Advanced Real Analysis I, HT2025

Benjamin Andersson

August 2025

Contents

Chapter 1

Lecture 1, 26/8, 13:00

- Lecture 10:00-12:00 usually.
- Course has 2 parts:
  - Measure and Integration theory.
  - Functional analysis.

Examination: Written exam in two parts, part A and part B., both giving 16 points. Can get at most 4 bonus points, all counted (as far as we can tell) on part A.

There are 2 Homework assignments, with 4-5-6 problems, we grade one of the problems randomly (and everyone gets graded on the same problem).

Regarding `ChatGPT` - do not just plug in the question in `ChatGPT` and copy the answer. (In private conversation with Kristian: more ok to use mainly for say negative feedback on solutions, or it is atleast up to the individual student).

Exercise sessions on fridays.

P. Halmos "Measure Theory" is followed by the book (atleast when it comes to measure theory).

1.1 Some motivation

Example 1.1.1. Let $f_n(x) = \lim_{m\to\infty}(\cos(n!\pi x))^{2m}$. Then we have that

$$f_n(x) = \begin{cases} 1, & \text{if } n!x \in \mathbb{Z} \\ 0, & \text{if } n!x \notin \mathbb{Z} \end{cases}$$

since if $n!x \notin \mathbb{Z}$ then $|\cos(n!x\pi)| < 1$, then $f_n(x) = 0$ since the limit of a real number t strictly between -1 and 1 is 0.

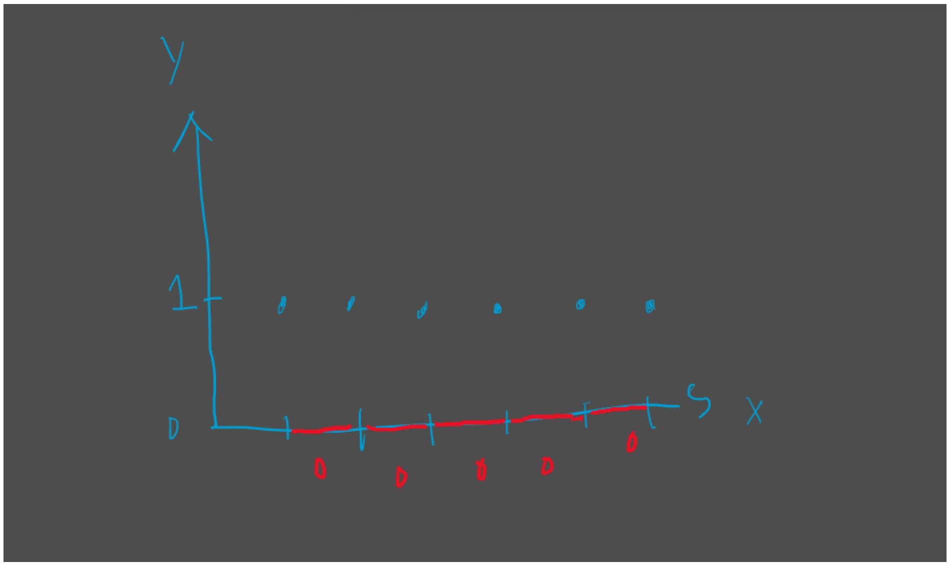

Figure 1.1: Picture 1.1 - sketch of function $f_n(x)$

We claim that $f_n(x)$ is Riemann-integrable over $[0, 1]$ and that

$$\int_0^1 f_n(x)\ dx = 0.$$

Define $f(x) := \lim_{n\to\infty} f_n(x)$. Then we claim that

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \\ 0, & \text{if } x \in \mathbb{Q} \end{cases}$$

f is *not* Riemann-integrable. We want to take $\lim_{n\to\infty} \int_0^1 f_n(x)\ dx$, but if we put the *limit* inside the integral, then the integral is *no longer defined*.

1.2 Lebesgues idea

"Def": Riemann-integrable. Think of $y = f(x)$ as some *continuous* function.

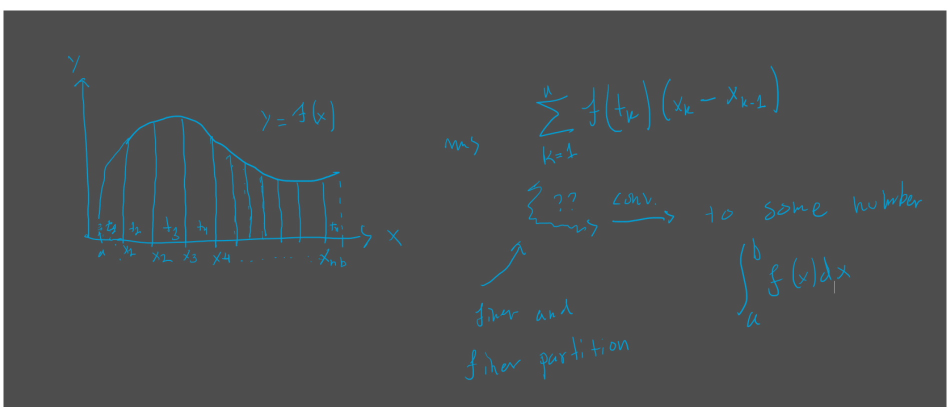

Figure 1.2: Picture 1.2 - Lebesgues idea

Lebesgues Idea: Let's make a "partition of the y-axis" (cf. picture (1.2).

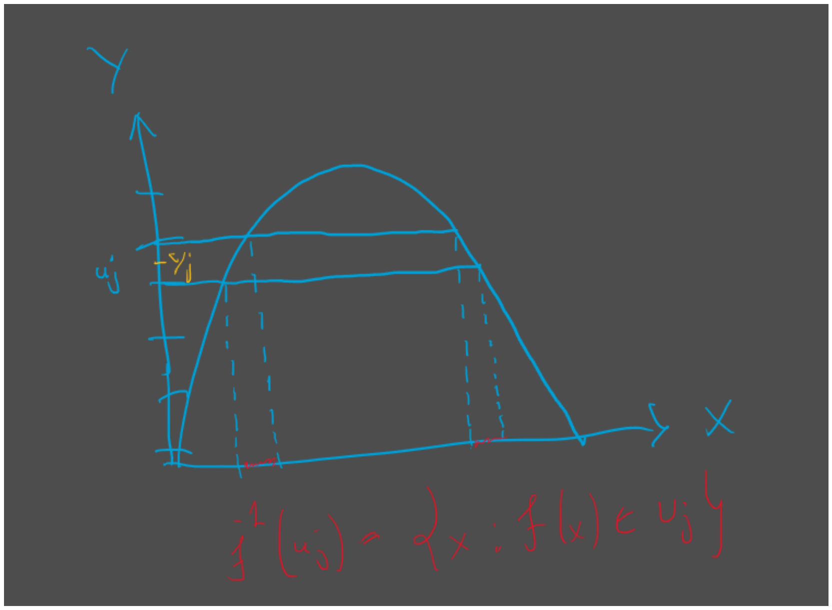

Figure 1.3: Picture 1.3

From picture 1.3 $\rightsquigarrow \sum \underbrace{y_j}_{\text{"height"}} \underbrace{\mu(f^{-1}(u_j))}_{\text{"base"}}$. We have that $\mu_j(\cdot) =$"Length" of U_j, which we need to *define*,

1.3 What would we like from μ

We would like:

- $\mu([c, d]) = d - c$.
- $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ (note that the codomain is here considered as a subset of the *extended* real numbers), where $\mathcal{P}(A)$ is the power-set of a set A (in our case $A = \mathbb{R}$).
- $\mu(U + a) = \mu(U)$ for any subset U of $\mathbb{R}$ and any real number a.
- $\mu\left(\bigcup_{i=1}^{\infty} U_j\right) = \sum_{j=1}^{\infty} \mu(U_j)$ if $U_i \cap U_j = \varnothing$ and $i \neq j$ - "Lengths of total should be the som of them". This last condition on μ is called **completely additive**.

In fact, the above conditions on μ would lead to a contradiction (see [Fri03, §1.6]). One can construct sets of both measure zero and one, if one assumes the Axiom of Choice.

The key, for our purposes, is putting restrictions on the condition that $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$. We need to define/stipulate what our *measurable sets* i.e. we want to find "Good sets for μ".

1.4 Rings and algebra

Let $\underline{\overline{X}}$ be a non-empty set. We may thinl of $\underline{\overline{X}}$ as our "space". We call $x \in \underline{\overline{X}}$ **points**.

Remark 1.4.1. [Remark on notation]

- By $\subset$ as in $E \subset \underline{\overline{X}}$ we really mean $E \subseteq X$, i.e. $E \subset \underline{\overline{X}}$ does not necessarily imply that E is a *proper* subset of $\underline{\overline{X}}$, in this course (and [Fri03]). E.g. $\underline{\overline{X}} \subset \underline{\overline{X}}$ is valid.
- E^c - complement.
- $E \cup F$ - union.
- $\underbrace{E - F}_{\text{[Fri03]}} = \underbrace{E \setminus F}_{\text{Kristian}}$ - difference.
- $E \cap F$ - intersection.

1.5 Limit supremum and limit inferior

Limit supremum

Definition 1.5.1. Let (E_n) be a (countable) sequence of subsets of a space $\underline{\overline{X}}$. Then $\overline{\lim}_n E_n$ is the set of all points which belongs to *infinitely many* E_n. It is called **limit supremum** of (E_n).

Limit infimum

Definition 1.5.2. Let (E_n) be a (countable) sequence of subsets of a space $\underline{\overline{X}}$. Then $\underline{\lim} E_n$ is the set of all points which belongs to *all except at most finitely many* E_n. It is called **limit infimum** of (E_n).

Example 1.5.3. Let $\underline{\overline{X}} = \mathbb{Z}$, and let $E_n = \{(-1)^n\}$. Then $\overline{\lim}_n \; E_n = \{1, -1\}$ and $\underline{\lim}_n \; E_n = \varnothing$.

Since in any sequence (E_n) there are infinitely many E_n, it follows that $\underline{\lim}_n E_n \subset \overline{\lim}_n \; E_n$.

Ring and algebra

Definition 1.5.4. A collection $\mathcal{R}$ of subsets of a space $\underline{\overline{X}}$ is an **Ring** if:
1) $\varnothing \in \mathcal{R}$.
2) $A, B \in \mathcal{R} \rightsquigarrow A \setminus B \in \mathcal{R}$ for all $A, B \in \mathcal{R}$.
3) $A, B \in \mathcal{R} \rightsquigarrow A \cup B \in \mathcal{R}$ for all $A, B \in \mathcal{R}$.

If we add another assumption,
4) $\underline{\overline{X}} \in \mathcal{R}$,

then we call $\mathcal{R}$ an **algebra**.

Remark 1.5.5. If $A \in \mathcal{R} \rightsquigarrow A^c \in \mathcal{R}$ since $A^c = \underline{\overline{X}} \setminus A$ and $\underline{\overline{X}}, A \in \mathcal{R} \overset{2)}{\rightsquigarrow} \underline{\overline{X}} \setminus A = A^c \in \mathcal{R}$.

- If $A, B \in \mathcal{R} \rightsquigarrow A \cap B \in \mathcal{R}$ since $A \cap B = \underbrace{A}_{\in \mathcal{R}} \setminus \underbrace{(A \setminus B)}_{\in \mathcal{R}}$.

Remark 1.5.6. There is a typo with regard to the above point in [Fri03, approximately p.2]

We prove that $A \cap B = A \setminus (A \setminus B)$.

Proof. First let $x \in A \cap B$. Then $x \in A$, and x is not in $A \setminus B$, since $A \setminus B$ are elements in A not in B. Hence $x \in A \setminus (A \setminus B)$, since $A \setminus (A \setminus B)$ are elements in A not in $A \setminus B$.

On the other hand, let $x \in A \setminus (A \setminus B)$. Then x is in A, but is not in $A \setminus B$. But $A \setminus B$ are elements in A not in B, so x being in A but not in $A \setminus B$ means that x must be in B, so $x \in A \cap B$. Another way to formulate this last point is that if x was in A and x was not in B, then x would be in $A \setminus B$. The contrapositive of this is that if x is not in $A \setminus B$ then x is in B. This shows both inclusions and so indeed we have that $A \cap B = A \setminus (A \setminus B)$. ☐

1.6 σ-algebras

Slogan: "Whatever you do, you always want to take limits".

σ-algebra

A collection $\mathcal{A}$ of subsets of $\underline{X}$ is a **σ-algebra** if:
1) $\underline{X} \in \mathcal{A}$.
2) $A, B \in \mathcal{A} \rightsquigarrow A \setminus B \in \mathcal{A}$.
3) If A_n for $n = 1, 2, \ldots$ is in $\mathcal{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ (*closed under countable unions*).

Remark 1.6.1. Notice that 3) is a difference between an *algebra* and a σ-algebra.

Remark 1.6.2. Note that since $\varnothing = \underline{X}^c$ and $\underline{X}^c = \underline{X} \setminus \underline{X}$, and $\underline{X} \in \mathcal{A}$ by 1), then by 2) we have that $\underline{X} \setminus \underline{X} = \varnothing \in \mathcal{A}$.

- Given some sets $A_n \in \mathcal{A}$ for $n = 1, \ldots, m$ (where m is a positve natural number), is $\bigcup_{n=1}^{m} A_n \in \mathcal{A}$? Yes! If we let $A_n = \varnothing$ for $n = m+1, m+2, \ldots$ (and by the previous remark these are all in $\mathcal{A}$), then we have that $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{m} A_n \in \mathcal{A}$.

Example 1.6.3. [Example of σ-algebras]

- $\mathcal{P}(\underline{X})$ is a σ-algebra for a space $\underline{X}$ ("largest"; can for example be "events" in probability theory).
- $\{\underline{X}, \varnothing\}$ is a σ-algebra ("smallest").

Exercise: If $\mathcal{A}$ is a σ-algebra and $A_n \in \mathcal{A}$ for $n = 1, 2, \ldots$ then

$$\bigcap_{i=1}^{\infty} A_n \in \mathcal{A}.$$

We claim that if one shows the above statements, then as in the case of going from an infinite but countable union to a finite union, one may show that also $\bigcap_{n=1}^{m} A_n \in \mathcal{A}$ for all $m \geqslant 1$ and $A_n \in \mathcal{A}$.

We want to be able to measure the lengths of the $A_n \rightsquigarrow$ Definition of measure [Fri03, § 1.2].

1.7 Extended real numbers, measures,properties

$-\infty < x < \infty$ real number whenever $x \in \mathbb{R}$ (circular?). But we may extend $\mathbb{R}$ by $\{\pm\infty\}$ so that we get the **extended real numbers**, which some laws in place, .e.g. $x + \infty = \infty$, $x \cdot \infty = \infty$, if $x > 0$ and $-\infty$ if $x < 0$. Note however that $-\infty + \infty$ is not defined (notice that this means that $\mathbb{R} \cup \{\pm\infty\}$ is not a field/ring).

Also $0 \cdot (\pm\infty)$ is undefined.

Central concept: Measure

Definition 1.7.1. Let $\mathcal{A}$ be a σ-algebra. A function $\mu : \sigma \to [0, \infty]$, Set $\mapsto$ Measure is a (positive) **measure** if:

1. $\mu(\varnothing) = 0$.
2. μ is *completely additive*, i.e.

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$$

if all the E_n:s are measurable, $E_n \in \mathcal{A}$ for all $n = 1, 2, \dots$, and $E_i \cap E_j = \varnothing$ if $i \neq j$.

Properties:

Proposition 1.7.2. *If $E, F \in \mathcal{A}$ ($\mathcal{A}$ a σ-algebra) and $E \subset F$, then $\mu(E) \leqslant \mu(F)$.*

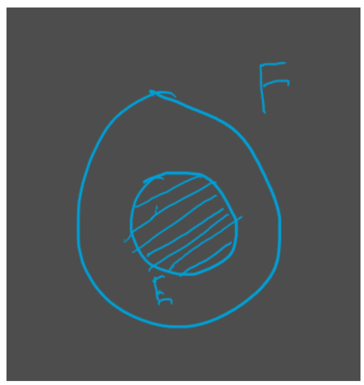

Figure 1.4: $E \subset F$

Proof.

Lemma 1.7.3. *If $E \subset F$ are sets, then $F = E \cup (F \setminus E)$.*

Proof. Let $x \in F$. Then since $E \subset F$ we have that either x is in E, or x is an element of F not in E. But this is precisely $E \cup (F \setminus E)$. On the other hand, if $x \in E \cup (F \setminus E)$, then either x is in E, in which case $x \in F$ since $E \subset F$, or x is in $F \setminus E \subset F$ so that $x \in F$. It follows that $F = E \cup (F \setminus E)$. □

Since E and $(F \setminus E)$ are disjoint, and both E and $F \setminus E$ are in $\mathcal{A}$, it follows that

$$\begin{aligned}\mu(F) &= \mu\left(E \cup (F \setminus E)\right) \\ &= \mu(E) + \mu(F \setminus E)\end{aligned}$$

and since the codomain of $\mu(\cdot)$ consists of *non-negative* real numbers together with ∞, it follows that $\mu(E) \leqslant \mu(E) + \mu(F \setminus E) = \mu(F)$, which is what we wanted to show. □

Example 1.7.4. Let $\mathcal{A} = \mathcal{P}(\underline{\overline{X}})$, the *power-set* of a space $\underline{\overline{X}}$, and take any point $x \in \underline{\overline{X}}$. Let

$$\mu(A) = \begin{cases} 0, & \text{if } x \notin A, \\ 1, & \text{if } x \in A \end{cases}$$

Is μ a measure?

Sol: $\mathcal{A}$ is a sigma algebra, and we have that:

1) $\mu(\varnothing) = 0$ since for fixed x, we have that $x \notin \varnothing$.
2) Assume that (E_n) are pair-wise disjoint elements in $\mathcal{A}$. Then x can be in *at most one* E_k. We treat the cases separately. x not in any E_k: If we have that $x \notin \bigcup_{n=1}^{\infty} E_n$ then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = 0$$

and

$$\sum_{n=1}^{\infty} \mu(E_n) = 0$$

since $x \notin E_n$ for $n \geqslant 1$. x is in exactly on E_k: On the other hand, if $x \in E_k$ for some $k \geqslant 1$ then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = 1$$

and

$$\begin{aligned}\sum_{n=1}^{\infty} \mu(E_n) &= 0 + 0 + \ldots + 1 + 0 + 0 + \ldots \\ &= 1.\end{aligned}$$

We conclude that μ is *completely additive*, so that μ is a measure (1.7.1).

Remark 1.7.5. The above μ is called the **Dirac Measure**, and is usually denoted by δ_x.

Next time $\rightsquigarrow$ Outer measures, Lebesgue measures.

Chapter 2

Lecture 2

<u>Last time:</u>

- Space $\underline{\overline{X}} \neq \varnothing$.
- $\mathcal{A}$ a σ-algebra of subsets of $\underline{\overline{X}}$.
- μ a measure defined on $\mathcal{A}$.

Measure space $(\underline{\overline{X}}, \mathcal{A}, \mu)$

Definition 2.0.1. The triple $(\underline{\overline{X}}, \mathcal{A}, \mu)$ is a **measure space**.
.

<u>Question</u>: What is the "length" of $\mathbb{Q} \subset \mathbb{R}$. We may think of $\mathbb{Q}$ as "dust in $\mathbb{R}$".

- $\mathbb{Q}$ is *dense* in $\mathbb{R}$.
- $\mathbb{Q}$ are *countable*.

Remark 2.0.2. The above question motivates the *Lebesgue measure*.

Let $(r_n)_n$ be an enumeration of $\mathbb{Q}$ (since $\mathbb{Q}$ is countable this is perfectly fine).

Take $\varepsilon > 0$, and cover each r_n by an open interval I_n of length $\frac{\varepsilon}{2^n}$. Then

$$\begin{aligned}
\underbrace{\lambda(\mathbb{Q})}_{\text{“length”}} &\leqslant \lambda\left(\bigcup_{n=1}^{\infty} I_n\right) \\
&\leqslant_{\text{“subadditivity”}} \sum_{n=1}^{\infty} \lambda(I_n) \\
&\leqslant \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \\
&= \varepsilon
\end{aligned}$$

where we used that the partial sums

$$\begin{aligned} S'_m &= \sum_{n=1}^{m} \frac{\varepsilon}{2^m} \\ &= \varepsilon \sum_{n=1}^{m} \frac{1}{2^n} \\ &= \varepsilon \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \\ &= \varepsilon \cdot \left(1 - \left(\frac{1}{2}\right)^m\right). \end{aligned}$$

Since the limit of the partial sums is the limit of the series, and we may think of this as $\lim_{n\to\infty}(a_n b_n)$ with $a_n = \varepsilon$ for each $n \geqslant 1$ and $b_n = \left(\frac{1}{2}\right)^n$ for $n \geqslant 1$, we see that by the *product-law* for limits and $\frac{1}{2^n} \to 0$ (and using the summation-law for limits) we see that this

$$\sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we should have that $\lambda(\mathbb{Q}) = 0$, i.e. that the "length" of $\mathbb{Q}$ is zero.

2.1 Outer measure ([Fri03, §1.3]

Outer measure, μ^*

Definition 2.1.1. The function $\mu^* : \mathcal{P}(\underline{\overline{X}}) \to [0, \infty]$ is an **outer measure** if:

1) $\mu^*(\varnothing) = 0$.
2) μ^* is *monotone*, i.e. if $E \subset F \subset \mathcal{P}(\underline{\overline{X}})$ then $\mu^*(E) \leqslant \mu^*(F)$.
3) μ^* is *countably subadditive*, i.e.

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(E_n)$$

for all $E_n \subset \underline{\overline{X}}$.

One can use outer measures to define *measures*, be restricting the domain (i.e. with our notation, restricting the domain of μ^*).

Measurable sets

Definition 2.1.2. Let μ^* be an outer measure (2.1.1). Then, we say that a set $E \subset \underline{\overline{X}}$ (or $E \subset \mathcal{P}(\underline{\overline{X}})$) is **measurable** if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E)$$

for all subsets $A \subset \underline{\overline{X}}$.

Remark 2.1.3. Note that $A = (A \cap E) \cup)(A \setminus E)$.

Since μ^* is *subadditive* (see 2.1.1), $A = (A \cap E) \cup (A \setminus E)$ we have that

$$\begin{aligned}\mu^*(A) &= \mu^*((A \cap E) \cup (A \setminus E)) \\ &\leqslant \mu^*(A \cap E) + \mu^*(A \setminus E).\end{aligned}$$

Hence it is enough to prove that $\mu^*(A) \geqslant \mu^*(A \cap E) + \mu^*(A \setminus E)$ holds for *every* set $A \subset \underline{X}$ (or perhaps $\mathcal{P}(\underline{X})$) to conclude that

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E).$$

Theorem 2.1.4 (Carathéodory)**.** *Let μ^* be an outer measure. Let $\mathcal{A}$ denote the (μ^*)-measurable sets. Then $\mathcal{A}$ is a σ-algebra and if we take $\mu := \mu^*|_{\mathcal{A}}$ then μ is a measure.*

Proof. See [Fri03, Theorem 1.3.1]. We will only prove that if $\mu^*(E) = 0$ for some subset E then E is measurable. Take any $A \subset \underline{X}$. Then we have that

$$\begin{aligned}\mu^*(\underbrace{A \cap E}_{\subset E}) + \mu^*(\underbrace{A \setminus E}_{\subset A}) &\leqslant \underbrace{\mu^*(E)}_{=0} + \mu^*(A) \\ &= \mu^*(A),\end{aligned}$$

where we used the *monotonicity* of μ^* and that μ^* only give us back *non-negative* extended real numbers, so that since $\mu^*(A \cap E) \leqslant \mu^*(E) = 0$ implies that $\mu^*(A \cap E) = 0$. □

2.2 Construction of outer measure ([Fri03, §1.4])

Sequential covering class $\mathscr{K}$ (s.c.c)

Definition 2.2.1. Let $\mathscr{K}$ be a collection of subsets of $\underline{X}$ (i.e. $\mathscr{K} \subset \mathcal{P}(\underline{X})$). Then $\mathscr{K}$ is a **sequential covering class** (at $\underline{X}$) if:

1) $\varnothing \in \mathscr{K}$.
2) For every $A \subset \underline{X}$ there exists a sequence (E_n) in $\mathscr{K}$ (i.e. $E_n \in \mathscr{K}$ for all $n \geqslant 1$) such that
$$A \subset \bigcup_{n=1}^{\infty} E_n,$$
i.e. A can be *covered* by (E_n).

Example 2.2.2. Let $\underline{X} = \mathbb{R}$. Then the collection of all *open and bounded intervals*, together with $\varnothing$ is a sequential covering class.

Let $\mathscr{K}$ be a sequential covering class (2.2.1). Let $\lambda : \mathscr{K} \to [0, \infty]$ be such that $\lambda(\varnothing) = 0$. Then for any subset $A \subset \underline{X}$, let

$$\lambda^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \lambda(E_k) \ : \ E_k \in \mathscr{K} \text{ and } \bigcup_{n=1}^{\infty} E_n \supset A \right\}. \tag{2.2.1}$$

Theorem 2.2.3 ([Fri03, Theorem 1.4.1])**.** *$\lambda^*(\cdot)$ in definition 2.2.1 is an outer measure.*

Proof. See book (in particular, *subadditivity*) □

2.3 Completion of measures ([Fri03, §1.5]

The role played by *sets of measure zero*, is "negligible" (In probability theory, such sets have probability zero).

Let $(\underline{X}, \mathcal{A}, \mu)$ be a measure space (2.0.1). We could have that there exists some set $E \in \mathcal{A}$, where $\mu(E) = 0$, i.e. have "measure zero", and $N \subset E$ such that $N \in \mathcal{A}$.

Complete measure

Definition 2.3.1. A measure (1.7.1) μ on a measure space say $(\underline{X}, \mathcal{A}, \mu)$ is **complete** if whenever $E \in \mathcal{A}$, $\mu(E) = 0$ and $N \subset E$ implies that $N \in \mathcal{A}$.

Remark 2.3.2. Note that by *monotonicity* of μ and the codomain of μ being assumed to be $[0, \infty]$, it follows from $\mu(E) = 0$ and $N \subset E$ that $\mu(N) = 0$.

It is always possible to extend a measure to a complete measure:

Theorem 2.3.3 ([Fri03, Theorem 1.5.1])**.** *Let $\overline{\mathcal{A}}$ be the collection of sets of the form $E \cup N$ where $A \in \mathcal{A}$ and N is a any subset of a set of $\mathcal{A}$ of measure zero. Then $\overline{\mathcal{A}}$ is a σ-algebra and $\overline{\mu}$ defined by*

$$\overline{\mu}(E \cup N) = \mu(E)$$

is a measure.

Proof. See book. □

Remark 2.3.4. We call $\overline{\mu}$ the **completion of** μ.

Remark 2.3.5. Note that the measure constructed in theorem (2.1.4) is complete (2.3.1): If $\mu^*(E) = 0$ and $N \subset E$ then $\mu^*(N) \leqslant \mu^*(E) = 0$ by monotonicity, which implies that $\mu^*(N) = 0$. By the construction in the proof ([Fri03, Theorem 1.3.1]) this implies that N is measurable.

2.4 The Lebesgue Measure ([Fri03, §1.6]

The Lebesgue *outer* measure

Definition 2.4.1. Let $\mathscr{K}$ be the sequential covering class for $\mathbb{R}$ consisting of all the bounded an open intervals (a, b) with $a < b$, $a, b \in \mathbb{R}$ together with $\varnothing$ (c.f. example 2.2.2).
Define $\lambda : \mathscr{K} \to [0, \infty)$ by $\lambda((a, b)) = b - a$ and $\lambda(\varnothing) = 0$. Then the **Lebesgue *outer* measure** is the measure in definition 2.2.1.

The Lebesgue measure; Lebesgue measurable sets

Definition 2.4.2. The *complete* measure λ determined through theorem 2.1.4 by $\lambda^*(\cdot)$ is the **Lebesgue measure** on $\mathbb{R}$. The *measurable sets* are the **Lebesgue measurable sets**.

For a proof that intervals are Lebesgue measurable (2.4.2), see Canvas page, Lecture 1-7, Lecture 2, "a note on intervals"

Remark 2.4.3. By [Fri03, Exc. 1.6.3-1.6.4] the Lebesgue outer measure (2.4.1) of an interval is the *length* of the interval.

Chapter 3

Lecture 3

Remark 3.0.1. Thanks to `ccczm` (discord) for providing the hand-written notes.

Recall: Let $M \subset \mathbb{R}$ be a non-empty open subset. Then $M = \bigcup_{n=1}^{\infty} I_n$ where the I_n are pairwise disjoint (open subsets are countable unions of pairwise disjoint intervals).

Measurable real-valued functions

Definition 3.0.2. Let $(\underline{X}, \mathcal{A}, \mu)$ be a measure space and take $\underline{X}_0 \in \mathcal{A}$. Then $F : \underline{X}_0 \to \mathbb{R}$ is said to be **measurable** if $F^{-1}(M) \in \mathcal{A}$ for all open $M \subset \mathbb{R}$.

Measurable extended real-valued functions

Definition 3.0.3. If $(\underline{X}, \mathcal{A}, \mu)$ is a measure space and $\underline{X}_0 \in \mathcal{A}$. Then we say that an extended real-valued function $F : \underline{X}_0 \to [-\infty, \infty]$ is **measurable** if:

1. For any open $M \subset \mathbb{R}$ we have $F^{-1}(M) \in \mathcal{A}$.
2. $F^{-1}(\pm\infty) \in \mathcal{A}$.

Theorem 3.0.4. *$F : \underline{X}_0 \to \mathbb{R}$ is measurable $\Leftrightarrow F^{-1}((-\infty, c)) \in \mathcal{A}$ for all $c \in \mathbb{R}$.*

Proof. $\underline{\Leftarrow}$: Notice that
$$F^{-1}((-\infty, c)) = \{x \in \underline{X}_0 \ : \ F(x) < c\}.$$
We may write
$$\{x \in \underline{X}_0 \ : \ F(x) \leqslant c\} = \bigcap_{n=1}^{\infty} \left\{x \in \underline{X}_0 \ : \ F(x) < c + \frac{1}{n}\right\}.$$
Since the right-hand side is a countable intersection of measurable sets (measurable by assumption) it follows that its intersection is measurable. Furthermore, we have that
$$F^{-1}((a, b)) = \{x \in \underline{X}_0 \ : \ F(x) < b\} \setminus \{x \in \underline{X}_0 \ : \ F(x) \leqslant a\}$$
and the (relative) complement of two measurable sets is measurable, hence this is measurable.

Furthermore, for any $M \subset \mathbb{R}$ open we have that $M = \bigcup_{n=1}^{\infty} I_n$ and preimages distribute over arbitrary unions, hence
$$F^{-1}(M) = \bigcup_{n=1}^{\infty} F^{-1}(I_n),$$

□

which is a countable union of measurable sets, hence measurable.

$\underline{\Rightarrow}$: Since $(-\infty, c)$ is open in $\mathbb{R}$, this follows by definition 3.0.2.

Proposition 3.0.5. *$F : \overline{\underline{X}}_0 \to \mathbb{R}$ is measurable $\Leftrightarrow F^{-1}((-\infty, c])$ is measurable.*

Proof. $\underline{\Leftarrow}$:

$$F^{-1}((-\infty, c)) = \bigcup_{n=1}^{\infty} \underbrace{F^{-1}\left(\left(-\infty, c - \frac{1}{n}\right]\right)}_{\text{measurable}}.$$

$\underline{\Rightarrow}$: This follows from $(-\infty, c] = \bigcap_{n=1}^{\infty} \left(-\infty, c + \frac{1}{n}\right)$, that preimages distribute over arbitrary intersections, that since F is measurable, the preimage of each $\left(-\infty, c + \frac{1}{n}\right)$ is measurable, and that countable intersections of measurable sets are measurable. □

Corollary 3.0.6. *$F : \overline{\underline{X}}_0 \to \mathbb{R}$ is measurable $\Leftrightarrow F^{-1}((c, \infty))$ is measurable for all $c \in \mathbb{R}$.*

Proof. $\underline{\Leftarrow}$: Notice that

$$\begin{aligned} F^{-1}((-\infty, c]) &= F^{-1}(\mathbb{R} \setminus (c, \infty)) \\ &= F^{-1}(\mathbb{R}) \setminus F^{-1}((-\infty, c)) \\ &= \overline{\underline{X}}_0 \setminus F^{-1}((-\infty, c)) \end{aligned}$$

is the (relative) complement of two elements in $\mathcal{A}$, hence is measurable.

$\underline{\Rightarrow}$: $(c, \infty) = \bigcup_{n=1}^{\infty} (c, n)$ is a union of open subsets, hence open, hence its preimage is measurable. □

Example 3.0.7. Let

$$F^+(x) = \begin{cases} F(x), & F(x) > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Theorem 3.0.8. *If F is measurable, then $F^+(x)$ is measurable.*

Proof. Note that

$$\begin{aligned} (F^+)^{-1}((-\infty, c)) &= \{x \in \overline{\underline{X}}_0 \ : \ F^+(x) < c\} \\ &= \begin{cases} \varnothing, \ \text{if } c \leqslant 0 \\ F^{-1}(-\infty, c), \ \text{if } c > 0 \end{cases}. \end{aligned}$$

This follows from the fact that

$$(F^+)^{-1}((-\infty, c)) = \{F \leq 0\} \cup \{0 < F < c\}$$

in the latter case, where $F \leqslant 0$ are all $x \in \overline{\underline{X}}_0$ such that $F^+(x) = 0$). The former is easier to see; if $c \leqslant 0$ then we would like to find $x \in \overline{\underline{X}}_0$ such that $F^+(x) < c$, which is impossible, since $F^+(x) \geqslant 0$).

But both $\varnothing$ and $F^{-1}(-\infty, c)$ are in $\mathcal{A}$ (the latter by assumpton on the measurability of F), hence by theorem 3.0.4 F is measurable. □

Example 3.0.9. Similar to F^+ we may define

$$F^-(x) = \begin{cases} -F(x), & \text{if } F(x) < 0 \\ 0, & \text{otherwise} \end{cases} .$$

Lemma 3.0.10. *If f, g are measurable then $\{x : f < g\}$ is a measurable set.*

Proof.

$$\{x : f < g\} = \bigcup_{n=1}^{\infty} (\{x : f < r_n\} \cap \{x : r_n < g\})$$

and the right-hand side is measurable. ☐

Theorem 3.0.11. *If f, g are measurable then $f \pm g$ and fg are measurable.*

Proof.

$$\{x : (f+g)(x) < c\} = \{x : f(x) < c - g(x)\}.$$

☐

3.1 Measurability is closed under taking limits

$(\sup_n F_n)(x)$ and $(\overline{\lim}_n F_n)$

Let (f_n) be a sequence of functions. Then we define

$$(\sup_n F_n)(x) := \sup_n F_n(x)$$

and

$$(\overline{\lim}_n F_n)(x) = \overline{\lim}_n F_n(x).$$

Remark 3.1.1. Defining $(\inf_n F_n)(x)$ and $(\underline{\lim}_n F_n)(x)$ is similar.

Theorem 3.1.2. *Given a sequence (F_n) of measurable functions, we have that $\overline{\lim}_n F_n, \underline{\lim}_n F_n, \sup F_n, \inf F_n$ are all measurable.*

Proof. Note: $\{x : (\sup_n F_n)(x) \leqslant c\} = \bigcap_{n=1}^{\infty} \{x : F_n(x) \leqslant c\}$ and the right-hand side is a countable intersection of measurable sets $(F_n^{-1}((-\infty, c])$ which are measurable by proposition 3.0.5, hence measurable. Hence the left-hand side is measurable, but the left-hand side is $(\sup_n F_n)^{-1}((-\infty, c])$ and c was arbitrary, so by applying 3.0.5 again we see that $(\sup_n F_n)(x)$ is measurable.

Note that $(\inf_n F_n(x) = -(\sup_n -F_n)(x)$ and

$$\begin{cases} \overline{\lim}_n F_n = \inf_n \sup_{k \geqslant n} F_k \\ \underline{\lim}_n F_n = \sup_n \inf_{k \geqslant n} F_k. \end{cases}$$

☐

Almost everywhere

Definition 3.1.3. A property $P : \underline{\overline{X}} \to \{\mathsf{true}, \mathsf{false}\}$ is said to hold **almost everywhere** (a.e.) if $\mu(P^{-1}(\{\mathsf{false}\}) = 0$ whenever μ is *complete*. If the underlying measure space $(X, \mathcal{A}, \mu)$ is *not complete*, then P is said to hold **almost everywhere** if there exists an $N \in \mathcal{A}$ such that $P^{-1}(\{\mathsf{false}\}) \subset N$ and $\mu(N) = 0$.

Example 3.1.4. Let $F, g : \underline{\overline{X}}_0 \rightrightarrows \mathbb{R}$ and let F be measurable and $F = g$ a.e., and μ is complete, then g is measurable.

Proof. Consider $B = \{x \ : \ g(x) \leqslant c\}$ and $A = \{x \ : \ F(x) \leqslant c\}$.

Then $B \setminus A = \{x \ : \ g(x) \leqslant c \text{ and } F(x) > c\}$ while $A \setminus B = \{x \ : \ F(x) \leqslant c \text{ and } g(x) > c\}$. Since μ is complete, the set $E = \{x \ : \ F(x) \neq g(x)\}$ is such that $\mu(E) = 0$ (so measurable). Furthermore, we notice that $B \setminus A \subset E$ and $A \setminus B \subset E$. By completeness we have that $B \setminus A$ and $A \setminus B$ are measurable. Since F is measurable, A is measurable by proposition 3.0.5.

We have that

$$\begin{aligned}
(B \setminus A) \cup (A \setminus (A \setminus B)) &= (B \cap A^c) \cup (A \cap (A \cap B^c)^c) \\
&= (B \cap A^c) \cup (A \cap (A^c \cup B)) \\
&= (B \cap A^c) \cup ((A \cap A^c) \cup (A \cap B)) \\
&= (B \cap A^c) \cup (B \cap A) \\
&= B \cap (A^c \cup A) \\
&= B \cap \underline{\overline{X}}_0 \\
&= B.
\end{aligned}$$

Since $A, A \setminus B$ are measurable, $A \setminus (A \setminus B)$ is measurable. Hence $(B \setminus A) \cup (A \setminus (A \setminus B)) = B$ is measurable. Since c was arbitrary, it follows by proposition 3.0.5 that g is measurable. ☐

Pointwise convergent almost everywhere

Definition 3.1.5. A sequence (F_n) is **pointwise convergent almost everywhere** if there exists a function g such that $\lim_n (F_n) = g$ holds almost everywhere.

Theorem 3.1.6. *If a sequence (F_n) of measurable functions converges (pointwise) everywhere to g, then g is measurable. If a sequence (F_n) of measurable functions converges almost everywhere to g, and μ is complete, then g is measurable.*

Remark 3.1.7. Note the weaker requirement in the case when μ is complete; i.e. we *only* require that the sequence $(F_n) \to g$ *almost* everywhere, to be able to say that g is measurable.

Proof. $\overline{\lim}_n F_n = \tilde{g}$ is measurable. But if (F_n) goes to g everywhere, then $\tilde{g} = g$ and hence g is measurable.

On the other hand, if (F_n) converges to g almost everywhere, this means that $\tilde{g} = g$ almost everywhere, i.e. $E = \{x \ : \ \tilde{g}(x) \neq g(x)\}$ is a measurable set with measure zero.

Lemma 3.1.8. *If $(F_n) \to g$ almost everywhere, and F_n are measurable functions and μ is complete, then g is measurable.*

Proof. Let $E = \{x \ : \ f(x) \neq g(x)\}$, with $\mu(E) = 0$. Fix $x \in \mathbb{R}$ and put $A = f^{-1}((-\infty, c])$ and $B = g^{-1}((-\infty, c])$. If $x \in E^c$ then $x \in A \Leftrightarrow x \in B$. Thence $A \cap E^c = B \cap E^c$. Therefore, we have

$$\begin{aligned}(A \cap E^c) \cup (B \cap E) &= (B \cap E^c) \cup (B \cap E) \\ &= B \cap (E \cup E^c) \\ &= B \cap \underline{\overline{X}} \\ &= B.\end{aligned}$$

But A is measurable by assumption since f is measurable, and E^c is measurable since E is measurable, hence $A \cap E^c$ is measurable. Since $B \cap E \subset E$ and E is measurable, it follows by completeness of μ that $B \cap E$ is measurable. Hence since B is a union of measurable sets, it is measurable. Since c was arbitrary, it follows by proposition 3.0.5 that g is measurable. ☐

By applying the above reasoning to $\tilde{g}$ and g, we see that g must be measurable, since $\tilde{g} = g$ almost everywhere, and $\tilde{g}$ is measurable by 3.1.2. ☐

Simple functions

Definition 3.1.9. A function $F : \underline{\overline{X}} \to \mathbb{R}$ is **simple** if there exists *mutually disjoint measurable sets* $E_1, E_2, \ldots, E_n$ and real numbers $\alpha_1, \ldots, \alpha_n$ such that

$$F(x) = \sum_{k=1}^{n} \alpha_k \chi_{E_k}(x)$$

where

$$\chi_{E_k}(x) = \begin{cases} 1, & \text{if } x \in E_k \\ 0, & \text{otherwise} \end{cases} .$$

Theorem 3.1.10. *Let $F : \underline{\overline{X}} \to [0, \infty]$ be non-negative and measurable. Then there exists a monotone increasing sequence (F_n) of simple (definition 3.1.9) non-negative functions that converges to F everywhere.*

Remark 3.1.11. As in the proof (see [Fri03, proof of Theorem 2.2.5]) one may formulate *monotone* as: For every $n \geqslant 1$ and every $x \in \underline{\overline{X}}$, it holds that $f_{n+1}(x) \geqslant f_n(x)$.

Chapter 4

Lecture 4

Integrals:

- $\int_{-\infty}^{\infty} \cos(x)\ dx$ can not be defined.
-
$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\sin(x)}{x}\ dx &= \lim_{R\to\infty} \int_{-R}^{R} \frac{\sin(x)}{x}\ dx \\ &= \pi\end{aligned}$$

  as a Riemann integral but $\int_{-\infty}^{\infty} \left|\frac{\sin(x)}{dx}\right|\ dx$ diverges.

We will not be able to compute $\int_{-\infty}^{\infty} \frac{\sin(x)}{x}\ dx$ as a Lebesgue integral.

4.1 Integrals of simple functions ([Fri03, §2.5])

Let $(\underline{\overline{X}}, \mathcal{A}, \mu)$ be a measure space. Recall that a **simple function** $f(x)$ is on the form

$$f(x) = \sum_i \alpha_i \chi_{E_i}(x)$$

with $E_i \cap E_j = \varnothing$ for $i \neq j$.

Example 4.1.1.

$$\begin{aligned}f(x) &= \begin{cases} 1, & \text{if } x \in \mathbb{R}\setminus\mathbb{Q} \\ 0, & \text{otherwise}\end{cases} \\ &= \chi_{\mathbb{R}\setminus\mathbb{Q}}\end{aligned}$$

is a simple function.

Integrable simple function and associated integral

Definition 4.1.2. A simple function $f = \sum_i \alpha_i \chi_{E_i}$ is **integrable** if $\mu(E_i) < \infty$ whenever $\alpha_i \neq 0$, with associated **integral**

$$\begin{aligned}\int f(x)\ d\mu(x) &= \int f\ d\mu \\ &= \sum_{i=1}^{m} \alpha_i \mu(E_i).\end{aligned}$$

Remark 4.1.3. By convention, we have that $\alpha_i \mu(E_i) = 0$ if $\alpha_i = 0$ or $\mu(E_i) = \infty$.

Theorem 4.1.4. *If $f = \sum_i^n \beta_j \chi_{F_j}$ is another representation of f, then*

$$\sum_{i=1}^{m} \alpha_i \mu(E_i) = \sum_{j=1}^{n} \beta_j \mu(F_j),$$

i.e. the integral is well-defined.

Integral over a measurable set, for a simple function

Definition 4.1.5. If f is simple, then the integral over the measurable set $E \in \mathcal{A}$ is defined as

$$\int_E f\ d\mu = \int \chi_E \cdot f\ d\mu$$

Example 4.1.6. Let $f = \chi_{\mathbb{R}\setminus\mathbb{Q}}$ and let $(\mathbb{R}, \mathcal{A}, \mu)$ be the Lebesgue measure space. Then

$$\begin{aligned}\int_{[0,1]} f\ d\mu &= \int \chi_{[0,1]} \cdot f\ d\mu \\ &= \int \chi_{[0,1]} \cdot \chi_{\mathbb{R}\setminus\mathbb{Q}}\ d\mu \\ &= \int \chi_{[0,1]\cap\mathbb{R}\setminus\mathbb{Q}}\ d\mu \\ &= \mu([0,1] \setminus \mathbb{Q}).\end{aligned}$$

Notice that $[0,1] = ([0,1] \setminus \mathbb{Q}) \cup ([0,1] \cap \mathbb{Q})$ is a disjoint set so by complete additivity we have that

$$\mu([0,1]) = \mu([0,1] \setminus \mathbb{Q}) + \mu([0,1] \cap \mathbb{Q}).$$

Claim: $\mu([0,1] \cap \mathbb{Q}) = 0$.

Proof. Set $I_n = \left[x - \frac{1}{n}, x + \frac{1}{n}\right]$ and note that $I_{n+1} \subset I_n$. Hence (I_n) is a decreasing sequence of measurable set, and has a limit:

$$\begin{aligned}\overline{\lim}_n I_n &= \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} I_n \\ &= \bigcap_{k=1} I_k.\end{aligned}$$

Also note that $\mu(I_n) = \frac{2}{n} < \infty$ for all n so $\lim_n \mu(I_n) = \mu(\lim I_n)$. This is a countable intersection of measurable sets, hence is measurable. We may also see that

$$\bigcap_{k=1}^{\infty} I_k = \{x\}.$$

Proof. $\subseteq$: x is in all I_n. If y is in all I_n then $x - \frac{1}{n} \leqslant y \leqslant x + \frac{1}{n}$ for all n. If $y \neq x$ then without loss of generality let $y < x \Leftrightarrow 0 < x - y$. By denseness of $\mathbb{Q}$ there is some $0 < \frac{m}{n} < x - y$ hence $0 < \frac{1}{n} < x - y$. Thus $x + \frac{1}{n} < x + (x - y) = y$, contradiction!

$\supseteq$: Direct, since $x \in I_n$ for all n. □

We then have that

$$\begin{aligned} \mu(\{x\}) &= \mu\left(\bigcap_{i=1}^{\infty} I_n\right) \\ &= \mu(\lim_n I_n) \\ &= \lim_n \mu(I_n) \\ &= \lim_n \frac{2}{n} \\ &= 0. \end{aligned}$$

Since $\mathbb{Q}$ is countable, we have that we can write $\mu([0,1] \cap \mathbb{Q})$ as a *countable disjoint union* $\{x_n\}_{n\in\mathbb{N}}$. Therefore,

$$\begin{aligned} \mu([0,1] \cap \mathbb{Q}) &= \mu\left(\bigcup_{n\in\mathbb{N}} \{x_n\}\right) \\ &= \sum_{n\in\mathbb{N}} \mu(\{x_n\}) \qquad (\mu \text{ completely additive}) \\ &= 0. \end{aligned}$$

□

It follows that

$$\begin{aligned} \mu([0,1] \setminus \mathbb{Q}) &= \mu([0,1]) \\ &= 1. \end{aligned}$$

Theorem 4.1.7. *The integral has the "usual properties" (see [Fri03, Theorem 2.5.1]). For example, f integrable $\Rightarrow |f|$ integrable.*

Proof. See book. Note that for "f integrable $\Rightarrow |f|$ integrable" we have that $\int |f|\, d\mu = \sum_{i=1}^{m} |\alpha_i| \mu(E_i)$ is a finite sum. □

4.2 Definition of the integral ([Fri03, § 2.6])

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

Cauchy sequence in the mean

Definition 4.2.1. A sequence of integrable simple functions (f_n) is said to be **Cauchy sequence in the mean** if

$$\int |f_n - f_m| \; d\mu \to 0 \quad \text{as } n, m \to \infty.$$

Integrable (measurable) function

Definition 4.2.2. A measurable function $f : \underline{X} \to [0, \infty]$ is **integrable** if there is a sequence (f_n) of integrable simple functions (f_n) such that

(1) (f_n) is a Cauchy sequence in the mean.
(2) $\lim_n f_n(x) = f(x)$ a.e.

Remark 4.2.3. Note that the condition on being Cauchy in the mean implies that $\lim_n \int f_n \; d\mu$ exists, since

$$\left| \underbrace{\int f_n \; d\mu}_{a_n} - \underbrace{\int f_m \; d\mu}_{a_m} \right| = \left| \int (f_n - f_m) \; d\mu \right|$$
$$\leqslant \int |f_n - f_m| \; d\mu \quad \text{([Fri03, Theorem 2.5.1.(d)])}.$$

and $\int |f_n - f_m| \; d\mu \to 0 \quad \text{as } n, m \to \infty$. Thus $\left(\int f_n \; d\mu \right)$ is a Cauchy-sequence in $\mathbb{R}$, so has a limit (since $\mathbb{R}$ is a complete metric space).

Definition of integral for integrable function f (general)

Definition 4.2.4. Assume that f is an integrable function. Then $\int f \; d\mu = \lim_{n \to \infty} \int f_n \; d\mu$.

Theorem 4.2.5. *The integral is independent of the choice of sequence of simple integral functions (f_n) satisfying (1) and (2) in definition 4.2.2.*

Integral over a measurable set (general)

Definition 4.2.6. f integrable and $E \in \mathcal{A}$. Then

$$\int_E f \; d\mu = \chi_E f \; d\mu.$$

Remark 4.2.7. $\chi_E \cdot f$ integrable.

Integral of measurable but not integrable function

Definition 4.2.8. If $f \geqslant 0$ is measurable and f is not integrable, then we write $\int f \; d\mu = \infty$.

Example 4.2.9. Assume that $\mu(\underline{X}) < \infty$ (finite measure space) and $f : \underline{X} \to [0, M]$ where $M < \infty$ and f is measurable. Then f is integrable.

Solution: By [Fri03, Theorem 2.2.5] there exists a sequence (f_n) of monotone increasing simple functions, and (see the proof) such that

$$0 \leqslant f(x) - f_n(x) \leqslant \frac{1}{2^n}, \qquad \text{if } f(x) > n,$$

in particular, when $f(x) > M$, for all $x \in \underline{X}$. Hence, $\lim_n f_n(x) = f(x)$ for all $x \in \underline{X}$. We then have that

$$\begin{aligned} |f_n(x) - f_m(x)| &\leqslant |f_n(x) - f(x)| + |f(x) - f_m| \\ &\leqslant \frac{1}{2^n} + \frac{1}{2^n} \qquad \text{for all } x \in \underline{X}, \text{ whenever } m, n > M \end{aligned}$$

Therefore,

$$\int |f_n - f_m| d\mu \leqslant \left(\frac{1}{2^n} + \frac{1}{2^m}\right) \mu(\underline{X}) \to 0 \quad \text{as } m, n \to \infty.$$

Proposition 4.2.10 (Problem 2.6.1)**.** *If f is integrable and g is measurable and $g = f$ a.e., then g is integrable and $\int f \, d\mu = \int g \, d\mu$.*

Proposition 4.2.11 (Problem 2.6.3)**.** *Assume that f is integrable. Then f is integrable $\Leftrightarrow$ $|f|$ is integrable.*

Remark 4.2.12. Therefore, $\int_{\mathbb{R}} \frac{\sin(x)}{x} \, dx$ does not exist.

Proof. $\Rightarrow$: There exists integrable simple functions (f_n) such that (a) and (b) holds. Let $g_n := |f_n|$, which are simple, integrable. Then $\lim_n g_n = |f|$ a.e and

$$\begin{aligned} \int |g_n - g_m| \, d\mu &= \int ||f_n| - |f_m|| \, d\mu \\ &\leqslant \int |f_n - f_m| \, d\mu \to 0 \qquad \text{as } m, n \to \infty. \end{aligned}$$

Thus, $|f|$ is integrable. □

Recall: f is **integrable** if there exists a sequence (f_n) of integrable, simple functions such that:

1. (f_n) is Cauchy in the mean.
2. $\lim_{n\to\infty} f_n(x) = f(x)$ a.e.

or alternatively (a) and (b'), where (b') is the statement that the given sequence (f_n) converges *in measure* to f.

Definition 4.2.13. If (f_n) is a sequence of a.e. real-valued, and measurable functions f_n, then we say that (f_n) **converges in measure** to f if for all $\varepsilon > 0$ it holds that

$$\lim_{n\to\infty} \mu(\{x \,:\, |f_n(x) - f(x)| \geqslant \varepsilon\}) \to 0.$$

One may also say "converges in probability".

Theorem 4.2.14 (2.6.1)**.** *f is integrable according to (a),(b) if and only if it is integrable according to (a) and (b').*

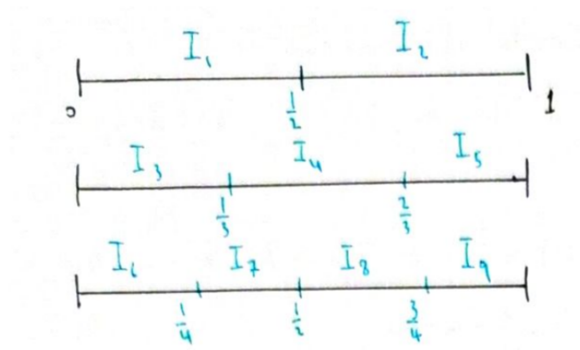

Figure 4.1: I_n

Example 4.2.15. Note that $f_n \to f$ a.e. does not imply that $f_n \to f$ in measure (even if f_n is a.e. real valued). It does however hold if $\mu(\underline{X}) < \infty$ according to [Fri03, Theorem 2.3.2, 2.4.1].

Let

$$f_n(x) = \begin{cases} 1, & \text{if } n \leqslant x < n+1 \\ 0, & \text{otherwise} \end{cases} .$$

Then $\lim_{n\to\infty} f_n(x) = 0$ for all x (it eventually becomes zero for all $n > x$). But

$$\begin{aligned} \mu(\{x \;:\; |f_n(x)| \geqslant \varepsilon\} &= \mu([n, n+1)) \\ &= 1 \qquad \forall n \geqslant 1, \text{and } 0 < \varepsilon \leqslant 1 \end{aligned}$$

with the Lebesgue measure, i.e. it converges almost everywhere to zero, but it does not converge in measure to zero.

Example 4.2.16 ("Skrivmaskinsekvensen)**.** Note: $f_n \to f$ in measure does not imply $f_n \to f$ a.e. Consider $\underline{X} = [0, 1]$ with the Lebesgue measure. Let $f_n = \chi_{I_n}$ where $I_n = \left[-\frac{1}{n}, \frac{1}{n}\right]$. Then $f_n \to 0$ in measure ($\mu(I_\infty) = 0$), or that is, for $0 < \varepsilon \leqslant 1$ we have that $\{x \;:\; |f_n(x)| \geqslant \varepsilon\} = I_n$ and $\mu(I_n) = \frac{1}{n} \to 0$ as $n \to \infty$. For $\varepsilon > 1$ we get $\{x \;:\; |f_n(x)| \geqslant \varepsilon\} = \varnothing$ and $\mu(\varnothing) = 0$.

I.e. $I_1 = \left[0, \frac{1}{2}\right], I_2 = \left[\frac{1}{2}, 1\right], I_3 = \left[0, \frac{1}{3}\right], \ldots$.

Definition/Theorem: Convergence in the mean to $f \Rightarrow$ convergence in measure. If (f_n) is a sequence of integrable functions, then it **converges in the mean** to an integrable function f if

$$\int |f_n - f| d\mu \to 0 \quad \text{as } n \to \infty.$$

Proof.

Remark 4.2.17. I.e. the claim we prove below is that if (f_n) is a sequence of integrable functions convering in the mean to f, then (f_n) converges in measure to f.

Let $\varepsilon > 0$ and let $E_n := \{x \;:\; |f_n(x) - f(x)| \geqslant \varepsilon\}$. Then

$$\begin{aligned} \underbrace{\int |f_n - f| \; d\mu}_{\to 0 \text{ as } n \to \infty} &\geqslant \int_{E_n} |f_n - f| \; d\mu \\ &\geqslant \varepsilon\mu(E_n). \end{aligned}$$

Hence $\lim_{n\to\infty} \mu(E_n) = 0$. $\square$

Chapter 5

Lecture 5

5.1 Summary of properties of the integral

Recall: A sequence (f_n) of integrable functions is **Cauchy sequence in the mean** if

$$\int |f_n - f_m| \, d\mu \xrightarrow{m,n\to\infty} 0.$$

If there exists an integrable function f such that

$$\int |f_n - f| \, d\mu \to 0 \quad \text{as } n \to \infty$$

then (f_n) **converges in the mean to** f.

Furthermore, recall that (f_n) **converges in measure** to f if for all $\varepsilon > 0$ we have that

$$\mu(\{x \ : \ |f_n(x) - f(x)| \geqslant \varepsilon\}) \to 0 \quad \text{as } n \to \infty.$$

Lemma 5.1.1. *If (f_n) converges to both f and g in measure, then $f = g$ a.e.*

Theorem 5.1.2 (2.7.2)**.** *$f_n \to f$ in the mean $\Rightarrow$ $f_n \to f$ in measure.*

Furthermore, we have the following **Completeness result** (recall that $\mathbb{R}$ with the euclidian metric is a *complete* metric space).

Theorem 5.1.3 (2.8.3)**.** *If (f_n) is Cauchy in the mean, then there exists an integrable function f such that $f_n \to f$ in the mean.*

Remark 5.1.4. There is a typo in the proof; last sentence should be "by Theorem 2.8.2".

Theorem 5.1.5 (2.8.4)**.** *Assume that f is integrable. Then for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $E \in \mathcal{A}$ such that $\mu(E) < \delta$, we have that*

$$\left| \int_E f \, d\mu \right| < \varepsilon.$$

Proof. Idea: Approximate f by simple functions. $\square$

Corollary 5.1.6 (Cor. 2.8.5). *If f is integrable and $\lim_{n\to\infty} E_n = E$ (for measurable $E_n \in \mathcal{A}$), such that $\mu(E) = 0$, then*

$$\lim_{n\to\infty} \int_{E_n} f \; d\mu = 0.$$

Example 5.1.7. Let $E_n = [n, \infty)$. Then $\lim_{n\to\infty} E_n = \varnothing$. Thus by the corollary

$$\lim_{n\to\infty} \int_n^\infty f d\mu = 0$$

for integrable f.

Proof. Idea of proof of Cor. 2.8.5: Assume that f is non-negative (f^+, f^{-1}). So $f \geqslant 0$. Introduce $\nu(F) = \int_F f \; d\mu$, this is a (finite) measure. And use [Fri03, Cor. 1.2.3] with ν as μ (with the books notation). □

5.2 Lebesgue Bounded Convergence Theorem

We will talk about **Lebesgue Bounded Convergence Theorem** or sometimes called **Dominated Convergence Theorem**.

Important: When can you put a limit *inside the integral*.

Example 5.2.1. Let

$$f_n(x) = \begin{cases} n, & \text{if } \frac{1}{n} \leqslant x \leqslant \frac{2}{n} \\ 0, & \text{otherwise} \end{cases} .$$

Then $\lim_{n\to\infty} f_n(x) = \underbrace{0}_{=f(x)}$ for all $x \in \mathbb{R}$. Furthermore, we have

$$\begin{aligned} \int_{\mathbb{R}} f_n \; dx &= \int_{\frac{1}{n}}^{\frac{2}{n}} n \; dx \\ &= n\left(\frac{2}{n} - \frac{1}{n}\right) \\ &= 1 \qquad \forall n \geqslant 1 \\ \Rightarrow \lim_{n\to\infty} \left(\int f_n \; dx\right) &= 1. \end{aligned}$$

However, if we put $A_n(\varepsilon) = \{x \;:\; |f_n(x)| \geqslant \varepsilon\}$ then we see that if $\varepsilon > n$ then $A_n(\varepsilon) = \varnothing$ with measure 0, and if $\varepsilon \leqslant n$ then $A_n(\varepsilon) = \left[\frac{1}{n}, \frac{2}{n}\right]$ with measure $\frac{1}{n} \to 0$ as $n \to \infty$. Therefore, we have that

$$\mu(A_n(\varepsilon)) \to 0 \qquad \text{as } n \to \infty \text{ for all } \varepsilon > 0.$$

Slogan: "Converges in mean is stronger than in measure".

Theorem 5.2.2 (LBCT). *Let (f_n) be a sequence of integrable functions, such that $f_n \to f$ a.e. (or in measure), where f is measurable. If $|f_n(x)| \leqslant g(x)$ a.e. for all $n \geqslant 1$ for some integrable function g, then f is integrable and*

$$\lim_{n\to\infty} \int |f_n - f| \; d\mu = 0 \qquad (\textit{i.e. } f_n \to f \textit{ in the mean}).$$

Remark 5.2.3. Note that

$$\left|\int f_n \, d\mu - \int f \, d\mu\right| \leqslant \int |f_n - f| \, d\mu \xrightarrow{n\to\infty} 0$$

implies that

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Proof. See [Fri03] or the appendix. □

Theorem 5.2.4 (Lebesgue monotone convergence theorem)**.** *Assume that* $0 \leqslant f_1 \leqslant f_2 \leqslant \dots$ *where each* f_n *is integrable and let* $f(x) = \lim_{n\to\infty} f_n(x)$*. Then*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Remark 5.2.5. If f is <u>not</u> integrable, then $\int f \, d\mu = \infty$ since $f \geqslant 0$ (see [Fri03, def. 2.6.6]).

Theorem 5.2.6 (Fatou's lemma)**.** *Assume* $f_n \geqslant 0$ *are integrable and let* $f(x) = \underline{\lim}_{n\to\infty} f_n(x)$*. Then*

$$\int f \, d\mu \leqslant \underline{\lim}_{n\to\infty} \int f_n \, d\mu.$$

Remark 5.2.7 (Application of Fatou's lemma)*.* If

$$\begin{aligned} \underline{\lim}_{n\to\infty} \int f_n d\mu &< \infty \\ \Rightarrow \int f \, d\mu &< \infty \\ &\Rightarrow f \text{ integrable.} \end{aligned}$$

Remark 5.2.8. Did not cover during class, but also see Chap. 2.11 in the book for the Riemann integrable ⇒ Lebesgue integrable and same value.

Chapter 6

Lecture 6: Ergodic theory

Motivation: X is a *phase space/state space*.

Definition 6.0.1. Let $T : X \to X$ be a map. If we take $A \subset X$, then we let

$$\begin{aligned} N_A(x,n) &:= \{k \in [0, n-1] \ : \ T^k(x) \in A\} \\ &= \sum_{k=0}^{n-1} \chi_A(T^k(x)). \end{aligned}$$

Assume that $\lim_{n\to\infty} \frac{1}{n} N_A(x,n) = F_A(x)$ exists as a limit, and assume that $F_A(x) = \mu(A)$ where μ is the **probability measure** $\mu(X) = 1$.

Consider the following: We have that

$$\begin{aligned} \chi_{T^{-1}(A)}(x) &= \begin{cases} 1, & \text{if } T(x) \in A, \\ 0, & \text{if } T(x) \notin A \end{cases} \\ &= \chi_A(Tx). \end{aligned} \tag{6.0.1}$$

Note that

$$\begin{aligned} N_{T^{-1}(A)}(x,n) &= \sum_{k=0}^{n-1} \chi_{T^{-1}(A)}(T^k(x)) \\ &= \sum_{k=0}^{n-1} \chi_A(T^{k+1}(x)) \\ &= N_A(x,n) + \chi_A(T^n(x)) - \chi_A(x), \end{aligned}$$

where we used 6.0.1. If we divide this last line by n and take the limit, then even though $\chi_A(T^n(x))$ depends on n, it only takes values zero and one, so it goes to 0, and $\chi_A(x)$ is either zero or one, so it also goes to zero. Thus we get that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} N_{T^{-1}A}(x,n) &= \lim_{n\to\infty} N_A(x,n) \\ \Rightarrow F_{T^{-1}(A)}(x) &= F_A(x) \\ \Rightarrow \mu(A) &= \mu(T^{-1}(A)). \end{aligned}$$

Setting: Let $(X, \mathcal{A}, \mu)$ be a measure space such that $\mu(X) = 1$. Assume that $T : X \to X$ is **measure preserving** , i.e. that $T^{-1}(A) \in \mathcal{A}$ for every measurable $A \in \mathcal{A}$, and that $\mu(T^{-1}(A)) = \mu(A)$ for $A \in \mathcal{A}$.

Remark 6.0.2. If T *preserves* μ, then μ is T**-invariant**

Definition 6.0.3 (T-invariant set)**.** A set $B \in \mathcal{A}$ is T**-invariant** if $T^{-1}(B) = B$.

Remark 6.0.4. If $x \in B = T^{-1}(B)$ then $T(x) \in B$, and if $x \notin B \Rightarrow x \notin T^{-1}(B)$ so that $T(x) \notin B$.

If $T_0 = T|_B : B \to B$ then T_0 is measure preserving.

Definition 6.0.5 (Ergodic: definition I)**.** T is **ergodic** with respect to μ, if the preimage of A under T equals A implies that A is either a null set or has measure one. Or more formally: T is ergodic if $T^{-1}(A) = A \Rightarrow \mu(A) = 0$ or $\mu(A) = 1$.

We provide an equivalent definition of ergodic below.

Definition 6.0.6 (Ergodic: definition II)**.** We say that a map $T : X \to X$ is **ergodic** if $\mu(T^{-1}(A)\triangle A) = 0$ implies that $\mu(A) = 0$ or $\mu(A) = 1$, where we recall that the **symmetric difference** $A\triangle B$ of two sets A, B is defined as

$$A\triangle B = (A \setminus B) \cup (B \setminus A).$$

Example 6.0.7. A concrete example: Let $X = [0, 1]$, with μ the Lebesgue measure, and let

$$T(x) = \begin{cases} 2x, \text{ if } 0 \leqslant x \leqslant \frac{1}{2} \\ 2 - 2x, \text{ if } \frac{1}{2} < x \leqslant 1 \end{cases}$$

We claim that this map is ergodic with respect to μ.

Definition 6.0.8 (Essentially T-invariant)**.** A measurable function $f : X \to \mathbb{R}$ is **essentially** T**-invariant** if $f(T(x)) = f(x)$ a.e. ($f \circ T$ is measurable).

Proposition 6.0.9. *T is ergodic if and only if every essentially T-invariant function is constant a.e.*

Proof. $\underline{\Rightarrow}$: Assume that $f \circ T = f$ everywhere for simplicity. For $r \in \mathbb{R}$, let $E_r := \{x \ : \ f(x) \geqslant r\}$. Note that $E_r \in \mathcal{A}$ since f is measurable by definition. Then

$$\begin{aligned} T^{-1}(E_r) &= \{x \ : \ T(x) \in E_r\} \\ &= \{x \ : \ (f \circ T)(x) \geqslant r\} \\ &= \{x \ : \ f(x) \geqslant r\} \\ &= E_r. \end{aligned}$$

Thus, E_r is T-invariant, so by ergodicity of T, we have that $\mu(E_r) = 0$ or $\mu(E_r) = 1$. Notice that if $r_1 < r_2$ then $f(x) \geqslant r_2 \Rightarrow f(x) \geqslant r_1$. Hence $E_{r_2} \subset E_{r_1}$. Set $S := \{r \ : \ u(E_r) = 1\}$. If $r \in S$ and $s < r$ then $E_r \subset E_s$ which implies that $\mu(E_s) \geqslant \mu(E_r) = 1$. Since we (I presume) are working with a measure μ such that $\mu(X) = 1$ we have that $1 = \mu(X) \geqslant \mu(E_s) \geqslant \mu(E_r) = 1$, thus $\mu(E_s) = 1$. This means that S is *downward* closed, i.e. if $r \in S$ and $s < r \Rightarrow s \in S$. Set $r_0 := \sup S$. This means that for any $\varepsilon > 0$, by definition of the supremum, there is some $r \in S$ such that $r_0 - \varepsilon < r \leqslant r_0$. Since S is downward closed we have that $r_0 - \varepsilon \in S$ and so $\mu(E_{r_0-\varepsilon}) = 1$. If $\mu(E_{r_0+\varepsilon}) = 1$ then $r_0 + \varepsilon \in S$, contradicting that $r_0 = \sup S$. Since we showed that $\mu(E_r) = 0$ or $\mu(E_r) = 1$ for every $r \in \mathbb{R}$, it follows that $\mu(E_{r_0+\varepsilon}) = 0$.

Set $A_m = \left\{f \geqslant r_0 + \frac{1}{m}\right\}$ for $m \geqslant 1$. Then we have that

$$\bigcup_{m=1}^{\infty} A_m = \{x \ : \ f > r_0\} \tag{6.0.2}$$

by countable subadditivity we get that $\mu(\{x \ : \ f > r_0\}) = 0$. Now let $B_m = \left\{x \ : \ f \geqslant r_0 - \frac{1}{m}\right\}$ for $m \geqslant 1$. By downward closedness of S we have $\mu(B_m) = 0$ for all m. Furthermore, we have

$$\bigcap_{m=1}^{\infty} B_m = \{x \ : \ f(x) \geqslant r_0\}.$$

Furthermore, note that $B_m \supset B_{m+1}$ is a decreasing sequence, and the space is σ-finite. We can thus conclude that 1): (B_m) has a limit which equals $\bigcap_{m=1}^{\infty} B_m$, and that

$$\begin{aligned}
\lim_{m\to\infty} \mu(B_m) &= \mu\left(\lim_{m\to\infty} B_m\right) \\
&= \mu\left(\bigcap_{m=1}^{\infty} B_m\right) \\
&= \mu(\{x \ : \ f(x) \geqslant r_0\}) \\
&= 1.
\end{aligned}$$

Summarizing, we have found that $\mu(\{x \ : \ f(x) > r_0\}) = 0$ and $\mu(\{x \ : \ f(x) \geqslant r_0\}) = 1$. We may write $\{x \ : \ f(x) \geqslant r_0\} = \{x \ : \ f(x) > r_0\} \sqcup \{x \ : \ f(x) = r_0\}$. Thus we have that

$$\begin{aligned}
\mu(\{x \ : \ f(x) \geqslant r_0\}) &= \underbrace{\mu(\{x \ : \ f(x) > r_0\})}_{=0} + \mu(\{x \ : \ f(x) = r_0\}) \\
&= 1 \\
\Rightarrow \mu(\{x \ : \ f(x) = r_0\}) &= 1.
\end{aligned}$$

Let $N := \{x \ : \ f(x) \neq r_0\}$. Then $X = \{x \ : \ f(x) = r_0\} \sqcup N$ and $\mu(N) = 0$ follows since $\mu(X) = 1$, $\mu(\{x \ : \ f(x) = r_0\}) = 1$ and $\mu(N) \geqslant 0$. Thus $f = r_0$ a.e.

Remark 6.0.10. Set $N := \{x \ : \ f(T(x)) \neq f(x)\}$. Then $\mu(N) = 0$. One may show that $E_r \setminus N = T^{-1}(E_r) \setminus N$ for each $r \in \mathbb{R}$. But then, letting $A = E_r$ and $B = T^{-1}(E_r)$, we have that

$$\begin{aligned}
(A \triangle B) \cap N^c &= ((A \cap B^c) \cup (B \cap A^c)) \cap N^c \\
&= (A \cap B^c \cap N^c) \cup (B \cap A^c \cap N^c) \\
&= \varnothing
\end{aligned}$$

since $A \cap N^c = B \cap N^c$. Thus $T^{-1}(E_r) \triangle E_r \subset N$ so $\mu(T^{-1}(E_r) \triangle E_r) = 0$, which by ergodicity (definition II) implies that $\mu(E_r) = 0$ or $\mu(E_r) = 1$. Then the proof proceeds as above, but without the assumption that $f(T(x)) = f(x)$ a.e. in the beginning, showing that $f(T(x)) = f(x)$ everywhere is stronger than what is needed for (atleast this direction) of the proof to go through.

$\underline{\Leftarrow}$: Exercise: Take any set $A \in \mathcal{A}$ such that $T^{-1}(A) = A$. Set $f = \chi_A$. Then $f \circ T = \chi_A \circ T = \chi_{T^{-1}(A)} = \chi_A$. Hence χ_A is essentially T-invariant and so $\chi_A = c$ a.e., but $\chi_A \cdot \chi_A = c^2 = c$, which implies that $c = 0, 1$. If $\chi_A = 0$ then $\int \chi_A = \mu(A) = 0$ and otherwise $\mu(A) = 1$, which is what we wanted to show. $\square$

6.1 Birkhoff's ergodic theorem

Theorem 6.1.1. *Let $T : X \to X$ be measure preserving and let f be integrable. Then*

$$\overline{f} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x))$$

exists a.e. Furthermore, $\overline{f}$ is integrable and

$$\int \overline{f}\ d\mu = \int f\ d\mu.$$

Remark 6.1.2. We have

$$\begin{aligned}\overline{f}(T(x)) &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^{k+1}(x)) \\ &= \lim_{n\to\infty} \frac{n+1}{n} \left(\frac{1}{n+1} \sum_{k=0}^{n} f(T^k(x)) - f(x) \right) \\ &= \overline{f}(x),\end{aligned}$$

a.e., i.e. $\overline{f}$ is essentially T-invariant.

Thus, if T is ergodic with respect to μ, then by proposition 6.0.9 we have that $\overline{f} = c \in \mathbb{R}$ a.e. By theorem 6.1.1

$$\begin{aligned}\int f\ d\mu &= \int \overline{f}\ d\mu \\ &= \int c\ d\mu \\ &= \mu(X)c \\ &= c \\ &= \overline{f},\end{aligned}$$

Therefore, we have that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) = \int f\ d\mu \quad \text{a.e.}$$

If $f = \chi_A$ for $A \in \mathcal{A}$ then

$$\begin{aligned}\int \chi_A &= \mu(A) \\ &= \lim_{n\to\infty} \frac{1}{n} \chi_A(T^k(x)) \quad \text{a.e.}\end{aligned}$$

We prove Birkhoff's theorem (6.1.1).

Proof. For simplicity, let f be bounded. Let

$$f_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) \quad \text{(measurable)},$$

where measurability comes from the fact that in this case the compositions works out so that T^k measurable and then $f \circ T^k$ measurable. Multiplying by a constant does not change measurability, and measurability is preserved under finite summation.

Step 1: Let $\overline{f}(x) = \underline{\lim}_n f_n(x)$ (which is measurable). □

Prove that $\lim_{n\to\infty} f_n(x)$ exists a.e. using Maximal ergodic theorem; Yosida and Kakotani, 1939.

Step 2: Prove that $\overline{f}$ is integrable and $\int \overline{f}\ d\mu = \int f\ d\mu$.

By Fatou's lemma (5.2.6), we have

$$\begin{aligned}\int \underline{\lim}_{n\to\infty} |f_n|\ d\mu &\leqslant \underline{\lim}_{n\to\infty} \int |f_n|\ d\mu \\ &= \underline{\lim}_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int \left| f(T^k(x)) \right|\ d\mu \\ &= \underline{\lim}_{n\to\infty} \int |f|\ d\mu \\ &= \int |f|\ d\mu \\ &< M,\end{aligned}$$

where in the last equality we used the we believe equivalent definition that f is integrable if $\int |f|\ d\mu < \infty$.

We also used the following **Claim**: $\int g \circ T^k\ d\mu = \int f\ d\mu$ if $f \geqslant 0$ is integrable and T is measure preserving (note that we believe this holds even if $f \geqslant 0$ does not necessarily hold; but this assumption, which is the only thing we need [since we want to show it for $|f|$] seem to make the proof a tad bit easier):

Proof. We start by simple functions. Let $s = \sum_i^n \alpha_i \chi_{E_i}$ we then get (since $T^k(x) \in E_i \Leftrightarrow x \in (T^k)^{-1}(E_i)$) that

$$\begin{aligned}(s \circ T^k)(x) &= \sum_{i=1}^n \alpha_i \chi_{E_i}(T^k(x)) \\ &= \sum_{i=1}^n \alpha_i \chi_{(T^k)^{-1}(E_i)}(x),\end{aligned}$$

so that

$$\begin{aligned}\int s \circ T^k\ d\mu &= \sum_{i=1}^n \alpha_i \int \chi_{(T^k)^{-1}(E_i)}\ d\mu \\ &= \sum_{i=1}^n \alpha_i \mu((T^k)^{-1}(E_i)) \\ &= \sum_{i=1}^n \alpha_i \mu(E_i) \\ &= \int s\ d\mu\end{aligned}$$

where we used that if T is measure preserving so that $\mu(T^{-1}(A)) = \mu(A)$ then it holds that $\mu(T^{-k}(A)) = \mu(A)$ for all $k \geqslant 1$.

For the case $|f| \geqslant 0$, note that f is integrable so $|f|$ is. We get a sequence of simple monotone increasing functions that converges to $|f|$ everywhere. By [Fri03, Theorem 2.10.1] this sequence consists of integrable functions. We have that $s_n \circ T^k \to |f| \circ T^k$ pointwise. By the Lebesgue monotone convergence theorem it follows that

$$\begin{aligned}\int |f| \circ T^k \, d\mu &= \lim_{n\to\infty} \int s_n \circ T^k \, d\mu \\ &= \lim_{n\to\infty} \int s_n \, d\mu \\ &= \int |f| \, d\mu,\end{aligned}$$

where we used the previous simple case to say that $\int s_n \circ T^k \, d\mu = \int s_n \, d\mu$ for $n \geqslant 1$. But $|f| \circ T^k = |f \circ T^k|$ so indeed

$$\int |f| \, d\mu = \int |f \circ T^k| \, d\mu.$$

Moving on, since $\lim_{n\to\infty} f_n = \overline{f}$ a.e. we have that $\lim_{n\to\infty} |f_n| = |\overline{f}|$ (e.g. use reverse triangle inequality), so indeed $\underline{\lim}_n f_n = \overline{f}$, and so $\int \overline{f} \, d\mu < \infty$ by the above, hence by the equivalent characterization of integrable (assuming this is true) it follows that $\overline{f}$ is integrable. It remains to prove that $\int f \, d\mu = \int \overline{f} \, d\mu$. We have assumed that $|f(x)| \leqslant M < \infty$ everywhere, so we find that

$$\begin{aligned}|f_n(x)| &= \left| \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) \right| \\ &\leqslant \frac{1}{n} \sum_{k=0}^{n-1} |f(T^k(x))| \\ &\leqslant M.\end{aligned}$$

Hence, we have that

$$\begin{aligned}\int |f_n| \, d\mu &\leqslant \int M \, d\mu \\ &= M\mu(X) \\ &= M, \quad \forall n \geqslant 1.\end{aligned}$$

By an application of Lebesgue bounded convergence theorem we find that

$$\lim_{n\to\infty} \int f_n \, d\mu = \int \overline{f} \, d\mu.$$

But we also have

$$\int f_n \, d\mu = \frac{1}{n} \sum_{k=0}^{n-1} \int f(T^k(x)) \, d\mu.$$

One may show similarly as one shows that $\int |f \circ T^k| \, d\mu = \int |f| \, d\mu$ (we believe) that

$$\int f \circ T^k \, d\mu = \int d\mu$$

thus $\int f_n \, d\mu = \int f \, d\mu$ for all $n \geqslant 1$. It follows that $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$ so $\int \overline{f} \, d\mu = \int f \, d\mu$. $\square$

Chapter 7

Lecture 7: The Lebesgue integral on $\mathbb{R}$

Notice that

$$\int \psi_{(c,d)} \cdot f(t)\ dt = \int_c^d f(t)\ dt$$

is the **Lebesgue integral of** f **over** (c, d).

We have

- $\int_d^c f(t)\ dt = -\int_c^d f(t)\ dt$.
- $\int_c^c f\ dt = 0$.

Definition 7.0.1. Assume that f is integrable on (a, b). Then

$$g(x) = \int_a^x f(t)\ dt + C, \qquad (x \in (a, b))$$

is the **indefinite** integral of f.

Remark 7.0.2. It follows from [Fri03, Theorem 2.8.4] that g is absolutely continuous ($\lambda(a, x]) = g(x)$).

Theorem 7.0.3 (2.8.4)**.** *Assume that f is integrable and let*

$$\lambda(E) = \int_E f\ d\mu$$

for $E \in \mathcal{A}$. Then for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $E \in \mathcal{A}$, if $\mu(E) < \delta$ then

$$\left| \int_E f\ d\mu \right| < \varepsilon.$$

Theorem 7.0.4.

$$g'(x) = f(x) \qquad \textit{almost everywhere.}$$

Proof. Assume that f is Lebesgue integrable on (a, b).

Lemma 7.0.5 (Lemma 1)**.** *If $\int_a^x f(t)\ dt = 0$ for all x, then $f(x) = 0$ almost everywhere.*

Proof. For any interval (c, d), we have that

$$\begin{aligned}\int_c^d f(t)\ dt &= \int_a^d f(t)\ dt - \int_a^c f(t)\ dt \\ &= 0.\end{aligned}$$

Lemma 7.0.6 (Added by us - not in lecture)**.** *If $E \subset \mathbb{R}$ is an open subset, then E can be written as a* countable, disjoint union *of intervals (a, b) [possibly with $a = -\infty$ or $b = \infty$].*

Proof. For $x \in E$, let

$$I_x := \bigcup \{(a, b)\ :\ x \in (a, b) \text{ and } (a, b) \subset E\}.$$

Clearly I_x is open. Furthermore, each I_x is a union of connected subsets, each sharing a point x in common, hence connected $\Rightarrow I_x$ is an interval (the only non-empty connected sets in $\mathbb{R}$ are intervals). We conclude that I_x is an open interval in $\mathbb{R}$ for each $x \in E$. Furthermore, assume that I is an interval containing x and contained in E. Then by construction $I \subset I_x$, so that I_x is *maximal* (under this property). Assume that $x \neq x'$ and that $I_x \neq I_{x'}$, but that $I_x \cap I_{x'}$ is non-empty, i.e. contains some element z. Then I_z contains both I_x and $I_{x'}$ but is strictly bigger since $I_x \neq I_{x'}$, contradiction! Therefore either $I_x = I_{x'}$ or they are pairwise disjoint. Each interval I_x contains a rational number r by denseness of $\mathbb{Q}$ in $\mathbb{R}$. Label each interval I_x with some rational number r_x. Then the r_x are distinct for distinct $x \in E$. Therefore, if we let $\Phi : \{I_x\}_{x \in E} \to \mathbb{Q}$ be defined by $\Phi(I_x) = r_x$, this is an injection. Hence

$$\bigsqcup_{x \in E} I_x = E$$

is a *countable* disjoint union of open intervals. □

Recall that $\Lambda(\cdot)$ is completely additive (the *indefinite integral* of f). This means that for any open $E \subset (a, b)$, we have that

$$\begin{aligned}\int_E f\ d\mu &= \sum_{i=1}^{\infty} \underbrace{\int_{a_i}^{b_i} f(t)\ dt}_{=0} \\ &= 0.\end{aligned}$$

Since E could be written as a disjoint union of intervals, which are Lebesgue, measurable, it follows that $(a, b) \setminus E$ is Lebesgue-measurable, and so by complete additivity of $\lambda(\cdot)$ we have

$$\int_{(a,b)\setminus E} f(t)\ dt = \int_{(a,b)} f(t)\ dt - \underbrace{\int_E f(t)\ dt}_{=0}, \tag{7.0.1}$$

for any open set $E \subset (a, b)$ (in particular $\int_a^b f(t)\ dt = 0$).

Assume there was some Lebesgue-measurable $F \subset (a, b)$ such that $\mu(F) > 0$ and $f(x) > 0$ on F. We have

$$\begin{aligned}\mu((a, b) \setminus F) &= \mu((a, b)) - \mu(F) \\ &= (b - a) - \mu(F) \\ &< b - a.\end{aligned}$$

Recall that $\mu(\cdot)$ is the *completion* of the Lebesgue *outer* measure

$$\mu^*(A) := \inf \left\{ \sum_{k=1}^{\infty} \lambda(I_k) \ : \ \bigcup_{k=1}^{\infty} I_k \supset A \right\}.$$

Thus, consider that $(a, b) \setminus F$ is measurable, and take ε such that $0 < \varepsilon < \mu(F)$. Since $(b - a) - \mu(F)$ is the infimum, there must be exist some family $(I_k)_k$ of intervals such that

$$\sum_{k=1}^{\infty} \lambda(I_k) < (b - a) - \mu(F) + \varepsilon$$

and $\bigcup_k I_k \supset (a, b) \setminus F$ where I_k are open (bounded) intervals in $\mathbb{R}$. Set $E := \bigcup_{k=1}^{\infty} I_k$. Then first, note that E is open and measurable, and

$$\begin{aligned} \mu(E) &\leqslant \sum_{k=1}^{\infty} \mu(I_k) \\ &= \sum_{k=1}^{\infty} \lambda(I_k) \\ &< (b - a) - \mu(F) + \varepsilon \\ &< b - a. \end{aligned}$$

We have $E \supset (a, b) \setminus F$. Set $G := (a, b) \setminus E$. Then

$$\begin{aligned} (a, b) \setminus E &\subset (a, b) \setminus ((a, b) \setminus F) \\ &= F. \end{aligned}$$

Observe that $G \cup E = (a, b)$ is a disjoint union, hence

$$\begin{aligned} b - a &= \mu(G) + \mu(E) \\ \Leftrightarrow (b - a) - \mu(E) &= \mu(G) \\ \Rightarrow \mu(G) &> 0 \qquad \text{(since } \mu(E) < b - a\text{)}. \end{aligned}$$

We have that

$$\begin{aligned} \int_{G \cup E} f \ d\mu &= \int_G f \ d\mu + \int_E f \ d\mu \\ &= \int_a^b f \ d\mu \\ \Leftrightarrow \int_G f \ d\mu &= \int_a^b f \ d\mu - \int_E f \ d\mu \\ &= 0. \end{aligned}$$

Furthermore, $f > 0$ on G since it holds on F.

Lemma 7.0.7 ([Fri03, Theorem 2.7.5])**.** *If f is an integrable function that is positive everywhere on a measurable set E, then $\int_E f \ d\mu = 0$ implies that $\mu(E) = 0$.*

By the lemma, we find that $\mu(G) = 0$, contradiction! Therefore, there can be no Lebesgue measurable set $F \subset (a, b)$ such that $\mu(F) > 0$ and $f(x) > 0$ on F. In particular, since f is measurable, we see that $\{f > 0\}$ must have measure zero. Observe that $\{f < 0\} = \{f^- > 0\}$. By applying the same reasoning on $\{f^- > 0\}$ we get a contradiction, which implies that also $\{f < 0\}$ must have measure zero, so that $\{f \neq 0\} = \{f > 0\} \sqcup \{f < 0\}$ must have measure zero $\Rightarrow$ $f = 0$ almost everywhere. $\square$

Lemma 7.0.8 (Lemma 2). *If $|f| \leqslant M < \infty$ then the statement of theorem 7.0.4 holds.*

Recall that

$$f^+(x) = \begin{cases} f(x), & \text{if } f(x) > 0 \\ 0, & \text{if } f(x) \leqslant 0 \end{cases}$$

and

$$f^-(x) = \begin{cases} -f(x), & \text{if } f(x) < 0 \\ 0, & \text{if } f(x) \geqslant 0 \end{cases} .$$

We can write

$$\begin{aligned} g(x) &= \int_a^x f(t)\ dt \\ &= \int_a^x (f^+(t) - f^-(t))\ dt \qquad (\text{since } f(t) = f^+(t) - f^-(t)) \\ &= \underbrace{\int_a^x f^+(t)\ dt}_{h^+(t)} - \underbrace{\int_a^x f^-(t)\ dt}_{h^-(t)}. \end{aligned}$$

Then the functions $h^{\pm}(t)$ are increasing (non-decreasing). There is a lemma (proof not covered in class) which states that a non-decreasing function $f : [a, b] \to \mathbb{R}$ is differentiable almost everywhere. Assuming the result (black-boxing it for now), we find that $(h^{\pm})'$ exists almost everywhere. Let

$$A(\delta) = \frac{h^+(x+\delta) - h^+(x)}{\delta}$$

and let

$$B(\delta) = \frac{h^-(x+\delta) - h^-(x)}{\delta}.$$

Then $\lim_{\delta\to 0} A(\delta) = (h^+)'(x) := a$ and $\lim_{\delta\to 0} B(\delta) = (h^-)'(x) := b$. For any $\varepsilon > 0$ there exists $\eta_1, \eta_2 > 0$ such that if $0 < |\delta| < \eta_1$ then $|A(\delta) - a| < \frac{\varepsilon}{2}$ and similarly $|B(\delta) - b| < \frac{\varepsilon}{2}$ whenever $0 < |\delta| < \eta_2$. Set $\eta := \min\{\eta_1, \eta_2\}$. Then we find that whenever $0 < |\delta| < \eta$ we have that

$$\begin{aligned} |A(\delta) - B(\delta) - (a - b)| &\leqslant |A(\delta) - a| + |B(\delta) - b| \\ &= \varepsilon. \end{aligned}$$

Therefore, $\lim_{\delta\to 0}(A(\delta) - B(\delta)) = (h^+)'(x) - (h^-)'(x)$. But $A(\delta) - B(\delta) = \frac{g(x+\delta)-g(x)}{\delta}$ so this means that $g'(x) = (h^+)'(x) - (h^-)'(x)$ exists a.e.

We are left with proving that $g'(x) = f(x)$. For $0 < h < b - x$, we have that

$$\begin{aligned} \left|\frac{g(x+h) - g(x)}{h}\right| &= \left|\frac{1}{h}\int_x^{x+h} f(t)\ dt\right| \\ &\leqslant M, \end{aligned}$$

and

$$\lim_{h\to 0} \frac{g(x+h) - g(x)}{h} = g'(x)$$

almost everywhere, where $\frac{g(x+h)-g(x)}{h}$ as a function of x is measurable. Notice that $x \mapsto x+h$ is a continuous map (hence measurable), and $g(\cdot)$ is measurable, so that $g(x+h)$ is measurable, and since $g(x)$ is measurable we have that $g(x+h)-g(x)$ is measurable. Furthermore, multiplying by the constant function $\frac{1}{h}$ (fixing h and letting x vary within appropriate bounds) we find that indeed $x \mapsto \frac{g(x+h)-g(x)}{h}$ is measurable, for each fixed h. Set $h_n = \frac{1}{n}$, and $q_n(x) = \frac{g(x+h_n)-g(x)}{h_n}$. Then we get a sequence of measurable functions (q_n) such that $|q_n(x)| \leqslant M$ for all n, where M is integrable over the finite measure space (a,b). By Lebesgue bounded convergence theorem we find that

$$\begin{aligned} \lim_{h\to 0} \int_a^x \frac{g(t+h)-g(t)}{h}\ dt &= \int_a^x \lim_{h\to 0} \frac{g(t+h)-g(t)}{h}\ dt \\ &= \int_a^x g'(t)\ dt. \end{aligned}$$

Remark 7.0.9. Note that *naively*, our functions q_n are not defined on the same domain. We may amend this by redefining q_n as

$$q_n(x) := \begin{cases} \underbrace{\frac{g(x+h_n)-g(x)}{h_n}}_{:=r_n}, & \text{if } x \in (a, b-h_n) \\ 0, \quad \text{if } x \in [b-h_n, b) \end{cases}$$

so that $q_n(x)$ is now defined on (a,b) for all $n \geqslant 1$. To see that this is still a measurable function, observe that, for $U \subset \mathbb{R}$ open, we have that

$$q_n^{-1}(U) = ((a, b-h_n) \cap r_n^{-1}(U)) \cup ([b_n - h, b))$$

if $0 \in U$ and otherwise this is just $(a, b-h_n) \cap r_n^{-1}(U)$, which in either way is measurable. To see that $q_n(x)$ as defined above is integrable, note that $|q_n|$ is essentially bounded by $M < \infty$ on (a,b), so is indeed integrable.

$$\int_a^x \frac{g(t+h)-g(t)}{h}\ dt = \frac{1}{h}\int_a^x g(t+h)\ dt - \frac{1}{h}\int_a^x g(t)\ dt \tag{7.0.2}$$

$$= \underbrace{\frac{1}{h}\int_{a+h}^{x+h} g(t)\ dt}_{I_1} - \underbrace{\frac{1}{h}\int_a^x g(t)\ dt}_{I_2} \qquad \text{(change of variable } u = t+h) \tag{7.0.3}$$

$$= \frac{1}{h}\int_x^{x+h} g(t)\ dt - \frac{1}{h}\int_a^{a+h} g(t)\ dt \tag{7.0.4}$$

where the picture below tries to illustrate what happened in the last step:

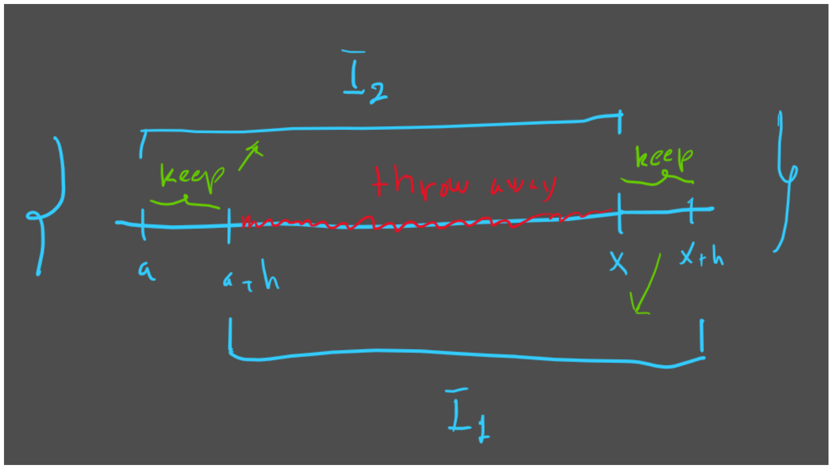

Figure 7.1: Illustration of last step in the above computation.

To be a bit more formal: First, we recall the following:

Lemma 7.0.10 ([Fri03, Problem 1.6.5])**.** *Consider the transformation $Tx = \alpha x + \beta$ from the real line onto itself, where α, β are real numbers and $\alpha \neq 0$. It maps sets E onto sets $T(E)$. Denote by μ (μ^*) the Lebesgue-measure (outer measure) on the real line. Then*

(a) For any set E, $\mu^*(T(E)) = |\alpha| \cdot \mu^*(E)$.

(b) E is Lebesgue-measurable if and only if $T(E)$ is Lebesgue-measurable.

(c) If E is Lebesgue-measurable, then $\mu(T(E)) = |\alpha| \cdot \mu(E)$.

Let $T(t) = t + h$. For fixed $0 < h < b - x$, and $E = (a + h, x + h)$ we then have that

$$\begin{aligned} \mu(T(E)) &= \mu(E) \\ \Leftrightarrow \mu((a + h, x + h)) &= \mu(a, x), \end{aligned}$$

(note that this part can be seen directly by definition of the Lebesgue-measure). In particular, by the lemma ((b) and (c)), any such map T is *measure preserving*.

Theorem 7.0.11 (Exercise 3, compendium on ergodic theory)**.** *If f is integrable, and T is measure-preserving, then*

$$\int (f \circ T)\ d\mu = \int f\ d\mu.$$

Lemma 7.0.12. *If $s = \sum_i \alpha_i \chi_{E_i}$ is a simple integrable function, and T is measure preserving, then*

$$\int (s \circ T)\ d\mu = \int s\ d\mu.$$

Proof. Observe that

$$\begin{aligned} (s \circ T)(x) &= \sum_i \alpha_i \chi_{E_i}(Tx) \\ &= \sum_i \alpha_i \chi_{T^{-1}(E_i)}(x) \qquad (Tx \in E_i \Leftrightarrow x \in T^{-1}(E_i)). \end{aligned}$$

Therefore, by definition 4.1, we have that

$$\begin{aligned}
\int (s \circ T)\ d\mu &= \sum_i \alpha_i \mu(T^{-1}(E_i)) \\
&= \sum_i \alpha_i \mu(E_i) \qquad \text{(since } T \text{ is measure-preserving)} \\
&= \int s\ d\mu.
\end{aligned}$$

Assume that $f(x)$ is integrable, and write $f = f^+ - f^-$. Then f^+ and f^- are integrable , *positive functions*. By [Fri03, Theorem 2.2.5] we have a sequences of non-negative, measurable, monotone-increasing functions $(s_n^\pm)$ that converges to $f^\pm$ everywhere. One may show (use [Fri03, Theorem 2.10.1]) that $s_n^\pm$ is integrable for each $n \geqslant 1$. Notice that $s_n^\pm \circ T \leqslant s_{n+1}^\pm \circ T$. For each fixed x we then have that $s_n^\pm(Tx) \to f^\pm(Tx) = (f^\pm \circ T)(x)$ (where $f^\pm \circ T$ is measurable by exercise 2 on ergodic theory), i.e. $(s_n^\pm \circ T) \to f^\pm \circ T$ everywhere, and each term in the sequence, $s_n^\pm \circ T$ is integrable, by the proof of lemma 7.0.12. Then we find that $f^\pm \circ T$ is integrable, since by the Lebesgue monotone convergence theorem we have

$$\begin{aligned}
\int (f^\pm \circ T)\ d\mu &= \lim_{n\to\infty} \int (s_n^\pm \circ T)\ d\mu \\
&= \lim_{n\to\infty} \int s_n^\pm\ d\mu \\
&= \int f^\pm\ d\mu \\
&< \infty \qquad \text{(exercise 1, HW2, since } f^\pm \geqslant 0)
\end{aligned}$$

Take a simple function $s = \sum_i \alpha_i \chi_{E_i}$ such that $0 \leqslant s \leqslant f$ everywhere. Then on E_i we must have that $f^\pm \geqslant \alpha_i$. It follows that $E_i \subset \{f^\pm \geqslant \alpha_i\}$ (the latter measurable) but $\mu(\{f^\pm \geqslant \alpha_i\}) < \infty$ by [Fri03, Theorem 2.7.1.(h)] so that (by monotonicity) $\mu(E_i) < \infty$ for all i such that $\alpha_i \neq 0$. Therefore, s is integrable, and so

$$I(s) = \int s\ d\mu \leqslant \int f^\pm d\mu = M < \infty$$

for all such simple functions s. Therefore, $I(f^\pm) \leqslant M$, which by exercise 1, homework 2 implies that $f^\pm$ is integrable.

We have that $f \circ T = (f^+ \circ T) - (f^- \circ T)$. We showed above that $f^+ \circ T$ and $f^- \circ T$ were integrable, and so it follows that $f \circ T$ is integrable, with integral

$$\begin{aligned}
\int (f \circ T)\ d\mu &= \int ((f^+ \circ T) - (f^- \circ T))\ d\mu \\
&= \int f^+ \circ T\ d\mu - \int f^- \circ T\ d\mu \\
&= \int f^+\ d\mu - \int f^-\ d\mu \\
&= \int (f^+ - f^-)\ d\mu \\
&= \int f\ d\mu,
\end{aligned}$$

which is what we wanted to show. $\square$

One may show that since $|g(t)| \leqslant M(b-a)$, it is integrable on (a,b). Therefore, if we let $T(t) = t+h$, and note that $f := g \cdot \chi_{(a+h,x+h)}$ is integrable, and that

$$\begin{aligned}(f \circ T)(t) &= g \cdot \chi_{(a+h,x+h)}(t+h) \\ &= g(t+h) \cdot \chi_{(a+h,x+h)}(t+h) \\ &= g(t+h) \cdot \chi_{T^{-1}(a,x)}(t) \\ &= g(t+h)\chi_{(a,x)}(t).\end{aligned}$$

Therefore, it follows by theorem 7.0.11 that

$$\begin{aligned}\int (f \circ T)\ dt &= \int f\ dt \\ \Leftrightarrow \int_a^x g(t+h)\ dt &= \int_{a+h}^{x+h} g(t).\end{aligned}$$

This motivates equation 7.0.3. Furthermore, since g is integrable, we have that

$$\begin{aligned}\int_{a+h}^{x+h} g(t)\ dt - \int_a^x g(t)\ dt &= \left(\int_{a+h}^{x} g(t)\ dt + \int_x^{x+h}\right) - \left(\int_a^{a+h} g(t)\ dt + \int_{a+h}^{x} g(t)\ dt\right) \\ &= \int_{a+h}^{x+h} g(t)\ dt - \int_a^x g(t)\ dt,\end{aligned}$$

where we used that $(a+h,x) \sqcup (x,x+h) = (a+h,x+h)$ and $(a,x) = (a,a+h) \sqcup (a+h,x)$ are disjoint unions and g is integrable so that the indefinite integral $\Lambda(E) = \int_E g(t)\ dt$ is *completely additive* on (a,b). This motivates equation 7.0.4. Moving on from said equation: If we take the limit as $h \to 0$ we find that

$$\lim_{h\to 0^+} \frac{1}{h}\int_x^{x+h} g(t)\ dt \to g(x)$$

by continuity of g. By using that the indefinite integral $\Lambda_{|f|}(E)$ of $|f|$ is absolutely continuous for measurable $E \subset (a,b)$ one finds that

$$\lim_{h\to 0^+} \frac{1}{h}\int_a^{a+h} g(t)\ dt = g(a).$$

Together this gives us that

$$\int_a^x g'(t)\ dt = g(x) - g(a).$$

But $g(a) = 0$, i.e. we have that

$$\begin{aligned}\int_a^x g'(t)\ dt - g(x) &= \int_a^x g'(t) - \int_a^x f(t)\ dt \\ &= \int_a^x (g'(t) - f(t))\ dt \\ &= 0.\end{aligned}$$

By 7.0.5 it follows that $g'(t) = f(t)$ a.e., which is what we wanted to show. $\square$

Example 7.0.13. Let $D = \{(x,y) \in \mathbb{R}^2\ :\ 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$. Then

$$\begin{aligned}\iint_D (x^2+y^2) \underbrace{dxdy}_{\text{product measure}} &= \int_0^1\int_0^1 (x^2+y^2)\ dxdy \\ &= \int_0^1\int_0^1 (x^2+y^2)\ dxdy,\end{aligned}$$

where the last two integrals are *iterated integrals*.

7.1 Product measures

Let $(\underline{\overline{X}}, \mathcal{A}, \mu)$ and $(\underline{\overline{Y}}, \beta, \nu)$ be σ-finite (most common spaces) measure spaces, and consider the cartesian product $\underline{\overline{X}} \times \underline{\overline{Y}} = \{(x, y) \,:\, x \in \underline{\overline{X}}, y \in \underline{\overline{Y}}\}$. We want to define a measure on this.

Definition 7.1.1. $\mathcal{A} \times \beta$ is the σ-algebra *generated* by the sets $A \times B$ such that $A \in \mathcal{A}$ and $B \in \beta$.

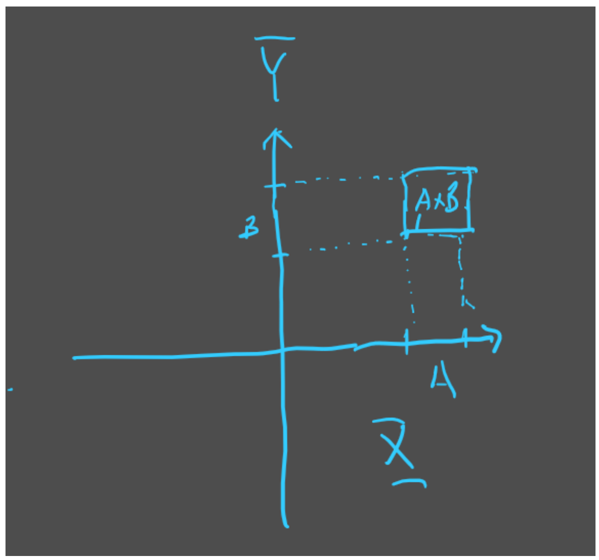

Figure 7.2: Illustration of $A \times B$

Remark 7.1.2. Generation: Smallest σ-algebra *containing* these sets.

Remark 7.1.3. We want a measure λ on $\mathcal{A} \times \beta$ such that $\lambda(A \times B) = \mu(A) \cdot \nu(B)$ for all $A \in \mathcal{A}, B \in \beta$.

The claim is that this will become the *unique* σ-finite measure on $\mathcal{A} \times \beta$.

Definition 7.1.4. If $E \subset \underline{\overline{X}} \times \underline{\overline{Y}}$ then we call any set on the form $E_x := \{y \,:\, (x, y) \in E\}$ an $\underline{\overline{X}}$**-section**, and similarly we call any set on the form $E^y := \{x \,:\, (x, y) \in E\}$ an $\underline{\overline{Y}}$**-section**.

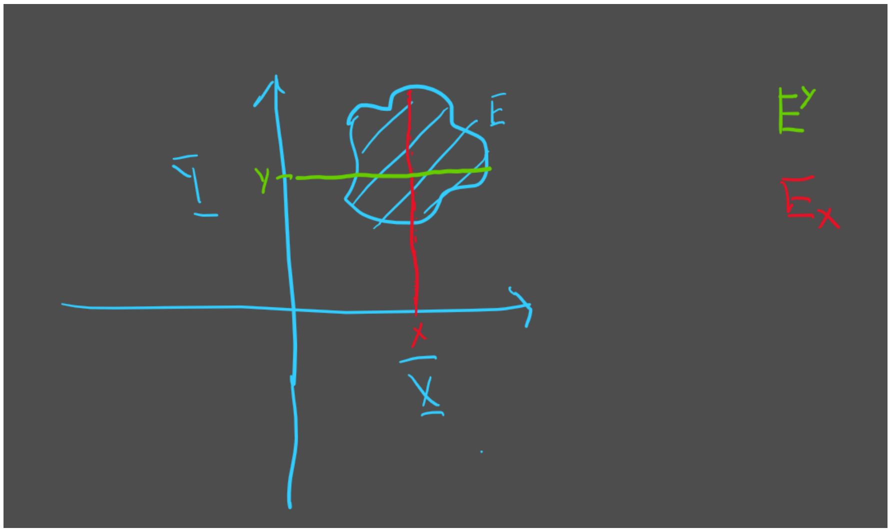

Figure 7.3: Illustration of E_x and E^y

Lemma 7.1.5. *If $E \in \mathcal{A} \times \beta$ then every $\underline{\overline{X}}$-section $E_x \in \beta$ and every $\underline{\overline{Y}}$-section $E^y \in \mathcal{A}$.*

Proof. See book. ☐

7.1.1 Main theorems

Theorem 7.1.6 (Theorem 1)**.** *Let $(\underline{\overline{X}}, \mathcal{A}, \mu)$ and $(\underline{\overline{Y}}, \beta, \nu)$ be σ-finite measure spaces. Assume that $E \in \mathcal{A} \times \beta$ ($\mathcal{A} \times \beta$ "smallest") and let*

$$f(x) = \nu(E_x) \qquad \forall x \in \underline{\overline{X}}$$

for and let

$$g(y) = \mu(E^y) \qquad \forall y \in \underline{\overline{Y}}.$$

Then f and g are measurable and takes values in $[0, \infty]$ (or say, non-negative) and

$$\int f \, d\mu = \int g \, d\nu.$$

Remark 7.1.7. Note that by lemma 7.1.5 we have that E_x, E^y are measurable.

Theorem 7.1.8 (Theorem 2)**.** *Let*

$$\begin{aligned} \lambda(E) &= \int \underbrace{\nu(E_x)}_{f(x)} \, d\mu \\ &= \int \underbrace{\mu(E^y)}_{g(y)} \, d\nu. \end{aligned}$$

Then λ is a σ-finite measure on $\mathcal{A} \times \beta$ such that

$$\lambda(A \times B) = \mu(A) \cdot \nu(B) \qquad \text{for all } A \in \mathcal{A} \text{ and } B \in \beta,$$

and λ is unique.

Remark 7.1.9. If we don't assume σ-finite then the claim is that one can make several examples - i.e. we lose *uniqueness* of λ.

Definition 7.1.10. λ in 7.1.8 is called **the product measure** of μ and ν and we write $\lambda = \mu \times \nu$.

7.1.2 Fubinis theorem (2.16)

Assume that $h : \overline{\underline{X}} \times \overline{\underline{Y}} \to \mathbb{R}$ is measurable. For $x \in \overline{\underline{X}}$ we let $h_x : \overline{\underline{Y}} \to \mathbb{R}$ be given by $h_x(y) = h(x, y)$. For $y \in \overline{\underline{Y}}$ let $h^y : \overline{\underline{X}} \to \mathbb{R}$ be given by $h^y(x) = h(x, y)$.

Lemma 7.1.11. *h_x is measurable in $(\overline{\underline{Y}}, \beta, \nu)$ and h^y is measurable in $(\overline{\underline{X}}, \mathcal{A}, \mu)$.*

Theorem 7.1.12 (Fubini's theorem)**.** *Assume that $h : \overline{\underline{X}} \times \overline{\underline{Y}} \to [0, \infty]$ is measurable (with respect to the product measure $\lambda = \mu \times \nu$ 7.1.10). Then*

$$\begin{aligned}
\int h \, d(\mu \times \nu) &= \int \left(\int h^y \, d\mu \right) \, d\nu \\
&= \int \left(\int h_x \, d\nu \right) \, d\mu.
\end{aligned}$$

Chapter 8

Lecture 8: Functional Analysis

In functional analysis, "vectors" are functions $v : X \to F$ where maybe X has a structure too. "scalars" F are usually $F = \mathbb{R}$ or $F = \mathbb{C}$. Maybe $\mathbb{Q}_p$, the p-adic numbers or $\mathbb{F}_\alpha$, boolean.

Compare this with Linear Algebra, where we have "vectors" $\vec{0}, u, v, u - v$ and "scalars" from any field F.

- For $\gamma \in F$ and v vector, also $\gamma \cdot v$ is a vector, and $0 \cdot v = \vec{0} = \gamma \cdot \vec{0}$, we also have $(\gamma + t) \cdot v = \gamma v + tv$ and $\gamma(v + w) = \gamma v + \gamma w$.

Remark 8.0.1. There is also something about "dimension of ten or something to control" with what looks like ∞ somewhere on the board (we were not present at this particular time).

<u>Goals</u>:

- Linear transformations
- Eigenvalues/Eigenvectors
- Orthonormal basis.

<u>Applications</u>:

- math of quantum.
- "Kernel methods" (ML).
- "Big Data" (High dimension).

8.0.1 Several kinds of spaces

We have **Hillbert spaces**, which are complete inner product spaces with lengths and angles (& completeness). Usually (?) $\mathbb{R}^n$ with norm

$$\begin{aligned} \|x\| &= \sqrt{\langle x, x \rangle} \\ &= \sqrt{\sum_i x_i^2} \end{aligned}$$

where

$$\begin{aligned}\langle x, y\rangle &= \sum_j x_j y_j \\ &= \|x\|\|y\| \cos\theta.\end{aligned}$$

is the inner (real) product.

We also have **Banach Spaces**, which is a "complete" normed space, e.g. $\mathbb{R}^n$ with

$$\|x\|_p := \left(\sum_j |x_j|^p\right)^{\frac{1}{p}}$$

which is the p-norm for $1 \leqslant p < \infty$ (or perhaps $p = \infty$ as well).

Furthermore, we have **metric spaces**, with distances between points, but not lengths of vectors, e.g. $X \subset \mathbb{R}^n$ not vector subspace, and usually distance $d(x, y) = \|x - y\|$.

We also have **topological spaces** with "nearness without distance". E.g. $X = \mathbb{R}^{\mathbb{R}}$, i.e. all functions $f : \mathbb{R} \to \mathbb{R}$, where a sequence (f_n) is near if $f_n(x) \to f(x)$ for all $x \in \mathbb{R}$.

Second part currently missing.

Chapter 9

Lecture 9

<u>Last time</u>: Defined topological and metric spaces, "completeness".

Definition 9.0.1 (Closed set)**.** In a topological space X, a subset $S \subseteq X$ is **closed** if its complement $X \setminus S$ is open.

Remark 9.0.2. $S^c = X \setminus S$ is simpler (notation-wise) if it is understood from the context that we are working inside X.

Theorem 9.0.3 (De Morgan's laws)**.**

- $$\left(\bigcup_{i \in I} S_i\right)^c = \bigcap_{i \in I} S_i^c.$$

- $$\left(\bigcap_{i \in I} S_i\right)^c = \bigcup_{i \in I} S_i^c.$$

<u>Open sets</u>: For a topological space X we have the following:

(1) $\varnothing, X$ are open.

(2) If V_i are open for $i \in I$, where I is an arbitrary index-set, then

$$\bigcup_{i \in I} V_i$$

is open.

(3) If I is a **finite** index-set and V_i is open for all $i \in I$, then

$$\bigcap_{i \in I} V_i$$

is open.

<u>Closed set</u>: The following statement holds with respect to a topological space X:

(1) $X = X \setminus \varnothing$ and $\varnothing = X \setminus X$ are closed.

(2) If A_i is closed for elements $i \in I$, where I is an **arbitrary** index-set, then

$$\bigcap_{i\in I} V_i$$

is closed.

(3) If I is a **finite** index-set, and V_i for elements i in I are closed, then

$$\bigcup_{i\in I} V_i$$

is closed.

Interior and closure: For a subset S of a topological space X, we define

- The **interior** of S as

$$\begin{aligned} S^\circ &= \operatorname{int}(S) \\ &= \bigcup_{\substack{V\subseteq X \text{ open} \\ V\subseteq S}} V. \end{aligned}$$

- The **closure** of S as

$$\begin{aligned} \overline{S} &= \operatorname{cl}_X(S) \\ &= \bigcap_{\substack{F\subseteq X \text{ closed} \\ F\supseteq S}} F. \end{aligned}$$

S° is open (it is a union of open sets) and $\overline{S}$ is closed (it is an intersection of closed sets).

Remark 9.0.4. If it is understood from the context that we are taking the closure in X, then one may write $\operatorname{cl}(S)$ for $\operatorname{cl}_X(S)$.

Definition 9.0.5 (Dense)**.** Given a topological space X, we say that a subset $S \subseteq X$ is **dense** in X if $\operatorname{cl}(S) = X$.

Example 9.0.6. $\mathbb{Q}$ is dense in $\mathbb{R}$.

Example 9.0.7. $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$.

Example 9.0.8 (non-example I)**.** $[0, 1]$ is not dense in $\mathbb{R}$.

Example 9.0.9 (non-example II)**.** $\mathbb{R}$ is not dense in $\mathbb{C}$.

Question from audience: Is X dense in X? Yes! We have that $\operatorname{cl}_X(X) = X$.

Definition 9.0.10 (Continuous function)**.** A function $f : X \to Y$ between topological spaces X and Y is **continuous** if for every open subset $V \subseteq Y$ we have that the preimage $f^{-1}(V) = \{x \in X : f(x) \in V\}$ is open in X.

Equivalently, f is continuous if for every closed $U \subseteq Y$ we have that $f^{-1}(U)$ is closed.

In metric spaces: We have "wiggle room". $V \subseteq X$ is open if for all $x \in V$ there is a positive real number $r > 0$ such that for all $y \in X$, if $\rho(x, y) < r$ then $y \in V$ (where ρ is the distance/metric of our space, say (X, ρ)).

$F \subseteq X$ is closed if F is "closed" under limits of sequences, i.e. if (x_n) is a sequence in F (i.e. $(x_n) \subset F$) such that $(x_n) \to x$ as n goes to ∞, then $x \in F$.

$S \subseteq X$ is dense if for all $x \in X$ and all $\varepsilon > 0$ there exist an $y \in S$ with $\rho(y, x) < \varepsilon$.

A map $f : (X, \rho_X) \to (Y, \rho_Y)$ between metric spaces is continuous at $x_1 \in X$ if for all $\varepsilon > 0$ there exist a $\delta > 0$ such that if $\rho_X(x_1, x) < \delta$ then $\rho_Y(f(x_1), f(x)) < \varepsilon$. If f is continuous at every point $x_1 \in X$ then we say that f is continuous.

Definition 9.0.11 (Isometry)**.** Given metric spaces (X, ρ_X) and (Y, ρ_Y), a map $f : X \to Y$ is an **isometry** if it preserves the metric, i.e. for all $x_1, x_2 \in X$ we have that

$$\rho_Y(f(x_1), f(x_2)) = \rho_X(x_1, x_2).$$

Example 9.0.12 (Example of continuous function in a metric space)**.** Given $x_0 \in X$ for a metric space (X, ρ) we have that the map $x \mapsto \rho(x, x_0)$ is continuous.

Metric spaces are **Hausdorff**: For any pair of points $x_1 \neq x_2$ in X, there are open sets $V_1 \ni x_1$ and $V_2 \ni x_2$ such that $V_1 \cap V_2 = \varnothing$.

Example 9.0.13 (Non-Hausdorff space)**.** Take the real line with two copies of the origin.

Metric spaces are **normal** : We can separate disjoint closed sets by disjoint open sets.

Consequences:

- $\{x\}$ is closed for all $x \in X$ if (X, ρ) is a metric space.
- Limits in (X, ρ) are unique, i.e. if $(x_n) \to x$ and $(x_n) \to y$ then $x = y$.

9.0.1 Main theorems about completeness

(1) Existence and uniqueness about **completions**. If X is a metric space then there is a complete metric space $\hat{X}$ and an isometry $\iota : X \hookrightarrow \hat{X}$ (note that isometries are injective) with *dense* image $\iota(X)$ in $\hat{X}$, i.e. such that $\mathrm{cl}_{\hat{X}}(\iota(X)) = \hat{X}$.

(2) Given two such $(\hat{X}, \iota_1)$ and $(\hat{X}_2, \iota_2)$ you can "complete the triangle": There is a surjective isometry $r : \hat{X}_1 \twoheadrightarrow \hat{X}_2$:

$$\begin{tikzcd} X \arrow[r, hook, "\iota_1"'] \arrow[d, hook, "\iota_2"] & \hat{X}_1 \arrow[dl, dotted, two heads, "r"] \\ \hat{X}_2 & \end{tikzcd}$$

Since r is an isometry, it is also injective, so in fact we have an isomorphy between any two completions $\hat{X}_1$ and $\hat{X}_2$ of X.

9.0.2 Baire's Category Theorem

There are two versions of **Baire's category theorem**, one for open sets and one for closed sets.

Theorem 9.0.14 (Open version)**.** *If X is complete, and $V_n \subseteq X$ is open and dense, for $n = 1, 2, 3, \ldots$ then*

$$\bigcap_{n=1}^{\infty} V_n$$

is dense.

Theorem 9.0.15 (Closed version)**.** *Let X be a complete metric space and let $F_n \subseteq X$ be closed for $n = 1, 2, 3, \ldots$ such that $Int(F_n) = \varnothing$. Then*

$$X \setminus \left(\bigcup_{n=1}^{\infty} F_n \right)$$

is dense in X.

Remark 9.0.16. Notice that since F_n is closed this is the same as $\mathrm{Int}_X(\mathrm{cl}_X(F_n)) = \varnothing$, i.e. F_n is **nowhere dense**.

Example 9.0.17 (Why only countably many open)**.** Let $X = \mathbb{R}$ and for each $x \in \mathbb{R}$, let $V_x = \mathbb{R} \setminus \{x\}$. This is open since $\{x\}$ is closed in $\mathbb{R}$. Note that $x + \frac{1}{n} \to x$ as $n \to \infty$ but $x + \frac{1}{n}$ is not in $\{x\}$ for any $n = 1, 2, 3, \ldots$, hence $\overline{V_x} = \mathbb{R}$ for any $x \in \mathbb{R}$ and so V_x is dense for all $x \in \mathbb{R}$.

We have that

$$\bigcap_{x \in \mathbb{R}} V_x = \varnothing,$$

which is not dense in $\mathbb{R}$.

Example 9.0.18 (Why complete)**.** Let $X = \mathbb{Q}$ (with the euclidian subspace topology induced from $\mathbb{R}$) and enumerate $\mathbb{Q}$ as $\mathbb{Q} = \{r_n\}_{n=1}^{\infty}$. Let $V_n = \mathbb{Q} \setminus \{r_n\}$. Then again for any $r \in \mathbb{Q}$ we have that $r + \frac{1}{n} \notin V_n$ but the limit goes to r so that $r_n \in \overline{V_n}$, hence $\overline{V}_n = \mathbb{Q}$. Again we have that

$$\bigcap_{n=1}^{\infty} V_n = \varnothing$$

which is not dense in $\mathbb{Q}$.

Example 9.0.19. Consider the metric

$$\rho(x, y) = \begin{cases} 1, & \text{if } x \neq y \\ 0, & \text{if } x = y \end{cases}.$$

Then Cauchy-sequences are in one-to-one bijection onto eventually constant sequences, i.e. convergent sequences, so any set X with this metric is complete. But generally, most complete metric spaces are *infinite* (this illustrates that not all complete metric spaces need to be infinite; we can take X to be finite in this example).

We prove 9.0.14:

Proof. Let X be a complete metric space (without loss of generality, assume $X \neq \varnothing$), and let $V_n \subseteq X$ be open and dense for $n = 1, 2, 3, \ldots$. To show: For any open $W \neq \varnothing$ in X, we have that

$$W \cap \left(\bigcap_{n=1}^{\infty} V_n \right) \neq \varnothing.$$

We proceed by induction, by finding a sequence of points $(x_n) \subset X$ and a sequence of radii $r_n > 0$ such that $\overline{B}(x_n, r_n) \subseteq V_n \cap B(x_{n-1}, r_{n-1})$ with $r_n \xrightarrow{n \to \infty} 0$.

<u>Base case</u>: V_1 meets W by denseness of V_1, so there is some $x_1 \in W \cap V_1$. Since W and V_1 are open so is their intersection, hence there is some ball $B(x_1, \varepsilon_1)$ contained in $W \cap V_1$. Let $r_1 = \frac{\varepsilon_1}{2}$. Then we see that $\overline{B}(x_1, r_1) \subseteq B(x_1, \varepsilon_1) \subseteq V_1 \cap W$.

Since V_2 is open dense and $B(x_1, r_1)$ is open we have that $V_2 \cap B(x_1, r_1)$ is open and non-empty. Hence there is some $x_2 \in B(x_1, r_1) \cap V_2$ and ε_2 such that $B(x_2, \varepsilon_2) \subseteq V_2 \cap B(x_1, r_1)$. Let $r_2 = \min\left\{\frac{\varepsilon_2}{2}, \frac{r_{n-1}}{2}\right\}$ so that $\overline{B}(x_2, r_2) \subseteq B(x_2, \varepsilon_2) \subseteq B(x_1, r_1) \cap V_2$. Inductively, we get points x_n in $W \cap (V_1 \cap \ldots \cap V_n)$. Assume by induction that we have constructed points $x_1, \ldots, x_{n-1}$ and radii $r_1, \ldots, r_{n-1}$ as above. Since V_n is dense there is a point $x_n \in V_n \cap B(x_{n-1}, r_{n-1})$. By openness there is some $\varepsilon_n > 0$ such that

$$B(x_n, \varepsilon_n) \subseteq V_n \cap B(x_{n-1}, r_{n-1}).$$

Taking $r_n = \min\left\{\frac{\varepsilon_n}{2}, \frac{r_{n-1}}{2}\right\}$ we find that

$$\overline{B}(x_n, r_n) \subseteq B(x_n, \varepsilon_n) \subseteq V_n \cap B(x_{n-1}, r_{n-1}).$$

<u>Claim</u>: (x_n) is a Cauchy sequence: If we take $i, j > n$ then both x_i and x_j are in $B(x_n, r_n)$ with diameter $2r_n < r_{n-1} < \ldots < \frac{r_1}{2^{n-2}}$. Hence as $n \to \infty$ we see that the diameter of $B(x_n, r_n)$ goes to $\frac{r_1}{2^{n-2}} \xrightarrow{n\to\infty} 0$. I.e. for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that if $i, j \geqslant N$ are of distance less than ε from eachother so that (x_n) is a Cauchy-sequence in X. Since X is complete, it follows that $\lim_{n\to\infty} x_n = x \in X$.

For any $n \geqslant 1$ we have that the sequence $(x_m)_{m \geqslant n}$ goes to x, so that $x \in \overline{B}(x_n, r_n) \subseteq V_n \cap B(x_{n-1}, r_{n-1})$ so that indeed $x \in V_n$ for all $n \geqslant 1$. Furthermore, $\overline{B}(x_1, r_1) \subseteq W \cap V_1$ so that $x \in W$, so indeed

$$x \in W \cap \left(\bigcap_{n=1}^{\infty} V_n\right),$$

which shows that

$$\bigcap_{n=1}^{\infty} V_n$$

is dense in X. □

<u>Baire's theorem for closed sets</u> Baire's theorem for closed sets (9.0.15) follow from Baire's theorem for open sets: If V_n is dense and open then $X \setminus V_n$ is closed and has empty interior. On the other hand, if F_n is closed and has empty interior, then $X \setminus F_n$ is open and since F_n has empty interior, it follows that $X \setminus F_n$ is dense.

We then see that by the open version of Baire's category theorem, we have that

$$\begin{aligned}\bigcap_{n=1}^{\infty} F_n^c &= \left(\bigcup_{n=1}^{\infty} F_n\right)^c \\ &= X \setminus \left(\bigcup_{n=1}^{\infty} F_n\right)\end{aligned}$$

is dense, where we used De Morgan's laws.

Baire's theorem is already interesting for $X = \mathbb{R}$. Recall Thomae's function:

$$\theta(x) = \begin{cases} \frac{1}{b}, & \text{if } x = \frac{a}{b} \in \mathbb{Q}, \\ 0, & \text{if } x \notin \mathbb{Q} \end{cases}.$$

Claim: Thomae's function is continuous at any $x \notin \mathbb{Q}$ and discontinuous at any $r \in \mathbb{Q}$.

Q: Is there a function $f : \mathbb{R} \to \mathbb{R}$ continuous on $\mathbb{Q}$ and discontinuous on $\mathbb{R} \setminus \mathbb{Q}$. The claim is that by Baire's theorem the answer to this question is negative.

Definition 9.0.20 (Nowhere dense)**.** A set $S \subseteq X$ is called **nowhere dense** if $\mathrm{int}_X(\mathrm{cl}_X(S)) = \varnothing$.

Example 9.0.21. $S = \{x\}$

Definition 9.0.22 (1^{st} and 2^{nd} Category)**.** A topological space if of $1^{\textbf{st}}$ **category** if it is a union of *countably* many nowhere dense sets, also called **meager**. Otherwise, it is called of $2^{\textbf{nd}}$ **category**.

If $X \setminus S$ is meager then S is called **residual**.

Remark 9.0.23. $X \setminus S$ is analogous to null-sets in measure-theory, if $X \setminus S$ is meager (not the same thing, though).

Remark 9.0.24. One way people often state Baire's category theorem is that if $X \neq \varnothing$ is complete, then X is not meagre.

9.0.3 Completions

- The proof in the coursebook is by equivalence-classes of Cauchy-sequences.
- Another proof is by contemplating an injection from X to a space of bounded functions $X \to \mathbb{R}$, called $\mathcal{B}(X, \mathbb{R})$, call this injection $\iota : X \hookrightarrow \mathcal{B}(X, \mathbb{R})$, defined by $\iota(x)(y) = \rho(x, y) - \rho(x_0, y)$ for some reference-point $x_0 \in X$. The claim is that this makes $\iota(x)$ bounded: By the triangle-inequality we have that

$$\begin{aligned} &\rho(x, y) \leqslant \rho(x, x_0) + \rho(x_0, y) \\ \Leftrightarrow\ &\rho(x, y) - \rho(x_0, y) \leqslant \rho(x, x_0), \end{aligned}$$

  and

$$\begin{aligned} &\rho(x_0, y) \leqslant \rho(x_0, x) + \rho(x, y) \\ \Leftrightarrow\ &\rho(x_0, y) - \rho(x, y) \leqslant \rho(x_0, x) \end{aligned}$$

  so that

$$|\rho(x, y) - \rho(x_0, y)| \leqslant \rho(x, x_0)$$

  so indeed bounded by $\rho(x, x_0)$ for all $y \in X$.

To check: That $\mathcal{B}(X, \mathbb{R})$ is complete, that $X \overset{\iota}{\hookrightarrow} \iota(X)$ is an <u>isometry</u> unto its image, and that $\iota(X)$ is dense in suitable subset: The claim is that $\iota(X)$ is dense in $\overline{\iota(X)}$ (this is clear by definition; any set S is dense in $\overline{S}$).

<u>Fact</u>: If S is complete and $S \subseteq X$ is closed, then S is also complete.

Proof. Let (x_n) be a Cauchy-sequence in S. Then there is some $x \in X$ such that $(x_n) \xrightarrow{n \to \infty} x \in X$. But $x \in S$ since S is closed.

Using this, with the assumption (not proven) that $\mathcal{B}(X, \mathbb{R})$ is indeed complete, we find that also $\mathrm{cl}_{\mathcal{B}(X,\mathbb{R})}(\iota(X))$ is complete. $\square$

Chapter 10

Lecture 10: “Completeness meets linearity”

We are starting chapter 4 on Friedman. There are however some things in chapter 3 not covered during the lectures (so far) that one needs to know:

- L^p-spaces $\hookrightarrow$ exercises. In particular, it is important that L^p is a complete metric space for $p \geqslant 1$ (note however that this is not true however for e.g. $\ell^p = L(\mathbb{N})^p$ [with counting measure] where $0 < p < 1$ [our comment]).
- [Fri03, Chap. 3.5-3.6] - compactness.
- [Fri03, Chap. 3.7 - 3.8].

Last time, we proved Baire's theorem. There are three consequences of Baire's theorem:

1. Banach-Steinhaus theorem (“Uniform-boundedness theorem”).
2. Open mapping theorem.
3. Closed graph theorem (“checking whether a map is continuous”).

Remark 10.0.1 (Something from last time). We want to point out that $\overline{B}(x, r)$ is just notation for the *closed* ball $\{y \in X \ : \ \rho(x, y) \leqslant r\}$ but <u>not necessarily closure</u> of the open ball $B(x, r)$, i.e. $\overline{B(x, r)} \neq \overline{B}(x, r)$, in general. For <u>example</u>, we we take any set X with the discrete metric, then $\overline{B}(x, r)$ for $r = \frac{1}{2}$ equals $\{x\}$, while $\overline{B(x, r)} = X$.

Remark 10.0.2. Zorn's lemma $\Leftrightarrow$ Axiom of Choice. Observe that for Baire's theorem, only the weaker countable choice is needed (i.e. every countable collection of non-empty sets has a choice function), we believe. This is, as far as we can tell, because we are only working with countable collections in the proof.

<u>Recall from linear algebra</u>:

- Linear independence.
- Zorn's lemma.
- Spanning, basis, dimension.
- As far as we can tell, a definition of a vector-space V over some field F being infinite-dimensional, is that *no finite set spans* the whole set V of vectors.

- Every vector space has a basis (assuming Zorn's lemma).

Recall the definition of a linear transformation: If X, Y are vector spaces and $T : X \to Y$ is map, then T is a **linear transformation** if

$$T(ax + by) = aT(x) + bT(y)$$

for all $x, y \in X$ and all $a, b \in F$, where F is a field over which X, Y are vector spaces.

New twist: *Partially* defined operators.

Let $D_T \subseteq X$ be a linear subspace. Then $T : D_T \to Y$ is not necessarily defined on all of X.

Example 10.0.3 (Partially defined operators)**.** $T = \frac{d}{dx}$ not defined on all of $\mathscr{C}[0,1] := \{f : [0,1] \to \mathbb{R} \,:\, f \text{ continuous function}\}$.

Definition 10.0.4 (Banach Space)**.** A Banach space $(X, \|\cdot\|)$ is a normed space such that the resulting metric is a complete metric space. Observe that X here is a vector space over some field F, usually $F = \mathbb{R}$ or $F = \mathbb{C}$, and norm $\|\cdot\| : X \to F$. This norm in turn induces a metric $\rho(x, y) := \|x - y\|$. That this is a metric is quite immediate from the definition of $\|\cdot\|$ being a norm (in particular, it satisfies the triangle-inequality $\|x + y\| \leqslant \|x\| + \|y\|$).

Theorem 10.0.5. *If X is a banach space such that $x_n \in X$ for $n = 1, 2, 3, \ldots$ and $\sum_{n=1}^{\infty} \|x_n\|$ converges in $\mathbb{R}$, then*

$$\sum_{n=1}^{\infty} x_n$$

converges in X.

Proof. Consider the partial sums $S_m := \sum_{n=1}^{m} x_n$.

Claim: (S_m) is Cauchy in X.

Proof. Let $m > \ell$. Then $S_m - S_\ell = \sum_{n=\ell+1}^{m} x_n$. Therefore, we have that

$$\begin{aligned} \|\|S_m - S_\ell\| &= \left\| \sum_{n=\ell+1}^{m} x_n \right\| \\ &\leqslant \sum_{n=\ell+1}^{m} \|x_n\| \xrightarrow{m,\ell\to\infty} 0. \end{aligned}$$

Here we used that $\sum_{n=1}^{\infty} \|x\| = S < \infty$ converges: For any $\varepsilon > 0$ there exist an $N \in \mathbb{N}_{\geqslant 1}$ such that if $\ell \geqslant N$ then

$$\begin{aligned} \sum_{n=\ell+1}^{m} \|x_n\| \leqslant S - S_\ell& \\ &= \sum_{n=\ell+1}^{\infty} \|x_n\| \\ &< \varepsilon \end{aligned}$$

(since the sum consist of only non-negative elements). ☐

Since (S_m) is Cauchy, and X is complete, (S_m) converges in X. ☐

Example 10.0.6. Let $X = \mathbb{R}$ or $X = \mathbb{C}$. Then absolute convergence implies convergence.

Example 10.0.7. Let $X = \mathscr{C}[0,1]$ with supremum-norm $\rightsquigarrow$ Weierstrass M-test.

10.0.1 Operator norms

Definition 10.0.8 (Operator norm $\|T\|_{\text{op}}$)**.** Let $T : X \to Y$ be linear map between normed vector spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ over $F = \mathbb{R}$ or $\mathbb{C}$. Then we define the **operator norm** $\|T\|_{\text{op}} : \mathcal{B}(X, Y) \to \mathbb{R}$ of T, as

$$\|T\|_{\text{op}} = \sup \left\{ \frac{\|Tx\|_Y}{\|x\|_X} \ : \ 0 \neq x \in X \right\}.$$

Observe that

$$\begin{aligned} &\frac{\|Tx\|_Y}{\|x\|_X} \leqslant \|T\|_{\text{op}} \\ \Leftrightarrow\ &\|Tx\|_Y \leqslant \|T\|_{\text{op}} \cdot \|x\|_X \qquad \text{for all non-zero } x \in X. \end{aligned}$$

In particular, if $C := \|T\|_{\text{op}} < \infty$ then

$$\|Tx\|_Y \leqslant C \cdot \|x\|_X \qquad \forall x \in X,$$

for some constant $C \in \mathbb{R}_{\geqslant 0}$.

If $(Y, \|\cdot\|)$ is complete, then $\|\cdot\|_{\text{op}}$ is complete on $\mathcal{B}(X, Y)$.

Theorem 10.0.9. *Given normed spaces X, Y and linear map $T : X \to Y$, the following are equivalent:*

1) *T is continuous on X.*
2) *T is continuous at a point of X.*
3) *$\|T\|_{op} < \infty$.*

Remark 10.0.10. Observe that if there is no constant $C \in \mathbb{R}_{\geqslant 0}$ such that $\|Tx\|_Y \leqslant C \cdot \|x\|_X$ for all $x \in X$, then $\|T\|_{\text{op}} = \infty$ (this is straightforward from taking the contrapositive of this statement together with the definition of the operator norm).

Proof. (1)⇒(2): Immediate.

(3)⇒(1): We have the inequality

$$\begin{aligned} \|Tp - Tq\| &= \|T(p - q)\| \\ &\leqslant \|T\|_{\text{op}} \cdot \|p - q\|_X. \end{aligned}$$

Hence, let $\varepsilon > 0$ be arbitrary. If $\|p - q\|_X < \frac{\varepsilon}{\|T\|_{\text{op}}}$ (assuming that $\|T\|_{\text{op}} \neq 0$, i.e. it seems to us this is equivalent to assuming that $T \neq 0$) we see that $\|Tp - Tq\| < \varepsilon$, so that $\delta = \frac{\varepsilon}{\|T\|_{\text{op}}}$ exists for each $\varepsilon > 0$. If $T = 0$ then let $\delta = \varepsilon$. Observe here also that we mean continuous with respect to the induced metric topologies $\rho_X(x, y) = \|x - y\|$ with ρ_Y defined similarly, i.e. the metric topologies induced from the norms on X and Y.

(2) ⇒(3): Suppose that T is continuous at a point x. For any $\varepsilon > 0$ there exists by assumption $\delta > 0$ such that if $\|x - x_0\|_X \leqslant \delta$ then $\|Tx - Ty\| \leqslant \varepsilon$.

Claim: $\|T\|_{\text{op}} \leqslant \frac{\varepsilon}{\delta}$.

Proof. If $\|(x + x_0) - x_0\| = \|x\| \leqslant \delta$ then $\|T(x_0 + x) - Tx_0\| \leqslant \varepsilon$. Put $y = x - x_0$. Then

$$\begin{aligned} \|Ty\| &= \frac{\|y\|}{\delta} \left\| T\left(\frac{\delta y}{\|y\|} \right) \right\| \\ &\leqslant \frac{\|y\|}{\delta} \cdot \varepsilon, \end{aligned}$$

where we used that $\frac{y}{\|y\|}$ has norm 1, so that $\left\|\frac{y}{\|y\|} \cdot \delta\right\| = \delta$. □

We see that $\|Ty\| \leqslant C \cdot \|y\|$ with $C = \frac{\varepsilon}{\delta}$. Since any element $y \in X$ can be written as $x - x_0$ with $x = (y + x_0)$, this becomes an upper bound for $\|Ty\|$ for all $y \in X$, i.e.

$$\frac{\|Ty\|}{\|y\|} \leqslant C$$

for $y \neq 0 \in X$. Hence C is an upper bound to the set

$$\left\{\frac{\|Tx\|}{\|x\|} \ : \ 0 \neq x \in X\right\}$$

and so by definition $\|T\|_{\text{op}}$ as the least upper bound of this set we find that $\|T\|_{\text{op}} \leqslant C < \infty$. □

Theorem 10.0.11. *$\|\cdot\|_{op}$ is a norm on $\mathcal{B}(X,Y) = \{T : X \to Y \ : \ T$ continuous and linear$\}$.*

Remark 10.0.12. Completeness in the image (Y) $\rightsquigarrow$ completeness of functions into that space also follows (?).

Remark 10.0.13. If $X = Y$, then there is *extra structure*:

$$X \xrightarrow{S} X \xrightarrow{T} X$$

Then the composition $T \circ S$ is such that $\|T \circ S\| \leqslant \|T\| \cdot \|S\|$ for the operator-norm. If X is also complete, then $\mathcal{B}(X,X)$ becomes a Banach-algebra A in the sense of a vector space over $F = \mathbb{R}$ or $\mathscr{F} = \mathbb{C}$ with a binary map $A \times A \to A$ that is linear in each argument (i.e. bilinear) and associative.

Baire's theorem implies that if $X \neq \varnothing$ is complete then X is not meager, i.e. can not be written as a countable union of nowhere dense sets, where we again recall that nowhere dense (in some ambient space X) means that $\text{int}(\text{cl}(F)) = \varnothing$. (in particular, if F is already closed then it has empty interior whenever it is nowhere dense).

Theorem 10.0.14 (Banach-Steinhaus theorem ("Uniform boundedness principle"))**.** *If X is a Banach space, and Y is a normed space, $\mathscr{C} \subset \mathcal{B}(X,Y)$, and $\sup_{T \in \mathscr{C}} \|Tx\|_Y < \infty$ for all x in some non-meager subset of X, then $\sup_{T \in \mathscr{C}} \|T\|_{op} < \infty$.*

Remark 10.0.15. Note how the theorem gives us uniform boundedness from point-wise boundedness in some non-meager subset of X.

Remark 10.0.16. Observe that if $X \neq \varnothing$ is complete then we may take the non-meagre subset as X itself.

Proof. For $n = 1, 2, 3, \ldots$ let

$$\begin{aligned} F_n &:= \{x \in X \ : \ \sup_{T \in \mathscr{C}} \|Tx\| \leqslant n\} \\ &= \bigcap_{T \in \mathscr{C}} \{x \in X \ : \ \|Tx\| \leqslant n\}. \end{aligned}$$

We claim that $x \mapsto \|x\|$ is continuous (if $\|x - y\| < \delta := \varepsilon$ then $|\|x\| - \|y\|| < \varepsilon$ by the fact that $\|x\| = \|(x-y)+y\| \leqslant \|x-y\| + \|y\|$ and then doing the same thing with x and y "swapped"). Hence, the composition $f_T := \|\cdot\| \circ T$ is continuous for any $T \in \mathscr{C}$. We note that $f_T^{-1}([0,n]) = \{x \in X \ :$

$\|Tx\| \leqslant n\}$ (since the norm is always non-negative). But f is continuous and $[0, n] \subset \mathbb{R}$ is closed so its pre-image is closed. Since

$$F_n = \bigcap_{T \in \mathscr{C}} f_T^{-1}([0, n])$$

it follows that F_n is closed.

Set $E := \{x \in X \ : \ \sup_{T \in \mathscr{C}} \|Tx\| < \infty\}$. Then it is (almost) immediate that

$$E = \bigcup_{n=1}^{\infty} F_n.$$

If E was meagre then the hypothesis of the theorem would fail, hence E must be non-meagre. Therefore $\operatorname{int}(\operatorname{cl}(F_n)) = \operatorname{int}(F_n) \neq \varnothing$ for some F_n. Since the topology on F_n is the metric topology, that F_n has an interior point, say x_0, means that there is a ball $B(x_0, r')$ around x_0 contained in F_n. We may shrink r' to $r < r'$ so that $\overline{B}(x_0, r) \subset F_n$. Hence if $\rho_X(x_0, x) = \|x - x_0\| \leqslant r$ then $x \in F_n$. If $x \in X$ is such that $\|x\| < r$ then $\|(x + x_0) - x\| = \|x\| \leqslant r$ which implies that $x + x_0 \in F_n$, which in turn implies that $\sup_{T \in \mathscr{C}} \|T(x + x_0)\| \leqslant n$.

But then we see that

$$\begin{aligned} \|Tx\| &= \|T(x + x_0) - Tx_0\| \\ &\leqslant \|T(x + x_0)\| + \|Tx_0\| \\ &\leqslant n + n \\ &= 2n, \end{aligned}$$

where we used that $x_0 \in F_n$.

We then have that

$$\begin{aligned} \|Tx\| &= \frac{\|x\|}{r} \left\| T\left(\frac{x}{\|x\|} \cdot r\right) \right\| \\ &\leqslant \frac{\|x\|}{r} \cdot 2n, \end{aligned}$$

for any $0 \neq x \in X$ where we used that $\left\| \frac{x}{\|x\|} \cdot r \right\| = \frac{r}{\|x\|} \|x\| = r$ so that indeed

$$\left\| T\left(\frac{x}{\|x\|} \cdot r\right) \right\| \leqslant 2n.$$

Since $x \neq 0$ was arbitrary, we conclude that $\|T\|_{\text{op}} \leqslant \frac{2n}{r}$. Since T was arbitrary, we have that $\sup_{T \in \mathscr{C}} \|T\|_{\text{op}} \leqslant \frac{2n}{r}$ for any $x \neq 0$ in X. □

Theorem 10.0.17 (Frigyes Lemma). *If $Y \subsetneq X$ is a closed linear subspace of a normed space X, then for all $\varepsilon > 0$ there is a $z \in X$, with $\|z\| = 1$ and $\|z - y\| > 1 - \varepsilon$, for all $y \in Y$.*

Proof. Let $x_0 \in X - Y$ (possible since $Y \subsetneq X$).

Let $d := \inf_{y \in Y} \|x_0 - y\| > 0$. The reason $d > 0$ is that Y is closed. If it was zero then we could construct a sequence (y_n) in Y converging to x_0 (since the closure of Y, i.e. Y since Y is closed, equals Y together with points x such that there is a sequence in y converging to x whenever we are working with metric spaces), and then x_0 would have to be in Y, contradiction!

Since we are taking the infimum: For any $\eta > 0$ there exists $y_0 \in Y$ such that $d \leqslant \|x_0 - y_0\| \leqslant d + \eta$ since if not then $d + \eta$ would be a greater lower bound.

Put $z := \frac{x_0 - y_0}{\|x_0 - y_0\|}$ where y_0 is chosen such that its associated η is "small enough".

For all $y \in Y$ we have that

$$\begin{aligned} z - y &= \frac{x_0 - y_0}{\|x_0 - y_0\|} - y \\ &= \frac{x_0 - y_0 - \|x_0 - y_0\| y}{\|x_0 - y_0\|}. \end{aligned}$$

Therefore, we have that

$$\begin{aligned} \|z - y\| &= \frac{1}{\|x_0 - y_0\|} \|x_0 - (y_0 + \|x_0 - y_0\| y)\| \\ &\geqslant \frac{1}{d + \eta} \|x_0 - y'\| \qquad (y' \in Y) \\ &\geqslant \frac{d}{d + \eta}. \end{aligned}$$

It remains to choose η such that $\frac{d}{d+\eta} > 1 - \varepsilon$. If $\varepsilon \geqslant 1$ then this trivially holds. If $0 < \varepsilon < 1$ then choose η such that

$$\frac{d}{1 - \varepsilon} - d > \eta > 0.$$

□

Remark 10.0.18. Comment: "Compactness is rare in infinite-dimensional spaces."

Theorem 10.0.19. *If X is a normed space, $Y \subseteq X$ is a finite-dimensional linear subspace, then Y is closed.*

Proof. Pick a basis $E = \{e_1, \ldots, e_n\}$ for Y. Let $(y_m) \subset Y$ converge to $x \in X$. We want to show that $x \in Y$. Write y_m in the basis E for all $m \in \mathbb{N}_{\geqslant 1}$ as

$$y_m = \sum_{i=1}^{n} c_{i,m} e_i.$$

Then

1) Extract a subsequence.
2) Show that there exists $\lim_{m \to \infty} c_{i,m} = c_i$ for every i and that $x = \sum_i c_i e_i \in Y$ (since $Y = \operatorname{span}(E)$).

See [Fri03, Theorem 4.3.2] for further details. □

Appendix A

Lebesgue bounded convergence theorem

Theorem A.0.1. *Let (f_n) be a sequence of integrable functions that converges either in measure or a.e. to a measurable function f. If there is an integrable function g such that $|f_n(x)| \leqslant g(x)$ a.e., for all $n \geqslant 1$, then f is integrable, and*

$$\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0,$$

i.e. then (f_n) converges in the mean to f.

Proof. We divide the proof by cases, depending on whether (f_n) converges to f in measure or a.e.

(f_n) converges to f in measure: Assume that (f_n) converges to f in measure. We shall prove that (f_n) is a Cauchy sequence in the mean. Let

$$E = \bigcup_{n=1}^{\infty} \{x \, : \, f_n(x) \neq 0\}.$$

Lemma A.0.2 (Problem 2.6.2)**.** *If f is integrable, then the set $N(f) = \{x \, : \, f(x) \neq 0\}$ is σ-finite.*

Proof. See Kristian Berjklövs note on the canvas course-page ("note on sigma-finiteness"). □

This means that we can write $E = \bigcup_{n=1}^{\infty} A_n$ with $A_n \in \mathcal{A}$ and $\mu(A_n) < \infty$. If we set $E_k = \bigcup_{n=1}^{k} A_k$ then $\mu(E_k) \leqslant \sum_{n=1}^{k} \mu(A_k) < \infty$ and $E_k \subset E_{k+1}$. Then the limit exists and $\lim_n E_n = \bigcup_{n=1}^{\infty} E_n = E$. Let $F_k = E - E_k$. Then, by [Fri03, Theorem 2.7.1] we have

$$\begin{aligned} \int_{F_k} |f_n - f_m| \, d\mu &\leqslant \int_{F_k} (|f_n| + |f_m|) \, d\mu \\ &\leqslant 2 \int_{F_k} g \, d\mu. \end{aligned}$$

Remark A.0.3. Also recall f integrable $\Leftrightarrow$ $|f|$ integrable.

Since $E_k \subset E_{k+1}$ we have that $F_k = E - E_k \supset E - E_{k+1} = F_{k+1}$. Furthermore, we find that since (F_k) is a decreasing sequence of sets, its limit exists, and

$$\begin{aligned}\lim_n F_n &= \bigcap_{n=1}^{\infty} F_n \\ &= \bigcap_{n=1}^{\infty} (E - E_n) \\ &= E - \left(\bigcup_{n=1}^{\infty} E_n \right) \\ &= E - E \\ &= \varnothing.\end{aligned}$$

By [Fri03, Cor. 2.8.5] we have that (since f_n, f_m integrable $\rightsquigarrow |f_n - f_m|$ integrable; see earlier remark)

$$\lim_{n\to\infty} \int_{F_n} |f_n - f_m| \; d\mu = 0.$$

Therefore, for each $\eta > 0$, there is some $k_0 \in \mathbb{N}$ such that if $k \geqslant k_0$ then

$$\int_{F_k} |f_n - f_m| \; d\mu \leqslant \eta. \tag{A.0.1}$$

Remark A.0.4. Notice that k_0 depends on η *but not on* m, n.

For any $\varepsilon > 0$, we define the sets

$$G_{m,n} := \{x \;:\; |f_m(x) - f_n(x)| \geqslant \varepsilon\}.$$

We then have

$$\begin{aligned}\int_{E_k} |f_n - f_m| \; d\mu &= \int_{E_k - G_{m,n}} |f_n - f_m| \; d\mu + \int_{E_k \cap G_{m,n}} |f_n - f_m| \; d\mu \\ &\leqslant \varepsilon\mu(E_k - G_{m,n}) + 2\int_{E_k \cap G_{m,n}} g \; d\mu \\ &\leqslant \varepsilon\mu(E_k) + 2\int_{E_k \cap G_{m,n}} g \; d\mu. \end{aligned} \tag{A.0.2}$$

By assumption, (f_m) converges in measure to f. Therefore, $\mu(G_{n,m}) \to 0$ as $m, n \to \infty$. Since g is integrable, by [Fri03, Theorem 2.8.4] we have that $\lambda(E) = \int_E g \; d\mu$ is *absolutely continuous*. Therefore, for any $\varepsilon > 0$ there exists a $\delta > 0$ so that if $\mu(E) < \delta$ for measurable set E, then $|\lambda(E)| < \varepsilon$. Notice that $\mu(E_k \cap G_{m,n}) \leqslant \mu(G_{m,n}) \to 0$ as $m, n \to \infty$. Therefore, for any $\varepsilon > 0$, there is some $n_0 \in \mathbb{N}$ such that if $m, n \geqslant n_0$ then

$$|\lambda(E_k \cap G_{m,n})| = \left| \int_{E_k \cap G_{m,n}} g \; d\mu \right| < \varepsilon.$$

Notice that since $0 \leqslant |f_n| \leqslant g$ a.e., it follows that $g \geqslant 0$ a.e., so that $\int g \; d\mu \geqslant 0$, hence

$$\left| \int g \; d\mu \right| = \int g \; d\mu.$$

Take the measurable set $G_{m,n}$. We know that $g \geqslant 0$ a.e., i.e. $N = \mu(\{x \ : \ g(x) < 0\}) = 0$. Then $A_{m,n} := N \cap G_{m,n}$ has measure zero, and $g \geqslant 0$ holds on $G_{m,n} - A_{m,n}$, i.e. $g \geqslant 0$ a.e. holds on $G_{m,n}$. Therefore, we have that

$$\left| \int_{G_{m,n}} g \ d\mu \right| = \int_{G_{m,n}} g \ d\mu < \varepsilon$$

for $m, n \geqslant n_0$.

Therefore, choosing appropriate n_0, we have that

$$\begin{aligned} 2 \int_{E_k \cap G_{m,n}} g \ d\mu &\leqslant 2 \int_{G_{m,n}} g \ d\mu \\ &< \varepsilon. \end{aligned} \tag{A.0.3}$$

Remark A.0.5. Notice that n_0 depends only on g and ε, and *not* on m, n. Above we also used that $|f_m - f_n| \leqslant |f_m| + |f_n| \leqslant 2g$ a.e., so if we take $N = \{x \ : \ |f_n(x)| > g(x)\} \cup \{x \ : \ |f_m(x)| > g(x)\}$ which is a null-set, then on $E_k \cap G_{m,n} - N$ we have that $|f_n(x)| \leqslant g(x)$ and $|f_m(x)| \leqslant g(x)$, i.e. both of these inequalities hold a.e. on $E_k \cap G_{m,n}$.

Combining equations A.0.1, A.0.2 and A.0.3, we get that (using that F_k and E_k are *disjoint*)

Remark A.0.6. Notice that

$$\begin{aligned} X - E &= X - \left(\bigcup_{n=1}^{\infty} \{x \ : \ f_n(x) \neq 0\} \right) \\ &= \bigcap_{n=1}^{\infty} (X - \{x \ : \ f_n(x) \neq 0\}) \\ &= \bigcap_{n=1}^{\infty} \{x \in X \ : \ f_n(x) = 0\}. \end{aligned}$$

Therefore $f_n(x) = f_m(x) = 0$ for anything in $X - E$, and so $\int_{X \setminus E} |f_n - f_m| \ d\mu = 0$ so that the only interesting part of the integral (for *any* $m, n \geqslant 1$ happens on the E-part)

$$\begin{aligned} \int |f_n - f_m| \ d\mu &= \int_E |f_n - f_m| \ d\mu + \underbrace{\int_{X \setminus E} |f_n - f_m| \ d\mu}_{=0} \\ &= \int_{E = E_k \sqcup F_k} |f_n - f_m| \ d\mu \\ &= \int_{F_k} |f_n - f_m| \ d\mu + \int_{E_k} |f_n - f_m| \ d\mu \\ &= \int_{F_k} |f_n - f_m| \ d\mu + \left(\int_{E_k \cap G_{m,n}} |f_n - f_m| \ d\mu + \int_{E_k - G_{m,n}} |f_n - f_m| \ d\mu \right) \\ &\leqslant \eta + \varepsilon + \varepsilon \mu(E_k) \qquad \text{for } m, n \geqslant n_0 \text{ and } k \geqslant k_0. \end{aligned}$$

We set $S_{m,n} := \int |f_m - f_n| \ d\mu$ and $B_N = \sup_{m,n \geqslant N} S_{m,n}$ and let $m, n \to \infty$. Then note that (B_N) is a *decreasing* sequence that is bounded below by 0, and so by the *monotone convergence theorem* it converges to its infimum, that is, we have that

$$\lim_{N \to \infty} B_N = \inf_N B_N = \overline{\lim}_{m,n \to \infty} S_{m,n}.$$

We show that $\inf_N B_N = 0$: We want to show that for any $\delta > 0$ there exists some N such that $B_N < \delta$. Let $\eta \in (0, \delta)$ and take $k \geqslant k_0(\eta)$ so that $\int_{F_k} |f_n - f_m| \, d\mu < \eta$. For such a k fixed, we also have that $C := 1 + \mu(E_k)$ is fixed. Let $\varepsilon = \frac{\delta - \eta}{C} > 0$ (since $\eta < \delta$) with $n_0(\varepsilon)$. Then note that

$$\int_{E_k} |f_n - f_m| \, d\mu \leqslant \varepsilon C \quad \text{for } m, n \geqslant n_0(\varepsilon).$$

Together, this gives us that

$$\begin{aligned} S_{m,n} &< \eta + \varepsilon C \\ &= \eta + (\delta - \eta) \\ &= \delta, \end{aligned}$$

whenever $m, n \geqslant n_0(\varepsilon)$ and $k \geqslant k_0(\eta)$. Therefore, for $N \geqslant n_0(\varepsilon)$ we have that $B_N = \sup_{m,n \geqslant N} S_{m,n} < \delta$ since δ is an upper bound for $S_{m,n}$ for all $N \geqslant n_0(\varepsilon)$. Then note that no $\delta > 0$ can be a lower bound, i.e. we can not have a $\delta > 0$ so that $\delta \leqslant B_N$ for all $N \geqslant 1$. Since $B_N \geqslant 0$ for all $N \geqslant 1$, we have that 0 is a *lower bound* to the set $S := \{B_N \ : \ N \geqslant 1\}$. Therefore $\inf_N B_N = \inf S \geqslant 0$. Since $\inf S \leqslant s$ for all $s \in S$, we get that $\inf S \leqslant s < \delta$ for all $\delta > 0$. Hence $\inf S \leqslant 0$ (since $\delta > 0$ was arbitrary). By combining this, we see that $\inf S = 0$. Thus

$$\begin{aligned} \lim_{N \to \infty} B_N &= \inf_N B_N \\ &= \overline{\lim}_{m,n \to \infty} S_{m,n} \\ &= 0. \end{aligned}$$

Now take arbitrary $\varepsilon > 0$. Then there exists N_0 so that if $N \geqslant N_0$ we have that $B_N = \sup_{m,n \geqslant N} S_{m,n} < \varepsilon$. In particular, this means that

$$\int |f_m - f_n| \, d\mu < \varepsilon \quad \text{for all } m, n \geqslant N.$$

Since $\varepsilon > 0$ was arbitrary, it follows that

$$\int |f_m - f_n| \, d\mu = 0 \quad \text{as } m, n \to \infty,$$

i.e. that (f_n) is Cauchy in the mean.

By [Fri03, Theorem 2.8.3] there exists an integrable function h such that

$$\lim_{n \to \infty} \int |f_n - h| \, d\mu = 0.$$

By [Fri03, Theorem 2.7.2] we have that (f_n) converges to h in measure. By lemma 5.1.1 we have that $f = h$ a.e. Since by hypothesis, f is measurable, by [Fri03, Problem 2.6.1] we have that f is integrable and $\int f \, d\mu = \int h \, d\mu$. Therefore, we have that

$$\begin{aligned} \int |f_n - f| \, d\mu &\leqslant \int |f_n - h| \, d\mu + \underbrace{\int |f - h| \, d\mu}_{=0} \\ &= \int |f_n - h| \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

We have shown that f is integrable, and that $\lim_{n \to \infty} |f_n - f| \, d\mu \to 0$, which is what we wanted to show.

$\underline{f_n \to f \text{ a.e.}}$: Suppose that (f_n) converges to f a.e. If we can show that (f_n) converges to f in measure, then we are done, since the same reasoning as in the previous case then applies. Let

$$N := \{x \,:\, |f(x)| > g(x)\} \cup \left(\bigcup_{n=1}^{\infty} \{x \,:\, |f_n(x)| > g(x)\} \right).$$

Notice that by hypothesis and countable sub-additivity of μ, we have that $(\bigcup_{n=1}^{\infty}\{x \,:\, |f_n(x)| > g(x)\})$ has measure zero. Furthermore, since by hypothesis we have that $|f_n(x)| \leqslant g(x)$ a.e. for all $n \geqslant 1$ we also have that

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} \{x \,:\, |f_n(x)| > g(x)\} \right) &\leqslant \sum_{n=1}^{\infty} \underbrace{\mu(\{x \,:\, |f_n(x)| > g(x)\})}_{=0} \\ &= 0\end{aligned}$$

so that N has measure zero. Assume that $x \notin \left\{x \,:\, g(x) \geqslant \frac{\varepsilon}{2}\right\} \cup N$. Then $g(x) < \frac{\varepsilon}{2}$ and

$$\begin{aligned}|f_j(x) - f(x)| &\leqslant |f_j(x)| + |f(x)| \\ &\leqslant 2g(x) \\ &< \varepsilon\end{aligned}$$

so that if we define $E_n := \bigcup_{j=n}^{\infty}\{x \,:\, |f_j(x) - f(x)| \geqslant \varepsilon\}$ □

then $x \notin E_n$, for any $n \geqslant 1$. Hence $E_n \subset \left\{x \,:\, g(x) \geqslant \frac{\varepsilon}{2}\right\} \cup N$ for any $n \geqslant 1$.

Remark A.0.7. As far as we can tell, there is a typo in the proof here (when they cite problem 2.7.2; we don't see why this problem is relevant; perhaps we are mistaken).

g is integrable, and so by [Fri03, 2.7.h], since we have that $g(x) \geqslant \frac{\varepsilon}{2} > 0$ on $A = \left\{x \,:\, g(x) \geqslant \frac{\varepsilon}{2}\right\}$, it follows that $\mu(A) < \infty$. By hypothesis, we have that $\lim_{n\to\infty} f_n = f$ a.e.

Set $A_j := \{x \,:\, |f_j(x) - f(x)| \geqslant \varepsilon\}$ so that $E_n = \bigcup_{j \geqslant n} A_j$. Then $E_{n+1} \subset E_n$ is a decreasing sequence of sets so has a limit $\lim_n E_n = \bigcap_{n=1}^{\infty} E_n$. If $x \in \bigcap_{n=1}^{\infty} E_n$ then $x \lim_n f_n(x) \neq f(x)$ so that $x \in N_0$ for the set $N_0 = \{x \,:\, f(x) \not\to f(x)\}$, which has measure zero. Hence $\lim_n E_n \subset N_0$, so that

$$\begin{aligned}\mu(\lim_n E_n) &= \mu(N_0) \\ &= 0.\end{aligned}$$

Notice that $\mu(E_n) \leqslant \mu(A) < \infty$ so by [Fri03, Theorem 1.2.1(iv)] we have that

$$\begin{aligned}\lim_n \mu(E_n) &= \mu(\lim_n E_n) \\ &= \mu\left(\bigcap_{n=1}^{\infty} E_n \right) \\ &= 0.\end{aligned}$$

We have that $A_n \subset E_n \Rightarrow \mu(A_n) \leqslant \mu(E_n)$ so that $\lim_n \mu(A_n) \leqslant \lim_n \mu(E_n) = 0$. Thus, we conclude that (f_n) converges to f in measure.

Appendix B

Lebesgue monotone convergence theorem

Theorem B.0.1 (Lebesgue monotone convergence theorem)**.** *Let* (f_n) *be a monotone increasing sequence of non-negative integrable functions, and let* $\lim_n f_n(x) = f(x)$*. Then*

$$\lim_{n\to\infty} \int f_n \; d\mu = \int f \; d\mu.$$

Proof. If f is integrable, then for any $n \geqslant 1$ we have that $0 \leqslant f_n \leqslant f$ which by [Fri03, Theorem 2.7.1] implies that

$$\int f_n \; d\mu \leqslant \int f \; d\mu. \tag{B.0.1}$$

If f is not integrable, then we have the same inequality, since $\int f \; d\mu = \infty$ by convention, while f_n is integrable so that $\int f_n \; d\mu < \infty$. Hence either way, equation B.0.1 holds. By taking the limit as n goes to infinity, we obtain

$$\lim_{n\to\infty} \int f_n \; d\mu \leqslant \int f \; d\mu \tag{B.0.2}$$

It remains to show that if $\lim_{n\to\infty} \int f_n \; d\mu < \infty$ then f is integrable, and we have equality in equation B.0.2. Note that if $m \geqslant n$ then $f_m \geqslant f_n$ so that

$$\int |f_m - f_n| \; d\mu = \int f_m \; d\mu - \int f_n \; d\mu \to 0 \quad \text{if } m \geqslant n \to \infty$$

Another way to write this, that is perhaps clearer, is that by assumption $\lim_n \int f_n \; d\mu = L < \infty$. Then we may write

$$\begin{aligned}\int |f_m - f_n| \; d\mu &= \int f_m \; d\mu - \int f_n \; d\mu \\ &\leqslant L - \int f_n \; d\mu,\end{aligned}$$

now we may choose N large enough so that if $n \geqslant N$ we have that $L - \int f_n \; d\mu < \varepsilon$ for any $\varepsilon > 0$, as long as $m \geqslant n \geqslant N$. Notice that if $n > m \geqslant N$ then we would have that $\int |f_m - f_n| \; d\mu = \int |f_n - f_m| \; d\mu$ and then the same argument would work, with n and m interchanged. Hence we see that for $m, n \geqslant N$ we have that $\int |f_m - f_n| \; d\mu < \varepsilon$, for any $\varepsilon > 0$. Thus it follows that (f_m) is Cauchy in the mean.

By hypothesis, we also have $(f_m) \to f$ a.e. (since pointwise everywhere convergence implies a.e. convergence). By [Fri03, Theorem 2.8.2] f is integrable, and

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

□

Appendix C

Fatou's lemma

Lemma C.0.1 (Fatou's lemma)**.** *Let* (f_n) *be a sequence of non-negative integrable functions, and let* $f = \underline{\lim}_n f_n(x)$*. Then*

$$\int f \; d\mu \leqslant \underline{\lim}_n \int f_n \; d\mu.$$

In particular, if $\underline{\lim}_n \int f_n \; d\mu < \infty$ *then* f *is integrable.*

Proof. If $\underline{\lim}_n \int f_n \; d\mu = \infty$ then the theorem is obviously true, so it remains to prove that it holds whenever $\underline{\lim}_n \int f_n \; d\mu < \infty$.

Let $g_n(x) = \inf_{j \geqslant n} f_j(x)$. Then g_n is measurable for all n by [Fri03, Theorem 2.2.3], and we have that $g_n(x) \leqslant f_k(x)$ for all $k \geqslant n$ and for all $x \in X$. Since f_k for $k \geqslant n$ is integrable, it follows by [Fri03, Theorem 2.10.1] that g_n is integrable. We have that

$$\begin{aligned} g &:= \lim_{n \to \infty} g_n(x) \\ &= \lim_{n \to \infty} \inf_{j \geqslant n} f_j(x) \\ &= \sup_n \inf_{j \geqslant n} f_j(x) \\ &= \underline{\lim}_n f_n(x) \\ &= f(x). \end{aligned}$$

We have a sequence of integrable functions g_n that is monotone increasing pointwise, and since $f_j(x) \geqslant 0$ we have that 0 is a lower bound to $\{f_j(x) \; : \; j \geqslant n\}$ for each fixed x, thus $g_n(x) \geqslant 0$. By Lebesgue's monotone convergence theorem it follows that

$$\begin{aligned} \int g \; d\mu &= \int \lim_{n \to \infty} g_n \; d\mu \\ &= \lim_{n \to \infty} \int g_n \; d\mu \\ &= \sup_n \int g_n \; d\mu \qquad \text{(since } g_n \leqslant g_{n+1} \text{ for all } n\text{).} \end{aligned}$$

Since $g_n \leqslant f_k$ for $k \geqslant n$ it follows that

$$\int g_n \; d\mu \leqslant \inf_{k \geqslant n} \int f_k \; d\mu.$$

By the previous reasoning, we then find that

$$\begin{aligned}\int g\ d\mu &= \sup_n \int g_n d\mu \\ &\leqslant \sup_n \inf_{k\geqslant n} \int f_k\ d\mu \\ &= \underline{\lim}_n \int f_n d\mu\end{aligned}$$

But

$$\begin{aligned}\int g\ d\mu &= \int f\ d\mu \\ \Rightarrow \int f\ d\mu &\leqslant \underline{\lim}_n \int f_n\ d\mu,\end{aligned}$$

which is what we wanted to show. $\square$

Appendix D

Some versions of Fubini's theorem

Theorem D.0.1 (Fubini's theorem: Version I)**.** *If h is an integrable function on $X \times Y$, then almost every almost every section of h is integrable, the functions*

$$\begin{cases} f(x) & = \int h_x(y)\ d\nu(y) \\ g(y) & = \int h^y(x)\ d\mu(x) \end{cases}$$

are integrable, and

$$\begin{aligned} \iint h\ d(\mu \times \nu) &= \iint h(x,y)\ d\mu(x)\ d\nu(y) \\ &= \iint h(x,y)\ d\nu(y)\ d\mu(x). \end{aligned}$$

Theorem D.0.2 (Fubini's theorem: Version II)**.** *If h is a measurable function on $X \times Y$ and either*

$$\iint |h|\ d\mu\ d\nu < \infty \qquad \textit{or} \qquad \iint |h|\ d\nu\ d\mu < \infty$$

holds, then h is integrable on $X \times Y$ and the assertions of D.0.1 hold.

Bibliography

[Fri03] Avner Friedman. *Foundations of Modern Analysis*. Dover Books on Mathematics. Dover Publications Inc., Mar. 17, 2003, p. 272. ISBN: 0486640620.

Index