# Functional- and Fourieranalysis

March 17, 2020

This scriptum is by no means intended to be original: It is a note on the things I do in the lecture. In some parts, it is very similar to the excellent books by Christian Blatter [?] and Hans Wilhelm Alt [?]. In particular, I try to adapt – if possible – to Blatter's notation, making it easier for readers to read the book. (This book can be found on Blatter's website `https://people.math.ethz.ch/~blatter/dlp.html` in German, it is also available in UHasselt's library in English.)
This scriptum is only intented for use in my 2020 class!
Please mail mistakes to jochen.schuetz@uhasselt.be.

## 1 Preliminaries, Banach- and Hilbert spaces

In this lecture, we consider Banach- and Hilbert spaces. Those spaces are always vector spaces over a field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with a norm and some special properties.

**Definition 1** (Normed space)**.** *Let $\mathcal{X}$ be a vector space. A function $\|\cdot\| : \mathcal{X} \to \mathbb{R}^{\geq 0}$ is called a* norm*, if there holds for all $x, y \in \mathcal{X}$, $\lambda \in \mathbb{K}$,*

- $\|\lambda x\| = |\lambda|\|x\|$ *(homogeneity),*
- $\|x\| = 0$ *if and only if* $x = 0$*,*
- $\|x + y\| \leq \|x\| + \|y\|$ *(triangle inequality).*

*The pairing $(\mathcal{X}, \|\cdot\|)$ is called a normed space.*

**Example 1.** *The following pairs are normed spaces:*

- $(\mathbb{R}^n, \|\cdot\|_p)$ *for $p \in [1, \infty]$ with the usual definition $\|x\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}$ for $p < \infty$ and $\|x\|_\infty := \max_{1\leq i\leq n} |x_i|$. It is easy to show that these norms fulfill the homogeneity property and $\|x\|_p = 0$ if and only if $x = 0$, however, save for $p \in \{1, \infty\}$, the triangle inequality is difficult to show, it will be shown below, see Lemma 1.*
- *The space $C^0([a, b])$ of continuous functions $[a, b] \to \mathbb{R}$, equipped with the supremum-norm*

$$\|f\|_\infty := \sup_{x\in[a,b]} |f(x)|$$

*is a normed space. Addition and scalar multiplication are to be understood componentwise, so $f + g$ is the function $x \mapsto f(x) + g(x)$.*

- *The space of functions that are Lebesgue-integrable to power p,*

$$L^p([a,b]) := \{f : [a,b] \to \mathbb{R} \mid f \text{ is integrable}, \int_a^b |f(x)|^p \mathrm{d}x < \infty\},$$

  *together with the norm*

$$\|f\|_p := \sqrt[p]{\int_a^b |f(x)|^p \mathrm{d}x}$$

  *is a normed space if $1 \leq p \leq \infty$. (Note that for $p \equiv \infty$, the norm has to be replaced by the* $\sup -$*norm.) Note that the notation $\|\cdot\|_p$ is ambiguous, depending on whether $\mathbb{R}^n$ or $L^p$ is considered.*

This course on *functional analysis* is the extension of linear algebra (which dealt with finite-dimensional spaces) to infinite-dimensional spaces. Some unexpected behavior shows up then. Before we show some examples, we start with the following important theorems:

**Theorem 1** (Young's inequality)**.** *Let $p, q \in \mathbb{R}^{>1}$ such that $\frac{1}{p} + \frac{1}{q} = 1$, and let $a, b \in \mathbb{R}^{\geq 0}$. Then, there holds*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

*Proof.* We may assume that $a, b > 0$, as the inequality becomes trivial otherwise. The logarithm function is concave ($\ln''(x) = \frac{-1}{x^2} < 0$), so there holds

$$\ln(x + \tau(y-x)) \geq (1-\tau)\ln(x) + \tau \ln(y), \qquad \tau \in [0,1].$$

From this and some elementary rules on the logarithm, we get (set $\tau := \frac{1}{q}$ and hence $(1-\tau) = \frac{1}{p}$)

$$\ln(ab) = \ln(a) + \ln(b) = \frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q) \leq \ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right).$$

Monotonicity of ln concludes the proof. □

From this, we can derive the discrete Hölder inequality:

**Theorem 2** (Hölder's inequality)**.** *Let $p, q \in \mathbb{R}^{\geq 1} \cup \{\infty\}$ such that $\frac{1}{p} + \frac{1}{q} = 1$. (We set $\frac{1}{\infty} := 0$ for this.) Then, there holds for all $x, y \in \mathbb{R}^n$ that*

$$\sum_{i=1}^n |x_i||y_i| \leq \|x\|_p \|y\|_q.$$

*Proof.* We only consider $x, y \neq 0$. The claim is easy to prove for $p = 1$ and $q = \infty$ (or the other way around), so we assume that $p, q \notin \{1, \infty\}$. Then, one can write

$$\sum_{i=1}^n \frac{|x_i|}{\|x\|_p} \frac{|y_i|}{\|y\|_q} \overset{\text{Young}}{\leq} \sum_{i=1}^n \left(\frac{|x_i|^p}{p\|x\|_p^p} + \frac{|y_i|^q}{q\|y\|_q^q}\right) = \frac{1}{p} + \frac{1}{q} = 1.$$

Multiplying by $\|x\|_p \|y\|_q$ concludes the proof. □

Let us, for the sake of completeness, give the Hölder inequality also for Lebesgue functions:

**Theorem 3** (Hölder's inequality for Lebesgue functions)**.** *Let $p, q \in \mathbb{R}^{\geq 1} \cup \{\infty\}$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Note the above convention for $p$ or $q$ equal to infinity. Let $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is an open set. Then, there holds*

$$\int_\Omega |u||v| \mathrm{d}x \leq \|u\|_p \|v\|_q.$$

An important application of Hölder's inequality is the triangle inequality:

**Lemma 1** (Triangle inequality of $\|\cdot\|_p$). *For $x, y \in \mathbb{R}^n$, there holds*

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p$$

*for $p \in [1, \infty]$.*

*Proof.* The proof is easy for $p \in \{1, \infty\}$, as one only has to rely on the triangle inequality for the absolute value. Hence, we restrict ourselves to $1 < p < \infty$. Note that we will define $q := (1 - \frac{1}{p})^{-1} = \frac{p}{p-1}$, so that $\frac{1}{p} + \frac{1}{q} = 1$. There holds:

$$\begin{aligned}
\|x+y\|_p^p = \sum_{i=1}^n |x_i + y_i|^p &= \sum_{i=1}^n |x_i + y_i| |x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^n |x_i| |x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i| |x_i + y_i|^{p-1} \\
&\overset{\text{Hölder}}{\leq} \|x\|_p \left( \sum_{i=1}^n |x_i + y_i|^{(p-1)\cdot q} \right)^{1/q} + \|y\|_p \left( \sum_{i=1}^n |x_i + y_i|^{(p-1)\cdot q} \right)^{1/q} \\
&= (\|x\|_p + \|y\|_p) \left( \sum_{i=1}^n |x_i + y_i|^{(p-1)\cdot q} \right)^{1/q} \\
&= (\|x\|_p + \|y\|_p) \left( \sum_{i=1}^n |x_i + y_i|^{p} \right)^{\frac{p-1}{p}} = (\|x\|_p + \|y\|_p) \|x+y\|_p^{p-1}
\end{aligned}$$

If $x + y \neq 0$, then division by $\|x+y\|_p^{p-1}$ gives the desired result. For $x + y = 0$, the inequality is trivial. □

Having discussed these very fundamental properties, we will now look into infinite-dimensional spaces.

**Example 2.** *We give some examples of infinite-dimensional spaces.*

- *By $\mathbb{K}^{\mathbb{N}}$, we denote the set of all sequences with values in $\mathbb{K}$, so*

  $$\mathbb{K}^{\mathbb{N}} := \{(a_i)_{i \in \mathbb{N}} \mid a_i \in \mathbb{K} \, \forall i \in \mathbb{N}\}.$$

  *Similar to norms on $\mathbb{R}^n$, define for $p \in [1, \infty)$ the functions*

  $$\|a\|_{l^p} := \sqrt[p]{\sum_{i \in \mathbb{N}} |a_i|^p}, \qquad \|a\|_{l^\infty} := \sup_{i \in \mathbb{N}} |a_i|.$$

  *Note that those are not norms on $\mathbb{K}^{\mathbb{N}}$. (Why not?) They are norms on the subspaces*

  $$l^p(\mathbb{K}) := \{a \in \mathbb{K}^{\mathbb{N}} \mid \|a\|_{l^p} < \infty\},$$

  *$l^p(\mathbb{K})$ is thus a normed space. For $i \in \mathbb{N}$, define $e_i \in \mathbb{K}^{\mathbb{N}}$ as the sequence having one at position $i$, and zeros elsewhere. Obviously, $e_i \in l^p(\mathbb{K})$ for any $p \in [1, \infty]$, and all $e_i$ are linearly independent. Therefore, $l^p(\mathbb{K})$ is indeed an infinite-dimensional space.*

- *The space $C^0([a, b])$ is also infinite-dimensional. To see this, choose $n \in \mathbb{N}$ and select $n$ pairwise distinct points $x_i \in [a, b]$, $1 \leq i \leq n$. Denote $\varepsilon := \frac{1}{2} \min_{i \neq j} |x_i - x_j| > 0$. There exist continuous functions $f_i$ with the following properties:*

- $\text{supp}(f_i) \subset (x_i - \varepsilon, x_i + \varepsilon)$,
- $f_i(x_i) = 1$.

*(Show this!) As a consequence, $f_i(x_j) = 0$ for any $j \neq i$. Also these $f_i$ must be linearly independent. As $n \in \mathbb{N}$ was arbitrary, $C^0([a,b])$ cannot have a finite dimension.*

**Example 3.** *Let $\mathcal{X}$ be a finite-dimensional vector space, $\dim(\mathcal{X}) = n$. We know that there exists a basis $e_i \in \mathcal{X}$, $i = 1, \ldots, n$, such that any $x \in \mathcal{X}$ can be written as a linear combination of these $e_i$. One consequence is that any two norms $\|\cdot\|$ and $\|\cdot\|_*$ that can be defined on $\mathcal{X}$ are equivalent. As a reminder, being equivalent means that there are constants $c, C \in \mathbb{R}^{>0}$ such that*

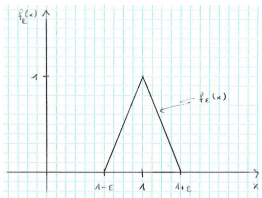
$$c\|x\| \leq \|x\|_* \leq C\|x\|, \qquad \forall x \in \mathcal{X}\,.$$

*Norms are not necessarily equivalent any more in infinite-dimensional spaces! Consider $C^0([a,b])$ with the supremum-norm $\|\cdot\|_\infty$ and the integral norm $\|\cdot\|_1$, $\|f\|_1 := \int_a^b |f| \mathrm{d}x$. There holds*

$$\|f\|_1 \leq |b-a| \|f\|_\infty,$$

*however, $c\|f\|_\infty \leq \|f\|_1$ does not hold for any $c \in \mathbb{R}^{>0}$. To see this, consider for simplicity $[a,b] = [0,2]$ (idea is the same on any other $[a,b]$), and define the continuous, piecewise linear parametric function*

$$f_\varepsilon(x) := \begin{cases} \frac{x-1+\varepsilon}{\varepsilon}, & x \in [1-\varepsilon, 1], \\ \frac{1-x+\varepsilon}{\varepsilon}, & x \in (1, 1+\varepsilon], \\ 0, & \text{otherwise.} \end{cases}$$




*For any $0 < \varepsilon < 1$, there holds $\|f_\varepsilon\|_\infty = 1$, but $\|f_\varepsilon\|_1 = \varepsilon$.*

As a reminder: A sequence $(a_i)_{i \in \mathbb{N}}$ is called a Cauchy sequence if for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$, such that for all $i, j \geq n_0$, there holds

$$|a_i - a_j| \leq \varepsilon.$$

Cauchy series do not necessarily have to converge. As an example, consider the space $\mathbb{Q}$ and the sequence $a_i$ with $i \mapsto \left(1 + \frac{1}{i}\right)^i \in \mathbb{Q}$. We know that (in $\mathbb{R}$) this sequence converges to $e$, so there exists for any $\varepsilon > 0$ an $n_0$ such that $|e - a_i| \leq \frac{\varepsilon}{2}$ for any $i \geq n_0$. Consequently,

$$|a_i - a_j| \leq |a_i - e| + |a_j - e| \leq \varepsilon, \qquad i, j \geq n_0$$

and $(a_i)$ is a Cauchy sequence in $\mathbb{Q}$. Still, it has no limit in $\mathbb{Q}$, because $\mathbb{Q}$ is not *complete*:

**Definition 2** (Completeness)**.** *A normed space $(\mathcal{X}, \|\cdot\|)$ is called* complete*, if every Cauch sequence has a limit in $\mathcal{X}$.*

Having said this, we come to the very important definition of what is a Banach space.

**Definition 3** (Banach space)**.** *A normed vector space $(\mathcal{X}, \|\cdot\|)$ that is complete (w.r.t. to the norm) is called a* Banach space*.*

**Remark 1.** *Let us note that many of the completeness proofs have the following structure:*

- *First, one identifies a suitable 'weaker' limit. (These could be pointwise limits for functions, or element-wise limits for sequences. )*

- *Then, one shows that this 'weaker' limit is an actual limit, and also that it belongs to $\mathcal{X}$.*

**Example 4.** *Here, we show some examples of (non-)complete vector spaces.*

- *We begin with the space $l^p(\mathbb{K})$ and the associated norm $\|\cdot\|_{l^p}$. (We consider $p<\infty$, $p=\infty$ is similar!) This is a complete normed space. To see this, assume that we have a sequence $(x_n)_{n\in\mathbb{N}} \subset l^p(\mathbb{K})$. (So the $x_n \in l^p(\mathbb{K})$ are themselves sequences!) Being a Cauchy sequence means that*

$$\sum_{i\in\mathbb{N}} |x_{n,i}-x_{m,i}|^p = \|x_n - x_m\|_{l^p}^p \le \varepsilon,$$

*for all $n,m$ above some given threshold. (Note that power of $p$. We will do that frequently in order not having to work with the root.) In particular, this implies that $(x_{n,i})_{n\in\mathbb{N}} \subset \mathbb{K}$ is a Cauchy sequence. Even more, for all $i$ and a given $\varepsilon$, we can find $n_0$ such that for all $n,m>n_0$, there holds*

$$|x_{n,i}-x_{m,i}| \le \varepsilon_i$$

*such that $\sum_{i\in\mathbb{N}} \varepsilon_i^p \le \varepsilon$. Because $\varepsilon_i^p \le \varepsilon$ and $n_0$ does not depend on $i$, this is of course uniform convergence. As a field is always complete, there exists a limit $\overline{x}_i$ to the sequence $(x_{n,i})_{n\in\mathbb{N}}$. Passing to the limit in the above inequality, one can also conclude that*

$$|x_{n,i}-\overline{x}_i| \le \varepsilon_i.$$

*Define the sequence $(\overline{x}_i)_{i\in\mathbb{N}}$. We have to show that it is in $l^p(\mathbb{K})$ and that it is the limit of $(x_n)_{n\in\mathbb{N}}$. So, there holds due to La. 1 (assume $n$ is chosen sufficiently large)*

$$\left(\sum_{i\in\mathbb{N}} |\overline{x}_i|^p\right)^{\frac{1}{p}} = \left(\sum_{i\in\mathbb{N}} |x_{n,i}-x_{n,i}+\overline{x}_i|^p\right)^{\frac{1}{p}} \le \left(\sum_{i\in\mathbb{N}} |x_{n,i}-\overline{x}_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i\in\mathbb{N}} |x_{n,i}|^p\right)^{\frac{1}{p}} < \infty,$$

*we therefore have $(\overline{x}_i)_{i\in\mathbb{N}} \in l^p(\mathbb{K})$. Furthermore,*

$$\sum_{i\in\mathbb{N}} |x_{n,i}-\overline{x}_i|^p \le \sum_{i\in\mathbb{N}} \varepsilon_i^p \le \varepsilon,$$

*so $(\overline{x}_i)_{i\in\mathbb{N}}$ is in fact the limit of $(x_n)_{n\in\mathbb{N}}$.*

- *We now look at $C^0([0,1])$ paired with two different norms.*
  - *First, we take the $\|\cdot\|_\infty$ norm. Then, the space is complete. To see this, consider a sequence of continuous functions, $(f_n)_{n\in\mathbb{N}} \subset C^0([0,1])$. Assuming that it is Cauchy means that*

$$|f_n(x)-f_m(x)| \le \|f_n - f_m\|_\infty \le \varepsilon$$

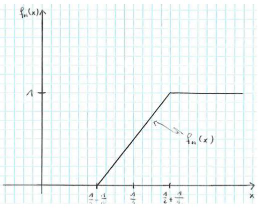
*for $n,m$ sufficiently large and $x\in[0,1]$. This implies that for a fixed $x$, $(f_n(x))_{n\in\mathbb{N}}$ converges to some limit $f(x)$. We have to show that this function is then again in $C^0([0,1])$, we only have to show continuity of course. We compute, for $|\delta|$ sufficiently small such that $x+\delta\in[0,1]$, and $x\in(0,1)$,*

$$\begin{aligned}|f(x+\delta)-f(x)| &= |f(x+\delta)-f_n(x+\delta)+f_n(x+\delta)-f_n(x)+f_n(x)-f(x)| \\ &\le |f(x+\delta)-f_n(x+\delta)|+|f_n(x+\delta)-f_n(x)|+|f_n(x)-f(x)| \\ &\le \varepsilon + |f_n(x+\delta)-f_n(x)| + \varepsilon.\end{aligned}$$

*We have assumed that $n$ is so large that $|f(x)-f_n(x)|\le\varepsilon$. As $f_n$ is continuous, $|f_n(x+\delta)-f_n(x)|\to 0$ with $\delta\to 0$. $f(x)$ is therefore in $C^0([0,1])$ and the space is complete.*

– *Taking the $\|\cdot\|_1$-norm shows a different picture, the space $C^0([0,1])$ will not be complete anymore under this norm. Take the sequence of functions*

$$f_n(x) := \begin{cases} 0, & x < \frac{1}{2} - \frac{1}{2n} \\ n\left(x - \frac{1}{2}\right) + \frac{1}{2} & \frac{1}{2} - \frac{1}{2n} \leq x \leq \frac{1}{2} + \frac{1}{2n} \\ 1, & \text{otherwise} \end{cases}$$




*The $f_n$ are continuous and form a Cauchy sequence w.r.t. to the $\|\cdot\|_1$ norm (try this as an exercise!). The limit, however, is the function*

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

*which is not continuous any more.*

Most of the examples we will be considering in this lecture are indeed Banach spaces. Even more, they are typically Banach spaces with a norm induced by a scalar product.

**Definition 4** (Scalar product)**.** *A function $(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{K}$ is called a* scalar product *if it is sesquilinear, which means for $\alpha, \beta \in \mathbb{K}$ and $x, y, z \in \mathcal{X}$ that*

$$(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z),$$
$$(z, \alpha x + \beta y) = \overline{\alpha}(z, x) + \overline{\beta}(z, y),$$

*and if additionally, there holds:*

- $(u, v) = \overline{(v, u)}$. *(This is the complex conjugate; in case $\mathbb{K} = \mathbb{R}$ there holds $(u, v) = (v, u)$ of course.)*
- $(u, u) > 0$ *for all $u \in \mathcal{X} \setminus \{0\}$.*

**Remark 2.**
- *Note that $(0, 0) = 0$ follows easily due to linearity.*
- *The condition $(u, u) > 0$ makes only sense if $(u, u) \in \mathbb{R}$. This, however, is a consequence of $(u, v) = \overline{(v, u)}$ and hence $(u, u) = \overline{(u, u)} \in \mathbb{R}$.*

**Lemma 2.** *If the vector space $\mathcal{X}$ is equipped with a scalar product $(\cdot,\cdot)$, then this scalar product induces a norm through*

$$\|\cdot\| : \quad x \mapsto \sqrt{(x, x)}.$$

Before proving this, we will first discuss Cauchy-Schwarz's inequality. For vectors $u, v \in \mathbb{R}^2$, it is known that

$$(u, v) = \|u\|\|v\| \cos(\theta),$$

with $\theta$ the angle between $u$ and $v$. Consequently, as $|\cos(\cdot)| \leq 1$, there holds $|(u, v)| \leq \|u\|\|v\|$. The extension of this to general linear spaces is the core of the following theorem:

**Theorem 4** (Cauchy-Schwarz inequality)**.** *Let the vector space $\mathcal{X}$ be equipped with a scalar product $(\cdot,\cdot)$, and let $\|\cdot\|$ be the induced norm. Then, there holds*

$$|(u,v)| \leq \|u\|\|v\|, \qquad \forall u,v \in \mathcal{X}\,.$$

*There is equality if and only if $u$ and $v$ are linearly dependent.*

*Proof.* The theorem is obviously true if $u$ or $v$ are zero. Let us therefore assume $u \neq 0 \neq v$. We define the orthogonal projection of $v$ onto $u$ as $\frac{(v,u)}{\|u\|^2}u$. Then, there holds

$$\begin{aligned} 0 \leq \left\| v - \frac{(v,u)}{\|u\|^2}u \right\|^2 &= (v,v) + \frac{|(v,u)|^2(u,u)}{\|u\|^4} - \left(v, \frac{(v,u)}{\|u\|^2}u\right) - \left(\frac{(v,u)}{\|u\|^2}u, v\right) \\ &= \|v\|^2 + \frac{|(v,u)|^2}{\|u\|^2} - 2\frac{|(v,u)|^2}{\|u\|^2} = \|v\|^2 - \frac{|(v,u)|^2}{\|u\|^2}. \end{aligned}$$

(You should try and do the intermediate steps. Note that $(\cdot,\cdot)$ is sesquilinear.) Multiplied by $\|u\|^2$, this yields

$$|(v,u)|^2 \leq \|v\|^2\|u\|^2,$$

which is the inequality claimed. There is equality if and only if $v - \frac{(v,u)}{\|u\|^2}u = 0$, which of course means that $u$ and $v$ are linearly dependent. □

**Corollary 1.** *Assume the same conditions as in Thm. 4. Then, there exists a $\theta$ (unique in $[0,\pi]$) such that*

$$(u,v) = \|u\|\|v\|\cos(\theta).$$

We now prove that $x \mapsto \sqrt{(x,x)}$ does indeed define a norm.

*Proof of La. 2.* The only non-trivial part to show is the triangle inequality, so let $x,y \in \mathcal{X}$, then

$$\begin{aligned} \|x+y\|^2 = (x+y,x+y) &= \|x\|^2 + (x,y) + (y,x) + \|y\|^2 \\ &= \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}((x,y)) \\ &\leq \|x\|^2 + \|y\|^2 + 2|(x,y)| \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2. \end{aligned}$$

This concludes the proof. □

Having said this, we can define Hilbert spaces:

**Definition 5** (Hilbert space)**.** *A Banach space with scalar-product-induced norm is called a* Hilbert space.

**Example 5.** *Let us show two examples of Hilbert spaces.*

- *Obviously, $\mathbb{R}^n$ together with the 'standard' scalar product $(x,y) := \sum_{i=1}^n x_i y_i$ is a Hilbert space. The induced norm is $\|\cdot\|_2$. Note that for any $p \neq 2$, $\|\cdot\|_p$ is not a norm induced by a scalar product. This will become clear later.*

- *Because the following space will be of utmost importance to Fourier analysis, we will have a definition within an example:*

**Definition 6.** *We define $L^2_\circ$ to be the space of square-integrable periodic functions with periodicity $2\pi$, i.e., for any $f \in L^2_\circ$, there holds $f(x+2\pi) = f(x)$ almost everywhere, and the integral*

$$\|f\|^2_{L^2_\circ} := \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 \mathrm{d}x$$

*is finite. (We use the scaling of $\frac{1}{2\pi}$ for our convenience, in Fourier analysis, it will turn out that then, a lot of constants are one.)*

*Then, $L^2_\circ$ is a Hilbert space together with the scalar product*

$$(f,g)_{L^2_\circ} := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g}(x)\mathrm{d}x.$$

**Lemma 3** (Parallelogram identity)**.** *Let $\mathcal{X}$ be a Hilbert space. Then, there holds for all $u, v \in \mathcal{X}$,*

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2. \tag{1}$$

*Proof.* Straightforward computation, convince yourself. □

**Remark 3.** *With La. 3, one can prove that on $\mathbb{R}^n$ $(n > 1)$, all $p-$norms, except for $p = 2$, are not induced by scalar products. (Exercise!)*

Orthogonality plays already an important role in the finite-dimensional case, and so does it in the infinite-dimensional one. In particular, projections onto subspaces are based on orthogonal vectors.

**Definition 7** (Orthogonality and projection)**.** *Let $\mathcal{X}$ be a Hilbert space, and let $\mathcal{Y}$ be a subspace of $\mathcal{X}$. Two elements $x, y \in \mathcal{X}$ are said to be orthogonal, if there holds*

$$(x, y) = 0.$$

*The projection of some $x \in \mathcal{X}$ onto $\mathcal{Y}$ is defined as the value $\Pi_\mathcal{Y}(x)$ for which*

$$\|x - \Pi_\mathcal{Y}(x)\| = \min_{y\in\mathcal{Y}} \|x - y\|.$$

*If the context is clear, we will often write $\Pi$ instead of $\Pi_\mathcal{Y}$.*

**Lemma 4.** *If $\mathcal{Y}$ is non-empty and closed, then $\Pi_\mathcal{Y}(x)$ is well-defined. In particular, the solution of the minimalization problem exists and is unique.*

*Proof.* If $x \in \mathcal{Y}$, then $\Pi_\mathcal{Y}(x) = x$, and there is nothing to prove. We restrict ourselves therefore to the case $x \notin \mathcal{Y}$. Let $\tau := \inf\limits_{y\in\mathcal{Y}} \|x - y\|$. As $\mathcal{Y}$ is closed and $x \notin \mathcal{Y}$, there must hold $\tau > 0$. (Prove this!) Let $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$ be a sequence such that $\|y_n - x\| \to \tau$. We want to prove that $y_n$ is a Cauchy sequence. To this end, observe that for two values $z_1, z_2 \in \mathcal{Y}$, there holds due to the parallelogram identity (1):

$$\|z_1 - z_2\|^2 = \|(z_1 - x) - (z_2 - x)\|^2 = 2\|z_1 - x\|^2 + 2\|z_2 - x\|^2 - \|(z_1 - x) + (z_2 - x)\|^2.$$

Furthermore,

$$\|(z_1 - x) + (z_2 - x)\|^2 = 4\left\|\frac{z_1 + z_2}{2} - x\right\|^2 \geq 4\tau^2,$$

because $\frac{z_1+z_2}{2} \in \mathcal{Y}$ and therefore, the distance to $x$ must be $\geq \tau$. With that being said, there is

$$\|z_1 - z_2\|^2 \leq 2\|z_1 - x\|^2 + 2\|z_2 - x\|^2 - 4\tau^2. \tag{2}$$

Now returning to the sequence $(y_n)_{n\in\mathbb{N}}$. Given an $\varepsilon$, choose $n_0$ such that for all $n \geq n_0$, there holds $|\|y_n - x\| - \tau| \leq \varepsilon$. Then, for all $n, m \geq n_0$, there holds with (2)

$$\|y_n - y_m\| \leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\tau^2 \leq 2(\tau+\varepsilon)^2 + 2(\tau+\varepsilon)^2 - 4\tau^2 = 8\tau\varepsilon + 4\varepsilon^2.$$

$(y_n)_{n\in\mathbb{N}}$ is hence a Cauchy sequence. The limit $\overline{y}$ fulfills $\|\overline{y} - x\| = \tau$, we thereby have existence.

Assume now that there are two $\overline{y}_1, \overline{y}_2 \in \mathcal{Y}$ such that $\|\overline{y}_i - x\| = \tau$. Then, again due to (2), there holds

$$\|\overline{y}_1 - \overline{y}_2\|^2 \leq 2\|\overline{y}_1 - x\|^2 + 2\|\overline{y}_2 - x\|^2 - 4\tau^2 = 0,$$

hence $\overline{y}_1 = \overline{y}_2$, and hence uniqueness. ☐

**Remark 4.**
- *One can define a projection also for a subset $\mathcal{Y} \subset \mathcal{X}$ rather than a subspace. Then La. 4 remains true if $\mathcal{Y}$ is non-empty, closed and convex.*
- *If $\mathcal{Y}$ is a subspace, then the projection operator is linear - this is an easy consequence of the following lemma.*

**Lemma 5.** *Consider the assumptions as in La. 4. Then, there holds:*

$$\overline{y} = \Pi_{\mathcal{Y}}(x) \qquad \Leftrightarrow \qquad (x - \overline{y}, y) = 0, \quad \forall y \in \mathcal{Y}\,.$$

*Proof.* Without loss of generality, we consider the case that $\overline{y} = 0$. This is possible, because of the following: Assume that $\overline{y} = \Pi_{\mathcal{Y}}(x)$, i.e.,

$$\inf_{y\in\mathcal{Y}} \|x - y\| = \|x - \overline{y}\|.$$

Furthermore, as $\overline{y} \in \mathcal{Y}$, there holds

$$\inf_{y\in\mathcal{Y}} \|x - y\| = \inf_{y\in\mathcal{Y}} \|(x - \overline{y}) - (y - \overline{y})\| = \inf_{y\in\mathcal{Y}} \|(x - \overline{y}) - y\| = \|(x - \overline{y}) - 0\|.$$

Hence, $\Pi_{\mathcal{Y}}(x - \overline{y}) = 0$, so we only consider a linear shift.

"⇒": There holds

$$\|x - \overline{y}\| = \|x - 0\| \leq \|x - y\|, \qquad \forall y \in \mathcal{Y}$$

due to the definition of the projection operator. Now fix some $y \in \mathcal{Y}$, and consider, for $t \in \mathbb{R}$, $ty \in \mathcal{Y}$. Consequently,

$$\|x\|^2 \leq \|x - ty\|^2 = \|x\|^2 + t^2\|y\|^2 - 2t\,\mathrm{Re}(x, y)$$

and hence

$$2t\,\mathrm{Re}(x, y) \leq t^2\|y\|^2, \qquad \forall t \in \mathbb{R}\,.$$

Now let first $t$ be positive, and let $t \to 0$. Thus, $\mathrm{Re}(x, y) \leq 0$. For $t$ negative and $t \to 0$, one gets $\mathrm{Re}(x, y) \geq 0$. Thus,

$$\mathrm{Re}(x, y) = 0.$$

In case $\mathbb{K} = \mathbb{C}$, note that with $y \in \mathcal{Y}$ there also holds $iy \in \mathcal{Y}$ and $\mathrm{Im}(x, y) = -\mathrm{Re}(x, iy)$. With the same arguments as before, one gets that also the imaginary part is zero. This concludes the first part.

"⇐": We can compute:

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\,\mathrm{Re}(x, y) = \|x\|^2 + \|y\|^2 \geq \|x\|^2 + \|0\|^2,$$

hence $\Pi_Y(x) = 0$. ☐

**Corollary 2.** *Assume that* $\mathcal{Y} = \mathrm{span}(y_i, 1 \leq i \leq N)$ *with* $y_i \in \mathcal{X}$*,* $\mathcal{X}$ *Hilbert space,* $N \in \mathbb{N}$*, and* $(y_i, y_j) = 0$ *if* $i \neq j$*. Then, there holds*

$$\Pi_{\mathcal{Y}}(x) = \sum_{i=1}^{N} \frac{(x, y_i)}{(y_i, y_i)} y_i.$$

**Example 6.** *There is a very close link to Fourier analysis here, see also Def. 6. Remember that*

$$L^2_\circ := \{f : \mathbb{R} \to \mathbb{C} \ \textit{is measurable} : \|f\|_{L^2_\circ} < \infty, \ f(x + 2\pi) = f(x) \ \textit{almost everywhere}\}.$$

*Periodic functions are best represented as functions 'on a circle', which itself can in the complex plane be described as* $\{e^{it} \mid t \in [0, 2\pi)\}$*. Roughly speaking, it can therefore be said that the role of* $x$ *(and* $x^2, x^3, ...$*) for non-periodic functions is the same as* $e^{it}$ *(and* $e^{2it}, e^{3it}, ...$*) for periodic ones. Define therefore* $\mathbf{e}_k(t) := e^{ikt}$ *for* $k \in \mathbb{Z}$*, and*

$$U_N := \mathrm{span}(\mathbf{e}_{-N}, \dots, 1, \dots, \mathbf{e}_N) \subset L^2_\circ.$$

*It can be easily shown that* $(\mathbf{e}_k, \mathbf{e}_l)_{L^2_\circ} = \delta_{k,l}$ *(try it!), and so the* $\mathbf{e}_k$ *form an orthonormal basis of the finite-dimensional space* $U_N$*.*

*The truncated Fourier series* $s_N$ *of a function* $f \in L^2_\circ$ *is defined as the orthogonal projection of* $f$ *onto* $U_N$*, hence*

$$s_N := \Pi_{U_N}(f). \tag{3}$$

*By construction, it is obvious that* $s_N$ *is a linear combination of the* $\mathbf{e}_k$*. The corresponding coefficients are denoted by* $\widehat{f}(k)$*, hence*

$$s_N =: \sum_{k=-N}^{N} \widehat{f}(k) \mathbf{e}_k.$$

$\widehat{f}(k)$ *are called Fourier coefficients, and can be computed through*

$$\widehat{f}(k) = (f, \mathbf{e}_k) \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} \mathrm{d}x,$$

*see Cor. 2. Also due to construction, and this shows the importance of the truncated Fourier series* $s_N$*, there holds*

$$\|f - s_N\|_{L^2_\circ} = \min_{s \in U_N} \|f - s\|_{L^2_\circ}.$$

## 2 Periodic functions: Fourier series

Before continuing, we remind the reader of the following important property which holds for all $\varphi \in \mathbb{R}$:

$$e^{i\varphi} = \cos(\varphi) + i\sin(\varphi).$$

Furthermore, we will make frequent use of

$$\cos(x) = \cos(-x), \qquad \sin(x) = -\sin(-x).$$

### 2.1 Fourier transform

We have already seen the definition of the Fourier coefficients $\widehat{f}(k)$ and the truncated Fourier series $s_N$. An important guiding question is whether, in what sense, and under what conditions, there holds $\lim_{N\to\infty} s_N = f$. We postpone this discussion for a moment, and start with some simple properties of $\widehat{f}(k)$:

**Lemma 6.** *The following properties hold:*

- *If $f$ is a real-valued function, then $\widehat{f}(k) = \overline{\widehat{f}(-k)}$.*
- *If $f$ is an even function, then $\widehat{f}(k) \in \mathbb{R}$.*
- *If $f$ is an odd function, then $\widehat{f}(k) \in i\,\mathbb{R}$.*

*Proof.* Prove this! $\square$

**Remark 5.** *'Traditionally', the Fourier series has been introduced as*

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

*rather than $\sum_{k\in\mathbb{Z}} \widehat{f}(k)e^{ikx}$. With real $a_k$ and $b_k$, the former is confined to functions $f:\mathbb{R}\to\mathbb{R}$, while the latter also works for functions $f:\mathbb{R}\to\mathbb{C}$. One can observe that there is indeed a simple bijection between the 'traditional' representation and the one with $\mathbf{e}_k$ involved:*

$$\sum_{k=-\infty}^{\infty} \widehat{f}(k)\mathbf{e}_k = \widehat{f}(0) + \sum_{k=1}^{\infty}\left(\left(\widehat{f}(k)+\widehat{f}(-k)\right)\cos(kx) + \left(\widehat{f}(k)-\widehat{f}(-k)\right)i\sin(kx)\right).$$

*Upon defining $a_k := \widehat{f}(k)+\widehat{f}(-k)$ and $b_k := i\left(\widehat{f}(k)-\widehat{f}(-k)\right)$, one observes that there is a simple bijection between the 'traditional' representation and the one with $\mathbf{e}_k$ involved. Note that because of La. 6, $a_k$ and $b_k$ are indeed real if $f$ is a real-valued function. Stated a bit simpler, one can say that $a_0 = 2\operatorname{Re}(\widehat{f}(0))$; and for $k>0$, $a_k = 2\operatorname{Re}(\widehat{f}(k))$ and $b_k = -2\operatorname{Im}(\widehat{f}(k))$.*

Before starting to prove under what circumstances a Fourier series exists - and converges toward $f$ - we show some instructive examples.

**Example 7.**

- *$f(x) = \sin(x)$. Obviously, the 'traditional' Fourier transform has $b_1 = 1$, and all other parameters are zero. Thus, $\widehat{f}(1) = -\widehat{f}(-1) = -\frac{i}{2}$, $\widehat{f}(k) = 0$ for $k \neq \pm 1$.*
- *$f(x) = \sin^2(x)$. Because $\sin^2(x) = \frac{1}{2} - \frac{\cos(2x)}{2}$ (try it!), the Fourier coefficients are $\widehat{f}(0) = \frac{1}{2}$, $\widehat{f}(-2) = \widehat{f}(2) = -\frac{1}{4}$, and $\widehat{f}(k) = 0$ for $k \notin \{0, \pm 2\}$.*

- *Consider the rectangle function* $f(x) = \chi_{(-\frac{1}{2},\frac{1}{2})}(x)$*. Its Fourier transform is given by* $\widehat{f}(0) = \frac{1}{2\pi}$ *and*

$$\widehat{f}(k) = \frac{1}{2\pi}\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-ikx}\mathrm{d}x = \frac{-1}{ik2\pi}\left(e^{-i\frac{k}{2}} - e^{i\frac{k}{2}}\right) = \frac{1}{k\pi}\sin\left(\frac{k}{2}\right) =: \frac{1}{2\pi}\operatorname{sinc}\left(\frac{k}{2}\right)$$

  *The* $\operatorname{sinc}(x) := \frac{\sin(x)}{x}$ *(*sinus cardinalis*) function occurs very often in Fourier analysis, which is why it got an extra name. Note that sometimes, the* sinc *function also comes in the normalized way, unfortunately, there is no standard in literature. Whenever we talk about the* sinc *function, we use it in the way presented above.*

- $f(x) = \frac{x}{\pi}$ *on* $[-\pi, \pi]$*. Note that this function is not smooth! Its Fourier transform can be computed as* $\widehat{f}(0) = 0$ *and,* $k \neq 0$*,*

$$\widehat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{x}{\pi}e^{-ikx}\mathrm{d}x = \frac{1}{2\pi}\left(\frac{-xe^{-ikx}}{\pi ik} + \frac{e^{-ikx}}{\pi k^2}\right)\Bigg|_{-\pi}^{\pi} = \frac{e^{-ikx}}{2\pi^2}\left(\frac{ix}{k} + \frac{1}{k^2}\right)\Bigg|_{-\pi}^{\pi} = \frac{i}{\pi k}\cos(k\pi).$$

So far, we didn't talk about whether the Fourier series exists, and whether it approximates the original function - and in what sense. However, for discontinuous functions, the Fourier series shows a very peculiar behavior at the discontinuity - in particular, the overshoots do not die out as $N \to \infty$. This phenomenon is known as *Gibbs* phenomenon (dutch: *Gibbs-verschijnsel*). Please read about this phenomenon here: `https://en.wikipedia.org/wiki/Gibbs_phenomenon`. Read until (including!) "Explanation". In particular, the first three images are important!

We now embark on the journey of proving that $f \in L^2_\circ$ is indeed approximated (in the $L^2_\circ$−norm) by its Fourier transform. We begin with a direct consequence of the fact that $s_N$ is the orthogonal projection onto $U_N$:

**Theorem 5** (Bessel's (in)equality)**.** *For* $f \in L^2_\circ$ *and* $s_N$ *as defined in* (3)*, there holds:*

$$\|s_N\|^2_{L^2_\circ} + \|f - s_N\|^2_{L^2_\circ} = \|f\|^2_{L^2_\circ}. \tag{4}$$

*From this follows* $\|s_N\|^2_{L^2_\circ} \leq \|f\|^2_{L^2_\circ}$ *easily.*

*Proof.* This is a special case of Ex. 8! □

The following lemma is one of the most important ones in the context of Fourier analysis. It goes back to the French mathematician Marc-Antoine Parseval:

**Lemma 7** (Parseval's identity)**.** *For any* $f \in L^2_\circ$*, there holds:*

$$\|f\|^2_{L^2_\circ} = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)|^2. \tag{5}$$

**Theorem 6.** *With this theorem, one directly obtains for an* $f \in L^2_\circ$*:*

$$\|f - s_N\|_{L^2_\circ} \to 0, \quad N \to \infty. \tag{6}$$

*Proof.* There holds $\|s_N\|^2_{L^2_\circ} = \sum_{k=-N}^{N} |\widehat{f}(k)|^2$. Plugging this into Bessel's equation (4) and exploiting (5) yields the statement. □

*Proof of Parseval's identity, La. 7.* The proof is a bit tedious. W.l.o.g., we only consider functions $f : \mathbb{R} \to \mathbb{R}$, the extension to complex functions is trivial, as any complex function is $f(x) = \mathrm{Re}(f(x)) + i\,\mathrm{Im}(f(x))$.

The proof consists of three steps:

1. We show that the statement holds for a rectangle function

$$f(x) = \begin{cases} 1, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}. \tag{7}$$

2. Then, we show that it holds for a finite linear combination of rectangle functions.

3. Once this is done, we know that for any $f \in L^2_\circ$, one can find a step function $f_\varepsilon$, such that $\|f - f_\varepsilon\|_{L^2_\circ} \le \varepsilon$. Let $s_N$ denote the truncated Fourier series for $f$, and $s_{\varepsilon,N}$ the one for $f_\varepsilon$. Furthermore, let $N$ be such that $\|f_\varepsilon - s_{\varepsilon,N}\| \le \varepsilon$. (We can choose such an $N$, because for step functions, Thm. 6 already holds.) Then, one can conclude:

$$\|f - s_N\|_{L^2_\circ} = \|f_\varepsilon - s_{\varepsilon,N} + f - f_\varepsilon - s_N + s_{\varepsilon,N}\| \le \|f_\varepsilon - s_{\varepsilon,N}\| + \|(f - f_\varepsilon) - (s_N - s_{\varepsilon,N})\|.$$

   Now $\|f_\varepsilon - s_{\varepsilon,N}\| \le \varepsilon$ by assumption, and

$$\|(f - f_\varepsilon) - (s_N - s_{\varepsilon,N})\|^2 \le \|f - f_\varepsilon\|^2 \le \varepsilon^2$$

   because of (4). Now let $\varepsilon \to 0$ and use (4) again.

Now let us work out steps 1 and 2. We remind the reader of the following infinite sums:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}, \qquad \sum_{k=1}^{\infty} \frac{\cos(kb)}{k^2} = \frac{(b-\pi)^2}{4} - \frac{\pi^2}{12}$$

(For an interesting proof of the first one, look up, e.g., *Basel problem* on the internet. The second one can be found in Analysis 1 by O. Forster.)

1. W.l.o.g. we consider $a = 0$ and $0 < b < \pi$. Then, for $f$ as in (7), there holds $\widehat{f}(0) = \frac{b}{2\pi}$, and for $k \neq 0$, $\widehat{f}(k) = \frac{1 - e^{-ikb}}{2\pi i k}$. Summing this up yields:

$$\left(\frac{b}{2\pi}\right)^2 + \sum_{k \in \mathbb{Z}^{\neq 0}} \left|\frac{1 - e^{-ikb}}{2\pi i k}\right|^2 = \frac{b^2}{(2\pi)^2} + \sum_{k \in \mathbb{Z}^{\neq 0}} \left(\frac{(1 - \cos(kb))^2 + \sin(kb)^2}{(2\pi)^2 k^2}\right)$$

$$\frac{b^2}{(2\pi)^2} + \sum_{k \in \mathbb{Z}^{\neq 0}} \frac{1 - 2\cos(kb) + \cos^2(kb) + \sin^2(kb)}{(2\pi k)^2} = \frac{b^2}{(2\pi)^2} + \sum_{k \in \mathbb{Z}^{\neq 0}} \frac{1}{2\pi^2} \frac{1 - \cos(kb)}{k^2}$$

$$= \frac{b^2}{(2\pi)^2} + \frac{2}{2\pi^2}\left(\frac{\pi^2}{6} + \frac{\pi^2}{12} - \frac{(b-\pi)^2}{4}\right) = \frac{2b\pi}{4\pi^2} = \frac{b}{2\pi} = \|f\|^2_{L^2_\circ}$$

2. Now if $f$ is a finite linear combination of rectangular functions, i.e., $f = \sum_{j=1}^{m} c_j f_j$ with $f_j$ being a function of form (7), then because of linearity, one can compute

$$\|f - s_N\| = \|\sum_{j=1}^{m} c_j (f_j - s_{j,N})\| \le \sum_{j=1}^{m} |c_j| \|f_j - s_{j,N}\|.$$

   We know that $\|f_j - s_{j,N}\|$ converges toward zero, which implicitly proves the statement.

□

**Remark 6.** *There is a word of warning necessary here: Convergence in $L^2$ and pointwise convergence is not the same! Let us clarify when one follows from the other:*

- *If the domain $\Omega$ where the function is defined is bounded, and some sequence of functions $f_N$ converges uniformly toward $f$, then it also converges in $L^2$, as one can compute*

$$2\pi \|f - f_N\|^2_{L^2_\circ} = \int_\Omega |f(x) - f_N(x)|^2 \mathrm{d}x \le |\Omega| \|f(x) - f_N(x)\|^2_\infty \to 0.$$

- *All other implications do in general not hold. As an example, let $f_N(x) := N \cdot \chi_{(0,\frac{1}{N})}(x)$. Obviously, $f_N$ converges pointwise toward $f \equiv 0$, however, not uniformly. Furthermore,*

$$2\pi \|f - f_N\|_{L^2_\circ}^2 = N \to \infty.$$

- *However, there is the following statement: If a sequence of functions $(f_N)_{N\in\mathbb{N}}$ converges in $L^2$ to a function $f$, then one can extract a subsequence that converges pointwise almost everywhere.*

**Lemma 8** (Riemann-Lebesgue-Lemma)**.** *For $f \in L^2_\circ$, there holds $\widehat{f}(k) \to 0$ for $k \to \pm\infty$.*

*Proof.* As the $\mathbf{e}_k$ form an orthonormal basis, there holds $\|s_N\|_{L^2_\circ}^2 = \sum_{k=-N}^{N} |\widehat{f}(k)|^2$. From Bessel's inequality one can directly conclude that $\widehat{f}(k) \to 0$ for $k \to \pm\infty$, because otherwise, the infinite sum diverges. □

It is the more fascinating that there hold very general statements for the pointwise convergence for Fourier series. Without a proof, we start with the following important theorem, proved by *Carleson* in 1966:

**Theorem 7.** *Let $f \in L^2_\circ$. Then, $s_N(x)$ converges almost everywhere toward $f(x)$, i.e., there holds $\lim_{N\to\infty} s_N(x) = f(x)$ for almost every $x \in (-\pi, \pi)$.*

This is less precise than what one aims at. Uniform convergence can be achieved if the function is of *bounded variation*:

**Definition 8.** *Let an arbitrary subdivision of $(-\pi, \pi)$ be given as*

$$\mathcal{T} : -\pi = t_0 < t_1 < \ldots < t_n = \pi.$$

*If there holds*

$$\sup_{\mathcal{T}} \sum_{j=1}^{n} |f(t_j) - f(t_{j-1})| < \infty$$

*then $f$ is said to be of bounded variation. The supremum is called the* (total) variation *$V(f)$ of $f$.*

To get an idea of the concept of variation, we state the following lemma:

**Lemma 9.** *Let $f \in C^1$. Then,*

$$V(f) = \int_{-\pi}^{\pi} |f'(x)| \mathrm{d}x.$$

*Proof.* Exercise. Exploit that $f(t_k) - f(t_{k-1}) = f'(\xi)(t_k - t_{k-1})$, where $\xi \in (t_{k-1}, t_k)$. □

However, not only differentiable functions have a variation. As an example, the function $\chi_{(-\pi,0)}$ has variation of two. (Convince yourself!)

With this being said, one can state the following Theorem:

**Theorem 8.** *Let $f \in L^2_\circ$ be of bounded variation. Then, for any $x$, there holds:*

$$s_N(x) \to \frac{f(x^+) + f(x^-)}{2}.$$

*If $f$ is, in addition, continuous, then convergence is uniform.*

*Proof.* The proof needs much more preparation than we have the time fore. A core ingredient, however, is the decomposition of a bounded-variation function $f$ into the difference of two monotonically increasing functions $g_1$ and $g_2$. (This is called Jordan decomposition.) □