Grad Analysis Notes

Pramana Saldin

Fall 2024

Real analysis concentrating on measures, integration, and differentiation and including an introduction to Hilbert spaces. Textbook: [Fol99].
Professor: Betsy Stovall.

Contents

1. Measure theory

1.1. Motivation: why measure theory? Measuring $\mathbb{R}$

September 04, 2024

We've encountered normed, complete spaces like $C^0(\mathbb{R})$ and $C^1(\mathbb{R})$. What we want is normed complete spaces based on integration. For example, $\int f$ and $\left(\int |f|^2\right)^2$. Riemann integration is insufficient for this, e.g.

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q}, \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

Definition 1.1

A **measure space** is a triple $(X, \mathcal{M}, \mu)$, where

1. X is a set,
2. $\mathcal{M} \subseteq \mathcal{P}(X)$ (the power set of X), that is,
   a) nonempty,
   b) closed under complements,
   c) closed under countable unions.

   Equivalently, $\mathcal{M}$ is a **σ-algebra** (of sets).
3. $\mu \colon \mathcal{M} \to [0, \infty]$ satisfying
   a) $\mu(\varnothing) = 0$,
   b) (*countable additivity*) if $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ disjoint sets, then

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i).$$

The sets $A \in \mathcal{M}$ are called **measurable sets**, and $A \notin \mathcal{M}$ are called **non-measurable sets**.

Example 1.1 –

1. $\{\varnothing, X\}$,
2. $\mathcal{P}(X)$, where

$$\mu(E) := \begin{cases} \#(E) & \text{if } E \text{ finite}, \\ \infty & \text{otherwise}. \end{cases}$$

3. Consider

$$\mathcal{M} := \left\{E \subseteq X : E \text{ or } E^C \text{ is countable}\right\},$$

   with

$$\mu(E) := \begin{cases} 0 & \text{if } E \text{ countable}, \\ \infty & \text{otherwise}. \end{cases}$$

Example 1.2 (Non-examples) –

1. intervals in $\mathbb{R}$,
2. open subset of $\mathbb{R}$ (not closed under complements),
3. open or closed sets (half-open intervals are not included).

Here's the solution for $\mathbb{R}$: let G_δ be countable intersections of open sets, let F_σ be countable unions of closed sets, let $G_{\delta\sigma}$ be countable unions of G_δ sets, let $F_{\sigma\delta}$ be countable intersections of F_σ sets, and so on. Taking the unions of all these sets, we get a σ-algebra on $\mathbb{R}$, which we call the **Borel σ-algebra of $\mathbb{R}$**.

Theorem 1.1 (Non-existence of power set measure on $\mathbb{R}$)
There does not exist a measure μ on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ such that

1. $\mu([0,1)) = 1$,
2. $\mu(E + \{x\}) = \mu(E)$.

This is E translated by x.

Proof. Suppose such a μ existed. Define an equivalence relation $x \sim y$ if $x - y \in \mathbb{Q}$. Consider its equivalence classes $[x] \in \mathbb{R}/\sim$. Every $[x]$ contains an element of $[0,1)$. By the axiom of choice, there exists a set $N \subseteq [0,1)$ containing 1 element from each equivalence class.

For $r \in \mathbb{Q} \cap [0,1)$, define

$$N_r := \underbrace{((N + \{r\}) \cap [0,1))}_{N'_r} \cup \underbrace{((N + \{r-1\}) \cap [0,1))}_{N''_r}.$$

N'_r and N''_r are disjoint, so

$$\begin{aligned}
\mu(N_r) &= \mu(N'_r) + \mu(N''_r) \\
&= \mu(N'_r - \{r\}) + \mu(N''_r - \{r-1\}) \\
&= \mu((N'_r - \{r\}) \cup (N''_r - \{r-1\})) && \text{(additivity)} \\
&= \mu(N).
\end{aligned}$$

Furthermore, the N_r's are disjoint, otherwise N would have 2 elements differing by a rational. Finally,

$$\bigcup_{r \in [0,1)} N_r = [0,1).$$

Hence,

$$1 = \mu([0,1)) = \mu\left(\bigcup_{r \in [0,1)} N_r\right) = \sum_{r \in [0,1)} \mu(N - r) = \sum_{r \in [0,1)} \mu(N). \qquad \square$$

1.2. The Lebesgue measure on $\mathbb{R}$

1.3. h-intervals

We start by defining $\mathcal{E}$ as the set of all *finite* unions of intervals open on the left, i.e. intervals of the form

$$(a_n, b_n], (a_n, \infty), (-\infty, b_n].$$

We call these **h-intervals**.

Proposition 1.2 (Properties of $\mathcal{E}$)

1. $\varnothing \in \mathcal{E}$,
2. $\mathcal{E}$ is closed under finite unions,
3. $\mathcal{E}$ is closed under complements.

This makes $\mathcal{E}$ an **algebra**.

It will be very useful for practical purposes to note that every element of $\mathcal{E}$ can be written as the disjoint union of h-intervals $I_n = (a_n, b_n]$ with $b_n < a_{n+1}$ for all n. Moreover, this expression is *unique*.

September 06, 2024

Define

$$\lambda_0\left(\bigcup_{n=1}^N I_n\right) = \sum_{n=1}^N \ell(I_n) = \sum_{n=1}^N b_n - a_n,$$

where $I_n = (a_n, b_n]$ are disjoint h-intervals from the representation described above. This is well-defined by the uniqueness of the representation.

λ_0 has two useful properties: (1) *monotonicity*, i.e., if $E \subseteq E'$, then $\lambda_0(E) \leq \lambda_0(E')$, and (2) *finitely additivity*, i.e., if $E = \bigcup_{j=1}^N E_j$, where $E_j \in \mathcal{E}$ are disjoint, then

$$\lambda_0(E) = \sum_{j=1}^N \lambda_0(E_j).$$

In fact, we can do better:

Proposition 1.3 (λ_0 is countably additive)

If $\{E_n\}_{n=1}^\infty \subseteq \mathcal{E}$ are disjoint, and

$$E := \bigcup_{n=1}^\infty E_n \in \mathcal{E},$$

then

$$\lambda_0\left(\bigcup_{n=1}^\infty E_n\right) = \sum_{n=1}^\infty \lambda_0(E_n).$$

We say λ_0 is *countably additive*.

Proof. By finite additivity and monotonicity, we know

$$\lambda_0\left(\bigcup_{n=1}^N E_n\right) = \sum_{n=1}^N \lambda_0(E_n) \leq \lambda_0(E).$$

Sending $N \to \infty$, we see that $\lambda_0(E) \geq \sum_{n=1}^\infty \lambda_0(E_n)$.

We need to show the other inequality now. By finite additivity, we may assume that E_n and E are h-intervals. We split into cases where E is finite and infinite.

Case I: $E = (a, b]$. Then $E_n = (a_n, b_n]$. Let $\varepsilon, \delta > 0$. Define $b_n' = b_n + \delta 2^{-n}$. Then (a_n, b_n') is an open cover of $[a + \varepsilon, b]$, which is compact. Hence, there are $n_1, \ldots, n_k$ such that $(a_{n_1}, b_{n_1}'), \ldots, (a_{n_k}, b_{n_k}')$ cover $[a + \varepsilon, b]$. We may reorder and discard any open intervals in this finite subcover that are contained in a larger interval, or don't intersect $[a+\varepsilon, b]$.

Now we can assume that $b \in (a_{n_1}, b'_{n_1})$, $a_{n_1} \in (a_{n_2}, b'_{n_2})$, etc. to $a_{n_j} \in (a_{n_{j+1}}, b'_{n_{k+1}})$ for $1 \leq j < k$. This makes $(a_{n_j}, b'_{n_j}]$ h-intervals that cover $(a+\varepsilon, b]$. So

$$\begin{aligned}
\lambda_0((a+\varepsilon, b]) &= (b-a)-\varepsilon \\
&\leq \sum_{j=1}^{k} \lambda_0((a_{n_j}, b'_{n_j}]) \\
&= \sum_{j=1}^{k} \ell((a_{n_j}, b'_{n_j}]) \\
&= \sum_{j=1}^{k} b'_{n_j} - a_{n_j} \\
&= \sum_{j=1}^{k} \delta_{n_j} + \sum_{j=1}^{k} b_{n_j} - a_{n_j} \\
&\leq \sum_{n=1}^{\infty} \delta_n + \sum_{n=1}^{\infty} b_n - a_n \\
&= \delta + \sum_{n=1}^{\infty} \lambda_0(E_n).
\end{aligned}$$

Taking $\varepsilon, \delta \to 0$ finishes.

Case II: E is infinite. We need to show that $\lambda_0(E) \leq \sum_{n=1}^{\infty} \lambda_0(E_n)$. Define

$$E_M := E \cap (-M, M], \qquad E_{n,M} := E_n \cap (-M, M].$$

By Case I,

$$\lambda_0(E_M) \leq \sum_{n=1}^{N} \lambda_0(E_{n,M}) \leq \sum_{n=1}^{N} \lambda_0(E_n).$$

Taking $N \to \infty$ and $M \to \infty$ finishes. □

1.3.1. Extending λ_0

This function λ_0 is a good candidate for a measure, but $\mathcal{E}$ is not a σ-algebra. We may extend this to any set $A \in \mathcal{P}(\mathbb{R})$ by defining

$$\lambda^*(A) := \inf\left\{ \sum_{n=1}^{\infty} \lambda_0(E_n) : A \subseteq \bigcup_{n=1}^{\infty} E_n, \{E_n\} \subseteq \mathcal{E} \right\}. \tag{1.1}$$

In other words, $\lambda^*(A)$ is the infimum over all countable $\mathcal{E}$-covers of A.

Proposition 1.4 (Properties of λ^*)

1. $\lambda^*(\varnothing) = 0$,
2. (*monotonicity*) $A \subseteq B \implies \lambda^*(A) \leq \lambda^*(B)$,
3. (*countable subadditivity*) given $A = \bigcup_{n=1}^{\infty} A_n$,
$$\lambda^*(A) \leq \sum_{n=1}^{\infty} \lambda^*(A_n).$$
4. $\lambda^*|_{\mathcal{E}} = \lambda_0$.

Properties (1-3) make λ^* an **outer measure**.

Proof. (1) is clear. (2) follows from the fact that a cover of A is also a cover of B.

(3) If $\{E_{n,j}\}_j$ covers A_n and

$$\sum_{j=1}^{\infty} \lambda_0(E_{n,j}) \leq \lambda^*(A_n) + \delta 2^{-n},$$

then $\bigcup_n \{E_{n,j}\}_j$ covers A, and

$$\sum_{n,j} \lambda_0(E_{n,j}) \leq \sum_n \lambda^*(A_n) + \delta,$$

and then send $\delta \to 0$.

(4) If $E \in \mathcal{E}$, and $\{E_j\} \subseteq \mathcal{E}$ covers E, then the sets

$$E_j' \coloneqq \left(E_j \setminus \bigcup_{k<j} E_k \right) \cap E$$

has union E, and are disjoint. So by countable additivity,

$$\lambda_0(E) = \sum_{j=1}^{\infty} \lambda_0(E_j') \leq \sum_{i=1}^{\infty} \lambda_0(E_j) \implies \lambda_0(E) \leq \lambda^*(E).$$

On the other hand, E covers E, so

$$\lambda_0(E) \geq \lambda^*(E).$$ □

1.3.2. Measurability

Definition 1.2

$A \subseteq \mathbb{R}$ is **measurable**, denoted $A \in \mathcal{L}$ if A **decomposes** B, i.e.

$$\lambda^*(B) = \lambda^*(B \cap A) + \lambda^*(B \cap A^C)$$

for all $B \in \mathcal{P}(\mathbb{R})$.

Remark 1.5. "$\leq$" is done by countable subadditivity, so proving a set is measurable amounts to proving "$\geq$". This always holds if $\lambda^*(B) = \infty$, so to prove a set A is measurable, we simply need to show

$$\lambda^*(B) \geq \lambda^*(B \cap A) + \lambda^*(B \cap A^C)$$

for all $B \in \mathcal{P}(\mathbb{R})$ such that $\lambda^*(B) < \infty$.

We call $\lambda := \lambda^*|_{\mathcal{L}}$ the **Lebesgue measure**.

Proposition 1.6
The h-interval $(-\infty, a]$ for $a \in \mathbb{R}$ is measurable.

Proof. For $E \in \mathcal{E}$,

$$\lambda_0(E) = \lambda_0((-\infty, a] \cap E) + \lambda_0((a, \infty) \cap E),$$

which implies

$$\lambda^*(B) = \lambda^*((-\infty, a] \cap B) + \lambda^*((a, \infty) \cap B). \qquad \square$$

September 09, 2024

Theorem 1.7

1. $\mathcal{L}$ is a σ-algebra,
2. λ is a measure on $\mathcal{L}$,
3. $\lambda^*(Z) = 0 \implies Z \in \mathcal{L}$.

A corollary of (1) is that $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{L}$, since the Borel sets are the smallest σ-algebra generated by the topology on $\mathbb{R}$, and $\mathcal{L}$ is the smallest σ-algebra generated by $(-\infty, a]$.

If $Z \in \mathcal{L}$ and $\lambda(Z) = 0$ implies that

$$\mathcal{P}(Z) \subseteq \mathcal{L},$$

we say that λ is **complete**. As a consequence of this, $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{L}$.

Proof of Theorem 1.7. (3) If $\lambda^*(Z) = 0$ and $B \in \mathcal{P}(\mathbb{R})$, then

$$\lambda^*(B) \geq \lambda^*(B \cap Z^C) \geq \lambda^*(B \cap Z) + \lambda^*(B \cap Z^C).$$

(1) We need to show that $\mathcal{L}$ is (a) closed under complements, and (b) closed under countable unions.

(a) is easy because the decomposition condition is the same when replacing A with A^C.

(b) First, we show this is true for *finite* unions. Let

$$\{A_n\}_{n=1}^N \subseteq \mathcal{L}.$$

Define

$$A'_N := \bigcup_{n=1}^N A_n, \qquad A''_N := A_N \setminus A'_{N-1}.$$

Let $B \in \mathcal{P}(\mathbb{R})$. By drawing a Venn diagram and using the fact that A_1 and A_2 individually are measurable,

$$\begin{aligned}\lambda^*(B\cap(A_1\cup A_2))+\lambda^*(B\cap(A_1\cup A_2)^C) &\le \lambda^*(B\cap A_1\cap A_2)+\lambda^*(B\cap A_1\cap A_2^C)\\ &\quad+\lambda^*(B\cap A_1^C\cap A_2)+\lambda^*(B\cap A_1^C\cap A_2^C)\\ &=\lambda^*(B\cap A_1)+\lambda^*(B\cap A_1^C)\\ &=\lambda^*(B).\end{aligned}$$

We induct to conclude that $A'_N \in \mathcal{L}$ for all $N \in \mathbb{N}$. By (a) and DeMorgan's laws, we also have $A''_N \in \mathcal{L}$.

Set $A = \bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty A'_n = \bigcup_{n=1}^\infty A''_n$. Let $B \in \mathcal{P}(\mathbb{R})$. Then

$$\begin{aligned}\lambda^*(B) &= \lambda^*(B\cap A'_N)+\lambda^*(B\cap A'^C_N)\\ &\ge \lambda^*(B\cap A'_N\cap A''_N)+\lambda^*(B\cap A'_N\cap A''^C_N)+\lambda^*(B\cap A^C)\\ &=\lambda^*(B\cap A''_N)+\lambda^*(B\cap A'_{N-1})+\lambda^*(B\cap A^C).\end{aligned}$$

By induction,

$$\lambda^*(B)\ge\sum_{n=1}^N \lambda^*(B\cap A''_n)+\lambda^*(B\cap A^C)$$

for all N. Sending $N\to\infty$,

$$\begin{aligned}\lambda^*(B) &\ge \sum_{n=1}^\infty \lambda^*(B\cap A''_n)+\lambda^*(B\cap A^C)\\ &\ge \lambda^*\left(\bigcup_{n=1}^\infty (B\cap A''_n)\right)+\lambda^*(B\cap A^C)\\ &=\lambda^*(B\cap A)+\lambda^*(B\cap A^C).\end{aligned}$$

(2) $\lambda(\varnothing)=\lambda^*(\varnothing)=0$, so it remains to check countable additivity. Let $\{A_n\}_{n=1}^\infty \subseteq \mathcal{L}$ be pairwise disjoint. By countable subadditivity, we simply need to show

$$\lambda\left(\bigcup_{n=1}^\infty A_n\right)\ge\sum_{n=1}^\infty \lambda(A_n).$$

For finite additivity, we have

$$\begin{aligned}\lambda\left(\bigcup_{n=1}^N A_n\right) &= \lambda^*\left(\bigcup_{n=1}^N A_n\right)\\ &=\lambda^*\left(\bigcup_{n=1}^N A_n\cap A_N\right)+\lambda^*\left(\bigcup_{n=1}^N A_n\cap A_N^C\right)\\ &=\lambda(A_N)+\lambda\left(\bigcup_{n=1}^{N-1} A_n\right).\end{aligned}$$

By induction, we have equality. For the countable case,

$$\lambda\left(\bigcup_{n=1}^\infty A_n\right)=\lambda\left(\bigcup_{n=1}^N A_n\right)+\lambda\left(\bigcup_{n=N+1}^\infty A_n\right)\ge\sum_{n=1}^N \lambda(A_n).$$

Sending $N \to \infty$ finishes. $\square$

Notice that in this proof, we did not need the fact that we were working with subsets of $\mathbb{R}$. Now is a good time to generalize what we have worked with so far.

1.4. General measure space constructions

1.4.1. Constructing σ-algebras

Definition 1.3

For $\mathcal{E} \subseteq \mathcal{P}(X)$, define the **$\sigma$-algebra generated by $\mathcal{E}$**, denoted $\mathcal{M}(\mathcal{E})$ as the smallest σ-algebra containing $\mathcal{E}$, or (equivalently) the intersection of every σ-algebra containing $\mathcal{E}$.

Example 1.3 – The Borel sets are the σ-algebra generated by open intervals, or by intervals of the form $(-\infty, a]$, where $a \in \mathbb{R}$:

$$\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\{(a, b) : a < b\}) = \mathcal{M}(\{(-\infty, a] : a \in \mathbb{R}\}).$$

September 11, 2024

Definition 1.4

Let $\{(X_\alpha, \mathcal{M}_\alpha)\}_{\alpha \in \mathcal{A}}$ be a collection of measure spaces for an indexing set $\mathcal{A}$ (where $X_\alpha \neq \varnothing$, $\mathcal{M}_\alpha$ is a σ-algebra). Set $X = \prod_\alpha X_\alpha$. The **product σ-algebra** on X is

$$\bigotimes_{\alpha \in \mathcal{A}} \mathcal{M}_\alpha := \mathcal{M}\left(\left\{\pi_\alpha^{-1}(E_\alpha) : \alpha \in \mathcal{A}, E_\alpha \in \mathcal{M}_\alpha\right\}\right),$$

where $\pi_\alpha : X \to X_\alpha$ is the coordinate projection.

This has nothing to do with the tensor product, we just needed a symbol that wasn't the normal product.

For $\mathcal{A}$ countable, we have the expected definition:

$$\bigotimes_{\alpha \in \mathcal{A}} \mathcal{M}_\alpha = \mathcal{M}\left(\left\{\prod_{\alpha \in \mathcal{A}} E_\alpha : E_\alpha \in \mathcal{M}_\alpha, \text{ for all } \alpha.\right\}\right) \tag{1.2}$$

For uncountable $\mathcal{A}$, if E is a nonempty set in $\bigotimes_{\alpha \in \mathcal{A}} \mathcal{M}_\alpha$, then $\pi_\alpha(E) = X_\alpha$ for all except countably many α (c.f. the product vs the box topology).

If $\mathcal{M}_\alpha = \mathcal{M}(\mathcal{E}_\alpha)$, then we can replace $\mathcal{M}_\alpha$ with $\mathcal{E}_\alpha$ in the LHS of the definition and Equation 1.2.

Proposition 1.8

$\mathcal{B}_{\mathbb{R}^n} = \bigotimes_{j=1}^n \mathcal{B}_{\mathbb{R}}$.

Proof. The open intervals with rational endpoints countably generate the topology on $\mathbb{R}$ ($\mathcal{T}_{\mathbb{R}}$) (every open subset of $\mathbb{R}$ is a countable union of such intervals), and products of intervals with rational endpoints generate $\mathcal{T}_{\mathbb{R}^n}$. By the first remark,

$$\bigotimes_{j=1}^n \mathcal{B}_{\mathbb{R}} = \mathcal{M}\left(\left\{\prod_{j=1}^n (a_j, b_j) : a_j, b_j \in \mathbb{Q}\right\}\right) = \mathcal{B}_{\mathbb{R}^n}. \qquad \square$$

Following the same proof shows that

$$\mathcal{B}_{\prod_{j=1}^{n} X_j} = \bigotimes_{j=1}^{n} \mathcal{B}_{X_j},$$

provided that $\{X_j\}$ are second countable. The special case is when X is a *separable metric space* (a metric space with a countable, dense subset, e.g., $\mathbb{Q} \subseteq \mathbb{R}$).

On the other hand, if X_1 and X_2 are not second countable,

$$\{U_1 \times U_2 : U_j \subseteq X_j, j = 1, 2\}$$

does not countably generate $\mathcal{T}_{X_1 \times X_2}$.

1.4.2. Constructing elementary families

Definition 1.5

$\mathcal{E} \subseteq \mathcal{P}(X)$ is an **elementary family** if

1. $\varnothing \in \mathcal{E}$,
2. $E, F \in \mathcal{E}$ implies $E \cap F \in \mathcal{E}$ and $E \in \mathcal{E} \implies E^C$ is a finite disjoint union of elements of $\mathcal{E}$.

Example 1.4 – The h-intervals in $\mathbb{R}$ form an elementary family.

Proposition 1.9

If $\mathcal{E}$ is an elementary family, then finite unions of elements of $\mathcal{E}$ form an algebra.

The proof for this is omitted, but the idea is to play with DeMorgan's laws. We will call the algebra created by Proposition 1.9, $\mathcal{A}$.

1.4.3. Constructing premeasures

Definition 1.6

If $\mathcal{A} \subseteq \mathcal{P}(X)$ is an algebra, a **premeasure** on $\mathcal{A}$ is a function

$$\mu_0 : \mathcal{A} \to [0, \infty]$$

such that

1. $\mu_0(\varnothing) = 0$,
2. (*countable additivity*) if $\{A_n\} \subseteq \mathcal{A}$ are disjoint,[1] and $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, then

$$\mu_0\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

[1] Actually, we get disjointness for free because $\mathcal{A}$ is an algebra.

Proposition 1.10
If $\mathcal{E}$ is an elementary family and

$$\rho\colon \mathcal{E} \to [0,\infty]$$

obeys

1. $\rho(\varnothing) = 0$,
2. $\rho\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \rho(E_n)$, where $\{E_n\} \subseteq \mathcal{E}$ and $\bigcup_{n=1}^{\infty} E_n \in \mathcal{E}$,

then if

$$\mu_0\left(\bigcup_{n=1}^{N} E_n\right) = \sum_{n=1}^{N} \rho(E_n),$$

whenever $\{E_n\} \subseteq \mathcal{E}$ are disjoint, μ_0 is a premeasure on $\mathcal{A}$.

Proof (sketch). We have to check that μ_0 is well-defined. Then we check $\mu_0(\varnothing) = 0$, which is immediate. Then we may use the restricted countable additivity on ρ to get countable additivity for μ_0. □

1.4.4. Constructing outer measures

Definition 1.7
An **outer measure** on $X \neq \varnothing$ is a function

$$\mu^*\colon \mathcal{P}(X) \to [0,\infty]$$

such that

1. $\mu^*(\varnothing) = 0$,
2. (*monotonicity*) $E \subseteq F \implies \mu^*(E) \leq \mu^*(F)$,
3. (*countable subadditivity*) $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu^*(E_n)$.

Proposition 1.11 (Constructing an outer measure)
Let $\mathcal{E} \subseteq \mathcal{P}(X)$ contain the sets $\varnothing$ and X, and let

$$\mu_0\colon \mathcal{E} \to [0,\infty]$$

satisfy $\mu_0(\varnothing) = 0$.[1] Then

$$\mu^*(A) \coloneqq \inf\left\{\sum_{n=1}^{\infty} \mu_0(E_n) : \{E_n\} \subseteq \mathcal{E},\ \bigcup_{n=1}^{\infty} E_n \supseteq A\right\}$$

is an outer measure on X.

[1] Notice that we don't require that μ_0 is a premeasure.

Proof. We didn't use any of the properties of $\mathbb{R}$ in the proof of Proposition 1.4, so we can the same proof as used there. □

1.4.5. Constructing measures

Similar to how we got λ from λ^*, we can get a measure from an outer measure.

Theorem 1.12 (Caratheodory's theorem)
Let μ^* be an outer measure on $X \neq \varnothing$, and say $A \in \mathcal{M}$, or A is **measurable**, if A **decomposes** μ^*, i.e.,

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

for all $B \in \mathcal{P}(X)$. Then $\mu := \mu^*|_{\mathcal{M}}$ is a measure.[1,2]

[1] In fact, μ is a *complete* measure, i.e., subsets of null sets are measurable.
[2] Just like before, "$\leq$" is immediate since μ^* is countably subadditive.

Proof. Copy the proof of the Lebesgue measure on $\mathbb{R}$. □

Theorem 1.13
If $\mathcal{A} \subseteq \mathcal{P}(X)$ is an algebra and $\mu_0 \colon \mathcal{A} \to [0, \infty]$ is a premeasure, then $\mathcal{A} \subseteq \mathcal{M}$, and $\mu^*|_{\mathcal{A}} = \mu_0$.

Proof. Let $A \in \mathcal{A}$ To prove $A \in \mathcal{M}$, suppose that $E \in \mathcal{A}$. Then

$$\mu_0(E) = \mu_0(E \cap A) + \mu_0(E \setminus A)$$

by countable additivity. Let $B \in \mathcal{P}(X)$. Suppose that B was covered by $\{E_n\} \subseteq \mathcal{A}$. For each E_n, we have

$$\mu_0(E_n) = \mu_0(E_n \cap A) + \mu_0(E_n \setminus A).$$

So

$$\sum_{n=1}^{\infty} \mu_0(E_n) = \sum_{n=1}^{\infty} \mu_0(E_n \cap A) + \sum_{n=1}^{\infty} \mu_0(E_n \setminus A) \geq \mu^*(B \cap A) + \mu^*(B \setminus A).$$

Taking the infimum over all such covers of B yields

$$\mu^*(B) \geq \mu^*(B \cap A) + \mu^*(B \setminus A),$$

which is enough to satisfy Theorem 1.12.

Suppose A is covered by $\{E_n\} \subseteq \mathcal{A}$. We may create new, disjoint sets

$$E_n' := \left(E_n \setminus \bigcup_{k<n} E_k\right) \cap A,$$

which belong to A because $\mathcal{A}$ is an algebra, and whose union is A. So by the countable additivity of the premeasure,

$$\mu_0(A) = \mu_0\left(\bigcup_{n=1}^{\infty} E_n'\right) = \sum_{n=1}^{\infty} \mu_0(E_n') \leq \sum_{n=1}^{\infty} \mu_0(E_n) \implies \mu_0(A) \leq \mu^*(A).$$

On the other hand, A covers A, so

$$\mu_0(A) \geq \mu^*(A).$$

□

1.5. Measure properties

September 13, 2024

Definition 1.8

Here are some common descriptors of the "finiteness" of measures:

1. μ is **finite** if $\mu(X) < \infty$.
2. μ is **σ-finite** if $X = \bigcup_n E_n$ with $E_n \in \mathcal{M}$ and $\mu(E_n) < \infty$ for all n.
3. μ is **semifinite** if every measurable set with $\mu(E) = \infty$ contains an $F \in \mathcal{M}$ with nonzero, finite measure ($0 < \mu(F) < \infty$).

Remark 1.14. (1) is nice because we avoid doing "$\infty - \infty$." Thus,

$$\mu(E) = \mu(X) - \mu(E^C)$$

for all $E \in \mathcal{M}$.

A non-example for (2) is the counting measure (which counts the number of elements in a set) on $\mathbb{R}$ or any uncountable set.

A non-example for (3) is letting $X = \mathbb{R}$ and $\mathcal{M}$ be the set of countable/co-countable subsets of $\mathbb{R}$, and defining

$$\mu(E) := \begin{cases} 0 & E \text{ is countable,} \\ \infty & E \text{ is uncountable.} \end{cases}$$

Theorem 1.15

For $(X, \mathcal{M}, \mu)$ a measure space,

1. (*monotonicity*) $E, F \in \mathcal{M}$ and $E \subseteq F$ implies $\mu(E) \leq \mu(F)$,
2. (*subadditivity*) $\{E_n\} \subseteq \mathcal{M}$ implies
$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu(E_n),$$
3. (*continuity from below*) if $\{E_n\} \subseteq \mathcal{M}$, then
$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} E_n\right),$$
4. (*continuity from above*) if $\{E_n\} \subseteq \mathcal{M}$, then
$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{N\to\infty} \mu\left(\bigcap_{n=1}^{N} E_n\right),$$
provided that the RHS is **finite**.

Proof. (1) $\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E)$.

(2) Let $E'_n := E_n \setminus \bigcup_{k<n} E_k$. These are disjoint and have union $\bigcup_n E_n$, so

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \mu\left(\bigcup_{n=1}^{\infty} E'_n\right) = \sum_{n=1}^{\infty} \mu(E'_n) \leq \sum_{n=1}^{\infty} \mu(E_n).$$

(3) Define E'_n the same as before.

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} E_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} E'_n\right) \\ &= \sum_{n=1}^{\infty} \mu(E'_n) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \mu(E'_n) \\ &= \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} E'_n\right) \\ &= \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} E_n\right).\end{aligned}$$

(4) WLOG, $E_n = \bigcap_{k\leq n} E_k$. By the hypothesis, there exists an n_0 such that

$$\mu(E_{n_0}) < \infty.$$

WLOG, $n_0 = 1$. Then

$$\begin{aligned}\mu\left(\bigcap_{n=1}^{\infty} E_n\right) &= \mu(E_1) - \mu\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right) \\ &= \mu(E_1) - \mu\left(\bigcup_{n=1}^{\infty} E_1 \setminus E_n\right) && (\mu(E_1) < \infty) \\ &= \mu(E_1) - \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} E \setminus E_n\right) && \text{(part (3))} \\ &= \lim_{N\to\infty} \mu\left(\bigcap_{n=1}^{N} E_n\right).\end{aligned}$$

□

We mentioned completeness in Theorem 1.12, but we provide the formal definition now.

Definition 1.9

A measure μ is **complete** if $Z \in \mathcal{M}$ and $\mu(Z) = 0$ (i.e. Z is **null**) implies $\mathcal{P}(Z) \subseteq \mathcal{M}$. In other words, all subset of null sets are measurable.

Example 1.5 – Here's a sketch for how to prove that $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \lambda)$ is *not* complete (this is fine to skip over):

1. Consider the *middle-third Cantor set* $K = \bigcap_{n=1}^{\infty} K_n$, where each K_n is a union of 2^n intervals of size 3^{-n}. The Cantor measure is the (unique) Borel measure μ on $\mathbb{R}$ satisfying $\mu(I) = 2n$ for each interval $I \subseteq K_n$ of length 3^n and $\mu(I) = 0$ for each interval disjoint from K. Thus, $\mu(K) = 1$. Define the *Cantor function* as

$$f(x) := \begin{cases} \mu([0, x)) & x \geq 0, \\ -\mu([x, 0)) & x < 0. \end{cases}$$

2. We can check that f is continuous by a δ-ε argument or by looking at the left and right continuity at each point by sequences of points on the left or right. Further, we can show that f maps K to $[0, 1]$ because it is constant outside of the Cantor set.

3. $g(x) := f(x) + x$ is a bijection from $[0, 1]$ to $[0, 2]$ because it is strictly increasing.

4. We can check the Lebesgue measure of $g(K)$ is 1 by, e.g., looking at complements.

5. Finally, we can prove that g is a homeomorphism, so a set is Borel in $[0, 2]$ if and only if its preimage under g is a Borel set in $[0, 1]$. Choose some non-measurable subset of $g(K)$ (the existence of a non (Lebesgue)-measurable set inside a positive measure set is also provable). Its preimage under g must not be Borel, but it is contained in K, so it is Lebesgue measurable.

Theorem 1.16 (Completion of a measure space)
If $(X, \mathcal{M}, \mu)$ is a measure space,

$$\overline{\mathcal{M}} := \{E \cup Z' : E \in \mathcal{M}, Z' \subseteq Z \in \mathcal{M},\ Z \text{ a null set}\},$$

is a σ-algebra and the measure $\overline{\mu}(E \cup Z') := \mu(E)$, where E and Z' are as above, is a complete measure.

Proof (key observations). To check complements, we use the fact that

$$(E \cup Z')^C = (E \cup Z)^C \cup (Z \setminus (Z' \setminus E)) = \underbrace{E^C \cap Z^C}_{\in \mathcal{M}} \cup \underbrace{(Z \setminus (Z' \setminus E))}_{\subseteq Z}.$$

Checking that this is closed under countable unions uses the fact that countable unions of null sets are null. □

Proposition 1.17
If $E \in \mathcal{L} \subseteq \mathcal{P}(\mathbb{R})$, then for all $\varepsilon > 0$, there exists closed F and open U such that

$$F \subseteq E \subseteq U$$

and $\lambda(U \setminus F) < \varepsilon$.

Remark 1.18. Here are some consequences of Proposition 1.17:

1. $E \in \mathcal{L} \iff E = G \setminus Z$, where $G \in G_\delta$ (the countable intersection of open sets) and Z is null $\iff E = F \cap Z'$, where $F \in F_\sigma$ (the countable union of closed sets) and Z' is null,

2. $E \in \mathcal{L} \implies \lambda(E) = \sup\{\mu(K) : K \subseteq E,\ K \text{ is compact}\}$,
3. Given $E \in \mathcal{L}$ and $\varepsilon > 0$, there exists a set U, the finite union of open intervals, such that

$$\mu((U \setminus E) \cup (E \setminus U)) < \varepsilon.$$

Proof of Proposition 1.17. Case I: $E \subseteq I$, I compact interval. We sketch a proof. Similar to the homework, there exists an open U such that $U \supseteq E$ and $\lambda(U) < \lambda(E) + \varepsilon$. Likewise, there exists an open $V \supseteq I \setminus E$ with

$$\lambda(V) < \lambda(I) - \lambda(E) + \varepsilon.$$

Let $F = I \setminus V$.

Case II: general case. Let $I_n \coloneqq [n, n+1]$, and $E_n \coloneqq E \cap I_n$. By case (1), we can choose $F_n \subseteq E_n$ and $U_n \supseteq E_n$ such that

$$\mu(U_n \setminus F_n) < \varepsilon 2^{-n}.$$

Set $U \coloneqq \bigcup_{n=1}^{\infty} U_n$, which is open, and we can show that $F \coloneqq \bigcup_{n=1}^{\infty} F_n$ is closed. Then

$$\begin{aligned} \mu(U \setminus F) &= \mu\left(\bigcup_{n=1}^{\infty} U_n \setminus F_n\right) \\ &\leq \sum_{n=1}^{\infty} \mu(U_n \setminus F_n) < \varepsilon. \end{aligned}$$

□

1.6. Measurable functions

September 16, 2024

Definition 1.10

Let $(X, \mathcal{M})$, $(Y, \mathcal{N})$ be measurable spaces. Say $f\colon X \to Y$ is **$(\mathcal{M}, \mathcal{N})$-measurable** if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{M}$.

Proposition 1.19

If $\mathcal{M}$ is generated by $\mathcal{E}$, then $f\colon X \to Y$ is measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$.

Proof. ($\implies$) Is immediate because $\mathcal{E} \subseteq \mathcal{M}$.

($\impliedby$)

Claim 1.1. $\left\{E \subseteq Y : f^{-1}(E) \in \mathcal{M}\right\}$ is a σ-algebra.

> **Proof.** Use the facts that $f^{-1}(E^C) = f^{-1}(E)^C$, and $f^{-1}\left(\bigcup_{n=1}^{\infty} E_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(E_n)$ from set theory, combined with the fact that $\mathcal{M}$ is a σ-algebra. ■

Since the set in the claim is a σ-algebra containing $\mathcal{E}$, it has to contain $\mathcal{N}$. □

Definition 1.11

We say $f\colon X \to \mathbb{R}$ (or $\mathbb{C}$, $\mathbb{R}^n$, or a manifold) is **measurable** if f is $(\mathcal{M}, \mathcal{N})$-measurable, where $\mathcal{N} = \mathcal{B}_{\mathbb{R}}$ (or $\mathcal{B}_{\mathbb{C}}$, etc.).

Remark 1.20. "f is a $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable" is a logically weaker condition than "f is a $(\mathcal{M}, \mathcal{L})$-measurable." We use the former because the latter rules out many (even continuous!) functions. The rest of class will be basically explaining why this makes sense.

Corollary 1.21
$f\colon X \to \mathbb{R}$ is measurable $\iff$ any of the following hold:

1. $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$,
2. $f^{-1}((-\infty, b]) \in \mathcal{M}$ for all $b \in \mathbb{R}$,
3. $f^{-1}((c, \infty)) \in \mathcal{M}$ for all $c \in \mathbb{R}$.

Corollary 1.22
Let $(X, \mathcal{M}), \{(Y_\alpha, \mathcal{M}_\alpha)\}_{\alpha \in \mathcal{A}}$ be measurable spaces. Let

$$Y := \prod_\alpha Y_\alpha, \qquad \mathcal{M} := \bigotimes_\alpha M_\alpha.$$

Then $f\colon X \to Y$ is measurable $\iff$ $f_\alpha := \pi_\alpha \circ f\colon X \to Y_\alpha$ is measurable for every $\alpha \in \mathcal{A}$.

Corollary 1.23
$f\colon X \to \mathbb{R}^n$ or $\mathbb{C}$ is measurable $\iff$ each coordinate function is measurable (for $\mathbb{C}$ this is if the real and imaginary parts are measurable).

Definition 1.12
$f\colon \mathbb{R} \to \mathbb{R}$ is **Borel** (resp. **Lebesgue**) **measurable** if f is $(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{R}})$-measurable (resp. $(\mathcal{L}, \mathcal{B}_{\mathbb{R}})$-measurable).

Remark 1.24. Notice now that Lebesgue measurable is logically *weaker* than Borel measurable.

Corollary 1.25
Continuous functions on $\mathbb{R}$ are Borel measurable.

Proposition 1.26
Let $M = \mathbb{R}, \mathbb{C}, \mathbb{R}^n$, or a manifold. If $F\colon M \to \mathbb{R}$ is continuous, and $g\colon X \to M$ is measurable, then $F \circ g$ is measurable.

Proof. Let $a \in \mathbb{R}$. Then

$$(F \circ g)^{-1}((-\infty, a)) = g^{-1}(F^{-1}(-\infty, a)).$$

$F^{-1}(-\infty, a)$ is open by F being continuous, and $g^{-1}(F^{-1}(-\infty, a)) \in \mathcal{M}$ by g being measurable. □

Remark 1.27. This works when F is Borel measurable as well.

Corollary 1.28
Let $f, g\colon X \to \mathbb{R}$ be measurable. Then so are

1. $f + g$,
2. $f \cdot g$,
3. f/g (provided that $g \neq 0$)
4. $\max(f, g)$,
5. $|f|$,
6. $\sqrt{f^2 + g^2}$.

1.6.1. Creating new measurable functions

Definition 1.13
Let $(X, \mathcal{M})$ be a measurable space and $A \in \mathcal{M}$. Then

$$\mathcal{M}_A \coloneqq \{A \cap B : B \in \mathcal{M}\}.$$

Proposition 1.29
$\mathcal{M}_A$ is a σ-algebra on A, and if $f\colon X \to Y$ is measurable, then $f|_A\colon A \to Y$ is measurable.

Proposition 1.30
If $f, g\colon X \to \mathbb{R}$ are measurable and $g \neq 0$, then $\{x : g(x) \neq 0\}$ is measurable and $f/g|_{\{x:g(x)\neq 0\}}$ is a measurable function.

Proof. We may write

$$\{x : g(x) \neq 0\} = g^{-1}((-\infty, 0)) \cup g^{-1}((0, \infty)).$$

Since each set on the RHS are measurable, we are done.

It suffices to show that $(f/g)^{-1}((-\infty, a))$ is measurable for all $a \in \mathbb{R}$. Of course, this is equivalent to

$$\{x : f(x) < g(x) \cdot a\}.$$

Notice that $f(x) < g(x) \cdot a$ if and only if there exists a rational number $r \in \mathbb{Q}$ such that $f(x) < r < g(x) \cdot a$. So if $a > 0$,

$$\{x : f(x) < g(x) \cdot a\} = \bigcup_{r \in \mathbb{Q}} \{x : f(x) < r\} \cup \left\{x : \frac{r}{a} < g(x)\right\},$$

if $a < 0$,

$$\{x : f(x) < g(x) \cdot a\} = \bigcup_{r \in \mathbb{Q}} \{x : f(x) < r\} \cup \left\{x : \frac{r}{a} > g(x)\right\},$$

and if $a = 0$,

$$\{x : f(x) < g(x) \cdot a\} = \{x : f(x) < 0\}.$$

All of these sets are measurable, so $f/g|_{\{x:g(x)\neq 0\}}$ is measurable. ☐

Proposition 1.31

If $\{f_n : X \to \overline{\mathbb{R}} \coloneqq \mathbb{R} \cup \{-\infty, \infty\}\}$ are measurable functions, then so are

(a) $\sup f_n$,

(b) $\inf f_n$,

(c) $\limsup f_n$,

(d) $\liminf f_n$,

(e) $G \coloneqq \{x : \lim_{n\to\infty} f_n(x) \text{ exists}\}$,

(f) $\lim_{n\to\infty} f_n|_G$.

September 18, 2024

Proof. (a) $\sup f_n \leq a \iff f_n(x) \leq a$ for all x. Therefore,

$$(\sup f_n)^{-1}([-\infty, a]) = \bigcap_n f_n^{-1}([-\infty, a]) \in \mathcal{M}.$$ [1]

(b) $\inf f_n = -\sup(-f_n)$.
(c) $\limsup f_n = \inf_k(\sup_{n\geq k} f_n)$, and a similar idea for (d).
(e)

$$\begin{aligned}\left\{x : \lim_{n\to\infty} f_n(x) \text{ exists}\right\} &= \{x : \limsup f_n(x) = \liminf f_n(x)\} \\ &= ((\limsup f_n - \liminf f_n)^{-1}(\{0\}) \cap \{\limsup f_n \in \mathbb{R}\}) \\ &\quad \cup (\{\limsup f_n = \infty\} \cap \{\liminf f_n = \infty\}) \\ &\quad \cup (\{\limsup f_n = -\infty\} \cap \{\liminf f_n = -\infty\}).\end{aligned}$$

(f) $\lim f_n|_G = \limsup f_n|_G$. By Proposition 1.29, we can restrict measurable functions and get measurable functions. ☐

[1] It's crucial that we have a *countable* sup, because *any* function can be written as an arbitrary supremum of measurable functions.

1.7. Measurability and measure (zero)

Definition 1.14

Let $(X, \mathcal{M}, \mu)$ be a measure space. We say a property $P(x)$ for $x \in X$ holds **almost everywhere (a.e.)** if there exists a null set Z such that $P(x)$ holds for all $x \in Z^C$.

Example 1.6 – Almost every point of $\mathbb{R}$ is transcendental.

Remark 1.32. If $(X, \mathcal{M}, \mu)$ is complete, $P(x)$ holds a.e. if $\{x : \neg P(x)\}$ is null.

Proposition 1.33
Let $(X, \mathcal{M}, \mu)$ be complete.

(a) If f is measurable and $f = g$ a.e., then g is measurable.

(b) If $\{f_n\}$ are measurable functions and $f_n \to f$ a.e. (i.e. the set of all points where $f_n(x) \not\to f(x)$ is a null set), then f is measurable.

Proof. (a) Let E be any measurable set. Then

$$\begin{aligned} g^{-1}(E) &= \{x : g(x) \in E\} \\ &= (\{x : f(x) \in E\} \setminus \{x : f(x) \neq g(x), f(x) \in E, g(x) \notin E\}) \\ &\quad \cup \{x : f(x) \neq g(x), f(x) \notin E, g(x) \in E\}. \end{aligned}$$

Notice that the last set is measurable since the set being removed and being added are subsets of a null set $(\{f(x) \neq g(x)\})$ and $(X, \mathcal{M}, \mu)$ is complete.

(b) Write $f^{-1}(-\infty, a) = \bigcup_m (\limsup f_n)^{-1}(-\infty, a - \frac{1}{m})$. □

Proposition 1.34
Let $(X, \overline{\mathcal{M}}, \overline{\mu})$ be the completion of $(X, \mathcal{M}, \mu)$. Every $\overline{\mu}$-measurable function $f \colon X \to \overline{\mathbb{R}}$ has a μ-measurable version $g \colon X \to \overline{\mathbb{R}}$ such that $\overline{\mu}(\{f \neq g\}) = 0$ (i.e. $f = g$ $\overline{\mu}$-a.e.).

Proof. For $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, let $E_{n,k} \coloneqq f\left([\frac{k}{n}, \frac{k+1}{n})\right) \in \overline{\mathcal{M}}$, which means that

$$E_{n,k} = F_{n,k} \cap Z'_{n,k}$$

for some $F_{n,k} \in \mathcal{M}$ and $Z'_{n,k} \subseteq Z_{n,k}$ where $Z_{n,k}$ is a null set in $\mathcal{M}$.

We have

$$f_n \coloneqq \sum_{k \in \mathbb{Z}} \frac{k}{n} \chi_{E_{n,k}} \to f$$

and the $\mathcal{M}$-measurable function

$$h_n \coloneqq \sum_{k \in \mathbb{Z}} \frac{k}{n} \chi_{F_{n,k}} \to f$$

on $(\bigcup Z_{n,k})^C$. Notice that $Z \coloneqq \bigcup Z_{n,k}$ is null. Then

$$G_n \coloneqq h_n \chi_{Z^C} \to f \chi_{Z^C},$$

which is $\mathcal{M}$-measurable. □

2. Integration

2.1. Simple functions

Definition 2.1
Let $E \subseteq X$ be any subset. The **characteristic function** of E is

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

Note that χ_E is measurable $\iff$ E is measurable.

Definition 2.2
A **simple function** is a finite linear combination of characteristic functions, e.g.,

$$\varphi = \sum_{j=1}^{N} a_j \chi_{E_j}, \qquad a_j \in \mathbb{R}.$$

Observe that φ is simple $\iff$ φ has finite range. Indeed, ($\implies$) is clear. For ($\impliedby$), we have that

$$\varphi = \sum_{k=1}^{M} b_k \chi_{\varphi^{-1}(\{b_k\})}, \tag{2.1}$$

where the b_k are distinct and $\{b_1, \ldots, b_M\} \subseteq \varphi(\mathbb{R}) - \{0\}$.
We call Equation 2.1 the **standard representation** of φ.

Remark 2.1. If φ is simple and $\varphi = \sum_{k=1}^{M} b_k \chi_{E_k}$ is the standard representation, then φ is measurable $\iff$ each E_k is measurable.

2.1.1. Approximating functions with simple functions

Proposition 2.2 (Dyadic simple estimation of f)
Let $f\colon X \to [-\infty, \infty]$ be measurable, and let $n \in \mathbb{N}$. Define

$$\begin{aligned} E_{n,-\infty} &:= \{f < -2^n\} \\ E_{n,\infty} &:= \{f \geq 2^n\} \\ E_{n,k} &:= \left\{ \frac{k}{2^n} \leq f < \frac{k+1}{2^n} \right\}. \end{aligned}$$

Set

$$f_n := -2^n \chi_{E_{n,-\infty}} + 2^n \chi_{E_{n,\infty}} + \sum_{k=-2^{2n}}^{2^{2n}-1} \frac{k}{2^n} \chi_{E_{n,k}}.$$

Then $\{f_n\}$ satisfy the following properties:

(a) $f_n \to f$ pointwise,

(b) f_n is increasing eventually on $f^{-1}([-M, \infty])$ for all $M > 0$,

(c) $f_n \to f$ uniformly on $f^{-1}([-M, M])$ for all $M > 0$.

We omit the proof of this.

2.2. Integrating non-negative functions

September 20, 2024 Let $(X, \mathcal{M}, \mu)$ be a measure space. We first begin with an incorrect definition for integration.

Definition 2.3 (Wrong definition of integration)
For $\varphi = \sum_{j=1}^{N} a_j \chi_{E_j} : X \to \mathbb{R}$ measurable and simple, define

$$\int \varphi \, d\mu := \sum_{j=1}^{N} a_j \mu(E_j). \tag{2.2}$$

The integral of a measurable function $f \colon X \to \overline{\mathbb{R}}$ is defined as

$$\int f \, d\mu := \lim_{n \to \infty} \int \varphi_n \, d\mu, \tag{2.3}$$

where φ_n is a sequence of measurable functions that converge to f.

Here are the issues: we need to prove Equation 2.3 is well-defined. Moreover, Equation 2.2 could include $\infty - \infty$, which is not defined.

In analysis, we have the example where

$$n\chi_{\left(\frac{1}{n}, \frac{2}{n}\right)} \to 0$$

pointwise, but $\int n\chi_{\left(\frac{1}{n}, \frac{2}{n}\right)} = 1$. This next definition resolved both of these issues.

Definition 2.4 (Correct definition of integration for non-negative functions)
Let

$$L^+ := \{f \colon X \to [0, \infty] : f \text{ is measurable}\}.$$

For $\varphi \in L^+$ that are simple, use the standard form $\sum_{j=1}^{n} a_j \chi_{E_j}$. We may define

$$\int \varphi \, d\mu := \sum_{j=1}^{N} a_j \mu(E_j). \tag{2.4}$$

For $f \in L^+$, we define

$$\int f \, d\mu := \sup \left\{ \int \varphi \, d\mu : \varphi \in L^+ \text{ simple and } \varphi \leq f \right\}. \tag{2.5}$$

We need to check that Equation 2.4 and Equation 2.5 agree when f is simple.

Proposition 2.3
Let $\varphi, \psi \in L^+$ be simple functions.

(a)
$$\int c\varphi \, d\mu = c \int \varphi \, d\mu, \qquad \text{for all } c \in [0, \infty).$$

(b)
$$\int \varphi + \psi \, d\mu = \int \varphi \, d\mu + \int \psi \, d\mu.$$

(c) If $\varphi \leq \psi$, then
$$\int \varphi \, d\mu \leq \int \psi \, d\mu.$$

(d)
$$\mu_\varphi(A) = \int_A \varphi \, d\mu := \int \varphi \chi_A \, d\mu$$

is a measure on $(X, \mathcal{M})$.

Remark 2.4. In (a) (and in the future), we use the convention $0 \cdot \infty = 0$.

(b) means that we may use any representation of $\varphi = \sum_j a_j \chi_{E_j}$, so long as the values $a_j \geq 0$.

(d) gives the fact that Equation 2.4 and Equation 2.5 agree on simple functions.

Proof of Proposition 2.3. (a) is immediate by the standard representation.

(b) Say $\varphi = \sum_{j=1}^N a_j \chi_{E_j}$, $\psi = \sum_{k=0}^M b_k \chi_{F_k}$ so that $a_0 = b_0 = 0$ (so $E_0 = \{\varphi = 0\}$, $F_0 = \{\psi = 0\}$). We have
$$\varphi + \psi = \sum_{j,k} (a_j + b_k) \chi_{E_j \cap F_k}.$$

The sets these characteristic functions are defined over are pairwise disjoint. Taking the union of sets $E_j \cap F_k$ where $a_j + b_k$ are equal, we can recover the standard representation. Hence,

$$\begin{aligned}
\int \varphi \, d\mu + \int \psi \, d\mu &= \sum_j a_j \mu(E_j) + \sum_k \mu(E_k) \\
&= \sum_j a_j \sum_k \mu(E_j \cap E_k) + \sum_k b_k \sum_j \mu(E_j \cap F_k) \\
&= \sum_{j,k} (a_j + b_k) \mu(E_j \cap F_k) \\
&= \sum_{\substack{c_\ell \in \{a_j + b_k\} \\ c_\ell \text{ distinct}}} c_\ell \sum_{\substack{j,k \\ a_j + b_k = c_\ell}} \mu(E_j \cap F_k) \\
&= \int \varphi + \psi \, d\mu.
\end{aligned}$$

(c) $\varphi \leq \psi \implies \psi - \varphi \in L^+$ and is simple. Notice
$$\int \varphi \, d\mu \leq \int \varphi \, d\mu + \int \psi - \varphi \, d\mu = \int \psi \, d\mu.$$

(d) $\mu_\varphi \geq 0$ is clear. Now we check countable additivity. Suppose

$$A = \bigsqcup_k A_k.$$

Then

$$\mu_\varphi(A) = \sum_{j=0}^{N} a_j \mu(A \cap E_j) = \sum_{j,k} a_j \mu(A_k \cap E_j)$$

□

Corollary 2.5

For $f \in L^+$ and $c \in [0, \infty)$,

$$\int cf \, d\mu = c \int f \, d\mu.$$

Proposition 2.6

For $f \in L^+$, $\int f \, d\mu = 0 \iff f = 0$ a.e.

Proof. ($\Longleftarrow$) If f simple, then this is true. If f is measurable, and $\varphi \leq f$ is simple, then $\varphi = 0$ a.e., so $\int \varphi \, d\mu = 0$. Taking the supremum, we get $\int f \, d\mu = 0$.

($\Longrightarrow$) Suppose $\{f \neq 0\}$ is not null. Equivalently,

$$0 < \mu(\{f > 0\}) = \mu\left(\bigcup_n \left\{f > \frac{1}{n}\right\}\right) = \lim_{n \to \infty} \mu\left(\left\{f > \frac{1}{n}\right\}\right)$$

by continuity from below, since this is an increasing union. So there exists an n such that

$$\mu\left(\left\{f > \frac{1}{n}\right\}\right) > 0.$$

Define the simple function

$$\varphi = \frac{1}{n}\chi_{\left\{f > \frac{1}{n}\right\}} \leq f.$$

Hence,

$$\int f \, d\mu \geq \int \varphi \, d\mu = \frac{1}{n}\mu\left(\left\{f > \frac{1}{n}\right\}\right) > 0.$$

□

Corollary 2.7

If $f, g \in L^+$ and $f \leq g$, then $\int f \, d\mu \leq \int g \, d\mu$. Moreover, if equality holds and both sides are finite, then $f = g$ a.e.

Theorem 2.8 (Monotone convergence theorem)

Let $\{f_n\} \subseteq L^+$ be an increasing (pointwise) sequence of functions. Then

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

Proof. Let $f = \lim_{n\to\infty} f_n$.

($\geq$) Notice that $f_n \leq f$ for all n, so $\int f_n \, d\mu \leq \int f \, d\mu$ for all n, so

$$\sup \int f_n \, d\mu \leq \int f \, d\mu.$$

($\leq$) Let $\varphi \in L^+$ be simple and satisfy $\varphi \leq f$. Let $0 < \alpha < 1$, and set $E_n^\alpha := \{x : f_n(x) \geq \alpha\varphi(x)\}$. Since the f_n are increasing, we have $f_m \geq \alpha\varphi$ on E_n^α for all $m \geq n$. Furthermore,

$$\bigcup_n E_n^\alpha = X.$$

Notice that for all $m \geq n$,

$$\int_{E_n^\alpha} f_m \, d\mu \geq \int_{E_n^\alpha} \alpha\varphi \, d\mu.$$

Hence,

$$\lim_{m\to\infty} \int_{E_n^\alpha} f_m \, d\mu \geq \lim_{n\to\infty} \int_{E_n^\alpha} \alpha\varphi \, d\mu.$$

Recall that integration over measurable sets of a simple function gives a measure. Hence, the RHS is equal to (by continuity from below)

$$\int \alpha\varphi \, d\mu.$$

Sending $\alpha \to 1$, we have

$$\int \lim_{m\to\infty} f_m \, d\mu \geq \int \varphi \, d\mu$$

for all simple functions $\varphi \in L^+$ bounded by f. Taking the sup over all simple functions, we get

$$\lim_{m\to\infty} \int f_m \, d\mu \geq \int f \, d\mu. \qquad \square$$

September 23, 2024

Theorem 2.9 (Swapping sum and integral)
Let $\{f_n\} \subseteq L^+$. Then $\int \sum_n f_n \, d\mu = \sum_n \int f_n \, d\mu$.

Proof. Case I: two functions. Let $\{\varphi_n\}, \{\psi_n\} \subseteq L^+$ be simple functions such that $\varphi_n \nearrow f_1$, $\psi_n \nearrow f_2$. Then $\varphi_n + \psi_n \nearrow f_1 + f_2$. Hence,

$$\begin{aligned}
\int f + g \, d\mu = \lim_{n\to\infty} \int \varphi_n + \psi_n \, d\mu &= \lim_{n\to\infty} \left(\int \varphi_n \, d\mu + \int \psi_n \, d\mu \right) \\
&= \lim_{n\to\infty} \left(\int \varphi_n \, d\mu \right) + \lim_{n\to\infty} \left(\int \psi_n \, d\mu \right) \\
&= \int f_1 \, d\mu + \int f_2 \, d\mu.
\end{aligned}$$

Case II: general case. Set $F_N := \sum_{n=1}^N f_n$. Notice that the F_N are increasing because

$f_n \in L^+$. Then

$$\sum_n \int f_n \, d\mu = \lim_{N\to\infty} \sum_{n=1}^{N} \int f_n \, d\mu = \lim_{N\to\infty} \int F_N \, d\mu \overset{(2.8)}{=} \int \lim_{n\to\infty} F_N \, d\mu = \int \sum_n f_n \, d\mu.$$

□

Proposition 2.10

If $\{f_n\} \subseteq L^+$ decreasing and $\int f_n \, d\mu < \infty$, then $\lim_{n\to\infty} \int f_n \, d\mu = \int \lim_{n\to\infty} f_n \, d\mu$.

We won't prove this because it will eventually be a consequence of the dominated convergence theorem.

Corollary 2.11 (Fatou's lemma)

If $\{f_n\} \subseteq L^+$, then

$$\int \liminf_{n\to\infty} f_n \, d\mu \le \liminf_{n\to\infty} \int f_n \, d\mu.$$

Remark 2.12. To remember the direction of this lemma, note the following inequality for sequences: $\liminf a_n + \liminf b_n \le \liminf(a_n + b_n)$.

Proof of Corollary 2.11.

$$\begin{aligned} \int \liminf_{n\to\infty} f_n \, d\mu = \int \lim_{n\to\infty} \left(\inf_{k\ge n} f_k \right) d\mu &= \lim_{n\to\infty} \int \inf_{k\ge n} f_k \, d\mu \\ &\le \lim_{n\to\infty} \left(\inf_{k\ge n} \int f_k \, d\mu \right) \\ &= \liminf_{n\to\infty} \int f_n \, d\mu. \end{aligned}$$

□

2.3. Integration in general

Definition 2.5

$f \colon X \to \overline{\mathbb{R}}$ is **integrable** if it is (1) measurable and (2) $\int f_+ \, d\mu < \infty$ and $\int f_- \, d\mu < \infty$, where

$$f_+ := \max\{f, 0\}, \qquad f_- := \max\{-f, 0\},$$

(so $f = f_+ - f_-$). We define

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

$f \colon X \to \mathbb{C}$ is **integrable** if $\operatorname{Re} f$ and $\operatorname{Im} f$ are, in which case,

$$\int f \, d\mu := \int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu.$$

We now try to make a space of functions out of this definition. Note that this definition will change later as we allow functions that are equal up to a null set to be considered the "same thing."

Define

$$L^1(X) = L^1(X, d\mu) = L^1(X, \mu) = L^1(\mu) = L^1 := \left\{f \colon X \to \mathbb{C} \text{ or } \overline{\mathbb{R}} : f \text{ integrable}\right\}.$$

Proposition 2.13

(a) If $f \in L^1$, $\{f \neq 0\}$ is σ-finite.

(b) If $f, g \in L^1$, the following are equivalent:

(i) $\int_E f\,d\mu = \int_E g\,d\mu$ for every measurable set E,

(ii) $f = g$ a.e.,

(iii) $\int |f - g|\,d\mu = 0$.

Proof. (a) Case I: $f \in L^+$. Then $\{f \neq 0\} = \bigcup_n \left\{f \geq \frac{1}{n}\right\}$. So

$$\infty > \int f\,d\mu \geq \int_{\left\{f \geq \frac{1}{n}\right\}} \frac{1}{n}\,d\mu = \frac{1}{n}\mu\left(\left\{f \geq \frac{1}{n}\right\}\right).$$

So $\{f \neq 0\}$ is the countable union of finite measure sets.

General case: This follows from case 1 applied to

$$(\operatorname{Re} f)_\pm, \qquad (\operatorname{Im} f)_\pm.$$

(b) ((ii) $\iff$ (iii)) we proved in class, since $f = g$ a.e. if and only if $|f - g| = 0$ a.e.

((ii) $\implies$ (i)) is clear by taking the positive negative parts.

((i) $\implies$ (ii)) WLOG, assume f and g are real-valued. We prove the contrapositive by assuming $\mu(\{f \neq g\}) > 0$. Notice that we can rewrite

$$\{f \neq g\} = \{f > g\} \cup \{g > f\} = \left(\bigcup_n \left\{f \geq g + \frac{1}{n}\right\}\right) \cup \left(\bigcup_n \left\{g \geq f + \frac{1}{n}\right\}\right).$$

This countable union has positive measure, so the sets in the union are not all null. Suppose $E := \left\{f \geq g + \frac{1}{n}\right\}$ has positive measure. Then

$$\int_E f\,d\mu \geq \int_E g + \frac{1}{n}\,d\mu = \int_E g\,d\mu + \frac{1}{n}\mu(E) \neq 0.$$

So $\int_E f\,d\mu \neq \int_E g\,d\mu$. $\square$

We define an equivalence relation $f \sim g$ if $f = g$ a.e. Indeed, this is an equivalence relation, and $f \sim f$ and $f \sim g \implies g \sim f$ are clear. Transitivity is also done since the union of null sets is null.

Proposition 2.14
$L^1(X)/\sim$ is a vector space over $\mathbb{C}$, and

$$\|[f]\|_{L^1} \coloneqq \|f\|_{L^1} \coloneqq \int |f| \, d\mu$$

is a norm on a vector space, and

$$f \mapsto \int f \, d\mu$$

is a continuous linear map of $L^1(X)/\sim$ into $\mathbb{C}$.

Proof. Notice that this the equivalence classes are preserved under linear operations: if $f_1 \sim f_2$ and $g_1 \sim g_2$, and $\alpha, \beta \in \mathbb{C}$, then

$$\alpha f_1 + \beta g_1 \sim \alpha f_2 + \beta g_2.$$

So it makes sense to define

$$\alpha[f] + \beta[g] \coloneqq [\alpha f + \beta g],$$

making $L^1(X)/\sim$ a $\mathbb{C}$-vector space.

The norm is well-defined by Proposition 2.13. We check the properties of a norm: (i) $\|[f]\|_{L^1} = 0 \implies f \equiv 0$ (by the previous proposition). (ii) $\|\alpha[f]\|_{L^1} = |\alpha| \, \|[f]\|_{L^1}$ by linearity of the integral L^1 and $|\alpha f| = |\alpha||f|$. (iii) $\|[f+g]\|_{L^1} \leq \|[f]\|_{L^1} + \|[g]\|_{L^1}$, since $|f+g| \leq |f| + |g|$ holds pointwise.

($f \mapsto \int f \, d\mu$) For linearity, we may split it into real and imaginary parts, and then positive and negative parts, and use linearity in each part.

For continuity, we use the following lemma:

Lemma 2.15
$\left|\int f \, d\mu\right| \leq \int |f| \, d\mu$.

Proof. If $\int f \, d\mu = 0$, then we are done. Otherwise, let $\alpha = \frac{\overline{\int f \, d\mu}}{\left|\int f \, d\mu\right|}$. Then

$$\left|\int f \, d\mu\right| = \operatorname{Re}\left(\alpha \int f \, d\mu\right) = \int \operatorname{Re}(\alpha f) \, d\mu \overset{|\alpha|=1}{\leq} \int |f| \, d\mu. \qquad \blacksquare$$

Combining this lemma with linearity, we get that the map is Lipschitz. $\square$

Proposition 2.16
If $(X, \overline{\mathcal{M}}, \overline{\mu})$ is the completion of $(X, \mathcal{M}, \mu)$, then $L^1(\mu) \cong L^1(\overline{\mu})$.

3. Modes of convergence

Definition 3.1

Let $\{f_n\}$ be a sequence of real or complex-valued measurable functions on $(X, \mathcal{M}, \mu)$.

1. $f_n \to f$ **almost everywhere (a.e.)** if $f_n(x) \to f(x)$ for all $x \in E$ for some E such that $\mu(E^C) = 0$. In other words, the set of all x where $f_n(x) \not\to f(x)$ is contained in a null set.

2. $f_n \to f$ **in measure** if for any chosen $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu(\{|f_n - f| > \varepsilon\}) = 0.$$

3. $f_n \to f$ **in L^1** if

$$\lim_{n\to\infty} \|f_n - f\|_{L^1(\mu)} = 0.$$

We will now present a slew of counterexamples.

Example 3.1 –

1. (Convergence a.e. and in measure, but not in L^1) Let $f_n := n\chi_{\left(\frac{1}{n},\frac{2}{n}\right)}$. It pointwise converges to 0 everywhere and is only nonzero in a segment of measure $\frac{1}{n}$, so it converges a.e. and in measure, but $\|f_n\|_{L^1} = 1$ for all n.

2. (Convergence a.e. only) $f_n := \chi_{[n,\infty)}$.

3. (Convergence in L^1 and in measure, but not a.e.) Consider the sequence $f_1 = \chi_{[0,1]}$, $f_2 = \chi_{[0,\frac{1}{2}]}$, $f_3 = \chi_{[\frac{1}{2},1]}$, $f_4 = \chi_{[0,\frac{1}{3}]}$ and so on. No point in $[0, 1]$ will ever eventually always be zero, but the integral and set where the measure is greater than any ε converges like $\frac{1}{n}$.

4. (Convergence a.e., in measure, and in L^1, but not uniformly) This is the classic analysis example: $f_n := \chi_{[0,\frac{1}{n}]}$.

3.1. The dominated convergence theorem and related results

Theorem 3.1 (Dominated convergence theorem)

Let (f_n) be a sequence in L^1 and assume that $f_n \to f$ a.e. *and* $\sup_n |f_n| \in L^1(\mu)$. Then $f_n \to f$ in L^1.

Remark 3.2. We note two things. The first is that $\sup_n |f_n| \in L^1$ is a *much* stronger assumption than $\sup_n \|f_n\|_{L^1} < \infty$. The second is that if the f_n's are measurable, then $\sup_n |f_n| \in L^1(\mu)$ if and only if there exists $g \in L^1(\mu)$ such that $|f_n| \leq g$ for all n. This motivates the name "dominated," since g dominates f.

Proof of Theorem 3.1. Taking the real and imaginary parts, we may assume that f_n's are real. Let $g = \sup_n |f_n|$. Consider the function

$$h_n := 2g - |f_n - f|.$$

Corollary 2.11 gives us

$$\begin{aligned} &\int \liminf h_n \, d\mu \le \liminf \int h_n \, d\mu \\ \implies &\int \liminf 2g \, d\mu \le \liminf \int 2g - |f_n - f| \, d\mu \\ \implies &\int 2g \, d\mu \le \liminf \int 2g - |f_n - f| \, d\mu \\ \implies &0 \le -\limsup \int |f_n - f| \, d\mu. \end{aligned}$$

Hence, $\lim_{n\to\infty} \int |f_n - f| \, d\mu = \|f_n - f\|_{L^1} = 0$. □

Proposition 3.3 (Swapping sum and integral in L^1)
If $\{f_j\} \subseteq L^1$ and $\sum_j \|f_j\|_{L^1} < \infty$, then $\sum_j f_j$ (the limit of partial sums) converges in L^1, and

$$\int \sum_j f_j \, d\mu = \sum_j \int f_j \, d\mu.$$

Proof. Let $g = \sum_j |f_j|$. By Theorem 2.8 applied to partial sums,

$$\int \sum_j |f_j| \, d\mu = \sum_j \int |f_j| \, d\mu.$$

So $g \in L^1$. In particular, g is finite a.e., so $\sum_j f_j$ converges *absolutely* a.e. Let $f = \sum_j f_j$. Partial sums of $\sum_j f_j$ are dominated by g, so apply Theorem 3.1 to see that $\sum_j f_j$ converges in L^1. □

Theorem 3.4
L^1 is complete.

Recall that a metric space being **complete** means that every Cauchy sequence converges.

September 27, 2024

Proof of Theorem 3.4. Let $\{f_n\} \subseteq L^1(\mu)$ be a Cauchy sequence. Then for all m, there exists an N_m such that $n_1, n_2 \ge N_m \implies \|f_{n_1} - f_{n_2}\| < 2^{-m}$. WLOG, N_m are increasing.

Consider the subsequence

$$f_{N_m} = f_{N_1} + \sum_{n=1}^{m-1} f_{N_{n+1}} - f_{N_n}.$$

This converges in L^1 by the Proposition 3.3. Let f be the limit. We have

$$f - f_{N_m} = \lim_{k\to\infty} f_{N_m} + \sum_{j=m}^{k-1} f_{N_{j+1}} - f_{N_j} - f_{N_m} = \sum_{j=m}^{\infty} f_{N_{j+1}} - f_{N_j}.$$

Taking the norms on both sides and using triangle inequality, we have

$$\|f - f_{N_m}\| \leq \sum_{j=m}^{\infty} 2^{-j}.$$

For any $n \geq N_m$,

$$\|f - f_n\| \leq \|f - f_{N_m}\| + \|f_{N_m} - f_n\| \leq \sum_{j=m}^{\infty} 2^{-j} + 2^{-m} \leq 2^{-m+2} \to 0. \qquad \square$$

Proposition 3.5

$I = [0, 1]$

Let $f \colon X \times I \to \mathbb{C}$. Assume $f(\,\cdot\,, t) \in L^1(\mu)$ for all t and define

$$F(t) := \int_X f(x, t)\, d\mu(x).$$

(a) If $f(x, \cdot\,)$ is continuous in t for all x and there exists a dominating function $g \in L^1(X)$ such that $|f(x, t)| \leq g(x)$ for all t, then F is continuous in t.

(b) (Differentiating under the sign) If $\frac{\partial f}{\partial t}(x, t)$ exists for every $(x, t) \in X \times I$ and there exists $h \in L^1(X)$ such that $\left|\frac{\partial f}{\partial t}(x, t)\right| \leq h(x)$ for all x, t, then F is differentiable and

$$F'(t) = \int \frac{\partial f}{\partial t}(x, t)\, d\mu(x).$$

Proof. (a) We need to show that for all $t_0 \in I$,

$$\lim_{\substack{t \to t_0 \\ t \in I}} F(t) = F(t_0).$$

Equivalently, for all sequences $\{t_n\} \subseteq I \setminus \{t_0\}$ with $t_n \to t_0$,

$$\lim_{n \to \infty} F(t_n) = F(t_0).$$

This is a direct consequence of Theorem 3.1.

(b) We need to show for all sequences $\{t_n\} \subseteq I \setminus \{t_0\}$ with $t_n \to t_0$,

$$\lim_{n \to \infty} \frac{F(t_n) - F(t_0)}{t_n - t_0} = \int \frac{f(x, t_n) - f(x, t_0)}{t_n - t_0}\, d\mu(x).$$

Now we use Theorem 3.1 and the mean value theorem to show for all x,

$$\left|\frac{f(x, t_n) - f(x, t_0)}{t_n - t_0}\right| \leq \sup_{\substack{s \in [t_0, t_n] \\ (\text{or } [t_n, t_0])}} \left|\frac{\partial f}{\partial t}(x, s)\right| \leq h(x). \qquad \square$$

3.1.1. Application: Fourier transforms

For $f \in L^1(\mathbb{R})$, define the **Fourier transform** as

$$\widehat{f}(\xi) := \int_{\mathbb{R}} e^{-2\pi i x \xi} f(x)\, dx.$$

For the following proof, notice that Proposition 3.5 certainly holds for replacing I with any closed, bounded interval.

Proposition 3.6
$\widehat{f}(\xi)$ is a bounded continuous function. Moreover, if $f, xf \in L^1(\mathbb{R})$, then $\widehat{f}$ is differentiable, and

$$\widehat{f}'(\xi) = -2\pi i\widehat{(xf)}(\xi).$$

Proof. $\widehat{f}(\xi)$ is bounded because

$$\left|\int_{\mathbb{R}} e^{-2\pi i x\xi} f(x)\,dx\right| \le \int_{\mathbb{R}} \left|e^{-2\pi i x\xi} f(x)\right| dx \le \int_{\mathbb{R}} |f(x)|\,dx < \infty.$$

$e^{-2\pi i x\xi} f(x)$ is clearly continuous in ξ on $[-M, M]$. Moreover,

$$\left|e^{-2\pi i x\xi} f(x)\right| \le |f(x)|,$$

so $|f(x)|$ dominates $e^{-2\pi i x\xi} f(x)$ for all ξ. So using part (a) of Proposition 3.5, we have that $\widetilde{f}(\xi)$ is continuous for all $\xi \in [-M, M]$. Taking $M \to \infty$ finishes.

On $[-M, M]$, we have that

$$\frac{\partial}{\partial \xi} e^{-2\pi i x\xi} f(x) = -2\pi i \left(e^{-2\pi i x\xi} x f(x)\right).$$

We dominate this function by

$$\left|-2\pi i \left(e^{-2\pi i x\xi} x f(x)\right)\right| \le 2\pi |xf(x)|,$$

which is in L^1 by assumption. So

$$\widehat{f}'(\xi) = \int_{\mathbb{R}} -2\pi i e^{-2\pi i x\xi} x f(x)\,dx = -2\pi i\widehat{(xf)}(\xi).$$

Taking $M \to \infty$ finishes. ■

3.1.2. Density of some classes of functions

Proposition 3.7
Simple functions are dense in $L^1(\mu)$.

Proof. Taking real and imaginary parts, then positive and negative parts, and using the triangle inequality, it suffices to show that for every $f \in L^+(\mu) \cap L^1(\mu)$, there exist simple functions $\{\varphi_n\}$ such that $\varphi_n \to f$ in L^1.

Recall the definition

$$\int f\,d\mu = \sup\left\{\int \varphi\,d\mu : \varphi \in L^+ \text{ simple}, \varphi \le f\right\}.$$

Take φ_n such that $\int \varphi_n \to \int f$. Hence,

$$\int f - \varphi_n \, d\mu \to 0 \implies \int |\varphi_n - f| \, d\mu \to 0. \qquad \square$$

Define the **support** of $f\colon X \to \mathbb{R}$ to be the closure of the set where f is nonzero, i.e., $\operatorname{supp}(f) := \overline{\{x : f(x) \neq 0\}}$. We are interested in functions with *compact* support. By Heine-Borel, this just means that the support is bounded; i.e. the function is equivalently 0 after some bounded interval.

Proposition 3.8

Continuous functions with compact support are dense in $L^1(\mathbb{R})$.

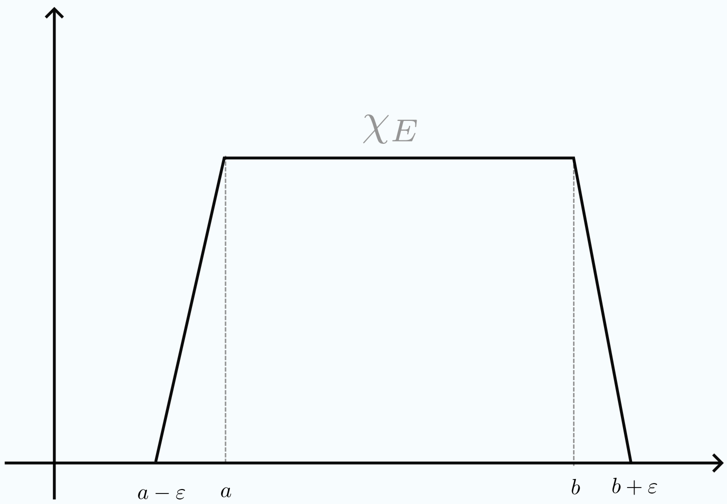
Proof. By the previous proposition and the triangle inequality, it suffices to show that for all $\varepsilon > 0$ and $E \subseteq \mathbb{R}$ such that $\lambda(E) < \infty$, there exist φ continuous with compact support such that

$$\|\varphi - \chi_E\|_{L^1} < \varepsilon.$$

We proved that $\lambda(E) < \infty$ implies that there exists a finite union of open intervals U such that

$$\lambda(U \triangle E) = \lambda(U \setminus E \cup E \setminus U) < \varepsilon.$$

Notice that this value is $\|\chi_U - \chi_E\|_{L^1}$, so we may assume that E is a finite union of bounded open intervals. By linearity, we may assume E is just one open interval, (a, b). We conclude by a picture:



$\square$

Remark 3.9. In fact, we can "smooth" out the ends to prove the previous proposition for C^∞ functions with compact support.

3.2. Other relations amongst modes of convergence

Proposition 3.10

If $f_n \to f$ in L^1, then $f_n \to f$ in measure.

Proof. Let $\varepsilon > 0$. Then

$$\varepsilon\mu\{|f_n - f| > \varepsilon\} = \varepsilon \int \chi_{\{|f_n - f| > \varepsilon\}} \, d\mu \leq \|f_n - f\|_{L^1} \xrightarrow{n\to\infty} 0. \qquad \square$$

September 30, 2024

Proposition 3.11
If $f_n \to f$ in measure, then some subsequence converges a.e.

Proof. We want to find a subsequence f_{n_k} such that $\mu(\{f_{n_k} \not\to f\}) > 0$.

$$\{f_{n_k} \not\to f\} = \bigcup_{m\in\mathbb{N}} \bigcap_{N\in\mathbb{N}} \bigcup_{k\geq N} \left\{|f_{n_k} - f| > \frac{1}{m}\right\}.$$

Recall that convergence in measure means that for all m,

$$\lim_{n\to\infty} \mu\left(\left\{|f_n - f| > \frac{1}{m}\right\}\right) = 0.$$

Equivalently, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\mu\left(\left\{|f_n - f| > \frac{1}{m}\right\}\right) < \varepsilon.$$

Equivalently, for all $\ell \in \mathbb{N}$, there exists an N_ℓ such that for all $n \geq N_\ell$,

$$\mu\left(\left\{|f_n - f| > 2^{-\ell}\right\}\right) < 2^{-\ell}.$$

WLOG, assume $N_1 < N_2 < \cdots$.
Consider the subsequence $\{f_{N_\ell}\}$. Given $\varepsilon = 2^{-\ell}$,

$$\mu\left(\bigcup_{\ell\geq L} \left\{|f_{n_\ell} - f| > 2^{-L}\right\}\right) \leq \sum_{\ell\geq L} 2^{-\ell} = 2^{-L+1}.$$

So

$$\{f_{n_\ell} \not\to f\} \subseteq \bigcap_{L\in\mathbb{N}} \bigcup_{\ell\geq L} \left\{|f_{n_\ell} - f| > 2^{-L}\right\}.$$

This set is null by the outer continuity of measure (the RHS has finite measure). $\square$

Corollary 3.12
Convergence in L^1 implies a.e. convergence in a subsequence.

Definition 3.2
$\{f_n\}$ is **Cauchy in measure** if for all $\varepsilon, \delta > 0$, there exists an $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$\mu\{|f_n - f_m| > \varepsilon\} < \delta.$$

We prove something similar to showing that $\mathbb{R}$ is complete.

Proposition 3.13

$\{f_n\}$ converges in measure $\iff$ $\{f_n\}$ is Cauchy in measure. Moreover, if $f_n \to f$ and $f_n \to g$ in measure, then $f = g$ a.e.

Proof. $(\implies)$

$$\begin{aligned} \{|f_n - f_m| > \varepsilon\} &\subseteq \left\{|f_n - f| > \frac{\varepsilon}{2}\right\} \cup \left\{|f_m - f| > \frac{\varepsilon}{2}\right\} \\ \implies \mu\{|f_n - f_m| > \varepsilon\} &\leq \mu\left\{|f_n - f| > \frac{\varepsilon}{2}\right\} + \mu\left\{|f_m - f| > \frac{\varepsilon}{2}\right\} \end{aligned}$$

$(\impliedby)$ We adapt the proof of the previous proposition [exercise]. If $\{f_n\}$ is Cauchy in measure, then along a subsequence $\{f_{n_k}\}_k$ it is Cauchy a.e., which implies convergence a.e. Let f be the limit.

Claim 3.1. $f_n \to f$ in measure.

Proof. Let $\varepsilon, \delta > 0$. Choose N such that $n, m \geq N$ implies $\mu\{|f_n - f_m| > \varepsilon\} < \delta$. In particular, $\mu\{|f_n - f_{n_k}| > \varepsilon\} < \delta$ for all n, k sufficiently large. Therefore,

$$\mu\{|f_n - f| > \varepsilon\} \leq \mu\left(\{f_{n_k} \not\to f\} \cup \{|f_n - f_{n_k}| > \varepsilon\}\right) < \delta. \qquad \blacksquare$$

For the second part, if $f_n \to f$ in measure, then $f_{n_k} \to f$ a.e. along a subsequence. Since $f_{n_k} \to g$ in measure, $f_{n_{k_j}} \to g$ a.e. But $f_{n_{k_j}} \to f$ a.e., so $f = g$ a.e. $\square$

Proposition 3.14

If μ is finite, and $f_n \to f$ a.e., then $f_n \to f$ in measure.

Proof. A.e. convergence is equivalent to

$$\bigcup_{m \in \mathbb{N}} \bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} \left\{|f_n - f| > \frac{1}{m}\right\}$$

being null. Equivalently, for all m,

$$\bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} \left\{|f_n - f| > \frac{1}{m}\right\}$$

is null. This is a decreasing set as $N \to \infty$. Since $\mu(X) < \infty$, continuity from above gives us

$$\begin{aligned} \mu\left(\bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} \left\{|f_n - f| > \frac{1}{m}\right\}\right) &= \lim_{N \to \infty} \mu\left(\bigcup_{n \geq N} \left\{|f_n - f| > \frac{1}{m}\right\}\right) \\ &\geq \lim_{N \to \infty} \left(\left\{|f_n - f| > \frac{1}{m}\right\}\right). \end{aligned} \qquad \square$$

In summary:

$$\begin{tikzcd} L^1 \arrow[r, Rightarrow] & \text{in measure} \arrow[r, bend left, Rightarrow, "\text{in some subsequence}"] & \text{a.e.} \arrow[l, bend left, Rightarrow, "\text{if } \mu(X)<\infty"] \end{tikzcd}$$

Theorem 3.15 (Egoroff's theorem)
If $f_n \to f$ a.e. and $\mu(X) < \infty$, then for all $\varepsilon > 0$, there exists $B_\varepsilon \subseteq X$ with $\mu(B_\varepsilon) < \varepsilon$ such that $f_n \to f$ uniformly on B_ε^C.

October 2, 2024

Proof. We know, since $\mu(X) < \infty$ for all m, there exists an N_m such that

$$\mu\left(\bigcup_{n \geq N_m} \left\{|f_n - f| > \frac{1}{m}\right\}\right) < 2^{-m}.$$

Set $B_{2^{-M+1}} := \bigcup_{m=M}^{\infty} \bigcup_{n \geq N_m} \left\{|f_n - f| > \frac{1}{m}\right\}$. Then

$$\mu(B_{2^{-M+1}}) < 2^{-M+1},$$

and if $x \notin B_{2^{-M+1}}$ and $n \geq N_m$, then $|f_n(x) - f(x)| \leq \frac{1}{m}$. □

4. Product measures

4.1. Construction

Given two measure space $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$, we already can create a measurable space $(X \times Y, \mathcal{M} \otimes \mathcal{N})$. Our goal with this section is to create a product measure $\mu \times \nu$.

Proposition 4.1

Let $\mathcal{A} := \left\{\bigcup_{j=1}^n A_j \times B_j : A_j \in \mathcal{M}, B_j \in \mathcal{N}\right\}$. Then $\mathcal{A}$ is an algebra, and if we define

$$\pi\left(\bigsqcup_{j=1}^n A_j \times B_j\right) := \sum_{j=1}^n \mu(A_i)\nu(B_j),$$

then π is well-defined premeasure on $\mathcal{A}$.

Proof. ($\mathcal{A}$ is an algebra) It's clear by a picture that $\mathcal{A}$ is closed under complements.

(Well-defined and premeasure) Suppose $A \times B = \bigsqcup_{j=1}^\infty A_j \times B_j$. We need to prove $\pi(A \times B) = \sum_{j=1}^\infty \pi(A_j \times B_j)$. We have

$$\begin{aligned}
\pi(A \times B) = \mu(A)\nu(B) &= \int_Y \mu(A)\chi_B(y)\, d\nu(y) \\
&= \int_Y \left(\int_X \chi_A(x)\chi_B(y)\, d\mu(x)\right) d\nu(y) \\
&= \int_Y \left(\int_X \chi_{A\times B}(x,y)\, d\mu(x)\right) d\nu(y) \\
&= \int_Y \left(\int_X \sum_{j=1}^\infty \chi_{A_j}(x)\chi_{B_j}(y)\, d\mu(x)\right) d\nu(y) \\
&= \sum_{j=1}^\infty \left(\int_Y \left(\int_X \chi_{A_j}(x)\chi_{B_j}(y)\, d\mu(x)\right) d\nu(y)\right) \\
&= \sum_{j=1}^\infty \mu(A_j)\nu(B_j),
\end{aligned}$$

where we moved the sum outside the integral by Theorem 2.8. $\square$

Corollary 4.2

π generates an outer measure $(\mu \times \nu)^*$ on $\mathcal{P}(X \times Y)$ such that $(\mu \times \nu)^*|_{\mathcal{M}\otimes\mathcal{N}} =: \mu \times \nu$ is a measure and $(\mu \times \nu)|_{\mathcal{A}} = \pi$.

Further, if μ, ν are σ-finite, then $\mu \times \nu$ is σ-finite and $\mu \times \nu$ is the *unique* measure on $\mathcal{M} \otimes \mathcal{N}$ extending π.

Proposition 4.3 (Slicing)
For $E \in \mathcal{M} \otimes \mathcal{N}$, and $f\colon X \times Y \to \mathbb{C}$ that is $\mathcal{M} \otimes \mathcal{N}$ measurable,

1. each **x-slice** $E_x := \{y : (x, y) \in E\}$ is in $\mathcal{N}$ for all x, and each **y-slice** $E^y := \{x : (x, y) \in E\}$ is in $\mathcal{M}$ for all y.
2. f_x is $\mathcal{N}$-measurable and f^y is $\mathcal{M}$-measurable, where

$$f_x(y) := f^y(x) := f(x, y).$$

Proof. (a) Let

$$\mathscr{P} := \{E \subseteq X \times Y : \text{all } x\text{-slices and } y\text{-slices are measurable}\}.$$

Notice that $\mathscr{P}$ contains measurable rectangles $A \times B$, and $\mathscr{P}$ is a σ-algebra, hence $\mathscr{P} \supseteq \mathcal{M} \otimes \mathcal{N}$.

This containment is usually strict.

(b) follows directly from (a). □

Theorem 4.4
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces. For $E \in \mathcal{M} \otimes \mathcal{N}$, the maps

$$x \mapsto \nu(E_x),$$
$$y \mapsto \mu(E^y),$$

are measurable and

$$(\mu \times \nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y).$$

October 4, 2024

Proof. Let $\mathcal{C} = \{E \subseteq X \times Y : \text{the theorem holds for } E\}$. We want to show that $\mathcal{C} \supseteq \mathcal{M} \otimes \mathcal{N}$. It is clear that $\mathcal{C}$ contains the algebra generated by measurable rectangles. We now show $\mathcal{C}$ is a σ-algebra. For now, assume that μ and ν are finite measures. If $E = \bigcup_n E_n$ *with* E_n *increasing*, then

$$\nu(E_x) = \nu\left(\bigcup_n (E_n)_x\right) = \lim_{n\to\infty} \nu((E_n)_x).$$

Similarly,

$$(\mu \times \nu)(E) = \lim_{n\to\infty} (\mu \times \nu)(E_n) = \lim_{n\to\infty} \int \nu((E_n)_x)\, d\mu(x).$$

Now, since the integrand is bounded by $\nu(Y)$, we may swap the limit and the integral by Theorem 3.1 to obtain

$$(\mu \times \nu)(E_x) = \int \lim_{n\to\infty} \nu((E_n)_x)\, d\mu(x) = \int \nu(E_x)\, d\mu(x).$$

A similar argument shows that $\mathcal{C}$ is closed under countable *deceasing intersections*. These two properties make $\mathcal{C}$ a **monotone class**. By the monotone class lemma (coming after this), $\mathcal{C}$ is a σ-algebra. The general σ-finite case follows by taking countable increasing unions of finite measure subsets. □

Lemma 4.5 (Monotone class lemma)
If $\mathcal{A} \subseteq \mathcal{P}(X)$ is an algebra, then the monotone class $\mathcal{C}$ generated by $\mathcal{A}$ coincides with the σ-algebra $\mathcal{M}$ generated by $\mathcal{A}$.

I omit the proof because it is mainly set theoretic. It is in our textbook.

4.2. Fubini-Tonelli

October 7, 2024

Theorem 4.6 (Fubini-Tonelli)
If $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite:

(a) (Tonelli) $f \in L^+(X \times Y)$ implies $\int_Y f_x(y)\, d\nu(y)$, $\int_X f^y(x)\, d\mu(x)$ are in $L^+(\mu)$ and $L^+(\nu)$ respectively, and

$$\int f\, d(\mu \times \nu) = \int_X \left(\int_Y f_x(y)\, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f^y(x)\, d\mu(x) \right) d\nu(y). \tag{4.1}$$

(b) (Fubini) $f \in L^1(X \times Y)$ implies $f_x \in L^1(Y, \nu)$ μ-a.e. and $f^y \in L^1(X, \mu)$ ν-a.e. Let $g(x) := \int_Y f_x(y)\, d\nu(y)$ and $h(x) := \int_X f^y(x)\, d\mu(x)$. Then $g \in L^1(\mu)$, $h \in L^1(\nu)$ and Equation 4.1 holds.

Proof. (a) is done for characteristic functions by Theorem 4.4. So we also have it for linear combinations of characteristic functions with positive coefficients, because all statements are linear. Finally, limits of increasing positive sequences of positive simple functions follows from the fact that limits of measurable sequences are measurable and Theorem 2.8.

(b) Apply (a) to $(\operatorname{Re} f)_\pm$ and $(\operatorname{Im} f)_\pm$. Hence, it suffices to prove the statements for $f \in L^+(X \times Y) \cap L^1(X \times Y)$. The new statements are that $f_x \in L^1(\nu)$ μ-a.e. and $f^y \in L^1(\mu)$ ν-a.e. But both follow from (a), which imply $g \in L^1(\mu)$, $h \in L^1(\nu)$, and

$$\begin{aligned} g \in L^1(\mu) &\implies g(x) < \infty, \mu\text{-a.e.} \\ h \in L^1(\mu) &\implies h(x) < \infty, \mu\text{-a.e.} \end{aligned}$$

□

Theorem 4.7 (Fubini-Tonelli for complete measures)
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be complete, σ-finite measures. If f is $\overline{\mathcal{M} \otimes \mathcal{N}}$-measurable and $f \in L^+$ or $f \in L^1(\overline{\mu \times \nu})$, then f_x is ν-measurable for μ-a.e. x, and f^y is μ-measurable for ν-a.e. y. Moreover, g and h from Theorem 4.6 are μ and ν-measurable respectively, and

$$\int f\, d(\overline{\mu \times \nu}(x, y)) = \int g\, d\mu(x) = \int h\, d\nu(y).$$

4.2.1. Application: distribution functions

Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space. Let $f \in L^1(\mu)$. Define the **distribution function** of f as

$$\lambda_f(\alpha) := \mu\{|f| > \alpha\}, \qquad \alpha \in [0, \infty).$$

This is a decreasing function, hence measurable. Indeed,

$$\lambda_f^{-1}([0, m])$$

is an interval, hence Borel for all m.

Proposition 4.8
$\|f\|_{L^1(\mu)} = \int_0^\infty \lambda_f(t)\,dt$.

Proof.

$$\begin{aligned}\int_X |f(x)|\,d\mu(x) &= \int_X \int_{[0,\infty)} \chi_{[0,|f(x)|)}(t)\,dt\,d\mu(x)\\ &= \iint_{X\times[0,\infty)} \chi_{\{(x,t):t<|f(x)|\}}(x,t)\,d(\mu\times\lambda)(x,t)\\ &\overset{(4.6)}{=} \int_{[0,\infty)} \int_X \chi_{\{x:|f(x)|>t\}}\,d\mu(x)\,dt\\ &= \int_{[0,\infty)} \lambda_f(t)\,dt.\end{aligned}$$

□

4.3. Change of variables in Lebesgue measure

Definition 4.1 (n-dimensional Lebesgue measure)
Let $(\mathbb{R}^n, \mathcal{L}^n, \lambda^n) := (\mathbb{R}\times\cdots\times\mathbb{R}, \mathcal{L}\otimes\cdots\otimes\mathcal{L}, \lambda\times\cdots\times\lambda)$. Suppose f is in $L^+(\lambda^n)\cup L^1(\lambda^n)$, and $E\in\mathcal{L}^n$. Say

$$\int f(x)\chi_E(x)\,d\lambda^n(x) =: \int_E f(x)\,dx,$$

and $\lambda^n(E) =: |E|$ when $E\in\mathcal{L}^n$ and the notation is clear from the context.

Definition 4.2
Let $\Omega\subseteq\mathbb{R}^n$ be open. $G\colon\Omega\to\mathbb{R}^n$ (or $U=G(\Omega)\subseteq\mathbb{R}^n$) is a **$C^1$-diffeomorphism** if

(a) G is a bijection to $G(\Omega)$,

(b) G and its derivatives are defined and continuous on Ω,

(c) $\det(DG)\neq 0$ on Ω.

This implies that G^{-1} exists and has continuous partials on $G(\Omega)$.

Theorem 4.9 (Change of variables)
If $G\colon \Omega \to \mathbb{R}^n$ is a C^1-diffeomorphism,

(a) If $f\colon G(\Omega) \to \mathbb{C}$ is measurable, then $G^*f := f \circ G\colon \Omega \to \mathbb{C}$ is measurable. If $f \in L^+(dx) \cup L^1(dx)$, then

$$\int_{G(\Omega)} f(x)\,dx = \int_\Omega (f \circ G)(x)\,|\det DG(x)|\,dx.$$

(b) If $E \subseteq \Omega$ is measurable, then $G(E) \subseteq G(\Omega)$ is measurable, and the measure of $G(E)$ can be expressed in any of the following ways:

$$\begin{aligned}|G(E)| = \int_{G(\Omega)} \chi_{G(E)}(y)\,dy &= \int_\Omega (\chi_{G(E)} \circ G)(x)\,|\det DG(x)|\,dx \\ &= \int_E |\det DG(x)|\,dx.\end{aligned}$$

Proof (outline). We present the main steps in the proof. The full proof is in the textbook.

Lemma 4.10
For $E \subseteq \mathbb{R}^n$, λ^n-measurable,

(i)

$$\begin{aligned}|E| &= \inf\left\{\sum_j |R_j| : E \subseteq \bigcup_{j=1}^\infty R_j, R_j \text{ a (proper) rectangle}\right\} \\ &= \inf\{|U| : E \subseteq U \text{ open}\} \\ &= \sup\{|K| : E \supseteq K \text{ compact}\}.\end{aligned}$$

(ii) We can write $E = F \cup Z_1 = G \setminus Z_2$, where $F \in F_\sigma$, $G \in G_\delta$, and Z_1, Z_2 are null sets.

(iii) If $|E| < \infty$, for all $\varepsilon > 0$, there exists a finite collection of disjoint rectangles $\{R_j\}_{j=1}^N$ such that $\left|E \triangle \bigcup_j R_j\right| < \varepsilon$.

Proof. (i) Use the fact that rectangles generate $\mathcal{B}_{\mathbb{R}^n}$ and $\mathcal{L}^n = \mathcal{B}_{\mathbb{R}^n}$ with the uniqueness of Carathéodory.

For (ii) and (iii), these facts are true for $\mathbb{R}^n$, and the product of open (resp. compact) sets are open (resp. compact). ■

Lemma 4.11
The theorem holds for translations: $G(x) = x + x_0$.

Proof. G preserves the measure of rectangles, so apply the previous proposition. □

Lemma 4.12

The theorem holds for $G \in GL_n(\mathbb{R})$.

Proof. It suffices to prove the result for when G swaps entries i and j, when G multiplies entry j by $\lambda \in \mathbb{R}$, and when G adds entry j to entry k. The first two work for rectangles. The last works for rectangles using Fubini (4.6). ■

Lemma 4.13

For all $K \subseteq \Omega$ compact, $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $\rho < \delta$, $b \in K$ and $a \in \mathbb{R}^n$ with $\rho/2 < b_j - a_j < \rho$,

$$\begin{aligned} G(b) + (1-\varepsilon)DG(b)\prod_j (a_j - b_j, 0] \\ \subseteq G\left(\prod_j (a_j, b_j]\right) \\ \subseteq G(b) + (1+\varepsilon)DG(b)\prod_j (a_j - b_j, 0]. \end{aligned}$$

Lemma 4.14

If $|E| = 0$, then $|G(E)| = 0$.

Proof. Cover E with small rectangles. Then $G(E)$ is covered by small n-parallelepipeds. ■

Lemma 4.15

If $E \in \mathcal{L}^n$, $G(E) \in \mathcal{L}^n$.

Proof. If $E = B \cup Z$ for Borel B and null Z, then $G(B) \in \mathcal{B}_{\mathbb{R}^n}$, since G is a homeomorphism and $G(Z)$ is null by the previous lemma. ■

Using the last lemma and Lemma 4.13, we know the theorem holds for finite unions of rectangles.

Adding Lemma 4.10 implies (b). This gives (a) for simple functions. But then (a) holds for monotone limits of simple functions in L^+, which implies the result in (a). □

4.3.1. Application: polar coordinates on $\mathbb{R}^n$

Let $\|\cdot\|$ be the Euclidean norm.

$$U = \left\{\mathbf{u} \in \mathbb{R}^{n-1} : \|\mathbf{u}\| < 1\right\}$$
$$S_+^{n-1} = \{\omega \in \mathbb{R}^n : \|\omega\| = 1, \omega_n > 0\}.$$

We may define a measure on $(S_+^{n-1}, \mathcal{B}_{S_+^{n-1}})$ by

$$\sigma_+^{n-1}(E) := \int_U \chi_E\left(\mathbf{u}, \sqrt{1-\|\mathbf{u}\|^2}\right) \frac{d\mathbf{u}}{\sqrt{1-\|\mathbf{u}\|^2}}.$$

Define

$$\begin{aligned} G\colon U \times (0,\infty) &\to \mathbb{R}^n_+ := \{\mathbf{x} : x_n > 0\}, \\ (\mathbf{u}, \rho) &\mapsto (\rho\mathbf{u}, \rho\sqrt{1-\|\mathbf{u}\|^2}). \end{aligned}$$

Since G is continuously differentiable, we may compute

$$\begin{aligned} \det DG = \det \begin{bmatrix} \rho\mathbb{I}_{n-1} & \frac{-\rho\mathbf{u}}{\sqrt{1-\|\mathbf{u}\|^2}} \\ \mathbf{u} & \sqrt{1-\|\mathbf{u}\|^2} \end{bmatrix} &= \det \begin{bmatrix} \rho\mathbb{I}_{n-1} & \text{unimportant} \\ 0 & \sqrt{1-\|\mathbf{u}\|^2} + \frac{\|\mathbf{u}\|^2}{\sqrt{1-\|\mathbf{u}\|^2}} \end{bmatrix} \\ &= \frac{\rho^{n-1}}{\sqrt{1-\|\mathbf{u}\|^2}}. \end{aligned}$$

So, for $E \subseteq \mathbb{R}^n_+$,

$$\begin{aligned} |E| &= \int_0^\infty \int_U \chi_E(\rho\mathbf{u}, \rho\sqrt{1-\|\mathbf{u}\|^2}) \frac{\rho^{n-1}}{\sqrt{1-\|\mathbf{u}\|^2}} \, d\mathbf{u} \, d\rho \\ &= \int_0^\infty \int_{S^{n-1}_+} \chi_E(\rho\omega)\rho^{n-1} \, d\sigma^{n-1}_+(\omega) \, d\rho. \end{aligned}$$

Since the Lebesgue measure is rotationally invariant, we may extend σ^{n-1}_+ to a rotationally invariant measure σ^{n-1} on S^{n-1}. For all measurable f,

$$\|f\|_{L^1} = \int_0^\infty \int_{S^{n-1}} |f(\rho\omega)|\rho^{n-1} \, d\sigma(\omega) \, d\rho.$$

These are **polar coordinates**.

> **Example 4.1** – We may use polar coordinates to prove that
>
> $$\int_{\mathbb{R}^n} \frac{1}{(1+\|x\|^2)^{p/2}} \, dx < \infty \iff p > n.$$

5. Signed measures

Definition 5.1

A **signed measure** on $(X, \mathcal{M})$ is a function

$$\nu \colon \mathcal{M} \to \overline{\mathbb{R}}$$

such that

(a) $\nu(\varnothing) = 0$,

(b) $\{E_j\} \subseteq \mathcal{M}$ disjoint implies $\nu\left(\bigcup_j E_j\right) = \sum_j \nu(E_j)$,

(c) ν cannot take both $+\infty$ and $-\infty$ values (hence, the above sum is always ambiguous).

Example 5.1 – Let $f \in L^1(\mu, \overline{\mathbb{R}})$. Define a signed measure

$$\mu_f(A) \coloneqq \int_A f \, d\mu, \qquad A \in \mathcal{M}.$$

Example 5.2 – If μ_+ and μ_- are two measures on $(X, \mathcal{M})$, and one is finite, then

$$\mu \coloneqq \mu_+ - \mu_-$$

is a signed measure.

The main goals of this section are to reduce signed measure to measures via our various decomposition theorems. We want to decompose some signed measures into the forms of the two above examples.

Definition 5.2

If ν is a signed measure on $(X, \mathcal{M})$, then $E \in \mathcal{M}$ is **positive** (resp. **negative**, **null**) if $\nu|_E$ is a *(positive) measure* (i.e. for all $F \in \mathcal{M}$ such that $F \subseteq E$, $\nu(F) \geq 0$), resp. $(-\nu)|_E$ is a positive measure, resp. $\nu|_E \equiv 0$.

We put positive in parentheses because a measure is indeed positive.

5.1. Various decomposition theorems

Theorem 5.1 (Hahn decomposition theorem)

Let ν be a signed measure on $(X, \mathcal{M})$. Then there exists a decomposition $X = P \sqcup N$ such that P is positive and N is negative for ν. Moreover, this decomposition is unique in the sense that if $P' \sqcup N'$ is another decomposition, $P' \vartriangle P = N' \vartriangle N$ is null.

October 11, 2024

Proof. (Uniqueness) $P' \vartriangle P \subseteq P' \cup P$, so positive for ν, and $N' \vartriangle N \subseteq N' \cup N$, so negative for ν. Hence, these sets are null for ν.

(Existence) WLOG (otherwise consider $-\nu$), ν never takes the value $+\infty$. Define

$$\mathcal{P} := \{E \in \mathcal{M} : E \text{ positive for } \nu\}.$$

Notice that $\mathcal{P}$ is closed under subsets and unions. Define $M_+ := \sup\{\nu(E) : E \in \mathcal{P}\}$. So there exists a sequence $\{P_n\} \subseteq \mathcal{P}$ such that $\nu(P_n) \to M_+$. Take $P := \bigcup_n P_n$.

Since ν is a (positive) measure on P, $\nu(P) = M_+$ (this implies M_+ is finite since $\nu < +\infty$).

Lemma 5.2
$N := P^C$ is negative for ν.

Proof. Suppose not. Then there exists $E \subseteq N$ ($E \in \mathcal{M}$) such that $\nu(E) > 0$.

Case I: E is positive for ν. Then $E \cup P \in \mathcal{P}$, but $\nu(E \cup P) = \nu(E) + \nu(P) > M_+$.

Case II: There exists $B \subsetneq E$ such that $\nu(B) < 0$. Let $E_1 = E$ and suppose $B_1 \subsetneq E$ satisfies $\nu(B_1) < 0$. WLOG, suppose

$$-\nu(B_1) \geq \min\left\{1, \frac{1}{2}\sup\left\{-\nu(\tilde{B}) : \tilde{B} \subseteq E_1, \tilde{B} \in \mathcal{M}\right\}\right\}.$$

Iterate this process on $E_2 := E_1 \setminus B_1$, so

$$\nu(E_2) = \nu(E_1) - \nu(B_1) > \nu(E_1).$$

Given $E_n \subsetneq \cdots \subsetneq E_1$, if E_n positive for ν, we use Case I. Otherwise, we have a sequence $B_n \subsetneq E_n$ such that

$$-\nu(B_n) \geq \min\left\{1, \frac{1}{2}\sup\left\{-\nu(\tilde{B}) : \tilde{B} \subseteq E_n, \tilde{B} \in \mathcal{M}\right\}\right\}.$$

Let $E' = \bigcap_n E_n = E \setminus (\bigcup_n B_n)$. By countable additivity,

$$\nu(E') = \nu(E) - \sum_n \nu(B_n) \implies \sum_n \nu(B_n) < \infty,$$

which means

$$\lim_n \left(\sup\left\{-\nu(\tilde{B}) : \tilde{B} \subseteq E_n, \tilde{B} \in \mathcal{M}\right\}\right) = 0,$$

so $E \in \mathcal{P}$. But $\nu(E') > \nu(E) > 0$, so this is Case I. ■

Thus, we have a decomposition $X = P \sqcup N$. □

Definition 5.3

October 14, 2024

Let μ, ν be signed measures on $(X, \mathcal{M})$. μ and ν are **mutually singular**, denoted $\mu \perp \nu$, if there exist $E, F \in \mathcal{M}$ such that there is a decomposition $X = E \sqcup F$ where E is μ-null and F is ν-null.

Corollary 5.3 (Jordan decomposition theorem)
If ν is a signed measure on $(X, \mathcal{M})$, then there exists unique measures ν_+ and ν_-, one finite, such that $\nu = \nu_+ - \nu_-$ and $\nu_+ \perp \nu_-$.

Proof. Define for $E \in \mathcal{M}$,

$$\begin{aligned}\nu_+(E) &:= \nu(E \cap P),\\ \nu_-(E) &:= -\nu(E \cap N).\end{aligned}$$

The rest of the proof is left as an exercise. $\square$

Definition 5.4
$|\nu| := \nu_+ + \nu_-$ is called the **total variation** of ν.

Definition 5.5
Let ν be a signed measure and let μ be a (positive) measure on $(X, \mathcal{M})$. We say ν is **absolutely continuous** with respect to μ, denoted $\nu \ll \mu$ if $\mu(E) = 0$ implies $\nu(E) = 0$ for all such $E \in \mathcal{M}$.

Proposition 5.4
$\nu \ll \mu \iff |\nu| \ll \mu$.

Proof. ($\implies$) We decompose $\nu = \nu_+ - \nu_-$ by Corollary 5.3. Suppose $\mu(E) = 0$, so $\nu(E) = 0 \implies \nu_+(E) = \nu_-(E)$. Notice that $\nu_+(E) = \nu(P \cap E)$. But notice that $P \cap E \subseteq E$, so $\mu(P \cap E) \leq \mu(E) = 0$, hence $\nu(P \cap E) = \nu_+(P \cap E) = 0$, which implies the result.

($\impliedby$) $\mu(E) = 0 \implies \nu(E) = \nu_+(E) + \nu_-(E) = 0 \implies \nu_+(E) = \nu_-(E) = 0 \implies \nu(E) = \nu_+(E) - \nu_-(E) = 0$. $\square$

Proposition 5.5
If ν is finite, then $\nu \ll \mu \iff$ for all $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\mu(E) < \delta \implies |\nu(E)| < \varepsilon.$$

Here are some non-examples explaining why the finiteness assumption is important.

Non-Example 5.3 – Let μ be any measure. Define

$$\nu(E) = \begin{cases} 0 & \text{if } \mu(E) = 0,\\ \infty & \text{if } \mu(E) \neq 0.\end{cases}$$

This $\nu \ll \mu$, but it does not satisfy the delta-epsilon condition in the proposition. Of course, ν is not finite either.

Non-Example 5.4 – Let $f(x) := |x|$ for $x \in \mathbb{R}$ and $\nu := \lambda_f$, where the measure is given by $\lambda_f(A) := \int_A |f| \, d\nu$, and let μ be the Lebesgue measure. Then $\nu \ll \mu$, but given any set with nonzero μ measure, we can translate it to make it have arbitrarily large measure.

Proof of Proposition 5.5. By the Jordan decomposition theorem (5.3), we may assume that $\nu \geq 0$.

($\Longleftarrow$) is obvious.

($\Longrightarrow$) Suppose there exists $\varepsilon > 0$ such that for all n, there exists E_n such that $\mu(E_n) < 2^{-n}$ and $\nu(E_n) \geq \varepsilon$. Let

$$E \coloneqq \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m.$$

Then

$$\mu(E) \leq \mu\left(\bigcup_{m=n}^{\infty} E_m\right) \leq \sum_{m=n}^{\infty} 2^{-m} = 2^{-n+1}$$

for all n. So $\mu(E) = 0$. However,

$$\nu(E) = \lim_{n\to\infty} \nu\left(\bigcup_{m=n}^{\infty} E_m\right) \geq \limsup_{n\to\infty} \nu(E_n) \geq \varepsilon. \qquad \square$$

Theorem 5.6 (Lebesgue decomposition theorem)
Let ν be a σ-finite signed measure, and μ a measure on $(X, \mathcal{M})$. Then there exists a decomposition of ν into an absolutely continuous part and a mutually singular part:

$$\nu = \nu_a + \nu_s,$$

where ν_a and ν_s are σ-finite signed measures on $(X, \mathcal{M})$ such that $\nu_a \ll \mu$ and $\nu_s \perp \mu$.

Proof. Main case: $\nu \geq 0$ and μ and ν are finite. Consider the measures

$$\rho_n \coloneqq \mu - \frac{1}{n}\nu.$$

On each of these measures, we do the Hahn decomposition to get

$$X = P_n \sqcup N_n$$

for all n, up to sets of measure zero. Since it becomes easier to be a positive set as n increases, the P_n are increasing and the N_n are decreasing. Let

$$P \coloneqq \bigcup_n P_n, \qquad N \coloneqq \bigcap_n N_n = P^C,$$

and define

$$\nu_a(E) \coloneqq \nu(E \cap P), \qquad \nu_s(E) \coloneqq \nu(E \cap N).$$

Then $\nu = \nu_a + \nu_s$ and $\nu_a \perp \nu_s$.

Claim 5.1. $\nu_s \perp \mu$.

> **Proof.** It suffices to show that N is μ-null. Let $E \subseteq N$, hence $E \subseteq N_n$ for all n. Then $\rho(E) = \mu(E) - \frac{1}{n}\nu(E) \leq 0$, so $\mu(E) \leq \frac{1}{n}\nu(E) \leq \frac{1}{n}\nu(X) \to 0$. $\blacksquare$

Claim 5.2. $\nu_a \ll \mu$.

Proof. Suppose $\mu(E) = 0$. Then $0 = \mu(E \cap P) = \mu(\bigcup_n E \cap P_n)$. So $\mu(E \cap P_n) = 0$ for all n. $E \cap P_n$ is positive for all n, therefore

$$\nu(E \cap P_n) \leq n\mu(E \cap P_n) = 0$$

for all n. Hence,

$$\nu_a(E) = \nu(E \cap P) = \lim_{n \to \infty} \nu(E \cap P_n) = 0. \qquad \blacksquare$$

General case. Outlined in the book/exercise. □

5.2. The Radon-Nikodym theorem and related results

From the homework, we defined a measure coming from any measure space $(X, \mathcal{M}, \mu)$ as follows: given a measurable function $f \colon X \to [0, \infty]$ and $A \in \mathcal{M}$,

$$\mu_f(A) := \int f\chi_A \, d\mu.$$

Theorem 5.7 (Radon-Nikodym theorem)
Let ν be a σ-finite signed measure and μ be a σ-finite measure on $(X, \mathcal{M})$ with $\nu \ll \mu$. Then there exists, *modulo a.e. equivalence* an f, which is *extended μ-integrable*[1] such that $\nu = \mu_f$.

[1] That is, $f \in L^1 \pm L^+$, so either the positive or negative apart is integrable.

October 16, 2024

Proof. Main case: μ and ν are both positive, finite measures. Let

$$\mathcal{F} := \{f \in L^+ : \mu_f \leq \nu\}$$

(equivalently, $\nu - \mu_f$ is a measure). Let $\{f_n\} \subseteq \mathcal{F}$ be a sequence of functions that approach the supremum of the integrals of $f \in \mathcal{F}$, i.e.

$$\mu_{f_n}(X) = \int f_n \, d\mu \xrightarrow{n \to \infty} \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\}.$$

Define a new sequence $\{g_n\}$ by $g_1 = f_1$ and $g_n = \max\{f_n, g_{n-1}\}$. Then g_n are increasing, measurable, and in L^+.

Claim 5.3. $\{g_n\} \subseteq \mathcal{F}$.

Proof. We do this by induction. It clearly holds for $g_1 = f_1$. Then

$$\begin{aligned}
\mu_{g_{n+1}} &= \mu_{f_{n+1}\chi_{\{g_{n+1} = f_{n+1}\}}} + \mu_{g_n \chi_{\{g_{n+1} = g_n > f_{n+1}\}}} \\
&\leq \nu_{\chi\{g_{n+1} = f_{n+1}\}} + \nu_{\chi\{g_{n+1} = g_n > f_{n+1}\}} \\
&\leq \nu. \qquad \blacksquare
\end{aligned}$$

Since the g_n are increasing, they have a limit $g \in L^+$. $g \in \mathcal{F}$ by the monotone convergence theorem (2.8), and

$$\int g\,d\mu = \sup\left\{\int f\,d\mu : f \in \mathcal{F}\right\}.$$

The finiteness of μ is used here!

Set $\rho := \nu - \mu_g \geq 0$. Decompose $\mu = \mu_a + \mu_s$ with respect to ρ.

Case 1: $\mu_a = 0$. Then $\rho \ll \mu$, and $\mu \perp \rho$, so $\rho = 0$, i.e. $\nu = \mu_g$, as desired.

Case 2: $\mu_a \neq 0$. By the proof of the Lebesgue decomposition theorem (5.6), there exists an n and P_n, where P_n is a positive set for the signed measure $\rho - \frac{1}{n}\mu$ with $\mu(P_n) > 0$. In particular,

$$\begin{aligned}
\frac{1}{n}\mu_{\chi_{P_n}} \leq \rho &\implies \frac{1}{n}\mu_{\chi_{P_n}} \leq \nu - \mu_g \\
&\implies \frac{1}{n}\mu_{\chi_{P_n}} + \mu_g = \mu_{(g+\frac{1}{n}\chi_{P_n})} \leq \nu \\
&\implies g + \frac{1}{n}\chi_{P_n} \in \mathcal{F}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\int g + \frac{1}{n}\chi_{P_n}\,d\mu &= \sup\left\{\int f\,d\mu : f \in \mathcal{F}\right\} + \frac{1}{n}\mu(P_n) \\
&> \sup\left\{\int f\,d\mu : f \in \mathcal{F}\right\},
\end{aligned}$$

The finiteness of ν is used here!

a contradiction.

General case and uniqueness: Exercise. □

Notice that we may view μ as the antiderivative of f in the Radon-Nikodym theorem. Suppose ν and μ satisfy the conditions of the Radon-Nikodym theorem. We define the **Radon-Nikodym derivative** of ν with respect to μ as

$$\frac{d\nu}{d\mu} := f.$$

Informally, we have

$$\int gf\,d\mu = \int g\frac{d\nu}{\cancel{d\mu}}\cancel{d\mu} = \int g\,d\nu.$$

We also have an analogue of the chain rule:

$$\frac{d\rho}{d\mu} = \frac{d\rho}{d\nu}\frac{d\nu}{d\mu}.$$

In particular, if $\mu \ll \nu \ll \mu$, then

$$\frac{d\mu}{d\nu}\frac{d\nu}{d\mu} = 1 \qquad \text{a.e.}$$

5.2.1. Finding a representative in $\mathbb{R}^n$

The issue with the result of the Radon-Nikodym theorem (5.7) is that our derivative is defined up to a.e. equivalence. This means it is hard to find a single representative for f. We'll show that it is possible to find a representative in a certain case in the Lebesgue measure.

Definition 5.6

Let μ be a measure on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ and $(\mathbb{R}^n, \mathcal{L}^n)$. μ is **locally finite** if $\mu(K) < \infty$ for every compact K.

By Radon-Nikodym, if μ is locally finite and $\mu \ll \lambda$, then

$$\frac{d\mu}{d\lambda} \in L^1_{\mathrm{loc}}(\mathbb{R}^n),$$

i.e. $\int_K \left|\frac{d\mu}{d\lambda}\right| d\lambda < \infty$.

The goal of this section is to find a fixed a.e. representative g for an equivalence class $[g] \in L^1_{\mathrm{loc}(\mathbb{R}^d)}$.

Define the **average operator** as

$$A_r g(x) \coloneqq \frac{1}{|B_r(x)|} \int_{B_r(x)} g(y)\, dy \eqqcolon ⨍_{B_r(x)} g(y)\, dy.$$

Lemma 5.8

If $f \in C(\mathbb{R}^d)$ (continuous functions on $\mathbb{R}^d$), then $A_r f \to f$ locally uniformly as $r \to 0$.

Proof. Let $R > 0$. With respect to uniform convergence on $B_R(0)$, it suffices to show that

$$\lim_{\substack{r\to 0^+ \\ r \le 1}} A_r f \to f \qquad \text{uniformly on } B_R(0).$$

But notice that the limit only sees f on $B_{R+1}(0)$, and f is uniformly continuous there. □

We want to strengthen this to say that $A_r f \to f$ a.e. To do this we will first prove a.e. convergence for functions in some dense class in $L^1(dx)$, and then prove continuity of an associated maximal function:

$$Mf(x) \coloneqq \sup_{r>0} (A_r |f|)(x).$$

We call M the **Hardy-Littlewood maximal operator**.

Proposition 5.9 (Chebyshev's inequality)

October 18, 2024

For $f \in L^1(\mu)$,

$$\alpha\mu(\{|f| \ge \alpha\}) \le \|f\|_{L^1}.$$

Theorem 5.10

There exists a $C_n > 0$ such that for all $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$,

$$\alpha\, |\{Mf \ge \alpha\}| \le C_n \|f\|_{L^1}.$$

We say that $f \in L^{1,\infty}(\mu)$, or in **weak** L^1 if $\sup_\alpha \alpha\mu(\{|f| \ge \alpha\}) < \infty$. Proposition 5.9 implies that $L^1 \subseteq L^{1,\infty}$, but the converse is generally not true, since $f(x) = \frac{1}{|x|^n}$ is not in L^1, but is in $L^{1,\infty}$.

Another way of stating Theorem 5.10 is that M maps L^1 into $L^{1,\infty}$, or M is of **weak type** $(1,1)$.

Corollary 5.11 (Corollary of Theorem 5.10)

For $f \in L^1_{\mathrm{loc}}$, $A_r f \to f$ a.e.

October 23, 2024

Lemma 5.12 (Vitali covering lemma)
If $\mathcal{Q} \subseteq \mathcal{P}(\mathbb{R}^n)$ is a collection of balls such that $U := \bigcup_{B \in \mathcal{Q}} B$ and $A < |U|$ for some positive $A \in \mathbb{R}$, then there exists a finite subcollection $\mathcal{Q}' \subseteq \mathcal{Q}$ of disjoint balls such that $U' := \bigcup_{B \in \mathcal{Q}'} B$ satisfies $C_n A < |U'|$ for some positive C_n.

In fact, we can get a constant C that does not depend on n, but we omit a proof.

Non-Example 5.5 (We cannot replace "balls" with "sets") – Let U be the area under the graph of $y = \frac{1}{x}$ for $x \in (0, \infty)$, and let $\mathcal{Q}$ be the collection of all rectangles of the form $(0, x) \times (0, \frac{1}{x})$. Then $\bigcup_{x \in (0,\infty)} [0, x] \times [0, \frac{1}{x}] = U$, but any two rectangles intersect, and the measure of U is ∞.

Non-Example 5.6 – Cover $B_1(0)$ with one rectangle in each direction. Then they all overlap in the center. Moreover, they union to the ball, but any disjoint set can have arbitrarily small area. This is related to the *Kakeya set*.

Proof of Lemma 5.12. There exists $K \subseteq U$ such that $|K| > A$. K is covered by $\mathcal{Q}'' \subseteq \mathcal{Q}$, a finite subcollection. Choose $B_1 \in \mathcal{Q}''$ of maximal radius. To select future balls, we now perform the following process to construct $\mathcal{Q}'$.

If $K \subseteq \bigcup_{j=1}^m 3B_j$, then

$3B_j$ means the ball formed by dilating B_j by a scale factor of 3.

$$A < |K| \le \sum_{j=1}^m |3B_j| = 3^n \sum_{j=1}^m |B_j|,$$

which mean we may use $C = 3^{-n}$ to finish.

While there exists $x \in K$ such that $x \notin 3B_j$ for all j, Choose B_{m+1} of maximal radius such that $B_{m+1} \cap \bigcup_{j=1}^m B_j = \varnothing$. □

The reason for choosing a constant 3 follows from something like Figure 1. While 3 is not sharp, it is "good enough" for us.

Let's get to proving the weak type $(1, 1)$ bound.

Proof of Theorem 5.10. For $\alpha > 0$, let

$$E_\alpha := \{Mf \ge \alpha\} \subseteq \left\{Mf > \frac{\alpha}{2}\right\}.$$

If $x \in E_\alpha$, then there exists $r > 0$ such that

$$\fint_{B_r(x)} |f| > \frac{\alpha}{2} \implies \int_{B_r(x)} |f| > \frac{\alpha}{2} |B_r(x)|.$$

Using $\{B_r(x)\}_{x \in E_\alpha}$ as an open cover for E_α, the Vitali covering lemma (5.12) tells us that given a constant $A < |E_\alpha|$ there exist balls $B_{x_1}, \ldots, B_{x_N}$ that are disjoint with

$$\left| \bigcup_{j=1}^N B_{x_j} \right| > 3^{-n} A.$$

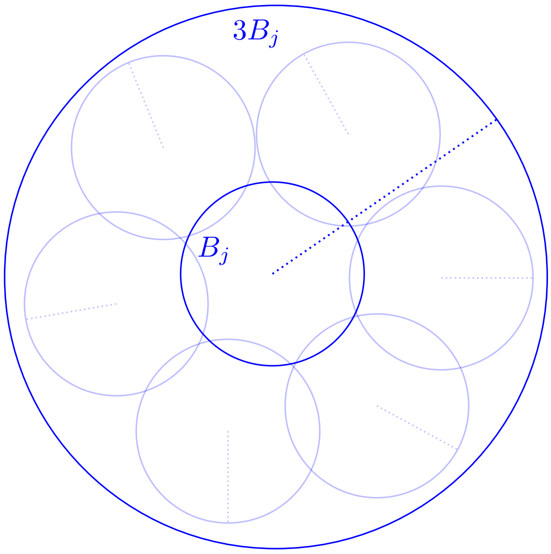


Figure 1: Geometric reason for picking $3B_j$.

This means

$$\begin{aligned}
3^{-n}A &< \left| \bigcup_{j=1}^{N} B_{x_j} \right| \\
&= \sum_{j=1}^{N} |B_{x_j}| \\
&< \frac{2}{\alpha} \left(\sum_{j=1}^{N} \int_{B_{x_j}} |f(x)|\, dx \right) \\
&= \frac{2}{\alpha} \int_{\bigcup_{j=1}^{N} B_{x_j}} |f(x)|\, dx \\
&\leq \frac{2}{\alpha} \|f\|_{L^1} .
\end{aligned}$$

□

Theorem 5.13 (Lebesgue differentiation theorem (variations))
For $f \in L^1_{\text{loc}}$,

(a)
$$\lim_{r \to 0^+} \fint_{B_r(x)} |f(y) - f(x)|\, dy = 0 \qquad \text{a.e.}$$

(b)
$$\lim_{r \to 0^+} \fint_{B_r(x)} |f(y) - A_r f(x)|\, dy = 0 \qquad \text{a.e.}$$

(c) $A_r f(x) \to f(x)$ a.e.

We proved (c) already. Moreover, (a) is the strongest and implies (b) and (c). Using the Lebesgue differentiation theorem, we conclude that $\lim_{r \to 0^+} A_r f(x)$ is an a.e. defined member of [f].

Definition 5.7

ν is a **regular Borel measure on $\mathbb{R}^n$** if

1. ν is a Borel measure,
2. $\nu(K) < \infty$ for all compact $K \subseteq \mathbb{R}^n$,
3. $\nu(E) = \inf\{\nu(U) : U \subseteq E, U \text{ open}\}$ for all E.

Theorem 5.14 (Radon-Nikodym on $\mathbb{R}^n$)

If ν and its absolutely continuous part (from the Lebesgue decomposition theorem (5.6): $\nu = \nu_a + \nu_s$) are regular Borel measures on $\mathbb{R}^n$, write $\nu_a = f dx$. Then

$$f(x) = \lim_{r \to 0^+} \frac{\nu(B_r(x))}{|B_r(x)|} \qquad \text{Lebesgue-a.e.}$$

October 25, 2024

Proof. This works for ν_a by the Lebesgue differentiation theorem (5.13), so we may assume that $\nu = \nu_s \perp dx$, and it suffices to show that the RHS of the formula is zero λ-a.e.

There exists $E \in \mathcal{B}_{\mathbb{R}^n}$ such that $\nu(E) = 0$, $|E^C| = 0$. Define

These sets are named confusingly, because $\varepsilon < \varepsilon'$ implies $E_\varepsilon \supseteq E_{\varepsilon'}$.

$$E_\varepsilon \coloneqq \left\{ x \in E : \limsup_{r \to 0^+} \frac{\nu(B_r(x))}{|B_r(x)|} \geq \varepsilon \right\}.$$

We want to show that $|E_\varepsilon| < \varepsilon$ for all ε (then $\left|\bigcup_{\varepsilon = \frac{1}{n}} E_\varepsilon\right| = \lim_{\varepsilon \to 0} |E_\varepsilon| = 0$).

Therefore, let $\delta > 0$ (which we will choose at the end). By regularity and the fact that $\nu(E) = 0$, there exists an open set $U_\delta \supseteq E$ such that

$$\nu(U_\delta) < \delta.$$

$x \in E_\varepsilon \implies x \in U_\delta \implies$ there exists $r_x > 0$ such that $B_x \coloneqq B_{r_x}(x) \subseteq U_\delta$ and $\frac{\nu(B_x)}{|B_x|} > \varepsilon$. So

$$U'_\delta \coloneqq \bigcup_{x \in E_\varepsilon} B_x \subseteq U_\delta$$

is a union of balls. By the Vitali covering lemma (5.12), for $A < |E_\varepsilon|$, there exist $x_1, \ldots, x_N$

such that B_{x_j} disjoint and $\left|\bigcup_{j=1}^N B_{x_j}\right| > CA$ for some constant $C > 0$. So

$$\begin{aligned}
\delta > \nu(U_\delta) \geq \nu(U'_\delta) \geq \nu\left(\bigcup_{j=1}^N B_{x_j}\right) \\
&= \sum_{j=1}^N \nu(B_{x_j}) \\
&> \varepsilon \sum_{j=1}^N |B_{x_j}| \\
&\geq \varepsilon \left|\bigcup_{j=1}^N B_{x_j}\right| \\
&\geq C\varepsilon A.
\end{aligned}$$

So setting $\delta < \frac{\varepsilon^2}{C}$ finishes. $\square$

6. Banach spaces

November 6, 2024

Definition 6.1
Let X be a normed vector space over $\mathbb{R}$ or $\mathbb{C}$. $\|\cdot\| : X \to [0, \infty)$ is called a **seminorm** on X if

1. $\|\lambda x\| = |\lambda| \cdot \|x\|$,
2. $\|x + y\| \leq \|x\| + \|y\|$.

If we add the condition

3. $\|x\| \geq 0$ for all x and $\|x\| = 0 \iff x = 0$,

$\|\cdot\|$ is called a **norm**. The normed metric $d(x, y) := \|x - y\|$ induces a topology. X is a **Banach space** if X is complete with respect to the metric above.

Recall that $\sum_{n=1}^{\infty} x_n$ is said to **converge** to x if the partial sums $\sum_{n=1}^{N} x_n$ converge to x (in norm) as $N \to \infty$. The series is said to **absolutely converge** if $\sum_{n=1}^{\infty} \|x_n\|$ is finite.

Theorem 6.1
$(X, \|\cdot\|)$ is a Banach space $\iff$ absolute convergence implies convergence for a series in X.

Proof (sketch). ($\implies$) Assume X is complete and $\{x_n\}$ is an absolutely convergent series. By the triangle inequality, the partial sums $s_N := \sum_{n=1}^{N} x_n$ are Cauchy, hence convergent.

($\impliedby$) Let $\{x_n\}$ be Cauchy. We may iteratively construct a subsequence $\{x_{n_k}\}$ such that

$$\|x_{n_k} - x_{n_{k+1}}\| < 2^{-k},$$

so

$$x = x_{n_1} + \underbrace{\sum_{k=1}^{\infty} x_{n_{k+1}} - x_{n_k}}_{\text{abs. conv, hence conv.}}.$$

Since x_n is Cauchy, $x = \lim_{k \to \infty} x_{n_k}$. □

An immediate consequence of the above fact with the dominated convergence theorem is:

Corollary 6.2
L^1 is a Banach space.

6.1. Linear operators and continuity

Theorem 6.3
Let $T\colon X \to Y$ be a linear map between normed vector spaces X and Y. The following are equivalent:

1. T is continuous on X.
2. T is continuous at 0 (or any other point).
3. There exists a constant C such that for all $x \in X$,

$$\|Tx\|_Y \leq C\,\|x\|_X\,.$$

Proof. ((1) $\implies$ (2)) is immediate.
((2) $\implies$ (3)) Let $\varepsilon = 1$. So there exists $\delta > 0$ such that

$$T(B_\delta^X(0)) \subseteq B_1^Y(0).$$

Let $x \in B_\delta^X(0)$. If $x = 0$, it is clear. Otherwise,

$$\frac{\delta}{2}\frac{x}{\|x\|} \in B_\delta^X(0) \implies T\left(\frac{\delta}{2}\frac{x}{\|x\|}\right) = \frac{\delta}{2\,\|x\|}Tx \in B_1^Y(0) \implies \|Tx\| \leq \frac{2}{\delta}\,\|x\|\,.$$

So set $C = \frac{2}{\delta}$.
((3) $\implies$ (1)) If $C = 0$, then $T \equiv 0$ and we are done. Otherwise, $C \neq 0$. Then if $\|x_1 - x_2\| < \frac{\varepsilon}{C}$, then

$$\|T(x_1) - T(x_2)\| = \|T(x_1 - x_2)\| \leq C\,\|x_1 - x_2\| < \varepsilon.$$

□

Definition 6.2
$T\colon X \to Y$ is a **bounded linear operator** if any of the statements in Theorem 6.3 hold. The **operator norm** of T is defined as

$$\begin{aligned}\|T\|_{X\to Y} &:= \inf\left\{C : \|Tx\|_Y \leq C\,\|x\|_X\right\} = \min\left\{C : \|Tx\|_Y \leq C\,\|x\|_X\right\}\\ &= \sup\{\|Tx\| : \|x\| = 1\}\\ &= \sup\{\|Tx\| : \|x\| \leq 1\}\\ &= \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\}.\end{aligned}$$

Proposition 6.4
$\mathcal{L}(X,Y)$ is a normed vector space with respect to the operator norm $\|\cdot\|_{X\to Y}$, and is complete if Y is.

Proof. The first two norm properties are clear. For triangle inequality, if $\|x\| = 1$, then

$$\|(T+S)x\| = \|Tx + Sx\| \leq \|Tx\| + \|Sx\| \leq \|T\| + \|S\|\,.$$

Now take the supremum over all x with $\|x\| = 1$.
For completeness, if $\sum_n \|T_n\| < \infty$, then for all $x \in X$, $\sum_n T_n x$ is absolutely conver-

gent, so it converges. So

$$\sum_n \|T_n x\| \leq \sum_n \|T_n\| \cdot \|x\|,$$

so

$$x \mapsto \sum_n T_n x$$

is linear and bounded by $C = \sum_n \|T_n\|$. □

6.1.1. Compositions and inverses

If $S \colon X \to Y$ and $T \colon Y \to Z$ are bounded linear operators, then $T \circ S$ is a bounded linear operator with $\|T \circ S\| \leq \|T\| \cdot \|S\|$. In particular, $\mathcal{L}(X, X)$ is an algebra with composition as multiplication. We call it a **Banach algebra** if X is a Banach space (hence $\mathcal{L}(X, X)$ is too).

We say $T \colon X \to Y$ is **invertible** if T is a bijection and T^{-1} is a bounded linear operator.

Example 6.1 (HW8 Problem 3) – Let X be a Banach space and $V \subseteq X$ a closed vector subspace with $V \neq X$. Let $Y := X/V$

$$\begin{aligned}\pi \colon X &\to Y\\ x &\mapsto x + V.\end{aligned}$$

We have a norm on Y given by

$$\|x+V\| := \inf\{\|x+y\| : y \in V\} = \inf\{\|x-y\| : y \in V\}.$$

Clearly the norm is non-negative. If $x \in V$, then $\|x+V\| \leq \|x+(-x)\| = \|0\| = 0$. So $\|x+V\| = 0$. Suppose $x \notin V$. Then, since V is closed, there exists a ball $B_\varepsilon(x)$ (w.r.t. the norm on X) disjoint from V. It follows that $\|x-v\| \geq \varepsilon$ for all $v \in V$, hence $\inf_{v\in V}\|x+v\| = \inf_{v \in V}\|x-v\| \geq \varepsilon > 0$.

If $\lambda = 0$, then $\|\lambda(x+V)\| = \|0+V\| = 0 = |0| \cdot \|x+V\|$. Now suppose $\lambda \neq 0$. Since scalar multiplication by a nonzero value is an automorphism of V,

$$\begin{aligned}\|\lambda(x+V)\| &= \|\lambda x + V\|\\ &= \inf_{v\in V}\|\lambda x + v\|\\ &= \inf_{v\in V}|\lambda|\left\|x+\frac{v}{\lambda}\right\|\\ &= |\lambda|\inf_{v'\in V}\|x+v'\|\\ &= |\lambda|\,\|x+V\|.\end{aligned}$$

For the triangle inequality, let $x+V, y+V \in Y$ and $v_0 \in V$. Then

$$\|x+y+V\| \leq \|x+y+v_0\| \leq \left\|x+\frac{v_0}{2}\right\| + \left\|y+\frac{v_0}{2}\right\|,$$

where the last line is from the triangle inequality on X. Taking the infimum over all $v_0 \in V$ (equivalent to taking the infimum over all $\frac{v_0}{2} \in V$), we have

$$\|x+y+V\| \leq \|x+V\| + \|y+V\|.$$

Further, we claim that the projection/quotient map π has operator norm 1. Let $\|x\| = 1$. Then

$$\|\pi(x)\| = \|x+V\| = \inf_{y\in V}\|x+y\| \leq \|x+0\| = 1.$$

Taking the supremum over all $x \in X$ where $\|x\| = 1$, we have that $\|\pi\| \leq 1$.

Let $x \notin V$. Pick $y_n \in V$ so that $\|x+y_n\| < \inf_{y\in V}\|x+y\| + \frac{1}{n}$. Since $x + y_n \in x+V$, we may choose it as a representative, and

$$\|\pi\| = \sup_{x\neq 0}\frac{\|x+V\|}{\|x\|} \geq \frac{\|x+V\|}{\|x+y_n\|} > \frac{\|x+V\|}{\|x+V\|+\frac{1}{n}}.$$

Taking the limit as $n \to \infty$, we have that $\|\pi\| \geq 1$. So $\|\pi\| = 1$.

6.2. Linear functionals and the Hahn-Banach theorem

November 8, 2024 **Functionals** are elements of $\mathcal{L}(X, k)$, where $k = \mathbb{R}$ or $\mathbb{C}$ and X is an k-vector space.

Example 6.2 – Let $X = L^1(\mu)$ and g a bounded, measurable function. We define a linear functional

$$\ell_g(f) := \int gf \, d\mu.$$

It is also a bounded linear functional:

$$\left| \int fg \, d\mu \right| \leq \int |fg| \, d\mu \leq M \int |f| \, d\mu \leq M \|f\|_{L^1}.$$

Proposition 6.5

If X is a vector space over $\mathbb{C}$, it is also a vector space over $\mathbb{R}$. A $\mathbb{C}$-linear functional ℓ can be turned into an $\mathbb{R}$-linear functional $\mathrm{Re}(\ell)$. In the other direction, an $\mathbb{R}$-linear functional u defines a $\mathbb{C}$-linear functional $f(x) = u(x) - iu(ix)$.

Proof (sketch). Use unimodular multiplication in X to show $\mathrm{Re}(\ell)$ is an $\mathbb{R}$-linear functional, and f being a $\mathbb{C}$-linear functional follows. □

Definition 6.3

Let X be a $\mathbb{R}$-vector space. A **sublinear functional** on X is a map $\rho \colon X \to \mathbb{R}$ such that

1. $\rho(x+y) \leq \rho(x) + \rho(y)$ for all $x, y \in X$,
2. $\rho(\lambda x) = \lambda \rho(x)$ for all $x \in X$ and $\lambda \geq 0$.

Example 6.3 – ρ linear $\implies$ ρ sublinear. Any seminorm (hence, norm) is sublinear.
The functional

$$\rho(f) := \sup_{\substack{B \subseteq \mathbb{R}^n \text{ a ball} \\ 1 \leq \mathrm{diam}(B)}} \fint_B |f(y)| \, d\mu(y)$$

is sublinear.
Let $X = C^\infty(\mathbb{R}^n)$. The functional

$$\rho_{N,M}(f) := \sup_{|x| \leq N} \max_{\alpha \leq M} |\partial^\alpha f(x)|$$

is sublinear.

Theorem 6.6 (Hahn-Banach theorem)

Let X be an $\mathbb{R}$-vector space and $V \subseteq X$ a subspace. Let ρ a sublinear functional on X obeying $f \leq \rho$. Then there exists a linear extension $F \colon X \to \mathbb{R}$ of f [1] also obeying $F \leq \rho$.

[1] i.e., $F|_V = f$

Proof. Case I: $V + \mathbb{R}x_0$ for some $x_0 \notin V$. Observe that for every $y_1, y_2 \in V$,

$$f(y_1) + f(y_2) = f(y_1 + y_2) \leq \rho(y_1 + y_2) \leq \rho(y_1 - x_0) + \rho(x_0 + y_2).$$

Rearranging,

$$f(y_1) - \rho(y_1 - x_0) \leq \rho(x_0 + y_2) - f(y_2).$$

Choose some

$$\alpha_{x_0} \in \left[\sup_{y_1 \in V} (f(y_1) - \rho(y_1 - x_0)), \inf_{y_2 \in V} (\rho(x_0 + y_2) - f(y_2))\right],$$

and set

$$F(y + \lambda x_0) \coloneqq f(y) + \lambda \alpha_{x_0}.$$

This is clearly linear, obeys $F \leq \rho$, and extends f. If $\lambda \geq 0$, then

$$\begin{aligned} F(y + \lambda x_0) &= f(y) + \lambda \alpha_{x_0} \\ &\leq f(y) + \lambda \left(\rho\left(x_0 + \frac{y}{\lambda}\right) - f\left(\frac{y}{\lambda}\right)\right) \\ &= \rho(\lambda x_0 + y). \end{aligned}$$

General case: We will use Zorn's lemma. The set

$$\mathcal{W} \coloneqq \{(W, F) : V \subseteq W \subseteq X, F \text{ extends } f\}$$

has a partial order given by $(W_1, F_1) \preceq (W_2, F_2)$ if $W_1 \subseteq W_2$ and f_2 extends F_1. Consider a toset $\{(W_\alpha, F_\alpha)\}_{\alpha \in \mathcal{A}}$. We create a maximal element by defining

$$W_0 \coloneqq \bigcup_{\alpha \in \mathcal{A}} W_\alpha,$$

and $F_0 \colon W_0 \to \mathbb{R}$ defined by the common functional to all the W_α (this is possible by the definition of the partial order). By Zorn's lemma, there is a maximal element $(\overline{W}, \overline{F}) \in \mathcal{W}$. If $\overline{W} \neq X$, then by case I, we can extend $\overline{F}$ to a linear functional on $\overline{W} + \mathbb{R}x_0$, which contradicts the fact that $(\overline{W}, \overline{F})$ is an upper bound. ☐

Corollary 6.7

Let X be a $\mathbb{R}$-vector space and $V \subseteq X$. If there exists a bounded linear functional $\ell \colon V \to \mathbb{R}$, then there exists a linear extension $F \colon X \to \mathbb{R}$ with

$$\|F\|_{X \to \mathbb{R}} = \|f\|_{V \to \mathbb{R}}.$$

November 11, 2024

Theorem 6.8 (Complex Hahn-Banach theorem)

Let X a $\mathbb{C}$-vector space and $V \subseteq X$ a subspace. Let ρ be a seminorm on X, and $f \colon V \to \mathbb{C}$ a functional with $|f| \leq \rho$. Then f extends to a linear functional $F \colon X \to \mathbb{C}$ linear with $|F| \leq \rho$.

Proof (ideas). Apply the real Hahn-Banach theorem (6.6) to $\mathrm{Re}(f) \eqqcolon u$ to extend it to a map $U \colon X \to \mathbb{R}$ with $|U| \leq \rho$. Then complexify by defining

$$F(x) \coloneqq U(x) - iU(ix).$$ ☐

Corollary 6.9

Let X be a normed vector space.

(a) If $V \subseteq X$ is a closed vector subspace and $x_0 \in X \setminus V$, there exists $f \in X^*$ such that $f(x_0) \neq 0$ and $f|_V \equiv 0$. Further, with $\delta := \inf_{y \in V} \|x_0 - y\|$, we can choose f such that $f(x_0) = \delta$ and $\|f\| = 1$.

(b) For all $x \in X$, $\|x\| = \min_{\substack{\|f\|=1 \\ f \in X^*}} |f(x)|$. In particular, there exists $f \in X^*$ such that $\|f\| = 1$ and $|f(x)| = \|x\|$.

(c) Bounded linear functional separate points, i.e., if $x \neq y$, then there exists an $f \in X^*$ such that $f(x) \neq f(y)$.

(d) X isometrically embeds into X^{**} by $x \mapsto [\widehat{x}\colon X^* \to \mathbb{C}\colon \widehat{x}(f) \mapsto f(x)]$.

Proof. (a) Set $W := V + \mathbb{C}x_0$. Define $f(\lambda x_0 + y) = \lambda\delta$. This is linear, and

$$\begin{aligned} |f(\lambda x_0 + y)| &= |\lambda| \cdot \left| f(x_0) + \frac{y}{\lambda} \right| \\ &= |\lambda| \cdot \inf_{\tilde{y} \in V} \|x_0 - \tilde{y}\| \\ &\leq |\lambda| \cdot \left\| x_0 - \frac{y}{\lambda} \right\| \\ &= \|\lambda x_0 - y\| . \end{aligned}$$

So $\|f\| \leq 1$. For $\|f\| \geq 1$, choose a sequence $\{y_n\} \subseteq V$ such that $\|x_0 - y_n\| \to \delta$. Then

$$\frac{|f(x_0 - y_n)|}{\|x_0 - y_n\|} \xrightarrow{n \to \infty} 1.$$

Then extend f to a norm 1 operator on X by complex Hahn-Banach (6.8).

(b) The "in particular" part follows from (a) with $V + \{0\}$. "$\|x\| \leq \min_{\substack{\|f\|=1 \\ f \in X^*}} |f(x)|$" follows from the definition of norm. "$\|x\| \geq \min_{\substack{\|f\|=1 \\ f \in X^*}} |f(x)|$" follows from the "in particular" part.

(c) By (b), if $x \neq y$, then there exists $f \in X^*$ such that $f(x - y) \neq 0$.

(d) Linearity is clear, for the isometric embedding, use (b). $\square$

6.2.1. Completion of X

Part (d) of the last corollary (6.9) motivates the idea of completing X. Since $\mathbb{C}$ is complete, so is $X^* = \mathcal{L}(X, \mathbb{C})$, and so is $X^{**} = \mathcal{L}(\mathcal{L}(X, \mathbb{C}), \mathbb{C})$. Therefore, X embeds into a complete space. Define the **completion** of X as the closure of this embedding: if $X \hookrightarrow \widehat{X} \subseteq X^{**}$, then the completion is $\overline{\widehat{X}}$.

Definition 6.4

X is **reflexive** if $\widehat{X} = X^{**}$, i.e. if $H\colon X \to X^{**}\colon x \mapsto \widehat{x}$ is an isometry.

6.2.2. The adjoint map

Let X, Y be normed vector spaces and $T \in \mathcal{L}(X, Y)$. The **adjoint** of T is the pullback

$$T^* \colon Y^* \to X^* \colon T^* f = f \circ T.$$

Proposition 6.10 (Properties of T^*)

1. $T^* \in \mathcal{L}(Y^*, X^*)$ and $\|T^*\| = \|T\|$.
2. T^* is injective $\iff \overline{TX} = Y$.
3. If $\overline{T^* Y^*} = X^*$, then T is injective.
4. $(T^*)^*(\widehat{x}) = \widehat{Tx}$.

Proposition 6.11
If X is a Banach space, X is reflexive $\iff$ X^* is.

Proof. Consider the subspace $V_0 \subseteq X^{***}$ given by

$$V_0 := \{F \in (X^{**})^* : F(\widehat{x}) = 0, \forall x \in X\}.$$

This is closed vector space.

Lemma 6.12
$X^{***} = V_0 \oplus \widehat{X^*}$.

Proof. Suppose $f \in X^*$ and $\widehat{f} \in V_0$. Then $\widehat{f}$ vanishes on $\widehat{X}$. But $\widehat{f}(\widehat{x}) = \widehat{x}(f) = f(x) = 0$, so f is the zero functional.

Let $F \in X^{***}$. Let $H \colon X \to X^{**}$ be defined as in the definition of reflexive. Consider the adjoint operator $H^* \colon X^{***} \to X^*$. Then $H^*(F) \in X^* \implies \widehat{H^*(F)} \in X^{***}$. It suffices to show that $F - \widehat{H^*}$ is zero on $\widehat{X}$. Indeed,

$$\begin{aligned}
(F - \widehat{H^*(F)})(\widehat{x}) &= F(\widehat{x}) - \widehat{x}(H^*(F)) \\
&= F(\widehat{x}) - H^* F(x) \\
&= F(\widehat{x}) - F(H(x)) \\
&= H(\widehat{x}) - F(\widehat{x}) = 0.
\end{aligned}$$

■

With this lemma, X is reflexive $\iff$ $\widehat{X} = X^{**}$ $\iff$ (because X is a Banach space) $\overline{\widehat{X}} = X^{**} \iff V_0 = 0 \iff \widehat{X^*} = (X^*)^{**} \iff X^*$ is reflexive. □

6.3. Baire category theorem and related results

Making up for the analysis notes I did not take on the BCT...

Definition 6.5
A set is **nowhere dense** if the interior of its closure is empty. Equivalently, a set E is nowhere dense if there is no ball B where $E \cap B$ is dense in B.

Example 6.4 –

- The integers are nowhere dense.
- A proper vector subspace or a positive codimension manifold inside $\mathbb{R}^d$ is nowhere dense.
- The middle-thirds Cantor set is nowhere dense.

Theorem 6.13 (Baire category theorem)
Let (X, d) be a complete metric space. Then

(a) X is not a countable union of nowhere dense sets. In other words, if $\{F_n\}$ are closed subsets of X and $F_n^\circ = \varnothing$, then $X \neq \bigcup_n F_n$.

(b) If $\{U_n\}$ is a collection of open and dense sets, then $\bigcap_n U_n$ is dense.

Proof. ((b) $\implies$ (a)) Suppose $\{F_n\}$ is a collection of closed subsets of X. It's clear that F_n^C is open. To show it is dense, assume that $x \notin F_n^C$ had a neighborhood B that did not intersect F_n^C. But then $B \subseteq F_n$, so $F_n^\circ \neq \varnothing$, a contradiction. By the assumption in (b), $\bigcap_n F_n^C$ is dense. In particular, it is nonempty, so $X \neq \left(\bigcap_n F_n^C\right)^C = \bigcup_n F_n$.

(b) Let $x_0 \in X$. Since U_1 is dense, $U_1 \cap B_{\varepsilon_0/2}(x_0)$ is nonempty. So there exists $x_1 \in U_1 \cap B_{\varepsilon_0/2}(x_0)$, $\varepsilon_1 > 0$ such that

$$B_{\varepsilon_1}(x_1) \subseteq U_1 \cap B_{\varepsilon_0/2}(x_0).$$

Now argue similarly with U_2 in place of U_1 to get a sequence $\{x_n\} \subseteq X$ and $\{\varepsilon_n > 0\}$ such that

$$B_{\varepsilon_n}(x_n) \subseteq U_n \cap B_{\varepsilon_{n-1}/2}(x_{n-1}).$$

WLOG, $\varepsilon_n \searrow 0$. Since $x_n \in B_{\varepsilon_{N-1}/2}(x_{N-1})$ for all $n \geq N$, this sequence is Cauchy by the triangle inequality. Define $x_\infty := \lim_{n\to\infty} x_n$. Then $x_\infty \in B_{\varepsilon_n}(x_n)$ for all n. Since $B_{\varepsilon_n}(x_n) \subseteq U_{n-1}$ for all n,

$$x \in \left(\bigcap_n U_n\right) \cap B_{\varepsilon_0}(x_0),$$

so $\bigcup_n U_n$ is dense. □

Definition 6.6
We define **1st category**/**meager** sets as countable unions of nowhere dense sets, and **2nd category** sets as any other sets. In particular, in a complete metric space, sets that are complements of meager sets are called **residual**/**generic**/**comeager**.

Remark 6.14. Terence Tao gives a useful comparison between measure-theoretic terms and Baire-category-theoretic terms:

Baire categorical	Measure theory
Complete non-empty metric space X	Measure space X of positive measure
First category (meager)	Zero measure (null)
Second category	Positive measure
Generic	Full measure (co-null)
Baire	measurable

Example 6.5 (Meager $\neq$ nowhere dense) – $\mathbb{Q}$ is meager, since it is the countable union of its elements.

Theorem 6.15

Let $\{f_n\}$ be a sequence of real/complex-valued continuous functions on a complete metric space X. Assume $\lim_{n\to\infty} f_n(x) = f(x)$ for all x. Then the set of continuity points is generic.

Example 6.6 – $\chi_{\mathbb{Q}}$ is not a pointwise limit of continuous functions.

Theorem 6.16 (Open mapping theorem)

Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$ surjective. Then T is an open mapping.

As a sanity check, the open mapping theorem is intuitively clear with finite-dimensional vector spaces, since a non surjective map will "squish" a unit ball to a lower dimension.

Proof of Theorem 6.16. Write $X = \bigcup_{n=1}^\infty B_n(0) \implies Y = \bigcup_{n=1}^\infty TB_n(0)$. Since Y is complete, there exists an $n \in \mathbb{N}$ such that

$$(\overline{TB_n(0)})^\circ = n(\overline{TB_1(0)})^\circ \neq \varnothing$$

by the Baire category theorem (6.13). So there exists $r > 0$ and $y_0 \in Y$ such that $B_{4r}(y_0) \subseteq \overline{TB_1(0)}$. In particular, $y_0 \in \overline{TB_1(0)}$ implies there is an $x_1 \in B_1(0)$ such that $y_1 := Tx_1 \in B_{2r}(y_0)$. Therefore, if $\|y\| < 2r$,

$$y = Tx_1 + (y - y_1) \in \overline{T(x_1 + B_1(0))} \subseteq \overline{TB_2(0)}.$$

Hence, if $\|y\| < r$, then $y \in \overline{TB_1(0)}$.

Now, if $\|y\| < r2^{-n}$, then $y \in \overline{T(B_{2^{-n}}(0))}$. In particular, if $\|y\| < \frac{r}{2}$, then there is an $x_1 \in B_{1/2}$ such that $\|y - Tx_1\| < \frac{r}{4}$. Continuing inductively, there are $x_1, \ldots, x_n$ such that

$$\left\| y - T\sum_{j=1}^n x_j \right\| < r2^{-n-1}.$$

Since X is complete, we can say $\sum_{i=1}^\infty x_n$ converges to some x. But $\|x\| < \sum_{j=1}^\infty 2^{-j} = 1$, and $y = Tx$. In other words, TB_1 contains all y with $\|y\| < \frac{r}{2}$, finishing the proof. ☐

November 15, 2024

Corollary 6.17

A bijection between Banach spaces is an isomorphism (i.e. $T^{-1} \in \mathcal{L}(Y, X)$).

Proposition 6.18

A Banach space cannot have a countably infinite basis.[1]

[1] In the algebraic sense, i.e., every element can be expressed as a finite linear combination of basis elements.

Proof. Suppose X is an infinite-dimensional Banach space with a countable basis $\{e_n\}$. Define $F_k := \operatorname{span}\{e_1, \ldots, e_k\}$, which is closed because it is a finite-dimensional subspace. Moreover, $F_k \neq \varnothing$, and $x \in F_k$ implies $x + \varepsilon e_{k+1} \notin F_k$ for all $\varepsilon > 0$. But by the hypothesis, $X = \bigcup_k F_k$, so X is meager. □

Theorem 6.19 (Closed graph theorem)

Let X, Y be Banach spaces and $T\colon X \to Y$ a linear map. Then T is continuous $\iff$ Γ_T is a closed subspace of $X \times Y$.[1]

[1] The norm on $X \times Y$ is given by $\|(x, y)\| = \|x\| + \|y\|$.

Proof. ($\implies$) $X \times Y$ is Hausdorff, so Γ_T is closed by a fact in elementary topology.

($\impliedby$) Since $\Gamma_T \subseteq X \times Y$ is closed, it is a Banach space. Since the projection $\pi_X(x, Tx) = x$ is a bijection, it is an isomorphism. Therefore, $T := \pi_Y \circ (\pi_X|_{\Gamma_T})^{-1}$ is bounded. □

The principle of uniform boundedness tells us that we just need to bound the norm of a collection of operators at each point individually to bound all the operators in the operator norm.

Theorem 6.20 (Principle of uniform boundedness)

Let X, Y be normed vector spaces and $\mathcal{A} \subseteq \mathcal{L}(X, Y)$. Then

(a) If $E := \left\{x : \sup_{T \in \mathcal{A}} \|Tx\| < \infty\right\}$ is not meager, then $\sup_{T \in \mathcal{A}} \|Tx\| < \infty$.

(b) If X is a Banach space and $\sup_{T \in \mathcal{A}} \|Tx\| < \infty$ for all $x \in X$, then $\sup_{T \in \mathcal{A}} \|T\| < \infty$.

Proof. (a) By the Baire category theorem (6.13), X is not meager. Define

$$E_n := \left\{ x : \sup_{T \in \mathcal{A}} T(x) \leq n \right\}.$$

Then $E \subseteq \bigcup_n E_n$, and the E_n are closed, being an intersection of closed sets. Since E is not meager, there exists $n \in \mathbb{N}$ such that E_n is not nowhere dense, i.e., there exist $r > 0$, and $x_0 \in X$ such that

$$B_r(x_0) \subseteq E_n.$$

Let $\|x\| < 1$. Then

$$\begin{aligned} \|Tx\| &= r^{-1} \|T(rx + x_0) - Tx_0\| \\ &\leq r^{-1}(\|T(rx + x_0)\| + \|Tx_0\|) \\ &\leq 2nr^{-1}. \end{aligned}$$

Then $\|T\| \leq 2nr^{-1}$ for all $T \in \mathcal{A}$.

(b) is a consequence of (a). □

Proposition 6.21

Suppose $\{T_n\} \subseteq \mathcal{L}(X, Y)$ converges pointwise in X. Then the pointwise limit of T_n, call it T, is in $\mathcal{L}(X, Y)$.

Remark 6.22. Note that $T_n \not\to T$ necessarily in the topology induced by $\|\cdot\|_{X\to Y}$.

Proposition 6.23

Let X, Y, Z be Banach spaces. Given a bilinear map $B\colon X\times Y \to Z$, B is continuous $\iff$ B is continuous in each variable.

Proof. We will show B continuous $\iff$ B is continuous in each variable.

($\implies$) This follows from a fact from elementary topology, where a continuous map $f\colon X\times Y\to Z$ being continuous implies that $f(x,-)$ and $f(-,y)$ are continuous for fixed $x\in X, y\in Y$.

($\impliedby$) Given some $y\in Y$, recall that $B(-,y)\colon X\to Z$ is a bounded linear operator by assumption. Let it be bounded by C_y. Then

$$\|B(x,y)\| \le C_y\,\|x\| \le C_y.$$

So

$$\sup_{\substack{B(x,-)\\ \|x\|\le 1}} \|B(x,y)\| \le C_y.$$

By the principle of uniform boundedness,

$$\sup_{\substack{B(x,-)\\ \|x\|\le 1}} \|B(x,-)\| \le C$$

for some finite C.

Let U_X, U_Y be the unit balls in X and Y respectively. It suffices to show that $B(U_X, U_Y)$ is bounded to show continuity [Exercise: show this]. Now, let $(x,y)\in U_x\times U_y$. Then

$$\|B(x,y)\| \le \sup_{\substack{B(x,-)\\ \|x\|\le 1}} \|B(x,-)\|\,\|y\| \le C\,\|y\| \le C. \qquad \square$$

6.4. Topological vector spaces

Definition 6.7

Let $(X,\mathcal{T})$ be a topological space with X also a vector space. We say $(X,\mathcal{T})$ is a **topological vector space (t.v.s.)** if "$\mathcal{T}$ is compatible with the vector space structure of X." More precisely, the addition and scaling maps

$$\begin{aligned} +\colon X\times X &\to X \\ (x,y) &\mapsto x+y \\ \cdot\colon k\times X &\to X \\ (\lambda,x) &\mapsto \lambda x \end{aligned}$$

are both continuous. In particular, for $x\in X$ and $\lambda\ne 0$, $[x\mapsto \lambda x]$ and $[y\mapsto y+x]$ is a homeomorphism, so it suffices to find a neighborhood base $\mathcal{N}_0$ around 0 so that addition and scalar multiplication are continuous.

Definition 6.8

A t.v.s. $(X, \mathcal{T})$ is **locally convex** if there exists a base for $\mathcal{T}$ whose elements are all convex (recall a subset E of a vector space is **convex** if for all $x, y \in E$ and $\theta \in [0,1]$, $\theta x + (1-\theta)y \in E$).

Example 6.7 – Every normed vector space is locally convex because of the triangle inequality.

Non-Example 6.8 – Define $\ell^p(\mathbb{N}) := \left\{(a_n)_{n\in\mathbb{N}} : \sum_n a_n^p < \infty\right\}$. When $0 < p < 1$, we get a metric given by

$$d(a,b) = \sum_n |a_n - b_n|^p .$$

But this is not locally convex because $B_1(0)$ is not. Indeed, let $a = e_1$ and $b = e_2$. Then

$$d(a,0) = d(b,0) = 1.$$

Let $c = \frac{1}{2}a + \frac{1}{2}b$. Then

$$d(c,0) = 2^{1-p} > 1.$$

6.4.1. Seminorm topologies

Definition 6.9

Let $\{\rho_\alpha\}_{\alpha\in\mathcal{A}}$ be a collection of seminorms on a vector space X. We define the **seminorm topology** induced by $\{\rho_\alpha\}$ to have neighborhood base at 0 given by balls in some finite subset of the seminorms, i.e.,

$$\left\{U_{\mathcal{B},\varepsilon} : \mathcal{B} \subseteq \mathcal{A}, \mathcal{B} \text{ finite}, \varepsilon > 0\right\},$$

where

$$U_{\mathcal{B},\varepsilon} = \{x : \rho_\alpha(x) < \varepsilon \text{ for all } \alpha \in \mathcal{B}\}.$$

November 18, 2024

Proposition 6.24

A vector space X equipped with the seminorm topology is a locally convex t.v.s.

Proof. $U_{\mathcal{T},\varepsilon}$ is convex by the triangle inequality for seminorms. Now we check this is a t.v.s.:

$$\rho_\alpha(x + y - (x' + y')) \le \rho_\alpha(x - x') + \rho_\alpha(y - y'),$$

implies addition is continuous, and

$$\begin{aligned}\rho_\alpha(\lambda x - \lambda' x') &\le \rho_\alpha((\lambda - \lambda')x) + \rho_\alpha(\lambda'(x - x')) \\ &\le \left|\lambda - \lambda'\right| \rho_\alpha(x) + |\lambda|\rho_\alpha(x - x') + \left|\lambda - \lambda'\right| \rho_\alpha(x - x')\end{aligned}$$

implies scalar multiplication is continuous. □

Proposition 6.25 (Continuous maps w.r.t. the seminorm topology)
Let $\{\rho_\alpha\}_{\alpha\in\mathcal{A}}$, $\{\sigma_\beta\}_{\beta\in\mathcal{B}}$ be seminorms on X, Y respectively, and $\mathcal{T}$, $\mathcal{S}$ be the induced seminorm topologies. Then $T\colon X \to Y$ is continuous $\iff$ for all $\beta \in \mathcal{B}$, there exists a finite collection of norms $\{\rho_{\alpha_1}, \dots, \rho_{\alpha_N}\}$ and a constant C such that

$$\sigma_\beta(Tx) \le C \sum_{j=1}^{N} \rho_{\alpha_j}(x), \qquad \forall x \in X. \tag{6.1}$$

Proof. ($\Longleftarrow$) Let $\mathcal{F} \subseteq \mathcal{B}$ be a finite set, $\varepsilon > 0$, and $x \in X$. By the assumption, for every $\beta \in \mathcal{B}$, there exists a finite subcollection $\mathcal{R} \subseteq \mathcal{A}$, a set of α's and a constant C_β that satisfies (6.1). Define

$$C \coloneqq \max_{\beta\in\mathcal{F}} C_\beta.$$

Set $\delta = \frac{\varepsilon}{C}$. If $x' \in U_{\mathcal{R},\delta}(x)$ and $\beta \in \mathcal{B}$, then

$$\sigma_\beta(Tx' - Tx) \le C_\beta \sum_{\alpha\in\mathcal{R}} \rho_\alpha(x' - x) < \varepsilon,$$

hence T is continuous.

($\Longrightarrow$) Assume T is continuous. Then for all $\beta \in \mathcal{B}$, there exists a finite subcollection $\mathcal{R} \subseteq \mathcal{A}$ and $\delta > 0$ such that

$$T\left(U_{\mathcal{R},\delta}(0)\right) \subseteq V_{\{\beta\},1}(0),$$

(where $V_{\beta,1}(0)$ is a basis open set in $\mathcal{S}$). For $x \in X$,

$$\frac{\delta}{2} \cdot \frac{x}{\sum_{\alpha\in\mathcal{R}} \rho_\alpha(x)} \in U_{\mathcal{R},\delta}(0).$$

Therefore,

$$\sigma_\beta\left(T\left(\frac{\delta}{2} \cdot \frac{x}{\sum_{\alpha\in\mathcal{R}} \rho_\alpha(x)}\right)\right) \le 1.$$

Pulling out the constants finishes. □

Example 6.9 – Consider the differentiation map $\frac{d}{dx}$ on the vector space $C^\infty([0,1])$. It is clearly a linear operator, but it is *not* a normed operator: let $f_\lambda = e^{\lambda x}$. Then for any λ,

$$\frac{d}{dx}(f_\lambda) = \lambda f_\lambda \implies \left\|\frac{d}{dx}\right\| \ge |\lambda|.$$

As an alternative, the seminorms

$$\rho_k(f) \coloneqq \sup_{x\in[0,1]} \left|f^{(k)}(x)\right|, \qquad k = 0, 1, \dots$$

make $C^\infty([0,1])$ a topological space. In fact, the space is complete in the sense that every Cauchy net converges (we will define this lower in the page).

Example 6.10 – Similar to the previous example, we have a collection of seminorms on $C^\infty(\mathbb{R}^n)$ given by
$$\rho_{N,M}(f) := \sup_{|x|\le N} \max_{\alpha\le M} |\nabla^\alpha f(x)|$$
for $N, M \ge 0$. For $v \in \mathbb{R}^n$, the linear map
$$\begin{aligned}(v\cdot\nabla)\colon C^\infty(\mathbb{R}^n) &\to C^\infty(\mathbb{R}^n)\\ f &\mapsto v\cdot\nabla f\end{aligned}$$
is continuous because
$$\rho_{N,M}(v\cdot\nabla f) \le \|v\|\,\rho_{N,M+1}(f).$$

We will show in the homework what the seminorm topology induced by $\{\rho_\alpha\}$ is Hausdorff and, if we also assume that if $\mathcal{A}$ is countable, metrizable. In particular, this means for all $x \ne 0$, there exists $\alpha \in \mathcal{A}$ such that $\rho_\alpha(x) \ne 0$.

Definition 6.10

Let X be a t.v.s. and $\langle x_\alpha\rangle$ a net. We say $\langle x_\alpha\rangle$ is **Cauchy** if
$$\langle x_\alpha - x_\beta\rangle_{(\alpha,\beta)\in\mathcal{A}^2} \to 0.$$
X is **complete** if Cauchy nets converge.

Definition 6.11

A **Fréchet space** is a complete, Hausdorff vector space X with a seminorm topology induced by a countable collection of seminorms.

Remark 6.26 (We can ignore nets!). If $\{\rho_\alpha\}$ is countable, it suffices to verify Cauchy-ness for sequences, not nets.

Example 6.11 (HW9 Problem 1) – $L^1_{loc}(\mathbb{R}^n)$ is a Fréchet space with its seminorm topology induced by the seminorms
$$\rho_k(f) := \int_{|x|\le k} |f(x)|\,dx.$$
We claim that this topology cannot be induced by a norm.

Suppose we could set a norm on $X = L^1_{loc}(\mathbb{R}^n)$. Let B be the 1-ball in any norm on X. We will show that no ball in the Fréchet topology is contained in B for contradiction. A ball in the Fréchet topology is one where finitely many of the norms $\{\rho_k\}$ are bounded by some ε. Suppose those norms are $\{\rho_{k_1}, \dots, \rho_{k_m}\}$. Let $k^* = \max\{k_1, \dots, k_m\}$. Then the function
$$f_\lambda(x) := \begin{cases} 0 & \text{if } |x| \le k^*,\\ \lambda & \text{if } k^* < |x| \le k^*+1,\\ 0 & \text{otherwise,}\end{cases}$$
is in the ball for any $\lambda \in \mathbb{R}$. Letting $\lambda \to \infty$, we still remain in the ball. But this is clearly a contradiction, since the norm of this function gets arbitrarily large.

6.5. Topologies and modes of convergence in X, X^*, and $\mathcal{L}(X, Y)$

November 20, 2024

Definition 6.12

The **weak topology** on X is the coarsest topology with respect to which all $f \in X^*$ are continuous. In other words, X has the seminorm topology induced by the seminorms

$$\rho_f(x) := |f(x)|\,.$$

Given a net $\langle x_\alpha \rangle$, we say $x_\alpha \rightharpoonup x$ **converges weakly** when $f(x_\alpha) \to f(x)$ for all $f \in X^*$.

Example 6.12 – The sequence

$$\left\{\phi_n \colon [0,1] \to \mathbb{C} \colon x \mapsto e^{2\pi i n x}\right\}_{n \in \mathbb{N}}$$

satisfies $\phi_n \rightharpoonup 0$ converges weakly in $L^1(\lambda)$. [Exercise: return to this example after reading about dual maps and try to prove it.]

Definition 6.13

The **weak*** topology on X^* is the weakest topology with respect to which all $\widehat{x} \in \widehat{X}$ are continuous. Notice that this is weaker than the weak topology if X is not reflexive. In other words, X^* has the seminorm topology induced by the seminorms

$$\rho_x(f) := |f(x)|.$$

Given a net $\langle f_\alpha \rangle$, we say $f_\alpha \overset{w^*}{\rightharpoonup} f$ if $f_\alpha(x) \to f(x)$ for all $x \in X$, so the weak* topology corresponds to the pointwise convergence of a net of functionals.

Theorem 6.27 (Alaoglu's theorem)
Let X be a normed vector space. Then the set

$$K := \{f \in X^* : \|f\| \leq 1\}$$

is weak* compact.

Proof. For $x \in X$, let $F_x := \{z \in \mathbb{C} : |z| \leq \|x\|\}$. Then

$$F := \prod_{x \in X} F_x$$

is compact, since each F_x is, using Tychonoff's theorem (A.16). Elements $f \in F$ can be viewed as functions $f \colon X \to \mathbb{C}$ with $|f(x)| \leq \|x\|$ for all $x \in X$. Notice that $K \subseteq F$; in fact, K is the collection of linear elements of F. Further, the product topology on F induces the weak* topology on K as a subspace of F (indeed, the product topology and the weak* topology both correspond to pointwise convergence in K). K is closed because if $\langle f_\alpha \rangle$ is a net in K and $f_\alpha \to f$, then $f_\alpha \to f$ pointwise, which implies f is linear. □

On the homework, we will prove that if X is separable and $\{f_n\}$ is a bounded linear sequence in X^*, then there exists a subsequence $\{f_{n_k}\}$ which converges weakly.

Definition 6.14

On the space $\mathcal{L}(X,Y)$, we have the norm topology, whose neighborhood base at 0 is given by $\{B_\varepsilon(0) : \varepsilon > 0\}$.

The **strong operator topology** is the weakest topology with respect to which $T \mapsto Tx$ is continuous with respect to $\|\cdot\|_Y$. Equivalently, it is the topology of pointwise convergence. The neighborhood base at 0 is given by

$$\left\{\left\{T : \|Tx_j\| < \varepsilon, j = 1, \ldots, n\right\} : n \in \mathbb{N}, \varepsilon > 0, x_i \in X\right\}.$$

The **weak operator topology** is the weakest topology with respect to which $T \mapsto Tx$ is continuous with respect to the weak topology on Y. Equivalently, it is the topology of *weak* pointwise convergence. The neighborhood base at 0 is given by

$$\left\{\left\{T : \left|f_i(Tx_j)\right| < \varepsilon : i = 1, \ldots, m, j = 1, \ldots, n\right\} : m, n \in \mathbb{N}, \varepsilon > 0, x_j \in X, f_i \in Y^*\right\}.$$

Proposition 6.28

If $\{T_n\} \subseteq \mathcal{L}(X,Y)$ is (norm) bounded and $T_n x \to Tx$ in norm for all $x \in \mathcal{D}$, where $\mathcal{D}$ is dense in X, then $T_n \to T$ in the strong operator topology.

Proof (sketch). Apply the triangle inequality: $T_n(x - y)$ is small for some $y \in \mathcal{D}$. $\square$

7. L^p spaces

Definition 7.1
Let $(X, \mathcal{M}, \mu)$ be a measure space.[1] For $f \colon X \to \mathbb{C}$ measurable and $0 < p < \infty$, define

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

For $p = \infty$, define

$$\|f\|_{L^\infty} := \inf\{\alpha : \mu\{|f| > \alpha\} = 0\} =: \operatorname{ess\,sup}(f),^{2}$$

with the convention that $\sup \varnothing = \infty$.
For $p = 0$, define

$$\|f\|_{L^0} := \mu\{f \neq 0\}.$$

[1]For this section, we will assume that μ is complete and that $\mu(X) > 0$.
[2]ess sup stands for "essential supremum." This definition is motivated by the fact that a supremum *should* still exist for a.e. equivalence classes, since changing the values of a function on a null set shouldn't really count as changing the sup of it. We'll verify this intuition in the next proposition.

This notation is misleading; we still haven't proved that this is a norm. In fact, sometimes it isn't!

Most of the time, we will write $\|f\|_p$ instead of $\|f\|_{L^p}$ for convenience.

Proposition 7.1

$$\begin{aligned}\|f\|_{L^\infty} &= \sup\{\alpha : \mu\{|f| \geq \alpha\} > 0\} \\ &= \inf\{\sup|g| : g = f \text{ a.e.}\}.\end{aligned}$$

Proof. Each element of $\{\alpha : \mu\{|f| > \alpha\} = 0\}$ is an upper bound for every element of $\sup\{\alpha : \mu\{|f| \geq \alpha\} > 0\}$, and every element of $\sup\{\alpha : \mu\{|f| \geq \alpha\} > 0\}$ is a lower bound for $\{\alpha : \mu\{|f| > \alpha\} = 0\}$, giving us the first equality.

($\leq$) follows in the second inequality because $\|f\|_{L^\infty} = \|g\|_{L^\infty} \leq \sup|g|$. For ($\geq$), let $\alpha := \|f\|_{L^\infty}$. Then $\mu\{|f| > \alpha\} \geq 0$ because

$$\{|f| > \alpha\} = \mu\left(\bigcup_n \left\{|f| > \alpha + \frac{1}{n}\right\}\right).$$

Define

$$g(x) := \begin{cases} f(x) & |f| \leq \alpha, \\ 0 & |f| > \alpha. \end{cases}$$

Then g satisfies $\sup|g| = \alpha$. □

Definition 7.2
Define

$$L^p(X) = L^p(X, \mu) = L^p := \left\{f \colon X \to \mathbb{C} : f \text{ measurable}, \|f\|_{L^p} < \infty\right\} / \text{a.e. equivalence}.$$

L^p is not a norm for $0 \leq p < 1$, however, for $0 < p < 1$, $\|f\|_{L^p}^p$ induces a metric:

$$d(f, g) := \|f - g\|_{L^p}^p.$$

But the topology is not locally convex, hence this space is not normable.

As for $p = 0$, L^0 is a terrible space! It is not a t.v.s., since scalar multiplication is not continuous.

7.1. The good L^p spaces

November 22, 2024

Here's the final outcome of this section:

Theorem 7.2

For $1 \leq p \leq \infty$, $\|\cdot\|_{L^p}$ is a norm and L^p is a Banach space.

7.1.1. Hölder's and Minkowski's inequality

Lemma 7.3 (Generalized AM-GM)

For $a, b \geq 0$ and $\theta \in [0, 1]$,

$$a^\theta b^{1-\theta} \leq \theta a + (1-\theta)b,$$

with equality $\iff$ $a = b$.

Proof. See my analysis II notes on page 31. □

Definition 7.3

For $1 \leq p \leq \infty$, the **dual exponent** of p is

$$p' \coloneqq \frac{p}{p-1}.$$

Note that $\frac{1}{p'} = 1 - \frac{1}{p}$, and we let $\infty' = 1$ and $1' = \infty$.

Lemma 7.4 (Hölder's inequality)

For f, g measurable and $1 \leq p \leq \infty$,

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'}.$$

In particular, if the RHS is finite, fg is integrable, $\int |fg| \, d\mu < \infty$ and equality holds $\iff$ there exists $\alpha, \beta \in \mathbb{R}$ with $\alpha\beta \neq 0$ such that $\alpha|f|^p = \beta|g|^{p'}$ a.e.

Proof. If $\|f\|_p$ or $\|g\|_{p'}$ are ∞, then we are done. Similarly, if $\|f\|_p$ or $\|g\|_{p'}$ are 0, we are done.

We may assume (by taking $\frac{f}{\|f\|_p}$ and $\frac{g}{\|g\|_{p'}}$) that $\|f\|_p = \|g\|_{p'} = 1$.

Case I: $p = 1, p' = \infty$. We may assume $\|g\|_\infty = \sup |g| = 1$. Then

$$\int |fg| \, d\mu \leq \int |f| \, d\mu = \|f\|_1.$$

The exact same argument works for $p = \infty$, $p' = 1$.

Case II: $1 < p < \infty$.

$$\begin{aligned}\int |fg|\, d\mu &= \int |f|^{p\cdot\frac{1}{p}}|g|^{p'\cdot\frac{1}{p'}}\, d\mu \\ &\overset{(7.3)}{\leq} \int \frac{1}{p}|f|^p + \frac{1}{p'}|g|^{p'}\, d\mu \\ &= \frac{1}{p} + \frac{1}{p'} \\ &= 1.\end{aligned}$$

Equality holds if and only if it holds for the generalized AM-GM (7.3), so it happens if and only if $|f|^p = |g|^{p'}$ a.e. Factoring in the normalization we did at the start, this implies the result. □

Now, we return to the proof of Theorem 7.2.

Lemma 7.5 (Minkowski's inequality)
For $1 \leq p \leq \infty$,

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p\,.$$

Proof. The case for $p = 1$ follows from the fact that we have the triangle inequality on the absolute value.

For $p = \infty$, pick representatives of f and g so that $\|f\|_\infty = \sup|f|$ and $\|g\| = \sup|g|$ and use properties of the supremeum.

If $1 < p < \infty$,

$$\begin{aligned}\|f+g\|_p^p &= \int |f+g|^p\, d\mu \\ &\leq \int (|f|+|g|)|f+g|^{p-1}\, d\mu \\ &\leq \|f\|_p \left\|(f+g)^{p-1}\right\|_{p'} + \|g\|_p \left\|(f+g)^{p-1}\right\|_{p'} \\ &\leq (\|f\|_p + \|g\|_p)\, \|f+g\|_p^{p-1}\,.\end{aligned}$$

Rearranging this gives us the result. □

Proof of Theorem 7.2. $\|\cdot\|_p$ distinguishes zero, since $\|f\|_p = 0 \iff f = 0$ a.e.

$\|\lambda f\|_p = |\lambda|\,\|f\|_p$ is clear, since $|\lambda f(x)| = |\lambda||f(x)|$ for all $x \in X$, so we can pull it out of the integral in the p-"norm."

Minkowski's inequality (7.5) gives the triangle inequality.

When $p = 1$, we already proved that L^1 is complete.

Case I: $p = \infty$. Let $\{f_n\}$ be a Cauchy sequence. The sequence $\{f_n\}$ being Cauchy in the ∞-norm is the same as it being uniformly Cauchy, which is the same as it being uniformly convergent (of course, up to a null set). Here, we use the fact that we can replace any function $f \in L^\infty$ with an a.e. representative g so that $\|f\|_\infty = \sup|g|$.

Case II: $1 < p < \infty$. If $\sum_n \|f_n\|_p < \infty$, we need to show that f_n converges in L^p, i.e.,

$$\lim_{N\to\infty} \sup_{m,n\geq N} \int \left|\sum_{k=m}^{n} f_k\right|^p d\mu = 0.$$

There are simple functions $\{\varphi_n\}$ such that $|\varphi_n| \leq |f|$ for all n and $\varphi_n \to f$ a.e. Then

$$|f - \varphi_n|^p \leq (2|f|)^p \leq 2^p |f|^p \in L^p.$$

So $\int |f - \varphi_n|^p \, d\mu$ converges to 0 by the dominated convergence theorem. □

Corollary 7.6
For $1 \leq p < \infty$, $C^0_{\text{cpct}}(\mathbb{R}^n)$ (continuous function with compact support) are dense in $L^p(\mathbb{R}^n)$.

Remark 7.7. The corollary is very false for $p = \infty$, because L^∞ convergence (after picking a "good" representative) implies uniform convergence.

Remark 7.8. Let

$$C^0(\mathbb{R}^n) := \{\text{bdd. continuous functions on } \mathbb{R}^n\},$$

$$C^0_0(\mathbb{R}^n) := \{\text{bdd. continuous functions on } \mathbb{R}^n \text{ that } \textit{vanish at zero}\},$$

where **vanishing at zero** means that $\lim_{|x| \to \infty} |f(x)| = 0$.

These are both closed subsets of L^∞. They also have closed interior, because any small perturbation will make any function in these spaces no longer continuous, hence they are both meager.

7.1.2. Sums and intersections of Banach spaces

Proposition 7.9 (Intersections and sums of Banach spaces are Banach)
Let $X, Y \subseteq Z$ be (vector) Banach subspaces of a vector space Z. Then

$$X + Y \subseteq Z, \quad \text{and} \quad X \cap Y \subseteq Z$$

are both Banach spaces with norms

$$\|f\|_{X \cap Y} := \|f\|_X + \|f\|_Y, \qquad \|f\|_{X+Y} := \inf_{\substack{g+h=f \\ g \in X, h \in Y}} \|g\|_X + \|h\|_Y.$$

We'll now adopt the notation $|f| \lesssim |g|$ to mean "$|f|$ is less than or equal to a constant times $|g|$."

Proposition 7.10 (Sum of L^p spaces)
For $0 \leq p \leq q \leq r \leq \infty$,

$$L^q \subseteq L^p + L^r.$$

In fact,

$$\|f\|_{L^p + L^r} \lesssim \|f\|_{L^q}.$$

Proof. Let $f \in L^q$. $f = 0$ a.e. is trivial. Otherwise, normalize f so that $\|f\|_r = 1$. Let $g := f\chi_{\{|f| \leq 1\}}$, and $h := f\chi_{\{|f| > 1\}}$. Then $g \in L^r$ for all $q \leq r \leq \infty$, and

$$\|g\|_r^r = \int |g|^{r-q} |g|^q \, d\mu \leq \int |f|^q \, d\mu = 1$$

(if $r \neq \infty$. Otherwise, $\|g\|_\infty \leq 1$ is clear). Also, $h \in L^p$ for all $0 \leq p \leq q$, and

$$\|h\|_p^p = \int |h|^{p-q}|h|^q \, d\mu \leq \int |f|^q \, d\mu = 1$$

(if $p \neq 0$. Otherwise, $\|h\|_0 \leq 1$ is clear). □

Proposition 7.11 (Intersection of L^p spaces)
If $0 < p \leq q \leq r \leq \infty$, then

$$L^p \cap L^r \subseteq L^q,$$

and if $p > 0$, then

$$\|f\|_q \leq \|f\|_p^{1-\theta} \|f\|_r^\theta \overset{(7.3)}{\leq} (1-\theta)\|f\|_p + \theta \|f\|_r,$$

where $\theta \in (0,1)$ is the value so that $q^{-1} = p^{-1}(1-\theta) + r^{-1}\theta$, i.e. $\theta = \frac{q^{-1}-r^{-1}}{p^{-1}-r^{-1}}$.

Proof. When $r = \infty$, we have $|f|^q \leq \|f\|_\infty^{q-p} |f|^p$, and our value of θ is $\frac{p}{q}$, so

$$\|f\|_q \leq \|f\|_p^{p/q} \|f\|_\infty^{1-(p/q)} = \|f\|_p^\theta \|f\|_\infty^{1-\theta}.$$

When $r < \infty$, we use Hölder's inequality (7.4) with conjugate exponents $\frac{r}{q\theta}$ and $\frac{p}{q(1-\theta)}$. Then

$$\begin{aligned}
\|f\|_q^q &= \int |f|q(1-\theta) + q\theta \, d\mu \\
&= \int |f|^{q(1-\theta)}|f|^{q\theta} \, d\mu \\
&\leq \int \left(|f|^{q(1-\theta)\frac{p}{q(1-\theta)}}\right)^{\frac{q(1-\theta)}{p}} \left(|f|^{q\theta\frac{r}{q\theta}}\right)^{\frac{q\theta}{r}} d\mu \\
&= \|f\|_r^{q(1-\theta)} \|f\|_r^{q\theta}.
\end{aligned}$$

□

7.1.3. Special cases: counting and finite measures

There are some special circumstances of what L^p spaces depending on the measure space $(X, \mathcal{M}, \mu)$. If μ is the counting measure on $\mathcal{A}$, for $0 \leq p \leq \infty$, $\ell^p(\mathcal{A}) = L^p(\mathcal{A}, \mathcal{P}(\mathcal{A}), \mu)$. For $0 < p < \infty$,

$$\sum_{\alpha \in \mathcal{A}} |f(\alpha)|^p < \infty \iff \{f \neq 0\} \text{ is countable.}$$

Proposition 7.12
For a set $\mathcal{A}$ and $0 \leq p \leq q \leq \infty$,

$$\ell^p(\mathcal{A}) \subseteq \ell^q(\mathcal{A}),$$

and if $p > 0$,

$$\|f\|_{\ell^q(\mathcal{A})} \lesssim \|f\|_{\ell^p(\mathcal{A})}.$$

Now we consider when μ is a finite measure.

Proposition 7.13
If $\mu(X) < \infty$ and $0 \le p \le q \le \infty$, then

$$L^q(X) \subseteq L^p(X),$$

and if $p > 0$,

$$\|f\|_p \le \mu(X)^{p^{-1}-q^{-1}} \|f\|_q .$$

Proof of inequality. When $p > 0$, the second claim implies the first. Notice that $|f|^p = |f|^p \cdot 1$. Using Hölder's inequality,

$$\begin{aligned}
\|f\|_p^p = \int |f|^p \, d\mu &\le \left(\int (|f|^p)^{\frac{q}{p}} \, d\mu\right)^{\frac{p}{q}} \left(\int 1^{\frac{q}{q-p}} \, d\mu\right)^{\frac{q-p}{q}} \\
&= \left(\int |f|^q \, d\mu\right)^{\frac{p}{q}} \mu(X)^{\frac{q-p}{q}} \\
&= \|f\|_q^p \, \mu(X)^{\frac{q-p}{q}} .
\end{aligned}$$

Taking the pth root of both sides gives the result. □

7.2. Dual L^p spaces

Proposition 7.14
For $g \in L^{p'}$, define

$$\Lambda_g(f) := \int fg \, d\mu, \quad \text{(when this makes sense).}$$

Then, in particular, $\Lambda_g \in (L^p)^*$, and $\|\Lambda_g\|_{(L^p)^*} \le \|g\|_{p'}$; if $1 < p \le \infty$ or μ is semifinite, we have equality.

Proof. ($\le$) Let $\|f\|_p = 1$. Then

$$|\Lambda_g f| = \left|\int fg \, d\mu\right| \le \int |fg| \, d\mu \overset{(7.4)}{\le} \|f\|_p \|g\|_{p'} = \|g\|_{p'} .$$

($\ge$) If $g = 0$ a.e. we are done. Otherwise,
Case I: $1 < p < \infty$. Take $f := \overline{g}\,|g|^{p'-2}$. Then

$$|f| = |g|^{p'-1} \implies |f|^p = (|g|^{p'-1})^p = |g|^{p'}.$$

So $f \in L^p$, $\|f\|_{L^p} = \|g\|_{p'}^{p'-1}$, and

$$|\Lambda_g f| = \|g\|_{p'}^{p'} = \|f\|_p \|g\|_{p'} .$$

Case II: $p = \infty$. Define $f := \overline{g}|g|^{-1} = \overline{\operatorname{sgn} g}$. Then $f \in L^\infty$, $\|f\|_\infty = 1$, and

$$|\Lambda_g f| = 1.$$

Case III: $p = 1$ and μ semi-finite. We may assume that $\|g\|_\infty = \sup |g|$ by choosing an appropriate a.e. representative. For $\varepsilon > 0$, there exists a set E_ε with positive, finite measure such that $|g| \geq \sup |g| - \varepsilon$ on E_ε. Define $f := \overline{g}\chi_{E_\varepsilon}$. Then

$$\mu(E_\varepsilon) \leq \|f\|_1 \leq \|g\|_\infty \mu(E_\varepsilon).$$

Hence,

$$\Lambda_g f = \int |g|^2 \chi_{E_\varepsilon} \, d\mu \leq \|g\|_\infty^2 \mu(E_\varepsilon),$$

and

$$\Lambda_g f \geq (\|g\|_\infty - \varepsilon)^2 \mu(E_\varepsilon).$$

Now take $\varepsilon \to 0$ to finish. □

November 27, 2024

Theorem 7.15

1. For $1 < p < \infty$, the map $\Lambda_\bullet : L^{p'} \to (L^p)^*$ given by $g \mapsto \Lambda_g$ is a surjective isometry. In particular, we have

$$\|f\|_{p'} = \|\Lambda_f\|_{(L^p)^*} \sup_{\|g\|=1} \left| \int fg \, d\mu \right|.$$

2. For $p = 1$, the above map from $L^\infty \to (L^1)^*$ is an isometry. It is surjective if and only if μ is σ-finite.

3. For $p = \infty$, the above map from $L^1 \to (L^\infty)^*$ is an isometry.

Proof. Proposition 7.14 gives us the isometry part for (1-3). Now we must show when $1 < p < \infty$ or $p = 1$, $\Lambda_\bullet$ is a surjection, so for $\varphi \in (L^p)^*$, we want $g \in L^{p'}$ such that $\varphi = \Lambda_g$.

Case I: $1 \leq p < \infty$ and μ finite. Let $\varphi \in (L^p)^*$. For $E \in \mathcal{M}$, define $\nu_\varphi := \varphi(\chi_E)$. Note that $\chi_E \in L^p$ because μ is finite. We can verify that ν_φ is a complex measure with $\nu_\varphi \ll \mu$; in particular, countable additivity comes from the continuity of φ. Let $g_\varphi = \frac{d\nu_\varphi}{d\mu}$ be the Radon-Nikodym derivative. By definition, Λ_{g_φ} is our candidate for φ.

We now show $\int |g_\varphi|^p \, d\mu < \infty$. Set $g_n := g_\varphi \chi_{|g_\varphi| \leq n}$. Then

$$\begin{aligned} \|g_n\|_{p'}^{p'} &= \int |g_n|^{p'} \, d\mu \\ &= \int g_n \overline{g_n} |g_n|^{p'-2} \, d\mu \\ &= \varphi(\overline{g_n} |g_n|^{p'-2}) \\ &\leq \|\varphi\|_{(L^p)^*} \left\| \overline{g_n} |g_n|^{p'-2} \right\|_p \\ &= \|\varphi\|_{(L^p)^*} \|g_n\|_{p'}^{p'-1}. \end{aligned}$$

We rearrange to get $\|g_n\|_{p'} \leq \|\varphi\|_{(L^p)^*}$. By the monotone convergence theorem, $g_\varphi \in L^p$.

Case II: $1 \le p < \infty$ and μ σ-finite. This follows using the same ideas as in the measure theory section to bring a proof with a finite measure space to one with a σ-finite measure space.

Case III: $1 < p < \infty$ and μ arbitrary. Let $E \in \mathcal{M}$ be σ-finite. By Case II, we can define a function $g_{\varphi,E}$ that is supported on E, but zero off E, with $g_{\varphi,E} \in L^{p'}$. Set

$$M := \sup_{\substack{E \in \mathcal{M} \\ E\ \sigma\text{-finite}}} \|g_{\varphi,E}\|_{p'} \le \|\varphi\|_{(L^p)^*}.$$

Choose a sequence $\{E_n\} \subseteq \mathcal{M}$ of σ-finite sets with $\|g_{\varphi,E_n}\|_{p'} \to M$. Consider the set $E := \bigcup_n E_n$. By the a.e. equivalence of g's, we see that $\|g_{\varphi,E}\|_{p'} = M$. Let $E \subseteq A \in \mathcal{M}$, where A is σ-finite. By maximality, we see that $g_{\varphi,A\setminus E} = 0$ a.e. Now, we show $\varphi = \Lambda_{g_{\varphi,E}}$. Indeed, if $f \in L^p(\mu)$, then we know $\{f \neq 0\}$ is σ-finite, so we can restrict to $A = \{f \neq 0\}$. $\square$

Corollary 7.16

For $1 < p < \infty$, L^p is reflexive.

Remark 7.17. $L^1 \to (L^\infty)^*$ is almost never surjective.

Non-Example 7.1 – Consider $\ell^\infty(\mathbb{N})$ and look at the functionals $\varphi_n(a) = a_n$. We know that $\|\varphi_n\|_{(\ell^\infty)^*} = 1$. [Exercise: show that there is no weak* convergent subsequence.] By Alaoglu's theorem (6.27), φ_n has a cluster point φ. We claim that $\varphi \notin \operatorname{im}(\Lambda_\bullet)$. Assume for contradiction that $\varphi = \Lambda_b$. Then we would have

$$\varphi(a) = \sum_n a_n b_n.$$

Since $\|\varphi\| = 1$, we would have $\sum_n |b_n| = 1$. But this is impossible since we cannot have a subsequence a_{n_k} such that

$$\lim_{k\to\infty} a_{n_k} = \sum_n a_n b_n.$$

Non-Example 7.2 – Here's another case of $L^1 \to (L^\infty)^*$ not being surjective. Let $X = [0,1]$ with the Lebesgue measure. We can view continuous functions on $[0,1]$, $C(X)$, as a subspace of L^∞, since there is only *one* a.e. representative for each continuous function on $[0,1]$. [Convince yourself this is true!] This makes the map $f \mapsto f(0)$ is a well-defined bounded linear functional in $C(X)^*$, and so we can extend with Hahn-Banach to map $\phi \in (L^\infty)^*$.

The step functions $f_n(x) := \max(1 - nx, 0)$ satisfy $\phi(f_n) = 1$, but $f_n \to 0$ for all $x > 0$, so $\int f_n g\,dx \to 0$ for all $g \in L^1$.

7.3. Distribution functions and weak L^p

December 2, 2024 Recall the *distribution function* of a measurable function $f\colon X \to \mathbb{C}$ is given by

$$\lambda_f(\alpha) := \mu\{|f| > \alpha\}.$$

Proposition 7.18 (Chebyshev's inequality)
Let $f\colon X \to \mathbb{C}$ be measurable, $\alpha \geq 0$, and $0 < p < \infty$. Then

$$\alpha^p \lambda_f(\alpha) \leq \|f\|_p^p .$$

Proof.

$$\alpha^p \lambda_f(\alpha) = \int \left(\alpha \chi_{\{|f|>\alpha\}}\right)^p \, d\mu \leq \int |f|^p \, d\mu = \|f\|_p^p . \qquad \square$$

Definition 7.4
$f \in L^{p,\infty}$ (**weak** L^p) if

$$[f]_{p,\infty} := \left(\sup_{\alpha>0} \alpha \lambda_f(\alpha)\right)^{\frac{1}{p}}$$

is finite.

$[\,\cdot\,]_{p,\infty}$ is *not* a norm because it doesn't obey the triangle inequality. However, it does induce a metric

$$d(f,g) := [f-g]_{p,\infty}$$

that makes $L^{p,\infty}$ a t.v.s. for $0 < p < \infty$.
It *is* a normable space for $1 < p < \infty$ with

$$\|f\|_{p,\infty} := \sup_{\substack{E \in \mathcal{M} \\ \mu(E)>0}} \int |f| \frac{1}{\mu(E)^{\frac{1}{p}}} \, d\mu.$$

Observe that $L^p \subseteq L^{p,\infty}$ because $[f]_{p,\infty} \leq \|f\|_p$ (by Chebyshev's inequality (7.18)).

Example 7.3 – This containment is normally strict. For example, if we use the Lebesgue measure on $\mathbb{R}^n$,

$$f(x) := \|x\|^{-n/p} \in L^{p,\infty} \setminus L^p.$$

Example 7.4 – The **p-maximal function**,

$$M_p f(x) := \sup_{r>0} \left(\fint |f(y)|^p \, d\mu(y) \right)^{\frac{1}{p}}$$

is in $L^{p,\infty}$ and satisfies the inequality

$$[M_p f]_{p,\infty} \leq C \, \|f\|_p$$

for some constant C by the same argument as when we proved the weak type-(1,1) bound for the Hardy-Littlewood maximal operator.

Notice that this is the Hardy-Littlewood maximal operator when $p = 1$.

7.4. Complex interpolation

We proved that given $1 \leq p < r \leq \infty$, any q satisfying $p < q < r$ has

$$(L^p \cap L^r) \subseteq L^q \subseteq (L^p + L^r).$$

The motivation for the following theorem is that we want to know if a linear operator T on $L^p + L^r$ that is bounded on L^p and L^r individually is also bounded on L^q.

Theorem 7.19 (Riesz-Thorin/complex interpolation)
Let $(X, \mathcal{M}, \mu)$, $(Y, \mathcal{N}, \nu)$ be measure spaces. Let $p_0, p_1, q_0, q_1 \in [1, \infty]$. If $q_0 = q_1 = \infty$, assume ν is semifinite. Let

$$T\colon L^{p_0}(\mu) + L^{p_1}(\mu) \to L^{q_0}(\nu) + L^{q_1}(\nu)$$

be a linear operator satisfying

$$\|Tf\|_{L^{q_j}(\nu)} \le M_j \|f\|_{L^{p_j}(\mu)}, \qquad j = 0, 1.$$

Set $p_t^{-1} := p_0^{-1}(1-t) + p_1^{-1}t$ and $q_t^{-1} := q_0^{-1}(1-t) + q_1^{-1}t$ for $t \in (0,1)$. Then

$$\|Tf\|_{L^{q_t}(\nu)} \le M_0^{1-t} M_1^t \|f\|_{L^{p_t}(\mu)},$$

for $f \in L^{p_t}$, $t \in (0,1)$.

Define $M_t := M_0^{1-t} M_1^t$ as the bounding constant in the Riesz-Thorin theorem. We will need a lemma from complex analysis.

Lemma 7.20 (The three lines lemma)
Let $S = \{z \in \mathbb{C} : 0 < \operatorname{Re} z < 1\}$. Let $\varphi\colon \overline{S} \to \mathbb{C}$ be (1) continuous and bounded on $\overline{S}$, (2) holomorphic on S, and (3) satisfy $|\varphi(z)| \le M_{\operatorname{Re} z}$ on $\partial S = \{z \in \mathbb{C} : \operatorname{Re} z \in \{0,1\}\}$. Then $|\varphi(z)| \le M_{\operatorname{Re} z}$ for all $z \in S$.

Proof. We may assume $M_0 \neq 0$ and $M_1 \neq 0$. Define

$$\varphi_\varepsilon(z) := M_z^{-1} e^{\varepsilon(z^2 - z)} \varphi(z), \qquad M_z := M_0^{1-z} M_1^z.$$

Then φ_ε is holomorphic on S, continuous on $\overline{S}$, and, letting $z = x + iy$,

$$|\varphi_\varepsilon(z)| = M_{\operatorname{Re} z}^{-1} e^{\varepsilon(x^2 - x - y^2)} |\varphi(z)|.$$

This is bounded on $\overline{S}$. In fact, it is bounded by 1 when $\operatorname{Re} z = x = 0, 1$. Since

$$\lim_{z \to i\infty} |\varphi_\varepsilon(z)| = \lim_{y \to \infty} M_{\operatorname{Re} z}^{-1} e^{\varepsilon(x^2 - x - y^2)} |\varphi(z)| = 0,$$

$|\varphi_\varepsilon(z)| \le 1$ on the boundary of the rectangle $R_A := [0,1] + i[-A, A]$ for sufficiently large A. By the maximum modulus principle, $|\varphi_\varepsilon| \le 1$ on all of R_A. Sending $A \to \infty$, $|\varphi_\varepsilon| \le 1$ on S. Sending $\varepsilon \searrow 0$, $|\varphi(z)| \le 1$ on S. $\square$

Proof of Theorem 7.19. Case I: $p_0 = p_1$. Let $p = p_0 = p_1$. For $f \in L^p$,

$$\begin{aligned} \|Tf\|_{q_t} &\le \|Tf\|_{q_0}^{1-t} \|Tf\|_{q_1}^t \\ &\le M_0^{1-t} \|f\|_p^{1-t} M_1^t \|f\|_p^t \end{aligned}$$

by Hölder's inequality, since $1 = \frac{(1-t)q}{q_0} + \frac{tq}{q_1}$.

Case II: $p_0 < p_1$. $p_0 < p_1$ implies $p_t \neq \infty$ for $t \in (0,1)$. Hence, the class Σ_X of measurable simple functions $f\colon X \to \mathbb{C}$ that vanish outside sets of finite measure is dense in L^{p_t} for $t \in (0,1)$.

By assumption, either ν is semifinite, or $q_t \neq \infty$ for all $t \in (0,1)$. Hence, for $g \in L^{q_t}$,

$$\|g\|_{q_t} = \sup\left\{\left|\int gh\,d\nu\right| : h \in \Sigma_X, \|h\|_{q'_t} = 1\right\}.$$

Lemma 7.21

Assume $p_0 < p_1$ and that ν is semi-finite, or q_i are not both ∞. Let $f \in \Sigma_X$, $g \in \Sigma_Y$, as defined above. Then $\left|\int (Tf)g\,d\nu\right| \leq M_t \|f\|_{p_t} \|g\|_{q'_t}$ for $t \in (0,1)$.

Proof. Let $f = \sum_j a_j \chi_{E_j}$, $g = \sum_k b_k \chi_{F_k}$. Write the coefficients in polar form: $a_j = |a_j| e^{i\theta_j}$, $b_k = |b_k| e^{i\vartheta_k}$. Let $p_z^{-1} := (1-z)p_0^{-1} + zp_1^{-1}$, $(q'_z)^{-1} := (1-z)(q'_0)^{-1} + z(q'_1)^{-1}$ for $z \in \overline{S}$. Define an operator $(\bullet)_z$ by

$$f_z = \left(\sum_j a_j \chi_{E_j}\right)_z := \sum_j |a_j|^{p_t/p_z} e^{i\theta_j} \chi_{E_j}$$
$$g_z = \left(\sum_k b_k \chi_{F_k}\right)_z := \sum_k |b_k|^{q'_t/q'_z} e^{i\vartheta_k} \chi_{F_k}$$

Define

$$\begin{aligned}\varphi(z) &:= \int (Tf_z) g_z\,d\nu,\\ &= \sum_{j,k} |a_j|^{p_t/p_z} |b_k|^{q'_t/q'_z} e^{i(\theta_j + \vartheta_k)} \int (T\chi_{E_j})\chi_{F_k}\,d\nu.\end{aligned}$$

Then φ is holomorphic on S and continuous and bounded on $\overline{S}$. Since $f_z \in L^{p_{\operatorname{Re} z}}$, $q_z \in L^{q'_{\operatorname{Re} z}}$,and $\|f_z\|_{r_{\operatorname{Re} z}} = \|g_z\|_{q'_{\operatorname{Re} z}}$,

$$|\varphi(z)| \leq M_j \quad \text{on } \{\operatorname{Re} z = j\} \text{ for } j = 0, 1.$$

By the three lines lemma (7.20), $|\varphi(z)| \leq M_{\operatorname{Re} z}$ for all $z \in S$. ∎

Using the above lemma and the fact that Σ_X is dense in L^{p_t} (since $p_t < \infty$), $T|_{\Sigma_X}$ extends uniquely to $T_t^{\text{ext}} \in \mathcal{L}(L^{p_t}, L^{q_t})$ with

$$\left\|T_t^{\text{ext}}\right\| \leq M_t.$$

It remains to show $T = T_t^{\text{ext}}$. Let $f \in L^{p_t}$. Take $\varphi_n \in \Sigma_X$ such that $\varphi_n \to f$ pointwise, $|\varphi_n| \leq |f|$ for all n, and $\varphi_n \to f$ uniformly on $A := \{|f| \leq 1\}$. We decompose f based on A: $f^{\leq 1} := f\chi_A$, $f^{>1} := f\chi_{A^C}$; so $f = f^{\leq 1} + f^{>1}$. Similarly decompose $\varphi_n = \varphi_n^{\leq 1} + \varphi_n^{>1}$. Then

$$\begin{aligned}\varphi_n^{\leq 1} &\to f^{\leq 1} \qquad \text{in } L^{p_1} \text{ and } L^{p_t} \text{ for } t \in (0,1),\\ \varphi_n^{>1} &\to f^{>1} \qquad \text{in } L^{p_0} \text{ and } L^{p_t} \text{ for } t \in (0,1),\end{aligned}$$

(in the case $p_1 = \infty$, use the fact that the convergence in uniform). Hence,

$$T\varphi_{\frac{\leq 1}{n}} \to \begin{cases} Tf^{\leq 1} & \text{in } L^{p_1}, \\ T^{\text{ext}} f^{\leq 1} & \text{in } L^{p_t} \text{ for } t \in (0,1). \end{cases}$$

So $T\varphi_{\frac{\leq 1}{n}} \to Tf^{\leq 1}$ and $T\varphi_{\frac{\leq 1}{n}} \to T^{\text{ext}} f^{\leq 1}$ in measure, so $Tf^{\leq 1} = T^{\text{ext}} f^{\leq 1}$. Likewise, $Tf^{>1} = T^{\text{ext}} f^{>1}$, which implies $T = T^{\text{ext}}$. □

7.5. Applications of complex interpolation and some Fourier analysis

December 6, 2024

Definition 7.5

For $f \in L^p(\mathbb{R}^n, dx)$, $g \in L^q(\mathbb{R}^n, dx)$, define the **convolution** of f and g to be the integral

$$\begin{aligned} f * g(x) &:= \int f(y) g(x-y)\, dy, \\ &= \int f(x-y) g(y)\, dy. \end{aligned}$$

Theorem 7.22 (Young's convolution inequality)
For $f \in L^p(\mathbb{R}^n, dx)$, $g \in L^q(\mathbb{R}^n, dx)$, where $1 \leq p < q \leq \infty$ satisfies $\frac{1}{p} + \frac{1}{q} \geq 1$,

(a) $f * g(x)$ converges for a.e. x,

(b) $f * g \in L^r$ where $r^{-1} = p^{-1} + q^{-1} - 1$

(c) We have the inequality
$$\|f * g\|_r \leq \|f\|_p \cdot \|g\|_q .$$

Proof. We seek to apply complex interpolation, so we only consider "endpoints."

Case 1: $(p, q, r) = (p, p', \infty)$. Is done by Hölder.

Case 2: $(p, q, r) = (p, 1, p)$. If $f \in L^p$, $g \in L^1$, then $f * g$ is defined for a.e. $x \in \mathbb{R}^n$. Assume $f, g \geq 0$. Let $h \in L^{p'}$. We may assume $h \geq 0$ (since $f * g \geq 0$). Applying Tonelli,

$$\begin{aligned} \int f * g(x) h(x)\, dx &= \iint f(x-y) g(y) h(x)\, dy\, dx \\ &\leq \int \|f\|_p \|h\|_{p'} g(y)\, dy \qquad \text{(Case I)} \\ &\leq \|f\|_p \|h\|_{p'} \|g\|_1 . \end{aligned}$$

Now the result follows.

General case. Let $f \in L^0$ and $g \in L^1$. By the previous case, this converges for a.e. x. Moreover,

$$\|f * g\|_p \leq \||f| * |g|\|_p \leq \|f\|_p \|g\|_1 .$$

We may now define an operator for fixed $f \in L^p$ by

$$T_f g := f * g.$$

Then T_f is a map from $L^{p'} + L^1$ to $L^\infty + L^p$, satisfying

$$\|T_f g\|_\infty \leq \|f\|_p \|g\|_{p'} \qquad \text{for } g \in L^{p'},$$
$$\|T_f g\|_p \leq \|f\|_p \|g\|_1 \qquad \text{for } g \in L^1.$$

Applying complex interpolation (7.19), we are done. □

This proof hints at a more general result.

Proposition 7.23 (Minkowski's integral inequality)
Assume $(X, \mathcal{M}, \mu)$, $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. Let K be $\mathcal{M} \otimes \mathcal{N}$-measurable and assume

$$\int \|K(-,y)\|_p \, d\nu(y) < \infty.$$[1]

Then $g(x) := \int K(x,y) \, d\nu(y)$ converges a.e., $g \in L^p(\mu)$, and

$$\left\| \int K(-,y) \, d\nu(y) \right\|_p \leq \int \|K(-,y)\|_p \, d\nu(y).$$

Notice this is a "triangle-like" inequality.

[1]For any $\mathcal{M} \otimes \mathcal{N}$-measurable function $K(x,y)$ in two variables, the notation $\|K(-,y)\|_p$ means taking the norm in the first coordinate (with the measure μ).

Proof (sketch). As before, we want to "test" K against a function h.
Case I: $K \geq 0$. Let $h \in L^{p'}(\mu)$ be non-negative. Then

$$\begin{aligned}
\int_X g(x)h(x) \, d\mu(x) &= \int_X \int_Y K(x,y) \, d\nu(y) h(x) \, d\mu(x) \\
&= \int_Y \int_X K(x,y) h(x) \, d\mu(x) \, d\nu(y) \\
&\leq \int_Y \|K(-,y)\|_p \|h\|_{p'} \, d\nu(y).
\end{aligned}$$

General case. The integral converges absolutely a.e. by Case I. The rest follows from

$$\left\| \int K(-,y) \, d\nu(y) \right\|_p \leq \left\| \int |K(-,y)| \, d\nu(y) \right\|_p$$

and Case I. □

We have another extension of this result.

Proposition 7.24 ($L^q_y(L^p_x) \subseteq L^p_x(L^q_y)$)
Let $1 \leq q \leq p \leq \infty$, and $(X, \mathcal{M}, \mu)$, $(Y, \mathcal{N}, \nu)$ σ-finite. Then, given an $\mathcal{M} \otimes \mathcal{N}$-measurable function K,

$$\left\| \left(\int |K(-,y)|^q \, d\nu(y) \right)^{\frac{1}{q}} \right\|_p \leq \left\| \left(\int |K(x,-)|^p \, d\mu(x) \right)^{\frac{1}{p}} \right\|_q .$$

Proof (sketch). The idea is similar to Minkowski's integral inequality (7.23); we want to prove the map

$$Q\colon x \mapsto \int |K(x,y)|^q \, d\nu(y)$$

is in $L^{p/q}(\mu)$. To do this, test against functions $h \in L^{(p/q)'}(\mu)$. □

Proposition 7.25

Let $(X, \mathcal{M}, \mu)$, $(Y, \mathcal{N}, \nu)$ be σ-finite. Let K be an $\mathcal{M} \otimes \mathcal{N}$-measurable function. Assume there exists a constant A such that for a.e. x, $\int |K(x,y)| \, d\nu(y) \le A$ and for a.e. y, $\int |K(x,y)| \, d\mu(x) \le A$. Then $Tf(x) := \int K(x,y) f(y) \, d\nu(y)$ converges a.e. for $f \in L^p(\nu)$, $T\colon L^p(\nu) \to L^p(\mu)$, and there exists a constant B_p such that $\|T\| \le CB_p$.

Proof (ideas). The main step is with $f \in L^p(\nu)$ non-negative. Let $h \in L^{p'}(\nu)$ be a non-negative "testing" function. Then

$$\begin{aligned}
\int (Tf(x)) h(x) \, d\mu(x) &= \int_X \int_Y K(x,y) f(y) h(x) \, d\nu(y) \, d\mu(x) \\
&= \int_X \int_Y K(x,y)^{\frac{1}{p}} f(y) K(x,y)^{\frac{1}{p'}} h(x) \, d\nu(y) \, d\mu(x) \\
&\le \underbrace{\left\| K(x,y)^{\frac{1}{p}} f(y) \right\|_{L^p(\mu \times \nu)}}_{\text{(A)}} \underbrace{\left\| K(x,y)^{\frac{1}{p'}} h(y) \right\|_{L^{p'}(\mu \times \nu)}}_{\text{(B)}},
\end{aligned}$$

where the last step is by Hölder's inequality. We now bound (A) and (B) separately:

$$\begin{aligned}
\left\| K(x,y)^{\frac{1}{p}} f(y) \right\|_{L^p(\mu \times \nu)} &= \left(\int_Y \int_X K(x,y) f(y)^p \, d\mu(x) \, d\nu(y) \right)^{\frac{1}{p}} \\
&= \left(\int_Y \int_X K(x,y) \, d\mu(x) f(y)^p \, d\nu(y) \right)^{\frac{1}{p}} \\
&\le A^{\frac{1}{p}} \|f\|_p, \\
\left\| K(x,y)^{\frac{1}{p'}} h(y) \right\|_{L^{p'}(\mu \times \nu)} &= \left(\int_Y \int_X K(x,y) h(y)^{p'} \, d\mu(x) \, d\nu(y) \right)^{\frac{1}{p'}} \\
&= \left(\int_Y \int_X K(x,y) \, d\mu(x) h(y)^{p'} \, d\nu(y) \right)^{\frac{1}{p'}} \\
&\le A^{\frac{1}{p'}} \|h\|_{p'}.
\end{aligned}$$

□

7.6. Lorentz spaces and real interpolation

December 9, 2024 Recall the Hardy-Littlewood maximal function:

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| \, dy,$$

and how we used it to prove for all $f \in L^1$, $\lambda |\{Mf > \lambda\}| \le C_n \|f\|_1$. Moreover, we have $g \in L^\infty$, $\|Mg\|_\infty \le \|g\|_\infty$. So we have bounds at the "endpoints" 1 and ∞ of the norms for the operator M. Unfortunately, M is not linear, and even worse, Mf is only weak L^p, so we cannot apply complex interpolation (7.19). Real interpolation gives us hope to interpolate for a different class of operators.

Definition 7.6

Let T be a map from some vector space $\mathcal{D}$ contained in the measurable functions on $(X, \mathcal{M}, \mu)$ to the measurable functions on $(Y, \mathcal{N}, \nu)$.

T is called **sublinear** if $|T(f+g)| \leq |Tf| + |Tg|$ and $|T(cf)| = c\,|Tf|$ for all $f, g \in \mathcal{D}$ and $c > 0$.

A sublinear map T is of **strong type** (p, q) $(1 \leq p, q \leq \infty)$ if $L^p(\mu) \subseteq \mathcal{D}$, $T\colon L^p(\mu) \to L^q(\nu)$, and there exists $C > 0$ such that $\|Tf\|_q \leq C\,\|f\|_p$ for all $f \in L^p$.

A sublinear map T is of **weak type** (p, q) $(1 \leq p, q \leq \infty)$ if $L^p(\mu) \subseteq \mathcal{D}$, $T\colon L^p(\mu) \to L^q(\nu)$, and there exists $C > 0$ such that $[Tf]_q \leq C\,\|f\|_p$ for all $f \in L^p$.

Definition 7.7

For $1 \leq p < \infty$ and $1 \leq q \leq \infty$, we say $f \in L^{p,q}$ if

$$\begin{aligned}
[f]_{p,q} &:= \left\| \lambda \left(\mu\{|f| > \lambda\}\right)^{\frac{1}{p}} \right\|_{L^q\left(\frac{d\lambda}{\lambda}\right)} \\
&:= \begin{cases} \sup \lambda \left(\mu\{|f| > \lambda\}\right)^{\frac{1}{p}} & q = \infty, \\ \left(\int_0^\infty \lambda^{q-1} \left(\mu\{|f| > \lambda\}\right)^{\frac{q}{p}}\, d\lambda\right)^{\frac{1}{q}} & 1 \leq q < \infty, \end{cases}
\end{aligned}$$

is finite. [Sanity check: check why the "$L^q\left(\frac{d\lambda}{\lambda}\right)$" norm makes sense for the formula in the $1 \leq q < \infty$ case.]

Some warnings:

- This is a quasi-norm (i.e. the triangle inequality is replaced by $\|x+y\| \leq K(\|x\| + \|y\|)$ for some constant $K > 1$).
- When $1 < p < \infty$, $L^{p,q}$ is normable.
- When $p = \infty$, $f \in L^{\infty,q}$ makes so little sense that some adopt the convention $f \in L^{\infty,q} \iff f \in L^\infty$ (but we won't).

Example 7.5 – $L^{p,\infty}$ is weak L^p. $[f]_{p,\infty} \leq \|f\|_p$ is Chebyshev's inequality.

Example 7.6 – When $p = q$,

$$\begin{aligned}
[f]_{p,p}^p &= \int_0^\infty \lambda^{p-1} \mu\{|f| > \lambda\}\, d\lambda \\
&= \int_X \int_0^{|f(x)|} \lambda^{p-1}\, d\lambda\, dx \\
&= \frac{1}{p} \|f\|_p^p.
\end{aligned}$$

Example 7.7 – When $q = 1$, let $f = \sum_j c_j \chi_{E_j}$ be a simple function for disjoint E_j and $|c_1| < \cdots |c_N|$. Then

$$\begin{aligned}\mu\{|f| > \lambda\} = \sum_{j=1}^{N} \mu(E_j)\chi_{[0,c_1)} + \sum_{j=2}^{N} \mu(E_j)\chi_{[c_1,c_2)} \\ + \cdots + \mu(E_N)\chi_{[c_{N-1},c_N)}.\end{aligned}$$

Hence,

$$\begin{aligned}[f]_{p,1} &= \int_0^\infty \mu\{|f| > \lambda\}^{\frac{1}{p}}\, d\lambda \\ &= \left(\sum_{j=1}^{N} \mu(E_j)\right)^{\frac{1}{p}} c_1 + \cdots + \mu(E_N)^{\frac{1}{p}}(c_N - c_{N-1}).\end{aligned}$$

Taking $N \to \infty$, there are strong restrictions on when this sum converges, so $L^{p,1}$ is small.

December 11, 2024

Theorem 7.26 (Hunt's version of Marcinkiewicz interpolation theorem)
Let $(X, \mathcal{M}, \mu)$, $(Y, \mathcal{N}, \nu)$ be measure spaces. Let $1 \le p_0, q_0, p_1, q_1 \le \infty$ and $p_0 \neq p_1$, $q_0 \neq q_1$. Let T be a sublinear map from measurable functions on X to measurable functions on Y satisfying a *restricted weak-type inequality*

$$\int |T\chi_E| \chi_F \, d\mu \lesssim |E|^{\frac{1}{p_j}} |F|^{\frac{1}{q_j'}}$$

Then $T \colon L^{p_\theta,r} \to L^{q_\theta,r}$ for fixed $\theta \in (0,1)$ for $1 \le r \le \infty$ is a bounded linear operator.

Remark 7.27. We get a strong type inequality when $p_\theta \le q_\theta$ by setting $r = q_\theta$:

$$\|Tf\|_{q_\theta} \sim [Tf]_{q_\theta,q_\theta} \lesssim [f]_{p_\theta,q_\theta} \le \|f\|_{p_\theta}.$$

Hunt's version is a generalization of the standard real interpolation theorem.

Theorem 7.28 (Marcinkiewicz/real interpolation)
Let $(X, \mathcal{M}, \mu)$, $(Y, \mathcal{N}, \nu)$ be measure spaces. Let $1 \le p_0, q_0, p_1, q_1 \le \infty$, $p_0 \le q_0$, $p_1 \le q_1$, $q_0 \neq q_1$, and

$$p^{-1} = (1-\theta)p_0^{-1} + \theta p_1^{-1}, \qquad q^{-1} = (1-\theta)q_0^{-1} + \theta q_1^{-1},$$

where $\theta \in (0,1)$. If T is a sublinear map from $L^{p_0}(\mu) + L^{p_1}(\mu)$ to the space of measurable functions on Y that is of weak types (p_0, q_0) and (p_1, q_1), then T is of strong type (p, q). More precisely, if $[Tf]_{q_i} \le C \|f\|_{p_i}$ for $i = 0, 1$, then $\|Tf\|_q \le B_p \|f\|_p$ where B_p depends only on p_i, q_i, C_i and p for $i = 0, 1$. $B_p |p - p_j|$ (resp. B_p) remains bounded as $p \to p_j$ if $p_j < \infty$ (resp. $p_j = \infty$).

For more details about how to prove these results, see these notes by Rowan Killip.

A. Topology

A.1. Definitions

Definition A.1
Let $X \neq \varnothing$. $\mathcal{T} \subseteq \mathcal{P}(\mathcal{X})$ is a **topology** if

1. $\varnothing \in \mathcal{T}$,
2. $X \in \mathcal{T}$,
3. $\mathcal{T}$ is closed under arbitrary unions,
4. $\mathcal{T}$ is closed under finite intersections.

The pair $(X, \mathcal{T})$ is called a **topological space**. Sets in $\mathcal{T}$ are called **open sets**, and their complements are **closed sets**.

- The **closure** of a set E, denoted $\overline{E}$ is the smallest closed set containing E.
- An **adherent point** is a point $x \in \overline{E}$. Equivalently, every neighborhood of x intersects E.
- An **accumulation point** is a point $x \in \overline{E \setminus \{x\}}$.
- $\mathcal{T}$ is **coarser** (resp. **weaker**) than $\mathcal{T}'$ if $\mathcal{T} \subseteq \mathcal{T}'$ (resp. $\mathcal{T}' \subseteq \mathcal{T}$).
- If $\mathcal{E} \subseteq \mathcal{P}(X)$, then the **topology generated by** $\mathcal{E}$ is denoted $\mathcal{T}(\mathcal{E})$, and is the coarsest topology containing $\mathcal{E}$. We call $\mathcal{E}$ a **subbase** for $\mathcal{T}(\mathcal{E})$.
- $\mathcal{N}_x \subseteq \mathcal{T}$ is a **neighborhood base** of x if $x \in U$ for all $U \in \mathcal{N}_x$ and for every $V \in \mathcal{T}$ that contains x, there exists $U \in \mathcal{N}_x$ such that $U \subseteq V$.
- $\mathcal{N}$ is a **base** if $\mathcal{N}_x := \{U : U \in \mathcal{N}, x \in U\}$ is a neighborhood base for all $x \in X$.

A.2. Countability axioms

A topological space $(X, \mathcal{T})$ is

- **1st countable** if each point has a countable neighborhood base.
- **2nd countable** if X has a countable base.
- **Separable** if X contains a countable, dense subset.

Proposition A.1
2nd countable $\implies$ separable.

Proof. Let Q be constructed by taking 1 point from every neighborhood. $\square$

Definition A.2
$x_n \to x$ if for all $U \in x$, there exists $N \in \mathbb{N}$ so that $n \geq N$ implies $x_n \in U$. In other words, x_n is "eventually in" U.

Proposition A.2
If x is first countable, then $x \in \overline{E} \iff$ there exists $\{x_n\} \subseteq E$ such that $x_n \to x$.

A.3. Separation axioms

A topological space $(X, \mathcal{T})$ is

- $\mathbf{T_0}$ if, given $x \neq y$, there exists $U \in \mathcal{T}$ so that $x \in U$ and $y \notin U$ *or* $y \in U$ and $x \notin U$.
- $\mathbf{T_1}$ if, there exists a $U \in \mathcal{T}$ so that $x \in U$ and $y \notin U$ (i.e. the T_0 property but symmetric).
- $\mathbf{T_2}$/**Hausdorff** if, given $x \neq y$, there exist disjoint neighborhoods $U \ni x$, $V \ni y$.
- $\mathbf{T_3}$/**regular** if X is T_1 and, given $x \in X$ and $A \subseteq X$ closed not containing x, there exist disjoint open sets U and V so that $x \in U$ and $A \subseteq V$.
- $\mathbf{T_4}$/**normal** if, given closed sets $A, B \subseteq X$, there are disjoint open sets U, V such that $A \subseteq U$ and $B \subseteq V$.

A.4. Weak topology

Definition A.3
The **weak topology** induced by a family of maps $\mathcal{F} = \{f_\alpha \colon X \to Y_\alpha\}_{\alpha \in \mathcal{A}}$ is the weakest topology making all the f_α's continuous. In other words, the subbase of the weak topology is
$$\bigcup_{\alpha \in \mathcal{A}} \left\{ f_\alpha^{-1}(U_\alpha) : U_\alpha \subseteq Y_\alpha, U_\alpha \text{ open} \right\}.$$

Remark A.3. We can also replace U_α with any subbase for the topology on Y_α.

Proposition A.4
If $\mathcal{T}$ is the weak topology induced by $\mathcal{F} = \{f_\alpha\}$, then $x_n \to x \iff f_\alpha(x_n) \to f_\alpha(x)$ for all $\alpha \in \mathcal{A}$. Moreover, $g \colon Z \to X$ is continuous $\iff f_\alpha \circ g$ is continuous for all α.

Example A.1 (Important) – Let $(X_\alpha, \mathcal{T}_\alpha)$ be a collection of topological spaces. Let
$$X \coloneqq \prod_{\alpha \in \mathcal{A}} X_\alpha$$
and $\pi_\alpha \colon X \to X_\alpha$ be the coordinate projections. The **product topology** on X is the weak topology with respect to the maps $\{\pi_\alpha\}$.

Proposition A.5
If $\mathcal{E}_\alpha$ is a base for $\mathcal{T}_\alpha$ for $\alpha \in \mathcal{A}$, then

$$\left\{ \prod_{\alpha \in \mathcal{A}} U_\alpha : U_\alpha = X_\alpha \text{ for all but finitely many } \alpha \right\}$$

is a base for the product topology on X.

Proposition A.6
If $\{X_\alpha\}$ is Hausdorff, so is X.

Proposition A.7
We have convergence in $X \iff$ there is convergence pointwise.

A.5. Compactness

Proposition A.8
If $K \subseteq X$ is a compact subset of a Hausdorff space X, then there exists $U \ni x$ open and $V \supseteq K$ open such that $U \cap V = \varnothing$.

Proof. For each $y \in C$, we know $x \neq y$, so there are neighborhoods $V_y \ni y$ and $U_y \ni x$ such that $U_y \cap V_y = \varnothing$. Hence, $\{V_y\}_{y \in C}$ is a cover of C. By compactness, there is a finite subcover $\{V_{y_i}\}_{i=1}^n$. Then $V := \bigcup_i V_i$, $U := \bigcap_i U_{y_i}$ works. □

Corollary A.9
Compact Hausdorff spaces are normal.

Corollary A.10
In a Hausdorff space, compact sets are closed.

Proposition A.11
The continuous image of a compact set is compact.

Corollary A.12
If $f \colon X \to Y$ is a continuous bijection and X is compact and Y is Hausdorff, then f is a homeomorphism.

A.6. Nets

The motivation of nets is that we want to generalize sequences with a potentially larger indexing set than $\mathbb{N}$.

Definition A.4

A **directed set** $\mathcal{A}$ is a set together with a binary relation "$\lesssim$" such that

1. $\alpha \lesssim \alpha$,
2. $\alpha \lesssim \beta$ and $\beta \lesssim \gamma$ implies $\alpha \lesssim \gamma$,
3. if $\alpha, \beta \in \mathcal{A}$, then there exists $\gamma \in \mathcal{A}$ so that $\alpha \lesssim \gamma$ and $\beta \lesssim \gamma$.

$\lesssim$ is called a **preorder**.

Example A.2 – Let $\mathcal{A} = \mathbb{R}^n$ with the preorder $x \lesssim y$ if $x_i \leq y_i$ for all $1 \leq i \leq n$.

Example A.3 (Important) – Let $(X, \mathcal{T})$ be a topological space and $x \in X$. Let $\mathcal{N}_x$ be a neighborhood base for $\mathcal{T}$ at x. Then define a preorder

$$U \lesssim V \text{ if } U \supseteq V.$$

This generalizes the idea of proofs using $B_{\frac{1}{n}}(x)$.

Definition A.5

A **net** in X is a map

$$\begin{aligned} \langle x_\alpha \rangle_{\alpha \in \mathcal{A}} : \mathcal{A} &\to X \\ \alpha &\mapsto x_\alpha, \end{aligned}$$

with $\mathcal{A}$ a directed set.

Let's redefine things from undergraduate analysis with nets.

Definition A.6

$\langle x_\alpha \rangle$ **converges to** x, denoted $x_\alpha \to x$ if for all $U \ni x$ open, x_α is "eventually in U," i.e., there exists $\alpha_0 \in \mathcal{A}$ such that $x_\alpha \in U$ for all $\alpha \gtrsim \alpha_0$.

x is a **cluster point** of $\langle x_\alpha \rangle$ if for all $U \ni x$ open, x_α "is frequently in U," i.e., for all $\alpha_0 \in \mathcal{A}$, there exists $\alpha \gtrsim \alpha_0$ so that $x_\alpha \in U$.

Proposition A.13

$x \in \overline{E}$ if and only if there exists a net $\langle x_\alpha \rangle$ with $x_\alpha \in E$ such that $x_\alpha \to x$.

Proof. ($\implies$) Let $x \in \overline{E}$ and $\mathcal{N}_x$ be a neighborhood base of $\mathcal{T}$ at x. Then pick $x_U \in$

$E \cap U \neq \varnothing$ for $U \in \mathcal{N}_x$. Then $x_U \to x$.
($\Longleftarrow$) $x_\alpha \to x$ implies for all $U \ni x$ open, $x_\alpha \in U$ for some α. □

Proposition A.14 (Convergence definition of limits)
$f\colon X \to Y$ is continuous $\iff$ for all nets $\langle x_\alpha \rangle$ such that $x_\alpha \to x$, the net $\langle f(x_\alpha) \rangle$ has $f(x_\alpha) \to f(x)$.

Proof. ($\Longrightarrow$) If f is continuous and $x_\alpha \to x$, let $V \ni f(x)$ be a neighborhood. By continuity, $x \in f^{-1}(V)$ implies x_α is eventually in $f^{-1}(V)$, which implies $f(x_\alpha)$ is eventually in V.

($\Longleftarrow$) Assume f is not continuous at x. So there exists a neighborhood $V \ni f(x)$ such that $f^{-1}(V)$ is not a neighborhood of x. Then

$$x \notin f^{-1}(V)^\circ = \overline{f^{-1}(V)^C}^C = (\overline{f^{-1}(V^C)})^C \implies x \in \overline{f^{-1}(V^C)}.$$

Let $\mathcal{N}_x$ be neighborhood base at x viewed as a directed set by "$\supseteq$" as in (A.3). Then for all $U \in \mathcal{N}_x$, there exists $x_U \in U \cap \overline{f^{-1}(V^C)}$. Consider the net $\langle x_U \rangle_{U \in \mathcal{N}_x}$. Then $x_U \to x$, but $f(x_U) \notin V$ for any U, so $f(x_U) \not\to f(x)$. □

Definition A.7
Let $\mathcal{A}$ and $\mathcal{B}$ be directed sets and $\langle \alpha_\beta \rangle : \mathcal{B} \to \mathcal{A}$ be a net. Say $\langle \alpha_\beta \rangle$ **diverges to infinity**, denoted $\alpha_\beta \to \infty$ if for all $\alpha_0 \in \mathcal{A}$, there exists $\beta_0 \in \mathcal{B}$ such that $\beta \gtrsim \beta_0$ implies $\alpha_\beta \gtrsim \alpha_0$.

If $\langle x_\alpha \rangle_{\alpha \in \mathcal{A}} : \mathcal{A} \to X$ is a net, a **subnet** of it is the composition

$$\left\langle x_{\alpha_\beta} \right\rangle_{\beta \in \mathcal{B}} : \mathcal{B} \to X$$

with $\langle \alpha_\beta \rangle : \mathcal{B} \to X$ a net such that $\alpha_\beta \to \infty$.

Example A.4 – Let $a_n = \frac{1}{n}$ and $n_k = \begin{cases} 3^k & k \text{ even,} \\ 2^k & k \text{ odd.} \end{cases}$. Then $\langle a_{n_k} \rangle$ is a subnet of $\langle a_n \rangle$, but not a subsequence, because $n_k \to \infty$, but not monotonically.

A.7. Compactness

Proposition A.15
The following are equivalent:

(a) X is compact.

(b) Every net X has a cluster point.

(c) Every net in X has a convergent subnet.

Proof. ((b) $\iff$ (c)) has already been done.

((b) $\implies$ (a)) Let $\{F_\alpha\}_{\alpha\in\mathcal{A}}$ be closed sets with the finite intersection property. Let $\mathcal{B}$ be the collection of finite subsets of $\mathcal{A}$. This forms a directed set by $B \lesssim B'$ if $B \subseteq B'$. To form a net, choose $x_B \in \bigcap_{\alpha\in B} F_\alpha$, $B \in \mathcal{B}$. By (b), we have a cluster point x of $\langle x_B\rangle$. For every neighborhood $U \ni x$ and $B \in \mathcal{B}$, there exists $B' \supseteq B$ such that $x_{B'} \in U$. So

$$x_{B'} \in \bigcap_{\alpha\in B'} F_\alpha \subseteq \bigcap_{\alpha\in B} F_\alpha.$$

Since every neighborhood of x intersects $\bigcap_{\alpha\in B} F_\alpha$,

$$x \in \overline{\bigcap_{\alpha\in\mathcal{B}} F_\alpha} = \bigcap_{\alpha\in\mathcal{B}} F_\alpha.$$

In particular, $x \in F_\alpha$ for all α. So $\bigcap_{\alpha\in\mathcal{A}} F_\alpha$ is nonempty, and X is compact.

((a) $\implies$ (b)) Let $\langle x_\alpha\rangle_{\alpha\in\mathcal{A}}$ be a net. Let $E_\alpha := \{x_\beta : \beta \gtrsim \alpha\}$. $\alpha_1, \dots, \alpha_N \in \mathcal{A}$ implies (by induction) that $\beta \gtrsim \alpha_n$ for all n. Consider the set of closed sets $\{F_\alpha := \overline{E_\alpha}\}$, which are closed sets with the finite intersection property. Therefore, there exists $x \in \bigcap_{\alpha\in\mathcal{A}} F_\alpha$. To show x is a cluster point, let $U \ni x$ be a neighborhood and $\alpha_0 \in \mathcal{A}$. Then

$$x \in F_{\alpha_0} = \overline{E_{\alpha_0}},$$

so $U \cap E_{\alpha_0} \neq \varnothing$, i.e. there exists $\beta \gtrsim \alpha_0$ such that $x_\beta \in U$, therefore $\langle x_\alpha\rangle$ visits U frequently. Since U arbitrary, x is a cluster point. □

We will use definition (b) for the proof of Tychonoff's theorem.

Theorem A.16 (Tychonoff's theorem)

If $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ are a collection of compact spaces, then $X := \prod_{\alpha\in\mathcal{A}} X_\alpha$ is compact (in the product topology).

Proof. Let $\langle x_i\rangle_{i\in\mathcal{I}}$ be a net. Define

$$\mathcal{P} := \left\{ (\mathcal{B}, x_\mathcal{B}) : \mathcal{B} \subseteq \mathcal{A}, x_\mathcal{B} \in \prod_{\beta\in\mathcal{B}} X_\beta, x_\mathcal{B} \text{ a cluster point of } \langle \pi_\mathcal{B}(x_i)\rangle_{i\in\mathcal{I}} \right\}.$$

Define a partial order on $\mathcal{P}$ by $(\mathcal{B}, x_\mathcal{B}) \lesssim (\mathcal{B}', y_{\mathcal{B}'})$ if $\mathcal{B} \subseteq \mathcal{B}'$ and $\pi_\mathcal{B}(y_{\mathcal{B}'}) = x_\mathcal{B}$ (say $y_{\mathcal{B}'}$ **extends** $x_\mathcal{B}$).

Our goal now is to invoke Zorn's lemma. Let $\mathcal{L} := \{(\mathcal{B}_\ell, x_{\mathcal{B}_\ell})\}$ be a chain (i.e. a totally ordered set/**toset**) in $\mathcal{P}$. Let $\mathcal{B}^* := \bigcup_\ell \mathcal{B}_\ell$. Then there exists a unique $x_{\mathcal{B}^*} \in \prod_{\alpha\in\mathcal{B}^*} X_\alpha$ extending all $x_{\mathcal{B}_\ell}$. We now need to show $x_{\mathcal{B}^*}$ is a cluster point of $\langle \pi_{\mathcal{B}^*}(x_i)\rangle_{i\in\mathcal{I}}$. Let $U_{\mathcal{B}^*} \ni x_{\mathcal{B}^*}$ be a neighborhood. WLOG,

$$U_{\mathcal{B}^*} = \prod_{\substack{\alpha\in\mathcal{F}\subseteq\mathcal{B}^* \\ \mathcal{F}\text{ finite}}} U_\alpha \times \prod_{\alpha\in\mathcal{B}^*\setminus\mathcal{F}} X_\alpha.$$

By construction, there exists $\mathcal{B}_\ell$ such that $\mathcal{F} \subseteq \mathcal{B}_\ell$. Since $x_{\mathcal{B}_\ell}$ is a cluster point of $\langle \pi_{\mathcal{B}_\ell}(x_i)\rangle$, $\pi_{\mathcal{B}_\ell}(x_i)$ is frequently in $U_{\mathcal{B}^*}$. Therefore, this chain has an upper bound $(\mathcal{B}^*, x_{\mathcal{B}^*})$. So by Zorn's lemma, $\mathcal{P}$ has a maximal element $(\overline{\mathcal{B}}, x_{\overline{\mathcal{B}}})$.

We claim $\overline{\mathcal{B}} = \mathcal{A}$. Indeed, if not, then there exists $\gamma \in \mathcal{A} \setminus \overline{\mathcal{B}}$. We claim that there exists $x_\gamma \in X_\gamma$ such that $(\overline{\mathcal{B}} \cup \gamma, x_{\overline{\mathcal{B}} \cup \gamma}) \in \mathcal{P}$. Indeed, $x_{\overline{\mathcal{B}}}$ is a cluster point of $\langle \pi_{\overline{\mathcal{B}}}(x_i) \rangle$ which means there is a subnet

$$\pi_{\overline{\mathcal{B}}}(x_{i_j}) \to x_{\overline{\mathcal{B}}}.$$

X_γ compact implies there exists a further subnet $\left\langle x_{i_{j_k}} \right\rangle$ such that $\pi_\gamma(x_{i_{j_k}})$ converges to some $x_\gamma \in X_\gamma$. Let $x_{\overline{\mathcal{B}} \cup \gamma}$ be the unique element extending $x_{\overline{\mathcal{B}}}$ and x_γ. But then the net $\left\langle \pi_{\overline{\mathcal{B}} \cup \gamma} \left(x_{i_{j_k}} \right) \right\rangle$ converges to $x_{\overline{\mathcal{B}} \cup \gamma}$, hence it is a cluster point of $\langle \pi_{\overline{\mathcal{B}}}(x_i) \rangle$, contradicting maximality. $\square$

References

[Fol99] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1999.