

# Measure Theory– Lecture Notes

Han-Miru Kim

July 16, 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 $\sigma$ -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

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

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

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)  $\liminf_{n \rightarrow \infty} A_n \subseteq \limsup_{n \rightarrow \infty} A_n$

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

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

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

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

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

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

It holds that

$$\begin{aligned} \limsup_{n \rightarrow \infty} A_n = A &\iff \limsup_{n \rightarrow \infty} \mathbb{1}_{A_n} = \mathbb{1}_A \\ \liminf_{n \rightarrow \infty} A_n = A &\iff \liminf_{n \rightarrow \infty} \mathbb{1}_{A_n} = \mathbb{1}_A \end{aligned}$$

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  **$\sigma$ -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 ( $\sigma$ -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  $\sigma$ -algebras containing  $\mathcal{K}$  forms again a  $\sigma$ -algebra.

We call this the  $\sigma$ -algebra **generated by  $\mathcal{K}$**  and it its the smallest  $\sigma$ -algebra that contains  $\mathcal{K}$ .

The algebra generated by the open sets of a topology is called the **Borel  $\sigma$ -Algebra** of  $X$ , denoted  $\mathcal{B}(X)$ .

## 1.2 Measures

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

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

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

- **$\sigma$ -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  $\sigma$ -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  $\mathcal{A}$ .

- If  $\mu$  is additive, then it is monotone with respect to inclusion, i.e.  $A \subseteq B \implies \mu(A) \leq \mu(B)$ .
- If  $\mu$  is additive and the sets  $A_k$  are mutually disjoint, then

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

- If  $\mu$  is  $\sigma$ -additive, then it is also  **$\sigma$ -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

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

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

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

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

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

is additive but not  $\sigma$ -additive because we can take the sequence

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

which is a sequence of mutually disjoint sets satisfying

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

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

- **finite**, if  $\mu(X) < \infty$
- **$\sigma$ -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,  $\mu$  finite  $\implies \mu$   $\sigma$ -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  $\sigma$ -subadditivity:

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

- (a)  $\mu(\emptyset) = 0$
- (b)  $\mu$  is  $\sigma$ -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  $\mu$  be a measure on  $X$  and  $A \subseteq X$ . We can *restrict*  $\mu$  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  **$\mu$ -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

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

<sup>1</sup>sometimes also called outer measure

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

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

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

forms a  $\sigma$ -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

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

- $A_1, A_2 \in \Sigma \implies A_1 \cup A_2 \in \Sigma$ :

Let  $B \subseteq X$ . From the previous remark, it is sufficient to just show the inequality

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

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

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

so with the decomposition

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

and subadditivity of the measure, we get

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

- $(A_n)_{n \in \mathbb{N}} \subseteq \Sigma \implies A = \bigcup_{n=1}^{\infty} A_n \in \Sigma$ :

We can assume without loss of generality that the sets are mutually disjoint. Otherwise, consider the sequence  $(\tilde{A}_n)_{n \in \mathbb{N}} \subseteq \Sigma$  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  $\mu$ -measureability of  $A_m$  with the test set  $B \cap \bigcup_{k=1}^m A_k$  to find that by induction on  $m$

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

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

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

for all  $m \in \mathbb{N}$ . Taking the limit  $m \rightarrow \infty$ , 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  $\mu$ -measureability of  $A$ .

□

**Definition 1.2.10.** A **measure space** is a tuple  $(X, \Sigma, \mu)$  consisting of measure  $\mu$  on a set  $X$  and the  $\sigma$ -algebra of  $\mu$ -measurable sets  $\Sigma$ .

**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  $\delta_x$ -measurable.

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

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

Every  $A$  is  $\mu$ -measurable.

- The **indiscrete measure** given by

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

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

The Carathéodory criterion of  $\mu$ -measurable sets and the  $\sigma$ -subadditivity of the measure give us some nice properties back.

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

(a)  $\mu$  is  $\sigma$ -additive.

(b) Continuity from below:

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

(c) Continuity from above:

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

*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  $\mu$ , we have

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

The other inequality (and thus equality) follow from  $\sigma$ -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  $\sigma$ -additivity, summation into a telescoping sum

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

- (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

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

□

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

### 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

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

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  $\lambda : \mathcal{K} \rightarrow [0, \infty]$  and any function with  $\lambda(\emptyset) = 0$ . Then this induces a measure  $\mu$  on  $X$  given by

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

*Proof.* Let  $A \subseteq \bigcup_{k=1}^{\infty} A_k$ . We show  $\sigma$ -subadditivity of  $\mu$ , 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  $\mu$ , 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.  $\square$

**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  $\lambda$  in the previous theorem only had minimal restrictions ( $\mathcal{K}$  had to be a covering and  $\lambda : \mathcal{K} \rightarrow [0, \infty]$  with  $\lambda(\emptyset) = 0$ ).

It turns out that if  $\lambda$  and  $\mathcal{K}$  are *nice enough*, then the induced measure is a  $\sigma$ -additive extension of  $\lambda$ . Nice-enough here means that  $\mathcal{K}$  is an algebra and  $\lambda$  is a pre-measure.

Recall that given an algebra  $\mathcal{A} \subseteq \mathcal{P}(X)$ , a function  $\lambda : \mathcal{A} \rightarrow [0, \infty]$  is called a **pre-measure** if it is  $\sigma$ -additive and satisfies  $\lambda(\emptyset) = 0$ .

Given a pre-measure  $\lambda$  on  $\mathcal{A}$ , we can obtain a measure  $\mu$  on  $\mathcal{P}(X)$  that coincides with  $\lambda$  on  $\mathcal{A}$ , i.e.  $\mu$  extends  $\lambda$ .

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

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

it holds that

- (a)  $\mu : \mathcal{P} \rightarrow [0, \infty]$  is a measure.
- (b)  $\mu(A) = \lambda(A), \forall A \in \mathcal{A}$
- (c) All  $A \in \mathcal{A}$  are  $\mu$ -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  $\sigma$ -additivity of the pre-measure  $\lambda$ , 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  $\mu$ , for every  $\varepsilon > 0$  we can choose 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  $(\sigma)$ -additivity of  $\lambda$  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  $\varepsilon \rightarrow 0$  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  $\lambda$  is called  **$\sigma$ -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  $\lambda : \mathcal{A} \rightarrow [0, \infty]$  be a  $\sigma$ -finite pre-measure on  $X$  and  $\mu$  the Carathéodory-Hahn extension of  $\lambda$  and let  $\Sigma$  be the  $\sigma$ -algebra of  $\mu$ -measurable sets.

If  $\tilde{\mu} : \mathcal{P}(X) \rightarrow [0, \infty]$  is another measure with  $\tilde{\mu}|_{\mathcal{A}} = \lambda$ , then  $\tilde{\mu}|_{\Sigma} = \mu$

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

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

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

For the first claim, let  $A \subseteq \bigcup_{k=1}^{\infty} A_k$  with  $A_k \in \mathcal{A}$ . By  $\sigma$ -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  $\mu$ , the inequality still holds:  $\tilde{\mu}(A) \leq \mu(A)$ . Note that we didn't have to use  $\sigma$ -finiteness of  $\lambda$  for this inequality.

For the second claim let  $A \in \Sigma$  be  $\mu$ -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  $\mu$ , it follows that

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

Since  $S \in \mathcal{A}$  is  $\mu$ -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 follows that  $\tilde{\mu}(A) = \mu(A) \leq \tilde{\mu}(A)$ .

In the more general case, we can use  $\sigma$ -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  $m \rightarrow \infty$ , we get

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

□

If we denote  $\tilde{\Sigma}$  to be the  $\sigma$ -algebra 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  $\Sigma$ .

**Example 1.3.8.** Let  $X = [0, 1]$ ,  $\mathcal{A} = \{\emptyset, X\}$  and set  $\lambda(\emptyset) = 0, \lambda(X) = 1$ .

The Carathéodory extension of  $\lambda$  has  $\mu(A)$  to be 0 or 1, depending on if  $A$  is empty or not. The  $\mu$ -measurable sets are  $\Sigma = \{\emptyset, X\}$ .

However, as we will see in the next section, the Lebesgue measure  $L^1$  is also an extension of  $\lambda$  with  $L^1|_{\Sigma} = \mu|_{\Sigma}$ , but they differ when measuring the interval  $[0, \frac{1}{2}]$ .

## 1.4 Lebesgue Measure

The Lebesgue measure is the Carathéodory-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 its **volume** to be

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

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

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

**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

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

*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

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

flipping the roles of  $I_k$  and  $J_j$ , we also get

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

which equals what we got before. □

We of course have to show that our attempt to use the Carathéodory-Hahn Extension of  $\text{vol}$  on the elementary sets is well defined. But it should be easy to see how the class of elementary sets forms an algebra and that the  $\text{vol}$  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

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

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

We call the cubes of the collection

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

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  $\text{vol}(Q) = 2^{-kn}$ .

**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<sup>2</sup>, i.e.

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

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  $\sigma$ -algebra  $\mathcal{B}(X)$  is the  $\sigma$ -algebra generated by open subsets of  $X$ .

**Definition 1.4.6.** A measure  $\mu$  on  $\mathbb{R}^n$  is called **Borel** (or a Borel measure), if every Borel set is  $\mu$ -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 collection of  $\mathcal{L}^n$ -measurable sets form a  $\sigma$ -algebra. So the Borel  $\sigma$ -algebra is contained in the  $\sigma$ -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 unwieldy. The following theorem gives a nicer characterisation.

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

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

<sup>2</sup>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

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

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  $\text{vol}(\tilde{I}_k) \leq \text{vol}(I_k) + \frac{\varepsilon}{2^k}$ .

Setting  $U := \bigcup_{k=1}^{\infty} \tilde{I}_k$ , we see that  $U$  is an open subset containing  $A$  and its volume is

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

since  $\varepsilon$  was arbitrary, the result follows.  $\square$

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

*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  $\varepsilon$  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 Comparison 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

$$\begin{aligned}\underline{\mu}(A) &:= \int_{\mathbb{R}^n} \chi_A d\mu := \sup\{\text{vol}(E) \mid E \subseteq A, E \text{ elementary set}\} \\ \overline{\mu}(A) &:= \int_{\mathbb{R}^n} \chi_A d\mu := \inf\{\text{vol}(E) \mid A \subseteq E, E \text{ elementary set}\}\end{aligned}$$

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  $f : I \rightarrow \mathbb{R}$  continuous,  $I \subseteq \mathbb{R}^n$  compact, its graph

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

is a Jordan measurable set.

The area under a function

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

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

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

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

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

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

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

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  $E^\varepsilon$  is open, so

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

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 Lebesgue measure is invariant under isometries of  $\mathbb{R}^n$ , which are maps

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

*Proof.* Missing □

**Definition 1.5.5.** A Borel measure  $\mu$  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 choose 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  $\sigma$ -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)

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

□

## 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  $f : I \rightarrow \bigcup_{i \in I} A_i$  such that  $f(i) \in A_i$ .

With this, we can construct the set  $\{f(i) \mid i \in I\}$  that contains exactly one element from each set  $A_i$ .

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

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

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

$$\begin{aligned} E_1 &:= E \cap [0, 1 - x) \oplus x \\ E_2 &:= E \cap [1 - x, 1) \oplus x \end{aligned}$$

which are disjoint.

If  $E$  is  $\mathcal{L}^1$ -measurable, 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  $\sigma$ -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

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

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 exist 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  $A = \bigcup_{i=0}^{\infty} B_i$

As we showed earlier, if  $B_i$  were  $\mathcal{L}^1$ -measurable, then  $\mathcal{L}^1(B_i) = 0$ , which contradicts  $\mathcal{L}^1(A) = \sum_{i=0}^{\infty} \mathcal{L}^1(B_i)$ .  $\square$

**Remark 1.6.2.** Because singletons  $\{\alpha\} \in \mathbb{R}$  are contained in the arbitrarily small interval  $(\alpha - \varepsilon, \alpha + \varepsilon)$  with Lebesgue measure  $2\varepsilon$ , 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.

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

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

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

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

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

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

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.

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

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.  $\square$

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  $\ell$ , 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  $\beta$  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

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

## 1.7 The Lebesgue-Stieltjes Measure

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

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

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  $\mathcal{K} = \{[a, b) \mid a, b \in \mathbb{R}\}$  does not form an Algebra (see 1.7.6), we cannot use the Carathéodory-Hahn extension theorem to produce a measure induced by  $\lambda_F$ .

However,  $\mathcal{K}$  constitutes a covering of  $\mathbb{R}$  as in Definition 1.3.1, so by Theorem 1.3.3, the function  $\lambda_F$  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  $\Lambda_F$  is nice, we will find the following definition useful.

**Definition 1.7.1.** A measure  $\mu$  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

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

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  $\mu$  on  $\mathbb{R}^n$  is Borel.

*Proof.* Let  $\mu$  be a metric measure on  $\mathbb{R}^n$ . Because the  $\mu$ -measurable subsets (1.2.9) form a  $\sigma$ -Algebra, it is sufficient to show that closed sets are  $\mu$ -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  $k = 1, 2, \dots$ , we define

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

It should be clear that

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

so since  $\mu$  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  $\lim_{k \rightarrow \infty} \mu(B \cap F_k) = \mu(B \setminus F)$ , 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<sup>3</sup> 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

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

<sup>3</sup>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

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

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

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

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

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

so even in the limit  $m \rightarrow \infty$ , 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  $\lim_{k \rightarrow \infty} \mu(B \setminus F_k) = \mu(B \setminus F)$ , and the result follows.  $\square$

**Theorem 1.7.3.** The Lebesgue-Stieltjes measure  $\Lambda_F$  is Borel regular.

*Proof.* (i)  **$\Lambda_F$  is Borel.** We show that it is metric and use the previous theorem, so let  $A, B \subseteq \mathbb{R}$  with  $\delta := \text{dist}(A, B) > 0$ . 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  $\Lambda_F$  is metric and thus Borel.

(ii)  **$\Lambda_F$  is Borel regular.**

To show that  $\Lambda_F$  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

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

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

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

in the limit  $n \rightarrow \infty$ , we get  $\Lambda_F(A) = \Lambda_F(B)$ , so  $\Lambda_F$  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  $\Lambda_F$ , 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

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

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

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

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  $\varepsilon \rightarrow 0$

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

□

#### Example 1.7.5.

- The Lebesgue measure is the special case when  $F(x) = x$ , so  $\Lambda_{\text{id}_{\mathbb{R}}} = \mathcal{L}^1$
- The Dirac measure  $\delta_0$  from Example 1.2.11 is the Lebesgue-Stieltjes measure  $\Lambda_{\Theta}$  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  $\mathcal{K} = \{[a, b) \mid a, b \in \mathbb{R}\}$  does not form an Algebra. If we set  $\tilde{\mathcal{K}}$  to be the collection of finite disjoint unions:

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

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

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

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 properly measure fractal sets (which can be thought 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

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

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

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

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

since every  $\delta_2$  covering is also a  $\delta_1$  covering. Therefore, the limit

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

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  $\mathcal{H}_\delta^0(A)$  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  $\varepsilon$  small enough. Therefore  $\mathcal{H}_\delta^0(A) = 1$  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  $\delta \rightarrow 0$ , 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  $\mathcal{H}_\delta^1(A) \leq 1 + \varepsilon$ .

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

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

One can convince themselves 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 described in the case  $s = 1$  gives us the upper bound

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

But unlike for  $s = 1$ , choosing bigger and bigger  $n$  means that we can make the  $\mathcal{H}^2$ -measure of  $A$  arbitrarily 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  $\mathcal{H}^0(A) \geq \mathcal{H}_{\frac{\delta}{2}}^0 \geq k$ . 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  $k \rightarrow \infty$ , the proof follows.

And if  $A$  is empty, by definition we have  $\mathcal{H}^0(\emptyset) = 0$ .  $\square$

**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  $\mathcal{H}_{\delta}^s$  is  $\sigma$ -subadditive for all  $\delta > 0$ , we get

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

by taking the limit  $\delta \rightarrow 0$  (as in the definition of  $\mathcal{H}^s$ ) we get the  $\sigma$ -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  $\delta_0 := \text{dist}(A, B) > 0$ . 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  $\mathcal{H}_{\delta}^s$  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

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

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

Therefore, we get

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

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

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

which, when taking the limit  $\delta \rightarrow 0$  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  $\mathcal{H}_{\delta}^s$ , this also means that  $\mathcal{H}_{\delta}^s(A) < \infty$  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

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

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

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

so in the limit  $m \rightarrow \infty$ , 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

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

Considering the infimum over all such coverings we get

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

so as  $\delta \rightarrow 0$ , 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

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

Equivalently, we could have defined it as

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

**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  $\mathcal{H}_\delta^0(\mathbb{S}^1)$ . The more and more we decreased  $\delta$ , the harder it became to cover all points, so in the limit  $\delta \rightarrow 0$ , 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 stretch a set  $A \subseteq \mathbb{R}^n$  by some factor  $\lambda > 0$ , then

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

If we let  $d := \mathcal{H}^d(C)$ , 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  $\mu$  on  $\mathbb{R}^n$  is called a **Radon** measure, if  $\mu$  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  $\delta_0$  is a Radon measure.
- If  $\mu$  is Borel regular and  $A \subset \mathbb{R}^n$  is  $\mu$ -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 analogous to Theorem 1.4.9 holds. For this we need the following Lemma:

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

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

*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  $A, \mu, \varepsilon$  as above.

Since  $\mu$  is Borel-regular, there exists a Borel set  $B \supseteq A$  with  $\mu(B) = \mu(A)$ . By  $\mu$ -measurability of  $A$ :

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

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  $\mu$ -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

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

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

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

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

**Theorem 1.9.4** (Approximation by open and compact sets). Let  $\mu$  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  $A \subseteq \mathbb{R}^n$   $\mu$ -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  $\mu$ -measurable. Since for all  $\varepsilon > 0$  there exists an open set  $U \supseteq A$  with  $\mu(U \setminus A) < \varepsilon$ . By  $\mu$ -measurability 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  $\mu$  is Borel regular, there exists a Borel set  $B \subseteq A$  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  $\mu$ -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

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

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  $\mu(A) \leq \mu(C) + \varepsilon$  and therefore

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

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  $\mu$ -measurable subsets, we also have  $\mu(C) = \lim_{m \rightarrow \infty} \mu(C_m)$ , which means

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

And therefore

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

$\mu(A) = \infty$  : In this case, set  $D_k := \{x \mid k-1 \leq |x| < k\}$ . These disjoint sets can be written as the union of a closed and 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  $\mu$  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

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

---

<sup>4</sup>Since  $\mathbb{R}^n \setminus A$  and  $C$  are also  $\mu$ -measurable, we could also have used equality here.

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  $\sigma$ -additive on the  $\sigma$ -algebra of  $\mu$ -measurable sets

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

This shows that

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

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

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

□

## 2 Measurable Functions

### 2.1 Basic definitions

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

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

And it's easy to show that

- (a)  $f^{-1}(A^c) = (f^{-1}(A))^c$
- (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  $\sigma$ -algebra, then

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

is a  $\sigma$ -algebra in  $X$ .

In the following, let  $\mu$  be a measure on  $\mathbb{R}^n$  and  $\Omega \subseteq \mathbb{R}^n$  be a  $\mu$ -measurable subset.

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

- (a)  $f^{-1}\{+\infty\}, f^{-1}\{-\infty\}$  are  $\mu$ -measurable.
- (b)  $f^{-1}(U)$  for every  $U \subseteq \mathbb{R}$  open is  $\mu$ -measurable.

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

- (c)  $f^{-1}(B)$  is  $\mu$ -measurable for each Borel set  $B \subseteq \mathbb{R}$
- (d)  $f^{-1}((-\infty, a))$  is  $\mu$ -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  $f : \Omega \rightarrow [-\infty, \infty]$ , the following are equivalent:

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

**Example 2.1.3.** Let  $f : \Omega \rightarrow \mathbb{R}$  be  $\mu$ -measurable and  $g : \mathbb{R} \rightarrow \mathbb{R}$  continuous. Then  $g \circ f$  is  $\mu$ -measurable.

**Theorem 2.1.4.**

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

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

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

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

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

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

It's easy to see that  $\chi_A$  is  $\mu$ -measurable if and only if  $A$  is  $\mu$ -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  $\mu$ -measurable, then  $f$  is called a  **$\mu$ -measurable simple function**.

Equivalently, a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $\mu$ -measurable simple function if and only if  $f$  is  $\mu$ -measurable and the image of  $f$  is a countable subset of  $\mathbb{R}$ .

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



**Theorem 2.1.6.** Let  $f : \Omega \rightarrow [0, \infty]$  be  $\mu$ -measurable. Then there exist  $\mu$ -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

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

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

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

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  $\sup\{k \mid x \in A_k\} = \infty$ , then

Missing Rest □

**Proposition 2.1.7.** Let  $f : \Omega \rightarrow \mathbb{R}$  be continuous and  $\mu$  a Borel measure. Then  $f$  is  $\mu$ -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  $\mu$  is Borel,  $f^{-1}(U)$  is  $\mu$ -measurable. □

From now on, we will say that a statement holds “ $\mu$ -a.e.” or “**almost everywhere with respect to  $\mu$** ” 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  $(f_k : [0, 1] \rightarrow \mathbb{R})_{k \in \mathbb{N}}$  given by

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

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  $(f_k : \Omega \rightarrow \overline{\mathbb{R}})_{k \in \mathbb{N}}$  **converges  $\mu$ -almost uniformly** on  $\Omega$  to a function  $f : \Omega \rightarrow \overline{\mathbb{R}}$ , if for all  $\delta > 0$  there exists a  $\mu$ -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:

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

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

- (a) If  $f_k(x) \rightarrow f(x)$  as  $k \rightarrow \infty$  for  $\mu$ -a.e.  $x \in \Omega$ , and  $f(x)$  finite  $\mu$ -a.e., then  $(f_k)_{k \in \mathbb{N}}$  converges  $\mu$ -almost uniformly to  $f$  on  $\Omega$ .
- (b) If additionally,  $\mu$  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

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

namely  $(f_k)_{k \in \mathbb{N}}$  converges uniformly to  $f$  on  $K$ .

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

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

which are  $\mu$ -measurable, since they are the pre-images of open subsets under a  $\mu$ -measurable function. They are also decreasing in  $j$  (i.e.  $C_{i,j+1} \subseteq C_{i,j} \forall i, j$ ) and since  $\mu(\Omega) < \infty$ , we can use continuity from above. Together with  $f_k(x) \rightarrow f(x)$  for  $\mu$ -a.e.  $x$ , we get

$$\lim_{j \rightarrow \infty} \mu(C_{ij}) = \mu \left( \bigcap_{j=1}^{\infty} C_{ij} \right) = 0, \quad \forall i \in \mathbb{N}$$

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

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

setting

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

which is  $\mu$ -measurable, we get the estimate

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

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  $(f_k) \rightarrow f$  on  $A$ .

- (b) By applying Theorem 1.9.4, there exists a compact subset  $K \subseteq A$  such that  $\mu(A \setminus K) < \frac{\delta}{2}$ , and so

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

□

**Remark 2.2.3.** The condition that  $\mu(\Omega) < \infty$  is necessary. Take for example the sequence of “bump functions moving to the right”:  $f_k = \chi_{[k, k+1]} : \mathbb{R} \rightarrow \mathbb{R}$ , 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 continuous functions are Riemann integrable and that Riemann-integrable functions only have measure-zero points of discontinuity.

Because  $\mu$  is assumed to be a Radon measure, all continuous functions are  $\mu$ -measurable. We now show the generalisation of the other fact.

**Theorem 2.2.4** (Lusin's Theorem). Let  $\mu$  be a Radon measure on  $\mathbb{R}^n$ ,  $\Omega \subseteq \mathbb{R}^n$  be  $\mu$ -measurable with  $\mu(\Omega) < \infty$  and  $f : \Omega \rightarrow \overline{\mathbb{R}}$   $\mu$ -measurable and finite  $\mu$ -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  $f|_K : K \rightarrow \overline{\mathbb{R}}$  is continuous and *not* that  $f$  is continuous 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

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

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

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

since  $\mu$  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

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

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  $g_i : D_i \rightarrow \mathbb{R}, g_i(x) = b_{ij}$  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

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

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

The set  $K$  also satisfies

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

and since  $f(x)$  is finite  $\mu$ -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  $\mu$  be an arbitrary measure on  $\mathbb{R}^n$  and  $\Omega \subseteq \mathbb{R}^n$   $\mu$ -measurable and let  $f, f_k : \Omega \rightarrow \overline{\mathbb{R}}$  be  $\mu$ -measurable and  $|f(x)| < \infty$   $\mu$ -a.e..

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

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

**Remark 2.3.2.** If the sequence converges uniformly, then also  $f_k \xrightarrow{\mu} f$ . However pointwise convergence is not enough to show  $f_k \xrightarrow{\mu} f$ , as the same Example as in 2.2.3 proves otherwise.

Also, the sequence  $f_k = \chi_{\{0\}} \xrightarrow{\mathcal{L}^1} \mathbf{0}$  shows that pointwise convergence does not necessarily follow from in-measure convergence.

**Theorem 2.3.3.** Let  $\mu(\Omega) < \infty$ . If  $f_k \rightarrow f$   $\mu$ -a.e. then  $f_k \xrightarrow{\mu} f$ .

*Proof.* By Egoroff's theorem,  $(f_k)_{k \in \mathbb{N}}$  converges  $\mu$ -almost uniformly on  $\Omega$ .

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. □

<sup>4</sup>In contrast to the lecturer, I will be using  $f_k \xrightarrow{\mu} f$  as shorthand for “ $f_k \xrightarrow{\mu} f$  as  $k \rightarrow \infty$ ” as it should be clear from context.

**Remark 2.3.4.** The converse of this theorem does not hold, so the statement  $f_k \rightarrow f$   $\mu$ -a.e. is stronger than  $f_k \xrightarrow{\mu} f$ .

To see this, take  $\Omega = [0, 1)$  with the measure  $\mathcal{L}^1$ . And set  $f_k = \chi_{A_k}$ , for

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

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

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

But the nowhere does the sequence converge pointwise to 0.

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

*Proof.* Since  $f_k \xrightarrow{\mu} f$ , it means that

□