Measure Theory Notes

Claudio Landim
Notes taken by Richard K. Yu

October 2025

Contents

1 Introduction: a non-measurable set

Is it possible to define a set function for measure on the real line $\mathbb{R}$ which adheres to our intuition about measure, that is, a function with the following properties

0. The function's domain is every subset of the $\mathbb{R}$, so $\lambda : \mathscr{P}(\mathbb{R}) \to [0, \infty]$
1. $\lambda((a, b]) = b - a$
2. *Translation invariance.* For all $A \subseteq \mathbb{R}$ and for all $x \in \mathbb{R}$, we have $\lambda(A + x) = \lambda(A)$.
3. *Countable additivity.* Let $A = \bigcup_j A_j$ and $A_j \cap A_k = \varnothing$ for $j \neq k$, then
$$\lambda(A) = \sum_{j=1}^{\infty} A_j$$
This is also sometimes called σ-additivity.

Note that monotonicity follows from countable additivity. For subsets $E \subseteq F \subseteq \mathbb{R}$, we have
$$\mu(E) \leq \mu(E) + \mu(F \setminus E) = \mu(F)$$
Today, we show that defining such a function λ with the above properties is not possible. First, define an equivalence relation $x \sim y$ for $x, y \in \mathbb{R}$ by $y - x \in \mathbb{Q}$. The equivalence class is given by
$$[x] = \{y \in \mathbb{R} : y - x \in \mathbb{Q}\}$$
Let Λ represent the set of equivalence classes, ${}^{\mathbb{R}}/_{\sim}$. We will represent the elements of Λ using Greek symbols, $\alpha, \beta, \ldots$ It is clear that Λ is not countable. If not, then the union $\bigcup \Lambda = \mathbb{R}$, but this shows that $\mathbb{R}$ is the countable union of countable sets. Being that such a union is countable and $\mathbb{R}$ is uncountable, this is a contradiction. Now, form the set $\Omega \subseteq \mathbb{R}$ by picking exactly one representative from each element of Λ. We invoke the Axiom of Choice in the construction of Ω. Further, we may assume that $\Omega \subseteq (0, 1)$. For any representative $x \in \Omega$ that is not in $(0, 1)$ initially, we may shift it by the largest integer smaller than it, that it is, $x - \lfloor x \rfloor$. It remains in the same equivalence class after such a shift.

Now, we will prove three properties involving Ω.

First, we claim that for all $p, q \in \mathbb{Q}$, either $\Omega + q = \Omega + p$, or $(\Omega + q) \cap (\Omega + p) = \varnothing$.
Proof. Suppose that the intersection of these sets is empty. Then we are done. Suppose that the intersection of these sets is nonempty. This means we may find some $x \in (\Omega + q) \cap (\Omega + p)$. We may write $x = a + q = b + p$ for some elements $a, b \in \Omega$. But then, we see that $a - b = p - q \in \mathbb{Q}$, which means that a, b are elements of the same equivalence class. We

cannot have $a \neq b$, owing to the construction of Ω as a set containing exactly one representative from each equivalence class. Hence $a = b$ which implies $p = q$ whence $\Omega + p = \Omega + q$. $\square$

So, if $q \neq p$, then $(\Omega + q) \cap (\Omega + p) = \varnothing$. Now, denoting $R = \mathbb{Q} \cap [-1, 1]$, observe that

$$\lambda\left(\bigsqcup_{r \in R} (\Omega + r)\right) = \sum_{r \in R} \lambda(\Omega + r) \leq \lambda([-1, 2]) = 3$$

By translation invariance $\sum_{r \in R} \lambda(\Omega) \leq 3$. We must have $\lambda(\Omega) = 0$, otherwise, the above sum would be unbounded. From translation invariance, we have $\lambda(\Omega) = \lambda(\Omega + r)$ for all $r \in R$, and it follows that $\lambda\left(\bigsqcup_{r \in R} (\Omega + r)\right) = 0$. $\square$

Finally, we claim that

$$(0, 1) \subseteq \bigsqcup_{r \in R} (\Omega + r)$$

Pick $x \in (0, 1)$. By construction, there exists some $\alpha \in \Omega$ so that $\alpha \in [x]$. So, $\alpha \in (0, 1)$. Observe that $x - \alpha = q \in \mathbb{Q}$ and that $q \in [-1, 1]$. Additionally, $x = \alpha + q \in \Omega + q$, so that $x \in \bigsqcup_{r \in R} (\Omega + r)$. Thus, the claimed subset relation follows. $\square$

But now

$$1 = \lambda((0, 1)) \leq \lambda\left(\bigsqcup_{r \in R} (\Omega + r)\right) = 0$$

This is a contradiction. So, such a function satisfying the four properties above is not possible. $\blacksquare$

To construct such a function, we must loosen one of the properties, it turns out the appropriate property to loosen is Property 0. We must accept that there are some subsets of $\mathbb{R}$ to which we cannot associate a measure. The goal of the next lectures is to extend a function with the desired properties to the largest possible collection of subsets of $\mathbb{R}$.

2 Classes of Subsets: Semi-algebras, algebras, σ-algebras

Definition. *Semi-algebra*
Let $\mathscr{S} \subseteq \mathscr{P}(\Omega)$. We call $\mathscr{S}$ a semi-algebra if it satisfies

1. $\Omega \in \mathscr{S}$.
2. If $A, B \in \mathscr{S}$, then $A \cap B \in \mathscr{S}$.
3. If $A \in \mathscr{S}$, then there exists some finite sequence of sets $E_1, \dots E_n \in \mathscr{S}$ so that $A^c = \bigsqcup_{i=1}^n E_i$

The motivation for the third property comes from considering that the complement of the interval $(a, b]$ is not an interval, but a union of disjoint intervals.

Example. Take $\Omega = \mathbb{R}$ and consider the collection

$$\{(a, b] \mid a \leq b, a, b \in \overline{\mathbb{R}}\}$$

We can also consider $\Omega = \mathbb{R}^n$ with sets of the form $(a_1, b_1] \times \dots \times (a_n, b_n]$. It may be verified in both cases that these are semi-algebras.

Definition. *Algebra*
The collection $\mathcal{A} \subseteq \mathscr{P}(\Omega)$ is an algebra if it satisfies

1. $\Omega \in \mathcal{A}$,
2. $A, B \in \mathcal{A} \implies A \cap B = \mathcal{A}$.
3. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$

Observe that it is the last condition that is stronger than that of a semi-algebra. Alternatively, equivalent conditions are that $\mathcal{A}$ are nonempty, closed under complements, and closed under finite unions.
Remark. If $\mathcal{A}$ is an algebra, then it is a semi-algebra.

Definition. *σ-algebra*
We will call a set $\mathcal{F} \subseteq \mathscr{P}(\Omega)$ a σ-algebra if it satisfies the following conditions

1. $\Omega \in \mathcal{F}$.
2. If $\{A_j\}_{j\in\mathbb{N}}$ is a family of sets where each set belongs to $\mathcal{F}$, then $\bigcap_{j\in\mathbb{N}} A_j \in \mathcal{F}$.
3. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

Remark. As before, if a set is a σ-algebra, then it is an algebra (and also a semi-algebra). In these three definitions we have a set of increasingly restrictive conditions.
Observation 1. Let Ω be a set and let $\mathcal{A}_\alpha \subseteq \mathscr{P}(\Omega)$ be such that $\mathcal{A}_\alpha$ is an algebra for all $\alpha \in I$, for an index set I. Then $\mathcal{A} = \bigcap_{\alpha\in I} \mathcal{A}_\alpha$ is an algebra. Verifying this is routine.
Observation 2. The same fact above holds for the intersection of σ-algebras of the same set. These observations are important because we want to speak of the σ-algebra generated by some collection of sets.

Let Ω be a set and $\mathscr{C} \subseteq \mathscr{P}(\Omega)$. To construct the algebra generated by $\mathscr{C}$, we want an algebra $\mathcal{A}$ with the following property: if $\mathcal{B} \supseteq \mathscr{C}$ and $\mathcal{B}$ is an algebra, then $\mathcal{B} \supseteq \mathcal{A}$. To achieve this, we may simply take

$$\mathcal{A} = \bigcap_{\mathcal{A}_\alpha \supseteq \mathcal{C}} \mathcal{A}_\alpha$$

This intersection has our desired properties. The same construction applies for the generated σ-algebra. Later, we will show that there is an explicit form for the algebra generated by the semi-algebra. It turns out that all elements of this generated algebra may be written as finite unions of elements in the semi-algebra.
One key difficulty is that there is no such explicit expression for elements in the σ-algebra which is generated by a semi-algebra. This difficulty manifests when we attempt to extend a measure from a semi-algebra to a σ-algebra generated by a semi-algebra.

Lemma. Let Ω be a set and let $\mathscr{S}$ be a semi-algebra with $\mathscr{S} \subseteq \mathscr{P}(\Omega)$. Denote $\mathcal{A}(\mathscr{S})$ as the algebra generated by $\mathscr{S}$. $A \in \mathcal{A}(\mathscr{S})$ if and only if there exist $E_1, \ldots, E_n \in \mathscr{S}$ so that $\bigsqcup_{j=1}^n E_j = A$.

Proof. First, observe that the reverse implication is true. If $E_1, \ldots, E_n \in \mathscr{S}$, then they are also in the algebra generated by $\mathscr{S}$. Since $\mathcal{A}(\mathscr{S})$ is closed under complementation and finite intersection, it is also closed under finite unions. So $A \in \mathcal{A}(\mathscr{S})$.

Conversely, suppose that $A \in \mathcal{A}(\mathscr{S})$. Define the set

$$\mathcal{B} = \left\{ \bigsqcup_{j=1}^{n} F_j \mid F_j \in \mathscr{S} \right\}$$

We claim that $\mathcal{B}$ is an algebra which contains $\mathscr{S}$. It is clear that $\mathcal{B} \supseteq \mathscr{S}$: we may write any element in $\mathscr{S}$ as a disjoint union of a single set in $\mathscr{S}$. We verify that $\mathcal{B}$ is an algebra.

1. $\Omega \in \mathcal{B}$ since $\Omega \in \mathscr{S}$.

2. Suppose that $A, B \in \mathcal{B}$. We may write $A = \bigsqcup_{j=1}^{n} E_j$ and $B = \bigsqcup_{k=1}^{m} F_k$, where $E_j, F_k \in \mathscr{S}$. Taking the intersection, we have

$$A \cap B = \left(\bigsqcup_{j=1}^{n} E_j \right) \cap \left(\bigsqcup_{k=1}^{m} F_k \right) = \bigsqcup_{j=1}^{n} \bigsqcup_{k=1}^{m} E_j \cap F_k$$

   where we have used the distributivity of set operations. The set still remains a disjoint union because it involves the intersections of disjoint sets. Further, $E_j \cap F_k \in \mathscr{S}$ since $\mathscr{S}$ is closed under intersection. Hence, since we have written $A \cap B$ as a finite, disjoint union of sets in $\mathscr{S}$, we have $A \cap B \in \mathcal{B}$.

3. Suppose that $A \in \mathcal{B}$. Using the same notation as in the previous point, we see that

$$A^c = \left(\bigsqcup_{j=1}^{n} E_j \right)^c = \bigcap_{j=1}^{n} E_j^c$$

   Since $E_j \in \mathscr{S}$, we know that E_j^c may be written as a finite, disjoint union of sets $\bigsqcup_{k_j=1}^{l_j} F_{j,k_j}$ in $\mathscr{S}$. As a result, we see that

$$\bigcap_{j=1}^{n} E_j^c = \bigcap_{j=1}^{n} \bigsqcup_{k_j=1}^{l_j} F_{j,k_j} = \bigsqcup_{k_1=1}^{l_1} \cdots \bigsqcup_{k_n=1}^{l_n} \left(\bigcap_{j=1}^{n} F_{j,k_j} \right)$$

   Since each set $F_{j,k_j} \in \mathscr{S}$, and $\mathscr{S}$ is a semi-algebra which is closed under intersection, we have $\left(\bigcap_{j=1}^{n} F_{j,k_j} \right) \in \mathscr{S}$. Therefore, we have written A^c as a finite disjoint union of elements in $\mathscr{S}$ and we conclude that $A^c \in \mathcal{B}$. The stated conclusion follows.

Definition. *Additive*
Let $\mathscr{C} \subseteq \mathscr{P}(\Omega)$, $\varnothing \in \mathscr{C}$, and $\mu : \mathscr{C} \to [0, \infty]$. We call μ *additive* if

1. $\mu(\varnothing) = 0$

2. If $E_1, \ldots, E_n \in \mathscr{C}$ and $E = \bigsqcup_{j=1}^{n} E_j \in \mathscr{C}$, then

$$\mu(E) = \sum_{j=1}^{n} \mu(E_j)$$

Observation. Assume $A \in \mathscr{C}$ so that $\mu(A) < \infty$. We may write

$$\mu(A) = \mu(A \cup \varnothing) = \mu(A) + \mu(\varnothing) \implies \mu(\varnothing) = 0$$

So, if there is a set where $\mu(A) < \infty$, the first condition is redundant.

Observation. Suppose that $E \subseteq F$, that $E, F, F \setminus E \in \mathscr{C}$, and that μ is an additive function on $\mathscr{C}$. We may always write $F = E \cup (F \setminus E)$. Using additivity, we may write

$$\mu(F) = \mu(E) + \mu(F \setminus E)$$

If $\mu(E)$ is infinite, then $\mu(F) = \infty$ by the equation above. Similarly, if $\mu(E)$ is finite, then we may subtract it from both sides to see that $\mu(F \setminus E) = \mu(F) - \mu(E)$. In both cases, $\mu(E) \leq \mu(F)$.

Example 1. Discrete measure. Let Ω be a set and $\mathscr{C} \subseteq \mathscr{P}(\Omega)$. Let $\{x_j\}_{j\in\mathbb{N}} \subseteq \Omega$ be a sequence and $\{P_j\}_{j\in\mathbb{N}}$ be a sequence of nonnegative numbers. Define

$$\mu(A) = \sum_{j\geq 1} P_j \cdot \mathbb{1}_{x_j \in A}$$

It may be checked that μ is additive.

Definition. *σ-additive*

With the same setting as the definition of additive, μ should be additive and also fulfill the condition that for a countable sequence of disjoint sets $\{E_j\}_{j\in\mathbb{N}} \subseteq \mathscr{C}$ with $\bigsqcup_{j\in\mathbb{N}} E_j \in \mathscr{C}$, we have

$$\mu(E) = \sum_{j\in\mathbb{N}} \mu(E_j)$$

Example 2. Let $\Omega = (0, 1)$ and define

$$\mathscr{C} = \{(a, b] \mid 0 \leq a \leq b < 1\}$$

Define a function $\mu : \mathscr{C} \to [0, \infty]$ by

$$\mu((a, b]) = \begin{cases} \infty & \text{if } a = 0 \\ b - a & \text{otherwise} \end{cases}$$

It may be shown that μ is additive, but not σ-additive.

3 Set functions

Aside. The notation $E_n \uparrow E$ means that $E_n \subseteq E_{n+1}$ for all $n \in \mathbb{N}$ and that $\bigcup_{n\in\mathbb{N}} E_n = E$. Likewise, the notation $E_n \downarrow E$ means that $E_n \supseteq E_{n+1}$ for all $n \in \mathbb{N}$ and that $\bigcap_{n\in\mathbb{N}} E_n = E$.

Definition. *Continuity of set functions*
Let $\mathscr{C} \subseteq \mathscr{P}(\Omega)$ and $\varnothing \in \mathscr{C}$ and $\mu : \mathscr{C} \to [0,\infty]$ and $E \in \mathscr{C}$.

1. We say that μ is continuous from below at E if for all $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathscr{C}$ with $E_n \uparrow E$, we have $\mu(E_n) \xrightarrow[n\to\infty]{} \mu(E)$.
2. We say that μ is continuous from above at E if for all $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathscr{C}$ with $E_n \downarrow E$ and where there exists some $\eta_0 \in \mathbb{N}$ so that $\mu(E_{\eta_0}) < \infty$, we have $\mu(E_n) \xrightarrow[n\to\infty]{} \mu(E)$.

In (2), we actually have $\mu(E_n) \downarrow \mu(E)$ when μ is monotone.

Remark. Why do we impose a finiteness condition in continuity from above? If not, we could consider function λ which returns the length of the interval, $E_n = [n, \infty)$. In this case, we have $\lambda(E_n) = \infty$ for all $n \in \mathbb{N}$ and that $E_n \downarrow \varnothing$. However, $\lim_{n\to\infty} \mu(E_n) = \infty$ and not 0. We require this finiteness condition to avoid this issue.

Lemma. Let $\mathcal{A} \subseteq \mathscr{P}(\Omega)$ be an algebra. Further, let $\mu : \mathcal{A} \to [0,\infty]$ be additive.

1. If μ is σ-additive, then μ is continuous at E, for all $E \in \mathcal{A}$.
2. If μ is continuous from below (at all sets in $\mathcal{A}$), then μ is σ-additive.
3. If μ is continuous from above at $\varnothing$ and that μ is finite, then μ is σ-additive.

Proof.

1. Assume that μ is σ-additive. Suppose that $E \in \mathcal{A}$ and $E_n \uparrow E$. We can make this sequence disjoint by setting $F_1 = E_1$ and $F_n = E_n \setminus E_{n-1}$. By construction, we have $\bigcup_{n\in\mathbb{N}} F_n = \bigcup_{n\in\mathbb{N}} E_n = E$. Further, each of the sets F_n belong to the algebra because they are defined using sets and their complements in the collection $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$. Using σ-additivity and disjointness

$$\mu(E) = \sum_{k=1}^{\infty} \mu(F_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(F_k) = \lim_{n\to\infty} \mu\left(\bigcup_{k=1}^{n} F_k\right) = \lim_{n\to\infty} \mu(E_n)$$

Now, we want to show that μ is continuous from above. Fix $E \in \mathcal{A}$ and let $E_n \in \mathcal{A}$ where $E_n \downarrow E$ and there exists some $\eta_0 \in \mathbb{N}$ so that $\mu(E_{\eta_0}) < \infty$. We want to show $\mu(E_n) \downarrow \mu(E)$. Our idea will be to create an increasing sequence of sets (i.e. we may imagine this as an annulus with a hole that shrinks). Define

$$G_1 = E_{\eta_0} \setminus E_{\eta_0+1}, \quad G_k = E_{\eta_0} \setminus E_{\eta_0+k}$$

Since $E_n \downarrow E$, by its construction, $G_k \uparrow (E_{\eta_0} \setminus E)$. Further, we know that $G_k \in \mathcal{A}$ since it is an intersection of sets (or complements of sets) in $\mathcal{A}$. Hence, applying our previous proof, $\mu(G_k) \uparrow \mu(E_{\eta_0} \setminus E)$. Now, using the finiteness of the measures of these sets, we have

$$\mu(E_{\eta_0} \setminus E) = \lim_k \mu(E_{\eta_0} \setminus E_{\eta_0+k})$$
$$\mu(E_{\eta_0}) - \mu(E) = \mu(E_{\eta_0}) - \lim_k \mu(E_{\eta_0+k})$$

Cancelling terms on both sides yields our desired identity.

2. Suppose that μ is continuous from below. Let $E = \bigsqcup_{k\in\mathbb{N}} E_k$, where $E, E_k \in \mathcal{A}$. Since $\bigsqcup_{k=1}^n E_k \subseteq E$, we see that

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

where we have used the additivity of μ and monotonicity. Since this holds for all $n \in \mathbb{N}$, we may pass to the limit to conclude $\sum_{k\in\mathbb{N}} \mu(E_k) \leq \mu(E)$.
Define a sequence $F_n = \bigcup_{k=1}^n E_k$. Note that $F_n \in \mathcal{A}$ and $F_n \uparrow E$. Since μ is continuous from below, we see that $\mu(F_n) \uparrow \mu(E)$. But by additivity, this means $\sum_{k=1}^n \mu(E_k) \uparrow \mu(E)$ and combined with the inequality from before, we have $\sum_{k\in\mathbb{N}} \mu(E_k) = \mu(E)$.

3. Suppose that μ is continuous from above at $\varnothing$ and that $\mu(\Omega) < \infty$ and $E = \bigsqcup_{k=1}^\infty E_k$ and $E, E_k \in \mathcal{A}$. Define

$$F_n = E \setminus \bigsqcup_{j=1}^{n-1} E_j$$

Then, $F_n \in \mathcal{A}$ and we see that $F_n \downarrow \varnothing$. Further $\infty > \mu(\Omega) \geq \mu(F_1)$. So, using continuity from above and the finiteness of $\mu(F_1)$, we see that $\mu(F_n) \downarrow 0$. We compute

$$\begin{aligned}\mu(E) &= \mu\left(\bigsqcup_{k=1}^n E_k \cup \bigsqcup_{k>n} E_k\right) \\ &= \sum_{k=1}^n \mu(E_k) + \mu(F_n) \xrightarrow[n\to\infty]{} \sum_{k=1}^\infty \mu(E_k) + 0\end{aligned}$$

which shows that μ is σ-additive.

Remark. Example 2 in the previous lecture is an example where the measure is continuous from above at $\varnothing$ and not finite and additive, and as a result, it is not σ-additive.

3.1 Extending the Set Function

We begin our extension now.

Theorem. *Extending an additive function on a semi-algebra*
Let $\mathscr{S} \subseteq \mathscr{P}(\Omega)$ be a semi-algebra. Let $\mu : \mathscr{S} \to [0, \infty]$ be additive. There exists $\nu : \mathcal{A}(\mathscr{S}) \to [0, \infty]$ with the following properties

1. ν is additive.
2. $\nu(A) = \mu(A)$ for all $A \in \mathscr{S}$. This means ν is an extension.
3. The extension ν is unique. To be explicit, let $\mu_1, \mu_2 : \mathcal{A}(\mathscr{S}) \to [0, \infty]$. If $\mu_1(A) = \mu_2(A)$ for all $A \in \mathscr{S}$, then $\mu_1(E) = \mu_2(E)$ for all $E \in \mathcal{A}(\mathscr{S})$.

Proof. By the previous lemma, we know that every set in the generated algebra may be written as a disjoint union of elements in the semi-algebra:

$$A \in \mathcal{A}(\mathscr{S}) \implies A = \bigsqcup_{j=1}^{n} E_j,\ E_j \in \mathscr{S}$$

How do we define our function ν? The natural thing is to use our identity above and the condition that ν must be additive to see that

$$\nu(A) = \sum_{j=1}^{n} \nu(E_j) = \sum_{j=1}^{n} \mu(E_j)$$

where the first identity should follow via additivity and the second identity should follow since ν is an extension. So, we propose the definition

$$\nu(A) := \sum_{j=1}^{n} \mu(E_j)$$

0. We must show that ν is well-defined, since there may be different ways to form a disjoint union of a set A. Assume that A may be written $A = \bigsqcup_{k=1}^{m} F_k$ for $F_k \in \mathscr{S}$. We must prove that $\nu(\bigsqcup_{k=1}^{m} F_k) = \nu(\bigsqcup_{j=1}^{n} E_j)$. Observe that $E_j \subseteq A = \bigsqcup_{k=1}^{m} F_k$. Consequently

$$E_j = E_j \cap \bigsqcup_{k=1}^{m} F_k = \bigsqcup_{k=1}^{m} E_j \cap F_k$$

Since $E_j, F_k \in \mathscr{S}$, so is their intersection. Note that we have written E_j as a disjoint union of sets in $\mathscr{S}$. By the additivity of μ, we know that $\mu(E_j) = \sum_{k=1}^{m} \mu(E_j \cap F_k)$. But now, we may substitute into the definition

$$\nu(A) = \sum_{j=1}^{n} \mu(E_j) = \sum_{j=1}^{n} \sum_{k=1}^{m} \mu(E_j \cap F_k) = \sum_{k=1}^{m} \mu(F_k)$$

where the last equality follows by the symmetricity of the expression (i.e. we may make a similar argument to expand $\mu(F_k)$). As a result, ν is well-defined.

1. We want to show ν is additive. Let A be as before and $B = \bigsqcup_{k=1}^m F_k$ for $F_k \in \mathscr{S}$ and $A \cap B = \varnothing$. Now, since A and B are disjoint, so are subsets of each. Hence $A \sqcup B = \bigsqcup_{j=1}^n E_j \sqcup \bigsqcup_{k=1}^m F_k$. By definition, using this representation of $A \sqcup B$ as a disjoint union of sets in $\mathscr{S}$, we may write
$$\nu(A \sqcup B) = \sum_{j=1}^n \mu(E_j) + \sum_{k=1}^m \mu(F_k) = \nu(A) + \nu(B)$$
2. It follows from our definition of ν that for all $A \in \mathscr{S}$, we have $\nu(A) = \mu(A)$ (each set in $\mathscr{S}$ is a disjoint union of one set). So ν is indeed an extension.
3. We know that ν is unique because we had no choice in how we defined it. However, we make the check here more formally. Suppose $\mu_1, \mu_2 : \mathcal{A}(\mathscr{S}) \to [0, \infty]$ are both additive and $\mu_1(A) = \mu_2(A)$ for all $A \in \mathscr{S}$. Let $B \in \mathcal{A}(\mathscr{S})$. Using the lemma again, we may write $B = \bigsqcup_{j=1}^n E_j$, with $E_j \in \mathscr{S}$. We compute
$$\mu_1(B) = \sum_{j=1}^n \mu_1(E_j) = \sum_{j=1}^n \mu_2(E_j) = \mu_2(B)$$
Thus, the extension ν is unique.

■

Proposition. Using the same setting as the theorem before, if μ is σ-additive on $\mathscr{S}$, then ν is σ-additive on $\mathcal{A}(\mathscr{S})$.

Proof. We assume that μ is σ-additive on $\mathscr{S}$. Let $A = \bigsqcup_{j=1}^\infty A_j$ with $A, A_j \in \mathcal{A}(\mathscr{S})$. We want to show $\nu(A) = \sum_{j=1}^\infty \mu(A_j)$. Since A belongs to the generated algebra, we may write $A = \bigsqcup_{j=1}^n E_j$ for $E_j \in \mathscr{S}$. Similarly, since $A_k \in \mathcal{A}(\mathscr{S})$, we may write
$$A_k = \bigsqcup_{l=1}^{m_k} E_{k,l}$$
where $E_{k,l} \in \mathscr{S}$. We have the following sequence of identities
$$\begin{aligned} E_j &= E_j \cap A = E_j \cap \left(\bigsqcup_{k=1}^\infty A_k \right) \\ &= E_j \cap \left(\bigsqcup_{k=1}^\infty \bigsqcup_{l=1}^{m_k} E_{k,l} \right) = \bigsqcup_{k=1}^\infty \bigsqcup_{l=1}^{m_k} E_j \cap E_{k,l} \end{aligned}$$

So, we have expressed an element of the semi-algebra as a countable, disjoint union of elements of the semi-algebra. Using the σ-additivity of μ, we have

$$\mu(E_j) = \sum_{k=1}^{\infty} \sum_{l=1}^{m_k} \mu(E_j \cap E_{k,l})$$

Also, we know that $\nu(A_k) = \sum_{l=1}^{m_k} \mu(E_{k,l})$. Further $E_{k,l} = E_{k,l} \cap A = \sum_{j=1}^{n} E_{k,l} \cap A_j$. Thus, by the additivity of μ, we see that $\mu(E_{k,l}) = \sum_{j=1}^{n} \mu(E_{k,l} \cap E_j)$. Putting everything together and using substitution

$$\begin{aligned} \nu(A) &= \sum_{j=1}^{n} \mu(E_j) = \sum_{j=1}^{n} \sum_{k=1}^{\infty} \sum_{l=1}^{m_k} \mu(E_j \cap E_{k,l}) \\ &= \sum_{k=1}^{\infty} \sum_{l=1}^{m_k} \mu(E_{k,l}) = \sum_{k=1}^{\infty} \nu(A_k) \end{aligned}$$

■

4 Carathéodory's Theorem

In the previous lesson, we showed how we may extend a function $\mu : \mathscr{S} \to [0,\infty]$ that is additive (σ-additive) on a semi-algebra to a function $\nu : \mathcal{A}(\mathscr{S}) \to [0,\infty]$ which is additive (σ-additive) on the algebra generated by $\mathscr{S}$. The goal of today's lecture is to extend ν to a σ-additive function $\pi : \sigma(\mathscr{S}) \to [0,\infty]$ on the σ-algebra $\sigma(\mathscr{S})$.
Recall that $\mathscr{S} \subseteq \mathscr{P}(\Omega)$, for some set Ω. There are few steps.

1. We begin by considering an *outer measure* $\pi^* : \mathscr{P}(\Omega) \to [0,\infty]$.
2. Then, we will introduce a set $\mathcal{M}$, the set of measurable subsets of Ω. We will show that $\mathcal{M}$ is a σ-algebra which contains the algebra $\mathcal{A}$ (and by definition, the σ-algebra generated by $\mathcal{A}$ as well).
3. Further $\pi^*\big|_{\mathcal{M}}$ is σ-additive and $\pi^*\big|_{\mathcal{A}} = \nu$ (i.e. it is an extension of ν).
4. Finally, we show that this extension of ν is unique under certain conditions.

4.1 Constructing π^*

Let $\mathcal{A}$ be an algebra. To construct $\pi^* : \mathscr{P}(\Omega) \to [0,\infty]$, for any $A \in \Omega$, define

$$\pi^*(A) = \inf\left\{ \sum_{i=1}^{\infty} \nu(E_i) \mid E_i \in \mathcal{A},\, A \subseteq \bigcup_{i\in\mathbb{N}} E_i \right\}$$

Note that $\sum_{i=1}^{\infty} \nu(E_i)$ is well-defined since each $E_i \in \mathcal{A}$. Further π^* is well-defined because the entire set $\Omega \in \mathcal{A}$, so there is always some set which covers A.

Definition. *Outer Measure*
Let $\mu : \mathscr{C} \to [0,\infty]$, $\mathscr{C} \subseteq \mathscr{P}(\Omega)$, and $\varnothing \in \mathscr{C}$. We call μ an outer measure if it satisfies

1. $\mu(\varnothing) = 0$
2. *Monotonicity.* $E \subseteq F$ and $E, F \in \mathscr{C} \implies \mu(E) \leq \mu(F)$.
3. *Countable sub-additivity.* If $E, E_i \in \mathscr{C}$ and $E \subseteq \bigcup_i E_i$, then $\mu(E) \leq \sum_i \mu(E_i)$

Remark. Countable additivity usually implies monotonicity. Countable sub-additivity is not enough to imply monotonicity, so monotonicity must be required as a separate condition.

Proposition. π^* is an outer measure.

Proof.

1. The collection $\{\varnothing\}_{i\in\mathbb{N}}$ is a collection of sets belonging to $\mathcal{A}$ and which covers $\varnothing$. But then
$$\pi^*(\varnothing) \leq \sum_{i=1}^{\infty} \nu(\varnothing) = 0$$
By definition, $\pi^*(A) \geq 0$ for any $A \subseteq \Omega$. Thus, $\pi^*(\varnothing) = 0$.

2. Consider $E \subseteq F$. Any covering of F is a covering of E. Hence, the set $\mathcal{F}$ of coverings of F is a subset of the set $\mathcal{E}$ of coverings of E. But then $\pi^*(E) \leq \pi^*(F)$ by the definition of infimum ($\pi^*(E)$ is a lower bound for the set of coverings of F).

3. Suppose that $E = E_i$ for $E, E_i \in \mathcal{A}$. We may assume that for all $i \in \mathbb{N}$, $\pi^*(E_i) < \infty$. If not, then the inequality of $\pi^*(E)$ and $\sum_i \pi^*(E_i)$ is immediate since the latter value will be infinite.
Fix an $\varepsilon > 0$. Since $\pi^*(E_i)$ is finite, by the definition of infimum, we may find a sequence of sets $\{H_{i,k}\}_{k\in\mathbb{N}}$ which covers E_i and satisfies
$$\pi^*(E_i) \leq \sum_{k\in\mathbb{N}} \nu(H_{i,k}) < \pi^*(E_i) + \frac{\varepsilon}{2^i}$$
But notice that $\{H_{i,k}\}_{(i,k)\in\mathbb{N}^2}$ is a countable covering of E. Hence
$$\pi^*(E) \leq \sum_{i,k} \nu(H_{i,k}) < \sum_i \pi^*(E_i) + \frac{\varepsilon}{2^i} = \sum_i \pi^*(E_i) + \varepsilon$$
Since this holds for every $\varepsilon > 0$, we have obtained our claim that $\pi^*(E) \leq \sum_{i=1}^{\infty} \pi^*(E_i)$

Therefore π^* is an outer measure. ■

4.2 Carathéodory's condition and the Carathéodory σ-algebra

Let $\mathcal{M}$ be a collection of subsets of $\mathscr{P}(\Omega)$. We say $A \in \mathcal{M}$ if for all $E \subseteq \Omega$, we have

$$\pi^*(E) = \pi^*(E \cap A) + \pi^*(E \cap A^c)$$

Remark. We always have $\pi^*(E) \leq \pi^*(E \cap A) + \pi^*(E \cap A^c)$ by the sub-additivity of the outer measure.

Proposition.

1. $\mathcal{M} \supseteq \mathcal{A}$
2. $\mathcal{M}$ is a σ-algebra.

Proof.

1. Let $A \in \mathcal{A}$ and $E \subseteq \Omega$. It suffices to prove $\pi^*(E) \geq \pi^*(E \cap A) + \pi^*(E \cap A^c)$. This inequality holds trivially if $\pi^*(E) = \infty$. Assume that $\pi^*(E)$ is finite. Let $\varepsilon > 0$. By the definition of infimum, we may find a sequence of sets $\{E_i\}_{i\in\mathbb{N}}$ which covers E and where $E_i \in \mathcal{A}$ and
$$\pi^*(E) \leq \sum_{i=1}^{\infty} \nu(E_i) < \pi^*(E) + \varepsilon$$
Since $E_i \in \mathcal{A}$ and $A \in \mathcal{A}$, we have $E_i \cap A \in \mathcal{A}$. Also, $E \cap A \subseteq \bigcup_{i=1}^{\infty}(E_i \cap A)$. Since these sets are a covering of $E \cap A$, we have
$$\pi^*(E \cap A) \leq \sum_{i=1}^{\infty} \nu(E_i \cap A)$$
Similarly, since $A^c \in \mathcal{A}$, it follows that $\pi^*(E \cap A^c) \leq \sum_{i=1}^{\infty} \nu(E_i \cap A^c)$. Summing the expressions and using the finite additivity of ν, we have
$$\pi^*(E \cap A) + \pi^*(E \cap A^c) \leq \sum_{i=1}^{\infty} \nu(E_i \cap A) + \nu(E_i \cap A^c) = \sum_{i=1}^{\infty} \nu(E_i) \leq \pi^*(E) + \varepsilon$$
Since this holds for all $\varepsilon > 0$, the desired inequality follows, and $A \in \mathcal{M}$. Thus, $\mathcal{M} \supseteq \mathcal{A}$, and every set in the algebra $\mathcal{A}$ is measurable.

2. Observe that $\Omega \in \mathcal{M}$ since $\Omega \in \mathcal{A}$, by the first statement. Further, if $A \in \mathcal{M}$, then $A^c \in \mathcal{M}$ because the Carathéodory condition is symmetric with respect to the complement:
$$\pi^*(E) = \pi^*(E \cap A) + \pi^*(E \cap A^c), \quad \forall E \subseteq \Omega$$
Lastly, we want to show that if $A_j \in \mathcal{M}$ for all $j \in \mathbb{N}$, then $\bigcup_j A_j \in \mathcal{M}$ as well. To proceed, we prove this step for finite unions and then pass to countable unions. Suppose that $A, B \in \mathcal{M}$. Applying the condition to E and $E \setminus A$ and then making a substitution in the last equation, we get:
$$\begin{aligned}
\pi^*(E) &= \pi^*(E \cap A) + \pi^*(E \cap A^c) \\
\pi^*(E \setminus A) &= \pi^*((E \setminus A) \cap B) + \pi^*((E \setminus A) \setminus B) \\
&= \pi^*((E \setminus A) \cap B) + \pi^*((E \setminus (A \cup B)) \\
\pi^*(E) &= \pi^*(E \cap A) + \pi^*((E \setminus A) \cap B) + \pi^*((E \setminus (A \cup B))
\end{aligned}$$

By countable sub-additivity of outer measure

$$\pi^*(E \cap (A \cup B)) \leq \pi^*(E \cap A) + \pi^*((E \setminus A) \cap B)$$

To elaborate, we claim that $E \cap (A \cup B) \subseteq (E \cap A) \cup ((E \setminus A) \cap B)$. We may compute

$$[[E \cap (A \cup B)] \cap A] \cup [[E \cap (A \cup B)] \cap A^c] = (E \cap A) \cup ((E \cap A^c)) \cap B)$$

So, the previous inequality holds, and together, we have

$$\begin{aligned} \pi^*(E) &= \pi^*(E \cap A) + \pi^*((E \setminus A) \cap B) + \pi^*((E \setminus (A \cup B)) \\ &\geq \pi^*(E \cap (A \cup B)) + \pi^*(E \cap (A \cup B)^c) \end{aligned}$$

Hence, $\mathcal{M}$ is closed under finite unions.
Assume that $A_j \in \mathcal{M}$ and set $A = \bigcup_j A_j$. We need to show

$$\pi^*(E) \geq \pi^*(E \cap A) + \pi^*(E \cap A^c)$$

Since $\mathcal{M}$ is closed under finite unions, we may write

$$\pi^*(E) \geq \pi^*\left(E \cap \bigcup_{j=1}^{n} A_j\right) + \underbrace{\pi^*\left(E \setminus \bigcup_{j=1}^{n} A_j\right)}_{\geq \pi^*\left(E \setminus \bigcup_{j=1}^{\infty} A_j\right)} \qquad (*)$$

for all $n \in \mathbb{N}$. Notice that $E \setminus \bigcup_{j=1}^{n} A_j \supseteq E \setminus \bigcup_{j=1}^{\infty} A_j$, so by the monotonicity of outer measure, $\pi^*\left(E \setminus \bigcup_{j=1}^{n} A_j\right) \geq \pi^*\left(E \setminus \bigcup_{j=1}^{\infty} A_j\right)$. So, it remains to show the inequality for the remaining summand. Define the sequence of disjoint sets $F_1 = A_1, \dots F_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j, \dots$ By construction, we see that $F_n \in \mathcal{M}$ since these are finite intersections, unions, and complements of sets in $\mathcal{M}$, and $\bigcup_{j=1}^{\infty} A_j = \bigsqcup_{j=1}^{\infty} F_j$. We claim that

$$\pi^*\left(E \cap \bigsqcup_{j=1}^{n} F_j\right) = \sum_{j=1}^{n} \pi^*(E \cap F_j)$$

We proceed by induction. The base case is because we are asserting that something is equal to itself. Assume the inductive hypothesis holds for F_n. Note that since $F_k \in \mathcal{M}$ for all $k \in \mathbb{N}$ (we have shown that $\mathcal{M}$ is an algebra), F_{n+1} satisfies Carathéodory's condition, that is

$$\begin{aligned} \pi^*\left(E \cap \bigsqcup_{j=1}^{n+1} F_j\right) &= \pi^*\left(E \cap \bigsqcup_{j=1}^{n+1} F_j \cap F_{n+1}\right) + \pi^*\left(E \cap \bigsqcup_{j=1}^{n+1} F_j \cap F_{n+1}^c\right) \\ &= \pi^*(E \cap F_{n+1}) + \pi^*\left(E \cap \bigsqcup_{j=1}^{n} F_j\right) \\ &= \sum_{j=1}^{n+1} \pi^*(E \cap F_j) \end{aligned}$$

where the last equality follows by the inductive hypothesis. The disjointness of F_j was important for this inductive argument. Using our claim, we now have

$$\pi^*\left(E \cap \bigcup_{j=1}^{n} A_j\right) = \pi^*\left(E \cap \bigsqcup_{j=1}^{n} F_j\right) = \sum_{j=1}^{n} \pi^*(E \cap F_j)$$

Putting this together with equation $(*)$, we have

$$\begin{aligned}\pi^*(E) &\geq \pi^*\left(E \cap \bigcup_{j=1}^{n} A_j\right) + \pi^*\left(E \setminus \bigcup_{j=1}^{\infty} A_j\right) \\ &= \sum_{j=1}^{n} \pi^*(E \cap F_j) + \pi^*(E \setminus A) \\ &\geq \pi^*(E \cap A) + \pi^*(E \setminus A)\end{aligned}$$

where we pass to the limit and use countable sub-additivity of the outer measure to obtain the last inequality. To be precise, since $\pi^*(E) \geq \sum_{j=1}^{n} \pi^*(E \cap F_j) + \pi^*(E \setminus A)$ for all $n \in \mathbb{N}$, we have

$$\pi^*(E) \geq \lim_{n\to\infty} \sum_{j=1}^{n} \pi(E \cap F_j) + \pi^*(E \setminus A)$$

We use countable sub-additivity to reach the last expression. Therefore, $\mathcal{M}$ is a σ-algebra. From these two assertions, we know that $\mathcal{M} \supseteq \sigma(\mathcal{A})$.

■

4.3 Showing extensions and σ-additivity

For our last step, we want to show that $\pi^* : \mathcal{M} \to [0, \infty]$ is σ-additive and that it extends ν.

Proposition. π^* extends ν.

Proof. To show that π^* extends ν, we must show that for all $A \in \mathcal{A}$, we have $\pi^*(A) = \nu(A)$. The inequality $\pi^*(A) \leq \nu(A)$ follows from considering the covering $\{A, \varnothing, \varnothing, \ldots\}$ and using the definition of the outer measure π^*.

$$\pi^*(A) \leq \sum_{j=1}^{\infty} \nu(E_j) = \nu(A)$$

Now, we must show the reverse inequality. Let $\{E_j\}_{j\in\mathbb{N}}$ be any cover of A with $E_j \in \mathcal{A}$. It suffices to show that $\nu(A) \leq \sum_{j=1}^{\infty} \nu(E_j)$. Form the collection of disjoint sets $\{F_j\}_{j\in\mathbb{N}}$ by setting $F_1 = E_1$ and $F_j = E_j \setminus \bigcup_{i=1}^{j-1} E_i$. Since every set F_j is formed by a finite sequence of set operations (union, intersection, complement) of sets in the algebra $\mathcal{A}$, we have $F_j \in \mathcal{A}$ for all $j \in \mathbb{N}$. Further $\bigsqcup_{j=1}^{\infty} F_j = \bigcup_{j=1}^{\infty} E_j$ by construction, so $A \subseteq \bigsqcup_{j=1}^{\infty} F_j$. From the set inclusion, we know $A = A \cap \bigsqcup_{j=1}^{\infty} F_j$. We use the σ-additivity of ν to obtain

$$\nu(A) = \nu\left(\bigsqcup_{j=1}^{\infty} A \cap F_j\right) = \sum_{j=1}^{\infty} \nu(A \cap F_j) \leq \sum_{j=1}^{\infty} \nu(E_j)$$

where the last inequality follows the from monotonicity of a σ-additive function ($A \cap F_j \subseteq F_j \subseteq E_j$). Since this holds for any cover of A, taking the infimum, we have $\nu(A) \leq \pi^*(A)$. Thus $\nu(A) = \pi^*(A)$ for all $A \in \mathcal{A}$, showing that π^* extends ν. ■

Proposition. π^* is σ-additive on $\mathcal{M}$.

Proof. Let $\{A_j\}_{j\in\mathbb{N}} \subseteq \mathcal{M}$ be a disjoint collection of sets and set $A = \bigcup_{j=1}^{\infty} A_j$. Since π^* is an outer measure, we have

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

In the previous proof that $\mathcal{M}$ is a σ-algebra, we showed via induction that

$$\pi^*\left(E \cap \bigsqcup_{j=1}^{n} F_j\right) = \sum_{j=1}^{n} \pi^*(E \cap F_j)$$

Taking $E = A$ and $F_j = A_j$, we have

$$\pi^*\left(\bigsqcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \pi^*(A_j)$$

But then by monotonicity, we know that $\pi^*\left(\bigcup_{j=1}^{\infty} A_j\right) \geq \pi^*\left(\bigsqcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \pi^*(A_j)$. Taking the limit as $n \to \infty$, we see that $\pi^*(A) \geq \sum_{j=1}^{\infty} \pi^*(A_j)$, which yields the reverse inequality. Hence $\pi^*\big|_{\mathcal{M}}$ is indeed σ-additive. ■

4.4 Uniqueness of π^* on $\sigma(\mathcal{A})$

Definition. *σ-finite*
Let $(\Omega, \mathcal{M}, \mu)$ be a measure space. The set Ω is *σ-finite* if there exists a collection $\{E_j\}_{j\in\mathbb{N}} \subseteq \mathcal{M}$ so that $E_j \uparrow \Omega$ and $\mu(E_j) < \infty$ for all $j \in \mathbb{N}$.

Definition. *Monotone class*
Let Ω be a set and $\mathcal{G} \subseteq \mathcal{P}(\Omega)$. We call $\mathcal{G}$ a *monotone class* if

1. If $A_j \in \mathcal{G}$ and $A_j \subseteq A_{j+1}$ for all $j \in \mathbb{N}$, then $A = \bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$.
2. If $B_j \in \mathcal{G}$ and $B_j \supseteq B_{j+1}$, then $\bigcap_{j=1}^{\infty} B_j \in \mathcal{G}$

Essentially, we are saying that $\mathcal{G}$ is closed under increasing and decreasing sequences of sets in $\mathcal{G}$.

Proposition. If $\{\mathcal{G}_\alpha\}_{\alpha \in I}$ is a collection of monotone classes with $\mathcal{G}_\alpha \subseteq \mathcal{P}(\Omega)$, then $\bigcap_{\alpha \in I} \mathcal{G}_\alpha$ is also a monotone class.

Remark. The above proposition means that it makes sense to define the smallest monotone class generated by some set $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ as

$$\mathcal{G}(\mathcal{C}) = \bigcap_{\mathcal{G}_\alpha \supseteq \mathcal{C}} \mathcal{G}_\alpha$$

Note that this expression is nonempty since $\mathcal{P}(\Omega)$ is a monotone class which contains $\mathcal{C}$.

Lemma. Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ be an algebra. Denote the monotone class generated by $\mathcal{A}$ by $\mathcal{M}(\mathcal{A})$, then

$$\mathcal{M}(\mathcal{A}) = \sigma(\mathcal{A})$$

We will prove this lemma in the next lecture.

Proposition. Suppose that $\mu_1, \mu_2 : \sigma(\mathcal{A}) \to [0, \infty]$ are two σ-additive functions and $\mu_1\big|_{\mathcal{A}} = \mu_2\big|_{\mathcal{A}}$ and Ω is σ-finite with a decomposition that belongs in the algebra $\mathcal{A}$. Then

$$\mu_1 = \mu_2$$

Remark. In the context of the previous proofs, we are considering the extension of $\nu : \mathcal{A} \to [0, \infty]$ to $\pi^*\big|_{\mathcal{M}}$ and claiming that the restriction of this measure to the smaller σ-algebra, $\pi^*\big|_{\sigma(\mathcal{A})}$, is unique when ν is σ-finite. In the wording of Folland, if the premeasure is σ-finite, then the measure which is extended to the generated σ-algebra (via Carathéodory's extension theorem) is unique.

Proof. Since Ω is σ-finite, there exists some $\{E_j\}_{j \in \mathbb{N}} \subseteq \mathcal{A}$ so that $\bigcup_{j \in \mathbb{N}} E_j = \Omega$ and $\mu_1(E_j) < \infty$ for all $j \in \mathbb{N}$. Fix some $n \in \mathbb{N}$ and define

$$\mathcal{B}_n = \{E \in \sigma(\mathcal{A}) \mid \mu_1(E \cap E_n) = \mu_2(E \cap E_n)\}$$

We claim that $\mathcal{B}_n = \sigma(\mathcal{A})$. Clearly, $\mathcal{B}_n \subseteq \sigma(\mathcal{A})$, by the definition. For the reverse inclusion, notice that $\mathcal{B}_n \supseteq \mathcal{A}$. This is because for any set $A \in \mathcal{A}$, we have $A \in \mathcal{A}$ and $E \cap A \in \mathcal{A}$, so $\mu_1(E \cap A) = \mu_2(E \cap A)$ by assumption and so $A \in \mathcal{B}_n$. Additionally, we claim that $\mathcal{B}_n$ is a monotone class. To show this, pick a sequence $\{A_j\}_{j \in \mathbb{N}} \subseteq \mathcal{B}_n$ so that $A_j \subseteq A_{j+1}$ and let $A = \bigcup_{j=1}^{\infty} A_j$. Since $A_j \in \mathcal{B}_n$, we have $\mu_1(A_j \cap E_n) = \mu_2(A_j \cap E_n)$. Since μ_1 and μ_2 are σ-additive, they are continuous from below, and by taking the limit and noting that each term of both sequences are equal

$$\mu_1(A \cap E_n) = \mu_2(A \cap E_n)$$

On the other hand, consider a sequence $\{B_j\}_{n \in \mathbb{N}} \subseteq \mathcal{B}_n$ where $B_n \supseteq B_{n+1}$ and $B = \bigcap_{j \in \mathbb{N}} B_j$. Since $B_j \in \mathcal{B}_n$, we have $\mu_1(B_j \cap E_n) = \mu_2(B_j \cap E_n)$. Also, since $\mu_1(B_j \cap E_n) \leq \mu_1(E_n) < \infty$ and μ_1 is σ-additive, it is continuous from below, meaning that

$$\lim_{j \to \infty} \mu_1(B_j \cap E_n) = \mu_1(B \cap E_n)$$

Since the same argument holds for μ_2, we have $\mu_1(B \cap E_n) = \mu_2(B \cap E_n)$ by comparing the terms of both sequences again. Thus $\mathcal{B}_n$ is a monotone class. Since $\mathcal{B}_n$ is a monotone class and $\mathcal{B}_n \supseteq \mathcal{A}$, we see that $\mathcal{B}_n \supseteq \mathcal{M}(\mathcal{A}) = \sigma(\mathcal{A})$ using the lemma about the equality of a monotone class and a σ-algebra which are both generated by the same algebra. Hence, $\mathcal{B}_n = \sigma(\mathcal{A})$.

Lastly, we wish to show that $\mu_1(A) = \mu_2(A)$ for all $A \in \sigma(\mathcal{A})$. Let $A \in \sigma(\mathcal{A})$. Then, $A \in \mathcal{B}_n$ and we have $\mu_1(A \cap E_n) = \mu_2(A \cap E_n)$. But since $E_n \uparrow \Omega$ and μ_1 and μ_2 are σ-additive, these functions are continuous from below and we may write

$$\mu_1(A) = \lim_{j \to \infty} \mu_1(A \cap E_n) = \lim_{j \to \infty} \mu_2(A \cap E_n) = \mu_2(A)$$

This shows the extension is unique. ■

We conclude with some remarks. The theorem or construction that we proved above is called Carathéodory's theorem. The point is that we are able to extend a function which is σ-additive in an algebra to a function which is σ-additive on the Carathéodory σ-algebra, and that the extension to the generated σ-algebra is unique under the condition of σ-finiteness of Ω (relative to ν and the algebra).

Note that the condition of σ-finiteness is crucial, and that the extension may not be unique without it. Further, note that the Carathéodory σ-algebra $\mathcal{M}$ depends on the outer measure π^*, which in turn depends on the premeasure ν on which we started.

5 Monotone Classes

Recall the definition of a monotone class. As a quick reminder, a monotone class $\mathcal{M} \subseteq \mathscr{P}(\Omega)$ satisfies the conditions below. Let $\{A_j\}_{j\in\mathbb{N}} \subseteq \mathcal{M}$.

1. If $A_j \uparrow A$, then $A \in \mathcal{M}$.
2. If $A_j \downarrow A$, then $A \in \mathcal{M}$

Remark. We make a few observations

1. If $\mathcal{F}$ is a σ-algebra, then it is a monotone class.
2. If $\mathcal{M}_\alpha \subseteq \mathscr{P}(\Omega)$ for all $\alpha \in I$ are all monotone classes, then the intersection $\bigcap_{\alpha\in I} \mathcal{M}_\alpha$ is a monotone class.

Recall that the second observation is what allows us to define the smallest monotone class which contains some set $\mathscr{C} \subseteq \mathscr{P}(\Omega)$. We will call this the monotone class generated by $\mathscr{C}$ and denote it $\mathcal{M}(\mathscr{C})$.

Theorem. *Generated monotone class equals generated σ-algebra*
Let $\mathcal{A} \subseteq \mathscr{P}(\Omega)$ be an algebra. Then

$$\mathcal{M}(\mathcal{A}) = \sigma(\mathcal{A})$$

Proof. Since $\sigma(\mathcal{A})$ is a σ-algebra, it is a monotone class. By definition, it contains $\mathcal{A}$. So, since $\sigma(\mathcal{A})$ is a monotone class containing $\mathcal{A}$, we have $\mathcal{M}(\mathcal{A}) \subseteq \sigma(\mathcal{A})$.
The reverse inclusion is the nontrivial part. By definition, we know that $\mathcal{A} \subseteq \mathcal{M}(\mathcal{A})$. Our strategy will be to show that $\mathcal{M}(\mathcal{A})$ is a σ-algebra. First, we show that $\mathcal{M}(\mathcal{A})$ is an algebra. Let $E \in \mathcal{M}(\mathcal{A})$ and define the set

$$\mathcal{G}(E) = \{F \in \mathcal{M}(\mathcal{A}) \mid (E \setminus F), (E \cap F), (F \setminus E) \in \mathcal{M}(\mathcal{A})\}$$

Claim. If $E \in \mathcal{A}$, then $\mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{A})$. To prove this, we show the two properties below

1. $\mathcal{G}(E) \supseteq \mathcal{A}$.
   Take some $H \in \mathcal{A}$. Since $\mathcal{A}$ is an algebra and we have assumed $E \in \mathcal{A}$, we know that $E \setminus H, E \cap H, H \setminus E$ are elements of $\mathcal{A}$, so they also belong to $\mathcal{M}(\mathcal{A})$. Hence, $H \in \mathcal{G}(E)$.
2. $\mathcal{G}(E)$ is a monotone class.
   Assume that $H_k \uparrow H$ and $\{H_k\}_{k\in\mathbb{N}} \subseteq \mathcal{G}(E)$. From this, we know that $H \in \mathcal{M}(\mathcal{A})$ since $\mathcal{M}(\mathcal{A})$ is a monotone class. By definition, $E \setminus H_k \in \mathcal{M}(\mathcal{A})$ and $(E \setminus H_k) \downarrow (E \setminus H)$. But since $\mathcal{M}(\mathcal{A})$ is a monotone class, it follows that $E \setminus H \in \mathcal{M}(\mathcal{A})$. Now, since $H_k \in \mathcal{G}(E)$, it follows that $E \cap H_k \in \mathcal{M}(\mathcal{A})$. But then $(E \cap H_k) \uparrow (E \cap H)$, and since $\mathcal{M}(\mathcal{A})$ is a

monotone class, $E \cap H \in \mathcal{M}(\mathcal{A})$. Similarly, $(H_k \setminus E) \uparrow (H \setminus E)$ and $H \setminus E \in \mathcal{M}(\mathcal{A})$. We conclude that $H \in \mathcal{G}(E)$ and that $\mathcal{G}(E)$ is closed under increasing sequences. A similar argument shows that when we consider a decreasing sequence $H_k \downarrow H$ with $\{H_k\}_{k \in \mathbb{N}} \subseteq \mathcal{G}(E)$, we also have $H \in \mathcal{G}(E)$. Thus $\mathcal{G}(E)$ is a monotone class as it is closed under increasing and decreasing sequences.

With these two properties above, we conclude that $\mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{A})$. Next, we make an analogous claim which relies on proving two similar properties.
Claim. If $E \in \mathcal{M}(\mathcal{A})$, then $E \in \mathcal{G}(E)$.

1. $\mathcal{G}(E)$ is a monotone class.
   The proof of this item is identical to that of the previous one (the membership of E does not change anything).

2. $\mathcal{G}(E) \supseteq \mathcal{A}$.
   Take $H \in \mathcal{A}$. It is clear that $H \in \mathcal{M}(\mathcal{A})$. But, by the first claim, we know that $\mathcal{G}(H) \supseteq \mathcal{M}(\mathcal{A})$. However, since $E \in \mathcal{M}(\mathcal{A})$ by assumption, then $E \in \mathcal{G}(H)$, which means that $E \setminus H, (E \cap H), H \setminus E \in \mathcal{M}(\mathcal{A})$. Thus, $H \in \mathcal{G}(E)$.

Now we must show that $\mathcal{M}(\mathcal{A})$ is an algebra.

1. We have $\Omega \in \mathcal{M}(\mathcal{A})$ since $\Omega \in \mathcal{A}$.

2. Suppose $E \in \mathcal{M}(\mathcal{A})$. By our first claim above, then $E \in \mathcal{G}(\Omega)$. But this means that $\Omega \setminus E = E^c \in \mathcal{M}(\mathcal{A})$.

3. Let $E, F \in \mathcal{M}(\mathcal{A})$. The second claim shows that $\mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{A})$. Also, $F \in \mathcal{M}(\mathcal{A})$, so $F \in \mathcal{G}(E)$ by the previous subset relation. But then by the definition of $\mathcal{G}(E)$, we have that $E \cap F \in \mathcal{M}(\mathcal{A})$.

Finally, we show that $\mathcal{M}(\mathcal{A})$ is a σ-algebra. It suffices to show that it is closed under countable unions. Let $\{A_j\}_{j \in \mathbb{N}} \subseteq \mathcal{M}(\mathcal{A})$. Since $\mathcal{M}(\mathcal{A})$ is an algebra, we know that $\bigcup_{j=1}^{n} A_j \in \mathcal{M}(\mathcal{A})$ for all $n \in \mathbb{N}$. But since $\mathcal{M}(\mathcal{A})$ is a monotone class, it is closed under increasing sequences, so it follows that $\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}(\mathcal{A})$. Hence $\mathcal{M}(\mathcal{A})$ is closed under countable unions and it is a σ-algebra. Since $\mathcal{M}(\mathcal{A})$ is a σ-algebra containing $\mathcal{A}$, we have $\mathcal{M}(\mathcal{A}) \supseteq \sigma(\mathcal{A})$. This proves our desired equality. ■

6 The Lebesgue Measure

The purpose of this lecture is to use Carathéodory's theorem to construct the Lebesgue measure. The Lebesgue measure is the measure which extends - to the σ-algebra generated by the intervals - the concept of length.

6.1 Constructing a length function on the semi-algebra

Let $\mathscr{S} \subseteq \mathscr{P}(\mathbb{R})$ be given by

$$\mathscr{S} = \{\varnothing, \mathbb{R}, (a,b], (a,\infty), (-\infty, b]\}$$

where $(a,b]$ are the half-open intervals with $a,b \in \mathbb{R}$ and $a < b$, and the other intervals are defined similarly.
Define $\mu : \mathscr{S} \to [0,\infty]$ by

$$\mu((a,b]) = \begin{cases} b-a, & \text{if } a,b \in \mathbb{R},\ a<b \\ 0, & \text{if } a,b \in \mathbb{R},\ a=b \\ \infty & \text{otherwise} \end{cases}$$

Above, we think of $a, b \in \overline{\mathbb{R}}$ and allow the case where $a = b$ so that we can get the empty set. Our goal will be to show that μ is σ-additive, that is, for all $\{A_j\}_{j\in\mathbb{N}} \subseteq \mathscr{S}$, if $A = \bigsqcup_{j=1}^{\infty} A_j \in \mathscr{S}$, then

$$\mu\left(\bigsqcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j)$$

We have seen that it is possible to uniquely extend μ to a σ-additive function $\nu : \mathcal{A}(\mathscr{S}) \to [0,\infty]$ on the generated algebra.

Proposition. μ is σ-additive.

Proof. Let $A_j, A \in \mathscr{S}$ with $A = \bigsqcup_{j=1}^{\infty} A_j$. We want to show that $\mu(A) = \sum_{j=1}^{\infty} \mu(A_j)$. It will be sufficient to show that $\nu(A) = \sum_{j=1}^{\infty} \nu(A_j)$, we prefer this identity since we know that the finite unions are defined on the algebra $\mathcal{A}(\mathscr{S})$, but they may not be defined on the semi-algebra $\mathscr{S}$. More explicitly, we will extend the (finitely) additive function μ to the additive function ν, prove that ν is σ-additive and this will imply that μ is σ-additive.
By monotonicity and finite additivity of μ, we have

$$\nu(A) \geq \nu\left(\bigsqcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \nu(A_j)$$

Taking the limit shows that $\nu(A) \geq \sum_{j=1}^{\infty} \nu(A_j)$. It remains to show the reverse inequality, $\nu(A) \leq \sum_{j=1}^{\infty} \nu(A_j)$.

First, assume that $A = (a, b]$ for $a, b \in \mathbb{R}$, and that $A_j = (a_j, b_j]$ and $A = \bigsqcup_{j=1}^{\infty} A_j$. Let $\varepsilon > 0$. By our assumptions, we may write $(a, b] = \bigsqcup_{j=1}^{\infty}(a_j, b_j]$. Observe that $[a+\varepsilon, b] \subseteq (a, b]$ and $\bigsqcup_{j=1}^{\infty}(a_j, b_j] \subseteq \bigcup_{j=1}^{\infty}(a_j, b_j + \frac{\varepsilon}{2^j})$. Using the fact that $[a+\varepsilon, b]$ is covered by the open sets $(a_j, b_j + \frac{\varepsilon}{2^j})$, we may obtain a finite subcover so that

$$(a+\varepsilon, b] \subseteq [a+\varepsilon, b] \subseteq \bigcup_{k=1}^{m}\left(a_{j_k}, b_{j_k} + \frac{\varepsilon}{2^{j_k}}\right) \subseteq \bigcup_{k=1}^{m}\left(a_{j_k}, b_{j_k} + \frac{\varepsilon}{2^{j_k}}\right]$$

But using the monotonicity of ν, we have

$$\nu((a+\varepsilon, b]) \leq \nu\left(\bigcup_{k=1}^{m}\left(a_{j_k}, b_{j_k} + \frac{\varepsilon}{2^{j_k}}\right]\right) \leq \sum_{k=1}^{m}\nu\left(\left(a_{j_k}, b_{j_k} + \frac{\varepsilon}{2^{j_k}}\right]\right)$$

We see that

$$\begin{aligned} b - a - \varepsilon &\leq \sum_{k=1}^{m}\left(b_{j_k} - a_{j_k} + \frac{\varepsilon}{2^{j_k}}\right) \leq \sum_{j=1}^{\infty}\left(b_j - a_j + \frac{\varepsilon}{2^j}\right) \\ &= \sum_{j=1}^{\infty}(b_j - a_j) + \varepsilon \end{aligned}$$

Since this holds for all ε, we have shown $\nu(A) \leq \sum_{j=1}^{\infty}\nu(A_j)$ for this case.

To extend this result to all cases, define $E_n = (-n, n]$ and note that $E_n \uparrow \mathbb{R}$. Fix an $A \in \mathscr{S}$. We claim that $A \cap E_n$ is also an element of the semi-algebra. This follows from casework.

1. $A \cap E_n = \varnothing$.
2. $A \subseteq E_n = A$ or $E_n \subseteq A = E_n$.
3. We have an interval of the form $(a, b]$ where the left endpoint is from A and the right is from E_n, or vice versa (including cases with infinity).

Also, we claim that

1. $\nu(A) = \lim_{n\to\infty}\nu(A \cap E_n)$
2. $A \cap E_n = \bigsqcup_{j=1}^{\infty} A_j \cap E_n$, note that $A_j \cap E_n \in \mathscr{S}$.

Now, we are under the same conditions as in the case we proved above, so we have

$$\nu(A \cap E_n) \leq \sum_{j=1}^{\infty}\nu(A_j \cap E_n) \leq \sum_{j=1}^{\infty}\nu(A_j)$$

Since this holds for all $n \in \mathbb{N}$, the limit is bounded by this quantity, so $\nu(A) \leq \sum_{j=1}^{\infty}\nu(A_j)$, as desired. Therefore μ is σ-additive. ■

So, starting from the length function μ we defined above on a semi-algebra $\mathscr{S}$, we are able to uniquely extend it to a σ-additive function ν on the algebra $\mathcal{A}(\mathscr{S})$. Using Carathéodory's extension theorem, we may further uniquely extend ν to π^* on the σ-algebra $\sigma(\mathcal{A})$. This is since $\mathbb{R}$ is σ-finite.

7 Complete Measures

In this lecture, we will see the relationship between complete measures and Carathéodory's extension theorem. Recall that in Carathéodory's theorem, we extended a set function to a σ-algebra that contains the generated σ-algebra. There is a relationship between the σ-algebra of measurable sets and complete measures.

Definition. *Complete Measure*
Let $\mathcal{F} \subseteq \mathscr{P}(\Omega)$ be a σ-algebra. Let $\mu : \mathcal{F} \to [0, \infty]$ be a measure. We call the pair $(\mu, \mathcal{F})$ *complete* when the below implication holds.
If $A \in \mathcal{F}$ and $\mu(A) = 0$ and $E \subseteq A$, then $E \in \mathcal{F}$.

Remark. By the monotonicity of the measure, we know that $\mu(E) = 0$. Notice also that the definition does not depend on the fact that $\mathcal{F}$ is a σ-algebra, so it is also alright, but uncommon, to call a collection of sets with this property complete.

7.1 Completing a σ-algebra and a measure

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Define

$$\overline{\mathcal{F}} = \{A \cup N \mid A \in \mathcal{F}, N \subseteq E \in \mathcal{F} \text{ and } \mu(E) = 0\}$$

Proposition. $\overline{\mathcal{F}}$ is a σ-algebra.

Proof.

1. We may write $\Omega = \Omega \cup \varnothing$, so $\Omega \in \overline{\mathcal{F}}$.

2. Assume that $A \in \overline{\mathcal{F}}$. We may write $A = E \cup N$ for $E \in \mathcal{F}$ and $N \subseteq H$ such that $\mu(H) = 0$. We see that

$$A^c = (E \cup N)^c = \underbrace{[(E \cup N)^c \cap H]}_{\subseteq H} \cup [(E \cup N)^c \cap H^c]$$

So, it suffices to show that $(E \cup N)^c \cap H^c$ belongs to $\mathcal{F}$. But $(E \cup N)^c \cap H^c = E^c \cap N^c \cap H^c = E^c \cap H^c$. But $E, H \in \mathcal{F}$, so $E^c \cap H^c \in \mathcal{F}$ and we have $A^c \in \overline{\mathcal{F}}$.

3. Lastly, we must show $\overline{\mathcal{F}}$ is closed under countable unions. Let $\{A_j\}_{j \in \mathbb{N}}$ be a sequence of sets in $\overline{\mathcal{F}}$. By definition, we may write $A_j = E_j \cup N_j$ for $E_j \in \mathcal{F}$ and $N_j \subseteq H_j$ where $H_j \in \mathcal{F}$ and $\mu(H_j) = 0$. We write

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} (E_j \cup N_j) = \bigcup_{j=1}^{\infty} E_j \cup \bigcup_{j=1}^{\infty} N_j$$

But since $E_j \in \mathcal{F}$ for all $j \in \mathbb{N}$, we have $\bigcup_{j=1}^{\infty} E_j \in \mathcal{F}$. Further, $\bigcup_{j=1}^{\infty} N_j \subseteq \bigcup_{j=1}^{\infty} H_j$. Since μ is a measure, we see that

$$\mu\left(\bigcup_{j=1}^{\infty} H_j\right) \le \sum_{j=1}^{\infty} \mu(H_j) = 0$$

It follows that $\overline{\mathcal{F}}$ is closed under countable unions.

■

Remark. It is easy to see that $\mathcal{F} \subseteq \overline{\mathcal{F}}$.

Now, we want to define an extension of the measure μ, denoted $\overline{\mu}$ on the σ-algebra $\overline{\mathcal{F}}$. Observe that we should have

$$\overline{\mu}(A \cup N) \le \overline{\mu}(A \cup E) = \mu(A \cup E) = \mu(A) + \mu(E) = \mu(A)$$

$$\mu(A) = \overline{\mu}(A) \le \overline{\mu}(A \cup N)$$

So, any reasonable extension $\overline{\mu}$ must be defined as $\overline{\mu}(A \cup N) = \mu(A)$.

Proposition. The extension $\overline{\mu}$ on $\overline{\mathcal{F}}$ defined by $\overline{\mu}(A \cup N) = \mu(A)$ is a measure.

Proof. First, we show that $\overline{\mu}$ is well-defined. Let $A, B \in \mathcal{F}$ so that $N \subseteq E \in \mathcal{F}$ with $\mu(E) = 0$ and $M \subseteq F \in \mathcal{F}$ with $\mu(F) = 0$ and $A \cup N = B \cup M$. Notice that

$$A \subseteq A \cup N = B \cup M \subseteq B \cup F$$

By monotonicity

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

A symmetrical argument shows that $\mu(B) \le \mu(A)$, hence $\mu(A) = \mu(B)$.

Second, we should verify that $\overline{\mu}$ is an extension of μ, this means that $\overline{\mu}\big|_{\mathcal{F}} = \mu$. But clearly

$$\overline{\mu}(A) = \overline{\mu}(A \cup \varnothing) = \mu(A)$$

Lastly, we want to show that $\overline{\mu}$ is σ-additive. Let $\{A_j\} \subseteq \overline{\mathcal{F}}$ is a sequence of disjoint sets and set $A = \bigsqcup_{j=1}^{\infty} A_j$. Using the definition of $\overline{\mathcal{F}}$, we may write $A_j = E_j \cup N_j$, where $N_j \subseteq H_j$ and $\mu(H_j) = 0$. Also

$$A = \bigsqcup_{j=1}^{\infty} A_j = \bigsqcup_{j=1}^{\infty} E_j \cup \bigsqcup_{j=1}^{\infty} N_j$$

But in the above, we have written A as a representation of a set in the σ-algebra and a null set, so

$$\overline{\mu}(A) = \mu\left(\bigsqcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j) = \sum_{j=1}^{\infty} \overline{\mu}(A_j)$$

■

Proposition. $(\overline{\mu}, \overline{\mathcal{F}})$ is complete. Sometimes we may say $\overline{\mathcal{F}}$ is $\overline{\mu}$-complete.

Proof. To show this, we need to show that if $\overline{\mu}(E) = 0$, then for every $A \subseteq E$, we have $A \in \overline{\mathcal{F}}$. Since $E \in \overline{\mathcal{F}}$, we may write it as $E = B \cup N$ where $B \in \mathcal{F}$ and $N \subseteq H$ with $H \in \mathcal{F}$ and $\mu(H) = 0$. But

$$A \subseteq E = (B \cup N) \subseteq (B \cup H) \in \mathcal{F}$$

So, by monotonicity and sub-additivity, we see that

$$\mu(B \cup H) \leq \mu(B) + \mu(H) = 0$$

For the last equality, note that $\overline{\mu}(E) = \mu(B) = 0$. So, we can write $A = \varnothing \cup A$. Since $\varnothing \in \mathcal{F}$ and A is contained in a null set of $\mathcal{F}$, it follows that $A \in \overline{\mathcal{F}}$. ■

Remark. This sequence of propositions shows that it is always possible to complete a measure.

Remark. It is important to notice that the extended σ-algebra $\overline{\mathcal{F}}$ depends on the measure μ. So we should denote this symbol $\overline{\mathcal{F}}_\mu$ to be explicit. If we extend using a different measure, we may get a completely different σ-algebra.

Proposition. The extension $\overline{\mu} : \overline{\mathcal{F}}_\mu \to [0, \infty]$ of $\mu : \mathcal{F} \to [0, \infty]$ is unique.

Proof. Let $\nu : \overline{\mathcal{F}} \to [0, \infty]$ be another extension of μ. Take $B \in \overline{\mathcal{F}}$, then we may write $B = E \cup N$ for $E \in \mathcal{F}$ and $N \subseteq H \in \mathcal{F}$ with $\mu(H) = 0$. Observe that

$$\overline{\mu}(B) = \mu(E) = \nu(E) \leq \nu(E \cup N) = \nu(B)$$

On the other hand

$$\begin{aligned} \nu(B) = \nu(E \cup N) &\leq \nu(E \cup H) \leq \nu(E) + \nu(H) \\ &= \overline{\mu}(E) \leq \overline{\mu}(B) \end{aligned}$$

■

Recall the construction of outer measure π^* and the Carathéodory σ-algebra $\mathcal{M}$ from the previous lectures, we obtain a measure by restricting π^* to $\mathcal{M}$.

Proposition. $\mathcal{M}$ is π^*-complete.

Proof. Let $A \subseteq B$ and $B \in \mathcal{M}$ and $\pi^*(B) = 0$. To show that $\mathcal{M}$ is complete, we must show that for all $E \subseteq \Omega$, we have

$$\pi^*(E) \geq \pi^*(E \cap A) + \pi^*(E \cap A^c)$$

But notice that $E \cap A \subseteq A \subseteq B$. Hence, by monotonicity

$$\pi^*(E \cap A) \leq \pi^*(B) = 0$$

Further $\pi^*(E) \geq \pi^*(E \cap A^c)$ by monotonicity. Summing these two relations gives us the desired inequality. Thus, $A \in \mathcal{M}$. ■

8 Approximation Theorems

Theorem. *Approximation by sets in the algebra*
Let $\mathcal{A} \subseteq \mathscr{P}(\Omega)$ be an algebra and let $\mathcal{F} := \sigma(\mathcal{A})$. Let $\mu : \mathcal{F} \to [0, \infty]$ be a measure, $A \in \mathcal{F}$, and $\mu(A) < \infty$. Then for all $\varepsilon > 0$, there exists $E \in \mathcal{A}$ such that

$$\mu(E \Delta A) < \varepsilon$$

Proof. Fix some $A \in \mathcal{F}$ and assume that $\mu(A) < \infty$. Let ν be the premeasure and recall that

$$\mu(A) = \pi^*(A) = \inf_{\{A_i\}} \sum_{i=1}^{\infty} \nu(A_i)$$

where $A_i \in \mathcal{A}$ and $A \subseteq \bigcup_i A_i$ for any collection $\{A_i\}_{i \in \mathbb{N}}$. By the definition of infimum and since $\mu(A) < \infty$, we may find a collection $\{A_i\}_{i \in \mathbb{N}}$ satisfying

$$\pi^*(A) \leq \sum_{i=1}^{\infty} \nu(A_i) < \pi^*(A) + \varepsilon$$

Thus, there exists an index n_0 so that $\sum_{i \geq n_0} \nu(A_i) < \varepsilon$.

Aside. To see this, just set $S = \sum_{i=1}^{\infty} \nu(A_i)$ and $S_n = \sum_{i=1}^{n} \nu(A_i)$. This will follow from the definition of convergence, since there exists some $N \in \mathbb{N}$ such that $|s - s_n| < \varepsilon$ for all $n \geq N$.

Based on this index n_0, we now have our candidate set as an estimate. Choose

$$E = \bigcup_{i=1}^{n_0} A_i$$

Since $A_i \in \mathcal{A}$ and E is a finite union, we see that $E \in \mathcal{A}$. We want to show that $\pi^*(E \setminus A) < \varepsilon$ and $\pi^*(A \setminus E) < \varepsilon$. For the first inequality, observe that

$$\begin{aligned} \pi^*(E \setminus A) &\leq \pi^* \left(\bigcup_{i=1}^{\infty} A_i \setminus A \right) = \pi^* \left(\bigcup_{i=1}^{\infty} A_i \right) - \pi^*(A) \\ &\leq \sum_{i=1}^{\infty} \pi^*(A_i) - \pi^*(A) < \varepsilon \end{aligned}$$

For the second term, observe that

$$\pi^*(A \setminus E) \leq \pi^* \left(\bigcup_{i=1}^{\infty} A_i \setminus \bigcup_{j=1}^{n_0} A_j \right) \leq \pi^* \left(\bigcup_{j \geq n_0 + 1} A_j \right) \leq \sum_{j=1}^{\infty} \pi^*(A_j) < \varepsilon$$

Note that $\pi^*\left(\bigcup_{i=1}^{\infty} A_i \setminus \bigcup_{j=1}^{n_0} A_j\right) \leq \pi^*\left(\bigcup_{j\geq n_0+1} A_j\right)$ is inequality and not equality since these sets may not be disjoint (i.e. suppose one of the A_j is repeated later on).
Since both of these quantities are bounded by ε, this concludes the theorem. Notationally, we remark that μ and π^* are the same thing (on the σ-algebra, and similarly for ν on the algebra).

■

Remark. If Ω is σ-finite with respect to μ, then recall that we are able to make the completion $\overline{\mu} : \overline{\mathcal{F}} \to [0, \infty]$. In this case, for any $A \in \overline{\mathcal{F}}$, for all $\varepsilon > 0$, there exists some $E \in \mathcal{A}$ for which the above approximation theorem applies. The assumption of σ-finiteness is needed to guarantee the uniqueness of the extension of μ from the algebra, and to $\overline{\mathcal{F}}$.

8.1 Regular Measure

Let Ω be a topological space and let $\mathscr{B}$ indicate the Borel σ-algebra. Let $\mu : \mathcal{F} \to [0, \infty]$ be a measure and $\mathcal{F} \supseteq \mathscr{B}$.

Definition. *Regular*
In the setting above, μ is *regular* if for all $A \in \mathcal{F}$ and for all $\varepsilon > 0$ there exists $F \subseteq A \subseteq G$, where F is closed, G is open such that

$$\mu(G \setminus F) < \varepsilon$$

Remark. Here, we do not assume that $\mu(A)$ is finite.
Remark. It is equivalent to require that $\mu(G \setminus A) < \varepsilon$ and $\mu(A \setminus F) < \varepsilon$.
To see this, first note that $G \setminus A$ and $A \setminus F$ are disjoint sets and $(G \setminus A) \cup (A \setminus F) = G \setminus F$. So, using the countable additivity of measure

$$\mu(G \setminus F) = \mu(G \setminus A) + \mu(A \setminus F)$$

Remark. Above, we assumed that $\mathscr{B} \subseteq \mathcal{F}$. If μ is regular, then it may be shown that $\mathcal{F} \subseteq \overline{\mathscr{B}}_\mu$.
Suppose that $A \in \mathcal{F}$. Since μ is regular, for $n \in \mathbb{N}$, we may find open sets

$$F_n \subseteq A \subseteq G_n$$

where F_n is closed, G_n is open (and thus both belong to $\mathscr{B}$ by definition). Also, we have that $\mu(G_n \setminus F_n) < \frac{1}{n}$. Denote $F = \bigcup_{n\in\mathbb{N}} F_n$ and $G = \bigcap_{n\in\mathbb{N}} G_n$, then $F, G \in \mathscr{B}$ and that $F \subseteq A \subseteq G$. Further

$$\mu(G \setminus F) \leq \mu(G_n \setminus F) \leq \mu(G_n \setminus F_n) \leq \frac{1}{n}$$

Now, since this is true for any $n \in \mathbb{N}$, we conclude that $\mu(G \setminus F) = 0$. Finally, we may write

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

where $F \in \mathscr{B}$ and $A \setminus F \subseteq G \setminus F$. Hence, this shows that A may be written as a union of a set in $\mathscr{B}$ and a null set relative to μ, that is, $A \in \overline{\mathscr{B}}_\mu$.

Theorem. *Regularity of the Lebesgue Measure*
Let $\mu : \mathscr{L} \to [0, \infty]$, where $\mathscr{L}$ is the Lebesgue σ-algebra. Then, μ is regular.

Proof. To reiterate, the regularity of μ means that for all $A \in \mathscr{L}$ and for all $\varepsilon > 0$, there exists $F \subseteq A \subseteq G$ with F closed and G open so that $\mu(G \setminus F) < \varepsilon$. Note that we do not require A to have finite measure.
Fix $\varepsilon > 0$ and some $A \in \mathscr{L}$. We aim to find an open set G so that $A \subseteq G$ and $\mu(G \setminus A) < \varepsilon$. Our first goal is to ensure the measure of A is finite. Set $E_n = [-n, n]$ and $A_n = A \cap E_n$. By monotonicity, we see that $\mu(A_n) \leq 2n < \infty$. Since $\mu(A_n)$ is finite, there exists a sequence $\{B_{n,k}\}_{k \in \mathbb{N}}$ so that $B_{n,k} \in \mathcal{A}$ and $A_n \subseteq \bigcup_{k \in \mathbb{N}} B_{n,k}$ and via the definition of infimum,

$$\mu(A_n) \leq \sum_{k=1}^{\infty} \mu(B_{n,k}) < \mu(A_n) + \frac{\varepsilon}{2^n}$$

Since $A = \bigcup_{n \in \mathbb{N}} A \cap E_n$, we have a cover for A via $B_{n,k}$, but we need this to consist of open sets. Since $B_{n,k} \in \mathcal{A}$ and is bounded, we may write

$$B_{n,k} = \bigsqcup_{j=1}^{l_{n,k}} I_{n,k,j}$$

where $I_{n,k,j} = (a_{n,k,j}, b_{n,k,j}]$. We may transform these intervals into open sets by picking some $\delta_{n,k,j} > 0$ and setting $c_{n,k,j} = b_{n,k,j} + \delta_{n,k,j}$. Let $J_{n,k,j} = (a_{n,k,j}, c_{n,k,j})$ and note that $I_{n,k,j} \subseteq J_{n,k,j}$. So, we have

$$B_{n,k} \subseteq G_{n,k} = \bigcup_{j=1}^{l_{n,k}} J_{n,k,j}$$

Thus, by monotonicity, we have

$$\mu(B_{n,k}) \leq \mu\left(G_{n,k}\right) \leq \sum_{j=1}^{l_{n,k}} I_{n,k,j} + \delta_{n,k,j} \leq \mu(B_{n,k}) + \frac{\varepsilon}{2^n 2^k}$$

Above, we have chosen $\sum_j \delta_{n,k,j} \leq \frac{\varepsilon}{2^n 2^k}$. Also, $G_{n,k}$ is open by construction. Next, we have

$$A_n \subseteq \bigcup_{k=1}^{\infty} G_{n,k} := G_n$$

where G_n is open since it is the union of open sets. We have

$$\mu(G_n) \leq \sum_{k=1}^{\infty} \mu(G_{n,k}) \leq \sum_{k=1}^{\infty} \left(\mu(B_{n,k}) + \frac{\varepsilon}{2^n 2^k}\right) = \sum_{k=1}^{\infty} \mu(B_{n,k}) + \frac{\varepsilon}{2^n} \leq \mu(A_n) + \frac{2\varepsilon}{2^n}$$

Define $G = \bigcup_{n=1}^{\infty} G_n$. Clearly, G is an open set. Since $A_n \subseteq G_n$ for all $n \in \mathbb{N}$, we have $A \subseteq G$. Finally, we compute

$$\begin{aligned}\mu(G \setminus A) &= \mu\left(\bigcup_{n\in\mathbb{N}} G_n \setminus \bigcup_{n\in\mathbb{N}} A_n\right) \leq \mu\left(\bigcup_{n\in\mathbb{N}} (G_n \setminus A_n)\right) \\ &\leq \sum_{n=1}^{\infty} \mu(G_n \setminus A_n) = \sum_{n=1}^{\infty} \mu(G_n) - \mu(A_n) \leq 2\varepsilon\end{aligned}$$

To finish, we must find a closed set $F \subseteq A$ such that F is closed and $\mu(A \setminus F) < \varepsilon$. To do this, consider $A^c \in \mathscr{L}$. By the first part of the proof, we may find an open set H so that $A^c \subseteq H$ and $\mu(H \setminus A^c) < \varepsilon$. So, $F = H^c$ is closed and $F \subseteq A$ by taking complements on both sides. We have

$$\mu(A \setminus F) = \mu(A \cap F^c) = \mu(H \cap A) = \mu(H \cap (A^c)^c) = \mu(H \setminus A^c) < \varepsilon$$

■

Notation. We may refer to the intersection of open sets as G_δ sets and the union of closed sets as F_σ sets.

Remark. Using the theorem above, for all $A \in \mathscr{L}$, we may find $R \in F_\sigma$ and $S \in G_\delta$ so that $\mu(S \setminus R) = 0$. So, if we want an exact approximation we must go to F_σ or G_δ sets.

9 Integration: measurable and simple functions

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $f : \Omega \to \overline{\mathbb{R}}$ be a function. We want to define the notion of an integral $I(f)$ with the following properties

1. Linearity. $I(\alpha f + g) = \alpha I(f) + I(g)$
2. $I(f)$ should be defined for the largest possible class of functions. Also, if $f \geq 0$, then we should have $I(f) \geq 0$.
3. If $f_n \uparrow f$ and $f_n \geq 0$ for all $n \in \mathbb{N}$, then we should have $I(f_n) \uparrow I(f)$.

We spend a few minutes discussing the difference between the intuitive notions of the Riemann and Lebesgue integrals, using the picture of the vertical and horizontal slices. One key idea is that in order for the picture of integration with respect to measures to make sense, the preimage of the function on measurable sets in the codomain must be measurable.
For instance, suppose that $(\Omega, \mathcal{F}, \mu)$ is a measure space and $f : \Omega \to \mathbb{R}$ where $\mathbb{R}$ is equipped with the Borel σ-algebra. Then, it would be natural to define a measurable function as one where the preimage of a Borel set under f belongs to $\mathcal{F}$.

9.1 Measurable functions

We fix a measure space $(\Omega, \mathcal{F}, \mu)$ as well as a function $f : \Omega \to \overline{\mathbb{R}}$. Note that given $\mathscr{B}$ as the Borel σ-algebra, we may define a σ-algebra on the extended reals by setting

$$\overline{\mathscr{B}} = \{A \cup B \mid A \in \mathscr{B}, B \subseteq \{-\infty, \infty\}\}$$

It may be verified that $\overline{\mathscr{B}}$ is a σ-algebra.

Definition. *Measurable function*
Using the above setting, a function $f : \Omega \to \overline{\mathbb{R}}$ is measurable if $f^{-1}(A) \in \mathcal{F}$ for all $A \in \overline{\mathscr{B}}$.

If we want to emphasize the σ-algebra of the domain, we may say that f is $\mathcal{F}$-measurable.
Remark. If $\mathcal{F} \subseteq \mathcal{G}$ are σ-algebras and f is $\mathcal{G}$-measurable, it is also $\mathcal{F}$-measurable. It may seem inconvenient to have to check every set in the σ-algebra, but it turns out this is not necessary.

Lemma. The following are equivalent.

1. f is measurable
2. $f^{-1}((-\infty, x]) \in \mathcal{F}$ for all $x \in \mathbb{R}$.

3. $f^{-1}((-\infty, x)) \in \mathcal{F}$ for all $x \in \mathbb{R}$.
4. $f^{-1}([x, \infty)) \in \mathcal{F}$ for all $x \in \mathbb{R}$.
5. $f^{-1}((x, \infty)) \in \mathcal{F}$ for all $x \in \mathbb{R}$.

Proof. We will just prove the first equivalence, as the proof of the others are analogous. The forward direction is trivial. If f is measurable, since $(-\infty, x]$ is a Borel set, it follows that $f^{-1}((-\infty, x]) \in \mathcal{F}$ by the definition of measurable.

Conversely, suppose that $f^{-1}((-\infty, x]) \in \mathcal{F}$ for all $x \in \mathbb{R}$. Define

$$\mathscr{C} = \{A \in \overline{\mathscr{B}} : f^{-1}(A) \in \mathcal{F}\}$$

We claim that

1. $\mathscr{C}$ is a σ-algebra
2. $\mathscr{C} \supseteq \mathcal{G}$, where $\mathcal{G}$ is some set which should generate $\overline{\mathscr{B}}$.

Then, from these two properties we know that $\mathscr{C} \supseteq \sigma(\mathcal{G})$. Now, if $\sigma(\mathcal{G})$ is larger than $\overline{\mathscr{B}}$, then we are done.

A natural choice is to set $\mathcal{G} = \{(-\infty, x] : x \in \mathbb{R}\}$, since by assumption these sets satisfy the property that their preimages are measurable. So, $\mathscr{C} \supseteq \mathcal{G}$. Further, these sets generate the extended Borel σ-algebra. It remains to show that $\mathscr{C}$ is a σ-algebra.

1. First, $\overline{\mathbb{R}} \in \mathscr{C}$. This is because $f^{-1}(\overline{\mathbb{R}}) = \Omega$ and $\Omega \in \mathcal{F}$.
2. Suppose $A \in \mathscr{C}$. Then $f^{-1}(A) \in \mathcal{F}$. But since $\mathcal{F}$ is a σ-algebra, we know that $f^{-1}(A^c) = f^{-1}(A)^c \in \mathcal{F}$. Since $A \in \overline{\mathscr{B}}$, we know $A^c \in \overline{\mathscr{B}}$ as well. Thus, $A^c \in \mathscr{C}$.
3. Take $A_j \in \mathscr{C}$ for $j \in \mathbb{N}$. The argument is similar. We know that $\bigcup_j A_j \in \overline{\mathscr{B}}$ since $\overline{\mathscr{B}}$ is a σ-algebra and $f^{-1}\left(\bigcup_j A_j\right) = \bigcup_j f^{-1}(A_j) \in \mathcal{F}$. Thus, $\bigcup_j A_j \in \mathscr{C}$.

■

The above technique will be used often in the ensuing lectures and proofs. The key idea is to create a collection which has the property which we want to prove, then prove that it is a σ-algebra which contains our desired generating set. Sometimes we will use the monotone class theorem, which tells us that the monotone class generated by an algebra is the same as the σ-algebra generated by that algebra.

9.2 Simple functions

Simple functions are examples of measurable functions. Let $(\Omega, \mathcal{F}, \mu)$ be measure space. Let $\{E_j\}_{j=1}^n$ be a collection of $\mathcal{F}$-measurable, pairwise disjoint sets. Let $\mathbf{1}_{E_j} : \Omega \to \overline{\mathbb{R}}$ be the indicator or characteristic function, that is

$$\mathbf{1}_{E_j}(w) = \begin{cases} 1 & \text{if } w \in E_j, \\ 0 & \text{otherwise} \end{cases}$$

Let $c_j \in \mathbb{R}$ be coefficients. Without loss of generality, we may assume that $\bigsqcup_{j=1}^n E_j = \Omega$. If not, we may just add a new set $E_0 = \Omega \setminus \bigsqcup_{j=1}^n E_j \in \mathcal{F}$ so that this condition is satisfied (i.e. we assume that E_j forms a partition of Ω).

Definition. *Simple function*
With the above setting, a simple function $f : \Omega \to \mathbb{R}$ has the representation

$$f = \sum_{j=1}^n c_j \mathbf{1}_{E_j}$$

Proposition. Every simple function is measurable.

Proof. Let $A \in \overline{\mathscr{B}}$. Observe that

$$f^{-1}(A) = \bigsqcup_{k \,:\, c_k \in A} E_k$$

Thus, $f^{-1}(A) \in \mathcal{F}$. ∎

9.3 Defining integration

Going back to our discussion at the beginning of the lecture, we now want a definition of the integral. A natural definition begins with considering simple functions. Let f be a simple function in the sense of the above definition. So, for a nonnegative, simple function f, we may define

$$I(f) = \sum_{j=1}^n c_j \cdot \mu(E_j)$$

Note that to avoid issues with adding and subtracting infinity, we assume $c_j \geq 0$ for all j. We must also check that this function is well-defined, i.e. that $I(f)$ produces the same value for some other representation $f = \sum_{k=1}^m d_k \mu(F_k)$. We may make a routine verification by considering the intersections $E_j \cap F_k$ (i.e. whenever $E_{j'} \cap F_{k'} \neq \varnothing$, then $c_{j'} = d_{k'}$ by assumption).

Now, we would like to extend our definition of integration to all measurable functions. To do this, we take a few steps

1. Let $f : \Omega \to [0, \infty]$ be measurable. Then there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of measurable, nonnegative, and monotonically increasing simple functions that approximate f. To be explicit, $f_n \geq 0$ and f_n is simple for all $n \in \mathbb{N}$ and $f_n \uparrow f$.

2. With the same f as before
$$I(f) := \lim_{n\to\infty} I(f_n)$$
and this definition is well-defined (does not depend on the sequence f_n).

3. Let $f : \Omega \to \overline{\mathbb{R}}$ be measurable. Define $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. It turns out that f^+, f^- are both measurable and $f = f^+ - f^-$. Note that we cannot have $\infty - \infty$ since $f^+(w) = \infty \implies f^-(w) = 0$ and vice versa.

   By part (2), since f^+ and f^- are both nonnegative and measurable, we know that $I(f^+)$ and $I(f^-)$ are well-defined. If at least one of these is finite, then the definition $I(f) = I(f^+) - I(f^-)$ makes sense.

Remark. Observe that with this definition, we are able to integrate functions that are not Riemann integrable. For instance, consider $\Omega = (0, 1]$, the Borel σ-algebra $\mathscr{B}$ and the Lebesgue measure λ. Then taking $E = \mathbb{Q} \cap \Omega$, we may integrate $f = \mathbf{1}_{E^c}$:

$$I(f) = \lambda(\Omega) = \lambda(E) + \lambda(E^c) = \lambda(E) = 1$$

10 More on Measurable Functions

Fix a measure space $(\Omega, \mathcal{F}, \mu)$ as our setting. Let $f, g : \Omega \to \overline{\mathbb{R}}$.

Lemma. Suppose g, f are $\mathcal{F}$-measurable and $\alpha \in \mathbb{R}$. Then, the following hold

1. αf is measurable.
2. $\alpha + f$ is measurable.
3. $f + g$ is measurable.
4. f^2 is measurable
5. $\frac{1}{f}$ is measurable, assuming it is defined.
6. $f^+, f^-, |f|$ are measurable.
7. fg is measurable.

Proof.

1. Consider the set $(-\infty, x]$. We want to show that its preimage is measurable, that is
$$f^{-1}((-\infty, x]) = \{\omega \in \Omega : \alpha f(\omega) \leq x\}$$
Suppose that $\alpha = 0$. Then $\alpha f = 0$, which is a constant function, so it is simple and we are done. Assume that $\alpha > 0$. Then
$$\{\omega \in \Omega : \alpha f(\omega) \leq x\} = \left\{\omega : f(\omega) \leq \frac{x}{\alpha}\right\} \in \mathcal{F}$$
If $\alpha < 0$, a similar argument holds, relying on the measurability of f.

3. Consider
$$\begin{aligned}(f+g)^{-1}((-\infty, x)) &= \{\omega \in \Omega : f(\omega) + g(\omega) < x\} \\ &= \bigcup_{r \in \mathbb{Q}} \{\omega : f(w) < r\} \cap \{\omega : g(w) < x - r\}\end{aligned}$$
One direction of the equality above is trivial. To show that the union contains the preimage, given $f(\omega) + g(\omega) < x$, by density we may choose $x - g(\omega) > r > f(\omega)$ for some $r \in \mathbb{Q}$. We have written the preimage as a countable union of measurable sets of $\mathcal{F}$; the union consists of intersections of measurable sets, which are themselves measurable. This also shows that (2) is true.

4. We consider the preimages
$$(f^2)^{-1}((-\infty, x)) = \{\omega \in \Omega : f(\omega)^2 < x\}$$

If $x \leq 0$, then there are no possible values of $f(\omega)^2$ that can be less than 0. So, the preimage is $\varnothing \in \mathcal{F}$. If $x > 0$, then we may rewrite the set as

$$\{\omega \in \Omega : -\sqrt{x} < f(\omega) < \sqrt{x}\}$$

This set is the preimage of an element of the Borel σ-algebra, so it belongs to $\mathcal{F}$ and f^2 is measurable.

7. We immediately obtain (7) as a consequence of the previous proofs since

$$fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$$

5. Again, we want to show that

$$(1/f)^{-1}((-\infty, x)) = \{\omega \in \Omega : \frac{1}{f(\omega)} < x\}$$

belongs to $\mathcal{F}$. Assume that $x > 0$. Then, we may rewrite the preimage as

$$\{\omega : f(\omega) < 0\} \cup \left\{\omega : f(\omega) > \frac{1}{x}\right\}$$

To interpret the point at 0 take the convention that $\frac{1}{0} = +\infty$. Since f is measurable, the above expression is a union of measurable sets, which is measurable.
Assume that $x = 0$. Then the set is simply $\{\omega : f(\omega) < 0\}$, which belongs to $\mathcal{F}$ since f is measurable. Finally, assume $x < 0$. Then we obtain the set

$$\left\{\omega : \frac{1}{x} < f(\omega) < 0\right\}$$

Since f is measurable, we see that the above set belongs to $\mathcal{F}$. Hence $\frac{1}{f}$ is measurable.

6. Recall that $f^+ = \max\{f, 0\}$. Consider the set

$$\{\omega \in \Omega : f^+(\omega) < x\}$$

If $x \leq 0$, then the above set is empty, and so it belongs in $\mathcal{F}$. On the other hand, if $x > 0$, then we have

$$\{\omega \in \Omega : f(\omega) < x\} \in \mathcal{F}$$

Continuing, consider $f^- = \max\{-f, 0\}$. We know that f is measurable, $-f$ is measurable. In the preceding argument we have shown that the maximum of a measurable function and 0 is measurable, hence f^- is measurable. Lastly, since $|f| = f^+ + f^-$, we see that it is measurable.

■

Lemma. Let $(f_n)_{n\geq 1}$ be a sequence of functions where $f_n : \Omega \to \overline{\mathbb{R}}$ and f_n is measurable. 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 $f_n \to f$, then f is measurable.

Note that these supremums and infimums are considered pointwise.

Proof.

1. First, we consider $\sup_n f_n$. Let $x \in \mathbb{R}$. Observe that

$$\{\omega \in \Omega : \sup_n f_n(\omega) > x\} = \bigcup_{n\geq 1} \{\omega : f_n(\omega) > x\}$$

   To elaborate, if $\sup_n f_n(\omega) > x$, then there exists some $k \in N$ so that $f_k(\omega) > x$, so the forward set inclusion holds. Conversely, if ω belongs to the set on the right, then $f_k(\omega) > x$ for some $k \in \mathbb{N}$. But then $\sup_n f_n(\omega) \geq f_k(\omega) > x$. Since the sets in the union are all measurable, so is the union. This proves the claim. To prove that the infimum is measurable, notice that $\inf\{f_n\} = -\sup_n\{-f_n\}$.

2. From the previous part, we know that $\sup_{n\geq k} f_n$ is measurable for all $k \in \mathbb{N}$. We write

$$\limsup_n f_n = \inf_k \sup_{n\geq k} f_n$$

   and apply the first part directly to conclude that the lim sup is measurable. An analogous argument holds for lim inf.

3. If $\lim_{n\to\infty} f_n = f$, then $\lim_n f_n = \limsup_n f_n$. By the previous part, we see that f is measurable.

■

Remark. It is important to notice that the proof above holds only for the supremum or infimum over a *countable* set.

Example. Let $(\mathbb{R}, \mathcal{L}, \lambda)$ be our measure space, where $\mathcal{L}$ is the Lebesgue σ-algebra and λ the Lebesgue measure. From the first lecture, we know that non-measurable sets exist, so take $A \notin \mathcal{L}$. It is clear that A is not countable because every countable set belongs to $\mathcal{L}$. Define $f_t = \mathbf{1}_{\{t\}}$ as the indicator function for the point $t \in \mathbb{R}$. We see that each f_t is measurable, however, $\sup_{t\in A} f_t = \mathbf{1}_A$ is not measurable, since $f^{-1}(\{1\}) = A$.

This example highlights a key issue of taking a supremum over an uncountable set. We will often need to find a smaller, countable set for which the supremum remains the same to deal with this. An example application where this problem occurs is with stochastic processes, where time is indexed by $\mathbb{R}$.

10.1 Measurable Functions on Topological Spaces

Let Ω be a topological space and $\mathscr{B}$ be the Borel σ-algebra associated to its open sets. Let $\mathcal{F} \supseteq \mathscr{B}$ and fix a measure space $(\Omega, \mathcal{F}, \mu)$. Let $f : \Omega \to \overline{\mathbb{R}}$.

Proposition. If f is continuous, then it is measurable.

Proof. Since f is continuous, this means that if A is open, then $f^{-1}(A)$ is open. But open sets are measurable in both Borel σ-algebras. So, $f^{-1}(A) \in \mathcal{F}$, and by a lemma from the previous lecture, f is measurable. ■

Observation. Another simple observation to make is that the measurability of f depends on the σ-algebra $\mathcal{F}$. For instance, if we know that f is measurable with respect to $\mathcal{F}$ and $\mathcal{F} \subseteq \mathcal{G}$, then f is measurable with respect to $\mathcal{G}$ as well.

10.2 Properties that Hold Almost Surely

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let P be a property. As an example of what this means, let $f : \Omega \to \overline{\mathbb{R}}$, we may think of a property as some subset of Ω where the points satisfy a condition:

$$\{\omega \in \Omega : f(\omega) = 0\}$$

Definition. *Almost surely, almost everywhere*
We say that the property P holds almost surely (a.s.) or almost everywhere (a.e.) if there exists $E \in \mathcal{F}$, the property P holds on E and $\mu(E^c) = 0$.

Example. We say that $f = 0$ almost surely if $f(\omega) = 0$ for all $\omega \in E$ and $\mu(E^c) = 0$.

We move to a new setting. Consider $\mathbb{R}$ and let $\mathscr{B}$ be the Borel σ-algebra on $\mathbb{R}$, let $\mathcal{L}$ be the Lebesgue σ-algebra and λ the Lebesgue measure.

Proposition. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be $\mathcal{L}$-measurable. Let $g : \mathbb{R} \to \overline{\mathbb{R}}$ and $g = f$ a.s. Then, g is also $\mathcal{L}$-measurable.

Proof. A key part to proving the above assertion is that λ is complete with respect to $\mathcal{L}$. Our goal is to show that g is measurable, so fix $A \in \overline{\mathscr{B}}$. Since $g = f$ a.s., by definition, this means that there exists $E \in \mathcal{L}$ such that $\lambda(E^c) = 0$ and $f(\omega) = g(\omega)$ for all $\omega \in \Omega$. Observe that

$$\begin{aligned} g^{-1}(A) &= \{\omega \in \Omega : g(\omega) \in A\} \\ &= (\{\omega \in \Omega : g(\omega) \in A\} \cap E) \cup (\{\omega \in \Omega : g(\omega) \in A\} \cap E^c) \\ &= (\{\omega \in \Omega : f(\omega) \in A\} \cap E) \cup (\{\omega \in \Omega : g(\omega) \in A\} \cap E^c) \\ &= (f^{-1}(A) \cap E) \cup (\{\omega \in \Omega : g(\omega) \in A\} \cap E^c) \end{aligned}$$

We have $f^{-1}(A) \cap E \in \mathcal{L}$ as it is the intersection of sets in $\mathcal{L}$. On the other hand, since E^c has measure zero and $\{\omega \in \Omega : g(\omega) \in A\} \cap E^c$ is a subset of E^c and $\mathcal{L}$ is complete, we see that $\{\omega \in \Omega : g(\omega) \in A\} \cap E^c \in \mathcal{L}$ as well. Therefore, $g^{-1}(A) \in \mathcal{L}$. ■

Remark. The same statement in the above proposition does not hold if we replace $\mathcal{L}$ with the Borel σ-algebra $\mathscr{B}$. This highlights the role of completeness. The previous statement will often be used to show that an unknown function which is equal to measurable function a.e. is also measurable. This is why we will often consider the Lebesgue σ-algebra for the measure space of the domain.

Proposition. Suppose $f : \Omega \to \overline{\mathbb{R}}$ is $(\mathcal{F}, \overline{\mathscr{B}})$-measurable and $g : \mathbb{R} \to \mathbb{R}$ is $(\overline{\mathscr{B}}, \overline{\mathscr{B}})$-measurable. Then the composition $g \circ f : \Omega \to \mathbb{R}$ is measurable.

Proof. The key idea is $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. ■

The above proof is useful for showing measurability. For instance, we could use this to show f^2 is measurable in an alternative way. Since f is measurable and g given by $x \mapsto x^2$ is continuous (thus measurable), their composition is measurable by the above proposition.

11 Definition of the Integral

Fix a measure space $(\Omega, \mathcal{F}, \mu)$. Recall our discussion on the symbol associated to the integral, $I(f)$. We give an outline of the lesson and this construction.

1. We may define a simple function on a finite partition $E_1, \ldots, E_n$ of $\mathcal{F}$-measurable sets in Ω with the representation
$$f = \sum_{j=1}^{n} c_j \mathbf{1}_{E_j}$$
Suppose that f is simple with $c_j \geq 0$. Then, we may define the integral by
$$I(f) = \sum_{j=1}^{n} c_j \cdot \mu(E_j)$$
It may happen that the sum is $+\infty$, but we will never run into issues with subtraction with the assumption of non-negativity.

2. Let $f : \Omega \to \mathbb{R}$ be a measurable function and $f \geq 0$. There exists a sequence of simple functions $f_n \uparrow f$. Then,
$$I(f) = \lim_{n\to\infty} I(f_n)$$
It will turn out that since the sequence of functions is monotone, the sequence of integrals will be monotone as well. So, the limit will exist. Further, this function is well-defined, it does not depend on the sequence.

3. Consider the measurable f as before. We will say that f is integrable if $I(f^+)$ is finite and $I(f^-)$ is finite. Then, we can define
$$I(f) = I(f^+) - I(f^-)$$

Lemma. With the previous setting, let f be measurable and nonnegative. Then, there exists a sequence $(f_n)_{n\in\mathbb{N}}$ which is nonnegative, simple, and monotonically increasing to f, that is, $f_n \leq f_{n+1}$ and $\lim_n f_n = f$.

Proof. We define a sequence of functions f_n by
$$f_n(x) = \begin{cases} n & \text{if } f(x) \geq n \\ \frac{k}{2^n} & \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}, \quad 0 \leq k \leq n2^n - 1 \end{cases}$$
Effectively, we are truncating the function along certain intervals in its codomain. Another way to write this function is
$$f_n(x) = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{\left\{\frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\right\}} + n\mathbf{1}_{\{f(x)\geq n\}}$$

Since f is measurable, the above function is measurable. To be explicit, we have

$$f^{-1}\left(\left[\frac{k}{2^n},\frac{k+1}{2^n}\right)\right)\in\mathcal{F},\quad f^{-1}([n,\infty))\in\mathcal{F}$$

So, with the above representation, we see that f_n is nonnegative and a simple function.

Next, we want to show that $f_n(x)\xrightarrow[n\to\infty]{} f(x)$. We proceed by cases.

1. Suppose that $f(x)=+\infty$. Then $f_n(x)=n$ for all $n\in\mathbb{N}$ and the limit holds.

2. Suppose that $f(x)<\infty$. This means that there exists some n_0 such that $f(x)<n_0$. Pick $n\geq n_0$. We are in the second case for the function f_n, since $f(x)<n$. Suppose $k\in\mathbb{N}$ satisfies $\frac{k}{2^n}\leq f(x)<\frac{k+1}{2^n}$. This inequality implies $k\leq 2^n f(x)<k+1$, hence $k=\lfloor 2^n f(x)\rfloor$. But then

$$\frac{2^n f(x)-1}{2^n}<f_n(x)=\frac{\lfloor 2^n f(x)\rfloor}{2^n}\leq f(x)$$

   Rearranging these inequalities, we have

$$f_n(x)\leq f(x)<f_n(x)+\frac{1}{2^n}$$

   Thus, $f_n(x)\xrightarrow[n\to\infty]{} f(x)$.

Finally, we will show that the sequence of functions f_n is monotonically increasing. We consider similar cases.

1. Suppose $f(x)=+\infty$. Then $f_n(x)=n<n+1=f_{n+1}(x)$.

2. Suppose $f(x)<+\infty$. Suppose that $f(x)\geq n+1$. Then, $f_n(x)=n$ and $f_{n+1}(x)=n+1$ and the inequality is satisfied again. Suppose that $n\leq f(x)<n+1$. Then since $f(x)\geq n$, we use the definition of the function to obtain

$$f_{n+1}(x)=\frac{\lfloor f(x)2^{n+1}\rfloor}{2^{n+1}}\geq\frac{\lfloor n2^{n+1}\rfloor}{2^{n+1}}=n=f_n(x)$$

   Now, suppose $f(x)<n$. Using the definition of the function as before, we have

$$f_n(x)=\frac{\lfloor f(x)2^n\rfloor}{2^n},\quad f_{n+1}(x)=\frac{\lfloor f(x)2^{n+1}\rfloor}{2^{n+1}}$$

   The geometric intuition here is that the function f_{n+1} splits the interval $[\frac{k}{2^n},\frac{k+1}{2^n})$ into two new intervals: $[\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}})$ and $[\frac{2k+1}{2^{n+1}},\frac{2(k+1)}{2^{n+1}})$. If $f(x)\in[\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}})$, then $f_{n+1}(x)=f_n(x)$. On the other hand, if $f(x)\in[\frac{2k+1}{2^{n+1}},\frac{2(k+1)}{2^{n+1}})$, then

$$f_{n+1}(x)=\frac{2k+1}{2^{n+1}}>\frac{2k}{2^{n+1}}=f_n(x)$$

Thus, monotonicity holds in every case and for all $x\in\Omega$, and the conclusion follows. ■

Proposition. The integral is monotone on simple functions. If f, g are nonnegative, simple and $f \leq g$, then $I(f) \leq I(g)$.

Proof. Since f is simple, there exists a finite partition $E_1, \ldots, E_n$ of sets in $\mathcal{F}$ and coefficients $c_j \in \mathbb{R}$ so that

$$f = \sum_{j=1}^{n} c_j \cdot \mathbf{1}_{E_j}$$

We have a similar representation for g

$$g = \sum_{k=1}^{m} d_k \cdot \mathbf{1}_{F_k}$$

Consider the sets $E_j \cap F_k$. Suppose that this set is nonempty. Then, take $x \in E_j \cap F_k$ and since $f(x) \leq g(x)$, we have $c_j \leq d_k$. But now

$$g = \sum_{j=1}^{n} \sum_{k=1}^{m} d_k \mathbf{1}_{E_j \cap F_k}$$
$$f = \sum_{j=1}^{n} \sum_{k=1}^{m} c_j \mathbf{1}_{E_j \cap F_k}$$

Since the value of the integral does not depend on the partition, we see that

$$I(g) = \sum_{j=1}^{n} \sum_{k=1}^{m} d_k \mu(E_j \cap F_k) \geq \sum_{j=1}^{n} \sum_{k=1}^{m} c_j \mu(E_j \cap F_k) = I(f)$$

Note that $d_k \mu(E_j \cap F_k) \geq c_j \mu(E_j \cap F_k)$ for all k, j since $\mu(E_j \cap F_k)$ is either 0 or if it is not zero, then $d_k \geq c_j$. ■

We now have two important pieces to generalizing the definition of the integral to nonnegative, measurable functions. If $f : \Omega \to \mathbb{R}$ is measurable and $f \geq 0$, we have shown that there exists a sequence of nonnegative, simple functions $f_n \uparrow f$. Further, we know that the integral is monotone for simple functions so that $I(f_n) \leq I(f_{n+1})$ for all $n \in \mathbb{N}$. But then $\lim_{n\to\infty} I(f_n)$ is a monotonically increasing sequence which is bounded in $\overline{\mathbb{R}}$ and so its limit exists. It remains to verify that this definition of the integral does not depend on the sequence $(f_n)_{n\in\mathbb{N}}$.

Lemma. Suppose that f_n are nonnegative, simple functions and $f_n \leq f_{n+1}$ and g is a nonnegative, simple function. If $g \leq \lim_n f_n$, then $I(g) \leq \lim_n I(f_n)$.

The main difference between this and the next theorem is that we compare of a limit of simple functions with a simple function.

Proof. The main tool we use is that a measure is continuous from below. First, assume that $g = c\mathbf{1}_E$ for $c \geq 0$ and $E \in \mathcal{F}$. Our plan is to prove the statement for this g and then extend using the linearity of the integral to all simple functions. There are two cases.

1. Suppose that $c = 0$. Then we have $I(g) = 0 \cdot \mu(E) = 0$. Note that this holds even if $\mu(E) = \infty$ since $0 \cdot \infty = 0$ by convention. The inequality holds since $I(f_n)$ must be nonnegative, and so must its limit.

2. Suppose $c > 0$. Also, fix some ε so that $0 < \varepsilon < c$. Consider that $\mathbf{1}_E f_n \leq f_n$ and

$$\lim_{n\to\infty} \mathbf{1}_E f_n \geq g$$

Since f_n is a simple function, we may write $f_n = \sum_{k=1}^{M_n} c_{n,k}\mathbf{1}_{F_{n,k}}$. Multiplying by $\mathbf{1}_E$, we have

$$\mathbf{1}_E f_n = \sum_{k=1}^{M_n} c_{n,k}\mathbf{1}_{F_{n,k}\cap E}$$

Denote

$$A_n = \{x \in E : f_n(x) \geq c - \varepsilon\}$$

There are some properties of A_n. First, $A_n \subseteq A_{n+1}$ since f_n is an increasing sequence. So if x satisfies $f_n(x) \geq c - \varepsilon$, it follows that $f_{n+1}(x) \geq f_n(x)$ will satisfy this condition as well. Second, we claim that $\bigcup_{n\in\mathbb{N}} A_n = E$. The only nontrivial set inclusion to check is the reverse inclusion. But, if $x \in E$, we know that $\lim_n f_n \geq g = c\mathbf{1}_E$. So, there exists some $n \in \mathbb{N}$ for which $f_n \geq c - \varepsilon$. Putting everything together, we have

$$\begin{aligned} I(f_n) &\geq I(f_n\mathbf{1}_E) \geq I(f_n\mathbf{1}_{A_n}) \geq I((c-\varepsilon)\mathbf{1}_{A_n}) \\ &= (c-\varepsilon)\mu(A_n) \end{aligned}$$

Since A_n is an increasing sequence of sets, we see that the limit on the left-hand side exists and

$$(c-\varepsilon)\mu(E) = (c-\varepsilon)\lim_{n\to\infty}\mu(A_n) \leq \lim_{n\to\infty} I(f_n)$$

We use continuity from below of the measure to obtain the first equality. Now, if $\mu(E) = +\infty$, then $\lim_{n\to\infty} I(f_n) = +\infty$ and our desired inequality is trivially satisfied. So, suppose that $\mu(E)$ is finite. Then, we have

$$I(g) - \varepsilon\mu(E) = c\mu(E) - \varepsilon\mu(E) \leq \lim_{n\to\infty} I(f_n)$$

Letting ε tend to zero, we obtain our desired inequality.

Finally, we may extend the result above to an arbitrary nonnegative simple function g. Say

$g = \sum_{k=1}^m c_k \mathbf{1}_{E_k}$. We compute

$$\begin{aligned} I(g) &= \sum_{k=1}^m I(c_k \mathbf{1}_{E_k}) \leq \sum_{k=1}^m \lim_n I(f_n \mathbf{1}_{E_k}) = \lim_{n\to\infty} \sum_{k=1}^m I(f_n \mathbf{1}_{E_k}) \\ &= \lim_{n\to\infty} I\left(\sum_{k=1}^m f_n \mathbf{1}_{E_k}\right) = \lim_n I(f_n) \end{aligned}$$

■

Theorem. *Integral is well-defined*
Let f be a measurable, nonnegative function. Then $I(f)$ is well-defined.

Proof. Suppose that (f_n) and (g_k) are two sequences of simple, nonnegative functions that both monotonically increase to f. We aim to prove that

$$\lim_{n\to\infty} I(f_n) = \lim_{n\to\infty} I(g_n)$$

Observe that since $g_k \uparrow f$, we have $g_k \leq f = \lim_n f_n$. Using the previous lemma, we may conclude that $I(g_k) \leq \lim_n I(f_n)$. Since this is true for all $k \in \mathbb{N}$, we may deduce that $\lim_k I(g_k) \leq \lim_n I(f_n)$. But we may make a symmetrical argument to see that $\lim_k I(g_k) \geq \lim_n I(f_n)$, so the conclusion follows immediately. ■

Definition. *Integrable*
Let $f : \Omega \to \mathbb{R}$ be a measurable function. We say that f is integrable if $I(f^+) < \infty$ and $I(f^-) < \infty$ and in this case we set

$$I(f) = I(f^+) - I(f^-)$$

12 Integral of Simple Functions

This material in this lecture should actually precede the definition of the integral, as we need some results such as additivity in the proofs.

12.1 Linearity of the Integral

Let f, g be nonnegative, measurable functions and $c \geq 0$. We would like to show that the integral is linear, that is

$$\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu$$
$$\int cf\, d\mu = c\int f\, d\mu$$

Proposition. Suppose that f, g are non-negative simple functions. Then

$$I(f+g) = I(f) + I(g), \quad I(kf) = kI(f),\, k \geq 0$$

Note that these integrals may be infinite.

Another way to state this is that we are proving that the integral on the space of simple functions is a linear functional.

Remark. From these properties, it will follow that $g \leq f$ means that $I(g) \leq I(f)$ since

$$f = g + (f-g), \quad I(f) = I(g + (f-g)) = I(g) + I(f-g) \geq I(g)$$

Proof. Since f is simple, we may write

$$f = \sum_{j=1}^{n} c_j \mathbf{1}_{E_j}, \quad c_j \geq 0$$
$$g = \sum_{k=1}^{m} d_k \mathbf{1}_{F_k}, \quad d_k \geq 0$$

Observe that we have the following decompositions: $E_j = \sum_{k=1}^m E_j \cap F_k$, $F_k = \sum_{j=1}^m F_k \cap E_j$, $\mathbf{1}_{E_j} = \sum_{k=1}^m \mathbf{1}_{E_j \cap F_k}$. As a result, we may write

$$f = \sum_{j,k} c_j \mathbf{1}_{E_j\cap F_k}, \quad g = \sum_{j,k} d_k \mathbf{1}_{E_j\cap F_k}, \quad f+g = \sum_{j,k}(c_j+d_k)\mathbf{1}_{E_j\cap F_k}$$

Now, by definition, we have

$$\begin{aligned} I(f+g) &= \sum_{j,k}(c_j+d_k)\mu(E_j\cap F_k) \\ &= \sum_j c_j \sum_k \mu(E_j \cap F_k) + \sum_k d_k \sum_j \mu(E_j\cap F_k) = I(f) + I(g) \end{aligned}$$

The proof of homogeneity, the second property, follows analogously. ■

13 Properties of the Integral

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be measurable. We go through several properties of the integral extending them from simple functions to nonnegative, measurable functions.

1. *Montonicity of the integral.* We want to show that if $0 \le g \le f$, then

$$I(g) \le I(f)$$

   From a previous lecture, we may find a nonnegative sequence of simple functions $f_n \uparrow f$ and $g_n \uparrow g$. Define $h_n = \min(f_n, g_n) \le f_n$. The minimum of simple functions is simple, so h_n is simple and nonnegative, further, $h_n \uparrow g$. Observe that

$$I(g) = \lim_n I(h_n) \le \lim_n I(f_n) = I(f)$$

   where the middle inequality follows from the monotonicity of the integral on simple functions.

2. *Linearity.* For $f \ge 0$ and $g \ge 0$ measurable and a real $c \ge 0$, we have

$$I(f+g) = I(f) + I(g), \quad I(cf) = cI(f)$$

   There exist sequences of simple functions $f_n \uparrow f$ and $g_n \uparrow g$. Notice that $f_n + g_n$ is a sequence of nonnegative, simple functions, being the sum of simple functions, and that $(f_n + g_n) \uparrow (f + g)$. Similarly, $cf_n \uparrow cf$. We have

$$I(f+g) = \lim_n I(f_n + g_n) = \lim_n I(f_n) + I(g_n) = \lim_n I(f_n) + \lim_n I(g_n) = I(f) + I(g)$$

   A similar argument holds for the remaining property. Note that we may use linearity to give an alternate argument for monotonicity.

Proposition. If f is integrable, then $\mathbf{1}_A f$ is integrable, for $A \in \mathcal{F}$.

Proof. First, note that

$$(\mathbf{1}_A f)^+ = \mathbf{1}_A f^+ \le f^+$$

Also, $\mathbf{1}_A f$ is measurable because it is the product of measurable functions. So, using the properties we have proven above, we see that $I((\mathbf{1}_A f)^+) \le I(f^+) < \infty$. An analogous argument holds for the negative part of the function, so $\mathbf{1}_A f$ is integrable. ■

Since we are dealing with measurable functions generally now, we have new notation

$$I(f) = \int f \, d\mu, \quad \int \mathbf{1}_A f \, d\mu = \int_A f \, d\mu$$

The expression above makes sense because of the previous proposition.

Proposition. If $f : \Omega \to \overline{\mathbb{R}}$ is measurable and $f \geq 0$ and $E \in \mathcal{F}$ with $\mu(E) = 0$, then f is integrable over E and $\int_E f \, d\mu = 0$.

Proof. Let $g_n \uparrow \mathbf{1}_E f$ be a sequence of simple functions. Suppose $x \notin E$. Then $g_n(x) \leq \mathbf{1}_E f(x) = 0$, hence $g_n(x) = 0$. This shows that $g_n = g_n \mathbf{1}_E$. Since g_n is a simple, nonnegative function, we may write

$$\mathbf{1}_E g_n = \sum_{k=1}^{M_n} c_{n,k} \mathbf{1}_{F_{n,k} \cap E}$$
$$I(\mathbf{1}_E g_n) = \sum_k c_{n,k} \cdot \mu(F_{n,k} \cap E) = 0$$

where the last equality follows since $\mu(F_{n,k} \cap E) \leq \mu(E) = 0$. Since $g_n = g_n \mathbf{1}_E$, it follows that $I(g_n) = 0$ for all $n \in \mathbb{N}$. Hence $\int_E f \, d\mu = 0$. ■

Remark. The proposition above in fact applies for all measurable f, we can apply the proposition to $|f|$ and use monotonicity to conclude that f is integrable from the integrability of $|f|$.

Proposition. Let f, g be measurable and integrable functions. Let $A, B \in \mathcal{F}$ and $A \cap B = \varnothing$.

1. $f + g$ is integrable, cf is integrable for $c \in \mathbb{R}$, and $\mathbf{1}_A f$ is integrable.
2. $\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu$
3. If f is integrable, then $|f| < \infty$ a.s.
4. $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$

Proof.

1. Observe that
$$(f + g)^+ \leq f^+ + g^+$$
Since f^+ and g^+ are nonnegative, we know that $I(f^+ + g^+) = I(f^+) + I(g^+) < \infty$. The finiteness of $I(f^+)$ and $I(g^+)$ follow from the fact that f and g are both integrable. Finally, the monotonicity of the integral shows us that $(f + g)^+$ is integrable. An analogous argument shows that $(f + g)^-$ is integrable. Therefore, $f + g$ is integrable.

   If $c = 0$, then we are done because the integral of the zero function is 0. Assume that $c > 0$. Then we have $(cf)^+ = cf^+$. Then we may argue using the linearity of the integral that cf^+ is integrable. A similar argument holds for $(cf)^-$. Assume $c < 0$.

Then, we have

$$(cf)^+ = -cf^-$$

and we use the same arguments from before. To obtain the equality above, we think about when cf is positive, this happens when f is negative. We need to add a negative sign at the front to make the expression positive.

2. By definition, we know that

$$\int_{A\cup B} f\, d\mu = \int \mathbf{1}_{A\cup B} f\, d\mu$$

We see that

$$(\mathbf{1}_{A\cup B} f)^+ = \mathbf{1}_{A\cup B} f^+ = \mathbf{1}_A f^+ + \mathbf{1}_B f^+$$

Since f is integrable, so is $\mathbf{1}_A f$ and $\mathbf{1}_B f$. Also, the sum of integrable functions is integrable from the previous part. We compute

$$\begin{aligned}\int \mathbf{1}_{A\cup B} f\, d\mu &= \int \left(\mathbf{1}_{A\cup B} f^+\right)^+ d\mu - \int (\mathbf{1}_{A\cup B} f)^-\, d\mu \\ &= \int \mathbf{1}_A f^+ d\mu + \int \mathbf{1}_B f^+ d\mu - \int \mathbf{1}_A f^- d\mu - \int \mathbf{1}_B f^- d\mu \\ &= \int \mathbf{1}_A f\, d\mu + \int \mathbf{1}_B f\, d\mu\end{aligned}$$

3. Recall that $f < \infty$ a.s. means that there exists a measurable $E \subseteq \{|f| < \infty\}$, we have $\mu(E^c) = 0$. Recall that if f is measurable, then $|f|$ is measurable. Since $|f| = f^+ + f^-$, it is the sum of integrable functions and thus integrable. Also,

$$\int n \cdot \mathbf{1}_{\{|f|\geq n\}}\, d\mu \leq \int_{\{|f|\geq n\}} |f|\, d\mu \leq \int |f|\, d\mu$$

The second inequality follows from the monotonicity of the integral on nonnegative, measurable functions. Moreover, since $n \cdot \mathbf{1}_{\{|f|\geq n\}}$ is simple, we have $\int n \cdot \mathbf{1}_{\{|f|\geq n\}} = n \cdot \mu(\{|f| \geq n\})$. Putting these results together

$$\mu(\{|f| \geq n\}) \leq \frac{1}{n} \int |f|\, d\mu$$

Call $A_n = \{|f| \geq n\}$. This is a decreasing sequence of sets since $A_n \supseteq A_{n+1}$ and $\bigcap_{n\geq 1} A_n = \{|f| = \infty\}$. Observe that A_n is finite for any $n \in \mathbb{N}$, based on the above inequality and the fact that $|f|$ is integrable. Using continuity from above, we see that

$$\mu(\{|f| = \infty\}) = \lim_n \mu(\{|f| \geq n\}) = \lim_{n\to\infty} \frac{1}{n} \int |f|\, d\mu = 0$$

Take $E^c = \{|f| = \infty\}$. This set is measurable as it is a countable intersection of measurable sets. Further, the complement of E^c is exactly the set of points where $|f|$ is finite, so our claim follows.

4. We omit the proof. See 38:11 in Lecture 14 for details. The key idea is the prove that for integrable $f, g \geq 0$, we have

$$\int (f - g)\, d\mu = \int f\, d\mu - \int g\, d\mu$$

After we have this, we use the decomposition $f + g = f^+ + g^+ - (f^- + g^-)$.

■

14 Properties of the Integral 2

We continue proving some common properties of the integral. We assume the setting of the previous proposition.

Proposition. Let f, g be integrable.

1. $|\int f\,d\mu| \le \int |f|\,d\mu$
2. For $c \in \mathbb{R}$, we have cf integrable and $\int cf\,d\mu = c\int f\,d\mu$.
3. The integral is monotone. So if $f \ge 0$, then $\int f\,d\mu \ge 0$ and if $f \ge g$, then $\int f\,d\mu \ge \int g\,d\mu$.
4. Suppose $f \ge 0$. If $\int f\,d\mu = 0$, then $f = 0$ a.e.
5. If $f = g$ a.e., then $\int f\,d\mu = \int g\,d\mu$.
6. Let $h : \Omega \to \overline{\mathbb{R}}$, where h is measurable and $|h| \le f$, then h is integrable.

Proof.

1. Observe that
$$\begin{aligned}\left|\int f\,d\mu\right| &= \left|\int f^+\,d\mu - \int f^-\,d\mu\right| \le \int f^+\,d\mu + \int f^-\,d\mu \\ &= \int (f^+ + f^-)\,d\mu = \int |f|\,d\mu\end{aligned}$$

2. To see that cf is integrable, notice that
$$\int |cf|\,d\mu = \int |c||f|\,d\mu = |c|\int |f|\,d\mu < \infty$$
Since f is integrable, $\int |f|\,d\mu$ is finite, hence cf is integrable. We proved this in the previous lecture as well. Now, we verify the identity. If $c = 0$, then both sides of the identity are 0. Assume $c > 0$. We know that
$$\begin{aligned}\int (cf)\,d\mu &= \int (cf)^+\,d\mu - \int (cf)^-\,d\mu = \int cf^+\,d\mu - \int cf^-\,d\mu \\ &= c\int f^+\,d\mu - c\int f^-\,d\mu = c\left(\int f^+\,d\mu - \int f^-\,d\mu\right) \\ &= c\int f\,d\mu\end{aligned}$$
The case where $c < 0$ is similar, there is just an extra negative sign that comes out.

3. Assume $f \geq 0$. We have
$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu = \int f^+\,d\mu \geq 0$$
One thing to notice is that the definition of an integrable function, when applied to a integrable function that is nonnegative, returns us the definition of the integral as the limit of the integral of a sequence of simple functions that converges to our target function.
For the second part, observe that $f - g \geq 0$, so from the first part we have
$$\int f\,d\mu - \int g\,d\mu = \int (f-g)\,d\mu \geq 0$$
The conclusion follows since f, g are integrable, so their integrals are real numbers and we can rearrange the above equation.

4. Denote the set $E_n = \{x \in \Omega : f(x) \geq \frac{1}{n}\}$. Now, we know that $\mathbf{1}_{E_n} f \leq f$, so by the monotonicity of the integral, we have
$$\frac{1}{n}\mu(E_n) = \int \frac{1}{n}\cdot \mathbf{1}_{E_n}\,d\mu \leq \int \mathbf{1}_{E_n} f\,d\mu \leq \int f\,d\mu = 0$$
The above shows us that $\frac{1}{n}\mu(E_n) \leq 0$, hence $\mu(E_n) = 0$. Call $E = \bigcup_{n\geq 1} E_n$. By the subadditivity of measure, we have $\mu(E) \leq \sum_{n\geq 1}\mu(E_n) = 0$. One may make a routine verification that
$$E = \{x \in \Omega : f(x) > 0\}$$
Since we proved $\mu(E) = 0$, we see that $f = 0$ a.e.

5. Since $f = g$ a.e., there exists a measurable set $E = \{\omega \in \Omega : f(\omega) = g(\omega)\}$ so that $\mu(E^c) = 0$. We may write
$$\begin{aligned}\int f\,d\mu &= \int (\mathbf{1}_E + \mathbf{1}_{E^c}) f\,d\mu = \int (\mathbf{1}_E f + \mathbf{1}_{E^c} f)\,d\mu \\ &= \int \mathbf{1}_E f\,d\mu + \int \mathbf{1}_{E^c} f\,d\mu \\ &= \int \mathbf{1}_E g\,d\mu\end{aligned}$$
Above, we have used a previous proposition that the integral over a set with measure zero is 0. Since $\mu(E^c) = 0$, we may rewrite the last expression as
$$\int \mathbf{1}_E g\,d\mu + \int \mathbf{1}_{E^c} g\,d\mu$$
Following our steps back, we see that the sum above is equal to $\int g\,d\mu$. Note that there is some imprecision in this proof. We should prove it on the positive and negative parts of f and g and then extend it.

6. To show that h is integrable, we must show that $\int h^+ \, d\mu < \infty$ and $\int h^- \, d\mu < \infty$. We have $h^+ \leq |h| \leq f$, so by monotonicity

$$\int h^+ \, d\mu \leq \int f \, d\mu < \infty$$

We know that the last integral is finite since we assume f is integrable. An analogous argument holds for h^-.

This concludes the proof. ■

Corollary. Suppose that $f : \Omega \to \overline{\mathbb{R}}$ is measurable and $|f| \leq c$ on E, with $\mu(E) < \infty$ and $|f| = 0$ on E^c. Then, f is integrable. Note that $c \in \mathbb{R}$ is some bounding constant.

Proof. Since $f^+ \leq |f|$ and $f^- \leq |f|$, it is enough to show that $\int |f| \, d\mu$ is finite. Observe that

$$\begin{aligned} \int |f| \, d\mu &= \int (\mathbf{1}_E + \mathbf{1}_{E^c})|f| \, d\mu = \int \mathbf{1}_E |f| \, d\mu + \int \mathbf{1}_{E^c} |f| \, d\mu \\ &\leq \int c\mathbf{1}_E \, d\mu + 0 = c \cdot \mu(E) < \infty \end{aligned}$$

So, since $|f|$ is integrable, so is f. ■

15 Theorems on the convergence of integrals

Theorem. *Monotone Convergence Theorem*
Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $f_n : \Omega \to \overline{\mathbb{R}}$ be a sequence of nonnegative, measurable functions so that $f_n \uparrow f$. Then

$$\int f_n \, d\mu \uparrow \int f \, d\mu$$

Remark. Convergence is in the sense that, if f is integrable, the integrals $\int f_n \, d\mu$ converge to $\int f \, d\mu$; while if f is not integrable either f_n is integrable for all n and $\int f_n \, d\mu \to +\infty$ as $n \to \infty$, or there is an integer N such that f_N is not integrable so that $\int f_n \, d\mu = +\infty$ for $n \geqslant N$.

Proof. We know that f must be measurable as it is the limit of measurable functions and $f \geq 0$. Since the sequence of functions f_n are all measurable and nonnegative, there exist sequences of nonnegative simple functions that increase monotonically to each f_n and to f. Denote each such sequence

$$g_{nk} \uparrow f_n$$

For instance, for f_1, we have the sequence $g_{11}, g_{12}, g_{13}, \dots$ and

$$\int f_1 \, d\mu = \lim_{k \to \infty} \int g_{1k} \, d\mu$$

These facts apply to each f_n and $\int f_n \, d\mu$, so generally,

$$\int f_n \, d\mu = \lim_{k \to \infty} \int g_{nk} \, d\mu$$

Now, our goal is to construct sequence of simple functions g_n which converges to f, so that we will be able to represent the integral of f using these functions. We may arrange our sequences of simple functions as below

$$\begin{matrix} g_{11} & g_{12} & g_{13} & \cdots \\ g_{21} & g_{22} & g_{23} & \cdots \\ g_{31} & g_{32} & g_{33} & \cdots \end{matrix}$$

To construct our sequence g_k, we take

$$g_1 = g_{11}, g_2 = \max\{g_{12}, g_{22}\}, g_3 = \max\{g_{13}, g_{23}, g_{33}\}, \dots$$

In general, we have

$$g_k = \max\{g_{nk} : 1 \leq n \leq k\}$$

Above, k is fixed and we imagine taking the maximum down a column. We must verify a few properties of our constructed sequence.

1. g_k are simple functions. Since each g_k is the maximum of as finite set of simple functions, it is simple.

2. We must have $g_k \leq g_{k+1}$. But this follows by definition since

$$g_{k+1} = \max(\{g_{n,k+1} : 1 \leq n \leq k\} \cup \{g_{k+1,k+1}\}) \geq \max\{g_{n,k} : 1 \leq n \leq k\} = g_k$$

3. We should have $\lim_{k\to\infty} g_k = f$. First, notice that $g_k \leq f_k$. If we fix movement along a row, for $n \leq k$ we have

$$g_{nk} \leq f_n \leq f_k$$

   These inequalities follow since $g_{nk} \uparrow f_n$ by construction and the sequence of f_n is monotonically increasing by assumption. Since this holds for all n, it follows that $g_k \leq f_k$.

Fix some $m \in \mathbb{N}$ so that $m \leq k$. Observe that we have

$$g_{mk} \leq g_k \leq f_k$$

We know that $f_k \uparrow f$. Since g_k is an increasing sequence that is bounded above, its limit must exist, so we may define g so that $g_k \uparrow g$. But notice that $\lim_{m\to\infty} g_{mk} = f_m$. Hence, in the limit, we have

$$f_m \leq g \leq f$$

Since these inequalities hold for any m (we may always choose a $k \geq m$ if we consider $\lim_{k\to\infty} g_k = g$), then taking the limit in m shows us that $g = f$. Thus, $g_k \uparrow f$. Now, since the integral independent of which sequence of simple functions is taken, we have

$$\int f\, d\mu = \lim_{k\to\infty} \int g_k\, d\mu$$

Now, since $f_n \leq f$, by monotonicity, we have

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

Since this holds for all $n \in \mathbb{N}$, we have $\limsup_n \int f_n\, d\mu \leq \int f\, d\mu$. The nontrivial part is to show

$$\int f\, d\mu \leq \lim \int f_n\, d\mu$$

Since we know that $g_k \leq f_k$, by monotonicity, we see that

$$\int f\, d\mu = \lim_{k\to\infty} \int g_k\, d\mu \leq \liminf_{k\to\infty} \int f_k\, d\mu$$

The conclusion follows. ■

Remark. The monotone convergence theorem also holds when $f_n \geq g$ for some integrable function g. This is because we may apply the monotone convergence theorem to $g_n := f_n - g \geq 0$ with $g_n \uparrow (f-g)$. We may obtain the identity

$$\lim_{n\to\infty} \int (f_n - g)\, d\mu = \int (f-g)\, d\mu$$

and then use the linearity of the integral to get rid of g.

Remark. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and measurable $f \geq 0$. Define

$$\mu_f(A) = \int_A f\, d\mu$$

We claim that μ_f is a measure. Observe that from a previous proposition, we have

$$\mu_f(\varnothing) = \int_\varnothing f\, d\mu = 0$$

Next, we must show that μ_f is σ-additive. Assume that $A = \bigsqcup_{i\geq 1} A_i$. Observe that

$$\begin{aligned}\mu_f(A) &= \int_A f\, d\mu = \int \mathbf{1}_A f\, d\mu \\ &= \int \sum_{i=1}^{\infty} \mathbf{1}_{A_i} f\, d\mu\end{aligned}$$

Setting $f_n := \sum_{i=1}^n \mathbf{1}_{A_i} f$, it is clear that $f_n \uparrow \mathbf{1}_A f$ (by the definition of a series). Using the monotone convergence theorem, we have

$$\begin{aligned}\int \sum_{i=1}^{\infty} \mathbf{1}_{A_i} f\, d\mu = \int \lim_{n\to\infty} \sum_{i=1}^{n} \mathbf{1}_{A_i} f\, d\mu &= \lim_{n\to\infty} \int \sum_{i=1}^{n} \mathbf{1}_{A_i} f\, d\mu \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \int \mathbf{1}_{A_i} f\, d\mu \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mu_f(A_i)\end{aligned}$$

where the penultimate equality where we exchange the finite sum and integral follows by the linearity of the integral. Therefore, μ_f is a measure.

Observation. In the above setting where f is nonnegative and measurable, let $B \in \mathcal{F}$ and $\mu(B) = 0$. We know that $\mu_f(B) = \int_B f\, d\mu = 0$. Hence if $\mu(B) = 0$, then $\mu_f(B) = 0$. This leads us to the definition of absolute continuity.

Definition. *Absolute continuity*
Let μ, ν be two measures defined on a measurable space $(\Omega, \mathcal{F})$. We say that μ is **absolutely continuous** with respect to ν if

$$\nu(A) = 0 \implies \mu(A) = 0$$

We write $\mu \ll \nu$ to indicate this.

Remark. In the previous observation, we have shown that $\mu_f \ll \mu$.

Aside. Absolute continuity is an important concept which leads into the Radon-Nikodym theorem. Let μ, ν be measures and ν be σ-finite. This theorem says that if $u \ll v$, then there exists nonnegative $g := \frac{d\mu}{d\nu}$ such that

$$\mu(B) = \int_B g \, d\nu$$

In the context of the previous observation, the Radon-Nikodym theorem gives us the existence of a function f so that

$$\mu_f(B) = \int_B f \, d\mu$$

Examples. We discuss some examples where the measures are not absolutely continuous. Suppose that we are working in $\mathbb{R}$ with a function $f > 0$. Define the measure

$$\mu_f(A) = \int_A f \, d\lambda$$

Define the Dirac measure δ_x, for $x \in \mathbb{R}$, by

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

- For $\{x\}$ we see that $\mu_f(\{x\}) = 0$, since countable sets have Lebesgue measure zero, and we have $\delta_x(\{x\}) = 1$. Hence, we see that δ_x is not absolutely continuous with respect to μ_f.
  On the other hand, take an interval $[a, b)$ such that $x \notin [a, b)$. Then $\delta_x([a, b)) = 0$ but $\mu_f([a, b)) > 0$, if we assume that f is bounded below by some constant. Hence, μ_F is not absolutely continuous with respect to δ_x. Here, we have an example of two measures that are both not absolutely continuous with respect to one another.

Remark. It may be shown that any function which is integrable is also uniformly integrable. Suppose that f is integrable. Consider the expression $\int_A f \, d\mu$. The precise statement of uniform integrability is that for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $\mu(A) < \delta$, then $\mu_f(A) < \varepsilon$. The uniform parameter is δ here, as we are proving this property for all sets of measure less than δ. We must obtain an estimate, pick some $n \in \mathbb{N}$ and observe that

$$\begin{aligned}\int_A f\,d\mu &= \int_A f\mathbf{1}_{\{f\geqslant n\}}\,d\mu + \int_A \mathbf{1}_{\{f<n\}} f\,d\mu \\ &\leq \int f\mathbf{1}_{\{f\geqslant n\}}\,d\mu + n\mu(A) \\ &= \int f\,d\mu - \int f\mathbf{1}_{\{f<n\}}\,d\mu + n\mu(A)\end{aligned}$$

Since $f\mathbf{1}_{\{f<n\}} \uparrow f$, by monotone convergence theorem, we may find some $n_0 \in \mathbb{N}$ so that these two expressions are arbitrarily close, that is $\int f\,d\mu - \int f\mathbf{1}_{\{f<n_0\}}\,d\mu \leq \frac{\varepsilon}{2}$. For the remaining summand $n_0\mu(A)$, we may simply choose δ so that $\mu(A) < \delta < \frac{\varepsilon}{2}$.

Theorem. *Fatou's lemma*
Suppose that we have a sequence of measurable functions $f_n \geq 0$. We emphasize that these functions do not have to be monotonic. Then

$$\int \liminf_n f_n\,d\mu \leq \liminf_n \int f_n\,d\mu$$

Example. We first consider an example where this limit may be strict. Define the function f_n by

$$f_n(x) = \begin{cases} 1 & \text{if } x \geq n \\ 0 & \text{otherwise} \end{cases}$$

We see that $\int f_n\,d\mu = +\infty$ for all $n \in \mathbb{N}$, but $\int \liminf_n f_n\,d\mu = 0$. So, we have a strict inequality.

Proof. Take $g_k := \inf_{n\geq k} f_n$. We see that $g_k \geq 0$ and $g_k \leq g_{k+1}$ since we are taking the infimum over smaller sets. Also, $g_k \uparrow \liminf_n f_n$. Now, by the monotone convergence theorem, we have

$$\int \liminf_n f_n\,d\mu = \int \lim_{k\to\infty} g_k\,d\mu = \lim_{k\to\infty} \int g_k\,d\mu \leq \liminf_{k\to\infty} \int f_k\,d\mu$$

The last inequality follows since $g_k \leq f_k$ by definition. To be explicit, for any $n \geq k$, we have $g_k \leq f_n$, so by the monotonicity of the integral,

$$\int g_k\,d\mu \leq \int f_n\,d\mu \ \ (\forall n \geq k) \implies \int g_k\,d\mu \leq \inf_{n\geq k} \int f_n\,d\mu$$

Finally taking the limit as $k \to \infty$, we obtain the claimed inequality. ■

Remark. Similarly to the monotone convergence theorem, to obtain the conclusion, it is enough to require that $f_n \geq g$ for some integrable function g.

Corollary. If $\{f_n\}$ is a sequence of measurable functions which is bounded above by an integrable function, then

$$\int \limsup_{n\to\infty} f_n \, d\mu \geqslant \limsup_{n\to\infty} \int f_n \, d\mu$$

Theorem. *Dominated Convergence Theorem*
Suppose that f_n is a sequence of measurable functions such that $f_n \to f$ and that $|f_n| \leq g$ where g is integrable. Then, f is integrable and

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

Proof. First, note that f_n is integrable since $|f_n| \leq g$ and g is integrable. Additionally, we know that f is integrable since $|f_n| \leq g$ for all $n \in \mathbb{N}$, which means $|f| = \lim_n |f_n| \leq g$. Since $|f_n| \leq g$, we have $f_n \leq g$ and $f_n \geq -g$. By the corollary, we have

$$\limsup_n \int f_n \, d\mu \leq \int \limsup_n f_n \, d\mu = \int f \, d\mu = \int \liminf_n f_n \, d\mu \leq \liminf_n \int f_n \, d\mu$$

■

16 Product Measures

Let $(\Omega_j, \mathcal{F}_j, \mu_j)$ be measure spaces. Suppose that $j = 1, 2$ for the moment. Define the Cartesian product as usual

$$\Omega = \Omega_1 \times \Omega_2 = \{(x, y) : x \in \Omega_1, y \in \Omega_2\}$$

We want to define a σ-algebra on Ω which is generated by rectangles.

Definition. *Rectangles*
Let $E_j \in \mathcal{F}_j$. Call the sets

$$E_1 \times E_2 = \{(x, y) : x \in E_1, y \in E_2\}$$

as rectangles.

Call $\mathcal{F}_1 \times \mathcal{F}_2$ as the set of all rectangles and define $\mu(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2)$. To be explicit

$$\mathcal{F}_1 \times \mathcal{F}_2 = \{E_1 \times E_2 : E_j \in \mathcal{F}_j\}$$

The goal will be to start from rectangles, define a measure μ through the preceding formula, and to extend this measure to the σ-algebra generated by the rectangles through Carathéodory's extension theorem.

Proposition. $\mathcal{F}_1 \times \mathcal{F}_2$ is a semi-algebra.

Proof. Clearly, $\Omega_1 \times \Omega_2 \in \mathcal{F}_1 \times \mathcal{F}_2$. Now, we must prove the second property of a semi-algebra holds, that is that the complement of every set may be written as the finite disjoint union of sets in the semi-algebra.

Suppose that $A \in \mathcal{F}_1 \times \mathcal{F}_2$. Then, $A = E_1 \times E_2$, for some $E_j \in \mathcal{F}_j$. Observe that

$$A^c = (E_1 \times E_2^c) \cup (E_1^c \times E_2) \cup (E_1^c \times E_2^c)$$

Thus, since $\mathcal{F}_j$ are σ-algebras, we see that A^c is indeed the finite, disjoint union of sets in $\mathcal{F}_1 \times \mathcal{F}_2$. ■

Remark. An intuitive way to understand this is to draw the rectangle in $\mathbb{R}^2$ and reason about sets that are outside of the rectangle. However, it is important to realize that the generic rectangle is a much more diverse class, since the sets E_1, E_2 in the case where $\Omega_j = \mathbb{R}$ may be any Lebesgue measurable or Borel measurable sets.

Above, we have shown that $\mathcal{F}_1 \times \mathcal{F}_2$ is a semi-algebra.

- We will denote the generated σ using the $\star$ operator

$$\mathcal{F}_1 \star \mathcal{F}_2 = \mathcal{F} = \sigma(\mathcal{F}_1 \times \mathcal{F}_2)$$

- We define a measure in the semi-algebra, $\mu : \mathcal{F}_1 \times \mathcal{F}_2 \to [0, \infty]$ by

$$E_1 \times E_2 \mapsto \mu_1(E_1)\mu_2(E_2)$$

  where we use the convention that $0 \cdot \infty = 0$.

- If we want to apply Carathéodory's extension theorem, we will need to show that μ is σ-additive on the semi-algebra.

Definition. *Section*
Suppose that $A \subseteq \Omega = (\Omega_1 \times \Omega_2)$. Fix $x \in \Omega_1$ and $y \in \Omega_2$. We may define the sections A_x and A^y of A by

$$A_x = \{y \in \Omega_2 : (x, y) \in A\} \subseteq \Omega_2$$
$$A^y = \{x \in \Omega_1 : (x, y) \in A\} \subseteq \Omega_1$$

Note that A is any subset of Ω, of course, if it belongs to the σ-algebra, we may say more about it.

Lemma. Let $A \in \mathcal{F}_1 \star \mathcal{F}_2 = \mathcal{F}$. Then, for all $x \in \Omega_1$, we have $A_x \in \mathcal{F}_2$ and for any $y \in \Omega_2$ we have $A^y \in \mathcal{F}_1$.

Proof. We want to show a certain property holds for any set in a σ-algebra. A common technique is to define a collection $\mathscr{C}$ of sets which satisfies the desired property, show that a simpler collection $\mathscr{D}$ of sets belongs to that collection $\mathscr{C}$, and then prove that our defined collection $\mathscr{C}$ is a σ-algebra so that the generated σ-algebra $\sigma(\mathscr{D})$ also has that property.
Call the first property we want to prove, that for all $x \in \Omega_1$ we have $A_x \in \mathcal{F}_2$, as property (α). Define the class of sets

$$\mathscr{C} = \{A \in \mathcal{F}_1 \star \mathcal{F}_2 : A \text{ satisfies } (\alpha)\}$$

We verify two properties about $\mathscr{C}$

1. We claim $\mathscr{C} \supseteq \mathcal{F}_1 \times \mathcal{F}_2$. Let $A \in \mathcal{F}_1 \times \mathcal{F}_2$, that is, A is a rectangle. So, we may write $A = E_1 \times E_2$ for sets $E_1 \in \mathcal{F}_1$ and $E_2 \in \mathcal{F}_2$. Let $x \in \Omega_1$. It may be verified that

$$A_x = \begin{cases} E_2 & x \in E_1 \\ \varnothing & x \in E_1^c \end{cases}$$

   So, regardless of the value of x, we always have $A_x \in \mathcal{F}_2$. Hence, $A \in \mathscr{C}$ and we see that $\mathscr{C}$ contains the rectangles.

2. $\mathscr{C}$ is a σ-algebra. First, we show that $\Omega \in \mathscr{C}$. Pick $x \in \Omega_1$ and consider the section Ω_x. But, based on the proof in (1), we see that

$$\Omega_x = (\Omega_1 \times \Omega_2)_x = \Omega_2 \in \mathcal{F}_2$$

Next, we need to show closure under complementation. Let $A \in \mathscr{C}$ and fix $x \in \Omega_1$. By definition, we have $A_x \in \mathcal{F}_2$. But since $\mathcal{F}_2$ is a σ-algebra, it follows that $(A_x)^c \in \mathcal{F}_2$. We are done if we can show that $(A^c)_x = (A_x)^c$. Observe that

$$\begin{aligned}(A^c)_x &= \{y \in \Omega_2 : (x,y) \in A^c\} = \{y \in \Omega_2 : (x,y) \notin A\} \\ &= \Omega_2 \setminus \{y \in \Omega_2 : (x,y) \in A\} = (A_x)^c\end{aligned}$$

It remains to show closure under countable unions. Let $A_j \in \mathscr{C}$ for $j \in \mathbb{N}$. Observe that

$$\begin{aligned}\left(\bigcup_j A_j\right)_x &= \left\{y \in \Omega_2 : (x,y) \in \bigcup_j A_j\right\} \\ &= \bigcup_{j\geq 1} \{y \in \Omega_2 : (x,y) \in A_j\} \\ &= \bigcup_{j \geq 1} (A_j)_x\end{aligned}$$

But, our assumption is that every A_j satisfies $(A_j)_x \in \mathcal{F}_2$. Hence, their countable union is in $\mathcal{F}_2$ and we conclude that $\bigcup_j A_j \in \mathscr{C}$.

Since $\mathscr{C} \supseteq \mathcal{F}_1 \times \mathcal{F}_2$ and $\mathscr{C}$ is a σ-algebra, we see that $\mathscr{C} \supseteq \mathcal{F}_1 \star \mathcal{F}_2$. Hence, any set $A \in \mathcal{F}_1 \star \mathcal{F}_2$ satisfies (α). An analogous proof holds for the sections A^y. ■

Remark. The lemma above is saying that if we pick a set A in the σ-algebra generated by the set of rectangles, the sections will be measurable in their respective σ-algebras.

Proposition. μ is additive in $\mathcal{F}_1 \times \mathcal{F}_2$.

Proof. We want to show additivity, given below

$$\begin{matrix} A, A_j \in \mathcal{F}_1 \times \mathcal{F}_2 \\ 1 \leq j \leq n \end{matrix} \quad A = \bigsqcup_{j=1}^{n} A_j \implies \mu(A) = \sum_{j=1}^{n} \mu(A_j)$$

Since A is a rectangle, we may write $A = E \times F$, and taking the section A_x, we see that

$$A_x = \begin{cases} \varnothing, & x \notin E \\ F, & x \in E \end{cases}$$

Decomposing A along A_j, we have

$$A_x = \left(\bigsqcup_{j=1}^{m} A_j\right)_x = \bigsqcup_{j=1}^{m} (A_j)_x$$

Since $A_j \in \mathcal{F}_1 \times \mathcal{F}_2$, we may write $A_j = E_j \times F_j$ for $E_j \in \mathcal{F}_1$ and $F_j \in \mathcal{F}_2$.

$$\bigsqcup_{j=1}^{m} (A_j)_x = \bigsqcup_{j=1}^{n} F_j \mathbf{1}_{E_j}(x)$$

If $x \in E$, then using the chain of equalities above, we have

$$F = \bigsqcup_{j=1}^{n} F_j \mathbf{1}_{E_j}(x)$$

Now, we see that

$$\mu_2(F) = \sum_{j=1}^{n} \mu_2(F_j) \cdot \mathbf{1}_{E_j}(x)$$

Note that if $x \notin E$, then we have a trivial identity $\varnothing = \varnothing$. Hence, to summarize both cases, we may just write

$$\mathbf{1}_E(x) \cdot \mu_2(F) = \sum_{j=1}^{n} \mu_2(F_j) \cdot \mathbf{1}_{E_j}(x)$$

Integrating both sides, we see that

$$\begin{aligned} \mu_1(E)\mu_2(F) &= \int \mathbf{1}_E(x) \cdot \mu_2(F)\, d\mu_1 = \int \sum_{j=1}^{n} \mu_2(F_j) \cdot \mathbf{1}_{E_j}(x)\, d\mu_1 \\ &= \sum_{j=1}^{n} \int \mu_2(F_j) \cdot \mathbf{1}_{E_j}(x)\, d\mu_1 = \sum_{j=1}^{n} \mu_1(E_j)\mu_2(F_j) \end{aligned}$$

which is precisely the identity we wanted to prove. ■

Remark. The argument for σ-additivity of the function μ is analogous. The only points that are changed are that we must consider a countable union of disjoint sets A_j, and that we must use the monotone convergence theorem rather than the linearity of the integral to exchange the sum and the integral in the last steps.

Assume now that μ_1 and μ_2 are σ-finite. This means that

$$\Omega_1 = \bigcup_{j\geq 1} E_j, \quad \Omega_2 = \bigcup_{j\geq 1} F_j$$

with $\mu_1(E_j) < \infty$ and $\mu_2(F_j) < \infty$ for all $j \in \mathbb{N}$ and where E_j and F_j are increasing sequences. We have

$$\Omega = \Omega_1 \times \Omega_2 = \bigcup_{j\geq 1} E_j \times F_j$$

Further

$$\mu(E_j \times F_j) = \mu_1(E_j) \cdot \mu_2(F_j) < \infty$$

Hence Ω is σ-finite with respect to μ. Finally, we are able to extend μ to a unique measure on $\mathcal{F}_1 \star \mathcal{F}_2$ via Carathéodory's extension theorem.

17 Measure on a countable product of spaces

Skip this lecture for now.

18 Fubini's Theorem

We will begin with two measure spaces $(\Omega_j, \mathcal{F}_j\, \mu_j)$ where Ω_j is σ-finite relative to μ_j. Previously, we constructed the measure $\mu = \mu_1 \star \mu_2$ on the generated σ-algebra of the sets of rectangles, which we denoted by $\mathcal{F}_1 \star \mathcal{F}_2$. Consider a function $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$. Assume $f \geq 0$ for now. Fix some $x \in \Omega_1$ and consider a function $f_x : \Omega_2 \to \overline{\mathbb{R}}$ given by $f_x(y) = f(x, y)$. We consider integration

$$\int f_x(y)\, d\mu_2(y)$$

Since this takes different values for each $x \in \Omega_1$, we may consider the above object as a function in x. As a function of x, we will show that if f is $\mathcal{F}$-measurable, then the integral is $\mathcal{F}_1$-measurable. Since $\int f_x(y)\, d\mu_2(y) \geq 0$ under our assumptions, we will see that

$$\int \left[\int f_x(y)\, d\mu_2(y) \right] d\mu_1(x) = \int f\, d\mu$$

Since this result will be symmetric, we may do this in the other order as well. We will need to prove some details:

1. For any fixed $x \in \Omega_1$, we will show that $f_x(y)$, which is a function of y, is $\mathcal{F}_2$-measurable.
2. We will also need to show that $x \mapsto \int f_x(y)\, d\mu_2(y)$ is $\mathcal{F}_1$-measurable as a function of x.
3. At the end, we must prove the identity.
4. The plan is to follow the above steps to show the result first for indicator functions, then nonnegative, and finally integrable functions.

18.1 Setup for Proving Fubini's Theorem

Let $f : \Omega \to \overline{\mathbb{R}}$ and suppose that f is $\mathcal{F}$-measurable.

Proposition. Let $x \in \Omega_1$ and consider $f_x : \Omega_2 \to \overline{\mathbb{R}}$ given by the map $y \mapsto f(x, y)$. Then f_x is an $\mathcal{F}_2$-measurable function.

Proof. To show that f_x is $\mathcal{F}_2$-measurable, we must show that $f_x^{-1}(B) \in \mathcal{F}_2$ for every $B \in \overline{\mathscr{B}}$. Notice that

$$\begin{aligned} f_x^{-1}(B) &= \{y \in \Omega_2 : f_x(y) \in B\} = \{y \in \Omega_2 : f(x, y) \in B\} \\ &= \left\{y \in \Omega_2 : (x, y) \in f^{-1}(B)\right\} = \left[f^{-1}(B)\right]_x \end{aligned}$$

Since f is $\mathcal{F}$-measurable and $B \in \overline{\mathscr{B}}$, this means that $f^{-1}(B) \in \mathcal{F}$. We have shown previously that the section of a measurable set belongs to its respective σ-algebra, hence $[f^{-1}(B)]_x \in \mathcal{F}_2$. ∎

Theorem. *Facts about measurability on sections*
Let $(\Omega_j, \mathcal{F}_j, \mu_j)$ be a measure space for $j = 1, 2$, where Ω_j are σ-finite with respect to μ_j. Let $E \in \mathcal{F} = \mathcal{F}_1 \star \mathcal{F}_2$. Then

1. The function $x \mapsto \mu_2(E_x)$ is $\mathcal{F}_1$-measurable. The function $y \mapsto \mu_1(E^y)$ is $\mathcal{F}_2$-measurable.

2. We have the following identity

$$\int \mu_2(E_x)\, d\mu_1 = \mu(E) = \int \mu_1(E^y)\, d\mu_2$$

Proof.

1. We want to show that $x \mapsto \mu_2(E_x)$ is $\mathcal{F}_1$-measurable for all $E \in \mathcal{F}$. First, assume that $\mu_1(\Omega_1) < \infty$ and $\mu_2(\Omega_2) < \infty$. We will first prove the assertion for rectangles. Then, we will extend by linearity to the algebra generated by the rectangles. Finally, we will use the monotone class theorem to extend this to all sets in the σ-algebra.

   (a) Assume $E = A \times B$ where $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$ (i.e. E is a rectangle). We have

   $$E_x = \begin{cases} B & x \in A \\ \varnothing & x \notin A \end{cases}$$

   Applying the measure, we see that

   $$\mu_2\left(E_x\right) = \begin{cases} \mu_2(B) & x \in A \\ 0 & x \notin A \end{cases} = \mu_2(B) \cdot \mathbf{1}_A(x)$$

   So, we see that $\mu_2(E_x)$ is some multiple of the indicator $\mathbf{1}_A(x)$. Since $A \in \mathcal{F}_1$, this means the indicator function is $\mathcal{F}_1$-measurable. Thus, $x \mapsto \mu_2(E_x)$ is $\mathcal{F}_1$-measurable.

   (b) Call $\mathscr{C}$ the set of rectangles. Call $\mathcal{A} = \mathcal{A}(\mathscr{C})$ the generated algebra. Taking $E \in \mathcal{A}$, we may write

   $$E = \bigsqcup_{j=1}^{n} E_j$$

   where $E_j \in \mathscr{C}$. Since E_j are rectangles, we may write $E_j = A_j \times B_j$ for $A_j \in \mathcal{F}_1$ and $B_j \in \mathcal{F}_2$. So

   $$E_x = \left(\bigsqcup_{j=1}^{n} E_j\right)_x = \bigsqcup_{j=1}^{n} (E_j)_x$$

Now, since $E \in \mathcal{F}$, we know that $E_x \in \mathcal{F}_2$. Hence, the expression $\mu_2(E_x)$ is defined and

$$\mu_2(E_x) = \sum_{j=1}^{n} \mu_2((E_j)_x) = \sum_{j=1}^{n} \mu_2(B_j)\mathbf{1}_{A_j}(x)$$

The last equality follows from the first part, where we have shown that

$$\mu_2((E_j)_x) = \mu_2(B_j)\mathbf{1}_{A_j}(x)$$

Since $\mu_2(E_x) = \sum_{j=1}^{n} \mu_2(B_j)\mathbf{1}_{A_j}(x)$ and $\mu_2(B_j) \geq 0$, it is a nonnegative, simple function. In particular, it is measurable as a function of x, hence $x \mapsto \mu_2(E_x)$ is $\mathcal{F}_1$-measurable for sets $E \in \mathcal{A}$.

(c) We now extend the result to all $E \in \mathcal{F}$. Denote

$$\mathcal{G} = \{E \in \mathcal{F} : x \mapsto \mu_2(E_x) \text{ is } \mathcal{F}_1\text{-measurable}\}$$

We have already shown that $\mathcal{G} \supseteq \mathscr{C}$. If we can show that $\mathcal{G}$ is a σ-algebra, then we are done since it will contain $\mathcal{F}$.

Aside. Showing closure under finite unions is difficult. Supposing $E, F \in \mathcal{G}$, this means that the functions $x \mapsto \mu_2(E_x)$ and $x \mapsto \mu_2(F_x)$ will both be $\mathcal{F}_1$-measurable, however, there is no way to show that

$$x \mapsto \mu_2((E \cup F)_x) = \mu_2(E_x \cup F_x)$$

will also be measurable, since E_x and F_x may have nontrivial intersection.

The way to sidestep this is to prove that $\mathcal{G}$ is a monotone class. Since $\mathcal{G} \supseteq \mathcal{A}(\mathscr{C})$, by the monotone class theorem, we will know that $\mathcal{G}$ also contains $\sigma(\mathcal{A})$.

We now show $\mathcal{G}$ is a monotone class. Assume that $E^{(n)} \in \mathcal{G}$ and $E^{(n)} \uparrow E$, where we use the upper index to avoid collision with the notation for sections. Since $E^{(n)} \in \mathcal{G}$, we see that

$$x \mapsto \mu_2(E_x^{(n)})$$

is $\mathcal{F}_1$-measurable. Further $E^{(n)} \uparrow E$ implies that $E_x^{(n)} \uparrow E_x$ for all $x \in \Omega_1$. Since μ_2 is a measure, it is σ-additive and it is continuous from below. This gives us the identity

$$\mu_2\left(E_x\right) = \lim_{n\to\infty} \mu_2\left(E_x^{(n)}\right)$$

But since $x \mapsto \mu_2(E_x^{(n)})$ are $\mathcal{F}_1$-measurable, we know that their limit is also $\mathcal{F}_1$-measurable. Hence, $E \in \mathcal{G}$.

Now, assume that $E^{(n)} \in \mathcal{G}$ and $E^{(n)} \downarrow E$. We may apply an analogous argument to see that $E \in \mathcal{G}$. Notice that we use the assumption that $\mu_2(\Omega_2) < \infty$ here in order to invoke continuity from above. Therefore, $\mathcal{G}$ is a monotone class and contains $\sigma(\mathscr{C}) = \mathcal{F}$ and this concludes the proof.

We may make an analogous proof to show $y \mapsto \mu_1(E^y)$ is $\mathcal{F}_2$-measurable. To complete the argument, we want to get rid of the assumptions that $\mu_1(\Omega_1) < \infty$ and $\mu_2(\Omega_2) < \infty$. To deal with this case, we need to use the σ-finiteness of the spaces.

Since Ω_1 is σ-finite with respect to μ_1, there exist $A_n \in \mathcal{A}_1$ (the algebra, but recall that rectangles involve sets from $\mathcal{F}_1$) so that $\Omega_1 = \bigcup A_n$ and $\mu_1(A_n) < \infty$ for all $n \in \mathbb{N}$. Similarly, there exist $B^{(n)} \in \mathcal{A}_2$ so that $\Omega_2 = \bigcup B^{(n)}$ and $\mu_2(B^{(n)}) < \infty$ for all $n \in \mathbb{N}$. In the next expressions, note that we may change the positioning of the index for n liberally: these are still the same sets. Thus, we have $F_n = A_n \times B_n \in \mathscr{C}$ and

$$\mu(F_n) = \mu_1(A_n)\mu_2(B^{(n)}), \quad \bigcup_{n \geq 1} F^{(n)} = \Omega_1 \times \Omega_2$$

We may assume that the sets A_n and B_n are increasing, and from this, it is easy to check the second equality above. Observe that $E \cap F^{(n)}$ has finite measure since $F^{(n)}$ has finite measure. By the first part of the proof, we know that

$$x \mapsto \mu_2((E \cap F^{(n)})_x]$$

is $\mathcal{F}_1$ measurable for all $n \in \mathbb{N}$. Since $\bigcup_{n \geq 1} F^{(n)} = \Omega_1 \times \Omega_2$, we see that $(E \cap F^{(n)})_x \uparrow E_x$. As a result, this means that

$$\mu_2((E \cap F^{(n)})_x) \uparrow \mu_2(E_x)$$

Since $x \mapsto \mu_2((E \cap F^{(n)})_x)$ are all $\mathcal{F}_1$-measurable, so is its limit $x \mapsto \mu_2(E_x)$.

2. We aim to prove

$$\int \mu_2(E_x)\, d\mu_1 = \mu(E)$$

From the first part, we know that $x \mapsto \mu_2(E_x)$ is a nonnegative, $\mathcal{F}_1$-measurable function, so this integral makes sense. We will proceed in steps again.

(a) Suppose that $E \in \mathscr{C}$ so that it is a rectangle, and we may write $E = A \times B$ for $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$. In the previous part, we have already shown

$$u_2(E_x) = \mu_2(B)\mathbf{1}_A(x)$$

Taking the integral

$$\int u_2(E_x)\, d\mu_1 = \int \mu_2(B)\mathbf{1}_A(x)\, d\mu_1 = \mu_2(B)\mu_1(A) = \mu(E)$$

(b) Suppose that $E \in \mathcal{A}(\mathscr{C})$. This means that $E = \bigsqcup_{j=1}^{n} E^{(j)}$, for $E^{(j)} \in \mathscr{C}$. Again, from the previous part, we have

$$\mu_2\left(E_x\right) = \sum_{f=1}^{n} \mu_2\left(E_x^{(j)}\right)$$

Integrating again

$$\int u_2(E_x)\,d\mu_1 = \int \sum_{j=1}^{n} \mu_2\left(E_x^{(j)}\right)\,d\mu_1 = \sum_{j=1}^{n} \int \mu_2\left(E_x^{(j)}\right)\,d\mu_1$$
$$= \sum_{j=1}^{n} \mu(E^{(j)}) = \mu(E)$$

where we use the linearity of the integral to interchange the sum and integral and part (a) to yield the last equalities.

(c) Now, we want to show that the identity holds for any $E \in \mathcal{F}$. Assume that $\mu_1(\Omega_1) < \infty$ and $\mu_2(\Omega_2) < \infty$. Define

$$\mathcal{G} = \left\{E \in \mathcal{F} : \int \mu_2(E_x)\,d\mu_1 = \mu(E)\right\}$$

We claim that $\mathcal{G}$ is a monotone class. Let $E_n \in \mathcal{G}$ and $E_n \uparrow E$. Since $E_n \in \mathcal{G}$, we see that

$$\int \mu_2(E_x^{(n)})\,d\mu_1 = \mu(E^{(n)})$$

for all $n \in \mathbb{N}$. Further, since μ_2 is a measure and $E_x^{(n)} \uparrow E_x$, we have $\lim_{n\to\infty} \mu_2(E_x^{(n)}) = \mu_2(E_x)$. By analogous reasoning, we have $\lim_{n\to\infty} \mu(E^{(n)}) = \mu(E)$. Now, we take the limit on both sides and apply the monotone convergence theorem to see that

$$\int \mu_2(E_x)\,d\mu_1 = \int \lim_{n\to\infty} \mu_2(E_x^{(n)})\,d\mu_1 = \lim_{n\to\infty} \int \mu_2(E_x^{(n)})\,d\mu_1 = \lim_{n\to\infty} \mu(E^{(n)}) = \mu(E)$$

A similar argument shows that $\mathcal{G}$ is closed under decreasing sequences $E_n \downarrow E$ and with the finiteness assumption to give the continuity from above of measures. Further, we will need to use the dominated convergence theorem with the function $\mu_2((\Omega_1 \times \Omega_2)_x)$ to pass to the limit inside. So, the identity holds for all $E \in \mathcal{F}$ through the monotone class theorem.

It remains to remove the assumptions of finiteness. By the σ-finiteness of each space, there exists $A^{(n)} \in \mathcal{A}_1$ so that $A^{(n)} \uparrow \Omega_1$ and $\mu_1(A^{(n)}) < \infty$ for all $n \in \mathbb{N}$. Similarly, there exists $B^{(n)} \in \mathcal{A}_2$ so that $B^{(n)} \uparrow \Omega_2$ and $\mu_2(B^{(n)}) < \infty$ for all $n \in \mathbb{N}$. Setting $F^{(n)} = A^{(n)} \times B^{(n)}$, we have $F^{(n)} \in \mathscr{C}$ and $\mu(F^{(n)}) < \infty$ and

$$\bigcup_{n\in\mathbb{N}} F^{(n)} = \Omega$$

Now, let $E \in \mathcal{F}$, since $E \cap F^{(n)} \in \mathcal{F}$ and this set has finite measure, by part (c), we have

$$\int \mu_2 \left(\left[E \cap F^{(n)} \right]_x \right) d\mu_1 = \mu \left(E \cap F^{(n)} \right)$$

Taking the limit on both sides, we see that

$$\begin{aligned} \int \mu_2 \left(E_x \right) d\mu_1 &= \int \lim_{n\to\infty} \mu_2 \left(\left[E \cap F^{(n)} \right]_x \right) d\mu_1 \\ &= \lim_{n\to\infty} \int \mu_2 \left(\left[E \cap F^{(n)} \right]_x \right) d\mu_1 = \lim_{n\to\infty} \mu \left(E \cap F^{(n)} \right) = \mu(E) \end{aligned}$$

where we use monotone convergence and also continuity from below of the measure.

■

Aside. In the above proof, we use the fact that if $E_n \uparrow E$, then $E_x^{(n)} \uparrow E_x$ several times. This is true since

$$E_x^{(n)} = \{y \in \Omega_2 : (x, y) \in E_n\} = E_n \cap (\{x\} \times \Omega_2)$$

18.2 Tonelli's Theorem and Fubini's Theorem

Theorem. *Tonelli's Theorem*
Let $(\Omega_j, \mathcal{F}_j, \mu_j)$ be measure spaces for $j = 1, 2$ and let Ω_j be σ-finite with respect to μ_j. Suppose that $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ is measurable. Then

$$\begin{aligned} \int \left[\int f_x(y) d\mu_2 \right] d\mu_1 &= \int f d\mu \\ &= \int \left[\int f^y(x)\, d\mu_1 \right] d\mu_2 \end{aligned}$$

Remark. In the language of Folland, these are functions in L^+.

Proof. First, we ensure these expressions make sense. From a previous proposition, we know that f_x is an $\mathcal{F}_2$-measurable function and since f is nonnegative, it must be nonnegative. Since it is nonnegative and measurable, it may be integrated.

We prove two claims now:

a. $\int f_x(y) d\mu_2$ is $\mathcal{F}_1$-measurable.

b. The stated identity holds.

The first claim still relates to the ensuring the expression is defined. To proceed, we will first consider f as an indicator, then as a simple function, and then pass to the limit to nonnegative, measurable functions.

1. Assume that $f = c\mathbf{1}_E$ for some $c \in [0, \infty)$ and $E \in \mathcal{F}$. Then

$$f_x = (c\mathbf{1}_E)_x = c(\mathbf{1}_E)_x = c\mathbf{1}_{E_x}$$

Integrating using this identity, we have

$$\int f_x(y)\, d\mu_2 = \int c\mathbf{1}_{E_x}(y)\, d\mu_2 = c\mu_2(E_x)$$

But, from the previous proof, we know that $x \mapsto c\mu_2(E_x)$ is an $\mathcal{F}_1$-measurable function. Evaluating the double integral, we have

$$\int \left[\int f_x(y) d\mu_2 \right] d\mu_1 = \int c\mu_2\left(E_x\right) d\mu_1 = c\mu(E) = \int f\, d\mu$$

where the last equality follows since our assumption was that $f = c\mathbf{1}_E$. So our desired identity holds.

2. Assume that f is a nonnegative simple function, that is, $f = \sum_{j=1}^n c_j \mathbf{1}_{E_j}$ for $c_j \geq 0$ and $E_j \in \mathcal{F}$. Fix some $x \in \Omega_1$ and call $f^{(j)} := c_j \mathbf{1}_{E_j}$, then

$$f_x = \left(\sum_{j=1}^n f^{(j)} \right)_x = \sum_{j=1}^n f_x^{(j)}$$

Since f is simple, it is $\mathcal{F}$-measurable, so by the first proposition in this lecture, we see that f_x is $\mathcal{F}_2$-measurable. Integrating, we have

$$\int f_x(y)\, d\mu_2 = \int \sum_{j=1}^n f_x^{(j)}(y)\, d\mu_2 = \sum_{j=1}^n \int f_x^{(j)}(y)\, d\mu_2(y)$$

By the previous part of this proof, we know that $x \mapsto \int f_x^{(j)}(y)\, d\mu_2(y)$ are $\mathcal{F}_1$-measurable. We see that $\int f_x(y)\, d\mu_2$ is $\mathcal{F}_1$-measurable since the sum of $\mathcal{F}_1$-measurable functions is $\mathcal{F}_1$-measurable. Now, we may integrate with respect to μ_1

$$\int \left[\int f_x(y) d\mu_2 \right] d\mu_1 = \sum_{j=1}^n \int \left[\int f_x^{(j)}(y) d\mu_2(y) \right] d\mu_1 = \sum_{j=1}^n \int f^{(j)} d\mu = \int f d\mu$$

where the second-to-last inequality follows from the first part of the proof applied to each j.

3. Assume that $f \geq 0$ and measurable. By our lemma on simple functions, there exists some sequence $f^{(j)}$ of nonnegative simple functions such that $f^{(j)} \uparrow f$. It follows that $f_x^{(j)} \uparrow f_x$ and by the monotone convergence theorem, we know that

$$\lim_{j \to \infty} \int f_x^{(j)}\, d\mu_2 = \int f_x\, d\mu_2$$

But, the limit of $\mathcal{F}_1$-measurable functions is $\mathcal{F}_1$-measurable, which shows that $x \mapsto \int f_x \, d\mu_2$ is $\mathcal{F}_1$-measurable. Since the above is a sequence of monotonically increasing, nonnegative functions, we may apply the monotone convergence theorem again to see that

$$\int \left[\int f_x^{(j)} \, d\mu_2 \right] d\mu_1 \uparrow \int \left[\int f_x \, d\mu_2 \right] d\mu_1$$

To be explicit, we may set $g_j := \left(x \mapsto \int f_x^{(j)} \, d\mu_2 \right)$ and $g := \left(x \mapsto \int f_x \, d\mu_2 \right)$ and apply the MCT to these. By the second step, since $f^{(j)}$ are simple functions, we know that

$$\int f^{(j)} \, d\mu = \int \left[\int f_x^{(j)} \, d\mu_2 \right] d\mu_1$$

But we may apply the monotone convergence theorem once again to see that

$$\lim_{j \to \infty} \int f^{(j)} \, d\mu = \int f \, d\mu$$

This concludes the proof of Tonelli's theorem. ■

Theorem. *Fubini's Theorem*
Assume the same setting from Tonelli's theorem, except we replace $f \geq 0$ with any integrable $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$. Then

$$\begin{aligned} \int f d\mu &= \int \left[\int f_x(y) d\mu_2 \right] d\mu_1 \\ &= \int \left[\int f^y(x) d\mu_1 \right] d\mu_2 \end{aligned}$$

where we are integrating with respect to an appropriately defined function $g(x)$ or $h(y)$.

Remark. We discuss the meaning of appropriately defining g. We may write $f = f^+ - f^-$. Since f is integrable, we know that f^+ is integrable, that is $\int f^+ \, d\mu < \infty$. Applying what we know from Tonelli's theorem, it follows that

$$x \mapsto \int f_x^+(y) \, d\mu_2$$

is an $\mathcal{F}_1$-measurable function with a finite integral. Further, it must be finite almost everywhere, if not, then $\int f^+ \, d\mu = \infty$. To reiterate, $\int f_x^+(y) \, d\mu_2 < \infty$ a.e. with respect to μ_1. An analogous argument may be made for f^-. The problem is that there may be some point $x \in \Omega_1$ where $\int f_x^+(y) \, d\mu_2 = \infty$ and $\int f_x^-(y) \, d\mu_2 = \infty$, and in this case we will not be able to define their difference. To handle this issue, we define a new function

$$g^+(x) = \begin{cases} \int f_x^+(y) d\mu_2 & \text{if int } < \infty \\ 0 & \text{otherwise} \end{cases}$$

We know that $(g^+)^{-1}(\mathbb{R})$ is $\mathcal{F}_1$-measurable since $x \mapsto \int f_x^+(y)\,d\mu_2$ is $\mathcal{F}_1$-measurable. Further, if we multiply it by a characteristic function, we will obtain g^+, so g^+ is $\mathcal{F}_1$-measurable as well. In the same way, we can define $g^-(x)$ and confirm its $\mathcal{F}_1$-measurable. Now, we may define $g := g^+ - g^-$.

Proof. We show that g is integrable with respect to μ_1 and

$$\int f\,d\mu = \int g(x)\,d\mu_1$$

Recall again that $f = f^+ - f^-$ where the integrability of f implies that f^+ and f^-. are integrable. By Tonelli's theorem, we have

$$\iint f_x^+ d\mu_2\,d\mu_1 = \int f^+ d\mu < \infty$$

From the statement above, we also know that $x \mapsto \int f_x^+\,d\mu_2$ is an $\mathcal{F}_1$-integrable function. As before, define

$$g^+(x) = \begin{cases} \int f_x^+(y)\,d\mu_2 & \text{if } \int f_x^+(y)\,d\mu_2 < \infty \\ 0 & \text{otherwise} \end{cases}$$

Call

$$E = \left\{ x \in \Omega_1 : \int f_x^+(y)\,d\mu_2 < \infty \right\} \in \mathcal{F}_1$$

We see that $g^+ = \left[\int f_x^+\,d\mu_2\right]\mathbf{1}_E$ and so g^+ is $\mathcal{F}_1$-measurable since $\mathbf{1}_E$ is measurable and $x \mapsto \int f_x^+\,d\mu_2$ is measurable by Tonelli. Since the set E^c has measure zero by assumption, we have

$$\int g^+ d\mu_1 = \int \left[\int f_x^+ d\mu_2\right] d\mu_1 < \infty$$

The same holds if we consider the argument with g^- and f^-. We compute

$$\begin{aligned} \int f\,d\mu &= \int f^+ d\mu - \int f^- d\mu \\ &= \int \left[\int f_x^+ d\mu_2\right] d\mu_1 - \int \left[\int f_x^- d\mu_2\right] d\mu_1 \\ &= \int g^+ d\mu_1 - \int g^- d\mu_1 = \int (g^+ - g^-)\,d\mu_1 = \int g\,d\mu_1 \end{aligned}$$

where we are allowed to invoke linearity of the integral for g because it is finite at all values of x. ■

Remark. Thus, we see that Fubini's theorem is actually a direct consequence of Tonelli's theorem. The only issue is that we need to handle the potential cases where $\infty - \infty$ may arise for values of x (or y) by defining a new function g (or h).

Remark. Another condition under which Fubini's Theorem holds is if

$$\int \left[\int |f_x| \, d\mu_2 \right] d\mu_1 < \infty$$

If the above holds, then we know that $f_x^+, f_x^- \leq |f_x|$. Hence

$$\int \left[\int f_x^+ \, d\mu_2 \right] d\mu_1 < \infty$$

and for f_x^- as well. But, from Tonelli's theorem, we know that

$$\int f^+ \, d\mu = \int \left[\int f_x^+ \, d\mu_2 \right] d\mu_1 < \infty$$

and the same for f^-, which shows that f is integrable.

19 The Hahn-Jordan Theorem

19.1 Signed Measures

Definition. *Signed measure*
A signed measure is a function $\nu : \mathcal{F} \to \overline{\mathbb{R}}$ such that

1. $\nu(\varnothing) = 0$
2. ν is σ-additive, so for $E_j \in \mathcal{F}$ which are pairwise disjoint, we have

$$\nu\left(\bigsqcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$$

We also have some implicit requirements (or consequences) that are important to state:

1. We must have that ν is well-defined as a function, hence, ν takes at most one of the values $+\infty$ and $-\infty$ over all measurable sets. This is equivalent to requiring that ν takes at most one of the values for all pairs E, F of disjoint, measurable sets.
2. Another requirement involves σ-additivity. Since ν is well-defined, the statement

$$\nu\left(\bigsqcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$$

means that the value of ν applied to the union does not depend on the order of summation. So, $\sum_{j=1}^{\infty} \nu(E_j)$ must be absolutely convergent or the sum must diverge in the sense that one of the two sums $\sum_{j:\nu(E_j)\geq 0} \nu(E_j)$ and $\sum_{j:\nu(E_j)<0} \nu(E_j)$ is infinite and the other is finite.

Lemma. We prove some properties for signed measures. Let $E, F \in \mathcal{F}$

1. Let $E \subseteq F$. If $|\nu(E)| < \infty$, then $\nu(F \setminus E) = \nu(F) - \nu(E)$. If $\nu(E) = \pm\infty$, then $\nu(F) = \pm\infty$.
2. Let $E \in \mathcal{F}$ and $\nu(E) = +\infty$. If $F \in \mathcal{F}$, then $\nu(F) > -\infty$.
3. ν is continuous from below and continuous from above.

Remark. When considering signed measures, we may assume without loss of generality that $\nu : \mathcal{F} \to (-\infty, \infty]$ or $\nu : \mathcal{F} \to [-\infty, \infty)$, since the function may only take on at most one value of $\pm\infty$.

Proof.

1. For the first property, observe that $F = E \cup (F \setminus E)$. Using σ-additivity, we see that
$$\nu(F) = \nu(E) + \nu(F \setminus E)$$
Since $|\nu(E)| < \infty$, we may subtract it from both sides to obtain the identity. On the other hand, suppose $\nu(E) = +\infty$. By the definition of signed measure, $\nu(F \setminus E)$ must be finite or $+\infty$. Referring the the equation above, in both cases, $\nu(F) = +\infty$. An analogous proof holds for the case where $\nu(E) = -\infty$.

2. Another way to say this is if there is one measurable set with the value $+\infty$, all other sets must take the value $+\infty$ or be finite. Assume $\nu(E) = +\infty$. Let F be another set. Assume, for contradiction, that $\nu(E \cap F) = -\infty$. Then $\nu(E) = -\infty$ by the first part. ($\lightning$) Thus $\nu(E \cap F)$ is a real number or $+\infty$. Recall that
$$\nu(E) = \nu(E \cap F) + \nu(E \setminus F)$$
Assume that $\nu(E \cap F)$ is a real number. Since $\nu(E) = +\infty$, for the above identity to hold, we must have $\nu(E \setminus F) = +\infty$. Since F is disjoint from $E \setminus F$, we cannot have $\mu(F) = -\infty$ (by our implicit requirement). Lastly, consider $\nu(E \cap F) = +\infty$. Since $E \cap F \subseteq F$, by the first part, we see that $\nu(F) = +\infty$. Hence, in all cases, $\nu(F) > -\infty$.

3. Suppose $E_n \uparrow E$, where $E_n \in \mathcal{F}$. In this case, we will take $E = \bigcup_{n=1}^{\infty} E_n$. Continuity from below means that
$$\nu(E_n) \xrightarrow[n\to\infty]{} \nu\left(\bigcup_{n=1}^{\infty} E_n\right)$$
Note that the limit is no longer monotonically increasing since we may have $\nu(E_n) > \nu(E_{n+1})$, depending on what sets are involved. Define the sets
$$F_1 = E_1, \quad F_k = E_k \setminus E_{k-1}$$
We see that $F_k \in \mathcal{F}$ as the countable intersection and complements of sets in $\mathcal{F}$. Further, $F_j \cap F_k = \varnothing$ for $j \neq k$. By definition, we see that
$$\begin{aligned}\nu\left(\bigcup_{k=1}^{\infty} E_k\right) &= \nu\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \nu(F_k) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \nu(F_k) = \lim_{n\to\infty} \nu\left(\bigcup_{k=1}^{n} F_k\right) \\ &= \lim_{n\to\infty} \nu(E_n)\end{aligned}$$
This shows continuity from below.
Now, we show continuity from above. Assume $E_n \downarrow E$, where $E_n \in \mathcal{F}$. Additionally,

we assume there exists some $n_0 \in \mathbb{N}$ so that $\nu(E_{n_0})$ is finite. Define

$$F_1 = E_{n_0} \setminus E_{n_0+1}, \quad F_k = E_{n_0} \setminus E_{n_0+k}$$

Notice that $\nu(E_{n_0+k}) \neq -\infty$, otherwise, $\nu(E_{n_0}) = -\infty$ by the first part. The same reasoning shows why $\nu(E_{n_0}) \neq \infty$. Since $E_{n_0+k} \downarrow \bigcap_{j\geq 1} E_j$, it follows that $F_k \uparrow E_{n_0} \setminus \bigcap_{j\geq 1} E_j$. By the continuity from below, we have

$$\nu(E_{n_0} \setminus E_{n_0+k}) = \nu(F_k) \xrightarrow[k\to\infty]{} \nu\left(E_{n_0} \setminus \bigcap_{j\geq 1} E_j\right)$$

Since $\nu(E_{n_0}) \in \mathbb{R}$, we may write

$$\lim_{k\to\infty} \nu(E_{n_0}) - \nu(E_{n_0+k}) = \nu(E_{n_0}) - \nu\left(\bigcap_{j\geq 1} E_j\right)$$

Rearranging gives us continuity from above.

Note how the proof of the first statement gives a sort of 'monotonicity' result. ■

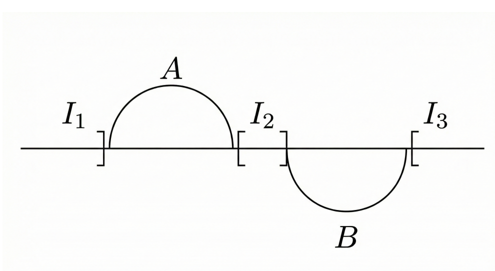
19.2 Decomposition theorems

Theorem. *Hahn decomposition of a measure*
Let $\nu : \mathcal{F} \to (-\infty, +\infty]$ be a signed measure. There exists a set $P \in \mathcal{F}$ and $N := P^c$ such that if $E \subseteq P$, then $\nu(E) \geq 0$ and if $F \subseteq N$, then $\nu(F) \leq 0$.

Remark. We discuss an example which shows that the sets P and N are not unique. Consider the measure space $(\mathbb{R}, \mathscr{B}, \lambda)$, consisting of the real line, the Borel sets and the Lebesgue measure. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and consider

$$\mu_f(E) = \int_E f \, d\lambda$$

Consider the graph of f to be given by the arcs over the sets A and B in the picture below, where it is zero everywhere else, as in the intervals I_1, I_2, I_3.



Notice that we may define $P = A \cup I$ as any of the intervals I_1, I_2, I_3 and take $N = P^c$. So decomposition is not unique. In the ensuing proof, we would like to find the sets A and B in which the positive and negative measures of space are 'concentrated'.

Proof. Let $(\Omega, \mathcal{F})$ be a measurable space and $\nu : \mathcal{F} \to (-\infty, \infty]$ be a signed measure. Define

$$\alpha = \inf_{A \in \mathcal{F}} \nu(A)$$

We claim that $\alpha > -\infty$. Note that α is defined since $\nu(\varnothing) = 0$. Assume, for contradiction, that $\alpha = -\infty$. We aim to construct a sequence of sets A_k for $k \geq 0$ satisfying $A_k \supseteq A_{k+1}$ and another sequence of sets B_k for $k \geq 1$ such that $\nu(B_k) \leq -k$.

We start our construction. Set $\Omega = A_0$. Define

$$\lambda(C) = \inf_{E \subseteq C} \nu(E)$$

This means $\lambda(A_0) = \lambda(\Omega) = \alpha = -\infty$. By the definition of infimum, there exists some set $B_1 \subseteq A_0$ such that $\nu(B_1) \leq -1$. Next, we want to choose a set A_1 so that $\lambda(A_1) = -\infty$. There are two cases.

1. Suppose that $\lambda(B_1) = -\infty$. Then set $A_1 = B_1$.
2. Otherwise, we have $\lambda(B_1) > -\infty$. Assume, for contradiction, that $\lambda(B_1 \setminus A_0) > -\infty$ as well. Suppose $E \subseteq A_0$, then

$$\begin{aligned} \nu(E) &= \nu\left(E \cap B_1\right) + \nu\left(E \cap (A_0 \setminus B_1)\right) \\ &\geq \lambda\left(B_1\right) + \lambda\left(A_0 \setminus B_1\right) \end{aligned}$$

   Since this argument holds for any E, this means that $\lambda\left(B_1\right) + \lambda\left(A_0 \setminus B_1\right)$ is a lower bound for ν. But then $\lambda(A_0) \geq \lambda\left(B_1\right) + \lambda\left(A_0 \setminus B_1\right) > -\infty$, which is impossible. Thus, $\lambda(B_1 \setminus A_0) = -\infty$ and we set $A_1 = B_1 \setminus A_0$.

In the situation where $A_1 = A_0 \setminus B_1$, we say that there exists a *bifurcation*. In the construction above, we have $A_1 \subseteq A_0$ and $\lambda(A_1) = -\infty$ and $B_1 \subseteq A_0$. We may now construct the sequences of sets A_k and B_k inductively following the process above. Notice that as a result of the construction, we have

$$A_k \supseteq A_{k+1}, \quad A_k \supseteq B_{k+1}, \quad \lambda(A_k) = -\infty, \quad \nu(B_k) \leq -k$$

Further, once there is a bifurcation, say at the k^{th} iteration. This means that $A_{k+1} = A_k \setminus B_{k+1}$. Then, from the relations above, for $j \geq k+2$ we have

$$B_{k+1} \cap B_j = \varnothing$$

Now, we will be able to construct our set B with infinite measure. There are two cases.

1. There are a finite number of bifurcations. So, there exists n_0 such that for all $n \geq n_0$ there are no bifurcations. Then, $B_n = A_n$, and since the sequence of sets A_n is decreasing, we have $B_n \supseteq B_{n-1}$. Recalling that $\nu(B_k) \leq -k$, we note that $\nu(B_k)$ must be a real number since $\alpha > -\infty$. Call $B = \bigcap_{k=1}^{\infty} B_n$, then $B_n \downarrow B$. By continuity of above, we know that
$$\nu(B) = \lim_{n\to\infty} \nu(B_n) = -\infty$$
But this contradicts our assumption regarding α.

2. There are an infinite number of bifurcations. We may arrange a sequence $n_0, n_1, n_2, \ldots$ of natural numbers at which a bifurcation occurs. By what we have shown before, $B_{n_j} \cap B_{n_k} = \varnothing$ for distinct j, k. Define $B = \bigcup_{j\geq 0} B_{n_j}$. By σ-additivity, we have again that
$$\nu(B) = \sum_{j\geq 0} \nu\left(B_k\right) = -\infty$$
Note that $\nu(B_k)$ must be negative real numbers. Hence, we obtain a contradiction.

From the above argument, we know that $\alpha = \inf_{A\in\mathcal{F}} \nu(A) > -\infty$. We will construct a set N so that $\nu(N) = \alpha$. By the definition of infimum, there exists $C_n \in \mathcal{F}$ so that
$$\alpha \leq \nu(C_n) < \alpha + \frac{1}{2^n}$$
By inclusion-exclusion, we may write
$$\begin{aligned}
\nu\left(C_n \cup C_{n+1}\right) &= \nu\left(C_n\right) + \nu\left(C_{n+1}\right) - \nu\left(C_n \cap C_{n+1}\right) \\
&< \alpha + \frac{1}{2^n} + \alpha + \frac{1}{2^{n+1}} - \alpha \\
&= \alpha + \frac{1}{2^n} + \frac{1}{2^{n+1}}
\end{aligned}$$
If we repeat this argument, we obtain
$$\nu\left(\bigcup_{k=n}^{n+q} C_k\right) \leq \alpha + \frac{1}{2^n} + \cdots + \frac{1}{2^{n+q}}$$
If we take the limit as $q \to \infty$, we have
$$\nu\left(\bigcup_{k\geq n} C_k\right) \leqslant \alpha + \frac{2}{2^n}$$
Set $D_n := \bigcup_{k\geq n} C_k$ and $D = \bigcap_{n\geq 1} D_n$. Since $\alpha + \frac{2}{2^n}$ is finite, we may apply continuity from above to see that
$$\nu(D) \leq \alpha$$
But by definition $\nu(D) \geq \alpha$, so $\nu(D) = \alpha$ and this is our desired set N.

Finally, we may represent $P = N^c$ and verify that if $E \subseteq P$, then $\nu(E) \geq 0$. Suppose, for contradiction that $\nu(E) < 0$. Then $F = N \cup E$ is a disjoint union and

$$\nu(F) = \nu(N) + \nu(E) < \alpha$$

which contradicts the fact that α is an infimum. On the other, hand suppose for contradiction that $F \subseteq N$ so that $\nu(F) > 0$. We cannot have $\nu(F) = +\infty$ since $\nu(N)$ is finite. Now

$$\nu(N \setminus F) = \nu(N) - \nu(F) < \alpha$$

which leads to a contradiction again. This proves the Hahn decomposition theorem. ■

Remark. Let ν be a signed measure. We may define two measures by

$$\nu_+(E) = \nu(E \cap P) \geq 0, \quad \nu_-(E) = -\nu(E \cap N) \geq 0$$

Further, we may always write

$$\nu = \nu^+ - \nu^-$$

This follows from σ-additivity and since P and N form a disjoint partition of Ω

$$\nu(E) = \nu(E \cap \Omega) = \nu(E \cap P) + \nu(E \cap N) = \nu_+(E) - \nu_-(E)$$

Thus, we are able to decompose a signed measure as two nonnegative measures. If we force the support of the two measures to be disjoint, then we can have uniqueness. We repeat this as a theorem from Folland below.

Theorem. *The Jordan Decomposition Theorem.*
If ν is a signed measure, there exist unique positive measures $\nu^+$and ν^-such that $\nu = \nu^+ - \nu^-$and $\nu^+ \perp \nu^-$.

Proof. Let $X = P \cup N$ be a Hahn decomposition for ν, and define $\nu^+(E) = \nu(E \cap P)$ and $\nu^-(E) = -\nu(E \cap N)$. Then clearly $\nu = \nu^+ - \nu^-$and $\nu^+ \perp \nu^-$. If also $\nu = \mu^+ - \mu^-$and $\mu^+ \perp \mu^-$, let $E, F \in \mathcal{M}$ be such that $E \cap F = \varnothing, E \cup F = X$, and $\mu^+(F) = \mu^-(E) = 0$. Then $X = E \cup F$ is another Hahn decomposition for ν, so $P \Delta E$ is ν-null. Therefore, for any $A \in \mathcal{M}$, $\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = \nu^+(A)$, and likewise $\nu^- = \mu^-$. ■

20 The Radon-Nikodym Theorem

Suppose that $(\Omega, \mathcal{F}, \mu)$ is a measure space and μ is σ-finite. Let $\nu : \mathcal{F} \to (-\infty, \infty]$ be σ-finite in the same measurable space as well.

Definition. *Absolutely continuous*
We say that ν is **absolutely continuous** with respect to μ, with the symbol $\nu \ll \mu$, if

$$\mu(A) = 0 \implies \nu(A) = 0$$

where $A \in \mathcal{F}$.

Example. Let f be an integrable function and define

$$\nu(A) = \int_A f \, d\mu$$

We know that if $\mu(A) = 0$, then $\int_A f \, d\mu = 0$, hence $\nu \ll \mu$. Recall our previous discussion on this. One consequence of the Radon-Nikodym theorem is that it will give us a converse to the above example: if $\nu \ll \mu$, we will be able to find a function f such that $\nu(A) = \int_A f \, d\mu$.

Definition. *Singular, mutually singular*
Let μ and ν be σ-finite measures. We say that ν is **singular** with respect to μ if there exists $A \in \mathcal{F}$ such that $\mu(A) = 0$ and $\nu(A^c) = 0$.

We use the symbol $\nu \perp \mu$ to denote that ν and μ are mutually singular. Note that this relation is symmetric.

Example. Let λ be the Lebesgue measure on the line. Let $\mathbb{Q} = \{q_j\}_{j \in \mathbb{N}}$ be an enumeration of the rationals and $c_j \geq 0$ with $\sum_{j=1}^{\infty} c_j < \infty$. Define

$$\nu = \sum_{j=1}^{\infty} c_j \delta_{q_j}$$

Recall that δ_x for $x \in \mathbb{R}$ is the Dirac measure, which is given by

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

For the measure ν, each δ_{q_j} has contribution c_j. The intuitive idea here is that the positive part of ν is concentrated on the rationals while the positive part of λ is concentrated on the irrationals. Taking $A = \mathbb{Q}$, we have

$$\nu(A^c) = 0, \quad \lambda(A) = 0$$

Indeed, it follows that $\nu \perp \lambda$.

Remark. Note that by requiring that μ and ν are (nonnegative) measures, it means that subsets of A^c must also have measure zero. In Folland, this is done by defining positive, negative, and null sets and mutually singularity for signed measures using those notions. We include these definitions for comparison.

Definition. *Positive, negative, null set*
If ν is a signed measure on $(X, \mathcal{M})$, a set $E \in \mathcal{M}$ is called positive (resp. negative, null) for ν if $\nu(F) \geq 0$ (resp. $\nu(F) \leq 0, \nu(F) = 0$) for all $F \in \mathcal{M}$ such that $F \subset E$.

Definition. *Mutually singular*
We say that two signed measures μ and ν on $(X, \mathcal{M})$ are mutually singular, or that ν is singular with respect to μ, or vice versa, if there exist $E, F \in \mathcal{M}$ such that $E \cap F = \varnothing, E \cup F = X, E$ is null for μ, and F is null for ν. Informally speaking, mutual singularity means that μ and ν "live on disjoint sets." We express this relationship symbolically with the perpendicularity sign:

$$\mu \perp \nu.$$

Theorem. *Radon-Nikodym*
Let μ be a σ-finite measure and $\nu : \mathcal{F} \to (-\infty, \infty]$ a signed measure. Then

1. There exist signed measures ν_1, ν_2 such that $\nu = \nu_1 + \nu_2$ and
$$\nu_1 \ll \mu \text{ and } \nu_2 \perp \mu$$
2. The decomposition is unique.
3. There exists a measurable function f so that
$$\nu_1(A) = \int_A f \, d\mu$$
Note that f does not have to be integrable.

Remark. We preface the proof with a sketch.

Step 1: Assuming that μ and ν are finite, construction of a function g that maximizes $\mu_g(\Omega)$ integral among all those such that $\mu_g(A) \leq \nu(A)$ for all $A \in \mathcal{F}$.

Step 2: Showing that $\nu_2 = \mu - \mu_g$ is singular with respect to μ (completing the proof when μ

and ν are finite).

Step 3: We will argue that the conclusions hold without the assumptions in Step 1.

Proof. First, assume that $\nu : \mathcal{F} \to [0, \infty]$ and that ν and μ are finite; we will prove the general case afterwards. Define

$$\mathscr{H} = \left\{ f \text{ measurable}, f \geq 0 : \int_A f \, d\mu \leq \nu(A), \ \ \forall A \in \mathcal{F} \right\}$$

Observe that $\mathscr{H}$ is nonempty since we may consider $f \equiv 0$. Define

$$\alpha = \sup_{f \in \mathscr{H}} \int_\Omega f \, d\mu$$

We want to find a maximal function in $\mathscr{H}$. Notice that the supremum will be a nonnegative function. We proceed in steps below.

1. We have $\alpha < \infty$. We know that $\nu(\Omega)$ is finite by assumption, so $\int_\Omega f \, d\mu$ will be bounded above by $\nu(\Omega)$ for all $f \in \mathscr{H}$. Hence, $\alpha \leq \nu(\Omega)$ and the supremum is finite.

2. By the definition of supremum, we may find a sequence of functions $(f_n)_{n \in \mathbb{N}}$ so that $f_n \in \mathscr{H}$ and
$$\alpha - \frac{1}{n} < \int_A f \, d\mu \leq \alpha$$
Define $g_n = \max\{f_1, \ldots, f_n\}$, which is an increasing sequence of functions. We claim that $g_n \in \mathscr{H}$. To see this, denote $E_{n,k} = \{x \in \Omega : f_k(x) = g_n(x)\}$ for $1 \leq k \leq n$ and observe that
$$\int_A g_n \, d\mu = \sum_{k=1}^n \int_{A \cap E_{n,k}} g_n \, d\mu = \sum_{k=1}^n \int_{A \cap E_{n,k}} f_k \, d\mu$$
$$\leq \sum_{k=1}^n \nu(A \cap E_{n,k}) = \nu(A)$$
The last equality follows since we may construct $E_{n,k}$ as a partition of Ω. Thus, $g_n \in \mathscr{H}$.

3. Since g_n is an increasing sequence, we may define its limit as g. We claim that $g \in \mathscr{H}$. Since $g_n \in \mathscr{H}$, we know that $\int_A g_n \, d\mu \leq \nu(A)$ for all $n \in \mathbb{N}$. By the monotone convergence theorem, we see that
$$\int_A g \, d\mu = \lim_{n \to \infty} \int_A g_n \, d\mu \leq \nu(A)$$

4. We will show $\int g \, d\mu = \alpha$. Since $g \in \mathscr{H}$ and α is an upper bound of $\{\int f \, d\mu : f \in \mathscr{H}\}$, we see that $\int g \, d\mu \leq \alpha$. On the other hand, we have
$$\int g \, d\mu \geq \int g_n \, d\mu \geq \int f_n \, d\mu \geq \alpha - \frac{1}{n}$$
Hence, we see that $\int g \, d\mu \geq \alpha$.

So, at the end of this process, we have $\mathscr{H}$ and g as a maximal element of $\mathscr{H}$. Set

$$\nu_1(A) := \int_A g \, d\mu \leq \nu(A), \quad \nu_2(A) := \nu(A) - \nu_1(A)$$

By construction, we see that $\nu_1 \ll \mu$. It remains to show that $\nu_2 \perp \mu$.

Aside. The motivation for defining σ_n below is to consider the inequality

$$\frac{\nu_2}{\mu} \geq \frac{1}{n}$$

Under the condition of mutual singularity, whenever ν_2 is positive, μ will be zero so the inequality above should hold. Rearranging gives us σ_n.

Define the signed measure

$$\sigma_n = \nu_2 - \frac{1}{n}\mu$$

Since this is a signed measure, σ_n has a Hahn decomposition into a positive set P_n and a negative set N_n which partition Ω. Our goal is to show $\mu(P_n) = 0$. Consider the function $g + \frac{1}{n}\mathbf{1}_{P_n}$, we claim that this function belongs to $\mathscr{H}$. To verify this, note that this function is nonnegative and measurable already, so it suffices to check that its integral is bounded by $\nu(A)$.

$$\int_A \left[g + \frac{1}{n}\mathbf{1}_{P_n}\right] d\mu = \nu_1(A) + \frac{1}{n}\mu(P_n \cap A)$$

Since $(P_n \cap A) \subseteq P_n$, it follows that $\sigma_n(P_n \cap A) \geq 0$. Plugging into the definition for σ_n, we have

$$\nu_2(P_n \cap A) \geq \frac{1}{n}\mu(P_n \cap A)$$

Continuing, we see that

$$\nu_1(A) + \frac{1}{n}\mu(P_n \cap A) \leq \nu_1(A) + \nu_2(P_n \cap A) \leq \nu_1(A) + \nu_2(A) = \nu(A)$$

where we have used the fact that ν_2 is a nonnegative measure to say that $\nu_2(P_n \cap A) \leq \nu_2(A)$. By this inequality, we obtain that $g + \frac{1}{n}\mathbf{1}_{P_n} \in \mathscr{H}$. Observe that

$$\int_A \left[g + \frac{1}{n}\mathbf{1}_{P_n}\right] d\mu = \alpha + \frac{1}{n}\mu(P_n)$$

Using the maximality of g, we must have $\mu(P_n) = 0$ for all $n \in \mathbb{N}$. Define $P = \bigcup_{n\in\mathbb{N}} P_n$, then using countable subadditivity, we have

$$\mu(P) \leq \mu\left(\bigcup_{n\in\mathbb{N}} P_n\right) = 0$$

Define $N = P^c$ and notice $P^c = \left(\bigcup_{n\in\mathbb{N}} P_n\right)^c = \bigcap_{n\in\mathbb{N}} N_n$. It remains to show that $\nu_2(N) = 0$. Since N_n is a negative set for σ_n, we know that $\sigma_n(N_n) = \nu_2(N_n) - \frac{1}{n}\mu(N_n) \leq 0$. Now, by monotonicity of measure, we have the following inequalities

$$\nu_2(N) \leq \nu_2(N_n) \leq \frac{1}{n}\mu(N_n) = \frac{1}{n}\mu(\Omega) < \infty$$

Since this inequality holds for all $n \in \mathbb{N}$, we have $\nu_2(N) = 0$.
To summarize, we have obtained two sets, P and N so that $\mu(P) = 0$ and $\nu_2(N) = 0$, hence $\mu \perp \nu_2$.

Now, we remove the finiteness assumption. We will assume that ν and μ are σ-finite instead. Since μ is σ-finite, we may find $E_n \in \mathcal{F}$ such that $E_n \uparrow \Omega$ and $\mu(E_n) < \infty$ for all $n \in \mathbb{N}$. Similarly, we may fine $F_n \in \mathcal{F}$ such that $F_n \uparrow \Omega$ and $\nu(F_n) < \infty$ for all $n \in \mathbb{N}$. Call $G_n = E_n \cap F_n$. Since E_n and F_n are increasing sequences, so is G_n, that is $G_n \subseteq G_{n+1}$. We claim that $\Omega = \bigcup_n G_n$. Take $x \in \Omega$. Since E_n is an increasing sequence, there exists some $n_0 \in \mathbb{N}$ so that $x \in E_n$ for all $n \geq n_0$. Similarly, there exists some $n_1 \in \mathbb{N}$ so that $x \in F_n$ for all $n \geq n_1$. Take $N = \max n_0, n_1$, then $x \in \bigcup_n G_n$ if $n \geq N$ and the equality follows. By monotonicity, we also see that $\mu(G_n) < \infty$ and $\nu(G_n) < \infty$ for all n.

So, there exists a sequence $G_n \uparrow \Omega$ so that $\mu(G_n) < \infty$ and $\nu(G_n) < \infty$. We construct the sets

$$H_1 = G_1, \ldots, H_k = G_k \setminus G_{k-1}$$

Since G_n is an increasing sequence, we see that $H_i \cap H_j = \varnothing$ for $i \neq j$, $\Omega = \bigsqcup_{j=1}^{\infty} H_j$, and $\mu(H_j) < \infty$ and $\nu(H_j) < \infty$. Define two finite measures

$$\mu_j(A) = \mu(A \cap H_j), \quad \nu_j(A) = \nu(A \cap H_j)$$

We may apply the Radon-Nikodym theorem for finite measures on these two measures to obtain a decomposition $\nu_j = \nu_j^1 + \nu_j^2$ where $\nu_j^1 \ll \mu_j$ and $\nu_j^2 \perp \mu_j$. Now, taking the sum of the measures

$$\nu^1 = \sum_j \nu_j^1, \quad \nu^2 = \sum_j \nu_j^2$$

It can be shown that $\nu^1 \ll \mu$ and $\nu^2 \perp \mu$; the intuitive idea is that we have shown the decomposition and the absolute continuity and mutual singularity on disjoint sets of Ω, so when we combine the domains H_j, these properties still carry over.

Lastly, we consider the original assumptions of the theorem. Let μ and ν be σ-finite and $\nu : \mathcal{F} \to (-\infty, \infty]$ be a signed measure. Since ν is a signed measure, it has a Jordan decomposition into nonnegative measures $\theta_1, \theta_2 : \mathcal{F} \to [0, \infty]$ so that $\nu = \theta_1 - \theta_2$. Since these

θ_j are nonnegative, we may apply the version of the theorem we proved above to obtain the decompositions

$$\theta_1 = \theta_1^1 + \theta_1^2, \quad \theta_2 = \theta_2^1 + \theta_2^2$$

where θ_1^1 and θ_2^1 are absolutely continuous with respect to μ and the other two measures are mutually singular with μ. Recall that in our Jordan decomposition for ν, we will have two sets P and N which are supports for θ_1 and θ_2 respectively. So, the supports of θ_1^1 and θ_2^1 are disjoint, and we can combine these measures to get a measure on the entire space Ω. The same applies to the remaining measures θ_1^2 and θ_2^2. Therefore, we define

$$\nu_1 = \theta_1^1 - \theta_2^1, \quad \nu_2 = \theta_1^2 - \theta_2^2$$

To conclude the proof, we verify the uniqueness assumption. Suppose we have the decompositions

$$\nu = \nu_1 + \nu_2 = \overline{\nu}_1 + \overline{\nu}_2$$

where $\nu_1, \overline{\nu}_1 \ll \mu$ and $\nu_2, \overline{\nu}_2 \perp \mu$. We may write the above identity as

$$\nu_1 - \overline{\nu}_1 = \overline{\nu}_2 - \nu_2$$

By the mutual singularity, we know that

$$\begin{aligned} &\exists A, \quad \mu(A) = 0 \quad \nu_2\left(A^c \cap F\right) = 0, \;\; \forall F \in \mathcal{F} \\ &\exists B, \quad \mu(B) = 0 \quad \overline{\nu}_2\left(B^c \cap G\right) = 0, \;\; \forall G \in \mathcal{F} \end{aligned}$$

Set $C = A \cup B$, so $\mu(C) = 0$. Consider the set $E \cap (A \cup B)$ and $E \cap (A \cup B)^c$.

1. Since $\mu(C) = 0$, it follows that $\mu(E \cap (A \cup B)) = 0$ by monotonicity. By absolute continuity, we know that

$$\nu_1(E \cap (A \cup B)) = 0, \quad \overline{\nu}_1(E \cap (A \cup B)) = 0$$

2. By mutual singularity, we have

$$\begin{aligned} \nu_2\left(E \cap A^C \cap B^C\right) = 0 \\ \overline{\nu}_2\left(E \cap A^C \cap B^C\right) = 0 \end{aligned}$$

Using the facts above, we may compute

$$\begin{aligned} (\nu_1 - \overline{\nu}_1)(E) &= \underbrace{(\nu_1 - \overline{\nu}_1)(E \cap (A \cup B))}_{0} + (\nu_1 - \overline{\nu}_1)(E \cap (A \cup B)^c) \\ &= (\nu_1 - \overline{\nu}_1)(E \cap (A \cup B)^c) = (\overline{\nu}_2 - \nu_2)(E \cap (A \cup B)^c) = 0 \end{aligned}$$

where the last equality follows from the mutual singularity of (2). Thus, $\nu_1 = \overline{\nu}_1$ which forces $\nu_2 = \overline{\nu}_2$. ■

Remark. The function f in the above proof is only unique up to sets of measure zero, since if $f = g$ μ-a.e., then

$$\nu_1(A) = \int_A f \, d\mu = \int_A g \, d\mu$$

21 Almost sure and almost uniform

We start a new chapter on the convergence of functions. Fix a measure space $(\Omega, \mathcal{F}, \mu)$ where Ω is σ-finite with respect to μ and $\mathcal{F}$ is μ-complete. Let $f, g : \Omega \to \overline{\mathbb{R}}$. Recall a previous proposition which states that if f is measurable, $\mathcal{F}$ is μ-complete, and $g = f$ a.s., then g is measurable. Based on this proposition, it makes sense to define an equivalence relation $f \sim g$ if $\mu(\{x \in \Omega : f(x) \neq g(x)\}) = 0$. The point is we define this equivalence relation on measurable functions.

Aside. Reflexivity and symmetry are obvious for the above relation $f \sim g$. Suppose that $f \sim g$ and $g \sim h$. Observe that

$$\begin{aligned}\mu(\{x \in \Omega : f(x) \neq h(x)\} &\leq \mu(\{f \neq g\} \cup \{g \neq h\}) \\ &\leq \mu(\{f \neq g\}) + \mu(\{g \neq h\}) = 0\end{aligned}$$

So, this is an equivalence relation.

We can phrase the above in the language of quotient sets. Consider the set

$$M = \{f : \Omega \to \overline{\mathbb{R}} : f \text{ is measurable}\}$$

We may take the quotient by the saying that functions which differ only on a measure zero set belong to the same equivalence class: $\mathcal{M} = {}^{M}/_{\sim}$. We will be working in the space $\mathcal{M}$, so that whenever we speak about a function, we are speaking about the representative of an equivalence class. This is the background of the chapter.

21.1 Types of Convergence

We will describe three types of convergence. The definitions are given below.

1. $f_n \to f$ a.e., this is called convergence almost everywhere.

   (a) Pointwise convergence. Let $f_n, f : E \to \overline{\mathbb{R}}$. We say that $f_n \to f$ pointwise if for all $x \in E$, $f_n(x) \xrightarrow[n\to\infty]{} f(x)$.

   (b) Almost sure or almost everywhere convergence. Let $f_n, f : \Omega \to \overline{\mathbb{R}}$. We write $f_n \to f$ a.s. or a.e. if there exists $E \in \mathcal{F}$ such that $f_n \to f$ pointwise on E and $\mu(E^c) = 0$.

2. Uniform convergence a.e.

   (a) Uniform convergence. Let $f_n, f : E \to \overline{\mathbb{R}}$. Recall that $f_n \to f$ uniformly if for $\varepsilon > 0$ there exists n_0 such that for all $n \geq n_0$, we have $|f_n(x) - f(x)| < \varepsilon$ for all $x \in E$.

Another way to say this is

$$\sup_{x\in E} |f_n(x) - f(x)| \le \varepsilon$$

The point is that we can choose one n_0 for which $|f_n(x) - f(x)|$ is bounded uniformly for all $x \in E$.

(b) Uniform convergence a.e. Let $f_n, f : \Omega \to \overline{\mathbb{R}}$. We say that $f_n \to f$ uniformly a.e. if there exists $E \in \mathcal{F}$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E. So, if we restrict these functions to E, then we have uniform convergence.

3. Almost uniform convergence. Let $f_n, f : \Omega \to \overline{\mathbb{R}}$. We say that $f_n \to f$ almost uniformly if for all $\varepsilon > 0$, there exists $E_\varepsilon \in \mathcal{F}$ such that $\mu(E_\varepsilon^c) \le \varepsilon$ and $f_n \to f$ uniformly on E_ε.

   The point of this definition is to take care of sequences of functions such as $f_n(x) = x^n$ on $[0, 1]$, where we can obtain uniform convergence if we choose some $\varepsilon > 0$ and consider that $f_n \to f$ uniformly on $[0, 1 - \varepsilon]$.

21.2 Information on convergence almost everywhere

Proposition. It is important to note that almost sure convergence is convergence in the equivalence class. To be precise, if $f_n \to f$ a.s. and $f_n \sim g_n$ and $f \sim g$, then $g_n \to g$ a.s.

Proof. Since $f_n \to f$ a.s., there exists some $E \in \mathcal{F}$ such that $f_n \to f$ pointwise on E and $\mu(E^c) = 0$. Since $f_n \sim g_n$, for all $n \in \mathbb{N}$, there exists some $F_n \in \mathcal{F}$ such that $f_n(x) = g_n(x)$ for all $x \in F_n$ and $\mu(F_n^c) = 0$. Similarly, since $f \sim g$, there exists some $F \in \mathcal{F}$ such that $f(x) = g(x)$ for all $x \in F$ and $\mu(F^c) = \varnothing$.

Take $H = F \cap E \cap \bigcap_{n\in\mathbb{N}} F_n$. We know that $H \in \mathcal{F}$ because it is the countable intersection of measurable sets. Further

$$\mu(H^c) = \mu\left(\Omega \setminus \left(F \cap E \cap \bigcap_{n\in\mathbb{N}} F_n\right)\right) = \mu\left(F^c \cup E^c \cup \bigcup_{n\in\mathbb{N}} F_n\right) \le 0$$

To conclude the proof, we observe that if $x \in H$, then $g_n(x) \to g(x)$. For all $n \in \mathbb{N}$, we have $g_n(x) = f_n(x)$ since $x \in F_n$. Next $f_n(x) \to f(x)$ since $x \in E$. Lastly, $f(x) = g(x)$ since $x \in F$. Therefore, the existence of H shows us that $g_n \to g$ a.s. ■

Remark. The idea of the proposition above is that when we have convergence a.s., we can change the functions to other representatives in the equivalence class while still ensuring that the new sequence converges a.s.

Proposition. If $f_n \to f$ a.s. and $f_n \to g$ a.s., then $f = g$ a.s.

Proof. By definition, there exists $E \in \mathcal{F}$ such that $\mu(E^c) = 0$ and $f_n \to f$ pointwise on E. Similarly, there exists $F \in \mathcal{F}$ such that $\mu(F^c) = 0$ and $f_n \to g$ pointwise on F.

Take $H = E \cap F$. We see that $\mu(H^c) = \mu(E^c \cup F^c) = 0$. Further, since $x \in H$, we have

$$g(x) = \lim_{n\to\infty} f_n(x) = f(x)$$

So, $f = g$ on H, that is $f = g$ a.s. ∎

Remark. The above shows the uniqueness of the limit when considered with respect to the equivalence class of functions.

Example. Consider the measure space $([0,1], \mathcal{L}, \lambda)$ where λ is the Lebesgue measure. Consider the sequence of functions $f_n(x) = x^n$. We know from analysis that $f_n(x) \to 0$ for $x \in [0,1)$ and $f_n(1) \to 1$. So, taking $f \equiv 0$, we have that $f_n \to f$ λ-a.s. But, it is not true that $f_n \to f$ pointwise. So this is a sequence of functions which converges almost surely but not pointwise, on $[0,1]$.

21.3 Information on uniform convergence a.e.

A useful notion for discussing uniform convergence almost everywhere will be the *essential supremum*. Let $f : \Omega \to \overline{\mathbb{R}}$. Assume that there exists $a > 0$ such that

$$\mu\{|f| > a\} = 0$$

Note that if $b > a$ and $\mu\{|f| > a\} = 0$, this implies that $\mu(\{|f| > b\} = 0$, since $\{|f| > b\} \subseteq \{|f| > a\}$.

Definition. *Essential supremum*
We define

$$\operatorname{ess\,sup}|f| = \inf\{a > 0 : \mu\{|f| > a\} = 0\}$$

Recall that

$$\sup|f| = \sup\{|f(x)| : x \in \Omega\}$$

It is clear that $\operatorname{ess\,sup}|f| \leq \sup|f|$. Strict inequality is possible. Consider the following example.

$$f(x) = \begin{cases} 1, & \text{if } x \neq 0 \\ \infty, & \text{if } x = 0 \end{cases}$$

We see that

$$\operatorname{ess\,sup}|f| = 1 < \infty = \sup|f|$$

We call the above notion the *essential* supremum because we are allowed to remove points in a set of measure zero. Next, we will discuss some common properties of the essential supremum.

Observation. If $\operatorname{ess\,sup}|f| = c$, then

$$\mu(\{x \in \Omega : |f(x)| > c\}) = 0$$

Since $\operatorname{ess\,sup}|f| = c$, for every $n \in \mathbb{N}$, there exists some $a_n \in \{a > 0 : \mu\{|f| > a\} = 0\}$ such that

$$c \leq a_n < c + \frac{1}{n}$$

So $\mu\{|f| > a_n\} = 0$ and we also have $\bigcup_{n\in\mathbb{N}}\{|f| > a_n\} = \{|f| > c\}$. Taking the measure, we see that

$$\mu\left(\bigcup_{n\in\mathbb{N}}\{|f| > a_n\}\right) = 0$$

by countable sub-additivity.

Proposition. The essential supremum is a class property. To be exact, if $f \sim g$, then $\operatorname{ess\,sup}|f| = \operatorname{ess\,sup}|g|$.

Proof. We may find a set $E \in \mathcal{F}$ so that $\mu(E^c) = 0$ and $f(x) = g(x)$ for all $x \in E$. Let $c = \operatorname{ess\,sup}|f|$ and consider the set $A := \{x \in \Omega : |g(x)| > c\}$. Observe that

$$\mu(A) \leq \mu(A \cap E^c) + \mu(A \cap E)$$

We have $\mu(A \cap E^c) = 0$ since $\mu(E^c) = 0$. Since $A \cap E \subseteq E$, we have $f(x) = g(x)$ on $A \cap E$. Hence, we see that

$$\begin{aligned}\mu(A) = \mu(\{x \in \Omega : |g(x)| > c\}) &\leq \mu(\{x \in \Omega : |f(x)| > c\} \cap E)\\ &\leq \mu(\{x \in \Omega : |f(x)| > c\}) = 0\end{aligned}$$

So, from the above, we see that $\mu(\{x \in \Omega : |g(x)| > c\}) = \mu(\{x \in \Omega : |f(x)| > c\}) = 0$. This means that $c \in \{a > 0 : \mu(\{|g| > a\}) = 0\}$. Thus, we see that $\operatorname{ess\,sup}|g| \leq c = \operatorname{ess\,sup}|f|$. Since the roles of f and g are symmetric, we may replace f and g in the argument above to obtain the reverse inequality, and from this we obtain $\operatorname{ess\,sup}|f| = \operatorname{ess\,sup}|g|$. Therefore, if two functions are equal a.e. (meaning they belong to the same equivalence class), then they have the same essential supremum. ■

The last property we will see for essential supremum is that it induces a distance (in fact a norm) on the equivalence class of functions and that uniform convergence a.e. corresponds to the convergence for the topology induced by this distance.

Proposition. Taking $d(f,g) = \operatorname{ess\,sup}|f-g|$ defines a distance function on $\mathcal{M}$.

Proof. Clearly, d is symmetric by definition. Continuing, we will verify that if $d(f,g)=0$, then $f=g$ a.s. It is sufficient to show that if $\operatorname{ess\,sup}|h|=0$, then $h=0$ a.s. By the definition of infimum, if $\operatorname{ess\,sup}|h|=0$, we must have

$$\mu\left(\left\{x\in\Omega : |h(x)| > \frac{1}{n}\right\}\right) = 0$$

for all $n\in\mathbb{N}$. But we have seen previously that

$$\bigcup_{n\in\mathbb{N}}\left\{x\in\Omega : |h(x)| > \frac{1}{n}\right\} = \{x\in\Omega : |h(x)|>0\}$$

So, using countable additivity, we conclude that

$$\mu(\{x\in\Omega : |h(x)|>0\}) = 0$$

But this means that $h=0$ a.s. (since the set where h takes positive and negative values has measure zero).
The remaining property to show is the triangle inequality, that is $d(f,h)\le d(f,g)+d(g,h)$. Denote $a := d(f,g)$ and $b := d(g,h)$. We want to show that

$$\mu\{x\in\Omega : |f(x)-h(x)| > a+b\} = 0$$

By the triangle inequality in $\mathbb{R}$, we have that $|f(x)-h(x)| \le |f(x)-g(x)| + |g(x)-h(x)|$. If we assume that the inequality we want is true, we have

$$a+b < |f(x)-h(x)| \le |f(x)-g(x)| + |g(x)-h(x)|$$

But this means that either $|f(x)-g(x)|>a$ or $|g(x)-h(x)|>b$, otherwise we have a contradiction. Hence

$$\begin{aligned}\mu\{x\in\Omega : |f(x)-h(x)|>a+b\} &\le \mu\{x : |f(x)-g(x)|>a\} + \mu\{x : |g(x)-h(x)|>b\}\\ &= 0+0\end{aligned}$$

Using the first property and that $a=\operatorname{ess\,sup}|f-g|$, we see that $\mu\{x : |f(x)-g(x)|>a\}=0$, and similarly we have $\mu\{x : |g(x)-h(x)|>b\}=0$. Thus, we have $\mu\{x\in\Omega : |f(x)-h(x)| > a+b\}=0$, in other words $\operatorname{ess\,sup}|f-h|\le a+b$, and the triangle inequality follows. ■

Aside. In the above proof, we use the fact that adding functions that belong to the same equivalence class is well-defined in several places. First, to define $d(f,g)=\operatorname{ess\,sup}|f-g|$ means that we are defining a metric on the equivalence classes using representatives from that class. Second, when we say it is sufficient to show that $\operatorname{ess\,sup}|h|=0$, we use that if $f - g = 0$ a.e., then $f = g$ a.e.

Proposition. $f_n \to f$ uniformly a.e. if and only if $d(f_n, f) \to 0$ using the distance function defined by the essential supremum.

Proof. Suppose that $f_n \to f$ uniformly a.e. Then, there exists some $E \in \mathcal{F}$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E. Fix $\varepsilon > 0$. We want to show that there exists some n_0 such that for all $n \geq n_0$, we have

$$\operatorname{ess\,sup} |f_n - f| = d(f_n, f) \leq \varepsilon$$

Since $f_n \to f$ uniformly on E, this means that for all $\varepsilon > 0$, there exists some n_0 such that for all $n_0 \geq n$ and for all $x \in E$ we have

$$|f_n(x) - f(x)| < \varepsilon$$

But this means $\sup_{x \in E} |f_n(x) - f(x)| \leq \varepsilon$. Hence, if $|f_n(x) - f(x)| > \varepsilon$, then $x \in E^c$. It follows that

$$\mu(\{x \in \Omega : |f_n(x) - f(x)| > \varepsilon\}) = 0$$

since the above set is a subset of E^c. From this, we see that $\varepsilon \in \{a > 0 : \mu(|f_n - f| > a) = 0\}$ and using the definition of ess sup as the infimum, we see that

$$d(f_n, f) = \operatorname{ess\,sup} |f_n - f| \leq \varepsilon$$

as desired.

Now, we show the converse: assume that $d(f_n, f) \xrightarrow[n \to \infty]{} 0$. This means that $\operatorname{ess\,sup} |f_n - f| \to 0$, equivalently, for all $\varepsilon > 0$, there exists some n_0 such that for all $n \geq n_0$ we have

$$\operatorname{ess\,sup} |f_n - f| \leq \varepsilon$$

It follows from a previous proposition that $\mu(\{x \in \Omega : |f_n(x) - f(x)| > \varepsilon\}) = 0$. Let $k \geq 1$, then there exists n_k such that for all $n \geq n_k$, we have $\mu(\{x \in \Omega : |f_n(x) - f(x)| > \frac{1}{2^k}\}) = 0$. Taking unions, we see that

$$\mu\left(\bigcup_{n \geq n_k} \left\{x \in \Omega : |f_n(x) - f(x)| > \frac{1}{2^k}\right\}\right) = 0$$

$$\mu\left(\bigcup_{k \geq 1} \bigcup_{n \geq n_k} \left\{x \in \Omega : |f_n(x) - f(x)| > \frac{1}{2^k}\right\}\right) = 0$$

Set $E^c = \bigcup_{k \geq 1} \bigcup_{n \geq n_k} \left\{x \in \Omega : |f_n(x) - f(x)| > \frac{1}{2^k}\right\}$. To conclude the proof, we must show that on E, we have uniform convergence. Observe that

$$E = \bigcap_{k \geq 1} \bigcap_{n \geq n_k} \left\{x \in \Omega : |f_n(x) - f(x)| \leq \frac{1}{2^k}\right\}$$

Let $\varepsilon > 0$, then, there exists some k_0 such that $\frac{1}{2^{k_0}} \leq \varepsilon$. Suppose that $x \in E$, then we know that for $n \geq n_{k_0}$, we have

$$|f_n(x) - f(x)| \leq \frac{1}{2^{k_0}} < \varepsilon$$

Since this is true for all $x \in E$, we have $\sup_{x\in E} |f_n(x) - f(x)| \leq \varepsilon$ and uniform convergence on E follows. ■

Remark. Consider the space

$$L^\infty = \{f \in \mathcal{M} : \operatorname{ess\,sup} |f| < \infty\}$$

Recall that this is well-defined because the essential supremum is the same for functions that are equal a.e. It turns out that L^∞ is a vector space. To see this, we need to verify that for $f, g \in L^\infty$ and $\alpha \in \mathbb{R}$, we have $\alpha f + g \in L^\infty$.

Remark. Recall our previous example where we take $f_n(x) = x^n$ on the measure space $([0,1], \mathcal{L}, \lambda)$. Notice that this sequence of functions converges a.e., but does not converge uniformly a.e. If we only remove a set of measure zero, then we will always be able to find a sequence of points converging to 1, and this prevents uniform convergence. It is in light of this example that we define almost uniform convergence.

21.4 Comparison of convergence for sequences of functions

We list some relationships between the different types of convergence.

- Uniform convergence a.e. implies convergence a.e. and uniform convergence a.e. implies almost uniform convergence.
- Almost uniform convergence implies convergence a.e.

Proposition. Almost uniform convergence implies convergence a.e.

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be our sequence of functions. Suppose that $f_n \to f$ almost uniformly. Then, for all $k \in \mathbb{N}$, we may find a set E_k such that $f_n \to f$ uniformly on E_k and $\mu(E_k^c) \leq \frac{1}{k}$. If we have uniform convergence on E_k, then we have pointwise convergence on E_k. Now, consider that

$$\mu\left(\bigcap_{k=1}^{\infty} E_k^c\right) \leq \mu(E_{k_0}) \leq \frac{1}{k_0}$$

Since this holds for all k_0, we have $\mu\left(\bigcap_{k=1}^{\infty} E_k^c\right) = 0$. Define $E = \bigcup_{k=1}^{n} E_k$, which is the complement of the intersection above. Observe that we have pointwise convergence on E. Pick an $x \in E$, then x belongs to E_k and we know that we have pointwise convergence on E_k. Hence $f_n \to f$ for all $x \in E$. This is exactly convergence a.e. ■

In light of the previous proposition, we also have a converse under certain conditions.

Theorem. *Egoroff's Theorem*
Suppose $\mu(\Omega) < \infty$. If $f_n \to f$ converges a.e., then $f_n \to f$ almost uniformly.

Proof. We aim to translate the information of convergence into that of a set. Consider the set

$$X = \bigcap_{k\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} \bigcap_{l\geq n} \left\{x \in \Omega : |f_l(x) - f(x)| \leq \frac{1}{k}\right\}$$

The set X are the points where $f_n \to f$ pointwise. So, the assumption that $f_n \to f$ converges a.e. is the same as saying $\mu(X^c) = 0$. Taking the complement, we have

$$X^c = \bigcup_{k\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} \bigcup_{l\geq n} \left\{x \in \Omega : |f_l(x) - f(x)| > \frac{1}{k}\right\}$$

But since $\mu(X^c) = 0$, this is the same as saying

$$\mu\left(\bigcap_{n\in\mathbb{N}} \bigcup_{l\geq n} \left\{x \in \Omega : |f_l(x) - f(x)| > \frac{1}{k}\right\}\right) = 0, \ \ \forall k$$

Since k is fixed in the above expression, we can ignore the dependence of the inner union on k and say that $A_n = \bigcup_{l\geq n} \left\{x \in \Omega : |f_l(x) - f(x)| > \frac{1}{k}\right\}$. Notice that $A_n \supseteq A_{n+1}$, so A_n is a decreasing sequence which decreases to the intersection, that is $A_n \downarrow \bigcap_n A_n$. Since $\mu(\Omega) < \infty$, we may use continuity from above to conclude that for all k

$$\lim_{n\to\infty} \mu\left(\bigcup_{l\geq n} \left\{x \in \Omega : |f_l(x) - f(x)| > \frac{1}{k}\right\}\right) = 0$$

Fix $\varepsilon > 0$. Using the definition of limit, there exists $n_{\varepsilon,k}$ (this means n depends on ε and k) such that for all $n \geq n_{\varepsilon,k}$ we have

$$\mu\left(\bigcup_{l\geq n} \left\{x : |f_l(x) - f(x)| > \frac{1}{k}\right\}\right) \leq \frac{\varepsilon}{2^k}$$

Since this is true for all $n \geq n_{\varepsilon,k}$, we have in particular $\mu\left(\bigcup_{l\geq n_{\varepsilon,k}} \left\{x : |f_l(x) - f(x)| > \frac{1}{k}\right\}\right) \leq \frac{\varepsilon}{2^k}$. Now, taking the union, we have

$$\mu\left(\bigcup_{k\geq 1} \bigcup_{l\geq n_{\varepsilon,k}} \left\{x : |f_l(x) - f(x)| > \frac{1}{k}\right\}\right) \leq \varepsilon$$

We may call the set inside the measure E_ε^c. Hence

$$E_\varepsilon = \bigcap_{k\geq 1} \bigcap_{l\geq n_{\varepsilon,k}} \left\{x : |f_l(x) - f(x)| \leq \frac{1}{k}\right\}$$

It remains to show that (f_n) converges uniformly to f on E_ε. Let $\delta > 0$ be given. Choose k so that $1/k \leq \delta$ and set $N := n_{\varepsilon,k}$. For any $n \geq N$ and any $x \in E_\varepsilon$, we have $n \geq n_{\varepsilon,k}$, and by the definition of E_ε,

$$|f_n(x) - f(x)| \leq \frac{1}{k} \leq \delta.$$

Thus

$$\sup_{x \in E_\varepsilon} |f_n(x) - f(x)| \leq \delta \quad \text{for all } n \geq N,$$

which shows that $f_n \to f$ uniformly on E_ε. Since $\mu(E_\varepsilon^c) \leq \varepsilon$, this proves that $f_n \to f$ almost uniformly. ∎

22 Convergence in measure

When we talk about convergence in measure or other notions of convergence which involve a.e., it is useful to assume the sequence of functions is measurable (and hence the limiting function as well).

Definition. *Convergence in measure*
Let $f_n, f : \Omega \to \overline{\mathbb{R}}$. We say that $f_n \to f$ in measure if for all $\varepsilon > 0$ we have

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) \xrightarrow[n\to\infty]{} 0$$

Example. We show an example of a sequence of functions f_n such that $f_n \to f$ in measure, but not pointwise. Consider the measure space $([0,1), \mathcal{L}, \lambda)$. We construct a sequence of functions in the following way

$$\begin{aligned} f_1(x) &= 1, \ \ \forall x \in [0,1) \\ f_2(x) &= \chi_{[0,\frac{1}{2})}, \quad f_3(x) = \chi_{[\frac{1}{2},1)} \\ &\vdots \\ f_{2^{k-1}}(x) &= \chi_{[0,\frac{1}{2^{k-1}})}, \ \ldots \end{aligned}$$

It is clear that $f_n(x) \not\to 0$ for all $x \in [0,1)$. The sequence $f_n(x)$ alternates between 0 and 1 for all x. To show convergence in measure, fix $\varepsilon > 0$. If $\varepsilon \geq 1$, then we are done since the functions f_n never take on values greater than 1, so the measure of the set $\{x : |f_n(x) - f(x)| > \varepsilon\}$ is always zero. So, assume $0 < \varepsilon < 1$. Pick $2^{k-1} \leq j \leq 2^k - 1$. From our inequality on j, we see that $j \leq 2 \cdot 2^{k-1} - 1 < 2 \cdot 2^{k-1} \implies \frac{1}{2^{k-1}} < \frac{2}{j}$. Using the construction of the function, we have

$$\lambda(\{x : |f_j(x) - f(x)| > \varepsilon\}) = \frac{1}{2^{k-1}} < \frac{2}{j} \xrightarrow[j\to\infty]{} 0$$

What is important is that we are estimating the measure of the sets, and this measure can tend to zero while the sequence of functions at individual points do not convergence pointwise.

Observation. If $f_n \to f$ in measure and $f_n \to g$ in measure, then $f = g$ a.s.

Proof. Fix $\delta > 0$. Using the triangle inequality, observe that

$$\mu(\{x : |f(x) - g(x)| > \delta\}) \leq \mu(\{x : |f_n(x) - f(x)| > \frac{\delta}{2}\}) + \mu(\{x : |f_n(x) - g(x)| > \frac{\delta}{2}\})$$

But convergence in measure means $\mu(\{x : |f_n(x) - f(x)| > \frac{\delta}{2}\}) \xrightarrow[n\to\infty]{} 0$ and $\mu(\{x : |f_n(x) - g(x)| > \frac{\delta}{2}\}) \xrightarrow[n\to\infty]{} 0$. Putting things together, we have $\mu(\{x : |f(x) - g(x)| > \delta\}) = 0$.

To conclude the proof, we can take $\delta = \frac{1}{k}$ for all $k \in \mathbb{N}$ and then show that $\bigcup_{k=1}^{\infty}\{x : |f(x) - g(x)| > \frac{1}{k}\}$ has zero measure to show that $f = g$ a.s. ■

Observation 2. If $f_n \to f$ in measure, $g_n \sim f_n$ for all $n \in \mathbb{N}$ and $g \sim f$, then $g_n \to g$ in measure.

Proof. Call $E_0 = \{x : f(x) = g(x)\}$ and $E_n = \{x : f_n(x) = g_n(x)\}$. Set $E = \bigcap_{n=0}^{\infty} E_n$. Since $f_n \sim g_n$ for all $n \in \mathbb{N}$ we see that $\mu(E^c) = 0$. Now, we have

$$\begin{aligned}\mu(\{x : |g_n(x) - g(x)| > \varepsilon\}) &= \mu(\{x : |g_n(x) - g(x)| > \varepsilon\} \cap E) + \underbrace{\mu(\{x : |g_n(x) - g(x)| > \varepsilon\} \cap E^c)}_{0} \\ &= \mu(\{x : |g_n(x) - g(x)| > \varepsilon\} \cap E) \\ &= \mu(\{x : |f_n(x) - f(x)| > \varepsilon\} \cap E) \\ &\leq \mu(\{x : |f_n(x) - f(x)| > \varepsilon\})\end{aligned}$$

Since $f_n \to f$ in measure, we see that the last expression in the above goes to 0. ■

22.1 Connecting a.e. convergence and convergence in measure

Lemma. If $f_n \to f$ in measure, then there exists a subsequence n_k such that $f_{n_k} \to f$ a.s.

Proof. Since $f_n \to f$ in measure, for all $\varepsilon > 0$ we have

$$\mu(\{x \in \Omega : |f_n(x) - f(x)| > \varepsilon\}) \xrightarrow[n\to\infty]{} 0$$

We will abbreviate by $g_n(x) := f_n(x) - f(x)$. We want to construct a subsequence n_k so that $g_{n_k} \to 0$ a.s. Taking $\varepsilon = \frac{1}{k}$, there exists n_k such that for all $n \geq n_k$ we have

$$\mu\left(\left\{x : |g_n(x)| > \frac{1}{k}\right\}\right) \leq \frac{1}{2} \cdot \frac{1}{2^k}$$

So, we may construct a sequence so that

$$\mu\left(\left\{x : |g_{n_k}(x)| > \frac{1}{k}\right\}\right) \leq \frac{1}{2^{k+1}}$$

where we choose n_k so that it is strictly greater than the previously chosen $n_1, n_2, \ldots, n_{k-1}$. Taking unions, we see that

$$\mu\left(\bigcup_{k\geq k_0}\left\{x : |g_{n_k}(x)| > \frac{1}{k}\right\}\right) \leq \frac{1}{2^{k_0}}$$

Taking the intersection, we have

$$\mu\left(\bigcap_{k_0\geq 1}\bigcup_{k\geq k_0}\left\{x : |g_{n_k}(x)| > \frac{1}{k}\right\}\right) = 0$$

Since the above measure is bounded by $\frac{1}{2^{k_0}}$ for every k_0. We set $E^c = \bigcap_{k_0 \geq 1} \bigcup_{k \geq k_0} \left\{x : \left|g_{n_k(x)}\right| > \frac{1}{k}\right\}$. It remains to show that for all $x \in E$, we have $g_{n_k}(x) \xrightarrow[k\to\infty]{} 0$. Let $x_0 \in E$, while also noticing that

$$E = \bigcup_{k_0 \geq 1} \bigcap_{k \geq k_0} \left\{x : |g_{n_k}(x)| \leq \frac{1}{k}\right\}$$

By the construction of E, there exists some k_0 such that $x_0 \in \bigcap_{k \geq k_0} \left\{x : \left|g_{n_k(x)}\right| \leq \frac{1}{k}\right\}$. However, this means that for all $k \geq k_0$, $|g_{n_k}(x_0)| \leq \frac{1}{k}$. Since this holds for all $k \geq k_0$, we see that $g_{n_k}(x_0) \xrightarrow[k\to\infty]{} 0$. To reiterate, we have shown the existence of a set E on which $g_{n_k} \to 0$ and $\mu(E^c) = 0$ and this means that $g_{n_k} \to 0$ a.s. ■

Example. In the example from the beginning of the lecture we saw how convergence in measure does not imply convergence a.e. However, if we choose the subsequence $f_1, f_2, f_4, \ldots, f_{2^k}, \ldots$, then we obtain a sequence of functions converging pointwise to 0. So, the subsequence is $1, 2, 4, 8, \ldots, 2^k, \ldots$.

Now, we discuss a converse of sorts to the above lemma. First, we show that the converse without any extra conditions is false, that is, there exists a sequence $f_n \to f$ a.s. but such that $f_n \not\to f$ in measure.

Example. Define

$$f_n(x) = \begin{cases} 1, & \text{if } x \geq n \\ 0, & \text{otherwise} \end{cases}$$

For $0 < \varepsilon < 1$, we have

$$\lambda(\{x : |f_n(x)| > \varepsilon\}) = +\infty$$

for all n. So, $f_n \not\to 0$ in measure. However, $f_n \to 0$ a.s.

Proposition. If $f_n \to f$ a.s. and $\mu(\Omega) < \infty$, then $f_n \to f$ in measure.

Proof. By Egoroff's theorem, since $f_n \to f$ a.s. and $\mu(\Omega) < \infty$, we know that $f_n \to f$ almost uniformly. This means that for all $\varepsilon > 0$, there exists some E_ε such that $f_n \to f$ uniformly on E_ε and $\mu(E_\varepsilon^c) \leq \varepsilon$. To show $f_n \to f$ in measure, we must show

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) \xrightarrow[n\to\infty]{} 0$$

To be precise, we will show that for all $\delta > 0$, there exists some n_δ such that for all $n \geq n_\delta$ we have $\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) \leq \delta$. Fix some $\delta > 0$. Using almost uniform convergence, we know that there exists some E_δ such that $\mu(E_\delta^c) \leq \delta$ and $f_n \to f$ uniformly on E_δ. By uniform convergence, we know that there exists some $n_{\varepsilon,\delta}$ such that for all $n \geq n_{\varepsilon,\delta}$ we have

$$\sup_{x \in E_\delta} |f_n(x) - f(x)| \leq \varepsilon$$

Now we compute

$$\begin{aligned}\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) &\leq \mu(\{x : |f_n(x) - f(x)| > \varepsilon\} \cap E_\delta) + \mu(E_\delta^c)\\ &\leq 0 + \delta\end{aligned}$$

We have that $\mu(\{x : |f_n(x) - f(x)| > \varepsilon\} \cap E_\delta) = 0$ because if $n \geq n_{\varepsilon,\delta}$, then $|f_n(x) - f(x)| \leq \varepsilon$, so it cannot be in $\{x : |f_n(x) - f(x)| > \varepsilon\}$. This shows that the measure of $\{x : |f_n(x) - f(x)| > \varepsilon\}$ is bounded by δ, which completes the proof that $f_n \to f$ in measure. ■

23 Hölder and Minkowski Inequalities

Let $1 \leq p < \infty$. Define

$$\|f\|_p := \left(\int |f(x)|^p \, \mu(dx)\right)^{\frac{1}{p}}$$

Theorem. *Hölder's inequality*
Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $p = 1$, then $q = \infty$.
If $\|f\|_p < \infty$ and $\|g\|_q < \infty$, then

$$\int |fg| \, d\mu \leq \|f\|_p \, \|g\|_q$$

Recall that $\|g\|_\infty = \operatorname{ess\,sup} |g|$.

To prove this theorem, we must explain the convex conjugate of functions.
Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be convex. One example of such a function is $F(x) = \frac{x^p}{p}$ for $p > 1$, and we will consider this function. Assume also that p, q are conjugate, that is, $\frac{1}{p} + \frac{1}{q} = 1$. We also assume $1 < p < \infty$. We will show that if $x, y > 0$, then $xy \leq \frac{x^p}{p} + \frac{x^q}{q}$. To prove this we need to introduce the Legendre transform of a convex function.

$$TF(\theta) = \sup_{x>0}\{\theta x - F(x)\} = \sup_{x>0}\left\{\theta x - \frac{x^p}{p}\right\}$$

Set $G_\theta(x) = \theta x - \frac{x^p}{p}$. Observe that $G_\theta(0) = 0$ and $G'_\theta(x) = \theta > 0$. Further, $\lim_{x\to\infty} G_\theta(x) = -\infty$, since the $\frac{x^p}{p}$ dominates. To determine the supremum, we need information about the critical points. Assume that $\theta - x^{p-1} = G'_\theta(x) = 0$, then we have $x = \theta^{\frac{1}{p-1}}$. Since $G_\theta(0) = 0$ and goes to $-\infty$ as $x \to \infty$, we see that there is a supremum at $x = \theta^{\frac{1}{p-1}}$. By substitution

$$\begin{aligned}\sup_{x>0}\left\{\theta x - \frac{x^p}{p}\right\} &= \theta^{1+\frac{1}{p-1}} - \frac{1}{p}\theta^{\frac{p}{p-1}} \\ &= \frac{1}{q}\theta^{\frac{p}{p-1}} = \frac{1}{q}\theta^q\end{aligned}$$

But now

$$\frac{\theta^q}{q} = \sup_{x>0}\left\{\theta x - \frac{x^p}{p}\right\} \geq \theta x - \frac{x^p}{p}$$

for all $\theta, x > 0$. But this is exactly the inequality we needed the prove.

Aside. In general, we have

$$\theta x \leq F(x) + (TF)(\theta)$$

where TF is the Legendre transform.

Proof. We now prove Hölder's inequality. Above, we have shown the inequality $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$ for $x, y > 0$. By the monotonicity and linearity of the integral, we have

$$\int |fg| \, d\mu \leq \int \frac{|f|^p}{p} \, d\mu + \int \frac{|g|^q}{q} \, d\mu$$

When $\|f\|_p$ and $\|g\|_q$ are finite, we see that the above equation immediately gives us the finiteness of $\int |fg| \, d\mu$. To obtain our desired inequality, we employ a trick. Pick $A > 0$ and observe

$$\int |fg| d\mu = \int \left| \frac{f}{A} gA \right| d\mu \leq \int \frac{|f|^p}{A^p p} d\mu + \int \frac{A|g|^q}{q} d\mu$$

On the left hand side, we have an expression that does not depend on A while on the right hand side, we have an expression that does, so we may optimize the right hand side to obtain

$$\int |fg| d\mu \leq \inf_A \left\{ \int \frac{|f|^p}{A^p p} d\mu + \int \frac{A|g|^q}{q} d\mu \right\}$$

Set $\alpha = \int |f|^p \, d\mu$ and $\beta = \int |g|^q \, d\mu$. We may write the above infimum as

$$\inf_{A>0} \left\{ \alpha \frac{1}{pA^p} + \beta \frac{A^q}{q} \right\}$$

Since this function approaches ∞ as $A \to \infty$, to find the infimum we should analyze the critical points again. Setting the derivative with respect to A to 0, we have the equation

$$\alpha A^{-p-1} = \beta A^{q-1}$$

Solving for A yields $A = \left(\frac{\alpha}{\beta}\right)^{\frac{1}{q+p}}$. If we plug this value of A back into the expression in the infimum, we derive

$$\begin{aligned} \frac{\alpha}{p} \left(\frac{\beta}{\alpha}\right)^{\frac{p}{q+p}} + \frac{\beta}{q} \left(\frac{\alpha}{\beta}\right)^{\frac{q}{p+q}} &= \frac{1}{p} \alpha^{\frac{q}{q+p}} \beta^{\frac{p}{q+p}} + \frac{1}{q} \beta^{\frac{p}{p+q}} \alpha^{\frac{q}{p+q}} \\ &= \alpha^{\frac{q}{q+p}} \beta^{\frac{p}{q+p}} \end{aligned}$$

Substituting for α and β and simplifying $\frac{q}{q+p} = \frac{1}{p}$ and $\frac{p}{q+p} = \frac{1}{q}$, we obtain our desired inequality.

Now, we prove the case where $p = 1$ and $q = \infty$, that is, we want to prove

$$\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$$

Recall that $\|f\|_1 = \int |f| \, d\mu$ and $\|g\|_\infty = \operatorname{ess\,sup} |g|$. Set $c = \operatorname{ess\,sup} |g|$ From a previous proposition, we know that $\mu\{x : |g(x)| > c\} = 0$. Using the linearity of the domain, we may write

$$\begin{aligned} \int |fg| \, d\mu &= \int_{\{x:|g(x)|>c\}} |fg| \, d\mu + \int_{\{x:|g(x)|\leq c\}} |fg| \, d\mu \\ &= \int_{\{x:|g(x)|\leq c\}} |fg| \, d\mu \leq c \int_{\{x:|g(x)|\leq c\}} |f(x)| \, d\mu \\ &\leq c \int f \, d\mu = \|g\|_\infty \|f\|_1 \end{aligned}$$

This completes the proof. ■

Theorem. *Minkowski's inequality*
Let $1 \le p < \infty$. Let f, g be so that $\|f\|_p < \infty$ and $\|g\|_p < \infty$. Then

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

Proof. Note that if $\|f\|_p$ or $\|g\|_p$ is infinite, the inequality is trivial, so assume both are finite. This theorem follows from the usual triangle inequality and the monotonicity of the integral for $p = 1$. So, assume $p > 1$. Observe that

$$\int |f+g|^p \, d\mu = \int |f+g|^{p-1}|f+g| \, d\mu \le \int |f+g|^{p-1}(|f|+|g|) \, d\mu$$

We estimate each of the parts of the right-hand expression. Using Hölder's inequality, we have

$$\int |f+g|^{p-1}|f| \, d\mu \le \left(\int |f+g|^{q(p-1)} \, d\mu \right)^{\frac{1}{q}} \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}}$$
$$\int |f+g|^{p-1}|g| \, d\mu \le \left(\int |f+g|^{q(p-1)} \, d\mu \right)^{\frac{1}{q}} \|g\|_p$$

where q is the conjugate of p. Hence

$$\begin{aligned} \int |f+g|^p \, d\mu &\le \left(\int |f+g|^{q(p-1)} \, d\mu \right)^{\frac{1}{q}} (\|f\|_p + \|g\|_p) \\ &= \left(\int |f+g|^p \, d\mu \right)^{\frac{1}{q}} (\|f\|_p + \|g\|_p) \end{aligned}$$

We may assume that $|f+g| \ne 0$, otherwise the inequality holds. In this case, we may divide to see that

$$\|f+g\|_p = \left(\int |f+g|^p \, d\mu \right)^{1-\frac{1}{q}} \le \|f\|_p + \|g\|_p$$

But $1 - \frac{1}{q} = \frac{1}{p}$, so our desired inequality follows. ■

Remark. In the case where $p = \infty$, we have already proven this when we demonstrated that $d(f,g) = \operatorname{ess\,sup} |f,g|$ is a metric. For instance, observe that

$$d(f+g, 0) \le d(f+g, g) + d(g, 0)$$
$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$$

So, Minkowski's inequality holds for all $1 \le p \le \infty$. Now, we may define the space

$$L^p(\mu) = \left\{ f : \Omega \to \overline{\mathbb{R}} : \int |f|^p \, d\mu < \infty \right\}$$

Observation. If $f, g \in L^p$ and $a \in \mathbb{R}$, then $af + g \in L^p$. So, L^p is a linear space.

Proof. The homogeneity condition follows immediately when we write $|af|^p = |a|^p|f|^p$. The additivity condition follows from Minkowski's inequality. If $f, g \in L^p$, this means that $\|f\|_p$ and $\|g\|_p$ are finite. So, we apply Minkowski's inequality to see that $\|f + g\|_p$ is bounded above by a finite number and so $f + g \in L^p$ as well. ■

Proposition. $\|f\|_p$ is a norm.

Proof. There are three properties to verify

1. If $a \in \mathbb{R}$ and $f \in L^p$, then we have $\|af\|_p = |a| \, \|f\|_p$. This follows immediately if we inspect the integrals.

2. If $\|f\|_p = 0$, then $f = 0$ a.s. This follows from the properties of the integral we have proven earlier.

3. $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. This is Minkowski's inequality, which we have proven above.

■

24 L^p Spaces

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $f : \Omega \to \overline{\mathbb{R}}$. Recall that for $1 \leq p < \infty$

$$\|f\|_p = \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}}$$

Also, recall that

$$L^p = \left\{ f : \Omega \to \overline{\mathbb{R}} : \|f\|_p < \infty \right\}$$

Further, as we have seen in the last lecture, L^p is a linear space. The goal of this lecture will be to show that the L^p norm defines a distance and that the vector space is complete with respect to that distance.

Proposition. Taking $d(f,g) = \|f - g\|_p$ defines a distance function, or metric.

Proof. Suppose that $d(f,g) = 0$. Then $\int |f-g|^p \, d\mu = 0$. But this means that $f = g$ a.s., showing that they belong to the same equivalence class. It is also clear that $d(f,g) = d(g,f)$ using the definition of absolute value. We may obtain the triangle inequality from Minkowski's inequality. Observe that

$$d(f,h) = \|f - h\|_p = \|f - g + g - h\|_p \leq \|f - g\|_p + \|g - h\|_p = d(f,g) + d(g,h)$$

Note that two functions that are equivalent $f \sim g$ are considered to be equal in L^p, as we are dealing with equivalence classes of functions. ■

Definition. *Convergence in L^p*
We say that $f_n \to f$ in L^p if $\|f_n - f\|_p \to 0$.

We list a few examples below. The key theme is that convergence a.e. or convergence in measure is not enough to imply convergence in L^p.

Example. Define

$$f_n(x) = \begin{cases} n^{\frac{1}{p}}, & \text{if } x \in [0, \frac{1}{n}] \\ 0, & \text{otherwise} \end{cases}$$

We see that $f_n \to 0$ a.s. For every point other than $x = 0$, we can find some $N \in \mathbb{N}$ such that $f_n(x) = 0$ for all $n \geq N$. We also have $f_n \to 0$ in measure. However, observe that

$$\|f_n - 0\|_p^p = \int |f_n|^p \, d\lambda = 1$$

Example. Define

$$f_n(x) = \begin{cases} \frac{1}{n^{\frac{1}{p}}}, & \text{if } x \in [0, n] \\ 0, & \text{otherwise} \end{cases}$$

This converges to 0 uniformly and in measure, but does not converge in L^p since

$$\|g_n - 0\|_p^p = \int |g_n|^p \, d\lambda = 1$$