Contents

1 Measures **1**
- 1.1 Semifinite Measures . . . 1
- 1.2 Complete Measures . . . 3
- 1.3 Outer Measures . . . 6
- 1.4 Carathéodory's Extension Theorem . . . 13
- 1.5 Dynkin's Uniqueness Theorem . . . 19

2 Point Set Topology **23**
- 2.1 Continuous Maps . . . 23
- 2.2 LCH Spaces . . . 26
- 2.3 σ-Compact LCH Spaces . . . 29

3 Functional Analysis **33**
- 3.1 Topological Vector Spaces . . . 33
- 3.2 Affine Hyperplanes . . . 40
- 3.3 Hahn-Banach Theorem . . . 42
- 3.4 Hilbert Spaces . . . 48
- 3.5 Differential Calculus . . . 53
- 3.6 Banach Algebra . . . 58
- 3.7 Inverse Function Theorem . . . 62

4 Manifold Theory **65**
- 4.1 Vector Bundles . . . 65

5 Tricks **69**
- 5.1 Upper and Lower Approximations . . . 69
- 5.2 Method of Successive Approximations . . . 70
- 5.3 Fixed Point Problem . . . 71

Chapter 1

Measures

Definition 1.1 (Measure)**.** *Let X be a set and $\mathcal{M}$ be a σ-algebra over X. A* ***measure*** *on $\mathcal{M}$ is a function $\mu : \mathcal{M} \to [0, \infty]$ satisfying*

1. $\mu(\{\}) = 0$
2. *If $\{E_j\}_1^\infty$ is a countable sequence of disjoint sets in $\mathcal{M}$, then* $\mu\left(\bigcup_{j=1}^\infty E_j\right) = \sum_{j=1}^\infty \mu(E_j)$.

If μ does not satisfy countable additivity, but does satisfy finite additivity, then μ is a ***finitely additive*** *measure.*

1.1 Semifinite Measures

Definition 1.2 (Semifinite Measure)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. μ is* ***semifinite*** *if*

$$\forall E \in \mathcal{M} : \mu(E) = \infty, \exists F \subset E : 0 < \mu(F) < \infty$$

Theorem 1.1. *Let μ be a semifinite measure and $E \in \mathcal{M}$, define*

$$\mathcal{F}_E = \{B \subset E : \mu(B) < \infty\}$$

then $\sup_{F \in \mathcal{F}(E)} \mu(F) = \mu(E)$.

Proof. If $\mu(E) < \infty$, then E itself satisfies the criterion.

Now suppose that $\mu(E) = \infty$. Suppose for the sake of contradiction that $M = \sup_{F \in \mathcal{F}(E)} \mu(F) < \infty$. Let $\{F_n\}_1^\infty$ such that $\mu(F_n) \nearrow M$, then

$$\mu\left(\bigcup_{n \in \mathbb{N}} F_n\right) = \lim_{n \to \infty} \mu\left(\bigcup_{k=1}^n F_n\right) = M$$

Let $F = \bigcup_{n\in\mathbb{N}} F_n$, then since $\mu(E) = \infty$, $\mu(E \setminus F) = \infty$, and there exists $G \subset E \setminus F : 0 < \mu(G) < \infty$. This gives $F \cup G \subset E$ and $M < \mu(F \cup G) < \infty$, which contradicts the fact that M is the supremum. □

Definition 1.3 (Semifinite Part). *Let $(X, \mathcal{M}, \mu)$ be a measure space and let*

$$\mu_0(E) = \sup \{\mu(F) : F \subset E, \mu(F) < \infty\}$$

*then μ_0 is a semifinite measure, known as the **semifinite part** of μ.*

Proof. **Basic Properties**

Let $E \in \mathcal{M}$ such that $\mu(E) < \infty$, then $\mu_0(E) = \mu(E)$.

Proof. By monotonicity, $\mu(F) \leq \mu(E)$ for all $F \subset E$. This means that $\mu_0(E) \leq \mu(E)$. Since E itself is finite, $\mu_0(E) \geq \mu(E)$.

Measure

μ_0 is a measure.

Proof. Firstly, $\mu_0(\emptyset) = \mu(\emptyset) = 0$.

Now, let $\{E_n\}_1^\infty$ be a sequence of disjoint sets. If there exists $n \in \mathbb{N}$ such that $\mu_0(E_n) = \infty$, then

$$\{F \subset E_n : \mu(F) < \infty\} \subset \left\{F \subset \bigcup_{n\in\mathbb{N}} E_n : \mu(F) < \infty\right\}$$

Therefore $\mu_0(E) \geq \mu_0(E_n) = \infty$, and $\mu_0(E) = \infty$.

Now suppose that $\mu_0(E_n) < \infty$ for all $n \in \mathbb{N}$ and $\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) < \infty$. Let $\varepsilon > 0$ and $F_n \subset E_n$ such that $\mu(F_n) > \mu_0(E_n) - \varepsilon/2^n$. This means that

$$\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) \geq \mu\left(\bigcup_{k\leq n} F_k\right) = \sum_{k\leq n} \mu(F_k) \quad \forall n \in \mathbb{N}$$

$$\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) \geq \sum_{n\in\mathbb{N}} \mu(F_n) > \sum_{n\in\mathbb{N}} [\mu_0(E_n) - \varepsilon/2^n]$$

$$\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) \geq \sum_{n\in\mathbb{N}} \mu_0(E_n) - \varepsilon$$

Since the above applies to all $\varepsilon > 0$, $\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) \geq \sum_{n\in\mathbb{N}} \mu_0(E_n)$.

Let $F \subset \bigcup_{n\in\mathbb{N}} E_n$ such that $\mu(F) < \infty$, then

$$\mu(F) \leq \sum_{n\in\mathbb{N}} \mu(F \cap E_n) \leq \sum_{n\in\mathbb{N}} \mu_0(E_n)$$

Since this applies to all $F \subset \bigcup_{n\in\mathbb{N}} E_n$, $\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) \leq \sum_{n\in\mathbb{N}} \mu_0(E_n)$.

Finally, suppose that $\mu_0(E_n) < \infty$ for all $n \in \mathbb{N}$, and $\mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \infty$. Let $\alpha \in \mathbb{R}$ and $F \subset \bigcup_{n\in\mathbb{N}} E_n$ such that $\alpha < \mu(F) < \infty$, then

$$\sum_{n\in\mathbb{N}} \mu_0(E_n \cap F) \geq \sum_{n\in\mathbb{N}} \mu(E_n \cap F) = \mu(F) > \alpha$$

Therefore $\sum_{n\in\mathbb{N}} \mu_0(E_n \cap F) = \mu_0\left(\bigcup_{n\in\mathbb{N}} E_n\right)$.

Semifinite

μ_0 is semifinite.

Proof. Let $E \in \mathcal{M} : \mu_0(E) \neq 0$, then

$$\mu_0(E) = \sup\{\mu(F) : F \subset E, \mu(F) < \infty\}$$

There exists $F \subset E$ such that $0 < \mu(F) < \infty$. □

1.2 Complete Measures

Definition 1.4 (Complete Measure)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space, then μ is **complete** if it is defined for all $\mathcal{P}(E)$ where $\mu(E) = 0$ (E is a null set).*

Theorem 1.2 (Completion of Measures)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$ be the collection of all null sets in $\mathcal{M}$, and*

$$\overline{\mathcal{M}} = \{E \cup F : E \in \mathcal{M}, F \subset N \text{ where } N \in \mathcal{N}\}$$

*Then $\overline{\mathcal{M}}$ is a sigma algebra, and there is a unique extension of $\overline{\mu}$ of μ to a **complete** measure on $\overline{\mathcal{M}}$.*

Proof.

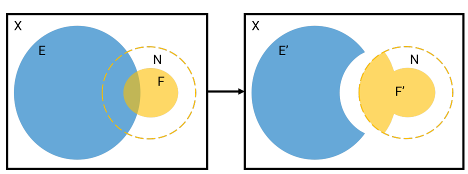

Figure 1.1: Disjoint Representation of Sets in $\overline{\mathcal{M}}$

Let $E \cup F \in \overline{\mathcal{M}}$ where $E \in \mathcal{M}$ is a measurable section, and $F \subseteq N \in \mathcal{N}$ is a "small" section. Rewrite $E \cup F = E' \cup F'$ where

$$E' = E \setminus N \quad F' = F \cup (E \cap N)$$

Then $E \cup F = E' \cup F'$ is a disjoint union. We can now decompose any element $S \in \overline{\mathcal{M}}$ of into a measurable part and a small part that are disjoint, and have the measurable E represent S.

For countable union, we can simply do the union by component

$$\bigcup_{i\in I}(E_i \cup F_i) = \left(\bigcup_{i\in I} E_i\right) \cup \left(\bigcup_{i\in I} F_i\right) \in \overline{\mathcal{M}}$$

where $\bigcup_{i\in I} E_i \in \mathcal{M}$ *and each* $F_i \subseteq N_i \in \mathcal{N}$*, then*

$$\bigcup_{i\in I} F_i \subseteq \bigcup_{i\in I} N_i \in \mathcal{N}$$

since a countable union of null sets is a null set, the small component remains small.

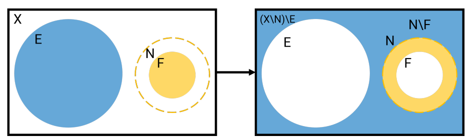

Figure 1.2: Complement of a Set in $\overline{\mathcal{M}}$

For complement, take any $E \cup F \in \overline{\mathcal{M}}$ *and assume that they are disjoint. Split* $X = N \cup (X \setminus N)$*, then*

$$\begin{aligned}(E \cup F)^c &= [N \cup (X \setminus N)] \setminus (E \cup F) \\ &= [N \setminus (E \cup F)] \cup [(X \setminus N) \setminus (E \cup F)] \\ &= [N \setminus F] \cup [(X \setminus N) \setminus E]\end{aligned}$$

where $N \setminus F \subseteq N$ *is still a small set, and* $(X \setminus N) \setminus E = N^c \cap E^c$ *is a measurable set.*

So we have $\overline{\mathcal{M}}$ *being a* σ*-algebra.*

Let $\overline{\mu} : \overline{\mathcal{M}} \to [0, \infty]$, $\overline{\mu}(E \cup F) = \mu(E)$*, and assume that* E *and* F *are disjoint. While* E *may be different based on the choice of representative,* $\mu(E)$ *is always the same. To see this, let* $S \in \mathcal{M}$,

$$S = E_1 \cup F_1 = E_2 \cup F_2 \quad F_1 \subseteq N_1, F_2 \subseteq N_2$$

be two different decompositions.

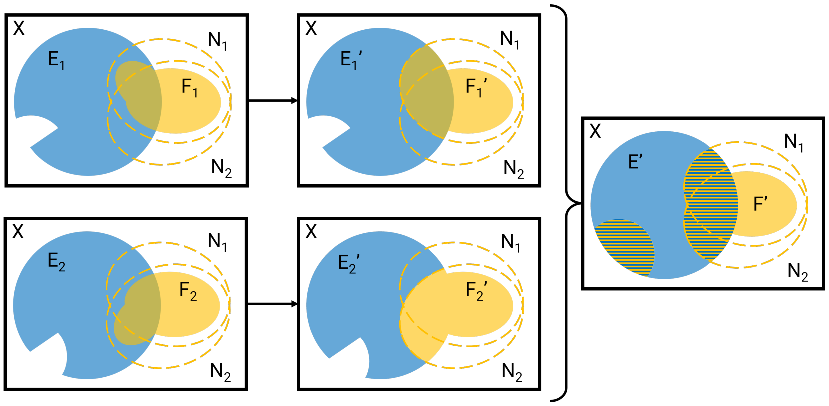

Figure 1.3: Different Representations Differ by a Null Set

Then the difference between E_1' *and* E_2' *must lie within* $N_1 \cup N_2$ *(so it's also a null set). Let* $x \in E_1'$ *but* $x \notin E_2'$*, then since* $E_1' \cup F_1' = E_2' \cup F_2'$*,* $x \in F_2' \subseteq N_2$*. Let* $x \in E_2'$ *but* $x \notin E_1'$*, then since* $E_1' \cup F_1' = E_2' \cup F_2'$*,* $x \in F_1' \subseteq N_1$*. So*

$$E_1' \setminus E_2' \subseteq N_2 \quad E_2' \setminus E_1' \subseteq N_1$$

So the choice of E *shouldn't change the value of the measure.*

$$\begin{aligned} E_1 \cup F_1 &= [(E_1 \cap E_2) \cup (E_1 \setminus E_2)] \cup F_1 \\ &= (E_1 \cap E_2) \cup (E_1 \setminus E_2) \cup F_1 \\ E_2 \cup F_2 &= [(E_2 \cap E_1) \cup (E_2 \setminus E_1)] \cup F_2 \\ &= (E_1 \cap E_2) \cup (E_2 \setminus E_1) \cup F_2 \end{aligned}$$

where $(E_1 \setminus E_2) \cup F_1, (E_2 \setminus E_1) \cup F_2 \subseteq N_1 \cup N_2$*, and*

$$\overline{\mu}(E_1 \cup F_1) = \mu(E_1 \cap E_2) = \overline{\mu}(E_2 \cup F_2)$$

therefore $\overline{\mu}$ *is well-defined.*

Since $\overline{\mu}$ *inherits properties from* μ *on its values on the measurable sets* $E \in \mathcal{M}$*, and that its behaviour is uniquely determined by the measurable component of each* $E \cup F$ *(it just ignores the small sets), it remains a measure by """project-ing""" everything in* $\overline{\mathcal{M}}$ *down to its* E *component.*

Now, $\overline{\mu}(\emptyset) = \mu(\emptyset) = 0$, and let $E_i \cup F_i$ be a countable family of sets in $\overline{\mathcal{M}}$, then

$$\begin{aligned}
\overline{\mu}\left(\bigcup_{i\in I} E_i \cup \bigcup_{i\in I} F_i\right) &= \overline{\mu}\left(\bigcup_{i\in I} E_i \cup \bigcup_{i\in I} F_i\right) \\
&= \mu\left(\bigcup_{i\in I} E_i\right) \\
&= \sum_{i\in I} \mu(E_i) \\
&= \sum_{i\in I} \overline{\mu}(E_i \cup F_i)
\end{aligned}$$

and $\overline{\mu}$ is a measure.

Let $E \cup F \in \overline{\mathcal{M}}$ be a null set and $F \subseteq N \in \mathcal{N}$, then $\overline{\mu}(E \cup F) = \mu(E) = 0$, and $E \in \mathcal{N}$ is also a null set. This means that $\mathcal{P}(E \cup N) \subseteq \overline{\mathcal{M}}$ and $\overline{\mu}$'s domain includes all subsets of its null sets. We achieved this by adding all small sets from $\mathcal{M}$ to $\overline{\mathcal{M}}$.

Finally, since the small component F of each $E \cup F \in \overline{\mathcal{M}}$ is a subset of a null set of μ, any extension μ' of μ into $\overline{\mathcal{M}}$ must have $0 = \mu'(N) \geq \mu'(F) = 0$. Extensions of μ into $\overline{\mathcal{M}}$ must entirely depend on the μ-measurable E component, and must agree with μ on them. Assuming that $E \cap F = \emptyset$,

$$\mu'(F) = 0 \Rightarrow \mu'(E \cup F) = \mu'(E) + \mu'(F) = \mu'(E) = \mu(E)$$

then $\mu' = \overline{\mu}$, and the extended measure is unique.

1.3 Outer Measures

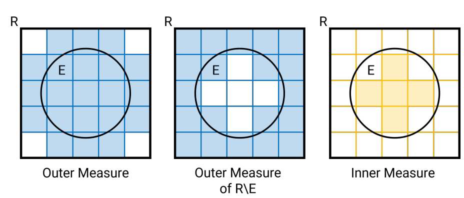

Figure 1.4: Outer and Inner Measures

To calculate the area of a specific region without a handy formula, we can cover it with shapes whose area we do know, from the inside and from the outside to approximate the actual area. With finer and finer approximations, we can take the area of the shape to be its outer and inner measure, if they are equal.

Definition 1.5 (Outer Measure)**.** *An **outer measure** on a non-empty set X is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ that satisfies:*

1. $\mu^*(\emptyset) = 0$
2. $F \subseteq E \Rightarrow \mu^*(F) \leq \mu^*(E)$
3. $\mu^*\left(\bigcup_{i\in I} E_i\right) \leq \sum_{i\in I} \mu^*(A_i), |I| \leq |\mathbb{N}|$

Outer measures approximate sets from above. See Upper and Lower Approximations for more details on the general idea.

Theorem 1.3 (Outer Approximation)**.** *Let $\mathcal{E} \subseteq \mathcal{P}(X)$ be a family of sets with $\emptyset \in \mathcal{E}$ and $X \in \mathcal{E}$, and $\rho : \mathcal{E} \to [0, \infty]$ be a notion of measure with $\rho(\emptyset) = 0$. Define for any $A \subseteq X$,*

$$\mu^*(A) = \inf\left\{\sum_{i\in I} \rho(E_i) : E_i \in \mathcal{E}, A \subseteq \bigcup_{i\in I} E_i\right\}$$

where I is a countable index set. Then μ^* is an outer measure.

Proof. Firstly, the set that we take the infimum from is non-empty, as for $E_1 = X, E_k = \emptyset \forall k > 1$ forms an elementary cover for any $A \in \mathcal{P}(X)$. So when $\mu^*(A) = \infty$, it comes from our notion of measure ρ, and not a lack of an elementary cover for A.

Now, let $A \subseteq B$ with $A, B \in \mathcal{P}(X)$, then any elementary cover of B also covers A, giving

$$\left\{I : E_i \in \mathcal{E}, A \subseteq \bigcup_{i\in I} E_i\right\} \supseteq \left\{I : E_i \in \mathcal{E}, B \subseteq \bigcup_{i\in I} E_i\right\}$$
$$\inf\left\{\sum_{i\in I} \rho(E_i) : A \subseteq \bigcup_{i\in I} E_i\right\} \leq \inf\left\{\sum_{i\in I} \rho(E_i) : B \subseteq \bigcup_{i\in I} E_i\right\}$$
$$\mu^*(A) \leq \mu^*(B)$$

Finally, let $\{A_j\}_1^\infty \subseteq \mathcal{P}(X)$ be a countable sequence of sets, and $\{E_i\}_{i\in I_j} \subseteq \mathcal{E}$ be an elementary cover of A_j, where $\bigcup_{i\in I_j} E_i \supseteq A_j$. This provides a way of writing A as an elementary cover from the elementary covers of each A_j:

$$A = \bigcup_{j\in\mathbb{N}} A_j \subseteq \bigcup_{j\in\mathbb{N}} \bigcup_{i\in I_j} E_i \quad E_i \in \mathcal{E}$$

Let $K = \bigcup_{j\in\mathbb{N}} I_j$, then K is countable, and

$$K \in \left\{I : A \subseteq \bigcup_{i\in I} E_i\right\}$$

$\bigcup_{k\in K} E_k$ is a countable cover of A with elementary sets, meaning that

$$\mu^*(A) \leq \sum_{k\in K} \rho(E_k) \leq \sum_{j\in\mathbb{N}} \sum_{i\in I_j} \rho(E_i)$$

for any choice of elementary covers for each A_j.

This allows us to construct a sequence as follows. Firstly, since each $\mu^*(A_j)$ is an infimum, we can approach each one of them as a sequence from above, where for each $x_{j,k}$ (j-th sequence, k-th entry)

$$\exists I_{j,k} : x_{j,k} = \sum_{i \in I_{j,k}} \rho(E_i)$$

and $\lim_{k\to\infty} x_{j,k} = \mu^*(A_j)$. This yields a countable family of sequences $\left\{ \{x_k\}_j \right\}$ where each $\{x_k\}_j$ approaches $\mu^*(A_j)$ from above.

Now, since for any sequence in the family, we can take entries that are arbitrarily close to the infimum

$$\forall \varepsilon_n > 0, j \in \mathbb{N}, \exists k_n \in \mathbb{N} : x_{j,k_n} < \mu^*(A_j) + \varepsilon_n 2^{-j}$$

and collect them such that their sum is arbitrarily close to the sum of the infimums. Using this, we can set $\varepsilon_n = 1/n$ and form a new sequence:

$$\begin{aligned} x_n = \sum_{j\in\mathbb{N}} x_{j,k_n} &< \sum_{j\in\mathbb{N}} \mu^*(A_j) + \varepsilon_n 2^{-j} \\ &= \sum_{j\in\mathbb{N}} \mu^*(A_j) + 2\varepsilon_n \end{aligned}$$

that approach $\sum_{j\in\mathbb{N}} \mu^*(A_j)$ from above. Since each x_n represents a countable elementary cover of A, each entry pushes $\mu^*(A)$ closer and closer to $\sum_{j\in\mathbb{N}} \mu^*(A_j)$, making it such that $\mu^*(A) \leq \sum_{j\in\mathbb{N}} \mu^*(A_j)$ at the limit. □

Definition 1.6 (Outer Measurable). *For any set $E \subseteq X$, use A to cut E into two pieces, one inside and one outside of A.*

If outer-measuring the pieces individually yields the same measure as outer-measuring the entire set, then the "borders" of A is "compatible" with the outer measure: the distribution of mass created by the cut can be approached by the outer measure.

If the outer measure uses an elementary collection of sets with measure to approximate the measure of a given set, then the distribution of mass around the "borders" of a measurable set should be able to be approached by those elementary sets.

*We formalise the idea as follows: Let X be a set and $\mu^* : \mathcal{P}(X) \to [0,\infty]$ be an outer measure on X. Then a set $A \subseteq X$ is **μ^*-measurable** if*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \forall E \in \mathcal{P}(X)$$

Theorem 1.4 (Carathéodory). *If μ^* is an outer measure on X, then the collection $\mathcal{M}$ of μ^* measurable sets is a sigma algebra, and the restriction of μ^* to $\mathcal{M}$ is a complete measure.*

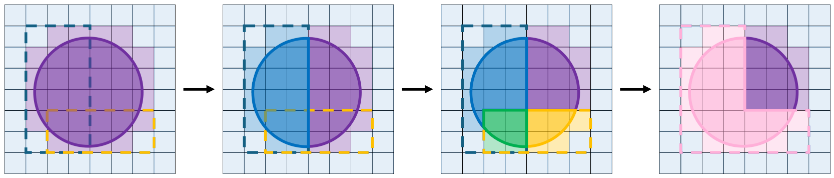

Figure 1.5: Finite Unions of Measurable Sets are Measurable

Proof. A set is μ^* measurable if the distribution of mass around its *border* can be approached by μ^*. For any $A \in \mathcal{M}$, the *cuts* made by A is the same as the cuts made by its complement:

$$\begin{aligned} A \in \mathcal{M} \Rightarrow \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \qquad \forall E \in \mathcal{P}(X) \\ &= \mu^*(E \cap A^c) + \mu^*(E \cap A^{cc}) \\ &\Rightarrow A^c \in \mathcal{M} \end{aligned}$$

It doesn't matter which side of the border is filled, so $\mathcal{M}$ is closed under complements.

Now, let $A, B \in \mathcal{M}$, then the border of $A \cup B$ is just a combination of the old ones. Cutting a set with $A \cup B$ is equivalent to using some part of A's border, and part of B's. Since each part of the new border is approachable by μ^*, the entire border is also approachable.

To see this, let $A, B \in \mathcal{M}$ and $E \in \mathcal{P}(X)$. First cut E with A, which

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

cleanly distributes the mass into two pieces. Then cut each piece with B:

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

which cleanly distributes the mass into four pieces in total, yielding

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

Since the cuts made by A and B are clean, we can "re-stitch" parts of them together along those cuts as follows.

First mend the blue and green piece together with the cut made by B.

$$\begin{aligned} A \cup B \supseteq A \Rightarrow [E \cap (A \cup B)] \cap A &= E \cap A \\ \mu^*([E \cap (A \cup B)] \cap A) &= \mu^*(E \cap A) \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) \end{aligned}$$

"Reformat" $E \cap A^c \cap B$ like so to isolate the cut made by A:

$$\begin{aligned}
[E \cap (A \cup B)] \cap A^c &= E \cap [(A \cup B) \cap A^c] \\
&= E \cap [(A \cap A^c) \cup (B \cap A^c)] \\
&= E \cap B \cap A^c \\
\mu^*([E \cap (A \cup B)] \cap A^c) &= \mu^*(E \cap B \cap A^c)
\end{aligned}$$

Then combine the two pieces like so by mending the cut made by A:

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

We "reformat" $E \cap A^c \cap B^c$ to combine the cuts made by A and B:

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

And finally, we combine $E \cap (A \cup B)$ and $E \cap (A \cup B)^c$ to find that

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

the cuts made by $A \cup B$ distribute mass in a way that can be approached by the outer measure, meaning that $A \cup B \in \mathcal{M}$. $\mathcal{M}$ is closed under finite union, and is an algebra.

Since the borders of μ^*-measurable sets cleanly distributes mass along them, there is no ambiguity when the borders are close to each other, as they "pull" the approached mass towards the inside of their sets. When two μ^* measurable sets share a border, that border also cleanly splits the mass between the two.

To see this, let $A, B \in \mathcal{M}, A \cap B$ and consider $\mu^*(A \cup B)$. We use A to cut $A \cup B$ back to A and B to obtain

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

Since we only used A to make the cut, this works even if $B \notin \mathcal{M}$. So μ^* is finitely additive on $\mathcal{M}$.

Now, since μ^* works by approaching the mass of sets from the outside, all sets in $\mathcal{M}$ are "limit points" of μ^*'s approximations. As we have all the "limit points", $\mathcal{M}$ should be "closed", and for any sequence of sets in $\mathcal{M}$, their "limit" can also be approached by μ^*.

To see this, let $\{A_j\}_1^\infty \subseteq \mathcal{M}$ be a sequence of pairwise disjoint sets.

$$\begin{aligned}
\mu^*\left(\bigcup_{i\in\mathbb{N}} A_i\right) &\geq \mu^*\left(\bigcup_{i\in I} A_i\right) \quad \forall I \subset \mathbb{N}, |I| < |\mathbb{N}| \\
&= \sum_{i\in I} \mu^*(A_i) \quad \forall I \subset \mathbb{N}, |I| < |\mathbb{N}| \\
\mu^*\left(\bigcup_{i\in\mathbb{N}} A_i\right) &\geq \sup\left\{\sum_{i\in I} \mu^*(A_i) : I \subset \mathbb{N}, |I| < |\mathbb{N}|\right\} \\
\mu^*\left(\bigcup_{i\in\mathbb{N}} A_i\right) &= \sum_{i\in\mathbb{N}} \mu^*(A_i)
\end{aligned}$$

The expanding borders of $\bigcup_{i\in I} A_i$ approaches $\bigcup_{i\in\mathbb{N}} A_i$, allowing μ^* to also approach the union. So μ^* is countably additive for sets in $\mathcal{M}$.

Now, let $E \in \mathcal{P}(X)$, $A = \bigcup_{i\in\mathbb{N}} A_i$, and $I \subset \mathbb{N}, |I| < |\mathbb{N}|$. Then we can cut E with $\bigcup_{i\in I} A_i$.

$$\begin{aligned}
\mu^*(E) = \mu^*&\left(E \cap \bigcup_{i\in I} A_i\right) \\
&+ \mu^*\left(E \cap \left(\bigcup_{i\in I} A_i\right)^c\right)
\end{aligned}$$

We can cut the left part with each A_i as

$$\begin{aligned}
&\mu^*\left(E \cap \bigcup_{i\in I} A_i\right) \\
&= \mu^*\left(E \cap \bigcup_{i\in I} A_i \cap A_{i_1}\right) + \mu^*\left(E \cap \bigcup_{i\in I} A_i \cap A_{i_1}^c\right) \\
&= \mu^*(E \cap A_{i_1}) + \mu^*\left(E \cap \bigcup_{i\in I, i\neq i_1} A_i\right) \\
&\vdots \\
&= \sum_{i\in I} \mu^*(E \cap A_i)
\end{aligned}$$

into $|I|$ disjoint pieces, preserving the total mass with the clean cuts of each A_i.

We now relax the borders of the complement and let $E \cap \bigcup_{i\in I} A_i$ approach

$E \cap \bigcup_{i\in\mathbb{N}} A_i$ from below.

$$\begin{aligned}\mu^*(E) &= \mu^*\left(E \cap \bigcup_{i\in I} A_i\right) + \mu^*\left(E \cap \left(\bigcup_{i\in I} A_i\right)^c\right)\\ &\geq \mu^*\left(E \cap \bigcup_{i\in I} A_i\right) + \mu^*\left(E \cap \left(\bigcup_{i\in \mathbb{N}} A_i\right)^c\right)\\ &= \sum_{i\in I} \mu^*(E \cap A_i) + \mu^*\left(E \cap \left(\bigcup_{i\in \mathbb{N}} A_i\right)^c\right)\end{aligned}$$

Mending the cuts made by each A_i yields a subset of the entire union. Since the cuts of each A_i cleanly distributes the mass among them, there is no double-counting and the mass can never exceed that of the union.

$$\sum_{i\in I} \mu^*(E \cap A_i) = \mu^*\left(E \cap \bigcup_{i\in I} A_i\right) \leq \mu^*\left(E \cap \bigcup_{i\in \mathbb{N}} A_i\right)$$
$$\sum_{i\in \mathbb{N}} \mu^*(E \cap A_i) \leq \mu^*\left(E \cap \bigcup_{i\in \mathbb{N}} A_i\right)$$

And in this process, we also cover the entirety of the union, so the measured mass can also be no less than the mass of the union.

$$\sum_{i\in \mathbb{N}} \mu^*(E \cap A_i) = \mu^*\left(E \cap \bigcup_{i\in \mathbb{N}} A_i\right)$$

which yields

$$\begin{aligned}\mu^*(E) &\geq \sum_{i\in \mathbb{N}} \mu^*(E \cap A_i) + \mu^*\left(E \cap \left(\bigcup_{i\in \mathbb{N}} A_i\right)^c\right)\\ &= \mu^*\left(E \cap \bigcup_{i\in \mathbb{N}} A_i\right) + \mu^*\left(E \cap \left(\bigcup_{i\in \mathbb{N}} A_i\right)^c\right)\\ &\geq \mu^*(E)\end{aligned}$$

and

$$\mu^*(E) = \mu^*\left(E \cap \bigcup_{i\in \mathbb{N}} A_i\right) + \mu^*\left(E \cap \left(\bigcup_{i\in \mathbb{N}}\right)^c\right)$$

As the distribution of mass around border of the union is approached by finite unions of $\{A_i\}$, it ended up also being one approachable by μ^*. So $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{M}$ and $\mathcal{M}$ is a σ-algebra, and μ^* is a measure on $\mathcal{M}$.

Finally, let $N \in \mathcal{M} : \mu^*(N) = 0$, then the distribution of mass around its borders can be approached by μ^*. However, the lack of mass inside the borders

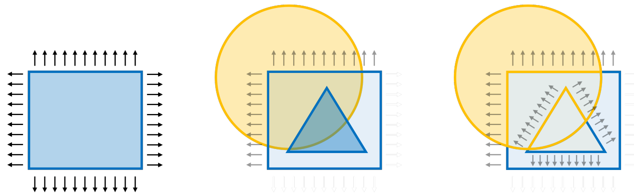

Figure 1.6: Outer Measure is Complete

of N means that there is no mass to "leak" into N. Any subsets of N *simply do not have any mass to distribute*, they just push all of the mass outside.

To see this, let $N \in \mathcal{M}$, $F \subset N$ and $E \in \mathcal{P}(X)$, then we first cut E with N

$$\mu^*(E) = \mu^*(E \cap N) + \mu^*(E \cap N^c)$$

and find that $\mu^*(E \cap N) \leq \mu^*(N) = 0$. Now use F to cut $E \cap N$ to obtain

$$\mu^*(E \cap N) = \mu^*(E \cap N \cap F) + \mu^*(E \cap N \cap F^c) = 0$$

as there is no mass to distribute. We mend the cut made by N to obtain

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap N) + \mu^*(E \cap N^c) \\
&= \mu^*(E \cap N \cap F) + \mu^*(E \cap N \cap F^c) + \mu^*(E \cap N^c) \\
&= \mu^*(E \cap F) + \mu^*(E \cap N \cap F^c) + \mu^*(E \cap N^c \cap F^c) \\
&= \mu^*(E \cap F) + \mu^*(E \cap F^c)
\end{aligned}$$

and $F \in \mathcal{M}$ is μ^* measurable. Since all subsets of null sets are measurable, $\mu^*|_{\mathcal{M}}$ is a complete measure. □

1.4 Carathéodory's Extension Theorem

Definition 1.7 (Elementary Family)**.** *Let X be a set, then $\mathcal{E} \subseteq \mathcal{P}(X)$ is an* ***elementary family*** *if*

1. $\emptyset \in \mathcal{E}$
2. $E, F \in \mathcal{E} \Rightarrow E \cap F \in \mathcal{E}$ *(closed under finite intersection)*
3. $E \in \mathcal{E}$ *implies that* E^c *is a finite disjoint union of members of* $\mathcal{E}$*.*

Theorem 1.5. *If $\mathcal{E}$ is an elementary family, then the collection*

$$\mathcal{A} = \left\{ \bigcup_{i=1}^{n} E_i : E_i \in \mathcal{E}, n \in \mathbb{N}, \{E_i\}_1^n \text{ disjoint} \right\}$$

of finite disjoint unions of members of $\mathcal{E}$ is an algebra.

Proof. Let $A, B \in \mathcal{E}$, then $B^c = \bigcup_{i=1}^n B_i$ is a finite disjoint union (FDU) of elements of $\mathcal{E}$, and since $\mathcal{E}$ is closed under intersections,

$$A \setminus B = A \cap B^c = A \cap \bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A \cap B_i$$

is a FDU of elementary sets. Since $A \setminus B \cup B$ is also a disjoint union, $A \setminus B, A \cup B \in \mathcal{A}$. Therefore $\bigcup_{i=1}^n E_i \in \mathcal{A}$ for any $E_i \in \mathcal{E}, n \in \mathbb{N}$.

Let $A = \bigcup_{i=1}^m A_i, B = \bigcup_{i=1}^n B_i \in \mathcal{A}$, then $A \cup B$ is a union of elementary sets, and therefore is also in $\mathcal{A}$.

Let $E = \bigcup_{i=1}^m E_i$, then since $E_i = \bigcup_{j=1}^n E_{i,j}$ is also a FDU of elementary sets.

$$E^c = \bigcap_{i=1}^m E_i^c = \bigcap_{i=1}^m \bigcup_{j=1}^n E_{i,j}$$

then applying the following identity

$$\left(\bigcup_{i=1}^m A_i\right) \cap \left(\bigcup_{j=1}^n B_j\right) = \bigcup_{i=1}^m \left[A_i \cap \bigcup_{j=1}^n B_j\right] = \bigcup_{i=1}^m \bigcup_{j=1}^n A_i \cap B_j$$

repeatedly yields

$$\begin{aligned} E^c &= \bigcap_{i=1}^m \bigcup_{j=1}^n E_{i,j} \\ &= \bigcup_{i_1=1}^n \cdots \bigcup_{i_m=1}^n E_{1,i_1} \cap \cdots \cap E_{m,i_m} \\ &= \bigcup_{\vec{i} \in [1,n]^m} \bigcap_{j=1}^m E_{j,i_j} \end{aligned}$$

Since $\mathcal{E}$ is closed under finite intersections, each $\bigcap_{j=1}^m E_{j,i_j} \in \mathcal{E}$, and E^c is a FDU of $\mathcal{E}$, and $E^c \in \mathcal{A}$. ☐

Definition 1.8. *Let X be a set and $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, then a function $\mu_0 : \mathcal{A} \to [0, \infty]$ is a **premeasure** if*

1. $\mu_0(\emptyset) = 0$
2. *If $\{A_i\}_1^\infty$ is a sequence of disjoint sets in $\mathcal{A}$ where $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$, then $\mu_0\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \mu_0(A_i)$.*

In other words, μ_0 is a finitely-additive measure on $\mathcal{A}$, with the occasional countable additivity when the union happens to also be in $\mathcal{A}$.

Theorem 1.6. *Let X be a set, $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, and μ_0 be a premeasure on $\mathcal{A}$, then μ_0 induces an outer measure*

$$\mu^*(A) = \inf \left\{ \sum_{i \in I} \mu_0(E_i) : E_i \in \mathcal{A}, \bigcup_{i \in I} E_i \supseteq A \right\}$$

on X.

Proof. Since μ_0 defines a notion of measure on a family of sets $\mathcal{E} = \mathcal{A}$, the outer approximation theorem applies. μ^* acts as an "least upper bound" on the mass of A, taking the infimum of all possible upper bounds on its mass.

Theorem 1.7. *Let X be a set, $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, μ_0 be a premeasure on $\mathcal{A}$, and μ^* be the induced outer measure, then*

1. *$\mu^*|\mathcal{A} = \mu_0$.*
2. *Every set in $\mathcal{A}$ is outer measurable.*

Proof. The first part essentially means that the elementary sets $\mathcal{A}$ and the notion of measure μ_0 is one that "makes sense". Since $\mathcal{A}$ is an algebra and μ_0 is a premeasure, there are no ways to chop any elementary sets into smaller elementary sets and somehow "lose" mass in the process.

To begin, each elementary set $E \in \mathcal{A}$ can simply be approximated by itself from the outside, giving $\mu^*(E) \leq \mu_0(E)$.

Let $\{E_i\}_1^\infty$ be an elementary cover of $E \in \mathcal{A}$. Since each part of the cover is in $\mathcal{A}$, we can remove any overlap. With E also being in $\mathcal{A}$, we can remove any "over-covering". So the only parts relevant to the outer measure is the tightest part of it, or simply its way of cutting E into disjoint pieces.

Now, let $\{E_i\}_1^\infty$ be a sequence of sets such that $\bigcup_{i \in \mathbb{N}} E_i \supseteq E$. Take $F_1 = E_1$ and $F_k = E \cap \left(E_k \setminus \bigcup_{i=1}^{k-1} E_i \right)$, then $\{F_i\}_1^\infty$ is a *disjoint* sequence of sets in $\mathcal{A}$ where $\bigcup_{i \in \mathbb{N}} F_i = E$. However, as μ_0 is a premeasure and each $F_i \in \mathcal{A}$, the mass of E is the combination of the masses of each F_i in the first place, so there is no mass lost with this partition:

$$\mu_0(E) = \sum_{i \in \mathbb{N}} \mu_0(F_i) \leq \sum_{i \in \mathbb{N}} \mu_0(E_i) \Rightarrow \mu^*(E) \geq \mu_0(E)$$

Therefore $\mu^*(E) = \mu_0(E)$ for any $E \in \mathcal{A}$, and $\mu^*|\mathcal{A} = \mu_0$.

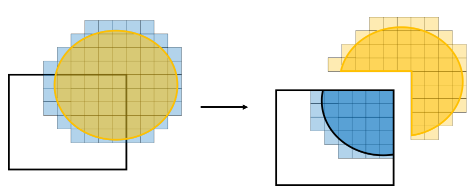

Figure 1.7: Sets in the algebra are outer-measurable

A set is outer-measurable if the distribution of mass along its borders is approachable by μ^*. Since for any elementary set $E \in \mathcal{A}$, μ^* uses E itself to approach it, the distribution of mass along the border is already precisely measured. E itself can be used to cut the approximating sets cleanly, which then separates the approximation into two parts with no ambiguity along its border.

To see this, let $A \in \mathcal{P}(X)$, $E \in \mathcal{A}$ and, $\{E_i\}_1^\infty$ be an elementary cover of A. Then we can take

$$I_i = E \cap E_i \quad O_i = E^c \cap E_i$$

then $\{I_i\}_1^\infty$ is an elementary cover of $E \cap A$ and $\{O_i\}_1^\infty$ is an elementary cover of $E^c \cap A$, and together they cover A.

$$(E \cap A) \cup (E^c \cap A) = \left(\bigcup_{i \in \mathbb{N}} E \cap E_i \right) \cup \left(\bigcup_{i \in \mathbb{N}} E^c \cap A \right)$$

Since $\{I_i\}_1^\infty$ and $\{O_i\}_1^\infty$ are elementary covers, we have

$$\mu^*(E \cap A) \leq \sum_{i \in \mathbb{N}} \mu_0(I_i)$$
$$\mu^*(E^c \cap A) \leq \sum_{i \in \mathbb{N}} \mu_0(O_i)$$

Moreover, $E_i = I_i \cup O_i$ is a disjoint union, so

$$\mu_0(E_i) = \mu_0(I_i) + \mu_0(O_i)$$

and we have

$$\begin{aligned} \mu^*(E \cap A) + \mu^*(E^c \cap A) &\leq \sum_{i \in \mathbb{N}} \mu_0(I_i) + \sum_{i \in \mathbb{N}} \mu_0(O_i) \\ &= \sum_{i \in \mathbb{N}} \mu_0(E_i) \end{aligned}$$

meaning that $\mu^*(E \cap A) + \mu^*(E^c \cap A)$ is a lower bound for

$$\left\{ \sum_{j \in J} \mu_0(E_j) : |J| = |\mathbb{N}|, E_j \in \mathcal{A}, \bigcup_{j \in J} \mu_0(E_j) \supseteq A \right\}$$

and

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

Therefore

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

and $E \in \mathcal{M}$ is μ^* measurable. □

Theorem 1.8 (Carathéodory's Extension Theorem)**.** *Let X be a set, $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, μ_0 be a premeasure on $\mathcal{A}$, and $\mathcal{M} = \mathcal{M}(\mathcal{A})$ be the sigma algebra generated by $\mathcal{A}$.*

Then there exists a measure on $\mathcal{M}$ whose restriction to $\mathcal{A}$ is μ_0, being the induced outer measure $\mu = \mu^*$.

If v is another measure on $\mathcal{M}$ that extends μ_0, then $v(E) \leq \mu(E)$ for any $E \in \mathcal{M}$, with equality when $\mu(E) < \infty$. If μ_0 is σ-finite, then μ is the unique extension of μ_0 to a measure on $\mathcal{M}$.

Proof. By Carathéodory's Theorem, the induced outer measure μ^* has a collection of measurable sets M being a sigma algebra. Since $\mathcal{A} \subseteq M$ are all outer measurable, $M \supseteq \mathcal{M}(\mathcal{A})$ contains a "spanning set" of $\mathcal{M}(\mathcal{A})$, and therefore contains $\mathcal{M} = \mathcal{M}(\mathcal{A})$. As shown earlier, the restriction of μ^* to $\mathcal{A}$ is equal to μ_0.

Let v be another measure on $\mathcal{M}$ that extends μ_0, $E \in \mathcal{M}$, and $\{E_i\}_1^\infty \supseteq E$ be an elementary cover of E. Then

$$v(E) = v\left(\bigcup_{i\in\mathbb{N}} E_i\right) \leq \sum_{i\in\mathbb{N}} v(E_i) = \sum_{i\in\mathbb{N}} \mu_0(E_i)$$

and $v(E)$ is a lower bound for

$$\left\{\sum_{i\in I} \mu_0(E_i) : |I| = |\mathbb{N}|, E_i \in \mathcal{A}, \bigcup_{i\in\mathbb{N}} E_i \supseteq E\right\}$$

and we have $v(E) \leq \mu^*(E)$ since $\mu^*(E)$ is the greatest lower bound.

We take advantage of the fact that v is also an extension of μ_0 by taking limits on sets that they agree on. Let $\{E_i\}_1^\infty$ be an elementary cover of E, and take

$$A_j = \bigcup_{i=1}^{j} E_i \quad A_j \in \mathcal{A}$$

then

$$v(A_j) = \mu^*(A_j) = \mu_0(A_j) \quad \forall j \in \mathbb{N}$$

and using continuity from below,

$$v\left(\bigcup_{i\in\mathbb{N}} E_i\right) = \lim_{i\to\infty} v(A_i) = \lim_{i\to\infty} \mu^*(A_i) = \mu^*\left(\bigcup_{i\in\mathbb{N}} E_i\right)$$

Suppose that $\mu^*(E) < \infty$, then since μ^* is the infimum,

$$\forall\varepsilon > 0, \exists I : |I| = |\mathbb{N}|, E_i \in \mathcal{A}, \bigcup_{i\in I} E_i \supseteq E, \sum_{i\in I} \mu_0(E_i) < \mu^*(E) + \varepsilon$$

Take $\{E_i\}_1^\infty$ such that

$$\mu^*(E) + \varepsilon > \sum_{i\in\mathbb{N}} \mu_0(E_i) \geq \mu^*\left(\bigcup_{i\in\mathbb{N}} E_i\right) = v\left(\bigcup_{i\in\mathbb{N}} E_i\right) \geq \mu^*(E)$$

with the inequality coming from subadditivity and monotonicity. Decompose

$$\mu^*\left(\bigcup_{i\in\mathbb{N}} E_i\right) = \mu^*(E) + \mu^*\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right) < \mu^*(E) + \varepsilon$$

$$\mu^*\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right) < \varepsilon$$

$$v\left(\bigcup_{i\in\mathbb{N}} E_i\right) = v(E) + v\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right)$$

$$v\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right) \leq \mu^*\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right) < \varepsilon$$

and we obtain

$$\mu^*(E) \leq v\left(\bigcup_{i\in\mathbb{N}} E_i\right) = v(E) + v\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right) < v(E) + \varepsilon$$

Rearranging the equation, and since we can find an elementary cover for any $\varepsilon > 0$,

$$\mu^*(E) - \varepsilon < \mu^*(E) - v\left(\bigcup_{i\in\mathbb{N}} E_i \setminus E\right) < v(E) \leq \mu^*(E)$$

we can approach $v(E)$ from below by arbitrarily squishing the overcover. This is possible because $\mathcal{M}$ is a sigma algebra generated by $\mathcal{A}$, and allows us to approach E and its measure using just the elementary sets.

Let $\varepsilon_n = \frac{1}{n}$ and I_n be such that $\bigcup_{i\in I_n} E_i \supseteq E, E_i \in \mathcal{A}$ and $\sum_{i\in I_n} \mu_0(E_i) < \mu^*(E) + \varepsilon_n$. This gives

$$x_n = \mu^*(E) - v\left(\bigcup_{i\in I_n} E_i \setminus E\right) > \mu^*(E) - \frac{1}{n}$$

then

$$x_n < v(E) \leq \mu^*(E) \forall n \in \mathbb{N} \quad \lim_{n\to\infty} x_n = \mu^*(E)$$

which gives

$$v(E) = \mu^*(E)$$

Suppose that μ^* is σ-finite, then we can decompose

$$X = \bigcup_{j\in\mathbb{N}} E_j \quad \mu^*(E_j) < \infty$$

further into

$$F_j = E_j \setminus \bigcup_{i=1}^{j-1} E_i \quad \bigcup_{i\in\mathbb{N}} F_j = X \quad v(E_j) \leq \mu^*(E_j) < \infty$$

and the F_js form a disjoint union.

Then taking any $E \in \mathcal{M}$,

$$\begin{aligned}
\mu^*(E) &= \mu^*\left(\bigcup_{i\in\mathbb{N}} E \cap F_i\right) \\
&= \sum_{i\in\mathbb{N}} \mu^*(E \cap F_i) \\
&= \sum_{i\in\mathbb{N}} v(E \cap F_i) \\
&= v\left(\bigcup_{i\in\mathbb{N}} E \cap F_i\right) \\
&= v(E)
\end{aligned}$$

and $\mu^* = v$ is unique. □

1.5 Dynkin's Uniqueness Theorem

Dynkin's uniqueness theorem is very useful for situations involving product measures and probabilistic independence. Proof taken from Professor Addario-Berry's notes.

Definition 1.9 (λ-System)**.** *Let Ω be a set and $\mathcal{A} \subset \mathcal{P}(\Omega)$ be a collection of subsets. We say that $\mathcal{A}$ is a λ-system if*

1. *The whole space $\Omega \in \mathcal{A}$.*
2. *For any $E, F \subset \mathcal{A}$ where $E \subset F$, we have $F \setminus E \in \mathcal{A}$.*
3. *For any increasing sequence $\{A_n\}_1^\infty \subset \mathcal{A}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.*

Theorem 1.9. *Let Ω be a ground set. If $\mathcal{A}$ is a π-system and a λ-system, then $\mathcal{A}$ is a σ-algebra.*

Proof. Since $\Omega \in \mathcal{A}$, $\mathcal{A}$ is closed under complements.

If a λ-system is closed under finite unions, then it is also closed under countable unions, so we can reduce this to simply finite unions. To this end, just use complements to write unions as intersections. For any $E, F \in \mathcal{A}$,

$$E \cup F = (E \cup F)^{cc} = (E^c \cap F^c)^c$$

□

Lemma 1.10. *Let $\mathcal{P}$ be a π-system, and*

$$\mathcal{M} = \{E \in \lambda(\mathcal{P}) : E \cap F \in \lambda(\mathcal{P}) \quad \forall F \in \mathcal{P}\}$$

be the collection of cooperative sets, then $\mathcal{N}$ is a λ-system with $\mathcal{N} = \lambda(\mathcal{P})$.

Proof. Let $E, F \in \mathcal{M}$ such that $F \supset E$, and $G \in \lambda(\mathcal{P})$, then

$$(F \setminus E) \cap G = \underbrace{(F \cap G)}_{\in \lambda(\mathcal{P})} \setminus \underbrace{(E \cap G)}_{\in \lambda(\mathcal{P})} \in \lambda(\mathcal{P})$$

and $F \setminus E \in \mathcal{M}$.

Let $\{E_n\}_1^\infty \subset \mathcal{M}$, then

$$\left(\bigcup_{n \in \mathbb{N}} E_n\right) \cap G = \bigcup_{n \in \mathbb{N}} \underbrace{(E_n \cap G)}_{\in \lambda(\mathcal{P})} \in \lambda(\mathcal{P})$$

and $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{M}$. Therefore $\mathcal{M}$ is a λ-system.

Since $\mathcal{P}$ is a π-system, $\mathcal{P} \subset \mathcal{M}$. As $\mathcal{M}$ is a λ-system containing $\mathcal{P}$, $\mathcal{M} = \lambda(\mathcal{P})$. □

Lemma 1.11. *Let $\mathcal{P}$ be a π-system, and*

$$\mathcal{N} = \{E \in \lambda(\mathcal{P}) : E \cap F \in \lambda(\mathcal{P}) \quad \forall F \in \lambda(\mathcal{P})\}$$

be the collection of helpful sets, then $\mathcal{N}$ is a λ-system with $\mathcal{N} = \lambda(\mathcal{P})$.

Proof. Let $E \in \mathcal{P}$, then since $\mathcal{M} = \lambda(\mathcal{P})$, $E \cap F \in \lambda(\mathcal{P})$ for all $F \in \mathcal{M} = \lambda(\mathcal{P})$. Therefore $E \in \mathcal{N}$.

By similar arguments as above, $\mathcal{N}$ is a λ-system.

Since $\mathcal{N} = \lambda(\mathcal{P})$, we have $E \cap F \in \lambda(\mathcal{P})$ for all $E, F \in \lambda(\mathcal{P})$, making it a π-system. □

Theorem 1.12 (Dynkin's π-System Lemma)**.** *Let $\mathcal{P}$ be a π-system over a ground set Ω, then*

$$\sigma(\mathcal{P}) = \bigcap_{\mathcal{F} \supset P : \mathcal{F} \ \sigma\text{-field}} \mathcal{F} = \bigcap_{\mathcal{F} \supset \mathcal{P} : \mathcal{F} \ \lambda\text{-system}} \mathcal{F} = \lambda(\mathcal{P})$$

In other words, the σ-algebra generated by $\mathcal{P}$ is the same as the λ-system generated by $\mathcal{P}$.

Proof Outline. Since every σ-field is a λ-system, $\sigma(\mathcal{P}) \supset \lambda(\mathcal{P})$, so it's sufficient to show that $\lambda(\mathcal{P})$ is a σ-field as well. As $\lambda(\mathcal{P})$ is a λ-system and a π-system, it is a σ-field. □

Theorem 1.13 (Dynkin's Uniqueness Theorem). *Let* $(X, \mathcal{F})$ *be a measurable space and* $\mathcal{P} \subset \mathcal{M}$ *be a* π*-system with* $\sigma(\mathcal{P}) = \mathcal{F}$. *Let* μ_1, μ_2 *be measures on* $\mathcal{F}$, *and suppose that 1.* $\mu_1(E) = \mu_2(E)$ *for all* $E \in \mathcal{P}$. *2. There exists a sequence* $\{X_n\}_1^\infty \subset \mathcal{P}$ *such that* $X_n \nearrow X$ *and* $\mu_1(X_n) < \infty$ *for all* $n \in \mathbb{N}$.

Then $\mu_1 = \mu_2$.

Proof. Let $G \in \mathcal{P}$ with $\mu_1(G) < \infty$, and let

$$\mathcal{A} = \{E \in \mathcal{F} : \mu_1(E \cap G) = \mu_2(E \cap G)\}$$

then $\mathcal{A} \supset \mathcal{P}$ with $\Omega \in \mathcal{A}$, and we claim that $\mathcal{A}$ is a lambda-system.

Let $\{E_n\}_1^\infty$ be an increasing sequence of sets in $\mathcal{A}$, then by continuity from below, $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{A}$. Let $E, F \in \mathcal{A}$ with $F \supset E$, then

$$\begin{aligned}
\mu_1(G \cap (F \setminus E)) &= \mu_1(G \cap F) - \mu_1(G \cap E) \\
&= \mu_2(G \cap F) - \mu_2(G \cap E) \\
&= \mu_2(G \cap (F \setminus E))
\end{aligned}$$

and $F \setminus E \in \mathcal{A}$. Then $\mathcal{A}$ is a λ-system. By Dynkin's π-System Lemma, $\mathcal{A} = \mathcal{F}$.

Let $E \in \mathcal{F}$, and $\{F_n\}_1^\infty \subset \mathcal{P}$ such that $\mu_1(F_n) = \mu_2(F_n) < \infty$ for all n, and $F_n \nearrow \Omega$. Now by continuity from below on both measures,

$$\mu_1(E) = \lim_{n \to \infty} \mu_1(E \cap F_n) = \lim_{n \to \infty} \mu_2(E \cap F_n) = \mu_2(E)$$

$\square$

Chapter 2

Point Set Topology

Definition 2.1 (Topology)**.** *Let X be a non-empty set of points. A family of subsets $\mathcal{T} \subseteq \mathcal{P}(X)$ forms a **topology** on X if:*

1. $\emptyset, X \in \mathcal{T}$
2. *Arbitrary unions of elements of ct is another element of $\mathcal{T}$.*
3. *Any finite intersection in $\mathcal{T}$ is another element of $\mathcal{T}$.*

*The pair $(X, \mathcal{T})$ is called a **topological space**, and the elements of $\mathcal{T}$ are the open sets of $\mathcal{T}$.*

2.1 Continuous Maps

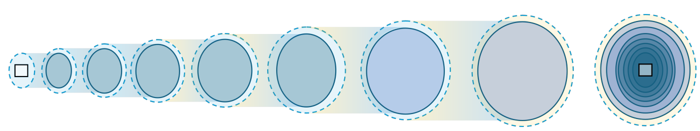

Figure 2.1: Nested Open Sets That Mirror ℝ's Topological Structure

Theorem 2.1. *Let $(X, \mathcal{T})$ be a normal space, and $A, B \subset X$ be closed sets. Let*

$$\Delta = \left\{ \frac{k}{2^n} : n \in \mathbb{N}, 0 < k < 2^n \right\}$$

be the collection of dyadic rational numbers in $(0, 1)$, then there is a family of open sets

$$\mathcal{F} = \{U_r : r \in \Delta\}$$

such that $A \subset U_r \subset B^c$ for all $r \in \Delta$, and that the closure $\overline{U_r} \subset U_s$ for all $r < s$.

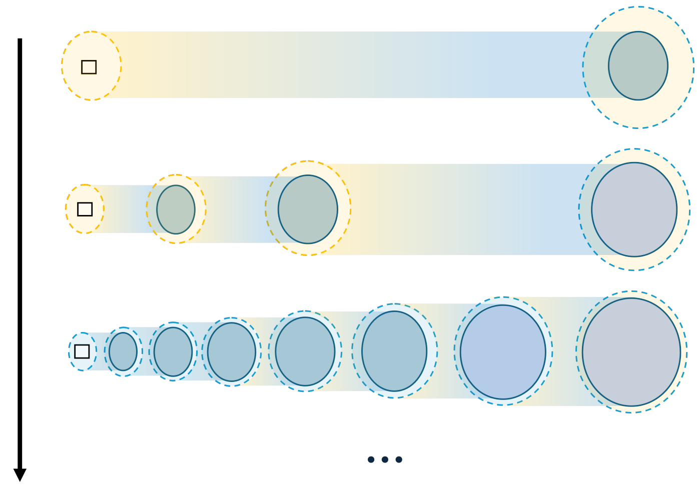

Figure 2.2: Dyadic Construction of Nested Open Sets

Proof. Denote $U_0 = A$ and $U_1 = B^c$, and let $\mathcal{F}_0 = \{U_0, U_1\}$, then $A \subset U_r \subset B^c$ for all $U_r \in \mathcal{F}_0$, and $\overline{U_0} = A \subset B^c = U_1$. Moreover, all sets in $\mathcal{F}_0$ except U_0 is open.

Let $n \in \mathbb{N}$, and suppose that $\mathcal{F}_{n-1}$ satisfies the above criteria (except U_0 being closed), and write

$$\mathcal{F}_{n-1} = \left\{U_{\frac{k}{2^{n-1}}} : 0 \leq k \leq 2^{n-1}\right\} = \left\{U_{\frac{2k}{2^n}} : 0 \leq k \leq 2^{n-1}\right\}$$

Let $k \in [0, 2^n]$ be an odd number, then $k+1$ and $k-1$ are even, $U_{k-1}, U_{k+1} \in \mathcal{F}_{n-1}$ are already defined. Since $\overline{U_{k-1}} \subset U_{k+1}$, $\overline{U_{k-1}}$ and U_{k+1}^c are disjoint closed sets. As X is normal,

$$\exists U \in \mathcal{T} : \overline{U_{k-1}} \subset U \subset \overline{U} \subset (U_{k+1}^c)^c = U_{k+1}$$

Let $U_k = U$ be such an open set, then $A \subset U_k \subset B^c$. Moreover, $\overline{U_k} \subset \overline{U_{k+1}} \subset U_s$ for all odd $s > k$, and $\overline{U_k} \subset \overline{U_{k+1}} \subset U_{s-1} \subset U_s$ for all even $s > k+1$. Define

$$\mathcal{F}_{n+1} = \left\{U_{\frac{k}{2^n}} : 0 \leq k \leq 2^{n-1}\right\}$$

then $\mathcal{F}_{n+1}$ also satisfies the desired criteria.

Take $\mathcal{F} = \bigcup_{n \in \mathbb{N}} \mathcal{F}_n \setminus \{U_0, U_1\}$, then every $U_r \in \mathcal{F}$ satisfies $A \subset U_r \subset B^c$, and with the exclusion of U_0, every set in $\mathcal{F}$ is open.

For any U_r, U_s, there exists $\mathcal{F}_n$ such that $U_r, U_s \in \mathcal{F}_n$. In which case, if $r < s$, $\overline{U_r} \subset U_s$. Moreover, for any dyadic number $r = \frac{k}{2^n}$, $U_r \in \mathcal{F}_n \subset \mathcal{F}$. Therefore $\mathcal{F}$ has the desired properties. $\square$

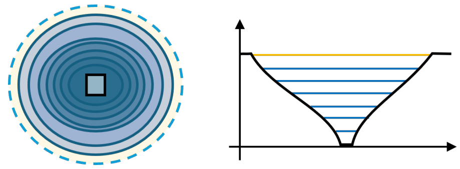

Figure 2.3: Layer Cake Construction of a Continuous Function

Theorem 2.2 (Urysohn's Lemma)**.** *Let $(X, \mathcal{T})$ be a normal topological space. If $A, B \subset X$ are disjoint closed sets, then there exists a continuous function $f : X \to [0,1]$ such that $f = 0$ on A and $f = 1$ on B.*

As a result, every normal space is completely regular.

Proof. Let $\Delta = \left\{ \frac{k}{2^n} : n \in \mathbb{N}, 0 < k < 2^n \right\}$ be the collection of dyadic rational numbers in $(0,1)$, then any number $x \in [0,1]$ can be approximated by elements of Δ. For any $r \in (0,1)$, we can write $r = \sup_{s \in \Delta : s < r} s = \inf_{s \in \Delta : s > r} s$.

Let

$$f : X \mapsto [0,1] \quad x \mapsto \inf \{ r : x \in U_r \}$$

where we assign x a value based on its stage of "interpolation" from A to B. Then since $A \subset U_r \subset B$ for all $r \in \Delta$, $f(A) = \{0\}$ and $f(B) = \{1\}$.

Let $\alpha \in [0,1]$, then since $\alpha = \sup_{s \in \Delta : s < \alpha} s$, $f(x) < \alpha$ if and only if there exists $s \in \Delta : s < \alpha$ such that $f(x) < s$, which only happens when $x \in \bigcup_{s \in \Delta : s < \alpha} U_s$. Therefore

$$f^{-1}((-\infty, \alpha)) = \bigcup_{s \in \Delta : s < \alpha} U_s$$

As each U_s is open, $\bigcup_{s < \alpha} U_s$ is open.

Since $\alpha = \inf_{s > \alpha} s$, $f(x) > \alpha$ if and only if there exists $s > \alpha$ such that $f(x) > s$ and $x \in U_s^c$. If $x \in U_s^c$, then there exists $t \in \Delta \cap (\alpha, s)$ such that $U_t \subset \overline{U_t} \subset U_s$ and $x \in \overline{U_t}^c \supset U_s^c$. If $x \in \overline{U_t}^c$ for some t, then there exists $s \in \Delta \cap (\alpha, t)$ such that $U_s \subset U_t \subset \overline{U_t}$ and $x \in U_s^c \supset \overline{U_t}^c$. Therefore

$$f^{-1}((\alpha, \infty)) = \bigcup_{s \in \Delta : s > \alpha} U_s^c = \bigcup_{t \in \Delta : t > \alpha} \overline{U_t}^c$$

As each $\overline{U_t}^c$ is open, $\bigcup_{t > \alpha} \overline{U_t}^c$ is open.

Since the open rays generate the standard topology on $\mathbb{R}$, and the preimage of those open rays are open, f is continuous. □

Theorem 2.3 (Tietze Extension Theorem)**.** *Let $(X, \mathcal{T})$ be a normal space, $A \subset X$ be a closed set and $f \in BC(A, [a,b])$ be a bounded continuous function. Then there exists a continuous extension $F \in BC(X, [a,b])$ of f to the whole space.*

Proof. First suppose that $f \in BC(A, [0,1])$.

Using the Method of Successive Approximations, take $BC(A, [0,1])$ equipped with the uniform norm as the space, with

$$T : BC(X, [0,1]) \to BC(A, [0,1]) \quad f \mapsto f|_A$$

as the continuous linear map, and $\Sigma = BC(X, [0,1])$ as the approximating sets.

Let $f \in BC(A, [0,1])$ and take

$$B = f^{-1}(||f|| \cdot [0, 1/3]) \quad C = f^{-1}(||f|| \cdot [2/3, 1])$$

As f is continuous, and B, C are closed in A under the relative topology. Since A is closed, B, C are closed in X as well.

By Urysohn's Lemma, there exists a continuous function $\phi : X \to [0, ||f|| /3]$ such that $\phi(B) = \{0\}$ and $\phi(C) = \{||f|| /3\}$. Since $\phi(x) \le ||f|| /3$ for all $x \in X$, $||\phi|| \le ||f||$. Now,

$$f(x) - \phi(x) \le ||f|| \left(1 - \frac{1}{3}\right) = \frac{2}{3} ||f|| \qquad \forall x \in B$$

$$0 \le f(x) - \phi(x) \le ||f|| \left(\frac{2}{3} - 0\right) \frac{2}{3} ||f|| \qquad \forall x \in B^c \cap A$$

with $f(x) \ge \frac{1}{3} \ge \phi(x)$ for all $x \in B^c \cap A$ and $f(x) \ge 0 = \phi(x)$ for all $x \in B$, we have $||f(x) - \phi|_A(x)|| \le \frac{2}{3} ||f||$.

Therefore Σ satisfies the criterion for the method of successive approximations, and for any $f \in BC(A, [0,1])$ there exists $\{\phi_n\}_1^\infty \subset BC(X, [0,1])$ such that $(\sum_{n=1}^\infty \phi_n)|_A = f$.

Now let $f \in BC(A, [a,b])$. Taking $g = \frac{f-a}{b-a}$ yields a function $g \in BC(A, [0,1])$. Reversing the transformation on the extension yields a continuous extension of the original f. □

2.2 LCH Spaces

Definition 2.2 (LCH Space)**.** *Let $(X, \mathcal{T})$ be a topological space. X is locally compact if every point has a compact neighbourhood. Locally compact Hausdorff spaces can be abbreviated as LCH.*

Theorem 2.4. *Let X be a LCH space and $E \subset X$, then E is closed if and only if $E \cap K$ is closed for every compact $K \subset X$.*

Proof. Since X is Hausdorff, all compact sets are closed and $E \cap K$ is closed for all compact sets K.

If E is not closed, then there exists an accumulation point $x \in \overline{E} \setminus E$ of E. Let K be a compact neighbourhood of x, then x is an accumulation point of $E \cap K$, not in $E \cap K$. Therefore $E \cap K$ is not closed. □

Theorem 2.5. *Let X be a LCH space, $x \in X$ be a point and $U \in \mathcal{N}(x)^o$ be an open neighbourhood of x, then there exists a compact neighbourhood $N \in \mathcal{N}(x)$ such that $x \in N \subset U$.*

Proof. Suppose that $\overline{U}$ is compact, then in the relative topology there exists disjoint open sets V and W such that $x \in V$ and $\partial U \subset W$. This gives $\overline{V} \subset W^c \subset U$ and as $\overline{U}$ is compact, so is $\overline{V}$ in the relative topolology. Since $\overline{U}$ is closed in X, $\overline{V}$ is also closed in X and therefore compact.

If $\overline{U}$ is not compact, take a compact neighbourhood $N \in \mathcal{N}(x)$ and replace U with $U \cap N^o$. □

Theorem 2.6. *Let X be a LCH space and $K \subset U \subset X$ where K is compact and U is open. Then there exists a precompact V such that $K \subset V \subset \overline{V} \subset U$.*

Proof. For any $x \in K$, let N_x be a compact neighbourhood of x such that $x \in N_x \subset U$. Then $\{U_x^o\}_{x \in K}$ is an open cover of K, and has a finite subcover $\{U_i\}_1^n$. Let $V = \bigcup_{i=1}^n U_i$, then $\overline{V} = \bigcup_{i=1}^n \overline{U_i} \subset U$ is a finite union of compact sets, and therefore compact. □

Theorem 2.7 (Urysohn's Lemma, LCH Version)**.** *Let X be a LCH space and $K \subset U \subset X$ where K is compact and U is open. Then there exists a continuous function $f \in C(X, [0,1])$ such that $f = 1$ on K and 0 outside of a compact subset of U.*

Therefore all LCH spaces are completely regular.

Proof. Let V be a precompact open neighbourhood of K with $K \subset V \subset \overline{V} \subset U$. Then $\overline{V}$ is a compact Hausdorff space, and therefore normal. Since K is compact in X, it is also compact in $\overline{V}$ and is closed.

By Urysohn's Lemma, there exists $f \in C(\overline{V}, [0,1])$ that is 1 on K and 0 on ∂V. Extend f to X by setting it to 0 on $\overline{V}^c$ and denote the extension as F. Let $E \subset [0,1]$ be closed, then

$$F^{-1}(E) = f^{-1}(E) \cup \left(F|_{\overline{V}^c}\right)^{-1}(E)$$

If $0 \notin E$, then $F^{-1}(E) = f^{-1}(E)$ is closed. If $0 \in E$, then $F^{-1}(E) = f^{-1}(E) \cup \overline{V}^c$. As $f^{-1}(0) \supset \partial V$,

$$F^{-1}(E) = f^{-1}(E) \cup \overline{V}^c = f^{-1}(E) \cup V^c$$

which is closed. Therefore f is continuous. □

Theorem 2.8 (Tietze Extension Theorem, LCH version)**.** *Let X be a LCH space and $K \subset X$ be compact. If $f \in C(K)$ is a continuous function, then there exists $F \in C(K)$ such that $F|_K = f$. Moreover, F can be taken to vanish outside of a compact set.*

Proof. Let V be a precompact open neighbourhood of K such that $K \subset V \subset \overline{V}$, then $\overline{V}$ is compact and normal. As K is compact in a compact Hausdorff space $\overline{V}$, K is closed.

By the Tietze Extension Theorem, there is an extension F of f in $\overline{V}$. Moreover, in the construction of ϕ in the proof, we can take $B' = B \cup \partial V$, which is still closed, and does not influence the approximation. Therefore we can take F to be 0 outside of $\overline{V}$.

Similar to Urysohn's Lemma, we can decompose any preimage

$$F^{-1}(E) = F|_{\overline{V}}^{-1}(E) \cup \overline{V}^c = F|_{\overline{V}}^{-1}(E) \cup V^c$$

with the V^c being optional. This way, the preimage of every closed set is closed, and the preimage of every open set is open. Therefore the extension is continuous. □

Theorem 2.9. *Let X be a LCH space, then the space $C_0(X)$ of functions that vanishes at infinity is the closure of the space $C_c(X)$ of compactly supported functions with respect to the uniform norm.*

Proof. Let $\{f_n\}_1^\infty \subset C_c(X)$ such that $f_n \to f$ uniformly, then for all $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $||f_n - f|| < \varepsilon$. As a result, for any $x \notin \operatorname{supp}(f_n)$, $|f(x)| < \varepsilon$ and $\{x : |f(x)| \geq \varepsilon\} \subset \operatorname{supp}(f_n)$ is a closed subset of a compact set, and therefore compact. So $\overline{C_c(X)} \subset C_0(X)$.

Now let $f \in C_0(X)$ and $n \in \mathbb{N}$ with $K_n = \{x : |f(x)| \geq 1/n\}$, then K_n is compact. By the LCH version of Urysohn's Lemma, there is an extension of f to a continuous function g_n that is 1 on K_n and 0 outside of some compact neighbourhood $\overline{V_n}$ of K_n. Take $f_n = fg_n$, then $f_n \in C_0(X)$ vanishes outside of $\overline{V_n}$ with

$$||f_n - f|| = ||f(1 - g_n)|| \leq \left|\left|f \cdot \chi_{K_n^c}\right|\right| < \frac{1}{n}$$

so $f_n \to f$ uniformly. □

Theorem 2.10. *Let X be a LCH space, then the space of continuous functions $C(X)$ is a closed subspace of $\mathbb{C}^X$ under the topology of uniform convergence on compact sets.*

Proof. Let $f \in \overline{C(X)}$, then f is an adherent point of $C(X)$, and cannot be separated from elements of $C(X)$ with the uniform norm on compact sets.

Let $K \subset X$ be compact, then for any $n \in \mathbb{N}$,

$$\exists f_n \in C(X) : \sup_{x \in K} |f_n(x) - f(x)| < \frac{1}{n}$$

meaning that $f_n|_K \to f|_K$ uniformly, and f is continuous on K.

Let $x \in X$ and $K \in \mathcal{N}(x)$ be a compact neighbourhood, then f is continuous on K. Since for any neighbourhood E of $f(x)$, $f^{-1}(E)$ is a neighbourhood of x in K, and $f^{-1}(E)^o \cap K^o$ is an open neighbourhood of x in X. Therefore f is continuous at x.

As f is continuous at x for all $x \in X$, f is continuous. □

2.3 σ-Compact LCH Spaces

Theorem 2.11. *Let X be a separable LCH space, then X is σ-compact.*

Proof. Let $\{x_n\}_1^\infty \subset X$ be a countable dense subset, and let $K_n \in \mathcal{N}^k(x)$, then K_n is compact and $X = \bigcup_{n\in\mathbb{N}} K_n$. □

Theorem 2.12. *Let X be a σ-compact LCH space, then there is a sequence $\{U_i\}_1^\infty$ of precompact open sets such that $\overline{U_n} \subset U_{n+1}$ and $X = \bigcup_{i\in\mathbb{N}} U_i$.*

Proof. Let $X = \bigcup_{n\in\mathbb{N}} K_n$ be where each K_n is compact.

Let $U_1 \in \mathcal{N}(K_1)^o$ be a precompact open neighbourhood of K_1. For any $n \in \mathbb{N}$, take a precompact open neighbourhood,

$$U_n \in \mathcal{N}\left(K_n \cup \bigcup_{i=1}^{n-1} \overline{U_i}\right)^o$$

then each $U_n \supset \bigcup_{i=1}^{n-1} \overline{U_i} = \overline{U_{n-1}}$ and $\bigcup_{n\in\mathbb{N}} U_n = X$. □

Theorem 2.13. *Let X be a σ-compact LCH space and $\{U_n\}_1^\infty$ be a sequence of precompact open sets such that $\overline{U_n} \subset U_{n+1}$ for all n and $\bigcup_{n\in\mathbb{N}} U_n = X$, then*

$$\mathcal{E} = \left\{ \left\{ g \in \mathbb{C}^X : \sup_{x\in\overline{U_m}} |f(x) - g(x)| < \frac{1}{n} \right\} \middle| m, n \in \mathbb{N} \right\}$$

is a neighbourhood base for $f \in \mathbb{C}^X$ in the topology of uniform convergence on compact sets.

Hence, this topology is first countable with $f_n \to f$ if and only if $f_n \to f$ uniformly on every $\overline{U_n}$.

Proof. **Neighbourhood Reduction**

Denote the generating sets of the topology as

$$B_K(f, n) = \left\{ g \in \mathbb{C}^X : \sup_{x\in K} |g(x) - f(x)| < \frac{1}{n} \right\}$$

then for any $U \in \mathcal{N}(f)^o$, there exists K compact and $n \in \mathbb{N}$ such that $f \in B_K(f, n) \subset U$.

Proof. Suppose that $B_K(g,n) \in \mathcal{N}(f)^o$ is an open neighbourhood of f, then there exists $m \in \mathbb{N}$ such that $f \in B_K(f,m) \subset B_K(g,n)$ and we can assume that neighbourhoods in the generating set are "centred" at f.

Let $U \in \mathcal{N}(f)^o$ be an open neighbourhood of f, then it contains a finite intersection of of the generating sets $\bigcap_{i=1}^n B_{K_i}(g_i, n_i)$, where each K_i is compact, and $B_{K_i}(g_i, n_i)$ is an open neighbourhood of f. Therefore we can write

$$f \in \bigcap_{i=1}^{n} B_{K_i}(f, m_i) \subset \bigcap_{i=1}^{n} B_{K_i}(g_i, n_i)$$

where taking $K = \bigcup_{i=1}^n K_i$ and $m = \max_{i=1}^n m_i$ yields

$$\bigcap_{i=1}^{n} B_{K_i}(f, m_i) \supset \bigcap_{i=1}^{n} B_{K_i}(f, m) = B_K(f, m)$$

and K is compact since it is a finite union of compact sets.

Therefore any open neighbourhood U can be reduced to $B_K(f,m)$ with $f \in B_K(f,m) \subset U$ where $m \in \mathbb{N}$ and K compact.

Compactness Reduction

Let $\{U_n\}_1^\infty$ be a sequence of precompact open sets such that $\overline{U_n} \subset U_{n+1}$ for all $n \in \mathbb{N}$ and $\bigcup_{n \in \mathbb{N}} U_n = X$. Then for any compact set K there exists $j \in \mathbb{N}$ such that $U_j \supset K$.

Proof. Since $\bigcup_{n \in \mathbb{N}} U_n = X$, $\{U_n\}_1^\infty$ forms an open cover of K, there exists a finite subcover $\{U_j\}_{j \in J}$ with $J \subset \mathbb{N}$ finite. Let $j = \max(J)$, as $U_j \supset U_k$ for all $j \leq j$, $U_j \supset K$. Therefore for any K compact there exists $j \in \mathbb{N}$ such that $U_j \supset K$.

Neighbourhood Base

The collection

$$\mathcal{E}(f) = \{B_K(f,n) : n \in \mathbb{N}\}$$

forms a neighbourhood base for $\mathcal{T}$ at f.

Proof. Let $U \in \mathcal{N}(f)^o$, then there exists $B_K(f,n)$ such that $f \in B_K(f,n) \subset U$. Since K is compact, there exists precompact U_j such that $U_j \supset K$. Take $B_{\overline{U_j}}(f,n) \in \mathcal{E}(f)$, then

$$f \in B_{\overline{U_j}}(f,n) \subset B_K(f,n) \subset U$$

Therefore $\mathcal{E}(f)$ forms a neighbourhood base at f. □

Theorem 2.14. *Let $(X, \mathcal{T})$ be a σ-compact LCH space, then X is paracompact.*

Proof. Let $\{U_n\}_1^\infty$ be a sequence of precompact open sets such that $\overline{U_n} \subset U_{n+1}$ and $\bigcup_{n \in \mathbb{N}} U_n = U$. Let $\mathcal{U}$ be $n \in \mathbb{N}$, then

$$V_n = \left\{E \cap (U_{n+2} \setminus \overline{U_{n-1}}) : E \in \mathcal{U}\right\}$$

is an open cover of $\overline{U_{n+1}} \setminus U_n$. As $\overline{U_{n+1}} \setminus U_n$ is compact, $\mathcal{V}_n$ has a finite subcover $\mathcal{V}_n$. Let $\mathcal{V} = \bigcup_{n\in\mathbb{N}} \mathcal{V}_n$, then $\mathcal{V}$ is a refinement of $\mathcal{U}$, and for any point $x \in U_n \setminus \overline{U_{n-1}}$, only sets in $\mathcal{V}_n$ can contain x. Therefore $\mathcal{V}$ is locally finite. □

Chapter 3

Functional Analysis

3.1 Topological Vector Spaces

Definition 3.1 (Topological Vector Space)**.** *Let $\mathcal{X}$ be a vector space over $F = \mathbb{R}|\mathbb{C}$, equipped with a topology $\mathcal{T}$. $\mathcal{X}$ is a **topological vector space** if the maps*

$$\mathcal{X} \times \mathcal{X} \to \mathcal{X} \quad (x, y) \mapsto x + y$$

and

$$F \times \mathcal{X} \to \mathcal{X} \quad (\lambda, x) \mapsto \lambda x$$

are continuous.

Definition 3.2 (Complete TVS)**.** *Let $\mathcal{X}$ be a topological vector space. $\mathcal{X}$ is **complete** if every Cauchy net in $\mathcal{X}$ converges.*

Theorem 3.1. *Let $\{p_i\}_{i \in I}$ be a family of seminorms, take*

$$U_{xi\varepsilon} = \{y \in \mathcal{X} : p_i(x - y) < \varepsilon\}$$

Let

$$\mathcal{E} = \{U_{xi\varepsilon} : x \in \mathcal{X}, i \in I, \varepsilon > 0\}$$

and $\mathcal{T} = \mathcal{T}(\mathcal{E})$ be the weak topology generated by the norms. Then

1. *For any $x \in \mathcal{X}$, the collection of finite intersections of $\mathcal{E}_x = \{U_{xi\varepsilon} : i \in I, \varepsilon > 0\}$ is a neighbourhood base at x.*
2. *If $\langle x_\alpha \rangle_{\alpha \in A}$ is a net in $\mathcal{X}$, then $x_\alpha \to x$ if an only if $p_i(x_\alpha - x) \to 0$ for all $i \in I$.*
3. *$(\mathcal{X}, \mathcal{T})$ is a locally convex topological vector space.*

Proof. **Centre Reduction**

Let $U_{yi\varepsilon} \in \mathcal{E}$ and $x \in U_{yi\varepsilon}$, then there exists $\delta > 0$ such that $U_{xi\delta} \subset U_{yi\varepsilon}$.

Proof. Let $\delta < \varepsilon - p_i(y-x)$, then for all $z \in U_{xi\delta}$,

$$p_i(y-z) \leq p_i(y-x) + p_i(z-x) < \varepsilon$$

and $U_{xi\delta} \subset U_{yi\varepsilon}$.

Intersection Reduction (Same Norm)

Let $U_{xi\varepsilon}, U_{yi\delta} \in \mathcal{E}$ such that $U_{xi\varepsilon} \cap U_{yi\delta} \neq \emptyset$, then for any $z \in U_{xi\varepsilon} \cap U_{yi\delta}$, there exists $\gamma > 0$ such that $U_{zi\gamma} \subset U_{xi\varepsilon} \cap U_{yi\delta}$.

Proof. Let $\gamma < \min(\varepsilon - p_i(x-z), \delta - p_i(y-z))$, then $U_{zi\gamma} \subset U_{xi\varepsilon}$ from the first term, and $U_{zi\gamma} \subset U_{yi\delta}$ from the second.

Intersection Reduction (Different Norms)

Let $\{U_{x_k,i_k,\varepsilon_k}\}_1^n \subset \mathcal{E}$ be a finite sequence of sets such that $\bigcap_{k=1}^n U_{x_k,i_k,\varepsilon_k} \neq \emptyset$. Then for any $x \in \bigcap_{k=1}^n U_{x_k,i_k,\varepsilon_k}$, there exists $\{U_{x,j_k,\delta_k}\}_1^m \subset \mathcal{E}_x$ such that $\bigcap_{k=1}^m U_{x,j_k,\delta_k} \subset \bigcap_{k=1}^n U_{x_k,i_k,\varepsilon_k}$ and the j_ks are distinct.

Proof. If there exists k, k' such that $i_k = i_{k'}$, then there exists $x' \in U_{x_k i_k \varepsilon_k} \cap U_{x_{k'} i_k \varepsilon_{k'}}$ and $\varepsilon' > 0$ such that $U_{x' i_k \varepsilon'} \subset U_{x_k i_k \varepsilon_k} \cap U_{x_{k'} i_k \varepsilon_{k'}}$. Replace the two balls with $U_{x' i_k \varepsilon'}$, then the new intersection is contained in the old intersection. Repeat this process until the j_ks are distinct.

Now suppose that the i_ks are distinct, then for each $U_{x_k,i_k,\varepsilon_k}$ choose $U_{x,i_k,\delta_k} \subset U_{x_k,i_k,\varepsilon_k}$, then $\bigcap_{k=1}^n U_{x,i_k,\delta_k} \subset \bigcap_{k=1}^n U_{x_k,i_k,\varepsilon_k}$.

Neighbourhood Base

Let $U \in \mathcal{T}$, then U is a union of finite intersections of $\mathcal{E}$ sets. If $x \in U$, then there exists $\{U_{x_k,i_k,\varepsilon_k}\}_1^n \subset \mathcal{E}$ such that

$$x \in \bigcap_{k=1}^n U_{x_k,i_k,\varepsilon_k} \subset U$$

Since there exists $\{U_{x,j_k,\delta_k}\}_1^m \subset \mathcal{E}_x$ such that

$$x \in \bigcap_{k=1}^m U_{x,j_k,\delta_k} \subset \bigcap_{k=1}^n U_{x_k,i_k,\varepsilon_k} \subset U$$

the collection of finite intersections of $\mathcal{E}_x$ sets forms a neighbourhood base at x.

Net

If $\langle x_\alpha \rangle_{\alpha \in A}$ is a net in $\mathcal{X}$, then $x_\alpha \to x$ if an only if $p_i(x_\alpha - x) \to 0$ for all $i \in I$.

Proof. Suppose that $p_i(x_\alpha - x) \to 0$ for all $i \in I$. Let $U \in \mathcal{T}$, then there exists $\{U_{x,i_k,\delta_k}\}_1^m \subset \mathcal{E}_x$ such that $\bigcap_{k=1}^m U_{x,j_k,\delta_k} \subset U$. Since $p_i(x_\alpha - x) \to 0$, there exists α_k for each k such that $p_i(x_{\alpha_k} - x) < \delta_k$. Choose α such that $\alpha \gtrsim a_k$ for all k, then $p_i(x_\alpha - x) < \delta_k$ for all k, and $x_\alpha \to x$.

Suppose that $x_\alpha \to x$, and fix $i \in I$. Then for all $\varepsilon > 0$, there exists $\alpha \in A$ such that $x_\beta \in U_{xi\varepsilon}$ for all $\beta \gtrsim \alpha$.

Intersection Reduction

Let $\mathcal{X}$ be a vector space. If $C, D \subset \mathcal{X}$ are convex, then $C \cap D$ is also convex.

Proof. Let $x, y \in C \cap D$, then

$$\{tx + (1-t)y : t \in [0,1]\} \subset C, D$$

Therefore $C \cap D$ is also convex.

Convex "Balls"

Let $\mathcal{X}$ be a vector space and p be a seminorm. Then sets of the form

$$U(x, \varepsilon) = \{y \in \mathcal{X} : p(x-y) < \varepsilon\} \quad x \in \mathcal{X}, \varepsilon > 0$$

are convex.

Proof. Let $x, y \in U(x, \varepsilon)$, then

$$\begin{aligned} p(tx + (1-t)y) &\leq tp(x) + (1-t)p(y) \\ &< t\varepsilon + (1-t)\varepsilon \\ &= \varepsilon \end{aligned}$$

Therefore $\{tx + (1-t)y : t \in [0,1]\} \subset U(x, \varepsilon)$.

Convex Base

The collection of finite intersections of

$$\mathcal{E} = \{U_{xi\varepsilon} : x \in \mathcal{X}, i \in I, \varepsilon > 0\}$$

forms a base of $\mathcal{T}$ consisting of convex sets.

Proof. Since finite intersections of $\mathcal{E}_x$ forms a neighbourhood base at x, and $\mathcal{E}_x \subset \mathcal{E}$ for all $x \in \mathcal{X}$, the finite intersections of $\mathcal{E}$ sets is a base for $\mathcal{T}$.

Let $\{U_{x_k, i_k, \varepsilon_k}\}_1^n \subset \mathcal{E}$, then since each $U_{x_k, i_k, \varepsilon_k}$ is convex, $\bigcap_{k=1}^n U_{x_k, i_k, \varepsilon_k}$ is also convex. Therefore finite intersections of $\mathcal{E}$ sets consists of convex sets.

Topological Vector Space

Let $\phi : \mathcal{X}^2 \to \mathcal{X}$ by $(x, y) \mapsto x + y$, then ϕ is continuous.

Proof. Let $U \in \mathcal{N}(x+y)^o$, then there exists $\{U_{x+y, i_k, \varepsilon_k}\}_1^n$ such that $\bigcap_{k=1}^n U_{x+y, i_k, \varepsilon_k} \subset U$.

Fix $i \in I$ and $\varepsilon > 0$, and consider $U_{x+y, i, \varepsilon}$, then $U_{x, i, \varepsilon/2} \times U_{y, i, \varepsilon/2}$ is open in $\mathcal{X}$ where for all $(x', y') \in U_{x, i, \varepsilon/2} \times U_{y, i, \varepsilon/2}$,

$$\begin{aligned} p_i((x'+y') - (x+y)) &\leq p_i(x'-x) + p_i(y'-y) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

and we have found $U_{x, i, \varepsilon/2} \times U_{y, i, \varepsilon/2} \in \mathcal{N}((x,y))^o$ such that $U_{x, i, \varepsilon/2} \times U_{y, i, \varepsilon/2} \subset \phi^{-1}(U_{x+y, i, \varepsilon})$.

Now, since for all $U_{x+y,i,\varepsilon}$, there exists $f^{-1}(U_{x+y,i,\varepsilon}) \in \mathcal{N}((x,y))$,

$$f^{-1}\left(\bigcap_{k=1}^{n} U_{x+y,i_k,\varepsilon_k}\right) = \bigcap_{k=1}^{n} f^{-1}(U_{x+y,i_k,\varepsilon_k}) \in \mathcal{N}((x,y))$$

□

Theorem 3.2. *Let $\mathcal{X}$ and $\mathcal{Y}$ be TVSs with topologies induced by the families of seminorms $\{p_i\}_{i\in I}$ and $\{q_j\}_{j\in J}$. A linear map $T : \mathcal{X} \to \mathcal{Y}$ is continuous if and only if for every $j \in J$, there exists $\{i_k\}_1^n$ and C such that*

$$q_j(Tx) \leq C \sum_{k=1}^{n} p_{i_k}(x) \quad \forall x \in \mathcal{X}$$

Proof. **Forward**

Relate the open unit ball with respect to a given norm in $\mathcal{Y}$ to an intersection of open balls in $\mathcal{X}$. Scale vectors in $\mathcal{X}$ such that they fit into the ball to establish the inequality.

Suppose that T is continuous, then for every $j \in J$, there exists $\{i_k\}_1^n$ and C such that

$$q_j(Tx) \leq C \sum_{k=1}^{n} p_{i_k}(x) \quad \forall x \in \mathcal{X}$$

Proof. Let $j \in J$, then $U_{0,j,1}$ is open in $\mathcal{Y}$ and $T^{-1}(U_{0,j,1})$ is open in $\mathcal{X}$. Since there exists $\{U_{0,i_k,\varepsilon_k}\}_1^n$ such that $\bigcap_{k=1}^n U_{0,i_k,\varepsilon_k} \subset T^{-1}(U_{0,j,1})$.

Let $\varepsilon < \varepsilon_k$ for all k, then $\bigcap_{k=1}^n U_{0,i_k,\varepsilon} \subset T^{-1}(U_{0,j,1})$. Let $x \in \mathcal{X}$. If $\sum_{k=1}^n p_{i_k}(x) > 0$, take $y = \frac{\varepsilon}{\sum_{k=1}^n p_{i_k}(x)} \cdot x$, then

$$p_{i_l}(y) = \frac{\varepsilon p_{i_l}(x)}{\sum_{k=1}^n p_{i_k}(x)} \leq \varepsilon \quad \forall l \in [1,n]$$

meaning that $y \in \bigcap_{k=1}^n U_{0,i_k,\varepsilon}$ and

$$\begin{aligned} q_j(Tx) &= \frac{\sum_{k=1}^n p_{i_k}(x)}{\varepsilon} \cdot \frac{\varepsilon}{\sum_{k=1}^n p_{i_k}(x)} \cdot q_j(Tx) \\ &= \frac{\sum_{k=1}^n p_{i_k}(x)}{\varepsilon} \cdot q_j(Ty) \\ &\leq \frac{1}{\varepsilon} \sum_{k=1}^{n} p_{i_k}(x) \end{aligned}$$

Backward

Suppose that for every $j \in J$, there exists $\{i_k\}_1^n$ and C such that

$$q_j(Tx) \leq C \sum_{k=1}^{n} p_{i_k}(x) \quad \forall x \in \mathcal{X}$$

then T is continuous.

Proof. Let $\langle x_\alpha \rangle_{\alpha \in A}$ be a net converging to x. Then $p_i(x_\alpha - x) \to 0$ for all $i \in I$. Since for each $j \in J$ we can bound $q_j(Tx)$ with a finite sum of p_is,

$$q_j(Tx_\alpha - Tx) \leq C \sum_{k=1}^{n} p_{i_k}(x_\alpha - x) \to 0$$

we have $q_j(Tx_\alpha - Tx) \to 0$ for all $j \in J$, and $Tx_\alpha \to Tx$. □

Theorem 3.3. *Let $\mathcal{X}$ be a vector space equipped with the topology defined by a family $\{p_i\}_{i \in I}$ of seminorms.*

1. *$\mathcal{X}$ is Hausdorff if and only if for each $x \neq 0$ there exists $i \in I$ such that $p_i(x) \neq 0$.*
2. *If $\mathcal{X}$ is Hausdorff and I is countable, then $\mathcal{X}$ is metrisable with a translation-invariant metric.*

Proof. **Hausdorff**

$\mathcal{X}$ is Hausdorff if and only if for each $x \neq 0$ there exists $i \in I$ such that $p_i(x) \neq 0$.

Proof. Suppose that $\mathcal{X}$ is Hausdorff, then for any $x \in \mathcal{X}$ there exists $U \in \mathcal{N}(x)^o$ and $V \in \mathcal{N}(0)^o$ such that $U \cap V = \emptyset$. Since U is open, there exists $\{U_{xi_k\varepsilon_k}\}_1^n$ such that $x \in \bigcap_{k=1}^n U_{xi_k\varepsilon_k} \subset U$. As the intersection does not contain 0, there exists k such that $0 \notin U_{xi_k\varepsilon_k}$, therefore $p_{i_k}(x-0) = p(i_k)(x) > \varepsilon_k > 0$.

Suppose that for each $x \neq 0$ there exists $i \in I$ such that $p_i(x) \neq 0$. Let $x, y \in \mathcal{X} : x \neq y$, then there exists $i \in I$ such that $p_i(x-y) \neq 0$. Let $\varepsilon < p_i(x-y)/2$, then $U = U_{xi\varepsilon}$ and $V = U_{yi\varepsilon}$ are disjoint open neighbourhoods that separate x and y.

Translation-Invariant Metric

If $\mathcal{X}$ is Hausdorff and I is countable, index the seminorms with $\mathbb{N}$ and define

$$\rho(x,y) = \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x-y))$$

where

$$\phi(t) = \begin{cases} \frac{t}{1+t} & t < \infty \\ 1 & t = \infty \end{cases}$$

then ρ is a translation-invariant metric on $\mathcal{X}$.

Metric

ρ is a metric on $\mathcal{X}$.

Proof. Let $x, y \in \mathcal{X}$. If $x = y$, then $p_n(x - y) = 0$ for all $n \in \mathbb{N}$, and $\rho(x, y) = 0$. If $x \neq y$, then there exists $n \in \mathbb{N}$ such that $p_n(x - y) > 0$, meaning that $\rho(x, y) > 0$.

Let $x, y, z \in \mathcal{X}$, then

$$\begin{aligned}
\rho(x, z) &= \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x - z)) \\
&\leq \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x - y) + p_n(y - z)) \\
&\leq \sum_{n \in \mathbb{N}} 2^{-n} \left[\phi(p_n(x - y)) + \phi(p_n(y - z))\right] \\
&= \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x - y)) + \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(y - z)) \\
&= \rho(x, y) + \rho(y, z)
\end{aligned}$$

and ρ satisfies the triangle inequality.

Lastly, let $x, y \in \mathcal{X}$, then

$$\begin{aligned}
\rho(x, y) &= \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x, y)) \\
&= \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(y - x)) \\
&= \rho(y, x)
\end{aligned}$$

Therefore ρ is a metric.

Translation-Invariant

Let $x, y, z \in \mathcal{X}$, then $\rho(x, y) = \rho(x + z, y + z)$.

Proof.

$$\begin{aligned}
\rho(x, y) &= \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x - y)) \\
&= \sum_{n \in \mathbb{N}} 2^{-n} \phi(p_n(x + z - (y + z))) \\
&= \rho(x + z, y + z)
\end{aligned}$$

Metrisable

If $\mathcal{X}$ is Hausdorff and I is countable, then ρ defined above induces the same topology as the topology defined by the seminorms.

Notations

Denote $\mathcal{T}_S$ as the topology defined by the seminorms, and $\mathcal{T}_\rho$ be the topology induced by the metric.

Let

$$B(x,n,r) = \{y \in \mathcal{X} : p_n(x-y) < r\}$$

and

$$B(x,r) = \{y \in \mathcal{X} : \rho(x,y) < r\}$$

Metric Generates Seminorm Topology

Let $x \in \mathcal{X}$, $n \in \mathbb{N}$, and $r > 0$, then there exists $\varepsilon > 0$ such that $B(x,\varepsilon) \subset B(x,n,r)$.

From this, for each $y \in B(x,n,r)$, there exists r_y and ε_y such that $B(y,\varepsilon_y) \subset B(y,n,r_y) \subset B(x,n,r)$ and

$$B(x,n,r) = \bigcup_{y \in B(x,n,r)} B(y,\varepsilon_y)$$

Therefore $B(x,n,r) \in \mathcal{T}_\rho$ and $\mathcal{T}_S \subset \mathcal{T}_\rho$.

Proof. Let $\varepsilon < 2^{-n}\phi(r)$, then when $\rho(x,y) < \varepsilon$,

$$\begin{aligned} \rho(x,y) &< \varepsilon \\ \sum_{k \in \mathbb{N}} 2^{-k}\phi(p_k(x-y)) &< 2^{-n}\phi(r) \\ 2^{-n}\phi(p_n(x-y)) &< 2^{-n}\phi(r) \\ p_n(x-y) &< r \end{aligned}$$

Therefore $B(x,\varepsilon) \subset B(x,n,r)$.

Seminorms Generate Metric Topology

Let $x \in \mathcal{X}$ and $r > 0$, then there exists $n \in \mathbb{N}$ and $\varepsilon > 0$ such that $B(x,n,\varepsilon) \subset B(x,r)$.

From this, for each $y \in B(x,r)$, there exists r_y, n_y, and ε_y such that $B(y,n_y,\varepsilon_y) \subset B(y,r_y) \subset B(x,r)$ and

$$B(x,r) = \bigcup_{y \in B(x,r)} B(y,n_y,\varepsilon_y)$$

Proof. Let $n \in \mathbb{N}$ such that $2^{-n} < r/2$. Let $\varepsilon > 0$ such that

$$\sum_{k=1}^{n} 2^{-k}\phi(\varepsilon) < r/2$$

Then for all $y \in B(x, n, \varepsilon)$,

$$\begin{aligned}
\rho(x,y) &= \sum_{k\in\mathbb{N}} 2^{-k}\phi(p_k(x-y)) \\
&= \sum_{k=1}^{n} 2^{-k}\phi(p_k(x-y)) + \sum_{k>n} 2^{-k}\phi(p_k(x-y)) \\
&\le \sum_{k=1}^{n} 2^{-k}\phi(p_n(x-y)) + \sum_{k>n} 2^{-k} \\
&< \sum_{k=1}^{n} 2^{-k}\phi(\varepsilon) + \sum_{k>n} 2^{-k} \\
&< r/2 + r/2 = r
\end{aligned}$$

Therefore $B(x, n, \varepsilon) \subset B(x, r)$. □

3.2 Affine Hyperplanes

Definition 3.3 (Affine Hyperplane)**.** *Let $\mathcal{X}$ be a normed vector space, an **affine hyperplane** is a set of the form*

$$H = [f = \alpha] = f^{-1}(\alpha)$$

where $f : \mathcal{X} \to \mathbb{R}$ is a linear functional that does not vanish identically.

Theorem 3.4. *The hyperplane $[f = \alpha]$ is closed if and only if f is continuous.*

Proof. **Forward**

Suppose that f is continuous, then for any convergent sequence $\{x_i\}_1^\infty \subset [f = \alpha]$ with $x_i \to x$, $f(x_i) \to f(x) = \alpha$, and $[f = \alpha]$ is closed.

Backwards

Leveraging the linearity of f, if the plane is closed, then we can find a convex neighbourhood above/under the plane for any point outside of it. If any point in there were to have a value "past the plane", then we can find somewhere in between that *should* be in the plane. However, the neighbourhood is supposed to be above/under the plane. Therefore this cannot happen.

Once we have continuity at one point on the linear functional, we can translate it back to the origin by simply shaving off the value of f at that point.

Now suppose that $[f = \alpha]$ is closed. As f does not vanish identically, $[f = \alpha]^c \neq \emptyset$. Let $x \in [f = \alpha]^c$ and suppose that $f(x) < \alpha$, then there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset [f = \alpha]^c$.

We claim that $f(y) < \alpha$ for all $y \in B(x, \varepsilon)$. Suppose that there exists $z \in B(x, \varepsilon)$ such that $f(z) > \alpha$, then in the line segment

$$\{(1-t)x + tz : t \in [0, 1]\} \subset B(x, \varepsilon)$$

there exists $t \in [0,1]$ such that $f((1-t)x + tz) = \alpha$, namely

$$t = \frac{\alpha - f(x)}{f(z) - f(x)}$$

which is in $[0,1]$ since $0 \leq \alpha - f(x) \leq f(z) - f(x)$. This contradicts the fact that $B(x,\varepsilon) \subset [f = \alpha]^c$. Therefore $f(z) < \alpha$ for all $z \in B(x,\varepsilon)$.

Let $x \in [f = \alpha]^c$ such that $f(x) < \alpha$ and $\varepsilon > 0$ such that $B(x,\varepsilon) \subset [f = \alpha]^c$. Let $y \in B(x,\varepsilon)$, then $y = x + y'$. Since

$$\begin{gathered} f(x+y') = f(x) + f(y') \\ f(y') < \alpha - f(x) \\ |f(y')| < \alpha - f(x) \end{gathered}$$

we have

$$|f(x+y')| \leq |f(x)| + |y'| < |f(x)| + |\alpha - f(x)|$$

and f is bounded on $B(x,\varepsilon)$. Since f is bounded on an open set, f is continuous. □

Definition 3.4 (Separation)**.** *Let $A, B \subset \mathcal{X}$ be two subsets. The hyperplane $[f = \alpha]$ **separates** A and B if*

$$f(x) \leq \alpha \forall x \in A \quad f(x) \geq \alpha \forall x \in B$$

*and **strictly separates** A and B if there exists some $\varepsilon > 0$ such that*

$$f(x) \leq \alpha - \varepsilon \forall x \in A \quad f(x) \geq \alpha + \varepsilon \forall x \in B$$

Definition 3.5 (Gauge)**.** *Let $\mathcal{X}$ be a normed space and $C \subset \mathcal{X}$ be an open convex set with $0 \in C$. For every $x \in \mathcal{X}$ define*

$$p(x) = \inf \left\{ \alpha > 0 : \frac{x}{\alpha} \in C \right\}$$

*as the **gauge** of C. Then*

1. *p is a sublinear functional.*
2. *There exists $M \geq 0$ such that $0 \leq p(x) \leq M\,||x||$ for all $x \in \mathcal{X}$.*
3. *$C = \{x \in \mathcal{X} : p(x) < 1\}$.*

Proof. (1, linear) First let $\lambda > 0$, and $x \in \mathcal{X}$, then

$$\begin{aligned} \lambda \left\{ \alpha > 0 : \frac{x}{\alpha} \in C \right\} &= \left\{ \lambda\alpha > 0 : \frac{x}{\alpha} \in C \right\} \\ &= \left\{ \lambda\alpha > 0 : \frac{\lambda x}{\lambda\alpha} \in C \right\} \\ &= \left\{ \alpha > 0 : \frac{\lambda x}{\alpha} \in C \right\} \end{aligned}$$

therefore $p(x) = \lambda p(x)$.

(2) Since C is open, we can find a neighbourhood centred at 0 such that any vectors of a certain length will be inside C. Using this we can create an upper bound on the scale-down factor in terms of $||x||$.

Let $r > 0$ such that $\overline{B(0,r)} \subset C$, then $r\frac{x}{||x||} \in C$ for any $x \neq 0$, and we have

$$p(x) \leq \frac{1}{r}\,||x|| \quad \forall x \in E$$

(3) Let $x \in C$, then since C is open $\exists \varepsilon > 0 : B(x, \varepsilon) \subset C$. Choose $\delta > 0$ such that $(1+\delta)\,||x|| < ||x|| + \varepsilon$, then $(1+\delta)x \in C$ and $p(x) \leq \frac{1}{1+\delta} < 1$.

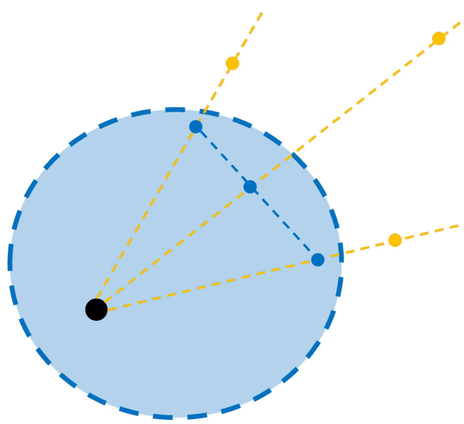

Figure 3.1: Scaling Vectors into C

(1, sublinear) Let $x, y \in \mathcal{X}$ and $\alpha, \beta > 0$ be estimates of $p(x)$ and $p(y)$ such that $\frac{x}{\alpha}$ and $\frac{y}{\beta}$ are in C. As C is convex, $\frac{tx}{\alpha} + \frac{(1-t)y}{\beta} \in C$ for all $t \in [0,1]$. Let $t_0 \in [0,1]$ such that

$$\frac{x+y}{\alpha+\beta} = \frac{t_0 x}{\alpha} + \frac{(1-t_0)y}{\beta}$$

then $\frac{x+y}{\alpha+\beta} \in C$, and $\alpha + \beta \geq p(x+y)$. Since this holds for all $\alpha > p(x)$ and $\beta > p(y)$, $p(x+y) \leq p(x) + p(y)$. □

3.3 Hahn-Banach Theorem

Theorem 3.5 (Hahn-Banach Theorem, Analytic Form)**.** *Let $\mathcal{X}$ be a vector space over $\mathbb{R}$, p a sublinear functional on $\mathcal{X}$. Let $\mathcal{M}$ be a subspace on $\mathcal{X}$, and f be a linear functional on $\mathcal{M}$ such that $f(x) \leq p(x)$ for all $x \in \mathcal{M}$. Then there is an extension F of f on $\mathcal{X}$ such that $F(x) \leq p(x)$ for all $x \in \mathcal{X}$.*

Proof. **Single-Dimensional Extensions**

First let $x \in \mathcal{X} \setminus \mathcal{M}$, then $\mathbb{R}x \cap \mathcal{M} = \{0\}$ and any vector $y \in \mathbb{R}x + \mathcal{M}$ can be written uniquely as $\lambda x + m$ for some $\lambda \in \mathbb{R}$ and $m \in \mathcal{M}$. Given a linear functional f on $\mathcal{M}$, f can be extended to be a linear functional g on $\mathbb{R}x + \mathcal{M}$ by

$$g_\alpha(y) = g_\alpha(\lambda x + m) = \lambda\alpha + f(m) \quad \alpha \in \mathbb{R}$$

where $g|_{\mathcal{M}} = f$.

Bounded by Sublinear Functional

Let $x \in \mathcal{X} \setminus \mathcal{M}$ and $g_\alpha : \mathbb{R}x + \mathcal{M}$ with $\lambda x + m \mapsto \lambda\alpha + f(m)$. Then there exists an $\alpha \in \mathbb{R}$ such that $g_\alpha(y) \leq p(y)$ for all $y \in \mathcal{X}$.

Proof. Let $y_1, y_2 \in \mathcal{M}$, then

$$f(y_1) + f(y_2) \leq p(y_1 + y_2) \leq p(y_1 - x) + p(y_2 + x)$$

and

$$f(y_1) - p(y_1 - x) \leq p(y_2 + x) - f(y_2)$$

meaning that

$$\underbrace{\sup_{y \in \mathcal{M}} [f(y) - p(y - x)]}_{a} \leq \underbrace{\inf_{y \in \mathcal{M}} [p(y + x) - f(y)]}_{A}$$

Let $\alpha \in [a, A]$, then for any $\lambda > 0$ and $y \in \mathcal{M}$, using the fact that $\alpha \leq A$,

$$\begin{aligned} g_\alpha(\lambda x + y) &= \lambda\alpha + f(y) \\ &= \lambda\left[\alpha + f\left(\frac{y}{\lambda}\right)\right] \\ &\leq \lambda\left[p\left(\frac{y}{\lambda} + x\right) - f\left(\frac{y}{\lambda}\right) + f\left(\frac{y}{\lambda}\right)\right] \\ &= \lambda p\left(\frac{y}{\lambda} + x\right) \\ &= p(y + \lambda x) \end{aligned}$$

Now let $\lambda = -\mu < 0$, then using the fact that $\alpha \geq a$,

$$\begin{aligned} g_\alpha(\lambda x + y) &= f(y) - \mu\alpha \\ &= \frac{1}{\mu}\left[f\left(\frac{y}{\mu}\right) - \alpha\right] \\ &\leq \frac{1}{\mu}\left[f\left(\frac{y}{\mu}\right) - f\left(\frac{y}{\mu}\right) + p\left(\frac{y}{\mu} - x\right)\right] \\ &= \frac{1}{\mu}p\left(\frac{y}{\mu} - x\right) \\ &= p(y - \mu x) \\ &= p(y + \lambda x) \end{aligned}$$

and we have $g_\alpha(\lambda x + y) \le p(\lambda x + y)$ for any $\lambda \in \mathbb{R}$ and $y \in \mathcal{M}$.

Union of Subspaces

Let

$$\mathcal{F} = \{g : \mathcal{X} \to \mathbb{R} : g|_{\mathcal{M}} = f, g(x) \le p(x)\ \forall x\}$$

be the family of all linear extensions of f to subspaces of $\mathcal{X}$ that are bounded by p. Order $\mathcal{F}$ by $g \le h$ if h is an extension of g, then $\le$ is a partial order.

Let C be a chain in $\mathcal{F}$. Denote $\mathrm{Dom}(g)$ as the domain of g, and let

$$\mathcal{N} = \bigcup_{g \in C} \mathrm{Dom}(g)$$

then $\mathcal{N}$ is a subspace of $\mathcal{M}$: For any $x, y \in \mathcal{N}$ and $\lambda \in \mathbb{R}$, there exists $g_1, g_2 \in C$ such that $x \in \mathrm{Dom}(g_1)$, $y \in \mathrm{Dom}(g_2)$. Since C is a chain, assume without loss of generality that $g_1 \le g_2$, then $\mathrm{Dom}(g_2) \supset \mathrm{Dom}(g_1)$ and $\lambda x, x{+}y \in \mathrm{Dom}(g_1) \subset \mathcal{N}$.

Maximal Extension

Let $x \in \mathcal{N}$, then there exists $g \in C$ such that $x \in \mathrm{Dom}(g)$. Define $F(x) = g(x)$. First verify that F is well-defined: let $g_1, g_2 \in C$ such that $x \in \mathrm{Dom}(g_1), \mathrm{Dom}(g_2)$. Since C is a chain, assume without loss of generality that $g_1 \le g_2$, then g_2 is an extension of g_1 and $g_1(x) = g_2(x)$. Therefore the value of F on x is independent of the choice of g.

Now, let $x, y \in \mathcal{N}$ and $\lambda \in \mathbb{R}$, then there exists $g_1, g_2 \in C$ such that $x \in \mathrm{Dom}(g_1)$ and $y \in \mathrm{Dom}(g_2)$. Assume again that $g_2 \ge g_1$, and we have

$$\begin{aligned} F(\lambda x + y) &= g_2(\lambda x + y) \\ &= \lambda g_2(x) + g_2(y) \\ &= \lambda F(x) + F(y) \\ F(x) &= g_2(x) \le p(x) \end{aligned}$$

that $F \in \mathcal{F}$ being an upper bound of C.

Since every chain C in $\mathcal{F}$ has an upper bound, by Zorn's Lemma, $\mathcal{F}$ has a maximal element F. If $\mathrm{Dom}(F) \subsetneq \mathcal{X}$, then F can be extended to $\mathrm{Dom}(F) + \mathbb{R}x$ for some $x \in \mathcal{X} \setminus \mathrm{Dom}(F)$, which contradicts the fact that F is a maximal element of $\mathcal{F}$. Therefore $\mathrm{Dom}(F) = \mathcal{X}$, and F is an extension of f such that $F(x) \le p(x)$ for all $x \in \mathcal{X}$. □

Theorem 3.6. *Let $\mathcal{X}$ be a real vector space, $||\cdot||$ be a seminorm on $\mathcal{X}$, $\mathcal{M}$ be a subspace of $\mathcal{X}$, and f be a linear functional on $\mathcal{M}$ such that $|f(x)| \le ||x||$ for all $x \in \mathcal{M}$. Then there is a linear extension F of f to $\mathcal{X}$ such that $|F(x)| \le ||x||$ for all $x \in \mathcal{X}$.*

Proof. Let $x \in \mathcal{X}$, then

$$f(x) \le ||x|| \Leftrightarrow f(-x) \le ||x|| \Leftrightarrow |f(x)| \le ||x||$$

Since the seminorm is sublinear, applying the Hahn-Banach theorem yields the desired extension. □

Theorem 3.7 (Complex Version). *Let $\mathcal{X}$ be a complex vector space, p a seminorm on $\mathcal{X}$, $\mathcal{M}$ a subspace of $\mathcal{X}$, and f a complex linear functional on $\mathcal{M}$ such that $|f(x)| \leq p(x)$ for $x \in \mathcal{M}$. Then there exists an extension F of f to $\mathcal{X}$ such that $|F(x)| \leq p(x)$ for all $x \in \mathcal{X}$.*

Proof. Let $u = \mathrm{Re}\,(f)$, then there is a linear extension U of u to $\mathcal{X}$ with $|U(x)| \leq p(x)$. Let $F = U(x) - iU(ix)$, then F is a linear extension of f. For any non-zero $x \in \mathcal{X}$, let $\alpha = \overline{\mathrm{sgn} F(x)}$, then

$$F(x) = F(\alpha x) = \alpha F(x) = U(x) \leq p(x)$$

□

Theorem 3.8. *Let $\mathcal{X}$ be a normed space over $\mathbb{C}$, then*

1. *If $\mathcal{M}$ is a closed subspace of $\mathcal{X}$ and $x \in \mathcal{X} \setminus \mathcal{M}$, there exists $f \in \mathcal{X}^*$ such that $f(x) \neq 0$ and $f|_M = 0$.*
2. *If $\delta = \inf_{y \in \mathcal{M}} ||x - y||$, then f can be taken to satisfy $||f|| = 1$ and $f(x) = \delta$.*
3. *If $x \in \mathcal{X} \backslash \{0\}$, then there exists $f \in \mathcal{X}^*$ such that $||f|| = 1$ and $f(x) = ||x||$.*
4. *The bounded linear functionals on $\mathcal{X}$ separate points.*
5. *If $x \in \mathcal{X}$, define $\hat{x} : \mathcal{X}^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$. Then the map $x \mapsto \hat{x}$ is a linear isometry from $\mathcal{X}$ to $\mathcal{X}^{**}$. Moreover, the closure of its image $\hat{\mathcal{X}} = \{\hat{x} : x \in \mathcal{X}\}$ is a Banach space, known as the **completion** of $\mathcal{X}$. If $\mathcal{X}$ is already a Banach space, then $\overline{\hat{\mathcal{X}}} = \hat{\mathcal{X}}$.*

Proof. **Point Indicator**

Let $\mathcal{M}$ be a closed subspace of $\mathcal{X}$ and $x \in \mathcal{X} \setminus \mathcal{M}$, then since $\mathcal{M}$ is closed, $\delta = \inf_{y \in \mathcal{M}} ||x - y|| > 0$.

Define $f : \mathcal{M} + \mathbb{C}x \to \mathbb{C}$ by $m + \lambda x \mapsto \lambda\delta$, then since for any $\lambda \neq 0$,

$$\begin{aligned} |f(m + \lambda x)| = |\lambda|\,\delta \leq |\lambda|\,||x + y|| & \qquad \forall y \in \mathcal{M} \\ |\lambda|\,\delta \leq |\lambda| \left|\left|x + \frac{m}{\lambda}\right|\right| & \\ |\lambda|\,\delta \leq ||\lambda x + m|| & \end{aligned}$$

f is a bounded linear functional on $\mathcal{M} + \mathbb{C}x$.

Extend f to F on $\mathcal{X}$ using the Hahn-Banach theorem with $p(y) = ||y||$, then $F(y) \leq ||y||$ implies that $||F|| \leq 1$. For any $\varepsilon > 0$, there exists $y \in \mathcal{M}$ such that $||x - y|| < \delta + \varepsilon$, in which case

$$\frac{F(x - y)}{||x - y||} > \frac{\delta}{\delta + \varepsilon}$$

where sending $\varepsilon \to 0$ yields $\delta/(\delta - \varepsilon) \to 1$. Therefore $||f|| \geq 1$.

Setting $\mathcal{M} = \{0\}$ yields the second result.

Separating Points

Let $x, y \in \mathcal{X}$, $x \neq y$, then there exists a continuous indicator functional on $\mathcal{X}$ such that $f(x-y) \neq 0$. Therefore $f(x) \neq f(y)$ and $\mathcal{X}^*$ separates points.

Completion

Let $x \in \mathcal{X}$ and define $\hat{x} : \mathcal{X}^* \to \mathbb{C}$ by $f \mapsto f(x)$, then the map $\phi : \mathcal{X} \to \mathcal{X}^{**}$ where $x \mapsto \hat{x}$ is a linear isometry. Moreover, the closure of its image $\overline{\hat{\mathcal{X}}} = \{\hat{x} : x \in \mathcal{X}\}$ is a Banach space. If $\mathcal{X}$ is already a Banach space, then $\overline{\hat{\mathcal{X}}} = \hat{\mathcal{X}}$.

Proof. Let $\lambda \in \mathbb{C}$ and $x, y \in \mathcal{X}$, then

$$\left(\widehat{\lambda x + y}\right)(f) = f(\lambda x + y) = \lambda f(x) + f(y) = \lambda \hat{x}(f) + \hat{y}(f)$$

and f is linear. Now let $x \in \mathcal{X}$, then

$$||\hat{x}|| = \sup_{||f||=1} |f(x)| \leq ||x||$$

and since there exists $f \in \mathcal{X}^*$ such that $f(x) = 1$, $||\hat{x}|| = ||x||$.

Since $\mathcal{X}^{**} = L(\mathcal{X}^*, \mathbb{C})$ and $\mathbb{C}$ is complete, $\mathcal{X}^{**}$ is also complete. As $\overline{\hat{\mathcal{X}}}$ is closed, $\overline{\hat{\mathcal{X}}}$ is also complete.

Lastly, suppose that $\mathcal{X}$ is a Banach space and let $\{x_n\}_1^\infty$ be a Cauchy sequence in $\mathcal{X}$, then since the natural map $x \mapsto \hat{x}$ is continuous, $\widehat{\lim_{n\to\infty} x_n} = \lim_{n\to\infty} \hat{x}_n$. Therefore $\hat{\mathcal{X}}$ is complete. □

Theorem 3.9 (Separation of Points and Open Convex Sets)**.** *Let $C \subset \mathcal{X}$ be a non-empty open convex set and $x \notin C$. Then there exists a continuous linear functional $f \in \mathcal{X}^*$ such that $f(y) < f(x)$ for all $y \in C$. In other words, $[f = f(x)]$ separates $\{x\}$ and C.*

Proof. The gauge serves to separate points inside the set from ones outside of it. Extending an indicator function for a point outside of C using the gauge as a constraint allows the function to also separate the point from C.

First suppose that $0 \in C$, and let $g : \mathbb{R}x \to \mathbb{R}$ with $tx \mapsto t$, then since $x \notin C$, $p(x) > 1$, and $g(tx) = t \leq tp(x) = p(tx)$. By the Hahn-Banach theorem, there is an extension f of g to $\mathcal{X}$ such that $f(x) = 1$ and $f(y) \leq p(y)$ for all $y \in \mathcal{X}$.

As $f(y) \leq p(y) \leq M\,||y||$ for some $M > 0$, f is continuous, and $[f = 1]$ is closed. With $f(y) \leq p(y) < 1$ for all $y \in C$, $[f = 1]$ is a closed hyperplane that separates x and C.

Suppose that $0 \notin C$. Choose any $y \in C$ so $0 \in C - y$. Then there exists $f : \mathcal{X} \to \mathbb{R}$ such that $f(x - y) = 1$ and $f(z) < 1$ for all $z \in C - y$. For any

$z \in C$,

$$f(z) = f(z-y) + f(y) < f(x-y) + f(y) = f(x)$$

and f separates x from C. □

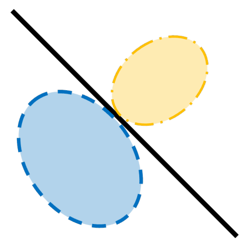

Figure 3.2: Separation of an Open Convex Set and a Convex Set

Theorem 3.10 (Hahn-Banach, First Geometric Form)**.** *Let $\mathcal{X}$ be a normed space over $\mathbb{R}$ and $A, B \subset X$ be non-empty convex sets such that $A \cap B \neq \emptyset$. If one of them is open, then there exists a closed hyperplane that separates A and B.*

Proof. Let $C = A - B$, then for any $a - b$ and $a' - b'$ in C,

$$t(a-b) + (1-t)(a'-b') = \underbrace{[ta + (1-t)a']}_{\in A} - \underbrace{[tb + (1-t)b']}_{\in B}$$

and since $C = \bigcup_{y \in B}(A - y)$, C is open. The collection of differences being convex allows it to be separated from 0 with a plane.

As $A \cap B = \emptyset$, $0 \notin C$. By the Hahn-Banach theorem, there exists $f \in \mathcal{X}^*$ that separates C from 0, i.e. $f(z) < 0 \forall z \in C$. Then

$$f(z) < f(y) \quad \forall x \in A, y \in B$$

Let $\alpha \in [\sup_{x \in A} f(x), \inf_{y \in B} f(y)]$, then

$$f(a) < \alpha < f(b) \quad \forall x \in A, y \in B$$

and $[f = \alpha]$ separates A from B. □

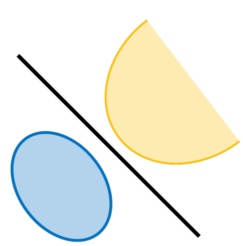

Figure 3.3: Separation of a Compact Convex Set and a Closed Convex Set

Theorem 3.11 (Hahn-Banach, Second Geometric Form). *Let $\mathcal{X}$ be a normed space over $\mathbb{R}$ and A, B be disjoint non-empty convex subsets. If A is closed and B is compact, then there exists a closed hyperplane that strictly separates A and B.*

Proof. Set $C = A - B$, then C is convex and $0 \notin C$.

Let $\{c_n\}_1^\infty \subset C$ be a convergent sequence in C with $c_n \to c$, write $c_n = a_n - b_n$ for some $a_n \in A$ and $b_n \in B$. Since B is compact, there exists a convergent subsequence $\{b_{n_k}\}_1^\infty$. This allows a subsequence $\{c_{n_k}\}_1^\infty$ with $c_{n_k} = a_{n_k} - b_{n_k}$ such that $c_{n_k} \to c$, $b_{n_k} \to b \in B$, and as addition is continuous, $a_{n_k} \to a \in A$. Therefore there exists $a \in A$ and $b \in B$ such that $c = a + b$, and C is closed.

The convex set being closed allows the existence of an open neighbourhood that separates 0 from C. This provides "room" for the function to interpolate between its value on C and on 0, and allows a strict separation.

Since $0 \notin C$, there exists $r > 0$ such that $B(0, r) \subset C^c$. Let $f \in \mathcal{X}^*$ such that

$$f(c) \leq f(rz) \quad \forall c \in C, z \in B(0,1)$$

As $f(c) \leq f(rz) = rf(z)$ for any $z \in B(0,1)$, $f(c) \leq -r\,||f||$. Take $\varepsilon = \frac{1}{2}r\,||f|| > 0$, then

$$\begin{aligned} f(x-y) &\leq -r\,||f|| \\ f(x) + \frac{1}{2}r\,||f|| &\leq f(y) - \frac{1}{2}r\,||f|| \\ f(x) + \varepsilon &\leq f(y) - \varepsilon \end{aligned}$$

Choose

$$\alpha \in [\sup f(x) + \varepsilon, \inf f(y) - \varepsilon]$$

then the hyperplane $[f = \alpha]$ strictly separates A from B. □

3.4 Hilbert Spaces

Definition 3.6 (Pre-Hilbert Space). *Let $\mathcal{H}$ be a complex vector space. $\mathcal{H}$ is a **pre-Hilbert space** if it is equipped with an inner product.*

Definition 3.7 (Hilbert Space). *Let $\mathcal{H}$ be a pre-Hilbert space, and define*

$$||x|| = \sqrt{\langle x, x \rangle}$$

*then $||x||$ is a norm on $\mathcal{H}$. If $\mathcal{H}$ is complete with respect to $||\cdot||$, then $\mathcal{H}$ is a **Hilbert space**.*

Proof. Since $\langle x, x \rangle = 0$ if and only if $x = 0$, $||x|| = 0$ if and only if $x = 0$. Let $\lambda \in \mathbb{C}$. then

$$\langle \lambda x, \lambda x \rangle = \lambda\overline{\lambda}\,\langle x, x \rangle = |\lambda|^2\,\langle x, x \rangle$$

Now let $x, y \in \mathcal{H}$, then

$$\begin{aligned}\langle x+y, x+y\rangle &= \langle x+y, x\rangle + \langle x+y, y\rangle \\ &= \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle \\ &= \langle x, x\rangle + \overline{\langle x, y\rangle} + \langle x, y\rangle + \langle y, y\rangle \\ &= \langle x, x\rangle + 2\mathrm{Re}\left(\langle x, y\rangle\right) + \langle y, y\rangle \\ &\leq \langle x, x\rangle + 2\left|\langle x, y\rangle\right| + \langle y, y\rangle \\ &\leq \langle x, x\rangle + 2\left|\left|x\right|\right|\left|\left|y\right|\right| + \langle y, y\rangle \\ &\leq \langle x, x\rangle + \langle y, y\rangle\end{aligned}$$

□

Theorem 3.12. *Let $\mathcal{H}$ be a pre-Hilbert space and $\{x_n\}_1^\infty, \{y_n\}_1^\infty \subset \mathcal{H}$. If $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Proof.

$$\begin{aligned}\left|\langle x_n, y_n\rangle - \langle x, y\rangle\right| &= \left|\langle x_n - x, y_n\rangle + \langle x, y_n\rangle - \langle x, y\rangle\right| \\ &= \left|\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle\right| \\ &\leq \left|\left|x_n - x\right|\right|\left|\left|y_n\right|\right| + \left|\left|x\right|\right|\left|\left|y_n - y\right|\right|\end{aligned}$$

which approaches 0 as $n \to \infty$. □

Theorem 3.13 (Projection onto Subspace)**.** *Let $\mathcal{H}$ be a Hilbert space, and $\mathcal{M} \subset \mathcal{H}$ be a closed subspace. Let*

$$\mathcal{M}^\perp = \{x \in \mathcal{H} : \langle x, y\rangle = 0 \quad \forall y \in \mathcal{M}\}$$

be its orthogonal complement, then $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ is an inner direct sum. Moreover, when decomposing $x \in \mathcal{H}$ to $y + z$ where $y \in \mathcal{M}$ and $z \in \mathcal{M}^\perp$, y and z are the unique elements of $\mathcal{M}$ and $\mathcal{M}^\perp$ whose distance to x is minimal.

Proof. In regular normed spaces, there is no good way of getting a sequence in $\mathcal{M}$ to approach a vector that minimises its distance from x. However, since the geometry of Hilbert spaces resemble Euclidean spaces thanks to the Pythagorean theorem and the Parallelogram law (and the space being complete), the exact vector that minimises distance may be constructed.

Decomposition

Let $x \in \mathcal{H}$, then there exists $y \in \mathcal{M}$ such that

$$\left|\left|x - y\right|\right| = \inf_{y' \in \mathcal{M}} \left|\left|x - y'\right|\right|$$

Proof. Let $x \in \mathcal{H}$, and let $\delta = \inf_{y \in \mathcal{M}} \left|\left|x - y\right|\right|$. Let $\{y_n\}_1^\infty \subset \mathcal{M}$ be a sequence such that $\left|\left|x - y_n\right|\right| \to \delta$. By the Parallelogram law,

$$2\left|\left|y_n - x\right|\right|^2 + 2\left|\left|y_m - x\right|\right|^2 = \left|\left|y_n - y_m\right|\right|^2 + \left|\left|y_n + y_m - 2x\right|\right|^2$$

where since $\frac{1}{2}(y_m + y_n) \in \mathcal{M}$,

$$\begin{aligned}||y_n - y_m||^2 &= 2\,||y_n - x||^2 + 2\,||y_m - x||^2 - ||y_n + y_m - 2x||^2 \\ &= 2\,||y_n - x||^2 + 2\,||y_m - x||^2 - 4\left|\left|\frac{1}{2}(y_n + y_m) - x\right|\right|^2 \\ &\leq 2\,||y_n - x||^2 + 2\,||y_m - x||^2 - 4\delta^2\end{aligned}$$

As $||y_n - x|| \searrow \delta$, there exists $N \in \mathbb{N}$ such that $||y_n - x|| < \delta + \varepsilon$ for all $n \geq N$, in which case

$$2\,||y_n - x||^2 + 2\,||y_m - x||^2 - 4\delta^2 < 4(\delta + \varepsilon)^2 - 4\delta^2$$

Therefore $||y_n - y_m||$ is a Cauchy sequence, and as $\mathcal{H}$ is complete, $y_n \to y \in \mathcal{H}$.

Orthogonal

Let $x \in \mathcal{H}$ and y as above, then $z = x - y \in \mathcal{M}^\perp$.

Proof. Let $z' \in \mathcal{M}$. If $\langle z, z' \rangle \notin \mathbb{R}$, then $\langle z, \operatorname{sgn}(\langle z, z' \rangle) z' \rangle \in \mathbb{R}$, so we can assume that $\langle z, z' \rangle \in \mathbb{R}$. Then the function

$$f(t) = ||z + tz'||^2 = \langle z, z \rangle + 2t\,\langle z, z' \rangle + t^2\,\langle z', z' \rangle$$

has a minimum at $t = 0$ (because $z + tz' = x - (y - tz')$, and y minimises $||x - y||$), meaning that $f'(0) = 2\,\langle z, z' \rangle = 0$ (bottom of parabola).

Unique

Let $x \in \mathcal{H}$ and y, z as above, then z is the unique vector in $\mathcal{M}^\perp$ and y is the unique vector in $\mathcal{M}$ that minimise distance to x.

Proof. Let $z' \in \mathcal{M}^\perp$, then by the Pythagorean theorem

$$||x - z'||^2 = \underbrace{||x - z||^2}_{=y\in\mathcal{M}} + \underbrace{||z - z'||^2}_{\in\mathcal{M}^\perp} \geq ||x - z||^2$$

with equality if and only if $z = z'$.

Let $y' \in \mathcal{M}$, then by the Pythagorean theorem again,

$$||x - y'||^2 = \underbrace{||x - y||^2}_{=z\in\mathcal{M}^\perp} + \underbrace{||y - y'||^2}_{\in\mathcal{M}} \geq ||x - y||^2$$

with equality if and only if $y = y'$.

Direct Sum

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$$

Proof. Since for each $x \in \mathcal{H}$ there exists $y \in \mathcal{M}$ and $z \in \mathcal{M}^\perp$ such that $x = y + z$, $\mathcal{H} = \mathcal{M} + \mathcal{M}^\perp$. Let $x \in \mathcal{M} \cap \mathcal{M}^\perp$, then $\langle x, x \rangle = 0$ and $x = 0$. $\square$

Definition 3.8 (Orthonormal Basis)**.** *Let $\mathcal{H}$ be a Hilbert space and $\{u_i\}_{i \in I} \subset \mathcal{H}$ be an orthonormal set, then the following are equivalent:*

1. ***Completeness**: If $\langle x, u_i \rangle = 0$ for all $i \in I$, then $x = 0$.*
2. *For each $x \in \mathcal{H}$, $x = \sum_{i \in I} \langle x, u_i \rangle u_i$, where the sum has countably many non-zero terms, and converges in the norm topology regardless of the ordering.*
3. ***Parseval's Identity**: $||x||^2 = \sum_{i \in I} |\langle x, u_i \rangle|^2$ for all $x \in \mathcal{H}$.*

*An orthonormal set satisfying these properties is an **orthonormal basis** of $\mathcal{H}$.*

Proof. **Countable Reduction**

Let $x \in \mathcal{H}$, then $\langle x, u_i \rangle \neq 0$ for only countably many is.

Proof. By Bessel's inequality, $||x||^2 \geq \sum_{i \in I} |\langle x, u_i \rangle|^2$ and $|\langle x, u_i \rangle|^2$ is non-zero for only countably many is.

Evaluating the Sum

Let $\{u_n\}_1^\infty$ be an enumeration of u_is where $\langle x, u_i \rangle \neq 0$, then

$$x = \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$$

Proof. By Bessel's inequality, the series $\sum_{n \in \mathbb{N}} |\langle x, u_n \rangle|^2$ converges. By the Pythagorean theorem,

$$\left|\left|\sum_{k=n}^{m} \langle x, u_k \rangle u_k\right|\right| = \sum_{k=n}^{m} |\langle x, u_k \rangle|^2 \leq \sum_{k=n}^{\infty} |\langle x, u_k \rangle|^2 \to 0$$

the series $\sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$ is Cauchy, and therefore converges. Let $y = x - \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$, then

$$\begin{aligned}
\langle y, u_j \rangle &= \left\langle x - \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n, u_j \right\rangle \\
&= \lim_{n \to \infty} \left\langle x - \sum_{k=1}^{n} \langle x, u_k \rangle u_k, u_j \right\rangle \\
&= \langle x, u_j \rangle - \langle \langle x, u_j \rangle u_j, u_j \rangle && (\forall n \geq j) \\
&= 0
\end{aligned}$$

for all $j \in \mathbb{N}$ and as $x, \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n \perp u_i$ for all $u_i \notin \{u_n : n \in \mathbb{N}\}$, $y \perp u_i$ for all $i \in I$. Therefore $y = 0$, and $x = \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$.

Decomposition to Parseval's Identity

Suppose that each $x \in \mathcal{H}$ can be written as $\sum_{i \in I} \langle x, u_i \rangle u_i$, then

$$||x||^2 = \sum_{i \in I} |\langle x, u_i \rangle|^2 \quad \forall x \in \mathcal{H}$$

Proof. Let $\{u_n\}_1^\infty$ be an enumeration of u_is such that $\langle x, u_i \rangle \neq 0$, then

$$\sum_{i \in I} |\langle x, u_i \rangle|^2 = \sum_{n \in \mathbb{N}} |\langle x, u_n \rangle|^2 = \sum_{n=1}^{\infty} |\langle x, u_n \rangle|^2$$

By the Pythagorean theorem,

$$||x||^2 = \left|\left| x - \sum_{k=1}^{n} \langle x, u_k \rangle u_k \right|\right|^2 + \left|\left| \sum_{k=1}^{n} \langle x, u_k \rangle u_k \right|\right|^2$$

Since $x = \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$,

$$||x||^2 - \left|\left| \sum_{k=1}^{n} \langle x, u_k \rangle u_k \right|\right|^2 = \left|\left| x - \sum_{k=1}^{n} \langle x, u_k \rangle u_k \right|\right|^2 \to 0$$

as $n \to \infty$. Therefore

$$||x||^2 = \left|\left| \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n \right|\right|^2 = \sum_{n=1}^{\infty} |\langle x, u_n \rangle|^2$$

Parseval's Identity to Completeness

Suppose that Parseval's identity holds, then $\langle x, u_i \rangle = 0$ for all $i \in I$ implies that $x = 0$.

Proof. Let $x \in \mathcal{H}$ and suppose that $\langle x, u_i \rangle = 0 \forall i \in I$, then

$$||x||^2 = \sum_{i \in I} |\langle x, u_i \rangle|^2 = 0$$

and $x = 0$.

□

Theorem 3.14. *Every Hilbert space has an orthonormal basis.*

Proof. A maximal orthonormal set is complete, and therefore an orthonormal basis.

Theorem 3.15. *Let $\mathcal{H}$ be a Hilbert space. $\mathcal{H}$ is separable if and only if $\mathcal{H}$ has a countable orthonormal basis, in which case every orthonormal basis is countable.*

Proof. **Separable to Countable Orthonormal Basis**

Suppose that $\mathcal{H}$ is separable, then $\mathcal{H}$ has a countable orthonormal basis.

Proof. Let $\{x_n\}_1^\infty$ be a countable dense subset of $\mathcal{H}$. Construct inductively a linearly independent subset of $\{x_n\}_1^\infty$, then apply the Gram-Schmidt Process to obtain an orthonormal set $\{u_n\}_1^\infty$, then the span of $\{u_n\}_1^\infty$ is dense in $\mathcal{H}$.

Let $x \in \mathcal{H}$, then there exists $\{x_n\}_1^\infty$ such that $x_n \to x$ with $x_n = \sum_{k=1}^\infty \langle x_n, u_k \rangle u_k$. For any ε, choose m and n such that

$$\sum_{k=1}^m \langle x, u_k \rangle u_k \approx \sum_{k=1}^m \langle x_n, u_k \rangle u_k \approx \sum_{k=1}^\infty \langle x_n, u_k \rangle u_k = x_n \approx x$$

where each $\approx$ represents a distance of less than $\varepsilon/4$, then $\sum_{k=1}^m \langle x, u_k \rangle u_k \to x$ as $m \to \infty$.

Countable Orthonormal Basis to Separable

If $\mathcal{H}$ has a countable orthonormal basis, then $\mathcal{H}$ is separable.

Proof. Let $F \subset \mathbb{C}$ be a countable dense subset, then the collection of finite linear combinations of $\{x_n\}_1^\infty$ with F coefficients would be our countable dense subset. □

3.5 Differential Calculus

Definition 3.9 (Fréchet Differentiable). *Let E, F be Banach/topological vector spaces, and $U \subset E, V \subset F$ be open sets. A function $f : U \to V$ is **Fréchet-differentiable** at a point $x \in U$, if there exists a continuous linear map $\lambda \in L(E, F)$ and ψ such that*

$$f(x + y) = f(x) + \lambda y + \psi(y)$$

where ψ is little-o of y/tangent to 0.

Theorem 3.16 (Fréchet Derivative). *Let $x \in E$ and $f : U \to V$ be differentiable at x, then define the **Fréchet derivative** $Df(x)$ to be the λ satisfying the above condition. λ is uniquely determined by f and x.*

Proof. Let $\lambda_1, \lambda_2 \in L(E, F)$ be continuous linear maps satisfying the desired property. For any $v \in E : ||v|| = 1$, let $t > 0$ such that $B(x, t) \subset U$, then

$$\begin{aligned} f(x + tv) - f(x) &= \lambda_1(tv) + o_1(tv) \\ &= \lambda_2(tv) + o_2(tv) \\ \lambda_1(tv) - \lambda_2(tv) &= o_2(tv) - o_1(tv) \\ t(\lambda_1 v - \lambda_2 v) &= o_2(tv) - o_1(tv) \\ (\lambda_1 - \lambda_2)v &= \frac{o_2(tv) - o_2(tv)}{||tv||} \end{aligned}$$

Taking the limit on both sides yields that

$$\lim_{t \to 0} (\lambda_1 - \lambda_2)(v) = \lim_{t \to 0} \frac{o_2(tv) - o_2(tv)}{||tv||} = 0$$

Therefore $\lambda_1 v = \lambda_2 v$ for all unit vectors v, which extends to all elements of E. □

Definition 3.10 (Second Derivative)**.** *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $U \subset \mathcal{X}$ be open, and $f : U \to \mathcal{Y}$ be a differentiable function, then*

$$Df : U \to L(\mathcal{X}, \mathcal{Y})$$

Since the space of bounded linear maps is complete, differentiating Df yields the ***second derivative***

$$D^2 f : U \to L(\mathcal{X}, L(\mathcal{X}, \mathcal{Y}))$$

As maps in $L(\mathcal{X}, L(\mathcal{X}, \mathcal{Y}))$ are separately continuous (with respect to x inputs), we identify $L(\mathcal{X}, L(\mathcal{X}, \mathcal{Y})) = L^2(\mathcal{X}, \mathcal{Y})$ as the space of continuous bilinear maps.

Definition 3.11 (Higher Derivatives)**.** *Let $p \in \mathbb{N}$, then define inductively the p-th derivative*

$$D^p f(x) = D(D^{p-1} f)(x)$$

with $D^p f(x) \in L^p(\mathcal{X}, \mathcal{Y})$ is continuous and multilinear. If

$$D^k f : U \to L^k(\mathcal{X}, \mathcal{Y})$$

exists and is continuous for each $k \leq p$, then $f \in C^p$.

Theorem 3.17. *Let $\{v_k\}_1^n$ be fixed elements of $\mathcal{X}$. If f is p times differentiable on U, and let*

$$g(x) = D^{n-1} f(x)(v_2, \cdots, v_n)$$

then g is differentiable on U, and

$$Dg(x)(v) = D^p f(x)(v, \cdots, v_n)$$

Proof. Consider g as the composition between

$$D^{n-1} f : U \to L^{n-1}(\mathcal{X}, \mathcal{Y}) \quad \lambda : L^{n-1}(\mathcal{X}, \mathcal{Y}) \to \mathcal{Y}$$

where λ is the evaluation map at $(v_2, \cdots, v_n)$. This makes λ continuous and linear, which allows differentiating the decomposition

$$D(\lambda \circ D^{n-1} f) = \lambda \circ D^n f$$

Therefore

$$Dg(x)(v) = (\lambda \circ D^n f)(x)(v) = (D^n f(x) v)(v_2, \cdots, v_n)$$

□

Theorem 3.18. *Let $f : U \to \mathcal{Y}$ be p times differentiable and $\lambda : \mathcal{Y} \to \mathcal{Z}$ be a bounded linear map. Then for any $x \in U$,*

$$D^p(\lambda \circ f)(x) = \lambda \circ D^p f(x)$$

Theorem 3.19 (The Second Derivative is Symmetric)**.** *Let $U \subset \mathcal{X}$ be open, $f : U \to \mathcal{Y}$ be twice differentiable with $D^2 f$ being continuous. Then for each $x \in U$, the bilinear map $D^2 f(x)$ is symmetric for all $v, w \in \mathcal{X}$.*

Proof. Let $r > 0$ such that $B(x, 2r) \subset U$. Let $v, w \in \mathcal{X}$ such that $||v||, ||w|| < r$. Denote

$$g(x) = f(x + v) - f(x)$$

Then

$$\begin{aligned}
&f(x + v + w) - f(x + w) - f(x + v) + f(x) \\
&= g(x + w) - g(x) \\
&= \int_0^1 Dg(x + tw)(w)dt \\
&= \int_0^1 \left[Df(x + v + tw) - Df(x + tw)\right](w)dt \\
&= \int_0^1 \int_0^1 D^2 f(x + sv + tw) \cdot (v)ds \cdot (w)dt
\end{aligned}$$

by applying the mean value theorem twice. Let

$$\psi(sv, tw) = D^2 f(x + sv + tw) - D^2 f(x)$$

then

$$\begin{aligned}
g(x + w) - g(x) &= \int_0^1 \int_0^1 D^2 f(x + sv + tw)(v, w)dsdt \\
&= \int_0^1 \int_0^1 D^2 f(x)(v, w)dsdt \\
&+ \int_0^1 \int_0^1 \psi(sv, tw)(v, w)dsdt \\
&= D^2 f(x)(v, w) + \underbrace{\int_0^1 \int_0^1 \psi(sv, tw)(v, w)dsdt}_{\phi(v,w)}
\end{aligned}$$

where

$$||\phi(v, w)|| \le \sup_{s,t} ||\phi(sv, tw)|| \cdot ||v|| \cdot ||w||$$

Swapping the role of v and w in the above example, we can work with

$$g_w(x) = f(x + w) - f(x)$$

and

$$\begin{aligned}
&f(x + v + w) - f(x + w) - f(x + v) + f(x) \\
&= g_w(x + v) - g_w(x) \\
&= D^2 f(x)(w, v) + \phi_w(v, w)
\end{aligned}$$

where

$$||\phi_w(v,w)|| \leq \sup_{s,t} ||\psi_w(sv,tw)|| \cdot ||v|| \cdot ||w||$$

The two separate ways of writing the same expression yields

$$D^2f(x)(v,w) - D^2f(x)(w,v) = \phi(v,w) - \phi_w(v,w)$$

where since D^2 is continuous,

$$\begin{aligned}\lim_{(v,w)\to 0} \phi(v,w) &= \lim_{(v,w)\to 0} \phi_w(v,w) \\ &= D^2f(x+v+w) - D^2f(x) \\ &= 0\end{aligned}$$

Meaning that $D^2f(x)(v,w) - D^2f(x)(w,v) = 0$. □

Theorem 3.20 (Higher Derivatives are Symmetric)**.** *Let $f \in C^p$ on U. Then for each $x \in U$, the map $D^pf(x)$ is symmetric.*

Proof. With induction on p. Suppose that $D^{p-1}f(x)$ is symmetric and let $g = D^{p-2}f$, then

$$D^2g(x)(v,w) = D^2g(x)(w,v)$$

Since $D^pf = D^2D^{p-2}F$,

$$\begin{aligned}D^pf(x)(v_1,\cdots,v_p) &= (D^2D^{p-2}f(x))(v_1,v_2)\cdot(v_3,\cdots,v_p) \\ &= (D^2D^{p-2}f(x))(v_2,v_1)\cdot(v_3,\cdots,v_p) \\ &= D^pf(x)(v_2,v_1,\cdots,v_p)\end{aligned}$$

we can swap the first two inputs to the function. By the inductive hypothesis, we can also permute the last $p-1$ inputs to the function.

As any permutation in S_p can be written as $(12)\cdot\sigma$ for some $\sigma \in S_{p-1}$, permutations do not affect the value of $D^pf(x)$. □

Theorem 3.21. *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, then the map*

$$L(\mathcal{X},\mathcal{Y}) \times \mathcal{X} \quad (T,x) \mapsto Tx$$

is bounded and bilinear. Let $J = [a,b]$ be a closed interval and $\alpha : J \to L(\mathcal{X},\mathcal{Y})$ be a continuous map, then we can integrate along the curve

$$\int_a^b \alpha(t)dt \in L(\mathcal{X},\mathcal{Y})$$

If α is differentiable, then $\frac{d\alpha}{dt}(t) \in L(\mathcal{X},\mathcal{Y})$ as well.

Theorem 3.22. *Let $\alpha : J \to L(\mathcal{X}, \mathcal{Y})$ be a continuous map and $x \in \mathcal{X}$, then*

$$\int_a^b \alpha(t)(x)dt = \int_a^b \alpha(t)dt \cdot x$$

Proof. The map $\lambda \mapsto \lambda x$ is linear and bounded. Therefore it can be taken out of the integral.

Theorem 3.23 (Mean Value Theorem)**.** *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $U \subset \mathcal{X}$ be open and $x \in U$. Let $y \in \mathcal{X}$, and $f : U \to \mathcal{Y}$ be a C^1 map. If the line $\{x + ty : t \in [0,1]\} \subset U$, then*

$$f(x+y) - f(x) = \int_0^1 Df(x+ty)(y)dt = \int_0^1 Df(x+ty)dt \cdot y$$

Proof. Let $g(t) = f(x+ty)$, then $Dg(t) = Df(x+ty) \circ y$ by the chain rule. Since g is continuous on $[0,1]$, by the Fundamental Theorem of Calculus,

$$\begin{aligned} g(1) - g(0) &= \int_0^1 Dg(t)dt \\ f(x+y) - f(x) &= \int_0^1 Df(x+ty)(t)dt \\ &= \int_0^1 Df(x+ty)dt \cdot y \end{aligned}$$

□

Theorem 3.24. *Let $U \subset \mathcal{X}$ be open and $x, z \in U$ such that the line segment*

$$L = \{tx + (1-t)z : t \in [0,1]\}$$

lies in U. If $f \in C^1$, then

$$||f(z) - f(x)|| \leq ||z - x|| \cdot \sup_{v \in L} ||Df(v)||$$

Proof. By the mean value theorem with $y = z - x$,

$$\begin{aligned} f(z) - f(x) &= \int_0^1 Df(x + t(z-x))dt \cdot (z-x) \\ ||f(z) - f(x)|| &= \left|\left| \int_0^1 Df(x + t(z-x)) \cdot (z-x)dt \right|\right| \\ &\leq \int_0^1 ||Df(x + t(z-x)) \cdot (z-x)||\, dt \\ &\leq \int_0^1 \sup_{t \in [0,1]} ||Df(x + t(z-x))|| \cdot ||z-x||\, dt \\ &= \sup_{z \in L} ||Df(z)|| \cdot ||z-x|| \end{aligned}$$

where the supremum exists since $f \in C^1$ and $Df \in C^0$. □

Theorem 3.25. *Let $U \subset \mathcal{X}$ be open and $x, z, x_0 \in U$ such that the line segment*

$$L = \{tx + (1-t)z : t \in [0,1]\}$$

lies in U. Then

$$||f(z) - f(x) - Df(x_0)(z-x)|| \leq ||z-x|| \sup_{v \in L} ||Df(v) - Df(x_0)||$$

Proof.

$$\begin{aligned}
&f(z) - f(x) \\
&= \int_0^1 Df(x + t(z-x))(z-x)dt \\
&= \int_0^1 \left[Df(x+t(z-x)) - Df(x_0) + Df(x_0)\right](z-x)dt \\
&= Df(x_0)(z-x) + \int_0^1 \left[Df(x+t(z-x)) - Df(x_0)\right](z-x)dt \\
&\leq Df(x_0) + \sup_{z \in L} ||Df(v) - Df(x_0)|| \cdot ||z-x||
\end{aligned}$$

□

3.6 Banach Algebra

Definition 3.12 (Banach Algebra)**.** *Let $(\mathcal{A}, \times)$ be an algebra over $\mathbb{C}$ and $||\cdot|| : \mathcal{A} \to [0, \infty)$ be a norm. $(\mathcal{A}, \times, ||\cdot||)$ is a* ***Banach algebra*** *if $(\mathcal{A}, ||\cdot||)$ is a Banach space and*

$$||x \times y|| \leq ||x|| \cdot ||y|| \quad \forall x, y \in \mathcal{A}$$

Definition 3.13 (Unital Banach Algebra)**.** *A Banach algebra $\mathcal{A}$ is* ***unital*** *if it has a multiplicative identity.*

Definition 3.14 (Invertible)**.** *Let $\mathcal{A}$ be a unital Banach algebra, then an element $x \in \mathcal{A}$ is* ***invertible*** *if $\exists x^{-1} \in \mathcal{A} : xx^{-1} = x^{-1}x = e$.*

Theorem 3.26 (Power Rule)**.** *Let $\mathcal{A}$ be a Banach algebra, then*

$$Dx^n(x_0)(h) = \sum_{k=1}^{n} x^{k-1} Dx(x_0)(h) x^{n-k}$$

*for all $n \in \mathbb{N}$. Since the derivative Dx is constant, $x \mapsto x^n$ is in C^∞ for all $n \in \mathbb{N}$. **

Proof. For $n = 1$, $x \mapsto x$ has constant derivative I everywhere, so it is also in C^∞.

Now suppose that x^n is C^∞ with

$$Dx^n(x_0)(h) = \sum_{k=1}^{n} x_0^{k-1} \cdot h \cdot x_0^{n-k}$$

then by the product rule,

$$\begin{aligned} Dx^{n+1}(x_0)(h) &= D(x \cdot x^n)(x_0)(h) \\ &= Dx(x_0)(h) \cdot x^n(x_0) + x(x_0) \cdot Dx^n(x_0)(h) \\ &= h \cdot x_0^{(n+1)-1} + \sum_{k=1}^{n} x_0^{k} \cdot h \cdot x_0^{n-k} \\ &= h \cdot x_0^{(n+1)-1} + \sum_{k=2}^{n+1} x_0^{k-1} \cdot h \cdot x_0^{(n+1)-k} \\ &= \sum_{k=1}^{n+1} x_0^{k-1} \cdot h \cdot x_0^{(n+1)-k} \end{aligned}$$

□

Theorem 3.27. *Let $\mathcal{A}$ be a unital Banach algebra. Then*

1. *If $|\lambda| > ||x||$, then $\lambda e - x$ is invertible where $\sum_{i \in \mathbb{N} \cup \{0\}} \lambda^{-n-i} x^i$ is the inverse.*
2. *If x is invertible and $||y|| < ||x^{-1}||^{-1}$, then $x - y$ is invertible where $x^{-1} \sum_{i \in \mathbb{N} \cup \{0\}} (yx^{-1})^i$ is the inverse.*
3. *If x is invertible and $||y|| < \frac{1}{2}||x^{-1}||^{-1}$, then $||(x-y)^{-1} - x^{-1}|| < 2||x^{-1}||^2||y||$.*
4. *The collection of invertible elements in $\mathcal{A}$ is open, and the mapping $x \mapsto x^{-1}$ is continuous on it.*

Proof. Let $\lambda \in \mathbb{C}$ and $x : ||x|| < |\lambda|$, then $y = \lambda^{-1} x$ satisfies $||y|| < 1$. We have

$$\begin{aligned} (e - y)^{-1} &= \sum_{i \in \mathbb{N} \cup \{0\}} y^i \\ (e - y)^{-1} &= \sum_{i \in \mathbb{N} \cup \{0\}} \lambda^{-n} x^i \\ (\lambda e - x)^{-1} &= \sum_{i \in \mathbb{N} \cup \{0\}} \lambda^{-n-1} x^i \end{aligned}$$

Now let $y : ||y|| < ||x^{-1}||^{-1}$, then

$$\begin{aligned} x - y &= (e - yx^{-1})x \quad ||yx^{-1}|| < 1 \\ (e - yx^{-1})^{-1} &= \sum_{i \in \mathbb{N} \cup \{0\}} (yx^{-1})^i \\ (x - y)^{-1} &= x^{-1} \sum_{i \in \mathbb{N} \cup \{0\}} (yx^{-1})^i \end{aligned}$$

Then let $y : ||y|| < \frac{1}{2}||x^{-1}||^{-1}$, and

$$\begin{aligned} &\left|\left| x^{-1} \sum_{i \in \mathbb{N} \cup \{0\}} (yx^{-1})^i - x^{-1} \right|\right| \\ &= \left|\left| x^{-1} \sum_{i \in \mathbb{N}^+} (yx^{-1})^i \right|\right| \\ &\leq \left|\left| x^{-1} \right|\right| \cdot \left|\left| \sum_{i \in \mathbb{N}^+} (yx^{-1})^i \right|\right| \\ &\leq \left|\left| x^{-1} \right|\right| \cdot \sum_{i \in \mathbb{N}^+} ||yx^{-1}||^i \\ &\leq \left|\left| x^{-1} \right|\right| \cdot \sum_{i \in \mathbb{N}^+} ||y||^i \cdot \left|\left| x^{-1} \right|\right|^i \\ &< \left|\left| x^{-1} \right|\right| \cdot ||y|| \ \cdot ||x||^{-1} \sum_{i \in \mathbb{N}} \frac{1}{2^i} \left|\left| x^{-1} \right|\right|^{-i} \cdot \left|\left| x^{-1} \right|\right|^i \\ &= \left|\left| x^{-1} \right|\right| \cdot ||y|| \ \cdot ||x||^{-1} \sum_{i \in \mathbb{N}} \frac{1}{2^i} \\ &= 2 \left|\left| x^{-1} \right|\right|^2 ||y|| \end{aligned}$$

Finally, let $\Omega \subset \mathcal{A}$ be the collection of all invertible elements in $\mathcal{A}$, and take $x \in \Omega$, and $y : ||y|| < ||x^{-1}||^{-1} \Leftrightarrow d(x, x - y) < ||x^{-1}||^{-1}$. Then

$$B\left(x, \left|\left| x^{-1} \right|\right|^{-1}\right) \subseteq \Omega$$

For continuity, let $\varepsilon > 0$ and $\delta < \frac{1}{2}\varepsilon \left|\left| x^{-1} \right|\right|^{-2}$, then

$$\begin{aligned} ||y|| < \frac{1}{2}\varepsilon^2 \left|\left| x^{-1} \right|\right|^{-1} \Leftrightarrow d(x, x - y) &< \frac{1}{2}\varepsilon^2 \left|\left| x^{-1} \right|\right|^{-1} \\ d(x^{-1}, (x - y)^{-1}) &< 2||x^{-1}||^2||y|| \\ &< \frac{1}{2}\varepsilon \left|\left| x^{-1} \right|\right|^{-2} \cdot 2||x^{-1}||^2 \\ &= \varepsilon \end{aligned}$$

□

Theorem 3.28 (Geometric Series)**.** *Let $\mathcal{A}$ be a unital Banach algebra and x such that $||e - x|| < 1$, then x is invertible with*

$$x^{-1} = \sum_{n=0}^{\infty}(e-x)^n$$

Proof. Firstly, since $||e - x|| < 1$, the series converges absolutely and has a limit. Now,

$$\begin{aligned} (e-x) \cdot \sum_{n=0}^{\infty}(e-x)^n &= \sum_{n=0}^{\infty}(e-x)^n - e \\ \sum_{n=0}^{\infty}(e-x)^n \cdot (e-x) &= \sum_{n=0}^{\infty}(e-x)^n - e \end{aligned}$$

we have $\sum_{n=0}^{\infty}(e-x)^n \cdot x = x \cdot \sum_{n=0}^{\infty}(e-x)^n = e$. Therefore x is invertible with the given inverse. □

Theorem 3.29. *Let $\mathcal{A}$ be a unital Banach algebra and x be an invertible element, then y is invertible for all $y \in B(x, \left|\left|x^{-1}\right|\right|^{-1})$ with $y^{-1} = (x^{-1}y)^{-1}x^{-1}$*

Proof. If $||x - y|| < \left|\left|x^{-1}\right|\right|^{-1}$, then

$$\left|\left|e - x^{-1}y\right|\right| = \left|\left|x^{-1}(x-y)\right|\right| \leq \left|\left|x^{-1}\right|\right| \cdot ||x-y|| < 1$$

and $x^{-1}y$ is invertible. In which case,

$$\left[(x^{-1}y)^{-1}x^{-1}\right] \cdot y = e$$

and

$$\begin{aligned} (x^{-1}y)(x^{-1}y)^{-1} &= e \\ y(x^{-1}y)^{-1} &= x \\ y \cdot \left[(x^{-1}y)^{-1}x^{-1}\right] &= e \end{aligned}$$

giving the inverse of y. □

Theorem 3.30 (Smoothness of the Inverse Near Identity)**.** *Let $\mathcal{A}$ be a unital Banach algebra, then the inversion map*

$$B(e,1) \to \mathcal{A} \quad x \mapsto x^{-1} = \sum_{n=0}^{\infty}(e-x)^n$$

is C^∞.

Proof. First note that the unit is not isolated, as for any $\varepsilon > 0$, $||e - (1+\varepsilon)e|| = \varepsilon$.

Since the coefficients on the terms is 1 for all n, the series has a radius of convergence of 1, and is therefore C^∞ on $B(e, 1)$. □

Theorem 3.31 (Smoothness of the Inverse)**.** *Let $\mathcal{A}$ be a unital Banach algebra, and y be an invertible element, then the inversion map is C^∞ at y.*

Proof. First note that no invertible element is isolated, as for any y invertible (so non-zero) and $\varepsilon > 0$, $z = y \cdot (1+\varepsilon/\,||y||)$ satisfies $||y - z|| = ||y \cdot \varepsilon/\,||y|||| = \varepsilon$.

Let $z \in \mathcal{A}$ with $||y - z|| < \left|\left|y^{-1}\right|\right|^{-1}$, then $\left|\left|e - y^{-1}z\right|\right| < 1$ and $z^{-1} = (y^{-1}z)^{-1} \cdot y^{-1}$. We can express the inversion map as $T \circ I \circ S$ where

$$\begin{aligned} S &: & B(y, \left|\left|y^{-1}\right|\right|^{-1}) \to B(e, 1) \quad & x \mapsto y^{-1}x \\ I &: & B(e, 1) \to \mathcal{A} \quad & x \mapsto x^{-1} \\ T &: & \mathcal{A} \to \mathcal{A} \quad & x \mapsto xy^{-1} \end{aligned}$$

Since the first and third map are linear, they are C^∞. As the inversion map is a composition of C^∞ functions, it is C^∞ on $B(y, \left|\left|y^{-1}\right|\right|^{-1})$. □

3.7 Inverse Function Theorem

Theorem 3.32 (Inverse Function Theorem (Simplified))**.** *Let $\mathcal{X}$ be a Banach space, $U \subset \mathcal{X}$ be an open set containing 0 and $f : U \to \mathcal{X}$ be a C^p map with $p \geq 1$. Suppose that $Df(0) = I$ is the identity and $f(0) = 0$, then f is a local C^p-isomorphism at 0.*

Proof. **Neighbourhood**

There exists $r > 0$ such that $Df(x)$ is invertible for all $x \in B(0, r)$. Moreover, if $g(x) = x - f(x)$, then $g(\overline{B(0, r)}) \subset \overline{B(0, r/2)}$.

Proof. Since $Dg(0) = 0$ and $f \in C^p$, by continuity, there exists $r > 0$ such that

$$\begin{aligned} ||Dg(x)|| &< \frac{1}{2} \\ ||Df(x) - I|| &< \frac{1}{2} \end{aligned}$$

for all x with $||x|| \leq r$. Since the space of isomorphisms is open, and $||Df(x) - I|| < \left|\left|I^{-1}\right|\right|^{-1}$, $Df(x)$ is invertible for all x with $||x|| < r$.

For any $x \in \overline{B(0, r)}$, by the mean value theorem,

$$||g(x)|| = ||g(x) - g(0)|| \leq ||x|| \cdot \sup_{t \in [0,1]} ||Dg(tx)|| \leq \frac{1}{2} ||x||$$

we have $x \in \overline{B(0, r/2)}$.

Contraction

The map $g : \overline{B(0,r)} \to \overline{B(0,r)}$ is a contractor.

Proof. Let $x_1, x_2 \in \overline{B(0,r)}$, then by the mean value theorem again

$$||g(x_1) - g(x_2)|| \leq ||x_1 - x_2|| \cdot \sup_{y \in \overline{B(0,r)}} ||Dg(y)|| < \frac{1}{2} ||x_1 - x_2||$$

Inverse

For any $y \in \overline{B(0, r/2)}$, there exists a unique $x \in \overline{B(0,r)}$ such that $f(x) = y$.

Proof (see Fixed Point Problem). Let $y \in \overline{B(0, r/2)}$ and $g_y(x) = x - f(x) + y$, then $f(x) = y$ if and only if $g_y(x) = x$. If $||y|| \leq r/2$ and $||x|| \leq r$,

$$||g_y(x)|| = ||g(x) + y|| \leq \frac{1}{2} ||x|| + ||y|| \leq r$$

and g_y maps $\overline{B(0,r)}$ to $\overline{B(0,r)}$. Now for any $x_1, x_2 \in \overline{B(0,r)}$,

$$||g_y(x_1) - g_y(x_2)|| = ||g(x_1) - g(x_2)||$$

Since g is a contractor, so is g_y. By the Banach fixed-point theorem, g_y has a unique fixed point $x \in \overline{B(0,r)}$, which is the solution to $f(x) = y$.

Continuity

Let $\phi : \overline{B(0, r/2)} \to \overline{B(0,r)}$ be the inverse of f, then ϕ is continuous.

Proof. Let $x_1, x_2 \in \overline{B(0,r)}$, then

$$x_1 - x_2 = f(x_1) + g(x_1) - f(x_2) - g(x_2)$$

and

$$\begin{aligned} ||x_1 - x_2|| &\leq ||f(x_1) - f(x_2)|| + ||g(x_1) - g(x_2)|| \\ &\leq 2 \, ||f(x_1) - f(x_2)|| \end{aligned}$$

So for any $y_1, y_2 \in \overline{B(0, r/2)}$,

$$||\phi(y_1) - \phi(y_2)|| \leq 2 \, ||y_1 - y_2||$$

and ϕ is continuous.

Differentiability

Let $\phi : \overline{B(0, r/2)} \to \overline{B(0,r)}$ be the inverse of f, then ϕ is differentiable with $D\phi(x) = Df(x)^{-1}$.

Proof. Let $x_1, x_2 \in \overline{B(0, r/2)}$, and $y_1 = f(x_1)$, $y_2 = f(x_2)$, then

$$\begin{aligned}&\left|\left|\phi(y_1) - \phi(y_2) - Df(x_2)^{-1}(y_1 - y_2)\right|\right| \\ &= \left|\left|x_1 - x_2 - Df(x_2)^{-1}(y_1 - y_2)\right|\right| \\ &\leq ||Df(x_2)|| \cdot ||Df(x_2)(x_1 - x_2) - (y_1 - y_2)||\end{aligned}$$

By the differentiability of f,

$$\begin{aligned}f(x_1) &= f(x_2) + Df(x_2)(x_1 - x_2) + o(x_1 - x_2) \\ f(x_1) - f(x_2) &= Df(x_2)(x_1 - x_2) + o(x_1 - x_2)\end{aligned}$$

so

$$\begin{aligned}&||Df(x_2)|| \cdot ||Df(x_2)(x_1 - x_2) - (y_1 - y_2)|| \\ &= ||Df(x_2)|| \cdot ||o(x_1 - x_2)||\end{aligned}$$

is little-o of $(x_1 - x_2)$. Since as shown before,

$$||x_1 - x_2|| \leq 2\,||y_1 - y_2||$$

meaning that

$$\begin{aligned}\frac{||o(x_1 - x_2)||}{||y_1 - y_2||} &= \frac{||o(x_1 - x_2)||}{||x_1 - x_2||} \cdot \frac{||x_1 - x_2||}{||y_1 - y_2||} \\ &\leq 2\frac{||o(x_1 - x_2)||}{||x_1 - x_2||}\end{aligned}$$

and the difference is little-o of $(y_1 - y_2)$ as well.

Smoothness

Let $\phi : \overline{B(0, r/2)} \to \overline{B(0, r)}$ be the inverse of f, then ϕ is of class C^p.

Proof. Let $\psi : \text{Laut}(\mathcal{X}) \to \text{Laut}(\mathcal{X})$ with $x \mapsto x^{-1}$ be the inversion map, then since $D\phi = \psi \circ Df$ and $\psi \in C^\infty$ (see Smoothness of the Inverse), $D\phi \in C^p$. □

Theorem 3.33 (Inverse Function Theorem). *Let $\mathcal{X}$ be a Banach space, $U \subset \mathcal{X}$ be an open set and $f : U \to \mathcal{X}$ be of class C^p with $p \geq 1$. If there exists $x_0 \in U$ such that $Df(x_0) \in Laut(\mathcal{X})$ is invertible, then f is a local C^p-isomorphism.*

Proof. Let $g(x) = Df(x_0)^{-1} \circ f(x + x_0)$, then $g : (U - x_0) \to \mathcal{X}$ is of class C^p with $Dg(0) = I$. By the simplified version, there exists $V, W \in \mathcal{N}(x_0)$ such that $g : V \to W$ is a C^p-isomorphism.

Since $f(x) = Df(x_0) \circ g(x - x_0)$, $f : (V + x_0) \to Df(x_0)(W)$ is a C^p-isomorphism. □

Chapter 4

Manifold Theory

4.1 Vector Bundles

Definition 4.1 (Total Space and Projection)**.** *Let X be an E-manifold of class C^p. If at each point $p \in X$ the set W_x is toplinearly isomorphic to a Banach space F, define*

$$W = \bigsqcup_{x \in X} W_x$$

*as their set-theoretic coproduct, and let $\pi : W \to X$ such that the **fibre** $\pi^{-1}(x) = W_x$. Then W forms the **total space** on the **canonical projection** π, and we place a manifold structure on it through the construction of a vector bundle.*

Alternatively, if W starts as a manifold, then $\pi : W \to X$ can be taken as a morphism instead, and each $\pi^{-1}(p)$ can be given the same Banach space structure.

Definition 4.2 (Trivialising Maps)**.** *Let U be an open set in X, and let*

$$\tau : \pi^{-1}(U) \to U \times F$$

If τ is a C^p-isomorphism commuting with the projection on U,

$$\pi^{-1}(U) \xrightarrow{\ \tau\ } U \times F \xrightarrow{\ \pi_1\ } U$$

*where $\pi = \tau \circ \pi_1$ on $\pi^{-1}(U)$, then τ is a **trivialising map** for π.*

In particular, by fixing $p \in X$, the trivialising map transports the Banach structure of F onto each fibre

$$\pi^{-1}(p) \xrightarrow{\ \tau\ } p \times F \xrightarrow{\ \pi_2\ } F$$

through the map $\tau_p = \pi_1 \circ \tau$.

Definition 4.3 (VB-Equivalence)**.** *Let (U, σ) and (V, τ) be two trivialising maps with intersecting domains. The two maps are **compatible (VB-equivalent)** if for any $p \in U \cap V$,*

$$\sigma_p \circ \tau_p^{-1} : F \to F$$

*the **transition map** is a toplinear isomorphism, and the map*

$$U \cap V \to Laut(F) \quad p \to (\sigma \circ \tau^{-1})_p$$

is a C^p morphism.

If W starts as a manifold, then each trivialising map can transport the Banach structure on F to $\pi^{-1}(p)$. The condition on the transition map ensures that the same Banach structure is given by different trivialising maps, which allows $\pi^{-1}(p)$ to be identified as $F_p \cong F$. The transition map condition can be reformulated as

$$\sigma_p : F_p \to F \quad \tau_p : F_p \to F$$

both being toplinear isomorphisms.

If F is finite-dimensional, then the morphism condition is implied by the transition condition.

Definition 4.4 (Trivialising Covering)**.** *Let $\{U_i\}_{i \in I}$ be an open cover of X. A family $\{(U_i, \tau_i)\}_{i \in I}$ of mutually compatible trivialising maps forms a **trivialising covering**.*

Definition 4.5 (Vector Bundle)**.** *Let $\{U_i\}_{i \in I}$ be an open cover of X and $\{(U_i, \tau_i)\}_{i \in I}$ be a trivialising covering. The maximal trivialising covering that is VB-equivalent to it forms the **vector bundle** structure on π.*

Definition 4.6 (Manifold Structure on a Bundle)**.** *Let X be an E-manifold of class C^p, $\pi : W \to X$ a mapping, and F be a Banach space. Let $\{U_i\}_{i \in I}$ be an open cover of X and for each $i \in I$, let*

$$\tau_i : \pi^{-1}(U_i) \to U_i \times F$$

such that for each pair i, j, τ_i and τ_j are VB-equivalent. Then there exists a unique $E \times F$-manifold structure on W such that:

1. *π is a morphism and a submersion.*
2. *Each τ_i is an isomorphism, making π into a vector bundle.*
3. *$\{(U_i, \tau_i)\}_{i \in I}$ is a trivialising covering.*

Proof. Let $\{(V_j, \psi_j)\}$ be an atlas on X. Let $i \in I$ and $j \in J$. Let

$$\widetilde{U}_i = \pi^{-1}(U_i) \quad \widetilde{U}_i^j = \pi^{-1}(U_i \cap V_j)$$

Define the **bundle chart** $(\widetilde{U}_i^j, \varphi_i^j)$ as the composition

$$\widetilde{U}_i^j \xrightarrow{\tau_i} U_i \times F \xrightarrow{\psi_j \times \mathrm{Id}} E \times F$$

where $\varphi_i^j = (\psi_j \times \mathrm{Id}) \circ \tau_i$.

By the VB-equivalent condition, the transition map between the trivialisations

$$\tau_j \circ \tau_i^{-1} : (U_i \cap U_j) \times F \to (U_i \cap U_j) \times F$$

is a C^p-isomorphism.

So the chart transition map is also a C^p-isomorphism, each τ_i is an isomorphism, and the family is a trivialising covering. As π at each point is the composition of two morphisms, π itself is also a morphism, which splits $E \times F$ as $E \times 0$ and $0 \times F$, making it a submersion. □

Definition 4.7. *Let X, Y be a manifold, $\pi : W \to X$ be a vector bundle with fibre F, and $f : Y \to W$ be a morphism. Let $(U, \psi) \in X$ be a chart, (U, τ) be a local trivialisation, and $(V, \phi) \in Y$ be a chart such that $f(V) \subset \pi^{-1}(U) = \widetilde{U}$, then*

$$\begin{array}{ccccc} V & \xrightarrow{f} & \widetilde{U} & \xrightarrow{\tau} & U \times F \\ & & {\scriptstyle\pi}\downarrow & & \downarrow \\ & & U & \xrightarrow{\psi} & \widehat{U} \end{array}$$

where the composite $\tau \circ f : V \to U \times F$ has two components,

$$f_U : V \to U \quad f_F : V \to F$$

*where f_F is the **vector component** of f in the bundle chart $U \times F$ over U.*

Chapter 5

Tricks

Special thanks to Anson Li for providing a summary of the following recurrent techniques.

5.1 Upper and Lower Approximations

Definition 5.1 (Upper and Lower Approximations)**.** *Let Ω be a set, $\Sigma \subseteq \Omega$ be a subset where a notion of measure $I : \Sigma \to [0, \infty]$ is defined.*

Let $x \in \Omega$ and suppose that there are measurable collections $\overline{\Sigma}_x, \underline{\Sigma}_x \subseteq \Sigma$ that approximate x from below and above, respectively. Then the supremum and infimum of the lower and upper collections

$$\overline{I}(x) = \sup_{\phi \in \overline{\Sigma}_x} I(\phi) \leq \inf_{\phi \in \underline{\Sigma}_x} I(\phi) = \underline{I}(x)$$

*are the **lower** and **upper** approximations of $I(x)$. Generally, If*

$$\overline{I}(x) = I(x) \text{ or } \underline{I}(x) = I(x) \quad \forall x \in \Sigma$$

*the approximations agree with I, then $\overline{I}$ and/or $\underline{I}$ **extend** I.*

The idea of lower and upper approximations is commonly used in analysis and measure theory. In particular, the Riemann Integral (Darboux's characterisation), the outer measure, Carathéodory's Theorem, Carathéodory's Extension Theorem, Integral for non-simple functions, the monotone convergence theorem and Fatou's Lemma all make use of this concept.

Theorem 5.1. *Let $x, y \in \Omega$, then $\overline{\Sigma}_x \subseteq \overline{\Sigma}_y \Rightarrow \overline{I}(x) \leq \overline{I}(y)$, and $\underline{\Sigma}_x \subseteq \underline{\Sigma}_y \Rightarrow \underline{I}(x) \geq \underline{I}(y)$.*

Theorem 5.2. *Let $x \in \Omega$ and $c \in [0, \infty]$. If $I(\phi) \leq c \forall \phi \in \overline{\Sigma}_x$, then $\overline{I}(x) \leq c$. If $I(\phi) \geq c \forall \phi \in \underline{\Sigma}_x$, then $\underline{I} \geq c$.*

Theorem 5.3 (Controlling the Lower Limit)**.** *Let $x \in \Omega$, $\{x_n\}_1^\infty$, and $\{\phi_n\}_1^\infty \subseteq \overline{\Sigma}_x$ be a supporting sequence such that $I(\phi_n) \nearrow \overline{I}(x)$. If for any $n \in \mathbb{N}$, $\exists \psi_n \in \Sigma_{x_n}$ such that $I(\phi_n) \leq I(\psi_n)$ eventually, then*

$$\overline{I}(x) = \lim_{n\to\infty} I(\phi_n) \leq \liminf_{n\to\infty} I(\psi_n) \leq \liminf \overline{I}(x_n)$$

Proof. Since $I(\phi_n) \leq I(\psi_n) \leq I(x_n)$ eventually, we have

$$\liminf_{n\to\infty} I(\phi_n) \leq \liminf_{n\to\infty} I(\psi_n) \leq \liminf_{n\to\infty} \overline{I}(x_n)$$

by the order limit theorem. Since $I(\psi_n) \nearrow \overline{I}(x)$,

$$\overline{I}(x) = \lim_{n\to\infty} I(\phi_n) \leq \liminf_{n\to\infty} I(\psi_n) \leq \liminf \overline{I}(x_n)$$

□

5.2 Method of Successive Approximations

Definition 5.2 (Reverse Lipschitz Criterion)**.** *Let X be a Banach space. For any $\rho \geq 0$ denote*

$$M_\rho = \{x \in X : ||x|| \leq \rho\}$$

A collection $\Sigma \subset X$ satisfies the reverse Lipschitz criterion if there exists $\gamma \in (0,1)$ and $C > 0$ such that for any $x \in M_\varepsilon$, there exists an approximating element $\phi \in \Sigma$ with $||x - \phi|| \leq \gamma\varepsilon$.

The precision of the approximation scales with the norm of the target, and the error must be in a way, consistently less than the norm of the target.

Theorem 5.4 (Method of Successive Approximations (Same Space))**.** *Let X be a Banach space, and $\Sigma \subset X$ be a set satisfying the reverse Lipschitz criterion, then for all $x \neq 0$, there exists a sequence $\{\phi_n\}_1^\infty \subset \Sigma$ such that $\sum_{n=1}^\infty \phi_n = x$.*

Proof. Let $x_0 = x$. For $n \in \mathbb{N}_0$, choose $\phi_{n+1} \in \Sigma$ such that $||x_n - \phi_{n+1}|| \leq \gamma\, ||x_n||$, and take the remainder $x_{n+1} = x_n - \phi_{n+1}$.

Leveraging the completeness of X, if we can control the norms of each ϕ_n, we can force the sum to converge. Since

$$||\phi_{n+1}|| \leq ||\phi_{n+1} - x_n|| + ||x_n|| \leq (1+\gamma)\, ||x_n|| < 2\, ||x_n||$$

and

$$||\phi_{n+1}|| \leq 2\, ||x_n|| \leq 2\gamma\, ||x_{n-1}|| \leq \cdots \leq 2\gamma^n\, ||x_0|| \quad \forall n \in \mathbb{N}_0$$

from induction, with $\gamma \in (0,1)$, $\sum_{n\in\mathbb{N}} ||\phi_n|| < \infty$ and $\sum_{n=1}^\infty \phi_n$ converges.

Now to bound the remainder, since $||x_n|| \leq \gamma^n\, ||x_0||$,

$$\left|\left| x_0 - \sum_{n=1}^\infty \phi_n \right|\right| = \lim_{n\to\infty} \left|\left| x - \sum_{n=1}^n \phi_n \right|\right| \leq \lim_{n\to\infty} \gamma^n\, ||x_0|| = 0$$

yielding $x = \sum_{n=1}^\infty \phi_n$. □

Theorem 5.5 (Method of Successive Approximations (Two Spaces))**.** *Let X and Y be Banach spaces and $T : Y \to X$ be a continuous linear map.*

Let $\Sigma \subseteq Y$. If there exists $C > 0$, $\gamma \in (0,1)$ such that for any $x \in X$, $||x|| \leq \varepsilon$, there exists $\phi \in \Sigma$ such that

$$||x - T(\phi)||_X \leq \gamma\varepsilon \quad ||\phi||_Y \leq C\varepsilon$$

Then for any $x \in X, x \neq 0$, there exists $\{\phi_n\}_1^\infty$ such that $x = T\left(\sum_{n=1}^\infty \phi_n\right)$.

Proof. Using a similar inductive construction: Let $x_0 = x$. For any $n \in \mathbb{N}_0$, choose $\phi_{n+1} \in \Sigma$ such that $||\phi_{n+1}|| \leq C\,||x_n||$ and $||x_n - T(\phi_{n+1})|| \leq \gamma\,||x_n||$, and take $x_{n+1} = x_n - \phi_n$.

Since
$$||\phi_{n+1}|| \leq C\,||x_n|| \leq \gamma\,||x_{n-1}|| \leq \cdots \leq \gamma^n\,||x_0|| \quad \forall n \in \mathbb{N}_0$$
with $\gamma \in (0,1)$, $\sum_{n=1}^\infty ||\phi_n|| < \infty$ and $\sum_{n=1}^\infty \phi_n$ converges.

As T is linear and continuous, with $||x_n|| \leq \gamma^n\,||x_0||$ we obtain

$$\begin{aligned}
\left|\left|x_0 - T\left(\sum_{n=1}^\infty \phi_n\right)\right|\right| &= \lim_{n\to\infty}\left|\left|x_0 - T\left(\sum_{k=1}^n \phi_k\right)\right|\right| \\
&= \lim_{n\to\infty}\left|\left|x_0 - \sum_{k=1}^n \phi_k\right|\right| \\
&\leq \lim_{n\to\infty}||x_n|| = \lim_{n\to\infty}\gamma^n\,||x_0|| \\
&= 0
\end{aligned}$$

□

5.3 Fixed Point Problem

Theorem 5.6. *Let $\mathcal{X}$ be a Banach space and $F : \mathcal{X} \to \mathcal{X}$ be any mapping. Suppose that $Id + F$ is a contraction, then F is invertible.*

Proof. Let $y \in F$, and take $g_y(x) = (\mathrm{Id} + F)(x) - y$, then $F(x) = y$ if and only if $g_y(x) = x$, and g_y is also a contraction. By the Banach fixed-point theorem, g_y has a unique fixed point. Therefore there exists $x \in E$ such that $y = F(x)$. □

Theorem 5.7. *Let $\mathcal{X}$ be a Banach space, $F : \mathcal{X} \to \mathcal{X}$ be any mapping. Suppose that there exists a homeomorphism $H : \mathcal{X} \to \mathcal{X}$ such that $Id + HF$ is a contraction, then F is invertible.*

Proof. Since HF is invertible and H is a homeomorphism, F is invertible as well. □