Abstract Analysis

January 17, 2025

Contents

1 Introduction

We would like to state and portray the uses of some of the important and highly usable results of integration theory, elucidating in the process the analytical thought which is of paramount importance in any route of exploration in this field[1]. We give bare-bone proofs as all this is standard, but we will highlight the main part of the proof by ♡ or if there are many main parts, then by ♡♡ . . . (!) Let us first begin with some motivation behind modern measure theory.

We know that the class of all Riemann integrable functions on $[a, b]$, denoted $R([a, b])$, is not complete under pointwise limit (a sequential approximation of Dirichlet's function shows that). Further, motivated by Weierstrass approximation, one would like to have commutability results between lim and $\int$, which again $R([a, b])$ lacks. Consequently, one is motivated to find a larger class of "integrable" functions for which these defects would be rectified.

The idea that H. Lebesgue had was quite simple. He continued the idea of Riemann (that is, of partitions) but made sure that the function under investigation is much more intertwined in with it. Indeed, for a bounded function $f : [a, b] \to \mathbb{R}$, we contain the image $\operatorname{Im}(f) \subseteq [\alpha, \beta]$ and then consider a partition $\mathcal{P} = \{I_i\}_{i=1}^n$ where I_i is an interval. Now choose $\xi_i \in f^{-1}(I_i) =: J_i$ for each i. Consequently, we may naturally define *Lebesgue sum of f w.r.t. $\mathcal{P}$* as follows

$$L(f, \mathcal{P}) := \sum_{i=1}^{n} f(\xi_i) m(J_i),$$

where $m(J_i)$ is supposed to be some sort of measure of J_i. Note that J_i in general might be very bad (may not even be an interval!). To complete this idea of "integration", we are naturally led to considering more general notions of measures. Indeed, this is what we will pursue in this course.

Remark 1.0.1. (***Pseudo-definition of measure***) First, what do we expect from a notion of measure on $\mathbb{R}$? Perhaps the following is the minimum conditions we would require to call a function "measure": A function $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ is said to be a *pseudo-measure* if it satisfies the following

1. (*measure of intervals*) for any interval I, the measure $\mu(I) = l(I)$ where l is the length function,
2. (*measure of disjoint unions*) for any disjoint sequence of subsets $\{A_n\}$, $\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$,

[1] One may argue, instead, in whole of mathematics.

3. (*translation invariance*) for any subset A and $x \in \mathbb{R}$, we have $\mu(A+x) = \mu(A)$.

We will call such a function a *pseudo-measure on* $\mathbb{R}$. Observe that for $A \subseteq B$, we obtain $\mu(A) \leq \mu(B)$ by breaking $B = A \cup B \setminus A$. We call μ a pseudo-measure because it does not exists!

Theorem 1.0.2. *(Vitali set) There exists no pseudo-measure on $\mathbb{R}$. In paritcular, there exists a set $V \subseteq \mathbb{R}$ such that for a pseudo-measure μ, $\mu(V) \notin [0, \infty]$.*

Proof. We will construct such a set V. Begin with the closed interval $J = [0,1]$. Define an equivalence relation $\sim$ on J given as follows:

$$x \sim y \iff x - y \in \mathbb{Q}.$$

This can easily be seen to be an equivalence relation on J. We have first some observations to make about this equivalence relation and the consequent partition of J that it entails.

1. Observe that the class of any rational r in J under $\sim$ is simply $[0]$, as $r - 0 \in \mathbb{Q}$.
2. Every equivalence class is countable in size. Indeed, for any $x \in J$, the class $[x]$ is just translate of x by rationals, which is countable.
3. There are uncountably many equivalence classes. Indeed, if there were atmost countably many equivalence classes, then by statement 2 above, it would follow there are atmost countably many elements in J, which is a contradiction.

Consequently, this equivalence relation partitions J into following classes:

$$J = \bigcup_{\alpha \in \mathcal{I}} [\alpha]$$

where $\mathcal{I}$ is an uncountable set.

We would now construct the set V as follows. First, let us assume axiom of choice, so that for each class $[\alpha]$, we may pick an element $r_\alpha \in [\alpha]$ and would thus obtain a subset of J, denoted $V = \{r_\alpha \mid \alpha \in I\}$. We call this the Vitali set.

Consider the set $Q = [-1,1] \cap \mathbb{Q}$. Since it is countable so consider an enumeration $Q = \{q_n\}$. Now consider the translates $V + q_n$ for all $n \in \mathbb{N}$ and their union $X = \bigcup_n V + q_n$. We now observe the following two facts about X.

1. If $n \neq m$, then $(V + q_n) \cap (V + q_m) = \emptyset$. Indeed, if $x \in (V + q_n) \cap (V + q_m)$, then $x = r_a + q_n = r_b + q_m$. Consequently, $r_a - r_b \in \mathbb{Q}$ and hence $[a] = [b]$. But by single choice of r_c for each $c \in \mathcal{I}$, we get $r_a = r_b$ and thus $q_n = q_m$ from above, which is a contradiction.
2. $J = [0,1] \subseteq X$. Indeed, for any $x \in [0,1]$, consider the class $[a]$ in which x is present. Consequently we have a unique $r_a \in V$ corresponding to x which satisfies $x \in [r_a]$. Thus, $x = r_a + t$ where $t \in \mathbb{Q}$. We may write $t = q_n$ to obtain that $x \in V + q_n$, as desired.
3. $X \subseteq [-1,2]$. Indeed, this follows immediately since $X = \bigcup_n V + q_n$ where q_ns are rationals in $[-1,1]$ and $V \subseteq [0,1]$.

With the above three observations, we obtain the following inclusions:

$$[0,1] \subseteq \bigcup_n V + q_n \subseteq [-1,2].$$

Now, if we apply the pseudo-measure μ on the above inclusions, we will obtain the following:

$$1 \leq \sum_n \mu(V) \leq 3.$$

If $\mu(V) = 0, \infty$, then we have an immediate contradiction. Else if $0 < \mu(V) < \infty$, then $\sum_n \mu(V) = \infty$ and we again have a contradiction. Thus, $\mu(V) \notin [0, \infty]$, a contradiction. □

Remark 1.0.3. The main issue in pseudo-measures is that we trying to get a measure on *all* of the subsets of $\mathbb{R}$. By Theorem 1.0.2, this is hopeless. What we shall now do instead is to obtain a measure not on all of the subsets of $\mathbb{R}$, but rather on only a subcollection of subsets of $\mathbb{R}$, and we shall choose this subcollection in a manner so that we don't allow sets like Vitali sets. Indeed, this becomes our point of departure for the abstract definition of σ-algebras and measure/measurable spaces, the need for the right domain of a measure function.

1.1 Few introductory notions

These are few of the basic definitions that one might remember from real analysis.

- **Limit Points** : $x \in X$ is called a *limit point* of a subset $S \subseteq X$ if $\forall\, r > 0$, $\exists\, a \neq x$ such that $a \in S \cap B_r(x)$. That is, ball of any size r around x contains atleast one point of S.
- **Isolated Points** : $y \in S$ is called an *isolated point* of a subset $S \subseteq X$ if $\exists\, r > 0$ such that $(B_r(y) \setminus \{y\}) \cap S = \Phi$. That is, $B_r(y)$ contains no other point of S apart from y.
  - Also note that every point of closure $\overline{S}$ is either a limit point or an isolated point of S.
  - More specifically, *any subset of $\mathbb{R}^d$ is closed if and only if it contains all of it's limit points.*
- **Perfect Set** : A is called a perfect set if $A = A'$ where A' is the set of all *limit points* of A. More conveniently, if A does not contain any isolated points then it is a perfect set. $\mathbb{R}$ is a perfect set.
- **Symmetric Difference** : A and B are two sets then symmetric difference is $A \Delta B = (A \setminus B) \cup (B \setminus A)$.
- **Power Set** : Collection of all subsets of a set S, written as $P(S)$.
- **Lower Bound** : A *lower bound* of a subset S of a poset $(P, \leq)$ is an element $a \in P$ such that $a \leq x$ for all $x \in S$.
- **Infimum** : A lower bound $p \in P$ is called an *infimum* of S if for all lower bounds y of S in P, $y \leq p$.
- **Limit Infimum** : For a sequence $\{x_n\}$, limit inferior is defined by:

$$\begin{aligned}
\liminf_{n\to\infty} x_n &= \underleftarrow{\lim}_{n\to\infty} \left(\inf_{m\geq n} x_m \right) \\
&= \sup_{n\geq 0} \inf_{m\geq n} x_m \\
&= \sup\{\inf\{x_m \mid m \geq n\} \mid n \geq 0\}.
\end{aligned} \tag{1}$$

- **Upper Bound** : An *upper bound* of a subset S of a poset P is an element $b \in P$ such that $b \geq x$ for all $x \in S$.
- **Supremum** : An upper bound $u \in P$ is called a *supremum* of S if for all upper bounds z of S in P, $z \geq u$.

- **Limit Supremum** : For a sequence $\{x_n\}$, limit supremum is defined by:

$$\begin{aligned}\limsup_{n\to\infty} x_n &= \varinjlim_{n\to\infty}\left(\sup_{m\geq n} x_m\right)\\ &= \inf_{n\geq 0}\sup_{m\geq n} x_m \\ &= \inf\{\sup\{x_m \mid m\geq n\} \mid n\geq 0\}\end{aligned} \tag{2}$$

- **Limit** : Consider the sequence $\{x_n\}$ in $[-\infty, +\infty]$, then $\varinjlim_n x_n$ is defined as

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n := \varinjlim_{n\to\infty} x_n.$$

- **Lower Sum** : $l(f,\mathcal{P})$ is the sum of the minimum functional values at the partition. That is,

$$l(f,\mathcal{P}) = \sum_{i=0}^{n-1} m_i(a_{i+1}-a_i)$$

  where $m_i = \inf\{f(x) \mid x\in[a_{i-1},a_i]\}$.
- **Upper Sum** : Similarly,

$$u(f,\mathcal{P}) = \sum_{i=0}^{n-1} M_i(a_{i+1}-a_i)$$

  where $M_i = \sup\{f(x) \mid x\in[a_{i-1},a_i]\}$.
  Remember that the function is *Riemann Integrable* if $l(f,\mathcal{P}) = u(f,\mathcal{P})$.
- **Countable Sets** : Note the following,
  1. *Cardinality* : Sets X and Y have the same cardinality if there exists a bijection from X to Y.
  2. *Finite Set* : A set is finite if it is empty or it has the same cardinality as $\{1,2,\ldots,n\}$ for some $n\in\mathbb{N}$.
  3. *Countably Infinite* : If the set has the same cardinality as $\mathbb{N}$.
  4. *Enumeration* : An enumeration of a countably infinite set X is a bijection of $\mathbb{N}$ onto X. That is, an enumeration is an infinite sequence $\{x_n\}$ such that each of the x_i's are in X and each element of X is x_i for some i.
  5. *Countable* : A set is countable if it is finite or countably infinite. For example, $\mathbb{N}$ is countable, $\mathbb{Q}$ is also countable (!), $\mathbb{R}\setminus\mathbb{Q}$ (irrationals) is not countable, $\mathbb{R}$ is not countable.
- **Totally Bounded** : A subset $B\subseteq X$ is *totally bounded* when it can be covered by a finite number of r-balls for all $r>0$. That is,

$$\forall\, r>0,\ \exists N\in\mathbb{N},\ \exists a_1,\ldots,a_N\in X \text{ such that } B\subseteq\bigcup_{n=1}^{N} B_r(a_n)$$

- **Compact Set** : A set K is said to be *compact* when given any cover of balls of possibly unequal radii, there is a finite sub-collection of them that still covers the set K. That is,

$$K\subseteq\bigcup_i B_{r_i}(a_i) \implies \exists\, i_1,\ldots,i_N,\ \ K\subseteq\bigcup_{n=1}^{N} B_{r_{i_n}}(a_{i_n})$$

  Note that *compact metric spaces are totally bounded (!)*. Also, *compact sets are closed.*

The problem begins with Riemann Integrable functions when we see that functions like Dirichlet function (1 on irrational and 0 on rational points) can become *measurable* even when the function is not continuous! This motivates the need of a formal notion of a *measure.*

We begin with some recollections from classical analysis of one real variable.

1. Every open set in $\mathbb{R}$ can be written as disjoint union of open intervals.

   Proof. Let $G \subseteq \mathbb{R}$ be a open subset. Now by definition of an open subset, we have that for any $x \in G$, there exists atleast one open subset U such that $x \in U \subseteq G$. Now consider the following union of all such open subsets of x,

   $$U_x = \bigcup_{x \in U \subseteq G} U$$

   It's now easy to see that U_x is the largest such subset of G, as any other $V \subseteq G$ such that $x \in V$ is by definition contained in U_x. Moreover, U_x is an interval as it is an arbitrary union of open intervals. Now, define the following relation on G:

   $$y \sim x \iff y \in U_x$$

   Now we clearly have that $x \in U_x$ (reflexive); for $y \sim U_x$ we have $U \subseteq U_x$ such that $x, y \in U$, hence $x \in U_y$ (symmetric); for $x \in U_y$ and $y \in U_z$, we have that $x, y, z \in U_y$, since $z \in U_y \subseteq G$ so $U_y \subseteq U_z$, so $x \in U_z$ (transitive). Hence $\sim$ is an equivalence relation, hence $\sim$ partitions the set G. Denote the set of all equivalence classes as $\mathcal{I}$ so we get

   $$G = \bigcup_{I \in \mathcal{I}} I$$

   such that $I_1 \cap I_2 = \Phi$ for any $I_1, I_2 \in \mathcal{I}$. Now note that for any $I \in \mathcal{I}$ is open because each I is generated by the relation $\sim$ such that $y \sim x$ iff $y \in U_x$. Hence for any $z \in I$, we have $z \in U_x \subseteq G$ where U_x is open. Therefore, we have $G = \cup_{I \in \mathcal{I}} I$ for disjoint open intervals in $\mathcal{I}$. ☐

2. Prove that every non-empty perfect subset of $\mathbb{R}$ (or $\mathbb{R}^n$) is uncountable. That is, if $A = A'$ then A is uncountable.

   Proof. Take $A \subseteq \mathbb{R}$ to be a perfect subset. Since A it is perfect, therefore, it must contain all of it's limit points or, equivalently, contains no isolated points. Clearly, then, A cannot be finite, but can only be countably infinite or uncountable. If it is uncountable, then the proof is over. If A is countably infinite, then we can write A as the following :

   $$A = \{a_1, a_2, \dots\}.$$

   Construct a ball around a_{i_1} of any radius $r_1 > 0$. Since A is perfect, therefore $\exists\ a_{i_2} \in B_{r_1}(a_{i_1}) \cap A = C_1$. Similarly, for some $r_2 > 0$, we have $a_{i_3} \in B_{r_2}(a_{i_2}) \cap B_{r_1}(a_{i_1}) \cap A = C_2$ such that $a_{i_1} \notin C_2$ and so on. In general, we would have the following,

   $$a_{i_{n+1}} \in \left(\bigcap_{j=1}^{n} B_{r_j}(a_{i_j}) \right) \cap A = C_n.$$

   Now, consider $C = \cap_n C_n$. Since $C_{n+1} \subseteq C_n$, therefore $C \neq \Phi$. But, $a_i \notin C$ for any $i \in \mathbb{N}$ as $a_i \notin C_{i+1}$. Therefore we have a contradiction. Hence A cannot by countably infinite, it must only be uncountable. ☐

3. In the definition of Lebesgue Outer measure on $\mathbb{R}$, one can instead take $\mathcal{C}_A$ to be collection of infinite sequences of the any form from $\{[a_n, b_n]\}, \{(a_n, b_n)\}$ or $\{(a_n, b_n]\}$.

   Proof. *Refer Proof of Proposition 2.5.3.* □

4. Show the following:
$$\bigcup_{n=1}^{N} E_n = \bigcup_{n=1}^{N} \left(E_n \cap \left(\bigcup_{k<n} E_k \right)^{c} \right)$$

   Proof. Take $x \in \bigcup_{n=1}^{N} E_n$. Then $\exists\ E_k$ for some a such that $x \in E_a$. Now, clearly, $x \in E_a \subseteq (\bigcup_{k<a} E_k)^c$, hence $x \in (E_a \cap (\bigcup_{k<a} E_k)^c)$. Hence, we have $\bigcup_{n=1}^{N} E_n \subseteq \bigcup_{n=1}^{N} (E_n \cap (\bigcup_{k<n} E_k)^c)$. The converse is easy to see too. □

2 Measures

2.1 Algebras & σ-algebras

Definition 2.1.1. (**Algebra/Field**) Let X be an arbitrary set. A collection $\mathcal{A} \subseteq P(X)$ of subsets of X is an algebra on X if:

- $X \in \mathcal{A}$.
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$.
- For each finite sequence $A_1, A_2, \ldots, A_n \in \mathcal{A}$ implies that
$$\bigcup_{i=1}^{n} A_i \in \mathcal{A}$$
- For each finite sequence $A_1, A_2, \ldots, A_n \in \mathcal{A}$ implies that
$$\bigcap_{i=1}^{n} A_i \in \mathcal{A}$$

Definition 2.1.2. (**σ-Algebra/σ-Field**) Let X be an arbitrary set. A collection $\mathcal{A} \subseteq P(X)$ of subsets of X is a σ-algebra on X if:

- $X \in \mathcal{A}$.
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$.
- For each infinite sequence $\{A_i\}$ such that $A_i \in \mathcal{A}$, it implies that
$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$$
- For each infinite sequence $\{A_i\}$ such that $A_i \in \mathcal{A}$, it implies that
$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$$

Proposition 2.1.3. *Let X be a set. Then the intersection of an arbitrary non-empty collection of σ-algebras on X is a σ-algebra on X.*

Proof. Consider a collection $\mathcal{C}$ of σ-algebras on X. Denote $\mathcal{A} = \bigcap \mathcal{C}$ as intersection of all σ-algebras in $\mathcal{C}$. We can now easily see that any subset in $\mathcal{A}$ would be present in every σ-algebra present in collection $\mathcal{C}$, hence, it would obey all properties of a σ-algebras. Therefore, $\mathcal{A}$ is a σ-algebra. □

Corollary 2.1.4. *Let X be a set and let $\mathcal{F} \subseteq P(X)$ be a family of subsets of X. Then there exists a smallest σ-algebra on X that includes $\mathcal{F}$.*

Proof. Consider any given family $\mathcal{F} \subseteq P(X)$ and just take intersection of the family $\mathcal{C}$ of all σ-algebras which contains $\mathcal{F}$ to construct this smallest σ-algebra. □

Definition 2.1.5. (**Generated σ-algebra**) The smallest σ-algebra on X containing a given family $\mathcal{F} \subseteq P(X)$ of subsets is called the σ-algebra *generated* by $\mathcal{F}$, denoted as $\sigma(\mathcal{F})$.

Definition 2.1.6. (**Borel σ-algebra on $\mathbb{R}^d$**) It is the σ-algebra on $\mathbb{R}^d$ generated by the collection of all open subsets of $\mathbb{R}^d$, denoted as $\mathcal{B}(\mathbb{R}^d)$.

Definition 2.1.7. (**Borel Subsets of $\mathbb{R}^d$**) Any $A \subseteq \mathbb{R}^d$ is called a Borel subset of $\mathbb{R}^d$ if $A \in \mathcal{B}(\mathbb{R}^d)$.

Proposition 2.1.8. *The Borel σ-algebra on $\mathbb{R}$, $\mathcal{B}(\mathbb{R})$, of Borel subsets of $\mathbb{R}$ is generated by each of the following collection of sets:*

1. *The collection of all closed subsets of $\mathbb{R}$.*
2. *The collection of all subintervals of $\mathbb{R}$ of the form $(-\infty, b]$.*
3. *The collection of all subintervals of $\mathbb{R}$ of the form $(a, b]$.*

Proof. To show all of these, consider the three σ-algebras $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ corresponding to conditions 1,2 & 3 respectively and try to prove $\mathcal{A}_3 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_1 \subseteq \mathcal{B}(\mathbb{R})$ together with $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_3$. The first three inclusions are trivial to see. For the case that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_3$, simply note that any open subset can be made by unions of the sets of form $(a, b]$ and by Homework-I,1, each open set is union of open subsets. □

Proposition 2.1.9. *The σ-algebra $\mathcal{B}(\mathbb{R}^d)$ of Borel subsets of $\mathbb{R}^d$ is generated by each of the following collections:*

1. *The collection of all closed subsets of $\mathbb{R}^d$.*
2. *The collection of all closed half-spaces in $\mathbb{R}^d$ that have the form $\{(x_1, \ldots, x_d) \mid x_i \leq b\}$ for some index i and some $b \in \mathbb{R}$.*
3. *The collection of all rectangles in $\mathbb{R}^d$ that have the form*

$$\{(x_1, \ldots, x_d) \mid a_i < x_i \leq b_i \text{ for } i = 1, \ldots, d\}$$

Proof. Almost the same as in Proposition 2.1.8. $\mathcal{A}_1 \subseteq \mathcal{B}(\mathbb{R}^d)$ trivially by definition. $\mathcal{A}_2 \subseteq \mathcal{A}_1$ as $\{(x_1, \ldots, x_d) \mid x_i \leq b\}$ is closed itself. $\mathcal{A}_3 \subseteq \mathcal{A}_2$ by the observation that $\{(x_1, \ldots, x_d) \mid a_i < x_i \leq b_i\}$ is made by the difference of two subsets of the form $\{(x_1, \ldots, x_d) \mid x_i \leq b_i\}$ and $\{(x_1, \ldots, x_d) \mid x_i > a_i\}$, the latter is the complement of a certain subset in $\mathcal{A}_2$, moreover, $\{(x_1, \ldots, x_d) \mid a_i < x_i \leq b_i \text{ for } i = 1, \ldots, d\}$ is then constructed by intersection of d such subsets. Finally, $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{A}_3$ can be seen via the fact that open subsets in $\mathbb{R}^d$ are made by union of rectangles of type 3 and as such, they are called open subsets. □

Lemma 2.1.10. *Let X be a set and $S \subseteq P(X)$ a class of subsets of X. Let $A \subseteq X$ be a subset. Denote by $S \cap A = \{B \cap A \mid B \in S\}$. Then,*

$$\sigma_A(S \cap A) = \sigma(S) \cap A.$$

where $\sigma_A(S \cap A)$ denotes the smallest σ-algebra over A generated by the class $S \cap A \subseteq P(A)$.

Proof. It is easy to see that $\sigma_A(S\cap A) \hookrightarrow \sigma(S)\cap A$ by considering that $S\cap A \subseteq \sigma(S)\cap A$. Conversely, we use the generating set principle. That is, since we wish to show that for any $B \in \sigma(S)$, we have $B \cap A \in \sigma_A(S \cap A)$, therefore we define

$$\mathcal{S} := \{B \in \sigma(S) \mid B \cap A \in \sigma_A(S \cap A)\}$$

and then observe quite easily that $\mathcal{S}$ is a σ-algebra over X inside $\sigma(S)$ containing S. Thus $\mathcal{S} = \sigma(S)$, as needed. □

The following are some conditions for an algebra to become a σ-algebra.

Proposition 2.1.11. *Let X be a set and let $\mathcal{A}$ be an algebra on X. Then, $\mathcal{A}$ is a σ-algebra on X if* ***either***

- *$\mathcal{A}$ is closed under the formation of* ***unions*** *of* increasing *sequence of sets, or,*
- *$\mathcal{A}$ is closed under the formation of* ***intersections*** *of* decreasing *sequence of sets.*

Proof. Take any countably infinite collection of subsets $A_1, A_2, \dots \in \mathcal{A}$ where $\mathcal{A}$ is an algebra. Due to the definition of an algebra, we have that $C_n = \bigcup_{i=1}^{n} A_i \in \mathcal{A}$ for any $n \geq 1 \in \mathbb{Z}_+$. Now note that $C_1 \subseteq C_2 \subseteq \dots$, that is, the sequence $\{C_n\}$ forms an increasing sequence of sets. Hence, by the requirement of the question, we have that $\bigcup_{i=1}^{\infty} C_i \in \mathcal{A}$. But then we also have that $\bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i=1}^{\infty} C_i \in \mathcal{A}$. Hence we have the required condition for part 1. For part 2, we can see that $C_1^c \supseteq C_2^c \supseteq \dots$ is a decreasing sequence of sets. Then we must have, by the requirement of the question, that $\bigcap_{i=1}^{\infty} C_i^c = \left(\bigcup_{i=1}^{\infty} C_i\right)^c \in \mathcal{A}$. But then by definition of algebra, we must have $\bigcup_{i=1}^{\infty} C_i \in \mathcal{A}$, which already contains the countably infinite union $\bigcup_{i=1}^{\infty} A_i$. □

The following are some finiteness conditions we would like to have on measure spaces.

Definition 2.1.12. (**Finiteness conditions**) Let $(X, \mathcal{A}, \mu)$ be a measure space. Then,

1. X is said to be *finite* if $\mu(X) < \infty$,
2. X is said to be *σ-finite* if there exists $\{A_n\} \subseteq \mathcal{A}$ such that $\bigcup_n A_n = X$ and $\mu(A_n) < \infty$,
3. X is said to be *semi-finite* if for all $A \in \mathcal{A}$ such that $\mu(A) = \infty$, there exists $B \subseteq A$ such that $B \in \mathcal{A}$ and $\mu(B) < \infty$.

2.1.1 X-indexed $\mathbb{R}$-series

We would now like to make sense of the sum $\sum_{x\in X} f(x)$ where $f : X \to [0, \infty]$ is an arbitrary function.

Definition 2.1.13. (X**-indexed** $\mathbb{R}$**-series**) Let $f : X \to [0, \infty]$ be a function where X is a set. We define the series $\sum_{x\in X} f(x)$ as follows:

$$\sum_{x\in X} f(x) = \sup\left\{\sum_{x\in F} f(x) \mid F \subseteq X \text{ is finite}\right\}.$$

The following are some basic properties of X-indexed $\mathbb{R}$-series.

Proposition 2.1.14. *Let X be a set and $f : X \to [0,\infty]$ be a function. Denote $S = \{x \in X \mid f(x) > 0\}$.*

1. *If S is uncountable then $\sum_{x\in X} f(x) = \infty$.*
2. *If S is countably finite then for any bijection $\varphi : \mathbb{N} \to S$, we have*

$$\sum_{x\in X} f(x) = \sum_{n\in\mathbb{N}} f(\varphi(n)).$$

Proof. 1. Write $S = \bigcup_n S_n$ where $S_n = \{f(x) > 1/n\}$. Note that S_n forms an increasing sequence of sets. As S is uncountable, there exists $N \in \mathbb{N}$ such that S_N is uncountable. Consequently, for any finite set $F \subseteq S_N$, we have $\sum_{x\in F} f(x) \geq \frac{|F|}{N}$. As $\sum_{x\in F} f(x) \leq \sum_{x\in X} f(x)$, therefore

$$\frac{|F|}{N} \leq \sum_{x\in X} f(x). \tag{$\heartsuit$}$$

As $F \subseteq S_N$ is arbitrary finite set and S_N is uncountable, therefore we get the desired result.

2. Pick any bijection $\varphi : \mathbb{N} \to S$ and pick a finite set $F \subseteq X$. We have $\sum_{x\in F} f(x) = \sum_{x\in F\cap S} f(x)$, so replace $F \subseteq X$ by a finite set $F \subseteq S$. Let $n \in \mathbb{N}$ be large enough so that $\varphi(\{1,\ldots,n\}) \supseteq F$. Consequently, we have

$$\sum_{x\in F} f(x) \leq \sum_{k=1}^{n} f(\varphi(k)) \leq \sum_{x\in X} f(x). \tag{$\heartsuit$}$$

Take $n \to \infty$ in the above inequality to obtain

$$\sum_{x\in F} f(x) \leq \sum_{k=1}^{\infty} f(\varphi(k)) \leq \sum_{x\in X} f(x).$$

Take sup over all finite subsets F of X in the above inequality to obtain

$$\sum_{x\in X} f(x) \leq \sum_{n\in\mathbb{N}} f(\varphi(n)) \leq \sum_{x\in X} f(x),$$

which yields the desired result. □

2.2 Measures

Definition 2.2.1. (**Countably additive function**) Let X be a set and $\mathcal{A}$ be a σ-algebra on X. Function $\mu : \mathcal{A} \longrightarrow [0,+\infty]$ is said to be *countably additive* **if** it satisfies:

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

for **each** infinite sequence $\{A_i\}$ of **disjoint** sets in $\mathcal{A}$.

Definition 2.2.2. (**Measure**) A *measure* on $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ that is **countably additive** and satisfies:

$$\mu(\Phi) = 0.$$

Remark 2.2.3. This is sometimes also referred as *countably additive measure* on $\mathcal{A}$.

Definition 2.2.4. We have following definitions to compactly represent above definitions:

1. (***Measure space***) If X is a set, $\mathcal{A}$ is a σ-algebra on X and if μ is a measure on $\mathcal{A}$, then the triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.
2. (***Measurable Space***) If X is a set and $\mathcal{A}$ is a σ-algebra on X, then the pair $(X, \mathcal{A})$ is called a *measurable space*.

Proposition 2.2.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $A, B \in \mathcal{A}$ such that $A \subseteq B$. Then,*

- *We have $\mu(A) \leq \mu(B)$.*
- *Additionally, if A satisfies that $\mu(A) < +\infty$, then:*

$$\mu(B - A) = \mu(B) - \mu(A).$$

Proof. Note that A and $B \cap A^c$ are disjoint sets in the sigma algebra $\mathcal{A}$. Hence we can write, by countably additive property of μ, that:

$$\begin{aligned} \mu(A \cup (B \cap A^c)) &= \mu(B) \\ &= \mu(A) + \mu(B \cap A^c) \end{aligned}$$

Since $\mu(B \cap A^c) \geq 0$, hence $\mu(A) \leq \mu(B)$. Moreover, if $\mu(A) < \infty$, then we can additionally write $\mu(B \cap A^c) = \mu(B) - \mu(A)$. □

Definition 2.2.6. Let μ be a measure on a measurable space $(X, \mathcal{A})$. Then,

- (***Finite measure***) If $\mu(X) < +\infty$.
- (***σ-Finite measure***) If $X = \bigcup_i A_i$ where $A_i \in \mathcal{A}$ such that $\mu(A_i) < +\infty$ for all $i \in \mathbb{N}$.

Remark 2.2.7. In other words, a subset $A \in \mathcal{A}$ is *σ-finite* if it is a union of a countable sequence of sets that are in $\mathcal{A}$ and are of finite measure under μ.

2.2.1 Elementary properties of measures

Proposition 2.2.8. *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $\{A_k\}$ is an* ***arbitrary*** *sequence of sets that belong to $\mathcal{A}$, then,*

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

Proof. Denote $B_1 = A_1$ and $B_i = A_i \cap \left(\bigcup_{k=1}^{i-1} A_k\right)^c$. Note that B_i and B_j are disjoint for distinct i and j. Since $\{A_k\} \in \mathcal{A}$, therefore $\{B_i\} \in \mathcal{A}$. Moreover, $\bigcup_{i=1}^{\infty} B_i = \bigcup_{k=1}^{\infty} A_k$ by construction. We then get,

$$\begin{aligned} \mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) &= \sum_{i=1}^{\infty} \mu(B_i) \\ &\leq \sum_{i=1}^{\infty} \mu(A_i) \quad (\because\ B_i \subseteq A_i \text{ by construction.}) \end{aligned}$$

Hence proved. □

2.3 Basic results on measure spaces

We have the following first result.

Proposition 2.3.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

1. *If $A, B \in \mathcal{A}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*
2. *If $A, B \in \mathcal{A}$ and $A \subseteq B$ where $\mu(A) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.*
3. *For any sequence $\{A_n\} \subseteq \mathcal{A}$, we have*
$$\mu\left(\bigcup_n A_n\right) \leq \sum_n \mu(A_n).$$
4. *If $\{A_n\} \subseteq \mathcal{A}$ is an increasing sequence of measurable sets, then*
$$\mu\left(\bigcup_n A_n\right) = \lim_n \mu(A_n).$$
5. *If $\{A_n\} \subseteq \mathcal{A}$ is a decreasing sequence of measurable sets where $\mu(A_1) < \infty$, then*
$$\mu\left(\bigcap_n A_n\right) = \lim_n \mu(A_n).$$
6. *If X is σ-finite, then X is semi-finite.*

Proof. Statements 1. and 2. are immediate from the disjoint decomposition $B = A \amalg (B \setminus A)$. For 3. note that for any $\{A_n\} \subseteq \mathcal{A}$, we can form a disjoint sequence $\{B_n\} \subseteq \mathcal{A}$ such that $\bigcup_n A_n = \coprod_n B_n$ and $\mu(B_n) \leq \mu(A_n)$. Statement 4. also follows from similar reasons, where we can now let $B_n = A_n \setminus A_{n-1}$. Let us do statement 5. in some detail.

Observe that the sequence $C_1 = \emptyset$ and $C_n = A_1 \setminus A_n$ is an increasing sequence of sets. Thus, we have by statement 4. that

$$\mu\left(\bigcup_n C_n\right) = \lim_n \mu\left(C_n\right). \tag{$\heartsuit$}$$

We can write $A_1 = (A_1 \setminus A_n) \amalg A_n$. Using statement 2. we obtain that

$$\begin{aligned} \mu(A_1) &= \mu(C_n) + \mu(A_n) \\ \mu(A_1) - \mu(A_n) &= \mu(C_n). \end{aligned} \tag{$\heartsuit\heartsuit$}$$

We now claim that $\bigcap_n A_n = A_1 \setminus \bigcup_n C_n$. Indeed, for $x \in \bigcap_n A_n$, $x \in A_n \subseteq A_1$ for all n and thus $x \in A_1$. But if $x \in C_n$ for some n, then $x \notin A_n$, consequently a contradiction. Hence $x \in A_1 \setminus \bigcup_n C_n$. Conversely, for $x \in A_1 \setminus \bigcup_n C_n$ and any $n \in \mathbb{N}$, we have that if $x \notin A_n$, then $x \in A_1 \setminus A_n = C_n$, a contradiction. Hence the claim is proved.

As each $C_n \subseteq A_1$, thus $\bigcup_n C_n \subseteq A_1$. Consequently, by statement 2. and above claim we obtain that

$$\begin{aligned} \mu\left(\bigcap A_n\right) &= \mu(A_1) - \mu\left(\bigcup_n C_n\right) \\ &= \mu(A_1) - \lim_n \mu(C_n) \\ &= \mu(A_1) - \lim_n \left(\mu(A_1) - \mu(A_n)\right) \\ &= \lim_n \mu(A_n). \end{aligned}$$

This proves statement 5.

For statement 6. pick any $A \in \mathcal{A}$ with $\mu(A) = \infty$. We wish to construct a subset $B \subseteq A$ with $B \in \mathcal{A}$ and $0 < \mu(B) < \infty$. Let $\{D_n\} \subseteq \mathcal{A}$ be a collection of finite measure sets such that $\bigcup_n D_n = X$. Note that we can assume D_n are disjoint by suitably replacing D_n by $D_n \setminus D_1 \cup \dots \cup D_{n-1}$. Assume to the contrary, so that for each $B \subseteq A$ with $B \in \mathcal{A}$, either $\mu(B) = 0$ or $\mu(B) = \infty$. Let $D_n \cap A$ be such that $D_n \cap A \neq \emptyset$. Consequently, $\mu(D_n \cap A) = 0$ or ∞. The latter isn't possible, therefore $\mu(D_n \cap A) = 0$ for all $n \in \mathbb{N}$.

Since we have $A = \coprod_n D_n \cap A$, therefore $\mu(A) = \sum_n \mu(D_n \cap A) = 0$, a contradiction to the fact that $\mu(A) = \infty$. $\square$

We now cover an important example of a measure.

Construction 2.3.2. (*Measures from a positive function*) Let $(X, \mathcal{A})$ be a measurable space and $f : X \to [0, \infty]$ be a function. We construct the following map

$$\{\text{All functions } X \to [0,\infty]\} \longrightarrow \{\text{measures on } (X, \mathcal{A})\}.$$

Indeed, define

$$\begin{aligned} \mu_f : \mathcal{A} &\longrightarrow [0,\infty] \\ A &\longmapsto \sum_{x\in A} f(x). \end{aligned}$$

We claim that μ_f forms a measure.

It is clear that $\mu_f(\emptyset) = 0$. Consequently we need to show that for a disjoint collection $\{A_n\} \subseteq \mathcal{A}$, we have

$$\mu_f\left(\coprod_n A_n\right) = \sum_n \mu_f(A_n).$$

We first have that

$$\mu_f\left(\coprod_n A_n\right) = \sup\left\{\sum_{x\in F} f(x) \mid F \subseteq \coprod_n A_n \text{ is finite}\right\} \tag{1}$$

and

$$\sum_n \mu_f(A_n) = \sum_n \sup\left\{\sum_{x\in G} f(x) \mid G \subseteq A_n \text{ is finite}\right\}. \tag{2}$$

We first show that (1) $\leq$ (2). We need only show that for a finite set $F \subseteq \amalg_n A_n$, we have $\sum_{x\in F} f(x) \leq$ (2). Indeed, as $F_n := F \cap A_n$ is a collection of disjoint finite set where $F_n \subseteq A_n$ and only for finitely many n is F_n non-empty, therefore $\sum_{x\in F} f(x) = \sum_n \sum_{x\in F_n} f(x) \leq$ (2).

Conversely, we now wish to show that (2) $\leq$ (1). We use a standard technique for this. Pick any $\epsilon > 0$. For each $n \in \mathbb{N}$, we obtain a finite set $G_n \subseteq A_n$ such that

$$\mu_f(A_n) - \frac{\epsilon}{2^n} \leq \sum_{x\in G_n} f(x). \tag{$\heartsuit$}$$

Summing this till $N \in \mathbb{N}$, we obtain

$$\sum_{n=1}^{N} \left(\mu_f(A_n) - \frac{\epsilon}{2^n}\right) \leq \sum_{n=1}^{N} \sum_{x \in G_n} f(x) = \sum_{x \in \amalg_{n=1}^{N} G_n} f(x) \leq (1).$$

Now take $N \to \infty$ and $\epsilon \to 0$ to obtain the result[2].

Observe that the map defined above in Construction 2.3.2 is neither injective nor surjective, and that's good, otherwise measure theory would have been redundant. We now study completions of a measure space.

Remark 2.3.3. The goal of next few sections is to establish a good measure on $\mathbb{R}^n$ through which we can proceed to a theory of integration of measurable functions. Indeed, this goal was achieved by Lebesgue and he constructed what will be called the Lebesgue measure on $\mathbb{R}^n$. Hence, one should view the goal of the next few sections as to construct this measure space $(\mathbb{R}^n, \mathcal{M}, m)$, which is highly usable (as we will see in the integration theory) and is the gold standard of modern analysis.

2.4 Completion of a measure space

Definition 2.4.1. (**Null sets and complete measure spaces**) Let $(X, \mathcal{A}, \mu)$ be a measure space. A null set is an element $A \in \mathcal{A}$ such that $\mu(A) = 0$. The collection of all null sets is written as $\mathrm{Null}(\mathcal{A}) \subseteq \mathcal{A}$. A measure space $(X, \mathcal{A}, \mu)$ is said to be complete if for all $A \in \mathrm{Null}(\mathcal{A})$, $\mathcal{P}(A) \subseteq \mathcal{A}$.

Remark 2.4.2. Note that for a measure space $(X, \mathcal{A}, \mu)$, the collection of all null sets $\mathrm{Null}(\mathcal{A})$ contains $\emptyset$ and is closed under countable union. Indeed, for $\{A_n\} \subseteq \mathrm{Null}(\mathcal{A})$, we have $\mu(\cup_n A_n) \leq \sum_n \mu(A_n) = 0$ by Proposition 2.3.1, 3.

Definition 2.4.3. (**Extension of measure spaces**) Let $(X, \mathcal{A}, \mu)$ and $(X, \mathcal{A}', \mu')$ be two measure spaces. Then we say that $(X, \mathcal{A}', \mu')$ is an extension of $(X, \mathcal{A}, \mu)$ if $\mathcal{A}' \supseteq \mathcal{A}$ and $\mu'|_{\mathcal{A}} = \mu$.

We will now for each measure space $(X, \mathcal{A}, \mu)$ will construct an extension of it which will be complete.

Construction 2.4.4. Let $(X, \mathcal{A}, \mu)$ be a measure space. Consider the following collection

$$\hat{\mathcal{A}} := \{A \cup B \mid A \in \mathcal{A}, B \subseteq N, N \in \mathrm{Null}(\mathcal{A})\}.$$

Define $\hat{\mu} : \hat{\mathcal{A}} \to [0, \infty]$ as $A \cup B \mapsto \mu(A)$.

Theorem 2.4.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then, $(X, \hat{\mathcal{A}}, \hat{\mu})$ is a complete measure space extending $(X, \mathcal{A}, \mu)$. We call it the completion of $(X, \mathcal{A}, \mu)$.*

Proof. We need to show the following things.

1. $\hat{\mathcal{A}}$ is a σ-algebra,
2. $\hat{\mu}$ is a measure,
3. $\hat{\mu}|_{\mathcal{A}} = \mu$,
4. $(X, \hat{\mathcal{A}}, \hat{\mu})$ is complete.

[2] We call this the *ϵ-wiggle around* inf *and* sup *technique.*

The first three are straightforward. We show 4. in some detail.

Pick $A \cup B \in \hat{\mathcal{A}}$ such that $\hat{\mu}(A \cup B) = \mu(A) = 0$. Then $A \in \mathrm{Null}(\mathcal{A})$. Further, $B \subseteq N$ where $N \in \mathrm{Null}(\mathcal{A})$. Let $C \subseteq A \cup B$. Then $C = (C \cap A) \cup (C \cap B)$. Since $C \cap A \subseteq A$ and $C \cap B \subseteq N$, therefore $C \subseteq A \cup N$ where $A \cup N \in \mathrm{Null}(\mathcal{A})$. Consequently, we may write $C = \emptyset \cup C$ where C is a subset of a null set. Hence $C \in \hat{\mathcal{A}}$. □

Example 2.4.6. Let $X = \{1, 2, 3\}$ and $\mathcal{A} = \{\emptyset, X, \{1\}, \{2, 3\}\}$. Define $\mu : \mathcal{A} \to [0, \infty]$ by $\mu(\emptyset) = 0 = \mu(\{2, 3\})$ and $\mu(\{1\}) = \mu(X)$. Clearly, $(X, \mathcal{A}, \mu)$ is a measure space which is not complete. We calculate its completion $(X, \hat{\mathcal{A}}, \hat{\mu})$. By Construction 2.4.4, as the only null set is $\{2, 3\}$, we have

$$\hat{\mathcal{A}} = \{\emptyset, X, \{1\}, \{2, 3\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}\}.$$

Hence $\hat{\mathcal{A}} = \mathcal{P}(X)$. Similarly, $\hat{\mu}$ is easy to find by the definition in Construction 2.4.4.

2.5 Outer measures

Definition 2.5.1. (**Outer measure**) Let X be a set and let $\mathcal{P}(X)$ be the collection of all subsets of X. An *outer measure* on X is a function $\mu^* : \mathcal{P}(X) \longrightarrow [0, +\infty]$ such that:

- For the empty set Φ,
$$\mu^*(\Phi) = 0$$
- If $A \subseteq B \subseteq X$, then
$$\mu^*(A) \leq \mu^*(B).$$
- If $\{A_n\}$ is an infinite sequence of subsets of X, then
$$\mu^*\left(\bigcup_n A_n\right) \leq \sum_n \mu^*(A_n)$$

Definition 2.5.2. (**Lebesgue outer measure on** $\mathbb{R}$) For each subset $A \subseteq \mathbb{R}$, let $\mathcal{C}_A$ be the *set of all infinite sequences* $\{(a_i, b_i)\}$ of *bounded open intervals* such that $A \subseteq \bigcup_i (a_i, b_i)$. That is,

$$\mathcal{C}_A = \{\{(a_i, b_i)\} \mid A \subseteq \cup_i (a_i, b_i) \text{ and } a_i, b_i \in \mathbb{R}\}$$

Then, $\lambda^* : \mathcal{P}(\mathbb{R}) \longrightarrow [0, +\infty]$ is the Lebesgue outer measure, defined by:

$$\lambda^*(A) = \inf\left\{\sum_i (b_i - a_i) \;\middle|\; \{(a_i, b_i)\} \in \mathcal{C}_A\right\} \tag{3}$$

To verify that λ^* is indeed an outer measure.

Proposition 2.5.3. *Lebesgue outer measure on $\mathbb{R}$ is an outer measure and it assigns to each subinterval of $\mathbb{R}$ it's length.*

Proof. Denote $\mathcal{C}_A = \{\{(a_i, b_i)\} \mid A \subseteq \cup_i (a_i, b_i)\}$. To show that λ^* is an outer measure, we first need to show that $\lambda^*(\Phi) = 0$. For that, consider the set of all infinite sequences $\{(a_i, b_i)\} \in \mathcal{C}_\Phi$, that is (trivially) $\Phi \subseteq \cup_i (a_i, b_i)$, such that $\sum_i (b_i - a_i) < \epsilon$ for all $\epsilon > 0$. Then, if we denote $\mathcal{L}_A = \{\sum_i (b_i - a_i) \mid \{(a_i, b_i)\} \in \mathcal{C}_A\}$, then $\inf \mathcal{L}_\Phi = 0$ as for any lower bound l of $\mathcal{L}_A$, if $l > 0$ then $\exists\, \{(a_i, b_i)\} \in \mathcal{C}_\Phi$ such that $\sum_i (b_i - a_i) < l$, hence $l \leq 0$, or $\inf \mathcal{L}_\Phi = 0$.

Second, we need to show that if $A \subseteq B \subseteq X$, then $\lambda^*(A) \leq \lambda^*(B)$. For this, consider $A \subseteq B$. Clearly, we have that $\mathcal{C}_B \subseteq \mathcal{C}_A$, therefore $\mathcal{L}_B \subseteq \mathcal{L}_A$ and hence $\inf \mathcal{L}_B \geq \inf \mathcal{L}_A$.
Third, we need to show that for any infinite sequence $\{A_n\}$ of subsets of X,

$$\lambda^*\left(\bigcup_n A_n\right) \leq \sum_n \lambda^*(A_n)$$

For this, consider the Lebesgue outer measure of A_n, that is, $\lambda^*(A_n)$. We must have, that for any infinite sequence $\{(a_{n,i}, b_{n,i})\} \in \mathcal{C}_{A_n}$, that

$$\sum_{i=1}^{\infty}(b_{n,i} - a_{n,i}) \geq \lambda^*(A_n).$$

Hence, consider that the difference is upper bounded according to n, that is the sequence $\{(a_{n,i}, b_{n,i})\} \in \mathcal{C}_{A_n}$ is such that,

$$\sum_{i=1}^{\infty}(b_{n,i} - a_{n,i}) - \lambda^*(A_n) \leq \epsilon/2^n.$$

Now, we can cover the entire $\bigcup_i A_i$ by the union of the above intervals, that is,

$$\bigcup_i A_i \subseteq \bigcup_n \bigcup_i (a_{n,i}, b_{n,i}).$$

Now, we know that

$$\lambda^*\left(\bigcup_i A_i\right) = \inf \mathcal{L}_{\cup_i A_i}.$$

But since

$$\sum_n \sum_i (b_{n,i} - a_{n,i}) \in \mathcal{L}_{\cup_i A_i},$$

and

$$\sum_n \left(\sum_i (b_{n,i} - a_{n,i}) - \lambda^*(A_n)\right) \leq \sum_n \epsilon/2^n$$

which is equal to

$$\sum_n \sum_i (b_{n,i} - a_{n,i}) - \sum_n \lambda^*(A_n) \leq \epsilon \times 1$$

or,

$$\sum_n \sum_i (b_{n,i} - a_{n,i}) \leq \sum_n \lambda^*(A_n) + \epsilon$$

and since $\lambda^*(\bigcup_i A_i) = \inf \mathcal{L}_{\bigcup_i A_i}$, therefore,

$$\lambda^*\left(\bigcup_i A_i\right) \leq \sum_n \sum_i (b_{n,i} - a_{n,i}) \leq \sum_n \lambda^*(A_n)$$

Hence proved.

Now, we need to show that λ^* assigns each subinterval it's length.
For this first show that $\lambda^*([a,b]) \leq b-a$. This is easy to show if we take,

$$[a,b] = \bigcup_i (a_i, b_i)$$

where $(a_1,b_1) = (a,b)$, $(a_i,b_i) = (a-\epsilon/2^i, a)$ for all even i and $(a_j,b_j) = (b, b+\epsilon/2^j)$ for all odd j.
Now,

$$\begin{aligned}\sum_i (b_i - a_i) &= (b-a) + \sum_{i=2,4,\dots} \epsilon/2^i + \sum_{i=3,5,\dots} \epsilon/2^i \\ &= b-a + \sum_{i=1,2,\dots} \epsilon/2^i \\ &= b-a+\epsilon\end{aligned}$$

therefore $\lambda^*([a,b]) = \inf \mathcal{L}_{[a,b]} \leq b-a+\epsilon$ for all $\epsilon > 0$, hence $\lambda^*([a,b]) \leq b-a$.
Now, to show the converse that $b-a \leq \lambda^*([a,b])$, we first note that $[a,b]$ is compact, so for any infinite cover $\{(a_i,b_i)\} \in \mathcal{C}_{[a,b]}$, there exists a finite subcover $\{(a_i,b_i)\}_{i=1}^n$ of $[a,b]$. Now, since λ^* is an outer measure, therefore,

$$b-a \leq \sum_{i=1}^{n} \lambda^*((a_i,b_i)) \leq \sum_{i=1}^{\infty} \lambda^*((a_i,b_i)) \in \mathcal{L}_{[a,b]}$$

Therefore, $b-a$ is a lower bound of $\mathcal{L}_{[a,b]}$ and hence $b-a \leq \inf \mathcal{L}_{[a,b]} = \lambda^*([a,b])$.
Hence $\lambda^*([a,b]) = b-a$.
Now since, one can construct subintervals of the form $(a,b]$ or $[a,b)$ from the following manner:

$$(a,b] \subseteq (a,b) \bigcup \left(\bigcup_n [b, b+\epsilon/2^n] \right)$$

from which we get that $\lambda^*((a,b]) \leq b-a$ and also,

$$[a,b] \subseteq (a,b] \bigcup \left(\bigcup_n [a-\epsilon/2^n, a] \right)$$

which yields $b-a \leq \lambda^*((a,b])$. Similarly for $(-\infty,b]$ to show that $\lambda^*((-\infty,b]) = +\infty$. □

Construction 2.5.4. (*Lebesgue outer measure on* $\mathbb{R}^n$) Consider $\mathbb{R}^m$ and for an box $I \subseteq \mathbb{R}^n$, by which we mean a product of interval $I = I_1 \times \cdots \times I_m$ for $I_i \subseteq \mathbb{R}$, denote $v(I)$ to be its volume; $v(I) = \prod_{i=1}^m l(I_i)$. For any $A \subseteq \mathbb{R}^n$, we define

$$\mu^*(A) = \inf \left\{ \sum_n v(I_n) \mid \bigcup_n I_n \supseteq A,\ I_n \text{ are boxes} \right\}.$$

We claim that μ^* forms an outer measure on $\mathbb{R}^n$.

Indeed, $\mu^*(\emptyset) = 0$ as $\emptyset \subseteq (-1/k, 1/k)^m$ for all $n \in \mathbb{N}$ so we have $\mu^*(A) \leq 2^m/n^m$. Taking $n \to \infty$ does the job.

Let $A \subseteq B$ in $\mathbb{R}^m$. Observe that to show $\mu^*(A) \leq \mu^*(B)$ we need only show that $\{\sum_n v(I_n) \mid \bigcup_n I_n \supseteq A,\ I_n$ are b

$\{\sum_n v(I_n) \mid \bigcup_n I_n \supseteq B,\ I_n \text{ are boxes}\}$. But this is trivial as and sequence of boxes $\{I_n\}$ covering B also covers A.

Finally we wish to show countable subadditivity. Pick $\{A_n\} \subseteq \mathcal{P}(\mathbb{R}^m)$. We wish to show that

$$\mu^*\left(\bigcup_n A_n\right) \leq \sum_n \mu^*(A_n).$$

We use the ϵ-wiggle around sup and inf technique to show this, as discussed earlier in Construction 2.3.2. Pick any $\epsilon > 0$ and observe that we have a sequence of boxes $\{I_{n,k}\}_k$ for each $n \in \mathbb{N}$ such that $\bigcup_k I_{n,k} \supseteq A_n$ and

$$\mu^*(A_n) + \frac{\epsilon}{2^n} \geq \sum_k v(I_{n,k}). \tag{$\heartsuit$}$$

Observe further that $\bigcup_n \bigcup_k I_{n,k} \supseteq \bigcup_n A_n$. Consequently, we have $\sum_n \sum_k v(I_{n,k}) \geq \mu^*(\bigcup_n A_n)$. Hence,

$$\sum_n \left(\mu^*(A_n) + \frac{\epsilon}{2^n}\right) \geq \sum_n \sum_k v(I_{n,k}) \geq \mu^*\left(\bigcup_n A_n\right).$$

Hence μ^* is an outer measure on $\mathbb{R}^n$.

Note that the only place we required knowledge about boxes explicitly was only to show that $\mu^*(\emptyset) = 0$. This motivates the following simple result

Theorem 2.5.5. *Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$ be a collection of sets containing $\emptyset$ and X. Let $l : \mathcal{S} \to [0, \infty]$ be a function such that $l(\emptyset) = 0$. Then μ^* defined by*

$$\begin{aligned} \mu^* : \mathcal{P}(X) &\longrightarrow [0, \infty] \\ A &\longmapsto \inf\left\{\sum_n l(I_n) \mid \bigcup_n I_n \supseteq A,\ I_n \in \mathcal{S}\right\} \end{aligned}$$

is an outer measure on X.

Proof. Verbatim to Construction 2.5.4, except that $\mu^*(\emptyset) = 0$ follows now by the assumption that $l(\emptyset) = 0$ and $\emptyset \in \mathcal{P}(X)$ so that $\emptyset$ forms its own covering. □

2.6 Lebesgue measurability & Carathéodory's theorem

Definition 2.6.1. (**μ^*-measurable subset**) Let X be a set and let μ^* be an *outer measure* on X. A subset $B \subseteq X$ is μ^*-measurable if:

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

holds **for all** subsets $A \subseteq X$.

Definition 2.6.2. (**Lebesgue measurable subset of $\mathbb{R}$**) A subset $B \subseteq \mathbb{R}$ is called a Lebesgue measurable subset of $\mathbb{R}$ if B **is λ^*-measurable.** That is, for **any** $A \subseteq \mathbb{R}$, we must have:

$$\lambda^*(A) = \lambda^*(A \cap B) + \lambda^*(A \cap B^c)$$

Remark 2.6.3. Important to note are the following:

- Due to sub-additivity of μ^* and $A \subseteq (A \cap B) \cup (A \cap B^c)$, we already have that

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

for *any* subsets $A, B \subseteq X$.

★ Due to the above fact, *all that remains to be shown to ascertain that $B \subseteq \mathbb{R}$ is μ^*-measurable* is to show the following converse:

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

for all $A \subseteq X$.

Proposition 2.6.4. *Let X be a set and let μ^* be an outer measure on X. Then each subset $B \subseteq X$ that satisfies $\mu^*(B) = 0$* ***or*** *that satisfies $\mu^*(B^c) = 0$ is μ^*-measurable.*

Proof. This result actually proves that for subset $B \subseteq X$ which has zero outer measure under μ^*, any other subset $A \subseteq X$ would be such that $\mu^*(A \cap B) = 0(!)$ After proving this, and from the remark above, we would just be left to show that if $\mu^*(B) = 0$, then $\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c)$. We show the former here, from which the latter follows naturally.

Consider $B \subseteq X$ such that $\mu^*(B) = 0$. It's true that $A \cap B \subseteq B$. Now since μ^* is an outer measure on X, therefore, we must have $\mu^*(A \cap B) \leq \mu^*(B) = 0$. This implies that $\mu^*(A \cap B) = 0$. Now, we would see that the required condition follows naturally from the previous. First, note the following:

$$A \cap B \subseteq A \text{ and } A \cap B^c \subseteq A.$$

Hence, we can write:

$$\mu^*(A \cap B) \leq \mu^*(A) \text{ and } \mu^*(A \cap B^c) \leq \mu^*(A).$$

Now if $\mu^*(B) = 0$, then $\mu^*(A \cap B) = 0$ and then in the second inequality, we would have:

$$\mu^*(A \cap B^c) + \mu^*(A \cap B) \leq \mu^*(A) + 0$$

Or, if $\mu^*(B^c) = 0$, then $\mu^*(A \cap B^c) = 0$ and then in the first inequality, we would have:

$$\mu^*(A \cap B) + \mu^*(A \cap B^c) \leq \mu^*(A) + 0.$$

Hence, B is μ^*-measurable for any $B \subseteq X$ which satisfies that either $\mu^*(B) = 0$ or $\mu^*(B^c) = 0$. □

The following theorem is a fundamental fact about outer measures.

Theorem 2.6.5 (Carathéodory). *Let X be a set, let μ^* be an outer measure on X and let $\mathcal{M}_{\mu^*}$ be the collection of all μ^*-measurable subsets of X. Then,*

- *$\mathcal{M}_{\mu^*}$ is a σ-algebra.*
- *The restriction of μ^* to $\mathcal{M}_{\mu^*}$ is a measure on $\mathcal{M}_{\mu^*}$.*

Proof. **Act 1.** *$\mathcal{M}_{\mu^*}$ is an algebra.*
First, it is clear that $X, \Phi \in \mathcal{M}_{\mu^*}$ from Proposition 2.6.4, because $\mu^*(\Phi) = \mu^*(X^c) = 0$. Now, if $B \in \mathcal{M}_{\mu^*}$, then $\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)\ \forall\ A \subseteq X$. But if we replace B by B^c in the above, we would get the same equation, hence $B^c \in \mathcal{M}_{\mu^*}$. So $\mathcal{M}_{\mu^*}$ is closed under complements. Now, to show closed nature under finite unions, we take any two subsets $B_1, B_2 \in \mathcal{M}_{\mu^*}$ and show that $A \cup B \in \mathcal{M}_{\mu^*}$. First we have

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c) \\ &= \mu^*(A \cap B_2) + \mu^*(A \cap B_2^c)\end{aligned}$$

for any $A \subseteq X$. Now, we see that from the fact that $B_1 \in \mathcal{M}_{\mu^*}$,

$$\begin{aligned}\mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^c) \\ &= \mu^*(A \cap B_1) + \mu^*(A \cap B_2 \cap B_1^c)\end{aligned}$$

Similarly, we have from the fact $B_2 \in \mathcal{M}_{\mu^*}$,

$$\begin{aligned}\mu^*(A \cap (B_1 \cup B_2)^c) &= \mu^*(A \cap (B_1 \cup B_2)^c \cap B_2) + \mu^*(A \cap (B_1 \cup B_2)^c \cap B_2^c) \\ &= \mu^*(A \cap B_1^c \cap B_2^c \cap B_2) + \mu^*(A \cap B_1^c \cap B_2^c \cap B_2^c) \\ &= \mu^*(\Phi) + \mu^*(A \cap B_1^c \cap B_2^c) \\ &= \mu^*(A \cap (B_1 \cup B_2)^c)\end{aligned}$$

Now, adding the above results yield,

$$\begin{aligned}\mu^*(A \cap (B_1 \cup B_2)^c) + \mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2)^c) + \mu^*(A \cap B_1) + \mu^*(A \cap B_2 \cap B_1^c) \\ &= \mu^*(A \cap B_1^c \cap B_2^c) + \mu^*(A \cap B_1^c \cap B_2) + \mu^*(A \cap B_1) \\ &= \mu^*(A \cap B_1^c) + \mu^*(A \cap B_1) \\ &= \mu^*(A).\end{aligned}$$

Hence, $B_1 \cup B_2$ is μ^*-measurable, so $B_1 \cup B_2 \in \mathcal{M}_{\mu^*}$. Now, we can, for a finite collection of subsets in $\mathcal{M}_{\mu^*}$, we can proceed like above, to show that $\mathcal{M}_{\mu^*}$ is closed under finite union, hence showing that $\mathcal{M}_{\mu^*}$ is an algebra.

Act 2. *$\mathcal{M}_{\mu^*}$ is a σ-algebra.*
All that is left to show that $\mathcal{M}_{\mu^*}$ is a σ-algebra is to show that it is closed under countable union. We have already proved closed nature under finite union. We extend it via induction principle. Suppose $\{B_i\}$ is a sequence of disjoint subsets in $\mathcal{M}_{\mu^*}$. For this, we first prove[3] using induction that, for all $A \subseteq X$ and $n \in \mathbb{N}$,

$$\text{To Prove}: \mu^*(A) = \sum_{i=1}^{n} \mu^*(A \cap B_i) + \mu^*\left(A \cap \left(\bigcap_{i=1}^{n} B_i^c\right)\right) \tag{4}$$

[3] But why to prove Eq. 4? The motivation for Eq. 4 comes from Part 1. More specifically, notice in the equation where we added $\mu^*(A \cap (B_1 \cup B_2)^c)$ and $\mu^*(A \cap (B_1 \cup B_2))$. Note it's 2^{nd} line, this is the case when $n = 2$ in Eq. 4 combined with the fact that B_i's are disjoint. Now why to take B_i's to be disjoint? The reason for this comes from the fact that for any infinite sequence of subsets $\{A_i\}$, one can construct infinite sequence of disjoint subsets, that is : $A_1, A_2 \cap A_1^c, A_3 \cap (A_1 \cup A_2)^c, \ldots$ and it's union is again $\bigcup_n A_n$. Hence if we prove that a disjoint infinite sequence is closed under union, then we could prove that any infinite sequence of subsets is closed under union too!

For the case when $n=1$, we see that it Eq. 4 reduces to $\mu^*(A) = \mu^*(A\cap B_1) + \mu^*(A \cap B_1^{\mathrm{c}})$. But since $B_i \in \mathcal{M}_{\mu^*}\ \forall\, i \in \mathbb{N}$, therefore this is trivially true. Now, by the induction principle, we assume that Eq. 4 is true uptill n and then we try to prove it for $n+1$ step. For this, since $B_{n+1} \in \mathcal{M}_{\mu^*}$ is disjoint to all other B_i's, we have,

$$
\begin{aligned}
\mu^*\left(A \cap \bigcap_{i=1}^{n} B_i^{\mathrm{c}}\right) &= \mu^*\left(\left(A \cap \bigcap_{i=1}^{n} B_i^{\mathrm{c}}\right) \cap B_{n+1}\right) + \mu^*\left(\left(A \cap \bigcap_{i=1}^{n} B_i^{\mathrm{c}}\right) \cap B_{n+1}^{\mathrm{c}}\right) \\
&= \mu^*(A \cap B_{n+1}) + \mu^*\left(A \cap \bigcap_{i=1}^{n+1} B_i^{\mathrm{c}}\right)
\end{aligned}
$$

where the last line follows from the fact that each B_i is disjoint to other B_j's, hence each B_j^{c} would contain B_i and therefore $B_{n+1} \subseteq \bigcap_{i=1}^{n} B_i^{\mathrm{c}}$. Now, substituting the above equation in Eq. 4 gives,

$$
\begin{aligned}
\mu^*(A) &= \sum_{i=1}^{n} \mu^*(A \cap B_i) + \mu^*(A \cap B_{n+1}) + \mu^*\left(A \cap \bigcap_{i=1}^{n+1} B_i^{\mathrm{c}}\right) \\
&= \sum_{i=1}^{n+1} \mu^*(A \cap B_i) + \mu^*\left(A \cap \bigcap_{i=1}^{n+1} B_i^{\mathrm{c}}\right)
\end{aligned}
$$

Hence, by induction principle, Eq. 4 is true for all $n \in \mathbb{N}$. Hence, now we can write,

$$
\begin{aligned}
\mu^*(A) &\geq \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \bigcap_{i=1}^{\infty} B_i^{\mathrm{c}}\right) \\
&= \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^{\mathrm{c}}\right)
\end{aligned}
$$

Now, to prove that $\bigcup_i B_i \in \mathcal{M}_{\mu^*}$, we need to show

$$
\text{To Show}: \ \mu^*(A) \geq \mu^*\left(A \cap \bigcup_i B_i\right) + \mu^*\left(A \cap \left(\bigcup_i B_i\right)^{\mathrm{c}}\right)
$$

This comes from previous result as follows:

$$
\begin{aligned}
\mu^*(A) &\geq \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^{c}\right) \\
&\geq \mu^*\left(\bigcup_{i=1}^{\infty}(A \cap B_i)\right) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^{c}\right) \\
&= \mu^*\left(A \cap \bigcup_{i=1}^{\infty} B_i\right) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^{c}\right)
\end{aligned} \tag{5}
$$

Therefore, $\bigcup_i B_i \in \mathcal{M}_{\mu^*}$. Now, as the previous footnote mentions, for every infinite sequence $\{C_i\}$ in $\mathcal{M}_{\mu^*}$, we have a disjoint sequence of subsets as $C_1, C_2 \cap C_1^{\mathrm{c}}, C_3 \cap C_2^{\mathrm{c}} \cap C_1, \ldots$. Now, this disjoint sequence is closed under union as we just showed and since union of this disjoint sequence is equal to the union of $\{C_i\}$, hence $\bigcup_i C_i \in \mathcal{M}_{\mu^*}$ for any sequence $\{C_i\}$ in $\mathcal{M}_{\mu^*}$. Thus, $\mathcal{M}_{\mu^*}$ is a σ-algebra.

Act 3. *μ^* restricted to $\mathcal{M}_{\mu^*}$ is a measure.*
Consider $\{B_n\}$ be an infinite sequence of subsets in $\mathcal{M}_{\mu^*}$. Now, by finite subadditivity, we trivially have

$$\mu^*\left(\bigcup_i B_i\right) \leq \sum_i \mu^*(B_i)$$

Moreover, from Part 2 and setting $A = \cup_i B_i$, we get:

$$\begin{aligned}
\mu^*\left(\bigcup_i B_i\right) &\geq \sum_j \mu^*\left(\bigcup_i B_i \cap B_j\right) + \mu^*\left(\bigcup_i B_i \cap \left(\bigcup_i B_i\right)^c\right) \\
&= \sum_j \mu^*(B_j) + \mu^*(\Phi) \\
&= \sum_j \mu^*(B_j).
\end{aligned}$$

We hence have the complete proof. □

Definition 2.6.6. (**Lebesgue measure**) The restriction of Lebesgue outer measure on $\mathbb{R}$ to the collection $\mathcal{M}_{\lambda^*}$ of Lebesgue measurable subsets of $\mathbb{R}$ is called *Lebesgue measure.* It would be denoted by λ. Hence, we would work with the measure space $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$[4].

2.7 Does $\lambda^*(E) = 0$ implies E is countable?

We would construct today a set which has measure 0, but not countable(!).

1. Take $E_0 = [0, 1]$.
2. Remove $(1/3, 2/3)$ from E_0 to form $E_1 = [0, 1/3] \cup [2/3, 1]$.
3. Proceed in the same way to form $E_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$.
4. At n^{th} step, E_n contains 2^n subintervals and each of which is of length $\frac{1}{3^n}$.
5. We clearly have $E_0 \supset E_1 \supset E_2 \supset \dots$.
6. Here, note that each E_n is a closed and compact subset of $\mathbb{R}$.
7. The set

$$P = \bigcap_{n=0}^{\infty} E_n \text{ is known as } \textbf{Cantor Set}.$$

2.7.1 Properties of Cantor set

Proposition 2.7.1. *Lebesgue measure of Cantor Set is 0.*

Proof. Note that Cantor Set is Lebesgue measurable as it is countable intersection of closed sets, hence it is present in the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ and hence is also in $\mathcal{M}_{\lambda^*}$. Hence, instead of λ^*, we

[4]From this point on-wards, whenever this text mentions that a given set is measurable in space $(X, \mathcal{A}, \mu)$, it must be assumed that the given set is in $\mathcal{A}$, given that there is no ambiguity.

can now write λ as $P \in \mathcal{M}_{\lambda^*}$. Now, measure of Cantor set P can be written as:

$$\begin{aligned}\lambda(P) &= \lambda\left(\bigcap_n E_n\right)\\ &= \lim_{\overleftarrow{n\to\infty}} \lambda(E_n)\\ &= \lim_{\overleftarrow{n\to\infty}} \frac{2^n}{3^n}\\ &= \lim_{\overleftarrow{n\to\infty}} \frac{1}{1.5^n}\\ &= 0.\end{aligned}$$

□

Proposition 2.7.2. *Cantor set is uncountable(!)*

Proof. We will show that there exists a bijection between Cantor Set and an uncountable set, specifically ternary system. For this, consider the ternary representation of every number in $[0, 1]$. What this means is that every number in $[0, 1]$ can be represented only using the numbers $0, 1$ and 2. Hence, one write $\frac{1}{3}$ as 0.1 and $\frac{2}{3}$ as 0.2. Now, $(1/3, 2/3) = E_1^c \cap [0, 1]$ is the set that has been removed from the process of creating E_1 from E_0. Clearly, every number in this $E_1^c \cap [0, 1]$ is of the form $0.1\dots$ where $\dots$ are all combinations of $0, 1$ and 2. Therefore, we are now left with the E_1 that has all the numbers represented as $0.0\dots$ or $0.2\dots$.
As we saw in the generation of E_1, the generation of E_2 from E_1 would hence involve removing numbers of the forms $0.01\dots$ and $0.21\dots$. And hence E_2 would then be the set of numbers whose first two decimal places are restricted to NOT have the digit 1; that is, E_2 would be of form $0.02\dots, 0.00\dots, 0.20\dots, 0.22\dots$.
Continuing like this, we see that E_n would have in ternary representation, all those numbers whose first n digits are NOT 1. Hence, for any $p \in P$, p would have the ternary representation constructed only from 0 and 2, but NOT 1.
Now, consider the map $f : P \to [0, 1]$ such that $f(p)$ replaces each occurence of 2 by 1 in the ternary representation of p. We now show that this map is surjective(!) so that P has atleast as many elements as $[0, 1]$. To show this, take any $x \in [0, 1]$ in it's ternary form, and replace all 1 by 2 and denote it as x'. Clearly, x' would be in P as x' has all decimal digits generated by 0 and 2. But $f(x')$ would be opposite action and would be equal to x. Therefore, we showed that for any $x \in [0, 1], \exists\, x' \in P$ such that $f(x') = x$. Hence f is surjective. Therefore P has atleast as many elements as $[0, 1]$. But since $P \subseteq [0, 1]$ therefore P has atmost as many elements as $[0, 1]$. This dichotomy suggests that

Cantor Set has as many elements as in $[0, 1]$ *(!)*

But since $[0, 1]$ is uncountable, therefore, P is uncountable. □

With this, we conclude that **for any set $E \subseteq \mathbb{R}$, if $\lambda^*(E) = 0$, then it's NOT necessarily true that E is countable**.
We now see an extremely interesting example of a Non-measurable set.

2.8 A non-measurable set

Theorem 2.8.1. *There is a subset of $\mathbb{R}$ that is not Lebesgue measurable*[5].

Proof. We construct the proof in the following *Acts*:

Act 1. *Equivalence Relation on $\mathbb{R}$.*
Construct the following relation $\sim$ on $\mathbb{R}$:

$$x \sim y \equiv x - y \in \mathbb{Q}.$$

Clearly, $\sim$ is reflexive as $x - x = 0$ is rational; it is also symmetric as negative of a rational is also a rational number; and it is also transitive as if $x - y$ and $y - z$ is rational, then $x - y + y - z = x - z$ is sum of two rationals, which is also rational. Hence $\sim$ is an equivalence relation. Therefore $\sim$ partitions the whole $\mathbb{R}$ into equivalence classes. Note that each equivalence class of x would consist elements of the form $\mathbb{Q} + x$. But since $\mathbb{Q}$ is dense in $\mathbb{R}$, therefore $\mathbb{Q} + x$, that is each equivalence class, is dense in $\mathbb{R}$.
Now, each equivalence class clearly intersects $(0, 1)$, therefore, inducing the Axiom of Choice on the set of all equivalence classes, we can form a subset $E \subset (0, 1)$ which contains exactly one element from each of the equivalence classes. We will later prove that E is not Lebesgue measurable.

Act 2. *E satisfies certain properties.*
Consider the set $\mathbb{Q} \cap (-1, 1)$. Clearly, this is countable as it's subset of $\mathbb{Q}$. Then, consider $\{r_n\}$ to be the enumeration of $\mathbb{Q} \cap (-1, 1)$. Construct the sequence of subsets $E_n = E + r_n$. We now verify that $\{E_n\}$ satisfies the following properties:

1. The sets E_n are disjoint.
2. $\bigcup_n E_n$ is a subset of the interval $(-1, 2)$.
3. The interval $(0, 1)$ is included in $\bigcup_n E_n$.

Property 1 : Assume that $E_n \cap E_m \neq \Phi$ for some $n, m \in \mathbb{N}$ such that $n \neq m$. Then $\exists e_1, e_2 \in E$ such that $e_1 + r_n = e_2 + r_m$ which means that $e_1 - e_2 = r_m - r_n \in \mathbb{Q}$. But this cannot happen as e_1, e_2 are elements of E and E contains exactly one element from the equivalence class of $\sim$ intersected with $(0, 1)$. Therefore $e_1 - e_2 \notin \mathbb{Q}$. Which is a contradiction. Hence $E_n \cap E_m = \Phi$ for all $n, m \in \mathbb{N}$ such that $n \neq m$.
Property 2 : Take $x \in \bigcup_n E_n$. This implies that $x \in E_m$ for some $m \in \mathbb{N}$. But $E_m = E + r_m = \{e + r_m \mid e \in E\}$. Since $E \subset (0, 1)$ and $r_m \in \mathbb{Q} \cap (-1, 1) \subset (-1, 1)$, therefore $x \in E + r_m \subseteq (-1, 2)$. Hence $\bigcup_n E_n \subseteq (-1, 2)$.
Property 3 : Take any $x \in (0, 1)$. Now take the $e \in E$ such that $x \sim e$, or $x - e \in \mathbb{Q}$. Hence $x \in \mathbb{Q} + e$. That is $x = r + e$. But since $0 < e < 1$ and $0 < x < 1$, therefore $r = x - e \in \mathbb{Q} \cap (-1, 1)$. Hence $x \in E + r$ and if we denote $r = r_n$ for some $n \in \mathbb{N}$, we get $x \in E + r_n = E_n$, therefore $x \in \bigcup_i E_i$. Hence $(0, 1) \subseteq \bigcup_i E_i$.

Act 3. *E is Not Lebesgue measurable.*
Assume that E is in-fact Lebesgue measurable. Now since E_n are disjoint (Property 1), therefore we can write:

$$\lambda\left(\bigcup_n E_n\right) = \sum_n \lambda\left(E_n\right).$$

[5]See [**Solovay70**] for more information.

Now, since **Lebesgue measure is translation invariant**[6], therefore $\lambda(E_n) = \lambda(E + r_n) = \lambda(E)$. Two cases now arise for $\lambda(\bigcup_n E_n)$:

1. If $\lambda(E) = 0$: Then $\lambda(\bigcup_n E_n) = 0$. But

$$\lambda((-1,2)) = 3 \leq \lambda\left(\bigcup_n E_n\right) \quad \text{(Property 3).}$$

   Therefore we have a contradiction.

2. If $\lambda(E) \neq 0$: Then $\lambda(\bigcup_n E_n) = \sum_n \lambda(E) = +\infty$. But

$$\lambda\left(\bigcup_n E_n\right) \leq \lambda((-1,2)) = 3 \quad \text{(Property 2).}$$

   We again have a contradiction.

Hence, the set E is just not Lebesgue measurable! □

2.9 Regularity

First consider the following proposition.

Proposition 2.9.1. *Consider $E \subseteq \mathbb{R}$. The following statements are equivalent:*

1. *E is Lebesgue measurable.*
2. *$\forall \epsilon > 0$, $\exists$ an open set O such that*

$$E \subseteq O \text{ and } \lambda^*(O \setminus E) < \epsilon.$$

3. *$\exists$ a G_δ set G such that*

$$E \subseteq G \text{ and } \lambda^*(G \setminus E) = 0.$$

Proof. The equivalence of each statement is as follows:
1 $\implies$ 2. Consider $E \subseteq \mathbb{R}$ to be Lebesgue measurable. By above, for any $E \subseteq \mathbb{R}$ and any $\epsilon > 0$, there exists open set U such that $E \subseteq U$ which satisfies

$$\lambda^*(U) \leq \lambda^*(E) + \epsilon.$$

Now since $E \subseteq U$, therefore,

$$\begin{aligned}\lambda^*(U \setminus E) &= \lambda^*(U) - \lambda^*(E) \\ &\leq \epsilon\end{aligned}$$

2 $\implies$ 3. Similarly, the above shows that there exists a G_δ set G such that $E \subseteq G$ which satisfies $\lambda^*(E) = \lambda^*(G)$. This directly means that $\lambda^*(G \setminus E) = 0$ because $E \subseteq G$ so $\lambda^*(G \setminus E) = \lambda^*(G) - \lambda^*(E)$.

3 $\implies$ 1. Since G is G_δ set therefore it is intersection of open sets in $\mathbb{R}$. Now since any

[6] Proof?

open set in $\mathbb{R}$ is an union of open intervals (Homework I, 1) which is Lebesgue measurable and therefore G is Lebesgue measurable. Now, we can write E as

$$E = G \setminus (G \setminus E)$$

where $G \setminus E$ is such that (from Statement 3) $\lambda^* (G \setminus E) = 0$, therefore, by Proposition 2.6.4, $G \setminus E$ is Lebesgue measurable. Hence E is also Lebesgue measurable.

□

Now, consider the next proposition, which is dual of the above.

Proposition 2.9.2. *Consider $E \subseteq \mathbb{R}$. The following statements are equivalent:*

1. *E is Lebesgue measurable.*
2. *$\forall\, \epsilon > 0$, $\exists$ closed set C such that*

$$C \subseteq E \text{ and } \lambda^* (E \setminus C) < \epsilon.$$

3. *$\exists$ a F_σ set F such that*

$$F \subseteq E \text{ and } \lambda^* (E \setminus F) = 0.$$

Proof. Implications are as follows:
1 $\implies$ 2. Suppose $E \subseteq \mathbb{R}$ is Lebesgue measurable. Note that if E is Lebesgue measurable (that is $E \in \mathcal{M}_{\lambda^*}$), then E^c is also Lebesgue measurable as $\mathcal{M}_{\lambda^*}$ is a σ-algebra (Theorem 2.6.5). Hence, using Proposition 2.9.1 on E^c gives us an open set O for all $\epsilon > 0$ such that $E^c \subseteq O$ and $\lambda^* (O \setminus E^c) < \epsilon$. Now let's take it's complement. Therefore, $C = O^c \subseteq E$ where C is clearly closed. Now, $E \setminus O^c = O \setminus E^c$[7]. Now,

$$\begin{aligned}\lambda^* (E \setminus O^c) &= \lambda^* (O \setminus E^c) \\ &< \epsilon\end{aligned}$$

which proves the first implication.

2 $\implies$ 3. From Proposition 2.9.1, we have that $\exists$ a G_δ set G such that $E^c \subseteq G$ and $\lambda^* (G \setminus E^c) = 0$. Note that the complement of countable intersection of open sets is countable union of closed sets. Therefore, $F = G^c$ is an F_σ set. Now, $G^c \subseteq (E^c)^c = E$. Now, we know that $E \setminus G^c = G \setminus E^c$. Therefore, we have the result as follows:

$$\begin{aligned}\lambda^* (E \setminus G^c) &= \lambda^* (G \setminus E^c) \\ &= 0.\end{aligned}$$

3 $\implies$ 1. Since F is an F_σ set, therefore, $F \in \mathcal{M}_{\lambda^*}$. Moreover, as Statement 2 show, $\lambda^* (E \setminus F) = 0$, thus by Proposition 2.6.4, $E \setminus F \in \mathcal{M}_{\lambda^*}$. Since,

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

that is E is union of two Lebesgue measurable sets, therefore $E \in \mathcal{M}_{\lambda^*}$, completing the proof. □

[7] It's not difficult to see as for any $x \in E \setminus O^c$, $x \in E$ but $x \notin O^c$. Therefore, $x \in O$ but $x \notin E^c$, that is $x \in O \setminus E^c$. Similarly for the converse.

Definition 2.9.3. (**Complete measure Space**) The measure space $(X, \mathcal{A}, \mu)$ is complete if the for any $A \in \mathcal{A}$ such that $\mu(A) = 0$ implies that for any subset $B \subseteq A$,

$$\mu(B) = 0.$$

Remark 2.9.4. Trivial to see are the following:
- Hence, if μ^* is an outer measure defined on X, then the space $(X, \mathcal{M}_{\mu^*}, \mu^*)$ is complete (follows from Proposition 2.6.4).
- This means that the Lebesgue outer measure restricted to Lebesgue measurable subsets of $\mathbb{R}$, $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ is complete.

Definition 2.9.5. (**Completion of a measure Space**) Let $(X, \mathcal{A})$ be a measurable space and let μ be a measure on $\mathcal{A}$. The completion of $\mathcal{A}$ under μ is the collection $\mathcal{A}_\mu$ of subsets $A \subseteq X$ for which there are sets E and F in $\mathcal{A}$ such that

$$E \subseteq A \subseteq F$$

and

$$\mu(F - E) = 0$$[8].

3 Measurable functions

We now see the definition and basic properties of measurable Functions, which would later be used to define Lebesgue integral.

Definition 3.0.1. (**Measurable function**) Let $(X, \mathcal{A})$ be a measurable space and let $A \subseteq X$ which is in $\mathcal{A}$. The function $f : A \to [-\infty, +\infty]$, is called a measurable function[9] if

$$\{x \mid f(x) > \alpha\} \text{ for any } \alpha \in \mathbb{R} \text{ is measurable (belongs in } \mathcal{A}).$$

Remark 3.0.2. Please note that the function f defined above has a measurable domain.

Proposition 3.0.3. *Let $(X, \mathcal{A})$ be a measurable space and $A \in \mathcal{A}$. Let $f : A \to [-\infty, +\infty]$ be a function. Then, the following statements are equivalent:*
1. *f is a measurable function.*
2. *For all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) \geq \alpha\} \in \mathcal{A}$.*
3. *For all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) < \alpha\} \in \mathcal{A}$.*
4. *For all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) \leq \alpha\} \in \mathcal{A}$.*

Proof. The equivalence is shown as follows:
1 $\implies$ 2. Since f is a measurable, therefore for all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) > \alpha\} \in \mathcal{A}$. This means that $C_{\alpha - \frac{1}{n}} = \{x \mid f(x) > \alpha - \frac{1}{n}\} \in \mathcal{A}$ for all $n \in \mathbb{N}$. Now, the following set

$$C = \bigcap_n C_{\alpha - \frac{1}{n}} = \{x \mid f(x) \geq \alpha\}.$$

[8]Note that, in Exercise III, Q. 2, we proved that for any $A \in \mathcal{A}$, this is trivially true. That is, all $\mathcal{A}$-measurable subsets are $\mathcal{A}_\mu$-measurable. In particular, E was a F_σ set and F was a G_δ set.

[9]One writes f as $\mathcal{A}$-measurable function to denote the σ-algebra over whose subset the function f is defined.

is measurable as $C \in \mathcal{A}$ because $C_{\alpha - \frac{1}{n}} \in \mathcal{A}$ for any $n \in \mathbb{N}$, hence the countable intersection would also be in $\mathcal{A}$, hence measurable.
2 $\implies$ **3.** Since $\{x \mid f(x) \geq \alpha\} \in \mathcal{A}$, therefore it's complement $\{x \mid f(x) < \alpha\} \in \mathcal{A}$ for any $\alpha \in \mathbb{R}$.
3 $\implies$ **4.** Since $\{x \mid f(x) < \alpha\} \in \mathcal{A}$ for any $\alpha \in \mathbb{R}$, thus, $C_{\alpha + \frac{1}{n}} = \{x \mid f(x) < \alpha + \frac{1}{n}\} \in \mathcal{A}$ for all $n \in \mathbb{N}$, hence

$$C = \bigcap_n C_{\alpha + \frac{1}{n}} = \{x \mid f(x) \leq \alpha\}$$

and since each $C_{\alpha + \frac{1}{n}} \in \mathcal{A}$, therefore $C \in \mathcal{A}$.
4 $\implies$ **1.** Since $\{x \mid f(x) \leq \alpha\} \in \mathcal{A}$ then it's complement $\{x \mid f(x) > \alpha\}$ for all $\alpha \in \mathbb{R}$, making f measurable. ☐

Proposition 3.0.4. *The following are basic examples of measurable functions:*

- *If f is a measurable function, then the set $\{x \mid f(x) = \alpha\}$ is measurable for all $\alpha \in R$.*
- *Constant functions are measurable.*
- *The characteristic function χ_A defined by:*

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

  is measurable if and only if A is measurable.
- *Continuous functions are measurable.*
- *Let $(X, \mathcal{A})$ be a measurable space. If f and g are measurable functions on X, then the sets*

$$\{x \in X \mid f(x) \neq g(x)\}$$
$$\{x \in X \mid f(x) < g(x)\}$$

  are measurable (belongs to $\mathcal{A}$).
- [10] *Monotone functions are measurable.*
- [11]*Consider $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function. Then f' is a λ-measurable function.*

Proof. The **first** example is trivial to see in light of Proposition 3.0.3 by taking intersection of $\{x \mid f(x) \leq \alpha\}$ and $\{x \mid f(x) \geq \alpha\}$, both of which are measurable.
For **second**, consider the constant function $f(x) = b \;\forall\; x \in \mathbb{R}$. Now, for all $\alpha \in \mathbb{R}$, consider the set $f^{-1}((\alpha, \infty)) = \{x \mid f(x) > \alpha\}$. If $b > \alpha$, then we are done, if $b \leq \alpha$, then by previous result, $\{x \mid f(x) \leq \alpha\}$ is also measurable (equal to $\mathbb{R}$ and $\mathbb{R} \in \mathcal{A}$).
For **third** example, consider the set ${\chi_A}^{-1}(\alpha, \infty) = \{x \mid \chi_A(x) > \alpha\}$ for any $\alpha \in \mathbb{R}$. If $\alpha > 1$, then $f^{-1}(\alpha, \infty) = \Phi \in \mathcal{A}$. If $\alpha = 1$, then $f^{-1}[\alpha, \infty) = A$, since $\chi_A(x)$ is given measurable, hence A is measurable. Now, Assume that A is measurable. Then consider the set ${\chi_A}^{-1}(\alpha, \infty)$ for any $\alpha \in \mathbb{R}$. As we saw previously, the case for $\alpha > 1$ is trivial. For $0 < \alpha \leq 1$, ${\chi_A}^{-1}(\alpha, \infty) = A \in \mathcal{A}$. Finally, for $\alpha \leq 0$, ${\chi_A}^{-1}(-\infty, \alpha] = \Phi \in \mathcal{A}$. Thus, χ_A is measurable.
For **fourth**, since f is continuous (so inverse of open sets is open, by definition), therefore $f^{-1}(\alpha, \infty)$ is open in $\mathbb{R}$, hence it must be Borel, hence measurable for any $\alpha \in \mathbb{R}$.

[10]Question 3 of Exercise 3.
[11]Question 4 of Exercise 3.

For **fifth**, since f and g are measurable. Then due to next Proposition 3.0.5, we know that $f-g$ is also measurable. This means that for any $\alpha \in \mathbb{R}$,

$$\{x \in X \mid f(x) - g(x) < \alpha\}$$

is measurable. Now set $\alpha = 0$ to get the result. Moreover, from this, we also get that $\{x \in X \mid f(x) - g(x) > 0\}$ is also measurable. Hence,

$$\{x \in X \mid f(x) - g(x) \neq 0\} = \{x \in X \mid f(x) - g(x) < 0\} \bigcup \{x \in X \mid f(x) - g(x) > 0\}$$

is also measurable.
For **sixth**, we proceed as follows:
Consider the function $f : A \to \mathbb{R}$ where $A \in \mathcal{M}_{\lambda^*}$ to be monotone. Now, consider the following two sets for any $\alpha \in \mathbb{R}$:

$$\begin{aligned} A_1 &= \{x \in A \mid f(x) > \alpha\} \\ A_2 &= \left(f^{-1}(\alpha), \infty\right) \cap A \end{aligned}$$

Now, take any $x \in A_1$, then $f(x) > \alpha \implies x > f^{-1}(\alpha)$. Now if $f^{-1}(\alpha) \cap A = \Phi$, then $\{x \in A \mid f(x) > \alpha\} = \Phi$ which is trivially measurable and we would be done. If however $f^{-1}(\alpha) \cap A \neq \Phi$, then $f^{-1}(\alpha) = \{y \in A \mid f(y) = \alpha\}$ so that $f(y) > \alpha$ implies that $y > f^{-1}(\alpha)$ so that $f(y) > f(f^{-1}(\alpha)) = \alpha$. Therefore, $x > f^{-1}(\alpha)$, that is $x \in A_2$, proving that $A_1 \subseteq A_2$.
Similarly, take $x \in A_2$, therefore

$$\begin{aligned} x &> f^{-1}(\alpha) \\ f(x) &> f\left(f^{-1}(\alpha)\right) \\ f(x) &> \alpha \\ x &\in \{x \mid f(x) > \alpha\} \\ x \in A_1. \end{aligned}$$

Therefore $A_2 \subseteq A_1$. Hence, $A_1 = A_2$. But since $\left(f^{-1}(\alpha), \infty\right)$ is an interval, hence measurable and A is given measurable, therefore $A_2 = \left(f^{-1}(\alpha), \infty\right) \cap A$ is measurable, which makes $A_1 = \{x \mid f(x) > \alpha\} = A_2$ measurable for all $\alpha \in \mathbb{R}$.
For **seventh**, the result is simple to see since we are given that f is λ-measurable due to continuity (see Statement 4). Therefore, we can define the sequence of functions $\{f_n\}$ as follows:

$$f_n(x) = \frac{f\left(x + \frac{1}{n}\right) - f(x)}{\frac{1}{n}} \forall\, x \in \mathbb{R}.$$

As we can see, f_n is λ-measurable due to Proposition 3.0.5. Hence, we can see that because $f'(x) = \varinjlim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ for any $x \in \mathbb{R}$, and since $\varinjlim_{n\to\infty} f_n(x) = f(x)$, therefore $f_n \to f'$ is λ-measurable (Proposition 3.0.9). $\square$

Proposition 3.0.5. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$. Consider two measurable functions $f, g : A \longrightarrow [0, +\infty]$ and $c \in \mathbb{R}$. Then,*

1. $f + c$,

2. $f \pm g$
3. cf,
4. fg
are also measurable.

Proof. **1.** Since f is measurable, therefore the set $\{x \mid f(x) > \alpha - c\} = \{x \mid f(x) + c \geq \alpha\}$ is measurable for any $\alpha \in \mathbb{R}$.
2. Both f and g are given measurable. The set $(f+g)^{-1}(\alpha, \infty)$ can be written as:

$$\begin{aligned}(f+g)^{-1}(\alpha, \infty) &= \{x \mid f(x) + g(x) > \alpha\} \\ &= \{x \mid f(x) > \alpha - g(x)\} \\ &= \{x \mid f(x) > b\}\end{aligned}$$

where $b \in [-\infty, \alpha]$. Note that the case where $g(x) = +\infty$ is trivial as $f(x) > \alpha - (+\infty) \equiv f(x) > -\infty$, which is by definition of co-domain of f. Now since $\{x \mid f(x) > b\}$ is measurable for any $b \in \mathbb{R} \supset (-\infty, \alpha]$ for any $\alpha \in \mathbb{R}$, therefore $(f+g)^{-1}(\alpha, \infty)$ is measurable for any $\alpha \in \mathbb{R}$.
3. Note that for $c = 0$, the function becomes constant and hence measurable (Proposition 3.0.4). Consider the set $(cf)^{-1}(\alpha, \infty)$. We can write this as follows,

$$\begin{aligned}(cf)^{-1}(\alpha, \infty) &= \{x \mid cf(x) > \alpha\} \\ &= \{x \mid f(x) > \alpha/c\}\end{aligned}$$

where $c \neq 0$. Since f is measurable, therefore $\{x \mid f(x) > \alpha/c\}$ is also measurable for any $\alpha \in \mathbb{R}$. Hence cf is measurable.
4. Consider the set $(f^2)^{-1}(-\infty, \alpha)$ for any $\alpha \in \mathbb{R}$.

$$\begin{aligned}(f^2)^{-1}(-\infty, \alpha) &= \{x \mid f^2(x) < \alpha\} \\ &= \{x \mid -\sqrt{\alpha} < f(x) < \sqrt{\alpha}\} \\ &= \{x \mid f(x) < \sqrt{\alpha}\} \bigcap \{x \mid f(x) > -\sqrt{\alpha}\}\end{aligned}$$

Therefore if f is measurable, then f^2 is measurable. With this, we can simply write fg as:

$$fg = \frac{(f+g)^2 - (f-g)^2}{4}$$

which, by previous results (2 & 3), is measurable. □

Proposition 3.0.6. [12] *Let $(X, \mathcal{A})$ be a measurable space. Consider a function $f : A \to \mathbb{R}$ where $A \in \mathcal{A}$. Then the following are equivalent:*
1. *f is a $\mathcal{A}$-measurable function.*
2. *$f^{-1}(U)$ is a measurable set $\forall$ open sets $U \subseteq \mathbb{R}$.*
3. *$f^{-1}(C)$ is a measurable set $\forall$ closed sets $C \subseteq \mathbb{R}$.*
4. *$f^{-1}(B)$ is a measurable set $\forall$ borel sets $B \in \mathcal{B}(R)$.*

Proof. *The proof is exactly the same as of Proposition 3.2.2.* □

[12] Question 1 of Exercise 3.

Definition 3.0.7. (**sequence of fuctions**) If $\{f_n\}$ is a sequence of $[-\infty, +\infty]$ valued functions on A, then $\sup_n f_n : A \to [-\infty, +\infty]$ is defined by

$$\left(\sup_n f_n\right)(x) = \sup\{f_n(x) \mid n \in \mathbb{N}\}.$$

Remark 3.0.8. One similarly defines the following:

- The infimum:
$$\left(\inf_n f_n\right)(x) = \inf\{f_n(x) \mid n \in \mathbb{N}\}.$$
- The limit supremum:
$$\left(\limsup_n f_n\right)(x) = \limsup\{f_n(x) \mid n \in \mathbb{N}\}.$$
- The limit infimum:
$$\left(\liminf_n f_n\right)(x) = \liminf\{f_n(x) \mid n \in \mathbb{N}\}.$$
- The limit:
$$\left(\varprojlim_n f_n\right)(x) = \varprojlim\{f_n(x) \mid n \in \mathbb{N}\}.$$

Proposition 3.0.9. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$. Consider $\{f_n\}$ be a sequence of $[-\infty, +\infty]$-valued measurable functions on A. Then,*

1. *The functions $\sup_n f_n$ and $\inf_n f_n$ are measurable.*
2. *The functions $\limsup_n f_n$ and $\liminf_n f_n$ are measurable.*
3. *The function $\varprojlim_n f_n$ (whose domain is $\{x \in A \mid \limsup_n f_n(x) = \liminf_n f_n(x)\}$) is measurable.*

Proof. Note that the set $(\sup_n f_n)^{-1}(-\infty, \alpha] = \{x \in A \mid (\sup_n f_n)(x) \leq \alpha\} = \bigcap_n\{x \in A \mid f_n(x) \leq \alpha\}$. Therefore $\sup_n f_n$ is measurable. Similarly, $(\inf_n f_n)^{-1}(-\infty, \alpha) = \{x \in A \mid (\inf_n f_n)(x) < \alpha\} = \bigcup_n\{x \in A \mid f_n(x) < \alpha\}$. Now, denote $g_k = \sup_{n\geq k} f_n$ and $h_k = \inf_{n\geq k} f_n$. But since $\limsup_n f_n = \inf_{n\geq 0}\sup_{k\geq n} f_k = \inf_{n\geq 0} g_n$ and $\{g_n\}$ is measurable by 1^{st} property, therefore $\limsup_n f_n$ is also measurable, similarly for $\liminf_n$. $\square$

3.1 Almost everywhere property.

Definition 3.1.1. (**μ-almost everywhere**) Let $(X, \mathcal{A}, \mu)$ be a measure space. A property P of points of X is said to hold μ-almost everywhere if the set

$$N = \{x \in X \mid P \text{ does not hold for } x\}$$

has measure zero. That is,

$$\mu(N) = 0.$$

Remark 3.1.2. Note that it's not necessary for the set N to belong in $\mathcal{A}$. The only requirement is for the set N to be contained in a set $F \in \mathcal{A}$ and then $\mu(F) = 0$ (which automatically implies that $\mu^*(N) = 0$).

But, if μ is complete then $N \in \mathcal{A}$. See Definition 2.9.3.

Definition 3.1.3. (**Almost everywhere convergence**) If $\{f_n\}$ is a sequence of functions on X and f is a function on X, then

$$\{f_n\} \longrightarrow f \text{ almost everywhere.}$$

if the set

$$\{x \in X \mid f(x) \neq \varinjlim_n f_n(x)\}$$

is of measure zero.

Proposition 3.1.4. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let f and g be extended real valued functions on X that are equal almost everywhere. If μ is **complete** and if f is measurable, then g is also measurable.*

Proof. Consider the region of non-equality as

$$N = \{x \mid f(x) \neq g(x)\}.$$

Given to us is the fact that $\mu^*(N) = 0$ and since μ is complete, so $N \in \mathcal{A}$. Now, consider the following for any $\alpha \in \mathbb{R}$:

$$\{x \mid g(x) \geq \alpha\} = (\{x \mid g(x) \geq \alpha\} \cap N) \bigcup (\{x \mid g(x) \geq \alpha\} \cap N^c) \, .$$

Denote the set $A = \{x \mid g(x) \geq \alpha\} \cap N$ and $B = \{x \mid g(x) \geq \alpha\} \cap N^c$. Since for any $x \in (\{x \mid g(x) \geq \alpha\} \cap N^c)$, $f(x) = g(x)$, therefore, we can equivalently write $B = (\{x \mid f(x) \geq \alpha\} \cap N^c)$. Now $N^c \in \mathcal{A}$ and due to Measurability of f, $\{x \mid f(x) \geq \alpha\} \in \mathcal{A}$. Hence $B \in \mathcal{A}$. Finally, due to $\{x \mid g(x) \geq \alpha\} \cap N \subseteq N$ and μ being complete with $\mu(N) = 0$, we get $\{x \mid g(x) \geq \alpha\} \cap N \in \mathcal{A}$, completing the proof. □

Proposition 3.1.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space, let $\{f_n\}$ be sequence of extended real valued functions on X and let f be an extended real valued function on X such that*

$$\{f_n\} \longrightarrow f \textit{ almost everywhere.}$$

*If μ is **complete** and if each f_n is measurable, then f is measurable.*

Proof. As Proposition 3.0.9 shows, $\liminf_n f_n$ and $\limsup_n f_n$ are measurable. As the given condition shows, $\liminf_n f_n$ is equal to f for almost all X. Hence Proposition 3.1.4 implies that f is also measurable. □

3.2 Cantor set

With the new tool in hand (measurable functions), we now turn back to the ever-interesting Cantor set, this time, to prove the sheer size of the σ-algebra $\mathcal{M}_{\lambda^*}$ in comparison to the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. In particular we show that $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{M}_{\lambda^*}$.
But before that, we look at following results:

Proposition 3.2.1. *The function ϕ defined by*

$$\begin{aligned}\phi : [0,1] &\longrightarrow P \\ \phi(\alpha) &= \sum_{n=1}^{\infty} \frac{2b_n}{3^n} \text{ for } \alpha \in [0,1]\end{aligned}$$

where $b_n \in \{0,1\}\ \forall\, n \in \mathbb{N}$ is measurable in $\mathcal{M}_{\lambda^*}$.

Proof. Note that $\phi(\alpha)$ thus maps a decimal number to it's binary representation $\{b_n\}$. First, we define the following function:

$$\begin{aligned}\phi_n : [0,1] &\longrightarrow \{0,1\} \\ \phi_n(\alpha) &= b_n.\end{aligned}$$

That is, ϕ_n maps α to it's n^{th} binary digit. We can see that $\phi_n(\alpha)$ can be written as the following:

$$\phi_n(\alpha) = \chi_{E_n} = \begin{cases} 1 & \text{if } \alpha \in E_n \\ 0 & \text{otherwise.} \end{cases}$$

where E_n is the intersection of countable sequence of sub-intervals of $[0,1]$. Hence E_n is a Lebesgue measurable subset of $\mathbb{R}$, so it is in $\mathcal{M}_{\lambda^*}$. But, as Proposition 3.0.4, statement 3 shows, $\chi_{E_n} = \phi_n$ is then a measurable function.
Now, the following arguments:

$$\begin{aligned}\frac{2}{3^n}\phi_n(\alpha) = \frac{2b_n}{3^n} & \text{ is measurable (Proposition 3.0.5).} \\ \left\{\frac{2\phi_n(\alpha)}{3^n}\right\} & \text{ is a sequence of measurale functions.} \\ \left\{\sum_{k=1}^{n} \frac{2\phi_n(\alpha)}{3^n}\right\} & \text{ is also a sequence of measurable functions (Proposition 3.0.5).} \\ \varprojlim_{n\to\infty} \sum_{k=1}^{n} \frac{2\phi_n(\alpha)}{3^n} & \text{ is a measurable function (Proposition 3.0.9).}\end{aligned}$$

Hence the function which maps each real from $[0,1]$ to it's binary representation is measurable. $\square$

Proposition 3.2.2. *Let $(X, \mathcal{A})$ be a measurable space. If f is a $\mathcal{A}$-measurable function on A and $B \in \mathcal{B}(\mathbb{R})$, then $f^{-1}(B) \in \mathcal{A}$.*

Proof. Denote $\mathcal{D}$ be the following set:

$$\mathcal{D} = \{B \subseteq \mathbb{R} \mid f^{-1}(B) \in \mathcal{A}\}.$$

Now, note that,

1. Since $f^{-1}(\mathbb{R}) = A \in \mathcal{A}$, therefore $\mathbb{R} \in \mathcal{D}$.

2. If $B \in \mathcal{D}$, then
$$B^{c} = \mathbb{R} \cap B^{c}$$
and
$$\begin{aligned} f^{-1}(B^{c}) &= f^{-1}(\mathbb{R} \cap B^{c}) \\ &= f^{-1}(\mathbb{R}) \cap f^{-1}(B^{c}) \\ &= A \cap \left(f^{-1}(B)\right)^{c} \end{aligned}$$
Now since $A \in \mathcal{A}$ and $f^{-1}(B) \in \mathcal{A}$ because $B \in \mathcal{D}$, therefore $f^{-1}(B^{c}) \in \mathcal{A}$ so that $B^{c} \in \mathcal{D}$.
3. We know that from the basic results of set functions that
$$f^{-1}\left(\bigcup_n B_n\right) = \bigcup_n f^{-1}(B_n)$$

Hence $\mathcal{D}$ is a σ-algebra on $\mathbb{R}$ (!) Now, due to measurability of f, we know that the set $\{x \mid f(x) > \alpha\}$ is in $\mathcal{A}$, which is equivalent to saying that $f^{-1}(\alpha, \infty) \in \mathcal{A}$. This hence means that $(\alpha, \infty) \in \mathcal{D}$ for any $\alpha \in \mathbb{R}$. Proposition 2.1.8 showed that a σ-algebra generated by such subsets of $\mathbb{R}$ is $\mathcal{B}(\mathbb{R})$. Hence, for any $B \in \mathcal{B}(\mathbb{R})$, we have that $B \in \mathcal{D}$. Therefore for any Borel set B, $f^{-1}(B) \in \mathcal{A}$[13]. □

3.3 Sequence of functions approximating a measurable function.

We now show that any measurable function can be defined in terms of a simple function and a step function. For this, we first define what we mean by simple functions in Definition 3.3.4. Before that, let's see few more interesting-but-basic properties of measurable functions.

Proposition 3.3.1. *Let $(X, \mathcal{A})$ be a measurable space and f be an extended real valued function on $A \in \mathcal{A}$. Define the following:*
$$f^{+}(x) = \max(f(x), 0) \text{ and } f^{-}(x) = -\min(f(x), 0).$$
Then, f is measurable if and only if f^{+} and f^{-} both are measurable on A.

Proof. If f is measurable, then $\{x \mid f(x) \geq \alpha\}$ is measurable. Note that $f^{+}(x) \geq 0$. Hence, for the case when $\alpha > 0$, the set $\{x \mid f^{+}(x) \geq \alpha\} = \{x \mid f(x) > \alpha\}$ which is measurable due to measurability of f. Similarly, if $\alpha = 0$, then $\{x \mid f^{+}(x) \geq 0\} = \{x \mid f(x) > 0\} \bigcup \{x \mid f(x) = 0\}$ in which both sets are measurable in view of Proposition 3.0.4. Finally, for $\alpha < 0$, we have $\{x \mid f^{+}(x) > \alpha\} = \{x \mid f^{+}(x) \geq 0\}$ which again is measurable. Now, $f^{-}(x) = -\min(f(x), 0) = \max(-f(x), 0)$ and since $-f$ is also measurable (Proposition 3.0.5), therefore if f is a measurable function, then f^{+} and f^{-} are both measurable functions too.
To show the converse, note that $f = f^{+} - f^{-}$ and since both are measurable, therefore f is also measurable (Proposition 3.0.5). □

Remark 3.3.2. Due to the above result, we can hence deduce that if f is a $\mathcal{A}$-measurable function then,
$$|f| = f^{+} + f^{-} \text{ is a measurable function on } A.$$

[13] This is a very interesting way to prove such a statement. Notice how we analyzed the set of all possible subsets of $\mathbb{R}$ for which $f^{-1}(B) \in \mathcal{A}$ right from the start!

Proposition 3.3.3. *Let $(X, \mathcal{A})$ be a measurable space and $A \in \mathcal{A}$. Let $f : A \to [-\infty, +\infty]$. Then,*

1. *If f is $\mathcal{A}$-measurable and if B is a subset of A, then the restriction f_B of f to B is also $\mathcal{A}$-measurable.*
2. *If $\{B_n\}$ is a sequence of sets that belong to $\mathcal{A}$ such that $A = \bigcup_n B_n$ and f_{B_n} is $\mathcal{A}$-measurable for each n, then f is also $\mathcal{A}$-measurable.*

Proof. The first result follows directly from the following observation:

$$\{x \in B \mid f_b(x) > \alpha\} = B \bigcap \{x \in A \mid f(x) > \alpha\}$$

and the second result follows from the following:

$$\{x \in A \mid f(x) > \alpha\} = \bigcup_n \{x \in B_n \mid f_{B_n}(x) > \alpha\}.$$

both for any $\alpha \in \mathbb{R}$. □

Definition 3.3.4. (**Simple Function**) A function is called simple if it has only finitely many values. Equivalently, we say that f is simple if we can write it as the following:

$$f = \sum_{k=1}^{N} \alpha_k \chi_{E_k} \ , \quad \alpha_k \in \mathbb{R}$$

where each E_k is a measurable set of finite measure.

Remark 3.3.5. Note that

- If E_k are intervals, then we say f to be a **step function**.

*The following Proposition asserts that **any measurable function can be approximated by an increasing sequence of simple functions.***

Proposition 3.3.6. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$ with $f : A \to [0, +\infty]$ be a measurable function on A. Then there exists a sequence $\{f_n\}$ of **simple** $[0, +\infty)$-valued measurable functions on A that satisfy*

$$f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots$$

and

$$\underleftarrow{\lim}_{n \to \infty} f_n(x) = f(x)$$

for any $x \in A$.

Proof. For the proof, construct the following sequence of sets, by dividing the whole interval $[0, n]$ for any $n \in \mathbb{N}$ into $n2^n$ number of intervals each of length $\frac{1}{2^n}$ and denote the following set:

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

for any $n \in \mathbb{N}$ and $k = 1, 2, \dots, n2^n$. With this construction, we can now define the following function for each n:

$$\phi_n : A \to [0, \infty), \text{ defined as}$$

$$\phi_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } x \in A_{n,k} \text{ for any } k = 1, 2, \dots, n2^n \\ n & \text{if } x \in A - \bigcup_k A_{n,k}. \end{cases}$$

Note that we can alternatively write $\phi_n(x)$ as the following (with more clarity):

$$\phi_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } f(x) \leq n \ , \quad \text{where } \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \text{ for some k} \in \{1, 2, \ldots, n2^n\} \\ n & \text{if } f(x) > n. \end{cases}$$

We now show that $\phi_n(x) \leq \phi_{n+1}(x) \ \forall \ x \in A$. Let's first show this for $f(x) \leq n$.
If $f(x) \leq n$, then,

$$\phi_n(x) = \frac{k_0 - 1}{2^n} \text{ for some } k_0 \in \{1, 2, \ldots, n2^n\}$$

such that $\frac{k_0-1}{2^n} \leq f(x) < \frac{k_0}{2^n}$. Now, two cases arises:

- **If** $\frac{k_0-1}{2^n} \leq f(x) < \frac{2k_0-1}{2^{n+1}}$: This is just the case that $f(x)$ lies in the first half of the interval $\left[\frac{k_0-1}{2^n}, \frac{k_0}{2^n}\right]$. Hence, in this case we get that:
$$\frac{k_0 - 1}{2^n} = \frac{2k_0 - 2}{2^{n+1}} \leq f(x) < \frac{2k_0 - 1}{2^{n+1}}$$
such that $\phi_n(x) = \frac{k_0-1}{2^n} = \phi_{n+1}(x)$.
- **If** $\frac{2k_0-1}{2^{n+1}} \leq f(x) < \frac{k_0}{2^n}$: This is the case when $f(x)$ lies in the second half of the interval. In this case, we see that,
$$\frac{2k_0 - 1}{2^{n+1}} \leq f(x) < \frac{2k_0}{2^{n+1}} = \frac{k_0}{2^n}$$
so that $\phi_n(x) = \frac{k_0-1}{2^n} = \frac{2k_0-2}{2^{n+1}} < \frac{2k_0-1}{2^{n+1}} = \phi_{n+1}(x)$.

Hence from both the cases, we have $\phi_n(x) \leq \phi_{n+1}(x)$ for all $x \in A$ such that $f(x) \leq n$. One can similarly see the same result for $n < f(x) \leq n + 1$ and for $f(x) > n + 1$, $\phi_n(x) \leq \phi_{n+1}(x)$ follows trivially. Hence, we have proved that $\forall n \in \mathbb{N}$ and $x \in A$,

$$\phi_n(x) \leq \phi_{n+1}(x). \tag{6}$$

Now, one can write the function ϕ_n as the following combination too:

$$\phi_n(x) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{A_{n,k}} + n\chi_{A - \bigcup_k A_{n,k}} \tag{7}$$

Due to the above representation of ϕ_n, the following steps becomes easier (& interesting) to see. Now, first note that $A_{n,k}$ is a measurable set because it's intersection of two measurable sets. Moreover, $A - \bigcup_k A_{n,k}$ is also a measurable set. Hence, in view of Proposition 3.0.4, Statement 3, we get that $\phi_n(x)$ is a measurable function for any $n \in \mathbb{N}$. Therefore, $\{\phi_n\}$ is a sequence of measurable functions adhering (6). We again find two cases:

- **If f is finite** : Now since f is finite, therefore $\exists n_0 \in \mathbb{N}$ such that $f(x) \leq n_0$. Hence, one can further deduce the following for all $n \geq n_0$ (hence $f(x) \leq n_0 \leq n$),
$$f(x) - \phi_n(x) = f(x) - \frac{k-1}{2^n} \text{ for some } k \in \{1, 2, \ldots, n2^n\} \text{ such that } \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$$
$$< \frac{1}{2^n}$$
Hence, as $n \to \infty$, $|f(x) - \phi_n(x)| \to 0$.

- **If f is infinite for some $x \in A$:** If f is infinite, then $\forall n \in \mathbb{N}$, $f(x) > n$. Hence,

$$\phi_n(x) = n \text{ for all } n \in N.$$

Therefore $\varinjlim_{n\to\infty} \phi_n(x) = +\infty = f(x)$ for particular $x \in A$ where f is infinity.
Hence, in both cases, $\{\phi_n\}$ converges to f. The proof is therefore complete. □

The following can be considered as an ***important corollary of the above Proposition.***

Proposition 3.3.7. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$ with $f : A \to [-\infty, +\infty]$ be a measurable function on A. Then there exists a sequence $\{f_n\}$ of* ***simple*** *$(-\infty, +\infty)$-valued measurable functions on A that satisfy*

$$|f_1(x)| \leq |f_2(x)| \leq |f_3(x)| \leq \dots$$

and

$$\varinjlim_{n\to\infty} f_n(x) = f(x)$$

for any $x \in A$.

Proof. Since f is a measurable function, therefore f^+ and f^- are measurable functions too (Proposition 3.3.1). Now, since any function f can be written as

$$f = f^+ - f^-$$

therefore, by Proposition 3.3.6, we have two sequences $\{f_n^{(1)}\}$ and $\{f_n^{(2)}\}$ such that

$$f_n^{(1)} \longrightarrow f^+ \text{ and } f_n^{(2)} \longrightarrow f^-$$

where $f_1^{(1)}(x) \leq f_2^{(1)}(x) \leq \dots$ and $f_1^{(2)}(x) \leq f_2^{(2)}(x) \leq \dots$. Denote

$$f_n(x) = f_n^{(1)}(x) - f_n^{(2)}(x)$$

Therefore, we see that

$$|f_n(x)| = f_n^{(1)}(x) + f_n^{(2)}(x) \leq f_{n+1}^{(1)}(x) + f_{n+1}^{(2)}(x) = |f_{n+1}(x)|$$

Now, we can deduce that

$$\begin{aligned}
|f(x) - f_n(x)| &= \left| f^+(x) - f^-(x) - f_n^{(1)}(x) + f_n^{(2)}(x) \right| \\
&= \left| f^+(x) - f_n^{(1)}(x) - \left(f^-(x) - f_n^{(2)}(x) \right) \right| \\
&\leq \left| f^+(x) - f_n^{(1)}(x) \right| + \left| f^-(x) - f_n^{(2)}(x) \right| \\
&\to 0 + 0
\end{aligned}$$

Hence proved. □

3.3.1 Replacing *simple* functions by *step* functions

We now prove a similar result akin to Proposition 3.3.6, where we show that any measurable function can be approximated by a sequence of step functions, almost everywhere. But before that, we prove a basic fact about Lebesgue measurable sets with finite measure.

Proposition 3.3.8. *For any λ-measurable set E of finite measure and a given $\epsilon > 0$, there exists a finite sequence of open intervals $\{I_n\}_{n=1}^{N}$ such that*

$$\lambda\left(E\Delta\left(\bigcup_{n=1}^{N} I_n\right)\right) < \epsilon.$$

Proof. Take any $\epsilon > 0$, then we have for any set $E \subseteq \mathbb{R}$, a sequence of open intervals $\{I_n\}$ such that $E \subseteq \bigcup_n I_n$ and $\lambda\left(\bigcup_n I_n\right) \leq \lambda(E) + \epsilon$ or $\lambda\left(\bigcup_n I_n \setminus E\right) \leq \epsilon < 2\epsilon$. Now since $\{I_n\}$ is a disjoint sequence, therefore, $\lambda\left(\bigcup_n I_n\right) = \sum_n \lambda(I_n)$ and due to the fact that $\lambda(E) < +\infty$, we get that $\sum_n \lambda(I_n) < +\infty$.
Now, since $\lambda(E) < +\infty$, therefore the sum $\sum_n \lambda(I_n) < +\infty$, hence, $\exists N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} \lambda(I_n) < \epsilon$. With this N, we now see that:

$$\begin{aligned}
\lambda\left(E\Delta\bigcup_{n=1}^{N} I_n\right) &= \lambda\left(E\setminus\bigcup_{n=1}^{N} I_n\right) + \lambda\left(\bigcup_{n=1}^{N} I_n \setminus E\right) \qquad \text{(both are disjoint.)}\\
&\leq \lambda\left(E\setminus\bigcup_{n=1}^{N} I_n\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right)\\
&= \lambda\left(E\setminus\bigcup_{n=1}^{N} I_n\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right)\\
&= \lambda\left(E\cap\left(\bigcup_{n=1}^{N} I_n\right)^{c}\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right)\\
&\leq \lambda\left(\bigcup_{n=N+1}^{\infty} I_n\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right) \qquad \because\ E\cap\left(\bigcup_{n=1}^{N} I_n\right)^{c} \subseteq \bigcup_{n=N+1}^{\infty} I_n\\
&\leq \epsilon + \epsilon = 2\epsilon
\end{aligned}$$

Hence, we get that for any finite Lebesgue measurable set E, for all $\epsilon > 0$, $\exists$ a sequence of open intervals $\{I_n\}_{n=1}^{N}$ such that their symmetric difference is a set with measure $\leq \epsilon$. □

Proposition 3.3.9. *Consider $(\mathbb{R}, \mathcal{M}_{\lambda^*})$ to be the Lebesgue measurable space and $A \in \mathcal{M}_{\lambda^*}$. Let $f : A \to [-\infty, +\infty]$ be a λ-measurable function. Then there exists a sequence of step functions $\{\phi_k\}$ such that*

$$\phi_k \longrightarrow f \text{ almost everywhere.}$$

Proof. We will prove first that for any characteristic function, there exists a sequence of step functions converging to it. Let $g = \chi_A$ be the characteristic function on A. Continuing from Proposition 3.3.8, we see that if we write the **step-function** ϕ as

$$\psi = \sum_{k=1}^{N} \chi_{I_k}$$

where $\{I_k\}$ is the set of open intervals such that $\lambda\left(A\Delta\left(\bigcup_{n=1}^{N} I_n\right)\right) < \epsilon$ for a given $\epsilon > 0$, from Proposition 3.3.8, then we get that the set $\{x \mid g(x) \neq \psi(x)\}$ has upper bound on it's measure given as follows:

$$\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) = \psi(x)\right\} \subseteq A \cap \bigcup_{k=1}^{N} I_k \quad \because g(x) = \psi(x) \text{ iff } x \in \cup_{k=1}^{N} I_k \text{ and } 0 \text{ for other } x \in A$$
$$\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\} \supseteq \left(A \cap \bigcup_{k=1}^{N} I_k\right)^{\mathrm{c}} \supseteq A\Delta \bigcup_{k=1}^{N} I_k.$$

Similarly, it's easy to see that for any x such that $g(x) \neq \psi(x)$, we have $x \in A\Delta \bigcup_{k=1}^{N} I_k$ so that we get,

$$\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\} \subseteq A\Delta \bigcup_{k=1}^{N} I_k.$$

Hence,

$$\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\} = A\Delta \bigcup_{k=1}^{N} I_k.$$

Therefore,

$$\lambda\left(\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\}\right) < \epsilon$$

Therefore, for every $n \geq 1$, there exists a step-function ψ_n so that the set $E_n = \left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi_n(x)\right\}$ is such that

$$\lambda(E_n) < \frac{1}{2^n}.$$

Now, **define** the following two sets:

$$F_n = \bigcup_{j=n+1}^{\infty} E_j \quad \text{(a decreasing sequence)}$$
$$F = \bigcap_{k=1}^{\infty} F_k.$$

For the set F_n, observe that

$$\begin{aligned}\lambda(F_n) = \lambda\left(\bigcup_{j=n+1}^{\infty}\right) &\leq \sum_{j=n+1}^{\infty} \lambda(E_j) \\ &< \sum_{j=n+1}^{\infty} \frac{1}{2^j} \\ &= \frac{1}{2^n}\end{aligned}$$

and for set F,

$$\lambda(F) = \lambda\left(\bigcap_{k=1}^{\infty} F_k\right) = \varprojlim_{k\to\infty} \lambda(F_k) \quad \because \{F_k\} \text{ is measurable \& decreasing.}$$
$$= 0.$$

Note that $\{F_k\}$ is measurable because any E_i is itself measurable because of Proposition 3.0.4, Statement 5.
Now,

$$\psi_n(x) \longrightarrow g(x)\ \forall\ x \in F^{\mathrm{c}}$$

because $F^{\mathrm{c}} = F_1^{\mathrm{c}}$ since $\{F_k\}$ is a decreasing sequence, therefore F^{c} is the set where $g(x)$ satisfies with the limit step function.
Finally, $\psi_n \not\to f\ \forall\ x \in F$, but since $\lambda(F) = 0$, hence

$$\psi_n \longrightarrow g \text{ almost everywhere}$$

Now, what we have proved so far is that for any characteristic function $g = \chi_A$ on a measurable set, there exists a sequence of step functions converging to it point-wise almost everywhere. Since from Proposition 3.3.6, there exists a sequence of simple functions converging to f, and since a simple function $h = \sum_{i=1}^{M} \alpha_i \chi_{E_i}$ is a finite combination of characteristics functions over measurable sets, therefore there exists a sequence of step functions converging to f almost everywhere.
In particular if $\psi_n^i \longrightarrow \chi_{E_i}$ almost everywhere, then $\sum_{i=1}^{M} \alpha_i \psi_n^i \longrightarrow \sum_{i=1}^{M} \alpha_i \chi_{E_i} = h$. Now by Proposition 3.3.6, there exists the sequence $\{h_n\}$ of simple functions converging to f. Since

$$K_n = \left\{\sum_{i=1}^{M_n} \alpha_i \psi_n^i\right\} \longrightarrow h_n \text{ almost everywhere,}$$

where note that $K_n = \sum_{i=1}^{M_n} \alpha_i \psi_n^i$ is a step function because ψ_n^i is a step function and there are finitely many (M_n) of them, and

$$\{h_n\} \longrightarrow f$$

therefore,

$$K_n \longrightarrow f \text{ almost everywhere.}$$

Hence proved. $\square$

3.4 Egorov's theorem

We now discuss a very important result in the theory of measurable functions named after Dmitri Fyodorovich Egorov, who published this result in 1911, thus establishing **a condition required for uniform convergence of a point-wise convergent sequence of measurable functions.**

Theorem 3.4.1. ***(Egorov's theorem)*** *Let $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ be the Lebesgue measure space on $\mathbb{R}$. Suppose $\{f_k\}$ is a sequence of real-valued, Lebesgue measurable functions on $E \in \mathcal{M}_{\lambda^*}$ where $\lambda(E) < +\infty$. If*

$$f_k \longrightarrow f \text{ pointwise on } E,$$

[14] *Then for each $\epsilon > 0$, there exists a* ***closed*** *set $A_\epsilon \subset E$ such that*

[14] From Proposition 3.0.9, the limit of a sequence of measurable functions is also measurable, hence there's no point in writing extraneously the requirement for f to be also measurable.

1. $\lambda(E \setminus A_\epsilon) < \epsilon$*, and*
2. $f_k \longrightarrow f$ ***uniformly*** *on* A_ϵ*.*

Proof. We break down the proof in the following 3 parts.

Act 1. *A Basic Construction.*
For each pair of integers n, k, construct the following set:

$$E_k^n = \left\{ x \in E \ : \ |f_j(x) - f(x)| < \frac{1}{n}\,, \ \ \forall\, j > k \right\}.$$

Now, fix n, so that we have the following observations:

$$E_k^n \subseteq E_{k+1}^n \tag{8}$$

and since $f_k \longrightarrow f$ point-wise, therefore

$$\lim_{\overleftarrow{k\to\infty}} \bigcup_{i=1}^{k} E_i^n = E. \tag{9}$$

Hence

$$\lambda(E \setminus E_k^n) \longrightarrow 0 \text{ as } k \to \infty.$$

Note that the above result utilizes the fact that $\lambda(E) < +\infty$. Now by the above, we can say that $\exists\, k_n$ such that

$$\lambda\left(E \setminus E_{k_n}^n\right) < \frac{1}{2^n}$$

which, by definition of E_k^n implies that

$$|f_j(x) - f(x)| < \frac{1}{n} \text{ whenever } j > k_n \text{ and } \underline{x \in E_{k_n}^n}.$$

Act 2. *Constructing* A_ϵ*.*
Now choose $N \in \mathbb{N}$ such that

$$\sum_{n=N}^{\infty} \frac{1}{2^n} < \frac{\epsilon}{2}$$

and define

$$\tilde{A}_\epsilon = \bigcap_{n=N}^{\infty} E_{k_n}^n \tag{10}$$

We now observe that

$$\begin{aligned}\lambda\left(E \setminus \tilde{A}_\epsilon\right) &= \lambda\left(E \cap \bigcup_{n=N}^{\infty}\left(E_{k_n}^n\right)^{\mathrm{c}}\right)\\ &= \lambda\left(\bigcup_{n=N}^{\infty} E \cap \left(E_{k_n}^n\right)^{\mathrm{c}}\right)\\ &\leq \sum_{n=N}^{\infty}\lambda\left(E \cap \left(E_{k_n}^n\right)^{\mathrm{c}}\right)\\ &= \sum_{n=N}^{\infty}\lambda\left(E \setminus E_{k_n}^n\right)\\ &< \sum_{n=N}^{\infty}\frac{1}{2^n}\\ &< \frac{\epsilon}{2}\end{aligned}$$

Act 3. *Finalé.*
We now claim and prove the following:

$$\text{Claim}: \; f_k \longrightarrow f \textit{ uniformly on } \tilde{A}_\epsilon.$$

For this, let $\delta > 0$ and choose $n' \geq N$ such that $\frac{1}{n'} < \delta$. Then

$$\text{if } x \in \tilde{A}_\epsilon \implies x \in E_{k_{n'}}^{n'} \implies |f_j(x) - f(x)| < \frac{1}{n'} < \delta \;, \;\; \forall\, j > k_{n'}. \tag{11}$$

Note that this is just the definition of uniform convergence.
Finally, note that E_k^n is a Lebesgue measurable set due to Proposition 3.0.4, Statement 5. Hence, $\tilde{A}_\epsilon$ is measurable. Now, by Proposition 2.9.2, Statement 2, there exists a closed set $A_\epsilon \subset \tilde{A}_\epsilon$ such that

$$\lambda\left(\tilde{A}_\epsilon \setminus A_\epsilon\right) < \frac{\epsilon}{2}.$$

Now,

$$\begin{aligned}\epsilon &> \lambda\left(E \setminus \tilde{A}_\epsilon\right) + \lambda\left(\tilde{A}_\epsilon \setminus A_\epsilon\right)\\ &\geq \lambda\left(E \setminus \tilde{A}_\epsilon \bigcup \tilde{A}_\epsilon \setminus A_\epsilon\right)\\ &= \lambda\left(E \setminus A_\epsilon\right).\end{aligned}$$

Now, by (11), we see that $f_k \longrightarrow f$ uniformly for all $x \in A_\epsilon \subset \tilde{A}_\epsilon$ such that $\lambda\left(E \setminus A_\epsilon\right) < \epsilon$ and A_ϵ is closed. Proof is now complete. $\square$

3.5 Lusin's theorem

The following is the final important result on the basic theory of measurable functions, attributed to Nikolai Nikolaevich Luzin who penned this theorem around 1912.

Theorem 3.5.1. ***(Lusin's theorem)*** *Consider the Lebesgue measure space* $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$. *Suppose* f *is a real-valued, Lebesgue measurable function defined over a Lebesgue measurable set* E *with finite measure. Then for all* $\epsilon > 0$, *there exists a* ***closed*** *set* $F_\epsilon \subset E$ *with*

1. $\lambda(E \setminus F_\epsilon) < \epsilon$, *and*
2. *The restriction* $f|_{F_\epsilon}$ *of* f *over* F_ϵ *is* ***continuous***.

Proof. From the Proposition 3.3.9, we have a sequence $\{f_n\}$ of step functions such that

$$f_n \longrightarrow f \text{ almost everywhere.}$$

Now, consider, for example the characteristic function over an interval $\chi_{[a,b]}$. Then, we can define a function $\phi(x)$ for any $\delta > 0$ as follows:

$$\phi(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{\delta/2} & a \le x \le a + \frac{\delta}{2} \\ 1 & a + \frac{\delta}{2} \le x \le b - \frac{\delta}{2} \\ \frac{b-x}{\delta/2} & b - \frac{\delta}{2} \le x \le b \\ 0 & x > b \end{cases}$$

which then satisfies

$$\{x \in \mathbb{R} \mid \phi(x) \neq \chi_{[a,b]}\} = \left(a, a + \frac{\delta}{2}\right) \bigcup \left(b - \frac{\delta}{2}, b\right)$$

which then implies that,

$$\lambda\left(\{x \in \mathbb{R} \mid \phi(x) \neq \chi_{[a,b]}\}\right) = \lambda\left(\left(a, a + \frac{\delta}{2}\right) \bigcup \left(b - \frac{\delta}{2}, b\right)\right) = \delta.$$

Note that $\phi(x)$ is also continuous over all $\mathbb{R}$. Hence, for any step function (finite sum of $\chi_{[a,b]}$-type functions) and $\delta > 0$, one can construct a continuous function which does not satisfies with the step function on a set with measure $< \delta$.

Hence, for step-functions $\{f_n\}$, corresponding to each f_n, $\exists$ a continuous function ϕ_n and a set E_n such that

$$E_n = \{x \mid \phi_n(x) \neq f_n(x)\} \text{ and } \lambda(E_n) < \frac{1}{2^n}.$$

Now, for all $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that

$$\sum_{n \ge N} \frac{1}{2^n} < \frac{\epsilon}{3}.$$

With the above fact, construct the set F' as follows:

$$F' = \left(A_{\frac{\epsilon}{3}} \setminus \bigcup_{n \ge N} E_n\right)$$

where $A_{\frac{\epsilon}{3}}$ is the closed subset $A_{\frac{\epsilon}{3}} \subset E$ such that **1.** $\lambda\left(E \setminus A_{\frac{\epsilon}{3}}\right) < \frac{\epsilon}{3}$ and **2.** $f_n \longrightarrow f$ **uniformly** on $A_{\frac{\epsilon}{3}}$. This is guaranteed by Theorem 3.4.1 (Egorov's Theorem).

Note that $f_n|_{F'}$ is **continuous** $\forall\, n \geq N$ because for any $x \in F' \implies \phi_n(x) = f_n(x)\ \forall\, n \geq N$ and since ϕ_n are already continuous $\forall n \in \mathbb{N}$.

Furthermore, since $F' \subset E$ and $f_n \longrightarrow f$ uniformly, then the restriction $f_n|_{F'}$ is continuous and converges uniformly to $f|_{F'}$, which **due to uniform convergence, is also continuous!**

Now, note that E_n's are measurable sets due to Proposition 3.0.4, Statement 5. Similarly, since $A_{\frac{\epsilon}{3}}$ is closed, therefore it is also measurable. Hence, F' is measurable.

Now by Proposition 2.9.2, there exists a closed set $F_\epsilon \subset F'$ such that $\lambda(F' \setminus F_\epsilon) < \frac{\epsilon}{3}$. Note that because $F_\epsilon \subset F'$ and $f|_{F'}$ is continuous, therefore the restriction $f|_{F_\epsilon}$ is also continuous.

Finally, combining
1. $\sum_{n\geq N} \frac{1}{2^n} < \frac{\epsilon}{3}$,
2. $\lambda\left(E \setminus A_{\frac{\epsilon}{3}}\right) < \frac{\epsilon}{3}$,
3. $\lambda(F' \setminus F_\epsilon) < \frac{\epsilon}{3}$

it can be easily seen that

$$\left(E \setminus A_{\frac{\epsilon}{3}}\right) \bigcup (F' \setminus F_\epsilon) = E \setminus F_\epsilon$$

Hence,

$$\begin{aligned}
\lambda(E \setminus F_\epsilon) &= \lambda\left(\left(E \setminus A_{\frac{\epsilon}{3}}\right) \bigcup (F' \setminus F_\epsilon)\right) \\
&\leq \lambda\left(\left(E \setminus A_{\frac{\epsilon}{3}}\right)\right) + \lambda\left((F' \setminus F_\epsilon)\right) \\
&< \frac{2\epsilon}{3} \\
&< \epsilon.
\end{aligned}$$

which completes the proof. $\square$

3.6 Applications-I : Measure spaces and measurable functions

We now present applications of the above theory. This is, in particular, to showcase the true strength of abstract analysis. This can also be used to strengthen one's intuition about the topic.

3.6.1 σ-algebras and measure spaces

Lemma 3.6.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Prove that μ is σ-finite if and only if there exists a countable disjoint family of measurable sets $\{A_n\}$ such that $X = \coprod_n A_n$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.*

Proof. Note that R $\implies$ L is immediate from definition. Let μ be σ-finite. Then there exists $\{B_n\} \subseteq \mathcal{A}$ such that $\mu(B_n) < \infty$ and $\bigcup_n B_n = X$. Define $A_1 = B_1$ and $A_n = B_n \setminus B_1 \cup \cdots \cup B_{n-1}$.

As $\mathcal{A}$ is a σ-algebra, so $\{A_n\} \subseteq \mathcal{A}$. Moreover, $A_n \cap A_m = \emptyset$ for all $n \neq m$ because if $m > n$[15] and $x \in A_m \cap A_n$, then $x \in B_m \setminus B_1 \cup \cdots \cup B_n \cup \ldots B_{m-1}$ and $x \in B_n \setminus B_1 \cup \ldots B_{n-1}$, a contradiction. As $A_n \subseteq B_n$, therefore $\mu(A_n) \leq \mu(B_n) < \infty$. To complete the proof, we need only show that $\bigcup_n A_n = \bigcup_n B_n$.

Pick any $x \in \bigcup_n A_n$. Then $x \in B_n \setminus B_1 \cup \cdots \cup B_{n-1}$ for some $n \in \mathbb{N}$. Thus, $x \in B_n$ and hence $x \in \bigcup_n B_n$. Conversely, pick $x \in \bigcup_n B_n$. Then $x \in B_n$ for some $n \in \mathbb{N}$. Now, either $x \in B_n \setminus B_1 \cup \cdots \cup B_{n-1}$ or $x \in B_n \cap (B_1 \cup \cdots \cup B_{n-1})$. If the former is true, then $x \in A_n$ and we are done. If the latter is true, then we may assume $x \in B_{n-1} \cap B_n$. Now again either $x \in B_{n-1} \setminus B_1 \cup \cdots \cup B_{n-2}$ or $x \in B_n \cap B_{n-1} \cap (B_1 \cup \cdots \cup B_{n-2})$. Repeating this process inductively, we will end up in either of the following cases:

1. $x \in A_k$ for some $1 \leq k \leq n$,
2. $x \in B_1 \cap \cdots \cap B_n$.

As $B_1 = A_1$ by construction, therefore in either case we are done. $\square$

Lemma 3.6.2. *Given $\mathcal{S} \subseteq \mathcal{P}(X)$, denote by $\mathcal{A}(\mathcal{S})$ the σ-algebra generated by $\mathcal{S}$. Then,*

$$\mathcal{A}(\mathcal{S}) = \mathcal{A}(\mathcal{A}(\mathcal{S})).$$

Proof. Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$ be an arbitrary collection of subsets of X. If X is empty then the statement is vacuously true, so let X be non-empty. Since the σ-algebra generated by $\mathcal{A}(\mathcal{S})$ is the intersection of all σ-algebras containing $\mathcal{A}(\mathcal{S})$, therefore we have that $\mathcal{A}(\mathcal{A}(\mathcal{S})) = \bigcap_{\mathcal{C} \supseteq \mathcal{A}(\mathcal{S})} \mathcal{C}$. Since $\mathcal{A}(S)$ is a σ-algebra containing $\mathcal{A}(\mathcal{S})$, therefore $\mathcal{A}(\mathcal{A}(\mathcal{S})) \subseteq \mathcal{A}(\mathcal{S})$. Since $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{C}$ for all σ-algebras $\mathcal{C}$ containing $\mathcal{A}(\mathcal{S})$, therefore $\mathcal{A}(\mathcal{A}(\mathcal{S})) \supseteq \mathcal{A}(\mathcal{S})$. $\square$

Lemma 3.6.3. *Let $\mathcal{A}(\mathcal{S})$ be the σ-algebra generated by a set $\mathcal{S} \subseteq \mathcal{P}(X)$. Then, $\mathcal{A}(\mathcal{S})$ is the union of the σ-algebras generated by Y as Y ranges over all countable subsets of $\mathcal{S}$.*

Proof. Let X be a non-empty set and $\mathcal{S} \subseteq \mathcal{P}(X)$. We wish to show that

$$\mathcal{A}(\mathcal{S}) = \bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y}).$$

Let $\mathcal{Y} \subseteq \mathcal{S}$ be a countable subcollection. Then, $\mathcal{A}(\mathcal{Y}) \subseteq \mathcal{A}(\mathcal{S})$. Consequently, $\bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y}) \subseteq \mathcal{A}(\mathcal{S})$. Conversely, we need to show that

$$\mathcal{A}(\mathcal{S}) \subseteq \bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y}).$$

We claim that $\bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y})$ is a σ-algebra containing $\mathcal{S}$. This would complete the proof as $\mathcal{A}(\mathcal{S})$ is the smallest σ-algebra containing $\mathcal{S}$.

Denote $\mathcal{Z} = \bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y})$. As $\mathcal{A}(\mathcal{Y})$s are σ-algebras, therefore $\mathcal{Z}$ contains X and $\emptyset$. Let $A \in \mathcal{Z}$. Then $A \in \mathcal{A}(\mathcal{Y})$ for some $\mathcal{Y} \subseteq \mathcal{S}$ countable. Consequently, $A^c \in \mathcal{A}(\mathcal{Y})$ and thus $A^c \in \mathcal{Z}$. Let $\{A_n\} \subseteq \mathcal{Z}$ be a countable collection of sets. Then $A_n \in \mathcal{A}(\mathcal{Y}_n)$ for all $n \in \mathbb{N}$. Further, we have that $\mathcal{Y}_k \subseteq \mathcal{A}\left(\bigcup_n \mathcal{Y}_n\right)$ for all $k \in \mathbb{N}$ as $\mathcal{Y}_k \subseteq \bigcup_n \mathcal{Y}_n$. As $\mathcal{Y}_k$ are countable and countable union of countable sets is countable, therefore $\bigcup_n \mathcal{Y}_n$ is countable. Thus, we have

$$A_k \in \mathcal{A}(\mathcal{Y}_k) \subseteq \mathcal{A}\left(\bigcup_n \mathcal{Y}_n\right) \subseteq \mathcal{Z}\ \forall k \in \mathbb{N}.$$

[15] which we may assume wlog.

Thus from above, we obtain that

$$\bigcup_k A_k \in \left(\bigcup_n \mathcal{Y}_n\right) \subseteq \mathcal{Z}.$$

Hence, $\mathcal{Z}$ is a σ-algebra. To complete the proof, we need only show that $\mathcal{Z}$ contains $\mathcal{S}$.

Let $A \in \mathcal{S}$. Then since $\{A\}$ is a countable subset of $\mathcal{S}$, therefore $\mathcal{A}(\{A\})$ is contained in $\mathcal{Z}$ and thus $A \in \mathcal{Z}$. $\square$

Lemma 3.6.4. *The σ-algebra generated by*

1. $\mathcal{S} = \{(a, b] \mid a < b \in \mathbb{Q}\}$,
2. $\mathcal{S} = \{(a, n] \mid a \in \mathbb{Q}, n \in \mathbb{Z}\}$,

is the Borel σ-algebra on $\mathbb{R}$.

Proof. 1. Let $\mathcal{S} = \{(a, b] \mid a, b \in \mathbb{Q}\}$. We wish to show that $\mathcal{A}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the Borel σ-algebra of $\mathbb{R}$. Since $(a, b]$ for $a, b \in \mathbb{Q}$ is contained in $\mathcal{B}$ as $(a, b] = (a, b) \cup \bigcap_{n\in\mathbb{N}}(b - 1/n, b + 1/n)$, therefore $\mathcal{S} \subseteq \mathcal{B}$. Consequently, $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{B}$ as $\mathcal{B}$ is the smallest σ-algebra containing open intervals.

Since we also know that $\mathcal{B}$ is generated by the collection of all closed intervals $[a, b]$ in $\mathbb{R}$, therefore to show that $\mathcal{B} \subseteq \mathcal{A}(\mathcal{S})$, it would suffice to show $[a, b] \in \mathcal{A}(\mathcal{S})$ where $a < b$ in $\mathbb{R}$. Pick $a < b$ in $\mathbb{R}$. By density of $\mathbb{Q}$, we may pick $\{a_n\}$ to be an increasing sequence such that $a_n \in \mathbb{Q}$, $a_n < a$ and $\lim_{n\to\infty} a_n = a$. Similarly, we may pick a decreasing sequence $\{b_n\}$ such that $b_n \in \mathbb{Q}$, $b_n > b$ and $\lim_{n\to\infty} b_n = b$. Consequently, we claim that

$$[a, b] = \bigcap_n (a_n, b_n]$$

where $(a_n, b_n] \in \mathcal{S}$. Indeed, $(\subseteq)$ is clear. For $(\supseteq)$, take $x \in \bigcap_n (a_n, b_n]$. Hence $a_n < x \leq b_n$. Taking $n \to \infty$, we get $a \leq x \leq b$ as desired. Thus, $[a, b] \in \mathcal{A}(\mathcal{S})$.

2. Let $\mathcal{S} = \{(a, n] \mid a \in \mathbb{Q},\ n \in \mathbb{N}\}$. We wish to show that $\mathcal{A}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the Borel σ-algebra of $\mathbb{R}$. Since $(a, n]$ for $a \in \mathbb{Q}$ and $n \in \mathbb{N}$ is contained in $\mathcal{B}$ as $(a, n] = (a, n) \cup \bigcap_{k\in\mathbb{N}}(n - 1/k, n + 1/k)$, therefore $\mathcal{S} \subseteq \mathcal{B}$. Consequently, $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{B}$.

Since we also know that $\mathcal{B}$ is generated by the collection of all open intervals of the form (a, ∞), $a \in \mathbb{R}$, therefore to show that $\mathcal{B} \subseteq \mathcal{A}(\mathcal{S})$, it would suffice to show $(a, \infty) \in \mathcal{A}(\mathcal{S})$ for all $a \in \mathbb{R}$. Pick (a, ∞) for some $a \in \mathbb{R}$. By density of $\mathbb{Q}$, there exists a decreasing sequence $\{a_n\}$ in $\mathbb{R}$ such that $a_n \in \mathbb{Q}$, $a_n > a$ and $\lim_{n\to\infty} a_n = a$. Consequently, we claim that

$$(a, \infty) = \bigcup_n (a_n, n]$$

where $(a_n, n] \in \mathcal{S}$. Indeed, for $(\subseteq)$, take $x \in (a, \infty)$. We therefore have $a < x < \infty$. As $\lim_{n\to\infty} a_n = a$ and $a_n > a$ for all $n \in \mathbb{N}$, therefore there exists $N \in \mathbb{N}$ such that $a < a_n \leq a_N < x$ for all $n \geq N$. Consequently, for some large $n \in \mathbb{N}$ greater than N such that $x \leq n$, we obtain $a_n < x \leq n$ and hence $x \in (a_n, n]$. For $(\supseteq)$, take $x \in \bigcup_n (a_n, n]$ and thus we get $a < a_n < x \leq n < \infty$. Thus, $(a, \infty) \in \mathcal{A}(\mathcal{S})$. $\square$

Lemma 3.6.5. *The Borel σ-algebra on $\mathbb{R}^2$ is generated by*

$$\{(I \times \mathbb{R}) \cup (\mathbb{R} \times J) \mid I, J \subseteq \mathbb{R},\ \textit{open intervals}\}.$$

Proof. Let $\mathcal{S} = \{(I \times \mathbb{R}) \cup (\mathbb{R} \times J) \mid I, J \subseteq \mathbb{R} \text{ is open}\}$. We wish to show that $\mathcal{A}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the σ-algebra of $\mathbb{R}^2$.

As $\mathcal{S}$ is a collection of open sets of $\mathbb{R}^2$ and $\mathcal{B}$ is generated by all open sets of $\mathbb{R}^2$, therefore $\mathcal{S} \subseteq \mathcal{B}$ and thus $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{B}$.

We now wish to show that $\mathcal{B} \subseteq \mathcal{A}(\mathcal{S})$. It would suffice to show that any open set $U \subseteq \mathbb{R}^2$ is in $\mathcal{A}(\mathcal{S})$. Note that $\mathcal{A}(\mathcal{S})$ consists of all open rectanlges $I \times J = (I \times \mathbb{R}) \cap (\mathbb{R} \times J)$. Thus, it would suffice to show that U can be written as countable union of open rectangles. Recall that open rectangles forms a basis for the usual topology on $\mathbb{R}^2$. Consider the collection of all open rectangles K inside U whose vertices have both rational coordinates. We claim that the union of such open rectangles is equal to U. Indeed, their union is inside U and for any $x \in U$, there exists an open ball $x \in B \subseteq U$, so there exists an open rectangle K inside B which contains x and has vertices which have both rational coordinates. Thus U is equal to the union of all such rectangles. Since there are only countably many such open rectangles as they are parameterized by choice of 4 points in $\mathbb{Q}^2 \cap U$ which is atmost countably many, therefore we have obtained a countable cover of U by open rectangles. This completes the proof. $\square$

Lemma 3.6.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $A, B \in \mathcal{A}$. Then,*

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

Proof. Observe that we can write

$$A \cup B = (A \setminus (A \cap B)) \cup B$$

where the right side is a disjoint union. Consequently, we have

$$\mu(A \cup B) = \mu(A \setminus A \cap B) + \mu(B). \tag{6.1}$$

We now have two cases. If $\mu(A \cap B) = \infty$, then since $\mu(A \cap B) \leq \mu(A), \mu(B)$ and $\mu(A) \leq \mu(A \cup B)$, therefore we get $\mu(A \cup B) = \mu(A \cap B) = \mu(A) = \mu(B) = \infty$, so that the statement to be proven is a tautology. Else if $\mu(A \cap B) < \infty$, then we can write

$$\mu(A \setminus A \cap B) = \mu(A) - \mu(A \cap B).$$

Consequently, by Eq. (6.1) and the fact that $\mu(A \cap B) < \infty$, we have

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

This completes the proof. $\square$

Lemma 3.6.7. *Let $x \in \mathbb{R}$ and let B be a Borel subset of $\mathbb{R}$. Then, $x + B$ and xB are Borel subsets of $\mathbb{R}$ (that is, Borel subsets of $\mathbb{R}$ are translation and dilation invariant).*

Proof. 1. Let $x \in \mathbb{R}$ and $\mathcal{B}$ be the Borel σ-algebra of $\mathbb{R}$. We wish to show that for all $B \in \mathcal{B}$, the translate $x + B \in \mathcal{B}$. Consequently, we wish to show

$$x + \mathcal{B} \subseteq \mathcal{B}$$

where $x + \mathcal{B} = \{x + B \mid B \in \mathcal{B}\}$. We use the following standard technique to show this.

Consider the following collection

$$\mathcal{C} = \{B \in \mathcal{B} \mid x + B \in \mathcal{B}\}.$$

Our goal is to show that $\mathcal{C} = \mathcal{B}$. Note that $\mathcal{C} \subseteq \mathcal{B}$. Conversely, we wish to show that $\mathcal{B} \subseteq \mathcal{C}$. This would follow immediately if we show that $\mathcal{C}$ is a σ-algebra containing all open intervals, as $\mathcal{B}$ is the σ-algebra generated by all open intervals.

We now establish that $\mathcal{C}$ is a σ-algebra. Since $x + \mathbb{R} = \mathbb{R}$ and $x + \emptyset = \emptyset$, therefore $\mathbb{R}, \emptyset \in \mathcal{C}$. Let $A \in \mathcal{C}$. We wish to show that $A^c \in \mathcal{C}$. Since $x + A \in \mathcal{B}$, therefore $(x + A)^c \in \mathcal{B}$. Thus it suffices to show that $(x + A)^c = x + A^c$. Indeed, we have the following equalities

$$\begin{aligned}
(x + A)^c &= \{y \in \mathbb{R} \mid y \notin x + A\} \\
&= \{y \in \mathbb{R} \mid y - x \notin A\} \\
&= \{y \in \mathbb{R} \mid y - x \in A^c\} \\
&= \{y \in \mathbb{R} \mid y \in x + A^c\} \\
&= x + A^c.
\end{aligned}$$

Let $\{A_n\} \subseteq \mathcal{C}$. We wish to show that $\bigcup_n A_n \in \mathcal{C}$. We have that for each $n \in \mathbb{N}$, $x + A_n \in \mathcal{B}$. It would thus suffice to show that

$$x + \bigcup_n A_n = \bigcup_n (x + A_n).$$

Indeed, take $x + a \in x + \bigcup_n A_n$. Hence $a \in A_n$ for some $n \in \mathbb{N}$. Consequently, $x + a \in x + A_n$. Thus $x + a \in \bigcup_n (x + A_n)$. Conversely, let $z \in \bigcup_n (x + A_n)$. Then $z = x + y_n$ for $y_n \in A_n$. Consequently, $z \in x + \bigcup_n A_n$. This show that $\mathcal{C}$ is a σ-algebra.

To complete the proof, we now need only show that $\mathcal{C}$ has all open intervals. This is immediate, as we show now. Take any $(a, b) \subseteq \mathbb{R}$. Since $x + (a, b) = (x + a, x + b) \in \mathcal{B}$, therefore $(a, b) \in \mathcal{C}$.

2. Let $x \in \mathbb{R}$ and $\mathcal{B}$ be the Borel σ-algebra of $\mathbb{R}$. We wish to show that for all $B \in \mathcal{B}$, the dilate $x \cdot B \in \mathcal{B}$. Note that $x \cdot B = \{xb \mid b \in B\}$. Consequently, we wish to show

$$x \cdot \mathcal{B} \subseteq \mathcal{B}$$

where $x \cdot \mathcal{B} = \{x \cdot B \mid B \in \mathcal{B}\}$. If $x = 0$, then $x \cdot \mathcal{B} = \{0\}$ and that is trivially inside $\mathcal{B}$ as $\{0\} = \bigcap_n (-1/n, 1/n)$. Thus we now assume that $x \neq 0$. We use the following standard technique to show the above inclusion.

Consider the following collection

$$\mathcal{C} = \{B \in \mathcal{B} \mid x \cdot B \in \mathcal{B}\}.$$

Our goal is to show that $\mathcal{C} = \mathcal{B}$. Note that $\mathcal{C} \subseteq \mathcal{B}$. Conversely, we wish to show that $\mathcal{B} \subseteq \mathcal{C}$. This would follow immediately if we show that $\mathcal{C}$ is a σ-algebra containing all open intervals, as $\mathcal{B}$ is the σ-algebra generated by all open intervals.

We now establish that $\mathcal{C}$ is a σ-algebra. Observe that $x \cdot \mathbb{R} = \mathbb{R}$. Indeed, as $x \cdot \mathbb{R} \subseteq \mathbb{R}$ is clear, we can also write any $a \in \mathbb{R}$ as $x \cdot x^{-1} a$. We also have $x \cdot \emptyset = \emptyset$. Therefore $\mathbb{R}, \emptyset \in \mathcal{C}$. Let $A \in \mathcal{C}$.

We wish to show that $A^c \in \mathcal{C}$. Since $x \cdot A \in \mathcal{B}$, therefore $(x \cdot A)^c \in \mathcal{B}$. Thus it suffices to show that $(x \cdot A)^c = x \cdot A^c$. Indeed, we have the following equalities

$$\begin{aligned}(x \cdot A)^c &= \{y \in \mathbb{R} \mid y \notin x \cdot A\} \\ &= \{y \in \mathbb{R} \mid x^{-1}y \notin A\} \\ &= \{y \in \mathbb{R} \mid x^{-1}y \in A^c\} \\ &= \{y \in \mathbb{R} \mid y \in x \cdot A^c\} \\ &= x \cdot A^c.\end{aligned}$$

Let $\{A_n\} \subseteq \mathcal{C}$. We wish to show that $\bigcup_n A_n \in \mathcal{C}$. We have that for each $n \in \mathbb{N}$, $x \cdot A_n \in \mathcal{B}$. It would thus suffice to show that

$$x \cdot \bigcup_n A_n = \bigcup_n (x \cdot A_n).$$

Indeed, take $x \cdot a \in x \cdot \bigcup_n A_n$. Hence $a \in A_n$ for some $n \in \mathbb{N}$. Consequently, $x \cdot a \in x \cdot A_n$. Thus $x \cdot a \in \bigcup_n (x \cdot A_n)$. Conversely, let $z \in \bigcup_n (x \cdot A_n)$. Then $z = x \cdot y_n$ for $y_n \in A_n$. Consequently, $z \in x \cdot \bigcup_n A_n$. This show that $\mathcal{C}$ is a σ-algebra.

To complete the proof, we now need only show that $\mathcal{C}$ has all open intervals. This is immediate, as we show now. Take any $(a, b) \subseteq \mathbb{R}$. If $x > 0$, then we have $x \cdot (a, b) = (x \cdot a, x \cdot b) \in \mathcal{B}$, therefore $(a, b) \in \mathcal{C}$. If $x < 0$, then we have $x \cdot (a, b) = (x \cdot b, x \cdot a) \in \mathcal{B}$, therefore $(a, b) \in \mathcal{C}$. □

Lemma 3.6.8. *Let $(X, \mathcal{A})$ be a measurable space and let $\{\mu_i\}_{i=1}^n$ be a finite collection of measures on $(X, \mathcal{A})$. If $r_1, \dots, r_n \in \mathbb{R}_{\geq 0}$, then $\sum_i r_i \mu_i$ is a measure on $(X, \mathcal{A})$ (that is, positive linear combination of measures is a measure).*

Proof. Let $(X, \mathcal{A})$ be a measurable space and $\{\mu_i\}_{i=1}^n$ be a collection of measures on it. Let $\{r_i\}_{i=1}^n \subseteq \mathbb{R}_{\geq 0}$. We wish to show that $\mu = \sum_{i=1}^n r_i \mu_i$ is a measure on $(X, \mathcal{A})$. First we may assume that each $r_i > 0$ as if any $r_j = 0$, then $\mu(A) = \sum_{i=1}^n r_i \mu_i(A) = \sum_{i \neq j} r_i \mu_i(A) + r_j \mu_j(A)$, therefore if $\mu_j(A) < \infty$, then $r_j \mu_j(A) = 0$ and if $\mu_j(A) = \infty$, then since $0 \cdot \infty = 0$, therefore still $r_j \mu_j(A) = 0$. Further, if all $r_i = 0$, then $\mu = 0$, which is the trivial measure. Consequently, we assume that $r_i > 0$ for all $i = 1, \dots, n$.

We now show that μ is a measure on $(X, \mathcal{A})$. We have $\mu(\emptyset) = \sum_{i=1}^n r_i \mu_i(\emptyset) = \sum_{i=1}^n r_i \cdot 0 = 0$. Let $\{A_n\} \subseteq \mathcal{A}$ be a collection of disjoint measurable sets. We wish to show that

$$\mu\left(\coprod_k A_k\right) = \sum_k \mu(A_k).$$

We have

$$\begin{aligned}\mu\left(\coprod_k A_k\right) &= \sum_{i=1}^n r_i \mu_i \left(\coprod_k A_k\right) \\ &= \sum_{i=1}^n r_i \sum_{k=1}^\infty \mu_i(A_k).\end{aligned}$$

We now claim that

$$\sum_{i=1}^n r_i \sum_{k=1}^\infty \mu_i(A_k) = \sum_{k=1}^\infty \sum_{i=1}^n r_i \mu_i(A_k) \tag{8.1}$$

and showing this will complete the proof as

$$\sum_{k=1}^{\infty}\sum_{i=1}^{n} r_i\mu_i(A_k) = \sum_{k=1}^{\infty}\mu(A_k).$$

We have few cases for establishing Eq. (8.1).

1. *If for all $i = 1,\ldots,n$, the series $\sum_{k=1}^{\infty}\mu_i(A_k)$ is finite.* Then, $\sum_{i=1}^{n} r_i\sum_{k=1}^{\infty}\mu_i(A_k) = \sum_{i=1}^{n}\sum_{k=1}^{\infty} r_i\mu_i(A_k)$. Now, if $\sum_n x_n, \sum_n y_n$ are two convergent positive series, then their linear combination $c\sum_n x_n + d\sum_n y_n$ is equal to $\sum_n cx_n + dy_n$, where $c, d \in \mathbb{R}_{\geq 0}$. Indeed, this follows at once from the equality $c\lim_{n\to\infty}\sum_{k=1}^{n} x_k + d\lim_{n\to\infty}\sum_{k=1}^{n} y_k = \lim_{n\to\infty}\sum_{k=1}^{n} cx_k + dy_k$, which follows from the fact that both the limit exists and $c, d \in \mathbb{R}$. Consequently, we have
$$\sum_{i=1}^{n}\sum_{k=1}^{\infty} r_i\mu_i(A_k) = \sum_{k=1}^{\infty}\sum_{i=1}^{n} r_i\mu_i(A_k),$$
which is what we needed.
2. *If there exists $i_0 = 1,\ldots,n$ such that the series $\sum_{k=1}^{\infty}\mu_{i_0}(A_k) = \infty$.* In this case, in the Eq. (8.1), the left side is ∞. The right side is also infinity as shown below:
$$\begin{aligned}\sum_{k=1}^{\infty}\sum_{i=1}^{n} r_i\mu_i(A_k) &\geq \sum_{k=1}^{\infty} r_{i_0}\mu_{i_0}(A_k)\\ &= \infty\end{aligned}$$
where the first inequality follows from $r_i > 0$ for all $i = 1,\ldots,n$ and measure being positive by definition. Consequently, Eq. (8.1) follows in this case as well.

This completes the proof. □

Lemma 3.6.9. *For any set X and a subset $S \subseteq X$, the collection*

$$\mathcal{A}_S = \{A \subseteq X \mid A \subseteq S \text{ or } A^c \subseteq S\}$$

is a σ-algebra on X.

Proof. Let X be a non-empty set, $S \subseteq X$ and define

$$\mathcal{A}_S := \{A \subseteq X \mid A \subseteq S \text{ or } A^c \subseteq S\}.$$

We claim that this forms a σ-algebra on X. As $X^c = \emptyset \subseteq S$, therefore $X \in \mathcal{A}_S$ and $\emptyset \in \mathcal{A}_S$. Let $A \in \mathcal{A}_S$. If $A \subseteq S$, then A^c is such that $(A^c)^c = A \subseteq S$, so $A^c \in \mathcal{A}_S$. If $A^c \subseteq S$, then A^c is such that $A^c \subseteq S$, so $A^c \in \mathcal{A}_S$. So in both cases $\mathcal{A}_S$ is closed uncer complements.

Let $\{A_n\} \subseteq \mathcal{A}_S$ be a collection of subsets. We wish to show that $\bigcup_n A_n \in \mathcal{A}_S$. We have three cases.

C1. *$A_n \subseteq S$ for all $n \in \mathbb{N}$.* Then $\bigcup_n A_n \subseteq S$ and thus $\bigcup_n A_n \in \mathcal{A}_S$.

C2. *$\exists A_m$ such that $A_m \not\subseteq S$.* Then $A_m^c \subseteq S$. We then observe by De-Morgan's law that
$$\left(\bigcup_n A_n\right)^c = \bigcap_n A_n^c \subseteq A_m^c \subseteq S.$$
Consequently, $\bigcup_n A_n \in \mathcal{A}_S$.

C3. *$A_n \not\subseteq S$ for all $n \in \mathbb{N}$.* Then $A_n^c \subseteq S$ for all $n \in \mathbb{N}$. We again observe by De-Morgan's law that

$$\left(\bigcup_n A_n\right)^c = \bigcap_n A_n^c \subseteq A_m^c \subseteq S \; \forall m \in \mathbb{N}.$$

Consequently, $\bigcup_n A_n \in \mathcal{A}_S$.

In all three cases, $\bigcup_n A_n \in \mathcal{A}_S$. Hence $\mathcal{A}_S$ is a σ-algebra. □

Lemma 3.6.10. *Let $(X, \mathcal{A}, \mu)$ be a semifinite measure space, and let $\mu(A) = \infty$ for some $A \in \mathcal{A}$. If $M > 0$, then there exists $B \subseteq A$ such that $M < \mu(B) < \infty$.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a semi-finite measure space and $A \in \mathcal{A}$ such that $\mu(A) = \infty$. We wish to show that for all $M > 0$, there exists a subset $B \subseteq A$ such that $B \in \mathcal{A}$ and $M < \mu(B) < \infty$.

We wish to show that there exists measurable subsets of A of arbitrarily large size. Therefore, consider the collection

$$S = \{\mu(B) \mid B \subseteq A, B \in \mathcal{A}, \mu(B) < \infty\}.$$

Denote $l = \sup S$. We wish to show that $l = \infty$. Pick a sequence $\{B_n\} \subseteq S$ such that $\lim_{n\to\infty}\mu(B_n) = l$. We first claim that

$$\mu\left(\bigcup_n B_n\right) = l \tag{10.1}$$

Clearly, $\bigcup_n B_n \in \mathcal{A}$. Observe that since

$$\mu(B_k) \leq \mu\left(\bigcup_n B_n\right)$$

for all $k \in \mathbb{N}$, therefore taking $k \to \infty$, we easily obtain

$$l \leq \mu\left(\bigcup_n B_n\right).$$

Conversely, we wish to show that

$$\mu\left(\bigcup_n B_n\right) \leq l.$$

Let $D_1 = B_1$, $D_2 = B_1 \cup B_2$ and in general $D_n = B_1 \cup \cdots \cup B_n$. Then we observe that $\{D_n\} \subseteq \mathcal{A}$ forms an increasing sequence of sets with $\bigcup_n D_n = \bigcup_n B_n$. Consequently,

$$\mu\left(\bigcup_n B_n\right) = \mu\left(\bigcup_k D_k\right) = \lim_{k\to\infty}\mu(D_k).$$

Since $D_k \subseteq A$ is such that $\mu(D_k) \leq \sum_{i=1}^k \mu(B_i) < \infty$ (by subadditivity), therefore $\mu(D_k) \in S$ for all $k \in \mathbb{N}$. Consequently,

$$\lim_{k\to\infty}\mu(D_k) \leq l.$$

Therefore we obtain $\mu\left(\bigcup_n B_n\right) \leq l$. Hence this completes the proof of Eq. (10.1).

Since we wish to show that $l = \infty$, so assume to the contrary that $l < \infty$. It follows from Eq. (10.1) that $\mu\left(\bigcup_n B_n\right) < \infty$ and therefore $\bigcup_n B_n \in S$. Let $C = \bigcup_n B_n$. Then consider $A_1 = A \setminus C$. Since $\mu(A_1) = \mu(A) - \mu(C)$ as $\mu(C) < \infty$, therefore we have $\mu(A_1) = \infty - \mu(C) = \infty$. It follows from semifiniteness that there exists $C_1 \subseteq A_1$ such that $C_1 \in \mathcal{A}$ and $0 < \mu(C_1) < \infty$. Note that C_1 and C are disjoint. It follows that the disjoint union $C_1 \cup C \subseteq A$ is such that $\mu(C \cup C_1) \in S$. But since $\mu(C_1 \cup C) = \mu(C_1) + \mu(C) > \mu(C) = l$, therefore S contains an element which is strictly larger than its supremum, a contradiction. Hence $l = \infty$ and this completes the proof. □

3.6.2 Lebesgue measure on $\mathbb{R}$

In this section $(\mathbb{R}, \mathcal{M}, m)$ denotes the Lebesgue measure space on $\mathbb{R}$ and m^* denotes the Lebesgue outer measure on $\mathbb{R}$.

Lemma 3.6.11. *Every Borel subset of $\mathbb{R}$ is Lebesgue measurable.*

Proof. Let $(\mathbb{R}, \mathcal{M}, m)$ be the Lebesgue measure space over $\mathbb{R}$. We wish to show that the σ-algebra of Borel sets denoted $\mathcal{B}$ is in $\mathcal{M}$. Denote by $\mathcal{A}$ the following:

$$\mathcal{A} = \{\text{disjoint finite union of intervals of form } (-\infty, a], (b, \infty), (a, b] \text{ for } a < b \in \mathbb{R}\}. \tag{1.1}$$

By construction of Lebesgue measure, we know that $\mathcal{A} \subseteq \mathcal{M}$. We thus claim that the σ-algebra generated by $\mathcal{A}$ contains $\mathcal{B}$, that is, $\langle \mathcal{A} \rangle \supseteq \mathcal{B}$. This will conclude the proof.

Indeed, as we know that $\mathcal{B}$ is generated by all closed intervals of the form $(-\infty, a]$ for all $a \in \mathbb{R}$, therefore it suffices to show that $(-\infty, a] \in \langle \mathcal{A} \rangle$, but that is a tautology as $(-\infty, a]$ is in $\mathcal{A}$. Hence $\mathcal{B} \subseteq \langle A \rangle$. □

Lemma 3.6.12. *Let A be a subset of $\mathbb{R}$ and $c \in \mathbb{R}$. Then,*

1. *$m^*(A + c) = m^*(A)$,*
2. *$A \in \mathcal{M}$ if and only if $A + c \in \mathcal{M}$,*
3. *if $A \in \mathcal{M}$, then $m(A + c) = m(A)$.*

Proof. Consider the Lebesgue measure space $(\mathbb{R}, \mathcal{M}, m)$. Take $A \subseteq \mathbb{R}$ and for $c \in \mathbb{R}$ define $A + c = \{a + c \in \mathbb{R} \mid a \in A\}$. Let us set up some notation. For any $E \subseteq \mathbb{R}$, we denote

$$C(E) = \left\{\{I_n\} \mid \bigcup_n I_n \supseteq A,\ I_n = (a_n, b_n] \in \mathcal{A}\right\} \tag{$*$}$$

where $\mathcal{A}$ is the algebra defined in Eq. (1.1). Further, let us denote

$$\Sigma C(E) = \left\{\sum_n l(I_n) \in [0, \infty] \mid \{I_n\} \in C(E)\right\} \tag{$**$}$$

where $l((a, b]) = b - a$ is the length function. By definition, we have $m^*(E) := \inf \Sigma C(E)$.

(i) : We first wish to show that the Lebesgue outer measure m^* is translation invariant. That is, $m^*(A + c) = m^*(A)$. We first show $m^*(A + c) \geq m^*(A)$. Pick any $\{I_n\} \in C(A)$. Then we claim that $\{I_n + c\}$ is an element of $C(A + c)$. Indeed, denoting $I_n = (a_n, b_n]$, we immediately

get $I_n + c = (a_n + c, b_n + c]$. Now to see that $\bigcup_n (I_n + c) \supseteq A + c$, pick any $a + c \in A + c$ where $a \in A$. Then, as $\bigcup_n I_n \supseteq A$, therefore $a \in I_n$ for some n and thus $a + c \in I_n + c$. It follows that $\{I_n + c\} \in C(A + c)$. Further note that $l(I_n) = l(I_n + c)$ by definition. Consequently, we have

$$\Sigma C(A) \subseteq \Sigma C(A + c).$$

Taking infima, we yield $m^*(A) = \inf \Sigma C(A) \leq \inf \Sigma C(A + c) = m^*(A + c)$, that is $m^*(A) \leq m^*(A + c)$.

Conversely, we wish to show that $m^*(A) \geq m^*(A+c)$. For this, we use the standard technique of *ϵ-wiggle around* inf. Fix $\epsilon > 0$. By definition of $m^*(A)$, there exists $\{I_n\} \in C(A)$ where $I_n = (a_n, b_n]$ such that

$$m^*(A) + \epsilon > \sum_n b_n - a_n. \tag{2.1}$$

Note that we can write the above as

$$\begin{aligned} m^*(A) + \epsilon &> \sum_n (b_n + c) - (a_n + c) \\ &= \sum_n l((a_n + c, b_n + c]) \\ &= \sum_n l(I_n + c). \end{aligned}$$

We have $\{I_n + c\} \in C(A + c)$ as shown previously, therefore we obtain

$$m^*(A) + \epsilon > \sum_n l(I_n + c) \geq \inf \Sigma C(A + c) = m^*(A + c).$$

Hence we have $m^*(A) + \epsilon > m^*(A + c)$. Taking $\epsilon \to 0$, we obtain $m^*(A) \geq m^*(A + c)$. This completes the proof.

(ii) : We next wish to show that $A + c \in \mathcal{M}$ if and only if $A \in \mathcal{M}$. Observe that it suffices to show that $A \in \mathcal{M} \implies A + c \in \mathcal{M}$. Indeed, for the converse, take $B = A + c \in \mathcal{M}$. To show that $A \in \mathcal{M}$, it would suffice to show that $B - c \in \mathcal{M}$, which would follow at once by previous. Hence, we may only show that $A \in \mathcal{M} \implies A + c \in \mathcal{M}$.

Pick $A \in \mathcal{M}$. Fix $\epsilon > 0$. By regularity theorems, there exists open $U \supseteq A$ such that $m^*(U \setminus A) < \epsilon$. We now claim the following three statements:

1. *$U + c$ is open* : Indeed, pick any $x + c \in U + c$ where $x \in U$. As U is open, there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq U$. Consequently, $(x - \delta + c, x + \delta + c) \subseteq U + c$, hence $U + c$ is open.
2. *$U + c$ contains $A + c$* : Pick any $a + c \in A + c$ where $a \in A$. As $U \supseteq A$, therefore $a + c \in U + c$.
3. *$(U + c) \setminus (A + c)$ equals $(U \setminus A) + c$* : We first show $(U + c) \setminus (A + c) \subseteq (U \setminus A) + c$. Pick any $x + c \in (U + c) \setminus (A + c)$. Then $x + c \in U + c$ and $x + c \notin A + c$. Thus, $x \in U$ and $x \notin A$. Hence $x \in U \setminus A$ and thus $x + c \in U \setminus A + c$.
   Conversely, pick $x + c \in (U \setminus A) + c$. Then $x \in U \setminus A$ and thus $x + c \in U + c$ and $x + c \notin A + c$. Thus $x + c \in (U + c) \setminus (A + c)$. This completes the proof of this claim.

By above three claims, we conclude that $U + c$ is an open set containing $A + c$ such that

$$m^*(U + c \setminus A + c) = m^*((U \setminus A) + c) \overset{(ii)}{=} m^*(U \setminus A) < \epsilon.$$

By regularity theorems, we conclude the proof.

(iii) : We wish to show that if $A \in \mathcal{M}$, then $m(A+c) = m(A)$. This is immediate from *(i)* as $m = m^*|_{\mathcal{M}}$. □

Lemma 3.6.13. *Let A be a subset of $\mathbb{R}$ and $c \in \mathbb{R}$. Then,*

1. *$m^*(cA) = |c|\, m^*(A)$,*
2. *for $c \neq 0$, $A \in \mathcal{M}$ if and only if $cA \in \mathcal{M}$,*
3. *if $A \in \mathcal{M}$, then $m(cA) = |c|\, m(A)$.*

Proof. Let $(\mathbb{R}, \mathcal{M}, m)$ be the Lebesgue measure space. Take any $A \subseteq \mathbb{R}$ and $c \in \mathbb{R}$.

(i) : We first wish to show that $m^*(cA) = |c|\, m^*(A)$. If $c = 0$, then the equality is immediate as $cA = \{0\}$ and we know that $m^*(\{0\}) = 0$ as $0 \in (-1/n, 1/n]$ for all $n \in \mathbb{N}$ so that $m^*(\{0\}) \leq 2/n$. Taking $n \to \infty$, we get that $m^*(\{0\}) = 0$. So we assume from now on that $c \neq 0$. We first immediately reduce to showing either one of

$$m^*(cA) \geq |c|\, m^*(A) \text{ or } m^*(cA) \leq |c|\, m^*(A)$$

Indeed, the other side follows by replacing A by cA and replacing c by $1/c$ in either of the above. We now have two cases based on c being positive or negative.

If $c > 0$, then we proceed as follows. We follow the convention of Eqns $(*)$ and $(**)$ as set up in Q2. We use the standard technique of ϵ-wiggle around inf. Fix $\epsilon > 0$. By definition of outer measure, there exists $\{I_n\} \in C(cA)$ where $I_n = (a_n, b_n]$ such that

$$m^*(cA) + \epsilon > \sum_n l(I_n). \tag{3.1}$$

As $\bigcup_n I_n \supseteq cA$ and $c > 0$, therefore we claim that $\bigcup_n (\frac{1}{c} I_n) \supseteq A$. Indeed, for any $a \in A$, $cA \in I_n$. Thus $ca \in (a_n, b_n]$. Consequently, $a \in (a_n/c, b_n/c] = (\frac{1}{c} I_n)$. Thus, $\{\frac{1}{c} I_n\} \in C(A)$. Consequently, we have

$$\sum_n l\left(\frac{1}{c} I_n\right) = \sum_n \frac{1}{c} l(I_n) \geq m^*(A).$$

Consequently, $\sum_n l(I_n) \geq c m^*(A)$. Using this in Eq. (3.1), we thus obtain

$$m^*(cA) + \epsilon > \sum_n l(I_n) \geq c m^*(A).$$

Taking $\epsilon \to 0$, we obtain $m^*(cA) \geq c m^*(A)$, as required.

If $c < 0$, then we begin similarly to the previous case. Fix $\epsilon > 0$. There exists $\{I_n\} \in C(A)$ where $I_n = (a_n, b_n]$ such that

$$m^*(A) + \epsilon > \sum_n l(I_n). \tag{3.2}$$

Note that $cI_n = c(a_n, b_n] = [cb_n, ca_n)$ as $c < 0$ and this type of set is not half-open and is thus not in $\mathcal{A}$, the algebra of half-opens of Eq. (1.1). Consequently, we have to use ϵ-wiggle to find a new collection of intervals obtained via cI_n which are half open but their sum of lengths in only in ϵ-neighborhood of those $\{cI_n\}$. Indeed, for each $n \in \mathbb{N}$, we may construct

$$J_n = \left(cb_n - \frac{\epsilon}{2^{n+1}}, ca_n + \frac{\epsilon}{2^{n+1}}\right].$$

Note that $J_n \supseteq cI_n$. As $\bigcup_n cI_n \supseteq cA$, therefore $\bigcup_n J_n \supseteq cA$. Thus $\{J_n\} \in C(cA)$. Consequently,

$$\begin{aligned} m^*(cA) &\le \sum_n l(J_n) \\ &= \sum_n c(a_n - b_n) + \frac{2\epsilon}{2^{n+1}} \\ &= \sum_n -c(b_n - a_n) + \sum_n \frac{\epsilon}{2^n} \\ &= -c\sum_n (b_n - a_n) + \epsilon \\ &= -c\sum_n l(I_n) + \epsilon \end{aligned}$$

where the third line follows from the series being positive and thus we can rearrange such a series. It thus follows by Eq. (3.2) and above that

$$\begin{aligned} m^*(cA) &< -c(m^*(A) + \epsilon) + \epsilon \\ &= -cm^*(A) + \epsilon(1 - c). \end{aligned}$$

Taking $\epsilon \to 0$, we obtain ($-c = |c|$ as $c < 0$)

$$m^*(cA) \le |c|\, m^*(A)$$

as required. This completes the proof.

(ii) : We now wish to show that for $c \neq 0$, $A \in \mathcal{M}$ if and only if $cA \in \mathcal{M}$. Note that this is not true for $c = 0$ as if we take a non-measurable set $V \subseteq \mathbb{R}$, then $cV = \{0\}$ is measurable but V is not.

Pick $c \neq 0$. We first note that showing only $A \in \mathcal{M} \implies cA \in \mathcal{M}$ is sufficient. Indeed, the other side follows by replacing c by $1/c$ in the above. So we reduce to showing $A \in \mathcal{M} \implies cA \in \mathcal{M}$.

Pick $A \in \mathcal{M}$ and $c \neq 0$ in $\mathbb{R}$. Fix $\epsilon > 0$. By regularity theorems, there exists open $U \setminus A$ such that $m^*(U \setminus A) < \frac{\epsilon}{|c|}$. We now claim the following statements:

1. *cU is open* : Pick $cx \in cU$ where $x \in U$. As U is open therefore there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq U$. Consequently, $c(x - \delta, x + \delta) = (c(x + \delta), c(x - \delta)) \subseteq cU$ and contain cx. Hence cU is open.
2. *cU contains cA* : Pick any cx in cA. Then $x \in A$. As $U \subseteq A$, therefore $x \in U$ and hence $cx \in cU$.
3. *$cU \setminus cA$ equals $c(U \setminus A)$* : For ($\subseteq$), pick any $cx \in cU \setminus cA$. Then $cx \in cU$ and $cx \notin cA$. Thus, $x \in U$ and $x \notin A$, that is $\in U \setminus A$ and thus $cx \in c(U \setminus A)$. Conversely to show ($\supseteq$), pick any $cx \in c(U \setminus A)$ where $x \in U \setminus A$. Thus, $x \in U$ and $x \notin A$. Thus $cx \in cU$ and $cx \notin cA$. Thus $cx \in cU \setminus cA$.

Following the above three lemmas, we conclude that cU is an open set containing cA such that

$$m^*(cU \setminus cA) = m^*(c(U \setminus A)) \overset{(i)}{=} |c|\, m^*(U \setminus A) < |c| \frac{\epsilon}{|c|} = \epsilon.$$

Thus by regularity theorems, $cA \in \mathcal{M}$ as well.

(iii) : We wish to show that if $A \in \mathcal{M}$, then $m(cA) = |c|\, m(A)$. But this is immediate from *(i)* as $m = m^*|_{\mathcal{M}}$. This completes the whole proof. □

Lemma 3.6.14. *For each subseteq $A \subseteq \mathbb{R}$, there exists a Borel subset $B \supseteq A$ such that*

$$m^*(A) = m(B).$$

Proof. We wish to show that for each $A \subseteq \mathbb{R}$, there exists a Borel set $B \supseteq A$ such that $m(B) = m^*(A)$. We divide into two cases based on outer measure of A. We will follow the notations of Eq. $(*)$ and $(**)$.

If $m^*(A) = \infty$. In this case, we claim that $B = \mathbb{R}$ will work. Indeed $\mathbb{R}$ is open and thus Borel. We thus claim that $m(\mathbb{R}) = \infty$. Indeed, for $I_n = (n, n+1]$, $n \in \mathbb{Z}$, we have that $\{I_n\}$ are disjoint and $\coprod_n I_n = \mathbb{R}$. As m is a measure and I_n are measurable, therefore

$$m(\mathbb{R}) = \sum_n m(I_n) = \sum_n 1 = \infty.$$

Hence $B = \mathbb{R}$ will work.

If $m^*(A) < \infty$, then we proceed as follows. For each $N \in \mathbb{N}$, there exists $\{I_n^N\} \in C(A)$ such that

$$m^*(A) + \frac{1}{N} > \sum_n l(I_n^N).$$

Define $U_N = \bigcup_n I_n^N$. As each half open interval $(a, b] = \bigcap_{n \in \mathbb{N}} (a, b + 1/n)$ is a Borel set, therefore U_N is a Borel set. Observe that

$$m(U_N) \leq \sum_n m^*(I_n^N) = \sum_n l(I_n^N) < m^*(A) + \frac{1}{N}.$$

Note that in the above we have used the fact that Lebesgue measure restricted to half opens is exactly the length function. We thus have for each $N \in \mathbb{N}$ a Borel set U_N containing A such that

$$m(U_N) < m^*(A) + \frac{1}{N}. \tag{4.1}$$

Denote $B_K = \bigcap_{N=1}^K U_N$. Then each B_K is Borel and $\{B_K\}$ is a decreasing sequence of sets. Furthermore, $\bigcap_{K=1}^\infty B_K = \bigcap_{N=1}^\infty U_N$. Denote $B = \bigcap_{K=1}^\infty B_K$. Observe that $B \supseteq A$ as $B_K \supseteq A$ for each $K \in \mathbb{N}$. Consequently, by continuity of m^* we have

$$m(B) \geq m^*(A).$$

For the converse, first note that by Eq. (4.1), $m(U_1) < \infty$. Thus by monotone convergence property of measures, we obtain that $\lim_{K\to\infty} m(B_K) = m\left(\bigcap_{K=1}^{\infty} B_K\right)$. It follows from above, $B_K \subseteq U_K$ and Eq. (4.1) that

$$\begin{aligned} m(B) &= m\left(\bigcap_{K=1}^{\infty} B_K\right) \\ &= \lim_{K\to\infty} m(B_K) \\ &\leq \lim_{K\to\infty} m(U_K) \\ &\overset{(4.1)}{<} \lim_{K\to\infty}\left(m^*(A) + \frac{1}{K}\right) \\ &\leq m^*(A). \end{aligned}$$

Thus $m(B) \leq m^*(A)$ and we are done. □

Lemma 3.6.15. *A bounded set $E \subseteq \mathbb{R}$ is measurable if and only if $m^*(A) = m^*(A\cap E)+m^*(A\cap E^c)$ for all bounded subsets $A \subseteq \mathbb{R}$.*

Proof. Let E be a bounded set of $\mathbb{R}$. We wish to show that E is measurable if and only if for all bounded sets $A \subseteq \mathbb{R}$, we get $m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c)$.

The ($\Rightarrow$) is immediate from definitions. For ($\Leftarrow$), we proceed as follows. We wish to show that for any $F \subseteq \mathbb{R}$, we have

$$m^*(F) \geq m^*(F \cap E) + m^*(F \cap E^c).$$

Indeed, if $m^*(F) = \infty$, then there is nothing to show. So we assume $m^*(F) < \infty$. Observe then that $m^*(F \cap E), m^*(F \cap E^c) \leq m^*(F) < \infty$. Fix $\epsilon > 0$. There exists a sequence $\{I_n\}$ of half-opens such that $\bigcup_n I_n \supseteq F$ and

$$m^*(F) + \epsilon > \sum_n m^*(I_n)$$

where we are using the fact that measure of a half-open interval is its length. Observe that for each $n \in \mathbb{N}$, we have $m^*(F) + \epsilon > m^*(I_n)$, thus each I_n is a half-open interval with bounded length, hence I_n is bounded as a set. Consequently, we have

$$\begin{aligned} m^*(F) + \epsilon &> \sum_n m^*(I_n) \\ \text{(by hypothesis)} \quad &\geq \sum_n m^*(I_n \cap E) + m^*(I_n \cap E^c) \\ \text{(by rearrangement of +ve series)} \quad &= \sum_n m^*(I_n \cap E) + \sum_n m^*(I_n \cap E^c) \\ \text{(by subadditivity)} \quad &\geq m^*\left(\bigcup_n I_n \cap E\right) + m^*\left(\bigcup_n I_n \cap E^c\right) \\ \text{(by } \cup_n I_n \supseteq F) \quad &\geq m^*(F \cap E) + m^*(F \cap E^c). \end{aligned}$$

This completes the proof. □

3.6.3 Measurable functions

Notation 3.6.16. At times, we will write a subset of X as follows:

$$\{x \in X \mid \mathcal{P}_x \text{ is true}\} = \{\mathcal{P}_x \text{ is true}\}.$$

This makes some constructions much more clearer to see and interpret.

Lemma 3.6.17. *Let $f : X \to Y$ be a function and $\mathcal{A}$ be an algebra on Y. Then,*

$$\langle f^{-1}(\mathcal{A})\rangle = f^{-1}(\langle \mathcal{A}\rangle).$$

Proof. Let $f : X \to Y$ be a function and $\mathcal{A}$ be an algebra over Y. We wish to show that

$$\langle f^{-1}(\mathcal{A})\rangle = f^{-1}(\langle \mathcal{A}\rangle). \tag{2.1}$$

We first claim that $f^{-1}(\langle \mathcal{A}\rangle)$ is a σ-algebra over X. Indeed, as $Y, \emptyset \in \langle \mathcal{A}\rangle$, we have $f^{-1}(Y) = X$ and $f^{-1}(\emptyset) = \emptyset$. Further, if $B \in f^{-1}(\langle \mathcal{A}\rangle)$, then $B = f^{-1}(A)$ for some $A \in \langle \mathcal{A}\rangle$. Hence $B^c = f^{-1}(A)^c = f^{-1}(A^c)$ and $A^c \in \langle A\rangle$ as $\langle A\rangle$ is a σ-algebra. Finally, pick $\{B_n\} \subseteq f^{-1}(\langle A\rangle)$. Then $B_n = f^{-1}(A_n)$ for $A_n \in \langle \mathcal{A}\rangle$. Consequently, $\bigcup_n B_n = \bigcup_n f^{-1}(A_n) = f^{-1}\left(\bigcup_n A_n\right)$ and since $\bigcup_n A_n \in \langle A\rangle$, hence this proves that $f^{-1}(\langle \mathcal{A}\rangle)$ is a σ-algebra.

We now show $(\subseteq)$ part of Eq. (2.1). Indeed, by above, it would suffice to show that $f^{-1}(\mathcal{A})$ is contained in the σ-algebra $f^{-1}(\langle \mathcal{A}\rangle)$. Pick any $B \in f^{-1}(\mathcal{A})$, so that $B = f^{-1}(A)$ where $A \in \mathcal{A}$. As $\mathcal{A} \subseteq \langle \mathcal{A}\rangle$, therefore $A \in \langle \mathcal{A}\rangle$. It follows that $B = f^{-1}(A) \in f^{-1}(\langle \mathcal{A}\rangle)$. This shows that $\langle f^{-1}(\mathcal{A})\rangle \subseteq f^{-1}(\langle \mathcal{A}\rangle)$.

We now show $(\supseteq)$ part of Eq. (2.1). We will use the standard technique of *good sets* for this. Consider

$$\mathcal{C} := \{A \in \langle \mathcal{A}\rangle \mid f^{-1}(A) \in \langle f^{-1}(\mathcal{A})\rangle\} \subseteq \langle \mathcal{A}\rangle$$

We now claim the following two statements:

1. *$\mathcal{C}$ is a σ-algebra on Y* : Indeed, $Y = f^{-1}(X)$ and $\emptyset = f^{-1}(\emptyset)$ where $X, \emptyset \in \langle \mathcal{A}\rangle$ and $X, \emptyset \in \langle f^{-1}(\mathcal{A})\rangle$. Further, for $A \in \mathcal{C}$, we have $f^{-1}(A) \in \langle f^{-1}(\mathcal{A})\rangle$ and thus $(f^{-1}(A))^c = f^{-1}(A^c) \in \langle f^{-1}(\mathcal{A})\rangle$. Thus $A^c \in \mathcal{C}$. Finally, pick $\{A_n\} \subseteq \mathcal{C}$. Then $f^{-1}(A_n) \in \langle f^{-1}(\mathcal{A})\rangle$ for each $n \in \mathbb{N}$. Thus, $\bigcup_n f^{-1}(A_n) = f^{-1}\left(\bigcup_n A_n\right) \in \langle f^{-1}(\mathcal{A})\rangle$. It then follows that $\bigcup_n A_n \in \mathcal{C}$. This shows that $\mathcal{C}$ is a σ-algebra.
2. $\mathcal{C} \supseteq \mathcal{A}$: Pick any $A \in \mathcal{A}$. As $\langle f^{-1}(\mathcal{A})\rangle$ contains $f^{-1}(\mathcal{A})$, so $f^{-1}(A) \in \langle f^{-1}(\mathcal{A})\rangle$.

We now conclude the proof. As $\mathcal{C}$ is a σ-algebra containing $\mathcal{A}$ and inside $\langle \mathcal{A}\rangle$, therefore $\mathcal{C} = \langle \mathcal{A}\rangle$. It follows that for each $A \in \langle \mathcal{A}\rangle$, we have $f^{-1}(A) \in \langle f^{-1}(\mathcal{A})\rangle$, that is $f^{-1}(\langle \mathcal{A}\rangle) \subseteq \langle f^{-1}(\mathcal{A})\rangle$, as required. This completes the proof. □

Lemma 3.6.18. *Let $(X, \mathcal{M}, m)$ be the Lebesgue measure space. Let $A \in \mathcal{M}$ be a bounded set such that $0 < m(A) < \infty$. For each $0 < M < m(A)$, there exists a $B \subsetneq A$ such that $B \in \mathcal{M}$ and $m(B) = M$.*

Proof. There are two proofs that we wish to present, one uses Lemma 3.6.19 and other is independent. The latter uses a nice technique which we would like to write down concretely.

Method 1 : *(Using Lemma 3.6.19)* Consider the map

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto m(A \cap (-\infty, x]). \end{aligned}$$

As A is a bounded set, therefore $m(A) < \infty$ as there exists a bounded interval $I \supseteq A$ where $I = [c, d]$. By Lemma 3.6.19, the map f is a continuous map. Let $a \in \mathbb{R}$ be such that $a < c$. Then $f(a) = m(A \cap (-\infty, a]) = m(\emptyset) = 0$. Let $b \in \mathbb{R}$ such that $b > d$. Then, $f(b) = m(A \cap (-\infty, b]) = m(A)$. On the interval $J = [a, b]$ we have $f(a) = 0$ and $f(b) = m(A)$. By intermediate value property of f, there exists $c \in J$ such that $f(c) = M$. Consequently, $A \cap (-\infty, c]$ is a measurable subset of A whose measure is M.

Method 2 : *(Exponential subdivision technique)* We shall explicitly construct $B \subsetneq A$ such that $m(B) = M$. First, we observe that the question is invariant under translation and dilation. Hence we may, after suitable dilation and translation, assume that $A \subseteq [0, 1)$. For each $n \in \mathbb{N}$, consider the following partition of $[0, 1)$

$$P_n : 0 < x_1 = \frac{1}{2^n} < x_2 = 2 \cdot \frac{1}{2^n} < \cdots < x_{2^n - 1} = (2^n - 1) \cdot \frac{1}{2^n} < 1.$$

Denote $I_{n,j} = [\frac{j-1}{2^n}, \frac{j}{2^n})$ for each $j = 1, \ldots, 2^n$. Observe that $I_{n,j}$ are disjoint and, denoting $A_{n,j} = A \cap I_{n,j}$, we further have a disjoint collection $\{A_{n,j}\}$ of measurable subsets[16] of A such that

$$\coprod_{j=1}^{2^n} A_{n,j} = A.$$

Further, we have that

$$\begin{aligned} \sum_{j=1}^{2^n} m(A_{n,j}) &= m\left(\coprod_{j=1}^{2^n} A_{n,j}\right) \\ &= m(A) \end{aligned}$$

and that

$$m(A_{n,j}) \leq m(I_{n,j}) = \frac{1}{2^n}.$$

Now, for each $n \in \mathbb{N}$, let N_n be the largest index such that

$$\sum_{j=1}^{N_n} m(A_{n,j}) \leq M.$$

[16] measurable because A and $I_{n,j}$ are measurable

By the choice of index N_n, we observe that

$$\begin{aligned} M &< \sum_{j=1}^{N_n+1} m(A_{n,j}) \\ &= \sum_{j=1}^{N_n} m(A_{n,j}) + m(A_{n,N_n+1}) \\ &\leq \sum_{j=1}^{N_n} m(A_{n,j}) + \frac{1}{2^n}. \end{aligned}$$

Denoting $C_n = \amalg_{j=1}^{N_n} A_{n,j}$, we obtain,

$$M - \frac{1}{2^n} < \sum_{j=1}^{N_n} m(A_{n,j}) = m\left(C_n\right) \leq M. \tag{3.1}$$

We now claim that $\{C_n\}$ is an increasing sequence of measurable subsets of A. First observe that for each $n \in \mathbb{N}$, we have that N_{n+1} is either $2N_n - 1$ or $2N_n$. Indeed, pick any $x \in C_n$. Then $x \in A_{n,j}$ where $j = 1, \ldots, N_n$. Expanding this, we have

$$\begin{aligned} x &\in A_{n,j} \\ &= A \cap \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right) \\ &= A \cap \left(\left[\frac{2(j-1)}{2^{n+1}}, \frac{2j-1}{2^{n+1}}\right) \amalg \left[\frac{2j-1}{2^{n+1}}, \frac{2j}{2^{n+1}}\right)\right) \\ &= A_{n+1,2j-1} \amalg A_{n+1,2j}. \end{aligned} \tag{3.2}$$

As $N_{n+1} = 2N_n - 1$ or $2N_n$, therefore for $j = 1, \ldots, N_n$, $2j = 2, \ldots, 2N_n$, hence in Eq. (3.2), we obtain that $x \in A_{n+1,2j-1}$ or $x \in A_{n+1,2j}$ and as $2j \leq 2N_n$, hence $x \in C_{n+1}$. This shows that $C_n \subseteq C_{n+1}$.

Applying $\lim_{n\to\infty}$ on Eq. (3.1), we thus obtain

$$M \leq \lim_{n\to\infty} m(C_n) \leq M.$$

Thus, by monotone convergence of measures, we conclude

$$\begin{aligned} M &= \lim_{n\to\infty} m(C_n) \\ &= m\left(\bigcup_n C_n\right). \end{aligned}$$

As $C_n \subseteq A$ for each $n \in \mathbb{N}$, therefore $\bigcup_n C_n \subseteq A$. Consequently we have obtained a subset of A whose measure is M. □

Lemma 3.6.19. *Let $(X, \mathcal{M}, \mu)$ be the Lebesgue measure space and $A \in \mathcal{M}$ be a bounded set. Then the function*

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto m(A \cap (-\infty, x]) \end{aligned}$$

is continuous.

Proof. Let $A \in \mathcal{M}$ which has finite measure. We wish to show that

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto m(A \cap (-\infty, x]) \end{aligned}$$

is continuous. Pick any $a \in \mathbb{R}$ and any $\epsilon > 0$. We wish to find a $\delta > 0$ such that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$. We now have two cases with respect to the position of x and a in $\mathbb{R}$.

1. *If* $a \leq x$: then $f(x) - f(a)$ can be rewritten as follows:

$$\begin{aligned} |f(x) - f(a)| &= m(A \cap (-\infty, x]) - m(A \cap (-\infty, a]) \\ &= m(A \cap (-\infty, x] \setminus A \cap (-\infty, a]) \\ &= m(A \cap (a, x]) \\ &\leq m((a, x]) \\ &= x - a. \end{aligned}$$

   Therefore taking $\delta = \epsilon$, we would be done.

2. *If* $a > x$: then $|f(x) - f(a)|$ can be written as

$$\begin{aligned} |f(x) - f(a)| &= |f(a) - f(x)| \\ &= m(A \cap (-\infty, a]) - m(A \cap (-\infty, x]) \\ &= m(A \cap (x, a]) \\ &\leq m((x, a]) \\ &= a - x. \end{aligned}$$

   Thus, again, taking $\delta = \epsilon$ would do the job.

This completes the proof. □

Lemma 3.6.20. *Let X be a measurable space and let $f : X \to \mathbb{R}$ be a function. Suppose $\{x \in X \mid a \leq f(x) < b\}$ is measurable for all $a < b$. Then f is a measurable function.*

Proof. As the Borel σ-algebra on $\mathbb{R}$ is generated by sets of the form $[a, \infty)$ for $a \in \mathbb{R}$, therefore for a fixed $a \in \mathbb{R}$ we need only show that $f^{-1}([a, \infty))$ is measurable in X.

We can write

$$\begin{aligned} f^{-1}([a, \infty)) &= \{a \leq f(x)\} \\ &= \bigcup_{n > a \text{ in } \mathbb{N}} \{a \leq f(x) < n\}. \end{aligned}$$

As we are given that $\{a \leq f(x) < b\}$ are measurable for all $a < b \in \mathbb{R}$ and countable union of measurable sets is measurable, therefore $f^{-1}([a, \infty))$ is measurable. □

Lemma 3.6.21. *All monotone functions $f : \mathbb{R} \to \mathbb{R}$ are measurable.*

Proof. We wish to show that all monotone functions $f : \mathbb{R} \to \mathbb{R}$ are measurable. Note that we may first reduce to assuming that f is non-decreasing as if f is non-increasing, then $-f$ will be non-decreasing.

Hence let $f : \mathbb{R} \to \mathbb{R}$ is non-decreasing. As Borel σ-algebra of $\mathbb{R}$ is generated by intervals of the form $[a, \infty)$, $a \in \mathbb{R}$, therefore it suffices to check that $f^{-1}([a, \infty))$ is measurable in $\mathbb{R}$. Observe

$$f^{-1}([a, \infty)) = \{a \leq f(x)\}.$$

We now have two cases to handle.

1. *If $a \in f(\mathbb{R})$:* Then there exists $b \in \mathbb{R}$ such that $f(b) = a$. We may write

$$\{a \leq f(x)\} = \{a < f(x)\} \amalg \{a = f(x)\}.$$

   Now since f is non-decreasing, therefore $f(x) > f(y)$ implies $x > y$. Further, we have that $f^{-1}(a) = \{a = f(x)\}$ is measurable as singletons are Borel. Consequently, we have

$$\begin{aligned}\{a \leq f(x)\} &= \{f(b) < f(x)\} \amalg \{a = f(x)\} \\ &= (b, \infty) \amalg f^{-1}(a).\end{aligned}$$

   Hence $f^{-1}([a, \infty))$ is measurable.
2. *If $a \notin f(\mathbb{R})$:* We further have two cases.
   (a) *If there exists $b \in \mathbb{R}$ such that $b \notin \{a \leq f(x)\}$:* Observe first that in this case $f(b) < a$. We claim that in this case $\{a \leq f(x)\}$ is lower bounded by b. Indeed, suppose not. Then there exists $y < b$ such that $y \in \{a \leq f(x)\}$. Then $a \leq f(y) \leq f(b) < a$, a contradiction. Hence $\{a \leq f(x)\}$ is bounded below.
   Let $c = \inf\{a \leq f(x)\}$, which now exists. Consequently, we have two more cases:
   - *If $f(c) \geq a$:* That is, if $c \in \{a \leq f(x)\}$. Then we claim

$$\{a \leq f(x)\} = [c, \infty).$$

     which is clearly a measurable. Indeed, for some $x \in \mathbb{R}$ such that $f(x) \geq a$, then $x \geq c$. Conversely, if $b \geq c$ in $\mathbb{R}$, then $f(b) \geq f(c) \geq a$, so $b \in \{a \leq f(x)\}$. This proves the claim.
   - *If $f(c) < a$:* That is, if $c \notin \{a \leq f(x)\}$. Then we claim

$$\{a \leq f(x)\} = (c, \infty)$$

     which is clearly a measurbale set. Indeed, for $x \in \mathbb{R}$ such that $f(x) \geq a$, $x > c$. Further $x \neq c$ as otherwise $f(x) < a$. Conversely, if $b > c$, then there exists $d \in \{a \leq f(x)\}$ such that $c < d < b$ as c is the infimum. Consequently, $a \leq f(d) \leq f(b)$. Hence $b \in \{a \leq f(x)\}$. This proves the claim.

   (b) *If there doesn't exists any $b \in \mathbb{R}$ such that $f(b) < a$:* Then for all $b \in \mathbb{R}$ we have $f(b) \geq a$. Consequently, $f^{-1}([a, \infty)) = \{a \leq f(x)\} = \mathbb{R}$, which is measurable.

Hence in all cases $f^{-1}([a, \infty))$ is a measurable set. This completes the proof. □

Lemma 3.6.22. *Let $f : X \to \mathbb{C}$ be a complex measurable function on a measurable space X. Then, there exists a complex measurable function $g : X \to \mathbb{C}$ such that $|g| = 1$ and $f = g\,|f|$.*

Proof. Let $f : X \to \mathbb{C}$ be a measurable function. We wish to find a measurable function $g : X \to \mathbb{C}$ such that $|g| = 1$ and $f = g\,|f|$.

As $|f| = f\chi_{\{f(x)\geq 0\}} - f\chi_{\{f(x)<0\}}$, therefore $|f|$ is a measurable function. Denote $E = \{|f(x)| = 0\}$. Consequently, we define g as follows:

$$g(x) = \begin{cases} \frac{f(x)}{|f(x)|} & \text{if } x \in E^c \\ 1 & \text{if } x \in E. \end{cases}$$

We first wish to show that g is measurable. For this, we shall use the fact that measurability of g can be checked on a cover $\{E_\alpha\}$ of X such that $g|_{E_\alpha}$ is measurable. Thus in our case, we need only show that $g|_E$ and $g|_{E^c}$ are measurable. On E, g is a constant, hence measurable. On E^c, g is $f/|f|$. As $|f|$ is not zero on E^c, therefore by Lemma 3.6.24, $f/|f|$ is measurable. Hence, g is measurable.

We now see that $|g|(x) = \left|\frac{f(x)}{|f(x)|}\right| = 1$ on E^c and $|g(x)| = 1$ on E. Thus $|g| = 1$ on X. Further, if $x \in E$, then $f(x) = 0 = g(x)|f|(x)$. If $x \in E^c$, then $g(x) = \frac{f(x)}{|f(x)|}$ which implies $|f(x)|g(x) = f(x)$. This shows that in all cases, $f = g|f|$. □

Example 3.6.23. It is not true that if $f : [0,1] \to \mathbb{R}$ is a function whose each fibre is measurable, then f is measurable.

Consider the following function

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} x & \text{if } x \in V^c \\ x+N & \text{if } x \in V \end{cases} \end{aligned}$$

where $V \subseteq [0,1]$ denotes the Vitali set and $N = 3$. Then, for each $y \in \mathbb{R}$, we have that $f^{-1}(y)$ is atmost a singleton, which is measurable in $[0,1]$. However, for any $1 < b < N$, we see that $f^{-1}((b,\infty)) = V$, which is not measurable. Hence f is a non-measurable function whose fibres are measurable.

Lemma 3.6.24. *Let $f, g : X \to \mathbb{C}$ be a measurable function such that $\{g(x) \neq 0\} = X$. Then f/g is measurable.*

Proof. Let $f, g : X \to \mathbb{C}$ be a measurable function such that $\{g(x) \neq 0\} = X$. Then we wish to show that f/g is measurable.

We first have that $(f,g) : X \to \mathbb{R}^2$ given by $x \mapsto (f(x), g(x))$ is measurable. Consequently, we consider the composite

$$X \xrightarrow{(f,g)} \mathbb{R}^2 \setminus \{y = 0\} \xrightarrow{\Phi} \mathbb{R}$$

where $\Phi(x,y) = \frac{x}{y}$. As Φ is continuous, therefore the composite $\Phi \circ (f,g)$ is measurable. Consequently, we obtain that the map $x \mapsto \frac{f(x)}{g(x)}$ is measurable, but this is exactly f/g over X. This completes the proof. □

Lemma 3.6.25. *Let $f, g : X \to \overline{\mathbb{R}}$ be measurable functions and pick any $r_0 \in \overline{\mathbb{R}}$. Then the map*

$$\begin{aligned} h : X &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto \begin{cases} r_0 & \textit{if } f(x) = -g(x) = \pm\infty \\ f(x) + g(x) & \textit{else} \end{cases} \end{aligned}$$

is measurable[17].

Proof. Let $f, g : X \to \overline{\mathbb{R}}$ be measurable functions and pick any $r_0 \in \overline{\mathbb{R}}$. Then we wish to show that the map

$$\begin{aligned} h : X &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto \begin{cases} r_0 & \text{if } f(x) = -g(x) = \pm\infty \\ f(x) + g(x) & \text{else} \end{cases} \end{aligned}$$

is measurable.

Define the following sets

$$\begin{aligned} E &= \{f(x) = \infty = -g(x)\} \\ F &= \{f(x) = -\infty = -g(x)\}. \end{aligned}$$

As $E = f^{-1}(\infty) \cap g^{-1}(-\infty)$ and $F = f^{-1}(-\infty) \cap g^{-1}(\infty)$, therefore they are measurable. Observe that E and F are disjoint. We thus need only show that h restricted to E, F and $X \setminus (E \amalg F)$ is measurable.

1. On E : As $h|_E$ is constant r_0, therefore $h|_E$ is measurable.
2. On F : As $h|_F$ is again constant r_0, therefore $h|_F$ is measurable.
3. On $X \setminus (E \amalg F)$: We first deduce that

$$\begin{aligned} X \setminus (E \amalg F) &= X \cap E^c \cap F^c \\ &= E^c \cap F^c \\ &= (\{f(x) \neq \infty\} \cup \{g(x) \neq -\infty\}) \bigcap (\{f(x) \neq -\infty\} \cup \{g(x) \neq \infty\}) \end{aligned}$$

Let $G = \{f(x) \in \mathbb{R}\}$ and $H = \{g(x) \in \mathbb{R}\}$. Then we may write $X \setminus (E \amalg F)$ as

$$\begin{aligned} X \setminus (E \amalg F) &= (G \cup H \cup \{f(x) = -\infty\} \cup \{g(x) = \infty\}) \bigcap (G \cup H \cup \{f(x) = \infty\} \cup \{g(x) = -\infty\}) \\ &= (G \cup H) \cup \Big((\{f(x) = -\infty\} \cup \{g(x) = \infty\}) \bigcap (\{f(x) = \infty\} \cup \{g(x) = -\infty\}) \Big) \\ &= (G \cup H) \cup \underbrace{\{f(x) = -\infty = g(x)\}}_{=:A} \cup \underbrace{\{f(x) = \infty = g(x)\}}_{=:B}. \end{aligned}$$

As $h|_{G\cup H}$ is $(f+g)|_{G\cup H}$ and on $G \cup H$, $f + g : G \cup H \to \mathbb{R}$, therefore h is measurable. We thus reduce to checking that $h|_A$ and $h|_B$ are measurable. On both of them, one immediately observes that h is constant $-\infty$ and ∞ respectively. Hence, $h|_A$ and $h|_B$ are measurable. As h restricted to $G \cup H$, A and B is measurable therefore h restricted to $X \setminus (E \amalg F)$ is measurable.

This completes the proof. ☐

Example 3.6.26. It is not true in general that if for a function $f : X \to \mathbb{R}$, the $|f| : X \to [0, \infty]$ is measurable then f is measurable.

[17] This question in particular shows that modifying a measurable function at a single point doesn't affect measurability at all.

Indeed, consider the following function where $V \subseteq [0,1]$ denotes the Vitali set:

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} -x & \text{if } x \in V \\ x & \text{if } x \in V^c. \end{cases} \end{aligned}$$

Then, $|f| = \mathrm{id}_{[0,1]}$ which is measurable whereas f is not measurable as $f^{-1}((-\infty, 0)) = V$, which is not a measurable set.

Lemma 3.6.27. *Let $(X_1, \mathcal{A}_{X_1}, \mu_1)$ be a measure space, $(X_2, \mathcal{A}_{X_2})$ be a measurable space and $f : X_1 \to X_2$ be a measurable function. Then*

$$\begin{aligned} \mu_2 : \mathcal{A}_{X_2} &\longrightarrow [0,\infty] \\ B &\longmapsto \mu_1(f^{-1}(B)) \end{aligned}$$

is a measure on $(X_2, \mathcal{A}_{X_2})$.

Proof. We first immediately observe that $\mu_2(\emptyset) = \mu_1(f^{-1}(\emptyset)) = \mu_1(\emptyset) = 0$. We thus reduce to showing that for any disjoint collection $\{B_n\} \subseteq \mathcal{A}_{X_2}$, we have $\mu_2(\coprod_n B_n) = \sum_n \mu_2(B_n)$. To this end, observe that

$$\begin{aligned} \mu_2\left(\coprod_n B_n\right) &= \mu_1\left(f^{-1}\left(\coprod_n B_n\right)\right) \\ &= \mu_1\left(\coprod_n f^{-1}(B_n)\right) \\ &= \sum_n \mu_1(f^{-1}(B_n)) \\ &= \sum_n \mu_2(B_n). \end{aligned}$$

This completes the proof. $\square$

Lemma 3.6.28. *Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{R}$ be a measurable function such that $\mu(\{|f(x)| \geq \epsilon\}) = 0$ for all $\epsilon > 0$. Then $f = 0$ almost everywhere.*

Proof. We first claim that it suffices to show that $\{|f(x)| > 0\}$ is a null set. Indeed, this is because $\{f(x) \neq 0\} = \{|f(x)| > 0\}$. Hence it suffices to show that $|f| = 0$ a.e.

Define for each $n \in \mathbb{N}$ the following subset of X

$$E_n = \{|f(x)| > 1/n\}.$$

We claim that

$$\{|f(x)| > 0\} = \bigcup_{n \in \mathbb{N}} E_n.$$

Indeed, for $(\subseteq)$, pick $x \in X$ such that $|f(x)| > 0$. Then there exists $n \in \mathbb{N}$ such that $|f(x)| > 1/n$. Hence $x \in E_n$. Conversely pick $x \in E_n$, then by way of construction of E_n, we have $|f(x)| > 1/n >$

0.

Observe that $\{E_n\}$ is an increasing sequence of sets as if $x \in E_n$ then $|f(x)| > \frac{1}{n} > \frac{1}{n+1}$, so $x \in E_{n+1}$. It then follows by monotone convergence property of measures that

$$\mu(\{|f(x) > 0|\}) = \mu\left(\bigcup_n E_n\right) = \lim_{n\to\infty}\mu(E_n) = \lim_{n\to\infty}0 = 0.$$

This completes the proof. □

Example 3.6.29. The statement of Egoroff's theorem depends crucially on the fact that each function in the sequence $\{f_n\}$ is measurable. Indeed, we show by the way of an example that the conclusion of Egoroff's theorem is not true when f_n's are not measurable.

We wish to show that the statement of Egoroff's theorem fails if we drop the condition that functions be measurable.

Consider the measure space $(\mathbb{Z}, \mathcal{A}, \mu)$ where $\mathcal{A} = \{\emptyset, \mathbb{Z}, 2\mathbb{Z}, \mathbb{Z} \setminus 2\mathbb{Z}\}$ and $\mu(\emptyset) = 0 = \mu(2\mathbb{Z})$, $\mu(\mathbb{Z}) = 1 = \mu(\mathbb{Z} \setminus 2\mathbb{Z})$. Consider the functions $f_n : (\mathbb{Z}, \mathcal{A}, \mu) \to \mathbb{R}$ where $\mathbb{R}$ has the Borel measure, given by

$$f_n(k) = \frac{k}{n}$$

for all $k \in \mathbb{Z}$. Observe that $\{f_n\}$ pointwise converges to the constant 0 function at all points of $\mathbb{Z}$. Further note that f_n is not measurable as $f_n^{-1}(\{k/n\}) = \{k\}$ is not a measurable set in $\mathcal{A}$ but $\{k/n\}$ is Borel measurable.

To show that this is a counterexample, it would suffice to show that there exists an $\epsilon_0 > 0$ such that for all measurable sets $E \in \mathcal{A}$, either $\mu(E^c) \geq \epsilon_0$ or f_n does not converges uniformly to 0 on E. We claim that in our situation, $\epsilon_0 = 1/2$ works. Indeed, for $E = \emptyset, 2\mathbb{Z}$, we have $\mu(E^c) = 1 > 1/2$. Thus we reduce to showing that f_n does not converges uniformly on $\mathbb{Z}$ and $\mathbb{Z} \setminus 2\mathbb{Z}$. Indeed, observe that $\sup_{k\in\mathbb{Z}} |f_n(k)| = \sup_{k\in\mathbb{Z}} k/n = \infty$ for each $n \in \mathbb{N}$. As f_n converges uniformly if and only if $\sup_{k\in\mathbb{Z}} |f_n(k)| \to 0$ as $n \to \infty$, therefore we deduce that f_n does not converge uniformly over $\mathbb{Z}$. Similarly, it doesn't converge uniformly over $\mathbb{Z} \setminus 2\mathbb{Z}$.

Lemma 3.6.30. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions. If $f = g$ almost everywhere, then $f = g$.*

Proof. Indeed, consider $h = f - g$. Suppose $h \neq 0$, therefore there exists $x_0 \in \mathbb{R}$ such that $h(x_0) \neq 0$. By continuity of h, there exists $\epsilon > 0$ such that $(x_0 - \epsilon, x_0 + \epsilon) \subseteq \{h(x) \neq 0\}$. Hence, $2\epsilon < m(\{h(x) \neq 0\}) = 0$, which yields $0 < 2\epsilon \leq 0$, a contradiction. □

Lemma 3.6.31. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $f_n, f : X \to \overline{\mathbb{R}}$ be measurable functions such that $f_n \to f$ pointwise almost everywhere. Then, there exists measurable functions $g_n : X \to \overline{\mathbb{R}}$ such that $f_n = g_n$ almost everywhere and $g_n \to f$ pointwise.*

Proof. Indeed, as f_n converges pointwise to f almost everywhere, therefore the set $E = \{\lim_{n\to\infty} f_n(x) \neq f(x)\}$ is a zero measure set. Consequently, we may define

$$g_n : X \longrightarrow \overline{\mathbb{R}}$$
$$x \longmapsto \begin{cases} f_n(x) & \text{if } x \notin E \\ f(x) & \text{if } x \in E. \end{cases}$$

We then observe that $\{g_n(x) \neq f_n(x)\} = E$, which is of measure zero, hence $g_n = f_n$ almost everywhere. Furthermore, we see that for any $x \in X$,

$$\lim_{n\to\infty} g_n(x) = \begin{cases} \lim_{n\to\infty} f_n(x) = f(x) & \text{if } x \notin E \\ \lim_{n\to\infty} f(x) = f(x) & \text{if } x \in E. \end{cases}$$

Thus, $\lim_{n\to\infty} g_n = f$ pointwise. This completes the proof. □

Example 3.6.32. We wish to show that there exists continuous function $f : \mathbb{R} \to \mathbb{R}$ and a Lebesgue measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $g \circ f : \mathbb{R} \to \mathbb{R}$ is not Lebesgue measurable.

While learning about the existence of a non-Borel measurable set, one learns about the existence of a homeomorphism $\varphi : [0, 1] \to [0, 2]$ such that $m(\varphi(C)) = 1 > 0$ where $C \subseteq [0, 1]$ is the Cantor set. Indeed, if $\mathcal{C} : [0, 1] \to [0, 1]$ denotes the Cantor function, then φ is constructed by defining $\varphi(x) = \mathcal{C}(x) + x$. As, $\mathcal{C}(0) = 0$ and $\mathcal{C}(1) = 1$, therefore $\varphi(0) = 0$ and $\varphi(1) = 2$. Consequently, we may define a continuous function $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f(x) = \begin{cases} x - 1 & \text{if } x > 2 \\ \varphi^{-1}(x) & \text{if } x \in [0, 2] \\ x & \text{if } x < 0. \end{cases}$$

Observe that f is continuous as f is obtained by gluing three continuous functions at points where they agree.

As $m(\varphi(C)) = 1 > 0$ for Cantor set C, therefore there exists a non-measurable set $V \subseteq \varphi(C) \subseteq [0, 2]$. But since $f(V) = \varphi^{-1}(V) \subseteq \varphi^{-1}(\varphi(C)) = C$ and C is a null set, therefore by completeness of Lebesgue measure, it follows that $f(V)$ is a Lebesgue measurable set. Consequently, we may define $g = \chi_{f(V)} : \mathbb{R} \to \mathbb{R}$, which is Lebesgue measurable as $f(V)$ is Lebesgue measurable. We thus have

$$\mathbb{R} \xrightarrow[\text{cont.}]{f} \mathbb{R} \xrightarrow[\text{Leb. msble}]{g} \mathbb{R} .$$

We claim that $h := g \circ f$ is not Lebesgue measurable. Indeed, observe that $h^{-1}(\{1\}) = (g \circ f)^{-1}(\{1\}) = f^{-1}(g^{-1}(\{1\})) = f^{-1}(f(V))$. But as f restricted to $[0, 2]$ is a homeomorphism from $[0, 2]$ to $[0, 1]$ because on $[0, 2]$, f is equal to φ^{-1}, hence $f^{-1}(f(V)) = V$. Hence $h^{-1}(\{1\}) = V$, where $\{1\}$ is measurable but $V \subseteq [0, 2]$ is non-measurable. This shows that h is not measurable. This completes the proof.

4 Integration of measurable functions

Let's first remind ourselves of the basic definition of a Riemann Integrable function. If we say that the function $f : \mathbb{R} \to \mathbb{R}$ is Riemann Integrable, then the integral of f on $[a, b]$, written as $\int_a^b f$, is given by the following two constructions on a partition P of $[a, b]$,

- Lower Sum :
$$L(f, P) = \sum_i m_i(a_i - a_{i-1}) \text{ where } m_i = \inf_{x\in[a_{i-1},a_i]} f(x)$$
- Upper Sum:
$$U(f, P) = \sum_i M_i(a_i - a_{i-1}) \text{ where } M_i = \sup_{x\in[a_{i-1},a_i]} f(x)$$

so that

$$\int_a^b f = L(f, P) = U(f, P).$$

The chain of observations begins now. One can easily write the Lower and Upper Sum as the following simple functions (remember that the partition is finitely many)

$$L(f, P) = \sum_i m_i \lambda\left([a_{i-1}, a_i]\right)$$
$$U(f, P) = \sum_i M_i \lambda\left([a_{i-1}, a_i]\right)$$

Or, equivalently, we can define a lower step function as follows:

$$\phi_P = \sum_i m_i \chi_{[a_{i-1}, a_i]}$$

so that the Riemann integral is simply

$$\int_a^b f(x)dx = \sup_P \int_a^b \phi_P(x)dx$$

where supremum is defined over all partitions. But since, by definition, $\phi_P(x) \leq f(x)\ \forall\ x \in \mathbb{R}$, we can alternatively define Riemann integral as

$$\int_a^b f(x)dx = \sup_{\phi_P \leq f} \int_a^b \phi_P(x)dx \tag{12}$$

where the supremum is defined for all step functions on any partition P.

This definition presented in (12) provides the motivation for extending the definition of Integration from Riemann to Lebesgue. In particular, note the definition of ϕ_P, usual measure on the intervals is applied in Riemann's definition. But, since we know that Borel σ-algebra is a proper subset of $\mathcal{M}_{\lambda^*}$, then it just makes sense to replace $a_{i-1} - a_i$ by $\lambda\left([a_{i-1}, a_i]\right)$ in the motivation that it might generalize the notion of integration.

4.1 Integration of non-negative measurable functions

Definition 4.1.1. (**Lebesgue integral of a simple function**) Consider $\phi : \mathbb{R} \to [0, +\infty)$ be a simple function as

$$\phi = \sum_{i=1}^{N} \alpha_i \chi_{E_i} \text{ where } \alpha_i \geq 0 \text{ and } \lambda\left(E_i\right) < +\infty$$

Then, the Lebesgue integral of ϕ is defined as

$$\int \phi dx = \sum_{i=1}^{N} \alpha_i \lambda\left(E_i\right)$$

Definition 4.1.2. (**Lebesgue integral of a measurable function**) Suppose $f : \mathbb{R} \to [0, +\infty)$ is a λ-measurable function, then the Lebesgue integral of f is defined as

$$\bigstar \quad \boxed{\int f dx = \sup_{\phi \leq f} \int \phi dx \text{ where } \phi \text{ are the } \textbf{simple} \text{ functions} \leq f.} \quad \bigstar$$

Definition 4.1.3. (**Lebesgue integral over a measurable set**) Consider $f : \mathbb{R} \to [0, +\infty)$ to be a λ-measurable function and $E \subseteq \mathbb{R}$ is Lebesgue measurable. Then,

$$\int_E f dx = \int f \cdot \chi_E dx$$

Remark 4.1.4. Therefore, the integral of a non-negative measurable function over a measurable set is given by the integral[18] of restriction of f to it and zero otherwise.

Proposition 4.1.5. *Consider the two λ-measurable functions $f, g : \mathbb{R} \to [0, +\infty)$ and $\phi : \mathbb{R} \to [0, +\infty)$ be a simple-function, then the Lebesgue integral has the following properties:*

1. *Consider two Lebesgue measurable subsets A and B of $\mathbb{R}$ such that $A \cap B = \Phi$. Then,*

$$\int_{A \cup B} \phi dx = \int_A \phi dx + \int_B \phi dx.$$

2. *For any $\alpha \in \mathbb{R}$,*

$$\int \alpha f dx = \alpha \int f dx.$$

3. *Integration for positive valued measurable functions is therefore distributive:*

$$\int (f + g) dx = \int f dx + \int g dx.$$

4. *If $f(x) \leq g(x)$ holds for all $x \in \mathbb{R}$, then*

$$\int f dx \leq \int g dx.$$

5. *Consider A and B be Lebesgue measurable subsets of $\mathbb{R}$ such that $A \subseteq B$. Then,*

$$\int_A f dx \leq \int_B f dx.$$

Proof. **Part 1** : Since ϕ is simple, therefore we can write

$$\phi = \sum_{i=1}^{N} \alpha_i \chi_{E_i}.$$

[18] From now on, any instance of *integral* should be presupposed by Lebesgue integral, of-course, unless otherwise stated, in this text.

Now, by definition

$$\begin{aligned}\int_{A\cup B}\phi dx &= \sum_{i=1}^{N}\alpha_i\lambda\left(E_i\cap(A\cup B)\right)\\ &= \sum_{i=1}^{N}\alpha_i\lambda\left((E_i\cap A)\cup(E_i\cap B)\right)\\ &= \sum_{i=1}^{N}\alpha_i\lambda\left(E_i\cap A\right)+\alpha_i\lambda\left(E_i\cap B\right) && \because E_i\cap A \text{ and } E_i\cap B \text{ are disjoint.}\\ &= \int_A\phi dx+\int_B\phi dx\end{aligned}$$

Part 2 & 3 : Can be seen easily from Theorem 4.2.1.

Part 4 : Note that we define

$$\int f dx = \sup_{\phi\le f}\int\phi dx \text{ where } \phi \text{ are simple functions.}$$

Therefore, for any $\phi \le f$, due to given condition $f \le g$, we would have $\phi \le g$. Hence,

$$\int\phi dx \le \int g dx$$

Since this is true for all simple $\phi \le f$, therefore $\sup_{\phi\le f}\int\phi dx \le \int g dx$, proving the result.

Part 5 : Consider the following:

$$\begin{aligned}\int_A f dx &= \int f\chi_A dx\\ &\le \int f\chi_B dx && \because \chi_A\le\chi_B, \text{ then apply 4.}\\ &= \int_B f dx\end{aligned}$$

Hence proved. □

4.2 Monotone convergence theorem

This is arguably one of the most important theorem in Integration theory,

Theorem 4.2.1. *(**Monotone Convergence Theorem**) Consider a sequence $\{f_n\}$ of $\mathbb{R}\to[0,+\infty)$ of λ-measurable functions which satisfies*

$$f_n(x)\le f_{n+1}(x)\ \forall\ x\in\mathbb{R} \text{ and } n$$

and suppose $\varinjlim_{n\to\infty} f_n$ exists. Then,

$$\int \varinjlim_{n\to\infty} f_n = \varinjlim_{n\to\infty}\int f_n.$$

Proof. Since $f_n \leq f_{n+1}$, therefore,

$$\int f_n \leq \int f_{n+1} \leq \int \varprojlim_{n\to\infty} f_n.$$

Hence,

$$\varprojlim_{n\to\infty} \int f_n \leq \int \varprojlim_{n\to\infty} f_n.$$

Therefore we have proved one inequality.

Now to prove the other inequality, consider any simple function $\phi \leq \varprojlim_{n\to\infty} f$. If we can show that $\int f_n \geq \int \phi$ for any $n \in \mathbb{N}$, then we are done. To this goal, consider $\alpha \in (0,1)$. Construct the set[19]

$$\boxed{E_n = \{x \mid f_n(x) \geq \alpha\phi(x)\}}.$$

Clearly,

$$E_n \subseteq E_{n+1} \; \forall \, n \in \mathbb{N}.$$

Now,

$$\int f_n \geq \int_{E_n} f_n \geq \alpha \int_{E_n} \phi. \tag{13}$$

Moreover, we can see that

$$\text{Claim 1 : } \bigcup_n E_n = \mathbb{R}.$$

This is easy to see as follows:

$$\begin{aligned}
\text{Take } x \in \bigcup_n E_n &\implies x \in E_{i_0} \text{ for some } i_0 \in \mathbb{N}. \\
&\implies x \in \mathbb{R} && \because \; E_n \text{ are subse} \\
\text{Take } x \in \mathbb{R} &\implies \text{Either (1) } x \in \{x \mid f_n(x) - \alpha\phi(x) \geq 0\} \text{ or (2) } x \in \{x \mid f_n(x) - \alpha\phi(x) < 0\} && \text{for any } n \in \mathbb{N}. \\
&\implies \text{If (1), then } x \in E_n, \text{ else if (2), then } \because \phi \leq \varprojlim_{n\to\infty} f_n, \; \exists \, n' \; s.t. \; x \in E_{n'} \\
&\implies x \in \bigcup_n E_n. && \text{Hence Claim 1.}
\end{aligned}$$

Next, we can also see that

$$\text{Claim 2 : } \int_{E_n} \phi \longrightarrow \int \phi$$

[19] After reading the proof, it should appear striking to the reader on actually how much the proof depends on this construction. Both the claims in the following page utilizes this construction E_n to full extent! Hence, it is advised (by Instructor) to purse such effective constructions in the problem sheets and your own proofs.

This can be seen by expanding the *simplicity* of ϕ as follows:

$$
\begin{aligned}
\underleftarrow{\lim}_{n\to\infty} \int_{E_n} \phi &= \sum_{i=1}^{N} a_i \underleftarrow{\lim}_{n\to\infty} \lambda\left(A_i \cap E_n\right) \\
&= \sum_{i=1}^{N} a_i \lambda\left(\bigcup_n A_i \cap E_n\right) && \because\ \{A_i \cap E_n\}_n \text{ is increasing.} \\
&= \sum_{i=1}^{N} a_i \lambda\left(A_i \cap \bigcup_n E_n\right) \\
&= \sum_{i=1}^{N} a_i \lambda\left(A_i \cap \mathbb{R}\right) && \text{Claim 1.} \\
&= \sum_{i=1}^{N} a_i \lambda\left(A_i\right) = \int \phi && \text{Hence Claim 2.}
\end{aligned}
$$

Finally, take limit in (13) to get:

$$
\begin{aligned}
\underleftarrow{\lim}_{n\to\infty} \int f_n \geq \underleftarrow{\lim}_{n\to\infty} \int_{E_n} f_n &\geq \alpha \underleftarrow{\lim}_{n\to\infty} \int_{E_n} \phi \\
&= \alpha \int \phi && \text{Claim 2.} \\
\underleftarrow{\lim}_{n\to\infty} \int f_n \geq \int \phi & && \because\ 0 < \alpha < 1 \text{ is arbitrary.}
\end{aligned}
$$

Hence, **for any** simple function $\phi \leq \underleftarrow{\lim}_{n\to\infty} f_n$, we have concluded that $\int \phi \leq \underleftarrow{\lim}_{n\to\infty} \int f_n$, hence it must be true that

$$\int \underleftarrow{\lim}_{n\to\infty} f_n = \sup_{\phi \leq \underleftarrow{\lim}_n f_n} \int \phi \leq \underleftarrow{\lim}_{n\to\infty} \int f_n.$$

Combining the converse inequality at the beginning, we hence get the desired result. $\square$

Proposition 4.2.2. *Consider a Lebesgue measurable function $f : \mathbb{R} \to [0, +\infty)$. Then,*

$$\int f dx = 0 \Longleftrightarrow f \equiv 0 \textit{ almost everywhere.}$$

Proof. $\mathbf{L} \implies \mathbf{R}$: Consider f is a non-negative real-valued function whose integral is zero. Construct the set,

$$E_n = \left\{x \mid f(x) \geq \frac{1}{n}\right\}.$$

In order to show that $f \equiv 0$ almost everywhere, it is hence sufficient to show that $\lambda\left(E_n\right) = 0\ \forall\ n \in \mathbb{N}$ because it equivalently proves that the measure of the set where f is greater than zero is zero. Now, consider the following function

$$g_n = \frac{1}{n} \chi_{E_n}.$$

Clearly, because g_n is a simple function and

$$\frac{1}{n}\chi_{E_n}(x) = \begin{cases} \frac{1}{n} & \text{if } f(x) \geq \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

which clearly means that $\frac{1}{n}\chi_{E_n} \leq f$, therefore,

$$\begin{aligned} \int f = 0 &= \sup_{\phi \leq f} \int \phi \\ &\geq \int \frac{1}{n}\chi_{E_n} \\ &= \frac{1}{n}\lambda(E_n) \\ \implies \lambda(E_n) &= 0 \;\forall\, n \in \mathbb{N}. \end{aligned}$$

R $\implies$ L : If a non-negative real-valued measurable function f is 0 almost everywhere, then for any simple function $\phi \leq f$, ϕ must also be 0 almost everywhere, so that

$$\begin{aligned} \int \phi &= \sum_{i=1}^{N} \alpha_i \lambda(E_i) \\ &= 0 \end{aligned}$$

Since this is true for any simple $\phi \leq f$, therefore the supremum of all such $\int \phi$ must also be zero, to make $\int f = 0$. □

A simple corollary of the MCT tells us an equivalent story for decreasing sequence of maps where first term is L^1, as compared to the statement of MCT.

Corollary 4.2.3. *Let (X, M, μ) be a measure space and let $f_n : X \to \mathbb{R}$ be a sequence of positive measurable maps. Suppose*

1. $\lim_n f_n(x)$ *exists and is equal to $f(x)$ for some measurable $f : X \to \mathbb{R}$,*
2. $f_n(x) \geq f_{n+1}(x)$ *for all $x \in X$ and $n \in \mathbb{N}$,*
3. $f_1(x) \in L^1$.

Then,

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

Proof. Since $f \leq f_n \leq f_1$, therefore $f \in L^1$. Now, consider the (not necessarily positive!) measurable sequence $g_n = f - f_n$. Since f_n decreases, therefore g_n increases. Now, $\lim_n g_n = 0$ as $\lim_n f_n = f$. Since $0 \in L^1$, therefore Hence, by MCT, we get that $\lim_n \int_X g_n dm = \int_X \lim_n g_n dm$. Expanding it and using the fact that f is in L^1 (so you can cancel $\int_X f dm$ both sides!) gives the desired result. □

Another important result which is of tremendous usability is the fact that Riemann and Lebesgue agree on compact domains(!)

Theorem 4.2.4. *(Riemann = Lebesgue on $[a, b]$) Let $[a, b] \subseteq \mathbb{R}$ be a closed bounded interval and $f : [a, b] \to \mathbb{R}$ be a Riemann integrable map. Then, the Riemann integral and Lebesgue integral of f agrees on $[a, b]$. That is,*

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

where m is the Lebesgue measure of $\mathbb{R}$.

4.3 Fatou's lemma

Theorem 4.3.1. *If $\{f_n\}$ is a sequence of Lebesgue measurable functions from $\mathbb{R}$ to $[0, +\infty)$, then,*

$$\int \liminf_n f_n \leq \liminf_n \int f_n.$$

Proof. We will use Monotone Convergence Theorem to prove this result. Define

$$g_k = \inf_{n \geq k} f_n$$

Therefore $g_k \leq g_{k+1}$ with $g_k \leq f_n \forall\, n \geq k$. Then,

$$\int g_k \leq \int f_n \forall\, n \geq k.$$

This implies that

$$\int g_k \leq \inf_{n \geq k} \int f_n.$$

Now, by MCT,

$$\int \varinjlim_k g_k = \varinjlim_k \int g_k$$

Therefore

$$\begin{aligned} \varinjlim_{k \to \infty} \inf_{n \geq k} \int f_n = \liminf_k \int f_k \geq \varinjlim_{k \to \infty} \int g_k &= \int \varinjlim_{k \to \infty} g_k \\ &= \int \varinjlim_{k \to \infty} \inf_{n \geq k} f_n \\ &= \int \liminf_k f_k \end{aligned}$$

Hence Proved. $\square$

Remark 4.3.2. In fact,

Fatou's Lemma $\iff$ Monotone Convergence Theorem.

4.4 Integration of general $\mathbb{R}$-Valued measurable functions

With the notion of integration of non-negative measurable function in place, it's not difficult to see how can one extend the same notion to measurable functions which takes value in the whole real line.

Definition 4.4.1. (**Lebesgue integral of a Real-Valued measurable Function**) Consider $f : \mathbb{R} \to \mathbb{R}$ to be a measurable function such that

1. $\int f^+ dx < \infty$, and
2. $\int f^- dx < \infty$.

If the above two conditions are satisfied, then f is called Lebesgue Integrable. Then, the Lebesgue integral of f is defined as

$$\star \boxed{\int f dx = \int f^+ dx - \int f^- dx} \star$$

Remark 4.4.2. It's important to note that the integral $\int f dx = \int f^+ dx - \int f^- dx$ is easily defined for any measurable function, but f is called Lebesgue integral only when it's value is finite!

Definition 4.4.3. (**Lebesgue integral over a measurable set**) Consider $f : \mathbb{R} \to \mathbb{R}$ is measurable, $f \cdot \chi_E$ is an Lebesgue Integrable function and $E \subseteq \mathbb{R}$ is also measurable. Then,

$$\int_E f dx = \int f \cdot \chi_E dx.$$

4.4.1 Basic properties of general Lebesgue integral

The following properties are direct extensions of Proposition 4.1.5 to the bigger class of Lebesgue Integrable functions.

Proposition 4.4.4. *Consider $f, g : \mathbb{R} \to \mathbb{R}$ to be Lebesgue Integrable functions. Then,*

1. *For any $\alpha \in \mathbb{R}$, we have:*
$$\int \alpha f dx = \alpha \int f dx.$$
2. *$f + g$ is also Lebesgue Integrable, with*
$$\int (f+g) dx = \int f dx + \int g dx.$$
3. *If $f \equiv 0$ almost everywhere on $\mathbb{R}$, then,*
$$\int f dx = 0.$$
4. *If $f \leq g$ almost everywhere on $\mathbb{R}$, then,*
$$\int f dx \leq \int g dx.$$
5. *If A and B are measurable sets such that $A \cap B = \Phi$, then,*
$$\int_{A \cup B} f dx = \int_A f dx + \int_B f dx.$$

Proof. **S1** : Consider the case that $\alpha \geq 0$. Then,

$$(\alpha f)^+ = \max(\alpha f, 0) = \alpha \max(f, 0) = \alpha f^+$$
$$(\alpha f)^- = -\min(\alpha f, 0) = -\alpha \min(f, 0 = \alpha f^-$$

and since $\int f^+ dx < \infty$ and $\int f^- dx < \infty$, therefore αf is also Lebesgue Integrable, with the integral given as

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

Now consider that $\alpha < 0$, then

$$(\alpha f)^+ = \max(\alpha f, 0) = -|\alpha| \min(f, 0) = |\alpha| f^-$$
$$(\alpha f)^- = -\min(\alpha f, 0) = |\alpha| \max(f, 0) = |\alpha| f^+$$

Hence, αf is again Lebesgue Integrable, with the integral calculated as:

$$\int \alpha f = \int (\alpha f)^+ - \int (\alpha f)^- = |\alpha| \left(\int f^- - \int f^+ \right) = -|\alpha| \int f = \alpha \int f.$$

S2 : First,

$$(f+g)^+ \leq f^+ + g^+$$
$$(f+g)^- \leq f^- + g^-$$

for all $x \in \mathbb{R}$, so that $f + g$ is Lebesgue Integrable. Now,

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

therefore,

$$\begin{aligned}(f+g)^+ - (f+g)^- &= f^+ - f^- + g^+ - g^- \\ (f+g)^+ + f^- + g^- &= (f+g)^- + f^+ + g^+ \\ \int (f+g)^+ + f^- + g^- &= \int (f+g)^- + f^+ + g^+ \\ \int (f+g)^+ + \int f^- + \int g^- &= \int (f+g)^- + \int f^+ + \int g^+ \qquad (\because \text{ of Proposition 4.1.5, S3.}) \\ \int (f+g)^+ - \int (f+g)^- &= \int f^+ - \int f^- + \int g^+ - \int g^- \\ \int (f+g)^+ - \int (f+g)^- &= \int f + \int g\end{aligned}$$

S3 : Given to us is that $f \equiv 0$ almost everywhere. This means that

$$\{x \in \mathbb{R} \mid f(x) \neq 0\} \text{ is of measure } 0.$$

We can write it equivalently as the union of the following two disjoint sets

$$\begin{aligned}\{x \in \mathbb{R} \mid f(x) \neq 0\} &= \{x \mid f(x) > 0\} \cup \{x \in \mathbb{R} \mid f(x) < 0\}.\\ \lambda\left(\{x \in \mathbb{R} \mid f(x) \neq 0\}\right) &= \lambda\left(\{x \mid f^+(x) > 0\}\right) + \lambda\left(\{x \mid f^-(x) < 0\}\right)\\ &= 0\end{aligned}$$

Since measure is positive valued by definition, therefore, these two have to be individually be zero. That is,

$$\lambda\left(\{x \mid f^+(x) > 0\}\right) = \lambda\left(\{x \mid f^-(x) < 0\}\right) = 0$$

Now, by Proposition 4.2.2, we get that

$$\int f^+ = \int f^- = 0$$

which implies that

$$\int f = \int f^+ - \int f^- = 0.$$

S4 : □

Proposition 4.4.5. *If $f : \mathbb{R} \to \mathbb{R}$ is a Lebesgue Integrable Function, then,*

$$\left|\int f dx\right| \leq \int |f| \, dx$$

Proof. Simply note the following:

$$\begin{aligned}\left|\int f dx\right| &= \left|\int \left(f^+ - f^-\right) dx\right|\\ &= \left|\int f^+ dx - \int f^- dx\right|\\ &\leq \left|\int f^+ dx\right| + \left|\int f^- dx\right|\\ &= \int f^+ dx + \int f^- dx\\ &= \int \left(f^+ + f^-\right) dx\\ &= \int |f| \, dx.\end{aligned}$$

□

4.5 Dominated convergence theorem

Theorem 4.5.1. *Let $\{f_n\}$ be a sequence of measurable functions such that there exists a Lebesgue Integrable function g which satisfies*

$$|f_n| \leq g \; \forall \, n.$$

Suppose that the limit $\varprojlim_{n\to\infty} f_n$ exists. Then $\varprojlim_{n\to\infty} f_n$ is Lebesgue Integrable and,

$$\varprojlim_{n\to\infty} \int f_n dx = \int \varprojlim_{n\to\infty} f_n dx.$$

Proof. Since $\{f_n\}$ is a sequence of measurable functions, therefore, $\varprojlim_{n\to\infty} f_n = f$ is also measurable and $|f|$ is bounded by g. But since g is Lebesgue Integrable, and f_n and f are bounded by g, then each f_n and also f are also Lebesgue Integrable (Trivial to see).
Now, $\{g + f_n\}$ is a sequence of measurable functions. Moreover, since $f_n \leq g$ for all n, therefore $f_n + g \geq 0$, so that $\{f_n + g\}$ is a sequence of non-negative measurable functions.
Now using Fatou's Lemma (Theorem 4.3.1), we get,

$$\begin{aligned}
\int \liminf_n (g + f_n) dx &\leq \liminf_n \int (g + f_n) dx \\
\int \left(g + \liminf_n f_n \right) dx &\leq \int g dx + \liminf_n \int f_n dx \\
\int g dx + \int \liminf_n f_n dx &\leq \int g dx + \liminf_n \int f_n dx \\
\int \liminf_n f_n dx &\leq \liminf_n \int f_n dx && \because g \text{ is L.I., so } \int g dx < \infty \\
\int f dx &\leq \liminf_n \int f_n dx && \because \limsup_n x_n = \liminf_n x_n = \varprojlim_n x_n.
\end{aligned}$$

Similarly, since $\{g - f_n\}$ is also a sequence of non-negative measurable functions, therefore we can use Fatou's Lemma to conclude:

$$\begin{aligned}
\int \liminf_n (g - f_n) dx &\leq \liminf_n \int (g - f_n) dx \\
\int \liminf_n \left(-f_n\right) dx &\leq \liminf_n \left(- \int f_n dx \right) \\
- \int \limsup_n f_n dx &\leq - \limsup_n \int f_n dx && \because \liminf_n (-x_n) = -\limsup_n x_n. \\
\int f dx &\geq \limsup_n \int f_n dx
\end{aligned}$$

We hence have that

$$\limsup_n \int f_n dx \leq \int f dx \leq \liminf_n \int f_n dx$$

But it is also true that

$$\liminf_n \int f_n dx \leq \limsup_n \int f_n dx.$$

Hence,

$$\liminf_n \int f_n dx = \limsup_n \int f_n dx = \varprojlim_n \int f_n dx = \int f dx$$

Hence proved. ∎

Proposition 4.5.2. *Consider $\{f_n\}$ to be a sequence of Lebesgue Integrable functions such that*

$$\sum_{n=1}^{\infty}\int |f_n|\,dx < \infty.$$

Then,

1. *The series*
$$\sum_{n=1}^{\infty} f_n(x) \text{ converges almost everywhere on } \mathbb{R}.$$
2. *The sum*
$$f = \sum_{n=1}^{\infty} f_n \text{ is Lebesgue Integrable.}$$
3. *The integral is*
$$\int \sum_{n=1}^{\infty} f_n dx = \sum_{n=1}^{\infty}\int f_n dx.$$

Proof. **S1**. Denote the following:

$$\varphi = \sum_{n=1}^{\infty} |f_n|$$

Clearly, φ is a non-negative measurable function. Since we know that Lebesgue integral for non-negative functions is countably additive, therefore,

$$\int \varphi dx = \sum_{n=1}^{\infty}\int |f_n|\,dx < \infty.$$

Now, if $\int \varphi < \infty$, then φ is finite almost everywhere on $\mathbb{R}$[20]. Now, since $\sum_{n=1}^{\infty} f_n$ is absolutely convergent almost everywhere (last line), hence it is convergent almost everywhere too on $\mathbb{R}$.

S2. Since $\left|\sum_{n=1}^{\infty} f_n\right| \leq \sum_{n=1}^{\infty}|f_n| = \varphi < \infty$ (almost everywhere) and since we can modify the set where φ is not defined (infinite) arbitrarily to make a new function which would be measurable and equal to $\sum_{n=1}^{\infty} f_n$ almost everywhere, therefore $\sum_{n=1}^{\infty} f_n$ would be measurable.

S3. Define

$$\phi_n = \sum_{i=1}^{n} f_i.$$

Clearly, $\phi_n \leq \left|\sum_{i=1}^{n} f_i\right| \leq \sum_{i=1}^{n}|f_i| \leq \sum_{i=1}^{\infty}|f_i| = \varphi$. Therefore, ϕ_n is a sequence of measurable functions and $\phi_n \leq \varphi$ where φ is an Integrable function (given). Therefore, using Dominated

[20]For a non-negative measurable function f with given that $\int f dx < \infty$, the set $E = \{x \in \mathbb{R} \mid f(x) = \infty\}$ together with supposition that $\lambda(E) > 0$ is such that; since $\int f dx = \sup_{\phi \leq f}\int \phi$, therefore, if we take $\phi = n\chi_E$ for any $n > 0$ then $n\chi E < f$. Hence $\int f dx > n\lambda(E)$ for all n, so that $\int f dx = \infty$. But it's a contradiction to $\int f dx < \infty$. Therefore $\lambda(E) = 0$.

Convergence Theorem (4.5.1), we get,

$$\begin{aligned}
\varinjlim_n \int \phi_n dx &= \int \varinjlim_n \phi_n dx \\
\varinjlim_n \int \sum_{i=1}^{n} f_i dx &= \int \varinjlim_n \sum_{i=1}^{n} f_i dx \\
\varinjlim_n \sum_{i=1}^{n} \int f_i dx &= \int \sum_{i=1}^{\infty} f_i dx && \because \text{Proposition 4.4.4, S2} \\
\sum_{i=1}^{\infty} \int f_i dx &= \int \sum_{i=1}^{\infty} f_i dx.
\end{aligned}$$

□

Hence proved.

4.6 Applications-II : Integration

We present important applications of the above results, showcasing the power of their usage. At parts here, we are proving results from Folland's exercises.

Lemma 4.6.1. *The Lebesgue integral*

$$\int_0^1 \frac{x^p - 1}{\log x} dx$$

exists for $p > -1$.

Proof. The first idea is to break p into cases. In some cases, it is obvious why the above integral exists, in others, we have to work. Denote $f_p(x) = \frac{x^p - 1}{\log x}$.

Act 1 : $p > 0$

In this regime, we can bound the $\int_0^1 f_p(x)dx$ by a fixed quantity. Indeed, since $f_p(x)$ is positive, it will suffice. Observe that

$$\frac{x^p - 1}{\log x} = \frac{1 - x^p}{-\log x} \leq \frac{1}{-\log x}.$$

Now, $-\log x$ can be lower bounded by $1 - ax$ for some $0 < a < 1$ by an easy graphical observation. Hence, continuing above, we get

$$\frac{x^p - 1}{\log x} \leq \frac{1}{1 - ax}.$$

The integral then translates to

$$\int_0^1 \frac{x^p - 1}{\log x} dx \leq -\int_0^1 \frac{1}{1 - ax} dx = -\frac{\log(1 - a)}{a} < \infty.$$

Act 2 : $-1 < p < 0$

This is the regime in which we got to work a bit. First, from some graphical observations about $x^p - 1$ and $\log x$, we conclude the following:

1. $x^p - 1$ is positive and $\log x$ is negative, so that $\frac{x^p-1}{\log x}$ is negative.
2. Viewing $1/\log x$ as an **attenuating factor**[21], we see that $0 < 1/\log x < -1$ for $0 < x < 1/e$ and $1/\log x \leq -1$ for $1/e \leq x < 1$.
3. On $1/e < x < 1$, $\log x > 1 - x$. Hence $1/\log x < 1/1 - x$.

With this, we write our integral as

$$\begin{aligned}\int_0^1 \frac{x^p - 1}{\log x} dx &= \int_0^{1/e} \frac{x^p - 1}{\log x} dx + \int_{1/e}^1 \frac{x^p - 1}{\log x} dx \\ &< \int_0^{1/e} (1 - x^p) dx + \int_0^1 \frac{x^p - 1}{x - 1} dx\end{aligned}$$

Now the first integral is bounded while the second is bounded as the derivative of x^p exists at $x = 1$. □

Lemma 4.6.2. *Let $f : \mathbb{R} \to \mathbb{R} \cup \{\infty, -\infty\}$ be a measurable map with $(\mathbb{R}, M, m)$ be a measure structure on $\mathbb{R}$. If there exists $M > 0$ such that for all $E \in M$ such that $0 < m(E) < \infty$ we have that*

$$\left| \frac{1}{m(E)} \int_E f dm \right| < M,$$

then

$$|f(x)| \leq M \ \ a.e..$$

Proof. Let $A = \{x \in \mathbb{R} \mid |f(x)| > M\}$. We can write it as $A = A_+ \cup A_-$ where $A_+ = \{x \in \mathbb{R} \mid f(x) > M\}$ and $A_- = \{x \in \mathbb{R} \mid f(x) < -M\}$. Clearly these are disjoint and covers A. Hence, we wish to show

$$m(A) = m(A_+) + m(A_-) = 0$$

which is equivalent to showing that $m(A_+) = m(A_-) = 0$ as measures are always positive.

$$\textbf{Act 1} : m(A_+) = 0.$$

The way A_+ and A_- are defined, it is natural for the next step to be a consideration of integral of f over these. Indeed, we observe that, due to the fact that $f \in L^1$ and $A_+ \subseteq \mathbb{R}$

$$Mm(A_+) = \int_{A_+} M \leq \int_{A_+} |f| \leq \int_{\mathbb{R}} |f| \, dm < \infty.$$

Thus, $\infty > \int_{A_+} f dm \geq Mm(A_+)$. Note we dropped the absolute sign as f is positive on A_+. Hence $m(A_+) \neq \infty$.

[21] we view $1/\log x$ as an attenuating factor instead of $x^p - 1$ as if we remove $1/\log x$, then we would be left with $x^p - 1$, whose integral is easy to find.

Now suppose $0 < m(A_+) < \infty$. Then by hypothesis, we can write

$$\int_{A_+} f dm < Mm(A_+),$$

which is a contradiction. Hence $m(A_+) = 0$.

Act 2 : $m(A_-) = 0$.

Again using $f \in L^1$ and $A_- \subseteq \mathbb{R}$, we get

$$\int_{A_-} |f|\, dm \leq \int_{\mathbb{R}} |f|\, dm < \infty.$$

Since $\left|\int_{A_-} f dm\right| \leq \int_{A_-} |f|\, dm$ and since $\int_{A_-} f dm < \int_{A_-} -M dm = -Mm(A_-)$ so that $\left|\int_{A_-} f dm\right| > Mm(A_-)$, therefore we get

$$Mm(A_-) < \left|\int_{A_-} f dm\right| \leq \int_{A_-} |f|\, dm < \infty.$$

Hence $m(A_-) \neq \infty$. Now with this, if we assume $\infty > m(A_-) > 0$, then by hypothesis, we obtain

$$\left|\int_{A_-} f dm\right| \leq m(A_-)M,$$

which contradicts the above inequality. Hence $m(A_-) = 0$.

$\square$

Lemma 4.6.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable map where the domain $\mathbb{R}$ has a measure structure $(\mathbb{R}, M, m)$. If $f \in L^1$ and $f \geq 0$, then for all $E \in M$*

$$\lim_{n\to\infty} \int_E f^{\frac{1}{n}} dm = m(E).$$

Proof. The fundamental observation that one has to make here is that if $y \in [0, \infty)$, then $y^{1/n}$ increases to 1 on $(0, 1]$ and $y^{1/n}$ decreases to 1 on $(1, \infty)$. Indeed, pick any $E \in M$ and define

$$E_{\leq} := E \cap \{x \in \mathbb{R} \mid f(x) \leq 1\}$$
$$E_{>} := E \cap \{x \in \mathbb{R} \mid f(x) > 1\}.$$

We thus have a disjoint measurable cover of E and hence $m(E) = m(E_{\leq}) + m(E_{>})$. Hence we get that

$$\lim_{n\to\infty} \int_E f^{\frac{1}{n}} dm = \lim_{n\to\infty} \int_{E_{\leq}} f^{\frac{1}{n}} dm + \lim_{n\to\infty} \int_{E_{>}} f^{\frac{1}{n}} dm.$$

Now, we have two integrals to consider.

Act 1 : $\lim_{n\to\infty} \int_{E_{\leq}} f^{\frac{1}{n}} dm = m(E_{\leq})$.

Since $f^{\frac{1}{n}}$ is a sequence of positive measurable maps increasing to 1, therefore by MCT, we get that

$$\begin{aligned}\lim_{n\to\infty}\int_{E_{\leq}} f^{\frac{1}{n}}dm &= \int_{E_{\leq}} \lim_{n\to\infty} f^{\frac{1}{n}}dm\\ &= \int_{E_{\leq}} 1dm\\ &= m(E_{\leq}).\end{aligned}$$

$$\textbf{Act 2}: \lim_{n\to\infty}\textstyle\int_{E_>} f^{\frac{1}{n}}dm = m(E_>).$$

It is this place where we will have to use the fact that $f \in L^1$. Since $f^{\frac{1}{n}}$ is a sequence of positive measurable maps decreasing to 1 where f is L^1. Hence, by Corollary 4.2.3 of MCT, we get that

$$\begin{aligned}\lim_{n\to\infty}\int_{E_>} f^{\frac{1}{n}}dm &= \int_{E_>} \lim_{n\to\infty} f^{\frac{1}{n}}dm\\ &= \int_{E_>} 1dm\\ &= m(E_>).\end{aligned}$$

This completes the proof, as we have showed $\lim_{n\to\infty}\int_E f^{\frac{1}{n}}dm = \lim_{n\to\infty}\int_{E_{\leq}} f^{\frac{1}{n}}dm + \lim_{n\to\infty}\int_{E_>} f^{\frac{1}{n}}dm = m(E_{\leq}) + m(E_>) = m(E)$. □

Lemma 4.6.4. *Let (X, M, m) be a measure space and $f : X \times [a, b] \to \mathbb{C}$ be a function such that $f(x, t) : X \to \mathbb{C}$ is measurable for all $t \in [a, b]$. Let $F(t) := \int_X f(x, t)dm$. Suppose there exists $g \in L^1$ such that*

$$|f(x,t)| \leq |g(x)|\, \forall x \in X$$

for every $t \in [a, b]$. If $\lim_{t\to t_0} f(x, t) = f(x, t_0)$ for every $x \in X$, then

$$\lim_{t\to t_0} F(t) = F(t_0).$$

Proof. Clearly we should use DCT. However, we first need to get a sequence of functions for it. Indeed, since we know that $\lim_{t_n\to t_0} f(x, t) = f(x, t_0)$, thus for any sequence $t_n \to t_0$, we have $\lim_{n\to\infty} f(x, t_n) = f(x, t_0)$. Hence we may define $f_n(x) = f(x, t_n)$ which are by definition measurable. Moreover, we have $|f_n(x)| \leq |g(x)|$ for all $x \in X$ where $g \in L^1$. Hence, by DCT, we obtain

$$\begin{aligned}\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty}\int_X f_n(x)dm &= \int_X \lim_{n\to\infty} f_n(x)dm\\ &= \int_X f(x, t_0)dm\\ &= F(t_0).\end{aligned}$$

Since $t_n \to t_0$ is arbitrary, therefore $\lim_{t\to t_0} F(t) = F(t_0)$. □

5 The L^p spaces

We now turn into some more abstract formulation for analysis of measurable functions, by analyzing their class formed under certain definitions.

Definition 5.0.1. (**L^p norm of a function**) Consider any function f and $p > 0$. The L^p norm of f, denoted $\|f\|_p$, is defined as:

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

Definition 5.0.2. (**The L^p Space**) Consider $(X, \mathcal{A}, \mu)$ to be a measure space. Suppose $p > 0$. Then, the class of measurable functions defined as:

$$L^p\left(X, \mathcal{A}, \mu\right) = \{f : X \to \mathbb{R} \mid \|f\|_p < \infty\} / (f \sim g \iff f = g \text{ a.e.}).$$

Moreover, two measurable functions $f, g \in L^p\left(X, \mathcal{A}, \mu\right)$ are said to be equal if and only if:

$$f = g \text{ almost everywhere on } \mathbb{R}.$$

Remark 5.0.3. Note that $L^p\left(X, \mathcal{A}, \mu\right)$ is just the class of Integrable functions when $p = 1$.

Remark 5.0.4. Note carefully the use of word *class* rather than set. It is because that an element of $L^p\left(X, \mathcal{A}, \mu\right)$ is not a function, but a class of functions identified by the relation $f \sim g$ if and only if $f = g$ almost everywhere. But for out purposes, one can get away by writing $f \in L^p\left(X, \mathcal{A}, \mu\right)$ to mean that f is measurable and $\|f\|_p < \infty$ so that $|f|^p$ is Integrable.

5.1 Algebraic properties of L^p space

We will now see some of the properties of L^p Spaces which reflects it's algebraic nature. In particular, we would prove that L^p is a vector space for any $p > 0$. But proving that $\|\cdot\|_p$ is actually the norm for functions in L^p $(p \geq 1)$ would require a lot of construction.

5.1.1 L^p is a vector space

Proposition 5.1.1. *Consider $(X, \mathcal{A}, \mu)$ to be a measure space. Then, the L^p space*

$$L^p\left(X, \mathcal{A}, \mu\right) \text{ is a } \textbf{Vector Space}.$$

Proof. First, let's deal with the scalar multiplication. Note that the ground field here is $\mathbb{R}$. For any $a, b \in \mathbb{R}$ and $f, g \in L^p\left(X, \mathcal{A}, \mu\right)$, we trivially have:

$$\begin{aligned}(ab)f &= a(bf)\\ 1f &= f\\ a(f+g) &= af + ag\\ (a+b)f &= af + bf\end{aligned}$$

Now, to show that $L^p(X, \mathcal{A}, \mu)$ is an abelian group under addition, the associativity, commutativity, identity (f such that $f = 0$ a.e.) and inverse (for f, $-f$ is the inverse) follows trivially. What remains to be shown is that for $f, g \in L^p(X, \mathcal{A}, \mu)$, $f + g \in L^p(X, \mathcal{A}, \mu)$ too. To see this, note that we know already, that $f + g$ is measurable, what we need to then show is that

$$\text{To Show : } \|f + g\|_p < \infty$$

for any $p > 0$. All we need to show is therefore,

$$\int |f + g|^p < \infty$$

To see this, note:

$$\begin{aligned} |f + g|^p &\leq (|f| + |g|)^p \\ &\leq 2^p \max(|f|^p, |g|^p) \\ &\leq 2^p (|f|^p + |g|^p) \end{aligned}$$

By Proposition 4.4.4 S4,

$$\int |f + g|^p \leq 2^p \left(\int |f|^p + \int |g|^p \right) < \infty$$

Therefore, $L^p(X, \mathcal{A}, \mu)$ is a Vector Space. $\square$

5.1.2 norm on L^p vector space

We first see that the norm defined at the beginning is actually not a norm in the case when $p < 1$. Therefore, L^p Vector Space with norm $\|\cdot\|_p$ would make sense only when $p \geq 1$.

Definition 5.1.2. (**Norm on a vector space**) Consider a Vector Space $(V, \mathbb{R})$. A norm $\|\cdot\|$ on V is a function

$$\|\cdot\| : (V, \mathbb{R}) \to [0, \infty)$$

satisfying following three conditions:

1. For any $x \in (V, \mathbb{R})$,
$$\|x\| = 0 \iff x = 0_V$$
2. For any $x \in (V, R)$ and $\alpha \in \mathbb{R}$,
$$\|\alpha x\| = |\alpha| \, \|x\|$$
3. For any $x, y \in (V, \mathbb{R})$
$$\|x + y\| \leq \|x\| + \|y\|$$

Now suppose $0 < p < 1$, then, it is simple to see $\|\cdot\|_p$ does not follow Triangle Inequality on $L^p(X, \mathcal{A}, \mu)$. To see this, note that for any $a, b > 0$ and $p \in (0, 1)$, we have:

$$a^p + b^p > (a + b)^p \tag{14}$$

This comes naturally from the relation:

$$t^{p-1} > (a + t)^{1-p}$$

and then it's integration.
Using (14), we can see that for any two sets $E, F \in \mathcal{A}$ such that $E \cap F = \Phi$, if we write

$$a = \mu(E)^{1/p} = \left(\int |\chi_E|^p\right)^{\frac{1}{p}} = \|\chi_E\|_p$$
$$b = \mu(F)^{1/p} = \left(\int |\chi_F|^p\right)^{\frac{1}{p}} = \|\chi_F\|_p$$

then,

$$\begin{aligned}
\|\chi_E + \chi_F\|_p &= \left(\int |\chi_E + \chi_F|^p\right)^{\frac{1}{p}} \\
&= \left(\int (\chi_E + \chi_F)^p\right)^{\frac{1}{p}} \\
&= \left(\int (\chi_E^p + \chi_F^p)\right)^{\frac{1}{p}} && \because \chi_E \cdot \chi_F = \chi_{E\cap F} = \chi_\Phi = 0 \\
&= \left(\int \chi_E^p + \int \chi_F^p\right)^{\frac{1}{p}} \\
&= (a^p + b^p)^{\frac{1}{p}} \\
&> a + b && \text{Take power } \frac{1}{p} \text{ both sides of Eq. (14)} \\
&= a + b \\
&= \|\chi_E\|_p + \|\chi_F\|_p
\end{aligned}$$

Hence, there exists functions in Vector Space $L^p(X, \mathcal{A}, \mu)$ for $p \in (0,1)$ such that $\|\cdot\|_p$ does not satisfies the Δ-Inequality, hence $\|\cdot\|_p$ is not a norm on the vector space $L^p(X, \mathcal{A}, \mu)$ for $p \in (0,1)$.

But what about $p \geq 1$**?** It turns out we need more revelations, in terms of results, to prove that for $p \geq 1$, $\|\cdot\|_p$ is a norm on the vector space $L^p(X, \mathcal{A}, \mu)$. We now discuss those *revelations*.

Lemma 5.1.3. *Consider $a \geq 0$, $b \geq 0$ and $0 < \lambda < 1$, then*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda) b.$$

Proof. Consider the convex function e^x. Since it is convex, therefore,

$$\begin{aligned}
a^\lambda b^{1-\lambda} &= e^{\lambda \ln a + (1-\lambda)\ln b} \\
&\leq \lambda e^{\ln a} + (1-\lambda) e^{\ln b} \\
&= \lambda a + (1-\lambda) b
\end{aligned}$$

□

5.1.3 Hölder's inequality

One of the important & frequently used inequalities which would be a stepping stone to show that $\|\cdot\|_p$ is a norm on $L^p(X, \mathcal{A}, \mu)$ for $p \geq 1$.

Theorem 5.1.4. ***(Hölder's inequality)*** *Consider* $1 < p, q < \infty$ *such that*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then, for any $f \in L^p(X, \mathcal{A}, \mu)$ *and* $g \in L^q(X, \mathcal{A}, \mu)$

1. $$fg \in L^1(X, \mathcal{A}, \mu)$$

2. $$\int |fg| \leq \left(\int |f|^p\right)^{\frac{1}{p}} \cdot \left(\int |g|^q\right)^{\frac{1}{q}}$$

   .

   OR

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$$

Proof. From the above Lemma, we have that for any $a > 0$ and $b > 0$, the following holds:

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}$$

Now, if we set

$$a = \frac{|f|^p}{(\|f\|_p)^p}$$
$$b = \frac{|g|^q}{(\|g\|_q)^q}$$

and then use the inequality in above lemma, we get:

$$\frac{|f|\,|g|}{\|f\|_p \|g\|_q} \leq \frac{1}{p} \cdot \frac{|f|^p}{(\|f\|_p)^p} + \frac{1}{q} \cdot \frac{|g|^q}{(\|g\|_q)^q}.$$

Now, because $|f|\,|g| = |fg|$, therefore from above inequality, we see that

$$\int |fg| < \infty$$

hence $fg \in L^1(X, \mathcal{A}, \mu)$. Furthermore, since we know that inequality is preserved in Integration, therefore integrating the above inequality leads to the following:

$$\int \frac{|f|\,|g|}{\|f\|_p \|g\|_q} \leq \int \frac{1}{p} \cdot \frac{|f|^p}{(\|f\|_p)^p} + \int \frac{1}{q} \cdot \frac{|g|^q}{(\|g\|_q)^q}$$
$$\frac{1}{\|f\|_p \|g\|_q} \int |fg| \leq \frac{1}{p\,(\|f\|_p)^p} \int |f|^p + \frac{1}{q\,(\|g\|_q)^q} \int |g|^q$$
$$\frac{\|fg\|_1^1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$
$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Hence proved. $\square$

Remark 5.1.5. With Hölder's Inequality, we are one step closer to proving that $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ for any $f, g \in L^p(X, \mathcal{A}, \mu)$ where $1 \leq p < \infty$, to formally make $\|\cdot\|_p$ a norm on the vector space $L^p(X, \mathcal{A}, \mu)$. **This is finally proved by Minkowski's Inequality** which we prove now:

5.1.4 Minkowski's inequality

Theorem 5.1.6. *(**Minkowski's inequality**) : Consider any $f, g \in L^p(X, \mathcal{A}, \mu)$ and $1 \leq p < \infty$. Then*

$$\left(\int |f+g|^p\right)^{\frac{1}{p}} \leq \left(\int |f|^p\right)^{\frac{1}{p}} + \left(\int |g|^p\right)^{\frac{1}{p}}$$

OR,

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. Since $|f+g| \leq |f| + |g|$, therefore if $p = 1$, then the result follows immediately. Now consider $p > 1$. Moreover, suppose that $q > 1$ is such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Note that this also leads to following equations

$$\begin{aligned} (p-1)q &= p \\ p\left(1 - \frac{1}{q}\right) &= 1 \end{aligned}$$

Now, with this, we can bound $\|f+g\|_p^p$ as follows:

$$\begin{aligned}
\|f+g\|_p^p &= \int |f+g|^p \\
&= \int |f+g| \cdot |f+g|^{p-1} \\
&\leq \int (|f|+|g|) \cdot |f+g|^{p-1} \\
&= \int |f| \cdot |f+g|^{p-1} + \int |g| \cdot |f+g|^{p-1} \\
&= \int \left||f| \cdot |f+g|^{p-1}\right| + \int \left||g| \cdot |f+g|^{p-1}\right| \\
&= \int \left|f \cdot (f+g)^{p-1}\right| + \int \left|g \cdot (f+g)^{p-1}\right| \\
&= \|f \cdot (f+g)^{p-1}\|_1 + \|g \cdot (f+g)^{p-1}\|_1 \\
&\leq \|f\|_p \cdot \|(f+g)^{p-1}\|_q + \|g\|_p \cdot \|(f+g)^{p-1}\|_q \qquad \text{Hölder's Inequality (5.1.4)} \\
&= (\|f\|_p + \|g\|_p) \cdot \|(f+g)^{p-1}\|_q \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int \left|(f+g)^{p-1}\right|^q\right)^{\frac{1}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int (f+g)^{(p-1)q}\right)^{\frac{1}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int (f+g)^{p}\right)^{\frac{1}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int (f+g)^{p}\right)^{\frac{1}{p}\cdot\frac{p}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \|f+g\|_p^{\frac{p}{q}} \\
\frac{\|f+g\|_p^p}{\|f+g\|_p^{\frac{p}{q}}} &\leq (\|f\|_p + \|g\|_p) \\
\|f+g\|_p^{p\left(1-\frac{1}{q}\right)} &\leq (\|f\|_p + \|g\|_p) \\
\|f+g\|_p &\leq \|f\|_p + \|g\|_p
\end{aligned}$$

Hence proved. □

Remark 5.1.7. ★ Hence, in continuation of our effort to prove that $\|\cdot\|_p$ **is a norm on the vector space** $L^p(X, \mathcal{A}, \mu)$ **for** $1 \leq p < \infty$, we can now satisfactorily say that it is indeed such, especially by Minkowski's Inequality just proved. One also calls a vector space with norm a norm space.

5.2 Properties of L^1 maps

We would in this section quickly portray some of the easy properties of L^1-maps which are good to keep in mind. The first tells us that a high schooler's dream of claiming a map to be zero if integral is zero is *almost* true for L^1 maps.

Lemma 5.2.1. *Let $f : X \to \mathbb{C}$ be a measurable map where (X, M, m) is a measure space. Suppose $f \in L^1$. Then, $\int_F f dm = 0$ for all $F \in M$ if and only if $f = 0$ almost everywhere.*

Proof. One side is trivial. For the other, we may reduce to the case when f is real valued. Let $A = \{x \in X \mid f^-(x) > 0\}$. As f^- is measurable, therefore $A \in M$. Since $\int_A f dm = 0$, therefore $\int_A f^+ - f^- dm = \int_A f^+ = \int_A f^- dm$. If $x \in A$, then $f^-(x) > 0$, and hence $f^+(x) = 0$. Hence $\int_A f^+ dm = 0$ and hence $\int_A f^- dm = 0$. Since $f^- \geq 0$, therefore $f^- = 0$ almost everywhere. We thus have $\int_X f dm = \int_X f^+ dm = 0$ as $X \in M$. Since $f^+ \geq 0$, therefore $f^+ = 0$ almost everywhere. □

Lemma 5.2.2. *Let (X, M, m) be a measure space and $f : X \to \mathbb{R}$ be a measurable map with $f \geq 0$. Then,*

$$m(\{x \in X \mid f(x) = \infty\}) = 0.$$

Proof. This again uses the standard idea of breaking the set which we wish to measure into sets whose bounds on measure is known. Indeed, observe that

$$E := \{f(x) = \infty\} = \bigcap_{n \in \mathbb{N}} \{f(x) > n\} =: \bigcap_{n \in \mathbb{N}} E_n.$$

Moreover, $\{E_n\}$ is decreasing. Thus,

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

Now we obtain bound on $m(E_n)$. Indeed,

$$nm(E_n) = \int_{E_n} ndm \leq \int_{E_n} f(x)dm \leq \int_X f(x)dm =: I < \infty.$$

Thus $m(E_n) \leq I/n$. Hence $\lim_{n \to \infty} m(E_n) = 0$. □

5.3 Completeness of norm space $L^p(X, \mathcal{A}, \mu)$

We now see that the norm space $L^p(X, \mathcal{A}, \mu)$ is actually a complete metric space on the metric induced by the norm! But before stating the result, let us revisit the definitions of *series*, *Cauchy sequences* & *completeness* for any arbitrary norm space $(V, \mathbb{R}, \|\cdot\|)$.

5.3.1 General definitions and results in normed spaces

Definition 5.3.1. (**Convergent sequence**) Let $(V, \mathbb{R}, \|\cdot\|)$ be a norm space and $\{x_n\}$ be a sequence in it. Then $\{x_n\}$ is said to converge to $x \in (V, \mathbb{R}, \|\cdot\|)$ if

$$\|x_n - x\| \longrightarrow 0 \text{ as } n \to \infty.$$

Definition 5.3.2. (**Cauchy sequence**) Let $(V, \mathbb{R}, \|\cdot\|)$ be a norm space and $\{x_n\}$ be a sequence in it. Then $\{x_n\}$ is said to be a Cauchy sequence in $(V, \mathbb{R}, \|\cdot\|)$ if

$$\forall \epsilon > 0, \; \exists\, N \in \mathbb{N} \text{ such that } \|x_n - x_m\| < \epsilon \; \forall\, n, m \geq N.$$

Definition 5.3.3. (**Complete norm Space** or **Banach Space**) A norm space $(V, \mathbb{R}, \|\cdot\|)$ is called a Complete Metric Space or a Banach Space if

$$\text{Every Cauchy sequence in } (V, \mathbb{R}, \|\cdot\|) \text{ is convergent in } (V, \mathbb{R}, \|\cdot\|).$$

Definition 5.3.4. (**Series in a norm Space**) A series in a norm space $(V, \mathbb{R}, \|\cdot\|)$ is defined as

$$\sum_{n=1}^{\infty} x_n \text{ where } x_n \in (V, \mathbb{R}, \|\cdot\|).$$

Definition 5.3.5. (**Convergent series in a norm space**) A series $\sum_{n=1}^{\infty} x_n$ in a norm space $(V, \mathbb{R}, \|\cdot\|)$ is said to be convergent if the sequence

$$\{S_n\} \text{ where } S_n = \sum_{i=1}^{n} x_i \text{ is convergent in } (V, \mathbb{R}, \|\cdot\|).$$

Definition 5.3.6. (**Absolutely convergent series**) Consider a series $\sum_{n=1}^{\infty} x_n$ in a norm space $(V, \mathbb{R}, \|\cdot\|)$. Then it is called absolutely convergent if and only if

$$\sum_{n=1}^{\infty} \|x_n\| < \infty.$$

We now see the equivalent condition needed for a norm space to become a complete norm space:

Theorem 5.3.7. *(***Equivalent condition for a Banach space***) Suppose that $(V, \mathbb{R}, \|\cdot\|)$ is a norm space. Then,*

$$(V, \mathbb{R}, \|\cdot\|) \textit{ is a Complete norm Space (or Banach Space)} \iff \textit{Every Absolutely Convergent Series is also Con}$$

Proof. **L** $\implies$ **R** : Suppose $(V, \mathbb{R}, \|\cdot\|)$ is a Banach Space. Hence any Cauchy sequence in it converges at a point within it. Now, take any Absolutely Convergent series, say,

$$\sum_{n=1}^{\infty} x_n$$

in $(V, \mathbb{R}, \|\cdot\|)$. This means that

$$\sum_{n=1}^{\infty} \|x_n\| < \infty.$$

Now this also means that if we write $S_n = \sum_{i=1}^{n} x_i$, then

$$\begin{aligned} \|S_n - S_m\| &= \|\sum_{i=1}^{n} x_i - \sum_{i=1}^{m} x_i\| \\ &= \|\sum_{i=n}^{m} x_i\| \\ &\leq \sum_{i=n}^{m} \|x_i\| \end{aligned}$$

Now since $\sum_{n=1}^{\infty} x_n$ is absolutely convergent, therefore,

$$\sum_{i=n}^{m} \|x_i\| \leq \sum_{i=1}^{\infty} \|x_i\| < \infty \; \forall \, n \leq m \in \mathbb{N}$$

also with the fact that $\sum_{i=n}^{m} \|x_i\| \longrightarrow 0$ as $n, m \to \infty$. Hence,

$$\|S_n - S_m\| < \infty \text{ and } \|S_n - S_m\| \to 0 \text{ as } n, m \to \infty.$$

Hence, $\{S_n\}$ is a Cauchy sequence in $(V, \mathbb{R}, \|\cdot\|)$ and thus is convergent. Therefore, $\sum_{i=1}^{\infty} x_i$ is also convergent.

R $\implies$ L : Suppose that $(V, \mathbb{R}, \|\cdot\|)$ is a norm space with given that every absolutely convergent series converges. Since we need to show that $(V, \mathbb{R}, \|\cdot\|)$ is then a Banach space, hence we now consider any arbitrary Cauchy sequence, say, $\{x_n\}$.
Now, construct a new sequence from the taken Cauchy sequence $\{x_n\}$ as $\{y_n\}$ defined by the following:

$$\begin{aligned}
y_1 &= x_{N_1} \text{ where } N_1 \text{ is such that } \|x_n - x_m\| < \frac{1}{2^1} \; \forall \, n, m \geq N_1 \\
y_2 &= x_{N_2} - x_{N_1} \text{ where } N_2 \text{ is such that } \|x_n - x_m\| < \frac{1}{2^2} \; \forall \, n, m \geq N_2 > N_1 \\
\vdots &= \vdots \\
y_k &= x_{N_k} - x_{N_{k-1}} \text{ where } N_k \text{ is such that } \|x_n - x_m\| < \frac{1}{2^k} \; \forall \, n, m \geq N_k > N_{k-1}.
\end{aligned}$$

Now, with $\{y_n\}$ in hand, we see some peculiar properties of it, such as:

$$\sum_{j=1}^{k} y_j = x_{N_k}.$$

and especially, we see that $\sum y_n$ is absolutely convergent(!) as follows:

$$\begin{aligned}
\sum_{j=1}^{\infty} \|y_j\| &\leq \|y_1\| + \sum_{j=1}^{\infty} \|y_j\| \\
&\leq \|x_{N_1}\| + \sum_{j=1}^{\infty} \frac{1}{2^j} \\
&= \|x_{N_1}\| + 1 < \infty \qquad\qquad \because \{x_n\} \text{ is Cauchy, so } \|x_i\| < \infty \forall i.
\end{aligned}$$

Now, since we are given that every absolutely convergent series in $(V, \mathbb{R}, \|\cdot\|)$ converges, therefore $\sum y_n$ also converges in $(V, \mathbb{R}, \|\cdot\|)$. But convergence of a series means convergence of it's sequence of partial sums $S_n = \sum_{i=1}^{n} y_i$ and $S_n = x_{N_n}$ as shown above. Therefore, we have

$$\{x_{N_n}\} \text{ converges in } (V, \mathbb{R}, \|\cdot\|).$$

Since $\{x_{N_n}\}$ converges, therefore, if we suppose $x_{N_n} \to x$, then:

$$\begin{aligned} \|x_n - x\| &= \|x_n - x_{N_n} + x_{N_n} - x\| \\ &\leq \|x_n - x_{N_n}\| + \|x_{N_n} - x\| && 2^{nd} \text{ term} \to 0. \\ &= \|x_n - x_{N_n}\| \end{aligned}$$

Now, we know that $n < N_n$, therefore, $\exists p \in \mathbb{N}$ such that $N_n > n \geq N_{n-p}$. Hence

$$\|x_n - x_{N_n}\| < \frac{1}{2^{n-p}}$$

and as $n \to \infty$, $\|x_n - x_{N_n}\| \to 0$ too. Therefore,

$$\|x_n - x\| \to 0.$$

Hence, $\{x_n\}$ is a convergent sequence, apart from being Cauchy. Since the choice of $\{x_n\}$ was arbitrary, therefore all Cauchy sequences are convergent. Hence $(V, \mathbb{R}, \|\cdot\|)$ is a Complete norm Space or Banach Space. □

5.3.2 $L^p(X, \mathcal{A}, \mu)$ is a Banach space!

We now see that $L^p(X, \mathcal{A}, \mu)$ is a Complete norm Space.

Theorem 5.3.8. *The normed vector space $L^p(X, \mathcal{A}, \mu)$ for $1 \leq p < \infty$ is a Banach Space.*

Proof. From the Theorem 5.3.7, we just need to equivalently show that any absolutely convergent series is convergent.
Now consider $\{f_k\}$ in $L^p(X, \mathcal{A}, \mu)$ to be absolutely convergent, so that

$$\sum_{k=1}^{\infty} \|f_k\|_p = B < \infty.$$

Also consider the following sequence:

$$G_n = \sum_{k=1}^{n} |f_k| \text{ and } G = \sum_{k=1}^{\infty} |f_k|.$$

Clearly, for all $n \in \mathbb{N}$ we have

$$\begin{aligned} \|G_n\|_p &= \|\sum_{k=1}^{n} |f_k|\,\| \\ &\leq \sum_{k=1}^{n} \|f_k\|_p \leq B < \infty. \end{aligned}$$

Also note that $\{G_n\}$ is an increasing sequence of positive-valued measurable functions.
Since $\varinjlim_n G_n$ exists, therefore, by the Monotone convergence theorem (Theorem 4.2.1), we have:

$$\begin{aligned} \int \varinjlim_n G_n^p &= \int G^p \\ &= \varinjlim_n \int G_n^p \\ &\leq B^p \end{aligned}$$

Therefore, we have that $\int G^p$ is finite almost everywhere on $\mathbb{R}$ (just consider the above result that $\int (G^p - \chi_X) = 0$ where $\mu(X) = B^p$.)
Since G^p is finite almost everywhere, therefore G is finite almost everywhere. Hence, we get

$$\begin{aligned} \sum_{k=1}^{\infty} f_k &\leq \sum_{k=1}^{\infty} |f_k| \\ &= G < \infty \ \text{ almost everywhere.} \end{aligned}$$

Now write

$$F = \sum_{k=1}^{\infty} f_k.$$

Clearly, we have

$$\begin{aligned} |F| &= \left| \sum_{k=1}^{\infty} f_k \right| \\ &\leq \sum_{k=1}^{\infty} |f_k| \\ &= G < \infty \end{aligned}$$

and since f_k are members of the vector space $L^p(X, \mathcal{A}, \mu)$, we also have that $F \in L^p(X, \mathcal{A}, \mu)$.
Now, we see that

$$\begin{aligned} \left| F - \sum_{k=1}^{n} f_k \right|^p &\leq |F| + \left| \sum_{k=1}^{n} f_k \right| \\ &\leq G + G = 2G \\ &\leq (2G)^p \ \because 1 \leq p < \infty. \\ &< \infty \end{aligned}$$

Now since $|F - \sum_{k=1}^{n} f_k|^p < \infty$, hence it is in $L^1(X, \mathcal{A}, \mu)$.
With the above inequality, we see that $|F - \sum_{k=1}^{n}|$ is finite and is absolutely bounded by another measurable function for each n, hence, we can now use the Dominated Convergence Theorem (Theorem 4.5.1) to write

$$\begin{aligned} \left(\varinjlim_n \int \left| F - \sum_{k=1}^{n} f_k \right|^p \right)^{\frac{1}{p}} &= \left(\int \varinjlim_n \left| F - \sum_{k=1}^{n} f_k \right|^p \right)^{\frac{1}{p}} \\ \varinjlim_n \| F - \sum_{k=1}^{n} f_k \|_p = 0 & \qquad \text{Note that } F = \sum_{k=1}^{\infty} f_k \in L^p(X, \mathcal{A}, \mu). \end{aligned}$$

Hence, we have

$$\sum_{k=1}^{\infty} f_k = F \in L^p(X, \mathcal{A}, \mu)$$

that is, the absolutely convergent series $\sum_{k=1}^{f_k}$ is also convergent in the same space! Therefore, $L^p(X, \mathcal{A}, \mu)$ is a Banach Space. $\square$

6 Product measure

We now turn to product measure spaces. This concept would help us to formalize the notion of double (or higher) integration over the so defined *product measure spaces.* In fact, this concept actually shows the generality of the concept of measure spaces, which we might discuss afterwards.

To introduce formal notion of product measure space, we need a definition based framework to work in, which we learn now:

Definition 6.0.1. (**Premeasure**) Consider an **algebra**[22] $\mathcal{A}$ over a set X. The map

$$\mu_0 : \mathcal{A} \longrightarrow [0, +\infty]$$

is called a premeasure if it satisfies:

1. $\mu_0(\Phi) = 0$, and
2. For $A_1, A_2, \ldots$ a sequence of disjoint sets from $\mathcal{A}$,

$$\mu_0\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu_0(A_i)$$

Definition 6.0.2. (**Outer measure by Premeasure**) Consider an algebra $\mathcal{A}$ defined on set X. Suppose $\mu_0 : \mathcal{A} \to [0, +\infty]$ is a premeasure on it. We then define μ_* as the following:

$$\mu_* : \mathcal{A} \longrightarrow [0, +\infty]$$

defined by, for $A \subseteq X$:

$$\mu_*(A) = \inf\left\{\sum_{n=1}^{\infty} \mu_0(E_n) \mid A \subseteq \bigcup_{n=1}^{\infty} E_n \text{ where } \{E_n\} \text{ is a sequence in } \mathcal{A}\right\}$$

Proposition 6.0.3. *For an algebra $\mathcal{A}$ on X, μ_* satisfies the following:*

1. *μ_* is an Outer measure.*
2. *The collection of μ_* measurable sets, $\mathcal{M}_{\mu_*}$, is a σ-algebra.*
3. *The σ-algebra generated by algebra $\mathcal{A}$, $\mathcal{B}$, is a proper subset of $\mathcal{M}_{\mu_*}$. That is,*

$$\mathcal{B} \subsetneq \mathcal{M}_{\mu_*}$$

Proof. Clearly, $\mu_*(\Phi) = 0$ as, the ∞ sequence of Φ, $\{A_n\}$ where $A_i = \Phi\ \forall\, i$, is such that

$$\Phi \subseteq \bigcup_i A_i$$

and

$$\sum_i \mu_0(A_i) = 0.$$

The next parts has proof similar to one done for Lebesgue Outer measure. ∎

[22] Note that this just an algebra, not a σ-algebra.

6.1 Some set theoretic concepts

We would need this concepts for later discussions.

Definition 6.1.1. (**Elementary class/family**) A collection of sets denoted by $\mathcal{E}$ is called an elementary class if:

1. $\Phi \in \mathcal{E}$,
2. For any $E, F \in \mathcal{E}$, then
$$E \cap F \in \mathcal{E}$$
3. If $E \in \mathcal{E}$, then
$$\exists \left\{F_n\right\}_{n=1}^{N} \text{ where } F_n\text{'s are disjoint and in } \mathcal{E} \text{ such that } E^c = \bigcup_{n=1}^{N} F_n$$

Proposition 6.1.2. *If $\mathcal{E}$ is an elementary class, then the collection $\mathcal{A}$ defined as:*

$$\textit{For any } A \in \mathcal{A},\ \exists \left\{E_n\right\}_{n=1}^{N} \textit{ where } E_n\textit{'s are disjoint and in } \mathcal{E} \textit{ such that } A = \bigcup_{n=1}^{N} E_n$$

is an Algebra.

Definition 6.1.3. (**Monotone class**) A collection of subsets of a set X denoted $\mathcal{C} \subseteq \mathcal{P}(X)$ is called a monotone class if:

1. For if $\{E_n\}$ is a sequence of monotonically increasing sets from $\mathcal{C}$, that is,
$$E_1 \subseteq E_2 \subseteq \dots$$
then,
$$\bigcup_{n=1}^{\infty} \in \mathcal{C}.$$
2. For if $\{E_n\}$ is a sequence of monotonically decreasing sets from $\mathcal{C}$, that is,
$$E_1 \supseteq E_2 \supseteq \dots$$
then,
$$\bigcap_{n=1}^{\infty} \in \mathcal{C}.$$

Proposition 6.1.4. *Consider a family of monotone class given as $\{\mathcal{C}_n\}$. Then,*

$$\bigcap_{n} \mathcal{C}_n \textit{ is a monotone class.}$$

Proof. Take any sequence of sets $\{I_n\}$ from $\bigcap_{n=1}^{\infty} \mathcal{C}_n$ such that they are monotonically increasing,

$$I_1 \subseteq I_2 \subseteq \dots.$$

Now, because each $I_i \in \mathcal{C}_n \forall\, n$ and it is a monotonically increasing sequence, therefore $\bigcup_{i=1}^{\infty} I_i \in \mathcal{C}_n \;\forall\, n$. Hence,

$$\bigcup_{i=1}^{\infty} I_i \in \bigcap_n \mathcal{C}_n.$$

Similarly, suppose $\{J_n\}$ is a monotonically decreasing sequence of sets from $\bigcap_n \mathcal{C}_n$,

$$J_1 \supseteq J_2 \supseteq \dots$$

Hence, each $J_i \in \mathcal{C}_n \;\forall\, n$. Since each $\mathcal{C}_n$ is a monotone class, therefore, $\bigcap_{i=1}^{\infty} J_i \in \mathcal{C}_n \;\forall\, n$. Hence,

$$\bigcap_{i=1}^{\infty} J_i \in \bigcap_n \mathcal{C}_n.$$

Hence proved. □

Definition 6.1.5. (**Generated monotone class**) Consider any $\mathcal{S} \subset \mathcal{P}(X)$. Then, $\mathcal{C}(\mathcal{S})$ is called the monotone class generated by $\mathcal{S}$ if $\mathcal{C}(\mathcal{S})$ is the smallest monotone class containing $\mathcal{S}$.

Proposition 6.1.6. *Let $\mathcal{A}$ be an Algebra. Suppose*

- *$\mathcal{C}(\mathcal{A})$ is the Monotone Class generated by $\mathcal{A}$, and*
- *$\mathcal{M}$ is the σ-Algebra generated by $\mathcal{A}$.*

Then,

$$\mathcal{M} = \mathcal{C}(\mathcal{A}).$$

6.2 Product measure space

Definition 6.2.1. (**Measurable rectangle**) Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure spaces. Suppose $X \times Y$ is the Cartesian Product of the sets X and Y. Then, $A \times B \subseteq X \times Y$ is called a measurable Rectangle if

$$A \in \mathcal{A} \text{ and } B \in \mathcal{B}$$

Definition 6.2.2. (**Elementary rectangles**) Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are measure spaces. Denote by $\mathcal{K}$ the collection of all measurable Rectangles. Then, we define Elementary Rectangles, $\mathcal{E}$, as the collection :

For any $A \in \mathcal{E}$, $\exists\, \{E_n\}_{n=1}^{N}$ where E_n's are disjoint measurable rectangles in $\mathcal{K}$ such that $A = \bigcup_{n=1}^{N} E_n$.

Remark 6.2.3. ★ It is important to note that **elementary rectangles** $\mathcal{E}$ **is an algebra**, due to Proposition 6.1.2.

Definition 6.2.4. (**Product of measurable spaces**) Denote $\mathcal{A} \times \mathcal{B}$ to be the σ**-Algebra generated by** $\mathcal{E}$. Then,

$$(X \times Y, \mathcal{A} \times \mathcal{B})$$

is the product of measurable Spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$.

Definition 6.2.5. (**Product measure space**) Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure Spaces. The product of these two measure Spaces is defined as the following triple:

$$(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$$

where

1. $X \times Y$ is the Cartesian Product of X and Y.
2. $\mathcal{A} \times \mathcal{B}$ is the σ-Algebra generated by Elementary Rectangles of the product $X \times Y$ under respective measure spaces.
3. $\mu \times \nu$ is defined as:
$$(\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B)$$
where $A \times B$ is a measurable Rectangle.

Remark 6.2.6. ★ Note the following:

- $\mu \times \nu$ defines a premeasure on Elementary Rectangles, $\mathcal{E}$, which is an Algebra.
- With the premeasure $\mu \times \nu$ on $\mathcal{E}$, we then construct the outer measure $\mu_* \times \nu_*$ by premeasure as done in Definition 6.0.2.
- As Proposition 6.0.3 shows, the collection of $\mu_* \times \nu_*$ measurable sets from $\mathcal{E}$ forms a σ-Algebra, that is, the σ-Algebra generated from all Elementary Rectangles. This is exactly what we did now.

6.2.1 Properties of product measure space

Definition 6.2.7. (**x & y sections**) Suppose $E \subseteq X \times Y$. Then we define

1. x-section as all y available in E if x is fixed:
$$E_x = \{y \in Y \mid (x, y) \in E\}$$
2. y-section as all x available in E if y is fixed:
$$E^y = \{x \in X \mid (x, y) \in E\}$$

Definition 6.2.8. (**x & y sections of a function**) Suppose f is a function on $X \times Y$. Then,

1. x-Section of f given $x \in X$ is just $f_x(y) = f(x, y)$.
2. y-Section of f given $y \in Y$ is just $f^y(x) = f(x, y)$

Proposition 6.2.9. *Suppose $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are two measurable spaces and $E \subseteq \mathcal{A} \times \mathcal{B}$. Then,*

1. *$E_x \in \mathcal{B}\ \forall x \in X$, and*
2. *$E^y \in \mathcal{A}\ \forall y \in Y$.*

That is, each section of a subset of product of measurable spaces, $\mathcal{A} \times \mathcal{B}$, is itself measurable.

Proof. Omitted. □

Proposition 6.2.10. *Suppose $f : \mathcal{A} \times \mathcal{B} \longrightarrow \mathbb{R}$ is a $\mathcal{A} \times \mathcal{B}$-measurable function. Then,*

1. *$f_x : \mathcal{B} \longrightarrow \mathbb{R}$ is a $\mathcal{B}$-measurable function $\forall x \in X$.*
2. *$f^y : \mathcal{A} \longrightarrow \mathbb{R}$ is a $\mathcal{A}$-measurable function $\forall y \in Y$.*

That is, each section of a measurable Function on product measurable space is itself a measurable function.

Proof. Trivial, same as Proposition 6.2.9. □

6.3 The Fubini-Tonelli theorem

This is perhaps the most important result of this course, whose proof can be found in any course book, available on webpage.

Theorem 6.3.1. ***(Tonelli's theorem)*** *Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two σ-finite*[23] *measure spaces. Consider an $\mathcal{A} \times \mathcal{B}$-measurable function*

$$f : X \times Y \longrightarrow [0, +\infty].$$

Then,

1. *The function:*
   - *$g : X \to [0, +\infty]$ given by:*
   $$g(x) = \int_Y f_x d\nu$$
   is $\mathcal{A}$-measurable.
   - *$h : Y \to [0, +\infty]$ given by:*
   $$h(y) = \int_X f_y d\mu$$
   is $\mathcal{B}$-measurable.
2. *f satisfies:*

$$\begin{aligned}\int_{X\times Y} f d(\mu \times \nu) &= \int_X \left(\int_Y f_x d\nu \right) d\mu \\ &= \int_Y \left(\int_X f_y d\mu \right) d\nu\end{aligned}$$

Theorem 6.3.2. ***(Fubini's theorem)*** *Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measurable spaces. Consider an $\mathcal{A} \times \mathcal{B}$-measurable function which is also $\mu \times \nu$-Integrable given as:*

$$f : X \times Y \longrightarrow [-\infty, +\infty]^{24}.$$

Then,

1. *We have that*
   - *f_x is ν-Integrable* ***almost everywhere*** *on Y.*
   - *f_y is μ-Integrable* ***almost everywhere*** *on X.*
2. *The following relation holds:*

$$\begin{aligned}\int_{X\times Y} f d(\mu \times \nu) &= \int_X \left(\int_Y f_x d\nu \right) d\mu \\ &= \int_Y \left(\int_X f_y d\mu \right) d\nu\end{aligned}$$

[23] This means that there are finite $\{A_n\}$ sets in $\mathcal{A}$ with finite measure such that $\bigcup_n A_n = X$. Similarly for $(Y, \mathcal{B}, \nu)$.
[24] **Note the target set here!**

6.4 Applications-III : Product and Fubini-Tonelli

Lemma 6.4.1. *Let (X, Σ_1, μ) and (Y, Σ_2, ν) be two σ-finite measure space with $f \in \mathcal{L}^1(\mu)$ and $g \in \mathcal{L}^1(\nu)$. Then the function $h(x,y) = f(x)g(y)$ is in $\mathcal{L}^1(\mu \times \nu)$ and that*

$$\int_{X\times Y} h d\mu \times \nu = \left(\int_X f d\mu\right)\left(\int_Y g d\nu\right). \tag{1.1}$$

Proof. We first show that h is measurable. Indeed, as $X \times Y \to \mathbb{C}$ given by $(x,y) \mapsto f(x)$ and $X \times Y \to \mathbb{C}$ given by $(x,y) \mapsto g(y)$ are measurable as they are composites $X \times Y \xrightarrow{\pi_1} X \xrightarrow{f} \mathbb{C}$ and $X \times Y \xrightarrow{\pi_2} X \xrightarrow{g} \mathbb{C}$ respectively, where we know that the projection π_i are measurable, therefore their pointwise product $h(x,y) = f(x)g(y)$ is measurable as well. This shows that h is measurable.

Now note that we have $\int_X |f|\, d\mu = M < \infty$ and $\int_Y |g|\, d\nu = N < \infty$. Furthermore, we have $|h|_x = (|f|\,|g|)_x = |f(x)|\,|g|$ and similarly $|h|^y = (|f|\,|g|)^y = |g(y)|\,|f|$. Consequently by Fubini-Tonelli for $L^+(\mu \times \nu)$, we obtain

$$\begin{aligned}
\int_{X\times Y} |h|\, d\mu \times \nu &= \int_X \int_Y |h|\, d\nu d\mu \\
&= \int_X \int_Y |f|\,|g|\, d\nu d\mu \\
&= \int_X |f| \left(\int_Y |g|\, d\nu\right) d\mu \\
&= \int_X N\,|f|\, d\mu \\
&= NM < \infty.
\end{aligned}$$

Hence, $h \in \mathcal{L}^1(\mu \times \nu)$.

We now wish to show Eq. (1.1). Indeed, as $h \in \mathcal{L}^1(\mu \times \nu)$, therefore by Fubini-Tonelli for $\mathcal{L}^1(\mu \times \nu)$, we obtain

$$\begin{aligned}
\int_{X\times Y} h d\mu \times \nu &= \int_X \int_Y h_x d\nu d\mu \\
&= \int_X \int_Y f(x) g d\nu d\mu \\
&= \int_X f(x) \left(\int_Y g d\nu\right) d\mu \\
&= \left(\int_X f d\mu\right)\left(\int_Y g d\nu\right)
\end{aligned}$$

as needed. □

Example 6.4.2. For $X = Y = \mathbb{N}$, $\Sigma_1 = \Sigma_2 = \mathcal{P}(\mathbb{N})$ and $\mu = \nu = \#$ the counting measure, we wish to restate the Fubini-Tonelli theorem in this setting.

First of all, we observe that both the spaces (X, Σ_1, μ) and (Y, Σ_2, ν) are σ-finite as $\mathbb{N}$ can be covered by $\{E_n\}$ where $E_n = \{n\}$ is a finite measure subset. Hence the Fubini-Tonelli applies.

For any measurable $h : X \to \mathbb{C}$, we first claim that the integral $\int_X h d\mu = \sum_n h(n)$. Indeed, we first have by definition

$$\int_X h d\mu = \int_X \Re(h)^+ d\mu - \int_X \Re(h)^- d\mu + i\left(\int_X \Im(h)^+ - \int_X \Im(h)^- d\mu\right)$$

where each $\Re(h)^{\pm}, \Im(h)^{\pm}$ are measurable functions $X \to [0, \infty)$. Hence we reduce to assuming h is a non-negative measurable function. In this case, we observe the following. Consider $g_n = \sum_{k=1}^n h(k)\chi_{\{k\}}$. Observe that g_n are increasing and converges to f pointwise. Then by MCT, we have

$$\begin{aligned}\int_X h d\mu &= \lim_{n\to\infty} \int_X g_n d\mu \\ &= \lim_{n\to\infty} \int_X \sum_{k=1}^n h(k)\chi_{\{k\}} d\mu \\ &= \lim_{n\to\infty} \sum_{k=1}^n \int_X h(k)\chi_{\{k\}} d\mu \\ &= \lim_{n\to\infty} \sum_{k=1}^n h(k) \\ &= \sum_{k=1}^\infty h(k)\end{aligned}$$

as needed.

Now pick any $h \in L^+(\mu\times\nu)$. We first claim that $\int_{X\times Y} hd\mu\times\nu = \sum_{n,m} h(n,m)$. Indeed, we claim that $\int_{X\times Y} hd\mu \times \nu = \sup\{\int_{X\times Y} \varphi d\mu \times \nu \mid 0 \le \varphi \le h,\ \varphi \text{ is simple}\} = \sup\{\sum_{(n,m)\in F} h(n,m) \mid F \subseteq \mathbb{N}\times\mathbb{N} \text{ is finite}\} = \sum_{n,m} h(n,m)$, as needed. Let $A = \{\int_{X\times Y} \varphi d\mu \times \nu \mid 0 \le \varphi \le h,\ \varphi \text{ is simple}\}$ and $B = \{\sum_{(n,m)\in F} h(n,m) \mid F \subseteq \mathbb{N}\times\mathbb{N} \text{ is finite}\}$. To show the above claim, we need only show that

$$\sup A = \sup B.$$

First suppose that B is not bounded. Then there exists a sequence $b_k \in B$ such that $b_k \to \infty$ as $k \to \infty$. Let $b_k = \sum_{(n,m)\in F_k} h(n,m) \to \infty$ as $k \to \infty$, where F_k are finite sets. Hence, construct $\varphi_k = \sum_{(n,m)\in F_k} h(n,m)\chi_{\{(n,m)\}}$. Clearly, $\varphi_k \in A$ is a simple function below h. As $\int_{X\times Y} \varphi_k d\mu \times \nu = \sum_{(n,m)\in F_k} h(n,m) = b_k$, therefore we get that A is unbounded as well.

Now suppose B is bounded. Then, A is bounded as well because for any simple function $0 \le \varphi \le h$, φ cannot be supported on an infinite cardinality set as otherwise B will be unbounded. Hence both $\sup A$ and $\sup B$ exists and we wish to show that they are equal. Note that the above argument shows that for any simple function $0 \le \varphi \le h$ given by $\varphi = \sum_{k=1}^n a_k \chi_{E_k}$, the integral $\int_{X\times Y} \varphi d\mu \times \nu = \sum_{k=1}^n a_k \#(E_k)$ is finite. Hence for any $\varphi \in A$, there exists a finite set F such that $\int_{X\times Y} \varphi d\mu \times \nu \le \sum_{(n,m)\in F} h(n,m)$. Thus, $\sup A \le \sup B$. Conversely, pick any $\sum_{(n,m)\in F} h(n,m) \in B$ for some finite F. Then, the simple function $\varphi = \sum_{(n,m)\in F} h(n,m)\chi_{\{(n,m)\}} \in A$ is such that $\int_{X\times Y} \varphi d\mu \times \nu = \sum_{(n,m)\in F} h(n,m)$. Hence $\sup B \le \sup A$. This completes the proof that integral of h over $X \times Y$ is just the double sum.

Now by Fubini-Tonelli for L^+, we obtain that

$$\begin{aligned}\int_{X\times Y} hd(\mu\times\nu) &= \sum_{n,m} h(n,m)\\ &= \int_X\int_Y h_n d\nu d\mu\\ &= \int_X \sum_m h(n,m)d\mu\\ \text{(by MCT)} &= \sum_m \int_X h(n,m)d\mu\\ &= \sum_m\sum_n h(n,m).\end{aligned}$$

Similarly, we also yield by an application of MCT that

$$\begin{aligned}\int_{X\times Y} hd(\mu\times\nu) &= \sum_{n,m} h(n,m)\\ &= \int_Y\int_X h^m d\mu d\nu\\ &= \sum_n\sum_m h(n,m).\end{aligned}$$

Now suppose $h \in \mathcal{L}^1(\mu\times\nu)$. Then by Fubini-Tonelli, we yield that

$$\begin{aligned}\int_{X\times Y} hd\mu\times\nu &= \sum_{n,m} h(n,m)\\ &= \int_X\int_Y h_n d\nu d\mu\\ &= \int_X \sum_m h(n,m)d\mu\\ \text{(by DCT as each } h^m\in\mathcal{L}^1(\mu) \text{ by Fubini)} &= \sum_m \int_X h(n,m)d\mu\\ &= \sum_m\sum_n h(n,m).\end{aligned}$$

Similarly, we yield

$$\int_{X\times Y} hd\mu\times\nu = \sum_n\sum_m h(n,m).$$

Hence, we yield the following two statements from this discussion:

1. *Let $\sum_{n,m} a_{n,m}$ be a double series of non-negative real numbers. Then,*

$$\sum_{n,m} a_{n,m} = \sum_n\sum_m a_{n,m} = \sum_m\sum_n a_{n,m}.$$

2. *Let $\sum_{n,m} a_{n,m}$ be a double series of complex numbers such that*

$$\sum_{n,m} |a_{n,m}| < \infty.$$

Then,

$$\sum_{n,m} a_{n,m} = \sum_n \sum_m a_{n,m} = \sum_m \sum_n a_{n,m}.$$

This completes the analysis.

Example 6.4.3. Let $c \in \mathbb{R}$ and define $f : [0,\infty) \to \mathbb{R}$ a map given by

$$f(x) = \frac{\sin x^2}{x} + \frac{cx}{1+x}.$$

Let $a > 0$. Then we wish to show that

$$\lim_{n\to\infty} \int_0^a f(nx)dx = ac.$$

We claim that $\frac{\sin x}{x}$ is a bounded function over $[0,\infty)$. Indeed, fix $\epsilon > 0$. As $\lim_{x\to 0} \frac{\sin x^2}{x} = 0$, therefore there exists $\delta > 0$ such that for $x \in (0,\delta)$, we have $\left|\frac{\sin x^2}{x}\right| < \epsilon$. Furthermore, for $x \geq \delta$ we have $\left|\frac{\sin x^2}{x}\right| \leq \frac{1}{|x|} \leq \frac{1}{\delta}$. Hence taking $M = \max\{\epsilon, 1/\delta\}$, we see that $\left|\frac{\sin x^2}{x}\right| \leq M$ over $[0,\infty)$. Consequently, over $[0,\infty)$, we have

$$\begin{aligned} |f(x)| &= \left|\frac{\sin x^2}{x} + \frac{cx}{1+x}\right| \\ &\leq |M| + \left|\frac{cx}{1+x}\right| \\ &\leq M + |c|. \end{aligned}$$

Thus, the sequence of measurable functions $|f(nx)|$ is upper bounded by $|g(x)| = M + |c|$ over $[0,a]$, which is $\mathcal{L}^1$ over $[0,a]$. Furthermore, we see that $f(nx) \to c$ over $(0,a]$ pointwise as $n \to \infty$. Hence, by DCT, we obtain

$$\begin{aligned} \lim_{n\to\infty} \int_0^a f(nx)dx &= \int_0^a \lim_{n\to\infty} f(nx)dx \\ &= \int_0^a cdx \\ &= ca \end{aligned}$$

as needed.

Example 6.4.4. Let $X = Y = [0,1]$, $\Sigma_1 = \Sigma_2 = \mathcal{B}_{[0,1]}$ the Borel σ-algebra on $[0,1]$ and $\mu =$ Lebesgue measure over $[0,1]$ and $\nu =$ counting measure over $[0,1]$. We wish to show that Fubini-Tonelli doesn't holds here for the function $\chi_D : X \times Y \to \mathbb{R}$ where $D = \{(x,x) \mid x \in X\}$.

Let us first calculate $\int_{X\times Y} \chi_D d\mu \times \nu$. As χ_D is just a characteristic function, therefore we simply have

$$\int_{X\times Y} \chi_D d\mu \times \nu = \mu \times \nu(D).$$

1. We claim that $\mu \times \nu(D) = \infty$. Indeed, by definition, we have

$$\mu \times \nu(D) = \inf \left\{ \sum_n \mu(I_n)\nu(J_n) \mid \bigcup_n I_n \times J_n \supseteq D, I_n \times J_n \in \mathcal{R} \right\}$$

where $\mathcal{R}$ is the elementary family of all rectangles. We claim that for any such cover $D \subseteq \bigcup_n I_n \times J_n$, we have $\sum_n \mu(I_n)\nu(J_n) = \infty$. Indeed, it suffices to show that there is an $n \in \mathbb{N}$ such that $\mu(I_n) \neq 0$ and J_n is infinite. Suppose there is no such n. It then follows that if $\mu(I_n) \neq 0$, then J_n is finite. Further, if $\mu(I_n) = 0$, then J_n can be finite or infinite. Let

$$K := \{n \in \mathbb{N} \mid \mu(I_n) \neq 0\}$$

and

$$L := \{n \in \mathbb{N} \mid \mu(I_n) = 0\}.$$

Consequently, $K \cup L = \mathbb{N}$.

Pick $n \in K$. Then, $\mu(I_n) \neq 0$ and J_n is finite. It follows that $(I_n \times J_n) \cap D$ is atmost a finite set. Thus, $\bigcup_{n \in K} I_n \times J_n$ covers atmost a countable subset of D. Hence, it follows that $\bigcup_{n \in L} I_n \times J_n$ covers an uncountable subset of D. Furthermore,

$$\begin{aligned} V &:= D \setminus \left(\bigcup_{n \in L} (I_n \times J_n) \cap D \right) \\ &= \bigcup_{n \in K} (I_n \times J_n) \cap D \ \text{ is countable.} \end{aligned} \tag{4.1}$$

For any $n \in \mathbb{N}$, observe that

$$(I_n \times J_n) \cap D = \{(x, x) \in D \mid x \in I_n \cap J_n\}. \tag{4.2}$$

From the preceding remark, it is thus clear that the set $\bigcup_{n \in L} (I_n \times J_n) \cap D = \{(x, x) \in D \mid x \in I_n \cap J_n \text{ for some } n \in L\}$ is uncountable, which further makes $A := \bigcup_{n \in L} I_n \cap J_n \subseteq [0, 1]$ uncountable. We claim that $[0, 1] \setminus A$ is countable. Indeed, by (4.1), we first see that

$$\begin{aligned} V &= \{(x, x) \mid x \in I_n \cap J_n \text{ for some } n \in K\} \\ &\cong \bigcup_{n \in K} I_n \cap J_n. \end{aligned}$$

Thus, $\bigcup_{n \in K} I_n \cap J_n$ is countable.

Observe that

$$[0, 1] = \left(\bigcup_{n \in K} I_n \cap J_n \right) \cup \left(\bigcup_{n \in L} I_n \cap J_n \right)$$

because $\{I_n \times J_n\}_{n \in \mathbb{N}}$ covers D. Consequently, as A is uncountable, therefore

$$[0, 1] \setminus A \subseteq \bigcup_{n \in K} I_n \cap J_n$$

is countable by Eq. (4.3), as required.

As $A \subseteq [0,1]$ is such that $[0,1] \setminus A$ is countable therefore $\mu(A) = 1$. But, $A \subseteq \bigcup_{n\in L} I_n$, therefore $1 = \mu(A) = \sum_{n\in L} m(I_n) = \sum_n 0 = 0$ as I_n for $n \in L$ is of measure 0. Hence we have $1 = \mu(A) \le 0$, a contradiction. This shows that $\sum_n \mu(I_n)\nu(J_n) = \infty$ for each $\{I_n \times J_n\} \subseteq \mathcal{R}$ such that $\bigcup_n I_n \times J_n \supseteq D$. Thus,

$$\mu \times \nu(D) = \infty.$$

2. We claim that $\int_Y \int_X \chi_D d\mu d\nu = 0$. Indeed, we have

$$\begin{aligned}\int_Y \int_X (\chi_D)^y \, d\mu d\nu &= \int_Y \int_X \chi_{D^y} d\mu d\nu \\ &= \int_Y \mu(\{(y,y)\}) d\nu \\ &= \int_Y 0 d\nu \\ &= 0,\end{aligned}$$

as required.

3. We claim that $\int_X \int_Y \chi_D d\nu d\mu = 1$. Indeed, we have

$$\begin{aligned}\int_X \int_Y (\chi_D)_x d\nu d\mu &= \int_X \int_Y \chi_{D_x} d\nu d\mu \\ &= \int_X \nu(\{(x,x)\}) d\mu \\ &= \int_X 1 d\mu \\ &= \mu(X) \\ &= 1,\end{aligned}$$

as needed.

Hence, we have shown that for Fubini-Tonelli to work, we require both spaces to be σ-finite (which is not the case here as Y is not σ-finite).

Example 6.4.5. We wish to construct an example of a monotone class of subsets of a non-empty set X such that it is not a σ-algebra. Indeed, consider $X = \{1., 2, 3\}$. Define $\mathcal{C} := \{\emptyset, \{1\}, X\}$. Then $\mathcal{C}$ is a monotone class as the only non-trivial increasing sequence of sets is $\emptyset \subseteq \{1\}$ and their union is clearly $\{1\}$ which is in $\mathcal{C}$. Furthermore the only non-trivial decreasing sequence is $X \supseteq \{1\}$, whose intersection is $\{1\}$, which is in $\mathcal{C}$. However, $\mathcal{C}$ is not a σ-algebra as $\{1\}^c = \{2,3\} \notin \mathcal{C}$.

Lemma 6.4.6. *Let (X, Σ, μ) be a measure space and $f : X \to \mathbb{C}$ be an $\mathcal{L}^1(\mu)$ map. For each $E \in \Sigma$, define*

$$\nu(E) = \int_E f d\mu.$$

1. *If $\mu(E) = 0$, then $\nu(E) = 0$.*

2. If $\{E_n\} \subseteq \Sigma$ is a disjoint collection, then

$$\nu\left(\coprod_n E_n\right) = \sum_n \nu(E_n).$$

3. For all $\epsilon > 0$, there exists $\delta > 0$ such that

$$\mu(E) < \delta \implies |\nu(E)| < \epsilon.$$

Proof. 1. Note that $\nu(E) = 0$ iff $|\nu(E)| = 0$. Consequently, we see that

$$\begin{aligned} |\nu(E)| = \left|\int_E f d\mu\right| &\leq \int_E |f|\, d\mu \\ &\leq \infty \cdot \int_E d\mu \\ &= \infty\mu(E) \\ &= \infty \cdot 0 = 0, \end{aligned}$$

as needed.

2. Pick $\{E_n\} \subseteq \Sigma$ to be a disjoint collection. Consider the sequence of measurable functions $g_n = f\chi_{\coprod_{k=1}^n E_k}$. Observe that $g_n \to f\chi_{\coprod_{k=1}^\infty E_k}$ pointwise as $n \to \infty$. Furthermore, observe that $|g_n| \leq |f|$ and as $f \in \mathcal{L}^1(\mu)$, therefore we may apply DCT on $\{g_n\}$.

Applying DCT, we yield

$$\begin{aligned} \int_{\coprod_{k=1}^\infty E_k} f d\mu = \int_X f\chi_{\coprod_{k=1}^\infty E_k} d\mu &= \lim_{n\to\infty} \int_X f\chi_{\coprod_{k=1}^n E_k} d\mu \\ &= \lim_{n\to\infty} \int_{\coprod_{k=1}^n E_k} f d\mu \\ &= \lim_{n\to\infty} \sum_{k=1}^n \int_{E_k} f d\mu \\ &= \sum_{k=1}^\infty \int_{E_k} f d\mu \\ &= \sum_{k=1}^\infty \nu(E_k), \end{aligned}$$

as needed.

3. As $f \in \mathcal{L}^1(\mu)$, therefore there exists a sequence of bounded functions $g_n \in \mathcal{L}^1(\mu)$ such that $g_n \to f$ pointwise as $n \to \infty$ and $|g_n| \leq |f|$ over X. Fix $E \in \Sigma$ of finite measure. It follows from DCT applied on g_n over X that

$$\lim_{n\to\infty} \int_E |f - g_n|\, d\mu \leq \lim_{n\to\infty} \int_X |f - g_n|\, d\mu = 0.$$

Fix $\epsilon > 0$. The convergence of above limit yields that there exists $N \in \mathbb{N}$ such that

$$\int_E |f - g_n| \, d\mu < \epsilon/2$$

for all $n \geq N$. Thus, in particular,

$$\int_E |f| - |g_N| \, d\mu \leq \int_E |f - g_N| \, d\mu < \epsilon/2.$$

Now, from above, we yield that

$$\begin{aligned} |\nu(E)| = \left| \int_E f d\mu \right| &\leq \int_E |f| \, d\mu \\ &\leq \epsilon/2 + \int_E |g_N| \, d\mu. \end{aligned}$$

As $|g_n|$ is bounded, therefore let $|g_N| \leq M_n$ for some $M_N \in [0, \infty)$. Consequently,

$$\begin{aligned} |\nu(E)| &\leq \epsilon/2 + \int_E |g_N| \, d\mu \\ &\leq \epsilon/2 + \int_E M_N d\mu \\ &\leq \epsilon/2 + M_N \mu(E). \end{aligned}$$

Hence, letting $\delta = \epsilon/2M_N$, we yield that for any $E \in \Sigma$ such that $\mu(E) < \delta$ we have

$$\begin{aligned} |\nu(E)| &< \epsilon/2 + \epsilon/2 \\ &= \epsilon. \end{aligned}$$

This completes the proof. □

Example 6.4.7. Let $X = Y = \mathbb{N}$, $\Sigma_1 = \Sigma_2 = \mathcal{P}(\mathbb{N})$ and $\mu = \nu =$ counting measure. Further, define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ given by

$$f(m,n) = \begin{cases} 1 & \text{if } m = n, \\ -1 & \text{if } m = n+1, \\ 0 & \text{otherwise.} \end{cases}$$

We wish to show that
1. $\int_{X \times Y} |f| \, d(\mu \times \nu) = \infty$,
2. $\int_X \int_Y f d\nu d\mu = 1$,
3. $\int_Y \int_X f d\mu d\nu = 0$.

Before proving, we would first like to show that f is indeed measurable. Indeed, we may write $f = \chi_D - \chi_S$ where $D = \{(m,m) \mid m \in \mathbb{N}\}$ is the diagonal and $S = \{(n+1,n) \mid n \in \mathbb{N}\}$. Both are measurable subsets of $\Sigma_1 \otimes \Sigma_2$ as $D = \bigcup_m \{(m,m)\}$ and $S = \bigcup_n \{(n+1,n)\}$. Note that singletons of $X \times Y$ are measurable as singletons in X and Y are measurable.

1. Observe that $|f| = \chi_D + \chi_S = \chi_{D\amalg S}$ where D and S are two disjoint subsets as defined above. Consequently

$$\begin{aligned}\int_{X\times Y} |f|\, d\mu\times\nu &= \mu\times\nu(D\amalg S)\\ &= \mu\times\nu(D)+\mu\times\nu(S).\end{aligned}$$

We claim that both $\mu\times\nu(D)$ and $\mu\times\nu(S)$ are ∞.

Indeed, for any $\{I_n\times J_n\}$ for $I_n\times J_n\in\mathcal{R}$ rectangles such that $\bigcup_n I_n\times J_n\supseteq D$, we see that if $(I_n\times J_n)\cap D\neq\emptyset$, then $\mu(I_n)\nu(J_n)\geq 1$ as in this case $I_n\cap J_n\neq\emptyset$. As D is an infinite set and $\bigcup_n(I_n\times J_n)\cap D=D$, therefore $\sum_n\mu(I_n)\nu(J_n)\geq\mu\times\nu\left(\bigcup_n I_n\times J_n\right)\geq\mu\times\nu(D)=\infty$. This shows $\mu\times\nu(D)=\infty$.

Similarly, if $\{I_n\times J_n\}$ for $I_n\times J_n\in\mathcal{R}$ is a collection of rectangles such that $\bigcup_n I_n\times J_n\supseteq S$, then for each n for which $(I_n\times J_n)\cap D\neq\emptyset$ we deduce that $\mu(I_n)\nu(J_n)\geq 1$. Hence, as above, we again get that $\sum_n\mu(I_n)\nu(J_n)=\infty$. This proves that $\int_{X\times Y}|f|\,d\mu\times\nu=\infty$.

2. We simply observe that by definition we have $D_m=\{m\}$ and $S_m=\{m-1\}$. Consequently,

$$\begin{aligned}\int_X\int_Y f_m d\nu d\mu &= \int_X\int_Y \chi_{D_m}-\chi_{S_m} d\nu d\mu\\ &= \int_X \nu(D_m)-\nu(S_m)d\mu\\ &= \int_{X\setminus\{1\}}(1-1)d\mu+\int_{\{1\}}(1-0)d\mu\\ &= 1.\end{aligned}$$

3. We simply observe that by definition $D^n=\{n\}$ and $S^n=\{n+1\}$. Consequently,

$$\begin{aligned}\int_Y\int_X f^n d\mu d\nu &= \int_Y\int_X(\chi_{D^n}-\chi_{S^n})d\mu d\nu\\ &= \int_Y \mu(D^n)-\mu(S^n)d\nu\\ &= \int_Y 1-1d\nu\\ &= 0.\end{aligned}$$

This completes the proof.

7 Differentiation

We now study some of the interconnections between integration and differentiation and related notions.

7.1 Differentiability

Definition 7.1.1. (**Upper/Lower left & Upper/Lower right Derivatives**) Suppose $f:\mathbb{R}\longrightarrow[-\infty,+\infty]$ is a function such that for all $x\in\mathbb{R}$, f is defined on some open interval around x, then we define the following quantities:

- ***Upper Right Derivative*** :

$$D^+f(x) = \limsup_{h\to 0^+} \frac{f(x+h)-f(x)}{h}$$

- ***Lower Right Derivative*** :

$$D_+f(x) = \liminf_{h\to 0^+} \frac{f(x+h)-f(x)}{h}$$

- ***Upper Left Derivative*** :

$$D^-f(x) = \limsup_{h\to 0^-} \frac{f(x+h)-f(x)}{h}$$

- ***Lower Left Derivative*** :

$$D_-f(x) = \liminf_{h\to 0^-} \frac{f(x+h)-f(x)}{h}$$

Definition 7.1.2. (**Differentiable function**) The function f is said to be differentiable at x if and only if:

$$D^+f(x) = D_+f(x) = D^-f(x) = D_-f(x).$$

Hence, a function is said to be differentiable if it is differentiable at all points of it's domain.

7.2 Functions of bounded variation

We now study those functions which do not change too erratically over an interval. We already have the notion of differentiability for the same, so we would see connections between such type of functions and there differential character.

Definition 7.2.1. (**Variations of a Function**) Suppose we are given a function on an interval

$$f : [a,b] \longrightarrow \mathbb{R}$$

and any partition $\mathcal{P}_{[a,b]} = \{a = x_0, x_1, x_2, \ldots, x_{k-1}, x_k = b\}$ where $x_i < x_{i+1}$. Now, define the following the following three quantities:

$$\begin{aligned} p_{\mathcal{P}} &= \sum_{i=1}^{k} (f(x_i) - f(x_{i-1}))^+ \\ n_{\mathcal{P}} &= \sum_{i=1}^{k} (f(x_i) - f(x_{i-1}))^- \\ t_{\mathcal{P}} &= p_{\mathcal{P}} + n_{\mathcal{P}} = \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| \end{aligned}$$

where $_{\mathcal{P}}$ denotes the partition over which the sum is defined and it's simple to observe that $p_{\mathcal{P}} - n_{\mathcal{P}} = f(b) - f(a)$. Also, $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$.
Then, we finally define the following three quantities:

- ***Positive Variation of*** f :
$$P_f = \sup_{\mathcal{P}} p_{\mathcal{P}}.$$
- ***Negative Variation of*** f :
$$N_f = \sup_{\mathcal{P}} n_{\mathcal{P}}.$$
- ***Total Variation of*** f :
$$T_f = \sup_{\mathcal{P}} t_{\mathcal{P}}.$$

Definition 7.2.2. (**Function of bounded variation**) Suppose $f : \mathbb{R} \to \mathbb{R}$ is a given function. Then f is said to be of bounded variation over interval $[a, b]$ if

$$T_f[a, b] = T_f < \infty.$$

The class of functions on a given interval $[a, b]$ which are of bounded variation is denoted by:

$$\mathcal{BV}\left([a, b]\right).$$

So that for any $f \in \mathcal{BV}\left([a, b]\right)$, $T_f < \infty$.

Remark 7.2.3. A function f is said to belong to $\mathcal{BV}\left((-\infty, \infty)\right)$ if f belongs to each $\mathcal{BV}\left([a, b]\right)$ for each interval $[a, b]$. Clearly, in this case $T_f(-\infty, \infty) = \sup_{[a,b]} T_f[a, b]$.

Proposition 7.2.4. *Suppose* $f \in \mathcal{BV}\left([a, b]\right)$. *Then,*
1. $f(b) - f(a) = P_f - N_f$.
2. $T_f = P_f + N_f$.

Proof. Take any $f \in \mathcal{BV}\left([a, b]\right)$. Then we have $T_f[a, b] < \infty$. Now, we know that for any partition $\mathcal{P}$ of $[a, b]$, $f(b) - f(a) = p_{\mathcal{P}} - n_{\mathcal{P}}$. Now, take supremum over all partitions of $[a, b]$, both sides of the above, to write:

$$\begin{aligned} \sup_{\mathcal{P}} \left(f(b) - f(a)\right) &= \sup_{\mathcal{P}} \left(p_{\mathcal{P}} - n_{\mathcal{P}}\right) \\ f(b) - f(a) &= \sup_{\mathcal{P}} p_{\mathcal{P}} - \sup_{\mathcal{P}} n_{\mathcal{P}} \qquad \text{Known result : } \sup_n \left(x_n - y_n\right) = \sup_n x_n - \sup_n y_n. \\ &= P_f - N_f \end{aligned}$$

For the 2^{nd} part, we have

$$\begin{aligned} T_f &= \sup_{\mathcal{P}} t_{\mathcal{P}} \\ &= \sup_{\mathcal{P}} (p_{\mathcal{P}} + n_{\mathcal{P}}) \\ &= \sup_{\mathcal{P}} p_{\mathcal{P}} + n_{\mathcal{P}} \qquad \text{Known result : } \sup_n \left(x_n + y_n\right) = \sup_n x_n + \sup_n y_n. \\ &= P_f + N_f \end{aligned}$$

Hence proved. ☐

The following theorem is important as it characterizes the functions in $\mathcal{BV}\left([a, b]\right)$.

Proposition 7.2.5. *The following result holds:*

$f \in \mathcal{BV}([a,b]) \iff \exists\, g,\ h$ *which are* ***monotonically increasing*** *and* ***finite*** *on* $[a,b]$*, such that* $f = g - h$*.*

Proof. $\mathbf{L} \implies \mathbf{R}$: Consider the functions

$$\begin{aligned} g(x) &= P_f[a,x] + f(a) \\ h(x) &= N_f[a,x] \end{aligned}$$

For any $a \le x_0 \le x_1 \le b$, we observe that:

$$\begin{aligned} g(x_0) &= P_f[a,x_0] + f(a) \le P_f[a,x_1] + f(a) = g(x_1) \\ h(x_0) &= N_f[a,x_0] \le N_f[a,x_1] = h(x_1) \end{aligned}$$

because we are adding another partitioning point. Hence, g, h are monotonically increasing functions on $[a,b]$. Hence, $g(b) = P_f[a,b] + f(a) < \infty$ as f is of bounded variation, so g is finite. Similarly h is finite on $[a,b]$. Finally, we note that:

$$\begin{aligned} g(x) - h(x) &= P_f[a,x] + f(a) - N_f[a,x] \\ &= P_f[a,x] - N_f[a,x] + f(a) \\ &= f(x) - f(a) + f(a) \\ &= f(x) \end{aligned}$$

Hence proved that if $f \in \mathcal{BV}([a,b])$, then there exists two monotonically increasing, finite functions on $[a,b]$ such that f is their difference.

$\mathbf{R} \implies \mathbf{L}$: Take any partition $\mathcal{P}[a,b] = a = x_0 < x_1 < x_2 < \cdots < x_k = b$. Now we see that

$$\begin{aligned} t_{\mathcal{P}}^f &= \sum_{i=1}^k |f(x_i) - f(x_{i-1})| \\ &= \sum_{i=1}^k |g(x_i) - h(x_i) - g(x_{i-1}) + h(x_{i-1})| \\ &\le \sum_{i=1}^k |g(x_i) - g(x_{i-1})| + \sum_{i=1}^k |h(x_i) - h(x_{i-1})| \\ &= t_{\mathcal{P}}^g + t_{\mathcal{P}}^h \\ &< \infty \end{aligned}$$

Hence $T_f = \sup_{\mathcal{P}} t_{\mathcal{P}}^f < \infty$. So $f \in \mathcal{BV}([a,b])$. □

7.3 Differentiability of monotone functions & Lebesgue's theorem

Definition 7.3.1. (**Vitali covering**) A collection $\mathcal{C}$ of closed, bounded, nondegenerate[25] intervals is said to cover a given set E in the sense of Vitali if:

$$\text{For any } x \in E \text{ and any } \epsilon > 0,\ \exists\, I \in \mathcal{C} \text{ such that } x \in I \ \&\ \lambda(I) < \epsilon.$$

[25] An interval $[a,b]$ is said to be nondegenerate if $a < b$.

Theorem 7.3.2. *(**The Vitali covering lemma**) Suppose $E \subset \mathbb{R}$ is of finite outer measure, that is $\lambda^*(E) < \infty$. Also consider a collection $\mathcal{C}$ of closed, bounded intervals which covers E in the sense of Vitali. Then,*

$$\forall\, \epsilon > 0,\ \exists \text{ disjoint \& finite subcollection } \{I_k\}_{k=1}^n \text{ of } \mathcal{C} \text{ such that}$$

$$\lambda^*\left(E \setminus \bigcup_{k=1}^n I_k\right) < \epsilon.$$

The following is a generalization of mean value theorem from Calculus.

Proposition 7.3.3. *Let f be an increasing function on a closed, bounded interval $[a,b]$. Then, $\forall\, \alpha > 0$, we have*

$$\lambda^*\left(\{x \in (a,b) \mid D^+f(x) = D^-f(x) \geq \alpha\}\right) \leq \frac{1}{\alpha} \cdot (f(b) - f(a))$$

Proof. We first begin be denoting $E_\alpha = \{x \in (a,b) \mid D^+f(x) = D^-f(x) \geq \alpha\}$. Now, construct the following collection $\mathcal{F}$ of closed and bounded intervals $[c,d]$ for which,

$$f(d) - f(c) \geq \alpha'(d-c)$$

where $0 < \alpha' \leq \alpha$. Now take any $x \in E_\alpha$. We hence see that $D^+f(x) \geq \alpha$. Now for any $\epsilon > 0$, we can construct a closed bounded interval $I = \left[x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}\right]$ for which $\lambda^*(I) = \epsilon$ with $x \in I$. But moreover, we have that

$$\begin{aligned} Df(x) = \frac{f(x+\epsilon/2) - f(x-\epsilon/2)}{\epsilon} &\geq \alpha \\ f(x+\epsilon/2) - f(x-\epsilon/2) &\geq \epsilon\alpha \geq \epsilon\alpha' \end{aligned}$$

Hence the interval $\left[x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}\right] \in \mathcal{F}$. Therefore, $\mathcal{F}$ ***covers*** E_α ***in the sense of Vitali***(!)

Now, by Vitali Covering Lemma (Theorem 7.3.2), we get that

$$\forall\, \epsilon > 0,\ \exists \text{ finite disjoint } \{I_k\}_{k=1}^n \text{ from } \mathcal{F} \text{ such that } \lambda^*\left(E_\alpha \setminus \bigcup_{k=1}^n I_k\right) < \epsilon$$

Now, observe that

$$E_\alpha \subseteq \bigcup_{k=1}^n I_k \cup \left(E_\alpha \setminus \bigcup_{k=1}^n I_k\right)$$

Hence, by finite sub-additivity of outer measures, we get the following:

$$\begin{aligned} \lambda^*(E_\alpha) &\leq \lambda^*\left(E_\alpha \setminus \bigcup_{k=1}^n\right) + \lambda^*\left(\bigcup_{k=1}^n I_k\right) \\ &< \epsilon + \sum_{k=1}^n \lambda^*(I_k) \\ &\leq \epsilon + \sum_{k=1}^n \frac{f(d_k) - f(c_k)}{\alpha'} && \text{Suppose } I_k = [c_k, d_k]. \end{aligned}$$

Now since f is an increasing function and $I_k \subset [a,b]\ \forall\ k$, therefore we have:

$$\sum_{k=1}^{n} f(d_k) - f(c_k) \leq f(b) - f(a)$$

That is,

$$\lambda^*\left(E_\alpha\right) < \epsilon + \frac{1}{\alpha'} \cdot (f(b) - f(a))$$

But since $\epsilon > 0$ and $\alpha' \in (0, \alpha]$ are arbitrary, therefore,

$$\lambda^*\left(E_\alpha\right) \leq \frac{1}{\alpha} \cdot (f(b) - f(a))$$

Hence proved. □

7.3.1 Lebesgue's differentiation theorem

This is also one of the most important theorems of this course. This theorem portrays that monotonicity of a function is much better attribute of *niceness* of a function than the usual belief of continuity, because we know example of continuous functions which is not differentiable, that is the **Weierstrass function**. But with this theorem, if we are given a monotone function on an open interval, then it ought to be differentiable almost everywhere on that interval. The same is obviously not true for just continuous functions.

Theorem 7.3.4. *(Lebesgue's Differentiation Theorem) Suppose f is a* ***monotone*** *function on open interval (a,b) to $\mathbb{R}$. Then,*

$$f \textit{ is differentiable on } (a,b) \textit{ almost everywhere (!)}$$

Corollary 7.3.5. *A function f of bounded variation over an interval $[a,b]$ is differentiable almost everywhere in (a,b).*

Proof. Lebesgue's Differentiation Theorem (7.3.4) and the fact that any function of bounded variation is a difference of two increasing functions (Proposition 7.2.5). □

7.4 Integration & differentiation in context

We now learn some relationships between differentiation and integration. But let us begin with the following basic proposition.

Proposition 7.4.1. *Let $f : X \to [0, +\infty)$ be a measurable function which is Lebesgue Integrable on a set $E \subseteq X$. Then,*

$$\forall\ \epsilon > 0,\ \exists\ \delta > 0 \ \textit{ such that } \forall\ A \subset E \ \textit{ with } \ \lambda(A) < \delta\ ,\ \int_A f < \epsilon.$$

Proof. Consider the following sequence $\{f_n\}$ of functions:

$$f_n(x) = \begin{cases} f(x) & \text{if } f(x) \leq n \\ n & \text{if } f(x) \geq n \end{cases}$$

Now, we see that $f_n(x) \leq f_{n+1}(x)$ because if $f_n(x) = f(x)$ then $f_n(x) = f(x) \leq n < n+1$ so that $f_{n+1}(x) = f(x)$. Hence $\{f_n\}$ is an increasing sequence. Therefore

$$f_n \longrightarrow f \text{ almost everywhere}$$

Hence, by Monotone Convergence Theorem (4.2.1), we get that

$$\int \varprojlim_n f_n = \varprojlim_n \int f_n.$$

Now, observe the following:

$$\begin{aligned} \int_E f - \varprojlim_n \int f_n &= 0 \\ \varprojlim_n \int_E f - \varprojlim_n \int f_n &= 0 \\ \varprojlim_n \int_E (f - f_n) &= 0 \end{aligned}$$

where we see that

$$(f - f_n)(x) = \begin{cases} 0 & \text{if } f(x) - n \leq 0 \\ f(x) - n & \text{if } f(x) - n \geq 0. \end{cases}$$

From this, we can construct the following sequence of sets:

$$E_n = \{x \in E \mid f(x) - n \geq 0\}.$$

Again, we see that for any $x \in E_n$, we would have $f(x) \geq n > n - 1$, so that $x \in E_{n-1}$. Hence $\{E_n\}$ is a decreasing sequence of subsets of E.
We now observe that

$$\int_{E_n} f \geq \int_{E_n} n = n\lambda(E_n).$$

Hence, we get that, for any $n \in \mathbb{N}$,

$$\lambda(E_n) \leq \frac{1}{n} \int_{E_n} f$$

So that, we can choose n corresponding to any $\delta = \epsilon/n$ such that

$$\lambda(E_n) \leq \frac{1}{n} \int_{E_n} f < \delta = \epsilon/n$$

and

$$\int_{E_n} f < n\delta = \epsilon.$$

Hence proved. □

7.4.1 Indefinite integral

The Indefinite integral of a Lebesgue Integrable function forms a sort of *bridge* between Integration and Differentiation.

Definition 7.4.2. (**Indefinite integral**) Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function. We then define the indefinite integral of f as the following function:

$$F(x) = \int_a^x f(t)dt$$

Proposition 7.4.3. *Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function. Then,*

1. *$F(x)$ is a* ***continuous*** *function on $[a, b]$.*
2. *$F(x)$ is of* ***bounded variation*** *on $[a, b]$.*

Proof. **1.** Take any $\epsilon > 0$. By Proposition 7.4.1, we get that

$$F(x) = \int_a^x f(t)dt < \epsilon \implies \exists\, \delta > 0 \; \& \; \exists A \subset [a, x] \text{ such that } \lambda^* (A) < \delta.$$

In more precise words, $\forall \epsilon > 0$, $\exists \delta > 0$ such that whenever

$$|a - x_0| = \lambda^* ([a, x_0]) < \delta$$

then we would have

$$\begin{aligned} |F(x_0) - F(a)| &< \epsilon \\ \left| \int_a^{x_0} f(t)dt - \int_a^a f(t)dt \right| &< \epsilon \\ \left| \int_a^{x_0} f(t)dt \right| &< \epsilon \end{aligned}$$

which is just the definition of continuity.

2. Take any partition of $[a, b]$, say, $\mathcal{P}([a, b]) = a = x_0 < x_1 < x_2 < \ldots x_{k-1} < x_k = b$. Now, we see that,

$$\begin{aligned} t_{\mathcal{P}} &= \sum_{i=1}^{k} |F(x_i) - F(x_{i-1})| \\ &= \sum_{i=1}^{k} \left| \int_a^{x_i} f(t)dt - \int_a^{x_{i-1}} f(t)dt \right| \\ &= \sum_{i=1}^{k} \left| \int_{x_{i-1}}^{x_i} f(t)dt \right| \\ &< \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} |f(t)|\, dt \\ &< \infty \end{aligned}$$

where last line follows because f is Lebesgue Integrable. Since our choice of partition $\mathcal{P}$ was arbitrary, therefore $\sup_{\mathcal{P}} t_{\mathcal{P}} < \infty$ hence $F(x)$ is of bounded variation. $\square$

Corollary 7.4.4. *The Indefinite integral of a Lebesgue Integrable function is Differentiable almost everywhere.*

Proof. By-product of Lebesgue's Differentiation Theorem (7.3.4), or more succinctly, Corollary 7.3.5. □

Proposition 7.4.5. *Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function and*

$$F(x) = \int_a^{x_0} f(t)dt = 0 \;\forall\; x \in (a, b).$$

Then, $f = 0$ almost everywhere on (a, b).

Proof. Suppose to the contrary that $\exists\; E \subset [a, b]$ such that $f(x) \neq 0 \;\forall\; x \in E$ with $\lambda^*(E) > 0$. By Proposition 2.9.2, we get that $\exists\; G \subset E$ which is closed such that $\lambda^*(G) > 0$ and $\lambda^*(E \setminus G) = 0$. Hence $(a, b) \setminus G$ is open. Now consider the integral:

$$\int_G f = \int_{(a,b)} f - \int_{(a,b)\setminus G} f$$

We know that $\int_{(a,b)} f = 0$ as $F(x) = 0 \;\forall\; x \in (a, b)$. In a similar tone, we have $f|_{(a,b)\setminus G} = 0$ almost everywhere because $\lambda^*(E \setminus G) = 0$ and $f = 0$ on $(a, b) \setminus E$ anyways. Therefore, we have $\int_G f = 0$. But $f|_G \neq 0$ by definition of $G \subset E$. Hence we have a contradiction. Therefore such a set E cannot exist. Thus $f = 0$ almost everywhere on (a, b) if $F = 0 \;\forall\; x \in (a, b)$. □

Theorem 7.4.6. *Let $[a, b]$ be a finite interval and let $f : [a, b] \longrightarrow \mathbb{R}$ be a Lebesgue Integrable function over it. Then,*

$$F' = f \text{ almost everywhere in } [a, b].$$

Proof. Omitted □

7.4.2 Absolutely continuous functions

This is a more general form of continuity, and since it has connections with indefinite integral, we then learn them here.

Definition 7.4.7. (**Absolutely Continuous Function**) A function $f : [a, b] \longrightarrow \mathbb{R}$ is said to be Absolutely Continuous if

$\forall\; \epsilon > 0\,,\;\; \exists\; \delta > 0$ such that $\forall$ **finite** & **disjoint** collection of open intervals $\{(a_k, b_k)\}_{k=1}^{n}$ each subset of (a, b) whi

$$\sum_{k=1}^{n} (b_k - a_k) < \delta,$$

also satisfies

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \epsilon.$$

Remark 7.4.8. Some straightforward results are:

- Any Absolutely Continuous function is Continuous in usual sense. This follows trivially from their definitions.
- Any Absolutely Continuous function is Uniformly Continuous. This also follows from the definition.

Proposition 7.4.9. *Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is an Absolutely Continuous function. Then, f is of Bounded Variation over $[a, b]$.*

Proof. Since f is an absolutely continuous function, therefore, for any fixed $\epsilon > 0$, we can construct the partition of $[a, b]$, say $\mathcal{P}' = a = x_0 < x_1 < \cdots < x_k = b$ such that each (x_{i-1}, x_i) is of length $< \delta$. Clearly, we would then have that

$$\sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| < k\epsilon.$$

Now, consider any arbitrary partition say $\mathcal{P} \equiv a = y_0 < y_1 < \cdots < y_N = b$ of $[a, b]$. Collect the partition points of $\mathcal{P}$ as the open disjoint intervals $\{(y_{i-1}, y_i)\}_{i=1}^{N}$. Then, for each i^{th} interval in this partition, we can further partition it into k_i open disjoint intervals such that each has length $< \delta$. In particular, we would have the following partition of $[y_{i-1}, y_i]$:

$$\left\{(z_{j-1}^i, z_j^i)\right\}_{j=1}^{k_i} \text{ where } z_0^i = y_{i-1}\ ,\ \ z_{k_i}^i = y_i.$$

Now, note that the variation of f over the $[y_{i-1}, y_i]$ would then be:

$$\begin{aligned} |f(y_i) - f(y_{i-1})| &= \left|\sum_{j=1}^{k_i} f(z_j^i) - f(z_{j-1}^i)\right| \\ &\leq \sum_{j=1}^{k_i} \left|f(z_j^i) - f(z_{j-1}^i)\right| \\ &< \sum_{j=1}^{k_i} \epsilon = k_i\epsilon & \because z_j^i - z_{j-1}^i < \delta \text{ by construction} \end{aligned}$$

Now, the variation over whole of $\mathcal{P}$ would then be:

$$\begin{aligned} t_{\mathcal{P}} &= \sum_{i=1}^{N} |f(y_i) - f(y_{i-1})| \\ &< \sum_{i=1}^{N} k_i\epsilon \\ &< \infty \end{aligned}$$

as k_i is finite for all i. Hence proved. □

This theorem relates Indefinite integral of a Lebesgue integral and Absolute Continuity.

Theorem 7.4.10. *Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function and it's Indefinite integral is denoted by the function $F(x)$. Then,*

$$F \textit{ is an Indefinite integral} \iff F \textit{ is Absolutlely Continuous.}$$

Proof. Omitted. □

8 Signed measures and derivatives

The concept of signed measures is the next generalization that we seek to understand. So far we have only encountered measures on a space which maps subsets to $[0, \infty]$. But what happens when we *increase* the co-domain to the whole $[-\infty, \infty]$? First of all, we can see clearly that a measure shall not map some subsets to $+\infty$ and some other subset to $-\infty$ as we would then have the problem of $\infty - \infty$, and since we are not doing set theory here, hence we would refrain ourselves only to such *signed* measures which either maps to $(-\infty, \infty]$ or $[-\infty, \infty)$ but not both.

Later we would see that having such a notion of *signed* measure actually leads to some very striking results!

Definition 8.0.1. (**Signed measure**) Suppose $(X, \mathcal{M})$ is a measurable Space. A function

$$\nu : \mathcal{M} \longrightarrow [-\infty, +\infty) \;\; \textbf{OR} \; \nu : \mathcal{M} \longrightarrow (-\infty, +\infty]$$

is called a Signed measure if it satisfies:

1. ν atmost maps sets either to $+\infty$ or $-\infty$, but not both[26].
2. ν maps null-set to 0:
$$\nu(\Phi) = 0$$
3. ν follows countable additivity:
$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i)$$
where $\{A_i\}$ is any sequence of disjoint sets from $\mathcal{M}$.

Definition 8.0.2. (**Positive set**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then a set $A \in \mathcal{M}$ is said to be a positive set w.r.t. ν if:

$$\forall\, S \subseteq A \text{ such that } S \in \mathcal{M}\,, \;\; \nu(S) \geq 0.$$

Definition 8.0.3. (**Negative set**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then a set $B \in \mathcal{M}$ is said to be a negative set w.r.t. ν if:

$$\forall\, S \subseteq B \text{ such that } S \in \mathcal{M}\,, \;\; \nu(S) \leq 0.$$

Remark 8.0.4. One could alternatively say that a set is a negative set if it is positive w.r.t. $-\nu$.

Definition 8.0.5. (**Null set**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then a set $N \in \mathcal{M}$ is said to be a null set w.r.t. ν if

$$N \text{ is both a Positive and Negative set w.r.t. } \nu$$

Proposition 8.0.6. *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Let $\{A_i\}$ be a sequence of positive sets w.r.t. ν. Then,*

$$A = \bigcup_i A_i \textit{ is a Positive Set w.r.t. } \nu.$$

[26]Hence the two possible choices for the ν above.

Proof. We know that we can write the sequence $\{A_i\}$ in the following form:

$$\{B_i\} \text{ where } B_i\text{'s are disjoint } \& \ B_i \subseteq A_i.$$

This can be easily be seen by $B_1 = A_1$ and $B_i = A_i \setminus B_{i-1}$. Hence $\{B_i\}$ is a sequence of disjoint positive sets. Moreover, we can see that

$$A = \bigcup_i A_i = \bigcup_i B_i.$$

Now, take any subset $E \subseteq A$, which we can simply write as:

$$\begin{aligned} \nu(E) &= \nu(E \cap A) \\ &= \nu\left(\bigcup_i E \cap B_i\right) \\ &= \sum_i \nu(E \cap B_i) \\ &> 0 \end{aligned}$$

because $E \cap A_i \subseteq B_i$ and B_i is a positive set. Hence proved. □

Remark 8.0.7. This is clearly also true for negative sets and null sets. That is, countable union of negative (null) sets is also a negative (null) set.

Proposition 8.0.8. *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. If $E \in \mathcal{M}$ is such that $\nu(E) \geq 0$, then*

$$\exists\, A \subseteq E \text{ such that } A \text{ is a positive Set w.r.t. } \nu,\ A \in \mathcal{M} \ \& \ \nu(A) > 0.$$

Proof. Written in Diary at 26^{th} September, 2018. Typeset it here when time allows. □

8.1 The Hahn decomposition theorem

Theorem 8.1.1. ***(Hahn decomposition theorem)*** *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then,*

$$\exists \text{ positive Set } A \in \mathcal{M} \text{ and negative Set } B \in \mathcal{M} \text{ such that } \boxed{A \cup B = X \ \& \ A \cap B = \Phi}$$

Moreover, any two such pairs (A, B) and (A', B') are unique upto the fact that

$$A \Delta A' \ \& \ B \Delta B' \text{ are } \nu\text{-Null Sets}$$

8.2 The Jordan decomposition of a signed measure

We now, in a sense, generalize the Hahn Decomposition Theorem (8.1.1), but to the signed measure ν itself. As usual, let's first familiarize ourselves with some definitions.

Definition 8.2.1. (**Mutual singularity of signed measures**) Let ν_1 and ν_2 be two measures (NOT signed!) over measurable space $(X, \mathcal{M})$. Then ν_1 and ν_2 are called mutually singular if

$$\exists\, A \in \mathcal{M} \text{ such that } \nu_1(A) = \nu_2(A^c) = 0$$

and is then denoted by:

$$\nu_1 \perp \nu_2.$$

Theorem 8.2.2. ***(Jordan decomposition theorem)*** *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then,*

$$\exists \text{ measures } \nu^+ \,\&\, \nu^- \text{ on } (X, \mathcal{M}) \text{ such that } \boxed{\nu = \nu^+ - \nu^- \,\&\, \nu^+ \perp \nu^-}$$

and such a decomposition of ν is unique.

8.3 The Radon-Nikodym theorem

This is one of the final and most important theorems of this course. As we will see, this theorem gives us a notion of the derivative of a signed measure. However, we would not go more deeper into that fact.
As usual, we first introduce some definitions.

Definition 8.3.1. (**Total variation of a signed measure**) The total variation of a signed measure ν over some measurable space is defined by

$$|\nu| = \nu^+ + \nu^-$$

where $\nu = \nu^+ - \nu^-$ is the Jordan Decomposition (Theorem 8.2.2) of ν.

Remark 8.3.2. Since ν^+ and ν^+ are the usual measures on the measurable space, therefore $|\nu|$ is also a usual measure on the same measurable space.

Definition 8.3.3. (**σ-finite signed measure**) Suppose ν is a signed measure on measurable space $(X, \mathcal{M})$. Then ν is called σ-Finite if

$$\exists\, \{X_n\}_{n=1}^{\infty} \text{ where } X_i \in \mathcal{M} \text{ and } |\nu(X_i)| < \infty \text{ such that } \bigcup_{n=1}^{\infty} X_n = X$$

Remark 8.3.4. ν is σ-Finite $\iff$ $|\nu|$ is σ-Finite.

Definition 8.3.5. (**Absolute continuity of usual measures**) Suppose λ and γ are usual measures over a measurable space $(X, \mathcal{M})$. If,

$$\lambda(E) = 0 \text{ for some } E \in \mathcal{M} \implies \gamma(E) = 0$$

always, then γ is said to be absolutely continuous w.r.t. λ. This is denoted by $\gamma \ll \lambda$.

Definition 8.3.6. (**Absolute continuity of signed measures**) Suppose μ and ν are signed measures over a measurable space $(X, \mathcal{M})$. If,

$$|\mu|(E) = 0 \text{ for some } E \in \mathcal{M} \implies \nu(E) = 0$$

always, then ν is called absolutely continuous w.r.t. μ. This is denoted by $\nu \ll \mu$.

Theorem 8.3.7. *(**Radon-Nikodym theorem**) Suppose $(X, \mathcal{M})$ is a measurable space and λ & γ are two σ-finite measures on it such that $\gamma \ll \lambda$. Then,*

$$\exists \text{ measurable Function w.r.t. } \lambda\ f : X \longrightarrow [0, \infty) \text{ such that}$$
$$\boxed{\gamma(E) = \int_E f d\lambda \ \forall\ E \in \mathcal{M}}$$

Moreover, f is unique up-to almost everywhere equality, that is, if $\gamma(E) = \int_E g d\lambda \ \forall\ E \in \mathcal{M}$, then,

$$f = g \text{ almost everywhere on } X \text{ w.r.t. } \lambda.$$

8.4 Applications-IV : Signed spaces

Lemma 8.4.1. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two signed measures on it. Then $\nu \ll \mu$ and $\mu \perp \nu$ if and only if $\nu = 0$.*

Proof. ($\Rightarrow$) As $\mu \perp \nu$, therefore there exists a μ-null set A and a ν-null set B such that $A \amalg B = X$. For any measurable set $E \subseteq X$, we have $E = (E \cap A) \amalg (E \cap B)$. As $E \cap A \subseteq A$, therefore $\mu(E \cap A) = 0$. As $\nu \ll \mu$, therefore $\nu(E \cap A) = 0$. Furthermore, since $E \cap B \subseteq B$, therefore $\nu(E \cap B) = 0$. Hence,

$$\begin{aligned}\nu(E) &= \nu(E \cap A) + \nu(E \cap B) \\ &= 0,\end{aligned}$$

as needed.

($\Leftarrow$) As for any measurable set $E \subseteq X$, we have $\nu(E) = 0$, hence $\nu \ll \mu$. Further, as X is now ν-null and $\emptyset$ is μ-null, therefore $X = X \amalg \emptyset$ gives us the required decomposition to claim that $\mu \perp \nu$. □

Lemma 8.4.2. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two positive measures on it. The following are equivalent.*

1. *$\nu \perp \mu$,*
2. *there exists a sequence $\{E_n\} \subseteq \mathcal{A}$ such that $\mu(E_n) \to 0$ and $\nu(X \setminus E_n) \to 0$ as $n \to \infty$.*

Proof. (1. $\Rightarrow$ 2.) As $\nu \perp \mu$, therefore there exists a ν-null set A and a μ-null set B such that $X = A \amalg B$. Hence, we may take $E_n = A$ and $X \setminus E_n = B$ for each $n \in \mathbb{N}$. This provides the required sequence.

(2. $\Rightarrow$ 1.) We wish to construct $A, B \subseteq X$ such that $A \amalg B = X$ and A is ν-null and B is μ-null. To construct A and B, we proceed as follows.

We first observe that since $\mu(E_n) \to 0$, therefore there exists a subsequence of $\mu(E_n)$ say $\mu(E_{n_k})$ such that $\sum_k \mu(E_{n_k}) < \infty$. Indeed, this is a consequence of a general result : for any positive sequence $\{a_n\}$ such that $\lim_n a_n = 0$, we have that there exists a subsequence $\{a_{n_k}\}$ such that $\sum_k a_{n_k} < \infty$. Indeed, for each $k \in \mathbb{N}$ there exists an $n_k \in \mathbb{N}$ such that $a_n \leq 1/2^k$ for all $n \geq n_k$. Consequently, we see that $\sum_{k=1}^{\infty} a_{n_k} \leq \sum_{k=1}^{\infty} 1/2^k < \infty$, as required.

We apply the above result to $\{\mu(E_n)\}$ to obtain a subsequence $\{E_{n_k}\}$. We now replace $\{E_n\}$

by $\{E_{n_k}\}$ so that we may assume $\sum_n \mu(E_n) < \infty$.

Consider the sequence

$$\begin{aligned} F_n &= X \setminus \bigcup_{k=n}^{\infty} E_k \\ &= \bigcap_{k=n}^{\infty} X \setminus E_k. \end{aligned}$$

Observe that F_n is an increasing sequence and that each $F_n \subseteq X \setminus E_n$. Hence,

$$\nu(F_n) \leq \nu(X \setminus E_n). \tag{2.1}$$

Moreover, observe that since $\lim_{n\to\infty}\nu(X \setminus E_n) = 0$, therefore $\lim_{n\to\infty}\nu(F_n) = 0$. Hence, we deduce by monotone property of measures that

$$\lim_{n\to\infty}\nu(F_n) = \nu\left(\bigcup_n F_n\right).$$

Hence, by previous discussion, we further deduce that

$$\lim_{n\to\infty}\nu(F_n) = 0 = \nu\left(\bigcup_n F_n\right).$$

Thus $A := \bigcup_n F_n$ is a ν-null set. It now suffices to show that $X \setminus A$ is a μ-null set.

Observe that $X \setminus A$ can be written as

$$\begin{aligned} X \setminus A &= \bigcap_{n=1}^{\infty} X \setminus F_n \\ &= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k. \end{aligned}$$

We claim that $X \setminus A$ is a μ-null set. Indeed, denote

$$S_n = \bigcup_{k=n}^{\infty} E_k.$$

We wish to show that

$$\mu\left(\bigcap_{n=1}^{\infty} S_n\right) = \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k\right) = 0.$$

Observe that S_n is a decreasing sequence. Furthermore, as $\mu(S_n) \leq \sum_{k=n}^{\infty} \mu(E_k) < \infty$, therefore we may apply the monotone property of measures. Consequently, we yield the following

$$\lim_{n\to\infty}\mu(S_n) = \mu\left(\bigcap_{n=1}^{\infty} S_n\right). \tag{2.2}$$

We now show that $\lim_{n\to\infty}\mu(S_n) = 0$. Indeed, denoting $l = \sum_k \mu(E_k)$ and $l_n = \sum_{k=1}^{n-1} \mu(E_k)$, we first see that $\lim_{n\to\infty} l_n = l$. Now observe that

$$\mu(S_n) \leq \sum_{k=n}^{\infty} \mu(E_k) = l - l_{n-1}$$

where the last equality follows from rearrangement of a positive convergent series. Hence, taking $\lim_{n\to\infty}$, we obtain that

$$\lim_{n\to\infty}\mu(S_n) \leq l - l = 0,$$

that is, $\lim_{n\to\infty}\mu(S_n) = 0$. By Eq. (2.2), we deduce that $X \setminus A = \bigcap_n S_n$ is a μ-null set, as needed. This completes the proof. $\square$

Example 8.4.3. We wish to find Lebesgue decomposition of $\nu = m + \delta_0$ where m is the Lebesgue measure and δ_0 is the Dirac delta measure at $0 \in \mathbb{R}$.

Indeed, as $(\mathbb{R}, \mathcal{M})$ is a σ-algebra, therefore the Lebesgue decomposition theorem holds. We see an immediate candidate for Lebesgue decomposition of ν with respect to m as follows:

$$\nu = \nu_a + \nu_s$$

where we set $\nu_a = m$ and $\nu_s = \delta_0$. Indeed, this works as $m \ll m$ holds trivially and $\delta_0 \perp m$ because of the decomposition $\mathbb{R} = \{0\} \amalg (\mathbb{R}^2 \setminus \{0\})$ where we see immediately that $\{0\}$ is m-null and $\mathbb{R}^2 \setminus \{0\}$ is δ_0-null.

Example 8.4.4. Let $p(x) = x^2 - 6x + 1$ be a function $\mathbb{R} \to \mathbb{R}$. Consider the signed measure

$$\nu(E) = \int_E p dm$$

on $(\mathbb{R}, \mathcal{M})$.

1. We first wish to show that $(\mathbb{R}, \mathcal{M}, \nu)$ is σ-finite. Indeed, let $X_n = [n, n+1]$. We claim that $-\infty < \nu(X_n) < \infty$ for each $n \in \mathbb{N}$. Now, observe that over X_n, the polynomial is a continuous function supported on a compact interval, hence it achieves a maxima and a minima, say M_n and m_n respectively. Consequently, we have $m_n \leq p \leq M_n$ over X_n.

$$\int_{X_n} m_n dm \leq \int_{X_n} p dm \leq \int_{X_n} M_n dm$$

and thus $-\infty < m_n \leq \nu(X_n) \leq M_n < \infty$ for each n. Hence ν is σ-finite.

2. We wish to find the Hahn-decomposition of $\mathbb{R}$ w.r.t. ν. That is, we wish to find a decomposition $\mathbb{R} = P \amalg N$ such that P is a ν-positive set and N is a ν-negative set.

Observe that $p(x)$ has two real roots $c_1, c_2 \in \mathbb{R}$. Consequently, we see that over $N = [c_1, c_2]$ the polynomial $p(x)$ is negative and hence $\nu(E) = \int_E p dm \leq 0$ for any measurable $E \subseteq N$. Thus N is a negative set. Similarly, define $P = (-\infty, c_1) \cup (c_2, \infty)$. Then observe that $p(x)$ is positive over p, thus $\nu(E) \geq 0$ for any measurable $E \subseteq P$.

3. We now wish to find the Jordan decomposition of ν. Indeed, define $\nu^+(E) := \nu(E \cap P)$ and

$\nu^-(E) := -\nu(E \cap N)$ where $X = P \amalg N$ is the Hahn decomposition. These are positive measures such that $\nu = \nu^+ - \nu^-$. Furthermore, $\nu^+ \perp \nu^-$ as P is ν^--null and N is ν^+-null by construction.

4. We wish to find the Lebesgue decomposition of ν with respect to the Lebesgue measure m. Indeed, we claim that $\nu \ll m$, which will immediately show that the Lebesgue decomposition of ν with respect to m is simply $\nu = \nu + 0$ where $\nu \ll m$ and $0 \perp m$. Indeed, take any measurable set $E \subseteq X$ such that $m(E) = 0$. As p is measurable therefore

$$\nu(E) = \int_E p dm = 0.$$

Hence $\nu \ll m$, completing the proof.

Lemma 8.4.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space, $\{E_n\}_{n=1}^N \subseteq \mathcal{A}$ and $\{c_n\}_{n=1}^N \subseteq \mathbb{R}_{\geq 0}$. Consider the positive measure*

$$\nu(E) = \sum_{n=1}^N c_n \mu(E \cap E_n)$$

for some fixed $E_n \in \mathcal{A}$. Then,

1. *$\nu \ll \mu$,*
2. *$d\nu/d\mu = \sum_{n=1}^N c_n \chi_{E_n}$.*

Proof. 1. We wish to show that $\nu \ll \mu$. Indeed, pick any $E \in \mathcal{A}$ such that $\mu(E) = 0$. As μ is positive, consequently $\mu(E \cap E_n) = 0$ for each $n = 1, \ldots, N$ as $E \cap E_n \subseteq E$. Hence, we deduce that $\nu(E) = 0$. Thus $\nu \ll \mu$.

2. We now wish to find the Radon-Nikodym derivative $d\nu/d\mu$, which exists as $\nu \ll \mu$. Indeed, this means we need to find a measurable function $f : X \to [0, \infty]$ such that

$$\nu(E) = \int_E f d\mu$$

for each $E \in \mathcal{A}$. We claim that the following simple function

$$f = \sum_{n=1}^N c_n \chi_{E_n}$$

is the required derivative. Indeed, observe that

$$\begin{aligned}
\int_E f d\mu &= \int_E \sum_{n=1}^N c_n \chi_{E_n} d\mu \\
&= \sum_{n=1}^N c_n \int_E \chi_{E_n} d\mu \\
&= \sum_{n=1}^N c_n \int_X \chi_{E_n \cap E} d\mu \\
&= \sum_{n=1}^N c_n \mu(E \cap E_n) \\
&= \nu(E),
\end{aligned}$$

as required. $\square$

Lemma 8.4.6. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two positive measures. Suppose $\nu \ll \mu$. Then,*

1. *if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$,*
2. *Assuming both μ and ν are σ-finite, if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$ and*

$$\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1} \text{ μ-almost everywhere.}$$

Proof. 1. We first wish to show that if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$.

Denote $f = d\nu/d\mu$. Suppose $E \in \mathcal{A}$ is such that $\nu(E) = 0$. Thus $\nu(E) = \int_E f d\mu = 0$. As $f > 0$ μ-almost everywhere, therefore consider the sequence $E_n = \{f(x) \geq 1/n\}$. Clearly, $\bigcup_n E_n = X \setminus N$ as $f > 0$ over X, where $N = \{f(x) = 0\}$ is a μ-null set. Hence, $\bigcup_n E \cap E_n = E \setminus N$. Thus, $\mu(E \setminus N) \leq \sum_n \mu(E \cap E_n)$. Now,

$$\frac{1}{n}\mu(E \cap E_n) \leq \int_{E \cap E_n} f d\mu \leq \int_E f d\mu = 0.$$

Thus, $\mu(E \cap E_n) = 0$ for each $n \in \mathbb{N}$. Hence,

$$\mu(E \setminus N) \leq \sum_n \nu(E \cap E_n) = 0$$

and thus $\mu(E) = \mu(E \cap N) + \mu(E \setminus N) = 0 + 0 = 0$.

2. Assuming both μ and ν are σ-finite, we now wish to show that if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$ and

$$\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1} \mu\text{-almost everywhere.}$$

We have shown that $\mu \ll \nu$ in the item 1 above. By Radon-Nikodym theorem, we have the derivative $g = d\mu/d\nu$ which is a measurable function $g : X \to [0, \infty]$ such that

$$\mu(E) = \int_E g d\nu.$$

Denote $f = d\nu/d\mu : X \to [0, \infty]$ which is such that

$$\nu(E) = \int_E f d\mu.$$

We are given that $f > 0$ μ-almost everywhere. We wish to show that $g = 1/f$ μ-almost everywhere.

As we have seen that for an L^+ function h, we obtain a positive measure given by $\mu_h = \int_E h d\mu$, therefore we deduce that $\mu = \nu_g$ and $\nu = \mu_f$. Consequently, denoting $N = \{f(x) = 0\}$ to be the μ-null set, we obtain

$$\begin{aligned}
\int_E \frac{1}{f} d\nu &= \int_{E \setminus N} \frac{1}{f} f d\mu + \int_{E \cap N} \frac{1}{f} d\nu \\
&= \int_{E \setminus N} d\mu + 0 \\
&= \mu(E \setminus N).
\end{aligned}$$

As $\mu(E \cap N) = 0$, therefore adding this to above we add

$$\int_E \frac{1}{f} d\nu = \mu(E \setminus N) + \mu(E \cap N) = \mu(E).$$

Thus by almost everywhere uniqueness of Radon-Nikodym derivative of μ w.r.t. ν, we see that $1/f = g$ μ-almost everywhere. $\square$

Lemma 8.4.7. *Let $(X, \mathcal{A})$ be a measurable space with μ and ν be two finite positive measures. Suppose*

$$f = \frac{d\nu}{d(\mu + \nu)}.$$

Assume that $1 - f > 0$. Then,

$$\nu(E) = \int_E \frac{f}{1-f} d\mu,$$

equivalently, that

$$\frac{d\nu}{d\mu} = \frac{f}{1-f}.$$

Proof. We first show that $g := 1 - f$ is equal to the derivative $d\mu/d(\mu + \nu)$. Observe that $g > 0$. Indeed, for this, we need to show that for any $E \in \mathcal{A}$, we have

$$\mu(E) = \int_E g d(\mu + \nu).$$

To this end, we see that by definition of f and finiteness of μ, ν and thus $\mu + \nu$ as measures, we may deduce

$$\begin{aligned}
\int_E g d(\mu + \nu) &= \int_E (1 - f) d(\mu + \nu) \\
&= \int_E d(\mu + \nu) - \int_E f d(\mu + \nu) \\
&= \mu(E) + \nu(E) - \nu(E) \\
&= \mu(E),
\end{aligned}$$

as required. We may therefore write $\mu = (\mu + \nu)_g$ as the notation introduced in the class for positive measures defined by positive measurable functions.

Next, we claim that the function f/g is the derivative $d\nu/d\mu$. For this, we wish to show that for any measurable $E \in \mathcal{A}$, we have that

$$\nu(E) = \int_E \frac{f}{g} d\mu.$$

As $\mu = (\mu + \nu)_g$, hence we see that

$$\begin{aligned}
\int_E \frac{f}{g} d\mu &= \int_E \frac{f}{g} g d(\mu + \nu) \\
&= \int_E f d(\mu + \nu) \\
&= \nu(E),
\end{aligned}$$

as required. $\square$

Example 8.4.8. Let $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ be a measurable space and ν be a σ-finite signed measure. Further, let μ be the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

1. We wish to show that $\nu \ll \mu$. This is immediate, as $\mu(E) = 0$ if and only if $E = \emptyset$, and hence $\nu(E) = 0$ by definition.

2. We now wish to compute the derivative $d\nu/d\mu$. This is straightforward, for we first observe that the following function

$$\begin{aligned} f : \mathbb{N} &\longrightarrow [0, \infty] \\ n &\longmapsto \nu(\{n\}) \end{aligned}$$

is measurable. Indeed, this is because the σ-algebra on $\mathbb{N}$ is the power set $\mathcal{P}(\mathbb{N})$. We thus claim that

$$f = \frac{d\nu}{d\mu}.$$

Indeed, pick any measurable set $E \subseteq \mathbb{N}$. Note that it is countable in size. Observe that

$$\begin{aligned} \int_E f d\mu &= \sum_{n \in E} f(n) \\ &= \sum_{n \in E} \nu(\{n\}) \\ &= \nu\left(\coprod_{n \in E} \{n\}\right) \\ &= \nu(E) \end{aligned}$$

where the second-to-last equality is obtained from the fact ν is a measure. This completes the proof.

Lemma 8.4.9. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two σ-finite positive measures on $(X, \mathcal{A})$. Let $\lambda = \mu + \nu$. Then the following are equivalent*

1. *$\mu \perp \nu$,*
2. *if $f = d\mu/d\lambda$ and $g = d\nu/d\lambda$, then*

$$fg = 0 \ \lambda\text{-almost everywhere.}$$

Proof. (1. $\Rightarrow$ 2.) As $\mu \perp \nu$, therefore there exists a ν-null set A and a μ-null set B such that

$$A \amalg B = X. \tag{9.1}$$

For any measurable $E \subseteq X$, we have

$$\begin{aligned} \mu(E) &= \int_E f d\lambda \\ \nu(E) &= \int_E g d\lambda. \end{aligned}$$

We first observe that if any of the μ or ν is the zero measure, then we are done. Indeed, for if $\mu = 0$, then we deduce that $\mu(X) = 0$ and hence $f = 0$ λ-a.e. Consequently, $fg = 0$ λ-a.e. Hence, we may now assume that none of the μ and ν are 0 measures.

Observe that since $\mu(B) = 0$, therefore $\int_B f d\lambda = 0$. As $\lambda(B) = \mu(B) + \nu(B) = \nu(B)$, therefore we deduce from the fact that $\nu \neq 0$ and $\nu(A) = \nu(X \setminus B) = 0$ that $\nu(B) \neq 0$. Hence,

$$\lambda(B) \neq 0. \tag{9.2}$$

For exactly the same reasoning applied on $\nu(A) = 0$, we deduce that

$$\lambda(A) \neq 0. \tag{9.3}$$

Hence, we have that $\int_B f d\lambda = 0 = \int_A g d\lambda$. By Eqns (9.2) and (9.3), we conclude that $f = 0$ λ-a.e. over B and $g = 0$ λ-a.e. over A.

Consider the set $N = \{f(x) \neq 0\} \cap \{g(x) \neq 0\}$. Writing

$$N = (N \cap A) \amalg (N \cap B),$$

we observe that

1. $N \cap A$ is ν-null as A is ν-null,
2. $N \cap A$ is μ-null as $\{g(x) \neq 0\} \cap A$ is λ-null and over A, we have $\lambda = \mu$,
3. $N \cap B$ is μ-null as B is μ-null,
4. $N \cap B$ is ν-null as $\{f(x) \neq 0\} \cap B$ is λ-null and over B, $\lambda = \nu$.

Hence, we see that $N \cap A$ and $N \cap B$ both are λ-null. Consequently, N is λ-null.

(2. $\Rightarrow$ 1.) For any measurable $E \subseteq X$, we have

$$\begin{aligned} \mu(E) &= \int_E f d\lambda \\ \nu(E) &= \int_E g d\lambda. \end{aligned}$$

Consider the following measurable sets

$$\begin{aligned} A &= \{g(x) = 0\} \\ B &= \{g(x) \neq 0\} \cap \{f(x) = 0\} \\ N &= \{g(x) \neq 0\} \cap \{f(x) \neq 0\}. \end{aligned}$$

Clearly, $X = A \amalg B \amalg N$. Furthermore, as $fg = 0$ λ-a.e, therefore N is λ-null. Over A we see that ν is 0 and over B we see that μ is 0. As N is λ-null, therefore it is both μ and ν-null as well. Consequently, we have

$$X = A \amalg (B \amalg N)$$

where A is ν-null and $B \amalg N$ is μ-null, as required. □

Lemma 8.4.10. *Let $(X, \mathcal{A}, \nu)$ be a signed space. Then,*

1. $\frac{d\nu^+}{d|\nu|} = \chi_P$,

2. $\frac{d\nu^-}{d|\nu|} = \chi_N$.

Proof. First, observe that these derivatives exists because $\nu^+ \ll |\nu|$ and $\nu^- \ll |\nu|$. By Jordan decomposition of ν, we have

$$\nu = \nu^+ - \nu^-$$

where $\nu^+(E) = \nu(P \cap E)$ and $\nu^-(E) = -\nu(N \cap E)$, where $X = P \amalg N$ is the Hahn-decomposition of X into a positive set P and a negative set N obtained by ν and $E \in \mathcal{A}$.

1. We claim that $\frac{d\nu^+}{d|\nu|}$ is given by χ_P. To this end, we need only show that

$$\nu^+(E) = \int_E \chi_P d\,|\nu|$$

as by Radon-Nikodym theorem, we know that the derivatives are unique $|\nu|$-almost everywhere, and therefore ν-almost everywhere.

Now, we see that

$$\begin{aligned}\int_E \chi_P d\,|\nu| &= |\nu|\,(E \cap P)\\ &= \nu^+(E \cap P) + \nu^-(E \cap P)\\ &= \nu(E \cap P \cap P) - \nu(E \cap P \cap N)\\ &= \nu(E \cap P) - \nu(\emptyset)\\ &= \nu^+(E),\end{aligned}$$

as needed.

2. We proceed similarly as above and claim that χ_N is the derivative $\frac{d\nu^-}{d|\nu|}$. Indeed, we see that

$$\begin{aligned}\int_E \chi_N d\,|\nu| &= |\nu|\,(E \cap N)\\ &= \nu^+(E \cap N) + \nu^-(E \cap N)\\ &= \nu(E \cap N \cap P) - \nu(E \cap N \cap N)\\ &= \nu(\emptyset) - \nu(E \cap N)\\ &= \nu^-(E),\end{aligned}$$

as required. $\square$

Lemma 8.4.11. *Let $(X, \mathcal{A}, \nu)$ be a signed space and let $f : X \to \mathbb{C}$ be a measurable function. Define*

$$\int_X f d\nu = \int_X f d\nu^+ - \int_X f d\nu^-$$

where $\nu = \nu^+ - \nu^-$ is the Jordan decomposition of ν. Then, *1. we have*

$$\left|\int_X f d\nu\right| \leq \int_X |f| \, d\,|\nu|\,,$$

2. for any $E \in \mathcal{A}$, we have

$$|\nu|\,(E) = \sup\left\{\left|\int_E f d\nu\right| \mid |f| \leq 1\right\}.$$

Proof. 1. We may write

$$\begin{aligned}\left|\int_X f d\nu\right| &= \left|\int_X f d\nu^+ - \int_X f d\nu^-\right| \\ &\leq \left|\int_X f d\nu^+\right| + \left|\int_X f d\nu^-\right| \\ &\leq \int_X |f| \, d\nu^+ + \int_X |f| \, d\nu^-. \end{aligned} \tag{11.1}$$

We now claim that $\int_X |f| \, d\nu^+ + \int_X |f| \, d\nu^- = \int_X |f| \, d\,|\nu|$. Indeed, we first observe that for any $E \in \mathcal{A}$, we have $\nu^+(E) = \int_E \chi_P d\,|\nu|$ and $\nu^-(E) = \int_E \chi_N d\,|\nu|$. Consequently, we get

$$\begin{aligned}\int_X |f| \, d\nu^+ + \int_X |f| \, d\nu^- &= \int_X |f| \, \chi_P d\,|\nu| + \int_X |f| \, \chi_N d\,|\nu| \\ &= \int_X |f| \, (\chi_P + \chi_N) d\,|\nu| \\ &= \int_X |f| \cdot 1 d\,|\nu| \\ &= \int_X |f| \, d\,|\nu|\,, \end{aligned} \tag{11.2}$$

as required. Hence we conclude by Eqns (11.1) and (11.2).

2. Let $\mathcal{Z} := \{|\int_E f d\nu| \mid |f| \leq 1\}$. We first see that for any measurable $f : X \to \mathbb{C}$ with $|f| \leq 1$, we have the following by item 1 above

$$\begin{aligned}\left|\int_E f d\nu\right| &\leq \int_E |f| \, d\,|\nu| \\ &\leq \int_E d\,|\nu| \\ &\leq |\nu|\,(E). \end{aligned}$$

Hence, $\sup \mathcal{Z} \leq |\nu|\,(E)$.

For the converse, we wish to show that $|\nu|\,(E) \leq \sup \mathcal{Z}$. If $\sup \mathcal{Z} = \infty$, then there is nothing to be shown. So we may assume $\sup \mathcal{Z} < \infty$. As the constant function 1 is in the collection, therefore

$$|\nu(E)| \leq \sup \mathcal{Z} < \infty. \tag{11.3}$$

In order to show $|\nu|\,(E) \leq \sup \mathcal{Z}$, it suffices to find a measurable function $f : X \to \mathbb{C}$ such that $|f| \leq 1$ and $|\nu|\,(E) \leq |\int_E f d\nu|$. Indeed, denoting by $X = P \amalg N$ to be the Hahn-decomposition of

X obtained by ν, we consider $f = \chi_P - \chi_N$. Clearly, image of f is $\{-1, 0, 1\}$ as $A \cap B = \emptyset$, hence $|f| \leq 1$. Moreover, we observe that

$$\begin{aligned}\left|\int_E f d\nu\right| &= \left|\int_E f d\nu^+ - \int_E f d\nu^-\right| \\ &= \left|\int_E (\chi_P - \chi_N) d\nu^+ - \int_E (\chi_P - \chi_N) d\nu^-\right|. \end{aligned} \tag{11.4}$$

By Eq. (11.3), we deduce that $\nu^+(E)$ and $\nu^-(E)$ are finite. Furthermore, over E we have that χ_P and χ_N are both in $\mathcal{L}^1(\nu^+)$ and $\mathcal{L}^1(\nu^-)$. With this, we may continue Eq. (11.4) as follows

$$\begin{aligned}\left|\int_E f d\nu\right| &= \left|\int_E \chi_P d\nu^+ - \int_E \chi_N d\nu^+ - \int_E \chi_P d\nu^- + \int_E \chi_N d\nu^-\right| \\ &= \left|\int_E \chi_P d\nu^+ - 0 - 0 + \int_E \chi_N d\nu^-\right| \\ &= \nu^+(E \cap P) + \nu^-(E \cap N) \\ &= \nu^+(E) + \nu^-(E) \\ &= |\nu|(E). \end{aligned}$$

where in the second equality we have used the fact the fact that $\nu^+(E) := \nu(E \cap P)$ and $\nu^-(E) := \nu(E \cap N)$. This shows that for some $f : X \to \mathbb{C}$ measurable with $|f| \leq 1$ we have $|\int_E f d\nu| = |\nu|(E)$, which consequently shows that $|\nu|(E) \leq \sup \mathcal{Z}$. This completes the proof. □

Example 8.4.12. We wish to find those signed spaces $(X, \mathcal{A}, \nu)$ which satisfies property 1 below. Further, we also wish to find those which satisfies 2 as below:

1. For c the counting measure on $(X, \mathcal{A})$, we have $c \ll \nu$.
2. For $x_0 \in X$ and the Dirac measure δ_{x_0}, we have $\delta_{x_0} \ll \nu$.

1. Let $E \in \mathcal{A}$. We know that $c(E) = 0$ iff $E = \emptyset$. Consequently, if $\nu(E) = 0$, then $c(E) = 0$ iff $E = \emptyset$. That is, $\nu(E) = 0$ iff $E = \emptyset$. Hence all those signed spaces $(X, \mathcal{A}, \nu)$ whose only null set is $\emptyset$ can only be such that $c \ll \nu$.

2. Let $E \in \mathcal{A}$. We know that $\delta_{x_0}(E) = 0$ iff $x_0 \notin E$. Thus if $\nu(E) = 0$ and $\delta_{x_0} \ll \nu$, then $x_0 \notin E$. Hence, $(X, \mathcal{A}, \nu)$ is a signed space such that all its null sets does not contain x_0. This completes the characterizations.

Lemma 8.4.13. *Let $(X, \mathcal{A}, \nu)$ be a signed space. Then,*

1. *If $\{E_n\} \subseteq \mathcal{A}$ be an increasing collection of measurable sets, then*

$$\nu\left(\bigcup_n E_n\right) = \lim_{n\to\infty} \nu(E_n).$$

2. *If $\{E_n\} \subseteq \mathcal{A}$ be a decreasing collection of measurable sets such that $\nu(A_1)$ is finite, then*

$$\nu\left(\bigcap_n E_n\right) = \lim_{n\to\infty} \nu(E_n).$$

Proof. 1. Denote $F_1 = E_1$ and $F_n = E_n \setminus E_{n-1}$ for $n \geq 2$. Observe that $\{F_n\}$ are disjoint, but

$$\bigcup_n E_n = \coprod_n F_n. \tag{9.1}$$

Now observe that $E_n = F_n \amalg E_{n-1}$. This is recursive relation, which when unravelled, yields

$$E_n = F_n \amalg F_{n-1} \amalg \cdots \amalg F_1.$$

Applying ν yields

$$\nu(E_n) = \sum_{k=1}^{n} \nu(F_k). \tag{9.2}$$

It follows from Eqns (9.1) and (9.2) that

$$\begin{aligned}
\nu\left(\bigcup_n E_n\right) &= \nu\left(\coprod_n F_n\right) \\
&= \sum_{k=1}^{\infty} \nu(F_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \nu(F_k) \\
&= \lim_{n\to\infty} \nu(E_n),
\end{aligned}$$

as needed.

2. Consider the sequence $F_1 = E_1$ and $F_n = E_1 \setminus E_n$ for $n \geq 2$. Note that $\{F_n\}$ is increasing. Hence by item 1, we have

$$\nu\left(\bigcup_n F_n\right) = \lim_{n\to\infty} \nu(F_n). \tag{9.3}$$

Now observe that

$$E_1 = F_n \amalg E_n$$

for each $n \in \mathbb{N}$. Hence, applying ν we yield

$$\nu(E_1) = \nu(F_n) + \nu(E_n).$$

As $\nu(E_1)$ is finite, therefore the RHS in above equation is finite. Consequently, each term in the above equation is finite. Hence we may write it as

$$\nu(E_1) - \nu(F_n) = \nu(E_n).$$

Taking $n \to \infty$ yields

$$\nu(E_1) - \lim_{n\to\infty} \nu(F_n) = \lim_{n\to\infty} \nu(E_n)$$

which by Eq. (9.3), yields

$$\nu(E_1) - \nu\left(\bigcup_n F_n\right) = \lim_{n\to\infty}\nu(E_n). \tag{9.4}$$

We now claim that if $A \in \mathcal{A}$ and $B \subseteq A$ in $\mathcal{A}$ is such that $\nu(B)$ is finite, then $\nu(A\backslash B) = \nu(A)-\nu(B)$. Indeed, we may write $A = (A \setminus B) \amalg B$ where $A \setminus B$ is measurable as well. Applying ν, we yield $\nu(A) = \nu(A \setminus B) + \nu(B)$. As $\nu(B)$ is finite, therefore we may subtract both sides by $\nu(B)$ to yield $\nu(A \setminus B) = \nu(A) - \nu(B)$, as desired.

Using the above proved statement on Eq. (9.4), we obtain

$$\begin{aligned}\lim_{n\to\infty} &= \nu\left(E_1 \setminus \bigcup_n F_n\right)\\ &= \nu\left(E_1 \cap \bigcap_n F_n^c\right)\\ &= \nu\left(\bigcap_n E_1 \cap F_n^c\right)\\ &= \nu\left(\bigcap_n E_n\right),\end{aligned}$$

as desired. □

Lemma 8.4.14. *Let (X, Σ, μ) be a measure space and $f, g : X \to [0, \infty)$ be two non-negative measurable functions such that $f(x)g(x) = 0$ for almost all $x \in X$. Suppose for each $E \in \Sigma$ we have*

$$\mu(E) = \int_E f d\mu.$$

Define for each $E \in \Sigma$

$$\nu(E) = \int_E g d\mu.$$

Then $\mu \perp \nu$.

Proof. We know that ν as defined is a positive measure. Let $N = \{f(x)g(x) \neq 0\}$. This is a null-set. Consequently, we wish to find A and B measurable subsets such that $X = A \amalg B$ with A being μ-null and B being ν-null.

Define $A = \{f(x) = 0\}$ and $B = \{g(x) = 0 \ \&\ f(x) \neq 0\}$. Observe that $X = A \amalg B \amalg N$. Let $X_1 = A \amalg N$ and $X_2 = B$. Consequently $X = X_1 \amalg X_2$. Now, for any measurable $A' \subseteq X_1$, we may write $A' = (A' \cap A) \amalg (A' \cap N)$

$$\mu(A') = \int_{A'\cap A} f d\mu + \int_{A'\cap N} f d\mu = \int_{A'\cap A} 0 d\mu + \int_{A'\cap N} f d\mu = 0 + 0 = 0$$

where the latter term is zero because it is an integral over a measure 0 subset. Similarly, for any measurable $B' \subseteq X_2$, we see that

$$\nu(B') = \int_{B'} g d\mu = \int_{B'} 0 d\mu = 0.$$

Hence we have shown that X_1 is μ-null and X_2 is ν-null, as required. □

Lemma 8.4.15. *Let $(X, \mathcal{A}, \nu)$ be a signed space. Then,*

1. *If $A \in \mathcal{A}$ is a positive set, then $B \subseteq A$ such that $B \in \mathcal{A}$ is also a positive set.*
2. *If $\{A_n\} \subseteq \mathcal{A}$ is a sequence of positive sets, then $\bigcup_n A_n$ is a positive set.*

Proof. 1. Pick any $C \subseteq B$. As $B \subseteq A$, therefore $C \subseteq A$. As A is positive, thus $\nu(C) \geq 0$, as needed.

2. Let $B_1 = A_1$ and $B_n = A_n \setminus (A_1 \cup \dots \cup A_{n-1})$. Then observe that $\{B_n\}$ is a disjoint sequence of sets in $\mathcal{A}$, each positive as well by item 1. Furthermore, observe that

$$\coprod_n B_n = \bigcup_n A_n.$$

Now pick any $E \subseteq \bigcup_n A_n$ and denote $E_n = E \cap B_n$. Then, since B_n are disjoint, thus so is $\{E \cap B_n\}$. Furthermore $E = \coprod_n E_n$. Hence, we obtain

$$\mu(E) = \mu\left(\coprod_n E_n\right) = \sum_n \mu(E_n).$$

As each B_n is a positive set, so $E_n = E \cap B_n$ is a positive set as well by item 1. Consequently, $\mu(E_n) \geq 0$ for all $n \in \mathbb{N}$. Hence, from above, we deduce that

$$\mu(E) = \sum_n \mu(E_n) \geq 0,$$

as needed. $\square$

9 The dual of $L^p(\mathbb{R}^n)$: Riesz Representation theorem

Definition 9.0.1. (**Linear Functional**) Suppose $(V, \mathbb{R}, \|\cdot\|)$ is a Banach Space. A **linear**[27] map $f : V \longrightarrow \mathbb{R}$ is called a linear functional.

Definition 9.0.2. (**Bounded linear functional**) A linear functional $\phi : V \longrightarrow \mathbb{R}$ where $(V, \mathbb{R}, \|\cdot\|)$ is a Banach space is called bounded if

$$\exists\, c \geq 0 \text{ such that } |\phi(x)| \leq c\|x\| \;\forall\, x \in V.$$

The space of all such bounded linear functionals is denoted by

$$\mathcal{B}(V).$$

That is, any $\phi \in \mathcal{B}(V)$ is a bounded linear functional.

Proposition 9.0.3. *Suppose $(V, \mathbb{R}, \|\cdot\|_V)$ is a Banach Space and $\mathcal{B}(V)$ is the space of bounded linear functionals over V. Then,*

$$\mathcal{B}(V) \text{ forms a Vector Space}$$

[27] $f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2)\ \forall\, v_1, v_2 \in V$ and $\alpha, \beta \in \mathbb{R}$. Or more simply, a morphism in the Category of Vector Spaces **Vect** :):

and the map

$$\|\cdot\| : \mathcal{B}(V) \longrightarrow [0,\infty)$$

defined by

$$\begin{aligned}\|\phi\| &= \sup\left\{\frac{|\phi(x)|}{\|x\|_V} \ : \ x \in V\right\} \\ &= \inf\{c \ : \ |\phi(x)| \leq c\|x\|_V \ , \ \ x \in V\}\end{aligned}$$

for any $\phi \in \mathcal{B}(V)$ forms a norm on the Vector Space $\mathcal{B}(V)$.

Proof. Take any two $\phi_1, \phi_2 \in \mathcal{B}(V)$ and $\alpha, \beta \in \mathbb{R}$. By the very nature of their existence, ϕ_1 & ϕ_2 have to be bounded linear functionals. Suppose

$$\begin{aligned}|\phi_1(x)| \leq c_1\|x\|_V \\ |\phi_2(x)| \leq c_2\|x\|_V\end{aligned}$$

$\forall\, x \in V$. Then $\alpha\phi_1$ is a function such that:

$$|\alpha\phi_1(x)| \leq \alpha c_1\|x\|_V$$

Hence $\alpha\phi_1 \in \mathcal{B}(V)$. Similarly, $\beta\phi_2 \in \mathcal{B}(V)$. Now, since we have that

$$|\phi_1 + \phi_2| \leq |\phi_1| + |\phi_2|$$

Therefore, $\phi_1 + \phi_2 \in \mathcal{B}(V)$. Hence, $\mathcal{B}(V)$ is a Vector Space.

To see that $\|\cdot\|$ is a norm over $\mathcal{B}(V)$, we see that for any $\alpha \in \mathbb{R}$ and $\phi \in \mathcal{B}(V)$, we trivially get that

$$\|\alpha\phi\| = |\alpha|\,\|\phi\|$$

and, for $f_1, f_2 \in \mathcal{B}(V)$, we also note that

$$\|f_1 + f_2\| \leq \|f_1\| + \|f_2\|$$

Hence, $\|\cdot\|$ is a norm on Vector Space $\mathcal{B}(V)$. □

9.1 $\mathcal{B}(V)$ is a Banach Space

Proposition 9.1.1. *Suppose V is a Banach Space. Then $\mathcal{B}(V)$ is a Banach Space.*

Proof. Take any Cauchy sequence $\{\phi_n\}$ in $\mathcal{B}(V)$. Now, since ϕ_n's are bounded linear functionals, therefore,

$$\exists\, c_n \geq 0 \text{ such that } |\phi_n(x)| \leq c_n\|x\|_V \;\forall\, x \in V$$

Now take any $x \in V$. Since $\phi_n(x) \in V$, we therefore have a sequence $\{\phi_n(x)\}$ in $\mathbb{R}$. We now note that

$$\begin{aligned}|\phi_n(x) - \phi_m(x)| &\leq \|x\|_V \times \sup\left\{\frac{|\phi_n - \phi_m|}{\|x\|_V} \ : \ x \in V\right\} \\ &= \|\phi_n - \phi_m\|\|x\|_V.\end{aligned}$$

Hence, $\{\phi_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$. Now write

$$\phi(x) = \varinjlim_n \phi_n(x).$$

Since $\mathbb{R}$ is complete, therefore $\phi(x) \in \mathbb{R}$. But our choice of x was arbitrary, hence $\phi(x) = \varinjlim_n \phi_n(x) < \infty \; \forall \, x \in V$. Hence $\phi \in \mathcal{B}(V)$.

Moreover,

$$\begin{aligned}
\varinjlim_n \|\phi_n\| &= \varinjlim_n \sup\left\{\frac{|\phi_n(x)|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \sup\left\{\varinjlim_n \frac{|\phi_n(x)|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \sup\left\{\frac{\left|\varinjlim_n \phi_n(x)\right|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \sup\left\{\frac{|\phi(x)|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \|\phi\|
\end{aligned}$$

Hence proved. □

10 Remarks on Banach spaces

Following are some exercises, examples and remarks on Banach spaces.

10.1 Normed linear spaces

Remark 10.1.1. a) We claim that any linear space could be normed. Let X be a linear space and $\{b_j\}$ be a Hamel basis. Then for each $x \in X$ there are unique finitely many non-zero elements $c_{x_1}, \ldots, c_{x_k} \in \mathbb{K}$ such that $x = c_{x_1} b_{j_1} + \ldots c_{x_k} b_{j_k}$. Define the following map

$$\begin{aligned}
\| - \| : X &\longrightarrow \mathbb{R}_{\geq 0} \\
x &\longmapsto \max\{|c_{x_1}|, \ldots, |c_{x_k}|\}.
\end{aligned}$$

We claim that $\| - \|$ is a norm. Indeed, if $\|x\| = 0$, then $c_{x_i} = 0$ for all $i = 1, \ldots, k$. Consequently, $x = 0$. If $x = 0$, then it is clear by uniqueness of c_{x_i} that all $c_{x_i} = 0$.

Consider $c \in \mathbb{K}$ and $x \in X$. Then $\|cx\| = \max |cc_{x_1}|, \ldots, |cc_{x_k}| = |c| \max\{|c_{x_1}|, \ldots, |c_{x_k}|\} = |c| \, \|x\|$.

We finally wish to show triangle inequality. Pick $x, y \in X$. Then, (we allow c_{x_i} and c_{y_i} to be zero)

$$\begin{aligned}
\|x + y\| &= \max\{|c_{x_1} + c_{y_1}|, \ldots, |c_{x_k} + c_{y_k}|\} \\
&\leq \max\{|c_{x_1}| + |c_{y_1}|, \ldots, |c_{x_k}| + |c_{y_k}|\} \\
&\leq \max\{|c_{x_1}|, \ldots, |c_{x_k}|\} + \max\{|c_{y_1}|, \ldots, |c_{y_k}|\} \\
&= \|x\| + \|y\|.
\end{aligned}$$

Hence every linear space is normable.

b) We claim that not all metric on a linear space X comes from a norm on X. Indeed, consider the following metric:

$$\begin{aligned} d : X \times X &\longrightarrow \mathbb{R}_{\geq 0} \\ (x, y) &\longmapsto \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases} \end{aligned}$$

Indeed it is clear that $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$. For triangle inequality, we need only consider the case when $x \neq y$ and to show that for any $z \in X$ we have

$$1 = d(x, y) \leq d(x, z) + d(y, z).$$

It is clear that we need only show that $d(x, z)$ and $d(y, z)$ are both not simultaneously 0. Indeed, if both are simultaneously 0, then $x = z = y$, a contradiction. Hence d is indeed a metric.

We claim that d is not induced by any norm. Indeed, assume to the contrary it is induced by a norm $\| - \|$. It follows that

$$\|x\| = d(x, 0) = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Since $\| - \|$ is a norm, it follows that for any $c \neq 1$ in $\mathbb{K}$ and $x \neq 0$ in X, we must have $\|cx\| = 1$ as $cx \neq 0$. We now have the following contradiction

$$1 = \|cx\| = |c| \, \|x\| = |c| \neq 1.$$

This completes the proof.

Remark 10.1.2. We wish to show that the following are equivalent for a linear space X with a function $\| : \|X \to \mathbb{R}_{\geq 0}$ satisfying $\|x\| = 0$ iff $x = 0$ and $\|cx\| = |c| \, \|x\|$ for all $c \in \mathbb{K}$ and $x \in X$:

1. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.
2. The closed unit ball $B_1[0] = \{x \in X \mid \|x\| \leq 1\}$ is convex.

(1. $\Rightarrow$ 2.) Pick $x, y \in B_1[0]$ and $c \in [0, 1]$. We wish to show that $cx + (1 - c)y \in B_1[0]$. Indeed, since $\|x\|, \|y\| \leq 1$, therefore we have

$$\begin{aligned} |cx + (1 - c)y| &\leq |c| \, \|x\| + |1 - c| \, \|y\| \\ &\leq c + (1 - c) \\ &= 1. \end{aligned}$$

(2. $\Rightarrow$ 1.) Pick $x, y \in X$. If any of the x or y is 0, then triangle inequality is immediate. Hence we may assume x and y are both not 0. Then $\frac{x}{\|x\|}, \frac{y}{\|y\|} \in B_1[0]$. Let $c = \frac{\|x\|}{\|x\| + \|y\|}$ so that $1 - c = \frac{\|y\|}{\|x\| + \|y\|}$. It is clear that $c \in [0, 1]$. By convexity of $B_1[0]$, it follows that

$$c \frac{x}{\|x\|} + (1 - c) \frac{y}{\|y\|} \in B_1[0].$$

But we have

$$c\frac{x}{\|x\|}+(1-c)\frac{y}{\|y\|}=\frac{x}{\|x\|+\|y\|}+\frac{y}{\|x\|+\|y\|}$$

hence the RHS above is in $B_1[0]$. Taking norm, we see

$$\|\frac{x}{\|x\|+\|y\|}+\frac{y}{\|x\|+\|y\|}\|=\frac{\|x+y\|}{\|x\|+\|y\|}\leq 1$$

from which we get

$$\|x+y\|\leq\|x\|+\|y\|,$$

as required.

Example 10.1.3. Consider $C[a,b]$ be the $\mathbb{R}$-vector space of all continuous functions on $[a,b]$. Define for any $1\leq p<\infty$

$$\|f\|_p:=\left(\int_a^b |f(t)|^p\,dt\right)^{\frac{1}{p}}.$$

a) We wish to show that $\|-\|_p$ is a norm on $C[a,b]$. Indeed, if $f\in C[a,b]$ such that $\|f\|_p=0$, then we have

$$\int_a^b |f(t)^p|\,dt=0.$$

We wish to show that $f=0$. Suppose not, so that $f(t_0)\neq 0$ at a point $t_0\in[a,b]$. If $t_0=a$ or b, then by continuity there is a point in (a,b) where f is non-zero. Replace t by that point in (a,b). It follows by continuity that there exists $\delta>0$ such that f is non-zero on $I=[t_0-\delta,t_0+\delta]\subseteq(a,b)$. Let $m=\min_{t\in I}|f(t)|^p>0$ which exists as f is continuous on compact I and $f\neq 0$ on I. Then

$$0=\int_a^b |f(t)|^p\,dt\geq\int_{t_0-\delta}^{t_0+\delta} m dt=m\cdot(2\delta)>0,$$

a contradiction. It follows that $f=0$ on $[a,b]$.

We now wish to show triangle inequality. For this, we invoke the fact that $C[a,b]$ is contained inside the linear space $L^p[a,b]$ of $\mathbb{R}$-valued Lebesgue measurable functions on $[a,b]$. Moreover, the function

$$\|f\|_p:=\left(\int_{[a,b]}|f|^p\,dm\right)^{1/p}$$

for $f\in L^p[a,b]$ defines a norm. Moreover if f is continuous, then the above Lebesgue integral on $[a,b]$ agrees with the usual Riemann integral. So we may conclude that there is an inclusion of linear spaces

$$(C[a,b],\|-\|_p)\subseteq(L^p[a,b],\|-\|_p).$$

We know that $(L^p[a,b], \|-\|_p)$ forms a normed linear space, where triangle inequality is established by Minkowski's inequality. Using the same theorem on the subspace $(C[a,b], \|-\|_p)$, we get the desired result.

b) We claim that $(C[0,2], \|-\|_1)$ is not complete. It suffices to show a Cauchy sequence which is not convergent. Indeed consider $f_n(x)$ as follows:

$$f_n(x) = \begin{cases} x^n & \text{if } x \in [0,1] \\ 1 & \text{if } x \in (1,2]. \end{cases}$$

We first claim that (f_n) is Cauchy in $C[0,2]$. Indeed, for $n \geq m$, we have

$$\begin{aligned} \|f_n - f_m\|_1 &= \int_0^2 |f_n(x) - f_m(x)|\, dx \\ &= \int_0^1 |x^n - x^m|\, dx \\ &= \int_0^1 x^m - x^n dx \\ &= \int_0^1 x^m dx - \int_0^1 x^n dx \\ &= \frac{1}{m+1} - \frac{1}{n+1} \\ &\leq \frac{1}{m+1}. \end{aligned}$$

So for a fixed $\epsilon > 0$, let $N \in \mathbb{N}$ be such that $1/N < \epsilon$. Then for all $n, m \geq N$, we have

$$\|f_n - f_m\|_1 \leq \frac{1}{m+1} \leq \frac{1}{N+1} < \epsilon,$$

as needed. Next, we claim that (f_n) doesn't converge in $C[0,2]$. Indeed, it would suffice to show that it converges in $L^1[0,2]$ to a non-continuous function. Indeed, consider the following simple function

$$f = \chi_{[1,2]}.$$

This is not continuous in $[0,2]$. We claim that $f_n \to f$ in $L^1[0,2]$. Indeed, we have

$$\begin{aligned} \|f_n - f\|_1 = \int_{[0,2]} |f_n - f|\, dm &= \int_{[0,1]} |f_n - f|\, dm + \int_{[1,2]} |f_n - f|\, dm \\ = \int_{[0,1]} |f_n - f|\, dm &= \int_0^1 |x^n|\, dx, \end{aligned}$$

where the last equality comes from Riemann and Lebesgue integrals being equal on compact intervals for Riemann integrable functions. Consequently, we have

$$\|f_n - f\|_1 = \frac{1}{n+1}$$

which converges to 0 as $n \to \infty$. Hence in $L^1[0,2]$, $f_n \to f$. As $C[0,2] \subseteq L^1[0,2]$ with the given norm, it follows that $(f_n) \subseteq C[0,2]$ does not converge in $C[0,2]$.

Example 10.1.4. Let $X = (C[0,1], \|\cdot\|_\infty)$. We wish to calculate the following:

1. $d(f_1, C)$ where $f_1(t) = t$ and C is the linear subspace of all constant functions,
2. $d(f_2, P)$ where $f_2(t) = t^2$ and P is the linear subspace of polynomials of degree at most 1.

1. We claim that $d(f_1, C) = 1/2$. Indeed, we have

$$\begin{aligned} d(f_1, C) &= \inf_{c\in C} \|f_1 - c\| = \inf_{c\in C} \sup_{t\in[0,1]} |t-c| \\ &= \inf_{c\in C} \begin{cases} c & \text{if } \frac{1}{2} \le c < \infty \\ 1-c & \text{if } -\infty < c < \frac{1}{2}. \end{cases} \\ &= \frac{1}{2}, \end{aligned}$$

as needed.

2. We claim that $d(f_2, P) = 1/8$. Pick any $at + b \in P$ for $a, b \in \mathbb{R}$. We first show that

$$\sup_{t\in[0,1]} \left|t^2 - at - b\right| = \max\left\{-b, 1-a-b, \frac{a^2}{4} + b\right\}. \tag{$*$}$$

Indeed, consider the discriminant $a^2 + 4b$ of $f(t) = t^2 - at - b$. There are two cases to be had here:

1. If $a^2 + 4b \le 0$: Then the maximum of $|f(t)|$ is equal to that of $f(t)$ and is achieved only on the boundary at $t = 0$ or 1 because $f(t) \ge 0$ for all $t \in [0,1]$. Consequently, $\sup_{t\in[0,1]} |f(t)| = -b$ or $1 - a - b$.
2. If $a^2 + 4b > 0$: Then the maximum of $|f(t)|$ is either on boundary at $t = 0, 1$ or at the point of minima of $f(t)$ at $t = a/2$, which thus becomes a point of maxima for $|f(t)|$. It follows that $\sup_{t\in[0,1]} |f(t)| = -b, 1-a-b$ or $\frac{a^2}{4} + b$.

These two cases shows the claim in Eqn $(*)$.

Consider now $f(a,b) = \max\left\{-b, 1-a-b, \frac{a^2}{4} + b\right\}$ as a function $f : \mathbb{R}^2 \to \mathbb{R}$. We wish to find $\inf_{(a,b)\in\mathbb{R}^2} f(a,b)$. First we observe the following three regions:

1. The region R_1 : This is

$$R_1 = \{(a,b) \in \mathbb{R}^2 \mid f(a,b) = -b\}.$$

2. The region R_2 : This is

$$R_2 = \{(a,b) \in \mathbb{R}^2 \mid f(a,b) = 1-a-b\}.$$

3. The region R_3 : This is

$$R_3 = \left\{(a,b) \in \mathbb{R}^2 \mid f(a,b) = \frac{a^2}{4} + b\right\}.$$

We now analyze bounds on a point $(a,b) \in R_i$ as follows.

1. If $(a,b) \in R_1$: Then we have

$$\begin{aligned} -b &> 1-a-b \\ -b &> a^2/4 + b \end{aligned}$$

solving which, we get bounds

$$\begin{aligned} a &< 1 \\ b &< -\frac{a^2}{8}. \end{aligned}$$

Hence, to minimize b, we need to maximize a, thus to get that $b < -1/8$. So we have $(a,b) = (1,-1/8)$ as a point of minima for $-b$.

2. If $(a,b) \in R_2$: Then we have

$$\begin{aligned} 1-a-b &> -b \\ 1-a-b &> \frac{a^2}{4}+b \end{aligned}$$

solving which we get bounds

$$\begin{aligned} a &< 1 \\ b &< \frac{1}{2}-\frac{a}{2}-\frac{a^2}{8}. \end{aligned}$$

Hence to minimize $1-a-b$, we have to maximize a and b. Doing so yields $a=1$ and $b=-1/8$. Hence $(a,b) = (1,-1/8)$ is a point of minima for $1-a-b$.

3. If $(a,b) \in R_3$: Then we have

$$\begin{aligned} \frac{a^2}{4}+b &< -b \\ \frac{a^2}{4}+b &< 1-a-b \end{aligned}$$

solving which, we get bounds

$$\begin{aligned} b &> -\frac{a^2}{8} \\ b &> -\frac{a^2}{8}-\frac{a}{2}+\frac{1}{2}. \end{aligned}$$

Hence to minimize $\frac{a^2}{4}+b$, we have to minimize b and a. Doing so, we obtain $b = -a^2/8$ which thus yields

$$a > 1.$$

Hence to minimize a, we have to take $a = 1$. Consequently, $(a,b) = (1,-1/8)$ is a point of minima for $a^2/4+b$.

From all the three cases above, we see that f minimizes at the point $(a,b) = (1,-1/8)$. Indeed, we see that $(1,-1/8) \in R_1 \cap R_2 \cap R_3$ as all three functions $-b, 1-a-b$ and $a^2/4+b$ are equal at it. Consequently, the $\inf_{(a,b)\in\mathbb{R}^2} f(a,b) = 1/8$, as required.

10.2 Properties

Proposition 10.2.1. *Let X be a normed linear space. The following are equivalent:*

1. *X is a Banach space.*
2. *$S^1(X) = \{x \in X \mid \|x\| = 1\}$ is a complete subset of X.*

Proof. content... $\square$

Proposition 10.2.2. *Let X be a normed linear space. Then the following are equivalent:*

1. *X is a Banach space.*
2. *Absolutely convergent series in X are convergent in X.*

Proof. 1. $\Rightarrow$ 2. Pick an absolutely convergent series $\sum_n x_n$ in X so that

$$\sum_n \|x_n\| < \infty.$$

It follows that $T_n = \sum_{k=1}^n \|x_k\|$ is a Cauchy sequence in $\mathbb{R}$. We wish to show that $\sum_n x_n$ converges in X. It suffices to show that the sequence $S_n = \sum_{k=1}^n x_k$ converges in X. We reduce to showing that (S_n) is Cauchy. Fix $\epsilon > 0$. For any $n \geq m$, we have

$$\begin{aligned}
\|S_n - S_m\| &= \|x_{m+1} + \cdots + x_n\| \\
&\leq \|x_{m+1}\| + \cdots + \|x_n\| \\
&= \left| \left(\sum_{k=1}^{n} \|x_k\| \right) - \left(\sum_{k=1}^{m} \|x_k\| \right) \right| \\
&= |T_n - T_m| < \epsilon
\end{aligned}$$

some $N \in \mathbb{N}$ and $n, m \geq N$ since (T_n) is Cauchy in $\mathbb{R}$. This shows that (S_n) is Cauchy, as required.

2. $\Rightarrow$ 1. Pick a Cauchy sequence $(x_n) \subseteq X$. We wish to show that there is a convergent subsequence of (x_n). We first find a subsequence of (x_n) which is better behaved. Indeed, by Cauchy condition, we find for each $k \geq 0$ a positive integer $N_k \in \mathbb{N}$ such that

$$\|x_n - x_m\| < \frac{1}{2^k}$$

for all $n, m \geq N_k$. We may assume N_k to be the least such possible by well-ordering on $\mathbb{N}$. Then we see that $N_{k+1} \geq N_k$ by minimality hypothesis. Thus consider the subsequence (x_{N_k}) of (x_n). Observe that

$$\|x_{N_{k+1}} - x_{N_k}\| < \frac{1}{2^k}$$

as $N_{k+1}, N_k \geq N_k$. We replace (x_n) by the subsequence (x_{N_k}) so that we may assume

$$\|x_{n+1} - x_n\| < \frac{1}{2^n} \; \forall n \in \mathbb{N}. \tag{3}$$

We now find the limit to which (x_n) converges. Indeed, define the following sequence in X:

$$\begin{aligned}
y_{n-1} &= \sum_{k=1}^{n-1} x_{k+1} - x_k \\
&= x_n - x_1.
\end{aligned}$$

We claim that $\sum_n x_{n+1} - x_n$ is an absolutely convergent series. Indeed, denote

$$\begin{aligned} S_{n-1} &:= \sum_{k=1}^{n-1} \|x_{k+1} - x_k\| \\ &\leq \sum_{k=1}^{n-1} \frac{1}{2^k} \end{aligned}$$

where the latter bound follows from Eqn. (3). Then, we see that for any $n \in \mathbb{N}$

$$S_n \leq \sum_{k=1}^{n} \frac{1}{2^k} < \sum_{k=1}^{\infty} \frac{1}{2^k} = M < \infty. \tag{4}$$

That is, (S_n) is a monotonically increasing positive bounded sequence in $\mathbb{R}$, therefore (S_n) is convergent. This shows that the series $\sum_n x_{n+1} - x_n$ is absolutely convergent. By our hypothesis, it follows that $\sum_n x_{n+1} - x_n$ is convergent in X. That is, the sequence

$$y_{n-1} = \sum_{k=1}^{n-1} x_k$$

of partial sums is convergent in X. But since $y_{n-1} = x_n - x_1$, it follows that (x_n) is a convergent sequence in X, as required. □

10.3 Bases & quotients

Lemma 10.3.1. *If X is a normed linear space with a Schauder basis, then X is separable.*

Proof. Let $(b_n) \subseteq X$ be a Schauder basis. Consider the following subset

$$D = \left\{ \sum_{k=1}^{n} q_k b_k \mid q_k \in E,\ n \in \mathbb{N} \right\}$$

where $E \subseteq \mathbb{K}$ is a countable dense subset. It is clear that D is countable. We claim that D is dense in X.

Pick any point $x \in X$. Since (b_n) is a Schauder basis, there exists $(c_k) \subseteq \mathbb{K}$ such that

$$x = \sum_{k=1}^{\infty} c_k b_k$$

where the series converges in X. Pick a ball $B_\epsilon(x)$ around x. We wish to show that $B_\epsilon(x) \cap D \neq \emptyset$. Indeed, consider $N \in \mathbb{N}$ large enough such that

$$\|x - \sum_{k=1}^{N} c_k b_k\| < \frac{\epsilon}{2}. \tag{9}$$

Moreover, for each $k = 1, \ldots, N$, consider $q_k \in E$ such that

$$|c_k - q_k| < \frac{\epsilon}{2 \cdot 2^k \|b_k\|} \tag{10}$$

which exists by density of E in $\mathbb{K}$. Hence, we have by Eqns (9) and (10) the following inequalities:

$$\begin{aligned}
\|x - \sum_{k=1}^{N} q_k b_k\| &\leq \|x - \sum_{k=1}^{N} c_k b_k\| + \|\sum_{k=1}^{N}(c_k - q_k) b_k\| \\
&< \frac{\epsilon}{2} + \sum_{k=1}^{N} |c_k - q_k| \, \|b_k\| \\
&< \frac{\epsilon}{2} + \frac{1}{2}\sum_{k=1}^{N} \frac{\epsilon}{2^k} \\
&= \frac{\epsilon}{2} + \frac{\epsilon}{2}\left(1 - \frac{1}{2^N}\right) \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,
\end{aligned}$$

as needed. This shows that $\sum_{k=1}^{N} q_k b_k \in B_\epsilon(x)$, that is D is dense in X. $\square$

Proposition 10.3.2 (2 out of 3 property)**.** *Let X be a normed linear space and $Y \subseteq X$ be a closed linear subspace. Then,*

1. *X, Y Banach implies X/Y Banach.*
2. *$X, X/Y$ Banach implies Y Banach.*
3. *$Y, X/Y$ Banach implies X Banach.*

Proof. 1. We have done in class that if X is Banach then for any closed linear subspace Y, X/Y is Banach.

2. We need the following lemma here:

Lemma 10.3.3. *Let X be a Banach space and $Y \subseteq X$ be a linear subspace. Then the following are equivalent:*

1. *Y is complete.*
2. *Y is closed.*

Proof of Lemma 10.3.3. 1. $\Rightarrow$ 2. Take $(y_n) \subseteq Y$ be a convergent sequence in X such that it converges to $x \in X$. We wish to show that $x \in Y$. Indeed, as $(y_n) \subseteq Y$ is convergent, so it is Cauchy. Since Y is complete, it follows that (y_n) converges to a point in Y. By uniqueness of point of convergence in a Hausdorff space, $x \in Y$.

2. $\Rightarrow$ 1. Pick a Cauchy sequence $(y_n) \subseteq Y$. We wish to show that it converges in Y. Indeed, (y_n) as a sequence in X is Cauchy and thus by completeness of X, we deduce that $y_n \to x$ in X. But since Y is closed, therefore by uniqueness of point of convergence, we must have $x \in Y$, as required. $\square$

Since X is Banach and Y is closed, it follows from Lemma 10.3.3 that Y is complete.

3. Pick a Cauchy sequence $(x_n) \subseteq X$. We wish to show that that it converges. We have a

sequence $(x_n + Y) \subseteq X/Y$. We first claim that $(x_n + Y)$ is Cauchy. Indeed, we have

$$\begin{aligned} \|x_n - x_m + Y\| &= \inf_{y \in Y} \|x_n - x_m + y\| \\ &\leq \|x_n - x_m\| < \epsilon \end{aligned}$$

for all $n, m \geq N$ for some $N \in \mathbb{N}$ as $(x_n) \subseteq X$ is Cauchy. As X/Y is Banach, it follows that $(x_n + Y) \to (x + Y)$ in X/Y. Consequently, for a fixed $\epsilon > 0$, we get

$$\|x_n - x + Y\| = \inf_{y \in Y} \|x_n - x + y\| < \epsilon/2 < \epsilon$$

for all $n \geq N$ for some $N \in \mathbb{N}$. It follows from above that there is a sequence $(y_n) \subseteq Y$ such that

$$\|x_n - x + y_n\| \leq \epsilon/2 < \epsilon. \tag{1}$$

We claim that $(y_n) \subseteq Y$ is Cauchy. Indeed, we first see from Eqn. (1) that

$$\|x_n + y_n - x\| < \epsilon$$

for all $n \geq N$. Consequently, the sequence $(x_n + y_n) \subseteq X$ converges to $x \in X$. Hence, $(x_n + y_n) \subseteq X$ is Cauchy, from which we get $N \in \mathbb{N}$ such that

$$\|x_n + y_n - x_m - y_m\| = \|x_n - x_m - (y_m - y_n)\| < \epsilon$$

for each $n, m \geq N$. We may write by triangle inequality the following:

$$|\|x_n - x_m\| - \|y_n - y_m\|| \leq \|x_n - x_m - (y_m - y_n)\| < \epsilon$$

so that

$$\|y_n - y_m\| < \|x_n - x_m\| + \epsilon \tag{2}$$

for all $n, m \geq N$. As $(x_n) \subseteq X$ is Cauchy, so for some $N' \in \mathbb{N}$ we have $\|x_n - x_m\| < \epsilon$ for all $n, m \geq N'$. Replacing N by maximum of N' and N, we obtain from Eqn. (2) the following:

$$\|y_n - y_m\| < 2\epsilon \; \forall n, m \geq N.$$

This shows that $(y_n) \subseteq Y$ is Cauchy. As Y is complete, therefore $y_n \to y \in Y$. As $x_n + y_n \to x$ in X, therefore $x_n \to x - y$ in X, thus showing that X is complete. $\square$

Proposition 10.3.4. *The Banach space ℓ^p is separable for all $1 \leq p < \infty$.*

Proof. Recall that

$$\ell^p = \left\{ (x_n) \mid x_n \in \mathbb{K} \;\&\; \sum_n |x_n|^p < \infty \right\}$$

with the norm being $\|(x_n)\|_p = \left(\sum_n |x_n|^p\right)^{1/p}$. Let $D \subseteq \mathbb{K}$ be a countable dense subset of $\mathbb{K}$ (which exists as $\mathbb{R}$ and $\mathbb{C}$ are separable in their usual topology). Using D we will construct a countable dense subset $F \subseteq \ell^p$. Indeed, consider the following subset of ℓ^p:

$$F = \bigcup_{N \geq 0} F_N$$

where

$$F_N = \{(x_n) \in \ell^p \mid x_n \in D,\ x_n = 0\ \forall n \geq N\}.$$

We see that F_N is countable as finite product of countable sets is countable and thus F is a countable union of countable sets, showing that F is countable. We next claim that F is dense in ℓ^p.

Pick any open set $B_r(y) \subseteq \ell^p$. Note that

$$B_r(y) = \left\{(x_n) \in \ell^p \mid \sum_n |x_n - y_n|^p < r^p\right\}.$$

As $y = (y_n) \in \ell^p$, therefore $\sum_n |y_n|^p = M < \infty$. Now observe that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\sum_{k=n}^{\infty} |y_n|^p < \epsilon \tag{5}$$

for all $n \geq N$. As $D \subseteq \mathbb{K}$ is dense and each $y_n \in \mathbb{K}$, therefore choose

$$x_n \in B_{r_n}(y_n) \cap D \subseteq \mathbb{K}$$

where $r_n = \frac{r}{2^{\frac{n+1}{p}}}$ for all $n \in \mathbb{N}$. Hence,

$$|x_n - y_n|^p < \frac{r^p}{2 \cdot 2^n}$$

for all $n \in \mathbb{N}$. As $\sum_{n=1}^{\infty} r^p/2^{n+1} = r^p/2$, therefore

$$\sum_{n\geq 1} |x_n - y_n|^p < \frac{r^p}{2}. \tag{6}$$

This shows that the element $(x_n) \in B_r(y) \subseteq \ell^p$.

Now, fix $\epsilon > 0$ so that there exists $K \in \mathbb{N}$ large enough using Eqn. (5) such that

$$\sum_{n=K}^{\infty} |y_n|^p < \epsilon. \tag{7}$$

Using Eqn. (6) and (7), we can write

$$\begin{aligned}
\sum_{n=1}^{K-1} |x_n - y_n|^p + \sum_{n=K}^{\infty} |y_n|^p &< \sum_{n=1}^{K-1} \frac{r^p}{2 \cdot 2^n} + \sum_{n=K}^{\infty} |y_n|^p \\
&< \frac{r^p}{2}\left(1 - \frac{1}{2^K}\right) + \sum_{n=K}^{\infty} |y_n|^p \\
&< \frac{r^p}{2}\left(1 - \frac{1}{2^K}\right) + \epsilon \\
&< \frac{r^p}{2}\left(1 - \frac{1}{2^N}\right) + \epsilon
\end{aligned}$$

for all $N \geq K$. So let $N \to \infty$ so that we obtain

$$\sum_{n=1}^{K-1} |x_n - y_n|^p + \sum_{n=K}^{\infty} |y_n|^p \leq \frac{r^p}{2} + \epsilon.$$

Thus taking $\epsilon = \frac{r^p}{4}$, we get $\tilde{K} \in \mathbb{N}$ such that

$$\sum_{n=1}^{\tilde{K}-1} |x_n - y_n|^p + \sum_{n=\tilde{K}}^{\infty} |y_n|^p \leq \frac{3r^p}{4} < r^p. \tag{8}$$

Define $\tilde{x} \in \ell^p$ as follows:

$$\tilde{x}_n = \begin{cases} x_n & \text{if } n \leq \tilde{K} - 1 \\ 0 & \text{if } n \geq \tilde{K}. \end{cases}$$

Then $\tilde{x} \in F_{\tilde{K}}$ and by Eqn. (8) it follows that

$$\sum_{n=1}^{\infty} |\tilde{x}_n - y_n|^p < r^p.$$

Consequently, $\tilde{x} \in F \cap B_r(y)$, as needed. □

Example 10.3.5 (ℓ^∞ is not separable)**.** We wish to show that ℓ^∞ does not have a Schauder basis. By Lemma 10.3.1, it suffices to show that ℓ^∞ is not separable. Suppose to the contrary that $D \subseteq \ell^\infty$ is a countable dense set. We will derive a contradiction to countability of D. Indeed, consider $\kappa = \{0, 1\}$ and the subset $\kappa^\infty \subseteq \ell^\infty$ of all sequences formed by 1 and 0. Observe that κ^∞ is uncountable.

Pick any $x \in \kappa^\infty$. We first claim that $B_{1/2}(x) \cap \kappa^\infty = \{x\}$. Indeed, if $y \in B_{1/2}(x)$, then $\sup_n |x_n - y_n| < 1/2$. It follows that there exists $0 < \epsilon < 1/2$ such that

$$|x_n - y_n| < \epsilon \ \forall n \in \mathbb{N}.$$

As $x_n = 0$ or 1, therefore

$$\begin{cases} -\epsilon < y_n < \epsilon & \text{if } x_n = 0 \\ 1 - \epsilon < y_n < 1 + \epsilon & \text{if } x_n = 1. \end{cases} \tag{9}$$

Hence, if $y \in \kappa^\infty$, then by Eqn. (9) it follows that $y_n = x_n$ for all $n \in \mathbb{N}$ and thus $x = y$.

We next show that for $x \neq x' \in \kappa^\infty$, the open balls $B_{1/2}(x) \cap B_{1/2}(x') = \emptyset$. Since $x \neq x'$, we may assume WLOG that there exists $m \in \mathbb{N}$ such that $x_m = 0$ and $x'_m = 1$. Thus, if $y \in B_{1/2}(x) \cap B_{1/2}(x')$, then by Eqn. (9), it follows that

$$\begin{gathered} -\epsilon < y_m < \epsilon \\ 1 - \epsilon < y_m < 1 + \epsilon. \end{gathered}$$

Since $\epsilon = 1/2$, therefore the above inequalities give a contradiction. Hence $B_{1/2}(x) \cap B_{1/2}(x') = \emptyset$.

We now complete the proof. As $D \subseteq \ell^\infty$ is dense, therefore $D \cap B_{1/2}(x) \neq \emptyset$ for all $x \in \kappa^\infty$. Pick one $d_x \in D \cap B_{1/2}(x)$ for each $x \in \kappa^\infty$. By above two claims, it follows that we have an injective map

$$f : \kappa^\infty \hookrightarrow D,$$

but κ^∞ is uncountable and D is countable, a contradiction. This completes the proof.

10.4 Continuous linear transformations

Example 10.4.1. We wish to show that the inverse of a bounded linear operator may not be bounded. Indeed consider $X = (P[0,1]_1, \|\cdot\|_{\sup})$ to be the normed linear space of all polynomials whose least degree term is of degree 1. Similarly, consider $Y = (P[0,1]_2, \|\cdot\|_{\sup})$ to be the normed linear space of all polynomials whose least degree term is of degree 2. We consider the following linear map

$$\begin{aligned} T : X &\longrightarrow Y \\ p &\longmapsto \int p dx \end{aligned}$$

so that if $p(x) = a_n x^n + \ldots a_1 x$, then $T(p) = \frac{a_n}{n+1} x^{n+1} + \cdots + \frac{a_1}{2} x^2$. We claim that T is bounded. Indeed,

$$\begin{aligned} \|T(p)\| &= \|\frac{a_n}{n+1} x^{n+1} + \cdots + \frac{a_1}{2} x^2\| \\ &= \sup_{x\in[0,1]} \left| \frac{a_n}{n+1} x^{n+1} + \cdots + \frac{a_1}{2} x^2 \right| \\ &= \sup_{x\in[0,1]} \left| x \cdot \left(\frac{a_n}{n+1} x^n + \cdots + \frac{a_1}{2} x \right) \right| \\ &\leq \sup_{x\in[0,1]} |x| \sup_{x\in[0,1]} \left| \left(\frac{a_n}{n+1} x^n + \cdots + \frac{a_1}{2} x \right) \right| \\ &\leq 1 \cdot \sup_{x\in[0,1]} |a_n x^n + \ldots a_1 x| \\ &= \sup_{x\in[0,1]} |p(x)| \\ &= \|p\|. \end{aligned}$$

Thus indeed, T is a bounded linear transformation. We next claim that the following linear transform is an inverse of T:

$$\begin{aligned} U : Y &\longrightarrow X \\ q &\longmapsto q' \end{aligned}$$

so that if $q(x) = a_n x^n + \ldots a_2 x^2$, then $U(q) = n a_n x^{n-1} + \cdots + 2 a_2 x$. Indeed, we see that

$$\begin{aligned} T \circ U(q) &= T\left(n a_n x^{n-1} + \cdots + 2 a_2 x\right) \\ &= n a_n \frac{x^n}{n} + \ldots 2 a_2 \frac{x^2}{2} \\ &= q. \end{aligned}$$

Similarly, for $p(x) = a_n x^n + \ldots a_1 x$, we see that

$$\begin{aligned} U \circ T(p) &= U\left(\frac{a_n}{n+1} x^{n+1} + \cdots + \frac{a_1}{2} x^2\right) \\ &= \frac{a_n}{n+1}(n+1) x^n + \cdots + \frac{a_1}{2}(2) x \\ &= p. \end{aligned}$$

This shows that U is inverse of T. We now show that U is unbounded. Indeed,

$$\begin{aligned}\|U(x^n)\| &= \|nx^{n-1}\| \\ &= \sup_{x\in[0,1]} \left|nx^{n-1}\right| \\ &= n\cdot 1 \\ &= n\cdot \|x^n\|.\end{aligned}$$

This shows that for all $n \geq 2$, there exists $q_n(x) \in Y$ given by $q_n(x) = x^{n+1}$ such that

$$\|U(q_n)\| = n+1 > n = n\|q_n\|,$$

making U unbounded. This completes the proof.

10.5 Miscellaneous applications

Example 10.5.1. We wish to construct an additive function $f : \mathbb{R} \to \mathbb{R}$ which is not continuous. Indeed, consider the Hamel basis of $\mathbb{R}$ over $\mathbb{Q}$ and denote it by $\mathcal{B}$. We know that $\mathcal{B}$ is not finite. Observe that any additive map $f : \mathbb{R} \to \mathbb{R}$ is $\mathbb{Q}$-linear as

$$f\left(\frac{p}{q}x\right) = pf\left(\frac{1}{q}x\right)$$

and since $qf\left(\frac{1}{q}x\right) = f(x)$, thus,

$$f\left(\frac{p}{q}x\right) = \frac{p}{q}f(x).$$

Since any function $\mathcal{B} \to \mathbb{R}$ can be extended $\mathbb{Q}$-linearly to $\mathbb{R} \to \mathbb{R}$, therefore we now construct a function $f : \mathcal{B} \to \mathbb{R}$ and show that its $\mathbb{Q}$-linear extension $\tilde{f} : \mathbb{R} \to \mathbb{R}$ cannot be continuous at 0.

Pick any sequence $(b_n) \subseteq \mathcal{B}$ and consider the following sequence in $\mathbb{R}$

$$x_n = \frac{b_n}{n^{\lceil |b_n| \rceil + n}}$$

where $\lceil z \rceil$ is the ceiling function (smallest integer larger than z). Note that the denominator of x_n is a positive integer. Observe that $x_n \to 0$ as $n \to \infty$.

Define the following function $f : \mathcal{B} \to \mathbb{R}$:

$$f(b) = \begin{cases} n^{\lceil |b_n| \rceil + n} & \text{if } b = b_n \\ 1 & \text{else.} \end{cases}$$

Extend this function to a $\mathbb{Q}$-linear map $\tilde{f} : \mathbb{R} \to \mathbb{R}$, so that it is additive. We claim that $\tilde{f}$ is not continuous at 0. Indeed, we have $x_n \to 0$ as $n \to \infty$, but

$$\tilde{f}(x_n) = \tilde{f}\left(\frac{b_n}{n^{\lceil |b_n| \rceil + n}}\right) = \frac{1}{n^{\lceil |b_n| \rceil + n}}\tilde{f}(b_n) = \frac{1}{n^{\lceil |b_n| \rceil + n}} n^{\lceil |b_n| \rceil + n} = 1$$

and thus $\tilde{f}(x_n) = 1 \not\to \tilde{f}(0) = 0$ as $n \to \infty$, making $\tilde{f}$ discontinuous at 0, as needed.

Proposition 10.5.2. *Let X be a normed linear space over field $\mathbb{K}$ and $T : X \to \mathbb{K}$ be a linear functional. If T is unbounded, then* $\operatorname{Ker}(T) \subseteq X$ *is dense.*

Proof. Since T is unbounded, therefore we first claim that T is unbounded on each $B_{1/n}[0]$. Indeed, if there exists $n_0 \in \mathbb{N}$ such that T is bounded on $B_{1/n_0}[0]$, then for any $x \in X$, we have $\frac{x}{n_0\|x\|} \in B_{1/n_0}[0]$. Thus, by boundedness of T on $B_{1/n_0}[0]$, there exists $K \in \mathbb{R}_{>0}$ such that

$$\left|T\left(\frac{x}{n_0\|x\|}\right)\right| \leq K.$$

By linearity it follows from above that

$$|Tx| \leq Kn_0\|x\|$$

for all $x \in X$. This makes T bounded, a contradiction. Hence T is unbounded on each $B_{1/n}[0]$.

Consequently, for each $n \in \mathbb{N}$, there exists $y_n \in B_{1/n}[0]$ such that $\|Ty_n\| \geq n$. It follows that $y_n \to 0$ as $n \to \infty$ since $y_n \in B_{1/n}[0]$. Further, observe that

$$z_n = \frac{y_n}{Ty_n} - \frac{x}{Tx} \in \operatorname{Ker}(T).$$

Now we claim that $z_n \to \frac{x}{Tx}$ as $n \to \infty$. Indeed, since

$$\|\frac{y_n}{Ty_n}\| = \frac{1}{|Ty_n|}\|y_n\| \leq \frac{1}{n}\|y_n\| < \|y_n\|$$

and since $\|y_n\| \to 0$ as $n \to \infty$, therefore this shows that $\frac{y_n}{Ty_n} \to 0$ as $n \to \infty$. It follows that $z_n \to \frac{x}{Tx}$ as $n \to \infty$, as required.

As $z_n \in \operatorname{Ker}(T)$, therefore $T(x)z_n \in \operatorname{Ker}(T)$ by linearity. Thus $T(x)z_n \to x$ as $n \to \infty$. This shows the density of $\operatorname{Ker}(T)$, thus completing the proof. □

The following is a generalization of Riesz lemma to $r = 1$.

Proposition 10.5.3. *Let X be a normed linear space and $Y \subseteq X$ be a finite dimensional proper linear subspace. Then there exists $x_1 \in S^1(X) = \{x \in X \mid \|x\| = 1\}$ such that*

$$d(x_1, Y) = 1.$$

Proof. Pick $x \in X \setminus Y$. As Y is finite-dimensional in X, therefore it is closed in X. Hence, $d(x, Y) > 0$. We first claim that there exists $\tilde{y} \in Y$ such that

$$d(x, Y) = d(x, \tilde{y}). \tag{10}$$

Indeed, since $d(x, Y) = \inf_{y \in Y} d(x, y) = M$, therefore there exists a sequence $(y_n) \subseteq Y$ such that $d(x, y_n) \to M$ as $n \to \infty$. Fix $\epsilon > 0$. Thus, there exists $N \in \mathbb{N}$ such that $|d(x, y_n) - M| < \epsilon$ for all $n \geq N$. That is, $0 < d(x, y_n) < M + \epsilon$ for all $n \geq N$. Since $g(y) := d(x, y)$ is a continuous map on Y, therefore we have that

$$(y_n)_{n \geq N} \subseteq K = g^{-1}([0, M + \epsilon])$$

where $K \subseteq Y$ is a closed subset of Y. We now claim that K is bounded. Pick $y \in K$. Then

$$\begin{aligned}\|y\| = d(y,0) &\leq d(y,x) + d(x,0) \\ &< M + \epsilon + d(x,0).\end{aligned}$$

This shows that K is bounded. As Y is finite-dimensional normed linear space, therefore generalized Heine-Borel holds and we deduce that K is a compact subset of Y. Since in a metric space compactness is equivalent to sequentially compact, therefore K is sequentially compact. It follows that $(y_n)_{n \geq N} \subseteq K$ has a subsequence which converges, say to $\tilde{y} \in K \subseteq Y$. Replace (y_n) by that subsequence so that we may write $y_n \to \tilde{y}$ and $d(x, y_n) \to M$. By continuity of g, it follows that $g(y_n) = d(x, y_n) \to g(\tilde{y}) = d(x, \tilde{y})$, but $d(x, y_n) \to M$, thus by uniqueness of limits in a Hausdorff space, it follows that $d(x, \tilde{y}) = M$, as needed. This completes the proof of claim in Eqn. (10).

We now complete the proof. Consider the vector

$$x_1 = \frac{x - \tilde{y}}{\|x - \tilde{y}\|} \in X.$$

We claim that $d(x_1, Y) = 1$. Indeed,

$$\begin{aligned}d(x_1, Y) &= \inf_{y \in Y} \left\| \frac{x - \tilde{y}}{\|x - \tilde{y}\|} - y \right\| \\ &= \frac{1}{\|x - \tilde{y}\|} \inf_{y \in Y} \|x - (\tilde{y} + \|x - \tilde{y}\| y)\| \\ &= \frac{1}{\|x - \tilde{y}\|} \inf_{y \in Y} \|x - y\|\end{aligned}$$

where the last equality follows from the bijection provided by affine transformations $Y \to Y$ mapping as $y \mapsto ay + x$ for $a \in \mathbb{K}$ and $x \in Y$, using the linearity of Y. From above equalities, it follows from Eqn. (10) that

$$\begin{aligned}d(x_1, Y) &= \frac{1}{\|x - \tilde{y}\|} \inf_{y \in Y} \|x - y\| \\ &= \frac{1}{\|x - \tilde{y}\|} d(x, Y) \\ &= \frac{1}{d(x, \tilde{y})} d(x, Y) = 1,\end{aligned}$$

as required to complete the proof. □

11 Main theorems of functional analysis

There are four major theorems in basic functional analysis, which we discuss now.

Theorem 11.0.1 (Uniform boundedness principle). *Let X be a Banach space and Y be a normed linear space. Consider a collection of bounded linear transformations $(T_i)_{i \in I} \subseteq B(X, Y)$ such that for each $x \in X$, the subset $(T_i x)_{i \in I} \subseteq Y$ is bounded. Then, $(\|T_i\|)_{i \in I}$ is bounded in $\mathbb{R}$, that is, $(T_i)_{i \in I} \subseteq B(X, Y)$ is a bounded set.*

Theorem 11.0.2 (Open mapping & bounded inverse theorem). *Let X and Y be Banach spaces and $T : X \to Y$ be a surjective bounded linear map. Then,*

1. *T is an open map.*
2. *If T is a bijection, then T is a homeomorphism.*

Theorem 11.0.3 (Closed graph theorem). *Let X and Y be Banach spaces and $T : X \to Y$ be a linear transformation. Then the following are equivalent:*

1. *T is continuous/bounded.*
2. *The graph $\Gamma(T) = \{(x, Tx) \in X \times Y \mid x \in X\}$ is closed in $X \times Y$.*

We now show how all three are equivalent.

Theorem 11.0.4. *Let X and Y be Banach spaces. Then the following implications are true:*

1. *CGT $\implies$ UBP.*
2. *BIT $\implies$ OMP.*
3. *CGT $\implies$ OMP.*

Proof. 1. Closed graph theorem (CGT) states that a linear map $T : X \to Y$ is bounded if and only if $\Gamma(T) = \{(x, Tx) \in X \times Y \mid x \in X\}$ is a closed set in $X \times Y$. We wish to show that uniform boundedness principle (UBP) holds, that is, if $(T_i)_{i\in I}$ is a non-empty collection of bounded linear maps from X to Y such that for each $x \in X$, the set $(T_i(x))_{i\in I} \subseteq Y$ is bounded, then the set $(\|T_i\|)_{i\in I} \subseteq \mathbb{R}$ is a bounded set.

Pick any collection $(T_i)_{i\in I} \subseteq B(X, Y)$ such that for all $x \in X$, there exists $M_x \in \mathbb{R}_+$ such that $\sup_{i\in I} \|T_i x\| \leq M_x$. We wish to show that $(\|T_i\|)_{i\in I}$ is bounded. Indeed, to this end, we contstruct a new norm on X, using which, we will show the above.

Define the following for each $x \in X$:

$$\|x\|_1 := \|x\| + \sup_{i\in I} \|T_i x\|.$$

This is well-defined, as $(T_i x)$ is a bounded set in Y. We now make the following claims:

C1. $(X, \|\cdot\|_1)$ is a normed linear space.
C2. $(X, \|\cdot\|_1)$ is a Banach space.

Assuming the above two claims to be true, let us first show how this will complete the proof. We consider the map

$$\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|_1).$$

We claim that this is a continuous linear transformation. Indeed, by CGT, we need only show that $\Gamma(\mathrm{id})$ is closed. That is, (denote $X_1 = (X, \|\cdot\|_1)$)

$$\Gamma(\mathrm{id}) = \{(x, x) \in X \times X_1 \mid x \in X\} \subseteq X \times X_1$$

is closed. Indeed, consider any sequence $(x_n, x_n) \subseteq \Gamma(\mathrm{id})$ which is convergent in $X \times X_1$. Then suppose $x_n \to x$ in X and $x_n \to x'$ in X_1. We claim that $x = x'$, so that $(x_n, x_n) \to (x, x)$ and since $(x, x) \in \Gamma(\mathrm{id})$, so this will show that $\Gamma(\mathrm{id})$ is closed.

Indeed, we have $x_n \to x$ in X, so $\|x_n - x\| \to 0$ as $n \to \infty$. Similarly, $\|x_n - x'\|_1 \to 0$ as $n \to \infty$. Since

$$\|x_n - x'\|_1 = \|x_n - x'\| + \sup_{i\in I} \|T_i x_n - T_i x'\| \to 0$$

as $n \to \infty$, therefore $\sup_{i\in I} \|T_i x_n - T_i x'\| \to 0$ and $\|x_n - x'\| \to 0$ as well. The latter says that $x_n \to x$ in X. By uniqueness of limits, we conclude that $x = x'$, as required. This shows that $\mathrm{id} : X \to X_1$ is continuous linear transform by CGT, hence bounded.

We wish to bound $\sup_{\|x\|=1} \|T_i x\|$. Pick any $x \in X$ with $\|x\| = 1$. Then we have for each $i \in I$ that

$$\begin{aligned}\|x\|_1 &= \|x\| + \sup_{i\in I} \|T_i x\| \\ &\geq 1 + \|T_i x\|.\end{aligned}$$

Thus, for each $i \in I$, we have

$$\|T_i x\| \leq \|x\|_1 - 1 \leq \|x\|_1.$$

It follows that

$$\sup_{\|x\|=1} \|T_i x\| \leq \sup_{\|x\|=1} \|x\|_1 = \|\mathrm{id}\| < \infty,$$

as required. Hence we now need only prove the claims C1 and C2.

To see claim C1, proceed as follows. Observe that if $\|x\|_1 = 0$, then $\|x\| = 0$, so $x = 0$. Further we have for any $c \in \mathbb{K}$ that $\|cx\|_1 = \|cx\| + \sup_{i\in I} \|T_i(cx)\| = |c|\,\|x\| + |c| \sup_{i\in I} \|T_i x\| = |c|\,\|x\|_1$. Finally, to see triangle inequality, we see that

$$\begin{aligned}\|x + y\|_1 &= \|x + y\| + \sup_{i\in I} \|T_i x + T_i y\| \\ &\leq \|x\| + \|y\| + \sup_{i\in I} (\|T_i x\| + \|T_i y\|) \\ &\leq \|x\| + \|y\| + \sup_{i\in I} \|T_i x\| + \sup_{i\in I} \|T_i y\| \\ &= \|x\|_1 + \|y\|_1,\end{aligned}$$

as required. This shows claim C1.

To see claim C2, proceed as follows. Take any Cauchy sequence $(x_n) \subseteq X_1$. We wish to show that it converges. We claim that (x_n) is Cauchy in X. Indeed, for any $\epsilon > 0$, we have $N \in \mathbb{N}$ such that for any $n, m \geq N$ we have

$$\|x_n - x_m\| \leq \|x_n - x_m\|_1 < \epsilon$$

and for each $j \in I$, we have

$$\|T_j x_n - T_j x_m\| \leq \sup_{i\in I} \|T_i x_n - T_i x_m\| \leq \|x_n - x_m\|_1 < \epsilon/2.$$

Thus, we get by former that (x_n) is Cauchy, so convergent to say $x \in X$. We claim that (x_n) converges to x in X_1. In the latter, by letting $m \to \infty$, we obtain that for each $j \in I$ and each $n \geq N$, we have

$$\|T_j x_n - T_j x\| \leq \epsilon/2 < \epsilon.$$

Thus, taking $\sup_{i\in I}$, we further obtain that for each $n \geq N$ we have

$$\sup_{i\in I} \|T_i x_n - T_i x\| \leq \epsilon/2 < \epsilon.$$

Now, we may write

$$\begin{aligned} \|x_n - x\|_1 &= \|x_n - x\| + \sup_{i\in I} \|T_i x_n - T_i x\| \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

for $n \geq N$, as requird. This completes the proof.

2. Consider any bounded linear map $T : X \to Y$ which is surjective. We wish to show that T is an open mapping using bounded inverse theorem. Indeed, as T is bounded, therefore $\mathrm{Ker}\,(T)$ is a closed linear subspace. Going modulo $\mathrm{Ker}\,(T)$, we get a linear transformation $\tilde{T} : X/\mathrm{Ker}\,(T) \to Y$ such that the following commutes:

$$\begin{tikzcd} X \arrow[r, two heads, "T"] \arrow[d, two heads, "\pi"'] & Y \\ X/\mathrm{Ker}\,(T) \arrow[ur, two heads, "\tilde{T}"'] & \end{tikzcd}$$

We first claim that $\tilde{T}$ is bounded. Indeed, as for any $x + \mathrm{Ker}\,(T) \in X/\mathrm{Ker}\,(T)$ we have $\tilde{T}(x + \mathrm{Ker}\,(T)) = Tx$, therefore

$$\|\tilde{T}(x + \mathrm{Ker}\,(T))\| = \inf_{z\in \mathrm{Ker}(T)} \|T(x+z)\| = \inf_{z\in \mathrm{Ker}(T)} \|Tx\| = \|Tx\|.$$

This shows that $\tilde{T}$ is a bounded linear map which is injective and surjective. Thus, $\tilde{T}$ is a bijection and thus by BIT, we get that $\tilde{T}$ is a homeomorphism. In particular, we see that $\tilde{T}$ is an open map. Now consider the map $\pi : X \to X/\mathrm{Ker}\,(T)$. We wish to show that π is an open map. Let $U \subseteq X$ be an open set and pick any point $x + \mathrm{Ker}\,(T) \in \pi(U) \subseteq X/\mathrm{Ker}\,(T)$ where $x \in U$. As there exists $B_\epsilon(x) \subseteq U$, thus we claim that $B_\epsilon(x + \mathrm{Ker}\,(T)) \subseteq \pi(U)$. Indeed, if $y + \mathrm{Ker}\,(T) \in B_\epsilon(x + \mathrm{Ker}\,(T))$, then $\|x - y + \mathrm{Ker}\,(T)\| < \epsilon$. As

$$\|x - y + \mathrm{Ker}\,(T)\| = \inf_{z\in \mathrm{Ker}(T)} \|x - y + z\| < \epsilon,$$

thus there exists $z \in Z$ such that $\|x - y + z\| < \epsilon$. Thus, $y - z \in B_\epsilon(x) \subseteq U$. Hence, $y - z + \mathrm{Ker}\,(T) = y + \mathrm{Ker}\,(T) \subseteq \pi(U)$, as needed.

3. We first show that closed graph theorem (CGT) implies bounded inverse theorem (BIT). Indeed, this combined with item 2 above will show that CGT $\implies$ OMP. Let $T : X \twoheadrightarrow Y$ be a surjective bounded linear transformation which is a bijection. We then wish to show that the inverse linear transformation of T, $T^{-1} : Y \to X$, is also bounded. By CGT, it is equivalent to showing that the graph $\Gamma(T^{-1}) \subseteq Y \times X$ is a closed set. Since T is a bijection, we get

$$\begin{aligned} \Gamma(T^{-1}) &= \{(y, T^{-1}y) \in Y \times X \mid y \in Y\} \\ &= \{(Tx, x) \in Y \times X \mid x \in X\} \\ &\cong \{(x, Tx) \in X \times Y \mid x \in X\} \end{aligned}$$

where the last homeomorphism is induced by restricting the natural homeomorphism $Y \times X \to X \times Y$. It follows that $\Gamma(T^{-1})$ is closed in $Y \times X$ since $\Gamma(T)$ is closed in $X \times Y$ by CGT (as it is continuous), as required. $\square$

We next see that it is important in closed graph theorem for X and Y to be Banach.

Example 11.0.5. We wish to show that there exists a linear map $T : X \to Y$ where X and Y are normed linear spaces such that T is unbounded and the graph $\Gamma(T) \subseteq X \times Y$ is closed.

Indeed, consider $X = C^1[0,1]^*$ to be the subspace of $C^1[0,1]$ of those functions f such that $f(a) = 0$ and $Y = C[0,1]^*$ both with sup norm. Define

$$\begin{aligned} T : X &\longrightarrow Y \\ f(x) &\longmapsto f'(x) \end{aligned}$$

to be the derivative map. We know that T is unbounded as $f_n(x) = x^n \in C^1[0,1]$ has norm 1 but its derivative has unbounded norm. We wish to show that $\Gamma(T)$ is closed in $X \times Y$. Indeed, consider any sequence $(f_n) \subseteq X$ such that $(f_n, Tf_n) \subseteq \Gamma(T)$ is convergent in $X \times Y$. As projection map are continuous, it follows that $(f_n) \subseteq X$ and $(Tf_n) = (f_n') \subseteq Y$ are convergent. Let $f_n \to f$ in X and $f_n' \to g$ in Y. As X and Y are in sup norm, it follows that $f_n \to f$ and $f_n' \to g$ uniformly. As $f_n(0) = 0$, it follows by the theorem on uniform convergence and derivatives that f_n converges uniformly to a differentiable function which we know is f and $f' = g$. That is $Tf = g$. This shows that $(f_n, Tf_n) \to (f, Tf)$ in $X \times Y$, that is, (f_n, Tf_n) converges in $\Gamma(T)$. This shows that $\Gamma(T)$ is closed. Yet, T is unbounded, as required.

Similarly, the hypothesis of completeness is essential in uniform boudnedness principle.

Example 11.0.6. We wish to show that the hypothesis of completeness of the domain in uniform boundedness principle is essential.

Indeed, let $X = \mathbb{R}^\infty \subseteq (\ell^2, \|\cdot\|_2)$ of all eventually zero sequences in ℓ^2 with the induced norm. Then X is not Banach as $(x_k^{(n)}) = (1, 1/2, \dots, 1/n, 0, \dots)$ is a sequence in X which is Cauchy but it is not convergent. We now construct a sequence of functionals $f_n : X \to \mathbb{K}$ such that for all $(x_k) \in X$, the sequence $(f_n((x_k)))_n$ is bounded in $\mathbb{K}$ but still $(\|f_n\|)_n \subseteq \mathbb{R}$ is unbounded.

Consider

$$\begin{aligned} f_n : X &\longrightarrow \mathbb{K} \\ (x_k) &\longmapsto \sum_{k=1}^{n} x_k. \end{aligned}$$

Pick any $(x_k) \in X$. Then,

$$|f_n((x_k))| = \left|\sum_{k=1}^{n} x_k\right| \leq \left|\sum_{k=1}^{\infty} x_k\right| < \infty$$

as there are only finitely many non-zero elements, thus for each $(x_k) \in X$, $(f_n((x_k)))_n$ is bounded. Moreover,

$$\|f_n\| = \sup_{(x_k) \in X} \frac{|f_n((x_k))|}{\|(x_k)\|} \geq \frac{|\sum_{k=1}^{n} x_k|}{\left(\sum_{k=1}^{\infty} |x_k|^2\right)^{1/2}}$$

for any (x_k) in X. We claim that $\|f_n\| \to \infty$ as $n \to \infty$. Indeed, consider $(x_k^{(n)}) = (1, 1/2, \ldots, 1/n, 0, \ldots)$. Then, $\|(x_k^{(n)})\| = 1 + 1/2^2 + \ldots 1/n^2 < M$ for a fixed $M > 0$ and for all n. Further, by above we have

$$\begin{aligned} \|f_n\| &\geq \frac{|\sum_{k=1}^n 1/k|}{\|(x_k^{(n)})\|} \\ &> \frac{1}{M} \sum_{k=1}^n \frac{1}{k} \to \infty \end{aligned}$$

as $n \to \infty$, as required.

We wish to next prove the main theorems using an important technical lemma.

Theorem 11.0.7 (Zabreiko's lemma). *Let X be a Banach space and $p : X \to \mathbb{R}_{\geq 0}$ be a seminorm. If p is countably subadditive, then p is continuous.*

Proof. Let us first define a seminorm on a Banach space.

Definition 11.0.8 (Seminorm and countably subadditive functions). Let X be a normed linear space. A function $p : X \to \mathbb{R}_{\geq 0}$ is said to be a seminorm if $p(\alpha x) = |\alpha| \, p(x)$ for all $\alpha \in \mathbb{K}$ and $x \in X$ and $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.

The function p is said to be countably subadditive if for every convergent series $\sum_n x_n$ in X, we have

$$p\left(\sum_{n=1}^{\infty} x_n\right) \leq \sum_{n=1}^{\infty} p(x_n)$$

In proving Zabreiko's lemma, we would need a notion of absorbing sets.

Definition 11.0.9 (Absorbing set). Let X be a normed linear space. A subset $A \subseteq X$ is said to be absorbing if for all $x \in X$, there exists $s_x \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_x$.

Note that if A is absorbing, then $-A$ is also absorbing. We now state the following proposition, which will be used in proving Zabreiko's lemma.

Proposition 11.0.10. *Let X be a normed linear space, $p : X \to \mathbb{R}_{\geq 0}$ be a function and $A \subseteq X$.*

1. *If A is absorbing, then $0 \in A$.*
2. *If X is Banach and A is closed convex and absorbing, then A contains a neighborhood of 0.*
3. *If p is a seminorm, then if p is continuous at 0, then p is continuous on X.*

Proof. 1. As A is absorbing, therefore for all $x \in X$, there exists $s_x \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_x$. Let $x = 0$. Then, there exists $s_0 \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_0$. Pick any $t \geq s_0$, we get $0 = ta$ for some $a \in A$. As $t \neq 0$, it follows that $a = 0$, as required.

2. Let $A \subseteq X$ be closed convex and absorbing. Then first observe that

$$D = A \cap (-A) \subseteq A$$

is non-empty as $0 \in A$ (and thus so is in $-A$). We claim that for any $S \subseteq D$, we have

$$\frac{1}{2}S + \frac{1}{2}(-S) \subseteq D.$$

Indeed, pick any $\frac{s_1 - s_2}{2} \in \frac{1}{2}S + \frac{1}{2}(-S)$. We wish to show that $\frac{s_1-s_2}{2} \in A$ and $\frac{s_1-s_2}{2} \in -A$. Thus, we reduce to showing that $\frac{s_1-s_2}{2}, \frac{s_2-s_1}{2} \in A$. It is easy to see that $A \cap -A$ is convex as A and $-A$ are convex. As $s_1, s_2 \in S \subseteq A \cap -A$ thus $-s_1, -s_2 \in S \subseteq A \cap -A$ as well. Now, by convexity of $A \cap -A$, we get

$$\frac{s_1 - s_2}{2}, \frac{s_2 - s_1}{2} \in A \cap -A,$$

as required.

We claim that D° is non-empty. This will complete the proof as by above we will have that $\frac{1}{2}D^\circ + \frac{1}{2}(-D^\circ) \subseteq D$ is open in D and since it contains 0, we would have shown that A contains an open set containing 0.

Suppose to the contrary that $D^\circ = \emptyset$. We wish to derive a contradiction to the fact that A is an absorbing set. Indeed, first observe that for all $n \in \mathbb{N}$, we have $(nD)^\circ = \emptyset$ and nD is closed. This gives us that for each $n \in \mathbb{N}$, the set $Y_n = X - (nD)$ is an open dense subset of X. Pick any $x \in X - D$. As $X - D$ is open, there exists $B_1 = \overline{B_{r_1}(x)} \subseteq X - D$ where $r_1 < 1$. As $X - 2D$ is dense, therefore $(X - 2D) \cap (B_1)^\circ$ is non-empty and thus we get a closed ball B_2 of radius less than $1/2$ in B_1. Continuing this, we have a sequence of closed balls $B_1 \supseteq B_2 \supseteq \dots \supseteq B_n \supseteq \dots$ with radius of B_n less than $1/n$ and $B_n \cap nD = \emptyset$. Let x_n be the center of B_n. We claim that (x_n) is a Cauchy sequence. Indeed, for any $1/k$ we have

$$\|x_n - x_m\| < 2/k$$

for all $n, m \geq k$. As X is complete therefore there exists $x \in X$ such that $x_n \to x$. Thus $x \in B_n$ for all $n \in \mathbb{N}$, that is, $x \notin nD$ for all $n \in \mathbb{N}$. As A is absorbing, therefore there exists $s_x \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_x$. As $-A$ is also absorbing, thus we get $s'_x \in \mathbb{R}_{>0}$ such that $x \in -tA$ for all $t \geq s'_x$. Let $n \in \mathbb{N}$ be larger than both s_x, s'_x. Then we have that $x \in nA$ and $x \in -nA$. It follows that $x \in nA \cap (-nA) = nD$, a contradiction to the fact that $x \notin D$. This completes the proof of item 2.

3. Let $x_n \to x$ in X where $x \neq 0$. We wish to show that $p(x_n) \to p(x)$. Indeed, since $x_n - x \to 0$ and p is continuous at 0, we get $p(x_n - x) \to p(0) = 0$. Thus for any $\epsilon > 0$, we have $p(x_n - x) = |p(x_n - x)| < \epsilon$ for all $n \geq N$. As $p(x_n) - p(x) \leq p(x_n - x)$ by seminorm crieterion, we get $|p(x_n) - p(x)| < \epsilon$ for all $n \geq N$. It follows that $p(x_n) \to p(x)$, as required. □

Using the above proposition, we now prove Zabreiko's lemma.

Proof of Theorem 11.0.7. By Proposition 11.0.10, 3, we reduce to proving that p is continuous at 0. We claim that it is sufficient to show that there is an open ball $B_r(0)$ of radius $r > 0$ at 0 such that $p(B_r(0))$ is a bounded set in $\mathbb{R}_{\geq 0}$. Indeed, for any sequence (x_n) in X converging to 0, which we may assume to be contained in $B_r(0)$, we get that $p(x_n) \in p(B_r(0))$ for all $n \in \mathbb{N}$. We wish to show that $p(x_n) \to p(0) = 0$. Indeed, if $p(B_r(0))$ is upper bounded by $M > 0$, we thus get for any $x \in B_r(0)$ the following bound:

$$p(x) = \|x\| p\left(\frac{x}{\|x\|}\right) \leq M\|x\|.$$

Consequently, we have

$$p(x_n) \leq M\|x_n\|.$$

As $\|x_n\| \to 0$ as $n \to \infty$, it follows by above that $p(x_n) \to 0$ as $n \to \infty$, as required.

So we reduce to showing that there exists an open $B_r(0)$ of 0 in X such that $p(B_r(0))$ is a bounded set. Consider $A_! = \{x \in X \mid p(x) < 1\}$. We claim that A is an absorbing set. Indeed, for any $x \in X$, we have $\|x\|$ such that for all $t \geq \|x\|$ we have $x \in tA_!$ since $p(x/t) = p(x)/t < 1/t$, so $p(t \cdot x/t) < 1$, as required. This shows that $A_!$ is absorbing. We claim that $A = \overline{A_!}$ is absorbing as well. Indeed, observe that since A contains an absorbing set, namely $A_!$, then A is absorbing as well.

We next show that A is convex. Note that since closure of convex set is convex and $A_!$ is convex since if $x, y \in A_!$, then $p((1-t)x + ty) \leq (1-t)p(x) + tp(y) < (1-t) + t = 1$, therefore A is convex. Thus, A is closed convex absorbing set in a Banach space. By Proposition 11.0.10, 2, it follows that A has a neighborhood of 0.

We now find the required ball $B_r(0)$ so that $p(B_r(0))$ is bounded. Indeed consider $r > 0$ such that $\overline{B_r(0)} \subseteq \bar{A}$ and fix a point $x \in B_r(0)$. Pick a point $x_1 \in A$ such that $\|x - x_1\| < r/2$, that is, $x_1 \in B_{r/2}(x) \cap B_r(0) \subseteq \frac{1}{2}A$. Thus $x - x_1 \in \frac{1}{2}B_r(0) \subseteq \frac{1}{2}A \subseteq \frac{1}{2}\bar{A}$. Now there exists $x_2 \in \frac{1}{2}A$ such that $\|x - x_1 - x_2\| \leq r/2^2$, that is, $x_2 \in B_{r/2^2}(x - x_1) \cap B_{r/2}(0) \subseteq \frac{1}{2^2}A$. Continuing this, we get a sequence (x_n) in A such that $x_n \in \frac{1}{2^{n-1}}A$ and $\|x - \sum_{k=1}^n x_k\| < \frac{r}{2^n}$. It follows that $\sum_{k=1}^n x_k \to x$ as $n \to \infty$.

By countable sub-additivity of p, it follows that

$$p(x) = p\left(\sum_{k=1}^{\infty} x_k\right) \leq \sum_{k=1}^{\infty} p\left(x_k\right).$$

As $x_k \in \frac{1}{2^k}A$, therefore $p(x_k) < \frac{1}{2^k}$ by definition of A. Thus, $\sum_{k=1}^{\infty} p(x_k) \leq 1$, and thus $p(x) \leq 1$. As $x \in B_r(0)$ was arbitrary, we have thus shown that $p(B_r(0)) \leq 2$, as required. $\square$

Theorem 11.0.11. *One can derive OMT, UBP, CGT from Zabreiko's lemma (Theorem 11.0.7).*

Proof. (Zabreiko $\Rightarrow$ OMT) Let $T : X \twoheadrightarrow Y$ be a surjective linear transformation between Banach spaces. By translation and scaling homeomorphism, we reduce to showing that $T(B_1(0))$ is open. Define

$$\begin{aligned} p : Y &\longrightarrow \mathbb{R}_{\geq 0} \\ y &\longmapsto \inf\{\|x\| \mid Tx = y\}. \end{aligned}$$

We claim that p is a countably subadditive semi-norm, so that by Theorem 11.0.7, we will get p is continuous. This is sufficient as

$$T(B_1(0)) = p^{-1}([0, 1))$$

which is easy to see. So we reduce to showing that p is a countably subadditive seminorm.

1. p is countably subadditive : Let $\sum_n y_n$ be a covergent series in Y. We wish to show that $p\left(\sum_n y_n\right) \leq \sum_n p(y_n)$. Indeed, fix $\epsilon > 0$. We get the following
$$p(y_n) + \frac{\epsilon}{2^n} \geq \|x_n\|$$
for each $n \in \mathbb{N}$ where $x_n \in X$ is such that $Tx_n = y_n$. Summing till N we get
$$\sum_{n=1}^{N} p(y_n) + \sum_{n=1}^{N} \frac{\epsilon}{2^n} \geq \sum_{n=1}^{N} \|x_n\| \geq \|\sum_{n=1}^{N} x_n\|$$
and since $T(x_1 + \cdots + x_n) = \sum_{n=1}^{N} y_n$, we get that $\|x_1 + \cdots + x_n\| \geq p(\sum_{n=1}^{N} y_n)$. This yields that
$$\sum_{n=1}^{N} p(y_n) + \sum_{n=1}^{N} \frac{\epsilon}{2^n} \geq p\left(\sum_{n=1}^{N} y_n\right).$$
Taking $N \to \infty$ and then $\epsilon \to 0$, the result follows.
2. p is a seminorm : Fact that $p(cy) = |c|\, y$ is immediate from definition. Subadditivity follows from item 1.

(Zabreiko $\Rightarrow$ UBP) Let X, Y be Banach and $(T_i)_{i \in I} \subseteq B(X, Y)$ be a family of bounded linear transformations such that for all $x \in X$, the set $(T_i(x))_{i \in I} \subseteq Y$ is bounded. We wish to show that $(\|T_i\|)_{i \in I}$ is bounded in $\mathbb{R}$.

Consider
$$\begin{aligned} p : X &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \sup_{i \in I} \|T_i(x)\|. \end{aligned}$$
We claim that p is a countably subadditive seminorm. Indeed, then it would follow by Theorem 11.0.7 that p is continuous. Then there exists $\delta > 0$ such that $\|x\| < \delta$ implies $|p(x)| \leq 1$. As p is a seminorm, therefore we would obtain
$$\|x\| < 1 \implies p(x) < 1/\delta.$$
As $\|T_i\| = \sup_{\|x\|<1} \|T_i x\|$ and $p(x) < 1/\delta$ for $\|x\| < 1$ where
$$p(x) = \sup_{i \in I} \|T_i x\| < 1/\delta$$
therefore $\|T_i x\| < 1/\delta$ for all $\|x\| < 1$, which would thus tield $\|T_i\| \leq 1/\delta$, as required. So we reduce to showing that p is a countably subadditive seminorm.

1. p is countably subadditive : Let $\sum_n x_n$ be a convergent series in X. We wish to show that $p\left(\sum_n x_n\right) \leq \sum_n p(x_n)$. Indeed, we have
$$p\left(\sum_n x_n\right) = \sup_{i \in I} \|T_i\left(\sum_n x_n\right)\| = \sup_{i \in I} \|\sum_n T_i x_n\| \leq \sup_{i \in I} \sum_n \|T_i x_n\| \leq \sum_n \sup_{i \in I} \|T_i x_n\| = \sum_n p(x_n)$$
where $\sup_{i \in I} \|T_i x_n\|$ exists and is bounded as by hypothesis, the set $(T_i x)_{i \in I}$ is bounded for any $x \in X$. This shows that p is countably subadditive.

2. p is a seminorm : Fact that $p(cy) = |c|\, y$ is immediate from definition. Subadditivity follows from item 1.

(Zabreiko $\Rightarrow$ CGT) Let $T : X \to Y$ be a linear transformation between Banach spaces. We wish to show that T is bounded if and only if $\Gamma(T) \subseteq X \times Y$ is closed.

($\Rightarrow$) is immediate by considering the inverse image at 0 of $X \times Y \to Y$ of $(x, y) \mapsto Tx - y$.

($\Leftarrow$) Consider the following function

$$\begin{aligned} p : X &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \|Tx\|. \end{aligned}$$

We claim that p is a countably subadditive seminorm. Indeed, this would imply that p is continuous by Theorem 11.0.7. Note that it is sufficient to show that $\{\|Tx\| \mid \|x\| < 1\}$ is bounded. But this set is same as $p(B_1(0))$. Thus, we reduce to showing that $p(B_1(0))$ is bounded. Indeed, this follows as there exists $\delta > 0$ such that

$$\|x\| < \delta \implies p(x) < 1$$

which by seminorm property is equivalent to

$$\|x\| < 1 \implies p(x) < 1/\delta.$$

This shows that $p(B_1(0)) < 1/\delta$, as needed. We thus reduce to showing that p is a countably subadditive seminorm.

1. p is countably subadditive : Let $\sum_n x_n$ be a convergent series in X. We wish to show that $p\left(\sum_n x_n\right) \leq \sum_n p(x_n)$. Indeed, we have

$$p\left(\sum_n x_n\right) = \|T\left(\sum_n x_n\right)\|$$

   where since $(\sum_{k=1}^n x_k, \sum_{k=1}^n Tx_k)$ is in the graph and is convergent where graph is closed, therefore $T(\sum_n x_n) = \sum_n Tx_n$. Thus,

$$\|T\left(\sum_n x_n\right)\| = \|\sum_n Tx_n\| \leq \sum_n \|Tx_n\| = \sum_n p(x_n).$$

   This shows that p is countably subadditive.
2. p is a seminorm : Fact that $p(cy) = |c|\, y$ is immediate from definition. Subadditivity follows from item 1.

This completes the proof of Theorem 11.0.7. □

This completes the proof. □

12 Strong & weak convergence

These are important definitions as these protray that how fundamental importance this topic gives to functionals, anyways, its *functional* analysis so we must be very comfortable with constructing and manipulating functionals on a normed linear space.

Definition 12.0.1 (**Strong & weak convergence**)**.** Let X be a normed linear space and $(x_n) \subseteq X$ be a sequence in X. Then, (x_n) is said to be strongly convergent if there exists $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$. Further (x_n) is said to be weakly convergent if there exists $x \in X$ such that for all functionals $f \in X^*$, the sequence $(f(x_n)) \to f(x)$ in $\mathbb{K}$. In the former case x is said to be the strong limit and in the latter case x is said to be the weak limit.

The following showcases a nice property of weak convergence.

Proposition 12.0.2. *Let X be a normed linear space and $x_n \to x$ weakly in X. Then*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

Proof. As $f(x_n) \to f(x)$ for all $f \in X^*$, therefore we will construct a functional using Hahn-Banach through which the desrired inequality is straightforward. Indeed, by separation theorem applied on point x, we get that there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$. Consequntly, we get by weak convergence that

$$f(x_n) \to f(x) = \|x\|.$$

Now, for each $n \in \mathbb{N}$ we have

$$|f(x_n)| \leq \|f\|\|x_n\| = \|x_n\|.$$

Taking liminf both sides, we obtain

$$\liminf_{n\to\infty} |f(x_n)| \leq \liminf_{n\to\infty} \|x_n\|.$$

As $f(x_n) \to f(x)$, therefore $\liminf_{n\to\infty} |f(x_n)| = |f(x)| = \|x\|$. Thus we get

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|,$$

as required. □

Definition 12.0.3 (**Weakly Cauchy and complete**)**.** A normed linear space X is *weakly complete* if every weakly Cauchy sequence is weakly convergent, where a sequence (x_n) in X is weakly Cauchy if for all $f \in X^*$, the sequence $(f(x_n))$ is Cauchy. Thus, unravelling this, we have that X is weakly complete if for any sequence (x_n) in X such that $(f(x_n))$ is Cauchy in $\mathbb{K}$ for each $f \in X^*$, there exists $x \in X$ such that $f(x_n) \to f(x)$ for each $f \in X^*$.

Proposition 12.0.4. *Any reflexive normed linear space X is weakly complete.*

Proof. Recall X is reflexive if the James map $\mathrm{ev} : X \to X^{**}$ is surjective. Since we have seen that ev is an isometric embedding, therefore reflexivity tells us ev is an isometric isomorphism.

To show that X is weakly complete, pick any weakly Cauchy sequence (x_n) in X. Then, for each $f \in X^*$, the sequence $f(x_n)$ is Cauchy in $\mathbb{K}$. As $\mathbb{K}$ is complete, it follows that $f(x_n)$ converges and let $f(x_n) \to c_f$ where $c_f \in \mathbb{K}$. We claim that the mapping

$$\begin{aligned} \varphi : X^* &\longrightarrow \mathbb{K} \\ f &\longmapsto c_f \end{aligned}$$

is a bounded linear map. This will complete the proof as by reflexivity we will have a unique $x \in X$ such that $\mathrm{ev}_x = \varphi$ and thus $\mathrm{ev}_x(f) = f(x) = c_f = \varphi(f)$, that is,

$$f(x) = \varprojlim_{n\to\infty} f(x_n)$$

for all $f \in X^*$, which shows that (x_n) weakly convergent, as required. We thus reduce to proving that φ is a bounded linear map.

To see linearity, pick any $f, g \in X^*$ and $\alpha \in \mathbb{K}$ to observe that

$$\varphi(f + \alpha g) = c_{f+\alpha g} = \varprojlim_n (f + \alpha g)(x_n) = \varprojlim_n f(x_n) + \alpha \varprojlim_n g(x_n) = c_f + \alpha c_g$$

since each $f(x_n)$ and $g(x_n)$ converges because they are Cauchy. To see boundedness, we first show that the set $\{x_n\} \subseteq X$ is a bounded set. Indeed, by a corollary of uniform boundedness principle we have that a set $Y \subseteq X$ is bounded if and only if $f(Y) \subseteq \mathbb{K}$ is bounded for each $f \in X^*$. For $Y = \{x_n\}$ and any $f \in X^*$, we see that $f(Y) = (f(x_n))$ is bounded as $f(x_n) \to c_f$. It follows that $\{x_n\}$ is a bounded set, as required. Consequently, let $\|x_n\| \le M$ for all $n \in \mathbb{N}$. We thus have

$$|\varphi(f)| = |c_f| = \varprojlim_n |f(x_n)| \le \limsup_n \|f\|\|x_n\| \le \|f\| \cdot M.$$

Hence, φ is a bounded linear map, as required. $\square$