AN INTRODUCTION TO FOURIER ANALYSIS AND SIGNAL PROCESSING

J. M. ALMIRA

Abstract. These notes are a proposal of material for an introductory course on Fourier Analysis with applications in Signals Theory.

Contents

Departamento de Matemáticas. Universidad de Jaén. E.U.P. Linares. 23700 Linares (Jaén), Spain. email: jmalmira@ujaen.es.

I would like to express my gratitude to Prof. B. Silbermann, Prof. P. Junghanns and Dr. U. Luther for their kind invitation to speak about this subject at their University.

1. Signals and systems: an introduction

1.1. Basic concepts.

Definition 1. *A signal is a physical quantity which depends on the time, the space, or any other variable or variables. In other words: a signal is nothing but a function of one or several variables. There are also random signals (i.e., random variables depending on one or several parameters), but we will not take care of them for these notes. We will also assume that our signals depend on a unique variable (time).*

The mathematical tool we will need for the study of signals is Mathematical Analysis. In particular, Fourier Analysis.

Definition 2. *We say that $x(t)$ is an analog signal if t can take its values from an interval of $\mathbb{R}$ with nonempty interior. We say that our signal is discrete if t take its values from $\mathbb{Z}$ or any other countable subset of $\mathbb{R}$.*

In all what follows we use the following standard notation: $\mathbf{x}(t)$, $\mathbf{y}(t)$, etc. will denote analog signals and $\mathbf{x}(n)$, $\mathbf{y}(n)$, etc. will denote discrete signals.

Signals serve as model of many physical processes. For example, the velocity of a particle is an analog signal and the results of an electrocardiogram are a discrete signal. In fact, signals are of interest to people from many different areas (such as mathematics, physics, engineering, etc) because of their many applications.

Definition 3. *Let $x(t)$ be a signal defined for all points of the real line. We call energy of the signal to the quantity given by*

$$\mathbf{E}(\mathbf{x}) = \int_{-\infty}^{\infty} |x(t)|^2 dt$$

(we will say that $\mathbf{x}$ *is an energy signal if* $\mathbf{E}(\mathbf{x}) < \infty$*) and we call mean power of the signal to the quantity given by*

$$\mathbf{P}(\mathbf{x}) = \lim_{M\to+\infty} \frac{1}{2M} \int_{-M}^{M} |x(t)|^2 dt$$

(we will say that $\mathbf{x}$ *is a power signal if* $\mathbf{P}(\mathbf{x}) < \infty$*).*

The above concepts are justified since in many situations where energy and power are meaningful, the defining formulas we have given, make the work. It is clear that there are analogous definitions for the discrete setting ($\mathbf{x} = \{\mathbf{x}(n)\}_{n\in\mathbb{Z}}$).

Definition 4. *The energy and the mean power of a discrete signal* $\mathbf{x} = \{\mathbf{x}(n)\}_{n\in\mathbb{Z}}$ *are given by the formulas:*

$$\mathbf{E}(\mathbf{x}) = \sum_{n=-\infty}^{+\infty} |\mathbf{x}(n)|^2 \text{ and } \mathbf{P}(\mathbf{x}) = \lim_{N\to+\infty} \frac{1}{2N+1} \sum_{n=-N}^{+N} |\mathbf{x}(n)|^2.$$

We say that $\mathbf{x} = \{\mathbf{x}(n)\}_{n\in\mathbb{Z}}$ *has finite energy if* $\mathbf{E}(\mathbf{x}) < \infty$ *and that has finite mean power if* $\mathbf{P}(\mathbf{x}) < \infty$.

Definition 5. *A system is an arbitrary process which transforms a certain class of (input) signals in another one (output). In more mathematical words, a system is an operator* $L : C_1 \to C_2$ *(*$x \to Lx$*), where* C_1 *and* C_2 *are two fixed classes of signals (e.g.,* $C_1 = C_2 =$ $\{$*analog energy signals* $\}$*).*

Some elementary properties of systems are the following:

Definition 6. *We say that* L *is time invariant if the identity*

$$L(\mathbf{x}(\cdot - s))(t) = (L\mathbf{x})(t - s)$$

holds for all $t, s \in \mathbb{R}$*. As it is natural, the corresponding definition for the discrete setting is:* $L(x(\cdot - k))(n) = (Lx)(n - k)$*;* $\forall n, k \in \mathbb{Z}$*.*

We say that the system L is causal if

$$\mathbf{x}_{|(-\infty,t]} = \mathbf{y}_{|(-\infty,t]} \Rightarrow (L\mathbf{x})(t) = (L\mathbf{y})(t), \text{ for all } t \in \mathbb{R}$$

Exercise: write down the corresponding definition for the discrete setting.

Definition 7. *We say that the system* L *has memory if the value* $(Lx)(t)$ *depends on the values that attain the signal* $x(t)$ *for some* $t' < t$ *(otherwise we say that the system has no memory).*

Definition 8. *We say that L is linear if $\mathbf{C}_1 = \mathbf{E}_1$ and $\mathbf{C}_2 = \mathbf{E}_2$ are vector spaces and $L : \mathbf{E}_1 \to \mathbf{E}_2$ is a linear operator. Moreover, if $\mathbf{E}_1$, $\mathbf{E}_2$ are normed spaces and L is a bounded operator, we say that the system is stable.*

We will pay special attention to the class of stable linear time invariant systems (LTI systems), since they are very useful in practice[1].

Definition 9. *The stable LTI systems are also called filters. A filter is called discrete if it transforms discrete signals into discrete signals. A filter is called continuous (or analog) if it sends analog signals into analog signals.*

1.2. **Some useful signals. Elementary transformations of signals.** Some elementary signals are specially useful for applications and, in fact, they are also very useful to represent other signals. Some of them are the following:

Example 10. *The* ***Heaviside's unit step signal*** *(analog) is given by*

$$u(t) = \begin{cases} 1 & t > 0 \\ 1/2 & t = 0 \\ 0 & t < 0 \end{cases}$$

and it can be used to define the ***sign signal,***

$$\mathbf{sgn}(t) = -1 + 2u(t) = \begin{cases} 1 & t > 0 \\ 0 & t = 0 \\ -1 & t < 0 \end{cases}$$

and the ***rectangular impulse***

$$\mathbf{rect}(t) = u(2t+1) - u(2t-1) = \begin{cases} 1 & t \in [-1,1] \\ 0 & t \notin [-1,1] \end{cases}$$

We can also use the unit step function to represent the ***ramp*** *signal*

$$\mathbf{r}(t) = \int_{-\infty}^{t} u(s)ds = \begin{cases} t & t > 0 \\ 0 & t \leq 0 \end{cases} .$$

[1]The discussion above lead us to conclude that, from a mathematical point of view, Signals Theory can be formulated in terms of Functional Analysis: We should focus our interest on the study of certain linear operators (usually, defined between two Hilbert spaces) which have some mathematical properties, like shift-invartiance, with an special meaning in Signals Theory. We also will be interested on the basis problem for these function spaces.

Example 11. *The **sampling signal** is given by*

$$\mathbf{sam}(t) := \begin{cases} \frac{\sin t}{t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

This signal will be very useful for the representation of a special type of analog signals (the so called band limited signals) in terms of their values on a discrete set of points (and that's the reason because we call 'sampling signal ' to $\mathbf{sam}(t)$*). The following is a graph of* $\mathbf{sam}(t)$*.*

Example 12. *Let us consider the signal* $p_\varepsilon(t) = \frac{1}{\varepsilon}\left(\frac{\sin(\pi t/\varepsilon)}{\pi t/\varepsilon}\right)^2$*. For the choice* $\varepsilon = 0.5$*, the graph of this signal is as follows:*

If we make $\varepsilon \to 0^+$*, then we obtain that*

$$\lim_{\varepsilon \to 0^+} p_\varepsilon(t) = \begin{cases} 0 & t \neq 0 \\ \infty & \end{cases}$$

On the other hand, for each $\varepsilon > 0$*, we have that*

$$\begin{aligned} \int_{-\infty}^{\infty} p_\varepsilon(t)dt &= \int_{-\infty}^{\infty} \frac{1}{\varepsilon}\left(\frac{\sin(\pi t/\varepsilon)}{\pi t/\varepsilon}\right)^2 dt \\ &= \pi \int_{-\infty}^{\infty} \left(\frac{\sin s}{s}\right)^2 ds = 1 \text{ (where we have made } s = \pi t/\varepsilon) \end{aligned}$$

This may motivate us to believe that there exists a certain 'limit function' $\delta(t) = \lim_{\varepsilon\to 0^+} p_\varepsilon(t)$ *such that* $\delta(t) = 0$ *for all* $t \neq 0$ *but, being* $\delta(0) = \infty$*, the relation*

$$\int_{-\infty}^{\infty} \delta(t)dt := \lim_{\varepsilon\to 0^+} \int_{-\infty}^{\infty} p_\varepsilon(t)dt = \lim_{\varepsilon\to 0^+} 1 = 1.$$

holds. We call the ***unit impulse*** *or* ***Dirac's delta*** *to this limit signal. It is clear that* $\delta(t)$ *is not a function in the usual sense of the word, since functions can only take finite values and, on the other hand, the integral of a function over the real line can not depends on its value at any concrete point. To give some meaning to this signal, what we make is thinking about it as a functional* $\delta : \mathbb{S} \to \mathbb{C}$ *which maps certain special functions* ϕ *(the Schwartz's functions* $\phi \in \mathbb{S}$*) onto their evaluation at zero. More precisely, we will have* $\delta(\phi) = \phi(0)$ *for all* $\phi \in \mathbb{S}$*. [see the next paragraph on generalized functions].*

The elementary transformations to which the tittle of this section refers are: time translation (also called shift) $x(t) \to y(t) = x(t-a)$, scaling $x(t) \to y(t) = x(at)$, and derivation $x(t) \to x'(t)$ of a signal. Of course, if the signals are in correspondence with functions in the ordinary sense of the word, these transformations are easy to handle. On the other hand, if we try to make these transformations to signals like the Dirac's delta, we must be careful with computations. To treat with this question, we must firstly introduce some formalism, and that's our main goal for the rest of the section.

1.2.1. *Generalized functions.* Ordinary functions map points to points (i.e., they map $\mathbb{R}$ into $\mathbb{R}$, $\mathbb{R}$ into $\mathbb{C}$ or $\mathbb{C}$ into $\mathbb{C}$). To introduce our new concept, we will treat as functions (in fact, they are functions) to certain functionals $L : \mathbb{S} \to \mathbb{C}$, where $\mathbb{S}$ denotes the space of fast decreasing functions (or Schwartz functions),

$$\mathbb{S} = \left\{ \phi \in \mathbf{C}^{(\infty)}(\mathbb{R}) : \lim_{t\to\pm\infty} t^n \phi^{(m)}(t) = 0,\, n, m \in \mathbb{N} \right\}.$$

We would like to note that $\mathbb{S}$ contains non trivial elements ($0 \in \mathbb{S}$ is clear), since $x(t) = e^{-t^2} \in \mathbb{S}$ (exercise!). Another interesting signal which belongs to the Schwartz class is given by$x(t) = v(t-a)v(b-t)$, where $a < b$ and

$$v(t) = \begin{cases} 0 & t \leq 0 \\ e^{-1/t} & t > 0. \end{cases}$$

(see the graphics: this is a $C^\infty(\mathbb{R})$ function with compact support). In fact, $\mathbb{S}$ is a vector subspace of $\mathbf{C}(\mathbb{R})$ (i.e., if $x(t), y(t) \in \mathbb{S}$ and $a, b \in \mathbb{C}$, then $ax(t)+by(t) \in \mathbb{S}$). What is more: the Schwartz's class is also closed with respect to many other operations like: scaling, time translation, multiplication by the powers t^n and the exponentials $e^{\mathbf{i}\theta t}$, derivation, etc.

We say that the signal $x(t)$ is slowly growing if the relation

$$\lim_{t\to\pm\infty}\frac{x(t)}{t^n}=0$$

holds for a certain choice of $n\in\mathbb{N}$ (this happens, for example, for all algebraic polynomials). We denote by **CSG** the class of continuous slowly growing signals.

The following is an easy-to-prove technical result:

Proposition 13. *Let us assume that $x(t)\in\mathbf{CSG}$ and $\phi(t)\in\mathbb{S}$. Then $h(t)=x(t)\phi(t)$ is bounded, continuous, absolutely integrable on $\mathbb{R}$ and vanishes at $\pm\infty$.*

Now, taking in consideration the result above, for each signal $x(t)\in\mathbf{CSG}$ we can define a functional $L_x:\mathbb{S}\to\mathbb{C}$, by

$$L_x(\phi)=\int_{-\infty}^{\infty}x(t)\phi(t)dt.$$

In a certain sense, this functional represents the signal x. Moreover, L_x has many interesting properties. For example, if $x(t)\in\mathbf{CSG}$ has a derivative then we may integrate by parts, to obtain that the relation

$$\begin{aligned}\int_{-\infty}^{\infty}x'(t)\phi(t)dt &= \lim_{\substack{a\to-\infty\\ b\to\infty}}\int_a^b x'(t)\phi(t)dt\\ &= \lim_{\substack{a\to-\infty\\ b\to\infty}}\left(x(t)\phi(t)]_a^b-\int_a^b x(t)\phi'(t)dt\right)\\ &= -\int_{-\infty}^{\infty}x(t)\phi'(t)dt\end{aligned}$$

holds for all $\phi\in\mathbb{S}$. This means that we can represent the derivative $x'(t)$ by using the functional $L_{x'}:\mathbb{S}\to\mathbb{C}$,

$$L_{x'}(\phi)=-\int_{-\infty}^{\infty}x(t)\phi'(t)dt.$$

But the second member of the identity above makes sense for all choice of $x(t)\in\mathbf{CSG}$ and $\phi(t)\in\mathbb{S}$, with no assumptions about existence of the derivative of $x(t)$. This motivate us to define a generalized derivative that exists for arbitrary signals $x(t)\in\mathbf{CSG}$ via the functional $L_{x'}(\phi)=-\int_{-\infty}^{\infty}x(t)\phi'(t)dt$. Of course, this functional will not represent an ordinary function in general, but it can be treated as a function- a generalized one!. The precise definitions run as follows:

Definition 14. *The **generalized derivative** of the signal $x(t) \in$ **CSG** is given by the functional*

$$x'\{\phi\} = L_{x'}(\phi) = -\int_{-\infty}^{\infty} x(t)\phi'(t)dt.$$

Definition 15. *The n-th generalized derivative of the signal $x(t) \in$ **CSG** is given by the functional:*

$$x^{(n)}\{\phi\} = (-1)^n \int_{-\infty}^{\infty} x(t)\phi^{(n)}(t)dt.$$

Definition 16. *We say that h is a **generalized function** if it is the n-th derivative of a certain signal $x(t) \in$ **CSG**. We denote by $\mathbb{G}$ the space of generalized functions.*

Example 17. ***(Heaviside's unit step signal and Dirac's delta signal, viewed as generalized functions)***. *Now we are able to define in a completely formal way the Dirac's delta signal, as a generalized function. We will also prove that the Heaviside's unit step signal is a generalized function. With this objective in mind, we note that the ramp signal*

$$\mathbf{r}(t) = \int_{-\infty}^{t} u(s)ds = \begin{cases} t & t > 0 \\ 0 & t \leq 0 \end{cases}$$

belongs to **CSG**, *so that it has a representation as a generalized function:*

$$\mathbf{r}\{\phi\} = \int_{-\infty}^{\infty} \mathbf{r}(t)\phi(t)dt = \int_{0}^{\infty} t\phi(t)dt.$$

Its generalized derivative is given by

$$\begin{aligned} \mathbf{r}'\{\phi\} &= -\int_{-\infty}^{\infty} \mathbf{r}(t)\phi'(t)dt = -\int_{0}^{\infty} t\phi'(t)dt \\ &= -\,t\phi(t)]_{-\infty}^{\infty} + \int_{0}^{\infty} \phi(t)dt \\ &= \int_{0}^{\infty} \phi(t)dt. \end{aligned}$$

This clearly means that $\mathbf{r}'\{\phi\} = \mathbf{u}\{\phi\}$ for all $\phi \in \mathbb{S}$, where $\mathbf{u}(t)$ denotes the Heaviside's unit step signal. Hence $\mathbf{u}$ is a generalized function.

If we take the second generalized derivative of the ramp function, we obtain that

$$\begin{aligned} \mathbf{r}''\{\phi\} &= \int_{-\infty}^{\infty} \mathbf{r}(t)\phi''(t)dt = \int_{0}^{\infty} t\phi''(t)dt \\ &= \int_{0}^{\infty} (t\phi' - \phi)'(t)dt = \lim_{t\to\infty} (t\phi' - \phi)(t) + \phi(0) \\ &= \phi(0) \end{aligned}$$

holds for all $\phi \in \mathbb{S}$ *(in the computations we have used that the relation*

$$(t\phi' - \phi)'(t) = \phi'(t) + t\phi''(t) - \phi'(t) = t\phi''(t)$$

holds for all $\phi \in \mathbb{S}$*). This means that the Dirac's delta function can be formally defined as a generalized function by the relation* $\delta = \mathbf{r}''$*, where* $\mathbf{r}$ *is the ramp signal and the derivative has been computed in the generalized sense.*

Example 18. ***(Derivative of Dirac's delta signal).*** *The representation* $\delta = \mathbf{r}''$*, can be used to compute the n-th derivative of* δ*, as follows:*

$$\begin{aligned}
\delta^{(n)}\{\phi\} &= \mathbf{r}^{(n+2)}\{\phi\} = (-1)^{n+2}\int_{-\infty}^{\infty} \mathbf{r}(t)\phi^{(n+2)}(t)dt \\
&= (-1)^n \int_0^{\infty} t\phi^{(n+2)}(t)dt \\
&= (-1)^n \int_0^{\infty} (t\phi^{(n+1)} - \phi^{(n)})'(t)dt \\
&= (-1)^n \left(\lim_{t\to\infty}(t\phi^{(n+1)} - \phi^{(n)})(t) + \phi^{(n)}(0)\right) \\
&= (-1)^n\phi^{(n)}(0),
\end{aligned}$$

for all $\phi \in \mathbb{S}$*. This, by abuse of notation, is usually written as:*

$$\int_{-\infty}^{\infty} \delta^{(n)}(t)\phi(t)dt = (-1)^n\phi^{(n)}(0), \text{ for all } \phi \in \mathbb{S}.$$

Example 19. ***(The Comb function III)****. An interesting example of generalized function is the so called comb function* **III***, which sends each signal* $\phi \in \mathbb{S}$ *to the number:*

$$\mathbf{III}\{\phi\} = \sum_{n=-\infty}^{\infty} \phi(n).$$

We can easily check that **III** *is a generalized function. To make this, we start by considering the integral part function* $[t]$*, defined by* $[t] = m$*, if* $t \in [m, m+1)$ *and* $m \in \mathbb{Z}$*, and we set*

$$q(t) = \int_0^t [s]ds \ (t \in \mathbb{R}).$$

Clearly, $q(t) \in \mathbf{CSG}$*, which means that* $q(t)$ *is also a generalized function,*

$$q\{\phi\} = \int_{-\infty}^{\infty} q(t)\phi(t)dt$$

Now we will prove that $\mathbf{III} = q''$, *where the derivative must to be considered in the generalized sense. By definition,*

$$\begin{aligned} q''\{\phi\} &= (-1)^2 \int_{-\infty}^{\infty} q(t)\phi''(t)dt \\ &= q(t)\phi'(t)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} q'(t)\phi'(t)dt \text{ (the ordinary derivative of } q, \text{ exists)} \\ &= -\int_{-\infty}^{\infty} q'(t)\phi'(t)dt \\ &= -\sum_{m=-\infty}^{\infty} m \int_{m}^{m+1} \phi'(t)dt \\ &= \sum_{m=-\infty}^{\infty} m\left\{\phi(m+1) - \phi(m)\right\} \\ &= \sum_{m=-\infty}^{\infty} \phi(m) + \sum_{m=-\infty}^{\infty} \{(m-1)\phi(m) - m\phi(m+1)\} \\ &= \sum_{m=-\infty}^{\infty} \phi(m) = \mathbf{III}\{\phi\}. \end{aligned}$$

1.2.2. *Elementary transformations of generalized functions.* In the section above we called 'space' to the class $\mathbb{G}$ of generalized functions but yet now, we have not proved that $\mathbb{G}$ is a vector space. Since the elements of $\mathbb{G}$ are functionals $x : \mathbb{S} \to \mathbb{C}$, the natural operations on $\mathbb{G}$ are defined by the formulas

$$(x+y)\{\phi\} = x\{\phi\} + y\{\phi\} \text{ and } (ax)\{\phi\} = a(x\{\phi\}),$$

respectively. Given $x, y \in \mathbb{G}$ and $a \in \mathbb{C}$, why can we claim that $x+y, ax \in \mathbb{G}$?. (exercise!).

There are other elementary transformations of signals that need a new definition for the setting of generalized functions. For example, this is the case for scaling and time translation[2]. The new concepts are fortunately not so dificult to obtain: they are usually based in an apropiate change of variable in the integral representation of the generalized function. For example, if we take in consideration the identities

$$\int_{-\infty}^{\infty} x(t+a)\phi(t)dt = \int_{-\infty}^{\infty} x(t)\phi(t-a)dt = x\{\phi(t-a)\}$$

[2]Generalized functions are very useful in Signal Theory. It is beacuse of this that all elementary operations of Fourier Analysis admit an interpretation for the class of generalized functions. Moreover, some of the functions we defined in the above section have special properties. For example, we will prove that the comb signal **III** is an eigenvector of the generalized Fourier transform $\mathcal{F}$.

and

$$\int_{-\infty}^{\infty} x(at)\phi(t)dt = \frac{1}{|a|}\int_{-\infty}^{\infty} x(s)\phi(s/a)ds = \frac{1}{|a|}x\{\phi(t/a)\} \text{ (for } a \neq 0)$$

then it will be natural to define translation in time and scaling of a generalized function x by the use of the formulas

$$x(\cdot + a)\{\phi\} = x\{\phi(t-a)\};\ x \in \mathbb{G},\ a \in \mathbb{R},\ \phi \in \mathbb{S}$$

and

$$x(\cdot a)\{\phi\} = \frac{1}{|a|}x\{\phi(t/a)\};\ x \in \mathbb{G},\ a \in \mathbb{R},\ \phi \in \mathbb{S}$$

Example 20. ***(Shift property of Dirac's delta signal).*** *The functional associated to time translation of the Dirac's delta signal, $\delta(\cdot - a)$ is given by:*

$$\delta(\cdot - a)\{\phi\} = \delta\{\phi(t+a)\} = \phi(a)$$

This is usually denoted (by abuse of notation) by

$$\int_{-\infty}^{\infty} \delta(t-a)\phi(t)dt = \phi(a),$$

and it is named the 'Shift property of Dirac's delta signal '.

Example 21. ***(Scaling of Dirac's delta signal).*** *For the scaling of Dirac's delta signal, $\delta(a\cdot)$ the following relation holds:*

$$\delta(\cdot a)\{\phi\} = \frac{1}{|a|}\delta\{\phi(t/a)\} = \frac{1}{|a|}\phi(0)$$

In intregal notation,

$$\int_{-\infty}^{\infty} \delta(ta)\phi(t)dt = \frac{1}{|a|}\phi(0).$$

1.3. **Filters and convolution.**

1.3.1. *The discrete case.* Given $\mathbf{x} = \{x(n)\}_{n=-\infty}^{+\infty}$ an arbitrary discrete signal, we can write

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k),\ n \in \mathbb{Z}$$

since the sum above only contains a finite number of terms (it only contains the term $k = n$, since $n - k \neq 0$ implies $\delta(n-k) = 0$). This means that the discrete signal $\mathbf{x}$ admits a representation

$$\mathbf{x} = \sum_{k=-\infty}^{\infty} x(k)\delta_k,$$

where $\delta_k = \{\delta(n-k)\}_{n=-\infty}^{+\infty}$, $k \in \mathbb{Z}$. What about the convergence of the representation above?. We have already proved the pointwise convergence of $\mathbf{x}_N = \sum_{k=-N}^{N} x(k)\delta_k$ to $\mathbf{x}$ (whenever $N \to \infty$) for arbitrary choices of the signal $\mathbf{x}$. If $\mathbf{x} \in l_p(\mathbb{Z})$ for a certain $p < \infty$, then

$$\begin{aligned} \left\| \mathbf{x} - \sum_{k=-N}^{N} x(k)\delta_k \right\|_{l_p(\mathbb{Z})} &= \|\{\cdots, x(-N-1), 0, \cdots, 0, x(N+1), \cdots\}\|_{l_p(\mathbb{Z})} \\ &= \left\{ \sum_{|k|>N} |x(k)|^p \right\}^{1/p} \to 0 \text{ for } N \to \infty. \end{aligned}$$

The same property holds if we change $l_p(\mathbb{Z})$ by $c_0(\mathbb{Z})$.

This property of discrete signals can be used to characterize the discrete filters as follows: Let $L : \mathbf{X} \to \mathbf{Y}$ be a discrete filter, where $\mathbf{X}$ and $\mathbf{Y}$ are supposed to be one of the spaces $l_p(\mathbb{Z})$ or $c_0(\mathbb{Z})$, and let

$$\{h(n)\}_{n=-\infty}^{+\infty} = L(\{\delta(n)\}_{n=-\infty}^{+\infty})$$

be the image of the unit impulse through our system L. Let $\mathbf{x} = \{x(n)\}_{n=-\infty}^{+\infty} \in \mathbf{X}$ be arbitrarely chosen. Then[3]

$$\begin{aligned} y(n) &= L(\mathbf{x})(n) = L\left(\mathbf{X}\text{-}\lim_{N\to\infty} \sum_{k=-N}^{N} x(k)\delta_k\right)(n) \\ &= \left(\mathbf{Y}\text{-}\lim_{N\to\infty} \sum_{k=-N}^{N} x(k)L(\delta_k)\right)(n) \\ &= \lim_{N\to\infty} \sum_{k=-N}^{N} x(k)L(\delta)(n-k) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \end{aligned}$$

This means that once we know $\{h(n)\}_{n=-\infty}^{+\infty}$, the output corresponding to the unit impulse, the output of the system $\mathbf{y} = L\mathbf{x}$ asociated to an arbitrary input $\mathbf{x} \in \mathbf{X}$, is given by

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k),\ n \in \mathbb{Z}.$$

Now we think we have enough motivation to introduce the discrete convolution concept:

[3]Recall that filters are continuous maps!

Definition 22. *Let* $\mathbf{x} \in l_\infty(\mathbb{Z})$, $\mathbf{y} \in l_1(\mathbb{Z})$ *be arbitrarely chosen. The convolution signal* $\mathbf{x} * \mathbf{y} \in l_\infty(\mathbb{Z})$ *is defined by the formula* $\mathbf{x} * \mathbf{y} = \{(x * y)(n)\}_{n=-\infty}^{\infty}$, *where*

$$(x * y)(n) = \sum_{k=-\infty}^{\infty} x(k)y(n-k),\ n \in \mathbb{Z}.$$

Theorem 23. *If* $L : l_\infty(\mathbb{Z}) \to l_1(\mathbb{Z})$ *is a discrete filter, the relation* $L\mathbf{x} = \mathbf{x} * L\delta$ *holds for all* $\mathbf{x} \in l_\infty(\mathbb{Z})$. *Moreover, if* $h \in l_1(\mathbb{Z})$ *is arbitrarely chosen then* $L\mathbf{x} := \mathbf{x} * h$ *is a filter* $L : l_\infty(\mathbb{Z}) \to l_1(\mathbb{Z})$ *with* $L\delta = h$. *The same conclusions hold true for discrete filters* $L : l_1(\mathbb{Z}) \to l_\infty(\mathbb{Z})$.

1.3.2. *The analog case.* In the analog case it is also possible to prove that filters are essentially those which appear by convolution. We will make no proof of this fact here, since it is a little bit technical. Instead of this, we will make a draft of proof, with heuristic arguments just to convince ourselves what's happening. Of course, a definition for convolution of analog signals is neccesary:

Definition 24. *Let* $\mathbf{x} \in L^\infty(\mathbb{R})$, $\mathbf{y} \in L^1(\mathbb{R})$ *be arbitrarely chosen. The convolution signal* $\mathbf{x} * \mathbf{y} \in L^\infty(\mathbb{R})$ *is defined by the formula*

$$(\mathbf{x} * \mathbf{y})(t) = \int_{-\infty}^{\infty} \mathbf{x}(s)\mathbf{y}(t-s)ds$$

Let us now assume that our signal $x(t)$ belongs to the Schwartz class $\mathbb{S}$, then we can use the shift property of Dirac's delta signal to claim that

$$x(t) = (x * \delta)(t) = \int_{-\infty}^{\infty} x(s)\delta(t-s)ds$$

Using the continuity of the filter L we can guarantee (this is the difficult step of the proof!) that it is no problem to interchange the actions of the integral operator and the filter L, to compute:

$$\begin{aligned} y(t) &= (Lx)(t) = L\left(\int_{-\infty}^{\infty} x(s)\delta(.-s)ds\right)(t) \\ &= \int_{-\infty}^{\infty} x(s)(L\delta(.-s))(t)ds \\ &= \int_{-\infty}^{\infty} x(s)(L\delta)(t-s)ds \\ &= (x * L\delta)(t) \end{aligned}$$

This may motivate us to think that analog filters are of the form $L\mathbf{x} = \mathbf{x} * L\delta$. In fact, this is (almost) the case, but we must be careful with the form in which the result is stated:

Theorem 25. *Let us assume that $L : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ is a filter. Then there exists a bounded variation signal $H \in \mathbf{BV}(\mathbb{R})$ such that the relation*

$$(L\mathbf{x})(t) = \int_{-\infty}^{\infty} \mathbf{x}(t-s)dH(s)$$

holds for all $\mathbf{x} \in L^1(\mathbb{R})$. Of course, if H is differentiable with derivarive $H' = h$, we obtain that

$$(L\mathbf{x})(t) = \int_{-\infty}^{\infty} \mathbf{x}(t-s)h(s)ds = \int_{-\infty}^{\infty} \mathbf{x}(s)h(t-s)ds = (\mathbf{x} * h)(t)$$

(See [5] for a proof).

Definition 26. *We say that L is a convolution filter if there exists a signal h such that $L\mathbf{x} = \mathbf{x} * h$ for all signal $\mathbf{x}$ which belongs to the domain of the filter.*

The result above claims that a very wide class of filters are convolution filters. Moreover, these filters have many nice properties. For example, the following proposition holds true:

Proposition 27. *The action of a convolution filter L on an arbitrary signal only depends on the output $h = L\delta$ associated to the unit impulse.*

If L is a convolution filter, it makes sense to study the properties of L in terms of the properties of $h = L\delta$. For example: under which conditions on $h = L\delta$ the filter is causal, has no memory, sends energy signals to energy signals, sends power signals to power signals, etc.? Moreover, this kind of problems are also of interest for general convolution operators $L_h(x) = x * h$ (i.e., to study these questions is (in principle) not neccesary restricting ourselves to continuous LTI systems). Of course, some results are easy to state (and to prove!):

Proposition 28. *The LTI system $L_h(x) = x * h$ is causal if and only if $h(t) = 0$ for all $t < 0$.*

Proposition 29. *The LTI system $L_h(x) = x * h$ is a filter $L_h : L^\infty(\mathbb{R}) \to L^\infty(\mathbb{R})$ if and only if $h(t) \in L^1(\mathbb{R})$.*

Exercise: Prove the two propositions above.

Remark 1. *Engineers use the connections we have stablished between filters and signals to translate to the context of signals some concepts which belong to the context of systems. For example, they say that an analog signal $h(t)$ is causal if $h(t) = 0$ for all $t < 0$.*

2. Fourier Series I

2.1. **The frecuency approach.** Everybody knows that polynomials are useful for the approximation of functions (hence signals). For example, in any introductory Calculus course, people learn to use Taylor series for the local approximation of functions, Newton's method (where functions are linearized) for the approximation of zeroes of functions, interpolatory quadrature formulae for the approximate computation of integrals, etc. With such a focus, many students have the impression that polynomials are the best tool for the approximation of functions. Of course, they also learn in other courses that other approximation processes are better for certain special purposes. For example, they learn something about spline approximation, rational approximation (usually Padé approximation), etc.

Now, from a certain point of view, we can claim that trigonometric approximation (i.e., approximation by trigonometric polynomials

$$\sum_{k=0}^{N}\left\{a_k\cos(kTt)+b_k\sin(kTt)\right\}, N\in\mathbb{N})$$

is more useful than polynomial approximation for the main objectives of Signals Theory. A key concept which is related to this claim is that of eigenvalue:

Definition 30. *Let $L:\mathbf{X}\to\mathbf{X}$ be a linear operator. We say that $\mathbf{x}\in\mathbf{X}$ ($\mathbf{x}\neq 0$) is an eigenvector which is associated to the eigenvalue $\lambda\in\mathbb{C}$ of L if $L\mathbf{x}=\lambda\mathbf{x}$.*

Eigenvectors of L are useful because the action of L on a linear combination of eigenvectors of L is very simple (the behaviour is exactly the same as the action of a diagonal matrix operating on a vector which is expressed in terms of the standard basis of $\mathbb{R}^n$): if $\mathbf{x}_1,...,\mathbf{x}_m$ are eigenvectors of L and the corresponding eigenvalues are given by $\lambda_1,...,\lambda_n$, then

$$L\left(\sum_{k=1}^{m}a_k\mathbf{x}_k\right)=\sum_{k=1}^{m}a_k\lambda_k\mathbf{x}_k.$$

Now: recall that a filter is a linear operator which maps certain (input) signals into another (output) signals. If we were able to represent the input signals as a linear combination of eigenvectors of the filter, then our work with the signals should be greatly simplyfied. So, why trigonometric approximation should help us to treat with signals?. The following theorem helps to understand what's happening:

Theorem 31. *The complex exponential signals $\exp(\mathbf{i}wt)=\cos(wt)+\mathbf{i}\sin(wt)$ are eigenvectors of any convolution filter.*

Proof. Let $L_h(\mathbf{x}) = \mathbf{x} * \mathbf{h}$ be a convolution filter. Then

$$\begin{aligned}
L_h(\exp(\mathbf{i}wt)) &= \exp(\mathbf{i}wt)*\mathbf{h} = \int_{-\infty}^{\infty} \exp(\mathbf{i}ws)\mathbf{h}(t-s)ds \\
&= -\int_{\infty}^{-\infty} \exp(\mathbf{i}w(t-\zeta))\mathbf{h}(\zeta)d\zeta \\
&= \left(\int_{-\infty}^{\infty} \exp(-\mathbf{i}w\zeta)\mathbf{h}(\zeta)d\zeta\right)\exp(\mathbf{i}wt)
\end{aligned}$$

So that, $\exp(\mathbf{i}wt)$ is an eigenvector of L_h associated to the eigenvalue

$$\lambda(w,\mathbf{h}) = \int_{-\infty}^{\infty} \exp(-\mathbf{i}w\zeta)\mathbf{h}(\zeta)d\zeta$$

and the proof ends. □

We will prove that, under very reasonable hyphotheses on a T-periodic analog signal $\mathbf{x}(t)$, it is possible to represent it as an infinite linear combination of exponentials,

$$\mathbf{x}(t) = \sum_{k=-\infty}^{\infty} c_k \exp(2\pi\mathbf{i}kt/T).$$

There are other important reasons because complex exponential functions are of interest in Signals Theory:

- The frequency response of a system is usually easy to meassure at a Lavoratory. This is not the case for the impulse response of the system.[4]
- The frequency components of a signal are usually possible to aproximate in practice.
- Several signals with different frequency components can share the same carrier whithout disturb to each other.

2.2. **Orthogonality: the concept.** Let us assume that $\mathbf{u} = (u_1, ..., u_n)$, $\mathbf{v} = (v_1, ..., v_n)$ are orthogonal vectors in the standard Euclidean space $\mathbb{R}^n$. Then, using Pitagoras Theorem, we obtain that

$$\begin{aligned}
\sum_{i=1}^{n}(u_i^2+v_i^2) &= ||\mathbf{u}||^2 + ||\mathbf{v}||^2 = ||\mathbf{u}+\mathbf{v}||^2 = \\
&= \sum_{i=1}^{n}(u_i+v_i)^2 = \sum_{i=1}^{n}(u_i^2+v_i^2) + 2\sum_{i=1}^{n}u_iv_i
\end{aligned}$$

[4]Given a convolution system, we can re-write the system at the frequency domain by taking the Fourier transform at both sides of the formula $y(=Lx) = x * h$, so that, $\widehat{y} = \widehat{x}\cdot\widehat{h}$ represents the system at the frequency domain. We call frequency response of the system to the signal $\widehat{h}$. (These concepts will be introduced later on in these notes).

Hence

$$\sum_{i=1}^{n} u_i v_i = 0.$$

Moreover, if $\sum_{i=1}^{n} u_i v_i = 0$ then $\mathbf{u} = (u_1, ..., u_n)$, $\mathbf{v} = (v_1, ..., v_n)$ are orthogonal vectors in $\mathbb{R}^n$. [Of course, for complex vectors $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ we should obtain that orthogonality means $\sum_{i=1}^{n} u_i \overline{v_i} = 0$]. We can use this characterization of orthogonality to generalize the concept to other settings. In particular, we are interested in an orthogonality concept for functions. To do this, we think about vectors $\mathbf{u} = (u_1, ..., u_n)$ as a sampling of a function $\mathbf{u}(t)$ ($t \in [a, b]$), where the sampling is made at the knots $\{t_i\}_{i=1}^{n}$ of some partition $a < t_1 < t_2 < \cdots < t_n < b$ (i.e., we set $u_i = \mathbf{u}(t_i)$, $i = 1, ..., n$). Furthermore, we will assume that the knots are equidistant: $t_i = a + hi$, $i = 1, ..., n$. With this interpretation at hand, we think about $h \sum_{i=1}^{n} u_i v_i$ as a Riemann sum associated to the integral $\int_a^b \mathbf{u}(t)\mathbf{v}(t)dt$, and define:

Definition 32. *The signals $\mathbf{u}, \mathbf{v}$ are said to be orthogonal in $[a, b]$ with respect to the weight function $\mathbf{w}(t)$ if*

$$\int_a^b \mathbf{u}(t)\overline{\mathbf{v}(t)}\mathbf{w}(t)dt = 0.$$

If the weight function is a nonzero constant, we just say that $\mathbf{u}, \mathbf{v}$ are orthogonal in $[a, b]$.

Definition 33. *We say that $\mathbf{W}$ is an orthogonal set in $[a, b]$ with respect to the weight function $\mathbf{w}(t)$ if for all $\mathbf{u}, \mathbf{v} \in \mathbf{W}$, $\mathbf{u} \neq \mathbf{v}$, we have that*

$$\int_a^b \mathbf{u}(t)\overline{\mathbf{v}(t)}\mathbf{w}(t)dt = 0.$$

If the weight function is a nonzero constant, we just say that $\mathbf{W}$ is an orthogonal set in $[a, b]$.

The exponential functions $\{\exp(\mathbf{i}nt)\}_{n=-\infty}^{+\infty}$ form an orthogonal system in $[-\pi, \pi]$. Another important example is given by the trigonometric system

$$\{1, \cos t, \cos 2t, ..., \sin t, \sin 2t, ...\}.$$

Indeed the following identities are satisfied:

$$\begin{aligned}
\int_{-\pi}^{\pi} \cos(nt)\cos(mt)dt &= \begin{cases} 0 & \text{if } n \neq m \\ \pi & \text{if } n = m \neq 0 \\ 2\pi & \text{if } n = m = 0 \end{cases} \\
\int_{-\pi}^{\pi} \sin(nt)\sin(mt)dt &= \begin{cases} \pi & \text{if } n = m \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
\int_{-\pi}^{\pi} \sin(nt)\cos(mt)dt &= 0 \text{ for all } n, m.
\end{aligned}$$

We check here the claim above by making the computation of

$$\int_{-\pi}^{\pi} \cos(nt)\cos(mt)dt$$

for all choices of n, m. (The other integrals imply analogous computations). Let us assume that $n \neq m$. Taking in consideration the well known trigonometric formula

$$\cos(nt)\cos(mt) = \frac{1}{2}(\cos((n+m)t) + \cos((m-n)t))$$

we compute a primitive of $\cos(nt)\cos(mt)$, which is given by:

$$\frac{1}{2}\left(\frac{\sin((n+m)t)}{n+m} + \frac{\sin((m-n)t)}{m-n}\right)$$

Hence

$$\int_{-\pi}^{\pi} \cos(nt)\cos(mt)dt = \frac{1}{2}\left(\frac{\sin((n+m)t)}{n+m} + \frac{\sin((m-n)t)}{m-n}\right)\Bigg]_{-\pi}^{\pi} = 0$$

If $n = m \neq 0$ we use the identity

$$\cos^2(nt) = \frac{1}{2}(1 + \cos(2nt))$$

to compute

$$\int_{-\pi}^{\pi} \cos^2(nt)dt = \frac{1}{2}\left(t + \frac{\sin(2nt)}{2n}\right)\Bigg]_{-\pi}^{\pi} = \pi.$$

Finally, if $n = m = 0$ then we are integrating 1 in the interval $[-\pi, \pi]$ so that the corresponding value is equal to 2π.

Definition 34. *We say that the orthogonal set* $\mathbf{W}$ *in the interval* $[a, b]$ *with respect to the weight function* $\mathbf{w}(t)$ *is complete if*

$$\int_a^b \mathbf{x}(t)\overline{\mathbf{v}(t)}\mathbf{w}(t)dt = 0 \textit{ for all } \mathbf{v} \in \mathbf{W} \textit{ implies that } \mathbf{x}(t) \equiv 0$$

2.3. **Formal Fourier Series.** We devote this section to the formal representation of analog signals $x(t)$, $t \in (-\pi, \pi]$ (equivalently, of 2π-periodic analog signals $x(t)$, $t \in \mathbb{R}$) as trigonometric series. We will use heuristic arguments to determine the coefficients that must appear in the expression

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos(kt) + b_k \sin(kt)) \tag{1}$$

(we multiply a_0 by $1/2$ because of certain technical reasons that will become clear in the near future). So, we want to know a formula for the computation of a_k and b_k in the expression above. To deduce such a formula, we will use that $\{1, \cos t, \cos 2t, ..., \sin t, \sin 2t, ...\}$ is an orthogonal system. We multiply both sides of (1) by $\cos(mt)$ (for a certain fixed m) and we integrate between $-\pi$ and π. If we assume good enough convergence of the series (e.g., if there exists uniform convergence in $[-\pi, \pi]$), [and that's what we will assume by the moment] then we will be able to interchange the sum with the integral, so that:

$$\begin{aligned}
\int_{-\pi}^{\pi} x(t)\cos(mt)dt &= \frac{a_0}{2}\int_{-\pi}^{\pi}\cos(mt)dt \\
&\quad + \sum_{k=1}^{\infty}\left(a_k \int_{-\pi}^{\pi}\cos(mt)\cos(kt)dt + b_k \int_{-\pi}^{\pi}\cos(mt)\sin(kt)dt\right) \\
&= \pi a_m.
\end{aligned}$$

Hence the relation

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos(mt)dt$$

holds for all $m = 0, 1, 2, ...$(Now is clear why we multiplied the coefficient a_0 by $1/2$: just to obtain the same formula for all a_m!). Analogously, multiplying both sides of (1) by $\sin(nt)$ (for a certain fixed n) and integrating between $-\pi$ and π, we obtain

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin(nt)dt$$

for all $n = 1, 2, ...$

Definition 35. *In all what follows, the expression*

$$x(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos(kt) + b_k \sin(kt))$$

means that the coefficients a_k, b_h are given by $a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos(kt)dt$ and $b_h = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin(ht)dt$ for all $k \in \mathbb{N}$ and all $h \in \mathbb{N}\setminus\{0\}$. Moreover, in such a case we say that

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k\cos(kt) + b_k\sin(kt))$$

is the (formal) Fourier series associated to the 2π-periodic analog signal $x(t)$, and that a_k, b_h are the corresponding Fourier coefficients.

Note that the definition above assumes nothing about convergence of the Fourier series. In fact, it is possible to obtain analog signals whose Fourier series is not summable.

Example 36. *Let $x(t)$ denote the 2π-periodic extension of the signal*

$$g(t) = \begin{cases} 1/2 & 0 \leq t < \pi \\ -1/2 & -\pi \leq t < 0 \end{cases},$$

We compute its Fourier series as follows: We use that $g(-t) = -g(t)$ and $\cos(kt) =$ cos$(-kt)$, to claim that the products $g(t)\cos(kt)$ $(k \geq 0)$ are odd functions, so that $a_k = 0$ for all k. The other Fourier coefficients are computed by hand:

$$\begin{aligned} b_k &= \frac{1}{\pi}\left(-\int_{-\pi}^{0}\frac{\sin(kt)}{2}dt + \int_0^{\pi}\frac{\sin(kt)}{2}dt\right) \\ &= \frac{1}{\pi}\left(\int_0^{\pi}\frac{\sin(kt)}{2}dt + \int_0^{\pi}\frac{\sin(kt)}{2}dt\right) \\ &= \frac{1}{\pi}\int_0^{\pi}\sin(kt)dt \\ &= \frac{1}{\pi}\left.\frac{\cos(kt)}{k}\right]_0^{\pi} \\ &= \frac{1}{\pi}\frac{1+(-1)^{k+1}}{k} \\ &= \begin{cases} \frac{2}{\pi k} & \textit{if } k \textit{ is odd} \\ 0 & \textit{otherwise} \end{cases} \end{aligned}$$

Hence we have

$$x(t) \sim \frac{2}{\pi}\sum_{k=0}^{\infty}\frac{1}{2k+1}\sin((2k+1)t)$$

2.3.1. *Odd and Even signals.* Odd and even signals are so frequent that their Fourier series merit special attention. Perhaps there exists a basic reason for this, and it is the fact that all analog signals $x(t)$ admit a decomposition as a sum of an even signal $x^+(t)$ and an odd signal $x^-(t)$. This is easy to check, since we can make:

$$x(t) = \underbrace{\frac{1}{2}(x(t)+x(-t))}_{x^+(t)} + \underbrace{\frac{1}{2}(x(t)+x(-t))}_{x^-(t)}$$

Even signals satisfy that $b_k = 0$ ($k \geq 0$) and

$$a_k = \frac{2}{\pi}\int_0^\pi x(t)\cos(kt)dt$$

for all $k \geq 1$. On the other hand, if $x(t)$ is odd, we have that $a_k = 0$ ($k \geq 1$) and

$$b_k = \frac{2}{\pi}\int_0^\pi x(t)\sin(kt)dt$$

for all $k \geq 0$. If we make use of certain symmetries with respect to $t = \pi/2$ we can also get signals whose Fourier series only contain coefficients indexed by even numbers (odd numbers, respectively). Let us see an example:

Example 37. *Let $x(t)$ denotes the (2π-periodic) sawtooth signal described by the figure:*

Since our signal is even, its Fourier series only contains cosines. Of course, $a_0 = \pi/4$. Moreover, when we integrate against $\cos(kt)$ we have that

$$\begin{aligned} a_k &= \frac{2}{\pi}\int_0^\pi x(t)\cos(kt)dt \\ &= \frac{2}{\pi}\int_0^{\pi/2} t\cos(kt)dt + \frac{2}{\pi}\int_{\pi/2}^\pi (\pi - t)\cos(kt)dt \end{aligned}$$

Now, making the change of variable $s = \pi - t$ for the second integral which appears at the righthand side of the identity above, we obtain

$$a_k = \frac{2}{\pi}\int_0^{\pi/2} t\cos(kt)dt + \frac{2}{\pi}\int_0^{\pi/2} s\cos(k(\pi - s))ds.$$

On the other hand, it is well known that

$$\cos(k(\pi - s)) = \cos k\pi \cos ks + \sin k\pi \sin ks = (-1)^k \cos(ks)$$

This implies that $a_k = 0$ *if* k *is odd and for* k *even,* $k = 2m$*, we obtain that*

$$a_{2m} = \frac{4}{\pi} \int_0^{\pi/2} t \cos(2mt) dt.$$

Integrating by parts,

$$\begin{aligned} a_{2m} &= \frac{4}{\pi} \left\{ t \frac{\sin(2mt)}{2m} \right]_0^{\pi/2} - \int_0^{\pi/2} \frac{\sin(2mt)}{2m} dt \right\} \\ &= -\frac{4}{\pi} \left\{ \frac{-\cos(2mt)}{(2m)^2} \right]_0^{\pi/2} \right\} = -\frac{4}{\pi} \left\{ \frac{(-1)^{m+1} + 1}{(2m)^2} \right\} \\ &= \frac{1}{\pi} \frac{(-1)^m - 1}{m^2} = \begin{cases} \frac{-2}{\pi(2h+1)^2} & m = 2h+1 \\ 0 & m = 2h \end{cases} \end{aligned}$$

It follows that for $m = 2k$*, the coefficient* $a_{2m} = a_{4k}$ *vanishes. Hence the unique coefficients which efectively appear in the formal Fourier series are* a_0 *and the* $a_{4k+2} = \frac{-2}{\pi(2k+1)^2}$*,* $k = 0, 1, \ldots$ *and the Fourier series of the sawtooth signal is*

$$x(t) \sim \pi/4 - 2 \sum_{h=0}^{\infty} \frac{1}{\pi(2h+1)^2} \cos((4h+2)t).$$

It is of interest to note that convergence is fast. For example, the figure shows the truncated Fourier series for $N = 22$*,*

$$\pi/4 - 2 \textstyle\sum_{h=0}^{5} \frac{1}{\pi(2h+1)^2} \cos((4h+2)t)$$

Of course, adding more symmetries, it is possible to cancel more and more terms in the formal Fourier series of a signal. On the other hand, just looking to the formal Fourier series of a signal, it is sometimes possible to take conclussions about its symmetries.

2.3.2. *The complex Fourier series.* There exists a strong relation between the complex exponential function $\exp(\mathbf{i}\theta)$ and the trigonometric functions: the well known Euler's formula: $e^{\mathbf{i}\theta} = \cos\theta + \mathbf{i}\sin\theta$. This suggest us the posibility of creating a complex version of the Fourier series. To start with this aim, we recall that $\sin(-\theta) = -\sin(\theta)$ and $\cos(-\theta) = \cos\theta$, so that Euler formula is equivalent to

$$\begin{aligned} \cos\theta &= \frac{1}{2}(e^{\mathbf{i}\theta} + e^{-\mathbf{i}\theta}) \\ \sin\theta &= \frac{1}{2\mathbf{i}}(e^{\mathbf{i}\theta} - e^{-\mathbf{i}\theta}) = \frac{\mathbf{i}}{2}(e^{-\mathbf{i}\theta} - e^{\mathbf{i}\theta}) \end{aligned}$$

This can be rephrased as:

$$\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ -\mathbf{i}/2 & \mathbf{i}/2 \end{pmatrix} \begin{pmatrix} e^{\mathbf{i}\theta} \\ e^{-\mathbf{i}\theta} \end{pmatrix}$$

Now, the signals $\sin\theta$ and $\cos\theta$ are linearly independent and

$$\det \begin{pmatrix} 1/2 & 1/2 \\ -\mathbf{i}/2 & \mathbf{i}/2 \end{pmatrix} = \mathbf{i}/2 \neq 0.$$

This clearly implies that $\{e^{\mathbf{i}\theta}, e^{-\mathbf{i}\theta}\}$ forms a free system (Note hat, by this way, each frecuency θ in the formal Fourier series of a signal will have two frecuencies (θ and $-\theta$) for its complex version).

Of course, we know a little bit more: we know that the sequence of signals $\left\{e^{\mathbf{i}n\theta}\right\}_{n=-\infty}^{\infty}$ is not only a free system but also forms an orthogonal system in the Hilbert space

$$L^2(-\pi,\pi) = \{x : \mathbb{R} \to \mathbb{C} : \int_{-\infty}^{\infty} |x(t)|^2 dt < \infty\},$$

with respect to the inner product given by

$$(x(t), y(t))_{L^2(-\pi,\pi)} = \int_{-\pi}^{\pi} x(t)\overline{y(t)}dt.$$

Indeed it is not difficult to check the identities

$$(e^{\mathbf{i}n\theta}, e^{\mathbf{i}m\theta})_{L^2(-\pi,\pi)} = \int_{-\pi}^{\pi} e^{\mathbf{i}n\theta}\overline{e^{\mathbf{i}m\theta}}d\theta = \begin{cases} 0 & \text{si } n \neq m \\ 2\pi & \text{si } n = m \end{cases}$$

This means that, given an arbitrary signal $x(t) \in L^2(-\pi,\pi)$, we are able to associate to $x(t)$ a (formal) complex Fourier series,

$$x(t) \sim \sum_{n=-\infty}^{\infty} c_n e^{\mathbf{i}nt}$$

just defining its Fourier coefficients by the formula:

$$\begin{aligned} c_n &= \frac{(x(t), e^{\mathbf{i}nt})_{L^2(-\pi,\pi)}}{(e^{\mathbf{i}nt}, e^{\mathbf{i}nt})_{L^2(-\pi,\pi)}} = \frac{(x(t), e^{\mathbf{i}nt})_{L^2(-\pi,\pi)}}{2\pi} = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\overline{e^{\mathbf{i}nt}}dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-\mathbf{i}nt}dt \text{ (for all } n \in \mathbb{Z}). \end{aligned}$$

Remark 2. *The formal Fourier series $x(t) \sim \sum_{n=-\infty}^{\infty} c_n e^{\mathbf{i}nt}$ is also senseful for arbitrary absolutely integrable signals $x(t) \in L^1(-\pi,\pi)$, since $\left|e^{\mathbf{i}t}\right| = 1$ for all $t \in \mathbb{R}$.*

2.4. **The complex Fourier series as an operator $\mathcal{F} : L^1(-\pi,\pi) \to c_0(\mathbb{Z})$.** Given an absolutely integrable signal $x(t) \in L^1(-\pi,\pi)$, we can define its Fourier coefficients by

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-\mathbf{i}nt}dt,\ n = 0, \pm 1, \pm 2, ...$$

To stand out the dependence of Fourier coefficients with respect of the signal $x(t)$, it is usual to denote to $c_n = \widehat{x}(n)$ (e.g., see [10]). This allow us to define an operator $\mathcal{F}$ by $\mathcal{F}(x) = \{\widehat{x}(n)\}_{n=-\infty}^{\infty}$. In this section, we will study some interesting properties of $\mathcal{F}$.

Of course, from the relation $x(t) \in L^1(-\pi,\pi)$, it follows (inmediately!) that the sequence of Fourier coefficients of $x(t)$, satisfies $\{\widehat{x}(n)\}_{n=-\infty}^{\infty} \in l_\infty(\mathbb{Z})$. Indeed it is clear that $|\widehat{x}(n)| \leq ||x(t)||_{L^1(-\pi,\pi)}$ for all $n \in \mathbb{Z}$. Moreover, $\mathcal{F}$ is behaves linearly, since the relation $\widehat{ax(t)+by(t)}(n) = a\widehat{x}(n) + b\widehat{y}(n)$ holds for all $a, b \in \mathbb{C}$ and $x(t), y(t) \in L^1(-\pi,\pi)$. This means that $\mathcal{F}$:$L^1(-\pi,\pi) \to l_\infty(\mathbb{Z})$ is a bounded linear operator.

A very important result, which will be useful for the study of convergence of Fourier series the next section, is the following classical theorem:

Theorem 38. ***(Riemann-Lebesgue's Lemma).*** *Let $x(t) \in L^1(-\pi,\pi)$ be an absolutely integrable signal. Then its Fourier coefficients satisfy $\{\widehat{x}(n)\}_{n=-\infty}^{\infty} \in c_0(\mathbb{Z})$.*

Proof. We must prove that, for any choice of $x(t) \in L^1(-\pi,\pi)$, we have that

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} x(t)e^{-\mathbf{i}nt}dt = \lim_{n\to\infty}\int_{-\pi}^{\pi} x(t)(\cos nt - \mathbf{i}\sin nt)dt = 0.$$

This is clearly equivalent to prove that

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} x(t)\cos nt = \lim_{n\to\infty}\int_{-\pi}^{\pi} x(t)\sin nt dt = 0.$$

holds true for all $x(t) \in L^1(-\pi,\pi)$.

We make the proof here for Riemann integrable signals, although the result holds true for the more general setting of Lebesgue integrable signals.

From a purely geometric point of view, the result seems to be true, since the multiplication of $x(t)$ by $\cos(nt)$ produces a signal which oscillates more and more (for $n \to \infty$)

between the graphs of $|x(t)|$ and $-|x(t)|$, so that the area enclosed by $x(t)\cos nt$, $n = 1, 2...$ results from 'eating ' more and more from the area enclosed by $\pm|x(t)|$ (which is finite, by hypothesis). The situation reflected in the following pictures:

Anyway, we need a rigorous proof of the result. So, let $-\pi = t_0 < t_1 < \cdots < t_n = \pi$ be a partition of the interval $[-\pi, \pi]$, and set $m_i = \inf_{t\in(t_{i-1},t_i)} x(t)$. Then the step function

$$g(t) = \sum_{i=1}^{n} m_i \chi_{(t_{i-1},t_i)}(t) + \sum_{i=1}^{n} x(t_i)\chi_{\{t_i\}}(t).$$

satisfies $g(t) \leq x(t)$ for all $t \in (-\pi, \pi)$. On the other hand, given $\varepsilon > 0$, if $n \in \mathbb{N}$ is big enough, then (recall we are using the Riemann integral)

$$\begin{aligned}
\left|\int_{-\pi}^{\pi} x(t)\sin(Nt)dt - \int_{-\pi}^{\pi} g(t)\sin(Nt)dt\right| &\leq \int_{-\pi}^{\pi} |(x(t)-g(t))\sin(Nt)|\,dt \\
&\leq \int_{-\pi}^{\pi} [x(t)-g(t)]\,dt \\
&= \int_{-\pi}^{\pi} x(t)dt - \sum_{i=1}^{n} m_i(t_i - t_{i-1}) \leq \frac{\varepsilon}{2}
\end{aligned}$$

Hence

$$\begin{aligned}
\left|\int_{-\pi}^{\pi} x(t)\sin(Nt)dt\right| &\leq \left|\int_{-\pi}^{\pi} g(t)\sin(Nt)dt\right| + \left|\int_{-\pi}^{\pi} x(t)\sin(Nt)dt - \int_{-\pi}^{\pi} g(t)\sin(Nt)dt\right| \\
&\leq \left|\int_{-\pi}^{\pi} g(t)\sin(Nt)dt\right| + \frac{\varepsilon}{2}
\end{aligned}$$

But

$$\begin{aligned}
\left|\int_{-\pi}^{\pi} g(t)\sin(Nt)dt\right| &= \left|\sum_{i=1}^{n} m_i \int_{t_{i-1}}^{t_i} \sin(Nt)dt\right| \\
&= \frac{1}{N}\left|\sum_{i=1}^{n} m_i(\cos(Nt_i) - \cos(Nt_{i-1}))dt\right| \\
&\leq \frac{2}{N}\left|\sum_{i=1}^{n} m_i\right|
\end{aligned}$$

Taking N big enough, we conclude that

$$\left|\int_{-\pi}^{\pi} g(t)\sin(Nt)dt\right| \leq \frac{\varepsilon}{2},$$

since n is fixed. Hence

$$\left|\int_{-\pi}^{\pi} x(t)\sin(Nt)dt\right| \leq \varepsilon$$

for all N big enough. This ends the proof (for $\int_{-\pi}^{\pi} x(t)\cos(Nt)dt$ the computations are analogous). $\square$

Alternative Proof. All deep mathematical results admit several proofs, and the Riemann-Lebesgue Theorem is not an exception!. Here we would like to remark a very easy proof of the result, based in Approximation Theory. As everybody knows, Weierstrass Approximation theorem claims that polynomials are a dense subspace of $\mathbf{C}[a,b]$, with the uniform norm. An equivalent formulation of the result claims that trigonometric polynomials are dense in $\mathbf{C}_{2\pi}(\mathbb{R})$. There are many proofs of this result (several of them based on Fourier Analysis techniques). One of them is via a more general result, which is named the Stone-Weierstrass Theorem and characterizes the dense sub-algebras of $\mathbf{C}[a,b]$ via an easy-to-check separation property. So, with no knowledge of Fourier series, we can claim that trigonometric polynomials are a dense subset of $\mathbf{C}_{2\pi}(\mathbb{R})$ with the uniform norm. On the other hand, it is clear that the continuous 2π-periodic functions are dense in $L^1(-\pi,\pi)$. It follows that trigonometric polynomials are dense in $L^1(-\pi,\pi)$. If $p(t)$ is a trigonometric polinomial then it is clear that $\widehat{p}(n)=0$ for all $|n|>\deg p$. Now, let $x(t)\in L^1(-\pi,\pi)$ and let $\varepsilon>0$ be arbitrarely fixed. Let us take $p(t)$ a trigonometric polynomial such that $\|x(t)-p(t)\|_{L^1(-\pi,\pi)}<\varepsilon$. Then

$$\begin{aligned}
|\widehat{x}(n)| &\leq \left|\widehat{x-p}(n)\right| + |\widehat{p}(n)| \\
&\leq \|x(t)-p(t)\|_{L^1(-\pi,\pi)} + |\widehat{p}(n)| \\
&< \varepsilon + 0 = \varepsilon, \text{ for all } n > \deg p = N_0(\varepsilon).
\end{aligned}$$

This ends the proof. □

Corollary 39. $\mathcal{F}: L^1(-\pi,\pi)\to c_0(\mathbb{Z})$, $\mathcal{F}(x)=\{\widehat{x}(n)\}_{n=-\infty}^{\infty}$ *is a bounded linear operator.*

Other elementary properties of Fourier coefficients are the following:

- If $\overline{x}(t):=\overline{x(t)}$ for all $t\in(-\pi,\pi)$, then $\widehat{\overline{x}}(n)=\overline{\widehat{x}(-n)}$, for all $n\in\mathbb{Z}$.
- Let us assume that $x(t)$ belongs to $L^1(-\pi,\pi)$ and is a 2π-periodic signal. For each $\delta\in\mathbb{R}$ we define $x_\delta(t):=x(t-\delta)$. Then $\widehat{x_\delta}(n)=\widehat{x}(n)e^{-\mathbf{i}n\delta}$ for all $n\in\mathbb{Z}$.
- Let us assume that $x(t)\in L^1(-\pi,\pi)$ is a 2π-periodic signal and $\widehat{x}(0)=0$ (i.e., its mean value $\int_{-\pi}^{\pi}x(t)dt$ vanishes), and let us define $X(t)=\int_0^t x(s)ds$. Then $X(t)$ is a 2π-periodic signal, $X(t)\in L^1(-\pi,\pi)$ and $\widehat{X}(n)=\frac{1}{\mathbf{i}n}\widehat{x}(n)$ for all $n\neq 0$.

Proof. We will only make the proof for the last claim (the other are easy exercises!). The proof is easy but it uses integration by parts, a very standard technique for the computation of Fourier series.

It follows from the definition of $X(t)$ that

$$X(t+2\pi)-X(t)=\int_t^{t+2\pi}x(s)ds=\int_{-\pi}^{\pi}x(\eta)d\eta=0 \text{ (where } \eta=s-t-\pi),$$

so that $X(t)$ is 2π-periodic. On the other hand,

$$\begin{aligned}\widehat{X}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X(t)e^{-\mathbf{i}nt}dt \\ &= \frac{1}{2\pi}\left(X(t)\frac{e^{-\mathbf{i}nt}}{-\mathbf{i}n}\right]_{-\pi}^{\pi} + \frac{1}{\mathbf{i}n}\int_{-\pi}^{\pi} X'(t)e^{-\mathbf{i}nt}dt\right) \\ &= \frac{1}{\mathbf{i}n}\frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-\mathbf{i}nt}dt = \frac{1}{\mathbf{i}n}\widehat{x}(n);\ n = \pm 1, \pm 2, ...,\end{aligned}$$

which is what we wanted to prove. □

Obviously, the signal $X(t) = \int_0^t x(s)ds$ is smoother than the signal $x(t)$ (whatever $x(t)$ is). We have proved that, under the assumption that $\widehat{x}(0) = 0$ (which is what makes $X(t)$ a 2π-periodic signal), the rate of convergence to zero of the Fourier coefficients of $X(t)$ is faster than the rate of convergence of the Fourier coefficients of $x(t)$. This is the first (trivial) step in a difficult theory which essentially proves that there exists a strong relation between the rate of convergence to zero of the Fourier coefficients of a signal and its smoothness properties. The general claim is of philosophical character and can be stated as follows: as smoother the signal is, as faster the rate of convergence of its Fourier coefficients is (see [13]).

Theorem 40. ***(Convolution and Fourier coefficients)*** *Let us assume that $x(t), y(t) \in L^1(-\pi,\pi)$ are 2π-periodic functions. Then the convolution*

$$h(t) = (x * y)(t) := \frac{1}{2\pi}\int_{-\pi}^{\pi} x(s)y(t-s)ds$$

satisfies $h(t) \in L^1(-\pi,\pi)$ and $||h||_{L^1(-\pi,\pi)} \leq \frac{1}{2\pi}||x||_{L^1(-\pi,\pi)}||y||_{L^1(-\pi,\pi)}$. Moreover, the relation

$$\widehat{h}(n) = \widehat{x}(n)\widehat{y}(n)$$

holds true for all $n \in \mathbb{Z}$. Finally, the relation

$$\widehat{xy}(n) = \sum_{k=-\infty}^{\infty} \widehat{x}(k)\widehat{y}(n-k)$$

also holds true for all n (i.e., $\mathcal{F}(xy) = \mathcal{F}(x) * \mathcal{F}(y)$).

Proof. We make a draft of proof (we do not make all details since this implies the use of a little bit of Meassure Theory, which is outside from the scope of the course. The complete proof can be seen, for example, in the very nice classical monograph [10, pg.

5]). To prove that $h(t) \in L^1(-\pi,\pi)$ and $||h||_{L^1(-\pi,\pi)} \leq \frac{1}{2\pi}||x||_{L^1(-\pi,\pi)}||y||_{L^1(-\pi,\pi)}$ we make the following estimation:

$$\begin{aligned}
||h||_{L^1(-\pi,\pi)} &= \int_{-\pi}^{\pi} |h(t)|dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\left|\int_{-\pi}^{\pi} x(s)y(t-s)ds\right|\right)dt \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{\pi} |x(s)y(t-s)|\,ds\right)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{\pi} |x(s)y(t-s)|\,dt\right)ds \text{ (Fubini Theor.)} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} |x(s)|\left(\int_{-\pi}^{\pi} |y(t-s)|\,dt\right)ds \;= \frac{1}{2\pi}||y||_{L^1(-\pi,\pi)}\int_{-\pi}^{\pi}|x(s)|ds \\
&= \frac{1}{2\pi}||y||_{L^1(-\pi,\pi)}||x||_{L^1(-\pi,\pi)}.
\end{aligned}$$

Now, we will compute the Fourier coefficients of h:

$$\begin{aligned}
\widehat{h}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} h(t)e^{-\mathbf{i}nt}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} x(s)y(t-s)ds\right)e^{-\mathbf{i}nt}dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} x(s)y(t-s)e^{-\mathbf{i}nt}dt\right)ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} x(s)e^{-\mathbf{i}ns}y(t-s)e^{-\mathbf{i}n(t-s)}dt\right)ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} y(t-s)e^{-\mathbf{i}n(t-s)}dt\right)x(s)e^{-\mathbf{i}ns}ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \widehat{y}(n)x(s)e^{-\mathbf{i}ns}ds = \widehat{y}(n)\widehat{x}(n).
\end{aligned}$$

Finally, the property $\mathcal{F}(xy) = \mathcal{F}(x) * \mathcal{F}(y)$ essentially follows from the fact that:

$$\left(\sum_{n=-\infty}^{\infty}\widehat{x}(n)e^{\mathbf{i}nt}\right)\left(\sum_{n=-\infty}^{\infty}\widehat{y}(n)e^{\mathbf{i}nt}\right) = \sum_{n=-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty}\widehat{x}(k)\widehat{y}(n-k)\right)e^{\mathbf{i}nt}. \;\square$$

Taking into account the relationship that exists between rate of convergence to zero of the Fourier coefficients of a signal and its smoothness properties, and the fact that the identities $\widehat{x*y}(n) = \widehat{x}(n)\widehat{y}(n)$ $(n \in \mathbb{Z})$ hold for arbitrary choices of signals $x(t), y(t) \in L^1(-\pi,\pi)$, we conclude the following (philosophical) claim: 'If we make the convolution of a signal $x(t)$ against another signal $y(t)$ which is smooth up to certain fixed order (i.e., Fourier coefficients of y approach zero at a certain pre-fixed rate of convergence), the result is a signal at least as smooth as y is' [**Proof.** If is enough to note that $|(\widehat{x*y})(n)| = |\widehat{x}(n)||\widehat{y}(n)| \leq ||\mathcal{F}(x)||_{l_\infty(\mathbb{Z})}|\widehat{y}(n)| = C|\widehat{y}(n)|$, $n \in \mathbb{Z}$]. This means that convolution operation usually has an smoothing effect on the signals we treat with.

2.5. **Fourier series and least square approximation.** Let $x(t), y(t)$ be 2π-periodic signals. We can think about $\int_{-\pi}^{\pi}|x(t)-y(t)|^2dt$ (the least square approximation error) as a meassure of the kindness of one of the signals as an approximation of the other. This point of view has many computational advantages (if we compare with the uniform distance $||x-y||_\infty = \sup_{t\in[-\pi,\pi]}|x(t)-y(t)|$), since

$$L^2(-\pi,\pi) = \{x : \|x\|_{L^2(-\pi,\pi)} = \left\{\int_{-\pi}^{\pi}|x(t)|^2dt\right\}^{\frac{1}{2}} < \infty\}$$

is a Hilbert space, which implies that best approximation errors from finitely dimensional subspaces are easy to compute, thanks to the geometry of the space. In this section we do not develope the general theory but we study some concrete cases of least square (and weighted least square) approximation of signals.

Let us assume that $x(t)$ is a 2π-periodic signal and we are interested in its approximation by a finite linear combination of sines and cosines,

$$g_N(t) = \frac{A_0}{2} + \sum_{k=1}^{N}(A_k\cos(kt) + B_k\sin(kt)),$$

in the least square sense. Then we are looking for a set of coefficients A_k, B_h ($k \in \{0,1,...,N\}$, $h \in \{1,...,N\}$) such that they minimize the error function:

$$E(A_0,..,A_N,B_1,...,B_N) = \int_{-\pi}^{\pi}\left[x(t) - \frac{A_0}{2} + \sum_{k=1}^{N}(A_k\cos(kt)+B_k\sin(kt))\right]^2 dt$$

Clearly, E is a positive differentiable function with domain $\mathbb{R}^{2N+1}$. This implies that its global minimun is also a local minimun, so that we can compute it by solving the system of equations $\frac{\partial E}{\partial A_m} = \frac{\partial E}{\partial B_h} = 0$, $0 \le m \le N$, $0 < h \le N$. But

$$\begin{aligned}
\frac{\partial E}{\partial A_m} &= \begin{cases} \frac{\partial}{\partial A_m}\left(\frac{A_0^2}{2}\pi - A_0\int_{-\pi}^{\pi}x(t)dt\right) & \text{if } m=0 \\ \frac{\partial}{\partial A_m}\left(A_m^2\int_{-\pi}^{\pi}\cos^2(mt)dt - 2A_m\int_{-\pi}^{\pi}x(t)\cos(mt)dt\right) & \text{if } m>0 \end{cases} \\
&= 2\left(\pi A_m - \int_{-\pi}^{\pi}x(t)\cos(mt)dt\right) \\
\frac{\partial E}{\partial B_h} &= \frac{\partial}{\partial B_h}\left(B_h^2\int_{-\pi}^{\pi}\sin^2(ht)dt - 2B_h\int_{-\pi}^{\pi}x(t)\sin(ht)dt\right) \\
&= 2\left(\pi B_h - \int_{-\pi}^{\pi}x(t)\sin(ht)dt\right)
\end{aligned}$$

Thus, the unique critical point of E produces the approximant $g_N = S_Nx$, the N-th partial sum of the Fourier series of our signal $x(t)$. Now, to compute the global minimun of E it is enough to restrict ourselves to the closed ball centered in the origin of coordinates of

$\mathbb{R}^{2N+1}$ and radius $R = 2\int_{-\pi}^{\pi}|x(t)|^2dt$, so that we can use the famous Weierstrass theorem on existence of global extrema for continuous functions defined on bounded closed subsets of $\mathbb{R}^n$ to guarantee the existence of a global minimun for E. Hence $S_N x$ is the best least square approximation we wanted to compute.

2.5.1. *Some complementary remarks.* Some remarks are neccesary. For example, the same proof can be used for all system of functions $\mathbf{W} = \{u_k\}_{k=0}^N$ which is orthogonal with respect to a certain weight function $\mathbf{w}(t)$ in an interval (a,b), $-\infty \leq a < b \leq \infty$. To check this, we define

$$E(A_0,..,A_N) = \int_a^b \left[x(t) - \sum_{k=1}^N A_k u_k(t)\right]^2 \mathbf{w}(t)dt$$

and we compute its critical points. Clearly,

$$\begin{aligned}\frac{\partial E}{\partial A_m} &= \frac{\partial}{\partial A_m}\left(A_m^2\int_a^b u_m^2(t)\mathbf{w}(t)dt - 2A_m\int_a^b x(t)u_m(t)\mathbf{w}(t)dt\right)\\ &= 2\left\{\left(\int_a^b u_m^2(t)\mathbf{w}(t)dt\right)A_m - \int_a^b x(t)u_m(t)\mathbf{w}(t)dt\right\},\end{aligned}$$

so that there exists a unique critical point, which is given by

$$A_m = \frac{\int_a^b x(t)u_m(t)\mathbf{w}(t)dt}{\int_a^b u_m^2(t)\mathbf{w}(t)dt} = \frac{(x(t),u_m(t))_{L^2_\mathbf{w}}}{(u_m(t),u_m(t))_{L^2_\mathbf{w}}},\ m = 0,1,...,N.$$

Hence the best approximation of $x(t)$ by elements of $\mathbf{span}(\mathbf{W})$, is given by

$$S_N^{\mathbf{W}}(t) = \sum_{k=0}^N \frac{(x(t),u_m(t))_{L^2_\mathbf{w}}}{(u_m(t),u_m(t))_{L^2_\mathbf{w}}}u_m(t).$$

This was proved by Gram in 1883 and has many important consequences. We stand up here only a few of them:

Corollary 41. *Let us assume that* $x(t) \in L^2_\mathbf{w}(a,b)$ *and* $\mathbf{W} = \{u_k\}_{k=0}^\infty$ *is an orthogonal system in* (a,b) *with respect to the weight function* $\mathbf{w}(t)$*. The following claims are equivalent:*

i) The partial sums $S_N^{\mathbf{W}}(x)(t) = \sum_{k=0}^N c_k u_k(t)$, $c_k = \frac{(x(t),u_k(t))_{L^2_\mathbf{w}}}{(u_k(t),u_k(t))_{L^2_\mathbf{w}}}$ *of the Fourier series of a signal* $x(t)$ *with respect to the system* $\mathbf{W}$ *converge to* $x(t)$ *in* $L^2_\mathbf{w}(a,b)$*.*

ii) Parseval's identity holds:

$$||x(t)||^2_{L^2_\mathbf{w}(a,b)} = \sum_{k=0}^\infty c_k^2||u_k||^2_{L^2_\mathbf{w}(a,b)}$$

Corollary 42. ***(Bessel's Inequality)*** *If $x(t) \in L^2_{\mathbf{w}}(a,b)$ and $\mathbf{W} = \{u_k\}_{k=0}^{\infty}$ is an orthogonal system in (a,b) with respect to the weight function $\mathbf{w}(t)$, then*

$$\sum_{k=0}^{\infty} c_k^2 ||u_k||^2_{L^2_{\mathbf{w}}(a,b)} \leq ||x(t)||^2_{L^2_{\mathbf{w}}(a,b)}.$$

Corollary 43. *The following claims are equivalent:*

i) $\mathbf{W} = \{u_k\}_{k=0}^{\infty}$ is a complete orthogonal system in (a,b) with respect to the weight function $\mathbf{w}(t)$.

ii) For all $x(t) \in L^2_{\mathbf{w}}(a,b)$ Parseval's identity holds.

iii) $\mathbf{span}\{u_k\}_{k=0}^{\infty}$ *is a dense subspace of $L^2_{\mathbf{w}}(a,b)$.*

2.5.2. *The complex Fourier series as an isometry $\mathcal{F} : L^2(-\pi,\pi) \to l_2(\mathbb{Z})$.* Bessel's inequality implies that $\mathcal{F}(L^2(-\pi,\pi)) \subseteq l_2(\mathbb{Z})$. We can ask about the suryectivity of $\mathcal{F}$ (i.e., we would like to know if $\mathcal{F}(L^2(-\pi,\pi)) = l_2(\mathbb{Z})$). If $\mathcal{F} : L^2(-\pi,\pi) \to l_2(\mathbb{Z})$ were onto, by considering on $L^2(-\pi,\pi)$ the norm

$$\|x(t)\|_2 = \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} |x(t)|^2 \, dt \right\}^{\frac{1}{2}},$$

the map $\mathcal{F} : L^2(-\pi,\pi) \to l_2(\mathbb{Z})$ should be an isometry of Hilbert spaces. This is in fact the case. The corresponding result was proved by Riesz and, independently by Fischer, in 1907.

Theorem 44. ***(Riesz-Fischer, 1907)*** *$\mathcal{F} : L^2(-\pi,\pi) \to l_2(\mathbb{Z})$ is an isometry of Hilbert spaces.*

The proof of this result is outside of the scope of the present course, since it is strongly based on the concept of Lebesgue integral.

Riesz-Fischer theorem does not hold if we replace the space of square Lebesgue integrable functions by the space of square Riemann integrable functions [**Proof.** We can justify the claim as follows: the set of square Riemann integrable signals is not complete under the metric induced by the norm $\|\cdot\|_2$ (so that, it is not a Hilbert space) but $l_2(\mathbb{Z})$ is a Hilbert space]. This result was a key argument for the impulse of the (new) concept of Lebesgue integral at the beguining of the XX century.

From the point of view of signals theory, we think about the operator $\mathcal{F} : L^2(-\pi,\pi) \to l_2(\mathbb{Z})$ as a transform, which sends analog energy signals $x(t) \in L^2(-\pi,\pi)$ to discrete energy signals $\mathcal{F}(x) = \{\widehat{x}(n)\}_{n=-\infty}^{\infty} \in l_2(\mathbb{Z})$. Moreover, Parseval's identity say us that $\mathcal{F}$ preserves the energy of our signals and Riesz-Fischer theorem states that the process is invertible. Hence we can save all the information of our analog signal $x(t)$ (which is a continuoum of values, $\{x(t) : t \in (-\pi, \pi)\})$ in terms of a discrete set of values: the Fourier coefficients of the signal. We say that $\{\widehat{x}(n)\}_{n=-\infty}^{\infty}$ represents the signal $x(t)$ in the frequency domain and $\{x(t) : t \in (-\pi, \pi)\}$ represents it in the time domain.

From a more mathematical point of view, Riesz-Fischer theorem is also a little bit surprising (and very important) result, since it shows that certain function spaces are indistinguishable of certain sequence spaces. But sequence spaces are allways easier to handle!. In fact, Riesz-Fisher theorem proves the existence of a unique prototipe of separable Hilbert space: $l_2(\mathbb{Z})$. It is, of course, one of the first important results that exists in Functional Analysis.

Remark 3. *Been careful with definitions, we can not say that 2π-periodic analog signals $x(t) \in L^2(-\pi, \pi)$ are energy signals, since they do not belong to $L^2(\mathbb{R})$. But we can say that they are power signals, since*

$$\mathbf{P}(x) = ||x(t)||_2^2$$

Proof. *Let us set $M = 2k\pi + T \in [2k\pi, 2(k+1)\pi)$, $k \in \mathbb{N}$. Then:*

$$\begin{aligned}
&\frac{1}{2M}\int_{-M}^{M} |x(t)|^2\, dt \\
= &\frac{1}{2M}\left\{\int_{-2k\pi-T}^{-2k\pi} |x(t)|^2\, dt + 2\sum_{h=0}^{k-1}\int_{2h\pi}^{2(h+1)\pi} |x(t)|^2\, dt + \int_{2k\pi}^{2k\pi+T} |x(t)|^2\, dt\right\} \\
= &\frac{1}{4k\pi+2T}\left\{\int_{-T}^{0} |x(t)|^2\, dt + \int_{0}^{T} |x(t)|^2\, dt + 2k\int_{0}^{2\pi} |x(t)|^2\, dt\right\}
\end{aligned}$$

so that

$$\begin{aligned}
\mathbf{P}(x) &= \lim_{M\to\infty}\frac{1}{2M}\int_{-M}^{M} |x(t)|^2\, dt \\
&= \lim_{k\to\infty}\frac{1}{4k\pi+2T}\left\{\int_{-T}^{0} |x(t)|^2\, dt + \int_{0}^{T} |x(t)|^2\, dt + 2k\int_{0}^{2\pi} |x(t)|^2\, dt\right\} \\
&= \lim_{k\to\infty}\frac{2k}{4k\pi+2T}\int_{0}^{2\pi} |x(t)|^2\, dt \\
&= \frac{1}{2\pi}\int_{0}^{2\pi} |x(t)|^2\, dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} |x(t)|^2\, dt = ||x(t)||_2^2.
\end{aligned}$$

This ends the proof. □

Moreover, the proof above lead us to the following interpretation of Parseval's identity: If

$$x(t) \sim \sum_{n=-\infty}^{\infty} \widehat{x}(n)e^{n\mathbf{i}t},\ x(t) \in L^2(-\pi,\pi),$$

then

$$\mathbf{P}(x) = ||x(t)||_2^2 = ||\left\{\widehat{x}(n)\right\}_{n=-\infty}^{\infty}||_{l_2(\mathbb{Z})}^2 = \sum_{n=-\infty}^{\infty} |\widehat{x}(n)|^2$$

On the other hand, we know that

$$||\widehat{x}(n)e^{n\mathbf{i}t}||_2^2 = |\widehat{x}(n)|^2 .$$

Hence we can think about the Fourier coefficients of the signal give us a description of the contribution to the power of our signal for each frequency component of the signal. It is because of this interpretation that engineers call the sequence $\left\{|\widehat{x}(n)|^2\right\}_{n=-\infty}^{\infty}$ *the spectral power density of the signal.*

3. Fourier Series II

3.1. **Rate of convergence of Fourier series and smoothness properties of signals.** Last section we proved that $x(t) \in L^2(-\pi,\pi)$ implies that $\{\widehat{x}(n)\}_{n=-\infty}^{\infty} \in l_2(\mathbb{Z})$. This, of course, does not implies convergence of the Fourier series $\sum_{n=-\infty}^{\infty} \widehat{x}(n)e^{\mathbf{i}nt}$. Moreover, also for summable Fourier series we have not proved (till now) that $x(t) = \sum_{n=-\infty}^{\infty} \widehat{x}(n)e^{\mathbf{i}nt}$ (pointwise).

To guarantee the convergence of a Fourier series $\sum_{n=-\infty}^{\infty} \widehat{x}(n)e^{\mathbf{i}nt}$ to $x(t)$, we will need to assume certain smoothness on the signal. For our purposes, we will assume much more than neccesary: our signals will be piecewise smooth of class $\mathbf{C}^{\infty}$. This means that, except for a finite number of points of the interval $[-\pi,\pi]$, $(-\pi = \theta_0 < \theta_1 < \theta_2 < \cdots < \theta_J = \pi)$, the signals have derivatives of all orders. Moreover, we will also assume the existence of right and left derivatives $x^{(k)}(\theta_i^+)$ and $x^{(k)}(\theta_i^-)$ for all $i \leq J$ and for all $k \in \mathbb{N}$ (hence the derivatives only have jump discontinuities).

Before making the computation of the Fourier coefficientes of such a signal, we make the following remark:

Remark 4. *Let us assume that $u(t)$ has jump discontinuities at the points $\{a, b\}$. Then to carry the integration by parts of $\int_a^b udv$, we should make:*

$$\begin{aligned}\int_a^b udv &= \lim_{t\to 0^+}\int_{a+t}^{b-t} udv = \lim_{t\to 0^+}\left(uv\Big]_{a+t}^{b-t} - \int_{a+t}^{b-t} udv\right)\\ &= uv\Big]_{a^+}^{b^-} - \int_a^b udv;\end{aligned}$$

where $uv\Big]_{a^+}^{b^-} = u(b^-)v(b) - u(a^+)v(a)$.

Let $x(t)$ be a piecewise $\mathbf{C}^\infty$ signal. Then its Fourier coefficients are given by:

$$\begin{aligned}a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos(kt)dt\\ &= \frac{1}{\pi}\sum_{j=1}^{J}\int_{\theta_{j-1}}^{\theta_J} x(t)\cos(kt)dt\\ &= \frac{1}{\pi}\sum_{j=1}^{J}\left(x(t)\frac{\sin(kt)}{k}\right]_{\theta_{j-1}^+}^{\theta_j^-} - \int_{\theta_{j-1}}^{\theta_j} x'(t)\frac{\sin(kt)}{k}dt\Bigg)\end{aligned}$$

If the signal is continuous, the sum $\sum_{j=1}^{J}\left\{ x(t)\frac{\sin(kt)}{k}\Big]_{\theta_{j-1}^+}^{\theta_j^-}\right\}$ vanish, so that we must only estimate the sum $\sum_{j=1}^{J}\left\{ -\int_{\theta_{j-1}}^{\theta_J} x'(t)\frac{\sin(kt)}{k}dt\right\}$. Otherwise, the Fourier coefficients should satisfy $|a_k| = \mathbf{O}(1/k)$. Now, if the first part of the sum vanishes and we make again an integration by parts, we obtain

$$a_k = \frac{1}{\pi}\sum_{j=1}^{J}\left(x'(t)\frac{\cos(kt)}{k^2}\right]_{\theta_{j-1}^+}^{\theta_j^-} - \int_{\theta_{j-1}}^{\theta_J} x''(t)\frac{\cos(kt)}{k^2}dt\Bigg)$$

Hence, $|a_k| = \mathbf{O}(1/k^2)$ for continuous piecewise smooth signals.

If the first derivative $x'(t)$ is also continuous then

$$\frac{1}{\pi}\sum_{j=1}^{J}\left(x'(t)\frac{\cos(kt)}{k^2}\right]_{\theta_{j-1}^+}^{\theta_j^-}\Bigg)$$

vanishes and, integrating again by parts, we obtain that $|a_k| = \mathbf{O}(1/k^3)$. The same argument can be repeated several times.

We have proved that $|a_k| = \mathbf{O}(1/k^2)$ holds for continuous signals with first derivative piecewise continuous[5]. Hence the corresponding Fourier series is summable, since

[5] In fact, to make the second integration by parts, we need to assume that $x'(t)$ is piecewise continuous and piecewise derivable, but this is part of the general assumptions we imposed at the beginning of this section.

$|\sin t|, |\cos t| \leq 1$ for all t. In fact, the Fourier series is absolutely and uniformly sumable[6]. Moreover, the Fourier series will converge to the original signal $x(t)$ [7]. This theorem was proved by Heine (1870), who proved it as a consequence of Bessel's inequality.

The discussion above lead us to the following main conclusion: under the general assumptions of this section, we only must be worry about pointwise convergence for Fourier series of non continuous signals. Hence, given a signal $x(t)$ with a finite number of jump discontinuities, we should study $\lim_{N\to\infty}(S_N x)(t_0)$ for the following two cases: i) t_0 is a continuity point of the signal and ii) t_0 is not a continuity point of the signal.

3.2. **Convergence at the continuity points.** In this section we assume that $x(t)$ is a noncontinuous piecewise $\mathbf{C}^{(\infty)}$ signal, with discontinuity points $-\pi = \theta_0 < \theta_1 < \theta_2 < \cdots < \theta_J = \pi$ and we ask the following question:

(Q1) Does the Fourier series of $x(t)$ converges for the values $t \in (\theta_{j-1}, \theta_j)$? If it is so, does it converges to $x(t)$?

We know that

$$(S_N x)(t) = \frac{a_0}{2} + \sum_{k=1}^{N} (a_k \cos(kt) + b_k \sin(kt)), \tag{2}$$

where

$$\begin{cases} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x(s)\cos(ks)ds \text{ ; for all } k \geq 0 \\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x(s)\sin(ks)ds\text{; for all } k \geq 1 \end{cases}$$

Now, if we insert the expressions of the coefficients a_k, b_k in (2), and we interchange the sum (which is finite) and the integral, we obtain

$$\begin{aligned} (S_N x)(t) &= \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\left\{\frac{1}{2} + \sum_{k=1}^{N}\{\cos(ks)\cos(kt) + \sin(ks)\sin(kt)\}\right\} ds \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\left\{\frac{1}{2} + \sum_{k=1}^{N}\cos(k(s-t))\right\} ds, \end{aligned}$$

and, if we set $u = s - t$, we have that

$$(S_N x)(t) = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t+u)\left\{\frac{1}{2} + \sum_{k=1}^{N}\cos(ku)\right\} du.$$

[6]This is nothing but Weierstrass M-test.

[7]Yet now, we have not proved this. To prove that the claim is true it will be enough to take into account the first Corollary of Fejèr's Theorem we will prove next section.

It follows that, to prove results about the convergence of $S_N x$, we need to analyze the properties of $k_n(u) = \frac{1}{2} + \sum_{k=1}^{N}\cos(ku)$. To do this, we first compute a closed formula for $k_n(u)$, as follows:

If we insert $a = ku$ and $b = u/2$ in the trigonometric formula

$$\frac{\sin(a+b) - \sin(a-b)}{2} = \cos a \sin b.$$

we obtain,

$$\sin\frac{(2k+1)u}{2} - \sin\frac{(2k-1)u}{2} = \cos(ku)\sin\frac{u}{2}$$

so that

$$\begin{aligned} k_n(u)\sin\frac{u}{2} &= \frac{1}{2}\left[\sin\frac{u}{2} + \sum_{k=1}^{N}\left(\sin\frac{(2k+1)u}{2} - \sin\frac{(2k-1)u}{2}\right)\right] \\ &= \frac{1}{2}\sin\frac{(2N+1)u}{2} \end{aligned}$$

and

$$(S_N x)(t) = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t+u)\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2}.$$

Now, it follows from the least squares best approximation property of $S_N x$ that, if we set $x(t) = 1$ then $(S_N x)(t) = 1$, so that,

$$1 = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2}.$$

We multiply both sides of the identity above by $x(t)$ and we put $x(t)$ inside of the integral, in the righthand side of the identity (recall we are taking the integration with respect to the variable u), to obtain

$$x(t) = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2}$$

It follows that

$$x(t) - (S_N x)(t) = \frac{1}{\pi}\int_{-\pi}^{\pi} (x(t) - x(t+u))\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2}.$$

so that, to estimate the difference $x(t_0) - (S_N x)(t_0)$ we have got an expresssion which depends on $\frac{x(t_0)-x(t_0+u)}{2\sin(u/2)}$. This should be useful for signals $x(t)$ derivable in t_0, since they satisfy $\lim_{u\to 0}\frac{x(t_0)-x(t_0+u)}{u} = x'(t_0) \in \mathbb{R}$ and, if we decompose $\Phi(u) = \frac{x(t_0)-x(t_0+u)}{2\sin(u/2)}$ as a product,

$$\Phi(u) = \frac{x(t_0) - x(t_0+u)}{u}\frac{1}{\sin(u/2)/(u/2)},$$

we can claim that $\Phi(u)$ is a bounded function, since $\lim_{u\to 0}\frac{\sin u}{u}=1$.

We want to estimate the limit

$$\lim_{N\to\infty}\frac{1}{\pi}\int_{-\pi}^{\pi}\Phi(u)\sin(\frac{(2N+1)u}{2})du$$

To do this, we use the following result (which is a variation of Riemann-Lebesgue theorem):

Lemma 45. *Let us assume that $h(u)$ is a 2π-periodic absolutely integrable (on $(-\pi,\pi)$) signal. Then*

$$\lim_{n\to\infty}\frac{1}{\pi}\int_{-\pi}^{\pi}h(u)\sin(\frac{(2n+1)u}{2})du=0$$

Proof. If we take into account the formula

$$\sin\frac{(2n+1)u}{2}=\sin((n+\frac{1}{2})u)=\sin(nu)\cos(u/2)+\cos(nu)\sin(u/2)$$

and we set $h_1(u)=\cos(u/2)h(u)$ y $h_2(u)=h(u)\sin(u/2)$, we obtain that

$$\frac{1}{\pi}\int_{-\pi}^{\pi}h(u)\sin((n+\frac{1}{2})u)du=b_n(h_1)+a_n(h_2),$$

where $b_n(h_1)$ denotes the n-th Fourier coefficient (from those associated to the sines functions) of h_1 and $a_n(h_2)$ denotes the n-th Fourier coefficient (from those associated to the cosines functions) of h_2. Now, $h(u)$, $h_1(u)$ and $h_2(u)$ are absolutely integrable on $(-\pi,\pi)$, so that we can use the Riemann-Lebesgue Lemma to conclude that their Fourier coefficients converge to zero. This ends the proof. □

It follows from the discussion above that the following result holds true:

Theorem 46. ***(Dirichlet)*** *Let us assume that $\Psi(u)=\frac{x(t_0)-x(t_0+u)}{u}$ is absolutely integrable on $[-\varepsilon,\varepsilon]$ for a certain $\varepsilon>0$. Then $\lim_{N\to\infty}(S_Nx)(t_0)=x(t_0)$.*

Proof. It follows from the absolutely integrability of $\Psi(u)=\frac{x(t_0)-x(t_0+u)}{u}$ on $[-\varepsilon,\varepsilon]$ (for a certain $\varepsilon>0$) that $\Psi(u)\in L^1(-\pi,\pi)$. Hence $\Phi(u)=\Psi(u)\frac{1}{\sin(u/2)/(u/2)}\in L^1(-\pi,\pi)$, and we can use the Lemma above with the righthand side of

$$\begin{aligned}(S_Nx)(t_0)-x(t_0) &= \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{x(t_0)-x(t_0+u)}{u}\frac{1}{\sin(u/2)/(u/2)}\sin((2N+1)u/2)\frac{du}{2}\\ &= \frac{1}{\pi}\int_{-\pi}^{\pi}\Phi(u)\sin((N+\frac{1}{2})u)du,\end{aligned}$$

so that

$$\lim_{N\to\infty}(S_Nx)(t_0)=x(t_0).$$

This ends the proof. □

Corollary 47. *Let us assume that there exists the derivative $x'(t_0)$. Then $\lim_{N\to\infty}(S_N x)(t_0) = x(t_0)$.*

Proof. By hyphothesis, $x'(t_0) = \lim_{u\to 0} \frac{x(t_0)-x(t_0+u)}{u}$. This clearly implies that $\Psi(u)$ is continuous at $u = 0$, which is the unique conflictive point. Hence $\Psi(u) \in L^1(-\varepsilon, \varepsilon)$ for a certain $\varepsilon > 0$ and we can use Dirichlet's Theorem. □

3.3. **Convergence at the discontinuity points.** In this section we will try to modify the arguments of the section above to study the convergence of $(S_N x)(t)$ at the points $t = \theta_j$ $(j = 1, ..., J-1)$ [i.e., at the discontinuity points of the signal $x(t)$].

By hyphothesis, at these points our signal has right and left derivative. We look for a closed expression of $S_N x$ for which this information is useful.

We decompose the integral form of $S_N x$ as follows:

$$\begin{aligned}(S_N x)(t) &= \frac{1}{\pi}\int_{-\pi}^{\pi} x(t+u)\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2} \\ &= \frac{1}{\pi}\left(\int_{-\pi}^{0} + \int_{0}^{\pi}\right) x(t+u)\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2}\end{aligned}$$

Now, we make the change of variable $v = -u$ in the first integral which appears at the righthand side of the expression above, to obtain that

$$\begin{aligned}\frac{1}{\pi}\int_{-\pi}^{0} x(t+u)\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2} &= \frac{-1}{\pi}\int_{\pi}^{0} x(t-v)\frac{\sin((2N+1)v/2)}{\sin(v/2)}\frac{dv}{2} \\ &= \frac{1}{\pi}\int_{0}^{\pi} x(t-v)\frac{\sin((2N+1)v/2)}{\sin(v/2)}\frac{dv}{2}.\end{aligned}$$

This allows us to claim that

$$(S_N x)(t) = \frac{1}{\pi}\int_0^{\pi}(x(t+u)+x(t-u))\frac{\sin((2N+1)u/2)}{\sin(u/2)}\frac{du}{2},$$

so that the difference

$$\begin{aligned}&(S_N x)(\theta_j) - \tfrac{x(\theta_j+)+x(\theta_j-)}{2} \\ &= \tfrac{1}{\pi}\int_0^{\pi}\left(\tfrac{x(\theta_j+u)+x(\theta_j-u)}{2} - \tfrac{x(\theta_j+)+x(\theta_j-)}{2}\right)\tfrac{\sin((2N+1)u/2)}{\sin(u/2)}\tfrac{du}{2}\end{aligned}$$

can be estimated by analogous arguments to those given for the estimation of $(S_N x)(t_0) - x(t_0)$ in the section above. By this way we can prove the following:

Theorem 48. *Let us assume that $x(t)$ is a noncontinuous piecewise smooth signal, with discontinuity points given by $t = \theta_j$ $(j = 1, ..., J-1)$. If*

$$\Psi(u) = \frac{x(\theta_j+u)+x(\theta_j-u)}{2u} - \frac{x(\theta_j+)+x(\theta_j-)}{2}$$

is absolutely integrable on $(-\varepsilon,\varepsilon)$*, for a certain* $\varepsilon>0$*. Then* $\lim_{N\to\infty}(S_Nx)(\theta_j)=\widetilde{x}(\theta_j):=\frac{x(\theta_j^+)+x(\theta_j^-))}{2}$.

3.4. **Gibbs phenomenon.** Let us consider the step function:

$$x(t)=\begin{cases}-1/2 & \text{if } t\in(-\pi,0]\\ 0 & \text{if } t=0\\ 1/2 & \text{if } t\in(0,\pi]\end{cases}.$$

We have already computed its Fourier series before. It is:

$$x(t)\sim\frac{2}{\pi}\sum_{k=0}^{\infty}\frac{\sin((2k+1)t)}{2k+1}$$

The $2N+1$-th partial sum, is then given by

$$(g_Nx)(t):=(S_{2N+1}x)(t)=\frac{2}{\pi}\sum_{k=0}^{N}\int_0^t\cos((2k+1)u)du=\frac{2}{\pi}\int_0^t\left\{\sum_{k=0}^{N}\cos((2k+1)u)\right\}du$$

Now, by the use of some elementary trigonometric formulae, we can compute a closed expression for the sum $\sum_{k=0}^{N}\cos((2k+1)u)$. To be more precise, we use that

$$\frac{\sin(a+b)-\sin(a-b)}{2}=\cos a\sin b$$

to obtain

$$\cos((2k+1)u)\sin u=\frac{\sin(2(k+1)u)-\sin(2ku)}{2},$$

and, taking into account the last formula, we obtain

$$\sum_{k=0}^{N}\cos((2k+1)u)\sin u=\frac{1}{2}\sin(2(N+1)u)$$

so that

$$(g_Nx)(t)=\frac{1}{\pi}\int_0^t\frac{\sin(2(N+1)u)}{\sin u}du.$$

Clearly, g_Nx does not converge uniformly to $x(t)$, since $x(t)$ is not continuous. On the other hand, we know that g_Nx converges pointwise to $x(t)$. Moreover, it converges uniformly on compact subsets of $\mathbb{R}\setminus 2\pi\mathbb{Z}$. Hence it is natural trying to compute how much $(g_Nx)(t)$ is greater (lower) than $x(t)$ for $t\approx 0$. To study this problem, we should compute the relative extrema of $(g_Nx)(t)$ and the closed formula we have just obtained for $(g_Nx)(t)$ will be helpful, since it can easily be derivated.

It was J. Willard Gibbs the first mathematician who noticed the importance of the following phenomenon (later called Gibbs phenomenon): the graphs of the partial sums of the Fourier series, $S_N x$, have the tendency to exceed the meassure of the jump discontinuity of the signal $x(t)$ in a fixed percent, which is independent of N, and represents approximatively the 9% of the jump.

We compute (for our example), the derivative

$$(g_N x)'(t) = (S_{2N+1}x)'(t) = \frac{1}{\pi}\frac{\sin(2(N+1)t)}{\sin t}$$

and, solving the equation $(g_N x)'(t) = 0$ we obtain that the first local maximun of $g_N x$ in the interval $(0, \pi]$ is attained for $t = \frac{\pi}{2(N+1)}$, where $g_N x$ equals to

$$(g_N x)(\frac{\pi}{2(N+1)}) = \frac{1}{\pi}\int_0^{\frac{\pi}{2(N+1)}} \frac{\sin(2(N+1)u)}{\sin u} du.$$

We make the change of variable $v = 2(N+1)u$ to obtain that

$$\begin{aligned} (g_N x)(\frac{\pi}{2(N+1)}) &= \frac{1}{\pi}\int_0^{\pi} \frac{\sin v}{\sin(v/(2(N+1)))}\frac{dv}{2(N+1)} \\ &= \frac{1}{\pi}\int_0^{\pi} \frac{\sin v}{v}\frac{v/(2(N+1))}{\sin(v/(2(N+1)))}dv. \end{aligned}$$

Now, $\frac{\sin v}{v}$ is a bounded function on $[0, \pi]$, and

$$\lim_{N\to\infty}\left\{\sup_{v\in(0,\pi)} |1 - \frac{v/(2(N+1))}{\sin(v/(2(N+1)))}|\right\} = 0.$$

Hence

$$\begin{aligned} &\lim_{N\to\infty}\left(\frac{1}{\pi}\int_0^{\pi}\frac{\sin v}{v}dv - (S_N x)(\frac{\pi}{2(N+1)})\right) \\ = &\lim_{N\to\infty}\left(\frac{1}{\pi}\int_0^{\pi}\frac{\sin v}{v}dv - \frac{1}{\pi}\int_0^{\pi}\frac{\sin v}{v}\frac{v/(2(N+1))}{\sin(v/(2(N+1)))}dv\right) \\ = &\lim_{N\to\infty}\left(\frac{1}{\pi}\int_0^{\pi}\frac{\sin v}{v}\left\{1 - \frac{v/(2(N+1))}{\sin(v/(2(N+1)))}\right\}dv\right) = 0 \end{aligned}$$

This means that

$$\lim_{N\to\infty}(g_N x)(\frac{\pi}{2(N+1)}) = \frac{1}{\pi}\int_0^{\pi}\frac{\sin v}{v}dv \approx 0.58949 = 0.5 + 0.08949,$$

which is what we wanted to prove.

The following pictures represent the graphs of the partial sums $S_3(x)$, $S_{10}(x)$ and $S_{40}(x)$, respectively. They clearly show the Gibbs phenomenon:

The Gibbs phenomenon appears whenever we truncate a Fourier series and it has special importance for the study of filters, since practice forzes the truncation of Fourier series of signals.

3.5. **Fejèr sums: Weierstrass approximation theorem.** Under the unique assumption of continuity of a signal $x(t)$, it is well known that its Fourier series may be no convergent. (This is proved by the use of Banach-Steinhauss bounded map theorem, plus the computation of the norms of the operators S_N, which are of order $\log n$). On the other hand, we also know that, under very weak hiphotheses (e.g., for signals $x(t) \in L^2(-\pi,\pi)$), the partial sums $S_N(x)$ converge to the signal $x(t)$ in the least square sense (i.e., $\|S_N(x) - x\|_{L^2(-\pi,\pi)} \to 0$ for $N \to \infty$). Of course, this kind of convergence is weaker that uniform convergence. For example, it does not imply pointwise convergence. But it help us to believe that '$S_N(x)$ must be not very far from x'. So, we would like to find some method to kill the many oscillations of Fourier series which (as we already know) cause the divergence of Fourier series. O.K. such a method exists!!. Much more than that: there are many forms to modify a Fourier sequence of approximants $\{S_n(x)\}_{n=0}^{\infty}$, which produce a new sequence of approximants that (now, yes) converge uniformly to the signal $x(t)$ under the only hyphothesis that $x(t)$ is 2π-periodic and continuous. Now, we will prove the most famous result that exists on the subject: Fejèr theorem on Cèsaro summability of Fourier series:

Theorem 49. ***(Fejèr, 1903)*** *Let $x(t) \in \mathbf{C}_{2\pi}(\mathbb{R})$ be a 2π-periodic continuous signal, and let*

$$S_N(x)(t) = \sum_{k=-N}^{N} c_k e^{\mathbf{i}kt}$$

be the N-th partial sum of its Fourier series. Then

$$\lim_{n\to\infty} \|x - \sigma_N(x)\|_\infty = 0,$$

where

$$\sigma_N(x) := \frac{1}{N+1}\sum_{k=0}^{N} S_k(x),\ N = 0, 1, 2, ...$$

Before proving this important result, we would like to recall some of its consequences:

Corollary 50. ***(Weierstrass approximation theorem)*** *Trigonometric polynomials are a dense subset of $\mathbf{C}_{2\pi}(\mathbb{R})$.*

Corollary 51. *Let $x(t) \in \mathbf{C}_{2\pi}(\mathbb{R})$ and let $t_0 \in \mathbb{R}$, then at least one of the following claims holds true:*

i) The sequence of numbers $\{S_N(x)(t_0)\}_{N=0}^{\infty}$ converges to $x(t_0)$.

ii) The sequence of numbers $\{S_N(x)(t_0)\}_{N=0}^{\infty}$ does not converges in $\mathbb{R}$.

Proof of Fejèr's Theorem. We are going to prove a stronger version of the result. We prove that, if $x \in L^1(-\pi,\pi)$ is 2π-periodic and the limit

$$\widetilde{x}(t) = \lim_{h\to 0^+} \frac{x(t+h)+x(t-h)}{2}$$

exists, then $\sigma_N(x)(t) \to \widetilde{x}(t)$ for $N\to\infty$. Moreover, if $x \in \mathbf{C}_{2\pi}(\mathbb{R})$, then $\widetilde{x}(t) = x(t)$ for all t, and $\sigma_N(x)(t) \to x(t)$, $N\to\infty$ uniformly on $\mathbb{R}$.

Let us start looking for a closed formula for the Fejèr means $\sigma_N(x)$. This we make by taking in consideration that

$$S_k(x)(t) = \frac{1}{\pi}\int_0^{\pi} (x(t+u)+x(t-u))\frac{\sin((2k+1)u/2)}{\sin(u/2)}\frac{du}{2},$$

and the summation formula (that we left as exercise!)

$$\sum_{k=0}^{N-1} \sin((2k+1)t) = \frac{\sin^2(Nt)}{\sin t},$$

to obtain that

$$\begin{aligned} \sigma_N(x) &= \frac{1}{N+1}\sum_{k=0}^{N} S_k(x)(t) \\ &= \frac{1}{\pi}\frac{1}{N+1}\int_0^{\pi} \frac{x(t+u)+x(t-u)}{2}\left(\frac{\sin((N+1)u/2)}{\sin(u/2)}\right)^2 du \end{aligned}$$

On the other hand, it is clear that $\sigma_N(1) = 1$ since $S_k(1) = 1$ for all $k \geq 0$. Hence we have that

$$\begin{aligned} &|\sigma_N(x)(t) - \widetilde{x}(t)| \\ &= \frac{1}{\pi}\frac{1}{N+1}\left|\int_0^{\pi}\left(\frac{x(t+u)+x(t-u)}{2} - \widetilde{x}(t)\right)\left(\frac{\sin((N+1)u/2)}{\sin(u/2)}\right)^2 du\right| \end{aligned} \tag{3}$$

Let us set $g(t,u) = \frac{x(t+u)+x(t-u)}{2} - \widetilde{x}(t)$. It follows from $\lim_{u\to 0^+}|g(t,u)| = 0$ that, for all $\varepsilon > 0$, there exists a certain $\delta > 0$ such that $|g(t,u)| < \varepsilon/2$ for all $u \in (0,\delta)$ (we can assume that $\delta < \pi$ without loss of generality). In principle, $\delta = \delta(t,\varepsilon)$ but, if $x(t)$ is a continuous signal then $\delta = \delta(\varepsilon)$ (i.e., it does not depend on t), since in such a case

$x(t)$ will be uniformly continuous in $[0, 2\pi]$. Thus, we can decompose the integral which appears in the righthand side of (3) as a sum, as follows

$$
\begin{aligned}
|\sigma_N(x)(t) - \widetilde{x}(t)| \;\leq\; & \underbrace{\frac{1}{\pi}\frac{1}{N+1}\int_0^{\delta} |g(t,u)| \left(\frac{\sin((N+1)u/2)}{\sin(u/2)}\right)^2 du}_{\mathbf{I}} \\
& + \underbrace{\frac{1}{\pi}\frac{1}{N+1}\int_{\delta}^{\pi} |g(t,u)| \left(\frac{\sin((N+1)u/2)}{\sin(u/2)}\right)^2 du}_{\mathbf{II}}
\end{aligned}
$$

and we can make the estimations:

$$
\begin{aligned}
\mathbf{I} \;&\leq\; \frac{\varepsilon}{2(N+1)\pi}\int_0^{\delta} \left(\frac{\sin((N+1)u/2)}{\sin(u/2)}\right)^2 du \\
&\leq\; \frac{\varepsilon}{2(N+1)\pi}\int_0^{\pi} \left(\frac{\sin((N+1)u/2)}{\sin(u/2)}\right)^2 du = \frac{\varepsilon}{2}\sigma_N(1) = \frac{\varepsilon}{2}
\end{aligned}
$$

and (taking into account that $1/\sin^2(u/2)$ is strictly non-increasing on $(0, \pi]$)

$$
\begin{aligned}
\mathbf{II} \;&\leq\; \frac{1}{\pi(N+1)}\frac{1}{\sin^2(\delta/2)}\int_{\delta}^{\pi} |g(t,u)| \left(\sin((N+1)u/2)\right)^2 du \\
&\leq\; \frac{1}{\pi(N+1)}\frac{1}{\sin^2(\delta/2)}\int_0^{\pi} |g(t,u)|^2 du = \frac{1}{\pi(N+1)}\frac{1}{\sin^2(\delta/2)}H(t),
\end{aligned}
$$

where $H(t) = \int_0^{\pi} |g(t,u)|\, du < \infty$.

Let us fix the value t. Then $\frac{1}{\pi(N+1)}\frac{1}{\sin^2(\delta/2)}H(t)$ converges to zero for $N \to \infty$. This allow us to guarantee that, if $N \geq N_0$ ($N_0 \in \mathbb{N}$ big enough) then $|\sigma_N(x)(t) - \widetilde{x}(t)| \leq \varepsilon$. On the other hand, if $x(t)$ is continuous, then there exists a constant $C > 0$ such that $\int_0^{\pi} |g(t,u)|^2 du < C$ independtly of t. Hence there exists $N_0 \in \mathbb{N}$ such that $||\sigma_N(x)(t) - x(t)||_{\infty} \leq \varepsilon$ for all $N \geq N_0$. This ends the proof. □

3.5.1. *Some remarks: the Gibbs phenomenon again.* One of the most remarkable properties of Fejèr means is that they smooth the Gibbs phenomenon (exercise!). Another method which is more frequently used by engineers, was proposed by Lanczos in 1956. It is based on the observation that the riple produced by Gibbs phenomenon is concentrated on an interval of lenght the period of the last term we put in the truncated Fourier series or the period of the first term we did not sum. Thus, if we change the value $(S_N x)(t)$ by the mean value of $S_N x$ in the interval $(t - \pi/N, t + \pi/N)$, we will smooth the riple. Using this idea, Lanczos defined the series:

$$\begin{aligned}
h_N(t) &= \frac{1}{2\pi/N}\int_{t-\frac{\pi}{N}}^{t+\frac{\pi}{N}}(S_Nx)(u)du \\
&= \frac{a_0}{2}\frac{N}{2\pi}\int_{t-\frac{\pi}{N}}^{t+\frac{\pi}{N}}du+\frac{N}{2\pi}\sum_{k=1}^{N}\left\{a_k\int_{t-\frac{\pi}{N}}^{t+\frac{\pi}{N}}\cos(ku)du+b_k\int_{t-\frac{\pi}{N}}^{t+\frac{\pi}{N}}\sin(ku)du\right\} \\
&= \frac{a_0}{2}+\frac{N}{2\pi}\sum_{k=1}^{N}a_k\frac{\sin(k(t+\pi/N))-\sin(k(t-\pi/N))}{k} \\
&\quad -\frac{N}{2\pi}\sum_{k=1}^{N}b_k\frac{\cos(k(t+\pi/N))-\cos(k(t-\pi/N))}{k} \\
&= \frac{a_0}{2}+\sum_{k=1}^{N}h(N,k)\left\{a_k\cos(kt)+b_k\sin(kt)\right\};
\end{aligned}$$

where $h(N,k)=\frac{\sin(k\pi/N)}{k\pi/N}$, $k=1,...,N$.

If you make the computations, you will check that this new form of summation of the Fourier series Sx reduces the excess we computed before for the Gibbs phenomenon. The old excess was 0.08949... and the new one is 0.01187... This means a reduction of more than 1/8 the former.

We can now interpret Fejèr sums from a new perspective, as follows: they are another form of weighting the terms $a_k\cos(kt)+b_k\sin(kt)$ of the Fourier series of $x(t)$. This can be easily proved, since, if we expand the formula $\sigma_Nx=\frac{1}{N+1}(S_0x+...+S_Nx)$, we obtain:

$$(\sigma_Nx)(t)=\frac{a_0}{2}+\sum_{k=1}^{N}\sigma(N,k)\left\{a_k\cos(kt)+b_k\sin(kt)\right\};$$

where $\sigma(N,k)=\frac{N-k}{N}$, $k=1,...,N$.

The following pictures show the smoothing efect of Lanczos and Fejèr means, on the Fourier series of the step signal

$$x(t)=\begin{cases}-1/2 & \text{if } t\in(-\pi,0] \\ 0 & \text{if } t=0 \\ 1/2 & \text{if } t\in(0,\pi]\end{cases}\sim\frac{2}{\pi}\sum_{k=0}^{\infty}\frac{\sin((2k+1)t)}{2k+1}$$

3.6. **More results on convergence of Fourier series.** The study of convergence of Fourier series is a huge subject, so that it contains many kinds of results and the theorems we have proved yet now, are only a small piece of the theory. In this section we would like to compile some more results about pointwise convergence but also to develope a little bit the theory of convergence in norm.

3.6.1. *Lebesgue theorem on pointwise convergence of Fejèr means.* Recall that Fejèr's theorem says that, if

$$\widetilde{x}(t_0) = \lim_{h\to 0} \frac{x(t_0+h)+x(t_0-h)}{2}$$

exists and $x(t) \in L^1(-\pi,\pi)$, then $(\sigma_n x)(t_0) \to \widetilde{x}(t_0)$. A much less restrictive condition on the signal than Fejèr's condition, is to assume that there exists a certain number $\widetilde{x}(t_0)$ such that

$$\lim_{h\to 0} \frac{1}{h} \int_0^h \left| \frac{x(t_0+s)+x(t_0-s)}{2} - \widetilde{x}(t_0) \right| ds = 0 \tag{4}$$

In fact, Fejèr's condition implies (4). Moreover, (4) does not change if we modify the signal $x(t)$ in a set of measure zero. Moreover, (4) holds with $\widetilde{x}(t_0) = x(t_0)$ almost everywhere, for all signals $x(t) \in L^1(-\pi,\pi)$ and, on the other hand, there are examples where Fejèr's condition does not hold at every point (see [14, Vol I, p. 65]).

The following theorem is clearly of interest:

Theorem 52. ***(Lebesgue)*** *If (4) holds then $(\sigma_n x)(t_0) \to \widetilde{x}(t_0)$. In particular, $(\sigma_n x)(t) \to x(t)$ almost everywhere.*

Corollary 53. *If the Fourier series of $x(t) \in L^1(-\pi,\pi)$ converges on a set of positive measure E then its sum coincides with $x(t)$ almost everywhere on E.*

Corollary 54. ***(Unicity theorem for Fourier coefficients)*** *Let $x(t), y(t) \in L^1(-\pi,\pi)$ be such that $\widehat{x}(n) = \widehat{y}(n)$ for all $n \in \mathbb{Z}$. Then $x(t) = y(t)$ almost everywhere.*

3.6.2. *Convergence in norm of Fourier series.* Let $\mathbf{X}$ be a Banach subspace of $L^1(-\pi,\pi)$ which contains all the trigonometric polynomials. Under which conditions we can guarantee that the operators $S_n : \mathbf{X} \to \mathbf{X}$, $x \to S_n x$ (the n-th partial sum of the Fourier series associated to the signal x) satisfy $\lim_{n\to\infty} \|S_n x - x\|_{\mathbf{X}} = 0$? (To simplify assertions, we say that $\mathbf{X}$ admits convergence in norm whenever $\lim_{n\to\infty} \|S_n x - x\|_{\mathbf{X}} = 0$ for all $x \in \mathbf{X}$). This is the problem we will study now.

Let us denote by $L^1(\mathbb{T})$ the space of 2π-periodic signals which belong to $L^1(-\pi,\pi)$, with the norm of $L^1(-\pi,\pi)$ (This allow us to make translations, etc.) The class of Banach subspaces of $L^1(\mathbb{T})$ best adapted to the problem of convergence in norm is the class of homogeneous Banach spaces:

Definition 55. *Let* $\mathbf{X}$ *be a Banach subspace of* $L^1(\mathbb{T})$. *We say that* $\mathbf{B}$ *is homogeneous if the following claims hold:*

i) $\|\bullet\|_{\mathbf{B}} \geq \|\bullet\|_{L^1(\mathbb{T})}$

ii) If $x(t) \in \mathbf{B}$ *and* $\alpha \in \mathbb{R}$ *then* $x_\alpha(t) := x(t-\alpha) \in \mathbf{B}$ *and* $\|x\|_{\mathbf{B}} = \|x_\alpha\|_{\mathbf{B}}$

iii) For all $x(t) \in \mathbf{B}$, *we have that* $\lim_{\alpha \to \alpha_0} \|x_\alpha - x_{\alpha_0}\|_{\mathbf{B}} = 0$.

Some examples of homogeneous Banach spaces are:

- $\mathbf{C}_{2\pi}(\mathbb{R})$, the space of 2π-periodic continuous functions, with the uniform norm.
- $\mathbf{C}_{2\pi}^{(k)}(\mathbb{R})$, the space of 2π-periodic k-times continuously differentiable functions, with the norm: $\|x(t)\|_{\mathbf{C}_{2\pi}^{(k)}} = \sum_{j=0}^{k} \frac{1}{j!} \max_{t\in[-\pi,\pi]} |x^{(k)}(t)|$
- $L^p(\mathbb{T})$, $1 \leq p < \infty$, the space of 2π-periodic functions with finite norm $\|x(t)\|_p = \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |x(t)|^p\, dt\right\}^{\frac{1}{p}}$.

Definition 56. *A summability kernel is a sequence* $\{k_n\}$ *of continuous* 2π*-periodic functions satisfying the following three conditions:*

i)

$$\frac{1}{2\pi}\int_0^{2\pi} k_n(t)dt = 1,\ n = 0, 1, ...$$

ii)

$$\frac{1}{2\pi}\int_0^{2\pi} |k_n(t)|\, dt = \mathbf{O}(1)$$

iii) For all $\delta \in (0, \pi)$,

$$\lim_{n\to\infty} \int_\delta^{2\pi-\delta} |k_n(t)|\, dt = 0$$

With this notation at hand, we can prove the following

Theorem 57. *Let* $\mathbf{B}$ *be an homogeneous Banach space and let* $\{k_n\}$ *be a summability kernel. Then*

$$\lim_{n\to\infty} \left\| x(t) - \frac{1}{2\pi}\int_0^{2\pi} k_n(s)x(t-s)ds \right\|_{\mathbf{B}} = 0$$

In particular,

$$\lim_{n\to\infty} \|x(t) - (\sigma_n x)(t)\|_{\mathbf{B}} = 0.$$

Proof. Clearly, for all $\delta \in [0, \pi)$, we have that

$$\frac{1}{2\pi}\int_0^{2\pi} k_n(s)x(t-s)ds - x(t) = \frac{1}{2\pi}\int_0^{2\pi} k_n(s)(x(t-s)-x(t))ds$$
$$= \frac{1}{2\pi}\int_{-\delta}^{\delta} k_n(s)(x(t-s)-x(t))ds + \frac{1}{2\pi}\int_{\delta}^{2\pi-\delta} k_n(s)(x(t-s)-x(t))ds$$

And

$$\begin{aligned}\left\|\frac{1}{2\pi}\int_{-\delta}^{\delta} k_n(s)(x(t-s)-x(t))ds\right\|_{\mathbf{B}} &\leq \sup_{|h|\leq\delta}\|x_h(t)-x(t)\|_{\mathbf{B}}\,\|k_n(t)\|_{L^1(-\pi,\pi)} \\ &\leq C\sup_{|h|\leq\delta}\|x_h(t)-x(t)\|_{\mathbf{B}}\end{aligned}$$

where $x_h(t) := x(t-h)$. Now, it follows from the homogeneity of $\mathbf{B}$ that, given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\sup_{|h|\leq\delta}\|x_h(t)-x(t)\|_{\mathbf{B}} \leq \frac{\varepsilon}{2C}$$

On the other hand,

$$\begin{aligned}&\left\|\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta} k_n(s)(x(t-s)-x(t))ds\right\|_{\mathbf{B}} \\ \leq\ & \sup_{h\in(\delta,2\pi-\delta)}\|x_h(t)-x(t)\|_{\mathbf{B}}\,\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta}|k_n(s)|\,ds \\ \leq\ & C'\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta}|k_n(s)|\,ds \to 0 \text{ for } n\to\infty.\end{aligned}$$

This ends the proof. □

Corollary 58. *If* $\mathbf{B}$ *is an homogeneous Banach space then the trigonometric polynomials are a dense subset of* $\mathbf{B}$. *In particular, the Weierstrass Approximation theorem holds true. (Also for the simultaneous approximation of functions).*

Corollary 59. *The unicity theorem for Fourier coefficients holds true.*

The problem of convergence in norm of the partial sums of the Fourier series $S_n x$, for $x \in \mathbf{B}$ (an homogeneous Banach space) is very delicate: much more than the corresponding problem for good summability kernels, like Fejèr's kernel. But something can be proved:

Theorem 60. *Let* $\mathbf{B}$ *be an homogeneous Banach space. The following claims are equivalent:*

i) $\mathbf{B}$ *admits convergence in norm.*

ii) The sequence of operators $S_n : \mathbf{B} \to \mathbf{B}$, $n \in \mathbb{N}$ *is uniformly bounded.*

Corollary 61. *There are continuous signals* $x \in \mathbf{C}_{2\pi}(\mathbb{R})$ *such that* $S_n x$ *does not converge uniformly to* x.

We already know that $L^2(\mathbb{T})$ admits convergence in norm (the theorem above gives us a new proof, since $\|S_n\|_{L^2} = 1$ for all n) What about convergence in norm for the spaces $L^p(\mathbb{T})$, $1 \leq p < \infty$, $p \neq 2$? This question, although natural, was not so easy to solve. We start with the answer, which is the following theorem, although we will only say a little bit about the guidelines of the proof.

Theorem 62. *The spaces $L^p(\mathbb{T})$, $1 < p < \infty$, admit convergence in norm although this is not the case for $p = 1$.*

The proof of this theorem is an easy corollary to the following technical results:

Lemma 63. *Let us assume that $\mathbf{B}$ is an homogeneous Banach space such that for all $x(t) \in \mathbf{B}$ and all $n \in \mathbb{N}$, the signal $e^{\mathbf{i}nt}x(t)$ belongs to $\mathbf{B}$ and $\left\|e^{\mathbf{i}nt}x(t)\right\|_{\mathbf{B}} = \|x(t)\|_{\mathbf{B}}$. Then there are equivalent claims:*

i) $\mathbf{B}$ admits convergence in norm.

ii) For all $x(t) \in \mathbf{B}$, there exists a signal $\widetilde{x}(t) \in \mathbf{B}$ such that

$$\widetilde{x}(t) \sim -\sum_{n=-\infty}^{\infty} \widehat{x}(n)\mathbf{sign}(n)e^{\mathbf{i}nt}$$

(we say that $\mathbf{B}$ admits conjugation).

Lemma 64. *The spaces $L^p(\mathbb{T})$, $1 < p < \infty$, admit conjugation although this is not the case for $p = 1$.*

The very difficult thing for the proof of convergence in norm of Fourier series for $L^p(\mathbb{T})$, $1 < p < \infty$, is precisely to prove the Lemma above. This was the starting point for Interpolation Theory!

4. Fourier transform

4.1. **From Fourier series to Fourier Transform: basic considerations.** Fourier series are useful for the representation of periodic signals but unfortunately these signals are not so frequent in practice as non-periodic signals. Thus a new (more ambitious) mathematical tool is needed. To this aim, we devote this section.

Let $x(t)$ be a nonperiodic signal defined on all the real line and let us denote by $x_T(t)$ ($T > 0$) the $2T$-periodic signal which is obtained from $x(t)$ by seting $x_T(t) = x(t)$ for $t \in (-T, T]$ and extending it $2T$-periodically. Under the aditional assumption that $x(t)$

is smooth enough (e.g., $x \in \mathbf{C}^1(\mathbb{R})$), we have

$$x(t) = x_T(t) = \frac{1}{2T}\sum_{k=-\infty}^{\infty}\left[\int_{-T}^{T} x(s)e^{-(\pi\mathbf{i}/T)ks}ds\right]e^{(\pi\mathbf{i}/T)kt}, \text{ for all } t \in (-T,T] \tag{5}$$

We would like to make $T \to \infty$ in the equality above. Of course, for any fixed value $t \in \mathbb{R}$, we will have $x(t) = x_T(t)$ for all T big enough, so that, if the limit exists for the righthand side of the equality, it will be equal to the former signal $x(t)$.

Now, what's happening with the righthand side of (5), for $T \to \infty$? Let us set $\Delta f = 1/(2T)$ and $f_k = k\Delta f$. Then we can write (5) as follows:

$$x(t) = \sum_{k=-\infty}^{\infty} \Delta f\left[\int_{-T}^{T} x(s)e^{-2\pi\mathbf{i}f_k s}ds\right]e^{2\pi\mathbf{i}f_k t}, \text{ for all } t \in (-T,T]$$

Now, $|f_{k+1} - f_k| = \Delta f = 1/2T$ ($k \in \mathbb{Z}$) converges to zero for $T \to \infty$. Thus, we can think about $\{f_k\}_{k\in\mathbb{Z}}$ as a partition of those used for the definition of Riemann's integral, so that the righthandside of the formula above should approach to the (improper) Riemann integral

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x(s)e^{-2\pi\mathbf{i}fs}ds\right)e^{2\pi\mathbf{i}ft}df$$

Hence, under certain reasonable hyphotheses about the smoothness of the nonperiodic signal $x(t)$, we conclude that the following identity holds:

$$x(t) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x(s)e^{-2\pi\mathbf{i}fs}ds\right)e^{2\pi\mathbf{i}ft}df$$

(this identity is usually called the Fourier integral Theorem). Now, making the change of variable $\xi = 2\pi f$, we obtain that:

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x(s)e^{-\mathbf{i}\xi s}ds\right)e^{\mathbf{i}\xi t}d\xi$$

Definition 65. *We say that*

$$\mathcal{F}(x)(\xi) := \widehat{x}(\xi) := \int_{-\infty}^{\infty} x(s)e^{-\mathbf{i}\xi s}ds$$

is the Fourier transform of the (nonperiodic) signal $x(t)$.

Definition 66. *We say that*

$$\mathcal{F}^{-1}(y)(t) := \frac{1}{2\pi}\int_{-\infty}^{\infty} y(\xi)e^{\mathbf{i}\xi t}d\xi$$

is the inverse Fourier transform of the (nonperiodic) signal $y(\xi)$.

Remark 5. *Of course, the Fourier integral theorem claims that*

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x(s)e^{-\mathbf{i}\xi s}ds\right)e^{\mathbf{i}\xi t}d\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{F}(x)(\xi)e^{\mathbf{i}\xi t}d\xi = \mathcal{F}^{-1}(\mathcal{F}(x))(t)$$

and, by analogous arguments, that

$$\mathcal{F}(\mathcal{F}^{-1}(y))(\xi) = y(\xi).$$

This means that the names 'Fourier transform' and 'Inverse Fourier transform' are coherent.

At least the following is true: under certain hyphotheses (that will be soon particularized), the knowledge of the Fourier transform of a signal is equivalent to the knowledge of the signal. This is quite similar to the results we proved for Fourier series: that (under certain aditional conditions on the signal $x(t)$) if we know the Fourier coefficients of the signal, then we know the signal, since it can be recuperated by summing its Fourier series. Thus, the role played by Fourier series in the study of periodic signals is now played by Fourier transforms for the study of nonperiodic signals.

To avoid problems with the definition of Fourier transform, we will assume that our signal is absolutely integrable on the real line (i.e., we assume that $||x||_{L^1(\mathbb{R})} = \int_{-\infty}^{\infty}|x(t)|dt < \infty$). This clearly implies that its Fourier transform $\widehat{x}(\xi)$ exists and is uniformly bounded on the real line, since $|e^{-2\pi\mathbf{i}\xi s}| = 1$ for all $\xi, s \in \mathbb{R}$ implies that $|\widehat{x}(\xi)| \leq ||x||_{L^1(\mathbb{R})}$ for all $\xi \in \mathbb{R}$. In fact, we can say more about $\widehat{x}(\xi)$, since the Riemann-Lebesgue theorem can be rephrased as follows:

Theorem 67. *If $x \in L^1(\mathbb{R})$ then $\lim_{\xi\to\pm\infty}\mathcal{F}(x)(\xi) = 0$.*

4.2. **Basic properties of the Fourier transform.** The word 'transform' is an advice that we are using a certain tool to transform a problem into another. This is precisely the case when we deal with Fourier transforms. For example, it is well known that they are an useful tool for the study of some special (but very important, from the physical point of view) type of differential equations. Many of these applications are possible thanks to the operational properties of the transform. In this section we state several of these basic properties.

Theorem 68. *Let us assume that $x \in L^1(\mathbb{R})$. The following holds true:*

a) For all $a \in \mathbb{R}$, we have that

$$\mathcal{F}[x(t-a)](\xi) = e^{-\mathbf{i}a\xi}\mathcal{F}[x(t)](\xi) \text{ and } \mathcal{F}[e^{\mathbf{i}a\xi}x(t)](\xi) = \mathcal{F}[x(t)](\xi - a)$$

b) If $\delta > 0$ and $x_\delta(t) = \delta^{-1}x(t/\delta)$, then

$$\mathcal{F}[x_\delta](\xi) = \mathcal{F}[x](\delta\xi) \text{ and } \mathcal{F}[x(\delta t)](\delta\xi) = \mathcal{F}(x)_\delta(\xi)$$

c) If x is continuous and piecewise derivable, with $x'(t) \in L^1(\mathbb{R})$, then

$$\mathcal{F}(x')(\xi) = \mathbf{i}\xi\mathcal{F}(x)(\xi).$$

Moreover, if $tx(t)$ is absolutely integrable then

$$\mathcal{F}(tx(t))(\xi) = \mathbf{i}\mathcal{F}(x)'(\xi).$$

d) The identity

$$\mathcal{F}((x * y)(t))(\xi) = \mathcal{F}(x)(\xi)\mathcal{F}(y)(\xi).$$

holds for all $y \in L^1(\mathbb{R})$.

Example 69. *Let us consider the step signal*

$$u_a(t) = \begin{cases} 1 & \text{if } |t| < a \\ 0 & \text{otherwise} \end{cases}$$

Its Fourier transform is:

$$\begin{aligned} \mathcal{F}(u_a)(\xi) &= \int_{-\infty}^{\infty} u_a(s)e^{-\mathbf{i}\xi s}ds \\ &= \int_{-a}^{a} e^{-2\pi\mathbf{i}\xi s}ds = \frac{e^{-\mathbf{i}\xi s}}{-\mathbf{i}\xi}\Big]_{s=-a}^{s=a} = \frac{e^{-\mathbf{i}a\xi} - e^{\mathbf{i}a\xi}}{-\mathbf{i}\xi} \\ &= \frac{\cos(-a\xi) + \mathbf{i}\sin(-a\xi) - \cos(a\xi) - \mathbf{i}\sin(a\xi)}{-\mathbf{i}\xi} \\ &= 2\frac{\sin(a\xi)}{\xi} \end{aligned}$$

4.3. **Fourier inversion formula.** We have already suggested, by the use of certain heuristic arguments, that (under certain reasonable assumptions on the smoothness of the signal $x(t)$) the identity

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{F}(x)(\xi)e^{\mathbf{i}\xi t}d\xi.$$

holds. In this section we will try to put the formula above in context. Under which assumptions (and in which sense) is it true?. Of course, likewise the study of convergence of Fourier series, there exists no easy answer to our question. Indeed, it does not exists a unique answer to the question. For example, we can consider the improper integral just as its principal value i.e. the limit

$$\mathbf{p.v.}\int_{-\infty}^{\infty}\mathcal{F}(x)(\xi)e^{\mathbf{i}\xi t}d\xi=\lim_{M\to\infty}\int_{-M}^{M}\mathcal{F}(x)(\xi)e^{\mathbf{i}\xi t}d\xi,$$

but we can consider also this integral in the proper sense. This should limit the validity of the Fourier inversion formula, since principal values exists also for divergent integrals. Another problem appears if $\mathcal{F}(x)=\widehat{x}\notin L^1(\mathbb{R})$, which implies that $\int_{-\infty}^{\infty}\mathcal{F}(x)(\xi)e^{\mathbf{i}\xi t}d\xi$ diverges. For example, this happens for the pulse signal u_a, since (as we already proved) $\mathcal{F}(u_a)(\xi)=2\frac{\sin(a\xi)}{\xi}$. On the other hand, also for signals $\widehat{x}\in L^1(\mathbb{R})$ we can have certain problems to check the formula

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}x(s)e^{-\mathbf{i}\xi s}ds\right)e^{\mathbf{i}\xi t}d\xi$$

since some operations (like the interchange of the integrals in the righthand side) can not be made without some extra assumptions on the signal $x(s)$.

Anyway, we can try the following:

- Firstly we multiply $\widehat{x}(\xi)$ by a signal $\Phi_\varepsilon(\xi)$ such that it is fast decreasing for $\xi\to\pm\infty$ and converges uniformly to 1 for $\varepsilon\to 0$. Then, fixed $\varepsilon>0$ we can make all the operations we need (interchange the order of integration, etc.), since the double integral is absolutely convergent.
- Secondly, we try to get our result by seting $\varepsilon\to 0$.

So, let us take $\Phi_\varepsilon(\xi)=e^{-\varepsilon^2\xi^2/2}$, and consider the integral:

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}x(s)e^{-\mathbf{i}\xi s}ds\right)\Phi_\varepsilon(\xi)e^{\mathbf{i}\xi t}d\xi\\
=\;&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}x(s)e^{-\mathbf{i}\xi s}ds\right)e^{-\varepsilon^2\xi^2/2}e^{\mathbf{i}\xi t}d\xi\\
=\;&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}x(s)e^{\mathbf{i}\xi(t-s)}e^{-\varepsilon^2\xi^2/2}d\xi\right)ds\\
=\;&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}e^{\mathbf{i}\xi(t-s)}e^{-\varepsilon^2\xi^2/2}d\xi\right)x(s)ds\\
=\;&\frac{1}{2\pi}\int_{-\infty}^{\infty}x(s)\mathcal{F}\left(e^{-\varepsilon^2\xi^2/2}\right)(t-s)ds\\
=\;&\frac{1}{2\pi}\frac{\sqrt{2\pi}}{\varepsilon}\int_{-\infty}^{\infty}x(s)e^{-(t-s)^2/(2\varepsilon^2)}ds\\
=\;&(x*\phi_\varepsilon)(t),
\end{aligned}$$

where

$$\phi_\varepsilon(t) = \frac{1}{\sqrt{2\pi}\varepsilon} e^{-\frac{t^2}{2\varepsilon^2}}.$$

The following technical Lemma holds true:

Lemma 70. *Let us assume that $x(t)$ is piecewise continuous. The relation*

$$\lim_{\varepsilon \to 0} (x * \phi_\varepsilon)(t) = \frac{x(t+) + x(t-)}{2}$$

holds true for all $t \in \mathbb{R}$.

Proof. See [4]. □

Theorem 71. ***(Integral Fourier Theorem)*** *Let us assume that $x(t) \in L^1(\mathbb{R})$ is piecewise continuous and $x(t) = (x(t+) + x(t-))/2$ for all $t \in \mathbb{R}$. Then*

$$x(t) = \lim_{\varepsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\xi t} \widehat{x}(\xi) e^{-\varepsilon^2 \xi^2/2} d\xi \ \ (t \in \mathbb{R})$$

Moreover, if $\widehat{x}(\xi) \in L^1(\mathbb{R})$, then x is continuous and

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\xi t} \widehat{x}(\xi) d\xi \ \ (t \in \mathbb{R})$$

There are many signals $x(t)$ such that both x and $\widehat{x}$ are absolutely integrable functions. In fact, to guarantee that the Fourier transform of $x \in L^1(\mathbb{R})$ belongs to $L^1(\mathbb{R})$ it will be enough to assume some smoothness on the signal x. This is the content of the following result:

Proposition 72. *Let us assume that $x \in L^1(\mathbb{R})$ is twice derivable and $x', x'' \in L^1(\mathbb{R})$. Then $\widehat{x} \in L^1(\mathbb{R})$.*

Proof. $x'' \in L^1(\mathbb{R})$ implies that $\mathcal{F}(x'') \in L^\infty(\mathbb{R})$. On the other hand, $\mathcal{F}(x'') = -\xi^2 \mathcal{F}(x)$. Hence there exists a constant $C < \infty$ such that $|\mathcal{F}(x)(\xi)| \leq C/(1+\xi^2)$. This proves that $\mathcal{F}(x) = \widehat{x} \in L^1(\mathbb{R})$. □

There are several forms of the Fourier integral Theorem. For example, if we use the method we have just explained but we change the function $\Phi_\varepsilon(\xi) = e^{-\varepsilon^2\xi^2/2}$ by the characteristic function $\chi_{[-1/\varepsilon, 1/\varepsilon]}$ then we will change the limit $\lim_{\varepsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\xi t} \widehat{x}(\xi) e^{-\varepsilon^2\xi^2/2} d\xi$

in the theorem above by the principal value of $\int_{-\infty}^{\infty} e^{\mathbf{i}\xi t}\widehat{x}(\xi)d\xi$,

$$\begin{aligned}\mathbf{v.p}\left(\int_{-\infty}^{\infty} e^{\mathbf{i}\xi t}\widehat{x}(\xi)d\xi\right) &= \lim_{r\to\infty}\int_{-r}^{r} e^{\mathbf{i}\xi t}\widehat{x}(\xi)d\xi \\ &= \lim_{\varepsilon\to 0}\int_{-1/\varepsilon}^{1/\varepsilon} e^{\mathbf{i}\xi t}\widehat{x}(\xi)d\xi \\ &= \lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} e^{\mathbf{i}\xi t}\widehat{x}(\xi)\chi_{[-1/\varepsilon,1/\varepsilon]}(\xi)d\xi\end{aligned}$$

With this notation at hand, we canit is possible to prove the following theorem (see [4]):

Theorem 73. ***(Fourier Integral Theorem, II)*** *Let us assume that $x(t)$ is absolutely integrable, piecewise differentiable on $\mathbb{R}$, and $x(t) = (x(t+)+x(t-))/2$ for all $t \in \mathbb{R}$. Then*

$$x(t) = \mathbf{v.p}\left(\int_{-\infty}^{\infty} e^{\mathbf{i}\xi t}\widehat{x}(\xi)d\xi\right) \quad (t \in \mathbb{R}).$$

4.4. **Fourier transform of functions from $L^2(\mathbb{R})$.** It may happen that the Fourier transform of an energy signal $x(t) \in L^2(\mathbb{R})$ does not exists, since the integral $\int_{-\infty}^{\infty} x(t)e^{\mathbf{i}\xi t}d\xi$ may be divergent if $x(t) \notin L^1(\mathbb{R})$. On the other hand, it should be nice to have an (adecuate) extension of the Fourier transform to the setting of energy signals (recall that $L^2(\mathbb{R})$ is the analog substitute of $l^2(\mathbb{Z})$). Thus, motivated by the very nice properties of Fourier series in connection with the space $l^2(\mathbb{Z})$, we may guess that some similar thing is perhaps possible for the Fourier Transform. In this section we prove that such an extension exists.

Lemma 74. *Let us assume that $x(t), y(t) \in L^1(\mathbb{R})$ and $\widehat{x}(t), \widehat{y}(t) \in L^1(\mathbb{R})$. Then $x(t)$, $y(t)$, $\widehat{x}(t)$, $\widehat{y}(t) \in L^2(\mathbb{R})$ and*

$$(x,y)_{L^2(\mathbb{R})} = \frac{1}{2\pi}(\widehat{x},\widehat{y})_{L^2(\mathbb{R})}$$

In particular, if $x(t) = y(t)$, then we have the following 'Parseval's formula':

$$||x||_{L^2(\mathbb{R})} = \frac{1}{2\pi}||\widehat{x}||_{L^2(\mathbb{R})}$$

Proof. We will only prove that $(x,y)_{L^2(\mathbb{R})} = \frac{1}{2\pi}(\widehat{x},\widehat{y})_{L^2(\mathbb{R})}$. Thus, we compute as follows:

$$\begin{aligned}
(x,y)_{L^2(\mathbb{R})} &= \int_{-\infty}^{\infty} x(t)\overline{y(t)}dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)\left(\overline{\int_{-\infty}^{\infty}\widehat{y}(\xi)e^{\mathbf{i}\xi t}d\xi}\right)dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\overline{\widehat{y}(\xi)}x(t)e^{-\mathbf{i}\xi t}dt\right)d\xi \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{\widehat{y}(\xi)}\left(\int_{-\infty}^{\infty}x(t)e^{-\mathbf{i}\xi t}dt\right)d\xi \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{\widehat{y}(\xi)}\widehat{x}(\xi)d\xi = \frac{1}{2\pi}(\widehat{x},\widehat{y})_{L^2(\mathbb{R})}
\end{aligned}$$

(Note that we have used the Fourier integral theorem). This ends the proof. □

For any $x(t) \in L^2(\mathbb{R})$ it is possible to find a sequence of signals $\{x_n(t)\}_{n=0}^{\infty}$ such that $x_n(t), \widehat{x}_n(\xi) \in L^1(\mathbb{R})$ for all $n \geq 0$ and $\lim_{n\to\infty}||x-x_n||_{L^2(\mathbb{R})} = 0$. It follows (use the Lemma above) that $\{\widehat{x}_n(\xi)\}_{n=0}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{R})$, since $\{x_n(t)\}_{n=0}^{\infty}$ is convergent (hence it is also a Cauchy sequence) and $||\widehat{x}_n(\xi) - \widehat{x}_m(\xi)||_{L^2(\mathbb{R})} = 2\pi||x_n(t) - x_m(t)||_{L^2(\mathbb{R})}$ for all $n, m \geq 0$. It follows that exists a signal $z(\xi) \in L^2(\mathbb{R})$ such that

$$\lim_{n\to\infty} ||\widehat{x}_n(\xi) - z(\xi)||_{L^2(\mathbb{R})} = 0.$$

Let us prove that $z(\xi)$ does not depend on the approximating sequence $\{x_n(t)\}_{n=0}^{\infty}$: if $\{y_n(t)\}_{n=0}^{\infty}$ is another sequence of signals such that $y_n(t), \widehat{y}_n(\xi) \in L^1(\mathbb{R})$ and $\lim_{n\to\infty}||x - y_n||_{L^2(\mathbb{R})} = 0$, then we should have

$$\lim_{n\to\infty} ||\widehat{x}_n(\xi) - \widehat{y}_n(\xi)||_{L^2(\mathbb{R})} = 2\pi \lim_{n\to\infty} ||x_n(t) - y_n(t)||_{L^2(\mathbb{R})} = 0$$

Thus, $\lim_{n\to\infty}||\widehat{x}_n(\xi) - z(\xi)||_{L^2(\mathbb{R})} = 0$. This proves that $z(\xi)$ only depends on the former signal $x(t) \in L^2(\mathbb{R})$.

Definition 75. *We call the Fourier transform of $x(t) \in L^2(\mathbb{R})$ to the unique signal $\widehat{x}(\xi) \in L^2(\mathbb{R})$ that is an $L^2(\mathbb{R})$-limit of a sequence of Fourier transforms $\{\widehat{x}_n(\xi)\}_{n=0}^{\infty}$, where $\{x_n(t)\}_{n=0}^{\infty} \subset L^2(\mathbb{R})$ satisfies $x_n(t), \widehat{x}_n(\xi) \in L^1(\mathbb{R})$ for all $n \geq 0$ and $\lim_{n\to\infty}||x - x_n||_{L^2(\mathbb{R})} = 0$.*

He have already proved the following result:

Theorem 76. ***(Plancherel)*** *There exists a (unique) extension of the Fourier transform which maps $L^2(\mathbb{R})$ into $L^2(\mathbb{R})$, $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ ($\mathcal{F}(x(t)) = \widehat{x}(\xi)$). Moreover, the* *relation*

$$(x,y)_{L^2(\mathbb{R})} = \frac{1}{2\pi}(\widehat{x},\widehat{y})_{L^2(\mathbb{R})} \ \ y \ \ ||x||_{L^2(\mathbb{R})} = \frac{1}{2\pi}||\widehat{x}||_{L^2(\mathbb{R})},$$

holds for all $x, y \in L^2(\mathbb{R})$.

Remark 6. *The new Fourier transform* $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ *has the same algebraic properties than the old one.*

4.5. **Fourier transform for Generalized Functions.** We already know that many elementary operations on signals (e.g., shifts, etc) can be re-defined for the context of generalized functions. So, it is natural to ask: what about the Fourier transform?. In this section we introduce the Fourier transform of generalized functions.

Definition 77. *Let* $x \in \mathbb{G}$ *be a generalized function. We define its Fourier transform through its action on the class of Schwartz functions,*

$$\mathcal{F}(x)\{\phi\} = \int_{-\infty}^{\infty} x(s)\mathcal{F}(\phi)(s)ds$$

(i.e., we set $\mathcal{F}(x)\{\phi\} = x\{\mathcal{F}(\phi)\}$ *for all* $\phi \in \mathbb{S}$*, where the Fourier transform in the righthand side of the equality, is the classical).*

Remark 7. *That* $\mathcal{F}(x)\{\phi\} = \int_{-\infty}^{\infty} x(s)\mathcal{F}(\phi)(s)ds$ *defines a generalized function (for each* $x \in \mathbb{G}$*), needs a proof! (i.e., we must prove that for each* $x \in \mathbb{G}$*, there exists a signal* $z(t) \in$ **CSG** *and a natural number* $n \in \mathbb{N}$ *such that*

$$x\{\mathcal{F}(\phi)\} = (-1)^n \int_{-\infty}^{\infty} z(s)\phi^{(n)}(s)ds$$

holds true for all $\phi \in \mathbb{S}$*). And this is not completely obvious...(exercise!).*

Example 78. *(Fourier transform of Dirac's delta function). Let us take* $x = \delta$*, the Dirac's delta function. Then we have that*

$$\mathcal{F}(\delta)\{\phi\} = \delta\{\mathcal{F}(\phi)\} = \mathcal{F}(\phi)(0) = \int_{-\infty}^{\infty} \phi(s)ds = 1\{\phi\}$$

holds for all $\phi \in \mathbb{S}$ *. Hence,* $\mathcal{F}(\delta) = 1$.

This is an interesting example, since it shows a generalized (not ordinary) function such that its Fourier transform is an ordinary function.

Exercise: Prove that the comb signal **III** satisfies $\mathcal{F}(\mathbf{III}) = \mathbf{III}$.

5. Aplications

5.1. **Fourier transform and filters: ideal filters.** We already know that the class of convolution filters is very large and, in fact, represents the majority of the useful filters that exists. So, filters are represented in the time domain, by the expresion $Lx = x * h$, where $h = L\delta$. If we apply the Fourier transform to both sides of $Lx = x * h$, we obtain that filters can be represented in the frequency domain, by the formula:

$$\widehat{y} = \mathcal{F}(Lx) = \mathcal{F}(x)\mathcal{F}(h) = \widehat{x} \cdot \widehat{h}$$

We use in this context the folllowing (standard) notation: $X(w) = \widehat{x}(w)$ is the frequency input of the system, $Y(w) = \widehat{y}(w)$ is the (frequency) output of the system, and $H(w) = \widehat{h}(w)$ is called the transfer function of the system. Of course, the transfer function $H(w)$ completely characterizes the system, since the formula

$$Y(w) = X(w)H(w), \tag{6}$$

holds for arbitrary inputs $X(w)$. This has some advantages in comparison with the characterization of filters in the time domain, since multiplication is simpler than convolution. Moreover, formula (6) is also useful for the design of filters, as we will see in this section.

Let us consider the truncated exponential function,

$$x_T(t) = e^{iwt}\chi_{[-T,T]}(t) = \begin{cases} e^{\mathbf{i}wt} & |t| \leq T \\ 0 & \text{otherwise} \end{cases}$$

and let L be a filter with impulse response $(L\delta)(t) = h(t)$. Then

$$\begin{aligned} (Lx_T)(t) &= (x_T * h)(t) = \int_{-\infty}^{\infty} x_T(t-s)h(s)ds \\ &= \int_{T-t}^{T+t} e^{\mathbf{i}w(t-s)}h(s)ds \\ &= e^{iwt}\int_{T-t}^{T+t} e^{-\mathbf{i}ws}h(s)ds \end{aligned}$$

Let us now set $T \to \infty$ in both sides of the equality. Then we obtain that

$$L(e^{iwt})(t) = \widehat{h}(w)e^{iwt}.$$

This means that e^{iwt} is an eigenvector of L associated to the eigenvalue $H(w) = \widehat{h}(w)$. Hence the following concept is senseful:

Definition 79. *Let $H(w) = A(w)e^{-\mathbf{i}\Phi(w)}$ be the transfer function of the filter L, where we assume that $A(w) \geq 0$ and $\Phi(w) \in \mathbb{R}$ for all w. We say that $A(w)$ is the amplitude frequency spectrum and $\Phi(w)$.is the phase frequency spectrum of the system. We say that* *the system does not produce amplitude distorsion if $A(w) = A_0 > 0$ is constant. We say that L is an ideal filter if it does not produce phase distorsion (Of course, we allow a phase shift, which is obtained whenever $\Phi(w) = wt_0$ is a linear function) . Hence, the filter is ideal if and only if $\Phi(w) = wt_0$ for a certain constant t_0.*

An easy property of the amplitude frequency and phase frequency spectrum of a filter L, is the following:

Theorem 80. *Let us assume that $h(t) = (L\delta)(t)$, the impulse response of the filter L is a real signal. Then $A(w)$ is an even function and $\Phi(w)$ is an odd function.*

Proof. It follows from $H(w) = A(w)e^{-\mathbf{i}\Phi(w)} = \widehat{h}(w)$, that

$$\overline{H(w)} = \overline{\int_{-\infty}^{\infty} e^{-\mathbf{i}ws}h(s)ds} = \int_{-\infty}^{\infty} e^{\mathbf{i}ws}h(s)ds = H(-w).$$

Hence

$$\overline{A(w)e^{-\mathbf{i}\Phi(w)}} = A(w)e^{\mathbf{i}\Phi(w)} = \overline{H(w)} = H(-w) = A(-w)e^{-\mathbf{i}\Phi(-w)},$$

and

$$\begin{cases} A(w) = A(-w) & \text{(i.e., } A \text{ is even)} \\ \Phi(-w) = -\Phi(w) & \text{(i.e., } \Phi \text{ is odd)} \end{cases}$$

This ends the proof. □

Now we describe some typical ideal filters:

5.1.1. *Low-pass ideal filters.* For a low-pass ideal filter, we are interested in the elimination of the influence of all the frequency components of the signal with a frequency amplitude greater than a fixed amplitude $w_0 > 0$. On the other hand we want not to disturb the frequencies w with $|w| \le w_0$. Let us denote by H_{w_0} the transfer function of the filter. Taking in consideration that $Y(w) = X(w)H_{w_0}(w)$ describes the system and $\Phi(w) = wt_0$ for a certain fixed t_0 (since the filter is ideal) we conclude that

$$H_{w_0}(w) = \begin{cases} e^{-\mathbf{i}wt_0} & |w| \le w_0 \\ 0 & \text{otherwise} \end{cases}.$$

Now, we use that $H_{w_0}(w) = \widehat{h_{w_0}}(w)$, and the Fourier integral theorem to compute the impulse response of the system,

$$\begin{aligned} h_{w_0}(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H_{w_0}(w)e^{\mathbf{i}wt}dw = \frac{1}{2\pi}\int_{-w_0}^{w_0} e^{\mathbf{i}w(t-t_0)}dw \\ &= \frac{\sin(w_0(t-t_0))}{\pi(t-t_0)} \end{aligned}$$

Hence, the filter can be described in the time domain by the formula:

$$L_{w_0}(x)(t) = (x * h_{w_0})(t) = \int_{-\infty}^{\infty} x(s)\frac{\sin(w_0(t-s-t_0))}{\pi(t-s-t_0)}ds$$

The word **ideal** reflects perfectly what's happening: the system is physically unrealizable, since the impulse response is not causal!.

5.1.2. *Ideal all-pass filters.* For an ideal all-pass filter, we do not allow any distorsion in both the amplitude frequency spectrum and the phase frequency spectrum. This means that we may consider our filter L as the limit situation we obtain by seting $w_0 \to \infty$ for the low-pass filter L_{w_0}above. Thus,

$$\begin{aligned}(Lx)(t) &= \lim_{w_0\to\infty}\int_{-\infty}^{\infty} x(s)\frac{\sin(w_0(t-s-t_0))}{\pi(t-s-t_0)}ds \\ &= \int_{-\infty}^{\infty} x(s)\delta(t-t_0-s)ds = x(t-t_0)\end{aligned}$$

Which proves that the only ideal all-pass filters that exists are the time-translations.

5.1.3. *Ideal high-pass filters.* Let us denote by S_{w_0} an ideal filter which eliminates the frequency components of a signal which satisfy $|w| < w_0$ for a certain w_0, but does not produce distorsion on the components of frequency $|w| \geq w_0$ (we say that S_{w_0} is a high-pass filter) and let L_{w_0} be the ideal low-pass associated with w_0. Then $S_{w_0} + L_{w_0}$ must be an ideal all-pass filter. Hence $(S_{w_0}x)(t) = x(t-t_0) - (L_{w_0}x)(t)$ for a certain t_0.

5.1.4. *Ideal narrow-band filters.* It is very usual that we want to eliminate all frequency components of a signal, except those which belong to a certain small neighbourhood of a fixed frequency $w_c \neq 0$. So, we fix a certain small value $w_0 << w_c$ and we say that a frequency w is near to w_c if $|w-w_c| < w_0$. Of course, we must take in consideration that the amplitude spectrum of our system is an even function. Hence the behaviour near $-w_c$ is fixed by its behaviour near w_c and the transfer function of the system must be of the form

$$H_{w_c,w_0}(w) = \begin{cases} e^{-\mathbf{i}wt_0} & \text{if } \max\{|w-w_c|, |w+w_c|\} \leq w_0 \\ 0 & \text{otherwise}\end{cases}$$

Hence, $h_{w_c,w_0}(t)=\mathcal{F}^{-1}(H_{w_c,w_0})$ is given by

$$\begin{aligned}
h_{w_c,w_0}(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H_{w_c,w_0}(w)e^{\mathbf{i}wt}dw \\
&= \frac{1}{2\pi}\int_{-w_c-w_0}^{-w_c+w_0} e^{\mathbf{i}w(t-t_0)}dw + \frac{1}{2\pi}\int_{w_c-w_0}^{w_c+w_0} e^{\mathbf{i}w(t-t_0)}dw \\
&= \frac{1}{2\pi\mathbf{i}(t-t_0)}\left\{e^{-\mathbf{i}(w_c-w_0)(t-t_0)} - e^{-\mathbf{i}(w_c+w_0)(t-t_0)}\right\} \\
&\quad + \frac{1}{2\pi\mathbf{i}(t-t_0)}\left\{e^{\mathbf{i}(w_c+w_0)(t-t_0)} - e^{\mathbf{i}(w_c-w_0)(t-t_0)}\right\} \\
&= \frac{1}{\pi(t-t_0)}\left\{\sin((w_c+w_0)(t-t_0)) - \sin((w_c-w_0)(t-t_0))\right\} \\
&= \frac{1}{\pi(t-t_0)}\sin(w_c(t-t_0))\cos(w_c(t-t_0)).
\end{aligned}$$

5.1.5. *Real Filters.* Ideal filters are physically unrealizable, since they imply noncausal filters. For a low-pass real filter, we would like to have a transfer function as near as possible to the ideal low-pass transfer function $H_{w_0}(w)$, but we also need two more things: i) The impulse response of the associated filter, must be causal and ii) We should be able to construct a mechanism (that's the real filter) for which the mathematical filter we have at hands is a good modell of the physical phenomenon. We only mention here two examples of real low-pass filters:

Example 81. *The second order butterworth filter has transfer function* $H(w)=\frac{1}{1-w^2LC+\mathbf{i}wL/R}$ *Hence its magnitude frequency response is given by:*

$$\begin{aligned}
|H(w)| &= \frac{\frac{1}{LC}}{\sqrt{w^4+(1/LC)^2+w^2\left[\frac{1}{(RC)^2}-\frac{2}{LC}\right]}} \\
&= \frac{w_c^2}{\sqrt{w^4+w_c^4}}, \text{ where } w_c^2=\frac{1}{LC}
\end{aligned}$$

Example 82. *The amplitude frequency response of the n-th Chebyshev filter, is given by:*

$$|H_n(w)|=\frac{1}{\sqrt{1+\epsilon^2\cos^2(n\arccos\frac{w}{w_c})}},$$

where $\epsilon\in(0,1)$ *is constant.*

5.2. **The sampling theorem.** Let us assume that we are sampling the analog signal $x(t)$ by the meassure of its values at the discrete set of time values $\{t_1 < t_2 < \cdots < t_m < ...\}$. Is it the information we get by this way really useful for the description of the signal $x(t)$? Clearly, the answer is: no (at least for arbitrary signals), since a discrete set of values $\{x(t_1), ..., x(t_m), ...\}$ does not determine any specific analog signal $x(t)$. We can interpolate the set of values with infinitely many distinct signals $x(t)$. Indeed, we can make this with very smooth signals. But: be careful!: if the sequence $\{t_1 < t_2 < \cdots < t_m < ...\}$ has some accumulation point, we can interpolate the data set with at most one analytic signal! (this is the identity principle, from an arbitrary introductory course on complex analysis). Hence, under certain aditional hyphotheses on the signal $x(t)$ it is sometimes possible to completely determine it by sampling.

The most famous theorem in this subject was proved by Whittaker in 1935 and rediscovered by Shannon in 1949. Shannon used the result for applications to communications theory. It works for band-limited signals:

Definition 83. *We say that $x(t) \in L^2(\mathbb{R})$ is a band-limited signal if its Fourier transform $\widehat{x}(\xi)$ has compact support. This means that there exists a certain $\Omega > 0$ such that $\widehat{x}(\xi) = 0$ for all $|\xi| > \Omega$.*

There exists an intuitive motivation for the theorem: if our signal is band-limited with $\mathbf{supp}(\widehat{\mathbf{x}}(\xi)) \subset [-\Omega, \Omega]$, then if we take in consideration the Fourier integral theorem and the fact that the exponential function $e^{\mathbf{i}\xi t}$ has small variation on intervals of small length (e.g., $\Delta t << \Omega^{-1}$), then our signal $x(t)$ should not change very much its behaviour between the values t and $t + \Delta t$. Of course, this intuitive argument needs a formal proof (and a formal statement). That's our next objective:

Theorem 84. ***(Sampling Theorem)*** *Let us assume that $x(t) \in L^2(\mathbb{R})$ and $\widehat{x}(\xi) = 0$ for all $|\xi| > \Omega$. Then $x(t)$ can be recovered from its values at the discrete set of points $\{t_n = n\pi/\Omega\}_{n\in\mathbb{Z}}$. Indeed, the following formula holds:*

$$x(t) = \sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})\frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}.$$

Proof. Clearly, $x(t) \in L^2(\mathbb{R})$ implies that $\widehat{x}(\xi) \in L^2(\mathbb{R})$ and $\widehat{x}(t) \in L^1(\mathbb{R})$, since it has compact support. Hence $\widehat{x}(t) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and we can use the Fourier integral theorem. On the other hand, we can take an L^2-convergent fourier series for $\widehat{x}(\xi)$ on the interval $[-\Omega, \Omega]$, since the complex exponential signals $\{e^{\mathbf{i}n\pi\xi/\Omega}\}_{n\in\mathbb{Z}}$ are an orthogonal basis of $L^2(-\Omega, \Omega)$. Concretely, the corresponding Fourier coefficients of $\widehat{x}(\xi)$ are given by

$$c_n = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} \widehat{x}(\xi)e^{-\mathbf{i}n\pi\xi/\Omega}d\xi$$

The re-write the formula above as follows:

$$c_{-n} = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} \widehat{x}(\xi)e^{\mathbf{i}n\pi\xi/\Omega}d\xi$$

We can substitute the integration interval $[-\Omega,\Omega]$ in the expression above by the real line, since $\widehat{x}(\xi)=0$ for all $|\xi|>\Omega$. Hence

$$c_{-n} = \frac{1}{2\Omega}\int_{-\infty}^{\infty} \widehat{x}(\xi)e^{\mathbf{i}n\pi\xi/\Omega}d\xi = \frac{\pi}{\Omega}\frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{x}(\xi)e^{\mathbf{i}\xi\frac{n\pi}{\Omega}}d\xi = \frac{\pi}{\Omega}x(\frac{n\pi}{\Omega}),\ n\in\mathbb{Z}.$$

(For the last equality, we used the Fourier integral theorem). This means that

$$\widehat{x}(\xi) = \frac{\pi}{\Omega}\sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})e^{-\mathbf{i}\xi\frac{n\pi}{\Omega}}.$$

We use once more the Fourier integral theorem and that $\widehat{x}(\xi)$ vanishes outside $[-\Omega,\Omega]$, and we also use the expresion we have just computed for the Fourier series of $\widehat{x}(\xi)$, to obtain that

$$\begin{aligned}
x(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\widehat{x}(\xi)e^{\mathbf{i}\xi t}d\xi = \frac{1}{2\pi}\int_{-\Omega}^{\Omega}\widehat{x}(\xi)e^{\mathbf{i}\xi t}d\xi \\
&= \frac{1}{2\pi}\frac{\pi}{\Omega}\int_{-\Omega}^{\Omega}\left(\sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})e^{-\mathbf{i}\xi\frac{n\pi}{\Omega}}\right)e^{\mathbf{i}\xi t}d\xi \\
&= \frac{1}{2\Omega}\int_{-\Omega}^{\Omega}\left(\sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})e^{\mathbf{i}\xi(t-\frac{n\pi}{\Omega})}\right)d\xi \\
&= \frac{1}{2\Omega}\sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})\frac{e^{\mathbf{i}\xi(t-\frac{n\pi}{\Omega})}}{\mathbf{i}(t-\frac{n\pi}{\Omega})}\Bigg]_{\xi=-\Omega}^{\xi=\Omega} \\
&= \frac{1}{2\Omega}\sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})\frac{e^{\mathbf{i}\xi(\Omega t-n\pi)/\Omega}}{\mathbf{i}(\Omega t-n\pi)/\Omega}\Bigg]_{\xi=-\Omega}^{\xi=\Omega} \\
&= \frac{1}{2\Omega}\sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})\frac{\sin(\Omega t-n\pi)}{\Omega t-n\pi},
\end{aligned}$$

which is what we wanted to prove. (We were able to compute the term by term integration in the expression above because the Fourier series of $\widehat{x}(\xi)$ converges in the norm of $L^2(-\Omega,\Omega)$, and the integral we wanted to compute was essentially the inner product (in $L^2(-\Omega,\Omega)$) of the signal $e^{\mathbf{i}\xi t}$ against the Fourier series of $\widehat{x}(\xi)$). $\square$

Of course, there exists a dual version of the sampling theorem, which works for time-limited signals which are sampled in the frequency domain. (The proof is analogous to the above: just take in consideration the simmetries which exists between the Fourier transform and its inverse)

Theorem 85. *Le us assume that $\widehat{x}(\xi) \in L^2(\mathbb{R})$ and $x(t) = 0$ for all $|t| > L$. Then $\widehat{x}(\xi)$ can be recovered from its values at the points $\{\xi_n = n\pi/L\}_{n\in\mathbb{Z}}$. Indeed, the following formula holds:*

$$\widehat{x}(\xi) = \sum_{n=-\infty}^{\infty} \widehat{x}(\frac{n\pi}{L})\frac{\sin(L\xi - n\pi)}{L\xi - n\pi}.$$

5.2.1. *Some remarks.* The following remarks to the sampling theorem are interesting:

- The class of band-limited signals $\mathbf{BL}(\Omega) = \{x \in L^2(\mathbb{R}) : \widehat{x}(\xi) = 0 \text{ for all } |\xi| > \Omega\}$ is a Hilbert subspace of $L^2(\mathbb{R})$ (with the same inner product than $L^2(\mathbb{R})$). Moreover, the signals
$$s_n(t) = \frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}, \; n \in \mathbb{Z}$$
are an orthogonal basis of $\mathbf{BL}(\Omega)$ and
$$x(t) = \sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\Omega})\frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi} \tag{7}$$
is the Fourier expansion associated to the band-limited signal $x(t)$ with respect to this basis. A proof of these claims is based on the use of Plancherel's Formula and the fact that
$$s_n(t) = \mathcal{F}^{-1}(\chi_{[-\Omega,\Omega]}(\xi)\frac{\pi}{\Omega}e^{-\mathbf{i}n\pi\xi/\Omega}), \; n \in \mathbb{Z}$$
(we proved it in the sampling theorem) where the signals
$$\left\{\chi_{[-\Omega,\Omega]}(\xi)\frac{\pi}{\Omega}e^{-\mathbf{i}n\pi\xi/\Omega}\right\}_{n\in\mathbb{Z}}$$
form a basis of $L^2(-\Omega, \Omega)$.
- From the computational point of view, the decomposition (7) has the disadvantaje that (in general) the series is slowly convergent, since $(\sin t)/t$, converges (for $t \to \infty$) slowly to zero. We can accelerate the process by oversampling the interval. to make this, we change our sampling knots $\left\{\frac{n\pi}{\Omega}\right\}_{n=-\infty}^{\infty}$ by the knots $\left\{\frac{n\pi}{\lambda\Omega}\right\}_{n=-\infty}^{\infty}$, for an arbitrary choice of the constant $\lambda > 1$. If we make so, we will change the signal $(\sin t)/t$ (at the new series) by another signal which approaches to zero like t^{-2} for $t \to \infty$. Concretely, the process is as follows:

I) Fix $\lambda > 1$ and let $x(t) \in \mathbf{BL}(\Omega)$, we follows the first steps of the proof of the sampling theorem (but now we change Ω by $\lambda\Omega$), to conclude that

$$\widehat{x}(\xi) = \frac{\pi}{\lambda\Omega} \sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\lambda\Omega}) e^{-\mathbf{i}\xi \frac{n\pi}{\lambda\Omega}}$$

holds for all $|\xi| \leq \lambda\Omega$.

II) Denote by $\widehat{g}_\lambda$ the piecewise linear function which vanishes outside $[-\lambda\Omega, \lambda\Omega]$ and is equal to 1 at $[-\Omega, \Omega]$. We can compute the inverse Fourier transform of $\widehat{g}_\lambda$. It is

$$g_\lambda(t) = \mathcal{F}^{-1}(\widehat{g}_\lambda) = \frac{\cos(\Omega t) - \cos(\lambda\Omega t)}{\pi(\lambda - 1)\Omega t^2}.$$

III) Taking in consideration that $\widehat{x}(\xi) = \widehat{x}(\xi)\widehat{g}_\lambda(\xi)$, we insert the series we obtained in I) at the Fourier inversion formula, to obtain that

$$\widehat{x}(\xi) = \frac{1}{2\pi} \int_{-\lambda\Omega}^{\lambda\Omega} \widehat{x}(\xi)\widehat{g}_\lambda(\xi) e^{\mathbf{i}\xi t} d\xi = \cdots = \frac{\pi}{\lambda\Omega} \sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\lambda\Omega}) g_\lambda(t - \frac{n\pi}{\lambda\Omega}),$$

which is what we wanted to prove.

5.3. **The uncertainty principle.** The following result is very easy to prove, just using some basic tools from Complex Analysis.

Theorem 86. *If the signal $x(t) \in L^2(\mathbb{R})$ is time-limited and band-limited, then it is equal to the zero signal.*

Proof. If $\widehat{x}(\xi) = 0$ for all $|\xi| > \Omega$, and $\widehat{x}(\xi) \in L^2(\mathbb{R})$ then the function

$$h(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi z} \widehat{x}(\xi) d\xi$$

is well defined for all $z \in \mathbb{C}$ and can be derivated under the integral sign. This implies that $h(z)$ is an entire function. Now, taking into account the Fourier integral theorem, we check that $h(z)$ is an extension of the signal $x(t) \in L^2(\mathbb{R})$ to the whole complex plane. Now, it follows that $h(z)$ has only isolated zeros. Hence $x(t)$ has no compact support, so that it is not a time-limited signal. An analogous argument proves that nonzero time-limited signals can not be band-limited. □

The result above is a weak form of a more general principle which, in the context of quantum mechanics, is called the uncertainty principle, and essentially claims the following: if a signal $x(t)$ vanishes or it is very small outside a certain small neighbourhood of a point, then its Fourier transform $\widehat{x}(\xi)$ must be very sparse in a large interval of the real line (this means that the energy of the signal $\mathbf{E}(\mathbf{x}) = \int_{-\infty}^{\infty} |\widehat{x}(\xi)|^2 d\xi$ can not be approximated by integration of the energy function $|\widehat{x}(\xi)|^2$ in any small interval of the real line). Of course, a dual result also holds (i.e., we can interchange the role of x and $\widehat{x}$ in the claim above). We state here the quantitative version of this result, although we will no make the proof.

A meassure of the sparsity of the signal $x(t)$ with respect to a certain point $a \in \mathbb{R}$, can be defined as follows:

$$\Delta_a x = \frac{\int_{-\infty}^{\infty} (t-a)^2 |x(t)|^2 dt}{\int_{-\infty}^{\infty} |x(t)|^2 dt}$$

If $x(t)$ has small modulus in a neighbourhood of $a \in \mathbb{R}$, the factor $(t-a)^2$ makes $\int_{-\infty}^{\infty} (t-a)^2 |x(t)|^2 dt$ small in comparison with $\int_{-\infty}^{\infty} |x(t)|^2 dt$, so that $\Delta_a x \approx 0$. On the other hand, if $x(t)$ is sparsed far away from a the factor $(t-a)^2$ makes $\int_{-\infty}^{\infty} (t-a)^2 |x(t)|^2 dt$ large in comparison with $\int_{-\infty}^{\infty} |x(t)|^2 dt$, so that $\Delta_a x >> 0$.

Theorem 87. ***(Heissenberg's inequality)*** *Let $x \in L^2(\mathbb{R})$ be an energy signal. Then*

$$\frac{1}{4} \leq (\Delta_a x)(\Delta_b \widehat{x})$$

for all $(a, b) \in \mathbb{R}^2$.

References

[1] **R. M. Bracewell,** *The Fourier transform and its applications,* McGraw-Hill Electrical and Electronic Engeenering Series (1965).

[2] **D. C. Champeney,** *A handbook of Fourier Theorems,* Cambridge university Press (1987).

[3] **R. E. Edwards,** *Fourier series (A modern introduction) Vol 1,* Graduate Texts in Mathematics **64***,* Springer (1979).

[4] **G. B. Folland,** *Fourier analysis and its applications,* Brooks-Cole Publishing Co. (1992).

[5] **R. A. Goldberg,** *Fourier transforms*, Cambridge University Press (1965).

[6] **E. A. González-Velasco,** *Fourier Analysis and Boundary Value Problems,* Academic-Press (1995).

[7] **R. W. Hamming,** *Digital Filters (Third Edition),* Dover (1998).

[8] **G. H. Hardy, W. W. Rogosinski,** *Fourier Series,* Dover (1999).

[9] **D. W. Kammler,** *A first course in Fourier Analysis,* Prentice-Hall, New Jersey (2000).

[10] **Y. Katznelson,** *An introduction to Harmonic Analysis,* Dover (1976).

[11] **S. G. Krantz,** *A panorama of Harmonic Analysis, The* Carus Math. Monographs **27** Math. Asoc. Amer. (1999)

[12] **P. Lancaster, K. Salkauskas,** *Transform Methods in Applied Mathematics,* Canadia math. Soc. series of NMonographs and Advanced Texts, John-Wiley & Sons (1996).

[13] **S. Mallat,** *A wavelet tour of signal processing,* 2^aEd., Academic Press (1999).

[14] **A. Zygmund,** Trigonometric Series, $2^{\underline{a}}$ Ed. Cambridge Univ. Press (1988).