F^{-1}/\Delta k$$

or equivalently

$$F^* = F^{-1}\frac{\Delta x}{\Delta k}$$

We define the inner product $(\cdot\mid\cdot)_X$ as

$$\left(\vec{f}_1|\vec{f}_2\right)_X := \vec{f}_1^{\,*} H_X \vec{f}_2$$

for any pair of sampling $\vec{f}_1$ and $\vec{f}_2$ of some functions defined on the x-space and where

$$H_X = \Delta x\, id$$

is hemitian positive definite. Similarly, we define the inner product $(\cdot\mid\cdot)_K$ as

$$(\vec{g}_1|\vec{g}_2)_K := \vec{g}_1^{\,*} H_K \vec{g}_2$$

for any pair of sampling $\vec{g}_1$ and $\vec{g}_2$ of some functions defined on the k-space and where

$$H_K = \Delta k \; id$$

is hemitian positive definite too. With these two inner products at hand, the DFT F and the inverse DFT F^{-1} become isomorphisms of inner-product-spaces. The adjoint of F is then given by

$$F^{\dagger} = H_X^{-1} F^* H_K = \frac{1}{\Delta x} F^{-1} \frac{\Delta x}{\Delta k} \Delta k = F^{-1}$$

The DFT is therefore unitary:

$$F^{\dagger} = F^{-1}$$

It is now easy to verify

$$\left(\vec{f}_1 \middle| \vec{f}_2\right)_X = \left(F^{-1} F \vec{f}_1 \middle| \vec{f}_2\right)_X = \left(F^{\dagger} F \vec{f}_1 \middle| \vec{f}_2\right)_X = \left(F \vec{f}_1 \middle| F \vec{f}_2\right)_K$$

and

$$(\vec{g}_1 | \vec{g}_2)_K = (F\, F^{-1} \vec{g}_1 | \vec{g}_2)_K = \left(F^{-1} \vec{g}_1 \middle| F^{\dagger} \vec{g}_2\right)_X = (F^{-1} \vec{g}_1 | F^{-1} \vec{g}_2)_X$$

which proves the Parseval-Plancherel identity for the discrete Fourier transform. The 2-norms defined by

$$\|\vec{f}\|_{X,2}^2 \coloneqq \left(\vec{f} \middle| \vec{f}\right)_X$$

$$\|\vec{g}\|_{K,2}^2 \coloneqq (\vec{g} | \vec{g}\,)_K$$

leads to

$$\|\vec{f}\|_{X,2} = \|F\vec{f}\|_{K,2}$$

$$\|\vec{g}\|_{K,2} = \|F^{-1}\vec{g}\|_{X,2}$$

The DFT and its inverse are isometries of inner-product vector spaces.

We finally give the relation between F^{-1} and Ω^{-1}. Matrix Ω^{-1} is formally the inverse matrix of Ω . It is given component wise by

$$(\Omega^{-1})_{mn} = \frac{1}{N} e^{i 2\pi \frac{n\, m}{N}} \qquad \text{with} \qquad n, m \in \{0, \dots, N-1\}$$

It follows from those definitions that

$$\Omega^{-1} = \frac{1}{N}\,\Omega^*$$

We also note that, as a permutation, the matrix Σ verifies:

$$\Sigma^* = \Sigma^{-1}$$

We deduce

$$F^{-1} = F^* \frac{\Delta k}{\Delta x} = (\Delta x\,\Sigma^{-1}\Omega\,\Sigma)^* \frac{\Delta k}{\Delta x} = \Delta k\,\Sigma^{-1}\Omega^*\,\Sigma = \Delta k\,\Sigma^{-1}N\,\Omega^{-1}\,\Sigma = \Delta k N\;\Sigma^{-1}\,\Omega^{-1}\,\Sigma$$

We have shown

$$F^{-1} = \Delta k N\;\Sigma^{-1}\,\Omega^{-1}\,\Sigma$$

It follows

$$F^{-1}\vec{g} = \Delta k N\,\Sigma^{-1}\Omega^{-1}\,\Sigma\,\vec{g}$$

This gives the receipt how to implement the inverse DFT:

- Apply the circular shift to $\vec{g}$ (i.e. multiplication by Σ),
- Perform an inverse FFT (i.e. multiplication by Ω^{-1}),
- Apply the inverse circular shift (i.e. multiplication by Σ^{-1}),
- Don't forget the scaling factor $\Delta k N$.

In Matlab for example, for a column vector, it reads something like

```
iFg = delta_k*N*fftshift(ifft(ifftshift(g, 1), [], 1), 1);
```

If the vector is a raw vector, apply the transforms in the second dimension:

```
iFf = delta_k*N*fftshift(ifft(ifftshift(g, 2), [], 2), 2);
```

The function `ifft` is the Matlab implementation of the multiplication by Ω^{-1}.

Unlike the DFT, the inverse DFT has a nice interpretation in term of an asymmetric Fourier serie. It holds by definition

$$(F^{-1}\vec{g})_n = \sum_{m=-N/2}^{N/2-1} \Delta k\; e^{i2\pi\,k_m\,x_n}\; g_m = \left[\sum_{m=-N/2}^{N/2-1} \mathrm{c}_m\; e^{i2\pi\,k_m\,x}\right]_{x=x_n}$$

where

$$c_m = \Delta k \; g_m$$

In fact is $F^{-1}\vec{g}$ equal to an asymmetric Fourier serie evaluated on the standard x-grid.

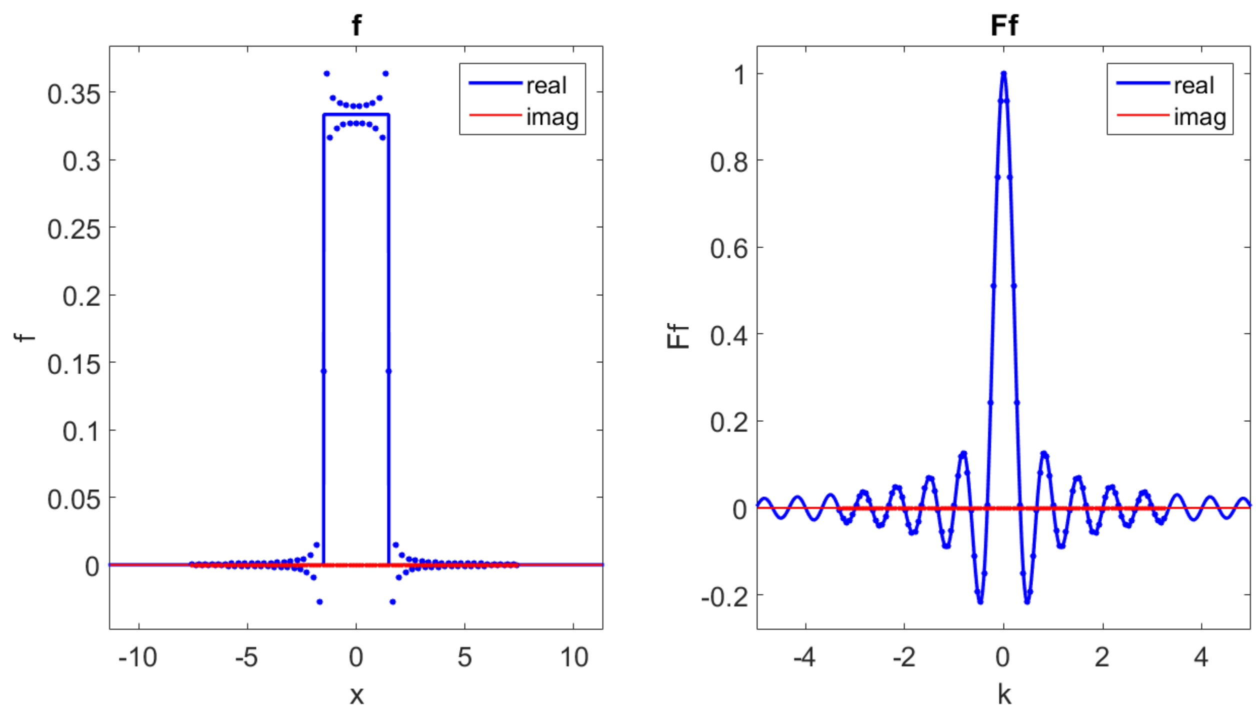
Similarly to the DFT, it can be shown that holds the approximation

$$\overrightarrow{\mathcal{F}^{-1}g} \approx F^{-1}\vec{g}$$

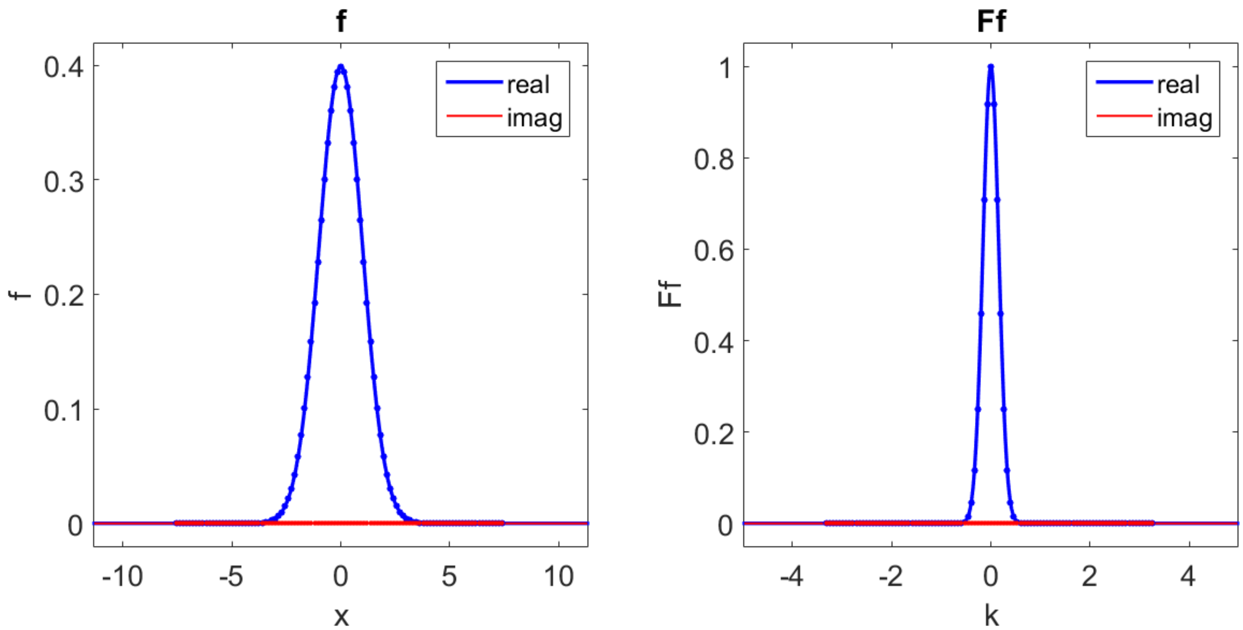
while the exact relationship between the inverse DFT and the inverse Fourier transform is

$$(F^{-1}\vec{g})_n = \int_{-\infty}^{+\infty} \mathcal{F}^{-1}g(x) \, \mathcal{A}_{\Delta k}^{N/2}(x_n - x) \; dx$$

The following figure displays on the right the graph of a $sinc(\cdot)$ function together with its sampling on the standard k-grid. On the left is its Fourier transform displayed together if the inverse DFT of the function sampled on the right.



The following figure displays the same for a Gaussian function.



Part III: Convolutions

Introduction to the convolution product

Given two functions $h(\cdot)$ and $f(\cdot)$, their "**convolution product**" $\{h * f\}(\cdot)$, or just "**convolution**" for short, is a new function $\{h * f\}(\cdot)$ defined by

$$\{h * f\}(x) \coloneqq \int_{-\infty}^{+\infty} h(x')\, f(x - x')\, dx'$$

In order to build an intuition about $\{h * f\}(\cdot)$, we set an example where $h(\cdot)$ is non-negative, is 0 outside a interval of length U centered in 0, and is normalized in the sense that

$$\int_{-\infty}^{+\infty} h(x)\ \ dx = 1$$

We choose then an integer $N > 0$ and we define

$$\Delta x \coloneqq U/N$$

We also define the weights

$$h_k \coloneqq \Delta x \cdot h(k \cdot \Delta x), \qquad k \in \{-N/2, \dots, N/2 - 1\}$$

The weights $h_{-N/2}, \dots, h_{N/2-1}$ are approximately normalized because

$$h_{-N/2} + \cdots + h_{N/2-1} = \sum_{k=-N/2}^{N/2-1} \Delta x \cdot h(k \cdot \Delta x) \approx \int_{-\infty}^{+\infty} h(x)\, dx = 1$$

The convolution $\{h * f\}(\cdot)$ evaluated in x can then be approximated as

$$\{h * f\}(x) = \int_{-\infty}^{+\infty} h(x')\, f(x - x')\, dx'$$

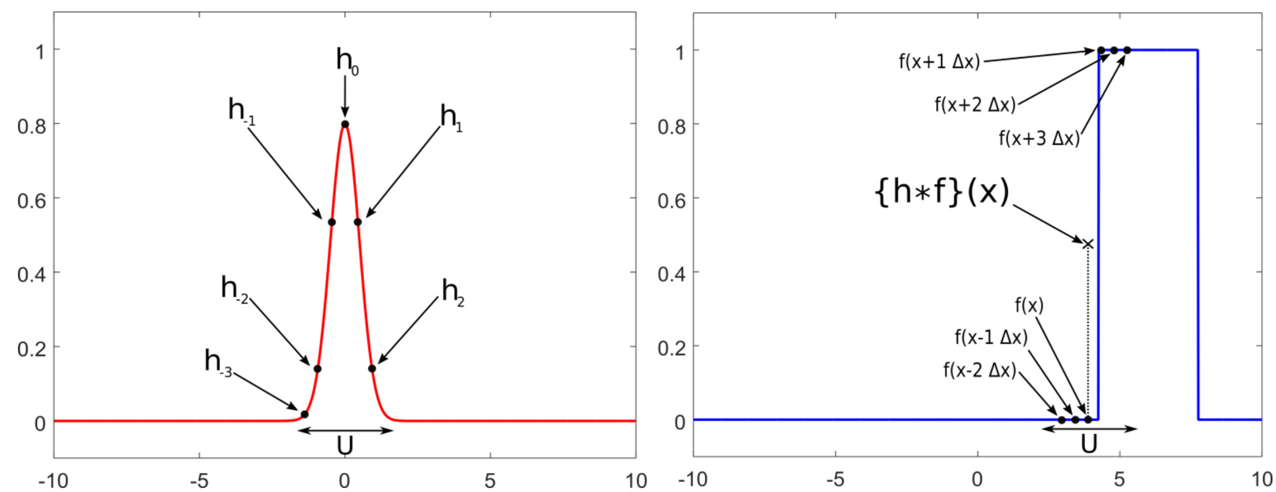
$$= \int_{-U/2}^{+U/2} h(x')\, f(x - x')\, dx'$$

$$\approx \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \Delta x \cdot h(k \cdot \Delta x)\ f(x - \ k \cdot \Delta x)$$

$$= h_{-N/2}\, f(x + N/2\ \ \Delta x) + \cdots + \ h_{N/2-1}\, f(x - (N/2 - 1)\, \Delta x)$$

This is a weighted average of the values of $f(\cdot)$ taken in an interval of length U centered in x, where the weights are given by $h(\cdot)$.

The next figure presents an example where $N = 6$, function $h(\cdot)$ is a Gaussian function and function $f(\cdot)$ is a door function.

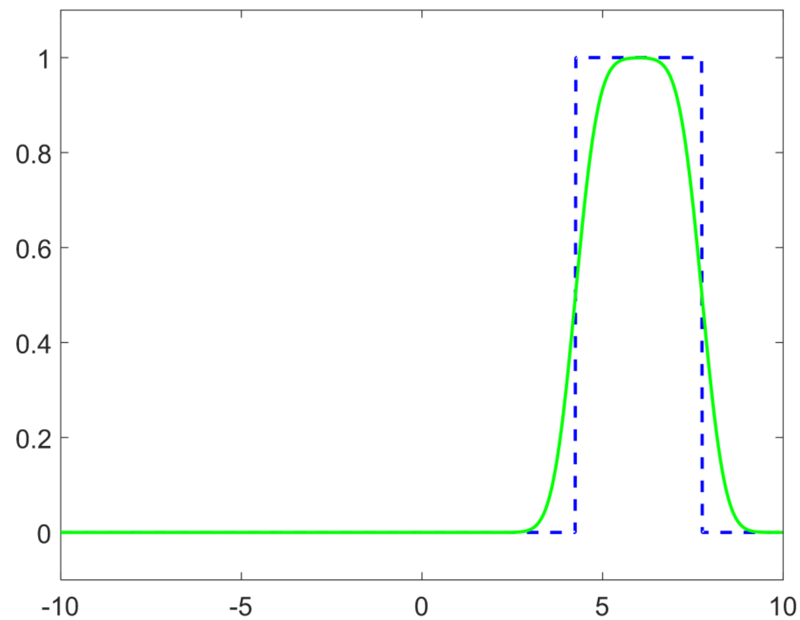


The convolution $\{h * f\}(\cdot)$ evaluated in x is approximately given by

$$h_{-3}\, f(x + 3\,\Delta x) + \cdots + \ h_2\, f(x - 2\,\Delta x)$$

It is a weighted average of the values of $f(\cdot)$ taken in an interval of size U centered in x, where the weights are given by $h(\cdot)$.

The next figure presents the graph of $\{h * f\}(\cdot)$ together with the graph of $f(\cdot)$ as a dashed line.



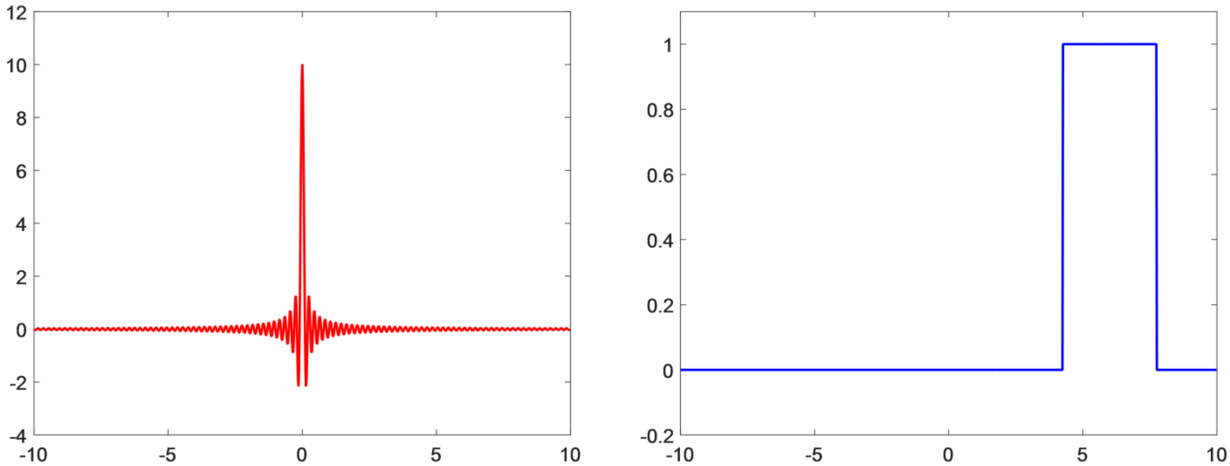
As another example, we chose $h(\cdot)$ to be a $sinc(\cdot)$ function (sinus cardinal) defined as follows:

$$h : x \longmapsto sinc(10 * x)/A$$

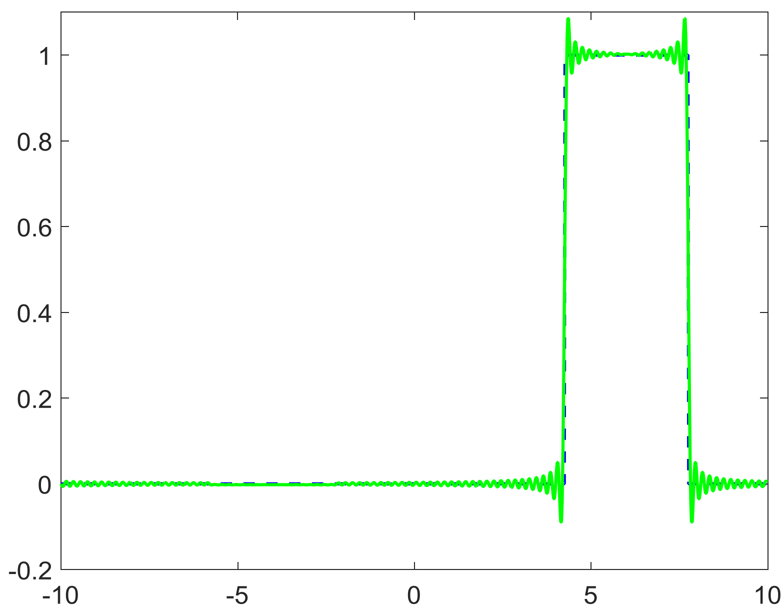
where A is given by

$$A := \int_{-\infty}^{+\infty} sinc(10 * x)\, dx$$

so that $h(\cdot)$ is normalized. We choose $f(\cdot)$ to be the same door function as previously. The next figure shows the graph of $h(\cdot)$ in red and the graph of $f(\cdot)$ on the right in blue.



The graph of $\{h * f\}(\cdot)$ is displayed in the next figure in green while the graph of $f(\cdot)$ is displayed as a dashed line.



In this example is $h(\cdot)$not zero outside an interval, and is neither non-negative. But $h(\cdot)$ is normalized and it still holds

$$\{h * f\}(x) \approx \sum_{k \in \mathbb{Z}} h_k \ f(x - \ k \cdot \Delta x)$$

The convolution $\{h * f\}(\cdot)$ evaluated in x is a weighted average of the values of function $f(\cdot)$ where the weights are given by function $h(\cdot)$.

The convolution of $h(\cdot)$ and $f(\cdot)$ acts like a smoothing of function $f(\cdot)$ and that the shape of this smoothing is given by the shape of $h(\cdot)$.

We now summarize a few rules that the convolution follows. The convolution is **symmetric**:

$$\{h * f\}(\cdot) = \{f * h\}(\cdot)$$

In fact it holds

$$\{h * f\}(x) = \int_{-\infty}^{+\infty} h(x')\, f(x - x')\, dx'$$

$$= \int_{+\infty}^{-\infty} h(x - z)\, f(z)\, (-dz)$$

$$= \int_{-\infty}^{+\infty} h(x - x')\, f(x')\, dx' \; = \; \{f * h\}(x)$$

where the substitution

$$z \coloneqq x - x';\; -dz = dx'$$

was used in the second line.

We omit the “dote-bracket” notation from now on in order to simplify the notation. That mean that we will write any function $f(\cdot)$ just with the symbol f. The convolution $\{h * f\}(\cdot)$ will be written $h * f$.

Another property of the convolution is

$$h * S_a f = S_a h * f$$

We leave the proof as an exercise. The convolution is **linear** in both arguments:

$$h * (\alpha f + \beta g) = \alpha\,(h * f) + \beta\,(h * g)$$

and

$$(\alpha h + \beta g) * f = \alpha\,(h * f) + \beta\,(g * f)$$

By the previous two rules follows that

$$[S_{-a}h + h + S_a h] * f = S_{-a}h * f + h * f + S_a h * f = h * [S_{-a}f + f + S_a f]$$

We conclude

$$P_L^S h * f = \sum_{z \in \mathbb{Z}} S_{zL} h * f = \sum_{z \in \mathbb{Z}} h * S_{zL} f = h * P_L^S f$$

If h is periodic holds

$$h = P_L^E h = P_L^S \Pi_L h$$

It follows

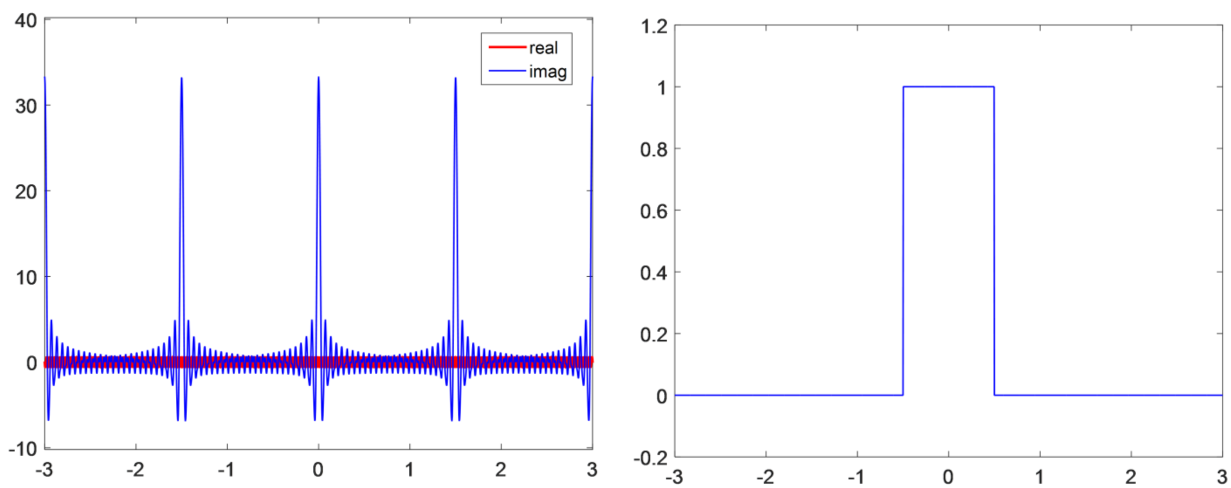
$$h * f = P_L^S \Pi_L h * f = \Pi_L h * P_L^S f$$

The convolution with an L-periodic convolution kernel h verifies

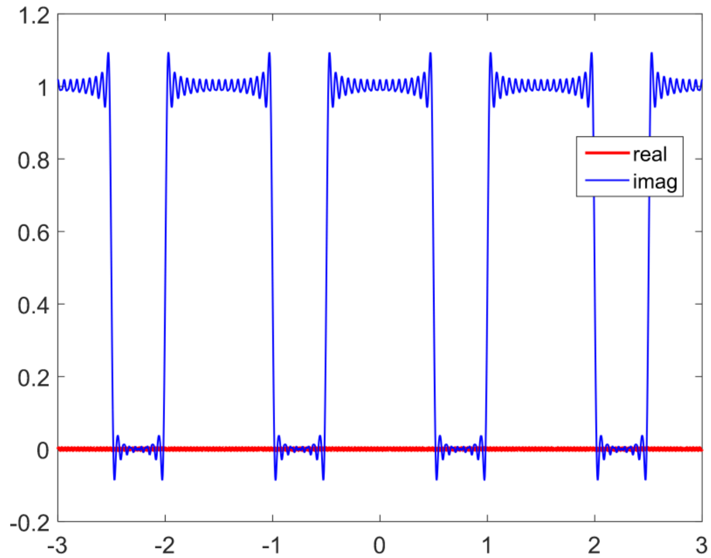
$$h * f = \Pi_L h * P_L^S f$$

It is the periodic summation of f convoluted with the "a single period of h", meaning the crope of h of width L.

Finally we display an example where h is an asymmetric Dirichlet (left on the next figure) kernel and where f is a door function centered in 0 (right on the next figure).



The graph of the convolution of both is displayed in the following figure.



This gives an intuitive explanation about what a convolution by a periodic function h does. We can know give an intuitive interpretation of the result of the DFT. We have shown in the previous chapter that

$$\left(F\vec{f}\right)_m = \int_{-\infty}^{+\infty} \mathcal{F}f(k)\, \mathcal{A}_{\Delta x}^{N/2}(k_m - k)\ dk$$

In the language of convolutions, it reads

$$\left(F\vec{f}\right)_m = \left\{\mathcal{A}_{\Delta x}^{N/2} * \mathcal{F}f\right\}(k_m)$$

Here is the convolution kernel

$$h(k) = \mathcal{A}_{\Delta x}^{N/2}(k)$$

which is W-periodic. We conclude that

$$\mathcal{A}_{\Delta x}^{N/2} * \mathcal{F}f = \Pi_W \mathcal{A}_{\Delta x}^{N/2} * P_W^S \mathcal{F}f$$

evaluated on the standard k-grid gives $F\vec{f}$. It is equal to the periodic summation of $\mathcal{F}f$ convoluted by the kernel $\Pi_W \mathcal{A}_{\Delta x}^{N/2}$ (which oscillate and is 0 outside the k-FoV W).

Similarly one can obtain

$$(F^{-1}\vec{g})_n = \left\{\mathcal{A}_{\Delta k}^{N/2} * \mathcal{F}^{-1}g\right\}(x_n)$$

The convolution kernel is in that case $\mathcal{A}_{\Delta k}^{N/2}$, which is L periodic. We conclude that

$$\mathcal{A}_{\Delta k}^{N/2} * \mathcal{F}^{-1}g = \Pi_L \mathcal{A}_{\Delta k}^{N/2} * P_L^S \mathcal{F}^{-1}g$$

evaluated on the standard x-grid gives $F^{-1}\vec{g}$. It is obtained by the periodic summation of $\mathcal{F}^{-1}g$ convoluted by the kernel $\Pi_L \mathcal{A}_{\Delta k}^{N/2}$ (which oscillate and is 0 outside the FoV L) and evaluated on the standard x-grid.

In the case one want to evaluate $\mathcal{F}^{-1}g$ by measuring $\vec{g}$ (which is what is done in MRI reconstruction), what we obtain by the inverse DFT is the periodic summation of $\mathcal{F}^{-1}g$ (which is responsible for fold-over artefacts) convoluted by $\Pi_L \mathcal{A}_{\Delta k}^{N/2}$ (which explains the oscillations that we call Gibbs artefacts).