Contents

Chapter 1

Introduction to Measures

1.1 Lebesgue's Problem of measure

(This section is taken from Hendrik Grundling's 2011 lecture.) Assigning a measure (length, area, volume, ...) to a subset S of $\mathbb{R}^d$ is necessary in mathematics, e.g. for any type of integral. A reasonable measure $m(S)$ of S should satisfy the following requirements:

1. If $S \subset \mathbb{R}^d$ is congruent (after shifts, rotations & reflections) to $T \subset \mathbb{R}^d$, then $m(S) = m(T)$
2. If $S = \bigcup_{i=1}^{\infty} S_i$ where $S_i \cap S_j = \emptyset$ for $i \neq j$, then $m(S) = \sum_{i=1}^{\infty} m(S)$ (countable additivity)
3. $m(I) = 1$ for $I = [0,1]^d \subset \mathbb{R}^d$ (normedness)
4. $m(S) \geq 0$

However this humble attempt is still too much to ask as the following shall show. Below, we construct a set S which cannot satisfy all four requirements.

For $x, y \in [0,1]$ define $x \sim y$ if and only if $x - y \in \mathbb{Q}$. This is an equivalence relation, so $[0,1]$ is partitioned into (disjoint) equivalence classes. Define the "Vitali set" $S \subset [0,1]$ by choosing one point from each equivalence class (this assumes the Axiom of Choice).

Lemma 1.1. If $r, q \in \mathbb{Q} \cap [0,1]$, $r \neq q$ then $(S + r) \cap (S + q) = \emptyset$.

Proof. If $x \in (S + r) \cap (S + q)$ then $x = p + r = t + q$ for $p, t \in S$. Then $p - t = q - r \neq 0$ and $q - r \in \mathbb{Q}$. So $p \sim t$, $p \neq t$. Since $p, t \in S$, this violates the definition of S. $\square$

Lemma 1.2. $[0,1] \subseteq \bigcup\{S + r | r \in \mathbb{Q} \cap [-1,1]\} =: T \subseteq [-1,2]$.

Proof. Let $x \in [0,1)$. Then $x \sim p$ for some $p \in S$. Thus $x - p =: r \in \mathbb{Q}$ and as $x, p \in [0,1]$ it follows that $r \in \mathbb{Q} \cap [-1,1]$. So $x \in S + r$, $r \in \mathbb{Q} \cap [-1,1]$. $\square$

Proposition 1.3. In $\mathbb{R}$, Lebesgue's problem of measure has no solution.

Proof. Let S be as above. We show a contradiction to the requirement $m(S+r) = m(S)$ for all r.

$$\begin{aligned} 1 = m([0,1]) = m(T) - m(T \setminus [0,1]) \leq m(T) \\ = \sum_{r \in \mathbb{Q} \cap [-1,1]} m(S+r) = \sum_{r \in \mathbb{Q} \cap [-1,1]} m(S) \end{aligned}$$

$m(S)$ can't be 0, hence the right-hand side must be $+\infty$. But this contradicts

$$\begin{aligned} \infty = m(T) &= m([-1,2]) - m([-1,2) \setminus T) \leq m([-1,2]) \\ &= m([-1,0]) + m([0,1]) + m([1,2]) - m(\{0,1\}) \leq 3m([0,1]) = 3 \end{aligned}$$

$\square$

In higher dimensions, things are even worse:

Theorem 1.4 (Banach-Tarski)**.** Let $S, T \subset \mathbb{R}^d$ be bounded with non-empty interiors, $d \geq 3$. Then there is a $k \in N$ and partitions $\{E_i, \ldots, E_k\}$, $\{F_i, \ldots F_k\}$ of S, T respectively such that E_i is congruent to F_i $\forall i$. (That is, S and T are equidecomposable.)

Thus we can take e.g. a unit sphere S, cut it up into finitely many pieces, and reassemble it to a set T of two unit spheres. If Lebesgue's problem of measure had a solution in $\mathbb{R}^d$, $d \geq 3$, then by congruence $m(S) = m(T)$ which contradicts $m(T) = 2m(S)$ and $m(S) > 0$.

A way out of this situation is to restrict attention to a smaller set of subsets, on which Lebesgue's problem of measure does have a solution.

1.2 σ-algebras

Definition 1.5. Let X be any non-empty set. A system $\mathcal{A}$ of subsets of X is called an algebra on X if it has the following properties:

1. $X \in \mathcal{A}$,
2. If $\{A, B\} \subset \mathcal{A}$ then $A \cup B \in \mathcal{A}$
3. If $A \in \mathcal{A}$ then $A^{\complement} \in \mathcal{A}$ (where $A^{\complement} = X \setminus A$).

If additionally the following condition holds, then $\mathcal{A}$ is called a σ-algebra:

4. If $\{A_1, A_2, \ldots\} \subset \mathcal{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

A measurable space is any pair $(X, \mathcal{A})$ where $\mathcal{A}$ is a σ-algebra on X.

Algebras are closed under finite intersections: If $\{A, B\} \subset \mathcal{A}$, then

$$(A \cap B)^{\complement} = A^{\complement} \cup B^{\complement} \in \mathcal{A}.$$

Similarly, σ-algebras are closed under countably infinite intersections.

Definition 1.6. Let X be a non-empty set. A system $\mathcal{C} \subset 2^X$ of subsets of X is called a ring if $\emptyset \in \mathcal{C}$ and $A \cup B, A \setminus B \in \mathcal{C}$ for any $A, B \in \mathcal{C}$.

Every σ-algebra is an algebra; every finite algebra is a σ-algebra.

Example 1.7. Let X be any non-empty set, and consider the following subsets of 2^X:

1. $\mathcal{A} = \{\emptyset, X\}$. This is the "coarsest", "smallest" or "trivial" σ-algebra on X (and thus also an algebra and a ring).
2. $\mathcal{A} = \{\emptyset, A, A^\complement, X\}$, where $A \subset X$. This is a σ-algebra, an algebra and a ring.
3. The power set 2^X, i.e. the family of all subsets of X, where X is any set (the "finest" or "largest" σ-algebra on X).

4. Let $X = \mathbb{R}$. Then

$$\mathcal{C} := \left\{ \bigcup_{k=1}^{n} (a_k, b_k] : n \in \mathbb{N}, a_k, b_k \in \mathbb{R}, a_k \leq b_k, b_k \leq a_{k+1} \right\} \subset 2^{\mathbb{R}}. \tag{1.8}$$

is a ring, but not an algebra. $\mathcal{C} \cup \mathbb{R}$ is an algebra, but not a σ-algebra. We shall later see that the σ-algebra generated by $\mathcal{C}$ is the Borel-σ-algebra on $\mathbb{R}$. □

If two algebras (or σ-algebras) $\mathcal{A}_1, \mathcal{A}_2$ are defined on the same set X, then their intersection (not necessarily their union) is again an algebra (or σ-algebra) (try a simple example). This even holds for *arbitrary* (uncountably infinite) intersections.

Lemma 1.9. Let X be any set, and let $\mathcal{E}$ be a system of subsets of X. Then there exist

1. a smallest algebra $\alpha(\mathcal{E}) \supset \mathcal{E}$
2. a smallest σ-algebra $\sigma(\mathcal{E}) \supset \mathcal{E}$

Proof. Let $\mathcal{S}_\alpha$ be the set of algebras which contain $\mathcal{E}$, and let $\mathcal{S}_\sigma$ be the set of σ-algebras which contain $\mathcal{E}$. Then

$$2^X \in \mathcal{S}_\sigma \subset \mathcal{S}_\alpha,$$

where 2^X is the power set of X (the set of all possible subsets). In particular neither $\mathcal{S}_\alpha$ nor $\mathcal{S}_\sigma$ are empty. Then

$$\alpha(\mathcal{E}) = \bigcap_{\mathcal{A}\in\mathcal{S}_\alpha} \mathcal{A} \qquad\qquad \sigma(\mathcal{E}) = \bigcap_{\mathcal{A}\in\mathcal{S}_\sigma} \mathcal{A}$$

are as required. □

The algebra and σ-algebra $\alpha(\mathcal{E})$ and $\sigma(\mathcal{E})$ are said to be generated by $\mathcal{E}$.

1.3 Measures

Definition 1.10. A countably additive function μ from a σ-algebra $\mathcal{A}$ of subsets of X into $[0, \infty]$ is called a measure. Then $(X, \mathcal{A}, \mu)$ is called a measure space.

Example 1.11.

1. For a set X with σ-algebra $\mathcal{A}$, define

$$\begin{aligned} \mu : \mathcal{A} &\to [0, \infty] \\ A &\mapsto \mu(A) = \text{ number of elements in A} \end{aligned}$$

   Such μ is called the *counting measure*.

2. For a σ-algebra $\mathcal{A}$ of X, fix $x \in X$ and define

$$\begin{aligned} \delta_x : \mathcal{A} &\to [0, \infty] \\ A &\mapsto \delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \end{aligned}$$

   Then δ_x is called the *Dirac measure* concentrated at x.

3. For a σ-algebra $\mathcal{A}$ on X let $\mu(\emptyset) = 0$ and $\mu(A) = \infty$ if $A \neq \emptyset$, $A \in \mathcal{A}$. Then μ is called the trivial measure.

The aim of this section is to construct a measure on a suitable σ-algebra of $\mathbb{R}$ which solves Lebesgue's problem of measure. Other measurable spaces which are important in probability theory are:

$(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$: This space is used for probabilistic models of experiments with infinitely many steps, that is, stochastic processes in discrete time steps.

$(D(\mathbb{R}^d), \mathcal{D})$: Here $D(\mathbb{R}^d)$ denotes the set of càdlàg (right-continuous with left limits) paths (mappings from $[0, \infty)$ to $\mathbb{R}^d$) with countably many jumps. This set can be equipped with a metric and hence a topology. The corresponding Borel-σ-algebra is $\mathcal{D}$. A probability measure ($\mu(D(\mathbb{R}^d)) = 1$) on this measurable space then governs the behaviour of a stochastic process in continuous time.

Theorem 1.12 (Basic properties of measures)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Then

1. $\mu(\emptyset) = 0$
2. $A \subset B \Rightarrow \mu(A) \leq \mu(B)$
3. If $A_1 \subset A_2 \subset A_3 \subset \ldots$, then
$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right)$$
4. If $A_1 \supset A_2 \supset A_3 \supset \ldots$ where $\mu(A_1) < \infty$, then
$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_{n=1}^{\infty} A_n\right)$$

Outer Measure

Definition 1.13. Let X be a non-empty set. Let $\mathcal{A} \subset 2^X$, $\emptyset \in \mathcal{A}$, and for any $E \in 2^X$, there exists a covering $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$ so that $E \subset \bigcup_{n=1}^{\infty} A_n$. Let $\mu : \mathcal{A} \to [0, \infty]$ be any map such that $\mu(\emptyset) = 0$. Then

$$\begin{aligned} \mu^* : 2^X &\to [0, \infty] \\ E &\mapsto \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) | A_n \in \mathcal{A}, E \subset \bigcup_{n=1}^{\infty} A_n \right\} \end{aligned}$$

is called the outer measure defined by μ and $\mathcal{A}$. (A μ as above is called a *pre-measure*.)

If $A \in 2^X$ satisfies

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \setminus A) \quad \forall B \in 2^X,$$

then A is called μ^*-measurable.

Lemma 1.14 (Subadditivity of μ^*)**.** For any sets E and $E_n \subset X$, if $E \subset \bigcup_n E_n$, then $\mu^*(E) \leq \sum_n \mu^*(E_n)$.

Proof. WLOG assume the right-hand side is not $+\infty$. Let $\varepsilon > 0$, and suppose $A_{nm} \in \mathcal{A}$ are such that $\bigcup_m A_{nm} \supset E_n$, and

$$\sum_m \mu(A_{nm}) < \mu^*(E_n) + \frac{\varepsilon}{2^n}.$$

Then $E \subset \bigcup_n \bigcup_m A_{nm}$, and

$$\mu^*(E) \leq \sum_n \sum_m \mu(A_{nm}) < \sum_n \left(\mu^*(E_n) + \frac{\varepsilon}{2^n}\right) = \sum_n \mu^*(E_n) + \varepsilon$$

Since the above holds for arbitrarily small ε, the statement follows. $\square$

Theorem 1.15 (Carathéodory)**.** The μ^*-measurable sets $\mathcal{M}(\mu^*)$ form a σ-algebra on X, and μ^* is a measure on $\mathcal{M}(\mu^*)$.

Proof. Check that $F \in \mathcal{M}(\mu^*)$ iff $F^{\complement} \in \mathcal{M}(\mu^*)$. Suppose that $A, B \in \mathcal{M}(\mu^*)$. For any $E \in 2^X$, we have

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \setminus A) \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \setminus B) + \mu^*(E \setminus A) \\ &= \mu^*(E \cap (A \cap B)) + \mu^*(E \setminus (A \cap B)),\end{aligned}$$

using that $A \in \mathcal{M}(\mu^*)$ in equality 1 & 3 and that $B \in \mathcal{M}(\mu^*)$ in equality 2. Thus $\mathcal{M}(\mu^*)$ is an algebra (right?). Now let $E_n \in \mathcal{M}(\mu^*)$ for $n = 1, 2, \ldots, F := \bigcup_{j=1}^{\infty} E_j$, and $F_n := \bigcup_{j=1}^{n} E_j \in \mathcal{M}(\mu^*)$. Since $E_n \setminus \bigcup_{j<n} E_j \in \mathcal{M}(\mu^*)$ for all n, we may assume E_n disjoint in proving F measurable. For

any $E \subset X$ we have

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \setminus F_n) + \mu^*(E \cap F_n) \\
&= \mu^*(E \setminus F_n) + \mu^*(E \cap F_n \cap E_n) + \mu^*(E \cap F_n \setminus E_n) \\
&= \mu^*(E \setminus F_n) + \mu^*(E \cap E_n) + \mu^*(E \cap F_{n-1})
\end{aligned}$$

by μ^*-measurability of F_n and E_n. We have shown

$$\mu^*(E \cap F_n) = \mu^*(E \cap E_n) + \mu^*(E \cap F_{n-1})$$

which inductively means $\mu^*(E \cap F_n) = \sum_{j=1}^n \mu^*(E \cap E_j)$. Thus

$$\mu^*(E) = \mu^*(E \setminus F_n) + \sum_{j=1}^{n} \mu^*(E \cap E_j) \geq \mu^*(E \setminus F) + \sum_{j=1}^{n} \mu^*(E \cap E_j).$$

As this holds for every n, we have

$$\mu^*(E) \geq \mu^*(E \setminus F) + \sum_{j=1}^{\infty} \mu^*(E \cap E_j) \geq \mu^*(E \setminus F) + \mu^*(E \cap F) \quad (1.16)$$

using Lemma 1.14. Again using Lemma 1.14, we have $\mu^*(E) = \mu^*(E \setminus F) + \mu^*(E \cap F)$, and thus $F \in \mathcal{M}(\mu^*)$. Hence $\mathcal{M}(\mu^*)$ is a σ-algebra. Setting $E = F$ in (1.16) shows $\mu^*(F) = \sum_{j=1}^{n} \mu^*(F \cap E_j)$ and hence that μ^* is a measure on F. $\square$

1.4 Lebesgue Measure

Consider the ring of subsets

$$\mathcal{C} = \{(a, b] : a, b \in \mathbb{R}, a \leq b\} \tag{1.17}$$

and the pre-measure $\gamma : \mathcal{C} \to [0, \infty)$ given by

$$\gamma((a, b]) = b - a. \tag{1.18}$$

Then $\mathcal{C}$ and γ define an outer measure λ^* on $X = \mathbb{R}$. This outer measure is called the Lebesgue outer measure. The Lebesgue measure can then be defined as the measure λ on the measurable space $(\mathbb{R}, \mathcal{M}(\lambda^*))$, which is constructed as in Theorem 1.15. The elements of $\mathcal{M}(\lambda^*)$ are called the Lebesgue-measurable sets.

Below we define the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, the "standard" σ-algebra on $\mathbb{R}$. Its elements are the Borel-measurable sets, or Borel sets, or simply "measurable" sets. It is possible to develop an intuition for the Borel-σ-algebra, since it is generated by the open intervals: We can construct Borel sets through countable unions and intersections of open and closed intervals. As it turns out (in this section and the next), $\mathcal{M}(\lambda^*)$ is almost equal to $\mathcal{B}(\mathbb{R})$.

Definition 1.19 (Topology)**.** Let X be a non-empty set, and let $\mathcal{T} \subset 2^X$ be a system of subsets of X with the following properties:

1. $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$.
2. If $A, B \in \mathcal{T}$, then $A \cap B \in \mathcal{T}$.
3. If $A_i \in \mathcal{T}$ for all $i \in I$ (I is not necessarily countable), then $\bigcup_{i \in I} A_i \in \mathcal{T}$

Then $\mathcal{T}$ is called a topology on X. The pair $(X, \mathcal{T})$ is called a topological space, and the elements of $\mathcal{T}$ are called the open sets in X. A set $A \subset X$ is called closed if $A^{\complement}$ is open. The σ-algebra $\sigma(\mathcal{T})$ is called the Borel σ-field on X. The elements of $\sigma(\mathcal{T})$ are called Borel sets.

Example 1.20. Let $X = \mathbb{R}$, and let $\mathcal{T}$ be the system of countably infinite unions of intervals $\{(a, b) : a, b \in \mathbb{R} \cup \{-\infty, +\infty\}\}$. Then $(\mathbb{R}, \mathcal{T})$ is a topological space, and $\sigma(\mathcal{T})$ is the Borel σ-field $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$. The Borel σ-algebra is generated by any of the following systems:

$$\mathcal{E}_1 = \{(a, b) : a, b \in \mathbb{R}\}, \quad \mathcal{E}_2 = \{[a, b] : a, b \in \mathbb{R}\},$$
$$\mathcal{E}_3 = \{(a, b] : a, b \in \mathbb{R}\}, \quad \mathcal{E}_4 = \{[a, b) : a, b \in \mathbb{R}\}$$

Theorem 1.21. Borel sets are Lebesgue-measurable: $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\lambda^*)$.

Proof. Since we know that $\mathcal{M}(\mu^*)$ is a σ-algebra, it suffices to show that it contains a generator of $\mathcal{B}(\mathbb{R})$, such as the ring $\mathcal{C}$ of half-open intervals $(a, b]$ (1.17). That is, we show

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

for any interval $I = (a, b] \in \mathcal{C}$ and any subset $B \in 2^X$.

Let $\{I_k\}_{k\in\mathbb{N}} \subset \mathcal{C}$ be a covering of B such that $\sum_k \gamma(I_k) \leq \mu^*(B)+\varepsilon$, where γ is the pre-measure (1.18) and ε is any positive small number. For each k, let $\{E_{k\ell}\}_{\ell\in\mathbb{N}} \subset \mathcal{C}$ and $\{F_{k\ell}\}_{\ell\in\mathbb{N}} \subset \mathcal{C}$ be a covering of $B \cap I$ and $B \setminus I$, respectively, such that

$$\sum_{\ell} \gamma(E_{k\ell}) \leq \mu^*(B \cap I_k) + 2^{-k}\varepsilon \quad \sum_{\ell} \gamma(F_{k\ell}) \leq \mu^*(B \setminus I_k) + 2^{-k}\varepsilon.$$

Then

$$\begin{aligned}\mu^*(B \cap I) + \mu^*(B \setminus I) &\leq \sum_k \sum_\ell \left[\gamma(E_{k\ell}) + \gamma(F_{k\ell})\right] \\ &\leq \sum_k \left[\gamma(I_k \cap I) + 2^{-k}\varepsilon + \gamma(I_k \setminus I) + 2^{-k}\varepsilon\right] \\ &= \sum_k \left[\gamma(I_k) + 2^{-k+1}\right] \leq \mu^*(B) + 2\varepsilon.\end{aligned}$$

and as the above inequality holds for arbitrary $\varepsilon > 0$, the result follows. $\square$

In the above proof, the decisive step was the equality $\gamma(I_k) = \gamma(I_k \cap I) + \gamma(I_k \setminus I)$, which relies on the assumption that $\mathcal{C}$ is a ring; the remaining steps work for any system $\mathcal{C}$ and any pre-measure γ.

The Lebesgue measure $\lambda(A)$ is defined for all Lebesgue sets $A \in \mathcal{M}(\mu^*)$; by restricting the domain of λ from $\mathcal{M}(\mu^*)$ to $\mathcal{B}(\mathbb{R})$, we define the Lebesgue measure for the Borel sets. This is a different measure (since the domains are different), but we still denote it by λ.

Uniqueness of Lebesgue Measure

To uniquely determine a measure μ on a measurable space $(X, \mathcal{A})$, it is not necessary to know all values $\{\mu(A) : A \in \mathcal{A}\}$. As it turns out, it suffices to determine $\mu(A)$ only for A from a much smaller subset of $\mathcal{A}$. This result is proved using the *monotone class theorem* due to Dynkin. It is a powerful tool for the characterisation of measures, and we will use it to prove the uniqueness of Lebesgue measure λ.

Definition 1.22. Let X be a non-empty set. A d-class on X is a system of subsets of X, $\mathcal{D} \subset 2^X$, which satisfies the following:

1. $X \in \mathcal{D}$
2. If $A, B \in \mathcal{D}$ such that $B \subset A$, then $A \setminus B \in \mathcal{D}$
3. For any increasing sequence $A_1 \subset A_2 \subset \dots$ where $A_j \in \mathcal{D}$, one has $\bigcup_{j=1}^{\infty} A_j \in \mathcal{D}$

A π-class on X is a system of subsets of X, $\pi \subset 2^X$, which is closed with respect to finite intersections. That is, it satisfies $A, B \in \pi \Rightarrow A \cap B \in \pi$.

Any σ-algebra is a d-class. Similarly to algebras and σ-algebras, the intersection of arbitrarily many d-classs on X is again a d-class. Hence we can say that a d-class $\mathcal{D} = d(\mathcal{C})$ is generated by $\mathcal{C} \subset 2^X$ if $\mathcal{D}$ is the smallest d-class containing $\mathcal{C}$.

Theorem 1.23 (Monotone Class Theorem, Dynkin)**.** Let X be a non-empty set, and let $\mathcal{C}$ be a π-class on X. Then $\sigma(\mathcal{C}) = d(\mathcal{C})$.

Proof. Since $\sigma(\mathcal{C})$ is a d-class, we have $\sigma(\mathcal{C}) \supset d(\mathcal{C})$. It remains to show $\sigma(\mathcal{C}) \subset d(\mathcal{C})$, which follows if we can show that $d(\mathcal{C})$ is a σ-algebra. We check that $d(\mathcal{C})$ indeed satisfies the first two properties required of a σ-algebra. For the last property, let $E_1, E_2, \ldots$ be a sequence (not necessarily increasing) of sets in $d(\mathcal{C})$. Since

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{k=1}^{\infty} F_k$$

where $F_k = \bigcup_{j=1}^{k} E_j$, we can write the countable union of E_j as an *increasing* countable union of F_j. Since $d(\mathcal{C})$ is a d-class, $\bigcup_{k=1}^{\infty} F_k$ will lie in $d(\mathcal{C})$; but we need to check that $F_k \in d(\mathcal{C})$. Since $F_k^\complement = \bigcap_{j=1}^{k} E_j^\complement$, it hence suffices to show that $d(\mathcal{C})$ is closed under *finite* intersections. First, define

$$\mathcal{D}_1 := \{E \in d(\mathcal{C}) : E \cap C \in d(\mathcal{C})\, \forall C \in \mathcal{C}\}$$

By definition of $\mathcal{C}$, we have $\mathcal{C} \subset \mathcal{D}_1$. The identities

$$(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C)$$
$$\left(\bigcup_{n=1}^{\infty} A_n\right) \cap C = \bigcup_{n=1}^{\infty} (A_n \cap C)$$

show that $\mathcal{D}_1$ is closed with respect to proper set differences and countable unions of increasing sets, that is, $\mathcal{D}_1$ is a d-class containing $\mathcal{C}$ and $\mathcal{D}_1 \supset d(\mathcal{C})$. By definition, $\mathcal{D}_1 \subset d(\mathcal{C})$, and so $\mathcal{D}_1 = d(\mathcal{C})$. Next, define

$$\mathcal{D}_2 := \{E \in d(\mathcal{C}) : E \cap F \in d(\mathcal{C}) \, \forall F \in d(\mathcal{C})\}$$

We see that $\mathcal{C} \subset \mathcal{D}_2$ and $X \in \mathcal{D}_2$. By the same arguments as above, $\mathcal{D}_2 = d(\mathcal{C})$. This shows that $d(\mathcal{C})$ is closed with respect to finite intersections. $\square$

Definition 1.24. Let $(X, \mathcal{A}, \mu)$ be a measure space. If there is a sequence $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $X = \bigcup_n E_n$ and $\mu(E_n) < \infty$ for all n, then the measure μ is said to be σ-finite.

For instance, Lebesgue measure is σ-finite on the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$; See this by setting $E_n = (-n, n]$. Moreover, any finite measure space with $\mu(X) < \infty$ is σ-finite.

Theorem 1.25. Let $(X, \mathcal{A})$ be a measurable space, and let $\mathcal{C}$ be a π-system on X such that $\mathcal{A} = \sigma(\mathcal{C})$. If μ and ν are measures on $(X, \mathcal{A})$ that agree on $\mathcal{C}$ and if there is an increasing sequence $\{C_n\}$ of sets that belong to $\mathcal{C}$, have finite measure under μ and ν, and satisfy $\bigcup_n C_n = X$, then $\mu = \nu$.

Proof. First, let's assume that μ and ν are finite measures, that is, $\mu(X) < \infty$ and $\nu(X) < \infty$. Define $\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$. We see $\mathcal{C} \subset \mathcal{D}$. Then $\mathcal{D}$ is a d-class because:

1. $\mu(X) = \nu(X) \Rightarrow X \in \mathcal{D}$
2. Let $A, B \in \mathcal{D}$ such that $A \subset B$. Then $\mu(A \setminus B) = \mu(A) - \mu(B) = \nu(A) - \nu(B) = \nu(A \setminus B)$, and thus $A \setminus B \in \mathcal{D}$. (This step assumes that the measures are finite, since $\infty - \infty$ is to be avoided.)
3. Let $E_n \in \mathcal{D}$ be an increasing sequence in $\mathcal{D}$. By Theorem 1.12,
$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n) = \lim_{n\to\infty} \nu(E_n) = \nu\left(\bigcup_{n=1}^{\infty} E_n\right)$$

and hence $\bigcup_{n=1}^{\infty} E_n \in \mathcal{D}$.

By Theorem 1.23, $\mathcal{A} = \mathcal{D}$, thus $\mu = \nu$.

Now assume that μ and ν are σ-finite with respect to the sequences $\{E_n\}_{n\in\mathbb{N}}$ and $\{F_n\}_{n\in\mathbb{N}}$; then both are σ-finite with respect to the the sequence $\{C_n\}_{n\in\mathbb{N}}$ where $C_n = E_n \cap F_n$. For each n, define measures μ_n and ν_n on $(X, \mathcal{A})$ by $\mu_n(A) = \mu(A \cap C_n)$ and $\nu_n(A) = \nu(A \cap C_n)$. From the first part of the proof, we have $\mu_n = \nu_n$, and thus using Th 1.12

$$\mu(A) = \lim_n \mu_n(A) = \lim_n \nu_n(A) = \nu(A)$$

for every $A \in \mathcal{A}$. $\square$

Theorem 1.26. The Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is the unique measure which satisfies $\lambda((a, b]) = b - a$ for all $a, b \in \mathbb{R}$, $a \leq b$.

Proof. First, we show $\lambda((a, b]) = b - a$. Indeed, consider an interval $I = (a, b]$. It is covered by the sequence $I \cup \emptyset \cup \emptyset \cup \ldots$ of elements in the ring $\mathcal{C}$, which shows $\lambda^*(I) \leq \gamma(I) = b - a$. It is intuitively clear (but technically requires a proof) that this covering is optimal. That is: *any* countable covering of I with sets $\{C_n\}_{n \in \mathbb{N}} \subset \mathcal{C}$ is such that $\sum_n \gamma(C_n) \geq b - a$. But this means $\lambda^*(I) \geq b - a$, and so $\lambda^*(I) = b - a$. I is a Borel set, and thus also a Lebesgue set, hence $\lambda(I) = \lambda^*(I) = b - a$. We know that λ is σ-finite; any other measure which maps $(a, b]$ to $b - a$ is also σ-finite. Th 1.25 completes the proof. $\square$

Since the pre-measure γ and thus the outer Lebesgue-measure λ^* are translation invariant, it follows that Lebesgue measure is translation invariant. In fact, up to a multiplicative constant, Lebesgue measure is the only translation invariant measure on $\mathbb{R}$:

Theorem 1.27. Let ν be a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ which is finite on bounded sets and translation invariant (that is, $\nu(A+x) = \nu(A)\,\forall A \in \mathcal{B}(\mathbb{R}), x \in \mathbb{R}$). Then ν is a multiple of the Lebesgue measure.

Proof. The interval $I = (0,1]$ is the pairwise disjoint union of n intervals $((k-1)/n, k/n]$, $k = 1, \dots, n$. These are all translates of each other and hence all have the same measure as any fixed one, say B. Set $c := \nu(I)$, then

$$n\nu(B) = c = c \times 1 = c\lambda(I) = c\,n\,\lambda(B),$$

and $\nu(B) = c\,\lambda(B)$. Any $C \in \mathcal{C}$ can be represented as a countable union of intervals of length $1/n$ $(n \in \mathbb{N})$. The family of such intervals hence generates $\mathcal{B}(\mathbb{R})$. Hence by Th 1.25, $\nu = c\lambda$. $\square$

Finally, here are a few properties of Lebesgue measure:

Theorem 1.28. Lebesgue measure on $\mathbb{R}$ has the following properties:

1. If $A \in 2^X$ is countable, then it is a Borel set, and $\lambda(A) = 0$
2. If $A \in \mathcal{M}(\lambda^*)$ is bounded, then $\lambda(A) < \infty$
3. For every $A \in \mathcal{M}(\lambda^*)$ we have $\lambda(A) = \inf\{\lambda(U) : U \text{ open }, U \supset A\}$
4. For every $A \in \mathcal{M}(\lambda^*)$ we have $\lambda(A) = \sup\{\lambda(K) : K \text{ compact }, K \subset A\}$. (Recall that a set $K \in 2^X$ is compact iff it is both closed and bounded.)

Proof. Properties 1. and 2. are left as an exercise. For 3. and 4, note that monotonicity of measures ($A \subset B \Rightarrow \lambda(A) \leq \lambda(B)$) implies

$$\lambda(A) \leq \inf\{\lambda(U) : U \text{ open }, U \supset A\},$$
$$\lambda(A) \geq \sup\{\lambda(K) : K \text{ compact }, K \subset A\}$$

and it remains to show the reverse inequalities. For part 3, we may assume $\lambda(A) < \infty$. Let $\varepsilon > 0$, and find a covering $\bigcup C_k$ of A where $C_k \in \mathcal{C}$ are half open intervals, such that

$$\sum_k \gamma(C_k) < \lambda(A) + \varepsilon.$$

Replace $C_k = (a_k, b_k]$ by $R_k = (a_k, b_k + 2^{-k}\varepsilon]$ to get a covering such that

$$\sum_k \gamma(R_k) < \lambda(A) + 2\varepsilon.$$

The open set $U = \bigcup_k (a_k, b_k + 2^{-k}\varepsilon)$ contains A, and $\lambda(U) \leq \lambda(A) + 2\varepsilon$. As ε was arbitrary, 3. is proved.

For part 4., assume first that A is bounded. Let C be a closed and bounded set that includes A, and let ε be an arbitrary positive number. Use part 3. to choose an open set U that includes $C \setminus A$ and satisfies

$$\lambda(U) < \lambda(C \setminus A) + \varepsilon.$$

Let $K = C \setminus U$. (Draw a picture!) Then K is a closed and bounded (and hence compact) subset of A; it satisfies $C \subset K \cup U$, and so satisfies

$$\lambda(C) \leq \lambda(K) + \lambda(U)$$

The two inequalities (together with $\lambda(C \setminus A) = \lambda(C) - \lambda(A)$ now imply that $\lambda(A) - \varepsilon < \lambda(K)$. Since ε was arbitrary, part 4 is proved in the case where A is bounded. Finally, consider the case where A is not bounded. Suppose that b is a real number that satisfies $b < \lambda(A)$; we shall produce a compact subset K of A such that $b < \lambda(K)$. Let $\{A_j\}$ be an increasing sequence of bounded measurable subsets of A such that $A = \bigcup_j A_j$ (for example, A_j could be defined to be $A \cap [-j, +j]$. By Th 1.12, $\lambda(A) = \lim_j \lambda(A_j)$, and so we can choose j_0 so that $\lambda(A_{j_0}) > b$. Now apply to A_{j_0} what was shown above for bounded A, obtaining a compact subset of A_{j_0} such that $\lambda(K) > b$. Since $K \subset A$ and since b was an arbitrary number less than $\lambda(A)$, the proof is complete. □

1.5 Completion of Measures

In this section, we establish a result which shows that Lebesgue sets can be very well approximated by Borel sets: For every Lebesgue set, there is a Borel subset and a Borel superset of the same measure. We show that the Lebesgue σ-algebra results from the Borel σ-algebra by completion. Completed probability spaces are of interest in probability theory.

Definition 1.29. Let $(X, \mathcal{A}, \mu)$ be a measure space.

1. A subset $B \in 2^X$ is called μ-null if there is $A \in \mathcal{A}$ such that $B \subset A$ and $\mu(A) = 0$.

2. If $\mathcal{A}$ contains all μ-null sets, then μ and $(X, \mathcal{A}, \mu)$ are called complete.

3. A property holds μ-almost everywhere if the set of $x \in X$ for which it doesn't hold is a μ-null set.

A measure space $(X, \mathcal{A}, \mu)$ which is not complete can be completed, in a unique well-defined way. First, define the completion of $\mathcal{A}$ under μ as the system $\mathcal{A}_\mu$ of all subsets $A \subset X$ for which there exist $E, F \in \mathcal{A}$ such that $E \subset A \subset F$ and $\mu(F \setminus E) = 0$. Members of $\mathcal{A}_\mu$ are sometimes called μ-measurable.

To define a candidate for a measure $\bar{\mu}$ on $\mathcal{A}_\mu$, let $A \in \mathcal{A}_\mu$, where $E, F \in \mathcal{A}$ are such that $E \subset A \subset F$, and set $\bar{\mu}(A) := \mu(E)$. This does not depend on the choice of E of F, hence $\bar{\mu}$ is well defined.

Proposition 1.30. Let $(X, \mathcal{A}, \mu)$ be a measure space. Then $\mathcal{A}_\mu$ is a σ-algebra on X, and $\bar{\mu}$ is a measure on $(X, \mathcal{A}_\mu)$. The restriction of $\bar{\mu}$ from $\mathcal{A}_\mu$ to $\mathcal{A}$ is μ.

Proof. This is left as an exercise. $\square$

Proposition 1.31. Suppose $(X, \mathcal{M}(\mu^*), \mu)$ is the measure space obtained from an outer measure by the Carathéodory construction (Th 1.15). That is, μ is the restriction of μ^* from 2^X to $\mathcal{M}(\mu^*)$. Then the measure μ is complete.

Proof. If $B \in 2^X$ is μ-null, then $B \subset A$ for some $A \in \mathcal{M}(\mu^*)$ with $\mu^*(A) = 0$. Outer measures are monotone, thus $\mu^*(B) = 0$. For arbitrary $C \in 2^X$,

$$\mu^*(C) \leq \mu^*(C \cap B) + \mu^*(C \setminus B) \leq \mu^*(C \setminus B) \leq \mu^*(C);$$

the first inequality follows from subadditivity of outer measures, the second from $\mu^*(C \cap B) \leq \mu^*(B) = 0$ and the third from $C \setminus B \subset C$. It follows that $B \in \mathcal{M}(\mu^*)$. $\square$

Lemma 1.32. Let A be a Lebesgue set in $\mathbb{R}$. Then there exist Borel subsets E and F such that $E \subset A \subset F$ and $\lambda(F \setminus E) = 0$.

Proof. First suppose that A is a Lebesgue set such that $\lambda(A) < \infty$. For each positive integer n use Th 1.28 to choose a compact set K_n that satisfies $K_n \subset A$ and $\lambda(A) - 1/n < \lambda(K_n)$ and an open set U_n that satisfies $A \subset U_n$ and $\lambda(U_n) \leq \lambda(A) + 1/n$. Let $E = \bigcup_n K_n$ and $F = \bigcap_n U_n$. Then E and F belong to $\mathcal{B}(\mathbb{R})$ and satisfy $E \subset A \subset F$. The relation

$$\lambda(F \setminus E) \leq \lambda(U_n \setminus K_n) = \lambda(U_n \setminus A) + \lambda(A \setminus K_n) \leq 2/n$$

holds for each n, and so $\lambda(F \setminus E) = 0$. Thus the lemma is proved in the case where $\lambda(A) < \infty$. If A is an arbitrary Lebesgue set, then A is the union of a sequence $\{A_n\}$ of Lebesgue measurable sets each of which satisfies $\lambda(A_n) < \infty$. For each n we can choose Borel sets E_n and F_n such that $E_n \subset A_n \subset F_n$ and $\lambda(F_n \setminus E_n) = 0$. The sets E and F defined by $E = \bigcup_n E_n$ and $F = \bigcup_n F_n$ then satisfy $E \subset A \subset F$ and $\lambda(F \setminus E) = 0$ (note that $F \setminus E \subset \bigcup_n (F_n \setminus E_n)$). □

Theorem 1.33. The measure space $(\mathbb{R}, \mathcal{M}(\lambda^*), \lambda)$ is the completion of the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.

Proof. Let λ be Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, let $\bar{\lambda}$ be the completion of λ, and let λ_m be Lebesgue measure on $(\mathbb{R}, \mathcal{M}(\lambda^*))$. Lemma 1.32 implies that $\mathcal{M}(\lambda^*)$ is included in the completion of $\mathcal{B}(\mathbb{R})$ under λ and that λ_m is the restriction of $\bar{\lambda}$ to $\mathcal{M}(\lambda^*)$. Thus we need only check that each set A that belongs to the completion of $\mathcal{B}(\mathbb{R})$ under $\bar{\lambda}$ is Lebesgue measurable. For such a set A there exist Borel sets E and F such that $E \subset A \subset F$ and $\lambda(F \setminus E) = 0$. Since $A \setminus E \subset F \setminus E$ and $\lambda_m(F \setminus E) = \lambda(F \setminus E) = 0$, the completeness of Lebesgue measure on $\mathcal{M}(\lambda^*)$ (Prop 1.31) implies that $A \setminus E \in \mathcal{M}(\lambda^*)$. Thus A, since it is the union of $A \setminus E$ and E, must belong to $\mathcal{M}(\lambda^*)$. $\square$

Remark 1.34. This completes our discussion of Lebesgue measure. We have only constructed and studied it on $\mathbb{R}^d$ for $d = 1$, to keep notation simple. The entire theory above remains correct if the ring $\mathcal{C}$ of half-open intervals $(a, b]$ in $\mathbb{R}$ is replaced by the ring of "blocks" $(a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_d, b_d]$ in $\mathbb{R}^d$, and the function $\gamma((a, b]) = b - a$ is replaced by $\gamma((a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_d, b_d]) = \prod_{k=1}^{d}(b_k - a_k)$. This results in Lebesgue measure on $\mathbb{R}^d$, Lebesgue sets $\mathcal{M}(\lambda^*)$ in $\mathbb{R}^d$ and Borel sets $\mathcal{B}(\mathbb{R}^d)$ in $\mathbb{R}^d$. The same results regarding completeness remain true.

Chapter 2

Lebesgue Integration

Starting from a measure space $(X, \mathcal{A}, \mu)$ we want to define integrals by generalising Riemann sums. The key point will be that for a function $f : X \to \mathbb{R}$ we split up the *range* space into intervals $\left(\frac{j-1}{k}, \frac{j}{k}\right]$ and then approximate an integral by sums: $\int f \; d\mu \approx \sum_{j=1}^{\infty} \left(\frac{j-1}{k}\right) \mu(E_j)$ where

$$E_j = f^{-1}\left(\left[\frac{j-1}{k}, \frac{j}{k}\right)\right) = \left\{x \in X \;\middle|\; f(x) \in \left[\frac{j-1}{k}, \frac{j}{k}\right)\right\}$$

For a *measurable function*, E_j will lie in $\mathcal{A}$.

2.1 Measurable functions

Definition 2.1. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces. A map $f : X \to Y$ is called $\mathcal{A}/\mathcal{B}$-measurable, or simply measurable, if

$$\forall\, B \in \mathcal{B} : \quad f^{-1}(B) \in \mathcal{A}.$$

If f is real-valued, it is implicit that measurability statements are with respect to the Borel σ-algebra, unless otherwise specified, and f is then called $\mathcal{A}$-measurable. Recall that for a function $f : X \to Y$, the pre-image is a function

$$\begin{aligned} f^{-1} : 2^Y &\to 2^X \\ B &\mapsto f^{-1}(B) = \{x \in X : f(x) \in B\} \end{aligned}$$

One can verify that compositions of measurable maps are measurable.

Example 2.2. Let f map from $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ as follows:

$$f(x) = \begin{cases} 1 & \text{if } x \in [0,1] \\ 0 & \text{else} \end{cases}$$

Then as B ranges over all possible subsets $\in \mathcal{B}(\mathbb{R}^d)$, the pre-image $f^{-1}(B)$ is always one of the sets $\emptyset, \mathbb{R}, [0,1]$ or $[0,1]^\complement$. These are Borel sets, and hence f is measurable.

Theorem 2.3. Let $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ be two measurable spaces, and assume that $\mathcal{G}_0 \subset 2^Y$ is such that $\sigma(\mathcal{G}_0) = \mathcal{G}$. Then $f : X \to Y$ is measurable iff

$$f^{-1}(G) \in \mathcal{F} \quad \forall G \in \mathcal{G}_0. \tag{2.4}$$

Proof. The above is a necessary condition for measurability of f; we only need to show that it is sufficient. Hence assume that (2.4) holds and define

$$\mathcal{D} = \{G \in \mathcal{G} : f^{-1}(G) \in \mathcal{F}\}.$$

For any collection $\{G_i\}_{i \in I} \subset \mathcal{G}$, we have the relations

$$f^{-1}\left(\bigcup_{i\in I} G_i\right) = \bigcup_{i\in I} f^{-1}(G_i), \qquad f^{-1}\left(\bigcap_{i\in I} G_i\right) = \bigcap_{i\in I} f^{-1}(G_i),$$
$$\left(f^{-1}(G_i)\right)^{\complement} = f^{-1}(G_i^{\complement}).$$

This shows that $\mathcal{D}$ is a σ-algebra on Y, and by definition $\mathcal{D} \subset \mathcal{G}$. If (2.4) holds, then $\mathcal{G}_0 \subset \mathcal{D}$, and then

$$\mathcal{G} = \sigma(\mathcal{G}_0) \subset \sigma(\mathcal{D}) = \mathcal{D} \subset \mathcal{G}$$

which shows $\mathcal{D} = \mathcal{G}$, that is, f is measurable. $\square$

As an example, suppose $Y = \mathbb{R}$ above. For f to be measurable, it suffices to show (2.4) for $\mathcal{G}_0$ being all open intervals, or all intervals of the form $(-\infty, b)$, $\ldots$.

Let $(X, \mathcal{T}_1)$ and $(Y, \mathcal{T}_2)$ be two *topological* spaces. The ε-δ definition of continuity of mappings in $\mathbb{R}$ is equivalent to the following more general definition of continuity:

$$f : X \to Y \text{ is continuous iff } \forall U \in \mathcal{T}_2 : \quad f^{-1}(U) \in \mathcal{T}_1,$$

that is, pre-images of open sets are open. We can hence say:

Corollary 2.5. Continuous functions are measurable.

Proposition 2.6. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let f and g be real-valued $\mathcal{A}$-measurable functions.

1. Then the functions $f + g$, αf for every $\alpha \in \mathbb{R}$, fg, f/g, $f \wedge g(x) = \min(f(x), g(x))$ and $f \vee g(x) = \max(f(x), g(x))$ are measurable functions.

2. The sets $\{x : f(x) < g(x)\}$, $\{x : f(x) \leq g(x)\}$ and $\{x : f(x) = g(x)\}$ are measurable.

3. Finally, if f_n is a sequence of $\mathcal{A}$-measurable functions, then $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$ and $\liminf_n f_n$ are measurable functions.

Here the sup function is defined by $\sup_n f_n(x) = \sup\{f_n(x) : n \in \mathbb{N}\}$, and the inf function is defined similarly. The limsup function is defined by

$$\limsup_n f_n(x) = \limsup\{f_n(x) : n \in \mathbb{N}\} = \inf_k \sup_{n \geq k} f_n(x),$$

and the liminf function is defined similarly:

$$\liminf_n f_n(x) = \liminf\{f_n(x) : n \in \mathbb{N}\} = \sup_k \inf_{n \geq k} f_n(x).$$

Proof. We only give hints: $f(x) + g(x) < t \Leftrightarrow f(x) < r$ and $g(x) < t - r$ for some $r \in \mathbb{Q}$. Show that f^2 is measurable, then use $fg = (f+g)^2 - f^2 - g^2)/2$. For the sets, note that $f(x) < g(x)$ iff there exists a rational r such that $f(x) < r < g(x)$. Take complements for the second statement about sets, and differences for the third. For f/g, see that $\{x : g(x) > 0\}$ is measurable, and see that $f(x)/g(x) < t \Leftrightarrow f(x) < tg(x)$. For min and max, note that $f \wedge g(x) > t \Leftrightarrow f(x) > t$ and $g(x) > t$, and $f \vee g(x) < t \Leftrightarrow f(x) < t$ and $g(x) < t$. Proceed similarly for the inf and sup of countably many functions. $\square$

Exercise 2.7. Show that any non-decreasing function $f : \mathbb{R} \to \mathbb{R}$ is (Borel-) measurable.

2.2 Integrals of simple functions

Definition 2.8. Let X be a non-empty set. Given $A \subset X$, the indicator function for A is the function χ_A defined on X by $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$. A function f defined on X is called a simple function if it is a linear combination of indicator functions: that is, there exist $n \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that $f(x) = \sum_{i=1}^{n} \alpha_i \, \chi_{A_i}(x)$

Different linear combinations can give the same simple function; but any simple function can be represented by a "canonical" linear combination as follows. If f is simple, then it has finite range, say $\{y_1, \ldots, y_n\}$. Then we can write

$$f(x) = \sum_{i=1}^{n} y_i \, \chi_{A_i}(x)$$

where $A_i = f^{-1}(\{y_i\})$. Note that the A_i are disjoint. Moreover, if X is endowed with the σ-field $\mathcal{A}$, then f is $\mathcal{A}$-measurable iff $\forall i : A_i \in \mathcal{A}$.

Definition 2.9 (Integral of a simple function)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ be a measurable simple function on X. The Lebesgue Integral of f is defined as

$$\int f \, d\mu = \int f(x)\mu(dx) = \sum_{i=1}^{n} \alpha_i \mu(A_i)$$

whenever the above linear combination is well defined.

Recall that we have defined $0 \cdot \infty := 0$, but we have left $\infty - \infty$ undefined. This simple definition of the Lebesgue integral can be extended to a larger class (larger than Riemann-integrable) of functions, see below. But first, we note a few properties of the Lebesgue integral:

Proposition 2.10. The Lebesgue integral is linear and monotonic, that is:

1. $\int \alpha f \, d\mu = \alpha \int f \, d\mu$ for any simple function f and $\alpha \in \mathbb{R}$
2. $\int (f+g) d\mu = \int f \, d\mu + \int g \, d\mu$ for any two simple functions f and g
3. If f and g are simple functions such that $f(x) \leq g(x)$ holds at each $x \in X$, then $\int f \, d\mu \leq \int g \, d\mu$.

Proof. Suppose $f = \sum_i a_i \chi_{A_i}$ and $g = \sum_j b_j \chi_{B_j}$, in canonical representation. Then

$$\int \alpha f \, d\mu = \sum_i \alpha a_i \mu(A_i) = \alpha \sum_i a_i \mu(A_i) = \alpha \int f \, d\mu.$$

For the sum, see that

$$f + g = \sum_i \sum_j (a_i + b_j) \chi_{A_i \cap B_j}$$

and so

$$\begin{aligned}\int (f+g)d\mu &= \sum_i \sum_j (a_i+b_j)\mu(A_i \cap B_j)\\ &= \sum_i \sum_j a_i\mu(A_i \cap B_j) + \sum_i \sum_j b_j\mu(A_i \cap B_j)\\ &= \sum_i a_i\mu(A_i) + \sum_j b_j\mu(B_j) = \int f d\mu + \int g d\mu\end{aligned}$$

due to the additivity property of the measure μ. Finally, $f \leq g$, then $g - f$ is a non-negative simple function, and $\int (g-f)d\mu \geq 0$. But then

$$\int g d\mu = \int (f+(g-f))d\mu = \int f d\mu + \int (g-f)d\mu \geq \int f d\mu.$$

$\square$

Theorem 2.11. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let f be any measurable function defined on X with values in $[0, \infty]$. Then there exists a sequence of simple functions $s_1, s_2, \dots$ on X such that

1. $0 \leq s_1 \leq s_2 \leq \dots$
2. $\forall x : \lim_{n \to \infty} s_n(x) = f(x)$.

Proof. For each $n \in \mathbb{N}$ and $k \in \{1, 2, \dots, n2^n\}$ define

$$A_{n,k} := f^{-1}\left(\left(\frac{k-1}{2^n}, \frac{k}{2^n}\right]\right), \qquad A_n := f^{-1}((n, \infty])$$

(see Figure 2.1).

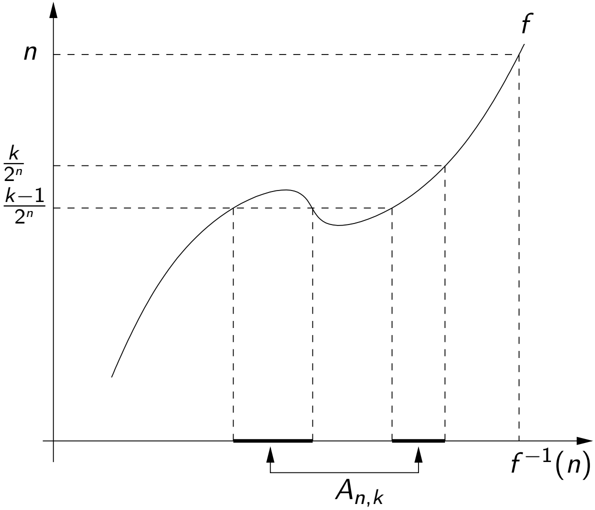

Figure 2.1: Each $A_{n,k}$ for $n \in \mathbb{N}$ and $k \in \{1, 2, \ldots, n \cdot 2^n\}$

Then $A_{n,k} \in \mathcal{A}$ as f is measurable. Define

$$s_n := \sum_{k=1}^{n2^n-1} \frac{k-1}{2^n} \cdot \chi_{A_{n,k}} + n \cdot \chi_{A_n}.$$

Then on A_{nk}, we have $s_n(x) = (k-1)/2^n$. As n increases by 1, A_{nk} is divided into two disjoint halves:

$$A_{n,k} = A_{n+1,2k-1} \cup A_{n+1,2k}$$

and thus $s_{n+1}(x) = (2k-1-1)/2^{n+1} = s_n(x)$ or $s_{n+1}(x) = (2k-1)/2^{n+1} > s_n(x)$ for $x \in A_{nk}$. Similarly, one sees that $s_n \leq s_{n+1}$ on E_{n+1} and on $f^{-1}((n, n+1])$, and the first statement is proven.

Now fix $x \in X$ and $n > x$, and let n and k be such that $x \in A_{nk}$. Then $f(x) \in ((k-1)/2^n, k/2^n]$ and $s_n(x) = (k-1)/2^n$, thus $f(x) - s_n(x) \leq 2^{-n}$. This shows the second statement. $\square$

Corollary 2.12. Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f : X \to [0, \infty]$ is measurable iff it is the pointwise limit of simple functions.

Proof. This follows from Prop 2.6 and Th 2.11. $\square$

2.3 Integrals of positive functions

Definition 2.13 (Integral of a positive function)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to [0, \infty]$ measurable. The Lebesgue Integral of f with respect to μ is

$$\int f d\mu := \sup \left\{ \int s d\mu : s \text{ is simple, measurable and } 0 \leq s \leq f \right\}.$$

For $E \in \mathcal{A}$ define

$$\int_E f \; d\mu := \int \chi_E \cdot f \; d\mu.$$

The supremum is not taken over the empty set, and thus $\int f d\mu$ is well defined; though it may well equal ∞.

Proposition 2.14. Let $(X, \mathcal{A}, \mu)$ be a measure space and f and g non-negative measurable functions on X such that $f \leq g$. Then

$$\int_E f \; d\mu \leq \int_E g \; d\mu \qquad \forall\, E \in \mathcal{A}.$$

Proof. Since $0 \leq f \leq g$ any simple measurable non-negative function s such that $0 \leq s \leq f$ also satisfies $0 \leq s \leq g$, hence as the integral is the supremum over these,

$$\int f d\mu \leq \int g d\mu.$$

Since $\chi_E f \leq \chi_E g \;\; \forall\, E \in \mathcal{A}$ the claim follows. $\square$

Note: If $f(x) \geq m > 0 \;\; \forall\, x \in E$ then $g := m \cdot \chi_E \leq f \cdot \chi_E$ so

$$\int_E f \; d\mu \geq \int_E m \cdot \chi_E \; d\mu = m\, \mu(E).$$

Theorem 2.15 (Monotone Convergence Theorem, Beppo Levi)**.** For a measure space $(X, \mathcal{A}, \mu)$ let f_n be non-negative measurable functions ($n \in \mathbb{N}$) such that $0 \leq f_1 \leq f_2 \leq \dots$ and $f(x) := \lim_n f_n(x)$ exists for every $x \in X$. Then the thusly defined function f is also non-negative measurable, and

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

That is, limit and integration may be interchanged.

Proof. That f is measurable follows from Prop 2.6 and the fact that $f = \limsup_n f_n$ ($= \liminf_n f_n = \sup_n f_n$). Prop 2.14 implies

$$0 \leq \int f_1 \; d\mu \leq \int f_2 \; d\mu \leq \dots \text{ and } \int f_n d\mu \leq \int f \, d\mu \, \forall n \in \mathbb{N}$$

and hence

$$\alpha := \lim_{n \to \infty} \int f_n \; d\mu \leq \int f \; d\mu. \tag{2.16}$$

It remains to show the opposite inequality "$\geq$" in (2.16). Let $c \in (0,1)$ and s be a simple function such that $0 \leq s \leq f$. Define

$$A_n := \{x \in X : f_n(x) \geq cs(x)\},$$

then note that $A_n \in \mathcal{A}$ due to Prop 2.6. Since $f_n \leq f_{n+1}$, we have $A_1 \subseteq A_2 \subseteq \ldots$. Since $c < 1$, we have $cs \leq cf \leq f = \lim_n f_n$, and so $\bigcup_n A_n = X$. Then $f_n \geq f_n \chi_{A_n} \geq cs\chi_{A_n}$, and so

$$\int f_n \, d\mu \geq \int_{A_n} f_n \, d\mu \geq \int_{A_n} cs \, d\mu. \tag{2.17}$$

Suppose that s has the representation $s = \sum_{i=1}^{k} \alpha_i \chi_{B_i}$; then

$$\lim_n \int_{A_n} cs \, d\mu = \lim_n \sum_{i=1}^{k} c\alpha_i \mu(B_i \cap A_n) = \sum_{i=1}^{k} c\alpha_i \mu(B_i) = c \int s \, d\mu$$

using linearity of sums and Th 1.12 part 3. The rightmost term in (2.17) hence converges to $c \int s \, d\mu$, and the leftmost term to α. This shows

$\alpha \geq c \int s \, d\mu$. This inequality holds for every $c \in (0,1)$, and hence $\alpha \geq \int s \, d\mu$. The latter inequality holds for every simple measurable s, and thus by Def 2.13, $\alpha \geq \int f \, d\mu$. $\square$

Using Beppo Levi's theorem, we derive the usual properties of the Lebesgue Integral:

Theorem 2.18. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let f and g be non-negative measurable functions on X. Then

(1) $\int c \cdot f \; d\mu = c \int f \; d\mu \qquad \forall \, c \in [0, \infty)$;

(2) $\int (f+g) \; d\mu = \int f \; d\mu + \int g \; d\mu$;

(3) $\int f \; d\mu = 0$ iff $f = 0$ μ-a.e.;

(4) The map $\nu : \mathcal{A} \to [0, \infty]$ by $\nu(A) := \int_A f \; d\mu \quad \forall \, A \in \mathcal{A}$ defines a measure ν. (The notation $d\nu = f \, d\mu$ is then common.)

Proof. (1) and (2) are left as an exercise. For (3), let $A_n = \{f > 1/n\}$ which is shorthand for $\{x \in X : f(x) > 1/n\}$. We have

$$\chi_{A_n} \leq n f$$

(check this by plugging in $x \in A_n$ and $x \notin A_n$). Then

$$\mu(A_n) = \int \chi_{A_n} d\mu \leq n \int f d\mu$$

by Prop 2.14. Suppose now that $\int f\, d\mu = 0$. Then $\mu(A_n) = 0$ for every n, and note that $\bigcup_n A_n = \{f > 0\}$; hence by subadditivity

$$\mu(\{f > 0\}) = \mu\left(\bigcup_n A_n\right) \leq \sum_n \mu(A_n) = \sum_n 0 = 0,$$

that is, $f = 0$ μ-a.e. On the other hand, suppose that $\int f\, d\mu > 0$. Then by Def 2.13 there exists a simple, non-negative measurable s such that $0 < \int s\, d\mu$. Suppose this s has representation $\sum_k \alpha_k \chi_{B_k}$. Since $\sum_k \alpha_k \mu(B_k)$ is positive, for at least one pair (α_k, B_k) we have both $\alpha_k > 0$ and $\mu(B_k) > 0$. But then f is not equal to 0 μ-a.e. since $f(x) \geq \alpha_k > 0$ for all $x \in B_k$ which has positive measure.

For (4), let $\nu : \mathcal{A} \to [0, \infty]$ by $\nu(A) := \int_A f \; d\mu$. Since $\nu(\emptyset) = 0$ we only need to check that ν satisfies countable additivity. Let $A_n \in \mathcal{A}$ be pairwise disjoint and $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Then $\chi_A \cdot f = \sum_{n=1}^{\infty} \chi_{A_n} \cdot f$. Let $f_k := \sum_{n=1}^{k} \chi_{A_n} f$; then $0 \leq f_1 \leq f_2 \leq \dots$ and $\lim_{k\to\infty} f_k(x) = f(x)\chi_A(x)$. So

$$\nu(A) = \int_A f \; d\mu = \int \chi_A f \, d\mu = \int \lim_k f_k \, d\mu = \lim_k \int f_k \; d\mu$$

$$= \lim_{k\to\infty} \sum_{n=1}^{k} \int_{A_n} f \, d\mu = \sum_{n=1}^{\infty} \nu(A_n).$$

$\square$

We have established the Lebesgue integral of positive measurable functions. In order to extend our work to functions with positive and negative values, we just one more step:

Definition 2.19 (Lebesgue integral of measurable functions)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ a measurable function. The positive and negative parts of f are given by $f_+ = f \vee 0$ and $f_- = -(f \wedge 0)$. If both $\int f_+ \, d\mu < \infty$ and $\int f_- \, d\mu < \infty$, then the Lebesgue Integral of f with respect to μ is

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

Unsurprisingly, this integral is linear, that is $\int \alpha f \, d\mu = \alpha \int f \, d\mu$ and $\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu$. Moreover, we have

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

In $X = \mathbb{R}^d$, the Lebesgue integral with μ being Lebesgue measure generalises the Riemann integral. The main advantage however is that it extends

e.g. to any (separable) metric space, which is needed for many applications
analysis and probability.