# Analysis: Convergence and Duality

Philippe Jaming

Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France. CNRS, IMB, UMR 5251, F-33400 Talence, France.

*Email address*: Philippe.Jaming@math.u-bordeaux.fr

CHAPTER 1

# Basics of Banach spaces

The aim of this first Lecture is to recall the basic definitions on Banach spaces and to illustrate them with some examples.

## 1. Definitions

DEFINITION 1.1. Let $E$ be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

- A *norm* on $E$ is a mapping $\|.\|$ from $E$ to the non-negative reals $\mathbb{R}^+$ such that for every $\lambda \in \mathbb{K}$ and every $x, y \in E$
  (1) $\|0\| = 0$, $\|x\| \geq 0$ and $\|x\| = 0$ implies that $x = 0$;
  (2) $\|\lambda x\| = |\lambda|\|x\|$;
  (3) $\|x + y\| \leq \|x\| + \|y\|$.
  We then say that $(E, \|.\|)$ is a *normed vector space.*
- A sequence $(e_n)$ in $E$ *converges to* $e \in E$ if $\|e_n - e\| \to 0$, that is if:

  *for every $\varepsilon > 0$ there exists $N$ (depending on $\varepsilon$) such that, for every $n \geq N$, $\|e_n - e\| \leq \varepsilon$.*

  In this case, we write $e = \lim e_n$.
- A sequence $(e_n)$ in $E$ is a *Cauchy sequence* if

  *for every $\varepsilon > 0$ there exists $N$ (depending on $\varepsilon$) such that, for every $m, n \geq N$, $\|e_n - e_m\| \leq \varepsilon$.*

EXAMPLE 1.2. Let $E$ be a finite dimensional vector space, $(e_k)_{k=1,\ldots,d}$ a basis of $E$. Each $e \in E$ can then be uniquely written as $e = \displaystyle\sum_{k=1}^{d} x_k e_k$. We can thus define $\|x\|_\infty = \sum_{k=1}^{d} |x_k|$.

It is easy to check that this defines a norm on $E$.

Moreover, if $(f_n)$ is a sequence in $E$, we can write each $f_n$ as $f_n = \displaystyle\sum_{k=1}^{d} x_n^{(k)} e_k$. It is easy to check that $f_n$ converges (resp. if a Cauchy sequence) if and only if each $(x_n^{(k)})$ converges (resp. is a Cauchy sequence).

But $(x_n^{(k)})$ is a sequence in $\mathbb{R}$ and in $\mathbb{R}$ (either by definition or after some cumbersome work depending on how $\mathbb{R}$ is constructed) Cauchy sequences converge.

Note that this norm depends on the chosen basis.

REMARK 1.3. It is an easy exercice to show that every convergent sequence is a Cauchy sequence. The converse may *not* be true. This shows that the chosen norm is not well adapted to the vector space $E$.

DEFINITION 1.4. Let $E$ be a vector space and $\|.\|$, $|||.|||$ be two norms on $E$. We say that they are equivalent if there exists a constant $C$ such that, for every $e \in E$,

$$\frac{1}{C}\|e\| \le |||e||| \le C\|e\|.$$

This of course defines an equivalence relation on norms which simply means that if $\|.\|_1$ is equivalent to $\|.\|_2$ which in turn is equivalent to $\|.\|_3$, then $\|.\|_1$ is equivalent to $\|.\|_3$.

The following lemma is a simple exercice left to the reader:

LEMMA 1.5. *Let $E$ be a vector space and $\|.\|$, $|||.|||$ be two equivalent norms on $E$. Then every Cauchy sequence (resp. convergent sequence) for one norm is also a Cauchy sequence (resp. convergent sequence) for the second one.*

A more subtil fact is the following:

THEOREM 1.6. *Let $E$ be a finite dimensional vector space. Then any two norms on $E$ are equivalent.*

SKETCH OF PROOF. It is enough to fix a basis $(e_k)$ of $E$ and to show that any norm on $E$ is equivalent to the $\|.\|_\infty$ norm defined in the above example.

Take $e \in E$, write $e = \sum x_k e_k$ and notice that

$$\|e\| = \left\|\sum x_k e_k\right\| \le \sum |x_k|\|e_k\| \le \left(\sum \|e_k\|\right)\max |x_k| = \left(\sum \|e_k\|\right)\|e\|_\infty.$$

The converse is more subtle and uses the following facts:

- from the above inequality, we get that $e \to \|e\|$ is a continuous mapping from $E, \|.\|_\infty$ to $\mathbb{R}$
- the unit sphere of $E, \|.|_\infty$ $S_E = \{e \in E : \|e\|_\infty = 1\}$ is *compact* (we will come back to this notion in the second semester and prove this fact)
- a continuous function over a compact set is bounded.

The remaining of the proof is then simple: there exists $C > 0$ such that, for every $e \in S_E$, $\|e\| \le C$. And, if $e' \in E$

– either $e' = 0$ and $\|e'\| = 0 = \|e'\|_\infty$

– or $e' \neq 0$ and then $e := \frac{1}{\|e'\|_\infty}e' \in S_E$ thus $\frac{\|e'\|}{\|e'\|_\infty} = \|e\| \le C$ therefore $\|e'\| \le C\|e\|$. □

DEFINITION 1.7. Let $E, \|.\|$ be a normed vector space. Then $E$ is a *complete* if every Cauchy sequence in $E$ is convergent. We then aso say that $E$ is a *Banach space.*

REMARK 1.8. This is a key property in analysis. It allows to define an object as a limit of a sequence.

Indeed, contrary to the definition of a convergent sequence, the definition of a Cauchy sequence does not require to know (or guess) the limit.

Many examples will occur in this course.

THEOREM 1.9. *Every finite dimensional normed vector space is a Banach space.*

PROOF. We have already seen that $E, \|.\|_\infty$ (with some fixed basis) is complete. Since all norms are equivalent, Cauchy sequences for any other norm $\|.\|$ are also Cauchy sequences for the $\|.\|_\infty$ norm thus are convergent for the $\|.\|_\infty$ and, by equivalence of norms again, are also convergent for $\|.\|$. □

## 2. An example: the space of continuous functions

THEOREM 2.1. *Let* $E = \mathcal{C}([0,1]) = \{f\ [0,1] \to \mathbb{C} : f$ *is continuous on* $[0,1]\}$. *Endow* $E$ *with the norm defined by*

$$\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|.$$

*Then* $E$ *is a Banach space.*

REMARK 2.2. It is a basic calculus exercice to show that $\|f\|_\infty$ is well-defined (continous functions over the compact set $[0,1]$ are bounded) and that it is a norm.

Before we prove this result, let us recall the following fact from calculus:

LEMMA 2.3. *Let* $(f_n)$ *be a sequence* $\mathcal{C}([0,1])$ *and let* $f$ *be a function* $[0,1] \to \mathbb{C}$. *If* $\sup_{x\in[0,1]} |f_n(x) - f(x)| \to 0$ *then* $f$ *is continous over* $[0,1]$ *so that* $f_n \to f$ *in* $E$

PROOF. Fix $x_0 \in [0,1]$ and let $\varepsilon > 0$. Then there exists $n$ such that, for every $x \in [0,1]$, $|f_n(x) - f(x)| \le \varepsilon$. As $f_n$ is continuous in $x_0$, there exists a neighborhood $V$ of $x_0$ such that, for every $x \in V$ $|f_n(x) - f_n(x_0)| \le \varepsilon$. But then

$$|f(x) - f(x_0)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \le 3\varepsilon$$

which shows continuity of $f$ in $x_0$. □

We can now prove the theorem.

PROOF. Let $(f_n)$ be a Cauchy sequence: for every $\varepsilon > 0$ there is an $N$ such that, if $m, n \ge N$ and $x \in [0,1]$ $|f_n(x) - f_m(x)| < \varepsilon$.

First fix $x \in [0,1]$, the above shows that $\big(f_n(x)\big)_n$ is a Cauchy sequence in $\mathbb{C}$, which is complete. Therefore $\big(f_n(x)\big)_n$ has a limit that we denote by $f(x)$. This defines a function $f\ [0,1] \to \mathbb{C}$. It remains to show that this $f$ is also a limit in the sense of $\|f\|_\infty$. Once this is done, a result in the calculus course shows that $f$ is then continuous.

But this is simple: fix $\varepsilon > 0$, take $N$ such that if $n \ge N$ and $x \in [0,1]$, for every $m \ge N$, $|f_n(x) - f_m(x)| < \varepsilon$. Fix $x$ and $n \ge N$ and let $m$ go to infinity then $|f_n(x) - f(x)| \le \varepsilon$, thus, for every $n \ge N$, $\|f_n - f\|_\infty := \sup_{x\in[0,1]} |f(x)| \le \varepsilon$ which precisely means that $f_n \to f$ in $E$. □

The following proof has highlited the fact that there are two convergens in $\mathcal{C}([0,1])$:

DEFINITION 2.4. Let $(f_n)$ be a sequence $\mathcal{C}([0,1])$ and let $f$ be a function $[0,1] \to \mathbb{C}$. We say that

- $f_n \to f$ pointwise (or simply) if, for every $x \in [0,1]$, $f_n(x) \to f(x)$ in $\mathbb{C}$
- $f_n \to f$ uniformly if $f_n \to f$ for the $\|\cdot\|_\infty$ norm.

The first step in the above proof amounts to showing that uniform convergence implies pointwise convergence. The converse is *false* as the example $f_n(x) = x^n$ shows. Indeed $f_n$ converges pointwise to the function $f$ defined by $f(x) = \begin{cases} 0 & \text{if } x \ne \\ 1 & \text{if } x = 1 \end{cases}$. As $f$ is not continuous while $f_n$ is, the convergence can not be uniform.

Note that one can define other norms on $\mathcal{C}([0,1])$, for instance, one can show that the following is also a norm:

$$\|\varphi\|_1 = \int_0^1 |\varphi(x)|\,\mathrm{d}x.$$

As $f_n \to f$ for this norm we can deduce that:

– the $\|\cdot\|_1$ is *not equivalent* to the $\|\cdot\|_\infty$ norm
– as $f$ is not continuous, $\mathcal{C}([0,1]), \|\cdot\|_1$ is not complete.
– uniform convergence is a stronger convergence than the convergence in $\|\cdot\|_1$-norm. The convergence in this last norm will lead us to introduce a larger space.

REMARK 2.5. $[0,1]$ plays no particular role here: we only use its compacity to show that $\|f\|_\infty$ is well defined. In particular
– everything in this section is valid if $\mathcal{C}([0,1])$ is replaced by $\mathcal{C}(K)$ where $K$ is a compact set
– everything in this section is valid if $\mathcal{C}([0,1])$ is replaced by $\mathcal{C}_b(X)$ the set of *bounded* continuous functions on a metric space $X$.

## 3. A second example: bounded linear mappings

In this section, $E, F$ will be two Banach spaces.
Let us recall the following fact about linear mappings:

THEOREM 3.1. *Let $T : E \to F$ be linear. Then the following are equivalent:*
1. *$T$ is continuous on $E$;*
2. *$T$ is continuous in $0$;*
3. *$T$ is bounded over the unit ball of $E$: there exists $C$ such that, for every $x \in E$ with $\|x\| \leq 1$, $\|Tx\| \leq C$.*
4. *there exists $K$ such that, for every $x \in E$, $\|Tx\| \leq K\|x\|$.*

SKETCH OF PROOF. The equivalence of (1) and (2) follows from the fact that $T(x) - T(x_0) = T(x - x_0) = T(x - x_0) - T(0)$.
(4) clearly implies (3) with $C = K$. For the converse, if $x \neq 0$, then $\frac{x}{\|x\|}$ has norm 1 thus $\frac{\|Tx\|}{\|x\|} = \left\|T\frac{x}{\|x\|}\right\| \leq C$ thus (4) holds with $K = C$.
Finally (4) clearly imples (2). Finally, if (2) holds, take $\varepsilon = 1$, there exists $\eta > 0$ such that, if $\|x\| \leq \eta$, $\|Tx\| \leq 1$. It follows that, when $\|x\| \leq 1$, then $\|\eta x\| \leq \eta$ thus

$$\|Tx\| = \frac{1}{\eta}\|T(\eta x)\| \leq \frac{1}{\eta}$$

which completes the proof. □

LEMMA 3.2. *If $E$ has finite dimension, then every linear map $E \to F$ is continuous.*

SKETCH OF PROOF. This is simple to show for the $\|\cdot\|_\infty$ norm on $E$ (with a fixed basis) since, writting $x = \sum x_i e_i$ we get $Tx = \sum x_i Te_i$ thus

$$\|Tx\| \leq \sum |x_i|\|Te_i\| \leq \left(\sum \|Te_i\|\right) \|x\|_\infty.$$

Then usuing the fact that all norms are equivalent in $E$, we conclude tha $T$ is also contiuous for any other norm. □

We can now define the following for a *continous (bounded)* linear mapping $T : E \to F$:

$$\|T\| = \sup_{x\in E, \|x\|_E \leq 1} \|Tx\|_F = \sup_{x\in E, \|x\|_E = 1} \|Tx\|_F \sup_{x\in E, x\neq 0} \frac{\|Tx\|_F}{\|x\|_E}.$$

It is an exercice to show that this three quantities are indeed equal and that they define a norm on the set $B(E,F) := \{T : E \to F \mathrm{linear and continuous}\}$. Such a norm on $B(E,F)$ is called a *subordinate norm*.

THEOREM 3.3. *The space $B(E,F)$ endowed with this norm is a Banach space.*

PROOF. We will just outline the proof, the details are left to the reader as they are essentally the same as for $\mathcal{C}([0,1])$.

Take $T_n$ to be a Cauchy sequence in $B(E,F)$.

Step 1: observe that for each $x \in E$, $T_n(x)$ is a Cauchy sequence in $F$. We can then define $T(x) = \lim T_n(x)$.

Step 2: $T$ is easily seen to be linear.

Step 3: $T_n$ is Cauchy thus a bonded sequence, that is, there exists $C$ such that $\|T_n\| \leq C$. In other words, for every $x \in E$ with $\|x\| \leq 1$, $\|T_n(x)\| \leq C$.

As the norm is a continuous function, $\|T_n(x)\| \to \|T(x)\|$ thus for every $x \in E$ with $\|x\| \leq 1$, $\|T(x)\| \leq C$ and $T$ is bounded, thus continuous.

Step 4: $\|T_n - T\| \to 0$ so that $T_n \to T$ in $B(E,F)$. □

A particular case is the dual space of $E$:

DEFINITION 3.4. Let $E$ be a Banach space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. The dual of $E$ is the Banach space $E' = B(E, \mathbb{K})$.

If $(E')' = E$ then $E$ is said to be reflexive.

The key example is the case when $E$ is a Hilbert space. Further examples will be given in this course:

THEOREM 3.5 (Riesz). *If $H$ is a Hilbert space over $\mathbb{K}$ then $H'$ can be identified with $H$ in the sense that, for every element $\ell \in H'$ there exists a unique $a \in H$ such that $\ell(x) = \langle x, a \rangle$ where $\langle ., . \rangle$ is the scalar product on $H$.*

Note that Cauchy-Schwarz implies that $x \to \langle x, a \rangle$ is indeed bounded linear mapping on $H$ (note that when $H$ is a complex Hilbert space, the scalar product is anti-linear in the second variable).

We can now define a new convergence:

DEFINITION 3.6. Let $E$ be a Banach space and $E'$ its dual. Let $(x_n)$ be a sequence in $E$ and $x \in E$. We say that $x_n \to x$ *weakly* and write $w - \lim x_n = x$ if, for every $\ell \in E'$, $\ell(x_n) \to \ell(x)$.

The term "weak" convergence is justified by the fact that convergence implies weak convergence (and is thus also called strong convergence). Indeed, if $(x_n)$ converges to $x$, as $\ell$ is *continuous*, $\ell(x_n) \to \ell(x)$.

It turns out that, if $E$ is finite dimensional, then the converse is true as well. Indeed if $(e_k)_{k=1,\ldots,d}$ is a basis of $E$ then every $x \in E$ can be uniquely written as $x = \sum x_k e_k$. We can then define the mapping $l_k : E \to \mathbb{K}$ by $l_k(x) = x_k$. The uniqueness of the $x_k$ also implies that each $l_k$ is linear. As $E$ is finite dimensional, we also get that $l_k$ is continuous, that is $l_k \in E'$. But now, if $(x^{(n)})$ converges weakly to $x$, for each $k = 1, \ldots, d$ $l_k(x^{(n)}) \to_{n\to\infty} l_k(x)$ and then

$$x^{(n)} = \sum_{k=1}^{d} l_k(x^{(n)}) e_k \to_{n\to\infty} \sum_{k=1}^{d} l_k(x) e_k = x.$$

In infinite dimensions, the situation is different, even in separable Hilbert spaces. For instance, such a space has an orthonormal basis $(e_n)$. Such a sequence can *not* converge (and even has no sub-sequence that is convergeant) since, for every $m \neq n$

$$\|e_n - e_m\|^2 = \|e_n\|^2 + \|e_m\|^2 + 2\Re\langle e_n, e_m \rangle = 2.$$

On the other hand, if $\ell \in H'$ then, as seen above, there exists $a \in H$ such that $\ell(x) = \langle x, a \rangle$. But
$$\sum_{n \geq 0} |\ell(e_n)|^2 = \sum_{n \geq 0} |\langle e_n, a \rangle|^2 = \|a\|^2.$$
In particular, this series is convergent, thus its general term goes to 0: $|\ell(e_n)|^2 \to 0$ thus $\ell(e_n) \to 0$ which is precisely $w - \lim e_n = 0$.

The notion of weak convergence is essential in analysis in infinite dimensions. The main reason (seen in the next semester) is that closed bounded sets will no longer be compact in that case, but will still be weakly compact.

The notion of weak-convergence should not be confused with Cesaro-convergence which is an other usefull convergence that is weaker than classical convergence:

Definition 3.7. Let $E$ be a Banach space, $x \in E$ and $(x_n)_{n \geq 0}$ a sequence in $E$. We say that $(x_n)$ Cesaro-converges to $x$ if the mean of its $n$ first terms
$$\frac{x_1 + x_2 + \cdots + x_n}{n} \to x.$$

It is not hard to see that a sequence can be Cesaro convergent without being convergent. For instance, if we consider the real sequence given by $x_n = (-1)^n$, then
$$\frac{x_1 + x_2 + \cdots + x_n}{n} = \begin{cases} -1/n & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases} \to 0.$$

However, the converse is true

Lemma 3.8. *Assume that $(x_n)$ converges to $x$ then $(x_n)$ Cesaro-converges to $x$.*

Proof. First note that
$$\frac{x_1 + x_2 + \cdots + x_n}{n} - x = \frac{(x_1 - x) + (x_2 - x) + \cdots + (x_n - x)}{n}.$$
By replacing $x_n$ by $x_n - x$ we may thus assume without loss of generality that $x = 0$.

Let $\varepsilon > 0$, and take $N$ such that, for all $n \geq N$, $\|x_n\| \leq \varepsilon$. Then
$$\begin{aligned} \left\| \frac{x_1 + x_2 + \cdots + x_n}{n} \right\| &\leq \frac{\|x_1\| + \cdots + \|x_{N-1}\| + \|x_N\| + \cdots + \|x_n\|}{n} \\ &\leq \frac{C}{n} + \frac{(n - N + 1)}{n}\varepsilon \end{aligned}$$
where $C = \|x_1\| + \cdots + \|x_{N-1}\|$. Now chose $N' \geq C/\varepsilon$, then for $n \geq N'$,
$$\left\| \frac{x_1 + x_2 + \cdots + x_n}{n} \right\| \leq \varepsilon + \epsilon$$
which shows the convergence. □

## 4. Convergence of series

In this section, we gather a few facts that will be used in this course.

Let $E$ be a Banach space. To a sequence $(x_n)_{n \geq 0}$ in $E$ we associate the sequence $(S_N)_{N \geq 0}$ of partial sums $S_N = \sum_{n=0}^{N}$.

Definition 4.1. We say that the series $\sum_{n=0}^{\infty} x_n$ *converges* if the sequence of partial sums $S_N$ converges and in this case, $\sum_{n=0}^{\infty} x_n = \lim S_N$.

We will say that $\sum_{n=0}^{\infty} x_n$ is *normally convergent* if the (real) series $\sum_{n=0}^{\infty} \|x_n\|$ converges

Note that in $\mathbb{R}$ or $\mathbb{C}$, normal convergence is called absolute convergence.

We will need the following two lemmas:

LEMMA 4.2. *Let $(x_n)$ be a sequence in a Banach space $E$. If $\sum_{n=0}^{\infty} x_n$ is normally convergent it is convergent. The converse is false*

PROOF. The converse is already false in $\mathbb{R}$ as shows the classical example $x_n = (-1)^n/n$.

The other direction is a direct copy of the proof in the real case. Assume that $\sum_{n=0}^{\infty} \|x_n\|$ is convergent and consider the partial sums $S_N = \sum_{n=0}^{N} x_n$ then for $N \geq M$,

$$\|S_N - S_M\| = \left\| \sum_{n=M+1}^{N} x_n \right\| \leq \sum_{n=M+1}^{N} \|x_n\|$$

from which one deduces that $S_N$ is a Cauchy sequence and is thus convergent. □

It turns out that this property characterizes Banach spaces, a fact that is sometimes convenient to show that a space is complete:

LEMMA 4.3. *Let $E, \|.\|$ be a normed vector space in which every normally convergent series is convergent. Then $E$ is a Banach space.*

PROOF. Assume $E$ has this property and let $(x_n)$ be a Cauchy sequence.

First, using the definition of a Cauchy sequence, we can construct an increasing sequence $n_k$ of integers such that $\left\|x_{n_{k+1}} - x_{n_k}\right\| \leq 10^{-k}$.

Then define $u_k = x_{n_{k+1}} - x_{n_k}$ and note that

— the series $\sum u_k$ is normally convergent, thus convergent

— as $x_{n_N} = \sum_{k=0}^{N-1} u_k + x_{n_0}$, the sequence $x_{n_k}$ is therefore convergent.

Finally, one uses the follwing simple fact

LEMMA 4.4. *A Cauchy sequence with a convergent sub-sequence is convergent.*

□

PROOF OF LEMMA 4.4. Let $(f_k)$ be a Cauchy sequence and assume that a subsequence converges: $f_{k_j} \to f$ when $j \to \infty$.

Let $\varepsilon > 0$, there exists $N$ such that, if $k, l \geq N$, $\|f_k - f_l\|_p \leq \varepsilon/2$. There exists $J$ such that, if $j \geq J$, $k_j \geq N$ (since $k_j \to \infty$ by definition of a subsequence) and $\left\|f_{k_j} - f\right\|_p \leq \varepsilon/2$ (the sequence $(f_{k_j})_j$ converges to $f$). But then $\|f_k - f\|_p \leq \left\|f_k - f_{k_j}\right\|_p + \left\|f_{k_j} - f\right\|_p \leq \varepsilon$ which shows that $f_k \to f$. □

CHAPTER 2

# $L^p$ spaces

## 1. Basics of integration theory

In this lecture, $(\Omega, \mathcal{B}, \mu)$ will be a $\sigma$-finite measure space. Recall that this means that there exists a countable family $(\Omega_n)_{n\geq 1}$ such that

(1) $n \geq 1$, $\Omega_{n+1} \subset \Omega_n$;
(2) $\bigcup_{n\geq 1} \Omega_n = \Omega$;
(3) $\mu(\Omega_n) < +\infty$.

Note that if we consider $\tilde{\Omega}_n = \Omega_{n+1} \setminus \Omega_n$, we obtain a family that satifies

(1) $n \neq m \geq 1$, $\tilde{\Omega}_m \cap \tilde{\Omega}_n = \emptyset$;
(2) $\bigcup_{n\geq 1} \tilde{\Omega}_n = \Omega$;
(3) $\mu(\tilde{\Omega}_n) < +\infty$.

Example 1.1. The three basic examples one should keep in mind are the following:

(1) $\Omega$ an open subset of $\mathbb{R}^d$ endowed with the Lebesgue measure $\mathrm{d}x$, or more generally, a measure of the form $\omega(x)\mathrm{d}x$ for some (measurable) weight function $\omega : \Omega \to \mathbb{R}^+$.
 Note that if $\omega = 1$ then this measure space is finite if and only if $\Omega$ is bounded.
(2) $\Omega = \{1, \ldots, d\}$ endowed with the counting measure $\mu(\{k\}) = 1$. Recall that a function $f$ on $\Omega$ is just a vector $f = \big(f(1), \ldots, f(d)\big) \in \mathbb{C}^d$ and that
$$\int f(x)\,\mathrm{d}\mu(x) = \sum_{k=1}^{d} f(k).$$
(3) $\Omega = \mathbb{N}$ endowed with the counting measure $\mu(\{k\}) = 1$. Recall that a function $f$ on $\Omega$ is now a sequence $f = \big(f(k)\big)_{k\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ and that
$$\int f(x)\,\mathrm{d}\mu(x) = \sum_{k=0}^{\infty} f(k).$$

We will need a few results from your course on integration.

Theorem 1.2 (Fubini).
*Let $(\Omega_1, \mathcal{B}_1, \mu_1)$ and $(\Omega_2, \mathcal{B}_2, \mu_2)$ be two $\sigma$-*finite *measure spaces and let $f$ be measurable on $\Omega_1 \times \Omega_2$.*

*If $f \geq 0$, then the three following integrals are equal (eventually $+\infty$)*
$$\int_{\Omega_1\times\Omega_2} f(x,y)\,d\mu_1\otimes\mu_2(x,y), \qquad \int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)\,d\mu_2(y)\right)d\mu_1(x),$$
*and*
$$\int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)\,d\mu_1(x)\right)d\mu_2(y).$$

*In particular, the functions* $x \to \int_{\Omega_2} f(x,y)\,d\mu_2(y)$ *and* $y \to \int_{\Omega_2} f(x,y)\,d\mu_1(x)$ *are mesurable.*

*If $f$ takes real or complex values, the same conclusion holds provided*

$$\int_{\Omega_1\times\Omega_2} |f(x,y)|\,d\mu_1\otimes\mu_2(x,y) < +\infty.$$

*In this case, the three integrals are finite*

THEOREM 1.3 (Monotone Convergence – Beppo-Levi).
*Let $(\Omega,\mathcal{B},\mu)$ be a measure space. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions on $\Omega$ such that*

(i) *for $\mu$-almost every $t\in\Omega$, $f_n(t)$ is increasing;*
(ii) *$f_n \geq 0$ $\mu$-almost everywhere.*

*Then* $\lim f_n$ *exists $\mu$-almost everywhere (eventually* $\lim f_n(t) = +\infty$*) and*

$$\lim_{n\to+\infty}\int_\Omega f_n(t)\,d\mu(t) = \int_\Omega \lim_{n\to+\infty} f_n(t)\,d\mu(t)$$

*(in particular, the left hand side and the right hand side are simultaneously finite or infinite).*

THEOREM 1.4 (Dominated Convergence – Lebesgue).
*Let $(\Omega,\mathcal{B},\mu)$ be a measure space. Let $(f_n)_{n\in\mathbb{N}}$ and $f$ bemesurable functions on $\Omega$ such that*

(i) *for $\mu$-almost every $t\in\Omega$, $f_n(t)\to f(t)$ when $n\to+\infty$*
(ii) *there exists a measure function $\varphi$ on $\Omega$ such that,*
 (a) *$\varphi$ is non-negative;*
 (b) *for $\mu$-almost every $t\in\Omega$, $|f_n(t)|\leq\varphi(t)$, thus $|f(t)|\leq\varphi(t)$;*
 (c) *$\varphi$ is integrable: $\int_\Omega \varphi(t)\,\mathrm{d}\mu(t) < +\infty$.*

*Then* $\int_\Omega f_n(t)\,d\mu(t) \to \int_\Omega f(t)\,d\mu(t)$ *when $n\to+\infty$.*

Note that, contrary to Beppo-Levi's theorem, $f_n$ is not required to be non-negative. This hypothesis is replaced by the domination hypothesis (the function $\varphi$).

Two important consequences of Lebesgue's theorem are the following results about continuity and differentiability of integrals depending on a parameter

COROLLARY 1.5 (Continuity of intégrals depending on a parameter – Lebesgue).
*Let $(\Omega,\mathcal{B},\mu)$ be a measure space and $(X,d)$ a metric space. Lett $F:\Omega\times X\to\mathbb{C}$ be a function such that:*

(i) *for $\mu$-almost every $t\in\Omega$, $x\mapsto F(t,x)$ is continuous;*
(ii) *there exists an integrable function $\varphi$ on $\Omega$ such that, for every $x\in X$ and $\mu$-almost every $t\in\Omega$, $|F(t,x)|\leq|\varphi(t)|$.*

*Then $f:X\to\mathbb{C}$ defined by* $f(x) = \int_\Omega F(t,x)\,d\mu(t)$ *is continuous on $X$.*

PROOF. Let $x$ et $x_0\in X$, then

$$f(x)-f(x_0) = \int_\Omega F(t,x)\,\mathrm{d}\mu(t) - \int_\Omega F(t,x_0)\,\mathrm{d}\mu(t) = \int_\Omega \big(F(t,x)-F(t,x_0)\big)\,\mathrm{d}\mu(t).$$

But, for $\mu$ almost every $t$, $|F(t,x)-F(t,x_0)| \leq 2\varphi(t)\in L^1$ and, by continuity of $F$ in $x_0$, $F(t,x)-F(t,x_0)\to 0$ when $x\to x_0$. By dominated convergence, $f(x)\to f(x_0)$ when $x\to x_0$, thus $f$ is continuous in $x_0$. □

COROLLARY 1.6 (Differentiability of intégrals depending on a parameter – Lebesgue).

*Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and $X \subset \mathbb{R}^d$ be open. Let $F : \Omega \times X \to \mathbb{C}$ a function such that:*

(i) *for $\mu$-almost every $t \in \Omega$, $x \mapsto F(t, x)$ is differentiable;*
(ii) *there exists an integrable function $\varphi$ on $\Omega$ such that, for every $x \in X$ and $\mu$-almost every $t \in \Omega$, $|F(t, x)| \le |\varphi(t)|$;*
(iii) *there exists an integrable function $\psi$ on $\Omega$ such that, for every $x \in X$ and $\mu$-almost every $t \in \Omega$, $\left|\dfrac{\partial F}{\partial x}(t, x)\right| \le |\psi(t)|$;*

*Then $f : X \to \mathbb{C}$ defined by $f(x) = \displaystyle\int_\Omega F(t, x)\, d\mu(t)$ is differentiable on $X$.*

We leave the proof of this fact as an exercice.

## 2. $L^p$-spaces: defintion

Let $1 \le p < +\infty$ be a real number and $(\Omega, \mathcal{B}, \mu)$ a $\sigma$-finite measure space. We define

$$L^p(\Omega, \mu) = \left\{ f : \Omega \to \mathbb{C},\ f\ \mu - \text{measurable},\ \int_\Omega |f(x)|^p \,\mathrm{d}\mu(x) < +\infty \right\}.$$

This space is endowed with the "norm" défined by

$$\|f\|_p = \left( \int_\Omega |f(x)|^p \,\mathrm{d}\mu(x) \right)^{\frac{1}{p}}.$$

This defines a norm in the following sense:

(i) For $f \in L^p(\Omega, \mu)$, we havea $\|f\|_p \ge 0$ and $\|f\|_p = 0$ if and only if $f = 0$ $\mu$-almost everywhere.
(ii) For $f \in L^p(\Omega, \mu)$ and $\lambda \in \mathbb{C}$, we have $\lambda f \in L^p(\Omega, \mu)$ and $\|\lambda f\|_p = |\lambda| \|f\|_p$.
(iii) For $f, g \in L^p(\Omega, \mu)$, and $f + g \in L^p(\Omega, \mu)$ we have $\|f + g\|_p \le \|f\|_p + \|g\|_p$.

REMARK 2.1. It is important to notice that, in the particular case $p = 2$, $L^2\Omega, \mu)$ is a Hilbert space and that the norm is associated to the scalar product

$$\langle f, g \rangle_{L^2(\Omega,\mu)} = \int_\Omega f(x)\overline{g(x)} \,\mathrm{d}\mu(x).$$

SKETCH OF PROOF. Note that (ii) is obvious. For (i), we use the fact that if a non-negative function has 0 integral, then it vanishes almost everywhere. The last point, (iii) is much more subtle (excepted when $p = 1$ and when $p = 2$ when it is a consequence of the Cauchy-Schwarz inequality) and will be proved later. However, it is easy to prove that $L^p(\Omega, \mu)$ is a vector space. Indeed

$$|f + g|^p \le (|f| + |g|)^p = 2^p \left( \frac{|f| + |g|}{2} \right)^p \le 2^{p-1}(|f|^p + |g|^p)$$

since $x \mapsto x^p$ is convex on $[0, +\infty)$ when $p \ge 1$. Thus, if $f, g \in L^p(\Omega, \mu)$ then $f + g \in L^p(\Omega, \mu)$. $\square$

For $p = +\infty$, one defines

$$L^\infty(\Omega, \mu) = \{ f : \Omega \to \mathbb{C},\ f\ \mu\text{−measurable},\ \text{il existe } K > 0 \text{ telle que } |f(x)| \le K,\ \mu\text{−}p.p. \}.$$

One then endows this space with the "norm" (with the same meaning as previously)

$$\|f\|_\infty = \inf\{K \ \ |f(x)| \le K \ \mu - a.e.\}.$$

REMARK 2.2.
— We will simply write $L^p(\Omega) = L^p(\mu) = L^p := L^p(\Omega,\mu)$ when either $\Omega,\mu$ or both are implicitly fixed When $\mu$ is the counting measure, one often write $\ell^p(\Omega) = L^p(\Omega,\mu)$.
— It is important to understand that $\|f\|_\infty$ is *not the supremum* of $f$ but the *essential supremum*. The two quantities don't coincide in general. For instance, if $f$ is defined on $\mathbb{R}$ by $f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{sinon} \end{cases}$ then $\sup|f| = 1$ while $\|f\|_\infty = 0$ ($\mathbb{Q}$ has measure 0 so that $f(x) = 0$ almost everywhere for the Lebesgue measure).

Of course, if $f$ is continuous, then $\sup|f| = \|f\|_\infty$.

The notation $L^\infty$ is justified by the following fact that we leave as an exercice:

EXERCICE 2.3.

(1) Show that, if $f \in L^q \cap L^\infty$ then, for $p > q$, $f \in L^p$.
(2) Show that, moreover, $\|f\|_p \to \|f\|_\infty$ when $p \to +\infty$.

The "norms" that we have just defined do not distinguish *all* measurable functions in the sense that $\|f-g\|_p = 0$ only implies that $f = g$ $\mu$-*almost* everywhere and not everywhere.

To avoid that nuisance, one can re-define $L^p(\Omega,\mu)$ so that its elements are "equivalence classes" of functions. More precisely, if $f \in L^p(\Omega,\mu)$, we may define $\tilde f$ as the set of all functions $h$ such that $f-h = 0$ $\mu$-almost everywhere and we then write $h \sim f$. In particular, $h \in L^p(\Omega,\mu)$ and $\|h\|_p = \|f\|_p$. Moreover, if $f \sim h$ and $h \sim g$ then $f \sim g$ (a finite union – and even a countable one– of sets of measure zero still has measure zero), thus $\tilde f = \tilde h$. Finally, we define

$$\|\tilde f\|_p = \|f\|_p$$

and this does not depend on the choice of $f$ in $\tilde f$.

Finally, if $\tilde f$ and $\tilde g$ are the equivalence classes of some $f$ and $g \in L^p(\Omega,\mu)$ and $\lambda,\mu \in \mathbb{C}$, we define $\lambda\tilde f + \mu\tilde g = (\lambda f + \mu g)\tilde{}$. It is easy to see that this does not depend on the choice of $f$ and $g$ in $\tilde f$ and $\tilde g$.

We now have two vector spaces. One consists of functions while the other one is a set of equivalence classes of functions. In the first one, the norm is not really a norm ($\|f\| = 0$ does not imply $f = 0$) while in the second one, it is!

Both spaces are denoted by $L^p(\Omega,\mu)$ and we will use the following *abuse of language*: *let $f$ be a function in $L^p(\Omega,\mu)$* to mean *let $\tilde f \in L^p(\Omega,\mu)$ and let $f \in \tilde f$*. This is very confortable, but one has to keep in mind that $f = g$ means $f = g$ $\mu$-almost everywhere. Further, if $x_0 \in \Omega$, then $f(x_0)$ does not make sense if $\mu(\{x_0\}) = 0$.

Of course if, for some reason, one knows that $\tilde f$ contains a (necessarily unique) continuous function, then the $f \in \tilde f$ we chose is this continuous function and $f(x_0)$ has its usual meaning.

Also, if $\mu$ is the counting measure, *i.e.* in $\ell^p$, this problem does not occur since $\mu(\{x_0\}) = 1$.

REMARK 2.4. We would like to stress that $f$ is continuous almost everywhere is not the same thing as $\tilde f$ contains a continuous function.

For instance if $f = \mathbf{1}_{\mathbb{Q}}$ is the function defined by $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{otherwise} \end{cases}$ then $f$ is nowhere continuous but $f = 0$ almost everywhere and the 0 function is of course continuous.

An other example is the following: $f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 1/x & \text{otherwise} \end{cases}$, then $f$ is continous almost everywhere but there is no continuous function $g$ on $\mathbb{R}$ such that $f(x) = g(x)$ for almost every $x$.

## 3. $L^p$-spaces: Hölder and Minkowski

Theorem 3.1 (Hölder's Inequality).
*Let $(\Omega, \mathcal{B}, \mu)$ be a measured space. Let $1 \le p, p' \le +\infty$ be such that $\dfrac{1}{p} + \dfrac{1}{p'} = 1$ (with the convention $p' = +\infty$ when $p = 1$ and vice versa). Let $f \in L^p(\Omega, \mu)$ and $g \in L^{p'}(\Omega, \mu)$, then $fg \in L^1(\Omega, \mu)$ and*

$$\left| \int_\Omega f(x)g(x)\, d\mu(x) \right| \le \int_\Omega |f(x)g(x)|\, d\mu(x) \le \left( \int_\Omega |f(x)|^p\, d\mu(x) \right)^{\frac{1}{p}} \left( \int_\Omega |g(x)|^{p'}\, d\mu(x) \right)^{\frac{1}{p'}}.$$

*Moreover,*

*— equality occurs and the first inequality if and only if there is a $\theta \in \mathbb{R}$ such that $f(x)g(x) = e^{i\theta}|f(x)g(x)|$.*

*— if $f \ne 0$ equality occurs and the second inequality if and only if there is a real $\lambda \ge 0$ such that*

(i) *when $1 < p < +\infty$, $|g(x)| = \lambda |f(x)|^{p-1}$ $\mu$-almost everywhere;*
(ii) *when $p = 1$, $|g(x)| \le \lambda$ $\mu$-almost everywhere and $|g(x)| = \lambda$ for $\mu$-almost every $x$ for which $f(x) \ne 0$;*
(iii) *when $p = +\infty$, $|f(x)| \le \lambda$ $\mu$-almost everywhere and $|f(x)| = \lambda$ for $\mu$-almost every $x$ for which $g(x) \ne 0$;*

Remark 3.2.
— When $p = 2$ then $q = 2$ then Hölder's Inequality is the well known Cauchy-Schwarz Inequality.

— The first inequality and its equality case are trivial facts following from the positivity of the integral of positive functions.

Exercice 3.3.
Show that, if $1 \le p_i \le +\infty$ are such that $\displaystyle\sum_{i=1}^m \frac{1}{p_i} = 1$, then

$$\left| \int_\Omega \prod_{i=1}^m f_i(x)\, \mathrm{d}\mu(x) \right| \le \prod_{i=1}^m \|f_i\|_{p_i}.$$

Exercice 3.4.
Endow $\mathbb{R}$ with the Lebesgue measure $\mathrm{d}x$. For a function $f$ on $\mathbb{R}$ and $\lambda > 0$, define a new function $f_\lambda$ on $\mathbb{R}$ through $f_\lambda(x) = f(\lambda x)$.

(1) Express $\|f_\lambda\|_p$ in terms of $\|f\|_p$.
(2) Assume that there is a constant $C$ such that, for every $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$, we have $fg \in L^r(\mathbb{R})$ and

$$\|fg\|_r \le C \|f\|_p \|g\|_q \tag{3.1}$$

Replacing $f, g$ by $f_\lambda$, $g_\lambda$ in (3.1) and letting $\lambda$ vary between 0 and $+\infty$, which relation can you deduce on $p, q, r$.

(3) Assume now that $p, q, r$ satisfy this relation. Prove (3.1) using Hölder.

Proof. When either $f = 0$ or $g = 0$, there is notinhg to prove.

The cases $p = 1$ and $p = +\infty$, are trivial. We may thus assume that $1 < p < \infty$ thus $1 < p' < \infty$ and $f, g \neq 0$. This allows to define

$$u = \left(\frac{|f|}{\|f\|_p}\right)^p \quad \text{and} \quad v = \left(\frac{|g|}{\|g\|_{p'}}\right)^{p'}.$$

As log is concave, one gets that, for $0 < \alpha < 1$, $u^\alpha v^{1-\alpha} \leq \alpha u + (1-\alpha)v$. In particular, taking $\alpha = 1/p$, we get

$$\frac{|f|}{\|f\|_p}\frac{|g|}{\|g\|_{p'}} \leq \frac{1}{p}\frac{|f|^p}{\|f\|_p^p} + \frac{1}{p'}\frac{|g|^{p'}}{\|g\|_{p'}^{p'}}.$$

Integrating with respect to $\mu$, we get the result.

For the equality case, log is actually strictly concave, which implies that $0 < \alpha < 1$, $u^\alpha v^{1-\alpha} < \alpha u + (1-\alpha)v$ unless $u = v$. The result follows. □

Theorem 3.5 (Jensen's Inequality).
*Let $(\Omega, \mathcal{B}, \mu)$ be a measure space with $\mu$ a* finite *measure. Let $J : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1$ convex function. For $f \in L^1(\Omega, \mu)$, write*

$$\langle f \rangle = \frac{1}{\mu(\Omega)} \int_\Omega f(x)\, d\mu(x)$$

*for its* mean *over $\Omega$. Then*

(i) *$[J \circ f]_-$, the negative part of $J \circ f$ is in $L^1(\Omega, \mu)$, thus $\displaystyle\int_\Omega J \circ f(x)\, d\mu(x)$ is well defined (possibly $+\infty$);*

(ii) *$J(\langle f \rangle) \leq \langle J \circ f \rangle$, that is*

$$J\left(\frac{1}{\mu(\Omega)} \int_\Omega f(x)\, d\mu(x)\right) \leq \frac{1}{\mu(\Omega)} \int_\Omega J\big(f(x)\big)\, d\mu(x).$$

Proof. As $J$ is convex and $\mathcal{C}^1$, for $a, t \in \mathbb{R}$,

$$J(t) \geq J(a) + J'(a)(t-a).$$

Taking $t = f(x)$ and $a = \langle f \rangle$, this leads to

$$J\big(f(x)\big)_+ - J\big(f(x)\big)_- = J\big(f(x)\big) \geq J(\langle f \rangle) + J'(\langle f \rangle) f(x) - J'(\langle f \rangle)\langle f \rangle. \tag{3.2}$$

In particular, let $x$ be such that $J\big(f(x)\big)_- \neq 0$ thus $J\big(f(x)\big)_+ = 0$, and

$$\begin{aligned} 0 \leq J\big(f(x)\big)_- &\leq -J'(\langle f \rangle) f(x) + J'(\langle f \rangle)\langle f \rangle - J(\langle f \rangle) \\ &\leq |J'(\langle f \rangle)||f(x)| + |J'(\langle f \rangle)\langle f \rangle - J(\langle f \rangle)|. \end{aligned}$$

As $f \in L^1$, $|J'(\langle f \rangle)||f(x)| \in L^1$ and $\mu$ is a *finite* measure, constant functions are integrables, thus $|J'(\langle f \rangle)\langle f \rangle - J(\langle f \rangle)| \in L^1$. the first part of the theorem is prove.

Finally, integrating (3.2), we get

$$\frac{1}{\mu(\Omega)} \int_\Omega J\big(f(x)\big)\, \mathrm{d}\mu(x) \geq \frac{1}{\mu(\Omega)} \int_\Omega J(\langle f \rangle)\, \mathrm{d}\mu(x) + \frac{J'(\langle f \rangle)}{\mu(\Omega)} \int_\Omega f(x) - \langle f \rangle\, \mathrm{d}\mu(x).$$

But
$$\frac{1}{\mu(\Omega)}\int_\Omega f(x)-\langle f\rangle\,\mathrm{d}\mu(x)=0$$
from which Jensen's inequality follows. □

The result is valid for any non-negative convex function $J$. This comes from the fact that such a function always satisfies an inequality of the form $J(t)\geq J(a)+c(t-a)$. Of course $c=J'(a)$ only when $J$ is differentiable in $a$.

From this, we can deduce a second proof of Hölder's inequality

Second proof of Hölder. Up to replacing $f$, $g$ by $|f|$, $|g|$, we may assume that $f,g\geq 0$. Again, the cases $p=1$ and $p=+\infty$ are obvious so that we also assume that $1<p<+\infty$.

Let $\Omega'=\{x\in\Omega\,:\,g(x)>0\}$. Then
$$\int_\Omega f^p(x)\,\mathrm{d}\mu(x)=\int_{\Omega'}f^p(x)\,\mathrm{d}\mu(x)+\int_{\Omega\setminus\Omega'}f^p(x)\,\mathrm{d}\mu(x)\geq\int_{\Omega'}f^p(x)\,\mathrm{d}\mu(x)$$
while
$$\int_\Omega f(x)g(x)\,\mathrm{d}\mu(x)=\int_{\Omega'}f(x)g(x)\,\mathrm{d}\mu(x)\quad\text{et}\quad\int_\Omega g(x)^{p'}\,\mathrm{d}\mu(x)=\int_{\Omega'}g(x)^{p'}\,\mathrm{d}\mu(x).$$
It follows that it is enough to prove Hölder's inequality with $\Omega'$ replacing $\Omega$, that is to say that $g$ does not vanish on $\Omega$.

We can then define the measure $\mathrm{d}\nu(x)=g(x)^{p'}\,\mathrm{d}\mu(x)$ and the function $F(x)=f(x)g(x)^{p'/p}$. Note that
$$\nu(\Omega)=\int_\Omega 1\,\mathrm{d}\nu(x)=\int_\Omega g(x)^{p'}\,\mathrm{d}\mu(x)$$
thus $\nu$ is a finite measure. Moreover
$$\begin{aligned}\frac{1}{\nu(\Omega)}\int_\Omega F(x)\,\mathrm{d}\nu(x)&=\frac{1}{\displaystyle\int_\Omega g(x)^{p'}\,\mathrm{d}\mu(x)}\int_\Omega f(x)g(x)^{-p'/p}g(x)^{p'}\,\mathrm{d}\nu(x)\\&=\frac{\displaystyle\int_\Omega f(x)g(x)\,\mathrm{d}\mu(x)}{\displaystyle\int_\Omega g(x)^{p'}\,\mathrm{d}\mu(x)}\end{aligned}$$
since $-\dfrac{p'}{p}+p'=p'\left(1-\dfrac{1}{p}\right)=1$. Finally, Jensen's Inequality with $J(t)=|t|^p$ yields
$$\left(\frac{\displaystyle\int_\Omega f(x)g(x)\,\mathrm{d}\mu(x)}{\displaystyle\int_\Omega g(x)^{p'}\,\mathrm{d}\mu(x)}\right)^p\leq\frac{\displaystyle\int_\Omega f(x)^pg(x)^{-p'}g(x)^{p'}\,\mathrm{d}\mu(x)}{\displaystyle\int_\Omega g(x)^{p'}\,\mathrm{d}\mu(x)}$$
which is what we wanted to prove. □

Exercice 3.6.

(1) Show that, if $J$ is strictly convex, then equality in Jensen's inequality only occurs when $f$ is constant.
(2) Deduce the equality cases in Hölder's Inequality from this.

Exercice 3.7. (Inclusion of $L^p$ spaces)

(1) Let $1 \leq p_1 < p_2 \leq +\infty$. Show that, if $f \in L^{p_1} \cap L^{p_2}$ then,
   (a) for all $p_1 < p < p_2$, $f \in L^p$,
   (b) the mapping $p \mapsto \log \int |f|^p$ is convex on $[p_1, p_2]$.
(2) For each $p \in [1, \infty]$, find $f \in L^p(\mathbb{R})$ such that $f \notin L^q(\mathbb{R})$ fpr $q \in [1, \infty]$, $q \neq p$.
(3) Under which condition on $p, q$ does the inclusion $\ell^p \subset \ell^q$ hold ?
(4) Let $(\Omega, \mathcal{B}, \mu)$ be a measured space and $1 \leq p < \infty$. show that $f \in L^p(\Omega, \mu)$, $d_f(t) := \mu(\{x \in \Omega : |f(x)| > \lambda\}) \leq \dfrac{\|f\|_p^p}{t^p}$.

THEOREM 3.8 (Minkowski's Inequality).
*Let $(\Omega, \mathcal{B}, \mu)$ and $(\Gamma, \tilde{\mathcal{B}}, \gamma)$ be two $\sigma$-finite measured and let $1 \leq p < +\infty$. Then, for every $\gamma \otimes \mu$-measurable function $f$,*

$$\left(\int_\Gamma \left(\int_\Omega |f(x,y)|\, d\mu(y)\right)^p d\gamma(x)\right)^{\frac{1}{p}} \leq \int_\Omega \left(\int_\Gamma |f(x,y)|^p\, d\gamma(x)\right)^{\frac{1}{p}} d\mu(y). \tag{3.3}$$

*In particular, if the right hand side is finite so is the left-hand side. Moreover, equality holds if and only if $f$ has separable variables, that is, $f$ is of the form $f(x,y) = \alpha(x)\beta(y)$.*

In other words

$$\left\| x \to \int_\Omega |f(x,y)|\, \mathrm{d}\mu(y) \right\|_p \leq \int_\Omega \|x \to f(x,y)\|_p\, \mathrm{d}\mu(y),$$

which extends the inequality $\left|\int_\Omega f(t)\, \mathrm{d}\mu(t)\right| \leq \int_\Omega |f(t)|\, \mathrm{d}\mu(t)$ (which is the particular case of $\Gamma$ consisting of a single element).

PROOF. It is enough to assume that $f \geq 0$ and that $f > 0$ on a set of positive measure.

Note also that Fubini's theorem implies that $y \to \displaystyle\int f(x,y)^p\, \mathrm{d}\gamma(x)$ are $H : x \to \displaystyle\int_\Omega f(x,y)\, \mathrm{d}\mu(y)$ measurable functions.

Assume the right hand side is finite, otherwise there is nothing to prove.

Let $f_n = f\chi_{E_n}$ with $E_n = F_n \cap \{(x,y) \in \Gamma \times \Omega : |f(x,y)| \leq n\}$ and $F_n$ an increasing sequence of finite measure subsets of $\Gamma \times \Omega$ that cover $\Gamma \times \Omega$: $\bigcup F_n = \Gamma \times \Omega$. For $f_n$, the left hand side of (3.3),

$$\left(\int_\Gamma \left(\int_\Omega |f_n(x,y)|\, \mathrm{d}\mu(y)\right)^p \mathrm{d}\gamma(x)\right)^{\frac{1}{p}}$$

is finite.

On the other hand, the monotone convergence theorem shows that this converges to

$$\left(\int_\Gamma \left(\int_\Omega |f(x,y)|\, \mathrm{d}\mu(y)\right)^p \mathrm{d}\gamma(x)\right)^{\frac{1}{p}}.$$

We may thus assume that this quantity is also finite.

According to Fubini (Beppo-Levi),

$$\begin{aligned}
\int_\Gamma H(x)^p\, \mathrm{d}\gamma(x) &= \int_\Gamma \left(\int_\Omega f(x,y)\, \mathrm{d}\mu(y)\right) H(x)^{p-1}\, \mathrm{d}\gamma(x) \\
&= \int_\Omega \int_\Gamma f(x,y) H(x)^{p-1}\, \mathrm{d}\gamma(x)\, \mathrm{d}\mu(y).
\end{aligned}$$

But, Hölder $(1/p + 1/p' = 1)$ implies

$$\begin{aligned}\int_\Gamma f(x,y)H(x)^{p-1}\,\mathrm{d}\gamma(x) &\le \left(\int_\Gamma f(x,y)^p\,\mathrm{d}\gamma(x)\right)^{1/p}\left(\int_\Gamma H(x)^{(p-1)p'}\,\mathrm{d}\gamma(x)\right)^{1/p'}\\ &= \left(\int_\Gamma f(x,y)^p\,\mathrm{d}\gamma(x)\right)^{1/p}\left(\int_\Gamma H(x)^{p}\,\mathrm{d}\gamma(x)\right)^{1-1/p}.\end{aligned}$$

It follows that

$$\int_\Gamma H(x)^p\,\mathrm{d}\gamma(x) \le \int_\Omega\left(\int_\Gamma f(x,y)^p\,\mathrm{d}\gamma(x)\right)^{1/p}\mathrm{d}\mu(y)\left(\int_\Gamma H(x)^p\,\mathrm{d}\gamma(x)\right)^{1-1/p}.$$

As we have assumed that $\int_\Gamma H(x)^p\,\mathrm{d}\gamma(x) \neq 0, +\infty$, we can divide both sides by

$$\left(\int_\Gamma H(x)^p\,\mathrm{d}\gamma(x)\right)^{1-1/p}$$

from which the result follows. □

Corollary 3.9 (Triangular Inequality for $L^p$ norms).
*Let $(\Omega, \mathcal{B}, \mu)$ be a measure space with $\mu$ a $\sigma$-finite measure. Let $1 \le p \le +\infty$ and $f, g \in L^p(\Omega, \mu)$. Then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

*and equality holds if and only if $g = \lambda f$ with $\lambda \ge 0$.*

Proof. Take $\Gamma = \{1,2\}$ with the coounting measure and define $F(1,y) = f(y)$, $F(2,y) = g(y)$. Minkowski's inequality then reduces to the triangular inequality □

Sketch of a second proof. There is a simpler more direct argument:

$$\begin{aligned}\int_\Omega |f(x)+g(x)|^p\,\mathrm{d}\mu(x) &= \int_\Omega |f(x)+g(x)|^{p-1}|f(x)+g(x)|\,\mathrm{d}\mu(x)\\ &\le \int_\Omega |f(x)+g(x)|^{p-1}|f(x)|\,\mathrm{d}\mu(x)\\ &\quad+\int_\Omega |f(x)+g(x)|^{p-1}|g(x)|\,\mathrm{d}\mu(x).\end{aligned}$$

Hölder's Inequality ($\frac{1}{p} + \frac{1}{p'} = 1$ that is $p' = \frac{p}{p-1}$) leads to

$$\begin{aligned}&\int_\Omega |f(x)+g(x)|^{p-1}|f(x)|\,\mathrm{d}\mu(x)\\ &\qquad\le \left(\int_\Omega |f(x)+g(x)|^{(p-1)\frac{p}{p-1}}\,\mathrm{d}\mu(x)\right)^{1-\frac{1}{p}}\left(\int_\Omega |f(x)|^p\,\mathrm{d}\mu(x)\right)^{\frac{1}{p}}.\end{aligned}$$

The second integral is treated the same way. It follows that

$$\int_\Omega |f(x)+g(x)|^p\,\mathrm{d}\mu(x) \le \left(\int_\Omega |f(x)+g(x)|^p\,\mathrm{d}\mu(x)\right)^{1-\frac{1}{p}}\left(\|f\|_p + \|g\|_p\right)$$

one concludes by dividing by $\|f+g\|_p^p$. Of course, one should first check the trivial case $\|f+g\|_p^p \neq 0$ but also prove that $\|f+g\|_p^p < +\infty$, which was done when we introduced the norm. □

## 4. Completeness of $L^p$ spaces

The aim of this section is to prove that $L^p$ is a Banach space. Before this, let us adapt dominated convergence to convergence in $L^p$:

LEMMA 4.1 ($L^p$-dominated convergence). *Let $(\Omega, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space. Let $1 \le p < +\infty$.*

*Let $(f_k)$ be a sequence in $L^p(\Omega, \mu)$ and $f, F$ be two functions in $L^p(\Omega, \mu)$. Assume that*

(i) *for every $k$, and $\mu$-almost every $x \in \Omega$, $|f_k(x)| \le F(x)$*
(ii) *for $\mu$-almost every $x \in \Omega$, $f_k(x) \to f(x)$ when $k \to +\infty$. In particular, $|f(x)| \le F(x)$ $\mu$-a.e.*

*Then $f_k \to f$ in $L^p(\Omega, \mu)$ i.e. $\|f_k - f\|_p \to 0$.*

PROOF. We have to prove that

$$\int_\Omega |f_k(x) - f(x)|^p \,\mathrm{d}\mu(x) \to 0.$$

But Condition (ii) implies that $|f_k(x) - f(x)|^p \to 0$ $\mu$-a.e.

Condition (ii) implies that

$$|f_k(x) - f(x)|^p \le (|f_k(x)| + |f(x)|)^p \le \big(2F(x)\big)^p$$

and the hypothesis $F \in L^p$ precisely means that $\int_\Omega F(x)\,\mathrm{d}\mu(x) < +\infty$. We can thus apply the dominated convergence theorem to obtain the result. □

THEOREM 4.2 ($L^p$ is complete).
*Let $(\Omega, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space. Let $1 \le p \le +\infty$. Then $L^p(\Omega, \mu)$ is complete (and thus a Banach space).*

*More precisely, if $(f_k)$ is a Cauchy sequence in $L^p(\Omega, \mu)$, then the exists a sub-sequence $(f_{k_j})_j$ and $F$ in $L^p(\Omega, \mu)$ such that*

(i) *for $j \ge 1$, $|f_{k_j}(x)| \le F(x)$ and $\mu$-almost every $x \in \Omega$;*
(ii) *for $\mu$-almost every $x \in \Omega$, $f_{k_j}(x) \to f(x)$ when $j \to +\infty$.*

PROOF. We will concentrate on the case $1 \le p < +\infty$. The case $p = +\infty$ follows mainly from the completeness of $\mathbb{C}$ and is left as an exercice.

As noted in the above lemma, the second part of the theorem implies that every Cauchy sequence in $L^p$ has a convergent sub-sequence. But, as we have already noticed in the first chapter, a Cauchy sequence with a convergent sub-sequence is convergent (Lemma 4.4).

The proof of the second part of the theorem is rather classical.

First, there exists $i_1$ such that, if $n \ge i_1$, $\|f_{i_1} - f_n\|_p \le 1/2$ ($\varepsilon = 1/2$ in the definition of a Cauchy sequence). There exists $i_2 > i_1$ such that, if $n \ge i_2$, $\|f_{i_2} - f_n\|_p \le 1/2^2$... This way, we inductively define $i_k > i_{k-1}$ such that, if $n \ge i_k$, $\|f_{i_k} - f_n\|_p \le 1/2^k$.

Consider the *non-decreasing positive* sequence defined by

$$F_l(x) = |f_{i_1}(x)| + \sum_{k=1}^{l} |f_{i_{k+1}}(x) - f_{i_k}(x)|.$$

The triangular inequality yields

$$\|F_l\|_p \le \|f_{i_1}\|_p + \sum_{k=1}^{l} \left\|f_{i_{k+1}} - f_{i_k}\right\|_p \le \|f_{i_1}\|_p + \sum_{k=1}^{+\infty} 2^{-k} = 1 + \|f_{i_1}\|_p < +\infty.$$

The monotone convergence theorem implies that $F_l$ converges almost everywhere to a function $F \in L^p$. In particular, $F(x)$ is finite for $\mu$-almost every $x \in \Omega$. For such an $x$, the series

$$f_{i_1}(x) + \sum_{k=1}^{l} \big(f_{i_{k+1}}(x) - f_{i_k}(x)\big)$$

is absolutely convergent, thus convergent. But this is a telescopic sequence:

$$f_{i_1}(x) + \sum_{k=1}^{l} \big(f_{i_{k+1}}(x) - f_{i_k}(x)\big) = f_{i_{l+1}}(x).$$

We have thus shown that $f_{i_{l+1}}$ is convergent and, with the triangular inequality, $|f_{i_{l+1}}| \leq F_l \leq F$ which completes the proof. $\square$

## 5. Separability of $L^p$ spaces

**5.1. About the Lebesgue integral.** Recall that a Banach space $X$ is said to be *separable* if there exists a countable subset $\mathcal{D} \subset X$ that is dense in $X$ *i.e.* such that, if $x \in X$ for every $\varepsilon > 0$, there exists $y \in \mathcal{D}$ such that $\|x - y\| < \varepsilon$.

Let $(\Omega, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space. We assume that this space has a further property: there exists a countable family $\tilde{\mathcal{B}} \subset \mathcal{B}$ such that, given $A \in \mathcal{B}$, for every $\varepsilon > 0$, there exists $B \in \tilde{\mathcal{B}}$ such that $\mu(A\Delta B) < \varepsilon$. Here $A\Delta B$ is the symetric difference $A\Delta B = (A \setminus B) \cup (B \setminus A)$.

Of course, this property is satified when $\Omega$ is countable. It is a bit more complicated when $\Omega$ is an open subset of $\mathbb{R}^d$. Actually, as a Borel set is optained from a countable number of operations (countable unions and intersections, complementary) from cubes (the $\sigma$-algebra generated by cubes is the Borel $\sigma$-algebra), it is enough to find a countable family such that the above property holds whenever $A$ is a cube. As cubes are finite products of intervals, it is enough to do so for intervals. But then, it is easy to see that

$$\tilde{B} = \left\{ \bigcup_{F \in \mathcal{F}} F \,:\, F \subset \{[k/2^j, (k+1)/2^j],\ k \in \mathbb{Z}, j \in \mathbb{N}\},\ F \text{ finite} \right\}$$

has the desired property. In other words: the Borel sets are generated by "dyadic intervals".

Next, recall that a simple function on $\Omega$ is a function of the form

$$s(x) = \sum_{k \in K} a_k \mathbf{1}_{A_k} \tag{5.4}$$

with $K$ finite and $0 < \mu(A_k) < \infty$. We can (and will) further assume that the $A_k$'s are disjoint. We can then definie the Lebesgue integral of a non negative function $f$ as

$$\int_\Omega f(x)\,\mathrm{d}\mu(x) = \sup \left\{ \sum_{k \in K} a_k \mu(A_k) \,:\, s \text{ given by (5.4) satisfies } s \leq f \right\}.$$

Note that this quantity can be $+\infty$. Note also that one can simply request $a_k$ to be rational and the $A_k$'s to be in $\tilde{B}$ so that a countable family $\mathcal{D}_0$ of simple non-negative functions suffices.

When $f$ is real valued, we then say that $f$ is Lebesgue integrable if $\int_\Omega |f(x)|\,\mathrm{d}\mu(x) < +\infty$. It follows that both its positive part $f_+ = \max(f, 0)$ and negative part $f_- =$

$\max(-f,0)$ are non-negative and have finite integral (since $s \le f_\pm \le |f|$). As $f = f_+ - f_-$ one of course defines

$$\int_\Omega f(x)\,\mathrm{d}\mu(x) = \int_\Omega f_+(x)\,\mathrm{d}\mu(x) - \int_\Omega f_-(x)\,\mathrm{d}\mu(x)$$

which makes sense since both quantities are finite (only one needs to be). For complex valued functions, one again requires $\int_\Omega |f(x)|\,\mathrm{d}\mu(x) < +\infty$. Then the real and imaginary parts of $f$ are integrale and

$$\int_\Omega f(x)\,\mathrm{d}\mu(x) = \int_\Omega \Re(f)(x)\,\mathrm{d}\mu(x) + i\int_\Omega \mathrm{Im}\,(f)(x)\,\mathrm{d}\mu(x).$$

From the above discution, the countable set

$$\mathcal{D} = \{f_1 - f_2 + i(f_3 - f_4), f_1, f_2, f_3, f_4 \in \mathcal{D}_0 \text{ with disjoint support}\}$$

has the following property: Given an integrable complex-valued function $f$, for every $\varepsilon > 0$, there exists $f_\varepsilon = f_\varepsilon^1 - f_\varepsilon^2 + i(f_\varepsilon^3 - f_\varepsilon^4) \in \mathcal{D}$ such that

$$\int_\Omega |f - f_\varepsilon|\,\mathrm{d}\mu(x) \le \varepsilon.$$

In other words, $\mathcal{D}$ is countable dense in $L^1(\Omega,\mu)$.

**5.2. Sepatability of $L^p(\Omega,\mu)$ for $1 \le p < \infty$.** Take $1 \le p < \infty$, fix $f \in L^p(\Omega,\mu)$, and $\varepsilon > 0$. First write $f = f_1 - f_2 + i(f_3 - f_4)$ with $f_i \ge 0$ and note that $f_i \le |f|$ so that each $f_i \in L^p(\Omega,\mu)$. Further, if we show that, for given $\varepsilon > 0$ there exists $f_\varepsilon^{(i)} \in \mathcal{S}_0$ such that $\left\|f_i - f_\varepsilon^{(i)}\right\|_p \le \varepsilon$ then the triangular inequality shows that, setting $f_\varepsilon = f_\varepsilon^{(1)} - f_\varepsilon^{(2)} + i(f_\varepsilon^{(3)} - f_\varepsilon^{(4)})$, we have $\|f - f_\varepsilon\|_p \le 4\varepsilon$. We can thus assume that $f \ge 0$.

Next, set $f_n = \mathbf{1}_{f \le n} f$ and note that $f_n \to f$ pointwise and $0 \le f_n \le f$ thus $f_n \to f$ in $L^p(\Omega,\mu)$. Note also that $f_n \le n$ so that $f_n$ is bounded.* Thus there exists and $N$ such that $\|f - f_N\|_p \le \varepsilon$. Up to replacing $f$ by $f_N$ we can thus assume that $f$ is bounded by some $N$.

Now let $(\Omega_n)$ be an increasing sequence of subsets of $\Omega$ of finite measure such that $\bigcup \Omega_n = \Omega$. Setting $f_n = \mathbf{1}_{\Omega_n} f_N$ we get $f_n \to f$ in $L^p(\Omega,\mu)$. We can thus assume that $f$ is supported in a set of finite measure $\tilde\Omega$ and write $\nu = \mu(\tilde\Omega)$.

Finally, $[0,N] \subset \bigcup_{j=0}^M I_j$ with $I_j = [(j/\nu)^{1/p}\varepsilon, \big((j+1)/\nu\big)^{1/p}\varepsilon[$. Note that the length of each interval is

$$\big((j+1)/\nu\big)^{1/p}\varepsilon - (j/\nu)^{1/p}\varepsilon \le \big((1+j)^{1/p} - j^{1/p}\big)\varepsilon/\nu^{1/p} \le \varepsilon/\nu^{1/p}$$

since $1 \le p < +\infty$ thus $0 < 1/p \le 1$. Also, let $a_j \in I_j \cap \mathbb{Q}$. Write $B_j = f^{-1}(I_j)$ and note that the $B_j$'s are a disjoint cover of $\tilde\Omega$. Further, for $x \in B_j$

$$-\varepsilon/\nu^{1/p} \le \big(j/\nu\big)^{1/p}\varepsilon - a_j \le f(x) - a_j \le \big((j+1)/\nu\big)^{1/p}\varepsilon - a_j \le \varepsilon/\nu^{1/p}.$$

In other words, for $x \in B_j$, $|f(x) - a_j|^p \le \varepsilon/\nu^{1/p}$. As the $B_j$'s cover the support of $f$, if we take an $x$ such that $f(x) \ne 0$, then there is a $j$ such that $x \in B_j$ and then

$$|f(x) - a_j|^p = |f(x)\mathbf{1}_{B_j}(x) - a_j\mathbf{1}_{B_j}(x)|^p \le \varepsilon^p/\nu\mathbf{1}_{B_j}(x).$$

*In otherwords, $L^\infty(\Omega,\mu) \cap L^p(\Omega,\mu)$ is dense in $L^p(\Omega,\mu)$.

Next, note that when $u$ and $v$ have disjoint support, then $|u+v|^p = |u|^p + |v|^p$. As the $B_j$'s are disjoint, we get that for $x$ in the support of $f$

$$\begin{aligned} |\sum_{j=1}^{N} f(x)\mathbf{1}_{B_j}(x) - \sum_{j=1}^{N} a_j\mathbf{1}_{B_j}(x)|^p &= \sum_{j=1}^{N} |f(x)\mathbf{1}_{B_j}(x) - a_j\mathbf{1}_{B_j}(x)|^p \leq \varepsilon^p/\nu \\ &\leq \sum_{j=1}^{N} \varepsilon^p/\nu\mathbf{1}_{B_j}(x) \\ &= \varepsilon^p/\nu\mathbf{1}_{\tilde{\Omega}}(x). \end{aligned}$$

Finally note that both sides of this inequality are 0 when $x$ is not in the support of $f$ so that it is valid as well for those $x$'s. Set $\tilde{f} = \sum_{j=1}^{N} a_j\mathbf{1}_{B_j}$ then, integrating over $\Omega$, gives

$$\int_\Omega |f(x) - \tilde{f}(x)|^p \,\mathrm{d}\mu(x) \leq \frac{\varepsilon^p}{\nu}\int_\Omega \mathbf{1}_{\tilde{\Omega}}(x)\,\mathrm{d}\mu(x) = \varepsilon^p$$

by definition of $\nu$.

Thus, after losing again an $\varepsilon$, we may replace $f$ by $\tilde{f}$, that is assume that $f$ is a (bounded) simple function with rational coefficients with support of finite measure

$$f = \sum_{j=1}^{M} a_j\mathbf{1}_{B_j}$$

The last step now amounts to approximate each of the sets $B_j$ by a set $A_j \in \tilde{\mathcal{B}}$ with $\mu(B_j \setminus A_j) < (\varepsilon/NM)^p$ and set

$$f_\varepsilon = \sum_{j=1}^{M} a_j\mathbf{1}_{A_j}$$

so that $f - f_\varepsilon = \sum_{j=1}^{M} a_j\mathbf{1}_{A_j\setminus B_j}$. Recall that $0 \leq a_j \leq N$ thus

$$\|f - f_\varepsilon\|_p \leq \sum_{j=1}^{M} a_j \left\|\mathbf{1}_{A_j\setminus B_j}\right\|_p \leq NM \max \mu(B_j \setminus A_j)^{1/p} \leq \varepsilon.$$

We have thus shown the following:

Theorem 5.1. *Let $1 \leq p < +\infty$ and $(\Omega, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space with a countable $\sigma$-algebra basis. Then $L^p(\Omega, \mu)$ is separable.*

*Moreover, simple functions (with rational coefficients) are dense in $L^p(\Omega, \mu)$.*

Remark 5.2. Note that $L^\infty$ is not separable in infinite dimensions, which is equivalent to the existance of a sequence $(\Omega_j)_{j\geq 1}$ of disjoints sets in $\Omega$ of positive measure.

For $J \subset \mathbb{N}$ we define $f_J(x) = \sum_{j\in I} \mathbf{1}_{\Omega_j}(x) = \begin{cases} 1 & \text{if there exists } j \in J \text{ such that } x \in \Omega_j \\ 0 & \text{otherwise} \end{cases}.$

Then if $J \neq K$, there exists $j \in K \setminus J$ or $j \in J \setminus K$. Up to exchanging $J$ and $K$, we can assume we are in the later case. Then for $x \in \Omega_j$, $f_J(x) = 1$ while $f_K(x) = 0$. Thus $\|f_J - f_K\|_\infty \geq 1$. Note also that $\mathcal{P}(\mathbb{N})$, the set of subsets of $\mathbb{N}$ is *not countable*.

Now consider the non-countable set of disjoint open balls $\{B(f_J, 1/2),\ J \subset \mathbb{N}\}$. If a set $S$ is dense in $L^\infty(\Omega)$, it contains at least one element of each of these balls and is therefore non countable.

**5.3. Continuity of translations in $L^p(\mathbb{R}^d)$.** For $a \in \mathbb{R}^d$, we define the *translation* operator $\tau_a$ acting on $f \in L^p(\mathbb{R}^d)$ by $\tau_a f(x) = f(x-a)$. Of course, for every $1 \le p \le +\infty$ this is a bounded linear $L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ with $\|\tau_a f\|_p = \|f\|_p$.

What we are interested in here is the continuity with respect to the translation parameter $a$:

PROPOSITION 5.3. *Let $1 \le p < \infty$ and $f \in L^p(\mathbb{R}^d)$ then the mapping $a \mapsto \tau_a f$ is continuous $\mathbb{R}^d \to L^p(\mathbb{R}^d)$. In other words*

$$\text{given } f \in L^p(\mathbb{R}^d), \text{ and } a_0 \in \mathbb{R}^d, \ \|\tau_a f - \tau_{a_0} f\|_p \to 0 \text{ when } a \to a_0.$$

PROOF. First $\|\tau_a f - \tau_{a_0} f\|_p = \|\tau_{a-a_0} f - f\|_p$ so that it is enough to consider $a_0 = 0$.
**Step 1.** *The property is true when $f = \mathbf{1}_A$ for $A$ a Borel set of finite measure.*

When $d = 1$ and $f = \mathbf{1}_{[\alpha,\beta]}$, $\tau_a f = \mathbf{1}_{[\alpha+a,\beta+a]}$, and then,

$$\tau_a f - f = \begin{cases} \mathbf{1}_{[\beta,\beta+a]} - \mathbf{1}_{[\alpha,\alpha+]a} & \text{if } a \ge 0 \\ -\mathbf{1}_{[\beta+a,\beta]} + \mathbf{1}_{[\alpha+a,\alpha]} & \text{if } a \le 0 \end{cases}.$$

Therefore $\|\tau_a f - f\|_p = (2a)^{1/p} \to 0$ when $a \to 0$.

Now if $Q \in \mathbb{R}^d$ a cube $Q = \prod_{j=1}^d [a_j, b_j]$ and $f = \mathbf{1}_Q$ the result follows directly. Further, if $U = \bigcup_{j=1}^N Q_j$ with the $Q_j$'s disjoint cubes and $f = \mathbf{1}_U$ then $f = \sum_{j=1}^N \mathbf{1}_{Q_j}$ thus $\|\tau_a f - f\|_p \le \sum_{j=1}^N \left\|\tau_a \mathbf{1}_{Q_j} - \mathbf{1}_{Q_j}\right\|_p \to 0$ when $a \to 0$.

If $U$ is a bounded open set then, for every $\varepsilon > 0$ there exists a family of disjoint cubes $Q_j \subset U$ such that† $|U \setminus \bigcup_{j=1}^N Q_j| \le \varepsilon$. Then, for $f = \mathbf{1}_U$ and $f_\varepsilon = \sum_{j=1}^N \mathbf{1}_{Q_j}$, we have

$$\|\tau_a f - \tau_a f_\varepsilon\|_p = \|f_\varepsilon - f\|_p = |U \setminus \bigcup_{j=1}^N Q_j|^{1/p}.$$

It follows that

$$\|\tau_a f - f\|_p \le \|\tau_a f - \tau_a f_\varepsilon\|_p + \|\tau_a f_\varepsilon - f_\varepsilon\|_p + \|f_\varepsilon - f\|_p \le 3\varepsilon^{1/p}$$

provided $a$ is small enough. The result is thus valid for such an $f$.

Finally, Let $E$ be a set of finite Lebesgue measure and $\varepsilon > 0$. There exists an open set $U$ such that $|E \Delta U| \le \varepsilon$. Set $f = \mathbf{1}_E$ and $f_\varepsilon = \mathbf{1}_U$, then $|f - f_\varepsilon| = \mathbf{1}_{E\Delta U}$, thus $\|f - f_\varepsilon\|_p = |E\Delta U|^{1/p} \le \varepsilon^{1/p}$ and one concludes as previsouly.
Step 2. *Conclusion.*

The first consequence is that, by linearity, if $f$ is a simple function, $f = \sum_{j=1}^N c_j \mathbf{1}_{E_j}$, then

$$\|\tau_a f - f\|_p \le \sum_{j=1}^N |c_j| \left\|\tau_a \mathbf{1}_{E_j} - \mathbf{1}_{E_j}\right\|_p \to 0$$

when $a \to 0$.

Finally, if $f \in L^p$ and $\varepsilon > 0$, there exists a simple function $f_\varepsilon$ such that $\|f - f_\varepsilon\|_p \le \varepsilon$. One then concludes as previously. □

REMARK 5.4. The result is false in $L^\infty$: take $f = \mathbf{1}_{[0,1]}$ and $a \ne 0$ then $\|\tau_a f - f\|_\infty = 1$. However, if $f \in \mathcal{C}_0(\mathbb{R}^d)$ then $\|\tau_a f - \tau_{a_0} f\|_\infty \to 0$ when $a \to a_0$.

Indeed, as previsouly, it is enough to consider $a_0 = 0$.

†We here adopt the notation $|E|$ for the Lebesgue measure of $E$

Now, let $a \in \mathbb{R}^d$ with $|a| \leq 1$ and $\varepsilon > 0$. Fix $R$ such that, for $|x| \geq R-1$, $|f(x)| \leq \varepsilon/2$. But then, if $|x| \geq 1$, $|x-a| \geq R-1$ thus $|f(x) - f(x-a)| \leq |f(x)| + |f(x-a)| \leq \varepsilon$.

On $B(0,R)$, $f$ is *uniformly* continuous, thus there exists $0 < \eta < 1$ such that, if $|a| \leq \varepsilon$, for every $x \in B(0,R)$, $|f(x) - f(x-a)| \leq \varepsilon$. It follows that $\sup_{x\in\mathbb{R}^d} |f(x) - f(x-a)| \leq \varepsilon$.

## 6. The projection Theorem

Projections play an essential role in Hilbert spaces. It turns out that a version of the projection theorem is still valid in $L^p$:

THEOREM 6.1. *Let $1 \leq p < +\infty$ and let $E$ be a closed vector space in $L^p(\Omega,\mu)$. For $f \in L^p(\Omega,\mu)$, let us write $d(f,E) = \inf_{g\in E} \|f-g\|_p$. Then there exists $g_0$ such that $d(f,E) = \|f-g_0\|_p$.*

REMARK 6.2. Not that, if $\|g\|_p > 2\|f\|_p$ then

$$\|f-g\|_p \geq \|g\|_p - \|f\|_p > \|f\|_p = \|f-0\|_p \geq d(f,E)$$

since $0 \in E$. Therefore $d(f,E) = \inf\{\|f-g\|_p : g \in E, \|g\|_p \leq 2\|f\|_p\}$.

If $E$ is finite dimensional, the $\{g \in E, \|g\|_p \leq 2\|f\|_p\}$ being bounded and closed, is compact. As $g \to \|f-g\|_p$ is continuous, the existence of $g_0$ follows.

In infinite dimensions, this argument is no longer valid.

PROOF WHEN $p \geq 2$. When $p = 2$ this follows from the parallelogram identity

$$\|u-v\|_2^2 + \|u+v\|_2^2 = 2\|u\|_2^2 + 2\|v\|_2^2.$$

Take $g_n \in E$ a sequence such that $\|f-g_n\|_2 \to d(f,E)$. Then the parallelogram identity applied to $u = \frac{f-g_m}{2}$, $v = \frac{f-g_n}{2}$ gives

$$\|g_n - g_m\|_2^2 = 4\left(\frac{1}{2}\|f-g_m\|_2^2 + \frac{1}{2}\|f-g_n\|_2^2 - \left\|\frac{g_n+g_m}{2} - f\right\|_2^2\right).$$

As $\frac{g_n+g_m}{2} \in E$, $\left\|\frac{g_n+g_m}{2} - f\right\|_2 \geq d(f,E)$ thus

$$\|g_n - g_m\|_2^2 \leq 2(\|f-g_m\|_2^2 - d(f,E)^2 + \|f-g_n\|_2^2 - d(f,E)^2)$$

from which one gets that $(g_n)$ is a Cauchy sequence. Thus $(g_n)$ is convergent and as $E$ is closed, the limit $g_0 \in E$. By continuity of the norm $\|f-g_n\|_2 \to \|f-g_0\|_2$ which is then the $g_0$ we were looking for.

When $p > 2$, the parallelogram identity is no longer valid. However, it is valid pointwise: if $f, g \in L^p(\Omega,\mu)$ and $x \in \Omega$ then

$$|f(x)-g(x)|^2 + |f(x)+g(x)|^2 = 2|f(x)|^2 + 2|g(x)|^2.$$

As $p > 2$, $r = p/2 > 1$. But, for $a, b > 0$

$$a^r + b^r \leq (a+b)^r \leq 2^{r-1}(a^r+b^r). \tag{6.5}$$

From this, we get

$$\begin{aligned}
|f(x)-g(x)|^p + |f(x)+g(x)|^p &= \big(|f(x)-g(x)|^2\big)^r + \big(|f(x)+g(x)|^2\big)^r \\
&\leq \big(|f(x)-g(x)|^2 + |f(x)+g(x)|^2\big)^r \\
&= 2^r\big(|f(x)|^2 + |g(x)|^2\big)^r \leq 2^{2r-1}\big(|f(x)|^{2r} + |g(x)|^{2r}\big) \\
&= 2^{p-1}\big(|f(x)|^p + |g(x)|^p\big).
\end{aligned}$$

Integrating with respect to $\mu$, we get
$$\|f-g\|_p^p+\|f+g\|_p^p\le 2^{p-1}\big(\|f\|_p^p+\|g|_p^p\big).$$
The remaining of the proof is exactly the same: take a sequence $g_n\in E$ such that $\|g_n-f\|_p\to d(f,E)$ and apply the inequality with $f$ replaced by $f-g_n$ and $g$ by $f-g_m$. We obtain
$$\begin{aligned}\|g_n-g_m\|_p^p &\le 2^{p-1}\big(\|f-g_n\|_p^p+\|f-g_m\|_p^p\big)-\|2f-g_n-g_m\|_p^p\\ &\le 2^{p-1}\big(\|f-g_n\|_p^p+\|f-g_m\|_p^p-2d(f,E)\big).\end{aligned}$$
We then deduce that $g_n$ is a Cauchy sequence, thus converges. As $E$ is closed, the limit is in $E$ and is the desired value. □

Proof of (6.5). Let us rewrite the inequality $a^r+b^r\le(a+b)^r$ in the form $1+(b/a)^r\le(1+b/a)^r$ that is, setting $t=b/a$, $1+t^r\le(1+t)^r$ for all $t>0$. For $t\ge0$ let $f(t)=(1+t)^r-(1+t^r)$. Clearly $f(0)=0$ and $f'(t)=r((1+t)^{r-1}-t^{r-1})\ge0$ since $r\ge1$ thus $s^{r-1}$ is increasing.

The other inequality uses convexity of $t\to t^r$:
$$(a+b)^r=2^r\left(\frac{a+b}{2}\right)^r\le 2^r\frac{a^r+b^r}{2}$$
which is the expected inequality. □

The proof for $p<2$ is more involved and requires the use of Hammer's inequality
$$|\|f+g\|_p+\|f-g\|_p|^p+|\|f+g\|_p-\|f-g\|_p|^p\le 2^p\big(\|f\|_p^p+\|g\|_p^p\big).$$
As we won't use the projection theorem in that case, we will not develop the proof here.

## 7. Duality

Thanks to Hölder's inequality, it is easy to construct continous linear functionals on $L^p(\Omega,\mu)$. Indeed,

Lemma 7.1. *Let $1\le p\le+\infty$ and let $p'$ be such that $\frac1p+\frac1{p'}=1$. Let $g\in L^{p'}(\Omega,\mu)$ and define*
$$\Phi_g(f)=\int_\Omega f(x)g(x)\,d\mu(x).$$
*Then $\Phi_g$ is a continous linear functional on $L^p(\Omega,\mu)$. Moreover*
$$\|\Phi_g\|:=\sup_{\|f\|_p\le1}\int_\Omega f(x)g(x)\,d\mu(x)=\|g\|_{p'}.$$

Proof. Hölder's inequality directly shows continuity with $\|\Phi_g\|\le\|g\|_{p'}$ while the equality follows from the equality case in Hölder's inequality. □

The key result of this section is the following converse of this lemma:

Theorem 7.2. *Let $1\le p<+\infty$ and let $p'$ be such that $\frac1p+\frac1{p'}=1$. Let $\Phi\in(L^p)'$ i.e. a bounded linear functional on $L^p(\Omega,\mu)$. Then there exists a unique $g\in L^{p'}(\Omega,\mu)$ such that $\Phi=\Phi_g$, that is*
$$\Phi(f)=\int_\Omega f(x)g(x)\,d\mu(x).$$
*for every $f\in L^p(\Omega,\mu)$.*

REMARK 7.3. It is important to notice that the result is false for $p=+\infty$. The dual of $L^\infty(\Omega,\mu)$ is much more difficult to describe and is out of scope of this course.

PROOF OF UNIQUENESS. The uniqueness is easy to prove: assume that $g_1,g_2\in L^{p'}$ are such that $\Phi_{g_1}=\Phi_{g_2}$ then, if $g=g_1-g_2$, for every $f\in L^p$, $\Phi_g(f)=0$.

If $p>1$, then $p'<+\infty$, take $f(x)=\begin{cases}|g(x)|^{p'-2}\overline{g(x)} & \text{if } g(x)\neq 0\\ 0 & \text{if } g(x)=0\end{cases}$. First $|f|^p=(|g|^{p'-1})^p=|g|^{p'}$ since $p=\frac{p'}{p'-1}$ when $\frac1p+\frac1{p'}=1$. Thus $f\in L^p$. Next,

$$0=\Phi_g(f)=\int_\Omega f(x)g(x)\,\mathrm{d}\mu(x)=\int_\Omega |g(x)|^{p'-2}\overline{g(x)}g(x)\,\mathrm{d}\mu(x)=\|g\|_{p'}^{p'}$$

thus $g=0$ as claimed.

If $p=1$, a slight modification is needed. Write $\Omega=\bigcup_{n\geq1}\Omega_n$ with $\mu(\Omega_n)<+\infty$ and $g(x)=e^{i\theta(x)}|g(x)|$. Then $f_n=e^{-i\theta}\mathbf{\Omega_n}\in L^1$ and

$$0=\Phi_g(f_n)=\int_\Omega f_n(x)g(x)\,\mathrm{d}\mu(x)=\int_{\Omega_n}|g(x)|\,\mathrm{d}\mu(x).$$

It follows that $g=0$ $\mu$-almost everywhere on $\Omega_n$ *i.e.* there is an $E_n\subset\Omega_n$ such that $g=0$ on $\Omega_n\setminus E_n$. Thus $g=0$ on $\bigcup_{n\geq1}\Omega_n\setminus\bigcup_{n\geq1}E_n=\Omega\setminus\bigcup_{n\geq1}E_n$. As $\bigcup_{n\geq1}E_n$ is a countable union of sets of mesure 0, it has measure 0 thus $g=0$ $\mu$-almost everywhere. □

Recall that $L^2(\Omega,\mu)$ is a Hilbert space so that the theorem follows from the more general theorem by Riesz. It turns out that the case $1\leq p<2$ can be deduced from it.

PROOF IN THE CASE $1\leq p<2$. First let $p'$ be the dual index, $\frac1p+\frac1{p'}=1$ and note that $p'>2$. Let $r$ be given by $\frac p2+\frac1s=1$ *i.e.* $s=\frac{2}{2-p}$ and $r=ps$. Note that $r$, $s$ have been chosen so that Hölder's inequality implies

(7.6)
$$\int_\Omega |f(x)|^p|g(x)|^p\,\mathrm{d}\mu(x)\leq\left(\int_\Omega|f(x)|^2\,\mathrm{d}\mu(x)\right)^{p/2}\left(\int_\Omega|g(x)|^{ps}\,\mathrm{d}\mu(x)\right)^{1/s}=\|f\|_2^p\|g\|_r^p.$$

Write $\Omega=\bigcup_{n\geq2}\Omega_n$ with $\mu(\Omega_n)<+\infty$ and the $\Omega_n$ being disjoint. Let us define $w$ through

$$w(x)=\sum_{n\geq1}\alpha_n\mathbf{1}_{\Omega_n}$$

where the $\alpha_n>0$ are chosen so that

(i) for every $n$, $\alpha_n>0$ and $\alpha_{n+1}\leq\alpha_n$,
(ii) $\|w\|_r^r=\sum_{n\geq1}\alpha_n^r\mu(\Omega_n)<+\infty$.

It follows from (7.6) that, for every $f\in L^2(\Omega,\mu)$, $fw\in L^p(\Omega,\mu)$ with $\|fw\|_p\leq\|w\|_r\|f\|_2$. In other words, the operator $T_w:L^2\to L^p$ defined by $T_wf=wf$ is bounded.

Now, let $\Phi\in(L^p)'$, that is let $\Phi$ be a bounded linear functional on $L^p(\Omega,\mu)$. It follows that $\Phi T_w$ is a bounded linear functional on $L^2(\Omega,\mu)$. According to Riesz's theorem, there exists $G\in L^2(\Omega,\mu)$ such that $\Phi T_w=\Phi_G$: *for every* $f\in L^2(\Omega,\mu)$,

$$\Phi T_wf=\Phi(fw)=\int_\Omega f(x)G(x)\,\mathrm{d}\mu(x).$$

Now consider the set $\mathcal{S}=\{\varphi\in L^p(\Omega,\mu)\,:\,\varphi/w\in L^2(\Omega,\mu)\}$. Note that $\mathcal{S}$ is dense in $L^p(\Omega,\mu)$. Indeed, if $f\in L^p(\Omega,\mu)$ and $\varepsilon>0$, there exists $N$ such that, writing $\Phi_N=\bigcup_{n\leq N}\Omega_n$ $f_N=f\mathbf{1}_{\Phi_N}\mathbf{1}_{|f|\leq N}$, then $\|f-f_N\|_p\leq\varepsilon$ (note that $f_N\to f$ a.e. and that

$|f_N| \leq f$ so that $f_N \to f$ in $L^p$). Further, for $x \in \Phi_N$, there is an $n \leq N$ such that $x \in \Omega_n$. Then $w(x) = \alpha_n \geq \alpha_N$ since the $\alpha_n$ have been chosen as a decreasing sequence. It follows that

$$\frac{|f_N(x)|}{w(x)} \leq \begin{cases} 0 & \text{if } x \notin \Phi_N \\ \frac{N}{\alpha_N} & \text{if } x \in \Phi_N \end{cases}.$$

Thus $f_N/w$ is bounded with support of finite measure and is thus in $L^2(\Omega, \mu)$ *i.e.* $f_N \in \mathcal{S}$.

Now, for $\varphi \in \mathcal{S}$, we can write $\varphi = fw$ with $f = \varphi/w \in L^2$. Therefore

$$\Phi(\varphi) = \Phi(fw) = \int_\Omega f(x)G(x)\,\mathrm{d}\mu(x) = \int_\Omega \varphi(x)\frac{G(x)}{w(x)}\,\mathrm{d}\mu(x) = \Phi_g(\varphi)$$

with $g := G/w$. If we are able to prove that $g \in L^{p'}(\Omega, \mu)$, then $\Phi_g$ is a continuous linear functional on $L^p$ as well. Therefore $\Phi = \Phi_g$ is an equality between two continuous functionals on $L^p$ on the dense set $\mathcal{S}$ of $L^p$. This equality is then true on all of $L^p$, which is what we wanted to prove.

It remains to prove that $g \in L^{p'}(\Omega, \mu)$. We need to distinguish two cases.

First consider the case $1 < p < 2$. Consider $\varphi_N = \overline{g}|g|^{p-2}\mathbf{1}_{|g|\leq N}\mathbf{1}_{\Phi_N}$ and observe that $|\varphi_N| = |g|^{p-1}\mathbf{1}_{|g|\leq N}\mathbf{1}_{\Phi_N}$. In particular $\varphi_N$ is bounded and has support of finite measure thus $\varphi_n \in L^p(\Omega, \mu)$ and on its support $w \geq \alpha_N$ so that $|\varphi_N/w| \leq |\varphi_N|/\alpha_N \in L^2(\Omega, \mu)$. In other words, $\varphi_N \in \mathcal{S}$. But then

$$\Phi(\varphi_N) = \Phi_g(\varphi_N) = \int_\Omega \varphi_N(x)g(x)\,\mathrm{d}\mu(x) = \int_\Omega |g(x)|^{p'}\mathbf{1}_{|g|\leq N}(x)\mathbf{1}_{\Phi_N}(x)\,\mathrm{d}\mu(x).$$

On the other hand, $\Phi$ is continuous on $L^p(\Omega, \mu)$ thus, for all $\varphi$, $|\Phi(\varphi)| \leq \|\Phi\|\|\varphi\|_p$, in particular

$$\begin{aligned} |\Phi(\varphi_N)| &\leq \|\Phi\|\|\varphi_N\|_p = \|\Phi\|\left(\int_\Omega |g|^{p(p-1)}(x)\mathbf{1}_{|g|\leq N}(x)\mathbf{1}_{\Phi_N}(x)\,\mathrm{d}\mu(x)\right)^{1/p} \\ &= \|\Phi\|\left(\int_\Omega |g|^{p'}(x)\mathbf{1}_{|g|\leq N}(x)\mathbf{1}_{\Phi_N}(x)\,\mathrm{d}\mu(x)\right)^{1/p}. \end{aligned}$$

Combining both identities shows that, for every $N$,

$$\left(\int_\Omega |g|^{p'}(x)\mathbf{1}_{|g|\leq N}(x)\mathbf{1}_{\Phi_N}(x)\,\mathrm{d}\mu(x)\right)^{1/p'} \leq C$$

Letting $N$ go to infinity and applying Beppo-Levi's Lemma, we get $\|g\|_{p'} \leq C$ so that $g \in L^{p'}(\Omega, \mu)$ as expected.

When $p = 1$ the argument needs to be modified. We write $g = e^{i\theta}|g|$ and consider $\varphi_N = e^{-i\theta}\mathbf{1}_{|g|>\|\Phi\|+1/N}\mathbf{1}_{\Phi_N}$. As previously, $\varphi_N \in \mathcal{S}$. But then

$$\begin{aligned} \Phi(\varphi_N) = \Phi_g(\varphi_N) = \int_\Omega \varphi_N(x)g(x)\,\mathrm{d}\mu(x) &= \int_\Omega |g(x)|\mathbf{1}_{|g|>\|\Phi\|+1/N}\mathbf{1}_{\Phi_N}\,\mathrm{d}\mu(x) \\ &\geq (\|\Phi\| + 1/N)|\{|g| > \|\Phi\| + 1/N\} \cap \Phi_N|. \end{aligned}$$

On the other hand

$$\begin{aligned} |\Phi(\varphi_N)| \leq \|\Phi\|\|\varphi_N\|_1 &= \|\Phi\|\int_\Omega \mathbf{1}_{|g|>\|\Phi\|+1/N}\mathbf{1}_{\Phi_N}\,\mathrm{d}\mu(x) \\ &= \|\Phi\||\{|g| > \|\Phi\| + 1/N\} \cap \Phi_N|. \end{aligned}$$

Combining both, we get that $|\{|g| > \|\Phi\| + 1/N\} \cap \Phi_N| = 0$. Finally, As $\{|g| > \|\Phi\|\} = \bigcup_{N\geq 1}\{|g| > \|\Phi\| + 1/N\} \cap \Phi_N$ we get that $|g| \leq \|\Phi\|$ almost everywhere. □

Proof using the projection theorem when $1 < p < \infty$. Let $\Phi$ be a continuous linear functional on $L^p(\Omega,\mu)$. We are looking for $g \in L^{p'}(\Omega,\mu)$ such that $\Phi = \Phi_g$. We can assume that $\Phi$ is not identically zero (otherwise take $g = 0$) so that there is an $f \in L^p(\Omega,\mu)$ with $L(f) \neq 0$. Up to replacing $f$ by $f/L(f)$ we can assume that $L(f) = 1$.

Let $E = \ker\Phi = \Phi^{-1}(0)$ and note that $E$ is a closed linear subspace of $L^p(\Omega,\mu)$. Therefore, there exists $g_0 \in E$ such that $\|f - g_0\|_p = d(f,E)$. Note that $L(f - g_0) = L(f) - L(g_0) = 1 - 0 = 0$ and that $\|f - g_0\|_p = \|(f - g_0) - 0\|_p = d(f,E)$. Up to replacing $f$ by $f - g_0$ we can assume that 0 is a projection of $f$ on $E$: $L(f) = 1$ and $\|f\|_p = d(f,E)$, that is, for all $g \in E$, $\|f\|_p \leq \|f - g\|_p$.

Now fix $w \in E$ and consider the function $\varphi$ defined on $(-1,1) \times \Omega$ by $\varphi(t,x) = |f(x) - tg(x)|^p$ and let $\Phi$ be defined on $\mathbb{R}$ by

$$\Phi(t) = \int_\Omega \varphi(t,x)\,\mathrm{d}x = \|f - tg\|_p^p.$$

First, observe that

– as $tg \in E$, $\Phi(t) = \|f - tg\|_p^p \geq \|f\|_p^p = \Phi(0)$. Thus $\Phi$ has a minimum at 0.

– $\varphi$ is continuous in $t$. Moreover,

$$\varphi(t,x) = \big(|f(x) - tg(x)|^2\big)^{p/2} = \big(|f(x)|^2 + t^2|g(x)|^2 + 2t\Re\overline{f(x)}g(x)\big)^{p/2}$$

thus

$$\begin{aligned}\frac{\partial\varphi}{\partial t} &= \frac{p}{2}\big(|f(x) - tg(x)|^2\big)^{p/2-1}\big(2t|g(x)|^2 + 2\Re\overline{f(x)}g(x)\big)\\ &= p|f(x) - tg(x)|^{p-2}\big(t|g(x)|^2 + \Re\overline{f(x)}g(x)\big)\end{aligned}$$

— for $|t| \leq 1$ and $x \in \Omega$,

$$\begin{aligned}|\varphi(t,x)| &= |f(x) - tg(x)|^p = 2^p\left|\frac{f(x) - tg(x)}{2}\right|^p \leq 2^p\left(\frac{|f(x)| + |g(x)|}{2}\right)^p\\ &\leq 2^{p-1}(|f(x)|^p + |g(x)|^p).\end{aligned}$$

Lebesgue's theorem on continuity of integrals then shows that $\Phi$ is continuous.

— for $|t| \leq 1$ and $x \in \Omega$,

$$\begin{aligned}\left|\frac{\partial\varphi}{\partial t}\right| &\leq p(|f(x)| + |g(x)|)^{p-2}\big(2|g(x)|^2 + |f(x)|^2\big)\\ &\leq p2^{p-3}(|f(x)|^{p-2} + |g(x)|^{p-2})\big(2|g(x)|^2 + |f(x)|^2\big)\\ &\leq p2^{p-2}(|f(x)|^p + |g(x)|^p + |f(x)|^{p-2}|g(x)|^2 + |g(x)|^{p-2}|f(x)|^2).\end{aligned}$$

As $f,g \in L^p$, $|f(x)|^p + |g(x)|^p$ is integrable. Further if $q = p/2$ and $q'$ is given by $\frac{1}{q} + \frac{1}{q'} = 1$ then $q' = \frac{q}{q-1} = \frac{p}{p-2}$ and Hölders inequality with these exponents gives

$$\int_\Omega |f(x)|^{p-2}|g(x)|^2\,\mathrm{d}\mu(x) \leq \left(\int_\Omega |f(x)|^p\,\mathrm{d}\mu(x)\right)^{(p-2)/p}\left(\int_\Omega |g(x)|^p\,\mathrm{d}\mu(x)\right)^{2/p}$$

thus $|f(x)|^{p-2}|g(x)|^2$ is also intergable. The same is true for $g(x)|^{p-2}|f(x)|^2$.

We can thus apply Lebegue's derivation theorem and see that $\Phi$ is differentiable on $(-1,1)$ and

$$\Phi'(t) = \int_\Omega \frac{\partial\varphi}{\partial t}(t,x)\,\mathrm{d}\mu(x).$$

In particular,

$$\Phi'(0) = p \int_\Omega |f(x)|^{p-2} \Re \overline{f(x)} g(x) \, \mathrm{d}\mu(x).$$

As $\Phi$ has a minimum at 0, we get $\Phi'(0) = 0$ that is, for every $g \in L^p$, with $L(g) = 0$,

$$\Re \int_\Omega |f(x)|^{p-2} \overline{f(x)} g(x) \, \mathrm{d}\mu(x) = 0.$$

Note that, if $g \in L^p$, with $L(g) = 0$ then $ig \in L^p$ and $L(ig) = 0$ so that

$$\Re i \int_\Omega |f(x)|^{p-2} \overline{f(x)} g(x) \, \mathrm{d}\mu(x) = 0.$$

Finally, define $\tilde{f}$ by $\tilde{f}(x) = |f(x)|^{p-2}\overline{f(x)}$ and note that $|\tilde{f}|^{p'} = |f|^{(p-1)p'} = |f|^p$ so that $\tilde{f} \in L^{p'}$ with $\left\|\tilde{f}\right\|_{p'} = \|f\|_p$. We have proved that *for every* $g \in L^p(\Omega, \mu)$ *with* $L(g) = 0$,

$$\int_\Omega \tilde{f}(x) g(x) \, \mathrm{d}\mu(x) = 0.$$

In otherwords, if $L(g) = 0$ then $\Phi_{\tilde{f}}(g) = 0$.

Now let $h \in L^p$ and consider $g = h - L(h)f \in L^p$. Note that $L(g) = L(h) - L(h)L(f) = 0$ since $L(f) = 1$ and that $Phi_{\tilde{f}}(f) = \|f\| - p^p$. Therefore $\Phi_{\tilde{f}}(g) = 0$. But

$$0 = \Phi_{\tilde{f}}(g) = \Phi_{\tilde{f}}\big(h - L(h)f\big) = \Phi_{\tilde{f}}(h) - L(h)\Phi_{\tilde{f}}(f) = \Phi_{\tilde{f}}(h) - L(h)\|f\|_p^p.$$

As $L(f) = 1$, $f \neq 0$ thus $\|f\|_p^p \neq 0$ and we conclude that

$$L(h) = \frac{1}{\|f\|_p^p} \Phi_{\tilde{f}}(h) = \Phi_{\tilde{f}/\|f\|_p^p}(h)$$

which is the expected result. $\square$

CHAPTER 3

# Convolution - regularization

### Multi-index notation

Before starting this section, we will introduce the multi-index notation:

A multi-index is a vector with integer coordinates: $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$. If $\beta = (\beta_1, \ldots, \beta_d) \in \mathbb{N}^d$, we will say that $\beta \leq \alpha$ if $\beta_j \leq \alpha_j$ for all $j \in \{1, \ldots, d\}$.

The length of a multi-index $\alpha$ is the sum of its coordinates: $|\alpha| = \alpha_1 + \cdots + \alpha_d$.

We will write $\alpha! = \alpha_1! \cdots \alpha_d!$, and the binomial coefficient for $\beta \leq \alpha$

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!} = \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_d}{\beta_d}.$$

For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, we write $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. For a function $f : \mathbb{R}^d \to \mathbb{C}$ we write

$$\partial^\alpha f = \frac{\partial^{\alpha_1}}{\partial_{x_1}^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial_{x_d}^{\alpha_d}} f.$$

With this notation, some classical one-variable formula are written in the same way for multi-variate functions:

– Leibnitz formula

$$\partial^\alpha (fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \partial^{\alpha-\beta} g$$

– Taylor formula

$$f(x_0 + h) = \sum_{|\alpha| \leq n} \partial^\alpha f(x_0) \frac{h^\alpha}{|\alpha|!} + o(h^N).$$

### 1. Definition and basic examples

Definition 1.1. Let $f, g$ be two functions on $\mathbb{R}^d$, we define the convolution of $f$ and $g$ as being the function on $\mathbb{R}^d$ given by

$$f * g(x) = \int_{\mathbb{R}^d} f(y) g(x-y) \, \mathrm{d}y. \tag{1.7}$$

Note that in the definition, we have said nothing about the existence of $f * g$. The aim of this chapter is precisely to give a meaning to $f * g$. However, there are a few basic examples for which this is easy:

Example 1.2. Let $f = \mathbf{1}_{[a,b]}$, $g = \mathbf{1}_{[c,d]}$.

First, the change of variable $t = x - y$ shows that $f * g = g * f$. On may thus assume that $b - a > d - c$, that is, the lenght of $[a, b]$ is bigger than the length of $[c, d]$.

It is obvious that, for $x$ fixed, $f(y)g(x - y) = \mathbf{1}_{I_x}(y)$ where $I_x$ is an intersection of two intervals and is thus an interval. It follows that $f * g(x) = |I_x|$ the length of this interval. Next $g(x - y) = 1$ is and only if $c \leq x - y \leq d$ that is $y \in [x - d, x - c]$ so that $I_x = [a, b] \cap [x - d, x - c]$. The length of this interval is clearly a piecewise affine function since $[a,b]$ is fixed and we "slide" a second interval $[-d,-c]$ at constant speed, *i.e.* the second interval is $[-d,-c]+x$.

It is enough to find the nodes and determine the length at those nodes.

There are 5 cases:

– the interval $[-d,-c]+x$ is entirely on the left of $[a,b]$ (up to the end point), that is $-c+x \leq a$ *i.e.* $x \leq a+c$. In this case $f*g(x) = |I_x| = 0$.

– the interval $[-d,-c]+x$ overlaps $[a,b]$ on the left side: $-d+x \leq a \leq -c+x$ *i.e.* $a+c \leq x \leq a+d$. In this case $I_x = [a,-c+x]$ and $f*g(x) = |I_x| = x-(a+c)$.

– the interval $[-d,-c]+x$ is entirely inside $[a,b]$: $a \leq -d+x \leq -c+x \leq b$ *i.e.* $a+d \leq x \leq b+c$. In this case $I_x = [-d,-c]+x$ and $f*g(x) = |I_x| = d-c$

– the interval $[-d,-c]+x$ overlaps $[a,b]$ on the right side: $-d+x \leq b \leq -c+x$ *i.e.* $b+c \leq x \leq b+d$. In this case $I_x = [-d+x,b]$ and $f*g(x) = |I_x| = b+d-x$.

– the interval $[-d,-c]+x$ is entirely on the left of $[a,b]$ (up to the end point), that is $b \leq -d+x$ *i.e.* $x \geq b+d$ and in this case again $f*g(x) = 0$.

We strongly advise the reader to draw the 5 cases and the graph of $f*g$. Once this is done, one can note for future use that $f*g$ is continuous and compactly supported with support $[a,b]+[c,d] = \{x+y, x\in[a,b], y\in[c,d]\} = [a+c,b+d]$.

Example 1.3. Assume that $f,g$ are tensors: $f(x_1,\ldots,x_d) = f_1(x_1)\cdots f_d(x_d)$ and $g(x_1,\ldots,x_d) = g_1(x_1)\cdots g_d(x_d)$. Then if $f_j*g_j$ are defined by (1.7), so if $f*g$ and

$$f*g(x_1,\ldots,x_d) = f_1*g_1(x_1)\cdots f_d*g_d(x_d).$$

An example of this are characteristic functions of cubes $Q = \prod_{j=1}^d I_j$ with $I_j$ intervals, then $\mathbf{1}_Q(x_1,\ldots,x_d) = \mathbf{1}_{I_1}(x_1)\cdots\mathbf{1}_{I_d}(x_d)$. This allows to compute $\mathbf{1}_Q*\mathbf{1}_{Q'}$ when $Q,Q'$ are cubes and shows that this function is continuous.

Lemma 1.4. *Let $f,g\in\mathcal{C}_c(\mathbb{R}^d)$, the space of compactly supported continuous functions. Then $f*g\in\mathcal{C}_c(\mathbb{R}^d)$ and $f*g=g*f$.*

*Morevoer, if $g\in\mathcal{C}_c^n(\mathbb{R}^d)$, then $f*g\in\mathcal{C}_c^n(\mathbb{R}^d)$ and for all $\alpha\in\mathbb{N}^d$, with $|\alpha|\leq n$, $\partial^\alpha(f*g) = f*(\partial^\alpha g) = (\partial^\alpha g)*f$.*

Note that $\partial^\alpha(f*g) = (\partial^\alpha g)*f$ implies that, if $g\in\mathcal{C}_c^n(\mathbb{R}^d)$ then $f*g$ is of class $\mathcal{C}^{n+m}$ and $\partial^{\alpha+\beta}(f*g) = (\partial^\alpha f)*(\partial^\beta g)$ as long as $|\alpha|\leq m$, $|\beta|\leq n$.

Proof. We will only prove the result in one variable, the proof for several variables is similar.

Consider $F(x,t) = f(t)g(x-t)$. Then

(1) $F$ is continuous in $t$ so that $f*g(x) = \int_{\mathbb{R}} F(x,t)\,\mathrm{d}t$ is well defined. Further, the change of varible $s = x-t$ shows that $f*g = g*f$.

(2) Write $I$ (resp. $J$) for an interval containing the support of $f$ (resp. of $g$). As $f,g$ are continuous with compact support, they are bounded, so we can take $C \geq \|f\|_\infty, \|g\|_\infty$. But then $|F(x,t)| \leq C^2\mathbf{1}_I(t)\mathbf{1}_J(x-t)$. It follows that

$$|f*g(x)| \leq C^2\int_{\mathbb{R}} \mathbf{1}_I(t)\mathbf{1}_J(x-t)\,\mathrm{d}t = C^2\mathbf{1}_I*\mathbf{1}_J(x).$$

The later one having compact support, $f*g$ has compact support. Further its support is included in $I+J = \{x+y, x\in I, y\in J\}$.

(3) Fix a bounded interval $K\subset\mathbb{R}$ and note that if $x\in K$ and $g(x-t)\neq 0$ then $t\in x-J\subset K-J = \{k-j, k\in K, j\in J\}$ (a bounded interval). It follows that $|F(x,t)| \leq C^2\mathbf{1}_I(t)\mathbf{1}_{K-J}(t)\in L^1(\mathbb{R})$. As $x\to F(x,t)$ is continuous for all

$t$, Lebesgue's continuity theorem shows that $f*g$ is continuous on $K$ and $K$ is arbitrary

The last part follows the same path noting that $\partial_x^\alpha F(x,t)=f(t)\partial^\alpha g(x-t)$ and then the same reasoning shows that this is bounded by an $L^1$ function independent of $x\in K$. It remains to apply Lebsgue's derivation theorem. □

## 2. Convolution between $L^p$ and its dual space

THEOREM 2.1. *Let $1\leq p\leq+\infty$ and $p'$ be such that $\dfrac{1}{p}+\dfrac{1}{p'}=1$. Let $f\in L^p(\mathbb{R}^d)$ and $g\in L^{p'}(\mathbb{R}^d)$ then*

$$f*g(x)=\int_{\mathbb{R}^d}f(t)g(x-t)\,dt \tag{2.8}$$

*is well defined for every $x\in\mathbb{R}^d$. The mapping $(f,g)\to f*g$ is bilinear and continuous $L^p\times L^{p'}\to L^\infty$ with $\|f*g\|_\infty\leq\|f\|_p\|g\|_{p'}$.*

*Moreover, if $1<p<+\infty$ $f*g\in\mathcal{C}_0(\mathbb{R}^d)$ so that $(f,g)\to f*g$ is a bounded bilinear mapping $L^p\times L^{p'}\to\mathcal{C}_0$.*

Recall that $\mathcal{C}_0(\mathbb{R}^d)$ is the space of continuous functions on $\mathbb{R}^d$ that go to 0 at infinity.

PROOF. First, if $g\in L^{p'}$ then $g_x:t\to g(x-t)$ is also in $L^{p'}$. Hölder's inequality then shows that $fg_x\in L^1$ thus $f*g$ is well defined through (2.8). Further, Hölder shows that $\|f*g\|_\infty\leq\|f\|_p\|g\|_{p'}$. As $(f;g)\to f*g$ is clearly bilinear, it follows that $(f,g)\to f*g$ is a bounded bilinear mapping $L^p\times L^{p'}\to L^\infty$.

The key observation is that $\mathcal{C}_0$ is a closed subspace of $L^\infty$. Indeed, if $(f_k)$ is a sequence of elements of $\mathcal{C}_0$ that converges to some $f$ in the $L^\infty$-norm, that is uniformly then

– the limit $f$ is continuous (uniform limits of continuous functions are continuous)

– for $\varepsilon>0$ there exists $n$ such that $\|f-f_n\|_\infty\leq\varepsilon$. But then, there exists $K$ such that, if $\|x\|\geq K$, $|f_n(x)|\leq\varepsilon$. Finally, for those $x$'s, $|f(x)|\leq|f_n(x)|+\|f-f_n\|_\infty\leq2\varepsilon$, so $f(x)\to0$ when $\|x\|\to+\infty$.

In conclusion $f\in\mathcal{C}_0$ and $\mathcal{C}_0$ is closed in $L^\infty$.

Now, Example 1.3 shows that, if $f,g$ are characteristic functions of cubes, $f*g$ is continuous compactly supported. By bilinearity, if $f,g$ are step functions, that is, finite linear combinations of characteristic functions of cubes, then $f*g\in\mathcal{C}_c(\mathbb{R}^d)\subset\mathcal{C}_0(\mathbb{R}^d)$.

Finally, let $f\in L^p(\mathbb{R}^d)$ and $g\in L^{p'}(\mathbb{R}^d)$. As $p\neq+\infty$, there exists a sequence $f_k$ of step functions such that $f_k\to f$ in $L^p(\mathbb{R}^d)$. As $p\neq1$, $p'\neq+\infty$, so there exists a sequence $g_k$ of step functions such that $g_k\to f$ in $L^p(\mathbb{R}^d)$.

But then

$$\begin{aligned}\|f*g-f_k*g_k\|_\infty &= \|(f-f_k)*g+f_k(g-g_k)\|_\infty\leq\|(f-f_k)*g\|_\infty+\|f_k(g-g_k)\|_\infty\\ &\leq \|f-f_k\|_p\|g\|_{p'}+\|f_k\|_p\|g-g_k\|_{p'}\to0\end{aligned}$$

since $\|f-f_k\|_p,\|g-g_k\|_{p'}\to0$ and $\|f_k\|_p$ is bounded since $f_k$ is convergent. □

## 3. Convolution of $L^1$ with itself

We want to make sense of

$$f*g(x)=\int_{\mathbb{R}^d}f(y)g(x-y)\,\mathrm{d}y. \tag{3.9}$$

This is possible as a Lebesgue integral when $\int_{\mathbb{R}^d} |f(y)g(x-y)|\,\mathrm{d}y$ is finite. But note that, integrating this quantity in the $x$ variable, we obtain, with Fubini

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(y)g(x-y)|\,\mathrm{d}y\,\mathrm{d}x &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(y)g(x-y)|\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\mathbb{R}^d} |f(y)| \left(\int_{\mathbb{R}^d} |g(x-y)|\,\mathrm{d}x\right)\mathrm{d}y\\ &= \int_{\mathbb{R}^d} |f(y)| \left(\int_{\mathbb{R}^d} |g(t)|\,\mathrm{d}t\right)\mathrm{d}y = \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

It follows that, if $f,g\in L^1(\mathbb{R}^d)$ then

$$\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |f(y)g(x-y)|\,\mathrm{d}y\right)\mathrm{d}x < +\infty$$

but then, for almost every $x$, $\int_{\mathbb{R}^d} |f(y)g(x-y)|\,\mathrm{d}y$ is finite. It follows that (3.9) is well defined for almost every $x$. Moreover, the resulting function is in $L^1(\mathbb{R}^d)$. Let us summarize this:

Proposition 3.1. *Let $f,g\in L^1(\mathbb{R}^d)$ then*

$$f*g(x) = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy$$

*is well defined for almost every $x\in\mathbb{R}^d$. Moreover, the mapping $(f,g)\to f*g$ is a bounded bilinear mapping $L^1(\mathbb{R}^d)\times L^1(\mathbb{R}^d)\to L^1(\mathbb{R}^d)$ and*

$$\|f*g\|_1 \le \||f|*|g|\|_1 = \|f\|_1\|g\|_1.$$

## 4. Extension principle

In this course, we will use the following general principle:

— $X$ and $Y$ are Banach spaces and $\mathcal{D}$ is a dense (vectorial) subspace of $X$;

— $T$ is a linear mapping $\mathcal{D}\to Y$;

— $T$ is bounded on $\mathcal{D}$, that is, there exists $C\ge 0$ such that, for all $x\in\mathcal{D}$, $\|Tx\|_Y \le C\|x\|_X$.

Then $T$ extends into a bounded linear mapping $\tilde{T}: X\to Y$ with same norm: for all $x\in\mathcal{D}$, $\tilde{T}x = Tx$ and for all $x\in X$, $\left\|\tilde{T}x\right\|_Y \le C\|x\|_X$.

Of course, we then write $\tilde{T} = T$.

Proof. Let us first extend $T$ and then show it is linear bounded:

Let $x\in X$. From the density of $\mathcal{D}$ in $X$, there exists a sequence $(x_n)_n\subset\mathcal{D}$ that converges to $x$ in $X$. In particular, it is a Cauchy sequence. Let us show that $(Tx_n)_n$ is also a Cauchy sequence. Indeed, let $\varepsilon>0$, there exists $N\ge 0$ such that, if $p,q\ge N$, then $\|x_p-x_q\|_X\le\varepsilon$. But then, as $x_p,x_q\in\mathcal{D}$ and $T$ is linear on $\mathcal{D}$,

$$\|Tx_p-Tx_q\|_Y = \|T(x_p-x_q)\|_Y \le C\|x_p-x_q\|_X \le C\epsilon$$

since $T$ is bounded on $\mathcal{D}$. Now, as $(Tx_n)_n$ is Cauchy in $Y$, a Banach space, $(Tx_n)_n$ has a limit that we denote by $a$.

We would of course like to call $a = Tx$. To do so, we need to show that, if $(y_n)_n$ is an other sequence of elements of $\mathcal{D}$ that converges to $x$ in $X$, then $Ty_n$ also converges to $a$. But, as $x_n, y_n \in \mathcal{D}$ and $T$ is linear on $\mathcal{D}$,

$$\|Tx_n - Ty_n\|_Y = \|T(x_n - y_n)\|_Y \leq C\|x_n - y_n\|_X \to C\|x - x\| = 0$$

since the norm is a continuous mapping. We thus write $a = \tilde{T}x$.

Further, if $x \in \mathcal{D}$ the sequence $x_n = x$ converges to $x$ so that $Tx = Tx_n \to \tilde{T}x$ and $\tilde{T}$ is an extension of $T$ from $\mathcal{D}$ to $X$. We will thus denote $\tilde{T} = T$.

Let us now show that $T$ is linear: let $x, y \in X$ and $\lambda, \mu \in \mathbb{K}$. By density, there exist sequences $(x_n), (y_n)$ in $\mathcal{D}$ that converge respectively to $x$ and $y$. But then $\lambda x_n + \mu y_n \to \lambda x + \mu y$ so $T(\lambda x_n + \mu y_n) \to T(\lambda x + \mu y)$. On the other hand, as $T$ is linear on $\mathcal{D}$, $T(\lambda x_n + \mu y_n) = \lambda T x_n + \mu T y_n \to \lambda T x + \mu T y$, so

$$T(\lambda x + \mu y) = \lambda T x + \mu T y.$$

Finally, if $x \in X$ and $(x_n)_n \subset \mathcal{D}$ converges to $x$, then $Tx_n \to Tx$ in $Y$ and $\|Tx_n\|_Y \leq C\|x_n\|_X$. As norms are continuous, $\|Tx\|_Y \leq C\|x\|_X$. So $T$ is a bounded linear mapping □

Let us illustrate this:

THEOREM 4.1. *Let $f \in L^1(\mathbb{R}^d)$ and $1 \leq p \leq +\infty$. Then the mapping $T_f : g \to f * g$ extends from $\mathcal{C}_c(\mathbb{R}^d) \to L^\infty$ to a mapping $L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$.*
*Moreover, this mapping commutes with the translations $\tau_a$.*

Recall that $\tau_a g(x) = g(x - a)$.

PROOF. Note that we have already seen that $f * g$ is well defined when $f \in L^1$ and $g \in L^\infty$. What we have to prove is that there is a $C > 0$ such that, for all $g \in \mathcal{C}_c(\mathbb{R}^d)$, $\|f * g\|_{L^p(\mathbb{R}^d)} \leq C\|g\|_{L^p(\mathbb{R}^d)}$.
But this follows from Minkowski's inequality :

$$\left\|\int_{\mathbb{R}^d} f(t)g(\cdot - t)\,\mathrm{d}t\right\|_r \leq \int_{\mathbb{R}^d} |f(t)|\|g(\cdot - t)\|_r\,\mathrm{d}t = \|f\|_1\|g\|_r.$$

Finally, when $p \neq +\infty$, $g \in \mathcal{C}_c(\mathbb{R}^d)$

$$\begin{aligned} T_f\tau_a g(x) = f * (\tau_a g)(x) &= \int_{\mathbb{R}^d} f(t)g(x - t - a)\,\mathrm{d}t \\ &= \int_{\mathbb{R}^d} f(t)g\big((x - a) - t\big)\,\mathrm{d}t = f * g(x - a) = \tau_a T_f g(x). \end{aligned}$$

Thus $T_f\tau_a = \tau_a T_f$ holds on the dense subspace $\mathcal{C}_c(\mathbb{R}^d)$ of $L^p(\mathbb{R}^d)$ and $T_f, \tau_a$ are continuous linear mappings on $L^p$ so the conclusion follows.
When $p = +\infty$, we can directly take $g \in L^\infty$ in the above computation. □

The extension principle works exactly the same way for bilinear mappings:
In this course, we will use the following general principle :
— $X_1, X_2$ and $Y$ are Banach spaces and $\mathcal{D}_1$ (resp. $\mathcal{D}_2$) is a dense (vectorial) subspace of $X_1$ (resp. $X_2$);
— $T$ is a bilinear mapping $\mathcal{D}_1 \times \mathcal{D}_2 \to Y$;
— $T$ is bounded on $\mathcal{D}_1 \times \mathcal{D}_2$, that is, there exists $C \geq 0$ such that, for all $x \in \mathcal{D}$, $\|T(x_1, x_2)\|_Y \leq C\|x_1\|_{X_1}\|x_2\|_{X_2}$.

Then $T$ extends into a bounded bilinear mapping $\tilde{T} : X_1 \times X_2 \to Y$ with same norm: for all $(x_1, x_2) \in \mathcal{D}_1 \times \mathcal{D}_2$, $\tilde{T}(x_1, x_2) = T(x_1, x_2)$ and for all $(x_1, x_2) \in X_1 \times X_2$, $\left\|\tilde{T}(x_1, x_2)\right\|_Y \leq C\|x_1\|_{X_1}\|x_2\|_{X_2}$.

Of course, we then write $\tilde{T} = T$.

## 5. Young's inequality

We would now like to extend the convolution to a bilinear mapping from $\mathcal{C}_c(\mathbb{R}^d) \times \mathcal{C}_c(\mathbb{R}^d) \to \mathcal{C}_c(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d) \times L^q(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$. For this to be possible, one needs to have a constant $C > 0$ such that the inequality

$$\|f * g\|_{L^r(\mathbb{R}^d)} \leq C\|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^q(\mathbb{R}^d)}. \tag{5.10}$$

To start, we will use a simple but common trick to check for which $p, q, r$ this is possible:

Fix $f, g \in \mathcal{C}_c(\mathbb{R}^d) \setminus \{0\}$ nand $f, g \geq 0$ so that $f * g \in \mathcal{C}_c(\mathbb{R}^d) \setminus \{0\}$ as well. Take a parameter $\lambda > 0$ and define $f_\lambda(x) = f(\lambda x)$, $g_\lambda(x) = g(\lambda x)$ then, changing variable $s = \lambda x$

$$f_\lambda * g_\lambda(x) = \int_{\mathbb{R}^d} f(\lambda t) g\big(\lambda(x-t)\big)\,\mathrm{d}t = \lambda^{-d}\int_{\mathbb{R}^d} f(s)g(\lambda x - s)\,\mathrm{d}s = \lambda^{-d} f * g(\lambda x).$$

On the other hand

$$\|f_\lambda\|_{L^p(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(\lambda t)|^p\,\mathrm{d}t\right)^{1/p} = \left(\lambda^{-d}\int_{\mathbb{R}^d} |f(s)|^p\,\mathrm{d}s\right)^{1/p} = \lambda^{-d/p}\|f\|_{L^p(\mathbb{R}^d)}.$$

The same way, we have

$$\|g_\lambda\|_{L^q(\mathbb{R}^d)} = \lambda^{-d/q}\|g\|_{L^q(\mathbb{R}^d)} \quad \text{and} \quad \|f_\lambda * g_\lambda\|_{L^r(\mathbb{R}^d)} = \lambda^{-d(1+1/r)}\|f * g\|_{L^r(\mathbb{R}^d)}.$$

Thus, if we replace $f, g$ by $f_\lambda, g_\lambda$ in (5.10), then

$$0 < \frac{\|f * g\|_{L^r(\mathbb{R}^d)}}{C\|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^q(\mathbb{R}^d)}} \leq \lambda^{d\left(1+\frac{1}{r}-\frac{1}{p}-\frac{1}{q}\right)}.$$

Letting $\lambda \to 0$, this implies that the power of $\lambda$ be $\leq 0$ while letting $\lambda \to +\infty$, this implies that the power of $\lambda$ be $\geq 0$. We have thus shown that (5.10) implies $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$. In other words, the conditions on $p, q, r$ in the following theorem are necessary.

Theorem 5.1 (Young's Inequality).
*Let $1 \leq p, q, r \leq +\infty$ be three real numbers such that $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$. Then, for all $f, g \in \mathcal{C}_c(\mathbb{R}^d)$,*

$$\|f * g\|_r \leq \|f\|_p\|g\|_q. \tag{5.11}$$

*It follows that the mapping $(f, g) \to f * g$ extends from $\mathcal{C}_c(\mathbb{R}^d) \times \mathcal{C}_c(\mathbb{R}^d) \to \mathcal{C}_c(\mathbb{R}^d)$ into a bounded bilinear mapping $L^p(\mathbb{R}^d) \times L^q(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$.*

*Further $f * g = g * f$.*

Proof. We only have to prove (5.11).

Note that several particular cases have already be proven: when $r = +\infty$, then $\frac{1}{p} + \frac{1}{q} = 1$ and this is (part of) Theorem 2.1.

When $r = 1$ then $\frac{1}{p} + \frac{1}{q} = 2$. As $p, q \geq 1$ this implies $p = q = 1$ and Young's inequality is Proposition 3.1. More generally, the case $p = 1$ was treated in Theorem 2.1 and, by symmetry $f * g = g * f$, so is the case $q = 1$.

Note finally that if $p=+\infty$, then as $1\leq q,r\leq+\infty$ then $1+1/r=1/q$ implies $q=1$ and $r=+\infty$ which is already covered. The same holds when $q=+\infty$. We can then assume that $1<p,q,r<+\infty$. Note that $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$ implies $r>p,q$ so that $0<p/r,q/r,1-p/r,1-q/r<1$.

We will use the following fact which comes from the duality of $L^r$-$L^{r'}$ when $\frac{1}{r}+\frac{1}{r'}=1$ (actually from Hölder's Inequality): *if* $\varphi\in L^r$ *then*

$$\|\varphi\|_r=\sup\{\int_{\mathbb{R}^d}\varphi(x)\psi(x)\,dx\ :\ \psi\in L^{r'},\|\psi\|_{r'}=1\}.$$

But now, if $f,g\in\mathcal{C}_c(\mathbb{R}^d)$, then $f*g\in\mathcal{C}_c(\mathbb{R}^d)\subset L^r(\mathbb{R}^d)$. Let $h\in L^{r'}$ with $r'=\frac{r}{r-1}$. We want to bound

$$I(f,g,h)=\int_{\mathbb{R}^d}f*g(x)h(x)\,\mathrm{d}x.$$

Obvously

$$|I(f,g,h)|\leq\int_{\mathbb{R}^d}|f*g(x)||h(x)|\,\mathrm{d}x\leq\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f|(t)|g|(x-t)|h|(x)\,\mathrm{d}x\,\mathrm{d}t=I(|f|,|g|,|h|)$$

with Fubini. We may thus replace $f,g,h$ with $|f|,|g|,|h|$, that is, we can now assume that $f,g,h\geq 0$. We have to prove that $I(f,g,h)\leq\|f\|_p\|g\|_q\|h\|_{r'}$.

Note that, as $f,g,h\geq0$, we may apply Fubini and get

$$I(f,g,h)=\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}f(t)g(x-t)h(x)\,\mathrm{d}x\,\mathrm{d}t.$$

To bound this quantity we will first isolate $h$ and apply Hölder's $L^r-L^{r'}$ inequality with $\dfrac{1}{r}+\dfrac{1}{r'}=1$ *i.e.* $r'=\frac{r}{r-1}$. We write $f(t)g(x-t)h(x)=F_1(x,t)F_2(x,t)$ with

$$F_1(x,t)=f(t)^{p/r}g(x-t)^{q/r}\quad\text{and}\quad F_2(x,t)=f(t)^{1-p/r}g(x-t)^{1-q/r}h(x)$$

so that

$$I(f,g,h)\leq\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}F_1(x,t)^r\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{r}}\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}F_2(x,t)^{r'}\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{r'}}. \tag{5.12}$$

Note that $F_1(x,t)^r,F_2(x,t)^{r'}\geq0$ so that we will be able to change the order of integration.

The first of these two integrals is rather simple to bound: using Fubini, we first integrate with respect to $x$,

$$\begin{aligned}\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}F_1(x,t)^r\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{r}} &= \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}f(t)^pg(x-t)^q\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{r}}\\ &= \left(\int_{\mathbb{R}^d}g(x)^q\,\mathrm{d}x\right)^{\frac{1}{q}\frac{q}{r}}\left(\int_{\mathbb{R}^d}f(t)^p\,\mathrm{d}t\right)^{\frac{1}{p}\frac{p}{r}}\\ &= \|f\|_p^{\frac{p}{r}}\|g\|_q^{\frac{q}{r}}. \end{aligned} \tag{5.13}$$

The second term is more involved. First

$$\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} F_2(x,t)^{r'}\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{r'}} = \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(t)^{(1-p/r)r'} g(x-t)^{(1-q/r)r'} h(x)^{r'}\,\mathrm{d}t\,\mathrm{d}x\right)^{\frac{1}{r'}}$$

$$\le \left(\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d} f(t)^{(1-p/r)r'} g(x-t)^{(1-q/r)r'}\,\mathrm{d}t\right)^{\frac{1}{r'}} \left(\int_{\mathbb{R}^d} h(x)^{r'}\,\mathrm{d}x\right)^{\frac{1}{r'}}$$

$$= \left\| f^{(1-p/r)r'} * g^{(1-q/r)r'} \right\|_{\infty}^{1/r'} \|h\|_{r'}. \tag{5.14}$$

We next introduce a parameter $s$ to be determined soon and $s'$ its dual index $\frac{1}{s}+\frac{1}{s'}=1$. Then from Theorem 2.1 we know that

$$\left\| f^{(1-p/r)r'} * g^{(1-q/r)r'} \right\|_{\infty} \le \left\| f^{(1-p/r)r'} \right\|_{s} \left\| g^{(1-q/r)r'} \right\|_{s'}. \tag{5.15}$$

As we want an estimate with $\|f\|_p$ this leads to the choice $s(1-p/r)r' = p$. As $r' = \dfrac{r}{r-1}$ we thus have $(1-p/r)r' = \frac{r-p}{r-1}$ so that

$$s = \frac{r-1}{r-p}p.$$

Note that $r > p > 1$ so $p < s < +\infty$. The dual index is then

$$s' = \frac{s}{s-1} = \frac{(r-1)p}{r(p-1)} = \frac{p'}{r'}$$

thus

$$(1-q/r)r's' = (1-q/r)p' = \left(1-\frac{q}{r}\right)p'.$$

But, multiplying $1+\dfrac{1}{r} = \dfrac{1}{p}+\dfrac{1}{q}$ by $q$ and rewriting it gives $1-\dfrac{q}{r} = q\left(1-\dfrac{1}{p}\right) = \dfrac{q}{p'}$. Finally

$$(1-q/r)r's' = q.$$

The choice of $s$ then implies that

$$\left\| f^{(1-p/r)r'} \right\|_{s} = \left(\int_{\mathbb{R}^d} f(x)^{(1-p/r)r's}\,\mathrm{d}x\right)^{\frac{1}{s}} = \left(\int_{\mathbb{R}^d} f(x)^{p}\,\mathrm{d}x\right)^{\frac{1}{s}} = \|f\|_p^{p/s}$$

while

$$\left\| g^{(1-q/r)r'} \right\|_{s'} = \left(\int_{\mathbb{R}^d} g(x)^{(1-q/r)r's'}\,\mathrm{d}x\right)^{\frac{1}{s'}} = \left(\int_{\mathbb{R}^d} g(x)^{q}\,\mathrm{d}x\right)^{\frac{1}{s'}} = \|g\|_p^{q/s'}.$$

Injecting this into (5.15), we get

$$\left\| f^{(1-p/r)r'} * g^{(1-q/r)r'} \right\|_{\infty} \le \|f\|_p^{p/s}\|g\|_p^{q/s'}.$$

From this, (5.14) reduces to

$$\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} F_2(x,t)^{r'}\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{r'}} \le \|f\|_p^{p/r's}\|g\|_p^{q/r's'}\|h\|_{r'}.$$

Finally, with (5.13), we get that (5.12) reduces to

$$I(f,g,h) \le \|f\|_p^{\frac{p}{r}+\frac{p}{r's}}\|g\|_p^{\frac{q}{r}+\frac{q}{r's'}}\|h\|_{r'}.$$

It remains to notice that

$$\frac{1}{r}+\frac{1}{r'}\frac{1}{s}=\frac{1}{r}+\frac{r-1}{r}\frac{r-p}{(r-1)p}=\frac{p+r-p}{rp}=\frac{1}{p}$$

and that $\frac{1}{r}+\frac{1}{r'}\frac{1}{s'}=\frac{1}{q}$. In conclusion we have

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(t)g(x-t)h(x)\,\mathrm{d}x\,\mathrm{d}t \le \|f\|_p\|g\|_q\|h\|_{r'}$$

hor all $h \in L^{r'}$. It follows that, for all $f,g\in\mathcal{C}_c(\mathbb{R}^d)$,

$$\|f*g\|_r \le \|f\|_p\|g\|_q.$$

The extension principle then shows that $f*g$ can be defined on $L^p\times L^q$. □

REMARK 5.2. If $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}>1$ then $f*g$ is a priori *not* defined by $\int_{\mathbb{R}^d} f(t)g(x-t)\,\mathrm{d}t$.

One needs to approximate $f$ and/or $g$ by a sequence of functions that converges to $f$ and $g$ in $L^p$ and $L^q$ respectively and for which the above definition makes sense.

To do so, write $f_k = f\mathbf{1}_{\|f\|\le k}$ so that $f_k\to f$ in $L^p$. Further, $f_k\in L^s$ for every $s\ge p$. But $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$ can be rewritten as $\frac{1}{p}-\frac{1}{r}=1-\frac{1}{q}=\frac{1}{q'}$ so that $q'>p$. In particular, $f_k\in L^{q'}$. But then $f_k*g(x)=\int_{\mathbb{R}^d} f_k(t)g(x-t)\,\mathrm{d}t$. As $f_k*g\to f*g$ in $L^r$ we conclude that

$$f*g(x)=\lim_{k\to+\infty}\int_{\mathbb{R}^d} f(t)\mathbf{1}_{|f|\le k}(t)g(x-t)\,\mathrm{d}t.$$

## 6. Regularization

### 6.1. Spaces of smooth functions: $\mathcal{C}_c^\infty(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$.

Spaces of smooth functions will play a key role in the sequel. The first space we consider is the following:

$$\mathcal{C}_c^\infty(\mathbb{R}^d)=\{f\in\mathcal{C}^\infty(\mathbb{R}^d)\ :\exists R>0 \text{ s.t. } f(x)=0 \text{ if } \|x\|\ge R\}$$

the space of smooth functions with compact support.

One may wonder if such functions actually exist so let us start by giving an example:

EXAMPLE 6.1. Let $g$ be defined on $\mathbb{R}$ by $g(x)=\begin{cases}0 & \text{if } x\le 0\\ e^{-1/x} & \text{if } x>0\end{cases}$. Then $g$ is clearly $\mathcal{C}^\infty$ on $\mathbb{R}\setminus\{0\}$. Moreover, for every $k$, there exists a polynomial $P_k$ such that $g^{(k)}(x)=\frac{P_k(x)}{x^{2k}}g(x)$ when $x\ne 0$.

Indeed, the formula is clearly true for $k=0$. For $k=1$, $g'(x)=0$ when $x\le 0$ while $g'(x)=-\frac{1}{x^2}e^{-1/x}$ so that the formula is also true for $k=1$. Assuming $g^{(k)}$ is of that form up to some rank $k\ge 1$ we get

$$g^{(k+1)}(x)=\frac{P_k'(x)}{x^{2k}}g(x)-\frac{2kP_k(x)}{x^{2k+1}}g(x)\frac{P_k(x)}{x^{2k}}g'(x)=\frac{x^2P_k'(x)-(2kx+1)P_k(x)}{x^{2k+2}}g(x)$$

and if $P_k$ is a polynomial, so is $P_{k+1}(x):=x^2P_k'(x)-(2kx+1)P_k(x)$.

Alternatively, one may also show that $g^{(k)}(x)=Q_k(1/x)g(x)$ with $Q_k$ a polynomial.

Next, it is clear that $g$ is continuous at 0. Assuming $g$ is of class $\mathcal{C}^{k-1}$ on $\mathbb{R}$, as $g^{(k)}(x)=\frac{P_k(x)}{x^{2k}}e^{-1/x}$ we get that $g^{(k)}(x)\to 0$ when $x\to 0^+$ and as $g^{(k)}(x)=0$ when

$x < 0$ we also get that $g^{(k)}(x) \to 0$ when $x \to 0^-$. It follows that $g^{(k)}$ extends by continuity at 0 so that $g^{(k-1)}$ is of class $\mathcal{C}^1$, thus $g$ is of class $\mathcal{C}^k$.

Finally define $f$ through $f(x) = g(1 - \|x-a\|^2/\eta^2)$ and note that $g$ is clearly $\mathcal{C}^\infty$ (taking the euclidean norm) and that $f(x) = 0$ when $1 - \|x-a\|^2/\eta^2 \leq 0$ that is, when $|x-a\| \geq \eta$. Thus $f$ is $\mathcal{C}^\infty$ supported in the ball $B(a,\eta)$.

EXAMPLE 6.2. We still consider $g$ defined on $\mathbb{R}$ by $g(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-1/x} & \text{if } x > 0 \end{cases}$. Next, we define

$$h(x) = \frac{g(x)}{g(x)+g(1-x)} = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{e^{-1/x}}{e^{-1/x}+e^{-1/(1-x)}} & \text{for } 0 < x < 1 \\ 1 & \text{for } x \geq 1 \end{cases}.$$

As $g(x)+g(1-x) \neq 0$ for all $x$, clearly $h$ is of class $\mathcal{C}^\infty$ on $\mathbb{R}$.

Next, we define $b(x) = h(2+x)h(2-x)$ which is clearly $\mathcal{C}^\infty$. Further, for $|x| \geq 2$, one of $2+x, 2-x$ is $\leq 0$ so $b(x) = 0$. For $|x| \leq 1$, both $2+x, 2-x$ are $\geq 1$ so that $h(2+x) = h(2-x) = 1$ and $b(x) = 1$. Finally as $0 \leq g \leq 1$, $0 \leq b \leq 1$. It follows that

*the function $b$ is a smooth bump function: $b$ is $\mathcal{C}^\infty$ with support $[-2,2]$ and $b(x) = 1$ for $x \in [-1,1]$ and $0 \leq b \leq 1$.*

Note that given $a < b < c < d$ there exists a function $B \in \mathcal{C}^\infty$ such that $b = 1$ on $[b,c]$, $b = 0$ outside $[a,b]$ and $0 \leq b \leq 1$. To do so, one choses $b(x) = h(\alpha+\beta x)h(\gamma-\delta x)$ with $\beta,\delta \geq 0$, $\gamma - \delta d = \alpha + \beta a = 0$ and $\alpha + \beta b = \gamma - \delta c = 1$. The choice is thus

$$\alpha = \frac{-a}{b-a}, \quad \beta = \frac{1}{b-a}, \quad \gamma = \frac{d}{d-c}, \quad \delta = \frac{1}{d-c}.$$

Note that one may tensor such functions: $b(x_1,\ldots,x_d) = \prod_{i=1}^d b_i(x_i)$. Then, if $Q_1, Q_2$ are two cubes with the closure of $Q_1$ in the interior of $Q_2$ (so that the boundaries don't touch) then there exists $b \in \mathcal{C}^\infty$ such that $b(x) = 1$ on $Q_1$, $b(x) = 0$ outside $Q_2$ and $0 \leq b \leq 1$.

It should be noted that once we have an element of $\mathcal{C}_c^\infty$, we get many others:

LEMMA 6.3. *Let $\varphi \in L^1(\mathbb{R}^d)$ and $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ then $\varphi * f \in \mathcal{C}^\infty(\mathbb{R}^d)$ and, if $\varphi$ is compactly supported then so if $\varphi * f \in \mathcal{C}_c(\mathbb{R}^d)$.*

We will define the support of $\varphi \in L^1(\mathbb{R}^d)$ in a precise way later on, here we simply mean that there is an $R > 0$ such that $\varphi(x) = 0$ whenever $\|x\| \geq R$.

PROOF. Indeed, if $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ then $f$ is bounded so that $\varphi * f(x) = \displaystyle\int_{\mathbb{R}^d} \varphi(t) f(x-t)\,\mathrm{d}t$. Set $F(x,t) = \varphi(t)f(x-t)$ and note that, for $t$ fixed, $x \to F(x,t)$ is $\mathcal{C}^\infty$ (unless $|\varphi(t)| = +\infty$ so this is true for almost every $t$). Further for every $\alpha \in \mathbb{N}^d$, $\partial_x^\alpha F(x,t) = \varphi(t)\partial^\alpha f(x-t)$. But $\partial^\alpha f$ is continuous with compact support so that it is bounded $|\partial^\alpha f(u)| \leq C_\alpha$ thus $|\partial_x^\alpha F(x,t)| \leq C_\alpha |\varphi(t)| \in L^1(\mathbb{R}^d)$. Lebesgue's derivation theorem then implies that $\varphi * f$ is of class $\mathcal{C}^\infty$ with $\partial^\alpha(\varphi * f) = \varphi * \partial^\alpha f$.

Finally, if $\varphi$ and $f$ are both compactly supported, there is an $R$ such that, if $\|t\| \geq R$ and $\|u\| \geq R$ then $\varphi(t) = 0$ and $f(u) = 0$. But then, if $\|x\| \geq 2R$ and $\|t\| \leq R$, $\|x-t\| \geq R$. It follows that, when $\|x\| \geq 2R$, $F(x,t) = \varphi(t)f(x-t) = 0$ for all $t \in \mathbb{R}^d$ thus $\varphi * f = \int F(x,t)\,\mathrm{d}t = 0$. □

Although $\mathcal{C}_c(\mathbb{R}^d)$ is a large class (we will even see that it is dense in every $L^p(\mathbb{R}^d)$ space with $p < +\infty$), this class is too small to contain a function like the Gaussian. We will thus define a larger class that has almost the same property. To do so, for $\alpha,\beta\in\mathbb{N}^d$ and $f:\mathbb{R}^d\to\mathbb{C}$, let

$$p_{\alpha,\beta}(f)=\sup_{x\in\mathbb{R}^d}|x^\alpha\partial^\beta f(x)|.$$

Definition 6.4. The Schwarz class is the set

$$\mathcal{S}(\mathbb{R}^d)=\{f\in\mathcal{C}^\infty(\mathbb{R}^d)\,:\forall\alpha,\beta\in\mathbb{N}^d,p_{\alpha,\beta}(f)<+\infty\}.$$

The Schwarz class is thus the space of all smooth functions such that all derivatives have fast decrease at infinity (*i.e.* faster than any polynomial). The class is not empty as obviously $\mathcal{C}_c^\infty(\mathbb{R}^d)\subset\mathcal{S}(\mathbb{R}^d)$.

Example 6.5. Let $f$ be a Gaussian, $f(x)=e^{-a\|x\|^2}$, $a>0$ (the norm is the Euclidean norm). Then $f\in\mathcal{S}(\mathbb{R}^d)$.

For simplicity, we will show this for $d=1$ and $a=1/2$ so $f(x)=e^{-x^2/2}$. Then, for every $k$, there exists a polynomial $P_k$ such that $f^{(k)}(x)=P_k(x)e^{-x^2/2}$. This is clear since $P_0=1$ and, by induction, $f^{(k+1)}(x)=\big(P_k'(x)-xP_k(x)\big)e^{-x^2/2}$ and $P_{k+1}=P_k'(x)-xP_k(x)$ is a polynomial if $P_k$ is. Finally, $x^NP_k(x)e^{-x^2/2}$ is clearly bounded.

It should be noted that the choice of $p_{\alpha,\beta}$ to define $\mathcal{S}(\mathbb{R}^d)$ is somewhat arbitrary. We may as well take $m,n$ two integers and define

$$\tilde{p}_{m,n}(f)=\sup_{x\in\mathbb{R}^d}(1+\|x\|^2)^m\sum_{|\beta|\leq n}\left|\frac{\partial^\beta f}{\partial x^\beta}f(x)\right|.$$

Then if we notice that $(1+\|x\|^2)^m$ is a polynomial of degree $2m$

$$(1+\|x\|^2)^m=\sum_{|\alpha|\leq 2m}c_\alpha x^\alpha$$

and $C=\max|c_\alpha|$ then

$$\tilde{p}_{m,n}(f)\leq\sum_{|\alpha|\leq 2m}|c_\alpha|\sum_{|\beta|\leq n}\sup_{x\in\mathbb{R}^d}|x^\alpha\partial^\beta f(x)|\leq C\sum_{|\alpha|\leq 2m}\sum_{|\beta|\leq n}p_{\alpha,\beta}(f).$$

On the other hand,

$$|x^\alpha|=|x_1|^{\alpha_1}\cdots|x_d|^{\alpha_d}\leq\|x\|_\infty^{|\alpha|}\leq\|x\|_2^{|\alpha|}\leq(1+\|x\|_2^2)^{|\alpha|}.$$

For the last inequality, one considers the cases $\|x\|_2\leq 1$ and $\|x\|_2\geq 1$. But then

$$p_{\alpha,\beta}(f)\leq\tilde{p}_{|\alpha|,|\beta|}(f).$$

It follows that

$$\mathcal{S}(\mathbb{R}^d)=\{f\in\mathcal{C}^\infty(\mathbb{R}^d)\,:\forall m,n\in\mathbb{N},\tilde{p}_{m,n}(f)<+\infty\}.$$

This change of "semi-norm" is sometimes convenient, for instance for the following lemma

Lemma 6.6. *For every $1\leq p\leq\infty$, $\mathcal{S}(\mathbb{R}^d)\subset L^p(\mathbb{R}^d)$.*

Proof. The lemma is trivial when $p=+\infty$ since $\tilde{p}_{0,0}(f)=\|f\|_\infty$.

For other $p$'s we will use the fact that, integrating in polar coordinates

$$\int_{\mathbb{R}^d}\frac{\mathrm{d}x}{(1+\|x\|^2)^\kappa}=\int_{\mathbb{S}^{d-1}}\int_0^{+\infty}\frac{r^{d-1}}{(1+r^2)^\kappa}\,\mathrm{d}r\,\mathrm{d}\sigma_{d-1}(\theta)$$
$$=\sigma_{d-1}(\mathbb{S}^{d-1})\int_0^{+\infty}\frac{r^{d-1}}{(1+r^2)^\kappa}\,\mathrm{d}r<+\infty$$

if $2\kappa > d$. It follows that, if

$$\int_{\mathbb{R}^d} |f(x)|^p \,\mathrm{d}x = \int_{\mathbb{R}^d} |(1+\|x\|^2)^d f(x)|^p \frac{\mathrm{d}x}{(1+\|x\|^2)^{dp}} \leq \tilde{p}_{d,0}(f) \int_{\mathbb{R}^d} \frac{\mathrm{d}x}{(1+\|x\|^2)^{dp}} < +\infty.$$

□

It is now easy to prove the following that we leave as an exercice

PROPOSITION 6.7. *Let $\alpha \in \mathbb{N}^d$, $\lambda, \mu \in \mathbb{C}$, $T \in GL(\mathbb{R}^d)$ an invertible linear transformation. Then*
– *if $f, g \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ so is $\lambda f + \mu g$, $f \circ T$, $fg$, $x^\alpha f$, $\partial^\alpha f$;*
– *if $f, g \in \mathcal{S}(\mathbb{R}^d)$ so is $\lambda f + \mu g$, $f \circ T$, $fg$, $x^\alpha f$, $\partial^\alpha f$.*

Let us now extend Lemma 6.3 which shows that we can add $f * g$ to the above list.

LEMMA 6.8. *Let $1 \leq p \leq \infty$, $\varphi \in L^p(\mathbb{R}^d)$ and $f \in \mathcal{S}(\mathbb{R}^d)$ then $\varphi * f \in \mathcal{C}^\infty(\mathbb{R}^d)$. Further if, for every $\alpha \in \mathbb{N}^d$, $t^\alpha \varphi \in L^p(\mathbb{R}^d)$ then $\varphi * f \in \mathcal{S}(\mathbb{R}^d)$.*

The second part of the lemma is satisfied if $\varphi$ is compactly supported or if $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

PROOF. The general scheme of proof is the same as for Lemma 6.3. Note that, as $\mathcal{S}(\mathbb{R}^d) \subset L^{p'}(\mathbb{R}^d)$, $1/p + 1/p' = 1$, we have $\varphi * f \in L^\infty(\mathbb{R}^d)$ and

$$\varphi * f(x) = \int_{\mathbb{R}^d} \varphi(t) f(x-t)\,\mathrm{d}t.$$

For $p = 1$, there is nothing to change: we again define $F(t,x) = \varphi(t)f(x-t)$ and, for every $\alpha \in \mathbb{N}^d$ $\partial_x^\alpha F(t,x) = \varphi(t)\partial^\alpha f(x-t)$ so that $|\partial_x^\alpha F(t,x)| \leq p_{\alpha,0}(f)|\varphi(t)| \in L^1(\mathbb{R}^d)$. By Lebesgue's Differentiation Theorem, $\varphi * f$ is of class $\mathcal{C}^\infty$ with $\partial^\alpha(\varphi * f) = \varphi * (\partial^\alpha f)$.

For $p > 1$, this can not work and we need to use the fact that $f$ has some extra decrease that can compensate the fact that $\varphi \notin L^1$. First, note that it is enough to show that $\varphi * f$ is of class $\mathcal{C}^\infty$ on the ball $B(0,R)$ with $R$ arbitrary. So assume that $\|x\| \leq R$

$$\partial_x^\alpha F(t,x) = \varphi(t)\partial^\alpha f(x-t) = \frac{\varphi(t)}{(1+\|t\|^2)^d} \frac{(1+\|t\|^2)^d}{(1+\|x-t\|^2)^d} (1+\|x-t\|^2)^d \partial^\alpha f(x-t).$$

First, as $(1+\|t\|^2)^{-d} \in L^{p'}(\mathbb{R}^d)$ (it is in all $L^q(\mathbb{R}^d)$ spaces, $q \geq 1$) and $\varphi \in L^p$, Hölder's inequality shows that $\Phi(t) := \dfrac{|\varphi(t)|}{(1+\|t\|^2)^d} \in L^1(\mathbb{R}^d)$.

Next $(1+\|x-t\|^2)^d |\partial^\alpha f(x-t)| \leq \tilde{p}_{d,|\alpha|}(f)$.

Finally if $|t| \geq 2R$, and $|x| \leq R$, $|x-t| \geq |t| - |x| \geq |t| - R \geq |t|/2$ so that

$$\frac{(1+\|t\|^2)^d}{(1+\|x-t\|^2)^d} \leq \left(\frac{1+\|t\|^2}{1+\|t\|^2/4}\right)^d \leq 4^d$$

while for $|t| \leq 2R$,

$$\frac{(1+\|t\|^2)^d}{(1+\|x-t\|^2)^d} \leq (1+2R)^d.$$

Assuming $R \geq 2$, we get that this bound also holds for $|t| \geq 2R$ and finally

$$|\partial_x^\alpha F(t,x)| \leq \tilde{p}_{d,|\alpha|}(f)(1+2R)^d \Phi(t) \in L^1(\mathbb{R}).$$

By Lebesgue's Differentiation Theorem, $\varphi * f$ is of class $\mathcal{C}^\infty$ with $\partial^\alpha(\varphi * f) = \varphi * (\partial^\alpha f)$ on $B(0,R)$ and as $R$ is arbitrary, the same holds on $\mathbb{R}^d$.

It remains to prove that, for all $\alpha, \beta$, $x^\alpha \partial^\beta(\varphi * f) = x^\alpha \varphi * (\partial^\beta f)$ is bounded. As $f \in \mathcal{S}(\mathbb{R}^d)$ implies that $\partial^\beta f \in \mathcal{S}(\mathbb{R}^d)$, it is enough to consider the case $\beta = 0$. But now, define $M_i\psi(t) = t_i\psi(t)$, then

$$\begin{aligned} x_i\varphi * f(x) = \int_{\mathbb{R}}^d \varphi(t)x_i f(x-t)\,\mathrm{d}t &= \int_{\mathbb{R}}^d \varphi(t)(x_i - t_i)f(x-t)\,\mathrm{d}t + \int_{\mathbb{R}}^d t_i\varphi(t)f(x-t)\,\mathrm{d}t \\ &= \varphi * M_i f + M_i\varphi * f \end{aligned}$$

which is bounded since $M_i\varphi \in L^p(\mathbb{R}^d)$ and $M_i f \in \mathcal{S}(\mathbb{R}^d)$. An induction on the length of $\alpha$ then shows that, for every $\alpha \in \mathbb{N}^d$, $x^\alpha\varphi * f$ is bounded. □

Remark 6.9. A careful examination of the above proofs shows that, for $\varphi \in L^p(\mathbb{R}^d)$ and $f \in \mathcal{C}^k(\mathbb{R}^d)$ such that for every $\alpha$ with $|\alpha| \leq k$ there is a $\kappa > 0$ such that $(1+|t|^2)^{-\kappa} \in L^{p'}$ (i.e. $2\kappa p' > d$) and $(1+|t|^2)^\kappa \partial^\alpha f \in L^\infty$, we have $\varphi * f \in \mathcal{C}^k$.

### 6.2. Regularization by convolution.

Theorem 6.10 (Approximation of unity).
*Let $1 \leq p < +\infty$ and $j \in \mathcal{S}(\mathbb{R}^d)$ be such that $j \geq 0$ and $\int_{\mathbb{R}^d} j(x)\,dx = 1$. For $s > 0$, denote by $j_s$ the function defined by $j_s(t) = s^{-d}j(t/s)$.*

*Then, for every $\varphi \in L^p(\mathbb{R}^d)$, $\varphi * j_s \in \mathcal{C}^\infty(\mathbb{R}^d)$ and $\varphi * j_s \to \varphi$ in $L^p$ when $s \to 0$.*

*For $p = +\infty$, $L^\infty$ has to be replaced by $\mathcal{C}_0(\mathbb{R}^d)$: for every $\varphi \in \mathcal{C}_0(\mathbb{R}^d)$, $\varphi * j_s \in \mathcal{C}^\infty(\mathbb{R}^d)$ and $\varphi * j_s \to \varphi$ uniformly when $s \to 0$.*

Proof. We will only give the proof for $1 \leq p < +\infty$. We leave to the reader the case $p = +\infty$. The only thing that one needs to use is the fact that functions in $\mathcal{C}_0(\mathbb{R}^d)$ are uniformly continuous.

Let us first note that $j_s \in \mathcal{S}(\mathbb{R}^d)$ and that

$$\int_{\mathbb{R}^d} j_s(t)\,\mathrm{d}t = \int_{\mathbb{R}^d} j(t/s)\,s^{-d}\mathrm{d}t = \int_{\mathbb{R}^d} j(r)\,\mathrm{d}r = 1$$

with a change of variable $r = t/s$.

We have thus already seen that $\varphi * j_s \in \mathcal{C}^\infty(\mathbb{R}^d)$.

Next, $j_s \in L^{p'}(\mathbb{R}^d)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, so that

$$\varphi * j_s(x) = \int_{\mathbb{R}^d} j_s(t)\varphi(x-t)\,\mathrm{d}t.$$

But then

$$\begin{aligned} \varphi(x) - \varphi * j_s(x) &= f(x)\int_{\mathbb{R}^d} j_s(t)\,\mathrm{d}t - \int_{\mathbb{R}^d} j_s(t)\varphi(x-t)\,\mathrm{d}t \\ &= \int_{\mathbb{R}^d} j_s(t)\big(\varphi(x) - \varphi(x-t)\big)\,\mathrm{d}t. \end{aligned}$$

From Minkowski's inequality we deduce that

$$\|\varphi - \varphi * j_s\|_p \leq \int_{\mathbb{R}^d} j_s(t)\|\varphi - \tau_t\varphi\|_p\,\mathrm{d}t.$$

Now fix $\varepsilon > 0$. As $p < +\infty$, we have seen that $\|\varphi - \tau_t\varphi\|_p \to 0$ when $t \to 0$ so that there exists $\eta > 0$ such that, if $|t| < \eta$, $\|\varphi - \tau_t\varphi\|_p \leq \varepsilon$. When $|t| \geq \eta$ we can simply use that $\|\varphi - \tau_t\varphi\|_p \leq 2\|\varphi\|_p$ We then write

$$\begin{aligned}\|\varphi - \varphi * j_s\|_p &\leq \int_{|t|\leq\eta} j_s(t)\|\varphi - \tau_t\varphi\|_p \,\mathrm{d}t + \int_{|t|\geq\eta} j_s(t)\|\varphi - \tau_t\varphi\|_p \,\mathrm{d}t \\ &\leq \varepsilon \int_{\mathbb{R}^d} j_s(t)\,\mathrm{d}t + 2\|\varphi\|_p \int_{|t|\geq\eta} j_s(t)\,\mathrm{d}t.\end{aligned}$$

It remains to notice that $\int_{\mathbb{R}^d} j_s(t)\,\mathrm{d}t = 1$ and that

$$\int_{|t|\geq\eta} j_s(t)\,\mathrm{d}t = \int_{|t|\geq\eta} s^{-d} j(t/s)\,\mathrm{d}t = \int_{r\geq\eta/s} j(r)\,\mathrm{d}r \to 0$$

when $s \to 0$. In particular, there is an $\eta'$ such that, if $s < \eta'$, $2\|\varphi\|_p \displaystyle\int_{|t|\geq\eta} j_s(t)\,\mathrm{d}t \leq \varepsilon$ and then $\|\varphi - \varphi * j_s\|_p \leq 2\varepsilon$. □

Again, the hypothesis can be weakend without changing the proof. To do so, we may assume that $(j_s)_{s\geq 0}$ is a family of $L^1(\mathbb{R}^d)$ functions such that

(1) there is a constant $C > 0$ such that, for all $s > 0$,

$$\int_{\mathbb{R}^d} j_s(x)\,\mathrm{d}x = 1 \quad \text{and} \quad \int_{\mathbb{R}^d} |j_s(x)|\,\mathrm{d}x \leq C.$$

(2) For every $\eta > 0$, $\int_{|x|\geq\eta} |j_s(x)|\,\mathrm{d}x \to 0$ when $s \to 0$.

Such a family is called an approximation of the identity (and sometimes a mollifier).

Corollary 6.11. *The space $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is dense in every $L^p(\mathbb{R}^d)$ space with $1 \leq p < +\infty$ and thus so is every space containing it like $\mathcal{C}_c(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$.*

Proof. Let $f \in L^p(\mathbb{R}^d)$ and $\varepsilon > 0$. Let $j \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $j_s(t) = s^{-d}j(t/s)$ First, for $R$ large enough $\left\|f - f\mathbf{1}_{|x|\leq R}\right\| \leq \varepsilon$. Next there exists $s$ such that $\left\|f\mathbf{1}_{|x|\leq R} - (f\mathbf{1}_{|x|\leq R}) * j_s\right\| \leq \varepsilon$. But then $\left\|f - (f\mathbf{1}_{|x|\leq R}) * j_s\right\| \leq 2\varepsilon$ and $(f\mathbf{1}_{|x|\leq R}) * j_s \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. □

Remark 6.12. One has to be careful with the density of $\mathcal{C}_c(\mathbb{R}^d)$ in $L^p(\mathbb{R}^d)$. The proof given here relies on approximation of unity. This in turn relies on the fact that translations are continuous.

We have proven this last fact by first proving it for characteristic functions of cubes, from which we deduced the fact for simple step functions. Then we concluded that translations are continuous by density of step functions in $L^p$. Our proof is thus not circular.

It turns out that it is simpler to prove that translations are continuous by first proving this fact for functions in $\mathcal{C}_c(\mathbb{R}^d)$ and then using the density of this last step. The approximation of unity theorem then allows to prove that $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is dense in $L^p$, but the density of $\mathcal{C}_c(\mathbb{R}^d)$ then needs a different proof.

## 7. Partition of unity

The aim of this section is to decompose the function 1 into a sum of smooth bum functions with controled support. Before doing so, we need a result from topology:

Proposition 7.1. *Let $\Omega \subset \mathbb{R}^d$ be an open set and let $\{\Omega_i, i \in I\}$ be an open cover of $\Omega$: each $\Omega_i$ is open and $\Omega \subset \bigcup_{i\in I} \Omega_i$. Then there exists a sequence $(x_k)_{k\in\mathbb{N}} \subset \Omega$ and a sequence $(r_k)_{k\in\mathbb{N}}$ of positive numbers such that*

(i) $\Omega = \displaystyle\bigcup_{k\geq 0} B(x_k, r_k)$;

(ii) *for each $k$ there is an $i \in I$ such that $B(x_k, 2r_k) \subset \Omega_i$;*
(iii) *each $x \in \Omega$ has an open neighbourhood $U$ such that $\{k : U \cap B(x_k, r_k) \neq \emptyset\}$ is finite.*

Let us postpone the proof of this proposition. We can now prove the following:

THEOREM 7.2 (Partition of unity). *Let $\Omega \subset \mathbb{R}^d$ be an open set and let $\{\Omega_i, i \in I\}$ be an open cover of $\Omega$. Then, for each $i \in I$, there exists $f_i \in \mathcal{C}^\infty(\mathbb{R}^d)$ such that*

(i) $0 \leq f_i \leq 1$*;*
(ii) $\operatorname{supp} f_i \subset \Omega_i$*;*
(iii) *each $x \in \Omega$ has an open neighborhood $U$ such that $\{i : f_i \neq 0$ on $U\}$ is finite;*
(iv) $\displaystyle\sum_{i \in I} f_i(x) = 1$ *for every $x \in \Omega$.*

DEFINITION 7.3. We say that $(f_i)_{i \in I}$ is a *partition of unity* subordinated to $(\Omega_i)_{i \in I}$.

PROOF. Let $x_k, r_k$ be given by Proposition 7.1. Let $g_k$ be a smooth function with $g_k > 0$ in $B(x_k, \varepsilon_k)$ and $g_k(x) = 0$ for $|x - x_k| \geq r_k$. Such a function was constructed in Example 6.1. Let $g = \sum_{k \geq 0} g_k$. As a sum of positive numbers, this always exists (but may be infinite for the moment)

Now, according to Proposition 7.1iii, for each $x \in \Omega$ there is an open neighborhood $U$ such that $J_U := \{k : B(x_k, 2\varepsilon_k) \cap U\}$ is finite. But then, if $y \in U$, $g_k(y) = 0$ unless $k \in J_U$. It follows that

$$\sum_{k \geq 0} g_k(y) = \sum_{k \in J_U} g_k(y)$$

on $U$. As this is a finite sum of smooth functions, it is a smooth function on $U$. As $x$ was arbitrary, it follows that $g$ is smooth in a neighborhood of each point in $\Omega$ thus it is smooth on $\Omega$. Further, given $x \in \Omega$, from Proposition 7.1i, there is at least one $k$ such that $x \in B(x_k, r_k)$ thus $g_k(x) > 0$. As $g_j(x) \geq 0$ for $j \neq k$, we get that $g(x) > 0$.

Next define $h_k = g_k/g$ which is $\mathcal{C}^\infty$, with support $B(x_k, r_k)$ and $\sum_{k \geq 0} h_k = 1$. Now take $i \in I$ and write $J_i = \{k : B(x_k, r_k) \subset \Omega_i\}$. Let $\tilde{f}_i = \sum_{k \in J_i} h_k$. One then repeats the previous proof and gets tat $\tilde{f}_i$ and $F := \sum \tilde{f}_i$ are smooth and $F > 0$ thus $f_i = \tilde{f}_i/F$ is smooth. Next, by definition, if $f_i(x) \neq 0$ then there is a $j \in J_i$ such that $h_k(x) \neq 0$ thus $x \in B(x_k, r_k) \subset \Omega_i$. Thus this $f_i$ have the required properties. $\square$

CHAPTER 4

# The Fourier transform

## 1. The $L^1$-theory

Definition 1.1. For $f \in L^1(\mathbb{R}^d)$ we define the *Fourier transform* of $f$, and denote it either by $\hat{f}$ or $\mathcal{F}f$, the function defined on $\mathbb{R}$ by

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x.$$

Let us start with a fundamental example:

Example 1.2. Let $a < b \in \mathbb{R}$ and $f = \mathbf{1}_{[a,b]}$. Then if $\xi \neq 0$,

$$\begin{aligned}
\hat{f}(\xi) &= \int_a^b e^{-2i\pi x\xi}\,\mathrm{d}x = \frac{-1}{2i\pi\xi}\big(e^{-2i\pi b\xi} - e^{-2i\pi a\xi}\big)\\
&= \frac{e^{2i\pi\frac{a+b}{2}\xi}}{\pi\xi}\,\frac{e^{2i\pi\frac{b-a}{2}\xi} - e^{-2i\pi\frac{b-a}{2}\xi}}{2i}\\
&= e^{2i\pi\frac{a+b}{2}\xi}\,\frac{\sin\pi(b-a)\xi}{\pi\xi}.
\end{aligned}$$

When $\xi = 0$, $\hat{f}(\xi) = \displaystyle\int_a^b \mathrm{d}x = b-a$.

It is convenient to introduce the function $\operatorname{sinc} t = \begin{cases} 1 & \text{if } t = 0\\ \frac{\sin t}{t} & \text{if } t \neq 0\end{cases}$. Note that this is an analytic function.

If we write $c = \frac{a+b}{2}$ for the center of the interval $[a,b]$ and $\ell$ for its length, $\ell = 2r$ then

$$\hat{f}(\xi) = \ell e^{2i\pi c\xi}\operatorname{sinc}\pi\ell\xi = 2re^{2i\pi c\xi}\operatorname{sinc}2\pi r\xi.$$

Let us now notice that, if $f$ is a tensor function $f(x_1,\ldots,x_d) = \displaystyle\prod_{j=1}^d f_j(x_j)$, then so does $\hat{f}$: $\hat{f}(\xi_1,\ldots,\xi_d) = \displaystyle\prod_{j=1}^d \hat{f}_j(\xi_j)$. This follows directly from Fubini's Theorem and the fact that $e^{-2i\pi\langle x,\xi\rangle} = e^{-2i\pi\sum_{j=1}^d x_j\xi_j} = \displaystyle\prod_{j=1}^d e^{-2i\pi x_j\xi_j}$.

Now, for $Q = \displaystyle\prod_{j=1}^d [a_j,b_j]$ is a cube, write $\ell_j = b_j - a_j$ for its side length, $|Q| = \displaystyle\prod_{j=1}^d \ell_j$ for its volume, $c = \left(\frac{a_1+b_1}{2},\ldots,\frac{a_d+b_d}{2}\right)$ for its center of gravity. Let $f(x) = \mathbf{1}_Q(x) =$

$\prod_{j=1}^{d} \mathbf{1}_{[a_j,b_j]}(x_j)$ then

$$\hat{f}(\xi) = |Q|\, e^{2i\pi\langle c,\xi\rangle} \prod_{j=1}^{d} \operatorname{sinc} \pi \ell_j \xi_j.$$

Note for future use that $\hat{f} \in \mathcal{C}_0(\mathbb{R}^d)$.

Let us now start detailing properties of the Fourier transform. First, it is well defined. Indeed, let $F(x,\xi) = f(x)e^{-2i\pi\langle x,\xi\rangle}$. Then, for $x$ fixed, $\xi \to F(x,\xi)$ is continuous. Moreover, $|F(x,\xi)| = |f(x)| \in L^1(\mathbb{R}^d)$, it follows that $\hat{f}(\xi) = \displaystyle\int_{\mathbb{R}^d} F(x,\xi)\,\mathrm{d}x$ is well defined and continuous. Further,

$$|\hat{f}(\xi)| \leq \int_{\mathbb{R}^d} |F(x,\xi)|\,\mathrm{d}x = \int_{\mathbb{R}^d} |f(x)|\,\mathrm{d}x = \|f\|_{L^1(\mathbb{R}^d)}.$$

As $f \to \hat{f}$ is clearly linear, this shows that this mapping is bounded $L^1(\mathbb{R}^d) \to \mathcal{C}_b(\mathbb{R}^d)$, the space of bounded continuous functions on $\mathbb{R}^d$. Actually, a bit more is true:

THEOREM 1.3 (Riemann-Lebesgue Lemma). *The Fourier transform $\mathcal{F}$ is a bounded linear mapping $L^1(\mathbb{R}^d) \to \mathcal{C}_0(\mathbb{R}^d)$ with $\|\mathcal{F}f\|_\infty \leq \|f\|_1$.*

PROOF. We have already seen that $\mathcal{F}$ is a bounded linear mapping $L^1(\mathbb{R}^d) \to \mathcal{C}_b(\mathbb{R}^d)$ with $\|\mathcal{F}f\|_\infty \leq \|f\|_1$. It remains to prove that $\mathcal{F}f \in \mathcal{C}_0(\mathbb{R}^d)$ when $f \in L^1(\mathbb{R}^d)$.

This is indeed the case when $f = \mathbf{1}_Q$, $Q$ a cube, thus also when $f$ is a (finite) linear combination of such functions, that is, when $f$ is a step function. But step functions are dense. Thus, if $f \in L^1(\mathbb{R}^d)$, there exists a sequence $f_k$ of step functions, such that $\|f_k - f\|_{L^1} \to 0$ when $k \to \infty$. But then

$$\|\mathcal{F}f - \mathcal{F}f_k\|_\infty = \|\mathcal{F}(f - f_k)\|_\infty \leq \|f - f_k\|_1 \to 0.$$

In other words, $\mathcal{F}f_k \to \mathcal{F}f$ in $\mathcal{C}_b(\mathbb{R}^d)$. As $\mathcal{F}f_k \in \mathcal{C}_0(\mathbb{R}^d)$ which is closed in $\mathcal{C}_b(\mathbb{R}^d)$ (see the chapter on convolutions for a proof), we get that $\mathcal{F}f \in \mathcal{C}_0(\mathbb{R}^d)$. □

A SECOND PROOF. In dimension 1, there is an alternative proof of the fact that $\hat{f}(\xi) \to 0$ when $\xi \to \pm\infty$. First note that $-1 = e^{-i\pi} = e^{-2i\pi\xi/2\xi}$ thus

$$\begin{aligned} 2\hat{f}(\xi) &= \int_{\mathbb{R}} f(t)e^{-2i\pi t\xi}\,\mathrm{d}t - \int_{\mathbb{R}} f(t)e^{2i\pi\xi/2\xi}e^{-2i\pi t\xi}\,\mathrm{d}t \\ &= \int_{\mathbb{R}} f(t)e^{-2i\pi t\xi}\,\mathrm{d}t - \int_{\mathbb{R}} f(t)e^{-2i\pi(t+1/2\xi)\xi}\,\mathrm{d}t \\ &= \int_{\mathbb{R}} \left[f(t) - f\left(t - \frac{1}{2\xi}\right)\right] e^{-2i\pi t\xi}\,\mathrm{d}t. \end{aligned}$$

In other words, $\hat{f}(\xi) = \dfrac{1}{2}\mathcal{F}[f - \tau_{1/2\xi}f](\xi)$. It follows that $|\hat{f}(\xi)| \leq \left\|f - \tau_{1/2\xi}f\right\|_1$. Now letting $\xi \to \pm\infty$ and using the continuity of $a \to \tau_a f$ from $\mathbb{R} \to L^1(\mathbb{R})$ shows that $|\hat{f}(\xi)| \to 0$.

Recall that this continuity required the same density argument. □

Let us now list the main properties of the Fourier transform. To do so, we need to introduce some notation. For $a, \omega \in \mathbb{R}^d$, $\lambda > 0$, $T \in GL_n(\mathbb{R}^d)$ (a $d \times d$ invertible matrix) and $f$ a function on $\mathbb{R}^d$, we define new functions on $\mathbb{R}^d$

$$\tau_a f(x) = f(x-a),\ M_\omega f(x) = e^{-2i\pi\langle\omega,x\rangle}f(x),\ \delta_\lambda f(x) = f(\lambda x),\ \Delta_T f(x) = f(T^{-1}x).$$

Note that $\tau_a, M_\omega, \delta_\lambda, \Delta_T$ are continuous linear mappings $L^p \to L^p$ for every $p$.

PROPOSITION 1.4. *Assume that $f \in L^1(\mathbb{R}^d)$ then*

– $\mathcal{F}[\tau_a f] = M_a\mathcal{F}[f]$, $\mathcal{F}[M_\omega f] = \tau_{-\omega}\mathcal{F}[f]$,

– $\mathcal{F}[\delta_\lambda f] = \lambda^{-d}\mathcal{F}[\delta_{1/\lambda} f]$ *and more generally* $\mathcal{F}[\Delta_T f] = |\det T|\Delta_{[T^{-1}]^t}\mathcal{F}[f]$.

– *If $\xi_j f \in L^1(\mathbb{R}^d)$ then $\hat{f}$ admits a continuous partial derivative in the $\xi_j$ direction with* $\dfrac{\partial \hat{f}}{\partial \xi_j}(\xi) = -2i\pi\mathcal{F}[x_j f](\xi)$.

– *If $f$ is $\mathcal{C}^1$ with* $\dfrac{\partial f}{\partial x_j} \in L^1(\mathbb{R}^d)$, *then* $\mathcal{F}\left[\dfrac{\partial f}{\partial x_j}\right](\xi) = 2i\pi\xi_j\mathcal{F}[f]\xi)$.

– *If $f, g \in L^1(\mathbb{R}^d)$ then* $\mathcal{F}[f*g] = \mathcal{F}[f]\mathcal{F}[g]$.

PROOF. The first 4 follow from a simple change of variable

– changing variable $y = x - a$,

$$\mathcal{F}[\tau_a f](\xi) = \int_{\mathbb{R}^d} f(x-a)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = \int_{\mathbb{R}^d} f(y)e^{-2i\pi\langle y+a,\xi\rangle}\,\mathrm{d}y$$
$$= e^{-2i\pi\langle a,\xi\rangle}\int_{\mathbb{R}^d} f(y)e^{-2i\pi\langle y,\xi\rangle}\,\mathrm{d}y = e^{-2i\pi\langle a,\xi\rangle}\hat{f}(\xi).$$

– the next one is even easier

$$\mathcal{F}[M_\omega f](\xi) = \int_{\mathbb{R}^d} f(x)e^{-2i\pi\langle\omega,\xi\rangle}e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)e^{-2i\pi\langle x+\omega,\xi\rangle}\,\mathrm{d}x = \hat{f}(\xi+\omega).$$

– changing variable $y = \lambda x$,

$$\mathcal{F}[\delta_\lambda f](\xi) = \int_{\mathbb{R}^d} f(\lambda x)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = \lambda^{-d}\int_{\mathbb{R}^d} f(y)e^{-2i\pi\langle y/\lambda,\xi\rangle}\,\mathrm{d}y$$
$$= \lambda^{-d}\int_{\mathbb{R}^d} f(y)e^{-2i\pi\langle y,\xi/\lambda\rangle}\,\mathrm{d}y = \lambda^{-d}\hat{f}(\xi/\lambda).$$

It is a particular case of the following:

– changing variable $y = T^{-1}x$, $x = Ty$

$$\mathcal{F}[\Delta_T f](\xi) = \int_{\mathbb{R}^d} f(T^{-1}x)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = |\det T|\int_{\mathbb{R}^d} f(y)e^{-2i\pi\langle Ty,\xi\rangle}\,\mathrm{d}y$$
$$= |\det T|\int_{\mathbb{R}^d} f(y)e^{-2i\pi\left\langle y,T^t\xi\right\rangle}\,\mathrm{d}y = |\det T|\hat{f}(T^t\xi).$$

–The next two ones are slightly more subtle. First assume that $x_j f \in L^1(\mathbb{R}^d)$ and consider again $F(x,\xi) = f(x)e^{-2i\pi\langle x,\xi\rangle}$. Then, for $x$ fixed, $\xi \to F(x,\xi)$ is of class $\mathcal{C}^1$, $|F(x,\xi)| = |f(x)| \in L^1(\mathbb{R}^d)$ and

$$\left|\frac{\partial F}{\partial \xi_j}(x,\xi)\right| = \left|-2i\pi x_j f(x)e^{-2i\pi\langle x,\xi\rangle}\right| = 2\pi|x_j f| \in L^1(\mathbb{R}^d).$$

It follows that $\hat{f}(\xi) = \displaystyle\int_{\mathbb{R}^d} F(x,\xi)\,\mathrm{d}x$ is differentiable with respect to $\xi_j$ with

$$\frac{\partial \hat{f}}{\partial \xi_j}(\xi) = \int_{\mathbb{R}^d}\frac{\partial F}{\partial \xi_j}(x,\xi) = \int_{\mathbb{R}^d} -2i\pi x_j f(x)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = \mathcal{F}[-2i\pi x_j f](\xi).$$

– Now, assume that $f \in \mathcal{C}^1$, $f, \dfrac{\partial f}{\partial \xi_j} \in L^1$. To simplify notation, we will take $j = 1$. Note that, from Fubini's Theorem,

$$\int_{\mathbb{R}^d} |f(x)|\,\mathrm{d}x = \int_{\mathbb{R}^{d-1}} \left( \int_{\mathbb{R}} |f(x_1, x_2, \ldots, x_d)|\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \cdots \mathrm{d}x_d < +\infty$$

so that $\displaystyle\int_{\mathbb{R}} |f(x_1, x_2, \ldots, x_d)|\,\mathrm{d}x_1 < +\infty$ for almost every $(x_2, \ldots, x_d)$. The same is true with $\frac{\partial f}{\partial \xi_j}$ replacing $f$. If two properties hold almost everywhere, they jointly hold almost everywhere. We may thus take an $(x_2, \ldots, x_d)$ such that

$$\int_{\mathbb{R}} |f(x_1, x_2, \ldots, x_d)|\,\mathrm{d}x_1 < +\infty \quad \text{and} \quad \int_{\mathbb{R}} \left| \frac{\partial f}{\partial \xi_j}(x_1, x_2, \ldots, x_d) \right| \mathrm{d}x_1 < +\infty$$

and almost every $(x_2, \ldots, x_d)$ is like that. The fundamental theorem of calculus then shows that

$$\begin{aligned} f(x_1, x_2, \ldots, x_d) = f(0, x_2, \ldots, x_d) + \int_0^{x_1} \frac{\partial f}{\partial \xi_j}(t, x_2, \ldots, x_d)\mathrm{d}t \\ \to f(0, x_2, \ldots, x_d) + \int_0^{\pm\infty} \frac{\partial f}{\partial \xi_j}(t, x_2, \ldots, x_d)\mathrm{d}t \end{aligned}$$

when $x_1 \to \pm\infty$. Thus $f(x_1, x_2, \ldots, x_d)$ has a limit in $\pm\infty$. But then $\displaystyle\int_{\mathbb{R}} |f(x_1, x_2, \ldots, x_d)|\,\mathrm{d}x_1 < +\infty$ implies that this limit is zero.

Next, write $x, \xi \in \mathbb{R}^d$ as $x = (x_1, \bar{x})$, $\xi = (\xi_1, \bar{\xi})$ with $\bar{x}, \bar{\xi} \in \mathbb{R}^{d-1}$. Integrating by parts,

$$\begin{aligned} \int_{\mathbb{R}} \frac{\partial f}{\partial \xi_1}(x_1, \bar{x}) e^{-2i\pi\langle x, \xi\rangle} \mathrm{d}x_1 &= \int_{\mathbb{R}} \frac{\partial f}{\partial \xi_1}(x_1, \bar{x}) e^{-2i\pi x_1 \xi_1} \mathrm{d}x_1 e^{-2i\pi\left\langle \bar{x}, \bar{\xi}\right\rangle} \\ &= e^{-2i\pi\left\langle \bar{x}, \bar{\xi}\right\rangle} \left[ f(x_1, \bar{x}) e^{-2i\pi x_1 \xi_1} \right]_{-\infty}^{+\infty} \\ &\quad +2i\pi\xi_1 \int_{\mathbb{R}} f(x_1, \bar{x}) e^{-2i\pi x_1 \xi_1} \mathrm{d}x_1 e^{-2i\pi\left\langle \bar{x}, \bar{\xi}\right\rangle} \\ &= 2i\pi\xi_1 \int_{\mathbb{R}} f(x_1, \bar{x}) e^{-2i\pi\langle x, \xi\rangle}\,\mathrm{d}x_1. \end{aligned}$$

It remains to integrate with respect to the $d-1$ remaining variables and to use Fubini.

The last property is a direct consequence of Fubini and the change of variable $u = x - y$

$$\begin{aligned} \mathcal{F}[f * g](\xi) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y) g(x-y)\,\mathrm{d}y\, e^{-2i\pi\langle x, \xi\rangle}\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f(y) \int_{\mathbb{R}^d} g(x-y) e^{-2i\pi\langle x, \xi\rangle}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d} f(y) \int_{\mathbb{R}^d} g(u) e^{-2i\pi\langle u+y, \xi\rangle}\,\mathrm{d}u\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d} f(y) \int_{\mathbb{R}^d} g(u) e^{-2i\pi\langle u, \xi\rangle}\,\mathrm{d}u\, e^{-2i\pi\langle y, \xi\rangle}\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d} f(y) \hat{g}(\xi)\, e^{-2i\pi\langle y, \xi\rangle}\,\mathrm{d}y = \hat{h}(\xi)\hat{g}(\xi) \end{aligned}$$

as claimed. $\square$

We can now give as a second example the case of the Gaussian:

EXAMPLE 1.5. Let $f$ be the Gaussian defined for $x \in \mathbb{R}$ by $f(x) = e^{-\pi x^2}$, then $\hat{f}(\xi) = e^{-\pi\xi^2}$.

Indeed, first note that $\hat{f}(0) = \displaystyle\int_{\mathbb{R}} e^{-\pi x^2}\,\mathrm{d}x$. But then, using Fubini in the first line and changing to polar coordinates:

$$\begin{aligned}
\hat{f}(0)^2 &= \int_{\mathbb{R}} e^{-\pi x^2}\,\mathrm{d}x \int_{\mathbb{R}} e^{-\pi y^2}\,\mathrm{d}y = \int_{\mathbb{R}^2} e^{-\pi(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^{+\infty}\int_0^{2\pi} e^{-\pi r^2}\,\mathrm{d}\theta r\,\mathrm{d}r \\
&= \int_0^{+\infty} 2\pi r e^{-\pi r^2}\,\mathrm{d}r = [-e^{-\pi r^2}]_0^{+\infty} = 1.
\end{aligned}$$

As $\hat{f}(0)$ is the integral of a positive function, $\hat{f}(0) \geq 0$ thus $\hat{f}(0) = 1$.

Next, note that $f$ satisfies the differential equation $f' = -2\pi x f$ thus $\mathcal{F}[f'] = -2\pi\mathcal{F}[xf]$. As clearly $f$ is $\mathcal{C}^1$ with $f, xf, f' \in L^1$ we can use the above properties: $\hat{f}' = -2i\pi\mathcal{F}[xf]$ $\mathcal{F}[f'] = 2i\pi\xi\hat{f}$. It follows that $\hat{f}$ satisfies the differential equation $(\hat{f})' = -2\pi\xi\hat{f}$ which is the same equation as the one satisfied by the Gaussian. Thus $\hat{f} = cf$. Comparing values at 0, we get $\hat{f} = f$.

In higher dimensions, we immediately get that, if $\gamma(x) = e^{-\pi|x|^2}$ then $\hat{\gamma}(\xi) = e^{-\pi|\xi|^2}$.

Now, let $A$ be a positive definite symetric matrix and $f(x) = e^{-\pi\langle Ax,x\rangle}$.

As $A$ is a real sumetric matrix, it is diagonalizable in an orthonormal matrix, $A = P\Delta P^t$ with $\Delta$ a diagonal martrix and $P$ an orthogonal matrix. Write $\Delta = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_d)$. As $A$ is positive definite, the $\lambda_j$'s are $> 0$ thus we can write $\lambda_j = \mu_j^2$. Then define $B = P\mathrm{diag}\,(\mu_1, \ldots, \mu_d)P^t$ and notice that $B^t = B$ and that $A = B^2 = B^tB$. It follows that $\langle Ax, x\rangle = \langle B^tBx, x\rangle = |Bx|^2$. As the $\mu_j$'s are $> 0$, $B$ is invertible thus $f(x) = \gamma(Bx)$. It follows that $f \in L^1(\mathbb{R}^d)$ and that $\hat{f}(x) = |\det B^{-1}|\gamma(B^{-1}x)$. But $B^{-1} = P\mathrm{diag}\,(1/\mu_1, \ldots, 1/\mu_d)P^t$ is symetric with $(B^{-1})^tB^{-1} = (B^{-1})^2 = A^{-1}$ thus $|\det B^{-1}| = \det(A)^{-1/2}$ and

$$|B^{-1}x|^2 = \left\langle B^{-1}x, B^{-1}x\right\rangle = \left\langle (B^{-1})^tB^{-1}x, x\right\rangle = \left\langle A^{-1}x, x\right\rangle$$

It follows that $\hat{f}(\xi) = \det(A)^{-1/2}e^{-\pi\left\langle A^{-1}x,x\right\rangle}$.

## 2. The inversion formula and the Fourier transform on $\mathcal{S}(\mathbb{R}^d)$

We are now going to show that the Fourier transform can be inverted and that it is (almost) its own inverse. To do so, let us start with the following simple observation:

Assume that $f, g \in L^1(\mathbb{R}^d)$, then $\hat{f}, \hat{g} \in \mathcal{C}_0(\mathbb{R}^d)$ so that $f\hat{g}$ and $g\hat{f}$ are both integrable. But as

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x)g(y)|\,\mathrm{d}y\,\mathrm{d}x = \|f\|_{L^1}\|g\|_{L^1} < +\infty,$$

Fubini's theorem shows that

$$\begin{aligned}
\int_{\mathbb{R}^d} f(x)\hat{g}(x)\,\mathrm{d}x &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x)g(y)e^{-2i\pi\langle x,y\rangle}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} g(y)\int_{\mathbb{R}^d} f(x)e^{-2i\pi\langle y,x\rangle}\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d} g(y)\hat{f}(y)\,\mathrm{d}y. \qquad (2.16)
\end{aligned}$$

Let us now replace $g$ by $M_\omega g$ so that $\hat{g}$ is replaced by $\tau_\omega \hat{g}$ (more precisely, if we replace $g$ by $M_{-\omega} g$). We get

$$\int_{\mathbb{R}^d} f(x)\hat{g}(x-\omega)\,\mathrm{d}x = \int_{\mathbb{R}^d} g(y)\hat{f}(y)e^{2i\pi\langle\omega,y\rangle}\,\mathrm{d}y. \tag{2.17}$$

The right hand side looks like a convolution and is indeed $g * \hat{f}$ when $g$ is even. Let us take as an example $g(y) = e^{-\pi|\lambda y|^2}$ so that $\hat{g}(x) = \lambda^{-d}e^{-\pi|x/\lambda|^2}$. Write $\gamma_\lambda(x) = \lambda^{-d}e^{-\pi|x/\lambda|^2}$. Then (2.17) reads

$$f * \gamma_\lambda(\omega) = \int_{\mathbb{R}^d} e^{-\pi|\lambda y|^2}\hat{f}(y)e^{2i\pi\langle\omega,y\rangle}\,\mathrm{d}y. \tag{2.18}$$

Now, since $\gamma \in \mathcal{S}(\mathbb{R}^d)$, according to Theorem 6.10, $f*\gamma_\lambda \to f$ in $L^1(\mathbb{R}^d)$. In particular, if $f_1, f_2 \in L^1(\mathbb{R}^d)$ are such that $\hat{f}_1 = \hat{f}_2$ then $f_1 * \gamma_\lambda(\omega) = f_2 * \gamma_\lambda(\omega)$. Letting $\lambda \to 0$ shows that $f_1 = f_2$. In other words, the Fourier transform is one-to-one.

What about the right hand side? Note that $e^{-\pi|\lambda y|^2}\hat{f}(y)e^{2i\pi\langle\omega,y\rangle} \to \hat{f}(y)e^{2i\pi\langle\omega,y\rangle}$ when $\lambda \to 0$. Further, as $|e^{-\pi|\lambda y|^2}\hat{f}(y)e^{2i\pi\langle\omega,y\rangle}| = |e^{-\pi|\lambda y|^2}\hat{f}(y)| \leq |\hat{f}(y)|$, if $\hat{f} \in L^1(\mathbb{R}^d)$, we can use dominated convergence and obtain the following theorem:

THEOREM 2.1 (Fourier inversion formula). *The Fourier transform is one-to-one $L^1(\mathbb{R}^d) \to \mathcal{C}_0(\mathbb{R}^d)$. Let $f \in L^1(\mathbb{R}^d)$ be such that $\hat{f} \in L^1(\mathbb{R}^d)$, then $f \in \mathcal{C}_0(\mathbb{R}^d)$ and*

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi)e^{2i\pi\langle\xi,x\rangle}\,\mathrm{d}\xi.$$

PROOF. We have not fully proven the above theorem, we have only shown that the inversion formula is valid in $L^1(\mathbb{R}^d)$. The observation is that the right hand side is $\mathcal{F}[\hat{f}](-x)$. As $\hat{f} \in L^1(\mathbb{R}^d)$, Riemann-Lebesgue's lemma implies that the right hand side is in $\mathcal{C}_0$. Now $f*\gamma_\lambda \to f$ in $L^1$ thus has a subsequence that converges almost-everywhere, thus $f$ is almost everywhere equal to $\mathcal{F}[\hat{f}](-x)$ i.e. is in the same class as a $\mathcal{C}_0$ function. Our convention is that we chose $f$ to be this $\mathcal{C}_0$ function. □

The Fourier inversion theorem shows that the Fourier transform is almost its own inverse, this explains the very symetric properties we have already observed in Proposition 1.4.

REMARK 2.2. If $f = \mathbf{1}_{[-1,1]}$ then $\hat{f} = \operatorname{sinc} 2\pi t \notin L^1(\mathbb{R})$. It follows that $\int_{\mathbb{R}} \hat{f}(\xi)e^{2i\pi\xi x}\,\mathrm{d}\xi$ does not make sense. We will see below that

$$\lim_{R,S\to+\infty} \int_{-R}^{S} \hat{f}(\xi)e^{2i\pi\xi x}\,\mathrm{d}\xi \to \mathbf{1}_{[-1,1]}(x)$$

in $L^2$. Actually,

$$\lim_{R\to+\infty} \int_{-R}^{R} \hat{f}(\xi)e^{2i\pi\xi x}\,\mathrm{d}\xi \to \mathbf{1}_{[-1,1]}(x)$$

is valid pointwise, excepted at the jumps $\pm 1$. Note that we now integrate over a symetric interval.

REMARK 2.3. It is important to have in mind that the Fourier transform is *not* a bijection $L^1(\mathbb{R}^d) \to \mathcal{C}_0(\mathbb{R}^d)$ as there are functions in $\mathcal{C}_0(\mathbb{R}^d)$ that are not Fourier transforms of $L^1$ functions.

Now, let $f \in \mathcal{S}(\mathbb{R}^d)$. For every $\alpha \in \mathbb{N}^d$, $x^\alpha f \in \mathcal{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$. It follows from Proposition 1.4 that $\hat{f} \in \mathcal{C}^\infty(\mathbb{R}^d)$ and $\partial^\alpha \hat{f} = (-2i\pi)^{|\alpha|}\mathcal{F}[x^\alpha f]$. Further as $x^\alpha f \in \mathcal{S}$, for every $\beta \in \mathbb{N}^d$, $\partial^\beta(x^\alpha f) \in \mathcal{S} \subset L^1(\mathbb{R}^d)$. Applying again Proposition 1.4 we obtain that $x^\beta \partial^\alpha \hat{f} = (-2i\pi)^{|\alpha|-|\beta|}\mathcal{F}[\partial^\beta(x^\alpha f)]$. But then, Riemann-Lebesgue's Lemma implies that $\mathcal{F}[\partial^\beta(x^\alpha f)]$ is in $\mathcal{C}_0$, in particular, it is bounded. We have just shown that, $\hat{f} \in \mathcal{C}^\infty(\mathbb{R}^d)$ and that, for every $\alpha, \beta \in \mathbb{N}^d$, $x^\beta \partial^\alpha \hat{f}$ is bounded, that is, that $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$.

Finally, as $\mathcal{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$, the Fourier inversion theorem applies to every $f \in \mathcal{S}(\mathbb{R}^d)$ and for such an $f$, $f(x) = \mathcal{F}[\hat{f}](-x)$. Writing $Z\hat{f}(y) = \hat{f}(-y)$ and noticing that $Z\hat{f} \in \mathcal{S}(\mathbb{R}^d)$ and that $\mathcal{F}[\hat{f}](-x) = \mathcal{F}[Zf](x)$, we see that every $f \in \mathcal{S}(\mathbb{R}^d)$ is the Fourier transform of a function in the Schwartz class. We have thus shown the following:

Theorem 2.4. *The Fourier transform is a bijection $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$. The inverse map is given by $\mathcal{F}^{-1}[f](\xi) = \mathcal{F}[f](-\xi)$.*

## 3. The $L^2$-theory

Our aim here is to extend the Fourier transform to other $L^p$ spaces. Let us recall that if $f, g \in \mathcal{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$ then

$$\int_{\mathbb{R}^d} f(x)\hat{g}(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} g(y)\hat{f}(y)\,\mathrm{d}y.$$

Now let $h \in \mathcal{S}(\mathbb{R}^d)$, then $\bar{h} \in \mathcal{S}(\mathbb{R}^d)$ and the Fourier inversion Formula reads

$$\bar{h}(x) = \overline{\int_{\mathbb{R}^d} \hat{h}(y)e^{2i\pi\langle y,x\rangle}\,\mathrm{d}y} = \int_{\mathbb{R}^d} \overline{\hat{h}(y)}e^{-2i\pi\langle y,x\rangle}\,\mathrm{d}y = \mathcal{F}[\overline{\hat{h}(y)}].$$

We now replace $g$ by $\overline{\hat{h}(y)} \in \mathcal{S}(\mathbb{R}^d)$ in the above formula. We thus obtain

$$\int_{\mathbb{R}^d} f(x)\overline{h(x)}\,\mathrm{d}x = \int_{\mathbb{R}^d} \hat{f}(y)\overline{\hat{h}(y)}\,\mathrm{d}y, \qquad f, h \in \mathcal{S}(\mathbb{R}^d).$$

In particular, taking $h = f$, we get $\|\mathcal{F}[f]\|_{L^2(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)}$ for every $f \in \mathcal{S}(\mathbb{R}^d)$. As $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, we can apply the Banach extension principle. It follows that $\mathcal{F}$ extends to a continuous linear mapping $L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. Further the mapping $\mathcal{F}^{-1}(f)(x) = \mathcal{F}(f)(-x)$ also extends from $\mathcal{S}(\mathbb{R}^d)$ to a continuous linear mapping $L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. As $\mathcal{F}^{-1}\big[\mathcal{F}[f]\big] = \mathcal{F}\big[\mathcal{F}^{-1}[f]\big] = f$ for all $f \in \mathcal{S}(\mathbb{R}^d)$, by density of $\mathcal{S}(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$, this identity stays true for $f \in L^2(\mathbb{R}^d)$. In particular, $\mathcal{F}$ is a bijection $L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ and its inverse map is $\mathcal{F}^{-1}$.

Finally, the mappings $\tau_a, M_\omega, \delta_\lambda, \Delta_T$ are all continuous on $L^2(\mathbb{R}^d)$, so the corresponding properties in Proposition 1.4 stay true in $L^2(\mathbb{R}^d)$.

In summary

Theorem 3.1. *The Fourier transform extends into a continuous linear mapping $L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ and the extended map is a bijection. The mapping is an isometry and satisfies*

– *Plancherel's identity: for all $f \in L^2(\mathbb{R}^d)$*

$$\int_{\mathbb{R}^d} |f(x)|^2\,dx = \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\,d\xi.$$

– *Parseval's identity: for all $f, g \in L^2(\mathbb{R}^d)$*

$$\int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^d} \hat{f}(\xi)\overline{g(\xi)}\,d\xi.$$

*Further, the identities* $\mathcal{F}[\tau_a f] = M_a\mathcal{F}[f]$, $\mathcal{F}[M_\omega f] = \tau_{-\omega}\mathcal{F}[f]$, $\mathcal{F}[\delta_\lambda f] = \lambda^{-d}\mathcal{F}[\delta_{1/\lambda} f]$ *and* $\mathcal{F}[\Delta_T f] = |\det T|\Delta_{[T^{-1}]^t}\mathcal{F}[f]$ *are all valid for* $f \in L^2(\mathbb{R}^d)$

Let us note that the convolution identity $\widehat{f*g} = \hat{f}\hat{g}$ does not extend to $f, g \in L^2(\mathbb{R}^d)$ as in this case $f * g \in \mathcal{C}_0(\mathbb{R}^d)$ and $\widehat{f*g}$ might not make sense. We will see later how to overvome this.

EXAMPLE 3.2. Let $a > 0$ and define $f$ on $\mathbb{R}$ as $e_a^+(t) = \mathbf{1}_{[0,+\infty)}e^{-at}$ and $e_a^-(t) = \mathbf{1}_{(-\infty,0]}e^{at}$. Note that $e_a^\pm \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ so that its Fourier transform is given by

$$\hat{e}_a^+(\xi) = \int_0^{+\infty} e^{-(a+2i\pi\xi)t}\,\mathrm{d}t = \frac{1}{a+2i\pi\xi}$$

while

$$\hat{e}_a^-(\xi) = \int_{-\infty}^{0} e^{(a-2i\pi\xi)t}\,\mathrm{d}t = \frac{1}{a-2i\pi\xi}.$$

Let $c_a^\pm$ be defined on $\mathbb{R}$ by $c_a^\pm(x) = \dfrac{1}{a \pm 2i\pi x}$. Note that $c_a^\pm \in L^2$ but *not* in $L^1$ so that it has an $L^2$-Fourier transform but not an $L^1$-Fourier transform. Never the less $c_a^\pm = \mathcal{F}[e_a^\pm]$ in $L^1$-sense thus also in the $L^2$-sense. Thus, the Fourier inversion theorem gives $\mathcal{F}[c_a^\pm](\xi) = \mathcal{F}\big[\mathcal{F}[e_a^\pm]\big](\xi) = \mathcal{F}^{-1}\big[\mathcal{F}[e_a^\pm]\big](-\xi) = e_a^\pm(-\xi) = e_a^\mp(\xi)$. This has to be understood in the $L^2$ sense, in particular, equalities hold only almost everywhere.

One may notice that $e_a^\pm$ is not continuous so that, according to Riemann-Lebesgue, they are not Fourier transforms of $L^1$ functions.

EXAMPLE 3.3. **An example of a function in $\mathcal{C}_0$ that is not a Fourier transform of an $L^1$ function.**

Let us define $f$ on $\mathbb{R}$ by $f(t) = \frac{\mathrm{sgn}(t)}{1+|t|}$. Note that $f \in L^2(\mathbb{R})$ but $f \notin L^1(\mathbb{R})$. The Fourier transform of $f$ can thus not be calculated via $\int f(t)e^{-2i\pi t\xi}\,\mathrm{d}t$ but only as an $L^2$ limit. To carry out this limit, we will need the following identity

$$\frac{1}{1+|t|} = \int_0^{+\infty} e^{-(1+|t|)x}\,\mathrm{d}x.$$

Using Fubini's Theorem, we see that

$$\begin{aligned}\int_{-R}^{R} \frac{\mathrm{sgn}(t)}{1+|t|}e^{-2i\pi t\xi}\,\mathrm{d}t &= \int_{-R}^{R}\int_0^{+\infty} \mathrm{sgn}(t)e^{-(1+|t|)x}\,\mathrm{d}x e^{-2i\pi t\xi}\,\mathrm{d}t \\ &= \int_0^{+\infty}\int_{-R}^{R} \mathrm{sgn}(t)e^{-(1+|t|)x}e^{-2i\pi t\xi}\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_0^{+\infty} e^{-x}\int_{-R}^{R} \mathrm{sgn}(t)e^{-|t|x}e^{-2i\pi t\xi}\,\mathrm{d}t\,\mathrm{d}x \qquad (3.19)\end{aligned}$$

To see that one is allowed to apply Fubini's theorem, one writes $|\operatorname{sgn}(t)e^{-(1+|t|)x}e^{-2i\pi t\xi}| = e^{-(1+|t|)x} \leq e^{-x} \in L^1([-R,R]\times\mathbb{R}, \mathrm{d}t\,\mathrm{d}x)$. But now, if $\xi \neq 0$, (or $x \neq 0$)

$$\begin{aligned}\int_{-R}^{R} e^{-|t|x}e^{-2i\pi t\xi}\,\mathrm{d}t &= -\int_{-R}^{0} e^{t(x-2i\pi\xi)}\,\mathrm{d}t + \int_0^R e^{-t(x+2i\pi\xi)}\,\mathrm{d}t\\ &= \left[-\frac{e^{t(x-2i\pi\xi)}}{x-2i\pi\xi}\right]_{-R}^{0} + \left[-\frac{e^{-t(x+2i\pi\xi)}}{x+2i\pi\xi}\right]_0^R\end{aligned}$$

$$= \frac{-1+e^{-R(x-2i\pi\xi)}}{x-2i\pi\xi} + \frac{1-e^{-R(x+2i\pi\xi)}}{x+2i\pi\xi} = \frac{-4i\pi\xi}{x^2+(2\pi\xi)^2} + \frac{e^{-R(x-2i\pi\xi)}}{x-2i\pi\xi} - \frac{e^{-R(x+2i\pi\xi)}}{x+2i\pi\xi}.$$

Inserting this into (3.19) gives

$$\int_{-R}^{R}\frac{e^{-2i\pi t\xi}}{1+|t|}\,\mathrm{d}t = 4i\pi\xi\int_0^{+\infty}\frac{e^{-x}}{x^2+(2\pi\xi)^2}\,\mathrm{d}x + \int_0^{+\infty}\left(\frac{e^{-R(x-2i\pi\xi)}}{x-2i\pi\xi} - \frac{e^{-R(x+2i\pi\xi)}}{x+2i\pi\xi}\right)e^{-x}\,\mathrm{d}x.$$

But, if $x > 0$

$$\frac{e^{-R(x-2i\pi\xi)}}{x-2i\pi\xi} - \frac{e^{-R(x+2i\pi\xi)}}{x+2i\pi\xi} \to 0$$

when $R \to +\infty$ while, if $\xi \neq 0$,

$$\begin{aligned}\left|\left(\frac{e^{-R(x-2i\pi\xi)}}{x-2i\pi\xi} - \frac{e^{-R(x+2i\pi\xi)}}{x+2i\pi\xi}\right)e^{-x}\right| &\leq \left(\frac{e^{-Rx}}{|x-2i\pi\xi|} + \frac{e^{-Rx}}{|x+2i\pi\xi|}\right)e^{-x}\\ &\leq \frac{e^{-x}}{\pi\xi} \in L^1(\mathbb{R}).\end{aligned}$$

We may thus apply domintated convergence and obtain that, for $\xi \neq 0$

$$\int_0^{+\infty}\left(\frac{e^{-R(x-2i\pi\xi)}}{x-2i\pi\xi} - \frac{e^{-R(x+2i\pi\xi)}}{x+2i\pi\xi}\right)e^{-x}\,\mathrm{d}x \to 0$$

when $R \to +\infty$ and thus

$$\lim_{R\to+\infty}\int_{-R}^{R}\frac{e^{-2i\pi t\xi}}{1+|t|}\,\mathrm{d}t = 4i\pi\xi\int_0^{+\infty}\frac{e^{-x}}{x^2+(2\pi\xi)^2}\,\mathrm{d}x.$$

But, the $L^2$-limit of this integral (seen as a function of $\xi$) is the Fourier transform of $f$. It follows that, for almost every $\xi$,

$$\hat{f}(\xi) = 4i\pi\xi\int_0^{+\infty}\frac{e^{-x}}{x^2+(2\pi\xi)^2}\,\mathrm{d}x = 2i\operatorname{sgn}(\xi)\int_0^{+\infty}\frac{e^{-2\pi|\xi|u}}{u^2+1}\,\mathrm{d}u$$

with a change of variable $x = 2\pi|\xi|u$.

One may observe that this function is continuous except at 0 where it has a jump discontinuity and goes to 0 at infinity. This follows immediately from Lebesgue's theorem: if we write $F(\xi,u) = \dfrac{e^{-2\pi|\xi|u}}{u^2+1}$ then

– $|F(\xi,u)| = \left|\dfrac{e^{-2\pi|\xi|u}}{u^2+1}\right| \leq \dfrac{1}{1+u^2} \in L^1(\mathbb{R}^+)$;

– if we fix $u$, $\xi \to F(\xi,u)$ thus $\xi \to \int_0^{+\infty} F(\xi,u)\,\mathrm{d}u$ is continuous over $\mathbb{R}$. In particular $\displaystyle\int_0^{+\infty} F(\xi,u)\,\mathrm{d}u = \int_0^{+\infty}\frac{1}{1+u^2}\,\mathrm{d}u = \frac{\pi}{2}$

– if we fix $u > 0$, $F(\xi,u) \to 0$ when $\xi \to \pm\infty$. Thus $\int_0^{+\infty} F(\xi,u)\,\mathrm{d}u \to 0$ as well.

Thus $\hat f(\xi) = 2i\,\mathrm{sgn}(\xi)\int_0^{+\infty} F(\xi,u)\,\mathrm{d}u$ has the properties we just announced with $\hat f(0^+) = -\hat f(0^-) = i\pi$.

One should also note that $\lim_{R\to+\infty}\int_{-R}^{R} f(t)\,\mathrm{d}t = 0$ since $f$ is odd.

Let us now erase the jump discontinuities with the help of the previous example. Let $g = f + i\pi(c_1^+ - c_1^-) = \frac{\mathrm{sgn}(t)}{1+|t|} + \frac{4\pi t}{1+(2\pi t)^2}$. Note that $g(t)\sim 3\,\mathrm{sgn}(t)/t$ in $\pm\infty$ so that $g\in L^2(\mathbb{R})$ but not in $L^1(\mathbb{R})$. By linerity $\hat g = \hat f - i\pi e_1^+ + i\pi e_1^-$. All three functions $\hat f, e_1^+, e_1^-$ are continuous outside 0 and $\hat f(0^\pm) = \pm i\pi$, $e_1^\pm(0^\mp) = 0$, $e_1^\pm(0^\pm) = 1$. Thus the jump discontinuities cancel.

## 4. Solving the heat equation

The aim of this section is to show how Fourier analysis can be used to solve some partial differential equations. As an example, we will here take the heat equation:

$$(E)\qquad \begin{cases}\partial_t u(x,t) &=& \Delta_x(x,t)\\ u(x,0) &=& u_0(x)\end{cases}$$

where $\Delta_x u(x,t) = (\partial_{x_1}^2 + \cdots + \partial_{x_d}^2)u(x,t)$. The unknown is a function $u$ on $\mathbb{R}^d\times(0,+\infty)$ and the variable $t$ represents time while the $x$ variable is a space variable. The meaning of $u(x,0) = u_0(x)$ has to be taken as $u(x,0)\to u_0(x)$ when $t\to 0$ in some sense that we will make precise later.

For the moment, we will leave a side all mathematical rogour and compute the space Fourier transform: for $\xi\in\mathbb{R}^d$, write $\hat u(\xi,t) = \displaystyle\int_{\mathbb{R}^d} u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x$.

If $u$ is a resonable function, then

$$(4.20)\qquad \partial_t \hat u(\xi,t) = \partial_t\int_{\mathbb{R}^d} u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = \int_{\mathbb{R}^d}\partial_t u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x.$$

Note that we are not trying to justify the fact that the differentiation can be entered into the integral. Further, under good circonstances, we may integrate by parts to obtain

$$\int_{\mathbb{R}^d}\partial_{x_i}u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = -\int_{\mathbb{R}^d}u(x,t)\partial_{x_j}e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = 2i\pi\xi_j\int_{\mathbb{R}^d}u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x$$

(at least if $u$ vaniashes at infinity). Repating this

$$\int_{\mathbb{R}^d}\partial_{x_i}^2 u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = -4\pi^2\xi_j^2\int_{\mathbb{R}^d}u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x$$

and summing up, we get

$$\int_{\mathbb{R}^d}\Delta_x u(x,t)e^{-2i\pi\langle x,\xi\rangle}\,\mathrm{d}x = -4\pi^2|\xi|^2\hat u(\xi,t).$$

Together with (4.20), this shows that $(E)$ implies

$$\begin{cases}\partial_t\hat u(\xi,t) &=& -4\pi^2|\xi|^2\hat u(\xi,t)\partial_t\hat u(\xi,0) &=& \hat u_0(\xi)\end{cases}.$$

Notice that, when $\xi$ is fixed, this is an ordinary differential equation which admits as a unique solution

$$\hat u(\xi,t) = e^{-4\pi^2|\xi|^2 t}\hat u_0(\xi).$$

Remmebering that $e^{-\pi|x|^2}$ is its own Fourier transform, a simple computation shows that, if $p_t(x) = (4\pi t)^{-d/2}e^{-|x|^2/4t}$, then $\hat p_t(\xi) = e^{-4\pi^2|\xi|^2 t}$ thus $\hat u(\xi,t) = \hat p_t(\xi)\hat u_0(\xi) = \mathcal{F}[p_t * u_0](\xi)$

with the convolution theorem. It remains to invert the Fourier transform and

$$u(x,t) = p_t * u_0(x) = \int_{\mathbb{R}^d} u_0(y) p_t(x-y)\,\mathrm{d}x. \tag{4.21}$$

So far, we have not been rigourous and have not been looking for any justification. There are two things to do

– justify that this is indeed a solution of $\partial_t u = \Delta_x u$. To do so, one first checks that $p(t,x) = p_t(x)$ satistifies the heat equation. Then $u(x,t) = p_t * u_0$ will also satisfy the heat equation if one can enter the differentiation operators inside the integral apearing in (4.21). This can be done with Lebesgue's differentiation theorem and the fact that, if $t_0 < t < t_1$, for every $\alpha \in \mathbb{N}^{d+1}$ there is a constant $C = C(\alpha, t_0, t_1)$ such that $|\partial^\alpha p(t,x)| \le Ce^{-|x|^2/C}$ ($\partial^\alpha$ means differentiation in space and time).

– Once one knows that $\partial_t u = \Delta_x u$, one notices that $p_t$ is an approximation of unity so that, if $u_0 \in L^p$ for some $1 \le p < \infty$, then $p_t * u_0 \to u_0$ in $L^p$.

– The only thing this method does not provide is the fact that there is no other solution. Of course, it is the only solution that can be obtained via a Fourier transform. We will see in the next chapter how to deal with that issue.

CHAPTER 5

# Distributions

## 1. Definition and examples

DEFINITION 1.1. A *distribution* $T \in \mathcal{D}'(\mathbb{R}^d)$ is a linear functional on $\mathcal{C}_c^\infty(\mathbb{R}^d)$ such that, for every $R > 0$, there exists $N \in \mathbb{N}$ and $C > 0$ such that, for every $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \subset B(0, R)$,

$$|\langle T, \varphi\rangle| \leq C \sup_{\alpha\in\mathbb{N}^d, |\alpha|\leq N} \sup_{x\in\mathbb{R}^d} |\partial^\alpha\varphi(x)|. \tag{1.22}$$

If $N$ can be chosen independent of $R$, we say that $T$ is of finite order and the order of $T$ is the smallest such $N$.

A *tempered distribution* $T \in \mathcal{S}'(\mathbb{R}^d)$ is a linear functional on $\mathcal{S}(\mathbb{R}^d)$ such that there exists $M, N \in \mathbb{N}$ and $D > 0$ such that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$|\langle T, \varphi\rangle| \leq D \sup_{\alpha\in\mathbb{N}^d, |\alpha|\leq N} \sup_{x\in\mathbb{R}^d} (1 + |x|)^M |\partial^\alpha\varphi(x)|. \tag{1.23}$$

If $N$ can be chosen independent of $R$, we say that $T$ is of finite order and the order of $T$ is the smallest such $N$.

In the previous definition, $\langle T, f\rangle$ stands for $T(f)$ and is here a more convenient notation.

Let us recall that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$. Further if we fix $R > 0$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \subset B(0, R)$, then for every $\alpha \in \mathbb{N}^d$, $\partial^\alpha\varphi(x) = 0$ if $|x| > R$. It follows that

$$(1 + |x|)^M |\partial^\alpha\varphi(x)| \leq (1 + R)^M |\partial^\alpha\varphi(x)|.$$

Thus, (1.23) implies (1.22) with $C = D(1 + R)^M$. In particular, $T \in \mathcal{D}'(\mathbb{R}^d)$ and $T$ has order of at most $N$. In other words: we just proved the following;

LEMMA 1.2. *Every tempered distribution is a distribution of finite order.*

Before developping the properties of distributions, let us first notice that every locally integrable function can be identified with a distribution.

EXAMPLE 1.3. **Locally integrable functions**
Let $f \in L^1_{loc}(\mathbb{R}^d)$. Recall that this means that, for every $R > 0$, $f\mathbf{1}_{B(0,R)} \in L^1(\mathbb{R}^d)$. Note also that, for every $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$,

$$\langle T_f, \varphi\rangle = \int_{\mathbb{R}^d} f(x)\varphi(x)\,\mathrm{d}x$$

is well defined. Also $T_f$ uniquely determines $f$. Indeed, fix $\varphi \in \mathcal{C}_c(\mathbb{R}^d)$ and let $R > 0$ such that $\operatorname{supp}\varphi \subset B(0, R)$. Write $\varphi_t(x) = t^{-d}\varphi(x/t)$ so that $\varphi_t(x) = 0$ if $|x| \geq tR$. Then for $0 < t < 1$ and $y \in B(0, S)$, $\varphi_t(y - x) = 0$ if $|x| \geq R + S$ since $|y - x| \geq |x| - |y| \geq R + S - S = R \geq tR$. It follows that

$$\mathbf{1}_{B(0,S)}(y) \int_{\mathbb{R}^d} f(x)\varphi_t(y-x)\,\mathrm{d}x = \mathbf{1}_{B(0,S)}(y) \int_{\mathbb{R}^d} \mathbf{1}_{B(0,R+S)}(x)f(x)\varphi_t(y-x)\,\mathrm{d}x = \mathbf{1}_{B(0,S)}(y)(\mathbf{1}_{B(0,R+S)}f)*\varphi_t(y).$$

Now $\mathbf{1}_{B(0,R+S)}f$ so that $(\mathbf{1}_{B(0,R+S)}f)*\varphi_t\to\mathbf{1}_{B(0,R+S)}f$ in $L^1$ as $t\to 0$. Therefore, there exists a sequence $t_j\to 0$ such that $\mathbf{1}_{B(0,S)}(y)(\mathbf{1}_{B(0,R+S)}f)*\varphi_t\to\mathbf{1}_{B(0,S)}(y)\mathbf{1}_{B(0,R+S)}f=\mathbf{1}_{B(0,S)}(y)f$ almost everywhere. It follows that $f$ is determined almost everywhere on $B(0,S)$. As $S$ is arbitrary, $f$ is determined almost everywhere on $\mathbb{R}^d$ by $T_f$.

Finally, if $\varphi\in\mathcal{C}_c^\infty$ with $\operatorname{supp}\varphi\subset B(0,R)$,

$$|\langle T_f,\varphi\rangle|=\left|\int_{\mathbb{R}^d}f(x)\varphi(x)\,\mathrm{d}x\right|=\left|\int_{B(0,R)}f(x)\varphi(x)\,\mathrm{d}x\right|\leq\int_{B(0,R)}|f(x)\varphi(x)|\,\mathrm{d}x\leq\int_{B(0,R)}|f(x)|\,\mathrm{d}x\sup_{x\in\mathbb{R}^d}|\varphi(x)|.$$

This shows that $T_f$ is a distribution of order 0.

However, it does not show that $T_f$ is a tempered distribution. For this, we require that $f$ be tempered in the sense that there exists an integer $m$ such that $\displaystyle\int_{\mathbb{R}^d}\frac{|f(x)|}{(1+|x|)^m}\,\mathrm{d}x<+\infty$ then $T_f$ is a tempered distribution since

$$|\langle T_f,\varphi\rangle|=\left|\int_{\mathbb{R}^d}f(x)\varphi(x)\,\mathrm{d}x\right|\leq\int_{\mathbb{R}^d}\frac{|f(x)|}{(1+|x|)^m}(1+|x|)^m|\varphi(x)|\,\mathrm{d}x\leq\int_{\mathbb{R}^d}\frac{|f(x)|}{(1+|x|)^m}\,\mathrm{d}x\sup_{x\in\mathbb{R}^d}(1+|x|)^m|\varphi(x)|.$$

In summary:

LEMMA 1.4. *For $f\in L^1_{loc}(\mathbb{R}^d)$, define $T_f:\varphi\to\langle T_f,\varphi\rangle=\displaystyle\int_{\mathbb{R}^d}f(x)\varphi(x)\,dx$.*

*Then $T_f\in\mathcal{D}'(\mathbb{R}^d)$, is of order $0$ and $f\to T_f$ is one-to-one.*

*Further, if $f$ is tempered in the sense that there exists an integer $m$ such that $\displaystyle\int_{\mathbb{R}^d}\frac{|f(x)|}{(1+|x|)^m}\,\mathrm{d}x<+\infty$ then $T_f\in\mathcal{S}'(\mathbb{R}^d)$.*

We leave as an exercice to extend the previous lemma to $\langle T_\mu,f\rangle=\displaystyle\int_{\mathbb{R}^d}\varphi(x)\,\mathrm{d}\mu(x)$ where $\mu$ is a locally finite measure, *i.e.* $\mu\big(B(0,R)\big)<+\infty$ for every $R>0$. In this case, $T_\mu\in\mathcal{D}'(\mathbb{R}^d)$. If further the measure $\mu$ is tempered: there exists an integer $m$ such that $\displaystyle\int_{\mathbb{R}^d}\frac{\mathrm{d}\mu(x)}{(1+|x|)^m}<+\infty$, then $T_\mu\in\mathcal{S}'(\mathbb{R}^d)$.

EXAMPLE 1.5. **Dirac and Dirac comb**
We can now give two further examples. The Dirac $\delta$ "function": for $x_0\in\mathbb{R}^d$ and $\varphi\in\mathcal{C}_c(\mathbb{R}^d)$, define $\langle\delta_{x_0},\varphi\rangle=\varphi(x_0)$ and notice that this is a finite measure, thus a tempered distribution of order 0.

More generally, we can define the Dirac comb as $\sum_{k\in\mathbb{Z}}\delta_k$. More precisely, for $\varphi\in\mathcal{C}_c(\mathbb{R})$, define $\left\langle\sum_{k\in\mathbb{Z}}\delta_k,\varphi\right\rangle=\sum_{k\in\mathbb{Z}}\varphi(k)$. Note that this sum is finite so that it is well defined and if $\operatorname{supp}\varphi\subset[-N,N]$ then

$$\left|\left\langle\sum_{k\in\mathbb{Z}}\delta_k,\varphi\right\rangle\right|=\left|\sum_{k\in\mathbb{Z}}\varphi(k)\right|\leq\sum_{k=-N}^{N}|\varphi(k)|\leq(2N+1)\sup_{x\in\mathbb{R}}|\varphi(x)|.$$

In particular, $\sum_{k\in\mathbb{Z}}\delta_k$ is a locally finite measure and thus a distribution of order 0. It is actually also a tempered distribution: if $\varphi\in\mathcal{S}(\mathbb{R})$ then $|\varphi(x)|\leq(1+|x|)^{-2}\sup_{x\in\mathbb{R}}(1+|x|)^2|\varphi(x)|$ thus

$$\left|\left\langle\sum_{k\in\mathbb{Z}}\delta_k,\varphi\right\rangle\right|=\left|\sum_{k\in\mathbb{Z}}\varphi(k)\right|\leq\sum_{k\in\mathbb{Z}}\frac{1}{(1+|k|)^2}\sup_{x\in\mathbb{R}}(1+|x|)^2|\varphi(x)|.$$

DEFINITION 1.6. A distribution $T \in \mathcal{D}'(\mathbb{R}^d)$ is said to be positive and we write $T \geq 0$ if, for every $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $\varphi \geq 0$, we have $\langle T, \varphi \rangle \geq 0$.

LEMMA 1.7. *A positive distribution is of order 0.*

PROOF. Let $T$ be a positive distribution and $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Let $R > 0$ be such that $\varphi(x) = 0$ when $|x| \geq R$ and let $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ be such that $\psi(x) = 1$ if $|x| \leq R$. Let $f_\pm = \|\varphi\|_\infty \psi \pm \varphi$ so that $f_\pm \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $f_\pm \geq 0$. It follows that

$$0 \leq \langle T, f_\pm \rangle = \langle T, \|\varphi\|_\infty \psi \pm \varphi \rangle = \|\varphi\|_\infty \langle T, \psi \rangle \pm \langle T, \varphi \rangle.$$

It follows that

$$|\langle T, \varphi \rangle| \leq \langle T, \psi \rangle \|\varphi\|_\infty$$

which is exactly saying that $T$ is of order 0. Note that the "constant" $\langle T, \psi \rangle$ depends on $\psi$ which only depends on $R$. $\square$

EXAMPLE 1.8. **Principle value** Note that $x \to 1/x$ is not in $L^1_{loc}(\mathbb{R}^d)$ so that it does not fall in the scope of Example 1.3. We will now propose a substitute for it.

Let us define the following, for $\varphi \in \mathcal{S}(\mathbb{R})$, let

$$\langle \mathrm{vp}\, \frac{1}{x}, \varphi \rangle = \lim_{\varepsilon \to 0} \int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x} \mathrm{d}x.$$

It is important to understand that this limit is *not* $\displaystyle\int_{\mathbb{R}} \frac{\varphi(x)}{x} \mathrm{d}x$ as this integral is divergent when $\varphi(0) \neq 0$. For the limit to exist, we will use in a crucial way that we are integrating over a symetric set $(-\infty, -\varepsilon] \sup [\varepsilon, +\infty)$. We then write

$$\int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x} \mathrm{d}x = \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x)}{x} \mathrm{d}x + \int_{|x| \geq 1} \frac{\varphi(x)}{x} \mathrm{d}x.$$

For the second integral, write $\varphi(x) = x^{-1} x \varphi(x)$ and use the fact that, as $\varphi \in \mathcal{S}(\mathbb{R})$, $x\varphi$ is bounded. Then

$$\left| \int_{|x| \geq 1} \frac{\varphi(x)}{x} \mathrm{d}x \right| \leq \int_{|x| \geq 1} \frac{\mathrm{d}x}{x^2} \sup_{x \in \mathbb{R}} |x \varphi(x)| = 2 \sup_{x \in \mathbb{R}} |x \varphi(x)|.$$

For the first integral, we use the fact that

$$\int_{\varepsilon \leq |x| \leq 1} \frac{\mathrm{d}x}{x} = \int_{-1}^{-\varepsilon} \frac{\mathrm{d}x}{x} + \int_{\varepsilon}^{1} \frac{\mathrm{d}x}{x} = 0.$$

Note that we are integrating an odd function over a symetric interval. It follows that

$$\begin{aligned} \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x)}{x} \mathrm{d}x &= \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x)}{x} \mathrm{d}x - \varphi(0) \int_{\varepsilon \leq |x| \leq 1} \frac{\mathrm{d}x}{x} \\ &= \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x) - \varphi(0)}{x} \mathrm{d}x. \end{aligned}$$

But, as $\varphi \in \mathcal{C}^1$, $x \to \dfrac{\varphi(x) - \varphi(0)}{x}$ is continuous. Further, according to the mean value theorem, $\left| \dfrac{\varphi(x) - \varphi(0)}{x} \right| \leq \sup_{t \in \mathbb{R}} |\varphi'(t)|$. It follows that

$$\lim_{\varepsilon \to 0} \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x)}{x} \mathrm{d}x = \int_{-1}^{1} \frac{\varphi(x) - \varphi(0)}{x} \mathrm{d}x.$$

and

$$\left|\int_{-1}^{1} \frac{\varphi(x)-\varphi(0)}{x}\mathrm{d}x\right| \leq 2\sup_{t\in\mathbb{R}}|\varphi'(t)|.$$

All together this shows that $\mathrm{vp}\,\dfrac{1}{x}$ is a well defined tempered distribution of order at most 1:

$$\langle \mathrm{vp}\,\frac{1}{x},\varphi\rangle = \int_{-1}^{1} \frac{\varphi(x)-\varphi(0)}{x}\mathrm{d}x + + \int_{|x|\geq 1}\frac{\varphi(x)}{x}\mathrm{d}x$$

ans

$$|\langle \mathrm{vp}\,\frac{1}{x},\varphi\rangle| \leq 2\big(\sup_{t\in\mathbb{R}}(1+|t|)|\varphi(t)| + \sup_{t\in\mathbb{R}}|\varphi'(t)|\big).$$

To check that the order is exactly 1, we consider functions sequence $\varphi_n \in \mathcal{C}_c^\infty(\mathbb{R})$ with $0\leq\varphi_n\leq 1$, $\operatorname{supp}\varphi_n\subset(0,2)$ (in particular $\varphi_n=0$ in a neighbourhood of 0) and $\varphi_n=1$ on $(1/n,1)$. Then

$$\langle \mathrm{vp}\,\frac{1}{x},\varphi_n\rangle = \int_0^{+\infty}\frac{\varphi_n(x)}{x}\,\mathrm{d}x \geq \int_{1/n}^{1}\frac{\varphi_n(x)}{x}\,\mathrm{d}x = \int_{1/n}^{1}\frac{1}{x}\,\mathrm{d}x = \ln n \to +\infty$$

while $\|\varphi_n\|_\infty = 1$. It follows that an inequality of the form $\langle \mathrm{vp}\,\frac{1}{x},\varphi\rangle \leq C\|\varphi\|_\infty$ can not hold for all functions $\varphi\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ with support in $[-2,2]$ and $\mathrm{vp}\,\frac{1}{x}$ is *not* of order 0.

In particular, this distribution is not of the form $T_f$ for any locally integrable function $f$.

## 2. Convergence de suites de distributions

Definition 2.1. Let $(T_n)_{n\geq 0}$ and $T$ be distributions, $T_n, T\in\mathcal{D}'(\mathbb{R}^d)$. We say that $T_n\to T$ in $\mathcal{D}'(\mathbb{R}^d)$ if, for every $\varphi\in\mathcal{C}_c^\infty(\mathbb{R}^d)$, $\langle T_n,\varphi\rangle\to\langle T,\varphi\rangle$.

Let $(T_n)_{n\geq 0}$ and $T$ be tempered distributions, $T_n, T\in\mathcal{S}'(\mathbb{R}^d)$. We say that $T_n\to T$ in $\mathcal{S}'(\mathbb{R}^d)$ if, for every $\varphi\in\mathcal{S}(\mathbb{R}^d)$, $\langle T_n,\varphi\rangle\to\langle T,\varphi\rangle$.

Of course, convergence in $\mathcal{S}'(\mathbb{R}^d)$ implies convergence in $\mathcal{D}'(\mathbb{R}^d)$.

Donnons quelques exemples

Example 2.2. Let $T_n=\sum_{k=-n}^n\delta_k$ and $T=\sum_{k\in\mathbb{Z}}\delta_k$, then $T_n\to T$ in $\mathcal{S}'(\mathbb{R}^d)$.

Indeed, this was already proved when we defined the Dirac comb and just amounts to saying that the series $\sum_{k\in\mathbb{Z}}\varphi(k)$ converges when $\varphi\in\mathcal{S}(\mathbb{R}^d)$. Note also that convergence in $\mathcal{D}'(\mathbb{R}^d)$ is much easier: if $\varphi\in\mathcal{C}_c(\mathbb{R}^d)$ then there exists an integer $N$ such that $\varphi(x)=0$ if $|x|\geq N+1$. But then $\langle T,\varphi\rangle=\langle T_n,\varphi\rangle$ for every $n\geq N$.

Example 2.3. **Weak convergence in $L^p$.**
Let $1<p,p'<+\infty$ with $\dfrac{1}{p}+\dfrac{1}{p'}=1$. Let $f_n, f\in L^p(\mathbb{R}^d)$ and assume that $f_n\rightharpoonup f$ *i.e.* weakly in $L^p$, that is $\displaystyle\int_{\mathbb{R}^d} f_n(x)\varphi(x)\,\mathrm{d}x\to\int_{\mathbb{R}^d} f(x)\varphi(x)\,\mathrm{d}x$ for every $\varphi\in L^{p'}(\mathbb{R}^d)$.

But $\mathcal{S}(\mathbb{R}^d)\subset L^{p'}(\mathbb{R}^d)$, it follows that $\langle T_{f_n},\varphi\rangle\to\langle T_f,\varphi\rangle$ for every $\varphi\in\mathcal{S}(\mathbb{R}^d)$, thus $T_{f_n}\to T_f$ in $\mathcal{S}'(\mathbb{R}^d)$.

Note that if $f_n\in L^p_{loc}$ converges in $L^p_{loc}$ to some $f\in L^p_{loc}$ then $f_n\to f$ in the sense of distributions *i.e.* in $\mathcal{D}'(\mathbb{R}^d)$. This is because convergence in $L^p_{loc}$ means that, for every $R>0$, $f_n\mathbf{1}_{B(0,R)}$ converges to $f\mathbf{1}_{B(0,R)}$ strongly in $L^p$ thus weakly, thus in $\mathcal{D}'(\mathbb{R}^d)$. But then, if $\varphi\in\mathcal{C}_c^\infty(\mathbb{R}^d)$, there exists $R>0$ such that $\operatorname{supp}\varphi\subset B(0,R)$. It follows that

$$\langle f_n,\varphi\rangle = \big\langle f_n\mathbf{1}_{B(0,R)},\varphi\big\rangle \to \big\langle f\mathbf{1}_{B(0,R)},\varphi\big\rangle = \langle f,\varphi\rangle.$$

EXAMPLE 2.4. **Approximation of unity** Let $g \in \mathcal{S}(\mathbb{R}^d)$ be a function with $\int_{\mathbb{R}^d} g(x)\,\mathrm{d}x = 1$ and let $g_n = n^d g(nx)$. We have seen in Theorem 6.10 that, if $\varphi \in \mathcal{S}(\mathbb{R}^d) \subset \mathcal{C}_0(\mathbb{R}^d)$ and $\check{\varphi}(x) = \varphi(-x)$ then $g_n * \check{\varphi} \to \check{\varphi}$ uniformly, in particular $g_n * \check{\varphi}(0) \to \check{\varphi}(0) = \varphi(0) = \langle \delta_0, \varphi\rangle$. But

$$g_n * \check{\varphi}(0) = \int_{\mathbb{R}^d} g_n(y)\check{\varphi}(0-y)\,\mathrm{d}y = \int_{\mathbb{R}^d} g_n(y)\varphi(y)\,\mathrm{d}y = \langle T_{g_n}, \varphi\rangle.$$

Thus $T_{g_n} \to \delta_0$ which is usually denoted by $g_n \to \delta_0$ in the sense of (tempered) distributions.

Note that $|g_n(x)| = n^d|g(nx)| \leq Cn^d(1+|nx|)^{-d-1} \to 0$ for every $x \neq 0$. So almost everywhere convergence does not imply convergence in the sense of distributions.

According to the remark following the proof of Theorem 6.10, the result stays valid for $g \in L^1(\mathbb{R}^d)$.

EXAMPLE 2.5. Let $\omega \in \mathbb{R}^d$ with $|\omega| = 1$. Consider the functions $f_n$ given on $\mathbb{R}^d$ by $f_n(x) = e^{-2i\pi\langle x, n\omega\rangle}$ and $T_n = T_{f_n}$. Then, for $\varphi \in \mathcal{S}(\mathbb{R}^n)$, $\langle T_n, \varphi\rangle = \displaystyle\int_{\mathbb{R}^d} \varphi(x)e^{-2i\pi\langle x, n\omega\rangle}\,\mathrm{d}x = \widehat{\varphi}(n\omega) \to 0$ according to the Riemann-Lebesgue Lemma. Thus $T_n \to 0$ in the sense of distributions.

Finally, we will accept without proof the following result (the proof relies on a refined verion of Banach-Steinhaus's theorem adapted to semi-norms and is beyond the scope of this course)

THEOREM 2.6. *Let $(T_n)$ be a sequence in $\mathcal{D}'(\mathbb{R}^d)$. Assume that, for every $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, $\langle T_n, \varphi\rangle$ has a limit. Define $\langle T, \varphi\rangle = \lim_{n\to+\infty} \langle T_n, \varphi\rangle$. Then $T \in \mathcal{D}'(\mathbb{R}^d)$.*

*Let $(T_n)$ be a sequence in $\mathcal{S}'(\mathbb{R}^d)$. Assume that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$, $\langle T_n, \varphi\rangle$ has a limit. Define $\langle T, \varphi\rangle = \lim_{n\to+\infty} \langle T_n, \varphi\rangle$. Then $T \in \mathcal{S}'(\mathbb{R}^d)$.*

## 3. Operations on distributions

Recall that, for $a, \omega \in \mathbb{R}^d$, $\lambda > 0$, $A \in GL_n(\mathbb{R}^d)$ (a $d \times d$ invertible matrix) and $f$ a function on $\mathbb{R}^d$, we defined new functions on $\mathbb{R}^d$

$$\tau_a f(x) = f(x-a),\ M_\omega f(x) = e^{-2i\pi\langle\omega, x\rangle} f(x),\ \delta_\lambda f(x) = f(\lambda x),\ \Delta_A f(x) = f(A^{-1}x).$$

Also, if $f$ is locally integrable, then $f$ is uniquely determined by $T_f$, that is, by $\langle T_f, \varphi\rangle$ for all $\varphi$. Thus, for instance $\tau_a f$ is determined by

$$\langle T_{\tau_a f}, \varphi\rangle = \int_{\mathbb{R}^d} f(x-a)\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(y)\varphi(y+a)\,\mathrm{d}y = \langle T_f, \tau_{-a} f\rangle.$$

This can then be used as definition for $\tau_a T$ for any distribution $T$. A similar reasoning applies the the 3 other transforms and leads to the following

DEFINITION 3.1. For $a, \omega \in \mathbb{R}^d$, $\lambda > 0$, $A \in GL_n(\mathbb{R}^d)$ and $T \in \mathcal{D}'(\mathbb{R}^d)$ –resp. $T \in \mathcal{S}'(\mathbb{R}^d)$—, define

– $\tau_a T$: $\langle \tau_a T, \varphi\rangle = \langle T, \tau_{-a}\varphi\rangle$ for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ –resp. $\varphi \in \mathcal{S}(\mathbb{R}^d)$;
– $M_\omega T$: $\langle M_\omega T, \varphi\rangle = \langle T, M_\omega\varphi\rangle$ for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ –resp. $\varphi \in \mathcal{S}(\mathbb{R}^d)$;
– $\delta_\lambda T$: $\langle \delta_\lambda T, \varphi\rangle = \lambda^{-d}\big\langle T, \delta_{1/\lambda}\varphi\big\rangle$ –resp. $\varphi \in \mathcal{S}(\mathbb{R}^d)$;
– $\Delta_A T$:$\langle \delta_A T, \varphi\rangle = \det(A)^{-1}\langle T, \delta_{A^{-1}}\varphi\rangle$ –resp. $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

Then $\tau_a T$, $M_\omega T$, $\delta_\lambda T$ and $\Delta_A T$ are in $\mathcal{D}'(\mathbb{R}^d)$ –resp. $\mathcal{S}'(\mathbb{R}^d)$– and have same order as $T$.

We leave as an exercice that the notation is coherent for all four operations when $T = T_f$. Also, we leave as an exercice that those operation indeed lead to new distributions.

Finally, if $b \in \mathcal{C}^\infty$ such that $\partial^\alpha b$ is bounded for every $\alpha \in \mathbb{N}^d$, and $T \in \mathcal{D}'(\mathbb{R}^d)$ or $\mathcal{S}'(\mathbb{R}^d)$ then $M_b T$ defined as $\langle M_b T, \varphi\rangle = \langle T, b\varphi\rangle$ also defines a distribution of same order as $T$.

DEFINITION 3.2. A distribution is said to be homogeneous if

EXAMPLE 3.3. $\delta_0$ is homogeneous of degree

**3.1. Elementary operations.**

**3.2. Differentiation.**

DEFINITION 3.4. Differentiation

LEMMA 3.5. *Let $T \in \mathcal{D}'(\mathbb{R})$ and $T_h = \frac{T - \tau_h T}{h}$. Then $T_h \to T'$ in $\mathcal{D}'(\mathbb{R}^d)$ when $h \to 0$. If $T \in \mathcal{S}'(\mathbb{R})$ then $T_h \to T'$ in $\mathcal{S}'(\mathbb{R}^d)$ when $h \to 0$.*

PROOF. □

!

LEMMA 3.6. *If $T$ is homogeneous of degree $\kappa$ then $\partial^\alpha T$ is homogeneous of degree $\kappa + |\alpha|$*