MEASURE AND INTEGRATION

Math 631

Author
Bill Runzhe Zhao
Duke University
2024 fall

Contents

1 Measure Theory and Integration

1.1 Elementary Sets and Measure

The motivation to start this study is to lay a foundation to assign a measure to arbitrary sets $E \subset \mathbb{R}^d$. The set can be nice set such as intervals, boxes, and polygons, but it can also be any generic or strange set that we wish to study. We want a measure that is well behaved under limits (e.g. Lebesgue measure is well behaved under pointwise limits).

Definition 1.1.1. *(Finite Intervals).* $(a, b), (a, b], [a, b), [a, b]$ *for* $a, b \in \mathbb{R}$. *The length is* $b - a$ *and it can be zero.*

Definition 1.1.2. *(Box).* $B \subset \mathbb{R}^d$ *is a Cartesian product of intervals* $B = I_1 \times ... \times I_d$

Definition 1.1.3. *(Elementary Set). Elementary Set is any subset of* $\mathbb{R}^d$ *that is a finite union of boxes.*

We denote the set of elementary sets as $\mathcal{E}(\mathbb{R}^d)$. It does not include unbounded sets.

Proposition 1.1.1. *(Boolean closure). The set* $\mathcal{E}(\mathbb{R}^d)$ *is closed taking unions* $(E \cup F)$*, intersects* $(E \cap F)$*, differences* $(E \setminus F)$ *and symmetric differences* $(E \Delta F = (E \setminus F) \cup (F \setminus E))$*, as well as translations:* $E \in \mathcal{E}(\mathbb{R}^d), x \in \mathbb{R}^d \Rightarrow E + x = \{y + x : y \in E\} \in \mathcal{E}(\mathbb{R}^d)$.

Proof. Let $E, F \in \mathcal{E}(\mathbb{R}^d)$, then each of them is a finite union of boxes. $E = \bigcup_{i=1}^{n} B_i$, $F = \bigcup_{i=1}^{m} B'_i$.
Denote the collection of $\{B_i\}_{i=1}^n$ and $\{B'_i\}_{i=1}^m$ as a new set of boxes: $\{C_i\}_{i=1}^{n+m} = \{B_1, ..., B_n, B'_1, ..., B'_m\}$
Then,

$$\begin{aligned} E \cup F &= (\bigcup_{i=1}^{n} B_i) \cup (\bigcup_{i=1}^{m} B'_i) \\ &= \bigcup_{i=1}^{n+m} C_i \end{aligned}$$

Thus, $E \cup F$ is a finite union of boxes and is an elementary set.

$$\begin{aligned} E \cap F &= (\bigcup_{i=1}^{n} B_i) \cap (\bigcup_{i=1}^{m} B'_i) \\ &= (B_1 \cap (\bigcup_{i=1}^{m} B'_i)) \cup ... \cup (B_n \cap (\bigcup_{i=1}^{m} B'_i) \\ &= ((B_1 \cap B'_1) \cup ... \cup (B_1 \cap B'_m)) \cup ... \cup ((B_n \cap B'_1) \cup ... \cup (B_n \cap B'_m)) \\ &= (B_1 \cap B'_1) \cup ... \cup (B_1 \cap B'_m) \cup ... \cup (B_n \cap B'_1) \cup ... \cup (B_n \cap B'_m) \end{aligned}$$

Now we prove that the intersection of two boxes is a box. Suppose $B_i = I_1^i \times ... \times I_d^i$, $B'_j = I_1^{'j} \times ... \times I_d^{'j}$, then $B_i \cap B'_j = (I_1^i \cap I_1^{'j}) \times ... \times (I_d^i \cap I_d^{'j})$. Since the intersection of two intervals is an interval (degenerate or non-degenerate), by definition, $B_i \cap B'_j$ is a box.
Then, $E \cap F$ is a finite union of $n \times m$ boxes, and it is an elementary set.

Now we look at difference.

$$\begin{aligned}
E \setminus F &= E \cap F^C \\
&= (\bigcup_{i=1}^{n} B_i) \cap (\bigcup_{i=1}^{m} B_i')^C \\
&= (\bigcup_{i=1}^{n} B_i) \cap (\bigcap_{i=1}^{m} B_i^{'C}) \\
&= (B_1 \cap (\bigcap_{i=1}^{m} B_i^{'C})) \cup ... \cup (B_n \cap (\bigcap_{i=1}^{m} B_i^{'C})) \\
&= ((B_1 \cap B_1^{'C}) \cap ... \cap (B_1 \cap B_m^{'C})) \cup ... \cup ((B_n \cap B_1^{'C}) \cap ... \cap (B_n \cap B_m^{'C})) \\
&= ((B_1 \setminus B_1^{'}) \cap ... \cap (B_1 \setminus B_m^{'})) \cup ... \cup ((B_n \setminus B_1^{'}) \cap ... \cap (B_n \setminus B_m^{'}))
\end{aligned}$$

Since B_i and $B_j^{'}$ are boxes, $B_i \setminus B_j^{'}$ are elementary. Then, by the first two parts that we just proved, each $((B_i \setminus B_1^{'}) \cap ... \cap (B_i \setminus B_m^{'}))$ is elementary, and their union is elementary. Thus, $E \setminus F$ is an elementary set.
Next, since E and F are elementary, $E \setminus F$ is elementary and $F \setminus E$ is also elementary. By the part we just proved, their union $(E \setminus F) \cup (F \setminus E)$ is also elementary. Thus, $E \Delta F$ is an elementary set.
Next,

$$\begin{aligned}
E + x &= \bigcup_{i=1}^{n} B_i + x \\
&= \bigcup_{i=1}^{n} (I_1^i + x) \times ... \times (I_d^i + x) \\
&= \bigcup_{i=1}^{n} I_1^{'i} \times ... \times I_d^{'i} \\
&= \bigcup_{i=1}^{b} B_i'
\end{aligned}$$

where we have translation of each interval by x in the second and third equality. Thus, $E + x$ is an elementary set.
Thus, we finished the proof that $\mathcal{E}(\mathbb{R}^d)$ is closed under union, intersection, difference, symmetric difference, and translation. □

Now it is time to give the elementary set a measure.

Lemma 1.1.1. *Let $E \subset \mathbb{R}^d$ be an elementary set.*
(1). E can be expressed (partitioned) as the finite union of disjoint boxes, that is, $E = \bigcup_{i=1}^{n} B_i$ for B_i pairwise disjoint.
(2). For any two such partitions $E = \bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{m} B_i'$, we have $\sum_{i=1}^{n} |B_i| = \sum_{i=1}^{m} |B_i'|$. We denote this value by $m(E) = m^d(E)$, the elementary measure of E. It is independent of partitions.

Proof. Begin with (i).

For the case $d = 1$, $E = \bigcup_{i=1}^n I_i$. Place the $2n$ endpoints of these intervals in increasing order (discarding repetitions) and relable them to be $c_1 \leq c_2 \leq ... \leq c_{2n}$.
Let $J_1, ...J_{4n-1}$ be disjoint intervals formed by these endpoints where

$$J_i = \{c_i\} \ for \ 1 \leq i \leq 2n$$

$$J_i = (c_{i-2n}, c_{i-2n+1}) \ for \ 2n+1 \leq i \leq 4n-1$$

Basically we have $2n$ endpoints and $2n-1$ open intervals between endpoints by doing this. Then, each I_i can be expressed as some subcollection of $J_1, ..., J_{4n-1}$. Then, $E = \bigcup_{k: J_k \cap E \neq \emptyset} J_k$.
For general cases $d \geq 2$ Now we have $E = \bigcup_{i=1}^n B_i$ where $B_i = I_i^1 \times ... \times I_i^d$. For each m such that $1 \leq m \leq d$, we apply the $d = 1$ case to get a family of disjoint intervals $\{J_k^m\}_{k=1}^{n_m}$ such that $\bigcup_{i=1}^n I_i^m = \bigcup_{i=1}^{n_m} J_k^m$.
Then, we can get $n_1 \times ... \times n_d$ pairwise disjoint boxes and each of them is represented as $\tilde{B}_{k_1 ... k_d} = J_{k_1}^1 \times ... \times J_{k_d}^d$ for $(k_1, ..., k_d) \in \{1, ..., n_1\} \times ... \times \{1, ..., n_d\}$.
Then, each B_i is a union of a subcollection of $\{\tilde{B}_{k_1 ... k_d}\}$, and thus E is expressed as a finite union of pairwise disjoint boxes.

Now let's prove (ii).
First we notice that for any interval I, we have that

$$|I| = \lim_{N \to +\infty} \frac{1}{N} \#(I \cap \frac{1}{N}\mathbb{Z})$$

where $\frac{1}{N}\mathbb{Z} = \{\frac{k}{N} : k \in \mathbb{Z}\}$.
One way to understand it is to take a sample of rational points in (a, b) with the distance between each adjacent pair to be $\frac{1}{N}$. Another way is through an example where $a = 2$, $b = 4$. Then, $(a, b) \cap \frac{1}{N}\mathbb{Z} = \{z : z \in (aN, bN), z \in \mathbb{Z})\}$. When $n = 1$, $\# = 1$; $n = 2$, $\# = 3$, ... Then the cardinality is equal to $(b - a)N - 1 = 2N - 1$ with each N. Then $\lim_{N \to +\infty} \frac{1}{N}((b - a)N - 1) = b - a = |(a, b)|$.
By taking Cartesian product,

$$|B| = \lim_{N \to +\infty} \frac{1}{N^d} \#(B \cap \frac{1}{N}\mathbb{Z}^d)$$

For $E = \cup_{i=1}^n B_i$ with pairwise disjoint B_i,

$$\begin{aligned} \frac{1}{N^d} \#(E \cap \frac{1}{N}\mathbb{Z}^d) &= \sum_{i=1}^n \frac{1}{N^d} \#(B_i \cap \frac{1}{N}\mathbb{Z}^d) \\ &\to \sum_{i=1}^n |B_i| \ as \ N \to +\infty \end{aligned}$$

For $E = \cup_{i=m}^{n} B_i'$ with pairwise disjoint B_i',

$$\begin{aligned}\frac{1}{N^d}\#(E \cap \frac{1}{N}\mathbb{Z}^d) &= \sum_{i=1}^{m} \frac{1}{N^d}\#(B_i' \cap \frac{1}{N}\mathbb{Z}^d) \\ &\to \sum_{i=1}^{m} |B_i'| \ as\ N \to +\infty\end{aligned}$$

The LHS is independent of partitions, thus $m(E) = \sum_{i=1}^{n} |B_i| = \sum_{i=1}^{m} |B_i'| =: m^d(E)$. $\square$

Theorem 1.1.1. *(Uniqueness of elementary measure). Let $d \geq 1$. Let $m' : \mathcal{E}(\mathbb{R}^d) \to \mathbb{R}^+$ be a map from the collection $\mathcal{E}(\mathbb{R}^d)$ of elementary subsets of $\mathbb{R}^d$ to the nonnegative reals that obeys the non-negativity, finite additivity, and translation invariance properties. Then there exists a constant $c \in \mathbb{R}^+$ such that $m'(E) = cm(E)$ for all elementary sets E. In particular, if we impose the additional normalisation $m'([0,1)^d) = 1$, then $m' \equiv m$. (Hint: Set $c := m'([0,1)^d)$, and then compute $m'([0, \frac{1}{n})^d)$ for any positive integer n.)*

Proof. In this proof we will use m to represent elementary measure.
First we observes that any $E \in \mathcal{E}(\mathbb{R}^d)$ can be expressed as a finite union of translated and stretched $[0,a)^d$ type sets together with with the boundary with zero measure. So we only need to work with type $[0,a)^d$ set.
Set $m'([0,1)^d) := c$. Then, it can be written as n^d finite disjoin unions of the translated $[0, \frac{1}{n})^d$. Therefore, by finite additivity, $m'([0, \frac{1}{n})^d) = \frac{m'([0,1)^d)}{n^d} = \frac{c}{n^d} = cm([0, \frac{1}{n})^d)$.
Without loss of generality, let the elementary set $E = \prod_{i=1}^{d} [a_i, b_i) \subset \mathbb{R}^d$. By translation invariance, $m'(E) = m'(E - (a_1, ..., a_d)) = m'(\prod_{i=1}^{d} [0, b_i - a_i))$.
First, consider the case where $b_i - a_i$ is rational. Then it can be represented in the form $\frac{p_1}{n}, ..., \frac{p_d}{n}$ using some common numerator n. Then, by partition into disjoint sets and boundaries,

$$\begin{aligned}\prod_{i=1}^{d} [0, b_i - a_i) &= \prod_{i=1}^{d} [0, \frac{p_i}{n}) \\ &= \prod_{i=1}^{d} \bigcup_{k_i \in \mathbb{Z}; 0 \leq k_i \leq p_i - 1} [\frac{k_i}{n}, \frac{k_i + 1}{n}) \\ &= \bigcup_{k \in \{(k_1, ..., k_d) : k_i \in \mathbb{Z}; 0 \leq k_i \leq p_i - 1\}} [0, \frac{1}{n})^d + k\end{aligned}$$

Then, by finite additivity and zero measure on boundary,

$$
\begin{aligned}
m'(\prod_{i=1}^{d}[0, b_i - a_i)) &= \sum_{k \in \{(k_1,...,k_d):k_i \in \mathbb{Z}; 0 \le k_i \le p_i - 1\}} m'([0, \frac{1}{n})^d + k) \\
&= \sum_{k \in \{(k_1,...,k_d):k_i \in \mathbb{Z}; 0 \le k_i \le p_i - 1\}} m'([0, \frac{1}{n})^d) \\
&= c\prod_{i=1}^{d}\frac{p_i}{n} \\
&= cm(\prod_{i=1}^{d}[0, \frac{p_i}{n}))
\end{aligned}
$$

Next, consider the case where $b_i - a_i$ is real. By the density of rationals, find two rational sequences $\{s_n^i\}_{n\in\mathbb{N}} < b_i - a_i < \{q_n^i\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} s_n^i = \lim_{i\to\infty} q_n^i = b_i - a_i$.
Then, we have $\prod_{i=1}^{d}[0, s_n^i) \subseteq \prod_{i=1}^{d}[0, b_i - a_i) \subseteq \prod_{i=1}^{d}[0, q_n^i)$.
Here we prove the monotonicity first. Let $A \subseteq B$, then, $m(B) = m((B \setminus A) \cup A) = m(B \setminus A) + m(A) \ge m(A)$. Then,

$$
m'(\prod_{i=1}^{d}[0, s_n^i)) \le m'(\prod_{i=1}^{d}[0, b_i - a_i)) \le m'(\prod_{i=1}^{d}[0, q_n^i))
$$

From the above calculation with the rationals we know that

$$
cs_n^1 s_n^2 ... s_n^d \le m'(\prod_{i=1}^{d}[0, b_i - a_i)) \le cq_n^1 q_n^2 ... q_n^d
$$

By the limiting and sandwich rule, we have

$$
\begin{aligned}
m'(\prod_{i=1}^{d}[0, b_i - a_i)) &= c(b_1 - a_1)(b_2 - a_2)...(b_d - a_d) \\
&= cm(\prod_{i=1}^{d}[0, b_i - a_i))
\end{aligned}
$$

We already proved that for elementary set $m'(E) = cm(E) = m'([0,1)^d)m(E)$. Then, if $m'([0,1)^d) = 1$, we have $m' \equiv m$. □

Now we have the property of the elementary measure:

Proposition 1.1.2. *(i).* $m(E) \ge 0$

(ii). Finite additivity: $m(E \cup F) = m(E) + m(F)$ *for disjoint* $E, F \in \mathcal{E}(\mathbb{R}^d)$. *By induction,* $m(E_1 \cup ... \cup E_n) = \sum_{i=1}^{n} m(E_i)$ *for disjoint* $\{E_i\}$.

(iii). $m(\emptyset) = 0$.

(iv). $m(B) = |B|$ $\forall$ *box* B.

(v). Monotonicity: $E \subset F$, *then* $m(E) \le m(F)$.

(vi). Sub-additivity: $m(E \cup F) \le m(E) + m(F)$ *for any* $E, F \in \mathcal{E}(\mathbb{R}^d)$. *By induction:* $m(E_1 \cup ... \cup E_n) \le \sum_{i=1}^{n} m(E_i)$.

(vii). $E \in \mathcal{E}(\mathbb{R}^d)$, $x \in \mathbb{R}$, then $m(E + x) = m(E)$.

Proof. (v). $F = E \cup (F \setminus E)$, so $m(F) \overset{\text{(ii)}}{=} m(E) + m(F \setminus E) \overset{\text{(i)}}{\geq} m(E)$.

(vi). Finite sub-additivity: $E \cup F = E \cup (F \setminus E)$, therefore $m(E \cup F) \overset{\text{(ii)}}{=} m(E) + m(F \setminus E) \overset{\text{(v)}}{\leq} m(E) + m(F)$.

(vii). E can be partitioned into finite disjoint boxes $E = \bigcup_{i=1}^n B_i$, where each $B_i = I_i^1 \times ... \times I_i^d$. Then, $E' = E + x = \bigcup_{i=1}^n B_i + x = \bigcup_{i=1}^n I_i^1 \times ... \times I_i^d + x = \bigcup_{i=1}^n (I_i^1 + x) \times ... \times (I_i^1 + x) = \bigcup_{i=1}^n B_i'$. $|B_i| = |B_i'|$. Then, because of disjoint, $m(E) = \sum_{i=1}^n |B_i| = \sum_{i=1}^n |B_i'| = m(E')$. $\square$

1.2 Jordan Measure

More complicated sets such as triangle, disk, or rotated boxes can be measured by approaching from outside and inside by elementary sets.

Definition 1.2.1. *(Jordan Measure). Let $E \subset \mathbb{R}^d$ be a bounded set.*

Its Jordan inner measure is $m_J(E) = \sup_{A \subset E,\ A \in \mathcal{E}(\mathbb{R}^d)} m(A)$.

Its Jordan outer measure is $m^J(E) = \inf_{E \subset B,\ B \in \mathcal{E}(\mathbb{R}^d)} m(B)$.

If $m_J(E) = m^J(E)$, then E is Jordan measurable.

Let $\mathcal{J}(\mathbb{R}^d)$ be class of Jordan measurable sets. For $E \in \mathcal{J}(\mathbb{R}^d)$, define Jordan measure as $m(E) := m_J(E) = m^J(E)$.

Note that (i). unbounded sets are not Jordan measurable. (ii). The Jordan measure of $E \in \mathcal{E}(\mathbb{R}^d)$ is equal to the elementary measure of E.

Here we give the characterization of Jordan Measure.

Proposition 1.2.1. *(Characterisation of Jordan measurability). Let $E \in \mathbb{R}^d$ be bounded. Then the following are equivalent: (TFAE):*
(1). $E \in \mathcal{J}(\mathbb{R}^d)$.
(2). $\forall \epsilon > 0$, $\exists A, B \in \mathcal{E}(\mathbb{R}^d)$ with $A \subset E \subset B$, such that $m(B \setminus A) < \epsilon$.
(3). $\forall \epsilon > 0$, $\exists A \in \mathcal{E}(\mathbb{R}^d)$ such that $m^J(A \Delta E) < \epsilon$.

Proof. (1) $\Rightarrow$ (2): Since E is Jordan measurable, $m_J(E) = m^J(E) = m(E)$. Then, $m(E) = \sup_{A \subset E,\ A \in \mathcal{E}(\mathbb{R}^d)} m(A) = \inf_{E \subset B,\ B \in \mathcal{E}(\mathbb{R}^d)} m(B)$. By definition, $\forall \epsilon > 0$, $\exists A' \subset E \subset B'$ such that $m(A') \geq m(E) - \frac{\epsilon}{2}$ and $m(B') \leq m(E) + \frac{\epsilon}{2}$. Since $A' \subset E \subset B'$, by finite additivity we have $m(B') = m(B' \cup A') = m((B' \setminus A') \cup A') = m(B' \setminus A') + m(A')$. Thus $m(B' \setminus A') = m(B') - m(A')$. By applying the two inequality we just got, $m(B' \setminus A') \leq \epsilon$.

(1) $\Rightarrow$ (3): Since E is Jordan measurable, $m(E) = m^J(E) = \inf_{E \subset A, A \in \mathcal{E}(\mathbb{R}^d)} m(A)$. Then $\forall \epsilon > 0$, $\exists A'$ such that $E \subset A'$ and $m(A') \leq m(E) + \epsilon$. Since $E \subset A'$, we have $m(A' \setminus E) = m(A') - m(E) \leq \epsilon$. Also we have $m(E \setminus A') = m(\emptyset) = 0$.
We also have $m^J(A' \Delta E) = \inf_{A' \Delta E \subset B, B \in \mathcal{E}(\mathbb{R}^d)} m(B)$. Take $B = A' \setminus E$ that we just found above. Clearly we have $A' \Delta E \subset A' \setminus E$. Then by the property of infimum, $m^J(A' \Delta E) \leq m(A' \setminus E) \leq \epsilon$. Thus we found an elementary set A' such that $m^J(A \Delta E) \leq \epsilon$.

(2) $\Rightarrow$ (1): Let $\epsilon > 0$ be arbitrary. Then, by (2), $\exists A, B \in \mathcal{E}(\mathbb{R}^d)$ and $A \subset E \subset B$, such that $m(B \setminus A) \leq \epsilon$. Since $m(B \setminus A) = m(B) - m(A)$, we have $m(B) \leq m(A) + \epsilon$. Since $\epsilon \geq 0$ is arbitrary, there exists A, B that satisfies all the above conditions and $m(B) \leq m(A)$.
Since A is elementary and $m(A)$ is elementary measure, we have $m(A)$ less than its least upper bound, that is,

$$m(A) \leq \sup_{A' \subset E, A' \in \mathcal{E}(\mathbb{R}^d)} m(A')$$

Also, since B is elementary and $m(B)$ is elementary measure, we have $m(B)$ less than its least upper bound, that is,

$$m(B) \geq \inf_{E \subset B', B' \in \mathcal{E}(\mathbb{R}^d)} m(B')$$

Combining $m(B) \leq m(A)$, we have

$$\inf_{E \subset B', \in \mathcal{E}(\mathbb{R}^d)} m(B') \leq \sup_{A' \subset E, \in \mathcal{E}(\mathbb{R}^d)} m(A')$$

Since for all elementary sets A' and B' such that $A' \subset E \subset B'$, by monotonicity of elementary measure, $m(A') \leq m(B')$. Then by the definition of supremum and infimum,

$$\sup_{A' \subset E, A' \in \mathcal{E}(\mathbb{R}^d)} m(A') \leq \inf_{E \subset B', B' \in \mathcal{E}(\mathbb{R}^d)} m(B')$$

Then we have

$$\sup_{A' \subset E, A' \in \mathcal{E}(\mathbb{R}^d)} m(A') = \inf_{E \subset B', B' \in \mathcal{E}(\mathbb{R}^d)} m(B')$$

And therefore E is Jordan measurable.

(3) $\Rightarrow$ (2): We know that $\forall \epsilon > 0$, $\exists A \in \mathcal{E}(\mathbb{R}^d)$ such that

$$m^J(A \Delta E) = \inf_{A \Delta E \subset C, C \in \mathcal{E}(\mathbb{R}^d)} m(C) \leq \epsilon$$

Notice that, by subdividing and regrouping the almost disjoint boxes that consists of the elementary sets we just encountered, C can be written as $C = D \setminus F$ where $F, D \in \mathcal{E}(\mathbb{R}^d)$, $F \subset E \subset D$, $F \subset A \subset D$. To see that $A \Delta E \subset C$ still holds, note that $A \Delta E = (A \setminus E) \cup (E \setminus A) \subset (D \setminus E) \cup (E \setminus F) \subset (D \setminus F) \cup (D \setminus F) = D \setminus F = C$.
From the definition of infimum and the inequality above, $\exists B' \in \mathcal{E}(\mathbb{R}^d)$ such that $A \Delta E \subset B'$ and $m(B') < m^J(A \Delta E) + \epsilon \leq 2\epsilon$. Also we have known that B' can be written as $B' = D' \setminus F'$ where $F', D' \in \mathcal{E}(\mathbb{R}^d)$, $F' \subset E \subset D'$, $F' \subset A \subset D'$. Therefore we have proved that (3) $\Rightarrow$ (2).

Since (1) $\Rightarrow$ (2), (2) $\Rightarrow$ (1), (1) $\Rightarrow$ (3), (3) $\Rightarrow$ (2), we have proved that these are equivalent. $\square$

Theorem 1.2.1. *(Regions under graphs are Jordan measurable). Let B be a closed box in $\mathbb{R}^d$, and let $f : B \to \mathbb{R}$ be a continuous function.*

1. *The graph $\{(x, f(x)) : x \in B\} \subset \mathbb{R}^{d+1}$ is Jordan measurable in $\mathbb{R}^{d+1}$ with Jordan measure zero. (Hint: on a compact metric space, continuous functions are uniformly continuous.)*

2. *The set $\{(x, t) : x \in B; 0 \leq t \leq f(x)\} \subset \mathbb{R}^{d+1}$ is Jordan measurable.*

Proof. We prove 1 first and then 2.

(1). Let $\epsilon > 0$ be arbitrary. Since f is continuous on the closed box B on a compact metric space, it is bounded and uniformly continuous. $\exists \delta > 0$, such that, $|x - c| \leq \delta \Rightarrow |f(x) - f(c)| \leq \epsilon$ for all $c, x \in B$.
Evenly subdivide B into n almost disjoint boxes $B = \bigcup_{i=1}^n B_i$, so that within each the euclidean distance between two points is less than δ (we can do this because B is compact and there exists a finite covering), that is, $|x_i - c_i| \leq \delta$ for $x_i, c_i \in B_i$. Then we have $|f(x_i) - f(c_i)| \leq \epsilon$. Then, $|B_i| = \frac{|B|}{n}$. Then, within each box B_i,

$$\{(x, f(x)) | x \in B_i\} \subset B_i \times [\min_{x \in B_i} f(x), \max_{x \in B_i} f(x)]$$

$$|B_i \times [\min_{x \in B_i} f(x), \max_{x \in B_i} f(x)]| \leq \frac{|B|}{n} \epsilon$$

Then,

$$\begin{aligned}\{(x,f(x))|x\in B\} &= \bigcup_{i=1}^{n}\{(x,f(x))|x\in B_i\}\\ &\subset \bigcup_{i=1}^{n} B_i\times[\min_{x\in B_i} f(x),\max_{x\in B_i} f(x)]\end{aligned}$$

Since all of those $\mathbb{R}^{d+1}$ boxes are disjoint,

$$\begin{aligned}m^J(\{(x,f(x))|x\in B\}) &= \inf m(\bigcup_{i=1}^{n} B_i\times[\min_{x\in B_i} f(x),\max_{x\in B_i} f(x)])\\ &\le m(\bigcup_{i=1}^{n} B_i\times[\min_{x\in B_i} f(x),\max_{x\in B_i} f(x)])\\ &= \sum_{i=1}^{n} m(B_i\times[\min_{x\in B_i} f(x),\max_{x\in B_i} f(x)])\\ &= \sum_{i=1}^{n} |B_i\times[\min_{x\in B_i} f(x),\max_{x\in B_i} f(x)]|\\ &= |B|\epsilon\end{aligned}$$

Since $\epsilon>0$ is arbitrary, we have $m^J(\{(x,f(x))|x\in B\})=0$.
Also since

$$\begin{aligned}m_J(\{(x,f(x))|x\in B\}) &= \sup_{A\subset\{(x,f(x))|x\in B\},A\in\mathcal{E}(\mathbb{R}^d)} m(A)\\ &= m(\emptyset)\\ &= 0\end{aligned}$$

We have that the set is Jordan measurable with measure zero.

(2). In the proof of this, we will use the conclusion from Proposition 1.2.1.
Let $\epsilon>0$ be arbitrary. Since f is continuous on the closed box B on a compact metric space, it is bounded and uniformly continuous. $\exists\delta>0$, such that, $|x-c|\le\delta\Rightarrow|f(x)-f(c)|\le\epsilon$ for all $x,c\in B$.
Evenly subdivide B into n almost disjoint boxes $B=\bigcup_{i=1}^{n}B_i$, so that within each the euclidean distance between two points is less than δ, that is, $|x_i-c_i|\le\delta$ for $x_i,c_i\in B_i$. Then we have $|f(x_i)-f(c_i)|\le\epsilon$. Then, $|B_i|=\frac{|B|}{n}$.
Let A,B be two elementary sets such that

$$A=\bigcup_{i=1}^{n} B_i\times[0,\min_{x\in B_i} f(x)]$$

$$C=\bigcup_{i=1}^{n} B_i\times[0,\max_{x\in B_i} f(x)]$$

Clearly, sub-boxes of A are almost disjoint, sub-boxes of C are almost disjoint, and $A\subset$

$\{(x,t)|x \in B, 0 \leq t \leq f(x)\} \subset C$. Then,

$$A \setminus C = \bigcup_{i=1}^{n} B_i \times [\min_{x\in B_i} f(x), \max_{x\in B_i} f(x)]$$

By finite additivity, we have

$$\begin{aligned} m(A \setminus C) &= m(\bigcup_{i=1}^{n} B_i \times [\min_{x\in B_i} f(x), \max_{x\in B_i} f(x)]) \\ &= \sum_{i=1}^{n} m(B_i \times [\min_{x\in B_i} f(x), \max_{x\in B_i} f(x)]) \\ &= \sum_{i=1}^{n} |B_i|| \max_{x\in B_i} f(x) - \min_{x\in B_i} f(x)| \\ &\leq \sum_{i=1}^{n} \frac{|B|}{n}\epsilon \\ &= |B|\epsilon \end{aligned}$$

Since $\epsilon > 0$ arbitrary, $A \subset \{(x,t)|x \in B, 0 \leq t \leq f(x)\} \subset C$, we have $\{(x,t)|x \in B, 0 \leq t \leq f(x)\}$ to be Jordan measurable. □

Proposition 1.2.2. *Let E, $F \in \mathcal{J}(\mathbb{R}^d)$, then,*
(1). $E \cup F$, $E \cap F$, $E \setminus F$, $E \Delta F \in \mathcal{J}(\mathbb{R}^d)$.
(2). $m(E) \geq 0$.
(3). Finite Additivity: If $E \cap F = \emptyset$, then $m(E \cup F) = m(E) + m(F)$.
(4). Monotonicity: If $E \subset F$, then $m(E) \leq m(F)$.
(5). Finite subadditivity: $m(E \cup F) \leq m(E) + m(F)$.
(6). Translation Invariabce: $m(E + x) = m(E)$.

Proof. $E \cap F \in \mathcal{J}(\mathbb{R}^d)$:
Using (3) from Proposition 1.2.1, let $A, B, C, D \in \mathcal{E}(\mathbb{R}^d)$ with $A \subset E \subset B$, $C \subset F \subset D$, such that $m^J(B \setminus A) < \epsilon$, $m(D \setminus C) < \epsilon$. Then we claim that $m((B \cap D) \setminus (A \cap C)) \leq 2\epsilon$. Indeed,

$$\begin{aligned} (B \cap D) \setminus (A \cap C) &= (B \cap D) \cap (A \cap C)^C \\ &= B \cap D \cap (A^C \cup C^C) \\ &= (B \cap D \cap A^C) \cup (B \cap D \cap C^C) \\ &\subset (B \cap A^C) \cup (D \cap C^C) \\ &= (B \setminus A) \cup (D \setminus C) \end{aligned}$$

Then, $m((B \cap D) \setminus (A \cap C)) \leq m((B \setminus A) \cup (D \setminus C)) \leq 2\epsilon$.
Since $A \cap C \subset E \cap F \subset B \cap D$, we have $E \cap F$ is also Jordan measurable. □

Theorem 1.2.2. *(Closure, interior, and topolitical boundary). Let $E \subset \mathbb{R}^d$ be a bounded set.*

1. *E and the closure $\overline{E}$ of E have the same Jordan outer measure.*

2. *E and the interior E° of E have the same Jordan inner measure.*

3. *E is Jordan measurable if and only if the topological boundary ∂E of E has Jordan outer measure zero.*

4. *The bullet-riddled square $[0,1]^2 \setminus \mathbb{Q}^2$, and set of bullets $[0,1]^2 \cap \mathbb{Q}^2$, both have Jordan inner measure zero and Jordan outer measure one. In particular, both sets are not Jordan measurable.*

Proof. (1). First, $E \subseteq \overline{E}$, which means that $m^J(E) \leq m^J(\overline{E})$. Thus we only need to prove that $m^J(E) \geq m^J(\overline{E})$. Since

$$m^J(E) = \inf_{E \subset B, B \in \mathcal{E}(\mathbb{R}^d)} m(B)$$

For $\epsilon > 0$ there exists an elementary set B that covers E and write it as a finite union of almost disjoint boxes $B = \bigcup_{i=1}^n B_i$, such that

$$\sum_{i=1}^n m(B_i) \leq m^J(E) + \epsilon$$

Since $\overline{E} \subseteq \overline{\bigcup_{i=1}^n B_i} \subseteq \bigcup_{i=1}^n \overline{B_i}$, we have $m^J(\overline{E}) \leq \sum_{i=1}^n m(\overline{B_i}) = \sum_{i=1}^n m(B_i) \leq m^J(E) + \epsilon$. Since $\epsilon > 0$ is arbitrary, we have $m^J(E) \geq m^J(\overline{E})$.

(2). First, $E^o \subseteq E$, which means that $m_J(E^o) \leq m_J(E)$. Thus we only need to prove that $m_J(E^o) \geq m_J(E)$. Since

$$m_J(E) = \sup_{B \subset E, B \in \mathcal{E}(\mathbb{R}^d)} m(B)$$

For $\epsilon > 0$ there exists an elementary set B that is covered by E and write it as a finite union of almost disjoint boxes $B = \bigcup_{i=1}^n B_i$, such that

$$\sum_{i=1}^n m(B_i) \geq m_J(E) - \epsilon$$

Since $\bigcup_{i=1}^n B_i^o \subseteq (\bigcup_{i=1}^n B_i)^o \subseteq E^o$, we have $m_J(E^o) \geq \sum_{i=1}^n m(B_i^o) = \sum_{i=1}^n m(B_i) \geq m_J(E) - \epsilon$. Since $\epsilon > 0$ is arbitrary, we have $m_J(E) \leq m_J(E^o)$.

(3). ($\Rightarrow$). Since E is Jordan measurable, let $\epsilon > 0$, then there exists $E, F \in \mathcal{E}(\mathbb{R}^d)$ such that $m(E \setminus F) \leq \epsilon$. Also we know that $E \setminus F \in \mathcal{E}(\mathbb{R}^d)$. Thus

$$\begin{aligned} m^J(\partial E) &= \inf_{\partial E \subset B, B \in \mathcal{E}(\mathbb{R}^d)} m(B) \\ &\leq m(E \setminus F) \\ &\leq \epsilon \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have $m^J(E) = 0$.

($\Leftarrow$). Notice that $\overline{E} \setminus E \subset \overline{E} \setminus E^o = \partial E$. Since $m^J(\partial E) = 0$, we have $m^J(\overline{E} \setminus E) = 0$. By definition

$$m^J(\overline{E} \setminus E) = \inf_{\overline{E} \setminus E \subset B, B \in \mathcal{E}(\mathbb{R}^d)} m(B)$$

we have $\forall \epsilon > 0$, $\exists B'$ elementary, $\overline{E} \setminus E \subset B'$ and $m^J(B') \leq \epsilon$.
Notice that

$$E \subset (\overline{E} \setminus E) \cup E \subset B' \cup E$$

and that $B' \cup E$ is an elementary set. Then,

$$\begin{aligned} m^J((B' \cup E)\Delta E) &= m^J(((B' \cup E) \setminus E) \cup (E \setminus (B' \cup E))) \\ &= m^J((B' \cup E) \setminus E) \\ &= m^J(B') \\ &\leq \epsilon \end{aligned}$$

Therefore, we found an elementary set $B' \cup E$ such that its symmetric difference with E has Jordan outer measure less or equal than ϵ. Then, E is Jordan measurable.

(4). We use the property that the rational numbers are dense in $\mathbb{R}$, thus the rational pairs are also dense in $\mathbb{R}^2$.
For $A = [0,1]^2 \setminus \mathcal{Q}^2$, we have $m_J(A) = \sup_{B \subset A, B \in \mathcal{E}(\mathbb{R}^d)} m(B)$. But since the rational pairs are dense, every non-empty elementary set $B \subset [0,1]^2$ contains rational pairs and thus is not contained in A. Thus, $m_J(A) = m(\emptyset) = 0$. $m^J(A) = \inf_{A \subset C, C \in \mathcal{E}(\mathbb{R}^d)} m(C)$. This value takes infimum when $C = [0,1]^2$, so $m^J(A) = 1$.
For $D = [0,1]^2 \cap \mathcal{Q}^2$, we have we have $m_J(D) = \sup_{E \subset A, E \in \mathcal{E}(\mathbb{R}^d)} m(E)$. But since $\mathbb{R}^2$ without rational pairs are also dense, every non-empty elementary set $E \subset [0,1]^2$ contains non-rational pairs and thus is not contained in D. Thus, $m_J(D) = m(\emptyset) = 0$. $m^J(D) = \inf_{D \subset F, F \in \mathcal{E}(\mathbb{R}^d)} m(F)$. This value takes infimum when $F = [0,1]^2$, so $m^J(D) = 1$.

□

Theorem 1.2.3. *(Equivalence of Riemann integral and Darboux integral). Let $[a,b]$ be an interval, and $f : [a,b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if it is Darboux integrable, in which case the Riemann integral and Darboux integrals are equal.*

Proof. Denote the set

$$E = \{(x,t) \mid x \in I, 0 \leq t \leq f(x)\}$$

First we prove that f is *Darboux Integrable* implies that the set E is Jordan measurable. From definition we know that

$$\overline{\int_a^b} f(x)dx = \inf_{E \subset B', B' \in \mathcal{E}(\mathbb{R}^d)} m(B')$$

Where B' is a collection of almost disjoint boxes partitioned by almost disjoin intervals $I = \bigcup_{i=1}^n I_i$.

$$B' = \bigcup_{i=1}^n \{(x_i,t) | x_i \in I_i, 0 \leq t \leq h(x_i), h(x_i) = d_i \geq f(x_i) \ \forall x_i \in I_i\}$$

From this we know that $E \subset B'$. Also, $\forall \epsilon > 0$, $\exists B$ satisfying the conditions above such that

$$m(B) \leq \overline{\int_a^b} f(x)dx + \epsilon$$

From the definition we also know that

$$\underline{\int_a^b} f(x)dx = \sup_{A' \subset E, A' \ elementary} m(A')$$

Where A' is a collection of almost disjoint boxes partitioned by almost disjoin intervals $I = \bigcup_{i=1}^{m} I_i'$.

$$A' = \bigcup_{i=1}^{m} \{(x_i, t) | x_i \in I_i', 0 \leq t \leq g(x_i), g(x_i) = c_i \leq f(x_i) \ \forall x_i \in I_i'\}$$

From this we know that $A' \subset E$. Also, $\forall \epsilon > 0$, $\exists A$ satisfying the conditions above such that

$$m(A) \geq \underline{\int_a^b} f(x)dx - \epsilon$$

Clearly, $A \subset E \subset B$. Since f is Darboux integrable, we have

$$\begin{aligned} m(B \setminus A) &= m(B) - m(A) \\ &= (\overline{\int_a^b} f(x)dx + \epsilon) - (\underline{\int_a^b} f(x)dx - \epsilon) \\ &= 2\epsilon \end{aligned}$$

Thus, we have shown that *Darboux Integrable* means that (2) in Proposition 1.2.1. has been satisfied. Thus, E is Jordan measurable. To deal with *Riemann Integrable*, we first collect the intervals that we just partitioned above $\{I_i\}_{i=1}^n$ and $\{I_i'\}_{i=1}^m$. Takes the endpoints of those intervals, order them and get a new (finer) subdivision of the interval $I = \bigcup_{i=1}^{n+m-3} I_i''$ that consists of $n + m - 3$ almost disjoint sub-intervals (this is clear to see). Then, take the set

$$F = \bigcup_{i=1}^{n+m-3} \{(x_i, t) | x_i \in I_i'', 0 \leq t \leq f(x_i^*) \ for\ some\ x_i^* \in I_i''\}$$

From its construction, it is clear to see that $A \subset F \subset B$. Then, $(E \setminus F) \subset (B \setminus A)$, $(F \setminus E) \subset (B \setminus A)$. Then, for $\epsilon > 0$, we have found an elementary set F such that $m^J(E \Delta F) \leq m(B \setminus A) \leq 2\epsilon$ by definition of Jordan outer measure and the monotonicity of elementary measure. Thus E is Jordan measurable

Notice that the Riemann sum of f on $I = [a, b]$ is just the elementary measure of F with partitions $\mathcal{P} : I = \bigcup_{i=1}^{n+m-3} I_i''$, and it equals its outer Jordan measure:

$$m^J(F) = m(F) = \sum_{i=1}^{n+m-3} f(x_i^*)|I_i''| = \mathcal{R}(f, \mathcal{P})$$

$$
\begin{aligned}
|m^J(F) - m^J(E)| &= |m^J(F) + m^J(E \setminus F) - (m^J(E \setminus F) + m^J(E))| \\
&= |m^J(F \cup E) - (m^J(E) + m^J(F \setminus E) - m^J(F \setminus E) + m^J(E \setminus F))| \\
&= |m^J(F \cup E) - (m^J(F \cup E) - m^J(F \setminus E) + m^J(E \setminus F))| \\
&= |m^J(F \setminus E) - m^J(E \setminus F)| \\
&\leq m^J(F \setminus E) + m^J(E \setminus F) \\
&= m^J(E \Delta F)
\end{aligned}
$$

Since we just proved that for $\epsilon > 0$ we can always find a set F such that $m^J(E\Delta F) \leq 2\epsilon$, and we established that $m^J(F) = \mathcal{R}(f, \mathcal{P})$, using the above inequality, we have $|\mathcal{R}(f, \mathcal{P}) - m^J(E)| \leq \epsilon$. Then, f is Riemann integrable when the set E is Jordan measurable.
Thus, we established the equivalence among Jordan measurability, Darboux integrability, and Riemann integrability. □

There are sets that are not Jordan measurable. For example, $E = [0,1] \cap \mathcal{Q}$ is not Jordan measurable. This is because E contains no open interval, any elementary set $A \in E$ can only be a finite union of singletons, therefore, $m_J(E) = 0$. However, $m^J(E) = 1$.

Similarly, there exists open and bounded sets that are not Jordan measurable, and there exists compact sets that are not Jordan measurable. So, we need Lebesgue measure.

1.3 Lebesgue Measure

Definition 1.3.1. *The Lebesgue outer measure (exterior measure) of* $E \subset \mathbb{R}^d$ *is*

$$m^*(E) := \inf\{\sum_{i=1}^{\infty} |B_i| : B_1, B_2, ... boxes, E \subset \bigcup_{i=1}^{\infty} B_i\}$$

Note that $m^*(E) \leq m^J(E)$ since it is the infimum over a bigger set.

For $E = [0,1] \cap \mathcal{Q} = \{q_1, q_2, ...\}$, by taking boxes $B_i = \{q_i\}$, we get the Lebesgue measure $m^*(E) \leq \sum_{i=1}^{\infty} |B_i| = \sum_{i=1}^{\infty} 0 = 0$.

Similarly, $m^*(E) = 0$ for any countable E, just take $\{x_i\}$ as boxes. Or, take boxes $B_i = (q_i - \frac{\epsilon}{2^i}, q_i + \frac{\epsilon}{2^i})$, then $m^*(E) \leq m(\bigcup_{i=1}^{\infty} B_i) \leq \sum_{i=1}^{\infty} |B_i| = 2\epsilon$. Since ϵ is arbitrary, we have $m^*(E) = 0$.

Now we gives the properties of the Lebesgue outer measure (the outer measure axioms).

1.3.1 Properties of Lebesgue Outer Measure

Proposition 1.3.1. *(The outer measure axioms).*

1. $m^*(\emptyset) = 0$.
2. *(Monotonicity). If* $E \subset F \subseteq \mathbb{R}^d$, *then* $m^*(E) \leq m^*(F)$.
3. *(σ-subadditivity). if* $E_1, E_2, ... \subset \mathbb{R}^d$ *is a countable sequence of sets, then,* $m^*(\bigcup_{n=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} m^*(E_i)$.

Proof. (3).
We know that,

$$m^*(E_i) = \inf\{\sum_{j=1}^{\infty} |B_j^i| : B_1^i, B_2^i, ... boxes, E_i \subset \bigcup_{j=1}^{\infty} B_j^i\}$$

Then, for each i, there exists $B_1^i, B_2^i, ...$ to be boxes, such that $E_i \subset \bigcup_{j=1}^{\infty} B_j^i$ and

$$\sum_{j=1}^{\infty} |B_j^i| \leq m^*(E_i) + \frac{\epsilon}{2^i}$$

Since $E_i \subset \bigcup_{j=1}^{\infty} B_j^i$, we have $\bigcup_{i=1}^{\infty} E_i \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} B_j^i$, then we have

$$\begin{aligned}
m^*(\bigcup_{i=1}^{\infty} E_i) &\leq m^*(\bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} B_j^i) \text{ by monotonicity} \\
&\leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |B_j^i| \text{ by definition of Lebesgue measure: infimum} \\
&= \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} |B_j^i|) \text{ by Tonelli's Theorem for series} \\
&= \sum_{i=1}^{\infty} (m^*(E_i) + \frac{\epsilon}{2^i}) \\
&= \sum_{i=1}^{\infty} m^*(E_i) + \epsilon
\end{aligned}$$

Since ϵ is arbitrary, we then have $m^*(\bigcup_{n=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} m^*(E_i)$.
Also, (3) and (1) together give the finite subadditivity property by just let $\emptyset = E_{k+1} = E_{k+2} = ...$, that is, $m^*(\bigcup_{i=1}^{k} E_i) \leq \sum_{i=1}^{k} m^*(E_i)$. ☐

Theorem 1.3.1. *(Distance of sets). Let $E, F \subset \mathbb{R}^d$ be disjoint closed sets, with at least one of E, F being compact. Then $dist(E, F) > 0$.*

Proof. Suppose, on the contrary, $d(E, F) = 0$. That is, $\forall \epsilon > 0$, $\exists x \in E$ and $y \in F$, such that $d(x, y) < \epsilon$.
Then we can construct sequences by axiom of countable choice, $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ such that $\forall n \in \mathbb{N}$, $d(x_n, y_n) \leq \frac{1}{n}$.
Suppose E is compact, then by Bolzano-Weierstrass Theorem (or, in a metric space, compactness is equivalent to sequential compactness), $\{x_n\}_{n \in \mathbb{N}}$ has a convergent subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ that converges to x_0. x_0 is a limit point of E. Since E is closed, $x_0 \in E$. Then $\forall \epsilon > 0$, $\exists k \in \mathbb{N}$ such that $d(x_0, x_{n_k}) \leq \frac{1}{n_k} \leq \frac{\epsilon}{2}$. Also, by the previous construction we know that $d(x_{n_k}, y_{n_k}) \leq \frac{\epsilon}{2}$. Then,

$$d(x_0, y_{n_k}) \leq d(x_0, x_{n_k}) + d(x_{n_k}, y_{n_k}) < \epsilon$$

which means that $y_{n_k} \to x_0$. x_0 is a limit point of F. Also by closedness of F, $x_0 \in F$.
Contradiction to $E \cap F = \emptyset$. Therefore $d(E, F) > 0$.

☐

Lemma 1.3.1. *(Finite additivity for separated sets). Let $E, F \subset \mathbb{R}^d$ be such that $dist(E, F) := \inf\{|x - y| : x \in E, y \in F\} > 0$, then $m^*(E \cup F) = m^*(E) + m^*(F)$.*

Proof. First, we prove $\leq$. This is natual from σ-additivity: $m^*(E \cup F) \leq m^*(E) + m^*(F)$.
Next, we prove $\geq$. Without loss of generality, assume that $m^*(E \cup F) < +\infty$.
Let $\epsilon > 0$ be arbitrary, then, by definition there exists a countable collection of boxes $B_1, B_2, ...$ such that,

$$E \cup F \subset \bigcup_{i=1}^{\infty} B_i$$

$$\sum_{i=1}^{\infty} |B_i| \leq m^*(E \cup F) + \epsilon$$

Fix $\delta \in (0, dist(E,F))$. By subdividing these boxes into finer boxes B_i', we may assume that $diam(B_i') < \delta$. Then, some of these boxes have intersection with E while others have intersection with F.
Let $I = \{i : B_i' \cap E \neq \emptyset\}$, $J = \{j : B_j' \cap F \neq \emptyset\}$. Then $B_i' \cap B_j' = \emptyset$ cause otherwise we would have a box with diameter bigger than δ.
Then, $m^*(E) \leq \sum_{i \in I} |B_i'|$, $m^*(F) \leq \sum_{j \in J} |B_j'|$.

$$\begin{aligned} m^*(E) + m^*(F) &\leq \sum_{i \in I \cup J} |B_i'| \\ &\leq \sum_{i=1}^{\infty} |B_i| \\ &\leq m^*(E \cup F) + \epsilon \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have $m^*(E) + m^*(F) \leq m^*(E \cup F)$.
Now that we have $\leq$ and $\geq$, we have $=$. □

Lemma 1.3.2. *(Outer measurability for elementary sets). Let $E \in \mathcal{E}(\mathbb{R}^d)$, then $m^*(E) = m(E)$, the elementary measure.*

Proof. First we prove $\leq$. We already know that $m^*(E) \leq m^J(E) = m(E)$, thus $\leq$ holds.
Next we prove $\geq$.
Consider first the case where E is closed, then, E is compact. Then we can use the Heine-Borel Theorem which states that any covering of a compact set by a collection of open sets contains a finite subcovering.
Take a covering of E by boxes: $E \subset \bigcup_{i=1}^{\infty} B_i$ such that $\sum_{i=1}^{\infty} |B_i| \leq m^*(E) + \epsilon$. For each box B_i, find an open box B_i' such that $B_i \subset B_i'$ and $|B_i'| \leq |B_i| + \frac{\epsilon}{2^i}$.
Then, $\sum_{i=1}^{\infty} |B_i'| \leq \sum_{i=1}^{\infty} |B_i| + \epsilon \leq m^*(E) + 2\epsilon$.
Using Heine-Borel Theory, there is a finite N such that $E \subset \bigcup_{i=1}^{N} B_i'$. Then,

$$\begin{aligned} m(E) &\leq \sum_{i=1}^{N} |B_i'| \\ &\leq \sum_{i=1}^{\infty} |B_i'| \\ &\leq m^*(E) + 2\epsilon \end{aligned}$$

Since $\epsilon > 0$ arbitrary, we have $\geq$.
Now consider the case where E is not closed. Then, write E as a finite union of disjoint boxes $E = \bigcup_{i=1}^{k} Q_i$, which need not be closed.
Let $\epsilon > 0$ be arbitrary, and for each $j \in \{1, ..., k\}$, find a closed sub-box $Q_j' \subset Q_j$ such that $|Q_j'| \geq |Q_j| - \frac{\epsilon}{2^k}$. Then, by the previous discussion and finite additivity of elementary

measure, we have

$$\begin{aligned} m^*(\bigcup_{j=1}^k Q'_j) &= m(\bigcup_{j=1}^k Q'_j) \\ &= \sum_{j=1}^k m(Q'_j) \\ &\geq \sum_{j=1}^k m(Q_j) - \epsilon \\ &= m(E) - \epsilon \end{aligned}$$

Also, $\bigcup_{j=1}^k Q'_j \subset E$, so by monotonicity, we have

$$\begin{aligned} m^*(E) &\geq m^*(\bigcup_{j=1}^k Q'_j) \\ &\geq m(E) - \epsilon \end{aligned}$$

Then $m(E) \leq m^*(E) + \epsilon$. Since $\epsilon > 0$ arbitrary, $\geq$ holds. $\square$

Lemma 1.3.3. *(Outer measure of countable unions of almost disjoint boxes). Let* $E = \bigcup_{i=1}^\infty B_i$ *be a countable union of almost disjoint boxes, then* $m^*(E) = \sum_{i=1}^\infty |B_i|$. *Almost disjoint means that* $B_i^o \cap B_j^o = \emptyset \ \forall i \neq j$ *(topological interior doesn't intersect).*

Proof. From countable sub-additivity and Lemma 1.3.2.,

$$m^*(E) \leq \sum_{i=1}^\infty m^*(B_i) = \sum_{i=1}^\infty |B_i|$$

Therefore, it suffices to show that

$$m^*(E) \geq \sum_{i=1}^\infty |B_i|$$

Notice that for each $N \in \mathbb{N}$,

$$E \supset \bigcup_{i=1}^N B_i$$

Then,

$$\begin{aligned} m^*(E) &\geq m^*(\bigcup_{i=1}^N B_i) \\ &= m(\bigcup_{i=1}^N B_i) \\ &= \sum_{i=1}^N |B_i| \end{aligned}$$

Let $N \to \infty$, we have $m^*(E) \geq \sum_{i=1}^\infty |B_i|$ Therefore we conclude the proof. $\square$

From this lemma we have a corollary.

Corollary 1.3.1. *If* $E = \bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} B_i'$, $(B_i)_{i\in\mathbb{N}}$ *and* $(B_i')_{i\in\mathbb{N}}$ *are almost disjoint boxes, then* $\sum_{i=1}^{\infty} |B_i| = \sum_{i=1}^{\infty} |B_i'|$.

Lemma 1.3.4. *An open set* $U \subseteq \mathbb{R}^d$ *is the countable union of almost disjoint boxes. (in fact, the countable union of almost disjoint closed boxes).*

Proof. For $n \in \mathbb{Z}$, let $\mathcal{Q}_n$ be the collection of all closed cubes of the form

$$[\frac{k_1}{2^n}, \frac{k_1+1}{2^n}] \times ... \times [\frac{k_d}{2^n}, \frac{k_d+1}{2^n}] \text{ for some integers } k_1, ..., k_d$$

Define $\mathcal{Q}_{\geq 0} := \bigcup_{n=1}^{\infty} \mathcal{Q}_n$ to be the union of all dyadic cubes of side length ≤ 1. Notice that $\mathcal{Q}_{\geq 0}$ has a tree structure, that is, for each $Q \in \mathcal{Q}_n$, $\exists! Q' \in \mathcal{Q}_{n-1}$ such that $Q \subset Q'$.
Given these, we have the dyadic nesting property: $\forall Q_1, Q_2 \in \mathcal{Q}_{\geq 0}$ with $Q_1^o \cap Q_2^o \neq \emptyset$, either $Q_1 \subseteq Q_2$ or $Q_2 \subseteq Q_1$.
Since U is open, $\forall x \in U$, $\exists$ open ball $B(x,r) \subset U$. Therefore, $\exists$ closed $Q \in \mathcal{Q}_{\geq 0}$ such that $x \in Q \subseteq U$. Then, let $\mathcal{Q}_U = \{Q \in \mathcal{Q}_{\geq 0} : Q \subseteq U\}$. Then,

$$U = \bigcup_{Q \in \mathcal{Q}_U} Q \text{ with } \mathcal{Q}_U \text{ being countable}$$

To get almost disjoint subcollection, take $\mathcal{Q}_U^* \subseteq \mathcal{Q}_U$ to be a subcollection of maximal elements with respect to set inclusion, which means that they are not contained in any other cube in $\mathcal{Q}_U$.

$$\mathcal{Q}_U^* := \{Q \in \mathcal{Q}_{\geq 0} : Q \subseteq U, \nexists Q' \text{ such that } Q \in \mathcal{Q}_{\geq 0} \text{ and } Q \supset Q'\}$$

First we see that if $Q \subseteq U$ then $Q \subseteq \mathcal{Q}_U$, then $\mathcal{Q}_U \subseteq \mathcal{Q}_U^*$. Together with the definition of $\mathcal{Q}_U^*$, we see that $\mathcal{Q}_U^* = \mathcal{Q}_U$. Second, by dyadic nesting property, every cube in $\mathcal{Q}$ is contained in exactly one maximal cube in $\mathcal{Q}^*$, and that any two such maximal cubes in $\mathcal{Q}^*$ are almost disjoint. Thus, $U = \bigcup_{Q \in \mathcal{Q}_U^*} Q$ are almost disjoint, and also countable. □

Lemma 1.3.5. *(Outer regularity). For any* $E \subseteq \mathbb{R}^d$,

$$m^*(E) = \inf_{E \subset U, U \text{ } open} m^*(U)$$

Proof. ($\leq$): it is easy to see from monotonicity that $\forall U \supset E$, $m^*(E) \leq m^*(U)$, thus

$$m^*(E) \leq \inf_{E \subset U, U \text{ } open} m^*(U)$$

Therefore it suffices to prove that

$$m^*(E) \geq \inf_{E \subset U, U \text{ } open} m^*(U)$$

By definition of the outer Lebesgue measure,

$$m^*(E) = \inf\{\sum_{i=1}^{\infty} |B_i| : E \subset \bigcup_{i=1}^{\infty} B_i, B_1, B_2, ...boxes\}$$

Then, $\forall \epsilon > 0$, $\exists B_1', B_2'$... such that

$$\sum_{i=1}^{\infty} |B_i'| \leq m^*(E) + \epsilon$$

Enlarge each box B_i' is contained in an open box $B_i' \subset B_i''$ such that

$$|B_i''| \leq |B_i'| + \frac{\epsilon}{2^i}$$

Thus, $E \subset \bigcup_{i=1}^{\infty} B_i''$ where $\bigcup_{i=1}^{\infty} B_i''$ is open.

$$\begin{aligned}\sum_{i=1}^{\infty} |B_i''| &\leq \sum_{i=1}^{\infty} |B_i'| + \epsilon \\ &\leq m^*(E) + 2\epsilon\end{aligned}$$

Since $\bigcup_{i=1}^{\infty} B_i''$ is open, by countable sub-additivity and the definition of infimum,

$$\inf_{E \subset U, U \text{ open}} m^*(U) \leq m^*(\bigcup_{i=1}^{\infty} B_i'') \leq \sum_{i=1}^{\infty} |B_i''| \leq m^*(E) + 2\epsilon$$

Since $\epsilon > 0$ is arbitrary, we have

$$m^*(E) \geq \inf_{E \subset U, U \text{ open}} m^*(U)$$

$\square$

1.3.2 Lebesgue Measurability

Definition 1.3.2. *(Lebesgue Measurability). A set $E \subset \mathbb{R}^d$ is Lebesgue measurable if $\forall \epsilon > 0$, $\exists$ an open set $U \subset \mathbb{R}^d$, $E \subseteq U$, such that $m^*(U \setminus E) < \epsilon$.*

We denote the class of all Lebesgue measurable sets by $\mathcal{L}(\mathbb{R}^d)$. For $E \in \mathcal{L}(\mathbb{R}^d)$, its Lebesgue measure is $m(E) := m^*(E)$.

There are plenty of Lebesgue measurable sets, as we can see from the following proposition.

Proposition 1.3.2. *(Existence of Lebesgue measurable sets). Let $E \subseteq \mathbb{R}^d$, then $E \in \mathcal{L}(\mathbb{R}^d)$ if*

1. *E is open.*
2. *E is closed.*
3. *E is a null set, i.e. $m^*(E) = 0$.*
4. *$E = \emptyset$.*
5. *if $E \in \mathcal{L}(\mathbb{R}^d)$, then $\mathbb{R}^d \setminus E \in \mathcal{L}(\mathbb{R}^d)$.*
6. *$E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{L}(\mathbb{R}^d)$ where $E_i \in \mathcal{L}(\mathbb{R}^d)$.*
7. *$E = \bigcap_{i=1}^{\infty} E_i \in \mathcal{L}(\mathbb{R}^d)$ where $E_i \in \mathcal{L}(\mathbb{R}^d)$.*

Proof. (1) is immediate from definition. By Lemma 1.3.4., write E as $E = \bigcup_{i=1}^{\infty} B_i$ where B_i are disjoint boxes. Expand each B_i to be an open box $B_i' \supset B_i$ such that $\forall \epsilon > 0$,

$$|B_i'| \leq |B_i| + \frac{\epsilon}{2^i}$$

Then, by σ-additivity and Lemma 1.3.3,

$$m^*(\bigcup_{i=1}^{\infty} B_i') \leq \sum_{i=1}^{\infty} |B_i'| \leq m(E) + \epsilon$$

Therefore

$$m^*(\bigcup_{i=1}^{\infty} B_i' \setminus E) \leq \epsilon$$

Thus, we found an open set $\bigcup_{i=1}^{\infty} B_i' \supset E$, such that $m^*(\bigcup_{i=1}^{\infty} B_i' \setminus E) \leq \epsilon$.
(3) and (4) are immediate.
(6). Since $E_i \in \mathcal{L}(\mathbb{R}^d)$, $\exists$ open set E_i' such that $E_i \subset E_i'$, $m^*(E_i') \leq m^*(E_i) + \frac{\epsilon}{2^i}$. Then, $\bigcup_{i=1}^{\infty} E_i \subset \bigcup_{i=1}^{\infty} E_i'$ where $\bigcup_{i=1}^{\infty} E_i'$ is open. Since

$$\begin{aligned}
\bigcup_{i=1}^{\infty} E_i' \setminus \bigcup_{i=1}^{\infty} E_i &= \bigcup_{i=1}^{\infty} E_i' \cap (\bigcap_{i=1}^{\infty} E_i^C) \\
&= \bigcup_{i=1}^{\infty} (E_i' \cap (\bigcap_{i=1}^{\infty} E_i^C)) \\
&\subset \bigcup_{i=1}^{\infty} (E_i' \cap E_i^C) \\
&= \bigcup_{i=1}^{\infty} (E_i' \setminus E_i)
\end{aligned}$$

Then, by monotonicity and σ-additivity,

$$\begin{aligned}
m^*(\bigcup_{i=1}^{\infty} E_i' \setminus \bigcup_{i=1}^{\infty} E_i) &\leq m^*(\bigcup_{i=1}^{\infty} (E_i' \setminus E_i)) \\
&\leq \sum_{i=1}^{\infty} m^*(E_i' \setminus E_i) \\
&= \sum_{i=1}^{\infty} m^*(E_i') - m^*(E_i) \\
&\leq \epsilon \qquad (1)
\end{aligned}$$

Therefore, $\bigcup_{i=1}^{\infty} E_i \in \mathcal{L}(\mathbb{R}^d)$.
(2). First, we can express each closed set as $E = \bigcup_{i=1}^{\infty} E_n$ for E_n closed and bounded (for example, $E_n = \overline{B(0, n)} \cap E$ for $n = 1, 2, ...$). Then by (6), it suffices to verify the claim when E is closed and bounded, hense compact.
By Lemma 1.3.5., $\exists U \supset E$ open, such that $m^*(U) \leq m^*(E) + \epsilon$. Therefore, it suffices to show that $m^*(U \setminus E) \leq \epsilon$.
If w have finite additivity for m^*, then we have $m^*(U \setminus E) + m^*(E) = m^*(U) \leq m^*(E) + \epsilon$ and then $m^*(U \setminus E) \leq \epsilon$. But we don't have it, so we should instead fo the following.

Since $U \setminus E$ is also open, by Lemma 1.3.4, $U \setminus E = \bigcup_{i=1}^{\infty} Q_i$ where Q_i are almost disjoint closed boxes. Then by Lemma 1.3.3, $m^*(U \setminus E) = \sum_{i=1}^{\infty} |Q_i|$.
We truncate the sum: for any finite $N \in \mathbb{N}$, $\bigcup_{i=1}^{N} Q_i$ is closed and disjoint from E. From Theorem 1.3.1., since E is compact and $\bigcup_{i=1}^{N} Q_i$ is closed, we have $dist(E, \bigcup_{i=1}^{N} Q_i) > 0$. Then by Lemma 1.3.1,

$$\begin{aligned} m^*(\bigcup_{i=1}^{N} Q_i) + m^*(E) &= m^*(E \cup \bigcup_{i=1}^{N} Q_i) \\ &\leq m^*(U) \\ &\leq m^*(E) + \epsilon \end{aligned}$$

$$\sum_{i=1}^{N} |Q_i| = m^*(\bigcup_{i=1}^{N} Q_i) \leq \epsilon$$

Let $N \to \infty$,

$$m^*(U \setminus E) = \sum_{i=1}^{\infty} |Q_i| = m^*(\bigcup_{i=1}^{\infty} Q_i) \leq \epsilon$$

Therefore $E \in \mathcal{L}(\mathbb{R}^d)$.
(5). Since $E \in \mathcal{L}(\mathbb{R}^d)$, for every $n \in \mathbb{N}$, $\exists U_n \supset E$ open such that $m^*(U_n \setminus E) < \frac{1}{n}$.
Let $F_n := U_n^C$, then $(\mathbb{R}^d \setminus E) \supset F_n$ for all n. Since

$$(\mathbb{R}^d \setminus E) \setminus F_n = (\mathbb{R}^d \setminus E) \cap F_n^C = (\mathbb{R}^d \setminus E) \cap U_n = U_n \setminus E$$

we have

$$m^*((\mathbb{R}^d \setminus E) \setminus F_n) < \frac{1}{n}$$

Let $F := \bigcup_{i=1}^{\infty} F_n$, then $(\mathbb{R}^d \setminus E) \supset F$. From monotonicity, we have

$$m^*((\mathbb{R}^d \setminus E) \setminus F) \leq m^*((\mathbb{R}^d \setminus E) \setminus F_n) < \frac{1}{n} \ \forall n \in \mathbb{N}$$

Taking $n \to \infty$, we have $m^*((\mathbb{R}^d \setminus E) \setminus F) = 0$, thus $(\mathbb{R}^d \setminus E) \setminus F$ is a null set, and is Lebesgue measurable. Therefore, $\mathbb{R}^d \setminus E$ is the union of this null set and F. Since by definition $F = \bigcup_{i=1}^{\infty} U_n^C$ where U_n^C is closed, F is Lebesgue measurable. Therefore, by (6), $\mathbb{R}^d \setminus E \in \mathcal{L}(\mathbb{R}^d)$.
(7). Since $E_i \in \mathcal{L}(\mathbb{R}^d)$, we have $E_i^C \in \mathcal{L}(\mathbb{R}^d)$ and $\bigcup_{i=1}^{\infty} E_i^C \in \mathcal{L}(\mathbb{R}^d)$. Therefore $(\bigcap_{i=1}^{\infty} E_i)^C \in \mathcal{L}(\mathbb{R}^d)$ and $\bigcap_{i=1}^{\infty} E_i \in \mathcal{L}(\mathbb{R}^d)$. □

For $E \in \mathcal{L}(\mathbb{R}^d)$, its Lebesgue measure is defined to be $m(E) := m^*(E)$, and it has the following properties, which is significantly better than Lebesgue outer measure.

Proposition 1.3.3. *(The measure axioms).*

1. $m(\emptyset) = 0$.

2. (σ-additivity) For a countable sequence of disjoint sets $E_1, E_2 ... \in \mathcal{L}(\mathbb{R}^d)$,

$$m(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m(E_n)$$

Proof. (1). is trivial.
(2). *Case 1.* E_n is compact.
Then, by Theorem 1.3.1., $dist(E_i, E_j) > 0$, and

$$m(\bigcup_{i=1}^{N} E_i) = \sum_{i=1}^{N} m(E_i)$$

By monotonicity,

$$m(\bigcup_{i=1}^{\infty} E_i) \geq m(\bigcup_{i=1}^{N} E_i) = \sum_{i=1}^{N} m(E_i)$$

Let $N \to \infty$,

$$m(\bigcup_{i=1}^{\infty} E_i) \geq \sum_{i=1}^{\infty} m(E_i)$$

Also from σ-subadditivity,

$$m(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} m(E_i)$$

Therefore we have

$$m(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m(E_i)$$

Case 2. E_n is not compact but bounded.
For each E_n, it can be written as the union of a compact set U_n and a set with outer measure $\frac{\epsilon}{2^n}$. Thus,

$$m(E_n) \leq m(U_n) + \frac{\epsilon}{2^n}$$

$$\sum_{n=1}^{\infty} m(E_n) \leq \sum_{n=1}^{\infty} m(U_n) + \epsilon$$

We just showed that for compact set,

$$\sum_{n=1}^{\infty} m(U_n) = m(\bigcup_{i=1}^{\infty} U_n)$$

and by monotonicity,

$$m(\bigcup_{i=1}^{\infty} U_n) \leq m(\bigcup_{i=1}^{\infty} E_n)$$

Thus,

$$\sum_{n=1}^{\infty} m(E_n) \leq m(\bigcup_{i=1}^{\infty} E_n) + \epsilon$$

Since $\epsilon > 0$ arbitrary, we have

$$\sum_{n=1}^{\infty} m(E_n) \leq m(\bigcup_{i=1}^{\infty} E_n)$$

Also from σ-subadditivity, we have

$$\sum_{n=1}^{\infty} m(E_n) \geq m(\bigcup_{i=1}^{\infty} E_n)$$

Thus

$$\sum_{n=1}^{\infty} m(E_n) = m(\bigcup_{i=1}^{\infty} E_n)$$

Case 3. E_n is not compact and not closed.
Decompose $\mathbb{R}^d$ into annulus, for $m = 1, 2, ...,$

$$A_m := \{x \in \mathbb{R}^d : m-1 \leq |x| \leq m\}$$

Then, each E_n can be written as $E_n = \bigcup_{m=1}^{\infty} E_n \cap A_m$ for $E_n \cap A_m$ bounded, measurable, and disjoint.
Then, by previous argument,

$$m(E_n) = \sum_{m=1}^{\infty} m(E_n \cap A_m)$$

Also, for $E_n \cap A_m$ bounded, measurable, and disjoint,

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} E_n \cap A_m$$

Then

$$m(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} m(E_n \cap A_m) = \sum_{n=1}^{\infty} m(E_n)$$

□

Theorem 1.3.2. *(Monotone convergence theorem for measurable sets).*

(i) (Upward monotone convergence). Let $E_1 \subset E_2 \subset \cdots \subset \mathbb{R}^d$ be a countable non-decreasing sequence of Lebesgue measurable sets. Then $m\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_n)$. (Hint: Express $\bigcup_{n=1}^{\infty} E_n$ as the countable union of the lacunae $E_n \setminus \bigcup_{n'=1}^{n-1} E_{n'}$.)

(ii) (Downward monotone convergence) Let $\mathbb{R}^d \supset E_1 \supset E_2 \supset \ldots$ be a countable non-increasing sequence of Lebesgue measurable sets. If at least one of the $m(E_n)$ is finite, Then $m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_n)$.

(iii) Give a counterexample to show that the hypothesis that at least one of the $m(E_n)$ is finite in the downward monotone convergence theorem cannot be dropped.

Proof. (1). Let $E_0 = \emptyset \subset E_1$. By expressing each finite union and countable union as the finite or countable union of the lacunae form,

$$\bigcup_{n=1}^{2} E_n = (E_2 \setminus E_1) \cup E_1$$

$$\bigcup_{n=1}^{3} E_n = (E_3 \setminus \bigcup_{n'=1}^{2} E_{n'}) \cup (E_2 \setminus E_1) \cup E_1$$

$$\bigcup_{n=1}^{4} E_n = (E_4 \setminus \bigcup_{n'=1}^{3} E_{n'}) \cup (E_3 \setminus \bigcup_{n'=1}^{2} E_{n'}) \cup (E_2 \setminus E_1) \cup E_1$$

$$\cdots\cdots$$

$$\bigcup_{n=1}^{N} E_n = \bigcup_{k=1}^{N} (E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'})$$

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{k=1}^{\infty} (E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'})$$

Since each $E_k \in \mathcal{L}(\mathbb{R}^d)$, we have $\bigcup_{n'=1}^{k-1} E_{n'} \in \mathcal{L}(\mathbb{R}^d)$ and $E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'} \in \mathcal{L}(\mathbb{R}^d)$, and any countable union of the latter is Lebesgue measurable as well. Also, $E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'}$ and $E_j \setminus \bigcup_{n'=1}^{j-1} E_{n'}$ are disjoint for any $k \neq j$.
Hence, from countable additivity,

$$\begin{aligned}
m(\bigcup_{n=1}^{\infty} E_n) &= m(\bigcup_{k=1}^{\infty} (E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'})) \\
&= \sum_{i=1}^{\infty} m(E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'}) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} m(E_k \setminus \bigcup_{n'=1}^{k-1} E_{n'}) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} (m(E_k) - m(\bigcup_{n'=1}^{k-1} E_{n'})) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} (m(E_k) - m(E_{k-1})) \\
&= \lim_{n\to\infty} (m(E_n) - m(E_0)) \\
&= \lim_{n\to\infty} m(E_n)
\end{aligned}$$

(2). Assume $m(E_k) < \infty$ for some k. Since $E_1, E_2, ...$ are all Lebesgue measurable, $E_k \setminus E_n$, $\bigcap_{n=k+1}^{\infty} E_n$ are Lebesgue measurable, then $\bigcup_{n=k+1}^{\infty} E_k \setminus E_n$ and $E_k \setminus \bigcap_{n=k+1}^{\infty} E_n$ are also Lebesgue measurable, and from basic set calculation we know that

$$\bigcup_{n=k+1}^{\infty} E_k \setminus E_n = E_k \setminus \bigcap_{n=2}^{\infty} E_n$$

Then, using conclusion from (1),

$$\begin{aligned}
m(E_k \setminus \bigcap_{n=k+1}^{\infty} E_n) &= m(\bigcup_{n=k+1}^{\infty} E_k \setminus E_n) \\
&= \lim_{n\to\infty} m(E_k \setminus E_n) \\
&= \lim_{n\to\infty} m(E_k) - m(E_n) \\
&= m(E_k) - \lim_{n\to\infty} m(E_n)
\end{aligned}$$

Since $\bigcap_{n=k+1}^{\infty} E_n \subset E_k$, we have

$$m(E_k \setminus \bigcap_{n=k+1}^{\infty} E_n) = m(E_k) - m(\bigcap_{n=k+1}^{\infty} E_n)$$

Now we prove $\bigcap_{n=k+1}^{\infty} E_n = \bigcap_{n=1}^{\infty} E_n$. $\forall x \in \bigcap_{n=k+1}^{\infty} E_n$, $x \in E_{k+1}$, since $E_{k+1} \subset E_1$, $x \in \bigcap_{n=1}^{\infty} E_n$. $\forall y \in \bigcap_{n=1}^{\infty} E_n$, $y \in E_{k+1}, E_{k+2}, \ldots$, so $y \in \bigcap_{n=k+1}^{\infty} E_n$. Then,

$$\lim_{n\to\infty} m(E_n) = m(\bigcap_{n=k+1}^{\infty} E_n) = m(\bigcap_{n=1}^{\infty} E_n)$$

(3). Consider the sequence $E_n := \mathbb{R}_+ \setminus [0, n]$. Clearly none of the $m(E_n)$ is finite. We have $m\left(\bigcap_{n=1}^{\infty} E_n\right) = m(\emptyset) = 0$. On the other hand,

$$\forall n \in \mathbb{N}, m(E_n) = \infty;$$

thus, the sequence of measures does not converge.

$\square$

Theorem 1.3.3. *(Dominated Convergence Theorem). We say that a sequence E_n of sets in $\mathbb{R}^d$ converges pointwise to another set E in $\mathbb{R}^d$ if the indicator functions 1_{E_n} converge pointwise to 1_E.*

(i) If the E_n are all Lebesgue measurable, and converge pointwise to E, then E is Lebesgue measurable also.

(Hint: use the identity $1_E(x) = \liminf_{n\to\infty} 1_{E_n}(x)$ or $1_E(x) = \limsup_{n\to\infty} 1_{E_n}(x)$ to write E in terms of countable unions and intersections of the E_n.)

(ii) (Dominated convergence theorem) Suppose that the E_n are all contained in another Lebesgue measurable set F of finite measure. Then $m(E_n)$ converges to $m(E)$.

(Hint: use the upward and downward monotone convergence theorems, Theorem 1.3.1.)

(iii) Give a counterexample to show that the dominated convergence theorem fails if the E_n are not contained in a set of finite measure, even if we assume that the $m(E_n)$ are all uniformly bounded.

Proof. (i). If $x \in E$, we have $\mathbf{1}_E(x) = 1$,

$$\lim_{n\to\infty} \inf_{k\geq n} \mathbf{1}_{E_k}(x) = 1$$

This means that, $\forall \epsilon > 0$, $\exists N$, when $n \geq N$, $|\inf_{n\geq N} \mathbf{1}_{E_n}(x) - 1| < \epsilon$. Since inf is non-decreasing and is less than or equal to 1, we have $\inf_{n\geq N} \mathbf{1}_{E_n}(x) > 1 - \epsilon$. Since $\epsilon > 0$ arbitrary, we have $\inf_{n\geq N} \mathbf{1}_{E_n}(x) = 1$. Then, $\mathbf{1}_{E_n}(x) = 1 \ \forall \ n \geq N$, which means that $x \in \bigcap_{n\geq N} E_n$. Since $\forall \epsilon > 0$ we can pick an N, we have $x \in \bigcap_{n\geq N} E_n \subset \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} E_n$.
If $x \in \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} E_n$, then $\exists N$, such that $x \in \bigcap_{n\geq N} E_n$. This means that $\mathbf{1}_{E_n}(x) = 1 \ \forall \ n \geq N$. Then, $\mathbf{1}_E(x) = \lim_{n\to\infty} \mathbf{1}_{E_n}(x) = 1$ because of pointwise convergence. Then $x \in E$.
Therefore, we have shown the following two sets are equivalent.

$$E = \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} E_n$$

Also, if $x \in E$, we have $\mathbf{1}_E(x) = 1$,

$$\lim_{n\to\infty} \sup_{k\geq n} \mathbf{1}_{E_k}(x) = 1$$

This means that $\forall \epsilon > 0$, $\forall N$, $\exists k$, such that when $k \geq N$, $\mathbf{1}_{E_k}(x) \geq 1 - \epsilon$. Since $\epsilon > 0$ is arbitrary, $\mathbf{1}_{E_k}(x) = 1$ and $x \in E_k$. Thus, $\forall \epsilon > 0$, $\forall N$, $x \in \bigcup_{k\geq N} E_k$ for some $k \geq N$. Thus, $x \in \bigcap_{N\in\mathbb{N}} \bigcup_{k\geq N} E_k$.
If $x \in \bigcap_{N\in\mathbb{N}} \bigcup_{k\geq N} E_k$, $x \in \bigcup_{k\geq N} E_k$ for all $N \in \mathbb{N}$. This means that $\forall N \in \mathbb{N}$, $\exists k_N \geq N$, such that $x \in E_{k_N}$ and $\mathbf{1}_{E_{k_N}}(x) = 1$. By pointwise convergence and the fact that $\mathbf{1}_{E_{k_N}}(x)$ is a subsequence of the convergent sequence $\mathbf{1}_{E_N}(x)$, we have $\mathbf{1}_E(x) = \lim_{N\to\infty} \mathbf{1}_{E_N}(x) = \lim_{N\to\infty} \mathbf{1}_{E_{k_N}}(x) = 1$. Then, $x \in E$. Therefre, we have shown the following two sets are equibalent,

$$E = \bigcap_{N\in\mathbb{N}} \bigcup_{k\geq N} E_k$$

Then we represented E as either a countable union or a countable intersection of Lebesgue measurable sets, and E is Lebesgue measurable.

(ii). Since

$$\bigcap_{n\geq 1} E_n \subset \bigcap_{n\geq 2} E_n \subset \bigcap_{n\geq 3} E_n \subset ...$$

ans they are all lebesgue measurable, we have

$$\begin{aligned} m(E) &= m(\bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} E_n) \\ &= \lim_{N\to\infty} m(\bigcap_{n\geq N} E_n)) \\ &\leq \lim_{N\to\infty} m(E_N) \end{aligned}$$

by monotonicity.
Similarly, since

$$\bigcup_{n\geq 1} E_n \supset \bigcup_{n\geq 2} E_n \supset \bigcup_{n\geq 3} E_n \supset ...$$

and $\bigcup_{n\geq k} E_n \subset F$ $\forall k \in \mathbb{N}$, F is a set with finite Lebesgue measure, we have $m(\bigcup_{n\geq k} E_n)$ is all finite for all k. Then,

$$\begin{aligned} m(E) &= m(\bigcap_{N\in\mathbb{N}} \bigcup_{k\geq N} E_k) \\ &= \lim_{N\to\infty} m(\bigcup_{n\geq N} E_n) \\ &\geq \lim_{N\to\infty} m(E_N) \end{aligned}$$

by monotonicity.
Now we have both $\leq$ and $\geq$, we conclude that $m(E) = \lim_{N\to\infty} m(E_N)$.

(iii). Consider the sequence $E_n := \mathbb{R}_+/[0,n]$. Clearly non of the E_n is contained in a set of finite measure. We have $E_n \to \emptyset$ pointwise. On the other hand,

$$\forall n \in \mathbb{N}, m(E_n) = \infty;$$

thus, the sequence of measures does not converge. □

Theorem 1.3.4. *(Inner regularity). Let $E \subset \mathbb{R}^d$ be Lebesgue measurable. Then*

$$m(E) = \sup_{K\subset E, K \ compact} m(K).$$

Proof. Since K is compact, K is Lebesgue measurable. Therefore, by monotonicity, for all $K \subset E$, $m(E) \geq m(K)$. Therefore,

$$m(E) \geq \sup_{K\subset E, K \text{ compact}} m(K)$$

Thus it suffices to prove that

$$m(E) \leq \sup_{K\subset E, K \text{ compact}} m(K)$$

Case 1: E is bounded. Take ∂E, then $E \setminus \partial E$ is open in $\mathbb{R}^d$ and can be written as $E \setminus \partial E = \bigcup_{n=1}^{\infty} B_i$ where B_i are almost disjoint.

$$\begin{aligned} m(E) &= m(E \setminus \partial E) \\ &= m(\bigcup_{i=1}^{\infty} B_i) \\ &= \sum_{i=1}^{\infty} |B_i|. \end{aligned}$$

Shrink each B_i to B_i' where $B_i' \subset B_i$ and $\partial B_i' \cap \partial B_i = \emptyset$ and $|B_i| \leq |B_i'| + \frac{\epsilon}{2^i}$ for arbigrary $\epsilon > 0$. Then,

$$\sum_{i=1}^{\infty} |B_i| \leq \sum_{i=1}^{\infty} |B_i'| + \epsilon$$

Also, $B_i' \cap B_j' = \emptyset$, so $m(\bigcup_{i=1}^{\infty} B_i') = \sum_{i=1}^{\infty} |B_i'|$.
Let $\overline{B_i'} = B_i' \cup \partial B_i'$ and it is a closed set. By our previous construction, $\overline{B_i'} \cap \overline{B_j'} = \emptyset$. Also, since boundaries has Lebesgue measure zero,

$$m(\bigcup_{i=1}^{\infty} B_i') = m(\overline{\bigcup_{n=1}^{\infty} B_i'}) = m(\bigcup_{i=1}^{\infty} \overline{B_i'})$$

Since $\overline{\bigcup_{i=1}^{\infty} B_i'}$ is closed and bounded, it is compact. Therefore, by the definition of supremum and that $\overline{\bigcup_{i=1}^{n} B_i'} \subset E$, we have

$$\begin{aligned}
m(E) &= \sum_{i=1}^{\infty} |B_i| \\
&\leq \sum_{i=1}^{\infty} |B_i'| + \epsilon \\
&= m(\bigcup_{i=1}^{\infty} \overline{B_i'}) + \epsilon \\
&= m(\overline{\bigcup_{n=1}^{\infty} B_i'}) + \epsilon \\
&\leq \sup_{K \subset E, K \text{ compact}} m(K) + \epsilon
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have $m(E) \leq \sup_{K \subset E, K \text{ compact}} m(K)$.
Now we have $\leq$ and $\geq$, we have $m(E) = \sup_{K \subset E, K \text{ compact}} m(K)$.

□

Theorem 1.3.5. *(Outer measure is not finitely additive). There exist disjoint bounded subsets E, F of the real line such that*

$$m^*(E \cup F) \neq m^*(E) + m^*(F).$$

(Hint: Show that the set constructed in the proof of the above proposition has positive outer measure.)

Proof. Consider the set that we will construct in the proof of Proposition 1.3.4.: $E := \{x_C : C \in \mathbb{R}/\mathbb{Q} \text{ and } x_C \in C \cap [0,1]\}$, and $\tilde{E} := \bigcup_{q \in \mathbb{Q} \cap [-1,1]} E + q$. For each q, $E + q$ is disjoint from another. We know by countable subadditivity that

$$\begin{aligned}
m^*(\tilde{E}) &\leq \sum_{q \in \mathbb{Q} \cap [-1,1]} m^*(E + q) \\
&= \sum_{q \in \mathbb{Q} \cap [-1,1]} m^*(E) \\
&= \begin{cases} 0 & \text{if } m^*(E) = 0, \\ +\infty & \text{if } m^*(E) > 0. \end{cases}
\end{aligned}$$

Also $m^*(\tilde{E}) \in [1, 3]$. This contradict with the case $m^*(E) = 0$, so it can only be that

$m^*(E) > 0$ and it equals to some positive real number.
Let $n \in \mathbb{N}$ large enough so that $m^*(E) \geq \frac{1}{n}$. If m^* is finitely additive, then for a subset $F \subset \mathbb{Q} \cap [-1, 1]$ with $\#F = 3n$, we have

$$\begin{aligned} m^*(\bigcup_{q \in F} E + q) &= \sum_{q \in F} m^*(E + q) \\ &= \sum_{q \in F} m^*(E) \\ &= 3n \times m^*(E) \\ &> 3 \end{aligned}$$

However, by monotonicity, $m^*(\bigcup_{q \in F} E + q) \leq m^*(\bigcup_{q \in \mathbb{Q} \cap [-1,1]} E + q) \leq 3$, contradiction!
Thus, Lebesgue outer measure is not finitely additive. $\square$

1.3.3 Non-Measurable Sets

Of course, there are non-measurable sets in $\mathbb{R}^d$.

Proposition 1.3.4. $\exists E \subset [0, 1]$, $E \notin \mathcal{L}(\mathbb{R}^d)$.

Proof. We use the fact that $(\mathbb{Q}, +)$ is a subgroup of $(\mathbb{R}, +)$, and it partitions $\mathbb{R}$ into disjoint cosets $x + \mathbb{Q}$. This create a quotient group $\mathbb{R}/\mathbb{Q} := \{x + \mathbb{Q} : x \in \mathbb{R}\}$. Each coset $C = x + \mathbb{Q}$ of $\mathbb{R}/\mathbb{Q}$ is dense in $\mathbb{R}$, so it has non-empty intersection with $[0, 1]$.
By axiom of choice, select $x_C \in C \cap [0, 1]$ from each $C \in \mathbb{R}/\mathbb{Q}$. Let $E := \{x_C : C \in \mathbb{R}/\mathbb{Q}, x_C \in C \cap [0, 1]\}$ be the collection of all these coset representatives. By construction, $E \subset [0, 1]$.
Claim 1. $[0, 1] \subseteq \bigcup_{q \in \mathbb{Q} \cap [-1,1]} E + q$. Indeed, for any $y \in [0, 1]$, $\exists C \in \mathbb{R}/\mathbb{Q}$ such that $y \in C$. Then, $y - x_C$ is rational. Since $y, x_C \in [0, 1]$, we have $|x_C - y| \leq 1$. Let $q = y - x_C$, since $x_C \in E$, we have $y \in E + q$.
Claim 2. For $q_1 \neq q_2 \in \mathbb{Q}$, $(E + q_1) \cap (E + q_2) = \emptyset$. Indeed, if $z \in (E + q_1) \cap (E + q_2)$, then $z = x_1 + q_1 = x_2 + q_2$ for $x_1, x_2 \in E$. Then, $x_1 = x_2 + (q_2 - q_1)$ where $q_2 - q_1$ is rational. Then, x_1 and x_2 are in the same coset C, then $x_1 = x_2 = x_C$, then $q_1 = q_2$, contradiction.
Suppose $E \in \mathcal{L}(\mathbb{R})$, then $E + q \in \mathcal{L}(\mathbb{R})$ $\forall q \in \mathbb{Q}$, and $\tilde{E} := \bigcup_{q \in \mathbb{Q} \cap [-1,1]} E + q \in \mathcal{L}(\mathbb{R})$. By monotonicity and *Claim 1*, $1 = m([0, 1]) \leq m(\tilde{E})$. Also since $\tilde{E} \subset [-1, 2]$, we have $m(\tilde{E}) \in [1, 3]$. By σ-additivity and *Claim 2* and transfomation invariant,

$$m(\tilde{E}) = \sum_{q \in \mathbb{Q} \cap [-1,1]} m(E + q) = \sum_{q \in \mathbb{Q} \cap [-1,1]} m(E)$$

If $m(E) = 0$ then $m(\tilde{E}) = 0$. If $m(E) > 0$ then $m(\tilde{E}) = +\infty$. Contradiction! Therefore, $E \notin \mathcal{L}(\mathbb{R})$. $\square$

1.4 Lebesgue Integral

1.4.1 Integration of Simple Functions

Definition 1.4.1. *(Simple function): A complex valued simple function $f : \mathbb{R}^d \to \mathbb{C}$ is a finite linear combination*

$$f = \sum_{k=1}^{n} c_k \mathbf{1}_{E_k}$$

for $E_k \in \mathcal{L}(\mathbb{R}^d)$, $c_k \in \mathbb{C}$.
An unsigned simple function just tales $c_k \in [0, +\infty)$. For an indicator function,

$$\int_{\mathbb{R}^d} \mathbf{1}_E(x)dx = m(E)$$

Definition 1.4.2. *(Integral of a simple function): for $f = \sum_{k=1}^{n} c_k \mathbf{1}_{E_k}$,*

$$Simp \int_{\mathbb{R}^d} f := \sum_{k=1}^{\infty} c_k m(E_k)$$

Here we check that this is well-defined.

Lemma 1.4.1. *Let $k, k' \geq 0$ be natual numbers, $c_1, ..., c_k, c'_1, ..., c'_{k'} \in [0, +\infty]$. Let $E_1, ...E_k, E'_1, ..., E_{k'} \subset \mathbb{R}^d$ be in $\mathcal{L}(\mathbb{R}^d)$ such that*

$$\sum_{i=1}^{k} c_i \mathbf{1}_{E_i} = \sum_{i=1}^{k'} c'_i \mathbf{1}_{E'_i} (*)$$

holds identically on $\mathbb{R}^d$. Then,

$$\sum_{i=1}^{k} c_i m(E_i) = \sum_{i=1}^{k'} c'_i m(E'_i)$$

Proof. First, $\{E_1, ..., E_k, E'_1, ..., E'_{k'}\}$ partitions $\mathbb{R}^d$ into $2^{k+k'}$ disjoint sets using finite Boolean algebra, each of which is an intersection of some of the $E_1, ..., E_k, E'_1, ..., E'_{k'}$ and their compliments. Letting go empty sets, we are left with m non-empty disjoint sets $A_1, ..., A_m$ for some $0 \leq m \leq 2^{k+k'}$. $A_i \in \mathcal{L}(\mathbb{R}^d)$ for $i \in \{1, ..., m\}$.
Then, $E_i = \bigcup_{j \in J_i} A_j$, $E'_{i'} = \bigcup_{j' \in J'_{i'}} A_{j'}$ for all $i = 1, ..., k$ and $j' = 1, ..., k'$ and some subsets $J_i, J'_{i'}$. By finite additivity, $m(E_i) = \sum_{j \in J_i} m(A_i)$, $m(E'_{i'}) = \sum_{j' \in J'_{i'}} m(A_{j'})$. Thus we need to prove that

$$\sum_{i=1}^{k} c_i \sum_{j \in J_i} m(A_i) = \sum_{i'=1}^{k'} c'_i \sum_{j \in J'_{i'}} m(A_{i'})$$

Fix $1 \leq j \leq m$, evaluate $(*)$ at a point x in the non-empty set A_j. Then, at such point,

$$\mathbf{1}_{E_i}(x) = \mathbf{1}_{J_i}(j)$$

$$\mathbf{1}_{E'_i}(x) = \mathbf{1}_{J'_{i'}}(j)$$

By $(*)$,

$$\sum_{i=1}^{k} c_i \mathbf{1}_{J_i}(j) = \sum_{i'=1}^{k'} c'_{i'} \mathbf{1}_{J'_{i'}}(j)$$

Multiply by $m(A_j)$,

$$\sum_{i=1}^{k} c_i \mathbf{1}_{J_i}(j) m(A_j) = \sum_{i'=1}^{k'} c'_{i'} \mathbf{1}_{J'_{i'}}(j) m(A_j)$$

Sum up $j = 1, ..., m$,

$$\sum_{i=1}^{k} c_i \sum_{j=1}^{m} \mathbf{1}_{J_i}(j) m(A_j) = \sum_{i'=1}^{k'} c'_{i'} \sum_{j=1}^{m} \mathbf{1}_{J'_{i'}}(j) m(A_j)$$

$$\sum_{i=1}^{k} c_i \sum_{j \in J_i} m(A_i) = \sum_{i'=1}^{k'} c'_i \sum_{j \in J'_{i'}} m(A_i)$$

□

Definition 1.4.3. *(Almost everywhere and support). A property $P(x)$ of $x \in \mathbb{R}^d$ holds (Lebesgue) almost everywhere (a.e.) if $\{x : P(x)$ does not hold$\}$ is a null set, that is, $m^*(\{x : P(x)$ does not hold$\}) = 0$.*
The support of a function f is $\{x \in \mathbb{R} : f(x) \neq 0\}$.

1.4.2 Measurable Functions

By extending the class of unsigned simple functions to the larger class of unsigned Lebesgue measurable functions, we can complete the unsigned simple integral to the unsigned Lebesgue integral.

Definition 1.4.4. *(Unsigned measurable function). An unsinged function f is Lebesgue measurable if it is the pointwise limit of unsigned simple functions, i.e., if $\exists f_1, f_2, f_3, ... : \mathbb{R}^d \to [0, +\infty]$ of unsigned simple functions such that $f_n(x) \to f(x)$ $\forall x \in \mathbb{R}^d$.*

This definition has some equivalent forms.

Lemma 1.4.2. *(Equivalent Notions of Measurability). Let $f : \mathbb{R}^d \to [0, +\infty]$, the followings are equivalent.*

1. *f is Lebesgue measurable .*
2. *f is the pointwise a.e. limit of unsigned simple functions f_n. Thus $\lim_{n\to\infty} f_n(x)$ exists and $f(x) = \lim_{n\to\infty} f_n(x)$ for a.e. $x \in \mathbb{R}^d$.*
3. *For every interval $I \subset [0, +\infty)$, the set $f^{-1}(I) := \{x \in \mathbb{R}^d : f(x) \in I\}$ is Lebesgue measurable. $f^{-1}(I) \in \mathcal{L}(\mathbb{R}^d)$.*

Proof. (i) $\Rightarrow$ (ii) is immediate from definition.
(ii) $\Rightarrow$ (iii): Assume that $\exists$ simple functions $f_n \to f$ pointwise a.e.. Then, for almost every $x \in \mathbb{R}^d$ and $N \in \mathbb{N}$,

$$f(x) = \lim_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x) = \inf_{N>0} \sup_{n \geq N} f_n(x) := \tilde{f}(x)$$

Let $\lambda > 0$ be arbitrary, and denote $\{g > \lambda\} := \{x \in \mathbb{R}^d : g(x) > \lambda\}$ for $g : \mathbb{R}^d \to [0, +\infty]$, we have for $M, N \in \mathbb{N}$,

$$\begin{aligned}\{\tilde{f} > \lambda\} &= \bigcup_{M>0} \bigcap_{N>0} \{x \in \mathbb{R}^d : \sup_{n \geq N} f_n(x) > \lambda + \frac{1}{M}\} \\ &= \bigcup_{M>0} \bigcap_{N>0} \bigcup_{n \geq N} \{x \in \mathbb{R}^d : f_n(x) > \lambda + \frac{1}{M}\}\end{aligned}$$

outside of a set of measure zero.
Since f_n is unsigned simple, $\{x \in \mathbb{R}^d : f_n(x) > \lambda\} \in \mathcal{L}(\mathbb{R}^d)$. By (6) and (7) of Proposition 1.3.2., $\{\tilde{f} > \lambda\} \in \mathcal{L}(\mathbb{R}^d)$. Also, $\{f > \lambda\}$ and $\{\tilde{f} > \lambda\}$ differs by a null set, so $\{f > \lambda\} \in \mathcal{L}(\mathbb{R}^d)$. Thus we have proved that $f^{-1}(I) \in \mathcal{L}(\mathbb{R}^d)$ for $I = (\lambda, +\infty)$.
Note that $\{f \geq \lambda\} = \bigcap_{\lambda' \in \mathbb{Q}, \lambda' < \lambda} \{f > \lambda'\}$. Then by (7) of Proposition 1.3.2., $\{f \geq \lambda\} \in \mathcal{L}(\mathbb{R}^d)$.
Note that

$$\begin{aligned}f^{-1}([a,b]) &= \{f \geq a\} \setminus \{f > b\} \\ f^{-1}([a,b)) &= \{f \geq a\} \setminus \{f \geq b\} \\ f^{-1}((a,b]) &= \{f > a\} \setminus \{f > b\} \\ f^{-1}((a,b)) &= \{f > a\} \setminus \{f \geq b\}\end{aligned}$$

By Proposition 1.3.2., they are all Lebesgue measurable.
(iii) $\Rightarrow$ (i): Let $f : \mathbb{R}^d \to [0, +\infty]$ with $f^{-1}(I) \in \mathcal{L}(\mathbb{R}^d)$ $\forall I \subset \mathbb{R}^d$ as an interval.
For each $n \geq 1$ and $x \in \mathbb{R}^d$, set

$$f_n(x) = \max_{m \in \mathbb{Z}} \{m 2^{-n} : m 2^{-n} \leq \min(f(x), n) \mathbf{1}_{\overline{B(0,n)}}(x)\}$$

Then, $f_1 \leq f_2 \leq ...$ pointwise, and $f(x) = \sup_{n \in \mathbb{N}} f_n$ $\forall x \in \mathbb{R}^d$. Each f_n takes finitely many values and for any $c \in [0, +\infty)$, $f_n^{-1}(c) = f^{-1}(I_c) \cap \overline{B(0,n)}$ for some interval I_c that is measurable. Thus f_n is simple, and is bounded and has finite measure support, and the claim follows. ☐

Theorem 1.4.1. *(Functions that are Measurable).*

1. *Every continuous function $f : \mathbb{R}^d \to [0, +\infty]$ is measurable.*
2. *The supremum, infimum, limit superior, or limit inferior of unsigned measurable functions is unsigned measurable.*

Proof. (1). First, divide the $\mathbb{R}^d$ into dyadic cubes:

$$\begin{aligned}\mathbb{R} &= \bigcup_{k_1 \in \mathbb{Z}} \bigcup_{k_2 \in \mathbb{Z}} ... \bigcup_{k_d \in \mathbb{Z}} [\frac{k_1}{2^n}, \frac{k_1 + 1}{2^n}) \times [\frac{k_2}{2^n}, \frac{k_2 + 1}{2^n}) \times ... \times [\frac{k_d}{2^n}, \frac{k_d + 1}{2^n}) \ \forall n \in \mathbb{N} \\ &= \bigcup_{k_1 \in \mathbb{Z}} \bigcup_{k_2 \in \mathbb{Z}} ... \bigcup_{k_d \in \mathbb{Z}} B_{k_1, k_2, ..., k_d} \ \forall n \in \mathbb{N}\end{aligned}$$

Then, define an unsigned simple function f_n can be written as

$$f_n(x) = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} ... \sum_{k_d \in \mathbb{Z}} \mathbf{1}_{B_{k_1,k_2,...,k_d}}(x) \inf_{B_{k_1,k_2,...,k_d}} f(x)$$

For all $x \in \mathbb{R}^d$, x has to be in one of the dyadic cubes. When $x \in B_{k_1,k_2,...,k_d}$, $f_n(x) = \inf_{B_{k_1,k_2,...,k_d}} f(x)$. Thus it suffices to show that $f_n \to f$ pointwise a.e. in this case.
It is easy to see that $f_n \geq 0$. By the definition of infimum,

$$\inf_{B_{k_1,k_2,...,k_d}} f(x) \leq f(x) \; \forall x \in B_{k_1,k_2,...,k_d}$$

$\forall \epsilon > 0$, $\exists x' \in B_{k_1,k_2,...,k_d}$, such that $f(x') \leq \inf_{B_{k_1,k_2,...,k_d}} f(x) + \epsilon$. Together with the last inequality we have

$$0 \leq f(x') - \inf_{B_{k_1,k_2,...,k_d}} f(x) \leq \epsilon$$

$$|f(x') - \inf_{B_{k_1,k_2,...,k_d}} f(x)| \leq \epsilon$$

Notice that, two points inside the closure of $B_{k_1,k_2,...,k_d}$ has distance less than $\frac{\sqrt{d}}{2^n}$. Thus, by continuity, $\exists \delta > 0$, pick n large enough that $n > log_2 \frac{\sqrt{d}}{\delta}$, so that $|x - x'| < \frac{\sqrt{d}}{2^n} < \delta$ implies $|f(x) - f(x')| \leq \epsilon$, $\forall x \in B_{k_1,k_2,...,k_d}$.
Therefore, $\forall x \in B_{k_1,k_2,...,k_d}$,

$$\begin{aligned} |f(x) - \inf_{B_{k_1,k_2,...,k_d}} f(x)| &\leq |f(x) - f(x')| + |f(x') - \inf_{B_{k_1,k_2,...,k_d}} f(x)| \\ &\leq 2\epsilon \end{aligned}$$

This means that, $\forall \epsilon > 0$, we found that when $n > log_2 \frac{\sqrt{d}}{\delta}$, $|f(x) - \inf_{B_{k_1,k_2,...,k_d}} f(x)| = |f(x) - f_n(x)| \leq 2\epsilon$. $f_n \to f$. Therefore, for all $x \in \mathbb{R}^d$ (equivalent to x in arbitrary cube), $f_n \to f$ pointwise a.s. Then, f is Lebesgue measurable.
(2). Since $\{f_n\}_{n \in \mathbb{N}}$ is unsigned measurable, $\{f_n \geq c\}$ and $\{f_n \leq c\}$ are measurable for all $c \geq 0$ by Lemma 1.4.2.. Thus, $\forall n \in \mathbb{N}$,

$$\begin{aligned} \{\sup_{k \geq n} f_k \leq c\} &= \{x \in \mathbb{R}^d : \sup_{k \geq n} f_k \leq c\} \\ &= \{x \in \mathbb{R}^d : f_k \leq c \; \forall k \geq n\} \\ &= \bigcap_{k \geq n} \{x \in \mathbb{R}^d : f_k \leq c\} \end{aligned}$$

$$\begin{aligned} \{\inf_{k \geq n} f_k \geq c\} &= \{x \in \mathbb{R}^d : \inf_{k \geq n} f_k \geq c\} \\ &= \{x \in \mathbb{R}^d : f_k \geq c \; \forall k \geq n\} \\ &= \bigcap_{k \geq n} \{x \in \mathbb{R}^d : f_k \geq c\} \end{aligned}$$

Since those are countable intersections of measurable sets, the functions

$$g_n := \sup_{k \geq n} f_k \; , \; h_n := \inf_{k \geq n} f_k$$

are unsigned measurable for all $n \in \mathbb{N}$. Take $n = 1$, then $\sup_{n\in\mathbb{N}} f_n$ and $\inf_{n\in\mathbb{N}} f_n$ are unsigned measurable.
Notice that g_n is monotonic decreasing in n and h_n is monotonic increasing in n. So

$$\lim_{n\to\infty} \sup_{k\geq n} f_k = \inf_{n\geq 1} \sup_{k\geq n} f_k = \inf_{n\geq 1} g_n$$

$$\lim_{n\to\infty} \inf_{k\geq n} f_k = \sup_{n\geq 1} \inf_{k\geq n} f_k = \sup_{n\geq 1} h_n$$

Since g_n and h_n are unsigned measurable, by the previous argument we have that the limsup and the liminf of f_n are unsigned measurable.

□

Theorem 1.4.2. *(Bounded unsigned measurable function is the uniform limit of bounded simple functions). Let $f : \mathbb{R}^d \to [0, +\infty]$. Then, f is a bounded unsigned measurable function if and only if f is the uniform limit of bounded simple functions.*

Proof. First we prove $\Leftarrow$.
Suppose f_n is a sequence of bounded simple functions and $f_n \to f$ uniformly. Then, $\forall \epsilon > 0$, $\exists N$ such that when $n \geq N$, $|f_n(x) - f(x)| \leq \epsilon \ \forall x \in \mathbb{R}^d$. By triangle inequality we have $|f(x)| \leq \epsilon + |f_n(x)|$. Since $f_n(x)$ is bounded, it is clear that f is also bounded. Also, uniform convergent induces pointwise convergence, and then f is Lebesgue measurable by Lemma 1.4.2.
Then we prove $\Rightarrow$. By Lemma 1.4.2., since f is unsigned Lebesgue measurable, f is the supremum $f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ of an increasing sequence $0 \leq f_1 \leq f_2 \leq ...$ of unsigned simple functions f_n, each of which are bounded with finite measure support.
By the definition of supremum, $\forall x$, $\forall \epsilon > 0$, $\exists n' \in \mathbb{N}$ such that $f_{n'}(x) \geq f(x) - \epsilon$. Also since f_n is monotonic increasing, we have $|f_{n'}(x) - f(x)| \leq \epsilon$.
Since $f_{n'}$ is a unsigned simple function, it can take finite many values $c_1, c_2, ..., c_m$. Based on these values, we divide the domain of $f_{n'}$ ($\mathbb{R}^d$) into m parts,

$$\mathbb{R}^d = f_{n'}^{-1}(c_1) \cup f_{n'}^{-1}(c_2) \cup ... \cup f_{n'}^{-1}(c_m)$$

For each $f_{n'}^{-1}(c_i)$, by supremum, $\forall \epsilon > 0$, $\exists n_i$, such that, when $n \geq n_i$, $|f_n(x) - f(x)| \leq \epsilon \ \forall x \in f_{n'}^{-1}(c_i)$. By this we got a collection of $\{n_1, n_2, ..., n_m\}$.
Set $N := \max\{n_1, n_2, ..., n_m\}$, then we have $\forall \epsilon > 0$, $\exists N = \max\{n_1, n_2, ..., n_m\}$ such that when $n \geq N$, $|f_n(x) - f(x)| \leq \epsilon \ \forall x \in \mathbb{R}^d$. Thus we have $f_n \to f$ uniformly.
Also since $0 \leq f_n \leq \sup_{n\in\mathbb{N}} f_n = f$, f is bounded, we have f_n is bounded. □

1.4.3 Unsigned Lebesgue Integrals

Now let's integrate unsigned measurable functions.

Definition 1.4.5. *(Lower and upper unsigned Lebesgue integral). Let $f : \mathbb{R}^d \to [0, +\infty]$. Define the lower unsigned Lebesgue integral of f as*

$$\underline{\int_{\mathbb{R}^d}} f := \sup_{0\leq g\leq f, g \ simple} Simp \int_{\mathbb{R}^d} g$$

and the upper unsigned Lebesgue integral of f as

$$\overline{\int_{\mathbb{R}^d}} f := \inf_{f \leq h, h \text{ simple}} \mathit{Simp} \int_{\mathbb{R}^d} h$$

Horizontal truncation property: as $n \to \infty$, $\underline{\int_{\mathbb{R}^d}} \min(f, n) \to \underline{\int_{\mathbb{R}^d}} f$. Vertical truncation property: as $n \to \infty$, $\underline{\int_{\mathbb{R}^d}} f \mathbf{1}_{|x| \leq n} \to \underline{\int_{\mathbb{R}^d}} f$. Subadditivity of upper integral: $\overline{\int_{\mathbb{R}^d}} f + g \leq \overline{\int_{\mathbb{R}^d}} f + \overline{\int_{\mathbb{R}^d}} g$. Superadditivity of lower integral: $\underline{\int_{\mathbb{R}^d}} f + g \geq \underline{\int_{\mathbb{R}^d}} f + \underline{\int_{\mathbb{R}^d}} g$.

Definition 1.4.6. *(Unsigned Lebesgue integral). Let $f : \mathbb{R}^d \to [0, \infty]$ be measurable. Define its unsigned Lebesgue integral as*

$$\int_{\mathbb{R}^d} f := \underline{\int_{\mathbb{R}^d}} f = \sup_{0 \leq g \leq f, g \text{ simple}} \mathit{Simp} \int_{\mathbb{R}^d} g$$

For $f : \mathbb{R}^d \to [0, \infty]$ measurable, bounded, and vanishing outside of a set of finite measure, the lower and upper Lebesgue integrals match.

Theorem 1.4.3. *Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable, bounded, and vanishing outside of a set of finite measure. Then the lower and upper Lebesgue integrals of f agree. (Hint: use Exercise 1.3.4.) There is a converse to this statement, but we will defer it to later notes. What happens if f is allowed to be unbounded, or is not supported inside a set of finite measure?*

Proof. The very obvious way to prove this is by first proving $\leq$ then proving $\geq$.
For $\leq$: by definition,

$$\underline{\int_{\mathbb{R}^d}} f = \sup_{0 \leq g \leq f, g \text{ simple}} \text{Simp} \int_{\mathbb{R}^d} g$$

$$\overline{\int_{\mathbb{R}^d}} f = \inf_{h \geq f, g \text{ simple}} \text{Simp} \int_{\mathbb{R}^d} h$$

Since $g \leq h$ for every g, h that satisfy the conditions, by definition we natually have $\underline{\int_{\mathbb{R}^d}} f \leq \overline{\int_{\mathbb{R}^d}} f$. Thus it suffices to show that $\underline{\int_{\mathbb{R}^d}} f \geq \overline{\int_{\mathbb{R}^d}} f$.

Let S be the finite measure support on which $f > 0$. By Theorem 1.4.2., we can find a sequence of unsigned simple functions $(g_n)_{n \in \mathbb{N}}$ such that (i) $0 \leq g_1 \leq g_2 \leq ...$ (ii) has finite measure support S (iii) $g_n \to f$ uniformly.

Pick a subsequence of this original sequence so that $\forall n \in \mathbb{N}$, $d_\infty(g_n, f) \leq \frac{1}{n}$. Furthermore, construct another sequence of unsigned simple functions $(h_n)_{n \in \mathbb{N}}$ such that $h_n = g_n + \frac{2}{n}$. Then, $h_n - f = g_n + \frac{2}{n} - f \geq \frac{1}{n} \geq 0$. $d_\infty(h_n, f) \leq d_\infty(h_n, g_n) + d_\infty(g_n, f) \leq \frac{3}{n}$.

Since now $\forall n \in \mathbb{N}$, $d_\infty(h_n, g_n) \leq \frac{2}{n}$, we have $d_\infty(h_n, g_n) \to 0$ as $n \to \infty$. They converge uniformly to each other.

Pick a simple function $g' \leq f$ that satisfies $\underline{\int_{\mathbb{R}^d}} f - \text{Simp} \int_{\mathbb{R}^d} g' \leq \frac{1}{n}$, $d_\infty(g', f) \leq \frac{1}{n}$.

$$\begin{aligned}
\underline{\int_{\mathbb{R}^d}} f - \text{Simp} \int_{\mathbb{R}^d} g_n &= \underline{\int_{\mathbb{R}^d}} f - \text{Simp} \int_{\mathbb{R}^d} g' + \text{Simp} \int_{\mathbb{R}^d} g' - \text{Simp} \int_{\mathbb{R}^d} g_n \\
&\leq \frac{1}{n} + \text{Simp} \int_{\mathbb{R}^d} |g' - g_n| \\
&\leq \frac{1}{n} + \text{Simp} \int_{\mathbb{R}^d} d_\infty(g', g_n) \mathbf{1}_S \\
&\leq \frac{1}{n} + \text{Simp} \int_{\mathbb{R}^d} (d_\infty(g', f) + d_\infty(f, g_n)) \mathbf{1}_S \\
&\leq \frac{1}{n} + \frac{2}{n} m(S) \\
&\to 0 \text{ as } n \to \infty
\end{aligned}$$

Thus $\lim_{n\to\infty} \text{Simp} \int_{\mathbb{R}^d} g_n = \underline{\int_{\mathbb{R}^d}} f$.
Pick another simple function $h' \geq f$ such that $\text{Simp} \int_{\mathbb{R}^d} h' - \overline{\int_{\mathbb{R}^d}} f \leq \frac{1}{n}$, $d_\infty(h', f) \leq \frac{1}{n}$.

$$\begin{aligned}
\text{Simp} \int_{\mathbb{R}^d} h_n - \overline{\int_{\mathbb{R}^d}} f &= \text{Simp} \int_{\mathbb{R}^d} h_n - \text{Simp} \int_{\mathbb{R}^d} h' + \text{Simp} \int_{\mathbb{R}^d} h' - \overline{\int_{\mathbb{R}^d}} f \\
&\leq \frac{1}{n} + \text{Simp} \int_{\mathbb{R}^d} |h' - h_n| \\
&\leq \frac{1}{n} + \text{Simp} \int_{\mathbb{R}^d} d_\infty(h', h_n) \mathbf{1}_S \\
&\leq \frac{1}{n} + \text{Simp} \int_{\mathbb{R}^d} (d_\infty(h', f) + d_\infty(f, h_n)) \mathbf{1}_S \\
&\leq \frac{1}{n} + \frac{2}{n} m(S) \\
&\to 0 \text{ as } n \to \infty
\end{aligned}$$

Thus $\lim_{n\to\infty} \text{Simp} \int_{\mathbb{R}^d} h_n = \overline{\int_{\mathbb{R}^d}} f$.
Since $\text{Simp} \int_{\mathbb{R}^d} (h_n - g_n) \leq \text{Simp} \int_{\mathbb{R}^d} d_\infty(h_n, g_n) \mathbf{1}_S \to 0$ as $n \to \infty$, we have $\lim_{n\to\infty} \text{Simp} \int_{\mathbb{R}^d} h_n = \lim_{n\to\infty} \text{Simp} \int_{\mathbb{R}^d} g_n$. Thus

$$\underline{\int_{\mathbb{R}^d}} f = \lim_{n\to\infty} \text{Simp} \int_{\mathbb{R}^d} g_n = \lim_{n\to\infty} \text{Simp} \int_{\mathbb{R}^d} h_n = \overline{\int_{\mathbb{R}^d}} f$$

□

Corollary 1.4.1. *(Finite additivity of Lebesgue integral). Let $f, g : \mathbb{R}^d \to [0, +\infty]$ be measurable. Then,*

$$\int_{\mathbb{R}^d} (f + g) = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$$

Proof. First, since f, g, $f + g$ are all measurable, their upper and lower Lebesgue integrals agree.

By horizontal truncation property and an limiting argument, we can assume that f, g are bounded. By vertical truncation property and an limiting argument, we can further assume that they have finite measure support.

By superadditivity and subadditivity,

$$\int_{\mathbb{R}^d} f+g = \overline{\int_{\mathbb{R}^d}} f+g \leq \overline{\int_{\mathbb{R}^d}} f + \overline{\int_{\mathbb{R}^d}} g = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$$
$$\int_{\mathbb{R}^d} f+g = \underline{\int_{\mathbb{R}^d}} f+g \geq \underline{\int_{\mathbb{R}^d}} f + \underline{\int_{\mathbb{R}^d}} g = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$$

Thus,

$$\int_{\mathbb{R}^d} f+g = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$$

□

We also have *Markov's Inequality*, which asserts that the Lebesgue integral of an unsigned measurable function controls how often that function can be large.

Lemma 1.4.3. *(Markov's Inequality). Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable. Then, for any $\lambda \in (0, +\infty)$,*

$$m(\{x \in \mathbb{R}^d : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbb{R}^d} f(x)dx$$

Proof. First, notice that

$$\lambda \mathbf{1}_{\{x \in \mathbb{R}^d : f(x) \geq \lambda\}} \leq f(x)$$

Then, since f is measurable,

$$\int_{\mathbb{R}^d} \lambda \mathbf{1}_{\{x \in \mathbb{R}^d : f(x) \geq \lambda\}} \leq \int_{\mathbb{R}^d} f(x)dx$$

By the definition of Lebesgue integral for simple function,

$$LHS = \lambda m(\{x \in \mathbb{R}^d : f(x) \geq \lambda\}) \leq \int_{\mathbb{R}^d} f(x)dx$$

Therefore we have

$$m(\{x \in \mathbb{R}^d : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbb{R}^d} f(x)dx$$

□

Corollary 1.4.2. *Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable.*

1. *If $\int_{\mathbb{R}^d} f(x)\, dx < \infty$, then f is finite almost everywhere. Give a counterexample to show that the converse statement is false.*
2. *$\int_{\mathbb{R}^d} f(x)\, dx = 0$ if and only if f is zero almost everywhere.*

Proof. (1). For $n \in \mathbb{N}$, we have

$$m(f \geq n) \leq \frac{1}{n} \int_{\mathbb{R}^d} f$$

Since $\int_{\mathbb{R}^d} f < \infty$, sending $n \to \infty$, we have

$$m(f \geq \infty) = 0$$

(2) First we prove $\Leftarrow$. Since $f = 0$ almost everywhere, $m(\{x \in \mathbb{R}^d : f(x) > 0\}) = 0$. Divide the domain into two disjoint measurable sets $\mathbb{R}^d = \{x \in \mathbb{R}^d : f(x) > 0\} \cup \{x \in \mathbb{R}^d : f(x) = 0\}$. Define a simple function $h = \sup_{x \in \{f>0\}} f \times \mathbf{1}_{\{f>0\}} + 0 \times \mathbf{1}_{\{f=0\}}$. Clearly $h \geq f$. Since f is measurable,

$$\begin{aligned}
\int_{\mathbb{R}^d} f &= \inf_{h \geq f, h \text{ simple}} \text{Simp} \int_{\mathbb{R}^d} h \\
&\leq \text{Simp} \int_{\mathbb{R}^d} (\sup_{x \in \{f>0\}} f \times \mathbf{1}_{\{f>0\}} + 0 \times \mathbf{1}_{\{f=0\}}) \\
&= \sup_{x \in \{f>0\}} f \times m(\{f > 0\}) + 0 \times m(\{f = 0\}) \\
&= 0
\end{aligned}$$

Also because of non-negativity, we have $\int_{\mathbb{R}^d} f = 0$.

Now we prove $\Rightarrow$. Since Markov's Inequality,

$$\begin{aligned}
m(\{x \in \mathbb{R}^d : f(x) \geq \lambda\}) &\leq \frac{1}{\lambda} \int_{\mathbb{R}^d} f \\
&= 0
\end{aligned} \tag{2}$$

This is true for arbitrary $\lambda > 0$. Therefore we conclude that $m(\{x \in \mathbb{R}^d : f(x) > 0\}) = 0$ is a null set. Therefore, $f = 0$ almost everywhere. $\square$

1.4.4 Absolute Integrability

Now we define the absolutely convergent Lebesgue interval.

Definition 1.4.7. *(Absolute integrability). An almost everywhere defined measurable function $f : \mathbb{R}^d \to \mathbb{C}$ is absolutely integrable if the unsigned integral*

$$||f||_{L^1(\mathbb{R}^d)} := \int_{\mathbb{R}^d} |f(x)| dx < \infty$$

Denote $L^1(\mathbb{R}^d)$ to be the space of absoluely integrable functions.
If f is real valued and $||f||_{L^1(\mathbb{R}^d)} < \infty$, define its Lebesgue integral as

$$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} f_+ - \int_{\mathbb{R}^d} f_-$$

where $f_+ = \max\{f, 0\} \geq 0$ and $f_- = \max\{-f, 0\} \geq 0$.
If f is complex valued and $||f||_{L^1(\mathbb{R}^d)} < \infty$,

$$\begin{aligned}
f &= Ref + iImf \\
&= (Ref)_+ - (Ref)_- + i[(Imf)_+ - (Imf)_-] \\
&= f_1 - f_2 + if_3 - if_4
\end{aligned}$$

where $f_1, f_2, f_3, f_4 : \mathbb{R}^d \to [0, +\infty]$.

When $||f||_{L^1(\mathbb{R}^d)} < \infty$, we can extend unsigned Lebesgue integral to such f by linearity:

$$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} f_1 - \int_{\mathbb{R}^d} f_2 + i \int_{\mathbb{R}^d} f_2 - i \int_{\mathbb{R}^d} f_4$$

Proposition 1.4.1. *(Integration is a linear operation from $L^1(\mathbb{R}^d)$ to $\mathbb{C}$).*

$$\int_{\mathbb{R}^d} (f+g) = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$$

$$\int_{\mathbb{R}^d} cf = c \int_{\mathbb{R}^d} f$$

for $c \in \mathbb{C}$.

From the pointwise triangle inequality

$$|f(x) + g(x)| \leq |f(x)| + |g(x)|$$

we have by monotonicity and linearity

$$\int_{\mathbb{R}^d} |f+g| \leq \int_{\mathbb{R}^d} (|f| + |g|) = \int_{\mathbb{R}^d} |f| + \int_{\mathbb{R}^d} |g|$$

that is,

$$||f+g||_{L^1(\mathbb{R}^d)} \leq ||f||_{L^1(\mathbb{R}^d)} + ||g||_{L^1(\mathbb{R}^d)}$$

Also, $\forall c \in \mathbb{C}$,

$$||cf||_{L^1(\mathbb{R}^d)} = |c| ||f||_{L^1(\mathbb{R}^d)}$$

Therefore we say $L^1(\mathbb{R}^d \to \mathcal{C})$ is a complex vector space.

$||.||_{L^1(\mathbb{R}^d)}$ is a *seminorm*, since $||f||_{L^1(\mathbb{R}^d)} = 0$ does not lead to $f \equiv 0$ (i.e. $f(x) = 0 \ \forall x \in \mathbb{R}^d$). Instead, lt leads to the following.

Proposition 1.4.2. $||f||_{L^1(\mathbb{R}^d)} = 0 \Rightarrow f = 0$ *a.e..*

Proof. From Markov's Inequality, for arbitrary $\lambda > 0$, we have pointwise

$$\lambda \mathbf{1}\{|f| \geq \lambda\} \leq |f|$$

Integrate both sides. Notice that LHS is a simple function.

$$\text{Simp} \int_{\mathbb{R}^d} \lambda \mathbf{1}\{|f| \geq \lambda\} \leq \sup_{0 \leq g \leq f|f|, g \text{ simple}} \text{Simp} \int_{\mathbb{R}^d} g = \underline{\int_{\mathbb{R}^d}} |f| = \int_{\mathbb{R}^d} f$$

$$m(\{|f| \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbb{R}^d} |f| = \frac{1}{\lambda} ||f||_{L^1(\mathbb{R}^d)}$$

If $||f||_{L^1(\mathbb{R}^d)} = 0$, then $m(\{|f| \geq \lambda\}) = 0$, which means that $f = 0$ a.e.. □

To be precise, $L^1(\mathbb{R}^d)$ is the normed space of equivalent functions.

Definition 1.4.8. *(Equivalent functions). Let $f, g \in L^1(\mathbb{R}^d \to \mathbb{C})$. $f \sim g$ if the L^1 distance*

$$d_{L^1}(f,g) := ||f-g||_{L^1(\mathbb{R}^d)} = 0$$

That is, $\{x \in \mathbb{R}^d : f(x) \neq g(x)\}$ is a null set.

Here we record another definition of distance.

Definition 1.4.9. *(Supremum distance or infinite norm distance). Let $f, g \in D$, then*

$$d_\infty(f,g) = \sup_{x \in D} |f(x) - g(x)|$$

It measures how close two functions are uniformly.

If $d_\infty(f, f_n) \to 0$ as $n \to \infty$, then $f_n \to f$ uniformly.
We also record another basic inequality.

Lemma 1.4.4. *(Triangle inequality). Let $f \in L^1(\mathbb{R}^d \to \mathbb{C})$, then*

$$|\int_{\mathbb{R}^d} f(x)dx| \leq \int_{\mathbb{R}^d} |f(x)|dx$$

Proof. If f is real valued, then by definition

$$\begin{aligned} |\int_{\mathbb{R}^d} f| &= |\int_{\mathbb{R}^d} f_+ - \int_{\mathbb{R}^d} f_-| \\ &\leq |\int_{\mathbb{R}^d} f_+| + |\int_{\mathbb{R}^d} f_-| \\ &= \int_{\mathbb{R}^d} |f_+| + \int_{\mathbb{R}^d} |f_-| \\ &= \int_{\mathbb{R}^d} |f| \end{aligned} \tag{3}$$

If f is complex valued, then we use the fact that $\forall z \in \mathbb{C}$, $z = |z|e^{i\theta}$ for some $\theta \in (-\pi, \pi]$. Then,

$$|\int_{\mathbb{R}^d} f| = e^{-i\theta} \int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} e^{-i\theta} f$$

Taking real parts of both sides, we get

$$|\int_{\mathbb{R}^d} f| = \int_{\mathbb{R}^d} Re(e^{-i\theta} f)$$

Since

$$Re(e^{-i\theta} f) \leq |e^{-i\theta} f| = |f|$$

we have

$$|\int_{\mathbb{R}^d} f| \leq \int_{\mathbb{R}^d} |f|$$

□

1.4.5 Littlewood's Three Principles

Littlewood's Three Principles gives informal heuristics about the basic intuition of Lebesgue measure theory.

1. Measurable sets are "almost open".

2. Absolutely integrable functions are "almost continuous".

3. Pointwise convergent sequences of f_n are "almost uniformly convergent".

Here we see an instance of the second principle. The following theorem says that simple functions, step functions (finite linear combination of indicator function on box), and continuous and compactly supported functions are dense subsets of $L^1(\mathbb{R}^d)$ w.r.t. $L^1(\mathbb{R}^d)$ semi-metric.

Theorem 1.4.4. *(Approximation of L^1 functions). Let $f \in L^1(\mathbb{R}^d)$, $\epsilon > 0$. Then,*

1. *$\exists$ simple function $g \in L^1(\mathbb{R}^d)$, such that $||f - g||_{L^1(\mathbb{R}^d)} \leq \epsilon$.*

2. *$\exists$ step function $g \in L^1(\mathbb{R}^d)$, such that $||f-g||_{L^1(\mathbb{R}^d)} \leq \epsilon$. (step function $g = \sum_{i=1}^N c_i \mathbf{1}_{B_i}$ where B_i are boxes).*

3. *$\exists$ $g \in C_c(\mathbb{R}^d)$, such that $||f - g||_{L^1(\mathbb{R}^d)} \leq \epsilon$.*
   ($C_c(\mathbb{R}^d) := \{g : \mathbb{R}^d \to \mathbb{C} : g$ is continuous and compactly supported$\}]$. Compactly supported means that $\{x : g(x) \neq 0\}$ has compact closure, that is, is contained in a ball).

Proof. (1). When f is unsigned, by definition of Lebesgue integral,

$$\int_{\mathbb{R}^d} f = \sup_{0 \leq g \leq f, g \text{ simple}} \text{Simp} \int_{\mathbb{R}^d} g = \sup_{0 \leq g \leq f, g \text{ simple}} \int_{\mathbb{R}^d} g$$

Then $\exists g$ simple, such that

$$\int_{\mathbb{R}^d} g \geq \int_{\mathbb{R}^d} f - \epsilon$$

Also since

$$\int_{\mathbb{R}^d} f \geq \int_{\mathbb{R}^d} g$$

we have by linearity

$$\int_{\mathbb{R}^d} |f - g| \leq \int_{\mathbb{R}^d} (f - g) = \int_{\mathbb{R}^d} f - \int_{\mathbb{R}^d} g \leq \epsilon$$

(2). It suffices to consider f simple, since by (1), for general $f \in L^1(\mathbb{R}^d)$, it is within ϵ distance with a simple function, so we can then apply the Triangle Inequality.
Let $f \in L^1(\mathbb{R}^d)$ be simple,

$$f = \sum_{i=1}^N c_i \mathbf{1}_{E_i}$$

We approximate each $\mathbf{1}_{E_i}$ by a step function g_i to have

$$||f - \sum_{i=1}^N c_i g_i||_{L^1(\mathbb{R}^d)} = ||\sum_{i=1}^N c_i(\mathbf{1}_{E_i} - g_i)||_{L^1(\mathbb{R}^d)} \leq \sum_{i=1}^N |c_i|\epsilon$$

So, it suffices to consider $f = \mathbf{1}_{E_i}$ for E_i measurable, and approximate it using an elementary set A_i.
By Exercise 1.2.16. in Textbook, E_i differs from an elementary set by a set of arbitrarily

small Lebesgue outer measure. Then, $\exists A_i$ such that

$$\epsilon \geq m(E_i \Delta A_i) = \int_{\mathbb{R}^d} |\mathbf{1}_{E_i} - \mathbf{1}_{A_i}| = ||\mathbf{1}_{E_i} - \mathbf{1}_{A_i}||_{L^1(\mathbb{R}^d)}$$

Thus we have

$$||\sum_{i=1}^{N} c_i(\mathbf{1}_{E_i} - g_i)||_{L^1(\mathbb{R}^d)} = \sum_{i=1}^{N} |c_i| ||\mathbf{1}_{E_i} - \mathbf{1}_{A_i}||_{L^1(\mathbb{R}^d)} \leq \sum_{i=1}^{N} |c_i| \epsilon$$

Therefore we have the claim (2).

(3). Again, by (1), (2), linearity, and triangle inequality, it suffices to show (iii) for $f = \mathbf{1}_B$ for box B. Set the continuous and compactly supported function as

$$g(x) = \max\{1 - R dist(x, B), 0\} \text{ for } R \text{ large enough}$$

so that

$$\int_{\mathbb{R}^d} |f - g| \leq \epsilon$$

and then we have the claim.

□

Before we see the instances of the third principle, we need first define local uniformity.

Definition 1.4.10. *(Local Uniformity). $f_n : \mathbb{R}^d \to \mathbb{C}$ converges to $f : \mathbb{R}^d \to \mathbb{C}$ locally uniformly if $\forall E \subset \mathbb{R}^d$ bounded, $f_n \to f$ uniformly on E.*

For example,
(1). $f_n(x) = \frac{x}{n}$, $n = 1, 2, ...$, $f_n \to f$ locally but not globally uniformly.
(2). $\sum_{i=1}^{N} \frac{x^n}{n!} \to e^x$ locally but not globally uniformly.
(3). $f_n(x) = \begin{cases} \frac{1}{nx}\mathbf{1}_{x>0} & x \neq 0 \\ 0 & x = 0 \end{cases}$, $f_n \to f$ pointwise, but neither locally nor globally uniformly.

Uniform converge $\Rightarrow$ locally uniform converge $\Rightarrow$ pointwise converge $\Rightarrow$ pointwise a.e. converge.

Theorem 1.4.5. *(Egorov's theorem). Let $f_n : \mathbb{R}^d \to \mathbb{C}$ converge pointwise a.e. to $f : \mathbb{R}^d \to \mathbb{C}$. f_n and f are measurable. Then, $\forall \epsilon > 0$, $\exists A \subset \mathbb{R}^d$ measurable with $m(A) < \epsilon$ and $f_n \to f$ locally uniformly on $\mathbb{R}^d \setminus A$.*

Proof. We may assume $f_n \to f$ pointwise everywhere by modifying them on a set of measure zero (absorbing $\{x : f_n(x) \not\to f(x)\}$ into A at the end of the statement).
Thus, $\forall x \in \mathbb{R}^d$, $\forall m > 0$, $\exists N(x, m) \in \mathbb{N}$ (this means that the choice of N depends on x and m) such that

$$|f_n(x) - f(x)| \leq \frac{1}{m} \ \forall n > N(x, m)$$

Write this set-theoretically, we have for each m,

$$E_{N,m} = \{x \in \mathbb{R}^d : |f_n(x) - f(x)| > \frac{1}{m} \text{ for some } n \geq N\}$$

as our "bad set", and

$$\bigcap_{N\in\mathbb{N}} E_{N,m} = \emptyset \ \forall m > 0$$

It is clear that $E_{N,m}$ is Lebesgue measurable and monotonic decreasing in N. Applying the downward monotone convergence theorem, for each m we have

$$m(E_{N,m} \cap B(0,R)) \to m(\emptyset) = 0 \ \forall R \in (0, +\infty)$$

That is, $\forall m \geq 1$, $\exists N_m$ such that

$$m(E_{N,m} \cap B(0,m)) \leq \frac{\epsilon}{2^m} \ \forall N \geq N_m$$

(This holds when $N = N_m$).
Let $A = \bigcup_{m=1}^{\infty} E_{N_m,m} \cap B(0,m)$, then

$$m(A) \leq \sum_{m=1}^{\infty} m(E_{N_m,m} \cap B(0,m)) = \epsilon$$

Now, we have for any $m \geq 1$ and $x \in B(0,R) \setminus A$, $x \neq E_{N_m,m}$ which means $|f_n(x) - f(x)| \leq \frac{1}{m}$. Since we have the choice of N_m not depending on x, $f_n \to f$ uniformly on $B(0,R) \setminus A$. Since every bounded set in $\mathbb{R}^d$ is in such a ball, we have the claim. □

Remark. This is not true if we don't have local uniformity. For example, "travelling bump" $f_n = \mathbf{1}_{n,n+1}$ converges to 0 pointwise, but not uniformly, not even if we delete any set of finite measure.

Now we look at another version of Littlewood's second principle (absolute integrable functions are almost continuous).

Theorem 1.4.6. *(Lusin's Theorem). Let $f : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable. Then $\forall \epsilon > 0$, $\exists E \subset \mathbb{R}^d$ with $m(E) \leq \epsilon$ such that the restriction of f to $\mathbb{R}^d \setminus E$ is continuous.*

Remark. We need the restriction of f rather than f itself. For example, $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ is not continuous on $\mathbb{R} \setminus E$ for $E \subset [0,1]$ with finite measure, but it is continuous on $\mathbb{R} \setminus E$ if we take $E := \mathbb{Q}$.

Proof. We use Egorov's Theorem and the third version of Littlewood's second principle in this proof.
By the density of $C_c(\mathbb{R}^d)$ in $L^1(\mathbb{R}^d)$, let $\epsilon > 0$, then $\forall n \in \mathbb{N}$, $\exists f_n \in C_c(\mathbb{R}^d)$ such that $||f_n - f||_{L^1(\mathbb{R}^d)} \leq \frac{\epsilon}{4^n}$.
By Markov's Inequality, for

$$E_n = \{x \in \mathbb{R}^d : |f_n(x) - f(x)| > \frac{1}{2^{n-1}}\}$$

we have

$$m(E_n) \leq 2^{n-1} ||f_n - f||_{L^1(\mathbb{R}^d)} \leq \frac{\epsilon}{2^{n+1}}$$

Let $E := \bigcup_{n=1}^{\infty} E_n$, then it is measurable and

$$m(E) \leq \sum_{i=1}^{\infty} m(E_n) = \frac{\epsilon}{2}$$

and $f_n \to f$ uniformly on $\mathbb{R}^d \setminus E$ by Egorov's theorem and compact-support.
Since the uniform limit of continuous functions is continuous, we have f is continuous on $\mathbb{R}^d \setminus E$. $\square$

Proposition 1.4.3. *(Littlewood-like principles). The following facts are not, strictly speaking, instances of any of Littlewood's three principles, but are in a similar spirit.*

1. *(Absolutely integrable functions almost have bounded support) Let $f : \mathbb{R}^d \to \mathbb{C}$ be an absolutely integrable function, and let $\varepsilon > 0$. Then there exists a ball $B(0,R)$ outside of which f has an L^1 norm of at most ε, or in other words that*

$$\int_{\mathbb{R}^d \setminus B(0,R)} |f(x)|\, dx \leq \varepsilon.$$

2. *(Measurable functions are almost locally bounded) Let $f : \mathbb{R}^d \to \mathbb{C}$ be a measurable function supported on a set of finite measure, and let $\varepsilon > 0$. Then there exists a measurable set $E \subset \mathbb{R}^d$ of measure at most ε outside of which f is locally bounded, or in other words that for every $R > 0$ there exists $M < \infty$ such that $|f(x)| \leq M$ for all $x \in B(0,R) \setminus E$.*

Proof. (1). Since f is absolutely integrable, there exists a continuous, compactly supported g such that $||f - g||_{L^1(\mathbb{R}^d)} \leq \epsilon$. Since g is compactly supported, $\exists R > 0$ such that $B(0,R)$ contains the support of g. Then, $g(x) = 0\ \forall x \in \mathbb{R}^d \setminus B(0,R)$. Therefore, by linearity and non-negativity of Lebesgue integral,

$$\begin{aligned} ||f - g||_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} |f - g| \\ &= \int_{\mathbb{R}^d \setminus B(0,R)} |f - g| + \int_{B(0,R)} |f - g| \\ &= \int_{\mathbb{R}^d \setminus B(0,R)} |f| + \int_{B(0,R)} |f - g| \\ &\geq \int_{\mathbb{R}^d \setminus B(0,R)} |f| \end{aligned}$$

Since the LHS $\leq \epsilon$, we have $\int_{\mathbb{R}^d \setminus B(0,R)} |f| \leq \epsilon$.

(2). **Exercise 1.3.23** says that the hypothesis in Lusin's theorem can be modified to be that f being measurable and finite everywhere (or a.e.). Then, there exists a measurable set $E \subset \mathbb{R}^d$ with $m(E) \leq \epsilon$, such that f is continuous on $\mathbb{R}^d \setminus E$.
We can enlarge E to an open set F such that $E \subset F$, $m(F) \leq m(E) + \epsilon \leq 2\epsilon$. Then $\mathbb{R}^d \setminus F \subset \mathbb{R}^d \setminus E$, and f is continuous on $\mathbb{R}^d \setminus F$.
$\forall R > 0$, $B(0,R) \setminus F = B(0,R) \cap (\mathbb{R}^d \setminus F)$ and we can see that it is closed and bounded by $B(0,R)$ itself. (Here the ball should be a closed ball and contains its boundary). Then, $B(0,R) \setminus F$ is compact. Since $B(0,R) \setminus F \subset \mathbb{R}^d \setminus F$, f is continuous on $B(0,R) \setminus F$, and

it is uniformly continuous because of restricting to compact sub-support. Since a uniformly continuous function is bounded, $\exists M > 0$, $|f| \leq M$. $\square$

1.5 Abstract Measure Spaces

Now we study the measure and integration on a general space $\mathcal{X}$. Generally, a measurable space is specified by the followings
(1). a set $\mathcal{X}$.
(2). $\mathcal{B}$ a collection of subsets of $\mathcal{X}$ that are "measuable".
(3). a mapping $\mu : \mathcal{B} \to [0, +\infty]$ as a measure.

1.5.1 σ-algebra.

Definition 1.5.1. *(Boolean Algebra). Let $\mathcal{X}$ be a set. A Boolean Algebra on $\mathcal{X}$ is a collection of subsets of $\mathcal{X}$ such that*
(1). $\emptyset \in \mathcal{B}$.
(2). $E \in \mathcal{B} \Rightarrow \mathcal{X} \setminus E \in \mathcal{B}$.
(3). (Finite Unions). $E, F \in \mathcal{B} \Rightarrow E \cup F \in \mathcal{B}$.

Generalize the (3) of Boolean Algebra, we have σ-algebra.

Definition 1.5.2. *(σ-algebra). Let $\mathcal{X}$ be a set. A σ-algebra on $\mathcal{X}$ is a collection os subsets of $\mathcal{X}$ such that*
(1). $\emptyset \in \mathcal{B}$.
(2). $E \in \mathcal{B} \Rightarrow \mathcal{X} \setminus E \in \mathcal{B}$.
(3). (Countable Unions). $E_1, E_2, ... \in \mathcal{B} \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$.

The pair $(\mathcal{X}, \mathcal{B})$ of a set $\mathcal{X}$ together with a σ-algebra on that set is a *measurable space*.

For σ-algebras $\mathcal{B}$ and $\mathcal{B}'$ with $\mathcal{B} \subseteq \mathcal{B}'$, we say that $\mathcal{B}'$ *refines* $\mathcal{B}$ or is a *refinement* of $\mathcal{B}$. Or we say $\mathcal{B}$ is *coarser* than $\mathcal{B}'$ or is a *coersening* of $\mathcal{B}'$.

Here we give some examples of σ-algebras.
(1). $\mathcal{L}(\mathbb{R}^d)$ is a σ-algebra on $\mathbb{R}^d$, but $\mathcal{J}(\mathbb{R}^d)$ and $\mathcal{E}(\mathbb{R}^d)$ are not.
(2). Trivial σ-algebra: $\{\emptyset, \mathcal{X}\}$
(3). Discrete σ-algebra: $\{E : E \subset \mathcal{X}\} = 2^{\mathcal{X}}$ as power set.
(4). Null algebra: $\mathcal{N}(\mathbb{R}^d) = \{E \subset \mathbb{R}^d : m(E) = 0 \text{ or } m(\mathbb{R}^d \setminus E) = 0\}$.
(5). Atomic σ-algebra: given a partition of $\mathcal{X}$ into disjoint sets $(A_\alpha)_{\alpha \in I}$ as "atoms", let $\mathcal{B} = \{\bigcup_{\alpha \in J} A_\alpha : J \subset I\}$.

Proposition 1.5.1. *(Intersection of σ-algebras). The intersection*

$$\bigwedge_{\alpha \in I} \mathcal{B}_\alpha := \bigcap_{\alpha \in I} \mathcal{B}_\alpha$$

of an arbitrary (and possibly infinite or uncountable) number of σ-algebras $\mathcal{B}_\alpha$ is again a σ-algebra, and is the finest σ-algebra that is coarser than all of the $\mathcal{B}_\alpha$.

Proof. Here we prove by verifying the definition.
(i). Since $\emptyset \in \mathcal{B}_\alpha\ \forall \alpha \in I$, we have $\emptyset \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$.
(ii). If $E \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$, then $E \in \mathcal{B}_\alpha\ \forall \alpha \in I$. Since each $\mathcal{B}_\alpha$ is a σ-algebra, we have $E^C \in \mathcal{B}_\alpha\ \forall \alpha \in I$. Therefore, $E^C \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$.
(iii). If $E_1, E_2, ... \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$, then $E_1, E_2, ... \in \mathcal{B}_\alpha\ \forall \alpha \in I$. Then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_\alpha\ \forall \alpha \in I$. Then $\bigcup_{n=1}^{\infty} E_n \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$.

Therefore, $\bigwedge_{\alpha\in I}\mathcal{B}_\alpha$ verifies the definition, and it is a σ-algebra.
Suppose that another σ-algebra $\mathcal{B}'$ is coarser than all of the $\mathcal{B}_\alpha$, that is, $\mathcal{B}' \subseteq \mathcal{B}_\alpha$ $\forall\alpha \in I$. Then, $\mathcal{B}' \subseteq \bigwedge_{\alpha\in I}\mathcal{B}_\alpha$. This way, $\bigwedge_{\alpha\in I}\mathcal{B}_\alpha$ is finer than any σ-algebra that is coarser than all of the $\mathcal{B}_\alpha$. □

Definition 1.5.3. *(Generation of σ-algebras). For a family of sets $\mathcal{F} \subset 2^{\mathcal{X}}$, define $\langle\mathcal{F}\rangle$, the σ-algebra generated by $\mathcal{F}$, as the intersection of all σ-algebras that contains $\mathcal{F}$.*

Remark. If $\mathcal{F}$ is a family of sets in $\mathcal{X}$. $P(E)$ is a property of sets $E \subset \mathcal{X}$ which obeys the following axioms
(1). $P(\emptyset)$ is true.
(2). $P(E)$ is true for all $E \in \mathcal{F}$.
(3). If $P(E)$ is true for some $E \in \mathcal{X}$, then $P(\mathcal{X} \setminus E)$ is true also.
(4). If $E_1, E_2, ... \in \mathcal{X}$ are such that $P(E_n)$ is true for all n, then $P(\bigcup_{n=1}^{\infty} E_n)$ is true also.
Then we conclude that $P(E)$ is true for all $E \in \langle\mathcal{F}\rangle$.

Definition 1.5.4. *(Borel σ-algebra). The Borel σ-algebra $\mathcal{B}[\mathcal{X}]$ on a topological space $(\mathcal{X}, \mathcal{T})$ is the σ-agebra generated by $\mathcal{T}$, the class of open sets in $\mathcal{X}$. Elements of $\mathcal{B}[\mathcal{X}]$ are called Borel measurable.*

Proposition 1.5.2. *(Generation of Borel σ-algebras). The Borel σ-algebra $\mathcal{B}[\mathbb{R}^d]$ of a Euclidean set is generated by any of the following collections of sets:*

(i) The open subsets of $\mathbb{R}^d$.

(ii) The closed subsets of $\mathbb{R}^d$.

(iii) The compact subsets of $\mathbb{R}^d$.

(iv) The open balls of $\mathbb{R}^d$.

(v) The boxes in $\mathbb{R}^d$.

(vi) The elementary sets in $\mathbb{R}^d$.

Proof. First, by definition, $\mathcal{B}[\mathbb{R}^d]$ is the σ-algebra generated by the open subsets of $\mathbb{R}^d$. This means that $\mathcal{B}[\mathbb{R}^d]$ is the intersection of all the σ-algebras that contains the open subsets of $\mathbb{R}^d$. To prove that $\mathcal{B}[\mathbb{R}^d]$ is generated by another family of sets is to prove that, every σ-algebra that contains the open subsets of $\mathbb{R}^d$ also contains that family of sets, and vice versa. This logic can be generalized to show that $\mathcal{B}[\mathbb{R}^d]$ can be generated by other family of sets from (ii) to (vi).
(i)-(ii). Let $\mathcal{B}$ be any σ-algebra containing all the open subsets of $\mathbb{R}^d$. Let E be any closed set in $\mathbb{R}^d$. Then E^C is open, and $E^C \in \mathcal{B}$. Then, $E = (E^C)^C \in \mathcal{B}$.
Let $\mathcal{B}$ be any σ-algebra containing all the closed subsets of $\mathbb{R}^d$. Let E be any open set in $\mathbb{R}^d$. Then E^C is closed, and $E^C \in \mathcal{B}$. Then, $E = (E^C)^C \in \mathcal{B}$.
(ii)-(iii). Let $\mathcal{B}$ be any σ-algebra containing all the closed subsets of $\mathbb{R}^d$. Let E be any compact set in $\mathbb{R}^d$. Then, by Heine-Borel Theorem, E is closed and $E \in \mathcal{B}$.
Let $\mathcal{B}$ be any σ-algebra containing all the compact subsets of $\mathbb{R}^d$. Let E be any closed set in $\mathbb{R}^d$. Then, $\forall n \in \mathbb{N}$, $E \cap \overline{B(0,n)}$ is bounded and closed and thus compact. Then $E \cap \overline{B(0,n)} \in \mathcal{B}$. By countable unions, $\bigcup_{n=1}^{\infty} E \cap \overline{B(0,n)} = E \in \mathcal{B}$.

(iii)-(iv). Let $\mathcal{B}$ be any σ-algebra containing all the compact subsets of $\mathbb{R}^d$. Let E be any open balls in $\mathbb{R}^d$. By Lemma 1.2.11, E can be written as the countable union of almost disjoint closed boxes $E = \bigcup_{n=1}^{\infty} B_n$. Since B_n $\forall n \in \mathbb{N}$ is closed and bounded, it is compact, then $B_n \in \mathcal{B}$ $\forall n \in \mathbb{N}$. By countable unions, $E = \bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$.
Let $\mathcal{B}$ be any σ-algebra containing all the open balls of $\mathbb{R}^d$. Let E be any compact set in $\mathbb{R}^d$. Then F^C is open. Since every open set can be written as a countable union of open balls, by countable unions, $F^C \in \mathcal{B}$, and then $F = (F^C)^C \in \mathcal{B}$.
(iv)-(v). Let $\mathcal{B}$ be any σ-algebra containing all the open balls of $\mathbb{R}^d$. Let B be any box in $\mathbb{R}^d$. Then, B can be written as the intersection of an open box B' and a closed box B'', $B = B' \cap B''$. Since $B^C = B'^C \cup B''^C$ and B'^C is closed and B''^C is open, which means that each of them is in $\mathcal{B}$ as we have shown in (i) and (ii), by countable unions, $B^C \in \mathcal{B}$. Therefore, $B = (B^C)^C \in \mathcal{B}$.
Let $\mathcal{B}$ be any σ-algebra containing all the boxes of $\mathbb{R}^d$. Let E be any open ball in $\mathbb{R}^d$. Then by Lemma 1.2.11, $E = \bigcup_{n=1}^{\infty}$ where $B_1, B_2, \ldots$ are almost disjoint boxes. Since $B_1, B_2, \ldots \in \mathcal{B}$, by countable unions, $E \in \mathcal{B}$.
(v)-(vi). Let $\mathcal{B}$ be any σ-algebra containing all the boxes of $\mathbb{R}^d$. Let E be any elementary set in $\mathbb{R}^d$. Then, by definition, $E = \bigcup_{i=1}^{n} B_i$ where B_i are boxes. Then, by countable unions, $E \in \mathcal{B}$.
Let $\mathcal{B}$ be any σ-algebra containing all the elementary sets in $\mathbb{R}^d$. Let B be any box in $\mathbb{R}^d$. Then B itself is an elementary set and then $B \in \mathcal{B}$. □

Proposition 1.5.3. *Let E be a Borel measurable subset of $\mathbb{R}^{d_1+d_2}$.*

(i) For any $x_1 \in \mathbb{R}^{d_1}$, the slice

$$\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\}$$

is a Borel measurable subset of $\mathbb{R}^{d_2}$. Similarly, for every $x_2 \in \mathbb{R}^{d_2}$, the slice

$$\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\}$$

is a Borel measurable subset of $\mathbb{R}^{d_1}$.

(ii) Give a counterexample to show that this claim is not *true if "Borel" is replaced with "Lebesgue" throughout.*
(Hint: the Cartesian product of any set with a point is a null set, even if the first set was not measurable.)

Proof. The proof uses Remark 1.4.15.
(i). Let $\mathcal{F}$ be the family of open sets in $\mathbb{R}^{d_1+d_2}$, let $E \subset \mathbb{R}^{d_1+d_2}$, then $P(E) : \forall x_1 \in \mathbb{R}^{d_1}, \{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ is a property of set E.
(i). We verify that $P(\emptyset) : \forall x_1 \in \mathbb{R}^{d_1}, \{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in \emptyset\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ is true. Since $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in \emptyset\} = \emptyset \in \mathcal{B}[\mathbb{R}^{d_2}]$ by definition of σ-algebra, $P(\emptyset)$ is true.
(ii). We verify that $P(E)$ is true for all $E \in \mathcal{F}$. Since $E \in \mathcal{F}$ is an open set, $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\}$ is an open set in $\mathbb{R}^{d_2}$, then $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ by the definition of Borel σ-algebra. Thus $P(E)$ is true for all $E \in \mathcal{F}$.
(iii). We verity that if $P(E)$ is true for some $E \subset \mathbb{R}^{d_1+d_2}$, then $P(\mathbb{R}^{d_1+d_2} \setminus E)$ is true also. Since $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ is true for some $E \in \mathbb{R}^{d_1+d_2}$, $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in$

$\mathbb{R}^{d_1+d_2} \setminus E\} = \mathbb{R}^{d_2} \setminus \{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ by the definition of σ-algebra.
(iv). We verify that if $E_1, E_2, ... \subset \mathbb{R}^{d_1+d_2}$ are such that $P(E_n)$ is true for all n, then $P(\bigcup_{n=1}^{\infty} E_n)$ is true also. Since $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E_n\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ is true for all n, then $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in \bigcup_{n=1}^{\infty} E_n\} = \bigcup_{n=1}^{\infty}\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E_n\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ by the definition of σ-algebra.
Thus, $P(E) : \{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_2}]$ is true for all $E \in \langle \mathcal{F} \rangle = \mathcal{B}[\mathbb{R}^{d_1+d_2}]$.

Let $\mathcal{F}$ be the family of open sets in $\mathbb{R}^{d_1+d_2}$, let $E \subset \mathbb{R}^{d_1+d_2}$, then $P(E) : \forall x_2 \in \mathbb{R}^{d_2}, \{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ is a property of sets E.
(i). We verify that $P(\emptyset) : \forall x_2 \in \mathbb{R}^{d_2}, \{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in \emptyset\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ is true. Since $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in \emptyset\} = \emptyset \in \mathcal{B}[\mathbb{R}^{d_1}]$ by definition of σ-algebra, $P(\emptyset)$ is true.
(ii). We verify that $P(E)$ is true for all $E \in \mathcal{F}$. Since $E \in \mathcal{F}$ is an open set, $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\}$ is an open set in $\mathbb{R}^{d_1}$, then $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ by the definition of Borel σ-algebra. Thus $P(E)$ is true for all $E \in \mathcal{F}$.
(iii). We verity that if $P(E)$ is true for some $E \subset \mathbb{R}^{d_1+d_2}$, then $P(\mathbb{R}^{d_1+d_2} \setminus E)$ is true also. Since $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ is true for some $E \subset \mathbb{R}^{d_1+d_2}$, $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in \mathbb{R}^{d_1+d_2} \setminus E\} = \mathbb{R}^{d_1} \setminus \{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d}]$ by the definition of σ-algebra.
(iv). We verify that if $E_1, E_2, ... \subset \mathbb{R}^{d_1+d_2}$ are such that $P(E_n)$ is true for all n, then $P(\bigcup_{n=1}^{\infty} E_n)$ is true also. Since $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E_n\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ is true for all n, then $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in \bigcup_{n=1}^{\infty} E_n\} = \bigcup_{n=1}^{\infty}\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E_n\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ by the definition of σ-algebra.
Thus, $P(E) : \{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\} \in \mathcal{B}[\mathbb{R}^{d_1}]$ is true for all $E \in \langle \mathcal{F} \rangle = \mathcal{B}[\mathbb{R}^{d_1+d_2}]$.

(ii). First, we construct a set F such that $F \notin \mathcal{L}(\mathbb{R})$, following Proposition 1.2.18.
$(\mathbb{Q}, +)$ is a subgroup of $(\mathbb{R}, +)$, and it partitions $\mathbb{R}$ into disjoint cosets $x + \mathbb{Q}$ for $x \in \mathbb{R}$. Then, this creates a quotient group $\mathbb{R}/\mathbb{Q} = \{x + \mathbb{Q} : x \in \mathbb{R}\}$. Each coset $C = x + \mathbb{Q} \in \mathbb{R}/\mathbb{Q}$ is dense in $\mathbb{R}$, so it has non-empty intersection with $[0, 1]$. By axiom of choice, select $x_c \in C \cap [0, 1]$ for each $C \in \mathbb{R}/\mathbb{Q}$ and put them together into a set $F = \{x_c : x_c \in C \cap [0, 1]\ \forall C \in \mathbb{R}/\mathbb{Q}\}$. We have already prove that $F \notin \mathcal{L}(\mathbb{R})$.
Since $\{0\} \times F$ is a null set, it is Lebesgue measurable and has measure zero. However, after slicing we are only left with F, which is not Lebesgue measurable. □

1.5.2 Countably Additive Measures and Measure Spaces

Definition 1.5.5. *(Measure and Measure Space). Let $(\mathcal{X}, \mathcal{B})$ be a measurable space. A measure on $(\mathcal{X}, \mathcal{B})$ is a map $\mu : \mathcal{B} \to [0, +\infty]$ that obeys*
(1). $\mu(\emptyset) = 0$.
(2). (σ-additivity): $E_1, E_2, ... \in \mathcal{B}$ are disjoint measuable sets, then

$$\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$$

Such a triple $(\mathcal{X}, \mathcal{B}, \mu)$ is called a measure space.

Here we give some examples of measure.
(1). Zero measure: $\mu(E) = 0\ \forall E \in \mathcal{B}$.
(2). Dirac measure: $\delta_x(E) := \mathbf{1}_E(x)$ for a given point $x \in \mathcal{X}$. It is a "point mass" for $E \in \mathcal{B}$.

(3). Counting measure: #:

$$\#(E) := \begin{cases} \text{cardinality of } E & \text{when finite} \\ +\infty & \text{else} \end{cases}$$

(4). Countable combination of measure μ_i

$$\mu = \sum_{i=1}^{\infty} c_i \mu_i \text{ for } c_i \in [0, +\infty]$$

where $(c_i\mu_i)(E) = c_i\mu_i(E)$, $\mu(E) = \sum_{i=1}^{\infty} c_i\mu_i(E)$.
Note that for countable $\mathcal{X}$, $\# = \sum_{x\in\mathcal{X}} \delta_x$, because $(\sum_{x\in\mathcal{X}} \delta_x)(E) = \sum_{x\in\mathcal{X}} \delta_x(E) = \sum_{x\in\mathcal{X}} \mathbf{1}_E(x) = \#(E)$
(5). Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable. For $f \in L^1(\mathbb{R}^d)$, set $\mu(E) = \int_E f dm$. For σ-additivity, we need to verify that for $E_1, E_2, ... \in \mathcal{L}(\mathbb{R}^d)$ disjoint,

$$\mu(\bigcup_{n=1}^{\infty} E_n) = \int_{\bigcup_{n=1}^{\infty} E_n} f dm = \sum_{n=1}^{\infty} \int_{E_n} f \mathbf{1}_{E_i} dm$$

Notice that

$$\int_{\bigcup_{n=1}^{\infty} E_n} f dm = \int_{\mathbb{R}^d} \sum_{i=1}^{\infty} f \mathbf{1}_{E_i} dm$$

We can exchange the sum according to the monotone convergence theorem later.

$$LHS = \sum_{i=1}^{\infty} \int_{\mathbb{R}^d} f \mathbf{1}_{E_i} dm$$

Proposition 1.5.4. *Let $(X, \mathcal{B}, \mu)$ be a measure space.*

(i). (Countable subadditivity) If $E_1, E_2, \dots$ are $\mathcal{B}$-measurable, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \le \sum_{n=1}^{\infty} \mu(E_n).$$

(ii). (Upwards monotone convergence) If $E_1 \subset E_2 \subset \cdots$ are $\mathcal{B}$-measurable, then

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

(iii). (Downwards monotone convergence) If $E_1 \supset E_2 \supset \cdots$ are $\mathcal{B}$-measurable, and $\mu(E_n) < \infty$ for at least one n, then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n) = \inf_n \mu(E_n).$$

Show that the downward monotone convergence claim can fail if the hypothesis that $\mu(E_n) < \infty$ for at least one n is dropped.

Proof. (i). It is obvious from De Morgan Law.

(ii).

$$\begin{aligned}
\mu(\bigcup_{n=1}^{\infty} E_n) &= \mu(E_1 \cup \bigcup_{n=2}^{\infty} (E_n \setminus E_{n-1})) \\
&= \mu(E_1) + \sum_{n=2}^{\infty} (\mu(E_n) - \mu(E_{n-1})) \text{ since disjoint} \\
&= \lim_{n \to \infty} \mu(E_n) \\
&= \sup_n \mu(E_n) \text{ since monotonicity}
\end{aligned}$$

(iii). Assume that $m(E_k) < \infty$ for some k. Since

$$E_k \setminus \bigcap_{n=k+1}^{\infty} E_n = \bigcup_{n=k+1}^{\infty} E_k \setminus E_n$$

we have

$$\begin{aligned}
\mu(E_k \setminus \bigcap_{n=k+1}^{\infty} E_n) &= \mu(\bigcup_{n=k+1}^{\infty} E_k \setminus E_n) \\
&= \lim_{n \to \infty} \mu(E_k) - \mu(E_n) \text{ by (ii)} \\
&= \mu(E_k) - \lim_{n \to \infty} \mu(E_n)
\end{aligned}$$

Since the LHS equals $\mu(E_k) - \mu(\bigcap_{n=k+1}^{\infty} E_n)$, we have

$$\mu(\bigcap_{n=k+1}^{\infty} E_n) = \lim_{n \to \infty} \mu(E_n)$$

Since $E_1 \supset E_2 \supset E_3 \supset ...$, we have $\bigcap_{n=k+1}^{\infty} E_n = \bigcap_{n=1}^{\infty} E_n$, and thus

$$\mu(\bigcap_{n=1}^{\infty} E_n) = \lim_{n \to \infty} \mu(E_n)$$

(iii). Consider the sequence $E_n := \mathbb{R}_+ \setminus [0, n]$ and the Lebesgue measure m. Clearly none of the $m(E_n)$ is finite. We have $m\left(\bigcap_{n=1}^{\infty} E_n\right) = m(\emptyset) = 0$. On the other hand,

$$\forall n \in \mathbb{N}, m(E_n) = \infty;$$

thus, the sequence of measures does not converge. $\square$

Proposition 1.5.5. *(Dominated convergence for sets). Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $E_1, E_2, \ldots$ be a sequence of $\mathcal{B}$-measurable sets that converge to another set E, in the sense that $\mathbf{1}_{E_n}$ converges pointwise to $\mathbf{1}_E$.*

(i). E is also $\mathcal{B}$-measurable.

(ii). If there exists a $\mathcal{B}$-measurable set F of finite measure (i.e., $\mu(F) < \infty$) that contains all of the E_n, then $\lim_{n \to \infty} \mu(E_n) = \mu(E)$. (Hint: Apply downward monotonicity to the sets $\bigcup_{n>N} (E_n \Delta E)$.)

(iii). The previous part of this exercise can fail if the hypothesis that all the E_n are contained in a set of finite measure is omitted.

Proof. (i). Since $\mathbf{1}_{E_n} \to \mathbf{1}_E$ pointwise, we have

$$\mathbf{1}_E = \lim_{n\to\infty} \mathbf{1}_{E_n} = \limsup_n \mathbf{1}_{E_n} = \inf_{N\in\mathbb{N}} \sup_{n\geq N} \mathbf{1}_{E_n}$$

we have

$$E = \bigcap_{N\in\mathbb{N}} \bigcup_{n=1}^{N} E_n$$

Since $E_1, E_2, ... \in \mathcal{B}$, by definition of σ-algebra we have $E \in \mathcal{B}$.

(ii). First, by finite additivity

$$\mu(E) = \mu(E \setminus E_n) + \mu(E \cap E_n)$$

$$\mu(E_n) = \mu(E_n \setminus E) + \mu(E_n \cap E)$$

Thus

$$\begin{aligned}\mu(E) - \mu(E_n) &= \mu(E \setminus E_n) - \mu(E_n \setminus E)\\ &\leq \mu(E \setminus E_n) + \mu(E_n \setminus E)\\ &= \mu(E\Delta E_n)\\ &\leq \mu(\bigcup_{k=n}^{\infty} E\Delta E_k)\end{aligned}$$

By downward convergence theorem,

$$\mu(E) - \lim_{n\to\infty} \mu(E_n) \leq \lim_{n\to\infty} \mu(\bigcup_{k=n}^{\infty} E\Delta E_k) = \mu(\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{n} E\Delta E_k)$$

Also,

$$\begin{aligned}\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{n} E\Delta E_k &= \bigcap_{n=1}^{\infty} (E \cap (\bigcup_{k=1}^{n} E_k^C)) \cup (\bigcup_{k=1}^{n} E_k \cap E^C)\\ &= (E \cap (\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{n} E_k^C)) \cup ((\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{n} E_k) \cap E^C)\\ &= (E \cap E^C) \cup (E \cap E^C)\\ &= \emptyset\end{aligned}$$

Therefore,

$$\mu(E) = \lim_{n\to\infty} \mu(E_n)$$

(iii). Consider the sequence $E_n := \mathbb{R}_+ \setminus [0, n]$ and the Lebesgue measure m. Clearly none of the $m(E_n)$ is finite. We have $m\left(\bigcap_{n=1}^{\infty} E_n\right) = m(\emptyset) = 0$. On the other hand,

$$\forall n \in \mathbb{N}, m(E_n) = \infty;$$

thus, the sequence of measures does not converge.

$\square$

Definition 1.5.6. *(Null, sub-null, and completeness) A null set of a measurable space $(\mathcal{X}, \mathcal{B}, \mu)$ is defined to be a $\mathcal{B}$-measurable set with measure zero. A sub-null set is any subset of a null set. A measure space is complete if every sub-null set is a null set.*

Proposition 1.5.6. *(Completion) Let $(\mathcal{X}, \mathcal{B}, \mu)$ be a measure space, then $\exists!$ refinement $(\mathcal{X}, \overline{\mathcal{B}}, \overline{\mu})$ known as the completion of $(\mathcal{X}, \mathcal{B}, \mu)$, which is the coarsest refinement of $(\mathcal{X}, \mathcal{B}, \mu)$ that is complete. $\overline{\mathcal{B}} := \{B \cup U : B \in \mathcal{B}, U \in \overline{U}\}$ where $\overline{U}$ is the collection of sub-null sets of μ measure.*

Proof. First, notice that $\forall B \in \mathcal{B}$, $B \subset B \cup U$ for some $U \in \overline{U}$. Thus $\mathcal{B} \subseteq \overline{\mathcal{B}}$.

Next, we prove that $\overline{\mathcal{B}}$ is a σ-algebra. First, $\emptyset \in \overline{\mathcal{B}}$ is trivial. Second, if $E \in \overline{\mathcal{B}}$, then $E = B \cup U$ where $B \in \mathcal{B}$ and $U \in \overline{U}$ is a null set. Then $E^C = B^C \cap U^C = B^C = B^C \cup \emptyset \in \overline{\mathcal{B}}$. Thirs, if $E_1, E_2, ... \in \overline{\mathcal{B}}$, then $E_1 = B_1 \cup U_1, E_2 = B_2 \cup U_2, ..., \bigcup_{i=1}^{\infty} E_1 = (\bigcup_{i=1}^{\infty} B_i) \cup (\bigcup_{i=1}^{\infty} U_i) \in \overline{\mathcal{B}}$.

Next, we prove that it is the coersest complete refinement. Suppose that there is another refinement $\mathcal{B}'$ such that $\mathcal{B} \subseteq \mathcal{B}'$. Then, $\forall B' \in \overline{\mathcal{B}}$, $B' = B \cup U$ where $B \in \mathcal{B}$ and U is a sub-null set. By completeness, U is a null set which is $\mathcal{B}$ measurable, and then $B' \in \mathcal{B} \subseteq \mathcal{B}'$. Therefore $\overline{\mathcal{B}} \subseteq \mathcal{B}'$.

Finally, we prove uniqueness. Suppose that $\exists \mathcal{B}''$ as another coersist complete refinement, then for any complete refinement including $\overline{\mathcal{B}}$, $\mathcal{B}'' \subseteq \overline{\mathcal{B}}$. However, we just proved that $\overline{\mathcal{B}} \subseteq \mathcal{B}'$ for all complete refinement $\mathcal{B}'$. Therefore, $\overline{\mathcal{B}} = \mathcal{B}''$. $\square$

Proposition 1.5.7. *The Lebesgue measure space $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d], m)$ is the completion of the Borel measure space $(\mathbb{R}^d, \mathcal{B}[\mathbb{R}^d], m)$.*

Proof. According to the last proposition, the completion of $(\mathbb{R}^d, \mathcal{B}[\mathbb{R}^d], m)$ is $(\mathbb{R}^d, \overline{\mathcal{B}[\mathbb{R}^d]}, m)$ where

$$\overline{\mathcal{B}[\mathbb{R}^d]} = \{B \cup U : B \in \mathcal{B}[\mathbb{R}^d], U \in U' \text{ where } m(U') = 0\}$$

Thus, we need to show that $\overline{\mathcal{B}[\mathbb{R}^d]} = \mathcal{L}(\mathbb{R}^d)$.

Let $E \in \overline{\mathcal{B}[\mathbb{R}^d]}$, then $E = B \cup U$ as is defined above. Since $B \in \mathcal{B}[\mathbb{R}^d]$ and the Borel σ-algebra is coarser than the Lebesgue σ-algebra, we have $B \in \mathcal{L}(\mathbb{R}^d)$. By monotonicity of Lebesgue outer measure, we have $m^*(U) \leq m(U') = 0$, thus U is a Lebesgue null set and $U \in \mathcal{L}(\mathbb{R}^d)$. Then, $E = B \cup U \in \mathcal{L}(\mathbb{R}^d)$.

Let $E \in \mathcal{L}(\mathbb{R}^d)$. Then, by Exercise 1.2.19, E can be written as $E = \bigcap_{i=1}^{\infty} U_n \setminus N$ where U_n are open sets and N is a Lebesgue null set. $\bigcap_{i=1}^{\infty} U_n \in \mathcal{B}[\mathbb{R}^d]$ since every U_n is an open set. Since N is a Lebesgue null set, we can find a sequence of Borel measurable set V_n such that (i) V_n is contained in another Borel-measurable set F with finite measure, (ii) V_n converge to N, and (iii) $m(V_n) \leq \frac{1}{n}$ $\forall n \in \mathbb{N}$. Then, by dominated convergence for sets, $N \in \mathcal{B}[\mathbb{R}^d]$ and $m(N) = \lim_{n \to \infty} m(V_n) = 0$. Thus, $E = \bigcap_{n=1}^{\infty} U_n \setminus N \in \mathcal{B}[\mathbb{R}^d]$ by finite Boolean operation.

Therefore $\overline{\mathcal{B}[\mathbb{R}^d]} = \mathcal{L}(\mathbb{R}^d)$.

$\square$

1.5.3 Measurable Functions and Integration.

Definition 1.5.7. *(Measurable function) Let $(\mathcal{X}, \mathcal{B})$ be a measurable space. Let $f : \mathcal{X} \to [0, +\infty]$ or $\mathbb{C}$ be an unsigned or complex-valued function. It is measurable if $f^{-1}(U)$ is $\mathcal{B}$-measurable for every open subset U of $[0, +\infty]$ or $\mathbb{C}$.*

Theorem 1.5.1. *(Egorov) Let $(\mathcal{X}, \mathcal{B}, \mu)$ be a finite measure space (so $\mu(\mathcal{X}) < \infty$), and let $f_n : X \to \mathbb{C}$ be a sequence of measurable functions that converge pointwise almost everywhere to a limit $f : \mathcal{X} \to \mathbb{C}$, and let $\varepsilon > 0$. Show that there exists a measurable set E of measure at most ε such that f_n converges uniformly to f outside of E. Give an example to show that the claim can fail when the measure μ is not finite.*

Proof. By modifying f_n and f on a set of measure zero (that can be absorbed into A at the end of the argument), we may assume that f_n converges pointwise everywhere on $\mathcal{X}$ to f. Thus, $\forall x \in \mathcal{X}, \forall m > 0, \exists N(x, m) \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| \leq \frac{1}{m} \ \forall n \geq N(x, m)$$

Write this set-theoretically, we have for each m:

$$E_{N,m} = \{x \in \mathcal{X} : |f_n(x) - f(x)| > \frac{1}{m} \text{ for some } n \geq N\}$$

and

$$\bigcap_{N \in \mathbb{N}} E_{N,m} = \emptyset$$

By Exercise 1.4.29 in textbook, we know that f is measurable, $|f_n - f|$ is also measurable. By the same exercise we also know that $E_{N,m}$ as a level set is also $\mathcal{B}$-measurable. Also, $E_{N,m}$ is monotonically decreasing in N. Also, by monotonicity we have $\mu(E_n) \leq \mu(\mathcal{X}) < \infty$. Then, by downwards monotone convergence, we have for each m,

$$\mu(E_{N,m}) \to 0 \text{ as } N \to \infty$$

This means that, $\forall m \geq 1, \exists N_m$, such that when $N \geq N_m$,

$$\mu(E_{N,m}) \leq \frac{\epsilon}{2^m} \ \forall \epsilon > 0$$

Let $A = \bigcup_{m=1}^{\infty} E_{N,m}$, then by σ-subadditivity,

$$\mu(A) \leq \sum_{m=1}^{\infty} \mu(E_{N_m,m}) \leq \epsilon$$

Then, $\forall \epsilon > 0$, take $m = \frac{1}{\epsilon}$, then $\exists N_m$ such that when $x \in \mathcal{X} \setminus A$ (this means that $x \notin A$ and then $x \notin E_{N_m,m}$), we have

$$|f_n(x) - f(x)| \leq \frac{1}{m} = \epsilon \text{ for all } n \geq N_m$$

Thus, $f_n \to f$ uniformly on $\mathcal{X} \setminus A$ where $\mu(A) \leq \epsilon$.
Example: $f = \mathbf{1}_{[n,n+1]}$ for $n = 0, 1, 2, ...$ This function is on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+)), m)$, and $m(\mathbb{R}^+) =$

∞. The set that f does not converges uniformly to f always have measure one. $\square$

Definition 1.5.8. *(Unsigned simple function) An unsigned simple function on a measurable space $(\mathcal{X}, \mathcal{B})$ is a measurable function $f : \mathcal{X} \to [0, +\infty]$ taking finitely many values (possibly $+\infty$).*

Definition 1.5.9. *(Integral of unsigned simple function) For f taking values $a_1, a_2, ..., a_k \in [0, +\infty]$, and a measure μ on $(\mathcal{X}, \mathcal{B})$, define*

$$Simp \int_{\mathcal{X}} f d\mu := \sum_{j=1}^{k} a_j \mu(f^{-1}(a_j))$$

Note that for $a_1, ..., a_k$ distinct, $\exists$ distinct open $U_1, ..., U_k$, $a_i \in U_i$, such that $f^{-1}(a_i) = f^{-1}(U_i)$. Therefore $f^{-1}(a_i)$ is measurable.

Remark. A Property $P(x)$ of $x \in \mathcal{X}$ holds μ- almost everywhere if it holds outside of a sub-null set.

Proposition 1.5.8. *(Inclusion-exclusion principle). Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $A_1, \ldots, A_n$ be $\mathcal{B}$-measurable sets of finite measure. Show that*

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{J \subseteq \{1,\ldots,n\} : J \neq \emptyset} (-1)^{|J|-1} \mu\left(\bigcap_{i \in J} A_i\right).$$

(Hint: Compute Simp $\int_X \left(1 - \prod_{i=1}^n (1 - \mathbf{1}_{A_i})\right) d\mu$ in two different ways.)

Proof. We prove the inclusion-exclusion principle by compute the below simple integral in two different ways.

$$\text{Simp} \int_{\mathcal{X}} (1 - \prod_{i=1}^{n} (1 - \mathbf{1}_{A_i})) d\mu$$

The first way is to observe that, $(1 - \prod_{i=1}^n (1 - \mathbf{1}_{A_i})) = 1$ if $x \in A_i$ for some $i \in \{1, ..., n\}$. Thus, the function is equivalent to a simple function $f : \mathcal{X} \to \{0, 1\}$ where

$$f = \mathbf{1}_{\bigcup_{i=1}^n A_i}$$

Since $\bigcup_{i=1}^n A_i \in \mathcal{B}$,

$$\begin{aligned} \text{Simp} \int_{\mathcal{X}} (1 - \prod_{i=1}^{n} (1 - \mathbf{1}_{A_i})) d\mu &= \text{Simp} \int_{\mathcal{X}} \mathbf{1}_{\bigcup_{i=1}^n A_i} d\mu \\ &= \mu(\bigcup_{i=1}^{n} A_i) \end{aligned}$$

The second way is to expand $f = (1 - \prod_{i=1}^{n}(1 - \mathbf{1}_{A_i}))$.

$$\begin{aligned}
f &= 1 - \prod_{i=1}^{n}(1 - \mathbf{1}_{A_i}) \\
&= 1 - (1 - \sum_{J\subseteq\{1,...,n\},|J|=1} \prod_{i\in J} \mathbf{1}_{A_i} + \sum_{J\subseteq\{1,...,n\},|J|=2} \prod_{i\in J} \mathbf{1}_{A_i} + ... + (-1)^n \sum_{J\subseteq\{1,...,n\},|J|=n} \prod_{i\in J} \mathbf{1}_{A_i}) \\
&= (-1)^{1+1} \sum_{J\subseteq\{1,...,n\},|J|=1} \mathbf{1}_{\bigcap_{i\in J} A_i} + (-1)^{2+1} \sum_{J\subseteq\{1,...,n\},|J|=2} \mathbf{1}_{\bigcap_{i\in J} A_i} + ... + (-1)^{n+1} \sum_{J\subseteq\{1,...,n\},|J|=n} \mathbf{1}_{\bigcap_{i\in J} A_i} \\
&= \sum_{J\subseteq\{1,...,n\},|J|=1} (-1)^{1+1}\mathbf{1}_{\bigcap_{i\in J} A_i} + \sum_{J\subseteq\{1,...,n\},|J|=2} (-1)^{2+1}\mathbf{1}_{\bigcap_{i\in J} A_i} + ... + \sum_{J\subseteq\{1,...,n\},|J|=n} (-1)^{n+1}\mathbf{1}_{\bigcap_{i\in J} A_i} \\
&= \sum_{J\subseteq\{1,...,n\},J\neq\emptyset} (-1)^{|J|+1}\mathbf{1}_{\bigcap_{i\in J} A_i}
\end{aligned}$$

Since $\bigcap_{i\in J} A_i \in \mathcal{B}$, we have

$$\text{Simp}\int_{\mathcal{X}} f d\mu = \sum_{J\subseteq\{1,...,n\},J\neq\emptyset} (-1)^{|J|+1}\mu(\bigcap_{i\in J} A_i)$$

Therefore,

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{J\subseteq\{1,...,n\}:J\neq\emptyset} (-1)^{|J|-1}\mu\left(\bigcap_{i\in J} A_i\right).$$

□

Definition 1.5.10. *(Unsigned integral) For general measurable $f : \mathcal{X} \to [0, +\infty]$ on $(\mathcal{X}, \mathcal{B}, \mu)$, then unsigned integral is defined as*

$$\int_{\mathcal{X}} f d\mu := \sup_{0\leq g\leq f, g \text{ simple}} Simp\int_{\mathcal{X}} g d\mu$$

Proposition 1.5.9. *(Easy properties of the unsigned integral). Let $(\mathcal{X}, \mathcal{B}, \mu)$ be a measure space, and let $f, g : \mathcal{X} \to [0, +\infty]$ be measurable.*

(iii) ***(Homogeneity)*** *We have $\int_{\mathcal{X}} cf\, d\mu = c\int_{\mathcal{X}} f\, d\mu$ for every $c \in [0, +\infty]$.*

(iv) ***(Superadditivity)*** *We have $\int_{\mathcal{X}}(f + g)\, d\mu \geq \int_{\mathcal{X}} f\, d\mu + \int_{\mathcal{X}} g\, d\mu$.*

(vi) ***(Markov's inequality)*** *For any $0 < \lambda < \infty$, one has*

$$\mu(\{x \in X : f(x) \geq \lambda\}) \leq \frac{1}{\lambda}\int_X f\, d\mu.$$

In particular, if $\int_X f\, d\mu < \infty$, then the sets $\{x \in X : f(x) \geq \lambda\}$ have finite measure for each $\lambda > 0$.

Proof. (iii). If $c > 0$,

$$\begin{aligned}
\int_{\mathcal{X}} cf d\mu &= \sup_{0 \le g \le cf, g \text{ simple}} \text{Simp} \int_{\mathcal{X}} g d\mu \\
&= \sup_{0 \le \frac{g}{c} \le f, g \text{ simple}} \text{Simp} \int_{\mathcal{X}} g d\mu \text{ (let } g' := \frac{g}{c}\text{, then it is also simple)} \\
&= \sup_{0 \le g' \le f, g \text{ simple}} \text{Simp} \int_{\mathcal{X}} c g' d\mu \\
&= \sup_{0 \le g' \le f, g \text{ simple}} c \text{ Simp} \int_{\mathcal{X}} g' d\mu \text{ (by Exercise 1.4.33 (iii))} \\
&= c \sup_{0 \le g' \le f, g \text{ simple}} \text{Simp} \int_{\mathcal{X}} g' d\mu \\
&= c \int_{\mathcal{X}} f d\mu
\end{aligned} \tag{4}$$

If $c = 0$, then $LHS = \int_{\mathcal{X}} 0 d\mu = 0 = RHS$.
(iv). By the definition of unsigned integral,

$$\int_{\mathcal{X}} f d\mu = \sup_{0 \le f' \le f, f' \text{ simple}} \text{Simp} \int_{\mathcal{X}} f' d\mu$$

$$\int_{\mathcal{X}} g d\mu = \sup_{0 \le g' \le f, g' \text{ simple}} \text{Simp} \int_{\mathcal{X}} g' d\mu$$

Therefore, let $\epsilon > 0$, then there exists simple functions f' and g', $0 \le f' \le f$ and $0 \le g' \le f$, such that

$$\int_{\mathcal{X}} f d\mu \le \text{Simp} \int_X f' d\mu + \frac{\epsilon}{2}$$

$$\int_{\mathcal{X}} g d\mu \le \text{Simp} \int_X g' d\mu + \frac{\epsilon}{2}$$

Then,

$$\begin{aligned}
\int_{\mathcal{X}} f d\mu + \int_{\mathcal{X}} g d\mu &= \text{Simp} \int_X f' d\mu + \text{Simp} \int_X g' d\mu + \epsilon \\
&= \text{Simp} \int_X (f' + g') d\mu + \epsilon \text{ (by Exercise 1.4.33 (iv))} \\
&\le \sup_{0 \le h \le f+g} \text{Simp} \int_X h d\mu + \epsilon \text{ } (h = f' + g' \text{ is also simple)} \\
&= \int_{\mathcal{X}} (f+g) d\mu + \epsilon
\end{aligned}$$

Note that h is also simple because it also takes on only finitely many values.
Since $\epsilon > 0$ is arbitrary, we have $\int_{\mathcal{X}} (f+g) d\mu \ge \int_{\mathcal{X}} f d\mu + \int_{\mathcal{X}} g d\mu$.
(vi). Since $f : \mathcal{X} \to [0, +\infty]$ is measurable, $\{x \in \mathcal{X} : f(x) \ge \lambda\}$ is measurable. Notice that

$$\lambda \mathbf{1}_{\{x \in \mathcal{X} : f(x) \ge \lambda\}} \le f(x)$$

and that the LHS is a simple function. Then, by monotonicity,

$$\int_{\mathcal{X}} \lambda \mathbf{1}_{\{x \in \mathcal{X} d\mu : f(x) \geq \lambda\}} \leq \int_{\mathcal{X}} f(x) d\mu$$

By the definition of unsigned integral of simple function,

$$\lambda \mu(\{x \in \mathcal{X} : f(x) \geq \lambda\}) \leq \int_{\mathcal{X}} f(x) d\mu$$

Since $\lambda > 0$,

$$\mu(\{x \in \mathcal{X} : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathcal{X}} f(x) d\mu$$

□

Theorem 1.5.2. *(Finite additivity). Let $(\mathcal{X}, \mathcal{B}, \mu)$ be a measure space, and let $f, g : \mathcal{X} \to [0, +\infty]$ be measurable. Then*

$$\int_{\mathcal{X}} (f + g) \, d\mu = \int_{\mathcal{X}} f \, d\mu + \int_{\mathcal{X}} g \, d\mu.$$

Proof. In view of superadditivity, it suffices to establish the subadditivity property

$$\int_{\mathcal{X}} (f + g) \, d\mu \leq \int_{\mathcal{X}} f \, d\mu + \int_{\mathcal{X}} g \, d\mu.$$

We establish this in stages. We first deal with the case when μ is a *finite* measure (which means that $\mu(\mathcal{X}) < \infty$) and f, g are bounded. Pick an $\varepsilon > 0$, and let f_ε be f rounded down to the nearest integer multiple of ε, and f^ε be f rounded up to the nearest integer multiple. Clearly, we have the pointwise bounds

$$f_\varepsilon(x) \leq f(x) \leq f^\varepsilon(x),$$

and

$$f^\varepsilon(x) - f_\varepsilon(x) \leq \varepsilon.$$

Since f is bounded, f_ε and f^ε are simple. Similarly, define g_ε and g^ε. We then have the pointwise bound

$$f + g \leq f^\varepsilon + g^\varepsilon \leq f_\varepsilon + g_\varepsilon + 2\varepsilon,$$

hence by Exercise 1.4.36 and the properties of the simple integral,

$$\int_{\mathcal{X}} (f + g) \, d\mu \leq \int_{\mathcal{X}} (f_\varepsilon + g_\varepsilon + 2\varepsilon) \, d\mu = \mathrm{Simp} \int_{\mathcal{X}} f_\varepsilon \, d\mu + \mathrm{Simp} \int_{\mathcal{X}} g_\varepsilon \, d\mu + 2\varepsilon\mu(\mathcal{X}).$$

From the definition, we conclude that

$$\int_{\mathcal{X}} (f + g) \, d\mu \leq \int_{\mathcal{X}} f \, d\mu + \int_{\mathcal{X}} g \, d\mu + 2\varepsilon\mu(\mathcal{X}).$$

Letting $\varepsilon \to 0$ and using the assumption that $\mu(\mathcal{X})$ is finite, we obtain the claim.

Now we continue to assume that μ is a finite measure, but now do not assume that f, g

are bounded. Then for any natural number n, we can use the previous case to deduce that

$$\int_X \min(f,n) + \min(g,n)\, d\mu \leq \int_X \min(f,n)\, d\mu + \int_X \min(g,n)\, d\mu.$$

Since $\min(f+g,n) \leq \min(f,n) + \min(g,n)$, we conclude that

$$\int_X \min(f+g,n)\, d\mu \leq \int_X \min(f,n)\, d\mu + \int_X \min(g,n)\, d\mu.$$

Taking limits as $n \to \infty$ using vertical truncation, we obtain the claim.

Finally, we no longer assume that μ is a finite measure, and also do not require f, g to be bounded. If either $\int_X f\, d\mu$ or $\int_X g\, d\mu$ is infinite, then by monotonicity, $\int_X (f+g)\, d\mu$ is infinite as well, and the claim follows; so we may assume that $\int_X f\, d\mu$ and $\int_X g\, d\mu$ are both finite. By Markov's inequality (Exercise 1.4.36 (vi)), we conclude that for each natural number n, the set

$$E_n := \{x \in X : f(x) > \frac{1}{n}\} \cup \{x \in X : g(x) > \frac{1}{n}\}$$

has finite measure. These sets are increasing in n, and for $f, g, f+g$ supported on $\bigcup_{n=1}^{\infty} E_n$, we have, by horizontal truncation,

$$\int_X (f+g)\, d\mu = \lim_{n\to\infty} \int_X (f+g)1_{E_n}\, d\mu.$$

From the previous case, we have

$$\int_X (f+g)1_{E_n}\, d\mu = \int_X f1_{E_n}\, d\mu + \int_X g1_{E_n}\, d\mu.$$

Letting $n \to \infty$ and using horizontal truncation, we obtain the claim. □

Definition 1.5.11. *(Absolutely integrable). Let $(\mathcal{X}, \mathcal{B}, \mu)$ be a measure space. A measurable function $f : \mathcal{X} \to \mathcal{C}$ is absolutely integrable if*

$$||f||_{L^1(\mathcal{X},\mathcal{B},\mu)} := \int_{\mathcal{X}} |f| d\mu < \infty$$

Definition 1.5.12. *(Integral for measurable function) for general measurable function $f : \mathcal{X} \to \mathcal{C}$, define $\int_{\mathcal{X}} f d\mu$ by*

$$\int_{\mathcal{X}} f d\mu = \int_{\mathcal{X}} Ref_+ d\mu - \int_{\mathcal{X}} Ref_- d\mu + i\int_{\mathcal{X}} Imf_+ d\mu - i\int_{\mathcal{X}} Imf_- d\mu$$

1.5.4 The convergence theorems

We want to know what is the sufficient condition such that

$$\int_{\mathcal{X}} f_n d\mu \to \int_{\mathcal{X}} f d\mu$$

Pointwise convergent is insufficient. For example, the travelling bump $f_n = \mathbf{1}_{[n,n+1]}$, the flattening bump $f_n = \frac{1}{n}\mathbf{1}_{[0,n]}$, and the narrowing spike $f_n = n\mathbf{1}_{[0,\frac{1}{n}]}$ all converges pointwise to zero, but $\int_{\mathcal{X}} f_n d\mu = 1$.

Theorem 1.5.3. *(Monotone convergence theorem). Let $0 \leq f_1 \leq f_2 \leq ...$ be a pointwise monotone non-decreasing sequence of unsigned functions on a measure space $(\mathcal{X}, \mathcal{B}, \mu)$. Then*

$$\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu = \int_{\mathcal{X}} \lim_{n\to\infty} f_n d\mu$$

Proof. Write $f = \lim_{n\to\infty} f_n = \sup_n f_n$. Then, $f : \mathcal{X} \to [0, +\infty]$ is measurable. Since f_n are monotonic non-decreasing to f, $f_n \leq f$ $\forall n$, by monotonicity property of unsigned integral, the sequence of integral $\int_{\mathcal{X}} f_n d\mu \in [0, +\infty]$ is monotone and bounded by $\int_{\mathcal{X}} f d\mu \in [0. +\infty]$. So, $\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$ exists and $\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu \leq \int_{\mathcal{X}} f d\mu$.

It remains to show $\int_{\mathcal{X}} f d\mu \leq \lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$.

By definition, it suffices to show that

$$\int_{\mathcal{X}} g d\mu \leq \lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu \ (*)$$

for any simple function $g : 0 \leq g \leq f$.

Fix $g = \sum_{i=1}^{k} a_i \mathbf{1}_{E_i}$, $a_i \in [0, +\infty]$.

Case 1, g takes value $+\infty$ on a set E with $\mu(E) > 0$.

Then, $f = +\infty$ pointwise on set E. We will show that the RHS of $*$ is $+\infty$. Let $M \geq 1$ be arbitrary. $F_n = \{x \in E : f_n(x) \geq M\}$, then F_n is measurable and $F_n \uparrow E$ (i.e. $F_n \subset F_{n+1}$ $\forall n$, $E = \bigcup_{n=1}^{\infty} F_n$). By monotone convergence theorem, $\mu(F_n) \uparrow \mu(E)$. Then, $\exists N_M$, $\mu(F_n) \geq \mu(F_{N_M}) \geq \min\{\frac{\mu(E)}{2}, M\}$ when $n \geq N_M$.

$$\begin{aligned}\int_{\mathcal{X}} f_n d\mu &\geq \int_{\mathcal{X}} M \mathbf{1}_{F_n} d\mu \\ &= M\mu(F_n) \\ &\geq M \min\{\frac{\mu(E)}{2}, M\} \ \forall n \geq N_M\end{aligned}$$

$$\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu \geq M \min\{\frac{\mu(E)}{2}, M\}$$

Since M is arbitrary, $\int_{\mathcal{X}} f_n d\mu \to \infty$ in Case 1.

Case 2, $g < +\infty$ everywhere.

Let $\epsilon > 0$ and $E_{i,n} := \{x \in E_i : f_n \geq (1-\epsilon)a_i\}$. Then $E_{i,n}$ is measurable, $E_{i,n} \uparrow E_i$. By monotone convergence theorem: $\mu(E_{i,n}) \uparrow \mu(E_i)$. Integrating the pointwise inequality

$$f_n \geq \sum_{i=1}^{k} (1-\epsilon) a_i \mathbf{1}_{E_{i,n}}$$

$$\int_{\mathcal{X}} f_n d\mu \geq (1-\epsilon) \sum_{i=1}^{k} a_i \mu(E_{i,n})$$

$$\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu \geq \lim_{n\to\infty} (1-\epsilon) \sum_{i=1}^{k} a_i \mu(E_{i,n}) = (1-\epsilon) \int_{\mathcal{X}} g d\mu$$

Since $\epsilon > 0$ is arbitrary,

$$\lim_{n\to\infty}\int_{\mathcal{X}} f_n d\mu \geq \int_{\mathcal{X}} g d\mu$$

for any simple g such that $0 \leq g \leq f$.

Therefore the theorem holds. □

Corollary 1.5.1. *(Tonelli's theorem for sums and integrals) For $f_1, f_2, ... : \mathcal{X} \to [0, +\infty]$ unsigned measurable functions on $(\mathcal{X}, \mathcal{B}, \mu)$,*

$$\int_{\mathcal{X}} \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int_{\mathcal{X}} f_n d\mu$$

Proof. Apply monotone convergence theorem to $F_N = \sum_{n=1}^{N} f_n$. See that F_N is monotonic non-decreasing, and $\lim_{N\to\infty} F_N = \sum_{n=1}^{\infty} f_n$.

$$\begin{aligned}
\int_{\mathcal{X}} \sum_{n=1}^{\infty} f_n d\mu &= \int_{\mathcal{X}} \lim_{N\to\infty} F_N d\mu \\
&= \lim_{N\to\infty} \int_{\mathcal{X}} \sum_{n=1}^{N} f_n d\mu \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\mathcal{X}} f_n d\mu \\
&= \sum_{n=1}^{\infty} \int_{\mathcal{X}} f_n d\mu
\end{aligned}$$

□

If we don't have monotonicity, then we have the Fatou's lemma.

Lemma 1.5.1. *(Fatou's lemma) Let $f_1, f_2, ... : \mathcal{X} \to [0, +\infty]$ be unsigned measurable functions on $(\mathcal{X}, \mathcal{B}, \mu)$. Then,*

$$\int_{\mathcal{X}} \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$$

Proof. Let $F_N = \inf_{n\geq N} f_n$, then $\liminf_{n\to\infty} f_n = \sup_{N>0} \inf_{n\geq N} f_n = \sup_{N>0} F_N$. Since F_N is monotonic non-decreasing with N, we have $\sup_{N>0} F_N = \lim_{N\to\infty} F_N$. Thus by monotone convergence theorem,

$$\int_{\mathcal{X}} \sup_{N>0} F_N d\mu = \int_{\mathcal{X}} \lim_{N\to\infty} F_N d\mu = \lim_{N\to\infty} \int_{\mathcal{X}} F_N d\mu = \sup_{N>0} \int_{\mathcal{X}} F_N d\mu$$

Since $F_N = \inf_{n\geq N} f_n \leq f_n \; \forall n \geq N$, we have

$$\int_{\mathcal{X}} F_N d\mu \leq \int_{\mathcal{X}} f_n \; \forall n \geq N$$

$$\int_{\mathcal{X}} F_N d\mu \leq \inf_{n\geq N} \int_{\mathcal{X}} f_n d\mu$$

Therefore

$$\int_{\mathcal{X}} \sup_{N>0} F_N d\mu = \int_{\mathcal{X}} \liminf_{n\to\infty} f_n d\mu \leq \sup_{N>0} \inf_{n\geq N} \int_{\mathcal{X}} f_n d\mu = \liminf_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$$

□

Theorem 1.5.4. *(Dominated convergence theorem) Let $f_1, f_2, ...$ be a sequence of measurable $\mathbb{C}$-valued function on $(\mathcal{X}, \mathcal{B}, \mu)$ such that $f_n \to f$ μ-a.e. for some $f : \mathcal{X} \to \mathbb{C}$. Suppose $\exists G \in L^1(\mathcal{X}, \mathcal{B}, \mu)$ such that $|f_n| \leq G$ μ-a.e. $\forall n \in \mathbb{N}$, then,*

$$\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu = \int_{\mathcal{X}} f d\mu$$

Proof. By modifying f_n and f on a null set, we may assume $f_n \to f$ everywhere and $|f_n| \leq G$ everywhere. By taking the real and the imaginary parts we may assume without loss of generality that f_n and f are real. Thus $-G \leq f_n \leq G$ pointwise $\forall n \in \mathbb{N}$, and also $-G \leq f \leq G$ pointwise. By Fatou's lemma,

$$\int_{\mathcal{X}} \liminf_{n\to\infty} f_n + G d\mu \leq \liminf_{n\to\infty} \int_{\mathcal{X}} f_n + G d\mu$$

$$\int_{\mathcal{X}} f d\mu \leq \liminf_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$$

Similarly,

$$\int_{\mathcal{X}} \liminf_{n\to\infty} G - f_n d\mu \leq \liminf_{n\to\infty} \int_{\mathcal{X}} G - f_n d\mu$$

$$-\int_{\mathcal{X}} f_n d\mu \leq \liminf_{n\to\infty}(-\int_{\mathcal{X}} f_n d\mu) \leq \limsup_{n\to\infty}(-\int_{\mathcal{X}} f_n d\mu) = -\limsup_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$$

Therefore

$$\limsup_{n\to\infty} \int_{\mathcal{X}} f_n d\mu \leq \int_{\mathcal{X}} f d\mu \leq \liminf_{n\to\infty} \int_{\mathcal{X}} f_n d\mu$$

$$\lim_{n\to\infty} \int_{\mathcal{X}} f_n d\mu = \int_{\mathcal{X}} f d\mu$$

□

Lemma 1.5.2. *(Borel-Cantelli lemma) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $E_1, E_2, E_3, \ldots$ be a sequence of $\mathcal{B}$-measurable sets such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Show that almost every $x \in X$ is contained in at most finitely many of the E_n (i.e. $\{n \in \mathbb{N} : x \in E_n\}$ is finite for almost every $x \in X$). (Hint: Apply Tonelli's theorem to the indicator functions 1_{E_n}.)*

Proof. The statement that $\{n \in \mathbb{N} : x \in E_n\}$ is finite for almost every $x \in \mathcal{X}$ is equivalent to saying $\{x \in \mathcal{X} : \{n \in \mathbb{N} : x \in E_n\}$ is infinite$\}$ is a null set.

Before the major argument, we first check that the set above is $\mathcal{B}$-measurable. Notice that if $x \in \{x \in \mathcal{X} : \{n \in \mathbb{N} : x \in E_n\}$ is infinite$\}$, then x is in some countable union of E_n. Also, if x is in some countable union of E_n then $x \in \{x \in \mathcal{X} : \{n \in \mathbb{N} : x \in E_n\}$ is infinite$\}$ automatically. Therefore, this set is equal to some countable union of E_n and thus $\mathcal{B}$-measurable.

Therefore it is equivalent to saying $\mu(\{x \in \mathcal{X} : \{n \in \mathbb{N} : x \in E_n\}$ is infinite$\}) = 0$. Suppose, on the contrary, $\mu(\{x \in \mathcal{X} : \{n \in \mathbb{N} : x \in E_n\}$ is infinite$\}) = \epsilon > 0$.

First, the statement that $\{n \in \mathbb{N} : x \in E_n\}$ is infinite is equivalent to saying: $\forall m \in \mathbb{N}$, $\exists n_m$ such that $x \in E_{n_m}$. Therefore,

$$\begin{aligned}
\{x \in \mathcal{X} : \{n \in \mathbb{N} : x \in E_n\} \text{ is infinite}\} &= \{x \in \mathcal{X} : x \in E_{n_m} \ \forall m \in \mathbb{N}\} \\
&= \{x \in \mathcal{X} : x \in \bigcap_{m=1}^{\infty} E_{n_m}\} \\
&= \bigcap_{m=1}^{\infty} E_{n_m}
\end{aligned}$$

Thus,

$$\mu(\bigcap_{m=1}^{\infty} E_{n_m}) = \epsilon > 0$$

Thus,

$$\begin{aligned}
\sum_{n=1}^{\infty} \mu(E_n) &= \sum_{n=1}^{\infty} \int_{\mathcal{X}} \mathbf{1}_{E_n} d\mu \\
&= \int_{\mathcal{X}} \sum_{n=1}^{\infty} \mathbf{1}_{E_n} d\mu \text{ (Tonelli's theorem)} \\
&\geq \int_{\mathcal{X}} \sum_{m=1}^{\infty} \mathbf{1}_{E_{n_m}} d\mu \ (\{n_m \in \mathbb{N} : x \in E_{n_m}\} \subseteq \{n \in \mathbb{N} : x \in E_n\}) \\
&= \sum_{m=1}^{\infty} \int_{\mathcal{X}} \mathbf{1}_{E_{n_m}} d\mu \text{ (Tonelli's theorem again)} \\
&= \sum_{m=1}^{\infty} \mu(E_{n_m}) \\
&\geq \sum_{m=1}^{\infty} \mu(\bigcap_{m=1}^{\infty} E_{n_m}) \text{ (monotonicity)} \\
&= \sum_{m=1}^{\infty} \epsilon \\
&= \infty \text{ (since } \epsilon > 0)
\end{aligned}$$

This contradicts with the condition that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Thus, $\{n \in \mathbb{N} : x \in E_n\}$ is finite for almost every $x \in \mathcal{X}$. □

Lemma 1.5.3. *(Defect version of Fatou's lemma). Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \cdots : X \to [0, +\infty]$ be a sequence of unsigned absolutely integrable functions that converges pointwise to an absolutely integrable limit f. Show that*

$$\int_X f_n \, d\mu - \int_X f \, d\mu - \|f - f_n\|_{L^1(\mu)} \to 0$$

as $n \to \infty$. (Hint: Apply the dominated convergence theorem (Theorem 1.4.49) to $\min(f_n, f)$.) Informally, this result (first established in [BrLi1983]) tells us that the gap between the left and right hand sides of Fatou's lemma can be measured by the quantity $\|f - f_n\|_{L^1(\mu)}$.

Proof. Since $f_n \to f$ pointwise, we have $\min(f_n, f) \to f$ pointwise. Also $|\min\{f_n, f\}| \leq f$

and g that are absolutely integrable. Therefore, using Dominated Convergence Theorem,

$$\lim_{n\to\infty}\int_{\mathcal{X}}\min(f_n,f)d\mu \to \int_{\mathcal{X}} f d\mu$$

$\forall\epsilon>0$, $\exists N_1$, when $n\geq N_1$,

$$|\int_{\mathcal{X}}\min(f_n,f)d\mu - \int_{\mathcal{X}} f d\mu| \leq \epsilon$$

$$|(\int_{\mathcal{X}}\min(f_n,f)d\mu - \int_{\mathcal{X}} f_n d\mu) + (\int_{\mathcal{X}} f_n d\mu - \int_{\mathcal{X}} f d\mu)| \leq \epsilon$$

$$|\int_{\mathcal{X}}\min(0,f-f_n)d\mu + (\int_{\mathcal{X}} f_n d\mu - \int_{\mathcal{X}} f d\mu)| \leq \epsilon$$

Since $|\min(0,f-f_n)|\leq|f-f_n|\leq f+f_n$ that are absolutely integrable, we have

$$\lim_{n\to\infty}\int_{\mathcal{X}}\min(0,f-f_n)d\mu = 0$$

thus, $\exists N_2$ such that when $n\geq N_2$,

$$|\int_{\mathcal{X}}\min(0,f-f_n)d\mu| \leq \epsilon$$

Thus,

$$|\int_{\mathcal{X}} f_n d\mu - \int_{\mathcal{X}} f d\mu| \leq 2\epsilon$$

Since $f_n\to f$ pointwise, for each x, $\forall\epsilon>0$, $\exists N$, when $n\geq N$,

$$|f-f_n|\leq\epsilon$$

$$||f-f_n|-0|\leq\epsilon$$

Therefore $|f-f_n|\to 0$ pointwise. Also, $|f-f_n|$ is absolutely integrable and is dominated by $\sup_n f+f_n$, so we use Dominated Convergence Theorem again,

$$\lim_{n\to\infty}\int_{\mathcal{X}}|f-f_n|d\mu = \lim_{n\to\infty}||f-f_n||_{L^1(\mu)} = 0$$

Therefore, $\forall\epsilon>0$, $\exists N_3$, when $n\geq N_2$,

$$||f-f_n||_{L^1(\mu)}\leq\epsilon$$

Therefore, when $n\geq\max(N_1,N_2,N_3)$,

$$|\int_{\mathcal{X}} f_n d\mu - \int_{\mathcal{X}} f d\mu - ||f-f_n||_{L^1(\mu)}| \leq |\int_{\mathcal{X}} f_n d\mu - \int_{\mathcal{X}} f d\mu| + ||f-f_n||_{L^1(\mu)} \leq 3\epsilon$$

and the claim naturally follows. □

1.5.5 Modes of convergence

Next we see seven versions of convergence.

Definition 1.5.13. *(L^∞ norm) For measurable function f on $(\mathcal{X}, \mathcal{B}, \mu)$, its L^∞ norm is*

$$||f||_{L^\infty(\mathcal{X},\mathcal{B},\mu)} := \inf\{m \in [0, +\infty] : \mu(|f(x)| > M) = 0\}$$

Definition 1.5.14. *(Modes of convergence). For measurable functions f, $(f_n)_{n\in\mathbb{N}}$ on $(\mathcal{X}, \mathcal{B}, \mu)$,*
*(1). $f_n \to f$ **pointwise** if $\forall x \in \mathcal{X}$, $\epsilon > 0$, $\exists N = N(x, \epsilon) \in \mathbb{N}$ such that $|f(x) - f_n(x)| \leq \epsilon$ $\forall n \geq N$.*
*(2). $f_n \to f$ **uniformly** if $\forall \epsilon > 0$, $\exists N = N(\epsilon) \in \mathbb{N}$ such that $|f_n(x) - f(x)| \leq \epsilon$ $\forall n \geq N$ $\forall x \in \mathcal{X}$.*
*(3). $f_n \to f$ **pointwise a.e.** (or **μ-a.e.**) if $f_n(x) \to f(x)$ for μ-a.e. $x \in \mathcal{X}$.*
*(4). $f_n \to f$ **almost uniformly** if $\forall \epsilon > 0$, $\exists E \in \mathcal{B}$ with $\mu(E) \leq \epsilon$, $f_n \to f$ uniformly on $\mathcal{X} \setminus E$.*
*(5). $f_n \to f$ **uniformly a.e.** (or **in L^∞ norm**) if $\forall \epsilon > 0$, $\exists N(\epsilon) \in \mathbb{N}$ such that $|f_n(x) - f(x)| \leq \epsilon$ $\forall n \geq N$ for μ-a.e. $x \in \mathcal{X}$.*
*(6). $f_n \to f$ **in L^1 norm** if $||f_n - f||_{L^1(\mathcal{X},\mathcal{B},\mu)} = \int_{\mathcal{X}} |f_n - f| d\mu \to 0$ as $n \to \infty$.*
*(7). $f_n \to f$ **in measure** if $\forall \epsilon > 0$, $\mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) \to 0$ as $n \to \infty$.*

Proposition 1.5.10. *(Easy implications). Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_n : X \to \mathbb{C}$ and $f : X \to \mathbb{C}$ be measurable functions.*

(iv) If f_n converges to f almost uniformly, then f_n converges to f pointwise almost everywhere.

(vi) If f_n converges to f in L^1 norm, then f_n converges to f in measure.

(vii) If f_n converges to f almost uniformly, then f_n converges to f in measure.

Proof. (iv). Since $f_n \to f$ *almost uniformly*, $\forall n \in \mathbb{N}$, $\exists E_n$ with $\mu(E_n) \leq \frac{1}{n}$, such that $f_n(x) \to f(x)$ uniformly for all $x \in \mathcal{X} \setminus E_n$. Therefore, $f_n(x) \to f(x)$ for all $x \in \bigcup_{n\in\mathbb{N}} \mathcal{X} \setminus E_n$.

$$\begin{aligned}\bigcup_{n\in\mathbb{N}} \mathcal{X} \setminus E_n &= \bigcup_{n\in\mathbb{N}} \mathcal{X} \cap E_n^C \\ &= \mathcal{X} \cap (\bigcup_{n\in\mathbb{N}} E_n^C) \\ &= \mathcal{X} \cap (\bigcap_{n\in\mathbb{N}} E_n)^C \\ &= \mathcal{X} \setminus \bigcap_{n\in\mathbb{N}} E_n\end{aligned}$$

Also we have by monotinicity and non-negativity,

$$0 \leq \mu(\bigcap_{n\in\mathbb{N}} E_n) \leq \mu(E_n) \leq \frac{1}{n} \ \forall n \in \mathbb{N}$$

By taking limits and sandwich law,

$$\mu(\bigcap_{n\in\mathbb{N}} E_n) = 0$$

Therefore we have $f_n(x) \to f(x)$ pointwise on $\mathcal{X} \setminus \bigcap_{n\in\mathbb{N}} E_n$ where $\bigcap_{n\in\mathbb{N}} E_n$ is a (sub) null set. (a.e. pointwise).
(vi). Since $f_n \to f$ in L^1 *norm*, $|f_n - f|$ is measurable, and with Markov's Inequality,

$$\mu(\{x \in \mathcal{X} : |f_n(x) - f(x)| \geq \epsilon\}) \leq \frac{1}{\epsilon}\int_{\mathcal{X}} |f_n(x) - f(x)| d\mu \to 0 \text{ as } n \to \infty$$

Therefore we have convergence in measure.
(vii). Since $f_n \to f$ *almost uniformly*, $\forall \epsilon > 0$, $\exists E \in \mathcal{B}$ with $\mu(E) \leq \epsilon$, $\exists N$ such that $|f_n(x) - f(x)| \leq \epsilon$ when $n \geq N$ for all $x \in \mathcal{X} \setminus E$.
Therefore, when $n \geq N$,

$$\begin{aligned}\mu(\{x \in \mathcal{X} : |f_n - f| \geq \epsilon\}) &= \mu(\{x \in \mathcal{X} \setminus E : |f_n - f| \geq \epsilon\}) + \mu(\{x \in E : |f_n - f| \geq \epsilon\}) \\ &\leq 0 + \mu(E) \\ &\leq \epsilon \end{aligned} \tag{5}$$

Therefore we have $\mu(\{x \in \mathcal{X} : |f_n - f| \geq \epsilon\}) \to 0$ as $n \to \infty$. ◻

As an interesting example, the typewriter sequence $f_n = \mathbf{1}_{[\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}]}$ where $k \geq 0$ and $2^k \leq n < 2^{k+1}$. ($f_1 = \mathbf{1}_{[0,1]}$, $f_2 = \mathbf{1}_{[0,\frac{1}{2}]}$, $f_3 = \mathbf{1}_{[\frac{1}{2},1]}$, etc.) converges to zero in measure and L^1 norm, but not all others.

Proposition 1.5.11. *(Uniqueness). Let f, g, f_n ($n \geq 1$) be $\mathbb{C}$-valued measurable functions on $(\mathcal{X}, \mathcal{B}, \mu)$. If $f_n \to f$ in one of the 7 modes above and $f_n \to g$ in one of the 7 modes above, then $f = g$ a.e..*

Proof. Here we prove the case where $f_n \to f$ pointwise a.e. and $f_n \to g$ in measure.
Let $\epsilon > 0$. WTS $|f(x) - g(x)| \leq \epsilon$ for μ-a.e. $x \in \mathcal{X}$.
Let $A := \{x \in \mathcal{X} : |f(x) - g(x)| > \epsilon\}$ and it is measurable. Suppose, on the contrary, that $\mu(A) > 0$.
Let $A_N := \{x \in A : |f_n(x) - f(x)| \leq \frac{\epsilon}{2}\ \forall n \geq N\}$ and it is measurable and increasing in N.
Since $f_n \to f$ pointwise a.e., we have $A \setminus \bigcup_{N\in\mathbb{N}} A_N$ is null, because

$$\begin{aligned} A \setminus \bigcup_{N\in\mathbb{N}} A_N &= A \cap (\bigcup_{N\in\mathbb{N}} A_N)^C \\ &= \underbrace{A \cap (\bigcap_{N\in\mathbb{N}} A_N^C)}_{\text{null}} \end{aligned}$$

Thus $\mu(\bigcup_{N\in\mathbb{N}} A_N) > 0$.
By Monotone Convergence Theorem we have $\mu(A_N) \to \mu(\bigcup_{N\in\mathbb{N}} A_N) > 0$. Then $\exists N' \in \mathbb{N}$, $\mu(A_{N'}) > 0$. Then $\forall x \in A_{N'}$, $|f_n(x) - f(x)| \leq \frac{\epsilon}{2}\ \forall n \geq N'$. Also, since $A_{N'} \subseteq A$, $|f(x) - g(x)| > \epsilon$.

Then, by triangle inequality $|f_n(x) - g(x)| > \frac{\epsilon}{2}$ $\forall x \in A_{N'}$ $\forall n \geq N'$ with $\mu(A_{N'}) > 0$. This contradicts with $f_n \to g$ in measure. $\square$

1.6 Differentiation Theorems

1.6.1 Lebesgue Differentiation Theorem

We want to show that

Theorem 1.6.1. *Let $f : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable. Then for almost every $x \in \mathbb{R}^d$,*

$$\frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y) - f(x)|dy \to 0$$

and hense

$$\lim_{r\to 0} \frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)dy = f(x)$$

as $r \to 0$, where $B(x,r) := \{y \in \mathbb{R}^d : |x-y| < r\}$. We say x is a Lebesgue point of f.

This result is a convergence theorem, because there are

1. an assertion that for all functions f in a given class (in this case, the class of absolutely integrable functions $f : \mathbb{R}^d \to \mathbb{C}$).
2. a certain sequence of linear expressions $T_r f$ (in this case, $T_r f(x) = \frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)dy$) .

To prove such convergence theorem, we use *density argument*:

1. establish claim for "dense" class of "nice" functions. By "dense" we means that a general function f in the original class can be approximated to arbitrary accuracy in a suitable sense by a function in the nice subclasses.
2. establish a maximal inequality to control errors.

First we see an example of using *density argument*:

Proposition 1.6.1. *(Translation is continuous in L^1). For $f : \mathbb{R}^d \to \mathbb{C}$, $h \in \mathbb{R}^d$, set the shift function $f_h : \mathbb{R}^d \to \mathbb{C}$ to be $f_h(x) := f(x-h)$. If $f \in L^1(\mathbb{R}^d)$, then $||f_h - f||_{L^1(\mathbb{R}^d)} \to 0$ as $h \to 0$.*

Proof. First verify this claim for a dense subclass of f. Consider first the case $f \in C_c(\mathbb{R}^d)$. Then $f_h \to f$ uniformly. Then

$$\begin{aligned}
||f - f_h||_{L^1} &= \int_{\mathbb{R}^d} |f - f_h| \\
&\leq \int_{Supp(f)\cup Supp(f_h)} ||f - f_h||_{L^\infty} \\
&\leq ||f - f_h||_{L^\infty}(m(Supp(f)) + m(Supp(f_h))) \\
&= ||f - f_h||_{L^\infty} 2m(Supp(f_h)) \\
&\to 0
\end{aligned}$$

Now, for general $f \in L^1(\mathbb{R}^d)$, let $g \in C_c(\mathbb{R}^d)$ such that $||f - g||_{L^1} < \epsilon$ where $\epsilon > 0$ is fixed.

By triangle inequality,

$$\begin{aligned}
||f - f_h||_{L^1} &\leq ||g - g_h||_{L^1} + ||f - g||_{L^1} + ||f_h - g_h||_{L^1} \\
&= ||g - g_h||_{L^1} + ||f - g||_{L^1} + ||(f - g)_h||_{L^1} \\
&= ||g - g_h||_{L^1} + 2||f - g||_{L^1} \\
&< ||g - g_h||_{L^1} + 2\epsilon \\
&\to 2\epsilon
\end{aligned}$$

Then $||f - f_h||_{L^1} \to 0$ as $h \to 0$. $\square$

Proposition 1.6.2. *(Convolution). Let $f : \mathbb{R}^d \to \mathbb{C}, g : \mathbb{R}^d \to \mathbb{C}$ be Lebesgue measurable functions such that f is absolutely integrable and g is essentially bounded (i.e. bounded outside of a null set). Show that the convolution $f * g : \mathbb{R}^d \to \mathbb{C}$ defined by the formula*

$$f * g(x) = \int_{\mathbb{R}^d} f(y)g(x - y)\, dy$$

is well-defined (in the sense that the integrand on the right-hand side is absolutely integrable) and that $f * g$ is a bounded, continuous function.

Proof. First, we prove that it is well defined. Since g is bounded outside of a null set E, we have for arbitrary $x \in \mathbb{R}^d$, $|g(x - y)| \leq M < \infty$ when $y \in \mathbb{R}^d \setminus E$.
Therefore, $\forall x \in \mathbb{R}^d$

$$\begin{aligned}
\int_{\mathbb{R}^d} |f(y)g(x - y)|dy &= \int_{\mathbb{R}^d \setminus E} |f(y)g(x - y)|dy + \int_E |f(y)g(x - y)|dy \\
&= \int_{\mathbb{R}^d \setminus E} |f(y)||g(x - y)|dy \\
&\leq M \int_{\mathbb{R}^d \setminus E} |f(y)|dy \\
&\leq M \int_{\mathbb{R}^d} |f(y)|dy \\
&< \infty
\end{aligned}$$

Therefore $f * g$ is well-defined.
Next, we prove that it is bounded. By triangle inequality, $\forall x \in \mathbb{R}^d$,

$$\begin{aligned}
|f * g(x)| &= |\int_{\mathbb{R}^d} f(y)g(x - y)dy| \\
&\leq \int_{\mathbb{R}^d} |f(y)g(x - y)|dy \\
&\leq M \int_{\mathbb{R}^d} |f(y)|dy \\
&< \infty
\end{aligned}$$

Finally, we show that it is continuous. Let $\delta > 0$.

$$\begin{aligned} f * g(x+\delta) - f * g(x) &= \int_{\mathbb{R}^d} f(y)g(x+\delta-y)dy - \int_{\mathbb{R}^d} f(y)g(x-y)dy \\ &= \int_{\mathbb{R}^d} f(y+\delta)g(x-y)dy - \int_{\mathbb{R}^d} f(y)g(x-y)dy \\ &= \int_{\mathbb{R}^d} g(x-y)(f(y+\delta)-f(y))dy \end{aligned}$$

Therefore, by triangle inequality

$$\begin{aligned} |f * g(x+\delta) - f * g(x)| &\leq \int_{\mathbb{R}^d} |g(x-y)||f(x+\delta)-f(x)|dy \\ &\leq M \int_{\mathbb{R}^d} |f(x+\delta)-f(x)|dy \\ &\to 0 \text{ as } \delta \to 0 \text{ by translation invariance} \end{aligned}$$

Therefore, $\forall \epsilon > 0$, $\exists \delta > 0$ such that when $|y-x| \leq \delta$, $|f * g(y) - f * g(x)| \leq \epsilon$. Convolution is continuous. $\square$

Theorem 1.6.2. *(Steinhaus theorem). Let $E \subset \mathbb{R}^d$ be a Lebesgue measurable set of positive measure. Show that the set $E - E := \{x - y : x, y \in E\}$ contains an open neighbourhood of the origin. (Hint: reduce to the case when E is bounded, and then apply the previous exercise to the convolution $1_E * 1_{-E}$, where $-E := \{-y : y \in E\}$.)*

Proof. First, we look at the case where E is bounded, $m(E) < \infty$.
Notice that if $\mathbf{1}_E * \mathbf{1}_{-E}(x) = \int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy > 0$, then by continuity $\exists E' \subset \mathbb{R}^d$ with $m(E') > 0$ such that $\int_{E'} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy > 0$. Therefore, $\exists y'$ such that $y' \in E' \cap E$ and $y' - x \in E' \cap E$. Therefore, $x = y' - (y' - x) \in E - E$.
Since the convolution is continuous, $\forall \epsilon > 0$, $\exists \delta > 0$, such that when $|x| < \delta$,

$$|\int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy - \int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(-y)dy| < \epsilon$$

Since $-y \in -E$ iff $y \in E$,

$$|\int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy - \int_{\mathbb{R}^d} \mathbf{1}_E(y)dy| < \epsilon$$

$$|\int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy - m(E)| < \epsilon$$

$$m(E) - \epsilon < \int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy < m(E) + \epsilon$$

Since $m(E) > 0$, we can take $\epsilon = \frac{m(E)}{2}$, then $\exists \delta' > 0$, such that when $|x| < \delta'$,

$$\int_{\mathbb{R}^d} \mathbf{1}_E(y)\mathbf{1}_{-E}(x-y)dy > 0$$

By our previous argument, $\{x : |x| < \delta'\} \subset E - E$. Therefore we have found a neighborhood around the origin with radius $\delta' > 0$, such that it is contained in $E - E$.

Next, we look at the case where E is unbounded, that is, $m(E) = \infty$. Then, we can found a $N \in \mathbb{N}$ sucn that $E_N = B(0,N) \cap E \neq \emptyset$, is Lebesgue measurable, and that $m(E_N) > 0$. Then, we apply the previous argument and found that $E_N - E_N$ cntains an open neighbourhood of the origin. Since $E_N - E_N \subset E - E$, we have that $E - E$ contains an open neighborhood of the origin. $\square$

To prove that Theorem 1.6.1. holds, we use the density argument. The first step which is the dense subclass case is easy.

Proposition 1.6.3. *Theorem 1.6.1. holds whenever f is continuous.*

Proof. Since f is continuous, $\forall \epsilon > 0$, $\exists \delta > 0$ such that when $\forall x, y \in \mathbb{R}^d$,

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$$

Therefore,

$$\begin{aligned} \frac{1}{m(B(x,\delta))} \int_{m(B(x,\delta))} |f(y) - f(x)| dy &< \frac{1}{m(B(x,\delta))} \int_{m(B(x,\delta))} \epsilon dy \\ &= \frac{1}{m(B(x,\delta))} m(B(x,\delta))\epsilon \\ &= \epsilon \end{aligned} \tag{6}$$

Therefore,

$$\lim_{\delta \to 0} \frac{1}{m(B(x,\delta))} \int_{m(B(x,\delta))} |f(y) - f(x)| dy = 0$$

$\square$

The quantitative estimate and controling of error needs is the following.

Theorem 1.6.3. *(Hardy-Littlewood Maximal Inequality). For $f : \mathbb{R}^d \to \mathbb{C}$ in $L^1(\mathbb{R}^d)$, and the Hardy-Littlewood maximal function*

$$Mf(x) := \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f|, \ x \in \mathbb{R}^d$$

for any $\lambda > 0$, we have

$$m(\{x \in \mathbb{R}^d : Mf(x) \geq \lambda\}) \leq \frac{C_d}{\lambda} ||f||_{L^1(\mathbb{R}^d)}$$

for some constant $C_d > 0$ depending only on d.

To prove the Hardy-Littlewood Maximal Inequality, we need the following Lemma.

Lemma 1.6.1. *(Vitali's covering lemma). For any finite collection of open balls $\mathcal{B} = \{B_1, B_2, ..., B_n\}$, $B_i \subset \mathbb{R}^d$, there is a sub-collection $\mathcal{B}' = \{B_1', B_2', ..., B_m'\}$ of pairwise disjoint balls, such that*

$$\bigcup_{i=1}^{n} B_i \subset \bigcup_{i=1}^{m} 3B_i'$$

where $3B_i$ is the ball with the same center but 3 times diameter of B_i. By finite additivity,

$$m(\bigcup_{i=1}^{n} B_i) \leq 3^d \sum_{i=1}^{m} m(B_i')$$

Proof. Here we prove this lemma using a greedy algorithm.
Take B_1' to be the largest ball among $\mathcal{B}$. Then we do the following induction. Now having chosen $\{B_1', B_2', ..., B_k'\}$. If the remaining balls each has non-empty intersections with $\bigcup_{i=1}^{k} B_i'$, then stop. Otherwise, take the B_{k+1}' to be the largest among $\mathcal{B} \setminus \{B_1', B_2', ..., B_k'\}$ that is disjoint from $\bigcup_{j=1}^{k} B_j'$.
Therefore, we have

1. we must stop at $\leq n$ rounds.

2. ending collection $\mathcal{B}' = \{B_1', ..., B_m'\}$ must be pairwise disjoint.

Then it remains to show that each ball $B_i \in \mathcal{B}$ is covered by the triples $3B_j'$ of the subcollection. That is, fix arbitrary B_i where $1 \leq i \leq n$, it is enough to show that $B_i \subset \bigcup_{j=1}^{m} 3B_j'$. First, notice that $B_i \cap \bigcup_{j=1}^{m} B_i' \neq \emptyset$, since otherwise the algorithm won't stop at $\mathcal{B}' = \{B_1', ..., B_m'\}$.
Then, let $j_0 := \min\{j : B_j' \cap B_i \neq \emptyset\}$. Then, B_{j_0}' is the first ball in the subcollection $\mathcal{B}'$ that has non-empty intersection with B_i, and we also have $B_i \cap (\bigcup_{j=1}^{j_0-1} B_j') = \emptyset$.
Then, $m(B_i) \leq m(B_{j_0}')$ since otherwise B_i would have been chosen. Therefore $diam(B_i) \leq diam(B_{j_0}')$, so B_i cannot intersect with both B_{j_0}' and $(3B_{j_0}')^C$. Therefore, $\forall i \in \{1, ..., n\}$,

$$B_i \subset 3B_{j_0}' \subset \bigcup_{j=1}^{m} 3B_j'$$

Therefore

$$\bigcup_{i=1}^{n} B_i \subset \bigcup_{j=1}^{m} 3B_j'$$

□

Then, we will prove the Hardy-Littlewood Maximal Inequality (HLMI). We will show that

$$m(\{Mf > \lambda\}) \leq \frac{3^d}{\lambda} ||f||_{L^1(\mathbb{R}^d)}$$

(to get the case of $\geq$, apply the above with $\lambda - \epsilon$ in place of λ, and let $\epsilon \to 0$).

Proof. By Inner Regularity, it is enough to show that for any compact $K \subseteq \{Mf > \lambda\}$,

$$m(K) \leq \frac{3^d}{\lambda} ||f||_{L^1(\mathbb{R}^d)}$$

Fix arbitrarily such K. By construction, for any $x \in K$, since $Mf(x) > \lambda$, $\exists r_x > 0$, such that

$$\int_{B(x,r_x)} |f| > \lambda m(B(x, r_x))$$

Therefore $\{B(x, r_x)\}$ is a cover of K by open balls. By compactness of K, we can also cover K by a finite number of sub-covering $B_1, B_2, ..., B_n$. By Vitali's covering lemma, $\exists$ a subcollection $B'_1, B'_2, ..., B'_m$, such that B'_i are pairwise disjoint and

$$m(\bigcup_{i=1}^{n} B_i) \leq 3^d \sum_{i=1}^{m} m(B'_i)$$

Therefore

$$\begin{aligned} m(K) &\leq m(\bigcup_{i=1}^{n} B_i) \\ &\leq 3^d \sum_{i=1}^{m} m(B'_i) \\ &< \frac{3^d}{\lambda} \sum_{i=1}^{m} \int_{B'_i} |f| \\ &= \frac{3^d}{\lambda} \int_{\bigcup_{i=1}^{m} B'_i} |f| \text{ since } B'_i \text{ are pairwise disjoint} \\ &\leq \frac{3^d}{\lambda} \int_{\mathbb{R}^d} |f| \end{aligned}$$

Therefore,

$$m(\{Mf > \lambda\}) = \sup_{K \subseteq \{Mf > \lambda\}, K \text{ compact}} m(K) \leq \frac{3^d}{\lambda} ||f||_{L^1(\mathbb{R}^d)}$$

$\square$

Proposition 1.6.4. *(Dyadic maximal inequality). If $f : \mathbb{R}^d \to \mathbb{C}$ is an absolutely integrable function, establish the dyadic Hardy-Littlewood maximal inequality*

$$m(\{x \in \mathbb{R}^d : \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\, dy \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbb{R}} |f(t)|\, dt$$

where the supremum ranges over all dyadic cubes Q that contain x.

Hint: the nesting property of dyadic cubes will be useful when it comes to the covering lemma stage of the argument, much as it was in Exercise 1.1.14.

Proof. It suffice to prove the case of $>$. To get the case of $\geq$, we can pertubate λ slightly and then taking limits.

By inner regularity, it suffices to show that, for any compact $K \subseteq \{\sup_{x \in Q} \frac{1}{|Q|} \int_Q |f| > \lambda\}$,

$$m(K) \leq \frac{1}{\lambda} \int_{\mathbb{R}^d} |f|$$

For any $x \in K$, $\exists Q_x$ as a dyadic cube, such that

$$|Q_x| < \frac{1}{\lambda} \int_{Q'_x} |f| \ (*)$$

Let $\mathcal{Q}$ be the collection of all the dyadic cubes Q that we found by $(*)$. By the nature of dyadic cubes, $\mathcal{Q}$ is countable. Take $\mathcal{Q}^* \subseteq \mathcal{Q}$ to be the subcollection of maximal cubes with respect to set inclusion, which means that each of the cube in $\mathcal{Q}^*$ is not contained in any other cube in $\mathcal{Q}^*$. Together with the dyadic nesting property,

1. $\forall Q \in \mathcal{Q}$, $\exists! Q^* \in \mathcal{Q}^*$ such that $Q \subset Q^*$.
2. $\forall Q^*, Q'^* \in \mathcal{Q}^*$, such that $Q^* \neq Q'^*$, Q^* and Q'^* are almost disjoint.

By this we get $\mathcal{Q}^* = \mathcal{Q}$. Therefore, by monotonicity and almost disjoint, we have

$$\begin{aligned} m(K) &\leq m(\bigcup_{Q \in \mathcal{Q}} Q) \\ &= \sum_{Q \in \mathcal{Q}} |Q| \\ &< \frac{1}{\lambda} \sum_{Q \in \mathcal{Q}} \int_Q |f| \\ &= \frac{1}{\lambda} \int_{\bigcup_{Q \in \mathcal{Q}}} |f| \\ &= \frac{1}{\lambda} \int_{\mathbb{R}^d} |f| \end{aligned}$$

Taking the supremum of K, we have

$$m(\{x \in \mathbb{R}^d : \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)| \, dy > \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbb{R}} |f(t)| \, dt$$

□

Now we turn to prove the Lebesgue Differentiation Theorem, that is, $\forall f \in L^1(\mathbb{R}^d)$ and a.e. $x \in \mathbb{R}^d$,

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| dy \to 0 \text{ as } r \to 0$$

and thus

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} f(y) \to f(x) \text{ as } r \to 0$$

Proof. We write the local average with a shorthand

$$A_{x,r} h := \frac{1}{m(B(x,r))} \int_{B(x,r)} h$$

Let $\epsilon > 0$. It suffices to show that

$$E_\lambda = \{x : \limsup_{r \to 0} A_{x,r} |f - f(x)| > \lambda\}$$

has measure zero, since this guarantees that $E = \bigcup_{n \in \mathbb{N}} E_{\frac{1}{n}}$ has measure zero, and then the limit holds outside E.

We already know that this theorem holds whenever $f \in C_C(\mathbb{R}^d)$.

Let $\epsilon > 0$ and $g \in C_C(\mathbb{R}^d)$ such that $||f - g||_{L^1(\mathbb{R}^d)} < \epsilon$.

Set $h = f - g$, then for any $x \in \mathbb{R}^d$, $r > 0$,

$$\begin{aligned} A_{x,r}|f - f(x)| &\leq A_{x,r}|h - h(x)| + A_{x,r}|g - g(x)| \\ &\leq A_{x,r}|h| + |h(x)| + A_{x,r}|g - g(x)| \end{aligned}$$

Let $\lambda > 0$ be arbitrary, then we have

$$\begin{aligned} m(x : \limsup_{r\to 0} A_{x,r}|f - f(x)| > \lambda) \leq m(\{x : \limsup_{r\to 0} A_{x,r}|h| > \frac{\lambda}{3}\}) + m(\{|h| > \frac{\lambda}{3}\}) \\ + m(\{x : \limsup_{r\to 0} A_{x,r}|g - g(x)| > \frac{\lambda}{3}\}) \end{aligned}$$

By continuity of g,

$$m(\{x : \limsup_{r\to 0} A_{x,r}|g - g(x)| > \frac{\lambda}{3}\}) = 0$$

By Markov's Inequality,

$$m(\{|h| > \frac{\lambda}{3}\}) < \frac{3}{\lambda}||h||_{L^1(\mathbb{R}^d)} < \frac{3\epsilon}{\lambda}$$

By HLMI:

$$\begin{aligned} m(\{x : \limsup_{r\to 0} A_{x,r}|h| > \frac{\lambda}{3}\}) &\leq m(\{x : \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |h| > \frac{\lambda}{3}\}) \\ &\leq \frac{3C_d}{\lambda}||h||_{L^1(\mathbb{R}^d)} \\ &= \frac{3C_d\epsilon}{\lambda} \end{aligned}$$

Therefore

$$m(E_\lambda) \leq \frac{3\epsilon}{\lambda} + \frac{3C_d\epsilon}{\lambda}$$

Since $\epsilon > 0$ is arbitrary, we have $m(E_\lambda) = 0$ and the Lebesgue Integration Theorem holds.

□

Now we see some consequences of Lebesgue differentiability. In LDT, we assume that f is absolutely integrable. However, differentiability is a local property, which motivates the following definition.

Definition 1.6.1. *(Locally integrable). $f : \mathbb{R}^d \to \mathbb{C}$ is locally integrable if $\forall x \in \mathbb{R}^d$, $\exists U$ open such that $x \in U$,*

$$\int_U |f| dm < +\infty$$

LDT can be generalized to locally integrable functions, just apply LDT to the restrictions of f. We write $L^1_{loc}(\mathbb{R}^d)$ for locally integrable functions. For example, any polynomials is in $L^1_{loc}(\mathbb{R}^d) \setminus L^1(\mathbb{R}^d)$.

Definition 1.6.2. *(Point of density). For $E \in \mathcal{L}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ is a point of density for E if*

$$\frac{m(E \cap B(x,r))}{m(B(x,r))} \to 1 \text{ as } r \to 0$$

x need not be in E.

For example, for $E = [-1, 1] \setminus \{0\}$, any $x \in (-1, 1)$ is a point of density. $-1, 1$ are not because the ratio $\to \frac{1}{2}$. Any $x \notin [-1, 1]$ is not because the ration $\to 0$.

Next we have a corollary of LDT for $f \in L^1_{loc}(\mathbb{R}^d)$.

Corollary 1.6.1. *Let $E \in \mathcal{L}(\mathbb{R}^d)$, then almost every $x \in E$ is a point of density for E, and almost every $x \notin E$ is not a point of density for E.*

Proof. Apply LDT to $\mathbf{1}_E$. For almost every $x \in E$,

$$\begin{aligned}
\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} \mathbf{1}_E(y)dy &= \lim_{r \to 0} \frac{1}{m(B(x,r))} (\int_{B(x,r) \cap E} \mathbf{1}_E(y)dy + \int_{B(x,r) \setminus E} \mathbf{1}_E(y)dy) \\
&= \lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r) \cap E} \mathbf{1}_E(y)dy \\
&= \lim_{r \to 0} \frac{m(E \cap B(x,r))}{B(x,r)} \\
&= 1
\end{aligned}$$

For almost every $x \notin E$, we have this limit equals zero. ☐

1.6.2 Almost Everywhere Differentiability

We have functions that are continuous but nowhere differentiable, such as the Weierstrass Function

$$F(x) := \sum_{n=1}^{\infty} 4^{-n} cos(16^n \pi x)$$

The sum converges absolutely at each $x \in \mathbb{R}$, and defines a bounded continuous function. But this function is nowhere differentiable.

Now we define differentiability.

Definition 1.6.3. *(Discrete derivative). For a function $F : \mathbb{R} \to \mathbb{R}$ and $x \in \mathbb{R}$, denote the discrete derivative as*

$$\Delta_h F(x) := \frac{1}{h}(F(x+h) - F(x)), \ h \in \mathbb{R}$$

Definition 1.6.4. *(Dini derivative). The Dini derivatives of F are*

$$\overline{D^+}F(x) = \limsup_{h \to 0^+} \Delta_h F(x)$$

$$\underline{D^+}F(x) = \liminf_{h \to 0^+} \Delta_h F(x)$$

$$\overline{D^-}F(x) = \limsup_{h \to 0^-} \Delta_h F(x)$$

$$\underline{D^-}F(x) = \liminf_{h \to 0^-} \Delta_h F(x)$$

It is easy to notice that $\underline{D^+}F(x) \leq \overline{D^+}F(x)$ and $\underline{D^-}F(x) \leq \overline{D^-}F(x)$.

Definition 1.6.5. *A function F is differentiable at x precisely when the four derivatives are equal and finite:*

$$\overline{D^+}F(x) = \underline{D^+}F(x) = \overline{D^-}F(x) = \underline{D^-}F(x) \in (-\infty, +\infty)$$

To show any monotone non-decreasing F is differentiable a.e., it is enough to show that $\overline{D^+}F(x) < \infty$ and $\overline{D^+}F(x) \leq \underline{D^-}F(x)$ for a.e. x. (in non increasing case, replace F with $-F$).

Our goal is to prove the following theorem.

Theorem 1.6.4. *(Monotone Differentiation Theorem). If $F : \mathbb{R} \to \mathbb{R}$ is monotone, then it is differentiable a.e..*

Before proving this theorem, we notices some facts and first prove some lemmas and propositions.

1. Monotone F is measurable. For example, if F is monotone non decreasing, then for all a, $\{x \in \mathbb{R} : F(x) > a\}$ is an interval, and is measurable. Similar logic applies to non increasing case.

2. If F is measurable, then the dini derivatives of F are measurable. For example, if if F is monotone non decreasing, then

$$\begin{aligned}\{x : \overline{D^+}F(x) > \lambda\} &= \{x : \lim_{n\to\infty} \sup_{0<\frac{1}{k}<\frac{1}{n}, k\in\mathbb{N}} \frac{F(x+\frac{1}{k}) - F(x)}{\frac{1}{k}} > \lambda\} \\ &= \bigcap_{n\in\mathbb{N}} \bigcup_{k=n+1}^{\infty} \{x : \frac{F(x+\frac{1}{k}) - F(x)}{\frac{1}{k}} > \lambda\}\end{aligned} \tag{7}$$

   and each set above are measurable.

First, we prove the MDT for continuous functions. For continuous functions, we have the rising sun lemma.

Lemma 1.6.2. *(Rising Sun Lemma). Let $[a, b]$ be a compact interval, and $G : [a, b] \to \mathbb{R}$ a continuous function. Then there exists a countable family of disjoint open intervals $(a_n, b_n) \subset [a, b]$ such that*

1. *$\forall n \in \mathbb{N}$, $G(a_n) = G(b_n)$ or else $a_n = a$ and $G(b_n) \geq G(a_n)$.*

2. *If $x \in [a, b] \setminus \bigcup_{n=1}^{\infty}(a_n, b_n)$, then $G(y) \leq G(x)$ $\forall y \in [x, b]$.*

Proof. We use the fact that any open $U \subset \mathbb{R}$ can be written as $U = \bigcup_{n=1}^{\infty}(a_n, b_n)$ with (a_n, b_n) disjoint, non-empty, and $a_n \notin U, b_n \notin U$.
Let $U := \{x \in (a, b) : \exists y \in (x, b) \ s.t. \ G(y) > G(x)\}$. As G is continuous, U is open, and so $U = \bigcup_{n=1}^{\infty}(a_n, b_n)$ as above.
Suppose that (a_n, b_n) is such that $a_n \neq a$. Since $a_n \notin U$, we have $G(y) \leq G(a_n)$ $\forall y \in [a_n, b]$. Similarly we have $G(y) \leq G(b_n)$ $\forall y \in [b_n, b]$. Since $a_n < b_n \leq b$, we have $G(b_n) \leq G(a_n)$. By the continuity of G, it suffices to show that $G(b_n) \geq G(t)$ for all $a_n < t < b_n$.
Suppose, for contradiction, there exists $a_n < t < b_n$ with $G(b_n) < G(t)$. Let $A := \{s \in [t, b] : G(s) \geq G(t)\}$, then A is closed and contains t but is disjoint from $[b_n, b]$ (Since if $s \in [b_n, b]$, then $G(s) \geq G(t) > G(b_n) \geq G(y)$ $\forall y \in [b_n, b]$, contradiction). Set $t_* := \sup(A)$, then $t_* \in [t, b_n) \subset (a_n, b_n) \subset U$. Then, $\exists y$ such that $t_* < y < b$, $G(y) > G(t_*)$. Since $G(t_*) \geq G(t) > G(b_n)$ and $G(b_n) \geq G(z)$ for all $b_n \leq z \leq b$, we have $y \in A$ ($t \in A$,

$t \leq t_* < y < b$, $G(y) > G(t)$) and thus t_* cannot be the supremum of A. Contradiction! Therefore, $G(b_n) \geq G(t)$ for all $a_n < t < b_n$.
Suppose that $a_n = a$, then the case is similar.
If $x \neq U$, then $\forall y \in [x, b]$, $G(y) \leq G(x)$. □

For both continuous and discontinuous function F, we have the One-sided Hardy-Littlewood Maximal Inequality (HLMI).

Lemma 1.6.3. *(One-sided Hardy-Littlewood Maximal Inequality). Let $F : [a, b] \to \mathbb{R}$ be a continuous monotone non-decreasing function, and let $\lambda > 0$. Then we have*

$$m(\{x \in [a, b] : \overline{D^+}F(x) \geq \lambda\}) \leq \frac{F(b) - F(a)}{\lambda}$$

and similarly, for the other three dini derivatives of F.

If F is not assumbed to be continuous, then we have the weaker inequality

$$m(\{x \in [a, b] : \overline{D^+}F(x) \geq \lambda\}) \leq C\frac{F(b) - F(a)}{\lambda}$$

for some absolute constant $C > 0$.

Proof. For F continuous, by modifying λ by an epsilon and dropping the endpoints from $[a, b]$ as they have measure zero, it is enough to show that

$$m(\{x \in (a, b) : \overline{D^+}F(x) > \lambda\}) \leq \frac{F(b) - F(a)}{\lambda}$$

We apply the rising sun lemma to the continuous function $G(x) := F(x) - \lambda x$ to get $I_n = (a_n, b_n) \subset (a, b)$ such that $G(a_n) \leq G(b_n)$ $\forall n$ and $G(x) \geq G(y)$ $\forall x \in [a, b] \setminus \bigcup_n I_n,\ y \in [x, b]$.
For $x \in (a, b) \setminus \bigcup_n I_n$, we have for all $h \in [0, b - x]$,

$$F(x) - \lambda x \geq F(x + h) - \lambda(x + h)$$

then

$$\overline{D^+}F(x) \leq \lambda$$

Therefore

$$\{x \in (a, b) : \overline{D^+}F(x) > \lambda\} \subset \bigcup_n I_n$$

$$\begin{aligned} m(\{x \in (a, b) : \overline{D^+}F(x) > \lambda\}) &\leq \sum_n b_n - a_n \\ &\leq \sum_n \frac{F(b_n) - F(a_n)}{\lambda} \text{ by rising sun} \\ &\leq \frac{F(b) - F(a)}{\lambda} \text{ by monotone and disjoint} \end{aligned}$$

For F non-continuous, It suffices to show the case

$$m(\{x \in [a, b] : D^+F(x) > \lambda\}) < C\frac{F(b) - F(a)}{\lambda}.$$

(to get the case of $\geq$, apply the above with $\lambda - \epsilon$ in place of λ, and then take $\epsilon \to 0$). From inner regularity, it suffices to show that for any compact $K \subseteq \{x \in [a,b] : \overline{D^+}F(x) > \lambda\}$,

$$m(K) \leq C\frac{F(b) - F(a)}{\lambda}$$

First,

$$\begin{aligned}
\overline{D^+}F(x) &= \limsup_{h \to 0^+} \frac{F(x+h) - F(x)}{h} \\
&= \lim_{\delta \to 0^+} \sup_{0<h<\delta} \frac{F(x+h) - F(x)}{h}
\end{aligned}$$

Fix arbitrary K satisfying the condition above, $\forall x \in K$, since $\overline{D^+}F(x) > \lambda$, we have $\exists \delta_x$ such that

$$\sup_{0<h<\delta_x} \frac{F(x+h) - F(x)}{h} > \lambda$$

Then, $\exists h_x$ such that $0 < h_x < \delta_x$ and

$$\frac{F(x+h_x) - F(x)}{h_x} > \lambda$$

$$h_x < \frac{F(x+h_x) - F(x)}{\lambda}$$

Let $\epsilon > 0$. We see that $\{(x - \epsilon, x + h_x + \epsilon)\}_{x \in K}$ is an open covering of compact K. Then, there exists a finite sub-covering $\{I_i\}_{i=1}^n = \{(x_i - \epsilon, x_i + h_{x_i} + \epsilon)\}_{i=1}^n$. By Vitali covering lemma, there exists a finite disjoint sub-sub-covering $\{I_i'\}_{i=1}^m = \{(x_i' - \epsilon, x_i' + h_{x_i'} + \epsilon)\}_{i=1}^m$ such that $\bigcup_{i=1}^n I_i \subset \bigcup_{i=1}^m 3I_i'$ and $m(\bigcup_{i=1}^n I_i) \leq 3\sum_{i=1}^m (h_{x_i'} + 2\epsilon)$.
Also, since $\{I_i'\}_{i=1}^m = \{(x_i' - \epsilon, x_i' + h_{x_i'} + \epsilon)\}_{i=1}^m$ are disjoint, we have $\{(x_i', x_i' + h_{x_i'})\}_{i=1}^m$ are disjoint.
Therefore, from monotonicity and disjointness,

$$\begin{aligned}
m(K) &\leq m(\bigcup_{i=1}^n I_i) \\
&\leq 3\sum_{i=1}^m (h_{x_i'} + 2\epsilon) \\
&< 3\frac{1}{\lambda}\sum_{i=1}^m (F(x_i' + h_{x_i'}) - F(x_i')) + 6m\epsilon \\
&\leq 3\frac{1}{\lambda}(F(b) - F(a)) + 6m\epsilon \text{ by monotonicity of } F \text{ and disjointness of intervals}
\end{aligned}$$

Since $\epsilon > 0$ arbitrary, we have

$$m(K) < \frac{3}{\lambda}(F(b) - F(a))$$

Then, by inner regularity, we have

$$m(\{x \in [a,b] : D^+F(x) > \lambda\}) < C\frac{F(b) - F(a)}{\lambda}.$$

and $C = 3$. □

Sending $\lambda \to \infty$ leads to the conclusion that the Dini derivatives of a continuous monotone non-decreasing function are finite almost everywhere. Therefore, to prove MDT, it suffices to show that

$$\overline{D^+}F(x) = \underline{D^+}F(x) = \overline{D^-}F(x) = \underline{D^-}F(x) \in (-\infty, +\infty)$$

holds for almost every x. Then it suffices to show that $\overline{D^+}F(x) \leq \underline{D^-}F(x)$ for almost every x. It suffices to show that for every pair $0 < r < R$ of real numbers, the set

$$E = E_{r,R} := \{x \in \mathbb{R} : \overline{D^+}F(x) > R > r > \underline{D^-}F(x)\}$$

is a null set, since by letting R, r range over rationals with $R > r > 0$ and taking countable unions, we would conclude that the set $\{x \in \mathbb{R} : \overline{D^+}F(x) > \underline{D^-}F(x)\}$ is a null set, and the claim follows.

To prove that it is null, we will establish the following estimate first.

Lemma 1.6.4. *(E has density less that one). For any interval $[a,b]$ and any $0 < r < R$,*

$$m(E_{r,R} \cap [a,b]) \leq \frac{r}{R}|b-a|$$

Proof. Applying the rising sun lemma to the following continuous function on $[-b, -a]$:

$$G(x) := rx + F(-x)$$

Then there exists a countable family of disjoint non-empty intervals $-I_n = (-b_n, -a_n)$ such that

1. $G(-a_n) \geq G(-b_n)\ \forall n$.
2. If $-x \in (-b, -a) \setminus \bigcup_n -I_n$, then $G(-x) \leq G(-y)\ \forall -x \leq -y \leq -a$.

Following 2, if $x \in (a,b) \setminus \bigcup_n I_n$ and $G(-x) \leq G(-y)\ \forall -x \leq -y \leq -a$, then

$$-rx + F(x) \leq -ry + F(y),\ a \leq y \leq x$$

$$\frac{F(y) - F(x)}{y - x} \geq r,\ a \leq y \leq x$$

Let $y \to x^-$ and take limit infimum, we have $\underline{D^-}F(x) \geq r$.
Therefore, $E_{r,R} = \{x \in \mathbb{R} : \overline{D^+}F(x) > R > r > \underline{D^-}F(x)\} \subset \bigcup_n I_n$. Thus it has finite measure.

Also, using the one-sided HDMI:

$$\begin{aligned} m(\{E_{r,R} \cap [a_n, b_n]\}) &= m(\{x \in [a_n, b_n] : \overline{D^+}F(x) > R > r > \underline{D^-}F(x)\}) \\ &\leq m(\{x \in [a_n, b_n] : \overline{D^+}F(x) > R\}) \\ &\leq \frac{F(b_n) - F(a_n)}{R} \end{aligned}$$

From 1, we have $F(b_n) - F(a_n) \leq r(b_n - a_n)$, therefore

$$m(\{E_{r,R} \cap [a_n, b_n]\}) \leq \frac{r}{R}(b_n - a_n)$$

By additivity and disjointness, we have

$$\begin{aligned} m(E_{r,R} \cap [a, b]) &= m(E_{r,R} \cap (a, b)) \\ &= \sum_n m(\{E_{r,R} \cap [a_n, b_n]\}) \\ &\leq \sum_n \frac{r}{R}(b_n - a_n) \\ &\leq \frac{r}{R}|b - a| \end{aligned}$$

□

This lemma implies that E has no points of density, since $\forall x, r'$,

$$\frac{m(E_{r,R} \cap B(x, r'))}{m(B(x, r'))} \leq \frac{r}{R} < 1$$

Together with Corollary 1.6.1. which states that almost every point $x \in E_{r,R}$ is a point of density of $E_{r,R}$ and almost every point in $E_{r,R}^C$ is not a point of density of $E_{r,R}$, they imply that $E_{r,R}$ is a null set. Thus,

$$\{x \in \mathbb{R} : \overline{D^+}F(x) > \underline{D^-}F(x)\} = \bigcap_{n=1}^{\infty} \bigcap_{m \geq n}^{\infty} E_{\frac{1}{m}, \frac{1}{n}}$$

is also a null set. Thus we proved the MDT for continuous functions.

Now we turn to prove the MDT for general F that is monotone non-decreasing, dropping the continuity assumption. We supplement the continuous monotone functions with another class of monotone functions, known as the jump functions.

Definition 1.6.6. *(Jump Function). A basic jump function is of the form*

$$J(x) = \begin{cases} 0 & x < x_0 \\ \theta & x = x_0 \\ 1 & x > x_0 \end{cases}$$

for $x_0 \in \mathbb{R}$, $\theta \in [0, 1]$.

A jump function is a function of the form

$$F = \sum_{n=1}^{\infty} c_n J_n$$

for $J_1, J_2, ...$ as basic jump functions and $c_1, c_2, ... > 0$ with $\sum_{i=1}^{\infty} c_n < \infty$.
If $c_n = 0$ for all but finitely many n, we call F a piecewise constant jump function.

We see the following facts:

1. all jump functions are monotone non-decreasing.
2. any jump function $\sum_{n=1}^{\infty} c_n J_n$ is the uniform limit of piecewise constant jump functions $\sum_{n=1}^{N} c_n J_n$.
3. from 2, the points of discontinuities of $\sum_{n=1}^{\infty} c_n J_n$ are the points x_n where each J_n jumps.

These jump functions, together with the continuous monotone functions, essentially generate all monotone functions, at least in the bounded case.

Lemma 1.6.5. *(Continuous-Singular Decomposition for Monotone Functions). Let $F : \mathbb{R} \to \mathbb{R}$ be a monotone non-decreasing function. Then*

1. *The only discontinuities are jump discontinuities, i.e. if F is not continuous at x, then*
$$\lim_{y \to x^-} F(y) < \lim_{y \to x^+} F(y)$$
and both limits exist.
2. *F has at most countably many points of discontinuity.*
3. *If F is bounded, then it can be expressed as*
$$F = F_c + F_{pp}$$
where F_c is monotone non-decreasing and continuous, F_{pp} is a jump function.

Proof. 1. By monotonicity, the limit $F_-(x) = \lim_{y \to x^-} F(y)$ and $F^+(x) = \lim_{y \to x^+} F(y)$ always exists, with $F_-(x) \le F(x) \le F_+(x)$ for all x.
2. By 1, whenever there is a discontinuity f of F, there is at least one rational number q_x such that $F_-(x) < q_x < F_+(x)$, and from monotonicity, each rational number can be assigned to at most one discontinuity.
3. Let $A := \{x : F \text{ is discontinuous at } x\}$, and by 2, A is at most countable. For each $x \in A$, define

$$c_x := F_+(x) - F_-(x) > 0$$
$$\theta_x := \frac{F(x) - F_-(x)}{F_+(x) - F_-(x)} \in [0, 1]$$

From these we see that

$$F_+(x) = F_-(x) + c_x$$

$$F(x) = F_-(x) + c_x\theta_x$$

Each c_x is the measure of interval $(F_-(x), F_+(x))$. By monotonicity of F, $\{(F_-(x), F_+(x))\}_{x\in A}$ are disjoint. By boundedness of F, $\bigcup_{x\in A}(F_-(x), F_+(x))$ are bounded. By σ-additivity, we have $\sum_{x\in A} c_x < \infty$.
Let J_x be the basic jump function with the jump at x and fraction θ_x. Then define a jump function

$$F_{pp} := \sum_{x\in A} c_x J_x$$

and we see that F_{pp} is discontinuous only at A.
For all $x \in A$,

$$(F_{pp})_+(x) = \lim_{y\to x^+} F_{pp}(x) = \sum_{x'\in A\cap(-\infty,x)} c_{x'} + c_x \lim_{y\to x^+} J_x = \sum_{x'\in A\cap(-\infty,x)} c_{x'} + c_x$$

$$(F_{pp})_-(x) = \lim_{y\to x_-} F_{pp}(x) = \sum_{x'\in A\cap(-\infty,x)} c_{x'} + c_x \lim_{y\to x^-} J_x = \sum_{x'\in A\cap(-\infty,x)} c_{x'}$$

$$F_{pp}(x) = \sum_{x'\in A\cap(-\infty,x)} c_{x'} + c_x\theta_x$$

Therefore we see that

$$(F_{pp})_+(x) = (F_{pp})_-(x) + c_x$$

$$F_{pp}(x) = (F_{pp})_-(x) + c_x\theta_x$$

Together with the previous equalities we have

$$F_+(x) - (F_{pp})_+(x) = F(x) - F_{pp}(x) = F_-(x) - (F_{pp})_-(x)$$

Therefore $F_c := F - F_{pp}$ is continuous.
Then, we also need to verify that F_c is monotone non-decreasing, that is, for all $a < b$,

$$F_{pp}(b) - F_{pp}(a) \le F(b) - F(a)$$

Notice that the LHS equals $\sum_{x\in A\cap[a,b]} c_x$ where $c_x = F_+(x) - F_-(x)$, and $(F_-(x), F_+(x))$ are disjoint intervals for $x \in A \cap [a, b]$ (by monotonicity) and lie in $(F(a), F(b))$. Therefore by countable additivity, LHS $\le$ RHS.
Therefore, 3 holds. $\square$

Now we are ready to prove the MDT for general monotone non-decreasing function F.

Proof. Since differentiability is a local property, we assume that F is bounded. Apply 3 in the Continuous-Singular Decomposition, we write $F = F_c + F_{pp}$. For continuous cases, the MDT has already been proven. Therefore, it suffices to show that jump functions are differentiable a.e..
First we use the piecewise constant jump function as the dense subclass. If F is a piecewise constant jump function, then $F' = 0$ everywhere except at finite many jump points, so F is differentiable a.e. in this case.
Next we run the density argument. Let $\epsilon > 0$ and $\lambda > 0$ be arbitrary. Then, by uniform

convergence, $\exists F_\epsilon$ piecewise constant jump function such that

$$\sup_x |F(x) - F_\epsilon(x)| < \epsilon$$

By taking F_ϵ to be a partial sum of the basic jump functions that make up F, we can ensure that $F - F_\epsilon$ is monotone non-decreasing. By one-sided HLMI for general monotone non-decreasing functions,

$$\begin{aligned} m(\{x \in \mathbb{R} : \overline{D^+}(F - F_\epsilon)(x) \geq \lambda\}) &\leq \frac{C}{\lambda}(F(\infty) - F_\epsilon(\infty) - F(-\infty) + F_\epsilon(-\infty)) \\ &\leq \frac{C}{\lambda} 2 \sup_x |F(x) - F_\epsilon(x)| \\ &\leq \frac{2C\epsilon}{\lambda} \end{aligned}$$

and silimarly for other three Dini derivatives.
Since F_ϵ has Dini derivative 0 almost everywhere, all four Dini derivatives of F are bounded in absolute value by λ outside a set of measure at most $\frac{8C\epsilon}{\lambda}$. Hence, by triangle inequality, all are withing 2λ distance of each other. Taking $\epsilon \to 0$, all 4 Dini derivatives are withing 2λ of each other a.e.. Taking $\lambda \to 0$, we have

$$\overline{D^+}F(x) = \underline{D^+}F(x) = \overline{D^-}F(x) = \underline{D^-}F(x) \in (-\infty, +\infty)$$

holds for almost every x. and hence F is differentiable a.e..

$\square$

We now use the differentiation theory of monotone functions to develop the differentiation theory for the class of functions of bounded variation.

Definition 1.6.7. *(Total variation). The total variation of $F : \mathbb{R} \to \mathbb{R}$ is*

$$||F||_{TV} := \sup_{x_0 < \ldots < x_n} \sum_{i=0}^{n-1} |F(x_i) - F(x_{i+1})| \in [0, +\infty]$$

where the supremum ranges over all finite increasing sequences $x_0, ..., x_n \in \mathbb{R}$.
F has bounded variation on $\mathbb{R}$ if $||F||_{TV} < \infty$. Given any interval $[a, b]$, for $F : [a, b] \to \mathbb{R}$,

$$||F||_{TV([a,b])} := \sup_{a \leq x_0 < \ldots < x_n \leq b} \sum_{i=0}^{n-1} |F(x_i) - F(x_{i+1})|$$

For examples,

1. For F monotone, F has bounded variation iff F is bounded (above and below). Also, for any $a < b$, $||F||_{TV([a,b])} = |F(b) - F(a)|$.

2. $F(x) = e^{-x^2}$, then $||F||_{TV([a,b])} = 2$. This is realized by taking $x_0 = -N, x_1 = 0, x_2 = N$, and let $N \to \infty$.

3. $||\sin||_{TV} = +\infty$, $||\sin||_{TV([0,N])} = \frac{2}{\pi}N + O(1)$.

We also have the following proposition.

Proposition 1.6.5. *(Triangle inequality, homogeneity, and constant). For any functions $F, G : \mathbb{R} \to \mathbb{R}$, establish the triangle property $\|F+G\|_{TV(\mathbb{R})} \leq \|F\|_{TV(\mathbb{R})} + \|G\|_{TV(\mathbb{R})}$ and the homogeneity property $\|cF\|_{TV(\mathbb{R})} = |c|\|F\|_{TV(\mathbb{R})}$ for any $c \in \mathbb{R}$. Also show that $\|F\|_{TV(\mathbb{R})} = 0$ if and only if F is constant.*

Proof. By definition,

$$||F+G||_{TV(\mathbb{R})} = \sup_{x_0<...<x_n} \sum_{i=1}^{n-1} |F(x_i) + G(x_i) - F(x_{i+1}) - G(x_{i+1})|$$

Let $\epsilon > 0$, then $\exists x_0' < ... < x_n'$, such that

$$\begin{aligned}
||F+G||_{TV(\mathbb{R})} &\leq \sum_{i=1}^{n-1} |F(x_i') + G(x_i') - F(x_{i+1}') - G(x_{i+1}')| + \epsilon \\
&\leq \sum_{i=1}^{n-1} |F(x_i') - F(x_{i+1}')| + \sum_{i=1}^{n-1} |G(x_i') - G(x_{i+1}')| + \epsilon \\
&\leq ||F||_{TV(\mathbb{R})} + ||G||_{TV(\mathbb{R})} + \epsilon
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have $\|F+G\|_{TV(\mathbb{R})} \leq \|F\|_{TV(\mathbb{R})} + \|G\|_{TV(\mathbb{R})}$.
For the case where either F or G is not of bounded variation, the right hand side equals $+\infty$, and the inequality holds. Next we show homogeneity. By definition,

$$\begin{aligned}
||cF||_{TV(\mathbb{R})} &= \sup_{x_0<...<x_n} \sum_{i=1}^{n-1} |cF(x_i) - cF(x_{i+1})| \\
&= \sup_{x_0<...<x_n} |c| \sum_{i=1}^{n-1} |F(x_i) - F(x_{i+1})| \\
&= |c| \sup_{x_0<...<x_n} \sum_{i=1}^{n-1} |F(x_i) - F(x_{i+1})| \\
&= |c|||F||_{TV(\mathbb{R})}
\end{aligned}$$

Finally we show that $||cF||_{TV(\mathbb{R})} = 0$ iff F is constant.
($\Leftarrow$): If $F(x) = c \; \forall x \in \mathbb{R}$, then, $\forall x_0 < ... < x_n$, $|F(x_i) - F(x_{i+1})| = 0$ and thus $||F||_{TV(\mathbb{R})} = 0$.
($\Rightarrow$): If F is not a constant, that is, suppose for $x_1 \neq x_2$, $F(x_1) \neq F(x_2)$. Then, by definition and supremum, $||F||_{TV(\mathbb{R})} \geq |F(x_1) - F(x_2)| > 0$. Therefore, if $||F||_{TV(\mathbb{R})} = 0$, then F is constant. □

A bounded variation function can be expressed as the difference of two bounded monotone functions.

Proposition 1.6.6. $F : \mathbb{R} \to \mathbb{R}$ *is of bounded variation if and only if it is the difference of two bounded monotone functions.*

Proof. Define the positive variation $F^+ : \mathbb{R} \to \mathbb{R}$:

$$F^+(x) := \sup_{x_0<...<x_n \leq x} \sum_{i=1}^{n} \max(0, F(x_i) - F(x_{i-1}))$$

Note that F^+ is non decreasing, and $F^+(x) \in [0, ||F||_{TV}]$, so F^+ is bounded monotone non decreasing. Then, it suffices to show that $F^+ - F$ is monotone non-decreasing. (It is bounded by triangle inequality and that F is bounded).
Let $a < b$ be arbitrary, it suffices to show that

$$F^+(b) \geq F^+(a) + F(b) - F(a)$$

If $F(b) \leq F(a)$, then the right hand side $\leq F^+(a)$ and then by non decreasing of F^+, $\leq F^+(b)$.
If $F(b) > F(a)$, then for any finite sequence $x_0 < ... < x_n \leq a$, we can add in points and a, b to get a new sequence $x'_0 < ... < x'_n \leq b$, with

$$\begin{aligned} F(b) - F(a) + \sum_{i=1}^{n} \max(0, F(x_i) - F(x_{i-1})) &\leq \sum_{i=1}^{n'} \max(0, F(x'_i) - F(x'_{i-1})) \\ &\leq F^+(b) \end{aligned}$$

Take suprema of $x_0 < ... < x_n \leq a$, we get $F^+(b) \geq F^+(a) + F(b) - F(a)$. ☐

Then, we immediately have the following theorem.

Theorem 1.6.5. *(BV Differentiation Theorem). $F : \mathbb{R} \to \mathbb{R}$ is of bounded variation, then F is differentiable a.e..*

Proof. First write F as the difference of two bounded monotone functions. Then apply the MDT and we get differentiability a.e.. ☐

Theorem 1.6.6. *(Lipschitz differentiation theorem, one-dimensional case). A function $f : \mathbb{R} \to \mathbb{R}$ is said to be Lipschitz continuous if there exists a constant $C > 0$ such that $|f(x) - f(y)| \leq C|x - y|$ for all $x, y \in \mathbb{R}$; the smallest C with this property is known as the Lipschitz constant of f. Show that every Lipschitz continuous function F is locally of bounded variation, and hence differentiable almost everywhere. Furthermore, show that the derivative F', when it exists, is bounded in magnitude by the Lipschitz constant of F.*

Proof. Since we are verifying f locally, let $[a, b]$ be a compact local interval with $a < b$. Then, by Lipschitz continuity,

$$\begin{aligned} ||f||_{TV([a,b])} &= \sup_{a \leq x_0 < x_1 < ... < x_n \leq b} \sum_{i=1}^{n-1} |F(x_i) - F(x_{i+1})| \\ &\leq \sup_{a \leq x_0 < x_1 < ... < x_n \leq b} C \sum_{i=1}^{n-1} |x_i - x_{i+1}| \\ &= C \sup_{a \leq x_0 < x_1 < ... < x_n \leq b} \sum_{i=1}^{n-1} |x_i - x_{i+1}| \end{aligned}$$

Let $\epsilon > 0$ be given, then $\exists a \leq x_0' < x_1' < ... < x_n' \leq b$ such that

$$\sup_{a \leq x_0 < x_1 < ... < x_n \leq b} \sum_{i=1}^{n-1} |x_i - x_{i+1}| \leq \sum_{i=1}^{n-1} |x_i' - x_{i+1}'| + \epsilon \leq |b-a| + \epsilon$$

Therefore

$$||f||_{TV([a,b])} \leq C|b-a| + C\epsilon$$

Since $\epsilon > 0$ is arbitrarily small, we have

$$||f||_{TV([a,b])} \leq C|b-a|$$

Then f is locally of bounded variation. We can write $\mathbb{R} = \bigcup_{n=1}^{\infty} [-n,n]$ and see that f is differentiable almost everywhere in each of such intervals by BV differentiation theorem. Therefore, by σ-additivity, it is differentiable almost everywhere.
For x such that F is differentiable, we have

$$\overline{D^+}F(x) = \underline{D^+}F(x) = \overline{D^-}F(x) = \underline{D^-}F(x) = F'(x)$$

and

$$F'(x) = \lim_{y \to x} \frac{F(y) - F(x)}{y - x}$$

Since F is Lipschitz continuous,

$$|F'(x)| = \lim_{y \to x} |\frac{F(y) - F(x)}{y - x}| \leq \lim_{y \to x} C = C$$

Therefore F' is bounded above by C. □

1.6.3 The second fundamental theorem of calculus

We want to know when we have $\int_{[a,b]} F'(x)dx = F(b) - F(a)$. It is clear that a.e. differentiability is not sufficient. For example,

$$H(x) := \mathbf{1}_{[0,+\infty)}(x)$$

is differentiable a.e. with $H'(x) = 0$ a.e.. However, $H(1) - H(-1) = 1 \neq \int_{[-1,1]} H'(x)dx = 0$. Even if F is continuous and differentiable a.e., it is not sufficient. One counter example is Cantor's function.

We begin with $F : [a,b] \to \mathbb{R}$ monotone non-decreasing, which implies that F is differentiable a.e. in $[a,b]$, so F' is defined a.e.. Also F' is non-negative, and F' is measurable.

Proposition 1.6.7. *(Upper bound for second fundamental theorem). Let $F : \mathbb{R} \to \mathbb{R}$ be monotone non-decreasing. Then,*

$$\int_{[a,b]} F'(x)dx \leq F(b) - F(a)$$

In particular, F' is absolutely integrable.

Proof. First, we extend F to be $\mathbb{R} \to \mathbb{R}$, such that $F(x) = F(a)$ if $x < a$, $F(x) = F(b)$ if

$x > b$. Then, F is bounded an non-decreasing in $\mathbb{R}$. $F'(x) = 0$ outside $[a, b]$.
Set

$$f_n(x) := \frac{F(x + \frac{1}{n}) - F(x)}{\frac{1}{n}}$$

Since F is differentiable a.e., we have $f_n \to F'$ a.e.. Applying Fatou's Lemma we have

$$\begin{aligned}
\int_{[a,b]} F'(x)dx &= \int_{[a,b]} \lim_{n\to\infty} f_n(x)dx \\
&\leq \liminf_{n\to\infty} \int_{[a,b]} f_n(x)dx \\
&= \liminf_{n\to\infty} n \int_{[a,b]} (F(x + \frac{1}{n}) - F(x))dx \\
&= \liminf_{n\to\infty} n(\int_{[a+\frac{1}{n}, b+\frac{1}{n}]} - \int_{[a,b]})F(x)dx \\
&= \liminf_{n\to\infty} n(\int_{[b,b+\frac{1}{n}]} - \int_{[a,a+\frac{1}{n}]})F(x)dx \\
&= \liminf_{n\to\infty}(F(b) - n\int_{[a,a+\frac{1}{n}]} F(x)dx) \\
&\leq \liminf_{n\to\infty} F(b) - F(a) \text{ by monotonicity} \\
&= F(b) - F(a)
\end{aligned}$$

□

Corollary 1.6.2. *If $F : [a, b] \to \mathbb{R}$ is monotone non-decreasing and bounded, then $F' \in L^1([a, b])$.*

Proof. From the previous proposition,

$$\int_{[a,b]} |F'(x)|dx = \int_{[a,b]} F'(x)dx \leq F(b) - F(a) < \infty$$

□

Theorem 1.6.7. *(Second fundamental theorem for Lipschitz functions). Let $F : [a, b] \to \mathbb{R}$ be Lipschitz continuous. Show that $\int_{[a,b]} F'(x)\,dx = F(b) - F(a)$. (Hint: Argue as in the proof of Proposition 1.6.37, but use the dominated convergence theorem (Theorem 1.4.49) in place of Fatou's lemma (Corollary 1.4.47).)*

Proof. First we verify that $F : [a, b] \to \mathbb{R}$ as a Lipschitz function is bounded. Suppose, on the contrary, that it is not bounded, i.e. $\exists x \in [a.b]$ such that $|F(x)| > M$ for all $M \in \mathbb{R}$, then, for all $y \in [a, b]$,

$$M - |F(y)| \leq |F(x)| - |f(y)| \leq |F(x) - F(y)| \leq C|x - y| \ \forall M \in \mathbb{R}$$

such a constant C cannot exists, and F is not Lipschitz continuous. Therefore, F is bounded. Then,

$$\int_{[a,b]} |f| \leq |b - a| \sup_{x\in[a,b]} |f| < \infty$$

and we can apply the Lebesgue Differentiation Theorem (LDT) to the locally integrable F later on.
Extend F and let $F(x) = F(a)$ for $x < a$, and $F(x) = F(b)$ for $x > b$. Then define

$$f_n(x) = \frac{F(x + \frac{1}{n}) - F(x)}{\frac{1}{n}}$$

Since F is differentiable a.e. shown in Theorem 1.6.6., we have $f_n \to F'$ a.e..
Also, for all n,

$$|f_n| \leq n|F(x + \frac{1}{n}) - F(x)| \leq nC|x + \frac{1}{n} - x| = C$$

and C as a function is in $L^1(([a,b], \mathcal{B}([a,b]), m))$. Therefore we can apply the Dominated Convergence Theorem.

$$\begin{aligned}
\int_{[a,b]} F'(x) &= \int_{[a,b]} \lim_{n\to\infty} f_n(x)dx \\
&= \lim_{n\to\infty} \int_{[a,b]} f_n(x)dx \\
&= \lim_{n\to\infty} \int_{[a,b]} n(F(x + \frac{1}{n}) - F(x))dx \\
&= \lim_{n\to\infty} n(\int_{[a+\frac{1}{n}, b+\frac{1}{n}]} - \int_{[a,b]})F(x)dx \\
&= \lim_{n\to\infty} n(\int_{[b, b+\frac{1}{n}]} - \int_{[a, a+\frac{1}{n}]})F(x)dx \\
&= \lim_{n\to\infty} (F(b) - n\int_{[a, a+\frac{1}{n}]} F(x)dx) \\
&= F(b) - \lim_{n\to\infty} n\int_{[a, a+\frac{1}{n}]} F(x)dx \\
&= F(b) - F(a) \text{ by LDT on } f \text{ restricted in local } [a,b]
\end{aligned}$$

□

The assumption of continuous monotone is not enough, since it is possible for all the fluctuation of a continuous monotone function to be concentrated in an uncountable set of zero measure. For example, the Cantor's Function.

Define the functions $F_0, F_1, F_2, \cdots : [0,1] \to \mathbb{R}$ recursively as follows:
1. Set $F_0(x) := x$ for all $x \in [0,1]$.
2. For each $n = 1, 2, \ldots$ in turn, define

$$F_n(x) := \begin{cases} \frac{1}{2}F_{n-1}(3x) & \text{if } x \in [0, 1/3]; \\ \frac{1}{2} & \text{if } x \in (1/3, 2/3); \\ \frac{1}{2} + \frac{1}{2}F_{n-1}(3x - 2) & \text{if } x \in [2/3, 1]. \end{cases}$$

(i) Graph F_0, F_1, F_2, and F_3 (preferably on a single graph).
(ii) Show that for each $n = 0, 1, \ldots, F_n$ is a continuous monotone non-decreasing function with $F_n(0) = 0$ and $F_n(1) = 1$. (*Hint: induct on n.*)

(iii) Show that for each $n = 0, 1, \dots,$ one has $|F_{n+1}(x) - F_n(x)| \le 2^{-n}$ for each $x \in [0,1]$. Conclude that the F_n converge uniformly to a limit $F : [0,1] \to \mathbb{R}$. This limit is known as the *Cantor function*.

(iv) Show that the Cantor function F is continuous and monotone non-decreasing, with $F(0) = 0$ and $F(1) = 1$.

(v) Show that if $x \in [0,1]$ lies outside the middle thirds Cantor set (Exercise 1.2.9), then F is constant in a neighbourhood of x, and in particular $F'(x) = 0$. Conclude that $\int_{[0,1]} F'(x)\,dx = 0 \ne 1 = F(1) - F(0)$, so that the second fundamental theorem of calculus fails for this function.

(vi) Show that $F\left(\sum_{n=1}^{\infty} a_n 3^{-n}\right) = \sum_{n=1}^{\infty} \frac{a_n}{2} 2^{-n}$ for any digits $a_1, a_2, \dots \in \{0,2\}$. Thus the Cantor function, in some sense, converts base three expansions to base two expansions.

(vii) Let $I = \left[\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n}\right]$ be one of the intervals used in the nth cover I_n of C (see Exercise 1.2.9), thus $n \ge 0$ and $a_1, \dots, a_n \in \{0,2\}$. Show that I is an interval of length 3^{-n}, but $F(I)$ is an interval of length 2^{-n}.

(viii) Show that F is not differentiable at any element of the Cantor set C.

Proof. (i). See Figure.1.

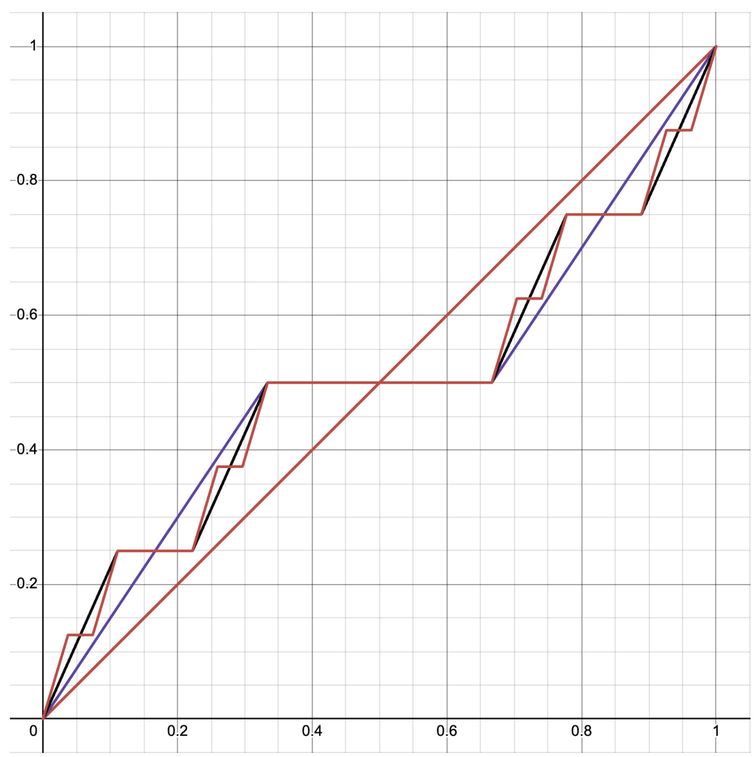

Figure 1: Graph of F_0, F_1, F_2, and F_3

(ii). Case 1: since $F_0(x) = x$, we can easily see that it is continuous monotone non-decreasing, with $F_0(0) = 0$ and $F_0(1) = 1$.

Case 2: suppose that, for $n \ge 1$, F_{n-1} is continuous monotone non-decreasing, with $F_{n-1}(0) = 0$ and $F_{n-1}(1) = 1$.

When $x \in [0, \frac{1}{3}]$, since $3x$ is continuous monotone non-decreasing and F_{n-1} is continuous monotone non-decreaing, we have $\frac{1}{2}F_{n-1}(3x)$ is continuous monotone non-decreaing, and

we have $\frac{1}{2}F_{n-1}(3 \times \frac{1}{3}) = \frac{1}{2}$. When $x \in (\frac{1}{3}, \frac{2}{3})$, $F_n(x) = \frac{1}{2}$ and thus continuous monotone non-decreasing. When $x \in [\frac{2}{3}, 1]$, since $3x - 2$ is continuous monotone-non-decreasing and F_{n-1} is continuous monotone non-decreasing, we have $\frac{1}{2} + \frac{1}{2}F_{n-1}(3x-2)$ is continuous monotone non-decreasing, and we also have $\frac{1}{2} + \frac{1}{2}F_{n-1}(3 \times \frac{2}{3} - 2) = \frac{1}{2}$.
Therefore, by induction, F_n is continuous monotone non-decreasing for every n.

(iii). Case 1: $|F_1(x) - F_0(x)| \le (\frac{1}{2})^0 = 1$ is obvious.
Case 2: suppose for $n \ge 2$, $|F_n(x) - F_{n-1}(x)| \le 2^{-(n-1)}$ for each $x \in [0,1]$. When $x \in [0, \frac{1}{3}]$, $F_{n+1}(x) = \frac{1}{2}F_n(3x)$, $F_n(x) = \frac{1}{2}F_{n-1}(3x)$, therefore $|F_{n+1}(x) - F_n(x)| = \frac{1}{2}|F_n(3x) - F_{n-1}(3x)| \le 2^{-n}$ for each $3x \in [0,1]$. When $x \in (\frac{1}{3}, \frac{2}{3})$, $|F_{n+1}(x) - F_n(x)| = 0$. Similarly, when $x \in [\frac{2}{3}, 1]$, $|F_{n+1}(x) - F_n(x)| = \frac{1}{2}|F_n(3x-2) - F_{n-1}(3x-2)| \le 2^{-n}$ for each $3x - 2 \in [0,1]$. Therefore $|F_{n+1}(x) - F_n(x)| \le 2^{-n}$ for each $x \in [0,1]$.
For every $x \in [0,1]$, without loss of generality, let $m > n$, $|F_m - F_n| \le |F_{n+1} - F_n| + |F_{n+2} = F_{n+1}| + ... + |F_m - F_{m-1}| \le 2^{-(m-1)} + ... + 2^{-n} = 2((\frac{1}{2})^{n+1} - (\frac{1}{2})^m)$ and for every $\epsilon > 0$ we can find an N, such that when $m > n \ge N$, $|F_m - F_n| \le \epsilon$, and F_n is a Cauchy sequence and thus converge uniformly to a limit F (since N applies to every $x \in [0,1]$).

(iv). Fist we show F is continuous. Since F_n is continuous and $F_n \to F$ uniformly on $[0,1]$. Let $x_0 \in [0,1]$ and $\epsilon > 0$. There exists an N so that $|F_n(x) - F(x)| \le \frac{\epsilon}{3}$ for all $n \ge N$ and all $x \in [0,1]$. Choose an $n_0 \ge N$.

$$\begin{aligned}|F(x) - F(x_0)| &\le |F(x) - F_{n_0}(x)| + |F_{n_0}(x) - F_{n_0}(x_0)| + |F_{n_0}(x_0) - F(x_0)| \\ &\le \frac{\epsilon}{3} + |F_{n_0}(x) - F_{n_0}(x_0)| + \frac{\epsilon}{3}\end{aligned}$$

Since F_{n_0} is continuous at x_0, there exists $\delta > 0$ so that $|x - x_0| \le \delta$ implies that $|F_{n_0}(x) - F_{n_0}(x_0)| \le \frac{\epsilon}{3}$. Therefore, when $|x - x_0| \le \delta$,

$$|F(x) - F(x_0)| \le \epsilon$$

and thus F is continuous at x_0. Since x_0 is arbitrary in $[0,1]$, F is continuous in $[0,1]$.
Now we show that F is monotone non-decreasing. Suppose, on the contrary, that $\exists 0 \le a < b \le 1$ such that $F(a) > F(b)$. Let $\epsilon = \frac{2F(a) - F(b)}{3} > 0$. Since uniform convergence, $\exists N$, when $n \ge N$, $|F_n(a) - F(a)| \le \epsilon$ and $|F_n(b) - F(b)| \le \epsilon$. Together with monotone non-decreasing of F_n, this implies that $F(a) - \epsilon \le F_n(a) \le F_n(b) \le F(b) + \epsilon$. Thus $F(a) \le F(b) + 2\epsilon$. Plug in $\epsilon = \frac{2F(a) - F(b)}{3} > 0$, we have $F(a) \le F(b)$. Contradiction! Therefore we have F is monotone non-decreasing.

(v). The middle third Cantor Set C is constructed as this way:

$$I_n := \bigcup_{a_1, \dots, a_n \in \{0,2\}} [\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n}]$$

$$C := \bigcap_{n=1}^{\infty} I_n$$

From the construction of Cantor function, we can clearly see that

$$C_1 := \{x \in [0,1] : F_1 \text{ is constant in a neighborhood of } x\} = I_1^C$$

......

$$C_n := \{x \in [0,1] : F_n \text{ is constant in a neighborhood of } x\} = I_n^C$$

......

$$C' := \{x \in [0,1] : F \text{ is constant in a neighborhood of } x\}$$

And from the construction of Cantor Set we can see that

$$I_1^C \subset I_2^C \subset I_3^C \subset ...$$

Since $F_n \to F$ uniformly, we have $C_n \to C$. That is, $I_n^C \to C'$, and $I_n^C \subset C'$ $\forall n$.
If $x \notin C$, then $x \in \bigcup_{n=1}^{\infty} I_n^C$, that is, $\exists n$, $x \in I_n^C$. Therefore, $x \in I_{n+1}^C \subset I_{n+2}^C, ...$ Since $I_n^C \to C'$ and $I_n^C \subset C'$ $\forall n$, we have $x \in C'$. Therefore, F is constant in a neighborhood of x if $x \notin C$. Therefore, in C', $F'(x) = 0$.
From this argument it is also very clear that $C \cup C' = [0,1]$, $F'(x) = 0$ if $x \in C' = [0,1] \setminus C$. Therefore,

$$\begin{aligned}\int_{[0,1]} F' dm &= \int_C F' dm + \int_{[0,1]\setminus C} F' dm \\ &= 0 + \int_{[0,1]\setminus C} 0 \; dm \text{ since Cantor Set has measure zero} \\ &= 0 \end{aligned} \tag{8}$$

(vi). Since $F_n \to F$ uniformly, it suffices to show that $\forall n$,

$$F_n(x) = \sum_{k=1}^{n} (\frac{a_n}{2}) 2^{-k}$$

where $x = \sum_{k=1}^{\infty} a_n 3^{-n}$.
The base case is $F_0(x) = x$, which does not depend on any a_n, and the base case trivially holds.
The inductive step is as follow. Assume for fixed $n \geq 1$, for all $x \in C$,

$$F_{n-1}(x) = \sum_{k=1}^{n-1} (\frac{a_k}{2}) 2^{-k}$$

We need to show that

$$F_n(x) = \sum_{k=1}^{n} (\frac{a_k}{2}) 2^{-k}$$

Case 1: $a_1 = 0$. In this case,

$$x = \sum_{k=1}^{\infty} a_k 3^{-k} = 0 + a_2 3^{-2} + a_3 3^{-3} + ... \in [0, \frac{1}{3}]$$

and $\frac{a_1}{2} = 0$, $F_n(x) = \frac{1}{2}F_{n-1}(3x)$. Also we have

$$3x = \sum_{k=1}^{\infty} a_{k+1}3^{-k}$$

and the induction assumption becomes

$$F_{n-1}(3x) = \sum_{k=1}^{n-1} (\frac{a_{k+1}}{2})2^{-k}$$

Therefore,

$$F_n(x) = \frac{1}{2}\sum_{k=1}^{n-1}(\frac{a_{k+1}}{2})2^{-k} = \sum_{k=1}^{n-1}(\frac{a_{k-1}}{2})2^{-(k+1)} = \sum_{k=1}^{n}(\frac{a_k}{2})2^{-k}$$

Case 2: $a_1 = 2$. In this case,

$$x = \sum_{k=1}^{\infty} a_k 3^{-k} = \frac{2}{3} + a_2 3^{-2} + a_3 3^{-3} + ... \in [0, \frac{1}{3}]$$

and $\frac{a_1}{2} = 1$, $F_n(x) = \frac{1}{2} + \frac{1}{2}F_{n-1}(3x-2)$. Also we have

$$3x - 2 = \sum_{k=2}^{\infty} a_k 3^{-(k-1)} = \sum_{k=1}^{\infty} a_{k+1}3^{-k}$$

and the induction assumption becomes

$$F_{n-1}(3x-2) = \sum_{k=1}^{n-1}(\frac{a_{k+1}}{2})2^{-k}$$

Therefore,

$$F_n(x) = \frac{1}{2} + \frac{1}{2}\sum_{k=1}^{n-1}(\frac{a_{k+1}}{2})2^{-k} = \frac{1}{2} + \sum_{k=1}^{n}(\frac{a_k}{2})2^{-k} = \sum_{k=1}^{n}(\frac{a_k}{2})2^{-k}$$

(vii).

$$|I| = |\sum_{i=1}^{n}\frac{a_i}{3^i} + \frac{1}{3^n} - \sum_{i=1}^{n}\frac{a_i}{3^i}| = \frac{1}{3^n}$$

We also have seen from the construction of Cantor function that, F_n is linear in corresponding $I_n = \sum_{a_1,..,a_i \in \{0,2\}} I$ and has

$$F_n' = (\frac{3}{2})^n$$

Then,

$$|F(I)| = \frac{1}{3^n}(\frac{3}{2})^n = 2^{-n}$$

(viii). Let $x \in \bigcap_{n=1}^{\infty} I_n$, then $x \in I_n$ $\forall n$. That is, x is either the left endpoint or the right endpoint of each of intervals consisting I_n for each n.

If x is a left endpoint, then

$$\lim_{n\to\infty} \frac{F_n(x+\frac{1}{3^n}) - F_n(x)}{\frac{1}{3^n}} = \lim_{n\to\infty} (\frac{3}{2})^n = \infty$$

$$\lim_{n\to\infty} \frac{F_n(x-\frac{1}{3^n}) - F_n(x)}{-\frac{1}{3^n}} = 0$$

Not differentiable. Similarly, if x is a right endpoint, then

$$\lim_{n\to\infty} \frac{F_n(x-\frac{1}{3^n}) - F_n(x)}{-\frac{1}{3^n}} = \lim_{n\to\infty} (\frac{3}{2})^n = \infty$$

$$\lim_{n\to\infty} \frac{F_n(x+\frac{1}{3^n}) - F_n(x)}{\frac{1}{3^n}} = 0$$

Also $F_n \to F$ uniformly, so F is not differentiable at $x \in C$.

To recover the second fundamental theorem, we need another hypothesis.

□

Definition 1.6.8. *(Absolute Continuity). For a possibly infinite interval I, a function $F : I \to \mathbb{C}$ is absolutely continuous if $\forall \epsilon > 0$, $\exists \delta > 0$ such that for any finite collection of disjoint open intervals $(a_i, b_i) \subset I$, $1 \leq i \leq n$, of total length $\leq \delta$,*

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| \leq \epsilon$$

Proposition 1.6.8. *(i) Every absolutely continuous function is uniformly continuous and therefore continuous.*

(ii) Every absolutely continuous function is of bounded variation on every compact interval $[a,b]$. (Hint: first show this is true for any sufficiently small interval.) In particular (by Exercise 1.6.40), absolutely continuous functions are differentiable almost everywhere.

(iii) Every Lipschitz continuous function is absolutely continuous.

(iv) The function $x \mapsto \sqrt{x}$ is absolutely continuous, but not Lipschitz continuous, on the interval $[0,1]$.

*(v) The Cantor function from Exercise 1.6.47 is continuous, monotone, and uniformly continuous, but **not** absolutely continuous, on $[0,1]$.*

(vi) If $f : \mathbb{R} \to \mathbb{R}$ is absolutely integrable, then the indefinite integral $F(x) := \int_{-\infty}^{x} f(y)\,dy$ is absolutely continuous, and that F is differentiable almost everywhere with $F'(x) = f(x)$ for almost every x.

Proof. (i). Let $\epsilon > 0$. Since F is absolutely continuous, take the finite collection of disjoint intervals to be just (a,b), a single interval with $|b-a| \leq \delta$. Then we have $|F(b) - F(a)| \leq \epsilon$. Therefore it is uniformly continuous and therefore continuous.

(ii). Let $\epsilon > 0$. Since F is absolutely continuous on $[a,b]$,, then there exists $\delta > 0$, for all finite collection of open disjoint intervals $\{(x_i, x_{i+1})\}_{i=1}^{n-1}$ on $[a,b]$ such that $\sum_{i=1}^{n-1} |x_{i+1} -$

$x_i| = |x_n - x_0| \leq \delta$, we have $\sum_{i=1}^{n-1} |F(x_{i+1}) - F(x_i)| \leq \epsilon$.
Chop (a, b) arbitrarily into a union of finite open disjoint intervals $(a, b) = \bigcup_{j=1}^{m-1}(x^j, x^{j+1})$ with $\sup_{j=1,\ldots,m-1} |x^j - x^{j+1}| \leq \delta$, $x^1 = a$, $x^m = b$. Withing each (x^j, x^{j+1}), chop it arbitrarily again into a union of finite open disjoint intervals $(x^j, x^{j+1}) = \bigcup_{i=1}^{n_j-1}(x_i^j, x_{i+1}^j)$ with $x_1^j = x^j$, $x_{n_j}^j = x^{j+1}$. By this arbitrary twice-chopping, we can get partition of (a, b) into arbitrary union of open disjoint intervals

$$(a, b) = \bigcup_{j=1}^{m} \bigcup_{i=1}^{n_j-1} (x_i^j, x_{i+1}^j)$$

Then by absolute continuity, we have $\sum_{i=1}^{n_j-1} |F(x_{i+1}^j) - F(x_i^j))| \leq \epsilon$ and then

$$\sum_{j=1}^{m} \sum_{i=1}^{n_j-1} |F(x_{i+1}^j) - F(x_i^j))| \leq m\epsilon < \infty$$

Therefore, for any partition of (a, b) into finite open disjoint unions of intervals, we have the sum of variations to be finite. Therefore, taking supremum to both LHS and RHS, we have

$$||F||_{TV([a,b])} = \sup_{m<\infty, n_j<\infty} \sum_{j=1}^{m} \sum_{i=1}^{n_j-1} |F(x_{i+1}^j) - F(x_i^j)| < \infty$$

Therefore F is of bounded variation on $[a, b]$, and hence differentiable a.e..

(iii). Let $\epsilon > 0$. Take arbitrary interval $I \subset \mathbb{R}$, such that $|I| \leq \epsilon$. Then, for any finite collection of disjoint open interval $\{(a_i, b_i)\}_{i=1}^n \subset I$, we have

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| \leq C \sum_{i=1}^{n} |b_i - a_i| \leq C\epsilon$$

Therefore, every Lipschitz continuous function is absolutely continuous.

(iv). Let $\epsilon > 0$. Since it is bounded continuous monotone non-decreasing in $[0, 1]$, by **Exercise 1.6.35.**, it is of bounded variation and

$$||F||_{TV((a,b))} = \sup_{a \leq x_1 < \ldots < x_n \leq b} \sum_{i=0}^{n-1} |F(x_i) - F(x_{i+1})| = |F(b) - F(a)|$$

Therefore, for any finite collection of disjoint open intervals $\{(a_i, b_i)\}_{i=1}^n \subset (a, b) \subset [0, 1]$,

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| \leq |F(b) - F(a)| = \sqrt{b} - \sqrt{a}$$

Therefore we can always find a $\delta > 0$ such that when $|b - a| \leq \delta$, $\sum_{i=1}^n |F(b_i) - F(a_i)| \leq \sqrt{b} - \sqrt{a} \leq \epsilon$. Therefore it is absolutely continuous.
Suppose, on the contrary, that is Lipschitz continuous, then the derivative of $F(x) = \sqrt{x}$ in $[0, 1]$ is bounded above by C, a constant. However,

$$\lim_{h \to 0} \frac{F(h) - F(0)}{h} = \lim_{h \to 0} \frac{1}{\sqrt{h}} = +\infty$$

Contradiction! Therefore, it is not Lipschitz continuous.

(v). Since $F_n \to F$ uniformly on $[0,1]$, and we have the theorem in mathematical analysis that says that continuous functions, if converge uniformly, converges to a uniformly continuous function, we have F to be uniformly continuous.

We use I'_n as a union of countable disjoint open intervals

$$I'_n := \bigcup_{a_1,...,a_n \in \{0,2\}} (\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n})$$

We see that it has 2^{n+1} endpoints and we rewrite $I'_n = \bigcup_{i=1}^{n}(a_i, b_i)$ where $0 = a_i < b_1 < a_2 < b_2 < ... < a_n < b_n = 1$, and we can see that

$$|I'_n| = (\frac{2}{3})^n$$

Let $\delta > 0$ be arbitrary, and we can always find $n \geq \log_{2/3} \delta$ such that $|I'_n| \leq \delta$.

However, from the construction of Cantor Function (that preserves the value at endpoints) we see that $F(a_i) = F_n(a_i) = F_n(b_{i-1}) = F(b_{i-1})\ \forall i = 2, ..., n,\ F(0) = F_n(0) = 0$, $F(1) = F_n(1) = 1$. Therefore,

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| = \sum_{i=1}^{n} |F_n(b_i) - F_n(a_i)| = 1$$

Therefore, no matter how hard we shrink the size δ, we have found a finite collection of open disjoint invervals I'_n (with $n \geq \log_{3/2} \delta$), with $\sum_{i=1}^{n} |F(b_i) - F(a_i)|$ always equaling one. Therefore, the Cantor function is not absolutely continuous.

(vi). Suppose, on the contrary, that it is not absolutely continuous. That is, $\exists \epsilon > 0$ such that $\forall \delta > 0$, there is a finite collection of disjoint open intervals $(a_i, b_i) \subset \mathbb{R}$, $1 \leq i \leq n$, of total length $\leq \delta$,

$$\sum_{i=1}^{n} |\int_{a_i}^{b_i} f(y)dy| \geq \epsilon$$

Also by triangle inequality,

$$\int_{\bigcup_{i=1}^{n}(a_i,b_i)} |f(y)|dy = \sum_{i=1}^{n} \int_{a_i}^{b_i} |f(y)|dy \geq \sum_{i=1}^{n} |\int_{a_i}^{b_i} f(y)dy| \geq \epsilon$$

Since for any $\delta > 0$ there is a finite $\bigcup_{i=1}^{n}(a_i, b_i)$ with total length $\leq \delta$ such that this is true, and $\mathbb{R}$ consists of infinitely many such finite collection of disjoint open intervals, we clearly have

$$\int_{\mathbb{R}} |f(y)|dy \geq \lim_{n\to\infty} n\epsilon = \infty$$

So f is not absolutely integrable. Therefore, if f is absolutely integrable, the indefinite integral $F(x) = \int_{-\infty}^{x} f(y)dy$ is absolutely continuous.

By (ii), $F(x) = \int_{-\infty}^{x} f(y)dy$ is differentiable a.e.

Therefore, since f is absolutely integrable and F is differentiable a.e., by Lebesgue Differentiation Theorem, for a.e. $x \in \mathbb{R}$,

$$F'(x) = \frac{1}{h} \lim_{h \to 0} \int_x^{x+h} f(y) dy = f(x)$$

□

For absolutely continuous functions, we can recover the second fundamental theorem of calculus.

We first prove Cousin's Theorem.

Theorem 1.6.8. *(Cousin's Theorem) Given any function $\delta : [a, b] \to (0, +\infty)$ on a compact interval $[a, b]$ of positive length, there exists a partition $a = t_0 < t_1 < ... < t_k = b$ with $k \geq 1$, together with real numbers $t_j^* \in [t_{j-1}, t_j]$ for each $1 \leq j \leq k$ and $t_j - t_{j-1} \leq \delta(t_j^*)$.*

Theorem 1.6.9. *Let $F : [a, b] \to \mathbb{R}$ be absolutely continuous. Then F' exists a.e., and*

$$\int_{[a,b]} F'(x) dx = F(b) - F(a)$$

Proof. Our main tool here will be *Cousin's theorem.*

By Exercise 1.6.43, F' is absolutely integrable. By Exercise 1.5.10, F' is thus uniformly integrable. Now let $\varepsilon > 0$. By Exercise 1.5.13, we can find $\kappa > 0$ such that

$$\int_U |F'(x)| \, dx \leq \varepsilon$$

whenever $U \subset [a, b]$ is a measurable set of measure at most κ. (Here we adopt the convention that F' vanishes outside of $[a, b]$.) By making κ small enough, we may also assume from absolute continuity that

$$\sum_{j=1}^n |F(b_j) - F(a_j)| \leq \varepsilon$$

whenever $(a_1, b_1), \ldots, (a_n, b_n)$ is a finite collection of disjoint intervals of total length $\sum_{j=1}^n b_j - a_j$ at most κ.

Let $E \subset [a, b]$ be the set of points x where F is not differentiable, together with the endpoints a, b, as well as the points where x is not a Lebesgue point of F'. Thus E is a null set. By outer regularity (or the definition of outer measure) we can find an open set U containing E of measure $m(U) < \kappa$. In particular,

$$\int_U |F'(x)| \, dx \leq \varepsilon.$$

Now define a gauge function $\delta : [a, b] \to (0, +\infty)$ as follows:

(i) If $x \in E$, we define $\delta(x) > 0$ to be small enough that the open interval $(x - \delta(x), x + \delta(x))$ lies in U.

(ii) If $x \notin E$, then F is differentiable at x and x is a Lebesgue point of F'. We let $\delta(x) > 0$ be small enough that

$$|F(y) - F(x) - (y - x)F'(x)| \leq \varepsilon |y - x|$$

holds whenever $|y-x| \leq \delta(x)$, and such that

$$\left| \frac{1}{|I|} \int_I F'(y)\,dy - F'(x) \right| \leq \varepsilon$$

whenever I is an interval containing x of length at most $\delta(x)$; such a $\delta(x)$ exists by the definition of differentiability, and of Lebesgue point. We rewrite these properties using big-O notation as

$$F(y) - F(x) = (y-x)F'(x) + O(\varepsilon|y-x|) \quad \text{and} \quad \int_I F'(y)\,dy = |I|F'(x) + O(\varepsilon|I|).$$

Applying Cousin's theorem, we can find a partition $a = t_0 < t_1 < \cdots < t_k = b$ with $k \geq 1$, together with real numbers $t_j^* \in [t_{j-1}, t_j]$ for each $1 \leq j \leq k$ and $t_j - t_{j-1} \leq \delta(t_j^*)$.

We can express $F(b) - F(a)$ as a telescoping series:

$$F(b) - F(a) = \sum_{j=1}^{k} \big(F(t_j) - F(t_{j-1})\big).$$

To estimate the size of this sum, let us first consider those j for which $t_j^* \in E$. Then, by construction, the intervals $[t_{j-1}, t_j]$ are disjoint in U. By construction of κ, we thus have

$$\sum_{j: t_j^* \in E} |F(t_j) - F(t_{j-1})| \leq \varepsilon$$

and thus

$$\sum_{j: t_j^* \in E} \big(F(t_j) - F(t_{j-1})\big) = O(\varepsilon).$$

Next, we consider those j for which $t_j^* \notin E$. By construction, for those j, we have

$$F(t_j) - F(t_j^*) = (t_j - t_j^*)F'(t_j^*) + O(\varepsilon|t_j - t_j^*|)$$

and

$$F(t_j^*) - F(t_{j-1}) = (t_j^* - t_{j-1})F'(t_j^*) + O(\varepsilon|t_j^* - t_{j-1}|),$$

and thus

$$F(t_j) - F(t_{j-1}) = (t_j - t_{j-1})F'(t_j^*) + O(\varepsilon|t_j - t_{j-1}|).$$

On the other hand, from construction again we have

$$\int_{[t_{j-1}, t_j]} F'(y)\,dy = (t_j - t_{j-1})F'(t_j^*) + O(\varepsilon|t_j - t_{j-1}|),$$

and thus

$$F(t_j) - F(t_{j-1}) = \int_{[t_{j-1}, t_j]} F'(y)\,dy + O(\varepsilon|t_j - t_{j-1}|).$$

Summing over j, we conclude that

$$\sum_{j: t_j^* \notin E} \big(F(t_j) - F(t_{j-1})\big) = \int_S F'(y)\,dy + O(\varepsilon(b-a)),$$

where S is the union of all the $[t_{j-1}, t_j]$ with $t_j^* \notin E$. By construction, this set is contained in $[a,b]$ and contains $[a,b] \setminus U$. Since

$$\int_U |F'(x)|\, dx \leq \varepsilon,$$

we conclude that

$$\int_S F'(y)\, dy = \int_{[a,b]} F'(y)\, dy + O(\varepsilon|b-a|).$$

Putting everything together, we conclude that

$$F(b) - F(a) = \int_{[a,b]} F'(y)\, dy + O(\varepsilon) + O(\varepsilon(b-a)).$$

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

$\square$

1.7 Construction of measures

In this section, we abstractify the construction of measure, from pre-measure to outer measure (exterior measure) to measure.

1.7.1 Outer measures and the Carathéodory extension theorem

Definition 1.7.1. *(Abstract outer measure). Let $\mathcal{X}$ be a set. An outer measure on $\mathcal{X}$ is a map $\mu^* : 2^{\mathcal{X}} \to [0, +\infty]$ satisfying*

1. $\mu^*(\emptyset) = 0$.
2. $E \subset F \Rightarrow \mu^*(E) \leq \mu^*(F)$.
3. $E_1, E_2, ... \subset \mathcal{X} \Rightarrow \mu^*(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu^*(E_n)$.

Outer measures can measure all subsets of $\mathcal{X}$, whereas measures can only measure a σ-algebra of measurable sets.

Definition 1.7.2. *(Carathéodory measurability). Let μ^* be an outer measure on $\mathcal{X}$. A set $E \subset \mathcal{X}$ is said to be Carathéodory measurable (or μ^*-measurable) if*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E) \ \forall A \subset \mathcal{X}$$

That is, E cuts every subset nicely into two.

Note that, by subadditivity, we always have $\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \setminus E) \ \forall A \subset \mathcal{X}$.

Lemma 1.7.1. *If E is null ($\mu^*(E) = 0$), then E is μ^*-measurable.*

Proof. We already have $\leq$. By monotonicity, we have $\mu^*(A \cap E) \leq \mu^*(E) = 0$ and $\mu^*(A \setminus E) \leq \mu^*(A)$. Therefore $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E)$. □

Now we construct measure from outer measure.

Definition 1.7.3. *(Complete measure) μ is a complete measure on $(\mathcal{X}, \mathcal{B})$ if $A \in \mathcal{B}$ and $\mu(A) = 0$, then for every $B \subseteq A$ we have $B \in \mathcal{B}$ and $\mu(B) = 0$.*

Theorem 1.7.1. *(Carathéodory extension theorem). If μ^* is an outer measure on $\mathcal{X}$, then the collection $\mathcal{B}$ of μ^*-measurable sets is a σ-algebra, and the restriction of μ^* to $\mathcal{B}$ ($\mu := \mu^*|_{\mathcal{B}}$) is a complete measure.*

Proof. We will show that $\mathcal{B}$ is a σ-algebra.
$\emptyset \in \mathcal{B}$ since $\emptyset$ is μ^*-null.
Since $A \setminus E^C = A \cap E$, $A \cap E^C = A \setminus E$, we have $\forall A \subset \mathcal{X}$, $\mu^*(A \setminus E^C) + \mu^*(A \cap E^C) = \mu^*(A \cap E) + \mu^*(A \setminus E) = \mu^*(A)$. Thus if $E \in \mathcal{B}$ then $E^C \in \mathcal{B}$.
To verify closedness under countable unions, we first check closedness under finite unions. Let $E, F \in \mathcal{B}$, $A \subset \mathcal{X}$. Let

$$A_{00} = A \setminus (E \cup F)$$

$$A_{10} = A \cap (E \setminus F)$$

$$A_{01} = A \cap (F \setminus E)$$

$$A_{11} = A \cap E \cap F$$

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap E) + \mu^*(A \setminus E) \\
&= \mu^*(A_{10}) + \mu^*(A_{11}) + \mu^*(A_{00}) + \mu^*(A_{01}) \\
&\geq \mu^*(A_{10} \cup A_{01} \cup A_{11}) + \mu^*(A_{00}) \\
&= \mu^*(A \cap (E \cup F)) + \mu^*(A \setminus (E \cup F))
\end{aligned}$$

Also by subadditivity we have

$$\mu^*(A) \leq \mu^*(A \cap (E \cup F)) + \mu^*(A \setminus (E \cup F))$$

Therefore,

$$\mu^*(A) = \mu^*(A \cap (E \cup F)) + \mu^*(A \setminus (E \cup F))$$

Then we proceed to check closure under countable unions. Let $E_1, E_2, ... \in \mathcal{B}$. We want to show that $\bigcup_{i=1}^{\infty} E_n \in \mathcal{B}$. By replacing E_n with $E_n \setminus (\bigcup_{i=1}^{n-1} E_i)$ which are also in $\mathcal{B}$ (by previous step) and have the same union, we may assume that E_n are disjoint.
Let $A \subset \mathcal{X}$ be arbitrary. It suffices to show that $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E)$ for $E = \bigcup_{n=1}^{\infty} E_n$.
Let $F_N := \bigcup_{n=1}^{N} E_n$, then $F_N \in \mathcal{B}$. By the same logic in previous steps,

$$\begin{aligned}
\mu^*(A) &\geq \mu^*(A \cap F_N) + \mu^*(A \setminus F_N) \\
&\geq \mu^*(A \cap F_N) + \mu^*(A \setminus E) \text{ by monotonicity since } F_N \subset E
\end{aligned}$$

Therefore, it suffices to show that

$$\lim_{N \to \infty} \mu^*(A \cap F_N) \geq \mu^*(A \cap E)$$

For any $N \in \mathbb{N}$, since $E_N \in \mathcal{B}$, we have

$$\begin{aligned}
\mu^*(A \cap F_N) &= \mu^*(A \cap F_N \cap E_N) + \mu^*(A \cap F_N \cap E_N^C) \\
&= \mu^*(A \cap E_N) + \mu^*(A \cap (\bigcup_{n=1}^{N-1} E_n)) \\
&= \mu^*(A \cap E_N) + \mu^*(A \cap F_{N-1})
\end{aligned}$$

By iteration, we get

$$\mu^*(A \cap F_N) = \sum_{n=1}^{N} \mu^*(A \cap E_n)$$

Therefore,

$$\begin{aligned}
\mu^*(A \cap E) &= \mu^*(\bigcup_{n=1}^{\infty} A \cap E_n) \\
&\leq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) \\
&= \lim_{N\to\infty} \sum_{i=1}^{N} \mu^*(A \cap E_N) \\
&= \lim_{N\to\infty} \mu^*(A \cap F_N)
\end{aligned}$$

Therefore, we have closure under countable unions.
Next we verify that the restriction of $\mu = \mu^*|_{\mathcal{B}}$ to $\mathcal{B}$ is a measure. Since we have shown that

$$\begin{aligned}
\mu^*(A) &\leq \mu^*(A \cap E) + \mu^*(A \setminus E) \\
&\leq \mu^*(A \setminus E) + \sum_{n=1}^{\infty} \mu^*(A \cap E_n) \\
&\leq \mu^*(A)
\end{aligned}$$

all the $\leq$ are $=$. Taking $A = E$, we immediately get σ-additivity,

$$\mu(E) = \mu^*(E) = \sum_{n=1}^{\infty} \mu^*(E_n) = \sum_{n=1}^{\infty} \mu(E_n)$$

Therefore μ is σ-additive on $\mathcal{B}$. Since $\emptyset \in \mathcal{B}$, we have μ also satisfies $\mu(\emptyset) = 0$.
Finally, we show that μ is a complete masure. Let $A \in \mathcal{B}$ such that $\mu(A) = 0$. Then, for every $B \subseteq A$, we want to verify that $\forall C \in \mathcal{B}$, $\mu^*(C) \geq \mu^*(C \cap B) + \mu^*(C \setminus B)$. This is indeed the case because

$$\mu^*(C \cap B) \leq \mu^*(C \cap A) \leq \mu^*(A) = 0$$

$$\mu^*(C \setminus B) \leq \mu^*(C)$$

Therefore, $B \in \mathcal{B}$ and $\mu^*(B) \leq \mu^*(A) = 0$. $\mu = \mu^*|_{\mathcal{B}}$ is a complete measure on $(\mathcal{X}, \mathcal{B})$.

□

1.7.2 Pre-measures

We now want to abstractify elementary measure to pre-measure. We want to extend a finitely additive measure μ_0 on a Boolean algebra $\mathcal{B}_0$ to a measure μ on a σ-algebra $\mathcal{B}$ such that $\mathcal{B} \supset \mathcal{B}_0$, $\mu|_{\mathcal{B}_0} = \mu_0$. For this, it is necessary that μ_0 is σ-additive on $\mathcal{B}_0$, i.e., whenever $E_n \in \mathcal{B}_0$ are disjoint,

$$\mu_0(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu_0(E_n)$$

Definition 1.7.4. *(Pre-measure). A pre-measure on a Boolean algebra $\mathcal{B}_0$ is a finitely additive measure $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ with the property that $\mu_0(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu_0(E_n)$ whenever $E_n \in \mathcal{B}_0$ are disjoint, and $\mu_0(\emptyset) = 0$.*

We can show that elementary measure on the Boolean algebra of all sets in $\mathbb{R}^d$ that are elementary or co-elementary (complement of an elementary set) is a pre-measure.

It turns out that the pre-measure condition is necessary and also sufficient.

Theorem 1.7.2. *(Hahn-Kolmogorov extension theorem). Let μ_0 be a pre-measure on a Boolean algebra $\mathcal{B}_0$ over a set $\mathcal{X}$. Then,*

$$\mu^*(E) := \inf\{\sum_{n=1}^{\infty} \mu_0(E_n) : E \subseteq \bigcup_{n=1}^{\infty} E_n; E_n \in \mathcal{B}_0 \ \forall n\}$$

is an outer measure on $\mathcal{X}$, and the restriction to the σ-algebra $\mathcal{B}$ of μ^*-measurable set is a measure, whose restriction to $\mathcal{B}_0$ is μ_0.

Proof. μ^* is an outer measure is easy to verify. First, we see that $\mu^*(\emptyset) = \mu_0(\emptyset) = 0$. Then, if $E \subset F$, then, for every $\{F_n\}$ such that $E \subset F \subset \bigcup_{n=1}^{\infty} F_n$,

$$\mu^*(E) \leq \sum_{n=1}^{\infty} \mu_0(F_n)$$

Therefore, $\mu^*(E) \leq \inf\{\sum_{n=1}^{\infty} \mu_0(F_n)\} = \mu^*(F)$.
Third, we want to verify that for $E_1, E_2, ... \in \mathcal{B}_0$, we have $\mu^*(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu^*(E_n)$. For each i, we have $\mu^*(E_i) = \inf\{\sum_{j=1}^{\infty} \mu_0(E_i^j) : E_i \subset \bigcup_{j=1}^{\infty} E_i^j; E_i^j \in \mathcal{B}, \forall j\}$. Therefore, there exists $E_i^1, E_i^2, ...$ such that $\sum_{j=1}^{\infty} \mu_0(E_i^j) \leq \mu^*(E_i) + \frac{\epsilon}{2^i}$. Since $\bigcup_{i=1} E_i \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} E_j^i$, we have

$$\begin{aligned}
\mu^*(\bigcup_{i=1}^{\infty} E_i) &\leq \mu^*(\bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} E_i^j) \\
&\leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu_0(E_i^j) \\
&= \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} \mu_0(E_i^j)) \\
&\leq \sum_{i=1}^{\infty} \mu^*(E_i) + \epsilon
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have $\mu^*(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$. Therefore $\mu^*(E)$ is an outer measure.
Let $\mathcal{B}$ be the collection of all sets $E \subset \mathcal{X}$ that are μ^*-measurable. and let $\mu = \mu^*|_{\mathcal{B}}$. Then by Carathéodory extension theorem, $\mathcal{B}$ is a σ-algebra, and μ is a complete measure.
Next, we verify $\mathcal{B}_0 \subset \mathcal{B}$. Let $E \in \mathcal{B}_0$, we want to show that E is μ^*-measurable, $\mu^*(E) = \mu_0(E)$. Let $A \in \mathcal{X}$ arbitrary, we need to show that $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E)$. By finite subadditivity, it suffices to show $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E)$. Assume that $\mu^*(A) < \infty$. Fix $\epsilon > 0$, then by definition of μ^*, there exists $E_1, E_2, ... \in \mathcal{B}_0$, $A \subset \bigcup_{n=1}^{\infty} E_n$, such that

$$\sum_{n=1}^{\infty} \mu_0(E_n) \leq \mu^*(A) + \epsilon$$

Since $E_n \cap E \in \mathcal{B}_0$ and $A \cap E \subset \bigcup_{n=1}^{\infty} E_n \cap E$, we have

$$\mu^*(A \cap E) \leq \sum_{n=1}^{\infty} \mu_0(E_n \cap E)$$

Similarly, $E_n \setminus E \in \mathcal{B}_0$, $A \setminus E \subset \bigcup_{n=1}^{\infty} E_n \setminus E$, we have

$$\mu^*(A \setminus E) \leq \sum_{n=1}^{\infty} \mu_0(E_n \setminus E)$$

Also, from finite additivity, $\mu_0(E_n \cap E) + \mu_0(E_n \setminus E) = \mu_0(E_n)$. Combining all those estimates,

$$\mu^*(A \cap E) + \mu^*(A \setminus E) \leq \sum_{n=1}^{\infty} \mu_0(E_n) \leq \mu^*(A) + \epsilon$$

Since $\epsilon > 0$ is arbitrary, we have $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E)$.
Finally, we have to show that $\mu^*(E) = \mu_0(E)$ if $E \in \mathcal{B}_0$. Since E covers itself, $\mu^*(E) \leq \mu_0(E)$. To show $\geq$, it is enough to show that $\forall E_1, E_2, ... \in \mathcal{B}_0$ and $E \subset \bigcup_{n=1}^{\infty} E_n$,

$$\sum_{n=1}^{\infty} \mu_0(E_n) \geq \mu_0(E)$$

By replacing E_n with $E_n \setminus \bigcup_{m=1}^{n-1} E_m$ (whose union still covers E), we may assume that E_n is disjoint. Further by replacing E_n with $E_n \cap E$, we may assume that $\bigcup_{n=1}^{\infty} E_n = E$. Then, by the hypothesis that μ_0 is a pre-measure, we have

$$\mu_0(E) = \mu_0(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu_0(E)$$

$\square$

The μ constructed in the above proof is the Hahn-Kolmogorov extension of the pre-measure μ_0. For instance, the Hahn-Kolmogorov extension of elementary measure is the Lebesgue measure.

μ_0 is σ-finite means that one can express the whole space $\mathcal{X}$ as the countable union of sets $E_1, E_2, ... \in \mathcal{B}_0$ for which $\mu_0(E_n) < \infty$ for all n.

We can use Hahn-Kolmogorov extension to construct many important class of measures, such as Lebesgue-Stieltjes measures, Product measures, and Hausdorff measures.

1.7.3 Lebesgue-Stieltjes measure

Theorem 1.7.3. *(Existence of Lebesgue-Stieltjes measure). Let $F : \mathbb{R} \to \mathbb{R}$ be monotone non-decreasing. Denote*

$$F_-(x) = \sup_{y<x} F(y)$$

$$F_+(x) = \inf_{y>x} F(y)$$

Then there exists a unique Borel measure $\mu_F : \mathcal{B}[\mathbb{R}] \to [0, +\infty]$, such that $\forall -\infty < a < b < +\infty$,

$$\mu_F([a,b]) = F_+(b) - F_-(a)$$

$$\mu_F([a,b)) = F_-(b) - F_-(a)$$

$$\mu_F((a,b]) = F_+(b) - F_+(a)$$

$$\mu_F((a,b)) = F_-(b) - F_+(a)$$

$$\mu_F(\{a\}) = F_+(a) - F_-(a)$$

Proof. Define the F-volume of any interval I as $|I|_F \in [0, +\infty]$ with the convention $F_-(+\infty) = \sup_{y\in\mathbb{R}} F(y)$ and $F_+(-\infty) = \inf_{y\in\mathbb{R}} F(y)$, $|\emptyset| = 0$.
Then, for I,J disjoint and shares a common endpoint, we have $|I \cup J|_F = |I|_F + |J|_F$. If $I = \bigcup_{i=1}^k I_i$ dosjoint, we have $|I| = |I_1| + ... + |I_k|$.
Let $\mathcal{B}_0$ be the Boolean algebra generated by the intervals, then $\mathcal{B}_0$ consists of those sets that can be expressed as a finite union of intervals. We can define a measure μ_0 on this algebra by declaring

$$\mu_0(E) = |I_1|_F + ... + |I_k|_F$$

where $E = I_1 cup... \cup I_k$ disjoint. This measure is well defined and is finitely additive.
We claim that μ_0 is a pre-measure. Indeed, suppose that $E \in \mathcal{B}_0$, $E = \bigcup_{i=1}^\infty E_i$ where $\{E_i\} \in \mathcal{B}_0$ disjoint. We want to show that

$$\mu_0(E) = \sum_{n=1}^{\infty} \mu(E_n)$$

By splitting up E into intervals and then intersecting each of the E_n with these intervals and using finite additivity, we may assume that E is a single interval. By splitting up the E_n into their component intervals and using finite additivity, we many assume that the E_n are also individual intervals.
By subadditivity, it suffices to show

$$\mu_0(E) = \sum_{n=1}^{\infty} \mu_0(E_n)$$

By definition of $\mu_0(E)$,

$$\mu_0(E) = \sup_{K\subset E, K \text{ compact}} \mu_0(K)$$

Thus it suffices to show that for $K \subset E$ compact,

$$\mu_0(K) \le \sum_{n=1}^{\infty} \mu_0(E)$$

In a similar spirit, we can show that

$$\mu_0(E_n) = \inf_{U_n \supset E_n, U_n \text{ open}} \mu_0(U_n)$$

Using the $\frac{\epsilon}{2^n}$ trick, it thus suffices to show that

$$\mu_0(K) \leq \sum_{n=1}^{\infty} \mu_0(U_n)$$

whenever U_n is an open interval containing E_n. By the Heine-Borel theorem, one can cover K by a finite number $\bigcup_{n=1}^{N} U_n$ of the U_n, hence by finite subadditivity,

$$\mu_0(K) \leq \sum_{n=1}^{N} \mu_0(U_n)$$

Then the claim holds.

Since now μ_0 is a pre measure, we can use Hahn-Kolmogorov extension theorem to extend it to a countably additive measure μ on a σ-algebra that contains $\mathcal{B}_0$. In particular, $\mathcal{B}$ contains all the elementary sets and hence contaihns the Borel σ-algebra. Restricting μ to the Borel σ-algebra we obtain the existence claim.

Now we establish uniqueness. If μ' is another Borel measure with the stated properties, then $\mu'(K) = |K|_F$ for every compact K, and hence by $\mu'(E) = \sup_{K \subset E} \mu'(K)$ and upward monotone convergence, one has $\mu'(I) = |I|_F$ for every interval$I$. Then μ' agrees with μ_0 on $\mathcal{B}_0$, and thus agrees with μ on Borel measurable sets. $\square$

For examples

$F(x) = x$, then $\mu_F([a, b]) = F(b) - F(a) = b - a = m([a, b])$.

$F(x) = cx + d$, then $\mu_F(E) = cm(E)$.

$H(x) = \mathbf{1}_{(-\infty,0]}(x)$, then $\mu_H(I) = 0$ whenever $0 \notin I$, $\mu_H(I) = 1$ if $0 \in I$.

We can show that this extends to any $E \in \mathcal{B}[\mathbb{R}]$ where $0 \in E$. Then $\mu_H(E) = \mathbf{1}_{0 \in E} = \delta_0(E)$.

$F(x) = \sum_{y \in E} \mathbf{1}_{[y,+\infty)}$ for some E countable, then μ_F is counting measure on E.

For F to be the Cantor function on $[0, 1]$ (extend to $\mathbb{R}$ by $F(x) = 0$ for $x < 0$, $F(x) = 1$ for $x > 1$), μ_F is the Cantor measure supported on Cantor set. F has no jumps $\Rightarrow \mu_F$ is "continuous", but it is singular with respect to Lebesgue measure.

Definition 1.7.5. *μ is absolutely continuous with respect to Lebesgue measure if*

$$\mu(E) = \int_E f dm$$

for some unsigned integrable f.

In one dimension, such μ is μ_F for F absolutely continuous with $F' = f$ a.e..

Given f, let F be such that $F' = f$, $F(x) = \int_{-\infty}^{x} f(y)dy$. Then $\mu_F([a, b]) = F(b) - F(a)$,

$$\begin{aligned}\mu_F(I) &= \int_a^b f(x)dx \\ &= \int_a^b F'(x)dx \\ &= F(b) - F(a)\end{aligned}$$

By second FTC of absolutely continuous functions. Sometimes we call F the (cumulative) distribution function for the Lebesgue-Stieltjes masure μ_F. And, if F is absolutely continuous, we call $f = F'$ the density for μ_F. Some important densities are

1. Gaussian
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

2. Cauchy (heavy tails)
$$\frac{1}{\pi} \frac{1}{1+x^2}$$

Proposition 1.7.1. *(Lebesgue-Stieltjes measure, absolutely continuous case).*

(i) If $F : \mathbb{R} \to \mathbb{R}$ is the identity function $F(x) = x$, show that μ_F is equal to Lebesgue measure m.

(ii) If $F : \mathbb{R} \to \mathbb{R}$ is monotone non-decreasing and absolutely continuous (which in particular implies that F' exists and is absolutely integrable), show that $\mu_F = m_{F'}$ in the sense of Exercise 1.4.49, thus

$$\mu_F(E) = \int_E F'(x)\, dx$$

for any Borel measurable E, and

$$\int_{\mathbb{R}} f(x)\, d\mu_F(x) = \int_{\mathbb{R}} f(x) F'(x)\, dx$$

for any unsigned Borel measurable $f : \mathbb{R} \to [0, +\infty]$.

Proof. (i). First we notice that for any $-\infty < b < a < +\infty$, since $F_-(x) = F(x) = F_+(x)$ everywhere,

$$\mu_F([a, b]) = F(b) - F(a) = m([b, a])$$

$$\mu_F([a, b)) = F(b) - F(a) = m([b, a))$$

$$\mu_F((a, b]) = F(b) - F(a) = m((b, a])$$

$$\mu_F((a, b)) = F(b) - F(a) = m((b, a))$$

$$\mu_F(\{a\}) = F(a) - F(a) = 0 = m(\{a\})$$

Also, by our previous exercise, $\mathcal{B}[\mathbb{R}]$ can be generated by both the open subsets and closed subsets of $\mathbb{R}$. Therefore, $\mu_F = m$ on $\mathcal{B}[\mathbb{R}]$.

(ii). Since F is absolutely continuous, F' exists a.e., and is absolutely integrable,

$$F(b) - F(a) = \int_{(a,b)} F'(x)dx = \int_{(a,b]} F'(x)dx = \int_{[a,b)} F'(x)dx = \int_{[a,b]} F'(x)dx$$

Also $F_-(x) = F(x) = F_+(x)$ everywhere. Therefore,

$$\mu_F([a, b]) = F(b) - F(a) = \int_{[a,b]} F'(x)dx$$

$$\mu_F([a, b)) = F(b) - F(a) = \int_{[a,b)} F'(x)dx$$

$$\mu_F((a,b]) = F(b) - F(a) = \int_{(a,b]} F'(x)dx$$

$$\mu_F((a,b)) = F(b) - F(a) = \int_{(a,b)} F'(x)dx$$

$$\mu_F(\{a\}) = F(a) - F(a) = 0 = \int_{\{a\}} F'(x)dx$$

Also, by our previous exercise, $\mathcal{B}[\mathbb{R}]$ can be generated by both the open subsets and closed subsets of $\mathbb{R}$. Therefore, $\mu_F(E) = \int_E F'(x)dx$ for any $E \in \mathcal{B}[\mathbb{R}]$.
Next, we prove that for any Borel measurable $f : \mathbb{R} \to [0, +\infty]$,

$$\int_{\mathbb{R}} f(x)d\mu_F(x) = \int_{\mathbb{R}} f(x)F'(x)dx$$

First, we look at unsigned simple functions. Suppose $f = \sum_{i=1}^n c_i \mathbf{1}_{E_i}$, $c_i \geq 0$, E_i is Borel measurable. WLOG, assume that E_i's are disjoint.

$$\begin{aligned}
\int_{\mathbb{R}} f(x)d\mu_F(x) &= Simp \int_{\mathbb{R}} f(x)d\mu_F(x) \\
&= \sum_{i=1}^n c_i \mu_F(E_i) \\
&= \sum_{i=1}^n c_i \int_{E_i} F'(x)dx \\
&= \sum_{i=1}^n \int_{E_i} c_i F'(x)dx \\
&= \sum_{i=1}^n \int_{\mathbb{R}} c_i \mathbf{1}_{E_i}(x)F'(x)dx \\
&= \int_{\mathbb{R}} (\sum_{i=1}^n c_i \mathbf{1}_{E_i}(x))F'(x)dx \\
&= \int_{\mathbb{R}} f(x)F'(x)dx
\end{aligned}$$

Next, we move to prove general case. For any Borel measurable f, $\exists! \{g_n\}$ unsigned simple functions such that $g_n \uparrow f$. Then, $g_n F' \uparrow fF'$ also. By monotone convergence theorem,

$$\begin{aligned}
\int_{\mathbb{R}} f(x)d\mu_F(x) &= \int_{\mathbb{R}} \lim_{n\to\infty} g_n(x)d\mu_F(x) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} g_n(x)d\mu_F(x) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} g_n(x)F'(x)dx \\
&= \int_{\mathbb{R}} \lim_{n\to\infty} g_n(x)F'(x)dx \\
&= \int_{\mathbb{R}} f(x)F'(x)dx
\end{aligned}$$

$\square$

Proposition 1.7.2. *(Lebesgue-Stieltjes measure, pure point case).*

(i) If $H : \mathbb{R} \to \mathbb{R}$ is the Heaviside function $H := 1_{[0,+\infty)}$, show that μ_H is equal to the Dirac measure δ_0 at the origin (defined in Example 1.4.22).

(ii) If $F = \sum_n c_n J_n$ is a jump function (as defined in Definition 1.6.30), show that μ_F is equal to the linear combination $\sum c_n \delta_{x_n}$ of Dirac measures (as defined in Exercise 1.4.22), where x_n is the point of discontinuity for the basic jump function J_n.

Proof. (i). To start with, let E be arbitrary interval on $\mathbb{R}$. Denote the start point and the end point as $s(E)$ and $e(E)$. Therefore, if $0 \in E$, we have $H_+(e(E)) = H_-(e(E)) = 1$, $H_+(s(E)) = H_-(s(E)) = 0$, and therefore $\mu_H(E) = 1$. If $0 \notin E$, we have $H_+(e(E)) = H_-(e(E)) = H_+(s(E)) = H_-(s(E))$, and therefore $\mu_H(E) = 0$. Also, by our previous exercise, $\mathcal{B}[\mathbb{R}]$ can be generated by both the open subsets and closed subsets of $\mathbb{R}$. Therefore, for arbitrary $E \in \mathcal{B}[\mathbb{R}]$,

$$\mu_H(E) = \begin{cases} 1 & \text{if } 0 \in E \\ 0 & \text{if } 0 \notin E \end{cases}$$

Therefore, $\mu_H = \delta_0$ where

$$\delta_0(E) = \begin{cases} 1 & \text{if } 0 \in E \\ 0 & \text{if } 0 \notin E \end{cases}$$

(ii). Again, Let E be arbitrary interval on $\mathbb{R}$. Denote the start point and the end point as $s(E)$ and $e(E)$. Suppose $x_n, x_{n+1}, ..., x_{n+m} \in E$, then

$$\sum_n C_n J_{n+}(e(E)) = \sum_n C_n J_{n-}(e(E)) = \sum_{i=1}^{n+m} c_i$$

$$\sum_n C_n J_{n+}(s(E)) = \sum_n C_n J_{n-}(s(E)) = \sum_{i=1}^{n-1} c_i$$

Therefore,

$$\mu_F(E) = \sum_{i=1}^{n+m} c_i - \sum_{i=1}^{n-1} c_i = \sum_{i=n}^{n+m} c_i = \sum_n c_n \delta_{x_n}(E)$$

where $\delta_{x_n} = \delta_{x_{n+1}} = ... = \delta_{x_{n+m}} = 1$, and all others are equal to zero.
Also, by our previous exercise, $\mathcal{B}[\mathbb{R}]$ can be generated by both the open subsets and closed subsets of $\mathbb{R}$. Therefore, for arbitrary $E \in \mathcal{B}[\mathbb{R}]$, $\mu_F(E) = \sum_n c_n \delta_{x_n}(E)$.

□

Proposition 1.7.3. *(Lebesgue-Stieltjes measure, singular continuous case).*

(i) If $F : \mathbb{R} \to \mathbb{R}$ is a monotone non-decreasing function, show that F is continuous if and only if $\mu_F(\{x\}) = 0$ for all $x \in \mathbb{R}$.

(ii) If F is the Cantor function (defined in Exercise 1.6.47), show that μ_F is a probability measure supported on the middle-thirds Cantor set (see Exercise 1.2.9) in the sense that $\mu_F(\mathbb{R} \setminus C) = 0$. The measure μ_F is known as Cantor measure.

(iii) If μ_F is Cantor measure, establish the self-similarity properties $\mu(\frac{1}{3} \cdot E) = \frac{1}{2}\mu(E)$ and $\mu(\frac{1}{3} \cdot E + \frac{2}{3}) = \frac{1}{2}\mu(E)$ for every Borel-measurable $E \subset [0,1]$, where $\frac{1}{3} \cdot E := \left\{\frac{1}{3}x : x \in E\right\}$.

Proof. ($\Rightarrow$). Since F is continuous, $F_+(x) = F_-(x)$ everywhere, and then $\mu_F(\{x\}) = F_+(x) - F_-(x) = 0$ everywhere.
($\Leftarrow$). Suppose that F is not continuous, then $\exists \epsilon > 0$ such that for all $\delta > 0$ when $|x - a| < \delta$, $|F(x) - F(a)| > \epsilon$. Since F is monotone non-decreasing, we have $F_+(a) > F(a) > F_-(a)$ and that $F_+(a) - F(a) > \epsilon$ and that $F(a) - F_-(a) > \epsilon$. Therefore $\mu_F(\{x\}) > 2\epsilon > 0$. Therefore, $\mu_F(\{x\}) = 0 \Rightarrow F\ \forall x$ is continuous.

(ii). The middle third Cantor Set C is constructed as this way:

$$I_n := \bigcup_{a_1,\ldots,a_n \in \{0,2\}} [\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n}]$$

$$C := \bigcap_{n=1}^{\infty} I_n$$

From the construction of Cantor function, we see that on each open interval that disjointly consists of I_n^C, the function value at the start point equals the function value at the end point. Also since F is continuous, we have the Lebesgue-Stieltjes measure of each of them equals zero, and thus by σ-additivity, we have $\mu_F(I_n^C) = 0$ for all n.
Since

$$C^C = (\bigcap_{n=1}^{\infty} I_n)^C = \bigcup_{n=1}^{\infty} I_n^C$$

and $I_1^C \subset I_2^C \subset \ldots \subset (0,1)$, by monotonic convergence theorem we have

$$\mu_F(C^C) = \lim_{n\to\infty} \mu_F(I_n^C) = \lim_{n\to\infty} 0 = 0$$

Also we have $\mu_F([0,1]) = 1$. Therefore it is a probability measure supported on C.

(iii). From the construction of Cantor's function, $F_n \to F$ uniformly, so it should also be that $\mu_{F_n} \to \mu_F$ uniformly. Again, let E be arbitrary intervals in $[0,1]$, and denote the start point and the end point as $s(E)$ and $e(E)$. Clearly, $[s(E), e(E)] \in [0,1]$.
Then, $[\frac{1}{3}s(E), \frac{1}{3}e(E)] \in [0, \frac{1}{3}]$. By the induction rule and that F is continuous,

$$\mu_{F_n}(\frac{1}{3}E) = F_n(\frac{1}{3}e(E)) - F_n(\frac{1}{3}s(E)) = \frac{1}{2}(F_{n-1}(e(E)) - F_{n-1}(s(E)))$$

Taking $n \to \infty$ on both sides, we have

$$\mu_F(\frac{1}{3}E) = \frac{1}{2}(F(e(E)) - F(s(E))) = \frac{1}{2}\mu_F(E)$$

Also, $[\frac{1}{3}s(E) + \frac{2}{3}, \frac{1}{3}e(E) + \frac{2}{3}] \in [\frac{2}{3}, 1]$. By the induction rule and that F is continuous,

$$\mu_{F_n}(\frac{1}{3}E) = F_n(\frac{1}{3}e(E)) - F_n(\frac{1}{3}s(E)) = \frac{1}{2}(F_{n-1}(e(E)) - F_{n-1}(s(E)))$$

Taking $n \to \infty$ on both sides, we have

$$\mu_F(\frac{1}{3}E) = \frac{1}{2}(F(e(E)) - F(s(E))) = \frac{1}{2}\mu_F(E)$$

Also, by our previous exercise, $\mathcal{B}[0,1]$ can be generated by both the open subsets and closed subsets of $[0,1]$. Therefore, for arbitrary $E \in \mathcal{B}[0,1]$, we have that our conclusion holds.

□

1.7.4 Product Measure

Given two measure spaces $(\mathcal{X}, \mathcal{A}, \mu)$ and $(\mathcal{Y}, \mathcal{B}, \nu)$, we want to construct a measure $\mu \times \nu$ (or $\mu \otimes \nu$) on $\mathcal{X} \times \mathcal{Y}$ such that $\mu \otimes \nu(E \times F) = \mu(E)\nu(F)$ $\forall E \in \mathcal{A}$, $F \in \mathcal{B}$.

We have the projection maps

$$\pi_{\mathcal{X}} : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}, \pi_{\mathcal{X}}((x,y)) = x$$

$$\pi_{\mathcal{Y}} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Y}, \pi_{\mathcal{X}}((x,y)) = y$$

Then we form the pullback σ-algebras

$$\pi^*_{\mathcal{X}}(\mathcal{A}) = \{\pi^{-1}_{\mathcal{X}}(E) : E \in \mathcal{A}\}$$

$$\pi^*_{\mathcal{Y}}(\mathcal{B}) = \{\pi^{-1}_{\mathcal{Y}}(F) : F \in \mathcal{B}\}$$

Definition 1.7.6. *(Product σ-algebra). The product σ-algebra is the σ-algebra generated by the union of these two:*

$$\mathcal{A} \times \mathcal{B} = \langle \pi^*_{\mathcal{X}}(\mathcal{A}) \cup \pi^*_{\mathcal{Y}}(\mathcal{B}) \rangle$$

To construct the measure mentioned above, we need the assumption that both spaces are σ-finite.

Definition 1.7.7. *(σ-finite).A measure epace $(\mathcal{X}, \mathcal{A}, \mu)$ is σ-finite if $\mathcal{X} = \bigcup_{n=1}^{\infty} \mathcal{X}_n$ with $\mu(\mathcal{X}_n) < \infty$.*

For example, $(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), m)$ is σ-finite, take $\mathcal{X}_n = B(0,n)$. $(\mathbb{R}^d, 2^{\mathbb{R}^d}, \#)$ is not.
With σ-finite spaces, product measure always exists and is unique.

Proposition 1.7.4. *Let $(\mathcal{X}, \mathcal{A}, \mu)$ and $(\mathcal{Y}, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then $\exists!$ product measure $\mu \otimes \nu$ on $\mathcal{A} \times \mathcal{B}$ such that $\mu \otimes \nu(E \times F) = \mu(E)\nu(F)$.*

Proof. Let $\mathcal{B}_0$ be the collection of all the sets of the form

$$S = (E_1 \times F_1) \cup ... \cup (E_n \times F_n)$$

for $n \in \mathbb{N}$, $E_1, ...E_n \in \mathcal{A}$, $F_1, ..., F_n \in \mathcal{B}$.
Claim 1: $\mathcal{B}_0$ is a Boolean algebra.

Claim 2: $\forall S \in \mathcal{B}_0$, S can be decomposed into a disjoint union of product sets $E_1 \times F_1$, ... $E_n \times F_n$. Then we define the quantity $\mu_0(S)$ by

$$\mu_0(S) = \sum_{i=1}^{n} \mu(E_i)\nu(F_i)$$

These two claims are verified similarly in elementary and Jordan measure chapters
Claim 3: μ_0 is well-defined and μ_0 is a pre-measure. That is, $\forall S \in \mathcal{B}_0$ and $S = \bigcup_{n=1}^{\infty} S_n$ disjoint, we have $\mu_0(S) = \sum_{n=1}^{\infty} \mu_0(S_n)$.
Splitting S up into disjoint product sets, and restricting the S_n to each of these product sets in turn, we may assume WLOG using the finite additivity of μ_0 that $S = E \times F$ for some $E \in \mathcal{A}$ and $F \in \mathcal{B}$. Similarly, by breaking each S_n up into component product sets and using finite additivity again, we may assume WLOG that each S_n takes the form $S_n = E_n \times F_n$ for some $E_n \in \mathcal{A}$, $F_n \in \mathcal{B}$. Then

$$\mu_0(S) = \mu(E)\nu(F)$$

$$\mu_0(S_n) = \mu(E_n)\nu(F_n)$$

Now it suffices to show

$$\mu(E)\nu(F) = \sum_{n=1}^{\infty} \mu(E_n)\nu(F_n)$$

First, by our construction,

$$\mathbf{1}_E(x)\mathbf{1}_F(y) = \sum_{n=1}^{\infty} \mathbf{1}_{E_n}(x)\mathbf{1}_{F_n}(y)$$

Fix $x \in \mathcal{X}$, we have

$$\int_{\mathcal{Y}} \mathbf{1}_E(x)\mathbf{1}_F(y)d\nu = \int_{\mathcal{Y}} \sum_{n=1}^{\infty} \mathbf{1}_{E_n}(x)\mathbf{1}_{F_n}(y)d\nu$$

Calculate the HLS, and use MCT to RHS, we have

$$\mathbf{1}_E(x)\nu(F) = \sum_{n=1}^{\infty} \mathbf{1}_{E_n}(x)\nu(F_n)$$

Using the MCT to integrate both sides with $\mathcal{X}$ we have

$$\mu(E)\nu(F) = \sum_{n=1}^{\infty} \mu(E_n)\nu(F_n)$$

Then, we use the Hahn-Kolmogorov extension theorem and obtain an extension of μ_0 to a complete measure $\mu \otimes \nu$ on a σ-algebra $\mathcal{B}'$ containing $\mathcal{B}_0$. Check that $\langle \mathcal{B}_0 \rangle = \mathcal{A} \times \mathcal{B} \Rightarrow \mathcal{B}' \supseteq \mathcal{A} \times \mathcal{B}$. The restriction of $\mu \otimes \nu$ to $\mathcal{A} \times \mathcal{B}$ gives existence.
To show uniqueness, observe from finite additivity that any measure $\mu \otimes \nu$ on $\mathcal{X} \times \mathcal{Y}$ that obeys $\mu \otimes \nu(E \times F) = \mu(E)\nu(F)$ $\forall E \in \mathcal{A}$, $F \in \mathcal{B}$ must extend μ_0, and so uniqueness follows. $\square$

Notice that m^2 on $\mathcal{L}(\mathbb{R})^2$ does not equal to, but is the completion of $m \otimes m$. $m \otimes m$ is defined on $\mathcal{L}(\mathbb{R}) \times \mathcal{L}(\mathbb{R})$, which is a strict sub-$\sigma$-algebra of $\mathcal{L}(\mathbb{R}^2)$. In fact, $\mathcal{L}(\mathbb{R}^2) = \overline{\mathcal{L}(\mathbb{R}) \times \mathcal{L}(\mathbb{R})}$, $m^2 = \overline{m \otimes m}$.

Before integrating using this product measure, we need the following technical definition and lemma.

Definition 1.7.8. *(Monotone class). A monotone class in $\mathcal{X}$ is a collection $\mathcal{B}$ of subsets of $\mathcal{X}$ that obeys*

1. *If $E_1 \subset E_2 \subset ...$, $E_n \in \mathcal{B}\ \forall n$, then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$.*
2. *If $E_1 \supset E_2 \supset ...$, $E_n \in \mathcal{B}\ \forall n$, then $\bigcap_{n=1}^{\infty} E_n \in \mathcal{B}$.*

Lemma 1.7.2. *(Monotone class lemma) Let $\mathcal{A}$ be a Boolean algebra on $\mathcal{X}$. Then $\langle \mathcal{A} \rangle$ is the smallest monotone class that contains $\mathcal{A}$.*

Proof. Let $\mathcal{B}$ be the intersection of all the monotone classes that contain $\mathcal{A}$. First we verify that $\langle \mathcal{A} \rangle$ is one such class. Since $\langle \mathcal{A} \rangle$ is also a Boolean Algebra, if $E_1, ... \in \langle \mathcal{A} \rangle$, $E_1 \subset E_2 \subset ...$, then by induction, $\bigcup_{n=1}^{\infty} E_n \in \langle \mathcal{A} \rangle$. Similarly for $\supset$. Therefore $\mathcal{B} \subset \langle \mathcal{A} \rangle$. It then suffices to show $\mathcal{B} \supset \langle \mathcal{A} \rangle$.
It is also clear that $\mathcal{B}$ is a monotone class that contains $\mathcal{A}$. By replacing all the elements of $\mathcal{B}$ with their complements, we see that $\mathcal{B}$ is closed under complements.
For any $E \in \mathcal{A}$, let $C_E := \{F \in \mathcal{B} : F \setminus E, E \setminus F, F \cap E, \mathcal{X} \setminus (E \cup F) \in \mathcal{B}\ \forall E \in \mathcal{A}\}$. First we verify that $\mathcal{A} \in C_E$. Indeed, if $G \in \mathcal{A}$, then by Boolean algebra and the definition of $\mathcal{B}$, $G \setminus E \in \mathcal{A} \in \mathcal{B}$, $E \setminus G \in \mathcal{A} \in \mathcal{B}$, $G \cap E \in \mathcal{A} \in \mathcal{B}$, $\mathcal{X} \setminus (G \cup E) \in \mathcal{A} \in \mathcal{B}$, and therefore $G \in C_E$. Since $\mathcal{B}$ is a monotone class, C_E is also a monotone class because all sets in C_E come from $\mathcal{B}$. Now since C_E is a monotone class that contains $\mathcal{A}$, we have $\mathcal{B} \subset C_E$. Also, all the sets in C_E come from $\mathcal{B}$ and we have $\mathcal{B} \subset C_E$. Therefore $\mathcal{B} = C_E$ for all $E \in \mathcal{A}$.
Next, let $D := \{E \in \mathcal{B} : F \setminus E, E \setminus F, F \cap E, \mathcal{X} \setminus (E \cup F) \in \mathcal{B}\ \forall F \in \mathcal{B}\}$. Then, by the previous discussion, $\mathcal{A} \subset D$, D is a monotone class, and $\mathcal{B} = D$. Since $\mathcal{B}$ is closed under complements, $\mathcal{B}$ is closed with respect to finite unions. Since this class also contains $\mathcal{A}$, which contains $\emptyset$, we conclude that $\mathcal{B}$ is a Boolean algebra. Since $\mathcal{B}$ is also closed under increasing countable unions, we conclude that it is closed under arbitrary countable unions, and is thus a σ-algebra. As it contains $\mathcal{A}$, it must also contains $\langle \mathcal{A} \rangle$.

□

Theorem 1.7.4. *(Tonelli's Theorem, incomplete version). Let $(\mathcal{X}, \mathcal{A}, \mu)$, $(\mathcal{Y}, \mathcal{B}, \nu)$ be σ-finite, $f : \mathcal{X} \times \mathcal{Y} \to [0, +\infty]$ is $\mathcal{A} \times \mathcal{B}$-measurable. Then*

1. *$\forall x \in \mathcal{X}$, $y \mapsto f(x, y)$ is $\mathcal{B}$-measurable, and $x \mapsto \int_{\mathcal{Y}} f(x, y) d\nu(y)$ is $\mathcal{A}$-measurable. Similarly with y in place of x.*
2. $$\begin{aligned} \int_{\mathcal{X} \times \mathcal{Y}} f d\mu \otimes \nu &= \int_{\mathcal{X}} (\int_{\mathcal{Y}} f d\nu(y)) d\mu(x) \\ &= \int_{\mathcal{Y}} (\int_{\mathcal{X}} f d\mu(x)) d\nu(y) \end{aligned}$$

Proof. Write $\mathcal{X} = \bigcup_{n=1}^{\infty} \mathcal{X}_n$ as an increasing union with $\mu(\mathcal{X}_n) < \infty\ \forall n$. Then by MCT, it suffices to show the claims with $\mathcal{X}$ replaced by $\mathcal{X}_n$. Thus assume WLOG $\mu(\mathcal{X}) < \infty$. Similarly WLOG assume $\nu(\mathcal{Y}) < \infty$. Therefore $\mu \otimes \nu(\mathcal{X} \times \mathcal{Y}) = \mu(\mathcal{X})\nu(\mathcal{Y}) < \infty$.
Unsigned measurable $f = \lim_{n\to\infty} f_n$ as an increasing limit of unsigned simple f_n. By MCT it suffices to verify the claim when f is a simple function. By linearity, it then suffices to verify the claim that f is an indicator function. $f = \mathbf{1}_S$, $S \in \mathcal{A} \times \mathcal{B}$.
Let $C := \{S \in \mathcal{A} \times \mathcal{B} : \text{the claims holds}\}$. By MCT and downwards MCT, C is a monotone class.
By $\mu \otimes \nu(E \times F) = \mu(E)\nu(F)$, C contains as an element any product $S = E \times F$ where $E \in \mathcal{A}$ and $F \in \mathcal{B}$. By finite addivity, C also contains as an element a disjoint finite union $S = (E_1 \times F_1) \cup ... \cup (E_k \times F_k)$. Then C contains the Boolean algebra $\mathcal{B}_0$ in the proof of the uniqueness and existence of product measure, as such sets can always be xpressed as the disjoint finite union of Cartesian products of measurable sets. Applying the monotone class lemma, C contains $\langle \mathcal{B}_0 \rangle = \mathcal{A} \times \mathcal{B}$. Then the claim follows.

□

Theorem 1.7.5. *Let $(\mathcal{X}, \mathcal{A}, \mu)$, $(\mathcal{Y}, \mathcal{B}, \nu)$ be σ-finite and complete, $f : \mathcal{X} \times \mathcal{Y} \to [0, +\infty]$ is $\overline{\mathcal{A} \times \mathcal{B}}$-measurable. Then*

1. *For μ a.e. $x \in \mathcal{X}$, $y \mapsto f(x, y)$ is $\mathcal{B}$-measurable, and $x \mapsto \int_{\mathcal{Y}} f(x, y) d\nu(y)$ is $\mathcal{A}$-measurable. Similar with y in place of x.*

2. $$\begin{aligned}\int_{\mathcal{X}\times\mathcal{Y}} f d\overline{\mu \otimes \nu} &= \int_{\mathcal{X}} (\int_{\mathcal{Y}} f d\nu(y)) d\mu(x) \\ &= \int_{\mathcal{Y}} (\int_{\mathcal{X}} f d\mu(x)) d\nu(y)\end{aligned}$$

Proof. Every measurable set in $\overline{\mathcal{A} \times \mathcal{B}}$ is equal to a measurable set in $\mathcal{A} \times \mathcal{B}$ outside a $\mu \otimes \nu$-null set. This implies that the $\overline{\mathcal{A} \times \mathcal{B}}$-measurable function f agrees with a $\mathcal{A} \times \mathcal{B}$-measurable function $\tilde{f}$ outside a $\mu \otimes \nu$-null set E (as can be seen by expressing f as the limit of simple functions). For μ-a.e. $x \in \mathcal{X}$, the function $y \mapsto f(x, y)$ agrees with $y \mapsto \tilde{f}(x, y)$ outside of a ν-null set, and is measurable since $(\mathcal{Y}, \mathcal{B}, \nu)$ is complete. Similarly, for ν-a.e. $y \in \mathcal{Y}$, the function $x \mapsto f(x, y)$ agrees with $x \mapsto \tilde{f}(x, y)$ outside of a μ-null set, and is measurable since $(\mathcal{X}, \mathcal{A}, \mu)$ is complete. Thus the claims follows. □

Theorem 1.7.6. *(Fubini's Theorem) Let $(\mathcal{X}, \mathcal{A}, \mu)$, $(\mathcal{Y}, \mathcal{B}, \nu)$ be σ-finite and complete, $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{C}$ is absolutely integrable with respect to $\overline{\mu \otimes \nu}$. Then,*

1. *For μ-a.e. $x \in \mathcal{X}$, $y \mapsto f(x, y) \in L^1(\mathcal{Y}, \mathcal{B}, \nu)$, and $x \mapsto \int_{\mathcal{Y}} f(x, y) d\nu(y) \in L^1(\mathcal{X}, \mathcal{A}, \mu)$. Similarly with y in place of x.*

2. $$\begin{aligned}\int_{\mathcal{X}\times\mathcal{Y}} f d\overline{\mu \otimes \nu} &= \int_{\mathcal{X}} (\int_{\mathcal{Y}} f d\nu(y)) d\mu(x) \\ &= \int_{\mathcal{Y}} (\int_{\mathcal{X}} f d\mu(x)) d\nu(y)\end{aligned}$$

Proof. Breaking f into real and imaginary parts and positive and negative parts, it suffices to consider f unsigned and in L^1.
(2) comes from Tonelli's Theorem part (2).

$$\begin{aligned}\infty &> \int_{\mathcal{X}\times\mathcal{Y}} f d\overline{\mu\otimes\nu} \\ &= \int_{\mathcal{X}} (\int_{\mathcal{Y}} f d\nu(y)) d\mu(x) \\ &= \int_{\mathcal{Y}} (\int_{\mathcal{X}} f d\mu(x)) d\nu(y)\end{aligned}$$

and then we have $\int_{\mathcal{Y}} f d\nu(y) < \infty$ for μ-a.e. x, and $\int_{\mathcal{X}} f d\mu(x) < \infty$ for ν-a.e. y. and then we have (1).

□

1.8 Signed Measures and Differentiation

1.8.1 Signed Measures

Definition 1.8.1. *(Signed measure). Let $(\mathcal{X}, \mathcal{B})$ be a measurable space. A signed measure on $(\mathcal{X}, \mathcal{B})$ is a mapping $\nu : \mathcal{B} \to [-\infty, +\infty]$ such that*

1. *$\nu(\emptyset) = 0$.*
2. *ν takes at most one values of $+\infty$ or $-\infty$.*
3. *$E_j \in \mathcal{B}$ for $j \geq 1$ disjoint, then $\nu(\bigcup_j E_j) = \sum_{j=1}^{\infty} \nu(E_j)$.*

All measures are signed measure.

Definition 1.8.2. *(Extended μ-measurable). If μ is a measure on $\mathcal{B}$, $f : \mathcal{X} \to [-\infty, +\infty]$ is measurable and at least one of $\int f^- d\mu$ or $\int f^+ d\mu$ is finite, then f is extended μ-measurable. $\nu(E) = \int_E f d\mu$ is a signed measure.*

Proposition 1.8.1. *Let ν be signed measure on $(\mathcal{X}, \mathcal{B})$. If $\{E_i\} \in \mathcal{B}$ is an increasing sequence, then $\nu(\bigcup_{i=1}^{\infty} E_i) = \lim_{j\to\infty} \nu(E_j)$. If it is a decreasing sequence, then $\nu(\bigcap_{j=1}^{\infty} E_i) = \lim_{j\to\infty} \nu(E_j)$.*

Proof. The proof is the same as the MCT of unsigned measure. □

Definition 1.8.3. *(Positive, negative, and null). $E \in \mathcal{B}$ is called positive (negative, or null) for ν if $\forall F \in \mathcal{B}$ and $F \subseteq E$, $\nu(F) \geq 0$ (≤ 0, or $= 0$).*

Lemma 1.8.1. *Any measurable subset of a positive set is positive, and the union of any countable family of positive sets is positive.*

Proof. The first statement is directly from definition.
For the second statement, let $Q_n = P_n \setminus \bigcup_{i=1}^{n-1} P_i$, then $Q_n \subseteq P_n$ and thus is positive. For

any $E \subseteq \bigcup_{i=1}^{\infty} P_i$,

$$\begin{aligned}\nu(E) &= \nu(E \cap (\bigcup_{i=1}^{\infty} P_i)) \\ &= \nu(\bigcup_{i=1}^{\infty} E \cap P_i) \\ &= \sum_{i=1}^{\infty} \nu(E \cap P_i) \\ &> 0\end{aligned}$$

since $E \cap P_i$ is a subset of P_i. Therefore the statement holds.

□

Theorem 1.8.1. *(Hahn decomposition theorem) Let ν be a signed measure on $(\mathcal{X}, \mathcal{B})$. Then, there exists a positive set P and a negative set N such that $P \cup N = \mathcal{X}$ and $P \cap N = \emptyset$. If P', N' is another pair, then $P \Delta P' = N \Delta N'$ is null for ν.*

Proof. WLOG, suppose $\nu \neq -\infty$. (Otherwise, consider $-\nu$).
Let $m = \sup_{E \text{ positive}} \nu(E)$. Then, there exists a positive sequence $\{P_j\}$ such that $\nu(P_j) \to m$. Let $P = \bigcup_{j=1}^{\infty} P_j$, then by previous lemma and MCT, P is positive, and $\nu(P) = m$. In particular, $m < \infty$.
We claim that $N = \mathcal{X} \setminus P$ is negative. Suppose, on the contrary, N is non-negative.
First, N cannot contain any non-null positive sets. Indeed, if $E \subset N$ is positive and $\nu(E) > 0$, then $E \cup P$ is positive and $\nu(E \cup P) = \nu(E) + \nu(P) > m$ which is impossible.
Second, if $A \subset N$ and $\nu(A) > 0$, then $\exists B \subset A$ with $\nu(B) > \nu(A)$. Indeed, since A cannot be positive, $\exists C \subset A$ with $\nu(C) < 0$. Thus, if $B = A \setminus C$, then $\nu(B) = \nu(A) - \nu(C) > \nu(A)$.
If N is non-negative, then, we can construct $\{A_j : A_j \subset N\}$ and $\{n_j : n_j \in \mathbb{N}\}$ by the following. n_1 is the smallest integer for which $\exists B \subset N$ with $\nu(B) > n_1^{-1}$, let $A_i = B$. By induction, n_j is the smallest integer for which $\exists B \subset A_{j-1}$ with $\nu(B) > \nu(A_{j-1}) + n_j^{-1}$, let $A_j = B$. Let $A = \bigcap_{j=1}^{\infty} A_j$, then $\infty > \nu(A) = \lim_{j\to\infty} \nu(A_j) > \sum_{j=1}^{\infty} n_j^{-1}$. So, $n_j \to \infty$ as $j \to \infty$. However, $\exists B \subset A$ with $\nu(B) > \nu(A) + n^{-1}$ for some n. For j sufficiently large, we have $n < n_j$, and $B \subset A_{j-1}$, which contradicts with the construction of n_j and A_j.
Therefore, N is negative.
If there is another pair P' and N', we have $P \setminus P' \subset P$, and $P \setminus P' = P \cap P'^C \subset P'^C \subset N'$, so $P \setminus P'$ is both positive and negative, thus null. Same for $P' \setminus P$. Thus $P \Delta P'$ is null. □

This decomposition is usually not unique.

Definition 1.8.4. *(Muturally singular) For another signed measure ν on $(\mathcal{X}, \mathcal{B})$, we say that ν and μ are muturally singular (or ν is singular with μ) if $\exists E, F \in \mathcal{B}$ such that $E \cup F = \mathcal{X}$, $E \cap F = \emptyset$, E is null for μ, F is null for ν. Write it as $\nu \perp \mu$.*

Theorem 1.8.2. *(Jordan decomposition theorem) If ν is a signed measure, then $\exists!$ positive measures ν^+, ν^-, with $\nu = \nu^+ - \nu^-$, $\nu^+ \perp \nu^-$.*

Proof. Let $\mathcal{X} = P \cup N$ be a Hahn decomposition of ν. Define $\nu^+(E) = \nu(E \cap P)$, $\nu^-(E) = -\nu(E \cap N)$. Then $\nu = \nu^+ - \nu_-$, $\nu^+ \perp \nu^-$ (since N is null for ν^+, P is null for ν^-).

If there is another decomposition of $\nu = \mu^+ - \mu^-$, $\mu^+ \perp \mu^-$, let $E, F \in \mathcal{B}$ such that $E \cap F = \emptyset$ $E \cup F = \mathcal{X}$, $\mu^+(F) = \mu^-(E) = 0$. Then, $\mathcal{X} = E \cup F$ is another Hahn decomposition for ν. So $P \Delta E$ is null. Therefore, $\forall A \in \mathcal{B}$, $\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = v^+(A)$. $P \Delta E$ is null. Same for $\nu^-(A) = \nu^-(A \cap F) = -\nu(A \cap F) = -\nu(A \cap N) = \nu^-(A)$. $F \Delta N$ is null. $\square$

Definition 1.8.5. *(Absolutely Continuous) If ν is positive, then we say ν is absolutely continuous with respect to μ if $E \in \mathcal{B}$, $\mu(E) = 0 \Rightarrow \nu(E) = 0$. Write t as $\nu << \mu$.*

For example, $\delta_0 \perp m$. Let $E = \mathbb{R} \setminus \{0\}$, $F = \{0\}$. F is null for m, E is null for δ_0.

For $f \in L^1(\mathbb{R}, \mathcal{L}(\mathbb{R}), m)$, $f : \mathbb{R} \to \mathbb{R}$, define

$$\nu(E) := \int_E f dm$$

Then, $\nu << m$, for if $m(E) = 0$, then $\nu(E) = 0$.

Define

$$L^1(\nu) = L^1(\nu^+) \cap L^1(\nu^-)$$

$$\int_{\mathcal{X}} f d\nu := \int_{\mathcal{X}} f d\nu^+ - \int_{\mathcal{X}} f d\nu^-$$

for $f \in L^1(\nu)$.

For μ positive, we say $f : \mathcal{X} \to [-\infty, +\infty]$ is extended μ-measurable if at least one of the terms above in RHS is finite. ν^+, ν^-, and $|\nu| := \nu^+ + \nu^-$ is called the positive, negative, and total variations of ν.

It is obvious to see that $\nu << \mu$ iff $|\nu| << \mu$.

1.8.2 The Lebesgue-Radon-Nikodym Theorem

Theorem 1.8.3. *Let ν be a finite signed measure, μ be a positive measure on $(\mathcal{X}, \mathcal{B})$. Then, $\nu << \mu$ iff $\forall \epsilon > 0$, $\exists \delta > 0$, such that $\mu(E) < \delta$ implies $|\nu(E)| < \epsilon$.*

Proof. Since $\nu << \mu$ iff $|\nu| << \mu$, $|\nu(E)| \leq |\nu|(E)$, it suffices to assume that $\nu = |\nu|$ is positive.
($\Leftarrow$) It is obvious from the right side that $\mu(E) = 0 \Rightarrow \nu(E) = 0$.
($\Rightarrow$) Suppose, contrapositively, $\exists > 0$, for all $n \in \mathbb{N}$, we can find $E_n \in \mathcal{B}$ with $\mu(E_n) < 2^{-n}$ and $\nu(E_n) \geq \epsilon$. Let $F_k = \bigcup_{n=k}^{\infty} E_n$, $F = \bigcap_{k=1}^{\infty} F_k$, then $\mu(F_k) \leq \sum_{n=k}^{\infty} 2^{-n} = 2^{1-k}$, by MCT, $\mu(F) = \lim_{k \to \infty} \mu(F_k) = 0$. However, since $\nu(F_k) \geq \epsilon$ for all k, since ν is finite, by MCT we have $\nu(F) = \lim_{k \to \infty} \nu(F_k) \geq \epsilon$. Then there is no $\nu << \mu$. $\square$

If μ is a (positive) measure and f is an extended μ-measurable function, the signed measure

$$\nu(E) = \int_E f d\mu$$

is $\nu << \mu$, and is finite iff $f \in L^1(\mu)$. Therefore we immediately have the following corollary.

Corollary 1.8.1. *If $f \in L^1(\mu)$, then $\forall \epsilon > 0,$, $\exists \delta > 0$, such that $|\int_E f d\mu| < \epsilon$ whenever $\mu(E) < \delta$.*

For $\nu(E) = \int_E f d\mu$, we express the relationship as $d\nu = f d\mu$.

Before we prove the main theorem, we prove a lemma.

Lemma 1.8.2. *Let ν and μ be finite measure on $(\mathcal{X}, \mathcal{B})$. Then, either $\nu \perp \mu$, or $\exists \epsilon > 0$, $E \in \mathcal{B}$, such that $\mu(E) > 0$ and $\nu \geq \epsilon\mu$ on E (E is a positive set for $\mu - \epsilon\nu$).*

Proof. Let $\mathcal{X} = P_n \cup N_n$ be a Hahn decomposition for $\nu - n^{-1}\mu$. Let $P = \bigcup_{n=1}^{\infty} P_n$, $N = \bigcap_{n=1}^{\infty} N_n = P^C$. Then, N is a negative set for $\nu - n^{-1}\mu$ for all n. That is, $0 \leq \nu(N) \leq n^{-1}\mu(N)$ for all n (the first inequality holds because ν μ are measures, not signed measures), so $\nu(N) = 0$.
If $\mu(P) = 0$, then $\nu \perp \mu$.
If $\mu(P) > 0$, then $\mu(P_n) > 0$ for some n, and P_n is a positive set for $\nu - n^{-1}\mu$. □

Now we prove the main theorem.

Theorem 1.8.4. *(The Lebesgue-Radon-Nikodym Theorem) Let ν be a σ-finite signed measure, μ be a σ-finite positive measure on $(\mathcal{X}, \mathcal{B})$. Then $\exists!$ σ-finite signed measures λ, ρ on $(\mathcal{X}, \mathcal{B})$ such that $\lambda \perp \mu$, $\rho << \mu$, and $\nu = \lambda + \rho$.*
Moreover, $\exists$ extended μ-integrable function $f : \mathcal{X} \to \mathbb{R}$ such that $d\rho = f d\mu$. Any two such functions are equal μ-a.e..

Proof. (Case 1) Suppose ν and μ are both finite positive measures. Let

$$\mathcal{F} := \{f : \mathcal{X} \to [0, +\infty] : \int_E f d\mu \leq \nu(E) \ \forall E \in \mathcal{B}\}$$

$\mathcal{F}$ is non-empty since $0 \in \mathcal{F}$. Also, if $f, g \in \mathcal{F}$, then $h = \max(f, g) \in \mathcal{F}$, for if $A := \{x : f(x) > g(x)\}$, then $\forall E \in \mathcal{B}$, we have $\int_E h d\mu = \int_{E \cap A} f d\mu + \int_{E \setminus A} g d\mu \leq \nu(E \cap A) + \nu(E \setminus A) = \nu(E)$.
Let $a := \sup\{\int f d\mu : f \in \mathcal{F}\}$, then $a \leq \nu(\mathcal{X}) < \infty$. Choose a sequence $\{f_n\} \subset \mathcal{F}$ such that $\int f_n d\mu \to a$. Let $g_n := \max(f_1, ..., f_n)$, $f = \sup_n f_n$. Then, by the previous argument, $g_n \in \mathcal{F}$. Also $g_n \uparrow f$ pointwise, and $\int g_n d\mu \geq \int f d\mu$. It follows that

$$\lim_{n\to\infty} \int g_n d\mu = \lim_{n\to\infty} \int f_n d\mu = a$$

Also by MCT,

$$\lim_{n\to\infty} \int g_n d\mu = \int \lim_{n\to\infty} g_n d\mu = \int f d\mu = a$$

and thus $f \in \mathcal{F}$, In particular $f < \infty$ a.e.
Claim: $d\lambda = d\nu - f d\mu$, as a positive measure, is singular with respect to μ. Indeed, $\forall E \in \mathcal{B}$,

$$\begin{aligned}\int_{\mathcal{X}} \mathbf{1}_E d\lambda &= \int_{\mathcal{X}} \mathbf{1}_E d\nu - \int_{\mathcal{X}} \mathbf{1}_E f d\mu \\ &= \nu(E) - \int_{\mathcal{X}} f d\mu \\ &\geq 0\end{aligned}$$

If not singular, then by the previous Lemma, $\exists E \in \mathcal{B}$ and $\epsilon > 0$ such that $\mu(E) > 0$ and $\lambda \geq \epsilon\mu$ on E. But then we will have $\epsilon \mathbf{1}_E d\mu \leq d\lambda = d\nu - f d\mu$, that is, $(f + \epsilon\nu(E))d\mu \leq d\nu$. So $f + \epsilon \mathbf{1}_E \in \mathcal{F}$, and

$$\int_{\mathcal{X}} (f + \epsilon \mathbf{1}_E) d\mu = a + \epsilon\nu(E) > a$$

this contradict with the definition of a.
Therefore, we have proved the existence of λ and ρ where $d\rho = f d\mu$, $d\lambda = d\nu - f d\mu$, $\rho << \mu$, $\lambda \perp \mu$.
As for uniqueness, if also $d\nu = d\lambda' + f' d\mu$, then $d\lambda - d\lambda' = (f' - f)d\mu$. But $\lambda - \lambda' \perp \mu$, while $(f - f')d\mu << d\mu$. Hence $d\lambda - d\lambda' = (f - f')d\mu = 0$, and therefore $\lambda = \lambda'$, $f = f'$ μ-a.e..
(Case 2) Suppose μ and μ are σ-finite measures. Then, $\mathcal{X}$ can be written as a countable disjoint union of μ-finite sets and a countable disjoint union of ν-finite sets. Taking intersections, we obtain $\{A_j\}_{j=1}^{\infty} \subset \mathcal{B}$ such that $\mu(A_j)$ and $\nu(A_j)$ are finite for all j and $\mathcal{X} = \bigcup_{j=1}^{\infty} A_j$.
Define $\mu_j(E) = \mu(E \cap A_j)$, $\nu_j(E) = \nu(E \cap A_j)$. Since they are finite measure, by the reasoning above, we have $\forall j$, $\exists d\nu_j = d\lambda_j + f_j d\mu_j$, where $\lambda_j \perp \mu_j$. Since $\mu_j(A_j^C) = \nu_j(A_j^C) = 0$, we have $\lambda_j(A_j^C) = \nu_j(A_j^C) - \int_{A_j^C} f d\mu_j = 0$ (this is not hard to show). And we may assume that $f_j = 0$ on A_j^C. Let $\lambda = \sum_{j=1}^{\infty} \lambda_j$, $f = \sum_{j=1}^{\infty} f_j$. Then $d\nu = d\lambda + f d\mu$, and $\lambda \perp \mu$, and $d\lambda$ and $f d\mu$ are σ-finite measure. Uniqueness follows as before.
(Case 3) If μ is a signed measure, apply the preceding argument to μ^+ and μ^- and subtract the results.

□

We call $d\nu = d\lambda + f d\mu$ the Lebesgue Decomposition of ν with respect to μ.

In particular, if $\nu << \mu$, then the theorem above says that $\exists f$ such that $d\nu = f d\mu$, f called the Radon-Nikodym derivative of ν with respect to μ. Denote it as $f = \frac{d\nu}{d\mu}$, thus

$$d\nu = \frac{d\nu}{d\mu} d\mu$$

Proposition 1.8.2. *Let ν be a σ-finite signed measure. μ, λ be σ-finite measures on $(\mathcal{X}, \mathcal{B})$ such that $\nu << \mu$ and $\mu << \lambda$.*

1. *if $g \in L^1(\mu)$, then $g\frac{d\nu}{d\mu} \in L^1(mu)$ and*

$$\int g d\nu = \int g \frac{d\nu}{d\mu} d\mu$$

2. *we have $\nu << \lambda$ and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$ λ-a.e..*

Proof. Considering ν^+ and ν^- separately, we may assume $\nu \geq 0$.
$\int g d\nu = \int g \frac{d\nu}{d\mu} d\mu$ is true if $g = \mathbf{1}_E$ by Lebesgue Radon Nikodym theorem. By linearity, it is true for simple functions. By MCT, it is true for nonnegative measurable functions. By linearity again, it is true for functions in $L^1(\nu)$.

Replacing ν, μ with μ, λ, let $g = \mathbf{1}_E \frac{d\nu}{d\mu}$. Then for all $E \in \mathcal{B}$, we have

$$\nu(E) = \int_E \frac{d\nu}{d\mu} d\mu = \int_E \frac{d\nu}{d\mu} \frac{d\mu}{d\lambda} d\lambda$$

Then $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \frac{d\mu}{d\lambda}$ λ-a.e.. $\square$

Corollary 1.8.2. *If $\mu << \lambda$ and $\lambda << \nu$, then $\frac{d\lambda}{d\mu} \frac{d\mu}{d\lambda} = 1$ a.e. w.r.t. both λ and μ.*

Proof. By the conditions we know that $\mu << \mu$. Then, from the previous theorem we have μ-a.e..

$$\frac{d\mu}{d\mu} = \frac{d\mu}{d\lambda} \frac{d\lambda}{d\mu}$$

Same with $\lambda << \lambda$. $\square$

Proposition 1.8.3. *If $\mu_1, ..., \mu_n$ are measures on $(\mathcal{X}, \mathcal{B})$, then $\exists \mu = \sum_{j=1}^n \mu_j$ such that $\mu_j << \mu\ \forall j$.*

Proof. trivial. $\square$

For example,

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$d\gamma = f dm$$

is the Gaussian measure. $\gamma << m << \gamma$,

$$\frac{d\gamma}{dm}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

For $\mathcal{X}$ countable and ν has the form $\nu(E) = \sum_{x \in E} f(x)$, then $\frac{d\nu}{d\#} = f$, $\#$ is the counting measure on $\mathcal{X}$. When ν is a probability measure, $\frac{d\nu}{d\#}$ is called the probability mass function for ν.

2 L^p Spaces

2.1 L^p Spaces and Banach Space

Definition 2.1.1. *On $(\mathcal{X}, \mathcal{B}, \mu)$, for $0 < p < \infty$, let $L^p = L^p(\mathcal{X}, \mathcal{B}, \mu)$ be the set of all measurable $f : \mathcal{X} \to \mathbb{C}$ such that*

$$\int_{\mathcal{X}} |f|^p d\mu < \infty$$

Define

$$||f||_{L^p} = (\int_{\mathcal{X}} |f|^p d\mu)^{\frac{1}{p}}$$

L^p is a vector space, for $f, g \in L^p$,

$$|f + g|^p \leq (2\max(|f|, |g|))^p \leq 2^p(|f|^p + |g|^p)$$

Integrate both sides we get $f + g \in L^p$.

As with L^1, two functions define the same element in L^p when they are equal a.e..

For the case $(\mathcal{X}, \mathcal{B}, \mu) = (\mathcal{X}, 2^{\mathcal{X}}, \#)$, we write $l^p(\mathcal{X}) = L^p(\mathcal{X}, \mathcal{B}, \mu)$. The most common examples are $\mathcal{X} = \mathbb{N}$ or $\mathbb{Z}$ or $[d] = \{1, ..., d\}$. In these cases,

$$l^p(\mathcal{X}) = \{(v_i)_{i\in\mathcal{X}} : \sum_{i\in\mathcal{X}} |v_i|^p < \infty\}$$

On $\mathbb{R}^d$, we have the norm $||v||_p = (\sum_{i\in\mathcal{X}} |v_i|^p)^{\frac{1}{p}}$.

Here we turn to the question of whether $||.||_{L^p}$ is a norm.

Definition 2.1.2. *(Semi norm). A semi norm on a vector space V over $K = \mathbb{C}$ or $\mathbb{R}$ is a function $v \to ||v||$ from $V \to [0, +\infty]$ such that*

1. *$||v + w|| \leq ||v|| + ||w||$.*
2. *$||cv|| = |c|||v||$ $\forall v \in V, c \in K$.*

Definition 2.1.3. *(Norm). A semi norm is a norm if: $||v|| = 0$ iff $v = 0$ everywhere.*

We already know that $||f||_{L^p} = 0$ iff $f = 0$ a.e.. (2) is also easy. Now we establish ths triangle inequality. It fails when $p < 1$. Indeed, notice that when $p \in (0, 1), a, b > 0$, we have $a^p + b^p > (a + b)^p$. WLOG, take $a = 1$ by dividing both sides by a^p, it suffices to prove that $1 + b^p > (1 + b)^p$ $\forall b > 0$. Then, consider $f = \mathbf{1}_E$ and $g = \mathbf{1}_F$ where E, F have positive measure and are disjoint. Then,

$$\begin{aligned}
||f||_{L^p} + ||g||_{L^p} &= (\int_{\mathcal{X}} \mathbf{1}_E^p d\mu)^{\frac{1}{p}} + (\int_{\mathcal{X}} \mathbf{1}_F^p d\mu)^{\frac{1}{p}} \\
&= ((\mu(E)^{\frac{1}{p}} + \mu(F)^{\frac{1}{p}})^p)^{\frac{1}{p}} \\
&< (\mu(E) + \mu(F))^{\frac{1}{p}} \\
&= (\int_{\mathcal{X}} |f|^p d\mu + \int_{\mathcal{X}} |g|^p d\mu)^{\frac{1}{p}} \\
&\leq (\int_{\mathcal{X}} |f + g|^p d\mu)^{\frac{1}{p}} \\
&= ||f + g||_{L^p}
\end{aligned}$$

Now we derive the Hölder's inequality.

Lemma 2.1.1. *For $a, b \geq 0$, $\lambda \in (0,1)$, we have*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b$$

with equality iff $a = b$.

Proof. This follows from the strict concavity of log on $(0, +\infty)$.

$$\log(\lambda a + (1-\lambda)b) > \lambda \log(a) + (1-\lambda)\log(b) = \log(a^\lambda b^{1-\lambda})$$

$\square$

Proposition 2.1.1. *(Hölder's Inequality). Let $p \in (1, +\infty)$, $q = \frac{p}{p-1}$ (so $\frac{1}{p} + \frac{1}{q} = 1$, q is the conjugate component of p). Then for measurable $f, g : \mathcal{X} \to \mathbb{C}$,*

$$||fg||_{L^1} \leq ||f||_{L^p} ||g||_{L^q}$$

with equality iff $\alpha|f|^p = \beta|g|^q$ a.e. for some constants α and β not both 0.

Proof. If either $f = 0$ a.e. or $g = 0$ a.e., then both sides equals zero. It is also immediate if $||f||_{L^p} = +\infty$ or $||g||_{L^p q} = +\infty$.
Assume $0 < ||f||_{L^p}, ||g||_{L^q} < \infty$. By scaling f, g by $||f||_{L^p}, ||g||_{L^q}$, the inequality doesn't change, and we may assume $||f||_{L^p} = ||g||_{L^q} = 1$. Using the above lemma, let $a = |f|^p, b = |g|^q, \lambda = \frac{1}{p}, \frac{1}{p} + \frac{1}{q} = 1$. We have

$$(|f|^p)^{\frac{1}{p}} (|g|^q)^{\frac{1}{q}} \leq \frac{1}{p}|f|^p + \frac{1}{q}|g|^q$$

Integrate both sides,

$$||fg||_{L^1} = \int_{\mathcal{X}} |fg| d\mu \leq \frac{1}{p} + \frac{1}{q} = 1 = ||f||_{L^p} ||g||_{L^q}$$

The equality holds when $|f|^p = |g|^q$ a.e..

$\square$

Remark: $p = q = 2$ is a special case of Cauchy-Schwartz inequality. The q in the condition is called the conjugate component of p.

Proposition 2.1.2. *(Minkowski's Inequality). For $p \in [1, +\infty)$, $f, g \in L^p$, we have*

$$||f + g||_{L^p} \leq ||f||_{L^p} + ||g||_{L^p}$$

Proof. We have proved when $p = 1$. It is also immediate if $f + g = 0$ a.e.. Therefore, assume otherwise. First, by triangle inequality of $|.|$, we have pointwise

$$|f + g|^p \leq (|f| + |g|)|f + g|^{p-1}$$

Set $q = \frac{p}{p-1}$ and integrate both sides

$$\begin{aligned}\int_{\mathcal{X}} |f+g|^p d\mu &\leq \int_{\mathcal{X}} |f||f+g|^{p-1} d\mu + \int_{\mathcal{X}} |g||f+g|^{p-1} d\mu \\ &\leq ||f||_{L^p} |||f+g|^{p-1}||_{L^q} + ||g||_{L^p} |||f+g|^{p-1}||_{L^q} \\ &\leq (||f||_{L^p} + ||g||_{L^q})(\int_{\mathcal{X}} |f+g|^p)^{\frac{1}{q}}\end{aligned}$$

Then since $|f+g| \neq 0$, we have

$$||f+g||_{L^p} \leq ||f||_{L^p} + ||g||_{L^p}$$

□

Therefore, for $p \geq 1$, L^p is a normed vector space.

Recall that a sequence (v_n) in a metric space (V, ρ) is Cauchy if $\rho(v_n, v_m) \to 0$ as $n, m \to \infty$. (V, ρ) is complete if any Cauchy sequence has limit in V. A normed space $(V, ||.||)$ (with metric $\rho(x,y) = ||x-y||$) is a Banach space if it is complete.

Lemma 2.1.2. *A normed vector space $(V, ||.||)$ is complete iff every absolutely convergent series in V converges, i.e. if $\sum_{n=1}^{\infty} ||v_n|| < \infty$ then $\exists v \in V$ such that $\sum_{n=1}^{N} v_n \to v$ as $N \to \infty$.*

Proof. ($\Rightarrow$). Suppose $(V, ||.||)$ is complete. Let $(v_n) \in V$ such that $\sum_{n=1}^{\infty} ||v_n|| < \infty$, with $S_N = \sum_{n=1}^{N} v_n$. Then $\forall N > M$, $||S_N - S_M|| \leq \sum_{n=M+1}^{N} ||v_n|| \to 0$ as $N, M \to \infty$. Therefore $(S_N)_{N \geq 1}$ is Cauchy, hence converges to some $v \in V$.
($\Leftarrow$). Suppose that every absolutely convergent series in V converges. Let (v_n) be Cauchy in V. Then,choose $n_1 < n_2 < n_3 < ...$ so that $||v_n - v_m|| < 2^{-j}$ when $n, m \geq n_j$.
Set $w_j = v_{n_j} - v_{n_{j-1}}$, $w_0 = 0$. Then, $v_{n_k} = \sum_{j=1}^{k} w_j$ and

$$\sum_{j=1}^{\infty} ||w_j|| \leq ||w_1|| + \sum_{j=1}^{\infty} 2^{-j} < \infty$$

Then the sequence $(v_{n_k})_{k \geq 1}$ has a limit $v \in V$.
We claim that $v_n \to v$. Indeed, let $\epsilon > 0$, $\exists N$ such that, when $n, m \geq N$, $||v_n - v_m|| < \epsilon$, and when $n_k \geq N$, $||v_{n_k} - v|| < \epsilon$. So, for any $n \geq N$, take k so that $n_k \geq N$,

$$||v_n - v|| \leq ||v_n - v_{n_k}|| + ||v_{v_k} - v|| < 2\epsilon$$

Therefore $v_n \to v$. □

Theorem 2.1.1. *$L^p(\mathcal{X}, \mathcal{B}, \mu)$ is a Banach space for $p \geq 1$.*

Proof. It only remains completeness to be verified.
Let $(f_k)_{k \geq 1}$ be arbitrary sequence in L^p such that $\sum_{k=1}^{\infty} ||f_k||_{L^p} < \infty$. (denote $B := \sum_{k=1}^{\infty} ||f_k||_{L^p}$). Want to show $\sum_{k=1}^{n} f_k$ converges to element of L^p.
Set $G_n := \sum_{k=1}^{n} |f_k|$, $G := \sum_{k=1}^{\infty} |f_k|$. $G_n, G : \mathcal{X} \to [0, +\infty]$. By Minkowski's Inequality,

$$||G_n||_{L^p} \leq \sum_{k=1}^{n} ||f_k||_{L^p} \leq B$$

By monotone convergence theorem,

$$\int_{\mathcal{X}} |G|^p d\mu = \lim_{n\to\infty} \int_{\mathcal{X}} |G_n|^p d\mu \le B^p$$

Therefore $G \in L^p$, $G < \infty$ a.e.. $F := \sum_{k=1}^{\infty} f_k$ converges a.e..
Since $|F| \le G$ pointwise, we have $F \in L^p$. Also

$$|F - \sum_{k=1}^{n} f_k|^p \le (|F| + \sum_{k=1}^{n} |f_k|)^p \le (2G)^p$$

So $F - \sum_{k=1}^{n} f_k \in L^p$. By dominated convergence theorem,

$$\lim_{n\to\infty} ||F - \sum_{k=1}^{n} f_k||_{L^p}^p = \lim_{n\to\infty} \int_{\mathcal{X}} |F - \sum_{k=1}^{n} f_k|^p d\mu = \int_{\mathcal{X}} \lim_{n\to\infty} |F - \sum_{k=1}^{n} f_k|^p d\mu = \int_{\mathcal{X}} 0 d\mu = 0$$

Therefore $\sum_{k=1}^{n} f_k \to F \in L^p$. □

We also define L^∞ spaces.

Definition 2.1.4. *(Essential Supremum). As in the Lebesgue setting,*

$$||f||_{L^\infty(\mathcal{X},\mathcal{B},\mu)} := \inf\{a \ge 0 : \mu(\{x : |f(x)| > a\}) = 0\}$$

Definition 2.1.5. *($L(\mathcal{X},\mathcal{Y})$). If $\mathcal{X}$ and $\mathcal{Y}$ are normed vector spaces, denoted the space of all bounded linear map $T : \mathcal{X} \to \mathcal{Y}$ by $L(\mathcal{X},\mathcal{Y})$.*

Definition 2.1.6. *(L^∞ space).*

$$L^\infty := \{f : \mathcal{X} \to \mathbb{C} \textit{ measurable}, ||f||_\infty < \infty\}$$

$||.||_\infty$ is a norm, the Hölder's Inequality holds $||fg||_{L^1} = ||f||_{L^1}||g||_{L^\infty}$ since $\frac{1}{1} + \frac{1}{\infty} = 1$, and L^∞ is a Banach space.

Folland 6.1.3 If $1 \le p < r \le \infty$, $L^p \cap L^r$ is a Banach space with norm

$$\|f\| = \|f\|_p + \|f\|_r,$$

and if $p < q < r$, the inclusion map $L^p \cap L^r \to L^q$ is continuous.

Proof. It suffices to show completeness. Let $(f_k)_{k\ge1}$ be such that

$$\sum_{k=1}^{\infty} ||f|| = \sum_{k=1}^{\infty} ||f||_{L^p} + \sum_{k=1}^{\infty} ||f||_{L^r} = B \le \infty$$

We want to show that $\sum_{k=1}^{n} f_k$ converges to something that is in $L^p \cap L^r$.

Let $G_n = \sum_{k=1}^n |f_k|$, $G = \sum_{k=1}^\infty |f_k|$. Then,

$$\begin{aligned}
||G_n|| &= ||\sum_{k=1}^n |f_k|||_{L^p} + ||\sum_{k=1}^n |f_k|||_{L^r} \\
&\leq \sum_{k=1}^n ||f_k||_{L^p} + \sum_{k=1}^n ||f_k||_{L^r} \text{ by Minkowski} \\
&= B
\end{aligned}$$

Therefore by monotone convergence theorem,

$$\int_{\mathcal{X}} |G|^p d\mu = \lim_{n\to\infty} \int_{\mathcal{X}} |G_n|^p d\mu \leq B^p$$

$$\int_{\mathcal{X}} |G|^r d\mu = \lim_{n\to\infty} \int_{\mathcal{X}} |G_n|^r d\mu \leq B^r$$

Therefore $G \in L^p \cap L^r$, $G < \infty$ a.e., and then $F = \sum_{k=1}^\infty f_k$ converges a.e..
Also since $|F| < G$ pointwise, we have $F \in L^p \cap L^r$.

$$|F - \sum_{k=1}^n f_k|^p \leq (|F| + \sum_{k=1}^n |f_k|)^p \leq (2G)^p$$

$$|F - \sum_{k=1}^n f_k|^r \leq (|F| + \sum_{k=1}^n |f_k|)^r \leq (2G)^r$$

Put integral on both sides and we find that $F - \sum_{k=1}^n f_k \in L^p \cap L^r$. Therefore, by dominated convergence theorem,

$$\begin{aligned}
\lim_{n\to\infty} ||F - \sum_{k=1}^n f_k|| &= \lim_{n\to\infty} ||F - \sum_{k=1}^n f_k||_{L^p} + \lim_{n\to\infty} ||F - \sum_{k=1}^n f_k||_{L^p} \\
&= \lim_{n\to\infty} (\int_{\mathcal{X}} |F - \sum_{k=1}^n f_k|^p d\mu)^{\frac{1}{p}} + \lim_{n\to\infty} (\int_{\mathcal{X}} |F - \sum_{k=1}^n f_k|^r d\mu)^{\frac{1}{r}} \\
&= (\lim_{n\to\infty} \int_{\mathcal{X}} |F - \sum_{k=1}^n f_k|^p d\mu)^{\frac{1}{p}} + (\lim_{n\to\infty} \int_{\mathcal{X}} |F - \sum_{k=1}^n f_k|^r d\mu)^{\frac{1}{r}} \\
&= (\int_{\mathcal{X}} \lim_{n\to\infty} |F - \sum_{k=1}^n f_k|^p d\mu)^{\frac{1}{p}} + (\int_{\mathcal{X}} \lim_{n\to\infty} |F - \sum_{k=1}^n f_k|^r d\mu)^{\frac{1}{r}} \\
&= 0
\end{aligned}$$

Therefore, $\sum_{k=1}^n \to F \in L^p \cap L^r$.
An inclusion map is $T : L^p \cap L^r \to L^q$: $T(f) = f$. It is enough to show that $\exists C$, such that $\forall f \in L^p \cap L^q$,

$$||f||_{L^q} \leq C(||f||_{L^p} + ||f||_{L^r})$$

Indeed, by Proposition 6.10., for $\lambda \in (0,1)$ and $\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r}$,

$$\begin{aligned}||f||_{L^q} &\leq ||f||_{L^p}^{\lambda}||f||_{L^r}^{1-\lambda}\\ &\leq \lambda||f||_{L^p} + (1-\lambda)||f||_{L^r}\\ &\leq ||f||_{L^p} + ||f||_{L^r}\end{aligned}$$

Then we have continuity. $\square$

Folland 6.1.4 If $1 \leq p < r \leq \infty$, $L^p + L^r$ is a Banach space with norm

$$\|f\| = \inf\{\|g\|_p + \|h\|_r : f = g + h\},$$

and if $p < q < r$, the inclusion map $L^q \to L^p + L^r$ is continuous.

Proof. Let (f_n) be Cauchy in $L^p + L^r$. Then $f_n = g_n + h_n$, $f_n - f_m = (g_n - g_m) + (h_n - h_m) \in L^p + L^r$. By the definition of infimum in the norm, for each $n.m \in \mathbb{N}$, we can select $g_n \in L^p$ and $h_n \in L^r$, such that

$$||g_n - g_m||_{L^p} + ||h_n - h_m||_{L^r} \leq ||f_n - f_m|| + \frac{1}{nm}$$

Also, by Triangle inequality applied to $f_n - f_m = (g_n - g_m) + (h_n - h_m)$, we have

$$||f_n - f_m|| \leq ||g_n - g_m||_{L^p} + ||h_n - h_m||_{L^r}$$

Since (f_n) is Cauchy, $||f_n - f_m|| \to 0$ as $n, m \to \infty$. Therefore as $n, m \to \infty$,

$$||g_n - g_m||_{L^p} + ||h_n - h_m||_{L^r} \to 0$$

Since both terms are non-negative, it can only be that both converges to zero, and both (g_n) and (h_n) we selected are Cauchy. Then, by completeness of L^p and L^r, $g_n \to g \in L^p$ in $||.||_{L^p}$ and $h_n \to h \in l^r$ in $||.||_{L^r}$. Let $f = g + h \in L^p + L^r$, by Triangle inequality again,

$$||f_n - f|| \leq ||g_n - g||_{L^p} + ||h_n - h||_{L^r} \to 0$$

as $n \to \infty$. Therefore, $f_n \to f \in L^p + L^r$ in $||.||$, and $L^p + L^r$ is Cauchy.
To prove continuity, it suffices to show that $\forall f \in L^q$, $\exists C$ such that

$$||f|| \leq C||f||_{L^q}$$

Let $f \in L^q$. For $\lambda > 0$, define $g = f\mathbf{1}_{|f|>\lambda}$ and $h = f\mathbf{1}_{|f|\leq\lambda}$.

Since $|f| > \lambda$ on the support of g, $p - q < 0$,

$$\begin{aligned}
||g||_{L^p}^p &= \int_{|f|>\lambda} |f|^p d\mu \\
&= \int_{|f|>\lambda} |f|^{p-q}|f|^q d\mu \\
&< \lambda^{p-q} \int_{|f|>\lambda} |f|^q d\mu \\
&= \lambda^{p-q}||f||_{L^q}^q \\
&< \infty
\end{aligned}$$

Therefore $g \in L^p$ and $||g||_{L^p} < \lambda^{\frac{p-q}{p}}||f||_{L^q}^{\frac{q}{p}}$.
Since $|f| \leq \lambda$ on the support of h, $r - q > 0$,

$$\begin{aligned}
||h||_{L^r}^r &= \int_{|f|>\lambda} |f|^r d\mu \\
&= \int_{|f|>\lambda} |f|^{r-q}|f|^r d\mu \\
&\leq \lambda^{r-q} \int_{|f|>\lambda} |f|^r d\mu \\
&= \lambda^{r-q}||f||_{L^r}^r \\
&< \infty
\end{aligned}$$

Therefore $h \in L^r$ and $||g||_{L^r} < \lambda^{\frac{r-q}{r}}||f||_{L^q}^{\frac{q}{r}}$.
Choose $\lambda = ||f||_{L^q}$, then $||g||_{L^p} \leq ||f||_{L^q}$, $||h||_{L^r} \leq ||f||_{L^q}$, and then we have $||g||_{L^p} + ||h||_{L^r} \leq 2||f||_{L^q}$,

$$||f|| = \inf\{||g||_{L^p} + ||h||_{L^r} : f = g + h\} \leq 2||f||_{L^q}$$

and the inclusion map is continuous.

$\square$

Folland 6.1.5 Suppose $0 < p < q < \infty$. Then $L^p \not\subset L^q$ iff X contains sets of arbitrarily small positive measure, and $L^q \not\subset L^p$ iff X contains sets of arbitrarily large finite measure.

(For the "if" implication: In the first case there is a disjoint sequence $\{E_n\}$ with $0 < \mu(E_n) < 2^{-n}$, and in the second case there is a disjoint sequence $\{E_n\}$ with $1 \leq \mu(E_n) < \infty$. Consider $f = \sum a_n \chi_{E_n}$ for suitable constants a_n.) What about the case $q = \infty$?

Proof. First, we prove Chebyshev's Inequality. For $E_n := \{x : |f(x)| \geq n\}$, we have

$$\begin{aligned}
\mu(E_n) &= \frac{1}{n^p} \int n^p \mathbf{1}_{E_n} d\mu \\
&\leq \frac{1}{n^p} \int |f|^p \mathbf{1}_{E_n} d\mu \\
&\leq \frac{1}{n^p} \int |f|^p d\mu \\
&= \frac{||f||_p^p}{n^p}
\end{aligned}$$

($\Rightarrow$). Let $f \in L^p$. Suppose, countrapositively, that $\exists \epsilon > 0$, such that $\forall E \in \mathcal{B}(\mathcal{X})$, $\mu(E) \neq (0, \epsilon)$. Since

$$\mu(E_n) \leq \frac{||f||_p^p}{n^p} \to 0$$

as $n \to \infty$, $\exists N$ such that $\forall n \geq N$, $\mu(E_n) = 0$.
Therefore, $|f| < N$ a.e.., and

$$\begin{aligned}
||f||_q^q &= \int |f|^q d\mu \\
&= \int |f|^{q-p} |f|^p d\mu \\
&< N^{q-p} \int |f|^p d\mu \\
&\leq \infty
\end{aligned}$$

Then $f \in L^q$. Therefore, $L^p \not\subset L^q \Rightarrow \mathcal{X}$ contains sets of arbitrarily small positive measure.
($\Leftarrow$). Suppose $\forall \epsilon > 0$, $\exists E \in \mathcal{B}(\mathcal{X})$ such that $\mu(E) \in (0, \epsilon)$. Let $\{E_n\}$ be such that $0 < \mu(E_n) < 2^{-n}$. By constructing $G_n := E_n \setminus \bigcup_{i=n+1}^{\infty} E_i$, we have G_n disjoint and by monotonicity

$$0 < \mu(G_n) < \mu(E_n) < 2^{-n}$$

Define $f := \sum_{n=1}^{\infty} \mu(G_n)^{-\frac{1}{q}} \mathbf{1}_{G_n} \geq 0$.

$$\begin{aligned}
||f||_p^p &= \int |f|^p d\mu \\
&= \int (\sum_{n=1}^{\infty} \mu(G_n)^{-\frac{1}{q}} \mathbf{1}_{G_n})^p d\mu \\
&= \lim_{m\to\infty} \int (\sum_{n=1}^{m} \mu(G_n)^{-\frac{1}{q}} \mathbf{1}_{G_n})^p d\mu \\
&= \lim_{m\to\infty} \int \sum_{n=1}^{m} \mu(G_n)^{-\frac{p}{q}} \mathbf{1}_{G_n} d\mu \text{ only these terms have contribution} \\
&= \lim_{m\to\infty} \sum_{n=1}^{m} \mu(G_n)^{1-\frac{p}{q}} \\
&\leq \lim_{m\to\infty} \sum_{n=1}^{m} \mu(G_n) \\
&< \infty
\end{aligned}$$

Therefore $f \in L^p$.

$$\begin{aligned}
||f||_q^q &= \int |f|^q d\mu \\
&= \int (\sum_{n=1}^{\infty} \mu(G_n)^{-\frac{1}{q}} \mathbf{1}_{G_n})^q d\mu \\
&= \lim_{m\to\infty} \int (\sum_{n=1}^{m} \mu(G_n)^{-\frac{1}{q}} \mathbf{1}_{G_n})^q d\mu \\
&= \lim_{m\to\infty} \int \sum_{n=1}^{m} \mu(G_n)^{-1} \mathbf{1}_{G_n} d\mu \text{ only these terms have contribution} \\
&= \lim_{m\to\infty} m \\
&= \infty
\end{aligned}$$

Therefore $f \notin L^q$. Therefore $\mathcal{X}$ contains arbitrarily small measure $\Rightarrow L^p \not\subset L^q$.

($\Rightarrow$). Let $f \in L^q$. Suppose, contrapositively, $\mathcal{X}$ does not contain sets or arbitrarily large measure. That is, $\exists M > 0$, such that $\forall E \in \mathcal{B}(\mathcal{X})$, we have $\mu(E) \leq M$.

Let $E_n := \{x : |f| \geq n\}$, and we have $\mu(E_n) \leq \frac{||f||_q^q}{n^q} \to 0$ as $n \to \infty$. Then $\exists N$ such that when $n \geq N$, $\mu(E_n) \leq \epsilon$. In E_n^C $\forall n \geq N$, we have $|f| < N$.

Since $f \in L^q$, we have $|f| < \infty$ a.e., and $\exists M'$ such that $|f| < M'$ a.e..

$$\begin{aligned}
||f||_p^p &= \int |f|^p d\mu \\
&= \int_{E_n} |f|^p d\mu + \int_{E_n^C} |f|^p d\mu \\
&\leq \int_{E_n} |f|^{q-p} |f|^p d\mu + N^p \mu(E_n^C) \\
&\leq M'^{q-p} \int_{E_n} |f|^q d\mu + MN^p \\
&< \infty
\end{aligned}$$

Therefore $f \in L^p$. Therefore $L^q \not\subset L^p \Rightarrow \mathcal{X}$ contains sets of arbitrarily large measure.

($\Leftarrow$). Suppose $\mathcal{X}$ contains sets of arbitrarily large measure. We can find a group of disjoint sets $\{G_n\}$ with $2^n < \mu(G_n) < \infty$ for each n. Set $f := \sum_{n=1}^{\infty} (\mu(G_n))^{-\frac{1}{p}} \mathbf{1}_{G_n}$.

$$\begin{aligned}
\int |f|^q d\mu &= \int (\sum_{n=1}^{\infty} (\mu(G_n))^{-\frac{1}{p}} \mathbf{1}_{G_n})^q d\mu \\
&= \lim_{m \to \infty} \int (\sum_{n=1}^{m} (\mu(G_n))^{-\frac{1}{p}} \mathbf{1}_{G_n})^q d\mu \\
&= \lim_{m \to \infty} \int (\sum_{n=1}^{m} (\mu(G_n))^{-\frac{q}{p}} \mathbf{1}_{G_n}) d\mu \text{ only these terms have contribution} \\
&= \lim_{m \to \infty} \sum_{n=1}^{m} \mu(G_n)^{1-\frac{q}{p}} \\
&< \lim_{m \to \infty} \sum_{n=1}^{m} (2^{\frac{q}{p}-1})^{-n} \\
&< \infty \text{ since } \frac{q}{p} - 1 > 1
\end{aligned}$$

Therefore $f \in L^q$.

$$\begin{aligned}
\int |f|^p d\mu &= \lim_{m \to \infty} \int (\sum_{n=1}^{m} (\mu(G_n))^{-\frac{1}{p}} \mathbf{1}_{G_n})^p d\mu \\
&= \lim_{m \to \infty} \sum_{n=1}^{m} (\mu(G_n))^{-1} \mu(G_n) \\
&= \lim_{m \to \infty} m \\
&= \infty
\end{aligned}$$

Therefore $f \notin L^p$. Therefore $\mathcal{X}$ contains sets of arbitrarily large measure $\Rightarrow L^q \not\subset L^p$.

When $q = \infty$, we have $L^p \not\subset L^\infty$ iff $\mathcal{X}$ contains sets of arbitrarily small measure.

($\Rightarrow$). If $\exists f \in L^p$ but $f \notin L^\infty$, then f is not bounded a.e.. According to Chebyshev's inequality, $\forall n > 0$,

$$\mu(\{x \in \mathcal{X} : |f(x)| \geq n\}) \leq \frac{||f||_p^p}{n^p} < \infty$$

and we can make $\{x \in \mathcal{X} : |f(x)| \geq n\}$ arbirarily small by taking n arbitrarily large.

($\Leftarrow$). Suppose $\forall \epsilon > 0$, $\exists E \in \mathcal{B}(\mathcal{X})$ such that $\mu(E) \in (0, \epsilon)$. Let $\{E_n\}$ be such that $0 < \mu(E_n) < 2^{-n}$. By constructing $G_n := E_n \setminus \bigcup_{i=n+1}^{\infty} E_i$, we have G_n disjoint and by monotonicity

$$0 < \mu(G_n) < \mu(E_n) < 2^{-n}$$

Define $f := \sum_{n=1}^{\infty} \mu(G_n)^{-\frac{1}{p+1}} \mathbf{1}_{G_n} \geq 0$. Then, we have verified in the previous part that $f \in L^p$, however

$$f \geq \sum_{n=1}^{\infty} 2^{\frac{n}{p+1}} \mathbf{1}_{G_n}$$

It is not bounded on a set of arbitrarily small but positive measure. Therefore $f \notin L^{\infty}$.

When $q = \infty$, we have $L^{\infty} \not\subset L^p$ iff $\mu(\mathcal{X}) = \infty$.

($\Rightarrow$). Let $f \in L^{\infty}$, then $\exists M$ such that $|f| <= M$ a.e.. Suppose, contrapositively, that $\mu(\mathcal{X}) < \infty$. Then

$$\int |f|^p d\mu \leq M\mu(\mathcal{X}) < \infty$$

amd $f \in L^p$. Therefore If $L^{\infty} \not\subset L^q \Rightarrow \mu(\mathcal{X}) = \infty$.

($\Leftarrow$). Suppose that $\mu(\mathcal{X}) = \infty$. Then, for any constant function $f = C$ that has support on the entire $\mathcal{X}$, $f \in L^{\infty}$. However,

$$||f||_p^p = \int |f|^p d\mu = C\mu(\mathcal{X}) = \infty$$

and then $f \notin L^p$.

Therefore $L^{\infty} \not\subset L^p$.

$\square$

Additionally. Show that for all $0 < p < q \leq +\infty$, $l^p(\mathbb{N}) \subset l^q(\mathbb{N})$, and $L^p([0,1]) \supset L^q([0,1])$, with strict containment in both cases.

Proof. For the first statement, when $q < +\infty$, we want to show that if $v \in l^p(\mathbb{N})$, then $v \in l^q(\mathbb{N})$. From $v \in l^p(\mathbb{N})$ we know that

$$||v||_{l^p}^p = \sum_{i \in \mathbb{N}} |v_i|^p < \infty$$

Thus, $|v_i| < \infty$ for all $i \in \mathbb{N}$. $\exists M > 0$ such that $M \geq \sup_{i \in \mathbb{N}} |v_i|$.

$$\begin{aligned}\sum_{i \in \mathbb{N}} |v_i|^p &= \sum_{i \in \mathbb{N}} |v_i|^{p-q} |v_i|^q \\ &= M^{p-q} \sum_{i \in \mathbb{N}} |v_i|^q \\ &< \infty\end{aligned}$$

Therefore

$$||v||_{l^q}^q = \sum_{i \in \mathbb{N}} |v_i|^q < \infty$$

Thus $v \in l^q$.
When $q = \infty$, if $v \in l^p(\mathbb{N})$, we have

$$\begin{aligned}||v||_{l^\infty(\mathbb{N})} &= \inf\{a \geq 0 : \#(\{i \in \mathbb{N} : |v_i| > a\}) = 0\} \\ &= \sup_{i\in\mathbb{N}} |v_i| \\ &< \infty\end{aligned}$$

Therefore $v \in l^\infty(\mathbb{N})$.

For the second statement, when $q \leq +\infty$, we want to show that if $f \in L^q([0,1])$, then $f \in L^p([0,1])$. From $f \in L^q([0,1])$ we know that

$$||f||^q_{L^q([0,1])} = \int_{[0,1]} |f(x)|^q dx < \infty$$

Since $0 < p < q$, we have

$$||f||^p_{L^p([0,1])} = \int_{[0,1]} |f(x)|^p dx < \int_{[0,1]} |f(x)|^q dx < \infty$$

Therefore $f \in L^p([0,1])$.
When $q = \infty$,

$$||f||_{L^\infty([0,1])} = \inf\{a > 0 : \mu(x \in [0,1] : |f(x)| > a) = 0\}$$

If $f \in L^\infty([0,1])$, then $||f||_{L^\infty([0,1])} < \infty$. Thus, $\exists M > 0$ such that $|f(x)| \leq M$ a.e. in $[0,1]$.

$$\begin{aligned}||f||^p_{L^p([0,1])} &= \int_{[0,1]} |f(x)|^p dx \\ &= \int_{[0,1]\setminus\{|f|>M\}} |f(x)|^p dx + \underbrace{\int_{[0,1]\cap\{|f|>M\}} |f(x)|^p dx}_{=0} \\ &\leq M^p m([0,1]) \\ &= M^p \\ &< \infty\end{aligned}$$

Therefore $f \in L^p([0,1])$.

□

Folland 6.1.9 Suppose $1 \leq p < \infty$. If $||f_n - f||_p \to 0$, then $f_n \to f$ in measure, and hence some subsequence converges to f a.e.. On the other hand, if $f_n \to f$ in measure and $|f_n| \leq g \in L^p$ for all n, then $||f_n - f||_p \to 0$.

Proof. For the first argument, let $\epsilon > 0$, and let $E := \{|f_n - f| \geq \epsilon\}$. Then

$$||f_n - f||^p_p = \int |f_n - f|^p d\mu \geq \int_E |f_n - f|^p d\mu \geq \epsilon^p \mu(E)$$

$$\mu(\{x : |f_n - f| \geq \epsilon\}) \leq \frac{||f_n - f||^p_p}{\epsilon^p}$$

Then, $\mu(\{x : |f_n - f| \geq \epsilon\}) \to 0$ if $||f_n - f||_p \to 0$, and $f_n \to f$ in measure. Given $\delta > 0$, we can choose n_1 such that $\mu(\{x : |f_{n_1} - f| \geq \epsilon\}) < \frac{\delta}{2^1}$, ..., choose n_k such that $\mu(\{x : |f_{n_k} - f| \geq \epsilon\}) < \frac{\delta}{2^k}$ ... Then we can find a subsequence $\{f_{n_k}\}$ such that $\mu(\bigcup_{k\in\mathbb{N}}\{x : |f_{n_k} - f| \geq \epsilon\}) < \sum_{k\in\mathbb{N}} \frac{\delta}{2^k} = \delta$. Since $\delta > 0$ is arbitrary, we have $\mu(\bigcup_{k\in\mathbb{N}}\{x : |f_{n_k} - f| \geq \epsilon\}) = 0$, and this subsequence converges to f a.e..

For the second argument, let $\epsilon > 0$, and notice that $\forall n$, $|f| \leq |f_n| + |f - f_n| \leq |g| + |f_n - f|$. Since

$$\mu(\{|f_n - f| \geq \epsilon\}) \to 0$$

we have

$$\mu(\cap_{n=1}^{\infty}\{|f_n - f| \geq \epsilon\}) = \lim_{m\to\infty} \mu(\cap_{n=1}^{m}\{|f_n - f| \geq \epsilon\}) \leq \lim_{m\to\infty} \mu(\{|f_m - f| \geq \epsilon\}) = 0$$

Therefore $|f| \leq |g| + \epsilon$ a.e., and $f \in L^p$.
Since $|f_n - f| \to 0$ in measure, we have $|f_n - f|^p \to 0$ in measure. Also it is dominated

$$|f_n - f|^p \leq 2^p(g^p + f^p) \in L^p$$

Thus we use the dominated convergence theorem,

$$\lim_{n\to\infty} \int |f_n - f|^p d\mu = \int \underbrace{\lim_{n\to\infty} |f_n - f|^p}_{\text{=0 outside an arbitrarily small measure set}} d\mu = 0$$

□

2.2 Linear Functionals on Banach Spaces

Definition 2.2.1. *(Bounded linear map). For normed vector spaces $(v_i, ||.||_i)$ $i = 1, 2$ and linear map $T : v_1 \to v_2$, we say T is a bounded linear map if*

$$||Tv||_2 \leq C||v||_1 \ \forall v \in V_1, C < \infty$$

The operator norm of T is

$$||T|| := \sup\{\frac{||Tv||_2}{||v||_1} : v \neq 0\} = \sup\{||Tu||_2 : ||u||_1 = 1\}$$

Definition 2.2.2. *(Linear functional and dual space). A linear map from $(V, ||.||)$ to K ($\mathbb{C}$ or $\mathbb{R}$) is called a linear functional, and the space of all bounded linear functionals on $(V, ||.||)$ is called the dual space $(V^*, ||.||)$, where $||.||$ is the operator norm, and this dual space is a Banach space.*

For example, on $V = \mathbb{C}^d$ with Euclidean norm $||v||_E = (\sum |v_i|^2)^{\frac{1}{2}}$. For fixed $w \in \mathbb{C}^d$,

$$\lambda_w(v) := \langle v, w\rangle = \sum_{i=1}^{d} v_i \overline{w_i}$$

is a bounded linear functional (if all bounded linear functionals have this form by representing linear transformation by matrix multiplication). By Cauchy-Schwartz inequality which we will prove later,

$$||\lambda_w(v)|| \leq ||v||_E||w||_E$$

Take $v = \frac{w}{||w||_E}$ where $w \neq 0$, we get

$$||\lambda_w|| \leq ||w||_E$$

So, V^* is isometrically isomorphic to V.

We can also equip $\mathbb{C}^d$ with $||v||_{l^p} = (\sum_{i=1}^d |v_i|^p)^{\frac{1}{p}}$, and verify that $(V^*, ||.||_{l^p}) \simeq (V, ||.||_{l^q})$ for $\frac{1}{p} + \frac{1}{q} = 1$.

For example, On $C([0,1] \to \mathbb{C})$ with $||.||_\infty = ||.||_{\sup}$, $\lambda(f) = f(x_0)$ for some fixed $x_0 \in [0,1]$ is a linear functional.

$$\lambda(f)| = |f(x_0)| \leq ||f||_{\sup}$$

$$||\lambda|| = \sup\{|f(x_0)| : ||f||_{\sup} = 1\} \leq 1$$

$=$ holds for f constant, in which case $||\lambda|| = 1$.

For example, on subspace V of l^∞, $V = \{(v_i)_{i\in\mathbb{Z}} : \sup_i |z_i| < \infty, \lim_{i\to\infty} v_i \text{ exists}\}$,

$$\lambda(v) := \lim_{i\to\infty} v_i$$

is a linear functional of norm 1.

$$||\lambda(v) = \sup\{\lambda(v) : ||v||_\infty = 1\} = \sup\{|\lim_{i\to\infty} v_i| : ||v||_\infty = 1\} = 1$$

Theorem 2.2.1. *(Riesz Representation Theorem, L^p setting) Let $p \in (1, +\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$ as conjugate component. For any $\lambda \in (L^p)^*$, $\exists g \in L^q$, such that $\lambda_g(f) := \lambda(f) = \int_{\mathcal{X}} fgd\mu \; \forall f \in L^p$, and $||\lambda|| = ||g||_{L^q}$. The identification $g \mapsto \lambda_g$ is an isometric isomorphism from L^q to $(L^p)^*$.*

Proof. See Folland P190. $\square$

3 Hilbert Spaces

Hilbert Spaces are a direct generalization of finite-dimensional Euclidean spaces.

3.1 Hilbert space and decomposition

Definition 3.1.1. *(Inner product). Let $\mathcal{H}$ be a complex vector space. An inner product on $\mathcal{H}$ is a map $(f, g) \mapsto \langle f, g\rangle$ from $\mathcal{H} \times \mathcal{H}$ to $\mathbb{C}$ such that*

1. $\langle af + bg, h\rangle = a\langle f, h\rangle + b\langle g, h\rangle$ $\forall a, b \in \mathbb{C}$ *and* $f, g \in \mathcal{H}$.
2. $\overline{\langle f, g\rangle} = \langle g, f\rangle$ $\forall f, g \in \mathcal{H}$.
3. $\langle f, f\rangle \in (0, +\infty)$ $\forall f \neq 0$.

One can immediately get

$$\begin{aligned}
\langle h, af + bg\rangle &= \overline{\langle af + bg, h\rangle} \\
&= \overline{a\langle f, h\rangle} + \overline{b\langle g, h\rangle} \\
&= \overline{a}\langle h, f\rangle + \overline{b}\langle h, g\rangle
\end{aligned}$$

Definition 3.1.2. *(pre-Hilbert Space) A complex vector space with an inner product is called a pre-Hilbert space. If $\mathcal{H}$ is a pre-Hilbert space, for $f \in \mathcal{H}$ we define*

$$||f|| := \sqrt{\langle f, f\rangle}$$

Theorem 3.1.1. *(Cauchy-Schwartz Inequality). $\forall f, g \in \mathcal{H}$,*

$$|\langle f, g\rangle| \leq ||f||||g||$$

with $=$ iff $f = ag$ for some $a \in \mathbb{C}$.

Proof. It is immediate if $\langle f, g\rangle = 0$. So, assume not.
Let $\alpha = sgn\langle f, g\rangle = \frac{\langle f,g\rangle}{|\langle f,g\rangle|}$, $h = \alpha g$. Then,

$$\begin{aligned}
\langle f, h\rangle &= \langle f, \frac{\langle f, g\rangle}{|\langle f, g\rangle|} g\rangle \\
&= \frac{\langle g, f\rangle}{|\langle f, g\rangle|}\langle f, g\rangle \\
&= |\langle f, g\rangle|
\end{aligned}$$

We know that $\forall t \in \mathbb{R}$,

$$0 \leq \langle f - th, f - th\rangle = ||f||^2 - 2t|\langle f, g\rangle| + t^2||g||^2$$

The minimum is reached at $t_0 = \frac{|\langle f,g\rangle|}{||g||^2}$, at which

$$\langle f - t_0h, f - t_0h\rangle = ||f||^2 - \frac{|\langle f, g\rangle|^2}{||g||^2} \geq 0$$

Then we have

$$|\langle f, g\rangle| \leq ||f||||g||$$

with equality iff $f - th = f - \alpha tg = 0$. $\square$

Proposition 3.1.1. $||.||$ *is a norm on* $\mathcal{H}$.

Proof. That $||f|| = 0$ iff $f = 0$ and $||cf|| = |c|||f||$ are immediate from the definition. We then need to check the triangle inequality.

$$\begin{aligned}
||f+g||^2 &= \langle f+g, f+g\rangle \\
&= ||f||^2 + 2Re(\langle f, g\rangle) + ||g||^2 \\
&\leq ||f||^2 + 2||f||||g|| + ||g||^2 \text{ by Cauchy-Schwartz Inequality} \\
&= (||f|| + ||g||)^2
\end{aligned}$$

Take square root and then we have $||f+g|| \leq ||f|| + ||g||$. $\square$

Definition 3.1.3. *(Hilbert Space) A Hilbert space is a pre-Hilbert space that is complete with respect to the norm* $||f|| = \sqrt{\langle f, f\rangle}$.

For example, it is easy to check that on $L^2(\mathcal{X}, \mathcal{B}, \mu)$, $\langle f, g\rangle = \int_{\mathcal{X}} f\overline{g}d\mu$ is an inner product, and we have shown that L^p is complete with respect to the norm $||f||_{L^2} = \sqrt{\langle f, f\rangle} = (\int_{\mathcal{X}} |f|^2 d\mu)^{\frac{1}{2}}$. Therefore any Cauchy sequence under the norm $\langle f, f\rangle = ||f||^2_{L^2}$ converges in L^2, and thus $L^2(\mathcal{X}, \mathcal{B}, \mu)$ is a Hilbert space.

Lemma 3.1.1. *If* $x_n \to x$, $y_n \to y$, *then* $\langle x_n, y_n\rangle \to \langle x, y\rangle$.

Proof. By Cauchy-Schwartz inequality,

$$\begin{aligned}
|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \\
&\leq ||x_n - x||||y_n|| + ||x||||y_n - y|| \\
&\to 0
\end{aligned}$$

as $x_n \to x$, $y_n \to y$. $\square$

Lemma 3.1.2. *(Parallelogram Law). For all* $f, g \in \mathcal{H}$,

$$||f+g||^2 + ||f-g||^2 = 2(||f||^2 + ||g||^2)$$

"The sum of the squares of the diagonals of a parallelogram is the sum of the squares of the four sides".

Proof. Sum up the following two identities

$$||f+g||^2 = ||f||^2 + 2Re\langle f, g\rangle + ||g||^2$$

$$||f-g||^2 = ||f||^2 - 2Re\langle f, g\rangle + ||g||^2$$

$\square$

Definition 3.1.4. *(Orthogonal). Say f is orthogonal to g, $f \perp g$, if $\langle f, g\rangle = 0$. For $E \subset \mathcal{H}$, denote $E^{\perp} = \{g \in \mathcal{H} : \langle g, f\rangle = 0 \ \forall f \in E\}$.*

Definition 3.1.5. *(Closedness). A subspace $S \subseteq \mathcal{H}$ is closed if $\forall \{f_n\} \subseteq S$ such that $f_n \to f \in \mathcal{H}$, we have $f \in S$.*

We can easily verify that $E^{\perp}$ is a closed subspace of $\mathcal{H}$, by the previous convergence lemma and the linearity of inner product. If $\{g_n\} \subseteq E^{\perp}$ such that $g_n \to g \in \mathcal{H}$, then

$$\langle g, f\rangle = \lim_{n\to\infty} \langle g_n, f\rangle = \lim_{n\to\infty} 0 = 0$$

Therefore $g_n \to g \in E^{\perp}$.

Theorem 3.1.2. *(Pythagorean Identity). If $f_1, ..., f_n \in \mathcal{H}$, $f_i \perp f_j \ \forall i \neq j$, then*

$$||\sum_{i=1}^{n} f_i||^2 = \sum_{i=1}^{n} ||f_i||^2$$

Proof.

$$\begin{aligned} ||\sum_{i=1}^{n} f_i||^2 &= \langle \sum_{i=1}^{n} f_i, \sum_{i=1}^{n} f_i \rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \langle f_i, f_j\rangle \\ &= \sum_{i=1}^{n} \langle f_i, f_i\rangle \\ &= \sum_{i=1}^{n} ||f_i||^2 \end{aligned}$$

□

Theorem 3.1.3. *Let V be a closed subspace of $\mathcal{H}$. Then, $\mathcal{H} = V \oplus V^{\perp}$, that is, $\forall h \in \mathcal{H}$, $\exists$ unique $f \in V, g \in V^{\perp}$ such that $h = f + g$. Moreover, f, g are the unique minimizers of $\{||v - h||\}_{v\in V}$ and $\{||w - h||\}_{w\in V^{\perp}}$, respectively.*

Proof. Fix $h \in \mathcal{H}$. Let $\delta = \inf\{||v - h|| : v \in V\}$. Let $\{v_n\} \subset V$ such that $||v_n - h|| \to \delta$. Claim 1: $\{v_n\}$ is Cauchy. Indeed, from Parallelogram Identity,

$$2(||v_n - h||^2 + ||v_m - h||^2) = ||v_n - v_m||^2 + ||v_n + v_m - 2h||^2$$

$$\begin{aligned} ||v_n - v_m||^2 &= 2||v_n - h||^2 + 2||v_m - h||^2 - 4||\underbrace{\frac{1}{2}(v_n + v_m)}_{\in V} - h||^2 \\ &\leq 2||v_n - h||^2 + 2||v_m - h||^2 - 4\delta^2 \\ &\to 0 \end{aligned}$$

as $n, m \to 0$. Therefore Claim 1 holds.

Since V is closed in $\mathcal{H}$ complete, $\exists f \in V$ such that $v_n \to f$. By Triangle Inequality,

$$||h-f|| \in [\underbrace{||h-v_n||}_{\to\delta} - \underbrace{||v_n-f||}_{\to 0}, \underbrace{||h-v_n||}_{\to\delta} + \underbrace{||v_n-f||}_{\to 0}]$$

Taking $n \to \infty$, we have $||h-f|| = \delta$. Thus we have found $f = \lim_{n\to\infty} v_n$.
Let $g = h - f$. Claim 2: $g \in V^{\perp}$. Indeed, let $v \in V$ arbitrary. We want to show that $\langle g, v\rangle = 0$. Scale v to assume that $\langle g, v\rangle \in \mathbb{R}$.
Let $F(t) = ||g+tv||^2 = ||g||^2 + 2t\langle g, v\rangle + t^2||v||^2$. Since $||g+tv|| = ||h-(f-tv)|| \geq \delta$, we have $F(t) \geq \delta^2$ $\forall t \in \mathbb{R}$, with $=$ at $t=0$. Thus F is minimized at $t=0$.

$$0 = F'(0) = 2\langle g, v\rangle$$

Therefore $g \in V^{\perp}$. Claim 2 holds.
Next we prove the uniqueness of g. For any other $g' \in V^{\perp}$, since $h-g = f \in V$, $g-g' \in V^{\perp}$, we have by Pythagoras

$$\begin{aligned} ||h-g'||^2 &= ||h-g||^2 + ||g-g'||^2 \\ &\geq ||h-g||^2 \end{aligned}$$

with $=$ iff $g = g'$.
Next we prove that the decomposition is unique: $h = f + g$. If for $f' \in V, g' \in V^{\perp}$, $h = f' + g'$, then $f - f' = g' - g \in V \cap V^{\perp}$, which are orthogonal to themselves, and then $f = f'$, $g = g'$. ☐

3.2 Linear Functionals on Hilbert Space

Definition 3.2.1. *(Linear functional) For $g \in \mathcal{H}$, define $\lambda_g : \mathcal{H} \to \mathbb{C}$ as $\lambda_g(h) = \langle h, g\rangle$ as a linear functional on $\mathcal{H}$.*

By Cauchy-Schwartz Inequality,

$$|\lambda_g(h)| \leq ||g||||h||$$

with $=$ for $h = \frac{g}{||g||}$ or $h = 0$. Therefore $||\lambda_g|| = ||g||$.
So, the mapping $\mathcal{H} \ni g \mapsto \lambda_g \in \mathcal{H}^*$ is isometry of $\mathcal{H}$ to $\mathcal{H}^*$ (norm preserving). It is also conjugate linear since $\lambda_{\alpha f + \beta g} = \overline{\alpha}\lambda_f + \overline{\beta}\lambda_g$. It is a fundamental fact that this mapping is surjective, as the following theorem goes.

Theorem 3.2.1. *(Riesz Representation Theorem for Hilbert Space) If $\lambda \in \mathcal{H}^*$, then $\exists! g \in \mathcal{H}$ such that $\lambda = \lambda_g$, i.e., $\lambda(h) = \langle h, g\rangle$ $\forall h \in \mathcal{H}$.*

Proof. We first prove uniqueness. If $\langle h, g\rangle = \langle h, g'\rangle$ $\forall h$, then take $h = g - g'$ and we have $\langle g-g', g\rangle = \langle g-g', g'\rangle$, that is, $||g-g'||^2 = 0$ and then $g = g'$.
We then prove existence. Let $\lambda \in \mathcal{H}^*$. If $\lambda = 0$, then $\lambda(h) = 0$ $\forall h \in \mathcal{H}$. $g = 0$ is obvious.
Otherwise, let $V = \{h \in \mathcal{H} : \lambda(h) = 0\}$. Then M is a proper closed subspace of $\mathcal{H}$ since λ is continuous. By Theorem 3.0.3, $\mathcal{H} = V \oplus V^{\perp}$. $V \neq \mathcal{H}$ since $\lambda \neq 0$. Then $V^{\perp} \neq \{0\}$. Pick

$f \in M^\perp$ with $||f|| = 1$. If $u = \lambda(h)f - \lambda(f)h$, then

$$\begin{aligned}\lambda(\lambda(h)f - \lambda(f)h) &= \langle \lambda(h)f - \lambda(f)h, g\rangle \\ &= \lambda(h)\langle f, g\rangle - \lambda(f)\langle h, g\rangle \\ &= 0\end{aligned}$$

Then $u \in M$. Then

$$\begin{aligned}0 &= \langle f, u\rangle \\ &= \lambda(h)||f||^2 - \lambda(f)\langle h, f\rangle \\ &= \lambda(h) - \langle h, \overline{\lambda(f)}f\rangle\end{aligned}$$

Therefore $\lambda(h) = \langle h, g\rangle$ where $g = \overline{\lambda(f)}f$. □

Folland 5.5.56 If E is a subset of a Hilbert space $\mathcal{H}$, $(E^\perp)^\perp$ is the smallest closed subspace of $\mathcal{H}$ containing E.

Proof. Let $E \subset \mathcal{H}$. WTS $(E^\perp)^\perp$ is closed, and $\forall S \subset \mathcal{H}$ such that $E \subseteq S$, $(E^\perp)^\perp \subseteq S$.

$$E^\perp = \{x \in \mathcal{H} : \langle x, y\rangle = 0 \forall y \in E\}$$

We first prove that $E^\perp$ is closed. Let $\{x_n\} \subset E^\perp$ so that $x_n \to x \in \mathcal{H}$. WTS $x \in E^\perp$. Since $\{x_n\} \in E^\perp$, we have $\langle x_n, y\rangle = 0\ \forall n\ \forall y \in E$. By a proposition above we have $\langle x_n, y\rangle \to \langle x, y\rangle$. Suppose $\exists y \in E$ sp that $\langle x, y\rangle \neq 0$. Then, $\langle x_n, y\rangle \to \langle x, y\rangle \neq 0$, contradict with $\langle x_n, y\rangle\ \forall n$. Therefore, $E^\perp$ is closed.
Since $E^\perp \subset \mathcal{H}$, we have $(E^\perp)^\perp$ is also closed.
Let $S \subset \mathcal{H}$ be a closed subset so that $E \subseteq S$. WTS $(E^\perp)^\perp \subseteq S$.
Claim 1: $S^\perp \subseteq E^\perp$. Let $x \in S^\perp$. Let $y \in E$ be arbitrary. Since $E \subseteq S$, $y \in S$. Since $x \in S^\perp$, we have $\langle x, y\rangle = 0$. Then, $x \in E^\perp$. Therefore $S^\perp \subseteq E^\perp$. Then, notice that $(E^\perp)^\perp \subseteq (S^\perp)^\perp$.
Since S is a closed subspace of $\mathcal{H}$, by the theorem above, $\mathcal{H} = S \oplus S^\perp$. Let $x \in (S^\perp)^\perp \subset \mathcal{H}$. Then, $x = y + z$ where $y \in S$ and $z \in S^\perp$. Note that clealy $\langle x, z\rangle = 0$, and $\langle x, z\rangle = \langle y + z, z\rangle = \langle y, z\rangle + \langle z, z\rangle = 0$. Since $y \in S$ and $z \in S^\perp$, then $\langle y, z\rangle = 0$. Then, $\langle z, z\rangle = 0$ which means $z = 0$. Then, $x = y \in S$. Therefore, $x \in (S^\perp)^\perp \Rightarrow x \in S$. $(S^\perp)^\perp \subseteq S$. Since $(E^\perp)^\perp \subseteq (S^\perp)^\perp \subseteq S$, we have $(E^\perp)^\perp$ is the smallest closed subspace of $\mathcal{H}$ that contains E. □

3.3 Orthonormal Set and Orthonormal Basis

Definition 3.3.1. *(Orthonormal sets)* $\{u_\alpha\}_{\alpha \in A}$ *is an orthonorml set if* $||u_\alpha|| = 1\ \forall \alpha$ *and* $u_\alpha \perp u_\beta\ \forall \alpha \neq \beta$.

The Gram-Schmidt process is used to converting a linearly independent sequence $\{x_n\}$ into an orthonormal sequence $\{u_n\}$, such that the linear span of $\{x_n\}_{n=1}^N$ coincides with the linear span of $\{u_n\}_{n=1}^N$.

1. Set $u_i = \frac{x_1}{||x_1||}$.

2. Having set $u_1, ..., u_{N-1}$, set $v_N = x_N - \sum_{n=1}^{N-1} \langle x_N, u_n \rangle u_n$. First $v_N \neq 0$ since x_N is not in the linear span of $x_1, ..., x_{N-1}$ and thus not in the linear span of $u_1, ..., u_{N-1}$. Then $\forall m < N$,

$$\begin{aligned} \langle v_N, u_m \rangle - \langle x_N - \sum_{n=1}^{N-1} \langle x_N, u_n \rangle u_n, u_m \rangle \\ = \langle x_N, u_m \rangle - \langle x_N, u_m \rangle \\ = 0 \end{aligned}$$

3. Set $u_N = \frac{v_N}{||V_N||}$.

Lemma 3.3.1. *(Bessel's Inequality) Let* $\{u_\alpha\}_{\alpha \in A}$ *be orthonormal in* $\mathcal{H}$*. Then* $\sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2 \leq ||x||^2 \; \forall x \in \mathcal{H}$*. In particular,* $\{\alpha : \langle x, u_\alpha \rangle \neq 0\}$ *is countable.*

Proof. It suffices to show that $\sup_{F \subset A, F \text{ finite}} \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 \leq ||x||^2$. Without loss of generality, let A be finite.

$$0 \leq ||x - \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha||^2 = ||x||^2 - 2Re(\langle x, \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha \rangle) + ||\sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha||^2$$

$$\begin{aligned} \langle x, \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha \rangle &= \sum_{\alpha \in A} \langle x, \langle x, u_\alpha \rangle u_\alpha \rangle \\ &= \sum_{\alpha \in A} \overline{\langle x, u_\alpha \rangle} \langle x, u_\alpha \rangle \\ &= \sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2 \end{aligned}$$

By Pythagorean Identity,

$$\begin{aligned} ||\sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha||^2 &= \sum_{\alpha \in A} ||\langle x, u_\alpha \rangle u_\alpha||^2 \\ &= \sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2 \end{aligned}$$

Combine them together, we have the Bessel's Inequality. □

Theorem 3.3.1. *Let* $\{u_\alpha\}_{\alpha \in A}$ *be an orthonormal set on* $\mathcal{H}$*. The followings are equivalent.*

1. *(Completeness). If* $\langle x, u_\alpha \rangle = 0 \; \forall \alpha \in A$*, then* $x = 0$*.*
2. *(Parseval Identity).* $||x||^2 = \sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2$*.*
3. $\forall x \in \mathcal{H}$*,* $x = \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha$*, where the sum in the right hand side has only countably many non-zero terms, and converges in the norm topology no matter how these terms are ordered.*

Before proving Theorem 3.1.2, we prove the following proposition.

Proposition 3.3.1. $\langle .,. \rangle$ *is continuous in norm topology, that is, if* $||x_n - x|| \to 0$*,* $||y_n - y|| \to 0$*, then* $\langle x_n, y_n \rangle \to \langle x, y \rangle$*.*

Proof.

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &\le |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\le ||x_n||||y_n - y|| + ||x_n - x||||y||\end{aligned}$$

Since $||y_n - y|| \to 0$, $||x_n - x|| \to 0$, we have $\langle x_n, y_n\rangle \to \langle x, y\rangle$. □

Now we prove Theorem 3.1.2.

Proof. (2) to (1) is immediate.
(1) to (3): By Bessel's Inequality, we can enumerate $\{\alpha : \langle x, u_\alpha\rangle \neq 0\}$ as $\{\alpha_j\}_{j=1}^\infty$. And moreover, $\sum_{j=1}^\infty |\langle x, u_{\alpha_j}\rangle|^2$ converges.
By Pythagorean Identity,

$$\begin{aligned}||\sum_{j=m}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j}||^2 &= \sum_{j=m}^n |\langle x, u_{\alpha_j}\rangle|^2 \\ &\to 0\end{aligned}$$

as $n, m \to \infty$. So, $(\sum_{j=1}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j})_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{H}$. Hence, it converges.
Let $y := x - \sum_{j=1}^\infty \langle x, u_{\alpha_j}\rangle u_{\alpha_j}$. By 1, to show $y = 0$, it is enough to show $\langle y, u_\beta\rangle = 0\ \forall \beta \in A$.
Let $\beta \in A$. By continuity of $\langle ., .\rangle$,

$$\begin{aligned}\langle y, u_\beta\rangle &= \langle x, u_\beta\rangle - \lim_{n\to\infty} \langle \sum_{j=1}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j}, u_\beta\rangle \\ &= \langle x, u_\beta\rangle - \lim_{n\to\infty} \sum_{j=1}^n \langle x, u_{\alpha_j}\rangle \mathbf{1}_{\alpha_j = \beta} \\ &= 0\end{aligned}$$

Therefore we have 3.
(3) to (2): Let $x \in \mathcal{H}$ be arbitrary. Let $\{\alpha_j\}_{j=1}^\infty = \{\alpha : \langle x, u_\alpha\rangle \neq 0\}$. Just as in the proof of Bessel's Inequality,

$$||x - \sum_{j=1}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j}||^2 = ||x||^2 - \sum_{j=1}^n |\langle x, u_{\alpha_j}\rangle|^2$$

and $LHS \to 0$ as $n \to \infty$ by 3. Therefore, $||x||^2 = \sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2$, as in 2.

□

Definition 3.3.2. *(Orthonormal basis). An orthonormal set satisfying the above 1 - 3 is called an orthonormal basis.*

For example, $\mathcal{H} = l^2(\mathbb{N}) = \{(a_n)_{n=1}^\infty : \sum_{k=1}^\infty |a_k|^2 < \infty\}$. Let $e_n \in l^2(\mathbb{N})$ be the "canonical basis vector" with $e_n = (\mathbf{1}_{k=n})_{k=1}^\infty$. For $a = (a_n)_{n=1}^\infty \in \mathcal{H}$, we have $\langle a, e_n\rangle = \sum_{k=1}^\infty a_k \mathbf{1}_{k=n} = a_n\ \forall n$. By 1 of the previous theorem, since $a_n = \langle a, e_n\rangle = 0\ \forall n \Rightarrow a = 0$, it folows that $(e_n)_{n=1}^\infty$ is an orthonormal basis for $l^2(\mathbb{N})$.

Proposition 3.3.2. *Every Hilbert space has an orthonormal basis.*

Before we prove this proposition, we recall some notions and Zorn's lemma.

Definition 3.3.3. *(Partially ordered set) Partially ordered set $(S, \leq)$ is a set S equipped with a relation $\leq$ that is a partial ordering, i.e., $\leq$ is reflexive ($x \leq x$), antisymmetric ($x \leq y$, $y \leq x \Rightarrow x = x$), transitive ($x \leq y$, $y \leq z \Rightarrow x \leq z$).*

Definition 3.3.4. *(Totally ordered set) $(S, \leq)$ is totally ordered if $\forall s_1, s_2 \in S$, either $s_1 \leq s_2$ or $s_2 \leq s_1$.*

Definition 3.3.5. *(Chain) $T \subseteq S$ is a chain if it is totally ordered.*

Definition 3.3.6. *(Upper bound). An upper bound for a subset $U \subset S$ is an element $s \in S$ such that $u \leq s$ $\forall u \in U$.*

Definition 3.3.7. *(Maximal element) A maximal element s^* of S is such that $\nexists s \in S$ s.t. $s^* \leq s$ and $s \neq s^*$, i.e. $s^* \leq s \Rightarrow s = s^*$.*

Lemma 3.3.2. *(Zorn's lemma). If a partially ordered set $(S, \leq)$ has the property that every chain in S has an upper bound in S, then S must contain a maximal element s^*.*

Now we can prove Proposition 3.1.2.

Proof. Consider the collection S of all orthonormal sets in $\mathcal{H}$ with partial ordering $\subseteq$. Let $T \subseteq S$ be an arbitrary chain. Then, $T^* = \bigcup_{t \in T}\{t\}$ is an orthonormal set and an upper bound for T. Hence, using Zorn's lemma, S contains a maximal element $S^* = \{u_\alpha\}_{\alpha \in A}$.
Claim: S^* is an orthonormal basis. Indeed, suppose otherwise, then 1 in Theorem 3.1.2. fails, i.e. $\exists x \in \mathcal{H}$, $\langle x, u_\alpha\rangle = 0$ $\forall \alpha \in A$, and $x \neq 0$. But then $S^* \bigcup \{\frac{x}{||x||}\}$ is an orthonormal set, and this contradicts with maximality of S^*. □

Definition 3.3.8. *(Separable) A topological space is separable if it has a countable dense subset.*

Proposition 3.3.3. *$\mathcal{H}$ is separable iff it has a countable orthonormal basis, in which case every orthonormal basis for $\mathcal{H}$ is countable.*

Proof. ($\Rightarrow$). If $\{x_n\}$ is a countable dense set in $\mathcal{H}$, by discarding recursively any x_n that is in the linear span of $x_1, ..., x_{n-1}$, we obtain a linearly independent sequence $\{y_n\}$, whose linear span is dense in $\mathcal{H}$.Application of the Gram-Schmidt process yields an orthonormal sequence $\{u_n\}$ whose linear span is dense in $\mathcal{H}$ and which is therefore a basis.
($\Leftarrow$). If $\{u_n\}$ is a countable orthonormal basis, the finite linear combinations of the u_n's with coefficients in a countable dense subset of $\mathbb{C}$ forms a countable dense set in $\mathcal{H}$. Moreover, if $\{v_\alpha\}_{\alpha \in A}$ is another orthonormal basis, for each n the set $A_n := \{\alpha \in A : \langle u_n, v_\alpha\rangle \neq 0\}$ is countable. By completeness of $\{u_n\}$, $\forall \alpha \in A$, $v_\alpha \neq 0 \Rightarrow \langle v_\alpha, u_n\rangle \neq 0$ $\exists n$, therefore $A \subset \bigcup_{n \in \mathbb{N}} A_n$. Therefore $A = \bigcup_{n \in \mathbb{N}} A_n$ and A is countable. □

Definition 3.3.9. *(Unitary map) A unitary map $U : \mathcal{H}_1 \to \mathcal{H}_2$ between Hilbert spaces $(\mathcal{H}_i, \langle ., . \rangle_i)$ is an invertible linear map that preserves inner products. $\forall x, y \in \mathcal{H}_1$, $\langle Ux, Uy\rangle_2 = \langle x, y\rangle_1$.*

Remark: by taking $x = y$, we see that every unitary map is isometry. $||Ux||_2 = ||x||_1$.

Lemma 3.3.3. *Any surjective linear isometry $U : \mathcal{H}_1 \to \mathcal{H}_2$ is unitary.*

Proof. (a) **(The polarization identity)** For any $x, y \in \mathcal{H}$,

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right).$$

(Completeness is not needed here.)

(b) If $\mathcal{H}'$ is another Hilbert space, a linear map from $\mathcal{H}$ to $\mathcal{H}'$ is unitary if and only if it is isometric and surjective.

(a).

$$\begin{aligned}
||x+y||^2 - ||x-y||^2 &= \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle - \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle - \langle y,y\rangle \\
&= \langle x,y\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,x\rangle
\end{aligned}$$

$$\begin{aligned}
i||x+iy||^2 - i||x-iy||^2 &= i\langle x+iy, x+iy\rangle - i\langle x-iy, x-iy\rangle \\
&= \langle x,y\rangle + \langle x,y\rangle - \langle y,x\rangle - \langle y,x\rangle
\end{aligned}$$

Combine them together and divide by 4, we get

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right)$$

(b). Let $T : \mathcal{H} \to \mathcal{H}'$.

($\Rightarrow$). Since T is unitary, it is linear, surjective, and $\langle Tx, Ty\rangle = \langle x, y\rangle$. Then,

$$\begin{aligned}
||Tx||^2 &= \langle Tx, Tx\rangle \\
&= \langle x, x\rangle \\
&= ||x||^2
\end{aligned}$$

Therefore $||Tx|| = ||x||$, and then T is isometric.

($\Leftarrow$). Since T is linear, surjective, and isometric,

$$\begin{aligned}
\langle Tx, Ty\rangle &= \frac{1}{4}\left(\|Tx+Ty\|^2 - \|Tx-Ty\|^2 + i\|Tx+Tiy\|^2 - i\|Tx-Tiy\|^2\right) \\
&= \frac{1}{4}\left(\|T(x+y)\|^2 - \|T(x-y)\|^2 + i\|T(x+iy)\|^2 - i\|T(x-iy)\|^2\right) \\
&= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) \\
&= \langle x, y\rangle
\end{aligned}$$

Also, surjective ensures full image. Therefore, T is unitary.

□

Unitary maps are the true "isomorphisms" in the category of Hilbert spaces. They preserve not only the linear structure and the topology, but also the norm and the inner product. From the pointview of this abstract structure, every Hilbert space looks like a l^2 space.

Proposition 3.3.4. *Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal basis for $\mathcal{H}$. Define $U : \mathcal{H} \to l^2(A) = L^2(A, 2^A, \#)$ by $U(x) = \widehat{x}$, $\widehat{x}(\alpha) := \langle x, u_\alpha \rangle$ (coordinate expansion), $\alpha \in A$. Then, U is unitary.*

Proof. Linearity follows from linearity of inner product. $U(x+y) = \widehat{x+y}$, $\widehat{x+y}(\alpha) = \langle x+y, u_\alpha \rangle = \langle x, u_\alpha \rangle, \langle x, u_\alpha \rangle$, therefore $\widehat{x+y} = \widehat{x} + \widehat{y}$.
Isometry follows from Parseval's Identity $||x||^2 = \sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2 = |\widehat{x}|^2_{l^2(A)}$.
Now we prove that U is surjective. Let $f \in l^2(A)$. Thus, $\sum_{\alpha \in A} |f(\alpha)|^2 < \infty$. Therefore, the non-zero terms are countable, and we enumerate $\{\alpha : |f(\alpha)| \neq 0\} = \{\alpha_j\}_{j=1}^{\infty}$. Moreover, $\sum_{j=1}^{\infty} |f(\alpha_j)|^2$ converges. By Pythagorean Identity,

$$|| \sum_{j=m}^{n} f(\alpha_j) u_{\alpha_j} ||^2 = \sum_{j=m}^{n} |f(\alpha_j)|^2 \to 0$$

as $n, m \to \infty$. So, $(\sum_{j=1}^{N} f(\alpha_j) u_{\alpha_j})$ is Cauchy, $\sum_{j=1}^{N} f(\alpha_j) u_{\alpha_j} \to \sum_{\alpha \in A} f(\alpha) u_\alpha = x \in \mathcal{H}$. $\widehat{x}(\alpha) = \langle \sum_{\beta \in A} f(\beta) u_\beta, u_\alpha \rangle = f(\alpha)$. Therefore U is surjective.
Following from the previous lemma, we have U is unitary.

$\square$

Corollary 3.3.1. *Any separable Hilbert space is (unitarily) isomorphic to $l^2(\mathbb{N})$.*

Proof. Let $\mathcal{H}$ be separable Hilbert space, then it has a countable orthonormal bases $\{u_n\}_{n \in \mathbb{N}}$. Then define $U : \mathcal{H} \to l^2(\mathcal{N})$ by $U(x) = \widehat{x}$, $\widehat{x}(n) = \langle x, u_\alpha \rangle$, then U is a unitary map between $\mathcal{H}$ and $l^2(\mathbb{N})$. Then it preserves the linear structure, the topology, the norm, and the inner product. $\square$

3.4 Citation

This is a citation[**?**].

References