UNIVERSITY OF WARWICK

MA359 MEASURE THEORY

Autumn Term 2002

Lectured by Doctor Omri Sarig

Typed by Tim Sullivan

INDEX

0. Introduction

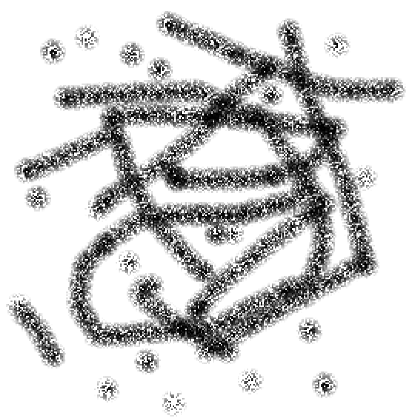
Problem. We wish to assign a numerical size $\mu(E)$ to subsets E of a given space X.

Geometry: $X = \mathbb{R}^2$; $\mu(E) = \text{Area}(E)$.

Mechanics: $X = \mathbb{R}^2 =$ model for thin sheet of metal with density $\rho(x, y)$; $\mu(E) =$ mass of E.

Probability Theory: $X =$ sample space, collection of possible outcomes; $E \subseteq X$ an "event"; $\mu(E) =$ probability that E happens.

However, there are problems. For instance, what is the area of this?



"Magnification" doesn't help. We encounter similar problems with the idea of mass – what is the mass of, say, a sponge? Or, in the case of probability, imagine picking a real number "at random" – what is the probability that the number chosen is rational?

What would we like to have for $\mathbb{R}$?

Idea. (1) A function μ with domain $2^{\mathbb{R}}$ and range $[0, +\infty]$; $\mu(\varnothing) = 0$, $\mu(\mathbb{R}) = \infty$.
(2) Translation-invariance: $\mu(x + E) = \mu(E)$ for all $x \in \mathbb{R}$, $E \subseteq \mathbb{R}$.
(3) "Consistency":

- (i) Monotonicity: $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$;
- (ii) $\mu(A \coprod B) = \mu(A) + \mu(B)$;
- (iii) σ-additivity: $\mu(\coprod_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mu(A_i)$;
- (iv) $\mu(\varnothing) = 0$;

Sad Fact. These requirements are *not* self-consistent!

Condition (1) is very strong because there are too many subsets:

$$\left|2^{\mathbb{R}}\right| = 2^{2^{\aleph_0}} >> 2^{\aleph_0} >> \aleph_0 = |\mathbb{N}|$$

σ-additivity is also asking for a lot because limits are involved.

The standard choice is to relax (1) and keep σ-additivity for the sake of practicality and an easy life. Relaxing (1) means that the domain for μ should only be a subcollection $\mathcal{F}$ of subsets of X. We would like $\mathcal{F}$ to have the following properties:
(1) $\varnothing, X \in \mathcal{F}$;
(2) $A, B \in \mathcal{F} \Rightarrow A \cup B, A \cap B, A \setminus B, A^{C} \in \mathcal{F}$;
(3) $\{A_i\}_{i=1}^{\infty}$ in $\mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i, \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

We do not require that $\mathcal{F}$ be closed under arbitrary unions and intersections, simply countable ones.

Definition. Let X be a set. A σ*-algebra* of subsets of X is a collection of subsets of X that contains $\varnothing$ and is closed under complementation and countable unions of its members.

Exercise. Prove that a σ-algebra satisfies the properties given above.

Definition. A *measurable space* is a pair $(X, \mathcal{F})$, where X is a set and $\mathcal{F}$ is a σ-algebra of subsets of X. Elements of $\mathcal{F}$ are called $\mathcal{F}$*-measurable*, or simply *measurable*.

Definition. Let $(X, \mathcal{F})$ be a measurable space. A *measure* on $(X, \mathcal{F})$ is a function $\mu : \mathcal{F} \to [0, +\infty]$ such that
(1) $\mu(\varnothing) = 0$;
(2) μ is σ*-additive*, i.e. if $A_i \in \mathcal{F}$ are pairwise disjoint then $\mu\left(\coprod_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i)$.
The triple $(X, \mathcal{F}, \mu)$ is called a *measure space*.

Proposition. (Basic properties of measures.) *Let $(X, \mathcal{F}, \mu)$ be a measure space.*
(1) Additivity: $A_1, \ldots, A_n \in \mathcal{F}$ pairwise disjoint $\Rightarrow \mu\left(\coprod_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$;
(2) Monotonicity: $A, B \in \mathcal{F}, A \subseteq B \Rightarrow \mu(A) \le \mu(B)$;
(3) Inclusion-Exclusion Principle: if $A, B \in \mathcal{F}$ and $\mu(A \cap B) < \infty$ then

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

(4) Difference Formula: if $A, B \in \mathcal{F}$, $A \subseteq B$, $\mu(A) < \infty$ then $\mu(B \setminus A) = \mu(B) - \mu(A)$.

Proof. (1) Define

$$B_i = \begin{cases} A_i & 1 \le i \le n \\ \varnothing & i > n \end{cases}$$

By σ-additivity,

$$\begin{aligned} \mu\left(\coprod_{i=1}^{n} A_i\right) &= \mu\left(\coprod_{i=1}^{\infty} B_i\right) \\ &= \sum_{i=1}^{\infty} \mu(B_i) \\ &= \sum_{i=1}^{n} \mu(A_i) \end{aligned}$$

(2) $\mu(B) = \mu(A \coprod (B \setminus A)) = \mu(A) + \mu(B \setminus A) \ge \mu(A)$.

(4) $\mu(B) = \mu(A) + \mu(B \setminus A)$. If $\mu(A) < \infty$ as given we can subtract $\mu(A)$ from both sides to get $\mu(B) - \mu(A) = \mu(B \setminus A)$.

(3)

$$\begin{aligned} \mu(A \cup B) &= \mu(A \coprod (B \setminus A)) \\ &= \mu(A) + \mu(B \setminus A) \\ &= \mu(A) + \mu(B \setminus (A \cap B)) \\ &= \mu(A) + \mu(B) - \mu(A \cap B) \end{aligned}$$

■

Definition. (Monotone sequences of sets.) We say that a sequence $\{E_i\}_{i=1}^{\infty}$ *increases* (respectively *decreases*) to a set E if

$$\forall i \in \mathbb{N},\ E_i \subseteq E_{i+1} \text{ and } \bigcup_{i=1}^{\infty} E_i = E,$$

and write $E_i \uparrow E$ (respectively

$$\forall i \in \mathbb{N},\ E_i \supseteq E_{i+1} \text{ and } \bigcap_{i=1}^{\infty} E_i = E,$$

and write $E_i \downarrow E$).

Proposition. (Continuity of measures.) *Let $(X, \mathcal{F}, \mu)$ be a measure space and let $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$.*

(1) If $E_i \uparrow E$ then $\mu(E) = \lim_{i \to \infty} \mu(E_i)$.

(2) If $E_i \downarrow E$ *and* $\exists i_0$ *such that* $\mu(E_{i_0}) < \infty$ *then* $\mu(E) = \lim_{i\to\infty} \mu(E_i)$.

Proof. (1) Assume first that $\mu(E_i) < \infty$ for all i. We claim that

$$E = \bigcup_{i=1}^{\infty} E_i = E_1 \amalg \left(\coprod_{i=2}^{\infty} E_i \setminus E_{i-1}\right).$$

Indeed, $\text{LHS} \subseteq \text{RHS}$ since if $x \in E$ then there is a minimal i such that $x \in E_i$. By minimality, $x \notin E_{i-1}$, so $x \in E_i \setminus E_{i-1} \subseteq \text{RHS}$. $\text{RHS} \subseteq \text{LHS}$ is trivial.

To see that the union on the RHS is pairwise disjoint suppose $i < j$. Then $E_j \setminus E_{j-1} \subseteq E_j \setminus E_i$ since $E_i \subseteq E_{i+1} \subseteq \ldots \subseteq E_{j-1} \subseteq E_j$. So, since $E_j \setminus E_{j-1} \subseteq E_i^C$ and $E_i \setminus E_{i-1} \subseteq E_i$, $(E_j \setminus E_{j-1}) \cap (E_i \setminus E_{i-1}) = \varnothing$.

We now use the σ-additivity of μ to calculate $\mu(E)$:

$$\begin{aligned}
\mu(E) &= \mu(E_1) + \sum_{i=2}^{\infty} \mu(E_i \setminus E_{i-1}) \\
&= \mu(E_1) + \sum_{i=2}^{\infty} (\mu(E_i) - \mu(E_{i-1})) \\
&= \lim_{n\to\infty} (\mu(E_1) + (\mu(E_2) - \mu(E_1)) + \ldots + (\mu(E_n) - \mu(E_{n-1}))) \\
&= \lim_{n\to\infty} \mu(E_n)
\end{aligned}$$

This proves (1) under the assumption that $\mu(E_i) < \infty$ for all i. If $\exists i_0$ such that $\mu(E_{i_0}) = \infty$ then by Monotonicity $i \geq i_0 \Rightarrow \mu(E_i) = \infty$ and $\mu(E) \geq \mu(E_i) = \infty$, hence

$$\mu(E) = \infty = \lim_{i\to\infty} \infty = \lim_{i\to\infty} \mu(E_i),$$

thus proving (1).

(2) By de Morgan,

$$E_{i_0} \setminus \bigcap_{i=i_0}^{\infty} E_i = \bigcup_{i=i_0+1}^{\infty} (E_{i_0} \setminus E_i)$$

and

$$(E_{i_0} \setminus E_i) \uparrow \left(E_{i_0} \setminus \bigcap_{i=i_0}^{\infty} E_i\right)$$

by (1). So

$$
\begin{aligned}
\mu\left(E_{i_0}\right)-\mu\left(\bigcap\nolimits_{i=i_0}^{\infty}\right) &= \mu\left(E_{i_0} \setminus \bigcap\nolimits_{i=i_0}^{\infty} E_i\right) \\
&= \mu\left(\bigcup\nolimits_{i=i_0}^{\infty} E_{i_0} \setminus E_i\right) \\
&= \lim_{i\to\infty} \mu\left(E_{i_0} \setminus E_i\right) \\
&= \mu\left(E_{i_0}\right) - \lim_{i\to\infty} \mu\left(E_i\right)
\end{aligned}
$$

$$
\Rightarrow \qquad \mu\left(\bigcap\nolimits_{i=i_0}^{\infty} E_i\right) = \lim_{i\to\infty} \mu\left(E_i\right)
$$

We observe that $\bigcap_{i=i_0}^{\infty} E_i = \bigcap_{i=1}^{\infty} E_i$.

■

1. LEBESGUE'S MEASURE

Aim. Define a measure m on $\mathbb{R}$ such that
(1) $m(\text{interval}) = \text{length}(\text{interval})$;
(2) translation-invariance, i.e. $m(x+E) = m(E)$.

Strategy. 'Do what we must.'
(1) Define the measure on certain basic sets for which the definition is obvious – the intervals.
(2) Extend the definition to all sets by approximating them by countable unions of basic sets.
(3) Restrict to a σ-algebra to get σ-additivity.

Step 1 – Defining the Measure on Intervals

We begin with some notation and definitions.

An *interval* is a set of the form (a,b), $(a,b]$, $[a,b)$, $[a,b]$ where $a \le b$. We also allow $\varnothing$, $\mathbb{R}$, (a,∞), $[a,\infty)$, $(-\infty,b)$, $(-\infty,b]$. We shall use $\langle a,b \rangle$ to represent all the possibilities above.

$$\text{Int} = \{ E \subseteq \mathbb{R} \mid E \text{ is an interval} \}$$

Let X be a set. A collection $\mathcal{S} \subseteq 2^X$ is called a *semi-algebra* (of subsets of X) if
(1) $\varnothing, X \in \mathcal{S}$;
(2) $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$;
(3) $\forall A \in \mathcal{S}$, A^{C} can be written as a finite union of pairwise disjoint elements of $\mathcal{S}$.

Proposition. Int *is a semi-algebra of subsets of* $\mathbb{R}$ *(but is not a* σ*-algebra).*

Proof. (1) $\varnothing = (0,0)$, $\mathbb{R} = (-\infty, +\infty)$.
(2) The intersection of two intervals is empty, an interval, or a point $[a,a]$.
(3) The complement of an interval is empty, an interval, or the union of two disjoint intervals.

■

Definition. Let $\ell : \text{Int} \to [0,\infty]$ be the *length function* given by

$$\ell(\langle a,b \rangle) = b - a .$$

Is ℓ σ-additive on Int? I.e. if $I = \coprod_{k=1}^{\infty} I_k$, where $I, I_k \in \text{Int}$, does it follow that

$$\ell(I) = \sum_{k=1}^{\infty} \ell(I_k)?$$

Lemma. *(1) ℓ is finitely additive: for $I, I_k \in \text{Int}$,*

$$I = \coprod_{k=1}^{N} I_k \Rightarrow \ell(I) = \sum_{k=1}^{N} \ell(I_k).$$

(2) ℓ is σ-sub-additive: for $I, I_k \in \text{Int}$,

$$I = \bigcup_{k=1}^{\infty} I_k \Rightarrow \ell(I) \leq \sum_{k=1}^{\infty} \ell(I_k).$$

Proof. (1) If I is not bounded then $\ell(I) = \infty$. In this case at least one of the I_k, I_j say, must be unbounded (because a finite union of bounded sets is bounded). Thus,

$$\ell(I) = \infty = \ell(I_j) \leq \sum_{k=1}^{N} \ell(I_k)$$

$$\Rightarrow \qquad \ell(I) = \infty = \sum_{k=1}^{N} \ell(I_k)$$

From now on, assume I, I_k are bounded; $I = \langle a, b \rangle$, $I_k = \langle a_k, b_k \rangle$. Re-order so that $a_1 < a_2 < \ldots < a_N$. Observe that

(i) $\forall k$, $b_k \leq a_{k+1}$ since if not there would be an overlap, $I_k \cap I_{k+1} \supseteq (a_{k+1}, a_{k+1} + \varepsilon]$ would be non-empty;
(ii) $\forall i \leq N-1$, $b_k = a_{k+1}$ since if not there would be a gap, (b_k, a_{k+1}), yet I can have no gaps.

By (ii),

$$\begin{aligned} \sum_{k=1}^{N} \ell(I_k) &= \sum_{k=1}^{N} (b_k - a_k) \\ &= \sum_{k=1}^{N-1} (a_{k+1} - a_k) + (b_N - a_N) \\ &= (a_N - a_1) + (b_N - a_N) \\ &= b_N - a_1 \end{aligned}$$

Since the a_k were ordered, $b_N = b$ and $a_1 = a$, so

$$\ell(I) = \sum_{k=1}^{N} \ell(I_k) = b - a.$$

(2) *Case 1: I is a compact interval $[a, b]$.*

Write $I_i = \langle a_i, b_i \rangle$, fix $\varepsilon > 0$ and set $J_i = \left(a_i - \varepsilon/2^{i+1}, b_i + \varepsilon/2^{i+1}\right)$. $\{J_i\}_{i=1}^{\infty}$ is an open cover of the compact interval I. Therefore, by the Heine-Borel Theorem, there exists a finite subcover; fix N and re-order so that $I \subseteq \bigcup_{i=1}^{N} J_i$. Write

$$\alpha_i = a_i - \varepsilon/2^{i+1}$$
$$\beta_i = b_i + \varepsilon/2^{i+1}$$

$a \in I \subseteq \bigcup_{i=1}^{N} J_i$ so $\exists i_1$ such that $a \in J_{i_1} \Leftrightarrow \alpha_{i_1} < a < \beta_{i_1}$. If $I \subseteq J_{i_1}$ we stop. If not, $\beta_{i_1} \in I$ and so $\exists i_2$ such that $a \in J_{i_2} \Leftrightarrow \alpha_{i_2} < \beta_{i_1} < \beta_{i_2}$. If $I \subseteq J_{i_1} \cup J_{i_2}$ we stop. If not, $\beta_{i_2} \in I$ and so $\exists i_3$ such that $a \in J_{i_3} \Leftrightarrow \alpha_{i_3} < \beta_{i_2} < \beta_{i_3}$. Continue this process; this process stops eventually and gives us $i_1, \ldots, i_{N'}$, $N' \le N$, such that

$$I \subseteq \bigcup_{k=1}^{N'} J_{i_k},$$
$$\alpha_{i_1} < a < \beta_{i_1},\ \alpha_{i_{k+1}} < \beta_{i_k} < \beta_{i_{k+1}},\ \alpha_{i_{N'}} < b < \beta_{i_{N'}}.$$

Therefore, because of overlaps,

$$\sum_{k=1}^{N'} \ell\left(J_{i_k}\right) \ge \ell([a,b]) = \ell(I).$$

Thus,

$$\begin{aligned}
\ell(I) &\le \sum_{k=1}^{N'} \ell\left(J_{i_k}\right) \\
&\le \sum_{i=1}^{\infty} \ell(J_i) \\
&= \sum_{i=1}^{\infty} \left(\ell(I_i) + 2\tfrac{\varepsilon}{2^{i+1}}\right) \\
&= \sum_{i=1}^{\infty} \ell(I_i) + \sum_{i=1}^{\infty} \varepsilon/2^i \\
&= \sum_{i=1}^{\infty} \ell(I_i) + \varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $\ell(I) \le \sum_{i=1}^{\infty} \ell(I_i)$. Thus, Case 1 is proved.

Case 2: I is a bounded interval, I_k arbitrary intervals.

If I_k bounded for all k we are in Case 1. If not $\exists k$ such that $\ell(I_k) = \infty$ and

$$\ell(I) < \infty = \ell(I_k) \le \sum_{i=1}^{\infty} \ell(I_k).$$

Case 3: General case.

Write $I = \langle a,b \rangle$, a,b possibly infinite. Choose $a_k \downarrow a$, $b_k \uparrow b$, $a_k \le b_k$. Clearly

$$[a_k, b_k] \subseteq \langle a_k, b_k \rangle \subseteq \bigcup_{k=1}^{\infty} I_k.$$

By Case 2, $\ell([a_k,b_k]) \le \sum_{k=1}^{\infty} \ell(I_k)$. But $\ell([a_k,b_k]) = b_k - a_k \underset{k\to\infty}{\to} \ell(I)$, so $\ell(I) \le \sum_{i=1}^{\infty} \ell(I_k)$.

■

Tricks. (1) $\varepsilon/2^i$ trick – we need summable errors.
(2) Approximating sets from within by compact sets.

We need to check that ℓ is σ-additive on Int otherwise we have no chance of getting a measure out of ℓ.

Proposition. $\ell : \mathrm{Int} \to [0,\infty]$ *is* σ*-additive on* Int. *I.e., if* $I, I_k \in \mathrm{Int}$ *and* $I = \coprod_{k=1}^{\infty} I_k$ *then* $\ell(I) = \sum_{k=1}^{\infty} \ell(I_k)$.

Proof. ($\le$) This follows from σ-sub-additivity.

($\ge$) It is enough to prove that $\ell(I) \ge \sum_{k=1}^{N} \ell(I_k)$ for each N, since we can then pass to the limit $N \to \infty$. The idea is to show that $I \setminus \bigcup_{k=1}^{N} I_k$ is a finite pairwise disjoint union of intervals. This will do because if

$$I = \left(\coprod_{k=1}^{N} I_k\right) \coprod \left(\coprod_{i=1}^{M} J_i\right), \quad \text{with } J_i \in \mathrm{Int} \qquad (*)$$

then the finite additivity of ℓ on Int would imply

$$\ell(I) = \sum_{k=1}^{N} \ell(I_k) + \underbrace{\sum_{i=1}^{M} \ell(J_i)}_{\ge 0} \ge \sum_{k=1}^{N} \ell(I_k).$$

To prove $(*)$ we can use the properties of Int as a semi-algebra. We claim that

$$I \setminus \bigcup_{k=1}^{N} I_k \underset{\text{de Morgan}}{=} I \cap I_1^{\mathrm{C}} \cap I_2^{\mathrm{C}} \cap \ldots \cap I_N^{\mathrm{C}}$$

is a finite disjoint union of intervals and prove this by induction on N.

Case $N=1$: Since Int is a semi-algebra $I_1^C = \coprod_{k=1}^n J_k$ for some $J_k \in \text{Int}$. Therefore, $I \cap I_1^C = \coprod_{k=1}^n I \cap J_k$, but $I \cap J_k \in \text{Int}$.

Induction Step: Assume $I \cap I_1^C \cap I_2^C \cap \ldots \cap I_N^C = \coprod_{k=1}^n J_k$ with $J_k \in \text{Int}$. Since Int is a semi-algebra, $I_{N+1}^C = \coprod_{i=1}^m K_i$, $K_i \in \text{Int}$. So,

$$\begin{aligned} I \cap I_1^C \cap I_2^C \cap \ldots \cap I_{N+1}^C &= \left(\coprod_{k=1}^n J_k\right) \cap \left(\coprod_{i=1}^m K_i\right) \\ &= \coprod_{\substack{1 \le i \le m \\ 1 \le k \le n}} \left(J_k \cap K_i\right) \end{aligned}$$

Exercise. Check that this union is pairwise disjoint.

Since Int is a semi-algebra, $(*)$ is proved, and so is the proposition.

■

Exercise. Modify the proof of $(\ge)$ to obtain the stronger statement: If $I, I_j \in \text{Int}$ and $I \supseteq \bigcup_{j=1}^\infty I_j$ then $\ell(I) \ge \sum_{j=1}^\infty \ell(I_j)$.

Step 2 – Outer Measures

Or, "Approximating the measure of an arbitrary set from above".

We want to bound the measure of a general set from above. We do this by considering countable covers by basic sets.

Definition. *Lebesgue's outer measure* is the function $m^* : 2^{\mathbb{R}} \to [0,\infty]$ given by

$$m^*(A) = \left\{ \sum_{k=1}^\infty \ell(I_k) \;\middle|\; A \subseteq \bigcup_{k=1}^\infty I_k, I_k \in \text{Int} \right\}.$$

Definition. In general, given a set X, an *outer measure* on X is a function $\mu^* : 2^X \to [0,\infty]$ such that

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

(2) monotonicity: $A \subseteq B \Rightarrow \mu^*(A) \le \mu^*(B)$;

(3) σ-sub-additivity: $\mu^*\left(\bigcup_{i=1}^\infty A_i\right) \le \sum_{i=1}^\infty \mu^*(A_i)$.

Remark. An outer measure is not, in general, a measure.

Proposition. *Lebesgue's outer measure is an outer measure.*

Proof. (1) $\varnothing \subseteq \bigcup_{i=1}^{\infty} I_i$ where $I_i = (1,1) = \varnothing$, and so $m^*(\varnothing) \leq \sum_{i=1}^{\infty} 0 = 0$. Obviously $m^*(A) \geq 0 \;\; \forall A \subseteq \mathbb{R}$ so $m^*(\varnothing) = 0$.

(2) Fix $\varepsilon > 0$ and find $\{I_i\}_{i=1}^{\infty} \subseteq \text{Int}$ for which

$$B \subseteq \bigcup_{i=1}^{\infty} I_i,$$
$$\sum_{i=1}^{\infty} \ell(I_i) \leq m^*(B) + \varepsilon.$$

(The cover exists by the definition of infimum.) Since $A \subseteq B \subseteq \bigcup_{i=1}^{\infty} I_i$ we must also have

$$m^*(A) \leq \sum_{i=1}^{\infty} \ell(I_i) \leq m^*(B) + \varepsilon,$$

but $\varepsilon > 0$ is arbitrary so $m^*(A) \leq m^*(B)$.

(3) Let $\varepsilon > 0$. For each i find $\left\{I_j^{(i)}\right\}_{j=1}^{\infty} \subseteq \text{Int}$ such that

$$A_i \subseteq \bigcup_{j=1}^{\infty} I_j^{(i)},$$
$$\sum_{j=1}^{\infty} \ell\left(I_j^{(i)}\right) \leq m^*(A_i) + \varepsilon/2^i.$$

Clearly,

$$\bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i=1}^{\infty}\left(\bigcup_{j=1}^{\infty} I_j^{(i)}\right).$$

The collection $\left\{I_j^{(i)} \;\middle|\; i, j \in \mathbb{N}\right\}$ is countable and so

$$\begin{aligned} m^*\left(\bigcup_{i=1}^{\infty} A_i\right) &\leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \ell\left(I_j^{(i)}\right) \\ &\leq \sum_{i=1}^{\infty}\left(m^*(A_i) + \varepsilon/2^i\right) \\ &= \sum_{i=1}^{\infty} m^*(A_i) + \varepsilon \end{aligned}$$

So m^* is σ-sub-additive.

■

Question. Is it true that $m^* : 2^{\mathbb{R}} \to [0,\infty]$ is an extension of the length function, i.e. $\forall I \in \text{Int}$, $m^*(I) = \ell(I)$?

Proposition. *The Lebesgue outer measure of an interval is its length. Moreover, if* $A = \coprod_{i=1}^{\infty} I_k$, $I_k \in \text{Int}$*, then* $m^*(A) = \sum_{k=1}^{\infty} \ell(I_k)$.

Proof. ($\leq$) $A \subseteq \coprod_{i=1}^{\infty} I_k$ so $m^*(A) \leq \sum_{k=1}^{\infty} \ell(I_k)$.

($\geq$) Fix $\varepsilon > 0$ and let $\{J_i\}_{i=1}^{\infty}$ be a cover of $\coprod_{k=1}^{\infty} I_k$ by $J_i \in \text{Int}$ such that

$$\sum_{i=1}^{\infty} \ell(J_i) \leq m^*\left(\coprod_{k=1}^{\infty} I_k\right) + \varepsilon .$$

Since $\coprod_{k=1}^{\infty} I_k \subseteq \bigcup_{i=1}^{\infty} J_i$, $I_k \subseteq \bigcup_{i=1}^{\infty} J_i \cap I_k \quad \forall k \in \mathbb{N}$. Since Int is closed under intersections, the sets $J_i \cap I_k$ are intervals. Therefore, since ℓ is σ-sub-additive on Int,

$$\ell(I_k) \leq \sum_{i=1}^{\infty} \ell(J_i \cap I_k) \;\; \forall k \in \mathbb{N} .$$

Sum over all $k \in \mathbb{N}$:

$$\sum_{k=1}^{\infty} \ell(I_k) \leq \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} \ell(J_i \cap I_k) = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \ell(J_i \cap I_k)$$

By assumption the I_k are pairwise disjoint, so $J_i \supseteq \coprod_{k=1}^{\infty} J_i \cap I_k$. By the earlier exercise, $\ell(J_i) \geq \sum_{k=1}^{\infty} \ell(J_i \cap I_k)$. It follows that

$$\begin{aligned} \sum_{k=1}^{\infty} \ell(I_k) &\leq \sum_{i=1}^{\infty} \ell(J_i) \\ &\leq m^*\left(\coprod_{k=1}^{\infty} I_k\right) + \varepsilon \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the proposition follows.

∎

Example. What is $m^*(\mathbb{Q})$?

(1) Direct method: enumerate $\mathbb{Q} = \{x_k\}_{k=1}^{\infty}$. Choose intervals

$$I_k = \left(x_k - \varepsilon/2^{k+1}, x_k + \varepsilon/2^{k+1}\right)$$

for a fixed $\varepsilon > 0$. $\{I_k\}_{k=1}^{\infty}$ is a cover of $\mathbb{Q}$.

$$0 \leq \sum_{k=1}^{\infty} \ell(I_k) = \sum_{k=1}^{\infty} \varepsilon/2^k = \varepsilon \underset{\varepsilon \to 0}{\longrightarrow} 0$$

(2) By σ-sub-additivity:

$$m^*(\mathbb{Q}) = m^*\left(\bigcup_{k=1}^{\infty} \{x_k\}\right) \leq \sum_{k=1}^{\infty} \ell([x_k, x_k]) = 0$$

Exercise. (1) Show that every countable set in $\mathbb{R}$ has zero outer measure.
(2) Show that a countable union of sets with outer measure zero also has outer measure zero.

Summary. We have extended $\ell : \text{Int} \to [0,\infty]$ to $m^* : 2^{\mathbb{R}} \to [0,\infty]$; m^* is defined for all subsets of $\mathbb{R}$; m^* is σ-sub-additive on $2^{\mathbb{R}}$; m^* is not σ-additive on $2^{\mathbb{R}}$ and so is not a measure.

Step 3 – Obtaining a Measure from m^* by Restricting its Domain

Definition. A set $E \subseteq \mathbb{R}$ is called *Lebesgue measurable* if it satisfies the condition

$$\forall T \subseteq \mathbb{R},\ m^*(T) = m^*(T \cap E) + m^*(T \cap E^{\mathrm{C}}).$$

This condition is called *Carathéodory's criterion*.

Remarks. (1) T is often called a "test set".
(2) Carathéodory's criterion is symmetric with respect to E, so if E is Lebesgue measurable so is E^{C}.
(3) The inequality

$$m^*(T) \leq m^*(T \cap E) + m^*(T \cap E^{\mathrm{C}})$$

is always true by the σ-sub-additivity of m^*.

Notation. $\mathcal{B}_0 = \{E \subseteq \mathbb{R} \mid E \text{ is Lebesgue measurable}\}$.

We aim to prove:
(1) $\mathcal{B}_0$ is a σ-algebra;
(2) $m^*\big|_{\mathcal{B}_0} : \mathcal{B}_0 \to [0,\infty]$ is σ-additive;
(3) $\mathcal{B}_0 \supseteq \text{Int}$.

Given these, $\left(\mathbb{R}, \mathcal{B}_0, m^*\big|_{\mathcal{B}_0}\right)$ will be a measure space.

Proposition. *Let* $\mathcal{B}_0 = \{E \subseteq \mathbb{R} \mid E \text{ is Lebesgue measurable}\}$. *Then*
(1) $\varnothing, \mathbb{R} \in \mathcal{B}_0$*;*
(2) $E \in \mathcal{B}_0 \Rightarrow E^{\mathrm{C}} \in \mathcal{B}_0$*;*

(3) $E_1, E_2 \in \mathcal{B}_0 \Rightarrow E_1 \cap E_2, E_1 \cup E_2, E_1 \setminus E_2 \in \mathcal{B}_0$.

Proof. (1) $\varnothing \in \mathcal{B}_0$ since $m^*(T) = m^*(T \cap \varnothing) + m^*(T \cap \varnothing^{\mathrm{C}}) = 0 + m^*(T)$; similarly for $\mathbb{R}$.

(2) If E is Lebesgue measurable then $\forall T \subseteq \mathbb{R}$,

$$m^*(T) = m^*(T \cap E) + m^*(T \cap E^{\mathrm{C}}) = m^*\left(T \cap \left(E^{\mathrm{C}}\right)^{\mathrm{C}}\right) + m^*(T \cap E^{\mathrm{C}})$$

(3) Let $E_1, E_2 \in \mathcal{B}_0$. Then

$$\begin{aligned} m^*(T \cap (E_1 \cup E_2)) &= m^*\left(T \cap \left(E_1 \amalg \left(E_2 \cap E_1^{\mathrm{C}}\right)\right)\right) \\ &\le m^*(T \cap E_1) + m^*\left(T \cap \left(E_2 \cap E_1^{\mathrm{C}}\right)\right) \\ m^*\left(T \cap (E_1 \cup E_2)^{\mathrm{C}}\right) &= m^*\left(T \cap \left(E_1^{\mathrm{C}} \cap E_2^{\mathrm{C}}\right)\right) \\ &= m^*\left(\left(T \cap E_1^{\mathrm{C}}\right) \cap E_2^{\mathrm{C}}\right) \\ &= m^*\left(T \cap E_1^{\mathrm{C}}\right) - m^*\left(T \cap E_1^{\mathrm{C}} \cap E_2\right) \end{aligned}$$

because if we apply Carathéodory's criterion for E_2 using the test set $T \cap E_1^{\mathrm{C}}$ we get

$$m^*\left(T \cap E_1^{\mathrm{C}}\right) = m^*\left(T \cap E_1^{\mathrm{C}} \cap E_2\right) + m^*\left(T \cap \left(E_1^{\mathrm{C}} \cap E_2^{\mathrm{C}}\right)\right)$$

Adding these estimates gives

$$m^*(T \cap (E_1 \cup E_2)) + m^*\left(T \cap (E_1 \cup E_2)^{\mathrm{C}}\right) \le m^*(T \cap E_1) + m^*\left(T \cap E_1^{\mathrm{C}}\right) = m^*(T)$$

The reverse inequality is always true (see the remark above) so $E_1 \cup E_2 \in \mathcal{B}_0$. The remainder (intersections, differences, etc.) follow from that already shown and de Morgan's laws.

■

Lemma. m^* *is finitely additive on* $\mathcal{B}_0$.

Proof. We require that given $E_1, \ldots, E_n \in \mathcal{B}_0$ pairwise disjoint, $m^*\left(\coprod_{i=1}^n E_i\right) = \sum_{i=1}^n m^*(E_i)$. Apply Carathéodory's criterion for E_n with test set $T = \coprod_{i=1}^n E_i$:

$$\begin{aligned} m^*(T) &= m^*(T \cap E_n) + m^*\left(T \cap E_n^{\mathrm{C}}\right) \\ &= m^*(E_n) + m^*\left(\coprod_{i=1}^{n-1} E_i\right) \\ &= \sum_{i=1}^n m^*(E_i) \end{aligned}$$

and the result follows by induction.

■

Exercise. Prove that if $E_1,\ldots,E_n \in \mathcal{B}_0$ are pairwise disjoint then $\forall T \subseteq \mathbb{R}$,

$$m^*\left(T \cap \coprod_{i=1}^n E_i\right) = \sum_{i=1}^n m^*\left(T \cap E_i\right).$$

Proposition. *$\mathcal{B}_0$ is a σ-algebra.*

Proof. We already know that $\mathcal{B}_0$ is an algebra so we need only check that given $\{E_i\}_{i=1}^\infty$ in $\mathcal{B}_0$ the union $\bigcup_{i=1}^\infty E_i \in \mathcal{B}_0$.

Step 1: It is possible to assume without loss of generality that this collection of sets is a collection of pairwise disjoint sets.

Proof: Let $E_i' = E_i \setminus \bigcup_{j=1}^{i-1} E_j$. These are in $\mathcal{B}_0$ because we know that $\mathcal{B}_0$ is closed under finite unions and set differences. Also, $\bigcup_{i=1}^\infty E_i = \bigcup_{i=1}^\infty E_i'$ and $E_i' \cap E_j' = \varnothing$ for $i \neq j$.

So we can assume that the E_i are pairwise disjoint.

Step 2: $\forall T \subseteq \mathbb{R}$, $m^*(T) = m^*\left(T \cap \coprod_{i=1}^\infty E_i\right) + m^*\left(T \setminus \coprod_{i=1}^\infty E_i\right)$.

Proof: $(\leq)$ follows from the σ-sub-additivity of m^*.

$(\geq)$ $\forall N \in \mathbb{N}$, $\coprod_{i=1}^N E_i \in \mathcal{B}_0$ since $\mathcal{B}_0$ is an algebra of sets. Therefore, $\forall T \subseteq \mathbb{R}$,

$$\begin{aligned} m^*(T) &= m^*\left(T \cap \coprod_{i=1}^N E_i\right) + m^*\left(T \setminus \coprod_{i=1}^N E_i\right) \\ &= \sum_{i=1}^N m^*\left(T \cap E_i\right) + m^*\left(T \setminus \coprod_{i=1}^N E_i\right) \\ &\geq \sum_{i=1}^N m^*\left(T \cap E_i\right) + m^*\left(T \setminus \coprod_{i=1}^\infty E_i\right) \end{aligned}$$

So for each $N \in \mathbb{N}$, $m^*(T) \geq \sum_{i=1}^N m^*\left(T \cap E_i\right) + m^*\left(T \setminus \coprod_{i=1}^\infty E_i\right)$. Passing to the limit as $N \to \infty$:

$$\begin{aligned} m^*(T) &\geq \sum_{i=1}^\infty m^*\left(T \cap E_i\right) + m^*\left(T \setminus \coprod_{i=1}^\infty E_i\right) \\ &\geq m^*\left(T \cap \coprod_{i=1}^\infty E_i\right) + m^*\left(T \setminus \coprod_{i=1}^\infty E_i\right) \end{aligned}$$

■

Proposition. *m^* is σ-additive on $\mathcal{B}_0$.*

Proof. We need to show that if $\{E_i\}_{i=1}^{\infty}$ in $\mathcal{B}_0$ are pairwise disjoint then

$$m^*\left(\coprod_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m^*(E_i).$$

($\le$) is the σ-sub-additivity of m^*.

($\ge$) By monotonicity, $\forall N \in \mathbb{N}$, $m^*\left(\coprod_{i=1}^{\infty} E_i\right) \ge m^*\left(\coprod_{i=1}^{N} E_i\right)$, which equals $\sum_{i=1}^{N} m^*(E_i)$ by finite additivity. Since this holds for all N, $m^*\left(\coprod_{i=1}^{\infty} E_i\right) \ge \sum_{i=1}^{\infty} m^*(E_i)$.

■

Proposition. *Intervals are Lebesgue measurable.*

Proof. Let $I \in \mathrm{Int}$. We need to show that $\forall T \subseteq \mathbb{R}$,

$$m^*(T) = m^*(T \cap I) + m^*(T \cap I^{\mathrm{C}}).$$

($\le$) is always true.

($\ge$) Fix $\varepsilon > 0$ and let $\{I_k\}_{k=1}^{\infty}$ be a cover of T by intervals I_k such that $m^*(T) \ge \sum_{k=1}^{\infty} \ell(I_k) - \varepsilon$. Then

$$m^*(T \cap I) \le m^*\left(\bigcup_{k=1}^{\infty} I_k \cap I\right) \le \sum_{k=1}^{\infty} m^*(I_k \cap I)$$
$$m^*(T \cap I^{\mathrm{C}}) \le \sum_{k=1}^{\infty} m^*(I_k \cap I^{\mathrm{C}})$$

So

$$m^*(T \cap I) + m^*(T \cap I^{\mathrm{C}}) \le \sum_{k=1}^{\infty} \left(m^*(I_k \cap I) + m^*(I_k \cap I^{\mathrm{C}})\right).$$

If we knew that $m^*(I_k \cap I) + m^*(I_k \cap I^{\mathrm{C}}) = m^*(I_k)$ then we would have

$$\mathrm{RHS} \le \sum_{k=1}^{\infty} m^*(I_k) \le m^*(T) + \varepsilon$$

Passing to the limit as $\varepsilon \to 0$ would prove ($\le$). So we need only prove that $\forall k \in \mathbb{N}$,

$$m^*(I_k \cap I) + m^*(I_k \cap I^{\mathrm{C}}) = m^*(I_k).$$

Int forms a semi-algebra and so $I_k \cap I \in \text{Int}$, $I \cap I^C = \coprod_{i=1}^n J_i$ with $J_i \in \text{Int}$. It follows that

$$\begin{aligned} m^*(I_k \cap I) + m^*(I_k \cap I^C) &= \ell(I_k \cap I) + \sum_{i=1}^n \ell(J_i) \\ &= \ell\left((I_k \cap I) \coprod \left(\coprod_{i=1}^n J_i\right)\right) \\ &= \ell(I_k) \\ &= m^*(I_k) \end{aligned}$$

■

Theorem. *There exists a* σ*-algebra* $\mathcal{B}_0$ *of subsets of* $\mathbb{R}$ *and a set function* $m : \mathcal{B}_0 \to [0,\infty]$ *such that*
(1) $\mathcal{B}_0 \supseteq \text{Int}$*;*
(2) m *is* σ*-additive;*
(3) $m(I) = \ell(I)$ *for* $I \in \text{Int}$.

Notation and Terminology. (1) $m = m^*\big|_{\mathcal{B}_0}$ is called *Lebesgue measure*.
(2) $\mathcal{B}_0$ is the σ-algebra of Lebesgue measurable subsets.
(3) $(\mathbb{R}, \mathcal{B}_0, m)$ is *Lebesgue's measure space* (on $\mathbb{R}$).

Properties of Lebesgue Measure

Proposition. (Translation-invariance) *If* $E \in \mathcal{B}_0$ *and* $x \in \mathbb{R}$ *then* $m(E) = m(x+E)$.

Proof. It is obvious that $\ell : \text{Int} \to [0,\infty]$ is translation-invariant. It follows that m^* is translation-invariant. We would like to now say that $m = m^*\big|_{\mathcal{B}_0}$ is also translation-invariant. However, we don't know that $E \in \mathcal{B}_0 \Rightarrow x + E \in \mathcal{B}_0$. Let $T \subseteq \mathbb{R}$.

$$\begin{aligned} m^*(T \cap (x+E)) + m^*(T \setminus (x+E)) &= m^*(x + ((-x+T) \cap E)) + m^*(x + ((-x+T) \cap E^C)) \\ &= m^*((-x+T) \cap E) + m^*((-x+T) \cap E^C) \\ &= m^*(-x+T) \\ &= m^*(T) \end{aligned}$$

So m is translation-invariant.

■

Definition. A measure space $(X, \mathcal{F}, \mu)$ is called σ*-finite* if $\exists F_i \in \mathcal{F}$ such that $X = \bigcup_{i=1}^\infty F_i$ and $\mu(F_i) < \infty$.

A measure space $(X, \mathcal{F}, \mu)$ is called a *probability space* if $\mu(X) = 1$.

Proposition. *$(\mathbb{R}, \mathcal{B}_0, m)$ is σ-finite.*

Proof. Take

$$F_n = \begin{cases} [k, k+1) & n = 2k \\ [-k, -k+1) & n = 2k+1 \end{cases}$$

■

Definition. Let $(X, \mathcal{F}, \mu)$ be a measure space. $A \subseteq X$ is called μ*-negligible* (or μ*-null*) if $\inf\{\mu(E) \mid A \subseteq E \in \mathcal{F}\} = 0$. $(X, \mathcal{F}, \mu)$ is called *complete* if every μ-null set is in $\mathcal{F}$.

Proposition. *$(\mathbb{R}, \mathcal{B}_0, m)$ is complete.*

Proof. Suppose A is a null set. Fix $\varepsilon > 0$ and find an $E \in \mathcal{B}_0$ such that $A \subseteq E$ and $\mu(E) < \varepsilon/2$. Since $m = m^*\big|_{\mathcal{B}_0}$ there must exist a cover of E by a countable collection of intervals $\{I_k\}_{k=1}^{\infty}$ such that

$$\sum_{k=1}^{\infty} \ell(I_k) \leq m^*(E) + \varepsilon/2 < \varepsilon .$$

Clearly $A \subseteq E$ is covered by $\bigcup_{k=1}^{\infty} I_k$ as well. So

$$m^*(A) \leq \sum_{k=1}^{\infty} \ell(I_k) < \varepsilon$$

Since $\varepsilon > 0$ was arbitrary, $m^*(A) = 0$. This implies that $A \in \mathcal{B}_0$ because

$$\begin{aligned} m^*(T) &\leq m^*(T \cap A) + m^*(T \setminus A) \\ &\leq m^*(A) + m^*(T) \\ &= m^*(T) \end{aligned}$$

and so $m^*(T) = m^*(T \cap A) = m^*(T \setminus A)$.

■

We will now work towards the following result:

Proposition. *Open and closed subsets of $\mathbb{R}$ are Lebesgue measurable.*

Lemma. (Lindelöf's Lemma) *Every open subset of* $\mathbb{R}$ *is a countable union of open intervals.*

Proof. Fix an open set $U \subseteq \mathbb{R}$. Every $x \in U$ has a neighbourhood such that $(x-\varepsilon_x, x+\varepsilon_x) \subseteq U$. Choose $x-\varepsilon_x < \alpha_x < x < \beta_x < x+\varepsilon_x$ such that $\alpha_x, \beta_x \in \mathbb{Q}$. Clearly $U = \bigcup_{x\in\mathbb{Q}} (\alpha_x, \beta_x)$. The cardinality of the cover $\{(\alpha_x, \beta_x)\}_{x\in U}$ is at most $|\mathbb{Q}\times\mathbb{Q}| = \aleph_0$.

■

Definition. The smallest σ-algebra of subsets of $\mathbb{R}$ containing all open sets is called the *Borel* σ*-algebra* and is denoted $\mathcal{B}(\mathbb{R})$. Elements of $\mathcal{B}(\mathbb{R})$ are called *Borel sets*.

Remark. Assignment 1 proves that the smallest σ-algebra containing a given collection actually exists.

Lemma. *All intervals, open sets and closed sets are Borel sets.*

Proof. Open sets are Borel because, by Lindelöf's Lemma, they are countable unions of open intervals. Closed sets are complements of open sets, and so are Borel. Intervals (a,b) and $[a,b]$ are Borel because they are open and closed respectively.

$$(a,b] = (a,c) \cup [c,b] \text{ for any } c \in (a,b)$$

so $(a,b]$ is Borel. Similarly for $[a,b)$.

■

Note. We can define the Borel σ-algebra for any topological space $(X, \mathcal{T})$: simply close $\mathcal{T}$ under complementation.

Proposition. $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}_0$*; i.e., every Borel set is Lebesgue measurable.*

Proof. Define $\mathcal{F} = \{ E \in \mathcal{B}(\mathbb{R}) \mid E \in \mathcal{B}_0 \}$. We prove that $\mathcal{F} \supseteq \mathcal{B}(\mathbb{R})$ by showing
(1) $\mathcal{F}$ contains the open sets;
(2) $\mathcal{F}$ is a σ-algebra.
The proposition then follows, since $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra contains the open sets.

(1) Suppose $U \subseteq \mathbb{R}$ is open. By Lindelöf's Lemma there are open intervals $\{I_k\}_{k=1}^{\infty}$ such that $U = \bigcup_{k=1}^{\infty} I_k$. Intervals are Lebesgue measurable and $\mathcal{B}_0$ is a σ-algebra. So $U \in \mathcal{B}_0$. $\mathcal{B}(\mathbb{R}) \supseteq \{\text{open sets}\}$ so $U \in \mathcal{B}(\mathbb{R})$ so $U \in \mathcal{F}$.

(2) $\varnothing \in \mathcal{F}$ is trivial.

$E \in \mathcal{F} \Rightarrow E^{C} \in \mathcal{F}$ since

$$\begin{aligned} E \in \mathcal{F} &\Leftrightarrow E \in \mathcal{B}(\mathbb{R}), E \in \mathcal{B}_0 \\ &\Leftrightarrow E^{C} \in \mathcal{B}(\mathbb{R}), E \in \mathcal{B}_0 \\ &\Leftrightarrow E^{C} \in \mathcal{F} \end{aligned}$$

$\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$ because

$$\begin{aligned} \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F} &\Leftrightarrow \forall i, E_i \in \mathcal{B}(\mathbb{R}), E_i \in \mathcal{B}_0 \\ &\Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}(\mathbb{R}), \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_0 \\ &\Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{F} \end{aligned}$$

■

Theorem. (Regularity Theorem) *If $E \in \mathcal{B}_0$, then $\varepsilon > 0$ there are sets $F \subseteq E \subseteq U$ such that F is closed, U is open, and $m(U \setminus F) < \varepsilon$. Moreover, if $m(E) < \infty$ then F can be chosen to be compact.*

Remark. Informally, this says that every measurable set is approximately open and approximately closed.

Proof. *Step 1: Constructing U.* Define $E_n = [n, n+1)$ for $n \in \mathbb{Z}$. By definition, $m(E_n) = m^*(E_n)$, so for every $\varepsilon_n > 0$ there is a countable cover of E_n by intervals $\langle a_k^{(n)}, b_k^{(n)} \rangle$ such that

$$\sum_{k=1}^{\infty} \ell\left(\langle a_k^{(n)}, b_k^{(n)} \rangle\right) \le m^*(E_n) + \varepsilon_n .$$

If $I_k^{(n)} = \left(a_k^{(n)} - \varepsilon_n / 2^{k+1}, b_k^{(n)} + \varepsilon_n / 2^{k+1}\right)$ then $E_n \subset \bigcup_{k=1}^{\infty} I_k^{(n)}$ and

$$\begin{aligned} m(E_n) &\ge \sum_{k=1}^{\infty} \left(\ell\left(I_k^{(n)}\right) - 2 \tfrac{\varepsilon_n}{2^{k+1}}\right) - \varepsilon_n \\ &= \sum_{k=1}^{\infty} \ell\left(I_k^{(n)}\right) - 2\varepsilon_n \end{aligned}.$$

Define $U_n = \bigcup_{k=1}^{\infty} I_k^{(n)}$ and $U = \bigcup_{n \in \mathbb{Z}} U_n$. Note that $U \supseteq E$ and that

$$\begin{aligned} m(U \setminus E) &\le \sum_{n \in \mathbb{Z}} m(U_n \setminus E) \\ &\le \sum_{n \in \mathbb{Z}} m(U_n \setminus E_n) \\ &= \sum_{n \in \mathbb{Z}} \left(m(U_n) - m(E_n)\right) \end{aligned}$$

$$\begin{aligned}&\leq \sum_{n\in\mathbb{Z}}\sum_{k=1}^{\infty}\ell\left(I_k^{(n)}\right)-m(E_n)\\&\leq \sum_{n\in\mathbb{Z}}2\varepsilon_n\\&=2\sum_{n\in\mathbb{Z}}\varepsilon_n\end{aligned}$$

Now choose $\varepsilon_n = \varepsilon/2^{|n|+10}$ to obtain

$$m(U \setminus E) \leq 2\varepsilon\left(\sum_{n=0}^{\infty}\tfrac{1}{2^{n+10}}+\sum_{n=1}^{\infty}\tfrac{1}{2^{n+10}}\right) < 2\varepsilon\left(\tfrac{1}{8}+\tfrac{1}{8}\right)=\varepsilon/2$$

This shows that $\forall E \in \mathcal{B}_0$, $\forall \varepsilon > 0$, $\exists U$ open, such that $U \supseteq E$ and $m(U \setminus E) < \varepsilon/2$.

Step 2: Constructing F. Apply Step 1 to E^{C} to obtain an open set $V \supseteq E^{\mathrm{C}}$ such that $m\left(V \setminus E^{\mathrm{C}}\right) < \varepsilon/2$. Then $F = V^{\mathrm{C}}$ satisfies F closed, $F \subseteq E$ and

$$m(E \setminus F) = m(E \cap V) = m\left(V \setminus E^{\mathrm{C}}\right) < \varepsilon/2\,.$$

Taken together, Steps 1 and 2 give $F \subseteq E \subseteq U$ such that

$$\begin{aligned}m(U \setminus F) &\leq m(U \setminus E) + m(E \setminus F)\\&< \varepsilon/2 + \varepsilon/2\\&= \varepsilon\end{aligned}$$

Step 3: Compact Lower Bound. Suppose $m(E) < \infty$. Use Steps 1 and 2 to obtain F closed such that $m(E \setminus F) < \varepsilon/2$. For each n set $F_n = F \cap [-n, n]$. This set is closed and bounded, hence compact. We also have $F_n \uparrow F$, so $E \setminus F_n \downarrow E \setminus F$. By assumption, $m(E) < \infty$, so $m(E \setminus F_n) \to m(E \setminus F)$ as $n \to \infty$, and $m(E \setminus F) < \varepsilon/2 < \infty$. It follows that we can choose n large enough so that $m(E \setminus F_n) < \varepsilon$. F_n is our suitable compact set.

■

Corollary. (Structure of Lebesgue Measurable Sets) *If* $E \in \mathcal{B}_0$ *then there is a* $B \in \mathcal{B}(\mathbb{R})$ *and a null set* N *such that* $E = B\Delta N = (B \setminus N) \cup (N \setminus B)$.

Proof. Take $U_n \supseteq E$ open such that $m(U_n \setminus E) < 1/n$ and consider $B = \bigcap_{n=1}^{\infty} U_n$. Then $E = B \setminus (B \setminus E)$, B is Borel, and $B \setminus E$ is null.

Alternatively, take $F_n \subseteq E$ closed such that $m(E \setminus F_n) < 1/n$ and consider $B = \bigcup_{n=1}^{\infty} U_n$.

■

Sets that are not Lebesgue Measurable

Theorem. (Vitali's Theorem) *There exists a subset of* $[0,1]$ *that is not Lebesgue measurable.*

Idea. Show that $[0,1]$ can be written as a countable pairwise disjoint union $[0,1] = \coprod_{k=1}^{\infty} A_k$ of sets A_k such that if A_1 is measurable then all A_k are measurable and have the same measure as A_1. This implies that A_1 is not measurable because otherwise

$$1 = m([0,1]) = m\left(\coprod_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} m(A_k) = m(A_1) + m(A_1) + \dots$$

which is impossible since if $m(A_1) > 0$ the RHS is infinite, and if $m(A_1) = 0$ then RHS is 0. We construct a suitable A_1 using the translation-invariance of $\mathcal{B}_0$ and m.

Proof. Define a binary operation on $[0,1)$ by

$$x \oplus y = \begin{cases} x + y & x + y \in [0,1) \\ x + y - 1 & x + y \geq 1 \end{cases}$$

Step 1. $([0,1), \oplus)$ is an Abelian group.

Proof. $x \oplus y$ is determined by the two conditions $x \oplus y \in [0,1)$ and $e^{2\pi i(x \oplus y)} = e^{2\pi i x} e^{2\pi i y}$.

Commutativity:

$$\begin{aligned} e^{2\pi i(x \oplus y)} &= e^{2\pi i x} e^{2\pi i y} \\ &= e^{2\pi i y} e^{2\pi i x} \\ &= e^{2\pi i(y \oplus x)} \end{aligned}$$

Associativity:

$$\begin{aligned} e^{2\pi i(x \oplus (y \oplus z))} &= e^{2\pi i x}\left(e^{2\pi i y} e^{2\pi i z}\right) \\ &= \left(e^{2\pi i x} e^{2\pi i y}\right) e^{2\pi i z} \\ &= e^{2\pi i((x \oplus y) \oplus z)} \end{aligned}$$

The neutral element is 0 and the inverse of x, $-x = 1 - x$.

Step 2. $\mathcal{B}_0$ and m are $\oplus$-invariant.

Proof. Write $A = A_1 \coprod A_2$ where $A_1 = A \cap [0, 1-x)$, $A_2 = A \cap [1-x, 1)$. If $A \in \mathcal{B}_0$ then $A_1, A_2 \in \mathcal{B}_0$ and

$$\begin{aligned} x \oplus A &= x \oplus (A_1 \amalg A_2) \\ &= (x \oplus A_1) \amalg (x \oplus A_2) \\ &= (x + A_1) \amalg ((x-1) + A_2) \end{aligned}$$

Since $\mathcal{B}_0$ is $+$-invariant, $x + A_1, (x-1) + A_2 \in \mathcal{B}_0$ and so $x \oplus A \in \mathcal{B}_0$. Since m is $+$-invariant, $m(x \oplus A) = m(A_1) + m(A_2) = m(A)$.

Step 3. Obtain a partition by defining an equivalence relation on $[0,1)$ by $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$ (it is easy to see $\sim$ is an equivalence relation). Let

$$[x] = \{ y \in [0,1) \mid x \sim y \}$$

be the equivalence class of $\sim$ containing x.

We wish to choose a representative from each class. Formally, let $\mathcal{E} = \{[x] \mid x \in [0,1)\}$ be the collection of equivalence classes. Using the Axiom of Choice find a choice function $f : \mathcal{E} \to [0,1)$ such that $f([x]) \in [x]$ (f is our law that "chooses" a representative).

Define $A = \{ f([x]) \mid x \in [0,1)\} =$ the collection of representatives, and enumerate $\mathbb{Q} \cap [0,1)$ as $\{q_k\}_{k=1}^{\infty}$. We claim that if $A_k = q_k \oplus A$ then $[0,1) = \coprod_{k=1}^{\infty} A_k$, i.e.

(1) $[0,1) = \bigcup_{k=1}^{\infty} A_k$;

(2) $i \neq j \Rightarrow A_i \cap A_j = \varnothing$.

Proof of (1). It is enough to show that $[0,1) \subseteq \bigcup_{k=1}^{\infty} A_k$. Suppose $y \in [0,1)$. y is equivalent to the representative of its class, whence $y - f([x]) = q' \in \mathbb{Q}$.

If $q' \in [0,1)$ set $q = q'$ and observe that $y = q + f([y]) = q \oplus f([y]) \in q \oplus A$.

If not then $q' = y - f([y]) \in (-1,0)$. In this case define $q = q' + 1 \in \mathbb{Q} \cap [0,1)$ and note that $y = q + f([y]) - 1 = q \oplus f([y]) \in q \oplus A$. In either case, $\exists q \in \mathbb{Q} \cap [0,1)$ such that $y \in q \oplus A$. By definition, $\exists k$ such that $q = q_k$ and $y \in q_k + A = A_k$.

Proof of (2). Suppose $y \in A_i \cap A_j$. Then $q_i \oplus f([y]) = y = q_j \oplus f([y])$. Add $1 - f([y])$ mod 1 to both sides to obtain $q_i = q_j$, i.e. $i = j$.

This argument implies that $A, A_1, A_2, \ldots$ are all non-measurable since if one of them were measurable then by Step 1 all would be measurable with equal measure. This contradicts σ-additivity because we have

$$1 = m([0,1)) = m\left(\coprod_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} m(A_k) = \sum_{k=1}^{\infty} m(A) \in \{0, \infty\}$$

■

Warning. Whenever the Axiom of Choice is used, measurability problems may arise.

Aim. To define the integral of a suitable function with respect to a measure on a measurable space.

Measurable Functions

Let $(X,\mathcal{F})$ be a measurable space.

Notation. We use the following shorthand notation for sets: if $f:X\to[-\infty,\infty]$ is a function then

$$
\begin{aligned}
[f<t]&=\{x\in X \mid f(x)<t\}\\
[f\in B]&=\{x\in X \mid f(x)\in B\}\\
[a\le f\le b]&=\{x\in X \mid a\le f(x)\le b\}
\end{aligned}
$$

etc.

Definition. Let $(X,\mathcal{F})$ be a measurable space and $f:X\to[-\infty,\infty]$ a function. We say that f is $\mathcal{F}$*-measurable* if $\forall t\in\mathbb{R}$, $[f<t]\in\mathcal{F}$.

Proposition. *The following are all equivalent:*
(1) $f:X\to\mathbb{R}$ *is* $\mathcal{F}$*-measurable;*
(2) $\forall t\in\mathbb{R}$, $[f>t]\in\mathcal{F}$*;*
(3) $\forall a,b\in\mathbb{R}$, $[a<f<b]\in\mathcal{F}$*;*
(4) $\forall B\in\mathcal{B}(\mathbb{R})$, $f^{-1}(B)\in\mathcal{F}$ *and* $f^{-1}(-\infty),f^{-1}(\infty)\in\mathcal{F}$.

Proof. (1)$\Rightarrow$(2). Fix t. Then

$$
\begin{aligned}
[f>t]&=[f\le t]^{C}\\
&=\left[\forall n,\ f<t+\tfrac{1}{n}\right]\\
&=\left(\textstyle\bigcap_{n=1}^{\infty}\left[f<t+\tfrac{1}{n}\right]\right)^{C}\\
&\in\mathcal{F}
\end{aligned}
$$

since $\mathcal{F}$ is a σ-algebra.

(2)$\Rightarrow$(3).

$$
\begin{aligned}
[a<f<b]&=[f>a]\cap[f<b]\\
&=[f>a]\cap[f\ge b]^{C}\\
&=[f>a]\cap\left[\forall n,\ f>b-\tfrac{1}{n}\right]^{C}
\end{aligned}
$$

$$= [f > a] \cap \left(\bigcap_{n=1}^{\infty} \left[f > b - \tfrac{1}{n} \right] \right)^C$$
$$\in \mathcal{F}$$

$(3) \Rightarrow (4)$. Define $\mathcal{C} = \{ B \in \mathcal{B}(\mathbb{R}) \mid f^{-1}(B) \in \mathcal{F} \}$. We show
(a) $\mathcal{C}$ contains all open sets;
(b) $\mathcal{C}$ is a σ-algebra;
(c) $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra containing the open sets, so $\mathcal{C} \supseteq \mathcal{B}(\mathbb{R})$, and we are done once we check that $f^{-1}(-\infty), f^{-1}(\infty) \in \mathcal{F}$.

(a) $(3) \Leftrightarrow f^{-1}(I) \in \mathcal{F}$ for all open intervals I. By Lindelöf, every open set $U \subseteq \mathbb{R}$ can be written as $U = \bigcup_{k=1}^{\infty} I_k$, I_k open intervals. Observe

$$\begin{aligned} f^{-1}(U) &= f^{-1}\left(\bigcup_{k=1}^{\infty} I_k \right) \\ &= \{ x \in X \mid f(x) \in I_k \text{ for some } k \} \\ &= \bigcup_{k=1}^{\infty} \{ x \in X \mid f(x) \in I_k \} \\ &= \bigcup_{k=1}^{\infty} f^{-1}(I_k) \\ &\in \mathcal{F} \end{aligned}$$

(b) Exercise.

(c) (a) and (b) imply $B \in \mathcal{B}(\mathbb{R}) \Rightarrow f^{-1}(B) \in \mathcal{F}$. We now need only check that $f^{-1}(-\infty)$ and $f^{-1}(\infty) \in \mathcal{F}$:

$$\begin{aligned} f^{-1}(\infty) &= [f = \infty] \\ &= [\forall n, f > n] \\ &= \bigcap_{n=1}^{\infty} [f > n] \\ &\in \mathcal{F} \end{aligned}$$

$$\begin{aligned} f^{-1}(-\infty) &= [f = -\infty] \\ &= [\forall n, f < -n] \\ &= \bigcap_{n=1}^{\infty} [f < -n] \\ &\in \mathcal{F} \end{aligned}$$

$(4) \Rightarrow (1)$ is trivial.

■

Proposition. *Suppose* $f_1, \ldots, f_n : X \to \mathbb{R}$ *are* $\mathcal{F}$*-measurable. If*

$$\{ (f_1(x), \ldots, f_n(x)) \in \mathbb{R}^n \mid x \in X \} \subseteq U$$

for some open set $U \subseteq \mathbb{R}^n$ *and if* $\phi : U \to \mathbb{R} : (t_1, \ldots, t_n) \mapsto \phi(t_1, \ldots, t_n)$ *is continuous then* $x \mapsto \phi(f_1(x), \ldots, f_n(x))$ *is* $\mathcal{F}$ *-measurable.*

Examples. (1) $f_1 + \ldots + f_n$ is $\mathcal{F}$ -measurable (take $\phi(t_1, \ldots, t_n) = t_1 + \ldots + t_n$).
(2) $f_1 f_2 \ldots f_n$ is $\mathcal{F}$ -measurable.
(3) $\arcsin \log \sqrt{f_1 + f_2 f_3^{1/2}} + 1$ is $\mathcal{F}$ -measurable provided

$$\{ (f_1(x), f_2(x), f_3(x)) \in \mathbb{R}^3 \mid x \in X \} \subseteq \text{domain}(\phi)$$

where $\phi(t_1, t_2, t_3) = \arcsin \log \sqrt{t_1 + t_2 t_3^{1/2}}$.

Proof. Suppose $V \subseteq \mathbb{R}^n$ is open. Then there is a countable collection of open boxes

$$C_k = \left(a_1^{(k)}, b_1^{(k)}\right) \times \ldots \times \left(a_n^{(k)}, b_n^{(k)}\right)$$

such that $V = \bigcup_{k=1}^{\infty} C_k$ (prove this). Fix $t \in \mathbb{R}$. We show that $[\phi(f_1, \ldots, f_n) > t] \in \mathcal{F}$. Since ϕ is continuous, $V = \{ (s_1, \ldots, s_n) \in \mathbb{R}^n \mid \phi(s_1, \ldots, s_n) > t \}$ is open. Find boxes $\{C_k\}_{k=1}^{\infty}$ such that $V = \bigcup_{k=1}^{\infty} C_k$. Now calculate

$$\begin{aligned} [\phi(f_1, \ldots, f_n) > t] &= \{ x \in X \mid (f_1(x), \ldots, f_n(x)) \in V \} \\ &= \{ x \in X \mid (f_1(x), \ldots, f_n(x)) \in \textstyle\bigcup_{k=1}^{\infty} C_k \} \\ &= \textstyle\bigcup_{k=1}^{\infty} \{ x \in X \mid f_i(x) \in \left(a_i^{(k)}, b_i^{(k)}\right), 1 \le i \le n \} \\ &= \textstyle\bigcup_{k=1}^{\infty} \bigcap_{i=1}^{\infty} \left[a_i^{(k)} < f_i < b_i^{(k)}\right] \\ &\in \mathcal{F} \end{aligned}$$

This shows that $\forall t \in \mathbb{R}$, $[\phi(f_1, \ldots, f_n) > t] \in \mathcal{F}$.

■

Recall. For a sequence $\{a_n\}_{n=1}^{\infty}$, a *limit point* is the limit of a subsequence $\{a_{n_k}\}_{k=1}^{\infty}$. The *limit superior* is $\limsup_{n\to\infty} a_n = \sup\{\text{limit points}\} = \max\{\text{limit points}\}$. The *limit inferior* is $\liminf_{n\to\infty} a_n = \inf\{\text{limit points}\} = \max\{\text{limit points}\}$. The sequence converges if and only if $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$. We also have the following formulae:

$$\limsup_{n\to\infty} a_n = \inf_N \sup_{n>N} a_n ,$$
$$\liminf_{n\to\infty} a_n = \sup_N \inf_{n>N} a_N .$$

Proposition. (Calculus of measurable functions.) *If* $\{f_n\}_{n=1}^{\infty}$ *are* $\mathcal{F}$*-measurable functions then the following functions are also* $\mathcal{F}$*-measurable:*
(1) $f(x) = \sup_n f_n(x)$*;* $f(x) = \inf_n f_n(x)$*; (It is essential that sup and inf are taken over a countable range of parameters.)*
(2) $f(x) = \limsup_{n\to\infty} f_n(x)$*;* $f(x) = \liminf_{n\to\infty} f_n(x)$*;*
(3) $f_c(x) = \lim_{n\to\infty} f_n(x)$ *if the limit exists,* c *otherwise;*
(4) $f(x) = \sum_{n=1}^{\infty} f_n(x)$ *if the sum converges;* $f(x) = \prod_{n=1}^{\infty} f_n(x)$ *if the product converges.*

Proof. (1) Suppose $f = \sup_n f_n$. Then $\forall t \in \mathbb{R}$,

$$\begin{aligned} [f < t] &= [\sup_n f_n < t] \\ &= [\exists k \text{ s.t. } \forall n, f_n < t - \tfrac{1}{k}] \\ &= \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} [f_n < t - \tfrac{1}{k}] \\ &\in \mathcal{F} \end{aligned}$$

since $[f_n < t - \frac{1}{k}] \in \mathcal{F}$ and $\mathcal{F}$ is a σ-algebra. Since this holds $\forall t \in \mathbb{R}$, $f = \sup_n f_n$ is $\mathcal{F}$-measurable.

In the case of $\inf_n f_n$, observe that $\inf_n f_n = -\sup_n(-f_n)$. If f_n is $\mathcal{F}$-measurable so is $-f_n$, thus the result follows from the above.

(2) Suppose that $f = \limsup_{n\to\infty} f_n$. Recall that $\limsup_{n\to\infty} a_n = \inf_N \sup_{n>N} a_n$. By (1), $\forall N \in \mathbb{N}$, $x \mapsto \sup_{n>N} f_n(x)$ is $\mathcal{F}$-measurable. Again by (1), $x \mapsto \inf_N \sup_{n>N} f_n(x)$ is $\mathcal{F}$-measurable. A similar proof shows that $\liminf_{n\to\infty} f_n$ is $\mathcal{F}$-measurable.

(3) We start by showing that $[\lim_{n\to\infty} f_n(x) \text{exists}] \in \mathcal{F}$. This is because

$$\begin{aligned} [\lim_{n\to\infty} f_n(x) \text{exists}] &= [\lim_{n\to\infty} f_n(x) \text{does not exist}]^{C} \\ &= [\limsup_{n\to\infty} f_n > \liminf_{n\to\infty} f_n]^{C} \\ &= [\exists q \in \mathbb{Q} \text{ s.t.} \limsup_{n\to\infty} f_n > q > \liminf_{n\to\infty} f_n]^{C} \end{aligned}$$

$$= X \setminus \bigcup_{q\in\mathbb{Q}} \left(\left[\limsup_{n\to\infty} f_n > q\right] \cap \left[\liminf_{n\to\infty} f_n < q\right] \right)$$

and by (2) and the fact that $\mathcal{F}$ is a σ-algebra. We now show that $[f_c < t] \in \mathcal{F} \quad \forall t \in \mathbb{R}$. Observe that

$$[f_c < t] = \begin{cases} \left([\lim \text{exists}] \cap [\limsup_{n\to\infty} f_n < t]\right) \cup [\lim \text{exists}]^C & c < t \\ [\lim \text{exists}] \cap [\limsup_{n\to\infty} f_n < t] & c \geq t \end{cases}$$

In both cases $[f_c < t] \in \mathcal{F}$.

(4) Exercise.

■

Examples of Measurable Functions

Definition. $f : E \to [-\infty, +\infty]$ is called *Lebesgue* (respectively *Borel*) *measurable* if it is measurable with respect to the Lebesgue (respectively Borel) σ-algebra.

Remarks. (1) Every Borel measurable function is Lebesgue measurable, since $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}_0$.
(2) Every continuous function $f : [a,b] \to \mathbb{R}$ is Borel (and hence Lebesgue) since $[f < t]$ is open for all t, and open sets are Borel.
(3) Highly discontinuous functions can be Borel, too, such as the Dirichlet function:

$$D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

This is Borel ($D = \mathbf{1}_{\mathbb{Q}}$) but is not continuous.

Definition. If A is a set the *indicator function* of A is

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

Proposition. *The indicator of A is Borel (respectively Lebesgue) measurable if and only if A is Borel (respectively Lebesgue).*

Proof.

$$[\mathbf{1}_A < t] = \begin{cases} \varnothing & t \le 0 \\ A^{\mathrm{C}} & 0 \le t \le 1 \\ \mathbb{R} & 1 < t \end{cases}$$

■

Definition. Let $(X,\mathcal{F})$ be a measurable space. A *simple function* is an $f : X \to \mathbb{R}$ of the form

$$f(x) = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i}(x)$$

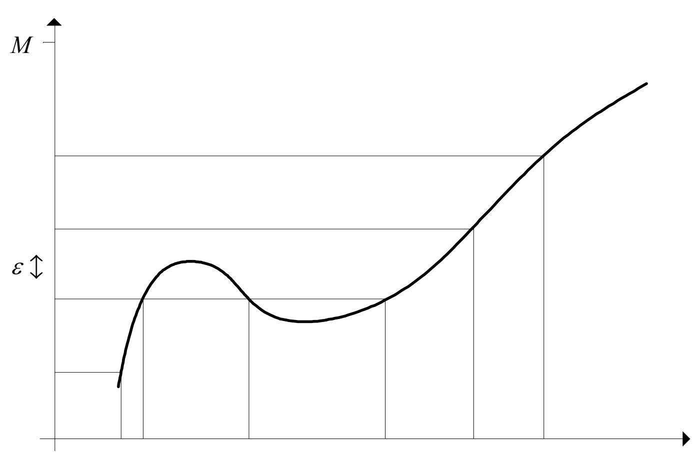
for some $n \in \mathbb{N}$, $\alpha_i \in \mathbb{R}$, $E_i \in \mathcal{F}$.

Warning. This is not the same as being a step function. A simple function is a step function if and only if all the E_i are intervals.

Theorem. (Basic Approximation) Let $(X,\mathcal{F})$ be a measurable space. Then $f : X \to \mathbb{R}$ is bounded and measurable if and only if $\forall \varepsilon > 0$ there is a simple function $\phi_\varepsilon : X \to \mathbb{R}$ such that $\forall x \in X$, $|f(x) - \phi_\varepsilon(x)| < \varepsilon$.

Proof. ($\Leftarrow$) Suppose $f : X \to \mathbb{R}$ is the uniform limit of simple functions, i.e. $\forall n \in \mathbb{N}$, $\exists \phi_n$ simple such that $\forall x \in X$, $|f(x) - \phi_n(x)| < 1/n$. Then
(1) f is measurable, because it is the limit of measurable functions.
(2) f is bounded, because $|f(x)| \le |\phi_1(x)| + 1 < \sup|\phi_1| + 1 < \infty$.

($\Rightarrow$) Suppose $f : X \to \mathbb{R}$ is bounded measurable.



(Divide the y-axis evenly by ε-strips, project on to x-axis.)

By assumption, f is bounded, so $M = \sup|f| < \infty$. Fix $\varepsilon > 0$ and choose $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$. Define

$$\phi_n(x) = \sum_{k=0}^{\lceil Mn \rceil} \frac{k}{n} \mathbf{1}_{\left[\frac{k}{n} \le f < \frac{k+1}{n}\right]}(x) + \sum_{k=-\lfloor Mn \rfloor}^{-1} \frac{k}{n} \mathbf{1}_{\left[\frac{k}{n} \le f < \frac{k+1}{n}\right]}(x).$$

This is a simple function. For all $x \in X$ there is a unique k such that $-\lfloor Mn \rfloor \le k < \lceil Mn \rceil$ and $\frac{k}{n} \le f(x) < \frac{k+1}{n}$, and for this k $f(x) \in \left[\frac{k}{n}, \frac{k+1}{n}\right]$, $\phi_n(x) = \frac{k}{n}$, and so $|f(x) - \phi_n(x)| < \frac{1}{n} < \varepsilon$.

■

Properties of Lebesgue Measurable Functions

Definition. Let $(X, \mathcal{F}, \mu)$ be a measure space and let P be a property of points $x \in X$. We say that P holds μ-*almost everywhere* (μ-*a.e.*) if $\mu(\{x \in X \mid \neg P(x)\}) = 0$

Examples. (1) $f_n \to f$ a.e. iff $\mu(\{x \in X \mid f_n(x) \not\to f(x)\}) = 0$.
(2) $f = g$ a.e. if $\mu([f \ne g]) = 0$.

Proposition. *If* $f : [a,b] \to [-\infty, \infty]$ *is Lebesgue measurable and finite* m*-a.e. then* $\forall \varepsilon, \delta > 0$ $\exists \phi \in C^1([a,b], \mathbb{R})$ *such that* $m\{x \in [a,b] \mid |f(x) - \phi(x)| > \varepsilon\} < \delta$.

Example. For the Dirichlet function $D = \mathbf{1}_{\mathbb{Q}}$ take $\phi(x) \equiv 0$ since $m[|D - \phi| > 0] = m(\mathbb{Q})$ $= 0$.

Proof. (1) First prove for indicators.
(2) Generalize to simple functions.
(3) Generalize to bounded measurable functions.
(4) Generalize to all a.e.-finite measurable functions.

Step 1 – Indicators. Suppose $f = \mathbf{1}_E$, $E \subseteq [a,b]$ Lebesgue measurable. Using the regularity of Lebesgue measure find $K \subseteq E \subseteq U$, K compact, U open, such that $m(U \setminus K) < \delta$ for a given fixed $\delta > 0$. By Urysohn's Lemma there exists a continuous function ϕ such that $\phi|_K \equiv 1$, $\phi|_{U^C} \equiv 0$.

If $x \in K$ then $|\mathbf{1}_E(x) - \phi(x)| = |1 - 1| = 0$; if $x \in U^C$ then $|\mathbf{1}_E(x) - \phi(x)| = |0 - 0| = 0$. Therefore,

$$m[|\mathbf{1}_E - \phi| > \varepsilon] \le m[\mathbf{1}_E \ne \phi] \le m(U \setminus K) < \delta.$$

Step 2 – Simple Functions. Let $f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i}$. Find using Step 1 continuous functions ϕ_i such that $m[|\mathbf{1}_{E_i} - \phi_i| > \varepsilon_i] < \delta_i$, where $\varepsilon_i, \delta_i > 0$ will be determined later.

$$\left[\left|\sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i} - \sum_{i=1}^{n} \alpha_i \phi_i\right| > \varepsilon\right] \subseteq \bigcup_{i=1}^{n} \left[|\mathbf{1}_{E_i} - \phi_i| > \frac{\varepsilon}{n|\alpha_i|}\right] \qquad (*)$$

since if $x \in \text{LHS}$ then $|\alpha_i \mathbf{1}_{E_i}(x) - \alpha_i \phi_i(x)| > \frac{\varepsilon}{n}$ for at least one i, for otherwise we would have

$$\left|\sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i}(x) - \sum_{i=1}^{n} \alpha_i \phi_i(x)\right| \le \sum_{i=1}^{n} |\alpha_i| |\mathbf{1}_{E_i}(x) - \phi_i(x)| < \sum_{i=1}^{n} |\alpha_i| \frac{\varepsilon}{n|\alpha_i|} = \varepsilon .$$

By $(*)$,

$$m\left[\left|\sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i} - \sum_{i=1}^{n} \alpha_i \phi_i\right| > \varepsilon\right] \le \sum_{i=1}^{n} m\left[|\mathbf{1}_{E_i} - \phi_i| > \frac{\varepsilon}{n|\alpha_i|}\right]$$

Therefore, if we choose $\varepsilon_i = \frac{\varepsilon}{n|\alpha_i|}$, $\delta_i = \frac{\delta}{n}$, then

$$m\left[\left|\sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i} - \sum_{i=1}^{n} \alpha_i \phi_i\right| > \varepsilon\right] < \delta .$$

Step 3 – Bounded Measurable Functions. Suppose f is bounded measurable. By the Basic Approximation Theorem there is a simple function ϕ_1 such that $|f(x) - \phi_1(x)| < \varepsilon/2$ for all x. Using Step 2 find a continuous ϕ such that $m[|\phi_1 - \phi| > \varepsilon/2] < \delta$. Now note that

$$[|f - \phi| > \varepsilon] \subseteq [|f - \phi_1| > \varepsilon/2] \cup [|\phi_1 - \phi| > \varepsilon/2]$$

and so

$$m[|f - \phi| > \varepsilon] \le \underbrace{m[|f - \phi_1| > \varepsilon/2]}_{=0} + \underbrace{m[|\phi_1 - \phi| > \varepsilon/2]}_{<\delta}.$$

Step 4 – General Case. Suppose $f : [a,b] \to [-\infty, \infty]$ is finite a.e. and measurable. Define

$$f_M(x) = \begin{cases} f(x) & |f(x)| \le M \\ 0 & \text{otherwise} \end{cases}$$

Note that $[f \neq f_M] = [|f| > M] \downarrow [|f| = \infty]$. Since $[f \neq f_1] \subseteq [a,b]$ has finite measure, $m[f \neq f_M] \underset{M\to\infty}{\to} m[|f| = \infty] = 0$. Now choose M so large that $m[f \neq f_M] < \delta/2$ and use Step 3 to obtain a continuous ϕ such that $m[|f_M - \phi| > \varepsilon] < \delta/2$. The inclusion

$$[|f - \phi| > \varepsilon] \subseteq [|f_M - \phi| > \varepsilon] \cup [f \neq f_M]$$

implies that ϕ provides a suitable approximation.

■

Theorem. (Egoroff's Theorem) *Suppose* $f_n : E \to \mathbb{R}$ *are Lebesgue measurable and that* $E \in \mathcal{B}_0$ *has finite measure. Assume that* $f_n \underset{n\to\infty}{\to} f$ *a.e. in* E. *Then* $\forall \varepsilon > 0 \;\; \exists K \subseteq E$ *compact such that* $m(E \setminus K) < \varepsilon$ *and* $f_n \underset{n\to\infty}{\to} f$ *uniformly on* K.

Proof. Define $G_{n,k} = \left\{x \in E \mid |f_k(x) - f(x)| > \frac{1}{n}\right\}$. Set

$$G_N^{(n)} = \bigcup_{k=N+1}^{n} G_{n,k} \equiv \left\{x \in E \mid \exists k > N \text{ s.t.} |f_k(x) - f(x)| > \tfrac{1}{n}\right\}.$$

Observe that $\left\{G_N^{(n)}\right\}_{N=1}^{\infty}$ is a decreasing sequence of sets and that

$$\bigcap_{N=1}^{\infty} G_N^{(n)} \subseteq [f_k \not\to f].$$

But $f_k \to f$ a.e. so $m\left(G_N^{(n)}\right) \underset{N\to\infty}{\to} m\left(\bigcap_{N=1}^{\infty} G_N^{(n)}\right) = 0$. Therefore, $\forall n \in \mathbb{N}$, $\exists N_n \in \mathbb{N}$ such that $m\left(G_{N_n}^{(n)}\right) < \varepsilon/2^{n+1}$. Now consider $E' = E \setminus \bigcup_{n=1}^{\infty} G_{N_n}^{(n)}$. Then

$$m(E \setminus E') \leq m\left(\bigcup_{n=1}^{\infty} G_{N_n}^{(n)}\right) < \sum_{n=1}^{\infty} \tfrac{\varepsilon}{2^{n+1}} = \tfrac{\varepsilon}{2}.$$

If $x \in E'$ then $x \notin G_{N_n}^{(n)}$ and so

$$k > N_n \Rightarrow |f_k(x) - f(x)| \leq \tfrac{1}{n}$$

and so $f_n \to f$ uniformly on E'. Now apply the regularity of m to obtain a compact $K \subseteq E'$ with $m(E' \setminus K) < \varepsilon/2$. Check that K still satisfies the requirements.

■

Theorem. (Lusin's Theorem) *If* $f : E \to \mathbb{R}$ *is Lebesgue measurable and* $E \in \mathcal{B}_0$ *has finite measure then* $\forall \varepsilon > 0$ *there is a compact* $K \subseteq E$ *such that* $f|_K : K \to \mathbb{R}$ *is uniformly continuous and* $m(E \setminus K) < \varepsilon$.

Proof. Exercise. ∎

Exercise. What happens when $f = D = \mathbf{1}_{\mathbb{Q}}$?

The Lebesgue Integral

What is $\int_0^1 f(t)dt$?

The regulated integral:

(1) step functions $\phi = \sum_{i=1}^n \alpha_i \mathbf{1}_{I_i}$, $\alpha_i \in \mathbb{R}$, $I_i \in \text{Int}$:

$$\int \phi = \sum_{i=1}^n \alpha_i \ell(I_i).$$

(2) regulated functions $f =$ uniform limit of step functions ϕ_n :

$$\int f = \lim_{n\to\infty} \int \phi_n .$$

The Lebesgue integral:

(1) simple functions $\phi = \sum_{i=1}^n \alpha_i \mathbf{1}_{E_i}$, $\alpha_i \in \mathbb{R}$, $E_i \in \mathcal{B}_0$:

$$\int \phi = \sum_{i=1}^n \alpha_i m(E_i).$$

(2) bounded measurable functions $f =$ uniform limit of simple functions ϕ_n :

$$\int f = \lim_{n\to\infty} \int \phi_n .$$

To make these ideas rigorous we start by integrating simple functions:

$$f = \sum_{i=1}^n \alpha_i \mathbf{1}_{E_i} \ , \ E_i \in \mathcal{B}_0$$

Natural definition:

$$\int_E f = \sum_{i=1}^n \alpha_i m(E_i \cap E)$$

Problem. There are many representations of f of the form $\sum_{i=1}^{n} \alpha_i \mathbf{1}_{E_i}$, for instance

$$\mathbf{1}_{[0,2]} + \mathbf{1}_{[1,3]} = \mathbf{1}_{[0,1)} + 2.\mathbf{1}_{[1,2]} + \mathbf{1}_{(2,3]}$$

Solution. Choose a canonical representation such as

$$f(x) = \sum_{y \in f(\mathbb{R})} y \mathbf{1}_{[f=y]}(x)$$

Example.

$$\operatorname{sgn}(x) = \begin{cases} x/|x| & x \neq 0 \\ 0 & x = 0 \end{cases}$$
$$\operatorname{sgn}(\mathbb{R}) = \{-1, 0, 1\}$$
$$[\operatorname{sgn} = -1] = (-\infty, 0)$$
$$[\operatorname{sgn} = 0] = \{0\}$$
$$[\operatorname{sgn} = 1] = (0, \infty)$$

Hence $\operatorname{sgn}(x) = -1.\mathbf{1}_{(-\infty,0)}(x) + 0.\mathbf{1}_{\{0\}}(x) + 1.\mathbf{1}_{(0,\infty)}(x)$.

Definition. If f is a (Lebesgue measurable) simple function and $E \in \mathcal{B}_0$ has finite measure the *integral of* f *over* E *with respect to* m is

$$\int_E f \, dm = \sum_{y \in f(\mathbb{R})} y.m(E \cap [f = y]).$$

Remark. We require $m(E) < \infty$ to prevent the following problem:

$$\begin{aligned} \int_{\mathbb{R}} \operatorname{sgn}(x) dm(x) &= -1.m((-\infty, 0)) + 0.m(\{0\}) + 1.m((0, \infty)) \\ &= \infty - \infty \\ &= \text{undefined} \end{aligned}$$

Proposition. *If f, g are simple functions, $\alpha, \beta \in \mathbb{R}$, $E \in \mathcal{B}_0$, $m(E) < \infty$, then*
(1) linearity: $\int_E \alpha f + \beta g = \alpha \int_E f + \beta \int_E g$ *;*
(2) monotonicity: $f \geq g \Rightarrow \int_E f \geq \int_E g$ *;*
(3) absolute value property: $\left| \int_E f \right| \leq \int_E |f|$ *;*
(4) $f = g$ *a.e.* $\Rightarrow \int_E f = \int_E g$ *;*

(5) $E = E_1 \coprod E_2,\ E_i \in \mathcal{B}_0 \Rightarrow \int_E f = \int_{E_1} f + \int_{E_2} f$.

Proof. (1) We start by showing $\int_E f + g = \int_E f + \int_E g$. By definition

$$\begin{aligned}
\int_E f + g &= \sum_{y \in (f+g)(\mathbb{R})} y.m(E \cap [f + g = y]) \\
&= \sum_{y \in f(\mathbb{R}) + g(\mathbb{R})} y.m(E \cap [f + g = y]) \\
&= \sum_{\substack{F \in f(\mathbb{R}),\, G \in g(\mathbb{R}) \\ F+G=y}} y.m(E \cap [f + g = y]) \\
&= \sum_{y \in f(\mathbb{R}) + g(\mathbb{R})} y \left(\sum_{\substack{F \in f(\mathbb{R}),\, G \in g(\mathbb{R}) \\ F+G=y}} m(E \cap [f = F, g = G]) \right) \\
&\left(\because [f + g = y] = \coprod_{\substack{F \in f(\mathbb{R}),\, G \in g(\mathbb{R}) \\ F+G=y}} [f = F, g = G] \right) \\
&= \sum_{F \in f(\mathbb{R}),\, G \in g(\mathbb{R})} (F + G) m(E \cap [f = F] \cap [g = G]) \\
&= \sum_{F \in f(\mathbb{R})} F \sum_{G \in g(\mathbb{R})} m(E \cap [f = F] \cap [g = G]) + \sum_{G \in g(\mathbb{R})} G \sum_{F \in f(\mathbb{R})} m(E \cap [f = F] \cap [g = G]) \\
&= \sum_{F \in f(\mathbb{R})} F.m(E \cap [f = F]) + \sum_{G \in g(\mathbb{R})} G.m(E \cap [g = G]) \\
&= \int_E f + \int_E g
\end{aligned}$$

Now show that $\int_E cf = c\int_E f$.

$$(cf)(\mathbb{R}) = \{ cy \mid y \in f(\mathbb{R}) \}$$

$$\Rightarrow \qquad \int_E cf = \sum_{y \in f(\mathbb{R})} cy.m(E \cap [f = y]) = c\int_E f$$

(2) If $f \geq g$, $f - g \geq 0$, so all values of $f - g$ are non-negative, so $\int_E f - g \geq 0$. Together with (1), this implies that $\int_E f - \int_E g \geq 0$, so $\int_E f \geq \int_E g$.

(3) $-|f| \leq f \leq |f|$, so $-\int_E |f| \leq \int_E f \leq \int_E |f|$, so $\left| \int_E f \right| \leq \int_E |f|$.

(4), (5). Exercises.

■

We now integrate bounded measurable functions.

Proposition. *Let* $f : E \to \mathbb{R}$ *be bounded measurable,* $E \in \mathcal{B}_0$, $m(E) < \infty$. *There exist simple functions* $\phi_n : E \to \mathbb{R}$ *such that* $\phi_n \underset{n\to\infty}{\to} f$ *uniformly on* E *and*

(1) $\{\int_E \phi_n\}$ *always converges;*

(2) its limit is independent of the choice of $\{\phi_n\}$.

Definition. The *integral of* f *over* E *with respect to* m is $\int_E f\, dm = \lim_{n\to\infty} \int_E \phi_n\, dm$.

Proof. The existence of $\phi_n \to f$ is the content of the Basic Approximation Theorem.

(1) By assumption, $\phi_n \underset{n\to\infty}{\to} f$ uniformly on E so $\varepsilon_n = \sup_E |\phi_n - f| \underset{n\to\infty}{\to} 0$. We check that the sequence $\{\int_E \phi_n\}$ is Cauchy:

$$\left|\int_E \phi_m - \int_E \phi_n\right| \le \int_E |\phi_m - \phi_n| \le m(E) \sup_E |\phi_m - \phi_n|.$$

But $\forall x \in E$,

$$|\phi_m(x) - \phi_n(x)| \le |\phi_m(x) - f(x)| + |f(x) - \phi_n(x)| \le \varepsilon_m + \varepsilon_n.$$

So

$$\left|\int_E \phi_m - \int_E \phi_n\right| \le m(E)(\varepsilon_m + \varepsilon_n) \underset{n\to\infty}{\to} 0.$$

(2) Exercise.

■

Proposition. (Properties of the integral.) *If* f, g *are bounded measurable functions,* $\alpha, \beta \in \mathbb{R}$, $E \in \mathcal{B}_0$, $m(E) < \infty$, *then*

(1) linearity: $\int_E \alpha f + \beta g = \alpha \int_E f + \beta \int_E g$;

(2) monotonicity: $f \ge g \Rightarrow \int_E f \ge \int_E g$;

(3) absolute value property: $\left|\int_E f\right| \le \int_E |f|$;

(4) $\forall G \in \mathcal{B}_0$, $\int_E f.\mathbf{1}_G = \int_{E\cap G} f$;

(5) $f = g$ *a.e.* $\Rightarrow \int_E f = \int_E g$;

(6) $E = E_1 \coprod E_2$, $E_i \in \mathcal{B}_0 \Rightarrow \int_E f = \int_{E_1} f + \int_{E_2} f$.

Proof. Let $\{\phi_n\}, \{\psi_n\}$ be sequences of simple functions tending to f, g uniformly on E as $n \to \infty$.

(1) Clearly $\alpha\phi_n + \beta\psi_n \underset{n\to\infty}{\to} \alpha f + \beta g$ uniformly on E since

$$|(\alpha\phi_n(x) + \beta\psi_n(x)) - (\alpha f(x) + \beta g(x))| \leq |\alpha||\phi_n(x) - f(x)| + |\beta||\psi_n(x) - g(x)| .$$

Thus

$$\begin{aligned} \int_E \alpha f + \beta g &= \int_E \lim_{n\to\infty}(\alpha\phi_n + \beta\psi_n) \\ &= \lim_{n\to\infty} \int_E \alpha\phi_n + \beta\psi_n \\ &= \alpha \lim_{n\to\infty} \int_E \phi_n + \beta \lim_{n\to\infty} \int_E \psi_n \\ &= \alpha \int_E f + \beta \int_E g \end{aligned}$$

(2) Suppose $f \geq g$. Fix $\varepsilon > 0$. By uniform convergence $\exists N$ such that $n > N \Rightarrow$ $\sup_E |f - \phi_n|, \sup_E |g - \psi_n| < \varepsilon$. In particular, for $n > N$,

$$\phi_n > f - \varepsilon \geq g - \varepsilon > \psi_n - 2\varepsilon .$$

By Monotonicity for simple functions,

$$\int_E \phi_n > \int_E (\psi_n - 2\varepsilon) = \int_E \psi_n = 2\varepsilon m(E)$$

$$\Rightarrow \qquad \int_E f = \lim_{n\to\infty} \int_E \phi_n \geq \lim_{n\to\infty} \int_E \psi_n - 2\varepsilon m(E) = \int_E g - 2\varepsilon m(E)$$

Pass to the limit as $\varepsilon \downarrow 0$; $\int_E f \geq \int_E g$.

(3) Corollary of (2).

(4) First observe that this property holds for simple functions, since

$$\begin{aligned}\int_E \left(\sum_{i=1}^n \alpha_i \mathbf{1}_{E_i}\right)\mathbf{1}_G &= \int_E \sum_{i=1}^n \alpha_i \mathbf{1}_{E_i \cap G} \\ &= \sum_{i=1}^n \alpha_i m(E \cap E_i \cap G) \\ &= \int_{E\cap G} \sum_{i=1}^n \alpha_i \mathbf{1}_{E_i}\end{aligned}$$

Since $\phi_n \to f$ uniformly on E, $\phi_n \to f$ uniformly on $E \cap G$. So

$$\begin{aligned}\int_{E\cap G} f &= \lim_{n\to\infty} \int_{E\cap G} \phi_n \\ &= \lim_{n\to\infty} \int_E \phi_n \mathbf{1}_G \\ &= \int_E f.\mathbf{1}_G\end{aligned}$$

(5) Let $M = \sup_E f \vee \sup_E g$.[†] By (1),

$$\int_E f = \int_E f.\mathbf{1}_{[f=g]} + \int_E f.\mathbf{1}_{[f\neq g]}.$$

But $\int_E f.\mathbf{1}_{[f\neq g]} = 0$ since $\left|\int_E f.\mathbf{1}_{[f\neq g]}\right| \le M.m(E \cap [f \neq g]) \le M.m[f \neq g] = 0$. So

$$\begin{aligned}\int_E f &= \int_E f.\mathbf{1}_{[f=g]} \\ &= \int_{E\cap[f=g]} f \\ &= \int_{E\cap[f=g]} g \\ &= \int_{E\cap[f=g]} g + \int_{E\cap[f\neq g]} g \\ &= \int_E \left(g.\mathbf{1}_{[f=g]} + g.\mathbf{1}_{[f\neq g]}\right) \\ &= \int_E g\end{aligned}$$

(6) Follows from (1) and (4).

■

Theorem. (Bounded Convergence Theorem) *Suppose* $\{f_n\}_{n=1}^\infty$ *are uniformly bounded measurable functions on* $E \in \mathcal{B}_0$ *of finite measure, i.e.* $\exists M > 0$ *such that* $\forall x \in E, n \in \mathbb{N}$, $|f_n(x)| \le M$. *If* $\forall x \in E$, $f_n(x) \underset{n\to\infty}{\to} f(x)$, *then* $\int_E f_n \underset{n\to\infty}{\to} \int_E f$.

[†] Shorthand notation: $(a \vee b)(x) = \max\{a(x), b(x)\}$; $(a \wedge b)(x) = \min\{a(x), b(x)\}$.

Proof. First note that $\int_E f$ is well-defined because
(1) f is measurable, since it is the limit of measurable functions;
(2) f is bounded (by M).

By Egoroff's Theorem, $\forall \varepsilon > 0$, $\exists E_\varepsilon \in \mathcal{B}_0$, $E_\varepsilon \subseteq E$, such that $m(E \setminus E_\varepsilon) < \varepsilon$ and $\sup_{E_\varepsilon} |f_n - f| \underset{n\to\infty}{\to} 0$. It follows that

$$\begin{aligned}\left|\int_E f_n - \int_E f\right| &\le \int_E |f_n - f| \\ &= \int_{E_\varepsilon} |f_n - f| + \int_{E\setminus E_\varepsilon} |f_n - f| \\ &\le \sup_{E_\varepsilon} |f_n - f| . m(E_\varepsilon) + \left(\sup_E |f| + \sup_E |f_n|\right) m(E \setminus E_\varepsilon) \\ &\le \underbrace{\sup_{E_\varepsilon} |f_n - f| . m(E_\varepsilon)}_{\to 0 \text{ as } n\to\infty} + 2M\varepsilon\end{aligned}$$

So $\limsup_{n\to\infty} \left|\int_E f_n - \int_E f\right| \le 2M\varepsilon \underset{\varepsilon\downarrow 0}{\to} 0$. Hence $\left|\int_E f_n - \int_E f\right| \underset{n\to\infty}{\to} 0$.

■

Exercise. Suppose f is continuous on $[a,b]$. Show that the Lebesgue and regulated integrals of f agree, i.e.

$$\int_{[a,b]} f\, dm = \int_a^b f(x)dx .$$

We now proceed to integrate unbounded functions, and begin with the non-negative case to prevent $\infty - \infty$ problems such as with $\int_{\mathbb{R}} \operatorname{sgn} dm$.

Definition. Suppose $f : E \to [0,\infty]$ is measurable and $E \in \mathcal{B}_0$ (not necessarily of finite measure). Then the *integral of f over E with respect to m* is

$$\int_E f\, dm = \sup\left\{ \int_{[h\neq 0]} h\, dm \;\middle|\; h \text{ bdd. meas.}, m[h \neq 0] < \infty, 0 \le h \le f.\mathbf{1}_E \right\}.$$

Proposition. (Properties of the integral.) *If $f, g : E \to [0,\infty]$, $\alpha, \beta > 0$, $E \in \mathcal{B}_0$, then*
(1) linearity: $\int_E \alpha f + \beta g = \alpha \int_E f + \beta \int_E g$ *;*
(2) monotonicity: $f \ge g \Rightarrow \int_E f \ge \int_E g$ *;*

(3) $\forall G \in \mathcal{B}_0$, $\int_E f.\mathbf{1}_G = \int_{E\cap G} f$;

(4) $f = g$ *a.e.* $\Rightarrow \int_E f = \int_E g$.

Proof. All trivial apart from $\int_E f + g = \int_E f + \int_E g$.

($\geq$) Take $0 \leq F \leq f.\mathbf{1}_E$, $0 \leq G \leq g.\mathbf{1}_E$ bounded measurable such that $m[F \neq 0]$, $m[G \neq 0] < \infty$ and

$$\int_{[F\neq 0]} F \geq \int_E f - \tfrac{\varepsilon}{2},$$
$$\int_{[G\neq 0]} G \geq \int_E g - \tfrac{\varepsilon}{2}.$$

The function $F + G$ is bounded measurable, satisfies $0 \leq F + G \leq (f + g).\mathbf{1}_E$ and $m[F + G \neq 0] = m([F \neq 0] \cup [G \neq 0]) < \infty$. Therefore,

$$\begin{aligned}\int_E f + g &\geq \int_{[F+G\neq 0]} F + G \\ &= \int_{[F+G\neq 0]} F + \int_{[F+G\neq 0]} G \\ &= \int_{[F\neq 0]} F + \int_{[G\neq 0]} G \\ &\geq \int_E f + \int_E g - \varepsilon\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $\int_E f + g \geq \int_E f + \int_E g$.

($\leq$) We show that for every bounded measurable function $0 \leq h \leq (f + g).\mathbf{1}_E$ such that $m[h \neq 0] < \infty$, $\int_{[h\neq 0]} h \leq \int_E f + \int_E g$. We then pass to the supremum over all such h to deduce ($\leq$).

The idea is to "split" h into two lower bounds, one for f and one for g . Define $F = h \wedge f$, $G = h - F$. Note that

$$h = 0 \Rightarrow F = 0, G = 0 ;$$
$$0 < h \leq f \Rightarrow F = h \leq f, G = 0 ;$$
$$f < h \leq f + g \Rightarrow F = f, G = h - f \leq g .$$

Therefore, $0 \leq F \leq f$ and $0 \leq G \leq g$, and since $F, G \leq h \leq f.\mathbf{1}_E$ we also have $0 \leq F \leq f.\mathbf{1}_E$, $0 \leq G \leq g.\mathbf{1}_E$. It is clear that F, G are bounded measurable, that $[F \neq 0]$, $[G \neq 0] \subseteq [h \neq 0]$ and that $F + G = h$. So

$$\begin{aligned}\int_{[h\neq 0]} h &= \int_{[h\neq 0]} F+G \\ &= \int_{[h\neq 0]} F + \int_{[h\neq 0]} G \\ &\leq \int_E f + \int_E g\end{aligned}$$

So $\int_E f+g \leq \int_E f + \int_E g$.

■

Theorem. (Fatou's Lemma) *Suppose* $\{f_n\}_{n=1}^{\infty}$ *are non-negative measurable functions on* $E \in \mathcal{B}_0$. *Then*

$$\int_E \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_E f_n .$$

Proof. It is enough to show that for every bounded measurable $0 \leq h \leq \liminf_{n\to\infty} f_n . \mathbf{1}_E$ such that $m[h \neq 0] < \infty$, $\int_E h \leq \liminf_{n\to\infty} \int_E f_n$ because we can take the supremum over all such h.

Define $h_n = h \wedge f_n$. For every $\varepsilon > 0$ $\exists N = N(x)$ such that $n > N(x) \Rightarrow f_n(x) > h(x) - \varepsilon$ because $h(x) \leq \liminf_{n\to\infty} f_n(x)$ $(*)$. Therefore, $h_n(x) \leq \liminf_{n\to\infty} f_n(x)$. But the h_n are bounded measurable and $m[h \neq 0] < \infty$ so $\int_E h_n \leq \int_E \liminf_{n\to\infty} f_n$.

Note that $h_n = h \wedge f_n \underset{n\to\infty}{\to} h$ by $(*)$. Since $\sup h_n \leq \sup h$ the Bounded Convergence Theorem tells us $\int_E h_n \underset{n\to\infty}{\to} \int_E h$. But $h_n \leq f_n$ by construction. So

$$\int_E h = \lim_{n\to\infty} \int_E h_n \leq \liminf_{n\to\infty} \int_E f_n .$$

■

Lemma. *If* $f : \mathbb{R} \to [0,\infty]$ *is measurable and* $\int_E f < \infty$ *then* f *is a.e. finite in* E.

Proof. Obviously, $\forall n \in \mathbb{N}$, $n.\mathbf{1}_{[f=\infty]} \leq f$. By monotonicity, $n.m(E \cap [f=\infty]) \leq \int_E f$ and so $m(E \cap [f=\infty]) \leq \frac{1}{n}\int_E f \underset{n\to\infty}{\to} 0$. So $m[f=\infty] = 0$.

■

Corollary. *Suppose* $f_n : E \to [0,\infty]$ *are measurable. If* $\liminf_{n\to\infty} \int_E f_n < \infty$ *then* $\liminf_{n\to\infty} f_n(x) < \infty$ *a.e. in* E.

Remark. The following example may aid in remembering Fatou's Lemma: consider $f_n = \mathbf{1}_{[n,n+1)}$. $\liminf_{n\to\infty} f_n \equiv 0$, so $\int_{\mathbb{R}} \liminf_{n\to\infty} f_n = 0$; $\int_{\mathbb{R}} f_n \equiv 1$, so $\liminf_{n\to\infty} \int_{\mathbb{R}} f_n = 1$.

Theorem. (Monotone Convergence Theorem) *(1) If* $f_n \geq 0$ *are measurable and* $f_n(x) \leq f_{n+1}(x)$ *for all* $n \in \mathbb{N}$ *and all* $x \in E \in \mathcal{B}_0$ *then*

$$\int_E \lim_{n\to\infty} f_n(x)\,dm(x) = \lim_{n\to\infty} \int_E f_n(x)\,dm(x).$$

(2) If $u_n \geq 0$ *are measurable then*

$$\int_E \sum_{n=1}^{\infty} u_n(x)\,dm(x) = \sum_{n=1}^{\infty} \int_E u_n(x)\,dm(x).$$

(3) If $E = \coprod_{n=1}^{\infty} E_n$, $E_n \in \mathcal{B}_0$, *then for all* $f \geq 0$ *measurable,*

$$\int_E f = \sum_{n=1}^{\infty} \int_{E_n} f.$$

Proof. (1) Define $f(x) = \lim_{n\to\infty} f_n(x) \in [0,\infty]$. This is well-defined because $\{f_n(x)\}_{n=1}^{\infty}$ is monotone. This f is a non-negative measurable function and so $\int_E f$ is well-defined. Since $f_n \leq f_{n+1}$, $\int_E f_n \leq \int_E f_{n+1}$ and so $\{\int_E f_n\}_{n=1}^{\infty}$ is monotone. It follows that $\lim_{n\to\infty} \int_E f_n$ exists.

($\leq$) Since $f_n \leq f$, $\int_E f_n \leq \int_E f$ and so $\lim_{n\to\infty} \int_E f_n \leq \int_E f$.

($\geq$) By Fatou's Lemma, $\int_E f = \int_E \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_E f_n$ and so $\lim_{n\to\infty} \int_E f_n \geq \int_E f$.

(2) Follows from (1) with $f_n = u_1 + \ldots + u_n$.

(3) Follows from (2) with $u_n = f.\mathbf{1}_{E_n}$.

■

We now proceed to the general case of integrating multi-signed unbounded functions. Recall that the problem here is one of avoiding $\infty - \infty$. For example,

$$\int_{-\infty}^{\infty} \operatorname{sgn} = \int_{-\infty}^{\infty} 1 + \int_{-\infty}^{\infty} (-1) = \infty - \infty .$$

Note that we cannot solve this by simply taking limits of domains of integration because

$$\int_{-\infty}^{\infty} \operatorname{sgn} = \lim_{n\to\infty} \int_{-n}^{n} \operatorname{sgn} = 0 \neq 1 = \lim_{n\to\infty} \int_{-n}^{n+1} \operatorname{sgn} .$$

Definition. Let f be a real-valued function. The *positive part* of f is $f^{+} = f \vee 0 = f.\mathbf{1}_{[f \geq 0]}$. The *negative part* of f is $f^{-} = -f \wedge 0 = -f.\mathbf{1}_{[f<0]}$. (Note that both f^{+} and f^{-} are positive.)

Exercise. Verify that $f = f^{+} - f^{-}$, $|f| = f^{+} + f^{-}$.

Definition. Let $f : E \to [-\infty, \infty]$ be some function, $E \in \mathcal{B}_0$. f is called *absolutely integrable on* E if it is measurable and $\int_E f^{+}, \int_E f^{-} < \infty$. f is called *one-sided integrable on* E if at least one of $\int_E f^{+}, \int_E f^{-}$ is finite. In each of these cases define the *integral of* f *over* E to be

$$\int_E f = \int_E f^{+} - \int_E f^{-} .$$

Exercise. Prove that a measurable function f is absolutely integrable on E if and only if $\int_E |f| < \infty$.

Proposition. (Properties of the integral.) *If* f, g *are absolutely integrable on* $E \in \mathcal{B}_0$, $\alpha, \beta \in \mathbb{R}$, *then*
(1) linearity: $\alpha f + \beta g$ *is absolutely integrable on* E *and* $\int_E \alpha f + \beta g = \alpha \int_E f + \beta \int_E g$;
(2) monotonicity: $f \geq g \Rightarrow \int_E f \geq \int_E g$;
(3) $\forall G \in \mathcal{B}_0$, $\int_E f.\mathbf{1}_G = \int_{E \cap G} f$;
(4) $f = g$ *a.e.* $\Rightarrow \int_E f = \int_E g$;
(5) $E = E_1 \coprod E_2$, $E_i \in \mathcal{B}_0 \Rightarrow \int_E f = \int_{E_1} f + \int_{E_2} f$.

Proof. As usual, everything is trivial except for $\int_E (f+g) = \int_E f + \int_E g$. First note that $f+g$ is absolutely integrable since $\int_E |f+g| \leq \int_E |f| + |g| = \int_E |f| + \int_E |g| < \infty$. By definition,

$$\int_E (f+g) = \int_E (f+g)^+ - \int_E (f+g)^-$$

$$\begin{aligned}\int_E (f+g)^+ &= \int_E (f+g).\mathbf{1}_{[(f^+-f^-)+(g^+-g^-)>0]} \\ &= \int_{E\cap[f^++g^+>f^-+g^-]} \left(f^+ + g^+ - f^- - g^-\right) \\ &= \int_{E\cap[f^++g^+>f^-+g^-]} \left(\left(f^+ + g^+\right) - \left(f^- + g^-\right)\right)\end{aligned}$$

Note that if $F, G, F-G > 0$ then $\int (F-G) + \int G = \int (F-G+G) = \int F$. Therefore,

$$\int_E (f+g)^+ = \int_{E\cap[f^++g^+>f^-+g^-]} \left(f^+ + g^+\right) - \int_{E\cap[f^++g^+>f^-+g^-]} \left(f^- + g^-\right).$$

In the same way we obtain

$$\int_E (f+g)^- = \int_{E\cap[f^++g^+<f^-+g^-]} \left(f^- + g^-\right) - \int_{E\cap[f^++g^+<f^-+g^-]} \left(f^+ + g^+\right).$$

Thus

$$\begin{aligned}\int_E (f+g) &= \int_E (f+g)^+ - \int_E (f+g)^- \\ &= \int_E \left(f^+ + g^+\right) - \int_E \left(f^- + g^-\right) \\ &= \int_E f^+ - \int_E f^- + \int_E g^+ - \int_E g^- \\ &= \int_E f + \int_E g\end{aligned}$$

■

Theorem. (Lebesgue's Dominated Convergence Theorem) *Let* $\{f_n\}_{n=1}^\infty$ *be a sequence of measurable functions such that* $f_n(x) \underset{n\to\infty}{\to} f(x)$ *for all* $x \in E \in \mathcal{B}_0$. *If* $|f_n(x)| \le g(x)$, *where* g *is absolutely integrable on* E, *then* $\int_E \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_E f_n$.

Proof. First note that $f(x) = \lim_{n\to\infty} f_n(x)$ is absolutely integrable since
(1) it is a limit of measurable functions, and so is measurable;
(2) $\int_E |f| \le \int_E |g| < \infty$ since $|f| = \lim_{n\to\infty} |f_n| \le |g|$.

We show that $\int_E f_n \underset{n\to\infty}{\to} \int_E f$ by proving

$$\int_E f \le \liminf_{n\to\infty} \int_E f_n \le \limsup_{n\to\infty} \int_E f_n \le \int_E f_n .$$

Since $|f_n| \le g$ we have $-g \le f_n \le g$. Therefore, $g + f_n \ge 0$ and by Fatou's Lemma,

$$\begin{aligned}\int_E g + \int_E f &= \int_E (g+f)\\ &= \int_E \liminf_{n\to\infty}(g+f_n)\\ &\le \liminf_{n\to\infty}\int_E (g+f_n)\\ &\le \int_E g + \liminf_{n\to\infty}\int_E f_n\end{aligned}$$

Since $\int_E |g| < \infty$, $\int_E f \le \liminf_{n\to\infty} \int_E f_n$. The functions $g - f_n$ are also non-negative.

$$\begin{aligned}\int_E g - \int_E f &= \int_E (g-f)\\ &= \int_E \liminf_{n\to\infty}(g-f_n)\\ &\le \liminf_{n\to\infty}\left(\int_E g - \int_E f_n\right)\\ &= \int_E g - \limsup_{n\to\infty}\int_E f_n\end{aligned}$$

So $\int_E f \ge \limsup_{n\to\infty} \int_E f_n$. Hence, $\int_E f = \lim_{n\to\infty} \int_E f_n$.

■

Exercise. Show that the Dominated Convergence Theorem implies both the Monotone Convergence Theorem and the Bounded Convergence Theorem.

Summary. (1) If f is absolutely integrable then $\int_E f = \int_E f^+ - \int_E f^-$, so calculating $\int_E f$ can be done by considering non-negative functions.

(2) Given a non-negative f we can find simple functions $0 \le h_n \le f$ such that $h_n(x) \uparrow f(x)$. Simply take

$$h_n(x) = \sum_{k=0}^{n^2} \tfrac{k}{n} \mathbf{1}_{\left[\frac{k}{n} \le f < \frac{k+1}{n}\right]}(x).$$

By the Dominated Convergence Theorem, $\int_E f = \lim_{n\to\infty} \int_E h_n$.

Remark. Although we worked in $(\mathbb{R}, \mathcal{B}_0, m)$ everything we did works almost verbatim for a general σ-finite measure space $(X, \mathcal{F}, \mu)$.

3. PRODUCT MEASURES

Question. How can we measure the size of subsets of $\mathbb{R}^n$? More generally, given two σ-finite measure spaces $(X,\mathcal{C},\mu)$ and $(Y,\mathcal{D},\nu)$ we seek a construction of a measure space on $X\times Y$.

Approach. (1) Define the measure of "basic sets" of the form $C\times D$, $C\in\mathcal{C}$, $D\in\mathcal{D}$ as $\lambda(C\times D)=\mu(C)\nu(D)$.
(2) Use the Carathéodory Extension Theorem to obtain a full measure out of this. (See Assignment 2.)

Definition. A *measurable rectangle* is a set of the form $C\times D$ where $C\in\mathcal{C}$, $D\in\mathcal{D}$. Denote by $\mathcal{R}$ the collection of all measurable rectangles.

Proposition. *$\mathcal{R}$ is a semi-algebra.*

Proof. (1) $\varnothing, X\times Y\in\mathcal{R}$ obvious.
(2) Intersections: $(C\times D)\cap(C'\times D')=(C\cap C')\times(D\cap D')$.
(3) Complements: $(X\times Y)\setminus(C\times D)=((X\setminus C)\times D)\coprod(C\times(Y\setminus D))$.

■

Proposition. *The algebra generated by $\mathcal{R}$ is*

$$\mathcal{A}(\mathcal{R})=\left\{\coprod_{i=1}^{n} C_i\times D_i \;\middle|\; C_i\in\mathcal{C}, D_i\in\mathcal{D}, n\in\mathbb{N}\right\}.$$

Proof. See Assignment 2.

■

Definition. $\lambda:\mathcal{A}(\mathcal{R})\to[0,+\infty]$ is defined as

$$\lambda\left(\coprod_{i=1}^{n} C_i\times D_i\right)=\sum_{i=1}^{n}\mu(C_i)\nu(D_i)$$

with the convention that $0\infty=\infty 0=0$.

Proposition. *This λ is*
(1) properly defined, i.e.

$$\coprod_{i=1}^{n} C_i\times D_i=\coprod_{i=1}^{n'} C_i'\times D_i' \Rightarrow \sum_{i=1}^{n}\mu(C_i)\nu(D_i)=\sum_{i=1}^{n'}\mu(C_i')\nu(D_i')$$

(2) σ-additive on $\mathcal{A}(\mathcal{R})$
(3) σ-finite if μ,ν are σ-finite.

Proof. (1) Suppose $\coprod_{i=1}^{n} C_i \times D_i = \coprod_{i=1}^{n'} C_i' \times D'$. Then

$$\sum_{i=1}^{n} \mathbf{1}_{C_i}(x)\mathbf{1}_{D_i}(y) \equiv \sum_{i=1}^{n'} \mathbf{1}_{C_i'}(x)\mathbf{1}_{D_i'}(y) \equiv \mathbf{1}_{\coprod_{i=1}^{n} C_i \times D_i}(x,y)$$

Fix y and integrate over all x w.r.t. μ. By linearity,

$$\sum_{i=1}^{n} \mu(C_i)\mathbf{1}_{D_i}(y) = \sum_{i=1}^{n'} \mu(C_i')\mathbf{1}_{D_i'}(y).$$

Integrate over y w.r.t. ν:

$$\sum_{i=1}^{n} \mu(C_i)\nu(D_i) = \sum_{i=1}^{n'} \mu(C_i')\nu(D_i')$$

(2) Suppose $\coprod_{i=1}^{n} C_i \times D_i = \coprod_{k=1}^{\infty} \coprod_{j=1}^{\infty} C_j^{(k)} \times D_j^{(k)}$. Again

$$\sum_{i=1}^{n} \mathbf{1}_{C_i}(x)\mathbf{1}_{D_i}(y) \equiv \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \mathbf{1}_{C_j^{(k)}}(x)\mathbf{1}_{D_j^{(k)}}(y).$$

Integrating over x w.r.t. μ we have

$$\begin{aligned}
\sum_{i=1}^{n} \mu(C_i)\mathbf{1}_{D_i}(y) &= \int_X \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \mathbf{1}_{C_j^{(k)}}(x)\mathbf{1}_{D_j^{(k)}}(y)\,d\mu(x) \\
&= \sum_{k=1}^{\infty} \int_X \sum_{j=1}^{\infty} \mathbf{1}_{C_j^{(k)}}(x)\mathbf{1}_{D_j^{(k)}}(y)\,d\mu(x) \\
&= \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \mu\left(C_j^{(k)}\right)\mathbf{1}_{D_j^{(k)}}(y)
\end{aligned}$$

Now integrate over y w.r.t ν. By the Monotone Convergence Theorem:

$$\sum_{i=1}^{n} \mu(C_i)\nu(D_i) = \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} \mu\left(C_j^{(k)}\right)\nu\left(D_j^{(k)}\right)$$

$$\text{LHS} = \lambda \coprod_{i=1}^{n} C_i \times D_i$$

$$\text{RHS} = \sum_{k=1}^{\infty} \lambda \coprod_{j=1}^{n_k} C_j^{(k)} \times D_j^{(k)}$$

$\Rightarrow$ σ-additivity on $\mathcal{A}(\mathcal{R})$.

(3) $(X,\mathcal{C},\mu)$ σ-finite $\Rightarrow$ $\exists X_i \in \mathcal{C}$ s.t. $X = \coprod_{i=1}^{\infty} X_i$ and $\mu(X_i) < \infty$. $(Y,\mathcal{D},\nu)$ σ-finite $\Rightarrow$ $\exists Y_j \in \mathcal{D}$ s.t. $Y = \coprod_{j=1}^{\infty} Y_j$ and $\mu(Y_j) < \infty$. Then $X \times Y = \coprod_{i,j=1}^{\infty} X_i \times Y_j$ where $X_i \times Y_j \in \mathcal{R}$ and $\lambda(X_i \times Y_j) = \mu(X_i)\nu(Y_j) < \infty$.

■

Definition. The *product* σ*-algebra* $\mathcal{C} \otimes \mathcal{D}$ is the minimal σ-algebra containing the collection $\{C \times D \mid C \in \mathcal{C}, D \in \mathcal{D}\}$.

Theorem. *Let* $(X,\mathcal{C},\mu)$, $(Y,\mathcal{D},\nu)$ *be two* σ*-finite measure spaces. There exists a unique measure, called the* product measure, $\mu \times \nu$ *on* $(X \times Y, \mathcal{C} \otimes \mathcal{D})$ *such that*

$$(\mu \times \nu)(C \times D) = \mu(C)\nu(D)$$

for all $C \in \mathcal{C}$ *and* $D \in \mathcal{D}$*, and this measure is* σ*-finite.*

Proof. Apply the Carathéodory Extension Theorem to $\lambda : \mathcal{A}(\mathcal{R}) \to [0,+\infty]$.

■

We would like to define the Lebesgue measure on $\mathbb{R}^2$ as $m \times m$ on $\mathcal{B}_0 \otimes \mathcal{B}_0$ but we have a technical problem:

The Completion Problem. Recall that a measure space $(X,\mathcal{F},\mu)$ is *complete* if every null set is measurable, a null set being an $A \subset X$ such that $\exists E \in \mathcal{F}$ such that $A \subset E$ and $\mu(E) = 0$.

The Lebesgue measure on $\mathbb{R}$ is complete, but $m \times m$ on $\mathcal{B}_0 \otimes \mathcal{B}_0$ is not complete: take $A \subset [0,1]$ not Lebesgue-measurable and consider $A_0 = A \times \{1\}$. This is a null set because $A_0 \subset [0,1] \times \{1\}$, which has measure zero, yet $A \notin \mathcal{B}_0 \otimes \mathcal{B}_0$.

Solution. The completion procedure:

Let $(\Omega,\mathcal{F},\mu)$ be a measure space. Define $\mu^* : 2^{\Omega} \to [0,+\infty]$ by

$$\mu^*(A) = \inf\{\, \mu(E) \mid A \subseteq E \in \mathcal{F} \,\}.$$

Definition. The *completion* of $(\Omega,\mathcal{F},\mu)$ is $(\Omega,\mathcal{F}_0,\mu_0)$, where
(1) $\mathcal{F}_0 = \{E_0 \subseteq \Omega \mid \exists E \in \mathcal{F} \text{ s.t. } \mu^*(E \Delta E_0) = 0\}$,
(2) $\mu_0 = \mu^*\big|_{\mathcal{F}_0}$.

Exercise. Prove that $(\Omega, \mathcal{F}_0, \mu_0)$ is a complete measure space.

Exercise. Show that $(\mathbb{R}, \mathcal{B}_0, m)$ is the completion of $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$. (Hint: Show that $\forall E_0 \in \mathcal{B}_0 \ \exists G = \bigcap_{n=1}^{\infty} U_n$ such that U_n open, $E_0 \subseteq G$, $m(G \setminus E_0) = 0$.)

Definition. (Lebesgue's Measure on $\mathbb{R}^n$). This is the measure space

$$(\mathbb{R}^n, (\underbrace{\mathcal{B}_0 \otimes \ldots \otimes \mathcal{B}_0}_{n})_0, (\underbrace{m \times \ldots \times m}_{n})_0)$$

i.e. the completion of the product space $\prod_{i=1}^{n} (\mathbb{R}, \mathcal{B}_0, m)$.

Definition. Given a product space $(X \times Y, \mathcal{C} \otimes \mathcal{D}, \mu \times \nu)$, $E \subseteq X \times Y$, $x \in X$, $y \in Y$,
(1) the x-*section* of E is $E_x = \{ y \in Y \mid (x, y) \in E \}$;
(2) the y-*section* of E is $E^y = \{ x \in X \mid (x, y) \in E \}$.

Exercise. Verify
(1) $(A \cup B)_x = A_x \cup B_x$;
(2) $(A \cap B)_x = A_x \cap B_x$;
(3) $(A^C)_x = (A_x)^C$;
(4) $(A \setminus B)_x = A_x \setminus B_x$;
(5) $\left(\bigcup_{\lambda \in \Lambda} A_\lambda\right)_x = \bigcup_{\lambda \in \Lambda} (A_\lambda)_x$
etc. and similarly for y-sections.

Our aim is to prove that

$$E \in \mathcal{C} \otimes \mathcal{D} \Rightarrow \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y)$$

Two problems:
(1) Is $E_x \in \mathcal{D}$? Is $E^y \in \mathcal{C}$?
(2) Is $x \mapsto \nu(E_x)$ $\mathcal{C}$-measurable? Is $y \mapsto \mu(E^y)$ $\mathcal{D}$-measurable?

Proposition. *If* $E \in \mathcal{C} \otimes \mathcal{D}$ *then for all* $x \in X$ *and* $y \in Y$, $E_x \in \mathcal{D}$ *and* $E^y \in \mathcal{C}$.

Proof. We only prove the result for x-sections: the proof for y-sections is analogous.

Recall that $\mathcal{C} \otimes \mathcal{D}$ is the minimal σ-algebra containing the collection $\mathcal{R} = \{ C \times D \mid C \in \mathcal{C}, D \in \mathcal{D} \}$. It is, therefore, enough to show that

$$\mathcal{F} = \left\{ E \in \mathcal{C} \otimes \mathcal{D} \mid \forall x \in X, E_x \in \mathcal{D} \right\}$$

is a σ-algebra containing $\mathcal{R}$.

Step 1: $\mathcal{F} \supseteq \mathcal{R}$.

Fix $C \times D \in \mathcal{R}$. Then

$$(C \times D)_x = \begin{cases} D & x \in C \\ \varnothing & x \notin C \end{cases}$$

and $\varnothing, D \in \mathcal{D}$.

Step 2: $\mathcal{F}$ is a σ-algebra.

(a) $\varnothing \in \mathcal{F}$ is trivial.
(b) $E \in \mathcal{F} \Rightarrow E^{\mathrm{C}} \in \mathcal{F}$ since $(E^{\mathrm{C}})_x = (E_x)^{\mathrm{C}} = Y \setminus E_x \in \mathcal{D}$.
(c) $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$ since $\left(\bigcup_{i=1}^{\infty} E_i\right)_x = \bigcup_{i=1}^{\infty} (E_i)_x \in \mathcal{D}$.

■

Remark. It is not true that $E \in (\mathcal{C} \otimes \mathcal{D})_0 \Rightarrow \forall x \in X, E_x \in \mathcal{D}$.

Example. Take $A \subset (0,1)$ non-measurable. Then $A \times \{1\} \in (\mathcal{B}_0 \otimes \mathcal{B}_0)_0$ since $(m \times m)(A \times \{1\}) = 0$, but $(A \times \{1\})^1 = A \notin \mathcal{B}_0$.

Proposition. *Let $(X, \mathcal{C}, \mu)$, $(Y, \mathcal{D}, \nu)$ be two σ-finite measure spaces. If $E \in \mathcal{C} \otimes \mathcal{D}$ then*
(1) $x \mapsto \nu(E_x)$ is $\mathcal{C}$-measurable and $\int_X \nu(E_x)\,d\mu(x) = (\mu \times \nu)(E)$;
(2) $y \mapsto \mu(E^y)$ is $\mathcal{D}$-measurable and $\int_Y \mu(E^y)\,d\nu(y) = (\mu \times \nu)(E)$.

Proof. Again, we prove (1) only.

Assume first that $\mu(X), \nu(Y) < \infty$. Define $\mathcal{F} = \left\{ E \in \mathcal{C} \otimes \mathcal{D} \mid (1) \text{ holds} \right\}$. We prove $\mathcal{F} \supseteq \mathcal{R}$ and $\mathcal{F}$ is a σ-algebra.

Step 1: $\mathcal{F} \supseteq \mathcal{R}$.

Fix $C \times D \in \mathcal{R}$. Then

$$(C \times D)_x = \begin{cases} D & x \in C \\ \varnothing & x \notin C \end{cases}$$

Hence $\nu\left((C\times D)_x\right)=\nu(D)\mathbf{1}_C(x)$. Obviously $x\mapsto \nu\left((C\times D)_x\right)$ is measurable and

$$\begin{aligned}\int_X \nu\left((C\times D)_x\right)d\mu(x) &= \nu(D)\int_X \mathbf{1}_C(x)\,d\mu(x)\\ &= \mu(C)\nu(D)\\ &= (\mu\times\nu)(C\times D)\end{aligned}$$

So (1) holds for $C\times D$.

Step 2: $\mathcal{F}$ is a σ-algebra.

Suppose $\{E_i\}_{i=1}^{\infty}\subseteq\mathcal{F}$. We show that $\nu((\bigcup_i E_i)_x)=\nu(\bigcup_i (E_i)_x)$ is measurable in x. If the union were disjoint we could use σ-additivity.

First attempt: Write $\bigcup_{i=1}^{\infty} E_i=\coprod_{i=1}^{\infty}\left(E_i\setminus\bigcup_{j=1}^{i-1}E_j\right)$ – but we can't be sure that $E_i\setminus\bigcup_{j=1}^{i-1}E_j\in\mathcal{F}$ because we still don't know that $\mathcal{F}$ is an algebra!

Second attempt: We show that $\mathcal{F}$ is a monotone class.

(a) $\varnothing\in\mathcal{F}$: trivial.

(b) $E_n\in\mathcal{F}, E_n\uparrow E\Rightarrow E\in\mathcal{F}$: if $E_n\in\mathcal{F}$ and $E_n\uparrow E$ then $(E_n)_x\uparrow E_x$. By the continuity of measures, $\nu\left((E_n)_x\right)\uparrow\nu(E_x)$. It follows that $x\mapsto\nu(E_x)$ is measurable because it is the limit of measurable functions.

By the Monotone Convergence Theorem,

$$\begin{aligned}\int\nu(E_x) &= \int\lim_{n\to\infty}\nu\left((E_n)_x\right)\\ &= \lim_{n\to\infty}\int\nu\left((E_n)_x\right)\\ &= \lim_{n\to\infty}(\mu\times\nu)(E_n)\\ &= (\mu\times\nu)(E)\end{aligned}$$

Thus $E\in\mathcal{F}$.

(c) $E_n\in\mathcal{F}, E_n\downarrow E\Rightarrow E\in\mathcal{F}$: Again $(E_n)_x\downarrow E_x$, so by the continuity of measures and since $\nu\left((E_n)_x\right)\le\nu(Y)<\infty$, $\nu\left((E_n)_x\right)\underset{n\to\infty}{\to}\nu(E_x)$. It follows that $x\mapsto\nu(E_x)$ is measurable.

Note $\nu\left((E_n)_x\right)\le\nu(E_x)$ and $\int_X\nu(Y)<\infty$. By the Dominated Convergence Theorem,

$$\begin{aligned}\int_X \nu(E_x)\,d\mu(x) &= \int_X \lim_{n\to\infty} \nu((E_n)_x)\,d\mu(x)\\ &= \lim_{n\to\infty}\int_X \nu((E_n)_x)\,d\mu(x)\\ &= \lim_{n\to\infty}(\mu\times\nu)(E_n)\\ &= (\mu\times\nu)(E)\end{aligned}$$

by the continuity of $\mu\times\nu$ and $(\mu\times\nu)(E_1)\le(\mu\times\nu)(X\times Y)=\mu(X)\nu(Y)<\infty$.

This shows that $\mathcal{F}$ is a monotone class. By the Monotone Class Theorem $\mathcal{F}$ contains a σ-algebra containing $\mathcal{R}$, so $\mathcal{F}\supseteq\mathcal{C}\otimes\mathcal{D}$.

We now treat the σ-finite case: assume $(X,\mathcal{C},\mu)$, $(Y,\mathcal{D},\nu)$ σ-finite. $\exists X_i\in\mathcal{C}$ s.t. $X=\coprod_{i=1}^{\infty}X_i$ and $\mu(X_i)<\infty$; $\exists Y_j\in\mathcal{D}$ s.t. $Y=\coprod_{j=1}^{\infty}Y_j$ and $\mu(Y_j)<\infty$. Define

$$\begin{aligned}\mathcal{C}_i &= \{E\cap X_i \mid E\in\mathcal{C}\}\\ \mu_i &= \mu|_{\mathcal{C}_i}\\ \mathcal{D}_j &= \{E\cap Y_j \mid E\in\mathcal{D}\}\\ \nu_j &= \nu|_{\mathcal{D}_j}\end{aligned}$$

Exercise. Prove
(1) $(X_i,\mathcal{C}_i,\mu_i)$, $(Y_j,\mathcal{D}_j,\nu_j)$ are finite measure spaces;
(2) $\mathcal{C}_i\otimes\mathcal{D}_j=\{E\cap(X_i\times Y_j)\mid E\in\mathcal{C}\otimes\mathcal{D}\}$;
(3) $\forall E\in\mathcal{C}\otimes\mathcal{D}$, $(\mu\times\nu)(E)=\sum_{i,j}(\mu_i\times\nu_j)(E\cap(X_i\times Y_j))$.

Now consider some set $E\in\mathcal{C}\otimes\mathcal{D}$.

$$\begin{aligned}\nu(E_x) &= \sum_{j=1}^{\infty}\nu(E_x\cap Y_j)\\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\mathbf{1}_{X_i}(x)\nu\left(E_x\cap(X_i\times Y_j)_x\right)\\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\mathbf{1}_{X_i}(x)\underbrace{\nu_i\left(E_x\cap(X_i\times Y_j)_x\right)}_{\mathcal{C}_i\text{-measurable by }\sigma\text{-finite case}}\end{aligned}$$

Since $\mathcal{C}_i\subseteq\mathcal{C}$ every $\mathcal{C}_i$-measurable function is $\mathcal{C}$-measurable. Therefore, since

$$\nu(E_x) = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} \mathbf{1}_{X_i}(x)\nu_i\left(\left(E \cap (X_i \times Y_j)\right)_x\right), \tag{$*$}$$

$x \mapsto \nu(E_x)$ is $\mathcal{C}$-measurable. Now integrate $(*)$ over $x \in X$. By the Monotone Convergence Theorem,

$$\begin{aligned}\int_X \nu(E_x)\,d\mu(x) &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\int_{X_i} \nu_i\left(\left(E \cap (X_i \times Y_j)\right)_x\right)d\mu(x) \\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}(\mu \times \nu)\left(E \cap (X_i \times Y_j)\right) \\ &= (\mu \times \nu)(E)\end{aligned}$$

The proof for y-sections is analogous.

■

Fubini's Theorem: We want to show that

$$\int_{X \times Y} f(x,y)\,d(\mu \times \nu)(x,y) = \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

Problems:
(1) Is $f(x,\cdot)$ ν-integrable for all x?
(2) Is $x \mapsto \int_Y f(x,y)\,d\nu(y)$ μ-integrable?

Theorem. (Fubini's Theorem) *Suppose $(X,\mathcal{C},\mu)$, $(Y,\mathcal{D},\nu)$ are σ-finite measure spaces and that $f: X \times Y \to [-\infty,+\infty]$ is $\mathcal{C} \otimes \mathcal{D}$-measurable with $\int_{X \times Y}|f|\,d(\mu \times \nu) < \infty$. Then*
(1) $\forall x \in X$, $y \mapsto f(x,y)$ is $\mathcal{D}$-measurable; $\forall y \in Y$, $x \mapsto f(x,y)$ is $\mathcal{C}$-measurable;
(2) for μ-a.e. $x \in X$, $y \mapsto f(x,y)$ is ν-absolutely integrable; for ν-a.e. $y \in Y$, $x \mapsto f(x,y)$ is μ-absolutely integrable;
(3) $x \mapsto \int_Y f(x,y)\,d\nu(y)$ is $\mathcal{C}$-measurable and μ-absolutely integrable;
$y \mapsto \int_X f(x,y)\,d\mu(x)$ is $\mathcal{D}$-measurable and ν-absolutely integrable;
(4)

$$\int_Y \int_X f(x,y)\,d\mu(x)\,d\nu(y) = \int_{X \times Y} f(x,y)\,d(\mu \times \nu)(x,y) = \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

Proof. We prove this first for indicator functions, then simple functions, then non-negative functions, then absolutely integrable functions.

Indicators: Suppose $f(x,y)=\mathbf{1}_E(x,y)$, $E\in\mathcal{C}\otimes\mathcal{D}$. $f(x,\cdot)=\mathbf{1}_{E_x(\cdot)}$. So, by the previous proposition, $f(x,\cdot)$ is $\mathcal{D}$-measurable since $E_x\in\mathcal{D}$. [(1) done.] The same proposition says that

$$\int_Y f(x,\cdot)d\nu(\cdot)=\nu(E_x)$$

is $\mathcal{C}$-measurable. [(2) done.] Again, the previous proposition gives

$$\begin{aligned}\int_X\int_Y f(x,\cdot)d\nu(\cdot)d\mu(x)&=\int_X \nu(E_x)d\mu(x)\\&=(\mu\times\nu)(E)\\&=\int_{X\times Y} f\,d(\mu\times\nu)\\&<\infty\end{aligned}$$

[(3), (4) done.] The proof is the same for y-sections.

Simple Functions: Follows from the previous case (indicators) by linearity.

Non-negative Measurable Functions: Suppose $f(x,y)\geq 0$ is $\mathcal{C}\otimes\mathcal{D}$-measurable. There exist simple functions $0\leq h_n(x,y)\leq h_{n+1}(x,y)$ such that $h_n(x,y)\uparrow f(x,y)$. For instance, take

$$h_n(x,y)=\sum_{k=1}^{n^2}\tfrac{k}{n}\mathbf{1}_{\left[\frac{k}{n}\leq f<\frac{k+1}{n}\right]}(x,y).$$

$f(x,\cdot)=\lim_{n\to\infty}h_n(x,\cdot)$ so $f(x,\cdot)$ is $\mathcal{D}$-measurable since the h_n are $\mathcal{D}$-measurable. [(1) done.] Since $h_n\uparrow f$, by the Monotone Convergence Theorem,

$$\begin{aligned}\int_Y f(x,\cdot)d\nu(\cdot)&=\int_Y \lim_{n\to\infty}h_n(x,\cdot)d\nu(\cdot)\\&=\lim_{n\to\infty}\int_Y h_n(x,\cdot)d\nu(\cdot)\end{aligned}$$

Since $h_n(x,y)$ simple, previous case shows $x\mapsto\int_Y h_n(x,\cdot)d\nu(\cdot)$ is $\mathcal{C}$-measurable, so $x\mapsto\int_Y f(x,\cdot)d\nu(\cdot)$ is $\mathcal{C}$-measurable. [(2) done.] By the Monotone Convergence Theorem, $\int_Y h_n(x,\cdot)d\nu(\cdot)\uparrow\int_Y f(x,\cdot)d\nu(\cdot)$. Again by the Monotone Convergence Theorem,

$$\begin{aligned}\int_X \int_Y f \, d\nu \, d\mu &= \int_X \lim_{n\to\infty} \int_Y h_n \, d\nu \, d\mu \\ &= \lim_{n\to\infty} \int_X \int_Y h_n \, d\nu \, d\mu \\ &= \lim_{n\to\infty} \int_{X\times Y} h_n \, d(\mu\times\nu) \\ &= \int_{X\times Y} \lim_{n\to\infty} h_n \, d(\mu\times\nu) \\ &= \int_{X\times Y} f \, d(\mu\times\nu)\end{aligned}$$

[(3), (4) done.] Similarly for the other order of integration.

General Case: Suppose f is $(\mu\times\nu)$-absolutely integrable. Decompose $f = f^+ - f^-$ and apply the previous case to f^+, f^-.

■

Remark. Remember this trick of approximating a non-negative function by a monotone sequence of simple functions.

Theorem. (Tonelli's Theorem) *Suppose $(X,\mathcal{C},\mu)$, $(Y,\mathcal{D},\nu)$ are σ-finite measure spaces and that $f : X\times Y \to [-\infty,+\infty]$ is $\mathcal{C}\otimes\mathcal{D}$-measurable. If $\int_X \int_Y |f| \, d\nu \, d\mu < \infty$ then $\int_{X\times Y} |f| \, d(\mu\times\nu) < \infty$.*

Proof. Since $(X,\mathcal{C},\mu)$, $(Y,\mathcal{D},\nu)$ are σ-finite so is $(X\times Y, \mathcal{C}\otimes\mathcal{D}, \mu\times\nu)$. So $\exists F_n \in \mathcal{C}\otimes\mathcal{D}$ with finite measure such that $F_n \uparrow X\times Y$. Now define $\phi_n = \mathbf{1}_{F_n} \cdot (|f| \wedge n)$. Check that $\int_{X\times Y} |\phi_n| \, d(\mu\times\nu) < \infty$ and that $\phi_n(x,y) \uparrow |f(x,y)|$. By the Monotone Convergence Theorem,

$$\begin{aligned}\int_{X\times Y} |f| \, d(\mu\times\nu) &= \int_{X\times Y} \lim_{n\to\infty} \phi_n \, d(\mu\times\nu) \\ &= \lim_{n\to\infty} \int_{X\times Y} \phi_n \, d(\mu\times\nu) \\ &= \lim_{n\to\infty} \int_X \int_Y \phi_n \, d\nu \, d\mu \\ &\le \int_X \int_Y |f| \, d\nu \, d\mu \\ &< \infty\end{aligned}$$

So f is $(\mu\times\nu)$-absolutely integrable.

■

Corollary. (The Fubini-Tonelli Theorem) *If one of the following integrals exists then all three exist and are equal:*

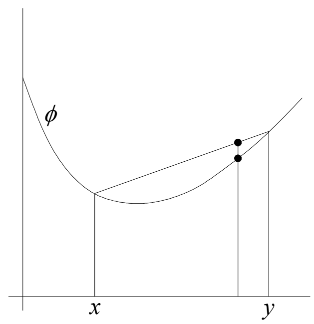
$$\int_Y \int_X f \, d\mu \, d\nu \, , \; \int_{X \times Y} f \, d(\mu \times \nu), \; \int_X \int_Y f \, d\nu \, d\mu \, .$$

4. L^p SPACES

Convexity Inequalities

Definition. A function $\phi : [a,b] \to \mathbb{R}$ is called *convex* if $\forall x, y \in [a,b]$, $\forall t \in [0,1]$,

$$\phi(tx + (1-t)y) \le t\phi(x) + (1-t)\phi(y).$$



Proposition. *ϕ is convex in $[a,b]$ iff for all $a \le x < y \le z < w \le b$,*

$$\frac{\phi(x)-\phi(y)}{x-y} \le \frac{\phi(z)-\phi(w)}{z-w}.$$

Proof. Enough to prove inequalities in the case $y = z$. Assume $a < x < z < w < b$. Solve for t in

$$\begin{aligned} z &= tx + (1-t)w \\ z &= t(x-w) + w \\ t &= \tfrac{z-w}{x-w} \\ 1-t &= \tfrac{x-w-z+w}{x-w} = \tfrac{x-z}{x-w} \end{aligned}$$

Observe,

$$\begin{aligned}
& \phi(tx + (1-t)w) \le t\phi(x) + (1-t)\phi(w) \\
\Leftrightarrow\quad & \phi(z) \le t\phi(x) + (1-t)\phi(w) \\
\Leftrightarrow\quad & 0 \le t(\phi(x) - \phi(z)) + (1-t)(\phi(w) - \phi(z)) \\
\Leftrightarrow\quad & 0 \le \tfrac{z-w}{x-w}(\phi(x) - \phi(z)) + \tfrac{x-z}{x-w}(\phi(w) - \phi(z)) \\
\Leftrightarrow\quad & 0 \ge (z-w)(\phi(x) - \phi(z)) + (x-z)(\phi(w) - \phi(z)) \\
\Leftrightarrow\quad & 0 \ge \tfrac{\phi(x)-\phi(z)}{x-z} + \tfrac{\phi(w)-\phi(z)}{z-w} \\
\because\quad & z = tx + (1-t)w, \\
& \phi(tx + (1-t)w) \le t\phi(x) + (1-t)\phi(w) \\
\Leftrightarrow\quad & \tfrac{\phi(x)-\phi(z)}{x-z} \le \tfrac{\phi(w)-\phi(z)}{w-z}
\end{aligned}$$

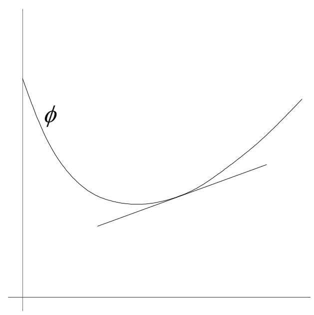
■

Corollary. *If* $\phi:[a,b]\to\mathbb{R}$ *is* C^2 *then* ϕ *is convex on* $[a,b]$ *iff* $\phi''(x)\geq 0$.

Examples. (1) $t\mapsto e^t$ is convex on $\mathbb{R}$.
(2) $t\mapsto t^\alpha$ is convex $[0,\infty)$ if $\alpha\geq 1$, but not convex if $\alpha<1$.

Proposition. (Supporting lines.) *If* $\phi:[a,b]\to\mathbb{R}$ *is convex then* $\forall x_0\in[a,b]$ $\exists m\in\mathbb{R}$ *such that* $\forall x\in[a,b]$

$$\phi(x)\geq\phi(x_0)+m(x-x_0).$$



Proof. Fix $x_0\in[a,b]$. Define $\hat{m}(x)=\frac{\phi(x)-\phi(x_0)}{x-x_0}$ for $x\neq x_0$. This is increasing in x by the previous proposition. So the following numbers exist and are finite:

$$m^+=\inf_{x>x_0}\hat{m}(x)$$
$$m^-=\sup_{x<x_0}\hat{m}(x)$$

Also, $m^-\leq m^+$. Now choose $m^-\leq m\leq m^+$ and note that

$$x>x_0\Rightarrow\frac{\phi(x)-\phi(x_0)}{x-x_0}\geq m^+\geq m\Rightarrow\phi(x)\geq\phi(x_0)+m(x-x_0)$$
$$x<x_0\Rightarrow\frac{\phi(x_0)-\phi(x)}{x_0-x}\leq m^-\leq m\Rightarrow\phi(x)\geq\phi(x_0)+m(x-x_0)$$

■

Theorem. (Jensen's Inequality) *If* $(X,\mathcal{F},\mu)$ *is a measure space,* $E\in\mathcal{F}$ *has finite measure and* $f:E\to\mathbb{R}$ *is absolutely integrable, then for any convex function* ϕ*,*

$$\phi\left(\frac{1}{\mu(E)}\int_E f\,d\mu\right)\leq\frac{1}{\mu(E)}\int_E\phi\circ f\,d\mu.$$

Proof. If $x_0=\frac{1}{\mu(E)}\int_E\phi\circ f\,d\mu$ and $m_{x_0}(x-x_0)+\phi(x_0)$ is a supporting line for ϕ then for all ξ,

$$\phi(f(\xi)) \geq \phi(x_0) + m_{x_0}\left(f(\xi) + \tfrac{1}{\mu(E)} \int_E f \, d\mu \right).$$

So,

$$\begin{aligned} \tfrac{1}{\mu(E)} \int_E \phi \circ f \, d\mu \geq \phi(x_0) & \\ &= \phi\left(\tfrac{1}{\mu(E)} \int_E f \, d\mu \right) + m_{x_0}\left(\tfrac{1}{\mu(E)} \int_E f \, d\mu - \tfrac{1}{\mu(E)} \int_E f \, d\mu \right) \\ &= \phi\left(\tfrac{1}{\mu(E)} \int_E f \, d\mu \right) \end{aligned}$$

■

Hölder's Inequality. *Let* $(X, \mathcal{F}, \mu)$ *be a measure space and suppose* $p, q > 1$*,* $\frac{1}{p} + \frac{1}{q} = 1$. *For every* $f, g : X \to [-\infty, +\infty]$ $\mathcal{F}$ *-measurable,*

$$\int_X |fg| \, d\mu \leq \left(\int_X |f|^p \, d\mu \right)^{1/p} \left(\int_X |g|^q \, d\mu \right)^{1/q}.$$

(Here we take $0\infty = 0$*.)*

Proof. We first prove *Young's Inequality*:

$$a, b \geq 0 \Rightarrow ab \leq \tfrac{a^p}{p} + \tfrac{b^q}{q}$$

$$\begin{aligned} ab &= \exp\left(\tfrac{1}{p} \log a^p + \tfrac{1}{q} \log b^q \right) \\ &\leq \tfrac{1}{p} \exp \log a^p + \tfrac{1}{q} \exp \log b^q \\ &= a^p / p + b^q / q \end{aligned}$$

by the convexity of exp. By Young's Inequality, if $A = \left(\int_X |f|^p \, d\mu \right)^{1/p}$, $B = \left(\int_X |g|^q \, d\mu \right)^{1/q}$ then $\forall x \in X$,

$$\tfrac{|f(x)|}{A} \tfrac{|g(x)|}{B} \leq \tfrac{1}{p} \tfrac{|f(x)|^p}{A^p} + \tfrac{1}{q} \tfrac{|g(x)|^q}{B^q}$$

Integrating w.r.t. μ over X gives, by monotonicity,

$$\frac{\int |fg| \, d\mu}{\left(\int |f|^p \, d\mu \right)^{1/p} \left(\int |g|^q \, d\mu \right)^{1/q}} \leq \frac{1}{p} \frac{\int |f|^p \, d\mu}{\int |f|^p \, d\mu} + \frac{1}{q} \frac{\int |g|^q \, d\mu}{\int |g|^q \, d\mu} = \frac{1}{p} + \frac{1}{q} = 1$$

■

Cauchy-Schwartz Inequality. *If* $(X,\mathcal{F},\mu)$ *is a measure space then for every* $f,g:X\to[-\infty,+\infty]$ $\mathcal{F}$*-measurable,*

$$\int_X |fg|\,d\mu \le \left(\int_X |f|^2\,d\mu\right)^{1/2}\left(\int_X |g|^2\,d\mu\right)^{1/2}.$$

Proof. Hölder with $p=q=2$.

■

Minkowski's Inequality. *Let* $(X,\mathcal{F},\mu)$ *be a measure space and* $p\ge 1$. *Then for every* $f,g:X\to[-\infty,+\infty]$ $\mathcal{F}$*-measurable,*

$$\left(\int_X |f+g|^p\,d\mu\right)^{1/p} \le \left(\int_X |f|^p\,d\mu\right)^{1/p} + \left(\int_X |g|^p\,d\mu\right)^{1/p}.$$

Proof. Define $A=\left(\int_X |f|^p\,d\mu\right)^{1/p}$, $B=\left(\int_X |g|^p\,d\mu\right)^{1/p}$. For every $x\in X$,

$$\begin{aligned}\frac{|f(x)+g(x)|^p}{(A+B)^p} &\le \left(\frac{|f(x)|+|g(x)|}{A+B}\right)^p \\ &= \left(\frac{A}{A+B}\frac{|f(x)|}{A}+\frac{B}{A+B}\frac{|g(x)|}{B}\right)^p \\ &\le \frac{A}{A+B}\frac{|f(x)|^p}{A^p}+\frac{B}{A+B}\frac{|g(x)|^p}{B^p}\end{aligned}$$

by the convexity of $t\mapsto t^p$. Integrate:

$$\frac{\left(\int|f+g|^p\,d\mu\right)^{1/p}}{\left(\int|f|^p\,d\mu\right)^{1/p}\left(\int|g|^p\,d\mu\right)^{1/p}} \le \frac{A}{A+B}\frac{\int|f|^p\,d\mu}{\int|f|^p\,d\mu}+\frac{B}{A+B}\frac{\int|g|^p\,d\mu}{\int|g|^p\,d\mu}=1$$

■

$\mathcal{L}^p$ and L^p Spaces

Minkowski's Inequality looks like a triangle inequality for a "norm":

$$\|f\|_p=\left(\int_X |f|^p\,d\mu\right)^{1/p}.$$

This suggests the following:

Definitions. Fix a measure space $(X,\mathcal{F},\mu)$ and $1\le p\le\infty$.
(1) For $f:X\to[-\infty,+\infty]$ define

$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{1/p} \text{ for } 1 \le p < \infty$$
$$\|f\|_\infty = \inf \left\{ \sup_{x \in X} |f'(x)| \;\middle|\; f' = f \ \mu\text{-a.e.} \right\}$$

(2) $\mathcal{L}^p(X, \mathcal{F}, \mu) = \{ f : X \to [-\infty, +\infty] \mid f \text{ is } \mathcal{F}\text{-measurable}, \|f\|_p < \infty \}$.

Example. The Dirichlet function

$$D(x) = \mathbf{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

has $\mathcal{L}^\infty$ norm $\|D\|_\infty = 0$ because $D = 0$ m-a.e.

Remark. $\|f\|_\infty$ is sometimes called the *essential supremum* of f and thus denoted $\operatorname{ess\,sup} f$.

Exercise. Show that if $\|f\|_\infty < \infty$ there is a bounded measurable function g such that $f = g$ a.e.

Proposition. *$\mathcal{L}^p(X, \mathcal{F}, \mu)$ is a vector space over $\mathbb{R}$ and $\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p(X, \mathcal{F}, \mu)$. I.e.*
(1) $\forall f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$, $\|f\|_p \ge 0$;
(2) $\forall f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$, $\lambda \in \mathbb{R}$, $\|\lambda f\|_p = |\lambda| \|f\|_p$;
(3) $\forall f, g \in \mathcal{L}^p(X, \mathcal{F}, \mu)$, $\|f + g\|_p \le \|f\|_p + \|g\|_p$.

Note. $\|\cdot\|_p$ is not a norm since $\exists f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$ such that $\|f\|_p = 0$ but $f \not\equiv 0$, for instance $f(x) = \mathbf{1}_{\{0\}}(x)$.

Proof. $\mathcal{L}^p(X, \mathcal{F}, \mu)$ is a vector space since $\forall f, g \in \mathcal{L}^p(X, \mathcal{F}, \mu)$, $\alpha, \beta \in \mathbb{R}$, Minkowski's Inequality says

$$\begin{aligned} \|\alpha f + \beta g\|_p &= \left(\int_X |\alpha f + \beta g|^p \, d\mu \right)^{1/p} \\ &\le \left(\int_X |\alpha f|^p \, d\mu \right)^{1/p} + \left(\int_X |\beta g|^p \, d\mu \right)^{1/p} \\ &= |\alpha| \|f\|_p + |\beta| \|g\|_p < \infty \end{aligned}$$

Properties (1) and (2) are trivial. Property (3) is Minkowski's Inequality.

■

Proposition. $\|f-g\|_p = 0 \Leftrightarrow f = g\ \mu$ - a.e.

Proof. $p=\infty$ is easy, so let $1 \le p < \infty$.

($\Leftarrow$) If $f=g$ a.e. then $|f-g|^p = 0$ a.e. and so

$$\|f-g\|_p = \left(\int_X |f-g|^p\, d\mu\right)^{1/p} = 0 .$$

($\Rightarrow$) Suppose $\|f-g\|_p = 0$. Then $\int_X |f-g|^p\, d\mu = 0$. Observe that

$$\mu[|f-g|^p \neq 0] \le \sum_{n=1}^{\infty} \mu[|f-g|^p > 1/n] \qquad (*)$$

Now,

$$\mu[|f-g|^p > 1/n] \le \int_{[|f-g|^p > 1/n]} n|f-g|^p\, d\mu$$

because the integrand is ≥ 1 on $[|f-g|^p > 1/n]$. Therefore,

$$\mu[|f-g|^p > 1/n] \le n\int_X |f-g|^p\, d\mu = n\|f-g\|_p^p = 0 .$$

By $(*)$, $f=g$ a.e.

■

This suggests that we consider $f, g \in \mathcal{L}^p$ with $\|f-g\|_p = 0$ as equivalent.

Define a relation $\sim$ on $\mathcal{L}^p(X,\mathcal{F},\mu)$ by $f \sim g$ iff $f=g$ μ-a.e. (iff $\|f-g\|_p = 0$).

Exercise. Prove that $\sim$ is an equivalence relation.

Define $[f] = \{g \in \mathcal{L}^p \mid f \sim g\}$.

Definition. The L^p *-space* of $(X,\mathcal{F},\mu)$ is

$$L^p(X,\mathcal{F},\mu) = \{[f] \mid f \in \mathcal{L}^p(X,\mathcal{F},\mu)\}$$

with the operations

$$[f]+[g]=[f+g]$$
$$\lambda[f]=[\lambda f]$$
$$\|[f]\|_p=\|f\|_p$$

Proposition. *This is a proper definition that makes* $L^p(X,\mathcal{F},\mu)$ *into a normed vector space over* $\mathbb{R}$.

Proof. (1) *Addition well-defined:* Suppose $f\sim f', g\sim g'$. Show $f+g=f'+g'$ μ-a.e.

$$\mu[f+g\neq f'+g']\leq \mu[f\neq f']+\mu[g\neq g']=0$$

(2) *Multiplication well-defined:* Exercise.
(3) *Norm well-defined:* Suppose $f\sim f'$. Then $|f|^p=|f'|^p$ μ-a.e. So

$$\|f\|_p=\left(\int|f|^p\right)^{1/p}=\left(\int|f'|^p\right)^{1/p}=\|f'\|_p .$$

(4) L^p *a vector space:* By Minkowski's Inequality.
(5) $\|\cdot\|_p$ *a norm:* $\|[f]-[g]\|_p=0\Rightarrow\|[f-g]\|_p=0\Rightarrow\|f-g\|_p=0\Rightarrow f=g$ a.e. $\Rightarrow[f]=[g]$.

■

Remark. It is traditional to drop the [ ] and refer to L^p-elements as "functions":
(1) L^p-functions are not functions;
(2) L^p-functions cannot be evaluated at a point;
(3) L^p-functions can be integrated against other functions, since if $f\sim f'$, $fg=f'g$ a.e., and so $\int_E fg\,d\mu=\int_E f'g\,d\mu$.

L^p *Convergence*

Definition. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of L^p-functions. We say that f_n *converges in* L^p to f, and write $f_n\underset{n\to\infty}{\overset{L^p}{\to}}f$, if $\|f_n-f\|_p\underset{n\to\infty}{\to}0$.

Remarks. (1) There is an analogous notion for $\mathcal{L}^p$.
(2) L^p-convergence neither implies nor is implied by convergence almost everywhere.

Examples. Take $f_n = \mathbf{1}_{[n,n+1]} \in L^p(\mathbb{R}, \mathcal{B}_0, m)$. $f_n(x) \underset{n\to\infty}{\longrightarrow} 0$ for all $x \in \mathbb{R}$, but $f_n \overset{L^p}{\underset{n\to\infty}{\not\to}} 0$ because $\|f_n\|_p = \left(\int_n^{n+1} dx\right)^{1/p} = 1$ so $\|f_n - 0\|_p \underset{n\to\infty}{\not\to} 0$.

However, consider the following array:

$$f_{1,0} = \mathbf{1}_{[0,1]}$$
$$f_{2,0} = \mathbf{1}_{[0,1/2]},\ f_{2,1} = \mathbf{1}_{[1/2,1]}$$
$$f_{3,0} = \mathbf{1}_{[0,1/3]},\ f_{3,1} = \mathbf{1}_{[1/3,2/3]},\ f_{3,2} = \mathbf{1}_{[2/3,1]}$$
$$\vdots$$
$$f_{n,k} = \mathbf{1}_{[k/n,(k+1)/n]} \text{ for } k = 0,1,\ldots,n-1$$

Let $\{g_n\}$ be the sequence $f_{1,0}, f_{2,0}, f_{2,1}, f_{3,0}, \ldots$. Then for all x

$$\liminf_{n\to\infty} g_n(x) = 0,\ \limsup_{n\to\infty} g_n(x) = 1$$

but $\|g_n\|_p \underset{n\to\infty}{\longrightarrow} 0$ for $1 \le p < \infty$ because

$$\|f_{n,k}\|_p = \left\|\mathbf{1}_{[k/n,(k+1)/n]}\right\|_p = \tfrac{1}{n} \underset{n\to\infty}{\longrightarrow} 0 .$$

Definitions. Let $(V, \|\cdot\|)$ be a normed vector space.

(1) A *Cauchy sequence* in V is a $\{v_n\}_{n=1}^{\infty} \subseteq V$ such that $\|v_m - v_n\| \underset{m,n\to\infty}{\longrightarrow} 0$.

(2) $(V, \|\cdot\|)$ is *complete* if every Cauchy sequence converges in norm to some $v \in V$.

(3) A complete normed vector space $(V, \|\cdot\|)$ is called a *Banach space*.

Theorem. (Riesz-Fischer) *If $(X, \mathcal{F}, \mu)$ is σ-finite then $L^p(X, \mathcal{F}, \mu)$ is complete with respect to $\|\cdot\|_p$, $1 \le p \le \infty$.*

Proof. We only consider the $1 \le p < \infty$ case as the $p = \infty$ case is easy and left as an exercise. Suppose $\{f_n\}_{n=1}^{\infty}$ is Cauchy, i.e. $\|f_m - f_n\|_p \underset{m,n\to\infty}{\longrightarrow} 0$.

The idea is first to find a candidate for a limit, then show that this limit lies in L^p, then establish convergence.

Consider the following identity:

$$f_{N_k} = f_{N_0} + \left(f_{N_1} - f_{N_0}\right) + \ldots + \left(f_{N_k} - f_{N_{k-1}}\right)$$

This suggests the "formula":

$$\text{“}\lim_{k\to\infty} f_{N_k} = f_{N_0} + \sum_{k=1}^{\infty}\left(f_{N_k} - f_{N_{k-1}}\right)\text{”}$$

The sum on the RHS is, in general, divergent. We find $N_k \uparrow \infty$ such that this sum converges absolutely a.e.. Choose N_i' so large that

$$m,n \ge N_i' \Rightarrow \left\|f_m - f_n\right\|_p < 1/2^i$$

We force the sequence to be increasing by setting $N_i = \max_{1\le j\le i} N_i'$. Now,

$$\begin{aligned}
\left(\int\left(\left|f_{N_0}\right| + \sum_{k=1}^{\infty}\left|f_{N_k} - f_{N_{k-1}}\right|\right)^p d\mu\right)^{1/p} &= \left(\int \lim_{n\to\infty}\left(\left|f_{N_0}\right| + \sum_{k=1}^{n}\left|f_{N_k} - f_{N_{k-1}}\right|\right)^p d\mu\right)^{1/p} \\
&\le \lim_{n\to\infty}\left(\int\left(\left|f_{N_0}\right| + \sum_{k=1}^{n}\left|f_{N_k} - f_{N_{k-1}}\right|\right)^p d\mu\right)^{1/p} \text{ by Fatou} \\
&= \lim_{n\to\infty}\left\|\left|f_{N_0}\right| + \sum_{k=1}^{n}\left|f_{N_k} - f_{N_{k-1}}\right|\right\|_p \\
&\le \lim_{n\to\infty}\left(\left\|f_{N_0}\right\|_p + \sum_{k=1}^{n}\left\|f_{N_k} - f_{N_{k-1}}\right\|_p\right) \text{by Minkowski} \\
&= \left\|f_{N_0}\right\|_p + \sum_{k=1}^{\infty}\left\|f_{N_k} - f_{N_{k-1}}\right\|_p \\
&< \infty
\end{aligned}$$

Therefore, the integrand $\left|f_{N_0}\right| + \sum_{k=1}^{\infty}\left|f_{N_k} - f_{N_{k-1}}\right|$ must be finite a.e., for otherwise its L^p norm would be ∞. This proves that

$$f(x) = \lim_{k\to\infty} f_{N_k}(x)$$

is well-defined a.e. This is our candidate.

This is an L^p function (i.e. the norm is finite) because

$$\begin{aligned}
\|f\|_p &= \left(\int \lim_{k\to\infty}\left|f_{N_k}\right|^p\right)^{1/p} \\
&\le \liminf_{k\to\infty}\left(\int\left|f_{N_k}\right|^p\right)^{1/p} \\
&= \liminf_{k\to\infty}\left\|f_{N_k}\right\|_p
\end{aligned}$$

and $\{\|f_{N_k}\|_p\}$ is bounded.

Our sequence converges to f in norm:

$$\begin{aligned}\|f - f_n\|_p &= \left\|\lim_{k\to\infty}\left(f_{N_k} - f_n\right)\right\|_p \\ &\le \liminf_{k\to\infty}\left\|f_{N_k} - f_n\right\|_p\end{aligned}$$

If $n > N_l$ then for $k > l$, $\left\|f_{N_k} - f_n\right\|_p < 1/2^l$. So $n > N_l \Rightarrow \|f - f_n\|_p < 1/2^l$. So $\|f_n - f\|_p \underset{n\to\infty}{\rightarrow} 0$.

■

Corollary. *If* $f_n \overset{L^p}{\underset{n\to\infty}{\rightarrow}} f$ *then* $\exists N_k \uparrow \infty$ *such that* $f_{N_k} \underset{k\to\infty}{\rightarrow} f$ *a.e.*

Corollary. $\sum_{n=1}^{\infty}\|f_n\|_p < \infty \Rightarrow \left\{\sum_{n=1}^{N} f_n\right\}_{N=1}^{\infty}$ *converges in* L^p. *(The limit is denoted by* $\sum_{n=1}^{\infty} f_n$.

Exercise. Reconcile the first corollary with the second example above.

The Case $p = 2$

Suppose $(X, \mathcal{F}, \mu)$ is σ-finite. We can define an inner product on $L^2(X, \mathcal{F}, \mu)$ via

$$(f, g) = \int_X fg\, d\mu .$$

This is well-defined:

(1) If $f = f'$ a.e. and $g = g'$ a.e. then $fg = f'g'$ a.e. and so $\int fg = \int f'g'$.

(2) fg is absolutely integrable since

$$\int_X |fg|\, d\mu \le \left(\int_X |f|^2\, d\mu\right)\left(\int_X |g|^2\, d\mu\right) = \|f\|_2 \|g\|_2 < \infty$$

and it is related to the L^2 norm in the right way: $\|f\|_2 = (f, f)^{1/2}$.

Definition. A complete inner product space $(V,(\cdot,\cdot))$ is called a *Hilbert space*.

Definition. A *bounded linear functional* on L^2 is a linear function $\phi : L^2 \to \mathbb{R}$ such that $\exists A > 0$ such that $|\phi(f)| \le A\|f\|_2$ for all $f \in L^2$.

Example. If $g_0 \in L^2$ is fixed then $\phi_{g_0}(f) = (f, g_0) = \int_X fg_0 \, d\mu$ is a bounded linear functional.

Proof. It is clear that $\phi_{g_0}(f)$ is independent of the choice of representative. It is also clear that ϕ_{g_0} is linear. It is bounded because the Cauchy-Schwarz Inequality says

$$|\phi_{g_0}(f)| \le \int_X |fg_0| \, d\mu = \|f\|_2 \|g_0\|_2 .$$

■

The converse to this is the Riesz Representation Theorem:

Theorem. (Riesz Representation Theorem) *If $(X, \mathcal{F}, \mu)$ is σ-finite then every bounded linear functional on $L^2(X, \mathcal{F}, \mu)$ is of the form $f \mapsto \int_X fg \, d\mu$ for some fixed $g \in L^2(X, \mathcal{F}, \mu)$.*

Absolute Continuity

Basic Example. Let $(X, \mathcal{F}, \nu)$ be a σ-finite measure space and suppose $f : X \to \mathbb{R}$ is non-negative and measurable. Define $\mu : \mathcal{F} \to \mathbb{R}$ by $\mu(E) = \int_E f \, d\nu$.

Exercise. (1) Prove that μ is a measure.
(2) Show that $(X, \mathcal{F}, \mu)$ is σ-finite.

Definition. Let μ, ν be two σ-finite measures on $(X, \mathcal{F})$. We say that μ is *absolutely continuous* with respect to ν, and write $\mu << \nu$, if $\nu(E) = 0 \Rightarrow \mu(E) = 0$.

Examples. (1) If $\mu(E) = \int_E f \, d\nu$ for all $E \in \mathcal{F}$ then $\mu << \nu$.

(2) Let ν be Lebesgue measure on $\mathbb{R}$, let $\mathbb{Q} = \{q_n\}_{n=1}^{\infty}$ be an enumeration of $\mathbb{Q}$, and define

$$\mu(E) = \sum_{n=1}^{\infty} \tfrac{1}{2^n} \mathbf{1}_E(q_n).$$

Exercise. (1) Show that μ is a finite measure.

(2) Show that μ is not absolutely continuous with respect to ν. (Hint: calculate $\mu(\mathbb{Q})$ and $\nu(\mathbb{Q})$.)

Theorem. (Radon-Nikodym) *Let μ,ν be two σ-finite measures on $(X,\mathcal{F})$. If $\mu << \nu$ then there exists a non-negative measurable function $f : X \to \mathbb{R}$ such that for all μ-absolutely integrable functions g,*

$$\int_X g\,d\mu = \int_X gf\,d\nu .$$

Moreover, f is unique up to μ-null sets.

Definition. In this case we write $d\mu = f\,d\nu$, $f = \frac{d\mu}{d\nu}$, and call $f = \frac{d\mu}{d\nu}$ the *Radon-Nikodym derivative* of μ with respect to ν.

Proof. We only treat the finite measure case $\mu(X), \nu(X) < \infty$; the general case is handled in the standard way by breaking X into $\aleph_0$ pieces of finite $(\mu+\nu)$-measure.

Define $\lambda = \mu + \nu$ and let $\phi : L^2(X,\mathcal{F},\lambda) \to \mathbb{R}$ be the functional

$$\phi(f) = \int_X f\,d\mu .$$

We need to show that ϕ is a well-defined bounded linear functional:

ϕ is well defined, for suppose $f_1 = f_2$ λ-a.e. Then

$$0 = \lambda[f_1 \neq f_2] = \mu[f_1 \neq f_2] + \nu[f_1 \neq f_2] \geq \mu[f_1 \neq f_2].$$

It follows that $f_1 = f_2$ μ-a.e. and so $\phi(f_1) = \int_X f_1\,d\mu = \int_X f_2\,d\mu = \phi(f_2)$. The linearity of ϕ is clear. To see that it is bounded use the Cauchy-Schwarz Inequality:

$$|\phi(f)| \leq \int |f|\,d\mu \leq \int |f|\,d\lambda \leq \sqrt{\int 1^2\,d\lambda}\sqrt{\int |f|^2\,d\lambda} = \sqrt{\lambda(X)}\|f\|_2 < \infty$$

By the Riesz Representation Theorem, $\exists h \in L^2(X,\mathcal{F},\lambda)$ such that $\forall f \in L^2(X,\mathcal{F},\lambda)$, $\phi(f) = (f,h)$. I.e., for $f \in L^2(X,\mathcal{F},\lambda)$,

$$\int f\,d\mu = \int fh\,d\lambda = \int fh\,d\mu + \int fh\,d\nu$$

$$\Rightarrow \qquad \int f(1-h)\,d\mu = \int fh\,d\nu \qquad (*)$$

This suggests $(1-h)d\mu = h\,d\nu$, and so $d\mu = \frac{h}{1-h}d\nu$. We make this manipulation rigorous:

Claim. $0 \le h < 1$ λ-a.e.

Proof. We start by noting that $\forall E \in \mathcal{F}$, $\mathbf{1}_E \in L^2(X,\mathcal{F},\lambda)$, and so

$$\mu(E) = \phi(\mathbf{1}_E) = \int_E h\,d\lambda \tag{A}$$

$$\nu(E) = \lambda(E) = \mu(E) = \int_E (1-h)\,d\lambda \tag{B}$$

Now

(A) $\Rightarrow$ $$0 \ge \int_{[h<0]} h\,d\lambda = \mu[h<0]k \ge 0$$

$\Rightarrow$ $$\int_{[h<0]} h\,d\lambda = 0$$

$\Rightarrow$ $$\lambda[h<0] = 0$$

Exercise. Prove the last implication.

(B) $\Rightarrow$ $$0 \ge \int_{[h\ge 1]} (1-h)\,d\lambda = \nu[h \ge 1] \ge 0$$

$\Rightarrow$ $$\nu[h\ge 1] = 0$$

Since $\mu << \nu$, $\mu[h \ge 1] = 0$ as well, so $\lambda[h\ge 1] = 0$. Hence $0 \le h < 1$ λ-a.e. This proves the claim.

Since $\lambda(X) < \infty$, every indicator function is in $L^2(X,\mathcal{F},\lambda)$, and so $(*)$ implies that for each $E \in \mathcal{F}$, $\int \mathbf{1}_E \cdot (1-h)\,d\mu = \int \mathbf{1}_E \cdot h\,d\nu$. By linearity, for all simple functions f,

$$\int f\cdot(1-h)\,d\mu = \int fh\,d\nu .$$

If f is non-negative measurable then choose simple $0 \le \phi_n \uparrow f$. Note that $0 \le \phi_n h \uparrow fh$, $0 \le \phi_n \cdot (1-h) \uparrow f\cdot(1-h)$. Then

$$\begin{aligned}\int f\cdot(1-h)\,d\mu &= \lim_{n\to\infty}\int \phi_n\cdot(1-h)\,d\mu \text{ by MCT}\\ &= \lim_{n\to\infty}\int \phi_n h\,d\nu \text{ since } \phi_n \text{ simple}\\ &= \int fh\,d\nu \text{ by MCT}\end{aligned}$$

This proves that for $f \ge 0$ measurable,

$$\int f \cdot (1-h) d\mu = \int fh \, d\nu \, . \qquad (**)$$

But if f is non-negative measurable then so is $f \cdot \frac{h}{1-h}$. Therefore, by $(**)$,

$$\begin{aligned} \int f \, d\mu &= \int \tfrac{f}{1-h} \cdot (1-h) d\mu \\ &= \int \tfrac{f}{1-h} h \, d\nu \\ &= \int f \tfrac{h}{1-h} d\nu \end{aligned}$$

for all f non-negative measurable.

Exercise. Generalize to all μ-absolutely integrable functions $f = f^+ - f^-$.

■

Exercise. Show the Radon-Nikodym derivative is unique modulo λ. (Hint: suppose $\int fg_i \, d\mu = \int fg_i \, d\nu$ for $i = 1,2$ and evaluate $\int (\text{sgn}(g_1 - g_2))(g_1 - g_2) d\nu$.)

Signed Measures

Imagine that electric charge is distributed in space. It makes sense to define $\mu(E)$ to be the total charge within the region E. However, this μ is not a measure because it is not, in general, non-negative.

Definition. Let $(X, \mathcal{F})$ be a measurable space. A *signed measure* is a set function $\mu : \mathcal{F} \to [-\infty, +\infty]$ such that
(1) μ attains at most one of the values $\pm\infty$;
(2) $\mu(\varnothing) = 0$;
(3) σ-additivity: if $E = \coprod_{i=1}^{\infty} E_i$ with $E_i \in \mathcal{F}$ then $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$, where

(a) if $|\mu(E)| = \infty$ then the convergence of $\sum_{i=1}^{\infty} \mu(E_i)$ to $\mu(E)$ is meant,

(b) if $|\mu(E)| < \infty$ then the absolute convergence of $\sum_{i=1}^{\infty} \mu(E_i)$ to $\mu(E)$ is meant.

Remarks. (1) was added to prevent the existence of two disjoint sets $A, B \in \mathcal{F}$ such that $\mu(A) = +\infty$, $\mu(B) = -\infty$, for in this case $\mu(A \coprod B) = \infty - \infty$.
(3) was strengthened to make sure all sums converge.

Definitions. Let $(X, \mathcal{F}, \mu)$ be a signed measure space.
(1) $E \in \mathcal{F}$ is called a *null set* if $\forall E' \subseteq E$ measurable, $\mu(E') = 0$. (This is stronger than saying $\mu(E) = 0$.)

(2) $E \in \mathcal{F}$ is *positive* if $\forall E' \subseteq E$ measurable, $\mu(E') \geq 0$.
(3) $E \in \mathcal{F}$ is *positive* if $\forall E' \subseteq E$ measurable, $\mu(E') \leq 0$.

Lemma. (Exhaustion Lemma) *Let $(X, \mathcal{F}, \mu)$ be a signed measure space. Every $E \in \mathcal{F}$ with $0 < \mu(E) < \infty$ contains a subset which is a positive set of strictly positive measure.*

Proof. Define $E_1 = E$ and set

$$\mu_1 = \sup\{|\mu(A)| \mid A \subseteq E_1 \text{ is measurable and } \mu(A) \leq 0\}.$$

Now choose $A_1 \subseteq E_1$ measurable such that

$$\mu(A_1) \leq -1 \text{ if } \mu_1 = +\infty,$$
$$\mu(A_1) \leq \mu_1/2 \text{ if } \mu_1 < +\infty.$$

Define $E_2 = E_1 \setminus A_1 = E \setminus A_1$. Set

$$\mu_2 = \sup\{|\mu(A)| \mid A \subseteq E_2 \text{ is measurable and } \mu(A) \leq 0\}.$$

(Note that if $\mu_1 < \infty$ then $\mu_2 \leq \mu_1/2$.) Choose $A_2 \subseteq E_2$ measurable such that

$$\mu(A_2) \leq -1 \text{ if } \mu_2 = +\infty,$$
$$\mu(A_2) \leq \mu_2/2 \text{ if } \mu_2 < +\infty.$$

Proceed by induction. Suppose $A_1, \dots A_n$ already defined. Set

$$\mu_{n+1} = \sup\{|\mu(A)| \mid A \subseteq E_n \text{ is measurable and } \mu(A) \leq 0\}.$$

Choose $A_{n+1} \subseteq E_{n+1}$ measurable such that

$$\mu(A_{n+1}) \leq -1 \text{ if } \mu_{n+1} = +\infty,$$
$$\mu(A_{n+1}) \leq \mu_{n+1}/2 \text{ if } \mu_{n+1} < +\infty.$$

We obtain in this way $\{A_i\}_{i=1}^{\infty}$ pairwise disjoint. Define $\widetilde{E} = E \setminus \coprod_{i=1}^{\infty} A_i$. We show that $\widetilde{E}$ is positive and $\mu(\widetilde{E}) > 0$.

Step 1. $\mu_n \underset{n\to\infty}{\to} 0$.

Proof. First note that $\exists N \in \mathbb{N}$ such that $\mu_N < +\infty$, for otherwise $\forall n \in \mathbb{N}$, $\mu_n = +\infty$ and so $\mu(A_n) \le -1$. Hence,

$$0 < \mu(E) = \mu\left(\tilde{E} \amalg \left(\amalg_{i=1}^{\infty} A_i\right)\right) = \underbrace{\mu(\tilde{E})}_{\in[-\infty,+\infty)} + \underbrace{\sum_{i=1}^{\infty} \mu(A_i)}_{=-\infty} = -\infty,$$

which is a contradiction. By the definition of μ_N, $\mu_{N+1} \le \mu_N/2$, so for $n > N$, $\mu_n \le \mu_N / 2^{n-N} \underset{n\to\infty}{\to} 0$.

Step 2. $\tilde{E}$ is positive.

Proof. $E' \subseteq \tilde{E}$ is measurable. By definition, $E' \subseteq E_n = E \setminus \amalg_{i=1}^{n-1} A_i$, so if $\mu(E') \le 0$ then $|\mu(E')| \le \mu_n \underset{n\to\infty}{\to} 0$. Thus, all subsets of $\tilde{E}$ with non-positive measure have measure zero, so $\tilde{E}$ is positive.

Step 3. $\mu(\tilde{E}) > 0$.

Proof. $0 < \mu(E) = \mu(\tilde{E}) + \underbrace{\sum_{i=1}^{\infty} \mu(A_i)}_{<0}$.

■

Theorem. (Hahn's Decomposition Theorem) *Let $(X, \mathcal{F}, \mu)$ be a signed measure space. There exist sets $X^{\pm} \in \mathcal{F}$ such that $X = X^+ \amalg X^-$, X^+ is positive and X^- is negative. If $X = X_1^+ \amalg X_1^-$ is another such decomposition then $X^+ \,\Delta\, X_1^+$ and $X^- \,\Delta\, X_1^-$ are null sets.*

Proof. Assume without loss of generality that μ omits the value $+\infty$ (else pass to $-\mu$). Define $m = \{\mu(E) \mid E \in \mathcal{F} \text{ is positive}\}$. This is finite because μ omits the value $+\infty$. Choose $E_n \in \mathcal{F}$ positive such that $\mu(E_n) \underset{n\to\infty}{\to} m$. Set $X^+ = \bigcup_{n=1}^{\infty} E_n$, $X^- = X \setminus X^+$.

Exercise. Show that a countable union of positive sets is positive.

So X^+ is positive. X^- is negative because otherwise there exists $E' \subseteq X^-$ measurable such that $\mu(E') > 0$. In this case $X^+ \amalg E'$ is a positive set with measure $m + \mu(E') > m$, which contradicts the maximality of m.

This proves the existence of a Hahn decomposition. We now prove uniqueness. Let $X = X_1^+ \amalg X_1^-$ be another Hahn decomposition.

Suppose $E \subseteq X^+ \,\Delta\, X_1^+$. Since

$$X^+ \Delta X_1^+ = (X^+ \setminus X_1^+) \cup (X_1^+ \setminus X^+) \subseteq X^+ \cup X_1^+,$$

$\mu(E') \geq 0$ since $X^+ \cup X_1^+$ is positive. But we also have

$$X^+ \Delta X_1^+ = (X^+ \cap X_1^-) \cup (X_1^+ \cap X^-) \subseteq X_1^- \cup X^-,$$

so $\mu(E') \leq 0$ since $X_1^- \cup X^-$ is negative. So $\mu(E') = 0$. Since E' was arbitrary, $X^+ \Delta X_1^+$ is null. The proof for $X^- \Delta X_1^-$ is similar.

■

Definition. Let μ, ν be two measures on $(X, \mathcal{F})$. We say μ, ν are *mutually singular* amd write $\mu \perp \nu$ if $X = X_\mu \coprod X_\nu$ where $\mu(X_\nu) = \nu(X_\mu) = 0$. (I.e. μ "lives on" X_μ, ν "lives on" X_ν, $X_\mu \cap X_\nu = \varnothing$).

Exercise. Show that if $\mu << \nu$ then μ and ν are not mutually singular.

Theorem. (Jordan's Decomposition Theorem) *Let $(X, \mathcal{F}, \mu)$ be a signed measure space. There exists a unique decomposition $\mu = \mu^+ - \mu^-$ where μ^+, μ^- are (proper) measures and $\mu^+ \perp \mu^-$.*

Proof. Let $X = X^+ \coprod X^-$ be a Hahn decomposition. Define two measures μ^+, μ^- by

$$\mu^+(E) = \mu(E \cap X^+),$$
$$\mu^-(E) = -\mu(E \cap X^-).$$

Exercise. Show μ^+, μ^- are measures.

To see that these measures are mutually singular take $X_{\mu^+} = X^+$, $X_{\mu^-} = X^-$ and observe that $\mu = \mu^+ - \mu^-$.

Suppose $\mu = \mu_1^+ - \mu_1^-$ is another Jordan decomposition. Let $X = X_1^+ \coprod X_1^-$ be a decomposition such that $\mu_1^+(X_1^-) = \mu_1^-(X_1^+) = 0$. We claim that this is a Hahn decomposition.

To see that X_1^+ is positive: $\forall E' \subseteq X_1^+$, $\mu(E') = \mu_1^+(E') - \mu_1^-(E') = \mu_1^+(E') \geq 0$. The negativity of X_1^- is shown similarly.

By the uniqueness part of the Hahn Decomposition Theorem, both $X^+ \Delta X_1^+$ and $X^- \Delta X_1^-$ are null sets. Therefore, for every $E \in \mathcal{F}$,

$$\begin{aligned}\mu_1^+(E) &= \mu_1^+(E \cap X_1^+) - \mu_1^-(E \cap X_1^+) \\ &= \mu(E \cap X_1^+) \\ &= \mu(E \cap X^+) \text{since } X^+ \Delta X_1^+ \text{ null} \\ &= \mu^+(E \cap X^+) \\ &= \mu^+(E)\end{aligned}$$

Similarly for μ_1^-, μ^-.

■