Functional- and Fourieranalysis

February 9, 2023

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 [2] and Hans Wilhelm Alt [1]. 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 2021 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. $\square$

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:*

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

(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 $e_i \in \mathcal{X}$, $i = 1, \dots, n$, 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*

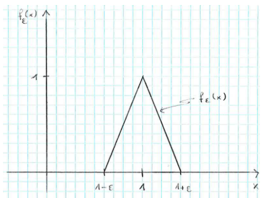
$$c\|x\| \leq \|x\|_* \leq C\|x\|, \qquad \forall x \in \mathcal{X}\,.$$

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

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



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*

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

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 ε, we can find n_0 such that for all $n, m > n_0$, there holds

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

such that $\sum_{i\in\mathbb{N}} \varepsilon_i^p \leq \varepsilon$. Because $\varepsilon_i^p \leq \varepsilon$ and n_0 does not depend on i, this is of course uniform convergence. As the fields $\{\mathbb{R}, \mathbb{C}\}$ are 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

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

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)

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

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

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

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*

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

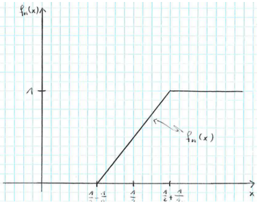
for n, m sufficiently large and $x \in [0,1]$. This implies that for a fixed x, $(f_n(x))_{n\in\mathbb{N}}$ converges to some limit $f(x)$. We have to show that this function is then again in $C^0([0,1])$, we only have to show continuity of course. We compute, for $|\delta|$ sufficiently small such that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

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

We have assumed that n is so large that $|f(x) - f_n(x)| \leq \varepsilon$. 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*

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



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 θ 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, and assume $\mathbb{K} = \mathbb{R}$. Then, there exists a θ (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 2π, 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 Π 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 $\tau := \inf_{y\in\mathcal{Y}} \|x-y\|$. 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 ε, 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.

"$\Rightarrow$": 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.

"$\Leftarrow$": 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*

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

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* $x^2, x^3, \ldots$*) for non-periodic functions is the same as* e^{it} *(and* $e^{2it}, e^{3it}, \ldots$*) for periodic ones. Define therefore* $\mathbf{e}_k(t) := e^{ikt}$ *for* $k \in \mathbb{Z}$, *and*

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

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 $\widehat{f}(k) \in \mathbb{R}$. (Even for complex functions means $f(x) = \overline{f(-x)}$, odd similarly.)*
- *If f is an odd function, then $\widehat{f}(k) \in i\,\mathbb{R}$.*

Proof. Prove this! □

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 $a_0 = 2\,\mathrm{Re}(\widehat{f}(0))$; and for $k > 0$, $a_k = 2\,\mathrm{Re}(\widehat{f}(k))$ and $b_k = -2\,\mathrm{Im}(\widehat{f}(k))$.

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 $f(x) = \operatorname{Re}(f(x)) + i\operatorname{Im}(f(x))$.

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

- *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 for. 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.) □

We make the following formal observations. If f had n derivatives, and the Fourier series approximated f pointwise, and one could differentiate in the sum, then one had

$$f^{(n)}(x) = \sum_{k\in\mathbb{Z}} (ik)^n \widehat{f}(k)\mathbf{e}_k, \tag{8}$$

so $\widehat{f^{(n)}}(k) = (ik)^n \widehat{f}(k)$. To have 'proper' convergence of the Fourier series toward $f^{(n)}$, one needs absolut convergence of $|(ik)^n \widehat{f}(k)|$, i.e., $\widehat{f}(k)$ should be asymptotically smaller than $\frac{1}{|k|^{n+1}}$. The converse should also be true. This means that the decay of the Fourier coefficients gives some insight into the regularity of f. The following theorem clarifies this:

Theorem 9. *1. Let $f^{(n)}$ be the $n-$th derivative of the complex periodic function $f \in L^2_\circ$. If $f^{(n)}$ is continuous and its variation $V(f^{(n)}) =: V < \infty$ is bounded, then*

$$|\widehat{f}(k)| \leq \frac{V}{2\pi|k|^{n+1}}, \quad k \neq 0.$$

2. Conversely, let for an $\varepsilon > 0$,

$$\widehat{f}(k) = \mathcal{O}\left(\frac{1}{|k|^{n+1+\varepsilon}}\right), \quad k \to \infty.$$

Then, $f(x) := \sum_{k\in\mathbb{Z}} \widehat{f}(k)\mathbf{e}_k$ is at least $n-$times continuously differentiable.

Proof. We only show the second part. Let $0 \leq p \leq n$. Then, there holds $\widehat{f}(k)(ik)^p = \mathcal{O}\left(\frac{1}{|k|^{n-p+1+\varepsilon}}\right)$ (which is at least $\mathcal{O}\left(\frac{1}{|k|^{1+\varepsilon}}\right)$), and therefore, the sum

$$\sum_{k\in\mathbb{Z}} \widehat{f}(k)(ik)^p \mathbf{e}_k$$

converges uniformly, and therefore, toward a continuous function. For $p = 0$, this is f, while for $1 \leq p \leq n$, this is the derivative. (Uniform convergence admits interchange of summation and differentiation.) □

2.2 Discrete Fourier transform

In 'practice' there is neither the possibility of storing all the frequencies nor storing functions. In general, a 'function' is only given at some fixed spatial instances (and is therefore merely a set of points). For simplicity, we assume that here, function values are given on the uniform grid

$$\left\{ x_j = -\pi + jh \mid h = \frac{2\pi}{\mathcal{M}}, 0 \leq j \leq \mathcal{M} \right\}.$$

Approximating the Fourier coefficients $\widehat{f}(k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\mathrm{d}x$ can be done using the summed trapezoidal rule in the following way:

$$\widehat{f}(k) \approx \widehat{f}_h(k) := \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_j)e^{-ikx_j}. \tag{9}$$

A couple of questions arise, the most prominent maybe whether one should compute $\widehat{f}_h(k)$ for all $k \in \mathbb{Z}$. Obviously, the answer to this questions is no, because with only a finite number of informations (values of f on x_j), it makes no sense to try to derive infinitely many information. The next lemma will shed some light into this, before, however, we make a small remark on the quality of the (summed) trapezoidal rule:

Remark 7. *Let $T_h(f)$ denote the application of the summed trapezoidal rule with mesh-size h to a function f. For a non-periodic domain $[a,b]$, it is well-known that*

$$\int_a^b f(x)\mathrm{d}x = T_h(f) + \mathcal{O}(h^2).$$

For a periodic domain, the situation is much better: One can show that if $f \in L^2_\circ$ is in C^s for an $s \in \mathbb{N}$, then there holds

$$\int_{-\pi}^{\pi} f(x)\mathrm{d}x = T_h(f) + \mathcal{O}(h^s).$$

As s is arbitrary, this phenomenon is called exponential convergence. It is very rare - appreciate it. In this sense, the summed trapezoidal rule is indeed the method of choice here.

As in the continuous case, we aim at using a scalar product, to put our investigations in the setting of Hilbert spaces. Here is the discrete analogue to the one in Def. 6:

Definition 9. *We define the following bilinear form:*

$$\ll f, g \gg := \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_j)\overline{g(x_j)}.$$

Lemma 10. *There holds that*

$$\ll e^{ikx}, e^{ilx} \gg = \begin{cases} (-1)^{\alpha\mathcal{M}}, & k,l \in \mathbb{Z} \text{ and } k-l = \alpha\mathcal{M}, \text{ for } \alpha \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}$$

In particular, this means that only the functions with frequencies in $-\mathcal{M}/2 \leq k \leq \mathcal{M}/2$ (if $\mathcal{M}$ is odd) or $-\mathcal{M}/2 < k \leq \mathcal{M}/2$ (if $\mathcal{M}$ is even) form a basis (and then, this basis is orthogonal).

Proof.

$$\begin{aligned} \ll e^{ikx}, e^{ilx} \gg &= \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} e^{ikx_j} e^{-ilx_j} = \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} e^{ik(-\pi+jh)} e^{-il(-\pi+jh)} \\ &= \frac{e^{i\pi(l-k)}}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} e^{i(k-l)jh} = \frac{e^{i\pi(l-k)}}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} \left(e^{i(k-l)h}\right)^j \end{aligned}$$

Now let $(k-l)$ not be a multiple of $\mathcal{M}$. (Note that $h = \frac{2\pi}{\mathcal{M}}$.) Then, the second exponential is different from one, and one can compute using the geometric sum:

$$= \frac{e^{i\pi(l-k)}}{\mathcal{M}} \frac{1 - \left(e^{i(k-l)h}\right)^{\mathcal{M}}}{1 - e^{i(k-l)h}} = \frac{e^{i\pi(l-k)}}{\mathcal{M}} \frac{1 - e^{i(k-l)2\pi}}{1 - e^{i(k-l)h}} = 0$$

If, indeed, $(k-l) = \alpha\mathcal{M}$, then $e^{i(k-l)h} = e^{i\alpha 2\pi} = 1$, and the sum cancels with the $\frac{1}{\mathcal{M}}$. It is straightforward to see that the factor $e^{i\pi(l-k)}$ is then equal to $(-1)^{\alpha\mathcal{M}}$. □

Finally, we are able to define the following trigonometric polynomial that mimics the continuous Fourier transform:

$$\begin{aligned} s_N^h(x) &= \sum_{k=-N}^{N} \widehat{f}_h(k) e^{ikx}, \qquad N < \mathcal{M}/2, \\ s_N^h(x) &= \sum_{k=-N+1}^{N} \widehat{f}_h(k) e^{ikx}, \qquad N = \mathcal{M}/2. \end{aligned}$$

In order to avoid the strange $(-1)^{\alpha\mathcal{M}}$ term from La. 10, we from now on assume that

$$\boxed{\mathcal{M} \text{ is even.}}$$

Note the close relationship to s_N defined previously as $s_N = \sum_{k=-N}^{N} \widehat{f}(k)e^{ikx}$. This is even more so because of the following lemma:

Lemma 11. *Let $f \in L^2_\circ$ be at least continuous, so one can talk abount pointwise values, and let t be any other trigonometric polynomial of form $t(x) = \sum_{k=-N+o}^{N} t_k e^{ikx}$, where $o = 1$ if $N = \mathcal{M}/2$ and $o = 0$ otherwise. Then, there holds:*

$$\ll f - s_N^h, f - s_N^h \gg \leq \ll f - t, f - t \gg .$$

Proof. Going back, one can see that $\widehat{f_h}(k) = \ll f, e^{ikx} \gg$. Furthermore, because of La. 10, the e^{ikx} are orthonormal in the range of k considered here. Then, the result follows easily from standard orthogonal projection. □

What is extremely important is the fact that for $N = \mathcal{M}/2$ (and so $\mathcal{M}$ even), s_N^h even *interpolates* f:

Theorem 10. *Let $N = \mathcal{M}/2$ and $\mathcal{M}$ even. Furthermore, let f be continuous. Then, there holds $\ll f - s_N^h, f - s_N^h \gg = 0$, i.e.,*

$$s_N^h(x_j) = f(x_j), \qquad 0 \leq j \leq \mathcal{M}.$$

Proof. Because the e^{ikx} are orthonormal for $-N < k \leq N$, one can define the vectors

$$y_k := [e^{ikx_0}, e^{ikx_1}, \ldots e^{ikx_{\mathcal{M}-1}}], \quad -N < k \leq N.$$

Necessarily, they must be linearly independent. Thus, there exists a solution $\vec{\alpha}$ to the linear system of equations

$$\sum_{k=-N+1}^{N} \alpha_k y_k = [f(x_0), \ldots, f(x_{\mathcal{M}-1})].$$

Defining the trigonometric polynomial $t(x) := \sum_{k=-N+1}^{N} \alpha_k e^{ikx}$ yields that $t(x_j) = f(x_j)$, and therefore

$$\ll f - t, f - t \gg = 0.$$

Obviously, this must be the best approximation, thus $t = s_N^h$. □

One problem that one comes across when only using discrete values of x is the so-called aliasing effect. The next remark will make this clear:

Remark 8.
- *Watch the video* https://www.youtube.com/watch?v=ByTsISFXUoY*. What do you see?*
- *Consider the 'clock' slides: Imagine you observe a clock over three minutes, and you photograph it every 45 seconds. You show only the photographs to someone. This means that the observer sees $\mathcal{M} = 4$. This means that he can only observe frequencies $-1 \leq k \leq 2$. However, what actually happens is $k = 3$ (three rounds in those three minutes). Because of $\mathcal{M} = 4$, our brain can't tell $k = 3$ from $k = -1$ from $k = -5$ from $k = 7$... It therefore chooses $k = -1$, which is one of the observable frequencies: Our brain knows discrete Fourier analysis!*

The essence of the next lemma is the quantification of the example above:

Lemma 12. *Let $f \in L^2_\circ$ be at least continuous. Then, there holds:*

$$\widehat{f}_h(k) = \sum_{l\in\mathbb{Z}} \widehat{f}(k + l\mathcal{M}), \qquad -\mathcal{M}/2 < k \leq \mathcal{M}/2.$$

Proof. One can easily compute

$$\begin{aligned}\widehat{f}_h(k) &= \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} \left(\sum_{l\in\mathbb{Z}} \widehat{f}(l) e^{ilx_j} \right) e^{-ikx_j} = \sum_{l\in\mathbb{Z}} \widehat{f}(l) \sum_{j=0}^{\mathcal{M}-1} \frac{e^{ilx_j} e^{-ikx_j}}{\mathcal{M}} \\ &= \sum_{l\in\mathbb{Z}} \widehat{f}(l) \ll e^{ilx}, e^{ikx} \gg = \sum_{l\in\mathbb{Z}} \widehat{f}(k + l\mathcal{M})\end{aligned}$$

because of La. 10 and the assumption that $\mathcal{M}$ is even. □

To conclude our investigations on the discrete Fourier analysis, we state and prove a theorem on the approximation quality of s_N^h.

Theorem 11. *Let $f \in L^2_\circ$ be sufficiently regular, i.e., $f \in C^s$ for an $s \geq 1$, and all derivatives of f and f itself are assumed to have bounded variation; and let s_N^h be the discrete Fourier transformation to f; with $N \leq \mathcal{M}/2$. Then, there holds*

$$\|f - s_N^h\|_{L^2_\circ} \leq CN^{-s}.$$

The constant C depends on s and the function f (and its derivatives), but not on N.

Remark 9. *The theorem means that with $\mathcal{M} \to \infty$ and $N \to \infty$, there is indeed convergence of s_N^h toward f in the L^2-norm. In the L^∞-norm, there is not necessarily guaranteed convergence.*

Proof of Theorem 11. For simplicity, we only show the proof for the case that $N = \mathcal{M}/2$. The difference $f - s_N^h$ has the Fourier representation

$$(f - s_N^h)(x) = \sum_{k\in\mathbb{Z}} \beta_k e^{ikx}, \qquad \beta_k = \begin{cases} \widehat{f}(k) - \widehat{f}_h(k), & -N < k \leq N \\ \widehat{f}(k), & \text{otherwise} \end{cases}.$$

From the aliasing lemma 12, we obtain that

$$\widehat{f}_h(k) - \widehat{f}(k) = \sum_{l\in\mathbb{Z}^{\neq 0}} \widehat{f}(k + 2lN), \quad 1 - N \leq k \leq N,$$

and consequently for $1 - N \leq k \leq N$:

$$\begin{aligned}&\left|\widehat{f}(k) - \widehat{f}_h(k)\right|^2 \leq \left(\sum_{l\in\mathbb{Z}^{\neq 0}} \left|\widehat{f}(k + 2lN)\right| \right)^2 \\ &= \left(\sum_{l\in\mathbb{Z}^{\neq 0}} |k + 2lN|^{-s} |k + 2lN|^s \left|\widehat{f}(k + 2lN)\right| \right)^2 \leq \sum_{l\in\mathbb{Z}^{\neq 0}} |k + 2lN|^{-2s} \sum_{l\in\mathbb{Z}^{\neq 0}} |k + 2lN|^{2s} \left|\widehat{f}(k + 2lN)\right|^2 \\ &\leq \sum_{l\in\mathbb{Z}^{\neq 0}} \left|2|l|N - N\right|^{-2s} \sum_{l\in\mathbb{Z}^{\neq 0}} |k + 2lN|^{2s} \left|\widehat{f}(k + 2lN)\right|^2 \\ &\leq N^{-2s} \sum_{l\in\mathbb{Z}^{\neq 0}} |2|l| - 1|^{-2s} \sum_{l\in\mathbb{Z}^{\neq 0}} |k + 2lN|^{2s} \left|\widehat{f}(k + 2lN)\right|^2\end{aligned}$$

From that follows that

$$\sum_{k=-N+1}^{N} |\widehat{f}(k) - \widehat{f}_h(k)|^2 \le CN^{-2s} \sum_{k=-N+1}^{N} \sum_{l\in\mathbb{Z}^{\neq 0}} |k+2lN|^{2s} |\widehat{f}(k+2lN)|^2$$
$$\le CN^{-2s} \sum_{\nu\in\mathbb{Z}} v^{2s} |\widehat{f}(\nu)|^2 = CN^{-2s} \|f^{(s)}\|^2_{L^2_\circ}.$$

Now, for $|k| \ge N$, we can use a similar, but simpler trick, to obtain:

$$\sum_{|k|=N}^{\infty} |\widehat{f}(k)|^2 \le \sum_{|k|=N}^{\infty} \left(\frac{|k|}{N}\right)^{2s} |\widehat{f}(k)|^2 \le N^{-2s} \sum_{k\in\mathbb{Z}} |k|^{2s} |\widehat{f}(k)|^2.$$

This concludes the proof.

□

We conclude this chapter with the famous *fast Fourier transform* (FFT), which is a way of efficiently computing the values $\widehat{f}_h(k)$ in (9). The FFT is an example of a *divide and conquer* algorithm, which means that it works by splitting up a 'larger' problem into 'smaller' subproblems. More precisely, we will split the problem of computing the discrete Fourier coefficients on a grid with $\mathcal{M}$ elements into computing twice the discrete Fourier coefficients on a grid with $\frac{\mathcal{M}}{2}$ elements. We will therefore assume that $\mathcal{M}$ is a power of two.

Let us first discuss the costs of computing $\widehat{f}_h(k)$ with a 'naive' approach. For a fixed k, this takes $\mathcal{O}(\mathcal{M})$ operations. One has to compute $\widehat{f}_h(k)$ for $k \in \{-\mathcal{M}/2+1, \dots, \mathcal{M}/2\}$, so $\mathcal{M}$ times. Summing up, this amounts to $\mathcal{O}(\mathcal{M}^2)$ operations. We will show that the FFT is able to perform the operation in $\mathcal{O}(\mathcal{M}\log(\mathcal{M}))$ operations, which is a substantial saving for large $\mathcal{M}$.

Let us consider the discrete Fourier coefficients on a grid with $2\mathcal{M}$ points, so the spacing between the points is $h/2$. Note that they have to be computed for $k \in \{-\mathcal{M}+1, \dots, \mathcal{M}\}$, so $2\mathcal{M}$ times.

$$\widehat{f}_{h/2}(k) = \frac{1}{2\mathcal{M}} \sum_{j=0}^{2\mathcal{M}-1} f(x_j) e^{-ikx_j} = \frac{1}{2} \left(\frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_{2j}) e^{-ikx_{2j}} + \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_{2j+1}) e^{-ikx_{2j+1}} \right)$$

The first sum is precisely the discrete Fourier coefficient on the grid with $\mathcal{M}$ points. For the second sum, this is not true, because the x_{2j+1} do not start at $-\pi$. So we have to translate them to get

$$= \frac{1}{2} \left(\widehat{f}_h(k) + \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_{2j}+h/2) e^{-ik(x_{2j}+h/2)} \right)$$
$$= \frac{1}{2} \left(\widehat{f}_h(k) + \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_{2j}+h/2) e^{-ikx_{2j}} e^{-ik\frac{\pi}{\mathcal{M}}} \right) = \frac{1}{2} \left(\widehat{f}_h(k) + e^{-ik\frac{\pi}{\mathcal{M}}} \widehat{g}_h(k) \right),$$

where $g(x) := f(x+h/2)$. We now have to compute the values of $\widehat{f}_h$ and $\widehat{g}_h$ for $k \in \{-\mathcal{M}/2+1, \dots, \mathcal{M}/2\}$, so $\mathcal{M}$ times. For k not in this range, we can use the following relation:

$$\widehat{f}_h(k+\mathcal{M}) = \frac{1}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_j) e^{-ikx_j - i\mathcal{M}x_j} = \frac{(-1)^{\mathcal{M}}}{\mathcal{M}} \sum_{j=0}^{\mathcal{M}-1} f(x_j) e^{-ikx_j} = (-1)^{\mathcal{M}} \widehat{f}_h(k).$$

So let us summarize: Instead of computing $2\mathcal{M}$ discrete Fourier coefficients on a grid with $2\mathcal{M}$ points, we compute two times $\mathcal{M}$ coefficients on a grid with $\mathcal{M}$ points. Of course we can now recursively (because $\mathcal{M}$ was supposed to be a power of 2) apply this algorithm to the coefficients on the smaller grid. If we define $E(\mathcal{M})$ to be the effort that we spend with the algorithm on a grid with $\mathcal{M}$ elements, we have the recursive relation

$$E(\mathcal{M}) = 2E\left(\frac{\mathcal{M}}{2}\right) + \mathcal{O}(\mathcal{M}).$$

Let us suppose that $\mathcal{M} = 2^p$. Then, we can compute

$$\begin{aligned} E(2^p) &= 2E(2^{p-1}) + \mathcal{O}(2^p) = 2\left(2E(2^{p-2}) + \mathcal{O}(2^{p-1})\right) + \mathcal{O}(2^p) \\ &= 2^2E(2^{p-2}) + \mathcal{O}(2^p) + \mathcal{O}(2^p) = \ldots = 2^pE(1) + \mathcal{O}(p2^p) \\ &= \mathcal{O}(\mathcal{M}) + \mathcal{O}(\log(\mathcal{M})\mathcal{M}) = \mathcal{O}(\log(\mathcal{M})\mathcal{M}). \end{aligned}$$

This is of course a large benefit in comparison to the original $\mathcal{O}(\mathcal{M}^2)$.

Remark 10. *There exist other ways of computing the FFT so that $\mathcal{M}$ doesn't necessarily have to be a power of two. For brevity, we are not going to discuss this here.*

3 Linear operators on Banach spaces

In this section, we discuss *continuous* linear operators on Banach spaces. If the space is finite-dimensional, linearity implies continuity:

Example 8. *Assume that* $\dim(\mathcal{X}) = n < \infty$ *and that* $T : \mathcal{X} \to \mathcal{Y}$ *is a linear operator. Obviously, this means that there is a basis* $e_1, \ldots, e_n$ *of* $\mathcal{X}$*, and that* Tx *is determined once all the* Te_i *are known. Now consider, for* $x, \delta \in \mathcal{X}$*,* $\delta = \sum_{i=1}^n \delta_i e_i$ $(\delta_i \in \mathbb{K})$

$$\|T(x+\delta) - T(x)\| = \|T(\delta)\| \leq \max_{i=1,\ldots,n} \|Te_i\| \sum_{i=1}^n |\delta_i| \leq C \sum_{i=1}^n |\delta_i|.$$

With $\delta \to 0$ *holds* $\delta_i \to 0$ *for all* i *(try to prove this!), and so the function* T *is continuous.*

We will show a bit below (see La. 15) that also the inverse is true: If all linear operators are continuous, then the dimension of $\mathcal{X}$ *is finite.*

It is evident that the example shown above does not hold for infinite-dimensional spaces, as the maximum of the Te_i does not necessarily have to exist. In fact, there are linear operators that are not continuous once $\mathcal{X}$ is infinite-dimensional.

Definition 10. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be normed spaces and* $T : \mathcal{X} \to \mathcal{Y}$ *a linear operator.*

- T *is called bounded, if there exists a* $\lambda \in \mathbb{R}$*, such that*

$$\|Tx\| \leq \lambda \|x\|, \qquad \forall x \in \mathcal{X}\,.$$

- *The operator norm for a bounded operator is defined as the minimal* λ*, so*

$$\|T\| := \inf\{\lambda \in \mathbb{R} : \|Tx\| \leq \lambda \|x\| \quad \forall x \in \mathcal{X}\}.$$

Lemma 13. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be normed spaces, and let* $T : \mathcal{X} \to \mathcal{Y}$ *be a bounded linear operator. There holds:*

$$\|Tx\| \leq \|T\| \|x\|$$

and

$$\|T\| = \sup_{x \in \mathcal{X} \setminus \{0\}} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

Proof. It is obvious that there holds $\|Tx\| \leq \|T\|\|x\|$ due to the definition of $\|T\|$. Furthermore, due to linearity, there holds

$$\sup_{x \in \mathcal{X} \setminus \{0\}} \frac{\|Tx\|}{\|x\|} = \sup_{x \in \mathcal{X} \setminus \{0\}} \left\| T\left(\frac{x}{\|x\|}\right) \right\| = \sup_{\|x\|=1} \|Tx\|.$$

Set $\tau := \sup_{x \in \mathcal{X} \setminus \{0\}} \frac{\|Tx\|}{\|x\|}$. Of course there holds $\|Tx\| \leq \tau \|x\|$, and therefore, $\tau \geq \|T\|$. On the other hand side, there holds, due to the definition of $\|T\|$ that

$$\frac{\|Tx\|}{\|x\|} \leq \|T\|$$

and hence $\tau \leq \|T\|$. This concludes the proof. $\square$

Theorem 12. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be normed spaces, and let* $T : \mathcal{X} \to \mathcal{Y}$ *be a linear operator. Then, the following are equivalent:*

1. T is absolutely continuous.

2. T is continuous at $x = 0$.

3. T is continuous for $x \in \mathcal{X}$.

4. T is bounded.

Proof. We show (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (1). Obviously, (1) $\Rightarrow$ (2) is trivial.

- (2) $\Rightarrow$ (3): Let $x, \delta \in \mathcal{X}$. There holds

$$T(x + \delta) - T(x) = T(\delta).$$

As T is continuous at 0, $T(\delta) \to 0$ for $\delta \to 0$, which means that T is continuous in x.

- (3) $\Rightarrow$ (4): Assume that T is not bounded. Due to La. 13, there exists $(x_n)_{n\in\mathbb{N}} \subset \mathcal{X}$, $\|x_n\| = 1$ such that $\|Tx_n\| \to \infty$. Let us choose $(x_n)_{n\in\mathbb{N}}$ (maybe by throwing out some elements) in such a way that

$$\|Tx_n\| \geq n.$$

Then, for $(\widetilde{x}_n)_{n\in\mathbb{N}} := (x_n/n)_{n\in\mathbb{N}}$, there holds

$$\|T\widetilde{x}_n\| \geq 1.$$

Furthermore, $\widetilde{x}_n \to 0$. But due to the continuity of T, there must hold

$$\|T(\widetilde{x}_n) - T(0)\| = \|T(\widetilde{x}_n)\| \to 0,$$

which is a contradiction.

- (4) $\Rightarrow$ (1): Because T is bounded, we can compute for $x, \delta \in \mathcal{X}$

$$\|T(x + \delta) - T(x)\| = \|T(\delta)\| \leq \|T\|\|\delta\|,$$

which is uniform continuity.

□

Before we continue, we consider the following important linear operator, the continuous Fourier transform.

Definition 11 (Continuous Fourier transform)**.** *Let X and Y be function spaces on $\mathbb{R}$ (think of $L^p(\mathbb{R})$), and let the linear operator $\mathcal{F}$ be defined as*

$$\mathcal{F} : X \to Y, \qquad f \mapsto \mathcal{F}f, \text{ with } \mathcal{F}f(\xi) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-i\xi x}\mathrm{d}x. \tag{10}$$

Typically, $\mathcal{F}f$ is denoted by $\widehat{f}$. $\widehat{f}$ is called the continuous Fourier transform *of f.*

Remark 11. *It is not clear that $\mathcal{F}f$ is well-defined (or even continuous), this depends on X and Y.*

From our previous investigations, it seems obvious to use $X = L^2(\mathbb{R})$ as underlying function space for $\mathcal{F}$. However, as $e^{i\xi x}$ is not in L^2, the integral (10) does not necessarily exist for a function $f \in L^2(\mathbb{R})$. (There will be an L^2-theory later, however.) Thus, we start with a function $f \in L^1(\mathbb{R})$ (there $\widehat{f}(\xi)$ exists because of Hölder), and we state the following lemma, which is the equivalent of the Riemann-Lebesgue lemma 8:

Lemma 14. *For an $f \in L^1(\mathbb{R})$, the Fourier transform $\mathcal{F} f \equiv \widehat{f}$ exists and is continuous (so we can take $Y = C^0(\mathbb{R})$ here). Furthermore, it vanishes for $\xi \to \pm\infty$.*

Proof. Hölders inequality for this case is $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$. Because $e^{-i\xi x}$ is bounded and $f \in L^1(\mathbb{R})$, existence follows easily. For continuity, we consider

$$\widehat{f}(\xi + \delta) - \widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \left(e^{-i(\xi+\delta)x} - e^{-i\xi x} \right) \mathrm{d}x.$$

Obviously, the part $f(x) \left(e^{-i(\xi+\delta)x} - e^{-i\xi x} \right)$ converges pointwise almost everywhere to zero as δ vanishes. Furthermore,

$$\left| f(x) \left(e^{-i(\xi+\delta)x} - e^{-i\xi x} \right) \right| \leq 2|f(x)|,$$

which is integrable. With Lebesgue's dominated convergence theorem, there follows that $\widehat{f}(\xi + \delta) \to \widehat{f}(\xi)$ for $\delta \to 0$, and hence $\widehat{f}$ is continuous. That the integral vanishes can be proven in a similar fashion as in La. 19, see below. □

We will see later (see Cor. 5) that for $X = L^2(\mathbb{R})$, the Fourier transform can be defined through a limit argument and that there holds

$$\| \mathcal{F} f \|_{L^2} = \|f\|_{L^2},$$

similar to Parseval's identity La. 7. Hence, Y can be taken as $L^2(\mathbb{R})$, and $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a continuous linear operator. We come back at $\mathcal{F}$ in more details in the next section.

Lemma 15. *Let $\mathcal{X}$ be a Hilbert space. If every linear operator $T : \mathcal{X} \to \mathbb{K}$ is continuous, then $\dim(\mathcal{X}) < \infty$.*

Proof. We prove this indirectly by showing that if $\dim(\mathcal{X}) = \infty$, then one can construct a linear operator that is not continuous.

So assume that $\dim(\mathcal{X}) = \infty$. Because $\mathcal{X}$ is not finite-dimensional, one can find a countable set of linearly independent functions $(e_i)_{i \in \mathbb{N}}$, and let $\mathcal{Y}$ be the span of these functions. Wlog, we can assume that $\|e_i\| = 1$. With means that we don't have at our disposal yet, see [1, p. 314] one can show that there exists a subspace $\mathcal{Z}$ such that

$$\mathcal{X} = \mathcal{Y} \oplus \mathcal{Z}.$$

Every value x can thus in a unique way be written as $x = y + z$, with $z \in \mathcal{Z}$ and

$$y = \sum_{i \in \mathbb{N}} \alpha_i e_i \in \mathcal{Y},$$

for $\alpha_i \in \mathbb{K}$. Now define an operator $T : \mathcal{X} \to \mathbb{K}$ by

$$Tx := T(y + z) := Ty := \sum_{i \in \mathbb{N}} i \alpha_i.$$

Obviously, T is linear. This operator is not continuous. To see this, note that

$$Te_i = i, \qquad \forall i \in \mathbb{N}\,.$$

Thus, there cannot exist a finite value $\|T\|$ such that $\|Tx\| \leq \|T\| \|x\|$ for all $x \in \mathcal{X}$. This means, due to the previous theorem, that T is not continuous. □

In the following, we seek answer to the question whether, and in what sense, pointwise convergence of continuous linear functions leads to continuous functions.

Theorem 13 (Uniform boundedness principle, Banach-Steinhauss)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, and A an index set. We consider a family $(T_\alpha)_{\alpha\in A}$ of linear operators from $\mathcal{X}$ to $\mathcal{Y}$ that is pointwise bounded, i.e.,*

$$\sup_{\alpha\in A}\|T_\alpha x\| < \infty, \qquad \forall x \in \mathcal{X}\,.$$

Then, there holds

$$\sup_{\alpha\in A}\|T_\alpha\| < \infty.$$

Proof. Let a bounded linear operator $T : \mathcal{X} \to \mathcal{Y}$ and an $x \in \mathcal{X}$, $r \in \mathbb{R}^{>0}$ be given. Then, there holds for any $h \in \mathcal{X}$ with $\|h\| \le r$:

$$\|Th\| \le \sup_{\delta\in\mathcal{X},\|\delta\|\le r}\|T(x+\delta)\|.$$

(Prove this!) Consequently, there holds for any $r > 0$

$$r\|T\| \le \sup_{y\in B(x,r)}\|Ty\|, \tag{11}$$

where we have defined $B(x,r) := \{z \in \mathcal{X} : \|z - x\| \le r\}$.

We give an indirect proof and assume that $\sup_{\alpha\in A}\|T_\alpha\| = \infty$. We will show that we can find a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ such that for the limit $\overline{x}$, there holds $\sup_{\alpha\in A}\|T_\alpha\overline{x}\| = \infty$, which of course contradicts the assumption. To this end, choose some sequence $(\alpha_n)_{n\in\mathbb{N}}$ such that

$$\|T_{\alpha_n}\| \ge 4^n.$$

Define $x_0 := 0$. Due to (11), we can find an x_1 with

$$\|x_1 - x_0\| \le \frac{1}{3}, \qquad \|T_{\alpha_1}x_1\| \ge \theta\frac{1}{3}\|T_{\alpha_1}\| \ge \theta\frac{4}{3}.$$

for some $\frac{1}{2} < \theta < 1$. Note that this factor θ is rather arbitrarily, it should be a bit less than one, so we don't have to prove that the supremum is being reached. We continue this process. Again, due to (11), there is x_2 such that

$$\|x_2 - x_1\| \le \frac{1}{3^2}, \qquad \|T_{\alpha_2}x_2\| \ge \theta\frac{1}{3^2}\|T_{\alpha_2}\| \ge \theta\frac{4^2}{3^2}.$$

Continuing, one finds

$$\|x_n - x_{n-1}\| \le \frac{1}{3^n}, \qquad \|T_{\alpha_n}x_n\| \ge \theta\frac{1}{3^n}\|T_{\alpha_n}\| \ge \theta\frac{4^n}{3^n}.$$

$(x_n)_{n\in\mathbb{N}}$ is Cauchy, because for $n < m$,

$$\|x_n - x_m\| = \left\|\sum_{j=n}^{m-1} x_{j+1} - x_j\right\| \le \sum_{j=n}^{m-1}\|x_{j+1} - x_j\| \le \sum_{j=n+1}^{m}\frac{1}{3^j} = \frac{1}{2}\left(\frac{1}{3^n} - \frac{1}{3^m}\right).$$

Hence, there exists a limit $\overline{x}$. For this limit, there holds

$$\|x_n - \overline{x}\| \le \frac{1}{2}\frac{1}{3^n}.$$

We can compute

$$
\begin{aligned}
\|T_{\alpha_n}\overline{x}\| &= \|T_{\alpha_n}(\overline{x}-x_n)+T_{\alpha_n}x_n\| \geq \|T_{\alpha_n}x_n\| - \|T_{\alpha_n}(\overline{x}-x_n)\| \\
&\geq \theta\frac{1}{3^n}\|T_{\alpha_n}\| - \|T_{\alpha_n}\|\|\overline{x}-x_n\| \\
&\geq \|T_{\alpha_n}\|\left(\theta\frac{1}{3^n} - \|\overline{x}-x_n\|\right) \\
&\geq \|T_{\alpha_n}\|\left(\theta\frac{1}{3^n} - \frac{1}{2}\frac{1}{3^n}\right) \geq 4^n\left(\theta\frac{1}{3^n} - \frac{1}{2}\frac{1}{3^n}\right) \to \infty.
\end{aligned}
$$

This of course contradicts the assumptions. □

With this, one can prove that pointwise convergence implies convergence of linear operators:

Corollary 3. *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, and let $(T_n)_{n\in\mathbb{N}}$, $T_n : \mathcal{X} \to \mathcal{Y}$, be a sequence of continuous linear functions with the pointwise limit existing, i.e.,*

$$
Tx := \lim_{n\to\infty} T_n x
$$

exists for all $x \in \mathcal{X}$. Then $T : \mathcal{X} \to \mathcal{Y}$ is a continuous linear operator with norm

$$
\|T\| \leq \liminf_{n\to\infty} \|T_n\| < \infty.
$$

Proof. Exercise. Use the previous theorem and the fact that $\|T_n\|$ must be bounded, and so the $\liminf_{n\in\mathbb{N}} \|T_n\|$ exits. □

As a last topic in this section, we consider linear operators from $\mathcal{X}$ to $\mathbb{K}$:

Definition 12 (Dual space)**.** *Let $\mathcal{X}$ be a normed vector space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then, the* dual space *is defined as all continuous, linear operators from $\mathcal{X}$ to $\mathbb{K}$, i.e.,*

$$
\mathcal{X}' := \{T : \mathcal{X} \to \mathbb{K},\ T \textit{ is continuous and linear}\}.
$$

Let us fix some $x \in \mathcal{X}$. Then, the operator

$$
\mathcal{X}' \to \mathbb{K} : f \mapsto \langle f, x\rangle := f(x)
$$

is linear. As a consequence, the operator

$$
\langle\cdot,\cdot\rangle : \qquad \mathcal{X}' \times \mathcal{X} \to \mathbb{K} : \qquad (f, x) \mapsto \langle f, x\rangle
$$

is bilinear. This operator is called the duality operator (or duality map).

Corollary 4. *There holds*

$$
|\langle f, x\rangle| \leq \|f\|\|x\|,
$$

which implies continuity of $\langle\cdot,\cdot\rangle$.

Example 9. *We show some examples of dual spaces.*

- *Consider the normed space $(\mathbb{K}^n, \|\cdot\|_p)$. A linear function from $\mathbb{K}^n$ to $\mathbb{K}$ can be represented as a colum vector, so there is an obvious bijection between $(\mathbb{K}^n)'$ and $\mathbb{K}^n$. (Prove this!) For every $f \in (\mathbb{K}^n)'$, there exists thus a unique $y_f \in \mathbb{K}^n$ such that there holds*

$$
|\langle f, x\rangle| = |y_f^T x| \leq \|y_f\|_q\|x\|_p,
$$

and hence $\|f\| \leq \|y_f\|_q$, *with* $\frac{1}{p}+\frac{1}{q}=1$. *We will now show that there also holds* $\|f\| = \|y_f\|_q$. *Let us assume* $p \in (1,\infty)$ *for a moment.*

Our goal is to find a vector x *such that, for a given* y *(please note: due to simplicity, we set* $y \equiv y_f$*), there holds* $|y^T x| = \|y\|_q \|x\|_p$. *To this end, define a set of values* $\lambda_i \in \mathbb{K}$, *with* $|\lambda_i| = 1$ *and* $\lambda_i y_i = |y_i|$. *(*y_i *is the* $i-th$ *component of* y*.) Based on this, one can define a vector* x *through*

$$x_i := \lambda_i |y_i|^{\frac{1}{p-1}}.$$

Note that $q = \frac{p}{p-1}$, *and*

$$\|x\|_p^p = \sum_{i=1}^n \left|\lambda_i |y_i|^{\frac{1}{p-1}}\right|^p = \sum_{i=1}^n |y_i|^{\frac{p}{p-1}} = \sum_{i=1}^n |y_i|^q = \|y\|_q^q.$$

From this equality, we can conclude that $\|y\|_q = \|x\|_p^{p/q}$ *and hence*

$$\|y\|_q^{q-1} = \|x\|_p^p \|x\|_p^{-p/q} = \|x\|_p. \tag{12}$$

Furthermore,

$$y^T x = \sum_{i=1}^n y_i x_i = \sum_{i=1}^n y_i \lambda_i |y_i|^{\frac{1}{p-1}} = \sum_{i=1}^n |y_i|^{\frac{p}{p-1}} = \sum_{i=1}^n |y_i|^q = \|y\|_q \|y\|_q^{q-1}.$$

This last term is, with (12), equivalent to

$$y^T x = \|y\|_q \|x\|_p.$$

Hence, $\|f\| \geq \|y\|_q$. *Together with* $\|f\| \leq \|y\|_q$, *this yields* $\|f\| = \|y\|_q$. *Similar proofs hold for* $p = 1$ *and* $p = \infty$. *We can thus state:*

$$(\mathbb{K}^n, \|\cdot\|_p)' \cong (\mathbb{K}^n, \|\cdot\|_q), \qquad 1 \leq p \leq \infty.$$

- *Something very similar holds for the* l^p *spaces, yet, with a very important restriction. We will show that for* $1 \leq p < \infty$ *(so* $p \neq \infty$*!) the dual of* l^p *is isometrically isomorph to* l^q, *again with* $\frac{1}{p}+\frac{1}{q}=1$. *For* $p=\infty$, *the dual space is larger than* l^1. *Again, as for the* $\mathbb{K}^n$ *case, we define linear operators as*

$$\langle f, x\rangle := \sum_{i\in\mathbb{N}} f_i x_i$$

for $x \in l^p$, *whenever the sum exists. Due to Hölder's inequality, this is the case for* $f \in l^q$. *So, we know that* $l^q \subset (l^p)'$. *Now let* $1 \leq p < \infty$ *and let* $f \in (l^p)'$ *be given. Define*

$$y_k := \langle f, e_k\rangle,$$

with e_k *the* $k-th$ *unit vector in* l^p. *We have to show that this* y *is in* l^q. *We aim to use the Banach-Steinhauss theorem (Thm. 13), and use what we already know about* $\mathbb{K}^n$. *For some given* $x \in l^p$, *define*

$$x^n := \sum_{k=0}^n x_k e_k.$$

It is an easy task to show that $x^n \to x$. *Similarly, define* $f^n \in (l^p)'$ *through*

$$\langle f^n, x\rangle := \sum_{k=0}^n y_k x_k.$$

Note that

$$\langle f, x^n \rangle = \langle f, \sum_{k=0}^{n} x_k e_k \rangle = \sum_{k=0}^{n} x_k \langle f, e_k \rangle = \sum_{k=0}^{n} x_k y_k = \langle f^n, x \rangle.$$

From this, we can conclude

$$\sum x_k y_k = \lim_{n\to\infty} \langle f^n, x \rangle = \lim_{n\to\infty} \langle f, x^n \rangle = \langle f, x \rangle.$$

This is pointwise convergence of f^n towards f. From our knowledge over $\mathbb{K}^n$, it is known that, for any fixed n, there holds

$$\|(y_0, \ldots, y_n, 0, \ldots)\|_{l^q} = \|f^n\|.$$

Due to Banach-Steinhauss, it is therefore evident that

$$\|y\|_{l^q} = \sup_{n\to\infty} \|(y_0, \ldots, y_n, 0, \ldots)\|_{l^q} = \sup_{n\in\mathbb{N}} \|f^n\| < \infty$$

and hence $y \in l^q$. That $\|f\| = \|y\|_{l^q}$ follows with similar arguments as for $\mathbb{K}^n$.

Let us point out that $(l^\infty)' \neq l^1$. To this end, we show that not every bounded linear function $f : l^\infty \to \mathbb{K}$ can be written as

$$f(x) = \sum_{k\in\mathbb{N}} x_k y_k.$$

We define such a function f first on the subset of convergent *series through $f(x) := \lim_{n\to\infty} x_n$. This is (on the subspace of convergent series) a continuous linear functional. Through means that we do not have at our disposal (cf. Hahn-Banach theorem, not covered here), it is possible to extend this to the whole of l^∞. Obviously, $f(e_k) = 0$, and thus $y_k = 0$ for all $k \in \mathbb{N}$ if we would assume a representation as above. However, $f(1, 1, 1, \ldots) = 1$. This contradicts $y_k = 0$.*

- *Without going into details, we only mention that the same holds for L^p spaces, i.e., for $1 \leq p < \infty$, there holds $(L^p)' = L^q$. The duality is via integration.*

Let us now additionally assume that $\mathcal{X}$ is a Hilbert space. Due to Cauchy-Schwarz, there holds for any $x, y \in \mathcal{X}$ that

$$|(x, y)| \leq \|x\|\|y\|,$$

hence the function

$$R : \mathcal{X} \to \mathcal{X}', \qquad x \mapsto (\cdot, x)$$

is well-defined. (Note that it must be $(\cdot, x)$ and not the other way around. Why?) The operator R is 'conjugate linear', meaning that there holds for all $x, y \in \mathcal{X}$, $\alpha \in \mathbb{K}$

$$R(x + y) = R(x) + R(y), \qquad R(\alpha x) = \overline{\alpha} R(x).$$

There holds:

Lemma 16. *The operator R defined above is continuous and there holds*

$$\|R(x)\| = \|x\|.$$

Proof. We first show the norm-conserving property. It is trivial for $x = 0$, so let $x \neq 0$ in the sequel. There holds

$$\|R(x)\| = \sup_{\|y\|=1} (y, x) \leq \sup_{\|y\|=1} \|y\|\|x\| = \|x\|,$$
$$\|R(x)\| = \sup_{\|y\|=1} (y, x) \geq \left(\frac{x}{\|x\|}, x \right) = \|x\|$$

and hence $\|R(x)\| = \|x\|$. Now,

$$\|R(x) - R(y)\| = \|R(x - y)\| = \|x - y\|, \tag{13}$$

R is thus continuous. □

Theorem 14. *The operator R is a conjugate linear isometric isomorphism. We call R the* Riesz-*isomorphism.*

Proof. We already know that R conserves the norm and is continuous. Furthermore, due to (13), R is injective. We only have to show that, given some $f \in \mathcal{X}'$, there exists a $y \in \mathcal{X}$, such that $R(y) = f$. We can safely assume that $f \neq 0$. f is by definition continuous, and hence the kernel of f, i.e.,

$$\text{kernel}(f) := \{x \in \mathcal{X}, f(x) = 0\}$$

is a non-empty closed linear subspace. For closed linear subspaces in Hilbert spaces, there holds

$$\mathcal{X} = \text{kernel}(f) \oplus \text{kernel}(f)^{\perp}.$$

Because $f \neq 0$, there must be an element $z \in \text{kernel}(f)^{\perp}$ with $\|z\| = 1$. Observe that there holds for all $x \in \mathcal{X}$:

$$-\langle f, x\rangle z + \langle f, z\rangle x \in \text{kernel}(f).$$

(Prove this!) Hence,

$$0 = (-\langle f, x\rangle z + \langle f, z\rangle x, z) = -\langle f, x\rangle (z, z) + \langle f, z\rangle (x, z) = -\langle f, x\rangle + \langle f, z\rangle (x, z).$$

Choose

$$y := \overline{\langle f, z\rangle} z.$$

Then, there holds $R(y) = f$. □

4 Functions on the real line: Fourier transform

We hae treated the case of periodic functions earlier, which is of course very special. Therefore, in this chapter, we extend the idea of Fourier *series* to Fourier *transformations*, defined for a suitably integrable function (we will define this more precisely later) on the real line.

4.1 Continuous Fourier transform

We have already defined the continuous Fourier transform in Def. 11 as the function $\mathcal{F}f \equiv \widehat{f}$, with

$$\widehat{f}(\xi) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-i\xi x}\mathrm{d}x$$

for $\xi \in \mathbb{R}$. Note also that of the factor $\frac{1}{2\pi}$, we chose to use the factor $\frac{1}{\sqrt{2\pi}}$. Reasons will become clearer in the sequel. In literature, there are also often different choices. Before going further, we start with an example and some easily obtainable technical results:

Example 10. *The function $f(x) := \chi_{[-a,a]}$ has the following Fourier transform:*

$$\begin{aligned}\widehat{f}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-i\xi x}\mathrm{d}x = \frac{-1}{\sqrt{2\pi}i\xi}\, e^{-i\xi x}\Big|_{-a}^{a} = \frac{e^{i\xi a} - e^{-i\xi a}}{\sqrt{2\pi}i\xi} = \frac{2\sin(\xi a)}{\sqrt{2\pi}\xi} = \frac{\sqrt{2}}{\sqrt{\pi}}\frac{\sin(\xi a)}{\xi}\\ &= \frac{\sqrt{2}a}{\sqrt{\pi}}\,\mathrm{sinc}(a\xi).\end{aligned}$$

Note that the function $\widehat{f}$ is continuous, yet it is not in $L^1(\mathbb{R})$.

Lemma 17. *Let $f \in L^1(\mathbb{R})$, and let a be a fixed constant in $\mathbb{R}$. Then, the following items hold:*

1. *Define $g(x) := f(ax)$. Then, $\widehat{g}(\xi) = \frac{\widehat{f}(\xi/a)}{|a|}$.*
2. *Define $g(x) := f(x)e^{iax}$. Then, $\widehat{g}(\xi) = \widehat{f}(\xi - a)$.*
3. *Let h be a function in $L^1(\mathbb{R})$, then $\int_{\mathbb{R}} f(\tau)\widehat{h}(\tau)\mathrm{d}\tau = \int_{\mathbb{R}} \widehat{f}(\tau)h(\tau)\mathrm{d}\tau$.*

Proof. Exercise. □

As an interesting example, and because we are going to need it in the sequel, we compute the Fourier transform of the following function:

Lemma 18. *The fourier transform of $f(x) := \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ is $\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\xi^2}{2}}$, i.e., the function is invariant under the Fourier transformation. f is hence a fixed point of $\mathcal{F}$, as there holds $\mathcal{F}f = f$.*

Proof. We refer to Fig. 1 for the used notation. Without loss of generality, we assume that $a > \xi > 0$ and define $g(z) := e^{-z^2/2}$, which is an analytic function. Cauchy integral theorem states

$$\int_{\partial R} g(z)\mathrm{d}z = 0,$$

and consequently

$$\int_{\sigma_1} g(z)\mathrm{d}z = \int_{\sigma_0} g(z)\mathrm{d}z + \int_{\gamma_+} g(z)\mathrm{d}z - \int_{\gamma_-} g(z)\mathrm{d}z.$$

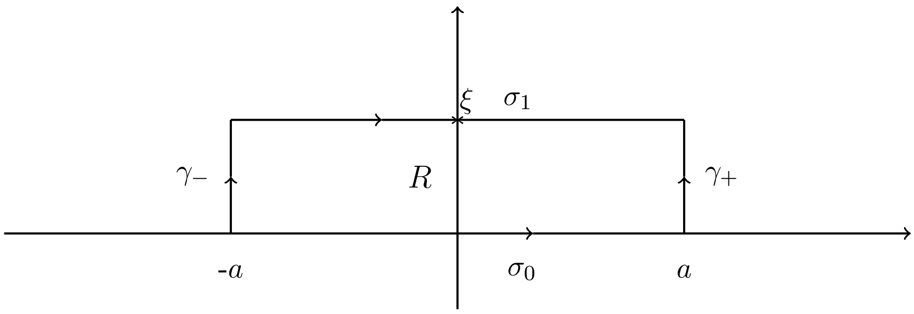


Figure 1: Notation for the proof of La. 18.

Now let us compute the integrals individually:

$$
\begin{aligned}
I_1 :&= \int_{\sigma_1} g(z)\mathrm{d}z = \int_{-a}^{a} e^{-(z+i\xi)^2/2}\mathrm{d}z = \int_{-a}^{a} e^{\xi^2/2}e^{-z^2/2}e^{-i\xi z}\mathrm{d}z \\
&= e^{\xi^2/2}\left(2\pi\widehat{f}(\xi) + o(1)\right), \\
I_0 :&= \int_{\sigma_0} g(z)\mathrm{d}z = \int_{-a}^{a} e^{-x^2/2}\mathrm{d}z = \sqrt{2\pi} + o(1), \\
I_\pm :&= \int_{\gamma_\pm} g(z)\mathrm{d}z = i\int_0^{\xi} e^{-(\pm a+it)^2/2}\mathrm{d}t = i\int_0^{\xi} e^{(-a^2+t^2)/2}e^{\pm iat}\mathrm{d}t \\
&= i\int_0^{\xi} e^{(t-a)(t+a)/2}e^{\pm iat}\mathrm{d}t, \\
|I_\pm| &\leq \int_0^{a} e^{(t-a)(t+a)/2}\mathrm{d}t \leq \int_0^{a} e^{(t-a)a/2}\mathrm{d}t = e^{-a^2/2}\frac{2}{a}e^{ta/2}\Big|_0^a = \frac{2}{a}\left(1-e^{-a^2/2}\right).
\end{aligned}
$$

Consequently, there holds for $a \to \infty$: $I_0 = I_1 + o(1)$, and therefore

$$
\widehat{f}(\xi) = \frac{e^{-\xi^2/2}}{\sqrt{2\pi}},
$$

which concludes the proof.

□

Remark 12. *Of course most of the lemmas, theorems and applications from Sec. 2 can be one-to-one transferred to the setting of the continuous Fourier transform. For example: Differentiation in function space becomes multiplication in Fourier space is still true, and can be similarly motivated as before. We omit the details.*

Obviously, the Fourier transform is well-defined in L^1, but not necessarily in L^2. To also obtain an L^2-theory, we begin with the equivalent to La. 7:

Lemma 19 (Plancherel theorem (also Parseval identity))**.** *Let $f \in C_c(\mathbb{R})$, i.e., f is a continuous function with compact support. Then, $\widehat{f} \in L^2(\mathbb{R})$ and there holds*

$$
\|f\|_{L^2} = \|\widehat{f}\|_{L^2}. \tag{14}
$$

Proof. With $f \in C_c(\mathbb{R})$, there holds $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, which means that both the Fourier transform $\widehat{f}$ and the L^2 norm $\|f\|_{L^2}$ exist.

Let us first suppose that $\operatorname{supp} f \subset [-\pi, \pi]$. Our aim is to apply the Fourier theory on a periodic domain; we can hence find coefficients c_k, $k \in \mathbb{Z}$, with

$$
c_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\mathrm{d}x = \frac{1}{2\pi}\int_{\mathbb{R}} f(x)e^{-ikx}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\widehat{f}(k)
$$

such that

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

Note that $\widehat{f}$ refers to the continuous Fourier transform here! Now consider for an $a\in\mathbb{R}$ the function $g(x) := f(x)e^{-iax}$. Note that $\|f\|_{L^2} = \|g\|_{L^2}$. Of course $\operatorname{supp} g \subset [-\pi,\pi]$, so $g\in C_c(\mathbb{R})$ still holds. Furthermore, applying the same trick, one ends up with

$$\|f\|_{L^2}^2 = \|g\|_{L^2}^2 = \sum_{k\in\mathbb{Z}}|\widehat{f}(k+a)|^2.$$

Hence, we can state that

$$\begin{aligned}\|f\|_{L^2}^2 = \int_0^1 \|f\|_{L^2}^2 \mathrm{d}a &= \int_0^1 \sum_{k\in\mathbb{Z}}|\widehat{f}(k+a)|^2 \mathrm{d}a\\ &= \sum_{k\in\mathbb{Z}}\int_0^1 |\widehat{f}(k+a)|^2\mathrm{d}a = \int_{\mathbb{R}}|\widehat{f}(a)|^2\mathrm{d}a = \|\widehat{f}\|_{L^2}^2.\end{aligned}$$

Interchange of summation and integration is due to Beppo-Levi's theorem on the monotone convergence.

If now $\operatorname{supp} f \subset [-a,a]$ for $a>\pi$, then we proceed with a scaling, so define $g(x) := f(ax)$, $\operatorname{supp} g\subset[-\pi,\pi]$. There holds (see La. 17) $\widehat{g}(\xi) = \frac{1}{a}\widehat{f}(\frac{\xi}{a})$. Hence,

$$\begin{aligned}\|f\|_{L^2}^2 &= \int_{\mathbb{R}}|f(x)|^2\mathrm{d}x = a\int_{\mathbb{R}}|f(az)|^2\mathrm{d}z = a\int_{\mathbb{R}}|g(z)|^2\mathrm{d}z\\ &= a\int_{\mathbb{R}}|\widehat{g}(\xi)|^2\mathrm{d}\xi = \frac{1}{a}\int_{\mathbb{R}}\left|\widehat{f}\left(\frac{\xi}{a}\right)\right|^2\mathrm{d}\xi = \int_{\mathbb{R}}|\widehat{f}(\xi)|^2\mathrm{d}\xi = \|\widehat{f}\|_{L^2}^2.\end{aligned}$$

□

Based on this remarkable fact, we can proceed to the following lemma:

Lemma 20. *The Fourier transform can be defined on the whole space $L^2(\mathbb{R})$.*

Proof. It is well-known that $C_c(\mathbb{R})\subset L^1(\mathbb{R})\cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$. Thus, for any $f\in L^2(\mathbb{R})$, there exists a sequence $(f_n)\subset C_c(\mathbb{R})$, such that $f_n\to f$ in L^2. For any such f_n, one can define the Fourier transform $\mathcal{F} f_n \equiv \widehat{f_n}$ as in (10). Because of Plancherel's theorem, which states that $\|\mathcal{F} f_n\|_{L^2} = \|f_n\|_{L^2}$, $(\mathcal{F} f_n)_{n\in\mathbb{N}}$ will also be a Cauchy sequence, and therefore have a limit which we call $\widehat{f}$. This limit is independent of the approximating sequence (f_n) (why?), and we define it to be $\mathcal{F} f$. □

Corollary 5 (Plancherel theorem (also Parseval identity))**.** *Let $f\in L^2(\mathbb{R})$. Then, $\widehat{f}\in L^2(\mathbb{R})$ and there holds*

$$\|f\|_{L^2} = \|\widehat{f}\|_{L^2}. \tag{15}$$

Having now defined the Fourier transformation, it is also important to know how to transform it back, i.e., given the Fourier coefficients, how can one reproduce the original function?

Theorem 15. *Let $f\in L^1(\mathbb{R})$ and $\widehat{f}\in L^1(\mathbb{R})$. Then,*

$$f(x) = \frac{1}{\sqrt{2\pi}}\int \widehat{f}(\xi)e^{i\xi x}\mathrm{d}\xi \tag{16}$$

holds almost everywhere, in particular, in the points of continuity of f.

Remark 13. *We chose the factor $\frac{1}{\sqrt{2\pi}}$ because it makes forward and backward formula coefficients equal.*

Proof of Theorem 15. The proof actually works with a couple of tricks, and the technical results we summarized in La. 17. Because $\widehat{f} \in L^1(\mathbb{R})$, we can conclude with the theorem of dominated convergence that the right-hand side of (16) is given by

$$\lim_{\varepsilon \to 0^+} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\varepsilon^2\xi^2/2+ix\xi} \widehat{f}(\xi) \mathrm{d}\xi.$$

We can further manipulate this integral by applying La. 17, in particular, we apply the change of Fourier transform: (Note that the Fourier transform refers to ξ this time.)

$$= \lim_{\varepsilon \to 0^+} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{e^{-\varepsilon^2\xi^2/2+ix\xi}} f(\tau) \mathrm{d}\tau.$$

Let us define

$$g_\varepsilon(\xi) := e^{-\varepsilon^2\xi^2/2+ix\xi} = e^{ix\xi} e^{-(\xi\varepsilon)^2/2}, \quad \text{and} \quad g(\xi) = e^{-\xi^2/2}.$$

Obviously, $g_\varepsilon(\xi) = e^{ix\xi} g(\xi\varepsilon)$. Again due to La. 17, the Fourier transform of g_ε is given by $\widehat{g_\varepsilon}(\tau) = \frac{1}{\varepsilon}\widehat{g}\left(\frac{\tau-x}{\varepsilon}\right)$. Therefore, we can continue computing the above integral as

$$= \lim_{\varepsilon \to 0^+} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{\varepsilon} \widehat{g}\left(\frac{\tau - x}{\varepsilon}\right) f(\tau) \mathrm{d}\tau.$$

Because $\widehat{g} = g$ (see La. 18) and because the integral of g along the real line is $\sqrt{2\pi}$, we know that the expression will converge almost everywhere toward $f(x)$. This ends the proof. □

Remark 14. *For the sake of completeness, let us define the Schwartz space $\mathcal{S}$ as the space of functions f that are arbitrarily often differentiable, and any derivative of f goes, for $|x| \to \infty$, faster to zero as any x^{-n}, $n \in \mathbb{N}$. Examples are $f(x) = e^{-cx^2}$ $(c > 0)$ or $f(x) = \frac{1}{\cosh(x)}$. If f is arbitrarily often differentiable, then also $\widehat{f}$ has a very fast decay. (Compare also for the periodic case.) Overall, this means that the Fourier transform is a bijection from $\mathcal{S}$ onto $\mathcal{S}$.*

Remark 15. *So far, we did only treat functions $f : \mathbb{R} \to \mathbb{C}$. It is also possible to extend the methodology of the previous chapters - with obvious modifications - to functions $f : \mathbb{R}^d \to \mathbb{C}$. In this case, one defines the Fourier transform as*

$$\widehat{f}(\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} f(x) e^{-i\xi\cdot x} \mathrm{d}x,$$

and $\widehat{f} : \mathbb{R}^d \to \mathbb{C}$. However, for the ease of presentation, we will stay in the one-dimensional setting.

4.2 Shannon's sampling theorem

From a practical point of view, it might be interesting to know *how many* information on f is needed to exactly reproduce it. As an example, suppose that one is given a set of discrete values $(f(kX)|k \in \mathbb{Z})$ for some value of X. Is it possible to reproduce f exactly? Without further assumption on f (and X), the answer is certainly no, because inbetween those values, the function might behave arbitrarily. However, there is hope for the special class of functions with limited bandwidth:

Definition 13. *A function $f \in L^1(\mathbb{R})$ is called Ω-bandlimited, if*

$$\widehat{f}(\xi) = 0 \quad \forall |\xi| > \Omega.$$

If a function f is Ω-bandlimited, then we can exactly reproduce it if we know the discrete values at instances kX for $X := \frac{\pi}{\Omega}$. This is the famous Shannon theorem (*Bemonsteringstheorema* van Nyquist-Shannon), which also yields a formula for f:

Theorem 16. *Let $f : \mathbb{R} \to \mathbb{C}$ be continuous and Ω-bandlimited. Furthermore, we assume that f vanishes sufficiently fast for $|x| \to \infty$, i.e.,*

$$f(x) = \mathcal{O}\left(\frac{1}{|x|^{1+\varepsilon}}\right)$$

for some positive value of ε. We define $X := \frac{\pi}{\Omega}$. Then, there holds:

$$f(x) = \sum_{k\in\mathbb{Z}} f(kX)\,\mathrm{sinc}(\Omega(x - kX)). \tag{17}$$

Before proving the theorem, we start with a remark:

Remark 16. *Because f is $\Omega-$bandlimited, the periodicity of the exponentials $e^{i\xi x}$ that play a role in the Fourier representation of f is bound from below by $\frac{2\pi}{\Omega}$. The theorem now states that more than two known values of f must be present per period.*

Proof. Due to the assumption on the decay behavior of f, we know that it is in $L^1 \cap L^2$. This implies that it has a continuous Fourier transform, which is also in L^1, as it has compact support. All requirements of Thm. 15 are therefore met, and we can write f as

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(\xi) e^{i\xi x} \mathrm{d}\xi = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \widehat{f}(\xi) e^{i\xi x} \mathrm{d}\xi. \tag{18}$$

Furthermore, as $\widehat{f}$ is continuous, we know that $\widehat{f}(-\Omega) = \widehat{f}(\Omega) = 0$. This means that we can consider $\widehat{f}$ to be a 2Ω periodic function (at least on $(-\Omega, \Omega)$), and we can therefore extend it into a Fourier series:

$$\widehat{f}(\xi) = \sum_{k\in\mathbb{Z}} c_k e^{ik\xi\frac{2\pi}{2\Omega}}, \quad \xi \in (-\Omega, \Omega) \tag{19}$$

and the c_k (in our earlier notation, it should be $c_k = \widehat{\widehat{f}}(k)$, which is a bit weird to use...) are computed to be

$$c_k = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \widehat{f}(\xi) e^{-ik\xi\frac{2\pi}{2\Omega}} \mathrm{d}\xi. \tag{20}$$

This is now where (18) comes into play: The c_k are exactly of this form, thus one can write

$$c_k = \frac{\sqrt{2\pi}}{2\Omega} f\left(-k\frac{2\pi}{2\Omega}\right) = \frac{\sqrt{2\pi}}{2\Omega} f\left(-kX\right). \tag{21}$$

Collecting previous fact, we may write $f(x)$ as

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \left(\sum_{k\in\mathbb{Z}} \frac{\sqrt{2\pi}}{2\Omega} f(-kX) e^{ik\xi\frac{2\pi}{2\Omega}} \right) e^{i\xi x} \mathrm{d}\xi = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \left(\sum_{k\in\mathbb{Z}} f(kX) e^{-ik\xi X} \right) e^{i\xi x} \mathrm{d}\xi. \tag{22}$$

Note that we have formally changed the summation index k to $-k$. The sum in (22) converges absolutely because of the decay conditions on f. Therefore, we may change summation and integration, and obtain

$$f(x) = \frac{1}{2\Omega} \sum_{k\in\mathbb{Z}} f(kX) \int_{-\Omega}^{\Omega} \left(e^{i\xi(-kX + x)} \right) \mathrm{d}\xi. \tag{23}$$

The integral in the sum can be computed as follows:

$$\int_{-\Omega}^{\Omega} e^{i\xi(x-kX)}\mathrm{d}\xi = \int_{-\Omega}^{\Omega} \cos(\xi(x-kX))\mathrm{d}\xi = \frac{2}{x-kX}\sin(\Omega(x-kX)), \quad x \neq kX. \tag{24}$$

It can be easily verified that this integral is - for all $x \in \mathbb{R}$ - equal to

$$\int_{-\Omega}^{\Omega} e^{i\xi(x-kX)}\mathrm{d}\xi = 2\Omega\,\mathrm{sinc}(\Omega(x-kX)). \tag{25}$$

Due to (23), we obtain

$$f(x) = \sum_{k\in\mathbb{Z}} f(kX)\,\mathrm{sinc}(\Omega(x-kX)), \tag{26}$$

which is the desired result. □

Remark 17. • *The quantity* $\frac{\pi}{X}$ *is called* Nyquist frequency *for a given sampling interval of size* X*.* $X^{-1} := \frac{\Omega}{\pi}$ *is called* Nyquist rate *for functions of bandwidth* Ω*.*

- *The right-hand side of* (17) *is called* cardinal series*.*

From the discrete Fourier transform, we already have some intuition on what happens if one uses the right-hand side of (17) for a function f that is *not* $\Omega-$bandlimited, but, e.g., $\Omega'-$bandlimited, with $\Omega' > \Omega$: One observes the phenomenon of *aliasing* again. If, for example, $\Omega < \Omega' < 3\Omega$, then the function $g \in L^2$ with Fourier transform

$$\widehat{g}(\xi) = \begin{cases} \widehat{f}(\xi) + \widehat{f}(\xi - 2\Omega) + \widehat{f}(\xi + 2\Omega), & -\Omega \leq \xi \leq \Omega \\ 0, & |\xi| > \Omega \end{cases}$$

has the same cardinal series as f. (Try this as an exercise!)

This basically means that undersampling is not a good idea. What, however, about oversampling? This means that a function that is Ω'-bandlimited, is treated as if it were $\Omega-$bandlimited, with some $\Omega > \Omega'$. In practice, this can yield a faster convergence of the infinite sum in (17). To demonstrate this, we define the help-function q via its Fourier transform,

$$\widehat{q}(\xi) := \begin{cases} 1, & |\xi| \leq \Omega' \\ \frac{1}{2}\left(1 - \sin\left(\frac{\pi(2|\xi| - \Omega - \Omega')}{2(\Omega - \Omega')}\right)\right), & \Omega' < |\xi| < \Omega\ . \\ 0, & |\xi| > \Omega \end{cases}$$

Obviously, $\widehat{q}(\xi)$ is continuous and integrable if only $\Omega' < \Omega$. Based on the proof of Thm. 16, see (22), we can compute

$$\begin{aligned} \frac{1}{\sqrt{2\pi}}\int_{-\Omega}^{\Omega} \widehat{f}(\xi)e^{i\xi x}\mathrm{d}\xi &= \frac{1}{\sqrt{2\pi}}\int_{-\Omega}^{\Omega} \widehat{f}(\xi)\widehat{q}(\xi)e^{i\xi x}\mathrm{d}\xi = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega}\left(\sum_{k\in\mathbb{Z}} f(kX)e^{-ikX\xi}\right)\widehat{q}(\xi)e^{ix\xi}\mathrm{d}\xi \\ &= \frac{1}{2\Omega}\sum_{k\in\mathbb{Z}} f(kX)\int_{-\Omega}^{\Omega}\widehat{q}(\xi)e^{i\xi(x-kX)}\mathrm{d}\xi. \end{aligned}$$

So upon defining $Q(s) := \frac{1}{2\Omega}\int_{-\Omega}^{\Omega}\widehat{q}(\xi)e^{i\xi s}\mathrm{d}\xi$, f is described by both (17) and

$$f(x) = \sum_{k\in\mathbb{Z}} f(kX)Q(x-kX).$$

The more preferable formulation is the one that has a faster decay of the weighting functions as $|k|$ tends to infinity. Obviously, the sinc function used in (17) has a decay rate of $\mathcal{O}\left(\frac{1}{k\pi}\right)$ (note that

$\Omega X = \pi$). Explicitly giving a formula for $Q(s)$ is tedious, but can be done in a straightforward way, it yields

$$Q(s) = \frac{\pi^2}{2\Omega s} \frac{\sin(\Omega' s) + \sin(\Omega s)}{\pi^2 - (\Omega - \Omega')^2 s^2}.$$

This yields that the convergence is of order $|s|^{-3}$. For $\Omega = 2\Omega'$, the expression $Q(x - kX)$ is of order $\mathcal{O}\left(\frac{1}{|k|^3}\right)$. This guarantees a much faster convergence if one evaluates the infinite sum only at finitely many instances.

4.3 The windowed Fourier transform

One of the drawbacks of the Fourier transform is that localization (in x) is lost when going to the frequency space. Certainly, this makes some sense because a 'frequency' is something that one can only measure if one has at least an environment of some point x at hand (cf. Heisenberg's uncertainty principle). However, one can introduce some sort of localization using the windowed Fourier transform. Choosing a (more or less arbitrary) function g with $\int_{\mathbb{R}} g(x)\mathrm{d}x = 1$ that is 'localized around zero' (this means that there should be a maximum at zero, maybe together with compact support), one can define the windowed Fourier transform Gf as a function of two variables as

$$Gf(\xi, s) := \frac{1}{\sqrt{2\pi}} \int_R f(x)g(x - s)e^{-i\xi x}\mathrm{d}x. \tag{27}$$

(Note that for $g \equiv 1$, this is precisely the Fourier transform.) Of course this is highly redundant, and therefore, there exist many different inversion formulae of this. We omit this for brevity. A very common choice of g is the following:

Definition 14. *Let $g(x)$, for a fixed value of $\sigma > 0$, be given by*

$$g(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}.$$

The associated transformation (27) is called Gabor*-transformation.*

The parameter σ controls how 'localized around zero' the function g is: g tends to zero monotonically and exponentially for $|x| \to \infty$, and $g(\pm 2\sigma) < \frac{g(0)}{5}$, so it is safe to say that the function has a width of about 2σ. (It has no compact support, however.)

The drawback of this transformation (and we use this as a motivation for wavelets in the sequel) can be seen for extremely large and extremely small frequencies. To motivate this, we assume that g has support in some interval $(-h, h)$ (the reasoning, however, also holds for the Gabor transformation).

- For extremely large (in magnitude) frequencies, there are many oscillations in the window $(s - h, s + h)$. Thus, it is not possible to determine where exactly the oscillation comes from - which is unfortunately not what we wanted.
- For extremely small (in magnitude) frequencies, the window $(s - h, s + h)$ doesn't even cover a complete oscillation, thus it has no way of determining such frequencies.

Remark 18. *The problem seems to be that the h (or the σ) is fixed. One should include a more clever logarithmic scale that is able to both resolve small and large frequencies... Wavelets do that!*

4.4 Other, similar transforms

There exist other sorts of transforms, tailor-suited to different 'needs' of the mathematical community. We have seen the Fourier transform as one way to solve PDEs on either a periodic grid or on the whole space. Often, also time t plays a role, which is inherently a positive quantity, i.e., the problem is posed on the half space. To treat problems like that, one can for example use the *Laplace transform*

$$\mathcal{L}_f(s) := \int_0^\infty e^{-st} f(t) \mathrm{d}t.$$

Unlike for the Fourier transform, there holds $s \in \mathbb{C}$ here. The treatment of this is, however, beyond the scope of this lecture.

References

[1] Hans Wilhelm Alt. *Lineare Funktionalanalysis*. Springer Berlin / Heidelberg, 5^{th} edition, 2006.

[2] Christian Blatter. *Wavelets: A Primer*. A K Peters, Natick, 1998.