MATH5825: Measure, Integration and Probability
Lecture Notes Semester 2 2011

Typed by Steven Nguyen

November 23, 2011

Contents

Chapter 1

Riemann Integration - Revision and Problems

These notes are based on lectures given by Dr. Hendrik Grundling at UNSW in Semester 2 2011.

Let $S \subset \mathbb{R}^n$ be a bounded set. If possible we want to finds its volume. Firstly we define an n-block (half open) as $C = [t_1, s_1) \times \cdots \times [t_n, s_n)$ and the volume of this n-block as $V(C) = (s_1 - t_1) \times \cdots \times (s_n - t_n)$. We want to approximate S by covering the set with n-blocks (See Figure 1.1). Let $\mathcal{C} = \{C_1, \ldots, C_N\}$ be a finite set of disjoint n-blocks covering S. Thus

$$S \subseteq \bigcup_{k=1}^{N} C_k \quad \text{and} \quad C_i \cap C_j = \emptyset \quad \text{if } i \neq j.$$

Then we approximate the volume of S by

$$V(\mathcal{C}) = \sum_{k=1}^{N} V(C_k).$$

To get the best approximation we define the **(Jordan) outer n-volume of** S by

$$\overline{V(S)} := \inf\{V(\mathcal{C}) \mid \mathcal{C} \text{ a finite set of n-blocks covering } S\}.$$

Note: $\overline{V(\emptyset)} = 0$. We take the volume of the smallest covering possible. We can also approximate the volume from the inside (See Figure 1.2): Let $\mathcal{C} = \{C_1, \ldots, C_N\}$ be a finte set of

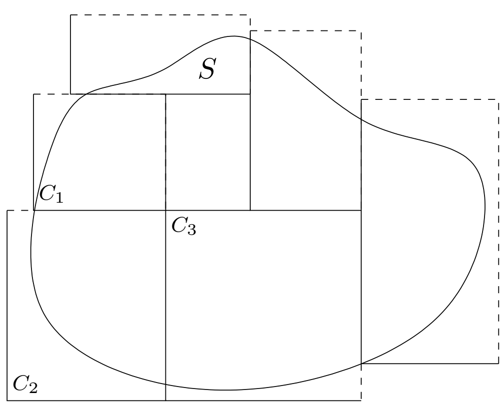

Figure 1.1: Covering of S

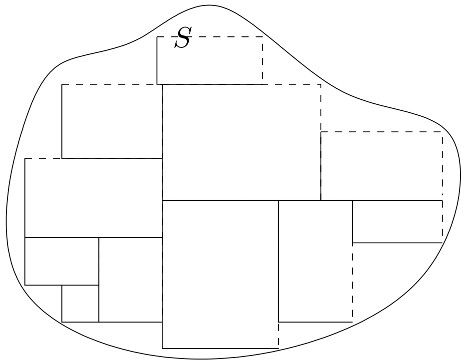

Figure 1.2: Approximating volume of S from the inside

disjoint n-blocks such that $\bigcup_{k=1}^{N} C_k \subseteq S$. We approximate the volume of S by $V(\mathcal{C}) = \sum_{k=1}^{N} V(C_k)$. The **(Jordan) inner n-volume of** S is

$$\underline{V(S)} := \sup\{V(\mathcal{C}) \mid \mathcal{C} \text{ is a finite set of disjoint n-blocks inside } S\}.$$

The convention is that $\underline{V(S)} = 0$ if S contains no n-blocks and with this it is obvious to that

$$0 \leq \underline{V(S)} \leq \overline{V(S)}.$$

Definition 1.1. A bounded set $S \subset \mathbb{R}^n$ is **Jordan measure** if $\overline{V(S)} = \underline{V(S)} = |S| =$ n-volume of S.

Theorem 1.1. A bounded set S is Jordan measurable if and only if its boundary ∂S is of Jordan measure zero i.e. for each $\varepsilon > 0$ $\exists$ covering $\mathcal{C} = \{C_1, \dots, C_N\}$ of ∂S by n-blocks such that $V(\mathcal{C}) < \varepsilon$.

Sets with "continuous" boundaries are Jordan measurable, however some natural sets are not.

Example 1.1. Let $S = \mathbb{Q} \cap [0, 1) \subset \mathbb{R}$ then there are no 1-blocks in S so $\underline{V(S)} = 0$. However if $\mathcal{C} = \{C_1, \dots, C_N\}$ is a disjoint covering by 1-blocks $[t, s)$ then $[0, 1) \subset \bigcup_{k=1}^{N} C_k$ (why?)[1] and $[0, 1)$ is itself a covering of S by a 1-block thus $\overline{V(S)} = V([0, 1)) = 1$. So S is not Jordan measurable.

Example 1.2. Some **open** bounded sets are **NOT** Jordan measurable. Let $\{q_k \mid k = 1, 2, \dots\} = \mathbb{Q} \cap (0, 1)$ be an enumeration of it. Fix $\varepsilon \in (0, 1)$. For each k define $(0, 1) \cap (q_k - \frac{\varepsilon}{2^{k+1}}, q_k + \frac{\varepsilon}{2^{k+1}}) = I_k \ni q_k$. Let $S := \bigcup_{k=1}^{\infty} I_k \subset (0, 1)$ then S is open. So $\underline{V(S)} \leq \sum_{k=1}^{\infty} V(I_k) \leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon < 1$. However each $q \in \mathbb{Q} \cap (0, 1)$ is in S, so by the same argument in Example 1.1, $\overline{V(S)} = 1 > \underline{V(S)}$. Thus S is not Jordan measurable.

[1] This is because the 1-blocks must appear in the form $[t_0, t_1), [t_1, t_2), \dots, [t_{n-1}, t_n)$ or else the covering will not be disjoint and cover all the rationals

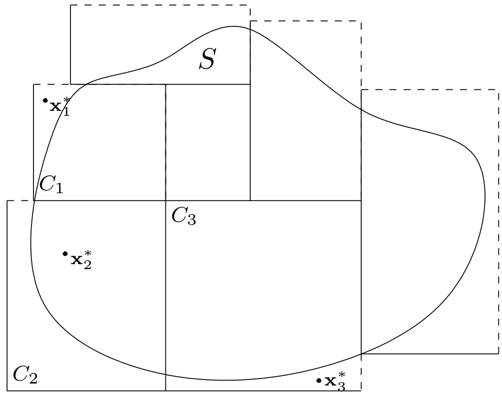

Figure 1.3: Finite Partition of R covering S

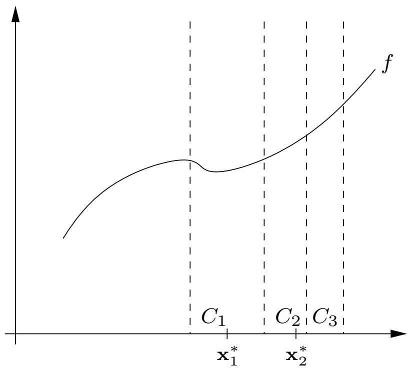

Figure 1.4: Riemann sum of $\mathcal{C}$ where $\mathbf{x}_k^*$ is a fixed choice of points

The Riemann integral of $f : S \subset \mathbb{R}^n \to \mathbb{R}$ for a Jordan measurable set S is defined by Riemann Sums: let $R \supset S$ be an n-block containing S (as S is bounded) and extend f to R by

$$F(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in S \\ 0 & \text{if } \mathbf{x} \notin S \end{cases}$$

let $\mathcal{C}$ be any partition of R by finitely many (disjoint) n-blocks. Then the Riemann sum of $\mathcal{C} = \{C_1, \ldots, C_N\}$ is $\mathcal{R}(\mathcal{C}) := \sum_{k=1}^{N} F(\mathbf{x}_k^*) \cdot V(C_k)$ where $\mathbf{x}_k^* \in C_k$ is a fixed choice of points. With this we can define the upper sum as:

$$\overline{\mathcal{R}(\mathcal{C})} = \sum_{k=1}^{N} M_k \cdot V(C_k) \quad \text{with } M_k = \sup\{F(\mathbf{x}) \mid \mathbf{x} \in C_k\}$$

and lower sum as:

$$\underline{\mathcal{R}(\mathcal{C})} = \sum_{k=1}^{N} m_k \cdot V(C_k) \quad \text{with } m_k = \inf\{F(\mathbf{x}) \mid \mathbf{x} \in C_k\}.$$

Clearly with these definitions we have $\underline{\mathcal{R}(\mathcal{C})} \leq \mathcal{R}(\mathcal{C}) \leq \overline{\mathcal{R}(\mathcal{C})}$. Thus we can define the **upper integral** by

$$\overline{I(f)} = \inf\{\overline{\mathcal{R}(\mathcal{C})} \mid \mathcal{C} \text{ any finite partition of } \mathcal{R} \text{ by disjoint n-blocks}\}$$

and **lower integral** by

$$\underline{I(f)} = \sup\{\underline{\mathcal{R}(\mathcal{C})} \mid \mathcal{C} \text{ any finite partition of } \mathcal{R} \text{ by disjoint n-blocks}\}.$$

The **norm** of the partition $\mathcal{C}$ of R is

$$\|\mathcal{C}\| = \max\{ \operatorname{diam}(C_k) \mid k = 1, \ldots, N\}$$

where $\operatorname{diam}(C_k) := \sup\{ \|\mathbf{x} - \mathbf{y}\| \mid \mathbf{x}, \mathbf{y} \in C_k\}$. The Riemann integral $I \equiv \int_S f \, d\mathbf{x}$ is defined by: For each $\varepsilon > 0 \;\; \exists \, \delta > 0$ such that $|I - \mathcal{R}(\mathcal{C})| < \varepsilon$ for all finite partitions $\mathcal{C}$ of $\mathcal{R}$ with $\|\mathcal{C}\| < \delta$ and all choices of points $\mathbf{x}_k^* \in C_k$. If I exists, we say f is Riemann integrable.

Theorem 1.2. Let S be a Jordan measurable set. Then:

(1) a function $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is Riemann integrable if and only if $\overline{I(f)} = \underline{I(f)}$ and in this case $\overline{I(f)} = I(f) = \underline{I(f)}$.

(2) a function f which is continuous on a **closed** Jordan measurable set S is Riemann integrable.

(3) χ_S is Riemann integrable and $I(\chi_S) = |S|$ integral is the usual one with all the familiar properties.

1.1 Problems with Riemann Integration

(1) Some natural bounded functions are not Riemann integrable.

(2) $\{f_n\}$ Riemann integrable, $f_n(x) \to f(x)$, $|f_n| \leq 1$, $|f| \leq 1$ need **not** imply that f is Riemann integrable.

(3) Some natural bounded sets are not Jordan measurable $\Rightarrow$ we cannot perform Riemann integration on them.

(4) Hard to generalise away from $\mathbb{R}^n$.

Example 1.3.

(1) An example of a natural bounded function which isn't Riemann integrable is the Dirichlet function

$$\chi_{\mathbb{Q}\cap[0,1)}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{if } x \notin \mathbb{Q} \cap [0,1]. \end{cases}$$

For any (disjoint) covering of $[0,1)$ must be of the form $[x_0, x_1) \cup [x_1, x_2) \cup \cdots \cup [x_{N-1}, x_N)$ with $x_0 = 0$ and $x_N = 1$. we can choose either all $x_k^* \in \mathbb{Q} \cap [x_{k-1}, x_k)$ or all $x_k^* \in [x_{k-1}, x_k) \setminus \mathbb{Q}$. So $\overline{\mathcal{R}(\mathcal{C})} = 1$, $\underline{\mathcal{R}(\mathcal{C})} = 0$. This holds $\forall \, \mathcal{C}$ hence $\overline{I(f)} = 1$, $\underline{I(f)} = 0$. So by Theorem 1.2, it is not Riemann integrable.

(2) We will now show an example of second problem listed above with Riemann integration. Let $\{q_k \mid k = 1, 2, \dots\} = \mathbb{Q} \cap [0, 1)$ be an enumeration of rationals in $[0, 1)$ and define

$$f_m := \sum_{i=1}^{m} \chi_{\{q_i\}}.$$

Then $f_m(x) \to f(x) = \chi_{\mathbb{Q}\cap[0,1)}(x)$. But f is not Riemann integrable. Even though $\int_0^1 f_m \, dx = \sum_{i=1}^{m} |\{q_i\}| = 0 \;\; \forall \, m$ (i.e. $\{f_m\}$ is Riemann integrable.)

1.2 Lebesgue's "Problem of measure in $\mathbb{R}^n$"

To negate the issues surrounding Riemann integration Lebesgue devised the following function: Assign to each bounded subset $S \subset \mathbb{R}^n$ a number $m(S) \geq 0$ (called the measure of S) such that:

(1) $m(S) = m(T)$ whenever S is congruent to T (i.e. $S = h(T)$ for an isometry h of $\mathbb{R}^n$).

(2) $S = \bigcup_{i=1}^{\infty} S_i, \; S_i \cap S_j = \emptyset$ if $i \neq j$ then $m(S) = \sum_{i=1}^{\infty} m(S_i)$ (countable additivity).

(3) $m(I) = 1$ when $I = [0, 1] \times \cdots \times [0, 1]$ (n times).

However this humble attempt is still too much to ask as the following shall show.

Proposition 1.3. In $\mathbb{R}$ the "problem of measure" has no solution.

Proof (Vitali)**.** We will construct a bounded set S for which $m(S)$ cannot be defined. 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 a Vitali set $S \subset [0, 1]$ by choosing one point from each equivalence class (this uses the Axiom of Choice).

Lemma 1.3.1. If $r, q \in \mathbb{Q} \cap [0, 1], \; r \neq q$ then $(S + r) \cap (S + q) = \emptyset$.

Proof (Lemma 1.3.1)**.** 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.

$\triangledown$

Lemma 1.3.2. $[0, 1] \subseteq \bigcup\{S + r \mid r \in \mathbb{Q} \cap [-1, 1]\} =: T \subseteq [-1, 2]$.

Proof (Lemma 1.3.2)**.** Let $x \in [0, 1)$ so $x \sim p$ 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 $x \in S + r, \; r \in \mathbb{Q} \cap [-1, 1]$.

$\triangledown$

If there is a measure m as defined above then by the first property $m(S+r) = m(S) \ \forall\, r$. So

$$\begin{aligned} 1 = m([0,1]) &= m(T) - m(T \setminus [0,1]) && \text{as } m(T) = m([0,1]) + m(T \setminus [0,1]) \\ &\leq m(T) && \text{by Lemma 1.3.1} \\ &= m\left(\bigcup\{S + r \mid r \in \mathbb{Q} \cap [-1,1]\}\right) \\ &= \sum_{r \in \mathbb{Q}\cap[-1,1]} m(S+r) && \text{by property (2) and Lemma 1.3.2} \\ &= \sum_{r \in \mathbb{Q}\cap[-1,1]} m(S) && \text{by property (1).} \end{aligned}$$

Thus $m(S) \neq 0$. So as we are summing over infinite points (rational points between $[-1,1]$) it follows that $\sum_{r\in\mathbb{Q}\cap[-1,1]} m(S) = \infty$. However

$$\begin{aligned} \infty = \sum_{r\in\mathbb{Q}\cap[-1,1]} m(S) &= m(T) \\ &= m([-1,2]) - m([-1,2) \setminus T) \\ &\leq m([-1,2]) && \text{by lemma 2} \\ &= m([-1,0]) + m([0,1]) + \\ &\quad\ m([1,2]) - m(\{0,1\}) \\ &\leq 3m([0,1]) \\ &= 3. \end{aligned}$$

Contradiction.

□

In higher dimensions the "problem of measure" has no solution:

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

Thus we can take a unit sphere and cut it into finitely many pieces which reassemble into a sphere of any size. Take S, T as above, then if a measure m exists for $\mathbb{R}^n$, $n \geq 3$ then

$$m(S) = \sum_{l=1}^{k} m(E_l) = \sum_{l=1}^{k} m(F_l) = m(T). \tag{1.1}$$

Let $S = I = [0,1) \times [0,1) \times \cdots \times [0,1)$ (n times) and $T = S \cup (S + \mathbf{a})$, $\|\mathbf{a}\| > 2$ then $S \cap (S + \mathbf{a}) = \emptyset$. Then

$$\begin{aligned} 2 = 2m(S) &= m(S) + m(S\mathbf{a}) \\ &= m(T) \\ &= m(S) && \text{by (1.1)} \\ &= 1. \end{aligned}$$

Thus m doesn't exist.

We will give a sketch proof of a subset of the Banach Tarski Theorem. Firstly let $G =$ Euclidean Group on $\mathbb{R}^n$, we define some appropriate definition that will be used in the proof:

Definition 1.2.

(1) Call $A, B \subset \mathbb{R}^n$ **equidecomposable** (with respect to G) if there are finite partitions:

$$A = \sum_{i=1}^{k} A_i, \quad B = \sum_{i=1}^{k} B_i$$

such that $B_i = g_i(A_i)\ \ \forall\ i$ and some $g_i \in G$. This is an equivalence relation.

(2) $E \subseteq \mathbb{R}^n$ has a **paradoxical decomposition** if $\exists\ A, B \subseteq E$ such that $A \cap B = \emptyset$ and A is equidecomposable with E and B is equidecomposable with E.

Theorem 1.5 (Banach Tarski Variant)**.** The unit ball $\overline{B(0,1)}$ in $\mathbb{R}^3$ has a paradoxical decomposition.

Proof (Sketch)**.** Let $H \subset G$ be a group generated by 2 rotations:

$$a = \text{ rotation by } \theta \text{ about } x\text{-axis}$$
$$b = \text{ rotation by } \theta \text{ about } z\text{-axis}$$

where $\theta =$ irrational multiple of π (fixed).

- Show $H \cong F_2 =$ free group of 2 generators i.e. all (reduced) strings of $\{a, a^{-1}, b, b^{-1}\}$ i.e. finite strings from which $aa^{-1}, bb^{-1}, a^{-1}a, b^{-1}b$ have been removed. $\emptyset =$ identity and concatenation is a group operation.

- Let $S(x) :=$ all reduced strings starting with x. Then

$$F_2 = \{\emptyset\} \cup S(a) \cup S(a^{-1}) \cup S(b) \cup S(b^{-1});$$
$$F_2 = aS(a^{-1}) \cup S(a);$$
$$F_2 = bS(b^{-1}) \cup S(b).$$

  Thus with $A = S(a^{-1}) \cup S(a)$ and $B = S(b^{-1}) \cup S(b)$ we have a "paradoxical decomposition" of F_2.

- Partition the unit sphere S^2 into H-orbits (equivalence classes where $x \sim y$ if $\exists\ h \in H$ such that $x = hy$). Let $T = \{x \in S^2 \mid gx = x \text{ for some } g \in H \backslash e\}$. Then T is countable, as H is countable and a rotation has 2 fixed points. Let $M \subset S^2 \setminus T$ be defined by taking one point from each orbit (using the Axiom of Choice).

- Show that if $x \in M,\ y \in Hx$ and $y \notin T$ then there is a unique $h \in H$ such that $y = hx$. Thus we get a partition of $S^2 \setminus T$ as follows

$$\begin{aligned} S^2 \setminus T &= HM \cap S^2 \setminus T \\ &= (M \cup S(a)M \cup S(a^{-1})M \cup S(b)M \cup S(b^{-1})M) \cap S^2 \setminus T \end{aligned}$$

  and

$$\begin{aligned} S^2 \setminus T &= (aS(a^{-1})M \cup S(a)M) \cap S^2 \setminus T \\ &= (bS(b^{-1})M \cup S(b)M) \cap S^2 \setminus T \end{aligned}$$

$\Rightarrow$ a paradoxical decomposition of $S^2 \setminus T$.

- Prove that $S^2 \setminus T$ equidecomposable with S^2.
- Connect each point on S^2 with the origin (the centre of the sphere) to get a paradoxical decomposition of $\overline{B(0,1)} \setminus 0$.
- Show $\overline{B(0,1)} \setminus 0$ is equidecomposable with $\overline{B(0,1)}$.

As the problem of measure has no solution on bounded sets of $\mathbb{R}^n$, we need to restrict to a smaller family of sets to define a measure.

Chapter 2

Abstract Measure Theory and the Lebesgue measure

Definition 2.1. Given a set $X = \emptyset$, then a σ**-algebra** of X is a set of sets $\mathcal{A} \subseteq \mathcal{P}(X)$ such that:

(1) $X \in \mathcal{A}$;

(2) if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$;

(3) if $A = \bigcup\limits_{n=1}^{\infty} A_n$ with $A_n \in \mathcal{A}$ then $A \in \mathcal{A}$. Say $A \in \mathcal{A}$ is $\mathcal{A}$-measurable.

From this definition it follows that $\emptyset \in \mathcal{A}$ since $X \in \mathcal{A}$ and from property (2) the result follows. $\mathcal{A}$ is closed with respect to finite unions and countable intersections. To see this consider $A_1, \dots, A_k \in \mathcal{A}$ then by taking $A_j = \emptyset$ for $j > k$ property (3) implies that $A_1 \cup \dots \cup A_k \in \mathcal{A}$. As for the latter from property (2) and (3) we have

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

Also if $A, B \in \mathcal{A}$ then $A \setminus B = B^c \cap A \in \mathcal{A}$.

Example 2.1.

(1) $\mathcal{P}(X)$ satisfies the conditions for being a σ-algebra.

(2) The smallest σ-algebra for X is $\mathcal{A} = \{\emptyset, X\}$.

Theorem 2.1. If $\mathcal{C} \subseteq \mathcal{P}(X)$ for a set $X \neq \emptyset$ then there exists a smallest σ-algebra $\mathcal{A}(\mathcal{C}) \subseteq \mathcal{P}(X)$ such that $\mathcal{C} \subseteq \mathcal{A}(\mathcal{C})$. Call $\mathcal{A}(\mathcal{C}) \equiv$ the σ**-algebra generated by** $\mathcal{C}$.

Proof. Let $\mathcal{A}$ = intersection of all σ-algebra of X containing $\mathcal{C}$ (this is non-empty since $\mathcal{C} \subset \mathcal{P}(X) = \sigma$-algebra).

- $X \in \mathcal{A}$ as it is in all $\sigma-$algebras.
- Let $A \in \mathcal{A} \Rightarrow A \in \mathcal{B}\ \ \forall\ \sigma$-algebra $\mathcal{B}$ such that $\mathcal{C} \subseteq \mathcal{B} \Rightarrow A^c \in \mathcal{B}$ with $\mathcal{C} \subset \mathcal{B} \Rightarrow A^c \in \mathcal{A}$.

- Let $\{A_1, A_2, \dots\} \subset \mathcal{A} \Rightarrow \{A_1, A_2, \dots\} \subset \mathcal{B} \ \ \forall \, \mathcal{B}$ with (σ-algebras' $\mathcal{C} \subset \mathcal{B} \Rightarrow A = \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{B} \ \ \forall \, \mathcal{B}$ with $\mathcal{C} \subset \mathcal{B} \Rightarrow A \in \mathcal{A}$.

$\square$

Definition 2.2. For a set X, a **topology** is a set of sets $\tau \subseteq \mathcal{P}(X)$ such that:

(1) $\emptyset \in \tau, \ X \in \tau$;

(2) if $V_i \in \tau, \ i = 1, \dots k$ then $\bigcap\limits_{i=1}^{k} V_i \in \tau$;

(3) if $V_\lambda \in \tau, \ \lambda \in \Lambda$ then $\bigcup\limits_{\lambda \in \Lambda} V_\lambda \in \tau$

A $V \in \tau$ is called an **open set**, and (X, τ) is a **topological space**. The **closed sets** are complements of open sets

Definition 2.3. Let (X, τ) be a topological space, then the **Borel σ-algebra** is $\mathcal{A}(\tau) = \mathcal{B}(X) \subseteq \mathcal{P}(X)$, i.e. the σ-algebra generated by τ. We say an $A \in \mathcal{B}(X)$ is a **Borel set**.

All open sets and closed sets are Borel, as well as all countable unions of closed sets (F_σ-sets) and all countable intersections of open sets (G_δ-sets).

Example 2.2. So more examples:

(1) For $\mathbb{R}$ with usual topology all intervals $(a, b), [a, b], [a, b), (a, b], (-\infty, a)$, $(-\infty, a], (b, \infty), [b, \infty)$ are borel.

(2) $\mathbb{Q}$ is an F_σ-set and hence Borel.

(3) All countable sets in $\mathbb{R}$ are Borel.

Definition 2.4 (Conventions for ∞)**.** The **extended real line** is the set $[-\infty, \infty] := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ together with:

- the ordering $-\infty < x < \infty \quad \forall \, x \in \mathbb{R}$ and usual ordering on $\mathbb{R}$. Define intervals $[a, \infty] := \{x \in [-\infty, \infty] \mid a \leq x \leq \infty\}$ etc.
- the topology $\tau(\xi)$ generated by $\xi = \{(a, b), \ [-\infty, a), \ (a, \infty] \mid a, b \in \mathbb{R}\}$.
- the arithmetic - usual rules for $\mathbb{R}$ together with:
  - $a + (\pm\infty) = (\pm\infty) + a = \pm\infty \ \ \forall \, a \in \mathbb{R}$, $\infty + \infty = \infty$ $\qquad$ ($\infty - \infty$ **undefined**).
  - $a.(\pm\infty) = (\pm\infty).a = \pm\infty$ if $a > 0$, 0 if $a = 0$ and $\mp\infty$ if $a < 0$;
  - $a/(\pm\infty) = 0 \ \ \forall \, a \in \mathbb{R}$;
  - $a/0 = +\infty$ if $a > 0$, $-\infty$ if $a < 0$ (∞/∞ and $0/0$ undefined) thus $[-\infty, \infty]$ is **not** a field.

Definition 2.5. For a given σ-algebra $\mathcal{A} \subseteq \mathcal{P}(X)$, a (**positive**) **measure** is a map $\mu : \mathcal{A} \to [0, \infty]$ such that

(1) it is **countably additive** i.e. if $A_i \in \mathcal{A},\ A_i \cap A_j = \emptyset$ if $i \neq j$ then

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

(2) $\mu(A) < \infty$ for some $A \in \mathcal{A}$.

A triple $(X, \mathcal{A}, \mu)$ is a **measure space**; if $\mu(X) < \infty$ it is **finite** and μ is **σ-finite** if $X = \bigcup_{i=1}^{\infty} A_i,\ A_i \in \mathcal{A},\ \mu(A_i) < \infty$. If $\mu(X) = 1$, say μ is a **probability measure**. If we let $\mu : \mathcal{A} \to \mathbb{C}$ (instead of $[0, \infty]$) it is a **complex** measure.

Example 2.3.

(1) For a set $X \neq \emptyset$ with $\mathcal{A}$ is a σ-algebra on X, define $\mu : \mathcal{A} \to [0, \infty]$ by

$$\mu(A) = \text{ number of elements in } A \in \mathcal{A}$$

then $\mu =$ **counting measure**.

(2) For a σ-algebra $\mathcal{A}$ of X, fix $x \in X$ and define $\delta_x : \mathcal{A} \to [0, \infty]$

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

So $\delta_x \equiv$ **point mass** of x is a measure.

(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 $\mu \equiv$ **trivial measure**.

(4) Below we will construct a measure on $\mathcal{B}(\mathbb{R}^n)$ which agrees with Riemannian measure on Jordan measurable sets.

Theorem 2.2. For a (positive) measure space $(X, \mathcal{A}, \mu)$ we have

(1) $\mu(\emptyset) = 0$;

(2) $\mu(A_1 \cup \cdots \cup A_n) = \mu(A_1) + \mu(A_2) + \cdots + \mu(A_n)$ if $A_i \in \mathcal{A}$ are disjoint;

(3) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)\ \ \forall\, A, B \in \mathcal{A}$;

(4) If $A_1 \subseteq A_2 \subseteq \dots,\ \ A_i \in \mathcal{A}$ and $A = \bigcup_{n=1}^{\infty} A_n$ then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

(5) If $A_1 \supseteq A_2 \supseteq \dots,\ \ A_i \in \mathcal{A}$ and $A = \bigcap_{n=1}^{\infty} A_n,\ \mu(A_1) < \infty$, then $\mu(A_n) \xrightarrow[\infty]{n} \mu(A)$

Proof.

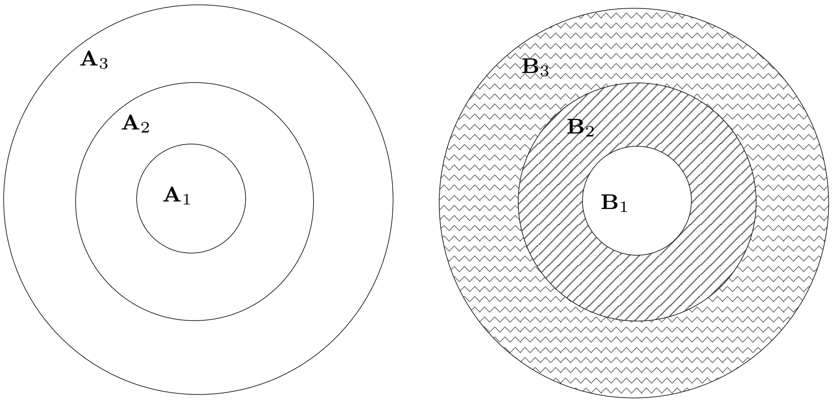

Figure 2.1: Creating collection of disjoint sets B_n

(1) Let $A_1 \in \mathcal{A}$ such that $\mu(A_1) < \infty$ and $A_2 = A_3 = \cdots = \emptyset$. Then countable additivity gives

$$\begin{aligned}\mu(A_1) &= \mu\left(\bigcup_{i=1}^{\infty} A_i\right) \\ &= \sum_{i=1}^{\infty} \mu(A_i) \\ &= \mu(A_1) + \mu(\emptyset) + \mu(\emptyset) + \ldots.\end{aligned}$$

Thus $\mu(\emptyset) = 0$.

(2) Take $A_{n+1} = A_{n+2} = \cdots = \emptyset$. Then with countable additivity the result follows.

(3)

$$\begin{aligned}\mu(B) &= \mu(A \cup (B \setminus A)) \\ &= \mu(A) + \mu((B \setminus A)) \\ &\geq \mu(A) && \text{since } \mu((B \setminus A)) \geq 0.\end{aligned}$$

(4) Given $A_1 \subseteq A_2 \subseteq \ldots$ define $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$. Then $B_i \cap B_j = \emptyset$ if $i \neq j$ with $A = \bigcup_{i=1}^{\infty} B_i$ and $A_n = B_1 \cup \cdots \cup B_n$ (see Figure 2.1). Thus

$$\begin{aligned}\mu(A) &= \sum_{i=1}^{\infty} \mu(B_i) && \text{by countable additivity} \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) \\ &= \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} B_i\right) && \text{by Theorem 2.2(2)} \\ &= \lim_{n\to\infty} \mu(A_n).\end{aligned}$$

(5) Given $A_1 \supseteq A_2 \supseteq \cdots \supseteq$ define $C_n = A_1 \setminus A_n = A_1 \cap A_n^c$. So we have $C_1 \subseteq C_2 \subseteq \ldots$ with $\mu(C_n) = \mu(A_1) - \mu(A_n)$ since $\mu(A_1) = \mu(C_n \cup A_n)$ and $\mu(A_1) < \infty$ (see Figure 2.2). Also we have $\bigcup_{n=1}^{\infty} C_n = A_1 \setminus A = A_1 \cap A^c = A_1 \cap \left(\bigcup_{i=1}^{\infty} A_n^c\right)$. So by Theorem 2.2(4):

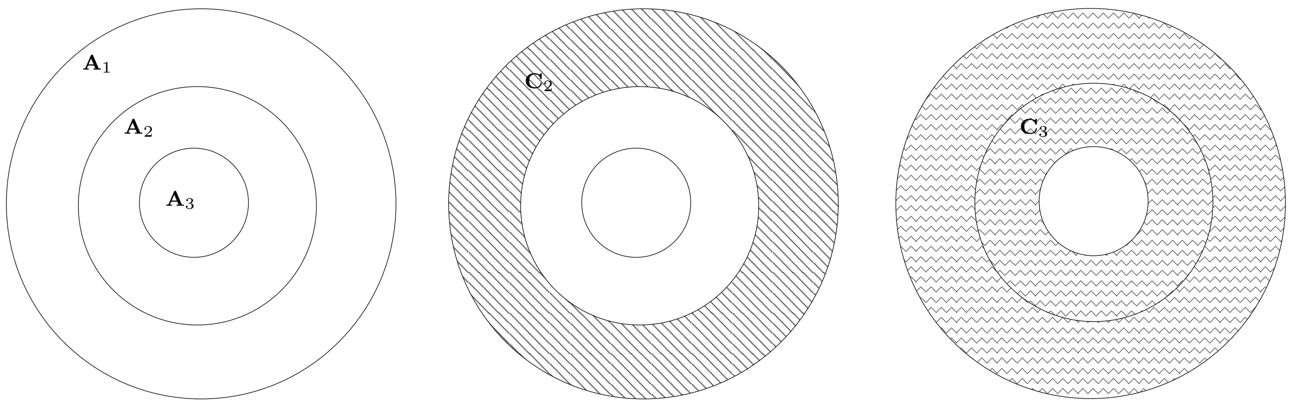

Figure 2.2: Creating sequence of set containments C_n

$$\begin{aligned}\mu(A_1 \setminus A) &= \mu(A_1) - \mu(A) \\ &= \lim_{n\to\infty} \mu(C_n) \\ &= \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)) \\ &= \mu(A_1) - \lim_{n\to\infty} \mu(A_n).\end{aligned}$$

Since $\mu(A_1) < \infty$ this implies

$$\mu(A) = \lim_{n\to\infty} \mu(A_n) < \infty.$$

$\square$

A general method to construct measures is via outer measure.

Definition 2.6. For a set $X \neq \emptyset$, an **outer measure** is a map $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that:

(1) $\mu^*(\emptyset) = 0$;

(2) if $A \subseteq B \subset X$ then $\mu^*(A) \leq \mu^*(B)$ (monotone property);

(3) if $A_n \subseteq X, n \in \mathbb{N}$ then
$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$$
(countable subadditivity)

For any collection of sets large enough and any positive function on them we can construct an outer measure.

Theorem 2.3. Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be a set of sets such that $\emptyset \in \mathcal{C}$, and for any $A \subseteq X$ $\exists \{C_j\}_{j=1}^{\infty} \subset \mathcal{C}$ such that $A \subseteq \bigcup\limits_{j=1}^{\infty} C_j$. Let $\gamma : \mathcal{C} \to [0, \infty]$ be any map such that $\gamma(\emptyset) = 0$. Define

$$\mu^*(A) := \inf \left\{ \sum_{j=1}^{\infty} \gamma(C_j) \;\middle|\; C_j \in \mathcal{C},\ A \subseteq \bigcup_{j=1}^{\infty} C_j \right\} \quad \forall\ A \subseteq X.$$

Then $\mu^*(A)$ is an outer measure.

Proof. We will prove that with these conditions μ^* satisfies the properties of outer measures:

(1) Let $A = \emptyset$ then $C = \emptyset \in \mathcal{C}$ covers A. So $0 \leq \mu^*(A) \leq \gamma(\emptyset) = 0 \Rightarrow \mu^*(\emptyset) = 0$.

(2) Let $A \subseteq B \subseteq X$, then each covering $\{C_j\}_{j=1}^{\infty} \subseteq \mathcal{C}$ of B is also a covering of A. Thus the infimum of $\mu^*(A)$ is over a larger set than the infimum of $\mu^*(B)$. Thus $\mu^*(A) \leq \mu^*(B)$.

(3) Let $A = \bigcup\limits_{j=1}^{\infty} A_j,\ A_j \subseteq X$ for $j = 1, 2 \ldots$ and let $\varepsilon > 0$. For each A_j $\exists$ covering $\{C_k^j\} \subset \mathcal{C}$ such that $A_j \subset \bigcup\limits_{k=1}^{\infty} C_k^j$. By the infimum in μ^* we have

$$\sum_{k=1}^{\infty} \gamma(C_k^j) \leq \mu^*(A_j) + \frac{\varepsilon}{2^j}. \tag{2.1}$$

Now since $A = \bigcup\limits_{j=1}^{\infty} A_j \Rightarrow A \subset \bigcup\limits_{j=1}^{\infty} \left(\bigcup\limits_{k=1}^{\infty} C_k^j \right)$. So

$$\begin{aligned} \mu^*(A) &= \sum_{j=1}^{\infty} \left(\sum_{k=1}^{\infty} \gamma(C_k^j) \right) \\ &\leq \sum_{j=1}^{\infty} \left(\mu^*(A_j) + \tfrac{\varepsilon}{2^j} \right) \quad \text{by (2.1)} \\ &= \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ and arbitrary we have

$$\mu^*(A) \leq \sum_{j=1}^{\infty} \mu^*(A_j).$$

□

Example 2.4. Let $X = \mathbb{R}^n$, $\mathcal{C} =$ all n-blocks and $\emptyset$. This satisfies the conditions of Theorem 2.3. Let $\gamma(C) = V(C) =$ n-volume of n-block $C \in \mathcal{C}$, $\gamma(\emptyset) = 0$. Then by Theorem 2.3 for $A \subseteq \mathbb{R}^n$

$$\mu^*(A) := \inf \left\{ \sum_{j=1}^{\infty} \gamma(C_j) \;\middle|\; C_j \in \mathcal{C},\ A \subseteq \bigcup_{j=1}^{\infty} C_j \right\}$$

is an outer measure, called the **Lebesgue outer measure** of $\mathbb{R}^n$.

Comparing the outer measure with the **outer volume** $\overline{V(A)}$ of Chapter 1, there are some noticable differences:

- A need not be bounded;
- coverings $\mathcal{C}$ can be infinite;
- coverings can overlap.

Claim. If $A \subset \mathbb{R}^n$ is bounded then $\mu^*(A) = \overline{V(A)}$.

Proof (Exercise)**.**

- If $\mathcal{A} \subset \mathcal{C}$ is a covering with overlaps, then $\exists$ partition (i.e. disjoint covering) $\mathcal{A}' \subset \mathcal{C}$ such that
$$\sum_k V(C'_k) \leq \sum_j V(C_k) \qquad C'_k \in \mathcal{A}',\ C_j \in \mathcal{A}.$$
- By convergence of the sums $\sum_{k=1}^{\infty} V(C_k)$ we can get within ε of limit with a finite covering.

$\square$

We have defined the outer measure on all of the subsets of a given set but to construct a measure from an outer measure we need to get an appropriate σ-algebra.

Definition 2.7.

(1) Let μ^* be an outer measure of $X \neq \emptyset$. Then a set $A \subseteq X$ is μ^***-measurable** if
$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c) \ \ \forall\ B \subseteq X. \tag{2.2}$$
Let $\mathcal{M}_{\mu^*} =$ set of μ^*-measurable subsets of X where $\mathcal{M}_{\mu^*} \subseteq \mathcal{P}(X)$.

(2) If $X = \mathbb{R}^n$, $\mu^* =$ Lebesgue outer measure, then a μ^*-measurable set A is called **Lebesgue measurable**.

Note that for an outer measure $\mu^*(B) \leq \mu^*(B \cap A) + \mu^*(B \cap A^c)$ by subadditivity of μ^*. So we only need to check that the converse inequality holds in proofs.

Theorem 2.4 (Caratheodory)**.** Let μ^* be an outer measure of $X \neq \emptyset$. Then

(1) $\mathcal{M}_{\mu^*}$ is a σ-algebra and

(2) μ^* restricted to $\mathcal{M}_{\mu^*}$ is a measure.

Proof.

- By symmetry of (2.2) if $A \in \mathcal{M}_{\mu^*}$ then $A^c \in M_{\mu^*}$. Trivially $\emptyset \in \mathcal{M}_{\mu^*} \Rightarrow X \in \mathcal{M}_{\mu^*}$. Thus $\mathcal{M}_{\mu^*}$ satisfies the first two conditions of a σ-algebra. We only need to show that it is closed with respect to countable unions. Firstly we will show that it is closed under finite unions.

  Let $A_1, A_2 \in \mathcal{M}_{\mu^*}$. Then $\forall\, B \subseteq X$

$$\begin{aligned}
\mu^*(B) &= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c) \\
&= \mu^*(B \cap \underbrace{(A_1 \cup A_2) \cap A_1}_{A_1}) + \mu^*(B \cap \underbrace{(A_1 \cup A_2) \cap A_1^c}_{A_1^c \cap A_2}) + \mu^*(B \cap (A_1 \cup A_2)^c) \\
&= \mu^*(B \cap (A_1 \cup A_2)) + \mu^*(B \cap (A_1 \cup A_2)^c) \qquad \text{as } A_1 \in \mathcal{M}_{\mu^*}.
\end{aligned}$$

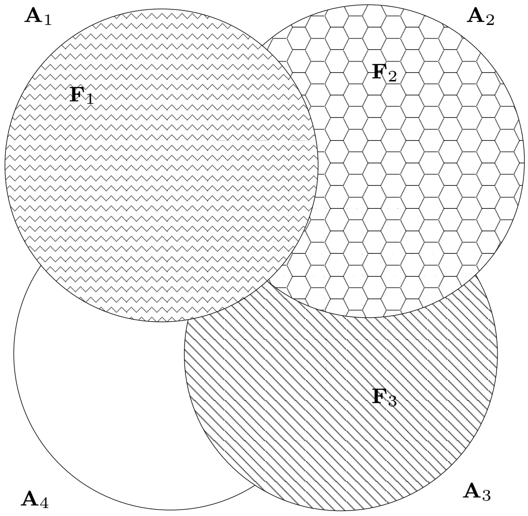

Figure 2.3: Creating collection of disjoint sets F_k

Thus $A_1 \cup A_2 \in \mathcal{M}_{\mu^*}$, inductively we can prove that this is true for finite unions. Now we need show that is is closed with respect to countable unions:

Let $A_j \in \mathcal{M}_{\mu^*},\ A = \bigcup\limits_{j=1}^{\infty} A_j$. We need to show that $A \in \mathcal{M}_{\mu^*}$). Firstly we "disjointify" A, let $F_1 := A_1,\ F_{k+1} := A_{k+1} \setminus \bigcup\limits_{j=1}^{k} A_k\ k \geq 1$. Then $F_i \cap F_j = \emptyset$ if $i \neq j$ (see Figure 2.3) and $F_i \in \mathcal{M}_{\mu^*}$ by what we have just proved above. So we have $A = \bigcup\limits_{k=1}^{\infty} F_k$.

Lemma 2.4.1.

$$\mu^* \left(B \cap \left(\bigcup_{j=1}^{n} F_j \right) \right) = \sum_{j=1}^{n} \mu^*(B \cap F_j) \ \ \forall B \subseteq X.$$

Proof.

$$\begin{aligned} \mu^* \left(B \cap \left(\bigcup_{j=1}^{n} F_j \right) \right) &= \mu^* \left(B \cap \left(\bigcup_{j=1}^{n} F_j \right) \cap F_n \right) + \\ &\quad \mu^* \left(B \cap \left(\bigcup_{j=1}^{n} F_j \right) \cap F_n^c \right) \\ &= \mu^*(B \cap F_n) + \mu^* \left(B \cap \bigcup_{j=1}^{n-1} F_j \right) \end{aligned}$$

by disjointness. Reapply this argument to the last term iteratively to obtain the desired result.

∇

Now

$$\begin{aligned}
\mu^*(B) &= \mu^*\left(B \cap \left(\bigcup_{j=1}^{n} F_j\right)\right) \\
&\qquad +\mu^*\left(B \cap \left(\bigcup_{j=1}^{n} F_j\right)^c\right) && \text{as } \mathcal{M}_{\mu^*} \text{ closed w.r.t. finite unions} \\
&\geq \mu^*\left(B \cup \left(\bigcup_{j=1}^{n} F_j\right)\right) + \mu^*(B \cap A^c) && \text{since } A^c \subseteq \left(\bigcup_{j=1}^{n} F_j\right)^c \\
&= \sum_{j=1}^{n} \mu^*(B \cap F_j) + \mu^*(B \cap A^c) && \text{by lemma 2.4.1.}
\end{aligned}$$

So

$$\begin{aligned}
\mu^*(B) &\geq \sum_{i=1}^{n} \mu^*(B \cap F_j) + \mu^*(B \cap A^c) \\
&\geq \mu^*(B \cap A) + \mu^*(B \cap A^c) && \text{by countable subadditivity of } \mu^* \\
&\geq \mu^*(B) && \text{by countable subadditivity of } \mu^*
\end{aligned}$$

Thus $\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c)\ \ \forall B \subseteq X$ i.e. $A \in \mathcal{M}_{\mu^*} \Rightarrow \mathcal{M}_{\mu^*}$ is a σ-algebra.

- We need to check countable additivity, Let $A = \bigcup_{j=1}^{\infty} A_j,\ A_j \in \mathcal{M}_{\mu^*}$ disjoint. Then

$$\begin{aligned}
\sum_{j=1}^{n} \mu^*(A_j) &= \mu^*\left(\bigcup_{j=1}^{n} A_j\right) && \text{by the proof of the lemma} \\
&\leq \mu^*(A) && \text{by } \bigcup_{j=1}^{n} A_j \subseteq A.
\end{aligned}$$

So

$$\begin{aligned}
\sum_{j=1}^{\infty} \mu^*(A_j) &\leq \mu^*(A) \\
&\leq \sum_{j=1}^{\infty} \mu^*(A_j) && \text{by countable subadditivity of } \mu^*.
\end{aligned}$$

Thus $\mu^*(A) = \sum_{j=1}^{\infty} \mu^*(A_j)$.

□

Thus from any outer measure (easily obtained from Theorem 2.3) we get a measure. If $\mu^* =$ Lebesgue outer measure on $\mathbb{R}^n$, then $\mu^* \restriction M_{\mu^*} = \mu$ is the **Lebesgue measure** of $\mathbb{R}^n$.

Theorem 2.5. Let μ^* be the Lebesgue outer measure on $\mathbb{R}^n$. Then $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\mu^*}$, i.e. Borel sets are Lebesgue measurable.

Proof. As $\mathcal{B}(\mathbb{R}^n)$ and $\mathcal{M}_{\mu^*}$ are σ-algebras, we only need to show all open sets are in $\mathcal{M}_{\mu^*}$. First note that all n-blocks are in $\mathcal{B}(\mathbb{R}^n)$ (an n-block can be constructed through intersections of open rectangles with closed rectangles see Figure 2.4). Thus each open set is a countable

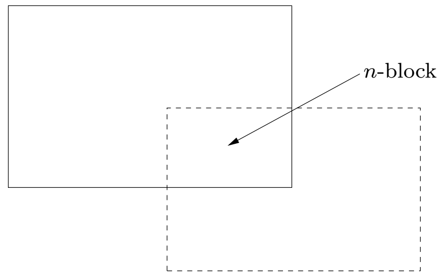

Figure 2.4: n-block from intersection of open and closed rectangle

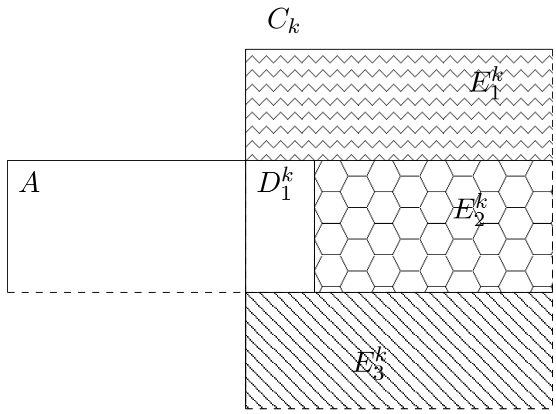

Figure 2.5: n-blocks that cover C_k

union of n-blocks[1]. So it suffices to show that all n-blocks are in $\mathcal{M}_{\mu^*}$ i.e. that if A is an n-block then

$$\mu^*(B) \geq \mu^*(B \cap A) + \mu^*(B \cap A^c).$$

This already holds if $\mu^*(B) = \infty$. So we assume $\mu^*(B) < \infty$. Let $\varepsilon > 0$, then $\exists$ a sequence of n-blocks $\{C_j\}$ such that $B \subseteq \bigcup_{j=1}^{\infty} C_j$ and

$$\sum_{k=1}^{\infty} V(C_k) \leq \mu^*(B) + \varepsilon.$$

For each k let $\{D_l^k\}$, $\{E_l^k\}$ be n-blocks such that

$$C_k \cap A \subseteq \bigcup_{l=1}^{\infty} D_l^k;$$
$$C_k \cap A^c \subseteq \bigcup_{l=1}^{\infty} E_l^k$$

and

$$\sum_{l=1}^{\infty} V(D_l^k) + \sum_{l=1}^{\infty} V(E_l^k) \leq V(C_k) + \tfrac{\varepsilon}{2^k}.$$

Thus

$$\mu^*(B \cap A) \leq \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} V(D_l^k) \qquad \text{and} \qquad \mu^*(B \cap A^c) \leq \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} V(E_l^k).$$

[1] Consider an open ball and take all the rational points inside the open ball. Construct n-blocks such that they are completely inside the open set using the argument that you can shrink the size of the n-blocks that are close to the boundary. For example $[q - \frac{1}{n}) \times [q + \frac{1}{n})$. Thus each open set is a countable union of n-blocks.

So

$$\begin{aligned}
\mu^*(B\cap A)+\mu^*(B\cap A^c) &\le \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left(V(D_l^k)+V(E_l^k)\right)\\
&\le \sum_{k=1}^{\infty}\left(V(C_k)+\tfrac{\varepsilon}{2^k}\right)\\
&= \sum_{k=1}^{\infty}V(C_k)+\varepsilon\\
&\le \mu^*(B)+\varepsilon.
\end{aligned}$$

Since ε arbitrary we have $B \in \mathcal{M}_{\mu^*}$.

□

Sometimes values of a measure are determined already on a small subset of its σ-algebra:

Theorem 2.6. Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be closed with respect to **finite** intersections. Let $\mathcal{A}(\mathcal{C}) = \sigma$-algebra generated by $\mathcal{C}$. If μ, ν are two finite measures on $\mathcal{A}(\mathcal{C})$ which coincide on $\mathcal{C}$ (i.e. $\mu(A) = \nu(A),\ A \in \mathcal{C}$) and $\mu(X) = \nu(X)$ then $\mu = \nu$ on $\mathcal{A}(\mathcal{C})$.

Proof. Define a **Dynkin class** as a family $\mathcal{D} \subseteq \mathcal{P}(X)$ if:

- $X \in \mathcal{D}$;
- $\mathcal{D}$ is closed with respect to proper set differences i.e. $B \subset A,\ A, B \in \mathcal{D} \Rightarrow A \setminus B \in \mathcal{D}$;
- $\mathcal{D}$ is closed with respect to countable unions of **increasing sets**.

Note that any σ-algebra is a Dynkin class. Now any intersection of Dynkin classes is a Dynkin class, hence there is a unique Dynkin class

$$\mathcal{D}(\mathcal{C}) := \cap\{\text{ Dynkin class } \mathcal{D} \mid \mathcal{C} \subseteq \mathcal{D}\}$$

which is said to be **generated** by $\mathcal{C} \subset \mathcal{P}(X)$.

Lemma 2.6.1. Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be closed with respect to finite intersections. Then $\mathcal{D}(\mathcal{C}) = \mathcal{A}(\mathcal{C}) = \sigma$-algebra generated by $\mathcal{C}$.

Proof. Since $\mathcal{A}(\mathcal{C})$ is a Dynkin class, $\mathcal{A}(\mathcal{C}) \supseteq \mathcal{D}(\mathcal{C})$. So we only need to show that $\mathcal{A}(\mathcal{C}) \subseteq \mathcal{D}(\mathcal{C})$. Since $X \in \mathcal{D}(\mathcal{C})$, let $A \in \mathcal{D}(\mathcal{C})$ and we have $X \setminus A = A^c \Rightarrow \mathcal{D}(\mathcal{C})$ is closed with respect to taking of complements. We need to show that $\mathcal{D}(\mathcal{C})$ is closed with respect to finite intersections. Any union $E = \bigcup_i E_i,\ E_i \in \mathcal{D}(\mathcal{C})$ can be written as a union of increasing sets:

$$F_k := \bigcup_{i=1}^{k} E_i \in \mathcal{D}(\mathcal{C})$$

so $E = \bigcup_{k=1}^{\infty} F_k \in \mathcal{D}(\mathcal{C})$. Now define

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

Since $\mathcal{C}$ is closed with respect to finite intersections, $\mathcal{C} \subset \mathcal{D}_1$. The identities

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

show $\mathcal{D}_1$ is closed with respect to proper set differences and countable unions of increasing sets (i.e. $\mathcal{D}_1$ is a Dynkin class containing $\mathcal{C}$). Thus $\mathcal{D}(\mathcal{C}) \subseteq \mathcal{D}_1 \subseteq \mathcal{D}(\mathcal{C}) \Rightarrow \mathcal{D}_1 = \mathcal{D}(\mathcal{C})$. Now define:

$$\mathcal{D}_2 = \{E \in \mathcal{D}(\mathcal{C}) \mid E \cap F \in \mathcal{D}(\mathcal{C}) \;\; \forall F \in \mathcal{D}(\mathcal{C})\}.$$

Since $\mathcal{D}_1 = \mathcal{D}(\mathcal{C})$ we get $\mathcal{C} \subset \mathcal{D}_2$ and $X \in \mathcal{D}_2$. By the same arguments as above, as $\mathcal{D}(\mathcal{C})$ is a Dynkin class, so is $\mathcal{D}_2$. Thus by $\mathcal{C} \subseteq \mathcal{D}_2 \subseteq \mathcal{D}(\mathcal{C})$ we get that $\mathcal{D}_2 = \mathcal{D}(\mathcal{C})$. Thus $\mathcal{D}(\mathcal{C})$ is closed with respect to with finite intersections and hence a σ-algebra.

∇

Let μ, ν be finite measures on $\mathcal{A}(\mathcal{C})$ as above. Let $\mathcal{D} = \{A \in \mathcal{A}(\mathcal{C}) \mid \mu(A) = \nu(A)\}$, $\mathcal{C} \subset \mathcal{D}$. Then $\mathcal{D}$ is a Dynkin class because

- $\mu(X) = \nu(X) \Rightarrow X \in \mathcal{D}$;
- If $A, B \in \mathcal{D}$ with $B \subset A$. Then

$$
\begin{aligned}
\mu(A \setminus B) &= \mu(A) - \mu(B) \\
&= \nu(A) - \nu(B) \\
&= \nu(A \setminus B).
\end{aligned}
$$

  Thus $A \setminus B \in \mathcal{D}$.

- If $E_n \in \mathcal{D}$ is an increasing sequence of sets, then by Theorem 2.2(4):

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

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

Thus $\mathcal{D}$ is a Dynkin class. Thus $\mathcal{D} \supseteq \mathcal{D}(\mathcal{C}) = \mathcal{A}(\mathcal{C})$ by lemma. Therefore $\mu = \nu$ on $\mathcal{A}(\mathcal{C})$.

$\square$

Corollary 2.7. Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be closed with respect to finite intersections and let μ, ν be σ-finite measures on $\mathcal{A}(\mathcal{C})$ which coincide on $\mathcal{C}$. If $\exists$ increasing sequencing $\{C_n\}_{n=1}^{\infty} \subset \mathcal{C}$ such that $X = \bigcup_{n=1}^{\infty} C_n$ and μ, ν are finite on each C_n then $\mu = \nu$.

Proof. Take increasing sequences $\{C_n\} \subset \mathcal{C}$ as above. For each n define μ_n, ν_n on $\mathcal{A}(\mathcal{C})$ by

$$\begin{aligned} \mu_n(A) &:= \mu(A \cap C_n) \\ \nu_n(A) &:= \nu(A \cap C_n) \qquad \forall\, A \in \mathcal{A}(\mathcal{C}). \end{aligned}$$

Then by Theorem 2.6, $\mu_n = \nu_n$ on $\mathcal{A}(\mathcal{C})$. But

$$\mu(A) = \lim_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \nu_n(A) = \nu(A) \qquad \forall\, A \in \mathcal{A}(\mathcal{C}).$$

Thus $\mu = \nu$ on $\mathcal{A}(\mathcal{C})$.

□

Corollary 2.8. The Lebesgue measure μ on $\mathbb{R}^n$ is the unique measure on $\mathcal{B}(\mathbb{R}^n)$ such that $\mu([t_1, s_1) \times \cdots \times [t_n, s_n)) = \prod_{i=1}^{n} (s_i - t_i)$ i.e. its n-volume.

Proof. The set $\mathcal{C}$ of n-blocks is closed with respect to finite intersections and $\mathcal{B}(\mathbb{R}^n) = \mathcal{A}(\mathcal{C})$. The sequence $C_k = [-k, k) \times \cdots \times [-k, k)$ is increasing, $\mathbb{R}^n = \bigcup_{k=1}^{\infty} C_k$ and μ is finite on each C_k. Thus by corollary 2.7 μ is uniquely determined by its values on $\mathcal{C}$ and by definition we have

$$\mu(C) = V(C) \qquad \forall\, C \in \mathcal{C}.$$

□

Theorem 2.9. Let ν be a measure on $\mathcal{B}(\mathbb{R}^n)$ which is finite on bounded sets in $\mathcal{B}(\mathbb{R}^n)$ and translation invariant (i.e. $\nu(A + \mathbf{x}) = \nu(A) \ \ \forall\, A \in \mathcal{B}(\mathbb{R}^n),\ \mathbf{x} \in \mathbb{R}^n$) Then $\nu = c.\mu$ for some $c \in (0, \infty)$ where μ is the Lebesgue measure on $\mathbb{R}^n$.

Proof. Let $I = [0, 1) \times \cdots \times [0, 1) =$ "unit cubes" then $\mu(I) = 1$. Set $\nu(I) = c$. Now I is the union of pair wise disjoint cubes of side length 2^{-k} and there are 2^{nk} of them. They are all translates of each other and hence have the same measure as any fixed one B. Thus

$$\nu(I) = 2^{nk}\nu(B) = c = c.\mu(I) = 2^{nk}c.\mu(B).$$

Thus $\nu(B) = c.\mu(B)$ $\forall$ cubes of length $2^{-k} \Rightarrow \nu(B) = c.\mu(B)$ $\forall$ cubes B of length 2^{-k}. Since the set of such cubes will generate all n-blocks

$$D = [t_1, s_1) \times \cdots \times [t_n, s_n)$$

by countable unions, hence $\nu(D) = c.\mu(D)$ for all n-blocks D, thus by corollary 2.8 $\nu(A) = c.\mu(A) \ \ \forall\, A \in \mathcal{B}(\mathbb{R}^n)$.

□

Theorem 2.10. For the Lebesgue measure μ on $\mathbb{R}^n$ we have:

(1) $\mu(A) = 0$ when $A \subset \mathbb{R}^n$ is countable;

(2) $\mu(K) < \infty$ when $K \in \mathcal{M}_{\mu^*}$ is bounded;

(3) $\mu(A) = \inf\{\mu(U) \mid U \text{ open},\ A \subseteq U\} \ \ \forall\, A \in \mathcal{M}_{\mu^*}$;

Figure 2.6: Cantor Set

(4) $\mu(A) = \sup\{\mu(K) \mid K \text{ compact},\ K \subseteq A\}\ \ \forall A \in \mathcal{M}_{\mu^*}$.

Proof.

(1) For a point $\{\mathbf{x}\} \subset \mathbb{R}^n$ for each $\varepsilon > 0\ \ \exists$ n-block $C \ni \mathbf{x}$ such that $V(C) < \varepsilon$. Thus $\mu^*(\{\mathbf{x}\}) = 0 \Rightarrow \mu(\{\mathbf{x}\})$. Let A be countable i.e. $A = \bigcup\limits_{i=1}^{\infty}(\{\mathbf{x}_i\})$. So

$$\mu(A) = \sum_{k=1}^{\infty} \mu(\{\mathbf{x}_k\}) = 0.$$

(2) $K \in \mathcal{M}_{\mu^*}$ such that K is bounded. This implies $K \subseteq C$ for an n-block $C = [-L, L) \times \cdots \times [-L, L)$ for L large enough. So $\mu(K) \leq \mu(C) = (2L)^n < \infty$.

(3) and (4) shall be proved later in Chapter 7 when we consider a larger class of measures (Radon measures).

$\square$

Example 2.5 (Cantor Set)**.** Let

$$\begin{aligned} K_0 &= [0,1] \\ K_1 &= [0,1] \setminus \left(\tfrac{1}{3}, \tfrac{2}{3}\right) \\ &\vdots \\ K_n &= K_{n-1} \text{ minus the open middle thirds of its intervals.} \end{aligned}$$

The Cantor set is $K := \bigcap\limits_{n=1}^{\infty} K_n$. This set is closed and bounded. K has no interior points (max length of intervals in $K_n = \left(\frac{1}{3}\right)^n$). To see K is uncountable, expand $x \in [0,1]$ in base 3:

$$x = 0.a_1a_2\cdots = \frac{a_1}{3} + \frac{a_2}{3^2} + \frac{a_3}{3^3} + \ldots$$

where $a_i \in \{0,1,2\}$. Note that we have

$$0.a_1a_2\ldots a_n22\cdots = 0.a_1\ldots a_{n-1}(a_n+1)00\ldots$$

and there are countably many points with this ambiguous expansions. Thus removing these will give us a well defined function. Then

$$\begin{aligned} K_1 &= \{x \in [0,1] \mid a_1 \neq 1\} \\ K_2 &= \{x \in [0,1] \mid a_1 \neq 1, a_2 \neq 1\} \\ &\vdots \end{aligned}$$

Thus $x \in K$ if all $a_i = 0$ or 2 in its tenary expansion. So there is a surjection $\varphi : K \to [0,1]$ by

$$\varphi(0.a_1a_2\dots) = 0.b_1b_2\dots = \frac{b_1}{2} + \frac{b_2}{2^2} + \dots$$

where

$$b_i = \frac{a_i}{2} = \begin{cases} a_i & \text{if } a_i = 0 \\ 1 & \text{if } a_i = 2 \end{cases}$$

and $x \in [0,1]$ is a binary expansion. Thus K is uncountable. Now $\mu(K_n) = \left(\frac{2}{3}\right)^n$ and $K_1 \supset K_2 \supset K_3 \supset \dots$. So by Theorem 2.2(5)

$$\mu(K) = \lim_{n\to\infty} \mu(K_n) = \lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0.$$

The inclusion $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\mu^*}$ is proper. Vitali sets are examples of non-measurable sets for the Lebesgue measurable on $\mathbb{R}$.

Definition 2.8. Let $(X, \mathcal{A}, \mu)$ is a measure space.

- Say $B \subset X$ is $\mu-$**null** if there is an $A \in \mathcal{A}$ with $B \subseteq A$ and $\mu(A) = 0$;
- If $\mathcal{A}$ contains **all** μ-null sets, we say that μ is **complete**;
- A property which holds for all $x \in X$ except for a μ-null set, $A \in \mathcal{A}$ is said to hold $\mu-$**almost everywhere**.

Example 2.6. Let $(X, \mathcal{M}_{\mu^*}, \mu)$ be the measure space obtained from an outer measure (by Theorem 2.4) then μ is complete.

Proof. If $B \subset X$ is μ-null, i.e. $B \subseteq A \in \mathcal{M}_{\mu^*}$ with $\mu^*(A) = 0$, then by monotone property $\mu^*(B) = 0$. Hence $\forall\, C \subseteq X$

$$\begin{aligned} \mu^*(C) &\leq \mu^*(C \cap B) + \mu^*(C \cap B^c) && \text{by countable subadditivity} \\ &= \mu^*(C \cap B^c) && \mu^*(B) = 0 \text{ and monotone property} \\ &\leq \mu^*(C). \end{aligned}$$

Thus $\mu^*(C) = \mu^*(C \cap B) + \mu^*(C \cap B^c)\ \ \forall\, C \subset X \Rightarrow B \in \mathcal{M}_{\mu^*}$.

□

Theorem 2.11. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\mathcal{N} := \{A \in \mathcal{A} \mid \mu(A) = 0\}$,

$$\overline{\mathcal{A}} := \{A \cup N \mid A \in \mathcal{A}, N \subset B \in \mathcal{N}\}$$

and define $\overline{\mu} : \overline{\mathcal{A}} \to [0, \infty]$ by $\overline{\mu}(A \cup N) := \mu(A)$ if $A \in \mathcal{A},\ N \subset B \in \mathcal{N}$. Then $(X, \overline{\mathcal{A}}, \overline{\mu})$ is a measure space and $\overline{\mu}$ is the unique extension of μ to a complete measure on $\overline{\mathcal{A}}$. Call $\overline{\mu}$ the **completion** of μ.

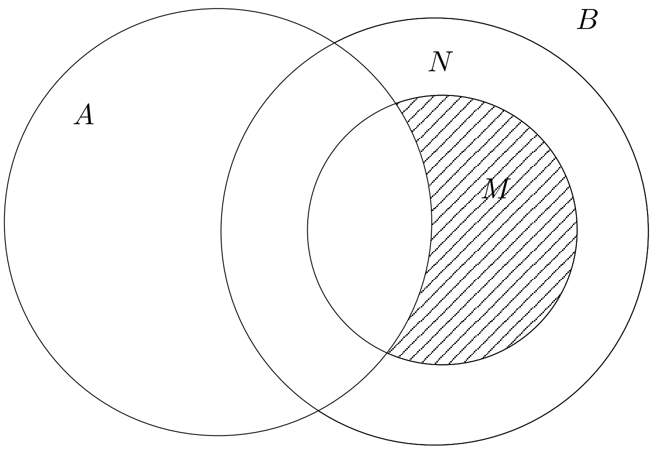

Figure 2.7: Creating a set M which is disjoint from A

Proof. As $\mathcal{A}$ and $\mathcal{N}$ are closed under countable unions, so is $\overline{\mathcal{A}}$. If $A \cup N \in \overline{\mathcal{A}}, A \in \mathcal{A}$ and $N \subset B \in \mathcal{N}$ we always have a disjoint union: $A \cup N = A \cup M \in \overline{\mathcal{A}}$ where $M = N \setminus A$ in which case $M \subset B \setminus A \in \mathcal{N}$ (see Figure 2.7). So we have $A \cap M = \emptyset = A \cap (B \setminus A)$ and $A \cup M = (A \cup (B \setminus A)) \cap ((B \setminus A)^c \cup M) \Rightarrow (A \cup M)^c = (A \cup (B \setminus A))^c \cup ((B \setminus A) \setminus M)$. Now $(A \cup (B \setminus A))^c \in \mathcal{A}$ and $(B \setminus A) \setminus M \subset B \setminus A \in \mathcal{N}$. Thus $(A \cup N)^c = (A \cup M)^c \in \overline{\mathcal{A}}$. So $\overline{\mathcal{A}}$ is a σ-algebra.

Now $\overline{\mu}$ is well defined since if $A_1 \cup N_1 = A_2 \cup N_2$ with $A_i \in \mathcal{A}$ and $N_i \subset B_i \in \mathcal{N}$ then

$$A_1 \subset A_2 \cup B_2 \in \mathcal{A} \Rightarrow \mu(A_1) \leq \mu(A_2) + \mu(B_2) = \mu(A_2).$$

Likewise $\mu(A_2) \leq \mu(A_1)$ i.e. $\mu(A_1) = \mu(A_2)$. Exercise - show $\overline{\mu}$ is complete and the only measure which extends μ to $\overline{\mathcal{A}}$.

□

The completion of a complete measure is just the original measure on the same σ-algebra.

Theorem 2.12. The Lebesgue measure $(\mathbb{R}^n, \mathcal{M}_{\mu^*}, \mu)$ is the completion of the Lebesgue measure on $\mathcal{B}(\mathbb{R}^n)$.

Proof. As $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\mu^*}$ and μ is complete on $\mathcal{M}_{\mu^*}$ it follows that $\overline{\mathcal{B}(\mathbb{R}^n)} \subseteq \mathcal{M}_{\mu^*}$. To show the converse. Let $A \in \mathcal{M}_{\mu^*}$ and $\varepsilon > 0$. Write A as a disjoint union of bounded sets $A_i \in \mathcal{M}_{\mu^*}$. For example

$$\begin{aligned} A_1 &:= \{\mathbf{x} \in A \mid \|\mathbf{x}\| \leq 1\} \\ &\vdots \\ A_k &:= \{\mathbf{x} \in A \mid k-1 \leq \|\mathbf{x}\| \leq k\}. \end{aligned}$$

Then $A = \bigcup_{k=1}^{\infty} A_k$ with $A_k \in \mathcal{M}_{\mu^*}$, $A_i \cap A_j = \emptyset$ if $i \neq j$ and $\mu(A_k) < \infty$. So we have

$$\mu(A) = \sum_{k=1}^{\infty} \mu(A_k).$$

By Theorem 2.10(4) for each j choose $K_j \subset A_j$ such that K_j is compact and

$$\mu(A_j) \leq \mu(K_j) + \frac{\varepsilon}{2^j} < \infty$$

i.e. $\mu(A_j \setminus K_j) \leq \frac{\varepsilon}{2^j}$. Let $K = \bigcup_{j=1}^{\infty} K_j$ then $K \in \mathcal{B}(\mathbb{R}^n)$ and

$$\begin{aligned} \mu(A \setminus K) &= \mu\left(\bigcup_{j=1}^{\infty}(A_j \setminus K_j)\right) \\ &= \sum_{j=1}^{\infty} \mu(A_j \setminus K_j) \\ &\leq \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} \\ &= \varepsilon. \end{aligned}$$

This can be done for every $\varepsilon > 0$. So $\exists$ an increasing sequence $K_{(1)} \subseteq K_{(2)} \subseteq K_{(3)} \subseteq \ldots (\Rightarrow A \setminus K_{(j)}$ decreasing) such that $\mu(A \setminus K_{(j)}) \leq \frac{1}{j}$. So

$$A \setminus \left(\bigcup_{j=1}^{\infty} K_{(j)}\right) = \bigcap_{j=1}^{\infty}(A \setminus K_j)$$

and hence by Theorem 2.2(5)

$$\mu\left(A \setminus \bigcup_{j=1}^{\infty} K_{(j)}\right) = \mu\left(\bigcap_{j=1}^{\infty}(A \setminus K_j)\right) = \lim_{j \to \infty} \mu(A \setminus K_j) \leq \lim_{j \to \infty} \frac{1}{j} = 0.$$

Thus $A = B \cup N$ where $B = \bigcup_{j=1}^{\infty} K_{(j)} \in \mathcal{B}(\mathbb{R}^n)$ and $N = A \setminus \bigcup_{j=1}^{\infty} K_{(j)}$ is μ-null i.e. $A \in \overline{\mathcal{B}(\mathbb{R}^n)}$

$\square$

Thus a Lebesgue measurable set is a union of a Borel set and a μ-null set.

Example 2.7. $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\mu^*}$ is proper for $n > 1$.

Proof. Let $\varphi : \mathbb{R} \to \mathbb{R}^n$ by $\varphi(t) = (t, 0, \ldots, 0)$. Then φ is continuous hence Borel. Let $V \subset \mathbb{R}$ be the Vitali set then $\varphi(V) \subset \mathbb{R}^n$ cannot be Borel (or else its inverse image $V \subset \mathbb{R}$ is Borel which is false). However $\varphi(V) \subset \{(t, 0, \ldots, 0) \mid t \in \mathbb{R}\} = x$-axis which is a null-set, hence $\varphi(V) \in \mathcal{M}_{\mu^*}$ and $\varphi(V) \notin \mathcal{B}(\mathbb{R}^n)$.

$\square$

Theorem 2.13. For μ the Lebesgue measure on $(\mathbb{R}^n, \mathcal{M}_{\mu^*})$, we have that all Jordan measurable sets are in $\mathcal{M}_{\mu^*}$ and on these μ coincides with the Jordan n-volumes.

Proof. Recall that on bounded sets $\mu^*(A) = \overline{V(A)}$. By Theorem 1.1, a bounded set A is Jordan measurable if and only if its boundary ∂A is of Jordan measure zero i.e. $V(\partial A) = 0$ $(= \mu^*(\partial A))$. Now A is disjoint union of its interior A° with its boundary ∂A ($\Rightarrow A^\circ$ is Borel). Thus A is a union of a Borel set A° with a null set, so by Theorem 2.12 $A \in \mathcal{M}_{\mu^*}$. Thus

$$|A| = \overline{V(A)} = \overline{V(A^\circ \cup \partial A)} = \overline{V(A^\circ)} = \mu^*(A^\circ) = \mu(A).$$

$\square$

Chapter 3

Measurable Functions

Starting from a measure space $(X, \mathcal{A}, \mu)$ we want to define integrals by generalising Riemann sums. The key point in generalising Riemann sums is 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 $\int f \, d\mu$ by sums:

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

The main problem we have we this approach is we need to ensure $f^{-1}\left(\left[\frac{j-1}{k}, \frac{j}{k}\right)\right) \in \mathcal{A}$.

Definition 3.1. Given σ-algebras $(X, \mathcal{A})$, $(Y, \mathcal{B})$ then a map $f : X \to Y$ is **measurable** if $f^{-1}(B) \in \mathcal{A} \;\; \forall B \in \mathcal{B}$. If $\mathcal{A}, \mathcal{B}$ are Borel σ-algebras, we say f is **Borel**. In the case $f : X \to [-\infty, \infty]$ assume that $[-\infty, \infty]$ has its standard Borel σ-algebra $\mathcal{B}([-\infty, \infty])$. Same convention for $f : X \to \mathbb{C}$ or $f : X \to \mathbb{C}^n$ unless otherwise specified.

From this definition it is clear that compositions of measurable maps are measurable.

Theorem 3.1. Let $\mathcal{A}$ be a σ-algebra for X. Let $f : X \to \mathbb{R}$ (or $[-\infty, \infty]$). The following are equivalent:

(1) f is $\mathcal{A}$-measurable;

(2) $f^{-1}(I) \in \mathcal{A} \;\; \forall$ intervals $I \subset \mathbb{R}$;

(3) $f^{-1}(U) \in \mathcal{A} \;\; \forall \; U \subseteq \mathbb{R}$;

(4) $f^{-1}([t, \infty)) \in \mathcal{A} \;\; \forall t \in \mathbb{R}$;

(5) $f^{-1}((-\infty, t)) \in \mathcal{A} \;\; \forall t \in \mathbb{R}$.

Proof. Clearly (1) $\Rightarrow$ all others, so we only need to prove that properties (2) – (5) $\Rightarrow$ (1). Note $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\mathbb{R}) = X$ and

$$\begin{aligned} f^{-1}(B^c) &= \{x \in X \mid f(x) \in B^c\} \\ &= \{x \in X \mid f(x) \in B\}^c \\ &= f^{-1}(B)^c. \end{aligned}$$

If $B_j \in \mathcal{B}(\mathbb{R})$, then

$$\begin{aligned} f^{-1}\left(\bigcup_{j=1}^{\infty} B_j\right) &= \{x \in X \mid f(x) \in \mathcal{B}_j \text{ for some } j\} \\ &= \bigcup_{j=1}^{\infty} \{x \in X \mid f(x) \in B_j\} \\ &= \bigcup_{j=1}^{\infty} f^{-1}(B_j). \end{aligned}$$

In each of the cases (2)-(5) we have a set of subsets of $\mathbb{R}$ which generates $\mathcal{B}(\mathbb{R})$. Thus any $B \in \mathcal{B}(\mathbb{R})$ can be obtained from these given sets by complements and countable unions. Since f^{-1} respects these operations it follows that $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(\mathbb{R})$. So f is $\mathcal{A}$-measurable.

$\square$

Exercise: Show that any non-decreasing function $f : [-\infty, \infty] \to [-\infty, \infty]$ is Borel.

Corollary 3.2. If X is a topological space then any continuous function $f : X \to \mathbb{R}$ is Borel.

Proof. f is continuous if and only if $f^{-1}(U)$ is open for each open set $U \subset \mathbb{R}$. Then by Theorem 3.1 (3) f is Borel. (Choose $\mathcal{A} = \mathcal{B}(X)$).

$\square$

Definition 3.2. A **simple function** $f : X \to \mathbb{R}$ is a finite linear combination of characteristic functions

$$f = \sum_{i=1}^{n} \alpha_i \cdot \chi_{A_i} \qquad \text{where } \alpha_i \in \mathbb{R}$$

which can always be written in the form

$$f = \sum_{i=1}^{m} \beta_i \cdot \chi_{B_i} \qquad \text{where } B_i \text{ are disjoint and } \beta_i \neq \beta_j \text{ if } i \neq j.$$

With this $f(X) = \{\beta_1, \ldots, \beta_m\}$ and f is $\mathcal{A}$-measurable if and only if $B_i \in \mathcal{A}$ and $A_i \in \mathcal{A}$ $\forall\, i$. Equivalently, f is a simple function if it has finite range.

Theorem 3.3. Let $\mathcal{A}$ be a σ-algebra on X and let $M_+(X) := \{f : X \to [0, \infty] \mid f \text{ is } \mathcal{A} - \text{measurable }\}$. Let $f \in M_+(X)$, Then there is a sequence of simple functions $s_1, s_2, \dots \in M_+(X)$ on X such that:

(1) $0 \leq s_1 \leq s_2 \leq \ldots$ (increasing);

(2) $\lim_{n\to\infty} s_n(x) = f(x) \qquad \forall\, x.$

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

$$A_{n,k} := f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right) = \left\{x \in X \;\middle|\; \frac{k-1}{2^n} \leq f(x) \leq \frac{k}{2^n}\right\}$$

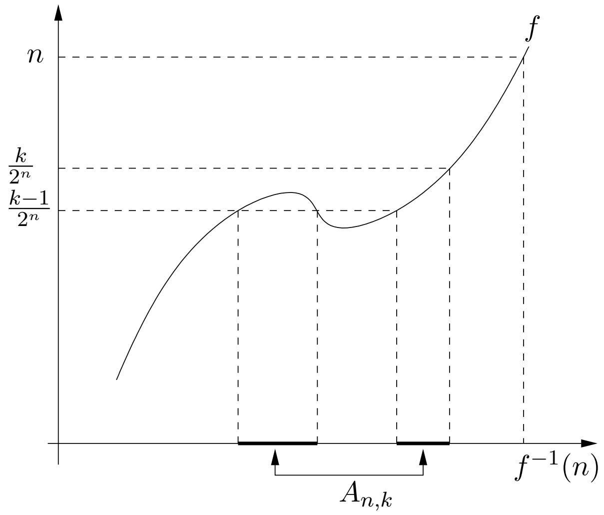

Figure 3.1: Each $A_{n,k}$ for $n \in \mathbb{N}$ and $k \in \{1, 2, \ldots, n \cdot 2^n\}$

Then $A_{n,k} \in \mathcal{A}$ as f is measurable (see Figure 3.1). Define

$$s_n := \sum_{k=1}^{n.2^n} \left(\frac{k-1}{2^n}\right) \cdot \chi_{A_{n,k}} + n \cdot \chi_{f^{-1}([n,\infty])}.$$

Then $s_n \leq s_{n+1}$ since if $x \in A_{n,k}$ then

$$s_n(x) = \left(\frac{k-1}{2^n}\right) \cdot \chi_{A_{n,k}}(x) \leq \inf\{f(y) \mid y \in A_{n,k}\}.$$

Since for any x $\exists\, N$ large enough such that

$$f(x) - 2^{-n} \leq s_n(x) \leq f(x) \qquad \forall\, n > N.$$

Therefore $\lim_{n\to\infty} s_n(x) = f(x)$.

$\square$

Note that s_n was constructed from $f^{-1}\left(\left[\frac{k-1}{j}, \frac{k}{j}\right)\right)$ i.e. we subdivide the **range** not the domain.

Recall

$$\limsup_n a_n := \inf_k \left(\sup_{n\geq k} a_n\right) = \lim_{k\to\infty} \left(\sup_{n\geq k} a_n\right)$$
$$\liminf_n a_n := \sup_k \left(\inf_{n\geq k} a_n\right) = \lim_{k\to\infty} \left(\inf_{n\geq k} a_n\right).$$

If $\lim_{n\to\infty} a_n$ $\exists$, then $\lim_{n\to\infty} a_n = \limsup_n a_n = \liminf_n a_n$.

Theorem 3.4. Let $\mathcal{A}$ be a σ-algebra on X and let $f_n : X \to [-\infty, \infty]$ be a sequence of $\mathcal{A}$-measurable functions on X. Then

(1) $\sup_n f_n$ and $\inf_n f_n$ are measurable;

(2) $\limsup_n f_n$ and $\liminf_n f_n$ are measurable;

(3) If $\lim_{n\to\infty} f_n(x)$ exists $\forall\, x \in X$ then $f := \lim_{n\to\infty} f_n$ is measurable.

Proof.

(1) Let $h(x) := \sup_n f_n(x)$, then

$$\begin{aligned} h^{-1}([-\infty,t]) &= \{x \in X \mid \sup_n f_n(x) \le t\} \\ &= \bigcap_n \{x \in X \mid f_n(x) \le t\} \\ &= \bigcap_n f_n^{-1}([-\infty,t]) \in \mathcal{A} \end{aligned}$$

as f_n are measurable. Since intervals $[-\infty,t]$ generate $\mathcal{B}([-\infty,\infty])$, it follows that h is measurable. As for $\inf_n f_n$ let $g(x) := \inf_n f_n(x)$, then

$$\begin{aligned} g^{-1}([-\infty,t)) &= \{x \in X \mid \inf_n f_n(x) < t\} \\ &= \bigcup_n \{x \in X \mid f_n(x) < t\} \\ &= \bigcup_n f_n^{-1}([-\infty,t)) \in \mathcal{A} \end{aligned}$$

$\Rightarrow g$ is measurable.

(2) Let $h_k(x) := \sup_{n\ge k} f_n(x)$, $g_k(x) := \inf_{n\ge k} f_n(x)$. Then these are both measurable by Theorem 3.4(1). So $\limsup_n f_n = \inf_k h_k$ and $\liminf_n f_n = \sup_k g_k$. By Theorem 3.4(1) it also follows that these two are measurable.

(3) If $\lim_{n\to\infty} f_n(x)$ exists $\forall\, x$, then $\lim_{n\to\infty} f_n = \limsup_n f_n$ which is measurable by Theorem 3.4(2).

Corollary. $f \in M_+(X)$ if and only if it is a pointwise limit of simple functions in $M_+(X)$.

As a consequence of Theorem 3.4(1), $f : X \to [-\infty,\infty]$ can be written in its positve and negative components:

$$f_+(x) := \sup\{f(x), 0\}$$
$$f_-(x) := -\inf\{f(x), 0\}$$

which are both measurable and $f = f_+ - f_-$ with $f_\pm \ge 0$, $(f_+)\cdot(f_-) = 0$ and $|f| = f_+ + f_-$ (see Figure 3.2). Thus f is $\mathcal{A}$-measurable if and only if $f_+, f_- \in M_+(X)$.

Theorem 3.5. Let $\mathcal{A}$ be a σ-algebra on X. Then

(1) A function $f : X \to \mathbb{R}^n$ is measurable if and only if the component functions $f_i : X \to \mathbb{R}$ are measurable $\forall\, i$, where $f(x) = (f_1(x), \ldots, f_n(x))$.

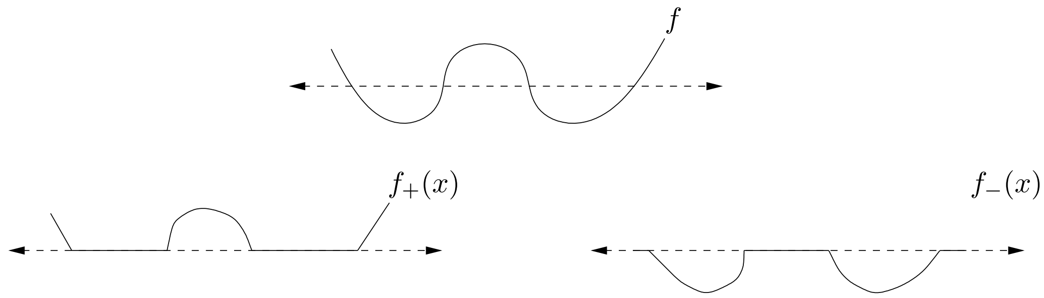

Figure 3.2: Decomposition of f into f_+ and f_-

(2) A function $f : X \to \mathbb{C}$ is measurable if and only if $u = \operatorname{Re}(f)$, $v = \operatorname{Im}(f)$ are measurable. Moreover $|f|$ is also measurable as is

$$\omega(x) = \begin{cases} \frac{f(x)}{|f(x)|} & \text{if } f(x) \neq 0 \\ 1 & \text{if } f(x) = 0. \end{cases}$$

So we get the polar decomposition $f = |f|w$ (note that $|w| = 1$).

(3) If $f, g : X \to \mathbb{C}$ are measurable and $\lambda \in \mathbb{C}$. Then so is λf, $f + g$ and $f \cdot g$.

(4) If $f : X \to \mathbb{C}$ is measurable then $\exists$ simple measurable functions $\varphi_n : X \to \mathbb{C}$ such that

$$0 \leq |\varphi_1| \leq |\varphi_2| \leq \cdots \leq |f| \qquad \text{and} \qquad \lim_{n\to\infty} \varphi_n(x) = f(x) \qquad \forall\, x \in X.$$

(5) If $f_n : X \to \mathbb{C}$ are measurable functions such that $\lim_{n\to\infty} f_n(x) = f(x)$ exists $\forall\, x$, then $f : X \to \mathbb{C}$ is measurable.

Proof.

(1) Let $C = (t_1, s_1) \times \cdots \times (t_n, s_n)$ be an open rectangular block in $\mathbb{R}^n$. The whole of $\mathcal{B}(\mathbb{R}^n)$ is generated by these. Let $f_1, \ldots, f_n$ be measurable, then

$$\begin{aligned} f^{-1}(C) &= \{x \in X \mid f(x) = (f_1(x), \ldots, f_n(x)) \in C\} \\ &= f_1^{-1}((t_1, s_1)) \cap \cdots \cap f_n^{-1}((t_n, s_n)) \in \mathcal{A} \end{aligned}$$

since f_i s' are measurable $\forall\, i$. Thus f is measurable. Conversely, let $f : X \to \mathbb{R}^n$ be measurable, then $f_i = p_i \circ f$ where $p_i : \mathbb{R}^n \to \mathbb{R}$ is the projection onto the i^{th} coordinate. Since p_i is continuous it is Borel, hence f_i is measurable (composition of measurable maps).

(2) The topology of $\mathbb{C}$ is just that of $\mathbb{R}^2$ so by Theorem 3.5(1), we have $f : X \to \mathbb{C}$ is measurable if and only if u, v are measurable. The map $T : \mathbb{C} \to \mathbb{R}$ by $T(z) = |z|$ is continuous, so since $|f| = T \circ f$ it follows that $|f|$ is measurable. The map $R : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ by $R(z) = \frac{z}{|z|}$ is continuous, so since

$$\omega(x) = R(f(x)) \cdot \chi_{(f^{-1}(0))^c}(x) + \chi_{f^{-1}(0)}(x)$$

it follows from Theorem 3.5(3) that ω is measurable.

(3) As f is measurable if and only if $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are measurable, it suffices to prove them for f, g real. For this case, define a measurable map $F : X \to \mathbb{R}^2$ by $F(x) = (f(x), g(x))$ and the maps $T : \mathbb{R}^2 \to \mathbb{R}$ by $T(s,t) = s+t$ and $S : \mathbb{R}^2 \to \mathbb{R}$ by $S(s,t) = s \cdot t$ are both continuous. So since $f + g = T \circ F$ and $f \cdot g = S \circ F$ it follows that $f + g$ and $f \cdot g$ are measurable. Likewise $M : \mathbb{R} \to \mathbb{R}$ by $M(t) = \lambda t$ is continuous, so $\lambda f = M \circ f$ is measurable.

(4) Let $f = g + ih = (g_+ - g_-) + i(h_+ - h_-)$ for g,h real-valued and apply Theorem 3.3 to obtain simple functions $\psi_n^+, \psi_n^-, \zeta_n^+$ and ζ_n^- increasing to limits g_+, g_-, h_+ and h_- respectively. Let $\varphi_n = (\psi_n^+ - \psi_n^-) + i(\zeta_n^+ - \zeta_n^-)$ then $\varphi_n \to f$ and

$$\begin{aligned}
|\varphi_n|^2 &= (\psi_n^+ - \psi_n^-)^2 + (\zeta_n^+ - \zeta_n^-)^2 \\
&= (\psi_n^+)^2 + (\psi_n^-)^2 + (\zeta_n^+)^2 + (\zeta_n^-)^2 \\
&\leq (\psi_{n+1}^+)^2 + (\psi_{n+1}^-)^2 + (\zeta_{n+1}^+)^2 + (\zeta_{n+1}^-)^2 \\
&= |\varphi_{n+1}|^2 \\
&\leq (g_+)^2 + (g_-)^2 + (h_+)^2 + (h_-)^2 \\
&= |f|^2.
\end{aligned}$$

(5) We have

$$f(x) = u(x) + iv(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} u_n(x) + i \lim_{n\to\infty} v_n(x)$$

i.e. $\lim_{n\to\infty} u_n(x) = u(x)$ and $\lim_{n\to\infty} v_n(x) = v(x)$ since u_n, v_n are measurable by Theorem 3.4, and so are u, v. Thus f is measurable by Theorem 3.5(2).

□

Chapter 4

Integration of Positive Functions

Definition 4.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $s = \sum_{i=1}^{n} \alpha_i \chi_{A_i} \in M_+(X)$ $(\alpha_i \geq 0)$ be a simple function. Its **integral** is

$$\int_X s \, d\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i)$$

(note that $\mu(A_i) = \infty$ is possible).

Definition 4.2. Let $(X, \mathcal{A}, \mu)$ be a measure space and functions $f \in M_+(X)$, then its **Lebesgue integral** with respect to μ is

$$\int_X f \, d\mu := \sup \left\{ \int_X s \, d\mu \mid s \in M_+(X) \text{ simple function such that } 0 \leq s \leq f \right\}.$$

By Theorem 3.3 this supremum exists but may be ∞. If $E \in \mathcal{A}$ define

$$\int_E f \, d\mu := \int_X \chi_E \cdot f \, d\mu.$$

If f is simple, the Lebesgue integral is the integral.

Proposition 4.1. For a measure space $(X, \mathcal{A}, \mu)$ consider $f, g \in M_+(X)$ 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 function $s \in M_+(X)$ such that $0 \leq s \leq f$ also satisfies $0 \leq s \leq g$, hence as the integral is the supremum over these:

$$\int_X f \, d\mu \leq \int_X f \, d\mu.$$

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

□

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 4.2 (Monotone Convergence Theorem - MCT / Beppo-Levi Theorem)**.** For a measure space $(X, \mathcal{A}, \mu)$ let $f_n \in M_+(X)$ such that

(1) $0 \le f_1(x) \le f_2(x) \le \cdots \le \infty \quad \forall\, x \in X$;

(2) $f(x) := \lim_{n\to\infty} f_n(x) \quad \forall\, x \in X$.

Then $f \in M_+(X)$ and

$$\int_X f \; d\mu = \lim_{n\to\infty} \int_X f_n \; d\mu,$$

i.e. $\int_X \lim_{n\to\infty} f_n(x) \; d\mu(x) = \lim_{n\to\infty} \int_X f_n \; d\mu$.

Proof. By Theorem 3.4(3) we have $f \in M_+(X)$. As f_n is increasing Proposition 4.1 $\Rightarrow$

$$0 \le \int_X f_1 \; d\mu \le \int_X f_2 \; d\mu \le \cdots \le \int_X f_n \; d\mu.$$

Thus $\exists\ \alpha \in [0, \infty]$ such that

$$\lim_{n\to\infty} \int_X f_n \; d\mu = \alpha \le \int_X f \; d\mu. \tag{4.1}$$

To prove the converse inequality in (4.1) let $c \in (0, 1)$ and $s \in M_+(X)$ simple function such that $0 \le s \le f$. Define

$$A_n := \{x \in X \mid f_n(x) \ge cs(x)\} = (f_n - cs)^{-1}([0, \infty]) \in \mathcal{A} \qquad \text{by Theorem 3.5(3).}$$

Since

- $f_n \le f_{n+1}$ we get $A_1 \subseteq A_2 \subseteq \ldots$ and $X = \bigcup_{n=1}^{\infty} A_n$;
- for each x, $f_n(x)$ eventually exceeds $cs(x) < f(x)$ and $f(x) = \lim_{n\to\infty} f_n(x)$;
- $f_n \ge \chi_{A_n} \cdot f_n \ge c \cdot \chi_{A_n} \cdot s$

we have

$$\int_X f_n \; d\mu \ge \int_{A_n} f_n \; d\mu \ge \int_{A_n} cs \; d\mu. \tag{4.2}$$

Let $s = \sum_{i=1}^{k} \alpha_i \chi_{B_i}$, then

$$\int_{A_n} cs \; d\mu = \sum_{i=1}^{k} c\alpha_i \mu(B_i \cap A_n) \xrightarrow[\infty]{n} \sum_{i=1}^{k} c\alpha_i \mu(B_i) = c \int_X s \; d\mu.$$

So by the inequalities in (4.2):

$$\alpha := \lim_{n\to\infty} \int_X f_n \; d\mu \ge c \int_X s \; d\mu \qquad \forall\, c \in (0, 1).$$

Thus $\alpha \ge \int_X s \; d\mu \;\; \forall\, s \in M_+(X)$ simple such that $0 \le s \le f$. So by definition $\alpha \ge \int_X f \; d\mu$ hence we get equality in Equation (4.1).

$\square$

By MCT we get the usual properties of the integral:

Theorem 4.3. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f, g \in M_+(X)$. Then

(1) $\int_E c \cdot f \; d\mu = c \int_E f \; d\mu \qquad \forall\, c \in [0, \infty),\ E \in \mathcal{A}$;

(2) $\int_E (f+g) \; d\mu = \int_E f \; d\mu + \int_E g \; d\mu \qquad \forall\, E \in \mathcal{A}$;

(3) $\int_E f \; d\mu = 0$ if and only if $\chi_E \cdot f = 0$ μ-a.e. $\forall\, E \in \mathcal{A}$;

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

Proof.

(1) By Theorem 3.3 we have simple functions $s_n, t_n \in M_+(X)$ with $0 \le s_1 \le s_2 \le \cdots \le f$, $0 \le t_1 \le t_2 \le \cdots \le g$ and $\lim\limits_{n\to\infty} s_n(x) = f(x)$, $\lim\limits_{n\to\infty} t_n(x) = g(x) \ \ \forall\, x$. Then cs_n is a similar increasing sequence with limit cf and $s_n + t_n$ is an increasing sequence with limit $f + g$. Thus by MCT

$$\int_E cf \; d\mu = \lim_{n\to\infty} \int_E cs_n \; d\mu = c \lim_{n\to\infty} \int_E s_n \; d\mu = c \int_E f \; d\mu$$

and

$$\int_E (f+g) \; d\mu = \lim_{n\to\infty} \int_E (s_n + t_n) \; d\mu \qquad \text{by MCT.} \tag{4.3}$$

(2) Let simple functions $s_n = \sum\limits_{i=1}^{k} \alpha_i \cdot \chi_{A_i}$, $t_n = \sum\limits_{i=1}^{l} \beta_j \cdot \chi_{B_j}$. Assume $\bigcup\limits_{i=1}^{k} A_i = \bigcup\limits_{j=1}^{l} B_j$ by adding zero terms (i.e. with some $\alpha_i = 0$ or $\beta_i = 0$), then

$$s_n = \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_i \cdot \chi_{A_i \cap B_j} \qquad t_n = \sum_{i=1}^{k} \sum_{j=1}^{l} \beta_j \cdot \chi_{A_i \cap B_j}.$$

So

$$\begin{aligned}
\textstyle\int_E (s_n + t_n) \; d\mu &= \textstyle\int_E \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{l} (\alpha_i + \beta_j) \cdot \chi_{A_i \cap B_j} \; d\mu \\
&= \sum_{i=1}^{k} \sum_{j=1}^{l} (\alpha_i + \beta_j) \cdot \mu(A_i \cap B_j \cap E) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_i \cdot \mu(A_i \cap B_j \cap E) + \sum_{i=1}^{k} \sum_{j=1}^{l} \beta_j \cdot \mu(A_i \cap B_j \cap E) \\
&= \textstyle\int_E s_n \; d\mu + \int_E t_n \; d\mu.
\end{aligned}$$

So Equation (4.3) $\Rightarrow$

$$\int_E (f+g) \; d\mu = \lim_{n\to\infty} \left(\int_E s_n \; d\mu + \int_E t_n \; d\mu \right) = \int_E f \; d\mu + \int_E g \; d\mu.$$

(3) Let $\chi_E \cdot f = 0$ μ-a.e., i.e. $\mu(F) = 0$ where $F = \{x \in X \mid \chi_E(x) \cdot f(x) \neq 0\} = (f^{-1}(0))^c \cap E$. Now by Theorem 4.3(2)

$$\int_E f \, d\mu = \int_F f \, d\mu + \int_{E \setminus F} f \, d\mu$$

and by MCT

$$\int_F f \, d\mu = \lim_{n \to \infty} \int_F s_n \, d\mu = 0$$

where $s_n \to f$ in an increasing fashion, since $\int_F s_n \, d\mu = \sum_{i=1}^{k} \alpha_i \cdot \mu(A_i \cap F) = 0$. Thus $\int_{E \setminus F} f \, d\mu = 0$ since $0 \leq \chi_{E \setminus F} \cdot s_n \leq \chi_{E \setminus F} \cdot f = 0$. So $\int_E f \, d\mu = 0$. Conversely let $\int_E f \, d\mu = 0$. Let $E_n = f^{-1}([\frac{1}{n}, \infty)) \cap E$. Then

$$\bigcup_{n=1}^{\infty} E_n = F = \{x \in E \mid f(x) > 0\}.$$

If $\mu(E_n) > 0$ for some n, then since $f \restriction E_n \geq \frac{1}{n}$

$$\int_E f \, d\mu \geq \int_{E_n} f \, d\mu \geq \frac{1}{n} \mu(E_n) > 0$$

which contradicts $\int_E f \, d\mu = 0$. Thus $\mu(E_n) = 0 \quad \forall n$, hence $\mu(F) = 0$. That is $\chi_E \cdot f = 0$ μ-a.e.

(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

$$0 \leq f_1 \leq f_2 \leq \ldots \qquad \text{and} \qquad \lim_{k \to \infty} f_k(x) = f(x) \quad \forall x \in A.$$

So

$$\begin{aligned} \nu(A) &= \textstyle\int_A f \, f\mu \\ &= \lim_{k \to \infty} \textstyle\int_A f_k \, d\mu \qquad \text{by MCT} \\ &= \lim_{k \to \infty} \textstyle\sum_{n=1}^{k} \int_{A_n} f \, d\mu \\ &= \textstyle\sum_{n=1}^{\infty} \nu(A_n). \end{aligned}$$

□

Theorem 4.4 (Fatou's Lemma)**.** For a measure space $(X, \mathcal{A}, \mu)$, let $f_n \in M_+(X)$ then

$$\int_X \liminf_n f_n \, d\mu \leq \liminf_n \int_X f_n \, d\mu.$$

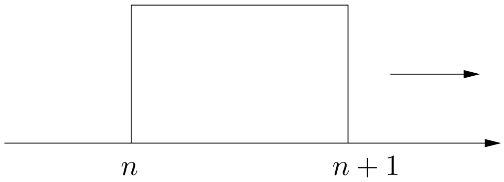

Figure 4.1: Characteristic function on the interval $[n, n+1]$

Proof. By definition

$$\liminf_n f_n = \lim_{k\to\infty} g_k \geq g_k = \inf_{n\geq k} f_n \in M_+(X)$$

and thus $g_1 \leq g_2 \leq \dots$ with $g_k \leq f_n \ \ \forall n \geq k$. So $\int_X g_k \, d\mu \leq \int_X f_n \, d\mu \ \ \forall n \geq k$.

$$\Rightarrow \int_X g_k \, d\mu \leq \inf_{n\geq k} \left(\int_X f_n \, d\mu \right). \tag{4.4}$$

Therefore by MCT

$$\begin{aligned} \lim_{k\to\infty} \textstyle\int_X g_n \, d\mu &= \textstyle\int_X \lim_{k\to\infty} g_k \, d\mu \\ &= \textstyle\int_X (\liminf_n f_n) \, d\mu \\ &\leq \liminf_n \left(\textstyle\int_X f_n \, d\mu \right) \quad \text{by inequality (4.4).} \end{aligned}$$

□

Example 4.1 (Moving Hump)**.** See Figure 4.1. The inequality in Fatou's Lemma may be a strict inequality. Let $f_n = \chi_{[n,n+1]}$. Then $\lim_{n\to\infty} f_n(x) = 0 \ \ \forall x$ and $\int_{\mathbb{R}} f_n \, d\mu = 1$. So

$$\int_{\mathbb{R}} \liminf_n f_n \, d\mu = 0 < \liminf_n \int_{\mathbb{R}} f_n \, d\mu = 1.$$

Chapter 5

Integration of Non-positive Functions

Definition 5.1. Let $(X, \mathcal{A}, \mu)$ be a measure space and define

$$\mathcal{L}^1(\mu) := \left\{ f : X \to [-\infty, \infty] \textbf{ or } \mathbb{C} \;\middle|\; f \text{ is measurable and } \int_X |f| \; d\mu < \infty \right\}.$$

A $f \in \mathcal{L}^1(\mu)$ is called **(absolutely) integrable** with respect to μ. For $f \in \mathcal{L}^1(\mu),\ f : X \to [-\infty, \infty]$ define

$$\int_E f \; d\mu := \int_E f_+ \; d\mu - \int_E f_- \; d\mu \qquad \text{where } f = f_+ - f_- \text{ (as above).}$$

As $f_\pm \leq |f| = f_+ + f_-$, we get $f \in \mathcal{L}^1(\mu)$ if and only if f_+ **and** $f_- \in \mathcal{L}^1(\mu)$. By using functions from $\mathcal{L}^1(\mu)$, we avoid undefined expressions $\infty - \infty$.

If $f : X \to \mathbb{C},\ f = u + iv$, as $|u|, |v| \leq |u + iv| \leq |u| + |v|$ we have that f is integrable if and only if u **and** v are integrable.

Definition 5.2. Let $f : X \to \mathbb{C}$ (or $[-\infty, \infty]$), $f \in \mathcal{L}^1(\mu),\ f = u + iv$ with u, v real-valued. Then

$$\int_E f \; d\mu := \int_E u_+ \; d\mu - \int_E u_- \; d\mu + i \int_E v_+ \; d\mu - i \int_E v_- \; d\mu \qquad \forall\, E \in \mathcal{A}.$$

Where $u = u_+ - u_- = \operatorname{Re}(f)$ and $v = v_+ - v_- = \operatorname{Im}(f)$.

Theorem 5.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, $f, g : X \to \mathbb{C}$ (similar for range $[-\infty, \infty]$)

(1) If $f, g \in \mathcal{L}^1(\mu),\ \alpha, \beta \in \mathbb{C}$ then

$$\int_X (\alpha f + \beta g) \; d\mu = \alpha \int_X f \; d\mu + \beta \int_X g \; d\mu;$$

(2) If $f \in \mathcal{L}^1(\mu)$ then

$$\left| \int_X f \; d\mu \right| \leq \int_X |f| \; d\mu;$$

(3) If $f \in \mathcal{L}^1(\mu)$ then $\int_E f \; d\mu = 0 \ \ \forall\, E \in \mathcal{A}$ if and only if $f = 0$ μ-a.e.

Proof.

(1) By the additivity of the real and imaginary parts, it suffices to prove that

$$\int_X (f+g)\ d\mu = \int_X f\ d\mu + \int_X g\ d\mu \qquad \text{and} \qquad \int_X \alpha f\ d\mu = \alpha \int_X f\ d\mu$$

for real-valued functions f, g and $\alpha \in \mathbb{C}$. Let $h = f+g$ then $h_+ - h_- = (f_+ - f_-) + (g_+ - g_-)$

$$\Rightarrow h_+ + f_- + g_- = f_+ + g_+ + h_- \geq 0.$$

Since we have transformed the equation to give us positive functions we can use our results from Chapter 4 on integration of positive functions.

$$\int h_+\ d\mu + \int f_-\ d\mu + \int g_-\ d\mu = \int f_+\ d\mu + \int g_+\ d\mu + \int h_-\ d\mu.$$

By Theorem 4.3(2) we have

$$\begin{aligned} \textstyle\int h\ d\mu &= \textstyle\int h_+\ d\mu - \int h_-\ d\mu \\ &= \textstyle\int f_+\ d\mu - \int f_-\ d\mu + \int g_+\ d\mu - \int g_-\ d\mu \\ &= \textstyle\int f\ d\mu - \int g\ d\mu. \end{aligned}$$

Now we prove $\int \alpha f\ d\mu = \alpha \int f\ d\mu \quad \alpha \in \mathbb{C}$. Theorem 4.3(1) proves the case that $\alpha \geq 0$, thus we can omit it. Consider $\alpha < 0$ i.e. $\alpha = -|\alpha|$, then by Theorem 3.4(1) we have

$$\begin{aligned} \textstyle\int \alpha f\ d\mu &= \textstyle|\alpha| \int -f\ d\mu \\ &= \textstyle|\alpha| \int (-u)_+\ d\mu - |\alpha| \int (-u)_-\ d\mu + i|\alpha| \int (-v)_+\ d\mu - i|\alpha| \int (-v)_-\ d\mu \\ &= \textstyle|\alpha| \int u_-\ d\mu - |\alpha| \int u_+\ d\mu + i|\alpha| \int v_-\ d\mu - i|\alpha| \int v_+\ d\mu \\ &= \textstyle-|\alpha| \int f\ d\mu \\ &= \textstyle\alpha \int f\ d\mu \end{aligned}$$

If $\alpha = i\lambda,\ \lambda \in \mathbb{R}$. Then

$$\begin{aligned} \textstyle\int \alpha f\ d\mu &= \textstyle\lambda \int if\ d\mu \\ &= \textstyle\lambda \int (iu - v)\ d\mu \\ &= \textstyle\lambda \left(-\int v\ d\mu + i \int u\ d\mu\right) \\ &= \textstyle i\lambda \left(\int u\ d\mu + i \int v\ d\mu\right) \\ &= \textstyle\alpha \int f\ d\mu. \end{aligned}$$

(2) We have that $\int_X f\ d\mu = e^{i\theta} \left|\int_X f\ d\mu\right|$ for some $\theta \in \mathbb{R}$. So

$$\begin{aligned} \left|\int_X f\ d\mu\right| &= e^{-i\theta} \int_X f\ d\mu \\ &= \int_X e^{-i\theta} f\ d\mu && \text{by Theorem 5.1(1)} \\ &\leq \int_X \mathrm{Re}(e^{-i\theta} f)\ d\mu. && (5.1) \end{aligned}$$

If $g \in \mathcal{L}^1(\mu)$ is $\mathbb{R}$-valued, then

$$\begin{aligned}
\left|\int_X g\, d\mu\right| &= \left|\int_X g_+\, d\mu + \int_X g_-\, d\mu\right| \\
&\leq \left|\int_X g_+\, d\mu\right| + \left|\int_X g_-\, d\mu\right| \\
&= \int_X g_+\, d\mu + \int_X g_-\, d\mu \\
&= \int_X |g|\, d\mu.
\end{aligned}$$

So by (5.1) we have that

$$\left|\int_X f\, d\mu\right| \leq \int_X |\operatorname{Re}(e^{-i\theta} f)|\, d\mu \leq \int_X |f|\, d\mu.$$

(3) If $f = 0$ $\mu-$a.e. then $u_\pm = 0$ $\mu-$a.e. and $v_\pm = 0$ $\mu-$a.e. So by Theorem 4.3(3) we obtain

$$\int_E f\, d\mu = 0 \qquad \forall\, E. \in \mathcal{A}.$$

Conversely, let $\int_E f\, d\mu = 0$ for all $E \in \mathcal{A}$. Then

$$\int_E u\, d\mu = 0 = \int_E v\, d\mu \qquad \forall\, E \in \mathcal{A}$$

$$\int_X u_+\, d\mu = \int_{u^{-1}([0,\infty))} u\, d\mu = 0$$

and

$$\int_X u_-\, d\mu = \int_{u^{-1}((-\infty,0])} u\, d\mu = 0.$$

Thus by Theorem 4.3(3) $u_\pm = 0$ μ-a.e. Likewise $v_\pm = 0$ μ-a.e. Thus $f = 0$ μ-a.e.

□

Theorem 5.2 (Lebesgue Dominated Convergence Theorem - DCT)**.** For a measure space $(X, \mathcal{A}, \mu)$, let $\{f_n\} \subset \mathcal{L}^1(\mu)$ be $\mathbb{C}$ or $[-\infty, \infty]-$valued such that

(1) $f(x) = \lim\limits_{n\to\infty} f_n(x)$ exists μ-a.e. and

(2) $|f_n(x)| \leq g(x)$ μ-a.e. for some $g \in \mathcal{L}^1(\mu)$ $\forall\, n$ (i.e. f_n is **dominated** by g).

Then

$$\lim_{n\to\infty} \int_X |f_n - f|\, d\mu = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu.$$

Proof. By Theorem 3.5(5), f is measurable after redefinition on a μ-null set. Since $|f| \leq g \in \mathcal{L}^1(\mu)$ a.e. $\Rightarrow f \in \mathcal{L}^1(\mu)$. Since $|f_n - f| \leq 2g$ a.e. Thus $2g - |f_n - f|$ is positive (after redefinition on μ-null set). So Fatou's Lemma applies i.e.

$$\int_X \liminf_n (2g - |f_n - f|)\, d\mu \leq \liminf_n \int_X (2g - |f_n - f|)\, d\mu$$

i.e.

$$\begin{aligned}\textstyle\int_X 2g\ d\mu &\le \liminf_n \textstyle\int_X (2g - |f_n - f|)\ d\mu \\ &= \textstyle\int_X 2g\ d\mu + \liminf_n \textstyle\int_X -|f_n - f|\ d\mu \\ &= \textstyle\int_X 2g\ d\mu - \limsup_n \textstyle\int_X |f_n - f|\ d\mu.\end{aligned}$$

Since $\int_X 2g\ d\mu < \infty$ this means

$$\limsup_n \int_X |f_n - f|\ d\mu = 0. \tag{5.2}$$

If a positive sequence $\alpha_k \in \mathbb{R}^+$ does **not** converge, then

$$\limsup_n \alpha_n = \lim_{k\to\infty}\left(\sup_{i\ge k}\alpha_i\right) > 0.$$

So (5.2) $\Rightarrow \lim\limits_{n\to\infty} \int_X |f_n - f|\ d\mu = 0$ and

$$0 \le \left|\int_X (f_n - f)\ d\mu\right| \le \int_X |f_n - f|\ d\mu \xrightarrow[\infty]{n} 0.$$

Thus

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

□

Example 5.1.

(1) Let $E_n \in \mathcal{A}$ and $E_1 \supseteq E_2 \supseteq \dots,\ E := \bigcap\limits_{i=1}^{\infty} E_i$. Let $f \in \mathcal{L}^1(\mu)$ and define $f_n = \chi_{E_n}\cdot f$. Then

$$\chi_E \cdot f = \lim_{n\to\infty} f_n \qquad \text{and} \qquad |f_n| \le |f| \in \mathcal{L}^1(\mu) \qquad \forall\, n.$$

So by DCT

$$\lim_{n\to\infty}\int_X f_n\ d\mu = \lim_{n\to\infty}\int_{E_n} f\ d\mu = \int_E f\ d\mu.$$

(2) The "dominated" part is important. Consider the moving hump example $f_n = \chi_{[n,n+1]}$ (see figure 4.1). Then

$$\lim_{n\to\infty} f_n(x) = 0 \qquad \forall\, x \qquad \text{and} \qquad |f_n| \le 1 \qquad (\textbf{not } \mathcal{L}^1(\mu)).$$

Then

$$\lim_{n\to\infty}\int_{\mathbb{R}} f_n\ d\mu = 1 \ne \int_{\mathbb{R}} \lim_{n\to\infty} f_n\ d\mu = 0.$$

Corollary 5.3. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \times [a,b] \to \mathbb{C}$ satisfying $f_t \in \mathcal{L}^1(\mu)\ \ \forall\, t \in [a,b]$ where $f_t(x) := f(x,t)$. Then

(1) If $\exists\, g \in \mathcal{L}^1(\mu)$ such that $|f_t| \leq g$ on $X \quad \forall\, t$, and $\lim_{t \to t_0} f(x,t) = f(x,t_0) \quad \forall\, x$, then

$$F(t) := \int_X f(x,t_0)\ d\mu(x)$$

is continuous at t_0 i.e. $\lim_{t \to t_0} \int_X f(x,t)\ d\mu(x) = \int_X f(x,t_0)\ d\mu(x)$.

(2) If $\frac{\partial f}{\partial t}\ \exists$ and $\exists\, g \in \mathcal{L}^1(\mu)$ such that $|\frac{\partial f}{\partial t}| \leq g \quad \forall\, x \in X,\ t \in (a,b)$, then $F'\ \exists$ and

$$F'(t) = \int_X \frac{\partial f}{\partial t}(x,t)\ d\mu(x) \qquad \text{for } t \in (a,b).$$

Proof.

(1) Apply DCT to $f_n(x) := f(x,t_n)$ where $t_n \to t_0$.

(2) Let

$$h_n(x) := \frac{f(x,t_n) - f(x,t_0)}{t_n - t_0} \qquad \text{where } t_n \to t_0.$$

Then

$$\frac{\partial f}{\partial t}(x,t_0) = \lim_{n \to \infty} h_n(x).$$

So $\frac{\partial f}{\partial t}$ is measurable with respect to x. So by the Mean Value Theorem (MVT) with respect to t we have

$$|h_n(x)| \leq \sup_{t \in (a,b)} \left| \frac{\partial f}{\partial t}(x,t) \right| \leq g(x).$$

So

$$\begin{aligned} F'(t_0) &= \lim_{n \to \infty} \tfrac{F(t_n) - F(t_0)}{t_n - t_0} \\ &= \lim_{n \to \infty} \textstyle\int_X h_n(x)\ d\mu(x) \\ &= \textstyle\int_X \frac{\partial f}{\partial t}(x,t_0)\ d\mu(x) \qquad \text{by DCT.} \end{aligned}$$

□

Theorem 5.4. Let $S = [a_1,b_1) \times \cdots \times [a_n,b_n) \subset \mathbb{R}^n$ be an n-block, and let $f : S \to \mathbb{R}$ be bounded. Then

(1) If f is Riemann integrable on S, then it is Lebesgue integrable on S with respect to the Lebesgue measure μ and

$$\int_S f\ d\mathbf{x} = \int_S f\ d\mu$$

i.e. the Riemann integral is equal to the Lebesgue integral.

(2) If f is Riemann integrable, then it is continuous μ-a.e. on S.

Proof.

(1) Let f be Riemann integrable over S. For each $p \in \mathbb{N}$ let $\mathcal{C}_p = \{C_1, \dots, C_{N_p}\}$ be a partition of S into n-blocks $C_k = [t_1, s_1) \times \cdots \times [t_n, s_n)$ such that $|s_i - t_i| < \frac{1}{p} \;\; \forall i$. Define

$$l_p := \sum_{k=1}^{N_p} \alpha_k \cdot \chi_{C_k} \qquad \text{and} \qquad u_p := \sum_{k=1}^{N_p} \beta_k \cdot \chi_{C_k}$$

where

$$\alpha_k := \inf\{f(x) \mid x \in C_k\} \qquad \text{and} \qquad \beta_k := \sup\{f(x) \mid x \in C_k\}.$$

So the lower Riemann sums is

$$\underline{\mathcal{R}(\mathcal{C}_p)} = \sum_{k=1}^{N_p} \alpha_k \mu(C_k) = \int_S l_p \; d\mu$$

and the upper Riemann sums is

$$\overline{\mathcal{R}(\mathcal{C}_p)} = \sum_{k=1}^{N_p} \beta_k \mu(C_k) = \int_S u_p \; d\mu.$$

Choose a sequence $\mathcal{C}_1, \mathcal{C}_2, \dots$ such that $\mathcal{C}_{p+1}$ is a refinement of $\mathcal{C}_p$, then $u_p - l_p \geq 0$ and

$$u_1 \geq u_2 \geq \cdots \geq f \geq \cdots \geq l_2 \geq l_1.$$

So $\exists$ limits

$$\lim_{p \to \infty} u_p = u \geq f \geq l = \lim_{l \to \infty} l_p.$$

Since f is Riemann integrable we have

$$\begin{aligned}
\int_S f \; d\mathbf{x} = \overline{I(f)} &= \inf \left\{ \overline{R(\mathcal{C}_p)} = \int_S u_p \; d\mu \;\middle|\; \mathcal{C}_p \text{ as above} \right\} \\
&= \lim_{p \to \infty} \int_S u_p \; d\mu \\
&= \underline{I(f)} = \sup \left\{ \underline{R(\mathcal{C}_p)} = \int_S l_p \; d\mu \;\middle|\; \mathcal{C}_p \text{as above} \right\} \\
&= \lim_{p \to \infty} \int_S l_p \; d\mu. \qquad (5.3)
\end{aligned}$$

Since f is bounded on $S \Rightarrow |f| < K$ on S for some $K < \infty$. So $|u_p| \leq K$ and $|l_p| < K$ on δ. So $|u_p| \leq K$ and $|l_p| < K$ on S for p large enough and $2K \cdot \chi_S \in \mathcal{L}^1(\mathbb{R}^n)$. So $|u_p - l_p| < 2K \cdot \chi_S$ on S. So

$$\begin{aligned}
0 &= \lim_{p \to \infty} \textstyle\int_S (u_p - l_p) \; d\mu \quad \text{by (5.3)} \\
&= \textstyle\int_S \lim_{p \to \infty} (u_p - l_p) \; d\mu \quad \text{by DCT} \\
&= \textstyle\int_S (u - l) \; d\mu.
\end{aligned}$$

$\Rightarrow u - l = 0$ μ-a.e. by Theorem 4.3(3) $\Rightarrow u = l = f$ μ-a.e. on S. So from (5.3) and DCT we have

$$\int_S f \; d\mathbf{x} = \int_S \lim_{p \to \infty} l_p \; d\mu = \int_S f \; d\mu. \qquad (5.4)$$

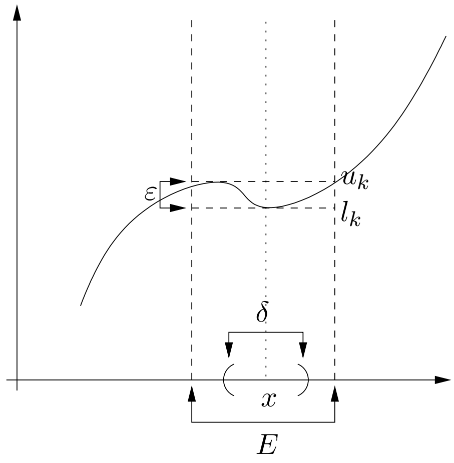

Figure 5.1: E is the interior of the n-block

(2) If f is Riemann integrable then $f = \lim_{p\to\infty} l_p$ μ-a.e. by Equation (5.4). Let B_p = boundaries of all n-blocks in $\mathcal{C}_p$. These are of lower dimension than n, so $|B_p| = 0$. So for the sequence $\mathcal{C}_1, \mathcal{C}_2, \ldots$ defined as above we have that $\mu\left(\bigcup_{p\in\mathcal{N}} B_p\right) = 0$ since $\mu(B_p) = 0$ and σ-additivity. Now $\lim_{p\to\infty} l_p = f = \lim_{p\to\infty} u_p$ μ-a.e. This means $\exists\ N \in \mathcal{A}$ such that $\mu(N) = 0$ and $\lim_{p\to\infty} l_p(x) = f(x) = \lim_{p\to\infty} u_p(x)$ for $x \in N^c$. If $x \notin N \cup \left(\bigcup_{n=1}^{\infty} B_p\right)$ then by the fact that $\lim_{p\to\infty} l_p(x) = f(x) = \lim_{p\to\infty} u_p(x)$ we have that $\forall\, \varepsilon > 0\ \ \exists\ k$ such that $u_k(x) - l_k(x) < \varepsilon$ (by $l_p \leq f \leq u_p$). Let E be the interior of the n-block C_k to which x belongs, then $u_k(y) - l_k(y) < \varepsilon\ \ \forall\, y \in E$ (see Figure 5.1). Thus since E is open and $x \in E$, $\exists\ \delta > 0$ such that $\|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Thus f is continuous at x since $\mu\left(N \cup \left(\bigcup_{p=1}^{\infty} B_p\right)\right) = 0$. The result follows.

$\square$

Remarks 5.1. In $\mathbb{R}$ we also have the converse of Theorem 5.4(2).

Example 5.2. Let $f = \chi_{\mathbb{Q}\cap[0,1]}$ then f is **not** Riemann integrable. However $f \restriction [0,1] \setminus \mathbb{Q} = 0$ = constant hence continuous on $[0,1] \setminus \mathbb{Q}$ but $\mu(\mathbb{Q}) = 0$. We cannot examine continuity within a restricted set! f is **not** continuous on points $[0,1] \setminus \mathbb{Q}$ inside $[0,1]$.

Chapter 6

Product Measures and the Fubini Theorem

Definition 6.1. Let $(X, \mathcal{A}),\ (Y, \mathcal{B})$ be σ-algebras. Define the **product σ-algebra** by

$$\mathcal{A} \otimes \mathcal{B} := \sigma\text{-algebra generated by } \{A \times B \mid A \in \mathcal{A},\ B \in \mathcal{B}\}.$$

We call a set $A \times B \subseteq X \times Y$ with $A \in \mathcal{A},\ B \in \mathcal{B}$ is a **(measurable) rectangle**.

Example 6.1. Note that $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Now $\mathcal{B}(\mathbb{R}^2) \subseteq \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ as $\mathcal{B}(\mathbb{R}^2)$ is generated by 2-blocks $[s_1, t_1) \times [s_2, t_2)$ and $\mathcal{B}(\mathbb{R})$ is generated by 1-blocks $[s, t)$. Conversely, let $P_i : \mathbb{R}^2 \to \mathbb{R}$ be projections $P_1(x, y) = x$ and $P_2(x, y) = y$. These are continuous and hence Borel. So if $A, B \in \mathcal{B}(\mathbb{R})$ then

$$A \times B = (A \times \mathbb{R}) \cap (\mathbb{R} \times B) = P_1^{-1}(A) \cap P_2^{-1}(B) \in \mathcal{B}(\mathbb{R}^2).$$

Thus we obtain reverse inclusion (see Figure 6.1). So $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$.

Definition 6.2. For a map $f : X \times Y \to Z$ define the **sections** f_x, f^y on Y, X respectively by

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

Theorem 6.1. Let $(X, \mathcal{A}),\ (Y, \mathcal{B})$ be σ-algebras. Then

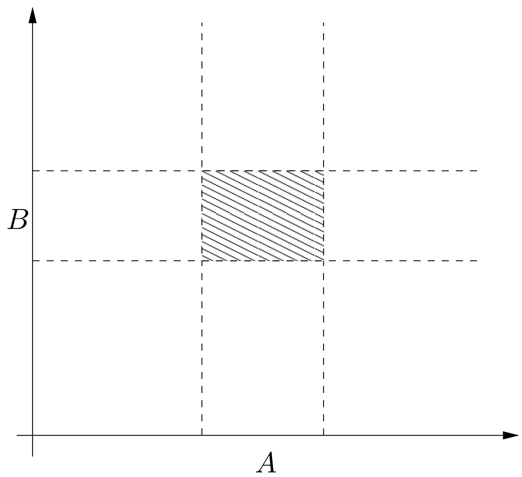

Figure 6.1: The intersection of A and B

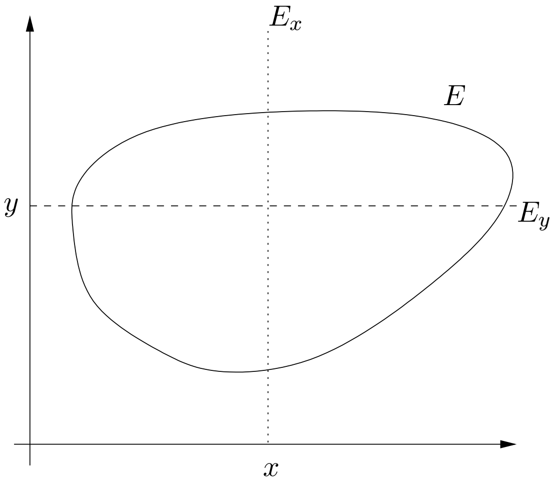

Figure 6.2: The intersection of E_x and E_y

(1) If $E \in \mathcal{A} \otimes \mathcal{B}$ then

$$E_x := \{y \in Y \mid (x, y) \in E\} \in \mathcal{B} \qquad \forall\, x \in X$$

and

$$E_y := \{x \in X \mid (x, y) \in E\} \in \mathcal{A} \qquad \forall\, y \in Y.$$

See Figure 6.2. Note that

$$(A \times B)_x = \begin{cases} B & \text{if } x \in A \\ \emptyset & \text{if } x \notin A \end{cases}$$

likewise for y.

(2) If $f : X \times Y \to \mathbb{C}$ or $[-\infty, \infty]$ is $\mathcal{A} \otimes \mathcal{B}$-measurable, then f_x is $\mathcal{B}$-measurable and f^y is $\mathcal{A}$-measurable.

Proof.

(1) Let $x \in X$ and

$$\mathcal{F} := \{E \subseteq X \times Y \mid E_x \in \mathcal{B}\}.$$

Then all rectangles $A \times B \in \mathcal{F}$ for $A \in \mathcal{A},\ B \in \mathcal{B}$. By $(E^c)_x = (E_x^c)$ and $\left(\bigcup_{n=1}^{\infty} E_n\right)_x = \bigcup_{n=1}^{\infty} (E_n)_x$ we have that $\mathcal{F}$ is closed with respect to taking complements and countable unions. Thus $\mathcal{F}$ is a σ-algebra containing $\mathcal{A} \otimes \mathcal{B}$ as $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{F}$. Thus by definition of $\mathcal{F}$, $E_x \in \mathcal{B}\ \ \forall\, x$ when $E \in \mathcal{A} \otimes \mathcal{B}$. Likewise for $E_y \in \mathcal{A}$ when $E \in \mathcal{A} \otimes \mathcal{B}$.

(2) Since

$$(f_x)^{-1}(c) = (f^{-1}(c))_x \qquad \text{and} \qquad (f^y)^{-1}(c) = (f^{-1}(c))^y$$

this follows from Theorem 6.1(1).

$\square$

Theorem 6.2. Let $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. If $E \in \mathcal{A} \otimes \mathcal{B}$, define $\varphi^E : X \to [0, \infty]$, $\psi^E : Y \to [0, \infty]$ by

$$\varphi^E(x) := \nu(E_x),\ x \in X \qquad \text{and} \qquad \psi^E(y) := \mu(E_y),\ y \in Y.$$

Then φ^E is $\mathcal{A}$-measurable and ψ^E is $\mathcal{B}$-measurable.

Proof. Assume first that ν is finite and observe that $\varphi^E(x) = \nu(E_x)$ is defined by Theorem 6.1 Define

$$\mathcal{F} := \{E \in \mathcal{A} \otimes \mathcal{B} \mid \varphi^E \text{ is } \mathcal{A}\text{-measurable}\}.$$

This contains the rectangle $\mathcal{R} := \{A \times B \mid A \in \mathcal{A},\ B \in \mathcal{B}\}$ because $\nu((A \times B)_x) = \nu(B) \cdot \chi_A(x)$ which is $\mathcal{A}$-measurable. Now $\mathcal{F}$ is a Dynkin Class because

- $X \times Y \in \mathcal{R} \subset \mathcal{F}$;
- If $E, F \in \mathcal{F}$ with $E \subset F$ then
$$\nu((F \setminus E)_x) = \nu(F_x \setminus E_x) = \nu(F_x) - \nu(E_x) \Rightarrow F \setminus E \in \mathcal{F}.$$
- If $\{E_n\}_{n=1}^{\infty} \subset \mathcal{F}$ is an increasing sequence then so is $\{(E_n)_x\}_{n=1}^{\infty} \subset \mathcal{B}$. Thus
$$\nu\left(\left(\bigcup_{n=1}^{\infty} E_n\right)\right) = \lim_{n\to\infty} \nu((E_n)_x) = \lim_{n\to\infty} \varphi^{E_n}(x)$$
is $\mathcal{A}$-measurable since it is the limit of measurable functions. Thus $\bigcup\limits_{n=1}^{\infty} E_n \in \mathcal{F}$.

Now $\mathcal{R}$ is closed with respect to finite intersections. So by Lemma 2.2 in proof of Theorem 2.6 $\mathcal{A} \otimes \mathcal{B} = \sigma$-algebra generated by $\mathcal{R}$ = Dynkin Class generated by $\mathcal{R} \subseteq \mathcal{F} \subseteq \mathcal{A} \otimes \mathcal{B}$. Hence $\mathcal{A} \otimes \mathcal{B} = \mathcal{F}$, so φ^E is measurable $\forall\, E \in \mathcal{A} \otimes \mathcal{B}$. Likewise ψ^E is measurable for all $E \in \mathcal{A} \otimes \mathcal{B}$.

Now let ν be σ-finite. So $Y = \bigcup\limits_{n=1}^{\infty} D_n$ with $\nu(D_n) < \infty$ and $D_n \cap D_m = \emptyset$ for all $n \neq m$. For each n define a measure $\nu_n : \mathcal{B} \to [0, \infty]$ by

$$\nu_n(B) := \nu(B \cap D_n) \leq \nu(D_n) \leq \infty$$

which is a finite measure since $\nu(D_n) < \infty$. So by the above, the functions $x \mapsto \nu_n(E_x) = \varphi_n^E(x)$ are $\mathcal{A}$-measurable so

$$\nu(E_x) = \sum_{n=1}^{\infty} \nu(E_x \cap D_n) = \sum_{n=1}^{\infty} \varphi_n^E(x)$$

is $\mathcal{A}$-measurable.

$\square$

Theorem 6.3. Let $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then there is a unique measure $\mu \times \nu : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ such that

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B) \qquad \forall\, A \in \mathcal{A},\ B \in \mathcal{B}.$$

For any $E \in \mathcal{A} \otimes \mathcal{B}$:

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

We say $\mu \times \nu \equiv$ **product measure** of μ and ν.

Proof. Theorem 6.2 $\Rightarrow$ the maps $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable. So define

$$(\mu \times \nu)_1(E) := \int_Y \mu(E^y)\, d\nu(y) \qquad \text{and} \qquad (\mu \times \nu)_2(E) := \int_X \mu(E_x)\, d\mu(x).$$

Then

$$\begin{aligned}(\mu \times \nu)_1(A \times B) &= \int_Y \mu(A) \cdot \chi_B(y)\, d\nu(y) \\ &= \mu(A)\nu(B) \\ &= (\mu \times \nu)_2(A \times B) \qquad (6.1)\end{aligned}$$

First we have that

$$(\mu \times \nu)_1(\emptyset) = 0 = (\mu \times \nu)_2(\emptyset).$$

Let $E = \bigcup_{n=1}^{\infty} E_n$ with disjoint $E_n \in \mathcal{A} \otimes \mathcal{B}$, then $\{E_n^y\} \in \mathcal{A}$ is disjoint and $E^y = \bigcup_{n=1}^{\infty} E_n^y$. So we have

$$\mu(E^y) = \sum_{n=1}^{\infty} \mu(E_n^y). \qquad (6.2)$$

Using this fact we obtain

$$\begin{aligned}(\mu \times \nu)_1(E) &= \textstyle\int_Y \mu(E^y)\, d\nu(y) \\ &= \textstyle\int_Y \sum_{n=1}^{\infty} \mu(E_n^y)\, d\nu(y) \quad \text{using Equation (6.2)} \\ &= \textstyle\sum_{n=1}^{\infty} \int_Y \mu(E_n^y)\, d\nu(y) \quad \text{by MCT} \\ &= \textstyle\sum_{n=1}^{\infty} (\mu \times \nu)_1(E_n).\end{aligned}$$

Thus $(\mu \times \nu)$ is a measure and likewise we have that $(\mu \times \nu)_2$ is a measure. Note that they are both σ-finite: If

$$X = \bigcup_{k=1}^{\infty} A_k, \quad \mu(A_k) < \infty \qquad \text{and} \qquad Y = \bigcup_{j=1}^{\infty} B_j, \quad \mu(B_j) < \infty.$$

Then

$$X \times Y = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} A_k \times B_j \qquad \text{and} \qquad (\mu \times \nu)(A_k \times B_j) < \infty.$$

As they agree on the rectangle $\mathcal{R}$ by (6.1) it follows from Corollary 2.7 that they are equal and uniquely specified.

□

Example 6.2. We have $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. If μ_n is the Lebesgue measure on $\mathbb{R}^n$, then

$$\mu_2([a,b) \times [c,d)) = (b-a)(d-c) = (\mu_1 \times \mu_1)([a,b) \times [c,d))$$

hence $\mu_2 = \mu_1 \times \mu_1$.

We now want to evaluate integrals with respect to iterated integrals, the following Theorem will allow us to do just that:

Theorem 6.4 (Tonelli-Fubini)**.** Let $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces.

(1) (Tonelli) If $f : X \times Y \to [0, \infty]$ is a $\mathcal{A} \otimes \mathcal{B}$-measurable function, then the functions g, h are measurable where

$$g(x) := \int_Y f_x \, d\nu \quad \forall\, x \in X \qquad \text{and} \qquad h(y) := \int_X f^y \, d\mu \quad \forall\, y \in Y.$$

Moreover

$$\begin{aligned} \int_{X\times Y} f \, d(\mu \times \nu) &= \int_X \left[\int_Y f(x,y) \, d\nu(y) \right] \, d\mu(x) \\ &= \int_Y \left[\int_X f(x,y) \, d\mu(x) \right] \, d\nu(y) \end{aligned} \tag{6.3}$$

(2) (Fubini) If $f \in \mathcal{L}^1(\mu \times \nu)$ then $f_x \in \mathcal{L}^1(\nu)$ μ-a.e. in x, $f^y \in \mathcal{L}^1(\mu)$ ν-a.e. in y and if $g(x) = \int_Y f_x \, d\nu$ then $g \in \mathcal{L}^1(\mu)$. If $h(y) = \int_X f^y \, d\mu$ then $h \in \mathcal{L}^1(\nu)$ and Equation (6.3) holds for the given function f.

Proof. When f is a characteristic function χ_E, then Theorem 6.4(1) holds for non-negative measurable simple functions by linearity of integral. For f measurable on $X \times Y$, let s_n be simple functions as in Theorem 3.3 such that

(1) $0 \le s_1 \le s_2 \le \ldots$

(2) $\lim\limits_{n\to\infty} s_n(x,y) = f(x,y) \qquad \forall\, x, y.$

Let

$$g_n(x) := \int_Y (s_n)_x(y) \, d\nu(y) \qquad \text{and} \qquad h_n(y) := \int_X (s_n)^y(x) \, d\mu(x).$$

Then by MCT, g_n increases to limit g and h_n increases to limit h. So g, h are measurable and

$$\begin{aligned} \textstyle\int_X g \, d\mu &= \lim_{n\to\infty} \textstyle\int_X g_n \, d\mu \\ &= \lim_{n\to\infty} \textstyle\int_X \left(\int_Y (s_n)_x \, d\nu \right) \, d\mu \\ &= \lim_{n\to\infty} \textstyle\int_{X\times Y} s_n \, d(\mu \times \nu) \\ &= \textstyle\int_{X\times Y} f \, d(\mu \times \nu) \end{aligned}$$

and

$$\begin{aligned} \textstyle\int_Y h \, d\nu &= \lim_{n\to\infty} \textstyle\int_Y h_n \, d\nu \\ &= \lim_{n\to\infty} \textstyle\int_Y \left(\int_X (s_n)^y \right) \, d\nu \\ &= \lim_{n\to\infty} \textstyle\int_{X\times Y} s_n \, d(\mu \times \nu) \\ &= \textstyle\int_{X\times Y} f \, d(\mu \times \nu). \end{aligned}$$

This establishes Theorem 6.4(1) and Tonelli's Theorem. It also shows that if $f \in \mathcal{L}^1(\mu \times \nu)$ and f non-negative, then

$$\int_{X\times Y} f \, d(\mu \times \nu) < \infty \Rightarrow g \in \mathcal{L}^1(\mu) \text{ and } h \in \mathcal{L}^1(\nu)$$

hence $g < \infty$ μ-a.e. and $h < \infty$ ν-a.e. i.e. $f_x \in \mathcal{L}^1(\nu)$ μ-a.e. in x and $f^y \in \mathcal{L}^1(\mu)$ ν-a.e. in y. Apply these results to the non-negative functions in the decomposition $f = u_+ - u_- + i(v_+ - v_-)$ to obtain Fubini's Theorem by linearity of integral.

This generalises to any finite product of σ-finite measure spaces.

Example 6.3. Let $\sum\limits_{n=1}^{\infty}\sum\limits_{m=1}^{\infty} a_{n,m}$ where

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} |a_{n,m}| < \infty. \tag{6.4}$$

Then

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m}.$$

Proof. Let $(X, \mathcal{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \delta) = (Y, \mathcal{B}, \nu)$ where δ = counting measure. Then

$$\int_{\mathbb{N}} f(n)\ d\mu(n) = \sum_{n=1}^{\infty} f(n).$$

Then

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \int_{\mathbb{N}}\int_{\mathbb{N}} a_{n,m}\ d_\mu(n)\ d\nu(m)$$

since Equation (6.4) $\Rightarrow a_{n,m}$ is $\mathcal{L}^1(\mu \times \nu)$ it follows from Fubini that

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m}.$$

□

Before we turn our attention to countable product measures, we revisit some definitions that were covered in MATH3611 - Higher Analysis that will be of use to us:

Definition 6.3. Let $\{X_\alpha \mid \alpha \in A\}$ be an index set of sets, i.e. A is a set and X_α is a set for each $\alpha \in A$.

(1) A map $g : A \to \dot{\bigcup\limits_{\alpha \in A}} X_\alpha$ and $g(\alpha) \in X_\alpha\ \ \forall\, \alpha \in A$ is called a **choice function**.

(2) The **Cartesian product** is

$$\prod_{\alpha \in A} X_\alpha := \left\{ g(A) \subset \dot{\bigcup_{\alpha \in A}} X_\alpha \;\middle|\; g : A \to \dot{\bigcup_{\alpha \in A}} X_\alpha \text{ is a choice function} \right\}.$$

Its elements are written $g(A) =: \prod\limits_{\alpha \in A} g(\alpha)$.

(3) The maps $p_{\alpha_0} : \prod\limits_{\alpha \in A} X_\alpha \to X_{\alpha_0}$ by $p_{\alpha_0}(g(A)) = g(\alpha_0) = g(A) \cap X_{\alpha_0}$ are the **coordinate projections**. Note that

$$p_\alpha^{-1}(U) = \{g \mid g(\alpha) \in U\} = \prod_{\beta \in A} U_\beta$$

where $U_\beta = X_\beta$ if $\beta \neq \alpha$ and $U_\alpha = U \subseteq X_\alpha$.

(4) Let (X_α, τ_α) be topological spaces. Then the **product topology** of $X := \prod_{\alpha\in A} X_\alpha$ is the weakest topology which makes all the p_α continuous and it is generated by

$$\mathcal{B} := \left\{ \bigcap_{i=1}^{n} p_{\alpha_i}^{-1}(U_{\alpha_i}) \;\middle|\; U_{\alpha_i} \in \tau_{\alpha_i},\ n \in \mathbb{N} \right\}.$$

Note $\bigcap_{i=1}^{n} p_{\alpha_i}^{-1}(U_{\alpha_i}) = \prod_{\alpha\in A} V_\alpha$ with $V_\alpha = X_\alpha$ if $\alpha \neq \alpha_i$ and $V_{\alpha_i} = U_{\alpha_i}\ \ \forall\, i$.

First note that for $a_n \in \mathbb{R}$, $\sum_{n=1}^{\infty} a_n < \infty \Rightarrow a_n \to 0$ so $\prod_{n=1}^{\infty} a_n < \infty$ and $\neq 0 \Rightarrow a_n \to 1$. Thus in our study of countable and infinite product measures we inevitably encounter probability measures.

Definition 6.4. Let $(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha)$, $\alpha \in A$ be probability spaces ($\mu_\alpha(X_\alpha) = 1$). A **cylinder with base** $C \in \mathcal{A}_{\alpha_1} \otimes \cdots \otimes \mathcal{A}_{\alpha_n}$ is the set

$$C \times \prod_{\alpha\in A,\ \alpha\neq\alpha_i} X_\alpha = \{g \in X \mid (g(\alpha_1), g(\alpha_2), \ldots, g(\alpha_n)) \in C\}.$$

Let $\bigotimes_{\alpha\in A} \mathcal{A}_\alpha$ be the σ-algebra generated by all cylinders, i.e. it is generated by

$$\mathcal{C} = \left\{ C \times \prod_{\alpha\in A,\ \alpha\notin A'} X_\alpha \;\middle|\; \alpha_i \in A,\ n \in \mathbb{N},\ C \in \mathcal{A}_{\alpha_1} \otimes \cdots \otimes \mathcal{A}_{\alpha_n} \right\}$$

where $A' = \{\alpha_1, \ldots, \alpha_n\}$. Note that $\mathcal{C}$ is closed with respect to complementation, **finite** unions and it contains X. If $\mathcal{A}_\alpha$ are Borel σ-algebras, $\mathcal{C}$ is generated by $\mathcal{B}$ and so is a Borel σ-algebra for the product topology.

First define the product measure for a countable index set A. Let $A = \mathbb{N}$, then $\mathcal{C}$ is a union of

$$\xi_n := \left\{ C \times X_{n+1} \times X_{n+2} \times \ldots \;\middle|\; C \in \bigotimes_{i=1}^{n} \mathcal{A}_i \right\} \qquad (\xi_n \subset \xi_{n+1})$$

then ξ_n is a σ-algebra and $\mathcal{C} = \bigcup_{n=1}^{\infty} \xi_n$. Define $\mu : \mathcal{C} \to [0, \infty)$ by

$$\mu(C \times X_{n+1} \times X_{n+2} \times \ldots) := (\mu_1 \times \cdots \times \mu_n)(C).$$

This is a well defined map. For example using the fact that $\xi_n \subset \xi_{n+1}$ gives $\mu(C \times X_{n+1} \times \ldots) = (\mu_1 \times \cdots \times \mu_n)(C)$ since

$$\begin{aligned} \mu((C \times X_{n+1}) \times X_{n+2} \times \ldots) &= (\mu_1 \times \cdots \times \mu_n \times \mu_{n+1})(C \times X_{n+1}) \\ &= (\mu_1 \times \cdots \times \mu_n)(C)\, \underbrace{\mu_{n+1}(X_{n+1})}_{=1}. \end{aligned}$$

Theorem 6.5. Given probability measures $(X_n, \mathcal{A}_n, \mu_n)$, $n \in \mathbb{N}$, define $\bigotimes_{n=1}^{\infty} \mathcal{A}_n$, $\mathcal{C}$, $\mu : \mathcal{C} \to [0, \infty)$ as above. Then there is a unique probability measure $\mu_\infty : \bigotimes_{n=1}^{\infty} \mathcal{A}_n \to [0, \infty)$ which coincides with μ on $\mathcal{C}$.

Proof. As $\mathcal{C}$ is closed with respect to finite intersections and $\bigotimes_{n=1}^{\infty} \mathcal{A}_n = \mathcal{A}(\mathcal{C})$ it follows from Theorem 2.6 that a probability measure on $\bigotimes_{n=1}^{\infty} \mathcal{A}_n$ is uniquely determined by its values on $\mathcal{C}$. So we only have to prove existence. (Note that as $X = \prod_{n=1}^{\infty} X_n \in \xi_k \subset \mathcal{C} \;\; \forall\, k \Rightarrow \mu(X) = \mu_1(X_1) = 1$. For example $k = 1$).

Lemma 6.5.1. $\mu : \mathcal{C} \to [0, \infty)$ is countably addivitive **on C** i.e. if $\{B_1, B_2, \dots\} \subset \mathcal{C}$ are disjoint and $\bigcup_{k=1}^{\infty} B_k \in \mathcal{C}$, then

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

Proof. We first prove that μ is continuous from above on $\mathcal{C}$, i.e. if $B_1, B_2, \dots \in \mathcal{C}$ such that $B_1 \supseteq B_2 \supseteq \dots$ and $\bigcap_{n=1}^{\infty} B_n = \emptyset$, then $\lim_{n\to\infty} \mu(B_n) = 0$.

Assume it is not true, i.e. $\exists$ decreasing sequences $\{B_k\}_{k=1}^{\infty} \subset \mathcal{C}$ with empty intersection such that $\lim_{n\to\infty} \mu(B_n) \neq 0$. Then $\exists\ \varepsilon > 0$ and an infinite subsequence $\{B_{k_n}\}_{n=1}^{\infty}$ such that $\mu(B_{k_n}) > \varepsilon \;\; \forall\, n$. If $B_{k_n} \in \xi_l$ then as μ is a measure on ξ_l we get $\mu(B_{k_n}) \to 0$, so this is not possible. So we may assume that $B_{k_n} \in \xi_{l_n}$ where $l_n > l_m$ if $n > m$ and without loss of generality we may take $B_{k_n} \in \xi_n$. So

$$B_{k_n} = C_n \times \prod_{i=n+1}^{\infty} X_i \qquad \text{with } C_n \in \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_n.$$

Now

$$\begin{aligned} \mu(B_{k_n}) &= (\mu_1 \times \dots \times \mu_n)(C_n) \\ &= \int_{X_1} \dots \int_{X_n} \chi_{C_n}(x_1, \dots, x_n)\, d\mu_1(x_1) \dots d\mu_n(x_n) \\ &= \int_{X_1} g_n(x_1)\, d\mu_1(x_1) \end{aligned}$$

where

$$g_n(x_1) := \int_{X_2} \dots \int_{X_n} \chi_{C_n}(x_1, \dots, x_n)\, d\mu_2(x_2) \dots d\mu_n(x_n).$$

As $B_{k_{n+1}} \subset B_{k_n}$ we have that $C_{n+1} \subset C_n \times X_{n+1}$. So

$$\chi_{C_{n+1}}(x_1, \dots, x_{n+1}) \leq \chi_{C_n}(x_1, \dots, x_n). \tag{6.5}$$

Thus $g_n : X_1 \to [0, \infty)$ is a decreasing sequence, so $\exists \lim_{n\to\infty} h_n =: h_1$. As μ_1 is a probability measure $g_n \leq 1$ where 1 is the constant function and so $g_n \in \mathcal{L}^1(\mu_1)$. So by

$$\begin{aligned} \lim_{n\to\infty} \mu(B_{k_n}) &= \lim_{n\to\infty} \int_{X_1} g_n(x_1)\, d\mu(x_1) \\ &= \int_{X_1} \lim_{n\to\infty} g_n(x_1)\, d\mu(x_1) && \text{by DCT} \\ &= \int_X h_1(x_1)\, d\mu(x_1) \\ &> \varepsilon \end{aligned} \tag{6.6}$$

$\Rightarrow \quad \exists\, x_1' \in X_1$ such that $h_1(x_1') > 0$ by Theorem 4.3. Then $x_1' \in C_1$ or else

$$\chi_{C_n}(x_1', x_2, \ldots, x_n) = 0 \;\; \forall\, n$$

by inequality (6.5) and $g_n(x_1') = 0$ i.e. $h_1(x_1') = 0$ which is not allowed by inequality (6.6). Now

$$g_n(x_1') = \int_{X_2} g_n^{(2)}(x_2)\, d\mu_{x_2} \qquad \text{for } n > 2$$

where

$$g_n^{(2)}(x_2) = \int_{X_3} \cdots \int_{X_n} X_{C_n}(x_1', x_2, \ldots, x_n)$$

and as above $g_n^{(2)} \searrow h_2$ (i.e. approaching in a decreasing fashion) as $n \to \infty$. So

$$g_n(x_1') \xrightarrow[\infty]{n} \int_{X_2} h_2\, d\mu_2 = h_1(x_1') > 0$$

$\Rightarrow \quad \exists\, x_2' \in X_2$ such that $h_2(x_2') > 0 \Rightarrow x_2' \in C_2$. Inductively we obtain $x_i' \in C_i' \;\; \forall\, i$ hence

$$(x_1', x_2', \ldots) \in \bigcap_{n=1}^{\infty} C_n \times \prod_{i=n+1}^{\infty} X_i = \bigcap_{n=1}^{\infty} B_{k_n} = \emptyset.$$

Contradiction. Thus μ is continuous from above on $\mathcal{C}$. Let $S_n \in \mathcal{C}$ be pairwise disjoint such that $S = \bigcup\limits_{n=1}^{\infty} S_n \in \mathcal{C}$. Let $A_n := S \setminus \bigcup\limits_{k=1}^{n} S_k$ then $A_{n+1} \subseteq A_n$ and $\bigcap\limits_{n=1}^{\infty} A_n = \emptyset$. Thus $\mu(A_k) \xrightarrow[\infty]{n} 0$ by the last part of the proof. Thus

$$\mu(A_n) = \mu(S) - \sum_{k=1}^{n} \mu(S_k) \xrightarrow[\infty]{n} 0$$

by finite additivity. Thus $\mu(S) = \lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \mu(S_k)$.

$\triangledown$

As $\mathcal{C} \ni X$, we can define an outer measure $\mu^* : \mathcal{P}(X) \to [0, \infty)$ via Theorem 2.3

$$\mu^*(A) := \inf \left\{ \sum_{j=1}^{\infty} \mu(B_j) \;\middle|\; B_j \in \mathcal{C} \text{ and } A \subseteq \bigcup_{j=1}^{\infty} B_j \right\}.$$

Lemma 6.5.2. Let $\mathcal{C}$, μ and μ^* be as above. Then

(1) $\mathcal{C} \subset \mathcal{M}_{\mu^*}$.

(2) $\mu(B) = \mu^*(B) \;\; \forall\, B \in \mathcal{C}$.

Proof.

(1) We need to show

$$\mu^*(S) \geq \mu^*(S \cap B) + \mu^*(S \cap B^c) \qquad \forall\, B \in \mathcal{C} \text{ and } S \subseteq X.$$

Let $A_n \in \mathcal{C}$ be a sequence covering the given S. Then

$$\begin{aligned} & \mu(A_n) = \mu(A_n \cap B) + \mu(A_n \cap B^c) \qquad (B \in \mathcal{C}) \\ \Rightarrow\ & \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu(A_n \cap B) + \sum_{n=1}^{\infty} \mu(A_n \cap B^c). \end{aligned}$$

The sequence $\{A_n \cap B\}_{n=1}^{\infty} \subset \mathcal{C}$ covers $S \cap B$ and $\{A_n \cap B^c\}_{n=1}^{\infty} \subset \mathcal{C}$ covers $S \cap B^c$.

$$\begin{aligned} \Rightarrow\ & \sum_{n=1}^{\infty} \mu(A_n) \geq \mu^*(S \cap B) + \mu^*(S \cap B^c) \\ \Rightarrow\ & \mu^*(S) \geq \mu^*(S \cap B) + \mu^*(S \cap B^c) \\ \Rightarrow\ & \mathcal{C} \subset \mathcal{M}_{\mu^*}. \end{aligned}$$

(2) By definition $\mu^*(B) \leq \mu(B)\ \ \forall\, B \in \mathcal{C}$ (using self-covering of B which is inf of $\mu^*(B)$). Let $B \subset \bigcup_{n=1}^{\infty} B_n$ with B and $B_n \in \mathcal{C}$. Then $B = \bigcup_{n=1}^{\infty} (B \cap B_n)$ so

$$\mu(B) \leq \sum_{n=1}^{\infty} \mu(B \cap B_n) \leq \sum_{n=1}^{\infty} \mu(B_n).$$

$\Rightarrow \mu(B) \leq \mu^*(B)$. Thus $\mu(B) = \mu^*(B)$.

$\triangledown$

As μ^* is a measure on $\mathcal{M}_{\mu^*} \supset \mathcal{C}$, it defines a measure on $\bigotimes_{n=1}^{\infty} \mathcal{A}_n = \mathcal{A}(\mathcal{C}) \subseteq \mathcal{M}_{\mu^*}$ which coincides with μ on $\mathcal{C}$. So by uniqueness part of the Theorem we have an extension of μ to $\bigotimes_{n=1}^{\infty} \mathcal{A}_n$ as a measure.

$\square$

Theorem 6.6. Let $(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha),\ \alpha \in A$ be probability spaces. Then

(1) $\mathcal{A} := \bigotimes_{\alpha \in A} \mathcal{A}_\alpha = \bigcup \left\{ \bigotimes_{\alpha \in B} \mathcal{A}_\alpha \,\middle|\, B \subseteq A \text{ and } B \text{ countable} \right\}$;

(2) There is a unique probability measure $\mu_\infty : \bigotimes_{\alpha \in A} \mathcal{A}_\alpha \to [0, \infty)$ which coincides with $\mu : \mathcal{C} \to [0, \infty)$ where

$$\mathcal{C} = \left\{ C \times \prod_{\alpha \in A,\ \alpha \notin A'} X_\alpha \,\middle|\, \alpha_i \in A,\ n \in \mathbb{N},\ C \in \mathcal{A}_{\alpha_1} \otimes \cdots \otimes \mathcal{A}_{\alpha_n} \right\}$$

with $A' = \{\alpha_1, \ldots, \alpha_n\}$ and

$$\mu\left(C \times \prod_{\alpha \in A,\ \alpha \notin A'} X_\alpha \right) = (\mu_{\alpha_1} \times \cdots \times \mu_{\alpha_n})(C).$$

Proof.

(1) By definition
$$\bigcup\left\{\bigotimes_{\alpha\in B}\mathcal{A}_\alpha \;\middle|\; B\subseteq A \text{ and } B \text{ countable}\right\}\subseteq \mathcal{A}.$$
So as it contains $\mathcal{C}$ and $\mathcal{A}=\mathcal{A}(\mathcal{C})$, it suffices to show that it is a σ-algebra. Closure with respect to taking complements is clear as $E\in\bigotimes_{\alpha\in B}\mathcal{A}_\alpha \Rightarrow E^c\in\bigotimes_{\alpha\in B}\mathcal{A}_\alpha$. Let
$$E_n\in\bigcup\left\{\bigotimes_{\alpha\in B}\mathcal{A}_\alpha \;\middle|\; B\subseteq A \text{ and } B \text{ countable}\right\}$$
then $E_n\in\bigotimes_{\alpha\in B_n}\mathcal{A}_n$ with $B_n\subseteq A$ and B_n is countable. Then
$$\{E_n\}_{n=1}^\infty\in\bigcup_{n=1}^\infty\bigotimes_{\alpha\in B_n}\mathcal{A}_\alpha\subseteq\bigotimes_{\alpha\in B}\mathcal{A}_\alpha$$
where $B=\bigcup_{n=1}^\infty B_n$. As B is countable, $\bigcup_{n=1}^\infty E_n$ is in
$$\bigcup\left\{\bigotimes_{\alpha\in B}\mathcal{A}_\alpha \;\middle|\; B\subseteq A \text{ and } B \text{ countable}\right\}$$

(2) By Theorem 2.5 we have that for each $\bigotimes_{\alpha\in B}\mathcal{A}_\alpha$ and B countable a measure μ_B : $\bigotimes_{\alpha\in B}\mathcal{A}_\alpha\to[0,\infty)$ which coincides with μ on $\mathcal{C}\cap\bigotimes_{\alpha\in B}\mathcal{A}_\alpha$. If $B_1\cap B_2\neq\emptyset$ then μ_{B_1} coincides with μ_{B_2} on $\bigotimes_{\alpha\in B_1\cap B_2}\mathcal{A}_\alpha$ (as they coincide on the cylinders). Thus we get a well-defined map
$$\mu:\underbrace{\bigcup\left\{\bigotimes_{\alpha\in B}\mathcal{A}_\alpha \;\middle|\; B\subseteq A \text{ and } B \text{ countable}\right\}}_{\mathcal{A}}\to[0,\infty)$$
which is σ-additive. So it is a measure and it is unqiue as it is unique on each $\bigotimes_{\alpha\in B}\mathcal{A}_\alpha$ with $B\subseteq\mathcal{A}$ countable.

□

Remarks 6.1.

(1) We can allow **finitely** many measures in the collection to be σ-finite, but the rest must be probability measures.

(2) If we partition A into **finitely many** pieces $A_1,\dots,A_n$ and construct product measures $\mu^{(i)}$ for each A_i, then we have Fubini's Theorem for measures $\mu^{(1)},\dots,\mu^{(n)}$ on $\prod_{\alpha\in A_1}X_\alpha,\dots,\prod_{\alpha\in A_n}X_\alpha$. A Fubini Theorem for an infinite sequence is problematic.

Chapter 7

Riesz Representation Theorem

Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Note that the integral defines a map $\varphi : \mathcal{L}^1(\mu) \to \mathbb{C}$ which is linear (i.e. it is a functional) and it is positive i.e. $\varphi(f) \geq 0$ if $f \geq 0$. For a wide class of topological spaces we will show that all positive functionals on the space of continuous functions of compact support correspond to integrals of Borel measures.

Definition 7.1. A topological space (X, τ) is **locally compact** if for each $x \in X$ there is an open neighbourhood $U \ni x$ such that $\overline{U}$ is compact. X is **Hausdorff** if any two distinct points $x, y \in X$ with $x \neq y$ have disjoint open neighbourhoods. We shorten the term "locally compact Hausdorff" to LCH.

Example 7.1.

(1) $\mathbb{R}^n$ is LCH, as is any discrete topology or any compact Hausdorff space. For example $[0,1]^{\mathbb{N}}$.

(2) l^2 (or any infinite-dimensional Hilbert space) is Hausdorff but **NOT** locally compact.

Lemma 7.1. Let X be a LCH space, then

(1) If $K, L \subset X$ are compact and disjoint, then there are disjoint open sets U, V such that $K \subset U$ and $L \subset V$ with $U \cap V = \emptyset$.

(2) Let $K \subset U$ with K compact, U open. Then there is an open set V with $\overline{V}$ compact such that $K \subset V \subseteq \overline{V} \subset U$.

Proof.

(1) Let K, L be non-empty (else it is trivally true). Let $x \in K$ then $\forall\, y \in L\ \exists$ disjoint open sets $U_y \ni x$ and $V_y \ni y$ as X is Hausdorff. So $L \subset \bigcup_{y \in L} V_y$ (see Figure 7.1). As L is compact, $\exists$ a finite set $\{y_1, \ldots, y_n\} \subset L$ such that $L \subset \bigcup_{i=1}^{n} V_{y_i}$. Then

$$U^{(x)} := \bigcap_{i=1}^{n} U_{y_i} \qquad \text{and} \qquad V^{(x)} := \bigcup_{i=1}^{n} V_{y_i}$$

are open and disjoint (see Figure 7.2). With this we have $K \subset \bigcup_{x \in K} U^{(x)}$. So as K is

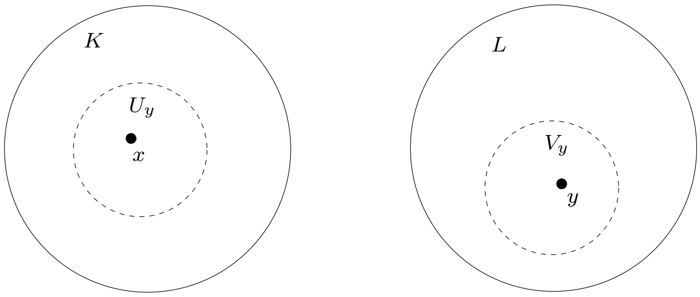

Figure 7.1: Disjoint sets K and L

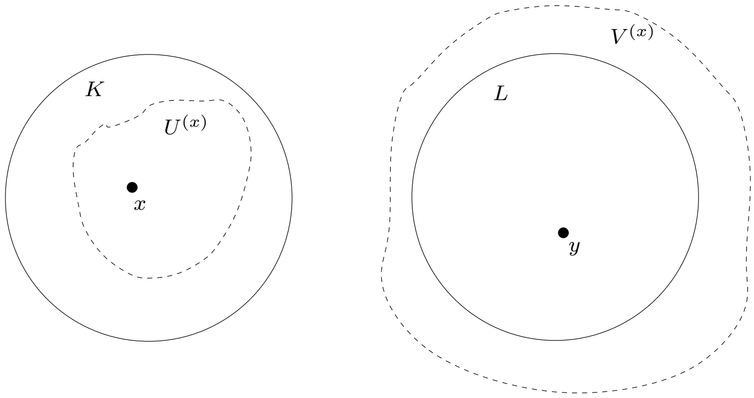

Figure 7.2: Disjoint sets $U^{(x)}$ and $V^{(x)}$

compact $\exists$ a finite set $\{x_1, \ldots, x_m\} \subset K$ such that $K \subset \bigcup_{j=1}^{m} U^{(x_j)}$. Define

$$U := \bigcup_{j=1}^{m} U^{(x_j)} \qquad \text{and} \qquad V := \bigcap_{j=1}^{m} V^{(x_j)}$$

then $K \subset U$ and $L \subset V$ with U, V open and $U \cap V = \emptyset$

(2) Let $x \in K$, then by locally compact property $\exists$ open set $W \ni x$ such that $\overline{W}$ compact. Let $Z = W \cap U \ni x$. Now $\{x\}$ and $\overline{Z} \setminus Z$ are compact and disjoint so by Lemma 7.1(1) $\exists$ open sets V_1, V_2 such that $x \in V_1$ and $\overline{Z} \setminus Z \subset V_2$ with $V_1 \cap V_2 = \emptyset$. Let $V^x := V_1 \cap Z$ then $V^x \ni x$ is open, $\overline{V^x}$ is compact (as $V^x \subset Z \subset W$) and $x \in V^x \subset \overline{V^x} \subset Z \subset U$. $\overline{V^x} \subset Z$ since $\overline{Z} \setminus Z$ is a subset of V_2 and V_2 is disjoint from $V_1 \supset V^x$. So $\overline{V^x} \cap (\overline{Z} \setminus Z) = \emptyset$). Now $K \subset \bigcup_{x \in K} V^x$. Since K is compact, $\exists$ a finite set $\{x_1, \ldots, x_p\} \subset K$ such that

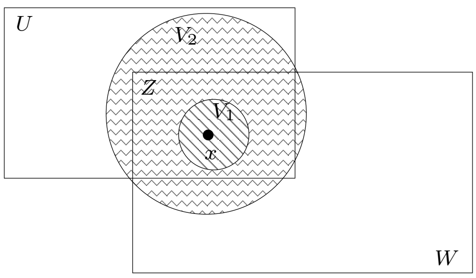

Figure 7.3: Disjoint sets V_1 and V_2

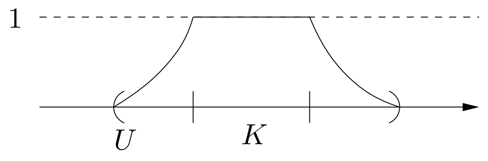

Figure 7.4: f with compact support $\mathrm{supp}(f) \subset U$

$K \subset \bigcup\limits_{i=1}^{p} V^{x_i}$. Let $V = \bigcup\limits_{i=1}^{p} V^{x_i}$, then $K \subset V \subset \overline{V} \subset U$ where V is open and $\overline{V}$ is compact.

$\square$

Definition 7.2. Let (X, τ) be a topological space. Then $V \subset X$ is **precompact** if V is open and $\overline{V}$ is compact.

Proposition 7.2 (Urysohn's Lemma)**.** Let $K \subset U \subset X$, where X is a LCH space, U is open and K is compact. Then $\exists$ continuous functions $f : X \to [0, 1]$ with compact support such that $f(x) = 1 \ \ \forall x \in K$ and $\mathrm{supp}(f) \subset U$ where $\mathrm{supp}(f) = \overline{\{x \mid f(x) \neq 0\}}$ (see Figure 7.4).

Proof. Let $K \neq \emptyset$ (else trivially true). By Lemma 7.1(2) $\exists$ a precompact set V_1 such that $K \subset V_1 \subset \overline{V_1} \subset U$. Hence there is a precompact set V_0 such that

$$K \subset V_0 \subset \overline{V_0} \subset V_1 \subset \overline{V_1} \subset U.$$

So there is a precompact $V_{\frac{1}{2}}$ such that

$$K \subset V_0 \subset \overline{V_0} \subset V_{\frac{1}{2}} \subset \overline{V_{\frac{1}{2}}} \subset V_1 \subset \overline{V_1} \subset U.$$

Continue the process, so for each dyadic

$$r \in D := \left\{ \frac{m}{2^n} \;\middle|\; 0 \leq m \leq 2^n, \ n \in \mathbb{N} \right\} \subset [0, 1]$$

we obtain a precompact set V_r such that $K \subset V_r \subset \overline{V_r} \subset V_s$ if $r < s$. If $t \in \mathbb{R}$ define

$$V_t := \begin{cases} \emptyset & \text{if } t < 0 \\ \bigcup\limits_{r<t} V_r & \text{if } 0 \leq t \leq 1 \text{ with } r \in D \\ X & \text{if } t > 1. \end{cases}$$

Then $\overline{V_t} \subset V_s$ if $t < s$ because if $0 \leq t < s \leq 1$ then $\exists \ r_1, r_2 \in D$ such that $t < r_1 < r_2 < s$ and so

$$\overline{V_t} \subset \overline{V_{r_1}} \subset \overline{V_{r_2}} \subset V_2. \tag{7.1}$$

Define $g : X \to [0, 1]$ by $g(x) := \inf\{t \in \mathbb{R} \mid x \in V_t\}$. Since $V_t = X$ if $t > 1 \Rightarrow g(X) \leq 1$ and $V_t = \emptyset \ \ \forall t < 0 \Rightarrow g \geq 0$. Note $x \in V_t$ if and only if $g(x) \leq t$. We show g is continuous. Let $\varepsilon > 0$ and $x_0 \in X$. Then

$$\begin{aligned} g^{-1}((g(x_0) - \varepsilon, g(x_0) + \varepsilon)) &= \{x \in X \mid |g(x) - g(x_0)| < \varepsilon\} \\ &= \{x \in X \mid g(x_0) - \varepsilon < g(x) < g(x_0) + \varepsilon\} \\ &= V_{g(x_0)+\varepsilon} \setminus \overline{V_{g(x_0)-\varepsilon}} \ni x_0. \end{aligned}$$

This is open. So g is continuous at x_0. Let $f := 1 - g$ to obtain the required function.

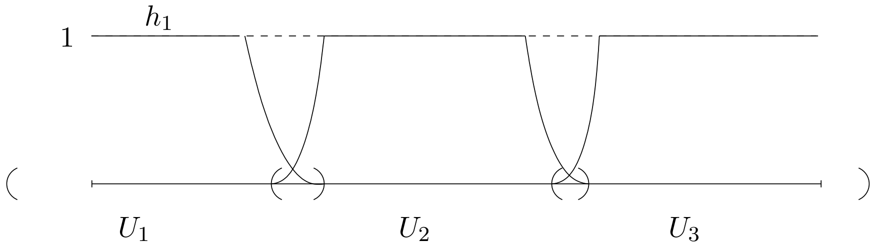

Figure 7.5: $h_j \in C_c(X)$ with $\text{supp}(h_j) \subset U_j \;\; \forall j$

$\square$

Notation. $C_c(X) :=$ continuous functions $f : X \to \mathbb{C}$ with compact support.

Proposition 7.3. Let (X, τ) be a LCH space, $K \subset X$ compact and let $\{U_1, \ldots, U_n\}$ be an open cover of K then $\exists\, h_j \in C_c(X)$ such that

(1) $\text{supp}(h_j) \subset U_j \;\; \forall j$;

(2) $h_j : X \to [0, 1]$;

(3) $\sum\limits_{j=1}^{n} h_j(x) = 1 \;\; \forall x \in K$.

See Figure 7.5.

Proof. Let $x \in K \Rightarrow x \in U_j$ for some j by Lemma 7.1(2) $\exists$ a precompact set N_x such that $x \in N_x \subset \overline{N_x} \subset U_j$. Since $K \subset \bigcup\limits_{x \in K} N_x$ with K compact $\exists$ a finite set $\{x_1, \ldots, x_m\} \subset K$ such that $K \subset \bigcup\limits_{i=1}^{m} N_{x_i}$. Define

$$F_j := \bigcup \left\{ \overline{N_{x_i}} \;\middle|\; \overline{N_{x_i}} \subset U_j \right\} \subset U_j$$

(which is compact). By Urysohn's Lemma $\exists\, g_1, \ldots, g_n \in C_c(X)$ such that $g_j = 1$ on F_j and $\text{supp}(g_j) \subset U_j$. Then $\sum\limits_{j=1}^{n} g_j \geq 1$ on K. By Urysohn's Lemma, $\exists\, f \in C_c(X)$ such that $f = 1$ on K and

$$\text{supp}(f) \subset \left\{ x \in X \;\middle|\; \sum_{j=1}^{n} g_j(x) > 0 \right\}.$$

Thus $\left(1 - f + \sum\limits_{j=1}^{n} g_j\right) > 0$. So we can define

$$h_j := \frac{g_j}{1 - f + \sum\limits_{j=1}^{n} g_j} \qquad \forall\, j = 1, \ldots, n.$$

Then $\text{supp}(h_j) = \text{supp}(g_j) \subset U_j$ and $\sum\limits_{j=1}^{n} h_j(x) = 1 \;\; \forall x \in K$.

$\square$

Notation. If $U \subset X$ is open, $f \in C_c(X)$ write $f \prec U$ if $0 \leq f \leq 1$ and $\text{supp}(f) \subset U$.

Definition 7.3. Let X be a LCH space. Then

(1) a **Radon measure** is a Borel measure $\mu : \mathcal{B}(X) \to [0, \infty]$ such that $\mu(K) < \infty$ if K is compact and

$$\mu(A) = \inf\{\mu(U) \mid U \supset A,\ U \text{ open}\} \qquad \forall\, A \in \mathcal{B}(X) = \text{ outer regularity (OR)}$$

and

$$\mu(V) = \sup\{\mu(K) \mid K \subset V,\ K \text{ compact}\} \qquad \forall\, V \textbf{ open } = \text{ inner regularity (IR)}$$

(2) a **positive linear functional** on $C_c(X)$ is a linear map $\omega : C_c(X) \to \mathbb{C}$ such that $\omega(f) \geq 0$ where $f \geq 0$.

By linearity and using $f = u_+ - u_- + iv_+ - iv_-$ we have $\omega(f) = \omega(u_+) - \omega(u_-) + i\omega(v_+) - i\omega(v_-)$. So a positive linear functional ω is uniquely determined by its values on positive functions. Moreover $\omega(f) \in \mathbb{R}$ if f is real-valued.

Theorem 7.4 (Riesz Representation Theorem)**.** Let X be LCH space. For each positive functional ω on $C_c(X)$ there is a unique Radon measure μ such that

$$\omega(f) = \int_X f\ d\mu \qquad \forall\, f \in C_c(X).$$

Moreover μ satisfies:

$$\mu(U) = \sup\{\omega(f) \mid f \in C_c(X),\ f \prec U\} \qquad \forall \text{open } U \tag{7.2}$$
$$\mu(K) = \inf\{\omega(f) \mid f \in C_c(X),\ f \geq \chi_K\} \qquad \forall \text{compact } K \tag{7.3}$$

Proof. We will first prove uniqueness. Let μ be a Radon measure such that

$$\omega(f) = \int_X f\ d\mu \qquad f \in C_c(X).$$

If U is open and $f \prec U$ then $\omega(f) = \int_X f\ d\mu \leq \mu(U)$ (as $f(X) = [0, 1]$ and $\text{supp}(f) \subset U$). If $K \subset U$ is compact, then Urysohn's Lemma $\Rightarrow$ $\exists\, f \in C_c(X)$ such that $f \prec U$ and $f = 1$ on K so

$$\mu(K) \leq \int_X f\ d\mu = \omega(f) \leq \mu(U).$$

By IR property for μ, we have just proven (7.2) for μ. Thus μ is determined on open sets by ω, hence it is uniquely determined on all $\mathcal{B}(X)$. Inspired by (7.2) we define

$$\mu(U) := \sup\{\omega(f) \mid f \in C_c(X),\ f \prec U\} \qquad \forall \text{open } U$$

and

$$\mu(K) := \inf\{\omega(f) \mid f \in C_c(X),\ f \geq \chi_k\} \qquad \forall\, E \subseteq X.$$

For open U, V with $U \subset V$ we have $\mu(U) \leq \mu(V)$ (from definition). So $\mu^*(U) = \mu(U)$ (U open). We proceed to prove the rest in a series of steps:

Step I: μ^* is an outer measure.

Proof. Let $U = \bigcup_{j=1}^{\infty} U_j$, U_j open, then by Urysohn's lemma $\exists\, f \prec U$ with compact support K. So $\exists\, n < \infty$ such that $K \subset \bigcup_{j=1}^{n} U_j$. Thus by Proposition 7.3 $\exists\, h_j \in C_c(X)$ such that $h_j \prec U_j$ and $\sum_{j=1}^{n} h_j = 1$ on K. But then $f = \sum_{j=1}^{n} f h_j$ such that $f h_j \prec U_j$. So

$$\omega(f) = \sum_{j=1}^{n} \omega(f h_j) \leq \sum_{j=1}^{n} \mu(U_j) \leq \sum_{j=1}^{\infty} \mu(U_j).$$

This is true for all $f \prec U$ so by definition $\mu(U) \leq \sum_{j=1}^{\infty} \mu(U_j)$. Then for any $E \subseteq X$ we have

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(U_j) \;\middle|\; U_j \text{ open, } E \subset \bigcup_{j=1}^{\infty} U_j \right\}.$$

From Theorem 2.3 it follows that μ^* is an outer measure.

$\triangledown$

Step II: $\mathcal{B}(X) \subset \mathcal{M}_{\mu^*}$.

Proof. We only need to show that all open sets are μ^*-measurable (by Theorem 2.3 $\Rightarrow \mathcal{M}_{\mu^*}$ a σ-algebra) i.e. if U open and $E \subseteq X$ with $\mu^*(E) < \infty$, then we need to show

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \cap U^c). \tag{7.4}$$

If E is open $\Rightarrow E \cap U$ is open. So $\forall\, \varepsilon > 0\ \exists\, f \in C_c(X)$ such that $f \prec E \cap U$ and $\omega(f) > \mu(E \cap U) - \varepsilon$ (by definition). Also $E \setminus \operatorname{supp}(f)$ open so $\exists\, g \in C_c(X)$ such that $g \prec E \setminus \operatorname{supp}(f)$ and $\omega(g) > \mu(E \setminus \operatorname{supp}(f)) - \varepsilon$. Then $f + g \prec E$. So

$$\begin{aligned} \mu(E) &\geq \omega(f) + \omega(g) \\ &> \mu(E \cap U) + \mu(E \setminus \operatorname{supp}(f)) - 2\varepsilon \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U) - 2\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ arbitrary we have that (7.4) holds for open E. For general $E \subseteq X$ with $\mu^*(E) < \infty$ we can find open $V \supset E$ such that $\mu(V) < \mu^*(E) + \varepsilon$ (by definition). So

$$\begin{aligned} \mu^*(E) + \varepsilon &> \mu(V) \\ &\geq \mu^*(V \cap U) + \mu^*(V \cap U^c \quad \text{by (7.4) for } V \text{ open} \\ &\geq \mu^*(E \cap U) + \mu^*(E \cap U^c. \end{aligned}$$

Since ε is arbitrary (7.4) holds $\forall\, E$.

$\triangledown$

Thus $\mu^* \restriction \mathcal{B}(X) =: \mu$ is a Borel measurable which satisfies (7.2) by construction.

Step III: μ satisfies (7.3).

Proof. Let $K \subset X$ compact, $f \in C_c(X)$ with $f \geq \chi_K$. Define $\forall \varepsilon > 0$ the open set

$$U_\varepsilon = f^{-1}((1-\varepsilon, \infty)) \supset K \qquad (= \{x \in X \mid f(x) > 1-\varepsilon\}).$$

If $g \prec U_\varepsilon$ then $g \leq \left(\frac{1}{1-\varepsilon}\right) f$ so $\omega(g) \leq \left(\frac{1}{1-\varepsilon}\right)\omega(f)$. Thus

$$\begin{aligned} \mu(K) &\leq \mu(U_\varepsilon) \\ &= \sup\{\omega(g) \mid g \prec U_\varepsilon\} \quad \text{by (7.2)} \\ &\leq \left(\tfrac{1}{1-\varepsilon}\right)\omega(f). \end{aligned}$$

But $\varepsilon > 0$ is arbitrary $\Rightarrow \mu(K) \leq \omega(f)$. Moreover if $U \supset K$ open then Urysohn's Lemma $\Rightarrow \exists\, f \in C_c(X)$ such that $f \prec U$ and $f \geq \chi_K$. Thus $\omega(f) \leq \mu(U)$ by (7.2). So $\mu(K) \leq \omega(f) \leq \mu(U)$. By definition of μ^*:

$$\mu(K) = \inf\{\omega(f) \mid f \in C_c(X),\ f \geq \chi_K\} \qquad \forall\, K \text{ compact}.$$

∇

By (7.3), $\mu(K) < \infty$ if K is compact. By definition of μ^*, μ is outer regular (OR). For inner regularity (IR), let $V \subseteq X$ open, $\alpha < \mu(V)$ and choose $f \in C_c(X)$ with $f \prec V$, $\omega(f) > \alpha$ (this is possible by (7.2)). If $K = \operatorname{supp}(f) \subset V$, $g \geq \chi_K$, $g \in C_c(X)$ then $g \geq f$. So $\omega(g) \geq \omega(f) > \alpha$. So by (7.3) $\mu(K) > \alpha$ i.e. $\alpha < \mu(K) < \mu(V)$, $\alpha < \mu(V)$ arbitrary. So

$$\mu(V) = \sup\{\mu(K) \mid K \subset V,\ K \text{ compact }\} = \text{ Inner Regularity (IR).}$$

Thus μ is a Radon measure.

Step IV: $\omega(f) = \int_X f\, d\mu \qquad \forall\, f \in C_c(X)$.

Proof. Since $C_c(X) = \operatorname{sp}\{f \in C_c(X) \mid f(x) \subseteq [0,1]\}$ we only need to prove that $\omega(f) = \int_X f\, d\mu$ for $f \in C_c(X)$ with $f(X) \subseteq [0,1]$. Define $K_j := f^{-1}([\frac{j}{N}, \infty))$ for $j = 1, 2, \ldots, N$, $K_0 = \operatorname{supp}(f)$ and

$$f_j(x) := \begin{cases} 0 & \text{if } x \notin K_{j-1} \\ f(x) - \frac{j-1}{N} & \text{if } x \in K_{j-1} \setminus K_j \\ \frac{1}{N} & \text{if } x \in K_j. \end{cases}$$

See Figure (7.6). So $f_j \in C_c(X)$ and $\frac{\chi_{K_j}}{N} \leq f_j \leq \frac{\chi_{K_{j-1}}}{N}$.

$$\Rightarrow \frac{\mu(K_j)}{N} \leq \int f_j\, d\mu \leq \frac{\mu(K_{j-1})}{N}.$$

If $U \supset K_{j-1}$ is open then $\omega(f_j) \leq \frac{\mu(U)}{N}$ $\forall\, U \supset K_{j-1}$ by (7.2) and $N f_j \prec U$. So by outer regularity (OR) $\omega(f_j) \leq \frac{\mu(K_{j-1})}{N}$. Thus (7.3)

$$\Rightarrow \frac{\mu(K_j)}{N} \leq \omega(f_j) \leq \frac{\mu(K_{j-1})}{N}.$$

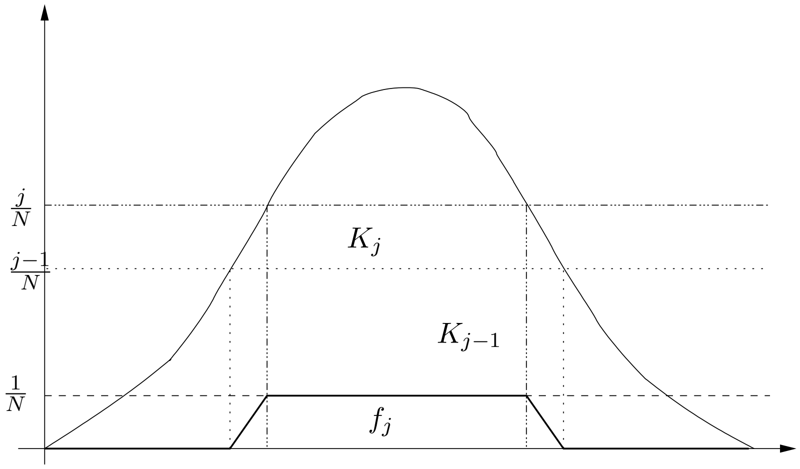

Figure 7.6: Defining K_j and f_j.

Now $f = \sum_{j=1}^{N} f_j$ so

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leq \int_X f\ d\mu \leq \frac{1}{N}\sum_{j=1}^{N-1}\mu(K_{j-1})$$

and

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leq \omega(f) \leq \frac{1}{N}\sum_{j=1}^{N-1}\mu(K_{j-1})$$

$$\Rightarrow \left|\omega(f) - \int_X f\ d\mu\right| \leq \frac{1}{N}\left[\mu(K_0) - \mu(K_n)\right] \leq \frac{\mu(\mathrm{supp}(f))}{N} < \infty.$$

As N is arbitrary we have $\omega(f) = \int_X f\ d\mu$.

$\triangledown$

$\square$

Example 7.2.

(1) On $C_c(X)$, the Riemann integral defines a positive function $\omega(f) = \int f\, d^n x$. It's associated Radon measure is the Lebesgue measure.

(2) Let X be LCH and choose points $x, y \in X$ and define $\omega(f) = f(x) + f(y)\ \ \forall f \in C_c(X)$ then $\omega : C_c(X) \to \mathbb{C}$ is a positive functional, hence has a Radon measure.

Theorem 7.5. Let X be LCH space such that every open set is σ-compact (this is the case if X is 2^{nd} countable). Then every Borel measure which is finite on compact sets is a Radon measure (i.e. it satisfies IR and OR).

Proof. Let μ be a Borel measure which is finite on compact sets, then $C_c(X) \subset \mathcal{L}^1(\mu)$ hence $\omega(f) := \int f\ d\mu$ is a positive linear functional on $C_c(X)$. So by Riesz Representation Theorem (Theorem 7.4) $\exists$ Radon measure ν such that $\omega(f) = \int f\ d\nu\ \ \forall f \in C_c(X)$. Let $U \subseteq X$ open by σ-compactness $U = \bigcup_{j=1}^{\infty} K_j$ with K_j compact. By Urysohn's Lemma, choose $f_1 \in C_c(X)$

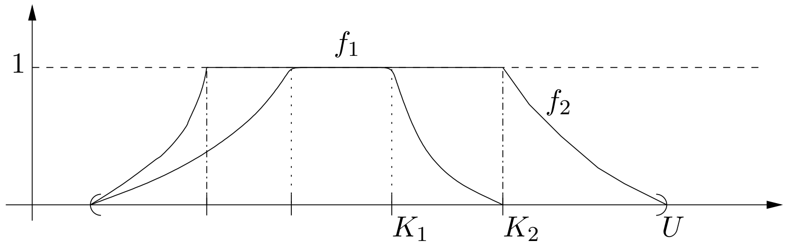

Figure 7.7: Defining K_j and f_j.

such that $f_1 \prec U$ and $f_1 = 1$ on K_1. For $n > 1$ let $f_n \in C_c(X)$ be chosen such that $f_n \prec U$ and $f_n = 1$ on $\bigcup_{j=1}^{n} K_j \cup \bigcup_{l=1}^{n-1} \operatorname{supp}(f_l)$. See Figure (7.7). Then f increases pointwise to χ_U as $n \to \infty$ so

$$\begin{aligned}
\mu(U) &= \int \chi_U \, d\mu \\
&= \lim_{n\to\infty} \int f_n \, d\mu \quad \text{by MCT} \\
&= \lim_{n\to\infty} \int f_n \, d\nu \\
&= \nu(U).
\end{aligned}$$

So μ, ν coincide on all open sets, hence on all Borel sets. So $\mu = \nu$.

$\square$

Chapter 8

L^p-spaces

For this chapter we assume a fixed measure space $(X, \mathcal{A}, \mu)$.

Definition 8.1. Let $p \in (0, \infty)$ and define

$$\mathcal{L}^p(\mu) := \left\{ f : X \to \mathbb{C} \;\middle|\; f \text{ measurable },\ \int_X |f|^p \, d\mu < \infty \right\}.$$

(Note that $\chi_E \in \mathcal{L}^p(\mu)\ \forall\, E \in \mathcal{A}$ with $\mu(E) < \infty$). On $\mathcal{L}^p(\mu)$ define an equivalence relation: $f \sim g$ if and only if $f = g$ μ-a.e. i.e. $\mu(\{x \in X \mid f(x) \neq g(x)\}) = 0$.

(1) Let $L^p(\mu) := \mathcal{L}^p(\mu)/\sim$ i.e. identify f, g if $f = g$ μ-a.e.

(2) If $\mu =$ counting measure on X, denote $l^p(X) := L^p(\mu)$,

$$l^p \equiv l^p(\mathbb{N}) = \left\{ \{a_n\}_{n=1}^{\infty} \;\middle|\; a_n \in \mathbb{C},\ \sum_{n=1}^{\infty} |a_n|^p < \infty \right\}.$$

(3) Define $\|\cdot\|_p := L^p(\mu) \to [0, \infty)$ by

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}} \qquad f \in L^p(\mu).$$

We usually do not indicate equivalence classes in $L^p(\mu)$. Note that $\|\lambda f\|_p = |\lambda| \cdot \|f\|_p\ \ \forall\, \lambda \in \mathbb{C}$ and $f \in L^p(\mu)$.

Proposition 8.1.

(1) $L^p(\mu)$ is a $\mathbb{C}$-linear space $\forall\, p \in (0, \infty)$.

(2) If X has two disjoint sets of non-zero finite measure with respect to μ, then $\|\cdot\|_p$ is NOT a norm when $0 < p < 1$.

Proof.

(1)

$$\begin{aligned} \textstyle\int |f+g|^p \, d\mu &\le \textstyle\int [2\max(|f|, |g|)]^p \ d\mu \\ &\le 2^p \textstyle\int (|f|^p + |g|^p) \, d\mu \\ &< \infty \qquad \forall\, f, g \in L^p(\mu). \end{aligned}$$

$\Rightarrow f + g \in L^p(\mu)$. By $\|\lambda f\|_p = |\lambda| \cdot \|f\|_p < \infty,\ f \in \mathbb{C},\ f \in L^p(\mu) \Rightarrow \lambda f \in L^p(\mu)$.

(2) Let $p \in (0,1)$, $E, F \in \mathcal{A}$ such that $E \cap F = \emptyset$ and $\infty > \mu(E) \neq 0 \neq \mu(F) < \infty$. Then $\chi_E, \chi_F \in L^p(\mu)$ and

$$\begin{aligned}
\|\chi_E + \chi_F\|_p &= \left(\int (\chi_E + \chi_F)^p \, d\mu\right)^{\frac{1}{p}} \\
&= \Big(\int \underbrace{(\chi_E + \chi_F)}_{\chi_{E\cup F}} \, d\mu\Big)^{\frac{1}{p}} \\
&> (\mu(E))^{\frac{1}{p}} + (\mu(F))^{\frac{1}{p}} \\
&= \|\chi_E\|_p + \|\chi_F\|_p
\end{aligned}$$

using $(a+b)^{\frac{1}{p}} > a^{\frac{1}{p}} + b^{\frac{1}{p}}$ if $a, b > 0$ and $p \in (0,1)$.

□

Theorem 8.2 (Hölder's inequality)**.** Let $p \in (1, \infty)$, $q = \frac{p}{(p-1)}$ (i.e. $p^{-1} + q^{-1} = 1$). If f, g are measurable functions on X, then

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q. \tag{8.1}$$

Thus if $f \in L^p(\mu)$, $f \in L^q(\mu)$ then $f \cdot g \in L^1(\mu)$. In this case we have equality in (8.1) if and only if $\alpha|f|^p = \beta|g|^q$ μ-a.e. for $\alpha, \beta > 0$.

Proof. The result is trivial if $\|f\|_p = 0$ or ∞ or $\|g\|_q = 0$ or ∞. So we assume otherwise.

Lemma 8.2.1. If $a, b \geq 0$ and $r \in (0,1)$ then

$$a^r b^{1-r} \leq ra + (1-r)b \tag{8.2}$$

with equality if and only if $a = b$.

Proof. Trivial if $b = 0$. So let $b \neq 0, t := \frac{a}{b}$. Let $f(t) = t^r - rt - (1-r)$. So $f'(t) = r(t^{r-1} - 1)$. So f has a global max of 0 at $t = 1$ i.e. $t^r \leq rt + (1-r)$ with equality at only $t = 1$. Multiply by b to get the inequality (8.2).

▽

Substitute

$$a = \left|\frac{f(x)}{\|f\|_p}\right|^p, \qquad b = \left|\frac{g(x)}{\|g\|_q}\right|^q \qquad \text{and} \qquad r = \frac{1}{p}$$

into the inequality (8.2) to obtain

$$\frac{|f(x)g(x)|}{\|f\|_p \cdot \|g\|_q} \leq \frac{|f(x)|^p}{p\int |f|^p \, d\mu} + \frac{|g(x)|^q}{q\int |g|^q \, d\mu}. \tag{8.3}$$

Integrate both sides of the inequality (8.3):

$$\frac{\|fg\|_1}{\|f\|_p \cdot \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

which gives the inequality (8.1). Equality holds if and only if it holds a.e. in the inequality (8.3) and this happens if $a = b$ i.e. $\|g\|_q^q \, |f|^p = \|f\|_p^p \, |g|^q$ a.e.

$\square$

Theorem 8.3. Let $1 \leq p < \infty$, then

(1) $\|\cdot\|_p$ is a norm on $L^p(\mu)$ i.e.

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p \qquad f, g \in L^p(\mu) \qquad (\textbf{Minkowski inequality}).$$

(2) $L^p(\mu)$ is a Banach space ($\equiv$ complete normed space) (**Riesz-Fischer Theorem**).

Proof.

(1) The result is obvious if $p = 1$ or if $f + g = 0$ μ-a.e. Otherwise

$$|f+g|^p \leq (|f| + |g|)\, |f+g|^{p-1}.$$

So

$$\begin{aligned} \textstyle\int |f+g|^p \, d\mu &\leq \textstyle\int (|f|+|g|)|f+g|^{p-1} \, d\mu \\ &\leq \|f\|_p \cdot \| \, |f+g|^{p-1} \, \|_q + \|g\|_p \cdot \| \, |f+g|^{p-1} \, \|_q \\ &= (\|f\|_p + \|g\|_q) \left(\textstyle\int |f+g|^p \, d\mu\right)^{\frac{1}{q}} \end{aligned}$$

by Hölder's inequality, $p^{-1} + q^{-1} = 1$. So

$$\|f+g\|_p = \left(\int |f+g|^p \, d\mu\right)^{\frac{1}{p}} \leq \|f\|_p + \|g\|_q.$$

(2) Let $\{f_n\}_{n=1}^{\infty} \subset L^p(\mu)$ be a Cauchy sequence. We may assume it has a subsequence $\{f_{n_i}\}_{i=1}^{\infty}$ with $n_1 < n_2 < \ldots$ such that

$$\|f_{n_{i+1}} - f_{n_i}\| < \frac{1}{2^i}$$

(for example by inserting terms $\frac{f_n + f_{n+1}}{2}$). Let $g_k := \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}|$, $g = \lim_{k\to\infty} g_k$ (may be ∞). So g_k and hence $(g_k)^p$ is increasing. By Theorem 8.3(1)

$$\|g_k\|_p \leq \sum_{i=1}^{k} \|f_{n_{i+1}} = f_{n_i}\|_p.$$

By Fatou's Lemma (Theorem 4.4) we have

$$\begin{aligned} \|g\|_p^p &= \textstyle\int \lim\limits_{k\to\infty} (g_k)^p \, d\mu && \text{by MCT} \\ &\leq \lim_{k\to\infty} \left(\inf_{i\geq k} \textstyle\int (g_i)^p \, d\mu\right) && \text{by Fatou's Lemma} \\ &= \lim_{k\to\infty} \left(\inf_{i\geq k} \|g_i\|^p\right) \\ &\leq 1. \end{aligned}$$

Hence $g(x) < \infty$ μ-a.e., hence the series

$$f_{n_k} = f_{n_1} + \sum_{i=1}^{k-1} (f_{n_{i+1}} - f_{n_i})$$

converges (absolutely) as $k \to \infty$ μ-a.e. So we can define $f(x) = \lim_{k\to\infty} f_{n_k}(x)$ μ-a.e. We have that

$$f = f_{n_1} + \sum_{i=1}^{\infty} (f_{n_{i+1}} - f_{n_i})$$

which implies

$$\|f\|_p \leq \|f_{n_1}\|_p + \sum_{i=1}^{\infty} \|f_{n_{i+1}} - f_{n_i}\|_p$$

so $f \in L^p(\mu)$. We want to show that f is the L^p-limit of f_n. Since $\{f_n\}_{n=1}^{\infty}$ is Cauchy, for $\varepsilon > 0$ $\exists N > 0$ such that $\|f_n - f_m\|_p < \varepsilon$ $\forall n, m > N$

$$\lim_{k\to\infty} |f_{n_k} - f_m| = |f - f_m| \qquad \mu\text{-a.e.}$$

So by Fatou's Lemma

$$\int |f - f_m|^p \, d\mu \leq \lim_{k\to\infty} \left(\inf_{i \geq k} \int |f_{n_i} - f_m|^p \, d\mu \right) \leq \varepsilon^p.$$

where $\int |f_{n_i} - f_m|^p \, d\mu = \|f_{n_i} - f_m\|_p^p$. Since ε is arbitrary, $\|f - f_m\|_p \xrightarrow[\infty]{n} 0$.

□

Remarks 8.1. For $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, we see from Hölder's inequality that for $g \in L^p(\mu)$ we can define a functional $\varphi_g : L^p(\mu) \to \mathbb{C}$ by $\varphi_g(f) := \int fg \, d\mu < \infty$ which is a functional with norm $\|\varphi_g\| \leq \|g\|_q$. So $\varphi_g \in L^p(\mu)^*$ (dual). In fact we have $L^p(\mu) \cong L^q(\mu)$ by this. The proof of this result is via the Radon-Nikodym Theorem in Chapter 9.

Theorem 8.4. Let $1 \leq p \leq \infty$, then

$$S = \left\{ \sum_{j=1}^{n} a_j \cdot \chi_{A_j} \;\middle|\; A_j \in \mathcal{A},\ \mu(A_j) < \infty,\ a_j \in \mathbb{C},\ n \in \mathbb{N} \right\} \equiv \text{ simple functions}$$

is dense in $L^p(\mu)$ with respect to $\|\cdot\|_p$.

Proof. Obviously we always have that such simple functions are in $L^p(\mu)$. Let $f \in L^p(\mu)$, choose a sequence of simple functions s_n as in Theorem 3.5(4) such that $s_n \to f$ and $0 \leq |s_1| \leq |s_2| \leq \cdots \leq |f|$. Since $\|s_n\|_p \leq \|f\|_p < \infty \Rightarrow s_n \in L^p(\mu)$. So if $s_n = \sum_{j=1}^{N} a_j \cdot \chi_{A_j}$ then

$$\|s_n\|_p^p = \int |s_n|^p \, d\mu = \sum_{j=1}^{N} |a_j|^p \mu(A_j) < \infty$$

$\Rightarrow \mu(A_j) < \infty$ $\forall j$, so s_n is of required type. Now

$$|s_n - f|^p \leq (|s_n| + |f|)^p \leq 2^p |f|^p \in L^1(\mu).$$

So by DCT $\|s_n - f\|_p \to 0$. So simple functions are dense in $L^p(\mu)$.

□

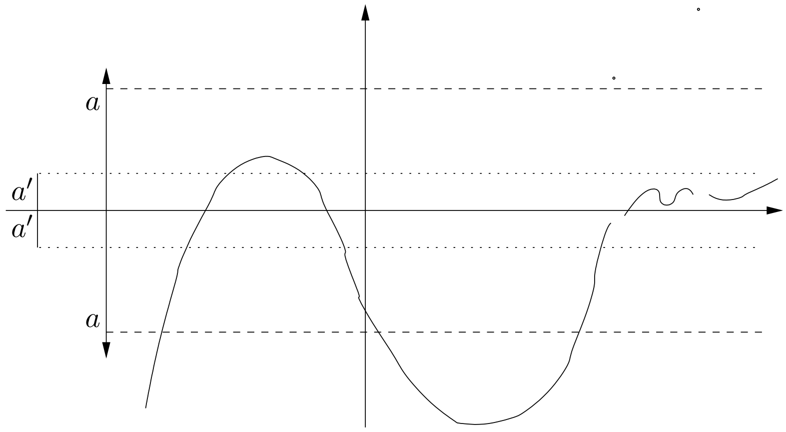

Figure 8.1: Essential Supremum

Remarks 8.2.

(1) $L^2(\mu)$ is special as it is a Hilbert space ($L^2(\mu)^* \cong L^2(\mu)$).

(2) In general we have $L^p \not\subset L^q$ for $p \neq q$.

Example 8.1. Let $(X, \mathcal{A}, \mu) = ((0,\infty), \mathcal{B}((0,\infty)), \mu)$ where μ is the Lebesgue measure. Let $f_\alpha(x) := \frac{1}{x^\alpha}$ then $f_\alpha \cdot \chi_{(0,1)} \in L^p(\mu)$ if and only if $p < \frac{1}{\alpha}$ and $f_\alpha \cdot \chi_{(1,\infty)} \in L^p(\mu)$ if and only if $p > \frac{1}{\alpha}$. Thus $L^p \not\subset L^q$ if $p \neq q$. Inclusion depends on the measure μ, For example $l^p \subset l^q$ if $1 \leq p < q \leq \infty$.

Definition 8.2. Let $f : X \to \mathbb{C}$ be measurable, then

(1) $\|f\|_\infty := \inf\{a \geq 0 \mid \mu(\{x \mid |f(x)| > a\}) = 0\} \equiv \text{ess. } \sup_{x\in X}|f(x)|$
$\equiv$ **essential supremum of f**, with convention $\inf \emptyset = \infty$. See Figure (8.1).

(2) $\mathcal{L}^\infty(\mu) := \{f : X \to \mathbb{C} \mid f \text{ is measurable and } \|f\|_\infty < \infty\}$.

(3) $L^\infty(\mu) := \mathcal{L}^\infty(\mu)/\sim$ where $f \sim g$ if and only if $f = g$ μ-a.e.

Note $f \in \mathcal{L}^\infty(\mu)$ if $\exists$ bounded measurable functions $g : X \to \mathbb{C}$ such that $f = g$ μ-a.e.

Theorem 8.5.

(1) $L^\infty(\mu)$ is a Banach space with norm $\|\cdot\|_\infty$.

(2) $S \equiv$ set of bounded simple functions, is dense in $L^\infty(\mu)$.

(3) If f, g are measurable functions on X, then $\|f \cdot g\|_1 \leq \|f\|_1 \cdot \|g\|_\infty$. If $f \in L^1(\mu)$, $g \in L^\infty(\mu)$ then $\|f \cdot g\|_1 = \|f\|_1 \cdot \|g\|_\infty$ if and only if $|g(x)| = \|g\|_\infty$ μ-a.e. on $(f^{-1}(\{0\}))^c$.

Proof. Exercise.

Theorem 8.6. If X is LCH space and μ is a Radon measure, then $C_c(X)$ is dense in $L^p(\mu)$ for $1 \leq p < \infty$.

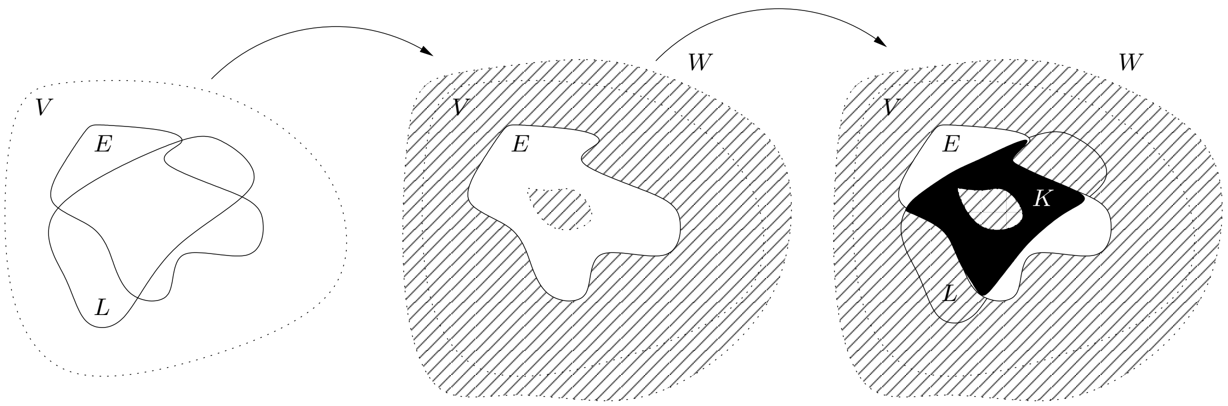

Figure 8.2: Construction of sets using IR and OR of Radon measure.

Proof. By Theorem 8.4 it suffices to show that in L^p-norm we can approximate any $\chi_E,\ E \in \mathcal{B}(X)$ such that $\mu(E) < \infty$ by $f \in C_c(X) \subset L^p(\mu)$.

Fix $\varepsilon > 0$ and $E \in \mathcal{B}(X)$ with $\mu(E) < \infty$. By OR $\exists$ open set $V \supset E$ such that $\mu(V) < \mu(E) + \varepsilon$. By IR $\exists$ compact $L \subset V$ such that $\mu(L) > \mu(V) - \varepsilon$. So

$$\mu(V \setminus E) < \varepsilon. \tag{8.4}$$

Thus we can choose by OR an open set $W \supset V \setminus E$ such that $\mu(W) < \varepsilon$. Let $K := L \setminus W$ (as L is compact and W is open $\Rightarrow K$ compact). $K \subset E$ since $V \setminus E = V \cap E^c \subset W \Rightarrow W^c \subset V^c \cup E$. Now, $K := L \setminus W = L \cap W^c \subset L \cap (V^c \cup E) = L \cap E \subset E$. See Figure (**??**). So

$$\mu(K) = \mu(L \setminus W) = \mu(L) - \underbrace{\mu(L \cap W)}_{<\varepsilon} > \mu(V) - 2\varepsilon \geq \mu(E) - \varepsilon. \tag{8.5}$$

By Urysohn's Lemma (Proposition 7.2) $\exists\, f \subset C_c(X),\ f : X \to [0,1]$ such that $\operatorname{supp}(f) \subset V$ and $f(x) = 1\ \ \forall\, x \in K$. Now $\{x \mid \chi_E(x) \neq f(x)\} \subseteq V \setminus K = (V \setminus E) \cup (E \setminus K)$ (disjoint). So

$$\mu(\{x \mid \chi_E(x) \neq f(x)\}) \leq \mu(V \setminus E) + \mu(E \setminus K) < 3\varepsilon$$

by inequalities (8.4) and (8.5). Thus

$$\begin{aligned} \|\chi_E - f\|_p^p &= \textstyle\int_X |\chi_E - f|^p \, d\mu \\ &\leq \mu(\{x \mid \chi_E(x) \neq f(x)\}) \\ &< 3\varepsilon \end{aligned}$$

using $|\chi_E - f| \in [0,1]$. Since ε is arbitrary, it follows we can approximate χ_E in l^p-norm by an $f \in C_c(X)$.

$\square$

Chapter 9

Non-positive measures and Radon-Nikodym Theorem

Definition 9.1. Let $(X, \mathcal{A})$ be a σ-algebra, then a signed-measure is a map $\nu : \mathcal{A} \to [-\infty, \infty]$ such that

(1) $\nu(\emptyset) = 0$;

(2) $\nu(\mathcal{A}) \not\supset \{-\infty, \infty\}$, where $\nu(\mathcal{A})$ is the range of ν;

(3) if $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ are disjoint then

$$\nu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \nu(A_j) \qquad \text{(Countable addivitivity)}$$

ν is **finite** if $\nu(\mathcal{A}) \cap \{-\infty, \infty\} = \emptyset$.

Example 9.1.

(1) Any positive measure is a signed measure.

(2) If μ_1, μ_2 are positive measures, one which is finite then $\nu = \mu_1 - \mu_2$ is a signed measure.

Theorem 9.1. Let $(X, \mathcal{A}, \nu)$ be a signed measure space, then

(1) If $A_j \in \mathcal{A}$ with $A_1 \subset A_2 \subset \dots$ then

$$\nu\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \nu(A_j)$$

(2) If $A_j \in \mathcal{A}$ with $A_1 \supset A_2 \supset \dots$ and if $\nu(A_1) < \infty$ then

$$\nu\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \nu(A_j).$$

Proof. Similar to Theorem 2.2

$\square$

Similar to the decomposition $f = f_+ - f_-$ for functions, we want to decompose measures in terms of a positive part and a negative part.

Definition 9.2. Let $(X, \mathcal{A}, \nu)$ be a signed measure space, then a set $A \in \mathcal{A}$ is **positive** (respectively **negative**, **null**) if $\forall B \in \mathcal{A}$ with $B \subseteq A$ we have $\nu(B) \geq 0$ (respectively $\nu(B) \leq 0,\ \nu(B) = 0$).

Theorem 9.2 (Hahn Decomposition Theorem)**.** Let $(X, \mathcal{A}, \nu)$ be a signed measure space. Then $\exists$ partition $X = P \cup N$ with $P \cap N = \emptyset$ such that P is a positive set and N is a negative set for ν. If P', N' is another such partition, then $P \Delta P' := (P \cup P') \setminus (P \cap P') = N \Delta N'$ is ν-null, where Δ is the symmetric difference.

Proof. Without loss of generality assume $+\infty \notin$ range(ν) (else consider $-\nu$). Let

$$\beta := \sup\{\nu(A) \mid A \text{ is a positive set for } \nu\}.$$

Then $\beta \geq 0$ (as $\emptyset$ is a positive set). Let $A_k \in \mathcal{A}$ be a sequence of positive sets such that $\beta = \lim\limits_{k\to\infty} \nu(A_k)$. Now $P := \bigcup\limits_{k=1}^{\infty} A_k$. We show P is positive.

Let $B_1 := A_1$ and $B_n := A_n \setminus \bigcup\limits_{k=1}^{n-1} A_k$ if $n \geq 2$. Then B_j are disjoint, $P = \bigcup\limits_{n=1}^{\infty} B_n$. So if $E \subseteq P,\ E \in \mathcal{A}$ then $B_n \cap E \subset A_n$ are disjoint and $\nu(B_n \cap E) \geq 0$ as A_n is positive. So

$$\nu(E) = \nu\left(\bigcup_{n=1}^{\infty}(B_n \cap E)\right) = \sum_{n=1}^{\infty} \nu(B_n \cap E) \geq 0$$

by countable additivity. Thus P is positive, hence $\nu(P) \leq \beta$ and $\nu(P \setminus A_k) \geq 0\ \ \forall k$. Thus

$$\nu(P) = \nu(A_k) + \nu(P \setminus A_k) \geq \nu(A_k).$$

Thus $\nu(P) \geq \lim\limits_{k\to\infty} \nu(A_k) = \beta$. Thus $\nu(P) = \beta < \infty$ (as $\infty \in$ range(ν)). Let $N = X \setminus P$ and $E \subseteq N,\ E \in \mathcal{A}$. If $\nu(E) > 0$ then we use:

Claim. For any $E \in \mathcal{A}$ with $0 < \nu(E) < \infty$, there is a set $F \subseteq E$ which is a positive set for ν and $\nu(F) > 0$.

Proof. Either E is a positive set (then take $E = F$) or it has sets of negative measure. In the latter cases, let $n_1 \in \mathbb{N}$ be the smallest number such that $\exists$ measurable set $E_1 \subset E$ with $\nu(E_1) = -\frac{1}{n_1}$. If $E \setminus E_1$ is not positive, let $n_2 \in \mathbb{N}$ be the smallest number such that $\exists$ measurable set $E_2 \subset E \setminus E_1$ with $\nu(E_2) < -\frac{1}{n_2}$. Continue inductively, to obtain sequence $\{n_k\}$ such that $E_k \subset E \setminus \bigcup\limits_{j=1}^{k-1} E_j$ with $n_k > 0$ being the smallest number such that $\nu(E_k) < -\frac{1}{n_k}$. If the series terminates, then the proof finishes. Thus assume it does not.

Let $F := E \setminus \bigcup\limits_{j=1}^{\infty} E_j$, so we have a disjoint union $E = F \cup \left(\bigcup\limits_{j=1}^{\infty} E_j\right)$, hence

$$0 < \nu(E) = \nu(F) + \sum_{j=1}^{\infty} \nu(E_j) < \infty.$$

So since $\nu(F) < \infty$, the series $\sum_{j=1}^{\infty} \nu(E_j)$ is convergent (absolutely convergent as all terms are negative) hence by $\frac{1}{n_j} < |\nu(E_j)|$ we get $\sum_{j=1}^{\infty} \frac{1}{n_j} < \infty$. So $\frac{1}{n_j} \to 0$. Obviously $\nu(F) > 0$. We show that F is positive.

Let $L \subseteq F,\ L \in \mathcal{A}$ and fix $\varepsilon > 0$. Choose n_j with $\frac{1}{n_j - 1} < \varepsilon$. So since $L \subseteq F \subset E \setminus \bigcup_{k=1}^{j} E_k$ we have

$$\mu(L) \geq -\frac{1}{n_j - 1} \geq -\varepsilon$$

(recall n_j is the **smallest** number in $\mathbb{N}$ such that $\exists\, E_j \subset E \setminus \bigcup_{k=1}^{j-1} E_j$ with $\mu(E_j) < -\frac{1}{n_j} \Rightarrow \mu(A) \geq -\frac{1}{n_j - 1} \quad \forall\, A \subset E \setminus \bigcup_{k=1}^{n_j - 1} E_k$). Since $\varepsilon > 0$ is arbitrary, $\mu(L) \geq 0$.

$\triangledown$

So using our claim $\exists\, F \subseteq E$ with $\nu(F) > 0$ and F being a positive set. But then $F \cup P$ is also positive, so

$$\beta \geq \nu(F \cup P) = \nu(F) + \nu(P) = \nu(F) + \beta.$$

Contradiction as $\nu(F) > 0$. Thus $\nu(E) < 0$ i.e. N is negative. If N', P' is another partition $X = P' \cup N'$, then $P \setminus P' \subset P$ and $P \setminus P' \subset N'$. So $P \setminus P'$ is both positive and negative, hence it is null.

$\square$

Definition 9.3.

(1) Two signed measures μ, ν on $(X, \mathcal{A})$ are **mutually singular** (denoted $\mu \perp \nu$) if $\exists$ partition $X = E \cup F$ with $E \cap F = \emptyset$ and $E, F \in \mathcal{A}$ such that E is null for μ and F is null for ν (μ, ν "live" on disjoint sets).

(2) Let $\nu = f\ d\mu$ given by $\nu(A) = \int_A f\ d\mu$ and $f = \chi_E$ to get restriction $d\mu_E := \chi_E\ d\mu$ to E.

Theorem 9.3 (Jordan Decomposition)**.** Let $(X, \mathcal{A}, \nu)$ be a signed measure space. Then $\exists$ unique positive measures ν^+, ν^- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.

Proof. Let $X = P \cup N$ for the Hahn decomposition of ν. Define $\nu^{\pm} := \mathcal{A} \to [0, \infty]$ by $\nu^+(A) := \nu(A \cap P),\ \nu^-(A) := -\nu(A \cap N) \quad \forall\, A \in \mathcal{A}$. Clearly $\nu = \nu^+ - \nu^-$and $\nu^+ \perp \nu^-$. Given another such decomposition $\nu = \mu^+ - \mu^-$ with $\mu^+ \perp \mu^-$ then $\exists\, E, F \in \mathcal{A}$ with $X = E \cup F,\ E \cap F = \emptyset$ and $\mu^+(F) = 0 = \mu^-(E)$. Then $X = E \cup F$ is another Hahn decomposition for ν. So by Theorem 9.2 $P \Delta E$ is ν-null. Thus

$$\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) \text{ (as } P\Delta E \text{ is } \nu\text{-null)} = \nu^+(A).$$

Likewise $\mu^- = \nu^-$.

$\square$

Definition 9.4. For a signed measure space $(X, \mathcal{A}, \nu)$ we define:

(1) $\mathcal{L}^1(\nu) := \mathcal{L}^1(\nu^+) \cap \mathcal{L}^1(\nu^-)$, $\int f \, d\nu := \int f \, d\nu^+ - \int f \, d\nu^-$ for $f \in \mathcal{L}^1(\nu)$.

(2) the **total variation** of ν is the positive measure $|\nu| := \nu^+ + \nu^-$.

(3) ν is finite (respectively σ-**finite**) if $|\nu|$ is finite (respectively σ-finite).

(4) let μ be a **positive** measure on $(X, \mathcal{A})$, then ν is **absolutely continuous** with respect to μ (written $\nu \ll \mu$) if $\mu(A) = 0$ for $A \in \mathcal{A}$ then $\nu(A) = 0$.

(5) Positive measures λ, μ are (**quasi-**) **equivalent** if $\lambda \ll \mu$ and $\mu \ll \lambda$ denoted $\lambda \sim \mu$.

Exercise. Prove the following:

(1) $A \in \mathcal{A}$ is ν-null if and only if $|\nu|(A) = 0$.

(2) If $\nu \ll \mu$ and $\nu \perp \mu$ then $\nu = 0$.

(3) $\nu \ll \mu$ if and only if $|\nu| \ll \mu$ if and only if $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.

(4) If $f \in \mathcal{L}^1(\mu)$, μ a positive measure and $\nu(A) = \int_A f \, d\mu$ then $\nu \ll \mu$.

Theorem 9.4. Let ν be a σ-finite signed measure on $(X, \mathcal{A})$ and let μ be a σ-finite positive measure on $(X, \mathcal{A})$. Then

(1) (Lebesgue decomposition) There are unique σ-finite signed measures λ, ρ such that $\nu = \lambda + \rho$ with $\lambda \perp \mu$ and $\rho \ll \mu$.

(2) (Radon-Nikodym Theorem) There is a unique $f \in L^1(\mu)$ such that

$$d\rho = f \, d\mu \qquad \text{i.e. } \rho(A) = \int_A f \, d\mu \qquad \forall \, A \in \mathcal{A}.$$

Proof. First assume μ, ν are both finite and positive. Define

$$\mathcal{F} := \left\{ f : X \to [0, \infty] \;\middle|\; f \text{ is measurable and } \int_A f \, d\mu \leq \nu(A) \;\; \forall \, A \in \mathcal{A} \right\}$$

($0 \in \mathcal{F}$, where 0 is the zero function).

Claim. If $f, g \in \mathcal{F}$, then $h := \max(f, g) \in \mathcal{F}$.

Proof. Let $B := \{x \in X \mid f(x) > g(x)\}$ and $h = \chi_B \cdot f + \chi_{B^c} \cdot g$. Then $\forall \, A \in \mathcal{A}$,

$$\int_A h \, d\mu = \int_{A \cap B} f \, d\mu + \int_{A \cap B^c} g \, d\mu \leq \nu(A \cap B) + \nu(A \cap B^c) = \nu(A).$$

Thus $h \in \mathcal{F}$.

$\triangledown$

Let $a = \sup\{\int_X f\ d\mu \mid f \in \mathcal{F}\}$. Note $0 \le a \le \nu(X) < \infty$. Choose a sequence $\{f_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ such that $\int_X f_n\ d\mu \to a$. By replacing f_n by $\max(f_1, \dots, f_n) \in \mathcal{F}$ we may assume that f_n is increasing (still have $\int_X f_n\ d\mu \to a$). Put $f := \lim\limits_{n\to\infty} f_n$ by MCT (Theorem 4.2), $f \in M_+(X)$ and $\int_A f\ d\mu = \lim\limits_{n\to\infty} \int_A f_n\ d\mu \le \nu(A)\ \ \forall A \in \mathcal{A}$. Thus $f \in \mathcal{F}$ and moreover

$$\int_X f\ d\mu = \lim_{n\to\infty} \int_X f_n\ d\mu = a < \infty. \tag{9.1}$$

Since $f \ge 0$, we see $f \in L^1(\mu)$. Define a measure $d\lambda := d\nu - f\ d\mu$ (i.e. $\lambda(A) = \nu(A) - \int_A f\ d\mu\ \ \forall A \in \mathcal{A}$). By definition of $\mathcal{F} \ni f$ we see λ is positive.

Claim. Either $\lambda \perp \mu$ or $\exists\, \varepsilon > 0$ and $E \in \mathcal{A}$ such that $\mu(E) > 0$ and E is a positive set for $\lambda - \varepsilon\mu$.

Proof. $\lambda - \frac{1}{n}\mu$ is a signed measure with Hahn decomposition $X = P_n \cup N_n$ with $P_n \cap N_n = \emptyset$. Define $P := \bigcup\limits_{n=1}^{\infty} P_n$ and $N = P^c = \bigcap\limits_{n=1}^{\infty} N_n$. Then N is a negative set for $\lambda - \frac{1}{n}\mu\ \ \forall n$ i.e.

$$0 \le \lambda(N) \le \frac{1}{n}\mu(N) < \infty \qquad \forall n$$

$\Rightarrow \lambda(N) = 0$. If $\mu(P) = 0$ then $\lambda \perp \mu$. Otherwise $\mu(P) > 0$, hence $\mu(P_n) > 0$ for some n. Now P_n is a positive set for $\lambda - \frac{1}{n}\mu$, so take $E = P_n$ and $\varepsilon = \frac{1}{n}$ to obtain the other alternative.

$\triangledown$

If μ, ν **not** mutually singular $\exists\, \varepsilon > 0$ and $E \in \mathcal{A}$ such that $\mu(E) > 0$ and $\lambda \ge \varepsilon_\mu$ on E. Then $\varepsilon \cdot \chi_E\ d\mu \le \chi_E\ d\lambda \le d\lambda = d\nu - f\ d\mu$ i.e.

$$\int_A (f + \varepsilon \cdot \chi_E)\ d\mu \le \int_A d\nu = \nu(A) \qquad \forall A \in \mathcal{A}$$

$\Rightarrow f + \varepsilon \cdot \chi_E \in \mathcal{F}$ and $\int_X (f + \varepsilon \cdot \chi_E)\ d\mu = a + \varepsilon\mu(E) > a$ by (9.1) contradicting the definition of a. Thus $\lambda \perp \mu$, and $d\rho := f\ d\mu$ (it is easy to see that $\rho \ll \mu$).

We will now prove uniqueness. If $d\nu = d\lambda' + f'\ d\mu$ is another such decomposition, then

$$d\lambda - d\lambda' = (f' - f)\ d\mu. \tag{9.2}$$

But $(\lambda - \lambda') \perp \mu$ [Proof: $\lambda \perp \mu \Rightarrow X = E \cup F$, $\lambda(E) = 0\mu(F)$ and $\lambda' \perp \mu \Rightarrow X = E' \cup F'$, $\lambda'(E') = 0\mu(F')$. Then $(\lambda - \lambda')(E \cap E') = 0$ and $\mu(F \cup F') = 0 \Rightarrow (\lambda - \lambda') \perp \mu$]. Moreover $(f' - f)\ d\mu \ll d\mu$. Thus Equation (9.2) $\Rightarrow d\lambda - d\lambda' = 0 = (f' - f)\ d\mu$ i.e.

$$\int_A (f' - f)\ d\mu = 0 \qquad \forall A.$$

So by Theorem 5.1(3) $f' - f$ μ-a.e.

We proceed to prove this result for the finite case. Let μ, ν be σ-finite i.e. $X = \bigcup\limits_{n=1}^{\infty} X_n = \bigcup\limits_{k=1}^{\infty} Y_k$ with $\mu(X_n) < \infty$, $\nu(Y_k) < \infty\ \ \forall n, k$ and $Z_{n,k} = X_n \cap Y_k$. Relabel $Z_{n,k}$ to obtain Z_j, $j \in \mathbb{N}$ such that

$$X = \bigcup_{j=1}^{\infty} Z_j \qquad \text{with } \mu(Z_j) < \infty \text{ and } \nu(Z_j) < \infty.$$

Define

$$\mu_j(A) := \mu(A \cap Z_j) \qquad \text{and} \qquad \nu_j(A) := \nu(A \cap Z_j) \qquad \forall\, A \in \mathcal{A}$$

to obtain finite positive measures μ_j, ν_j for which the Theorem holds by the proof above i.e. $d\nu_j = d\lambda_j + f_j\ d\mu_j$ where $\lambda_j(Z_j^c) = 0 = f_j \restriction Z_j^c$. Let

$$\lambda = \sum_{j=1}^{\infty} \lambda_j \qquad \text{and} \qquad f = \sum_{j=1}^{\infty} f_j.$$

Then $\lambda \perp \mu$ (check) and $d\nu = d\lambda + f\ d\mu$. As for uniquness the same argument used for the finite case is sufficient.

If ν is a signed measure, apply previous results to ν^+ and ν^- and subtract.

$\square$

Definition 9.5. Let $\nu \ll \mu$ for ν (respectively μ) a signed (respectively positive) measure on $(X, \mathcal{A})$. Then by Theorem 9.4, $d\nu = f\ d\mu$ for unique $f \in L^1(\mu)$. We call f the **Radon-Nikodym derivative** of ν with respect to μ and denote it by $f := \frac{d\nu}{d\mu}$.

Radon Nikodym is very useful. For example it is required to prove that

$$L^p(\mu)^* \cong L^q(\mu) \qquad \text{for } \frac{1}{p} + \frac{1}{q} = 1,\ 1 < p < \infty.$$

Remarks 9.1. We can extend these measures to complex measures on $(X, \mathcal{A})$ by $\mu = \mu_r + i\mu_i$ where r = real part, i = imaginary part and μ_r, μ_i are signed measures which are **finite**. Then Theorem 9.4 also holds when ν is a complex measure and μ is σ-finite positive measure.

Theorem 9.5 (Chain Rule)**.** Let μ, λ, ν be σ-finite measures on $(X, \mathcal{A})$ with ν a signed measure, λ and μ are positive measures such that $\nu \ll \mu$ and $\mu \ll \lambda$. Then

(1) $g\frac{d\nu}{d\mu} \in L^1(\mu)\ \ \forall\, g \in L^1(\nu)$ and

$$\int g\ d\nu = \int g \left(\frac{d\nu}{d\mu}\right)\ d\mu \tag{9.3}$$

(2) $\nu \ll \lambda$ and

$$\frac{d\nu}{d\lambda} = \left(\frac{d\nu}{d\mu}\right)\left(\frac{d\mu}{d\lambda}\right), \qquad \lambda\text{-a.e.}$$

Proof. By considering ν^+ and ν^- separately, we may assume ν is positive. By definition

$$\nu(A) = \int_X \chi_A\ d\nu = \int_A \frac{d\nu}{d\mu}\ d\mu = \int_X \chi_A \cdot \frac{d\nu}{d\mu}\ d\mu \qquad A \in \mathcal{A}.$$

So Equation (9.3) holds for all functions χ_A with $A \in \mathcal{A}$ and $\nu(A) < \infty$. Thus it holds for all positive simple functions in $L^1(\nu)$, hence by MCT for all non-negative functions in $L^1(\nu)$ (see Theorem 3.3) and hence $\forall\, g \in L^1(\nu)$ by linearity. Now,

$$\nu(E) = \int_X \chi_E\ d\nu = \int_E \frac{d\nu}{d\mu}\ d\mu = \int_E \frac{d\nu}{d\mu} \cdot \frac{d\mu}{d\lambda}\ d\lambda \qquad \forall\, E \in \mathcal{A}$$

by applying Equation (9.3) to $\mu \ll \lambda$ setting $g = \chi_E \cdot \frac{d\nu}{d\mu}$. Thus $\nu \ll \lambda$ and

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \cdot \frac{d\mu}{d\lambda} \qquad \lambda\text{-a.e.}$$

$\square$

If $\mu \sim \nu$ and μ, ν are positive then

$$\left(\frac{d\mu}{d\nu}\right) \cdot \left(\frac{d\nu}{d\mu}\right) = 1 \text{ a.e. on } X.$$

On $\mathbb{R}^n$, the Radon-Nikodym derivatives can be calculated explicitly.

Theorem 9.6 (Lebesgue Differentiation Theorem)**.** Let $\mu =$ Lebesgue measure on $\mathbb{R}^n$, let ν be a signed Radon measure on $\mathbb{R}^n$ with $d\nu = d\lambda + f\ d\mu$ its Lebesgue decomposition. Then

$$f(x) = \lim_{r \to 0} \left[\frac{\nu(B_r(x))}{\mu(B_r(x))} \right] \qquad \mu\text{-a.e.}$$

where $B_r(x) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$. In particular, if $\nu \ll \mu$, then

$$\frac{d\nu}{d\mu}(x) = \lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} \qquad \mu\text{-a.e..}$$

Proof. The proof is very technical. See Real Analysis, Modern Techniques and Their applications, Second Edition (1999) - Gerald B. Folland p95-98.

Definition 9.6. Let μ, ν be probability measure on $(X, \mathcal{A})$. Let $\gamma : \mathcal{A} \to [0, \infty]$ be a probability measure such that $\mu \ll \gamma$ and $\nu \ll \gamma$ (these always exist, for example $\gamma = \frac{1}{2}(\mu + \nu)$. Then $\frac{d\mu}{d\gamma}, \frac{d\nu}{d\gamma} \in L^1(\gamma)$ by Theorem 9.4. Hence

$$\text{Hellinger integral } \equiv H(\mu, \nu) := \int_X \left[\frac{d\mu}{d\gamma}(x) \cdot \frac{d\nu}{d\gamma}(x)\right]^{\frac{1}{2}} \ d\gamma(x) < \infty.$$

The following properties about Hellinger integrals can be easily proven and is left as an exercise:

Properties.

(1) $H(\mu, \nu)$ is independent of γ.

(2) If $\mu \sim \nu$ choose $\gamma = \mu$ to get

$$H(\mu, \nu) = \int_X \sqrt{\frac{d\nu}{d\mu}}\ d\mu.$$

(3) Applying Cauchy-Schwartz to the definition we obtain that $0 \leq H(\mu, \nu) \leq 1$.

(4) $H(\mu, \nu) = 1$ if and only if $\mu = \nu$.

(5) $H(\mu, \nu) = 0$ if and only if $\mu \perp \nu$.

(6) Thus $\mu \sim \nu \Rightarrow H(\mu, \nu) > 0$.

Theorem 9.7 (Kakutani 1948)**.** Given σ-algebras $(X_n, \mathcal{A}_n)$ for $n \in \mathbb{N}$. Let μ_n, ν_n be probability measures on $(X_n, \mathcal{A}_n)$ such that $\mu_n \sim \nu_n \ \ \forall n$. Let $\mu = \prod_{n=1}^{\infty} \mu_n$ and $\nu = \prod_{n=1}^{\infty} \nu_n$ on

$$(X, \mathcal{A}) := \left(\prod_{n=1}^{\infty} X_n, \bigotimes_{n=1}^{\infty} \mathcal{A}_n \right)$$

then

(1) If $\prod_{n=1}^{\infty} H(\mu_n, \nu_n) > 0$ then $\mu \sim \nu$ and

$$\frac{d\nu}{d\mu}(x) = \prod_{n=1}^{\infty} \frac{d\nu_n}{d\mu_n} \text{ a.e.}$$

(2) $\prod_{n=1}^{\infty} H(\mu_n, \nu_n) = 0$ if and only if $\mu \perp \nu$. Moreover $H(\mu, \nu) = \prod_{n=1}^{\infty} H(\mu_n, \nu_n)$. (Note that $\prod_{n=1}^{\infty} H(\mu_n, \nu_n)$ always exists as $\prod_{n=1}^{\infty} H(\mu_n, \nu_n)$ is a decreasing sequence).

Proof. Let $\psi_n(x) := \prod_{k=1}^{n} \sqrt{\frac{d\nu_k}{d\mu_k}(x_k)}$ where $x = (x_1, x_2, \dots) \in X = \prod_{k=1}^{\infty} X_k$. Note

$$\psi_n(x) = \underbrace{\varphi(x_1, \dots, x_n)}_{\prod_{k=1}^{n} \sqrt{\frac{d\nu_k}{d\mu_k}(x_k)}} \cdot \underbrace{1(x_{n+1}, \dots)}_{=1 \ \forall (x_{n+1}, \dots)}$$

and $\mu = (\mu_1 \times \dots \times \mu_n) \times \left(\prod_{k=n+1}^{\infty} \mu_k \right)$. So by Fubini's Theorem (Theorem 6.4(2))

$$\begin{aligned}
\|\psi_n\|_{L^2(\mu)}^2 &= \int_X |\psi_n(x)|^2 \, d\mu(x) \\
&= \int_{X_1 \times \dots \times X_n} \varphi(x_1, \dots, x_n) \, d(\mu_1 \times \dots \times \mu_n) \cdot \underbrace{\int_{X_{n+1} \times \dots} 1 \, d \prod_{k=n+1}^{\infty} \mu_k}_{=1} \\
&= \prod_{k=1}^{n} \int_{X_k} \frac{d\nu_k}{d\mu_k}(x_k) \, d\mu_k(x_k) \\
&= \prod_{k=1}^{n} \int_{X_k} d\nu_k = 1.
\end{aligned} \tag{9.4}$$

We show $\{\psi_n\}_{n=1}^{\infty} \subset L^2(\mu)$ is Cauchy (hence has limit $\psi \in L^2(\mu)$. For $m > n$ (by symmetry this choice is arbitrary)

$$
\begin{aligned}
\|\psi_n-\psi_m\|^2_{L^2(\mu)} &= \int_X \left(\prod_{k=1}^{n}\sqrt{\frac{d\nu_k}{d\mu_k}}-\prod_{k=1}^{m}\sqrt{\frac{d\nu_k}{d\mu_k}}\right)^2 \; d\mu \\
&= \int_X \prod_{k=1}^{n}\frac{d\nu_k}{d\mu_k}\left(1-\prod_{k=n+1}^{m}\sqrt{\frac{d\nu_k}{d\mu_k}}\right)^2 \; d\mu \\
&= \int_X \left(\underbrace{\prod_{k=1}^{n}\frac{d\nu_k}{d\mu_k}}_{|\psi_n|^2}+\underbrace{\prod_{k=1}^{m}\frac{d\nu_k}{d\mu_k}}_{|\psi_m|^2}-2\underbrace{\prod_{k=1}^{n}\frac{d\nu_k}{d\mu_k}}_{|\psi_n|^2}\prod_{l=n+1}^{m}\sqrt{\frac{d\nu_l}{d\mu_l}}\right) \; d\mu \\
&= 2\left(1-\prod_{k=n+1}^{m}\int_X\sqrt{\frac{d\nu_k}{d\mu_k}}\right) \qquad \text{by Equation (9.4)} \\
&= 2\left(1-\prod_{k=n+1}^{m}H(\mu_k,\nu_k)\right). \qquad (9.5)
\end{aligned}
$$

(1) If $\prod_{n=1}^{\infty} H(\mu_n,\nu_n)>0$ then

$$\lim_{m\to\infty}\lim_{n\to\infty}\prod_{k=n+1}^{m}H(\mu_k,\nu_k)=1.$$

Thus Equation (9.5) $\Rightarrow \{\psi_n\}_{n=1}^{\infty}$ is Cauchy. Note that by Cauchy Schwartz we have

$$\|\psi_m+\psi_n\|^2_{L^2(\mu)} \le (\|\psi_n\|-\|\psi_m\|)^2=4. \qquad (9.6)$$

Let $\psi := \lim_{n\to\infty}\psi_n \in L^2(\mu)$. Then

$$
\begin{aligned}
\left(\int_X |\psi_m^2-\psi_n^2| \; d\mu\right)^2 &= \left(\int_X |\psi_m+\psi_n|\cdot|\psi_m-\psi_n| \; d\mu\right)^2 \\
&\le \underbrace{\int_X |\psi_m+\psi_n|^2 \; d\mu}_{\|\psi_m+\psi_n\|^2_{L^2(\mu)}}\cdot\int_X |\psi_m-\psi_n|^2 \; d\mu \\
&\le 4\int_X |\psi_m-\psi_n|^2 \; d\mu \qquad \text{by Inequality (9.6)} \\
&= 8\left(1-\prod_{k=n+1}^{m}H(\mu_k,\nu_k)\right) \qquad \text{by Equation (9.5).}
\end{aligned}
$$

So $\left(\int_X |\psi_m^2-\psi_n^2| \; d\mu\right)^2 \xrightarrow[\infty]{n,m} 0$. Thus $\lim_{m\to\infty}\lim_{n\to\infty}\|\psi_n^2-\psi_m^2\|_{L_1(\mu)}=0 \Rightarrow \psi_n^2$ is Cauchy in $L^1(\mu) \Rightarrow \;\; \exists\,\xi := \lim_{n\to\infty}\psi_n^2$ in $L^1(\mu)$. Now $(\psi-\psi_n)^2=\psi^2+\psi_n^2-2\psi\psi_n$ where $(\psi-\psi_n)^2 \xrightarrow{L^1(\mu)} 0$, $\psi_n^2 \xrightarrow{L^1(\mu)} \xi$ and $\psi\psi_n \xrightarrow{L^1(\mu)} \psi^2$. Thus $\xi=\psi^2$ i.e.

$$\lim_{n\to\infty}\|\psi_n^2-\psi^2\|_{L^2(\mu)}=0. \qquad (9.7)$$

Let $A \in \mathcal{A}$ and $P_n x := (x_1, \dots, x_n, 0, 0, \dots)$. Then

$$\begin{aligned}
\textstyle\int_X \chi_A \cdot (P_n x)\ d\nu(x) &= \textstyle\int_{X_1 \times \cdots \times X_n} \chi_A(x_1, \dots, x_n, 0, 0, \dots)\ d\nu_1(x_1) \dots d\nu_n(x_n) \\
&= \textstyle\int_{X_1 \times \cdots \times X_n} \chi_A(P_n x) \prod\limits_{k=1}^{n} \frac{d\nu_k}{d\mu_k}(x_k)\ d\mu_1(x_1) \dots d\mu_n(x_n) \\
&= \textstyle\int_X \chi_A(P_n x) \cdot \psi_n^2(x)\ d\mu(x).
\end{aligned}$$

If A is a cyclinder set then $\chi_A \circ P_n$ is eventually constant and $\chi_A \circ P_n \to \chi_A$. For such an A,

$$\int_X \chi_A(P_n x)\ d\nu(x) \xrightarrow[\infty]{n} \int_X \chi_A\ d\nu = \nu(A).$$

As $\psi_n^2 \to \psi^2$ in $L^1(\mu)$ we obtain

$$\int_X \chi_A(P_n x) \cdot \psi_n^2(x)\ d\mu(x) \xrightarrow[\infty]{n} \int_A (\psi(x))^2\ dx.$$

Thus $\nu(A) = \int_A (\psi(x))^2\ d\mu(x)\ \ \forall$ cyclinders A, hence $\forall\, A \in \mathcal{A}$. Thus $\nu \ll \mu$. Likewise $\mu \ll \nu$, hence $\mu \sim \nu$ and $\psi^2 = \frac{d\nu}{d\mu}$ a.e.

As $\mu \sim \nu$ we obtain

$$\begin{aligned}
H(\mu, \nu) &= \textstyle\int_X \psi\ d\mu \\
&= \lim_{n \to \infty} \textstyle\int_X \psi_n\ d\mu \\
&= \lim_{n \to \infty} \prod\limits_{k=1}^{n} \textstyle\int_{X_k} \sqrt{\frac{d\nu_k}{d\mu_k}}\ d\mu_k \\
&= \lim_{n \to \infty} \prod\limits_{k=1}^{n} H(\mu_k, \nu_k).
\end{aligned}$$

(2) Let $\prod\limits_{k=1}^{\infty} H(\mu_k, \nu_k) = 0$. Thus for each $N \in \mathbb{N}\ \ \exists\, n \in \mathbb{N}$ such that $\prod\limits_{k=1}^{n} H(\mu_k, \nu_k) < \frac{1}{N}$ (this is also true for all $m > n$). Let

$$A_n := \left\{ (x_1, \dots, x_n) \in \prod_{k=1}^{n} X_k \ \middle|\ \prod_{k=1}^{n} \frac{d\nu_k}{d\mu_k} \in \bigotimes_{k=1}^{n} \mathcal{A}_k \text{ such that } \prod_{k=1}^{n} \frac{d\nu_k}{d\mu_k}(x) > 1 \right\}.$$

Then

$$\begin{aligned}
(\mu_1 \times \cdots \times \mu_n) A_n &= \textstyle\int_{A_n} 1\ d\mu_1 \dots d\mu_n \\
&< \textstyle\int_{A_n} \prod\limits_{k=1}^{n} \sqrt{\frac{d\nu_k}{d\mu_k}}\ d\mu_1 \dots d\mu_n \\
&= \prod\limits_{k=1}^{n} H(\mu_k, \nu_k) \\
&< \tfrac{1}{N}.
\end{aligned}$$

The same calculation gives

$$(\nu_1 \times \cdots \times \nu_n) \left(\prod_{k=1}^{n} X_k \setminus A_n \right) < \frac{1}{N}.$$

Let $A_{(N)} = A_n \times \prod\limits_{k=1}^{n} X_k$ then $\mu(A_{(N)}) < \frac{1}{N}$ and $\nu(X \setminus A_{(N)}) < \frac{1}{N}$. Since $\exists\, B \in \mathcal{A}$ such that $\mu(B) = 0 = \nu(X \setminus B)$ (Exercise) we have $\mu \perp \nu$.

Conversely, let $\mu \perp \nu \Rightarrow \exists\, A \in \mathcal{A}$ such that $\mu(A) = 0 = \nu(X \setminus A)$. Thus

$$
\begin{aligned}
H(\mu,\nu) &= \int_A \sqrt{\tfrac{d\mu}{d\gamma} \cdot \tfrac{d\nu}{d\gamma}}\; d\gamma(x) + \int_{X\setminus A} \sqrt{\tfrac{d\mu}{d\gamma} \cdot \tfrac{d\nu}{d\gamma}}\; d\gamma(x) \\
&\leq \left(\int_A \tfrac{d\mu}{d\gamma}\; d\gamma\right)^{\frac{1}{2}} \left(\int_A \tfrac{d\nu}{d\gamma}\; d\gamma\right)^{\frac{1}{2}} + \left(\int_{X\setminus A} \tfrac{d\mu}{d\gamma}\; d\gamma\right)^{\frac{1}{2}} \left(\int_{X\setminus A} \tfrac{d\nu}{d\gamma}\; d\gamma\right)^{\frac{1}{2}} \\
&= (\mu(A))^{\frac{1}{2}}(\nu(A))^{\frac{1}{2}} + (\mu(X \setminus A))^{\frac{1}{2}}(\nu(X \setminus A))^{\frac{1}{2}} \\
&= 0
\end{aligned}
$$

□

Chapter 10

Transformations and Isomorphisms of Measure Spaces

Definition 10.1. Let $(X, \mathcal{A}), (Y, \mathcal{B})$ be σ-algebras and let $T : X \to Y$ be $\mathcal{A} - \mathcal{B}$-measurable map. Let $\mu : \mathcal{A} \to [0, \infty]$ be $\mathcal{A}$-measurable and define $\mu_T := \mu \circ T^{-1} : \mathcal{B} \to [0, \infty]$ i.e. $\mu_T(B) := \mu(T^{-1}(B)) \ \ \forall B \in \mathcal{B}$. Then $T^{-1}(\mathcal{B}) = \{T^{-1}(B) \mid B \in \mathcal{B}\} \subseteq \mathcal{A}$ is a σ-algebra and μ_T is a measure, as T^{-1} commutes with unions and intersections. Say $\mu_T \equiv$ **image** of μ under T.

Theorem 10.1. With notation above, if $f : Y \to [-\infty, \infty]$ is measurable then

$$\int_Y f \ d\mu_T = \int_X (f \circ T) \ d\mu \tag{10.1}$$

whenever either side is defined.

Proof. If $f = \chi_B$ for $B \in \mathcal{B}$, then

$$\begin{aligned} \textstyle\int_Y \chi_B \ d\mu_T &= \mu_T(B) \\ &= \mu(T^{-1}(B)) \\ &= \textstyle\int_X \chi_{T^{-1}(B)}(x) \ d\mu(x) \\ &= \textstyle\int_X (\chi_B \circ T)(x) \ d\mu(x). \end{aligned}$$

Thus Equation (10.1) holds for all measurable characteristic functions, hence for all measurable simple functions and hence for all measurable functions by MCT, limits, etc.

□

Thus we can define measures on sets through maps from measure spaces.

Corollary 10.2. Given σ-finite measure spaces $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ and a measure space $T : X \to Y$ such that $\nu \ll \mu_T = \mu \circ T^{-1}$ (i.e. $\mu(A) = 0 \Rightarrow \nu(T(A)) = 0$). Then $\exists\, \varphi \in L^1_+(\nu)$ such that

$$\int_A f(y) \ d\nu(y) = \int_{T^{-1}(A)} f(T(x))\varphi(T(x)) \ d\mu(x) \qquad \forall\, A \in \mathcal{B}$$

and $f \in L^1(\nu)$ such that $f \circ T \in L^1(\mu)$.

Proof. As μ is finite, so is μ_T. So

$$\begin{aligned}
\int_Y f(y)\ d\nu(y) &= \int_Y f(y)\tfrac{d\nu}{d\mu_T}(y)\ d\mu_T(y) && \text{by Theorem 9.4} \\
&= \int_X f(T(x))\tfrac{d\nu}{d\mu_T}(T(x))\ d\mu(x) && \text{by Theorem 10.1.}
\end{aligned}$$

Let $\varphi(y) = \frac{d\nu}{d\mu_T}(y)$ and replace f by $\chi_A \cdot f$ to get the claim.

$\square$

Example 10.1. If $X = Y = \mathbb{R}^n$, T a diffeomorphism and $A \subset \mathbb{R}^n$ a Jordan measurable set, then Riemann integrals give

$$\int_{T(A)} f(y)\ d^n y = \int_A f(T(x)) \cdot |JT(x)|\ d^n x$$

where $JT =$ Jacobian of $T : \mathbb{R}^n \to \mathbb{R}^n$. Thus converting to Lebesgue measure (for f Riemann integrable) $\varphi = (JT) \circ T^{-1}$.

Definition 10.2. Two measures spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are **isomorphic** if $\exists$ bijection $T : X \to Y$ such that $T(\mathcal{A}) = \mathcal{B}$ and $\mu \circ T^{-1} = \nu$. Likewise two measurable spaces $(X, \mathcal{A}), (Y, \mathcal{B})$ are isomorphic if $\exists$ bijection $T : X \to Y$ such that $T(\mathcal{A}) = \mathcal{B}$.

If $\mathcal{A}, \mathcal{B}$ are Borel σ-algebras then we say an isomorphism $T : X \to Y$ is a Borel **isomorphism**.

We want to study a large convenient class of measurable spaces and their Borel isomorphisms.

Definition 10.3. A **Polish** space (X, τ) is a separable topological space which can be topologised by a **complete metric**.

Many Polish spaces do not have "natural" or "simple" metrics.

Example 10.2.

(1) $\mathbb{R}^n$ is a Polish space $\forall\, n$.

(2) Any Banach space (for example $L^2(\mathbb{R})$) is Polish.

(3) $\{0,1\}^{\mathbb{N}} = \prod_{i=1}^{\infty} \{0,1\}$ with product topology is Polish where the metric is

$$d(a,b) = \sum_{n=1}^{\infty} \frac{1}{2^n} |a_n - b_n| \qquad \text{with } a = (a_1, a_2, \dots).$$

(4) $\mathbb{Q} \subset \mathbb{R}$ with relative topology is **NOT** Polish. (By Baire Category Theorem, if it is Polish it is not a countable union of nowhere dense sets, but points in $\mathbb{Q}$ are nowhere dense and $\mathbb{Q}$ are countable so $\mathbb{Q}$ is not Polish).

(5) Any LCH space which is 2^{nd} countable is Polish.

(6) $(0,1)$ is Polish, but not complete with respect to natural metric.

Theorem 10.3. Let (X, τ) be Polish. Then

(1) Each closed subset and each open subset of X is Polish.

(2) A disjoint union of finite or infinite **sequences** of Polish spaces is Polish. (Given Polish spaces (X_i, τ_i) topologise the $\dot{\bigcup}_{n\in\mathbb{N}} X_n = X$ by $S \in \tau$ if and only if $S \cap X_i \in \tau_i \ \ \forall\, i$).

(3) Product of countably many Polish spaces is Polish.

Proof.

(1) Any closed subset of a complete metric space is complete. It follows that a closed subspace of a Polish space is Polish. Let $S \subset X$ be open and let d be a complete metric of X. Let
$$d_0(x, y) := d(x, y) + \left|\frac{1}{d(x, S^c)} - \frac{1}{d(y, S^c)}\right|.$$
Check this is a metric. We show this is equivalent to d on S. Since
$$|d(x, S^c) - d(y, S^c)| \leq d(x, y)$$
$x \mapsto d(x, S^c)$ is continuous. Thus $x_n \to x$ in S with respect to d if and only if $x_n \to x$ with respect to d_0. Thus $d \restriction S$ is equivalent to d_0,

Let $\{x_n\}_{n=1}^{\infty} \subset S$ be Cauchy with respect to $d_0 \Rightarrow \{x_n\}_{n=1}^{\infty}$ is Cauchy with respect to d, so x_n converges to $x \in X$ (as d is complete on X). Now $x \in S$ or else $\lim_{n\to\infty} d(x_n, S^c) = 0$, which implies $\limsup_{n,m} d_0(x_n, x_m) = \infty$ contradicting the assumption that $\{x_n\}_{n=1}^{\infty}$ is Cauchy with respect to d_0. Thus $x_n \to x \in S$ with respect to $d_0 \Rightarrow S$ is complete with respect to d_0. So S is Polish.

For (2) and (3) let $X_1, X_2, \ldots$ be Polish spaces, d_n be a complete metric for X_n and $D_n \subset X_n$ be a countable dense set.

(2) Then
$$\dot{\bigcup_{n\in\mathbb{N}}} D_n \subset \dot{\bigcup_{n\in\mathbb{N}}} X_n = X$$
is countable, dense and
$$d(x, y) := \begin{cases} d_n(x, y) & \text{if } x, y \in X_n \\ 1 & \text{otherwise} \end{cases}$$
is a complete metric for X.

(3) Assume $x, y \in \prod_{n=1}^{\infty} X_n$ with coordinates x_n, y_n then
$$d(x, y) = \sum_{n=1}^{\infty} \frac{d_n(x_n, y_n)}{2^n}$$
is a metric on $\prod_{n=1}^{\infty} X_n$ which metrizes the Tychonoff topology such that it is a complete metric.

$\square$

Theorem 10.4 (Urysohn + Alexandroff). Every Polish space is homeomorphic to a G_δ-subset of $[0,1]^{\mathbb{N}}$.

Proof (Sketch). The homeomorphism $\varphi : X \to [0,1]^{\mathbb{N}}$ is obtained by: Let d be a complete metric on X and $\{x_n\}_{n=1}^{\infty} \subset X$ be dense, then

$$\varphi(x) := (a_1, a_2, \dots),\ a_i \in [0,1] \qquad \text{and} \qquad a_n := \min(1, d(x, x_n)).$$

For full proof refer to D. L. Cohn: measure theory, Birkhauser 1993.

$\square$

Definition 10.4. A **standard** Borel space $(X, \mathcal{B}(X))$ is a pair where $X =$ Polish space and $\mathcal{B}(X) =$ Borel σ-algebra. A standard measure space is a measure space $(X, \mathcal{A}, \mu)$ such that $(X, \mathcal{A}) =$ standard Borel space.

Remarks 10.1. Any finite Borel measure on a Polish space is regular (= Radon).

Theorem 10.5 (Kuratowski's Isomorphism Theorem). Let $(X, \mathcal{B}(X)), (Y, \mathcal{B}(Y))$ be standard Borel spaces. Then they are Borel isomorphic if and only if they have the same cardinality.

- If $\text{card}(X) = n < \infty$ then X is Borel isomorphic to $(\{1, 2, \dots, n\}, \mathcal{P}(\mathbb{N}))$.
- If $\text{card}(X) = \text{card}(\mathbb{N})$ then X is Borel isomorphic to $(\mathbb{N}, \mathcal{P}(\mathbb{N})$.
- If $\text{card}(X) > \text{card}(\mathbb{N})$ then $\text{card}(X) = \text{card}(\mathbb{R})$ and X is Borel isomorphic to $(\{0,1\}^{\mathbb{N}}, \mathcal{B}(\{0,1\}^{\mathbb{N}}))$.

Proof. If $(X, \mathcal{B}(X), (Y, \mathcal{B}(Y))$ are isomorphic, then they have the same cardinality. So we only need to prove the converse. If $\text{card}(X) = \text{card}(Y) \leq \text{card}(\mathbb{N})$ then each subset if Borel (as points are closed and there are countably many, unions of points are Borel) i.e. $\mathcal{B}(X) = \mathcal{P}(X)$, so X, Y is Borel isomorphic.

If $\text{card}(X) = \text{card}(Y) > \text{card}(\mathbb{N})$. We provide a sketch proof for this case:

Show $[0,1]$ is Borel isomorphic to a Borel subset of $\{0,1\}^{\mathbb{N}}$ via binary expansions. Thus $[0,1]^{\mathbb{N}}$ is Borel isomorphic to a subset of $(\{0,1\}^{\mathbb{N}})^{\mathbb{N}}$. This is isomorphic to $\{0,1\}^{\mathbb{N}}$. So $[0,1]^{\mathbb{N}}$ is Borel isomorphic to $\{0,1\}^{\mathbb{N}}$. By Urysohn + Alexdandroff (Theorem 10.4), the Polish space X is homeomorphic to a G_δ-subset of $[0,1]^{\mathbb{N}}$, hence it is Borel isomorphic to a subset of $\{0,1\}^{\mathbb{N}}$. Thus $\text{card}(X) \leq \text{card}(\mathbb{R}) = \text{card}(\{0,1\}^{\mathbb{N}})$. Now show $\exists\, S \in \mathcal{B}(X)$ such that S is Borel isomorphic to $\{0,1\}^{\mathbb{N}} \Rightarrow \text{card}(X) = \text{card}(R)$.

Finally show for Polish spaces X, Y that if $\exists\, S_0 \in \mathcal{B}(X), S_1 \in \mathcal{B}(Y)$ such that X is Borel isomorphic to $S_1 \subset Y$ and Y is Borel isomorphic to $S_0 \subset X$, then X, Y are Borel isomorphic.

For full proof refer to K. R. Parthasarathy: Probability measures on metric spaces, AMS 1967 - Chapter 2 or D. L. Cohn: measure theory, Birkhauser 1993 - 8.3.6.

$\square$

Chapter 11

Disintegration of Measures

Fubini's Theorem allows us to construct product measures by iterated itegrals: Given σ-finite measure spaces $(X,\mathcal{A},\mu),(Y,\mathcal{B},\nu)$ we have

$$\int_{X\times Y} f\ d(\mu\times\nu) = \int_X \left[\int_Y f(x,y)\ d\nu(y)\right]\ d\mu(x) = \int_Y \left[\int_X f(x,y)\ d\mu(x)\right]\ d\nu(y)$$

for $\mathcal{A}\otimes\mathcal{B}$-measurable funcions f. Conversely, given a measure ρ on $Z = X\times Y$ one can ask is this a product measure. In general this is not the case. Let $d\rho = f\ d(\mu\times\nu)$ then

$$\rho(A) = \int_A f\ d(\mu\times\nu) = \int_{X\times Y} \chi_A\cdot f\ d(\mu\times\nu) = \int_X \left[\int_Y \chi_A(x,y)\cdot f(x,y)\ d\nu(y)\right]\ d\mu(x). \quad (11.1)$$

If we had $\rho = \alpha\times\beta$ then

$$\rho(A) = \int \chi_A\ d(\alpha\times\beta) = \int_X \left[\int_Y \chi_A(x,y)\ d\beta(y)\right]\ d\alpha(x). \quad (11.2)$$

We can find α,β if $f(x,y) = g(x)\cdot h(y)$ in which case $d\alpha = g\ d\mu$ and $d\beta = h\ d\nu$. However not all functions in $L^1(\mu\times\nu)$ are of this type. Moreover, they are not unique. For example

$$f(x,y) = g(x)\cdot h(y) = [\lambda g(x)]\left[\frac{1}{\lambda}h(y)\right] \qquad \text{for } \lambda\neq 0.$$

We need one of the factors to obtain the other. If f is not of this type, fix μ on X then Equation (11.1) gives

$$\rho(A) = \int_X \left[\int_Y \chi_A\cdot f(x,y)\ d\nu(y)\right]\ d\mu(x) = \int_X \left[\int_Y \chi_A(x,y)\ d\nu_x(y)\right]\ d\mu(x)$$

where $d\nu_x(y) = f(x,y)\ d\nu(y)$. That is, we have an x-dependent measure $d\nu_x$ on Y. We may regard $d\nu_x$ as a measure on $\{x\}\times Y = P^{-1}(\{x\})$ where $P : X\times Y\to X$ is a projection. Note that if ν is a probability measure then $\nu = (\mu\times\nu)\circ P^{-1} = (\mu\times\nu)_P$ as

$$(\mu\times\nu)(P^{-1}(A)) = (\mu\times\nu)(A\times Y) = \mu(A)\nu(Y) = \mu(A) \qquad \forall\, A\in\mathcal{A}.$$

Theorem 11.1 (Disintegration Theorem)**.** Let $(Z,\mathcal{A},\nu),(X,\mathcal{B},\mu)$ be standard probability spaces, and let $P : Z\to X$ be a measurable map such that $\mu = \nu_p = \nu\circ P^{-1}$. Then $\exists\, X_0\in\mathcal{B}$ such that $\mu(X_0) = 1$ and a map $X\mapsto$ probability measure on $(Z,\mathcal{A})$ denoted $x\mapsto\nu_x$ which is measurable in that $x\mapsto\nu_x(A)$ is measurable $\forall\, A\in\mathcal{A}$ and such that

(1) $\nu_x(P^{-1}(\{x\})) = 1 \ \ \forall\, x \in X_0$ $(\Rightarrow \nu_x(P^{-1}(\{x\})^c) = 0)$.

(2) $\nu(A \cap P^{-1}(B)) = \int_B \nu_x(A)\ d\mu(x) \ \ \forall\, A \in \mathcal{A}, B \in \mathcal{B}$.

(3) $\nu(E) = \int \nu_x(E)\ d\mu(x) \ \ \forall$ Borel sets E.

(4) $x \mapsto \nu_x$ is unique μ-a.e.

Proof. Note Theorem 11.1(2) $\Rightarrow$ Theorem 11.1(3) by setting $B = X$. Let $f \in L^1_+(\nu)$ and define $\nu_f : \mathcal{B} \to [0, \infty)$ by

$$\nu_f(B) := \int_{P^{-1}(B)} f\ d\nu \qquad \text{for } B \in \mathcal{B}$$

i.e. $d\nu_f = (f\ d\nu)_P$. Then $\nu_f \ll \mu = \nu_P$. So by Radon-Nikodym Theorem (Theorem 9.4):

$$\frac{d\nu_f}{d\mu} \in L^1(\mu) \qquad \text{and} \qquad \int_B \frac{d\nu_f}{d\mu}\ d\mu = \int_{P^{-1}(B)} f\ d\nu.$$

Let $\nu_x(A) := \frac{d\nu_{\chi_A}}{d\mu}(x)$, $A \in \mathcal{A}$ (μ-a.e.). Then

$$\int_B \nu_x(A)\ d\mu(x) = \int_{P^{-1}(B)} \chi_A\ d\nu = \nu(A \cap P^{-1}(B)). \tag{11.3}$$

Let $A_1, A_2, \dots \in \mathcal{A}$ be disjoint. Now $\chi_{\bigcup\limits_{k=1}^{\infty} A_k} = \sum\limits_{k=1}^{\infty} \chi_{A_k}$, so

$$\begin{aligned}
\nu_{\chi_{\bigcup\limits_{k=1}^{\infty} A_k}}(B) &= \int_{P^{-1}(B)} \chi_{\bigcup\limits_{k=1}^{\infty} A_k}\ d\nu && \text{by definition} \\
&= \int \sum_{k=1}^{\infty} \chi_{A_k}\ d\nu \\
&= \sum_{k=1}^{\infty} \int_{P^{-1}(B)} \chi_{A_k}\ d\nu && \text{by MCT} \\
&= \sum_{k=1}^{\infty} \int_B \frac{d\nu_{\chi_{A_k}}}{d\mu}\ d\mu \\
&= \sum_{k=1}^{\infty} \int_B \nu_x(A_k)\ d\mu(x) \\
&= \int_B \sum_{k=1}^{\infty} \nu_x(A_k)\ d\mu(x) && \text{by MCT.}
\end{aligned}$$

Thus by Equation (11.3)

$$\nu_x\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \nu_x(A_k) \quad \mu\text{-a.e. in } x. \tag{11.4}$$

A problem is that this holds μ-a.e. in X but the set on which it holds depends on A_k. We need to find X_0: If Z is countable then $\mathcal{B}(Z) = \mathcal{P}(Z)$ and μ is given as a sum of point measures, so the statement to prove becomes trivial.

Assume Z is uncountable, so as it is Polish, we have the Borel isomorphism $(Z, \mathcal{A}) \cong (\{0,1\}^{\mathbb{N}}, \mathcal{B}(\{0,1\}^{\mathbb{N}}))$ by Theorem 10.4. Assume we have this measure space. Let

$$\begin{aligned}
\mathcal{C} &= \text{cylinder sets} \\
&= \left\{ C \times \prod_{k\in\mathbb{N},\ k\notin N} \{0,1\} \;\middle|\; C \in \underbrace{\mathcal{B}(\{0,1\}^{\mathbb{N}}) \otimes \cdots \otimes \mathcal{B}(\{0,1\}^{\mathbb{N}})}_{m \text{ times}} \right\} \\
&= \mathcal{P}(\{0,1\}^m) \subset \mathcal{B}(\{0,1\}^{\mathbb{N}})
\end{aligned}$$

where $N = \{n_1, \ldots, n_m\}$ and $\mathcal{A}(\mathcal{C}) = \mathcal{B}(\{0,1\}^{\mathbb{N}})$. So $\operatorname{card}(\mathcal{C}) = \operatorname{card}(\mathbb{N})$. Moreover $D \in \mathcal{C}$ is open **and** compact ($\Rightarrow$ closed).

For each disjoint $D_1, \ldots, D_n \in \mathcal{C}$ by Equation (11.4) $\exists\, X_{D_1,\ldots,D_n} \in \mathcal{B}$ such that

$$\mu(X^c_{D_1,\ldots,D_n}) = 0 \qquad \text{and} \qquad \nu_x\left(\bigcup_{k=1}^n D_k\right) = \sum_{k=1}^n \nu_x(D_k) \qquad \forall\, x \in X_{D_1,\ldots,D_n}.$$

As $\mathcal{C}$ is countable, the set $\{X_{\mathcal{D}} \mid \mathcal{D} = \{D_1, \ldots, D_n\},\ D_i \in \mathcal{C} \text{ for } n \in \mathbb{N} \text{ with } D_i \text{ disjoint}\}$ is countable, so the intersection

$$X_0 = \cap\{X_{\mathcal{D}} \mid \mathcal{D} = \{D_1, \ldots, D_n\},\ D_i \in \mathcal{C} \text{ for } n \in \mathbb{N} \text{ with } D_i \text{ disjoint}\}$$

still satisfies $\mu(X_0^c) = 0$ and

$$\nu_x\left(\bigcup_{k=1}^n D_k\right) = \sum_{k=1}^n \nu_x(D_k) \qquad \forall \text{disjoint } D_i \in \mathcal{C} \text{ and } x \in X_0.$$

For $x \in X_0$ we construct a measure on $\mathcal{B}(\{0,1\}^{\mathbb{N}})$ via proof of Theorem 6.5:

Lemma 11.1.1. ν_x is countably additive on $\mathcal{C}$.

Proof. $D_1, D_2, \dots \in \mathcal{C}$ be disjoint and let $D = \bigcup_{i=1}^{\infty} D_i \in \mathcal{C}$. As $D \in \mathcal{C}$ is compact and D_i are open it follows that all but finitely many D_i are empty. So $D = \bigcup_{k=1}^{n} D_{i_k}$ and obviously

$$\nu_x(D) = \sum_{k=1}^n \nu_x(D_{i_k}) = \sum_{n=1}^{\infty} \nu_x(D_n).$$

$\triangledown$

Define outer measure

$$\nu_x^*(S) := \inf\left\{ \sum_{n=1}^{\infty} \nu_x(D_n) \;\middle|\; S \subset \bigcup_{n=1}^{\infty} D_n \text{ with disjoint } D_n \in \mathcal{C} \right\}.$$

In Lemma 6.5.2 of proof of Theorem 6.5 we only used countable addivitivity of μ on $\mathcal{C}$, so as we have it here by Lemma 11.1.1 we get that $\mathcal{C} \subset \mathcal{M}_{\nu_x^*}$ and $\nu_x(D) = \nu_x^*(D)\ \ \forall\, D \in \mathcal{C}$ also.

Thus $\nu_x^* \restriction \mathcal{B}(\{0,1\}^{\mathbb{N}})$ is a measure uniquely determined on $\mathcal{C}$. Thus Theorem 11.1(2) and (3) follwos from Equation (11.3).

To see that $\nu_x(P^{-1}(\{x\})) = 1$, fix $B \in \mathcal{B}$ (observe that we may replace X_0 by $X_1 \subset X_0$ such that $\mu(X_1) = 1$). So

$$\begin{aligned} \textstyle\int_B \nu_x(P^{-1}(B))\; d\mu(x) &= \nu(P^{-1}(B)) && \text{by Equation (11.3)} \\ &= \mu(B) && \text{since we assumed } \mu = \nu \circ P^{-1} \\ &= \textstyle\int \chi_B \; d\mu. \end{aligned}$$

Thus

$$\nu_x(P^{-1}(B)) = \chi_B(x) \; \mu\text{-a.e.} \tag{11.5}$$

Fix a metric on X (it is Polish) and let $\beta_n \subset \mathcal{B}$ be an increasing sequence of countable partitions of Y such that $\sup\{ \operatorname{diam}(b) \mid b \in \beta_n\} \xrightarrow[\infty]{n} 0$. For $x \in X$ write $x \in b_n(x) \in \beta_n$ to mean that $b_n(x) \in \beta_n$ is the elements of β_n such that $x \in b_n(x)$. Define $X_1 \in \{B \cap X_0 \mid B \in \mathcal{B}\}$ by $\nu_x(P^{-1}(b)) = 1 \;\; \forall\, x \in b \cap X_1, \; b \in \beta_n, \; n \geq 1$ (this can be done by Equation (11.5)). Let $x \in X_1$, then $\{x\} = \bigcap\limits_{n=1}^{\infty} b_n(x)$, so

$$\nu_x(P^{-1}(\{x\}) = \lim_{n \to \infty} \nu_x(P^{-1}(b_n(x))) = 1.$$

□

This generalises to:

Theorem 11.2. Let $(Z, \mathcal{A}, \nu), (X, \mathcal{B}, \mu)$ be standard σ-finite measure spaces. Let $P : Z \to X$ be a measurable function such that $\mu \sim \nu_P = \nu \circ P^{-1}$. Then $\exists$ map μ-a.e. $x \mapsto \nu_x$ where ν_x are σ-finite measure on Z, such that

(1) $x \mapsto \nu_x(E)$ is a Borel map $\forall\, E \in \mathcal{A}$.

(2) $\nu(E) = \int \nu_x(E) \; d\mu(x) \;\; \forall\, E \in \mathcal{A}$.

(3) $x \mapsto \nu_x$ is unique μ-a.e. in x.

(4) $\nu_x(P^{-1}(\{x\})^c) = 0$ μ-a.e. in x.

Proof. Refer to R. C. Fabec: Fundamentals of Infinite Dimensional Representation Theorey, Theorem I.27.

□