Measure Theory– Lecture Notes

Han-Miru Kim

August 30, 2021

About

Lecture notes taken from the Measure and Integration lecture given by Dr. Francesca Da Lio during Spring Semester 2021 at ETH Zürich.
With a focus on the theorems and their proofs, these notes have fewer examples than given in the lecture, but the proofs will be more explicit.
The exam will be a 20 minute **oral exam**. It will consist of two or three questions where we have to prove some results.
The lecture slides and a script is provided at the professors home page `https://people.math.ethz.ch/~fdalio/MASSundINTEGRALFS21`.

1 Measure spaces

If we naively try to define a notion of measure that has some intuitive properties, we can run into some problems that give paradoxical results. The **Riemann Integral** we saw in Analysis I/II also had some drawbacks of not being general enough. We can use measure theory to define a better definition of the integral.

1.1 Algebras and σ-Algebras

From now on, let X denote a non-empty set.

Definition 1.1.1. For a sequence of subsets $(A_n)_{n=1}^\infty$ in $\mathcal{P}(X)$. We define

$$\limsup_{n\to\infty} A_n := \bigcap_{n=1}^\infty \bigcup_{m=n}^\infty A_m$$
$$\liminf_{n\to\infty} A_n := \bigcup_{n=1}^\infty \bigcap_{m=n}^\infty A_m$$

And if they are equal, we say that the sequence $(A_n)_{n=1}^\infty$ converges to its limit $\lim_{n\to\infty} A_n$.

Informally, the $\limsup$ consists of elements of X that occur in infinitely many A_n, whereas the $\liminf$ consists of elements that occur for all but finitely many A_n.

Remark 1.1.2.

(a) $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$

(b) If $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$, then

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

(c) If $A_n \supseteq A_{n+1}$ for all $n \in \mathbb{N}$, then

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

The similarity in names with the lim sup and lim inf from Analysis can be seen using the characteristic function

$$\mathbb{1}_A : X \to \{0,1\}$$
$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

It holds that

$$\limsup_{n\to\infty} A_n = A \iff \limsup_{n\to\infty} \mathbb{1}_{A_n} = \mathbb{1}_A$$
$$\liminf_{n\to\infty} A_n = A \iff \liminf_{n\to\infty} \mathbb{1}_{A_n} = \mathbb{1}_A$$

where the lim sup and lim inf on the left are as in Definition 1.1.1 and the ones on the right are the ones from Analysis.

Definition 1.1.3 (Algebras of sets)**.** A collection of subsets $\mathcal{A} \subseteq \mathcal{P}(X)$ is called an **algebra in** X if

(a) $X \in \mathcal{A}$

(b) $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$

(c) $A \in \mathcal{A} \implies A^c \in \mathcal{A}$

An algebra $\mathcal{E}$ is called a **σ-algebra**, if for any sequence $(A_n)_{n=1}^{\infty}$ in $\mathcal{E}$ we have $\bigcup_{n=1}^{\infty} A_n \in \mathcal{E}$

Note that using the De Morgan's identity

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c$$

we can see that algebras (σ-algebras) are stable under finite (infinite) intersections aswell.

Definition 1.1.4. For a collection of sets $\mathcal{K} \subseteq \mathcal{P}(X)$, the intersection of all σ-algebras containing K forms again a σ-algebra.
We call this the σ-algebra **generated by** K and it its the smallest σ-algebra that contains K.
The algebra generated by the open sets of a topology is called the **Borel σ-Algebra** of X, denoted $\mathcal{B}(X)$.

1.2 Measures

Definition 1.2.1. Let $\mathcal{A}$ be an Algebra on X and $\mu : \mathcal{A} \to [0,\infty]$. We say that μ is

- **additive**, if for any *finite* family of disjoint sets $A_1, \dots, A_n \in \mathcal{A}$

$$\mu\left(\bigsqcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k)$$

- **σ-additive**, if for any *countable* family of disjoint sets $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that $\bigsqcup_{k=1}^{\infty} A_k \in \mathcal{A}$

$$\mu\left(\bigsqcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$$

- A **pre-measure**, if it is σ-additive and satisfies $\mu(\emptyset) = 0$.

Remark 1.2.2. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathcal{A}$ such that their union is again in $\mathcal{A}$.

(a) If μ is additive, then it is monotone with respect to incusion, i.e. $A \subseteq B \implies \mu(A) \leq \mu(B)$.

(b) If μ is additive and the sets A_k are mutually disjoint, then

$$\mu\left(\bigsqcup_{k=1}^{\infty} A_k\right) \geq \sum_{k=1}^{\infty} \mu(A_k)$$

(c) If μ is σ-additive, then it is also **σ-subadditive**, which means that for any sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ with $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k)$$

To see this, we can define the mutually disjoint sets

$$B_1 = A_1, \quad B_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k \in \mathcal{A}$$

Since $\bigsqcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} A_k$ and $\mu(B_k) \leq \mu(A_k)$ we have

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \mu\left(\bigsqcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} \mu(B_k) \leq \sum_{k=1}^{\infty} \mu(A_k)$$

It follows immediately from (b) and (c) that

$$\boxed{\mu \text{ is additive and } \sigma\text{-subadditive} \iff \mu \text{ is } \sigma\text{-additive}}$$

Example 1.2.3. Not all additive functions are σ-additive. For $X = \mathbb{N}$ and

$$\mathcal{A} = \{A \in \mathcal{P}(X) \,|\, A \text{ is finite or } A^c \text{ is finite}\}$$

the function $\nu : \mathcal{A} \to [0, \infty]$ with $\nu(\emptyset) = 0$ and

$$\nu(A) = \begin{cases} \sum\limits_{n \in A} \frac{1}{2^n} & \text{if } A \text{ is finite} \\ \infty & \text{if } A^c \text{ is finite} \end{cases}$$

is additive but not σ-additive because we can take the sequence

$$A_1 = \{1\}, A_2 = \{2\}, A_3 = \{3\}, \ldots, A_n = \{n\}, \ldots$$

which is a sequence of mutually disjoint sets satsfying

$$\nu(A_1) = \frac{1}{2}, \nu(A_2) = \frac{1}{4}, \ldots, \nu(A_n) = \frac{1}{2^n}$$
$$\implies \nu\left(\bigsqcup_{n=1}^{\infty} A_n\right) = \nu(\mathbb{N}) = \infty \not\leq \sum_{n=1}^{\infty} \nu(A_n) = 1$$

Definition 1.2.4. A σ-additive function $\mu : \mathcal{A} \to [0, \infty]$ is called

- **finite**, if $\mu(X) < \infty$
- **σ-finite**, if there exists a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that

$$\bigcup_{n=1}^{\infty} A_n = X \quad \text{and} \quad \mu(A_n) < \infty \quad \forall n \in \mathbb{N}$$

Clearly, μ finite $\implies$ μ σ-finite.
While pre-measures are only defined on algebras $\mathcal{A} \subseteq \mathcal{P}(X)$, we would like to extend the domain of such functions to $\mathcal{P}(X)$ without losing too many of its nice properties. In particular, we want to keep monotonicity and σ-subadditivity:

Definition 1.2.5. A function $\mu : \mathcal{P}(X) \to [0, \infty]$ is called a **measure**[1] on X, if

(a) $\mu(\emptyset) = 0$

(b) μ is σ-subadditive: If $A \subseteq \bigcup_{k=1}^{\infty} A_k$, then $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$

Note that subadditivity implies monotonicity with respect to inclusion, i.e. $A \subseteq B \implies \mu(A) \leq \mu(B)$.

Definition 1.2.6. Let μ be a measure on X and $A \subseteq X$. We can *restrict* μ to A (written $\mu \llcorner A$) defined by

$$(\mu \llcorner A)(B) := \mu(A \cap B) \quad \forall B \subseteq X$$

Definition 1.2.7 (Carathéodory criterion)**.** A subset $A \subseteq X$ is called **μ-measurable** if

$$\mu(B) = \mu(B \cap A) + \mu(B \setminus A), \quad \forall B \subseteq X$$

Remark 1.2.8. (a) By subadditivity of the measure, the definition is equivalent to

$$\boxed{\mu(B) \geq \mu(B \cap A) + \mu(B \setminus A), \quad \forall B \subseteq X}$$

[1]sometimes also called outer measure

(b) If $\mu(A) = 0$, then A is μ-measurable:

$$\mu(B \cap A) + \mu(B \setminus A) \le \mu(A) + \mu(B) = \mu(B)$$

Theorem 1.2.9. Let $\mu : \mathcal{P}(X) \to [0, \infty]$ be a measure. Then the collection of measurable sets

$$\Sigma = \{A \subseteq X | A \text{ is } \mu\text{-measurable}\}$$

forms a σ-algebra.

Proof.

- $X \in \Sigma$: Let $B \subseteq X$. It's trivial to see that

$$\mu(B \cap X) + \mu(B \setminus X) = \mu(B) + \mu(\emptyset) = \mu(B)$$

- $A \in \Sigma \implies A^c \in \Sigma$: With the equalities

$$B \cap A^c = B \setminus A, \quad \text{and} \quad B \setminus A^c = B \cap A$$

  we get

$$\mu(B \cap A^c) + \mu(B \setminus A^c) = \mu(B \setminus A) + \mu(B \cap A) \overset{A \in \Sigma}{=} \mu(B)$$

- $A_1, A_2 \in \Sigma \implies A_1 \cup A_2 \in \Sigma$:
  Let $B \subseteq X$. From the previous remark, it is sufficient to just show the inequality

$$\mu(B) \ge \mu(B \cap (A_1 \cup A_2)) + \mu(B \setminus (A_1 \cup A_2))$$

  Using μ-measurability for A_1 on the test set $B \setminus A_2$, we see

$$\begin{aligned} \mu(B \setminus A_2) &= \mu\big((B \setminus A_2) \cap A_1\big) + \mu\big((B \setminus A_2) \setminus A_1\big) \\ &= \mu\big((B \setminus A_2) \cap A_1\big) + \mu(B \setminus (A_2 \cup A_1)) \end{aligned}$$

  so with the decomposition

$$(B \cap A_2) \cup \big((B \setminus A_2) \cap A_1\big) = B \cap (A_1 \cup A_2)$$

  and subadditivity of the measure, we get

$$\begin{aligned} \mu(B) &= \mu(B \cap A_2) + \mu(B \setminus A_2) \\ &= \mu(B \cap A_2) + \mu\big((B \setminus A_2) \cap A_1\big) + \mu(B \setminus (A_2 \cup A_1)) \\ &\ge \mu(B \cap (A_2 \cup A_1)) + \mu(B \setminus (A_2 \cup A_1)) \end{aligned}$$

- $(A_n)_{n \in \mathbb{N}} \subseteq \Sigma \implies A = \bigcup_{n=1}^{\infty} A_n \in \Sigma$:
  We can assume without loss of generality that the sets are mutually disjoint. Otherwise, consider the sequence $\left(\tilde{A}_n\right)_{n \in \mathbb{N}} \subseteq \Sigma$ given by

$$\tilde{A}_1 := A_1, \tilde{A}_n := A_n \setminus \bigcup_{k=1}^{n-1} A_k \quad \text{which satisfy} \quad \bigsqcup_{k=1}^{\infty} \tilde{A}_k = \bigcup_{k=1}^{\infty} A_k.$$

We can use μ-measureability of A_m with the test set $B \cap \bigsqcup_{k=1}^m A_k$ to find that by induction on m

$$\begin{aligned}
\mu\left(B \cap \bigsqcup_{k=1}^m A_k\right) &= \mu\left(\left(B \cap \bigsqcup_{k=1}^m A_k\right) \cap A_m\right) + \mu\left(\left(B \cap \bigsqcup_{k=1}^m A_k\right) \setminus A_m\right) \\
&= \mu(B \cap A_m) + \mu\left(B \cap \bigsqcup_{k=1}^{m-1} A_k\right) \\
&= \sum_{k=1}^m \mu(B \cap A_k)
\end{aligned}$$

and using monotonicity of μ on the inclusion $\bigsqcup_{k=1}^m A_k \subseteq A$ it follows that

$$\begin{aligned}
\mu(B) &= \mu\left(B \cap \bigsqcup_{k=1}^m A_k\right) + \mu\left(B \setminus \bigsqcup_{k=1}^m A_k\right) \\
&\geq \sum_{k=1}^m \mu(B \cap A_k) + \mu(B \setminus A)
\end{aligned}$$

for all $m \in \mathbb{N}$. Taking the limit $m \to \infty$, we get

$$\begin{aligned}
\mu(B) &\geq \sum_{k=1}^\infty \mu(B \cap A_k) + \mu(B \setminus A) \\
&\geq \mu(B \cap A) + \mu(B \setminus A)
\end{aligned}$$

which shows μ-measurability of A.

□

Definition 1.2.10. A **measure space** is a tuple (X, Σ, μ) consisting of measure μ on a set X and the σ-algebra of μ-measurable sets Σ.

Example 1.2.11. The following are measure spaces:

- For every $x \in X, A \subseteq X$, define the **Dirac measure at** x

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

  Every A is δ_x-measurable.

- For every $A \in \mathcal{P}$, the **counting measure** is a measure, where every subset is μ-measurable:

$$\mu(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{otherwise} \end{cases}$$

  Every A is μ-measurable.

- The **indiscrete measure** given by

$$\mu(A) = \begin{cases} 1 & A \neq \emptyset \\ 0 & A = \emptyset \end{cases}$$

  only has $\emptyset, X$ as μ-measurable sets.

The Carathéodory criterion of μ-measurable sets and the σ-subadditivity of the measure give us some nice properties back.

Theorem 1.2.12. Let (X, Σ, μ) be a measure space and $(A_n)_{n\in\mathbb{N}} \subseteq \Sigma$. Then the following are true

(a) μ is σ-additive.

(b) Continuity from below:

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq A_{n+1} \subseteq \ldots \implies \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

(c) Continuity from above:

$$\mu(A_1) < \infty, \quad A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq A_{n+1} \supseteq \ldots \implies \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

Proof. (a) Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of mutually disjoint sets. In the proof of the previous theorem, we already saw

$$\mu\left(B \cap \bigsqcup_{k=1}^{m} A_k\right) = \sum_{k=1}^{m} \mu(B \cap A_k)$$

so in particular, for $B = X$, we see

$$\mu\left(\bigsqcup_{k=1}^{m} A_k\right) = \sum_{k=1}^{m} \mu(A_k)$$

By monotonicity of μ, we have

$$\mu\left(\bigsqcup_{k=1}^{\infty} A_k\right) \geq \lim_{m\to\infty} \mu\left(\bigsqcup_{k=1}^{m} A_k\right) = \lim_{m\to\infty} \sum_{k=1}^{m} \mu(A_k) = \sum_{k=1}^{\infty} \mu(A_k)$$

The other inequality (and thus equality) follow from σ-subadditivity of the measure.

(b) Let $(A_n)_{n\in\mathbb{N}}$ be an increasing sequence. Define the pairwise disjoint family

$$\tilde{A}_1 := A_1, \quad \tilde{A}_k := A_k \setminus A_{k-1} \implies \mu(\tilde{A}_k) = \mu(A_k) - \mu(A_{k-1}), \quad \bigsqcup_{k=1}^{\infty} \tilde{A}_k = \bigcup_{k=1}^{\infty} A_k,$$

from σ-additivity, summation into a telescoping sum

$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{\infty} A_k\right) &= \mu\left(\bigsqcup_{k=1}^{\infty} \tilde{A}_k\right) = \sum_{k=1}^{\infty} \mu(\tilde{A}_k) \\
&= \mu(\tilde{A}_1) + \lim_{m\to\infty} \sum_{k=2} \mu(A_k) - \mu(A_{k-1}) \\
&= \lim_{m\to\infty} \mu(A_m)
\end{aligned}$$

(c) Let $(A_n)_{n\in\mathbb{N}}$ be a decreasing sequence. Consider instead the increasing sequence $\tilde{A}_1 \subseteq \tilde{A}_2 \subseteq \dots$ given by

$$\tilde{A}_1 := \emptyset, \quad \tilde{A}_k := A_1 \setminus A_k \implies \mu(A_1) = \mu(A_k) + \mu(\tilde{A}_k), \quad \bigcup_{k=1}^{\infty} \tilde{A}_k = A_1 \setminus \bigcap_{k=1}^{\infty} A_k$$

by (b), we find

$$\begin{aligned}\mu(A_1) - \lim_{k\to\infty} \mu(A_k) &= \lim_{k\to\infty} \mu(\tilde{A}_k) \\ &\overset{(b)}{=} \mu\left(\bigcup_{k=1}^{\infty} \tilde{A}_k\right) = \mu\left(A_1 \setminus \bigcap_{k=1}^{\infty} A_k\right) \\ &= \mu(A_1) - \mu\left(\bigcap_{k=1}^{\infty} A_k\right)\end{aligned}$$

□

The condition $\mu(A_1)$ in (c) is necessary. Consider the example $X = \mathbb{N}$ with the counting-measure and the sequence $A_n := \{m \in \mathbb{N} | m \geq n\}$. The intersections converge to the emtpy set, but the $\mu(A_k)$ is always ∞.

1.3 Construction of Measures

Let X be non-empty set.

Definition 1.3.1. A collection of subsets $\mathcal{K} \subseteq \mathcal{P}(X)$ is called a **covering** of X if

$$\emptyset \in \mathcal{K} \quad \text{and} \quad \exists (K_j)_{j\in\mathbb{N}} \subseteq \mathcal{K}: \quad X = \bigcup_{j=1}^{\infty} K_j$$

Example 1.3.2. The collection of higher-dimensional open intervals

$$\{\prod_{k=1}^{n} (a_k, b_k) | a_k \leq b_k \in \mathbb{R}\}$$

are a covering of $\mathbb{R}^n$.
It is easy to see that every Algebra $\mathcal{A}$ of X is a covering since $\emptyset, X \in \mathcal{A}$.

Theorem 1.3.3. Let $\mathcal{K}$ be a covering of X and $\lambda : \mathcal{K} \to [0, \infty]$ and any function with $\lambda(\emptyset) = 0$.
Then this induces a measure μ on X given by

$$\mu(A) = \inf\left\{\sum_{j=1}^{\infty} \lambda(K_j) | K_j \in \mathcal{K}, A \subseteq \bigcup_{j=1}^{\infty} K_j\right\}$$

Proof. Let $A \subseteq \bigcup_{k=1}^{\infty} A_k$. We show σ-subadditivity of μ, i.e $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$.
If the right-hand side is infinite, then the inequality is trivial, so assume it is finite.
By definition of μ, for all $k \in \mathbb{N}$ and $\varepsilon > 0$ there exists a sequence $(K_{j,k})_{j\in\mathbb{N}}$ in $\mathcal{K}$ such that

$$A_k \subseteq \bigcup_{j=1}^{\infty} K_{j,k} \quad \text{and} \quad \sum_{j=1}^{\infty} \lambda(K_{j,k}) \leq \mu(A_k) + \frac{\varepsilon}{2^k}$$

Taking the union over all sequences for each k, we get

$$A \subseteq \bigcup_{j,k=1}^{\infty} K_{j,k} \quad \text{and} \quad \mu(A) \leq \sum_{j,k} \lambda(K_{j,k}) \leq \varepsilon + \sum_{k=1}^{\infty} \mu(A_k)$$

Since $\varepsilon > 0$ was arbitrary, subadditivity follows. □

Example 1.3.4. Set $\mathcal{K} = \{\emptyset, X\}$ and define $\lambda(\emptyset) = 0$, $\lambda(X) = 1$.
The induced measure is defined by $\mu(A) = 0$ if $A = \emptyset$ and $\mu(A) = 1$ if $A \neq \emptyset$.

The function λ in the previous theorem only had minimal restrictions ($\mathcal{K}$ had to be a covering and $\lambda : \mathcal{K} \to [0, \infty]$ with $\lambda(\emptyset) = \emptyset$).
It turns out that if λ and $\mathcal{K}$ are *nice enough*, then the induced measure is a σ-additive extension of λ.
Nice-enough here means that $\mathcal{K}$ is an algebra and λ is a pre-measure.
Recall that given an algebra $\mathcal{A} \subseteq \mathcal{P}(X)$, a function $\lambda : \mathcal{A} \to [0, \infty]$ is called a **pre-measure** if it is σ-additive and satisifes $\lambda(\emptyset) = 0$.
Given a pre-measure λ on $\mathcal{A}$, we can obtain a measure μ on $\mathcal{P}(X)$ that coincides with λ on $\mathcal{A}$, i.e. μ extends λ.

Theorem 1.3.5 (Carathéodory-Hahn extension)**.** Let $\lambda : \mathcal{A} \to [0, \infty]$ be a pre-measure on X. Then for

$$\mu : \mathcal{P}(X) \to [0, \infty], \quad \mu(A) := \inf \left\{ \sum_{k=1}^{\infty} \lambda(A_k) | A \subseteq \bigcup_{k=1}^{\infty} A_k, A_k \in \mathcal{A} \right\}$$

it holds that

(a) $\mu : \mathcal{P} \to [0, \infty]$ is a measure.

(b) $\mu(A) = \lambda(A), \forall A \in \mathcal{A}$

(c) All $A \in \mathcal{A}$ are μ-measurable, i.e. satisfy $\mu(B) \geq \mu(B \cap A) + \mu(B \setminus A), \forall B \subseteq X$.

Proof. (a) Because algebras are also coverings, we can just use the previous theorem.

(b) Let $A \in \mathcal{A}$. Since A itself contains A, the term $\lambda(A)$ is present in the right hand side, so $\mu(A) \leq \lambda(A)$.
Now assume there is some other collection $\bigcup_{k=1}^{\infty} A_k$ that contains A with $A_k \in \mathcal{A}$. By inductively defining the mutually disjoint sequence

$$B_1 = A_1, \quad B_k := A_k \setminus \bigcup_{i=1}^{k-1} B_i$$

we see $\sum_{k=1}^{\infty} \lambda(B_k) \leq \sum_{k=1}^{\infty} \lambda(A_k)$, so since we're taking the infimum, we can assume that WLOG the A_k are mutually disjoint.
Setting $\tilde{A}_k := A_k \cap A \in \mathcal{A}$, we see that they are also mutually disjoint and their union contains A.
By σ-additivity of the pre-measure λ, we get

$$\lambda(A) = \sum_{k=1}^{\infty} \lambda(\tilde{A}_k) \leq \sum_{k=1}^{\infty} \lambda(A_k)$$

since the collection $(A_k)_{k \in \mathbb{N}}$ was arbitrary, the inequality $\lambda(A) \leq \mu(A)$ follows.

(c) Let $A \in \mathcal{A}$ and $B \subseteq X$ be any test set. By definition of μ, for every $\varepsilon > 0$ we can chose a collection $(B_k)_{k\in\mathbb{N}} \subseteq \mathcal{A}$ that contains B and

$$\sum_{k=1}^{\infty} \lambda(B_k) \leq \mu(B) + \varepsilon$$

By (σ)-additivity of λ and $A, B_k \in \mathcal{A}$ we have

$$\lambda(B_k) = \lambda(B_k \cap A) + \lambda(B_k \setminus A) \quad \forall k$$

so since the $(B_k \cap A)_{k\in\mathbb{N}}$ and $(B_k \setminus A)_{k\in\mathbb{N}}$ contain $B \cap A$ and $B \setminus A$ each, we get

$$\begin{aligned}\mu(B \cap A) + \mu(B \setminus A) &\leq \sum_{k=1}^{\infty} \lambda(B_k \cap A) + \sum_{k=1}^{\infty} \lambda(B_k \setminus A) \\ &= \sum_{k=1}^{\infty} \lambda(B_k) \leq \mu(B) + \varepsilon\end{aligned}$$

and in the limit $\varepsilon \to 0$ the inequality follows.

□

Not only does such an extension exist, we can show that under certain assumptions it is unique:

Definition 1.3.6. A pre-measure λ is called **σ-finite** if there exists a covering $X = \bigcup_{k=1}^{\infty} S_k$, $S_k \in \mathcal{A}$ such that $\lambda(S_k) < \infty, \forall k$.

Theorem 1.3.7 (Uniqueness of Carathéodory-Hahn extension)**.** Let $\lambda : \mathcal{A} \to [0, \infty]$ be a σ-finite pre-measure on X and μ the Carathéodory-Hahn extension of λ and let Σ be the σ-algera of μ-measurable sets.
If $\tilde{\mu} : \mathcal{P}(X) \to [0, \infty]$ is another measure with $\tilde{\mu}|_{\mathcal{A}} = \lambda$, then $\tilde{\mu}|_{\Sigma} = \mu|_{\Sigma}$

Proof. Let $\tilde{\mu} : \mathcal{P}(X) \to [0, \infty]$ be a measure extending λ. We show

(i) $\forall A \in \mathcal{P}(X)$: $\tilde{\mu}(A) \leq \mu(A)$.

(ii) $\forall A \in \Sigma$: $\tilde{\mu}(A) \geq \mu(A)$.

For the first claim, let $A \subseteq \bigcup_{k=1}^{\infty} A_k$ with $A_k \in \mathcal{A}$. By σ-subadditivity of $\tilde{\mu}$ it follows that

$$\tilde{\mu}(A) \leq \sum_{k=1}^{\infty} \tilde{\mu}(A_k) = \sum_{k=1}^{\infty} \lambda(A_k)$$

So by taking the infimum over all such coverings $(A_k)_{k\in\mathbb{N}}$ as in the definition of μ, the inequality still holds: $\tilde{\mu}(A) \leq \mu(A)$. Note that we didn't have to use σ-finiteness of λ for this inquality.
For the second claim let $A \in \Sigma$ be μ-measurable. We then consider the simple case where there exists an $S \in \mathcal{A}$ such that

$$A \subseteq S \quad \text{and} \quad \lambda(S) < \infty$$

Then, using the first claim on $S \setminus A$ and monotonicity of μ, it follows that

$$\tilde{\mu}(S \setminus A) \leq \mu(S \setminus A) \leq \mu(S) = \lambda(S)$$

Since $A \in \mathcal{A}$ is μ-measurable and $A = S \cap A$ we get with $\mu|_{\mathcal{A}} = \lambda = \tilde{\mu}|_{\mathcal{A}}$ that

$$\begin{aligned}\tilde{\mu}(A) + \tilde{\mu}(S \setminus A) &\leq \mu(S \cap A) + \mu(S \setminus A) = \mu(S)\\ &= \lambda(S) = \tilde{\mu}(S)\\ &\leq \tilde{\mu}(A) + \tilde{\mu}(S \setminus A)\end{aligned}$$

where we used sub-additivity of $\tilde{\mu}$ in the last step. It follow sthat $\tilde{\mu}(A) = \mu(A) \leq \tilde{\mu}(A)$.
In the more general case, we can use σ-finiteness to get a covering

$$X = \bigcup_{k=1}^{\infty} S_k, S_k \in \mathcal{A}, \lambda(S_k) < \infty$$

As remarked in the proof of the last theorem, we can assume without loss of generality that the S_k are mutually disjoint.
Defining $A_k = A \cap S_k$ we get $A = \bigcup_{k=1}^{\infty} A_k$. Because $\mathcal{A}$ is closed under finite unions and $\tilde{\mu}|_{\mathcal{A}} = \mu|_{\mathcal{A}}$, we have that for all $m \in \mathbb{N}$:

$$\bigcup_{k=1}^{m} A_k \in \mathcal{A} \implies \tilde{\mu}\left(\bigcup_{k=1}^{m} A_k\right) = \mu\left(\bigcup_{k=1}^{m} A_k\right)$$

and by using monotonicity on the inclusion $A \supseteq \bigcup_{k=1}^{m} A_k$ and taking the limit $m \to \infty$, we get

$$\tilde{\mu}(A) \geq \lim_{m\to\infty} \tilde{\mu}\left(\bigcup_{k=1}^{m} A_k\right) = \lim_{m\to\infty} \mu\left(\bigcup_{k=1}^{m} A_k\right) = \mu(A)$$

$\square$

If we denote $\tilde{\Sigma}$ to be the σ-agebra of $\tilde{\mu}$-measurable sets, the theorem doesn't tell us if $\tilde{\Sigma} = \Sigma$. Moreover, it doesn't tell us anything about the behaviour of $\tilde{\mu}$ outside of Σ.

Example 1.3.8. Let $X = [0, 1], \mathcal{A} = \{\emptyset, X\}$ and set $\lambda(\emptyset) = 0, \lambda(X) = 1$.
The Carathéodory extension of λ has $\mu(A)$ to be 0 or 1, depending on if A is empty or not. The μ-measurable sets are $\Sigma = \{\emptyset, X\}$.
However, as we will see in the next section, the Lebesuge measure $Ł^1$ is also an extension of λ with $Ł^1|_{\Sigma} = \mu|_{\Sigma}$, but they differ when measuring the interval $[0, \frac{1}{2}]$.

1.4 Lebesgue Measure

The Lebesgue measure is the Carethéodory-Hahn extension of the pre-measure that corresponds to the "physical" notion of what a volume of simple objects such as n-dimensional hypercubes like $[0, 1]^n$ is.
We want to give a precise definition of what these "simple objects" are and define the pre-measure.

Definition 1.4.1. For $a = (a_1, \dots, a_d), b = (b_1, \dots, b_d) \in \mathbb{R}^d$ we define the d-dimensional **interval**

$$(a, b) := \begin{cases} \prod_{i=1}^{d}(a_i, b_i) & \text{if } a_i < b_i \quad \forall i \\ \emptyset & \text{otherwise} \end{cases} \subseteq \mathbb{R}^d$$

in an analogous way, we define the closed and half-open boxes $[a, b], [a, b)$ or $(a, b]$. Like on the real line, we also allow the open ends to be $\pm\infty$.

To each d-dimensional interval I (whether open, closed or half-open), we define it's **volume** to be

$$\mathrm{vol}(I) := \begin{cases} \prod_{i=1}^{d}(b_i - a_i) \in [0, +\infty] & \text{if } a_i < b_i, \quad \forall i \\ 0 & \text{otherwise} \end{cases}$$

An **elementary set** is the finite disjiont union of intervals and we define its volume to be

$$\mathrm{vol}\left(\bigsqcup_{k=1}^{d} I_k\right) := \sum_{k=1}^{d} \mathrm{vol}(I_k) \in [0, \infty]$$

Remark 1.4.2. We can check easily that the volume function is well defined. For example, the decomposition $[0,2] = [0,1) \sqcup [1,2] = [0,1) \sqcup [1,1.5) \sqcup [1.5,2]$ should all give the same volume.
More generally, if $I = \bigsqcup_{k=1}^{n} I_k = \bigsqcup_{j=1}^{m} J_j$ where I_k, J_j are Intervals, then

$$\sum_{k=1}^{n} \mathrm{vol}(I_k) = \sum_{j=1}^{m} \mathrm{vol}(J_j)$$

Proof. Let $(I_k)_{k\in\mathbb{N}}$ and $(J_j)_{j\in\mathbb{N}}$ be as above. Then

$$I_k = I \cap I_k = \bigcup_{j=1}^{m} J_j \cap I_k$$

taking the volume on both sides and summing over all k, we get

$$\sum_{k=1}^{n} \mathrm{vol}(I_k) = \sum_{k=1}^{n}\sum_{j=1}^{m} \mathrm{vol}(J_j \cap I_k)$$

flipping the roles of I_k and J_j, e also get

$$\sum_{j=1}^{m} \mathrm{vol}(J_j) = \sum_{j=1}^{m}\sum_{k=1}^{n} \mathrm{vol}(J_j \cap I_k)$$

which equals what we got before. □

We of course have to show that our attempt to use the Carathéodory-Hahn Extension of vol on the elementary sets is well defined. But it should be easy to see how the class of elmentary sets forms an algebra and that the vol function is a pre-measure on it. In our example above, we used half-open intervals of length $1, 2^{-1}$ to decompose the interval $[0,2] \subseteq \mathbb{R}$.
A direct generalisation for this in higher dimensions is to introduce finer and finer hypercubes that cover $\mathbb{R}^d$. For $k \in \mathbb{N}$ let $\mathcal{D}_k$ the collection of half open cubes

$$\mathcal{D}_k := \left\{ \prod_{i=1}^{d} [\frac{a_i}{2^k}, \frac{a_i + 1}{2^k}) | a_i \in \mathbb{Z} \right\}$$

In particular, $\mathcal{D}_0$ is the collection of hpyercubes of edge length 1 and vertices in $\mathbb{Z}^d$.
We call the cubes of the collection

$$\{Q | Q \in \mathcal{D}_k, k = 0, 1, 2, \dots\}$$

the **dyadic cubes**.

Remark 1.4.3. The dyadic cubes have the following properties:

(a) For all $k \in \mathbb{N}$, $\mathbb{R}^n = \bigsqcup_{Q \in \mathcal{D}_k} Q$.

(b) If $Q \in \mathcal{D}_k$ and $P \in \mathcal{D}_l$, with $l \leq k$, then either $Q \subseteq P$ or $P \cap Q = \emptyset$.

(c) Every $Q \in \mathcal{D}_k$ has volume $\mathrm{vol}(Q) = 2^{-kn}$.

Definition 1.4.4. The **Lebesgue measure** $\mathcal{L}^n$ is the Carathéodory Hahn extension of the volume defined on the algebra of elementary sets[2], i.e.

$$\mathcal{L}^n(A) := \inf \left\{ \sum_{k=1}^{\infty} \mathrm{vol}(E_k) | A \subseteq \bigcup_{k=1}^{\infty} E_k, E_k \text{ is an elementary set} \right\}$$

If we want to measure open subsets $U \subseteq \mathbb{R}^n$ with the Lebesgue-measure, we want to ensure that a countable covering of U with disjoint elementary sets E_k is possible, or else taking the infimum makes it so that U is not $\mathcal{L}^n$-measurable.

Lemma 1.4.5. Every open set in $\mathbb{R}^n$ can be written as a countable union of disjoint dyadic cubes.

Proof. Let $U \subseteq \mathbb{R}^n$ be a non-empty open subset.
Let $\mathcal{S}_0$ to be the collection of all cubes in $\mathcal{D}_0$ that lie entirely in U. Let $\mathcal{S}_1$ to be the collection of all cubes in $\mathcal{D}_1$ that lie entirely in U, but are not subcubes of $\mathcal{S}_0$, etc. Let $\mathcal{S}_k$ be the collection of cubes in $\mathcal{D}_k$ which are not subcubes of any cubes in $\mathcal{S}_0, \dots, \mathcal{S}_{k-1}$. Set $\mathcal{S} := \bigcup_{k \in \mathbb{N}} \mathcal{S}_k$.
Because each $\mathcal{D}_k$ is countable, $\mathcal{S}$ is countable. By construction, the cubes in $\mathcal{S}$ are also disjoint.
Since U is open and the cubes become arbitrarily small, every $x \in U$ will be covered by some $Q \in \mathcal{S}$, so $U = \bigsqcup_{Q \in \mathcal{S}} Q$.

□

Recall that the Borel σ-algebra $\mathcal{B}(X)$ is the σ-algebra generated by open subsets of X.

Definition 1.4.6. A measure μ on $\mathbb{R}^n$ is called **Borel** (or a Borel measure), if every Borel set is μ-measurable.

Remark 1.4.7. From Lemma 1.4.5, it follows that $\mathcal{L}^n$ is a Borel measure.
The lemma says that the open sets are $\mathcal{L}^n$-measurable. Moreover, by Theorem 1.2.9 the the collection of $\mathcal{L}^n$-measurable sets form a σ-algebra. So the Borel σ-algebra is contained in the σ-algebra of $\mathcal{L}^n$-measurable subsets.

When we want to characterize $\mathcal{L}^n(A)$ for some subset $A \subseteq \mathbb{R}^n$, the definition used in the Carathéodory-Hahn extension where we consider all countable coverings using elementary sets is quite unwiedly. The following theorem gives a nicer characterisation.

Theorem 1.4.8. For every $A \subseteq \mathbb{R}^n$ it holds

$$\boxed{\mathcal{L}^n(A) = \inf_{A \subseteq U} \mathcal{L}^n(U), \quad U \text{ open}}$$

[2] Because elementary sets are finite disjoint unions of intervals, we can replace E_k with intervals I_k

Proof. By monotonicity, $\mathcal{L}^n(A) \leq \mathcal{L}^n(U)$ follows directly.
For the other inequality, suppose that $\mathcal{L}^n(A) < \infty$ (or else the inequality is trivial). By definition, for any $\varepsilon > 0$ we can find intervals $(I_k)_{k\in\mathbb{N}}$ with

$$A \subseteq \bigcup_{k=1}^{\infty} I_k, \quad \sum_{k=1}^{\infty} \operatorname{vol}(I_k) \leq \mathcal{L}^n(A) + \varepsilon$$

Since $\mathcal{L}^n(A) < \infty$, every interval I_k must have finite volume and is thus bounded. So let $\tilde{I}_k \supseteq I_k$ be open bounded intervals with $\operatorname{vol}(\tilde{I}_k) \leq \operatorname{vol}(I_k) + \frac{\varepsilon}{2^k}$.
Setting $U := \bigcup_{k=1}^{\infty} \tilde{I}_k$, we see that U is an open subset containing A and it's volume is

$$\mathcal{L}^n(U) \leq \sum_{k=1}^{\infty} \operatorname{vol}(\tilde{I}_k) \leq \sum_{k=1}^{\infty} \operatorname{vol}(I_k) + \frac{\varepsilon}{e^k} \leq \mathcal{L}^n(A) + 2\varepsilon$$

since ε was arbitrary, the result follows. □

This alternative characterisation lets us find out what subsets $A \subseteq \mathbb{R}^n$ are $\mathcal{L}^n$-measurable.

Theorem 1.4.9. For any subset $A \subseteq \mathbb{R}^n$ the following are equivalent

(a) A is $\mathcal{L}^n$-measurable.

(b) $\forall \varepsilon > 0 \; \exists U \supseteq A$ open with $\mathcal{L}^n(U \setminus A) < \varepsilon$.

(c) A it can be "approximated" from the inside and outside: $\forall \varepsilon > 0 \; \exists F$ closed, U open with $F \subseteq A \subseteq U$ such that

$$\mathcal{L}^n(U \setminus A) + \mathcal{L}^n(A \setminus F) < \varepsilon$$

(d) $\forall \varepsilon > 0 \exists F$ closed, $\exists U$ open, such that $F \subseteq A \subseteq U$ and $\mathcal{L}^n(U \setminus F) < \varepsilon$.

Proof.

(a) $\implies$ (b): Let $\varepsilon > 0$, A be $\mathcal{L}^n$-measurable.

- If $\mathcal{L}^n(A) < \infty$, by the previous theorem, we can chose a $U \supseteq A$ open such that

$$\mathcal{L}^n(U) \leq \mathcal{L}^n(A) + \varepsilon$$

  Because A is $\mathcal{L}^n$-measurable we can use U as a test set and get

$$\begin{aligned}\mathcal{L}^n(U) &= \mathcal{L}^n(U \cap A) + \mathcal{L}^n(U \setminus A) \\ &= \mathcal{L}^n(A) + \mathcal{L}^n(U \setminus A)\end{aligned}$$

  which gives us

$$\mathcal{L}^n(U \setminus A) = \mathcal{L}^n(U) - \mathcal{L}^n(A) < \varepsilon$$

- If $\mathcal{L}^n(A) = \infty$, we set

$$A_k = A \cap [-k, k]^n \implies A = \bigcup_{k=1}^{\infty} A_k$$

since $\mathcal{L}^n(A_k) < \infty$, we are in the first case so we can find $U_k \supseteq A_k$ open with

$$\mathcal{L}^n(U_k \setminus A_k) < \frac{\varepsilon}{2^k} \quad \forall k \in \mathbb{N}$$

Then their union $U := \bigcup_{k=1}^{\infty} U_k$ is open and contains A. Moreover, we have

$$\begin{aligned} \mathcal{L}^n(U \setminus A) &= \mathcal{L}^n\left(\bigcup_{k=1}^{\infty}(U_k \setminus A)\right) \\ &\leq \mathcal{L}^n\left(\bigcup_{k=1}^{\infty}(U_k \setminus A_k)\right) \\ &\leq \sum_{k=1}^{\infty} \mathcal{L}^n(U_k \setminus A_k) < \varepsilon \end{aligned}$$

(b) $\implies$ (a): Let $B \subseteq \mathbb{R}^n$. For $\varepsilon > 0$, chose $U \supseteq A$ open with $\mathcal{L}^n(U \setminus A) < \varepsilon$. Then

$$B \setminus A \subseteq (B \setminus U) \cup (U \setminus A)$$

Since open subsets are $\mathcal{L}^n$-measurable, we have

$$\begin{aligned} \mathcal{L}^n(B) &= \mathcal{L}^n(B \cap U) + \mathcal{L}^n(B \setminus U) \\ &\geq \mathcal{L}^n(B \cap A) + \mathcal{L}^n(B \setminus A) - \mathcal{L}^n(U \setminus A) \\ &\geq \mathcal{L}^n(B \cap A) + \mathcal{L}^n(B \setminus A) - \varepsilon \end{aligned}$$

since ε was arbitrary, $\mathcal{L}^n$-measurability of A follows.

(b) $\iff$ (c): For $\varepsilon > 0$ use (b) for A^c to get an open set $V \supseteq A^c$ with $\mathcal{L}^n(V \setminus A^c) < \varepsilon$. Then $F = V^c \subseteq A$ is closed and

$$\mathcal{L}^n(A \setminus V^c) = \mathcal{L}^n(V \setminus A^c) < \varepsilon$$

The other implication is trivial.

(c) $\implies$ (d): Using (c), we get $F \subseteq A$ closed and $U \supseteq A$ open. Because $F \subseteq A \subseteq U$,

$$U \setminus F = (U \setminus A) \cup (A \setminus F)$$

it follows from subadditivity that

$$\mathcal{L}^n(U \setminus F) \leq \mathcal{L}^n(U \setminus A) + \mathcal{L}^n(A \setminus F) < \varepsilon$$

(d) $\implies$ (c): For $\varepsilon > 0$, use (d) to get $F \subseteq A$ closed, $U \supseteq A$ open with $\mathcal{L}^n(U \setminus F) < \varepsilon$. Because $F \subseteq A \subseteq U$

$$U \setminus A \subseteq U \setminus F, \quad A \setminus F \subseteq U \setminus F$$

so we get

$$\mathcal{L}^n(U \setminus A) + \mathcal{L}^n(A \setminus F) \leq 2\mathcal{L}^n(U \setminus F) < 2\varepsilon$$

□

1.5 Comparision between Lebesgue and Jordan Measure

Definition 1.5.1. A bounded subset $A \subseteq \mathbb{R}^n$ is **Jordan-measurable** if $\underline{\mu}(A) = \overline{\mu}(A)$, where

$$\underline{\mu}(A) := \underline{\int}_{\mathbb{R}^n} \chi_A d\mu := \sup\{\mathrm{vol}(E) \big| E \subseteq A, E \text{ elementary set}\}$$
$$\overline{\mu}(A) := \overline{\int}_{\mathbb{R}^n} \chi_A d\mu := \inf\{\mathrm{vol}(E) \big| A \subseteq E, E \text{ elementary set}\}$$

If that is the case, denote the Jordan measure of A with the common value $\mu(A)$.
We call $\underline{\mu}(A)$ the **Jordan inner measure** of A and $\overline{\mu}(A)$ the **Jordan outer measure** of A.

Example 1.5.2. For $f : I \to \mathbb{R}$ continuous, $I \subseteq \mathbb{R}^n$ compact, its graph

$$\Gamma = \{(x, f(x)) \big| x \in I\} \subseteq \mathbb{R}^{n+1}$$

is a Jordan measurable set.
The area under a function

$$G = \{(x,t) \in I \times \mathbb{R} \big| 0 \leq t \leq f(x)\}$$

is also Jordan-measurable

As the following theorem will show, the Lebesgue measure can measure more sets than the Jordan measure can.

Theorem 1.5.3. Let $A \subseteq \mathbb{R}^n$ be bounded, then

(a) $\underline{\mu}(A) \leq \mathcal{L}^n(A) \leq \overline{\mu}(A)$

(b) If A is Jordan-measurable, then A is $\mathcal{L}^n$-measurable and $\mathcal{L}^n(A) = \mu(A)$.

Proof. (a) Because elementary sets are finite disjoint unions of intervals, we have

$$\begin{aligned}
\mathcal{L}^n(A) &= \inf\left\{\sum_{k=1}^{\infty} \mathrm{vol}(I_k) \big| A \subseteq \bigcup_{k=1}^{\infty} I_k, I_k \text{ intervals}\right\} \\
&\leq \inf\left\{\sum_{k=1}^{m} \mathrm{vol}(I_k) \big| A \subseteq E = \bigsqcup_{k=1}^{m} I_k, I_k \text{ intervals}\right\} \\
&= \overline{\mu}(A)
\end{aligned}$$

For the other inequality, for every elementary set $E = \bigsqcup_{k=1}^{m} I_k \subseteq A$ we have

$$\mathrm{vol}(E) = \mathcal{L}^n(E) \leq \mathcal{L}^n(A)$$

so when taking the sup over such E, we get

$$\underline{\mu}(A) \leq \mathcal{L}^n(A)$$

(b) If A is Jordan measurable, then it follows from (i) that

$$\underline{\mu}(A) \leq \mathcal{L}^n(A) \leq \overline{\mu}(A) = \underline{\mu}(A)$$

To show that A is $\mathcal{L}^n$-measurable, we use characterisation (b) from Theorem 1.4.9

Because A is bounded, $\mathcal{L}^n(A) < \infty$ and because it is Jordan-measurable, we can find for all $\varepsilon > 0$ elementary sets $E_\varepsilon, E^\varepsilon$ such that

$$E_\varepsilon \subseteq A \subseteq E^\varepsilon \quad \text{and} \quad \mathrm{vol}(E^\varepsilon) - \varepsilon < \mu(A) < \mathrm{vol}(E_\varepsilon) + \varepsilon$$

Because the volume doesn't depend on whether the intervals comprising the elementary set are open, half-open or closed, we can assume WLOG that E^ε is open, so

$$\begin{aligned}\mathcal{L}^n(E^\varepsilon \setminus A) &\leq \mathcal{L}^n(E^\varepsilon \setminus E_\varepsilon) = \mathrm{vol}(E^\varepsilon \setminus E_\varepsilon)\\ &= \mathrm{vol}(E^\varepsilon) - \mathrm{vol}(E_\varepsilon) < 2\varepsilon\end{aligned}$$

which shows the condition from the previous theorem.

□

One would naturally think that the "physical" volume of an object should stay invariant under translation or rotation.

Theorem 1.5.4. The Lebegue measure is invariant under isometries of $\mathbb{R}^n$, which are maps

$$\Phi : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto x_0 + Rx, \quad R \in O(n)$$

Proof. `Missing` □

Definition 1.5.5. A Borel measure μ on $\mathbb{R}^n$ is called **Borel regular**, if for every $A \subseteq \mathbb{R}^n$ there exists a Borel set $B \supseteq A$ such that $\mu(A) = \mu(B)$.

Lemma 1.5.6. The Lebesgue measure is Borel regular.

Proof. If $\mathcal{L}^n(A) = \infty$, we can simply take $B = \mathbb{R}^n$, so assume $\mathcal{L}^n(A) < \infty$.
By the characterisation with open sets from Theorem 1.4.8, we can chose for every $k \in \mathbb{N}$ an open set $U_k \supseteq A$ open with

$$\mathcal{L}^n(U_k) < \mathcal{L}^n(A) + \frac{1}{k}, \quad k \in \mathbb{N}$$

by intersecting each U_k with the previous ones, we can also assume without loss of generality that the sequence $(U_k)_{k\in\mathbb{N}}$ is monotonously decreasing (i.e. $U_{k+1} \subseteq U_k$).
By Remark 1.4.7, the open sets U_k are in the σ-algebra of $\mathcal{L}^n$-measurable subsets. Setting $B := \bigcap_{k=1}^\infty U_k$ it follows from continuity from above (Theorem 1.2.12)

$$\mathcal{L}^n(B) \overset{\text{c.f.a}}{=} \lim_{k\to\infty} \mathcal{L}^n(U_k) = \mathcal{L}^n(A)$$

□

1.6 Special-Examples of sets

As we will see, not all subsets of $\mathbb{R}^n$ are $\mathcal{L}^n$-measurable.
To construct such a non-measurable set, we will use the Axiom of Choice, which states that for any family of non-empty disjoint sets $(A_i)_{i\in I}$, there exists a choice-function $f : I \to \bigcup_{i\in I} A_i$ such that $f(i) \in A_i$.
With this, we can construct the set $\{f(i) | i \in I\}$ that contains exactly one element from each set A_i.

For $x, y \in [0,1)$ we define $\oplus := (\mod 1 \circ +)$

$$x \oplus y = \begin{cases} x+y & \text{if } x+1 < 1 \\ x+y-1 & \text{if } x+y \geq 1 \end{cases}$$

So if we have a subset $E \subseteq [0,1)$, we can "shift" the set E by x, with $E \oplus x \subseteq [0,1)$.
Where some part $E \cap [0, 1-x)$ moves naturally to the right and the set $E \cap [1-x, 1)$ moves back to the left side. Set

$$E_1 := E \cap [0, 1-x) \oplus x$$
$$E_2 := E \cap [1-x, 1) \oplus x$$

which are disjoint.
If E is $\mathcal{L}^1$-meaurable, then the translated sets E_1, E_2 are also $\mathcal{L}^1$-measurable and

$$\begin{aligned}\mathcal{L}^1(E \oplus x) &= \mathcal{L}^1(E_1) + \mathcal{L}^1(E_2) \\ &= \mathcal{L}^1(E \cap [0, 1-x)) + \mathcal{L}^1(E \cap [1-x, 1)) \\ &= \mathcal{L}^1(E)\end{aligned}$$

A non-measurable set

Then we define the equivalence relation

$$x, y \in [0,1) \quad x \sim y \iff x - y \in \mathbb{Q}$$

by the axiom of choice, there exists a set $P \subseteq [0,1)$ that contains exactly one representative of each equivalence class.
By enumerating all rational points in $[0,1)$ by an index $Q \cap [0,1) = \{r_k\}_{k \in \mathbb{N}}$ with $r_0 = 0$ we define

$$P_k := P \oplus r_k$$

Then it is easy to see that

(a) The P_j are disjoint and $[0,1) = \bigsqcup_{j=0}^{\infty} P_j$.

Because if $x \in P_n \cap P_m$, then $x = p_n \oplus r_n = p_m \oplus r_m$. Since $r_n, r_m \in \mathbb{Q}$ it follows that also $p_n - p_m \in \mathbb{Q}$ so they must be of the same equivalence class.

It also covers $[0,1)$ because by construction, every $x \in [0,1)$ belongs to a unique equivalence class.

(b) If P were $\mathcal{L}^1$-measurable, then so is $P_j = P \oplus r_j$ and $\mathcal{L}^1(P) = \mathcal{L}^1(P_j)$.

We just showed this earlier.

But P cannot be $\mathcal{L}^1$-measurable, because by σ-additivity on $\mathcal{L}^1$-measurable subsets

$$1 = \mathcal{L}^1([0,1)) = \sum_{i=0}^{\infty} \mathcal{L}^1(P_j) = \sum_{i=0}^{\infty} \mathcal{L}^1(P)$$

and the right hand side is either 0 or infinite.
So since P is not $\mathcal{L}^1$-measurable there exists a set $B \subseteq \mathbb{R}$ with

$$\mathcal{L}^1(B) < \mathcal{L}^1(B \cap P) + \mathcal{L}^1(B \setminus P)$$

We also know that $\mathcal{L}^1(P)$ can't be zero, or else it would be $\mathcal{L}^1$-measurable.
Moreover, if $E \subseteq P$ is $\mathcal{L}^1$-measurable, then $\mathcal{L}^1(E) = 0$ because we can set

$$E_i := E \oplus r_i \implies F := \bigsqcup_{i=0}^{\infty} E_i \subseteq [0,1) \text{ is } \mathcal{L}^1\text{-measurable}$$

and we have

$$1 = \mathcal{L}^1([0,1)) \geq \mathcal{L}^1(F) = \sum_{i=0}^{\infty} \mathcal{L}^1(E_i) = \sum_{i=0}^{\infty} \mathcal{L}^1(E)$$

which can only be true if $\mathcal{L}^1(E) = 0$.
Not only does there exists a non-$\mathcal{L}^1$-measurable subset, we can construct more using P as a "template".

Proposition 1.6.1. For every $A \subseteq \mathbb{R}$ with $\mathcal{L}^1(A) > 0$, there exists a subset $B \subseteq A$ that is not $\mathcal{L}^1$-measurable.

Proof. Because we can shift and scale A or take subsets of A, we can assume without loss of generality that $A \subseteq (0,1)$.
Then set $B_i = A \cap P_i$. Then $A = \bigsqcup_{i=0}^{\infty} B_i$
As we showed earlier, if B_i were $\mathcal{L}^1$-measuralble, then $\mathcal{L}^1(B_i) = 0$, which contradicts $\mathcal{L}^1(A) = \sum_{i=0}^{\infty} \mathcal{L}^1(B_i)$. □

Remark 1.6.2. Because singletons $\{\alpha\} \in \mathbb{R}$ are contained in the arbitrarily small interval $(\alpha - \varepsilon, \alpha + \varepsilon)$ with Lebesgue measure 2ε, singletons have Lebesgue measure zero.
It follows that by subadditivity, every countable subset of $\mathbb{R}$ also has Lebesgue measure zero.

The Cantor tridadic set

The real numbers can be defined as the set of Cauchy-sequences in $\mathbb{Q}$ up to equivalence of Cauchy sequences. This gives for every $x \in \mathbb{R}$ and base $b > 2 \in \mathbb{N}$ a b-ary expansion with digits $d_i(x) \in \{0, \dots, b-1\}$.

$$x = \sum_{i=1}^{\infty} d_i(x) b^{-i}$$

Although the digit expansion is not always unique, the set of those with multiple expansions is countable and thus have measure zero.

Proposition 1.6.3. The **Cantor set** is the set of numbers whose 3-adic digits don't contain a 1.

$$C = \{x \in [0,1] | d_i(x) \in \{0,2\} \forall i\}$$

Then C is uncountable and $\mathcal{L}(C) = 0$.

Proof. We construct the Cantor set as $C := \bigcap_{n=1}^{\infty} C_n$, where

$$C_n = \{x \in [0,1] | d_i(x) \neq 1 \forall i \leq n$$

Then each C_n can be written as a finite union closed intervals. For example

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right], \text{etc.}$$

They are in particular Borel and have Lebesgue measure

$$\mathcal{L}^1(C_n) = \left(\frac{2}{3}\right)^n$$

because this sequence is decreasing, by continuity from above we have

$$\mathcal{L}^1(C) = \mathcal{L}^1\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} \mathcal{L}^1(C_n) = 0$$

To show that C is countable, we define a function that maps elements of C surjectively to $[0,1]$
Because C consists of numbers whose 3-ary sequence of digits don't contain a 1, i.e. only use the digits 0 and 2, we can map them to binary sequences of digits which use the digit 0 and 1, by converting every digit 2 to a 1 and look at it as a binary sequence.

$$f : C \to [0,1], \quad \sum_{i=1}^{\infty} \frac{d_i(x)}{3^i} \mapsto \sum_{i=1}^{\infty} \frac{d_i(x)}{2} \frac{1}{2^i}$$

For example, $\frac{8}{27} = 0.022_3 \mapsto 0.011_2 = \frac{3}{8}$.
Because this lets us generate any (even infinite) binary sequence of digits, the map is surjective. □

The construction of the Cantor set can be generalised to give us the so-called **fat Cantor sets**, where we start off with the interval $I_1 = [0,1]$, and for $n \in \mathbb{N}$, if some interval inside I_n has length ℓ, then we remove the centered subinterval of length $\beta\ell$ and let $I_{n+1} \subseteq I_n$ be the remaining pieces of this operation. The fat cantor set with parameter β is then $C_\beta := \bigcap_{n=1}^{\infty} I_n$.
We see that the "normal" Cantor set has the parameter $\beta = \frac{1}{3}$ and if $\beta < \frac{1}{3}$, then we have

$$\mathcal{L}^1(I_n \setminus I_{n+1}) = 2^{n-1}\beta^n \implies \mathcal{L}^1(I_1 \setminus C_\beta) = \sum_{n=1}^{\infty} 2^{n-1}\beta^n = \frac{\beta}{1-2\beta}$$

but this set is not Jordan-measurable as

$$\underline{\mu}(C_\beta) = 0 \quad \text{but} \quad \overline{\beta}(C_\beta) = 1 - \frac{\beta}{1-2\beta}$$

1.7 The Lebesgue-Stieltjes Measure

Let $F : \mathbb{R} \to \mathbb{R}$ be non-decreasing and continuous from the left, i.e:

$$F(x_0) = \lim_{x\to x_0^-} F(x) := \lim_{\substack{x\to x_0 \\ x<x_0}} F(x) \quad \forall x_0 \in \mathbb{R}$$

For a, b we define

$$\lambda_F[a,b) = \begin{cases} F(b) - F(a) & \text{if } a < b \\ 0 & \text{otherwise} \end{cases}$$

Because the collection of bounded half-open sets $\mathcal{K} = \{[a,b) | a, b \in \mathbb{R}\}$ does not form an Algebra (see 1.7.6), we cannot use the Carathéodory-Hahn extension theorem to produce a measure induced by λ_F.
However, $\mathcal{K}$ constitutes a covering of $\mathbb{R}$ as in Definition 1.3.1, so by Theorem 1.3.3, the function λ_F induces a measure

$$\Lambda_F(A) := \inf\left\{\sum_{k=1}^{\infty} \lambda_F[a_k, b_k), \quad A \subseteq \bigcup_{k=1}^{\infty} [a_k, b_k)\right\}$$

called the **Lebesgue-Stieltjes Measure** generated by F.
To find out if Λ_F is nice, we will find the following definition useful.

Definition 1.7.1. A measure μ on $\mathbb{R}^n$ is called **metric,** if the measure is additive on separated sets, i.e. for all $A, B \subseteq \mathbb{R}^n$ with

$$\operatorname{dist}(A,B) := \inf\{|a-b|, a \in A, b \in B\} > 0$$

it holds

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

By subadditivity, the inequality "$\geq$" is sufficient.

Theorem 1.7.2 (Carathéodory criterion for Borel measures)**.** A metric measure μ on $\mathbb{R}^n$ is Borel.

Proof. Let μ be a metric measure on $\mathbb{R}^n$. Because the μ-measurable subsets (1.2.9) form a σ-Algebra, it is sufficient to show that closed sets are μ-measurable.
Let $F \subseteq \mathbb{R}^n$ be closed and $B \subseteq \mathbb{R}^n$ be some test set. If $\mu(B) = \infty$, then the inequality

$$\mu(B) \geq \mu(B \cap F) + \mu(B \setminus F)$$

is trivial, so assume $\mu(B) < \infty$. For $k = 1, 2, \ldots$, we define

$$F_k := \{x \in \mathbb{R}^n | 0 \leq \operatorname{dist}(x, F) \leq \frac{1}{k}\}$$

It should be clear that

$$\operatorname{dist}(B \setminus F_k, B \cap F) \geq \frac{1}{k} > 0$$

so since μ is metric and monotonous, we have

$$\mu(B \cap F) + \mu(B \setminus F_k) = \mu((B \cap F) \cup (B \setminus F_k)) \leq \mu(B) \quad \forall k$$

If we can show that $\lim_{k\to\infty} \mu(B \cap F_k) = \mu(B \setminus F)$, then we are done.
To do so, first note that the $(F_k)_{k\in\mathbb{N}}$ form a decreasing sequence $F_{k+1} \subseteq F_k$. Moreover, we have $F = \bigcap_{k=1}^{\infty} F_k$, so we can write

$$B \setminus F = B \setminus \bigcap_{l=1}^{\infty} F_l = \bigcup_{l=1}^{\infty} (B \setminus F_l)$$

We can expand the union above in telescoping fashion [3] and use the fact that the $(B \setminus F_l)_{l\in\mathbb{N}}$ form an increasing sequence to get

$$\begin{aligned}\bigcup_{l=1}^{\infty}(B \setminus F_l) &= (B \setminus F_1) \cup \bigcup_{l=1}^{\infty}(B \setminus F_{l+1}) \setminus (B \setminus F_l) \\ &= (B \setminus F_k) \cup \bigcup_{l=k}^{\infty}(F_l \setminus F_{l+1}) \cap B\end{aligned}$$

Setting

$$R_l := (F_l \setminus F_{l+1}) \cap B = \{x \in B | \frac{1}{l+1} < d(x,F) \leq \frac{1}{l}\}$$

[3]For example, for any sequence $(A_l)_{l\in\mathbb{N}}$ we can write $\bigcup_{l=1}^{\infty} A_l = A_1 \cup \bigcup_{l=1}^{\infty} A_{l+1} \setminus A_l$.

we see that the $(R_l)_{l\in\mathbb{N}}$ are pairwise disjoint, so we have

$$B \setminus F = (B \setminus F_k) \cup \bigsqcup_{l=k}^{\infty} R_l$$

Therefore, for all $k \in \mathbb{N}$ it holds

$$\mu(B \setminus F_k) \leq \mu(B \setminus F) \leq \mu(B \setminus F_k) + \sum_{l=k}^{\infty} \mu(R_l)$$

Now we only need to show that

$$\lim_{k\to\infty} \sum_{l=k}^{\infty} \mu(R_l) = 0$$

Observe that R_i only "touches" its neighbors R_{i-1}, R_{i+1}, in other words

$$\operatorname{dist}(R_i, R_j) > 0, \quad \text{if } |i-j| \geq 2$$

decomposing the sum $\sum_{l=1}^{\infty} \mu(R_l)$ into the even and odd numbers, we can use the fact that μ is metric to get

$$\sum_{l=1}^{2m+1} \mu(R_l) = \left(\sum_{k=1}^{m} \mu(R_{2k})\right) + \left(\sum_{k=1}^{m} \mu(R_{2k})\right) = \mu\left(\bigcup_{k=1}^{m} R_{2k}\right) + \mu\left(\bigcup_{k=1}^{m} R_{2k}\right) \leq 2\mu(B) < \infty$$

so even in the limit $m \to \infty$, the series converges. But in the inequality we showed earlier

$$\mu(B \setminus F_k) \leq \mu(B \setminus F) \leq \mu(B \setminus F_k) + \sum_{l=k}^{\infty} \mu(R_l)$$

we were allowed to omit any number of (non-negative) starting terms $\mu(R_l)$ for $l < k$, so in the limit we get $\lim_{k\to\infty} \mu(B \setminus F_k) = \mu(B \setminus F)$, and the result follows. ☐

Theorem 1.7.3. The Lebesgue-Stieltjes measure Λ_F is Borel regular.

Proof. (i) **$\boldsymbol{\Lambda_F}$ is Borel.** We show that it is metric and use the previous theorem, so let $A, B \subseteq \mathbb{R}$ with $\delta := \operatorname{dist}(A, B) > 0$. We now show that for all $\varepsilon > 0$ we have

$$\Lambda_F(A) + \Lambda_F(B) \leq \Lambda_F(A \cup B) + \varepsilon$$

Per definition of the Lebesgue-Stieltjes measure, we can find a collection of half-open intervals with

$$A \cup B \subseteq \bigcup_{k=1}^{\infty} [a_k, b_k) \quad \text{and} \quad \sum_{k=1}^{\infty} < \Lambda_F(A \cup B) + \varepsilon$$

Because we can always subdivide any interval $[a_k, b_k)$ further, we may also assume that $|b_k - a_k| < \delta$ for all k.

Because A and B are separated, for each interval $[a_k, b_k)$ either

$$A \cap [a_k, b_k) = \emptyset \quad \text{or} \quad B \cap [a_k, b_k) = \emptyset$$

, so the covering of $A \cup B$ gives us a covering $\mathcal{A}$ of A and a covering $\mathcal{B}$ of B. Therefore

$$\begin{aligned}\Lambda_F(A) + \Lambda_F(B) &\leq \sum_{[a_k,b_k)\in\mathcal{A}} \Lambda_F([a_k,b_k)) + \sum_{[a_k,b_k)\in\mathcal{B}} \Lambda_F([a_k,b_k)) \\ &= \sum_{k\in\mathbb{N}} \lambda_F([a_k,b_k) \leq \Lambda_F(A\cup B) + \varepsilon\end{aligned}$$

This shows that Λ_F is metric and thus Borel.

(ii) **Λ_F is Borel regular.**
To show that Λ_F is Borel regular, let $A \subseteq \mathbb{R}$. Of course we can assume $\Lambda_F(A) < \infty$. Then for any $n \in \mathbb{N}$ we can find coverings

$$A \subseteq \bigcup_{k=1}^{\infty} \left[a_k^{(n)}, b_k^{(n)}\right) =: B_n \quad \text{with} \quad \sum_{k=1}^{\infty} \lambda_F\left(\left[a_k^{(n)}, b_k^{(n)}\right)\right) \leq \Lambda_F(A) + \frac{1}{n}$$

If we set $B := \bigcap_{n=1}^{\infty} B$, then B is Borel and $A \subseteq B \subseteq B_n$ and

$$\Lambda_F(A) \leq \Lambda_F(B) \leq \Lambda_F(B_n) \leq \sum_{k=1}^{\infty} \lambda_F\left(\left[a_k^{(n)}, b_k^{(n)}\right)\right) \leq \Lambda_F(A) + \frac{1}{n}$$

in the limit $n \to \infty$, we get $\Lambda_F(A) = \Lambda_F(B)$, so Λ_F is Borel regular.

□

The Carathéodory-Hahn extension had the property that it coincided with the pre-measure on the algebra, on which the pre-measure was defined. Despite not being such an extension, the Lebesgue-Stieltjes measure has a similar property.

Theorem 1.7.4. For $a < b \in \mathbb{R}$ it holds

$$\Lambda_F([a,b)) = \lambda_F([a,b)) = F(b) - F(a)$$

Proof. Let $a < b \in \mathbb{R}$. By definition of Λ_F, we already have $\Lambda_F([a,b)) \leq \lambda_F([a,b))$.
For the other inequality, let $([a_k,b_k))_{k\in\mathbb{N}}$ be a covering of $[a,b)$.
Since F is left-continuous, for every $\varepsilon > 0$ there exist $\delta, \delta_k > 0$ such that

$$F(b) - F(b-\delta) \leq \varepsilon, \quad \text{and} \quad F(a_k) - F(a_k - \delta_k) \leq 2^{-k}\varepsilon \quad \forall k \in \mathbb{N}$$

Because $[a, b-\delta]$ is compact and is covered by $\bigcup_{k=0}^{\infty}(a_k - \delta_k, b_k)$, there exists a finite subcovering

$$[a, b-\delta] \subseteq \bigcup_{k=0}^{m} (a_k - \delta_k, b_k)$$

By removing any redundant intervals, we can decrease the sum $\sum_{k=0}^{m} \lambda_F(a_k - \delta_k, b_k)$, so we can assume WLOG that they are ordered in such a way that

$$a_k - \delta_k < b_{k-1} \quad \text{for all} \quad k = 1, \ldots, m$$

Since F is increasing and $a_0 - \delta_0 < a < b < b_m$, we have

$$\begin{aligned}F(b-\delta) - F(a) &\leq F(b_m) - F(a_0 - \delta_0) \\ &\leq F(b_m) - F(a_1 - \delta_1) + F(b_0) - F(a_0 - \delta_0) \\ &\leq \ldots \leq \sum_{k=0}^{m} F(b_k) - F(a_k - \delta_k)\end{aligned}$$

so with the initial estimates, we have

$$\begin{aligned}
\lambda_F([a,b)) &= F(b) - F(a) = F(b) - F(b-\delta) + F(b-\delta) - F(a) \\
&\leq \varepsilon + \sum_{k=0}^{m} F(b_k) - F(a_k - \delta_k) \\
&= \varepsilon + \sum_{k=0}^{m} F(b_k) - F(a_k) + \sum_{k=0}^{m} F(a_k) - F(a_k - \delta_k) \\
&\leq \varepsilon + \sum_{k=0}^{\infty} F(b_k) - F(a_k) + \sum_{k=0}^{\infty} 2^{-k}\varepsilon \\
&= \sum_{k=0}^{\infty} \lambda_F([a_k, b_k)) + 3\varepsilon
\end{aligned}$$

Since this is true for all coverings $([a_k, b_k))_{k\in\mathbb{N}}$, we get in the limit $\varepsilon \to 0$

$$\lambda_F([a,b)) \leq \Lambda_F([a,b))$$

□

Example 1.7.5.

- The Lebesgue measure is the special case when $F(x) = x$, so $\Lambda_{\mathrm{id}_{\mathbb{R}}} = \mathcal{L}^1$
- The Dirac measure δ_0 from Example 1.2.11 is the Lebesgue-Stieltjes measure Λ_Θ for the Heaviside step function

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Remark 1.7.6. In the beginning of this section, we noted that the collection of bounded half-open sets $\mathcal{K} = \{[a,b) | a, b \in \mathbb{R}\}$ does not form an Algebra. If we set $\tilde{\mathcal{K}}$ to be the collection of finite disjoint unions:

$$\tilde{\mathcal{K}} = \left\{ \bigsqcup_{k=1}^{m} [a_k, b_k) | m \geq 1 \in \mathbb{N}, a_k, b_k \in \mathbb{R} \right\}$$

then the set $\tilde{\mathcal{K}}$ is stable under *intersection* and *difference*. And we say that $\tilde{\mathcal{K}}$ forms a **ring**.
That is: $\emptyset \in \tilde{\mathcal{K}}$ and $A, B \in \tilde{\mathcal{K}} \implies A \cap B, A \setminus B \in \tilde{\mathcal{K}}$
The same is not true for the collection of open and closed intervals.

1.8 Hausdorff Measures

Say we have the unit square $A := [0,1]^2$. It's $\mathcal{L}^2$-measure would of course be 1.
If we however were to embed the square into $\mathbb{R}^3$ with

$$\iota : \mathbb{R}^2 \to \mathbb{R}^3, \quad (x,y) \mapsto (x,y,0)$$

we would find that $\mathcal{L}^3(\iota(A)) = 0$. More generally, the Lebesgue measure $\mathcal{L}^n$ on subsets $A \subseteq \mathbb{R}^n$ that have "dimension" $< n$ is always going to be zero.
Moreover, the Lebesgue measure also has the weakness of failing to properply measure fractal sets (which can be though of as having non-integer "dimension")

It would be nice to define a collection of measures that are able to measure sets, regardless of whether they are embedded into a higher-dimensional space. The Hausdorff measures try to solve this.
We start by introducing an intermediate measure, where instead of covering a subset A with dyadic cubes (as was the case for the Lebesgue measure) we do this using open balls of radius smaller than some $\delta > 0$.

Definition 1.8.1. For $s \geq 0, \delta > 0$ and $A \subseteq \mathbb{R}^n$ non-empty, we set

$$\mathcal{H}^s_\delta(A) := \inf \left\{ \sum_{k\in I} r_k^s \middle| I \text{ at most countable}, A \subseteq \bigcup_{k\in I} B(x_k, r_k), 0 < r_k < \delta \right\}$$

where we set $\mathcal{H}^0_\delta(\emptyset) = 0$.

Remark 1.8.2. $\mathcal{H}^s_\delta$ defines a measure on $\mathbb{R}^n$ and for fixed s, A, the function $\delta \mapsto \mathcal{H}^s_\delta(A)$ is non-increasing:

$$\delta_2 \leq \delta_1 \implies \mathcal{H}^s_{\delta_1}(A) \leq \mathcal{H}^s_{\delta_2}(A)$$

since every δ_2 covering is also a δ_1 covering. Therefore, the limit

$$\boxed{\mathcal{H}^s(A) := \lim_{\delta \downarrow 0} \mathcal{H}^s_\delta(A) = \sup_{\delta > 0} \mathcal{H}^s_\delta(A)}$$

exists. We now use this as for our next definition.

Definition 1.8.3. We call $\mathcal{H}^s$ the **s-dimensional Hausdorff measure** on $\mathbb{R}^n$

As hinted at earlier with the case of fractals, notice that s may take on non-integer values.
To build some intuition, let's consider an example of a "one-dimensional" set $A \subseteq \mathbb{R}^2$.

Example 1.8.4. Let $A = \mathbb{S}^1 = \{x \in \mathbb{R}^2, \|x\| = 1\}$.

$s = 0$: We see that $\mathcal{H}^0_\delta(A)$ depends only on the number of balls covering A.

If $\delta > 1$, we see that A can be covered by the ball $B(0, 1+\varepsilon)$, for ε small enough. Therefore $\mathcal{H}^0_\delta(A) = 1$ for $\delta > 1$.

On the other hand, if $\delta < 1$ then we have to cover A by using multiple balls. It should be clear that in the limit $\delta \to 0$, we have $\mathcal{H}^s(A) = \infty$.

$s = 1$: Again, for $\delta > 1$, it's easy to see that the covering with the single ball $B(0, 1+\varepsilon)$ give us $\mathcal{H}^1_\delta(A) \leq 1+\varepsilon$.

But for arbitrary $\delta > 0$, let $n \in \mathbb{N}$ such that $\delta > \frac{\pi}{n}$. We then can place n equally spaced balls along the circle each with radius $\frac{\pi}{n} + \varepsilon$, for some $\varepsilon > 0$ small enough. This covers A entirely and gives us an upper bound

$$\mathcal{H}^1_\delta(A) \leq \sum_{i=1}^n \frac{\pi}{n} = \pi$$

One can convince themself that there is no "better" covering strategy resulting in a lower upper bound, so $\mathcal{H}^1(A) = \pi$.

$s = 2$: The same covering strategy as decribed in the case $s = 1$ gives us the upper bound

$$\mathcal{H}^s_\delta(A) \leq \sum_{i=1}^n \left(\frac{\pi}{n}\right)^2 = \frac{\pi^2}{n}$$

But unlike for $s = 1$, chosing bigger and bigger n means that we can make the $\mathcal{H}^2$-measure of A arbitrarliy small. So $\mathcal{H}^2(A) = 0$.

A similar argumentation also shows that $\mathcal{H}^s(A) = 0$ for all $s > 1$.

Before we prove that $\mathcal{H}^s$ is actually a measure, let's take a look at the case $s = 0$ more closely.

Remark 1.8.5. $\mathcal{H}^0$ is the counting measure from 1.2.11.

Proof. If A is finite and has k elements $A = \{a_1, \dots, a_k\}$, let $\delta > 0$ be the minimal distance between all elements.
It easily follows from the triangle inequality, that $\mathcal{H}^0(A) \geq \mathcal{H}^0_{\frac{\delta}{2}} \geq k$. The other inequality is also trivial.
If A is infinite, then for any $k \in \mathbb{N}$ we can find a subset A_k with at least k elements. By monotonicity we have $k = \mathcal{H}^0(A_k) \leq \mathcal{H}^0(A)$ and in the limit $k \to \infty$, the proof follows.
And if A is empty, by defintion we have $\mathcal{H}^0(\emptyset) = 0$. □

Theorem 1.8.6. For $s \geq 0$, $\mathcal{H}^s$ is a Borel regular measure on $\mathbb{R}^n$

Proof. Let $s \geq 0$.

(i) **$\mathcal{H}^s$ is a measure.** Clearly, $\mathcal{H}^s(\emptyset) = 0$. Let $(A_k)_{k \in \mathbb{N}}$ and $A \subseteq \bigcup_{k=1}^\infty A_k$. Since $\mathcal{H}^s_\delta$ is σ-subadditive for all $\delta > 0$, we get

$$\mathcal{H}^s_\delta(A) \leq \sum_k \mathcal{H}^s_\delta(A_k) \leq \sum_k \mathcal{H}^s(A_k) \quad \forall \delta > 0$$

by taking the limit $\delta \to 0$ (as in the definition of $\mathcal{H}^s$) we get the σ-subadditivity of $\mathcal{H}^s$.

(ii) **$\mathcal{H}^s$ is metric and therefore Borel.** The proof is more or less the same as for the Lebesgue-Stieltjes measure. Let $A, B \subseteq \mathbb{R}^n$ such that $\delta_0 := \operatorname{dist}(A, B) > 0$. We then take a covering $A \cup B$ of balls of size smaller than $\delta := \frac{\delta_0}{4}$ and claim that we can partition the covering into two non-overlapping coverings of A and B each.

Since $\mathcal{H}^s_\delta$ takes the infimum over all such coverings, suppose that $A \cup B = \bigcup_k B(x_k, r_k)$ with $r_k < \delta$. Then we set

$$\mathcal{A} = \{B(x_k, r_k) \big| B(x_k, r_k) \cap A \neq \emptyset\} \mathcal{B} = \{B(x_k, r_k) \big| B(x_k, r_k) \cap B \neq \emptyset\}$$

And it becomes obvious that these are non-overlapping coverings of A and B each (by using the triangle inequality).

Therefore, we get

$$\mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B) \leq \sum_k r_k^s$$

and taking the infimum of coverings of $A \cup B$, this means

$$\mathcal{H}^s_\delta(A \cup B) \geq \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B)$$

which, when taking the limit $\delta \to 0$ just states $\mathcal{H}^s(A \cup B) \geq \mathcal{H}^s(A) + \mathcal{H}^s(B)$. By ($\sigma$)-subadditivity of $\mathcal{H}^s$, the reverse inequality holds and so $\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B)$ shows that $\mathcal{H}^s$ is metric and thus also Borel.

(iii) **$\mathcal{H}^s$ is Borel regular.** Again, the proof follows the same structure as in the proof for the Lebesgue-Stieltjes measure. Let $A \subseteq \mathbb{R}^n$ and suppose $\mathcal{H}^s(A) < \infty$ (Otherwise, just take $B = \mathbb{R}^n$). By monotonicity of $\mathcal{H}^s_\delta$, this also means that $\mathcal{H}^s_\delta(A) < \infty$ for all $\delta > 0$.

For $\delta = \frac{1}{m}, m = 1, 2, \dots$, this gives us a covering $\bigcup_{k\in I} B(x_{k,m}, r_{k,m}) \supseteq A$ with $r_{k,m} < \frac{1}{m}$ and

$$\sum_{k\in I} r_{k,m}^s \leq \mathcal{H}^s_{\frac{1}{m}}(A) + \frac{1}{m}$$

Then set $A_m := \bigcup_{k\in I} B(x_{k,m}, r_{k,m})$ and $B = \cap_{m=1}^{\infty} A_m$. Then B is a Borel set containing A. Which by monotonicity of $\mathcal{H}^s_{\frac{1}{m}}$ lets us sandwich

$$\begin{aligned}\mathcal{H}^s_{\frac{1}{m}}(A) \leq \mathcal{H}^s_{\frac{1}{m}}(B) \leq \mathcal{H}^s_{\frac{1}{m}}(A_m) &\leq \sum_{k\in I} r_{k,l}^s \\ &\leq \mathcal{H}^s_{\frac{1}{m}}(A) + \frac{1}{m}\end{aligned}$$

so in the limit $m \to \infty$, we get $\mathcal{H}^s(B) = \mathcal{H}^s(A)$.

□

In the example, where we calculated $\mathcal{H}^s(\mathbb{S}^1)$, we saw that

$$\mathcal{H}^0(A) = \infty, \quad \mathcal{H}^1(A) = \pi, \quad \mathcal{H}^2(A) = 0$$

The following Lemma proves the general pattern.

Lemma 1.8.7. Let $A \subseteq \mathbb{R}^n$ and $0 \leq s < t < \infty$. Then

(a) $\mathcal{H}^s(A) < \infty \implies \mathcal{H}^t(A) = 0$

(b) $\mathcal{H}^t(A) > 0 \implies \mathcal{H}^s(A) = \infty$

Proof. Since (b) is just the contraposition of (a), it's enough to prove (a).
Let $0 \leq s < t \in \mathbb{R}$ and $A \subseteq \mathbb{R}^n$ with $\mathcal{H}^s(A) < \infty$.
For any covering $A \subseteq \bigcup_{k\in I} B(x_k, r_k)$ with $r_k < \delta$, we have

$$\mathcal{H}^t_\delta(A) \leq \sum_{k\in I} r_k^t = \sum_{k\in I} r_k^{t-s} r_k^s \leq \delta^{t-s} \sum_{k\in I} r_k^s$$

Considering the infimum over all such coverings we get

$$\mathcal{H}^t_\delta(A) \leq \delta^{t-s}\mathcal{H}^s_\delta(A)$$

so as $\delta \to 0$, we get $\mathcal{H}^t(A) = 0$. □

This Lemma makes the definition of "dimension" possible.

Definition 1.8.8. The **Hausdorff dimension** of a subset $A \subseteq \mathbb{R}^n$ is defined as

$$\boxed{\dim_{\mathcal{H}}(A) := \inf\{s \geq 0 | \mathcal{H}^s(A) = 0\}}$$

Equivalently, we could have defined it as

$$\dim_{\mathcal{H}}(A) := \sup\{t \geq 0 | \mathcal{H}^t(A) = \infty\}$$

Example 1.8.9. Let $Q = [-1, 1]^n \subseteq \mathbb{R}^n$. Then

$$2^{-n}\mathcal{L}^n(Q) \leq \mathcal{H}^n(Q) \leq 2^{-n} n^{\frac{n}{2}}(Q)$$

Proof. `Missing` □

Examples of non-integer Hausdorff-dimension sets

There is a famous problem of finding out what the coast-line length of England is. The problem is that depending on how many knicks and bumps in the coast-line we count, the length gets longer and longer. Although it seems to be a bit counter-intuitive that a coast-line does not have a length, it isn't as unbelievable as it seems, because this is exactly what happened when we tried to measure $\mathcal{H}^0_\delta(\mathbb{S}^1)$. The more and more we decreased δ, the harder it became to cover all points, so in the limit $\delta \to 0$, we found that $\mathcal{H}^0(\mathbb{S}^1) = \infty$.
What this points to is that the coast-line must have some Hausdorff dimension $\dim_{\mathcal{H}}(A) > 1$.

Example 1.8.10 (Triadic Cantor Set)**.** When talking about the Lebesgue, we found that the Triadic Cantor set C was an example of an uncountable set with $\mathcal{L}^1$-measure zero (see 1.6.3)
First we note that if we strech a set $A \subseteq \mathbb{R}^n$ by some factor $\lambda > 0\mathbb{R}$, then

$$\mathcal{H}^s(\lambda \cdot A) = \lambda^s \mathcal{H}^s(A), \quad \forall s > 0$$

If we let $d := \dim_{\mathcal{H}}(C)$, then we can take two copies of $\frac{1}{3} \cdot C$ and when we put them back together, we obtain C again. So

$$\mathcal{H}^d(C) = 2\mathcal{H}^d\left(\frac{1}{3} \cdot C\right) = \frac{2}{3^d}\mathcal{H}^d(C)$$

which gives us the result

$$d = \log_3 2 = \frac{\ln 2}{\ln 3}$$

Example 1.8.11 (Cantor Dust)**.** `Missing`

Example 1.8.12 (The Koch Curve)**.** `Missing`

1.9 Radon Measures

Definition 1.9.1. A measure μ on $\mathbb{R}^n$ is called a **Radon** measure, if μ is Borel regular and $\mu(K) < \infty$ for every compact $K \subseteq \mathbb{R}^n$.

Example 1.9.2.

- $\mathcal{L}^n$ is a Radon measure on $\mathbb{R}^n$
- $\mathcal{H}^s$ for $s < n$ is not a Radon measure.
- The Dirac measure δ_0 is a Radon measure.
- If μ is Borel regular and $A \subset \mathbb{R}^n$ is μ-measurable with $\mu(A) < \infty$, then the restriction measure

  $$(\mu \llcorner A)(B) := \mu(A \cap B)$$

  is a Radon measure.

We wish to show that for Radon measure, an analogue to Theorem 1.4.9 holds. For this we need the following Lemma:

Lemma 1.9.3. Let μ be a Radon measure. For every μ-measurable set $A \subseteq \mathbb{R}^n$ it holds

$$\forall \varepsilon > 0 \ \exists U \supseteq A, U \text{ open such that } \mu(U \setminus A) < \varepsilon \tag{$*$}$$

Proof Sketch. For the full proof, see Prof. Michael Struwe's notes "Analysis III – Mass und Integral" We only show that WLOG, A is a Borel set. Let A, μ, ε as above.
Since μ is Borel-regular, there exists a Borel set $B \supseteq A$ with $\mu(B) = \mu(A)$. By μ-measurability of A:

$$\mu(A) = \mu(B) = \mu(\underbrace{B \cap A}_{=A}) + \mu(B \setminus A) \implies \mu(B \setminus A) = 0$$

Now assume that the Lemma is true for Borel sets, so there exists an open set $U \supseteq B$ with $\mu(U \setminus B) < \varepsilon$. Since Borel sets are also μ-measurable, we apply the Carathéodory criterion on the test set $U \setminus A$ to get

$$\begin{aligned}\mu(U \setminus A) &= \mu((U \setminus A) \cap B) + \mu((U \setminus A) \setminus B) \\ &= \mu(B \setminus A) + \mu(U \setminus B) < \varepsilon\end{aligned}$$

For the rest, set

$$\mathcal{G} := \{B \subseteq \mathbb{R}^n \big| B \text{ Borel}, (*) \text{ is true for } B\}$$

and show that $\mathcal{G}$ contains the Borel algebra $\mathcal{B}$. To do so, we proceed as follows

- $\mathcal{G}$ contains all the open sets.
- $\mathcal{G}$ is closed under countable inclusion.
- $\mathcal{G}$ is closed under countable intersection.
- It therefore contains all closed sets.

Then we set

$$\mathcal{F} = \{B \subseteq \mathbb{R}^n \big| B \in \mathcal{G}, \text{ or } B^c \in \mathcal{G}\}$$

Using de Morgan's rule, we see that $\mathcal{F}$ is a σ-algebra that contains all open sets. □

Theorem 1.9.4 (Approximation by open and compact sets)**.** Let μ be a Radon measure on $\mathbb{R}^n$.

- For every $A \subseteq \mathbb{R}^n$ it holds

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

- For every $A \subseteq \mathbb{R}^n$ μ-measurable it holds

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

Proof.

(a) Suppose $\mu(A) < \infty$ (or else take $U = \mathbb{R}^n$).

We first show it assuming that A is μ-measurable. Since for all $\varepsilon > 0$ there exists an open set $U \supseteq A$ with $\mu(U \setminus A) < \varepsilon$. By μ-measurablility of A, we have

$$\mu(U) = \mu(U \cap A) + \mu(U \setminus A) = \mu(A) + \varepsilon$$

Now let A be an arbitrary set. Since μ is Borel regular, there exists a Borel set $B \supseteq A$ with $\mu(A) = \mu(B)$. Then

$$\begin{aligned}\mu(A) = \mu(B) &= \inf\{\mu(U), B \subseteq U \text{ open}\} \\ &\geq \inf\{\mu(U), A \subseteq U \text{ open}\}\end{aligned}$$

(b) Let A be μ-measurable. We consider two cases:

$\mu(A) < \infty$: Set $\nu := \mu \llcorner A$, which is also a Radon measure. By applying (a) on the set $\mathbb{R}^n \setminus A$, for all $\varepsilon > 0$ there exists an open set U with $\mathbb{R}^n \setminus A \subseteq U$ and

$$\nu(U) \leq^{4} \nu((\mathbb{R}^n \setminus A) \cap U) + \nu(\mathbb{R}^n \setminus U) = \mu((\mathbb{R}^n \setminus A) \cap A) + \varepsilon = \varepsilon$$

Then the set $C := \mathbb{R}^n \setminus U$ is closed and is contained in A and

$$\mu(A \setminus C) = \mu(A \cap (\mathbb{R}^n \setminus C)) = \nu(\mathbb{R}^n \setminus C) = \nu(G) < \varepsilon$$

Which gives $\mu(A) \leq^4 \mu(C) + \varepsilon$ and therefore

$$\mu(A) = \sup\{\mu(C) | C \subset A, C \text{ closed}\}$$

Notice that for any closed set C we can take the sequence of compact sets

$$C_m := C \cap \overline{B}(0,m) \implies \bigcup_{m \in \mathbb{N}} C_m = C$$

and by continuity from below for μ-measurable subsets, we also have $\mu(C) = \lim_{m\to\infty} \mu(C_m)$, which means

$$\forall \varepsilon > 0 \ \exists m_0 : m \geq m_0 \implies \mu(C) - \mu(C_m) < \varepsilon$$

And therefore

$$\sup\{\mu(K) | K \subseteq A, K \text{ compact}\} = \sup\{\mu(C) | C \subseteq A, C \text{ closed }\} = \mu(A)$$

$\mu(A) = \infty$: In ths case, set $D_k := \{x | k-1 \leq |x| < k\}$. These disjoint sets can be written as the union of a closed an an open set, and are thus Borel. Moreover, $A = \bigcup_{k=1}^{\infty} (D_k \cap A)$.
Because $D_k \cap A \subseteq \overline{D_k} \cap A$ and μ is Radon, $\mu(D_k \cap A) < \infty$.
But then we are in the first case, so there exists a closed set $C_k \subseteq D_k \cap A$ with

$$\mu(D_k \cap A) - \mu(C_k) \leq \frac{1}{2^k}$$

Because $(\bigcup_{k=1}^{m} C_k)_{m \in \mathbb{N}}$ is an increasing sequence, we can use continuity from below and the fact that measures are σ-additive on the σ-algebra of μ-measurable sets

$$\begin{aligned} \lim_{m\to\infty} \mu\left(\bigcup_{k=1}^{m} C_k\right) &= \mu\left(\bigcup_{k=1}^{\infty} C_k\right) = \sum_{k=1}^{\infty} \mu(C_k) \\ &\geq \sum_{k=1}^{\infty} \mu(D_k \cap A) - \frac{1}{2^k} \\ &= \mu(A) - 1 = \infty \end{aligned}$$

This shows that

$$\sup\{\mu(C) | C \subseteq A \text{ closed } |\} = \infty = \mu(A)$$

and by a similar argument as in the first case (writing $\mu(C) = \lim_{m\to\infty} \mu(C \cap \overline{B}(0,m))$) we get

$$\mu(A) = \sup\{\mu(K) | K \subseteq A \text{ compact}\}$$

$\square$

[4]Since $\mathbb{R}^n \setminus A$ and C are also μ-measurable, we could also have used equality here.

2 Measurable Functions

2.1 Basic definitions

For X, Y nonempty sets and $f : X \to Y$ with $A \subseteq Y$, the inverse image is defined as

$$f^{-1}(A) = \{x \in X | f(x) \in A\}$$

And it's easy to show that

(a) $f^{-1}(A^c) = \left(f^{-1}(A)\right)^c$

(b) For a sequence of subsets $(A_k)_k$ the following holds

$$f^{-1}\left(\bigcup_{k=1}^{\infty} A_k\right) = \bigcup_{k=1}^{\infty} f^{-1}(A_k)$$

(c) The analogue for countable intersections follows easily from de Morgan's rule, (a) and (b)

In particular, if $\mathcal{A} \subseteq \mathcal{P}(Y)$ is a σ-algebra, then

$$\Sigma := f^{-1}(\mathcal{A}) := \{f^{-1}(A) | A \in \mathcal{A}\}$$

is a σ-algebra in X.
In the following, let μ be a measure on $\mathbb{R}^n$ and $\Omega \subseteq \mathbb{R}^n$ be a μ-measurable subset.

Definition 2.1.1. A function $f : \Omega \to [-\infty, \infty]$ is called $\boldsymbol{\mu}$**-measurable** if in the sense of definition 1.2.7

(a) $f^{-1}\{+\infty\}, f^{-1}\{-\infty\}$ are μ-measurable.

(b) $f^{-1}(U)$ for every $U \subseteq \mathbb{R}$ open is μ-measurable.

Remark 2.1.2. The following two conditions are equivalent to (b)

(c) $f^{-1}(B)$ is μ-measurable for each Borel set $B \subseteq \mathbb{R}$

(d) $f^{-1}((-\infty, a))$ is μ-measurable for all $a \in \mathbb{R}$.

And if we consider $\overline{\mathbb{R}} = [-\infty, \infty]$ with the topology generated by the open sets of $\mathbb{R}$ and the neighborhoods $[-\infty, a), (a, \infty), a \in \mathbb{R}$ of $\pm\infty$, then for a function $f : \Omega \to [-\infty, \infty]$, the following are equivalent:

- f is μ-measurable.
- $f^{-1}(U)$ is μ-measurable, $\forall U \subseteq \overline{\mathbb{R}}$ open
- $f^{-1}([-\infty, a))$ is μ-measurable, $\forall a \in \mathbb{R}$.

Remark 2.1.3. Even preimages of Borel sets are μ-measurable!
By the properties discussed in the begnning of this chapter, we know that the inverse image of a Borel set can be written as some combination of complements, unions and intersections of preimages of open sets. And since μ-measurable sets form a σ-algebra (see Theorem 1.2.9), they are also μ-measurable.

Example 2.1.4. Let $f : \Omega \to \mathbb{R}$ be μ-measurable and $g : \mathbb{R} \to \mathbb{R}$ continuous. Then $g \circ f$ is μ-measurable.

Theorem 2.1.5.

(a) Let $f, g : \Omega \to \mathbb{R}$ be μ-measurable functions. Then: $f+g, f\cdot g, |f|, \operatorname{sgn}(f), \max\{f,g\}, \min\{f,g\}$ and (if g is never zero) $\frac{f}{g}$ are μ-measurable, where

$$(\operatorname{sgn} f)(x) := \begin{cases} \frac{f(x)}{|f(x)|} & \text{if } f(x) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) For a sequence of μ-measurable functions $(f_k : \Omega \to \overline{\mathbb{R}})_{k\in\mathbb{N}}$ the following are also μ-measurable

$$\inf_{k\in\mathbb{N}} f_k, \quad \sup_{k\in\mathbb{N}} f_k, \quad \liminf_{k\to\infty} f_k, \quad \limsup_{k\to\infty} f_k$$

Proof. `Missing` □

Recall that when defining the Riemann integral in Analysis I/II, we started by defining the Integral of *step functions* "by hand", which were functions that were constant when decomposing them into intervals. To define the Riemann integral of general types of functions, we defined the "Ober- und Untersummen", and if they coincided, took that as the value for the integral.
Our approach will be slightly more general

Definition 2.1.6. Given a subset $A \subseteq \mathbb{R}^n$, we define the **characteristic function** of the set A as

$$\chi_A : \mathbb{R}^n \to \mathbb{R}, \quad \chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

It's easy to see that χ_A is μ-measurable if and only if A is μ-measurable.
A **simple function** is a function of the form

$$f(x) = \sum_{i=1}^{\infty} c_i \chi_{A_i}(x), \quad c_i \in \mathbb{R}, A_i \subseteq \mathbb{R}^n, A_i \text{ mutually disjoint}$$

And if the A_i are μ-measurable, then f is called a $\boldsymbol{\mu}$**-measurable simple function**.
Equivalently, a function $f : \mathbb{R}^n \to \mathbb{R}$ is a μ-measurable simple function if and only if f is μ-measurable and the image of f is a countable subset of $\mathbb{R}$.

The following theorem lets us decompose any non-negative μ-measurable function into a simple function.

Theorem 2.1.7. Let $f : \Omega \to [0, \infty]$ be μ-measurable. Then there exist μ-measurable sets $A_k \subseteq \Omega$ such that

$$f = \sum_{k=1}^{\infty} \frac{1}{k} \chi_{A_k}$$

Proof. We define the sets A_k inductively, starting with

$$A_1 = \{x \in \Omega \big| f(x) \geq 1\} = f^{-1}[1, \infty]$$

which is μ-measurable. Then for all $k = 2, 3, \ldots$, we define

$$A_k = \{x \in \Omega \big| \frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j} \chi_{A_j}\}$$

To show that this produces the function f, we show both inequalities in

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k} \chi_{A_k}(x)$$

$\geq$: If $\sup\{k|x \in A_k\} = \infty$, then

`Missing Rest` □

Proposition 2.1.8. Let $f : \Omega \to \mathbb{R}$ be continuous and μ a Borel measure. Then f is μ-measurable.

Proof. For any open set $U \subseteq \mathbb{R}^n$, $f^{-1}(U) = O \cap \Omega$ for some open set $O \subseteq \mathbb{R}^n$.
Since μ is Borel, $f^{-1}(U)$ is μ-measurable. □

From now on, we will say that a statement holds "μ-a.e." or "**almost everywhere with respect to** μ" to mean that the set A, for which the statement does not hold, has $\mu(A) = 0$.

2.2 Lusin's and Egoroff's Theorems

Consider for example the sequence of ($\mathcal{L}^1$-measurable) functions $(f_k : [0,1] \to \mathbb{R})_{k\in\mathbb{N}}$ given by

$$f_k = \chi_{A_k}, \quad \text{for} \quad A_k = \left(1 - \frac{1}{2^k}, 1\right)$$

this sequence converges pointwise to the constant function $\mathbf{0}$, but not uniformly.
However we can say that for any $\delta > 0$, it converges uniformly on the (compact) subset $A := [0, 1 - \delta]$, which satisfies $\mathcal{L}^1([0,1] \setminus A) = \delta$.

Definition 2.2.1. Let $\Omega \subseteq \mathbb{R}^n$. We say that a sequence of functions $(f_k : \Omega \to \overline{\mathbb{R}})_{k\in\mathbb{N}}$ **converges μ-almost uniformly** on Ω to a function $f : \Omega \to \overline{\mathbb{R}}$, if for all $\delta > 0$ there exists a μ-measurable subset $A \subseteq \Omega$ with $\mu(\Omega \setminus A) < \delta$ such that $(f_k)_{k\in\mathbb{N}}$ converges uniformly on A. That is:

$$\sup_{x\in A}|f_k(x) - f(x)| \to 0 \text{ as } k \to \infty$$

Theorem 2.2.2 (Egoroff)**.** Let $\Omega \subseteq \mathbb{R}^n$ be μ-measurable with $\mu(\Omega) < \infty$, and let $f, (f_k)_{k\in\mathbb{N}} : \Omega \to \overline{\mathbb{R}}$ μ-measurable.

(a) If $f_k(x) \to f(x)$ as $k \to \infty$ for μ-a.e. $x \in \Omega$, and $f(x)$ finite μ-a.e., then $(f_k)_{k\in\mathbb{N}}$ converges μ-almost uniformally to f on Ω.

(b) If additionally, μ is a Radon measure we can also assume that the set on which $(f_k)_{k\in\mathbb{N}}$ converges uniformly is compact. That is:
$\forall \delta > 0\ \exists K \subseteq \Omega$ compact with $\mu(\Omega \setminus K) < \delta$ and

$$\sup_{x\in C}|f_k(x) - f(x)| \to 0 \text{ as } k \to \infty$$

Proof. (a) Let $\delta > 0$ and for $i, j \in \mathbb{N}$ define the sets

$$C_{ij} := \bigcup_{k=j}^{\infty} \left\{x \in \Omega \middle| |f_k(x) - f(x)| > \frac{1}{2^i}\right\}$$

which are μ-measurable, since they are the pre-images of open subsets under a μ-measurable function. They are also decreasing in j (i.e. $C_{i,j+1} \subseteq C_{i,j} \forall i, j$) and $\mu(C_{i,1}) \leq \mu(\Omega) < \infty$, so we can use continuity from above. Note how $\bigcap_{j=1}^{\infty} C_{ij}$ consists of points in Ω, on which $(f_k)_{k\in\mathbb{N}}$ does not converge to f. But as $f_k(x) \to f(x)$ for μ-a.e. x we get

$$\lim_{j\to\infty} \mu(C_{ij}) \overset{\text{c.f.a.}}{=} \mu\left(\bigcap_{j=1}^{\infty} C_{ij}\right) = 0, \quad \forall i \in \mathbb{N}$$

This means that for every i, there exists an $N(i) > 0$ such that

$$\mu\left(\bigcap_{j=1}^{N(i)} C_{i,j}\right) = \mu(C_{i,N(i)}) < \frac{\delta}{2^i}$$

setting

$$A := \Omega \setminus \bigcup_{i=1}^{\infty} C_{i,N(i)}$$

which is μ-measurable, we get the estimate

$$\mu(\Omega \setminus A) = \mu\left(\bigcup_{i=1}^{\infty} C_{i,N(i)}\right) < \sum_{i=1}^{\infty} \frac{\delta}{2^i} = \delta$$

Moreover, for all $x \in A, i \in \mathbb{N}, k \geq N(i)$ we have

$$|f_k(x) - f(x)| \leq \frac{1}{2^i}$$

which shows uniform convergence of $f_k \to f$ on A.

(b) By applying Theorem 1.9.4, there exists a compact subset $K \subseteq A$ such that $\mu(A \setminus K) < \delta$, and so

$$\mu(\Omega \setminus K) \leq \mu(\Omega \setminus A) + \mu(A \setminus K) \leq 2\delta$$

□

Remark 2.2.3. The condition that $\mu(\Omega) < \infty$ is necessary. Take for example the sequence of "bump functions moving to the right": $f_k = \chi_{[k,k+1]} : \mathbb{R} \to \mathbb{R}$, which converges pointwise to $\mathbf{0}$.
It is clear that any set $A \subseteq \mathbb{R}$ on which the sequence converges uniformly must be bounded from the right, which means that A cannot satisfy $\mu(\mathbb{R} \setminus A) < \delta$.

In Analysis I/II, we proved that continous functions are Riemann integrable and that Riemann-integrable functions only have measure-zero points of discontinuity.
Because μ is assumed to be a Radon measure, all continuous functions are μ-measurable. We now show the generalisation of the other fact.

Theorem 2.2.4 (Lusin's Theorem)**.** Let μ be a Radon measure on $\mathbb{R}^n$, $\Omega \subseteq \mathbb{R}^n$ be μ-measurable with $\mu(\Omega) < \infty$ and $f : \Omega \to \overline{\mathbb{R}}$ μ-measurable and finite μ.a.e..
Then $\forall \varepsilon > 0\ \exists K \subseteq \Omega$ compact with $\mu(\Omega \setminus K) < \varepsilon$ such that $f|_K$ is continuous.

Remark 2.2.5. Warning: The theorem states that the function $f|_K : K \to \overline{\mathbb{R}}$ is continous and *not* that f is continous at x for all $x \in K$.
If we drop the condition $\mu(\Omega) < \infty$, then we can still find such a set $C \subseteq \Omega$, which is closed, but not necessarily compact.

Proof. For each $i \in \mathbb{N}_{>0}$, let $\{B_{ij}\}_{j\in\mathbb{N}}$ be a collection of disjoint Borel sets such that

$$\mathbb{R} = \bigsqcup_{j=1}^{\infty} B_{ij} \quad \text{and} \quad \operatorname{diam}(B_{ij}) := \sup\{|x-y| | x, y \in B_{ij}\} < \frac{1}{i}$$

Then define $A_{ij} := f^{-1}(B_{ij})$ which are μ-measurable and let

$$\tilde{\Omega} := \bigcup_{j=1}^{\infty} A_{ij} \implies \Omega = \tilde{\Omega} \sqcup f^{-1}\{\pm\infty\}$$

since μ is a Radon measure, there exist compact sets $K_{ij} \subseteq A_{ij}$ such that

$$\mu(A_{ij} \setminus K_{ij}) < \frac{\varepsilon}{2^{i+j}}$$

Then

$$\begin{aligned}
\mu\left(\tilde{\Omega} \setminus \bigcup_{j=1}^{\infty} K_{ij}\right) &= \mu\left(\bigcup_{j=1}^{\infty} A_{ij} \setminus \bigcup_{j=1}^{\infty} K_{ij}\right) \\
&\leq \mu\left(\bigcup_{j=1}^{\infty} (A_{ij} \setminus K_{ij})\right) \\
&\leq \sum_{j=1}^{\infty} \mu(A_{ij} \setminus K_{ij}) \\
&< \sum_{j=1}^{\infty} \frac{\varepsilon}{2^{i+j}} = \frac{\varepsilon}{2^i}
\end{aligned}$$

This means that for all i, there exists a $N(i)$ such that

$$\mu\left(\tilde{\Omega} \setminus \bigcup_{j=1}^{N(i)} K_{ij} < \frac{\varepsilon}{2^i}\right)$$

Then define the compact sets

$$D_i := \bigcup_{j=1}^{N(i)} K_{ij} \quad \text{and} \quad K := \bigcap_{i=1}^{\infty} D_i$$

For each i, j chose some $b_{ij} \in B_{ij}$ and define $g_i : D_i \to \mathbb{R}, g_i(x) = b_{ij}$ if $x \in K_{ij}$ for all $j \leq N(i)$.
Since the sets $\{K_{ij}\}_{j\in\mathbb{N}}$ are compact disjoint sets, this means that g_i is locally constant and thus continuous.
Moreover, by construction of B_{ij}, we have

$$\operatorname{diam}(B_{ij}) < \frac{1}{i} \implies |f(x) - g_i(x)| < \frac{1}{i}, \quad \forall x \in D_i$$

which means that the sequence of continuous functions $(g_i|_K)_{i\in\mathbb{N}^{>0}} : K \to \mathbb{R}$ converges uniformly to a continous function $f|_K$.
The set K also satisifes

$$\mu(\tilde{\Omega} \setminus K) = \mu\left(\bigcup_{i=1}^{\infty} (\tilde{\Omega} \setminus D_i)\right) \leq \sum_{i=1}^{\infty} \mu(\Omega \tilde{\setminus} D_i) < \varepsilon$$

and since $f(x)$ is finite μ-a.e., we have

$$\mu(\Omega \setminus K) \leq \mu(\tilde{\Omega} \setminus K) + \mu(f^{-1}\{\pm\infty\} \setminus K) \leq \varepsilon + 0$$

□

2.3 Convergence in Measure

For this section, let μ be an arbitrary measure on $\mathbb{R}^n$ and $\Omega \subseteq \mathbb{R}^n$ μ-measurable and let $f, f_k : \Omega \to \overline{\mathbb{R}}$ be μ-measurable and $|f(x)| < \infty$ μ-a.e..

Definition 2.3.1. We say that the sequence $(f_k)_{k\in\mathbb{N}}$ converges **in measure** μ to f (written $f_k \overset{\mu}{\to} f$ as $k \to \infty$)[4], if $\forall \varepsilon > 0$

$$\boxed{\lim_{k\to\infty} \mu\left(\{x \in \Omega \mid |f(x) - f_k(x)| > \varepsilon\}\right) = 0}$$

Remark 2.3.2. If the sequence converges uniformly, then also $f_k \overset{\mu}{\to} f$. However pointwise convergence is not enough to show $f_k \overset{\mu}{\to} f$, as the same Example as in 2.2.3 proves otherwise.
Also, the sequence $f_k = \chi_{\{0\}} \overset{\mathcal{L}^1}{\to} \mathbf{0}$ shows that pointwise convergence does not necessarily follow from in-measure convergence.

Theorem 2.3.3. Let $\mu(\Omega) < \infty$. If $f_k \to f$ μ-a.e. then $f_k \overset{\mu}{\to} f$.

Proof. By Egoroff's theorem, $(f_k)_{k\in\mathbb{N}}$ converges μ-almost uniformally on Ω.
This means that for all $\varepsilon > 0$ there exists a set A with $\mu(\Omega \setminus A) < \delta$ such that for all $\varepsilon > 0 : \exists N \in \mathbb{N}$ with

$$\sup_{x\in A} |f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N$$

for such $n \geq N$ we have

$$\{x \in \Omega \mid |f_n(x) - f(x)| > \varepsilon\} \subseteq \Omega \setminus A$$

Taking $\mu(\cdot)$ on both sides gives the result. $\square$

Remark 2.3.4. The converse of this theorem does not hold, so the statement $f_k \to f$ μ-a.e. is stronger than $f_k \overset{\mu}{\to} f$.
To see this, take $\Omega = [0, 1)$ with the measure $\mathcal{L}^1$. And set $f_k = \chi_{A_k}$, for

$$A_1 = [0, 1), \quad A_2 = [0, \tfrac{1}{2}), A_3 = [\tfrac{1}{2}, 1), \quad A_4 = [0, \tfrac{1}{4}), A_5 = [\tfrac{1}{4}, \tfrac{2}{4}), \ldots, A_7 = [\tfrac{3}{4}, 1), \quad A_8 = [0, \tfrac{1}{8}), \ldots$$

and more generally for $k \geq 1$, chose n such that $2^n \leq k < 2^{n+1}$ and set

$$A_k = \left[\frac{k - 2^n}{2^n}, \frac{k - 2^n + 1}{2^n}\right)$$

Therefore

$$\mu\left(\{|f_k(x)| > 0\}\right) = \mu(A_k) = \frac{1}{2^n} < \frac{2}{k} \implies f_k \overset{\mu}{\to} \mathbf{0}$$

But the nowhere does the sequence converge pointwise to 0.

Theorem 2.3.5. Let $f_k \overset{\mu}{\to} f$. Then there exists a subsequence $(f_{k_n})_{n\in\mathbb{N}}$ that converges to f μ-a.e..

[4] In contrast to the lecturer, I will be using $f_k \overset{\mu}{\to} f$ as shorthand for "$f_k \overset{\mu}{\to} f$ as $k \to \infty$" as it should be clear from context.

Proof. Since $f_k \xrightarrow{\mu} f$, for all $n \in \mathbb{N}$ there exists a $k_n \in \mathbb{N}$ such that

$$\mu\left(\{x \in \Omega \mid |f_k(x) - f(x)| > 2^{-n}\}\right) < 2^{-n}, \quad \forall k \geq k_n$$

Define for $h \geq 1$

$$A_n := \{x \in \Omega \mid |f_{k_n}(x) - f(x)| > 2^{-n}\} \quad \text{and} \quad E_h := \bigcup_{n \geq h} A_n$$

by subadditivity, we have

$$\mu(E_h) \leq \sum_{n=h}^{\infty} \mu(A_n) < 2^{-h+1}$$

For any $x \in \Omega \setminus E_h$ we have that

$$\forall n \geq h : x \notin A_n, \implies \forall n \geq h : |f_{k_n}(x) - f(x)| \leq 2^{-n}$$

This means that for all $h \in \mathbb{N}$ the sequence $(f_{k_n})_{n\in\mathbb{N}}$ converges to $f(x)$ on $\Omega \setminus E_h$.
Because $\mu(E_1) \leq \mu(\Omega) < \infty$ and the sequence $(E_h)_{h\in\mathbb{N}}$ is decreasing, we have by continuity from above:

$$E := \bigcap_{h=1}^{\infty} E_h \implies \mu(E) = \lim_{h\to\infty} \mu(E_h) = 0$$

Since $(f_{k_n})_{n\in\mathbb{N}}$ converges to f on $\Omega \setminus E_h$, the sequence $(f_{k_n}|_E)_{n\in\mathbb{N}}$ converges to f on $\Omega \setminus E$.

□

3 Integration

Assume for this chapter, that μ is a Radon measure on $\mathbb{R}^n$ and $\Omega \subseteq \mathbb{R}^n$ is μ-measurable.

3.1 Definitions and Basic Properties

Definition 3.1.1. A function $g : \Omega \to \overline{\mathbb{R}}$ is called a **simple function**, if the image of g is at most countable.

Because summing up series with both positive and negative coefficients can lead to some convergence issues (for example the sequence $a_n = (-1)^n$), we split a function into its positive and negative parts. Define

$$f^+ := \max(f, 0), \quad f^- := \max(-f, 0)$$
$$\implies f = f^+ - f^-, \quad |f| = f^+ + f^-$$

Definition 3.1.2.

(a) For $g : \Omega \to [0, \infty]$ is a **non-negative, simple, μ-measurable** function, we define

$$\boxed{\int_\Omega g d\mu := \sum_{0\leq y\leq\infty} y \cdot \mu\left(g^{-1}\{y\}\right)}$$

where we use the convention $0 \cdot \infty = 0$. (We want the integral of the function **0** to be zero.)

(b) A simple, μ-measurable function $g : \Omega \to [-\infty, +\infty]$ is called a **μ-integrable simple function**, if either $\int_\Omega g^+ d\mu < \infty$ or $\int_\Omega g^- d\mu < \infty$. Then define

$$\int_\Omega g d\mu := \int_\Omega g^+ d\mu - \int_\Omega g^- d\mu = \sum_{-\infty \leq y \leq \infty} y \cdot \mu\left(g^{-1}\{y\}\right)$$

(c) For any function $f : \Omega \to [-\infty, \infty]$ define the **upper integral**

$$\overline{\int_\Omega} f d\mu := \inf\left\{\int_\Omega g d\mu \middle| g \geq f \mu\text{-a.e., } g \text{ is a } \mu\text{-integrable simple function}\right\}$$

aswell as the **lower integral**

$$\underline{\int_\Omega} f d\mu := \sup\left\{\int_\Omega e d\mu \middle| e \leq f \mu\text{-a.e., } e \text{ is a } \mu\text{-integrable simple function}\right\}$$

(d) A μ-measurable function $f : \Omega \to \overline{\mathbb{R}}$ is called **μ-integrable**, if the upper- and lower integral coincide, in which case we write

$$\int_\Omega f d\mu := \overline{\int_\Omega} f d\mu = \underline{\int_\Omega} f d\mu$$

Warning: in some contexts, functions with integral $\pm\infty$ are called "not integrable". For our purposes, they are.

Remark 3.1.3. It is easy to show that

$$\underline{\int_\Omega} f d\mu \leq \overline{\int_\Omega} f d\mu$$

In Exercise Sheet 09, we prove that the multiple definitions of an integral are "consistent", that is: for a μ-integrable simple fucntion we have

$$\underline{\int_\Omega} f d\mu = \overline{\int_\Omega} f d\mu = \int_\Omega f d\mu$$

where the last integral is understood as the definition used in (b)

Proposition 3.1.4. Let μ be Radon and let $f : \Omega \to [0, \infty]$ be μ-measurable. Then f is μ-integrable.

Proof. If $\underline{\int_\Omega} f d\mu = \infty$, then it is trivial. Now, if $\underline{\int_\Omega} f d\mu < \infty$, it means that $f(x) < \infty$ μ-a.e.

- **Case $\mu(\Omega) < \infty$:** For all $\varepsilon > 0$, set for $k \in \mathbb{N}$

$$A_k := \{x \in \Omega \big| k\varepsilon \leq f(x) < (k+1)\varepsilon\} = f^{-1}\left[k\varepsilon, (k+1)\varepsilon\right)$$

  Since $f(x) < \infty$ μ-a.e. it means that for $\tilde{\Omega} := \bigcup_{k\in\mathbb{N}} A_k = f^{-1}[0, \infty)$, we have $\mu(\Omega \setminus \tilde{\Omega}) = 0$.

To sandwich f between μ-integrable simple functions, we define

$$e(x) := \varepsilon \sum_{k=0}^{\infty} k\chi_{A_k}(x)$$
$$g(x) := \varepsilon \sum_{k=0}^{\infty} (k+1)\chi_{A_k}(x)$$

which gives us $e(x) \leq f(x) < g(x)$ μ-a.e. and

$$\int_\Omega e d\mu \leq \underline{\int_\Omega} f d\mu \leq \overline{\int_\Omega} f d\mu \leq \int_\Omega g d\mu = \varepsilon \sum_{k\in\mathbb{N}} (k+1)\mu(A_k)$$
$$= \varepsilon \sum_{k\in\mathbb{N}} k\mu(A_k) + \varepsilon \sum_{k\in\mathbb{N}} \mu(A_k) = \int_\Omega e d\mu + \varepsilon \cdot \mu(\tilde{\Omega})$$

where in the last step we used that the A_k were mutually disjoint and were preimages of Borel sets and thus μ-measurable. Because $\mu(\tilde{\Omega}) \leq \mu(\Omega) < \infty$, we can let $\varepsilon \to 0$ and get the result.

- **General Case:** Let $\Omega \subseteq \mathbb{R}^n$ be a μ-measurable set. Then take any countable covering of $\mathbb{R}^n$ with disjoint dyadic cubes $R^n = \bigcup_{l=1}^{\infty} Q_l$ and set $\Omega_l := \Omega \cap Q_l$.

  Since $\mu(\Omega_l) < \infty$, we are in the first case, so for all $\varepsilon > 0$ we can find μ-integrable simple functions $e_l, g_l : \Omega_l \to [0, \infty]$ with $e_l \leq f \leq g_l$ μ-a.e. and

  $$\int_{\Omega_l} e_l d\mu \leq \int_{\Omega_l} g_l d\mu \leq \int_{\Omega_l} e_l d\mu + \frac{\varepsilon}{2^l}$$

  We then define [5]

  $$e := \sum_{l=1}^{\infty} e_l \cdot \chi_{\Omega_l}, \quad \text{and} \quad g := \sum_{l=1}^{\infty} g_l \cdot \chi_{\Omega_l}$$

  which are again μ-integrable simple functions satisfying $e \leq f \leq g$ for μ-a.e. $x \in \Omega$ and

  $$\int_\Omega e d\mu \leq \underline{\int_\Omega} f d\mu \leq \overline{\int_\Omega} f d\mu \leq \int_\Omega g d\mu = \sum_{l=1}^{\infty} \int_{\Omega_l} g_l d\mu \leq \sum_{l=1}^{\infty} \int_{\Omega_l} e_l d\mu + \sum_{l=1}^{\infty} \frac{\varepsilon}{2^l} = \int_\Omega e d\mu + \varepsilon$$

  letting $\varepsilon \to 0$, we get the result.

$\square$

Proposition 3.1.5 (Monotonicity)**.** Let $f_1, f_2 : \Omega \to \overline{\mathbb{R}}$ be μ-integrable with $f_1 \leq f_2$ μ-a.e.. Then

$$\int_\Omega f_1 d\mu \leq \int_\Omega f_2 d\mu$$

Proof. If a μ-integrable simple function g satisfies $g \geq f_2$ μ-a.e., then it also satisfies $g \geq f_1$ μ-a.e.. Looking at the definition of μ-integrable functions,

$$\int f d\mu = \inf \left\{ \int_\Omega g d\mu | g \geq f \mu\text{-a.e., } g \text{ is a } \mu\text{-integrable simple function} \right\}$$

then we are taking the infimum over a larger set for f_1 than for f_2, so

$$\int_\Omega f_1 d\mu = \overline{\int_\Omega} f_1 d\mu \leq \overline{\int_\Omega} f_2 d\mu = \int_\Omega f_2 d\mu$$

$\square$

[5] There is some minor abuse of notation but it's easy to extend the domain of e_l, g_l to Ω.

As an immediate consequence, we have

Corollary 3.1.6. Let $f_1, f_2 : \Omega \to \overline{\mathbb{R}}$ μ-integrable with $f_1 = f_2$ μ-a.e.

$$\int_\Omega f_1 d\mu = \int_\Omega f_2 d\mu$$

Definition 3.1.7. Let $f : \Omega \to \overline{\mathbb{R}}$ be a function

- f is called **μ-summable**, if f is μ-measurable and

$$\int_\Omega |f| d\mu < \infty$$

- f is called **locally μ-summable** in Ω, if for all compact sets $K \subseteq \Omega$, $f|_K$ is μ-summable.

Proposition 3.1.8. Let $f : \Omega \to \overline{\mathbb{R}}$.

- If f is μ-summable, then it is μ-integrable.
- If $f(x) = 0$ μ-a.e., then f is μ-integrable and $\int_\Omega f d\mu = 0$.

Proof.

(a) We show that $\underline{\int}_\Omega f d\mu = \overline{\int}_\Omega f d\mu$. Since f is μ-measurable, $f^\pm = \max(\pm f, 0)$ are also μ-measurable. Moreover, since $0 \leq f^\pm \leq |f|$, by Proposition 3.1.4, they are μ-integrable and satisfy $\int_\Omega f^\pm d\mu < \infty$. This means that for all $\varepsilon > 0$, there exist μ-integrable simple functions $e_\pm \leq f^\pm \leq g_\pm$ μ-a.e. with

$$\int_\Omega e_\pm d\mu \leq \int_\Omega f^\pm d\mu \leq \int_\Omega g_\pm d\mu \leq \int_\Omega e_\pm d\mu + \frac{\varepsilon}{2}$$

Setting $e := e_+ - g_-$ and $g := g_+ - e_-$, we see that again $e \leq f \leq g$ μ-a.e. and

$$\begin{aligned} \int_\Omega e d\mu \leq \underline{\int}_\Omega f d\mu \leq \overline{\int}_\Omega f d\mu \leq \int_\Omega g d\mu = \int_\Omega g_+ d\mu - \int_\Omega e_- d\mu \\ \leq \int_\Omega e_+ d\mu + \frac{\varepsilon}{2} + \int_\Omega g_- d\mu + \frac{\varepsilon}{2} = \int_\Omega e d\mu + \varepsilon \end{aligned}$$

Letting $\varepsilon \to 0$ shows that f is μ-integrable.

(b) If $f(x) = 0$ μ-a.e. we can set $e = g = \mathbf{0}$ and see that $e \leq f \leq g$ μ-a.e., which shows

$$0 = \int_\Omega e d\mu \leq \underline{\int}_\Omega f d\mu \leq \overline{\int}_\Omega f d\mu \leq \int_\Omega g d\mu = 0$$

□

Proposition 3.1.9. Let $f : \Omega \to [0, \infty]$ be μ-measurable.

(a) $\int_\Omega f d\mu = 0 \implies f(x) = 0$ μ-a.e.

(b) $\int_\Omega f d\mu < \infty \implies f(x) < \infty$ μ-a.e.

Proof. (a) Contraposition: Assume $f(x)$ is not μ-a.e. zero. Then the μ-measurable sets

$$A_k = \left\{x \in \Omega \middle| f(x) \geq \frac{1}{k}\right\}, \quad k \geq 1$$

form an increasing sequence and their union

$$A_\infty := \bigcup_{k=1}^{\infty} A_k = \{x \in \Omega | f(x) > 0\}$$

has non-zero measure $0 < \mu(A_\infty) = \lim_{k\to\infty} \mu(A_k)$.

This means that there exists some $K \geq 1$ such that $\mu(A_K) > 0$. Then the simple function $s := \frac{1}{K}\chi_{A_K}$ satisfies $s \leq f$, which implies

$$0 < \frac{1}{K}\mu(A_K) = \int_\Omega s d\mu \leq \int_\Omega f d\mu$$

(b) Contraposition: Assume there exists an $A \subseteq \Omega$ with $f(x) = \infty, \forall x \in A, \mu(A) > 0$ then the simple function $s := \infty \cdot \chi_A$ satisfies $s \leq f$, and thus

$$\int_\Omega f d\mu \geq \int_\Omega s d\mu = \infty \cdot \mu(A) = \infty$$

□

3.2 More Properties of the Integral

Theorem 3.2.1 (Tchebyshev Inequality)**.** Let $f : \Omega \to \overline{\mathbb{R}}$ be μ-summable. Then for every $a > 0$:

$$\mu\left(\{x \in \Omega | |f(x)| > a\} \leq \frac{1}{a}\int_\Omega |f| d\mu\right)$$

Proof. Apply monotonicity (Proposition 3.1.5) on the functions

$$f_1 = a \cdot \chi_{\{x\in\Omega \big| |f(x)|>a\}} \leq f_2 = |f|$$

here, f_1 takes on the value a, whenever $|f(x)| > a$, and 0 elsewhere. □

Corollary 3.2.2. Let $f, f_k : \Omega \to \overline{\mathbb{R}}$ be μ-integrable with

$$\lim_{k\to\infty} \int_\Omega |f_k - f| d\mu = 0$$

Then $f_k \xrightarrow{\mu} f$ and there exists a subsequence $(f_{k_n})_{n\in\mathbb{N}}$ with $f_{k_n} \to f$ μ-a.e.

Proof. Applying Tchebyshev's inequality on the function $f_k - f$, it means that for all $\varepsilon > 0$

$$\mu\left(\{x \in \Omega | |f_k - f| > \varepsilon\}\right) \leq \frac{1}{\varepsilon}\int_\Omega |f_k - f| d\mu$$

since the right hand side converges to 0 as $k \to \infty$, it follows that $f_k \xrightarrow{\mu} f$ (as in Definition 2.3.1). The second part follows from Theorem 2.3.5. □

Lemma 3.2.3. This Lemma is taken from Exercise Sheet 09.

(a) For two μ-integrable simple functions f, g, there exist sequences $\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}} \subseteq \overline{\mathbb{R}}$, with μ-measurable, mutually disjoint sets $(C_n)_{n\in\mathbb{N}}$ such that

$$f = \sum_{n\in\mathbb{N}} a_n \chi_{C_n}, \quad g = \sum_{n\in\mathbb{N}} b_n \chi_{C_n}$$

(b) For a μ-integrable (simple) function of the form $f = \sum_{n\in\mathbb{N}} a_n \chi_{A_n}$ for A_n pairwise disjoint and μ-measurable, it holds

$$\int_\Omega f d\mu = \sum_{n\in\mathbb{N}} a_n \mu(A_n)$$

(c) Let $f, g : \Omega \to [-\infty, \infty]$ be μ-summable simple functions, $a, b \in \mathbb{R}$. Then $af + bg$ is a μ-summable simple function and

$$\int_\Omega (af + bg) d\mu = a \int_\Omega f d\mu + b \int_\Omega g d\mu$$

Proof Sketch. For a detailed proof, see Master solution of Exercise Sheet 09.

(a) Since the images of f and g are countable, let $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be their values. Set $A_n := f^{-1}\{a_n\}, B_n := g^{-1}\{b_n\}$. By taking all intersections of the form $C_{ij} := A_i \cap B_j$ and reindexing them (for example, with Cantor's Diagonal map) we get a sequence $(C_m)_{m\in\mathbb{N}}$ of disjoint, μ-measurable sets.

(b) Although the values $(a_n)_{n\in\mathbb{N}}$ might not be necessarily be different, we can take the union of all A_m corresponding to a value $c_m \in \overline{\mathbb{R}}$ to write

$$f^{-1}\{c_m\} = \bigcup_{n\in\mathbb{N}: a_n = c_m} A_n$$

using the fact that the A_n are μ-measurable and disjoint and by definition of the integral for simple functions, we have

$$\int_\Omega f d\mu = \sum_{m\in\mathbb{N}} c_m \mu\left(f^{-1}\{c_m\}\right) = \sum_{m\in\mathbb{N}} c_m \mu\left(\bigcup_{n: a_n = c_m} A_n\right) = \sum_{m\in\mathbb{N}} c_m \sum_{n\in\mathbb{N}: a_n = c_m} \mu(A_n) = \sum_{n\in\mathbb{N}} a_n \mu(A_n)$$

(c) From part (a), we can find μ-measurable subsets C_n and sequences a_n, b_n such that

$$f = \sum_{n\in\mathbb{N}} a_n \chi_{C_n}, \quad g = \sum_{n\in\mathbb{N}} b_n \chi_{C_n}$$

Since the C_n are mutually disjoint, the function $af + bg$, can be written as

$$af + bg = \sum_{n\in\mathbb{N}} (a a_n + b b_n) \chi_{C_n}$$

which shows that $af + bg$ is a simple function because the C_n are disjoint. By (b), we get

$$\int_\Omega (af + bg) d\mu = \sum_{n\in\mathbb{N}} (a a_n + b b_n) \mu(C_n) = a \sum_{n\in\mathbb{N}} a_n \mu(C_n) + b \sum_{n\in\mathbb{N}} b_n \mu(C_n) = a \int_\Omega f d\mu + b \int_\Omega g d\mu$$

To show that $af + bg$ is μ-summable, apply additivity to $|f|, |g|$ and use the triangle inequality:

$$\int_\Omega |f + g| d\mu \le \int_\Omega |f| + |g| d\mu = \int_\Omega |f| d\mu + \int_\Omega |g| d\mu < \infty$$

□

Remark 3.2.4. We can actually weaken the condition on f, g. The reason we used that f, g are μ-summable in the first place was so that the right hand side of the equation

$$\int_\Omega (af + bg) d\mu = a \int_\Omega f d\mu + b \int_\Omega g d\mu$$

is well defined. This becomes more clear when we write out

$$a \int_\Omega f d\mu + b \int_\Omega g d\mu = a \left[\int_\Omega f^+ d\mu - \int_\Omega f^- d\mu \right] + b \left[\int_\Omega g^+ d\mu - \int_\Omega g^- d\mu \right]$$

We can see that it is enough to require that either

$$\begin{aligned} &\int_\Omega f^+ d\mu < \infty \quad \text{and} \quad \int_\Omega g^+ d\mu < \infty \\ \text{or} \quad &\int_\Omega f^- d\mu < \infty \quad \text{and} \quad \int_\Omega g^- d\mu < \infty \end{aligned}$$

which is the case when f, g are μ-summable. Of course, that means that $af + bg$ may no longer be μ-summable.

Theorem 3.2.5. Let $f, g : \Omega \to \overline{\mathbb{R}}$ be μ-summable, $\lambda \in \mathbb{R}$. Then

(a) $f + g$ is μ-summable and

$$\int_\Omega (f + g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu$$

(b) λf is μ-summable and

$$\int_\Omega \lambda f d\mu = \lambda \int_\Omega f d\mu$$

Proof. (a) Let f, g as above. Then for any $\varepsilon > 0$, we can choose simple μ-integrable functions $f_\varepsilon, f^\varepsilon, g_\varepsilon, g^\varepsilon$ such that $f_\varepsilon \leq f \leq f^\varepsilon, g_\varepsilon \leq g \leq g^\varepsilon$ μ-a.e. and

$$\begin{aligned} \int_\Omega f^\varepsilon d\mu - \int_\Omega f d\mu < \varepsilon \quad & \int_\Omega f d\mu - \int_\Omega f_\varepsilon d\mu < \varepsilon \\ \int_\Omega g^\varepsilon d\mu - \int_\Omega g d\mu < \varepsilon \quad & \int_\Omega g d\mu - \int_\Omega g_\varepsilon d\mu < \varepsilon \end{aligned}$$

and since f, g are μ-summable, and $f \leq f^\varepsilon \implies (f^\varepsilon)^- \leq f^-$, we get

$$\begin{aligned} \int_\Omega (f_\varepsilon)^+ \leq \int_\Omega f^+ d\mu \leq \int_\Omega |f| d\mu < \infty \\ \int_\Omega (f^\varepsilon)^- \leq \int_\Omega f^- d\mu \leq \int_\Omega |f| d\mu < \infty \end{aligned}$$

Ditto for $(g_\varepsilon)^+$ and $(g^\varepsilon)^-$. By the previous Lemma (using the weakened condition), it follows that $f_\varepsilon + g_\varepsilon$ and $f^\varepsilon + g^\varepsilon$ are μ-integrable and

$$\begin{aligned} \int_\Omega (f^\varepsilon + g^\varepsilon) d\mu = \int_\Omega f^\varepsilon d\mu + \int_\Omega g^\varepsilon d\mu \\ \int_\Omega (f^\varepsilon + g^\varepsilon) d\mu = \int_\Omega f_\varepsilon d\mu + \int_\Omega g_\varepsilon d\mu \end{aligned}$$

which gives us the estimate

$$\begin{aligned}\int_\Omega f d\mu + \int_\Omega g d\mu - 2\varepsilon \leq \int_\Omega f_\varepsilon d\mu + \int_\Omega g_\varepsilon d\mu = \int_\Omega (f_\varepsilon + g_\varepsilon) d\mu \leq \underline{\int_\Omega} (f+g) d\mu \\ \leq \overline{\int_\Omega} (f+g) d\mu \leq \int_\Omega (f^\varepsilon + g^\varepsilon) d\mu = \int_\Omega f^\varepsilon d\mu + \int_\Omega g^\varepsilon d\mu \\ \leq \int_\Omega f d\mu + \int_\Omega g d\mu + 2\varepsilon\end{aligned}$$

Letting $\varepsilon \to 0$, it follows that $(f+g)$ is μ-integrable with

$$\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu$$

To show that $(f+g)$ is μ-summable, we apply the previous result to $|f|$ and $|g|$ with the triangle inequality:

$$\int_\Omega |f+g| d\mu \leq \int_\Omega |f| + |g| d\mu \leq = \int_\Omega |f| d\mu + \int_\Omega |g| d\mu < \infty$$

(b) Let $\lambda \in \mathbb{R}$. For all $\varepsilon > 0$, we can again find μ-integrable simple functions $f_\varepsilon \leq f \leq f^\varepsilon$ with

$$\int_\Omega f^\varepsilon d\mu - \int_\Omega f d\mu < \varepsilon \quad \text{and} \quad \int_\Omega f d\mu - \int_\Omega f_\varepsilon d\mu < \varepsilon$$

- If $\lambda = 0$, then it's trivial.
- If $\lambda \geq 0$, then

  $$\lambda f_\varepsilon \leq \lambda f \leq \lambda f^\varepsilon$$

  and using the previous Lemma, we can make the estimates

  $$\begin{aligned}\lambda \int_\Omega f d\mu - \lambda\varepsilon \leq \lambda \int_\Omega f_\varepsilon d\mu \leq \underline{\int_\Omega} (\lambda f) d\mu \\ \leq \overline{\int_\Omega} (\lambda f) d\mu \leq \int_\Omega \lambda f^\varepsilon d\mu = \lambda \int_\Omega f^\varepsilon d\mu \\ \leq \lambda \int_\Omega f d\mu + \lambda\varepsilon\end{aligned}$$

  and thus λf is μ-integrable with

  $$\int_\Omega (\lambda f) d\mu = \lambda \int_\Omega f d\mu$$

  and applying this result to $|\lambda f|$, we get

  $$\int_\Omega |\lambda f| d\mu = \int_\Omega |\lambda||f| d\mu = |\lambda| \int_\Omega |f| d\mu < \infty$$

  which shows that λf is μ-summable.
- If $\lambda < 0$, then we have $\lambda f_\varepsilon \geq \lambda f \leq \lambda f^\varepsilon$ and the proof is analogous to the case $\lambda > 0$.

□

Corollary 3.2.6 (Continuous triangle inequality)**.** Let $f : \Omega \to \overline{\mathbb{R}}$ be μ-summable. Then

$$|\int_\Omega f d\mu| \leq \int_\Omega |f| d\mu$$

Proof. From the previous theorem and monotonicity, it follows from $-|f| \leq f \leq |f|$ that

$$-\int_\Omega |f| d\mu = \int_\Omega -|f| d\mu \leq \int_\Omega f d\mu \leq \int_\Omega |f| d\mu$$

□

Given a function $f : \Omega \to \overline{\mathbb{R}}$, there are two ways to restrict a function to a subset $A \subseteq \Omega$.

Lemma 3.2.7. Let $f : \Omega \to \overline{\mathbb{R}}$ be μ-summable and $A \subseteq \Omega$ μ-measurable.
Then $f|_A$ and $f\chi_A$ are μ-summable (on A and Ω) and

$$\int_A f|_A d\mu = \int_\Omega f\chi_A d\mu$$

Proof. We first show this for simple functions and then generalize from there.

- Let $g : \Omega \to \overline{\mathbb{R}}$ be a μ-summable *simple* function. By Lemma 3.2.3, there exist mutually disjoint, μ-measurable sets $(A_i)_{i\in I}$ with values $(a_i)_{i\in\mathbb{N}}$ such that $g = \sum_{i\in\mathbb{N}} a_i\chi_{A_i}$. Moreover, we know (by the same Lemma) that the integrals of $g|_A$ and $g\chi_A$ are

$$\begin{aligned} g|_A = \sum_{i\in\mathbb{N}} a_i\chi_{A_i\cap A} \quad &\implies \int_A g|_A d\mu = \sum_{i\in\mathbb{N}} a_i\mu(A_i \cap A) \\ g\chi_A = \sum_{i\in\mathbb{N}} a_i\chi_{A_i}\chi_A = \sum_{i\in\mathbb{N}} a_i\chi_{A_i\cap A} \quad &\implies \int_\Omega g\chi_A d\mu = \sum_{i\in\mathbb{N}} a_i\mu(A_i \cap A) \end{aligned}$$

  And they are μ-summable because $\mu(A_i \cap A) \leq \mu(A_i)$ and $|g\chi_A| \leq |g|$.

- Let $f : \Omega \to \overline{\mathbb{R}}$ μ-summable. For all $\varepsilon > 0$ we can find simple μ-integrable functions g, h such that $g \leq f \leq h$ μ-a.e. (and thus also $g|_A \leq f|_A \leq h|_A$ μ-a.e.) such that

$$\int_\Omega h d\mu - \varepsilon \leq \int_\Omega f d\mu \leq \int_\Omega g d\mu + \varepsilon$$

  To show that $f|_A$ is μ-integrable, we can check (using linearity of the integral) that

$$\begin{aligned} 0 &\leq \overline{\int_A} f|_A d\mu - \underline{\int_A} f|_A d\mu \leq \int_A h|_A d\mu - \int_A g|_A d\mu \\ &= \int_\Omega h\chi_A d\mu - \int_\Omega g\chi_A d\mu = \int_\Omega (h-g)\chi_A d\mu \leq \int_\Omega (h-g) d\mu \\ &= \int_\Omega h d\mu - \int_\Omega g d\mu \leq \int_\Omega f d\mu - \int_\Omega f d\mu + 2\varepsilon = 2\varepsilon \end{aligned}$$

  Letting $\varepsilon \to 0$ shows that $f|_A$ is μ-integrable.

  To show that $f\chi_A$ is μ-integrable, check

$$\begin{aligned} 0 &\leq \overline{\int_\Omega} f\chi_A d\mu - \underline{\int_\Omega} f\chi_A d\mu \leq \int_\Omega h\chi_A d\mu - \int_\Omega g\chi_A d\mu = \int_\Omega (h-g)\chi_A d\mu \\ &\leq \int_\Omega (h-g) d\mu = \int_\Omega h d\mu - \int_\Omega g d\mu \leq 2\varepsilon \end{aligned}$$

Now we only need to show that their integrals are equal. On the one hand, we have

$$\int_A f|_A d\mu - \int_\Omega f\chi_A d\mu \leq \int_A h|_A d\mu - \int_\Omega g\chi_A d\mu = \int_\Omega h\chi_A d\mu - \int_\Omega g\chi_A d\mu = \int_\Omega (h-g)\chi_A d\mu \leq 2\varepsilon$$

and on the other hand, we have

$$\int_A f|_A d\mu - \int_\Omega f\chi_A d\mu \geq \int_A g|_A d\mu - \int_\Omega h\chi_A d\mu = \int_\Omega g\chi_A d\mu - \int_\Omega h\chi_A d\mu = \int_\Omega (g-h)\chi_A d\mu \geq -2\varepsilon$$

Which means

$$|\int_A f|_A d\mu - \int_\Omega f\chi_A d\mu| \leq 2\varepsilon$$

□

Since we have just showed that both ways of restricting a μ-summable function yield the same integral, we will from now on write

$$\int_A f d\mu := \int_A f|_A d\mu = \int_\Omega f\chi_A d\mu$$

Corollary 3.2.8. Let $f : \Omega \to \overline{\mathbb{R}}$ be a μ-summable function and $A \subseteq \Omega$ with $\mu(A) = 0$. Then

$$\int_A f d\mu = 0$$

Proof. Since $\mu(A) = 0$ it is μ-measurable and $f\chi_A = 0$ μ-a.e.. By Proposition 3.1.9 we have $\int_\Omega f\chi_A d\mu = 0$ and from the previous Lemma, the result follows. □

Proposition 3.2.9. Let $f : \Omega \to \overline{\mathbb{R}}$ be μ-summable, $\Omega = A \cup B$, for $A, B \subseteq \Omega$ μ-measurable.

(a) If $A \cap B = \emptyset$, then

$$\int_\Omega f d\mu = \int_A f d\mu + \int_B f d\mu$$

(b) And more generally, for $A \cap B \neq \emptyset$:

$$\int_\Omega f d\mu = \int_A f d\mu + \int_B f d\mu - \int_{A\cap B} f d\mu$$

Proof. (a) Clearly, $f = f\chi_A + f\chi_B$, So by linearity and Lemma 3.2.7 , we have

$$\int_\Omega f d\mu = \int_\Omega f\chi_A + f\chi_B d\mu = \int_\Omega f\chi_A d\mu + \int_\Omega f\chi_B d\mu = \int_A f d\mu + \int_B f d\mu$$

(b) We can partition Ω into the disjoint sets $\Omega = (A \setminus B) \sqcup (B \setminus A) \sqcup (A \cap B)$ together with

$$\chi_A = \chi_{A\setminus B} + \chi_{A\cap B} \quad \text{and} \quad \chi_B = \chi_{B\setminus A} + \chi_{A\cap B}$$

we can write

$$\begin{aligned} 1 = \chi_\Omega &= \chi_{A\setminus B} + \chi_{B\setminus A} + \chi_{A\cap B} \\ &= \chi_A - \chi_{A\cap B} + \chi_B - \chi_{A\cap B} + \chi_{B\cap A} \\ &= \chi_A + \chi_B - \chi_{A\cap B} \end{aligned}$$

So by (a), we have

□

3.3 Comparison between Lebesgue and Riemann-Integral

3.3.1 The Riemann Integral

Definition 3.3.1. Let $I = [a, b] \subseteq \mathbb{R}$ and $P = \{a = x_0, x_1, \ldots x_n = b\}$ a partition of I.
For a bounded function $f : I \to \mathbb{R}$ and P a partition on I, we define the **upper** and **lower Riemann sums** by

$$S(P, f) := \sum_{i=1}^{n} (x_i - x_{i-1}) \sup_{x \in (x_{i-1}, x_i]} f(x)$$
$$s(P, f) := \sum_{i=1}^{n} (x_i - x_{i-1}) \inf_{x \in (x_{i-1}, x_i]} f(x)$$

For $\mathcal{P}$ the set of all partitions of I, define

$$\mathcal{R} \overline{\int_a^b} f(x)dx := \inf\{S(P, f), P \in \mathcal{P}\}$$
$$\mathcal{R} \underline{\int_a^b} f(x)dx := \sup\{S(P, f), P \in \mathcal{P}\}$$

and we say that a bounded function $f : I \to \mathbb{R}$ is Riemann integrable (or short: $\mathcal{R}$-integrable) if

$$\mathcal{R} \underline{\int_a^b} f(x)dx = \mathcal{R} \overline{\int_a^b} f(x)dx =: \int_a^b f(x)dx$$

The following definition is neither standard nor part of the original lecture.

Definition 3.3.2. Let $I \subseteq \mathbb{R}$ be an interval. A **step-function** is a function $\varphi : I \to \mathbb{R}$ of the form

$$\varphi = \sum_{i=1}^{n} c_i \chi_{(x_{i-1}, x_i]}$$

where $P = \{x_1, \ldots, x_n\}$ is a partition of I and $c_1, \ldots, c_n \in \mathbb{R}$.

Since the interval $(x_{i-1}, x_i]$ is $\mathcal{L}^1$-measurable, step-functions are $\mathcal{L}^1$-integrable and

$$\int_{[a,b]} \varphi d\mathcal{L}^1 = \sum_{i=1}^{n} c_i (x_i - x_{i-1})$$

For any partition P, we can express the upper and lower Riemann sums $S(P, f)$ and $s(P, f)$ using the $\mathcal{L}^1$-integral of step functions. By defining[6]

$$\overline{P}(x) = \sum_{i=1}^{n} \sup_{x \in (x_{i-1}, x_i]} f(x) \chi_{(x_{i-1}, x_i]}$$
$$\underline{P}(x) = \sum_{i=1}^{n} \inf_{x \in (x_{i-1}, x_i]} f(x) \chi_{(x_{i-1}, x_i]}$$

it's easy to see that

$$S(P, f) = \int_{[a,b]} \overline{P}(x) d\mathcal{L}^1 \quad \text{and} \quad s(P, f) = \int_{[a,b]} \underline{P}(x) d\mathcal{L}^1$$

[6] The orginal lecture wrote $\overline{\varphi}, \underline{\varphi}$ instead of $\overline{P}, \underline{P}$. That confused me so I changed it.

However, we aren't quite satisfied. If you wanted to construct $\overline{P}$ or $\underline{P}$, we would have to calculate $\sup_{x\in(x_{i-1},x_i]} f(x)$ for each interval which isn't so handy.
It's clear that $\underline{P}(x) \leq f(x) \leq \overline{P}(x)$ and for any other step function $\varphi(x)$ to the same partition P will satisfy

$$\varphi \geq f \implies \varphi \geq \overline{P} \quad \text{and} \quad \varphi \leq f \implies \varphi \leq \underline{P}$$

which means

$$S(P,f) = \inf\left\{\int_{[a,b]} \varphi d\mathcal{L}^1 | \varphi : I \to \mathbb{R} \text{ is a step function to the partition } P, \varphi \geq f\right\}$$
$$s(P,f) = \sup\left\{\int_{[a,b]} \varphi d\mathcal{L}^1 | \varphi : I \to \mathbb{R} \text{ is a step function to the partition } P, \varphi \leq f\right\}$$

so we finally get the alternative characterisation of the Riemann-integral using the Lebesgue integral of step functions:

$$\mathcal{R}\overline{\int_a^b} f(x)dx = \inf\left\{\int_{[a,b]} \varphi d\mathcal{L}^1 | \varphi : I \to \mathbb{R} \text{ is a step-function, } \varphi \geq f\right\}$$
$$\mathcal{R}\underline{\int_a^b} f(x)dx = \sup\left\{\int_{[a,b]} \underline{\varphi} d\mathcal{L}^1 | \varphi : I \to \mathbb{R} \text{ is a step-function, } \varphi \geq f\right\}$$

One can visualize the difference between the Riemann and Lebesgue-integral as follows:

The Riemann integral considers the area under a curve as made out of vertical rectangles, whereas the Lebesgue integral considers horizontal slabs under the curve.

Example 3.3.3.
- The Dirichlet function $\chi_{\mathbb{Q}\cap[0,1]}$ is an example of a function that is not Riemann integrable. Because for any partition P (which is made up of *finitely* many sections!), one finds that $S(P,f) = 1$ and $s(P,f) = 0$.
- Moreover, the Riemann integral doesn't behave as nicely when considering limits of functions. For example, for some enumeration $\mathbb{Q} \cap [0,1] = \{r_1, r_2, r_3, \ldots\}$ we can define

  $$f_n = \chi_{r_1,\ldots,r_n}$$

  It is easy to see that f_n is $\mathcal{R}$-integrable with integral $\mathcal{R}\int_0^1 f_n(x)dx = 0$.

  But in the limit $n \to \infty$, the sequence of functions converges to the non $\mathcal{R}$-integrable Dirichlet function.

We will see later in this chapter that the Lebesgue integral behaves much more nicely, because the characteristic functions of $\mathcal{L}^1$-measurable sets are $\mathcal{L}^1$-integrable.

Proposition 3.3.4. Let $f : [a,b] \to \mathbb{R}$ be a bounded $\mathcal{R}$-integrable function. Then it is $\mathcal{L}^1$-integrable and

$$\mathcal{R}\int_a^b f(x)dx = \int_{[a,b]} f d\mathcal{L}^1$$

Proof. As noted earlier, step functions are $\mathcal{L}^1$-integrable (but not the other way around). So comparing the alternative characterisation above with Definition 3.1.2, where we are taking the supremum/infimum over a larger set, we have

$$\inf\left\{\int_{[a,b]}\varphi d\mathcal{L}^1 \middle| \varphi : I \to \mathbb{R} \text{ a step function, } \varphi \geq f\right\}$$
$$\geq \inf\left\{\int_{[a,b]} g d\mathcal{L}^1 \middle| g \text{ is a } \mathcal{L}^1\text{-integrable simple function, } g \geq f \ \mu\text{-a.e.}\right\}$$

and similarly: $\sup\left\{\int_{[a,b]}\varphi d\mathcal{L}^1|\ldots\right\} \leq \sup\left\{\int_{[a,b]} g d\mathcal{L}^1|\ldots\right\}$, so it follows that

$$\mathcal{R}\ \underline{\int_a^b} f(x)dx = \sup\left\{\int_{[a,b]}\varphi d\mathcal{L}^1|\ldots\right\} \leq \sup\left\{\int_{[a,b]} g d\mathcal{L}^1|\ldots\right\} \leq \underline{\int_{[a,b]}} f d\mu$$
$$\leq \overline{\int_{[a,b]}} f d\mu = \inf\left\{\int_{[a,b]} g d\mathcal{L}^1|\ldots\right\} \leq \inf\left\{\int_{[a,b]}\varphi d\mathcal{L}^1|\ldots\right\} = \mathcal{R}\ \overline{\int_a^b} f(x)dx$$

so f is is $\mathcal{R}$-integrable, the inequalities become equalities and $\underline{\int_{[a,b]}} f d\mu = \overline{\int_{[a,b]}} f d\mu$, showing that f is $\mathcal{L}^1$-integrable and that their integral coincides. ☐

3.4 Convergence Results

In this chapter, we want to see when we can exchange limits with integrals, i.e. when

$$\lim_{k\to\infty}\int_\Omega f_k d\mu = \int_\Omega \lim_{k\to\infty} f_k d\mu$$

It is not difficult to find counterexamples, but when this holds, it allows us to calculate many integrals. For this section, let μ be a measure on $\mathbb{R}^n$ and Ω μ-measurable.

Theorem 3.4.1 (Fatou's Lemma)**.** Let μ be a Radon measure on $\mathbb{R}^n$ and $f_k : \Omega \to [0,\infty]$ be a sequence of μ-measurable functions. Then

$$\int_\Omega \liminf_{k\to\infty} f_k d\mu \leq \liminf_{k\to\infty}\int_\Omega f_k d\mu$$

Proof. Since the f_k are μ-integrable, by Theorem 2.1.5 $f := \liminf_{k\to\infty} f_k$ is aswell.
We show that for every μ-integrable simple function $g \leq f$:

$$\int_\Omega g d\mu \leq \liminf_{k\to\infty}\int_\Omega f_k d\mu$$

Since $f_k \geq 0$ and by Lemma 3.2.3, we can assume without loss of generality that $g \geq 0$ is of the form $g = \sum_{j=0}^\infty a_j \chi_{A_j}$ with A_j μ-measurable and pairwise disjoint and $a_0 = 0, a_i > 0$ for $i > 0$[7].
For any factor $t \in (0,1)$, define

$$B_{j,k} := \{x \in A_j | f_l(x) > t a_j, \forall l \geq k\}$$

[7] We can't have it that all $a_i > 0$, or else the domain of g isn't Ω anymore. For example with $h = 0$.

then $A_j = \bigcup_{k=1}^{\infty} B_{j,k}$. Because the series is increasing in k, by continuity from below

$$\lim_{k\to\infty} \mu(B_{j,k}) = \mu(A_j)$$

For $J, k \in \mathbb{N}$ fix:

$$\int_\Omega f_k d\mu \geq \sum_{j=1}^{J} \int_{A_j} f_k d\mu \geq \sum_{j=1}^{J} \int_{B_{j,k}} f_k d\mu \geq t \cdot \sum_{j=1}^{J} a_j \mu(B_{j,k})$$

Fist, we let $k \to \infty$ and get

$$\liminf_{k\to\infty} \int_\Omega f_k d\mu \geq \lim_{k\to\infty} t \cdot \sum_{j=1}^{J} a_j \mu(B_{j,k}) = t \cdot \sum_{j=1}^{J} a_j \mu(A_j)$$

then let $J \to \infty$:

$$\liminf_{k\to\infty} \int_\Omega f_k d\mu \geq t \sum_{j=1}^{\infty} a_j \mu(A_j) = t \int_\Omega g d\mu$$

and if we let $t \to 1$, we get

$$\liminf_{k\to\infty} \int_\Omega f_k d\mu \geq \int_\Omega g d\mu$$

□

Example 3.4.2. The condition $f_k \geq 0$ in Fatou's Lemma is necessary. Take for example $\mu = \mathcal{L}^n, \Omega = \mathbb{R}^n$ and $f_k = -\frac{1}{k^n}\chi_{B_k(0)}$. Since the volume of the ball $B_k(0)$ is proportional to k^n, we have

$$\liminf_{k\to\infty} \int_\Omega f_k d\mu = -C$$

for some constant C (that depends on n). But the function sequence converges uniformly to $\liminf_{k\to\infty} f_k = 0$.

The following theorem is also sometimes known as Beppo Levi's Theorem. We will use the descriptive name Monotone Convergence Theorem, or MCT for short.

Theorem 3.4.3 (Monotone Convergence Theorem)**.** Let $f_k : \Omega \to [0, \infty]$ be a sequence of μ-measurable functions non-decreasing in k (i.e. $f_1 \leq \ldots \leq f_k \leq f_{k+1} \leq \ldots$). Then

$$\int_\Omega \lim_{k\to\infty} f_k d\mu = \lim_{k\to\infty} \int_\Omega f_k d\mu$$

Proof. We show inequalities in both directions.

$\geq$: Since f_k is non-decreasing, we have for all $j \in \mathbb{N}$ that $\int_\Omega f_j d\mu \leq \int_\Omega \lim_{k\to\infty} f_k d\mu$. So in the limit

$$\lim_{k\to\infty} \int_\Omega f_k d\mu \leq \int_\Omega \lim_{k\to\infty} f_k d\mu$$

$\leq$: By Fatou's Lemma

$$\int_\Omega \lim_{k\to\infty} f_k d\mu = \int_\Omega \liminf_{k\to\infty} f_k d\mu \overset{\text{Fatou}}{\leq} \liminf_{k\to\infty} \int_\Omega f_k d\mu = \lim_{k\to\infty} \int_\Omega f_k d\mu$$

□

Remark 3.4.4. We could also prove MCT first and use it to prove Fatou's Lemma.

Proof of Beppo Levi without Fatou's Lemma. Set $f := \lim_{k\to\infty} f_k$. Since $(f_k)_{k\in\mathbb{N}}$ is non-decreasing, we have $\lim_{k\to\infty} \int_\Omega f_k d\mu \leq \int_\Omega f d\mu$.
For any simple function h and $\varepsilon > 0$, set

$$\Omega_k := \{x \in \Omega \big| f_n(x) \geq (1-\varepsilon)h(x)\}$$

this forms an increasing sequence ($\Omega_k \subseteq \Omega_{k+1}$). We know that WLOG, h is of the form $h = \sum_{i\in\mathbb{N}} c_i \chi_{A_i}$ with $c_0 = 0, c_i > 0$ for $i > 0$ and A_i mutually disjoint. Then we can also show that $\bigcup_{k\in\mathbb{N}} \Omega_k = \Omega$.
Suppose there exists an $x_0 \in \Omega \setminus \bigcup_{k\in\mathbb{N}} \Omega$. This means $\forall k \in \mathbb{N} : f_k(x_0) < (1-\varepsilon)h(x_0)$. But then

$$\lim_{k\to\infty} f_k(x_0) = f(x_0) \leq (1-\varepsilon)h(x_0) < h(x_0) \lightning$$

Therefore, one finds

$$\int_\Omega f_k d\mu \geq \int_{\Omega_k} f_k d\mu \geq \int_{\Omega_k} (1-\varepsilon) h d\mu$$

in the limit $k \to \infty$, it follows

$$\lim_{k\to\infty} \int_\Omega f_k d\mu \geq \lim_{k\to\infty} \int_{\Omega_k} (1-\varepsilon) h d\mu = \int_\Omega (1-\varepsilon) h d\mu$$

letting $\varepsilon \to 0$ we obtain

$$\lim_{k\to\infty} \int_\Omega f_k d\mu \geq \int_\Omega h d\mu$$

since h was an arbitrary simple function $\leq f$, the proof follows.

□

Example 3.4.5. A neat application of this is the special case where the limit is sum of non-negative functions: Let $(f_k)_{k\in\mathbb{N}} : \Omega \to [0, \infty]$ be a sequence of μ-measurable functions. Then $\sum_{k=1}^{\infty} f_k$ is μ-measurable and

$$\int_\Omega \sum_{k=1}^{\infty} f_k d\mu = \sum_{k=1}^{\infty} \int_\Omega f_k d\mu$$

and to prove this, we apply MCT to the sequence of partial sums $s_n := \sum_{k=1}^{n} f_k$.

Theorem 3.4.6 (Dominated Convergence Theorem)**.** Let $f, (f_k)_{k\in\mathbb{N}} : \Omega \to \overline{\mathbb{R}}$ μ-measurable with $f_k \to f$ μ-a.e..
If there exists a μ-summable function $g : \Omega \to [0, \infty]$ with $|f_k| \leq g$ μ-a.e., then

$$\lim_{k\to\infty} \int_\Omega |f_k - k| d\mu = 0$$

and in particular

$$\lim_{k\to\infty} \int_\Omega f_k d\mu = \int_\Omega f d\mu$$

Proof. First note from $|f| = \lim_{k\to\infty}|f_k| \le g$ μ-a.e. it follows

$$|f_k - f| \le |f_k| + |f| \le 2g$$

so $f, |f_k - k|$ are μ-summable and

$$\liminf_{k\to\infty} 2g - |f_k - f| = 2g \quad \mu\text{-a.e.}$$

so with Corollary 3.1.6, Fatou's Lemma and linearity:

$$\begin{aligned}\int_\Omega 2g d\mu &= \int_\Omega \liminf_{k\to\infty}(2g - |f_k - f|)d\mu \\ &\overset{\text{Fatou}}{\le} \liminf_{k\to\infty}\int_\Omega (2g - |f_k - f|)d\mu \\ &= \int_\Omega 2g d\mu - \limsup_{k\to\infty}\int_\Omega |f_k - f| d\mu\end{aligned}$$

And since $|f_k - f| \ge 0$, we have

$$\lim_{k\to\infty}\int_\Omega |f_k - f| d\mu = 0$$

the second statement follows from linearity and the triangle inequality

$$0 \le |\int_\Omega f_k d\mu - \int_\Omega f d\mu| = |\int_\Omega f_k - f d\mu| \le \int_\Omega |f_k - f| d\mu \overset{k\to\infty}{\to} 0$$

□

We can also get a second proof of 3.2.2

Second Proof. `Skipped` □

3.5 Absolute Continuity of Integrals

Let μ be a Radon measure on $\mathbb{R}^n$ and Ω be μ-measurable.
For any μ-summable function $f : \Omega \to \overline{\mathbb{R}}$, we can define

$$\nu : \Sigma_\mu \to \mathbb{R}, \quad A \mapsto \nu(A) := \int_A f d\mu$$

where Σ_μ is the σ-algebra of μ-measurable sets.

Remark 3.5.1. By Corollary 3.2.8, we have

$$\mu(A) = 0 \implies \nu(A) = 0$$

Moreover, if $f \ge 0$ and μ is a Radon measure, then ν is σ-additive and again a Radon measure.
We then write

$$\nu := \mu \llcorner f$$

this notation coincides with the restriction of measures to subsets we saw in Section 1.9, as

$$\mu \llcorner \chi_A = \mu \llcorner A$$

Definition 3.5.2. Let Σ_μ, Σ_ν denote the σ-algebra of μ-(resp. ν)-measurable subsets. A measure ν such that $\Sigma_\mu \subseteq \Sigma_\nu$ with property

$$\mu(A) = 0 \implies \nu(A) = 0$$

is called **absolutely continuous** with respect to μ and we write $\nu \ll \mu$.

The name of the definition should be reminiscent of a definition from Analysis. Let's see why:

Theorem 3.5.3. Let $f : \Omega \to \overline{\mathbb{R}}$ be μ-summable. Then $\forall \varepsilon > 0\ \exists \delta > 0$ such that

$$\forall A \in \Sigma_\mu : \quad \mu(A) < \delta \implies \int_A |f| d\mu < \varepsilon$$

Proof by contradiction. Assume that there exists an $\varepsilon > 0$ and a sequence of μ-measurable subsets $A_k \subseteq \Omega$ with $\mu(A_k) < 2^{-k}$ that satisfy

$$\int_{A_k} |f| d\mu \geq \varepsilon \quad \forall k \in \mathbb{N}$$

for $m \geq 1$ we define

$$B_m := \bigcup_{k=m}^{\infty} A_k$$

This forms a decreasing sequence and by subadditivity: $\mu(B_m) \leq \sum_{k=m}^{\infty} < 2^{1-m}$ which means $\mu(B_1) < \infty$ and

$$\int_{B_m} |f| d\mu \geq \int_{A_k} |f| d\mu \geq \varepsilon, \forall k \geq m$$

By continuity from above, for $B := \bigcup_{l=m}^{\infty} B_m$ we have $\mu(B) = \lim_{m\to\infty} \mu(B_m) = 0$. And

$$f_m := |f| \chi_{B_m} \implies \lim_{m\to\infty} f_m = |f| \chi_B$$

Since $|f_k| \leq |f|$ and $|f|$ is μ-summable, we can use the dominated convergence theorem to show

$$\varepsilon \geq \lim_{m\to\infty} \int_{B_m} |f| d\mu = \lim_{m\to\infty} \int_\Omega f_m d\mu \overset{\text{DCT}}{=} \int_\Omega \lim_{m\to\infty} f_m d\mu = \int_A |f| d\mu = 0 \lightning$$

$\square$

3.6 Vitali's Theorem

Lebesgue's Theorem gives us a sufficient condition to pass exchange limits and integrals.
In this chapter, we will improve it and give a necessary condition to exchange limits and integrals.
In this section, let $f, f_k : \Omega \to \overline{\mathbb{R}}$ be μ-summable.

Definition 3.6.1. The family $\{f_k\}_{k\in\mathbb{N}}$ is called **uniformly μ-summable**, if $\forall \varepsilon > 0\ \exists \delta > 0 : \forall k \in \mathbb{N}, \forall A \subseteq \Omega$ μ-measurable:

$$\mu(A) < \delta \implies \int_A |f_k| d\mu < \varepsilon$$

Theorem 3.6.2 (Vitali's Theorem)**.** Let $f, (f_k)_{k\in\mathbb{N}} : \Omega \to \overline{\mathbb{R}}$. If $\mu(\Omega) < \infty$, the following are equivalent:

(a) $\{f_k\}_{k\in\mathbb{N}}$ is uniformly μ-summable and $f_k \xrightarrow{\mu} f$.

(b) $\lim_{k\to\infty} \int_\Omega |f_k - f| d\mu = 0$

Proof.

(b) $\implies$ (a) If $\int_\Omega |f_k - f| d\mu \to 0$, by Corollary 3.2.2 it follows $f_k \xrightarrow{\mu} f$. To show that $\{f_k\}_{k\in\mathbb{N}}$ is uniformly μ-summable, let $\varepsilon > 0$.

By our assumption (b), there exist a k_0 such that

$$\int_\Omega |f_k - f| d\mu < \varepsilon \quad \forall k \geq k_0$$

by the previous theorem (Theorem 3.5.3), there exists a $\delta > 0$ such that $\forall A \subseteq \Omega$ μ-measurable

$$\mu(A) < \delta \implies \int_A |f| d\mu < \varepsilon \quad \text{and} \quad \max_{1\leq k\leq k_0} \int_A |f_k| d\mu < \varepsilon$$

and for the remaining $k \geq k_0$, if $\mu(A) < \delta$, we get

$$\int_A |f_k| d\mu \leq \int_A |f| d\mu + \int_A |f_k - f| d\mu \leq 2\varepsilon$$

which shows that $\{f_k\}_{k\in\mathbb{N}}$ is uniformly μ-summable.

(a) $\implies$ (b) We first show that (b) holds for some subsequence $(f_{k_n})_{n\in\mathbb{N}}$, and then show it for the entire sequence.

By Theorem 2.3.5, $f_k \xrightarrow{\mu} f$ means that there exists a subsequence $(f_{k_n})_{n\in\mathbb{N}}$ that converges to f μ-a.e.

By assumption (a), for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\mu(A) < \delta \implies \int_A |f| d\mu < \varepsilon \quad \text{and} \quad \int_A |f_k| d\mu < \varepsilon, \quad \forall k \in \mathbb{N}$$

Feeding δ into Egoroff's Theorem, we get a $B \subseteq \Omega$ with

$$\mu(\Omega \setminus B) < \delta \quad \text{and} \quad \sup_{x\in B} |f_{k_n}(x) - f(x)| \to 0 \text{ as } n \to \infty$$

Now we can bound the integral $\int_\Omega |f_{k_n} - f| d\mu$. Because the above limit approaches limit, set $n_0 \in \mathbb{N}$ such that

$$\sup_{x\in B} |f_{k_n}(x) - f(x)| < \frac{\varepsilon}{\mu(\Omega)}$$

Thus, for all $n \geq n_0$, we have

$$\begin{aligned}
\int_\Omega |f_{k_n} - f| d\mu &= \int_{\Omega\setminus B} |f_{k_n} - f| d\mu + \int_B |f_{k_n} - f| d\mu \\
&\leq \int_{\Omega\setminus B} |f_{k_n}| + |f| d\mu + \int_B \sup_{x\in B} |f_{k_n} - f| d\mu \\
&< 2\varepsilon + \frac{\varepsilon}{\mu(\Omega)} \mu(B) \leq 3\varepsilon
\end{aligned}$$

Letting $\varepsilon \to 0$, we find $\lim_{n\to\infty} \int_\Omega |f_{k_n} - f| d\mu = 0$.

We will prove by contradiction that it also holds for the entire sequence.

Suppose $\limsup_{k\to\infty}\int_\Omega |f_k - f| d\mu > 0$. Then there exists a subsequence $(f_{k_n})_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty}\int_\Omega |f_{k_n} - f| d\mu = \limsup_{k\to\infty}\int_\Omega |f_k - f| d\mu > 0$$

But nonetheless, $\{f_{k_n}\}_{n\in\mathbb{N}}$ is still uniformly μ-summable and $f_{k_n} \xrightarrow{\mu} f$ as $n\to\infty$, repeating the argument from before, there exists a sub-subsequence $(f_{k_{n'}})_{n'\in\mathbb{N}}$ such that

$$\lim_{n'\to\infty}\int_\Omega |f_{k_{n'}} - f| d\mu = 0$$

which contradicts the previous equation.

$\square$

Remark 3.6.3. Since the proof in (a) $\implies$ (b) relied on Egoroff's Theorem, the condition $\mu(\Omega) < \infty$ is necessary. The other implication always holds.
Consider the sequence given by $f_k = \frac{1}{k}\chi_{[0,k]}$. Clearly, $f_k \xrightarrow{\mu} \mathbf{0}$, but

$$\int_{\mathbb{R}} |f_k - \mathbf{0}| d\mu = \int_{[0,k]} \frac{1}{k} d\mu = 1 \neq 0$$

As promised earlier, we can now improve Lebesgue's Theorem.

`Missing 30 minutes`

Theorem 3.6.4.

3.7 $L^p(\Omega,\mu)$ spaces

Again, let μ be a Radon measure on $\mathbb{R}^n$ and Ω μ-measurable.
Recall that the supremum/infimum of a set are defined in terms of upper/lower bounds.

$$\sup X := \min\{a\in\mathbb{R} | a \geq x, \forall x \in X\}$$

When talking about measure spaces, we want to ignore measure-zero sets:

Definition 3.7.1. Let $X\subseteq\mathbb{R}$. We say that $a\in\mathbb{R}$ is a ***µ*-essential upper bound** of X, if

$$\mu\left(\{x\in X | x > a\}\right) = 0$$

Similar to the defintion of sup, define the ***µ*-essential supremum** to be smallest μ-essential upper bound

$$\mu - \text{ess sup}\, X := \inf\{a\in\mathbb{R} | \mu\left(\{x\in X | x > a\}\right) = 0\}$$

Analogously, one can define the μ-essential infimum.

Definition 3.7.2. Let $f:\Omega\to\overline{\mathbb{R}}$ be μ-measurable. For $1\leq p<\infty$, define their ***p*-norm**

$$\boxed{\|f\|_{L^p(\Omega,\mu)} := \left(\int_\Omega |f|^p d\mu\right)^{1/p} \leq \infty}$$

and for $p = \infty$:

$$\|f\|_{L^\infty(\Omega,\mu)} := \mu - \text{ess sup}_{x\in\Omega}|f(x)| = \inf\{c \in [0,\infty] \,|\, |f| \le c\mu\text{-a.e.}\}$$

We call the space of functions, for which their p-norm is finite the $\boldsymbol{\mathcal{L}^p}$-space

$$\mathcal{L}^p(\Omega, \mu) := \{f : \Omega \to \overline{\mathbb{R}} \,|\, f\mu\text{-measurable}, \|f\|_{L^p(\Omega,\mu)} < \infty\}$$

If Ω and μ are clear from context, we will write $\|f\|_{L^p}$ or even $\|f\|_p$[8]instead of $\|f\|_{L^p(\Omega,\mu)}$.
We call a function f p**integrable**, if $\|f\|_p < \infty$.[9]

Remark 3.7.3. For $f \in \mathcal{L}^p(\Omega, \mu)$ we have $|f(x)| \le \|f\|_\infty$ for μ-almost all $x \in \Omega$

. `Missing` □

Remark 3.7.4. The space $\mathcal{L}^p(\Omega, \mu)$ is not necessarily closed under multiplication. Set $p = 1$, $\Omega = (0, 1]$ and $\mu = \mathcal{L}^1$ and

$$f(x) = g(x) = \frac{1}{\sqrt{x}} \in \mathcal{L}^1(\Omega, \mu)$$

then their product is $(fg)(x) = \frac{1}{x}$, which is not $\mathcal{L}^1$-integrable on $(0, 1]$.

Moreover, $\|\cdot\|_p$ is not a norm on $\mathcal{L}^p(\Omega, \mu)$ as it is not positive definite.
To remedy this, we consider equivalence classes of functions in $\mathcal{L}^p(\Omega, \mu)$ where two functions are considered equivalent if and only if they are equal μ-a.e..

Theorem 3.7.5. Let $L^p(\Omega, \mu)$ denote the quotient space $\mathcal{L}^p(\Omega, \mu)/\sim$, where

$$f \sim g \iff f = g \ \mu\text{-a.e.}$$

Then L^p with norm $\|\cdot\|_p$ is a Banach space. That is: a complete, and normed vector space.

For now, we will only prove that prove positive homogeneity.
The rest will be split into multiple Lemmas and Theorems.

Proof. `Missing` □

Lemma 3.7.6 (Young Inequality)**.** Let $1 < p, q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. (We say that p, q are **conjugated**). Then

$$\forall a, b \ge 0 \quad ab \le \frac{a^p}{p} + \frac{b^q}{q}$$

Proof. If b is zero, it's trivial. Fix $b > 0$ and consider the map

$$f : [0, \infty) \to \mathbb{R}, \quad a \mapsto ab - \frac{a^p}{p}$$

[8] The lecturer never used this, but I will.
[9] This was not used by the lecturer, but this was used in MMP I, for example.

Because $\lim_{a\to\infty} f(a) = -\infty$, f is bounded from above and must have a maximum at $a = a^*$ determined by $f'(a^*) = b - a^{p-1} = 0$.
After checking the second derivative, one finds that $a^* = b^{\frac{1}{p-1}}$ is indeed a maximum of f. So $\forall a \geq 0$

$$ab - \frac{a^p}{p} = f(a) \leq f(a^*) = b^{\frac{1}{p-1}} b - \frac{b^{\frac{p}{p-1}}}{p} = \left(1 - \frac{1}{p}\right) b^{\frac{p}{p-1}} = \frac{b^q}{q}$$

□

Corollary 3.7.7 (Hölder Inequality)**.** Let $1 \leq p, q \leq \infty$ be conjugate, $f \in L^p(\Omega, \mu), g \in L^q(\Omega, \mu)$. Then $fg \in L^1(\Omega, \mu)$ and

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Proof. WLOG we can assume $p \leq q$.

- Case $p = 1, q = \infty$. By Remark 3.7.3, we have $|fg| \leq |f| \|g\|_\infty$ μ-a.e., so by monotonicity of the integral

$$\int_\Omega |fg| d\mu \leq \int_\Omega |f| \|g\|_\infty d\mu \leq \|g\|_\infty \underbrace{\int_\Omega |f| d\mu}_{=\|f\|_1}$$

- Case $1 < p, q < \infty$. If $\|f\|_p = 0$ or $\|g\|_q = 0$, we can use Proposition 3.1.9 and we get zeros on both sides of the equality. So assume $\|f\|_p, \|g\|_q > 0$.

  Set $\tilde{f} := \frac{|f|}{\|f\|_p}$ and $\tilde{g} := \frac{|g|}{\|g\|_q}$ Notice that

$$\int_\Omega \tilde{f}^p d\mu = \frac{1}{(\|f\|_p)^p} \int_\Omega |f|^p d\mu = \frac{(\|f\|_p)^p}{(\|f\|_p)^p} = 1$$

  and likewise $\int_\Omega \tilde{g}^q d\mu = 1$. Apply the Young inequality to $\tilde{f}(x)$ and $\tilde{g}(x)$. We get

$$\tilde{f}\tilde{g} \leq \frac{\tilde{f}^p}{p} + \frac{\tilde{g}^q}{q}$$

  taking the integal on both sides:

$$\frac{1}{\|f\|_p \|g\|_q} \underbrace{\int_\Omega |fg| d\mu}_{=\|fg\|_1} = \int_\Omega \tilde{f}\tilde{g} d\mu \leq \frac{1}{p} \int_\Omega \tilde{f}^p d\mu + \frac{1}{q} \int_\Omega \tilde{g}^q d\mu = \frac{1}{p} + \frac{1}{q} = 1$$

  and the proof follows.

□

Say we know that a function $f : \Omega \to \overline{\mathbb{R}}$ lives in some L^p-space. It would be nice to know if $f \in L^r$ for some other $1 \leq r \leq \infty$.
The next Corollary of the Hölder inequality shows that the different L^p-spaces are nested inside eachother.

Corollary 3.7.8.

(a) Let $f \in L^p, g \in L^q$ with $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} \leq 1$. Then $fg \in L^r$ and $\|fg\|_r \leq \|f\|_p \|g\|_q$.

(b) If $\mu(\Omega) < \infty$, then

$$1 \leq r \leq s \leq \infty \implies L^s(\Omega, \mu) \subseteq L^r(\Omega, \mu)$$

Proof.

(a) If $p = q = r = \infty$, it's clear. If $r < \infty$, apply the Hölder inequality to $|f|^r \in L^{p/r}$ and $|g|^r \in L^{q/r}$. We can do this because $\frac{1}{p/r} + \frac{1}{q/r} = 1$.

(b) If $\mu(\Omega) < \infty$, the function $\mathbf{1}$ is p-integrable for any $1 \leq p \in \infty$ as it has integral $\mu(\Omega)$. Then note that $\frac{1}{r} = \frac{1}{s} + \frac{s-r}{rs}$. So appying (a), we get

$$f \in L^s(\Omega, \mu) \implies \|f\|_r = \|f \cdot \mathbf{1}\|_r \overset{(a)}{\leq} \|f\|_s \cdot \underbrace{\|\mathbf{1}\|_{\frac{rs}{s-r}}}_{=\mu(\Omega)} < \infty$$

which shows $f \in L^r(\Omega, \mu)$

□

Remark 3.7.9. In general, the inclusion in (b) is strict in the sense that there are functions in L^r that do not belong to L^s with $r < s$.
A simple example is

$$\Omega = (0,1), \quad f(x) = \log \frac{1}{x} \in L^p, f \notin L^\infty$$

In our quest to prove $L^p(\Omega, \mu)$ is a Banach Space, we still need to show the Triangle inequality and closure under addition:

Corollary 3.7.10 (Minkowski Inequality)**.** Let $1 \leq p \leq \infty$ and $f, g \in L^p(\Omega, \mu)$. Then $f + g \in L^p$ and

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Proof. The cases $p = 1$ and $p = \infty$ are easy:

$$|(f+g)(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$$

For $1 < p < \infty$, we use that the function $x \mapsto x^p$ is convex. In particular:

$$\left(\frac{a+b}{2}\right)^p \leq \frac{a^p + b^p}{2}$$

Setting $a = 2f(x)$ and $b = 2g(x)$, we get the estimate

$$|(f+g)(x)|^p \leq 2^{p-1} \left(|f(x)|^p + |g(x)|^p\right), \quad \forall x \in \Omega$$

This shows that $f + g \in L^p$ and $|f + g|^{p-1} \in L^{\frac{p}{p-1}}$. □

`Missing until rest of chapter`

`Missing Lecture`

Missing first half

4 Product Measure and Multiple Integrals

4.1 Fubini's Theorem

Let X, Y be sets.

Definition 4.1.1. Let μ be a measure on X and ν a measure on Y. We define their **product measure** $\mu \times \nu : \mathcal{P}(X \times Y) \to [0, \infty]$

$$(\mu \times \nu)(S) := \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) \middle| S \subseteq \bigcup_{i=1}^{\infty} A_i \times B_i, A_i \in \Sigma_\mu, B_i \in \Sigma_\nu \right\}$$

Remark 4.1.2. There is another way of obtaining the product measure. If we were to define for $A \in \Sigma_\mu, B \in \Sigma_\nu$ the function

$$\lambda(A \times B) := \mu(A)\nu(B)$$

then λ would be a pre-measure on the algebra of finite disjoint unions. The product measure would then be the Carathéodory-Hahn extension of λ.

Example 4.1.3. For $X = \mathbb{R}^m, Y = \mathbb{R}^n$, with $\mu = \mathcal{L}^m, \nu = \mathcal{L}^n$. We can show that $\mu \times \nu = \mathcal{L}^{m+n}$.

Theorem 4.1.4 (Fubini)**.** Let μ, ν be Radon measures on $X = R^m, Y = \mathbb{R}^n$, and $\mu \times \nu$ the product measure on $X \times Y = \mathbb{R}^{m+n}$. Then

(a) For every $A \in \Sigma_\mu, B \in \Sigma_\nu$ their product $A \times B$ is $\mu \times \nu$-measurable and

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

(b) For $S \in \Sigma_{\mu \times \nu}$ with $(\mu \times \nu)(S) < \infty$, for ν-almost all $y \in \pi_y(S)$, the set $S_y = \{x \big| (x, y) \in S\}$ is μ-measurable and the mapping

$$y \mapsto \mu(S_y) = \int_X \chi_{S_y} d\mu$$

is ν-summable with

$$\int_Y \mu(S_y) d\nu = \int_Y \int_X \chi_{S_y} d\mu d\nu = (\mu \times \nu)(S)$$

by symmetry in X, Y, the analogue also holds for S_x.

(c) $\mu \times \nu$ is a Radon measure.

(d) If $f : X \times Y \to \overline{\mathbb{R}}$ is $\mu \times \nu$-summable, then the mappings $f_y := f(-, y)$ and $f_x := f(x, -)$ are μ-summable for ν-a.a y and μ-summable, for ν-a.a. y, respectively. Moreover, the mappings

$$y \mapsto \int_X f(x, y) d\mu, \quad x \mapsto \int_Y f(x, y) d\nu$$

are ν- and μ-summable and

$$\begin{aligned} \int_{X \times Y} f d(\mu \times \nu) &= \int_Y \left(\int_X f(x, y) d\mu \right) d\mu \\ &= \int_X \left(\int_Y f(x, y) d\nu \right) d\mu \end{aligned}$$

Proof. `No Proof` □