Faculty of Mathematics
Institute for Analysis
JProf. Dr. Xian Liao

Englerstrasse 2, Geb. 20.30
76131 Karlsruhe

Telefon:	+49 (0) 721/608-42616
Fax:	+49 (0) 721/608-46530
E-Mail:	xian.liao@kit.edu
Web:	www.math.kit.edu/iana2/~liao
Datum:	February 17, 2023

Lecture notes for the lecture "Functional Analysis" (Lecture 0104800, Winter Semester 2022/2023) [1] [2]

Instructor:
JProf. Xian Liao (xian.liao@kit.edu)

Teaching assistant:
Ms. Rebekka Zimmermann (rebekka.zimmermann@kit.edu)

Time (weekly hours 4+2):

- Monday & Friday 11:30-13:00 (lecture)
- Wednesday 08:00-09:30 (problem class)

Place:
Seminar Room 1.067 (Mathematical Building 20.30)

Exam:
Writing exam (28.02.2023, 11am-13pm, Gaede Hörsaal)

[1] Comments are welcome to be sent to me by email.
[2] Updated version can be found here.

Contents

Main references:

- H. W. Alt, Linear functional analysis, Springer, 2016.
- H. Koch, PDE and functional analysis, lecture notes, WS2016/2017.
- P. Kunstmann, Functional analysis, lecture notes, WS2018.
- R. Schnaubelt, Functional analysis, lecture notes, WS2014.
- D. Werner, Funktionalanalysis, Springer, 2011.

[24.10.2022]

1 Introduction

In the Analysis courses we have dealt with real/complex-valued functions $f = f(x) : \mathbb{R}^n \supset \Omega \mapsto \mathbb{K}$, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, which are continuous in the sense that for any $x \in \Omega$ and $\varepsilon > 0$, there exists $\delta > 0$ such that $B_\delta^{\mathbb{R}^n}(x) \subset \Omega$ and

$$f(B_\delta^{\mathbb{R}^n}(x)) \subset B_\varepsilon^{\mathbb{K}}(f(x)).$$

Here the (open) ball $B_r(y)$, $r > 0$, $y \in \mathbb{R}^n$ denotes the set of points whose distance with y are smaller than r, either in $\mathbb{R}^n$ or in $\mathbb{K}$. From now on we take an abstract viewpoint and view functions f as points in a functional space X:

$$f \in X = \{f(x) \mid f \text{ satisfies some specific properties}\}.$$

In the lecture of Functional Analysis we are going to define and study various functional spaces, as well as the linear operators between them. In this introduction part we will introduce Banach spaces and give some intuitive examples, and then we will discuss Hilbert spaces, Lebesgue spaces, distribution spaces, Sobolev spaces, etc. Functional analysis is built on the structure of analysis and linear algebra, and it finds its applications to vast areas of mathematics including probability, partial differential equations and numerical analysis.
Before introducing Banach spaces, let us recall/introduce some abstract concepts.

1. Topological space. A system τ of subsets of a set X defines a topology in X if

- $\varnothing \in \tau$, $X \in \tau$;
- If $\tau' \subset \tau$, then $\bigcup_{U \in \tau'} U \in \tau$;
- If $U_1, U_2 \in \tau$, then $U_1 \bigcap U_2 \in \tau$.

The sets in τ are called the open sets of the topological space (X, τ). We will often omit τ and simply refer to X as a topological space.

A topological space is called Hausdorf space, if in addition the following separation axiom holds

- For any two points $x_1, x_2 \in X$ with $x_1 \neq x_2$, there exist two open sets $U_1, U_2 \in \tau$ such that $x_1 \in U_1$, $x_2 \in U_2$ with $U_1 \cap U_2 = \varnothing$.

A function $f : X \mapsto Y$ defined from a topological space X to another topological space Y is called continuous if the inverse image of every open set is open.

2. Metric space. A set X is called a metric space if there is a (distance) function $d : X \times X \mapsto [0, \infty)$ such that

   - $d(x, y) \geqslant 0$ and $d(x, y) = 0$ iff $x = y$;
   - $d(x, y) = d(y, x)$, $\forall x, y \in X$;
   - $d(x, z) \leqslant d(x, y) + d(y, z)$, $\forall x, y, z \in X$ (triangle inequality).

   A metric space is called complete if each Cauchy sequence in (X, d) converges in X. A sequence $(x_k)_{k \in \mathbb{N}}$ is called a Cauchy sequence in (X, d) if $d(x_k, x_l) \to 0$ for $k, l \to \infty$. A sequence $(x_k)_{k \in \mathbb{N}}$ is said to converge to x in (X, d) if $d(x_k, x) \to 0$ for $k \to \infty$.

3. Induced topology by a metric. Let X be a metric space and let $B_r(x) = \{y \in X \,|\, d(x, y) < r\}$ denote the ball centered at $x \in X$ with radius $r > 0$. For $A \subset X$ we define its interior as

   $$\mathring{A} = \{x \in X \,|\, B_\varepsilon(x) \subset A \text{ for some } \varepsilon > 0\} (\subset A)$$

   and its closure as

   $$\bar{A} = \{x \in X \,|\, B_\varepsilon(x) \cap A \neq \varnothing \text{ for all } \varepsilon > 0\} (\supset A).$$

   Then (X, τ), $\tau = \{A \subset X \,|\, \mathring{A} = A\}$ is a topology space, and this topology is induced by the metric. In particular all the balls of form $B_r(x)$ are open sets, and (X, τ) is a Hausdorff space, since for all $x, y \in X$ with $x \neq y$ and $d := d(x, y) > 0$, the two open sets $B_{\frac{d}{2}}(x)$ and $B_{\frac{d}{2}}(y)$ are disjoint.

A map $f : X \mapsto Y$ between two metric spaces is continuous if for any ball $B^Y_\varepsilon(f(x))$, there exists a ball $B^X_\delta(x)$ such that $f(B^X_\delta(x)) \subset B^Y_\varepsilon(f(x))$.

1.1 Normed spaces

We will focus on normed spaces, the most important class of topological vector spaces.

Definition 1.1 (Normed spaces)**.** *Let X be a $\mathbb{K}$ vector space. A map $\|.\| : X \to [0,\infty)$ is called norm if*

- $\|x\| = 0 \quad \iff \quad x = 0$*;*
- $\|x+y\| \leqslant \|x\| + \|y\|$, $\forall x, y \in X$ *(triangle inequality);*
- $\|\lambda x\| = |\lambda| \|x\|$, $\forall \lambda \in \mathbb{K}$, $x \in X$.

It is called normed space.
If (X,d) is a complete metric space, with the metric defined by $d(x,y) = \|x-y\|$, then X is called a Banach space.

Remark 1.2. *It is straightforward to check that $d(x,y) = \|x-y\|$ is a distance function defined in X. But not all the metrics are induced by the norms, e.g. the non-vector metric spaces.*

Example 1.3 (Normed spaces)**. *Exercise.***

1. *$\mathbb{R}^n$ and $\mathbb{C}^n$ equipped with the Euclidean norm are real resp. complex Banach spaces.*
2. *Let X be a set. The space of bounded functions $\mathbb{B}(X)$ equipped with the supremum norm is a Banach space.*
3. *Let (X,d) be a metric space. The space of bounded continuous functions $C_b(X)$ equipped with the supremum norm is a Banach space, or more precisely a closed sub vector space of $\mathbb{B}(X)$.*
4. *The space $C_0(\mathbb{R}^n) \subset C_b(\mathbb{R}^n)$ of functions converging to 0 at ∞ is a Banach space. Similarly the space c_0 of sequences converging to 0 equipped with the sup norm is a Banach space.*
5. *Let $U \subset \mathbb{R}^d$ be open. $C^k_b(U)$ is the vector space of k times differentiable functions on U which are together with there derivatives bounded. The norm*
$$\|u\|_{C^k(U)} = \max{}_{|\alpha| \leqslant k} \|\partial^\alpha u\|_{\mathbb{B}(U)}$$
turns C^k_b into a Banach space.

6. *Let $U \subset \mathbb{C}$ be open. The space of bounded holomorphic functions $H^\infty(U)$ is a Banach space when equipped with the supremum norm.*

A closed subvector space of a Banach space is also a Banach space **(Exercise)**:

Lemma 1.4. *Let X be a Banach space, and $U \subset X$ be a closed subvector space. Then U is a Banach space.*
Furthermore, X/U is a vector space,

$$\|\tilde{x}\|_{X/U} = \inf_{y \in U} \|y - x\|$$

defines a norm (here $\tilde{x}$ is the equivalence class of x) and X/U is a Banach space.

1.2 Sequence spaces

We study now sequence spaces as examples.

Definition 1.5 (Sequence spaces)**.** *Let $1 \leqslant p \leqslant \infty$. A $\mathbb{K}$ sequence $(x_j)_{j \in \mathbb{N}}$ is called p summable if*

$$\|(x_j)\|_{l^p} < \infty,$$

where the expression $\|(x_j)\|_{l^p}$ is given by

$$\|(x_j)\|_{l^p} := \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \quad \text{if } p < \infty, \qquad \|(x_j)\|_{l^\infty} := \sup{}_{j \in \mathbb{N}} |x_j| \quad \text{if } p = \infty.$$

The set of all p summable sequences is denoted by $l^p(\mathbb{N}) = l^p$.

Lemma 1.6 (Hölder's inequality)**.** *If $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leqslant p, q \leqslant \infty$ and $(x_j) \in l^p$, $(y_j) \in l^q$, then $(x_j y_j)$ is summable and Hölder's inequality holds:*

$$\left| \sum_{j=1}^{\infty} x_j y_j \right| \leqslant \sum_{j=1}^{\infty} |x_j y_j| \leqslant \|(x_j)\|_{l^p} \|(y_j)\|_{l^q}.$$

Proof. It holds obviously if $p = 1$ or $q = 1$.
For $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p, q < \infty$, we recall Young's inequality

$$|xy| \leqslant \frac{1}{p}|x|^p + \frac{1}{q}|y|^q$$

for all $x, y \in \mathbb{R}$. Without loss of generality we assume $x, y > 0$ and this can be proven by searching the maximum of

$$x \to xy - \frac{1}{p}x^p$$

for $y > 0$ which is attained at $x_0 = y^{1/(p-1)}$:

$$x_0 y - \frac{1}{p}x_0^p = \frac{p-1}{p}y^{\frac{p}{p-1}} = \frac{1}{q}y^q.$$

As a consequence

$$\sum_j |x_j y_j| \leqslant \sum_j \frac{1}{p}|x_j|^p + \frac{1}{q}|y_j|^q = \frac{1}{p}\|(x_j)\|_{l^p}^p + \frac{1}{q}\|(y_j)\|_{l^q}^q$$

and we obtain Hölder's inequality

$$\begin{aligned}\sum_j |x_j y_j| =& \|(x_k)\|_{l^p}\|(y_k)\|_{l^q} \sum_j |x_j| \|(x_k)\|_{l^p}^{-1} |y_j| \|(y_k)\|_{l^q}^{-1} \\ \leqslant& \|(x_k)\|_{l^p}\|(y_k)\|_{l^q}(\frac{1}{p} + \frac{1}{q}) \\ =& \|(x_j)\|_{l^p}\|(y_j)\|_{l^q}.\end{aligned}$$

□

[24.10.2022]

[28.10.2022]

Theorem 1.7. *For all $1 \leqslant p \leqslant \infty$, $(l^p, \|\cdot\|_{l^p})$ are Banach spaces.*

Remark 1.8. *We may replace $\mathbb{N}$ by $\mathbb{Z}$, by a finite set, or even an arbitrary set. Then $l^\infty(X) = \mathbb{B}(X)$. The triangle inequality is called Minkowski inequality.*

Proof. $l^\infty(\mathbb{N}) = \mathbb{B}(\mathbb{N})$ is showed to be Banach space in the exercise.
We show first the triangle inequality. It is obvious if $p = 1$. For $1 < p < \infty$

$$\begin{aligned}\sum_{j=1}^\infty |x_j + y_j|^p =& \sum_{j=1}^\infty |x_j + y_j|^{p-1}|x_j + y_j| \\ \leqslant& \sum_{j=1}^\infty |x_j + y_j|^{p-1}|x_j| + \sum_{j=1}^\infty |x_j + y_j|^{p-1}|y_j| \\ \leqslant& \|(|x_j + y_j|^{p-1})\|_{l^q}\Big(\|(x_j)\|_{l^p} + \|(y_j)\|_{l^p}\Big), \quad \frac{1}{p} + \frac{1}{q} = 1 \\ =& \|(x_j + y_j)\|_{l^p}^{p-1}\Big(\|(x)_j\|_{l^p} + \|(y_j)\|_{l^p}\Big),\end{aligned}$$

and hence

$$\|(x_j + y_j)\|_{l^p} \leqslant \|(x_j)\|_{l^p} + \|(y_j)\|_{l^p},$$

provided we can divide both sides by $\|(x_j + y_j)\|_{l^p}$. There is nothing to show if this quantity is 0. It is finite whenever we sum over a finite number of indices, and hence a limit argument completes the triangle inequality.

It is easy to check that l^p is a vector space. One easily sees that $\|(x_j)\|_{l^p} = 0$ implies $(x_j) = 0$ and

$$\|(\lambda x_j)\|_{l^p} = |\lambda| \|(x_j)\|_{l^p}.$$

Thus the spaces l^p are normed vector spaces.

Now suppose that $x_n = (x_{n,j})$ is a Cauchy sequence in l^p: $\|(x_{n,j} - x_{m,j})\|_{l^p} \to 0$ as $n, m \to \infty$. Then for every j, $n \to x_{n,j}$ is a Cauchy sequence in $\mathbb{K}$. Let $y_j = \lim_{n\to\infty} x_{n,j}$ and $y = (y_j)$. Then, for every $m > 1$ (assuming $p < \infty$)

$$\begin{aligned}
\|y - x_m\|_{l^p}^p &= \lim_{N\to\infty} \sum_{j=1}^{N} |y_j - x_{m,j}|^p \\
&= \lim_{N\to\infty} \lim_{n\to\infty} \sum_{j=1}^{N} |x_{n,j} - x_{m,j}|^p \\
&\leqslant \lim_{n\to\infty} \|x_n - x_m\|_{l^p}^p \\
&\to \quad 0 \qquad \text{as } m \to \infty.
\end{aligned}$$

$y = (y_j)$ is the limit of the Cauchy sequence $(x_{n,j})$ in l^p. □

Lemma 1.9. *Let X and Y be normed spaces. Their direct sum $X \oplus Y (= X \times Y)$ is a vector space. If $1 \leqslant p \leqslant \infty$ then*

$$\|(x, y)\|_p = \|(|x|_X, |y|_Y)\|_{l^p}$$

defines a norm with which $X \oplus Y$ becomes a Banach space.

Proof. Exercise. □

1.3 Continuous linear maps

We can define continuous maps between two topological spaces or two metric spaces. Between two normed spaces we can define continuous linear maps.

Definition 1.10 (Continuous linear maps)**.** *Let X and Y be normed spaces. We define $L(X, Y)$ as the set of all continuous linear maps from X to Y.*

Remark 1.11. *Let X be a normed space, then addition, scalar multiplication and the map to the norm are continuous. But the map to the norm is not linear.* ***(Exercise)***
For two normed spaces X, Y, $L(X, y)$ is a vector space with the obvious addition and multiplication. ***(Exercise)***

Theorem 1.12. *Let X, Y be two normed spaces.*
For any continuous linear map $T : X \to Y$,

$$\|T\|_{X\to Y} := \sup_{\|x\|_X \leqslant 1} \|T(x)\|_Y < \infty$$

and $\|.\|_{X\to Y}$ defines a norm on $L(X, Y)$.
For any linear operator $S : X \to Y$, S is continuous if and only if $\|S\|_{X\to Y}$ is finite.
If Y is a Banach space, then $L(X, Y)$ is a Banach space.

Remark 1.13. *For the linear map $T : X \mapsto Y$,*

- *$\|Tx\|_Y \leqslant \|T\|_{X\mapsto Y}\|x\|_X$ for all $x \in X$;*
- *if $\|Tx\|_Y \leqslant a\|x\|_X$ for all $x \in X$ and some $a < \infty$, then $\|T\|_{X\mapsto Y} \leqslant a$.*

Proof. Let $T : X \to Y$ be a continuous linear map. Then $T0 = 0$. For any ball $B_1(0)$ in Y, there exists $\delta > 0$ so that

$$\|Tx\|_Y < 1 \qquad \text{if } \|x\|_X < \delta.$$

In particular, if $x \in X$, $x \neq 0$, then $\left\|\frac{(\frac{1}{2}\delta)x}{\|x\|_X}\right\|_X = \frac{1}{2}\delta < \delta$ and

$$\|Tx\|_Y = \frac{\|x\|_X}{\frac{1}{2}\delta}\left\|T\frac{\frac{1}{2}\delta x}{\|x\|_X}\right\|_Y < \frac{\|x\|_X}{\frac{1}{2}\delta}$$

and thus

$$\|T\|_{X\to Y} \leqslant 2\delta^{-1}.$$

It is straightforward to check that $\|\|_{X\mapsto Y}$ defines a norm in $L(X, Y)$.
Vice versa: Let $S : X \to Y$ be linear so that $\|S\|_{X\to Y} < \infty$. If $\|S\|_{X\mapsto Y} = 0$, then $S = 0$ and it is obviously continuous. If $\|S\|_{X\mapsto Y} > 0$, then for any $\varepsilon > 0$ we choose $\delta = \varepsilon/\|S\|_{X\to Y}$. Then for any $y \in B^X_\delta(x)$, we have $Sy \in B^Y_\varepsilon(Sx)$:

$$\|Sx - Sy\|_Y = \|S(x - y)\|_Y \leqslant \|S\|_{X\to Y}\|x - y\|_X < \varepsilon.$$

In particular S is uniformly continuous.
Now assume that Y is a Banach space. The idea of the proof of the completeness of $L(X, Y)$ is the same as in the proof of Theorem 1.7 above. Let

$T_n \in L(X,Y)$ be a Cauchy sequence. For any x, $T_n x$ is a Cauchy sequence in Y since

$$\|T_n x - T_m x\|_Y \leqslant \|T_n - T_m\|_{X\to Y}\|x\|_X.$$

Let

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

The convergence is uniform on bounded sets, and hence the limit T is continuous and in $L(X,Y)$. Moreover

$$\begin{aligned}\|T - T_n\|_{X\to Y} = \sup{}_{\|x\|_X\leqslant 1}\|(T-T_n)x\|_Y &= \sup{}_{\|x\|_X\leqslant 1}\limsup_{m\to\infty}\|(T_m - T_n)x\|_Y \\ &\leqslant \limsup_{m\to\infty}\|T_m - T_n\|_{X\to Y} \to 0\end{aligned}$$

as $n \to \infty$. Here we used continuity of addition and the map to the norm. □

[28.10.2022]

[31.10.2022]

Definition 1.14. *Two norms $\|.\|$ and $|.|$ on a normed space X are called equivalent, if there exits $C \geqslant 1$ so that for all $x \in X$*

$$C^{-1}\|x\| \leqslant |x| \leqslant C\|x\|.$$

Remark 1.15. *If the two norms $\|\|_1$ and $\|\|_2$ as well as the two norms $\|\|_2$ and $\|\|_3$ are equivalent on a normed space X, then the two norms $\|\|_1$ and $\|\|_3$ are also equivalent.*

Theorem 1.16. *All norms on finite dimensional normed spaces are equivalent. Finite dimensional normed spaces are Banach spaces.*

Proof. We first prove the case $X = \mathbb{K}^d$. Let $|.|$ be the Euclidean norm on $\mathbb{K}^d$ and $\|.\|$ a second norm. Let $\{e_j\}_{j=1,\cdots,d}$ be the standard basis. Then

$$\Big\|\sum_{j=1}^d a_j e_j\Big\| \leqslant \sum_{j=1}^d |a_j| \max{}_k \|e_k\| \leqslant (\sqrt{d}\max{}_k\|e_k\|)\Big|\sum_{j=1}^d a_j e_j\Big|.$$

Thus $v \to \|.\|$ is continuous with respect to $|.|$. On the Euclidean unit sphere (which is compact), the continuous function $\|\cdot\| : \mathbb{K}^d \mapsto (0,\infty)$ attains its infimum, which is positive and we denote it by λ^{-1}. Then

$$|v| = |v||v/|v|| \leqslant \lambda|v|\|v/|v|\| = \lambda\|v\|.$$

The two inequalities imply the equivalence of the norms $\|.\|$ and $|.|$ by choosing

$$C = \max\left\{\sqrt{d}\max\|e_k\|, \lambda, 1\right\}.$$

Thus every norm on $\mathbb{K}^d$ is equivalent to the Euclidean norm, and any two norms are equivalent. A Cauchy sequence $v_m = (v_{m,j})$ with respect to $\|.\|$ is also a Cauchy sequence with respect to the Euclidean norm, hence it converges to a vector v with respect to $|\cdot|$, and hence also respect to $\|\cdot\|$. This proves the claim for $\mathbb{K}^d$.
Now let X be a $\mathbb{K}$-vector spaces of dimension d. Then there is a basis of d vectors $\{v_j\}_{j=1,\cdots,d}$, and a bijective linear map $\phi : \mathbb{K}^d \mapsto X$ maps $\sum_{j=1}^d a_j e_j$ to $\sum_{j=1}^d a_j v_j$. If $\|.\|_X$ is a norm on X then $x \to \|\phi(x)\|_X$ is a norm on $\mathbb{K}^d$. Thus the first part follows. Since $\phi(x_n)$ is a Cauchy sequence with respect to $\|.\|_X$ iff (x_n) is a Cauchy sequence in $\mathbb{K}^d$ with respect to the second metric, and one converges iff the second converges.This completes the proof. □

Remark 1.17. *The norms on infinite dimensional normed spaces are not necessarily equivalent. For example, the norm $\|\|_{\mathbb{B}(U)}$ and the norm $\|\|_{C^1(U)}$ are not equivalent on $C_b^1(U)$.*

Definition 1.18 (Dual space)**.** *Let X be a normed space. We define the dual (Banach) space as $X^* = L(X, \mathbb{K})$.*

Example 1.19. *Let $X = \mathbb{R}^d$ with the Euclidean norm. For any $y \in \mathbb{R}^d$, we can define a continuous linear map*

$$T_y : \mathbb{R}^d \mapsto \mathbb{R} \ \ \text{via } x \mapsto \sum_{j=1}^{d} x_j y_j.$$

The linear map

$$\phi : \mathbb{R}^d \ni y \to T_y \in (\mathbb{R}^d)^*$$

is

- *isometric: $|y| = \|T_y\|_{\mathbb{R}^d \mapsto \mathbb{R}}$;*
- *injective: $\phi^{-1}(\{0\}) = \{0\}$;*
- *surjective: For any $T : \mathbb{R}^d \mapsto \mathbb{R}$, there exists $y = (Te_1, \cdots, Te_d) \in \mathbb{R}^d$ such that $\phi(y) = T$.*

It allows to identify $\mathbb{R}^d$ and $(\mathbb{R}^d)^*$.

Question: Can we identify l^q and $(l^p)^*$, $\frac{1}{p} + \frac{1}{q} = 1$? In next chapter we will answer it when $p = q = 2$, and we will answer the general case later.

2 Hilbert spaces

We introduce in this chapter Hilbert spaces, which are equipped with inner products and are Banach spaces with the induced norms. With the structure of inner products, we can introduce the notion of orthogonality between two elements. The Hilbert projection theorem implies the existence of the projection operator from a Hilbert space to its closed subspace. Thanks to this fact, the celebrated Riesz representation theorem identifies a Hilbert space with its dual space. As a variant, the Lax-Milgram theorem becomes powerful in showing the existence of weak solutions of some partial differential equations of elliptic type.

2.1 Definition and first properties

2.1.1 Inner products and (pre-)Hilbert spaces

Definition 2.1. *Let X be a $\mathbb{K}$ vector space. A map $\langle .,. \rangle : X \times X \to \mathbb{K}$ is called inner product if*

- $\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle, \quad \forall x_1, x_2, y \in X,\ \forall \lambda_1, \lambda_2 \in \mathbb{K}$;
- $\langle x, y \rangle = \overline{\langle y, x \rangle} \qquad \forall x, y \in X$;
- $\langle x, x \rangle \geqslant 0,\ \forall x \in X$ *and* $\langle x, x \rangle = 0$ *iff* $x = 0$.

The vector space with inner product $(X, \langle \cdot, \cdot \rangle)$ is called pre-Hilbert space.

Remark 2.2. *It is straightforward to derive from the first and second axioms that*

- $\langle x, \lambda y \rangle = \bar{\lambda} \langle x, y \rangle$;
- $\langle x + y, x + y \rangle = \langle x, x \rangle + 2\mathrm{Re}\langle x, y \rangle + \langle y, y \rangle$.

Example 2.3.

1. *Euclidean vector spaces over $\mathbb{K}$, Euclidean inner product.*
2. *Real and complex square summable sequences space $l^2(\mathbb{N})$ with $\langle (x_j), (y_j) \rangle = \sum x_j \overline{y_j}$.*
3. *Let $U \subset \mathbb{R}^n$ be an open set, $\langle f, g \rangle = \int_U f \overline{g} dx$ defines an inner product in $C_b(U)$.*

Lemma 2.4 (Cauchy-Schwarz inequality)**.** *Let X be a pre-Hilbert space. Then*

$$|\langle x, y \rangle| \leqslant (\langle x, x \rangle \langle y, y \rangle)^{\frac{1}{2}}$$

for all $x, y \in X$.

Proof. Let $x, y \in X$ and $\lambda \in \mathbb{K}$. Then

$$\begin{aligned} 0 \leqslant & \langle x - \lambda y, x - \lambda y\rangle \\ = & \langle x, x\rangle - \lambda\langle y, x\rangle - \bar{\lambda}\langle x, y\rangle + |\lambda|^2\langle y, y\rangle. \end{aligned}$$

If $y = 0$ there is nothing to show, so we assume $y \neq 0$ and define $\lambda = \frac{\langle x,y\rangle}{\langle y,y\rangle}$. Then

$$0 \leqslant \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle}$$

which implies the Cauchy-Schwarz inequality. □

Lemma 2.5. *Let X be a pre-Hilbert space. The map*

$$x \to \|x\| := \sqrt{\langle x, x\rangle}$$

defines a norm on X.

With this notation the Cauchy-Schwarz inequality becomes

$$|\langle x, y\rangle| \leqslant \|x\|\|y\|.$$

Proof. Clearly $\|x\| \geqslant 0$, $\|x\| = 0$ iff $x = 0$. Moreover

$$\|\lambda x\|^2 = |\lambda|^2\|x\|^2$$

and by the Cauchy-Schwarz inequality

$$\begin{aligned} \|x + y\|^2 = & \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \\ \leqslant & \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ = & (\|x\| + \|y\|)^2. \end{aligned}$$

□

[31.10.2022]

[04.11.2022]

Lemma 2.6. *Let X be a pre-Hilbert space. Then the inner product defines a continuous map from $X \times X$ to $\mathbb{K}$.*

Proof. Exercise. □

Definition 2.7. *A pre-Hilbert space is a Hilbert space if it is a Banach space with respect to the metric-induced norm.*

Example: $\mathbb{R}^d$, $\mathbb{C}^d$, $l^2(\mathbb{N})$ are all Hilbert space. But $C_b(U)$ is not a Hilbert space, although it is Banach space with respect to the sup-norm.

2.1.2 Parallelogram identity

It is not hard to verify that the induced norm by the inner product $\langle \cdot, \cdot \rangle$ satisfies the parallelogram identity

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{2.1}$$

for $x, y \in X$ some pre-Hilbert space. On the other side, if the norm in a normed space satisfies the parallelogram identity, we can define the inner product correspondingly.

Theorem 2.8 (Jordan von Neumann)**.** *Let X be a normed $\mathbb{K}$ vector space with norm $\|.\|$ satisfying the parallelogram identity* (2.1)*. Then*

$$\langle x, y \rangle := \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2\Big) \tag{2.2}$$

if $\mathbb{K} = \mathbb{R}$ or

$$\langle x, y \rangle := \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2 = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\Big) \tag{2.3}$$

if $\mathbb{K} = \mathbb{C}$ defines an inner product on X, such that the norm $\|.\|$ is the norm of the pre-Hilbert space. Vice versa: The norm of a pre-Hilbert space satisfies the parallelogram identity.

Proof. We begin with a real normed spaces whose norm satisfies the parallelogram identity. We define the inner product by (2.2), then

$$\langle x, y \rangle = \langle y, x \rangle.$$

Since by the parallelogram identity

$$\begin{aligned} \|x+y+z\|^2 &= 2\|x+z\|^2 + 2\|y\|^2 - \|x-y+z\|^2 \\ &= 2\|y+z\|^2 + 2\|x\|^2 - \|y-x+z\|^2, \end{aligned}$$

we derive

$$\|x+y+z\|^2 = \|x\|^2 + \|y\|^2 + \|x+z\|^2 + \|y+z\|^2 - \frac{1}{2}\|x-y+z\|^2 - \frac{1}{2}\|y-x+z\|^2,$$

and hence by changing z into $-z$,

$$\|x+y-z\|^2 = \|x\|^2 + \|y\|^2 + \|x-z\|^2 + \|y-z\|^2 - \frac{1}{2}\|x-y-z\|^2 - \frac{1}{2}\|y-x-z\|^2.$$

We arrive at

$$\begin{aligned}\langle x+y,z\rangle =&\frac{1}{4}(\|x+y+z\|^2-\|x+y-z\|^2)\\ =&\frac{1}{4}(\|x+z\|^2-\|x-z\|^2)+\frac{1}{4}(\|y+z\|^2-\|y-z\|^2)\\ =&\langle x,z\rangle+\langle y,z\rangle.\end{aligned}$$

We claim

$$\langle \lambda x,y\rangle=\lambda\langle x,y\rangle,\quad \forall x,y\in X,\quad \forall\lambda\in\mathbb{R}\,.$$

It obviously holds for $\lambda=\pm1$, and hence for all $\lambda\in\mathbb{Z}$ by the previous step. Thus it holds for all rational λ and, by continuity of the map $\lambda\to\|\lambda x\pm y\|$, for $\lambda\in\mathbb{R}$.
We complete the proof for complex Hilbert spaces: We define

$$\langle x,y\rangle=\frac{1}{4}\sum_{k=0}^{3}i^k\|x+i^ky\|^2$$

and observe that $\langle ix,y\rangle=i\langle x,y\rangle$, $\langle x,y\rangle=\overline{\langle y,x\rangle}$ by definition, $\mathrm{Re}\langle x,y\rangle$ is the previous real inner product and $\mathrm{Im}\langle x,y\rangle=\mathrm{Re}\langle x,iy\rangle$. The above argument implies $\langle x+y,z\rangle=\langle x,z\rangle+\langle y,z\rangle$. □

Remark 2.9. *We could define a Hilbert spaces as a Banach space whose norm satisfies the paralellogram identity. By an abuse of notation we call a normed space pre-Hilbert space if it satisfies the parallelogram identity.*

Corollary 2.10. *A normed space is a pre-Hilbert space if and only if all two dimensional subspaces are pre-Hilbert spaces.*

Proof. It is a pre-Hilbert space if and only if its norm satisfies the parallelogram identity which holds if and only if the parallelogram identity holds for all two dimensional subspaces. □

2.1.3 Projection and orthogonality

Recall that a set A in a metric space is called closed if $\bar{A}=A$, and is called compact if any sequence in A has a subsequence which converges to a limit in A [3]. A set B is called convex if it holds $tx+(1-t)y\in B$ for all $x,y\in B$ and all $t\in[0,1]$.

[3] Equivalently, A is compact if any open cover of A contains a finite subcover, or if A is totally bounded and complete. In particular, totally boundedness implies boundedness and completeness implies closedness.

Lemma 2.11. *Let H be a Hilbert space, $K \subset H$ compact, and $C \subset H$ closed and convex, such that C and K disjoint. Then there exist $x \in K$ and $y \in C$ so that*

$$\|x - y\| = d(C, K),$$

where the distance between two sets is defined as $d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}$.

Proof. Let $(x_j) \subset K$ and $(y_j) \subset C$ be minimizing sequences. Since K is compact there is a subsequence which we denote again by (x_j) and $x \in K$ so that $x_j \to x$. We also denote the corresponding subsequence of (y_j) again by (y_j). By the triangle inequality

$$\|x - y_j\| \to d(C, K).$$

Then

$$\begin{aligned}\|y_n - y_m\|^2 =& \|(x - y_n) - (x - y_m)\|^2 \\ =& 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - (y_n + y_m)\|^2 \\ \leqslant& 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4d^2(C, K) \\ & \to 0 \qquad \text{as } n, m \to \infty\end{aligned}$$

since by convexity $\frac{1}{2}(y_n + y_m) \in C$. Thus (y_n) is a Cauchy sequence with limit $y \in C$. Moreover

$$d(C, K) = \lim_{n\to\infty} \|x - y_n\| = \|x - y\|.$$

□

Remark 2.12 (Definition of $p(x)$ for $x \notin C$)**.** *In particular, if $K = \{x\}$ in a Hilbert space H, then the closest point in the close and convex subset C to x is unique and we denote it by $p(x)$.*
The uniqueness is a consequence of the proof of Lemma 2.11: If there are two closest points y_1, y_2, then by denoting $d := d(x, C)$,

$$\|y_1 - y_2\|^2 = 2\|x - y_1\|^2 + 2\|x - y_2\|^2 - \|2x - (y_1 + y_2)\|^2 \leqslant 2d^2 + 2d^2 - 4d^2 = 0.$$

Definition 2.13. *Let H be a pre-Hilbert space.*
We call two elements $x, y \in H$ orthogonal if $\langle x, y\rangle = 0$.
We call two subsets M, N in H orthogonal if $\langle x, y\rangle = 0$ for all $x \in M$, $y \in N$.

Remark 2.14. ***(Exercise)*** *For any subset A in a pre-Hilbert space H, we can define $A^\perp$ as*

$$A^\perp = \{y \in H \mid \langle x, y\rangle = 0 \quad \forall x \in A\}.$$

Then $A^\perp$ is a closed subspace of H.
If B is a closed subspace in a Hilbert space H, then $(B^\perp)^\perp = B$.

Theorem 2.15 (Hilbert projection theorem)**.** *Let H be a Hilbert space and $x \in H$.*

1. *Suppose that C is a closed and convex subset of H such that $x \notin C$. Then for $y \in C$,*

$$\begin{aligned} & y = p(x) \\ \Leftrightarrow\ & \mathrm{Re}\langle x - y, z - y\rangle \leqslant 0, \quad \forall z \in C. \end{aligned} \tag{2.4}$$

2. *Suppose that C is a closed subspace such that $x \notin C$. Then for $y \in C$,*

$$\begin{aligned} & y = p(x) \\ \Leftrightarrow\ & \langle x - y, z\rangle = 0, \quad \forall z \in C. \end{aligned} \tag{2.5}$$

Moreover, in this case

$$\|x\|^2 = \|x - p(x)\|^2 + \|p(x)\|^2. \tag{2.6}$$

Proof. If C is a closed and convex subset, then by the convexity of C,

$$y + t(z - y) \in C, \quad \forall y, z \in C, \quad \forall t \in [0, 1].$$

If $y = p(x)$, then

$$\|x - y\|^2 \leqslant \|x - y - t(z - y)\|^2 = \|x - y\|^2 - 2t\,\mathrm{Re}\langle x - y, z - y\rangle + t^2\|z - y\|^2,$$

which implies (2.4). If conversely (2.4) holds, then for all $z \in C$ it holds

$$\|x - z\|^2 = \|x - y - (z - y)\|^2 = \|x - y\|^2 - 2\,\mathrm{Re}\langle x - y, z - y\rangle + \|z - y\|^2 \geqslant \|x - y\|^2,$$

and $y = p(x)$ is the unique closest point in C to x (by Remark 2.12).
If C is a closed subspace, then we claim that (2.4) and (2.5) are equivalent. Obviously (2.5) implies (2.4). Conversely, in the case that C is a closed subspace, then (2.4) implies $\mathrm{Re}\,\langle x - y, z\rangle \leqslant 0$ for all $z \in C$, and hence $\mathrm{Re}\,\langle x - y, z\rangle = 0$ for all $z \in C$, and thus $\mathrm{Im}\,\langle x - y, z\rangle = 0$ for all $z \in C$, such that the orthogonality relation (2.5) holds. In particular, (2.6) follows from (2.5) with $y = z = p(x)$. ☐

An intuitive example of Theorem 2.15 is $H = \mathbb{R}^2$ with $\langle x, y\rangle = x_1 y_1 + x_2 y_2$:

1. Suppose that C is a closed convex subset and $y \in C$. We can take y as the origin, and take the coordinate of x to be $(0, x_2)$ with $x_2 > 0$. Then

$$\begin{aligned} y = p(x) &\Leftrightarrow z = (z_1, z_2), \quad z_2 \leqslant 0, \quad \forall z \in C \\ &\Leftrightarrow \operatorname{Re}\langle x - y, z - y\rangle = x_2 z_2 \leqslant 0, \quad \forall z \in C. \end{aligned}$$

2. If C is a closed subspace and is not $\{0\}$ or $\mathbb{R}^2$, then without loss of generality we can take it to be the horizontal (real) axis. Then for $y = (y_1, 0)$,

$$y = p(x) \Leftrightarrow x_1 = y_1 \Leftrightarrow \langle x - y, z\rangle = 0, \quad \forall z = (z_1, z_2) \text{ with } z_2 = 0.$$

Remark 2.16 (Projection operator)**.** *If C is a closed subspace in a Hilbert space H, then the operator*

$$P : H \mapsto H \text{ via } Px = \begin{cases} p(x) & \text{if } x \notin C, \\ x & \text{if } x \in C, \end{cases}$$

is an projection operator, i.e. $P^2 = P$.
This projection operator P is linear and bounded, with the norm $\|P\|_{H \mapsto H} = 1$ if $C \neq \{0\}$. Indeed, the linearity of P follows from

$$(2.5) \overset{y \in C}{\Leftrightarrow} y = Px.$$

We calculate

$$\|Px\|^2 = \langle Px, Px\rangle = \langle x, Px\rangle - \langle x - Px, Px\rangle = \langle x, Px\rangle \leqslant \|x\|\|Px\|,$$

which implies $\|P\|_{H \mapsto H} \leqslant 1$. Obviously $Pz = z$ on C, and hence $\|P\| = 1$ if $C \neq \{0\}$. If $C = \{0\}$, then $P = 0$ is the zero operator in $L(H, H)$.

2.2 Orthonormality

2.2.1 Orthonormal system and Bessel's inequality

Definition 2.17. *Let H be a pre-Hilbert space. We call a subset $A \subset H$ an orthogonal set if any two different vectors in A are orthogonal. If, in addition, every vector x in A is of unit norm: $\|x\| = 1$, this orthogonal set A is called an orthonormal set.*

Let $(x_j)_{j=1,\cdots,n}$ be a linearly independent vectors in a pre-Hilbert space X. Then by the Gram-Schmidt procedure we obtain an orthonormal system $(y_j)_{j=1,\cdots,n}$ such that $\langle y_k, y_l\rangle = \delta_{kl},\ k,l = 1,\cdots,n$. More precisely, one can perform the following procedure recursively

$$y_1 = \frac{1}{\|x_1\|}x_1, \quad \tilde{y}_2 = x_2 - \langle x_2, y_1\rangle y_1, \quad y_2 = \frac{1}{\|\tilde{y}_2\|}\tilde{y}_2.$$

We can take $n = \infty$.

[07.11.2022]

[11.11.2022]

Lemma 2.18 (Bessel inequality). *Let $(x_j)_{j\in\mathbb{N}}$ be an orthonormal system in a pre-Hilbert space X. Then for any finite n and $x \in X$,*

$$\|\sum_{j=1}^{n}\langle x, x_j\rangle x_j\|^2 = \sum_{j=1}^{n}|\langle x, x_j\rangle|^2 \leqslant \|x\|^2,$$

and

$$\left\|x - \sum_{j=1}^{n}\langle x, x_j\rangle x_j\right\|^2 = \|x\|^2 - \sum_{j=1}^{n}|\langle x, x_j\rangle|^2 \geqslant 0.$$

In particular, $x = \sum_{j=1}^{n}\langle x, x_j\rangle x_j$ is equivalent to $\|x\|^2 = \sum_{j=1}^{n}|\langle x, x_j\rangle|^2$.

Proof. The first equality $\|\sum_{j=1}^{n}\langle x, x_j\rangle x_j\|^2 = \sum_{j=1}^{n}|\langle x, x_j\rangle|^2$ follows from the orthonormality.
For any $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$, we have

$$\begin{aligned}\left\|x - \sum_{j=1}^{n}\lambda_j x_j\right\|^2 &= \|x\|^2 - \sum_{j=1}^{n}\langle x, x_j\rangle\bar{\lambda}_j - \sum_{j=1}^{n}\langle x_j, x\rangle\lambda_j + \sum_{j=1}^{n}|\lambda_j|^2\\ &= \|x\|^2 - \sum_{j=1}^{n}|\langle x, x_j\rangle|^2 + \sum_{j=1}^{n}|\langle x, x_j\rangle - \lambda_j|^2.\end{aligned}$$

The minimum is obtained at $(\lambda_j) = (\langle x, x_j\rangle)$:

$$\left\|x - \sum_{j=1}^{n}\langle x, x_j\rangle x_j\right\|^2 = \|x\|^2 - \sum_{j=1}^{n}|\langle x, x_j\rangle|^2 \geqslant 0.$$

□

Example 2.19 (Legendre polynomials). *Let $X = C([-1,1];\mathbb{R})$ be equipped with the inner product $\langle f,g\rangle = \int_{-1}^{1} fg dx$. All the monomials x^n in the vector space X are linearly independent. We define the Legendre polynomials by the Rodrigues' formula*

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n].$$

Show that P_n is the unique polynomial of degree n such that $P_n(1) = 1$ and $\int_{-1}^{1} x^m P_n(x)dx = 0$ for all $0 \leqslant m < n$. Show that the family of functions

$$\Big(\sqrt{\frac{2n+1}{2}}P_n(x)\Big)_{n\geqslant 0}$$

is an orthonormal system in X. ***Exercise.***

2.2.2 Orthonormal basis and Parseval's equality

Recall that a subset A of a metrix space X is called dense if its closure is X. A metric space X is called separable, if there is a countable dense subset.

Example 2.20 (Countable sets and separable metric spaces).

1. *$\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}^N$ are countable.*
2. *Subsets of countable sets are countable; the product of countable sets is countable; If X_j are countable sets then there union is countable.*
3. *$\mathbb{R}^N$ is separable since Q^N is countable and dense.*
4. *$C([0,1])$ is separable by Weierstrass approximation theorem.*
5. *$l^p(\mathbb{N})$, $1 \leqslant p < \infty$ is separable.* ***Exercise.***

Definition 2.21. *An orthonormal set $(x_j)_{j\in\mathbb{N}}$ of a Hilbert space is called orthonormal basis if*

$$\langle x, x_j\rangle = 0, \quad \forall j \in \mathbb{N} \text{ implies } x = 0. \tag{2.7}$$

Theorem 2.22. *The following properties are equivalent for a Hilbert space H which is not finite dimensional.*

(1) The space H is separable.

(2) There exists an orthonormal basis $(x_j)_{j\in\mathbb{N}}$ such that the Parseval's equality holds

$$\|x\|^2 = \sum_{j\in\mathbb{N}} |\langle x, x_j\rangle|^2, \quad \forall x \in H.$$

(3) The exists a linear surjective isometry $l^2 \to H$.

Proof. "(1) $\Rightarrow$ (2)" Suppose that H is separable. Let $(y_n)_{n\in\mathbb{N}}$ be a dense sequence and let X_N be the span of $(y_1, \cdots, y_N)$. The dimension of X_N is at most N, and we can use the Gram-Schmidt procedure to find an orthonormal basis (x_n) of X_N. We do this recursively by increasing N. This leads to a countable orthonormal sequence $(x_n)_{n\in\mathbb{N}}$ so that its span is dense in H.
For any $x \in H$, the map

$$\varphi : N \to \|x - \sum_{j=1}^{N} \langle x, x_j \rangle x_j\|^2 = \|x\|^2 - \sum_{j=1}^{N} |\langle x, x_j \rangle|^2$$

is monotonically decreasing. We observe that if $x \in C_N := \operatorname{span}(\{x_n\}_{n\leqslant N})$, then $x = \sum_{n=1}^{N} \langle x, x_n \rangle x_n$, and if $x \notin C_N$, then $\sum_{n=1}^{N} \langle x, x_n \rangle x_n = p_N x$ is the closed point in C_N to x (by the proof of Bessel inequality). Since the span of $(x_n)_{n\in\mathbb{N}}$ is dense, $\varphi(N)$ converges to 0:

$$0 \leftarrow \operatorname{dist}(x, \operatorname{span}\{(x_j)_{1\leqslant j\leqslant N}\})^2 = \|x - \sum_{j=1}^{N} \langle x, x_j \rangle x_j\|^2 = \varphi(N),$$

that is,

$$\sum_{n=1}^{N} \langle x, x_n \rangle x_n \to x \text{ in } H \text{ as } N \to \infty.$$

This in turn implies $x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$ and

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2.$$

Hence (x_n) is an orthonormal basis.
"(2) $\Rightarrow$ (3)" Now suppose that (x_n) is an orthonormal basis such that the Parseval's equality holds. We claim that the following map is well-defined

$$\phi : l^2 \ni (a_n) \to \sum_{n=1}^{\infty} a_n x_n \in H.$$

Indeed, for $M > N$, due to the orthonormality, the norm of

$$\sum_{n=1}^{M} a_n x_n - \sum_{n=1}^{N} a_n x_n = \sum_{n=N+1}^{M} a_n x_n$$

is given by $\sqrt{\sum_{n=N+1}^{M}|a_j|^2}$. This implies that the partial sums are a Cauchy sequence in H and the map $\phi : l^2 \mapsto H$ is well-defined, and is isometric (by an approximation argument):

$$\langle \phi((a_n)), \phi((b_n))\rangle_H = \langle \sum_{n=1}^{\infty} a_n x_n, \sum_{n=1}^{\infty} b_n, x_n\rangle = \sum_{n=1}^{\infty} a_n \overline{b_n} = \langle (a_n), (b_n)\rangle_{l^2(\mathbb{N})}.$$

The map is clearly linear. The surjectivity follows from Parseval's identity: For any $x \in H$ we can take $a_n = \langle x, x_n\rangle$ such that $\phi((a_n)) = x$.
"(3) $\Rightarrow$ (1)" It follows from the fact that $l^2(\mathbb{N})$ is separable. □

[11.11.2022]

[14.11.2022]

Remark 2.23.

- *An isometric (anti-)linear map ϕ between two normed spaces X, Y is injective since if $\phi(x) = 0$, then*

$$\|x\|_X = \|\phi(x)\|_Y = 0 \Rightarrow x = 0.$$

- *A Hilbert space is either isomorphic (there exists an isometric surjective linear map) to $\mathbb{R}^d$ resp. $\mathbb{C}^d$, to $l^2(\mathbb{N})$, or it is not separable.*
- *One can show that if (x_n) is an orthonormal basis, then the Parseval's identity holds. Vice versa: An orthonormal set such that Parseval's identity holds is an orthonormal basis.* ***Exercise.***

Example 2.24. ***Exercise.***

- *The space $l^2(\mathbb{R})$ with inner product*

$$\langle f, g\rangle = \sum_{x\in\mathbb{R}} f(x)\overline{g(x)}$$

is not separable.

- *Let $C([0, 2\pi]; \mathbb{C})$ be equipped with the inner product*

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f\overline{g}dx.$$

Then the system $(e^{inx})_{n\in\mathbb{Z}}$ is an orthonormal set, such that Parseval's identity holds on $C([0, 2\pi]; \mathbb{C})$.

(The Lebesgue space $L^2([0, 2\pi]; \mathbb{C})$is the completeness of $C([0, 2\pi]; \mathbb{C})$ under the norm $\sqrt{\langle\cdot,\cdot\rangle}$, which is a Hilbert space and $(e^{inx})_{n\in\mathbb{Z}}$ is an orthonormal basis on it. In particular, $\hat{f}(n) := \langle f, e^{inx}\rangle$ is called Fourier coefficient of f, and $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$ is called Fourier series of f.)

2.3 The Riesz representation theorem

By use of Hilbert projection theorem, we can prove the Riesz representation theorem for Hilbert spaces. Recall the definition of dual spaces in Definition 1.18.

Theorem 2.25 (Riesz representation theorem)**.** *Let H be a Hilbert space. Then*

$$J : H \ni x \to (y \to \langle y, x \rangle) \in H^*$$

is an antilinear isometric bijective map (isomorphism).

Proof. By the Cauchy-Schwarz inequality

$$\|J(x)\|_{H^*} = \sup_{\|y\| \leqslant 1} |\langle y, x \rangle| \leqslant \|x\|_H$$

and the map is well defined and antilinear.
Since

$$\|x\|_H \|J(x)\|_{H^*} \geqslant |(J(x))(x)| = \langle x, x \rangle = \|x\|_H^2$$

we see that

$$\|J(x)\|_{H^*} \geqslant \|x\|_H.$$

Thus J is an isometry: $\|J(x)\|_{H^*} = \|x\|_H$. In particular J is injective.
To show that J is surjective we assume that $x^* \in H^*$ and try to find x_1 so that $x^* = J(x_1)$. Let

$$N = \{y \in H : x^*(y) = 0\}.$$

Then N is a closed subspace. Let p be the orthogonal projection to N as in Theorem 2.15. If $x^* \neq 0$, then there exists $y_1 \neq 0$ in H such that $x^*(y_1) \neq 0$. We take $y_0 = \frac{y_1}{x^*(y_1)} \in H$ such that $x^*(y_0) = 1$ and define

$$x_0 = y_0 - p(y_0).$$

Then $x^*(x_0) = 1$ and for all $y \in N$ by (2.5) $\langle y, x_0 \rangle = 0$. Moreover, for any $x \in H$, $x^*(x - x^*(x)x_0) = 0$, and hence $x - x^*(x)x_0 \in N$ and by (2.5)

$$\langle x - x^*(x)x_0, x_0 \rangle = 0.$$

This implies

$$\langle x, x_0 \rangle = \langle x^*(x)x_0, x_0 \rangle = x^*(x)\|x_0\|_H^2,$$

and hence

$$x^*(x) = \left\langle x, \frac{x_0}{\|x_0\|^2} \right\rangle = J(\frac{x_0}{\|x_0\|^2})(x), \quad \text{i.e. } x^* = J(x_1) \text{ with } x_1 = \frac{x_0}{\|x_0\|^2}.$$

□

Since $l^2(\mathbb{N})$ is a Hilbert space, Riesz representation theorem ensures that we can identify $(l^2(\mathbb{N}))^*$ and $l^2(\mathbb{N})$ in the sense that there exists an isomorphism between them.
The Lax-Milgram theorem can be viewed as a variant of the Riesz representation theorem.

Theorem 2.26 (Lax-Milgram). *Let H be a Hilbert space and*

$$Q : H \times H \ni (x, y) \to Q(x, y) \in \mathbb{K}$$

- *linear in x, antilinear in y;*
- *bounded in the sense that there exists $C > 0$ such that*

$$|Q(x, y)| \leqslant C\|x\|\|y\|;$$

- *coercive in the sense that there exists $\delta > 0$ so that*

$$\operatorname{Re} Q(x, x) \geqslant \delta\|x\|^2.$$

Then there exists a unique continuous linear map $A : H \to H$ with continuous inverse A^{-1} so that

$$Q(x, y) = \langle Ax, y\rangle.$$

Moreover

$$\|A\|_{H\to H} \leqslant C, \qquad \|A^{-1}\|_{H\to H} \leqslant \delta^{-1}.$$

Proof. Exercise. □

A little on operators between Hilbert spaces By use of Riesz representation theorem, the following holds true (**Exercise**):
For any linear continuous map $T \in L(H_1, H_2)$ between two Hilbert spaces, there exists a unique $T^* \in L(H_2, H_1)$ such that

$$\langle Tx, y\rangle_{H_2} = \langle x, T^*y\rangle_{H_1}.$$

The operator T^* is called the adjoint of T, and it satisfies

$$\|T\|_{H_1\mapsto H_2} = \|T^*\|_{H_2\mapsto H_1}.$$

It is easy to see that $T^{**} = T$. The situation will be more delicate when the unbounded operators are concerned. Spectral theory is devoted to the relevant study.
A map $U \in L(H_1, H_2)$ between two Hilbert spaces is called unitary, if

- It is invertible: There exists $V \in L(H_2, H_1)$ (denoted by U^{-1}) such that $VU = 1_{H_1}$ and $UV = 1_{H_2}$;
- It preserves the inner product: $\langle Ux, Uy\rangle_{H_2} = \langle x, y\rangle_{H_1}$, $\forall x, y \in H_1$.

Obviously the inverse of a unitary operator $U \in L(H_1, H_2)$ is also its adjoint $U^{-1} = U^*$, since

$$\langle Ux, y\rangle_{H_2} = \langle Ux, UU^{-1}y\rangle_{H_2} = \langle x, U^{-1}y\rangle_{H_1}.$$

The composition of unitary operators is unitary, and in particular the set of all unitary operator $L(H) := L(H, H)$ on the Hilbert space H is a group (called unitary group $U(H)$).

[14.11.2022]

[18.11.2022]

3 Lebesgue spaces

3.1 Review of measure spaces

3.1.1 Measure spaces

Definition 3.1 (Measure space)**.** *Let X be a set. A family of subset $\mathcal{A}$ is called a $\sigma-$algebra if*

1. $\{\} \in \mathcal{A}$
2. $A \in \mathcal{A}$ *implies* $X \backslash A \in \mathcal{A}$
3. $A_n \in \mathcal{A}$ *implies* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

A map $\mu : \mathcal{A} \to [0, \infty]$ is called a measure if whenever $A_n \in \mathcal{A}$ are pairwise disjoint then

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

The triple $(X, \mathcal{A}, \mu)$ is called a measure space.
A measure space $(X, \mathcal{A}, \mu)$ is called sigma finite if there exists a sequence of measurable sets A_n of finite measure so that $X = \bigcup_{n=1}^{\infty} A_n$.
A measure μ for a measure space $(X, \mathcal{A}, \mu)$ is called complete, if the σ algebra contains every subset of a set of measure zero.

Example 3.2. *1. X a set, $\mathcal{A} = 2^X$ the set of all subsets, and $\mu(A)$ the number of elements.*

2. X a set, $\mathcal{A} = 2^X$ the set of all subsets, $x \in X$ a fixed element, and $\delta_x(A) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A, \end{cases}$ for $A \subset X$. δ_x is called the point measure or Dirac measure.

3. If (X, d) is a metric space (or (X, τ) is a topological space), then there is a smallest σ algebra containing all open sets. It is called the Borel σ algebra of X, and a Borel measure is a measure on the Borel sets.

4. $X = \mathbb{R}^n$, $\mathcal{A}$ the Borel sets, μ the Lebesgue measure restricted to the Borel sets.

5. $X = \mathbb{R}^n$, $\mathcal{A}$ the Lebesgue sets, μ the Lebesgue measure.

Recall the construction of Lebesgue measures $m = m^n$ on $\mathbb{R}^n$:

Step 1. We define the measure of a coordinate rectangle as the product of the sidelengths;

Step 2. We define the measure of a union of countably many pairwise disjoint coordinate rectangles as the sum over the measures of the rectangles;

Step 3. We define the *outer measure* m^* of a general set as the infimum of all measures of coverings by unions of disjoint coordinate rectangles;

Step 4. All the sets A satisfying Caratheodory's criterium: $m^*(B) = m^*(A \cap B) + m^*(B \cap A^c)$, $\forall B \subset \mathbb{R}^n$ form a σ algebra $\mathcal{M}$;

Step 5. The triple $(\mathbb{R}^n, \mathcal{M}, m)$ is a measure space, where $m(A) = m^*(A)$ is the Lebesgue measure, and $A \in \mathcal{M}$ is a Lebesgue measurable set.

Measures are often constructed by restricting outer measures on Caratheodory measurable sets, by the same procedure above.

Definition 3.3. • *Let X be a set. A map $\mu : X \mapsto [0, \infty]$ is called an outer measure if*

1. $\mu(\{\,\}) = 0$.

2. $A \subset B$ implies $\mu(A) \leqslant \mu(B)$.

3. $\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \leqslant \sum_{j=1}^{\infty} \mu(A_j), \quad \forall A_j \subset X$.

- *Let μ be an outer measure on X. We call a subset $A \subset X$ Caratheodory measurable if for all $B \subset X$*

$$\mu(B) = \mu(B \cap A) + \mu(B \cap (X \backslash A)).$$

- *Let (X, d) be a metric space. We call μ an outer metric measure if it is an outer measure which satisfies*

$$\mu(A \cup B) = \mu(A) + \mu(B), \quad \forall A, B \subset X \text{ with } \operatorname{dist}(A, B) > 0.$$

We state Caratheodory's theorem which constructs a measure on the Caratheodory measurable sets, without proof.

Theorem 3.4 (Caratheodory)**.** *Let μ be an outer measure on the set X. Then the Caratheodory measurable sets $\mathcal{C}$ are a σ algebra and $(X, \mathcal{C}, \mu|_{\mathcal{C}})$ is a measure space. Moreover $\mathcal{C}$ contains all sets of exterior measure 0.*
If X is a metric space and μ is a metric outer measure, then $\mathcal{C}$ contains all open sets and hence all Borel sets.

Example 3.5 (Hausdorff measure)**.** *Let (X, d) be a metric space and $s \geqslant 0$. We define for any $r > 0$ and $A \subset X$,*

$$\mathcal{H}^s_r(A) = \inf \left\{ \sum_{n=1}^{\infty} 2^{-s} \frac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)} (diam A_n)^s \,\middle|\, \bigcup_{n=1}^{\infty} A_n \supset A \text{ with } diam(A_n) < r \right\}.$$

$\mathcal{H}^s_r(A)$ is nonincreasing in r and we define

$$\mathcal{H}^s(A) = \lim_{r \to 0_+} \mathcal{H}^s_r(A).$$

Then $\mathcal{H}^s$ is a metric outer measure, and by Caratheodory theorem its restriction on Caratheodory measurable sets is a measure, which is called Hausdorff measure.

3.1.2 Measurable and integrable functions

Definition 3.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

- *A map $f : X \to \mathbb{R} \cup \{-\infty, \infty\}$ is called measurable if*

$$f^{-1}((t, \infty]) \in \mathcal{A}, \quad \forall t \in \mathbb{R}\,.$$

- *For a measurable function $f : X \to [0, \infty]$, we define its Lebesgue integral by the Riemann integral*

$$\int_X f \, \mathrm{d}\mu = \int_0^\infty \mu(f^{-1}((t, \infty]))dt \in [0, \infty],$$

and we call it integrable if $\int_X f \, \mathrm{d}\mu < \infty$.

We call a measurable function $f : X \mapsto \mathbb{R} \cup \{-\infty, \infty\}$ *integrable if* $|f|$ *is integrable, and we define*

$$\int_X f \, \mathrm{d}\mu = \int_X f^+ \, \mathrm{d}\mu - \int_X f^- \, \mathrm{d}\mu \,,$$

where $f^+ = \max\{f, 0\}$ *denotes its positive part and* $f^- = \max\{-f, 0\}$ *denotes its negative part.*

We call a function $f : X \to \overline{\mathbb{C}}$ *integrable if the real and imaginary parts are both integrable, and we define*

$$\int_X f \, \mathrm{d}\mu = \int_X (\operatorname{Re} f) \, \mathrm{d}\mu + i \int_X (\operatorname{Im} f) \, \mathrm{d}\mu \,.$$

- *Let* $1 \leqslant p < \infty$. *We call a measurable function* f p *integrable if* $|f|^p$ *is integrable and denote*

  $$\|f\|_{L^p} = \left(\int_X |f|^p d\mu \right)^{1/p} .$$

- *We call a measurable function* ∞ *integrable or essentially bounded if there is a constant* C *so that*

  $$\mu(\{x : |f(x)| > C\}) = 0.$$

  The best constant is denoted by $\|f\|_{L^\infty}$.

- *We call two measurable functions equivalent (denoted by* $f \sim g$*), if they are the same almost everywhere (i.e.* $\mu(\{x \in X | f(x) \neq g(x)\}) = 0$*).*

  We define $L^p(\mu)$ *as the set of equivalence classes of* p *integrable functions.*

[18.11.2022]

[21.11.2022]

Remark 3.7. *We notice the following facts from the definitions:*

- *For measurable functions* $f, g : X \mapsto \mathbb{R}$ *such that* $f \leqslant g$ *almost everywhere, it holds* $\int_X f \, \mathrm{d}\mu \leqslant \int_X g \, \mathrm{d}\mu$. *It follows straightforward if* $f, g : X \mapsto [0, \infty]$,

  $$\int_X f \, \mathrm{d}\mu = \int_0^\infty \mu(\{f > t\}) dt \leqslant \int_0^\infty \mu(\{g > t\}) dt = \int_X g \, \mathrm{d}\mu \,.$$

  For general $f, g : X \mapsto \mathbb{R}$ *with* $f \leqslant g$, *one notices* $f^+ \leqslant g^+$ *and* $f^- \geqslant g^-$.

- *The scale multiplication commutes with the Lebesgue integral. It holds obviously true for $\lambda > 0$ and nonnegative integrable function $h : X \mapsto [0,\infty]$ by use of change of variables:*

$$\lambda \int_X h d\mu = \lambda \int_0^\infty \mu(\{h > t\})dt = \int_0^\infty \mu(\{h > t'/\lambda\})\,\mathrm{d}t' = \int_X \lambda h \,\mathrm{d}\mu \,.$$

  Other cases follow correspondingly.

- *If $\mu(X) < \infty$, then for any integrable function $f : X \mapsto \overline{C}$ and complex number a,*

$$\int_X (f+a)\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu + a\mu(X).$$

  Indeed, it is easy to check with $f : X \mapsto [0,\infty]$ and $a \geqslant 0$,

$$\begin{aligned}\int_X (f+a)\,\mathrm{d}\mu &= \int_0^\infty \mu(\{f+a > t\})dt = \int_0^\infty \mu(\{f > t-a\})dt \\ &= \int_0^\infty \mu(\{f > t\})dt + \int_{-a}^0 \mu(\{f > t\})dt = \int_X f\,\mathrm{d}\mu + a\mu(X).\end{aligned}$$

  Other cases follow correspondingly.

- *For any measurable function $f : X \mapsto \overline{C}$, it holds*

$$\Big|\int_X f\,\mathrm{d}\mu\Big| \leqslant \int_X |f|\,\mathrm{d}\mu \,.$$

  It holds obviously for real-valued functions:

$$\Big|\int_X f\,\mathrm{d}\mu\Big| = \Big|\int_X f^+\,\mathrm{d}\mu - \int_X f^-\,\mathrm{d}\mu\Big| \leqslant \max\left\{\int_X f^+\,\mathrm{d}\mu \,, \int_X f^-\,\mathrm{d}\mu\right\} \leqslant \int_X |f|\,\mathrm{d}\mu \,.$$

  For complex-valued function $f : X \mapsto \overline{C}$, there exists a complex number λ with modulus 1 such that

$$\begin{aligned}\Big|\int_X f\,\mathrm{d}\mu\Big| &= \lambda \int_X f\,\mathrm{d}\mu = \int_X (\lambda f)\,\mathrm{d}\mu = \mathrm{Re} \int_X (\lambda f)\,\mathrm{d}\mu = \int_X \mathrm{Re}\,(\lambda f)\,\mathrm{d}\mu \\ &\leqslant \int_X |\mathrm{Re}\,(\lambda f)|\,\mathrm{d}\mu \leqslant \int_X |\lambda f|\,\mathrm{d}\mu = \int_X |f|\,\mathrm{d}\mu \,.\end{aligned}$$

  It is easy to see from the above that the equality implies $f = \lambda|f|$ for some complex number λ with modulus 1.

We recall the convergence theorems about the relation between the limit of integrals, and the integral over limits.

Theorem 3.8 (Fatou's lemma)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a measure space, and* $f_n : X \mapsto [0, \infty]$*,* $n \in \mathbb{N}$ *be measurable functions. Then*

$$\int_X \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \;\leqslant \liminf_{n\to\infty} \int_X f_n \, \mathrm{d}\mu\,.$$

If furthermore $(f_n)_{n\in\mathbb{N}}$ *converges to a measurable function* $f : X \mapsto [0, \infty]$ *almost everywhere, then*

$$\int_X f \, \mathrm{d}\mu \;\leqslant \liminf_{n\to\infty} \int_X f_n \, \mathrm{d}\mu \;\leqslant \sup{}_{n\in\mathbb{N}} \int_X f_n \, \mathrm{d}\mu\,.$$

Theorem 3.9 (Lebesgue's dominated convergence theorem)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a measure space, and* $f_n, f : X \mapsto \mathbb{R} \cup \{-\infty, \infty\}$*,* $n \in \mathbb{N}$ *be measurable functions, such that* $f_n \to f$ *almost everywhere. Let* $g : X \mapsto [0, \infty]$ *be an integrable function, such that* $|f_n| \leqslant g$ *for all* $n \in \mathbb{N}$ *almost everywhere. Then*

$$\lim_{n\to\infty} \int_X f_n \, \mathrm{d}\mu \;= \int_X f \, \mathrm{d}\mu\,.$$

Theorem 3.10 (Beppo Levi's monotone convergence theorem)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a measure space, and* $f_n : X \mapsto [0, \infty]$*,* $n \in \mathbb{N}$ *be measurable functions, such that* $f_n \leqslant f_{n+1}$ *almost everywhere. Then there exists a measurable function* $f : X \mapsto [0, \infty]$ *such that* f_n *converges to* f *almost everywhere, and*

$$\lim_{n\to\infty} \int_X f_n \, \mathrm{d}\mu \;= \int_X f \, \mathrm{d}\mu\,.$$

Theorem 3.11 (Fubini-Tonelli)**.** *Let* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{B}, \nu)$ *be* σ*-finite measure spaces,* $\mathcal{A} \times \mathcal{B}$ *the product* σ *algebra and* $\mu \times \nu$ *the product measure. Let* f *be* $\mu \times \nu$ *integrable. Then for almost of* $x \in X$*,* $y \to f(x, y)$ *is* ν *integrable,* $x \to \int_Y f(x, y) d\nu(y)$ *is* μ *integrable and*

$$\int_{X\times Y} f(x, y) d\mu \times \nu = \int_X \int_Y f(x, y) d\nu(y) d\mu(x).$$

3.2 Preliminary inequalities

In this section we review some preliminary inequalities for p-integrable functions.

3.2.1 Jensen's inequality

The convex functions $F : \mathbb{R} \mapsto \mathbb{R}$ have left and right derivatives, and hence there are at most countably many points where F is not differentiable.

Lemma 3.12. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is convex. Then both one sided derivatives exist and if $x < y$ then*

$$\frac{df}{dx}^{+}(x) \leqslant \frac{df}{dy}^{-}(y) \leqslant \frac{df}{dy}^{+}(y)$$

and for all $z \in \mathbb{R}$,

$$f(z) \geqslant \max\Big\{f(x) + \frac{df}{dx}^{+}(x)(z-x), f(x) + \frac{df}{dx}^{-}(x)(z-x)\Big\}.$$

Proof. If $x_0 < x_1 < x_2$, then

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(x_2) - f(x_0)}{x_2 - x_0} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Hence

$$h \to \frac{f(x+h) - f(x)}{h}$$

is monotonically increasing. This implies the differentiability from the right, and similarly from the left and the relation between the derivatives. □

[21.11.2022]

[25.11.2022]

We have Jensen's inequality for the composition of convex functions and integrable functions.

Theorem 3.13 (Jensen's inequality)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, $\mu(X) = 1$, $F : \mathbb{R} \to \mathbb{R}$ convex and $f : X \mapsto \mathbb{R}$ integrable. Then $F \circ f$ is measurable and*

$$F \circ \int_X f \, \mathrm{d}\mu \leqslant \int_X F \circ f \, \mathrm{d}\mu .$$

Proof. Since $F : \mathbb{R} \to \mathbb{R}$ is convex, F is continuous and hence $F \circ f$ is measurable.
Let $t_0 = \int_X f d\mu$. By Lemma 3.12:

$$F(t) \geqslant F(t_0) + \frac{dF}{dt}^{+}(t_0)(t - t_0),$$

it holds

$$\int_X F\circ f d\mu \geqslant \int_X \Big(F(t_0) + \frac{dF^+}{dt}(t_0)(f-t_0)\Big)\,\mathrm{d}\mu$$
$$= F(t_0) + \frac{dF^+}{dt}(t_0)\left(\int_X f d\mu - t_0\right) = F(t_0).$$

□

Remark 3.14. *For a particular case where $X = \{x, y\}$, $\mu(\{x\}) = t$, $\mu(\{y\}) = 1-t$, for some $t \in [0,1]$, Jensen's inequality reads as*

$$F\Big(tf(x) + (1-t)f(y)\Big) \leqslant tF(f(x)) + (1-t)F(f(y)),$$

which follows also immediately from the definition of convexity.

3.2.2 Hölder's inequality

We prove the Hölder's inequality for general p-integrable functions. Recall Lemma 1.6 for the case of l^p, where $(X, \mathcal{A}, \mu) = (\mathbb{N}, 2^{\mathbb{N}}, \text{counting measure})$.

Lemma 3.15 (Hölder's inequality)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $1 \leqslant p, q \leqslant \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then fg is integrable and*

$$\left|\int_X fg d\mu\right| \leqslant \|f\|_{L^p}\|g\|_{L^q}. \tag{3.1}$$

If $1 < p < \infty$, $f \neq 0$ and $g \neq 0$, then the equality holds if and only if $\frac{g}{\|g\|_{L^q}} = \lambda|f|^{p-2}\bar{f}\|f\|_{L^p}^{-(p-1)}$ almost everywhere for some $\lambda \in \mathbb{K}$ with $|\lambda| = 1$ (in particular if $g = \lambda|f|^{p-2}\bar{f}$).

Proof. We follow the proof for Lemma 1.6 in the sequence space. If $\|f\|_{L^p} = 0$, then $f = 0$ almost everywhere, and the inequality (3.1) is trivial. The same holds for $\|g\|_{L^q} = 0$. In the following we assume $f \neq 0$ and $g \neq 0$.
If $(p, q) = (1, \infty)$, then $|g(x)| \leqslant \|g\|_{L^\infty} =: a > 0$ almost everywhere and the inequality (3.1) follows:

$$\int_X |fg| d\mu = \int_0^\infty \mu(\{|fg| > t\})dt \leqslant \int_0^\infty \mu(\{|f| > t/a\})dt$$
$$= a\int_0^\infty \mu(\{|f| > t\})dt = a\int_X |f| d\mu.$$

It is the same for $(p, q) = (\infty, 1)$.

Now show the inequality for the case $1 < p, q < \infty$. It suffices to consider f and g with $\|f\|_{L^p} = 1$ and $\|g\|_{L^q} = 1$, and prove

$$\int_X |f||g| d\mu \leqslant \int_X \Big(\frac{1}{p}|f|^p + \frac{1}{q}|g|^q\Big) d\mu = 1.$$

It follows immediately from Young's inequality and the inequality is strict unless

$$|fg(x)| = \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

almost everywhere, which implies $|g| = |f|^{p-1}$ almost everywhere. We have already shown

$$\left|\int_X fg d\mu\right| \leqslant \int_X |fg| d\mu \leqslant \|f\|_{L^p}\|g\|_{L^q}$$

and in the case of equality all inequalities must be equalities:

$$fg = \lambda|fg|, \quad |g| = |f|^{p-1}$$

almost everywhere for some complex number λ of modulus 1, by Remark 3.7 and the above consideration. This implies $fg = \lambda|f|^p$ and hence $g = \lambda|f|^{p-2}\bar{f}$. Back to our consideration, the equality implies $\tilde{g} = \lambda|\tilde{f}|^{p-2}\bar{\tilde{f}}$ for $\tilde{f} = \frac{f}{\|f\|_{L^p}}$ and $\tilde{g} = \frac{g}{\|g\|_{L^q}}$, that is, $g = \lambda|f|^{p-2}\bar{f}\|g\|_{L^q}\|f\|_{L^p}^{-(p-1)}$. This is also sufficient condition for the equality in (3.1). □

3.2.3 Minkowski's inequality

By use of Hölder's inequality one can show Minkowski's inequality.

Theorem 3.16 (Minkowski's inequality)**.** *Let $1 \leqslant p < \infty$ and let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ finite measurable spaces. Let f be $\mu \times \nu$ measurable. Then*

$$\left(\int_X \left|\int_Y |f(x,y)| d\nu(y)\right|^p d\mu(x)\right)^{1/p} \leqslant \int_Y \left(\int_X |f(x,y)|^p d\mu(x)\right)^{1/p} d\nu(y). \tag{3.2}$$

If $1 < p < \infty$, the integrals above are finite, and

$$\left(\int_X \left|\int_Y f(x,y) d\nu(y)\right|^p d\mu(x)\right)^{1/p} = \int_Y \left(\int_X |f(x,y)|^p d\mu(x)\right)^{1/p} d\nu(y), \tag{3.3}$$

then there exist a μ-measurable function α and a ν-measurable nonnegative function β so that

$$f(x,y) = \alpha(x)\beta(y)$$

almost everywhere.
A special case is the triangle inequality (which assumes $Y = \{0,1\}$, ν counting measure and no σ finiteness on X)

$$\|f+g\|_{L^p(\mu)} \leqslant \|f\|_{L^p(\mu)} + \|g\|_{L^p(\mu)}, \tag{3.4}$$

whenever f and g are p-integrable, with equality for $p > 1$ iff f and g are positively linearly dependent, i.e. $f = \lambda g$ for some $\lambda \geqslant 0$ or $g = 0$.

Proof. To show the inequality we assume first that f is nonnegative and omit the absolute value. The case $p = 1$ is trivial. We claim that

$$G(y) := \int_X f^p(x,y)d\mu(x) \qquad \text{and } H(x) := \int_Y f(x,y)d\nu(y)$$

are ν and μ measurable functions respectively. This follows from Fubini's Theorem 3.11 if f^p resp f are $\mu \times \nu$ integrable, and by an approximation argument in the general case. More precisely, we may assume without loss of generality $X = \cup_j A_j$, $Y = \cup_k B_k$ with (A_j) resp. (B_k) pairwise disjoint and of finite measure. Let $f_n = f 1_{\{f \leqslant n\}}$ be monotone increasing sequence, such that $f = \lim_{n\to\infty} f_n$, and f_n is integrable in the product measure space $(A_j \times B_k, (\mathcal{A} \times \mathcal{B})_{jk}, (\mu \times \nu)_{jk})$ restricted on $A_j \times B_k$, for any j,k. Hence $\int_{B_k} f_n(x,y)d\nu(y)$ is μ_j-measurable on A_j and hence μ-measurable on X. Thus $\int_Y f_n(x,y)d\nu(y)$ is μ-measurable and hence $H(x)$ as the limit is μ-measurable.

[25.11.2022]

[28.11.2022]

Then if $G^{1/p} \in L^1(\nu)$ and $H \in L^p(\mu)$ such that fH^{p-1} is $\mu \times \nu$ integrable, the Minkowski's inequality (3.2) follows:

$$\begin{aligned}
\int_X H^p(x)d\mu(x) &= \int_X \int_Y f(x,y)d\nu(y)H^{p-1}(x)d\mu(x) \\
&= \int_Y \int_X f(x,y)H^{p-1}(x)d\mu(x)d\nu(y) \\
&\leqslant \int_Y \left(\int_X f^p(x,y)d\mu(x)\right)^{1/p} \left(\int_X H^p d\mu(x)\right)^{\frac{p-1}{p}} d\nu(y) \\
&= \int_Y (G(y))^{1/p}\, d\nu(y) \left(\int_X H^p d\mu(x)\right)^{\frac{p-1}{p}} \\
\Rightarrow \left(\int_X H^p(x)d\mu(x)\right)^{1/p} &\leqslant \int_Y (G(y))^{1/p} d\nu(y),
\end{aligned}$$

where we used Hölder's inequality with $q = \frac{p}{p-1}$. An approximation argument as above implies the general case.

Now assume that $p > 1$, f is complex valued, such that the integrals in (3.2) are finite. In order to have equality in the application of Hölder's inequality it holds for almost all x and y

$$\frac{H(x)^{p-1}}{\|H(x)^{p-1}\|_{L^{\frac{p}{p-1}}}} = \Big(\frac{|f(x,y)|}{\|f(x,y)\|_{L^p_x}}\Big)^{p-1}, \text{ where } H(x) = \int_Y |f(x,y)| d\nu(y),$$

which implies

$$|f(x,y)| = \tilde{\alpha}(x)\beta(y)$$

for some measurable nonnegative functions $\tilde{\alpha}$ and β. In order to have f instead of $|f|$ on the left hand side of the equality (3.3), there exists $\lambda(x)$ with $|\lambda| = 1$ such that $f(x,y) = \lambda(x)|f(x,y)|$ almost everywhere.

For the triangle inequality we apply the first part with the counting measure on $Y = \{0,1\}$. The product measure is defined in the obvious fashion, even without assuming that μ is σ finite. If f is p integrable then by the definition of the integral

$$\mu(\{x : |f(x)| > t\}) \leqslant t^{-p}\|f\|^p_{L^p}.$$

Indeed, for $p = 1$, it follows from the definition that

$$\begin{aligned}\int_X |f| \,\mathrm{d}\mu &= \int_0^\infty \mu(\{x : |f(x)| > t\})dt \geqslant \int_0^{t_0} \mu(\{x : |f(x)| > t\})dt \\ &\geqslant \int_0^{t_0} \mu(\{x : |f(x)| > t_0\})dt = t_0\mu(\{x : |f(x)| > t_0\}).\end{aligned}$$

For general $1 \leqslant p < \infty$, we replace $|f|$ by $|f|^p$. Let

$$A = \bigcup_{j=1}^\infty \Big\{x : |f(x)| + |g(x)| > \frac{1}{j}\Big\}$$

which is a countable union of sets of finite measure. We replace X by A, take as σ algebra the sets in $\mathcal{A}$ which are contained in A, and μ restricted to this σ algebra as measure. This is σ finite measure space. □

Remark 3.17. *If $p = \infty$, then the inequality* (3.2) *holds trivially since*

$$\int_Y |f(x,y)| d\nu(y) \leqslant \int_Y \|f(x,y)\|_{L^\infty_x(\mu)} d\nu(y), \quad \mu - a.e.$$

The Minkowski's inequality (3.2) *can be easily generalized to*

$$\Big\|\|f\|_{L^q(\nu)}\Big\|_{L^p(\mu)} \leqslant \Big\|\|f\|_{L^p(\nu)}\Big\|_{L^q(\mu)},$$

whenever $1 \leqslant q \leqslant p \leqslant \infty$. Indeed, we apply (3.2) *with f, p replaced by $|f|^q$, $p/q \geqslant 1$ respectively.*

3.2.4 Hanner's inquality

There is an improvement of the triangle inequality.

Theorem 3.18 (Hanner's inequality)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and* f*,* g *be p-integrable functions,* $1 \leqslant p < \infty$*. If* $1 \leqslant p \leqslant 2$ *then*

$$\|f+g\|_{L^p}^p + \|f-g\|_{L^p}^p \geqslant (\|f\|_{L^p} + \|g\|_{L^p})^p + \big|\|f\|_{L^p} - \|g\|_{L^p}\big|^p, \tag{3.5}$$

$$(\|f+g\|_{L^p} + \|f-g\|_{L^p})^p + \big|\|f+g\|_{L^p} - \|f-g\|_{L^p}\big|^p \leqslant 2^p(\|f\|_{L^p}^p + \|g\|_{L^p}^p). \tag{3.6}$$

If $2 \leqslant p < \infty$ *all inequalities are reversed.*

Remark 3.19. *The inequalities reduce to the parallelogram identity if* $p = 2$*, (recalling* (2.1)*).*
Both (3.5) *and* (3.6) *are equivalent: The second is obtained from the first by replacing* f *by* $f + g$ *and* g *by* $f - g$*. It suffices to prove the first inequality.*
If $p = 1$*, then* (3.6) *reads as*

$$2\max\{\|f+g\|_{L^1}, \|f-g\|_{L^1}\} \leqslant 2(\|f\|_{L^1} + \|g\|_{L^1}),$$

which follows from the triangle inequality (3.4)*.*

Proof. As (3.6) follows from the triangle inequality (3.4), we restrict ourselves to the case $1 < p < \infty$. We may assume that $\|g\|_{L^p} \leqslant \|f\|_{L^p}$ (otherwise we exchange the two) and $\|f\|_{L^p} = 1$ (otherwise we multiply f and g by the inverse of the norm $\|f\|_{L^p}^{-1}$).
The first inequality follows from the following pointwise inequality: Let

$$\alpha(r) = (1+r)^{p-1} + (1-r)^{p-1}, \quad \beta(r) = [(1+r)^{p-1} - (1-r)^{p-1}]r^{1-p}.$$

We claim that

$$\alpha(r)|f|^p + \beta(r)|g|^p \leqslant |f+g|^p + |f-g|^p \tag{3.7}$$

for $1 < p \leqslant 2$, $0 \leqslant r \leqslant 1$ and complex numbers f and g (and the reverse inequality for $2 \leqslant p < \infty$). Indeed, (3.7) implies

$$\alpha(r)|f(x)|^p + \beta(r)|g(x)|^p \leqslant |f(x)+g(x)|^p + |f(x)-g(x)|^p$$

and by integration

$$\alpha(r)\|f\|_{L^p}^p + \beta(r)\|g\|_{L^p}^p \leqslant \|f+g\|_{L^p}^p + \|f-g\|_{L^p}^p.$$

We apply the inequality with $r = \|g\|_{L^p}$ and recall that $\|f\|_{L^p} = 1$. The left hand side becomes

$$\Big[(\|f\|_{L^p} + \|g\|_{L^p})^{p-1} + (\|f\|_{L^p} - \|g\|_{L^p})^{p-1}\Big]\|f\|_{L^p}$$
$$+ \Big[(\|f\|_{L^p} + \|g\|_{L^p})^{p-1} - (\|f\|_{L^p} - \|g\|_{L^p})^{p-1})\Big]\|g\|_{L^p}$$
$$=(\|f\|_{L^p} + \|g\|_{L^p})^p + (\|f\|_{L^p} - \|g\|_{L^p})^p.$$

[28.11.2022]

[02.12.2022]

It remains to prove (3.7): If f, g are both real-valued,

$$\alpha(r) + \beta(r)R^p \leqslant |1+R|^p + |1-R|^p, \quad \forall R \geqslant 0,$$

if $1 < p \leqslant 2$ (and the reverse inequality for $2 \leqslant p < \infty$). The proof is elementary. Let for $0 \leqslant R \leqslant 1$,

$$F_R(r) = \alpha(r) + \beta(r)R^p.$$

We claim that it attains its maximum at $r = R$ if $1 \leqslant p < 2$ and resp. its minimum if $p > 2$. We compute

$$\begin{aligned} F_R' &= \alpha' + \beta' R^p \\ &= (p-1)[(1+r)^{p-2} - (1-r)^{p-2}] - (p-1)[(1+r)^{p-2} - (1-r)^{p-2}]r^{-p}R^p \\ &= (p-1)[(1+r)^{p-2} - (1-r)^{p-2}](1-(R/r)^p) \end{aligned}$$

and the derivative vanishes only at $r = R$ inside $(0,1)$ and changes sign there. Thus

$$\alpha(r) + \beta(r)R^p \leqslant (1+R)^p + (1-R)^p$$

if $0 \leqslant R \leqslant 1$ and $p \leqslant 2$ with the opposite inequality if $p \geqslant 2$. Now let $R \geqslant 1$. Since $\alpha + \beta R^p$ is increasing on $[0,1]$ if $p \leqslant 2$, we obtain

$$\alpha(r) + \beta(r)R^p \leqslant 2^{p-1}(1+R^p) \leqslant (R+1)^p + (R-1)^p$$

where the second inequality is ensured by the relation between the derivatives: $2^{p-1}pR^{p-1} \leqslant p((R+1)^{p-1} + (R-1)^{p-1})$ which follows from $2^{p-1} \leqslant \alpha(\frac{1}{R}) = (1+\frac{1}{R})^{p-1} + (1-\frac{1}{R})^{p-1}$. The reverse inequality holds if $p > 2$. This implies (3.7) for real-valued f and g. We claim that (3.7) holds for complex f and g. It suffices to consider $f = a > 0$ and $g = be^{i\theta}$, $b > 0$ and it follows from the fact that

$$|f+g|^p + |f-g|^p = (a^2 + b^2 + 2ab\cos\theta)^{p/2} + (a^2 + b^2 - 2ab\cos\theta)^{p/2}$$

has its minimum at $\theta = 0$ (resp. its maximum if $p \geqslant 2$). It suffices to consider the function $h(\rho) = (a^2 + b^2 + 2ab\rho)^{p/2} + (a^2 + b^2 - 2ab\rho)^{p/2}$ on $\rho \in [0,1]$, and notice that $x \to x^{p/2}$ is concave if $p \leqslant 2$ (convex if $p \geqslant 2$) and $h'(0) = 0$. □

3.3 Lebesgue spaces $L^p(\mu)$

In this section we will always assume that $(X, \mathcal{A}, \mu)$ is a measurable space, and we will study the Lebesgue spaces $L^p(\mu)$.

3.3.1 Fischer-Riesz theorem

Theorem 3.20 (Fischer-Riesz). *$(L^p(\mu), \|\|_{L^p})$ is Banach space.*

Proof. The fact that $(L^p(\mu), \|\|_{L^p})$ is a normed space follows from Definition 3.6, Remark 3.7 and the triangle inequality (3.4).
It remains to show completeness. Let (f_n) be representatives of a Cauchy sequence. By taking subsequences if necessary we may assume

$$\|f_n - f_m\|_{L^p} \leqslant 2^{-\min\{m,n\}}.$$

We define the monotone sequence of functions

$$F_n(x) = |f_1(x)| + \sum_{m=1}^{n-1} |f_{m+1}(x) - f_m(x)|$$

and $F = \lim_{n\to\infty} F_n(x)$. F is measurable and by monotone convergence theorem for $1 \leqslant p < \infty$,

$$\Big(\int_X |F|^p d\mu\Big)^{1/p} = \lim_{n\to\infty}\Big(\int_X |F_n|^p d\mu\Big)^{1/p} \leqslant \|f_1\|_{L^p} + \sum_{m\geqslant 1} 2^{-m} = \|f_1\|_{L^p} + 1,$$

while for $p = \infty$ it holds also $\|F\|_{L^\infty} \leqslant \|f_1\|_{L^\infty} + 1$, and in particular F is finite almost everywhere. Thus

$$f_n = f_1 + \sum_{m=1}^{n-1} (f_{m+1} - f_m)$$

converges if $F(x) < \infty$. Let f be the limit if $F(x) < \infty$, and 0 otherwise. It is measurable. Since $\max\{f, f_n\} \leqslant F$ we obtain by dominated convergence

$$\|f - f_n\|_{L^p}^p = \int_X |f - f_n|^p d\mu \to 0,$$

for $1 \leqslant p < \infty$, and for $p = \infty$ it follows directly from $\| \sum_{m=n}^{\infty}(f_{m+1} - f_m)\|_{L^\infty} \to 0$ as $n \to \infty$. □

Remark 3.21. *The Lebesgue space $L^2(\mu)$ equipped with the inner produce*

$$\langle f, g\rangle_{L^2(\mu)} = \int_X f\bar{g}\, d\mu$$

is a Hilbert space.

3.3.2 Projection and duality

Recall the Hilbert projection Theorem 2.15 for the Hilbert space $L^2(\mu)$. We have similar projection result in $L^p(\mu)$, $1 < p < \infty$.

Lemma 3.22 (Projection in $L^p(\mu)$)**.** *Let $1 < p < \infty$ and let K be a closed convex set in $L^p(\mu)$. Let $f \in L^p(\mu)$. Then there exists a unqiue $g \in K$ with*

$$\|f - g\|_{L^p(\mu)} = \operatorname{dist}(f, K).$$

Moreover

$$\operatorname{Re} \int_X (h - g)(\bar{f} - \bar{g})|f - g|^{p-2} d\mu \leqslant 0, \quad \forall h \in K. \tag{3.8}$$

Proof. Let $\{h_n\}$ be a minimzing sequence. Since $\frac{1}{2}(h_n + h_m) \in K$ and $\|h_n - f + h_m - f\|_{L^p} \leqslant \|h_n - f\|_{L^p} + \|h_m - f\|_{L^p}$, we see that

$$\|h_n - f + h_m - f\|_{L^p} \to 2 \operatorname{dist}(f, K).$$

Now let $p \leqslant 2$, from the second Hanner's inequality we obtain

$$\begin{aligned} &\big(\|h_n - f + h_m - f\|_{L^p} + \|h_n - h_m\|_{L^p}\big)^p \\ &\qquad + \big|\|h_m - f + h_m - f\|_{L^p} - \|h_n - h_m\|_{L^p}\big|^p \\ \leqslant &2^p\big(\|h_n - f\|_{L^p}^p + \|h_m - f\|_{L^p}^p\big). \end{aligned}$$

Let $A = \limsup_{n,m\to\infty} \|h_n - h_m\|_{L^p}$. This limsup is obtained along two subsequences $n, m \to \infty$. Let $D = \operatorname{dist}(f, K)$. Then

$$(2D + A)^p + |2D - A|^p \leqslant 2^{p+1} D^p$$

which implies $A = 0$ by the strict convexity of $A \to (2D + A)^p + (2D - A)^p$ for $A \in [0, 2D)$.

[02.12.2022]

[05.12.2022]

If $p > 2$ we argue similarly with the first inequality:

$$\begin{aligned} &\|h_n - f + h_m - f\|_{L^p}^p + \|h_n - h_m\|_{L^p}^p \\ &\leqslant (\|h_n - f\|_{L^p} + \|h_m - f\|_{L^p})^p + |\|h_n - f\|_{L^p} - \|h_m - f\|_{L^p}|^p \to (2D)^p. \end{aligned}$$

Thus (h_n) is a Cauchy sequence. Up to a subsequence, (h_n) converges to some limit $g \in K$ in $L^p(\mu)$, and $g \in K$ is the point of minimal distance. The same argument implies that the closest point is unique.

Now let $h \in K$. Let

$$N(t) = \int_X |f - (g + t(h - g))|^p d\mu.$$

Then $N(t)$ attains its minimum at $t = 0$ on the interval $[0, 1]$. We claim that its derivative at $t = 0$ is

$$\frac{d}{dt} N|_{t=0} = p \operatorname{Re} \int_X |f(x) - g(x)|^{p-2} (f(x) - g(x))(\bar{g}(x) - \bar{h}(x)) d\mu.$$

This implies the assertion $N'(0) \geqslant 0$. To calculate the derivative we assume that $f, g \in L^p(\mu)$, then almost everywhere

$$\frac{d}{dt}|_{t=0} |f + tg|^p = p|f|^{p-2} \operatorname{Re} f\bar{g}$$

and the pth power is convex such that the composition of $G = G(z, \bar{z}) = |z|^p$ and the linear function $H(t) = f + tg$ is convex in $t \in [-1, 1]$:

$$|f|^p - |f - g|^p \leqslant \frac{1}{t}(|f + tg|^p - |f|^p) \leqslant |f + g|^p - |f|^p.$$

The formula follows by Lebesgue's dominated convergence theorem. □

Remark 3.23. *If K is a closed subspace of $L^p(\mu)$, $1 < p < \infty$, then the projection relation* (3.8) *can be improved into*

$$\int_X h(\bar{f} - \bar{g})|f - g|^{p-2} d\mu = 0, \quad \forall h \in K. \tag{3.9}$$

Theorem 3.24 (Duality of $L^p(\mu)$)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$j : L^q \ni g \to (f \to \int_X fg d\mu) \in (L^p(\mu))^*$$

is a linear isometric isomorphism.

Proof. The proof is the same as for Riesz representation theorem 2.25 in case of Hilbert spaces. By Hölder's inequality the map is well defined and

$$\|j(g)\|_{(L^p)^*} \leqslant \|g\|_{L^q}.$$

Since

$$\|j(g)\|_{(L^p)^*} \|g\|_{L^q}^{q-1} = \|j(g)\|_{(L^p)^*} \||g|^{q-2}\bar{g}\|_{L^p} \geqslant j(g)(|g|^{q-2}\bar{g}) = \int_X |g|^q d\mu = \|g\|_{L^q}^q,$$

we conclude as for Hilbert spaces that

$$\|j(f)\|_{(L^p)^*} = \|f\|_{L^q}.$$

Surjectivity is proven exactly as for Hilbert spaces (see Theorem 2.25). **Exercise.** □

Corollary 3.25 (Duality of $L^1(\mu)$)**.** *Suppose that $(X, \mathcal{A}, \mu)$ is σ finite measure space. Then*

$$L^\infty(\mu) \ni g \to (f \to \int_X fgd\mu) \in (L^1(\mu))^*$$

is a linear isometric isomorphism.

Proof. **Exercise.** □

The duality of the space $L^\infty(\mu)$ is more involved. We are going to show Riesz representation Theorem 3.41 below, which characterizes the linear functional on $C_c(X)$ where X is sigma compact metric space in terms of Radon measure.

3.3.3 Young's inequality on $\mathbb{R}^d$

Let $(X, \mathcal{A}, \mu)$ be a measure space and suppose that $1 \leqslant p, q, r \leqslant \infty$ and $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. If $f \in L^p(\mu)$, $g \in L^q(\mu)$ and $h \in L^r(\mu)$, then fgh is integrable and

$$\left| \int_X fghd\mu \right| \leqslant \|f\|_{L^p(\mu)} \|g\|_{L^q(\mu)} \|h\|_{L^r(\mu)}.$$

This is a consequence of a multiple application of Hölder's inequality:

$$\left| \int_X fghd\mu \right| \leqslant \|f\|_{L^p} \|gh\|_{L^{\frac{p}{p-1}}}$$

and if $p > 1$,

$$\int_X |g|^{\frac{p}{p-1}} |h|^{\frac{p}{p-1}} d\mu \leqslant \||g|^{\frac{p}{p-1}}\|_{L^{\frac{q(p-1)}{p}}} \||h|^{\frac{p}{p-1}}\|_{L^{\frac{r(p-1)}{p}}} = \|g\|_{L^q}^{\frac{p}{p-1}} \|h\|_{L^r}^{\frac{p}{p-1}},$$

since

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

We denote $L^p(\mathbb{R}^d)$ (or even L^p) for $L^p(m^d)$ where m^d is the Lebesgue measure.

Lemma 3.26. *Suppose that $1 \leqslant p, q, r \leqslant \infty$ satisfy*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$$

and that $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$ and $h \in L^r(\mathbb{R}^d)$. Then

$$\mathbb{R}^d \times \mathbb{R}^d \ni (x, y) \to f(-x)g(x-y)h(y)$$

is integrable and

$$I(f, g, h) := \int_{\mathbb{R}^d \times \mathbb{R}^d} f(-x)g(x-y)h(y)dm^{2d}(x, y)$$

satisfies

$$|I(f, g, h)| \leqslant \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}$$

and I is invariant by changing the orders of f, g or of g, h:

$$I(f, g, h) = I(g, f, h) = I(g, h, f) = I(h, g, f) = I(h, f, g) = I(f, h, g).$$

Proof. Measurability is a consequence of the theorem of Fubini and an approximation argument (similar as in the proof of Minkowski's inequality 3.2). It suffices to prove the estimate for nonnegative functions. We assume $1 < p, q, r < \infty$ since the limit cases are simpler, and follow by obvious modifications. We define p', q' and r' by $\frac{1}{p} + \frac{1}{p'} = 1$, i.e. $p' = \frac{p}{p-1}$ etc. Let

$$\begin{aligned} \alpha(x, y) &= |f(-x)|^{p/r'} |g(x-y)|^{q/r'}, \\ \beta(x, y) &= |f(-x)|^{p/q'} |h(y)|^{r/q'}, \\ \gamma(x, y) &= |g(x-y)|^{q/p'} |h(y)|^{r/p'}. \end{aligned}$$

Then $\frac{1}{p'} + \frac{1}{q'} + \frac{1}{r'} = 1$ and

$$\begin{aligned} |I| \leqslant & \int_{\mathbb{R}^d \times \mathbb{R}^d} \alpha(x, y)\beta(x, y)\gamma(x, y)dm^{2d} \\ \leqslant & \|\alpha\|_{L^{r'}} \|\beta\|_{L^{q'}} \|\gamma\|_{L^{p'}} \\ = & \|f\|_{L^p}^{\frac{p}{r'}} \|g\|_{L^q}^{\frac{q}{r'}} \|f\|_{L^p}^{\frac{p}{q'}} \|h\|_{L^r}^{\frac{r}{q'}} \|g\|_{L^q}^{\frac{q}{p'}} \|h\|_{L^r}^{\frac{r}{p'}} \\ = & \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}. \end{aligned}$$

The second last equality is a consequence of the theorem of Fubini. □

[05.12.2022]

[09.12.2022]

Theorem 3.27 (Young's inequality)**.** *Suppose that* $1 \leqslant p, q, r' \leqslant \infty$ *and*

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r'}.$$

If $f \in L^p$ *and* $g \in L^q$*, then for almost all* x

$$f(x-y)g(y)$$

is integrable and

$$f * g(x) := \begin{cases} \int_{\mathbb{R}^d} f(x-y)g(y)dm^d(y) & \text{if integrable} \\ 0 & \text{otherwise} \end{cases}$$

defines a unique element in $L^{r'}(\mathbb{R}^d)$ *and*

$$\|f * g\|_{L^{r'}(\mathbb{R}^d)} \leqslant \|f\|_{L^p}\|g\|_{L^q}.$$

In particular, if $q = 1$*, then* $p = r'$ *and*

$$\|f * g\|_{L^p} \leqslant \|f\|_{L^p}\|g\|_{L^1}.$$

Proof. We have $e^{-|x|^2} \in L^r$ for all $1 \leqslant r \leqslant \infty$. Then

$$e^{-|x|^2} f(x-y)g(y)$$

is m^{2d} integrable by Lemma 3.26. We apply Fubini to see that $\int_{\mathbb{R}^d} f(x-y)g(y)dm^d(y)$ exists for almost all x. If $r = \infty$, i.e. $p = q = r' = 1$, then it follows from Fubini Theorem.
For the case $1 \leqslant r < \infty$, by Theorem 3.24 and Corollary 3.25, the estimate follows once we prove

$$\left| \int_{\mathbb{R}^d} f * g h dm^d \right| \leqslant \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r}$$

for

$$\frac{1}{r} + \frac{1}{r'} = 1$$

and all $h \in L^r$. Since then

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$$

and, by Fubini and Lemma 3.26

$$\begin{aligned} \left| \int_{\mathbb{R}^d} f * gh(x) dm^d \right| \leqslant & \int_{\mathbb{R}^d} |f| * |g||h| dm^d \\ = & \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)||g(y)||h(x)| dm^{2d}(x,y) \\ \leqslant & \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r}. \end{aligned}$$

□

3.3.4 Schur's lemma

We have seen from Young's inequality that any function $g \in L^1(\mathbb{R}^d)$ defines a linear map from $L^p(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$, $p \in [1, \infty]$:

$$L^p(\mathbb{R}^d) \ni f \mapsto (f * g) \in L^p(\mathbb{R}^d).$$

However a general measure space $(X, \mathcal{A}, \mu)$ is not necessarily a vector space and the addition operator may make no sense. This motivates us to consider more general integral kernel $k(x, y)$ (instead of $g(x - y)$), and Schur's lemma gives a criterium for an integral kernel to define a linear map from $L^p(\nu)$ to $L^p(\mu)$ for $1 \leqslant p \leqslant \infty$.

Theorem 3.28 (Schur's lemma). *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ finite measure spaces and $k : X \times Y \to \mathbb{C}$ be $\mu \times \nu$ measurable. Suppose that $C_1, C_2 \in [0, \infty)$ and*

$$\sup{}_x \int_Y |k(x, y)| d\nu(y) \leqslant C_1, \qquad \sup{}_y \int_X |k(x, y)| d\mu(x) \leqslant C_2.$$

If $1 \leqslant p \leqslant \infty$ and $f \in L^p(\nu)$, then

$$\int_Y k(x, y) f(y) d\nu(y)$$

exists for almost all x and

$$\Big\| \int_Y k(x, y) f(y) d\nu(y) \Big\|_{L^p(\mu)} \leqslant C_1^{1-\frac{1}{p}} C_2^{\frac{1}{p}} \|f\|_{L^p(\nu)}.$$

The map

$$L^p(\nu) \ni f \to Tf := \int_Y k(x, y) f(y) d\nu(y) \in L^p(\mu)$$

is a continuous linear map which satisfies

$$\|T\|_{L^p(\nu) \to L^p(\nu)} \leqslant C_1^{1-\frac{1}{p}} C_2^{\frac{1}{p}}.$$

Proof. As in the proof of Young's inequality, it suffices to show the estimate. If $p = 1$ or $p = \infty$ this is an immediate consequence of the theorem of Fubini. Repeating the argument of Young's inequality we have to prove that

$$\int_{X \times Y} |g(x)||k(x, y)||f(y)| d(\mu \times \nu) \leqslant C_1^{1-\frac{1}{p}} C_2^{\frac{1}{p}} \tag{3.10}$$

where

$$\frac{1}{p}+\frac{1}{q}=1,$$

if $\|f\|_{L^p} = \|g\|_{L^q} = 1$ and $1 < p, q < \infty$. It suffices to prove (3.10) for nonnegative functions f, g and k, where f and g are bounded and 0 outside a set of finite measure, and the general case follows by an approximation argument.

For $z \in \mathbb{C}$ with $0 \leqslant \operatorname{Re} z \leqslant 1$, we define

$$f_z = \begin{cases} f^{pz-1}f & \text{if } f \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$g_z = \begin{cases} g^{\frac{p}{p-1}(1-z)-1}g & \text{if } g \neq 0 \\ 0 & \text{otherwise .} \end{cases}$$

Then for $\sigma \in \mathbb{R}$,

$$\|f_{i\sigma}\|_{L^\infty} = \|g_{1+i\sigma}\|_{L^\infty} = 1,$$

and

$$\|f_{1+i\sigma}\|_{L^1} = \|f\|^p_{L^p} = \|g_{i\sigma}\|_{L^1} = \|g\|^q_{L^q} = 1,$$

and hence

$$\int_{X\times Y} |g_{i\sigma}(x)||k(x,y)||f_{i\sigma}(y)|d\mu \times \nu \leqslant C_1$$

$$\int_{X\times Y} |g_{1+i\sigma}(x)||k(x,y)||f_{1+i\sigma}(y)|d\mu \times \nu \leqslant C_2.$$

Moreover

$$f_{\frac{1}{p}} = f, \qquad g_{\frac{1}{p}} = g.$$

Notice that f_z and g_z are bounded and zero outside a set of finite measure. By dominated convergence

$$z \to H(z) = \int_{X\times Y} g_z(x)k(x,y)f_z(y)d(\mu\times\nu)$$

is continuous in the strip $\mathcal{C} = \{z : 0 \leqslant \operatorname{Re} z \leqslant 1\}$, differentiable and satisfies the Cauchy-Riemann differential equations. The claim follows from the three lines inequality, which ensures that the maximum is attained at the boundary:

Lemma 3.29 (Three lines inequality)**.** *Suppose that $u \in C(\mathcal{C})$ is bounded and holomorphic in the interior. Then*

$$\sup{}_{\mathcal{C}}|u| = \sup{}_{\partial\mathcal{C}}|u|$$

We apply the lemma to

$$u(z) = C_1^{z-1} C_2^{-z} H(z),$$

such that

$$C_1^{-(1-\frac{1}{p})} C_2^{-\frac{1}{p}} \int_{X \times Y} g(x) k(x,y) f(y) d(\mu \times \nu) = u(\frac{1}{p}) \leqslant 1.$$

□

Proof of Lemma 3.29. (**Explained in the exercise class.**) 1) Let $U \subset \mathbb{C}$ be a *bounded* open connected set and $u \in C(\overline{U}; \mathbb{C})$ be a holomorphic function in the interior. We claim that then

$$\sup_{x \in \overline{U}} |u(x)| = \sup_{x \in \partial U} |u(x)|.$$

We prove this by contradiction. Suppose that $|u|$ attains its maximum M at some interior point z_0 and suppose that this is larger than $\sup_{\partial U} |u(z)|$. Then

$$f(z) = \text{Re}[u(z)/u(z_0)]$$

satisfies $0 \leqslant f \leqslant 1$ and $f(z_0) = 1$. Moreover f is harmonic. Let

$$f_\varepsilon(x + iy) = f(x + iy) + \varepsilon (x - \text{Re}\, z_0)^2$$

where ε is so small that $f_\varepsilon(z) < 1$ for $z \in \partial U$. Then f_ε has a maximum in an interior point z_1. At this point the Hessian is negative semidefinite by its trace $\Delta f_\varepsilon(z_1) = 2\varepsilon$. This is a contradiction.
2) Let u be defined in $\mathcal{C}$ as in the lemma, and let

$$u_\varepsilon(z) = e^{\varepsilon z^2} u(z).$$

Since $u_\varepsilon(z) \to 0$ as $|\,\text{Im}\, z| \to \infty$

$$\sup_{z \in \mathcal{C}} |u_\varepsilon(z)| = \sup_{z \in \partial \mathcal{C}} |u_\varepsilon| \leqslant e^\varepsilon \sup_{z \in \partial \mathcal{C}} |u(z)|.$$

Now we let ε tend to zero. □

3.4 Radon measures on metric spaces

In this section we always assume that (X, d) is a metric space, and we restrict ourselves to Borel measure space $(X, \mathcal{B}(X), \mu)$.
We recall that a subset K in a metric space (X, d) is *compact* (i.e. any open covering of K contains a finite subcovering) is equivalent to **(This will be explained shortly in the exercise class.)**

- K is *sequentially compact*, i.e. every sequence in K contains a subsequence which converges to a limit in K;

- K is *totally bounded* and *complete*, i.e. for any $\varepsilon > 0$ there exists a finite covering $\cup_{j=1}^{N} B(x_j, \varepsilon)$ of K, and any Cauchy sequence in K has a subsequence which converges to a limit in K.

Recall that X is *locally compact* if for every point x there is a neighborhood whose closure is compact. (X, d) is called σ *compact*[4] if it is locally compact and if it is a countable union of compact sets. If (X, d) is σ compact, then there exists a sequence of compact set K_j so that K_j is contained in the interior $\mathring{K}_{j+1}$ of K_{j+1} and $X = \bigcup_{j=1}^{\infty} K_j$ (**Exercise.**) We denote by $C_0(X) \subset C_b(X)$ the continuous functions f with limit 0 at ∞, i.e. for all $\varepsilon > 0$ there exists j so that f is at most of size ε outside K_j. We define $C_c(X)$ as the subspace of continuous functions with compact support.
We recall that the Borel sets $\mathcal{B}(X)$ are the smallest σ algebra containing all open sets. A Borel measure is a measure on the Borel sets. A Borel measure μ on a metric space (X, d) is called *inner regular* if for every Borel set A

$$\mu(A) = \sup \{\mu(K) : K \subset A, K \text{ compact}\},$$

while *outer regular* if for every Borel set A

$$\mu(A) = \inf\{\mu(U) : A \subset U, U \text{ open}\}.$$

It is called *locally finite* if for every $x \in X$ there exists an open neighborhood $U \ni x$ so that $\mu(U) < \infty$.

3.4.1 Radon measure

Definition 3.30 (Radon measure)**.** *A Borel measure μ on a metric space (X, d) is called a Radon measure, if it is locally finite and inner regular.*

Lemma 3.31 (Compact sets are of finite Radon-measure)**.** *Let μ be a Radon measure on a metric space (X, d). Then the measure of a compact set is finite. Furthermore for $\varepsilon > 0$ and K compact, there exists an open set $U \supset K$ of finite measure with $\mu(U) < \mu(K) + \varepsilon$.*

Proof. Let K be compact. For every $x \in K$ exists an open set U_x containing x with $\mu(U_x) < \infty$. Since K is compact and

$$K \subset \bigcup U_x,$$

[4] In some references σ compact metric space is not necessarily locally compact.

there exists a finite subcovering

$$K \subset \bigcup_{j=1}^{N} U_{x_j} =: \tilde{U}$$

such that

$$\mu(K) \leqslant \mu(\tilde{U}) \leqslant \sum_{j=1}^{N} \mu(U_{x_j}) < \infty.$$

We define $U_j = \tilde{U} \cap \{x : d(x, K) < \frac{1}{j}\}$, such that the finite-measure set $U_j \supset K$. Since $1_{U_j} \to 1_K$ pointwisely, by the theorem of Lebesgue

$$\int_X 1_{U_j} d\mu = \mu(U_j) \to \mu(K) = \int_X 1_K d\mu.$$

□

[09.12.2022]

[12.12.2022]

Vise Versa: If (X, d) is a σ compact metric space, then a Borel measure is also a Radon measure, if (and only if) any compact set is of finite measure.

Lemma 3.32 (Radon measure on σ compact metric space)**.** *Let (X, d) be a σ compact metric space.*
Then a Borel measure such that any compact set is of finite measure is Radon measure. And a Radon measure is σ finite and outer regular.

Proof. We claim first that for any Borel measure μ on a metric space (X, d), for any Borel set B with $\mu(B) < \infty$ and any $\varepsilon > 0$, there exists a closed set $C \subset B$ with $\mu(B \backslash C) < \varepsilon$. **(Exercise.)** If in addition any compact set if of finite μ-measure, then μ is locally finite since X is locally compact. Let K_j be compact subsets with $X = \bigcup K_j$ and K_j contained in the interior of K_{j+1}. Then for any closed set C with finite-measure

$$\mu(C \cap K_j) \to \mu(C)$$

and $C \cap K_j$ is compact. If $\mu(B) < \infty$, then its measure is the supremum of the measures of a sequence of compact sets $C_k \cup K_j$. If $\mu(B) = \infty$, then the same holds since the Borel set $K_j \cap B$ is of finite measure. Thus μ is inner regular and hence Radon measure.

Let now μ be Radon. Then it is σ finite since X is σ compact. For any Borel set B, $K_j \backslash B$ is Borel with $\mu(\mathring{K}_j \backslash B) < \infty$. Then there exists a closed set $C_j \subset \mathring{K}_j \backslash B$ with $\mu((\mathring{K}_j \backslash C_j) \backslash B) < \varepsilon 2^{-j}$. Let

$$U = \bigcup_{j=1}^{\infty} (\mathring{K}_j \backslash C_j).$$

It is open and

$$B = \bigcup_{j=1}^{\infty} (\mathring{K}_j \cap B) \subset \bigcup_{j=1}^{\infty} \mathring{K}_j \backslash C_j = U.$$

Moreover

$$\mu(U \backslash B) = \mu(\bigcup (\mathring{K}_j \backslash C_j) \backslash B) < \varepsilon.$$

Thus a Radon measure is outer regular. □

Remark 3.33. *The Lebesgue measure restricted to Borel sets is Radon measure. The counting measure on $\mathbb{R}$ is not a Radon measure.*
Continuous functions on compact metric spaces are $p-$integrable, $p \in [1, \infty]$ with respect to Radon measures, which are finite measures. In particular, continuous functions with compact supports are $p-$integrable with respect to the Radon measure μ on a metric space. However, without the assumption of compact supports, continuous functions are not necessarily in $L^p(\mu)$: For example, the constant function 1 is a continuous function defined on $\mathbb{R}^d$ with support $\mathbb{R}^d$, but is not in $L^p(\mathbb{R}^d)$ for any $p \in [1, \infty)$.

3.4.2 Density of Lip.-functions with comp. supp. in $L^p(\mu)$

Continuous functions with compact supports belong to $L^p(\mu)$, $p \in [1, \infty]$ with respect to Radon measure μ on a σ-compact metric space. They are indeed dense in $L^p(\mu)$ if $1 \leqslant p < \infty$.

Theorem 3.34 (Density of Lipschitz functions with compact supports in $L^p(\mu)$)**.** *Let (X, d) be a σ compact metric space, μ a Radon measure on X and $1 \leqslant p < \infty$. Then the set of Lipschitz-continuous functions with compact supports are dense in $L^p(\mu)$.*

Proof. **Step 1. Continuous functions with compact supports are dense.** We claim that the simple function set is dense in $L^p(\mu)$, $1 \leqslant p < \infty$, **(Exercise.)** and it suffices to approximate a characteristic function of a measurable set A of finite measure by a continuous function in $L^p(\mu)$. Let $\varepsilon > 0$. By the inner and outer regularity of a Radon measure on a σ compact metric space, there exists a compact set K and an open set U so that

$$K \subset A \subset U \qquad \mu(U) < \mu(K) + \varepsilon.$$

Then the distance between a compact set K and a disjoint closed set $X\backslash U$ is positive $d(K, X\backslash U) =: d_0 > 0$. For L sufficiently large such that $d_0 L \geqslant 1$, we define

$$f_L(x) = \max\{1 - Ld(x,K), 0\} = \begin{cases} 1 & \text{if } x \in K, \\ 1 - Ld(x,K) \in [0,1] & \text{if } x \in K_L, \\ 0 & \text{if } x \in X\backslash K_L, \end{cases}$$

where $K_L := \{x \in X | d(x,K) \leqslant \frac{1}{L}\} \subset U$ is compact. Then $f_L \in C_c(X)$ with $\operatorname{Supp}(f) = K_L$ and it approximates the characterisitic function χ_A in the sense of $L^p(\mu)$:

$$\|f_L - \chi_A\|_{L^p} \leqslant (\mu(U\backslash K))^{\frac{1}{p}} < \varepsilon^{\frac{1}{p}}.$$

Step 2. Lipschitz continuous functions with compact supports are dense. We prove that for every $\varepsilon > 0$ and $f \in C_c(X)$ with compact support, there exists g Lipschitz continuous with

$$\operatorname{supp} g \subset \operatorname{supp} f$$

and

$$\sup |g - f| < \varepsilon.$$

Then g approximates f also in $L^p(\mu)$-sense. It suffices to do this for $f \geqslant 0$. Since $\operatorname{supp} f$ is compact it is uniformly continuous: There exists $\delta > 0$ so that $|f(x) - f(y)| < \varepsilon$ if $d(x,y) < \delta$. With

$$\tilde{L} = \sup{}_{x,y} \left\{ \frac{|f(x) - f(y)|}{d(x,y)} : d(x,y) \geqslant \delta \right\},$$

which is finite since it is the supremum of a continuous function on a compact set, we obtain the inequality

$$|f(x) - f(y)| \leqslant \varepsilon + \tilde{L} d(x,y), \quad \forall x, y \in X.$$

We define

$$g(x) = \min_y \{f(y) + 2\tilde{L} d(x,y)\}.$$

One easily checks that

- The mimimum is attained in $B_\delta(x)$, since the values evaluated outside $B_\delta(x)$ is bigger than $f(x) - \tilde{L}d(x,y) + 2\tilde{L}d(x,y) \geqslant f(x) + \tilde{L}\delta$;
- $g \geqslant 0$ and takes value between $f - \varepsilon$ and f as

  $$\max\{0, f(x) - \varepsilon\} \leqslant g(x) \leqslant f(x).$$

  In particular, $\operatorname{Supp}(g) \subset \operatorname{Supp}(f)$ and $\sup |g - f| < \varepsilon$;

- g has Lipschitz constant $2\tilde{L}$, since for any $x_1, x_2 \in X$, if $g(x_1) = f(y_1) + 2\tilde{L}d(x_1, y_1)$ for some $y_1 \in B_\delta(x_1)$, then

  $$\begin{aligned} g(x_2) - g(x_1) &\leqslant f(x_2) - (f(y_1) + 2\tilde{L}d(x_1, y_1)) \\ &\leqslant f(y_1) + \tilde{L}d(x_2, y_1) - (f(y_1) + 2\tilde{L}d(x_1, y_1)) \leqslant 2\tilde{L}d(x_1, x_2), \end{aligned}$$

  if $d(x_2, y_1) \geqslant \delta$, and similarly for $d(x_2, y_1) < \delta$ (by bounding $g(x_2)$ by $f(y_1) + 2\tilde{L}d(x_2, y_1)$).

□

[12.12.2022]

[16.12.2022]

3.4.3 Arzela-Ascoli theorem and separability of $L^p(\mu)$

Theorem 3.35 (Arzela-Ascoli). *Let (X, d) be a compact metric space. Then a closed set $A \subset C_b(X)$ is compact if and only if*

1. *A is bounded.*
2. *A is equicontinuous, i.e. for $\varepsilon > 0$ there exists $\delta > 0$ so that*

   $$|f(x) - f(y)| < \varepsilon \qquad \text{if } f \in A \text{ and } d(x, y) < \delta.$$

Proof. Let A be compact. Since

$$C_b(X) \ni f \to \|f\|_{C_b(X)}$$

is continuous and hence attains its maximum in A, we deduce that A is bounded. Let $\varepsilon > 0$. For every continuous function $f \in C_b(X)$ defined on the compact metric space X, there exist $\delta_f > 0$ and an open neighborhood $U_f \subset C_b(X)$ so that

$$|g(x) - g(y)| < \varepsilon \qquad \text{if } g \in U_f \text{ and } d(x, y) < \delta_f.$$

Then $A \subset_{f \in A} U_f$, and since A is compact there is a finite subcovering, $A \subset_{j=1}^N U_{f_j}$. We define $\delta = \min \delta_{f_j}$.
Now assume that A is closed, bounded and equicontinuous. Let $(f_j) \subset A$ be a sequence. Let $\varepsilon > 0$ and $\delta > 0$ as in the second condition. Then there exist a finite number N of points x_k so that $B_\delta(x_k)$ cover X since X is compact. There exists a subsequence so that $f_{j_l}(x_k)$ converges for all

x_k. In particular, after relabeling, there are infinitely many $\{f_{j_l}\}_{l\in\mathbb{N}}$ so that $|f_{j_l}(x_k) - f_{j_m}(x_k)| < \varepsilon$. Then for any $x \in X$ there exists $B_\delta(x_k) \ni x$ such that

$$|f_{j_l}(x)-f_{j_1}(x)| \leqslant |f_{j_l}(x)-f_{j_l}(x_k)|+|f_{j_l}(x_k)-f_{j_1}(x_k)|+|f_{j_1}(x)-f_{j_1}(x_k)| < 3\varepsilon.$$

Hence $(f_{j_l}) \subset B_{3\varepsilon}(f_{j_1})$ and A is compact. □

Corollary 3.36. *If (X, d) is compact, then $C_b(X)$ is separable.*

Proof. By the proof of Theorem 3.34, the Lipschitz continuous functions are dense in $C_b(X)$. The countable union of separable sets is separable and its closure is separable. Hence it suffices to prove that

$$\{f \in C_b(X) : \|f(x)\|_{C_b(X)} \leqslant n, |f(x) - f(y)| \leqslant nd(x, y)\}$$

is separable. This set is compact by Theorem 3.35 and hence separable. □

Corollary 3.37. *Let (X, d) be σ compact metric space and μ a Radon measure. If $1 \leqslant p < \infty$ then $L^p(\mu)$ is separable.*

Proof. Since continuous functions with compact supports are dense, it suffices to show that $C_c(K_n)$ is separable in $L^p(\mu)$. It follows from Corollary 3.36. □

3.4.4 Compact sets in $L^p(\mathbb{R}^d)$

The space $\mathbb{R}^d$ is σ compact and the Lebesgue measure is σ finite Radon measure. We observe first the continuity of the translation operator on $L^p(\mathbb{R}^d)$, $p \in [1, \infty)$.

Lemma 3.38. *Suppose that $1 \leqslant p < \infty$, $f \in L^p(\mathbb{R}^d)$, $\varepsilon > 0$. Then there exist $\delta > 0$ and $R > 0$ so that for all $|h| < \delta$*

$$\|f(. + h) - f(.)\|_{L^p} < \varepsilon, \quad \|\chi_{\mathbb{R}^d\setminus B_R(0)} f\|_{L^p} < \varepsilon.$$

Proof. The second claim is a consequence of monotone convergence. For the first we approximate f by a Lipschitz continuous function g with compact support, $\|g - f\|_{L^p} < \varepsilon/4$ and estimate

$$\begin{aligned}\|f(. + h) - f\|_{L^p} \leqslant&\|f(. + h) - g(. + h)\|_{L^p} + \|f - g\|_{L^p} + \|g(\cdot + h) - g\|_{L^p}\\ \leqslant&\varepsilon/4 + \varepsilon/4 + |h|\|g\|_{Lip} \left(2m^d(\operatorname{supp} g)\right)^{1/p}\\ \leqslant&\varepsilon\end{aligned}$$

by choosing $|h| \leqslant \delta$ for some small δ. □

[16.12.2022]

[19.12.2022]

We hence can approximate $f \in L^p(\mathbb{R}^d)$, $p \in [1, \infty)$ by a sequence of smooth functions which are convolutions of f itself and the Dirac-sequence (which is a sequence of smooth functions with compact supports, and take values in $[0, 1]$ with the integral 1 on the whole space).

Corollary 3.39. *Let η be a radial smooth nonnegative function supported in the unit ball with unit integral: $\int_{\mathbb{R}^d} \eta dm^d = 1$, and let $\eta_r(x) = r^{-d}\eta(x/r)$, $r > 0$. Let $f \in L^p(\mathbb{R}^d)$, $p \in [1, \infty)$ and $f_r := \eta_r * f$ as in Young's Theorem 3.27. Then $f_r \in C^k(\mathbb{R}^d)$, $\forall k \in \mathbb{N}$, $\forall r > 0$, and for any $\varepsilon > 0$, there exists δ so that*

$$\|f_r - f\|_{L^p} < \varepsilon, \quad \forall r \leqslant \delta.$$

In particular $f_r \to f$ in L^p as $r \to 0$.

Proof. **Exercise (and give an example of such η).** It suffices to control $\|f_r - f\|_{L^p}$ by $\sup_{|h| \leqslant r} \|f(. + h) - f\|_{L^p}$. □

We characterize compact subsets of L^p spaces.

Theorem 3.40 (Kolmogorov)**.** *Let $1 \leqslant p < \infty$. A closed subset $C \subset L^p(\mathbb{R}^n)$ is compact iff*

1. *C is bounded.*
2. *For every $\varepsilon > 0$ there exists δ so that for all $|h| < \delta$ and all $f \in C$*

$$\|f(. + h) - f\|_{L^p(\mathbb{R}^d)} < \varepsilon$$

3. *For every $\varepsilon > 0$ there exists R so that for all $f \in C$*

$$\|\chi_{\mathbb{R}^d \setminus B_R(0)} f\|_{L^p} < \varepsilon.$$

Proof. Let C be compact. Since $f \to \|f\|_{L^p}$ is continuous it attains its maximum and hence C is bounded. We argue by contradiction to show the second and third parts. Suppose there exists $\varepsilon > 0$, $h_j \to 0$ and $f_j \in C$ so that

$$\|f_j(\cdot + h_j) - f_j\|_{L^p} \geqslant \varepsilon.$$

Since C is compact, there exists a subsequence $(f_{j_l}) \subset C$ which converges to the limit f in C. By Lemma 3.38 there exists $\delta > 0$ so that

$$\|f(\cdot + h) - f\|_{L^p} < \varepsilon/2$$

for $|h| < \delta$. This contradicts the previous inequality. Similarly we deduce the third part.
Vice versa: Suppose that $C \subset L^p(\mathbb{R}^d)$ is closed, bounded, and satisfies the three claims. Let ε, δ, R be as in the claims. Since a closed set in a Banach space is complete, it suffices to show the existence of finite 10ε-covering for C (i.e. there are finite covering balls of radius 10ε for C). By the third claim, the set $C_1 := \{f\chi_{B_{R+2\delta}(0)} \mid f \in C\}$ lies in ε-neighborhood of C, and vice versa. Hence it suffices to show the existence of finite 4ε-covering in $L^p(\mathbb{R}^d)$ for C_1. We claim that the set $C^\delta := \{f^\delta := \eta_\delta * f_1 \mid f_1 \in C_1\}$ is in 3ε-neighborhood of C_1. This claim follows from the following inequality (similarly as in the proof of Corollary 3.39)

$$\begin{aligned}
&\|\eta_\delta * (f\chi_{B_{R+2\delta}(0)}) - f\chi_{B_{R+2\delta}(0)}\|_{L^p} \\
&= \left\| \int_{B_1(0)} \big(f\chi_{B_{R+2\delta}(0)}(x-\delta z) - f1_{B_{R+2\delta}(0)}(x)\big)\eta(z)dz \right\|_{L^p_x} \\
&\leqslant \left\| \int_{B_1(0)} \big(f(x-\delta z) - f(x)\big)\eta(z)dz \right\|_{L^p_x(B_{R+\delta}(0))} \\
&\quad + \left\| \int_{\mathbb{R}^d} \big(f\chi_{B_{R+2\delta}(0)}(x-\delta z) - f\chi_{B_{R+2\delta}(0)}(x)\big)\eta(z)dz \right\|_{L^p_x(\mathbb{R}^d \setminus B_{R+\delta}(0))} \\
&\leqslant \sup{}_{|h|<\delta} \|f(\cdot + h) - f(\cdot)\|_{L^p} + 2\|f\chi_{\mathbb{R}^d \setminus B_R(0)}\|_{L^p}.
\end{aligned}$$

Hence it suffices to show the existence of ε-covering in $L^p(\mathbb{R}^d)$ for C^δ. By Corollary 3.39, the approximated functions f^δ are Lipschitz continuous with compact supports in $B_{R+3\delta}(0)$, and all the functions in C^δ have uniform supremum norm, Lipschitz constant and support. By Theorem 3.35 the set $\overline{C^\delta}$ is compact in $C_b(B_{R+3\delta}(0))$ and hence in $L^p(\mathbb{R}^d)$, and we can cover it by a finite ε-covering in $L^p(\mathbb{R}^d)$. $\square$

3.4.5 Riesz representation theorem for $C_c(X)$

Let $X = [0,1] \subset \mathbb{R}$ be a finite interval, and $C_b(X) = C_c(X) = C([0,1])$ be the set of all continuous functions on $[0,1]$, which is a separable Banach space. The ordinary Riemann integral

$$L(f) = \int_0^1 f(x)dx$$

- defines a linear map (also called linear functional) from $C_b(X)$ to $\mathbb{K}$;
- satisfies $L(f) \geqslant 0$ if $f \geqslant 0$;

- is bounded $|L(f)| \leqslant \|f\|_{C_b(X)}$.

We observe that for any open interval $(a,b) \subset [0,1]$, and any continuous function f with compact support inside (a,b) and with value $f \in [0,1]$, it holds

$$L(f) < b - a,$$

and since we can choose f such that $L(f)$ approximates $b-a$,

$$b - a = \sup\{L(f) : f \in C_c(X), \quad \mathrm{Supp}\,(f) \subset (a,b), \quad 0 \leqslant f \leqslant 1\}.$$

This reveals a connection between the (Lebesgue) measure of (a,b) and the values of the nonnegative linear map L on continuous functions with support in (a,b) and values in $[0,1]$.

[19.12.2022]

[23.12.2022]

More generally, let (X,d) be metric space, and let $L : C_c(X) \mapsto \mathbb{K}$ be a nonnegative linear map, then we define the variation measure of L by

$$\mu^*(U) = \sup\{L(f) : f \in C_c(X), \, \mathrm{supp}\, f \subset U, \, 0 \leqslant f \leqslant 1\}$$

for open sets U and for general sets

$$\mu^*(A) = \inf\{\mu^*(U) : A \subset U, U \text{ open}\}.$$

If we assume further the boundedness for L in the sense that for any compact set K, there exists a constant $C_K > 0$ such that

$$|L(f)| \leqslant C_K \|f\|_{C_b(X)} \tag{3.11}$$

holds for all $f \in C_c(X)$ with $\mathrm{Supp}\, f \subset K$, we have the following representation of this nonnegative linear map on $C_c(X)$ by the Lebesgue integral with respect to the Radon measure μ induced by μ^*.

Theorem 3.41 (Riesz representation theorem for nonnegative linear functionals on $C_c(X)$)**.** *Let (X,d) be a σ compact metric space and let the nonnegative linear map $L : C_c(X) \to \mathbb{K}$ satisfy* (3.11)*. Then there exists a unique Radon measure μ on (X,d) so that for all $f \in C_c(X)$,*

$$L(f) = \int_X f d\mu.$$

Remark 3.42. *One can relax the assumption that X is a σ compact metric space to that X is a locally compact Hausdorff space (with Radon measure defined more generally), and the condition* (3.11) *can be removed: The proof will be more delicate. In probability some interesting measures (e.g. the so-called Wiener measure) occur on topological spaces that are not locally compact, and Riesz representation theorem is not applicable.*
The reverse holds obviously: For any Radon measure μ on a σ-compact metric space X, the integral $L(f) := \int_X f d\mu$ defines a nonnegative linear map from $C_c(X)$ to $\mathbb{K}$ satisfying (3.11)*. That is, there exists a bijective map*

$$J : \{\textit{Radon measure } \mu \textit{ on } X\} \ni \mu \mapsto L \in \{L : C_c(X) \to \mathbb{K} \,|\, L \textit{ is nonnegative linear map}, \\ L \textit{ is bounded on } C_c(K),\ \forall K \subset X \textit{compact}\}$$

and for all open sets U,

$$\mu(U) = \sup\{L(f) \,|\, f \in C_c(X),\ \mathrm{supp}(f) \subset U,\ 0 \leqslant f \leqslant 1\}.$$

Notice that in general it is not natural to assume that L is a continuous/bounded linear map from $C_b(X)$ to $\mathbb{K}$: For example, for the Lebesgue measure on $\mathbb{R}^d$, the Lebesgue integral $\int_{\mathbb{R}^d} f dm^d$ is not bounded on the unit closed ball $\{\|f\|_{C_b(\mathbb{R}^d)} \leqslant 1\}$.
The Lebesgue measure on $\mathbb{R}^d$ can be constructed from the Lebesgue integral as in the theorem.

Proof. **Step 1: μ^* is an outer metric measure, and its restriction μ on Caratheodory measurable sets and then on Borel sets is a Radon measure.** It is easy to see that $\mu^*(A) \leqslant \mu^*(B)$ if $A \subset B$. To show that μ^* is an outer measure, let U_j be open sets, $U = \bigcup_j U_j$ and we have to show that

$$\mu^*(U) \leqslant \sum_j \mu^*(U_j),$$

which is equivalent to show for all $0 \leqslant f \leqslant 1$ with $\mathrm{Supp}\, f \subset U$,

$$L(f) \leqslant \sum_j \mu^*(U_j).$$

Let $K = \mathrm{Supp}\, f$ which is compact. Thus K is covered by finitely many $\bigcup_{j=1}^N U_j$ for some $N < \infty$. We claim that there exist g_j, $0 \leqslant g_j \leqslant 1$, $\mathrm{Supp}\, g_j \subset U_j$ and $\sum_{j=1}^N g_j = 1$ on K. **(Exercise.)** We define $f_j = g_j f$. Then $0 \leqslant f_j \leqslant 1$, $f = \sum_{j=1}^N f_j$ and

$$L(f) = \sum_{j=1}^N L(f_j) \leqslant \sum_{j=1}^N \mu^*(U_j).$$

For general set $A = \bigcup_j A_j$, for any $\varepsilon > 0$, there exist (U_j) such that $\mu^*(U_j) < \mu^*(A_j) + \varepsilon 2^{-j}$, and hence $A \subset \bigcup U_j$ and

$$\mu^*(A) \leqslant \mu^*(\bigcup_j U_j) \leqslant \sum_j \mu^*(U_j) \leqslant \sum_j \mu^*(A_j) + \varepsilon.$$

Thus $\mu^*(A) \leqslant \sum_j \mu^*(A_j)$.
To show that μ^* is an outer metric measure, let A, B be two disjoint sets with positive distance. If A, B are both open sets, then $\mu^*(A \cup B) \geqslant \mu^*(A) + \mu^*(B)$, since for all $f \in C_c(A; [0,1])$ and $g \in C_c(B; [0,1])$ the sum $f + g \in C_c(A \cup B; [0,1])$. For general disjoint sets A, B, there exist disjoint open sets V and W containing A resp. B. Then for all open set $U \supset A \cup B$, the two open sets $U \cap V$ and $U \cap W$ are disjoint so that

$$\mu^*(U) \geqslant \mu^*(U \cap V) + \mu^*(U \cap W) \geqslant \mu^*(A) + \mu^*(B).$$

Thus $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. Similarly it follows $\mu^*(\varnothing) = 0$.
Let μ be the measure constructed from the outer metric measure μ^* by the Caratheodory Theorem 3.4. Then

$$\mu(B) = \mu^*(B)$$

for Borel sets. By (3.11), μ is bounded on compact sets: For any $K \subset X$ compact, there exist an open set V and a compact set K_0 such that $K_0 \supset V \supset K$ (since X is σ compact metric space), and hence for any $f \in C_c(X; [0,1])$ with $f = 1$ on $\overline{V}$ and $\mathrm{supp}(f) \subset K_0$,

$$\mu(K) \leqslant \mu(V) \leqslant L(f) \leqslant C_{K_0} \|f\|_{C_b(X)} = C_{K_0} < \infty.$$

Thus by Lemma 3.32 its restriction to Borel sets is a Radon measure.
Step 2: Representation of the renormalized map $\Lambda(f)$ by the Lebesgue integral. [5] Let $f \in C_c(X)$ be nonnegative. We define

$$\Lambda(f) = \sup \{|L(g)| : g \in C_c(X), |g| \leqslant f\} \geqslant 0.$$

Clearly $0 \leqslant f_1 \leqslant f_2$ implies $\Lambda(f_1) \leqslant \Lambda(f_2)$ and for $c > 0$, $\Lambda(cf) = c\Lambda(f)$. We claim that

$$\Lambda(f_1 + f_2) = \Lambda(f_1) + \Lambda(f_2)$$

for $f_1, f_2 \in C_c(X)$ nonnegative. Indeed, if $|g_1| \leqslant f_1$ and $|g_2| \leqslant f_2$ then $|g_1 + g_2| \leqslant f_1 + f_2$, and, if in addition $L(g_1), L(g_2) \in [0, \infty)$,

$$|L(g_1) + L(g_2)| = |L(g_1)| + |L(g_2)| \leqslant \Lambda(f_1 + f_2).$$

[5] We can simply take Λ to be L itself, when it applies on nonnegative functions. For general case, we notice by linearity that $L(f) = L(\mathrm{Re}\, f) + iL(\mathrm{Im}\, f)$ and for real-valued function f, $L(f) = L(f^+) - L(f^-)$.

Since for any $g \in C_c(X)$, there exists $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ so that $|g| = |\alpha g|$ and $|L(g)| = \alpha L(g) = L(\alpha g)$, the above gives

$$|L(g_1)| + |L(g_2)| = L(\alpha_1 g_1) + L(\alpha_2 g_2) \leqslant \Lambda(f_1 + f_2),$$

and hence $\Lambda(f_1) + \Lambda(f_2) \leqslant \Lambda(f_1 + f_2)$. Now let $|g| \leqslant f_1 + f_2$. We define

$$g_1 = \begin{cases} \frac{f_1 g}{f_1+f_2} & \text{if } f_1 + f_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$

and similarly g_2. Then $|g_i| \leqslant f_i$ and hence

$$|L(g)| = |L(g_1) + L(g_2)| \leqslant \Lambda(f_1) + \Lambda(f_2)$$

which gives

$$\Lambda(f) = \Lambda(f_1) + \Lambda(f_2).$$

We claim that

$$\Lambda(f) = \int_X f d\mu.$$

It suffices to consider $f \in C_c(X)$ such that $0 \leqslant f \leqslant 1$. We approximate f by step function from below so that

$$0 \leqslant f - \frac{1}{N}\sum_{j=1}^{N-1} \chi_{U_j} \leqslant \frac{1}{N} \tag{3.12}$$

with

$$U_j = \{x : f(x) > \frac{j}{N}\}.$$

By continuity $U_{j+1} \subset U_j$. We approximate the characteristic function χ_{U_j} by continuous and compactly supported functions η_j so that $\operatorname{supp}\eta_j \subset U_{j-1}$, $\eta_j = 1$ on U_j, $\frac{1}{N}\sum_{j=1}^{N-1}\eta_j \leqslant f$ and $\mu(\operatorname{supp}\eta_j \backslash U_j) < 1/N$. Hence f can be approximated by $\frac{1}{N}\sum_{j=1}^{N-1}\eta_j$:

$$0 \leqslant f - \frac{1}{N}\sum_{j=1}^{N-1}\eta_j = \big(f - \frac{1}{N}\sum_{j=1}^{N-1}\chi_j\big) - \frac{1}{N}\sum_{j=1}^{N-1}\big(\eta_j - \chi_{U_j}\big) \leqslant \frac{1}{N}.$$

Since

$$\left|\Lambda(f) - \frac{1}{N}\sum_{j=1}^{N-1}\int_X \eta_j d\mu\right| \leqslant \left|\Lambda(f) - \frac{1}{N}\sum_{j=1}^{N-1}\Lambda(\eta_j)\right| + \frac{1}{N}\sum_{j=1}^{N-1}\left|\Lambda(\eta_j) - \int_X \eta_j d\mu\right|$$
$$\leqslant C_{\operatorname{supp}(f)}\frac{1}{N} + \sup{}_j \left|\Lambda(\eta_j) - \int_X \eta_j d\mu\right|,$$

it suffice to verify

$$\left|\Lambda(\eta_j) - \int_X \eta_j d\mu\right| \leqslant 1/N,$$

which follows from $\mu(\operatorname{supp}\eta_j \backslash U_j) < 1/N$ and

$$\mu(U_j) \leqslant \Lambda(\eta_j) \leqslant \mu(\operatorname{supp}\eta_j), \quad \mu(U_j) \leqslant \int_X \eta_j d\mu \leqslant \mu(\operatorname{supp}\eta_j).$$

Step 3: Final check. Now

$$|L(f)| \leqslant \Lambda(|f|) = \int_X |f| d\mu, \quad \forall f \in C_c(X).$$

Since $C_c(X)$ is dense in $L^1(\mu)$, we can extend L to an element in $(L^1(\mu))^*$: For any $g \in L^1(\mu)$, there exists a sequence $(f_n) \subset C_c(X)$ so that $f_n \to g$ in $L^1(\mu)$, and we define

$$L(g) = \lim_{n\to\infty} L(f_n).$$

This is welldefined since $|L(f_n - f_m)| \leqslant \int_X |f_n - f_m| d\mu$. It also holds $\|L\|_{(L^1(\mu))^*} \leqslant 1$ since

$$|L(g)| = \lim_{n\to\infty} |L(f_n)| \leqslant \lim_{n\to\infty} \int_X |f_n| d\mu = \int_X |g| d\mu.$$

By Corollary 3.25, there exists $\sigma \in L^\infty(\mu)$ such that $L(f) = \int_X f\sigma d\mu$ for $f \in L^1(\mu)$ and

$$\|\sigma\|_{L^\infty(\mu)} = \|L\|_{(L^1(\mu))^*} \leqslant 1.$$

We claim that $\sigma = 1$ almost everywhere. By definition

$$\mu(U) = \sup\left\{\int_X f\sigma d\mu = L(f) : f \in C_c(X), 0 \leqslant f \leqslant 1, \operatorname{supp} f \subset U\right\}.$$

We choose a sequence of functions with

$$\int_X f_j \sigma d\mu = L(f_j) \to \mu(U).$$

Since $\int_X f_j\sigma d\mu \leqslant \int_U |\sigma| d\mu \leqslant \mu(U)$, we deduce $\sigma = 1$ almost everywhere. To show the uniqueness of the Radon measure, it suffices to show that the measure of any compact set is unique: Let μ_1, μ_2 be two corresponding Radon measures. For any compact set K and open set $U \supset K$, there exists a function $f \in C_c(U; [0,1])$ with $f = 1$ on K, so that

$$\mu_1(K) \leqslant \int_X f d\mu_1 = L(f) = \int_X f d\mu_2 \leqslant \mu_2(U),$$

and hence $\mu_1(K) \leqslant \mu_2(K)$, which implies $\mu_1 = \mu_2$. □

3.4.6 Radon-Nikodym Theorem

In this section we restrict ourselves to the Lebesgue measure space.[6] The space $\mathbb{R}^d$ is σ compact and the Lebesgue measure is σ finite Radon measure.

Definition 3.43. *Let μ and ν be Radon measures on $\mathbb{R}^d$. For $x \in \mathbb{R}^d$ we define*

$$\overline{D_\mu}\nu(x) = \begin{cases} \limsup_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \text{ for all } r > 0 \\ \infty & \text{if for some } r > 0, \mu(B_r(x)) = 0 \end{cases}$$

$$\underline{D_\mu}\nu(x) = \begin{cases} \liminf_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \text{ for all } r > 0 \\ \infty & \text{if for some } r > 0, \mu(B_r(x)) = 0. \end{cases}$$

We say that ν is differentiable with respect to μ and x if $\overline{D_\mu}\nu(x) = \underline{D_\mu}\nu(x)$, for which we write $D_\mu\nu(x)$ and we call it the density of ν with respect to μ. We say the measure ν is absolutely continuous with respect to μ, denoted by $\nu << \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$ for any Borel set A. We say that the measure ν is singular with respect to μ, denoted by $\nu \perp \mu$, if there exists a Borel set B such that $\mu(X\backslash B) = \nu(B) = 0$.

Remark 3.44. *Let $f \in C_c(\mathbb{R}^d)$ and μ Radon measure. Then*

$$x \to \int f(y-x)d\mu(y)$$

is continuous and hence Borel measurable.
Since the characteristic function of open and closed balls can be obtained as pointwise limit of continuous functions with compact support, the map

$$x \to \nu(B_r(x)), \quad x \to \mu(B_r(x))$$

are measurable. Thus

$$x \to \begin{cases} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \\ \infty & \text{if } \mu(B_r(x)) = 0 \end{cases}$$

is Borel measurable. The map

$$r \to \begin{cases} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \\ \infty & \text{if } \mu(B_r(x)) = 0 \end{cases}$$

[6] This section is for interested readers, and the content will not be taken into account in the exam.

is continuous from the left and right if $\mu(B_r(x)) > 0$ by inner and outer regularity. Thus also $\overline{D_\mu}\nu(x)$ and $\underline{D_\mu}\nu(x)$ are Borel measurable since we can write them as inf's and sup's over rational radii. Moreover by inner and outer regularity we obtain the same $\overline{D_\mu}\nu(x)$ and $\underline{D_\mu}\nu(x)$ if we use closed balls.

Theorem 3.45 (Radon-Nikodym)**.** *Let ν and μ be Radon measures on $\mathbb{R}^d$. Then*

- *$D_\mu\nu(x)$ exists and is finite μ almost everywhere.*
- *If $\nu << \mu$, then for all Borel set A,*

$$\nu(A) = \int_A D_\mu\nu d\mu.$$

Proof. It suffices to consider the case $\mu(\mathbb{R}^d) < \infty$ and $\nu(\mathbb{R}^d) < \infty$. We claim without proof that for two Radon measures μ, ν on $\mathbb{R}^d$, for all Borel sets B and all $t > 0$,

$$\nu(B \cap \{x : \underline{D_\mu}\nu(x) < t\}) \leqslant t\mu(B \cap \{x : \underline{D_\mu}\nu(x) < t\})$$

and

$$\nu(B \cap \{x : \underline{D_\mu}\nu(x) > t\}) \geqslant t\mu(B \cap \{x : \underline{D_\mu}\nu(x) > t\}).$$

Similarly for $\overline{D_\mu\nu}$. It follows from a covering theorem in $\mathbb{R}^d$ and regularity of Radon measures.

We now show the existence of $D_\mu\nu(x)$. For $s < t$ we define

$$R(s,t) = \{x : \underline{D_\mu}\nu(x) < s < t < \overline{D_\mu}\nu(x)\}.$$

Then

$$t\mu(R(s,t)) \leqslant \nu(R(s,t)) \leqslant s\mu(R(s,t)),$$

which implies $\mu(R(s,t)) = 0$. Since

$$\{x : \underline{D_\mu}\nu(x) < \overline{D_\mu}\nu(x)\} = \bigcup_{s<t, s,t\in\mathbb{Q}} R(s,t),$$

we see that $\underline{D_\mu}\nu(x) = \overline{D_\mu}\nu(x)$ for μ almost all x. To show that $D_\mu\nu$ is finite μ-a.e., let $B = \{x : \overline{D_\mu}\nu(x) = \infty\}$. Then $\mu(B) = 0$ follows from

$$\nu(B) \geqslant t\mu(B), \quad \forall t > 0.$$

Now let $\nu << \mu$, then

$\nu(\{D_\mu\nu(x) = \infty\}) = \nu(\{D_\mu\nu(x) = 0\}) = \mu(\{D_\mu\nu(x) = \infty\}) = \mu(\{D_\mu\nu(x) = 0\}) = 0.$

Let A be a Borel set. For $t > 1$ we define

$$A_n = A \cap \{t^n \leqslant D_\mu\nu < t^{n+1}\}.$$

Then

$$\nu(A) = \sum_{m=-\infty}^{\infty} \nu(A_m) \leqslant \sum_{m=-\infty}^{\infty} t^{m+1}\mu(A_m) \leqslant t\int_0^\infty \mu(\{D_\mu\nu > s\})ds = t\int_A D_\mu\nu d\mu$$

and

$$\nu(A) = \sum_{m=-\infty}^{\infty} \nu(A_m) \geqslant \sum_{m=-\infty}^{\infty} t^{m}\mu(A_m) \geqslant t^{-1}\int_0^\infty \mu(\{D_\mu\nu > s\})ds = t^{-1}\int_A D_\mu\nu d\mu.$$

We let now $t \to 1$. □

Corollary 3.46 (Lebesgue points)**.** *Let μ be a Radon measure on $\mathbb{R}^d$ and $\tilde{f} \in L^1_{loc}(\mu)$. Then*

$$f(x) := \lim_{r\to 0} \mu(B_r(x))^{-1} \int_{B_r(x)} \tilde{f} d\mu$$

exists almost everywhere and we define $f(x) = 0$ if it does not exist. Then f is in the equivalence class of $\tilde{f}$.

Proof. It suffices to consider nonnegative $\tilde{f}$ and $\mu(\mathbb{R}^d) < \infty$. We define

$$\nu(A) = \int_A \tilde{f}\, d\mu.$$

This is a Radon measure which is absolutely continuous with respect to μ. Thus

$$\nu(A) = \int_A D_\mu\nu d\mu = \int_A \tilde{f} d\mu$$

and $D_\mu\nu$ lies in the equivalence class. Now the first claim follows from Theorem 3.45.

□

Remark 3.47. *Similarly if $\tilde{f} \in L^p_{loc}(\mu)$, $p \in [1, \infty)$, then $f \in L^p_{loc}$ and almost everywhere*

$$\lim_{r\to 0} \mu(B_r(x))^{-1} \int_{B_r(x)} |f(y) - f(x)|^p d\mu(y) = 0.$$

Then there exists a cannonical representative of the equivalence class.

[23.12.2022]

[09.01.2023]

4 Distributions and Sobolev spaces

In this chapter we restrict ourselves to Euclidean space $\mathbb{R}^d$, and we are going to introduce distributions and Sobolev spaces defined on open sets in $\mathbb{R}^d$. Recall the Riesz representation theorem describing the (continuous) dual space X^* consisting of (continuous) linear functionals $x^* \in X^*$ [7] from some topological $\mathbb{K}$-vector space X to $\mathbb{K}$:

- Theorem 2.25 for Hilbert spaces H: Any linear functional on H can be uniquely identified with an element in H itself, in the sense that $J : H \ni x \mapsto (y \mapsto \langle y, x\rangle) \in H^*$ is an antilinear isomorphism (isometric bijective map).

- Theorem 3.24 (and Corollary 3.25) for the Lebesgue spaces $L^p(\mathbb{R}^d)$, $1 \leqslant p < \infty$: Any function in $L^{p'}(\mathbb{R}^d)$ can be uniquely identified with a linear functional in $(L^p(\mathbb{R}^d))^*$, in the sense that $j : L^{p'}(\mathbb{R}^d) \ni g \mapsto (f \mapsto \int_{\mathbb{R}^d} fg dm^d) \in (L^p(\mathbb{R}^d))^*$ is a linear isomorphism.

  Here the Lebesgue spaces read

  $$L^p(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{K} \quad \text{is } p-\text{integrable}\}/\sim, \quad 1 \leqslant p \leqslant \infty,$$

  equipped with the norm $\|f\|_{L^p}$, where $\sim$ denotes the equivalence relation between two measurable functions which differ only on a measure-zero set. In particular, the Lebesgue space $L^2(\mathbb{R}^d)$ is a Hilbert space, equipped with the inner product $\langle f, g\rangle = \int_{\mathbb{R}^d} f\bar{g} dm^d$.

- Theorem 3.41 for the continuous compactly supported function space $C_c(\mathbb{R}^d)$: Any <u>nonnegative</u> continuous [8] linear map $L \in \big(C_c(\mathbb{R}^d)\big)^*$ is represented uniquely by a Radon measure μ on $\mathbb{R}^d$, in the sense that $L(f) = \int_{\mathbb{R}^d} f d\mu$ holds for all $f \in C_c(\mathbb{R}^d)$.

  Here the continuous bounded function space

  $$C_b(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{K} \text{ continuous and bounded}\},$$

  is equipped with the norm $\|f\|_{C_b} = \sup_{x\in\mathbb{R}^d}|f(x)|$, and its subspace

  $$C_c(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{K} \text{ continuous and with compact support}\},$$

[7] In the following we also denote by X' the dual space of X.

[8] If $f_n \to f$ in $C_c(\mathbb{R}^d)$ means that there exists a compact set $K \subset \mathbb{R}^d$ such that $\text{supp}(f_n) \subset K$ for all n and $\|f_n - f\|_{C_b(\mathbb{R}^d)} \to 0$, then a continuous map $L : C_c(\mathbb{R}^d) \mapsto \mathbb{K}$ means that $L(f_n) \to L(f)$ holds for all $f_n \to f$ in $C_c(\mathbb{R}^d)$. Thus a continuous linear map on $C_c(\mathbb{R}^d)$ is bounded on $C_b(K)$ for all compact set $K \subset \mathbb{R}^d$ (with the proof the same as Lemma 4.9 below).

is a vector subspace but not a closed subset of $C_b(\mathbb{R}^d)$. $C_c(\mathbb{R}^d)$ is a dense subset in $L^p(\mathbb{R}^d)$, $1 \leqslant p < \infty$, and any nonnegative $L^{p'}(\mathbb{R}^d)$-function g can then be viewed as a Radon measure gdm^d on $\mathbb{R}^d$, which defines (uniquely) a (continuous) linear functional on $C_c(\mathbb{R}^d)$.

This motivates to generalize the definition of "functions" as continuous linear maps on the test function space $C_c^\infty(\mathbb{R}^d)$. The vector space $C_c^\infty(\mathbb{R}^d)$ is not a Banach space, and its topology is complicated, which we do not discuss in detail here.
We start with some celebrated theorems for continuous linear maps between two Banach spaces in Section 4.1, and we will see that the ideas can apply in Section 4.2.

4.1 Baire category theorem, Banach-Steinhaus theorem and open mapping theorem

Theorem 4.1 (Baire category theorem)**.** *A countable intersection of dense open subsets of a complete metric space is dense.*

Proof. Let (X, d) be a complete metric space and A_j open dense sets. Let $x \in X$ and $\varepsilon > 0$. Take $x_1 \in A_1$ so that

$$d(x, x_1) < \delta_0 := \varepsilon,$$

and take $\delta_1 > 0$ with $2\delta_1 < \delta_0$ such that

$$B_{2\delta_1}(x_1) \subset A_1 \cap B_{\delta_0}(x).$$

Recursively given $(x_k, \delta_k)_{k=1}^{n-1}$, there are $x_n \in A_n$, $\delta_n > 0$ with $2\delta_n < \delta_{n-1}$ such that

$$d(x_{n-1}, x_n) < \delta_{n-1}, \quad B_{2\delta_n}(x_n) \subset A_n \cap B_{\delta_{n-1}}(x_{n-1}).$$

We arrive at a sequence (δ_n):

$$\delta_n < \frac{1}{2}\delta_{n-1} < \cdots < 2^{-n}\delta_0 = 2^{-n}\varepsilon,$$

and a sequence (x_n) such that for all $m > n$,

$$d(x_n, x_m) \leqslant \sum_{j=0}^{m-n-1} d(x_{n+j}, x_{n+j+1}) < \sum_{j=0}^{m-n-1} \delta_{n+j} < 2^{-n+1}\varepsilon,$$

$$\overline{B_{\delta_m}(x_m)} \subset B_{2\delta_m}(x_m) \subset A_m \cap B_{\delta_{m-1}(x_{m-1})} \subset \cdots \subset \bigcap_{j=n}^{m} A_j \cap B_{\delta_{n-1}}(x_{n-1}).$$

Thus (x_n) is a Cauchy sequence, which converges to a limit y, such that $y \in \overline{B_{\delta_n}(x_n)}$, $\forall n$, and

$$d(x,y) \leqslant d(x,x_1) + d(x_1,y) < 2\varepsilon, \quad y \in \bigcap_j A_j.$$

Thus $\bigcap_j A_j$ is dense in (X,d).

□

Recall the definition of Banach spaces (i.e. complete normed spaces) in Section 1.1, and the set $L(X,Y)$ consisting of all continuous linear maps between two normed spaces X, Y in Section 1.3.

Theorem 4.2 (Banach-Steinhaus, Principle of uniform boundedness)**.** *Let X and Y be Banach spaces, $\mathcal{F} \subset L(X,Y)$. If for each $x \in X$*

$$\sup\{\|Tx\|_Y : T \in \mathcal{F}\} < \infty,$$

then

$$\sup\{\|T\|_{X\to Y} : T \in \mathcal{F}\} < \infty.$$

In particular, for a sequence of continuous linear maps $\{T_n\} \subset L(X,Y)$ such that

$$T_n x \text{ converges in } Y, \quad \forall x \in X,$$

we have

$$\sup{}_n\{\|T_n\|_{X\to Y}\} < \infty.$$

Proof. Let

$$C_n = \{x \in X : \sup{}_{T\in\mathcal{F}}\|Tx\|_Y \leqslant n\}.$$

This set is closed since both T and $\|\|$ are continuous, and C_n is an intersection of closed sets. By assumption $\bigcup C_n = X$. We claim that some C_n has nonempty open interior. If not then the sets $U_n = X \backslash C_n$ are open and dense, and hence with nonempty intersection by Lemma 4.1. This is a contradiction to $\bigcup C_n = X$. Let $B_r(x_0) \subset C_{n_0}$ be an open ball. If $\|x\| < r$ then

$$\|Tx\|_Y \leqslant \|T(x-x_0)\|_Y + \|T(x_0)\|_Y \leqslant n_0 + \sup{}_{T\in\mathcal{F}}\|T(x_0)\|_Y =: R.$$

Then

$$\|T\|_{X\to Y} \leqslant R/r$$

for all $T \in \mathcal{F}$. □

Remark 4.3. *The assumption that Y is Banach space can be relaxed to normed space.*
Obviously for normed spaces X, Y, the uniform boundedness of the operator norms

$$\sup_{T \in \mathcal{F}} \{\|T\|_{X \to Y}\} \leqslant c_0 < \infty$$

implies the (uniform) boundedness of the application on some fixed element $x \in X$:

$$\sup_{T \in \mathcal{F}} \{\|Tx\|_Y\} \leqslant c_0 \|x\|_X < \infty.$$

The two boundedness are equivalent if X is Banach spaces. Given an example to show the failure of the equivalence if X is not complete. ***(Exercise)***

[09.01.2023]

[13.01.2023]

The Baire category theorem has further interesting consequences. The following open mapping theorem shows the equivalence between surjectivity and open property of a continuous linear map between two Banach spaces.

Theorem 4.4 (Open mapping theorem)**.** *Let X and Y be Banach spaces and $T \in L(X, Y)$. T is surjective if and only if it is open, i.e. if the image of open sets is open.*

Proof. Let T be open. Then $T(B_1^X(0))$ is open. In particular it contains a ball $B_r^Y(0)$. Then $Y = \bigcup_n B_n^Y(0) = \bigcup_n T(B_{n/r}^X(0))$ and T is surjective.
Now suppose that T is surjective. It suffices to show that $T(B_1^X(0))$ contains some ball $B_\varepsilon^Y(0)$, since for any open ball $U \subset X$ and any $x \in U$, there exists $B_\delta^X(x) \subset U$ such that $B_{\varepsilon\delta}^Y(Tx) \subset TU$:

$$B_{\varepsilon\delta}^Y(Tx) = Tx + B_{\varepsilon\delta}^Y(0) = Tx + \delta B_\varepsilon^Y(0) \subset Tx + \delta T(B_1^X(0)) = T(B_\delta^X(x)) \subset TU.$$

Let

$$Y_n := \overline{T(B_n^X(0))} = \overline{\{Tx | \|x\|_X < n\}} = \overline{n\{Tx \mid \|x\|_X < 1\}} = nY_1.$$

It is closed, $Y_n \subset Y_{n+1}$ and $Y = \bigcup Y_n$. As in the proof of Theorem 4.2 we conclude that one (and hence all) of the Y_n contains an open ball. Hence there exists $B_r^Y(y_0) \subset Y_1$, and thus $B_r^Y(0) \subset Y_2$, since

$$\begin{aligned} y \in B_r^Y(0) &\Longrightarrow y + y_0 \in B_r^Y(y_0) \subset Y_1 = \overline{T(B_1^X(0))} \\ &\Longrightarrow y \in \overline{T(B_1^X(0))} - \overline{T(B_1^X(0))} \subset 2\overline{T(B_1^X(0))} = Y_2. \end{aligned}$$

In particular, $B_r^Y(0) \subset T(B_4^X(0))$, since for any $y \in Y_2 = \overline{TB_2^X(0)}$,

$$\exists x_1 \in B_2^X(0) \text{ s.t. } y - Tx_1 \in B_{2^{-1}r}^Y(0) \subset \overline{TB_1^X(0)},$$

and inductively

$$\exists x_n \in B_{2^{2-n}}^X(0) \text{ s.t. } y - Tx_1 - \cdots - Tx_n \in B_{2^{-n}r}^Y(0) \subset \overline{TB_{2^{1-n}}^X(0)},$$

which implies $y = Tx$ with $x := \sum_n x_n \in B_4^X(0)$. □

The following corollary implies that the inverse map of an invertible continuous linear map is also continuous linear map.

Corollary 4.5. *Let X, Y be two Banach spaces. Suppose that $T \in L(X, Y)$ is bijective. Then $T^{-1} \in L(Y, X)$.*

Proof. Linearity of the inverse map follows immediately from the linearity and injectivity of T. By Theorem 4.4 T is open. So $T(B_1^X(0))$ contains a ball $B_r^Y(0)$ and hence

$$\forall y \in B_r^Y(0),\ \exists x \in B_1^X(0) \text{ s.t. } y = Tx, \text{ that is, } \|x\|_X = \|T^{-1}y\|_X \leqslant \frac{1}{r}\|y\|_Y.$$

T^{-1} is a linear and bounded map. □

Corollary 4.6. *The set of nowhere differentiable functions in $C_b(0,1)$ is dense.*

Proof. **Exercise.** □

4.2 Distributions

In this section, we are going to introduce the definition of the distribution space first, and then to generalize the definitions of product, differentiation, support and convolution from test function space to distribution space by the idea of duality. Finally, we are going to approximate a distribution by test functions.

4.2.1 Definitions

Definition 4.7 (Test function space and distribution space)**.** *Let $U \subset \mathbb{R}^d$ be open.*
We denote by $\mathcal{D}(U) = C_c^\infty(U)$ the vector space of all test functions on U, which are infinitely differentiable functions with compact support in U.

We call $\varphi_j \to \varphi$ *in* $C_c^\infty(U) = \mathcal{D}(U)$ *if there is a compact set* $K \subset U$ *so that* $\operatorname{supp}\varphi_j \subset K$ *for all* j*, and for all multiindices* α

$$\partial^\alpha \varphi_j \to \partial^\alpha \varphi \qquad \text{in } C_b(U).$$

A distribution T *on* U *is a continuous linear map from* $C_c^\infty(U) \to \mathbb{K}$*, where by continuity we mean*

$$T\varphi_j \to T\varphi, \quad \text{if } \varphi_j \to \varphi \text{ in } C_c^\infty(U).$$

We denote the space of distributions by $\mathcal{D}'(U)$*, (i.e. the dual space of* $\mathcal{D}(U)$*), which is a* $\mathbb{K}$*-vector space.*
We call $T_n \to T$ *in* $\mathcal{D}'(U)$ *if*

$$T_n\varphi \to T\varphi, \qquad \forall \varphi \in \mathcal{D}(U).$$

Remark 4.8. *The test function space* $\mathcal{D}(U)$ *is not a Banach space. Nevertheless, one can introduce a countable family of seminorms on* $\mathcal{D}(U)$ *such that it becomes a Fréchet space. For example, for* $U = \mathbb{R}$*, we define a sequence of seminorm (i.e. a function* $a \mapsto [a]$ *from a* $\mathbb{K}$*-vector space to* $\mathbb{R}$ *satisfying* $[a] \geqslant 0$*,* $[c \cdot a] = |c|[a]$ *and* $[a+b] \leqslant [a] + [b]$*)*

$$[\varphi]_{k,n} = \|\varphi^{(k)}\|_{C_b([-n,n])}, \quad k, n \in \mathbb{N} \cup \{0\},$$

and introduce a metric

$$d^1(\varphi, \psi) = \sum_{k,n} 2^{-(k+n)} \frac{[\varphi - \psi]_{k,n}}{1 + [\varphi - \psi]_{k,n}},$$

such that $(\mathcal{D}(\mathbb{R}), d^1)$ *becomes a metric space (but not complete). The convergence* $\varphi_j \to \varphi$ *in* $\mathcal{D}(\mathbb{R})$ *(defined above) is not induced by this metric, and hence* $T \in \mathcal{D}'(U)$ *is not continuous on* $(\mathcal{D}(U), d^1)$*.*
Fréchet spaces can be viewed as generalization of Banach spaces, but with more complicated topological structure. Nevertheless, Banach-Steinhaus theorem and the open mapping theorem still hold on Fréchet spaces.

When restricted on a compact set, the definition of a distribution becomes more clear. For any fixed compact set $K \subset U$, we consider the subspace of $\mathcal{D}(U)$:

$$X_K = \{\varphi \in \mathcal{D}(U) : \operatorname{supp}\varphi \subset K\}.$$

We define a metric on X_K

$$d_K(\varphi, \psi) = \sup{}_{k \geqslant 0} 2^{-k} \min\{1, \|\varphi - \psi\|_{C_b^k(U)}\}.$$

With this metric X_K is a complete metric space:

$$d_K(\varphi_j, \varphi) \to 0 \quad \text{iff} \quad \varphi_j \to \varphi \text{ in } C_b^k(U),\ \forall k \geqslant 0.$$

Hence if $T \in \mathcal{D}'(U)$ and $d_K(\varphi_j, \varphi) \to 0$, then $T\varphi_j \to T\varphi$, and T determines a continuous linear map on X_K.

[13.01.2023]

[16.01.2023]

Furthermore we have the following description.

Lemma 4.9. *Let $T \in \mathcal{D}'(U)$. For every compact set $K \subset U$ there exists k and $C > 0$ so that*

$$|T\varphi| \leqslant C\|\varphi\|_{C_b^k(U)}, \quad \forall \varphi \in X_K.$$

Proof. Since $T \in \mathcal{D}'(U)$ defines a continuous linear map from $(X_K, d) \to \mathbb{K}$,

$$|T\varphi| < \infty, \quad \forall \varphi \in X_K.$$

Now we argue as for the uniform boundedness principle of Banach-Steinhaus: There exists m so that the set

$$C_m := \{\varphi \in X_K : |T\varphi| \leqslant m\}$$

contains an open ball $B_r(\varphi_0)$. Hence the set C_{2m} contains an open ball $B_r(0)$, since

$$\forall \varphi \in B_r(0), \quad |T\varphi| \leqslant |T(\varphi - \varphi_0)| + |T\varphi_0| \leqslant m + m = 2m.$$

Let k_0 be such that $2^{-k_0} < r$. If $\varphi \in \mathcal{D}(U)$ such that $\|\varphi\|_{C_b^{k_0}(U)} \leqslant 2^{-k_0}$ and $\operatorname{supp} \varphi \subset K$, then

$$d(\varphi, 0) \leqslant \max\{\sup{}_{k \leqslant k_0} 2^{-k}\|\varphi\|_{C_b^{k_0}(U)}, \sup{}_{k > k_0} 2^{-k}\} < r,$$

and hence

$$|T\varphi| \leqslant 2m, \quad \forall \varphi \in X_K \text{ with } \|\varphi\|_{C_b^{k_0}(U)} \leqslant 2^{-k_0}.$$

The result follows with $C = m2^{1+k_0}$. □

Lemma 4.10. *Suppose that $T_n \to T$ in $\mathcal{D}'(U)$ and that $K \subset U$ is compact. Then there exists k and C so that*

$$\sup{}_n |T_n(\varphi)| \leqslant C\|\varphi\|_{C_b^k(U)}, \quad \forall \varphi \in X_K,$$

and

$$\sup\{|T_n(f) - T(f)| : f \in X_K, \|f\|_{C_b^k} \leqslant 1\} \to 0.$$

Proof. **Exercise.** □

$\mathcal{D}'(U)$ is a rather big space, which includes almost all the interesting "functions". For example, $C(U)$, $C^k(U)$, $L^p(U)$, $p \in [1, \infty]$ are all subspace of $L^1_{\text{loc}}(U)$, which is subspace of $\mathcal{D}'(U)$:

Example 4.11 ($L^1_{\text{loc}}(U) \subset \mathcal{D}'(U)$)**.** *Let*

$$L^1_{\text{loc}}(U) = \{f : U \to \mathbb{K} \textit{ measurable} \,|\, f \in L^1(K) \textit{ for all } K \subset U \textit{compact}\},$$

which is a $\mathbb{K}$-vector space. We call $f_j \to f$ in $L^1_{\text{loc}}(U)$ if

$$f_j \chi_K \to f \chi_K \textit{ in } L^1(K), \quad \forall \textit{ compact set } K \subset U.$$

(Exercise:) *Any $f \in L^1_{\text{loc}}(U)$ determines a unique distribution $T_f \in \mathcal{D}'(U)$ given by*

$$T_f(\varphi) = \int_U f \varphi \, dm^d, \quad \forall \varphi \in \mathcal{D}(U).$$

In the following we will identify an $L^1_{\text{loc}}(U)$ function f with the distribution $T_f \in \mathcal{D}'(U)$.
Similarly, any Radon measure μ on U defines a unique distribution $T^\mu \in \mathcal{D}'(U)$:

$$T^\mu(\varphi) = \int_U \varphi d\mu, \quad \forall \varphi \in \mathcal{D}(U).$$

4.2.2 Products, derivatives and supports

One can easily define the product between two smooth functions: If $\varphi \in \mathcal{D}(u)$ and $\phi \in C^\infty(U)$ then $\phi\varphi \in \mathcal{D}(U)$, and

$$\varphi_j \to \varphi \text{ in } \mathcal{D}(U) \overset{\phi \in C^\infty(U)}{\Longrightarrow} \phi\varphi_j \to \phi\varphi \text{ in } \mathcal{D}(U).$$

Hence for any $T \in \mathcal{D}'(U)$ and $\phi \in C^\infty(U)$, $T(\phi \,\cdot) : \mathcal{D}(U) \to \mathbb{K}$ determines a continuous linear map, which we define as the product between T and ϕ.
Similarly, if $\varphi \in \mathcal{D}(U)$, then $\partial_{x_k}\varphi \in \mathcal{D}(U)$, and $\varphi_j \to \varphi$ in $\mathcal{D}(U)$ implies $\partial_{x_k}\varphi_j \to \partial_{x_k}\varphi$ in $\mathcal{D}(U)$. Hence for any $T \in \mathcal{D}'(U)$, the map $T(-\partial_{x_k} \cdot) : \mathcal{D}(U) \to \mathbb{K}$ determines a distribution in $\mathcal{D}'(U)$, which we define as $\partial_{x_k} T$.
For $\varphi \in \mathcal{D}(U)$, its zero points are those where it vanishes, and we define its support as $\text{supp}(\varphi) = \overline{\{x \in U \,|\, \varphi(x) \neq 0\}}$. For $T \in \mathcal{D}'(U)$, we say that T vanishes near $x \in U$ if there exists $r > 0$ so that $T(\varphi) = 0$ for all $\varphi \in \mathcal{D}(U)$ with support in $B_r(x)$.

Definition 4.12 (Products, derivatives and supports)**.** *We define the product $\phi T \in \mathcal{D}'(U)$ of a distribution $T \in \mathcal{D}'(U)$ and a smooth function $\phi \in C^\infty(U)$ as*

$$(\phi T)(\varphi) = T(\phi\varphi), \quad \forall \varphi \in \mathcal{D}(U).$$

We define the derivative $\partial_{x_j} T \in \mathcal{D}'(U)$ of a distribution $T \in \mathcal{D}'(U)$ as

$$(\partial_{x_j} T)(\varphi) = T(-\partial_{x_j}\varphi), \quad \forall \varphi \in \mathcal{D}(U).$$

We define the support of $T \in \mathcal{D}'(U)$ as the complement of the points near which T vanishes. We denote by $\mathcal{E}'(U)$ the set of distributions with compact supports.

We can easily check

- Associative law: $\phi(\psi T) = \psi(\phi T) = (\phi\psi)T$, $\forall T \in \mathcal{D}'(U)$, $\phi, \psi \in C^\infty(U)$;
- Distributive law: $\phi(T + S) = \phi T + \phi S$, $\forall T, S \in \mathcal{D}'(U)$, $\forall \phi \in C^\infty(U)$;
- Schwarz theorem: $\partial_{x_j}\partial_{x_k} T = \partial_{x_k}\partial_{x_j} T$, $\forall T \in \mathcal{D}'(U)$;
- Leibniz' product rule: $\partial_{x_j}(\phi T) = (\partial_{x_j}\phi)T + \phi(\partial_{x_j} T)$, $\forall \phi \in C^\infty(U)$, $T \in \mathcal{D}'(U)$:

$$\begin{aligned}\partial_{x_j}(\phi T)(\varphi) = -T(\phi \partial_{x_j}\varphi) &= T((\partial_{x_j}\phi)\varphi) - T(\partial_{x_j}(\phi\varphi)) \\ &= [(\partial_{x_j}\phi)T](\varphi) + (\phi\partial_{x_j} T)(\varphi).\end{aligned}$$

- Coincidence with the classical definitions: If $T = T_\phi \in \mathcal{D}'(U)$, with $\phi \in \mathcal{D}(U) \subset L^1_{\text{loc}}(U)$, then

$$\begin{aligned}&\psi T_\phi = T_{\phi\psi} \text{ for } \psi \in C^\infty(U), \\ &T_{\partial_{x_j}\phi} = \partial_{x_j} T_\phi.\end{aligned}$$

We first have the following approximation result for the constant function 1 on compact sets.

Lemma 4.13 (Local approximation of 1)**.** *Let $U \subset \mathbb{R}^d$ be an open set, and K_j be a monotone sequence of compact sets so that K_j is contained in the interior of K_{j+1} and $U = \bigcup K_j$. Then there exists a sequence of test functions $(\phi_j) \subset \mathcal{D}(U)$ so that*

$$\operatorname{supp}\phi_j \subset K_{j+1}, \quad \phi_j = 1 \text{ on } K_{j-1}, \quad \text{and } \partial^\alpha\phi_j \to \partial^\alpha 1 \text{ on compact sets.}$$

In particular, for any $\varphi \in \mathcal{D}(U)$, $\phi_j\varphi \to \varphi$ in $\mathcal{D}(U)$, and hence for any $T \in \mathcal{D}'(U)$, $\phi_j T \to T$ in $\mathcal{D}'(U)$, where $\operatorname{supp}(\phi_j T) \subset \operatorname{supp}(\phi_j)$*.*

Proof. For any j, let $r_j > 0$ so that

$$\min\{\operatorname{dist}(K_{j-1}, \mathbb{R}^d \backslash K_j), \operatorname{dist}(K_j, \mathbb{R}^d \backslash K_{j+1})\} \geqslant r_j > 0.$$

Let $\eta \in C_c^\infty(B_1(0))$ be radial with $\int_{\mathbb{R}^d} \varphi dx = 1$, and $\eta_r(x) = r^{-d}\eta(x/r)$, as in Corollary 3.39. Then

$$\phi_j := \eta_{r_j} * \chi_{K_j} \in \mathcal{D}(U),$$

is the searched sequence: For any compact set $K \subset U$, there exists j_0 such that $K \subset K_{j_0-1}$ and $\phi_j = 1$ on K for all $j \geqslant j_0$.

□

Unlike the product of two continuous functions, the product of two distributions are not always well-defined. But one can always define the derivatives of a distribution.

Example 4.14 ($\partial_x H = \delta_0$)**.** *Let $U = \mathbb{R}$.*
The Heaviside function $H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geqslant 0 \end{cases}$ belongs to $L^\infty(U) \subset L^1_{\text{loc}}(U) \subset \mathcal{D}'(U)$. We calculate its derivative: For $\varphi \in \mathcal{D}(U)$,

$$(\partial_x H)(\varphi) = (\partial_x T_H)(\varphi) = -T_H(\varphi') = -\int_0^\infty H(x)\varphi'(x)dx = \varphi(0) = \delta_0(\varphi),$$

and hence

$$\partial_x H = \delta_0,$$

where $\delta_0 \in \mathcal{D}'(\mathbb{R})$ is the Dirac measure, which belongs to $\mathcal{E}'(\mathbb{R})$ with compact support $\operatorname{supp}(\delta_0) = \{0\}$.
One can further calculate $\partial_x \delta_0 \in \mathcal{D}'(\mathbb{R})$ as

$$(\partial_x \delta_0)(\varphi) = -\delta_0(\varphi') = -\varphi'(0), \quad \forall \varphi \in \mathcal{D}(\mathbb{R}),$$

and iteratively $(\partial_x^n \delta_0)(\varphi) = (-1)^n \varphi^{(n)}(0)$.

[16.01.2023]

[20.01.2023]

Lemma 4.15 (Disjoint supports)**.** *Let $\varphi \in \mathcal{D}(U)$ and $T \in \mathcal{D}'(U)$ with disjoint supports. Then*

$$T\varphi = 0.$$

Proof. Let K be the support of φ. Given $x \in K$ there exists r so that $T\psi = 0$ for every $\psi \in \mathcal{D}(U)$ supported in $B_r(x)$. Since K is compact there is a finite covering of such balls $B_{r_j}(x_j)$ with $1 \leqslant j \leqslant N$. We choose a partition of unity $\eta_j \in C^\infty(U)$ supported in $B_{r_j}(x_j)$ so that (recalling the proof of Theorem 3.41)

$$\sum_{j=1}^{N} \eta_j(x) = 1, \quad \forall x \in K,$$

Then

$$T\varphi = \sum_{j=1}^{N} T(\eta_j \varphi) = 0.$$

□

4.2.3 Convolution in $\mathbb{R}^d$ and density of $\mathcal{D}(U)$ in $\mathcal{D}'(U)$

Definition 4.16 (Convolution between $\mathcal{D}(\mathbb{R}^d)$ and $\mathcal{D}'(\mathbb{R}^d)$)**.** *Let $T \in \mathcal{D}'(\mathbb{R}^d)$ and $\phi \in \mathcal{D}(\mathbb{R}^d)$. We define their convolution by*

$$(\phi * T)(x) = T(\phi(x - \cdot)), \quad \forall x \in \mathbb{R}^d.$$

The righthand side denotes T acting on $\phi(x-y)$ as a function of y.

Lemma 4.17. *With the notation above, we know $\phi * T \in C^\infty(\mathbb{R}^d)$ and*

$$\partial_{x_j}(\phi * T) = (\partial_{x_j}\phi) * T = \phi * (\partial_{x_j} T).$$

If $\psi \in L^1(\mathbb{R}^d)$ then

$$\phi * T_\psi(x) = \phi * \psi(x).$$

Moreover, if $\operatorname{supp}\phi = K_1$ *and* $\operatorname{supp} T = K_2$ *then*

$$\operatorname{supp}(\phi * T) \subset K_1 + K_2 = \{x + y : x \in K_1 \text{ and } y \in K_2\}.$$

In particular, if $\phi \in \mathcal{D}(\mathbb{R}^d)$ and $T \in \mathcal{E}'(\mathbb{R}^d)$, then $\phi * T \in \mathcal{D}(\mathbb{R}^d)$.

Proof. **Exercise.**

□

Remark 4.18. *We can equivalently define the convolution of $\phi \in \mathcal{D}(\mathbb{R}^d)$ and $T \in \mathcal{D}'(\mathbb{R}^d)$ as the distribution:*

$$(\phi * T)(\varphi) = T(\tilde{\phi} * \varphi), \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d),$$

where $\tilde{\phi}(x) = \phi(-x)$*, since by the linearity of* T*,*

$$(\phi * T)(\varphi) = \int_{\mathbb{R}^d} T(\phi(x - \cdot))\varphi(x)dm^d = T\Big(\int_{\mathbb{R}^d} \tilde{\phi}(\cdot - x)\varphi(x)dm^d\Big) = T(\tilde{\phi} * \varphi).$$

In the same way we can easily define the convolution $\phi * T \in \mathcal{D}'(\mathbb{R}^d)$ *when* $\phi \in C^k(\mathbb{R}^d)$ *has compact support. By an abuse of notation we write the convolution evaluated at* x *whenever it is defined, even if it is not defined on all of* $\mathbb{R}^d$*.*

Definition 4.19 (Convolution between $\mathcal{D}'(\mathbb{R}^d)$ and $\mathcal{E}'(\mathbb{R}^d)$)**.** *Let* $T \in \mathcal{D}'(\mathbb{R}^d)$ *and let* $S \in \mathcal{E}'(\mathbb{R}^d)$ *which consists of distributions with compact support. We define their convolution by*

$$(S * T)(\phi) = T(\tilde{S} * \phi)$$

for $\phi \in \mathcal{D}(\mathbb{R}^d)$*. Here* $\tilde{S}(\psi) = S(\tilde{\psi})$*,* $\forall \psi \in \mathcal{D}(\mathbb{R}^d)$*.*

Example 4.20 (Convolution with δ_0)**.** *Recall the Dirac sequence* $(\eta_r)_{r>0}$ *in Corollary 3.39, such that for any* $f \in L^p$*,* $1 \leqslant p < \infty$*, we have* $\eta_r * f \to f$ *in* L^p*.*
For any $\varphi \in \mathcal{D}(U)$*,* $\eta_r * \varphi \to \varphi$ *in* $\mathcal{D}(U)$*. Similarly for any* $f \in C_c^k(U)$*, there exist compact set* $K \subset U$ *and* $\varepsilon > 0$ *such that* $\operatorname{supp}(\eta_r * f)$ *is contained in* K *for all* $r \leqslant \varepsilon$*, and* $\eta_r * f \to f$ *in* $C_b^k(U)$*.*
Show that $\delta_0 \in \mathcal{E}'(\mathbb{R}^d)$ *and* $\delta_0 * T = T$ *for all* $T \in \mathcal{D}'(\mathbb{R}^d)$*. In particular* $\eta_r \to \delta_0$*, and* $\eta_r * T \to T$ *in the sense of distribution.* ***(Exercise).***

Recall Lemma 4.13 that any distribution T can be approximated by a sequence of distributions with compact supports $\phi_j T$. Lemma 4.17 and the above example tells us that any distribution with compact support $S \in \mathcal{E}'(U)$ can be approximated by $(\eta_r * S)_r \subset \mathcal{D}(U)$ in the distribution sense. Hence we have proved

Theorem 4.21. *Let* $U \subset \mathbb{R}^d$ *be open. Then* $\mathcal{D}(U) \subset \mathcal{D}'(U)$ *is dense.*

By approximation argument we can also prove

Corollary 4.22. *Suppose that* $U \subset \mathbb{R}^d$ *is an open connected set and* $\partial_{x_j} T = 0$*,* $j = 1, \cdots, d$*. Then there eixsts a constant* c *so that* $T = T_c$*.*

Proof. **Exercise.** □

4.3 Schwartz functions and tempered distributions

We briefly cover the definitions of Schwartz functions and tempered distributions, which are the proper functional framework for the Fourier transform on the whole space $\mathbb{R}^d$.

4.3.1 Definitions

Definition 4.23. *The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ consists of all Schwartz functions, which are smooth functions $f \in C^\infty(\mathbb{R}^d)$ satisfying for any $k \in \mathbb{N}$*

$$\|f\|_{k,\mathcal{S}} := \sup{}_{x\in\mathbb{R}^d,\, |\alpha|\leqslant k}(1+|x|^k)|\partial^\alpha f(x)| < \infty,$$

which is equivalent to say

$$\mathcal{S}(\mathbb{R}^d) = \{f \in C^\infty(\mathbb{R}^d) \,|\, \sup{}_{x\in\mathbb{R}^d}|x^\alpha \partial^\beta f(x)| < \infty, \forall \text{ multiindices } \alpha, \beta\}.$$

We define $d : \mathcal{S} \times \mathcal{S} \to [0, \infty)$ by

$$d(f, g) = \sup{}_k 2^{-k} \min\{1, \sup{}_{|\alpha|+|\beta|=k}\|x^\alpha \partial^\beta (f-g)\|_{sup}\},$$

such that $(\mathcal{S}, d)$ is a complete metric space.
A tempered distribution is a continuous linear map from $(\mathcal{S}(\mathbb{R}^d), d)$ to $\mathbb{K}$. We denote the space of tempered distributions by $\mathcal{S}'(\mathbb{R}^d)$. We say that $T_j \to T$ in $\mathcal{S}'(\mathbb{R}^d)$ if $T_j(\phi) \to T(\phi)$ for all $\phi \in \mathcal{S}(\mathbb{R}^d)$.

It is straightforward to check that

- For $f \in \mathcal{S}$, $\|f\|_{k,\mathcal{S}}$ may depend on $k \in \mathbb{N}$, and for any N there exists C_N such that $\|f\|_{C_b^N} \leqslant C_N$ and $|f(x)| \leqslant C_N(1+|x|)^{-N}$;
- $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$;
- $\mathcal{S}(\mathbb{R}^d) \subset C^k(\mathbb{R}^d)$ for all $k \in \mathbb{N}$ and $\mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ for all $p \in [1, \infty]$;
- If $f \in \mathcal{S}(\mathbb{R}^d)$, then $x^\alpha f$, $\partial^\alpha f \in \mathcal{S}(\mathbb{R}^d)$ for any multiindex α;
- If $g \in C^\infty$ and for any multiindex α there exist $c_{|\alpha|}$ and $\kappa_{|\alpha|}$ so that

  $$|\partial^\alpha g| \leqslant c_{|\alpha|}(1+|x|)^{\kappa_{|\alpha|}},$$

  then $f \in \mathcal{S}(\mathbb{R}^d)$ implies $fg \in \mathcal{S}(\mathbb{R}^d)$;
- If $f, g \in \mathcal{S}(\mathbb{R}^d)$, then $f * g \in \mathcal{S}(\mathbb{R}^d)$;
- The Gaussian function $e^{-\frac{1}{2}|x|^2} \in \mathcal{S}(\mathbb{R}^d)$;
- $\mathcal{E}'(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$, since for any $S \in \mathcal{E}'(\mathbb{R}^d)$, $S(f) = S(\eta f)$, $\forall f \in C^\infty(\mathbb{R}^d)$ and $\eta \in \mathcal{D}(\mathbb{R}^d)$ with $\eta = 1$ on $\mathrm{supp}(S)$.

Similar as Lemma 4.9 and Lemma 4.10, we have **(Exercise)**

Lemma 4.24. *Let $T \in \mathcal{S}'(\mathbb{R}^d)$. Then there exist k and c so that*

$$|T\phi| \leqslant c\|\phi\|_{k,\mathcal{S}}, \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d).$$

If $T_n \to T$ in $\mathcal{S}'(\mathbb{R}^d)$ then there exist C and l so that

$$\sup{}_n |T_n\phi| \leqslant C\|\phi\|_{l,\mathcal{S}},$$

and

$$\sup\{|T_n(\phi) - T\phi| \,|\, \phi \in \mathcal{S}(\mathbb{R}^d),\ \|\phi\|_{l,\mathcal{S}} \leqslant 1\} \to 0.$$

[20.01.2023]

[23.01.2023]

For $T \in \mathcal{S}'(\mathbb{R}^d)$, we define the multiplication by a smooth function with controlled derivatives gT and the derivative $\partial_{x_j}T$ as we did it for distributions in Definition 4.12.
Similarly, since compactly supported distributions $S \in \mathcal{E}'(\mathbb{R}^d)$ can act on Schwartz functions $f \in \mathcal{S}(\mathbb{R}^d)$, we can define the convolution $(S * f)(x) = S(f(x - \cdot)) \in \mathcal{S}(\mathbb{R}^d)$. We then can define, as in Definition 4.16 and Definition 4.19, the convolution of a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ with Schwartz functions $f \in \mathcal{S}(\mathbb{R}^d)$: $(f * T)(x) = T(f(x - \cdot)) \in C^\infty(\mathbb{R}^d)$, and with compactly supported distributions $S \in \mathcal{E}'(\mathbb{R}^d)$: $(S * T)(\phi) = T(\tilde{S} * \phi)$, $\forall \phi \in \mathcal{S}(\mathbb{R}^d)$.
By approximation argument, one has **(Exercise)**

Theorem 4.25. *For $1 \leqslant p \leqslant \infty$, there are the embeddings*

$$\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d).$$

The embeddings are dense if $p < \infty$.

4.3.2 Fourier transformation

Recall the definition of the Fourier transform $\mathcal{F}(f)$ of a Lebesgue integrable function $f \in L^1(\mathbb{R}^d;\mathbb{C})$:

$$\hat{f}(\xi) := \mathcal{F}(f)(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\,\mathrm{d}x\,, \quad \forall \xi \in \mathbb{R}^d\,.$$

By the density of the embedding $\mathcal{D}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$, one can easily show Riemann-Lebesgue theorem:

$$\mathcal{F} : L^1(\mathbb{R}^d) \to C_0(\mathbb{R}^d) \subset C_b(\mathbb{R}^d).$$

The Fourier transform $\mathcal{F} : L^1(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ is a linear continuous map and

$$\|\mathcal{F}(f)\|_{C_b(\mathbb{R}^d)} \leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}\|f\|_{L^1(\mathbb{R}^d)}, \text{ i.e. } \|\mathcal{F}\|_{L^1(\mathbb{R}^d)\mapsto C_b(\mathbb{R}^d)} \leqslant \frac{1}{(2\pi)^{\frac{d}{2}}}.$$

Since $\mathcal{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$, we can define the Fourier transform on $\mathcal{S}(\mathbb{R}^d)$:

Theorem 4.26. *The Fourier transform maps continuously from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$: For any integer $k \in \mathbb{N}$, there exist a constant C and an integer $N \in \mathbb{N}$ such that*

$$\|\mathcal{F}(f)\|_{k,\mathcal{S}} \leqslant C\|f\|_{N,\mathcal{S}}, \quad \forall f \in \mathcal{S}(\mathbb{R}^d).$$

Furthermore, for any $f \in \mathcal{S}(\mathbb{R}^d)$,

$$\mathcal{F}((\frac{1}{i}\partial_x)^\alpha f) = \xi^\alpha \mathcal{F}(f), \quad \mathcal{F}(x^\alpha f) = (i\partial_\xi)^\alpha \mathcal{F}(f), \quad \forall \textit{multiindex } \alpha. \tag{4.1}$$

Proof. Exercise. □

Roughly speaking, the Schwartz functions have two properties: they have infinite bounded derivatives and they decay fast at infinity. In view of (4.1), the Fourier transform works well in the framework of Schwartz space, and hence of tempered distributions by duality: We define the Fourier transform of a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ as a tempered distribution $\mathcal{F}(T) \in \mathcal{S}'(\mathbb{R}^d)$:

$$\mathcal{F}(T)(f) = T(\mathcal{F}(f)), \quad \forall f \in \mathcal{S}(\mathbb{R}^d).$$

We have

Corollary 4.27. *The Fourier transform $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is a continuous linear map, and (4.1) holds on $\mathcal{S}'(\mathbb{R}^d)$.*

Example 4.28. *We calculate the Fourier transform $\hat{\delta}_0$ of the delta function $\delta_0 \in \mathcal{E}'(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$:*

$$\hat{\delta}_0(f) = \delta_0(\hat{f}) = \hat{f}(0) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(\xi)\,\mathrm{d}\xi = T_{(2\pi)^{-\frac{d}{2}}}(f), \quad \forall f \in \mathcal{S}(\mathbb{R}^d),$$

that is, the Fourier transform of the delta function is a constant function:

$$\mathcal{F}\delta_0 = (2\pi)^{-\frac{d}{2}}.$$

By (4.1), we have

$$\widehat{\partial^\alpha \delta_0} = (2\pi)^{-\frac{d}{2}}(i\xi)^\alpha.$$

Since (see Example 4.11) any function $f \in L^p(\mathbb{R}^d)$, $p \in [1, \infty]$ can be identified with a tempered distribution $T_f \in \mathcal{S}'(\mathbb{R}^d)$ which is defined as

$$T_f(\phi) = \int_{\mathbb{R}^d} f \phi \, \mathrm{d}x \, , \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d),$$

we can define the Fourier transform of f as the following tempered distribution

$$\mathcal{F}(f)(\phi) = T_f(\mathcal{F}(\phi)), \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d).$$

We have seen $\mathcal{F} : L^1(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d)$. We have further **(Exercise)**

Theorem 4.29. *The Fourier transform defines a unitary operator from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$, i.e. $\mathcal{F}$ is continuous linear map satisfying $\mathcal{F}^*\mathcal{F} = \mathcal{F}\mathcal{F}^* = \mathrm{Id}$. In particular, the Fourier transform of $f \in L^2(\mathbb{R}^d)$ is a function in $L^2(\mathbb{R}^d)$ which reads as the following limit (almost everywhere)*

$$\mathcal{F}(f)(\xi) = \lim_{R\to\infty} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{|x|<R} e^{-ix\cdot\xi} f(x) \, \mathrm{d}x \, . \tag{4.2}$$

By Riesz-Thorin Interpolation theorem (see Exercise sheet),

$$\mathcal{F} : L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d) \text{ is continuous and linear,} \quad \forall p \in [1,2],$$
$$\text{such that } \|\mathcal{F}(f)\|_{L^{p'}(\mathbb{R}^d)} \leqslant (2\pi)^{-\frac{d}{2}(\frac{2}{p}-1)} \|f\|_{L^p(\mathbb{R}^d)}.$$

In general the Fourier transform of a $L^p(\mathbb{R}^d)$, $p > 2$-function does not necessarily belong to Lebesgue spaces: We keep in mind that the Fourier transform of the constant function $f(x) = (2\pi)^{-\frac{d}{2}}$ is the Dirac-function $\delta_0 \in \mathcal{E}'(\mathbb{R}^d)$. We remark that one can define the inverse Fourier transform as $\mathcal{F}^{-1}(T) = \mathcal{F}(\widetilde{T})$ on $\mathcal{S}'(\mathbb{R}^d)$, such that $\mathcal{F}^{-1}\mathcal{F} = \mathcal{F}\mathcal{F}^{-1} = \mathrm{Id}$ on $\mathcal{S}'(\mathbb{R}^d)$.

[23.01.2023]

[27.01.2023]

4.4 Sobolev spaces $W^{k,p}$

4.4.1 Definitions

Definition 4.30. *Let $U \subset \mathbb{R}^d$ be open, $k \in \mathbb{N}$.*

- *For $1 \leqslant p \leqslant \infty$, the Sobolev space $W^{k,p}(U) \subset L^p(U)$ is the set of all $L^p(U)$ functions, so that for all multiindices α of length at most k there exists $f_\alpha \in L^p(U)$ so that*

$$\partial^\alpha T_f = T_{f_\alpha}.$$

By an abuse of notation we identify $\partial^\alpha T_f$ *with* f_α*, and hence* $g = \partial_{x_j} f$ *in* $L^p(U)$ *for* $f \in L^p(U)$ *means*

$$\int_U g\phi dm^d = -\int_U f\partial_{x_j}\phi dm^d, \quad \forall \phi \in \mathcal{D}(U). \tag{4.3}$$

For $p = 2$*, we denote* $H^k(U) = W^{k,2}(U)$*.*

- *We define for* $p \in [1, \infty)$

$$\|f\|_{W^{k,p}(U)} = \left(\sum_{|\alpha| \leqslant k} \|\partial^\alpha f\|^p_{L^p(U)} \right)^{\frac{1}{p}},$$

and

$$\|f\|_{W^{k,\infty}(U)} = \sup{}_{|\alpha|\leqslant k} \|\partial^\alpha f\|_{L^\infty(U)}.$$

Then $(W^{k,p}(U), \|\|_{W^{k,p}(U)})$ *is normed space.*

For $p = 2$*, we define*

$$\langle f, g\rangle_{H^k(U)} = \sum_{|\alpha|\leqslant k} \langle \partial^\alpha f, \partial^\alpha g\rangle_{L^2(U)}.$$

Then $(H^k(U), \langle , \rangle_{H^k(U)})$ *is pre-Hilbert space.*

- *For* $1 \leqslant p < \infty$*, we define* $W_0^{k,p}(U)$ *as the closure of* $C_c^\infty(U)$ *with respect to the norm* $\|\cdot\|_{W^{k,p}(U)}$*.*

- *For* $1 < p < \infty$*, we define* $W^{-k,p}(U) = (W_0^{k,p'}(U))^*$*. The norm is defined by*

$$\|T\|_{W^{-k,p}(U)} = \sup{}_{f \in W_0^{k,p'}(U),\, \|f\|_{W^{k,p'}(U)} \leqslant 1} |T(f)|.$$

Remark 4.31.

- *If* $V \subset U$ *open, then the restriction map*

$$\chi_V : W^{k,p}(U) \ni f \mapsto f|_V \in W^{k,p}(V)$$

is a continuous linear map with the norm less than 1*: We have* $\partial^\alpha(f|_V) = (\partial^\alpha f)|_V$ *and* $\|f|_V\|_{W^{k,p}(V)} \leqslant \|f\|_{W^{k,p}(U)}$ *for all* $f \in W^{k,p}(U)$*.*

The (trivial) extension by 0

$$W_0^{k,p}(V) \ni f \mapsto \tilde{f} = \begin{cases} f & \text{inside } V \\ 0 & \text{outside } V \end{cases} \in W_0^{k,p}(U)$$

defines a canonical (nonsurjective) isometry from $W_0^{k,p}(V)$ *to* $W_0^{k,p}(U)$*:*

$$\|f\|_{W^{k,p}(V)} = \|\tilde{f}\|_{W^{k,p}(U)}.$$

Hence the restriction map $\chi_V : W^{k,p}(U) \to W^{k,p}(V)$ *has indeed norm 1.*

- *We denote* $L^p = W^{0,p}$*. Obviously the embedding* $W^{k_1,p}(U) \subset W^{k_2,p}(U)$ *holds if* $k_2 \leqslant k_1$ *and* $p \in (1, \infty)$*.*
- *One can check* ***(Exercise)***
  - *If* $g \in C_b^k(U)$ *and* $f \in W^{k,p}(U)$*, then* $gf \in W^{k,p}(U)$ *and the Leibniz' rule holds;*
  - *For* $f \in W^{k,p}(U)$ *and* $\phi : V \to U$ *a* C_b^k *diffeomorphism, the chain rule holds and*

$$\|f \circ \phi\|_{W^{k,p}(V)} \leqslant C\|f\|_{W^{k,p}(U)}. \tag{4.4}$$

Theorem 4.32. *Let* $U \subset \mathbb{R}^d$ *be open,* $k \in \mathbb{N}$ *and* $1 \leqslant p \leqslant \infty$*.*
The Sobolev space $(W^{k,p}(U), \|\|_{W^{k,p}(\Omega)})$ *is a Banach space. In particular for* $p = 2$*,* $(H^k(\Omega), \langle,\rangle_{H^k(\Omega)})$ *is a Hilbert space.*
If $U = \mathbb{R}^d$ *and* $1 \leqslant p < \infty$*, then* $W_0^{k,p}(\mathbb{R}^d) = W^{k,p}(\mathbb{R}^d) = \overline{C_c^\infty(\mathbb{R}^d)}^{\|\|_{W^{k,p}(\mathbb{R}^d)}}$*.*

Proof. Let $\Sigma_k = \{\alpha : |\alpha| \leqslant k\}$. There is an obvious isometry (nonsurjective)

$$W^{k,p}(U) \ni f \to (\partial^\alpha f)_{|\alpha| \leqslant k} \in L^p(U \times \Sigma_k),$$

where

$$\|f\|_{W^{k,p}(U)} = \|(\partial^\alpha f)\|_{L^p(U \times \Sigma_k)}.$$

Let f_j be a Cauchy sequence in $W^{k,p}(U)$ with limit $f \in L^p(U)$. Then $\partial^\alpha f_j \to f_\alpha$ in $L^p(U)$ and hence in $\mathcal{D}'(U)$. It is easy to check that $f_\alpha = \partial^\alpha f$ in $\mathcal{D}'(U)$: For any $\phi \in \mathcal{D}(U)$,

$$T_{f_\alpha}(\phi) = \lim_{j \to \infty} (-1)^{|\alpha|} \int_U f_j \partial^\alpha \phi \, dm^d = (-1)^{|\alpha|} \int_U f \partial^\alpha \phi \, dm^d = T_{\partial^\alpha f}(\phi).$$

Thus $f \in W^{k,p}(U)$ and $W^{k,p}(U)$ is complete and we can identify $W^{k,p}(U)$ with a closed subspace of $L^p(U \times \Sigma_k)$.

[27.01.2023]

[30.01.2023]

Now let $1 \leqslant p < \infty$. Density of $\mathcal{D}(\mathbb{R}^d) \subset W^{k,p}(\mathbb{R}^d)$ follows by the same argument (first smooth cutoff and then convolution) as in the proof of Theorem 4.21. □

Remark 4.33. • *If $U \subset \mathbb{R}^d$ is bounded, we can also follow the argument as in the proof of Theorem 4.21 to show that any function $f \in W^{k,p}(U)$, $k \in \mathbb{N}$, $p \in [1, \infty)$ can be approximated by a sequence of test functions $\eta_{r_j} * (\phi_j f)$ in $W^{k,p}_{\mathrm{loc}}(U)$, i.e. in $W^{k,p}(K)$, for all compact sets $K \subset U$. The approximation holds only locally, otherwise we have shown $W^{k,p}(U) = W^{k,p}_0(U)$, which is in general not true for bounded open set U (keeping in mind of the constant function $1 \in W^{1,p}(U)$ but not in $W^{1,p}_0(U)$, $p > d$, where $W^{1,p}$-norm is finer than C_b-norm).*

- *In order to show the global approximation by smooth functions, we need some regularity assumption on the boundary of U. For example, if U is an open bounded Lipschitz domain, then $f \in W^{k,p}(U)$, $k \in \mathbb{N}$, $p \in [1, \infty)$ can be approximated by a sequence of smooth functions $f_n \in C^\infty(\overline{U})$ (globally) in $W^{k,p}(U)$, see Corollary 4.36 below.*

- *We may interpret $W^{k,p}_0(U)$ as the set of $f \in W^{k,p}(U)$ such that "$\partial^\alpha U = 0$ on ∂U, $\forall |\alpha| \leqslant k-1$." Let $U \subset \mathbb{R}^d$ be bounded and open and $k \in \mathbb{N}$. We assume that $f \in C^k_b(U)$ and its derivatives of order up to $k-1$ extend to continous functions in $\bar{U}$ which vanish at ∂U. Then $f \in W^{k,p}_0(U)$ for $1 \leqslant p < \infty$.*

- *The map $J : L^p(U \times \Sigma_k) \to W^{-k,p}(U)$ defined below has norm 1:*

$$J((f_\alpha))(u) = \sum_{|\alpha| \leqslant k} \int_U f_\alpha \partial^\alpha u \, dm^d, \quad \forall u \in W^{k,p'}_0(U),$$

where $u \in W^{k,p'}_0(U)$ can be identified with the element $(\partial^\alpha u) \in L^{p'}(U \times \Sigma_k)$. If U is a Lipschitz domain, then for any $f \in W^{-k,p}(U)$, $p \in (1, \infty)$, there exists a unique $F \in W^{k,p}_0(U)$, such that

$$f = \sum_{|\alpha| \leqslant k} (-1)^{|\alpha|} \partial^{2\alpha} F, \quad \text{i.e. } f(g) = \sum_{|\alpha| \leqslant k} \int_U \partial^\alpha F \partial^\alpha g dm^d, \quad \forall g \in W^{k,p'}_0(U),$$

and hence $\mathcal{D}(U)$ is dense in $W^{-k,p}(U)$.

4.4.2 Extension and traces

In order to extend a Sobolev function defined in a bounded domain U to the whole space $\mathbb{R}^d$ or to take the trace of it on the boundary, we need some regularity assumption of the boundary.

Definition 4.34 (Boundary regularity of a domain)**.** *A domain in $\mathbb{R}^d$ is a non-empty open connected set.*

A bounded domain $U \subset \mathbb{R}^d$ is said to have a Lipschitz boundary if for each point $x_0 \in \partial U$ there exists $r > 0$ and a one-to-one function $\phi : B_r(x_0) \to D \subset \mathbb{R}^d$ such that

- $\phi(B_r(x_0) \cap U) \subset \mathbb{R}^d_-$;
- $\phi(B_r(x_0) \cap \partial U) \subset \partial \mathbb{R}^d_-$;
- ϕ, ϕ^{-1} *are Lipschitz continuous.*

We then call U a bounded Lipschitz-domain. Similarly a bounded domain has C^k-boundary if the functions $\phi, \phi^{-1} \in C^k$.

Theorem 4.35 (Whitney extenstion theorem)**.** *Let $1 \leqslant p < \infty$ and $U \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Then there exists a linear (continuous) extension map*

$$W^{k,p}(U) \to W^{k,p}(\mathbb{R}^d).$$

Sketchy proof. We sketch the proof under the stronger assumption that ∂U is a C^k manifold. By use of compactness, we take the partition of unity (η_m) on $\overline{U}$ such that for any function f defined on U:

$$f = \sum_{m=0}^{N} (\eta_m f), \quad \mathrm{Supp}\,(\eta_0) \subset\subset U, \quad \mathrm{Supp}\,(\eta_m) \subset B_{r_m}(x_m),\, m = 1, \cdots, N.$$

By the one-to-one C^k-map $\phi_m : B_{r_m}(x_m) \to D_m \subset \mathbb{R}^d$ and the Leibniz' rule as well as the chain rule (4.4), it suffices to extend Sobolev functions $g_m := (\eta_m f) \circ (\phi_m^{-1}) \in W^{k,p}(\mathbb{R}^d_-)$ defined on the lower half space $\mathbb{R}^d_-$ continuously to Sobolev functions $\tilde{g}_m \in W^{k,p}(\mathbb{R}^d)$ defined the whole space $\mathbb{R}^d$, since then

$$\tilde{f} = (\eta_0 f) + \sum_{m=1}^{N} \tilde{g}_m \circ \phi_m \in W^{k,p}(\mathbb{R}^d),$$
$$\|\tilde{f}\|_{W^{k,p}(\mathbb{R}^d)} \leqslant C \|f\|_{W^{k,p}(U)}.$$

Case $g \in C^k$. Let $V := \overline{\mathbb{R}^d_-}$, and $g \in C^k(V)$. We make the Ansatz

$$\tilde{g}(x) = \begin{cases} g(x) & \text{if } x_d \leqslant 0 \\ \sum\limits_{j=1}^{k+1} a_j g(x_1, \dots x_{d-1}, -j x_d) & \text{if } x_d > 0. \end{cases}$$

If $g \in C^k(V)$ we want to choose the a_j so that $\partial^\alpha \tilde{g}$ is continuous for $|\alpha| \leqslant k$. It suffices to do this for $\partial^j_{x_d}$ for $0 \leqslant j \leqslant k$ when $\{x_d = 0\}$, which leads to the

Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 3 & \dots & k+1 \\ 1^2 & 2^2 & 3^2 & \dots & (k+1)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1^k & 2^k & 3^k & \dots & (k+1)^k \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{k+1} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ \dots \\ (-1)^k \end{pmatrix}.$$

The Vandermonde matrix is invertible and we can solve this system. The coefficients (a_j) hence exist and depend only on k. Then

$$\|\tilde{g}\|_{W^{k,p}(\mathbb{R}^d)} \leqslant C\|g\|_{W^{k,p}(V)}.$$

Case $g \in W^{k,p}(\mathbb{R}^d_-)$. We define the extension $\tilde{g}$ in the same way. Then we have to prove that for $|\alpha| \leqslant k$ the distributional derivative is given by the distribution defined by the distributional derivatives on both $\mathbb{R}^d_+$ and $\mathbb{R}^d_-$. By an application of the theorem of Fubini and using distributional derivatives in $d-1$ variables this is true for all α with $\alpha_d = 0$, and it suffices to consider $d = 1$. Let $1 \leqslant \kappa \leqslant k$ and $\varphi \in \mathcal{D}(\mathbb{R})$. Then, with

$$\tilde{g}_\kappa(x) := \begin{cases} \frac{d^\kappa}{dx^\kappa} g(x) & \text{if } x < 0 \\ \sum_{j=1}^{k+1} a_j (-j)^\kappa \big(\frac{d^\kappa}{dx^\kappa} g\big)(-jx) & \text{if } x > 0 \end{cases}$$

we have to check **(Exercise)**

$$(-1)^\kappa \int_{\mathbb{R}} \tilde{g} \frac{d^\kappa}{dx^\kappa} \varphi dx = \int_{\mathbb{R}} \tilde{g}_\kappa \varphi(x) dx.$$

□

[30.01.2023]

[03.02.2023]

We have the following density result for the embedding $C^\infty(\overline{U}) \subset W^{k,p}(U)$ from Theorem 4.32.

Corollary 4.36 (Density of smooth function space in Sobolev spaces)**.** *Suppose that U is a open, bounded with Lipschitz boundary, $k \in \mathbb{N}$ and $1 \leqslant p < \infty$. Then the restrictions of $C_c^\infty(\mathbb{R}^d)$ functions is dense in $W^{k,p}(U)$.*

We can take the trace $f|_{\partial U}$ of a continuous function $f \in C(\overline{U})$ on the boundary in the usual sense. Recall the Gaussian Integral Theorem: Any function $u \in C^1(\overline{U})$ defined in a C^1-domain U satisfies

$$\int_{\partial U} u \nu_j d\sigma = \int_U \partial_{x_j} u dx, \quad \forall 1 \leqslant j \leqslant d,$$

where ν denotes the outer normal vector of ∂U. Consequently for $f, F \in C^1(\overline{U})$,

$$\int_{\partial U} (fF)|_{\partial U} \nu_j d\sigma = \int_U \partial_{x_j}(fF)dx = \int_U (\partial_{x_j} fF + f\partial_{x_j} F)dx, \quad \forall 1 \leqslant j \leqslant d.$$

Hence if $F|_{\partial U} = 0$, then

$$\int_U (\partial_{x_j} f)F \,\mathrm{d}x = -\int_U (f\partial_{x_j} F)dx, \quad \forall 1 \leqslant j \leqslant d,$$

which coincides with (4.3). If $F^j \in C^1(\bar{U})$, $j = 1, \cdots, d$, then

$$\int_{\partial U} (f|_{\partial U})(\sum_{j=1}^d F^j \nu^j)d\sigma = \int_U \Big(f \sum_{j=1}^d \partial_{x_j} F^j + \sum_{j=1}^d \partial_{x_j} fF^j\Big)dx.$$

This motivates us to define the trace of $f \in W^{1,p}(U)$ as follows.

Theorem 4.37 (Traces)**.** *Let U be a bounded C^1 domain and let $f \in W^{1,p}(U)$, $1 \leqslant p < \infty$. Then there is a unique trace $g \in L^p(\partial U)$ so that*

$$\int_{\partial U} (\sum_{j=1}^d F^j \nu^j) g d\mathcal{H}^{d-1} = \int_U \Big(f \sum_{j=1}^d \partial_{x_j} F^j + \sum_{j=1}^d \partial_{x_j} fF^j\Big)dm^d, \tag{4.5}$$

where $F^j \in C^1(\overline{U})$ and ν denotes the outer normal vector of ∂U. We denote $f|_{\partial U} = g$, and it satisfies

$$\|g\|_{L^p(\partial U)} \leqslant c\|f\|_{L^p(U)}^{\frac{p-1}{p}} (\|f\|_{L^p(U)} + \|Df\|_{L^p})^{\frac{1}{p}}, \quad \|Df\|_{L^p} := \|\,|(\partial_{x_j} f)|\,\|_{L^p}. \tag{4.6}$$

Sketchy proof. **Case $f \in C^1(\overline{U})$.** If $f \in C^1(\overline{U})$ then $g = f|_{\partial U}$ satisfies (4.5). By a partition of unity and the one-to-one map (ϕ_m), we may assume that U is below $\partial\mathbb{R}^d_-$ and f has compact support. If $p < \infty$, then (4.6) follows:

$$\begin{aligned}\|g\|^p_{L^p(\partial U)} =& \int_{\mathbb{R}^{d-1}} |g|^p d\sigma \\ =& p \operatorname{Re} \int_{\mathbb{R}^d} |f|^{p-2} \bar{f} \partial_{x_d} f dm^d \\ \leqslant& p\|f\|^{p-1}_{L^p} \|\partial_{x_d} f\|_{L^p}.\end{aligned}$$

Case $f \in W^{1,p}(U)$. For $f \in W^{1,p}(U)$, we approximate f by smooth functions $f_n \in C^1(\overline{U})$ in $W^{1,p}(U)$. Since (f_n) is a Cauchy sequence in $W^{1,p}(U)$, $(f_n|_{\partial U})$ is a Cauchy sequence in $L^p(\partial U)$ and hence we obtain a limit function $g \in L^p(\partial U)$ at the boundary, which satisfies (4.5) and (4.6). The uniqueness follows from that fact that if $\int_{\partial U} hg = 0$ for all $h \in C^1(\overline{U})$, then $g = 0$ in $L^p(\partial U)$. □

Remark 4.38. *If U is a bounded C^1-domain and $1 \leqslant p < \infty$, then $W_0^{1,p}(U)$ can be characterized as*

$$W_0^{1,p}(U) = \{f \in W^{1,p}(U) \,|\, f|_{\partial U} = 0\}.$$

4.4.3 Poincaré inequality and Sobolev inequality

We have seen in the trace theorem that, with some extra regularity assumption on the Lebesgue functions, we can take the trace of a Sobolev function $f \in W^{1,p}(U)$ at the boundary (it does not make sense to take trace of $f \in L^p(U)$). We are interested in the relation between a Sobolev function itself and its derivatives, in the framework of Lebesgue spaces.
Recall the Newton-Leibniz' formular: $f(x) = \int_0^x f'(t)dt$, for $x \in [0,1]$, if $f \in C^1([0,1])$ and $f(0) = 0$. This implies $\|f\|_{C_b([0,1])} \leqslant \|f'\|_{L^\infty}$. We have also similar estimates for $f \in W_0^{1,p}(U)$.

Theorem 4.39 (Poincaré inequality)**.** *Let $U \subset \mathbb{R}^d$ be a bounded domain. Then*

$$\|f\|_{L^p(U)} \leqslant \operatorname{diam}(U)\|Df\|_{L^p(U)}, \quad \forall f \in W_0^{1,p}(U),\, p \in [1,\infty),$$

where $\operatorname{diam}(U) = \sup\{|x-y| \,|\, x,y \in U\}$.

Proof. Let $a := \operatorname{diam}(U) < \infty$. Without loss of generality we assume $U \subset \{x = (x', x_d) | 0 < x_d < a\}$. By the density of $C_c^\infty(U)$ in $W_0^{1,p}(U)$, it suffices to show the inequality for $f \in C_c^\infty(U)$. We extend f trivially to $\mathbb{R}^d$, then

$$\begin{aligned}
\int_U |f(x)|^p \,\mathrm{d}x &= \int_{\mathbb{R}^{d-1}} \int_0^a |f(x', x_d)|^p \,\mathrm{d}x_d \,\mathrm{d}x' \\
&= \int_{\mathbb{R}^{d-1}} \int_0^a \left| \int_0^{x_d} \partial_{x_d} f(x', y) \,\mathrm{d}y \right|^p \mathrm{d}x_d \,\mathrm{d}x' \\
&\leqslant \int_{\mathbb{R}^{d-1}} \int_0^a \int_0^{x_d} |\partial_{x_d} f(x', y)|^p \,\mathrm{d}y \,(x_d)^{p-1} \,\mathrm{d}x_d \,\mathrm{d}x', \text{ by Hölder's inequality,} \\
&\leqslant a^p \int_U |\partial_{x_d} f|^p \,\mathrm{d}x\,.
\end{aligned}$$

□

Remark 4.40. *We have assumed "zero boundary condition" in Theorem 4.39. Similarly, we can assume "zero average condition" and show the Poincaré inequality in a ball **(Exercise)**: If $1 \leqslant p \leqslant \infty$ and $f \in W^{1,p}(B_R(0))$ then*

$$\|f - f_B\|_{L^p(B_R(0))} \leqslant 2^{\frac{d}{p}} R \|Df\|_{L^p(B_R(0))},$$

where $f_B := (m^d(B_R(0)))^{-1} \int_{B_R(0)} f dm^d$ denotes the average of f in $B_R(0)$.

The following Sobolev inequality implies the Sobolev embedding $W^{1,p}(\mathbb{R}^d) \subset L^q(\mathbb{R}^d)$, for $1 \leqslant p < d$ and Sobolev conjugate $q = \frac{dp}{d-p} \in (p, \infty)$.

Theorem 4.41 (Sobolev inequality)**.** *Suppose that $1 \leqslant p < d$ and $q \in (p, \infty)$ such that*

$$\frac{1}{q} + \frac{1}{d} = \frac{1}{p}.$$

Then there exists $c > 0$ such that

$$\|f\|_{L^q(\mathbb{R}^d)} \leqslant c\|Df\|_{L^p(\mathbb{R}^d)}, \quad \forall f \in W^{1,p}(\mathbb{R}^d).$$

Proof. By the density embedding $\mathcal{D}(\mathbb{R}^d) \subset W^{1,p}(\mathbb{R}^d)$, it suffices to show the inequality for $f \in \mathcal{D}(\mathbb{R}^d)$. We prove the estimate first for $p = 1$ and $q = \frac{d}{d-1}$. More precisely we prove the estimate

$$\|f\|^d_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} \leqslant 2^{-d} \prod_{j=1}^{d} \|\partial_j f\|_{L^1(\mathbb{R}^d)} \tag{4.7}$$

by induction on the dimension. The case $d = 1$ follows straightforward from Newton-Leibniz' formular:

$$f(x) = \int_{-\infty}^{x} f'(t)dt = -\int_{x}^{\infty} f'(t)dt, \quad \forall f \in \mathcal{D}(\mathbb{R}).$$

Suppose we have proven the estimate for $d \leqslant k - 1$. Then by Fubini and Hölder's inequality (since $\frac{1}{k-1} + \frac{k-2}{k-1} = 1$)

$$\begin{aligned}
\|f\|^{\frac{k}{k-1}}_{L^{\frac{k}{k-1}}(\mathbb{R}^k)} &= \int_{\mathbb{R}} \int_{\mathbb{R}^{k-1}} |f|^{\frac{1}{k-1}} |f| dm^{k-1} dm^1(x_1) \\
&\leqslant \int_{\mathbb{R}} \|f(x_1, \dots)\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^{k-1})} \|f(x_1, \dots)\|_{L^{\frac{k-1}{k-2}}(\mathbb{R}^{k-1})} dm^1(x_1) \\
&\leqslant \sup{}_{x_1} \|f(x_1, .)\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^{k-1})} \int_{\mathbb{R}} \|f(x_1, .)\|_{L^{\frac{k-1}{k-2}}(\mathbb{R}^{k-1})} dm^1(x_1) \\
&\leqslant 2^{-\frac{1}{k-1}} \|\partial_{x_1} f\|_{L^1(\mathbb{R}^k)} \int_{\mathbb{R}} \left(2^{-(k-1)} \prod_{j=2}^{k} \|\partial_{x_j} f(x_1, .)\|_{L^1(\mathbb{R}^{k-1})} \right)^{\frac{1}{k-1}} dm^1(x_1)
\end{aligned}$$

We take the inequality to the power $k - 1$, and apply Hölder's inequality in the form

$$\left(\int \prod_{j=2}^{k} |g_j|^{\frac{1}{k-1}} dm^1 \right)^{k-1} \leqslant \prod_{j=2}^{k} \int |g_j| dm^1$$

to arrive at

$$\|f\|^k_{L^{\frac{k}{k-1}}(\mathbb{R}^k)} \leqslant 2^{-k} \prod_{j=1}^k \|\partial_j f\|_{L^1(\mathbb{R}^k)}.$$

Now let $1 < p < d$. We apply the above inequality (4.7) to $|f|^{\frac{(d-1)p}{d-p}}$. Then

$$\begin{aligned}
\|f\|^{\frac{(d-1)p}{d-p}}_{L^q(\mathbb{R}^d)} =& \||f|^{\frac{(d-1)p}{d-p}}\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} \\
\leqslant& \|D|f|^{\frac{(d-1)p}{d-p}}\|_{L^1(\mathbb{R}^d)} \leqslant \frac{(d-1)p}{d-p} \int |f|^{\frac{d(p-1)}{d-p}} |Df| dm^d \\
\leqslant& \frac{(d-1)p}{d-p} \|f\|^{\frac{d(p-1)}{d-p}}_{L^q(\mathbb{R}^d)} \|Df\|_{L^p(\mathbb{R}^d)}
\end{aligned}$$

where we first argue for smooth functions, and where we used Hölder's inequality in the last step. □

Remark 4.42. *For $p > d$, we have Morrey's inequality for $\tilde{f} \in W^{1,p}(U)$, $U \subset \mathbb{R}^d$ open: Every point is a Lebesgue point and the canonnical representative f is continuous, such that the following is true for all $x, y \in U$ with $|x-y| <$* $\operatorname{dist}(x, \mathbb{R}^d \backslash U)$

$$|f(x) - f(y)| \leqslant c|x-y|^{1-\frac{d}{p}} \|Df\|_{L^p(B_{|x-y|}(x))}.$$

This implies indeed the Hölder continuity $f \in C^{1-\frac{d}{p}}(U)$ for $p \in (d, \infty)$, which is almost everywhere differentiable, with the derivative the same as its weak derivative almost everywhere. If $p = \infty$, then Lipschitz continuous functions $f \in W^{1,\infty}(U)$ are almost everywhere differentiable.

[03.02.2023]

[06.02.2023]

5 Hahn-Banach Theorem, Reflexivity and Weak Topology

Recall the normed space of continuous linear maps $L(X, Y)$ between two normed spaces X, Y, where the norm in $L(X, Y)$ is given by

$$\|T\|_{X \to Y} = \sup{}_{\|x\|_X \leqslant 1} \|Tx\|_Y.$$

If $Y = \mathbb{K}$, then we denote $X^* = L(X, \mathbb{K})$ (as the dual space of X), which is a Banach space, equipped with the norm

$$\|x^*\|_{X^*} = \sup{}_{\|x\|_X \leqslant 1} |x^*(x)|.$$

We have seen that $(l^p(\mathbb{N}))^*$ is isomorphic to $l^{p'}(\mathbb{N})$, $\frac{1}{p} + \frac{1}{p'} = 1$ for $p \in [1, \infty)$ (see Theorem 3.24 and Corollary 3.25). In the exercise sheet $(c_0(\mathbb{N}))^*$ can be characterized with $l^1(\mathbb{N})$. We are going to show that $(l^\infty)^* \neq l^1(\mathbb{N})$, as a consequence of the celebrated Hahn-Banach Theorem. It is one of the fundamental theorems in functional analysis.

5.1 Hahn-Banach Theorems

In this section we will introduce Hahn-Banach theorems, first in the purely algebraic setting (for both real and complex cases), and then in the framework of normed spaces (which is referred as Hahn-Banach theorem by default). We will then give some interesting consequences of Hahn-Banach Theorem.

5.1.1 Hahn-Banach theorems

Definition 5.1. *Let X be a $\mathbb{K}$ vector space. A map $p : X \to \mathbb{R}$ is called sublinear if*

- $p(\lambda x) = \lambda p(x)$ *for* $x \in X$ *and* $\lambda \geqslant 0$,
- $p(x + y) \leqslant p(x) + p(y)$ *for* $x, y \in X$.

Example 5.2. 1. *The norm of a normed space is sublinear.*

2. *If $\mathbb{K} = \mathbb{R}$, any element in the dual space of a normed space $X^* = \{$continuous linear functional $T : X \to \mathbb{R}\}$ is sublinear.*
3. *Let $C \subset X$ be convex such that for every $x \in X$ there exists $\lambda > 0$ so that $\lambda x \in C$. The Minkowski functional*

$$p_C(x) = \inf\{\lambda > 0 : \frac{1}{\lambda} x \in C\} \in [0, \infty)$$

is sublinear. If X is a normed space, then its norm is the Minkowski functional of the unit ball.

Theorem 5.3 (Hahn-Banach, real case)**.** *Let X be a real vector space, $Y \subset X$ a linear subspace, $p : X \to \mathbb{R}$ sublinear and $l : Y \to \mathbb{R}$ linear such that*

$$l(y) \leqslant p(y) \qquad \text{for all } y \in Y.$$

Then there exists $L : X \to \mathbb{R}$ linear so that

1. *$L(y) = l(y)$ for all $y \in Y$*

2. $L(x) \leqslant p(x)$ for all $x \in X$.

Proof. There are two very different steps.
Step 1. Suppose that $Y \neq X$. Then there exists $x_0 \in X \backslash Y$. Let Y_1 be the space spanned by Y and x_0. Every element of Y_1 can uniquely be written as

$$y + sx_0, \quad y \in Y, s \in \mathbb{R}\,.$$

In Step 1 we aim to find a linear map $l_1 : Y_1 \to \mathbb{R}$ such that

- $l_1(y) = l(y)$ for $y \in Y$
- $l_1(y + sx_0) \leqslant p(y + sx_0)$ for $s \in \mathbb{R}$ and $y \in Y$.

It suffices to find $t = l_1(x_0)$ so that

$$l(y) + st \leqslant p(y + sx_0), \quad \forall y \in Y,\ s \in \mathbb{R}\,.$$

Notice the following two equivalence relations

$$\begin{aligned} &st \leqslant s(p(y/s + x_0) - l(y/s)), \quad \forall s > 0, y \in Y \\ &\iff t \leqslant \inf_{y \in Y}\{p(y + x_0) - l(y)\}, \end{aligned}$$

and

$$\begin{aligned} &-st \geqslant -s(l(-y/s) - p(-y/s - x_0)), \quad \forall s < 0, y \in Y \\ &\iff t \geqslant \sup{}_y\{l(y) - p(y - x_0)\}. \end{aligned}$$

Hence we can find t if and only if

$$l(y) - p(y - x_0) \leqslant p(\tilde{y} + x_0) - l(\tilde{y}) \qquad \text{for all } y, \tilde{y} \in Y$$

which follows from the inequality on Y and sublinearity of p:

$$l(y) + l(\tilde{y}) = l(y + \tilde{y}) \leqslant p(y + \tilde{y}) \leqslant p(y + x_0) + p(\tilde{y} - x_0).$$

Step 2. We need the axiom of choice in the form of Zorn's lemma: Let Z be a partially ordered set which contains an upper bound for every chain. Then there is a maximal element. A partially ordered set Z is a set equipped with the relation $\leqslant$, which satisfies $a \leqslant a$ for all $a \in Z$, $a \leqslant b \& b \leqslant a \Rightarrow a = b$ and $a \leqslant b \& b \leqslant c \Rightarrow a \leqslant c$. Here a chain is a totally ordered subset, that is a subset A where either $a \leqslant b$ or $b \leqslant a$ holds for any two elements $a, b \in A$. An element b is an upper bound for the chain A, if $a \leqslant b$ for all $a \in A$. An element $a \in Z$ is maximal if $b \in Z$, $b \geqslant a$ implies $b = a$.

We define

$$Z = \{(W, l_W) : Y \subset W \subset X,\, l_W|_Y = l,\, l_W(w) \leqslant p(w) \text{ for } w \in W\},$$

with the ordering

$$(W, l_W) \leqslant (V, l_V) \qquad \text{if } W \subset V \quad \text{ and } l_V|_W = l_W.$$

This is a partial order. If $\tilde{Z}$ is a chain then

$$V = \bigcup_{(W, l_W) \in \tilde{Z}} W$$

with the obvious l_V being an upper bound for the chain. Now let (V, l_V) be a maximal element. If $V = X$ we are done. Otherwise we obtain a contradiction by the first step. □

Theorem 5.4 (Hahn-Banach, complex version)**.** *Let X be a complex vector space, Y a subvector space, $p : X \to \mathbb{R}$ sublinear and $l : Y \to \mathbb{C}$ linear so that*

$$\operatorname{Re} l(y) \leqslant p(y) \text{ for } y \in Y.$$

Then there exists $L : X \to \mathbb{C}$ linear so that

1. $L|_Y = l$
2. $\operatorname{Re} L(x) \leqslant p(x)$.

Proof. We first notice that any linear map $T : X \to \mathbb{C}$ is uniquely determined by its real part $R = \operatorname{Re} T$:

$$T(x) = R(x) - iR(ix).$$

Let X be viewed as $\mathbb{R}$-vector space, and Y $\mathbb{R}$-subvector space. We apply Theorem 5.3 to the real part $\operatorname{Re} l : Y \to \mathbb{R}$ and derive $A : X \to \mathbb{R}$ such that

$$Ay = \operatorname{Re}(l(y)) \text{ for all } y \in Y \text{ and } Ax \leqslant p(x) \text{ for all } x \in X.$$

The linear map $L(x) = A(x) - iA(ix)$ is the searched extension map. □

The above Hahn-Banach theorems are purely algebraic. We formulate the consequences for normed vector spaces, making use of the fact that norms are sublinear.

Theorem 5.5 (Hahn-Banach)**.** *Let X be a normed $\mathbb{K}$ vector space, Y a subspace endowed with the norm of X, and $y^* \in Y^* = L(Y, \mathbb{K})$ a continuous linear map.*
Then there exists $x^* \in X^* = L(X, \mathbb{K})$ so that

1. $x^*|_Y = y^*$, *i.e.* $x^*(y) = y^*(y)$ *for all* $y \in Y$,

2. $\|x^*\|_{X^*} = \|y^*\|_{Y^*}$.

Proof. We define

$$p(x) = \|y^*\|_{Y^*}\|x\|_X,$$

such that

$$\operatorname{Re} y^*(y) \leqslant p(y), \quad \forall y \in Y.$$

We apply Theorem 5.4 to obtain $x^* : X \to \mathbb{K}$ a linear map so that $x^*|_Y = y^*$ and

$$\operatorname{Re} x^*(x) \leqslant p(x).$$

This implies that for all $x \in X$, there exists $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that

$$|x^*(x)| = \operatorname{Re} \alpha x^*(x) = \operatorname{Re} x^*(\alpha x) \leqslant p(\alpha x) = p(x) = \|y^*\|_{Y^*}\|x\|_X,$$

and hence $\|x^*\|_{X^*} \leqslant \|y^*\|_{Y^*}$ such that $x^* \in X^*$. Since $x^*|_Y = y^*$,

$$\|x^*\|_{X^*} = \|y^*\|_{Y^*}.$$

□

Example 5.6. $(l^\infty(\mathbb{N}))^* \neq l^1(\mathbb{N})$.
The space of converging sequences c *is a closed subspace of* l^∞. *Let* $l : c \to \mathbb{K}$ *be defined by*

$$l((x_j)) = \lim_{j\to\infty} x_j.$$

Then $l \in c^*$ *and*

$$|l((x_j))| \leqslant \|(x_j)\|_{l^\infty}$$

for every converging sequence. By Theorem 5.5 it has an extension $L \in (l^\infty)^*$. *Clearly* $L(e_j) = l(e_j) = 0$. *Hence* L *cannot be represented by a sequence* $y = (y_j) \in l^1$ *since otherwise*

$$L((x_j)) = \sum_{j=1}^{\infty} \overline{y_j} x_j \Longrightarrow L(e_n) = \overline{y_n} = 0 \text{ for all } n.$$

[06.02.2023]

[10.02.2023]

5.1.2 Consequences of Hahn-Banach theorem

We state some consequences of Hahn-Banach Theorem.

Lemma 5.7. *If X is a normed space, $Y \subset X$ is a closed subspace with $Y \neq X$ and $x_0 \in X \backslash Y$, then there exists $x^* \in X^*$, with $x^*|_Y = 0$, $\|x^*\|_{X^*} = 1$ and $x^*(x_0) = \operatorname{dist}(x_0, Y) > 0$.*

Proof. Let $\tilde{Y}$ be the span of Y and x_0. Since Y is closed, $a := \operatorname{dist}(x_0, Y) > 0$. We define

$$y^*(y + tx_0) = ta, \quad \forall (y + tx_0) \in \tilde{Y}.$$

We claim that $y^* \in \tilde{Y}^*$ with $\|y^*\|_{\tilde{Y}^*} = 1$. Obviously y^* is linear. For $t = 0$, $|y^*(y + tx_0)| = 0 \leqslant \|y + tx_0\|_X$. For $t \neq 0$,

$$|y^*(y + tx_0)| = |t|a = |t|a\frac{\|y + tx_0\|_X}{\|y + tx_0\|_X} = d\frac{\|y + tx_0\|_X}{\|\frac{y}{t} + x_0\|_X} \leqslant \|y + tx_0\|_X.$$

Now let $(y_n) \subset Y$ be a sequence such that $a = \lim_n \|x_0 - y_n\|_X$. Then

$$\|y^*\|_{\tilde{Y}^*} \geqslant \left| y^*\left(\frac{x_0 - y_n}{\|x_0 - y_n\|_X}\right)\right| = \frac{a}{\|x_0 - y_n\|_X} \to 1,$$

which implies $\|y^*\|_{\tilde{Y}^*} = 1$.
Hence by Theorem 5.5 we can extend it to $x^* \in X^*$, with $x^*|_Y = y^*|_Y = 0$, $x^*(x_0) = y^*(x_0) = a$ and $\|x^*\|_{X^*} = \|y^*\|_{\tilde{Y}^*} = 1$. □

We have the following interesting descriptions of some specific subsets of X by use of elements in X^*.

Corollary 5.8. *Let X be a normed space. Then*

(a) For any $x \in X$, there exists $x^* \in X^*$ with $\|x^*\|_{X^*} = 1$ and $x^*(x) = \|x\|_X$. Hence for $x_1, x_2 \in X$ with $x_1 \neq x_2$, there exists $x^* \in X^*$ with $\|x^*\|_{X^*} = 1$ and $x^*(x_1) \neq x^*(x_2)$.

(b) The norm can be characterized as

$$\|x\|_X = \sup\{\operatorname{Re} x^*(x) : x^* \in X^*, \|x^*\|_{X^*} = 1\}.$$

Hence there is a cannonical isometry

$$J : X \to X^{**} = (X^*)^*, \quad J(x)(x^*) = x^*(x) \text{ with } \|J(x)\|_{X^{**}} = \|x\|_X.$$

(c) For $A \subset X$,

$$A \text{ is bounded} \iff x^*(A) \text{ is bounded for all } x^* \in X^*.$$

(d) For $Y \subset X$ a subvector space,

$$Y \text{ is not dense in } X \iff \exists x^* \in X^* \backslash \{0\} \text{ with } x^*(y) = 0, \quad \forall y \in Y.$$

(e) If X^* is separable, then X is separable.

Proof. (a) Let $Y = \{0\}$. If $x \neq 0$, then by Lemma 5.7 there exists $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \operatorname{dist}(x, Y) = \|x\|$. If $x = 0$ we choose $x_0 \neq 0$ and find x^* for x_0. Consequently, if $x_1 \neq x_2$, we find x^* for $x_0 = x_1 - x_2$.
(b) The characterization for the norm on X follows from (a) and the fact $|Re x^*(x)| \leqslant \|x^*\|_{X^*}\|x\|_X$.
Correspondingly we compute

$$\|J(x)\|_{X^{**}} = \sup_{\|x^*\|_{X^*}=1} \operatorname{Re} J(x)(x^*) = \sup_{\|x^*\|_{X^*}=1} x^*(x) = \|x\|_X.$$

(c) "$\Longrightarrow$" If $A \subset X$ is bounded and $x^* \in X^*$, then obviously $x^*(A)$ is bounded in $\mathbb{K}$.
"$\Longleftarrow$" Now if $x^* \in X^*$ and $x^*(A)$ is bounded, then

$$\sup_{x \in A} |J(x)(x^*)| = \sup_{x \in A} |x^*(x)| < \infty.$$

As X^* is a Banach space, Banach-Steinhaus Theorem 4.2 implies

$$\sup_{x \in A} \|J(x)\|_{X^{**}} < \infty.$$

Hence A is bounded.
(d) "$\Longrightarrow$" It follows from Lemma 5.7.
"$\Longleftarrow$" If Y is dense, then $x^*|_Y = 0$ implies $x^*|_X = 0$ by continuity.
(e) Since X^* is separable, its subset is also separable, and in particular the unit sphere in X^* is separable: Let (x_j^*) be a dense sequence of unit vectors in the unit sphere. We can take a sequence $(x_j) \subset X$ with $\|x_j\|_X = 1$ such that $x_j^*(x_j) \geqslant \frac{1}{2}$. We claim that the span Y of (x_j) is dense in X. Indeed otherwise by (d) there exists $x^* \in X^*$ such that $x^*|_{\overline{Y}} = 0$ (by continuity) and $\|x^*\|_{X^*} = 1$ (by linearity). Hence

$$x^*(x_j) = x_j^*(x_j) + (x^* - x_j^*)(x_j),$$

which can be chosen bigger than $\frac{1}{4}$ in view of $x_j^*(x_j) \geqslant \frac{1}{2}$ and the density of (x_j^*) in the unit sphere. This is contradiction to $x^*|_{\overline{Y}} = 0$. Hence Y is dense in X and $\{qx_j \,|\, j \in \mathbb{N},\ q \in \mathbb{Q} \text{ or } \mathbb{Q} + i\mathbb{Q}\}$ is a countable dense set in X. □

5.2 Reflexivity, weak and weak-$*$ topology

5.2.1 Reflexive spaces

Recall the linear isometry

$$J : X \to X^{**}, \quad x \mapsto x^*(x), \quad \forall x^* \in X^*. \tag{5.1}$$

As $\overline{J(X)}^{\|\|_{X^{**}}}$ is a closed linear subspace of X^{**}, $\overline{J(X)}^{\|\|_{X^{**}}}$ is a Banach space. The map J is not necessarily surjective, as seen from the example that $X = l^1(\mathbb{N})$, $X^* = l^\infty(\mathbb{N})$, $X^{**} \neq l^1(\mathbb{N})$. This motivates us to define

Definition 5.9. *A normed space X is called reflexive if the evaluation map $J : X \to X^{**}$ is surjective.*

Hence a reflexive normed space is automatically Banach space. $l^p(\mathbb{N})$ is reflexive space for $1 < p < \infty$, while not reflexive if $p = 1$ or ∞.

Lemma 5.10. *The following hold:*

(a) Let X be reflexive Banach space, and $Y \subset X$ be a closed subspace. Then Y is reflexive.

(b) A Banach space X is reflexive if and only if X^* is reflexive.

Proof. (a) Let $Y \subset X$ be a closed subspace and $Y \neq X$. For any $x^* \in X^*$ the restriction on Y belongs to Y^* with

$$\|x^*|_Y\|_{Y^*} \leqslant \|x^*\|_{X^*}.$$

Let $y^{**} \in Y^{**}$, we define $x^{**} \in X^{**}$ by

$$x^{**}(x^*) = y^{**}(x^*|_Y), \quad \forall x^* \in X^*. \tag{5.2}$$

Since X is reflexive, there exists $x \in X$ such that

$$x^{**}(x^*) = x^*(x), \quad \forall x^* \in X^*. \tag{5.3}$$

We claim that $x \in Y$. If $x \notin Y$, then by Lemma 5.7 there exists $\tilde{x}^* \in X^*$ such that $\tilde{x}^*|_Y = 0$ and $\tilde{x}^*(x) = 1$. This is impossible due to (5.2) and (5.3). Thus J_Y is surjective: For any $y^* \in Y^*$, by Hahn-Banach Theorem 5.5 there exists $x_1^* \in X^*$ such that $x_1^*|_Y = y^*$, and hence

$$y^{**}(y^*) = x^{**}(x_1^*) = J(x)(x_1^*) = x_1^*(x) = y^*(x) = J(x)(y^*), \quad \forall y^* \in Y^*.$$

Thus Y is reflexive.

[10.02.2023]
[13.02.2023]

(b) "⇒" We observe that for $x^{***} \in X^{***}$, we can define $x^* \in X^*$ by

$$x^*(x) = x^{***}(J(x)), \quad \forall x \in X.$$

If $J : X \to X^{**}$ is surjective, then for any $y^{**} \in X^{**}$ there exists a unique $y \in X$ such that $y^{**} = J(y)$, and hence

$$x^{***}(y^{**}) = x^{***}(J(y)) = x^*(y) = y^{**}(x^*), \quad \forall y^{**} \in X^{**}.$$

Thus X^* is reflexive.
"⇐" Since X is a Banach space, $J(X)$ is a closed subspace of X^{**}. If X is not reflexive, then by Lemma 5.7 there exists $x^{***} \in X^{***} \backslash \{0\}$ such that

$$x^{***}(J(y)) = 0, \quad \forall y \in X.$$

Since X^* is reflexive, there exists $x^* \in X^*$ such that

$$x^{***}(J(y)) = J(y)(x^*) = x^*(y), \quad \forall y \in X.$$

Thus $x^* = 0$, which is contradiction to $x^{***} \neq 0$. Hence X is reflexive. □

Remark 5.11. *If $Y \subset X$ are two normed spaces such that $\|y\|_X \leqslant C\|y\|_Y$, then $X^* \subset Y^*$ and for any $x^* \in X^*$,*

$$\|x^*\|_{Y^*} = \sup_{\|y\|_Y \leqslant 1} |x^*(y)| \leqslant \sup_{y \in Y,\, \|y\|_X \leqslant C} |x^*(y)| \leqslant C\|x^*\|_{X^*}.$$

Correspondingly $Y^{**} \subset X^{**}$ with

$$\|y^{**}\|_{X^{**}} \leqslant C\|y^{**}\|_{Y^{**}}.$$

If $\|\|_X = \|\|_Y$, then $\|x^*\|_{Y^*} \leqslant \|x^*\|_{X^*}$ but not necessarily equal.

Corollary 5.12. *Let $U \subset \mathbb{R}^d$ be open, $k \in \mathbb{N}$, $p \in (1, \infty)$. Then $W^{k,p}(U)$ and $W_0^{k,p}(U)$ are reflexive.*

Proof. We first claim that reflexivity of a Banach space is preserved by an isomorphic map. Let $j : X \to Y$ between two normed spaces is isomorphic, then $j^* : Y^* \ni y^* \mapsto y^* \circ j \in X^*$ is also isomorphic, and hence $j^{**} : X^{**} \ni x^{**} \mapsto x^{**} \circ j^* \in X^{**}$ is isomorphic. Thus if X is a reflexive Banach space, then $J_X : X \to X^{**}$ is isomorphic, which implies $J_Y = j^{**} \circ J_X \circ j^{-1} : Y \to Y^{**}$ is surjective.
$W^{k,p}(U)$ is isomorphic to a closed subspace of the reflexive Banach space $L^p(U \times \Sigma_k)$:

$$W^{k,p}(U) \ni f \mapsto (\partial^\alpha f)_{|\alpha| \leqslant k} \in L^p(U \times \Sigma_k),$$

and is hence reflexive. $W_0^{k,p}(U)$ is a closed subspace of $W^{k,p}(U)$, and is hence reflexive. □

5.2.2 Separation theorems

Let X be a normed space. Recall the sublinear Minkowski functional

$$p_C(x) = \inf\{\lambda > 0 : \frac{1}{\lambda}x \in C\} \in [0, \infty]$$

where $C \subset X$ is convex and we define the infimum of an empty set to be ∞.

Lemma 5.13. *Let X be a normed space and $C \subset X$ convex.*

(a) If $0 \in \mathring{C}$, then there exists $\delta > 0$ such that $B_\delta(0) \subset C$, and hence $p_C(x) \leqslant \delta^{-1}\|x\|_X$ for all $x \in X$.

(b) If C is open with $0 \notin C$, then there exists $x^* \in X^*$ such that $\operatorname{Re} x^*(x) < 0$ for all $x \in C$.

Proof. (a) follows immediately from the definition. (b)**Exercise.** □

Theorem 5.14 (Separation theorem)**.** *Let X be normed space. Let V, W be disjoint convex sets.*

(a) If V, W are open, there exists $x^* \in X^*$ such that

$$\operatorname{Re} x^*(w) < \operatorname{Re} x^*(v), \quad \forall v \in V,\ w \in W.$$

We say that x^* separates W and V.

(b) If V is closed and W is compact. Then there exists $x^* \in X^*$ such that

$$\sup{}_{w \in W} x^*(w) < \inf_{v \in V} x^*(v).$$

We say that x^* separates V and W strictly.

Proof. (a) Let $U = W - V = \{w - v \,|\, v \in V, w \in W\}$, then U is convex and open, and $0 \notin U$. By Lemma 5.13, there exists $x^* \in X^*$ such that $\operatorname{Re} x^*|_U < 0$, that is, x^* separates W and V.
(b) Since V is closed, for any $w \in W$ there exists $\varepsilon > 0$ such that $B_\varepsilon(w)$ is disjoint with V. Since W is compact there are finite open covering balls $(B_{\varepsilon_j/2}(w^{(j)}))_{j=1}^N$.
If $N = 1$, by (a) there exists $x^* \in X^*$ such that

$$\operatorname{Re} x^*(w_1) < \operatorname{Re} x^*(v), \quad \forall w_1 \in B_\varepsilon(w), \quad \forall v \in V.$$

Since there exists $y \in B_1(0)$ such that

$$\operatorname{Re} x^*(y) > \frac{1}{2}\|x^*\|_{X^*} \Rightarrow \operatorname{Re} x^*(\frac{\varepsilon}{2}y) > \frac{\varepsilon}{4}\|x^*\|_{X^*} =: \delta > 0,$$

we have for all $w_2 \in B_{\varepsilon/2}(w)$,

$$\operatorname{Re} x^*(w_2) = \operatorname{Re} x^*(w_2 + \frac{\varepsilon}{2} y) - \operatorname{Re} x^*(\frac{\varepsilon}{2} y) < \operatorname{Re} x^*(v) - \delta, \quad \forall v \in V,$$

and hence x^* separates W and V strictly.
If $N > 1$, it follows by replacing $B_\varepsilon(w)$ by $\cup_{j=1}^N B_{\varepsilon_j}(w^{(j)})$, and taking $\delta = \min_{j=1}^N \frac{\varepsilon_j}{4} \|x^*\|_{X^*}$. □

5.2.3 Weak topology

Definition 5.15. *Let X be a normed space. We call a sequence $(x_n) \subset X$ converges weakly against $x \in X$, denoted by $x_n \rightharpoonup x$, if*

$$x^*(x_n) \to x^*(x), \quad \forall x^* \in X^*.$$

Lemma 5.16. *Let X be a normed space.*

(a) Convergence in norm implies weak convergence.

(b) Weakly convergent sequences are bounded.

(c) Let C be a closed convex set in X, and $(x_n) \subset C$ be a sequence converging weakly to $x \in X$. Then $x \in C$.

(d) If X is Banach space, and the unit ball is uniformly convex, that is, for any $\varepsilon > 0$,

$$\delta(\varepsilon) := \inf\{1 - \frac{1}{2}\|x + y\|_X \mid \|x\|_X, \|y\|_X \leqslant 1,\ \|x - y\|_X > \varepsilon\} > 0,$$

then weak convergence $x_n \rightharpoonup x$ and convergence in norm $\|x_n\| \to \|x\|_X$ imply strong convergence $\|x_n - x\|_X \to 0$.

Proof. (a) It follows from the continuity of the map $x^* \in X^*$.
(b) Recall $x^*(x) = J(x)(x^*)$ for all $x \in X$, $x^* \in X^*$. If $x_n \rightharpoonup x$ in X, then for any fixed $x^* \in X^*$ the continuous linear maps $J(x_n) : X^* \to \mathbb{K}$ between two Banach spaces converges:

$$J(x_n)(x^*) \to J(x)(x^*).$$

By Banach-Steinhaus Theorem 4.2, we have

$$\sup{}_n \|x_n\|_X = \sup{}_n \|J(x_n)\|_{X^{**}} < \infty.$$

(c) It suffices to consider $\mathbb{K} = \mathbb{R}$. We assume by contraction that $x \notin C$, then by Theorem 5.14 there exists $x^* \in X^*$ such that

$$x^*(x) < \inf_{y \in C} x^*(y).$$

This is a contradiction to the weak convergence of x_n to x.
(d) If $\|x\| = 0$, then $\|x_n\| \to \|x\|$ implies $x_n \to x$.
If $\|x\| > 0$, then without loss of generality we assume $\|x\|_X = 1$, and $x_n \rightharpoonup x$ with $\|x_n\|_X \to 1$. Replacing x_n by $\frac{x_n}{\|x_n\|_X}$ we may assume $\|x_n\|_X = 1$.

[13.02.2023]

[17.02.2023]

By applying Lemma 5.13-(b) with $C = B_1(x)$, there exists $x^* \in X^*$ with

$$\operatorname{Re} x^*(x - y) > 0 \Rightarrow \operatorname{Re} x^*(x) > \operatorname{Re} x^*(y), \quad \forall y \in B_1(0).$$

Hence $\operatorname{Re} x^*(x) > 0$, and without loss of generality one may assume $1 = \operatorname{Re} x^*(x)$, which implies $\|x^*\|_{X^*} = 1$.
We assume by contradiction that (x_n) is not a Cauchy sequence, and hence there exists $\varepsilon > 0$ such that for any $n \in \mathbb{N}$ there exists $m > n$ with $\|x_n - x_m\|_X > \varepsilon$. By uniform convexity of the unit ball, $\frac{1}{2}\|x_n + x_m\|_X \leqslant 1 - \delta(\varepsilon)$, and hence

$$\left| x^*(\frac{1}{2}(x_n + x_m)) \right| \leqslant 1 - \delta(\varepsilon).$$

This is a contradiction to $\operatorname{Re} x^*(\frac{1}{2}(x_n + x_m)) \to \operatorname{Re} x^*(x) = 1$.
Similarly, the strong convergence for (x_n) which converges weakly to x follows from the uniform convexity. □

Remark 5.17. *By Hanner's inequality, the unit ball of the Lebesgue space $L^p(U)$ or the Sobolev space $W^{k,p}(U)$ with $p \in (1, \infty)$ is uniformly convex, and hence (d) is applicable.* ***Exercise.***

There exist weak converging sequence in a Banach space whose norms do not converge. Nevertheless, by Mazur's Theorem below, there exist its convex linear combinations which converges strongly.

Theorem 5.18 (Mazur)**.** *Let X be a Banach space, and (x_n) converges weakly to $x \in X$. Then there exist real numbers $(\lambda_k^{(n)}) \subset [0, 1]$ and $(N(n)) \subset \mathbb{N}$ such that*

$$\sum_{k=n}^{N(n)} \lambda_k^{(n)} = 1 \textit{ and } \sum_{k=n}^{N(n)} \lambda_k^{(n)} x_k \to x.$$

Proof. For any $n \in \mathbb{N}$, let C_n be the convex hull of $\{x_k\}_{k \geqslant n}$:

$$C_n = \left\{ \sum_{k=n}^{\infty} \mu_k x_k \,\middle|\, 0 \leqslant \mu_k, \quad \sum_{k=n}^{\infty} \mu_k = 1, \quad \text{only finitely many } \mu_k \text{ are nonzero} \right\}.$$

Then $\overline{C_n}$ is closed and convex in X. Since $x_k \in C_n$ for all $k \geqslant n$, $x \in \overline{C_n}$ by Lemma 5.16-(c). Thus for any $n \in \mathbb{N}$, there exists $y_n \in C_n$ such that $\|x - y_n\| \leqslant \frac{1}{n}$. □

5.2.4 Weak-* topology

Definition 5.19. *Let X be a normed space. A sequence $(x_n^*) \subset X^*$ converges weakly* to $x^* \in X$ if*

$$x_n^*(x) \to x^*(x), \quad \forall x \in X.$$

By Banach-Steinhaus Theorem 4.2, any weak-* convergent sequence is bounded in X^*. Conversely, any bounded sequence in X^* has a weak-* convergent subsequence:

Theorem 5.20 (Banach-Alaoglu, a simplified version)**.** *Let X be a separable normed space. Then every bounded sequence $x_n^* \subset X^*$ contains a weak* convergent subsequence.*

Proof. Let (x_k) be dense in X, and (x_n^*) be a bounded sequence in X^*. Since $x_n^*(x_1)$ is bounded, there exists a converging subsequence $x^*_{n_j^{(1)}}(x_1)$. Iteratively we obtain subsequences $n^{(k)} \subset n^{(k-1)}$ such that $\left(x^*_{n_j^{(k)}}(x_k)\right)_j$ converges for each $k \in \mathbb{N}$. Let

$$y_m^* = x^*_{n_m^{(m)}}, \quad \forall m \in \mathbb{N},$$

then $(y_m^*(x_k))_m$ converges for all $k \in \mathbb{N}$. Since (x_k) is dense and $(y_m^*) \subset X^*$ (as a subsequence of (x_n^*)) is bounded, by continuity (y_m^*) weakly-* converges. □

Remark 5.21. *Without the separability assumption, and one can show that the closed unit ball in the dual space of a normed space $\overline{B_1^{X^*}(0)} \subset X^*$ is compact in the weak-* topology. The version of Banach-Alaoglu Theorem 5.20 can fail without separability assumption. Let $X = \ell^\infty$, and $\varphi_n(x) = x_n \in X^*$ with $\|\varphi_n\|_{X^*} = 1$. For any subsequence $(\varphi_{n_k})_k$, with $y := (-1)^1 e_{n_1} + (-1)^2 e_{n_2} + \cdots \in \ell^\infty$,*

$$\varphi_{n_k}(y) = (-1)^k \text{ diverges.}$$

Hence there is no weak convergent subsequence.*

If X is reflexive, then $J : X \to X^{**}$ is isometry, such that

$$x^*(x) = J(x)(x^*), \quad \forall x \in X, \quad \forall x^* \in X^*,$$

and hence the weak and weak-* convergence are the same, both in X and X^*. Thus we have immediately the following

Corollary 5.22. *If X is reflexive and separable Banach space, then any bounded sequence has a weakly converging subsequence.*

Hence any bounded sequence in L^p, $1 < p < \infty$ has weakly converging subsequence.

Remark 5.23. *The separability assumption can be dropped: Let (x_n) be a bounded sequence in a reflexive Banach space, then the closure Y of the span of (x_n) is reflexive and separable, which yields the existence of weakly converging subsequence by Corollary 5.22.*

Example 5.24. *1. Let (X, μ) be measurable space, and $p \in (1, \infty)$. Then $L^p(\mu)$ is reflexive Banach space.*

- *The strong convergence of a sequence $f_n \overset{L^p(\mu)}{\longrightarrow} f$: $\|f_n - f\|_{L^p(\mu)} \to 0$ is equivalent to*

$$f_n \overset{L^p(\mu)}{\rightharpoonup} f \quad \& \quad \|f_n\|_{L^p(\mu)} \to \|f\|_{L^p(\mu)};$$

- *If $(X, \mu) = (U, m^d)$, then the weak (i.e. weak-*) convergence of a sequence $f_n \overset{L^p(U)}{\rightharpoonup} f$: $\int_X f_n g d\mu \to \int_X f g d\mu$, $\forall g \in L^{p'}(\mu)$, is equivalent to*

$$\sup{}_n \|f_n\|_{L^p(U)} < \infty \quad \& \quad \int_U f_n \varphi d\mu \to \int_X f \varphi d\mu, \quad \forall \varphi \in \mathcal{D}(U),$$

that is, $\sup_n \|f_n\|_{L^p(U)} < \infty \quad \& \quad T_{f_n} \overset{\mathcal{D}'(U)}{\to} T_f$.

2. Let $U \subset \mathbb{R}^d$ be open, $k \in \mathbb{N}$ and $p \in (1, \infty)$. Then $W^{k,p}(U)$ is isometric to a closed subspace of $L^p(U \times \Sigma_k)$ and hence is reflexive, separable Banach space. The weak convergence $f_n \overset{W^{k,p}}{\rightharpoonup} f$ is equivalent to

$$\sup{}_n \|f_n\|_{W^{k,p}(U)} < \infty \quad \& \quad T_{f_n} \overset{\mathcal{D}'(U)}{\to} T_f.$$

5.3 Introduction to spectral theory

We introduce here some basic definitions and facts (without proof) in spectral theory, which will be generalized and discussed in more detail in the lecture "Spectral Theory". [9]

[9] This section is not taken into account in the exam.

5.3.1 Compact operators

Definition 5.25 (Kernel and Range)**.** *Let X, Y be two normed spaces and $T \in L(X, Y)$. We denote the kernel of T by*

$$N(T) = \{x \in X \,|\, Tx = 0\}$$

and the range of T by

$$R(T) = \{Tx \,|\, x \in X\}.$$

We have the following facts concerning the kernel and range of a continuous linear operator and its adjoint operator.

Lemma 5.26. *Let X, Y be normed spaces.*

- *For $T \in L(X, Y)$, let $T^* : Y^* \to X^*$ be a linear map given by*

  $$T^*(y^*)(x) = y^*(Tx), \quad \forall x \in X.$$

  Then $\|T^*\|_{Y^* \to X^*} = \|T\|_{X \to Y}$, and

  $$N(T^*) = R(T)^{\perp} := \{y^* \in Y^* \,|\, y^*(y) = 0, \, \forall y \in R(T)\},$$
  $$N(T) = R(T^*)_{\perp} := \{x \in X \,|\, x^*(x) = 0, \, \forall x^* \in R(T^*)\}$$

  We call $T^* : Y^* \to X^*$ the adjoint of T.

- *Let $P \in L(X, X)$ be a projection (i.e. $P^2 = P$). Then $I - P \in L(X, Y)$ is also a projection, such that*

  $$R(P) = N(I - P), \quad N(P) = R(I - P),$$

  and $X = R(P) \bigoplus N(P)$ (i.e. $N(P), R(P)$ are closed subspace such that $N(P) + R(P) = X$ and $N(P) \cap R(P) = \{0\}$).

Definition 5.27 (Compact operator)**.** *Let X, Y be Banach spaces and $T \in L(X, Y)$. We call T compact if for every bounded sequence $(x_j) \subset X$, $(T(x_j)) \subset Y$ has a convergent subsequence.*

Remark 5.28.

- *If X, Y are finite dimensional, then $T \in L(X, Y)$ is compact. More generally, if $R(T)$ is finite dimensional then T is compact.*

- *The composition of a compact operator and a continuous operator is still compact.*

- *$T \in L(X, Y)$ is compact $\Longleftrightarrow T^* \in L(Y^*, X^*)$ is compact.*
- *The embedding $W_0^{1,p}(U) \subset L^p(U)$ with $U \subset \mathbb{R}^d$ a bounded open set is compact.*

Theorem 5.29 (Riesz-Schauder)**.** *Let X be Banach space and $K \in L(X)$ be compact. Let $T = I - K$, then $\dim N(T) = \dim N(T^*) < \infty$ and $R(T)$ is closed. Furthermore, T is invertible iff T is injective.*

5.3.2 Spectrum of compact operators

We generalize the eigenvalues of a matrix (as linear map from $\mathbb{C}^n$ to $\mathbb{C}^n$) to the spectrum of a continuous linear operator.

Definition 5.30. *Let X be Banach space and $T \in L(X, X) =: L(X)$. The resolvent set of T is given by*

$$\rho(T) = \{\lambda \in \mathbb{C} \mid \lambda I - T : X \to X \text{ is bijective}\},$$

and its spectrum by

$$\sigma(T) = \mathbb{C} \setminus \rho(T).$$

We further define the point spectrum of T by

$$\sigma_p(T) = \{\lambda \in \mathbb{C} \mid \exists x \in X \setminus \{0\} \text{ with } \lambda x = Tx\} \subset \sigma(T),$$

where we call $\lambda \in \sigma_p(T)$ an eigenvalue of T and the corresponding x an eigenvector or eigenfunction of T. For $\lambda \in \rho(T)$ the operator

$$R(\lambda, T) := (\lambda I - T)^{-1} : X \to X$$

and the set $\{R(\lambda, T) \,|\, \lambda \in \rho(T)\}$ are called resolvent.

Notice here that unlike the case of a $n \times n$-matrix M,

$$(\lambda I - M) \text{ is bijective} \Longleftrightarrow \lambda \in \mathbb{C} \text{ is not an eigenvalue of } M,$$

the operator T defined on a (infinite-dimensional) Banach space may have $\lambda \in \sigma(T) \setminus \sigma_p(T)$, which is in the spectrum of T (i.e. $(\lambda I - T)$ is not bijective) but is not an eigenvalue. For example, the right and left shift operators $R, L : \ell^p \to \ell^p$

$$Rx = (0, x_1, x_2, \cdots) \text{ and } Lx = (x_2, x_3, \cdots)$$

are continuous linear operators with the norm $\|R\|_{\ell^p \to \ell^p} = \|L\|_{\ell^p \to \ell^p} = 1$. Obviously $LR = I$ and $RLx = (0, x_2, x_3, \cdots)$, and hence R is injective (not surjective) and has left inverse (but no right inverse), while L is surjective (not injective) and has right inverse (but no left inverse). The right shift operator R has no eigenvalue, since for any $\lambda \in \mathbb{C}$,

$$Rx = \lambda x \Longrightarrow x = 0,$$

while $0 \in \sigma(R)$, since R is not bijective.

Theorem 5.31 (Spectrum of compact operators)**.** *Let X be Banach space and $K \in L(X)$ be compact. Then $\sigma(K)$ is countable set with* 0 *as the only possible accumulation point.*