Analytical Methods for Engineering

Paolo Guiotto

Foreword

Advanced Engineering Modeling demand knowledge and use of sophisticated tools of Analysis and Probability. The goal of this course is to introduce to the most relevant of these tools in a practical and essential way. This means that we do not privilege the theoretical depth but, rather, we aim to put the focus on the tools and the methods they are used to solve problems.

Almost all the advanced tools we introduce here are based on the modern theory of Integration proposed by Lebesgue at the beginning of XXth century. This theory is based on the fundamental concept of *measure*, and this is our starting point. We will introduce the concepts of *abstract measure* and *abstract integrals*, illustrating them on the fundamental example of Lebesgue measure and integral (used in most of Analytical applications). Abstract measure and integral are also the fundamentals for modern Probability as conceived by Kolmogorov. [...]

As said, our goal is to help students to familiarize with tools and methods. Therefore, proofs are proposed only when their technical level is not excessive and they provide some insight into what the corresponding statements say. Sometimes, proofs are proposed with extra assumptions than what actually needed just to simplify them and to get quickly to the point. Other times proofs are just "sketch" of proofs, that is not formally rigorous proofs that could be made 100% true proof with some technical work (omitted here). Yet, the goal is to help to understand "why", rather than providing a complete view of the matter. I know this approach is controversial. In my experience, it works better for students who do not have a specific interest in the matter itself and that, nonetheless, need to learn tools to understand their curricular disciplines.

A good number of solved problems is proposed throughout the notes, as well as several exercises (without solution) at the end of each Chapter. The student is encouraged to try to solve problems right after the first few examples have been shown in class. A $*$ legend to distinguish between different levels for examples and exercises:

- $(*)$ denotes the **basic level**, that is the minimal and easiest level, where the focus is mostly on the understanding of the definitions, being able to apply them on simple cases without particular technical skills required.
- $(**)$ denotes the **intermediate level**, that is the level expected for the majority of the students. Here the focus is on applying the theory to solve complex problems that, however, require the application of standard procedures. A $(**+)$ indicates the presence of technical difficulties.
- $(***)$ denotes the **advanced level**, that is a level that denotes a deep comprehension of the main ideas behind the theory, including being able to organize an abstract argument (a proof) of a general property.

To help the student with conceptual maps, a number of "checklist" is proposed. They are useful to quickly remind "what to do" to respond to a certain question. The checklists are helpful to reach the intermediate level, but they cannot replace the critical approach which is always required.

A final note. The notes contain a large number of errors of any type (mathematics, english, typos, etc). Each student can participate to make these note better for the next students will come by pointing out these. Thanks, and good luck with your job! May the wind be always at your back!

Contents

LECTURE 1

Abstract Measures

A measure is a coherent way to assign positive numbers to sets. Coherent means that an empty set must have zero measure, and the measure of a disjoint union of sets is the sum of the measures. In a way, the measure of a set is an extension of the geometrical concept of area or volume. However, its applications go much beyond Geometry. A *probability measure* is a measure with total value equal to 1. This Lecture introduces the first important definitions and properties of abstract measures, illustrating some simple examples.

1.1. $\sigma-$algebras of sets

A measure is a *set function*. Its domain is a suitable family of subsets of a set X. According to the specific context, these subsets can have a geometrical interpretation, a stochastic interpretation (as in Probability Theory, where they are called *events*) or others (as in engineering, where they represent the *information*). In every case, this family must obey to a few elementary properties:

Definition 1.1.1

A family $\mathscr{F} \subset \mathscr{P}(X)$ is called $\sigma-$**algebra** if

i) $\varnothing, X \in \mathscr{F}$;

ii) if $E \in \mathscr{F}$ then also $E^c \in \mathscr{F}$;

iii) if $E_n \in \mathscr{F}$, $n \in \mathbb{N}$, then also $\bigcup_n E_n \in \mathscr{F}$.

Elements of a $\sigma-$algebra are called **measurable sets**.

If a family of sets satisfies iii) but only for finite unions, we say that $\mathscr{F}$ is an **algebra**. The σ is to remind of countable unions. Apparently, the definition 1.1 is simple. However, it is not so easy to exhibit non-trivial examples of $\sigma-$algebras. Let us start with some easy examples.

Example 1.1.2: $(*)$

Let X be any set.

- $\mathscr{F} := \{\varnothing, X\}$ is a $\sigma-$algebra (trivial $\sigma-$algebra).
- $\mathscr{F} = \mathscr{P}(X)$ (parts of X) is a $\sigma-$algebra.
- Let $A \subsetneq X$ and $A \neq \varnothing$, then $\mathscr{F} = \{\varnothing, A, A^c, X\}$ is a $\sigma-$algebra.

Example 1.1.3: $(**)$

Let X be any set, $\mathscr{F} := \{E \subset X \ : \ \text{at least one of } E, E^c \text{ is countable}\}$. Then, $\mathscr{F}$ is a $\sigma-$algebra.

PROOF. i) Clearly $\varnothing, X \in \mathscr{F}$ (for $\varnothing$, $\varnothing$ itself is finite having zero elements, thus countable; for X, $X^c = \varnothing$ is countable). ii) Suppose $E \in \mathscr{F}$. Then E or E^c is countable, thus E^c or $(E^c)^c$ is countable, and this means that $E^c \in \mathscr{F}$. iii) Suppose $(E_n) \subset \mathscr{F}$. For each E_n, one between E_n and E_n^c is countable. Let $E := \bigcup_n E_n$. We claim that one between E and E^c is countable. We may argue as follows: if all E_n are countable, then $E = \bigcup_n E_n$ is countable. Otherwise, there exists at least one of E_n, say E_N which is not countable. But then, E_N^c must be countable. Therefore

$$E^c = \left(\bigcup_n E_n\right)^c = \bigcap_n E_n^c \subset E_N^c,$$

is countable.

Let us see some remarkable example of families that are not $\sigma-$algebras.

Example 1.1.4: $(*)$

Let $X = \mathbb{R}^d$ and $\mathscr{F} := \{E \subset \mathbb{R}^d \ : \ E \text{ open}\}$. Then $\mathscr{F}$ is not a $\sigma-$algebra.

PROOF. Indeed, while it is always true that countable (and also uncountable) unions of open sets are open sets, it is in general false that if E is open, then E^c is open as well (indeed, this happens iff $E = \varnothing, \mathbb{R}^d$).

Example 1.1.5: $(*)$

Let $X = \mathbb{R}$, $\mathscr{F} := \{I \subset \mathbb{R} \ : \ I \text{ interval}\}$. Then $\mathscr{F}$ is not a $\sigma-$algebra.

PROOF. We may say that $\varnothing =]0,0[\in \mathscr{F}$ and $X = \mathbb{R} =]-\infty, +\infty[\in \mathscr{F}$. However, the complementary I^c of an interval is not an interval (in general, it is the union of two intervals) and the union of two or more intervals is not (in general) an interval.

Example 1.1.6: extension $(***)$

Let $X = \mathbb{R}$ and $\mathscr{F} := \{\bigcup_n I_n \ : \ I_n \text{ intervals}\}$. This $\mathscr{F}$ is not a $\sigma-$algebra.

PROOF. This time, $\varnothing, X \in \mathscr{F}$ and if $(E_n) \subset \mathscr{F}$ then $\bigcup_n E_n \in \mathscr{F}$ (countable union of countable unions is still a countable union). However, if $E \in \mathscr{F}$ it is not true (in general) that $E^c \in \mathscr{F}$. To check this is a bit tricky. Let $\mathbb{Q} = \{q_n \ : \ n \in \mathbb{N}\}$, set $I_n := [q_n, q_n]$ and take $E = \mathbb{Q} = \bigcup_n I_n \in \mathscr{F}$. However, $E^c = \mathbb{R}\backslash\mathbb{Q} \notin \mathscr{F}$. Indeed, if $E^c = \mathbb{R}\backslash\mathbb{Q} = \bigcup_n J_n$ with J_n intervals, then $J_n \subset \mathbb{R}\backslash\mathbb{Q}$. Because of the density of rationals and irrationals in reals, this would force J_n to have empty interior, that is $J_n = [x_n, x_n]$ for every n. But this would imply that $\mathbb{R}\backslash\mathbb{Q} = \{x_n \ : \ n \in \mathbb{N}\}$, that is $\mathbb{R}\backslash\mathbb{Q}$ would be countable, which is impossible.

A number of simple elementary properties follow from the definition of $\sigma-$algebra. In summary, we may say that a $\sigma-$algebra is closed for set operations.

Proposition 1.1.7

Let $\mathscr{F}$ be a $\sigma-$algebra of sets on X. Then

- if $E, D \in \mathscr{F}$, then $E \backslash D$ and $E \vartriangle D := (E\backslash D) \cup (D \backslash E)$ belong to $\mathscr{F}$;
- if $E_n \in \mathscr{F}$, $n \in \mathbb{N}$, then also $\bigcap_n E_n \in \mathscr{F}$.

The proof is left as exercise. A general method to construct a $\sigma-$algebra is the following. Let $\mathscr{S} \subset \mathscr{P}(X)$ be a family of subsets of X, not necessarily a $\sigma-$algebra (for example, $\mathscr{S} := \{A \subset \mathbb{R}^d : A \text{ open}\}$). Then, we look for the smallest (the most "cheap") $\sigma-$algebra of sets containing $\mathscr{S}$. We call this $\sigma-$algebra, the *$\sigma-$algebra generated by* $\mathscr{S}$. The existence and uniqueness of such $\sigma-$algebra is ensured by the

Proposition 1.1.8

Let X be any set, $\mathscr{S} \subset \mathscr{P}(X)$ a family of subsets of X. Then,

$$\sigma(\mathscr{S}) := \bigcap_{\mathscr{F} \supset \mathscr{S}} \mathscr{F}$$

is the smallest $\sigma-$algebra containing $\mathscr{S}$.

Proof. First, the intersection is not empty: among all $\mathscr{F} \supset \mathscr{S}$ there is $\mathscr{F} = \mathscr{P}(X)$. Thus, $\sigma(\mathscr{S})$ is well posed. It is now straightforward to check that it is also a $\sigma-$algebra (exercise) and, of course, it contains $\mathscr{S}$. Finally, by definition, if $\mathscr{F} \supset \mathscr{S}$, then $\mathscr{F} \supset \sigma(\mathscr{S})$.

We already noticed that the $\mathscr{S}$ = family of open sets of $\mathbb{R}^d$ is not a $\sigma-$algebra. However, $\sigma(\mathscr{S}) =: \mathscr{B}_{\mathbb{R}^d}$ it is. This is called **Borel $\sigma-$algebra**, its elements are called **Borel sets** or **borelians**.

1.2. Definition of Measure

Definition 1.2.1

Let $\mathscr{F}$ be a $\sigma-$algebra on X. A function $\mu : \mathscr{F} \longrightarrow [0, +\infty]$ is called **measure on** $\mathscr{F}$ if

i) $\mu(\varnothing) = 0$;
ii) if $E = \bigcup_n E_n$ with $E_n \in \mathscr{F}$, $n \in \mathbb{N}$ and $E_n \cap E_m = \varnothing$ for $n \neq m$ (disjoint union of measurable sets), then
$$\mu(E) = \sum_n \mu(E_n).$$

The triplet $(X, \mathscr{F}, \mu)$ is called **measure space**.

Property ii) is called **countable additivity**. We introduce a convenient notation:

$$\bigsqcup_n E_n := \bigcup_n E_n, \text{ if } E_n \cap E_m = \emptyset,\ n \neq m, \implies \mu\left(\bigsqcup_n E_n\right) = \sum_n \mu(E_n).$$

Here are some elementary and introductory examples.

Example 1.2.2: (∗) Dirac measure

Let X be any set, $\mathscr{F}$ a $\sigma-$algebra. Let $x \in X$ and define

$$\delta_x(E) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Then, δ_x is a measure on $\mathscr{F}$. The proof is left as an exercise. □

Example 1.2.3: (∗) counting measure

Let X be a countable set, for example $X = \mathbb{N}$ (but also $X = \mathbb{Q}$ or $\mathbb{Q}^N$). In other words, $X = \{x_n \ : \ n \in \mathbb{N}\}$. Let $\mathscr{F} = \mathscr{P}(X)$ and set

$$\mu(E) := \sum_{n \,:\, x_n \in E} 1 \equiv \sum_n \delta_{x_n}(E).$$

Then μ is a measure on X (interpretation: $\mu(E)$ counts the number of elements in the set E). Check left as exercise.

A fundamental example is the **Lebesgue measure on** $\mathbb{R}^d$. This is the topic of the next Lecture.

1.3. Basic properties

In this section we illustrate some of the most basic and commonly used properties of any generic measure. Proofs are generally easy and following natural ideas.

Proposition 1.3.1

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then,

i) (monotonicity) if $E, F \in \mathscr{F}$, $E \subset F$, then $\mu(E) \leqslant \mu(F)$.

ii) (subtractivity) if $E, F \in \mathscr{F}$, $E \subset F$ and $\mu(E) < +\infty$, then $\mu(F \backslash E) = \mu(F) - \mu(E)$ (with agreement that $+\infty - m = +\infty$ for every $m \in [0 + \infty[$).

iii) (finite additivity formula) if $E, F \in \mathscr{F}$ are non necessarily disjoint, then $\mu(E \cup F) = \mu(E) + \mu(F) - \mu(E \cap F)$ provided $\mu(E \cap F) < +\infty$.

$$E \subset F,\ E, F \in \mathscr{F}, \implies \mu(E) \leqslant \mu(F).$$

PROOF. i) Just notice that $F = E \sqcup (F \backslash E)$, and since $F \backslash E \in \mathscr{F}$ we have

$$\mu(F) = \mu(E) + \underbrace{\mu(F \backslash E)}_{\geqslant 0} \geqslant \mu(E).$$

ii) By previous relation, since $\mu(E) < +\infty$, we have the conclusion.
iii) Noticed that $E \cup F = E \backslash (E \cap F) \sqcup E \cap F \sqcup F \backslash (E \cap F)$, we have

$$\begin{aligned} \mu(E \cup F) &= \mu(E \backslash (E \cap F)) + \mu(E \cap F) + \mu(F \backslash (E \cap F)) \\ &\overset{ii)}{=} \mu(E) - \mu(E \cap F) + \mu(E \cap F) + \mu(F) - \mu(E \cap F), \end{aligned}$$

which is the conclusion.

Let $(E_n) \subset F$ be a sequence of measurable sets. We say that

$$E_n \nearrow, \iff E_0 \subset E_1 \subset \ldots \subset E_n \subset E_{n+1} \subset \ldots.$$

By monotonicity,

$$\mu(E_0) \leqslant \mu(E_1) \leqslant \ldots \leqslant \mu(E_n) \leqslant \mu(E_{n+1}) \leqslant \ldots,$$

that is, $(\mu(E_n)) \subset [0, +\infty]$ is an increasing sequence of numbers (accepting $+\infty$ as number). Therefore,

$$\exists \lim_n \mu(E_n).$$

The question is: *is the limit of the measures the measure of some limit set?* The answer is provided by the following.

Theorem 1.3.2: continuity form below

Let $(X, \mathscr{F}, \mu)$ measure space, $(E_n) \subset \mathscr{F})$, $E_n \nearrow$. Then

$$\exists \lim_n \mu(E_n) = \mu\left(\bigcup_n E_n\right).$$

PROOF. The existence of the limit has been already discussed in the premises. We show the identity. Let's start from $\mu\left(\bigcup_n E_n\right)$. Since sets are nested, the union is definitely not disjoint. However, we can transform it into a disjoint union:

$$E_0 \cup E_1 \cup E_2 \cup \cdots \cup E_n \cup \cdots = E_0 \sqcup (E_1 \backslash E_0) \sqcup (E_2 \backslash E_1) \sqcup \cdots \sqcup (E_n \backslash E_{n-1}) \sqcup \cdots$$

Therefore, setting $E_{-1} := \emptyset$, we have

$$\begin{aligned} \mu\left(\bigcup_n E_n\right) &= \mu\left(\bigsqcup_n (E_n \backslash E_{n-1})\right) = \sum_n \mu(E_n \backslash E_{n-1}) = \lim_{m \to +\infty} \sum_{n=0}^m \mu(E_n \backslash E_{n-1}) \\ &= \lim_{m \to +\infty} \mu\left(\bigsqcup_{n=0}^m E_n \backslash E_{n-1}\right) = \lim_{m \to +\infty} \mu(E_m). \quad \square \end{aligned}$$

Similarly, we say that

$$E_n \searrow, \iff E_0 \supset E_1 \supset \ldots \supset E_n \supset E_{n+1} \supset \ldots.$$

By monotonicity,

$$\mu(E_0) \geqslant \mu(E_1) \geqslant \ldots \geqslant \mu(E_n) \geqslant \mu(E_{n+1}) \geqslant \ldots,$$

that is $(\mu(E_n)) \subset [0, +\infty]$ is a decreasing sequence of numbers (accepting $+\infty$ as number). Therefore,

$$\exists \lim_n \mu(E_n).$$

As above, the point is: is this limit measure the measure of a limit set? The reasonable guess for this set is

$$\bigcap_n E_n.$$

Surprisingly, the answer is negative in general, as the following example shows.

Warning 1.3.3

Let $X = \mathbb{N}$, $\mathscr{F} = \mathscr{P}(\mathbb{N})$ and μ the counting measure. Let

$$E_n := \{n, n+1, n+2, \ldots\}.$$

It is clear that:

- $E_n \searrow$;
- $\mu(E_n) = +\infty$.

However,

$$E := \bigcap_n E_n = \varnothing, \quad \mu(E) = 0 \neq +\infty = \lim_n \mu(E_n). \quad \square$$

Nonetheless, the continuity from above becomes true as soon as we add a little (but fundamental) requirement:

Corollary 1.3.4: continuity from above

Let $(X, \mathscr{F}, \mu)$ measure space, $(E_n) \subset \mathscr{F})$, $E_n \searrow$. Suppose moreover that $\mu(E_1) < +\infty$. Then

$$\exists \lim_n \mu(E_n) = \mu\left(\bigcap_n E_n\right).$$

In particular: if μ is a **finite measure**, that is $\mu(X) < +\infty$, then continuity from above always holds true.

Proof. Call $E := \bigcap_n E_n$ and set

$$F_n := E_0 \backslash E_n, \; n \geqslant 1.$$

Since $E_n \searrow$, $F_n \nearrow$. Thus, by continuity from below,

$$\lim_n \mu(F_n) = \mu\left(\bigcup_n F_n\right) = \mu\left(E_0 \backslash \bigcap_n E_n\right) = \mu(E_0) - \mu(E),$$

by subtractivity. By the same property, $\mu(F_n) = \mu(E_0 \backslash E_n) = \mu(E_0) - \mu(E_n)$ thus, being everything finite, we have

$$\mu(E_0) - \lim_n \mu(E_n) = \mu(E_0) - \mu(E),$$

which is the conclusion.

1.4. Exercises

Exercise 1.4.1 ($*$). *Say whether the following are $\sigma-$algebras or not:*

i) $X = \mathbb{R}$ *with* $\mathscr{F} := \{I \subset \mathbb{R} \ : \ I \text{ interval}\}$.
ii) X *any countable set,* $\mathscr{F} := \{E \subset X \ : \ E \text{ is finite set}\}$.
iii) X *any infinite set,* $\mathscr{F} := \{E \subset X \ : \ E \text{ is countable set}\}$.

Exercise 1.4.2 ($*$). *Let* $X = \{a, b, c, d\}$, $\mathscr{S} = \{\{a\}, \{a, c\}\}$. *Determine* $\sigma(\mathcal{S})$.

Exercise 1.4.3 ($*$). *Let X be a non empty set,* $A, B \subset X$, $A \neq B$. *Determine* $\sigma(\{A, B\})$.

Exercise 1.4.4 ($*$). *Let X be an uncountable set,* $\mathscr{F} := \{A \subset X \ : \ \text{one of } A, A^c \text{ is countable}\}$. *Define,*

$$\mu(A) := \begin{cases} 0, & \text{if } A \text{ is countable}, \\ 1, & \text{if } A^c \text{ is countable}. \end{cases}$$

Determine if μ is a measure on $(X, \mathscr{F})$.

Exercise 1.4.5 ($**$). *Let* $\mathscr{S}_1, \mathscr{S}_2$ *any two families of subsets of X, that is* $\mathscr{S}_1, \mathscr{S}_2 \subset \mathscr{P}(X)$. *Prove that*

$$\sigma(\mathscr{S}_1 \cup \mathscr{S}_2) = \sigma\left(\sigma(\mathscr{S}_1) \cup \sigma(\mathscr{S}_2)\right).$$

Exercise 1.4.6. *Let* $\mathscr{F}_1, \mathscr{F}_2 \subset \mathscr{P}(X)$ *be two $\sigma-$algebras of sets.*

i) ($*$) *Prove that also* $\mathscr{F}_1 \cap \mathscr{F}_2$ *is a $\sigma-$algebra.*
ii) ($**$) *Is it true that also* $\mathscr{F}_1 \cup \mathscr{F}_2$ *is a $\sigma-$algebra? Provide a proof (if true) or a counterexample (if false).*

Exercise 1.4.7 ($**$). *Let* $(X, \mathscr{F}, \mu)$ *be a measure space. State and prove a formula for* $\mu(E \cup F \cup G)$ *where* $E, F, G \in \mathscr{F}$.

Exercise 1.4.8 ($***$). *Let* $X = [0, 1]$, $E \subset X$ *and define*

$$\mu(E) := \lim_{n \to +\infty} \frac{1}{n} \sharp \left\{ \frac{k}{n} \in A \ : \ k \in \mathbb{N} \right\},$$

provided the limit exists.

i) *Show that if* $E = [a, b] \subset [0, 1]$ *then* $\mu(E) = b - a$.
ii) *Show that μ is additive, that is if* $A \cap B = \emptyset$, *then* $\mu(A \cup B) = \mu(A) \cup \mu(B)$.
iii) *Take as A the set of dyadic numbers* $A := \{\frac{k}{2^m} \ : \ m \in \mathbb{N}, \ k = 0, 1, \dots, 2^m\}$. *What can be said about $\mu(A)$? Use the answer to respond to the question: is μ a measure?*

Exercise 1.4.9 ($***$). *Let* $(X, \mathscr{F}, \mu)$ *be a measure space,* $(E_n) \subset \mathscr{F}$. *Suppose that*

$$\sum_n \mu(E_n) < +\infty.$$

Prove that the set

$$S := \left\{ x \in X \ : \ x \in E_j, \text{ for infinitely many } j \right\}.$$

i) *By suitable set operations, express S in terms of sets* E_n*, deducing that* $S \in \mathscr{F}$.

ii) *Deduce the measure of S.*

LECTURE 2

Lebesgue Measure

Lebesgue measure is a fundamental tool of Mathematical Analysis. It answers to the aim of having a geometrical measure on sets of $\mathbb{R}^N$ fulfilling few basic properties:

- measure of a rectangle is its area;
- measure is invariant by rotations and translations;
- natural sets as open and closed sets are measurable.

In this Lecture we sketch the construction of such a measure. Most of the proofs are technical and too long, much beyond our scope, thus are omitted.

2.1. Outer Measure

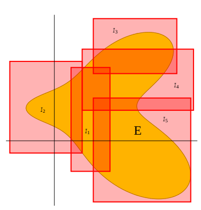
Definition 2.1.1

A set of type $I = [a_1, b_1] \times \cdots \times [a_d, b_d] \subset \mathbb{R}^d$ is called **(multi) interval**. Its **measure** is, by definition

$$|I|_d := (b_1 - a_1) \cdots (b_d - a_d).$$

If not expressedly needed, we will write just $|I|$ for $|I|_d$.

As we know, the family of intervals is not a $\sigma-$algebra. However, by exhaustion methods, we can use intervals to fill any set $E \subset \mathbb{R}^d$. We say that a family of intervals (I_n) is a **covering of** E if $E \subset \bigcup_n I_n$.



For each covering, the (possibly infinite) sum $\sum_n |I_n|$ represents an approximation by excess of the measure of E. Since there are infinitely many coverings of a set E, we have (in general) infinitely many approximations by excess of the measure of E. The best of these approximations is what we call **outer measure of** E.

Definition 2.1.2: outer measure

For $E \subset \mathbb{R}^d$ we set

$$\lambda_d^*(E) := \inf \left\{ \sum_j |I_j|_d \ : \ E \subset \bigcup_j I_j, \ I_j \text{ intervals} \subset \mathbb{R}^d \right\}.$$

By definition, $\lambda_d^*(\varnothing) = 0$. If not explicitly needed, we will write just $\lambda^*(E)$ for $\lambda_d^*(E)$.

Notice that the outer measure is defined for every set E. Here some properties of λ^*.

Proposition 2.1.3

The following properties hold true:

i) (coherence) $\lambda^*(I) = |I|$ for every I interval;
ii) (translation invariance) $\lambda^*(E + x) = \lambda^*(E)$, $\forall E \subset \mathbb{R}^d$, $\forall x \in \mathbb{R}^d$;
iii) (homogeneity) $\lambda^*(cE) = |c|^d \lambda^*(E)$;
iv) (monotonicity) $\lambda^*(E) \leqslant \lambda^*(F)$ if $E \subset F$;
v) (sub-additivity):

$$\lambda^* \left(\bigcup_n E_n \right) \leqslant \sum_n \lambda^*(E_n).$$

Proof. i) Among all coverings of I there is also that one made of I only, thus $\lambda^*(I) \leqslant |I|$ just by definition. In particular, this shows that $\lambda^*(I) < +\infty$. For the vice versa, let $\bigcup_j I_j$ be a covering of I made of rectangles such that

$$\sum_j |I_j| \leqslant \lambda^*(I) + \varepsilon.$$

Since $I \cap I_j$ is a rectangle (easy), and $I = I \cap \bigcup_j I_j = \bigcup_j I \cap I_j$, thus

$$|I| \leqslant \sum_j |I \cap I_j| \leqslant \sum_j |I_j| \leqslant \lambda^*(I) + \varepsilon.$$

Since ε can be take arbitrarily small, we conclude that $|I| \leqslant \lambda^*(I)$.
ii) Easy, just notice that every covering $\bigcup_j I_j$ of E corresponds to a covering $\bigcup_j (I_j + x)$ of $E + x$ and vice versa. And since it is easy to check that $|I + x| = |I|$, the conclusion easily follows (fill the details).
iii), iv) Exercise.
v) This is less easy. To begin, we notice that the conclusion is true if some of $\lambda^*(E_n) = +\infty$. Thus we may assume $\lambda^*(E_n) < +\infty$ for every n. Then, for every E_n, there is a covering $\bigcup_k I_{n,j}$ such that

$$\sum_j |I_{n,j}| \leqslant \lambda^*(E_n) + \frac{\varepsilon}{2^n}.$$

Then

$$\bigcup_n E_n \subset \bigcup_n \bigcup_j I_{n,j} = \bigcup_{n,j} I_{n,j},$$

thus this is a covering for $\bigcup_n E_n$ made of rectangles. Consequently,

$$\lambda^*\left(\bigcup_n E_n\right) \leqslant \sum_{n,j} |I_{n,j}| = \sum_n \sum_j |I_{n,j}| \leqslant \sum_n \left(\lambda^*(E_n) + \frac{\varepsilon}{2^n}\right) = \sum_n \lambda^*(E_n) + \varepsilon,$$

and since $\varepsilon > 0$ is arbitrary, the conclusion follows.

Sub-additivity is weaker than **countable additivity**. Unfortunately, this last is false in general. This is the consequence of the following difficult result:

Proposition 2.1.4: Vitali

λ^* is not countably additive.

Proof. (sketch for $d = 1$) The proof is based on showing that there exists a set $E \subset [-1, 1]$ such that

$$[-1, 1] \subset \bigsqcup_n (E + q_n) \subset [-2, 2],$$

where $(q_n) = \mathbb{Q} \cap [-1, 1]$. The existence of such a set is difficult and based on subtle logical arguments, so we omit here. Accepting this, by monotonicity we would have

$$2 = \lambda^*([-1, 1]) \leqslant \lambda^*\left(\bigsqcup_n E + x_n\right) \leqslant \lambda^*([-2, 2]) = 4.$$

If λ^* were countably additive, we would also have

$$\lambda^*\left(\bigsqcup_n E + q_n\right) = \sum_n \lambda^*(E + q_n) = \sum_n \lambda^*(E),$$

because the translation invariance. But then

$$2 \leqslant \sum_n \lambda^*(E) \leqslant 4,$$

and this would be impossible. Indeed, either $\lambda^*(E) > 0$ or $\lambda^*(E) = 0$. In the first case $\sum_n \lambda^*(E) = +\infty$, while in the second $\sum_n \lambda^*(E) = 0$. In both cases, the previous bound would be impossible. □

2.2. Lebesgue class and measure

The problem with outer measure is that it assigns a measure to each set of $\mathbb{R}^d$. Among these, there are "bad sets" as the Vitali's set, that makes countable additivity fail. To solve this issue, we restrict the class of sets to which we assign a measure. This class should be large enough to contain "natural" sets, such as open and closed sets of $\mathbb{R}^d$. The idea is to consider only sets which are "well approximated" (in the sense of measure) by an open set.

Definition 2.2.1: Lebesgue class

$$\mathscr{M}_d := \left\{E \subset \mathbb{R}^d \;:\; \forall \varepsilon > 0,\ \exists O_\varepsilon \supset E,\ O_\varepsilon \text{ open}, \lambda^*(O_\varepsilon \backslash E) \leqslant \varepsilon\right\}.$$

Here, some easy consequences of this definition:

Proposition 2.2.2

i) Open sets are Lebesgue measurable. In particular, $\emptyset, \mathbb{R}^d \in \mathscr{M}_d$.
ii) Intervals are Lebesgue measurable.
iii) Measure 0 sets are Lebesgue measurable.
iv) Any sets that differs by a measurable set by a measure 0 set is measurable.

Proof. i) If E is open, take $O_\varepsilon := E$, then $O_\varepsilon \backslash E = \emptyset$ hence $\lambda_d^*(O_\varepsilon \backslash E) = 0 \leqslant \varepsilon$, for every $\varepsilon > 0$.
ii) For simplicity, we show this in the case $d = 1$. Let $I = [a, b]$. For $\varepsilon > 0$, set $O_\varepsilon =]a - \varepsilon, b + \varepsilon[$. Therefore $O_\varepsilon \backslash I =]a - \varepsilon, a[\cup]b, b + \varepsilon[$, hence, by sub-additivity,

$$\lambda^*(O_\varepsilon \backslash I) \leqslant \lambda^*(]a - \varepsilon, a[) + \lambda^*(b, b + \varepsilon[) \leqslant 2\varepsilon.$$

iii) Suppose $\lambda^*(E) = 0$. There exists then a covering $\bigcup_j I_j$ such that

$$\sum_j |I_j| \leqslant \lambda^*(E) + \varepsilon = \varepsilon.$$

We may enlarge each I_j to become an open rectangle I_j^ε such that $|I_j^\varepsilon| \leqslant |I_j| + \frac{\varepsilon}{2^j}$. Then, setting

$$O_\varepsilon := \bigcup_j I_j^\varepsilon,$$

this is open (union of open sets), it contains E (because is contains one of its covering) and

$$\lambda^*(O_\varepsilon \backslash E) \leqslant \lambda^*(O_\varepsilon) \leqslant \sum_j |I_j^\varepsilon| \leqslant \sum_j \left(|I_j| + \frac{\varepsilon}{2^j}\right) = \sum_j |I_j| + \varepsilon = 2\varepsilon. \quad \square$$

iv) Indeed let $F = E \cup N$ with $\lambda^*(N) = 0$ and $E \in \mathscr{M}_d$; according to the definition, for every $\varepsilon > 0$ there exists an open set O_ε such that $\lambda_d(O_\varepsilon \backslash E) \leqslant \varepsilon$. Since also N is measurable (being a null set), there exists another open $\widetilde{O}_\varepsilon \supset N$ such that $\lambda^*(\widetilde{O}_\varepsilon \backslash N) \leqslant \varepsilon$. Then, $O_\varepsilon \cup \widetilde{O}_\varepsilon$ is open, contains $E \cup N$ and

$$\begin{aligned}
\lambda^*\left((O_\varepsilon \cup \widetilde{O}_\varepsilon) \backslash (E \cup N)\right) &= \lambda^*\left(O_\varepsilon \backslash (E \cup N) \cup \widetilde{O}_\varepsilon \backslash (E \cup N)\right) \\
&\leqslant \lambda^*(O_\varepsilon \backslash (E \cup N)) + \lambda^*(\widetilde{O}_\varepsilon \backslash (E \cup N)) \\
&\leqslant \lambda^*(O_\varepsilon \backslash E) + \lambda^*(\widetilde{O}_\varepsilon \backslash N) \leqslant 2\varepsilon. \quad \square
\end{aligned}$$

With some technical work it is possible to prove the

Theorem 2.2.3

The family $\mathscr{M}_d$ is a $\sigma-$algebra of sets called **Lebesgue class**. Sets of $\mathscr{M}_d$ are called **Lebesgue measurable sets**. The outer measure λ_d^* is a measure on $\mathscr{M}_d$, called **Lebesgue measure**. We denote the Lebesgue measure by λ_d.

Of course, since the Lebesgue measure is just the outer measure on a sub-family of sets, it inherit its properties. In particular, it is invariant by translations and it fulfills homogeneity. Actually, these properties are particular cases of the following one:

Proposition 2.2.4

Let T be a $d \times d$ invertible matrix. Then, if $E \in \mathscr{M}_d$, also $TE := \{Tx \ : \ x \in E\} \in \mathscr{M}_d$ and,

$$\lambda_d(TE + v) = |\det T| \lambda_d(E), \ \forall v \in \mathbb{R}^d. \tag{2.2.1}$$

2.3. Exercises

Exercise 2.3.1 ($*$)**.** *Let $N \subset \mathbb{R}$ be a null set. Show that, necessarily, N^c is dense in $\mathbb{R}$, that is the following property holds:*

$$\forall]a, b[\subset \mathbb{R}, \ N^c \cap]a, b[\neq \emptyset.$$

Exercise 2.3.2 ($**$ Cantor set)**.** *Define*

$$
\begin{aligned}
&C_0 := [0, 1],\\
&C_1 := [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, 1] = C_0 \backslash]\tfrac{1}{3}, \tfrac{2}{3}[,\\
&C_2 := [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{3}{9}] \cup [\tfrac{6}{9}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1] = C_1 \backslash \left( ]\tfrac{1}{9}, \tfrac{2}{9}[\cup]\tfrac{7}{9}, \tfrac{8}{9}[\right)\\
&\vdots\\
&C_n := [0, \tfrac{1}{3^n}] \cup [\tfrac{2}{3^n}, \tfrac{3}{3^n}] \cup \ldots \cup [\tfrac{3^n-1}{3^n}, 1] = C_{n-1} \backslash \bigcup_{k=0}^{2^{n-1}-1} \left] \frac{3k+1}{3^n}, \frac{3k+2}{3^n} \right[, \quad n \in \mathbb{N},
\end{aligned}
$$

and $C := \bigcap_{n \geqslant 0} C_n$. Then $C \in \mathscr{M}_1$ and $\lambda(C) = 0$.

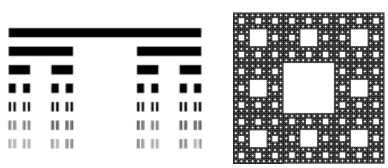


FIGURE 1. Cantor set (left), Sierpinski carpet (right).

Exercise 2.3.3 ($**$ Sierpinski carpet)**.** *The Sierpinki carpet is a bidimensional set of Cantor type. Let $T_0 := [0, 1] \times [0, 1]$ and define recursively T_n according to the following rule*

$$T_n := T_{n-1} \backslash \bigcup_{i,j=0}^{2^{n-1}-1} \left] \frac{3i+1}{3^n}, \frac{3i+2}{3^n} \right[\times \left] \frac{3j+1}{3^n}, \frac{3j+2}{3^n} \right[, \quad n \in \mathbb{N}.$$

Define finally $T := \bigcap_{n \geqslant 0} T_n$. Show that T is measurable and determine its measure.

Exercise 2.3.4 ($**$)**.** *Construct a Cantor like set by removing, at each step, the middle quarters. Show that the limit set is Lebesgue measurable and compute its measure.*

Exercise 2.3.5 ($**$)**.** *Let $E \subset \mathbb{R}^2$ be the set of points $(x, y) \in \mathbb{R}^2$ such that x and y are rationally dependent, that is*

$$E = \left\{(x, y) \in \mathbb{R}^2 \ : \ \exists (m, n) \in \mathbb{N} \times \mathbb{N}, \ (m, n) \neq (0, 0) \ : \ mx + ny = 0\right\{ .$$

Prove that $\lambda_2(E) = 0$.

Exercise 2.3.6 ($**$)**.** *Let $A \subset E \subset B$, $A, B \in \mathscr{M}_d$ with $\lambda(A) = \lambda(B) < +\infty$. Deduce that $E \in \mathscr{M}_d$. What about $\lambda_d(E)$?*

Exercise 2.3.7 ($***$)**.** *Let $A, B, C \subset [0, 1]$ be such that the following property holds:*

$$\forall x \in [0, 1] \ \text{ belongs to at least 2 sets among } A, B, C.$$

Prove that at least one of these sets has measure $\geqslant \frac{2}{3}$. (hint: $1 = \lambda((A \cap B) \cup (A \cap C) \cup (B \cap C))$ and argue by contradiction. . .)

Exercise 2.3.8 ($***$)**.** *Let $N \subset [0, 1]$ be such that $\lambda(N) = 0$. Prove that $\lambda(N^2) = 0$, where $N^2 = \{x^2 \ : \ x \in N\}$. What if $N \subset [-R, R]$ (that is N bounded)? What if $N \subset \mathbb{R}$ is generic?*

Exercise 2.3.9 ($***$)**.** *Let $S \subset [0, 1]$ the set of numbers which do not have the digit 5 in their decimal representation. Is S measurable? If yes, what is its measure?*

LECTURE 3

Measurable Functions

As we have now a definition of *measurable set*, we introduce a definition of *measurable function*. The idea is simple: we wish that natural sets as level sets of f, for instance $\{x \,:\, f(x) \geqslant a\}$ are measurable sets. Measurable functions are fundamental to define integrals (next Lecture). In Probability, measurable functions are called *random variables*.

3.1. Definiton and first properties

Definition 3.1.1

Let $\mathscr{F}$ be a $\sigma-$algebra on X. We say that $f : E \subset X \longrightarrow \mathbb{R}$ is $\mathscr{F}-$ **measurable** (notation $f \in L(E, \mathscr{F})$) if

$$\{f \in I\} \equiv \{x \in E \,:\, f(x) \in I\} \in \mathscr{F},\ \forall I \subset \mathbb{R},\ I \text{ interval.}$$

If the $\sigma-$algebra is understood we just write $f \in L(E)$.

Remark 3.1.2

If $f \in L(E)$ then, necessarily, $E \in \mathscr{F}$. Indeed: $E = \{f \in \mathbb{R}\} \in \mathscr{F}$.

Example 3.1.3: $(*)$

Every constant function is measurable.

Proof. Indeed, if $f \equiv c$, then

$$\{f \in I\} = \begin{cases} X, & \text{if } c \in I, \\ \varnothing, & \text{if } c \notin I. \end{cases} \qquad \square$$

Example 3.1.4: $(*)$ **indicator function**

Let

$$1_E := \begin{cases} 1, & x \in E, \\ 0, & x \notin E. \end{cases}$$

Then, $1_E \in L(X)$, iff $E \in \mathscr{F}$.

PROOF. Notice that $1_E \in \{0,1\}$. Therefore

$$\{1_E \in I\} = \begin{cases} X, & \text{if } 0,1 \in I, \\ E, & \text{if } 1 \in I,\ 0 \notin I, \\ E^c, & \text{if } 0 \in I,\ 1 \notin I, \\ \emptyset, & \text{if } 0,1 \notin I. \end{cases}$$

Thus, $\{1_E \in I\} \in \mathscr{F}$ iff $E, E^c \in \mathscr{F}$.

Definition 3.1.5

A **simple function** is a function assuming only a finite number of values. We may represent such a function as

$$s = \sum_{k=1}^{N} c_k 1_{E_k}, \quad \text{where } E_k = \{s = c_k\},\ \bigcup_k E_k = X.$$

It is easy to check that *a simple function* $s = \sum_n c_k 1_{E_k}$ *is measurable iff* $E_k \in \mathscr{F}$ *for every* $k = 1, \dots, N$ (exercise). It is sometimes useful to check measurability of a function through simplified conditions.

Proposition 3.1.6

The following properties are equivalent:

i) f is $\mathscr{F}-$measurable.
ii) $\{f \geqslant a\} \in \mathscr{F}, \forall a \in \mathbb{R}$.
iii) $\{f > a\} \in \mathscr{F}, \forall a \in \mathbb{R}$.

PROOF. By definition, i) $\Longrightarrow$ ii),iii). Let us check that ii) and iii) are equivalent, then that they imply i). Assume ii) for instance. Notice that

$$\{f > a\} = \bigcup_{n \geqslant 1} \left\{ f \geqslant a + \frac{1}{n} \right\}.$$

Indeed, clearly $\{f \geqslant a + \frac{1}{n}\} \subset \{f > a\}$ so the union is contained into $\{f > a\}$. Conversely, if $f(x) > a$, choosing n in such a way that $f(x) \geqslant a + \frac{1}{n}$ (we can do this because $a + \frac{1}{n} \longrightarrow a$), we have $x \in \{f \geqslant a + \frac{1}{n}\}$, so x belongs to the union. Now, since ii) holds, $\{f \geqslant a + \frac{1}{n}\} \in \mathscr{F}$ for every n, and since $\mathscr{F}$ is a $\sigma-$algebra, also the union belongs to $\mathscr{F}$, so $\{f > a\} \in \mathscr{F}$. This proves that ii) $\Longrightarrow$ iii). With a similar argument we prove that iii) $\Longrightarrow$ ii).
To finish, let us prove that ii) $\Longrightarrow$ i). Let I be an interval. If $I = [a, +\infty[$, then by ii) we have $\{f \in I\} + \{f \geqslant a\} \in \mathscr{F}$. If $I = [a,b]$ we can write

$$\{f \in I\} = \{a \leqslant f \leqslant b\} = \{f \geqslant a\} \backslash \{f > b\}.$$

Since $\{f \geqslant a\} \in \mathscr{F}$ (by ii)), $\{f > b\} \in \mathscr{F}$ (by iii), which is equivalent to ii)), and $\mathscr{F}$ is a $\sigma-$algebra, we deduce that also their difference belongs to $\mathscr{F}$, so $\{f \in I\} \in \mathscr{F}$. With similar arguments we discuss all possible types of intervals I.

The class of measurable functions is closed respect to the main algebraic operations:

Proposition 3.1.7

Let f and g be $\mathscr{F}-$measurable functions. Then

i) every linear combination $\alpha f + \beta g$ is $\mathscr{F}-$measurable.
ii) algebraic product $f \cdot g$ is $\mathscr{F}-$measurable.
iii) ratio f/g is $\mathscr{F}-$measurable if $g \neq 0$.

Proof. i) We will limit to prove that $f+g$ is $\mathscr{F}-$measurable. The remainder of the proof is left as exercise (see Exercise 3.4.4). We prove that $\{f+g>a\} \in \mathscr{F}$ for every $a \in \mathbb{R}$. We notice that

$$\{f+g><a\} = \{x \in E \ : \ f(x)+g(x)>a\} = \{x \in \mathbb{E} \ : \ f(x) > a - g(x)\}.$$

So, if $x \in \{f+g>a\}$, $f(x) > a - g(x)$, therefore there exists $q \in \mathbb{Q}$ such that

$$f(x) > q > a - g(x), \implies x \in \bigcup_{q \in \mathbb{Q}} \{f>q\} \cap \{g > a-q\}.$$

We proved that

$$\{f+g>a\} \subset \bigcup_{q \in \mathbb{Q}} \{f>q\} \cap \{g>a-q\}.$$

Vice versa: if x belongs to the union, then there exists $q \in \mathbb{Q}$ such that $f(x) > q > a - g(x)$, from which $f(x)+g(x) > a$. Therefore,

$$\{f+g>a\} = \bigcup_{q \in \mathbb{Q}} \{f>q\} \cap \{g>a-q\}.$$

The r.h.s. is a countable union fo intersections of measurable sets (because f, g are both measurable), so $\{f+g>a\} \in \mathscr{F}$, this for every $a \in \mathbb{R}$,from which the conclusion follows.

Another important operation is

Proposition 3.1.8

If f is $\mathscr{F}-$measurable and $\varphi \in \mathscr{C}(\mathbb{R})$, then $\varphi(f)$ is $\mathscr{F}-$measurable.

For the proof see the exercise 3.4.5. So, for instance, if f is $\mathscr{F}-$measurable, then also $|f|, f^2, f^3, e^f, \sin f, \dots$ are $\mathscr{F}-$measurable. But warning! $\varphi(f)$ might be measurable even if f it is not.

Warning 3.1.9

Let E be a **non measurable** set (that is a set $E \subset X$ but $E \notin \mathscr{F}$ [a]). Set

$$f := 1_E - 1_{E^c}.$$

Then, f is a simple function, but since $E, E^c \notin \mathscr{F}$, f cannot be measurable ($\{f \geqslant 0\} = E \notin \mathscr{F}$). Take now $\varphi(y) = |y|$. Then $\varphi \in \mathscr{C}(\mathbb{R})$ and $\varphi(f) = |f| \equiv 1 \in L(X)$. Thus, $f \notin L(X)$ but $\varphi(f) \in L(X)$. $\square$

[a] This depends on the $\sigma-$algebra $\mathscr{F}$. If for instance $\mathscr{F} = \mathscr{P}(X)$, then such a set cannot exist. If $\mathscr{F} = \mathscr{M}_d$ is the Lebesgue class, we know that there are sets $E \subset \mathbb{R}^d$, $E \notin \mathscr{F}$ (for example, Vitali's set).

If $X = \mathbb{R}^d$ and $\mathscr{F} = \mathscr{M}_d$ is the Lebesgue class, $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ is a numerical function of real variables. We have that:

Proposition 3.1.10

Every continuous function $f \in \mathscr{C}(\mathbb{R}^d)$ is Lebesgue measurable.

PROOF. The proof is based on a remarkable property of continuous functions: *for every open set* $A \subset \mathbb{R}$, $\{f \in A\} \subset \mathbb{R}^d$ *is open*. So,

$$\{f > a\} \text{ is open} \implies \{f > a\} \in \mathscr{F},\ \forall a \in \mathbb{R}.$$

3.2. Null sets, almost everywhere

Measure 0 sets, also called **null sets**, play an important role. Given a certain property $p(x)$ with $x \in X$, $(X, \mathscr{F}, \mu)$ measure space, we say that $p(x)$ *holds for* **almost every** $x \in E$ if

$$\boxed{\exists N \in \mathscr{F},\ \mu(N) = 0,\ :\ p(x)\ \text{ true } \forall x \in E \backslash N.}$$

So for example:

- a function f is such that $f = 0$ a.e. on E, if $f(x) = 0$, $\forall x \in E \backslash N$ with $\mu(N) = 0$.
- given f, g, we say $f = g$ a.e. on E if $f(x) = g(x)$, $\forall x \in E \backslash N$, $\mu(N) = 0$.

Example 3.2.1: $(*)$

Let $(X, \mathscr{F}, \mu) = (\mathbb{R}, \mathscr{M}_1, \lambda_1)$ and $p(x) = x$ is irrational. We have that $p(x)$ is true iff $x \in \mathbb{R} \backslash \mathbb{Q}$, and since $\lambda_1(\mathbb{Q}) = 0$, we conclude that $p(x)$ holds a.e. $x \in \mathbb{R}$.

We now introduce an important

Definition 3.2.2

We say that a measure space $(X, \mathscr{F}, \mu)$ is **complete** if, for every null set $N \in \mathscr{F}$ (that is $\mu(N) = 0$), we have

$$\forall E \subset N, \implies E \in \mathscr{F}.$$

Proposition 3.2.3

$(\mathbb{R}^d, \mathscr{M}_d, \lambda_d)$ is complete.

PROOF. Let N be a null set, $\lambda_d(N) = 0$. If $E \subset N$ then

$$\lambda_d^*(E) \leqslant \lambda_d^*(N) \equiv \lambda_d(N) = 0,$$

and since all (outer) measure 0 sets are Lebesgue measurable we have the conclusion.

Example 3.2.4: $(*)$

Build an example of non complete space.

PROOF. Let $X = \{a, b, c\}$, $\mathscr{F} = \{\varnothing, X, \{a\}, \{b, c\}\}$. It is easy to check that $\mathscr{F}$ is a $\sigma-$algebra. Define $\mu(\{a\}) = \mu(X) = 1$, $\mu(\{b, c\}) = \mu(\varnothing) = 0$. It is easy to check that μ is a measure on $(X, \mathscr{F})$. Now, $\{b\}, \{c\} \subset \{b, c\}$ but in this example they are not in $\mathscr{F}$.

Any measure space $(X, \mathscr{F}, \mu)$ can be made complete basically by "adding" subsets of null sets. This is called the completion of $(X, \mathscr{F}, \mu)$.

Proposition 3.2.5

Let $(X, \mathscr{F}, \mu)$ be a measure space. Define

$$\widetilde{\mathscr{F}} := \{E \subset X \ : \ \exists A, B \in \mathscr{F},\ A \subset E \subset B,\ \mu(B \backslash A) = 0\},\ \ \widetilde{\mu}(E) := \mu(A).$$

Then,

i) $\widetilde{\mathscr{F}}$ is a $\sigma-$algebra containing $\mathscr{F}$.
ii) $\widetilde{\mu}$ is a well defined measure on $\widetilde{\mathscr{F}}$.
iii) $\widetilde{\mu}(E) = \mu(E)$ for every $E \in \mathscr{F}$.

So, if needed, we can always assume that our working space is complete. In this case, for example, we can freely modify a measurable function on a measure zero set still obtaining a measurable function. This makes us to appreciate, once more, how weak is measurability: if you just modify the value of a continuous function in one single point, you loose continuity!

Proposition 3.2.6

Let $(X, \mathscr{F}, \mu)$ be a **complete** measure space. The following statements hold:

i) Let $f = g$ a.e. on E. Then $f \in L(E)$ iff $g \in L(E)$.
ii) if $f \in L(E \backslash N)$ with $\mu(N) = 0$, then $f \in L(E)$.

PROOF. i) Let $f \in L(E)$ and let's check that $g \in L(E)$. We can write,

$$\{g \in I\} = \{x \in E \ : \ g(x) \in I\} = \{x \in E \backslash N \ : \ g(x) \in I\} \cup \underbrace{\{x \in N \ : g(x) \in I\}}_{\widetilde{N}}$$

Now: $\widetilde{N} \subset N$ and because $\mathscr{F}$ is complete, $\widetilde{N} \in \mathscr{F}$ and $\mu(\widetilde{N}) = 0$. Moreover,

$$\begin{aligned} \{x \in E \backslash N \ : \ g(x) \in I\} &= \{x \in E \backslash N \ : \ f(x) \in I\} = (\{x \in E \ : \ f(x) \in I\}) \backslash (\{x \in N \ : \ f(x) \in I\}) \\ &= \{f \in I\} \backslash \widehat{N} \end{aligned}$$

where, again, $\widehat{N} \subset N$ (same argument used above for $\widetilde{N}$) has measure 0. Therefore, $\{f \in I\} \backslash \widehat{N} \in \mathscr{F}$ (difference of measurable sets) so, in conclusion,

$$\{g \in I\} = \underbrace{\{f \in I\} \backslash \widehat{N}}_{\in \mathscr{F}} \cup \widetilde{N} \in \mathscr{F}.$$

ii) It is similar to i):

$$\{f \in I\} = \underbrace{\{x \in E \backslash N \ : \ f(x) \in I\}}_{\in \mathscr{F}} \cup \underbrace{\{x \in N \ : \ f(x) \in I\}}_{\subset N} \in \mathscr{F}. \quad \square$$

Remark 3.2.7

In particular, if $f \equiv g$ a.e. on $E \subset \mathbb{R}^d$ and $f \in \mathscr{C}(E)$ then $g \in L(E)$. $\square$

3.3. Pointwise limit of measurable functions

Another remarkable feature of measurability is that it is preserved under very weak limit operations.

Definition 3.3.1

Let $(X, \mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L(E)$ be a sequence of measurable functions. We say that (f_n) **converges a.e. to** f **on** E (notation, $f_n \xrightarrow{a.e.} f$) if

$$\exists \lim_{n \to +\infty} f_n(x) = f(x), \ a.e. \, x \in E.$$

Almost everywhere limit of measurable functions is a measurable function:

Theorem 3.3.2

Let $(X, \mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L(E)$. The following statements hold:

i) if $f_n \longrightarrow f$ for every $x \in E \in \mathscr{F}$, then $f \in L(E)$.

ii) il $f_n \longrightarrow f$ a.e. $x \in E$ and $(X, \mathscr{F}, \mu)$ is complete, then $f \in L(E)$.

Proof. i) We prove that $\{f > a\} \in \mathscr{F}$ for every $a \in \mathbb{R}$, from this the conclusion follows . The idea is that $f > a$ means that, f_n must be definitely $> a$, thus we can connect $\{f > a\}$ to sets $\{f_n > a\}$ for which we have measurability by assumption. Let's see this precisely. First, the set of x for which sequence $(f_n(x))$ is definitely larger than a b is

$$\bigcup_k \bigcap_{n \geqslant k} \{f_n > b\} \in \mathscr{F}, \ \forall b \in \mathbb{R}.$$

Then, notice that the following identity holds:

$$\{f > a\} = \bigcup_{b \in \mathbb{Q}, \ b > a} \bigcup_k \bigcap_n \{f_n > b\} \in \mathscr{F}$$

Indeed: if $f(x) > a$ then, since $f_n(x) \longrightarrow f(x) > a$, taking $b \in \mathbb{Q}$ such that $a < b < f(x)$, by definition of limit $f_n(x) > b$ for all $n \geqslant k$ for a suitable k. Thus $\subset$ holds. Viceversa, if x belongs to the r.h.s., then

$f_n(x) > b$ for some $b > a$ and for every $n \geqslant k$ for some k. Thus $f(x) = \lim_n f_n(x) \geqslant b > a$, and this proves $\supset$.
ii) Let N such that $\mu(N) = 0$ and $f_n \longrightarrow f$ on $X \backslash N$. Applying the previous part we deduce that $f \in L(X \backslash N)$, and since the space is complete, by the Proposition 3.2.6 it follows that $f \in L(X)$.

Remark 3.3.3

Other important properties such as continuity or differentiability do not "pass" to the point-wise limit. For example, $f_n(x) = x^n$, $f_n \in \mathscr{C}([0,1])$ and

$$f_n(x) = x^n \longrightarrow \begin{cases} 0, & 0 \leqslant x < 1, \\ \\ 1, & x = 1 \end{cases} =: f(x),\ \forall x \in [0,1].$$

Clearly, $(f_n) \subset L([0,1])$ and $f = 1_1 \in L([0,1])$ (this confirms previous thm), but while $(f_n) \subset \mathscr{C}([0,1])$ we have $f \notin \mathscr{C}([0,1])$. $\square$

3.4. Exercises

Exercise 3.4.1 ($*$). *Show that any monotone function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is Lebesgue measurable.*

Exercise 3.4.2 ($**$). *Show that f is $\mathscr{F}-$measurable iff $\{f > a\} \in \mathscr{F}$ for every $a \in \mathbb{Q}$.*

Exercise 3.4.3 ($**$). *Let $\mathscr{F}$ be a measured space. The aim of this exercise if to prove that*

$$f \in L(E), \iff \{f \in B\} \in \mathscr{F},\ \forall B \in \mathscr{B}(\mathbb{R}^d).$$

i) *Check $\Longleftarrow$.*
ii) *Check $\Longrightarrow$ Define $\mathscr{B} \subset \mathscr{B}(\mathbb{R}^d)$ the class of sets $B \in \mathscr{B}(\mathbb{R}^d)$ such that $\{f \in B\} \in \mathscr{F}$. Show that $\mathscr{B}$ is a $\sigma-$algebra and $\mathscr{B} = \mathscr{B}(\mathbb{R}^d)$.*

Exercise 3.4.4 ($***$). *Adapt the ideas of the proof of the proposition 3.1.7 to check that if $f, g \in L(E)$ then also $f \cdot g \in L(E)$. (hint: if $g > 0$, $fg > a$ is equivalent to $f > \frac{a}{g} \dots$).*

Exercise 3.4.5 ($***$). *Prove the proposition 3.1.8. (hint: $\{\varphi(f) > a\} = f^{-1}\left(\varphi^{-1}(]a, +\infty[)\right)$, $\varphi^{-1}(]a, +\infty[)$ is open being φ continuous;)*

Exercise 3.4.6 ($*$). *For each of the following sequences of functions (f_n) on $(\mathbb{R}, \mathscr{M}_1, \lambda_1)$, determine if they are a.e. convergent and, in this case, to what.*

i) $f_n(x) = 1_{[n,n+1]}(x)$.
ii) $f_n(x) = 1_{[1/n,n]}(x)$.
iii) $f_n(x) := 1_{[\frac{1-(-1)^n}{4}, \frac{3-(-1)^n}{4}]}(x)$.

Exercise 3.4.7 ($**$). *Let $g \in \mathscr{C}(\mathbb{R}^d)$ be such that $g = 0$ a.e.. Deduce that, necessarily, $g \equiv 0$ that is $g(x) = 0$ for every $x \in \mathbb{R}^d$.*

Exercise 3.4.8 ($**$). *Let $(f_n) \subset L(E)$. Define*

$$f(x) := \inf_n f_n(x), \quad g(x) := \sup_n f_n(x).$$

Check that

i) $\{f = -\infty\}, \{g = +\infty\} \in \mathscr{F}$.

ii) *if* $F := E \setminus \{f = -\infty\}$ *and* $G := E \setminus \{g = +\infty\}$*, then* $f \in L(F)$ *and* $g \in L(G)$.

Exercise 3.4.9 ($***$)**.** *We recall that a sequence* (a_n) *of real numbers is convergent iff it is a Cauchy sequence, that is iff the following property holds:*

$$\forall \varepsilon > 0,\ \exists N = N(\varepsilon) \in \mathbb{N},\ :\ |a_n - a_m| \leqslant \varepsilon,\ \forall n, m \geqslant N.$$

With this in mind, let $(f_n) \subset L(X, \mathscr{F})$ *be a sequence of measurable functions on* X*. Check that the set*

$$S := \{x \in X\ :\ (f_n(x)) \text{ converges in } \mathbb{R}\}$$

is $\mathscr{F}-$*measurable. (hint: S is the set of* $x \in X$ *for which* $(f_n(x))$ *is a Cauhcy sequence... Use set operations to express S under this form).*

LECTURE 4

Abstract Integral

Measure allows a general definition of integral. The relevance of such definition is both in its versatility and in its power. Abstract integrals are used in Analysis and Geometry, they provides foundations to Probability and to Quantum Physics. Their tools as by far stronger than usual Riemann tools. This is why the integral introduced with this Lecture can be considered the ”true integration Theory”.

4.1. Lebesgue definition

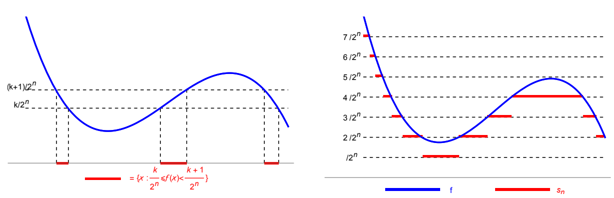
Let $(X, \mathscr{F}, \mu)$ be a measure space and let $f \in L(E)$ be a measurable function on $E \subset X$. The goal is to define the integral $\int_E f \, d\mu$. The main steps of this construction are: (a) the case of positive measurable functions; (b) extension to real-valued functions; (c) extension to complex-valued functions. This last extension is important in its own right and is particularly relevant because the Fourier transform (the characteristic function in probability) is the integral of a complex-valued function.

We begin with the case of positive measurable functions. The traditional Riemann approach to integration is based on partitioning the domain, whereas the Lebesgue approach is based on partitioning the co-domain. Fix n and divide the co-domain $[0, +\infty[$ as follows:

$$[0, +\infty[= \left[0, \frac{1}{2^n}\right[\cup \left[\frac{1}{2^n}, \frac{2}{2^n}\right[\cup \ldots \cup \left[\frac{2^{2n}-1}{2^n}, 2^n\right[\cup [2^n, +\infty[\,.$$

Then, we define simple functions $s_n(x)$

$$s_n(x) := \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{\frac{k}{2^n} \leqslant f < \frac{k+1}{2^n}} + 2^n 1_{f \geqslant 2^{2n}} \,. \tag{4.1.1}$$



These are measurable simple functions (because $f \in L(E)$) and, by their definition, their graphs are below that one of f. Letting $n \to +\infty$ we have a point-wise approximation of f:

Proposition 4.1.1

Let $(X, \mathscr{F}, \mu)$ be a measured space, $f \in L(E)$, $f \geqslant 0$ a positive measurable function. Let (s_n) the sequence of simple functions defined by (4.1.1). Then,

i) $s_0(x) \leqslant s_1(x) \leqslant \ldots \leqslant s_n(x) \leqslant f(x)$, for every $x \in E$;
ii) $\lim_{n\to+\infty} s_n(x) = f(x)$, for every $x \in E$.

PROOF. Define now sets

$$E_{k,n} := \left\{ \frac{k}{2^n} \leqslant f < \frac{k+1}{2^n} \right\},\ k = 0, \ldots, 2^{2n} - 1,\ \ E_{2^{2n},n} := \{f \geqslant 2^n\}.$$

Since f is measurable, sets $E_{k,n}$ are measurable. Define now

$$s_n(x) := \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{E_{k,n}} + 2^n 1_{E_{2^{2n},n}}.$$

Clearly s_n is a simple function and, by construction, $s_n(x) \leqslant f(x)$ for every $x \in E$. It holds $s_n \leqslant s_{n+1}$ (exercise). Let's prove ii) of statement. Fix $x \in E$ and pick N big enough in such a way that $f(x) < 2^N$. Then, for $n \geqslant N$, there exists a unique $E_{k,n} \ni x$. In particular, $s_n(x) = \frac{k}{2^n}$, whence hence

$$0 \leqslant f(x) - s_n(x) \leqslant \frac{k+1}{2^n} - \frac{k}{2^n} = \frac{1}{2^n}.$$

Letting $n \longrightarrow +\infty$ we get the conclusion.

We now set

$$\int_E s_n \, d\mu := \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \mu\left(\frac{k}{2^n} \leqslant f < \frac{k+1}{2^n} \right) + 2^n \mu(f \geqslant 2^n).$$

Since μ could take value $= +\infty$, we need to specify how to handle the algebra with infinities. We will assume the following natural algebraic rules:

$$c \cdot (+\infty) = \begin{cases} 0, & c = 0, \\ +\infty, & c > 0, \end{cases} \qquad (+\infty) + (+\infty) = +\infty.$$

With these agreements, the integral $\int_E s_n \, d\mu$ is well defined and it can be considered as an approximation by defect of the integral $\int_E f \, d\mu$. We will now prove that the sequence of integrals $\int_E s_n \, d\mu$ is actually convergent. The limit value will be, by definition, $\int_E f \, d\mu$.

Proposition 4.1.2

Let $(X, \mathscr{F}, \mu)$ be a measure space and $f \in L(E)$. Then,

$$\exists \lim_{n\to+\infty} \int_E s_n \, d\mu =: \int_E f \, d\mu \in [0 + \infty]. \tag{4.1.2}$$

PROOF. We prove that the sequence of integrals $\int_E s_n \, d\mu$ is increasing with n. From this, it will follow that the limit $\lim_n \int_E s_n \, d\mu$ exists, so the definition of $\int_E f \, d\mu$ makes sense and the conclusion follows. To prove this, we notice that

$$\mu\left(\frac{k}{2^n} \leqslant f < \frac{k+1}{2^n}\right) = \mu\left(\frac{2k}{2^{n+1}} \leqslant f < \frac{2(k+1)}{2^{n+1}}\right) = \mu\left(\frac{2k}{2^{n+1}} \leqslant f < \frac{2k+1}{2^{n+1}}\right) + \mu\left(\frac{2k+1}{2^{n+1}} \leqslant f < \frac{2k+2}{2^{n+1}}\right).$$

Therefore,

$$\begin{aligned}
\int_E s_n\, d\mu &= \textstyle\sum_{k=0}^{2^{2n}-1} \frac{k}{2^n}\left[\mu\left(\frac{2k}{2^{n+1}} \leqslant f < \frac{2k+1}{2^{n+1}}\right) + \mu\left(\frac{2k+1}{2^{n+1}} \leqslant f < \frac{2k+2}{2^{n+1}}\right)\right] + 2^n\mu(f \geqslant 2^n) \\
&\leqslant \underbrace{\sum_{k=0}^{2^{2n}-1} \frac{2k}{2^{n+1}}\mu\left(\frac{2k}{2^{n+1}} \leqslant f < \frac{2k+1}{2^{n+1}}\right) + \sum_{k=0}^{2^{2n}-1} \frac{2k+1}{2^{n+1}}\mu\left(\frac{2k+1}{2^{n+1}} \leqslant f < \frac{2k+2}{2^{n+1}}\right)}_{\sum_{j=0}^{2^{2n+1}-1} \frac{j}{2^{n+1}}\mu\left(\frac{j}{2^{n+1}} \leqslant f < \frac{j+1}{2^{n+1}}\right)} + \\
&\qquad + \underbrace{2^n\mu(2^n \leqslant f < 2^{n+1})}_{\frac{2^{2n+1}}{2^{n+1}}\mu\left(\frac{2^{2n+1}}{2^{n+1}} \leqslant f < \frac{2^{2n+1}+1}{2^{n+1}}\right)} + 2^n\mu(f \geqslant 2^{n+1}) \\
&\leqslant \textstyle\sum_{j=0}^{2^{2(n+1)}-1} \frac{j}{2^{n+1}}\mu\left(\frac{j}{2^{n+1}} \leqslant f < \frac{j+1}{2^{n+1}}\right) + 2^{n+1}\mu(f \geqslant 2^{n+1}) \\
&= \int_E s_{n+1}\, d\mu.
\end{aligned}$$

With some technical work, some first properties of this definition can be obtained

Proposition 4.1.3

Let $(X, \mathscr{F}, \mu)$ be a measure space. The following properties hold:

i) if $f, g \in L(E)$, $0 \leqslant f \leqslant g$, then $\int_E f\, d\mu \leqslant \int_E g\, d\mu$.

ii) if $f, g \in L(E)$, $f, g \geqslant 0$ and $\alpha, \beta \geqslant 0$ are constant, then

$$\int_E (\alpha f + \beta g)\, d\mu = \alpha \int_E f\, d\mu + \beta \int_E g\, d\mu.$$

Proofs are left in the exercises. We now extend the definition of integral to real valued functions. We introduce

$$f_+ := \max\{f, 0\}, \ \textbf{(positive part)} \quad f_- := \max\{-f, 0\}, \ \textbf{(negative part)}.$$

Since $\max\{y, 0\}$ and $\max\{-y, 0\}$ are continuous functions of y, both $f_\pm$ are measurable if f is measurable. Both are also non negative functions.

Definition 4.1.4

Let $(X, \mathscr{F}, \mu)$ be a measure space and $f \in L(E)$. We say that f is $\mu-$**integrable** if

$$\int_X |f| \, d\mu < +\infty.$$

In this case we set

$$\int_X f := \int_X f_+ \, d\mu - \int_X f_- \, d\mu.$$

We write $f \in L^1(X, \mathscr{F}, \mu)$, if the measure space is understood, we just write $f \in L^1(E)$.

Remark 4.1.5

The definition is well posed. Indeed: since $\int_X |f| \, d\mu$ is finite, being $0 \leqslant f_\pm \leqslant |f|$ we have that also

$$\int_X f_\pm \, d\mu \leqslant \int_X |f| \, d\mu < +\infty.$$

Thus the difference $\int_X f_+ \, d\mu - \int_X f_- \, d\mu$ makes sense (we do not have the indeterminate form $(+\infty) - (+\infty)$.

The final extension it to the case of complex valued functions.

Definition 4.1.6

Let $(X, \mathscr{F}, \mu)$ be a measured space. Given $f : X \longrightarrow \mathbb{C}$, we say that f is **measurable** (notation $f \in L_{\mathbb{C}}(X)$) if both $\mathrm{Re}\, f$ and $\mathrm{Im}\, f$ are measurable (that is $\mathrm{Re}\, f, \mathrm{Im}\, f \in L(X)$). We say that f is $\mu-$**integrable** (notation $f \in L^1_{\mathbb{C}}(X, \mathscr{F}, \mu)$) if

$$\int_X |f| \, d\mu < +\infty.$$

In this case we set

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

Remark 4.1.7: A

so for this case, the definition is well posed. Indeed, since $|\mathrm{Re}\, f|, |\mathrm{Im}\, f| \leqslant |f|$, we have $\int_X |\mathrm{Re}\, f| \, d\mu, \int_X |\mathrm{Im}\, f| \, d\mu \leqslant \int_X |f| \, d\mu < +\infty$, thus $\mathrm{Re}\, f, \mathrm{Im}\, f \in L^1(X)$ and both integrals $\int_X \mathrm{Re}\, f \, d\mu, \int_X \mathrm{Im}\, f \, d\mu \in \mathbb{R}$. Hence, the value of $\int_X f \, d\mu$ is well defined.

4.2. General properties

We summarize, in the next proposition, the main properties of the abstract integral. Proofs are omitted here.

Proposition 4.2.1

Let $(X, \mathscr{F}, \mu)$ be a measure space. The following properties hold:

i) (linearity) if $f, g \in L^1(E)$ and $\alpha, \beta \in \mathbb{R}$ $(\mathbb{C})$ then
$$\int_E (\alpha f + \beta g)\, d\mu = \alpha \int_E f\, d\mu + \beta \int_E g\, d\mu.$$

ii) (ordering) if $f, g \in L^1(E)$ are real valued and $f \leqslant g$, then
$$\int_E f\, d\mu \leqslant \int_E g\, d\mu.$$

iii) (triangular inequality) if $f \in L^1(E)$, then
$$\left| \int_E f\, d\mu \right| \leqslant \int_E |f|\, d\mu.$$

iv) (restriction) if $f \in L^1(E)$ and $F \in \mathscr{F}$, $F \subset E$, then $f \in L^1(F)$ and
$$\int_F f\, d\mu = \int_E f 1_F\, d\mu.$$

v) (decomposition) if $f \in L^1(E), L^1(F)$ with $E \cap F = \emptyset$, then $f \in L^1(E \sqcup F)$ and
$$\int_{E \sqcup F} f\, d\mu = \int_E f\, d\mu + \int_F f\, d\mu.$$

vi) (null sets) if $\mu(N) = 0$ then $\int_N f\, d\mu = 0$. In particular, if $f, g \in L^1(E)$ and $f = g$ a.e., then
$$\int_E f\, d\mu = \int_E g\, d\mu.$$

An important inequality is given in the following

Lemma 4.2.2: Chebyshev's inequality

Let $(X, \mathscr{F}, \mu)$ be a measure space, and $f \in L(X)$, $f \geqslant 0$ a positive measurable function. Then

$$\mu(f \geqslant \alpha) \leqslant \frac{1}{\alpha} \int_{f \geqslant \alpha} f\, d\mu \leqslant \frac{1}{\alpha} \int_X f\, d\mu,\ \forall \alpha > 0. \tag{4.2.1}$$

Proof. We have
$$\mu(f \geqslant \alpha) = \int_{f \geqslant \alpha} 1\, d\mu \overset{f \geqslant \alpha \iff \frac{f}{\alpha} \geqslant 1}{\leqslant} \int_{f \geqslant \alpha} \frac{f}{\alpha}\, d\mu = \frac{1}{\alpha} \int_{f \geqslant \alpha} f\, d\mu \leqslant \frac{1}{\alpha} \int_X f\, d\mu.$$

This apparently simple inequality has important consequences. A first example is the following:

Corollary 4.2.3

Let $(X, \mathscr{F}, \mu)$ be a measure space, and $f \in L(E)$, $f \geqslant 0$ a positive measurable function. If
$$\int_E f \, d\mu = 0,$$
then $f = 0$ a.e. on E.

PROOF. Since $\int_E f \, d\mu = 0$, by Chebyshev's inequality we have that
$$\mu(f \geqslant \alpha) = 0, \; \forall \alpha > 0.$$
Our goal is to prove that
$$\mu(f > 0) = 0,$$
from which the conclusion will follow. To this aim, notice that, if $E_n := \{f \geqslant \frac{1}{n}\}$, then $E_n \nearrow E := \{f > 0\}$. By continuity from below,
$$\mu(f > 0) = \lim_n \mu \left(f \geqslant \frac{1}{n} \right) = 0. \quad \square$$

Here is another application of Chebyshev's inequality.

Proposition 4.2.4

Let $(X, \mathscr{F}, \mu)$ be a measure space, $f \in L(X)$, $f \geqslant 0$. Then
$$\int_X f \, d\mu < +\infty, \implies f < +\infty, \; a.e.$$

PROOF. Just notice that $\{f = +\infty\} = \bigcap_n \{f \geqslant n\}$, and since
$$\mu(f \geqslant n) \leqslant \frac{1}{n} \int_X f \, d\mu = \frac{C}{n},$$
we get that $\mu(f = +\infty) \leqslant \frac{C}{n}$, for every $n \in \mathbb{N}$. Letting $n \longrightarrow +\infty$ we have the conclusion.

4.3. Exercises

Exercise 4.3.1 ($**$)**.** *Extend theorem 4.1 to any $f \in L(E)$. Prove that there exists (s_n) of simple functions such that*
$$\lim_{n \to +\infty} s_n(x) = f(x), \; \forall x \in E.$$
(you may start writing $f = f_+ - f_-$)

Exercise 4.3.2 ($**$)**.** *Let $(X, \mathscr{F}, \mu)$ be a measure space. Prove that*
$$\mu(E) = \int_X 1_E \, d\mu, \; \forall E \in \mathscr{F}.$$
(hint: distingush cases $\mu(E) = +\infty$ from $\mu(E) < +\infty$).

Exercise 4.3.3 ($***$)**.** *Let $f, g \in L(E)$, $0 \leqslant f \leqslant g$. Show that*

i) $s_n^f(x) \leqslant s_n^g(x)$ *(here, s_n^f, s_n^g are, respectively, the simple functions (4.1.1) built on f and g).*

ii) *from the definition (4.1.2), $\int_E f\, d\mu \leqslant \int_E g\, d\mu$.*

Exercise 4.3.4 (∗∗). *Given a measure space $(X.\mathscr{F}, \mu)$, let $f \in L(E)$, with $0 < \mu(E) < +\infty$. Suppose moreover that $0 \leqslant f(x) \leqslant M$ a.e. $x \in E$. Show that if $\int_E f\, d\mu = M\mu(E)$, then $f = M$ a.e. on E.*

Exercise 4.3.5 (∗∗). *Let $(X, \mathscr{F}, \mu)$ and let $f \in L^1$ be such that*

$$\left| \int_E f\, d\mu \right| = \int_E |f|\, d\mu,$$

for some $E \in \mathscr{F}$. Prove that f has constant sign on E $\mu-$a.e.

Exercise 4.3.6 (∗). *Let $(X, \mathscr{F}, \mu)$ and $f \in L(X)$, $f \geqslant 0$. Extend Chebyshev's inequality:*

$$\mu(f \geqslant \alpha) \leqslant \frac{1}{\alpha^p} \int_X f^p\, d\mu,\ \forall \alpha > 0,\ \forall p \geqslant 1.$$

Is this still true for $0 < p < 1$?

Exercise 4.3.7 (∗∗). *Let $(X, \mathscr{F}, \mu)$ and $f \in L(X)$, $f \geqslant 0$. Extend Chebyshev's inequality as follows: for $\phi = \phi(x) : [0, +\infty[\longrightarrow \mathbb{R}$increasing and convex function with $\phi(0) = 0$, prove that*

$$\mu(f \geqslant \alpha) \leqslant \frac{1}{\phi(\alpha)} \int_X \phi(f)\, d\mu,\ \forall \alpha > 0.$$

(hint: remind that ϕ is convex iff $\phi(tx + (1-t)y) \leqslant t\phi(x) + (1-t)\phi(y)$, $\forall x, y$, $\forall t \in [0, 1]$)

Exercise 4.3.8 (∗∗). *Let $(X, \mathscr{F}, \mu)$ and $f \in L^1(X)$ be such that*

$$\int_E f\, d\mu \geqslant 0,\ \forall E \in \mathscr{F}.$$

Show that $f \geqslant 0$ a.e.

Exercise 4.3.9 (∗∗+). *Let $f \in L^1(X, \mathscr{F}, \mu)$ and suppose that*

$$\left| \int_E f\, d\mu \right| \leqslant M\mu(E),\ \forall E \in \mathscr{F}.$$

Prove that $|f(x)| \leqslant M$ a.e. (hint: consider $E = \{f \geqslant M + \varepsilon\}$ with $\varepsilon > 0$, use Chebyshev's inequality to show that $\mu(E) < +\infty$ and the assumption to prove that $\mu(E) = 0$; from this, deduce the conclusion. . .)

LECTURE 5

Lebesgue Integral

Lebesgue integral is the integral respect to the Lebesgue measure. It is usually written as

$$\int_E f(x)\,dx.$$

This because there is an important relation with the familiar Riemann integral taught in Calculus courses. This Lecture focuses on this particular integral and on some of its main features.

5.1. Comparison with Riemann and Generalized Integrals

In dimension $d = 1$, the well known definition of integral are *Riemann's integral* and *generalized integral*. We now have a new definition of integral, Lebesgue's integral

$$\int_E f\,d\lambda_1.$$

While Riemann and generalized integrals are well defined integrations on intervals, Lebesgue's integral allows a large flexibility about the domain. However, when the integration domain is an interval, a comparison makes sense. We start with the case of Riemann's integral.

Theorem 5.1.1

If $f \in \mathscr{R}([a,b])$ (Riemann integrable) then $f \in L^1([a,b])$ and

$$\text{(Riemann)} \quad \int_a^b f(x)\,dx = \int_{[a,b]} f\,d\lambda_1 \quad \text{(Lebesgue)}.$$

Moreover, $f \in \mathscr{R}([a,b])$ iff $f \in L^1([a,b])$ and the set of discontinuities of f has measure 0.

In practice, to compute Lebesgue integrals in one variable we may use, where possible, tools from ordinary Calculus. Thus, for example, *if* $f \in \mathscr{C}^1([a,b])$ *then*

$$\int_{[a,b]} f'\,d\lambda_1 \equiv \int_a^b f'(x)\,dx = f(b) - f(a).$$

Warning 5.1.2

Lebesgue's integral is a huge extension of Riemann's integral. There are lot of functions which are not integrable in Riemann sense, while they are integrable in Lebesgue sense. An example is Dirichlet's function $1_{[0,1]\backslash\mathbb{Q}} \in L^1([0,1])\backslash\mathscr{R}([0,1])$.

Because of their identity on Riemann integrable functions, we will denote Lebesgue's integral on $[a,b]$ as $\int_a^b f(x)\,dx$. There is no ambiguity with this: when both are defined, they coincide; when only Lebesgue's integral makes sense, there is no risk of misunderstanding.

As known, Riemann integrable functions are necessarily *bounded* and defined on *closed and bounded intervals*. For many reasons, it is interesting to have an operation of integral on *unbounded intervals* and for *unbounded functions*, or for a combination of these two. This yields to the definition of generalized integral. For sake of simplicity, here we will focus on generalized integrals on unbounded intervals, but what we say here holds similarly for the other cases. We recall that

$$\int_a^{+\infty} f(x)\,dx := \lim_{b\to+\infty}\int_a^b f(x)\,dx.$$

Other cases work in the same way. In general, it may happens that a generalized integral exists but the corresponding Lebesgue's integral is not.

Example 5.1.3

Function $f(x) := \frac{\sin x}{x}$ is integrable in generalized sense on $[0,+\infty[$ but is not $L^1([0,+\infty[)$.

Proof. To simplify technical details, we consider a slight modification of such f defining

$$f(x) := \begin{cases} +1, & 0 \leqslant x < 1, \\ -1/2, & 1 \leqslant x < 2, \\ +1/3, & 2 \leqslant x < 3, \\ \vdots \end{cases} = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}1_{[n,n+1[}(x).$$

Then

$$\int_0^{+\infty} f(x)\,dx = \lim_{N\to+\infty}\int_0^N f(x)\,dx = \lim_{N\to+\infty}\sum_{n=1}^{N}(-1)^{n+1}\frac{1}{n} = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n},$$

which is, after Leibniz's test, a convergent series. Thus $\int_0^{+\infty} f(x)\,dx \in \mathbb{R}$. However,

$$\int_{[0,+\infty[}|f| = \sum_{n=1}^{\infty}\int_{[n,n+1[}\left|(-1)^{n+1}\frac{1}{n}\right| = \sum_{n=1}^{\infty}\frac{1}{n} = +\infty.$$

This means that $f \notin L^1([0,+\infty[)$.

In the previous example, $|f|$ is not integrable in generalized sense, that is f *is not absolutely integrable*. It turns out that, functions absolutely integrable in generalized sense are also integrable in Lebesgue sense and the two integrals coincide:

Theorem 5.1.4

If f is **absolutely integrable** in generalized sense on $[a,+\infty[$, that is $\int_a^{+\infty} |f(x)|\,dx < +\infty$, then $f \in L^1([a,+\infty[)$ and

$$\text{(generalized integral)} \int_a^{+\infty} f(x)\,dx = \int_{[a,+\infty[} f\,d\lambda_1 \ \text{(Lebesgue)}.$$

5.2. Reduction Formula

In dimension $d > 1$, the basic tool of calculus is **reduction formula**. As the name says, this formula allows to reduce an higher dimensional integral into a lower dimensional one. To illustrate the principle, we consider a function $f : E \subset \mathbb{R}^d \longrightarrow \mathbb{R}$ written as

$$f = f(x,y), \text{ where } (x,y) \in \mathbb{R}^k \times \mathbb{R}^h \equiv \mathbb{R}^d.$$

Theorem 5.2.1: Fubini

Assume $f = f(x,y) \in L^1(E)$. Then reduction formula holds true

$$\int_E f(x,y)\,dxdy = \int_{x\,:\,E_x \neq \varnothing} \left(\int_{E_x} f(x,y)\,dy \right) dx = \int_{y\,:\,E^y \neq \varnothing} \left(\int_{E^y} f(x,y)\,dx \right) dy. \tag{5.2.1}$$

where

$$E_x := \{y \in \mathbb{R}^h \ : \ (x,y) \in E\}, \quad E^y := \{x \in \mathbb{R}^k \ : \ (x,y) \in E\}.$$

To apply reduction formula in practical cases, we need to know $f \in L^1(E)$, that is

$$\int_E |f(x,y)|\,dxdy < +\infty.$$

Notice that, in this case, by (5.2.1) applied to $|f|$, we have

$$+\infty > \int_E |f(x,y)|\,dxdy = \int_{x\,:\,E_x \neq \varnothing} \left(\int_{E_x} |f(x,y)|\,dy \right) dx = \int_{y\,:\,E^y \neq \varnothing} \left(\int_{E^y} |f(x,y)|\,dx \right) dy.$$

A vice versa is also true:

Theorem 5.2.2: Tonelli

Assume $f \in L(E)$ be such that one of the integrals

$$\int_{x\,:\,E_x \neq \varnothing} \left(\int_{E_x} |f(x,y)|\,dy \right) dx, \quad \int_{y\,:\,E^y \neq \varnothing} \left(\int_{E^y} |f(x,y)|\,dx \right) dy$$

is finite. Then $f \in L^1(E)$.

Theorems 5.2 and 5.2 are usually applied together. Given $f \in L(E)$ we apply first Thm 5.2 to check if $f \in L^1$. If this is the case, then we may apply Thm 5.2 to compute the integral. The two statements joint are called also *Fubini-Tonelli Theorem*. But warning! It might happens that both iterated integral of (5.2.1) are finite and $f \notin L^1$.

Example 5.2.3

Let
$$f(x,y) = \frac{x-y}{(x+y)^3},\ (x,y) \in E := [0,1]^2.$$
Then $\int_{x\,:\,E_x \neq \emptyset} \left(\int_{E_x} f\, dy \right)\, dx \neq \int_{y\,:\,E^y \neq \emptyset} \left(\int_{E^y} f\, dx \right)\, dy$ (hence, in particular, $f \notin L^1([0,1]^2)$).

Proof. Notice first that
$$E^y = \{x \in \mathbb{R}\ :\ (x,y) \in [0,1]^2\} = \begin{cases} \emptyset, & y \notin [0,1], \\ [0,1] & y \in [0,1] \end{cases}$$
and similarly for E_x. Therefore
$$\int_{E^y} f(x,y)\,dx = \begin{cases} 0, & y \notin [0,1], \\ \displaystyle\int_0^1 \frac{x-y}{(x+y)^3}\,dx = \int_0^1 \frac{x+y-2y}{(x+y)^3}\,dx = \int_0^1 \frac{1}{(x+y)^2}\,dx - 2y\int_0^1 \frac{1}{(x+y)^3}\,dx. & y \in [0,1]. \end{cases}$$
Except for $y=0$ (a measure 0 set) both integrals are finite and their value is
$$\left[\frac{(x+y)^{-1}}{-1}\right]_{x=0}^{x=1} - 2y\left[\frac{(x+y)^{-2}}{-2}\right]_{x=0}^{x=1} = \frac{1}{y} - \frac{1}{y+1} + y\left(\frac{1}{(y+1)^2} - \frac{1}{y^2}\right) = -\frac{1}{(y+1)^2}.$$
Hence
$$\int_{BbbR} \left(\int_{E^y} f(x,y)\,dx \right)\,dy = \int_0^1 \left(-\frac{1}{(y+1)^2} \right)\,dy = \left[(y+1)^{-1}\right]_{y=0}^{y=1} = \frac{1}{2} - 1 = -\frac{1}{2}.$$
Exchanging x with y we obtain the same result except for the sign: $\int_{\mathbb{R}} \left(\int_{E_x} f(x,y)dy \right)\,dx = \frac{1}{2}$.

5.3. Change of variable formula

Change of variable is an important tool of calculus of integrals. Let $f \in L^1(E)$ and suppose we aim to compute
$$\int_E f(x)\,dx.$$
Suppose moreover that, to compute the integral, it looks to be convenient to introduce a new variable $y = \Phi(x)$. With this we mean that $\Phi : E \longrightarrow F$ is a good (regular) transformation and a bijection, so that we can also express x as function of y, $x = \Phi^{-1}(y)$. We wonder how the integral w.r.t. x variable transforms into an integral into the y variable.

Proceeding informally, imagine we may decompose E into "small" sub-domains E_n, that is $E = \bigsqcup_n E_n$. Then we may expect that
$$\int_E f(x)\,dx = \sum_n \int_{E_n} f(x)\,dx \approx \sum_n f(x_n)\lambda(E_n),$$

for suitable points $x_n \in E_n$ (of course this is a delicate point, but this is not a proof!). Now, since $F = \Phi(E)$ we may decompose F as

$$F = \Phi\left(\bigsqcup_n E_n\right) = \bigsqcup_n \Phi(E_n) =: \bigsqcup_n F_n.$$

Since Φ is regular, around $x_n \in E_n$ we may linearize it, that is

$$\Phi(x) \approx \Phi(x_n) + \Phi'(x_n)(x - x_n),$$

thus

$$F_n = \Phi(E_n) = \Phi(x_n) + \Phi'(x_n)(E_n - x_n),$$

and because of the invariance formula (2.2.1) we have

$$\begin{aligned}\lambda(F_n) &= \lambda\left(\Phi(x_n) + \Phi'(x_n)(E_n - x_n)\right) = \lambda\left(\Phi'(x_n)(E_n - x_n)\right) = |\det \Phi'(x_n)|\lambda(E_n - x_n)\\ &= |\det \Phi'(x_n)|\lambda(E_n),\end{aligned}$$

or

$$\lambda(E_n) = \frac{1}{|\det \Phi'(x_n)|}\lambda(F_n).$$

Thus, setting $y_n = \Phi(x_n) \in F_n$,

$$\begin{aligned}\textstyle\int_E f(x)\,dx &\approx \textstyle\sum_n f(x_n)\frac{1}{|\det \Phi'(x_n)|}\lambda(F_n) \approx \sum_n f(\Phi^{-1}(y_n))\frac{1}{|\det \Phi'(\Phi^{-1}(y_n)|}\lambda(F_n)\\ &\approx \textstyle\int_F f(\Phi^{-1}(y))\frac{1}{|\det \Phi'(\Phi^{-1}(y))|}\,dy.\end{aligned}$$

Finally, recalling that $\frac{1}{\det A} = \det A^{-1}$ and that $\left(\Phi'(\Phi^{-1}(y)\right)^{-1} = (\Phi^{-1})'(y)$ we have an idea for the following

Theorem 5.3.1

Let $\Phi : E \longrightarrow F = \Phi(E)$ be such that $\Phi \in \mathscr{C}^1$ with $\Phi^{-1} \in \mathscr{C}^1$ (we say Φ is a **diffeomorphism**). Then $f(x) \in L^1(E)$ iff $f(\Phi^{-1}(y)|\det(\Phi^{-1})'(y)| \in L^1(F)$ and

$$\int_E f(x)\,dx = \int_{\Phi(E)} f(\Phi^{-1}(y))|\det(\Phi^{-1})'(y)|\,dy \tag{5.3.1}$$

5.4. Exercises

Exercise 5.4.1 ($*$). *Determine for which values of the parameter(s) the following integrals exist in L^1 sense:*

$$i)\ \int_0^{+\infty} \frac{1}{x^\alpha(1+x^\beta)}\,dx. \quad ii)\ \int_{-\infty}^{+\infty} e^{i\alpha(x+iy)^2}\,dy. \quad iii)\ \int_0^{+\infty} \frac{e^{-\alpha x} - e^{-x}}{x^\beta}\,dx.$$

Exercise 5.4.2 ($*$). *Detemine if the following integrals exists in L^1 sense:*

$$i)\ \int_0^{+\infty} \frac{\sin x}{x(1+x)}\,dx. \quad ii)\ \int_0^{+\infty} \frac{1-\cos x}{x^2(1+x)}\,dx.$$

Exercise 5.4.3 ($**$). *Determine for which values of $\alpha \in \mathbb{R}$ the function*

$$f(x) := \sum_{n=1}^{\infty} n^{\alpha} 1_{\frac{1}{n+1} < x \leqslant \frac{1}{n}}(x),$$

belongs to $L^1([0,1])$.

Exercise 5.4.4 ($**$). *In this problem we assume that the value of the integral $I := \int_{\mathbb{R}} e^{-t^2}\, dt$ is not known, and we compute it. By using in a suitable way the Tonelli thm, prove that the function $f : \mathbb{R}^2 \longmapsto \mathbb{R}$ given by $f(x,y) = ye^{-(1+x^2)y^2}$ is in $L^1(\mathbb{R}^2)$ and by using Fubini compute its integral on $[0,+\infty[^2$. Deduce the value of I. Justify everything with care.*

Exercise 5.4.5 ($**$). *Compute*

$$\int_0^{+\infty} \left(\int_0^{2\pi} \frac{y}{x} e^{-y/x} \sin x \, dx \right) \, dy.$$

Exercise 5.4.6 ($**$). *Justifying all steps, compute*

$$\int_0^{+\infty} \left(\int_x^{+\infty} e^{-y^2} \, dy \right) \, dx.$$

Exercise 5.4.7 ($**+$). *Let $f(x) = g(\|x\|) \in L(\mathbb{R}^m)$. Check that $f \in L^1(\mathbb{R}^m)$ iff $r^{m-1} g(r) \in L^1([0,+\infty[)$ and*

$$\int_{\mathbb{R}^m} f(x) \, dx = m \lambda_{m-1}(\mathbb{B}^{m-1}) \int_0^{\infty} r^{m-1} g(r) \, dr.$$

$(\lambda_{m-1}(\mathbb{B}_{m-1}) = \lambda_{m-1}(\{u \in \mathbb{R}^{m-1} \, : \, \|u\| \leqslant 1\})$. (hint: by symmetry, $\int_{\mathbb{R}^m} f = 2 \int_{x_m \geqslant 0} f$, hence notice that $x = ru$ where $r = \|x\|$ and $u = \frac{x}{\|x\|} \in \{u_1^2 + \cdots + u_m^2 = 1\} = \{u_m^2 = 1 - \|(u_1, \ldots, u_{m-1})\|^2\}$, so $x = \Phi(r, u_1, \ldots, u_{m-1}) = (ru_1, \ldots, ru_{m-1}, r\sqrt{1 - \|(u_1, \ldots, u_{m-1})\|^2})$, then apply change of variable)

Exercise 5.4.8 ($**$). *Determine for which values of $\alpha > 0$ it holds $f(x,y) = \frac{1}{(1-xy)^{\alpha}} \in L^1([0,1]^2)$.*

Exercise 5.4.9 ($**+$). *Let $E_{p,q} := \{(x,y) \in \mathbb{R}^2 \, : \, |x|^p + |y|^q \leqslant 1\}$, where $p, q > 0$. Show that*

$$\int_{E_{p,q}} \frac{1}{|x|^p + |y|^q} \, dxdy < +\infty \quad \Longleftrightarrow \quad \frac{1}{p} + \frac{1}{q} > 1.$$

(hint: adapt polar coordinates)

LECTURE 6

Monotone and Dominated Convergence

One of the most important features of abstract measure and integral is the extreme flexibility with *passage to the limit* into integral. The problem is the following: let $(f_n) \subset L^1(X)$, under which conditions can we say that

$$\exists \lim_n \int_X f_n \, d\mu = \int_X \lim_n f_n \, d\mu \ ?$$

Lebesgue's integral shows properties without any precedent for the ordinary Riemann's integral.

6.1. Monotone convergence

Consider a sequence $(f_n) \subset L(X)$ of positive (that is $f_n \geqslant 0$) measurable functions. Suppose moreover that the sequence (f_n) increases with n, that is

$$f_n \leqslant f_{n+1}, \ \forall n \in \mathbb{N}.$$

It is clear that

$$\int_X f_n \, d\mu \leqslant \int_X f_{n+1} \, d\mu, \ \forall n \in \mathbb{N}.$$

Therefore, as well known,

$$\exists \lim_n \int_X f_n \, d\mu \in [0, +\infty].$$

At same time, since $(f_n(x))$ is increasing with n, also $\lim_n f_n(x) =: f(x)$ exists. Clearly $f \geqslant 0$ and $f \in L(X)$ (because pointwise limit of measurable functions is measurable, see Theorem 3.3). Thus $\int_X f \, d\mu \in [0, +\infty]$. What is the relation between the integral of $f = \lim_n f_n$ and limit of integrals $\lim_n \int_X f_n \, d\mu$?

Theorem 6.1.1: Beppo Levi

Let $(X, \mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L(E)$ be such that $0 \leqslant f_n \leqslant f_{n+1}$ on E, for every n. Then

$$\lim_n \int_E f_n \, d\mu = \int_E \lim_n f_n \, d\mu. \tag{6.1.1}$$

Proof. We call $f(x) := \lim_n f_n(x)$. By i) of Theorem 3.3, $f \in L(E)$. Now, since $(f_n(x)) \nearrow$,

$$f_n(x) \leqslant f(x), \ \forall x \in X, \implies \int_X f_n \, d\mu \leqslant \int_X f \, d\mu, \implies \lim_n \int_X f_n \, d\mu \leqslant \int_X f \, d\mu.$$

The goal is to prove the $\geqslant$. Notice that if $\alpha = \lim_n \int_X f_n \, d\mu = +\infty$ then the conclusion is true. Suppose then that $\alpha < +\infty$ and let $\beta < \int_X f \, d\mu$. Our goal is to show that $\alpha > \beta$. From this, being β arbitrary, it follows that $\alpha \geqslant \int_X f \, d\mu$, which is the conclusion.
To show this, let (s_k) the increasing sequence of simple functions (4.1.1) for f, and similarly (s_k^n) is for f_n. Since $\int_X s_k \, d\mu \longrightarrow \int_X f \, d\mu$, there exists N such that

$$\int_E s_N \, d\mu = \sum_{j=0}^{2^{2N}-1} \frac{j}{2^N} \mu\left(\frac{j}{2^N} \leqslant f < \frac{j+1}{2^N}\right) + 2^N \mu(f \geqslant 2^N) > \beta.$$

Consider now

$$\int_E s_N^n \, d\mu = \sum_{j=0}^{2^{2N}-1} \frac{j}{2^N} \mu\left(\frac{j}{2^N} \leqslant f_n < \frac{j+1}{2^N}\right) + 2^N \mu(f_n \geqslant 2^N).$$

We aim to take the limit $n \longrightarrow +\infty$ in previous identity. First notice that measures $\mu(f_n \geqslant a)$ are finite and bounded in n. This is a consequence of Čebišëv inequality because

$$\mu(f_n \geqslant a) \leqslant \frac{1}{a} \int_X f_n \, d\mu \leqslant \frac{\alpha}{a}, \ \forall n \in \mathbb{N}.$$

Therefore, we can write

$$\mu\left(\frac{j}{2^N} \leqslant f_n < \frac{j+1}{2^N}\right) = \mu\left(f_n \geqslant \frac{j}{2^N}\right) - \mu\left(f_n \geqslant \frac{j+1}{2^N}\right).$$

Now, since $F_n := \{f_n \geqslant a\} \subset \{f_{n+1} \geqslant a\} = F_{n+1}$, by continuity from below we have

$$\mu(f_n \geqslant a) \longrightarrow \mu(f \geqslant a),$$

and, by previous bound, it follows that $\mu(f \geqslant a) \leqslant \frac{\alpha}{a} < +\infty$ for every $a > 0$. Therefore,

$$\mu\left(\frac{j}{2^N} \leqslant f_n < \frac{j+1}{2^N}\right) \longrightarrow \mu\left(\frac{j}{2^N} \leqslant f < \frac{j+1}{2^N}\right)$$

From this it follows that

$$\lim_n \int_E s_N^n \, d\mu = \int_E s_N \, d\mu > \beta.$$

Thus, for M large enough

$$\int_E s_N^M \, d\mu > \beta,$$

and since

$$\alpha \geqslant \int_E f_M \, d\mu \geqslant \int_E s_N^M \, d\mu > \beta,$$

the conclusion follows.

Corollary 6.1.2: monotone convergence theorem

Let $(X, \mathscr{F}, \mu)$ be a measured space, $(f_n) \subset L(E)$ be such that $0 \leqslant f_n \leqslant f_{n+1}$ **a.e.** for every n. Then

$$\lim_n \int_E f_n \, d\mu = \int_E \lim_n f_n \, d\mu. \tag{6.1.2}$$

PROOF. The apparently minimal extension is $f_n \leqslant f_{n+1}$ almost everywhere. That is:

$$f_n(x) \leqslant f_{n+1}(x),\ \forall x \in X\backslash N_n,\ \text{with}\ \mu(N_n) = 0.$$

Notice that the set N_n can depend on n, that is it is not necessarily the same for all f_n. Now the trick is: let $N := \bigcup_n N_n$. By sub additivity

$$\mu(N) \leqslant \sum_n \mu(N_n) = 0,$$

that is N is a null set. Moreover, we can now say that

$$f_n(x) \leqslant f_{n+1}(x),\ \forall x \in X\backslash N,\ \forall n \in \mathbb{N}.$$

Thus, applying Theorem 6.1 on $X\backslash N$ we have

$$\lim_n \int_{X\backslash N} f_n\, d\mu = \int_{X\backslash N} \lim_n f_n\, d\mu.$$

Finally, since N is a null set and $f = \lim_n f_n \in L(E\backslash N)$, we have that $f \in L(E)$ and

$$\int_{X\backslash N} f_n\, d\mu = \int_X f_n\, d\mu, \quad \int_{X\backslash N} f\, d\mu = \int_X f\, d\mu.$$

From this (6.1.2) follows.

Monotone convergence theorem requires very particular assumptions, in general hardly verified. Functions f_n must be positive, so the theorem does not apply to real or complex valued sequences. Furthermore, functions f_n must be ordered in the sense that $f_n \leqslant f_{n+1}$ a.e.. Let's see some example of application of this theorem.

Example 6.1.3: $(**)$

Compute

$$\lim_{n\to+\infty} \int_0^{+\infty} n \log\left(1 + \frac{e^{-x}}{n}\right)\, dx.$$

PROOF. Let $f_n(x) := n\log\left(1 + \frac{e^{-x}}{n}\right) = \log\left(1 + \frac{e^{-x}}{n}\right)^n$. Clearly $(f_n) \subset \mathscr{C}([0,+\infty[) \subset L([0,+\infty[)$ and $f_n \geqslant 0$. Recalling that

$$\left(1 + \frac{y}{n}\right)^n \nearrow e^y,\ \ \forall y \geqslant 0,$$

we have that $f_n(x) \leqslant f_{n+1}(x)$ for every $x \in [0,+\infty[$. Thus, we can apply monotone convergence theorem and, by (6.1.2) we have

$$\begin{aligned}\lim_{n\to+\infty} \int_0^{+\infty} n\log\left(1+\tfrac{e^{-x}}{n}\right)\, dx &= \int_0^{+\infty} \lim_{n\to+\infty} \log\left(1+\tfrac{e^{-x}}{n}\right)^n dx = \int_0^{+\infty} \log e^{e^{-x}}\, dx = \\ &= \int_0^{+\infty} e^{-x}\, dx = \left[-e^{-x}\right]_{x=0}^{x=+\infty} = 1. \quad \square\end{aligned}$$

Monotone convergence applies to series:

Corollary 6.1.4

Let $(X, \mathscr{F}, \mu)$ be a measured space, $(f_n) \subset L(X)$, $f_n \geqslant 0$ a.e.. Then

(6.1.3) $$\sum_n \int_X f_n \, d\mu = \int_X \sum_n f_n \, d\mu.$$

PROOF. Set $g_n := \sum_{k=1}^n f_k$. Clearly $(g_n) \subset L(X)$ and $g_n \geqslant 0$. According to (6.1.2) we have

$$\lim_n \int_X g_n \, d\mu = \int_X \lim_n g_n \, d\mu.$$

Now,

$$\lim_n g_n = \lim_n \sum_{k=1}^n f_k = \sum_{k=1}^\infty f_k,$$

while

$$\lim_n \int_X g_n \, d\mu = \lim_n \int_X \sum_{k=1}^n f_k \, d\mu = \lim_n \sum_{k=1}^n \int_X f_k \, d\mu = \sum_{k=1}^\infty \int_X f_k \, d\mu,$$

from which conclusion follows.

Example 6.1.5: $(**)$

Compute

$$\int_{[0,1]^2} \frac{1}{1-xy} \, dxdy.$$

PROOF. Notice that $f(x,y) := \frac{1}{1-xy} \in \mathscr{C}([0,1]^2 \backslash \{(1,1)\}) \subset L([0,1]^2]$ and also $f \geqslant 0$ a.e. on $[0,1]^2$. Recalling of the geometric sum $\sum_{n=0}^\infty q^n = \frac{1}{1-q}$ for $|q| < 1$ we have that

$$\frac{1}{1-xy} = \sum_{n=0}^\infty (xy)^n, \ \forall (x,y) \in [0,1]^2 \backslash \{(1,1)\}, \quad \Longleftrightarrow \quad a.e. \ (x,y) \in [0,1]^2.$$

Therefore, by (6.1.3), we have

$$\int_{[0,1]^2} \frac{1}{1-xy} \, dxdy = \int_{[0,1]^2} \sum_{n=0}^\infty x^n y^n \, dxdy = \sum_{n=0}^\infty \int_{[0,1]^2} x^n y^n \, dxdy.$$

Now, by reduction formula,

$$\int_{[0,1]^2} x^n y^n \, dxdy = \int_0^1 x^n \int_0^1 y^n \, dy \, dx = \left[\frac{x^{n+1}}{n+1} \right]_{x=0}^{x=1} \left[\frac{y^{n+1}}{n+1} \right]_{y=0}^{y=1} = \frac{1}{(n+1)^2}.$$

Thus

$$\int_{[0,1]^2} \frac{1}{1-xy} \, dxdy = \sum_{n=0}^\infty \frac{1}{(n+1)^2} = \sum_{n=1}^\infty \frac{1}{n^2}. \ \left(= \frac{\pi^2}{6} \right).$$

Integral of positive measurable functions allows to define large classes of measures.

Corollary 6.1.6

Let $(X, \mathscr{F}, \mu)$ be a measured space and $f \in L(X)$, $f \geqslant 0$. Then,

$$\mu_f(E) := \int_E f \, d\mu, \; E \in \mathscr{F},$$

is a measure on $\mathscr{F}$. We write $d\mu_f = f d\mu$ (for the origin of this notation see Exercise 6.3.6).

Proof. Clearly $\mu_f(\emptyset) = 0$. If $E = \bigsqcup_n E_n$ we have

$$\mu_f(E) = \int_E f \, d\mu = \int_X f 1_E \, d\mu = \int_X \sum_n f 1_{E_n} \, d\mu = \sum_n \int_X f 1_{E_n} \, d\mu = \sum_n \mu_f(E_n).$$

What about monotone convergence for a decreasing sequence of functions? In general, this is false, as the following example shows.

Example 6.1.7: $(*)$

On $(\mathbb{R}, \mathscr{M}_1, \lambda_1)$ consider $f_n := 1_{[n,+\infty[}$. Easily $(f_n) \subset L(\mathbb{R})$ and $f_n \geqslant f_{n+1}$ a.e.. Furthermore, $\lim_n f_n = 0$ thus

$$\int_{\mathbb{R}} \lim_n f_n \, dx = 0.$$

However, $\int_{\mathbb{R}} f_n \, dx = \lambda_1([n, +\infty[) = +\infty$, thus

$$\lim_n \int_{\mathbb{R}} f_n \, dx = +\infty \neq 0 = \int_{\mathbb{R}} \lim_n f_n \, dx. \qquad \square$$

This is, of course, the same phenomenon of continuity from above. By adding a finiteness assumption, the conclusion holds:

Corollary 6.1.8: decreasing monotone convergence

Let $(X\mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L(E)$ be such that $f_n \geqslant f_{n+1} \geqslant 0$ **a.e.** for every n. Assume that $\int_X f_1 \, d\mu < +\infty$. Then (6.1.2) holds.

The proof is left for exercise.

6.2. Dominated convergence

Monotone convergence shows that, under suitable circumstances, pointwise convergence is sufficient to pass limit into the integral. The two assumptions, namely $f_n \geqslant 0$ and $f_n \leqslant f_{n+1}$ a.e., are too restrictive. Is it possible to weaken these assumption? The next result is perhaps one of the most powerful results of Lebesgue Theory.

Theorem 6.2.1: Lebesgue's dominated convergence

Let $(X, \mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L^1(E)$. Assume that

i) (f_n) converges a.e. on E, that is $\exists \lim_{n\to+\infty} f_n(x) =: f(x)$, a.e. $x \in E$;

ii) there exists $g \in L^1(E)$ such that

$$|f(x)| \leqslant g(x), \; a.e.\, x \in E.$$

(g is called **integrable dominant**).

Then, $f \in L^1(E)$ and

$$\lim_n \int_E f_n \, d\mu = \int_E \lim_n f_n \, d\mu. \tag{6.2.1}$$

Proof. Arguing as in monotone convergence thm, we may assume that i) and ii) hold everywhere for $x \in E$. Since f is, by definition, the point wise limit of (f_n), $f \in L(E)$ (Theorem 3.3). Furthermore,

$$|f(x)| \overset{i)}{=} \lim_n |f_n(x)| \overset{ii)}{\leqslant} |g(x)|, \implies \int_E |f| \, d\mu \leqslant \int_E |g| \, d\mu \overset{ii)}{<} +\infty,$$

thus $f \in L^1(E)$. We prove now (6.2.1) by proving a stronger fact. Indeed, (6.2.1) is equivalent to

$$\lim_n \int_E (f - f_n) \, d\mu = 0.$$

Since

$$\left| \int_E (f - f_n) \, d\mu \right| \overset{\triangle}{\leqslant} \int_E |f - f_n| \, d\mu,$$

the conclusion follows once we prove

$$\lim_n \int_E |f - f_n| \, d\mu = 0.$$

Define

$$\delta_n := \sup_{k \geqslant n} |f_k - f|.$$

Clearly $\delta_n \geqslant |f_n - f|$. Furthermore, $\delta_n \in L(E)$ (exercise) and

$$\delta_{n+1} = \sup_{k \geqslant n+1} |f_k - f| \leqslant \sup_{k \geqslant n} |f_k - f| = \delta_n.$$

Thus $\delta_n \searrow$ and since $f_n \longrightarrow f$ point-wise on E, we have $\delta_n \longrightarrow 0$ point-wise on E. Finally, since

$$\int_E |\delta_1| \, d\mu = \int_E \sup_{k \geqslant 1} |f_k - f| \, d\mu \leqslant \int_X \sup_{k \geqslant 1} (|f| + |f_k|) \, d\mu \leqslant \int_X 2|g| \, d\mu < +\infty.$$

Thus, we verify hypotheses of the decreasing monotone convergence Corollary 6.1. Therefore

$$\lim_n \int_E |f - f_n| \, d\mu \leqslant \lim_n \int_E \delta_n \, d\mu = \int_E \lim_n \delta_n \, d\mu = \int_E u \, d\mu = 0.$$

Example 6.2.2: $(**)$

Compute
$$\lim_{n\to+\infty}\int_0^{+\infty} n^2\left(1-\cos\frac{x}{n}\right)e^{-\frac{n}{n+1}x}\,dx.$$

PROOF. Let $f_n(x) := n^2\left(1-\cos\frac{x}{n}\right)e^{-\frac{n}{n+1}x}$. Clearly $f_n \in L^1([0,+\infty[)$ for every n. Moreover
$$f_n(x) \sim_{n\to+\infty} n^2\frac{x^2}{2n^2}e^{-\frac{n}{n+1}x} = \frac{x^2}{2}e^{-\frac{n}{n+1}x} \longrightarrow \frac{x^2}{2}e^{-x} =: f(x),\ \forall x \geqslant 0.$$
Now,
$$|f_n(x)| \leqslant n^2\frac{(x/n)^2}{2}e^{-\frac{n}{n+1}x} = x^2e^{-\frac{n}{n+1}x},\ \forall x \geqslant 0,$$
and since $\frac{n}{n+1} \nearrow 1$, $\frac{n}{n+1} \geqslant \frac{1}{2}$ for $n \geqslant 1$. Hence
$$|f_n(x)| \leqslant x^2e^{-\frac{x}{2}} := g(x),\ \forall n \geqslant 1,\ \forall x \in [0,+\infty[.$$
Clearly $g \in L^1([0,+\infty[)$, so it is an integrable dominant. By dominated convergence then
$$\lim_{n\to+\infty}\int_0^{+\infty} f_n(x)\,dx = \int_0^{+\infty}\frac{x^2}{2}e^{-x}\,dx = \left[-\frac{x^2}{2}e^{-x}\right]_0^{+\infty} + \int_0^{+\infty} xe^{-x}\,dx = \left[-xe^{-x}\right]_0^{+\infty} + \int_0^{+\infty} e^{-x}\,dx = 1. \qquad \square$$

Here is a version of dominated convergence for series:

Corollary 6.2.3

Let $(X,\mathscr{F},\mu)$ be a measure space and let $(f_n) \subset L^1(E)$ be such that
$$\sum_n \int_E |f_n|\,d\mu < +\infty.$$
Then, $\sum_n f_n$ converges a.e., the sum belongs to $L^1(E)$ and
$$\int_E \sum_n f_n = \sum_n \int_E f_n.$$

6.3. Exercises

Exercise 6.3.1 $(**)$**.** *Compute*
$$\lim_{n\to+\infty}\int_n^{+\infty}\frac{e^{-n(x-n)}}{1+x^2}\,dx.$$

Exercise 6.3.2 $(**)$**.** *Compute*
$$\lim_{n\to+\infty}\int_0^{+\infty} n\left(1+\frac{x}{n}\right)^{-n}\sin\frac{x}{n}\,dx.$$

Exercise 6.3.3 $(**)$**.** *Compute*
$$\lim_{n\to+\infty}\int_0^{+\infty}\frac{n}{x(1+x^2)}\sin\frac{x}{n}\,dx.$$

Exercise 6.3.4 ($**$)**.** *For which $n \in \mathbb{N}$ we have $f_n(x) := \frac{1+nx^2}{(1+x^2)^n} \in L^1([0,+\infty[)$? Compute*

$$\lim_{n\to+\infty} \int_0^{+\infty} f_n(x)\, dx.$$

Exercise 6.3.5 ($**$)**.** *Let $f \in L^1([0,+\infty[)$. Prove that*

$$\lim_{\lambda\to+\infty} \int_0^{+\infty} f(x)e^{-\lambda x}\, dx = 0.$$

Exercise 6.3.6 ($**$)**.** *Let $(X, \mathscr{F}, \mu)$ be a measured space and $f \in L(X)$ and $f \geqslant 0$. Let μ_f the measure*

$$\mu_f(E) := \int_E f\, d\mu.$$

The goal is to prove that

$$(\star) \quad \int_X g\, d\mu_f = \int_X gf\, d\mu,\ \forall g \in L^1(X,\nu).$$

- *check $(\star)$ for s simple and positive.*
- *extend $(\star)$ to $g \in L(X,\mu_f)$, $g \geqslant 0$ (use monotone convergence and Prop. 4.1).*
- *extend $(\star)$ to every $g \in L^1(X,\mu_f)$.*

Exercise 6.3.7 ($***$)**.** *Let $(X, \mathscr{F}, \mu)$ be a finite measure space. Show that*

$$f \in L^1,\ \iff\ \sum_n n\mu(n \leqslant |f| < n+1) < +\infty.$$

What happens if μ is not finite?

Exercise 6.3.8 ($***$)**.** *Let $(X, \mathscr{F}, \mu)$ a measure space. Suppose that $(f_n) \subset L(X)$ is such that*

$$\int_X |f_n|\, d\mu \leqslant \frac{C}{n^\alpha},$$

for some C and α constant, $\alpha > 1$. Prove that $f_n \xrightarrow{a.e.} 0$. (hint: use monotone convergence for series. . .)

Exercise 6.3.9 ($***$)**.** *Let $f \in L^1(X)$, $f \geqslant 0$, $(X, \mathscr{F}, \mu)$ measure space. Prove the following continuity property:*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ :\ \int_E f\, d\mu \leqslant \varepsilon,\ \forall E \in \mathscr{F},\ :\ \mu(E) \leqslant \delta.$$

(hint: start searching for $M > 0$ such that $\int_{f\geqslant M} f\, d\mu \leqslant \frac{\varepsilon}{2}$. . .)

LECTURE 7

Integrals depending on parameters

In several applications, we need to discuss how an integral depends on some parameter. Formally, let

$$F(\xi) := \int_E f(x,\xi)\, d\mu(x),\ \xi \in D \subset \mathbb{R}^k.$$

Here, the integral is respect to $x-$variable. We may expect that, under suitable assumptions on the dependence of f on ξ, integral I will be continuous, differentiable etc. Exploring this is the scope of this Lecture.

7.1. Continuity

Let $(X, \mathscr{F}, \mu)$ be a measure space, $f = f(x,\xi)$ be defined for $x \in E \subset X$ and $\xi \in D \subset \mathbb{R}^m$. We assume that

$$\boxed{f(\sharp,\xi) \in L^1(E),\ \forall \xi \in D}$$

In this way, the function

$$F(\xi) := \int_E f(x,\xi)\, d\mu(x),$$

is well defined for every $\xi \in D$. In this section we discuss its continuity.

Theorem 7.1.1

Let $(X, \mathscr{F}, \mu)$ be a measure space and $f : E \times D \longrightarrow \mathbb{R}$, $D \subset \mathbb{R}^d$. Assume that

i) $f(\sharp,\xi) \in L^1(E), \forall \xi \in D$.

ii) $f(x,\sharp) \in \mathscr{C}(D)$ a.e. $x \in E$.

iii) $\exists g \in L^1(E)$ such that $|f(x,\xi)| \leqslant g(x)$ for every $\xi \in D$, a.e. $x \in E$.

Then $F(\xi) := \int_E f(x,\xi)\, dx \in \mathscr{C}(D)$.

Proof. First, by i), F is well defined for $\xi \in D$. To prove continuity at point $\xi \in D$, we have to check that

$$\forall (\xi_n) \subset D,\ :\ \xi_n \longrightarrow \xi,\ \implies\ F(\xi_n) \longrightarrow F(\xi).$$

Notice that

$$F(\xi_n) = \int_E f(x,\xi_n)\, d\mu =: \int_E f_n(x)\, d\mu,\ \text{ where } f_n(x) := f(x,\xi_n).$$

The idea is to apply dominated convergence to (f_n). We have:

- by i) $f_n(x) = f(x,\xi_n) \longrightarrow f(x,\xi)$ a.e. $x \in D$;
- by ii) $|f_n(x)| = |f(x,\xi_n)| \leqslant g(x)$, a.e. $x \in D$.

Thus, according to Lebesgue's dominated convergence, we have

$$\lim_n F(\xi_n) = \lim_n \int_E f_n \, d\mu = \int_E \lim_n f_n \, d\mu = \int_E f(x,\xi) \, d\mu = F(\xi).$$

7.2. Differentiability

Here we provide a powerful sufficient condition under which the integral function

$$F(\xi) := \int_E f(x,\xi) \, d\mu(x)$$

be differentiable.

Theorem 7.2.1: Differentiability under the integral sign

Let $(X, \mathscr{F}, \mu)$ be a measure space, $f : E \times D \longrightarrow \mathbb{R}$, $D \subset \mathbb{R}^d$. Assume that

i) $f(\sharp, \xi) \in L^1(E)$, $\forall \xi \in D$.

ii) $\exists \partial_\xi f(x,\xi)$ for all $\xi \in D$ and a.e. $x \in E$;

iii) $\exists g \in L^1(E)$ such that $|\partial_\xi f(x,\xi)| \leqslant g(x)$ for every $\xi \in D$, a.e. $x \in E$.

Then

$$\exists \partial_\xi F(\xi) = \int_E \partial_\xi f(x,\xi) \, d\mu, \ \forall \xi \in D.$$

Proof. For simplicity on notations we consider $D \subset \mathbb{R}$. i) ensures that F is well defined for $\xi \in D$. Let's compute

$$\partial_\xi F(\xi) = \lim_{h \to 0} \frac{F(\xi + h) - F(\xi)}{h}.$$

Since limit $\lim_{h\to 0}$ can be computed sequentially, we take an arbitrary $(h_n) \subset \mathbb{R}\backslash\{0\}$, $h_n \longrightarrow 0$ and notice that

$$\frac{F(\xi + h_n) - F(\xi)}{h_n} = \int_E \frac{f(x, \xi + h_n) - f(x,\xi)}{h_n} \, d\mu =: \int_E f_n(x) \, d\mu.$$

Now, by ii) it follows that

$$f_n(x) = \frac{f(x, \xi + h_n) - f(x,\xi)}{h_n} \longrightarrow \partial_\xi f(x,\xi), \text{ a.e. } x \in E.$$

The difficult part is to find an integrable dominant for f_n. To this aim first notice that by Lagrange thm there exists η_n such that

$$|f(x, \xi + h_n) - f(x,\xi)| = |\partial_\xi f(x, \eta_n) h_n| \overset{ii)}{\leqslant} g(x)|h_n|, \text{ a.e. } x \in E,$$

thus

$$|f_n(x)| \leqslant g(x), \text{ a.e. } x \in E.$$

Therefore, by dominated convergence

$$\frac{F(\xi + h_n) - F(\xi)}{h_n} = \int_X f_n(x) \, d\mu \longrightarrow \int_X \partial_\xi f(x,\xi) \, d\mu.$$

Let's see a beautiful application of this result that will be important for the future:

Example 7.2.2:

Let

$$F(\xi) := \int_{\mathbb{R}} e^{-\frac{x^2}{2\sigma^2}} e^{-i\xi x}\, dx,\ \xi \in \mathbb{R}.$$

Show that $\partial_\xi F(\xi)$ is well defined for any $\xi \in \mathbb{R}$. Deduce a differential equation for F, solve it and show that

$$F(\xi) = \sqrt{2\pi\sigma^2} e^{-\frac{1}{2}\sigma^2\xi^2},\ \forall \xi \in \mathbb{R}. \tag{7.2.1}$$

Proof. Let $f(x,\xi) := e^{-\frac{x^2}{2\sigma^2}} e^{-i\xi x}$. The integral is well defined because

$$\int_{\mathbb{R}} \left| e^{-\frac{x^2}{2\sigma^2}} e^{-i\xi x} \right| dx = \int_{\mathbb{R}} e^{-\frac{x^2}{2\sigma^2}}\, dx = \sqrt{2\pi\sigma^2} < +\infty.$$

In other words, $f(\cdot,\xi) \in L^1(\mathbb{R})$ for every $\xi \in \mathbb{R}$, that is i) of differentiation thm 7.2 is fulfilled. Notice also that

ii) $\partial_\xi f(x,\xi) = (-ix)e^{-i\xi x} e^{-\frac{x^2}{2\sigma^2}}$, $\forall \xi \in \mathbb{R}$, $\forall x \in \mathbb{R}$;

iii) $|\partial_\xi f(x,\xi)| = \left|(-ix)e^{-i\xi x} e^{-\frac{x^2}{2\sigma^2}}\right| = |x| e^{-\frac{x^2}{2\sigma^2}} =: g(x) \in L^1(\mathbb{R})$, $\forall \xi \in \mathbb{R}$.

Hence, according to differentiation theorem,

$$\begin{aligned}
\partial_\xi F(\xi) &= -i \int_{\mathbb{R}} e^{-i\xi x} \left(x e^{-\frac{x^2}{2\sigma^2}} \right) dx = i\sigma^2 \int_{\mathbb{R}} e^{-i\xi x} \partial_x \left(e^{-\frac{x^2}{2\sigma^2}} \right) dx \\
&= i\sigma^2 \left(\left[e^{-i\xi x} e^{-\frac{x^2}{2\sigma^2}} \right]_{x=-\infty}^{x=+\infty} - \int_{\mathbb{R}} \partial_x \left(e^{-i\xi x} \right) e^{-\frac{x^2}{2\sigma^2}}\, dx \right) \\
&= i\sigma^2 \int_{\mathbb{R}} i\xi e^{-i\xi x} e^{-\frac{x^2}{2\sigma^2}}\, dx = -\sigma^2 \xi \int_{\mathbb{R}} e^{-i\xi x} e^{-\frac{x^2}{2\sigma^2}}\, dx = -\sigma^2 \xi F(\xi).
\end{aligned}$$

Therefore

$$\partial_\xi F(\xi) = -\sigma^2 \xi F(\xi), \implies F(\xi) = e^{-\frac{1}{2}\sigma^2\xi^2} F(0).$$

Finally, since $F(0) = \int_{\mathbb{R}} e^{-\frac{x^2}{2\sigma^2}}\, dx = \sqrt{2\pi\sigma^2}$ we conclude.

7.3. Exercises

Exercise 7.3.1 ($**$)**.** *Let*

$$F(x) := \int_0^{+\infty} e^{-y} \frac{\sin(xy)}{y}\, dy.$$

Show that F is well defined for any $x \in \mathbb{R}$, compute F' and determine F.

Exercise 7.3.2 ($**$)**.** *Let*

$$F(x) := \int_0^{+\infty} \frac{e^{-xt} - e^{-t}}{t}\, dt.$$

Show that $F(x)$ is well defined for any $x > 0$, is differentiable and compute F', hence deduce F.

Exercise 7.3.3 ($**$)**.** *Define*

$$F(\xi) := \int_0^1 \frac{x^\xi - 1}{\log x}\, dx.$$

Show that F is well defined for any $\xi \geqslant 0$. Show that $\exists \partial_\xi F$. Use this to find out F.

Exercise 7.3.4 ($**$)**.** *Let $(X, \mathscr{F}, \mu)$ be a finite measure space (that is a $\mu(X) < +\infty$) and let $E \in \mathscr{F}$. Show that*

$$\min_{y \in \mathbb{R}} \int_X (1_E(x) - y)^2\, d\mu$$

exists and find it.

Exercise 7.3.5 ($**$)**.** *Compute, for $a > 0$ and $b > 0$:*

$$\int_0^{+\infty} \left(e^{-\frac{a^2}{x^2}} - e^{-\frac{b^2}{x^2}} \right) dx.$$

Exercise 7.3.6 ($**$)**.** *Consider*

$$F(x) := \int_0^\pi \frac{\log(1 + x \cos y)}{\cos y}\, dy.$$

i) *Determine the domain of definition for F.*
ii) *Compute $F'(x)$ (where defined). Deduce $F(x)$.*

Exercise 7.3.7 ($**+$)**.** *Evaluate the integral*

$$\int_0^1 \frac{\log(1+x)}{1+x^2}\, dx$$

by using the parametric integral $F(\xi) := \int_0^1 \frac{\log(1+\xi x)}{1+x^2}\, dx$.

Exercise 7.3.8 ($**+$)**.** *Consider the function*

$$F(x) := \int_0^{+\infty} e^{-t-\frac{x}{t}}\, \frac{dt}{\sqrt{t}}.$$

i) *Determine the set of $x \in \mathbb{R}$ for which F is well defined.*
ii) *Discuss differentiability of F on its domain, and deduce a differential equation for F.*
iii) *Determine F explicitly.*

Exercise 7.3.9 ($**+$)**.** *Let*

$$F_n(\alpha) := \int_0^1 x^\alpha (\log x)^n\, dx.$$

Determine values of α for which $F(\alpha)$ is well defined and differentiable, compute F_n' and deduce a differential equation for F_n. Use this to explicitly determine F_n.

LECTURE 8

Normed Spaces

In many problems, the natural framework is a *vector space* equipped with a way to measure distance between vectors. This is needed to define limits and discuss convergence of sequence of vectors. The most natural way to measure distance is through a tool extending the concept of *modulus* for real numbers. This is called *norm* and it is the focus of this Lecture.

8.1. Definition of norm and examples

We recall that V is a **vector space** on $\mathbb{R}$ or $\mathbb{C}$ (field of **scalars**) if a **sum** $f+g \in V$ is defined for every $f, g \in V$ and a **product by scalars** $\alpha f \in V$ is defined, for every $\alpha \in \mathbb{R}$ (or $\mathbb{C}$) and $f \in V$. Sum and product by scalars verify a number of natural properties as:

i) (sum commutative) $f+g=g+f$, $\forall f, g \in V$;
ii) (sum is associative) $f+(g+h)=(f+g)+h$, $\forall f, g, h \in V$;
iii) (sum has zero) $\exists 0_V \in V$ such that $f+0_V=f$, $\forall f \in V$;
iv) (sum has opposite) $\forall f \in V$ there exists $g \in V$ such that $f+g=0_V$ (notation: $-f:=g$);
v) (product is associative) $(\alpha\beta)f=\alpha(\beta f)$, $\forall \alpha, \beta \in \mathbb{R}$ (or $\mathbb{C}$), $\forall f \in V$;
vi) (unit) $1f=f$, $\forall f \in V$;
vii) (distributivity) $(\alpha+\beta)f=\alpha f+\beta f$, $\alpha(f+g)=\alpha f+\beta g$, $\forall \alpha, \beta \in \mathbb{R}$ (or $\mathbb{C}$), $\forall f, g \in V$.

Definition 8.1.1

Let V be a vector space (on $\mathbb{R}$ or $\mathbb{C}$). A function $\|\cdot\| : V \longrightarrow [0,+\infty[$ is called **norm** on V if the following properties hold:

i) (vanishing) $\|f\|=0$ iff $f=0$;
ii) (homogeneity) $\|\alpha f\|=|\alpha|\|f\|$, $\forall \alpha \in \mathbb{R}$ (or $\mathbb{C}$), $\forall f \in V$;
iii) (triangular inequality) $\|f+g\| \leqslant \|f\|+\|g\|$, $\forall f, g \in V$.

We say that $(V, \|\cdot\|)$ is a **normed space**.

The remaining of this Section is devoted to show several important examples.

8.1.1. $\mathbb{R}^d$.

Clearly, $V=\mathbb{R}^d$ is a vector space on $\mathbb{R}$ with usual sum and product by scalars:

$$(x_1,\dots,x_d)+(y_1,\dots,y_d):=(x_1+y_1,\dots,x_d+y_d), \quad \alpha(x_1,\dots,x_d)=(\alpha x_1,\dots,\alpha x_d).$$

The most natural definition of norm is suggested by Euclidean geometry and, in particular, by Pythagorean theorem. It is the so called *euclidean norm*,

$$\|(x_1,\dots,x_d)\|_2 := \sqrt{\sum_{k=1}^d x_k^2}.$$

The check that this is a norm is non trivial. Vanishing and homogeneity are straightforward. The difficult part is the triangular inequality. Let $x = (x_1,\dots,x_d)$, $y = (y_1,\dots,y_d)$. Then

$$\|x+y\|_2^2 = \sum_k (x_k+y_k)^2 = \sum_k x_k^2 + \sum_k y_k^2 + 2\sum_k x_k y_k = \|x\|_2^2 + \|y\|_2^2 + 2\sum_k x_k y_k.$$

Now, to have $\|x+y\|_2 \leqslant \|x\|_2 + \|y\|_2$ we need the celebrate

Lemma 8.1.2: Cauchy-Schwarz inequality

(8.1.1)
$$\sum_k x_k y_k \leqslant \|x\|_2 \|y\|_2.$$

Identity holds true provided $x \propto y$.

PROOF. Conclusion is evident if $\|x\|_2 = 0$ or $\|y\|_2 = 0$ (it reduces to $0 \leqslant 0$). Assume $\|x\|_2, \|y\|_2 \neq 0$. Thus (8.1.1) is equivalent to

$$\sum_k \frac{x_k}{\|x\|_2}\frac{y_k}{\|y\|_2} \leqslant 1.$$

Now, since we have $ab \leqslant \frac{a^2+b^2}{2}$ (this comes from $(a-b)^2 \geqslant 0$), we have also

$$\sum_k \frac{x_k}{\|x\|_2}\frac{y_k}{\|y\|_2} \leqslant \frac{1}{2}\sum_k \left(\frac{x_k^2}{\|x\|_2^2} + \frac{y_k^2}{\|y\|_2^2} \right) = 1.$$

This proves (8.1.1). To finish the proof, we notice that $=$ holds provided $=$ hold in the elementary inequality $ab \leqslant \frac{a^2+b^2}{2}$, that is $a = b$, thus $\frac{x_k}{\|x\|_2} = \frac{y_k}{\|y\|_2}$ for all k, but this means $x \propto y$. □

On $\mathbb{R}^d$ other natural norms are defined, as, for instance,

$$\|(x_1,\dots,x_d)\|_\infty := \max_k |x_k|, \quad \|(x_1,\dots,x_d)\|_1 := \sum_{k=1}^d |x_k|.$$

Proofs are left in the exercises.

8.1.2. $\mathbb{C}^d$. As $\mathbb{R}^d$,$\mathbb{C}^d$ is the vector space of $d-$ples of complex numbers $(z_1,\dots,z_d)$. Sum and product by scalars are defined in the same way as for $\mathbb{R}^d$. We notice that in this case the field of scalars can be both $\mathbb{R}$ as well as $\mathbb{C}$. On $\mathbb{C}^d$ we may define similar norms as for $\mathbb{R}^d$:

$$\|(z_1,\dots,z_d)\|_1 := \sum_k |z_k|, \quad \|(z_1,\dots,z_d)\|_2 := \sqrt{\sum_k |z_k|^2}, \quad \|(z_1,\dots,z_d)\|_\infty := \max_k |z_k|.$$

The checks are left as exercise.

8.1.3. Uniform Norm. Let X be a generic set and set

$$B(X) := \left\{ f : X \longrightarrow \mathbb{R} \ : \ \|f\|_\infty := \sup_{x\in X} |f(x)| < +\infty \right\},$$

the set of all *real valued bounded functions* with usual sum of functions and product of a function by a scalar, that is

$$(f+g)(x) := f(x) + g(x),\ x \in X, \quad (\alpha f)(x) := \alpha f(x),\ x \in X,\ \alpha \in \mathbb{R}.$$

Here, we may consider also the case of $\mathbb{C}$ valued functions, with scalars $\mathbb{R}$ as well as $\mathbb{C}$. For sake of simplicity we will limit to the case of real valued functions.

Proposition 8.1.3

$(B(X), \|\cdot\|_\infty)$ is a normed space.

Proof. We start by checking that $B(X)$ is a vector space. Indeed if f, g are bounded, then clearly $f+g$ is bounded as well, and similarly for αf. Let's prove that $\|\cdot\|_\infty$ is a norm. We have:

- vanishing: $\|f\|_\infty = 0$ iff $\sup_{x\in X} |f(x)| = 0$, that is $|f(x)| \leqslant 0$ for all $x \in X$, but then $f \equiv 0$, which is the zero of $B(X)$;
- homogeneity: $\|\alpha f\|_\infty = \sup_{x\in X} |\alpha f(x)| = \sup_{x\in X} |\alpha||f(x)| = |\alpha| \sup_{x\in X} |f(x)| = |\alpha|\|f\|_\infty$ (by the way, this proves also that $\alpha f \in B(X)$ once $f \in B(X)$);
- triangular inequality: first notice that

$$|f(x)+g(x)| \leqslant |f(x)| + |g(x)| \leqslant \|f\|_\infty + \|g\|_\infty,\ \forall x \in X, \implies \|f+g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty.$$

This proves also that if $f, g \in B(X)$ also $f + g \in B(X)$.

Let $D \subset \mathbb{R}^d$. An important subset of $B(D)$ is that of *continuous and bounded functions on* D:

$$\mathscr{C}_b(D) := \{ f \in B(D) \ : \ f \in \mathscr{C}(D) \}.$$

Since sum and product by scalars of continuous functions are continuous, $\mathscr{C}_b(D)$ is itself a vector space. Equipping $\mathscr{C}_b(D)$ with $\|\cdot\|_\infty$ norm makes $(\mathscr{C}_b(D), \|\cdot\|_\infty)$ itself a normed space. In particular, if $D \subset \mathbb{R}^d$ is **compact** (that is, closed and bounded), then, according to Weierstrass theorem, any $f \in \mathscr{C}(D)$ is bounded. Thus

$$\mathscr{C}_b(D) \equiv \mathscr{C}(D),\ (D \text{ compact}).$$

Moreover, still by Weierstrass theorem, since $|f| \in \mathscr{C}(D)$, $|f|$ itself has maximum on D. This means that,

$$\|f\|_\infty = \sup_{x\in D} |f(x)| = \max_{x\in D} |f(x)|.$$

In other words, when D is compact, we may use $\max_D |f|$ as definition of $\|\cdot\|_\infty$ norm. As for the euclidean norm, other norms are possible on $\mathscr{C}(D)$. Here we illustrate an example of these.

Example 8.1.4: $(**)$

Let $V := \mathscr{C}([a,b])$ and set
$$\|f\|_1 := \int_a^b |f(x)|\, dx.$$
Then, $\|\cdot\|_1$ is a well defined norm on V.

PROOF. We check first that $\|\cdot\|_1$ is well defined: if $f \in V = \mathscr{C}([a,b])$, then $|f| \in \mathscr{C}([a,b])$, thus $|f|$ is integrable (even in Riemann sense). To check that it verifies the characteristic properties we start with vanishing. Suppose
$$\|f\|_1 = \int_a^b |f(x)|\, dx = 0.$$
Now, since $|f|$ is continuous we claim that the previous is possible iff $|f(x)| \equiv 0$. Indeed, if $|f(x_0)| > 0$ for some $x_0 \in [a,b]$ then, by continuity, $|f(x)| \geqslant \frac{|f(x_0)|}{2}$ for $x \in I_{x_0} \subset [a,b]$, neighbourhood of x_0. Thus
$$0 = \|f\|_1 = \int_a^b |f(x)|\, dx \geqslant \int_{I_{x_0}} |f(x)|\, dx \geqslant \frac{|f(x_0)|}{2}\lambda_1(I_{x_0}) > 0.$$
Thus, we get a contradiction and vanishing holds. Homogeneity and triangular inequality are straightforward. Indeed,
$$\|\alpha f\|_1 = \int_a^b |\alpha f(x)|\, dx = \int_a^b |\alpha||f(x)|\, dx = |\alpha|\|f\|_1,$$
and
$$\|f+g\|_1 = \int_a^b |f(x)+g(x)|\, dx \leqslant \int_a^b |f(x)| + |g(x)|\, dx = \|f\|_1 + \|g\|_1.$$

8.2. Norm comparison

As we have seen, in some cases, several norms can be defined on the same vector space V. It is important to have a way to compare two norms.

Definition 8.2.1

Let $\|\cdot\|$ and $\|\cdot\|_*$ be norms on V. We say that $\|\cdot\|$ **is stronger than** $\|\cdot\|_*$ if
$$\exists C > 0, : \ \|f\|_* \leqslant C\|f\|, \ \forall f \in V.$$
If each of the two is stronger than the other, we say that the two norm are **equivalent**.

Example 8.2.2: $(**)$

Norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ on $\mathbb{R}^d$ are equivalent.

PROOF. We will prove that
$$\|x\|_2 \leqslant C\|x\|_1 \leqslant C'\|x\|_\infty \leqslant C''\|x\|_2, \ \forall x \in \mathbb{R}^d,$$

from which the conclusion follows. Noticed that $|x_k| \leqslant \|x\|_1$ for every k,

$$\|x\|_2 = \left(\sum_k |x_k|^2 \right)^{1/2} \leqslant \left(\sum_k |x_k| \|x\|_1 \right)^{1/2} = \|x\|_1,$$

thus $C = 1$. Now, since $|x_k| \leqslant \|x\|_\infty$ for every k

$$\|x\|_1 = \sum_k |x_k| \leqslant d \|x\|_\infty,$$

thus $C' = d$. Finally, since also $|x_k| \leqslant \|x\|_2$ for every k, we have

$$\|x\|_\infty = \max_k |x_k| \leqslant \|x\|_2,$$

thus $C'' = d$.

This result is not incidental:

Theorem 8.2.3

If V is finite dimensional, then any two norms are equivalent.

Proof. We do the proof on $V = \mathbb{R}^d$ (but the same proof can be easily adapted to a generic vector space). Let $\| \cdot \|$ be a generic norm. We prove that $\| \cdot \|$ is equivalent to the Euclidean norm $\| \cdot \|_2$.
To this aim notice that, denoting by (e_k) the standard basis of $\mathbb{R}^d$,

$$\|x\| = \left\| \sum_k x_k e_k \right\| \leqslant \sum_k |x_k| \|e_k\| \leqslant \left(\max_k \|e_k\| \right) \sum_k |x_k| =: C \|x\|_1 \leqslant C d \|x\|_2.$$

This proves $\| \cdot \|_2$ is stronger than $\| \cdot \|$. To prove the vice versa, we need to prove

$$\exists C' \ : \ \|x\|_2 \leqslant C' \|x\|, \ \forall x \in \mathbb{R}^d.$$

Notice that, if this is true, then

$$\|u\| \geqslant \frac{1}{C'} =: C'' > 0, \ \forall u \in \mathbb{R}^d, \ \|u\|_2 = 1. \tag{8.2.1}$$

Vice versa, if this last is true, setting $u = \frac{x}{\|x\|_2}$, clearly $\|u\|_2 = 1$, then

$$\frac{1}{C'} \leqslant \left\| \frac{x}{\|x\|_2} \right\| = \frac{1}{\|x\|_2} \|x\|, \iff \|x\|_2 \leqslant C' \|x\|.$$

Thus, (8.2.1) is equivalent to the conclusion. Notice that (8.2.1) means that the function $T(u) := \|u\|$ has a positive lower bound on the surface of the sphere $\mathbb{S} := \{u \in \mathbb{R}^d \ : \ \|u\|_2 = 1\}$. To prove this we first prove that T *in continuous as function* $\mathbb{R}^d \longrightarrow \mathbb{R}$ *respect to the* $\| \cdot \|_2$ *norm*. Indeed,

$$|T(u) - T(v)| = |\|u\| - \|v\|| \leqslant \|u - v\| \leqslant C d \|u - v\|_2.$$

Since $\mathbb{S}$ is compact, by Weierstrass thm T has a minimum. Let $u^* \in \mathbb{S}$ be a minimum point: Then $T(u) = \|u\| \geqslant T(u^*) = \|u^*\| > 0$ (otherwise $\|u^*\| = 0$ would imply $u^* = 0$, but $u^* \in \mathbb{S}$ cannot be $= 0$). This completes the proof.

The previous fact is no longer true when the space is infinite dimensional, as the following example shows.

Example 8.2.4: $(**)$

On $V = \mathscr{C}([0,1])$, let us consider uniform norm $\|f\|_\infty := \sup_{[0,1]} |f(x)| \equiv \max_{[0,1]} |f(x)|$ and the norm

$$\|f\|_1 := \int_0^1 |f(x)|\,dx.$$

Then $\|\cdot\|_\infty$ is stronger than $\|\cdot\|_1$ but they are not equivalent.

Proof. We first prove that $\|\cdot\|_\infty$ is stronger than $\|\cdot\|_1$. This is easy,

$$\|f\|_1 = \int_0^1 |f(x)|\,dx \leqslant \int_0^1 \|f\|_\infty\,dx = \|f\|_\infty,\ \forall f \in V.$$

To prove that $\|\cdot\|_1$ is not stronger than $\|\cdot\|_\infty$ we need to prove that it does not exist C such that

$$\|f\|_\infty \leqslant C\|f\|_1,\ \forall f \in V.\ (\star)$$

To have $(\star)$ true, it means that if the area "under" f (the integral $\int_0^1 |f|$) is small, then, necessarily, $\max|f|$ must be small, so $|f|$ must be uniformly small. This seems to be false: we can have arbitrarily large functions with small area instead. To formalize this example, define

$$f_n(x) := \begin{cases} n - n^3 x, & 0 \leqslant x \leqslant \frac{1}{n^2}, \\ 0 & \frac{1}{n^2} \leqslant x \leqslant 1. \end{cases} \tag{8.2.2}$$

Notice that

$$\|f_n\|_1 = \int_0^1 |f_n(x)|\,dx = \frac{1}{2} n \cdot \frac{1}{n^2} = \frac{1}{2n}, \quad \|f_n\|_\infty = \sup_{x\in[0,1]} |f_n(x) - f(x)| = n,$$

thus, if $(\star)$ were true, we should have $n \leqslant C\frac{1}{2n}$, that is $2n^2 \leqslant C$ for every $n \in \mathbb{N}$. But this is impossible, thus $(\star)$ is false.

Definition 8.2.5

We say that $(V, \|\cdot\|_V)$ is **embedded** into $(W, \|\cdot\|_W)$ if

$$V \subset W, \text{ and } \|f\|_W \leqslant C\|f\|_V,\ \forall f \in V.$$

We write $(V, \|\cdot\|_V) \hookrightarrow (W, \|\cdot\|_W)$.

Example 8.2.6: $(**)$

Let $V = \mathscr{C}^1([0,1])$ equipped with $\|f\|_V := |f(0)| + \|f'\|_\infty$ and $W := \mathscr{C}([0,1])$ equipped with $\|f\|_W := \|f\|_\infty$. Check that $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are normed spaces and that $(V, \|\cdot\|_V) \hookrightarrow (W, \|\cdot\|_W)$.

Proof. We already know that $(W, \|\cdot\|_W)$ is a normed space. About V: easily, it is a vector space with usual sum and product by scalars. Let's check that $\|\cdot\|_V$ is a norm. Clearly, since $f \in V$ means $f \in \mathscr{C}^1$,

$\|f\|_V$ is well defined. Let's check the characteristic properties. We have

$$\|f\|_V = 0, \iff |f(0)| + \|f'\|_\infty = 0, \iff \begin{cases} |f(0)| = 0, \\ \\ \|f'\|_\infty = 0. \end{cases}$$

The first says that $f(0) = 0$. The second, because of vanishing for $\|\cdot\|_\infty$ norm, tells $f' \equiv 0$. But then, $f \equiv C$ and since $f(0) = 0$ we conclude $f \equiv 0$, that is vanishing holds. Homogeneity and triangular inequality are straightforward, we leave as exercise.

Let us come to the embedding. Clearly $V = \mathscr{C}^1([0,1]) \subset W = \mathscr{C}([0,1])$. To show that this inclusion is an embedding between the two spaces we need to check that there exists $C > 0$ such that

$$\|f\|_W \leqslant C\|f\|_V, \ \forall f \in V, \iff \|f\|_\infty \leqslant C\left(|f(0)| + \|f'\|_\infty\right), \ \forall f \in C^1([0,1]).$$

We need a way to express $f(x)$ in terms of $f'(x)$. This is provided by the fundamental formula of Integral Calculus, according which we have

$$f(x) - f(0) = \int_0^x f'(y)\, dy, \iff f(x) = f(0) + \int_0^x f'(y)\, dy.$$

Then

$$\begin{aligned} |f(x)| &\leqslant |f(0)| + \left|\int_0^x f'(y)\, dy\right| \overset{\triangle}{\leqslant} |f(0)| + \int_0^x |f'(y)|\, dy \leqslant |f(0)| + \|f'\|_\infty \int_0^x 1\, dy \\ &\leqslant |f(0)| + x\|f'\|_\infty, \ \forall x \in [0,1]. \end{aligned}$$

Therefore

$$\|f\|_W = \max_{[0,1]} |f(x)| \leqslant \max_{[0,1]} \left(|f(0)| + x\|f'\|_\infty\right) = |f(0)| + \|f'\|_\infty = \|f\|_V,$$

this being true for every $f \in V$. The embedding is now proved.

8.3. Exercises

Exercise 8.3.1 ($*$). *On* $\mathbb{R}^2$ *define*

$$\|(x,y)\|_* := (\sqrt{|x|} + \sqrt{|y|})^2.$$

Is $\|\cdot\|_*$ *a norm?*

Exercise 8.3.2 ($*$). *Check that* $\|\cdot\|_1$ *and* $\|\cdot\|_\infty$ *are norms for* $\mathbb{R}^d$.

Exercise 8.3.3 ($**$). *On* $V := \mathscr{C}^1([0,1])$ *we consider* $\|\cdot\|_\infty$ *norm and* $\|f\|_V := \|f\|_\infty + \|f'\|_\infty$. *Check that* $\|\cdot\|_V$ *is stronger than* $\|\cdot\|_\infty$ *norm, but they are not equivalent.*

Exercise 8.3.4 ($**$). *On* $V = \mathscr{C}^1([0,1])$ *consider the sup-norm* $\|\cdot\|_\infty$, *the total variation norm*

$$\|f\|_v := \|f\|_\infty + \|f'\|_1 \equiv \|f\|_\infty + \int_0^1 |f'(x)| dx,$$

and the $\mathscr{C}^1$ *norm* $\|f\|_* := \|f\|_\infty + \|f'\|_\infty$.

i) *Show that there exist* $c, C > 0$ *such that* $\|f\|_\infty \leqslant c\|f\|_v \leqslant C\|f\|_*$.
ii) *By using* $f_k(x) := c_k \sin(k\pi x)$ *or* $g_k(x) = c_k x^k$, *show that do not exist constants* $m, M > 0$ *such that* $\|f\|_v \leqslant m\|f\|_\infty$ *and* $\|f\|_* \leqslant M\|f\|_v$.

Exercise 8.3.5 $(**)$. *Let* $V := \{f \in \mathscr{C}^1([a,b]) \;:\; f(a) = 0\}$ *equipped with* $\|f\|_V := \int_a^b |f'(x)|\,dx$. *i) Check that* $\|\cdot\|_V$ *is a norm on* V. *ii) Prove that* $(V, \|\cdot\|_V) \hookrightarrow (\mathscr{C}([a,b]), \|\cdot\|_\infty)$.

Exercise 8.3.6 $(**)$. *Let* $V := \{f \in \mathscr{C}^2([a,b]) \;:\; f(a) = f(b) = 0\}$ *equpped with* $\|f\|_V := \int_a^b |f''(x)|\,dx$. *i) Check that* $\|\cdot\|_V$ *is a norm on* V. *ii) Show that* $(V, \|\cdot\|_V) \hookrightarrow (\mathscr{C}^1([a,b]), \|\cdot\|_{\mathscr{C}^1})$ *where* $\|f\|_{\mathscr{C}^1} := \|f\|_\infty + \|f'\|_\infty$.

Exercise 8.3.7 $(**)$. *Let* $\ell^1 := \{(x_n)_{n\in\mathbb{N}} \subset \mathbb{R} \,:\, \sum_n |x_n| < +\infty\}$ *with natural sum* $(x_n)+(y_n) := (x_n+y_n)$ *and* $\alpha(x_n) := (\alpha x_n)$. *Check that* ℓ^1 *is a vector space. We set*

$$\|(x_n)\|_{\ell^1} := \sum_n |x_n|.$$

Check that $\|\cdot\|_{\ell^1}$ *is a well defined norm on* ℓ^1.

Exercise 8.3.8 $(**+)$. *Let* $\ell^2 := \left\{(x_n)_{n\in\mathbb{N}} \;:\; \sum_n x_n^2 < +\infty\right\}$. *Show that* ℓ^2 *is a vector space with the same algebraic operations defined in the previous exercise for* ℓ^1. *Define then*

$$\|(x_n)\|_{\ell^2} := \sqrt{\sum_k x_k^2}.$$

Prove that $\|\cdot\|_{\ell^2}$ *is a well defined norm on* ℓ^2 *(hint: adapt ideas from the Euclidean norm of* $\mathbb{R}^d$*).*

Exercise 8.3.9 $(**)$. *Let* $V := \mathscr{C}([0,1])$ *equipped with*

$$\|f\|_V := \int_0^1 \frac{|f(x)|}{\sqrt{x}}\,dx.$$

i) *Check that* $\|\cdot\|_V$ *is well defined on* V *and it is a norm.*
ii) *Check that usual* $\|\cdot\|_\infty$ *norm is stronger than* $\|\cdot\|_V$ *norm.*
iii) *Define*

$$f_n(x) := \begin{cases} \sqrt[3]{n}, & 0 \leqslant x \leqslant \frac{1}{n}, \\[2ex] \frac{1}{\sqrt[3]{x}}, & \frac{1}{n} \leqslant x \leqslant 1. \end{cases}$$

Is $(f_n) \subset V$? *Compute* $\|f_n\|_V$. *What can you draw about the relation between* $\|\cdot\|_\infty$ *and* $\|\cdot\|_V$?

LECTURE 9

L^p spaces

Spaces of integrable functions are of paramount relevance in applications. In this Lecture we introduce these spaces with their natural norms. We introduce also the space of *essentially bounded* functions, namely the version of bounded functions for measurable functions.

9.1. L^1 space

L^1 is the space of Lebesgue integrable functions:

Definition 9.1.1

Let $(X, \mathscr{F}, \mu)$ be a measure space. We set

$$L^1(X) := \left\{ f \in L(X) \ : \ \|f\|_1 := \int_X |f| \, d\mu < +\infty \right\}$$

It is easy to check that $L^1(X)$ is a vector space with usual operations of sum and product by a scalar. Indeed, if $f, g \in L^1(X)$ then $f + g \in L^1(X)$ because

$$\int_X |f+g| \, d\mu \leqslant \int_X (|f| + |g|) \, d\mu = \int_X |f| \, d\mu + \int_X |g| \, d\mu < +\infty,$$

and similarly $\alpha f \in L^1(X)$ for every $\alpha \in \mathbb{R}$ (or $\mathbb{C}$). The quantity $\| \cdot \|_1$ seems to be a natural way to define the norm of a vector of $L^1(X)$. And indeed we have already shown that $\| \cdot \|_1$ is a true norm on a vector space like $\mathscr{C}([a,b])$ (see Example 8.2). When we consider the space $L^1(X)$, however, $\| \cdot \|_1$ verifies all the characteristic properties of a norm except for vanishing, which takes a mild form:

Proposition 9.1.2

$\| \cdot \|_1$ verifies homogeneity, triangular inequality and vanishing in the following weak form:

$$\|f\|_1 = 0, \quad \Longleftrightarrow \quad f = 0, \ a.e.$$

PROOF. Homogeneity and triangular inequality are easy and left to the reader. Let us focus on vanishing. It is evident that if $f = 0$ a.e. then $\|f\|_1 = \int_X |f| \, d\mu = 0$. Vice versa: if $\|f\|_1 = \int_X |f| \, d\mu = 0$ then $|f| = 0$ a.e. follows from Čebišëv Lemma 4.2 applied to $|f|$.

Warning 9.1.3

Strictly speaking, $\|\cdot\|_1$ is not a norm. To make it a true norm we should modify the structure of L^1. The idea is the following: we say that *f and g are equivalent is $f = g$ a.e.*. Given $f \in L(X)$, we define $\{f\}$ as the set of all functions equivalent to f. Then, we consider the so-called *quotient space*, that is the set V made of all possible equivalence classes $\{f\}$. These will be vectors of a new space, where

$$\{f\} + \{g\} := \{f + g\}, \;\; c\{f\} := \{cf\}.$$

One can verify that these operations are well posed and give to V the structure of vector space. The zero of the space is, of course, the class $0_V = \{0\}$. Finally, one set

$$\|\{f\}\|_1 := \int_X |f| \, d\mu.$$

Since $f \overset{a.e.}{=} g$ implies $|f| \overset{a.e.}{=} |g|$, thus in particular $\int_X |f| \, d\mu = \int_X |g| \, d\mu$, previous definition is independent of any particular element chosen from the class $\{f\}$. Furthermore

$$\|\{f\}\|_1 = 0, \iff \int_X |f| \, d\mu = 0, \iff |f| \overset{a.e.}{=} 0, \iff f \overset{a.e.}{=} 0, \iff \{f\} = \{0\} = 0_V.$$

So, this quotient space is a true normed space.

However, even if $\|\cdot\|_1$ is not a true norm, we consider such in all respects. The unique care is to remind that vanishing works in a slightly weaker form.

9.2. L^p space $(1 < p < +\infty)$

L^p space is just an extension of L^1 space:

Definition 9.2.1

Let $(X, \mathscr{F}, \mu)$ be a measure space, $1 < p < +\infty$. We set

$$L^p(X) := \left\{ f \in L(X) \; : \; \int_X |f|^p \, d\mu < +\infty \right\}.$$

We define

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p}.$$

For the sake of simplicity, we will work out in detail the fundamental case $p = 2$, assuming real scalars. For the general case we will limit to main statements, leaving the technical proofs in the exercises. The first step is the

Proposition 9.2.2

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then, $L^2(X)$ is a vector space, and $\|\cdot\|_2$ is a norm on $L^2(X)$ with vanishing $\|f\|_2 = 0$ iff $f = 0$ a.e.

PROOF. Let 's start proving that $L^2(X)$ is a vector space. Let $f, g \in L^2(X)$. To prove that $f+g \in L^2(X)$ we have to prove that $\int_X (f+g)^2 \, d\mu < +\infty$. Now, since $2ab \leqslant a^2 + b^2$,

$$(f+g)^2 = f^2 + g^2 + 2fg \leqslant f^2 + g^2 + f^2 + g^2 = 2(f^2 + g^2),$$

thus, integrating,

$$\int_X (f+g)^2 \, d\mu \leqslant 2 \int_X f^2 + g^2 \, d\mu = 2 \left(\int_X f^2 \, d\mu + \int_X g^2 \, d\mu \right) < +\infty.$$

By this $f + g \in L^2(X)$. The proof that $\alpha f \in L^2(X)$ is straightforward.
Let's now move on the properties of $\| \cdot \|_2$ norm. We start by the vanishing:

$$\|f\|_2 = 0, \iff \int_X |f|^2 \, d\mu = 0, \iff |f|^2 = 0, \ a.e., \iff f = 0, \ a.e.$$

Homogeneity is straightforward. Finally, the triangular inequality. This is similar to the proof of the same property for the euclidean norm on $\mathbb{R}^d$. We start computing

$$\|f+g\|_2^2 = \int_X (f+g)^2 \, d\mu = \int_X f^2 \, d\mu + \int_X g^2 \, d\mu + 2 \int_X fg \, d\mu = \|f\|_2^2 + \|g\|_2^2 + 2 \int_X fg \, d\mu.$$

To have $\|f+g\|_2 \leqslant \|f\|_2 + \|g\|_2$ we need the

Lemma 9.2.3: Cauchy–Schwarz inequality

$$\left| \int_X fg \, d\mu \right| \leqslant \|f\|_2 \|g\|_2, \ \forall f, g \in L^2(X). \tag{9.2.1}$$

Moreover, equality holds iff $f \propto g$ a.e.

PROOF. It is similar to the proof of CS (8.1.1) for euclidean norm. We first notice that if $\|f\|_2 = 0$ or $\|g\|_2 = 0$ the inequality is a trivial $0 \leqslant 0$. Thus we assume $\|f\|_2, \|g\|_2 \neq 0$. Dividing by l.h.s., the proof is reduced to

$$\left| \int_X \frac{f}{\|f\|_2} \frac{g}{\|g\|_2} \, d\mu \right| \leqslant 1.$$

Again, by $ab \leqslant \frac{a^2+b^2}{2}$,

$$\left| \int_X \frac{f}{\|f\|_2} \frac{g}{\|g\|_2} \, d\mu \right| \leqslant \int_X \frac{|f|}{\|f\|_2} \frac{|g|}{\|g\|_2} \, d\mu \leqslant \frac{1}{2} \left(\int_X \frac{f^2}{\|f\|_2^2} \, d\mu + \int_X \frac{g^2}{\|g\|_2^2} \, d\mu \right) = 1,$$

and this proves (9.2.1). About equality, all $\leqslant$ signs must be $=$. In particular, the last one implies $\frac{|f|}{\|f\|_2} = \frac{|g|}{\|g\|_2}$ a.e., and from this the conclusion follows.

Returning to the proof of triangular inequality, we have

$$\|f+g\|_2^2 \leqslant \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2 = (\|f\|_2 + \|g\|_2)^2,$$

from which the conclusion follows by taking roots.

For the general case $1 < p < +\infty$ the argument is similar. The key step is the extension of the Cauchy–Schwarz inequality to the important

Lemma 9.2.4: Hölder inequality

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then

$$\left|\int_X fg\, d\mu\right| \leqslant \|f\|_p \|g\|_q,$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

See exercises for the proof. With some non trivial work (left to exercises) it is possible to prove the

Proposition 9.2.5

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then, for every $1 < p < +\infty$, $L^p(X)$ is a vector space and $\|\cdot\|_p$ is a norm on $L^p(X)$ with vanishing in weak form.

9.3. L^∞ space

The concept of bounded function fights with that one measure. For example, measure consider the same any two functions which are a.e. equal. However, if we define, on $X = \mathbb{R}$,

$$f \equiv 0, \quad g = \sum_{n=1}^{\infty} n 1_{\{n\}},$$

we see that $f = g$ a.e. but while f is constant, g is unbounded. In other words, we cannot use

$$\|f\|_\infty = \sup_X |f(x)|$$

to define a norm on $L(X)$. However, from the point of view of measure, it is clear that we should consider g *essentially bounded*.

Definition 9.3.1

Let $(X, \mathscr{F}, \mu)$ be a measure space, $f \in L(X)$. We say that $f$vis **essentially bounded on** X if

$$\exists M > 0, : \ |f| \leqslant M, \ a.e.$$

The class of essentially bounded functions on X is denoted by $L^\infty(X)$.

It is not difficult to prove that

Proposition 9.3.2

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then, $L^\infty(X)$ is a vector space equipped with usual sum and multiplication by scalars.

The proof is left as exercise. We now introduce a suitable norm on $L^\infty(X)$. If $f \in L^\infty(X)$, there exists $M > 0$ such that $|f| \leqslant M$ a.e. We call this M an **essential upper bound**. It is clear that, every $K > M$

is an essential upper bound as well. So, the **essential best upper bound** is the smallest of the essential upper bounds,

$$\|f\|_\infty \equiv \text{ess sup}\,|f| := \inf\{K \geqslant 0 \,:\, |f| \leqslant K\ a.e.\}.$$

Notice that, by definition of $\|f\|_\infty$, we have

$$\forall \varepsilon > 0,\ |f| \leqslant \|f\|_\infty + \varepsilon,\ \text{a.e.}$$

Setting $\varepsilon = \frac{1}{n}$ we have

$$\{|f| > \|f\|_\infty\} = \bigcup_n \left\{|f| > \|f\|_\infty + \frac{1}{n}\right\}, \implies \mu\left(|f| > \|f\|_\infty\right) \leqslant \sum_n \mu\left(|f| > \|f\|_\infty + \frac{1}{n}\right) = 0,$$

that is

(9.3.1) $$\boxed{|f(x)| \leqslant \|f\|_\infty,\ \text{a.e. } x \in X.}$$

Definition 9.3.3

Let $(X, \mathscr{F}, \mu)$ be a measure space, $f \in L(X)$. We set

$$L^\infty(X) := \{f \in L(X) \,:\, \|f\|_\infty < +\infty\}.$$

Proposition 9.3.4

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then, $L^\infty(X)$ is a normed space equipped with $\|\cdot\|_\infty$, with vanishing in weak form.

Proof. Let's verify the characteristic properties of a norm.

- vanishing: if $\|f\|_\infty = 0$ then, by (9.3.1), $|f(x)| \leqslant 0$ a.e., that is $f = 0$ a.e..
- homogeneity: by (9.3.1), $|\alpha f(x)| = |\alpha||f(x)| \leqslant |\alpha|\|f\|_\infty$ a.e.. Therefore $\|\alpha f\|_\infty \leqslant |\alpha|\|f\|_\infty$. Now, since (by the same inequality), $\|f\|_\infty = \|\frac{1}{\alpha}(\alpha f)\|_\infty \leqslant \frac{1}{|\alpha|}\|\alpha f\|_\infty$ we get the conclusion;
- triangular inequality: by (9.3.1),

  $$|f| \leqslant \|f\|_\infty,\ \ |g| \leqslant \|g\|_\infty,\ a.e. \implies |f+g| \leqslant |f| + |g| \leqslant \|f\|_\infty + \|g\|_\infty,\ a.e.$$

  This says that $\|f+g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$. □

9.4. Exercises

Exercise 9.4.1 (∗∗)**.** *Let $\alpha > 0$. Determine for which $p \geqslant 1$ the function*

$$f(x) := \frac{1}{1 + \|x\|^\alpha},\ x \in \mathbb{R}^d,$$

belongs to $L^p(\mathbb{R}^d)$ w.r.t. Lebesgue's measure. (hint: look at Exercise 5.4.7)

Exercise 9.4.2 (∗∗)**.** *The goal is to show that $L^p(X)$ is a vector space. To this aim, show the following numerical inequality:*

$$\exists C > 0,\ :\ (u+v)^p \leqslant C(u^p + v^p),\ \forall u, v \geqslant 0.$$

Use this inequality to conclude.

Exercise 9.4.3 ($***$). *The goal is to prove Hölder inequality.*

i) *By using the concavity of function* $\log t$*, prove the Young inequality*

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q,\ \forall a, b \geqslant 0,$$

with $1 < p, q < +\infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$.

ii) *Himitating the proof of CS inequality and using i), prove the Hölder inequality.*

Exercise 9.4.4 ($***$). *The goal is to prove that* $\|\cdot\|_p$ *is a norm* $(1 < p < +\infty)$. *Vanishing follows from Chebyshev's inequality and homogeneity is straight forward. For the triangular inequality write*

$$\|f+g\|_p^p = \int_X |f+g|^p \, d\mu = \int_X |f+g||f+g|^{p-1} \, d\mu \leqslant \int_X |f||f+g|^{p-1} \, d\mu + \int_X |g||f+g|^{p-1} \, d\mu,$$

then apply Hölder inequality.

Exercise 9.4.5 ($**$). *Let* $(X, \mathscr{F}, \mu)$ *be a measure space. Check that:*

i) *if* $\mu(X) < +\infty$, *then* $\|\cdot\|_2$ *norm is stronger than* $\|\cdot\|_1$ *norm. (use CS inequality).*

ii) *if* $\mu(X) = +\infty$, *in general previous conclusion is false (consider* $X = [0, +\infty[$ *and* $\mu =$*Lebesgue measure...).*

iii) *even if* $\mu(X) < +\infty$ *in general* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *norms are not equivalent (take* $X = [0, 1]$ *with* $\mu =$*Lebesgue measure...).*

iv) *in certain cases* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *norm can be equivalent (take X finite set...).*

Exercise 9.4.6 ($**$). *Let* $f \in L^2(\mathbb{R})$.

i) *Is it true that* $f \in L^1(\mathbb{R})$*? Prove in general, if true, disprove with a counter example, if false.*

ii) *Show that if* $xf(x) \in L^2(\mathbb{R})$ *then, necessarily,* $f \in L^1(\mathbb{R})$, *proving also the bound*

$$\|f\|_1 \leqslant \sqrt{2}\left(\|f\|_2 + \|xf\|_2\right).$$

Exercise 9.4.7 ($**$). *Let* $(X, \mathscr{F}, \mu)$ *a measure space,* $f \in L^p(X)$. *Prove that*

i) $\mu(|f| \geqslant \alpha) \leqslant \frac{1}{\alpha^p}\|f\|_p^p$ *for every* $\alpha > 0$.

ii) $\lim_{\alpha \to +\infty} \alpha^p \mu(|f| \geqslant \alpha) = 0$.

Exercise 9.4.8 ($***$). *Let* $f \in L^2([0, +\infty[)$.

i) *Prove that* $\left(\int_0^x f(y)\, dy\right)^2 \leqslant 2\sqrt{x}\int_0^x \sqrt{y} f(y)^2 \, dy$, $\forall x \geqslant 0$ *(hint: Cauchy-Schwarz's inequality).*

ii) *Define* $g(x) := \frac{1}{x}\int_0^x f(y)\, dy$. *Check that* $g \in L^2([0, +\infty[)$ *and* $\|g\|_2 \leqslant 2\|f\|_2$.

Exercise 9.4.9 ($***$). *Extend Hölder's inequality: let* $f \in L^p$, $g \in L^q$ *and* $h \in L^r$ *with* $1 < p, q, r < +\infty$ *such that* $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. *Prove that*

$$\|fgh\|_1 \leqslant \|f\|_p \|g\|_q \|h\|_r.$$

LECTURE 10

Convergence

As for modulus in the real line, norm allows to define limits of sequences of vectors in a normed space. When the normed space is a space of functions, convergence is almost never an easy matter. This Lecture introduces to this topic through many examples.

10.1. Limit of a sequence

We start with the

Definition 10.1.1

Let $(V, \|\cdot\|)$ be a normed space. Given $(f_n) \subset V$ we say that

$$f_n \xrightarrow{\|\cdot\|} f, \quad \Longleftrightarrow \quad \|f_n - f\| \longrightarrow 0.$$

The first remark is that if a limit exists, it is unique:

Proposition 10.1.2

If $(f_n) \subset V$ has a limit, the limit is unique.

PROOF. If $f_n \xrightarrow{\|\cdot\|} f$ and $f_n \xrightarrow{\|\cdot\|} g$, then

$$\|f - g\| \leqslant \|f - f_n\| + \|f_n - g\| \longrightarrow 0, \quad \Longrightarrow \quad \|f - g\| = 0, \quad \Longleftrightarrow \quad f = g. \qquad \square$$

Remark 10.1.3

As the proof shows, uniqueness of a limit depends on vanishing of the norm. So, if we consider an L^p norm, the same proof leads to the following conclusion:

$$f_n \xrightarrow{\|\cdot\|_p} f, \; f_n \xrightarrow{\|\cdot\|_p} g, \quad \Longrightarrow \quad f \overset{a.e.}{=} g.$$

Thus, in principle (unless the unique measure zero set is $\varnothing$) the limit is not unique. This might be disturbing. However, since any two limits differ by a measure zero set, from the point of view of measure they are the same object, thus, we shouldn't be worried too much!

Example 10.1.4: $(*)$

Let $V = L^1([0,1])$, $f_n(x) := x^n$. Discuss convergence of (f_n) in $\|\cdot\|_1$ norm.

Proof. We first have to guess a potential limit. We may notice that, for $x \in [0,1]$ fixed,

$$\lim_n f_n(x) = \lim_n x^n = \begin{cases} 0, & 0 \leqslant x < 1, \\ 1, & x = 1 \end{cases} =: f(x),$$

thus we may guess that $f_n \xrightarrow{L^1} f$. Now, since $f = 0$ a.e. this means $f_n \xrightarrow{L^1} 0$. Let's check this: we have

$$\|f_n - 0\|_1 = \|f_n\|_1 = \int_0^1 |x^n|\,dx = \int_0^1 x^n\,dx = \left[\frac{x^{n+1}}{n+1}\right]_{x=0}^{x=1} = \frac{1}{n+1} \longrightarrow 0,$$

that is our guess was correct.

Convergence is not an intrinsic property of a sequence, but it always depends on a specific norm. It may well happen that the same sequence under different norms might have different behaviours. However, if convergence happens in a stronger norm, then it happens also for a weaker norm.

Proposition 10.1.5

If $\|\cdot\|$ is stronger than $\|\cdot\|_*$, then any sequence converging under $\|\cdot\|$ converges also under $\|\cdot\|_*$ to the same limit.

Proof. By assumption $\|f\|_* \leqslant C\|f\|$ for every $f \in V$ and for a suitable $C > 0$. If $f_n \xrightarrow{\|\cdot\|} f$ then

$$0 \leqslant \|f_n - f\|_* \leqslant C\|f_n - f\| \longrightarrow 0,$$

thus $\|f_n - f\|_* \longrightarrow 0$, that is $f_n \xrightarrow{\|\cdot\|_*} f$.

Example 10.1.6: $(**)$

On $V = \mathscr{C}([0,1])$ equipped with uniform norm $\|\cdot\|_\infty$ let us consider again the sequence $f_n(x) := x^n$. This sequence is not convergent in uniform norm.

Proof. We already proved that $f_n \xrightarrow{\|\cdot\|_1} 0$. Since

$$\|f\|_1 = \int_0^1 |f(x)|\,dx \leqslant \int_0^1 \|f\|_\infty\,dx = \|f\|_\infty,$$

the uniform norm on V is stronger than the L^1 norm. In particular, if $f_n \xrightarrow{\|\cdot\|_\infty} g \in V$ then also $f_n \xrightarrow{\|\cdot\|_1} g$. Since we already checked that $f_n \xrightarrow{\|\cdot\|_1} 0$, the unique possibility is $f_n \xrightarrow{\|\cdot\|_\infty} 0$. However,

$$\|f_n - 0\|_\infty = \|f_n\|_\infty = \max_{[0,1]} |x^n| = 1 \not\longrightarrow 0.$$

So, the sequence (f_n) cannot converge in uniform norm.

An useful fact to know is the

Proposition 10.1.7

Let $(V, \|\cdot\|)$ be a normed space. Then, every convergent sequence (f_n) is necessarily **bounded**, that is

$$\exists M \ :\ \|f_n\| \leqslant M,\ \forall n \in \mathbb{N}.$$

PROOF. Assume $f_n \xrightarrow{\|\cdot\|} f$: then

$$\exists N \ :\ \|f_n - f\| \leqslant 1,\ \forall n \geqslant N.$$

Thus in particular $\|f_n\| \leqslant \|f_n - f\| + \|f\| \leqslant \|f\| + 1$ for all $n \geqslant N$. Thus, if we define

$$M := \max\{\|f_0\|, \|f_1\|, \dots, \|f_{N-1}\|, \|f\| + 1\},$$

we conclude $\|f_n\| \leqslant M$ for every $n \in \mathbb{N}$.

10.2. Convergence in space of functions

Most relevant normed spaces are spaces of functions. The simplest way to converge for a sequence of functions (f_n) is the *point-wise convergence*, that is

$$\lim_{n \to +\infty} f_n(x) = f(x),\ \forall x \in X.$$

A natural problem arises: *what relation exists between point-wise convergence and principal norms convergence?* This is what with want to investigate here in some remarkable cases.

10.2.1. Uniform norm. We start with the uniform norm $\|\cdot\|_\infty$ which, according to the case, might be slightly different but more or less it works at the same manner.

Proposition 10.2.1

Let $(f_n) \subset B(X)$ be such that $f_n \xrightarrow{\|\cdot\|_\infty} f$. Then $f_n \longrightarrow f$ point wise.

PROOF. Since $\|f_n - f\|_\infty \longrightarrow 0$ means

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \ :\ \|f_n - f\|_\infty \leqslant \varepsilon,\ \forall n \geqslant N,$$

that is, according to the definition of uniform norm,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \ :\ \sup_{x \in X} |f_n(x) - f(x)| \leqslant \varepsilon,\ \forall n \geqslant N,$$

or, again,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \ :\ |f_n(x) - f(x)| \leqslant \varepsilon,\ \forall x \in X,\ \forall n \geqslant N,$$

This says $f_n(x) \longrightarrow f(x)$ for every $x \in X$, and this is precisely *point-wise convergence*.

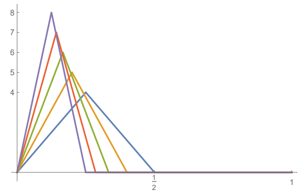
The converse is false,

Example 10.2.2: $(*)$

Let

$$f_n(x) := \begin{cases} n^2 x, & 0 \leqslant x \leqslant \frac{1}{n}, \\ -n^2 x + 2n, & \frac{1}{n} \leqslant x \leqslant \frac{2}{n}, \\ 0, & \frac{2}{n} \leqslant x \leqslant 1. \end{cases}$$

Discuss point-wise convergence and uniform convergence.



PROOF. First: $f_n(0) \equiv 0 \longrightarrow 0$ while for $x > 0$, since $\frac{2}{n} \longrightarrow 0$, for $n \geqslant N(x)$ we have $x > \frac{2}{n}$ thus $f_n(x) = 0$ for such n. This means $f_n(x) \longrightarrow 0$. Conclusion: $f_n \longrightarrow 0$ point-wise. Thus, if $f_n \overset{\|\cdot\|_\infty}{\longrightarrow} f$ then, necessarily, $f \equiv 0$. However, $\|f_n - f\|_\infty = \|f_n\|_\infty = n \longrightarrow +\infty$, so (f_n) cannot be convergent in uniform norm.

A similar conclusion holds for the L^∞ norm.

Proposition 10.2.3

Let $(X, \mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L^\infty(X)$ be such that $f_n \overset{\|\cdot\|_\infty}{\longrightarrow} f$. Then, $f_n \overset{a.e.}{\longrightarrow} f$, that is

$$f_n(x) \longrightarrow f(x),\ a.e.\, x \in X.$$

PROOF. It is very similar to the proof of previous proposition. Suppose that $f_n \overset{\|\cdot\|_\infty}{\longrightarrow} f$. Then,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N}\ :\ \|f_n - f\|_\infty \leqslant \varepsilon,\ \forall n \geqslant N.$$

Recalling that

$$|g(x)| \leqslant \|g\|_\infty,\ a.e.\, x \in X,$$

we have

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N}\ :\ |f_n(x) - f(x)| \leqslant \|f_n - f\|_\infty \leqslant \varepsilon,\ a.e.\, x \in X,\ \forall n \geqslant N.$$

Here there there is a subtle passage. Previous statement says that, for each $n \geqslant N$, there exists a null set E_n (that is $\mu(E_n) = 0$) such that

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N}\ :\ |f_n(x) - f(x)| \leqslant \|f_n - f\|_\infty \leqslant \varepsilon,\ \forall x \in X \backslash E_n,\ \forall n \geqslant N.$$

Let $E := \bigcup_n E_n$. By sub additivity $\mu(E) \leqslant \sum_n \mu(E_n) = 0$, thus E is a null set and of course

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N}\ :\ |f_n(x) - f(x)| \leqslant \|f_n - f\|_\infty \leqslant \varepsilon,\ \forall x \in X \backslash E,\ \forall n \geqslant N.$$

From this it follows that $f_n(x) \longrightarrow f(x)$ for every $x \in X \backslash E$, that is $f_n \overset{a.e.}{\longrightarrow} f$.

10.2.2. L^p norm $(1 \leqslant p < +\infty)$. Convergence in $\|\cdot\|_p$ norm is also called **convergence in $p-$mean**. Basically, saying $\|f_n - f\|_p$ becomes small it means that the area between f_n and f (weighted to power p) is small. In principle, the area could be small also with huge gap between functions. That's why we cannot expect that L^p convergence implies a.e. point-wise convergence. The next example shows dramatically this phenomenon.

Example 10.2.4: $(**)$

Let $(f_n) \subset L^1([0,1])$ defined as follows:

$$\begin{aligned} & f_1 := 1_{[0,1]}, \\ & f_2 := 1_{[0,1/2]}, \quad f_3 := 1_{[1/2,1]}, \\ & f_4 := 1_{[0,1/4]}, \quad f_5 := 1_{[1/4,2/4]}, \quad f_6 := 1_{[2/4,3/4]}, \quad f_7 := 1_{[3/4,1]}, \\ & \vdots f_{2^m} := 1_{[0,1/2^m]}, \quad f_{2^m+1} := 1_{[1/2^m,2/2^m]}, \dots, f_{2^m+k} := 1_{[k/2^m,(k+1)/2^m]}, \\ & \vdots \end{aligned}$$

Then $f_n \xrightarrow{L^1} 0$, but $(f_n(x))$ does not converge for every $x \in [0,1]$.

Proof. The first check is easy: $\|f_{2^m+k}\|_1 = \frac{1}{2^m} \longrightarrow 0$ for $m \longrightarrow +\infty$. About the second: take $x \in [0,1]$. Notice that, for every m fixed, there's just one k^* such that $[\frac{k^*}{2^m}, \frac{k^*+1}{2^m}[\ni x$. Thus $f_{2^m+k}(x) = 0$ for $k \neq k^*$ and $= 1$ for $k = k^*$. This means that the sequence $f_n(x)$ is infinitely many times $= 0$ and infinitely many times $= 1$. In particular, $(f_n(x))$ cannot be convergent, whatever is $x \in [0,1]$.

Nonetheless, we have the important

Theorem 10.2.5

Let $(X, \mathscr{F}, \mu)$ be a measure space, $(f_n) \subset L^p(X)$ $(1 \leqslant p < +\infty)$ such that $f_n \xrightarrow{L^p} f$. Then, there exists a sub-sequence $(f_{n_k}) \subset (f_n)$ such that $f_{n_k} \xrightarrow{a.e.} f$.

Proof. For simplicity, we consider $p = 1$. Replacing f_n with $f_n - f$ we may always assume $f = 0$. Thus we assume $f_n \xrightarrow{L^1} 0$, that is

$$\int_X |f_n| \, d\mu \longrightarrow 0.$$

For every k there's n_k such that

$$\int_X |f_{n_k}| \, d\mu \leqslant \frac{1}{2^k}.$$

We claim $f_{n_k} \xrightarrow{a.e.} 0$. Indeed, by monotone convergence (for series)

$$\int_X \sum_k |f_{n_k}|\, d\mu = \sum_k \int_X |f_{n_k}|\, d\mu \leqslant \sum_k \frac{1}{2^k} = 1 < +\infty,$$

thus $\sum_k |f_{n_k}(x)| < +\infty$ a.e. $x \in X$. Now, recalling that if $\sum_k a_k$ converges then, necessarily, $a_k \longrightarrow 0$, we deduce $f_{n_k}(x) \longrightarrow 0$ a.e. $x \in X$, which is the conclusion.

Remark 10.2.6

In the example 10.2.4, a sub-sequence of (f_n) that converges a.e. is, for example, (f_{2^m}). Indeed:

$$f_{2^m}(x) = 1_{[0,1/2^m]}(x) \longrightarrow 1_{\{0\}}(x), \ \forall x \in [0,1].$$

10.3. Exercises

Exercise 10.3.1 ($*$). *For each of the following sequences discuss: i) pointwise convergence on $[0,+\infty[$; ii) a.e. convergence on $[0,+\infty[$; iii) $L^1([0,+\infty[)$ convergence; iv) $L^2([0,+\infty[)$ convergence; v) $L^\infty([0,+\infty[)$ convergence.*

$$1.\ f_n := \frac{1}{n} 1_{[0,n]}.\ \ 2.\ f_n := n 1_{[0,1/n]}.\ \ 3.\ f_n := \sum_{k=0}^{n} \frac{1}{k} 1_{[k,k+\frac{1}{2^n}]}.$$

Exercise 10.3.2 ($*+$). *Let $V := \mathscr{C}([0,1])$ equipped with usual $\|\cdot\|_\infty$ norm. Let*

$$f_n(x) := x^n - x^{n+1}, \quad g_n(x) := x^n - x^{2n}.$$

Clearly $(f_n), (g_n) \subset V$.

i) *Discuss their convergence in V.*
ii) *What happens if we consider the $\|\cdot\|_1$ norm on V?*

Exercise 10.3.3 ($*$). *Let $f_n(x) = 1_{[-1,0]}(x) + 1_{]0,1/n]}(x)\sqrt{1-nx}$. Discuss convergence of (f_n) in $L^2([-1,1])$.*

Exercise 10.3.4 ($*$). *Let*

$$f_n(x) := \frac{1}{\sqrt{x+\frac{1}{n}}},\ x \in [0,1],\ n \in \mathbb{N}.$$

i) *Plot quickly the graph of f_n. Is $(f_n) \subset L^1([0,1])$? Is $(f_n) \subset L^2([0,1])$?*
ii) *Is (f_n) convergent in $L^1([0,1])$ and, in the case, to what? Is (f_n) convergent in $L^2([0,1])$ and, in the case, to what?*

Exercise 10.3.5 ($*$). *Let $f_n(x) := \frac{1}{1+x^n}$, $x \in [0,+\infty[$, $n \in \mathbb{N}$, $n \geqslant 2$. Plot quickly the graph of f_n. Is $f_n \in L^1([0,+\infty[)$? Is (f_n) convergent (and, in the case, to what) in $L^1([0,+\infty[)$? Justify your answer.*

Exercise 10.3.6 ($*$). *On $V = \mathscr{C}^1([0,1])$ let's consider*

$$i)\ \|f\|_* := \|f\|_\infty + \|f'\|_\infty.\ \ ii)\ \|f\|_{**} = |f(0)| + \|f'\|_\infty.\ \ iii)\ \|f\|_{***} := |f(1)| + \int_0^1 |f'(x)|\, dx.$$

Which among these are norms? For those who are norms, consider then the sequence $f_n(x) := \frac{1}{n}\sin(n^2 x)$. Discuss convergence of (f_n) in each of the norms. Discuss also relations among the norms.

Exercise 10.3.7 $(*+)$**.** *Let $f_n(x) := \frac{n}{1+n^9x^3}$, $x \in [0,1]$. Discuss convergence of (f_n) in $L^p([0,1])$ for $p = 1, 2, +\infty$.*

Exercise 10.3.8 $(**)$**.** *On $V := \{f \in \mathscr{C}^1([0,1]) \;:\; f(0) = 0\}$ we define*

$$\|f\| := \max_{t\in[0,1]} t^{1/2}|f'(t)|.$$

i) *Check that $\|\cdot\|$ is a norm on V.*
ii) *Show that $\|f\|$ is stronger than $\|f\|_\infty$ on V.*
iii) *Define $(f_n) \subset V$ as*

$$f_n(t) := \begin{cases} t^{1/4}, & t \in [\frac{1}{n}, 1], \\ \\ \frac{n^{3/4}}{4}t, & t \in [0, \frac{1}{n}[. \end{cases}$$

Compute $\|f_n\|$ and $\|f_n\|_\infty$. What can be deduced about equivalence of $\|\cdot\|$ and $\|\cdot\|_\infty$?

Exercise 10.3.9 $(**)$**.** *Let (E_n) be a sequence of Lebesgue measurable sets of $\mathbb{R}^d$. Suppose that $1_{E_n} \xrightarrow{\|\cdot\|_1} f$ for some $f \in L^1(\mathbb{R})$. Prove that there exists a measurable set E such that $f = 1_E$ a.e..*

LECTURE 11

Convolution

In general, Lebesgue integrable functions are very irregular and, among them, *regular functions* (such as continuous, differentiable, ...) are certainly not the prototype of an integrable function. In fact, we may think that regular function are pretty "rare" among integrable functions. Despite this intuition, in this Lecture we show that any integrable function can be approximated, in L^p norm, by a suitable sequence of extremely regular (that is $\mathscr{C}^\infty$) functions. This is possible because of a powerful operation called *convolution product* (or just convolution). The results of these chapter are very technical and most of the proofs will be omitted. Nonetheless, in many contexts of Analysis it is very important to know that we can always approximate any L^p function through a sequence of regular functions. Moreover, the convolution has several applications in Probability and Information Engineering.

11.1. Definition

A natural idea to approximate a generic (integrable) function f is to build a function whose value at point x is an average of values of f around x. For instance, fixed $\varepsilon > 0$ we might consider the function

$$f_\varepsilon(x) := \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon} f(y)\,dy = \frac{1}{2\varepsilon}\int_{\mathbb{R}} f(y)1_{[x-\varepsilon,x+\varepsilon]}(y)\,dy = \int_{\mathbb{R}} f(y)\frac{1}{2\varepsilon}1_{[-\varepsilon,\varepsilon]}(x-y)\,dy.$$

Calling $\delta_\varepsilon(x) := \frac{1}{2\varepsilon}1_{[-\varepsilon,\varepsilon]}(x)$ we have

$$f_\varepsilon(x) = \int_{\mathbb{R}} f(y)\delta_\varepsilon(x-y)\,dy.$$

The r.h.s. is called *convolution (product) of f with δ_ε*.

Definition 11.1.1

Let $f, g \in L(\mathbb{R}^m)$. We call **convolution** of f and g the function

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

(provided it makes sense)

We expect that some integrability on f and g is needed to have $f * g$ well defined. We have the

Theorem 11.1.2: (Young)

Let $f \in L^1(\mathbb{R}^m)$ and $g \in L^p(\mathbb{R}^m)$ $(1 \leqslant p \leqslant +\infty)$. Then, the convolution $f * g$ is well defined, $f * g \in L^p(\mathbb{R})$ and the **Young inequality** holds

$$\|f * g\|_p \leqslant \|f\|_1 \|g\|_p. \tag{11.1.1}$$

Proof. We accept $f * g \in L(\mathbb{R}^m)$. For sake of simplicity we do the proof of Young inequality in the case $p = 1$. We have:

$$\begin{aligned}
\|f * g\|_1 &= \int_{\mathbb{R}^m} |f * g(x)|\, dx = \int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} f(y) g(x-y)\, dy \right| dx \leqslant \int_{\mathbb{R}} \int_{\mathbb{R}^m} |f(y)||g(x-y)|\, dy\, dx \\
&\stackrel{RF}{=} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} |f(y)||g(x-y)|\, dx\, dy = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^m} |g(x-y)|\, dx \right) |f(y)|\, dy \\
&\stackrel{z=x-y}{=} \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^m} |g(z)|\, dz \right) |f(y)|\, dy = \|f\|_1 \|g\|_1.
\end{aligned}$$

Example 11.1.3: $(*)$

Compute the convolution $e^{-a|\cdot|} * e^{-b|\cdot|}$ with $a \neq b$ and $a, b > 0$.

Proof. Clearly $e^{-a|x|}, e^{-b|x|} \in L^1(\mathbb{R})$ for $a, b > 0$, so the convolution is well defined. We have

$$e^{-a|\cdot|} * e^{-b|\cdot|}(x) = \int_{\mathbb{R}} e^{-a|y|} e^{-b|x-y|}\, dy.$$

For $x \geqslant 0$, we split the integral into three parts:

$$\begin{aligned}
I_1 &= \int_{-\infty}^{0} e^{-a(x-y)} e^{-b(-y)}\, dy = e^{-ax} \int_{-\infty}^{0} e^{(a+b)y}\, dy = \frac{e^{-ax}}{a+b}, \\
I_2 &= \int_{0}^{x} e^{-a(x-y)} e^{-by}\, dy = e^{-ax} \int_{0}^{x} e^{(a-b)y}\, dy \stackrel{a \neq b}{=} e^{-ax} \frac{e^{(a-b)x} - 1}{a-b} = \frac{e^{-bx} - e^{-ax}}{a-b}., \\
I_3 &= \int_{x}^{+\infty} e^{-a(y-x)} e^{-by}\, dy = e^{ax} \int_{x}^{+\infty} e^{-(a+b)y}\, dy = \frac{e^{-bx}}{a+b}.
\end{aligned}$$

Summing $I_1 + I_2 + I_3$ gives

$$(f * g)(x) = \frac{2}{a^2 - b^2} \left(a\, e^{-bx} - b\, e^{-ax} \right).$$

For $x < 0$, the calculation is the same with $-x$ replacing x. Therefore,

$$\boxed{(f * g)(x) = \frac{2}{a^2 - b^2} \left(a\, e^{-b|x|} - b\, e^{-a|x|} \right).}$$

Convolution $f * g$ is an operation that produces a function given two functions f and g. Previous Theorem shows that if $f, g \in L^1$ then $f * g \in L^1$. This operation fulfils properties similar to algebraic product of numbers:

Proposition 11.1.4

Convolution product fulfils:

i) (commutativity) $f * g = g * f$, for $f, g \in L^1$;
ii) (associativity) $f * (g * h) = (f * g) * h$, for $f, g, h \in L^1$;
iii) (distributivity) $f * (g + h) = f * g + f * h$, for $f, g, h \in L^1$

Proof is left as exercise. We may wonder if a unit exists, namely a function $\delta \in L^1$ such that $f * \delta = f$ for every $f \in L^1$. This δ does not exists. We show this in dimension $m = 1$. Taking $f = 1_{[-\varepsilon,\varepsilon]}$, is a unit δ would exists, we would have

$$f * \delta = f, \iff \int_{-\varepsilon}^{\varepsilon} \delta(x-y)\, dy = 1_{[-\varepsilon,\varepsilon]}(x),$$

thus, in particular, for $x = 0$,

$$1 = \int_{-\varepsilon}^{\varepsilon} \delta(-y)\, dy \longrightarrow 0,\ \varepsilon \longrightarrow 0.$$

In other words, a unit should be the famous Dirac's delta function. Nonetheless, "approximate units" exists, and this is the content of next Section.

11.2. Approximate units

Even if there is not a unit for the convolution, there are "approximate units".

Definition 11.2.1

Let $\delta \in L^1(\mathbb{R}^m)$ such that

$$\delta \geqslant 0 \ \text{ a.e.}, \quad \int_{\mathbb{R}^m} \delta(x)\, dx = 1.$$

The family of functions $(\delta_\varepsilon)_{\varepsilon>0}$ defined as

$$\delta_\varepsilon(x) := \frac{1}{\varepsilon^m}\delta\left(\frac{x}{\varepsilon}\right),\ x \in \mathbb{R}^m$$

is called **approximate unit**.

Remark 11.2.2: B

definition, $\delta_\varepsilon \geqslant 0$ a.e. and

$$\int_{\mathbb{R}^m} \delta_\varepsilon(x)\, dx = \frac{1}{\varepsilon^m}\int_{\mathbb{R}^m} \delta\left(\frac{x}{\varepsilon}\right)\, dx \overset{y=x/\varepsilon}{=} \frac{1}{\varepsilon^m}\int_{\mathbb{R}^m} \delta(y)\varepsilon^m\, dy = 1.$$

Here some examples of approximate units.

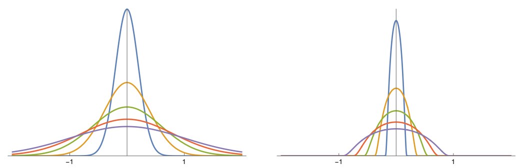
Example 11.2.3: step approximate unit

Let $\delta(x) = \frac{1}{2}1_{[-1,1]}$. Then $\delta_\varepsilon(x) = \frac{1}{2\varepsilon}1_{[-\varepsilon,\varepsilon]}(x)$.

The next example plays a very important role in many applications. Differently from the previous unit, the next one is also a very regular function (a $\mathscr{C}^\infty$ function with very fast decay at infinity).

Example 11.2.4: Gaussian approximate unit

Let
$$\delta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Longrightarrow \quad \delta_\varepsilon(x) = \frac{1}{\sqrt{2\pi\varepsilon^2}} e^{-\frac{x^2}{2\varepsilon^2}}.$$



We can also have units vanishing outside a compact interval.

Example 11.2.5

Let
$$\delta(x) := \begin{cases} e^{-\frac{1}{1-x^2}}, & |x| < 1, \\ \\ 0, & |x| \geqslant 1. \end{cases}$$
This is $\mathscr{C}^\infty(\mathbb{R})$ function. Indeed, the unique problem is at $x = \pm 1$. Easily we verify that δ is continuous at $x = \pm 1$. Computing the derivative we have
$$\delta'(x) = \begin{cases} e^{-\frac{1}{1-x^2}} \frac{-2x}{(1-x^2)^2}, & |x| < 1, \\ \\ 0, & |x| > 1. \end{cases}$$
Therefore,
$$\lim_{x \to 1-} \delta'(x) = -2 \lim_{x \to 1-} \frac{e^{-\frac{1}{1-x^2}}}{(1-x^2)^2} \overset{t = \frac{1}{1-x^2}}{=} -2 \lim_{t \to -\infty} t^2 e^{-t} = 0,$$
and, similarly, $\delta'(-1+) = 0$. Since $\delta'(1+) = \delta'(-1-) = 0$, we conclude that $\exists \delta'(\pm 1) = 0$, and, in particular $\delta \in \mathscr{C}^1(\mathbb{R})$. Iterating this argument we have the conclusion.
Clearly, $\delta \geqslant 0$ but $I := \int_{\mathbb{R}} \delta(x)\, dx$ is not necessarily $= 1$. However, rescaling δ by $\frac{1}{I}$, that is taking $\frac{1}{I}\delta$ we can define an approximate unit
$$\delta_\varepsilon(x) := \frac{1}{\varepsilon I} \delta\left(\frac{x}{\varepsilon}\right).$$
Notice that $\delta_\varepsilon(x) = 0$, if $|x| \geqslant \varepsilon$.

Approximate units deserve this name because $f * \delta_\varepsilon \approx f$ when $\varepsilon \approx 0$. Precisely, we have the

Proposition 11.2.6

Let $f \in L^p(\mathbb{R}^m)$ and $(\delta_\varepsilon) \subset L^1(\mathbb{R}^m)$ be an approximate unit. Then

$$f * \delta_\varepsilon \xrightarrow{L^p} f. \tag{11.2.1}$$

PROOF. (sketch, case $p = 1$, $m = 1$) To show that $f * \delta_\varepsilon \xrightarrow{L^1} f$, we start noticing that

$$f * \delta_\varepsilon(x) = \int_{\mathbb{R}} f(x-y)\delta_\varepsilon(y)\, dy = \int_{\mathbb{R}} f(x-y)\frac{1}{\varepsilon}\delta\left(\frac{y}{\varepsilon}\right)\, dy \overset{u:=\frac{y}{\varepsilon}}{=} \int_{\mathbb{R}} f(x-\varepsilon u)\delta(u) du.$$

Recalling that $\int_{\mathbb{R}} \delta(u)\, du = 1$, we have

$$f * \delta_\varepsilon(x) - f(x) = \int_{\mathbb{R}} (f(x-\varepsilon u) - f(x))\,\delta(u)\, du$$

so

$$\|f * \delta_\varepsilon - f\|_1 \leqslant \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-\varepsilon u) - f(x)|\,\delta(u)\, du \right) dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-\varepsilon u) - f(x)|\, dx \right) \delta(u)\, du$$

Introducing the *translation operator* $T_\eta f := f(\sharp + \eta)$, we may write

$$\int_{\mathbb{R}} |f(x-\varepsilon u) - f(x)|\, dx = \|T_{-\varepsilon u} f - f\|_1.$$

At this point we need the

Lemma 11.2.7

Let $f \in L^p(\mathbb{R}^m)$. Then

$$\lim_{h\to 0} \|T_h f - f\|_p = 0.$$

The proof of the Lemma is technical and is omitted here. We can now conclude the proof: since

$$\|f * \delta_\varepsilon - f\|_1 \leqslant \int_{\mathbb{R}} \|T_{-\varepsilon u} f - f\|_1 \delta(u)\, du,$$

to compute the limit for $\varepsilon \downarrow 0$, we apply dominated convergence. We have

- (pointwise a.e. limit) by the Lemma $\|T_{-\varepsilon u} f - f\|_1 \delta(u) \longrightarrow 0$, for every $u \in \mathbb{R}$.
- (integrable dominant) since $\|T_{-\varepsilon u} f - f\|_1 \leqslant \|T_{-\varepsilon u} f\|_1 + \|f\|_1 = 2\|f\|_1$,we have

$$\|T_{-\varepsilon u} f - f\|_1 \delta(u) \leqslant 2\|f\|_1 \delta(u) =: g(u) \in L^1(\mathbb{R}).$$

Thus, assumption of dominated convergence thm are fulfilled and the conclusion now follows. □

11.3. Mollification Theorem

Doing the convolution with an approximate unit introduces an approximation of any $f \in L^p$ function. Since

$$f * \delta_\varepsilon(x) = \int_{\mathbb{R}^m} f(y)\delta_\varepsilon(x-y)\, dy,$$

we see that $f * \delta_\varepsilon(x)$ depends by x through $\delta_\varepsilon(x-y)$ under the integral sign. This suggests the idea that if δ_ε (hence δ) is regular enough (that is, differentiable a certain number of times), also $f * \delta_\varepsilon$ could be differentiable, this because of the differentiability under integral sign theorem. This opens the way to approximate any $f \in L^p$ through regular functions. This operation is also named *mollification* of f.

There are many classes of regular functions and corresponding approximation results. Here, we will choose a particular class that plays an important role in the theory of Fourier Transform. Roughly speaking, this is the class of $\mathscr{C}^\infty$ functions decaying fast at ∞. To keep light notations, we will limit to the case of functions of one real variable, but the definitions and results extends in a straightforward way to functions of vector variable.

Definition 11.3.1: Schwartz class

$$\mathcal{S}(\mathbb{R}) := \left\{ f \in \mathscr{C}^\infty(\mathbb{R}) \; : \; \sup_{x \in \mathbb{R}} (1+|x|)^N |\partial_x^k f(x)| < +\infty, \; \forall N, \; k \in \mathbb{N} \right\}.$$

In words: a Schwartz function is a $\mathscr{C}^\infty(\mathbb{R})$ function decaying at infinity with its derivatives faster than any polynomial. For example:

- $e^{-x^2}, e^{-x^4}, x^2 e^{-x^2} \in \mathcal{S}(\mathbb{R})$;
- $e^{-|\sharp|} \notin \mathcal{S}(\mathbb{R})$ (problem: regularity at 0);
- $\frac{1}{1+x^2} \notin \mathcal{S}(\mathbb{R})$ (problem: not decaying fast enough at $\pm\infty$).

It is not difficult to prove that Schwartz class is contained in any L^p space:

Proposition 11.3.2

$$\mathcal{S}(\mathbb{R}) \subset L^p(\mathbb{R}), \; \forall 1 \leqslant p \leqslant +\infty.$$

Proof. $(p = 1)$ If $f \in \mathcal{S}$ then, in particular, $f \in \mathscr{C}$, thus f is integrable in every closed and bounded interval. For integrability on $\mathbb{R}$ in L^1 sense, we have to check the behaviour at $\pm\infty$. Since

$$(1+|x|)^2 |f(x)| \leqslant C, \implies |f(x)| \leqslant \frac{C}{(1+|x|)^2} \in L^1.$$

Theorem 11.3.3: mollification

Schwartz class is dense in $L^p(\mathbb{R})$ for every $1 \leqslant p \leqslant +\infty$. Precisely:

$$\forall f \in L^p(\mathbb{R}), \; \exists (f_n) \subset \mathcal{S}(\mathbb{R}) \; : \; f_n \xrightarrow{L^p} f.$$

Proof. $(p = 1)$ Let $f \in L^1(\mathbb{R})$. The idea is to take $f * \delta_\varepsilon$, where (δ_ε) is the Gaussian unit. To check that $f * \delta_\varepsilon \in \mathcal{S}(\mathbb{R})$ we need f be zero outside an interval. This is not true in general, so we need to approximate f by functions zero outside an interval. We start from this task.
Define $f_R := f 1_{[-R,R]}$. Clearly $f_R \in L^1(\mathbb{R})$. A straightforward application of dominated convergence (exercise) shows that

$$\|f - f_R\|_1 = \int_{\mathbb{R}} |f(x) - f_R(x)| \, dx = \int_{\mathbb{R}} f(x) 1_{[-R,R]^c}(x) \, dx, \;\; R \longrightarrow +\infty.$$

Define now,

$$f_{R,\varepsilon}(x) := f_R * \delta_\varepsilon(x) = \int_{\mathbb{R}} f_R(y) e^{-\frac{(x-y)^2}{2\varepsilon^2}} \frac{dy}{\sqrt{2\pi\varepsilon^2}} = \int_{-R}^{R} f(y) e^{-\frac{(x-y)^2}{2\varepsilon^2}} \frac{dy}{\sqrt{2\pi\varepsilon^2}}$$

Let's check that $f_{R,\varepsilon} \in \mathcal{S}(\mathbb{R})$ for every $\varepsilon > 0$ (fixed). Notice that, deriving under integral sign (if allowed), we would have

$$\partial_x^k f_{R,\varepsilon}(x) = \frac{1}{\sqrt{2\pi\varepsilon^2}} \int_{-R}^{R} f(y) \partial_x^k e^{-\frac{(x-y)^2}{2\varepsilon^2}} \, dy$$

We will verify in a moment that this passage is really allowed. Before, we notice that since $\partial_t^k e^{-\frac{t^2}{2\varepsilon^2}} = p_\varepsilon^k(t) e^{-\frac{t^2}{2\varepsilon^2}}$, where p_ε^l is a certain $k-$th degree polynomial (it is irrelevant here the particular form of this). Thus

$$\partial_x^k f_{R,\varepsilon}(x) = \frac{1}{\sqrt{2\pi\varepsilon^2}} \int_{-R}^{R} f(y) p_\varepsilon^k(x-y) e^{-\frac{(x-y)^2}{2\varepsilon^2}} \, dy.$$

Notice that since $|p_\varepsilon^k(t) e^{-t^2/2\varepsilon^2}| \leqslant C_\varepsilon^k$ for every $t \in \mathbb{R}$, we have a bound

$$\left| f(y) p_\varepsilon^k(x-y) e^{-\frac{(x-y)^2}{2\varepsilon^2}} \right| \leqslant C_\varepsilon^k |f(y)| =: g(y) \in L^1([-R,R]),$$

thus differentiation under integral sign is justified. We now prove the decay at ∞: we have to prove that

$$\lim_{x \to \pm\infty} |x|^N \partial_x^k f_{R,\varepsilon}(x) = 0. \tag{11.3.1}$$

If $|x| > R$, $|x-y| \geqslant |x| - |y| \geqslant |x| - R$, thus

$$e^{-\frac{(x-y)^2}{2\varepsilon^2}} \leqslant e^{-\frac{(|x|-R)^2}{2\varepsilon^2}}$$

while, being $p_\varepsilon^k(t) \leqslant a_\varepsilon^k |t|^k + b_\varepsilon^k$ (for suitable constants $a_\varepsilon^k, b_\varepsilon^k$), we have that

$$|p_\varepsilon^k(x-y)| \leqslant a_\varepsilon^k |x-y|^k + b_\varepsilon^k \leqslant a_\varepsilon^k (|x|+R)^k + b_\varepsilon^k.$$

Therefore

$$|\partial_x^k f_{R,\varepsilon}(x)| \leqslant \frac{1}{\sqrt{2\pi\varepsilon^2}} \int_{-R}^{R} f(y) \, dy \left(a_\varepsilon^k (|x|+R)^k + b_\varepsilon^k \right) e^{-\frac{(|x|-R)^2}{2\varepsilon^2}} \leqslant c_\varepsilon^k \|f\|_1 |x|^k e^{-\frac{x^2}{2\varepsilon^2}},$$

for a suitable constant c_ε^k. From this bound limit (11.3.1) easily follows.

Conclusion: we can now put together the two arguments. Fix $n \in \mathbb{N}$, $n \geqslant 1$. Choose R_n such that $\|f - f_{R_n}\|_1 \leqslant \frac{1}{n}$. By key property (11.2.1) of approximate units, choose now ε_n such that $\|f_{R_n,\varepsilon_n} - f_{R_n}\|_1 \leqslant \frac{1}{n}$. Therefore

$$\|f - f_{R_n,\varepsilon_n}\|_1 \leqslant \frac{2}{n},$$

hence $f_{R_n,\varepsilon_n} \xrightarrow{L^1} f$. According to what shows above, $(f_{R_n,\varepsilon_n}) \subset \mathcal{S}(\mathbb{R})$, and the proof is complete.

11.4. Exercises

Exercise 11.4.1 (∗). *Compute the convolutions* $1_{[-1,1]} * 1_{[-1,1]}$ *and* $1_{[0,1]} * 1_{[0,1]}$.

Exercise 11.4.2 (∗). *Check that the convolution integral of* $x1_{[0,+\infty[}$ *and* $x^2 1_{[0,+\infty[}$ *is well defined. What goes wrong with Young's theorem?*

Exercise 11.4.3 ($*$)**.** *Let $f, g \in L^1(\mathbb{R})$ be even functions, that is $f(-x) = f(x)$ and $g(-x) = g(x)$ a.e. Check that $f * g$ is even.*

Exercise 11.4.4 ($*$)**.** *Prove the Proposition 11.1.4.*

Exercise 11.4.5 ($**$)**.** *Let $f, g \in L^2(\mathbb{R})$. Check that $f * g$ is well defined and it belongs to $L^\infty(\mathbb{R})$ and $\|f * g\|_\infty \leqslant \|f\|_2 \|g\|_2$. Extend this to the case $f \in L^p$ and $g \in L^q$ with $1 < p, q < +\infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$.*

Exercise 11.4.6 ($***$)**.** *Goal: prove Young inequality.*

i) *prove the case $f \in L^1$ and $g \in L^\infty$.*
ii) *Prove the case $f \in L^1$ and $g \in L^2$ by using the following trick:*

$$|f(y)g(x-y)| = |f(y)|^{1/2} |f(y)|^{1/2} |g(x-y)|,$$

then use Cauchy-Schwarz inequality.
iii) *Extend trick shown at point ii) to prove the case $f \in L^1$ and $g \in L^p$ ($1 < p < +\infty$).*

Exercise 11.4.7 ($***$)**.** *Let $\delta(x) = \frac{1}{2} 1_{[-1,1]}(x)$ and $f \in L^1(\mathbb{R})$. Prove that $f * \delta_\varepsilon \in \mathscr{C}(\mathbb{R})$ for every $\varepsilon > 0$.*

Exercise 11.4.8 ($**$)**.** *Let $f, g \in L^1(\mathbb{R})$, $f \equiv g \equiv 0$ off $[-R, R]$. Is $f * g \equiv 0$ off a suitable interval?*

Exercise 11.4.9 ($**+$)**.** *Let $f \in L^1(\mathbb{R})$ and $g \in \mathscr{C}^1(\mathbb{R})$ with $g' \in L^1(\mathbb{R})$. Check that $f * g$ is differentiable and*

$$(f * g)' = f * g'.$$

Deduce the bound $\|(f * g)'\|_1 \leqslant \|f\|_1 \|g'\|_1$.

LECTURE 12

Completeness

Discussing convergence in a normed space might be complicate, particularly when the space in infinite dimensional. Often, the sequence is not explicitly given but is defined as the solution of a certain equation. A question arises: *is it possible to establish convergence of a sequence without explicitly determining its limit*? This is the focus of this Lecture.

12.1. Cauchy property

The Cauchy property is an intrinsic property fulfilled by any convergent sequence:

Proposition 12.1.1

Let $(V, \|\cdot\|)$ be a normed space. If $(f_n) \subset V$ is a convergent sequence to some $f \in V$, then fulfills the **Cauchy property**:

$$\forall \varepsilon > 0,\ \exists N\ :\ \|f_n - f_m\| \leqslant \varepsilon,\ \forall n \geqslant N.$$

Proof. If $\|f_n - f\| \longrightarrow 0$, according to the definition of limit,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ :\ \|f_n - f\| \leqslant \varepsilon,\ \forall n \geqslant N.$$

Then,

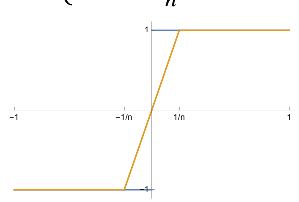
$$\|f_n - f_m\| \leqslant \|f_n - f\| + \|f - f_m\| \leqslant 2\varepsilon,\ \forall n, m \geqslant N.$$

Unfortunately, this condition is not always sufficient to ensure convergence.

Example 12.1.2: $(**)$

Let $V := \mathscr{C}([-1, 1])$ equipped with $\|\cdot\|_1$ norm. Define

$$f_n(x) = \begin{cases} -1, & -1 \leqslant x \leqslant -\frac{1}{n}, \\ nx, & -\frac{1}{n} \leqslant x \leqslant \frac{1}{n}, \\ 1, & \frac{1}{n} \leqslant x \leqslant 1. \end{cases}$$



Then $(f_n)$is a Cauchy sequence not convergent in V respect to $\|\cdot\|_1$.

PROOF. Cauchy property: just notice that, if $m > n$, $\|f_n - f_m\|_1 \leqslant \frac{1}{2}\frac{2}{n}\frac{1}{2} = \frac{1}{2n} \leqslant \varepsilon$ provided $m, n \geqslant N = \left[\frac{1}{\varepsilon}\right] + 1$. Let's prove now that (f_n) cannot be convergent in V. First notice that if we look at $(f_n) \subset L^1([-1,1])$. then easily $f_n \xrightarrow{\|\cdot\|_1} -1_{[-1,0]}+1_{[0,1]}$ (indeed: $\|f_n-(-1_{[-1,0]}+1_{[0,1]})\|_1 \leqslant \frac{1}{2n} \longrightarrow 0$). However, $-1_{[-1,0]} + 1_{[0,1]} \notin \mathscr{C}([-1,1])$ thus we cannot conclude that (f_n) converges in V. Actually, we may use this fact to just prove the opposite. Indeed: assume that $f_n \xrightarrow{\|\cdot\|_1} g$ for some $g \in \mathscr{C}([-1,1])$. Since $\mathscr{C}([0,1]) \hookrightarrow L^1([0,1])$, we have at once

$$f_n \xrightarrow{L^1} g, \wedge f_n \xrightarrow{L^1} -1_{[-1,0]} + 1_{[0,1]}, \implies g = -1_{[-1,0]} + 1_{[0,1]}, \text{ a.e..}$$

We claim that $g = -1_{[-1,0]} + 1_{[0,1]}$ on $[-1,1]\backslash\{0\}$. Indeed: take $x_0 < 0$. If $g(x_0) \neq -1$ then, by continuity of g, $g(x) \neq -1$ in a neighborhood I_{x_0} of x_0. But then there would be a positive measure set I_{x_0} on which $g \neq -1_{[-1,0]} + 1_{[0,1]}$ contradicting $g = -1_{[-1,0]} + 1_{[0,1]}$ a.e.. Similarly, $g(x) \equiv 1$ on $]0,1]$, so $g = -1_{[-1,0]} + 1_{[0,1]}$ on $[-1,1]\backslash\{0\}$. But then g cannot e continuous at $x = 0$, and this contradict $g \in \mathscr{C}([-1,1])$.

This example is quite "pathological". Indeed, $\|\cdot\|_1$ is not the natural norm for the set of continuous functions $\mathscr{C}([0,1])$ just because convergence in "mean" is too weak to ensure continuity to the limit.

12.2. Banach spaces

Fortunately, in most important normed spaces Cauchy sequences are convergent. This deserve a special

Definition 12.2.1

A normed space $(V, \|\cdot\|)$ is called **Banach space** (or **complete space**) if every Cauchy sequence $(f_n) \subset V$ is convergent.

We will now illustrate this Definition on the most important cases we considered in this course. We start with finite dimensional spaces. The following fact is know from Mathematical Analysis:

Theorem 12.2.2

$(\mathbb{R}, |\cdot|)$, $(\mathbb{C}, |\cdot|)$, $(\mathbb{R}^d, \|\cdot\|)$ (any norm), $(\mathbb{C}^d, \|\cdot\|)$ (any norm) are Banach spaces.

We do not prove this theorem, but we will use in the next results with focus on infinite dimensional spaces. We start with the simplest case,

Proposition 12.2.3

$(B(X), \|\cdot\|_\infty)$ is a Banach space.

Proof. Let $(f_n) \subset B(X)$ be a Cauchy sequence:

$$\forall \varepsilon > 0,\ \exists N\ :\ \|f_n - f_m\|_\infty = \sup_{x\in X} |f_n(x) - f_m(x)| \leqslant \varepsilon,\ \forall n, m \geqslant N,$$

or, equivalently,

$$\forall \varepsilon > 0,\ \exists N\ :\ |f_n(x) - f_m(x)| \leqslant \varepsilon,\ \forall n, m \geqslant N,\ \forall x \in X. \tag{12.2.1}$$

Thus, in particular, $(f_n(x)) \subset \mathbb{R}$ is a Cauchy sequence in $\mathbb{R}$, and being this last complete,

$$\exists\ f(x) := \lim_n f_n(x),\ \forall x \in X.$$

This defines a function. We prove that $f \in B(X)$ and $f_n \xrightarrow{\|\cdot\|_\infty} f$. About the first notice just that

$$|f(x)| \leqslant |f(x) - f_N(x)| + |f_N(x)| \leqslant \varepsilon + \|f_N\|_\infty,\ \forall x \in X,\ \Longrightarrow\ \|f\|_\infty \leqslant \varepsilon + \|f_N\|_\infty,$$

that is $f \in B(X)$. Finally, letting $m \longrightarrow +\infty$ in (12.2.1)we have

$$|f_n(x) - f(x)| \leqslant \varepsilon,\ \forall n \geqslant N,\ \forall x \in X,\ \Longrightarrow\ \|f_n - f\|_\infty = \sup_{x\in X} |f_n(x) - f(x)| \leqslant \varepsilon,\ \forall n \geqslant N$$

that means $\|f_n - f\|_\infty \longrightarrow 0$.

A particular case of space of bounded functions is $\mathscr{C}(K)$ where K is compact in $\mathbb{R}^d$.

Proposition 12.2.4

$(\mathscr{C}(K), \|\cdot\|_\infty)$ with $K \subset \mathbb{R}^d$ compact, is a Banach space.

Proof. We know that $\mathscr{C}(K) \subset B(K)$. Therefore, if $(f_n) \subset \mathscr{C}(K) \subset B(K)$ is a Cauchy sequence, according to previous Theorem, $f_n \xrightarrow{\|\cdot\|_\infty} f \in B(K)$. It remains to prove that $f \in \mathscr{C}(K)$, that is f is continuous on K,

$$\lim_{x\to x_0} f(x) = f(x_0),\ \forall x_0 \in K.$$

Let fix $x_0 \in K$. We have

$$\begin{aligned} |f(x) - f(x_0)| &\leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &\leqslant 2\|f_n - f\|_\infty + |f_n(x) - f_n(x_0)|. \end{aligned}$$

Now, since $\|f_n - f\|_\infty \longrightarrow 0$, for $\varepsilon > 0$ we find N such that $\|f_N - f\|_\infty \leqslant \varepsilon$. Thus

$$|f(x) - f(x_0)| \leqslant 2\varepsilon + |f_N(x) - f_N(x_0)|,$$

and since $f_N \in \mathscr{C}(K)$,

$$\lim_{x\to x_0} |f(x) - f(x_0)| \leqslant \varepsilon + \lim_{x\to x_0} |f_N(x) - f_N(x_0)| = \varepsilon,$$

and since ε is arbitrary, we conclude that $\lim_{x\to x_0} |f(x) - f(x_0)| = 0$.

This last result enlighten a general fact.

Proposition 12.2.5

Let $(V, \|\cdot\|)$ be a Banach space and $W \subset V$ a subspace of V. Suppose that W is **closed**, that is:

$$\forall (f_n) \subset W, \; : \; f_n \xrightarrow{\|\cdot\|} f, \quad \Longrightarrow \quad f \in W. \tag{12.2.2}$$

Then $(W, \|\cdot\|)$ is a Banach space.

Proof. Let $(f_n) \subset W \subset V$ be a Cauchy sequence. Since $(V, \|\cdot\|)$ is a Banach space, $f_n \xrightarrow{\|\cdot\|} f$ and, by (12.2.2), $f \in W$. Thus (f_n) converges in $(W, \|\cdot\|)$.

Also L^p spaces are important examples Banach spaces.

Theorem 12.2.6

Let $(X, \mathscr{F}, \mu)$ be a measure space. Then, $(L^p(X), \|\cdot\|_p)$ is a Banach space (for every $1 \leqslant p \leqslant +\infty$).

Proof. (sketch, $p = 1$) Let $(f_n) \subset L^1(X)$ be a Cauchy sequence, that is

$$\forall \varepsilon > 0, \; \exists N \; : \; \|f_n - f_m\|_1 \leqslant \varepsilon, \; \forall n, m \geqslant N.$$

We know that even if (f_n) were convergent in L^1, it would not necessarily converge pointwise. Nonetheless, we need to identify a candidate limit f. To do this, we will now extract a subsequence from (f_n) that converges pointwise. We proceed in the following way:

- for $\varepsilon = 1$, let $n_0 := N(1)$;
- for $\varepsilon = \frac{1}{2}$, let $n_1 > \max(n_0, N(1/2)) \geqslant n_0 = N(1)$. In this way $\|f_{n_1} - f_{n_0}\|_1 \leqslant 1$.
- for $\varepsilon = \frac{1}{2^2}$, let $n_2 > \max(n_1, N(1/2^2)) \geqslant n_1 = N(1/2)$. In this way $\|f_{n_2} - f_{n_1}\|_1 \leqslant \frac{1}{2}$.
- in general, for $\varepsilon = \frac{1}{2^k}$, let $n_k > \max(n_{k-1}, N(1/2^k)) \geqslant n_{k-1} = N(1/2^{k-1})$. In this way

$$\|f_{n_k} - f_{n_{k-1}}\|_1 \leqslant \frac{1}{2^{k-1}}.$$

We claim that (f_{n_k}) converges a.e.. Indeed, by monotone convergence for series

$$\int_X \sum_k |f_{n_k} - f_{n_{k-1}}| \, d\mu = \sum_k \int_X |f_{n_k} - f_{n_{k-1}}| \, d\mu = \sum_k \|f_{n_k} - f_{n_{k-1}}\|_1 \leqslant \sum_k \frac{1}{2^{k-1}} = 2.$$

In particular, $\mu\left(\sum_k |f_{n_k} - f_{n_{k-1}}| = +\infty\right) = 0$ (Chebyshev inequality), so $\sum_k |f_{n_k} - f_{n_{k-1}}| < +\infty$ a.e. This says that the series $\sum_k (f_{n_k} - f_{n_{k-1}})$ is absolutely convergent, hence convergent, and since its partial sums are

$$s_j := \sum_{k=1}^{j} (f_{n_k} - f_{n_{k-1}}) = f_{n_j} - f_{n_0},$$

it means that (f_{n_j}) must be convergent. We finally have a point-wise limit

$$f := \lim_j f_{n_j}.$$

Being this the pointwise limit of measurable functions, it is itself measurable.

. We now claim that $f_n \xrightarrow{L^1} f$. We begin showing that $f_{n_j} \xrightarrow{L^1} f$. Recalling that

$$f = \lim_j f_{n_j} = f_{n_0} + \lim_j \sum_{k=1}^{j} (f_{n_k} - f_{n_{k-1}}) = f_{n_0} + \sum_{k=1}^{\infty} (f_{n_k} - f_{n_{k-1}}),$$

we have

$$f - f_{n_j} = \sum_{k=j+1}^{\infty} (f_{n_k} - f_{n_{k-1}}),$$

so

$$\|f - f_{n_j}\|_1 = \left\| \sum_{k=j+1}^{\infty} (f_{n_k} - f_{n_{k-1}}) \right\|_1 \leqslant \sum_{k=j+1}^{\infty} \|f_{n_k} - f_{n_{k-1}}\|_1 \leqslant \sum_{k=j+1}^{\infty} \frac{1}{2^{k-1}} \leqslant \frac{1}{2^j} \longrightarrow 0,$$

from which $f_{n_j} \xrightarrow{L^1} f$. Finally, since

$$\|f - f_n\|_1 \leqslant \|f - f_{n_j}\|_1 + \|f_{n_j} - f_n\|_1,$$

by choosing n, n_j large enough in such a way $\|f - f_{n_j}\|_1 \leqslant \varepsilon$ (by the previous conclusion) and $\|f_{n_j} - f_n\|_1 \leqslant \varepsilon$ (by the Cauchy property), we get the conclusion.

12.3. Banach fixed point Theorem

A way used in many models to define a sequence is through a *recurrence equation* like

$$\begin{cases} f_{n+1} = T[f_n], \\ \\ f_0 \in V. \end{cases} \tag{12.3.1}$$

Here we assume that $(V, \|\cdot\|)$ be an underlying vector space, $f_0 \in V$ is known (first element of the sequence) and $T : V \longrightarrow V$ is just a map (function) from V to itself. Since T is a function on a normed space, a natural definition of continuity makes sense: T is continuous if

$$f_n \xrightarrow{\|\cdot\|} f, \implies T[f_n] \xrightarrow{\|\cdot\|} T[f].$$

Now, if $f_n \xrightarrow{\|\cdot\|} f$ and T is assumed continuous, then

$$f \longleftarrow f_{n+1} = T[f_n] \longrightarrow T[f], \implies f = T[f].$$

The possible limit is what is called a *fixed point* of the map T. Of course, this argument does not show in any way that a limit exists. The Banach's fixed point theorem provides an important sufficient condition to ensure existence of the limit for a recurrence sequence.

Theorem 12.3.1: (Banach)

Let $(V, \|\cdot\|)$ be a Banach space. Assume $T : V \longrightarrow V$ be a **contraction**, that is

$$\exists L < 1, : \|T[f] - T[g]\| \leqslant L\|f - g\|, \forall f, g \in V. \tag{12.3.2}$$

Then, for every $f_0 \in V$, the sequence (f_n) recursively defined by (12.3.1) converges to the unique fixed point $f \in V$ of T. The following bound holds:

$$\|f_n - f\| \leqslant \frac{L^n}{1-L}\|f_1 - f_0\|. \tag{12.3.3}$$

Proof. Existence. We prove that (f_n) is a Cauchy sequence. Notice first that

$$\|f_{n+1} - f_n\| = \|T[f_n] - T[f_{n-1}]\| \leqslant L\|f_n - f_{n-1}\| \leqslant L^2\|f_{n-1} - f_{n-2}\| \leqslant \ldots \leqslant L^n\|f_1 - f_0\|.$$

Thus, if $m > n$,

$$\|f_m - f_n\| \leqslant \sum_{k=n}^{m-1} \|f_{k+1} - f_k\| \leqslant \sum_{k=n}^{m-1} L^k\|f_1 - f_0\| \leqslant L^n \sum_{k=0}^{\infty} L^k\|f_1 - f_0\| = \frac{L^n}{1-L}\|f_1 - f_0\|.$$

Since $L < 1$, $L^n \longrightarrow 0$ and by this it follows that (f_n) is a Cauchy sequence. Being V a Banach space, $f_n \xrightarrow{\|\cdot\|} f$ for some $f \in V$. To conclude existence, just notice that since T is a contraction, it is also continuous, then

$$f \longleftarrow f_{n+1} = T[f_n] \longrightarrow T[f], \implies f = T[f].$$

This proves, at once, that (f_n) converges and that T has (at least) a fixed point.
Uniqueness: we show that T can has at most a fixed point. If $g = T[g]$ then $f - g = T[f] - T[g]$, thus $\|f - g\| = \|T[f] - T[g]\| \leqslant L\|f - g\|$ or $(1-L)\|f - g\| \leqslant 0$. But since $L < 1$ this is possible only if $\|f - g\| = 0$ that is $f = g$.

Banach's fixed point theorem can also be interpreted as an existence and uniqueness result for the solution of a fixed point equation

$$f = T[f].$$

Under the assumptions of the theorem, the solution is unique. Furthermore, it can be determined as limit of a sequence (f_n), recursively defined $(f_{n+1} = T[f_n])$ and with arbitrary initial point f_0. In this direction, a useful extension of the theorem is provided by the following

Corollary 12.3.2

Let $(V, \|\cdot\|)$ be a Banach space. Assume $T : V \longrightarrow V$ be such that some iterated $T^N = T \circ \cdots \circ T$ of T is a contraction on V. Then T has a unique fixed point $f \in V$.

Proof. By Banach thm, T^N has a unique fixed point f, that is $T^N[f] = f$. We claim that f is also the unique fixed point of T. First, it is a fixed point for T: indeed,

$$T[f] = T[T^N[f]] = T^{N+1}[f] = T^N[T[f]],$$

that is $T[f]$ is also a fixed point for T^N, but since this has a unique fixed point f it must be $T[f] = f$.
Second, f is unique. If $T[g] = g$ then $T^2[g] = T[g] = g$ and, in general, $T^N[g] = g$, thus g is a fixed point for T^N, so $g = f$ by uniqueness.

An important application of fixed point equations is to the*Cauchy problem for differential equations*. We consider the problem

$$\begin{cases} y'(t) = f(t, y(t)), \\ y(t_0) = y_0. \end{cases} \tag{12.3.4}$$

We assume, for the moment, minimal requirements as $f : D \subset \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ and $f \in \mathscr{C}(D)$. A striking remark is that, by integrating side by side the equation on $[t_0, t]$ we get

$$y(t) - y(t_0) = \int_{t_0}^{t} y'(s)\, ds = \int_{t_0}^{t} f(s, y(s))\, ds,$$

that is, because of the initial condition, y solves

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds. \tag{12.3.5}$$

Thus, if y is a solution of Cauchy problem (12.3.4), then y is a solution of the *integral equation* (12.3.5). Vice versa, assume that $y \in \mathscr{C}$ solves the integral equation (12.3.5) then y solves also the Cauchy problem (12.3.4). Indeed, clearly by (12.3.5) we have $y(t_0) = y_0$, thus the passage condition is verified. What is not immediately evident is that y is differentiable and solves the differential equation. By integral equation (12.3.5) we notice that y is a constant plus the integral function of a continuous function (namely, $f(s, y(s))$). According to the fundamental theorem of integral calculus, this last is differentiable and the derivative is just $f(t, y(t))$. Thus

$$\exists y'(t) = 0 + f(t, y(t)) = f(t, y(t)),$$

and this means that y solves the differential equation.

The conclusion is that solving the Cauchy problem (12.3.4) or the integral equation (12.3.5) is equivalent. Introducing the operator

$$T : \mathscr{C} \longrightarrow \mathscr{C}, \;\; T[y](t) := y_0 + \int_{t_0}^{t} f(s, y(s))\, ds,$$

we are led to show that

$$\exists y \in \mathscr{C} \;:\; y = T[y].$$

In this way we see that the sought solution y is a *fixed point of the operator T*.

Theorem 12.3.3: (global Cauchy–Lipschitz existence and uniqueness)

Assume $f : D = [a, b] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be such that

i) $f \in \mathscr{C}([a, b] \times \mathbb{R})$;

ii) f is Lipschitz continuous in y uniformly in t, that is,

$$|f(t, y_1) - f(t, y_2)| \leqslant C|y_1 - y_2|, \; \forall t \in [a, b], \; y_1, y_2 \in \mathbb{R}.$$

Then, for every passage condition $(t_0, y_0) \in [a, b] \times \mathbb{R}$, there exists a unique $y \in \mathscr{C}^1([a, b])$ solution of the Cauchy problem (12.3.4).

PROOF. Let $V := \mathscr{C}([a,b])$ equipped with usual $\|\cdot\|_\infty$ norm that makes V a Banach space. We apply Corollary 12.3 to

$$T[y] := y_0 + \int_{t_0}^t f(s, y(s))\, ds,$$

showing that some iterated of T is a contraction. Let's start by T itself noticing that

$$T[y](t) - T[\tilde{y}](t) = \int_{t_0}^t \left(f(s, y(s)) - f(s, \tilde{y}(s))\right)\, ds,$$

hence

$$|T[y](t) - T[\tilde{y}](t)| \leqslant \left| \int_{t_0}^t |f(s, y(s)) - f(s, \tilde{y}(s))|\, ds \right| \leqslant \left| \int_{t_0}^t C|y(s) - \tilde{y}(s)|\, ds \right|. \tag{12.3.6}$$

From this we have

$$|T[y](t) - T[\tilde{y}](t)| \leqslant C\|y - \tilde{y}\|_\infty (t - t_0) \leqslant C(b-a)\|y - \tilde{y}\|_\infty,$$

and taking $\max_t$ we have finally

$$\|T[y] - T[\tilde{y}]\|_\infty \leqslant C(b-a)\|y - \tilde{y}\|_\infty, \ \forall y, \tilde{y} \in V.$$

Thus, if $C(b-a) < 1$, T is a contraction on V. The conclusion would follows now by Banach theorem. This would prove existence and uniqueness in the case when $[a,b]$ is short enough, that is $b - a < \frac{1}{C}$. If this is not true we continue with the argument and we pass to T^2:

$$\begin{aligned} |T^2[y](t) - T^2[\tilde{y}](t)| &= |T\left[T[y]\right](t) - T\left[T[\tilde{y}]\right](t)| \overset{(12.3.6)}{\leqslant} \int_{t_0}^t C|T[f](s) - T[g](s)|\, ds \\ &\overset{(12.3.6)}{\leqslant} C^2 \int_{t_0}^t \int_{t_0}^s |y(r) - \tilde{y}(r)|\, dr\, ds. \end{aligned}$$

Iteraring this arugment, we get

$$\begin{aligned} \left|T^N[y](t) - T^N[\tilde{y}](t)\right| &\leqslant C^N \int_{t_0}^t \int_{t_0}^{s_1} \int_{t_0}^{s^2} \cdots \int_{t_0}^{s_{N-1}} |y(s_N) - \tilde{y}(s_N)|\, ds_N\, ds_{N-1} \cdots ds_1 \\ &\leqslant C^N \|y - \tilde{y}\|_\infty \tfrac{(t-t_0)^N}{N!}. \end{aligned}$$

from which

$$\|T^N[y] - T^N[\tilde{y}]\|_\infty \leqslant \frac{(C(b-a))^N}{N!}\|y - \tilde{y}\|_\infty.$$

Now, since $\frac{(C(b-a))^N}{N!} \longrightarrow 0$ when $N \to +\infty$, choosing N large enough, we can make $L := \frac{(C(b-a))^N}{N!} < 1$, so T^N is a contraction and Corollary 12.3 applies.

12.4. Exercises

Exercise 12.4.1 ($**$). *Let $\alpha > 0$ be fixed and define*

$$V_\alpha := \left\{ f \in \mathscr{C}([0, +\infty[) \ : \ \|f\| := \sup_{x \geqslant 0} e^{\alpha x}|f(x)| < +\infty \right\}.$$

i) *Check that V_α is a vector space and $\|\cdot\|$ is a well defined norm on V_α.*
ii) *Is V_α a Banach space?*

Exercise 12.4.2 ($**$). *Let*

$$V := \left\{ f \in \mathscr{C}([0,1]) \ : \ \|f\| := \sup_{t \in]0,1]} \frac{|f(t)|}{t} < +\infty \right\}.$$

i) *Check that $\|\cdot\|$ is a well defined norm on V.*
ii) *Let f_n be defined as*

$$f_n(t) := \begin{cases} nt, & 0 \leqslant t \leqslant \frac{1}{n^2}, \\ \\ \sqrt{t}, & \frac{1}{n^2} \leqslant t \leqslant 1. \end{cases}$$

Is $(f_n) \subset V$? If yes, is $f_n \xrightarrow{\|\cdot\|} f$ for some $f \in V$?

ii) *On V is also defined the $\|\cdot\|_\infty$ norm. Show that $\|\cdot\|$ is stronger than $\|\cdot\|_\infty$. Are the two also equivalent? (prove or disprove)*
iv) *Is V a Banach space under $\|\cdot\|$.*

Exercise 12.4.3 ($**+$). *Let $(X, \mathscr{F}, \mu)$ be a measure space. Prove that $L^\infty(X)$ is a Banach space. (hint: argue as in the proof of completeness of $B(X)$).*

Exercise 12.4.4 ($**$). *Let $V := \mathscr{C}^\infty([0,1])$ equipped with $\|\cdot\|_\infty$ norm. Consider the map $T : V \longrightarrow V$, defined by $T[f] := f'$. Is this map continuous on $(V, \|\cdot\|_\infty)$? Provide a proof, if true, a counterexample, if false.*

Exercise 12.4.5 ($**+$). *Let $V := \mathscr{C}^1([a,b])$ equipped with norm $\|f\| := \|f\|_\infty + \|f'\|_\infty$. Prove that $(V, \|\cdot\|)$ is a Banach space. (hint: take $(f_n) \subset V$ Cauchy sequence, check that both (f_n) and (f_n') are uniformly convergent, then use the fundamental theorem of integral calculus $f_n(x) = f_n(a) + \int_a^x f_n'(y)\, dy$ and pass to the limit . . .)*

Exercise 12.4.6 ($**+$). *Let $(V, \|\cdot\|)$ be a normed space. Show that V is a Banach space if and only if the following property holds:*

$$\forall (u_n) \subset S := \{f \in V \ : \ \|f\| = 1\},\ (u_n) \text{ Cauchy sequence} \implies u_n \xrightarrow{\|\cdot\|} u \in S.$$

Exercise 12.4.7 ($**$). *Let $(V, \|\cdot\|)$ be a Banach space. Let $(f_n) \subset V$. We say that the series $\sum_n f_n$ is convergent if*

$$\exists \lim_{n \to +\infty} \sum_{k=1}^{n} f_k.$$

Show that if $\sum_k \|f_k\|$ converges (in $\mathbb{R}$), then also $\sum_k f_k$ converges (in V).

Exercise 12.4.8 ($*$). *Determine the Banach's theorem recursive sequence (f_n) obtained to solve the Cauchy problem $y'(t) = ty(t)$ with initial condition $y(0) = 1$. (hint: take $f_0 = 0$). What is the conclusion?*

Exercise 12.4.9 ($**+$). *The following integral equation for $f : [-a,a] \longrightarrow \mathbb{R}$ arises in a model for the motion of gas particles on a line:*

$$f(x) = 1 + \frac{1}{\pi} \int_{-a}^{a} \frac{1}{1 + (x-y)^2} f(y)\, dy, \text{ for } -a \leqslant x \leqslant a.$$

For any fixed $a > 0$, show that this equation has a unique, bounded and continuous solution.

LECTURE 13

Hilbert Spaces

Hilbert spaces are particular Banach spaces in which the norm is induced by an *inner product*. The inner product add an euclidean flavour to the structure of normed space through the idea that we can define "angles" between vector.

13.1. Scalar and Hermitian products

There are little (but significant) algebraic differencies on inner products when the field of scalars is $\mathbb{R}$ or $\mathbb{C}$. We start with the real case, a bit simpler:

Definition 13.1.1: (scalar product)

Let V be a vector space on $\mathbb{R}$. A function $\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{R}$ is called **(real) scalar product** if

i) (positivity) $\langle f, f\rangle \geqslant 0$ for every $f \in V$;
ii) (vanishing) $\langle f, f\rangle = 0$ iff $f = 0$;
iii) (linearity) $\langle \alpha f + \beta g, h\rangle = \alpha\langle f, h\rangle + \beta\langle g, h\rangle$, $\forall f, g, h \in V$ and $\forall \alpha, \beta \in \mathbb{C}$;
iv) (symmetry) $\langle f, g\rangle = \langle g, f\rangle$, $\forall f, g \in V$.

We notice that, combining linearity with symmetry, $\langle\cdot,\cdot\rangle$ is linear also in the second argument:

$$\langle h, \alpha f + \beta g\rangle \stackrel{iv)}{=} \langle \alpha f + \beta g, h\rangle \stackrel{iii)}{=} \alpha\langle f, h\rangle + \beta\langle g, h\rangle \stackrel{iv)}{=} \alpha\langle h, f\rangle + \beta\langle h, g\rangle.$$

In other words, $\langle\cdot,\cdot\rangle$ is a *bilinear function* of its arguments.

Example 13.1.2

On $\mathbb{R}^d$,

$$(x_1, \dots, x_d) \cdot (y_1, \dots, y_d) = \sum_{k=1}^{d} x_k y_k$$

is a scalar product.

Example 13.1.3

On $L^2(X)$,

$$\langle f, g\rangle_2 := \int_X fg \, d\mu$$

is a scalar product, with vanishing in the weaker form $\langle f, f\rangle_2 = 0$ iff $f = 0$ a.e..

Proof. First notice that $\langle f, g\rangle_2$ is well defined for $f, g \in L^2(X)$. Indeed, according to CS inequality,

$$\int_X |fg|\, d\mu = \int_X |f||g|\, d\mu \leqslant \|f\|_2 \|g\|_2 < +\infty.$$

Positivity is evident. Vanishing:

$$\langle f, f\rangle_2 = 0, \iff \int_X f^2\, d\mu = 0, \iff f^2 = 0,\ a.e., \iff f = 0,\ a.e.$$

Linearity and simmetry are straightforward.

When V is vector space on $\mathbb{C}$, the previous definition leads to contradictions: indeed, according to positivity, $\langle if, if\rangle \geqslant 0$; however, by linearity $\langle if, if\rangle = i^2\langle f, f\rangle = -\langle f, f\rangle \leqslant 0$, thus $\langle f, f\rangle = 0$ for every $f \in V$. This explains why we have to adjust the definition:

Definition 13.1.4: (hermitian product)

Let V be a vector space over the scalar field $\mathbb{C}$. A function $\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{C}$ is called **hermitian product** if positivity, vanishing, linearity holds true, and moreover $\langle\cdot,\cdot\rangle$ is **anti–symmetric**:

$$\langle f, g\rangle = \overline{\langle g, f\rangle},\ \forall f, g \in V.$$

Easily $\langle\cdot,\cdot\rangle$ is additive in second variable because

$$\langle f, g_1 + g_2\rangle = \overline{\langle g_1 + g_2, f\rangle} \overset{lin.}{=} \overline{\langle g_1, f\rangle} + \overline{\langle g_2, f\rangle} = \langle f, g_1\rangle + \langle f, g_2\rangle.$$

However,

$$\langle f, \lambda g\rangle = \overline{\langle \lambda g, f\rangle} = \overline{\lambda}\overline{\langle g, f\rangle} = \overline{\lambda}\langle f, g\rangle.$$

Example 13.1.5

On $\mathbb{C}^d$

$$(z_1, \ldots, z_d) \cdot (w_1, \ldots, w_d) := \sum_{k=1}^{d} z_k \overline{w_k},$$

is an hermitian product.

Example 13.1.6

Let $(X, \mathscr{F}, \mu)$ be a measure space. On $L^2_{\mathbb{R}}(X)$ ($L^2_{\mathbb{C}}(X)$),

$$\langle f, g\rangle_2 := \int_X fg\, d\mu \quad \left(\langle f, g\rangle_2 := \int_X f\overline{g}\, d\mu\right)$$

is a scalar product (hermitian product), with vanishing in the weaker form $\langle f, f\rangle_2 = 0$ iff $f = 0$ a.e..

A remarkable particular case of L^2 space is the following:

Example 13.1.7: $(*)$

On
$$\ell^2 := \left\{ (x_n) \subset \mathbb{R} \;:\; \sum_n x_n^2 < +\infty \right\},$$
we define
$$\langle (x_n), (y_n) \rangle_{\ell^2} := \sum_n x_n y_n.$$
Then, $\langle \cdot, \cdot \rangle_{\ell^2}$ is a scalar product. Actually, $\ell^2 = L^2(\mathbb{N}, \mathscr{P}(\mathbb{N}), \nu)$ where ν is the counting measure. Here notice that vanishing holds in the strong form:
$$\langle f, f \rangle_{\ell^2} = 0, \quad \Longleftrightarrow \quad \sum_n f_n^2 = 0, \quad \Longleftrightarrow \quad f_n \equiv 0, \quad iff \quad f = 0.$$

Little ℓ^2 space is an interesting example. On one side, it provides a straightforward extension of the euclidean space $\mathbb{R}^m$. On the other side, it is a good space to build examples and counter examples. And finally, it turns out that ℓ^2 is basically the prototype of a generic Hilbert space (we will be more precise on this in the next chapters).

13.2. Norm induced by scalar/hermitian product

In the Euclidean space $\mathbb{R}^m$, the canonical scalar product
$$x \cdot y = \sum_{k=1}^m x_k y_k,$$
is tightly related to the Euclidean norm. Indeed,
$$x \cdot x = \sum_{k=1}^m x_k^2 = \|x\|^2., \quad \Longleftrightarrow \quad \|x\| = \sqrt{x \cdot x}.$$
The same happens in other cases, as for example, for the $L^2(X)$ norm. This is actually true in general: *every scalar/hermitian product induces a natural norm on the vector space where it is defined* setting
$$\|f\| := \sqrt{\langle f, f \rangle}.$$
To show that this is a true norm is the goal of the main result of this section. The proof follows an argument similar to the proof that the Euclidean norm is a norm on $\mathbb{R}^m$. The key ingredient is the abstract version of the Cauchy-Schwarz inequality.

Lemma 13.2.1: (abstract Cauchy–Schwarz inequality)

(13.2.1)
$$|\langle f, g \rangle| \leqslant \|f\| \|g\|, \ \forall f, g \in V.$$

PROOF. (Lemma) Let $g \neq 0$ (otherwise is trivial) and define $\varphi(\alpha) := \|f + \alpha g\|^2 \geqslant 0$. Notice that $\varphi(\alpha) = \langle f + \alpha g, f + \alpha g \rangle = \alpha^2 \|g\|^2 + 2\alpha\langle f, g\rangle + \|f\|^2$, attains its minimum at $\alpha^* = -\frac{\langle f,g\rangle}{\|g\|^2}$. Since $\varphi(\alpha^*) \geqslant 0$ we obtain

$$\frac{\langle f, g\rangle^2}{\|g\|^2} - 2\frac{\langle f, g\rangle^2}{\|g\|^2} + \|f\|^2 \geqslant 0,$$

and by rearranging this we obtain

$$\langle f, g\rangle \leqslant \|f\|\|g\|.$$

Exchanging f with $-f$ and using linearity and homogeneity, we get

$$-\langle f, g\rangle \leqslant \|f\|\|g\|.$$

Combining these two inequalities we have,

$$-\|f\|\|g\| \leqslant \langle f, g\rangle \leqslant \|f\|\|g\|,$$

which is the conclusion.

Proposition 13.2.2

Let $\langle\cdot,\cdot\rangle$ be a scalar/hermitian product on V. Then

$$\|f\| := \sqrt{\langle f, f\rangle},\ f \in V,$$

is a norm on V.

PROOF. (case of scalar product) Clearly, by positivity, $\|f\|$ is well defined and positive for every $f \in V$. Norm vanishing and homogeneity follows directly from vanishing and homogeneity of the scalar product. For triangular inequality we have

$$\begin{aligned}\|f+g\|^2 &= \langle f+g, f+g\rangle = \langle f,f\rangle + \langle f,g\rangle + \langle g,f\rangle + \langle g,g\rangle = \|f\|^2 + \|g\|^2 + 2\langle f,g\rangle \\ &\overset{CS}{\leqslant} \|f\|^2 + \|g\|^2 + 2\|f\|\|g\| = (\|f\| + \|g\|)^2,\end{aligned}$$

from which the conclusion follows.

According to the Cauchy-Schwarz inequality we have that, for $f, g \neq 0$,

$$\frac{|\langle f, g\rangle|}{\|f\|\|g\|} \leqslant 1, \quad \Longleftrightarrow \quad \frac{\langle f, g\rangle}{\|f\|\|g\|} \in [-1, 1].$$

It turns out that, if $V = \mathbb{R}^2$ and $\langle\cdot,\cdot\rangle$ is the canonical scalar product of $\mathbb{R}^2$, the previous quantity is the cosine of the angle $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ made by the two vectors f and g. This explains why we set

$$\cos\theta := \frac{\langle f, g\rangle}{\|f\|\|g\|},$$

and we call this *angle made by f and g* (provided $f, g \neq 0$). In this way the identity

$$\|f+g\|^2 = \|f\|^2 + \|g\|^2 + 2\langle f, g\rangle = \|f\|^2 + \|g\|^2 + 2\|f\|\|g\|\cos\theta,$$

is the general version of the *cosine theorem* of Trigonometry.

Definition 13.2.3

We say that f and g are **orthogonal** (notation $f \perp g$) if $\langle f, g\rangle = 0$.

Example 13.2.4

On $V = L^2([0, 2\pi])$ with usual scalar product $\sin \perp \cos$. Indeed:

$$\langle \sin, \cos\rangle_2 = \int_0^{2\pi} \sin x \cos x \, dx = \left[\frac{(\sin x)^2}{2}\right]_{x=0}^{x=2\pi} = 0.$$

For orthogonal vectors we have the general version of the **Pythagorean theorem**:

$$f \perp g, \implies \|f + g\|^2 = \|f\|^2 + \|g\|^2.$$

Another remarkable identity is the

Proposition 13.2.5: (parallelogram identity)

Let V be a vector space equipped with scalar/hermitian product $\langle \cdot, \cdot\rangle$. Then,

$$\|f + g\|^2 + \|f - g\|^2 = 2\left(\|f\|^2 + \|g\|^2\right), \ \forall f, g \in V. \tag{13.2.2}$$

Proof. We do the proof in the case of scalar product, leaving the case of hermitian product as exercise. We have

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2 + 2\langle f, g\rangle,$$

$$\|f - g\|^2 = \|f\|^2 + \|g\|^2 - 2\langle f, g\rangle,$$

and summing up these identities the conclusion follows.

Norm induces convergence for sequences of vectors. An important fact is the

Proposition 13.2.6

Let V be equipped with a scalar/hermitian product. Then, the scalar/hermitian product is continuous in each component respect to the natural norm, that is:

$$f_n \xrightarrow{\|\cdot\|} f, \implies \langle f_n, g\rangle \longrightarrow \langle f, g\rangle, \ \forall g \in V.$$

Proof. Just notice that

$$|\langle f_n, g\rangle - \langle f, g\rangle| = |\langle f_n - f, g\rangle| \overset{CS}{\leqslant} \|f_n - f\|\|g\| \longrightarrow 0.$$

Definition 13.2.7

A space V equipped with a scalar/hermitian product $\langle \cdot, \cdot\rangle$ is called **Hilbert space** if $(V, \|\cdot\|)$ is a Banach space.

13.3. Exercises

Exercise 13.3.1 ($**+$). *Let* $V := \{ f \in \mathscr{C}^1([0,1];\mathbb{R}) \ : \ f(0) = 0 \}$. *On* V *we define*

$$\langle f, g \rangle := \int_0^1 f'(x) g'(x) \, dx.$$

i) *Check that* $\langle \cdot, \cdot \rangle$ *is a scalar product on* V.
ii) *Determine if* V *is a Hilbert space.*

Exercise 13.3.2 ($*$). *Let* $V := \mathbb{R}[X]$ *be the set of all the polynomials with real coefficients. For* $p, q \in V$ *we define*

$$\langle p, q \rangle_B := \int_{\mathbb{R}} p(x) q(x) e^{-x^2} \, dx.$$

i) *Check that* $\langle \cdot, \cdot \rangle_V$ *is a well defined scalar product on* V.
ii) *Compute* $\langle x^m, x^n \rangle_V$. *For which* $m, n \in \mathbb{N}$ *is* $x^m \perp x^n$*?*
iii) *Solve*

$$\min_{a,b \in \mathbb{R}} \| x^2 - (ax + b) \|_V^2.$$

(hint: compute the norm and apply ordinary calculus tools)

Exercise 13.3.3 ($*$). *Let* $V := L^2_{\mathbb{C}}([0,1]^d)$ *equipped with the standard* L^2 *hermitian product*

$$\langle f, g \rangle := \int_{[0,1]^d} f(x) \overline{g(x)} \, dx.$$

Check that functions $f_n(x) := e^{i 2\pi n \cdot x}$, $n = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ *are orthogonal.*

Exercise 13.3.4 ($**$). *Let* $(V, \langle \cdot, \cdot \rangle)$ *be a scalar/hermitian product space. Show that*

$$\| f \| = \sup_{g \in V \, : \, \|g\| = 1} \langle f, g \rangle, \ \forall f \in V.$$

Exercise 13.3.5 ($**$). *Let* $V := \mathbb{R}^{m \times m}$ *be the set of* $m \times m$ *matrices with real entries and usual algebraic sum and product by scalars. Given* $A, B \in V$, *let*

$$\langle A, B \rangle := \mathrm{Tr}(A^* B),$$

where $\mathrm{Tr}(M)$ *is the trace of matrix* M *(sum of the elements of the diagonal),* A^* *is the transposed matrix of* A. *Check that* $\langle \cdot, \cdot \rangle$ *is a well defined scalar product.*

Exercise 13.3.6 ($**+$). *Prove the Cauchy-Schwarz inequality for an hermitian product. (hint: adapt the proof of the real case, but consider* $\varphi(\alpha) := \| f + \alpha e^{i\theta} g \|^2$ *for a suitable* $\theta \ldots$ *)*

Exercise 13.3.7 ($**+$). *Let* $(V, \langle \cdot, \cdot \rangle)$ *be an Hilbert space. Let* $(f_n) \subset V$ *be such that* $\| f_n \| = 1$ *and*

$$\lim_{n,m \to +\infty} \| f_n + f_m \| = 2.$$

Prove that (f_n) *connverges in* V.

Exercise 13.3.8 ($**+$)**.** *Let $(V, \|\cdot\|)$ be a normed space on real scalar, with norm verifying the parallelogram identity. We define*

$$\langle f, g\rangle := \frac{1}{2}\left(\|f+g\|^2 - (\|f\|^2 + \|g\|^2)\right).$$

Check that the product $\langle \cdot, \cdot \rangle$ verifies

i) *positivity, symmetry, and additivity in the first factor.*
ii) *homogeneity $\langle \lambda f, g\rangle = \lambda \langle f, g\rangle$ for $\lambda = -1$, $\lambda \in \mathbb{N}$, $\lambda \in \mathbb{Z}$, $\lambda \in \mathbb{Q}$ and, finally, extend to $\lambda \in \mathbb{R}$.*

Exercise 13.3.9 ($***$)**.** *Let $(V, \langle \cdot, \cdot, \rangle)$ be an Hilbert space and $(f_n) \subset V$ an orthogonal sequence of vectors, that is $\langle f_n, f_m\rangle = 0$ for all $n \neq m$. Prove that the following statements are equivalent:*

i) *$\sum_n f_n$ converges in V.*
ii) *$\sum_n \|f_n\|^2$ converges in $\mathbb{R}$.*
iii) *$\sum_n \langle f_n, g\rangle$ converges for every $g \in V$.*

LECTURE 14

Orthogonal Projection

Let $(V, \|\cdot\|)$ be a normed space, $U \subset V$ a linear subspace of V. Let $f \in V$. A very important applied problem is the following: *determine the best approximation of f by a vector of U*. Usually f is a function that we wish to approximate, in the best possible way, with a function of class $U \subset V$. Formally, we aim to determine

$$\min_{u \in U} \|f - u\|.$$

Of course, as for every optimization problem, the first issue concerns existence of a solution. Next, characterization of the solution would be welcome. In general, this is problem is very complex. In Hilbert spaces, however, it has a powerful and elegant solution. This is the focus of this Lecture.

14.1. Main Theorem

Theorem 14.1.1

Let $(H, \langle \cdot, \cdot \rangle)$ be an Hilbert space and $U \subset H$ be a closed subspace. Then, for every $f \in H$ there exists a unique $\Pi_U f \in U$ such that

$$\|f - \Pi_U f\| = \min_{u \in U} \|f - u\|. \tag{14.1.1}$$

$\Pi_U f$ is called **orthogonal projection of f on U** and it is characterized by the following orthogonality condition:

$$\langle f - \Pi_U f, u \rangle = 0, \quad \forall u \in U. \tag{14.1.2}$$

Proof. The proof is trivial if $u \in U$: in this case $\Pi_U f = f$. So, let $f \notin U$. The proof is organized as follows:

(1) we prove existence of min, hence of $\Pi_U f$.
(2) we prove that $\Pi_U f$ is unique.
(3) we prove the characterization (14.1.2).

(1) Let

$$\alpha := \inf_{u \in U} \|f - u\|.$$

According to the characterization of best lower bound,

$$\forall n \geqslant 1, \ \exists u_n \in U \ : \ \alpha \leqslant \|f - u_n\| \leqslant \alpha + \frac{1}{n}. \tag{14.1.3}$$

The goal is to prove that (u_n) converges and the limit is the minimum point for $\|f - u\|$. To prove (u_n) converges, we prove that (u_n) is a Cauchy sequence (then convergence follows by H Hilbert space).

.We need an estimate of $\|u_n - u_m\|$. The key ingredient is the parallelogram identity. Indeed,

$$\begin{aligned}\|u_n - u_m\|^2 &= \|(u_n - f) - (u_m - f)\|^2 = 2\left(\|u_n - f\|^2 + \|u_m - f\|^2\right) - \|(u_n - f) + (u_m - f)\|^2 \\ &\leqslant 2\left(\left(\alpha + \frac{1}{n}\right)^2 + \left(\alpha + \frac{1}{m}\right)^2\right) - \|(u_n + u_m - 2f)\|^2 \\ &= 4\alpha^2 + 4\alpha\left(\frac{1}{n} + \frac{1}{m}\right) + 2\left(\frac{1}{n^2} + \frac{1}{m^2}\right) - 4\left\|\frac{u_n + u_m}{2} - f\right\|^2.\end{aligned}$$

Since U is a linear space and $u_n, u_m \in U$, we have $\frac{u_n+u_m}{2} \in U$. Therefore $\left\|\frac{u_n+u_m}{2} - f\right\| \geqslant \alpha$. From this,

$$\|u_n - u_m\|^2 \leqslant 2\alpha\left(\frac{1}{n} + \frac{1}{m}\right) + 2\left(\frac{1}{n^2} + \frac{1}{m^2}\right) \leqslant \varepsilon^2, \ \forall n, m \geqslant N.$$

This means that (u_n) is a Cauchy sequence, thus it is convergent because H is complete by assumption. Now, let $u_n \longrightarrow u^*$. Since

$$\alpha \leqslant \|f - u_n\| \leqslant \alpha + \frac{1}{n}, \implies \alpha \leqslant \|f - u^*\| \leqslant \alpha,$$

that is $\|f - u^*\| = \alpha = \inf_{u \in U} \|f - u\|$. This means the inf is achieved at $u = u^*$, thus it is a minimum.

(2) We show that u^* is unique. Suppose $u^{**} \in U$ is such that $\|f - u^{**}\| = \alpha$. Again by parallelogram identity,

$$\|u^* - u^{**}\|^2 = 2\left(\|u^* - f\|^2 + \|u^{**} - f\|^2\right) + \|u^* + u^{**} - 2f\|^2 = 4\alpha^2 - 4\left\|\frac{u^* + u^{**}}{2} - f\right\|^2 \leqslant 4\alpha^2 - 4\alpha^2 = 0,$$

that is $u^* = u^{**}$. This authorizes to call this unique element $\Pi_U f$.

(3) For every $u \in U$,

$$\|f - \Pi_U f\|^2 \leqslant \|f - (\Pi_U f + u)\|^2 = \|(f - \Pi_U f) + u\|^2 = \|f - \Pi_U f\|^2 + \|u\|^2 + 2\langle f - \Pi_U f, u\rangle,$$

from which

$$\|u\|^2 + 2\langle f - \Pi_U f, u\rangle \geqslant 0, \ \forall u \in U.$$

Replacing u by tu (here $t \in \mathbb{R}$), we have

$$t^2\|u\|^2 + 2t\langle f - \Pi_U f, u\rangle \geqslant 0, \ \forall u \in U, \ \forall t \in \mathbb{R}.$$

Taking $t > 0$ and simplifying, letting $t \longrightarrow 0+$ we have

$$\langle f - \Pi_U f, u\rangle \geqslant 0, \ \forall u \in U.$$

Finally, replacing u with $-u$,

$$\langle f - \Pi_U f, u\rangle \leqslant 0,$$

and by this the conclusion follows.

The projection theorem enlighten the relevance of closed subspaces of an Hilbert space. We remind that a set $S \subset V$, $(V, \|\cdot\|)$ normed space, is closed if and only if it contains limits of convergent sequences of vectors of S, that is

$$\forall (f_n) \subset S, \ : \ f_n \xrightarrow{\|\cdot\|} f, \implies f \in S.$$

In concrete cases this property can be checked directly. A useful fact to know is the

Proposition 14.1.2

If $(V, \|\cdot\|)$ is a normed space and S is a finite dimensional subspace of V, then S is closed.

Proof. Let $v_1, \dots, v_N$ a basis for S. We can assume that vectors $v_1, \dots, v_N$ are linearly independent. For every $f \in S$, there exist a unique array $(f^1, \dots, f^N) \in \mathbb{R}^N$ such that $f = \sum_{k=1}^n f^k v_k$. The map $T : (f^1, \dots, f^N) \longmapsto f$ is linear, bijective (thus invertible) and

$$\|T(f^1, \dots, f^N)\| = \left\| \sum_{k=1}^N f^k v_k \right\| \leqslant \sum_{k=1}^N |f^k| \|v_k\| \leqslant K \|(f^1, \dots, f^N)\|_1,$$

having defined $K := \max_k \|v_k\|$. Now, set

$$\|(f^1, \dots, f^N)\|_* := \|T(f^1, \dots, f^N)\|.$$

It is easy to check that $\|\cdot\|_*$ is well defined, positive, homogeneous and it fulfils the triangular inequality. Moreover, $\|(f^1, \dots, f^N)\|_* = 0$ iff $\sum_{k=1}^N f^k v_k = 0$, and this happens iff $f^k = 0$ for every k. So, $\|\cdot\|_*$ fulfils also vanishing. In other words, $\|\cdot\|_*$ is a norm on $\mathbb{R}^N$, and since all the norms on $\mathbb{R}^N$ are equivalent,

$$\exists M > 0, \quad \|(f^1, \dots, f^N)\|_1 \leqslant M \|(f^1, \dots, f^N)\|_*.$$

Let now $(f_n) \subset S$ with $f_n \xrightarrow{\|\cdot\|} f$. The goal is to prove that $f \in S$. We start noticing that (f_n) is a Cauchy sequence w.r.t. $\|\cdot\|$. Since each $f_n \in S$, we can write

$$f_n = \sum_{k=1} f_n^k v_k = T \underbrace{(f_n^1, \dots f_n^N)}_{=:g_n \in \mathbb{R}^N}.$$

for suitable f_n^k. Therefore,

$$\|g_n - g_m\|_1 \leqslant M \|g_n - g_m\|_* = M \|T(g_n - g_m)\| = M \|Tg_n - Tg_m\| = M \|f_n - f_m\|,$$

from which (g_n) is a Cauchy sequence in $\mathbb{R}^N$ w.r.t. $\|\cdot\|_1$ norm, and since this is a Banach space, $g_n \xrightarrow{\|\cdot\|_1} g$, for some $g = (g^1, \dots, g^N) \in \mathbb{R}^N$. From this it follows that

$$\|f_n - Tg\| = \|Tg_n - Tg\| \leqslant K \|g_n - g\|_1 \longrightarrow 0, \quad \Longrightarrow \quad f_n \xrightarrow{\|\cdot\|} Tg = \sum_{k=1}^N g^k v_k,$$

and since the limit is unique, we conclude that $f = Tg = \sum_{k=1}^N g^k v_k \in S$.

Example 14.1.3

Determine the best approximation of x^2 through a first degree polynomial under $L^2([0,1])$ norm.

Proof. Let $V = L^2([0,1])$. We have to minimize the $L^2([0,1])$ distance between x^2 and $ax + b$,that is to determine

$$\min_{a,b \in \mathbb{R}} \left(\int_0^1 |x^2 - (ax + b)|^2 \, dx \right)^{1/2}.$$

Let $U := \{ax + b \ : \ a, b \in \mathbb{R}\} = \mathrm{Span}(1, x)$. U is finite dimensional, hence it is closed. Therefore, the best $L^2([0,1])$ approximation of x^2 through an element of U is $\Pi_U x^2$. Now, $\Pi_U x^2 = Ax + B$ for suitable

A and B, to be determined in such a way that orthogonality condition

$$\langle x^2 - \Pi_U x^2, w\rangle = \int_0^1 (x^2 - \Pi_U x^2)w(x)\,dx = 0,\ \forall u \in U.$$

Since U is generated by 1 and x through linear combinations, the orthogonality condition can be reduced to just two conditions

$$\begin{cases} \int_0^1 (x^2 - (Ax+B))1\,dx = 0, \\ \\ \int_0^1 (x^2 - (Ax+B))x\,dx = 0, \end{cases} \iff \begin{cases} \frac{1}{3} - \frac{A}{2} - B = 0, \\ \\ \frac{1}{4} - \frac{A}{3} - \frac{B}{2} = 0. \end{cases}$$

By solving the linear system we find $A = 1$ and $B = -\frac{1}{6}$, that is $\Pi_U x^2 = x - \frac{1}{6}$.

Warning 14.1.4

The assumption U closed is essential for the projection theorem. Indeed, on ℓ^2 take $U := \{(x_n) \ :\ \exists N,\ x_n \equiv 0,\ \geqslant N\}$. It is easy to check that U is a linear subspace of ℓ^2 and that it is not closed: the sequence (f_n), where $f_n = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, \ldots\right) \in U$, we have $f_n \stackrel{\ell^2}{\longrightarrow} f = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, \frac{1}{n+1}, \ldots\right)$ because

$$\|f_n - f\|_{\ell^2}^2 = \left\| \left(0, \ldots, 0, \frac{1}{n+1}, \frac{1}{n+2}, \ldots\right) \right\|_{\ell^2}^2 = \sum_{k=n+1}^{\infty} \frac{1}{k^2} \longrightarrow 0,\ n \to +\infty.$$

This also says that

$$\inf_{u\in U} \|f - u\|_{\ell^2} \leqslant \|f - f_n\|_{\ell^2} \longrightarrow 0, \quad \Longrightarrow \quad \inf_{u \in U} \|f - u\|_{\ell^2} = 0.$$

However, the minimum is not achieved otherwise there would be $u \in U$ such that $\|f - u\|_{\ell^2} = 0$, that is $f = u \in U$, but $f \notin U$.

14.2. Orthogonal complement

Definition 14.2.1

Let V be an inner product space, $U \subset V$ a subspace. We call **orthogonal complement** of U,

$$U^\perp := \{v \in V \ :\ \langle v, u\rangle = 0\}.$$

It is easy to check that $U^\perp$ is **always** a closed linear subspace of V, no matter whether U is closed or not. Indeed, if $(v_n) \subset U^\perp$ and $v_n \longrightarrow v$, then

$$0 = \langle v_n, u\rangle \longrightarrow \langle v, u\rangle,\ \forall u \in U, \quad \Longrightarrow \quad v \in U^\perp.$$

Therefore, if V is a Hilbert space, the orthogonal projection $\Pi_{U^\perp}$ is always well defined. It is sometimes useful to know that

Proposition 14.2.2

Let H be an Hilbert space, U a closed subspace. Then

$$\Pi_{U^\perp} f = f - \Pi_U f.$$

PROOF. Let's check that i) $f - \Pi_U f \in U^\perp$ and ii) $f - \Pi_U f$ verifies the othogonality condition (14.1.2) for $U^\perp$. Indeed, by (14.1.2), we have

$$\langle f - \Pi_U f, u \rangle = 0, \ \forall u \in U, \implies f - \Pi_U f \in U^\perp.$$

Moreover, since $\Pi_U f \in U$, we have

$$\langle f - (f - \Pi_U f), v \rangle = \langle \Pi_U f, v \rangle = 0, \ \forall v \in U^\perp.$$

Therefore, i) and ii) areverified and the conclusion follows.

Warning 14.2.3

From the previous proposition, apparently,

$$\Pi_U f = f - \Pi_{U^\perp} f,$$

is the orthogonal projection on U. We stress once more the fact that **this is true only if U is closed**. Indeed, take $H = \ell^2$ and the subspace $U = \{(f_n) : \exists N, \ f_n \equiv 0, \ n \geqslant N\}$. Notice that

$$v \in U^\perp, \iff \langle v, u \rangle = 0, \ \forall u \in U.$$

We claim that $v = 0 = (0, \dots)$. Indeed, if $v = (v_n)$ with $v_N \neq 0$ for some N, then taking $u = (\delta_{Nn} \dots) \in U$ we would have

$$0 = \langle v, u \rangle = \sum_n v_n \delta_{Nn} = v_N,$$

which contradicts $v_N \neq 0$. Therefore, $U^\perp = \{0\}$ and, as a consequence, $\Pi_{U^\perp} f \equiv 0$. Therefore, if $\Pi_U f = f - \Pi_{U^\perp} f = f$, but this is possible iff $f \in U \subsetneq \ell^2$. Therefore, $\Pi_U f$ is not defined unless $f \in U$.

14.3. Exercises

Exercise 14.3.1 ($*$). *Let $f(x) := \cos x \in L^2([0, 2\pi])$. Determine the best possible second degree polynomial closest to f in the $L^2([0, 2\pi])$ norm.*

Exercise 14.3.2 ($**$). *Solve*

$$\min_{a,b,c \in \mathbb{R}} \int_{-1}^{1} |x^3 + ax^2 + bx + c|^2 \, dx.$$

Exercise 14.3.3 ($**$). *Solve*

$$\min_{a,b \in \mathbb{R}} \int_0^{+\infty} |e^{-x} - (ae^{-2x} + be^{-3x})|^2 \, dx$$

Exercise 14.3.4 $(**)$. *Solve*

$$\max_{f \in L^2([0,1])\,:\,\int_0^1 f^2\,dx=1} \int_0^1 f(x)e^x\,dx.$$

Exercise 14.3.5 $(**+)$. Let $V := L^2(\mathbb{R})$ equipped with usual real scalar product. Consider

$$U := \{f \in V \;:\; f(-x) = f(x),\ \text{a.e. } x \in \mathbb{R}\}.$$

i) Show that U is closed (hint: recall that $f_n \xrightarrow{L^2} f$ does not imply that (f_n) converges pointwise but...).
ii) Check that $\Pi_U f(x) = \frac{1}{2}(f(x) + f(-x))$.

Exercise 14.3.6 $(**)$. *Let $H := L^2([0,1])$ equipped with usual scalar product and set*

$$U := \left\{ f \in H \;:\; \int_0^1 f(x)\,dx = 0 \right\}.$$

i) *Is U a closed subspace of H?*
ii) *Determine $U^{\perp}$.*

Exercise 14.3.7 $(**)$. Let $(V, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, $\phi, \psi \in V$ two linearly independent unit vectors (that is $\|\phi\| = \|\psi\| = 1$). Let also $W_1 := \{\alpha\phi \;:\; \alpha \in \mathbb{R}\}$, $W_2 := \{\beta\psi \;:\; \beta \in \mathbb{R}\}$ and $W_1 + W_2 = \{w_1 + w_2 \;:\; w_1 \in W_1,\ w_2 \in W_2\}$. Clearly, W_1 and W_2 are closed. We accept $W_1 + W_2$ is closed as well.

i) Determine the orthogonal projections Π_{W_1} and Π_{W_2}.
ii) Determine $\Pi_{W_1+W_2}$.
iii) Under which condition on ϕ, ψ is it true that $\Pi_{W_1+W_2} = \Pi_{W_1} + \Pi_{W_2}$?

Exercise 14.3.8 $(**+)$. *Let $(V, \langle\cdot,\cdot\rangle)$ be an Hilbert space, $U \subset V$ a linear subspace. Discuss under which conditions is $(U^{\perp})^{\perp} = U$, proving what true, and disproving by an example what false.*

Exercise 14.3.9 $(**+)$. *Let $(V, \langle\cdot,\cdot\rangle)$ be an Hilbert space, $U \subset V$ be a closed subspace of V. Let $\Pi_U f$ be the orthogonal projection of f on U.*

i) *Prove that $\Pi_U(f+g) = \Pi_U f + \Pi_U g$ and $\Pi_U(\alpha f) = \alpha \Pi_U f$ for every $f, g \in V$ and $\alpha \in \mathbb{R}$.*
ii) *$\Pi_U(\Pi_U f) = \Pi_U f$, for every $f \in V$.*
iii) *$\langle \Pi_U f, g\rangle = \langle f, \Pi_U g\rangle$, for every $f, g \in V$.*

LECTURE 15

Orthonormal bases

In Linear Algebra, a basis is a family of linearly independent vectors such that any other vector can be expressed as (finite) linear combination of vectors of the basis. For an infinite dimensional space, this definition implies an extremely large (uncountable) set of vectors. It is preferable to deal with an infinite but countable basis, accepting that every vector might be expressed as infinite linear combination of the basis' vectors. Since the spaces we work with are normed spaces (at least), it is not a problem to deal with infinite sums as limit of finite sums.

15.1. General definition and properties

We start by the

Definition 15.1.1

Let $(V, \|\cdot\|)$ be a normed space. Given a sequence of vectors $(f_n) \subset V$, we set

$$\sum_{n=0}^{\infty} f_n := \lim_{N\to+\infty} \sum_{n=0}^{N} f_n,$$

provided the limit exists in V.

A sufficient condition to ensure convergence of a series of vectors in a Banach space is the *normal convergence test*:

Proposition 15.1.2: (Weierstrass)

Let $(V, \|\cdot\|)$ be a Banach space. Then,

$$\sum_n \|f_n\| < +\infty, \implies \sum_n f_n \text{converges}.$$

Proof. We check that the sequence of partial sums $s_n := \sum_{k=0}^{n} f_k$ is a Cauchy sequence. Notice that, if $n > m$,

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} f_k \right\| \leqslant \sum_{k=m+1}^{n} \|f_k\| = \sigma_n - \sigma_m, \quad \left(\text{where } \sigma_n := \sum_{k=0}^{n} \|f_k\| \right). \tag{15.1.1}$$

Since, by assumption, $\sum_n \|f_n\|$ is convergent, the corresponding sequence of partial sums (σ_n) is a Cauchy sequence, so

$$\forall \varepsilon > 0,\ \exists N,\ :\ |\sigma_n - \sigma_m| \leqslant \varepsilon,\ \forall n, m \geqslant N.$$

In particular, for $n > m \geqslant N$, being $\sigma_n \geqslant \sigma_m$ we have

$$\|s_n - s_m\| \overset{(15.1.1)}{\leqslant} \sigma_n - \sigma_m = |\sigma_n - \sigma_m| \leqslant \varepsilon,\ \forall n > m \geqslant N,$$

and this is the Cauchy property for (s_n). Since the space $(V, \|\cdot\|)$ is complete, the sequence (s_n) is convergent and we have the conclusion.

In a Hilbert space, the Weierstrass test can be sharpened:

Proposition 15.1.3

Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then, if (f_n) is a sequence of orthogonal vectors,

$$\sum_n f_n \text{ converges} \iff \sum_n \|f_n\|^2 < +\infty. \tag{15.1.2}$$

Proof. Let $s_n := \sum_{k=0}^n f_k$ be the $n-$th partial sum of the series $\sum_n f_n$. By the Pythagorean theorm we have that, for $n > m$,

$$\|s_n - s_m\|^2 = \left\| \sum_{k=m+1}^n f_k \right\|^2 = \sum_{k=m+1}^n \|f_k\|^2 = \sigma_n^2 - \sigma_m^2, \quad \left(\text{where } \sigma_n^2 := \sum_{k=0}^n \|f_k\|^2\right).$$

It is therefore clear that (s_n) is a Cauchy sequence in H iff (σ_n^2) is a Cauchy sequence in $\mathbb{R}$. From this the conclusion follows.

We are now ready for the

Definition 15.1.4

Let $(V, \langle\cdot,\cdot\rangle)$ be a vector space equipped with an inner product. A set of vectors (e_n) is called
- **orthonormal system** if $\langle e_i, e_j\rangle = \delta_{ij}$;
- **orthonormal basis** if it is an orthonormal system and

$$\forall f \in V,\ \exists (c_n) \subset \mathbb{R}\ (\mathbb{C})\ :\ f = \sum_{n=0}^\infty c_n e_n.$$

Example 15.1.5: $(*)$

On $H = \ell^2$ (see Example 13.1.7 for the definition), let $e_n := (\delta_{nm})_m$. Then (e_n) is an orthonormal basis for ℓ^2.

Notice that, if $f = \sum_n c_n e_n$ then

$$\langle f, e_j \rangle = \langle \sum_n c_n e_n, e_j \rangle = \langle \lim_n \sum_{k=0}^n c_k e_k, e_j \rangle = \lim_n \langle \sum_{k=0}^n c_k e_k, e_j \rangle = \lim_n \sum_{k=0}^n c_k \langle e_k, e_j \rangle = \sum_n c_n \delta_{nj} = c_j.$$

Thus,

$$f = \sum_n \langle f, e_n \rangle e_n. \tag{15.1.3}$$

The series at r.h.s of (15.1.3) is called **abstract Fourier series**, $(\langle f, e_n \rangle)$ are called **Fourier coefficients**. Notice also that, in this case,

$$\|f\|^2 = \langle f, f \rangle = \sum_{n,m} \langle f, e_n \rangle \langle f, e_m \rangle \langle e_n, e_m \rangle = \sum_n |\langle f, e_n \rangle|^2,$$

which is called **Parseval identity**.

Given a finite or countable orthonormal set $(e_n) \subset H$, we call

$$\mathrm{Span}(e_n) := \left\{ \sum_n c_n e_n \in H \ : \ (c_n) \subset \mathbb{R}\ (\mathbb{C}) \right\}.$$

Notice that ,according to the convercenge test (15.1.2), we have

$$\sum_n c_n e_n \in H, \quad \Longleftrightarrow \quad \sum_n \|c_n e_n\|^2 = \sum_n |c_n|^2 < +\infty.$$

Therefore,

$$\mathrm{Span}(e_n) = \left\{ \sum_n c_n e_n \ : \ \sum_n |c_n|^2 < +\infty \right\}.$$

This shows an interesting fact:

Proposition 15.1.6

Let H be an Hilbert space and (e_n) an orthonormal system. Then $\mathrm{Span}(e_n)$ is isometrically equivalent to ℓ^2. In particular, $\mathrm{Span}(e_n)$ is closed in H. If (e_n) is an orthonormal basis, then H itself is isometrically equivalent to ℓ^2.

Proof. Consider the map $T : \mathrm{Span}(e_n) \longrightarrow \ell^2$ defined by

$$T\left(\sum_n c_n e_n \right) = (c_n).$$

Then

$$\left\| T\left(\sum_n c_n e_n \right) \right\|_{\ell^2}^2 = \|(c_n)\|_{\ell^2}^2 = \sum_n |c_n|^2 = \left\| \sum_n c_n e_n \right\|_H^2,$$

that is $\|Tf\|_{\ell^2} = \|f\|_H$ for every $f \in \mathrm{Span}(e_n)$. This means that T preserves norm, that is is an isometry, between $\mathrm{Span}(e_n)$ and ℓ^2. Since $\ell^2 = L^2(\mathbb{N}, \nu)$ (ν is the counting measure) it is Hilbert space, $\mathrm{Span}(e_n)$ is also an Hilbert space, in particular it is closed.

Remark 15.1.7

We may say that ℓ^2 is in fact the prototype of an Hilbert space with an orthonormal basis.

Orthonormal bases are useful to have a representation of orthogonal projection

Proposition 15.1.8

Let $(H, \langle\cdot,\cdot\rangle)$ be an Hilbert space and U a closed subspace of H. If (e_n) is an orthonormal basis for U then

$$\Pi_U f = \sum_n \langle f, e_n\rangle e_n. \tag{15.1.4}$$

PROOF. Clearly $\Pi_U f$ defined by (15.1.4) belongs to U. We prove that $\Pi_U f$ fulfils orthogonality condition (14.1.2). Let $u \in U$. Since (e_n) is an orthonormal basis for U, $u = \sum_n \langle w, e_n\rangle e_n$. Then,

$$\langle f - \Pi_U f, u\rangle = \langle f, u\rangle - \langle \Pi_U f, u\rangle = \sum_n \langle u, e_n\rangle\langle f, e_n\rangle - \sum_n \langle f, e_n\rangle\langle e_n, u\rangle = 0.$$

Corollary 15.1.9: Bessel inequality

Let (e_n) an orthonormal system of vectors for $(H, \langle\cdot,\cdot\rangle)$ Hilbert space. Then

$$\sum_n |\langle f, e_n\rangle|^2 \leqslant \|f\|^2, \ \forall f \in V. \tag{15.1.5}$$

PROOF. Let $U := \mathrm{Span}(e_n)$. U is a closed subspace of V, and

$$\Pi_U f = \sum_n \langle f, e_n\rangle e_n.$$

According Pythagorean theorem

$$f = \Pi_U f + (f - \Pi_U f), \implies \|f\|^2 = \|\Pi_U f\|^2 + \|f - \Pi_U f\|^2 \geqslant \|\Pi_U f\|^2 = \sum_n |\langle f, e_n\rangle|^2.$$

15.2. Test for orthonormal bases

Under which conditions an orthonormal system (e_n) is also a basis? Of course, if we can prove that every f is sum of its Fourier series under (e_n), we are done. Next proposition provides an intrinsic test:

Proposition 15.2.1

Let $(H, \langle\cdot,\cdot\rangle)$ be an Hilbert space. Necessary and sufficient condition for (e_n) orthonormal system to be a basis is

$$\langle f, e_n\rangle = 0, \ \forall n \in \mathbb{N}, \implies f = 0. \tag{15.2.1}$$

PROOF. Necessity: assume (e_n) is a basis. If $\langle f, e_n\rangle = 0$ for all n then, by (15.1.3), $f = \sum_n \langle f, e_n\rangle e_n = 0$.
Sufficiency: assume (15.2.1) holds. Let $U := \mathrm{Span}(e_n)$. It is not difficult to check that U is a closed subspace (we accept this). Let $\Pi_U f$ be the orthogonal projection over U. We have $\Pi_U f = \sum_n \langle f, e_n\rangle e_n$. Since

$$f = \Pi_U f + (f - \Pi_U f) = \sum_n \langle f, e_n\rangle e_n + (f - \Pi_U f),$$

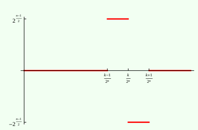
the conclusion follows once we prove $f - \Pi_U f = 0$. Now, since by orthogonality condition (14.1.2) we have $\langle f - \Pi_U f, e_n\rangle = 0$ for every n, by (15.2.1) this implies $f - \Pi_U f = 0$.

Here is an example how density test (15.2.1) works:

Proposition 15.2.2: Haar basis

On $L^2([0,1])$ equipped with usual scalar product and define the **Haar functions**

$$e_0(x) \equiv 1, \quad e_{k/2^n}(x) = \begin{cases} 2^{\frac{n-1}{2}}, & \frac{k-1}{2^n} \leqslant x < \frac{k}{2^n}, \\ -2^{\frac{n-1}{2}}, & \frac{k}{2^n} \leqslant x < \frac{k+1}{2^n}, \\ 0, & \text{otherwise}. \end{cases} \quad k = 1, \dots, 2^n - 1,\ k \text{ odd},\ n \geqslant 1,$$



Then $(e_0, e_{\frac{k}{2^n}})_{k,n}$ is a basis for $L^2([0,1])$. In particular,

$$f = \langle f, e_0\rangle + \sum_{n=0}^{\infty} \sum_{k=1,\ k\ odd}^{2^n-1} \langle f, e_{k/2^n}\rangle e_{k/2^n}, \ \forall f \in L^2([0,1]).$$

(this formula is among the simplest *wavelet reconstruction formula* of a function f).

PROOF. Orthonormality can be easily checked as exercise. Assume that $f \perp e_0, e_{k/2^n}$ for all k, n. Notice first that

$$0 = \langle f, e_{k/2^n}\rangle_2 = 2^{\frac{n-1}{2}} \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(x)\, dx - 2^{\frac{n-1}{2}} \int_{\frac{k}{2^n}}^{\frac{k+1}{2^n}} f(x)\, dx, \quad \Longrightarrow \quad \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(x)\, dx = \int_{\frac{k}{2^n}}^{\frac{k+1}{2^n}} f(x)\, dx.$$

Therefore

$$0 = \langle f, e_0\rangle_2 = \int_0^1 f(x)\, dx = 2\int_0^{\frac{1}{2}} f(x)\, dx = 4\int_0^{\frac{1}{4}} f(x)\, dx = \ldots = 2^n \int_0^{\frac{1}{2^n}} f(x)\, dx,$$

and again, by previous identity, $\int_{\frac{k}{2^n}}^{\frac{k+1}{2^n}} f(x)\, dx = 0$ for every (n,k). By this it is easy to deduce that

$$\int_a^b f(x)\, dx = 0, \ \forall a, b \in \left\{ \frac{k}{2^n} \ :\ n \in \mathbb{N},\ k \in \{0, 1, \dots, 2^n\} \right\} =: \mathbb{D},$$

the set of dyadic numbers. Now, it is known (we accept it) that $\mathbb{D}$ is dense in $[0,1]$. Therefore, the previous identity extends to every $a, b \in [0,1]$. It is now a standard job to conclude that $\int_E f = 0$ for every $E \subset [0,1]$ Lebesgue measurable, and by this the conclusion follows.

15.3. Gram–Schmidt orthogonalization algorithm

Does an orthonormal basis always exist? A first remark is the following: if H has an orthonormal basis then, the set

$$S := \left\{ \sum_n q_n e_n \ : \ (q_n) \subset \mathbb{Q} \right\},$$

is countable (same cardinality of $\mathbb{N}$) and every $f = \lim_n f_n$ with $(f_n) \in S$ (we skip the details). In other words, *there is a countable set $S \subset H$* **dense** *in H.*

Definition 15.3.1

We say that a normed space $(V, \|\cdot\|)$ is **separable** if there exists a countable set S dense in V.

Thus, to admit an orthonormal basis, the space must be separable. This condition is also sufficient:

Theorem 15.3.2: (Gram–Schmidt)

Let $(H, \langle \cdot, \cdot \rangle)$ be a **separable** Hilbert space. Then, H admits an orthonormal basis. This can be constructed in the following way: if (u_n) is any set of linearly independent vectors dense in H, defining

$$e_0 := \frac{u_0}{\|u_0\|}, \quad e_n = \frac{u_n - \sum_{j=0}^{n-1} \langle u_n, e_j \rangle e_j}{\left\| u_n - \sum_{j=0}^{n-1} \langle u_n, e_j \rangle e_j \right\|}, \ (n \geqslant 1), \tag{15.3.1}$$

we have that (e_n) is an orthonormal basis for H.

Proof. First step. Since H is separable, there exists a countable set (u_n) dense in H. We define S as the set of **finite** linear combinations of u_n. It is clear that $S \supset (u_n)$, thus S is dense in H and also we can eliminate u_n who are linearly dependent from others obtaining the same S. In other words, we have that there exists (u_n) of linearly independent vectors such that

$$S = \{\text{finite linear combinations of } (u_n)\}$$

is dense in H.

We now check that the definitions (15.3.1) are well posed and they are an orthonormal basis for H. We argue by induction. For $n = 0$, e_0 is well defined because $u_0 \neq 0$ (by linear independence of vectors of S). Furthermore, $\mathrm{Span}(e_0) = \mathrm{Span}(u_0)$. Assume now the check of good position and orthonormality has been done on $e_0, \dots, e_n$ and $\mathrm{Span}(e_0, \dots, e_n) = \mathrm{Span}(u_0, \dots, u_n)$. Since $U_n = \mathrm{Span}(u_0, \dots, u_n)$ is closed, orthogonal projection Π_{U_n} is well defined and

$$\Pi_{U_n} f = \sum_{k=0}^{n} \langle f, e_k \rangle e_k.$$

Now, we claim that $u_{n+1} - \Pi_{U_n} u_{n+1} \neq 0$. If not, $u_{n+1} = \Pi_{U_n} u_{n+1} = \sum_{k=0}^{n} \langle u_{n+1}, e_k \rangle e_k \in \mathrm{Span}(e_0, \ldots, e_n) = \mathrm{Span}(u_0, \ldots, u_n)$. But this is is in contradiction with linear independence. Thus $\|u_{n+1} - \Pi_{U_n} u_{n+1}\| > 0$ and vector e_{n+1} is well defined. Clearly $\|e_{n+1}\| = 1$, and since $e_{n+1} \propto u_{n+1} - \Pi_{U_n} u_{n+1} \perp U_n$, we have that $e_{n+1} \perp \mathrm{Span}(e_0, \ldots, e_n)$, thus $e_0, \ldots, e_{n+1}$ are orthonormal. Finally, $\mathrm{Span}(e_0, \ldots, e_{n+1}) = \mathrm{Span}(u_0, \ldots, u_{n+1})$. This proves that (e_n) is an orthonormal system. To check that it is also a basis for H we apply the test for orthonormal bases provided by Proposition 15.2. Let $f \in H$ be such that $\langle f, e_n \rangle = 0$ for every n. Since S is dense in H, $(s_n) \subset S$ such that $s_n \longrightarrow f$. Since $S \subset \mathrm{Span}(u_n) = \mathrm{Span}(e_n)$, we have that $\langle f, s_n \rangle = 0$ for every n. But then, letting $n \longrightarrow +\infty$, we have $\langle f, f \rangle = 0$, that is $\|f\|^2 = 0$, from which $f = 0$.

15.3.1. Hermite polynomials. $L^2(\mathbb{R})$ is a very common framework in many applied problems. In this Section, we will compute an orthonormal basis for it. To attack the problem, we start by changing slightly the setting by considering

$$H := \left\{ f : \mathbb{R} \longrightarrow \mathbb{R} \ : \ \int_{\mathbb{R}} |f(x)|^2 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx < +\infty \right\} = L^2(\mathbb{R}, \mathcal{N}),$$

that is *the L^2 space respect to the probability measure $d\mathcal{N} := \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx$ called also* **standard gaussian**. H is an Hilbert space with scalar product and norm

$$\langle f, g \rangle := \int_{\mathbb{R}} f(x) g(x) \, \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx, \quad \|f\|^2 = \int_{\mathbb{R}} |f(x)|^2 \, \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx.$$

We notice that,

$$1, x, x^2, \ldots, x^n, \ldots \in L^2(\mathbb{R}, \mathcal{N}).$$

In general, the x^n are not orthogonal because $\langle x^n, x^m \rangle = \int_{\mathbb{R}} x^{n+m} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = 0$ iff $n+m$ is odd. However, we can apply the Gram-Schmidt algorithm to "orthogonalize" powers. Set

$$e_n = \frac{1}{\alpha_n} \left(x^n - \sum_{j=0}^{n-1} \langle x^n, e_j \rangle e_j \right) =: \frac{H_n}{\|H_n\|}.$$

So, for instance

$$e_0 = \tfrac{1}{\alpha_0} 1, \qquad \alpha_0^2 = \int_{\mathbb{R}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy = 1,$$

$$e_1 = \tfrac{1}{\alpha_1} \left(x - \langle x, e_0 \rangle e_0 \right) = \tfrac{1}{\alpha_1} \left(x - \int_{\mathbb{R}} y \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \right) = \tfrac{1}{\alpha_1} x, \quad \alpha_1^2 = \int_{\mathbb{R}} x^2 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = 1.$$

Therefore $e_0(x) \equiv 1,\ e_1(x) = x$. Again

$$e_2 = \frac{1}{\alpha_2}\left(x^2 - \langle x^2, e_0\rangle e_0 - \langle x^2, e_1\rangle e_1\right) = \frac{1}{\alpha_2}\left(x^2 - \langle x^2, 1\rangle e_0 - \langle x^2, x\rangle e_1\right) = \frac{1}{\alpha_2}(x^2-1).$$

The value of α_2 is

$$\alpha_2^2 = \int_{\mathbb{R}} (x^2-1)^2 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx = \int_{\mathbb{R}} x^4 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx - 2\int_{\mathbb{R}} x^2 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx + \int_{\mathbb{R}} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx = 3-2+1 = 2.$$

In conclusion $e_2(x) = \frac{1}{\sqrt{2}}(x^2-1)$. It is clear that we can compute $e_3, e_4, \dots$ in this way, but it looks to be difficult to have a "quick" recipe to compute e_n for every n. To do this, notice first that the H_n are polynomials called *Hermite polynomials*. Precisely, $H_n(x) = x^n + p_{n-1}(x)$, where p_{n-1} is an $n-1$-th degree polynomial. In particular, H_n has degree $= n$ and

$$\mathrm{Span}(H_0, ..., H_n) = \mathrm{Span}(1, x, \dots, x^n).$$

Furthermore, by construction $H_n \perp H_m,\ n \neq m$. In particular,

$$H_n \perp \mathrm{Span}(H_0, ..., H_{n-1}) = \mathrm{Span}(1, \dots, x^{n-1}).$$

Let's see how to determine more efficiently the H_n. The first step is the

Proposition 15.3.3

$$H_n' = nH_{n-1}. \tag{15.3.2}$$

Proof. Notice that

$$H_n' = nx^{n-1} + p_{n-1}' = nH_{n-1} + q_{n-2} = nH_{n-1} + \sum_{j=0}^{n-2} c_j H_j.$$

Now, multiplying both sides by H_k in the scalar product, we obtain

$$\langle H_n', H_k\rangle = n\langle H_{n-1}, H_k\rangle + c_k\|H_k\|^2 = c_k\|H_k\|^2,\ \forall k \leqslant n-2.$$

On the other side

$$\langle H_n', H_k\rangle = \int_{\mathbb{R}} H_n' H_k \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\ dx \overset{parts}{=} -\int_{\mathbb{R}} H_n(H_k' - xH_k)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\ dx = \langle H_n, H_k' - xH_k\rangle = 0,$$

because $H_k' - xH_k \in \mathrm{Span}(1, x, \dots, x^{n-1})$ if $k \leqslant n-2$. The moral is $c_k = 0$ for every $k = 0, \dots, n-2$, from which the conclusion follows.

The (15.3.2) is not a good rule to compute H_n because even if we know H_{n-1} we should proceed with an integration, which is not a problem being H_{n-1} a polynomial but it involves a free constant to be determined by other conditions. Notice that, in proving the previous Proposition we proved the *integration by parts formula*

$$\langle p', q\rangle = \langle p, (xq - q')\rangle,\ \forall p, q \text{ polynomials}. \tag{15.3.3}$$

Indeed,

$$\begin{aligned}\langle p', q\rangle &= \int_{\mathbb{R}} p'(x)q(x)e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}} = -\int_{\mathbb{R}} p(x)\left(q(x)e^{-\frac{x^2}{2}}\right)'\frac{dx}{\sqrt{2\pi}}\\ &= -\int_{\mathbb{R}} p(x)\left(q'(x)-xq(x)\right)e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}} = \langle p, xq-q'\rangle.\end{aligned}$$

By this we obtain easily the

Proposition 15.3.4

$$H_{n+1} = xH_n - H_n'. \tag{15.3.4}$$

Proof. Consider the polynomial $xH_n - H_n'$: we already proved that $xH_n - H_n' \perp H_k$ for $k \leqslant n$. Hence, $xH_n - H_n' \perp e_k$ for $k \leqslant n$ and because this is an orthonormal base,

$$xH_n - H_n' = c_{n+1}e_{n+1} \equiv \widetilde{c}_{n+1}H_{n+1}.$$

Noticed that $xH_n - H_n' = x(x^n + p_{n-1}) - (x^n + p_{n-1})' = x^{n+1} + (xp_{n-1} - nx^{n-1} - p_{n-1}') \equiv x^{n+1} + p_n$, we deduce that $\widetilde{c}_{n+1} = 1$.

By recurrence relation (15.3.4) we have, for instance,

$$\begin{aligned}&H_3 = xH_2 - H_2' = x(x^2-1) - 2x = x^3 - 3x,\\ &H_4 = xH_3 - H_3' = x(x^3-3x) - (3x^2-3) = x^4 - 6x^2 + 3,\\ &H_5 = xH_4 - H_4' = x(x^4-6x^2+3) - (4x^3-12x) = x^5 - 10x^3 + 15x,\\ &\vdots\end{aligned}$$

definitely much easier than rule (15.3.2). Let's now compute the norm of H_n to determine the scaling factor of e_n. We have

$$\begin{aligned}\|H_n\|^2 = \langle H_n, H_n\rangle &= \langle xH_{n-1} - H_{n-1}', H_n\rangle = \langle H_{n-1}, xH_n\rangle - \langle H_{n-1}', H_n\rangle\\ &\overset{(15.3.3)}{=} \langle H_{n-1}, xH_n\rangle - \langle H_{n-1}, xH_n - H_n'\rangle\\ &\overset{(15.3.2)}{=} n\langle H_{n-1}, H_{n-1}\rangle = n\|H_{n-1}\|^2.\end{aligned}$$

Therefore,

$$\|H_n\|^2 = n\|H_{n-1}\|^2 = n(n-1)\|H_{n-2}\|^2 = \ldots = n!\|H_0\|^2 = n!$$

In conclusion

$$\left(\frac{1}{\sqrt{n!}}H_n(x)\right)_{n\in\mathbb{N}} \text{ is an orthonormal system for } L^2(\mathbb{R};\mathcal{N}(0,1)).$$

As a consequence,

$$\left(\frac{1}{\sqrt{2\pi n!}}H_n(x)e^{-\frac{x^2}{4}}\right)_n \text{ is an orthonormal system for } L^2(\mathbb{R}).$$

To verify that this is also an orthonormal basis, we will need Fourier transform.

15.4. Exercises

Exercise 15.4.1 (∗∗). *Discuss, in function of the real parameter $\alpha > 0$, the convergence for the series*

$$\sum_{n=0}^{\infty} \frac{1}{n^{\alpha}} \cos(nx)$$

is $L^1([0, 2\pi])$ and $L^2([0, 2\pi])$.

Exercise 15.4.2 (∗∗). *Let $H := L^2([0, +\infty[)$ equipped with usual real scalar product. Define*

$$e_n(x) := \sqrt{n} 1_{[n, n+\frac{1}{n}]}(x).$$

i) *Discuss point-wise convergence and H convergence of (e_n).*
ii) *Is (e_n) an orthonormal system? Is it a basis for H?*

Exercise 15.4.3 (∗). *Let H be a Hilbert space, (e_n) an orthonormal system such that the Parseval identity holds,*

$$\|f\|^2 = \sum_{n=0}^{\infty} |\langle f, e_n \rangle|^2, \ \forall f \in H.$$

Can we say that (e_n) is an orthonormal basis for H?

Exercise 15.4.4 (∗). *Let $H = L^2([0, 1])$ equipped with usual scalar product. Accepting that $e_0(x) \equiv 1$, $e_n(x) = \sqrt{2} \cos(n\pi x)$ is an orthonormal basis for H, apply the Parseval identity to*

$$f(x) := x 1_{[0,1/2]}(x) + (1 - x) 1_{]1/2,1]}(x),$$

to prove that

$$\frac{\pi^2}{8} = \sum_{k=0}^{\infty} \frac{1}{(2k + 1)^2}.$$

Exercise 15.4.5 (∗). *On $H = L^2([0, 1])$ equipped with the usual scalar product, apply the Gram-Schmidt algorithm to $v_n(x) := x^n$ $n \in \mathbb{N}$, to compute $e_0, \dots, e_4$.*

Exercise 15.4.6 (∗∗). *On $H := \{f \in L([-1, 1]) \ : \ \int_{-1}^{1} \frac{|f(x)|^2}{\sqrt{1-x^2}} \, dx < +\infty\}$ we define the scalar product*

$$\langle f, g \rangle := \int_{-1}^{1} \frac{f(x) g(x)}{\sqrt{1 - x^2}} \, dx.$$

We accept this is well defined, a scalar product on H (with weak vanishing) and $(H, \langle \cdot, \cdot \rangle)$ is a Hilbert space.

i) *Define $T_n(x) := \cos(n \arccos x)$. Find $(c_n) \subset \mathbb{R}$ such that $(c_n T_n)$ be an orthonotmal system.*
ii) *Compute T_0 and T_1, and prove that $T_{n+1} = 2x T_n - T_{n-1}$. Conclude that T_n are polynomials.*
iii) *Let $G(t, x) = \sum_{n=0}^{\infty} T_n(x) t^n$. Check that the series is convergent for $|t| < 1$ and use the recurrence relation to prove that*

$$G(t, x) = \frac{1 - tx}{1 - 2tx + t^2}, \ \forall (t, x) \in [-1, 1] \times [-1, 1].$$

Exercise 15.4.7 ($*{*}*$)**.** *Let H be a Hilbert space, (e_n) an orthonormal basis. Let $(\widetilde{e_n})$ be another orthonormal system. Check that if*

$$(\star)\ \sum_{n=0}^{\infty} \|e_n - \widetilde{e}_n\|^2 < +\infty$$

then also $(\widetilde{e}_n)$ is an orthonormal basis for H.

Exercise 15.4.8 (Legendre polynomials ($*{*}+$))**.** *Let $H := L^2([-1,1])$ and $v_n := x^n$, $n = 0, 1, 2, \ldots$.*

i) *Applying the Gram–Schmidt algorithm to $(v_n)_n$, compute e_0, e_1, e_2, e_3.*
ii) *Let $p_0(x) = 1$, $p_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2-1)^n$. Show that $\langle p_n, p_m \rangle = 0$ if $n \neq m$.*
iii) *Compute $\|p_n\|_2$.*

Exercise 15.4.9 ($*{*}+$)**.** *Let H be a separable infinite-dimensional Hilbert space. Show that there is no "Lebesgue measure" on H; that is, there is no measure μ on the Borel σ-algebra of H such that (i) μ is finite on bounded sets, and (ii) μ is translation-invariant. (hint. Let (e_n) be an orthonormal basis and choose $r > 0$ small enough so that the balls $B(e_n, r)$ are pairwise disjoint. Consider the set $E := \bigcup_{n=1}^{\infty} B(e_n, r]$. Use translation invariance and finiteness on bounded sets to derive a contradiction.)*

LECTURE 16

Classical Fourier Series

Classical Fourier Series arise from a very natural problem: *is it always possible to represent any $T-$periodic function as a (possibly infinite) linear combination of elementary $T-$periodic functions*? This Lecture discusses about this problem. As we will see, a natural way to look at this problem is the language of orthonormal bases we introduced in previous Lecture.

16.1. L^2 convergence

Let f be an arbitrary $T-$periodic function $f : \mathbb{R} \longrightarrow \mathbb{R}$ (that is, $f(x+T) = f(x)$ for every $x \in \mathbb{R}$). Classical examples of such a function are the *fundamental harmonics*

$$\sin\left(\frac{2\pi}{T}nx\right), \ \cos\left(\frac{2\pi}{T}nx\right), \ n \in \mathbb{N}.$$

For an arbitrary f we ask whether it is possible to determine coefficients a_n, b_n such that

$$f(x) = \sum_{n=0}^{\infty}\left(a_n\cos\left(\frac{2\pi}{T}nx\right) + b_n\sin\left(\frac{2\pi}{T}nx\right)\right) = a_0 + \sum_{n=1}^{\infty}\left(a_n\cos\left(\frac{2\pi}{T}nx\right) + b_n\sin\left(\frac{2\pi}{T}nx\right)\right).$$

The series on the right-hand side are called *trigonometric series*. This problem has a natural formulation in Hilbert space theory. Let us see how. The first step is to rearrange the form of a trigonometric series. Recalling the Euler identities

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}, \ \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}, \tag{16.1.1}$$

after straightforward calculations, we can write

$$\begin{aligned} a_0 + \textstyle\sum_{n=1}^{\infty}\left(a_n\cos\left(\frac{2\pi}{T}nx\right) + b_n\sin\left(\frac{2\pi}{T}nx\right)\right) &= a_0 + \textstyle\sum_{n=1}^{\infty}\left(\frac{a_n-ib_n}{2}e^{i\frac{2\pi}{T}nx} + \frac{a_n+ib_n}{2}e^{-i\frac{2\pi}{T}nx}\right) \\ &= \textstyle\sum_{n\in\mathbb{Z}} c_n e^{i\frac{2\pi}{T}nx}. \end{aligned}$$

Notice that we can pass from real form to complex form and vice-versa according to rules

$$c_n = \begin{cases} a_0, & n=0, \\ \frac{a_n-ib_n}{2}, & n \geqslant 1, \\ \frac{a_{-n}+ib_{-n}}{2}, & n \leqslant -1, \end{cases} \quad a_n = \begin{cases} c_0, & n=0, \\ a_n = 2\mathrm{Re}(c_n), & n \geqslant 1, \end{cases} \quad b_n = \begin{cases} 0, & n=0, \\ b_n = -2\mathrm{Im}(c_n), & n \geqslant 1. \end{cases}$$

We call **trigonometric series** a series of the form

$$\sum_{n\in\mathbb{Z}} c_n e^{i\frac{2\pi}{T}nx}, \ (c_n)_{n\in\mathbb{Z}} \subset \mathbb{C}.$$

Functions $e_n(x) := e^{i\frac{2\pi}{T}nx}$ are also called **characters**.

Proposition 16.1.1

Let $H := L^2_{\mathbb{C}}([0,T])$ be equipped with hermitian product

$$\langle f, g\rangle := \frac{1}{T}\int_0^T f(x)\overline{g(x)}\, dx.$$

Then $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal system in $L^2([0,T])$.

PROOF. It is just a simple calculation:

$$\langle e_n, e_m\rangle = \frac{1}{T}\int_0^T e^{i\frac{2\pi}{T}nx}\overline{e^{i\frac{2\pi}{T}mx}}\, dx = \frac{1}{T}\int_0^T e^{i\frac{2\pi}{T}(n-m)x}\, dx = \begin{cases} \frac{1}{T}\int_0^T 1\, dx = 1, & m = n, \\ \frac{1}{T}\left[\frac{e^{i\frac{2\pi}{T}(n-m)x}}{i\frac{2\pi}{T}(n-m)}\right]_{x=0}^{x=T} = 0, & m \neq m. \end{cases}$$

So, the identity

$$f = \sum_{n\in\mathbb{Z}} c_n e_n$$

becomes true in $L^2_{\mathbb{C}}([0,T])$ once we prove that $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal basis. This is true, however the proof is long and non trivial, and in fact consists showing directly that any f is sum of its abstract Fourier series. We will omit this proof here:

Theorem 16.1.2

The set of characters $e_n(x) := e^{i\frac{2\pi}{T}nx}$, $n \in \mathbb{Z}$, is an orthonormal basis for $L^2_{\mathbb{C}}([0,T])$. Thus, in particular,

$$f \overset{L^2}{=} \sum_{n\in\mathbb{Z}} \widehat{f}(n) e_n, \tag{16.1.2}$$

where

$$\widehat{f}(n) = \langle f, e_n\rangle = \frac{1}{T}\int_0^T f(x) e^{-i\frac{2\pi}{T}nx}\, dx,\ n \in \mathbb{Z}. \tag{16.1.3}$$

Identity (16.1.2) holds in L^2 sense. This doesn't necessarily mean that

$$f(x) = \sum_{n\in\mathbb{Z}} \widehat{f}(n) e^{i\frac{2\pi}{T}nx},\ \forall x \in [0,T].$$

This because, as well known, L^2 convergence does not imply point-wise convergence. More precisely: the infinite sum $\sum_{n\in\mathbb{Z}} \widehat{f}(n)e_n$ is a.e. equal to f. But when we take finite sums $s_N := \sum_{|n|\leqslant N} \widehat{f}(n)e_n$, we cannot say that $s_N(x) \longrightarrow f(x)$ a.e. x. Infact, there are examples for which it may happen that s_N is never point-wise convergent. These are quite "exotic" examples, however with some regularity requirement on f, the point wise convergence holds. Let's see a couple of examples.

Example 16.1.3: (square wave $*$)

Determine the FS of

$$f(x) := \begin{cases} 0, & x \in [0, \pi[, \\ 1, & x \in [\pi, 2\pi[, \end{cases}$$

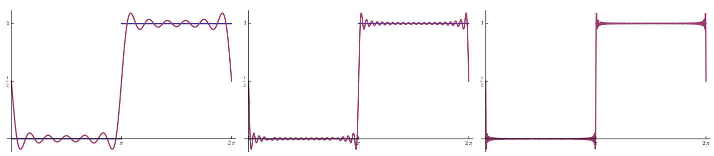
Proof. We have

$$\widehat{f}(n) = \frac{1}{2\pi}\int_\pi^{2\pi} e^{-inx}\,dx = \begin{cases} \frac{1}{2\pi}\int_\pi^{2\pi} dx = \frac{1}{2}, & n = 0, \\ \frac{1}{2\pi}\left[\frac{e^{-inx}}{-in}\right]_{x=\pi}^{x=2\pi} = \frac{1}{2\pi}\frac{1-e^{-in\pi}}{-in} = i\frac{1-(-1)^n}{2n\pi}, & n \neq 0. \end{cases}$$

Therefore the Fourier series for f is

$$\frac{1}{2} + \sum_{n\in\mathbb{Z}\setminus\{0\}} i\frac{1-(-1)^n}{2n\pi}e^{inx} = \frac{1}{2} + \sum_{k\in\mathbb{Z}} \frac{i}{(2k+1)\pi}e^{i(2k+1)x} = \frac{1}{2} - \sum_{k=1}^{\infty}\frac{2}{(2k+1)\pi}\sin((2k+1)x).$$

It is interesting to plot some graphs of partial sums of the series and to compare it with the graph of f. Here's the case of the partial sum of previous series of the first 5,20 and 100 terms respectively.



The picture seems to indicate at least a pointwise convergence for $x \in [0, 2\pi]$ except in the discontinuity points of f (that is on $x = 0, \pi, 2\pi$). For example, if $x = \pi$, the point-wise evaluation of the sum leads to

$$\frac{1}{2} - \sum_{k=1}^{\infty}\frac{2}{(2k+1)\pi}\sin((2k+1)\pi) = \frac{1}{2} - \sum_{k=1}^{\infty}\frac{2}{(2k+1)\pi}\cdot 0 = \frac{1}{2} \neq f(\pi) = 1.$$

Remark 16.1.4: ($*$)

The previous Example shows a remarkable application of the Parseval identity. According to this identity,

$$\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 = \frac{1}{4} + \sum_{n\ odd}\left|i\frac{1}{n\pi}\right|^2 = \frac{1}{4} + \frac{2}{\pi^2}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2},$$

and since $\|f\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi}|f(x)|^2\,dx = \frac{1}{2\pi}\int_\pi^{2\pi} 1\,dx = \frac{1}{2}$, we get

$$\frac{\pi^2}{8} = \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}.$$

Now, since

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\sum_{k=1}^{\infty}\frac{1}{(2k)^2}+\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}=\frac{1}{4}\sum_{k=1}^{\infty}\frac{1}{k^2}+\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2},$$

we obtain the remarkable harmonic sum

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{4}{3}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}=\frac{\pi^2}{6}. \qquad \square$$

Example 16.1.5: (triangular wave $*$)

Determine the FS of

$$f(x):=\begin{cases} x, & x\in[0,\pi[,\\ 2\pi-x, & x\in[\pi,2\pi[.\end{cases}$$

Proof. By definition,

$$\widehat{f}(0)=\frac{1}{2\pi}\int_0^{2\pi}f(x)\,dx=\frac{1}{2\pi}\frac{2\pi^2}{2}=\frac{\pi}{2},$$

while, as $n\neq 0$, integrating by parts,

$$\begin{aligned}\widehat{f}(n) &= \frac{1}{2\pi}\int_0^{2\pi}f(x)e^{-inx}\,dx=\frac{1}{2\pi}\left(\int_0^{\pi}xe^{-inx}\,dx+\int_{\pi}^{2\pi}(2\pi-x)e^{-inx}\,dx\right)\\ &= \frac{1}{2\pi}\left(\left[x\frac{e^{-inx}}{-in}\right]_{x=0}^{x=\pi}-\int_0^{\pi}\frac{e^{-inx}}{-in}\,dx+\left[(2\pi-x)\frac{e^{-inx}}{-in}\right]_{x=\pi}^{x=2\pi}+\int_{\pi}^{2\pi}\frac{e^{-inx}}{-in}\,dx\right)\\ &= \frac{1}{2\pi}\left(i\frac{(-1)^n\pi}{n}+\frac{1}{in}\left[\frac{e^{-inx}}{-in}\right]_{x=0}^{x=\pi}-i\frac{(-1)^n\pi}{n}-\frac{1}{in}\left[\frac{e^{-inx}}{-in}\right]_{x=\pi}^{x=2\pi}\right)\\ &= \frac{1}{2\pi}\left(\frac{1}{n^2}\left((-1)^n-1\right)-\frac{1}{n^2}\left(1-(-1)^n\right)\right)=\frac{(-1)^n-1}{\pi n^2}.\end{aligned}$$

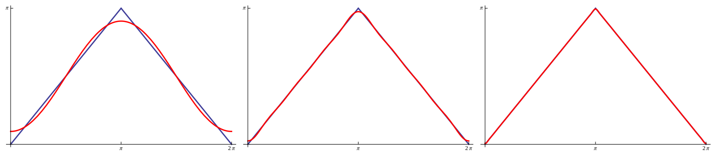
Therefore the Fourier series is

$$\frac{\pi}{2}+\sum_{n\neq 0}\frac{(-1)^n-1}{\pi n^2}e^{inx}=\frac{\pi}{2}-\frac{2}{\pi}\sum_{k\in\mathbb{Z}}\frac{1}{(2k+1)^2}e^{i(2k+1)x}.$$

For the real form we have $a_0=\frac{\pi}{2}$, $a_n=2\,\mathrm{Re}\,(\widehat{f}(n))=2\frac{(-1)^n-1}{\pi n^2}$ whereas $b_n=-2\,\mathrm{Im}\,(\widehat{f}(n))=0$ for any $n\geqslant 1$. Therefore the real form is

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

Also in this case let's see some plots of partial sums. The next picture shows sums of 1,4,15 terms respectively.



Here we clearly see that the approximation appears to converge to f pointwise, and even uniformly. Moreover, it seems much "better" than in the previous example. This indicates that regularity plays an important role in the rate of convergence. Finally, if—as it appears—the sum of the series is $f(x)$ for every x, then by taking $x = 0$ we obtain the remarkable identity

$$0 = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}, \quad \Longleftrightarrow \quad \frac{\pi^2}{8} = \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}.$$

16.2. Uniform convergence

The examples shown in the previous section suggest that pointwise convergence holds, provided that f is regular enough. This is a consequence of the following important fact: *the higher is the regularity of f, the faster Fourier coefficients decay to 0 as $n \longrightarrow \pm\infty$.* As we will see, this implies stronger convergence of the FS. As a general remark, we notice that

$$|\widehat{f}(n)| \leqslant \frac{1}{T}\int_0^T |f(x)|\,dx \overset{CS}{\leqslant} \frac{1}{T}\left(\int_0^T 1\,dx\right)^{1/2}\left(\int_0^T |f(x)|^2\,dx\right)^{1/2} = \left(\frac{1}{T}\int_0^T |f|^2\right)^{1/2} = \|f\|_2.$$

This bound shows that $|\widehat{f}(n)|$ is always bounded in n. With some extra regularity than simple L^2 measurability, we can get much more:

Proposition 16.2.1

Let $f \in \mathscr{C}^k([0,T])$, $f^{(j)}(0) = f^{(j)}(T)$, $j = 0,1,2,\dots,k-1$. then

$$\widehat{f^{(k)}}(n) = \left(i\frac{2\pi}{T}n\right)^k \widehat{f}(n), \ \forall n \in \mathbb{Z}. \tag{16.2.1}$$

In particular

$$|\widehat{f}(n)| \leqslant \frac{C\|f^{(k)}\|_\infty}{|n|^k}. \tag{16.2.2}$$

Proof. We limit to the case $f \in \mathscr{C}^1$ with $f(0) = f(T)$ (the general case follows similarly). Integrating by parts we have

$$\begin{aligned}\widehat{f'}(n) &= \frac{1}{T}\int_0^T f'(x)e^{-i\frac{2\pi}{T}nx}\,dx = \frac{1}{T}\left(\left[f(x)e^{-i\frac{2\pi}{T}nx}\right]_{x=0}^{x=T} - \int_0^T f(x)\left(-i\frac{2\pi}{T}n\right)e^{-i\frac{2\pi}{T}nx}\,dx\right)\\ &= \frac{1}{T}\left(f(T)e^{-i2\pi n} - f(0) + i\frac{2\pi}{T}n\int_0^T f(x)e^{-i\frac{2\pi n}{T}x}\,dx\right) = i\frac{2\pi}{T}n\widehat{f}(n).\end{aligned}$$

This proves the (16.2.1) for $k = 1$. Moreover, for $n \neq 0$,

$$|\widehat{f}(n)| = \frac{C}{|n|}|\widehat{f'}(n)|$$

and since

$$|\widehat{f'}(n)| \leqslant \frac{1}{T}\int_0^T \left|f'(x)e^{-i\frac{2\pi}{T}nx}\right| dx \leqslant \frac{1}{T}\int_0^T \|f'\|_\infty \, dx = \|f'\|_\infty,$$

we finally have

$$|\widehat{f}(n)| \leqslant \frac{C\|f'\|_\infty}{|n|}, \ \forall n \neq 0.$$

The fast decay of Fourier coefficients has implications on the way the FS converges.

Corollary 16.2.2

Let $f \in \mathscr{C}^2([0,T])$, $f^{(j)}(T) = f^{(j)}(0)$, $j = 0, 1$. Then, the FS of f converges uniformly to f.

PROOF. **First step: FS converges uniformly.** Let $s_N := \sum_{|n|\leqslant N} \widehat{f}(n)e_n$ be the $N-$th partial sum. Clearly $(s_N) \subset \mathscr{C}([0,T])$. We claim that (s_N) is convergent in $\|\cdot\|_\infty$ norm. To this aim, we apply the Weierstrass test: since $\|e_n\|_\infty = \max_{x\in[0,T]} |e^{i\frac{2\pi}{T}nx}| = 1$, we have

$$\sum_n \|\widehat{f}(n)e_n\|_\infty = \sum_n |\widehat{f}(n)|\|e_n\|_\infty = \sum_n |\widehat{f}(n)|.$$

By the assumptions on f and the bound (16.2.2), we have

$$|\widehat{f}(n)| \leqslant \frac{C\|f''\|_\infty}{|n|^2}, \ n \neq 0,$$

so

$$\sum_n \|\widehat{f}(n)e_n\|_\infty \leqslant |\widehat{f}(0)| + \sum_n \frac{K}{|n|^2} < +\infty.$$

Therefore, the Weierstrass test applies and we have the conclusion.
Second step: FS converges to f. By the first step we know there exists $g \in \mathscr{C}([0,T])$ such that

$$s_N \xrightarrow{\|\cdot\|_\infty} g.$$

Moreover, we already know that $s_N \xrightarrow{\|\cdot\|_2} f$. We claim $f \equiv g$ on $[0,T]$. Indeed, by the former we know that there exists a subsequence (s_{N_k}) such that $s_{N_k} \longrightarrow f$ a.e.. Since $s_{N_k} \xrightarrow{\|\cdot\|_\infty} g$, in particular $s_{N_k} \longrightarrow g$ pointwise on $[0,T]$. Thus $f = g$ a.e.. But both f and g are continuous functions (f by hypothesis, g being the uniform limit of continuous functions), thus $f \equiv g$ on $[0,T]$, and the proof is complete.

Remark 16.2.3

The two arguments in the previous proof are essentially independent of one another. In particular, the second argument shows that once we know that the FS of $f \in \mathscr{C}$ converges uniformly (in the $\|\cdot\|_\infty$ norm), then it necessarily converges to f itself.

The previous result is quite demanding on f. There are weaker versions of it with mild assumptions on f. For our purposes in this course, the previous corollary is sufficient.

16.3. Exercises

Exercise 16.3.1 ($*$). *Let $f(x) := x(1-x)$, $x \in [0,1]$. Compute the $L^2([0,1])$ FS of f and discuss its convergence in $L^\infty([0,1])$. Is the FS also point-wise convergent? If yes, what is the point-wise limit?*

Exercise 16.3.2 ($*$). *Let $f(x) = x$, $x \in [-\pi,\pi]$. Compute the FS of f. Use this series to prove (once more!) the formula $\frac{\pi^2}{6} = \sum_{n=1}^\infty \frac{1}{n^2}$.*

Exercise 16.3.3 ($*$). *Let $f(x) = x^2$, $x \in [-\pi,\pi]$. Compute the FS of f. Is the FS uniformly convergent? Use this FS to compute the value of the sum $\sum_{n=1}^\infty \frac{1}{n^4}$.*

Exercise 16.3.4 ($**$). *Use the FS of $f(x) = |\sin x|$ on $[-\pi,\pi]$ to determine a "cosine series" for $\sin x$ for $x \in [0,\pi]$, that is a representation of the form*

$$\sin x = \sum_{n=0}^\infty b_n \cos(nx),\ x \in [0,\pi].$$

Discuss carefully for which values of x such identity holds.

Exercise 16.3.5 ($**$). *Let $f(x) = \cosh x$, $x \in [-\pi,\pi]$. Compute the FS of f and discuss whether it converges to f or not on $[-\pi,\pi]$. Use this FS to compute the value of the sum $\sum_{n=0}^\infty \frac{1}{n^2+1}$.*

Exercise 16.3.6 ($**+$). *Let $b \in [0,1]$ be fixed and set $f_b(x) := \min\{x,b\}$, $x \in [0,1]$.*

i) *Compute the $L^2([0,1])$ FS of f_b. Is f_b the sum of its FS? Is the FS uniformly convergent to f_b?*
ii) *Deduce, by i), the formula*

$$\min\{x,y\} = \frac{4}{\pi^2}\sum_{n=0}^\infty \frac{1}{(2n+1)^2}\left(\sin\left(n+\frac{1}{2}\right)\pi x\right)\left(\sin\left(n+\frac{1}{2}\right)\pi y\right).$$

Exercise 16.3.7 ($**$). *Let $f \in L^2_{\mathbb{R}}([0,2T])$ be such that $f(x+T) = -f(x)$ a.e. What can you deduce on the Fourier coefficients for the real form of the FS?*

Exercise 16.3.8 ($**+$). *Let $f,g \in L^2([0,T])$. We define* **convolution product of** f **and** g *the function*

$$(f*g)(x) := \frac{1}{T}\int_0^T f(x-y)g(y)\,dy.$$

i) *Prove that the convolution is well defined and it belongs to $L^2([0,T])$, proving a Young inequality for $\|f*g\|_2$.*
ii) *Check that $\widehat{f*g}(n) = \widehat{f}(n)\widehat{g}(n)$, for every $n \in \mathbb{Z}$.*

Exercise 16.3.9 ($**$). *Let $f \in L^2([0,T])$ be such that*

$$\sum_n |n\widehat{f}(n)| < +\infty.$$

Prove that the FS of f is uniformly convergent.

LECTURE 17

L^1 Fourier Transform

In previous Lecture we have seen that, if $f \in L^2([0,T])$, then

$$f(x) \overset{L^2}{=} \sum_{n\in\mathbb{Z}} \widehat{f}(n) e^{i\frac{2\pi n}{T}x}. \tag{17.0.1}$$

This formula has remarkable applications in *signal processing*, especially with sound signals that exhibit a periodic behavior. Since the r.h.s. is just a $T-$periodic function, the period $[0,T]$ can be any interval of length T, as for example $[-\frac{T}{2},\frac{T}{2}]$. In particular then,

$$\widehat{f}(n) = \frac{1}{T}\int_{-T/2}^{T/2} f(y)e^{-i\frac{2\pi n}{T}x}\,dy,$$

and

$$f(x) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{i\frac{2\pi n}{T}x} = \sum_{n\in\mathbb{Z}}\left(\frac{1}{T}\int_{-T/2}^{T/2} f(y)e^{-i\frac{2\pi n}{T}y}\,dy\right)e^{i\frac{2\pi n}{T}x}$$

Now, suppose $T = 2\pi N$ with $N \longrightarrow +\infty$. Introducing points $\xi_n := \frac{n}{N}$, $n \in \mathbb{Z}$ as a subdivision of $\mathbb{R}$ in such a way that $d\xi_n = \xi_{n+1} - \xi_n = \frac{1}{N}$ we would have

$$f(x) = \sum_{n\in\mathbb{Z}}\left(\frac{1}{2\pi}\int_{-\pi N}^{\pi N} f(y)e^{-i\xi_n y}\,dy\right)e^{i\xi_n x}\,d\xi_n \longrightarrow \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} f(y)e^{-i\xi y}\,dy\right)e^{i\xi x}\,d\xi.$$

The integral

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(y)e^{-i\xi y}\,dy \tag{17.0.2}$$

is called **Fourier Transform of** f and the previous formula suggests that

$$f(x) = \frac{1}{2\pi}\widehat{\widehat{f}}(-x). \tag{17.0.3}$$

The (17.0.3) is the analogous of (17.0.1) and it suggests that $f \in L^2(]-\infty,+\infty[)$ can be "reconstructed" from its Fourier Transform. Of course, our argument was very informal. The scope of this and next chapters is to introduce the FT in a rigorous way, and see when formula (17.0.3), also named **inversion formula**, holds true.

17.1. Definition and first examples

A first problem with (17.0.2) is that the natural condition on f ensuring its well position is $f \in L^1(\mathbb{R})$ and not $f \in L^2(\mathbb{R})$. Indeed,

$$\int_{\mathbb{R}} |f(x)e^{-i\xi x}|\, dx = \int_{\mathbb{R}} |f(x)|\, dx < +\infty, \quad \Longleftrightarrow \quad f \in L^1(\mathbb{R}).$$

We introduce now the

Definition 17.1.1

Let $f \in L^1(\mathbb{R})$. The function

$$\widehat{f}(\xi) := \int_{\mathbb{R}} f(y)e^{-i\xi y}\, dy,\ \xi \in \mathbb{R} \tag{17.1.1}$$

is called **Fourier Transform (FT) of** f

Warning 17.1.2

In the literature, there are slightly different definitions of FT. The *mathematicians FT* is defined as

$$\widehat{f}(\xi) := \int_{\mathbb{R}} f(y)e^{-i2\pi\xi y}\, dy.$$

Basically, respect to our definition, it changes ξ with $2\pi\xi$. The advantage with this definition is that inversion formula (17.0.3) becomes slightly easier, $f(x) = \widehat{\widehat{f}}(-x)$. Furthermore, with mathematicians definition, the L^2 Fourier-Plancherel Transform becomes a true isometry (see Lecture on L^2 FT). On the other hand, the presence of fact 2π complicates formulas and make them a bit more difficult to memorize.
In Probability, as we will see, FT appears as *characteristic function*, which is an integral of type

$$\phi(\xi) := \int_{\mathbb{R}} f(y)e^{i\xi y}\, dy.$$

In this case, it is evident that we changed $-\xi$ into ξ.

Let's see some important examples. We already computed

Example 17.1.3: Gaussian distribution $(**)$

$$\widehat{e^{-\frac{\sharp^2}{2\sigma^2}}}(\xi) = \sqrt{2\pi\sigma^2}e^{-\frac{1}{2}\sigma^2\xi^2}, \quad (\sigma > 0). \tag{17.1.2}$$

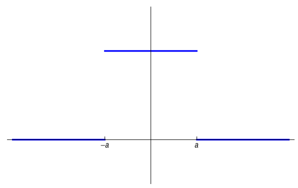
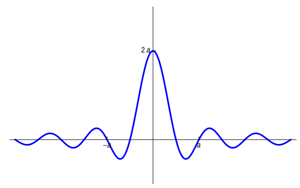
Proof. See (7.2.1).

Example 17.1.4: rectangle $(*)$

et $\mathrm{rect}_a := 1_{[-a,a]}$. Then

$$\widehat{\mathrm{rect}_a}(\xi) = 2a\frac{\sin(a\xi)}{a\xi} =: 2a\ \mathrm{sinc}(a\xi),\ \forall \xi \in \mathbb{R}. \tag{17.1.3}$$

(where $\mathrm{sinc}\ t := \frac{\sin t}{t}$, with the agreement that $\mathrm{sinc}\ 0 = 1$).

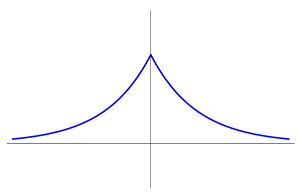
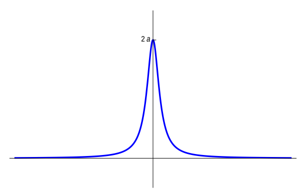
 

Proof. Clearly $\mathrm{rect}_a = 1_{[-a,a]} \in L^1(\mathbb{R})$. We have

$$\widehat{\mathrm{rect}_a}(\xi) = \int_{-a}^{a} e^{-i\xi y}\, dy = \begin{cases} \xi = 0, & = \int_{-a}^{a} dy = 2a, \\ \xi \neq 0, & = \left[\frac{e^{-i\xi y}}{-i\xi}\right]_{y=-a}^{y=a} = -\frac{1}{i\xi}\left(e^{-ia\xi} - e^{ia\xi}\right) = 2a\frac{\sin(a\xi)}{a\xi} \end{cases}$$

Example 17.1.5: exponential

$$\widehat{e^{-a|\sharp|}}(\xi) = \frac{2a}{a^2+\xi^2},\ \xi \in \mathbb{R},\ \ (a > 0). \tag{17.1.4}$$

Proof. Clearly $e^{-a|\sharp|} \in L^1(\mathbb{R})$ if $a > 0$. By definition

$$\begin{aligned}
\widehat{e^{-a|\sharp|}}(\xi) &= \int_{-\infty}^{+\infty} e^{-a|y|}e^{-i\xi y}\, dy = \int_{-\infty}^{0} e^{ay}e^{-i\xi y}\, dy + \int_{0}^{+\infty} e^{-ay}e^{-i\xi y}\, dy \\
&= \int_{-\infty}^{0} e^{(a-i\xi)y}\, dy + \int_{0}^{+\infty} e^{-(a+i\xi)y}\, dy = \left[\frac{e^{(a-i\xi)y}}{a-i\xi}\right]_{y=-\infty}^{y=0} + \left[-\frac{e^{-(a+i\xi)y}}{a+i\xi}\right]_{y=0}^{y=+\infty} \\
&= \frac{1}{a-i\xi} + \frac{1}{a+i\xi} = \frac{2a}{a^2+\xi^2}.
\end{aligned}$$

The Definition of FT extends naturally to multidimensional functions:

Definition 17.1.6

Let $f \in L^1(\mathbb{R}^d)$. The function

$$\widehat{f}(\xi) := \int_{\mathbb{R}} f(y)e^{-i\xi\cdot y}\,dy,\ \xi \in \mathbb{R}^d \tag{17.1.5}$$

is called **Fourier Transform (FT) of f**

Example 17.1.7: multivariate gaussian $(**)$

Let C be a strictly positive definite symmetric matrix (that is: $Cx\cdot x > 0$ for very $x \in \mathbb{R}^d \setminus \{0\}$, $C^t = C$). Then

$$\widehat{e^{-\frac{1}{2}C^{-1}\sharp\cdot\sharp}}(\xi) = \sqrt{(2\pi)^d \det C}\,e^{-\frac{1}{2}C\xi\cdot\xi}. \tag{17.1.6}$$

Proof. We notice first that C^{-1} is diagonalizable. Indeed: C is positive definite, and by this it follows that C is invertible. Since C is symmetric, C^{-1} it is. Thus C^{-1} is symmetric, hence it is diagonalizable, that is $C^{-1} = T^{-1}\Lambda^{-1}T$ for some T orthogonal matrix, that is $T^{-1} = T^t$ (transposed matrix), and $\Lambda^{-1} := \operatorname{diag}(\frac{1}{\sigma_1^2},\ldots,\frac{1}{\sigma_d^2})$ a diagonal matrix. Therefore

$$C^{-1}y\cdot y = T^t\Lambda^{-1}Ty\cdot y = \Lambda^{-1}Ty\cdot Ty.$$

Now, notice that

$$\widehat{e^{-\frac{1}{2}C^{-1}\sharp\cdot\sharp}}(\xi) = \int_{\mathbb{R}^d} e^{-\frac{1}{2}\Lambda^{-1}Ty\cdot Ty}e^{-i\xi\cdot T^tTy}\,dy \overset{x=Ty}{=} \int_{\mathbb{R}^d} e^{-\frac{1}{2}\Lambda^{-1}x\cdot x}e^{-iT\xi\cdot x}|\det T^t|\,dx,$$

and because $T^t = T^{-1}$ easily $|\det T^t| = 1$. Therefore

$$\widehat{e^{-\frac{1}{2}C^{-1}\sharp\cdot\sharp}}(\xi) \overset{F-T}{=} \prod_{j=1}^d \int_{\mathbb{R}} e^{-\frac{1}{2\sigma_j^2}x^2}e^{-i(T\xi)_j x_j}\,dx_j = \prod_j \widehat{e^{-\frac{\sharp^2}{2\sigma_j^2}}}((T\xi)_j) \overset{(17.1.2)}{=} \prod_j \sqrt{2\pi\sigma_j^2}\,e^{-\frac{1}{2}\sigma_j^2(T\xi)_j^2}.$$

To finish notice that

$$\prod_j \sigma_j^2 = \det\Lambda = \det(TC^{-1}T^{-1})^{-1}) = \det C,$$

and

$$\sum_j \sigma_j^2(T\xi)_j^2 = (\Lambda^{-1}T\xi)\cdot T\xi = T^t\Lambda T\xi\cdot\xi = C\xi\cdot\xi,$$

and by these identities the conclusion follows easily.

Let's finish this Section with few useful "algebraic" properties of the Fourier transform:

Proposition 17.1.8

Let $f \in L^1$. Then

i) $\widehat{f(\cdot - x_0)} = e^{-i\xi\cdot x_0}\widehat{f}$.

ii) $\widehat{e^{-i\sharp\cdot v}f}(\xi) = \widehat{f}(\xi + v)$.

iii) $\widehat{f(\lambda\sharp)}(\xi) = \frac{1}{|\lambda|^d}\widehat{f}\left(\frac{\xi}{\lambda}\right)$.

PROOF. We limit to prove the first one, the remaining being similar (exercise). We have

$$\begin{aligned}\widehat{f(\cdot - x_0)}(\xi) &= \int_{\mathbb{R}^d} f(x-x_0)e^{-i\xi\cdot x}\,dx = e^{-i\xi\cdot x_0}\int_{\mathbb{R}^d} f(x-x_0)e^{-i\xi\cdot(x-x_0)}\,dx \\ &= e^{-i\xi\cdot x_0}\widehat{f}(\xi).\end{aligned}$$

For future developments we will deal with 1-dimensional FT, the majority of the results extend to the general case with straightforward adjustments.

17.2. Exercises

Exercise 17.2.1 $(*)$**.** *Compute the Fourier transforms of the following functions:*

$1.\ x\,\mathrm{rect}_a(x).\ \ 2.\ (a-|x|)\mathrm{rect}_a(x).\ \ 3.\ (\cos x)\mathrm{rect}_{\pi/2}(x).\ \ 4.\ e^{-|x|}\mathrm{sgn}(x).\ \ 5.\ e^{-x}1_{[0,+\infty[}(x).$

Exercise 17.2.2 $(*)$**.** *Compute* $\widehat{1_{[-a,a]^d}}$.

Exercise 17.2.3 $(**+)$**.** *Compute the FT of* $f := 1_{x^2+y^2+z^2\leqslant r^2}$.

Exercise 17.2.4 $(**)$**.** *Show that if f is real valued and even (that is $f(-x)=f(x)$ a.e.), then $\widehat{f}$ is real valued.*

Exercise 17.2.5 $(**)$**.** *Let R be an orthogonal matrix, $RR^t = R^tR = \mathbb{I}$. Express the FT of $f(Rx)$ in terms of $\widehat{f}$.*

Exercise 17.2.6 $(**)$**.** *Prove the properties of the Proposition 17.1.8.*

Exercise 17.2.7 $(*)$**.** *Let $f(x_1,\dots,x_n) := \prod_{j=1}^n f_j(x_j) \in L^1(\mathbb{R}^n)$. Prove that $\widehat{f}(\xi_1,\dots,\xi_n) = \prod_{j=1}^n \widehat{f_j}(\xi_j)$.*

Exercise 17.2.8 $(**+)$**.** *Let $f\in L^1(\mathbb{R})$ be such that $f(x)>0$ a.e.. Prove that $|\widehat{f}(\xi)| < f(0)$, $\forall \xi\neq 0$.*

Exercise 17.2.9 $(**+)$**.** *Let $f\in L^1(\mathbb{R})$ be such that $f(x)\equiv 0$ for $|x|\geqslant R$. Prove that $\widehat{f}$ is a power series.*

LECTURE 18

Properties of L^1 FT

In this lecture we present some of the most important properties of the FT. Among others, a special role is played by differentiation: the FT converts a differential polynomial into multiplication by an algebraic polynomial. Concretely, this means converting certain differential equations into algebraic equations. Of course, this has many consequences in applications, some of which will be presented in subsequent lectures.

18.1. Riemann–Lebesgue Lemma

What can be said about the FT $\widehat{f}$ of $f \in L^1$? For example: *is $\widehat{f} \in L^1(\mathbb{R})$?* This question is particularly important in order the inversion formula

$$f(x) = \frac{1}{2\pi}\widehat{\widehat{f}}(-x),$$

makes sense. Unfortunately, the answer is negative.

Example 18.1.1

Let $f = \mathrm{rect}_1 \in L^1(\mathbb{R})$. Then, $\widehat{f}(\xi) = \frac{\sin\xi}{\xi} \notin L^1(\mathbb{R})$. □

In fact, $\widehat{f}$ is qualitatively very different from its original f:

Lemma 18.1.2: Riemann–Lebesgue

Let $f \in L^1(\mathbb{R})$. Then

i) $\widehat{f} \in \mathscr{C}(\mathbb{R}) \cap L^\infty(\mathbb{R})$ (continuous and bounded) and

$$\|\widehat{f}\|_\infty \leqslant \|f\|_1, \ \forall \xi \in \mathbb{R}. \tag{18.1.1}$$

ii) Furthermore,

$$\lim_{\xi \to \pm\infty} \widehat{f}(\xi) = 0. \tag{18.1.2}$$

Proof. The continuity follows by continuity of integrals depending on parameters. Indeed, setting $f(x,\xi) := f(x)e^{-i\xi x}$ we have $f(\sharp,\xi) \in L^1(\mathbb{R})$ for every $\xi \in \mathbb{R}$; $f(x,\sharp) \in \mathscr{C}(\mathbb{R})$, a.e. $x \in \mathbb{R}$; the following bound holds

$$|f(y)e^{-i\xi\cdot y}| \leqslant |f(y)| \in L^1(\mathbb{R}).$$

Thus, hypotheses of Thm 7.1 are fulfilled, $\widehat{f} \in \mathscr{C}(\mathbb{R})$. The bouns (18.1.1) follows from

$$|\widehat{f}(\xi)| \leqslant \int_{\mathbb{R}} \left|f(y)e^{-i\xi y}\right| \, dy = \int_{\mathbb{R}} |f(y)| \, dy = \|f\|_1.$$

The proof of (18.1.2) is more complex. We omit here this proof, we will obtain it under more restrictive assumptions in next results.

18.2. Fourier Transform of Derivative

The Riemann–Lebesgue (RL) lemma does not ensure that $\widehat{f} \in L^1$ (and indeed, this is false). Knowing that $\widehat{f} \in \mathscr{C}(\mathbb{R})$ ensures there are no integrability issues on any finite interval $[a, b]$. Thus, the issue concerns the behavior of $\widehat{f}$ at $\pm\infty$. According to the RL lemma, $\widehat{f}(\xi) \longrightarrow 0$ as $\xi \longrightarrow \pm\infty$, but this is insufficient to conclude. However, as with FS, by imposing some regularity on f we obtain more precise behavior at $\pm\infty$ and, ultimately, integrability. Before we attack the main result, we need to introduce the concept of weak derivative:

Definition 18.2.1: weak derivative

Let $f \in L^1(\mathbb{R})$. We say that $\exists \partial_x f \in L^1(\mathbb{R})$ if there exists a function $g \in L^1(\mathbb{R})$ such that

$$f(b) - f(a) = \int_a^b g(x)dx, \ a.e. \ a, b \in \mathbb{R}. \tag{18.2.1}$$

We set $\partial_x f := g$.

It can be proved that the definition is well posed modulo a.e. equivalence.

Example 18.2.2

If $f \in L^1(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$, then $\partial_x f$ exists pointwise in the ordinary sense and, according to the fundamental theorem of Integral Calculus, the identity (18.2.1) holds with $g = \partial_x f$. So, if $\partial_x f \in L^1(\mathbb{R})$, the ordinary derivative is also the weak derivative.

Example 18.2.3

If $f(x) = e^{-|x|}$, then $f \in L^1(\mathbb{R})$ has weak derivative $\partial_x f(x) = -\mathrm{sgn}(x)e^{-|x|}$ (defined for $x \neq 0$).

PROOF. Indeed,

- if $0 < a < b$
$$\int_a^b -\mathrm{sgn}(x)e^{-|x|} \, dx = \int_a^b -e^{-x} \, dx = [e^{-x}]_{x=a}^{x=b} = e^{-b} - e^{-a} = f(b) - f(a).$$
- if $a < b < 0$ the argument is similar.

- if $a < 0 < b$, we have

$$\begin{aligned}\int_a^b -\mathrm{sgn}(x)e^{-|x|}\,dx &= \int_a^0 e^x\,dx + \int_0^b -e^{-x}\,dx = [e^x]_{x=a}^{x=0} + [e^{-x}]_{x=0}^{x=b}\\ &= (1-e^a) + (e^{-b}-1) = e^{-b} - e^a = f(b) - f(a).\end{aligned}$$

Proposition 18.2.4

Let $f, \partial_x f \in L^1(\mathbb{R})$. Then

$$\widehat{\partial_x f}(\xi) = i\xi \widehat{f}(\xi). \tag{18.2.2}$$

In particular:

$$|\widehat{f}(\xi)| \leqslant \frac{\|\partial_x f\|_1}{|\xi|}. \tag{18.2.3}$$

More in general, if $\partial_x^k f \in L^1(\mathbb{R})$, $k = 0, 1, \dots, n$, then

$$\widehat{\partial_x^n f}(\xi) = (i\xi)^n \widehat{f}(\xi). \tag{18.2.4}$$

In particular:

$$|\widehat{f}(\xi)| \leqslant \frac{\|\partial_x^n f\|_1}{|\xi|^n}. \tag{18.2.5}$$

Proof. Integrating by parts,

$$\widehat{\partial_x f}(\xi) = \int_{\mathbb{R}} \partial_x f(x) e^{-i\xi x}\,dx = \left[f(x)e^{-i\xi x}\right]_{x=-\infty}^{x=+\infty} - \int_{\mathbb{R}} f(x)\partial_x\left(e^{-i\xi x}\right)\,dx.$$

The key remark is $f(x) \longrightarrow 0$ at $x \longrightarrow \pm\infty$ [a]. Indeed,

$$f(x) - f(0) = \int_0^x \partial_y f(y)\,dy \longrightarrow \int_0^{\pm\infty} \partial_y f(y)\,dy \in \mathbb{R}, \text{ because } \partial x f \in L^1(\mathbb{R}).$$

Therefore the $\lim_{x\to\pm\infty} f(x)$ exists finite. Being f is integrable, such a limit cannot be anything else than 0. Now, being $e^{-i\xi y}$ bounded function, we obtain that $\left[f(x)e^{-i\xi x}\right]_{x=-\infty}^{x=+\infty} = 0$. Hence

$$\widehat{\partial_x f}(\xi) = -\int_{\mathbb{R}} f(x)\partial_x\left(e^{-i\xi x}\right)\,dx = i\xi \int_{\mathbb{R}} f(x)e^{-i\xi x}\,dx = i\xi\widehat{f}(\xi).$$

This proves (18.2.2). For the bound (18.2.3) we have

$$|\widehat{f}(\xi)| = \frac{|\widehat{\partial_x f}(\xi)|}{|\xi|} \overset{(18.1.1)}{\leqslant} \frac{\|\partial_x f\|_1}{|\xi|}.$$

The general case of the formula (18.2.4) can be obtained iterating the formula (18.2.2).

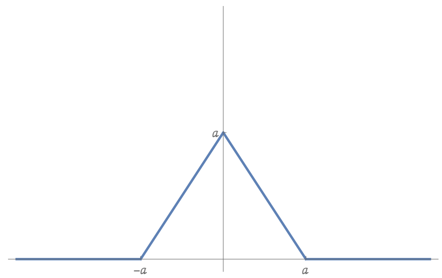
[a] Warning! This is not true in general for an $f \in L^1(\mathbb{R})$: for example $f(x) = x1_{\mathbb{Z}}(x) \in L^1(\mathbb{R})$ (clearly $f = 0$ a.e.) but f is not even bounded as $|x| \longrightarrow +\infty$.

Example 18.2.5

Let

$$f_a(x) := \begin{cases} 0, & x \leqslant -2a,\ x \geqslant 2a, \\ x + 2a, & -2a \leqslant x \leqslant 0, \\ -x + 2a, & 0 \leqslant x \leqslant 2a. \end{cases}$$

Compute $\partial_x f_a$ and deduce $\widehat{f}_a$.



PROOF. Clearly $f_a \in L^1(\mathbb{R})$. Moreover,

$$\partial_x f_a(x) := \begin{cases} 0, & x < -2a,\ x > 2a, \\ 1, & -2a < x < 0, \\ -1, & 0 < x < 2a. \end{cases} \qquad = 1_{[-2a,0]}(x) - 1_{[0,2a]}(x) \in L^1(\mathbb{R}).$$

By (18.2.2)

$$i\xi \widehat{f}_a(\xi) = \widehat{\partial_x f_a}(\xi) = \widehat{\chi_{[-a,0]}}(\xi) - \widehat{\chi_{[0,a]}}(\xi),$$

and since

$$1_{[-2a,0]}(\sharp) = 1_{[-a,a]}(\sharp + a), \implies \widehat{1_{[-2a,0]}}(\xi) = \widehat{\mathrm{rect}_a(\sharp + a)}(\xi) = e^{ia\xi}\frac{\sin(a\xi)}{\xi}.$$

Similarly

$$\widehat{1_{[0,2a]}}(\xi) = \widehat{\mathrm{rect}_a(\sharp - a)}(\xi) = e^{-ia\xi}\frac{\sin(a\xi)}{\xi}.$$

Thus

$$\widehat{\partial_x f_a}(\xi) = \left(e^{ia\xi} - e^{-ia\xi}\right)\frac{\sin(a\xi)}{\xi} = 2i\frac{(\sin(a\xi))^2}{\xi},$$

hence, finally

$$\widehat{f}_a(\xi) = \left(\frac{\sin(a\xi)}{\xi}\right)^2.$$

For $n \geqslant 2$, bound (18.2.5) ensures integrability. This yields a simple test for $\widehat{f} \in L^1$:

Corollary 18.2.6

If $f, \partial_x f, \partial_x^2 f \in L^1(\mathbb{R})$, then $\widehat{f} \in L^1(\mathbb{R})$.

PROOF. By (18.2.5)

$$|\widehat{f}(\xi)| \leqslant \frac{\|\partial_x^2 f\|_1}{\xi^2} =: \frac{K}{\xi^2}.$$

Now, by RL lemma $\widehat{f} \in \mathscr{C}(\mathbb{R})$, thus $\widehat{f}$ is integrable on any closed and bounded interval. To establish integrability on $\mathbb{R}$ we look at behaviour of $\widehat{f}$ at $\pm\infty$. Thanks to the previous bound, $|\widehat{f}|$ decays faster than $\frac{K}{\xi^2}$, which is integrable at $\pm\infty$. The conclusion now follows.

18.3. Derivative of Fourier Transform

The (18.2.2) shows a remarkable feature of the FT: *the FT converts "derivations" into "multiplications by $i\xi$"*. Reversing the order of the two operations—namely, FT and differentiation—the same phenomenon occurs: *the FT converts "multiplications by $-i\xi$" into "derivations"*. Here is the precise statement:

Proposition 18.3.1

Let $f \in L^1(\mathbb{R})$ such that $xf(x) \in L^1(\mathbb{R})$. Then

$$\exists \partial_\xi \widehat{f}(\xi) = [\widehat{(-i\sharp) f(\sharp)}](\xi), \ \forall \xi \in \mathbb{R}. \tag{18.3.1}$$

PROOF. It is an application of the differentiation under the integral sign. By definition

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x}\, dx.$$

Differentiating,

$$\partial_\xi(\widehat{f})(\xi) = \int_{\mathbb{R}} -ixf(x)e^{-i\xi x}\, dx = [\widehat{-i\sharp f(\sharp)}](\xi).$$

To justify this, we need to dominate $-ixf(x)e^{-i\xi x}$ uniformly in ξ with an L^1 function in x. But this follows immediately by our assumptions being

$$\left|-ixf(x)e^{-i\xi x}\right| \leqslant |xf(x)| \in L^1, \ \forall \xi \in \mathbb{R}.$$

Combining formulas (18.2.4) and (18.3.1) we obtain the relation

$$\widehat{(-i\sharp)^j \partial_x^k f} \equiv \partial_\xi^j \left[(i\xi)^k \widehat{f} \right]. \tag{18.3.2}$$

A remarkable consequence of this relation is the following

Proposition 18.3.2

FT applies $\mathscr{S}(\mathbb{R})$ into itself, that is: if $f \in \mathscr{S}(\mathbb{R})$ then $\widehat{f} \in \mathscr{S}(\mathbb{R})$.

PROOF. Let $f \in \mathscr{S}(\mathbb{R})$. To show that $\widehat{f} \in \mathscr{S}(\mathbb{R})$ have to check two facts:

i) $\widehat{f} \in \mathscr{C}^\infty$;
ii) $\widehat{f}^{(k)}$ is rapidly decaying at ∞.

Let's see how both are direct consequences of the *multiplication-derivation duality*. Indeed, since $f \in \mathscr{S}$ we have $x^k f \in L^1$. By Proposition 18.3 it follows that $\exists \partial_\xi^k \widehat{f} = \widehat{-i\sharp^k f} \in \mathscr{C}(\mathbb{R})$ (Riemann–Lebesgue). Conclusion: $\widehat{f} \in \mathscr{C}^k$ for every k, that is $\widehat{f} \in \mathscr{C}^\infty$.
Moreover, by formula (18.3.2)

$$|(i\xi)^h \partial^k \widehat{f}(\xi)| = |[\partial^h \widehat{(-i\sharp)^k f(\sharp)}](\xi)| \overset{(18.1.1)}{\leqslant} \|\partial^h (-i\sharp)^k f\|_1 =: C_{h,k} < +\infty.$$

Therefore

$$\sup_{\xi \in \mathbb{R}} (1 + |\xi|)^h |\partial_k \widehat{f}(\xi)| < +\infty, \ \forall h, k,$$

and this precisely means that $\widehat{f} \in \mathscr{S}(\mathbb{R})$.

18.4. Convolution

Another remarkable property of FT is that it converts convolution products into algebraic products:

Theorem 18.4.1

Let $f, g \in L^1(\mathbb{R})$. Then

$$\widehat{f * g} = \widehat{f}\widehat{g}. \tag{18.4.1}$$

PROOF. By Young inequality, $f * g \in L^1$ so FT makes sense. Computing its FT we get:

$$\begin{aligned} \widehat{f * g}(\xi) &= \int_{\mathbb{R}} (f * g)(y) e^{-i\xi y} \, dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y - x) g(x) \, dx \right) e^{-i\xi y} \, dy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y - x) g(x) e^{-i\xi y} \, dx \right) dy \\ &\overset{Fubini}{=} \int_{\mathbb{R}} g(x) e^{-i\xi x} \left(\int_{\mathbb{R}} f(y - x) e^{-i\xi(y-x)} \, dy \right) dx \\ &= \widehat{f}(\xi)\widehat{g}(\xi). \quad \square \end{aligned}$$

Remark 18.4.2

Here's another argument that shows that *convolution product has not units*, that is $\nexists \delta \in L^1$ such that $f * \delta = f$ for every $f \in L^1$. If a unit δ exists, then

$$\widehat{f * \delta} = \widehat{f}, \quad \Longleftrightarrow \quad \widehat{f}\widehat{\delta} = \widehat{f}, \ \forall f \in L^1.$$

So, for instance, taking f the Gaussian, $\widehat{f}$ is still a Gaussian, thus in particular, $\widehat{f} \neq 0$ always, we would obtain

$$\widehat{\delta}(\xi) \equiv 1.$$

But, according to RL lemma, $\widehat{\delta}(\xi) \longrightarrow 0$ for $\xi \longrightarrow \pm\infty$ and this is impossible.

18.5. Exercises

Exercise 18.5.1 $(*)$**.** *By using the multiplication-differentiation duality, compute* $\widehat{\sharp^2 1_{[-1,1]}}$.

Exercise 18.5.2 $(*)$**.** *Compute the FT of* $xe^{-x^2} * e^{-x^2}$.

Exercise 18.5.3 $(**)$**.** *Let* $a > 0$ *and define* $f_a(x) := e^{-ax}1_{[0,+\infty[}(x)$. *Compute the FT of* $f_a * f_b$ *(with* $a, b > 0$*).*

Exercise 18.5.4 $(**)$**.** *Let* $a, b > 0$, $a \neq b$ *and define*

$$f_{a,b}(x) := \frac{e^{-a|x|} - e^{-b|x|}}{x}.$$

Is $f_{a,b} \in L^1(\mathbb{R})$*? If yes, compute* $\widehat{f_{a,b}}$.

Exercise 18.5.5 $(**)$**.** *The scope of this exercise is to compute the FT of the standard gaussian* $f(x) = e^{-\frac{x^2}{2}}$ *in a "smart" way. Start noticing that* $f'(x) = -xf(x)$*, hence apply the FT both sides...*

Exercise 18.5.6 $(**)$**.** *Let* $f \in L^1(\mathbb{R})$ *be such that* $f', xf \in L^1(\mathbb{R})$. *Show that* $\widehat{f} \in L^1(\mathbb{R})$.

Exercise 18.5.7 $(**)$**.** *Let* $f \in \mathscr{S}(\mathbb{R}^n)$. *Compute* $\widehat{\Delta f}$ *where* $\Delta f = \sum_{j=1}^{n} \partial_j^2 f$ *((here* $\partial_j^2 \equiv \partial_{x_j}^2$*).*

Exercise 18.5.8 $(**+)$**.** *Let* $f, xf \in L^1(\mathbb{R})$ *be such that* $\int_{\mathbb{R}} f(x)\, dx = 0$. *Let*

$$g(x) := \int_{-\infty}^{x} f(y)\, dy.$$

i) *Check that g is well defined and belongs to* $L^1(\mathbb{R})$.
ii) *Determine the relation between the FT of g and that one of f.*

Exercise 18.5.9 $(***)$**.** *Let* $f \in L^1$ *and define*

$$f_\varepsilon(x) := \left(\widehat{e^{-\varepsilon^2\sharp^2}\widehat{f}}\right)(-x).$$

i) *Check that* f_ε *is well defined.*
ii) *Show that* $f_\varepsilon \xrightarrow{L^1} f$.

LECTURE 19

Inversion Formula

This Lecture is devoted to prove inversion formula

$$f(x) = \frac{1}{2\pi}\hat{\hat{f}}(-x).$$

19.1. Main result

Theorem 19.1.1

Let $f \in L^1(\mathbb{R})$ be such that $\hat{f} \in L^1(\mathbb{R})$. Then, inversion formula holds in the sense that

$$f(x) = \frac{1}{2\pi}\hat{\hat{f}}(-x), \ a.e.\ x \in \mathbb{R}. \tag{19.1.1}$$

Proof. A naïve attempt to prove inversion formula would start noticing that

$$\hat{\hat{f}}(-x) = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(y)e^{-i\xi y}\,dy\right)e^{i\xi x}\,d\xi = \int_{\mathbb{R}\times\mathbb{R}} f(y)e^{i\xi(x-y)}\,dyd\xi.$$

Here, there is a first trouble. Fubini's thm applies provided $f(y)e^{i\xi(x-y)}$ is $L^1(\mathbb{R}^2)$ in (y,ξ). However, unless $f = 0$ a.e., we have

$$\int_{\mathbb{R}^2}\left|f(y)e^{i\xi(x-y)}\right|\,dyd\xi = \int_{\mathbb{R}^2}|f(y)|\,dyd\xi = +\infty,$$

So, to make this false departure a true one, we introduce a weight $e^{-\frac{1}{2}\varepsilon^2\xi^2}$ that will be eliminated letting $\varepsilon \downarrow 0$. That is, let's consider the integral

$$I_\varepsilon(x) := \int_{\mathbb{R}^2} f(y)e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{i\xi(x-y)}\,dyd\xi$$

Notice that now $f(y)e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{i\xi(x-y)}$ is $L^1(\mathbb{R}^2)$ in (y,ξ) being

$$\int_{\mathbb{R}^2}\left|f(y)e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{i\xi(x-y)}\right|\,dyd\xi = \int_{\mathbb{R}^2}|f(y)|e^{-\frac{1}{2}\varepsilon^2\xi^2}\,dyd\xi \overset{RF}{=} \|f\|_1\sqrt{\frac{2\pi}{\varepsilon^2}}.$$

This allows to use RF on I_ε. We will do in two ways. On one side,

$$I_\varepsilon(x) = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(y)e^{-i\xi y}\,dy\right)\,e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{-i\xi x}\,d\xi = \int_{\mathbb{R}}\hat{f}(\xi)e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{i\xi x}\,d\xi.$$

To compute the limit when $\varepsilon \downarrow 0$, we apply dominated convergence. Notice that

- $\hat{f}(\xi)e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{i\xi x} \longrightarrow \hat{f}(\xi)e^{i\xi x}$, a.e. $\xi \in \mathbb{R}$;
- $|\hat{f}(\xi)e^{-\frac{1}{2}\varepsilon^2\xi^2}e^{i\xi x}| = |\hat{f}(\xi)|e^{-\frac{1}{2}\varepsilon^2\xi^2} \leqslant |\hat{f}(\xi)| \in L^1(\mathbb{R})$, $\forall\varepsilon > 0$.

Therefore,

$$I_\varepsilon(x) \longrightarrow \int_{\mathbb{R}} \widehat{f}(\xi)e^{i\xi x}\, d\xi = \widehat{\widehat{f}}(-x),\ \forall x \in \mathbb{R}. \tag{19.1.2}$$

On the other hand, we may also write

$$I_\varepsilon(x) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-\frac{1}{2}\varepsilon^2\xi^2} e^{i\xi(x-y)}\, d\xi \right) f(y)\, dy.$$

The innermost integral is the FT of gaussian $e^{-\frac{\xi^2}{2(1/\varepsilon^2)}}$ evaluated at $x-y$, thus

$$\int_{\mathbb{R}} e^{-\frac{1}{2}\varepsilon^2\xi^2} e^{i\xi(x-y)}\, d\xi = \sqrt{\frac{2\pi}{\varepsilon^2}} e^{-\frac{(x-y)^2}{2\varepsilon^2}} = \delta_\varepsilon(x-y),$$

where (δ_ε) is the Gaussian approximate unit. Thus

$$I_\varepsilon(x) = \int_{\mathbb{R}} f(y)\delta_\varepsilon(x-y)\, dy = f * \delta_\varepsilon(x).$$

Now, by (11.2.1), $f * \delta_\varepsilon \xrightarrow{L^1} f$ for $\varepsilon \downarrow 0$. We know that this is not necessarily a point wise limit, however, extracting a suitable subsequence, we may say that

$$I_\varepsilon(x) = f * \delta_\varepsilon(x) \longrightarrow f(x),\ a.e.\ x \in \mathbb{R}.$$

Combining this with (19.1.2) we get the conclusion.

Example 19.1.2: Cauchy distribution $(*)$

$$\widehat{\frac{1}{a^2+\sharp^2}}(\xi) = \frac{1}{2a}e^{-a|\xi|},\quad (a>0). \tag{19.1.3}$$

PROOF. This is a calculation that, to be done by the definition, requires non trivial techniques of calculus for integrals. He we illustrate how inversion formula provides a remarkable shortcut. Recall that, according to (17.1.4)

$$\widehat{e^{-a|\sharp|}}(\xi) = \frac{2a}{a^2+\xi^2},\ \xi \in \mathbb{R},\quad (a>0).$$

From this, it is evident that $\widehat{\frac{1}{2a}e^{-a|\sharp|}}(\xi) = \frac{1}{a^2+\xi^2} \in L^1(\mathbb{R})$ and since also $\frac{1}{2a}e^{-a|\sharp|} \in L^1(\mathbb{R})$, according to inversion formula we obtain

$$\widehat{\frac{1}{a^2+\sharp^2}}(\xi) = \widehat{\widehat{\frac{1}{2a}e^{-a|\sharp|}}}(\xi) = \frac{1}{2a}e^{-a|-\xi|} = \frac{1}{2a}e^{-a|\xi|}.$$

Example 19.1.3: $(**)$

Let $f(x) := \frac{1}{(1+x^2)^2}$.

i) Use multiplication-derivation duality to compute $\widehat{\sharp f(\sharp)}$ (hint: $xf(x) = \partial_x \ldots$).
ii) Use i) to determine $\widehat{f}$.

iii) Use $\widehat{f}$ to compute

$$\int_0^{+\infty} \frac{1}{(1+x^2)^2}\,dx, \quad \int_0^{+\infty} \frac{\sin x}{(1+x^2)^2}\,dx.$$

Proof. i) We have

$$xf(x) = \frac{x}{(1+x^2)^2} = -\frac{1}{2}\partial_x \frac{1}{1+x^2},$$

so

$$\widehat{\sharp f(\sharp)}(\xi) = -\frac{1}{2}\partial_x \widehat{\frac{1}{1+\sharp^2}}(\xi) = -\frac{1}{2}(i\xi)\widehat{\frac{1}{1+\sharp^2}}(\xi) = \frac{i}{4}\xi e^{-|\xi|}.$$

ii) Now, recalling that

$$\partial_\xi \widehat{f}(\xi) = \widehat{i\sharp f(\sharp)}(\xi) = i\frac{i}{4}\xi e^{-|\xi|} = -\frac{1}{4}\xi e^{-|\xi|}.$$

In particular $\widehat{f}$ is a primitive of $-\frac{1}{4}\xi e^{-|\xi|}$. Let's determine this. Because of the modulus, we distinguish $\xi \geqslant 0$ by $\xi \leqslant 0$. In the first case,

$$\widehat{f}(\xi) = -\frac{1}{4}\int \xi e^{-\xi}\,d\xi + c = \frac{1}{4}\left(\xi e^{-\xi} - \int e^{-\xi}\,d\xi\right) + c = \frac{1}{4}\left(\xi e^{-\xi} + e^{-\xi}\right) + c$$

In the second case

$$\widehat{f}(\xi) = -\frac{1}{4}\int \xi e^{\xi}\,d\xi + c = -\frac{1}{4}\left(\xi e^{\xi} - \int e^{\xi}\,d\xi\right) + c = -\frac{1}{4}\left(\xi e^{\xi} - e^{\xi}\right) + c'$$

To determine c, c' we notice that, since $f \in L^1(\mathbb{R})$, according to RL Lemma, $\widehat{f}(\xi) \longrightarrow 0$ as $|\xi| \longrightarrow +\infty$. In particular we get easily that $c = c' = 0$. The conclusion is

$$\widehat{f}(\xi) = \begin{cases} \xi \geqslant 0, & = \frac{1}{4}\left(\xi e^{-\xi} + e^{-\xi}\right), \\ \xi \leqslant 0, & = -\frac{1}{4}\left(\xi e^{\xi} - e^{\xi}\right) \end{cases} = \frac{1}{4}e^{-|\xi|}\left(|\xi| + 1\right).$$

iii) We can easily reduce the two integrals to suitable Fourier integrals

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \frac{1}{(1+x^2)^2} e^{-i\xi x}\,dx.$$

Indeed: in the first case we have

$$\int_0^{+\infty} \frac{1}{(1+x^2)^2}\,dx = \frac{1}{2}\int_{\mathbb{R}} \frac{1}{(1+x^2)^2}\,dx = \frac{1}{2}\widehat{f}(0) = \frac{1}{4}.$$

About the second, recalling that $\sin x = \frac{e^{ix} - e^{-ix}}{2i}$, we have

$$\begin{aligned} \int_0^{+\infty} \frac{\sin x}{(1+x^2)^2}\,dx &= \frac{1}{2i}\left(\int_0^{+\infty} \frac{e^{ix}}{(1+x^2)^2}\,dx + \int_0^{+\infty} \frac{e^{-ix}}{(1+x^2)^2}\,dx\right) = \frac{1}{2}\int_{\mathbb{R}} \frac{1}{(1+x^2)^2} e^{-ix}\,dx \\ &= \frac{1}{2}\widehat{f}(1) = \frac{1}{2e}. \end{aligned}$$

19.2. Inverse Fourier Transform

We can look at FT as a transformation that applies $f \in L^1$ into $\widehat{f}$. If the domain of this transformation is clear (L^1), not the same can be said for its co-domain. For instance, according to Riemann-Lebesgue's

Lemma 18.1.2, we know that $\widehat{f} \in \mathscr{C}(\mathbb{R})$ and $\widehat{f}(\pm\infty) = 0$, so we can say that

$$\widehat{\ }: L^1(\mathbb{R}) \longrightarrow \mathscr{C}_0(\mathbb{R}) =: \{g \in \mathscr{C}(\mathbb{R}) \ : \ g(\pm\infty) = 0\}.$$

Notice that the bound (18.1.1) tells that this mapping is continuous. Inversion theorem implies also injectivity:

Proposition 19.2.1

The FT is injective, that is

$$\widehat{f} = \widehat{g} \implies f = g \text{ a.e..}$$

PROOF. If $\widehat{f} = \widehat{g}$ then, by linearity $\widehat{f-g} = 0$. Now, since $f - g \in L^1$ and, trivially, $\widehat{f-g} = 0 \in L^1$, according to the inversion formula,

$$(f-g)(x) = \frac{1}{2\pi}\widehat{\widehat{f-g}}(-x) = \frac{1}{2\pi}\widehat{0}(-x) \equiv 0, \text{ a.e.}$$

We may wonder if the FT is also surjective, that is, a bijection from L^1 to $\mathscr{C}_0$. Unfortunately, this is false. This fact makes the *inversion problem* non trivial: *given a function $g = g(\xi)$, determine (if any) a* **Fourier original** *of g, that is a function f such that $\widehat{f} = g$.* A partial answer to this problem is provided by the following

Corollary 19.2.2

Let $g \in L^1(\mathbb{R})$ be such that $\widehat{g} \in L^1(\mathbb{R})$. Then, there exists a unique Fourier original for g,

$$f(x) = \frac{1}{2\pi}\widehat{g}(-x), \ a.e.\ x \in \mathbb{R}.$$

PROOF. If $g, \widehat{g} \in L^1(\mathbb{R})$ then

$$g(\xi) = \frac{1}{2\pi}\widehat{\widehat{g}}(-\xi).$$

Now, setting $f(x) := \frac{1}{2\pi}\widehat{g}(-\sharp)(x)$, recalling the properties of FT, we have $\widehat{f}(\xi) = \frac{1}{2\pi}\widehat{\widehat{g}}(-\xi) = g(\xi)$.

The "inverse" of FT is the operation

$$\check{g}(x) := \frac{1}{2\pi}\widehat{g}(-x) = \frac{1}{2\pi}\int_{\mathbb{R}} g(\xi)e^{ix\xi}\, d\xi.$$

Basically, this is again the FT again. This might leads to think that "perhaps" FT is a bijection on L^1. This is false! As we know, $\text{rect}_1 \in L^1$ but $\widehat{\text{rect}_1}(\xi) = \frac{\sin(\xi)}{\xi} \notin L^1$. In fact, the image of L^1 FT is difficult to be characterized. So, what we can say is that FT is a bijection on a subspace of L^1,

$$\{f \in L^1 \ : \ \widehat{f} \in L^1\}.$$

Unfortunately, since $\widehat{f}$ cannot be computed explicitly, it is hard to characterize condition $f, \widehat{f} \in L^1$. Nonetheless, noticed that, for $f \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$ we have also $\widehat{f} \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$, we can say that *the FT is a bijection on the Schwarz space $\mathcal{S}(\mathbb{R})$.*

19.3. Exercises

Exercise 19.3.1 (∗∗). *Let*

$$f(x) := \frac{\sin x}{x(1+x^2)}.$$

Compute the FT of f.

Exercise 19.3.2 (∗∗). *Let $a, b > 0$, $a \neq b$, and set*

$$g_{a,b}(\xi) := \frac{1}{(\xi^2+a^2)(\xi^2+b^2)},\ \xi \in \mathbb{R},$$

i) *Show that $g_{a,b}$ has a Fourier original in L^1 and compute it. (hint: split the fraction and recall that $\widehat{e^{-\lambda|\sharp|}}(\xi) = \dots$).*
ii) *Show that $\sharp g(\sharp)$ has a Fourier original in L^1 and find it in term of the original f of g. Justify carefully your answer.*

Exercise 19.3.3 (∗∗). *Let*

$$g(\xi) := \frac{\xi\cos\xi - \sin\xi}{\xi^2},\ \xi \neq 0.$$

i) *Is $g \in L^1$? Is $\widehat{g} \in L^1$? Justify carefully.*
ii) *Discuss the problem of determining a Fourier original for g and determine it (if any).*

Exercise 19.3.4 (∗∗+). *Let*

$$g(\xi) = \frac{1}{1+\xi^4},\ \xi \in \mathbb{R}.$$

Show that f admits an L^1 Fourier original and determine it. (hint: $(\xi^4+1) = (\xi^2+\sqrt{2}\xi+1)(\xi^2-\sqrt{2}\xi+1)$)

Exercise 19.3.5 (∗∗+). *Let $g(\xi) = \frac{\xi}{1+\xi^4}$.*

i) *Show that g has a Fourier original f.*
ii) *Compute $\int_{\mathbb{R}} xf(x)\,dx$ and $f'(0)$.*

Exercise 19.3.6 (∗∗). *Let $f \in L^1(\mathbb{R})$ be such that $\xi\widehat{f}(\xi) \in L^1(\mathbb{R})$.*

i) *Deduce that f is a.e. continuous. (hint: check that $\widehat{f} \in L^1(\mathbb{R})\dots$).*
ii) (+) *Show that f has weak derivative $g(x) := \widehat{i\sharp\widehat{f}}(-x)$.*

Exercise 19.3.7 (∗∗). *For $a > 0$, let $f_a(x) := \frac{1}{x^2+a^2}$. Use the FT to compute $f_a * f_b$ for $a, b > 0$.*

Exercise 19.3.8 (∗∗+). *Solve the equation*

$$\int_{\mathbb{R}} f(x-y)e^{-|y|}\,dy = e^{-2|x|},\ x \in \mathbb{R},$$

in the unknown $f \in L^1(\mathbb{R})$.

Exercise 19.3.9 (∗∗). *Let $f_a(x) := e^{-ax}1_{[0,+\infty[}(x)$.*

i) *Compute the FT of f_a.*
ii) *Let $g(\xi) := \frac{1}{(\xi+i)^2}$. Is $g \in \mathscr{C}_0$? If yes, discuss the problem of determining a Fourier original for g.*

LECTURE 20

L^2 Fourier Transform

Apparently, FT can be defined only for L^1 functions. The resulting operation has important features but also a number of limitations. The major of these is, perhaps, the fact that we cannot clearly characterize when inversion formula holds. In this Lecture we show that FT can be defined on $L^2(\mathbb{R})$. This (new) transform has a big pro: it is a bijection and, even more, an isometry, on $L^2(\mathbb{R})$. In particular, inversion formula holds for every L^2 function. There is, of course, something to pay, and this is with the definition of the Transform, for which we do not dispose a formula, unless the function is also in L^1, and in this case the L^2 transform coincides with the L^1 definition.

20.1. Duality lemma

Differently from the case when the measure of the domain is finite, there is no inclusion between two different L^p spaces. For instance:

- $f(x) = \frac{1}{1+|x|}$, $f \in L^2(\mathbb{R})$ but $f \notin L^1(\mathbb{R})$;
- $f(x) = \frac{1}{\sqrt{|x|}(1+|x|)}$, $f \in L^1(\mathbb{R})$ but $f \notin L^2(\mathbb{R})$.

Thus, it is not evident how can we define $\widehat{f}$ for $f \in L^2(\mathbb{R})$ when $f \notin L^1(\mathbb{R})$. Nonetheless, the following Lemma suggests that FT should have some remarkable property with the geometry of L^2:

Lemma 20.1.1: duality Lemma

Let $f, g \in L^1(\mathbb{R})$. Then

$$\int_{\mathbb{R}} f\widehat{g} = \int_{\mathbb{R}} \widehat{f} g. \tag{20.1.1}$$

Proof. First notice that both members of identity (20.1.1) are well defined: $f, g \in L^1(\mathbb{R})$ implies $\widehat{f}, \widehat{g} \in L^\infty$, so $\widehat{f}g, f\widehat{g} \in L^1$. The proof of(20.1.1) is just an easy computation:

$$\begin{aligned}\int_{\mathbb{R}} f(x)\widehat{g}(x)\,dx &= \int_{\mathbb{R}} f(x)\left(\int_{\mathbb{R}} g(y)e^{-ixy}\,dy\right)\,dx \overset{Fubini}{=} \int_{\mathbb{R}} g(y)\left(\int_{\mathbb{R}} f(x)e^{-ixy}\,dx\right)\,dy \\ &= \int_{\mathbb{R}} g(y)\widehat{f}(y)\,dy.\end{aligned}$$

Apparently, the duality formula (20.1.1) can be interpreted as

$$\langle f, \widehat{g}\rangle = \langle \widehat{f}, g\rangle.$$

However, this is not correct for two good reasons. First, in general none of $f, g, \widehat{f}, \widehat{g} \in L^2$, so the scalar product does not make any sense. Second, $\widehat{f}$ and $\widehat{g}$ are $\mathbb{C}$ valued functions. Thus the natural L^2 structure for this case should have hermitian product

$$\langle f, g\rangle = \int_{\mathbb{R}} f\overline{g}.$$

To cope these two objections, we consider now the L^1 FT restricted to Schwartz class $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ with respect to hermitian product. It holds:

Proposition 20.1.2

Let $\langle\cdot,\cdot\rangle$ the standard hermitian product of $L^2(\mathbb{R})$. Then

$$\langle \widehat{f}, \widehat{g}\rangle = \frac{1}{2\pi}\langle f, g\rangle, \quad \forall f, g \in \mathcal{S}(\mathbb{R}). \tag{20.1.2}$$

In particular,

$$\|\widehat{f}\|_2 = \frac{1}{2\pi}\|f\|_2, \ \forall f \in \mathcal{S}(\mathbb{R}). \tag{20.1.3}$$

Proof. Let $f, g \in \mathcal{S}(\mathbb{R})$. Then,

$$\langle \widehat{f}, \widehat{g}\rangle = \int_{\mathbb{R}} \widehat{f}\,\overline{\widehat{g}} \overset{(20.1.1)}{=} \int_{\mathbb{R}} f\,\widehat{\overline{\widehat{g}}}.$$

Now,

$$\overline{\widehat{g}(\xi)} = \overline{\int_{\mathbb{R}} g(y)e^{-i\xi y}\,dy} = \int_{\mathbb{R}} \overline{g(y)}e^{i\xi y}\,dy = \widehat{\overline{g}}(-\xi).$$

Therefore, since inversion formula holds for Schwartz functions,

$$\widehat{\overline{\widehat{g}}}(x) = \widehat{\widehat{\overline{g}}}(-x) = \frac{1}{2\pi}\overline{g}(-(-x)) = \frac{1}{2\pi}\overline{g}(x),$$

thus

$$\langle \widehat{f}, \widehat{g}\rangle = \frac{1}{2\pi}\int_{\mathbb{R}} f\overline{g} = \frac{1}{2\pi}\langle f, g\rangle.$$

The (20.1.3) follows taking $f = g$ in previous identity.

20.2. Plancherel Theorem

We already know that FT maps $\mathcal{S}(\mathbb{R})$ into itself. The proposition 20.1.2 says that if we look at $\mathcal{S}(\mathbb{R})$ as a subspace of $L^2(\mathbb{R})$, then

$$\widehat{\ }: \mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$$

preserves length of vectors and angles modulo a scaling factor. This is the key for the

Theorem 20.2.1: Plancherel

There exists a unique extension of FT to $L^2(\mathbb{R})$. This extension, called **Fourier-Plancherel Transform** (FPT),

i) coincides with usual L^1 FT for $f \in L^1 \cap L^2$;

ii) fulfils

$$\langle \widehat{f}, \widehat{g} \rangle = \frac{1}{2\pi} \langle f, g \rangle, \ \forall f, g \in L^2(\mathbb{R}). \tag{20.2.1}$$

and the **Parseval identity**,

$$\|\widehat{f}\|_2 = \frac{1}{\sqrt{2\pi}} \|f\|_2, \ \forall f \in L^2(\mathbb{R}). \tag{20.2.2}$$

iii) fulfils inversion formula:

$$f(x) \overset{a.e.}{=} \frac{1}{2\pi} \widehat{\widehat{f}}(-x), \ \forall f \in L^2(\mathbb{R}). \tag{20.2.3}$$

iv) is a bijection on $L^2(\mathbb{R})$ with inverse

$$\check{f}(x) = \frac{1}{2\pi} \widehat{f}(-x).$$

PROOF. We start proving the existence of the transform. Let $f \in L^2$: by density, there exists $(f_n) \subset \mathcal{S}$ such that $f_n \xrightarrow{L^2} f$. Consider $(\widehat{f_n}) \subset \mathcal{S} \subset L^2$. We claim that this is a Cauchy sequence. Indeed,

$$\|\widehat{f_n} - \widehat{f_m}\|_2 = \|\widehat{f_n - f_m}\|_2 \overset{(20.1.3)}{=} \frac{1}{2\pi} \|f_n - f_m\|_2.$$

Since (f_n) is L^2 convergent, (f_n) is a Cauchy sequence, hence also $(\widehat{f_n})$ it is. We define

$$\widehat{f} \overset{L^2}{:=} \lim_n \widehat{f_n}.$$

This definition is independent of any particular approximating sequence (f_n). Indeed, if $(g_n) \subset \mathcal{S}$, $g_n \xrightarrow{L^2} f$, then

$$\|\widehat{f_n} - \widehat{g_n}\|_2 = \|\widehat{f_n - g_n}\|_2 = \frac{1}{2\pi} \|f_n - g_n\|_2 \longrightarrow \frac{1}{2\pi} \|f - f\|_2 = 0,$$

that is $\lim_n \widehat{f_n} \overset{L^2}{=} \lim_n \widehat{g_n}$. With this the existence is shown. It remains to verify properties i) to iv).

i) Let $f \in L^1 \cap L^2$. We temporarily denote by $\widehat{f}$ the L^1 FT, and by $\widetilde{f}$ the "new" L^2 FPT. The claim is $\widehat{f} = \widetilde{f}$ a.e. Let $(f_n) \subset \mathcal{S}$ be such that $f_n \xrightarrow{L^1, L^2} f$. This is possible according to the mollification theorem. Therefore, on one side, $\widehat{f_n} \xrightarrow{L^2} \widetilde{f}$, so, modulo a subsequence, $\widehat{f_n} \xrightarrow{a.e.} \widetilde{f}$. On the other side, by RL Lemma bound $\|\widehat{f_n} - \widehat{f}\|_\infty \leqslant \|f_n - f\|_1$, so, in particular, $\widehat{f_n} \xrightarrow{pw} \widehat{f}$, therefore $\widehat{f_n} \xrightarrow{a.e.} \widehat{f}$. But then $\widehat{f} = \widetilde{f}$ a.e.

ii) Both isometry (20.2.1) and Parseval (20.2.2) identities follows from the definition of FPT. For instance: if $f, g \in L^2$ and $\widehat{f} \overset{L^2}{=} \lim_n \widehat{f_n}$, $\widehat{g} \overset{L^2}{=} \lim_n \widehat{g_n}$, with $(f_n), (g_n) \subset \mathcal{S}$, then, by the continuity of the inner product,

$$\langle \widehat{f}, \widehat{g} \rangle \longleftarrow \langle \widehat{f_n}, \widehat{g_n} \rangle = \frac{1}{2\pi} \langle f_n, g_n \rangle \longrightarrow \frac{1}{2\pi} \langle f, g \rangle.$$

iii) Let $f \in L^2$ and $\widehat{f} \overset{L^2}{=} \lim_n \widehat{f_n}$ with $(f_n) \subset \mathcal{S}$. Notice that, since FT applies $\mathcal{S}$ into itself, $(\widehat{f_n}) \subset \mathcal{S}$, thus according to the definition of FPT,

$$\widehat{\widehat{f}} \overset{L^2}{=} \lim_n \widehat{\widehat{f_n}}$$

Since inversion formula holds on Schwartz functions in strong form $\hat{\hat{f}}_n(x) \equiv 2\pi f_n(-x)$, we have

$$\widehat{\widehat{f}} \overset{L^2}{=} \lim_n 2\pi f_n(-\sharp) \overset{L^2}{=} 2\pi f(-\sharp), \implies \widehat{\widehat{f}(x)} \overset{a.e.}{2} \pi f(-x),$$

from which conclusion follows.
iv) exercise.

Here is a genuine example of FPT.

Example 20.2.2: $(*)$

Let $\mathrm{sinc}(x) := \frac{\sin x}{x}$. Then $\mathrm{sinc} \in L^2 \backslash L^1$ and

$$\widehat{\mathrm{sinc}}(\xi) = \frac{1}{2} 1_{[-2,2]}(x).$$

In particular, this example shows that, differently from L^1 FT, the FPT of f is not necessarily a continuous function.

Proof. First, $\mathrm{sinc} \in L^2$. Indeed,

$$\int_{\mathbb{R}} |\mathrm{sinc}\, x|^2 \, dx = \int_{\mathbb{R}} \left| \frac{\sin x}{x} \right|^2 \, dx.$$

Function $\left(\frac{\sin x}{x}\right)^2$ is continuous in $x = 0$ and it is also bounded by $\frac{1}{x^2}$ at $\pm\infty$. Therefore it is integrable on $\mathbb{R}$, that is $\mathrm{sinc} \in L^2$ and the L^2 FT can be computed. Notice, however, that since $\mathrm{sinc} \notin L^1$, the L^1 FT of sinc is not defined. Now, recalling that

$$\widehat{\mathrm{rect}_1}(\xi) = 2\, \mathrm{sinc}(\xi),$$

applying both sides the L^2 FT and recalling the inversion formula

$$\widehat{\mathrm{sinc}}(\xi) = \frac{1}{2}\widehat{\widehat{\mathrm{rect}_1}}(\xi) = \frac{1}{2}\mathrm{rect}_1(-x) = \frac{1}{2}\mathrm{rect}_1(x) = \frac{1}{2} 1_{[-2,2]}(x).$$

In the previous example, the FPT has been computed through a lucky trick. In general, for $f \in L^2$, there is no integral representation for FPT. The most closest to be a formula is provided by the

Proposition 20.2.3

Let $f \in L^2(\mathbb{R})$. Then

$$\hat{f}(\xi) \overset{L^2}{=} \lim_{R \to +\infty} \int_{-R}^{R} f(x) e^{-i\xi x} \, dx. \tag{20.2.4}$$

We leave the proof of formula (20.2.4) to exercises. In any case, this formula has been handled with care: indeed, the limit is in L^2 sense, and as we know, it is not necessarily a point wise limit. We might have infact that r.h.s. of (20.2.4) is never convergent!

20.3. Properties of FPT

FPT fulfills basically the same properties of FT. In particular, we have same formulas for FPT of derivative, derivative of FPT, FPT of a convolution. Of course, statements have to be adapted to the L^2 setup and proofs have to be redone from scratch because L^1 proofs rely on the integral representation of FT. We will leave most of the proofs for the exercises.

20.3.1. Duality Lemma.

Duality lemma extends to L^2 in a quite simple and natural way:

Lemma 20.3.1

Let $f, g \in L^2(\mathbb{R})$. Then

$$\int_{\mathbb{R}} f\widehat{g} = \int_{\mathbb{R}} \widehat{f} g$$

Proof. We have

$$\int_{\mathbb{R}} f\widehat{g} = \int_{\mathbb{R}} f\overline{\overline{\widehat{g}}} = \langle f, \overline{\widehat{g}}\rangle = \langle f, \widehat{\overline{g}}(-\sharp)\rangle = \frac{1}{2\pi}\langle \widehat{f}, \widehat{\widehat{\overline{g}}}(-\sharp)\rangle = \frac{1}{2\pi}\langle \widehat{f}, 2\pi\overline{g}(-(-\sharp))\rangle = \langle \widehat{f}, \overline{g}\rangle = \int_{\mathbb{R}} \widehat{f} g.$$

20.3.2. Derivative of FPT.

Respect to FT properties, for the FPT we need to start from the transform of derivative. We have the

Proposition 20.3.2

Assume $f, xf \in L^2(\mathbb{R})$. Then

$$\exists \partial_\xi \widehat{f}(\xi) = \widehat{-i\sharp f}(\xi). \tag{20.3.1}$$

Proof. The first remark is that, since $f, xf \in L^2$, it follows $f \in L^1$. Indeed,

$$\int_{\mathbb{R}} |f| = \int_{\mathbb{R}} \frac{1}{1+|x|}(1+|x|)f \overset{CS}{\leqslant} \left(\int_{\mathbb{R}} \frac{1}{(1+|x|)^2}\,dx\right)^{1/2} \left(\int_{\mathbb{R}} (1+|x|)^2|f(x)|^2\,dx\right)^{1/2} < +\infty.$$

Thus, $f \in L^1 \cap L^2$ and $\widehat{f}$ makes sense as FT too. Now, to prove (20.3.1), we have to verify that

$$\widehat{f}(\xi) - \widehat{f}(0) = \int_0^\xi \widehat{-i\sharp f}(\eta)\,d\eta.$$

To this aim, notice that, by duality,

$$\int_0^\xi \widehat{-i\sharp f} = \int_{\mathbb{R}} 1_{[0,\xi]}\widehat{-i\sharp f} = -i\int_{\mathbb{R}} \widehat{1_{[0,\xi]}}(x) x f(x)\,dx.$$

Now,

$$\widehat{1_{[0,\xi]}}(x) = \widehat{1_{[-\xi/2,\xi/2]}(\sharp - \xi/2)}(x) = e^{-i\frac{\xi}{2}x}\widehat{\mathrm{rect}_{\xi/2}}(x) = e^{-i\frac{\xi}{2}x} 2\frac{\xi}{2}\frac{\sin(\frac{\xi}{2}x)}{\frac{\xi}{2}x} = \frac{2}{x}e^{-i\frac{\xi x}{2}}\sin\frac{\xi x}{2},$$

and, recalling of Euler formula, $\sin\frac{\xi x}{2} = \frac{1}{2i}\left(e^{i\frac{\xi x}{2}} - e^{-i\frac{\xi x}{2}}\right)$, we obtain

$$\widehat{1_{[0,\xi]}}(x) = \frac{2}{x}e^{-i\frac{\xi x}{2}}\frac{1}{2i}\left(e^{i\frac{\xi x}{2}} - e^{-i\frac{\xi x}{2}}\right) = \frac{1}{ix}\left(1 - e^{i\xi x}\right),$$

thus

$$\int_0^{\xi} \widehat{-i\sharp f} = -i\int_{\mathbb{R}} \frac{1}{ix}\left(1 - e^{i\xi x}\right) x f(x)\, dx = \widehat{f}(\xi) - \widehat{f}(0).$$

20.3.3. FPT of the derivative. We can now show the FPT of the derivative formula:

Proposition 20.3.3

Assume $f, \partial_x f \in L^2(\mathbb{R})$. Then

$$\widehat{\partial_x f}(\xi) = i\xi\widehat{f}(\xi).$$

PROOF. Since $\mathcal{S}(\mathbb{R})$ is dense in L^2, we have that

$$\widehat{\partial_x f} = i\xi\widehat{f}, \quad \Longleftrightarrow \quad \langle\widehat{\partial_x f}, g\rangle = \langle i\sharp\widehat{f}, g\rangle,\ \forall g \in \mathcal{S}(\mathbb{R}).$$

Now, by duality,

$$\langle i\sharp\widehat{f}, g\rangle = \int_{\mathbb{R}} i\xi\widehat{f}(\xi)\overline{g(\xi)}\, d\xi = \int_{\mathbb{R}} f(x)\widehat{i\sharp\overline{g}}(x)\, dx = -\int_{\mathbb{R}} f(x)\partial_x\widehat{\overline{g}}(x)\, dx.$$

We now aim to integrate by parts. First notice that

$$f(b)^2 - f(a)^2 = \int_a^b \partial_x f^2(x)\, dx = \int_a^b 2f\partial_x f\, dx.$$

Being $f, \partial_x f \in L^2$, by Cauchy-Schwarz inequality we have $f\partial_x f \in L^1$, thus, letting $a \longrightarrow -\infty$ and $b \longrightarrow +\infty$ we get that $\exists f(\pm\infty)$, and since $f \in L^2$, necessarily $f(\pm\infty) = 0$. Therefore, recalling also that $\widehat{\overline{g}} \in \mathcal{S}$,

$$\int_{\mathbb{R}} f(x)\partial_x\widehat{\overline{g}}(x)\, dx = \left[f(x)\widehat{\overline{g}}(x)\right]_{x=-\infty}^{x=+\infty} - \int_{\mathbb{R}} \partial_x f(x)\widehat{\overline{g}}(x)\, dx = -\int_{\mathbb{R}} \partial_x f(x)\widehat{\overline{g}}(x)\, dx.$$

Therefore,

$$\langle i\sharp\widehat{f}, g\rangle = +\int_{\mathbb{R}} \partial_x f\widehat{\overline{g}} = \int_{\mathbb{R}} \widehat{\partial_x f}\overline{g} = \langle\widehat{\partial_x f}, g\rangle,$$

as desired.

20.3.4. FPT of convolution. Also for the convolution, the FPT converts convolution product into an algebraic product.

Proposition 20.3.4

Let $f \in L^1(\mathbb{R})$ and $g \in L^2(\mathbb{R})$. Then

$$\widehat{f * g} = \widehat{f}\widehat{g}.$$

PROOF. By Young inequality $f * g \in L^2$, thus its FPT is well defined. Let $(g_n) \subset \mathscr{S}$ such that $g_n \xrightarrow{L^2} g$ and consider $f * g_n$. According to Young inequality $(f * g_n) \subset L^2$. We claim that $f * g_n \xrightarrow{L^2} f * g$. This follows by Young inequality:

$$\|f * g - f * g_n\|_2 = \|f * (g - g_n)\|_2 \leqslant \|f\|_1 \|g - g_n\|_2 \longrightarrow 0.$$

Thus,

$$\widehat{f * g_n} \xrightarrow{L^2} \widehat{f * g}.$$

Now, since $(g_n) \subset \mathscr{S} \subset L^1$ we have also that

$$\widehat{f * g_n} = \widehat{f}\widehat{g_n}.$$

We claim that this converges in L^2 to $\widehat{f}\widehat{g}$. Indeed,

$$\|\widehat{f}\widehat{g} - \widehat{f}\widehat{g_n}\|_2 = \|\widehat{f}(\widehat{g} - \widehat{g_n})\|_2 \overset{RL}{\leqslant} \|\widehat{f}\|_\infty \|\widehat{g} - \widehat{g_n}\|_2 \longrightarrow 0.$$

In conclusion,

$$\widehat{f}\widehat{g} \xleftarrow{L^2} \widehat{f}\widehat{g_n} = \widehat{f * g_n} \xrightarrow{L^2} \widehat{f * g}.$$

20.4. Exercises

Exercise 20.4.1 ($*$)**.** *Compute the FPT of* $f(x) := \frac{x}{1+x^2}$.

Exercise 20.4.2 ($**$)**.** *Let* $f(x) := \frac{1}{x+i}$.

i) *Is* $f \in L^(\mathbb{R})$*? Is* $f \in L^2(\mathbb{R})$*?*
ii) *Show that at least one of FT or FPT of* f *exists and determine* $\widehat{f}$ *(hint:* $\frac{1}{x+i} = \frac{x-i}{(x+i)(x-i)} = \frac{x-i}{x^2+1} \dots$*)*
iii) *Show that* $\widehat{f}$ *can have sense only in one of FT or FPT.*

Exercise 20.4.3 ($**+$)**.** *Show that if* $f \in L^2(\mathbb{R})$ *then*

$$\widehat{f}(\xi) \overset{L^2}{=} \lim_{R \to +\infty} \int_{-R}^{R} f(y) e^{-i\xi y}\, dy. \tag{20.4.1}$$

Use this to compute the FPT of

$$f(x) := \operatorname{sinc}(ax).$$

Exercise 20.4.4 ($**$)**.** *Let* $a, b > 0$*. Check that the integral*

$$\int_{-\infty}^{+\infty} \frac{\sin(at)\sin(bt)}{t^2}\, dt$$

is well defined and use the FPT to compute its value.

Exercise 20.4.5 ($**$)**.** *Let* $g(\xi) := \frac{1}{|\xi|^3+1}$.

i) *Is* $g \in L^1$*? Is* $g \in L^2$*? Has* g *a Fourier original in* L^1*? And in* L^2*? Justify your answers.*
ii) *If* g *has a Fourier original* f*, compute*

$$\int_{\mathbb{R}} |f * f'|^2\, dx.$$

Exercise 20.4.6 ($**+$)**.** *Let $f \in L^2(\mathbb{R})$ and*

$$f_k(x) := \int_k^{k+1} e^{ix\xi}\,\widehat{f}(\xi)\, d\xi.$$

i) *Check that the f_k are well defined and $f_k \in L^2(\mathbb{R})$ for every $k \in \mathbb{Z}$.*
ii) *Check that*

$$f \overset{L^2}{=} \sum_{k=-\infty}^{+\infty} f_k.$$

Exercise 20.4.7 ($**+$)**.** *By using FT, compute the convolution $f_a * f_b$ of two Cauchy distributions where $f_a(x) = \frac{1}{a^2+x^2}$. Use this to compute*

$$\lim_n \sqrt{n} \int_{\sqrt{n}\alpha}^{\sqrt{n}\beta} \underbrace{f_1 * f_1 * \cdots * f_1}_{n-times}\, dx.$$

Exercise 20.4.8 ($***$)**.** *Let $f \in L^2(\mathbb{R})$ with weak derivative $f \in L^2(\mathbb{R})$. Prove the Heisenberg inequality*

$$\|xf\|_2 \|\xi \widehat{f}\|_2 \geqslant \frac{\|f\|_2^2}{2}.$$

(hint: $\|f\|_2^2 = \int_{\mathbb{R}} |f|^2\, dx = \int (x)'|f|^2\, dx = \ldots$ Justify with care) Can you determine when equality holds?

Exercise 20.4.9 ($**+$)**.** Let $T : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R})$ be defined as

$$Tf := \widehat{f},\ f \in \mathcal{S}(\mathbb{R}).$$

Look at T as an operator on functions.

i) Prove that $T^4 = \mathbb{I}$ (here $T^4 = T \circ T \circ T \circ T$).
ii) Prove that all possible $\lambda \in \mathbb{C}$ such that $Tf = \lambda f$ for some $f \in \mathcal{S}(\mathbb{R})$, $f \neq 0$ are $\lambda = \pm 1, \pm i$.
iii) Determine f such that $Tf = f$.
iv) Use the multiplication-derivation duality $T(\partial_x f) = -i2\pi\xi T f$ to determine f such that $Tf = \lambda f$ for $\lambda = -1, \pm i$.

LECTURE 21

Applications to Integro-Differential Equations

FT is a versatile tool that can be exploited to solve different problems. Among other features, FT properties of derivatives and convolution play an important role to solve certain equations. In fact, FT converts derivatives into multiplication by polynomials and convolution products into algebraic products. As a consequence, certain differential or integral or combined integro-differential equations can be converted into algebraic equations. The idea is that, given an equation

$$\mathscr{E}[u] = 0,$$

in the unknown $u = u(x)$, we apply FT (or FPT) to both sides, obtaining an algebraic equation for $\widehat{u}$,

$$\mathscr{F}[\widehat{u}] = 0.$$

The idea is that this equation is easier than the original one, leading to a solution $\widehat{u}$. At this point, to get u we need to invert FT (or FPT) to compute $\check{\widehat{u}}$. In this Lecture we illustrate this method and ideas on a number of significant cases.

21.1. An ODE

We start with an example for which we do not need FT.

Example 21.1.1. *Determine all $u \in L^1$ with $u', u'' \in L^1$ solutions of the equation*

$$u'' - u = e^{-|x|}, \quad x \in \mathbb{R}.$$

Proof. Since $u, u', u'' \in \mathscr{L}^1$ and also $e^{-|x|} \in L^1$, we can apply FT to the equation: we get

$$\widehat{u'' - u} = \widehat{e^{-|\sharp|}}, \iff \widehat{u''} - \widehat{u} = \frac{2}{1+\xi^2}.$$

Since $\widehat{u''}(\xi) = (i\xi)^2\widehat{u} = -\xi^2\widehat{u}(\xi)$, we obtain the equation

$$-(\xi^2+1)\widehat{u} = \frac{2}{1+\xi^2}.$$

Here you may appreciate how an ODE has become an algebraic equation. We can solve this obtaining

$$\widehat{u} = -\frac{2}{(1+\xi^2)^2}. \tag{21.1.1}$$

Thus: *if $u \in L^1$ with $u', u'' \in L^1$ is a solution of the ODE, then $\widehat{u}$ is given by previous formula*. Now, to go back to u we need to solve the inversion problem (21.1.1). Since we are here in L^1 context (the argument for L^2 solutions would be easier), we may notice that $g := -\frac{2}{(1+\xi^2)^2} \in L^1$ and clearly also $g', g'' \in L^1$, so $\widehat{g} \in L^1$. Thus, equation (21.1.1) has a unique solution

$$u(x) = \frac{1}{2\pi}\widehat{g}(-x).$$

Instead of proceeding in the calculation of $\widehat{g}$ we may observe that

$$\left(\frac{2}{1+\xi^2}\right)^2 = \left(\widehat{e^{-|\sharp|}}\right)^2 = \widehat{e^{-|\sharp|}}\widehat{e^{-|\sharp|}} = \widehat{e^{-|\sharp|} * e^{-|\sharp|}},$$

so,

$$u(x) = -\frac{1}{2}e^{-|\sharp|} * e^{-|\sharp|}(x) = -\frac{1}{2}\int_{\mathbb{R}} e^{-|x-y|}e^{-|y|}\,dy$$

Computing the convolution

$$\int_{\mathbb{R}} e^{-|x-y|}e^{-|y|}\,dy = \int_{-\infty}^{0} e^{-|x-y|}e^{y}\,dy + \int_{0}^{+\infty} e^{-|x-y|}e^{-y}\,dy.$$

If $x \geqslant 0$, previous integrals are

$$= \int_{-\infty}^{0} e^{-(x-y)}e^{y}\,dy + \int_{0}^{x} e^{-(x-y)}e^{-y}\,dy + \int_{x}^{+\infty} e^{-(y-x)}e^{-y}\,dy$$

$$= e^{-x}\int_{-\infty}^{0} e^{2y}\,dy + e^{-x}\int_{0}^{x}\,dy + e^{x}\int_{x}^{+\infty} e^{-2y}\,dy$$

$$= \tfrac{e^{-x}}{2} + xe^{-x} + e^{x}\tfrac{e^{-2x}}{2} = e^{-x}(1+x),$$

while, as $x < 0$,

$$= \int_{-\infty}^{x} e^{-(x-y)}e^{y}\,dy + \int_{x}^{0} e^{x-y}e^{y}\,dy + \int_{0}^{+\infty} e^{-(y-x)}e^{-y}\,dy$$

$$= e^{-x}\int_{-\infty}^{x} e^{2y}\,dy + e^{x}\int_{x}^{0}\,dy + e^{x}\int_{0}^{+\infty} e^{-2y}\,dy$$

$$= \tfrac{e^{x}}{2} - xe^{x} + e^{x}\tfrac{1}{2} = e^{x}(1-x),$$

thus $u(x) = -\frac{1}{2}e^{-|x|}(1+|x|)$.

The conclusion of this argument is: *if $u, u', u'' \in L^1$ is a solution of the ODE then,* **necessarily**, $u(x) = -\frac{1}{2}e^{-|x|}(1+|x|)$. This, however, does not prove yet that such u is a solution: we should now verify that u has really first and second derivatives a.e. (easy) and this will close the argument. □

21.2. Convolution Equations

Convolution property (18.4.1) can be useful to solve certain integral equations that arises in application, where the integral part has convolution form.

EXAMPLE 21.2.1. *Determine for which values of real parameter λ the equation*

$$u(x) = \lambda\int_{\mathbb{R}} e^{-|x-y|}u(y)\,dy + e^{-|x|},$$

in the unknown $u \in L^1$ has a unique solution and, in that case, find it.

PROOF. The equation can be rewritten as

$$u = \lambda e^{-|\sharp|} * u + e^{-|\sharp|}.$$

By applying to both members the FT we obtain

$$\widehat{u} = \lambda\widehat{e^{-|\sharp|}}\widehat{u} + \widehat{e^{-|\sharp|}}, \quad\Longleftrightarrow\quad \left(1 - \lambda\widehat{e^{-|\sharp|}}\right)\widehat{u} = \widehat{e^{-|\sharp|}}.$$

Now, recall that by (17.1.4)

$$\widehat{e^{-|\sharp|}}(\xi) = \frac{2}{1+\xi^2},$$

therefore

$$\widehat{u}(\xi) = \frac{1}{1-\lambda\frac{2}{1+\xi^2}}\frac{2}{1+\xi^2} = \frac{2}{(1-2\lambda)+\xi^2}.$$

Let's look at this $\widehat{u}$. If $1-2\lambda = 0$, that is $\lambda = \frac{1}{2}$,

$$\widehat{u} = \frac{2}{\xi^2}$$

and this function cannot be the FT of an $u \in L^1$ ($\widehat{u}$ is not continuous). The same happens if $1-2\lambda < 0$: in this case we could write

$$(1-2\lambda)+\xi^2 = (\xi - \sqrt{2\lambda-1})(\xi+\sqrt{2\lambda-1}),$$

hence

$$\widehat{u} = \frac{2}{(\xi-\sqrt{2\lambda-1})(\xi+\sqrt{2\lambda-1})}.$$

By this easily we deduce $u \notin L^1$.

The conclusion of these remarks is that a solution $u \in L^1$ is possible only if $1-2\lambda > 0$, that is $\lambda < \frac{1}{2}$. In this case

$$\widehat{u} = \frac{2}{(1-2\lambda)+\xi^2} \overset{(17.1.4)}{=} \frac{1}{\sqrt{1-2\lambda}}\frac{2\sqrt{1-2\lambda}}{(\sqrt{1-2\lambda})^2+\xi^2} = \frac{1}{\sqrt{1-2\lambda}}\widehat{e^{-\sqrt{1-2\lambda}|\sharp|}}$$

that is

$$u(x) = \frac{1}{\sqrt{1-2\lambda}}e^{-\sqrt{1-2\lambda}|x|}. \quad \square$$

21.3. Heat Equation

The classical equation describing the heat diffusion on an infinite volume is the PDE

$$\begin{cases} \partial_t u(t,x) = \frac{\sigma^2}{2}\partial_{xx}u(t,x),\ t \geqslant 0,\ x \in \mathbb{R}, \\ \\ u(0,x) = \varphi(x),\ x \in \mathbb{R}, \end{cases} \tag{21.3.1}$$

Here $u = u(t,x)$ represents the temperature at time $t \geqslant 0$ on each point x of an infinite and homogeneous rod with initial temperature φ. By a suitable use of FT we can easily determine a formula for the solution u. To this aim we introduce the $x-$FT defined as

$$v(t,\xi) := \widehat{u(t,\sharp)}(\xi) \equiv \int_{\mathbb{R}} u(t,x)e^{-i\xi x}\,dx,\ \xi \in \mathbb{R}.$$

Of course we should do some assumption like $u(t,\sharp) \in L^1$ to define this. Assuming also $\partial_t u(t,\sharp) \in L^1$ we may write,

$$\widehat{\partial_t u(t,\sharp)}(\xi) = \int_{\mathbb{R}} \partial_t u(t,y)e^{-i\xi y}\,dy = \partial_t \int_{\mathbb{R}} u(t,y)e^{-i\xi y}\,dy = \partial_t\widehat{u(t,\sharp)}(\xi) = \partial_t v(t,\xi).$$

Here, the switching between derivation and integration requires some assumption. Let us skip this for the moment, we proceed as if everything can be computed. According to properties of the FT

$$\widehat{\partial_{xx}u(t,\sharp)}(\xi) = (-i\xi)^2\widehat{u(t,\sharp)}(\xi) = -\xi^2 v(t,\xi).$$

So, in term of v the heat equation becomes

$$\partial_t v(t,\xi) = -\frac{\sigma^2}{2}\xi^2 v(t,\xi), \ \ t \geqslant 0, \ \xi \in \mathbb{R}.$$

Moreover,

$$v(0,\xi) = \widehat{u(0,\sharp)}(\xi) = \widehat{\varphi}(\xi),$$

thus, to find v we have to solve the Cauchy problem

$$\left\{\begin{array}{ll} \partial_t v(t,\xi) = -\frac{\sigma^2}{2}\xi^2 v(t,\xi), & t \geqslant 0, \ \xi \in \mathbb{R}. \\ \\ v(0,\xi) = \widehat{\varphi}(\xi), & \xi \in \mathbb{R}. \end{array}\right. \tag{21.3.2}$$

For ξ fixed, this is a simple Cauchy problem for an first order linear equation. This can be easily solved leading to

$$v(t,\xi) = \widehat{\varphi}(\xi)e^{-\frac{1}{2}\sigma^2\xi^2 t}, \ t \geqslant 0, \ \xi \in \mathbb{R}.$$

Now recalling the Fourier transform of the Gaussian

$$e^{-\frac{1}{2}\sigma^2 t\xi^2} = \widehat{\frac{e^{-\frac{\sharp^2}{2\sigma^2 t}}}{\sqrt{2\pi\sigma^2 t}}}(\xi),$$

we have

$$\widehat{u(t,\sharp)}(\xi) = v(t,\xi) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\widehat{\varphi}(\xi)\widehat{e^{-\frac{\sharp^2}{2\sigma^2 t}}}(\xi) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\left(\widehat{\varphi * e^{-\frac{\sharp^2}{2\sigma^2 t}}}\right)(\xi),$$

that gives, finally,

$$u(t,x) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\left(\varphi * e^{-\frac{\sharp^2}{2\sigma^2 t}}\right)(x) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\int_{\mathbb{R}} \varphi(y)e^{-\frac{(x-y)^2}{2\sigma^2 t}}\,dy, \ t > 0, \ x \in \mathbb{R}. \tag{21.3.3}$$

Some remarks on this solution. First: it is not difficult to show that *if* $\varphi \in L^1$, $u \in \mathscr{C}^\infty(]0,+\infty[\times\mathbb{R})$ *and* u *is a solution of heat equation*. This fact can be checked by direct verification, that is computing $\partial_t u$ and $\partial_{xx}u$ (through derivation under integral sign) and checking that $\partial_t u = \frac{\sigma^2}{2}\partial_{xx}u$ for every $t > 0$ and $x \in \mathbb{R}$. (this is a technical but nice exercise that the reader is invited to do) Second: the u given by formula (21.3.3) is not defined at $t = 0$. This poses a problem: *in what sense* u *verifies the initial condition* $u(0,x) = \varphi(x)$*?* We may provide the following justification. Since

$$\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{\sharp^2}{2\sigma^2 t}} =: \delta_{1/\sqrt{\sigma^2 t}},$$

is an approximate unit,

$$u(t,\sharp) = \varphi * \delta_{1/\sqrt{\sigma^2 t}} \xrightarrow{L^1} \varphi, \ t \downarrow 0, \ \text{if } \varphi \in L^1(\mathbb{R}).$$

Thus (21.3.3) fulfils the initial condition in a "weak form". Last remark: formula (21.3.3) makes sense for $\varphi \in L^p$, $1 \leqslant p \leqslant +\infty$. In particular, $\varphi \in L^\infty$ makes sense. This is apparently conflicting with the argument that led to (21.3.3), because at certain point we needed $\widehat{\varphi}$ and, in general, there is no $\widehat{\varphi}$ for a $\varphi \in L^1$. However, formula (21.3.3) makes sense. As pointed out in the first remark, it is not difficult to check that (also in case $\varphi \in L^\infty$), $u \in \mathscr{C}^\infty(]0,+\infty[\times\mathbb{R})$ and it solves the heat equation. In other words, formula (21.3.3), derived through FT, goes much beyond the original problem setup.

21.4. The Black–Sholes Equation

The *Black–Sholes equation* is a PDE describing the behavior of the *value* of a *financial derivative* written over a risky asset whose price is stochastic. We do not enter in the derivation of the equation, we will limit to a qualitative description of the model.

The *Black–Sholes model* describes a simple *market model* where two assets are available:

- a *risk free asset*, called *bank bond*, delivering a deterministic instantaneous return rate r, that is

$$\frac{dB(t)}{B(t)} = r\ dt.$$

- a *risky asset*, called *stock*, delivering a stochastic instantaneous return

$$\frac{dS(t)}{S(t)} = \lambda\ dt + \text{gaussian r.v. mean } 0 \text{ and variance } \sigma^2\ dt.$$

The uncertainty delivered by S makes uncertain any investment on it. In other words, while $B(T)$ is deterministic, $S(T)$ is a random variable. It is therefore natural to look to forms of *protection against financial risks* ensuring a final payment $F(S(T))$. This payment should be delivered by the issuer to the owner of such a contract at time T. The contract is written at time $t = 0$ (initial time), when the owner pays a sum to the issuer to receive such a right. A major question is: *how much should one pay for that contract?*

Before Black and Sholes, such price was determined on the base of historical data. Black and Sholes discovered that the price should be based on a rational mechanism. They introduce $V = V(t,x)$, price of the contract starting at time t knowing that the value of the underlie is x, namely $S(t) = x$, and expiring at time T delivering $F(S(T))$. They assume a simple financial principle: *market is arbitrage free, that is there're no financial assets delivering certain profit or certain loss*. By this it follows that, for example, $V(T,x) = F(x)$. Indeed, if $V(T,x) < F(x)$, that is if the *value of a contract written at time T when $S(T) = x$ paying $F(S(T)) = F(x)$ at time T itself* **is less than** $F(x)$, then the buyer has a money machine: they pay $V(T,x)$ to receive $F(x)$ realizing a risk free profit $F(x) - V(T,x)$. Similarly, $V(T,x) > F(x)$ is impossible.

By similar arguments, Black and Sholes derived a condition on $V(t,x)$ at any time $t < T$. They proved that V must fulfil the following PDE,

$$\partial_t V(t,x) + \frac{1}{2}\sigma^2 x^2 \partial_{xx} V(t,x) + rx\partial_x V(t,x) - rV(t,x) = 0. \tag{21.4.1}$$

This, together with above mentioned condition $V(T,x) = F(x)$ leads to the following problem:

$$\begin{cases} \partial_t V(t,x) + \frac{1}{2}\sigma^2 x^2 \partial_{xx} V(t,x) + rx\partial_x V(t,x) - rV(t,x) = 0, & 0 \leqslant t \leqslant T,\ x \geqslant 0, \\ V(T,x) = F(x), & x \geqslant 0. \end{cases} \tag{21.4.2}$$

The problem (21.4.2) is apparently similar to the (21.3.1). However, it is not evident how to use the FT being the spatial domain x asymmetric. Setting $y = \log x$, that is $x = e^y$ and

$$u(t,y) := V(t,e^y),$$

easily we have

$$\begin{cases} \partial_t u(t,y) + \frac{1}{2}\sigma^2 \partial_{yy} u(t,y) + r\partial_y u(t,y) - ru(t,y) = 0, & 0 \leqslant t \leqslant T,\ y \in \mathbb{R}, \\ u(T,y) = F(e^y), & y \in \mathbb{R}. \end{cases} \tag{21.4.3}$$

We can now use the FT to solve this problem. Let

$$v(t,\xi) := \widehat{u(t,\cdot)}(\xi).$$

Then

$$\partial_t v + \frac{1}{2}\sigma^2(-i\xi)^2 v + r(-i\xi)v - rv = 0,$$

or

$$\partial_t v = \left(\frac{1}{2}\sigma^2\xi^2 + ir\xi + r\right) v.$$

This is an ordinary differential equation in $v(t,\xi)$ (ξ fixed). The final condition on $u(T,y)$ becomes

$$v(T,\xi) = \widehat{u(T,\cdot)}(\xi) = \widehat{F(e^\sharp)}(\xi).$$

Therefore

$$v(t,\xi) = e^{\left(\frac{1}{2}\sigma^2\xi^2 + ir\xi + r\right)(t-T)} v(T,\xi) = e^{-r(T-t)} e^{-ir(T-t)\xi} e^{-\frac{1}{2}\sigma^2(T-t)\xi^2} \widehat{F(e^\sharp)}(\xi).$$

We can now return to y. First recall that

$$e^{-\frac{1}{2}\sigma^2(T-t)\xi^2} = \widehat{\frac{1}{\sqrt{2\sigma^2(T-t)}} e^{-\frac{\sharp^2}{2\sigma^2\sqrt{T-t}}}}$$

thuse

$$e^{-\frac{1}{2}\sigma^2(T-t)\xi^2}\widehat{F(e^\sharp)} = \widehat{\frac{1}{\sqrt{2\pi\sigma^2(T-t)}} e^{-\frac{\sharp^2}{2\sigma^2\sqrt{T-t}}} * F(e^\sharp)}.$$

Moreover, the multiplication by $e^{-ir(T-t)\xi}$ in the FT means a translation in the variable of $-r(T-t)$ in its original. Putting together these facts,

$$\begin{aligned} u(t,y) &= e^{-r(T-t)} \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} e^{-\frac{\sharp^2}{2\sigma^2\sqrt{T-t}}} * F(e^\sharp)(y - r(T-t)) \\ &= e^{-r(T-t)} \int_{\mathbb{R}} F\left(e^{y-r(T-t)-\eta}\right) \frac{e^{-\frac{\eta^2}{2\sigma^2(T-t)}}}{\sqrt{2\pi\sigma^2(T-t)}}\, d\eta. \end{aligned}$$

Returning to V we finally obtain

$$\begin{aligned} V(t,x) &= u(t,\log x) = e^{-r(T-t)} \int_{\mathbb{R}} F(e^{\log x - r(T-t) - \eta}) \frac{e^{-\frac{\eta^2}{2\sigma^2(T-t)}}}{\sqrt{2\sigma^2(T-t)}}\, d\eta \\ (z := -\tfrac{\eta}{\sigma\sqrt{T-t}}) \quad &= e^{-r(T-t)} \int_{\mathbb{R}} F\left(xe^{-r(T-t)+(\sigma\sqrt{T-t})z}\right) \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\, dz \end{aligned} \tag{21.4.4}$$

This is the famous **Black formula**, still used to price contracts. For instance, the *call option* is a contract with payoff function $F(x) = \max\{K, x\}$. The price at $t = 0$ if $S(0) = x$ is

$$\begin{aligned}
V(0,x) &= e^{-rT} \int_{\mathbb{R}} \max\left\{ xe^{-rT+(\sigma\sqrt{T})z}, K \right\} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \, dz \\
&= e^{-rT} \left(\int_{-\infty}^{\frac{1}{\sigma\sqrt{T}} \log \frac{K}{x} + \frac{r}{\sigma}\sqrt{T}} K \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \, dz + \int_{\frac{1}{\sigma\sqrt{T}} \log \frac{K}{x} + \frac{r}{\sigma}\sqrt{T}}^{+\infty} xe^{-rT+\sigma\sqrt{T}z} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \, dz \right) \\
&= e^{-rT} \left(K\Phi\left(\frac{1}{\sigma\sqrt{T}} \log \frac{K}{x} + \frac{r}{\sigma}\sqrt{T} \right) + xe^{-rT} \int_{-\infty}^{-\frac{1}{\sigma\sqrt{T}} \log \frac{K}{x} - \frac{r}{\sigma}\sqrt{T}} e^{-\sigma\sqrt{T}z} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \, dz \right) \\
&= e^{-rT} \left(K\Phi\left(\frac{1}{\sigma\sqrt{T}} \log \frac{K}{x} + \frac{r}{\sigma}\sqrt{T} \right) + xe^{\left(\frac{\sigma^2}{2} - r\right)T} \Phi\left(-\frac{1}{\sigma\sqrt{T}} \log \frac{K}{x} - \frac{r}{\sigma}\sqrt{T} + \sigma\sqrt{T} \right) \right).
\end{aligned}$$

where we denoted by

$$\Phi(w) := \int_{-\infty}^{w} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \, dz$$

the *distribution function of the standard gaussian.*

21.5. Exercises

Exercise 21.5.1. *Solve the following equation in the unknown $u \in L^1$*

$$\int_{\mathbb{R}} u(x-y)e^{-|y|} \, dy = 2e^{-|x|} - e^{-2|x|}.$$

Exercise 21.5.2 ($\star$)**.** *Solve the following equation in the unknown $u \in L^1(\mathbb{R})$ such that $u', u'' \in L^1(\mathbb{R})$:*

$$u''(x) - \frac{1}{2} \int_{\mathbb{R}} e^{-|y|} u(x-y) \, dy = e^{-|x|} \text{sgn}(x). \quad (\star)$$

Exercise 21.5.3 ($\star$)**.** *Solve the following equation in two cases: i) $u \in L^1(\mathbb{R})$, ii) $u \in L^2(\mathbb{R})$:*

$$\int_{\mathbb{R}} u(y)u(x-y) \, dy + u(x) = \frac{1}{1+x^2}.$$

Exercise 21.5.4. *Consider the Cauchy problem for the wave equation on an infinite interval*

$$\begin{cases} \partial_{tt} u(t,x) = c^2 \partial_{xx} u(t,x), \ t \geqslant 0, \ x \in \mathbb{R}, \\ u(0,x) = \varphi(x), \ x \in \mathbb{R}, \\ \partial_t u(0,x) = \psi(x), \ x \in \mathbb{R}. \end{cases}$$

Setting $v(t,\xi) := \widehat{u(t,\sharp)}(\xi)$, determine $v(t,\xi)$. Deduce D'Alembert formula

$$u(t,x) = \frac{1}{2} \left(\varphi(x+t) + \varphi(x-t) \right) + \frac{1}{2} \int_{x-t}^{x+t} \psi(y) \, dy.$$

Exercise 21.5.5. Find the solution of the following problem

$$\begin{cases} \partial_{xx}u(t,x) = \partial_{tx}u(t,x), & x \in \mathbb{R},\ t > 0, \\ u(0,x) = e^{-|x|}, & x \in \mathbb{R}. \end{cases}$$

Exercise 21.5.6. Find the solution of the following problem

$$\begin{cases} \partial_t u(t,x) + t\partial_x u(t,x) = 0, & x \in \mathbb{R},\ t > 0, \\ u(0,x) = f(x), & x \in \mathbb{R}. \end{cases}$$

Exercise 21.5.7. Find the solution of the following problem

$$\begin{cases} \partial_t u(t,x) = e^{-t}\partial_{xx}u(t,x), & x \in \mathbb{R},\ t > 0, \\ u(0,x) = e^{-|x|}, & x \in \mathbb{R}. \end{cases}$$

Exercise 21.5.8. Find the solution of the following problem

$$\begin{cases} \partial_{tt}u(t,x) + \partial_{xxxx}u(t,x) = 0, & x \in \mathbb{R},\ t > 0, \\ u(0,x) = \mathrm{rect}1, & x \in \mathbb{R}, \\ u_t(0,x) = 0, & x \in \mathbb{R}. \end{cases}$$

Exercise 21.5.9. The model of heat diffusion with convection is described by the Cauchy problem

$$\begin{cases} \partial_{tt}u = c^2\partial_{xx}u + k\partial_x u, & t > 0,\ x \in \mathbb{R}, \\ u(0,x) = f(x), & x \in \mathbb{R}. \end{cases}$$

Find the evolution of the temperature in the case $c = 1$, $k = \frac{1}{2}$ and initial temperature $f(x) = e^{-x^2}$.