Characterizing Amenability in Discrete Groups

A Tour of Algebra and Analysis

Avinash Iyer

March 2025

Contents

Chapter 0

Introduction

In the beginning, God created the heavens and the Earth,[1] and a lot of other things that are detailed in the book of Genesis. Unfortunately, those that wrote down and translated the book of Genesis failed to mention the most important feature of the universe that He (may or may not have) created — the axiom of choice. It may be remarked that the axiom of choice is not, strictly speaking, a God-given creation, but accepting it certainly requires a leap of faith — after all, Paul Cohen and Kurt Gödel showed that it is independent of the rest of the axioms of set theory — but since we are going to be working in the realm of analysis throughout this thesis, we will be accepting it as such.

Unfortunately, despite our best efforts, and the convenient results that the axiom of choice provides for (see Example B.1.3), the axiom of choice provides some counterintuitive and downright paradoxical results — not that it's worth throwing out, but it is certainly worth investigating and understanding. One of these counterintuitive results is detailed in Chapter 2, where we show implicitly that there does not exist a finitely additive measure on the three-dimensional real numbers that is also invariant under Euclidean isometry by proving the Banach–Tarski paradox in its most general form.

Using the Banach–Tarski paradox as motivation, we then go to proving various characterizations, definitions, and proofs of amenability in groups — i.e., we now want to understand when a group is well-behaved, rather than ill-behaved as in the case of the isometry group of $\mathbb{R}^3$. We first use some primarily group-theoretic techniques surrounding amenability, such as in the proof of Tarski's theorem in Chapter 3 and the establishment of amenability in subgroups and quotient groups in the first section of Chapter 4. Then, we use techniques from functional analysis in later chapters to prove amenability, first by establishing the equivalence between amenability and the existence of an invariant state on $\ell_\infty(G)$, in sections 2–4 of Chapter 4; we then expand on these techniques in Chapter 5 to understand Følner's condition and approximate means.

After a quick discussion of the application of Følner's condition to geometric group theory, we discuss representations of groups as bounded operators on Hilbert spaces, using yet more techniques from functional analysis and the theory of operator algebras to show, in Chapter 6, that a group is amenable if it satisfies certain criteria related to the left-regular representation of Γ on the space $\ell_2(\Gamma)$. Finally, in Chapter 7, we move from the representation of a group to representations of its group $*$-algebra, and show how properties of the group C^*-algebra inform us about properties of the group, and vice-versa.

This is a very, very ambitious project that I am not sure I would have been able to complete without the effort that professor Rainone helped with understanding some of the more heavy theory surrounding amenability, the Følner condition, and especially C^*-algebras. Additionally, his draft Functional Analysis textbook was an incredibly helpful resource in learning everything necessary to understand the more heavy analysis-based formulations of amenability. Of course, in addition to his help, I'm grateful for the opportunities, classes, and memorable times I have spent with the mathematics department at Occidental,

[1]Well, maybe not God specifically.

without whom I would never have believed myself capable of learning such intricate and complex topics.

I have always believed that a proper thesis, even an expository one, need not be a bland affair. When it is a topic the author is especially interested in, such as myself with the concepts, theories, and ideas surrounding amenability, I believe it is of paramount importance that the author make the subject as enjoyable for the reader as it is for them. As a result, I have included light humor throughout this thesis, with the purpose of bringing a smile to the readers' face just as this topic has brought many a smile to my own.

Most of the results and proofs in this thesis are not my own, but a compilation of various sources that I believe provide a broad and deep coverage of the subject; I have collected, simplified, and reordered them in order to understand not only for myself, but to potentially help others with the process of understanding amenability. Just because I may have forgotten to attribute them in the paper does not mean they are of my creation. All sources used to establish proofs used in this book are attributed in the bibliography.

Note that all groups in this thesis are endowed with the discrete topology. This is because I did not have enough time to learn about the more general case of topological groups, which are endowed with a locally compact topology that gives them interesting properties, and the topic of discrete groups is, evident in the length of this document, enough to discuss interesting properties.

Chapter 1

Categorical Constructions for the Unemployed Mathematician

In various sections of this book, we allude to various properties and ideas surrounding "free" categorical constructions[1] that admit particular universal properties. We will detail some of these constructions here.

1.1 Free Groups

Given a set A, we want to know how exactly we can create a group structure from the elements in A such that they extend from A to a group generated by A in a particularly "natural" way. This will be the free group, whose properties we will discuss in Chapter 2.

Definition 1.1.1. Let G be a group, and $S \subseteq G$ be a subset. We define the subgroup generated by S to be

$$\langle S \rangle_G = \bigcap \left\{ H \,\middle|\, S \subseteq H,\ H \text{ a subgroup} \right\}.$$

We say S generates G if $\langle S \rangle_G = G$.

Generated subgroups can be broadly characterized as follows:

$$\langle S \rangle_G = \left\{ s_1^{a_1} s_2^{a_2} \cdots s_n^{a_n} \,\middle|\, n \in \mathbb{N},\ s_1, \ldots, s_n \in S,\ a_1, \ldots, a_n \in \{-1, 1\} \right\}.$$

We say $\langle S \rangle_G$ is finitely generated if $\operatorname{card}(S) < \infty$.

If S is such that, for any $x \in S$, we have $x^{-1} \in S$, then we say S is symmetric.

To construct a free group, we begin by stating its universal property — that is, its innate nature as an "extension" of a set-function into a group structure. Then, we will show that a more constructive definition of the free group satisfies this universal property. The following section draws heavily from [Löh17], but we will mostly focus on the construction of the free group rather than the proof of uniqueness.

Definition 1.1.2. Let S be a set. A group F containing S is said to be freely generated if, for every group G, and every map $\phi\colon S \to G$, there is a unique group homomorphism $\varphi\colon F \to G$ that extends φ. The following diagram, where ι denotes the inclusion of S into F, commutes:

$$\begin{tikzcd} S \arrow[r, "\phi"] \arrow[d, hook, "\iota"'] & G \\ F \arrow[ur, "\varphi"'] & \end{tikzcd}$$

[1] Hence the name of this chapter.

We say F is the free group generated by S.

Intuitively, to construct the free group, if we have $a \mapsto \phi(a)$ between S and G, then we will define $\varphi(a^n) = \phi(a)^n$ inside $F(S)$. Uniqueness will follow from the fact that we can take two groups that satisfy the universal property, F and F′, and apply the universal property on set-valued functions between S and F and S and F′ respectively. The formal proof of the construction of the free group is as follows.

Proof. We will construct a group consisting of "words" made up of elements of S and their inverses. This starts by considering the alphabet $A = S \cup \widehat{S}$, where $\widehat{S}$ is a disjoint copy of S — every $\hat{s} \in \widehat{S}$ will play the role of an inverse to s in our group.

- Define A^* to be the set of all words over the alphabet A, including the empty word, ϵ. We define the operation $A^* \times A^* \to A^*$ by concatenating words, which is an associative operation with neutral element ϵ.
- Define the equivalence relation $\sim$ generated by the following two relations, where for all $x, y \in A^*$ and $s \in S$, we have

$$\begin{aligned} xs\hat{s}y &\sim xy \\ x\hat{s}sy &\sim xy. \end{aligned}$$

  The equivalence classes with respect to $\sim$ will be denoted $[\cdot]$.

  We have a well-defined composition $[x][y] = [xy]$ mapping $F(S) \times F(S) \to F(S)$ for all $x, y \in A^*$.

We show that $F(S)$ with the concatenation operation is a group. Here, we see that $[\epsilon]$ is the neutral element for the composition, and associativity is inherited from associativity of concatenation in A^*. To show the existence of inverses, we define the inverse map inductively by taking $I(\epsilon) = \epsilon$, and

$$\begin{aligned} I(sx) &= I(x)\hat{s} \\ I(\hat{s}x) &= I(x)s \end{aligned}$$

for all $x \in A^*$ and $s \in S$. Inductively, we can see that $I(I(x)) = x$ and

$$\begin{aligned} [I(x)][x] &= [I(x)x] \\ &= [\epsilon] \\ [x][I(x)] &= [xI(x)] \\ &= [\epsilon]. \end{aligned}$$

Thus, $F(S)$ is a group.

Now, we show $F(S)$ is freely generated. Let $i\colon S \to F(S)$ be the map that sends $s \mapsto [s]$. By our construction, we know that $i(S) \subseteq F(S)$ is a generating set for $F(S)$. We will show the universal property holds for $F(S)$.

To start, let $\phi\colon S \to G$ be a set-valued map between S and an arbitrary group G. We construct $\phi^*\colon A^* \to G$ by taking

$$\begin{aligned} \epsilon &\mapsto e \\ sx &\mapsto \phi(s)\phi^*(x) \\ \hat{s}x &\mapsto (\phi(s))^{-1}\phi^*(x) \end{aligned}$$

for all $x \in A^*$ and $s \in S$. This definition of ϕ^* is compatible with the equivalence relation on A^*, and we see that $\phi^*(xy) = \phi^*(x)\phi^*(y)$. Thus, we get a well-defined map $\varphi\colon F(S) \to G$, taking $[x] \mapsto [\phi^*(x)]$.

It remains to be shown that the map $i\colon S \to F(S)$ is injective, which will show that $F(S)$ is freely generated by S. Let $s_1, s_2 \in S$, and consider the set-function $\phi\colon S \to \mathbb{Z}$ given by $\phi(s_1) = 1$ and $\phi(s_2) = -1$. Then, we must have

$$\begin{aligned}\varphi(i(s_1)) &= \phi(s_1)\\ &= 1\\ &\neq -1\\ &= \phi(s_2)\\ &= \varphi(i(s_2)).\end{aligned}$$

Thus, we have $i(s_1) \neq i(s_2)$, so i is injective. □

Most of the definitions of the free group automatically default to the characterization of $F(S)$ as the set of reduced words in $S \cup S^{-1}$. This is the characterization we will be using in the future, but it is still important to understand where exactly the "free" in free group comes from, and how it relates to the particular universal property that actually characterizes $F(S)$ uniquely up to isomorphism.

1.2 Free Vector Spaces

Given a set A, just as we are able to construct a free group, we can take any set A and construct a "universal" vector space out of the set. The free vector space (as it is known) is the universal object that extends any set-valued function into a linear map, treating elements of the set as its basis (see Definition **??**). We are interested in the case of the free vector space over the complex numbers, but note that the following definition of the free vector space applies over any field.

Theorem 1.2.1. Let Γ be a nonempty set. There is a vector space, $\mathbb{C}[\Gamma]$, with $\dim(\mathbb{C}[\Gamma]) = \operatorname{card}(\Gamma)$, and an injective map $\delta\colon \Gamma \to \mathbb{C}[\Gamma]$ such that the following universal property holds: if V is a $\mathbb{C}$-vector space, and $\phi\colon \Gamma \to V$ is a set-valued function, then there is a unique linear map $T_\phi\colon \mathbb{C}[\Gamma] \to V$ such that $T_\phi \circ \delta = \phi$.

$$\begin{tikzcd} \Gamma \arrow[r, "\delta"] \arrow[dr, "\phi"'] & \mathbb{C}[\Gamma] \arrow[d, "T_\phi"] \\ & V \end{tikzcd}$$

Proof. Consider the linear subspace of finitely supported functions, $\mathbb{C}[\Gamma] \subseteq \mathcal{F}(\Gamma, \mathbb{C})$. For each $t \in \Gamma$, we define

$$\delta_t(s) = \begin{cases} 1 & s = t \\ 0 & \text{else} \end{cases}.$$

We see that $\delta_t \neq \delta_s$ whenever $s \neq t$, meaning that the map $\delta\colon \Gamma \to \mathbb{C}[\Gamma]$, defined by $s \mapsto \delta_s$, is injective.

We will show that $\{\delta_s\}_{s\in\Gamma}$ is a linear basis for $\mathbb{C}[\Gamma]$. If $f \in \mathbb{C}[\Gamma]$, with $\operatorname{supp}(f) = \{s_1, \dots, s_n\} \subseteq \Gamma$, we set $\alpha_j = f(t_j)$, and see that

$$f = \sum_{j=1}^{n} \alpha_j \delta_{s_j},$$

which shows that $\{\delta_s\}_{s\in\Gamma}$ is a spanning set.

To show that $\{\delta_s\}_{s\in\Gamma}$ is linearly independent, consider $g = \sum_{j=1}^{n} \alpha_j \delta_{s_j} \in \mathbb{C}[\Gamma]$ such that $g = 0$. Then, $g(t) = 0$ for all $t \in \Gamma$, and in particular, $g(s_i) = 0$ for every $1 \leqslant i \leqslant n$. Thus, we have

$$0 = g(s_j)$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} \alpha_j \delta_{s_j}(s_i) \\
&= \alpha_i,
\end{aligned}
$$

so $\alpha_j = 0$ for each j. Thus, $\{\delta_s\}_{s\in\Gamma}$ is linearly independent.

Turning to the universal property, we define $T_\phi : \mathbb{C}[\Gamma] \to V$ in terms of ϕ as follows:

$$T_\phi\left(\sum_{j=1}^{n} \alpha_j \delta_{s_j}\right) = \sum_{j=1}^{n} \alpha_j \phi(s_j).$$

This yields an expression of T_ϕ uniquely in terms of ϕ and δ, thereby satisfying the universal property. □

Example 1.2.1. Let z be an abstract variable, and consider the set of "formal powers" of z, $\left\{z^k\right\}_{k\in\mathbb{N}}$. Then, the free vector space generated by this set, $\mathbb{C}[z]$, is the set of all polynomials with coefficients in $\mathbb{C}$. By the universal property, we know that every polynomial $p \in \mathbb{C}[z]$ has a unique expression $p = \sum_{j=0}^{n} a_j z^j$.

One of the primary uses of the free vector space is that, via this construction, we can show that vector spaces are particularly nice algebraic objects. We often use these properties implicitly in linear algebra, but they are only provable by using the free vector space.

Theorem 1.2.2. Let X, Y, and Z be vector spaces.

(a) If $\iota \colon Y \hookrightarrow X$ is an injective linear map, and $\varphi \colon Y \to Z$ is a linear map, then there is a (not necessarily unique) map $T \colon X \to Y$ such that $T \circ \iota = \varphi$.

$$
\begin{tikzcd}
0 \arrow[r] & Y \arrow[r, hook, "\iota"] \arrow[d, "\varphi"'] & X \arrow[dl, "T"] \\
 & Z &
\end{tikzcd}
$$

This shows that vector spaces are injective objects — any linear map factors through an injective map.

(b) If $\pi \colon X \to Z$ is a surjective linear map, and $\varphi \colon Y \to Z$ is a linear map, then there is a (not necessarily unique) map $\delta \colon Y \to X$ such that $\pi \circ \delta = \varphi$.

$$
\begin{tikzcd}
 & Y \arrow[dl, "\delta"'] \arrow[d, "\varphi"] & \\
X \arrow[r, "\pi"'] & Z \arrow[r] & 0
\end{tikzcd}
$$

This shows that vector spaces are projective objects — any linear map factors through a surjective map.

Proof.

(a) Let $\mathcal{A}$ be a basis for Y. Then, since ι is an injective linear map, the set $\mathcal{B}_0 = \{\iota(y) \mid y \in \mathcal{A}\}$ can be extended to a basis $\mathcal{B}$ for X.

We set $t\colon \mathcal{B} \to Z$ to be

$$t(x) = \begin{cases} \varphi(x) & x \in \mathcal{B}_0 \\ 0 & x \in \mathcal{B} \setminus \mathcal{B}_0 \end{cases}.$$

By the universal property of the free vector space, this extends to a linear map $T\colon X \to Y$. Since $T \circ \iota$ agrees with φ on $\mathcal{A}$, the universal property of the free vector space states that $T \circ \iota$ agrees with φ on all of Y.

(b)

□

1.3 Free Algebras

Chapter 8 of this thesis will be focused on understanding the properties of the (reduced) group C^*-algebra. This will require some background in the theory of algebras, so we will understand the purely algebraic properties here before diving into the analytic properties in Chapter 5. Just as there are free groups and free vector spaces, we can also talk about free algebras. In Chapter 8, we will construct special norms on free algebras to elucidate properties of the underlying group.

Similar to a free group, the free algebra (or free $*$-algebra) is constructed by taking a certain collection of "words" over a set of symbols, and then, if desired, "modding out" by the ideal generated by a set of relations. We formalize this in steps.

Definition 1.3.1. Let $E = \{x_i\}_{i\in I}$ be a collection of symbols that may not commute. The space of all polynomials over E is the free vector space over the set of words formed by symbols in E,

$$\Gamma_E = \{x_{i_1}x_{i_2}\cdots x_{i_n} \mid n \in \mathbb{N}, i_1, \dots, i_n \in I\}.$$

We denote this space $\mathbb{C}\langle E\rangle$.

In the free vector space $\mathbb{C}\langle E\rangle$, we may define multiplication by concatenation:

$$(x_{i_1}x_{i_2}\cdots x_{i_n})(x_{j_1}x_{j_2}\cdots x_{j_m}) = x_{i_1}x_{i_2}\cdots x_{i_n}x_{j_1}x_{j_2}\cdots x_{j_m},$$

where $i_1, \dots, i_n, j_1, \dots, j_m \in I$. The space $\mathbb{C}\langle E\rangle$, equipped with multiplication by concatenation, is known as the free algebra on E.

To turn $\mathbb{C}\langle E\rangle$ into a $*$-algebra, we define the formal set $E^* = \left\{x_i^*\right\}_{i\in I}$, and define the involution on $\mathbb{C}\langle E \cup E^*\rangle$ by taking

$$\left(\alpha x_{i_1}^{\varepsilon_1}x_{i_2}^{\varepsilon_2}\cdots x_{i_n}^{\varepsilon_n}\right)^* = \overline{\alpha}x_{i_n}^{\delta_n}x_{i_{n-1}}^{\delta_{n-1}}\cdots x_{i_2}^{\delta_2}x_{i_1}^{\delta_1},$$

where

$$\delta_j = \begin{cases} * & \varepsilon_j = 1 \\ 1 & \varepsilon_j = * \end{cases}.$$

The set $\mathbb{C}\langle E \cup E^*\rangle$ with the involution defined above is known as the free $*$-algebra on E, and is usually denoted $\mathbb{A}^*(E)$.

If $R \subseteq \mathbb{A}^*(E)$ is a collection of relations, we let $I(R) = \text{ideal}(R)$. Then, the quotient algebra

$$\mathbb{A}^*(E|R) = \mathbb{A}^*(E)/I(R)$$

is known as the universal $*$-algebra on E with relations R.

Theorem 1.3.1 (Universal Properties)**.** Let $E = \{x_i\}_{i\in I}$ be a set of abstract symbols, and let B be a $*$-algebra. Let $\phi\colon E \to B$ be an injective map, and define $b_i = \phi(x_i)$.

- There is a unique $*$-homomorphism $\varphi\colon \mathbb{A}^*(E) \to B$ such that $x_i \mapsto b_i$. The following diagram commutes.

$$\begin{tikzcd} E \arrow[r, "\phi"] \arrow[d, "\iota"'] & B \\ \mathbb{A}^*(E) \arrow[ur, "\varphi"'] & \end{tikzcd}$$

- If $R \subseteq \mathbb{A}^*(E)$ is a set of relations, and $\{b_i\}_{i\in I}$ satisfies the relations R, then there is a unique $*$-homomorphism $\mathbb{A}^*(E|R) \to B$ such that $x_i + I(R) \mapsto b_i$. The following diagram commutes.

$$\begin{tikzcd} E \arrow[r, "\phi"] \arrow[d, "\iota"'] & B \\ \mathbb{A}^*(E|R) \arrow[ur, "\varphi"'] & \end{tikzcd}$$

One of the most important $*$-algebras we will study is generated from a group by taking the free vector space over the group.

Definition 1.3.2. Let Γ be a group with identity element e, and let $\mathbb{C}[\Gamma]$ be the free vector space generated by Γ. We define a multiplication $f * g$, where $f, g \in \mathbb{C}[\Gamma]$ are finitely supported functions, by convolution:

$$\begin{aligned} f * g(s) &= \sum_{t\in\Gamma} f(t)g\left(t^{-1}s\right) \\ &= \sum_{r\in\Gamma} f\left(sr^{-1}\right)g(r). \end{aligned}$$

The involution on $\mathbb{C}[\Gamma]$ is defined by $f^*(t) = \overline{f\left(t^{-1}\right)}$. The multiplicative identity is δ_e, and multiplication satisfies $\delta_s * \delta_t = \delta_{st}$.

Chapter 2

How to Feed 5,000 Hungry Mathematicians: Paradoxical Decompositions

In the Bible, one of the miracles of Jesus is the feeding of the five thousand,[I] where, despite only having five loaves of bread and two fishes, a large crowd splits these morsels among themselves and eats to satisfaction after Jesus calls upon the power of God to enable them to do so. Of course, we may not be able to fully replicate this without some divine intervention — but, mathematically, thanks to the power of the axiom of choice, we can show that something like the feeding of the five thousand is not only possible, but a fundamental feature of the isometry group of $\mathbb{R}^3$. This is exemplified in the most general form of the Banach–Tarski paradox.

> **Proposition 2.0.1** (Strong Banach–Tarski Paradox). Let A and B be bounded subsets of $\mathbb{R}^3$ with nonempty interior. There is a partition of A into finitely many disjoint subsets such that a sequence of isometries applied to these subsets yields B.

The Banach–Tarski paradox throws a wrench into a common belief that we have about $\mathbb{R}^3$ — specifically, that every subset of $\mathbb{R}^3$ has a *finitely additive* "volume" that is invariant under isometry.[II] This property does exist for $\mathbb{R}$ and $\mathbb{R}^2$, as their isometry groups have a property known as amenability — in Section 4.4, we will provide an outline for why this is true.

To develop paradoxical decompositions, we will begin with the Ping Pong Lemma, which will allow us to find freely generated subgroups. We will apply this to the case of SO(3) to find a freely generated subgroup. Then, we will use the fact that free groups on more than one generator have a property known as paradoxicality — this property will provide the germ of the proof of the Banach–Tarski paradox.

2.1 The Ping Pong Lemma

To move towards paradoxical decompositions, we need to find a simple and easily applicable criterion for knowing when an arbitrary group contains a freely generated subgroup. This is the domain of the Ping Pong Lemma, which we will prove in this section to show the existence of a freely generated subgroup of SO(3). Later, this freely generated subgroup will be indispensable in proving the Banach–Tarski paradox.

[I] Fun fact: the feeding of the five thousand is the only other miracle of Jesus (aside from the resurrection) that is in all four gospels.

[II] Note that if we desire countable additivity, the axiom of choice shows that there does not exist a countably additive measure on $P(\mathbb{R})$ that is also translation-invariant (see [Fol84, Section 1.1]). Finite additivity is a weaker condition than countable additivity that allows for the existence of well-behaved measures on $P(\mathbb{R})$ and $P(\mathbb{R}^2)$, but even this fails in $\mathbb{R}^3$ and above.

We begin by defining a free product of a family of groups $\{\Gamma_i\}_{i\in I}$. This will allow us to state the Ping Pong Lemma in its maximal generality.

Definition 2.1.1 (Free Product)**.** Let A be a set, and set $W(A)$ to be the set of words in A equipped with the operation of concatenation. This turns $W(A)$ into a construction known as the free monoid.

If $\{\Gamma_i\}_{i\in I}$ is a family of groups, and $A = \coprod_{i\in I} \Gamma_i$ is the coproduct (or disjoint union) of the groups Γ_i, then we define the equivalence relation $\sim$ generated by

$$\begin{aligned} we_iw' &\sim ww' \text{ where } e_i \text{ is the neutral element of } \Gamma_i \text{ for some } i \in I \\ wabw' &\sim wcw' \text{ where } a, b, c \in \Gamma_i \text{ and } c = ab \text{ for some } i \in I. \end{aligned}$$

Then, the quotient $W(A)/\sim$ is known as the free product of the groups $\{\Gamma_i\}_{i\in I}$, and is denoted

$$\bigstar_{i\in I}\Gamma_i.$$

Remark 2.1.1. The free group $F(S)$ is an instance of the free product where the Γ_i are the formal cyclic groups generated by each $s \in S$.

From the way we have defined the free product, it can be shown, as in [Har00, II.A.], that every element of the free product admits a unique reduced word in $W(A)$, along with the following universal property: if $\{\Gamma_i\}_{i\in I}$ is a family of groups, and $h_i\colon \Gamma_i \to \Gamma$ for some fixed group Γ, then there is a unique homomorphism $h\colon \bigstar_{i\in I}\Gamma_i \to \Gamma$ such that the following diagram commutes for each Γ_{i_0}.

$$\begin{tikzcd} \Gamma_{i_0} \arrow[r, "h_{i_0}"] \arrow[d, hook, "\iota_{i_0}"'] & \Gamma_i \\ \bigstar_{i\in I}\Gamma_i \arrow[ur, "h"'] & \end{tikzcd}$$

Theorem 2.1.1 (Ping Pong Lemma)**.** Let G be a group that acts on a set X, and let Γ_1, Γ_2 be subgroups of G. Let $\Gamma = \langle \Gamma_1, \Gamma_2 \rangle$. Assume Γ_1 contains at least 3 elements, and Γ_2 contains at least 2 elements.

Suppose there exist nonempty subsets $X_1, X_2 \subseteq X$ with $X_1 \triangle X_2 \neq \emptyset$ such that for all $\gamma \in \Gamma_1$ with $\gamma \neq e_G$,

$$\gamma(X_2) \subseteq X_1,$$

and for all $\gamma \in \Gamma_2$ with $\gamma \neq e_G$,

$$\gamma(X_1) \subseteq X_2.$$

Then, Γ is isomorphic to the free product $\Gamma_1 \star \Gamma_2$.

Proof. Let w be a nonempty reduced word with letters in the disjoint union of $\Gamma_1 \setminus \{e_G\}$ and $\Gamma_2 \setminus \{e_G\}$. We must show that the element of Γ defined by w is not the identity.

If $w = a_1b_1a_2b_2\cdots a_k$ with $a_1, \ldots, a_k \in \Gamma_1 \setminus \{e_G\}$ and $b_1, \ldots, b_{k-1} \in \Gamma_2 \setminus \{e_G\}$, then,

$$\begin{aligned} w(X_2) &= a_1b_1\cdots a_{k-1}b_{k-1}a_k(X_2) \\ &\subseteq a_1b_1\cdots a_{k-1}b_{k-1}(X_1) \\ &\subseteq a_1b_1\cdots a_{k-1}(X_2) \\ &\vdots \\ &\subseteq a_1(X_2) \\ &\subseteq X_1. \end{aligned}$$

Seeing as $X_2 \not\subseteq X_1$ (by the definition of symmetric difference), it is the case that $w \neq e_G$.

If $w = b_1 a_2 b_2 a_2 \cdots b_k$, we select $a \in \Gamma_1 \setminus \{e_G\}$, and we find that $awa^{-1} \neq e_G$, meaning $w \neq e_G$. Similarly, if $w = a_1 b_1 \cdots a_k b_k$, we select $a \in \Gamma_1 \setminus \left\{e_G, a_1^{-1}\right\}$, similarly finding that $awa^{-1} \neq e_G$. If $w = b_1 a_2 b_2 \cdots a_k$, then we select $a \in \Gamma_1 \setminus \{1, a_k\}$, and find $awa^{-1} \neq e_G$. □

We can refine Theorem 2.1.1 to the case of "doubles" wherein we find a different (yet more readily applicable) sufficient condition for a group that contains a copy of the free group on two generators.

Corollary 2.1.1 (Ping Pong Lemma for "Doubles")**.** Let G act on X, and let A_+, A_-, B_+, B_- be disjoint subsets of X whose union is not equal to X. Then, if

$$\begin{aligned} a \cdot (X \setminus A_-) &\subseteq A_+ \\ a^{-1} \cdot (X \setminus A_+) &\subseteq A_- \\ b \cdot (X \setminus B_-) &\subseteq B_+ \\ b^{-1} \cdot (X \setminus B_+) &\subseteq B_-, \end{aligned}$$

then it is the case that $\langle a, b\rangle$ is isomorphic to the free group on two generators.

Proof. We let $A = A_+ \sqcup A_-$, $B = B_+ \sqcup B_-$, $\Gamma_1 = \langle a\rangle$, and $\Gamma_2 = \langle b\rangle$. Then, A, B, Γ_1, Γ_2 satisfy the conditions for Theorem 2.1.1. □

Remark 2.1.2. Instead of typing out "the free group on two generators," we will henceforth use $F(a, b)$ to refer to the free group on two generators.

We can apply Theorem 2.1.1 to show the existence of a set of isometries of $\mathbb{R}^n$ that is isomorphic to $F(a, b)$.

Definition 2.1.2 (Special Orthogonal Group)**.** For $n \in \mathbb{N}$, we define $SO(n)$ to be the group of all real $n \times n$ matrices A such that $A^T = A^{-1}$ and $\det(A) = 1$.

In terms of an isometry of $\mathbb{R}^3$, the group $SO(3)$ denotes the set of all rotations about any line through the origin.

Theorem 2.1.2. There are elements $a, b \in SO(3)$ such that $\langle a, b\rangle_{SO(3)} \cong F(a, b)$.

Proof. We let

$$\begin{aligned} a &= \begin{pmatrix} 3/5 & 4/5 & 0 \\ -4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ a^{-1} &= \begin{pmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ b &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & -4/5 \\ 0 & 4/5 & 3/5 \end{pmatrix} \\ b^{-1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & 4/5 \\ 0 & -4/5 & 3/5 \end{pmatrix}. \end{aligned}$$

We specify

$$X = A_+ \sqcup A_- \sqcup B_+ \sqcup B_- \sqcup \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

where

$$A_+ = \left\{ \frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, k \in \mathbb{Z}, x \equiv 3y \text{ modulo } 5, z \equiv 0 \text{ modulo } 5 \right\}$$

$$A_- = \left\{ \frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, k \in \mathbb{Z}, x \equiv -3y \text{ modulo } 5, z \equiv 0 \text{ modulo } 5 \right\}$$

$$B_+ = \left\{ \frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, k \in \mathbb{Z}, z \equiv 3y \text{ modulo } 5, x \equiv 0 \text{ modulo } 5 \right\}$$

$$B_- = \left\{ \frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, k \in \mathbb{Z}, z \equiv -3y \text{ modulo } 5, x \equiv 0 \text{ modulo } 5 \right\}.$$

To verify that the conditions of Theorem 2.1.1 hold, we calculate

$$\begin{pmatrix} 3/5 & 4/5 & 0 \\ -4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix}\left(\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \frac{1}{5^{k+1}}\begin{pmatrix} 3x + 4y \\ -4x + 3y \\ 5z \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix}\left(\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \frac{1}{5^{k+1}}\begin{pmatrix} 3x - 4y \\ 4x + 3y \\ 5z \end{pmatrix} \tag{2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & -4/5 \\ 0 & 4/5 & 3/5 \end{pmatrix}\left(\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \frac{1}{5^{k+1}}\begin{pmatrix} 5x \\ 3y - 4z \\ 4y + 3z \end{pmatrix} \tag{3}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & 4/5 \\ 0 & -4/5 & 3/5 \end{pmatrix}\left(\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \frac{1}{5^{k+1}}\begin{pmatrix} 5x \\ 3y + 4z \\ -4y + 3z \end{pmatrix}. \tag{4}$$

We verify that the conditions for Corollary 2.1.1 hold for each of these four conditions.

(1) For any vector

$$\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \notin A_-,$$

we see that $k + 1 \in \mathbb{Z}$, $x' = 3x + 4y \equiv 3(-4x + 3y)$ modulo 5, and that $z' = 5z \equiv 0$ modulo 5.

(2) For any vector

$$\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \notin A_+,$$

we see that $k + 1 \in \mathbb{Z}$, $x' = 3x - 4y \equiv -3(4x + 3y)$ modulo 5, and $z' = 5z \equiv 0$ modulo 5.

(3) For any vector

$$\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \notin B_-,$$

we see that $k + 1 \in \mathbb{Z}$, $z' = 4y + 3z \equiv 3(3y - 4z)$ modulo 5, and $x' = 5x \equiv 0$ modulo 5.

(4) For any vector

$$\frac{1}{5^k}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \notin B_+,$$

we see that $k+1 \in \mathbb{Z}$, $z' = -4y + 3z \equiv -3(3y+4z)$ modulo 5, and $x' = 5x \equiv 0$ modulo 5.

Thus, by Theorem 2.1.1 and Corollary 2.1.1, it is the case that $\langle a, b \rangle \cong F(a,b)$. □

2.2 Introducing Paradoxical Decompositions

We now turn our attention towards "paradoxical" actions that seem to recreate a set by using disjoint proper subsets. This will allow us to use the result from Theorem 2.1.2 to move towards the Banach–Tarski paradox.

Definition 2.2.1 (Paradoxical Decompositions and Paradoxical Groups)**.** Let G be a group that acts on a set X, with $E \subseteq X$. We say E is G-paradoxical if there exist pairwise disjoint proper subsets $A_1, \ldots, A_n$ and $B_1, \ldots, B_m$ of E and group elements $g_1, \ldots, g_n, h_1, \ldots, h_m \in G$ such that

$$E = \bigcup_{j=1}^{n} g_j \cdot A_j$$

and

$$E = \bigcup_{j=1}^{m} h_j \cdot B_j.$$

If G acts on itself by left-multiplication, and G satisfies these conditions, we say G is a paradoxical group.

Example 2.2.1. The free group on two generators, $F(a,b)$, is a paradoxical group.

To see that $F(a,b)$ is a paradoxical group, we let $W(x)$ denote the set of words in $F(a,b)$ that start with $x \in \{a, b, a^{-1}, b^{-1}\}$. For instance, $ba^2ba^{-1} \in W(b)$.

Since every word in F is either the empty word, or starts with one of a, b, a^{-1}, b^{-1}, we see that

$$F(a,b) = \{e_{F(a,b)}\} \sqcup W(a) \sqcup W(b) \sqcup W\left(a^{-1}\right) \sqcup W\left(b^{-1}\right).$$

If $w \in F(a,b) \setminus W(a)$, we see that $a^{-1}w \in W(a^{-1})$. Thus, $w \in aW(a^{-1})$. For any $t \in F(a,b)$ either $t \in W(a)$ or $t \in F(a,b) \setminus W(a) = aW(a^{-1})$. Thus, $F(a,b)$ is equal to $W(a) \sqcup aW(a^{-1})$.

Similarly, if $t \in F(a,b)$ either $t \in W(b)$ or $t \in F(a,b) \setminus W(b) = bW(b^{-1})$, so $F(a,b)$ is equal to $W(b) \sqcup bW(b^{-1})$.

We have thus constructed

$$\begin{aligned} F(a,b) &= W(a) \sqcup aW\left(a^{-1}\right) \\ &= W(b) \sqcup bW\left(b^{-1}\right), \end{aligned}$$

a paradoxical decomposition of $F(a,b)$ with the action of left-multiplication.

Now that we understand a little more about paradoxical groups, we now want to understand the actions of paradoxical groups on sets.

Proposition 2.2.1. Let G be a paradoxical group that acts freely on X. Then, X is G-paradoxical.

Proof. Let $A_1, \dots, A_n, B_1, \dots, B_m \subset G$ be pairwise disjoint, and let $g_1, \dots, g_n, h_1, \dots, h_m \in G$ such that

$$\begin{aligned} G &= \bigcup_{j=1}^{n} g_j A_j \\ &= \bigcup_{j=1}^{m} h_j B_j. \end{aligned}$$

Let $M \subseteq X$ contain exactly one element from every orbit in X.

Claim. The set $\{g \cdot M \mid g \in G\}$ is a partition of X.

Proof of Claim: Since M contains exactly one element from every orbit in X, it is the case that $G \cdot M = X$, so

$$\bigcup_{g \in G} g \cdot M = X$$

Additionally, for $x, y \in M$, if $g \cdot x = h \cdot y$, then $\left(h^{-1} g\right) \cdot x = y$, meaning y is in the orbit of x and vice versa, implying $x = y$. Since G acts freely on X, we must have $h^{-1} g = e_G$.

Thus, we can see that $g_1 \cdot M \neq g_2 \cdot M$, implying $\{g \cdot M \mid g \in G\}$ is a partition of X. □

We define

$$A_j^* = \bigcup_{g \in A_j} g \cdot M,$$

and similarly define

$$B_j^* = \bigcup_{h \in B_j} h \cdot M.$$

As a useful shorthand, we can also write $A_j^* = A_j \cdot M$, and similarly, $B_j^* = B_j \cdot M$, to denote the union of the elements of A_j and B_j respectively acting on M.

Since $\{g \cdot M \mid g \in G\}$ is a partition of X, and $A_1, \dots, A_n, B_1, \dots, B_m \subset G$ are pairwise disjoint, it must be the case that $A_1^*, \dots, A_n^*, B_1^*, \dots, B_m^* \subset X$ are also pairwise disjoint.

For the original $g_1, \dots, g_n, h_1, \dots, h_m$ that defined the paradoxical decomposition of G, we thus have

$$\begin{aligned} \bigcup_{j=1}^{n} g_j \cdot A_j^* &= \bigcup_{j=1}^{n} \left(g_j A_j\right) \cdot M \\ &= G \cdot M \\ &= X, \end{aligned}$$

and

$$\bigcup_{j=1}^{m} h_j \cdot B_j^* = \bigcup_{j=1}^{m} \left(h_j B_j\right) \cdot M$$

$$= G \cdot M$$
$$= X.$$

Thus, X is G-paradoxical. □

Remark 2.2.1. This proof requires the axiom of choice, as we invoked it to define M to contain exactly one element from every orbit in X.

2.3 The Weak Banach–Tarski Paradox

Now that we have established $F(a, b)$ as being a paradoxical group, we wish to use it to construct paradoxical decompositions of the unit sphere $S^2 \subseteq \mathbb{R}^3$. Specifically, we will show a weak version of the Banach–Tarski paradox — one where you can break apart the unit ball into finitely many pieces and reconstitute it into two copies of itself.

Fact 2.3.1. If H is a paradoxical group, and $H \leqslant G$, then G is a paradoxical group.

With this fact in mind, we will show that SO(3) is a paradoxical group.

Theorem 2.3.1. There are rotations A and B that about lines through the origin in $\mathbb{R}^3$ that generate a subgroup of SO(3) isomorphic to $F(a, b)$

Proof. We take A and B as in the proof of Theorem 2.1.2. □

Remark 2.3.1. Since SO(n) contains a subgroup isomorphic to SO(3) for all $n \geqslant 3$, it is the case that SO(n) also contains a subgroup isomorphic to $F(a, b)$.

Since we have shown that SO(3) is paradoxical, as it contains a paradoxical subgroup, we can now begin to examine the action of SO(3) on subsets of $\mathbb{R}^3$.

Theorem 2.3.2 (Hausdorff Paradox)**.** There is a countable subset D of S^2 such that $S^2 \setminus D$ is SO(3)-paradoxical.

Proof. Let A and B be the rotations in SO(3) that serve as the generators of the subgroup isomorphic to $F(a, b)$ (as in 2.1.2).

Since A and B are rotations, so too is any element of $\langle A, B\rangle$. Thus, any such non-empty word contains two fixed points.

We let

$$F = \left\{x \in S^2 \mid x \text{ is a fixed point for some word } w\right\}.$$

Since $\langle A, B\rangle$ is countably infinite, so too is F. Thus, the union of all these fixed points under the action of all such words w is countable.

$$D = \bigcup_{w \in \langle A, B\rangle} w \cdot F.$$

Therefore, $\langle A, B\rangle$ acts freely on $S^2 \setminus D$, so $S^2 \setminus D$ is SO(3)-paradoxical. □

Unfortunately, the Hausdorff paradox is not enough for us to be able to prove the Banach–Tarski paradox. In order to do this, we need to be able to show that two sets are "similar" under the action of a group.

Definition 2.3.1 (Equidecomposable Sets)**.** Let G act on X, and let $A, B \subseteq X$. We say A and B are G-equidecomposable if there are partitions $\left\{A_j\right\}_{j=1}^{n}$ of A and $\left\{B_j\right\}_{j=1}^{n}$ of B, and elements $g_1, \ldots, g_n \in G$, such that for all j,

$$B_j = g_j \cdot A_j.$$

We write $A \sim_G B$ if A and B are G-equidecomposable.

Fact 2.3.2. The relation $\sim_G$ is an equivalence relation.

Proof. Let A, B, and C be sets.

To show reflexivity, we can select $g_1 = g_2 = \cdots = g_n = e_G$. Thus, $A \sim_G A$.

To show symmetry, let $A \sim_G B$. Set $\{A_j\}_{j=1}^n$ to be the partition of A, and set $\{B_j\}_{j=1}^n$ to be the partition of B, such that there exist $g_1, \ldots, g_n \in G$ with $g_j \cdot A_j = B_j$. Then,

$$\begin{aligned} g_j^{-1} \cdot (g_j \cdot A_j) &= g_j^{-1} \cdot B_j \\ A_j &= g_j^{-1} \cdot B_j, \end{aligned}$$

so $B_j \sim_G A_j$.

To show transitivity, let $A \sim_G B$ and $B \sim_G C$. Let $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^n$ be the partitions of A and B respectively and $g_1, \ldots, g_n \in G$ such that $g_i \cdot A_i = B_i$. Let $\{B_j\}_{j=1}^m$ and $\{C_j\}_{j=1}^m$ be partitions of B and C, and $h_1, \ldots, h_m \in G$, such that $h_j \cdot B_j = C_j$.

We refine the partition of A to A_{ij} by taking $A_{ij} = g_i^{-1}(B_i \cap B_j)$, where $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Then, $(h_j g_i) \cdot A_{ij}$ maps the refined partition of A to C, so A and C are G-equidecomposable. □

Fact 2.3.3. For $A \sim_G B$, there is a bijection $\phi \colon A \to B$ by taking $C_i = C \cap A_i$, and mapping $\phi(C_i) = g_i \cdot C_i$.

In particular, this means that for any subset $C \subseteq A$, it is the case that $C \sim \phi(C)$.

We can now use this equidecomposability to glean information about the existence of paradoxical decompositions.

Proposition 2.3.1. Let G act on X, with $E, E' \subseteq X$ such that $E \sim_G E'$. Then, if E is G-paradoxical, then so too is E′.

Proof. Let $A_1, \ldots, A_n, B_1, \ldots, B_m \subset E$ be pairwise disjoint, with $g_1, \ldots, g_n, h_1, \ldots, h_m \in G$ such that

$$\begin{aligned} E &= \bigcup_{i=1}^n g_i \cdot A_i \\ &= \bigcup_{j=1}^m h_j \cdot B_j. \end{aligned}$$

We let

$$A = \bigsqcup_{i=1}^n A_i$$

$$B = \bigsqcup_{j=1}^m B_j.$$

It follows that $A \sim_G E$ and $B \sim_G E$, since we can take the partition of A to be $A_1, \ldots, A_n$, and partition E by taking $g_i \cdot A_i$ for $i = 1, \ldots, n$, and similarly for B.

Since $E \sim_G E'$, and $\sim_G$ is an equivalence relation, it follows that $A \sim_G E'$ and $B \sim_G E'$. Thus, there is a paradoxical decomposition of E' in $A_1, \dots, A_n$ and $B_1, \dots, B_m$. ◻

We will now show that S^2 is SO(3) paradoxical.

Proposition 2.3.2. Let $D \subseteq S^2$ be countable. Then, S^2 and $S^2 \setminus D$ are SO(3)-equidecomposable.

Proof. Let L be a line in $\mathbb{R}^3$ such that $L \cap D = \emptyset$. Such an L must exist since S^2 is uncountable.

Define $\rho_\theta \in SO(3)$ to be a rotation about L by an angle of θ. For a fixed $n \in \mathbb{N}$ and fixed $\theta \in [0, 2\pi)$, define $R_{n,\theta} = \left\{x \in D \mid \rho_\theta^n \cdot x \in D\right\}$. Since D is countable, $R_{n,\theta}$ is necessarily countable.

We define $W_n = \{\theta \mid R_{n,\theta} \neq \emptyset\}$. Since the map $\theta \mapsto \rho_\theta^n \cdot x$ into D is injective, it is the case that W_n is countable. Therefore,

$$W = \bigcup_{n \in \mathbb{N}} W_n$$

is countable.

Thus, there must exist $\omega \in [0, 2\pi) \setminus W$. We define ρ_ω to be a rotation about L by ω. Then, for every $n, m \in \mathbb{N}$, we have

$$\rho_\omega^n \cdot D \cap \rho_\omega^m \cdot D = \emptyset.$$

We define $\widetilde{D} = \bigsqcup_{n=0}^{\infty} \rho_\omega^n D$. Note that

$$\begin{aligned}
\rho_\omega \cdot \widetilde{D} &= \rho_\omega \cdot \bigsqcup_{n=0}^{\infty} \rho_\omega^n \cdot D \\
&= \bigsqcup_{n=1}^{\infty} \rho_\omega^n \cdot D \\
&= \widetilde{D} \setminus D,
\end{aligned}$$

meaning $\widetilde{D}$ and D are SO(3)-equidecomposable.

Thus, we have

$$\begin{aligned}
S^2 &= \widetilde{D} \sqcup \left(S^2 \setminus \widetilde{D}\right) \\
&\sim_{SO(3)} \left(\rho_\omega \cdot \widetilde{D}\right) \sqcup \left(S^2 \setminus \widetilde{D}\right) \\
&= \left(\widetilde{D} \setminus D\right) \sqcup \left(S^2 \setminus \widetilde{D}\right) \\
&= S^2 \setminus D,
\end{aligned}$$

establishing S^2 and $S^2 \setminus D$ as SO(3)-equidecomposable.

In particular, this means S^2 is also SO(3)-paradoxical. ◻

To prove the Banach–Tarski paradox, we need a slightly larger group than SO(3) — one that includes translations in addition to the traditional rotations.

Definition 2.3.2 (Euclidean Group)**.** The Euclidean group, E(n), consists of all isometries of a Euclidean space. An isometry of a Euclidean space consists of translations, rotations, and reflections.

Corollary 2.3.1 (Weak Banach–Tarski Paradox)**.** Every closed ball in $\mathbb{R}^3$ is E(3)-paradoxical.

Proof. We only need to show that $B(0,1)$ is E(3)-paradoxical. To do this, we start by showing that $B(0,1) \setminus \{0\}$ is SO(3)-paradoxical.

Since S^2 is SO(3)-paradoxical, there exists pairwise disjoint subsets $A_1, \dots, A_n, B_1, \dots, B_m \subset S^2$ and elements $g_1, \dots, g_n, h_1, \dots, h_m \in$ SO(3) such that

$$\begin{aligned} S^2 &= \bigcup_{i=1}^{n} g_i \cdot A_i \\ &= \bigcup_{j=1}^{m} h_j \cdot B_j. \end{aligned}$$

Define

$$\begin{aligned} A_i^* &= \{tx \mid t \in (0,1], x \in A_i\} \\ B_j^* &= \left\{ty \mid t \in (0,1], y \in B_j\right\}. \end{aligned}$$

Then, $A_1^*, \dots, A_n^*, B_1^*, \dots, B_m^* \subset B(0,1) \setminus \{0\}$ are pairwise disjoint, and

$$\begin{aligned} B(0,1) \setminus \{0\} &= \bigcup_{i=1}^{n} g_i \cdot A_i^* \\ &= \bigcup_{j=1}^{m} h_j \cdot B_j^*. \end{aligned}$$

Thus, we have established that $B(0,1) \setminus \{0\}$ is E(3)-paradoxical.

Now, we want to show that $B(0,1) \setminus \{0\}$ and $B(0,1)$ are E(3)-equidecomposable. Let $x \in B(0,1) \setminus \{0\}$, and let ρ be a rotation through x by a line not through the origin such that $\rho^n \cdot 0 \neq \rho^m \cdot 0$ when $n \neq m$.

Let $D = \{\rho^n \cdot 0 \mid n \in \mathbb{N}\}$. We can see that $\rho \cdot D = D \setminus \{0\}$, and that D and $\rho \cdot D$ are E(3)-equidecomposable. Thus,

$$\begin{aligned} B(0,1) &= D \sqcup (B(0,1) \setminus D) \\ &\sim_{E(3)} (\rho \cdot D) \sqcup (B(0,1) \setminus D) \\ &= (D \setminus \{0\}) \sqcup (B(0,1) \setminus D) \\ &= B(0,1) \setminus \{0\}. \end{aligned}$$

Therefore, $B(0,1)$ is E(3)-paradoxical. □

2.4 The Strong Banach–Tarski Paradox

In order to prove the general case of the Banach–Tarski paradox, we need one more piece of mathematical machinery.

In Fact 2.3.2, we showed that the relation $A \sim_G B$ if and only if A and B are G-equidecomposable is an equivalence relation. Using the power of subsets, we may extend this to a preorder on any subsets A and B of X.

Definition 2.4.1. Let G act on a set X with $A, B \subseteq X$. We write $A \preceq_G B$ if A is equidecomposable with a subset of B.

Fact 2.4.1. The relation $\preceq_G$ is a reflexive and transitive relation.

Proof. To see reflexivity, we can see that since $A \sim_G A$, and $A \subseteq A$, $A \preceq_G A$.

To see transitivity, let $A \preceq_G B$ and $B \preceq_G C$. Then, there exist $g_1, \dots, g_n \in G$ such that $g_i \cdot A_i = B_{\alpha,i}$ for each i, where $A \sim_G B_\alpha \subseteq B$. Similarly, there exist $h_1, \dots, h_m \in G$ such that $h_j \cdot B_j = C_{\beta,j}$ for each j, where $B \sim_G C_\beta \subseteq C$.

We take a refinement of B by taking intersections $B_{\alpha,ij} = B_{\alpha,i} \cap B_j$, with $i = 1, \dots, n$ and $j = 1, \dots, m$. We define $C_{\beta,\alpha,ij} = h_j \cdot B_{\alpha,ij}$ for each $j = 1, \dots, m$. Then, $h_j g_i \cdot A_i = C_{\beta,\alpha,ij}$, meaning $A \sim_G C_{\beta,\alpha,ij} \subseteq C_\beta \subseteq C$, so $A \preceq_G C$. □

We know from Fact 2.3.3 that $A \preceq_G B$ implies the existence of a bijection $\phi\colon A \to B' \subseteq B$, meaning $\phi\colon A \hookrightarrow B$ is an injection. Similarly, if $B \preceq_G A$, then Fact 2.3.3 implies the existence of an injection $\psi\colon B \hookrightarrow A$.

One may ask if an analogue of the Cantor–Schröder–Bernstein theorem exists in the case of the relation $\preceq_G$, implying that the preorder established in Fact 2.4.1 is indeed a partial order. The following theorem establishes this result.

Theorem 2.4.1. Let G act on X, and let $A, B \subseteq X$. If $A \preceq_G B$ and $B \preceq_G A$, then $A \sim_G B$.

Proof. Let $B' \subseteq B$ with $A \sim_G B'$, and let $A' \subseteq A$ with $B \sim_G A'$. Then, we know from Fact 2.3.3 that there exist bijections $\phi\colon A \to B'$ and $\psi\colon B \to A'$.

Define $C_0 = A \setminus A'$, and $C_{n+1} = \psi(\phi(C_n))$. We set

$$C = \bigcup_{n \geqslant 0} C_n.$$

Since $\psi^{-1}(\psi(\phi(C_n))) = \phi(C_n)$, we have

$$\psi^{-1}(A \setminus C) = B \setminus \phi(C).$$

Having established in Fact 2.3.3 that for any subset of $C \subseteq A$, $C \sim_G \phi(C)$, we also see that $A \setminus C \sim_G B \setminus \phi(C)$.

Thus, we can see that

$$\begin{aligned} A &= (A \setminus C) \sqcup C \\ &\sim_G (B \setminus \phi(C)) \sqcup \phi(C) \\ &= B. \end{aligned}$$

□

Finally, we are able to prove Proposition 2.0.1. We restate the proposition here, followed by its proof.

Proposition 2.0.1 (Strong Banach–Tarski Paradox)**.** Let A and B be bounded subsets of $\mathbb{R}^3$ with nonempty interior. There is a partition of A into finitely many disjoint subsets such that a sequence of isometries applied to these subsets yields B.

Proof of Proposition 2.0.1: By symmetry, it is enough to show that $A \preceq_{E(3)} B$.

Since A is bounded, there exists $r > 0$ such that $A \subseteq B(0, r)$.

Let $x_0 \in B^\circ$. Then, there exists $\varepsilon > 0$ such that $B(x_0, \varepsilon) \subseteq B$.

Since $B(0, r)$ is compact (hence totally bounded), there are translations $g_1, \ldots, g_n$ such that

$$B(0, r) \subseteq g_1 \cdot B(x_0, \varepsilon) \cup \cdots \cup g_n \cdot B(x_0, \varepsilon).$$

We select translations $h_1, \ldots, h_n$ such that $h_j \cdot B(x_0, \varepsilon) \cap h_k \cdot B(x_0, \varepsilon) = \emptyset$ for $j \neq k$. We set

$$S = \bigcup_{j=1}^{n} h_j \cdot B(x_0, \varepsilon).$$

Each $h_j \cdot B(x_0, \varepsilon) \subseteq S$ is E(3)-equidecomposable with any arbitrary closed ball subset of $B(x_0, \varepsilon)$, it is the case that $S \preceq B(x_0, \varepsilon)$.

Thus, we have

$$\begin{aligned} A &\subseteq B(0, r) \\ &\subseteq g_1 \cdot B(x_0, \varepsilon) \cup \cdots \cup b_n \cdot B(x_0, \varepsilon) \\ &\preceq S \\ &\preceq B(x_0, \varepsilon) \\ &\preceq B. \end{aligned}$$

$\square$

Chapter 3

Resolving A Paradox: Tarski's Theorem

Ultimately, the reason the Banach–Tarski paradox "works" is because the paradoxical group $F(a, b)$, lacks a property known as amenability — specifically, that a group admitting a paradoxical decomposition is not amenable. Before we go further into the characterizations of amenability discussed in chapters 5 and 6, we will show that this statement reverses. Indeed, every amenable group is *non*-paradoxical.

> **Theorem 3.0.1** (Tarski's Theorem)**.** Let G be a group that acts on a set X, and let $E \subseteq X$ be nonempty. There is a finitely additive translation-invariant measure $\mu\colon P(X) \to [0, \infty]$ with $\mu(E) \in (0, \infty)$ if and only if E is not G-paradoxical.

We can prove one of the directions of Tarski's theorem now.

Proof of the Forward Direction of Theorem 3.0.1: Let E be G-paradoxical. Suppose toward contradiction that such a translation-invariant finitely additive ν existed with $\nu(E) \in (0, \infty)$.

Let $A_1, \ldots, A_n, B_1, \ldots, B_m \subseteq E$ be pairwise disjoint, and let $t_1, \ldots, t_n, s_1, \ldots, f_m \in G$ such that

$$\begin{aligned} E &= \bigsqcup_{i=1}^{n} t_i \cdot A_i \\ &= \bigsqcup_{j=1}^{m} s_j \cdot B_j. \end{aligned}$$

Then, it would be the case that

$$\begin{aligned} \nu(E) &= \nu\left(\bigsqcup_{i=1}^{n} t_i \cdot A_i\right) \\ &= \sum_{i=1}^{n} \nu(t_i \cdot A_i) \\ &= \sum_{i=1}^{n} \nu(A_i), \end{aligned}$$

and

$$\nu(E) = \sum_{j=1}^{m} \nu\big(B_j\big).$$

However, this also yields

$$\nu(E) = \nu\left(\left(\bigsqcup_{i=1}^{n} A_i\right) \sqcup \left(\bigsqcup_{j=1}^{m} B_j\right)\right)$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \nu(A_i) + \sum_{j=1}^{m} \nu(B_j) \\
&= \sum_{i=1}^{n} \nu(t_i \cdot A_i) + \sum_{j=1}^{m} \nu(x_j \cdot B_j) \\
&= \nu(E) + \nu(E) \\
&= 2\nu(E).
\end{aligned}
$$

implying that $\nu(E) = 0$ or $\nu(E) = \infty$. □

The opposite direction, unfortunately, will be significantly harder to prove. We will need to know some results from graph theory, understand the properties of the type semigroup of an action, and use some results on commutative semigroups to show the existence of a mean.

3.1 A Little Bit of Graph Theory

To prove the reverse direction of Tarski's theorem, we need to develop some machinery from graph theory that will allow us to prove that a certain semigroup we will construct in the next section satisfies the cancellation identity.

We start by defining graphs and paths, before proving a special case of Hall's theorem, ultimately extending to the infinite case with König's theorem.

Definition 3.1.1 (Graphs and Paths)**.** A graph is a triple (V, E, ϕ), with V, E nonempty sets and $\phi\colon E \to P_2(V)$ a map from E to the set of all unordered subset pairs of V.

For $e \in E$, if $\phi(e) = \{v, w\}$, then we say v and w are the endpoints of e, and e is incident on v and w.

A path in (V, E, ϕ) is a finite sequence $(e_1, \dots, e_n)$ of edges, with a finite sequence of vertices $(v_0, \dots, v_n)$, such that $\phi(e_k) = \{v_{k-1}, v_k\}$.

The degree of a vertex, $\deg(v)$, is the number of edges incident on v.

We define the neighbors of $S \subseteq V$ to be the set of all vertices $v \in V \setminus S$ such that v is an endpoint to an edge incident on S. We denote this set $N(S)$.

Definition 3.1.2 (Bipartite Graphs and k-Regularity)**.** Let (V, E, ϕ) be a graph, with $k \in \mathbb{N}$.

(i) If $\deg(v) = k$ for each $v \in V$, we say (V, E, ϕ) is k-regular.

(ii) If $V = X \sqcup Y$, with each edge in E having one endpoint in X and one endpoint in Y, then we say V is bipartite, and write (X, Y, E, ϕ).

Definition 3.1.3 (Perfect Matching)**.** Let (X, Y, E, ϕ) be a bipartite graph. Let $A \subseteq X$ and $B \subseteq Y$. A perfect matching of A and B is a subset $F \subseteq E$ with

(i) each element of $A \cup B$ is an endpoint of exactly one $f \in F$;

(ii) all endpoints of edges in F are in $A \cup B$.

Definition 3.1.4 (Hall Condition)**.** We say a bipartite graph (X, Y, E, ϕ) satisfies the Hall Condition on X if, for all $S \subseteq X$, $|N(S)| \geqslant |S|$.

Equivalently, we say a (finite) collection of not necessarily distinct finite sets $\mathcal{X} = \{X_i\}_{i=1}^{n}$ satisfies the

Hall Condition if and only if for all subcollections $\mathcal{Y}_k = \{X_{i_k}\}_{k=1}^m$,

$$|\mathcal{Y}_k| \leqslant \left|\bigcup_{k=1}^m X_{i_k}\right|.$$

Remark 3.1.1. These two formulations of the Hall condition are equivalent regarding an X-perfect matching.

Theorem 3.1.1 (Hall's Theorem for Finite k-Regular Bipartite Graphs)**.** Let (X, Y, E, ϕ) be a k-regular bipartite graph for some $k \in \mathbb{N}$, and let $V = X \sqcup E$ be finite. Then, there is a perfect matching of X and Y.

Proof. Note that since $|E| = k|K| = k|Y|$, it is the case that $|X| = |Y|$.

Let $M \subseteq V$ be any subset. We will show that $|N(M)| \geqslant |M|$ — that is, (X, Y, E, ϕ) satisfies the Hall condition.

Let $M_X = M \cap X$ and $M_Y = M \cap Y$, where $M = M_X \sqcup M_Y$. Let $[M_X, N(M_X)]$ be the set of edges with endpoints in M_X and $N(M_X)$, and $[M_Y, N(M_Y)]$ be the set of edges with endpoints in M_Y and $N(M_Y)$. We also let $[X, N(M_X)]$ denote the set of edges with endpoints in X and $N(M_X)$, and similarly, $[Y, N(M_Y)]$ is the set of edges with endpoints in Y and $N(M_Y)$.

We can see that $[M_X, N(M_X)] \subseteq [X, N(M_X)]$, and similarly, $[M_Y, N(M_Y)] \subseteq [Y, N(M_Y)]$.

Since $|[M_X, N(M_X)]| = k|M_X|$ and $|[X, N(M_X)]| = k|N(M_X)|$, we have

$$|M_X| \leqslant |N(M_X)|,$$

and similarly,

$$|M_Y| \leqslant |N(M_Y)|.$$

Thus, $|M| \leqslant |N(M)|$.

We will now show that there is an X-perfect matching. Suppose toward contradiction that F is a maximal perfect matching on $A \subseteq X$ and $B \subseteq Y$ with $X \setminus A \neq \emptyset$.

Then, there is $x \in X \setminus A$. Consider $Z \subseteq V$ consisting of all vertices z such that there exists a F-alternating path $(e_1, \ldots, e_n)$ between $z \in Z$ and x.

It cannot be the case that $Z \cap Y$ is empty, since the number of neighbors of x is greater than or equal to 1 by the Hall condition — if it were the case that $Z \cap Y$ were empty, we could add an edge to F consisting of x and one element of $N(\{x\})$, which would contradict the maximality of F.

Consider a path traversing along Z, $(e_1, \ldots, e_n)$. It must be the case that $e_n \in F$, or else we would be able to "flip" the matching F by exchanging e_i with e_{i+1} for $e_i \in F$, which would contradict the maximality of F yet again. Thus, every element of $Z \cap Y$ is satisfied by F, so $Z \cap Y \subseteq B$.

Since each element in $Z \cap Y$ is paired with exactly one element of $Z \cap X$ (with one left over), it is the case that $|Z \cap X| = |Z \cap Y| + 1$.

Suppose toward contradiction that there exists $y \in N(Z \cap X)$ with $y \notin Z \cap Y$. Then, there exists $v \in Z \cap X$ and $e \in E$ such that $\phi(e) = \{v, y\}$. However, this means v is connected via a path to x, meaning $y \in Z$, so $y \in Z \cap Y$. Thus, we must have $N(Z \cap X) = Z \cap Y$.

Therefore,

$$|Z \cap X| = |Z \cap Y| + 1$$

$$= |N(Z \cap X)| + 1,$$

which contradicts the fact that (X, Y, E, ϕ) satisfies the Hall condition. Therefore, $A = X$.

By symmetry, there is a perfect matching of X and Y in (X, Y, E, ϕ). □

Remark 3.1.2. An equivalent formulation to Hall's theorem states that there is a system of distinct representatives on the collection $\mathcal{X} = \{X_k\}_{k=1}^n$, which is a set $\{x_k\}_{k=1}^n$ such that $x_k \in X_k$ and $x_i \neq x_j$ for $i \neq j$.

This implies the existence of an injection $f\colon \mathcal{X} \hookrightarrow \bigcup_{k=1}^n X_k$, such that $f(X_k) \in X_k$.

Theorem 3.1.2 (Infinite Hall's Theorem)**.** Let $\mathcal{G} = \{X_i\}_{i \in I}$ be a collection of (not necessarily distinct) finite sets. If, for every finite subcollection $\mathcal{Y} = \{X_{i_k}\}_{k=1}^n$,

$$n \leqslant \left| \bigcup_{k=1}^n X_{i_k} \right|,$$

then there is a choice function on G.

Proof. We endow each $X_i \in \{X_i\}_{i \in I}$ with the discrete topology. Since each X_i is finite, each X_i is compact.

Thus, by Tychonoff's theorem, it is the case that $\prod_{i \in I} X_i$ is compact.

For every finite subset $Y \subseteq \mathcal{G}$, we define

$$S_Y = \left\{ f \in \prod_{i \in I} X_i \,\middle|\, f|_Y \text{ is injective} \right\}.$$

The injectivity of $f|_Y$ is equivalent to the existence of a system of distinct representatives on Y. Since Y satisfies the Hall condition, each S_Y is nonempty. Additionally, for any net of functions $f_\alpha \in S_Y$ with $\lim_\alpha f_\alpha = f$, it is the case that $f_\alpha|_Y$ is injective, so $f|_Y$ is injective, meaning S_Y is closed.

We define $F = \{S_Y \mid Y \subseteq \mathcal{G} \text{ finite}\}$. For finite $Y_1, Y_2 \subseteq \mathcal{G}$, every system of distinct representatives in $Y_1 \cup Y_2$ is necessarily a system of distinct representatives on Y_1 and a system of distinct representatives on Y_2, meaning $S_{Y_1 \cup Y_2} \subseteq S_{Y_1} \cap S_{Y_2}$. Thus, F has the finite intersection property.

Since $\prod_{i \in I} X_i$ is compact, $\bigcap F$ is nonempty, where the intersection is taken over all finite subsets of $\mathcal{G}$. For any $f \in \bigcap F$, f is necessarily a choice function. □

Remark 3.1.3. This is equivalent to the existence of an injection $f\colon \mathcal{G} \hookrightarrow \bigcup_{i \in I} X_i$.

We will use this infinite case of Hall's theorem to prove König's theorem.

Theorem 3.1.3 (König's Theorem)**.** Let (X, Y, E, ϕ) be a k-regular bipartite graph (not necessarily finite). Then, there is a perfect matching of X and Y.

Proof. If $k = 1$, it is clear that there is a perfect matching in (X, Y, E, ϕ) consisting of the edges in (X, Y, E, ϕ).

Let $k \geqslant 2$. Since any finite subset of X satisfies the Hall condition, as displayed in the proof of Theorem 3.1.1, there is some X-perfect matching in (X, Y, E, ϕ). We call this X-perfect matching F. There is an injection $f\colon X \hookrightarrow Y$ following the edges in F.

Similarly, since any finite subset of Y satisfies the Hall condition, there is some Y-perfect matching in

(X, Y, E, ϕ). We call this Y-perfect matching G. There is an injection $g\colon Y \hookrightarrow X$ following the edges of G.

Consider the subgraph $(X, Y, F \cup G, \phi|_{F \cup G})$. The injections f and g still hold in this graph. By the Cantor–Schröder–Bernstein theorem, there is a bijection $h\colon X \to Y$ in $(X, Y, F \cup G, \phi|_{F \cup G})$, which is equivalent to the existence of a perfect matching of X and Y. $\square$

3.2 Type Semigroups

Definition 3.2.1. Let G be a group that acts on a set X.

(i) We define $X^* = X \times \mathbb{N}_0$, and

$$G^* = \{(g, \pi) \mid g \in G, \pi \in \mathrm{Sym}(\mathbb{N}_0)\}.$$

(ii) If $A \subseteq X^*$, the values of n for which there is an element of A whose second coordinate is n are called the levels of A.

Fact 3.2.1. If $E_1, E_2 \subseteq X$, then $E_1 \sim_G E_2$ if and only if $E_1 \times \{n\} \sim_{G^*} E_2 \times \{m\}$ for all $m, n \in \mathbb{N}_0$.

Proof. Let $E_1 \sim_G E_2$. Then, there exist pairwise disjoint $A_1, \ldots, A_n \subset E_1$, pairwise disjoint $B_1, \ldots, B_n \subset E_2$, and elements $g_1, \ldots, g_n \in G$ such that $g_i \cdot A_i = B_i$. We select the permutation $\pi_i \in \mathrm{Sym}(\mathbb{N}_0)$ such that $\pi_i(n) = m$ and $\pi_i(m) = n$ for each i. Then,

$$(g_i, \pi_i) \cdot (A_i \cdot \{n\}) = B_i \cdot \{m\}.$$

Similarly, if $E_1 \times \{n\} \sim_{G^*} E_2 \times \{m\}$, then of the pairwise disjoint subsets

$$A_1 \times \{n\}, \ldots, A_n \times \{n\} \subset E_1 \times \{n\}$$

and

$$B_1 \times \{m\}, \ldots, B_n \times \{m\} \subset E_2 \times \{m\},$$

we set $A_1, \ldots, A_n \subset E_1$ and $B_1, \ldots, B_n \subset E_2$. Similarly, for

$$(g_1, \pi_1), \ldots, (g_n, \pi_n) \in G^*$$

such that

$$(g_i, \pi_i) \cdot A_i \times \{n\} = B_i \times \{m\},$$

we select $g_1, \ldots, g_n \in G$. Then, by definition,

$$g_i \cdot A_i = B_i$$

for each i. Thus, $E_1 \sim_G E_2$. $\square$

Definition 3.2.2. Let G be a group that acts on X, and let G^*, X^* be defined as in 3.2.1.

(i) A set $A \subseteq X^*$ is said to be bounded if it has finitely many levels.

(ii) If $A \subseteq X^*$ is bounded, the equivalence class of A with respect to G^*-equidecomposability is called the type of A, which is denoted $[A]$.

(iii) If $E \subseteq X$, we write $[E] = [E \times \{0\}]$.

(iv) Let $A, B \subseteq X^*$ be bounded with $k \in \mathbb{N}_0$ such that for

$$B' = \{(b, n + k) \mid (b, n) \in B\},$$

we have $B' \cap A = \emptyset$. Then, $[A] + [B] = [A \sqcup B']$. Note that $[B'] = [B]$.

(v) We define

$$\mathcal{S} = \{[A] \mid A \subseteq X^* \text{ bounded}\}$$

under the addition defined in (iv) to be the type semigroup of the action of G on X.

Fact 3.2.2. Addition is well-defined in $(\mathcal{S}, +)$, and $(\mathcal{S}, +)$ is a well-defined commutative semigroup with identity $[\emptyset]$.

Proof. To show that addition is well-defined, we let $[A_1] = [A_2]$, and $[B_1] = [B_2]$. Without loss of generality, $A_1 \cap B_1 = \emptyset$ and $A_2 \cap B_2 = \emptyset$.

By the definition of the type, $A_1 \sim_{G^*} A_2$ and $B_1 \sim_{G^*} B_2$, meaning

$$A_1 \sqcup B_1 \sim_{G^*} A_2 \sqcup B_2,$$

so

$$\begin{aligned} [A_1] + [B_1] &= [A_1 \sqcup B_1] \\ &= [A_2 \sqcup B_2] \\ &= [A_2] + [A_2], \end{aligned}$$

meaning addition is well-defined.

Since addition is well-defined, and $A \sqcup B = B \sqcup A$, we can see that addition is also commutative. We also have

$$\begin{aligned} [A] + [\emptyset] &= [A \sqcup \emptyset] \\ &= [A], \end{aligned}$$

so $[\emptyset]$ is the identity on $\mathcal{S}$.

Finally, since for any $[A], [B] \in \mathcal{S}$, A and B have finitely many levels, it is the case that $A \cup B$ has finitely many levels for any A and B, so $[A] + [B] \in \mathcal{S}$. ◻

Definition 3.2.3. For any commutative semigroup $\mathcal{S}$ with $\alpha \in S$ and $n \in \mathbb{N}$, we define

$$n\alpha = \underbrace{\alpha + \cdots + \alpha}_{n \text{ times}}$$

Definition 3.2.4. For $\alpha, \beta \in \mathcal{S}$, if there exists $\gamma \in \mathcal{S}$ such that $\alpha + \gamma = \beta$, we write $\alpha \leqslant \beta$.

Fact 3.2.3. If G is a group acting on X with corresponding type semigroup $\mathcal{S}$, then the following are true.

(i) If $\alpha, \beta \in \mathcal{S}$ with $\alpha \leqslant \beta$ and $\beta \leqslant \alpha$, then $\alpha = \beta$.

(ii) A set $E \subseteq X$ is G-paradoxical if and only if $[E] = 2[E]$.

Proof. Let G act on X, and let $\mathcal{S}$ be the corresponding type semigroup.

(i) If $[A] \leqslant [B]$, then there exists $C_1 \in \mathcal{S}$ such that $[A] + [C_1] = [B]$. Without loss of generality, $C_1 \cap A = \emptyset$, meaning $[B] = [A \sqcup C_1]$. Thus, $A \sqcup C_1 \sim_{G^*} B$, meaning $B \preceq_{G^*} A$.

Similarly, if $[B] \leqslant [A]$, then $B \preceq_{G^*} A$. By Theorem 2.4.1, it is thus the case that $A \sim_{G^*} B$.

(ii) Let E be G-paradoxical.

Then, $E \sim_G \bigsqcup_{i=1}^{n} A_i$ and $E \sim_G \bigsqcup_{j=1}^{m} B_j$ for pairwise disjoint subsets $A_1, \dots, A_n, B_1, \dots, B_m \subset E$. Thus, we have

$$
\begin{aligned}
[E] &= \left[\left(\bigsqcup_{i=1}^{n} A_i\right) \sqcup \left(\bigsqcup_{j=1}^{m} B_j\right)\right] \\
&= \left[\bigsqcup_{i=1}^{n} A_i\right] + \left[\bigsqcup_{j=1}^{m} B_j\right] \\
&= 2[E].
\end{aligned}
$$

Similarly, if $[E] = 2[E]$, then there exist A and B such that

$$
\begin{aligned}
[E] &= [A] + [B] \\
&= [A \sqcup B],
\end{aligned}
$$

meaning A and B are each G-equidecomposable with E, so E is G-paradoxical.

□

We can now prove the cancellation identity, which we will be useful as we construct our desired finitely additive measure.

Theorem 3.2.1 (Cancellation Identity on $\mathcal{S}$)**.** Let $\mathcal{S}$ be the type semigroup for some group action, and let $\alpha, \beta \in \mathcal{S}$, $n \in \mathbb{N}$ such that $n\alpha = n\beta$. Then, $\alpha = \beta$.

Proof. Let $n\alpha = n\beta$. Then, there are two disjoint bounded subsets $E, E' \subseteq X^*$ with $E \sim_{G^*} E'$, and pairwise disjoint subsets $A_1, \dots, A_n \subseteq E$, $B_1, \dots, B_n \subseteq E'$ such that

- $E = A_1 \cup \cdots \cup A_n$, $E' = B_1 \cup \cdots \cup B_n$
- $\left[A_j\right] = \alpha$ and $\left[B_j\right] = \beta$ for each $j = 1, \dots, n$.

Let $\chi \colon E \to E'$ be a bijection as in Fact 2.3.3, with $\phi_j : A_1 \to A_j$, $\psi_j : B_1 \to B_j$ also being bijections as in Fact 2.3.3; here we define ϕ_1 and ψ_1 to be the identity map.

For each $a \in A_1$ and $b \in B_1$, we define

$$
\begin{aligned}
\overline{a} &= \{a, \phi_2(a), \dots, \phi_n(a)\} \\
\overline{b} &= \{b, \psi_2(b), \dots, \psi_n(b)\}.
\end{aligned}
$$

We construct a graph by letting $X = \{\overline{a} \mid a \in A_1\}$ and $Y = \left\{\overline{b} \,\middle|\, b \in B_1\right\}$, and, for each j, define edges $\left\{\overline{a}, \overline{b}\right\}$ if $\chi\big(\phi_j(a)\big) \in \overline{b}$.

Since χ is a bijection, for each $j = 1, \dots, n$, $\chi\big(\phi_j(a)\big)$ must be an element of B_k for some k, and since $\{B_k\}_{k=1}^{n}$ are disjoint, $\chi\big(\phi_j(a)\big)$ is an element of exactly one B_k. Thus, the graph is n-regular.

By Theorem 3.1.3, this graph has a perfect matching F. As a result, for each $\overline{a} \in X$, there is a unique $\overline{b} \in Y$ and a unique edge $\left\{\overline{a}, \overline{b}\right\} \in F$ such that $\chi\big(\phi_j(a)\big) = \psi_k(b)$ for some $j, k \in \{1, \dots, n\}$.

We define

$$
\begin{aligned}
C_{j,k} &= \left\{a \in A_1 \,\middle|\, \left\{\overline{a}, \overline{b}\right\} \in F,\ \chi\big(\phi_j(a)\big) = \psi_k(b)\right\} \\
D_{j,k} &= \left\{b \in B_1 \,\middle|\, \left\{\overline{a}, \overline{b}\right\} \in F,\ \chi\big(\phi_j(a)\big) = \psi_k(b)\right\}.
\end{aligned}
$$

Therefore, we must have $\psi_k^{-1} \circ \chi \circ \phi_j$ is a bijection from $C_{j,k}$ to $D_{j,k}$, so $C_{j,k} \sim_{G^*} D_{j,k}$.

Since $C_{j,k}$ and $D_{j,k}$ are partitions of A_1 and B_1 respectively, it follows that $A_1 \sim_{G^*} B_1$, so $\alpha = \beta$. □

Corollary 3.2.1. Let $\mathcal{S}$ be the type semigroup of some group action, and let $\alpha \in \mathcal{S}$ and $n \in \mathbb{N}$ such that $(n+1)\alpha \leqslant n\alpha$. Then, $\alpha = 2\alpha$.

Proof. We have

$$\begin{aligned} 2\alpha + n\alpha &= (n+1)\alpha + \alpha \\ &\leqslant n\alpha + \alpha \\ &= (n+1)\alpha \\ &\leqslant n\alpha. \end{aligned}$$

Inductively repeating this argument, we get $n\alpha \geqslant 2n\alpha$; since $n\alpha \leqslant 2n\alpha$ by definition, we must have $n\alpha = 2n\alpha$, so $\alpha = 2\alpha$. □

Remark 3.2.1. We will call such an α a paradoxical element.

3.3 Two Results on Commutative Semigroups

Now that we are aware of paradoxical elements and the relationship between G-paradoxicality and the properties of the particular elements of the type semigroup (Fact 3.2.3), we will now relate these properties to finitely additive measures of sets by using the following lemma and theorem.

Lemma 3.3.1. Let $\mathcal{S}$ be a commutative semigroup, with $\mathcal{S}_0 \subseteq \mathcal{S}$ finite, and $\epsilon \in \mathcal{S}_0$ satisfying the following assumptions:

(a) $(n+1)\epsilon \nleqslant n\epsilon$ for all $n \in \mathbb{N}$ (i.e., that ϵ is non-paradoxical);

(b) for each $\alpha \in \mathcal{S}$, there is $n \in \mathbb{N}$ such that $\alpha \leqslant n\epsilon$.

Then, there is a set function $\nu\colon \mathcal{S}_0 \to [0,\infty]$ that satisfies the following conditions:

(i) $\nu(\epsilon) = 1$;

(ii) for $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in \mathcal{S}_0$ with $\alpha_1 + \cdots + \alpha_n \leqslant \beta_1 + \cdots \beta_m$,

$$\sum_{j=1}^{n} \nu(\alpha_j) \leqslant \sum_{j=1}^{m} \nu(\beta_j).$$

Proof. We will prove this result by inducting on the cardinality of $\mathcal{S}_0$.

We start with $|\mathcal{S}_0| = 1$. In that case, we define $\nu(\epsilon) = 1$, satisfying condition (i). To satisfy condition (ii), we see that for $n, m \in \mathbb{N}$ with $n\epsilon \leqslant m\epsilon$, if $n \geqslant m+1$, then $(m+1)\epsilon \leqslant n\epsilon \leqslant m\epsilon$, implying that $\epsilon = 2\epsilon$, which contradicts assumption (a).

Let $\alpha_0 \in \mathcal{S}_0 \setminus \{\epsilon\}$. The induction hypothesis says there is a set function satisfying conditions (i) and (ii), $\nu\colon \mathcal{S}_0 \setminus \{\alpha_0\} \to [0,\infty]$.

For $r \in \mathbb{N}$, there are $\gamma_1, \dots, \gamma_p, \delta_1, \dots, \delta_q \in \mathcal{S} \setminus \{\alpha_0\}$ such that

$$\delta_1 + \cdots + \delta_q + r\alpha_0 \leqslant \gamma_1 + \cdots + \gamma_p. \tag{†}$$

Consider the set N defined as follows:

$$N = \left\{ \frac{1}{r}\left(\sum_{j=1}^{p} \nu(\gamma_j) - \sum_{j=1}^{q} \nu(\delta_j) \right) \,\middle|\, \gamma_j, \delta_j \text{ satisfy } (\dagger) \right\}. \tag{‡}$$

We define the extension of ν as follows:

$$\nu(\alpha_0) = \inf N.$$

This infimum is well-defined since, by assumption (b), there is some $n \in \mathbb{N}$ such that $\alpha_0 \leqslant n\epsilon$, and $\nu(\epsilon)$ is defined.

Now, we must show that this extension of ν satisfies condition (ii).

Let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \in \mathcal{S}_0 \setminus \{\alpha_0\}$ and $s, t \in \mathbb{N}_0$ such that

$$\alpha_1 + \cdots + \alpha_n + s\alpha_0 \leqslant \beta_1 + \cdots + \beta_m + t\alpha_0. \tag{$*$}$$

We will verify condition (ii) in the three following cases.

Case 0: If $s = t = 0$, then the induction hypothesis states that $(*)$ satisfies condition (ii).

Case 1: Let $s = 0$ and $t > 0$. We want to show that

$$\sum_{j=1}^{n} \nu(\alpha_j) \leqslant t\nu(\alpha_0) + \sum_{j=1}^{m} \nu(\beta_j),$$

which implies that

$$\nu(\alpha_0) \geqslant \frac{1}{t}\left(\sum_{j=1}^{n} \nu(\alpha_j) - \sum_{j=1}^{m} \nu(\beta_j) \right).$$

By the definition of infimum, it suffices to show that for $r \in \mathbb{N}$ and $\delta_1, \ldots, \delta_q, \gamma_1, \ldots, \gamma_p \in \mathcal{S} \setminus \{\alpha_0\}$ satisfying $(\dagger)$, it is the case that

$$\frac{1}{r}\left(\sum_{j=1}^{p} \nu(\gamma_j) - \sum_{j=1}^{q} \nu(\delta_j) \right) \geqslant \frac{1}{t}\left(\sum_{j=1}^{n} \nu(\alpha_j) - \sum_{j=1}^{m} \nu(\beta_j) \right).$$

Multiplying $(*)$ by r on both sides, and adding $t\delta_1 + \cdots + t\delta_q$ to both sides, we have

$$r\alpha_1 + \cdots + r\alpha_n + t\delta_1 + \cdots + t\delta_q \leqslant r\beta_1 + \cdots + r\beta_m + t(r\alpha_0) + t\delta_1 + \cdots + t\delta_q.$$

Substituting $(\dagger)$, we find

$$r\alpha_1 + \cdots + r\alpha_n + t\delta_1 + \cdots + t\delta_q \leqslant r\beta_1 + \cdots + r\beta_m + t\gamma_1 + \cdots + t\gamma_p.$$

Applying the induction hypothesis, we have

$$r\sum_{j=1}^{n} \nu(\alpha_j) + t\sum_{j=1}^{q} \nu(\delta_j) \leqslant r\sum_{j=1}^{m} \nu(\beta_j) + t\sum_{j=1}^{p} \nu(\gamma_j),$$

yielding

$$\frac{1}{r}\left(\sum_{j=1}^{p} \nu(\gamma_j) - \sum_{j=1}^{q} \nu(\delta_j) \right) \geqslant \frac{1}{t}\left(\sum_{j=1}^{n} \nu(\alpha_j) - \sum_{j=1}^{m} \nu(\beta_j) \right).$$

Case 2: Let $s > 0$. For $z_1, \dots, z_t \in \mathbb{N}$ (‡), we need to show that

$$s\nu(\alpha_0) + \sum_{j=1}^{n} \nu(\alpha_j) \leqslant z_1 + \cdots + z_t + \sum_{j=1}^{n} \nu(\beta_j).$$

Without loss of generality, we can set $z_1, \dots, z_n = z$, as for each $z \in \mathbb{N}$, $z \geqslant \nu(\alpha_0)$.

As in Case 1, we multiply (∗) by r, add $t\delta_1 + \cdots + t\delta_q$ to both sides, and substitute with (†), yielding

$$\begin{aligned} r\alpha_1 + \cdots + r\alpha_n + rs\alpha_0 + t\delta_1 + \cdots + t\delta_q &\leqslant r\beta_1 + \cdots + r\beta_m + t(r\alpha_0) + t\delta_1 + \cdots + t\delta_q \\ r\alpha_1 + \cdots + r\alpha_n + t\delta_1 + \cdots + t\delta_q + rs\alpha_0 &\leqslant r\beta_1 + \cdots + r\beta_m + t\gamma_1 + \cdots + t\gamma_p. \end{aligned}$$

Defining

$$z = \frac{1}{r}\left(\sum_{j=1}^{p} \nu(\gamma_j) - \sum_{j=1}^{q} \nu(\delta_j)\right),$$

we get

$$\begin{aligned} s\nu(\alpha_0) + \sum_{j=1}^{n} \nu(\alpha_j) &\leqslant \sum_{j=1}^{n} \nu(\alpha_j) + \frac{s}{sr}\left(r\sum_{j=1}^{m} \nu(\beta_j) - r\sum_{j=1}^{n} \nu(\alpha_j) + t\sum_{j=1}^{p} \nu(\gamma_j) - t\sum_{j=1}^{q} \nu(\delta_j)\right) \\ &= tz + \sum_{j=1}^{m} \nu(\beta_j). \end{aligned}$$

Thus, we have shown that ν extends in a manner that satisfies conditions (i) and (ii). □

We can "upgrade" our finitely additive set function to a semigroup homomorphism as follows.

Theorem 3.3.1. Let $(\mathcal{S}, +)$ be a commutative semigroup with identity element 0, and let $\epsilon \in \mathcal{S}$. Then, the following are equivalent:

(i) $(n+1)\epsilon \leqslant n\epsilon$ for all $n \in \mathbb{N}$;

(ii) there is a semigroup homomorphism $\nu\colon (\mathcal{S}, +) \to ([0,\infty], +)$ such that $\nu(\epsilon) = 1$.

Proof. To show that (ii) implies (i), we let $\nu\colon (\mathcal{S}, +) \to ([0,\infty], +)$ be a semigroup homomorphism with $\nu(\epsilon) = 1$. Then,

$$\begin{aligned} \nu((n+1)\epsilon) &= (n+1)\nu(\epsilon) \\ &= n+1 \\ &> n \\ &= n\nu(\epsilon) \\ &= \nu(n\epsilon), \end{aligned}$$

meaning that $(n+1)\epsilon \nleqslant n\epsilon$.

To show that (i) implies (ii), we suppose that for each $\alpha \in \mathcal{S}$, there is $n \in \mathbb{N}$ such that $\alpha \leqslant n\epsilon$ — for any such α for which this is not the case, we define $\nu(\alpha) = \infty$.

For a finite subset $\mathcal{S}_0 \subseteq \mathcal{S}$ with $\epsilon \in \mathcal{S}_0$, we define $M_{\mathcal{S}_0}$ to be the set of all $\kappa\colon \mathcal{S} \to [0,\infty]$ such that

- $\kappa(\epsilon) = 1$;
- $\kappa(\alpha + \beta) = \kappa(\alpha) + \kappa(\beta)$ for $\alpha, \beta, \alpha + \beta \in \mathcal{S}_0$.

Since we assume condition (i), we know that such a κ with $\kappa(\epsilon) = 1$ exists. Additionally, since

$$\alpha + \beta \leqslant (\alpha + \beta)$$

and

$$(\alpha + \beta) \leqslant \alpha + \beta,$$

it is the case that

$$\kappa(\alpha + \beta) \leqslant \kappa(\alpha) + \kappa(\beta) \leqslant \kappa(\alpha + \beta),$$

meaning $\kappa(\alpha + \beta) = \kappa(\alpha) + \kappa(\beta)$. Thus, $M_{\mathcal{S}_0}$ is nonempty. It is also the case that $M_{\mathcal{S}_0}$ is closed, since any net of functions $\kappa_p : \mathcal{S} \to [0, \infty]$ with $\kappa_p(\epsilon) = 1$ and $\kappa_p(\alpha + \beta) = \kappa_p(\alpha) + \kappa_p(\beta)$ will necessarily satisfy these conditions in the limit.

We let $[0, \infty]^{\mathcal{S}} = \{\kappa \mid \kappa : \mathcal{S} \to [0, \infty]\}$ be equipped with the product topology. By Tychonoff's theorem, $[0, \infty]^{\mathcal{S}}$ is compact.

Since, for any $\mathcal{S}_1, \dots, \mathcal{S}_n$ finite, it is the case that

$$M_{\mathcal{S}_1 \cup \dots \cup \mathcal{S}_n} \subseteq M_{\mathcal{S}_1} \cap \dots \cap M_{\mathcal{S}_n},$$

since any such $\kappa \in M_{\mathcal{S}_1 \cup \dots \cup \mathcal{S}_n}$ must necessarily be in every $M_{\mathcal{S}_i}$. Thus, the family

$$\{M_{\mathcal{S}_0} \mid \mathcal{S}_0 \subseteq \mathcal{S} \text{ finite}\}$$

has the finite intersection property. Thus, by compactness, there is some ν such that

$$\nu \in \bigcap \{M_{\mathcal{S}_0} \mid \mathcal{S}_0 \subseteq \mathcal{S} \text{ finite}\},$$

with $\nu(\epsilon) = 1$ and, for all $\alpha, \beta \in \mathcal{S}$, since $\nu \in M_{\{\alpha, \beta, \alpha+\beta\}}$, $\nu(\alpha + \beta) = \nu(\alpha) + \nu(\beta)$. □

3.4 Proof of Tarski's Theorem

Finally, we are able to prove the reverse direction of Tarski's Theorem. We restate the theorem before giving its proof.

Theorem 3.0.1 (Tarski's Theorem)**.** Let G be a group that acts on a set X, and let $E \subseteq X$ be nonempty. There is a finitely additive translation-invariant measure $\mu\colon P(X) \to [0, \infty]$ with $\mu(E) \in (0, \infty)$ if and only if E is not G-paradoxical.

Proof of the Reverse Direction of Theorem 3.0.1: Let $\mathcal{S}$ be the type semigroup of the action of G on X.

Suppose E is not G-paradoxical. Then, $[E] \neq 2[E]$ by Fact 3.2.3, meaning $(n + 1)[E] \nleqslant n[E]$ for all $n \in \mathbb{N}$ by the contrapositive of Corollary 3.2.1.

Thus, by Theorem 3.3.1, there is a map $\nu\colon \mathcal{S} \to [0, \infty]$ with $\nu([E]) = 1$. The map $\mu\colon P(X) \to [0, \infty]$ defined by

$$\mu(A) = \nu([A])$$

is the desired finitely additive measure. □

Chapter 4

Amenability through Invariant States

The whole is greater than the sum of its parts.

Aristotle, who had yet to learn about amenability in groups.

In Chapter 4, we introduced amenability through Tarski's Theorem (Theorem 3.0.1), where we proved the existence of a translation-invariant finitely additive set function, $\mu\colon P(X) \to [0,\infty]$, and the non-paradoxicality of G's action on X. There, we used the type semigroup of G's action on X in order to prove the theorem. Now, we turn our attention towards other constructions from analysis — as well as the intrinsic properties of the group itself — to understand ahd prove other criteria for G's amenability.

In this section, we will use techniques from functional analysis to prove the equivalence between amenability and the existence of an invariant state $\mu\colon \ell_\infty(G) \to [0,1]$.

4.1 Amenability in Subgroups and Quotient Groups

We begin by defining a mean on G — note that this definition is slightly different from the one used in the proof in Theorem 3.0.1. However, one can show that they are equivalent by letting G act on itself by left-multiplication and taking $E = G$.

Definition 4.1.1. Let G be a group, with $P(G)$ denoting its power set.

An invariant mean on G is a set function $m\colon P(G) \to [0,1]$ which satisfies, for all $t \in G$ and $E, F \subseteq G$,

- $m(G) = 1$;
- $m(E \sqcup F) = m(E) + m(F)$;
- $m(tE) = m(E)$.

We say G is amenable if G admits a mean.

The mean m is a translation-invariant probability measure on the measurable space $(G, P(G))$.

We can establish some inheritance properties using the properties of a mean. In Proposition 4.1.1, we will show that subgroups of amenable groups are amenable and quotients of amenable groups are amenable.

Proposition 4.1.1. Let G be an amenable group with $H \leqslant G$. Then, the following are true:

(1) H is amenable;

(2) for $H \trianglelefteq G$, G/H is amenable.

Proof.

(1) Let R be a right transversal for H, wherein we select one element of each right coset of H to make up R.

If m is a mean for G, we set $\lambda \colon P(H) \to [0,1]$ defined by

$$\lambda(E) = m(ER).$$

We have

$$\begin{aligned}\lambda(H) &= m(HR)\\ &= m(G)\\ &= 1.\end{aligned}$$

We claim that if $E \cap F = \emptyset$, then $ER \cap FR = \emptyset$. Suppose toward contradiction this is not the case. Then, $xr_1 = yr_2$ for some $x \in E$, $y \in F$, and $r_1, r_2 \in R$. Then, we must have $r_2r_1^{-1} = y^{-1}x \in H$, meaning $r_1 = r_2$ as, by definition, R contains exactly one element of each right coset. Thus, $x = y$, so $E \cap F \neq \emptyset$.

We then have

$$\begin{aligned}\lambda(E \sqcup F) &= m((E \sqcup F)R)\\ &= m(ER \sqcup FR)\\ &= m(ER) + m(FR)\\ &= \lambda(E) + \lambda(F),\end{aligned}$$

and

$$\begin{aligned}\lambda(sE) &= m(sER)\\ &= m(ER)\\ &= \lambda(E).\end{aligned}$$

(2) Let $\pi \colon G \to G/H$ be the canonical projection, defined by $\pi(t) = tH$. We define

$$\lambda \colon P(G/H) \to [0,1]$$

by $\lambda(E) = m\big(\pi^{-1}(E)\big)$. We have

$$\begin{aligned}\lambda(G/H) &= m\Big(\pi^{-1}(G/H)\Big)\\ &= m(G)\\ &= 1,\end{aligned}$$

and

$$\begin{aligned}\lambda(E \sqcup F) &= m\Big(\pi^{-1}(E \sqcup F)\Big)\\ &= m\Big(\pi^{-1}(E) \sqcup \pi^{-1}(F)\Big)\\ &= m\Big(\pi^{-1}(E)\Big) + m\Big(\pi^{-1}(F)\Big)\\ &= \lambda(E) + \lambda(F).\end{aligned}$$

To show translation-invariance, we let $sH = \pi(s) \in G/H$, and $E \subseteq G/H$. Note that

$$\pi^{-1}(\pi(s)E) = s\pi^{-1}(E),$$

since for $r \in s\pi^{-1}(E)$, we have $r = st$ for $t \in \pi(E)$, so $\pi(r) = \pi(st) = \pi(s)\pi(t) \in \pi(s)E$.

Additionally, if $r \in \pi^{-1}(\pi(s)E)$, we have $\pi(r) \in \pi(s)E$, so $\pi\left(s^{-1}r\right) \in E$, meaning $s^{-1}r \in \pi^{-1}(E)$.

Thus,

$$\begin{aligned}
\lambda(\pi(s)E) &= m\left(\pi^{-1}(\pi(s)E)\right) \\
&= m\left(s\pi^{-1}(E)\right) \\
&= m\left(\pi^{-1}(E)\right) \\
&= \lambda(E).
\end{aligned}$$

□

4.2 Establishing Amenability through Functional Analysis

Now that we understand some useful properties of means in relation to groups and subgroups, we turn our attention toward finding means on groups. In order to do this, we turn our attention towards the space $\ell_\infty(G)$, which allows us to use theories from functional analysis to better understand means on G.

Definition 4.2.1. Let G be a group.

(1) The space $\mathcal{F}(G)$ is defined by

$$\mathcal{F}(G) = \{f \mid f\colon G \to \mathbb{C} \text{ is a function}\}.$$

(2) A function $f \in \mathcal{F}(G)$ is called positive if $f(x) \geqslant 0$ for all $x \in G$.

(3) A function $f \in \mathcal{F}(G)$ is called simple if $\operatorname{Ran}(f)$ is finite. We let

$$\Sigma = \left\{f \in \mathcal{F}(G) \;\middle|\; f \text{ is simple}\right\}.$$

Fact 4.2.1. It is the case that $\Sigma \subseteq \mathcal{F}(G)$ is a linear subspace.

Definition 4.2.2. For $E \subseteq G$, we define

$$\mathbb{1}_E\colon G \to \mathbb{C}$$

by

$$\mathbb{1}_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}.$$

This is the characteristic function of E.

Fact 4.2.2. We have

$$\operatorname{span}\{\mathbb{1}_E \mid E \subseteq G\} = \Sigma.$$

Proof. We see that $\mathbb{1}_E \in \Sigma$ for any $E \subseteq G$, and that Σ is a subspace.

If $\phi \in \Sigma$ with $\operatorname{Ran}(\phi) = \{t_1, \dots, t_n\}$, where t_i are distinct, we set

$$E_i = \phi^{-1}(\{t_i\}),$$

yielding

$$\phi = \sum_{i=1}^{n} t_i \mathbb{1}_{E_i}.$$

□

Definition 4.2.3.

(1) A function $f \in \mathcal{F}(G)$ is bounded if there exists $M > 0$ such that $|f(g)| \leqslant M$ for all $g \in G$.

(2) The space $\ell_\infty(G)$ is defined by

$$\ell_\infty(G) = \{f \in \mathcal{F}(G) \mid f \text{ is bounded}\}.$$

(3) The norm on $\ell_\infty(G)$ is defined by

$$\|f\|_{\ell_\infty} = \sup_{x \in G} |f(x)|.$$

Proposition 4.2.1. The space $\ell_\infty(G)$ is complete. Additionally, $\overline{\Sigma} = \ell_\infty(G)$.

Proof. Let $(f_n)_n$ be $\|\cdot\|$-Cauchy in $\ell_\infty(G)$. Then, for all $x \in G$, it is the case that

$$\begin{aligned} |f_n(x) - f_m(x)| &= |(f_n - f_m)(x)| \\ &\leqslant \|f_n - f_m\|_{\ell_\infty}, \end{aligned}$$

meaning $(f_n(x))_n$ is Cauchy in $\mathbb{C}$. We define $f(x) = \lim_{n\to\infty} f_n(x)$. We must show that $f \in \ell_\infty(G)$, and $\|f_n - f\|_{\ell_\infty} \to 0$.

We have

$$\begin{aligned} |f(x)| &= \left|\lim_{n\to\infty} f_n(x)\right| \\ &= \lim_{n\to\infty} |f_n(x)| \\ &\leqslant \limsup_{n\to\infty} \|f_n\|_{\ell_\infty} \\ &\leqslant C, \end{aligned}$$

as Cauchy sequences are always bounded. Thus, $\sup_{x\in G} |f(x)| \leqslant C$.

Given $\varepsilon > 0$, we find N such that for all $m, n \geqslant N$, $\|f_n - f_m\|_{\ell_\infty} \leqslant \varepsilon$. Thus, for $x \in G$, we have

$$\begin{aligned} |f_n(x) - f_m(x)| &\leqslant \|f_n - f_m\|_{\ell_\infty} \\ &\leqslant \varepsilon. \end{aligned}$$

Taking $m \to \infty$, we get $|f_n(x) - f(x)| \leqslant \varepsilon$, for all $n \geqslant N$, so $\|f_n - f\|_{\ell_\infty} \leqslant \varepsilon$ for all $n \geqslant N$.

For real-valued $f \in \ell_\infty(G)$, let $|f| \subseteq [-M, M]$ for some $M > 0$. Let $\varepsilon > 0$. Since $[-M, M]$ is compact, it is totally bounded, so we can find intervals $I_1, \dots, I_n$ with $[-M, M] = \bigsqcup_{k=1}^{n} I_k$, with the length of each I_k less than ε.

Set $E_k = f^{-1}(I_k)$. Pick some $t_k \in I_k$. We set

$$\phi = \sum_{i=1}^{n} t_k \mathbb{1}_{E_k}.$$

Then, it is the case that $\|\phi - f\|_{\ell_\infty} < \varepsilon$.

If $f \in \ell_\infty(G)$ is complex-valued, we apply this process separately to $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$. □

Corollary 4.2.1. For any $f \in \ell_\infty(G)$, there is a sequence $(\phi_n)_n$ of simple functions with $\|\phi_n - f\|_{\ell_\infty} \to 0$. If $f \geqslant 0$, then we can select $\phi_n \geqslant 0$.

Now that we understand how simple functions relate to $\ell_\infty(G)$, we start by defining a translation action on $\ell_\infty(G)$, from which we will be able to convert the idea of means into invariant elements of the state space of the dual of $\ell_\infty(G)$.

Proposition 4.2.2. Let G be a group. There is an action

$$\lambda\colon G \to \mathrm{Isom}(\ell_\infty(G)),$$

where $\lambda(s) = \lambda_s$, defined by

$$\lambda_s(f)(t) = f\left(s^{-1}t\right)$$

Proof. We have

$$\begin{aligned}
\lambda_s(f + \alpha g)(t) &= (f + \alpha g)\left(s^{-1}t\right) \\
&= f\left(s^{-1}t\right)\alpha g\left(s^{-1}t\right) \\
&= \lambda_s(f)(t) + \alpha\lambda_s(g)(t) \\
&= (\lambda_s(f) + \alpha\lambda_s(g))(t).
\end{aligned}$$

Thus, λ_s is linear. Additionally,

$$\begin{aligned}
\|\lambda_s(f)\|_{\ell_\infty} &= \sup_{t\in G}|\lambda_s(f)(t)| \\
&= \sup_{t\in G}\left|f\left(s^{-1}t\right)\right| \\
&= \|f\|_{\ell_\infty},
\end{aligned}$$

and

$$\begin{aligned}
\|\lambda_s(f) - \lambda_s(g)\|_{\ell_\infty} &= \|\lambda_s(f - g)\| \\
&= \|f - g\|_{\ell_\infty},
\end{aligned}$$

meaning λ_s is an isometry.

We have

$$\begin{aligned}
\lambda_s \circ \lambda_r(f)(t) &= \lambda_r(f)\left(s^{-1}t\right) \\
&= \lambda_r\left(r^{-1}s^{-1}t\right) \\
&= f\left((sr)^{-1}t\right) \\
&= \lambda_{sr}(f)(t),
\end{aligned}$$

establishing that $\lambda_s \circ \lambda_r = \lambda_{sr}$.

By a similar process, we find that $\lambda_s(\mathbb{1}_E) = \mathbb{1}_{sE}$ for any $E \subseteq G$ and $s \in G$. □

Definition 4.2.4. A state on $\ell_\infty(G)$ is a continuous linear functional $\mu \in \ell_\infty(G)^*$ such that the following are true:

- μ is positive;
- $\mu(\mathbb{1}_G) = 1$.

A state is called left-invariant if

$$\mu(\lambda_s(f)) = \mu(f).$$

Example 4.2.1. The evaluation functional, $\delta_x : \ell_\infty \to \mathbb{R}$, defined by

$$\delta_x(f) = f(x),$$

is a state. However, it is not necessarily invariant, as

$$\begin{aligned} \delta_x(\lambda_s(f)) &= \lambda_s(f)(x) \\ &= f\left(s^{-1}x\right) \\ &\neq f(x). \end{aligned}$$

However, we can use the evaluation functional to create an invariant state. If G is finite, we define

$$\mu = \frac{1}{|G|} \sum_{x \in G} \delta_x,$$

which is indeed an invariant state.

We can characterize states slightly differently, which will enable us to show the equivalence between invariant states and means.

Lemma 4.2.1.

(1) If μ is a state on $\ell_\infty(G)$, then

$$\|\mu\|_{op} = 1.$$

(2) If $\mu \in \ell_\infty(G)^*$ is such that

$$\begin{aligned} \|\mu\|_{op} &= \mu(\mathbb{1}_G) \\ &= 1, \end{aligned}$$

then μ is positive and a state.

Proof.

(1) Let μ be a state. Given $f \in \ell_\infty(G)$, we have

$$\begin{aligned} \|f\|_{\ell_\infty} \mathbb{1}_G - f \geqslant 0 \\ \|f\|_{\ell_\infty} \mathbb{1}_G + f \geqslant 0, \end{aligned}$$

so

$$0 \leqslant \mu\left(\|f\|_{\ell_\infty} \mathbb{1}_G - f\right)$$

$$= \|f\|_{\ell_\infty}\mu(\mathbb{1}_G) - \mu(f)$$

meaning

$$\mu(f) \leqslant \|f\|_{\ell_\infty}.$$

Additionally,

$$\begin{aligned} 0 &\leqslant \mu\big(\|f\|_{\ell_\infty}\mathbb{1}_G + f\big) \\ &= \|f\|_{\ell_\infty}\mu(\mathbb{1}_G) + \mu(f), \end{aligned}$$

meaning

$$-\mu(f) \leqslant \|f\|_{\ell_\infty}.$$

Thus, we have $|\mu(f)| \leqslant \|f\|_{\ell_\infty}$, so $\|\mu\|_{op} \leqslant 1$. However, since $\mu(\mathbb{1}_G) = 1$, we must have $\|\mu\|_{op} = 1$.

(2) Suppose $\|\mu\|_{op} = \mu(\mathbb{1}_G) = 1$. Let $f \geqslant 0$. Set $g = \frac{1}{\|f\|_{\ell_\infty}} f$.

Then, $\operatorname{Ran}(g) \subseteq [0, 1]$, and $\operatorname{Ran}(g - \mathbb{1}_G) \subseteq [-1, 1]$. Thus, $\|g - \mathbb{1}_G\|_{\ell_\infty} \leqslant 1$.

Since $\|\mu\|_{op} = 1$, we must have

$$\begin{aligned} |\mu(g - \mathbb{1}_G)| &\leqslant 1 \\ |\mu(g) - 1| &\leqslant 1, \end{aligned}$$

and since $\mu(\mathbb{1}_G) = 1$, we have $\mu(g) \in [0, 2]$. Thus, $\mu(f) = \|f\|_{\ell_\infty}\mu(g) \geqslant 0$.

□

Corollary 4.2.2. The set of states in $\ell_\infty(G)^*$ forms a w^*-compact subset of $B_{\ell_\infty(G)^*}$.

Proof. From the Banach–Alaoglu Theorem (Theorem D.4.4), we only need to show that the set of states, $S(\ell_\infty(G))$, is w^*-closed, as every element of $S(\ell_\infty(G))$ has norm 1.

Let $f \in \ell_\infty(G)$ be positive, and let $(\varphi_i)_i$ be a net in $S(\ell_\infty(G))$ with $(\varphi_i)_i \xrightarrow{w^*} \varphi \in \ell_\infty(G)^*$. From Lemma 4.2.1, we must show that φ is positive and $\varphi(\mathbb{1}_G) = 1$.

We start by seeing that, since each φ_i is a state, we have $\varphi_i(f) \geqslant 0$ for each $i \in I$, so we must have $\varphi(f) \geqslant 0$.

Similarly, since $\varphi_i(\mathbb{1}_G) = 1$ for each $i \in I$, and $(\varphi_i)_i \xrightarrow{w^*} \varphi$, we have $\varphi(\mathbb{1}_G) = 1$. Thus, by Lemma 4.2.1, we have that $S(\ell_\infty(G))$ is w^*-closed. □

Now, we may show the correspondence between invariant states and means.

Proposition 4.2.3. If $\mu \in \ell_\infty(G)^*$ is a state, then $m\colon P(G) \to [0, 1]$ defined by $m(E) = \mu(\mathbb{1}_E)$ is a finitely additive probability measure on G.

Moreover, if μ is invariant, then m is a mean.

Proof. We have

$$\begin{aligned} m(G) &= \mu(\mathbb{1}_G) \\ &= 1 \end{aligned}$$

$$\begin{aligned} m(\emptyset) &= \mu(0) \\ &= 0 \end{aligned}$$

$$\begin{aligned} m(E \sqcup F) &= \mu(\mathbb{1}_{E \sqcup F}) \\ &= \mu(\mathbb{1}_E + \mathbb{1}_F) \\ &= \mu(\mathbb{1}_E) + \mu(\mathbb{1}_F) \\ &= m(E) + m(F). \end{aligned}$$

Additionally, since $0 \leqslant \mathbb{1}_E \leqslant \mathbb{1}_G$, we have $0 \leqslant \mu(\mathbb{1}_E) \leqslant 1$, so $0 \leqslant m(E) \leqslant 1$.

If μ is invariant, then

$$\begin{aligned} m(sE) &= \mu(\mathbb{1}_{sE}) \\ &= \mu(\lambda_s(\mathbb{1}_E)) \\ &= \mu(\mathbb{1}_E) \\ &= m(E). \end{aligned}$$

□

Proposition 4.2.4. If G admits a mean, then $\ell_\infty(G)^*$ admits an invariant state.

Proof. Let m be a mean. Define $\mu_0 \colon \Sigma \to \mathbb{R}$ by

$$\mu_0\left(\sum_{k=1}^{n} t_k \mathbb{1}_{E_k}\right) = \sum_{k=1}^{n} t_k m(E_k).$$

Since m is finitely additive, it is the case that μ_0 is well-defined, linear, and positive, with $\mu_0(\mathbb{1}_G) = m(G) = 1$.

Additionally, since m is a mean, then for $f = \sum_{k=1}^{n} t_k \mathbb{1}_{E_k}$, we have

$$\begin{aligned} \mu_0(\lambda_s(f)) &= \mu_0\left(\lambda_s\left(\sum_{k=1}^{n} t_k \mathbb{1}_{E_k}\right)\right) \\ &= \mu_0\left(\sum_{k=1}^{n} t_k \mathbb{1}_{sE_k}\right) \\ &= \sum_{k=1}^{n} t_k m(sE_k) \\ &= \sum_{k=1}^{n} t_k m(E_k) \\ &= \mu_0(f). \end{aligned}$$

We see that

$$\begin{aligned} |\mu_0(f)| &= \left|\sum_{k=1}^{n} t_k m(E_k)\right| \\ &\leqslant \sum_{k=1}^{n} |t_k| m(E_k) \\ &\leqslant \sum_{k=1}^{n} \|f\|_{\ell_\infty} \sum_{k=1}^{n} m(E_k) \end{aligned}$$

$$
\begin{aligned}
&= \|f\|_{\ell_\infty} \sum_{k=1}^{n} m(E_k) \\
&\leqslant \|f\|_{\ell_\infty},
\end{aligned}
$$

meaning μ_0 is continuous, so μ_0 is uniformly continuous.

Since $\overline{\Sigma} = \ell_\infty(G)$, uniform continuity provides that μ_0 extends to a continuous linear functional $\mu\colon \ell_\infty(G) \to \mathbb{R}$ with $\mu(\mathbb{1}_G) = \mu_0(\mathbb{1}_G) = 1$.

For $f \geqslant 0$, we find a sequence $(\phi_n)_n$ in Σ with $\phi_n \geqslant 0$ and $\|\phi_n - f\|_{\ell_\infty} \xrightarrow{n\to\infty} 0$. We set

$$
\begin{aligned}
\mu(f) &= \lim_{n\to\infty} \mu(\phi_n) \\
&= \lim_{n\to\infty} \mu_0(\phi_n) \\
&\geqslant 0,
\end{aligned}
$$

so μ is a state.

If $f \in \ell_\infty(G)$, $s \in G$, and $(\phi_n)_n$ a sequence in Σ with $(\phi_n)_n \to f$, then

$$
\begin{aligned}
\|\lambda_s(\phi_n) - \lambda_s(f)\|_{\ell_\infty} &= \|\lambda_s(\phi_n - f)\|_{\ell_\infty} \\
&= \|\phi_n - f\|_{\ell_\infty} \\
&\to 0.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\mu(\lambda_s(\phi_n)) &= \mu_0(\lambda_s(\phi_n)) \\
&= \mu_0(\phi_n) \\
&= \mu(\phi_n) \\
&\to \mu(f),
\end{aligned}
$$

so $\mu(f) = \mu(\lambda_s(f))$. Thus, $\mu \in \ell_\infty(G)^*$ is an invariant state. □

4.3 Establishing Amenability using Invariant States

Owing to the correspondence between invariant states and means, we are now able to establish amenability for large classes of groups.

Proposition 4.3.1. The group of integers, $\mathbb{Z}$, is amenable.

Proof. We define the left shift, $\lambda_1\colon \ell_\infty(\mathbb{Z}) \to \ell_\infty(\mathbb{Z})$, by

$$
\lambda_1(f)(k) = f(k-1).
$$

This is an action as in Proposition 4.2.2.

We set $Y = \operatorname{Ran}(\mathrm{id} - \lambda_1) \subseteq \ell_\infty(\mathbb{Z})$. We claim that $\operatorname{dist}_Y(\mathbb{1}_\mathbb{Z}) \geqslant 1$.

Suppose toward contradiction that there is $y \in Y$ with $\|\mathbb{1}_\mathbb{Z} - y\|_{\ell_\infty} = r < 1$. Then, $y = f - \lambda_1 f$ for some $f \in \ell_\infty(\mathbb{Z})$, so

$$
\|\mathbb{1}_\mathbb{Z} - (f - \lambda_1(f))\|_{\ell_\infty} = r.
$$

Thus, for all $k \in \mathbb{Z}$, we have

$$|1 - (f(k) - f(k-1))| \leqslant r,$$

so $|f(k) - f(k-1)| \geqslant 1 - r > 0$. However, such an f cannot be bounded.

Since $\operatorname{dist}_{\overline{Y}}(\mathbb{1}_{\mathbb{Z}}) = \operatorname{dist}_Y(\mathbb{1}_{\mathbb{Z}})$, the Hahn–Banach separation theorems provide $\mu \in (\ell_\infty(\mathbb{Z}))^*$ with $\|\mu\|_{op} = 1$, $\mu|_{\overline{Y}} = 0$, and $\mu(\mathbb{1}_{\mathbb{Z}}) = \operatorname{dist}_Y(\mathbb{1}_{\mathbb{Z}}) \geqslant 1$.

Since $\|\mu\|_{op} = 1$ and $\mu(\mathbb{1}_{\mathbb{Z}}) \geqslant 1$, we must have $\mu(\mathbb{1}_{\mathbb{Z}}) = 1$.

Additionally, since $\|\mu\|_{op} = \mu(\mathbb{1}_{\mathbb{Z}}) = 1$, we have that μ is a state on $\ell_\infty(\mathbb{Z})$, and since $\mu(y) = 0$ for all $y \in Y$, we have

$$\begin{aligned} \mu(f - \lambda_1(f)) &= 0 \\ \mu(f) &= \mu(\lambda_1(f)). \end{aligned}$$

Inductively, this means that $\mu(f) = \mu(\lambda_k(f))$ for all $k \in \mathbb{Z}$, so μ is an invariant state on $\ell_\infty(\mathbb{Z})$. Thus, $\mathbb{Z}$ is amenable. □

Proposition 4.3.2. If $N \trianglelefteq G$ and G/N are amenable, then G is amenable.

Proof. Let $\rho \in (\ell_\infty(G/N))^*$ be an invariant state, and let $p\colon P(N) \to [0,1]$ be a mean. For $E \subseteq G$, we define $f_E\colon G/N \to \mathbb{R}$ by

$$f_E(tN) = p\left(N \cap t^{-1}E\right).$$

We start by verifying that f_E is well-defined. For $tN = sN$, we have $s^{-1}t \in N$, so

$$\begin{aligned} p\left(N \cap t^{-1}E\right) &= p\left(s^{-1}t\left(N \cap t^{-1}E\right)\right) \\ &= p\left(s^{-1}tN \cap s^{-1}E\right) \\ &= p\left(N \cap s^{-1}E\right). \end{aligned}$$

Since f_E is defined through p, we can see that f_E is bounded. Additionally,

$$\begin{aligned} f_{E \sqcup F}(tN) &= p\left(N \cap t^{-1}(E \sqcup F)\right) \\ &= p\left(N \cap \left(t^{-1}E \sqcup t^{-1}F\right)\right) \\ &= p\left(\left(N \cap t^{-1}E\right) \sqcup \left(N \cap t^{-1}F\right)\right) \\ &= p\left(N \cap t^{-1}E\right) + p\left(N \cap t^{-1}F\right) \\ &= f_E(tN) + f_F(tN) \\ &= (f_E + f_F)(tN), \end{aligned}$$

and

$$\begin{aligned} f(sE)(tN) &= p\left(N \cap t^{-1}sE\right) \\ &= f_E\left(s^{-1}tN\right) \\ &= \lambda_{sN}(f_E)(tN), \end{aligned}$$

so $f_{sE} = \lambda_{sN}(f_E)$. Finally,

$$\begin{aligned} f_G(tN) &= p\Big(N \cap t^{-1}G\Big) \\ &= p(N) \\ &= 1, \end{aligned}$$

meaning $f_G = \mathbb{1}_{G/N}$.

We define $m\colon P(G) \to [0,1]$ by

$$m(E) = \rho(f_E).$$

Then, we have

$$m(E \sqcup F) = m(E) + m(F)$$

$$m(G) = 1$$

$$\begin{aligned} m(sE) &= \rho(f_{sE}) \\ &= \rho(\lambda_{sN}(f_E)) \\ &= \rho(f_E) \\ &= m(E), \end{aligned}$$

so m is a mean. □

Corollary 4.3.1. The finite direct product of amenable groups is amenable.

Proof. If H and K are amenable, then $K \cong (H \times K)/H$ is amenable and H is amenable, so $H \times K$ is amenable by Proposition 4.3.2. Induction provides the general case. □

Corollary 4.3.2. Finitely generated abelian groups are amenable.

Proof. By the fundamental theorem of finitely generated abelian groups (see [DF04, p. 158]), all finitely generated abelian groups are isomorphic to $\mathbb{Z}^d \times \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$.

Since $\mathbb{Z}^d$ is a finite direct product of $\mathbb{Z}$, and the torsion subgroup $\mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$ is finite, we see that a finitely generated abelian group is a direct product of two amenable groups, hence amenable. □

Corollary 4.3.3. If $\{G_i\}_{i \in I}$ is a directed family of amenable groups, then the direct limit,

$$G = \bigcup_{i \in I} G_i,$$

is also amenable.

Proof. Let $\mu_i \in (\ell_\infty(G_i))^*$ be invariant states.

Set

$$M_i = \{\mu \in S(\ell_\infty(G)) \mid \mu(\lambda_s(f)) = \mu(f) \text{ for all } s \in G_i\}.$$

We set $\mu(f) = \mu_i(f|_{G_i})$. Since each G_i is amenable, it is the case that each M_i is nonempty. Similarly, seeing as we have established the state space as w^*-closed in $B_{\ell_\infty(G)^*}$, it is the case that each M_i is w^*-closed in $B_{\ell_\infty(G)^*}$.

For $i_1, \dots, i_n$, we find $G_j \supseteq G_{i_1}, \dots, G_{i_n}$, which exists since $\{G_i\}_{i \in I}$ is directed. We have that

$M_j \subseteq \bigcap_{k=1}^n M_{i_k}$, so $\{M_i\}_{i\in I}$ has the finite intersection property.

By compactness, there is $\mu \in \bigcap_{i\in I} M_i$ which is necessarily invariant on G. ◻

Corollary 4.3.4. All abelian groups are amenable.

Proof. Every abelian group is the direct limit of its finitely generated subgroups. ◻

Corollary 4.3.5. All solvable groups are amenable.

Proof. Let $e_G = G_0 \leqslant G_1 \leqslant \cdots \leqslant G_n \leqslant G$ be such that $G_{j-1} \trianglelefteq G_j$ for $j = 1, \ldots, n$, and G_i/G_j is abelian.

Since G_0 is abelian, it is amenable, as is G_1/G_0, so G_1 is amenable. We see then that G_2 is amenable as G_1 and G_2/G_1 are amenable.

Continuing in this fashion, we see that G is amenable. ◻

4.4 Remarks and Notes

The following proposition is, in a sense, a kind of converse to Proposition 4.1.1, in that if a subgroup is amenable, we can show that the original group is also amenable, but this is only a sufficient condition if the subgroup has finite index.

Proposition 4.4.1. Let G be a group, and let $H \leqslant G$ be amenable, with $[G : H] = n < \infty$. Then, G is amenable.

Proof. Let $H \leqslant G$ be amenable with $[G : H] = n$. Let μ be the mean on H, and let $\{g_i H\}_{i=1}^n$ be a partition of G by the left cosets of H. We define the mean on G by taking, for $A \subseteq G$,

$$\lambda(A) = \frac{1}{n}\sum_{i=1}^n \mu\left(g_i^{-1}A \cap H\right).$$

We begin by verifying that this is well-defined. Specifically, we will show that this definition is independent of the coset representatives. Suppose $g_j H = h_j H$. Then, $h_j^{-1} g_j \in H$. Now, we have $g_j^{-1}A \cap H \subseteq H$, so by left-multiplication, we get $\left(h_j^{-1}g_j\right)g_j^{-1}A \cap H \subseteq H$, so $h_j^{-1}A \cap H \subseteq H$. Since $\{g_i H\}_{i=1}^n$ is a partition, we get that this definition of the mean on G is independent of the choice of coset representatives.

Next, we show that this is a finitely additive measure. Let $A, B \subseteq G$ be such that $A \cap B = \emptyset$. Then, we get

$$\begin{aligned}
\lambda(A \sqcup B) &= \frac{1}{n}\sum_{i=1}^n \mu\left(g_i^{-1}(A \sqcup B) \cap H\right) \\
&= \frac{1}{n}\sum_{i=1}^n \mu\left(\left(g_i^{-1}A \cap H\right) \sqcup \left(g_i^{-1}B \cap H\right)\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mu\left(g_i^{-1}A \cap H\right) + \sum_{i=1}^n \mu\left(g_i^{-1}B \cap H\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^n \mu\left(g_i^{-1}A \cap H\right) + \frac{1}{n}\sum_{i=1}^n \mu\left(g_i^{-1}B \cap H\right) \\
&= \lambda(A) + \lambda(B).
\end{aligned}$$

It is relatively simple to see that λ is a probability measure, as

$$\lambda(G) = \frac{1}{n}\sum_{i=1}^n \mu\left(g_i^{-1}G \cap H\right)$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} \mu(G \cap H) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu(H) \\
&= 1.
\end{aligned}
$$

Now, we must show that λ is translation-invariant. Let $A \subseteq G$ and $t \in G$. Then, using the translation-invariance of μ, we get

$$
\begin{aligned}
\lambda(tA) &= \frac{1}{n}\sum_{i=1}^{n} \mu\left(g_i^{-1}tA \cap H\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu\left(g_i^{-1}(t(A \cap H))\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu\left(g_i^{-1}A \cap H\right) \\
&= \lambda(A).
\end{aligned}
$$

Thus, G is amenable. □

In [Tit72], Jacques Tits proved that in any subgroup of $GL_n(\mathbb{F})$ (where $\mathbb{F}$ is any field with characteristic zero), then the subgroup either admits a solvable subgroup of finite index (hence amenable by Corollary 4.3.5 and Proposition 4.4.1) or contains a non-abelian freely generated subgroup (which is necessarily not amenable by Theorem 3.0.1). This is also known as the Tits alternative.

Note that we have found a sufficient condition to find a (non-abelian) freely generated subgroup by Theorem 2.1.1, and showed in Theorem 2.1.2 that SO(3) contains a (necessarily) non-abelian freely generated subgroup — in this case, isomorphic to $F(a, b)$ — so we know by the Tits alternative that SO(3) does not admit a solvable subgroup with a finite index. In the introduction to Chapter 3, we stated that the Banach–Tarski paradox cannot hold for $\mathbb{R}$ and $\mathbb{R}^2$, because the rotation group SO(2) in $\mathbb{R}^2$ is abelian, and since the isometry group E(2) has the abelian subgroup SO(2) with finite index, E(2) is amenable by Proposition 4.4.1. Similarly, the isometry group E(1) contains an abelian subgroup SO(1)[I] with finite index.

[I] SO(1) = $\{1\}$.

Chapter 5

Følner's Condition and Approximate Means

While group amenability and the existence of an invariant state on the group are equivalent, as well as showing the group is non-paradoxical, we are interested in establishing a more combinatorial characterization for group amenability. This will eventually involve showing the equivalence between the existence of an invariant state, the Følner condition, and the existence of an approximate mean. We then take a tour of the essence of geometric group theory to apply Følner's condition to to the groups of subexponential growth.

5.1 Følner's Condition

Definition 5.1.1. A group is said to satisfy the Følner condition if, for every $\varepsilon > 0$ and $E \subseteq G$, there is a nonempty finite subset $F \subseteq G$ such that for all $t \in E$,

$$\frac{|tF \triangle F|}{|F|} \leqslant \varepsilon.$$

Equivalently, we can also say that the Følner condition is satisfied if and only if

$$\frac{|tF \cap F|}{|F|} \geqslant 1 - \varepsilon$$

for every $\varepsilon > 0$.

Lemma 5.1.1. A countable group G satisfies the Følner condition if and only if G admits a sequence $(F_n)_n$ with $F_n \subseteq G$ finite such that

$$\left(\frac{|tF_n \triangle F_n|}{|F_n|}\right)_n \xrightarrow{n\to\infty} 0$$

for all $t \in G$. Such a sequence is known as a Følner sequence.

Proof. Let G admit a Følner sequence, $(F_n)_n$. Given $\varepsilon > 0$ and $E \subseteq G$ finite, find N such that for all $s \in E$ and $n \geqslant N$,

$$\frac{|sF_n \triangle F_n|}{|F_n|} \leqslant \varepsilon.$$

We take $F = F_N$ in the definition of the Følner condition.

Let G satisfy the Følner condition. We write $G = \bigcup_{n\geqslant 1} E_n$, with $E_1 \subseteq E_2 \subseteq \cdots$, and define F_n such that for all $t \in E_n$,

$$\frac{|tF_n \triangle F_n|}{|F_n|} \leqslant \frac{1}{n}.$$

Given $t \in G$, then $t \in E_N$ for some N, so $t \in E_n$ For all $n \geqslant N$, so

$$\frac{|tF_n \triangle F_n|}{|F_n|} \leqslant \frac{1}{n}$$

for all $n \geqslant N$. Thus,

$$\left(\frac{|tF_n \triangle F_n|}{|F_n|}\right) \xrightarrow{n\to\infty} 0.$$

□

Lemma 5.1.2. Let G be a finitely generated group with generating set S (see Definition 1.1.1). If $(F_n)_n$ is a sequence of finite subsets such that, for all $s \in S$,

$$\left(\frac{|sF_n \triangle F_n|}{|F_n|}\right)_n \to 0,$$

then $(F_n)_n$ is a Følner sequence for G.

Proof. Note that

- $s(A\triangle B) = sA\triangle sB$;
- $A\triangle C \subseteq (A\triangle B) \cup (B\triangle C)$.

We see that for any $s \in S$,

$$\begin{aligned}\frac{\left|s^{-1}F_n \triangle F_n\right|}{|F_n|} &= \frac{\left|s^{-1}(F_n \triangle sF_n)\right|}{|F_n|}\\ &= \frac{|F_n \triangle sF_n|}{|F_n|}\\ &\to 0.\end{aligned}$$

Thus, we may assume that S is symmetric — i.e., that $\left\{s^{-1} \mid s \in S\right\} = \{s \mid s \in S\}$.

For any $s, t \in S$, we have

$$\begin{aligned}\frac{|stF_n \triangle F_n|}{|F_n|} &\leqslant \frac{|stF_n \triangle F_n|}{|F_n|} + \frac{|sF_n \triangle F_n|}{|F_n|}\\ &= \frac{|s(tF_n \triangle F_n)|}{|F_n|} + \frac{|sF_n \triangle F_n|}{|F_n|}\\ &= \frac{|tF_n \triangle F_n|}{|F_n|} + \frac{|sF_n \triangle F_n|}{|F_n|}\\ &\to 0.\end{aligned}$$

We use induction to find the general case. □

Example 5.1.1. Consider the group $\mathbb{Z}$. Since $\mathbb{Z}$ is generated by the element $\{1\}$, we see that for F_n =

$[-n, n]$, that

$$\begin{aligned}\frac{|(F_n + 1)\triangle F_n|}{|F_n|} &= \frac{2}{2n+1} \\ &\to 0,\end{aligned}$$

meaning that $\mathbb{Z}$ satisfies the Følner condition.

5.2 From Følner's Condition to Amenability: Enter Approximate Means

We have thus far proven that G satisfies the Følner condition if and only if G admits a Følner sequence, and that G is amenable if and only if G admits an invariant state.

We will now begin harmonizing these two characterizations through the use of approximate means, eventually showing that G satisfies the Følner condition if and only if G admits an approximate mean, and that G admits an approximate mean if and only if G is amenable.

Definition 5.2.1. For a group G, we define

$$\operatorname{Prob}(G) = \left\{ f\colon G \to [0,\infty) \;\middle|\; \operatorname{card}\big(\operatorname{supp}(f)\big) < \infty,\ \sum_{t \in G} f(t) = 1 \right\}.$$

Note that $\operatorname{Prob}(G) \subseteq B_{\ell_1(G)}$. For $f \in \operatorname{Prob}(G)$, we set $\varphi_f\colon \ell_\infty(G) \to \mathbb{C}$ defined by

$$\varphi_f(g) = \sum_{t \in G} g(t) f(t).$$

Fact 5.2.1. For $f \in \operatorname{Prob}(G)$, φ_f is a state on $\ell_\infty(G)$.

Proof. We can see that, by definition, φ_f is positive, linear, and has $\varphi_f(\mathbb{1}_G) = 1$.

We only need to show that $\|\varphi_f\| = 1$. We see that

$$\begin{aligned}|\varphi_f(g)| &= \left|\sum_{t \in G} g(t) f(t)\right| \\ &\leqslant \sum_{t \in G} |g(t)||f(t)| \\ &\leqslant \|g\|_\infty \sum_{t \in G} |f(t)| \\ &= \|g\|_\infty,\end{aligned}$$

so $\|\varphi_f\| \leqslant 1$. Since $\varphi_f(\mathbb{1}_G) = 1$, we must have $\|\varphi_f\| = 1$. □

Proposition 5.2.1. There is an action $\lambda\colon G \to \operatorname{Isom}(\ell_1(G))$ such that $\operatorname{Prob}(G)$ is invariant.

Proof. Let $\lambda_s(f)(t) = f\big(s^{-1}t\big)$. Then,

$$\begin{aligned}\|\lambda_s(f)\|_1 &= \sum_{t \in G} |\lambda_s(f)(t)| \\ &= \sum_{t \in G} \left|f\left(s^{-1}t\right)\right| \\ &= \sum_{r \in G} |f(r)|\end{aligned}$$

$$= \|f\|_1.$$

Just as in Proposition 4.2.2, it is the case that λ_s is linear. Additionally,

$$\begin{aligned}
\lambda_r \circ \lambda_s(f)(t) &= \lambda_s(f)\left(r^{-1}t\right) \\
&= f\left(s^{-1}r^{-1}(t)\right) \\
&= f\left((rs)^{-1}t\right) \\
&= \lambda_{rs}(f)(t).
\end{aligned}$$

We see that if $f \in \mathrm{Prob}(G)$, then for $f \geqslant 0$, we have $\lambda_s(f) \geqslant 0$, and

$$\begin{aligned}
\sum_{t \in G} \lambda_s(f)(t) &= \sum_{t \in G} f\left(s^{-1}t\right) \\
&= \sum_{r \in G} f(r) \\
&= 1
\end{aligned}$$

for any $f \in \mathrm{Prob}(G)$. □

Definition 5.2.2. For a countable group G, a sequence $(f_k)_k$ is called an approximate mean if, for all $s \in G$,

$$\|f_k - \lambda_s(f_k)\|_1 \xrightarrow{k \to \infty} 0.$$

To begin the forward direction regarding the equivalence between the Følner condition, approximate means, and means, we begin by showing that the existence of a Følner sequence implies the existence of an approximate mean. Then, we will show that the existence of an approximate mean implies the existence of an invariant state (hence mean).

Proposition 5.2.2. If G admits a Følner sequence $(F_k)_k$, then G admits an approximate mean.

Proof. Set $f_k = \frac{1}{|F_k|}\mathbb{1}_{F_k} \in \mathrm{Prob}(G)$. Then,

$$\begin{aligned}
\|f_k - \lambda_s(f_k)\|_1 &= \frac{1}{|f_k|}\|\mathbb{1}_{F_k} - \lambda_s(\mathbb{1}_{F_k})\| \\
&= \frac{1}{F_k}\|\mathbb{1}_{F_k} - \mathbb{1}_{sF_k}\| \\
&= \frac{|F_k \triangle sF_k|}{|F_k|} \\
&\to 0.
\end{aligned}$$

□

Proposition 5.2.3. If G admits an approximate mean, then G is amenable.

Proof. Let $(f_k)_k$ be an approximate mean. We define $\varphi_k = (\varphi_{f_k})_k$ (as in Definition 5.2.1) to be a sequence of states on $\ell_\infty(G)$.

Since the state space on $\ell_\infty(G)$ is w^*-compact, there is a state μ and a subnet $\left(\varphi_{k_j}\right)_j \xrightarrow{w^*} \mu$.

We only need to show that μ is invariant. Note that

$$|\mu(g) - \mu(\lambda_s(g))| \leqslant \left|\mu(g) - \varphi_{k_j}(g)\right| + \left|\varphi_{k_j}(g) - \varphi_{k_j}(\lambda_s(g))\right| + \left|\varphi_{k_j}(\lambda_s(g)) - \mu(\lambda_s(g))\right|$$

for all $g \in \ell_\infty(G)$, $s \in G$, and all j.

Given $\varepsilon > 0$, we find J such that for $j \geqslant J$,

$$\left|\mu(g) - \varphi_{k_j}(g)\right| < \varepsilon/3$$
$$\left|\mu(\lambda_s(g)) - \varphi_{k_j}(\lambda_s(g))\right| < \varepsilon/3.$$

We also see that

$$\begin{aligned}
\left|\varphi_{k_j}(g) - \varphi_{k_j}(\lambda_s(g))\right| &= \left|\sum_{t\in G} g(t) f_{k_j}(t) - \sum_{t\in G} g\left(s^{-1}t\right) f_{k_j}(t)\right| \\
&= \left|\sum_{t\in G} g(t) f_{k_j}(t) - \sum_{r\in G} g(r) f_{k_j}(sr)\right| && r = s^{-1}t \\
&= \left|\sum_{t\in G} g(t)\left(f_{k_j}(t) - \lambda_{s^{-1}}\left(f_{k_j}\right)(t)\right)\right| \\
&\leqslant \|g\|_\infty \sum_{t\in G} \left|f_{k_j}(t) - \lambda_{s^{-1}}\left(f_{k_j}\right)(t)\right| \\
&= \|g\|_\infty \left\|f_{k_j} - \lambda_{s^{-1}}\left(f_{k_j}\right)\right\|_1 \\
&< \varepsilon/3
\end{aligned}$$

for large j. Thus, we have

$$|\mu(g) - \mu(\lambda_s(g))| < \varepsilon,$$

for all $\varepsilon > 0$, so $\mu(g) = \mu(\lambda_s(g))$. □

We will now commence with the reverse direction, starting by showing that amenability implies the existence of an approximate mean, and then showing that the existence of an approximate mean implies that the Følner condition is satisfied.

Proposition 5.2.4. If G is amenable, then G admits an approximate mean.

Proof. Suppose G does not admit an approximate mean. Then, there exists a finite subset $E_0 \subseteq G$ and $\varepsilon_0 > 0$ such that for all $s \in E_0$ and all $f \in \mathrm{Prob}(G)$, we have $\|f - \lambda_s(f)\| \geqslant \varepsilon_0$.

Consider the set

$$X = \bigoplus_{|E_0|} \ell_1(G),$$

endowed with the norm

$$\begin{aligned}
\left\|(f_s)_{s\in E_0}\right\| &= \sum_{s\in E_0} \sum_{t\in G} |f_s(t)| \\
&= \sum_{s\in E_0} \|f_s\|_1,
\end{aligned}$$

and let

$$C = \left\{(f - \lambda_s(f))_{s\in E_0} \;\middle|\; f \in \mathrm{Prob}(G)\right\}.$$

Since $\mathrm{Prob}(G)$ is convex, it is the case that C is convex, and since $|E_0|$ is finite, C is necessarily bounded. Note that $0 \notin \overline{C}$.

By the Hahn–Banach separation for convex sets (Theorem D.4.3), there is a real-valued $\varphi \in X^*$ such that $\varphi(C) \geqslant 1$. Here,

$$
\begin{aligned}
X^* &\cong \bigoplus_{|E_0|} \ell_1(G)^* \\
&\cong \sum_{|E_0|} \ell_\infty(G),
\end{aligned}
$$

endowed with the norm

$$
\begin{aligned}
\left\| (g_s)_{s\in E_0} \right\| &= \max_{s\in E_0} \left(\sup_{t\in G} |g_s(t)| \right) \\
&= \max_{s\in E_0} \|g_s\|_\infty .
\end{aligned}
$$

We let $\varphi = \left(\varphi_{g_s}\right)_{s\in E_0}$, where $g_s \in \ell_\infty(G)$ is defined by the duality

$$
\varphi_{g_s}(f) = \sum_{t\in G} f(t) g_s(t).
$$

Thus, for all $f \in \operatorname{Prob}(G)$, we have

$$
\begin{aligned}
1 &\leqslant \varphi\left((f - \lambda_s(f))_{s\in E_0}\right) \\
&= \sum_{s\in E_0} \varphi_{g_s}(f - \lambda - s(f)) \\
&= \sum_{s\in E_0} \sum_{t\in G} (f - \lambda_s(f))(t) g_s(t) \\
&= \sum_{s\in E_0} \left(\sum_{t\in G} f(t) g_s(t) - \sum_{t\in G} f\left(s^{-1}t\right) g_s(t) \right) \\
&= \sum_{s\in E_0} \left(\sum_{t\in G} f(t) g_s(t) - \sum_{r\in G} f(r) g_s(sr) \right) \\
&= \sum_{s\in E_0} \left(\sum_{r\in G} f(r) g_s(r) - \sum_{r\in G} f(r) \lambda_{s^{-1}}(g)(r) \right) \\
&= \sum_{s\in E_0} \sum_{r\in G} f(r)(g_s - \lambda_{s^{-1}}(g_s))(r).
\end{aligned}
$$

Note that this holds for any $f \in \operatorname{Prob}(G)$, including the case of $f = \delta_t$ for a given $t \in G$. We must have

$$
\begin{aligned}
&= \sum_{s\in E_0} \sum_{r\in G} \delta_t(r)(g_s(r) - \lambda_{s^{-1}}(g_s))(r) \\
&= \sum_{s\in E_0} (g_s - \lambda_{s^{-1}}(g))(t),
\end{aligned}
$$

and in particular,

$$
\geqslant \mathbb{1}_G .
$$

Since G is amenable, there is a mean $\mu \colon \ell_\infty(G) \to \mathbb{C}$ with $\mu(g_s) = \mu(\lambda_{s^{-1}}(g_s))$, meaning

$$
\begin{aligned}
0 &= \mu\left(\sum_{s\in E_0} (g_s - \lambda_{s^{-1}}(g_s))(t) \right) \\
&\geqslant \mu(\mathbb{1}_G) \\
&= 1,
\end{aligned}
$$

which is a contradiction. ◻

To show that the existence of an approximate mean implies the Følner condition, we require the following lemma.

Lemma 5.2.1. Let $f\colon S \to \mathbb{R}$ be finitely supported with $\sum_{s\in S} f(s) = 1$. Then, there exist subsets $\{F_i\}_{i=1}^n$, where $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n$, and constants $\{c_i\}_{i=1}^n$, such that

$$f = \sum_{i=1}^n c_i \mathbb{1}_{F_i},$$

where

$$\sum_{i=1}^n c_i |F_i| = 1.$$

This is known as the layer cake representation for f.

Proof. We define $F_1 = \mathrm{supp}(f)$, and take $c_1 = \min(\mathrm{Ran}(f))$. Taking $E_1 = f^{-1}(c_1)$ (as a set-theoretic inverse), we define $F_2 = F_1 \setminus E_1$.

Take $d_1 = \min(f(F_2))$, and define $c_2 = d_1 - c_1$. Then, defining $E_2 = f^{-1}(d_1)$, $F_3 = F_2 \setminus E_2$, and $d_2 = \min(f(F_3))$, we define $c_3 = d_2 - c_2 - c_1$.

Continuing in this pattern, we find $d_{i-1} = \min(f(F_i))$, $E_i = f^{-1}(d_{i-1})$, and $c_i = d_{i-1} - \sum_{j=1}^{i-1} c_i$.

This yields a decomposition $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n$, where $\sum_{i=1}^n c_i \mathbb{1}_{F_i} = f$ by construction.

We now verify that $\sum_{i=1}^n c_i |F_i| = 1$.

$$\begin{aligned} 1 &= \sum_{s\in S} f(s) \\ &= \sum_{s\in S} \sum_{i=1}^n c_i \mathbb{1}_{F_i}(s). \end{aligned}$$

By definition, if $s \in F_j$ for some j, we see that c_j is summed for $\left|F_j\right|$ many times. Thus, we obtain

$$= \sum_{i=1}^n c_i |F_i|.$$

$\square$

Remark 5.2.1. Instead of using this construction where we take set-theoretic inverses and remove "residual" sets, there is an alternative method of construction that involves ordering the range as $r_1 < r_2 < \cdots < r_n$, and constructing the set $F_k = \{s \mid f(s) \geqslant r_k\}$.

We will use the layer cake decomposition to prove that if G admits an approximate mean, then G satisfies the Følner condition.

Proposition 5.2.5. Let G admit an approximate mean. Then, G satisfies the Følner condition.

Proof. Let $(f_k)_k$ be an approximate mean, as in Definition 5.2.2. Fix a finite nonempty set $S \subsetneq G$. Then, by the definition of an approximate mean, there must exist some $N \in \mathbb{N}$ such that for all $k \geqslant N$ and all $s \in G$,

$$\|f_k - \lambda_s(f_k)\|_1 \leqslant \frac{\varepsilon}{|S|}.$$

In particular, this holds for f_N and for all $s \in S$.

Since $f_N \in \mathrm{Prob}(G)$ is finitely supported and $\sum_{s\in G} f_N(s) = 1$, we may use Lemma 5.2.1 to rewrite f_N as

$$f_N = \sum_{i=1}^{n} c_i \mathbb{1}_{F_i},$$

where $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n$, and $\sum_{i=1}^{n} c_i|F_i| = 1$.

For a given $1 \leqslant i \leqslant n$, for each $s \in S$ and $t \in sF_i \triangle F_i$, we have

$$f_N(t) - f_N\left(s^{-1}t\right) = \begin{cases} c_i & t \in F_i \setminus sF_i \\ -c_i & t \in sF_i \setminus F_i \end{cases}.$$

Thus, we see that $|f_N(t) - \lambda_s(f_N)(t)| \geqslant c_i$ on $sF_i \triangle F_i$. Thus, for each $s \in S$,

$$\begin{aligned} \sum_{i=1}^{n} c_i|sF_i \triangle F_i| &\leqslant \sum_{t\in S} |f_N(t) - \lambda_s(f)(t)| \\ &< \frac{\varepsilon}{|S|} \\ &= \frac{\varepsilon}{|S|} \sum_{i=1}^{n} c_i|F_i|. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \sum_{s\in S}\sum_{i=1}^{n} c_i|sF_i \triangle F_i| &< \frac{\varepsilon}{|S|} \sum_{s\in S}\sum_{i=1}^{n} c_i|F_i| \\ &= \varepsilon \sum_{i=1}^{n} c_i|F_i|. \end{aligned}$$

Thus, by the pigeonhole principle, there must exist some $1 \leqslant i \leqslant n$ for which

$$\sum_{s\in S} c_i|sF_i \triangle F_i| < \varepsilon c_i|F_i|.$$

Setting $F = F_i$, we find that, for all $s \in S$,

$$\begin{aligned} \frac{|sF\triangle F|}{|F|} &\leqslant \sum_{s\in S} \frac{|sF\triangle F|}{|F|} \\ &< \varepsilon. \end{aligned}$$

□

Thus far, we have shown the following to be equivalent for a discrete group G:

(1) G is non-paradoxical;

(2) G is amenable;

(3) G admits an invariant state;

(4) G admits an approximate mean;

(5) G satisfies the Følner condition.

The equivalence between (1) and (2) follows from Theorem 3.0.1, the equivalence between (2) and (3) follows from Propositions 4.2.3 and 4.2.4, and the equivalence between (3), (4), and (5) follows from Propositions 5.2.2, 5.2.3, 5.2.4, and 5.2.5.

5.3 Applying Følner's Condition: Groups of Subexponential Growth

Before we move to Chapters 7 and 8 to discuss representations of groups inside the algebra of bounded operators on a Hilbert space, we will provide an application of Følner's condition by taking a tour into geometric group theory. In this section, we will establish the amenability of yet another wide class of groups (just as we established that all abelian groups are amenable in Chapter 5) — the groups of subexponential growth.

First, we construct a little bit of machinery to understand the growth rate of a group, then we prove that Følner's condition holds for these special classes of groups.

Definition 5.3.1. Let G be a group with finite symmetric generating set S (see Definition 1.1.1). Then, we define the word length of $g \in G$ with respect to S to be

$$\ell_{G,S}(g) = \min\{n \mid g = s_1 \dots s_n,\ s_i \in S\},$$

taking $\ell_{G,S}(e_G) = 0$. We define the word metric on G with respect to S by taking

$$d_S(g, h) = \ell_{G,S}\left(g^{-1}h\right).$$

Fact 5.3.1. If S and T are finite symmetric generating sets for G, then the respective word metrics d_S and d_T are equivalent (as in the sense of Definition B.2.1).

Proof. We start by showing that d_S is indeed a metric. Notice that the following facts necessarily hold by our definition of the word length:

- $\ell_{G,S}(g) = \ell_{G,S}(g^{-1})$;
- $\ell_{G,S}(gh) \leqslant \ell_{G,S}(g) + \ell_{G,S}(h)$.

We thus have:

$$\begin{aligned} d_S(g, h) &= \ell_{G,S}\left(g^{-1}h\right) \\ &= \ell_{G,S}\left(h^{-1}g\right) \\ &= d_S(h, g) \end{aligned}$$

$$\begin{aligned} d_S(g, h) &= \ell_{G,S}\left(g^{-1}h\right) \\ &= \ell_{G.S}\left(g^{-1}kk^{-1}h\right) \\ &\leqslant \ell_{G,S}\left(g^{-1}k\right) + \ell_{G,S}\left(k^{-1}h\right) \\ &= d_S(g, k) + d_S(k, h) \end{aligned}$$

$$\begin{aligned} d_S(g, g) &= \ell_{G,S}\left(g^{-1}g\right) \\ &= \ell_{G,S}(e_G) \\ &= 0 \end{aligned}$$

$$\begin{aligned} d_S(g,h) = 0 &\Leftrightarrow \ell_{G,S}\left(g^{-1}h\right) = 0 \\ &\Leftrightarrow g^{-1}h = e_G \\ &\Leftrightarrow g = h. \end{aligned}$$

Thus, d_S is indeed a metric.

Let S and T be finite symmetric generating sets for G. It is sufficient to show that there exists some $k \in \mathbb{N}$ such that, for all $g \in G$,

$$\frac{1}{k}\ell_{G,S}(g) \leqslant \ell_{G,T}(g) \leqslant k\ell_{G,S}(g).$$

Set

$$\begin{aligned} M &= \max\{\ell_{G,T}(s) \mid s \in S\} \\ N &= \max\{\ell_{G,S}(t) \mid t \in T\}. \end{aligned}$$

Now, let $n = \ell_{G,S}(g)$, such that $g = s_1 \cdots s_n$, where $s_i \in S$. Then, we have

$$\begin{aligned} \ell_{G,T}(g) &= \ell_{G,T}(s_1 \cdots s_n) \\ &\leqslant \ell_{G,T}(s_1) + \cdots + \ell_{G,T}(s_n) \\ &\leqslant M\ell_{G,S}(g), \end{aligned}$$

and similarly, $\ell_{G,S}(g) \leqslant N\ell_{G,T}(g)$. Setting $k = \max(M, N)$, we get

$$\frac{1}{k}\ell_{G,S}(g) \leqslant \ell_{G,T}(g) \leqslant k\ell_{G,S}(g).$$

□

Now, we may begin defining the growth rate of a group. We will use the fact that all word metrics with respect to a generating set are symmetric in order to show that the growth rate is well-defined (i.e., independent of the generating set for G).

Definition 5.3.2. Let G be a group with finite symmetric generating set S. We define

$$\begin{aligned} B_{G,S}(n) &= \{g \in G \mid \ell_{G,S}(g) \leqslant n\}; \\ \gamma_{G,S}(n) &= |B_{G,S}(n)|. \end{aligned}$$

The following facts hold for γ.

Fact 5.3.2. Let G be a finitely generated group. The following facts hold:

(1) $\gamma_{G,S}(n)$ is an increasing function;

(2) $\gamma_{G,S}(n+m) \leqslant \gamma_{G,S}(n)\gamma_{G,S}(m)$;

(3) $\lim\limits_{n\to\infty} (\gamma_{G,S}(n))^{1/n} = \rho_{G,S}$ exists;

(4) if S and T are finite symmetric generating sets for G, then there exists $k \in \mathbb{N}$ such that $\gamma_{G,T}(n) \leqslant \gamma_{G,S}(kn)$ for all $n \in \mathbb{N}$, and $\rho_{G,S} = \rho_{G,T}$.

Proof.

(1) Since $B_{G,S}(n) \subseteq B_{G,S}(n+1)$, we have $\gamma_{G,S}(n) \leqslant \gamma_{G,S}(n+1)$, so $\gamma_{G,S}$ is increasing.

(2) We start by showing that $B_{G,S}(n)B_{G,S}(m) = B_{G,S}(n+m)$. First, if $g \in B_{G,S}(n)$ and $h \in B_{G,S}(m)$, we know that $\ell_{G,S}(gh) \leqslant \ell_{G,S}(g) + \ell_{G,S}(h) \leqslant m + n$, so $B_{G,S}(n)B_{G,S}(n) \subseteq B_{G,S}(n+m)$. Additionally, if

$g \in B_{G,S}(n+m)$, we may write

$$g = \underbrace{s_1 \cdots s_\ell}_{g_1} \underbrace{s_{\ell+1} \cdots s_k}_{g_2},$$

where $k \leqslant n + m$, $\ell \leqslant n$, and $k - \ell \leqslant m$, so $g_1 \in B_{G,S}(n)$ and $g_2 \in B_{G,S}(m)$. Thus, we have $B_{G,S}(n)B_{G,S}(m) = B_{G,S}(n+m)$.

Now, we have

$$\begin{aligned}\gamma_{G,S}(n+m) &= |B_{G,S}(n+m)| \\ &= |B_{G,S}(n)B_{G,S}(m)| \\ &\leqslant |B_{G,S}(n)||B_{G,S}(m)| \\ &= \gamma_{G,S}(n)\gamma_{G,S}(m).\end{aligned}$$

(3) From (2), we know that $\gamma_{G,S}(n) \leqslant \gamma_{G,S}(1)^n$. Inductively, we have

$$\gamma_{G,S}(n+1) \leqslant \gamma_{G,S}(1)^{n+1},$$

and thus,

$$1 \leqslant \gamma_{G,S}(n)^{1/n} \leqslant \gamma_{G,S}(1).$$

(4) We know that there exists k such that $\frac{1}{k}\ell_{G,S} \leqslant \ell_{G,T} \leqslant k\ell_{G,S}$ by the proof of Fact 5.3.1. Thus, if $g \in B_{G,T}(n)$, then $\ell_{G,T}(g) \leqslant n$, so $\ell_{G,S}(g) \leqslant kn$, so $g \in B_{G,S}(kn)$ and $B_{G,T}(n) \subseteq B_{G,T}(kn)$. We have $\gamma_{G,T}(n) \leqslant \gamma_{G,S}(kn)$.

Similarly, if $g \in B_{G,S}(n)$, then $\ell_{G,S}(g) \leqslant n$, so $\ell_{G,T}(g) \leqslant kn$, and $g \in B_{G,T}(kn)$. Thus, we get $B_{G,S}(n) \subseteq B_{G,T}(kn)$, so $\gamma_{G,S}(n) \leqslant \gamma_{G,T}(kn)$.

It follows that

$$\gamma_{G,S}\left(\frac{n}{k}\right)^{1/n} \leqslant \gamma_{G,T}(n)^{1/n} \leqslant \left(\gamma_{G,S}(kn)^k\right)^{1/kn}.$$

Sending $n \to \infty$, we get $\rho_{G,S} \leqslant \rho_{G,T} \leqslant \rho_{G,S}$, so $\rho_{G,S} = \rho_{G,T}$.

□

Definition 5.3.3. Let G be a group with finite symmetric generating set S. The quantity

$$\rho_G = \limsup_{n\to\infty} \gamma_{G,S}(n)^{1/n}$$

is known as the growth rate of the group G. If we have $\rho = 1$, then we say G is of subexponential growth.

Fact 5.3.3. All finite groups are of subexponential growth.

Proof. Note that since ρ is independent of the generating set (as we proved in Fact 5.3.2), we can set $S = G$, and we have $\limsup_{n\to\infty} |G|^{1/n} = 1$. □

Fact 5.3.4. Let Γ be a finitely generated abelian group. Then, Γ is of subexponential growth.

Proof. We start by showing that $G = \mathbb{Z}^d$ is of subexponential growth. Notice that every element of $\mathbb{Z}^d$ is some linear combination of the set

$$S = \{e_1, e_2, \dots, e_d\}, \tag{$*$}$$

where

$$e_j = (0, 0, \dots, \underbrace{1}_{\text{position } j}, 0, 0, \dots).$$

Additionally, we see that any element of $B_{G,S}(n)$ is of the form $e_1^{i_1} e_2^{i_2} \dots e_d^{i_d}$, where $\sum_{j=1}^{d} i_j \leqslant n$. Thus, we must have $\gamma_{G,S}(n) \leqslant n^d$, meaning that

$$\begin{aligned}\rho &= \limsup_{n\to\infty} \gamma_{G,S}(n)^{1/n} \\ &= \limsup_{n\to\infty} n^{d/n} \\ &= 1,\end{aligned}$$

so $\mathbb{Z}^d$ is of subexponential growth.

Now, if $G' = \mathbb{Z}^d \times \mathbb{Z}/k_1\mathbb{Z} \times \cdots \times \mathbb{Z}/k_r\mathbb{Z}$, then since there is a torsion subgroup in G', we must have $\gamma_{G',S'}(n) \leqslant \gamma_{\mathbb{Z}^{d+r},T}(n)$ for any n, where T is a generating set for $\mathbb{Z}^{d+r}$ and S' is a generating set for G'. Since

$$\begin{aligned}\rho_{\mathbb{Z}^{d+r}} &= \limsup_{n\to\infty} \gamma_{\mathbb{Z}^{d+r},T}(n)^{1/n} \\ &= 1,\end{aligned}$$

and $1 \leqslant \gamma_{G',S'}(n)$, we must have $\rho_{G'} = 1$.

Since, by the fundamental theorem of finitely generated abelian groups, it is the case that $\Gamma \cong \mathbb{Z}^d \times \mathbb{Z}/k_1\mathbb{Z} \times \cdots \times \mathbb{Z}/k_r\mathbb{Z}$ for some $d, k_1, \dots, k_r \in \mathbb{N}$, Γ is of subexponential growth. □

To prove that the groups of subexponential growth are amenable, we use the following lemma from real analysis.

Lemma 5.3.1. Let $(a_n)_n$ be a sequence such that $a_n > 0$ for each n. Then,

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} (a_n)^{1/n}.$$

Similarly,

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = \limsup_{n\to\infty} (a_n)^{1/n}.$$

Theorem 5.3.1. Let Γ be a finitely generated group of subexponential growth. Then, Γ is amenable.

Proof. To prove that Γ is amenable, we show that it satisfies the Følner condition. From the results in Chapter 5, we know that this implies that Γ is amenable. Let S be a finite symmetric generating set for Γ.

For any $\varepsilon > 0$, we see that there is some $k \in \mathbb{N}$ such that

$$|B_{\Gamma,S}(k)|^{1/k} \leqslant 1 + \varepsilon.$$

Thus, by the lemma above, we must have

$$\frac{|B_{\Gamma,S}(k+1)|}{|B_{\Gamma,S}(k)|} \leqslant 1+\varepsilon.$$

Note that, by Lemma 5.1.2, we only need to verify that the Følner condition holds on S. For any $s \in S$, we have

$$\begin{aligned}\frac{|sB_{\Gamma,S}(k) \triangle B_{\Gamma,S}(k)|}{|B_{\Gamma,S}(k)|} &\leqslant \frac{2(|B_{G,S}(k+1)| - |B_{\Gamma,S}(k)|)}{|B_{\Gamma,S}(k)|} \\ &\leqslant 2\varepsilon.\end{aligned}$$

Therefore, Γ satisfies the Følner condition, hence is amenable. □

Remark 5.3.1. The result in Theorem 5.3.1 can be used along with Fact 5.3.4 and Corollary 4.3.3 to prove Corollary 4.3.4.

Chapter 6

Further Analysis: the Left-Regular Representation

Chapter 7

The Nuclear Option: Amenability in C^*-Algebras

Appendix A

Algebra and Linear Algebra

In general, as we progress through these appendices, we will consistently add additional structure to a set. First, we begin by developing groups, rings, and fields, vector spaces, and algebras. In the following appendices, we will apply metric structures, topologies, and measures, building up to the central structure of functional analysis: normed vector spaces and the operators on these normed vector spaces.

These appendices were largely written to provide essential background for the techniques and results that will appear in the main body of the text. As such, they do not include detailed proofs — occasionally, we will include outlines for certain proofs in the remarks. The proofs for many of these results can be found in relevant (and some not-as-relevant) texts.

We make heavy use of results from algebra and linear algebra in this thesis. Some excellent resources to learn more about algebra and linear algebra are [DF04] and [Alu09]. Most of the theorems are presented without proof, not because we do not want to state their proofs, but because this thesis is already long enough.

A.1 Group, Rings, (some) Fields

A.1.1 Groups

Definition A.1.1 (Groups)**.** Let A be a set, and let $\star$ be a binary operation on A. We say A is a group if

- A is closed under the operation $\star$;
- $\star$ is associative, such that for all $a, b, c \in A$, $(a \star b) \star c = a \star (b \star c)$;
- A has an identity element e_A, where $a \star e_A = e_A \star a = a$ for any $a \in A$;
- for any $a \in A$, there exists $a^{-1} \in A$ such that $a^{-1} \star a = a \star a^{-1} = e$.

If the operation $\star$ is such that $a \star b = b \star a$ for all $a, b \in A$, then we say A is an abelian group.

Generally, we abbreviate $a \star b := ab$.

Definition A.1.2 (Subgroups, Normal Subgroups, and Quotient Groups)**.** If G is a group, $H \subseteq G$ is a subgroup if H is closed under the group operation and inverses. We write $H \leqslant G$.

If H is a subgroup, a left coset of H is the set $gH := \{gh \mid h \in H\}$, where $g \in G$. Similarly, a right coset of H is the set $Hg := \{hg \mid h \in H\}$. The index of H, denoted $[G : H]$, is the number of left (or right) cosets of H.

If $H \leqslant G$ is also such that, for any $g \in G$ and $h \in H$, $ghg^{-1} \in H$, then we call H a normal subgroup of G. We write $H \trianglelefteq G$.

Defining the equivalence relation $g \sim g'$ if and only if $g^{-1}g' \in H$, the group of equivalence classes $gH := [g]$ is known as the quotient group G/H.

If the only normal subgroups of a group G are G itself and $\{e_G\}$, then we say the group G is simple.

Definition A.1.3. Let G and H be groups. A map $\varphi\colon G \to H$ is called a (group) homomorphism if, φ "preserves the group structure," in the sense that

$$\varphi(ab) = \varphi(a)\varphi(b)$$
$$\varphi\left(a^{-1}\right) = \varphi(a)^{-1}$$

for all $a, b \in G$.

We define $\ker(\varphi)$ to be the set of all $g \in G$ such that $\varphi(g) = e_H$.

If $H \trianglelefteq G$, the map $\pi\colon G \to G/H$ that sends $g \mapsto gH$ is known as the canonical projection.

If φ is a bijection, then φ is known as an isomorphism. We write $G \cong H$ if there exists an isomorphism $\varphi\colon G \to H$.

Theorem A.1.1 (First Isomorphism Theorem for Groups)**.** Let G and H be groups, and let $\varphi\colon G \to H$ be a group homomorphism. Then, $\ker(\varphi) \trianglelefteq G$ is a normal subgroup, and $G/\ker(\varphi) \cong \operatorname{im}(\varphi)$.

There is a complete classification of finitely generated (and finite) abelian groups.[1]

Theorem A.1.2. Let G be a finitely generated Abelian group. Then, there is some $d \in \mathbb{N}$ and some $k_1, \ldots, k_n$ such that

$$G \cong \underbrace{\mathbb{Z}^d}_{F} \times \underbrace{\mathbb{Z}/k_1\mathbb{Z} \times \mathbb{Z}/k_2\mathbb{Z} \times \cdots \times \mathbb{Z}/k_n\mathbb{Z}}_{T}.$$

The group F is known as the free subgroup of G, and the group T is known as the torsion subgroup of G. If G is finite, then G is isomorphic to some torsion group $\mathbb{Z}/k_1\mathbb{Z} \times \cdots \times \mathbb{Z}/k_1\mathbb{Z}$.

Furthermore, we may also take each of k_i in both cases to be equal to $p_i^{e_i}$ for some prime p_i and some $e_i \in \mathbb{N}$.

Definition A.1.4. A group G is solvable if if admits a finite series of normal subgroups

$$e_G = G_0 \trianglelefteq G_1 \trianglelefteq \cdots \trianglelefteq G_n \trianglelefteq G$$

such that G_j/G_{j-1} is abelian for each $j = 1, \ldots, n$.

Definition A.1.5 (Group Actions)**.** Let G be a group, and let A be a set. A (left) group action of G on A is a map $\rho\colon G \times A \to A$ such that, for all $a \in A$,

- $\rho(e_G, a) = a$;
- $\rho(g, \rho(h, a)) = \rho(gh, a)$.

We abbreviate $\rho(g, a) = g \cdot a$.

The permutation representation of the action ρ is a homomorphism $\varphi\colon G \to \operatorname{Sym}(A)$.

[1]There is also a complete classification of all the finite simple groups, but we absolutely do not have enough space for that one.

Definition A.1.6 (Kernels, Stabilizers, and Orbits)**.** Let G act on A, and let $a \in A$.

- The stabilizer of a under G is the set of elements in G that fix a:

$$G_a \coloneqq \{g \in G \mid g \cdot a = a\}.$$

- The kernel of the action of G on A is the intersection of the stabilizers of G:

$$\begin{aligned}\text{kernel} &\coloneqq \bigcap_{a \in A} G_a \\ &= \{g \in G \mid g \cdot a \text{ for all } a \in A\}.\end{aligned}$$

- The action is faithful if the kernel is e_G.
- The action is free if $G_a = \{e_G\}$ for all $a \in A$.
- The orbit of a is the equivalence class $[a]_\sim$ under the relation $a \sim b$ if there exists some $g \in G$ such that $a = g \cdot b$:

$$G \cdot a = \{b \in A \mid b = g \cdot a \text{ for some } g \in G\}.$$

Theorem A.1.3 (Orbit-Stabilizer Theorem)**.** If G acts on A, and $a \in A$, then the number of elements in the orbit of a is the index of the stabilizer of a. In symbolic form,

$$|G \cdot a| = [G : G_a].$$

Remark A.1.1. Various theorems such as the Sylow Theorems and Lagrange's Theorem fall out of the Orbit-Stabilizer theorem.

A.1.2 Rings and Fields

Definition A.1.7 (Rings)**.** Let A be a set. Specifically, let A be an abelian group, letting $+$ denote the operation on A and $0 \coloneqq e_A$. Then, A is a ring if A also admits a multiplication, $\cdot$, such that

- $a \cdot (b + c) = a \cdot b + a \cdot c$;
- $(a + b) \cdot c = a \cdot c + b \cdot c$;
- $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

If the multiplication on A is commutative, then we say A is a commutative ring. If A admits an element 1_A such that $a \cdot 1_A = 1_A \cdot a = a$, then we say A is a unital ring.

When referring to the abelian group of A under $+$, we often write $(A, +)$.

Definition A.1.8 (Subrings, Ideals, and Quotient Rings)**.** Let R be a ring. A subset $A \subseteq R$ is known as a subring if A is a subgroup of $(R, +)$, and A is closed under multiplication. In other words, for all $a, b \in A$, we have

- $a - b \in A$;
- $ab \in A$,

where $a - b = a + (-b)$.

If $I \subseteq R$ is a subring that also has the property that, for all $x \in I$ and $r \in R$, $rx \in I$ and $xr \in I$, then we say I is an ideal.

Similar to the case of groups and normal subgroups, we can form the quotient ring R/I by defining the equivalence relation $a \sim b$ if $a - b \in I$, and defining $a + I := [a]_\sim$.

Definition A.1.9 (Ring Homomorphism)**.** If R and S are rings, then a map $\varphi\colon R \to S$ is a ring homomorphism if φ "preserves the ring structure," in the sense that, for all $a, b \in R$,

- $\varphi(a + b) = \varphi(a) + \varphi(b)$;
- $\varphi(ab) = \varphi(a)\varphi(b)$.

The kernel of the ring homomorphism is defined to be the set of all elements $a \in R$ such that $\varphi(a) = 0_S$.

If $I \subseteq R$ is an ideal, then the map $\pi\colon R \to R/I$ that sends $a \mapsto a + I$ is known as the canonical projection.

If φ is a bijection, then φ is known as an isomorphism. We write $R \cong S$ if there exists an isomorphism $\varphi\colon R \to S$.

Analogously, there is a first isomorphism theorem for rings.

Theorem A.1.4 (First Isomorphism Theorem for Rings)**.** Let R and S be rings, and let $\varphi\colon R \to S$ be a ring homomorphism. Then, $\ker(\varphi) \subseteq R$ is an ideal, and $R/\ker(\varphi) \cong \operatorname{im}(\varphi)$.

The ideal structure of a ring R admits some more specification.

Definition A.1.10. An ideal $I \subsetneq R$ is said to be maximal if, for any other ideal $J \subseteq R$ with $I \subseteq J$, either $I = J$ or $I = R$.

Theorem A.1.5 (Krull's Theorem)**.** If $I \subseteq R$ is a proper ideal, then there exists some maximal ideal $M \subseteq R$ such that $I \subseteq M$.

Remark A.1.2. Krull's Theorem can be proven using Zorn's Lemma B.1.1 applied to the partially ordered set of proper ideals ordered by inclusion.

Definition A.1.11. Let R be a unital ring.

- If $a, b \in R$ are nonzero elements such that $ab = 0$, then we say a and b are zero divisors in R.
- If R is commutative and does not contain any zero divisors, then we say R is an integral domain.
- If an element $a \in R$ is such that there exists some b such that $ab = ba = 1_R$, then we call a a unit.
- If R is a such that every element of R is a unit, then we say R is a division algebra.
- If R is a division algebra that is commutative, then R is a field.

Remark A.1.3. Generally, when we deal with fields, we will usually be dealing with the complex numbers, $\mathbb{C}$, unless otherwise stated.

A.2 Linear Algebra

Certain constructions in linear algebra are extremely important in understanding functional analysis. We provide an overview of the theory of vector spaces and linear transformations in the purely algebraic context. Analytic properties that result from applying norms on these vector spaces will appear in Appendix D.

A.2.1 The Structure of Vector Spaces

Definition A.2.1. Let X be some set, and $\mathbb{F}$ some field (generally, we assume $\mathbb{F} = \mathbb{C}$). We say X is an $\mathbb{F}$-vector space if X is equipped with two operations:

- scalar multiplication: $m\colon \mathbb{F} \times X \to X$, which sends $(\alpha, x) \mapsto \alpha x$; and

- vector addition: $a\colon X \times X \to X$, which sends $(x, y) \mapsto x + y$.

In general, $(X, +)$ is an abelian group, and scalar multiplication satisfies the following identities, for all $\alpha, \beta \in \mathbb{F}$ and $x, y \in X$:

- $\alpha(\beta x) = (\alpha\beta)x$;
- $\alpha(x + y) = \alpha x + \alpha y$;
- $(\alpha + \beta)x = \alpha x + \beta x$;
- $1_{\mathbb{F}}x = x$;
- $0_{\mathbb{F}}x = 0_X$.

There are certain geometric properties of subsets of vector spaces that we will be using a lot, especially when we discuss locally convex topologies on these vector spaces.

Definition A.2.2. Let X be a $\mathbb{C}$-vector space.

- If $A, B \subseteq X$, then we define

$$A + B = \{x + y \mid x \in A, y \in B\}.$$

  If $A = \{x_0\}$, we abbreviate $\{x_0\} + B$ as $x_0 + B$, which is called the translation of B.

- If $A \subseteq X$, and $\alpha \in \mathbb{C}$, then

$$\alpha A = \{\alpha x \mid x \in A\}$$

  is the scaling of A by α. We write $(-1)A = -A$.

- A subset $A \subseteq X$ is called symmetric if $-A = A$.
- A subset $A \subseteq X$ is called balanced if $\alpha A \subseteq A$ for all $|\alpha| \leqslant 1$.
- A subset $C \subseteq X$ is called convex if for all $t \in [0, 1]$ and $x_1, x_2 \in C$, $(1 - t)x_1 + tx_2 \in C$.

We define

$$\begin{aligned}\operatorname{conv}(A) &= \bigcap\{C \mid A \subseteq C \subseteq X, C \text{ is convex}\}\\ &= \left\{\sum_{j=1}^{n} t_j a_j \,\middle|\, n \in \mathbb{N}, t_j \geqslant 0, \sum_{j=1}^{n} t_j = 1, a_j \in A\right\}.\end{aligned}$$

Definition A.2.3. If X is a vector space, and $M \subseteq X$ is a subset such that, for all $x, y \in M$ and $\alpha \in \mathbb{C}$, $\alpha x + y \in M$, then M is called a subspace of X.

If M is a subspace, we may define the equivalence relation $x \sim_M y$ if and only if $x - y \in M$. Equivalence classes under the relation $\sim_M$ are written $x + M := [x]_\sim$, and form the quotient space X/M.

Furthermore, if $\{M_i\}_{i \in I}$ is a family of subspaces of X and, for any $x \in X$, there is a unique sum

$$x = \sum_{i \in I} x_i,$$

where $x_i \in M_i$, then we say X is the internal direct sum of the family of subspaces $\{M_i\}_{i \in I}$,

$$X = \bigoplus_{i \in I} M_i.$$

Definition A.2.4. Let X be a vector space, and let $\{x_i\}_{i\in I} \subseteq X$ be a subset.

- The set $\{x_i\}_{i\in I}$ is called linearly independent if, for any finite linear combination such that

$$\sum_{i\in I} \alpha_i x_i = 0_X,$$

  it is the case that all $\alpha_i = 0$.

- The set $\{x_i\}_{i\in I}$ is called spanning for X if the set of all finite linear combinations $\sum_{i\in I} \alpha_i x_i$ is equal to X.

- The set $\{x_i\}_{i\in I}$ is called a basis for X if it is linearly independent and spanning.

Definition A.2.5. If X is a vector space, and $\mathcal{B} \subseteq X$ is a basis, then $\dim(X) := |\mathcal{B}|$.

If $M \subseteq X$ is a subspace, then the codimension of M is $\dim(X/M)$.

Remark A.2.1. Every vector space has a basis. This can be proven with Zorn's Lemma (Theorem B.1.1) applied on the partially ordered set (Definition B.1.1) of linearly independent subsets ordered by inclusion.

Additionally, not only does every vector space have a basis, but for any set, there is a vector space with the set as its basis (see Theorem 1.2.1).

A.2.2 Linear Maps

Linear algebra is not only the study of vector spaces, but also the study of linear maps on these vector spaces.

Definition A.2.6. Let X and Y be vector spaces. A map $T\colon X \to Y$ is called linear if for every $x, x_1, x_2 \in X$ and $\alpha \in \mathbb{C}$, we have

$$\begin{aligned} T(x_1 + x_2) &= T(x_1) + T(x_2) \\ T(\alpha x) &= \alpha T(x). \end{aligned}$$

We write $I_X := \mathrm{id}_X$.

The set of linear maps between X and Y is denoted $\mathcal{L}(X, Y)$. The set of all linear maps between X and X is abbreviated $\mathcal{L}(X)$.

A linear map $\varphi\colon X \to \mathbb{C}$ is called a linear functional on X. The collection of linear functionals on X is called the algebraic dual of X, written $X' := \mathcal{L}(X, \mathbb{C})$.

The space $\mathcal{L}(X, Y)$ equipped with pointwise operations is a vector space over $\mathbb{C}$.

Definition A.2.7 (Four Fundamental Subspaces)**.** Let $T\colon X \to Y$ be a linear map.

- The kernel of T, $\ker(T)$, is the set of all $x \in X$ such that $T(x) = 0_X$.
- The range of T, $\mathrm{Ran}(T)$, is the set of all $y \in Y$ such that there exists $x \in X$ with $T(x) = y$.
- The cokernel of T is $\mathrm{coker}(T) = Y/\mathrm{Ran}(T)$.
- The coimage of T is $\mathrm{coim}(T) = X/\ker(T)$.

Note that, by an analogous version of the first isomorphism theorem, we have $\mathrm{coim}(T) \cong \mathrm{Ran}(T)$.

Additionally, T is injective if and only if $\ker(T) = \{0\}$, and T is surjective if and only if $\mathrm{coker}(T) = \{0\}$.

We write $\dim(\operatorname{Ran}(T)) = \operatorname{rank}(T)$, and $\dim(\ker(T)) = \operatorname{null}(T)$.

Definition A.2.8. Let $T \in \mathcal{L}(X)$, and let $\lambda \in \mathbb{C}$. The eigenspace for λ is the subspace $E_\lambda(T) = \ker(T - \lambda I)$.

If $E_\lambda(T) \neq \{0\}$, then λ is called an eigenvalue for T The nonzero vectors in $E_\lambda(T)$ are called eigenvectors for T.

The set

$$\sigma_p(T) = \left\{\lambda \,\middle|\, \lambda \text{ is an eigenvalue for } T\right\}$$

is known as the point spectrum of T.

Theorem A.2.1 (Rank–Nullity)**.** If $T\colon X \to Y$ is a linear map between vector spaces, then $\operatorname{rank}(T)+\operatorname{null}(T) = \dim(Y)$.

The separation properties of linear functionals are used heavily in proofs of various results in analysis. The following theorem is refined via the Hahn–Banach theorems (see D.2.7), as in infinite dimensions, continuity becomes an issue that analysts are forced to deal with. However, we start with the algebraic case.

Proposition A.2.1. Let X be a vector space. If $0 \neq x_0 \in X$, then there is a $\varphi \in X'$ such that $\varphi(x_0) \neq 0$.

Geometrically, linear functionals are tied to hyperplanes within vector spaces.

Proposition A.2.2. Let X be a vector space with $\dim(X) \geqslant 2$, and let $H \subseteq X$ be a subspace. The following are equivalent:

(i) $H = \ker(\varphi)$ for some nonzero $\varphi \in X'$;

(ii) $H \subseteq X$ is a maximal *proper* subspace;

(iii) $\dim(X/H) = 1$ (i.e., H has codimension 1).

A subspace that satisfies any of these equivalent properties is called a hyperplane.

If $U = H + x_0$ for some fixed x_0, then U is known as an affine hyperplane.

A.3 Algebras

In our definition of vector spaces, we stated that they are akin to abelian groups, equipped with an operation of scalar multiplication. We may extend the analogy towards "rings" that include scalar multiplication, which are known as algebras.

A.3.1 The Structure of Algebras

Definition A.3.1. Let A be a $\mathbb{C}$-vector space. We say A is an algebra if A admits a multiplication, $(a, b) \mapsto a \cdot b$, that satisfies, for all $a, b, c \in A$ and $\alpha \in \mathbb{C}$,

- $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;
- $a \cdot (b + c) = a \cdot b + a \cdot c$;
- $(a + b) \cdot c = a \cdot c + b \cdot c$;
- $(\alpha a) \cdot b = \alpha(a \cdot b) = a \cdot (\alpha b)$.

If A, considered as a ring, is also unital, then we say A is a unital algebra. If $a \cdot b = b \cdot a$, then we say A is commutative.

If A also admits a unary operation $*\colon A \to A$ that satisfies, for all $a, b \in A$ and $\alpha \in \mathbb{C}$,

- $a^{**} = a$;
- $(ab)^* = b^* a^*$;
- $(\alpha a + b)^* = \overline{\alpha} a^* + b^*$,

then we say A is a $*$-algebra.

In an algebra, the ideal structure requires compatibility with the underlying field, and in a $*$-algebra, we may further specify compatibility with the star structure.

Definition A.3.2. Let A be an algebra, and let $J \subseteq A$. Then,

- we say J is a subalgebra of A if J is a linear subspace of A that is closed under multiplication of elements in J;
- we say J is an ideal of A if J is a subalgebra of A that is closed under multiplication by elements in A.

If A is a $*$-algebra, then

- we say J is $*$-closed if for any $t \in J$, $t^* \in J$;
- if J is a $*$-closed subalgebra, then we say J is a $*$-subalgebra;
- if J is a $*$-closed ideal, then we say J is a $*$-ideal.

There are a variety of important, named elements in any $*$-algebra. Most of these elements inherit their name from the fact that they are abstractions of elements of spaces of bounded linear operators (see D.5.5). However, despite their name and, their definitions are purely algebraic in nature.

Definition A.3.3. Let A be a $*$-algebra. Then,

- we say $e \in A$ is an idempotent if $e^2 = e$;
- we say $x \in A$ is invertible if there exists a unique $y \in A$ such that $xy = yx = 1_A$ — we write
$$\mathrm{GL}(A) := \{a \in A \mid a \text{ is invertible}\}$$
for the group of invertible elements in A;
- an element $x \in A$ is called Hermitian (or self-adjoint) if $x = x^*$ — we write
$$A_{\text{s.a.}} := \{x \in A \mid x = x^*\}$$
for the set of self-adjoint elements in A;
- an element $a \in A$ is called positive if there exists $b \in A$ such that $a = b^* b$ — we write
$$A_+ := \left\{a \in A \,\middle|\, a \text{ is positive}\right\}$$
for the set of positive elements in A;
- an element $p \in P$ is called a projection if it is self-adjoint and idempotent — we write
$$\mathcal{P}(A) := \left\{p \in A \,\middle|\, p = p^* = p^2\right\}$$
for the set of projections in A;
- if A is unital, then an element $u \in A$ is called unitary if $u^* u = uu^* = 1_A$ — we write
$$\mathcal{U}(A) = \{u \in A \mid u^* u = uu^* = 1_A\}$$
for the set of unitary elements in A;

- an element $z \in A$ is called normal if $z^*z = zz^*$ — we write $\mathrm{Nor}(A)$ for the set of normal elements in A.

Fact A.3.1. The following inclusions hold:

$$\mathcal{P}(A) \subseteq A_+ \subseteq A_{\text{s.a.}} \subseteq \mathrm{Nor}(A),$$

and

$$\mathcal{U}(A) \subseteq \mathrm{Nor}(A).$$

Furthermore, $\mathrm{span}(A_{\text{s.a.}}) = A$, where the self-adjoint elements

$$h = \frac{1}{2}(x + x^*)$$
$$k = \frac{i}{2}(x^* - x)$$

form the Cartesian decomposition $x = h + ik$.

A.3.2 Algebra Homomorphisms

Just as there are group homomorphisms, ring homomorphisms, and linear maps, there are also algebra homomorphisms.

Definition A.3.4. Let A and B be algebras over $\mathbb{C}$.

(1) An algebra homomorphism between A and B is a linear map $\varphi\colon A \to B$ that is also a ring homomorphism — i.e., $\varphi(ab) = \varphi(a)\varphi(b)$.

(2) If φ is an algebra homomorphism that is also bijective, then φ is called an algebra isomorphism. An automorphism is an algebra isomorphism $\alpha\colon A \to A$. We write

$$\mathrm{Aut}(A) := \left\{\alpha \,\middle|\, \alpha\colon A \to A \text{ is an automorphism}\right\}.$$

(3) A character on A is a nonzero homomorphism $h\colon A \to \mathbb{C}$. We write

$$\Omega(A) := \{h \mid h\colon A \to \mathbb{C} \text{ is a character}\}$$

to denote the character space of A.

(4) If A and B are $*$-algebras, then an algebra homomorphism $\varphi\colon A \to B$ that satisfies $\varphi(a^*) = \varphi(a)^*$ (known as $*$-preserving) is known as a $*$-homomorphism.

(5) If φ is a bijective $*$-homomorphism, then we say φ is a $*$-isomorphism.

(6) If A and B are $*$-algebras, a linear map $\phi\colon A \to B$ is said to be positive if $\phi(A_+) \subseteq B_+$. A positive linear map $\phi\colon A \to B$ between $*$-algebras is called faithful if $\ker(\phi) \cap A_+ = \{0\}$ — i.e., ϕ is faithful if it is injective on the positive elements.

One of the benefits of working with $*$-algebras (as opposed to algebraic objects with less structure) is that, even if our $*$-algebras aren't unital, we can extend them to contain units, and specifically in an "essential" manner. We formalize this below.

Definition A.3.5. Let A be an algebra. An ideal $I \subseteq A$ is said to be essential if, for any other ideal $J \subseteq A$, $J \cap I \neq \{0\}$.

In other words, essential ideals are "big" in the ideal structure of an algebra.

Theorem A.3.1 (Existence of a Unitization)**.** Let A be a nonunital ($*$-) algebra. Then, there exists a unital algebra $\widetilde{A}$ and an injection $\iota\colon A \hookrightarrow \widetilde{A}$ such that $\iota(A) \subseteq \widetilde{A}$ is an essential ideal.

Furthermore, if $\varphi\colon A \to B$ is a ($*$-) homomorphism between ($*$-) algebras, then φ extends to a unital ($*$-) homomorphism $\widetilde{\varphi}\colon \widetilde{A} \to \widetilde{B}$. If B is unital, then there exists a unital ($*$-) homomorphism $\overline{\varphi}\colon \widetilde{A} \to B$ that extends φ.

Similarly, if $h\colon A \to \Omega$ is a character, then h extends to a character $\widetilde{h}\colon \widetilde{A} \to \mathbb{C}$ on $\widetilde{A}$ that extends h.

Appendix B

Point-Set Topology

B.1 Ordering, the Axiom of Choice, and Zorn's Lemma

Definition B.1.1 (Preorders, Partial Orders, Total Orders, and Well-Orders)**.** Let X be a set, and $\preceq$ be a relation on X. We say a relation is a preorder if it is reflexive and transitive:

- $a \preceq a$
- $a \preceq b \wedge b \leqslant c \Rightarrow a \preceq c$.

We say X is a directed set if, for any $a, b \in X$, there is $c \in X$ such that $a \preceq c$ and $b \preceq c$.

If $\preceq$ is also antisymmetric — that is, $a \preceq b \wedge b \preceq a \Rightarrow a = b$ — then, we say $\preceq$ is a partial order.

We say $m \in X$ is a maximal element if, for any $x \in X$ with $m \preceq x$, $m = x$.

If X is partially ordered by $\preceq$ and, for any two elements $a, b \in X$, either $a \preceq b$ or $b \preceq a$, then we say $\preceq$ is a total order on X.

If X is a totally ordered set that has the property that, for any nonempty $A \subseteq X$, there is some $x \in A$ such that for any $y \in A$, $x \prec y$ for all $y \in A$ with $y \neq x$, then we say $\preceq$ is a well-order on X.

Example B.1.1.

- The set $\mathbb{N}$ with the usual ordering is a well-ordered set.
- If A is a set, then $P(A)$ with the ordering $A \preceq B$ if $A \supseteq B$ is a partially ordered set. This is known as the containment ordering.
- Similarly, if A is a set, then $P(A)$ with the ordering $A \preceq B$ if $A \subseteq B$ is a partially ordered set. This is known as the inclusion ordering.
- A collection of functions $\{\varphi_i \colon Z_i \to Y\}_{i \in I}$ ordered by $\varphi_i \preceq \varphi_j$ if $Z_i \subseteq Z_j$ and $\varphi_j|_{Z_i} = \varphi_i$, is a partially ordered set. This is often known as the extension ordering.

Example B.1.2. If V is a vector space, and $\leqslant$ is a partial order on the vector space that satisfies, for all $u, v, w \in V$ and $t \geqslant 0$ (specifically, $t \in \mathbb{R}^+ \subseteq \mathbb{C}$),

- for $u \leqslant v$, $u + w \leqslant v + w$,
- for $u \leqslant v$, $tu \leqslant tv$,

then we say V is an ordered vector space.

If V is an ordered vector space, then a cone in V is a subset $C \subseteq V$ that is "closed upwards," in the sense that if $x, y \in C$, then $x + y \in C$, and if $x \in C$ and $t \geqslant 0$, $tx \in C$, and that $C \cap (-C) = \{0\}$.

The set of all elements $v \in V$ such that $v \geqslant 0$ is known as the cone of positive elements, and denoted V_+.

The axiom of choice, stated below, is a load-bearing part of topology and analysis.

Definition B.1.2 (Axiom of Choice)**.** Let $\mathcal{A} = \{A_i\}_{i \in I}$ be an indexed collection of sets. There exists an indexed set $\{x_i\}_{i \in I}$ such that $x_i \in S_i$ for each $i \in I$.

Using some of the basic results in order theory, we may state an equivalent formulation of the axiom of choice that is also more readily used to prove important results in analysis.

Theorem B.1.1 (Zorn's Lemma)**.** If $(X, \preceq)$ is a nonempty partially ordered set with the property that for all $C \subseteq X$ with C totally ordered, C has an upper bound, then X has a maximal element.

Zorn's lemma can be used to prove the following theorems.

Example B.1.3.

- Every $\mathbb{F}$-vector space V has a basis $B \subseteq V$ such that the set of all finite linear combinations of elements of B over $\mathbb{F}$ is V.
- If φ is a continuous linear functional defined on a subspace $W \subseteq V$, there is an extension Φ such that $\Phi|_W = \varphi$. This is one of the Hahn–Banach theorems.
- The arbitrary product of compact spaces is compact. This is known as Tychonoff's Theorem (Theorem B.3.3).

B.2 Metric Spaces

Building upon the basics of sets and orders, we move towards understanding metric spaces.

B.2.1 Basics of Metric Spaces

Definition B.2.1 (Metrics)**.** Let X be a set. A distance metric is a function

$$d \colon X \times X \to [0, \infty)$$

such that the following three properties are satisfied:

- if $x, y \in X$ and $d(x, y) = 0$, then $x = y$;
- $d(x, y) = d(y, x)$ for all $x, y \in X$;
- $d(x, z) \leqslant d(x, y) + d(y, z)$ for all $x, y, z \in X$.

A function that satisfies the latter two properties is called a semimetric.

Two metrics d and ρ on X are equivalent if there exist constants $c_1, c_2 \geqslant 0$ such that

$$\begin{aligned} d(x, y) &\leqslant c_1 \rho(x, y) \\ \rho(x, y) &\leqslant c_2 d(x, y) \end{aligned}$$

for all $x, y \in X$.

A metric space is a pair (X, d), where d is a metric.

Example B.2.1 (Some Distance Metrics)**.**

- The discrete metric on any nonempty set is given by

$$d(x,y)\begin{cases}1 & x \neq y\\ 0 & x = y\end{cases}$$

- The Euclidean metric between $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ in $\mathbb{R}^n$ is

$$d_2(x,y) = \left(\sum_{j=1}^{n} \left|y_j - x_j\right|^2\right)^{1/2}.$$

- Other metrics on $\mathbb{R}^n$ include

$$d_1(x,y) = \sum_{j=1}^{n} \left|y_j - x_j\right|$$
$$d_\infty(x,y) = \max_{j=1}^{n} \left|y_j - x_j\right|.$$

  All of d_1, d_2, d_∞ are equivalent metrics.

- The Hamming distance between two strings of bits is

$$d_H \colon \{0,1\}^n \times \{0,1\}^n \to [0,\infty)$$
$$d_H\left(\left(x_j\right)_{j=1}^{n}, \left(y_j\right)_{j=1}^{n}\right) = \left|\{j \mid x_j \neq y_j\}\right|.$$

- The set $C([0,1],\mathbb{R})$ consisting of continuous real-valued functions from $[0,1]$ to $\mathbb{R}$ can be equipped with

$$d_u(f,g) = \sup_{t\in[0,1]} |f(t) - g(t)|,$$

  which is the uniform metric, or

$$d_1(f,g) = \int_0^1 |f(t) - g(t)|\, dt.$$

- All subsets of a metric space X equipped with the same metric is also a metric space.

- If ρ is a metric on X, then we can create a distance metric

$$d(x,y) = \frac{\rho(x,y)}{1+\rho(x,y)}$$

  that is bounded on $[0,1]$.

- If $d_1, \dots, d_n$ are metrics on X and $c_1, \dots, c_n > 0$ are constants, then

$$d(x,y) = \sum_{k=1}^{n} c_k d_k(x,y)$$

  defines a metric on X.

- If $(\rho_k)_k$ is a family of separating semimetrics for X — i.e., for $x, y \in X$ distinct, there is some ρ_j such that $\rho_j(x,y) \neq 0$ — then, we can obtain bounded semimetrics by taking

$$d_k(x,y) = \frac{\rho_k(x,y)}{1+\rho_k(x,y)}$$

for each k. From each d_k, we define

$$d(x,y) = \sum_{k=1}^{n} 2^{-k} d_k(x,y),$$

which is a metric on X.

- If $(X_k, \rho_k)_{k \geqslant 1}$ is a sequence of metric spaces, then we can form the product space

$$X = \prod_{k \geqslant 1} X_k$$

with the metric

$$D(f,g) = \sum_{k \geqslant 1} d_k(f(k), g(k)).$$

Here, $d_k = \frac{\rho_k}{1+\rho_k}$ is the corresponding bounded metric to ρ_k.

Definition B.2.2 (Open and Closed Sets)**.** Let (X, d) be a metric space.

(1) For $x \in X$ and $\delta > 0$, we define

(a) the open ball at x with radius $\delta > 0$

$$U(x,\delta) = \{y \in X \mid d(y,x) < \delta\};$$

(b) the closed ball centered at x with radius $\delta > 0$

$$B(x,\delta) = \{y \in X \mid d(y,x) \leqslant \delta\};$$

(c) the sphere centered at x with radius $\delta > 0$

$$S(x,\delta) = \{y \in X \mid d(y,x) = \delta\}.$$

(2) A set $V \subseteq X$ is open if, for all $x \in V$, there is $\delta > 0$ such that $U(x,\delta) \subseteq V$.

A subset $C \subseteq X$ is closed if C^c is open.

(3) If $x \in V$ and $V \subseteq X$ is open, then we say V is an open neighborhood of x. A neighborhood of x is any subset $N \subseteq X$ such that N contains an open neighborhood of x.

(4) If $A \subseteq X$ is any subset, the interior of A is

$$A^\circ := \bigcup \{V \mid V \text{ is open}, V \subseteq A\},$$

the closure of A is

$$\overline{A} = \bigcap \{C \mid C \text{ is closed}, A \subseteq C\},$$

and the boundary of A is

$$\partial A = \overline{A} \setminus A^\circ.$$

We can now talk about the topology of the metric space.

Fact B.2.1. Let (X, d) be a metric space, and let

$$\mathcal{U} = \{V \mid V \subseteq X \text{ open}\}.$$

Then, the following are true.

- $\emptyset \in \mathcal{U}, X \in \mathcal{U}$.
- If $\{V_i\}_{i \in I}$ is a family of open sets, then $\bigcup_{i \in I} V_i \in \mathcal{U}$.
- If $\{V_i\}_{i=1}^{n}$ is a finite collection of open sets, then $\bigcap_{i=1}^{n} V_i \in \mathcal{U}$.

Definition B.2.3. Let (X, d) be a metric space. Suppose $A \subseteq X$ is a nonempty subset.

(1) The distance from a point $x \in X$ to the set A is defined by

$$\operatorname{dist}_A(x) = \inf_{a \in A} d(x, a).$$

(2) The diameter of A is defined by

$$\operatorname{diam}(A) = \sup_{x,y \in A} d(x, y).$$

(3) If $\operatorname{diam}(A) < \infty$, then we say A is bounded.

(4) If, for every $\delta > 0$, there is a finite subset $F_\delta \subseteq X$ such that

$$A \subseteq \bigcup_{x \in F_\delta} U(x, \delta).$$

(5) For $A, B \subseteq X$, we define the Hausdorff distance between A and B to be

$$d_H(A, B) = \max\left\{\sup_{x \in A} \operatorname{dist}_B(x), \sup_{y \in B} \operatorname{dist}_A(y)\right\}.$$

Example B.2.2. Let Ω be a nonempty set, and (X, d) be a metric space. A function $f\colon \Omega \to X$ is said to be bounded if $\operatorname{diam}(\operatorname{Ran}(f)) < \infty$.

The collection $\operatorname{Bd}(\Omega, X)$ denotes all bounded functions with domain Ω and codomain X.

On $\operatorname{Bd}(\Omega, X)$, we define the uniform metric by

$$D_u(f, g) = \sup_{x \in \Omega} d(f(x), g(x)).$$

B.2.2 Convergence and Continuity in Metric Spaces

Definition B.2.4. Let (X, d) be a metric space.

(1) A sequence in X is a map $x\colon \mathbb{N} \to X$, which we call $(x_n)_n$ or $(x_n)_{n\geqslant 1}$.

(2) A natural sequence is a strictly increasing sequence of natural numbers $(n_k)_{k\geqslant 1}$ with $n_k \geqslant k$ and $n_k < n_{k+1}$.

(3) If $(n_k)_k$ is a natural sequence, the sequence $(x_{n_k})_k$ is called a subsequence of $(x_n)_n$.

(4) We say $(x_n)_n \to x$ if $d(x_n, x)_n \xrightarrow{n\to\infty} 0$. We say x is the limit of $(x_n)_n$.

Example B.2.3.

- If Ω is a nonempty set, and (X, d) is a metric space, the sequence of functions $f_n\colon \Omega \to X$ is said to converge pointwise to $f\colon \Omega \to X$ if
$$f_n(x) \xrightarrow{n\to\infty} f(x)$$
for each $x \in \Omega$.

- If $(f_n)_n \in \mathrm{Bd}(\Omega, X)$ is a sequence, we say $(f_n)_n \to f$ converges uniformly if
$$D_u(f_n, f) \xrightarrow{n\to\infty} 0,$$
or, equivalently,
$$\sup_{x\in\Omega} d(f_n(x), f(x)) \xrightarrow{n\to\infty} 0.$$

Definition B.2.5 (Sequential Criteria for Closure)**.** If (X, d) is a metric space, and $E \subseteq X$ is nonempty, then E is closed if and only if, for all $(x_n)_n \to x$ with $x_n \in E$, $x \in E$.

If $E \subseteq X$ is any nonempty set, then $\overline{E}$ is precisely the set of all $x \in X$ such that $(x_n)_n \to x$ for some $(x_n)_n \subseteq E$.

Definition B.2.6 (Completeness)**.** Let (X, d) be a metric space.

- If $(x_n)_n$ is a sequence in X such that for all $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that for all $m, n \geqslant N$, $d(x_m, x_n) < \varepsilon$, then we say the sequence is called Cauchy.
- If, for any $(x_n)_n$ Cauchy, $(x_n)_n \to x$ in X, then we say X is complete.
- If (X, d) is complete, then for any $A \subseteq X$ closed, A is also complete.
- If $A \subseteq X$ is complete as a metric space, then A is closed.

Example B.2.4. The metric space $\mathbb{Q}$ with the metric inherited from $\mathbb{R}$ is not complete. For instance, there is a sequence of rational numbers $(2, 2.7, 2.71, 2.718, \dots)$ converging to e, but $e \notin \mathbb{Q}$.

The space $\mathrm{Bd}(\Omega, X)$ is complete if X is complete.

Definition B.2.7 (Continuity)**.**

- Let (X, d) and (Y, ρ) be metric spaces, and let $f\colon X \to Y$ be a function. We say f is continuous at x if, for every $\varepsilon > 0$, there is $\delta > 0$ such that $z \in U(x, \delta) \Rightarrow \rho(f(x), f(z)) < \varepsilon$.
- If f is continuous at every point in X, then we say f is continuous.
- If f is bijective, continuous, and f^{-1} is continuous, then we say f is a homeomorphism.
- We say f is uniformly continuous on X if, for any $\varepsilon > 0$, there is $\delta > 0$ such that for any $y, z \in X$, $d(y, z) < \delta \Rightarrow \rho(f(y), f(z)) < \varepsilon$.

- We say f is Lipschitz if there exists $C > 0$ such that $d(x, y) \leqslant Cd(f(x), f(y))$ for all $x, y \in X$.
- We say f is an isometry if $d(x, y) = d(f(x), f(y))$ for all $x, y \in X$.

Fact B.2.2. Let $f \colon X \to Y$ be a map between metric spaces. The following are equivalent:

(i) f is continuous;

(ii) if $V \subseteq Y$ is open, then $f^{-1}(V) \subseteq X$ is open;

(iii) if $(x_n)_n \to x$ in X, then $(f(x_n))_n \to f(x)$ in Y.

Fact B.2.3. If M and N are metric spaces with N complete, and $A \subseteq M$ is dense, then if $f \colon A \to N$ is uniformly continuous, then there is a unique uniformly continuous map $\tilde{f} \colon M \to N$.

Definition B.2.8. Let (X, d) and (Y, ρ) be metric spaces.

(1) We say X and Y are homeomorphic if there is a homeomorphism $f \colon X \to Y$.

(2) We say X and Y are uniformly isomorphic if there is a uniformly continuous bijection $f \colon X \to Y$ with f^{-1} uniformly continuous. Such an f is called a metric space uniformism.

(3) We say X and Y are isometrically isomorphic if there is a bijective isometry $f \colon M \to N$.

Fact B.2.4. If X and Y are uniformly isomorphic metric spaces with X complete, then so too is Y.

If d and ρ are equivalent metrics on a set X, then the identity map

$$\mathrm{id}_X : (X, \rho) \to (X, d)$$

is a metric space uniformism.

B.3 Topological Spaces

We can now move from metric spaces to the more general setting of topological spaces. This will enable us to understand certain properties (like openness, continuity, etc.) separate from the metric structure (or lack thereof) that a certain set is endowed.

B.3.1 Definitions

Definition B.3.1. Let X be a nonempty set. A topology on X is a family of subsets τ satisfying

(1) $\emptyset \in \tau, X \in \tau$;

(2) if $\{V_i\}_{i \in I} \subseteq \tau$, then $\bigcup_{i \in I} V_i \in \tau$;

(3) if $\{V_i\}_{i=1}^{n} \subseteq \tau$, then $\bigcap_{i=1}^{n} V_i \in \tau$.

If τ is a topology on X, then (X, τ) is called a topological space. We call members of τ open sets.

If $C \subseteq X$ and $C^c \in \tau$, then C is called.

If E is closed and open, it is called clopen.

A countable union of closed sets is called an F_σ set, and a countable intersection of open sets is called a G_δ set.

Definition B.3.2. If X is a nonempty set, then the definition $\tau = P(X)$ is known as the discrete topology.

If X is a nonempty set, and $\tau = \{X, \emptyset\}$, then we call τ the indiscrete topology.

Definition B.3.3. Let X be a nonempty set. Suppose $\tau_1, \tau_2 \subseteq P(X)$ are two topologies on X. If $\tau_1 \subseteq \tau_2$, then we say τ_1 is weaker (or coarser) than τ_2. We say τ_2 is stronger (or finer) than τ_1.

Definition B.3.4. Let X be a nonempty set, and suppose $\mathcal{E} \subseteq P(X)$ is a family of subsets. We define the topology on X generated by $\mathcal{E}$ to be

$$\tau(\mathcal{E}) = \bigcap \left\{ \tau \;\middle|\; \tau \text{ is a topology on } X, \mathcal{E} \subseteq \tau \right\}.$$

In other words, $\tau(\mathcal{E})$ is the weakest topology that contains the family $\mathcal{E}$.

Definition B.3.5. Let (X, τ) be a topological space. If $Y \subseteq X$ is a subset, then the subspace topology on Y is defined by

$$\tau_Y = \{V \cap Y \mid V \in \tau\}.$$

Definition B.3.6. Let (X, τ) be a topological space, and let $A \subseteq X$ be a subset.

(1) The interior of A is the open set $A^\circ = \bigcup\{V \mid V \in \tau,\ V \subseteq A\}$.

(2) The closure of A is the closed set $\overline{A} = \bigcap\{C \mid C \text{ closed}, A \subseteq C\}$.

(3) We say A is dense if $\overline{A} = X$.

(4) We say A is nowhere dense if $\left(\overline{A}\right)^\circ = \emptyset$.

If X admits a countable dense subset, then we say X is separable.

If X is the countable union of nowhere dense subsets, then we say X is meager.

Remark B.3.1. A set A is dense if and only if, for any $U \in \tau$ with $U \neq \emptyset$, it is the case that $A \cap U \neq \emptyset$.

Fact B.3.1. If (M, d) is a separable metric space, and $E \subseteq M$ is a subset, then E with the subspace topology is also separable.

Definition B.3.7. Let (X, τ) be a topological space.

- An open neighborhood of x_0 is an open set $V \in \tau$ with $x_0 \in V$. We write

$$\mathcal{O}_{x_0} = \{V \mid V \in \tau, x_0 \in V\}$$

  to denote the family of all open neighborhoods of x_0.

- If $N \subseteq X$ is a subset with $x_0 \in V \subseteq N$, where $V \in \mathcal{O}_{x_0}$, then we say N is a neighborhood of x_0. We write $\mathcal{N}_{x_0}$ to be the collection of neighborhoods of x_0.

- A neighborhood base for τ at x_0 is a family $\mathcal{O} \subseteq \mathcal{O}_{x_0}$ with such that for all $U \in \mathcal{O}_{x_0}$, there is $V \in \mathcal{O}$ with $V \subseteq U$.

- We say (X, τ) is first countable if every $x \in X$ admits a countable neighborhood base.

- A base for τ is a family $\mathcal{B} \subseteq \tau$ that contains a neighborhood base for τ at x_0 For each $x_0 \in X$.

- We say (X, τ) is second countable if it admits a countable base.

Fact B.3.2. If $\mathcal{B}$ is a base for τ, then every $U \in \tau$ can be written as a union $U = \bigcup_{i \in I} B_i$, where $B_i \in \mathcal{B}$.

Fact B.3.3. All metric spaces are first-countable, with a neighborhood base of

$$\mathcal{O}_{x_0} = \{U(x_0, 1/n) \mid n \in \mathbb{N}\}$$

for each $x_0 \in X$.

A metric space (X, d) is second countable if and only if it is separable.

Fact B.3.4. If X is a topological space, and $x_0 \in X$ has a countable neighborhood base, then there is a neighborhood base $(V_n)_{n\geqslant 1}$ with $V_1 \supseteq V_2 \supseteq \cdots$.

B.3.2 Continuity in Topological Spaces

Definition B.3.8. Let (X, τ) and (Y, σ) be topological spaces, and let $f\colon X \to Y$ be a map.

(1) We say f is continuous at $x_0 \in X$ if, for every $U \in \mathcal{O}_{f(x_0)}$, there is $V \in \mathcal{O}_x$ with $f(V) \subseteq U$.

(2) We say f is continuous if f is continuous at every point in X.

(3) We say f is a homeomorphism if f is a continuous bijection with a continuous inverse.

(4) We say f is an open map if $U \in \tau$ implies $f(U) \in \sigma$. Similarly, we say f is a closed map if $C \subseteq X$ closed implies $f(C) \subseteq Y$ is closed.

(5) We say f is a quotient map if f is surjective with $V \subseteq Y$ open if and only if $f^{-1}(V) \subseteq X$ open.

(6) We say f is an embedding if $f\colon X \to \operatorname{Ran}(f)$ is a homeomorphism, where $\operatorname{Ran}(f)$ is endowed with the subspace topology.

(7) We write $C(X, Y)$ to be the continuous functions from X to Y. If $Y = \mathbb{C}$ with the regular topology, then we write $C(X)$.

Fact B.3.5. A function $f\colon X \to Y$ is continuous if and only if $f^{-1}(U) \subseteq X$ is open for every open $U \subseteq Y$. Equivalently, f is continuous if and only if $f^{-1}(C) \subseteq X$ is closed for every closed $C \subseteq Y$.

Definition B.3.9 (Separation Axioms)**.** Let (X, τ) be a topological space.

- We say X is T1 if $\{x\}$ is closed for every $x \in X$.
- We say X is T2 (or Hausdorff) if, for every $x, y \in X$ with $x \neq y$, there are $U, V \in \tau$ with $x \in U$, $y \in V$, and $U \cap V = \emptyset$.
- We say X is T3 if, for every $x \in X$ and $B \subseteq X$ closed with $x \notin B$, there are $U, V \in \tau$ with $x \in U$, $B \subseteq V$, and $U \cap V = \emptyset$. If X is T1 and T3, we say X is regular.
- We say X is T3.5 if, for every $x_0 \in X$ and closed $B \subseteq X$ with $x_0 \notin B$, there is a continuous function $f\colon X \to [0, 1]$ with $f(x_0) = 0$ and $f(B) = 1$. If X is T1 and T3.5, we say X is completely regular.
- We say X is T4 if, for every pair of closed subsets $A, B \subseteq X$ with $A \cap B = \emptyset$, there are $U, V \in \tau$ with $A \subseteq U$, $B \subseteq V$, and $U \cap V = \emptyset$. If X is T1 and T4, then we say X is normal.

Just as we defined completely regular spaces through the existence of certain continuous functions that act to separate points, we can completely classify normality through a separating family of continuous functions.

Theorem B.3.1 (Urysohn's Lemma)**.** Let (X, τ) be a topological space. It is the case that X is normal if and only if for every pair of disjoint closed subsets $A, B \subseteq X$, there is a continuous function $f\colon X \to [0, 1]$ with $f(A) = 0$ and $f(B) = 1$.

Remark B.3.2. Metric spaces are an example of normal spaces.

B.3.3 Initial and Final Topologies

Definition B.3.10. Let X be a set, and suppose $\{(Y_i, \tau_i)\}_{i \in I}$ is a family of topological spaces. Let $\{f_i : X \to Y_i\}$ be a family of maps. Setting

$$\varepsilon = \left\{ f_i^{-1}(V) \,\middle|\, V_i \in \tau_i \right\},$$

and letting $\tau = \tau(\varepsilon)$ be the topology on X generated by ε, we say τ is the initial topology on X induced by the maps $\{f_i\}_{i \in I}$.

Specifically, τ is the weakest topology on X such that each f_i is continuous.

Definition B.3.11 (Product Topology)**.** Let $\{(X_i, \tau_i)\}_{i \in I}$ be a family of topological spaces. The topology on the product $\prod_{i \in I} X_i$ is defined to be the initial topology induced by the family of projection maps,

$$\pi_j : \prod_{i \in I} X_i \to X_j,$$

defined by $\pi_j((x_i)_{i \in I}) = x_j$.

For each $U \subseteq X_i$ open, we have $\pi_j^{-1}(U) = \prod_{i \in I} U_I$, where $U_i = X_i$ for $i \neq j$, and $U_j = U$. A base for this topology is the collection

$$\mathcal{B} = \left\{ \prod_{i \in I} U_i \,\middle|\, U_i = X_i \text{ for all but finitely many open } U_i \subseteq X_i \right\}.$$

If we consider $X_i = X$ for all i, there is a bijection between $X^I := \{f \mid f : I \to X\}$, the set of all functions from I to X, and $\prod_{i \in I} X_i$, with the map $f \mapsto (f(i))_{i \in I}$. The product topology on X^I coincides with the topology of pointwise convergence.

Definition B.3.12 (Final Topology)**.** Let (X, τ) be a topological space, Y a nonempty set, and suppose $q : X \to Y$ is a surjection. Then, the collection

$$\tau_q := \left\{ V \subseteq Y \,\middle|\, q^{-1}(V) \in \tau \right\}$$

is what is known as the final (or quotient) topology on Y produced by q.

B.3.4 Convergence in Topological Spaces

Given a non-first-countable space X and a subset $A \subseteq X$, it is not necessarily the case that $x \in \overline{A}$ is the limit of a sequence $(x_n)_n$. However, we know that the sequential characterization of properties like closure, compactness (which will be covered in an upcoming section), and continuity is useful, so we want to generalize these ideas to non-first-countable spaces. This is where we can use nets.

Definition B.3.13 (Nets)**.** A net is a map $A \to X$, where $\alpha \mapsto x_\alpha$, where A is a directed set. We write nets as $(x_\alpha)_\alpha$.

Example B.3.1 (Some Directed Sets)**.**

(1) The natural numbers, $\mathbb{N}$, or the real numbers, $\mathbb{R}$, equipped with their usual ordering, are examples of directed sets. Every totally ordered set is directed.

(2) If S is any set, the collection $F(S)$ consisting of all finite subsets of S is directed by inclusion.

(3) The collection of finite partitions over a closed and bounded interval, $\mathcal{P}([a, b])$ is by the partition

norm. If $P = \left\{x_j\right\}_{j=0}^n$ and $Q = \left\{y_j\right\}_{j=0}^m$ are partitions, we define

$$\|P\| = \max_{1\leqslant j\leqslant n}\left|x_j - x_{j-1}\right|$$
$$\|Q\| = \max_{1\leqslant j\leqslant m}\left|y_j - y_{j-1}\right|,$$

and the preorder that $P \leqslant Q$ if and only if $\|P\| \geqslant \|Q\|$. In other words, we say $P \leqslant Q$ if Q is finer than P.

For any partitions P and Q, their common refinement is a supremum for both — $P \vee Q \geqslant P, Q$ for each partition.[a]

(4) Let (X, τ) be a topological space, and for every x, we order the $\mathcal{O}_x$ by containment. That is, for elements $U, V \in \mathcal{O}_x$, we set $U \leqslant V$ if and only if $U \supseteq V$. This is a directed ordering, as we can always take $U \cap V \subseteq U, V$ (since both U and V contain x). Similarly, the neighborhood system at x, $\mathcal{N}_x$, is also directed by containment.

(5) If A and B are directed sets, then $A \times B$ with the Cartesian ordering — $(\alpha_1, \beta 1) \leqslant (\alpha_2, \beta 2)$ if and only if $\alpha_1 \leqslant \alpha_2$ and $\beta_1 \leqslant \beta_2$ — is also a directed set.

[a] This is extremely useful in defining the Riemann integral.

Example B.3.2 (Some Nets)**.**

(1) Any sequence $(x_k)_{k\in\mathbb{N}}$ is a net.

(2) Let $F(\Omega)$ be the set of all finite subsets of Ω directed by inclusion. Let $f\colon \Omega \to \mathbb{C}$ be a map. Then, we have a net $(s_F)_{F\in F(\Omega)}$ defined by

$$s_F = \sum_{x\in F} f(x).$$

(3) Consider the collection of partitions $\mathcal{P}([a, b])$ directed by the partition norm. For a bounded function $f\colon [a, b] \to \mathbb{R}$ and a partition $P = \left\{x_j\right\}_{j=0}^n$, for each j we set

$$M_j(P) = \sup_{t\in\left[x_j, x_{j-1}\right]} f(t)$$
$$m_j(P) = \inf_{t\in\left[x_j, x_{j-1}\right]} f(t).$$

We obtain two nets, $U, L\colon \mathcal{P}([a, b]) \to \mathbb{R}$, defined by

$$U(P) = \sum_{j=1}^n M_j(P)\left(x_j - x_{j-1}\right)$$
$$L(P) = \sum_{j=1}^n m_j(P)\left(x_j - x_{j-1}\right).$$

These are known as the upper and lower Darboux sums.

Definition B.3.14. Let (X, τ) be a topological space, and let $(x_\alpha)_\alpha$ be a net in X.

(1) For a set $U \subseteq X$, we say $(x_\alpha)_\alpha$ is eventually in U if there is $\alpha_0 \in A$ such that $x_\alpha \in U$ for all $\alpha \geqslant \alpha_0$.

(2) We say the net $(x_\alpha)_\alpha$ converges to $x \in X$ if, for every $U \in \mathcal{O}_x$, $(x_\alpha)_\alpha$ is eventually in U. We write $(x_\alpha)_\alpha \xrightarrow{\tau} x$, though if the topology is clear from context the τ is not written.

(3) For a given $U \subseteq X$, we say $(x_\alpha)_\alpha$ is frequently in U if for every $\beta \in A$, there is $\alpha \in A$ with $\alpha \geqslant \beta$ and $x_\alpha \in U$.

(4) A point $x \in X$ is known as a cluster point of the net $(x_\alpha)_\alpha$ if for every $U \in \mathcal{O}_x$, $(x_\alpha)_\alpha$ is frequently in U. That is, for all $U \in \mathcal{O}_x$ and for all $\beta \in A$, there exists $\alpha \in A$ with $\alpha \geqslant \beta$ and $x_\alpha \in U$.

Fact B.3.6 (Characterizations Using Nets)**.** Let (X, τ) and (Y, σ) be topological spaces, $E \subseteq X$ a subset, and $f\colon X \to Y$ a map.

- It is the case that $x \in \overline{E}$ if and only if there is a net $(x_\alpha)_\alpha$ in E with $(x_\alpha)_\alpha \to x$.
- A map f is continuous if and only if for every convergent net $(x_\alpha)_\alpha \xrightarrow{\tau} x$, we have $(f(x_\alpha))_\alpha \xrightarrow{\sigma} f(x)$.
- If X is given by the initial topology induced by the family of maps $\{f_i\colon X \to (Y_i, \tau_i)\}_{i\in I}$, the net $(x_\alpha)_\alpha$ converges to x if and only if $(f_i(x_\alpha))_\alpha \xrightarrow{\tau_i} f_i(x)$ in Y_i for each $i \in I$.
- If $\{(X_i, \tau_i)\}_{i\in I}$ is a family of topological spaces, with $X = \prod_{i\in I} X_i$ equipped with the product topology, then a net $(x_\alpha)_\alpha$ in X converges to $x \in X$ if and only if $(x_\alpha(i))_\alpha \xrightarrow{\tau_i} x(i)$ in X_i for each $i \in I$.

Definition B.3.15. Let A and B be directed sets.

(1) A subset $J \subseteq A$ is said to be cofinal if for every $\alpha \in A$, there is $\gamma \in J$ with $\gamma \geqslant \alpha$.

(2) A map $\sigma\colon B \to A$ is monotone increasing if it preserves the order of the sets — i.e., if $\beta_1 \leqslant \beta_2$ in B, then $\sigma(\beta_1) \leqslant \sigma(\beta_2)$ in A.

(3) A map $\sigma\colon B \to A$ is called natural if it is monotone increasing and $\sigma(B) \subseteq A$ is cofinal.

If X is a topological space, and $(x_\alpha)_\alpha$ is a net in X, a subnet, $\left(x_{\sigma(\beta)}\right)_{\beta\in B}$, is a net where B is a directed set and $\sigma\colon B \to A$ is a natural map. We will write $\alpha_\beta = \sigma(\beta)$.

Fact B.3.7. Let $(x_\alpha)_\alpha$ be a net in (X, τ). A point $x \in X$ is a cluster point of $(x_\alpha)_\alpha$ if and only if $(x_\alpha)_\alpha$ admits a subnet converging to x.

An important fact about nets is that they can be used to characterize the topology of certain spaces.

Fact B.3.8. Let (X, τ) be a completely regular space, and $(x_\alpha)_\alpha$ a net in X. Then, $(x_\alpha)_\alpha \to x$ if and only if $(f(x_\alpha))_\alpha \to f(x)$ for every $f \in C(X, [0, 1])$.

B.3.5 Compactness

Definition B.3.16. Let (X, τ) be a topological space. A subspace $K \subseteq X$ is called compact if, for every collection $\{U_i\}_{i\in I} \subseteq \tau$ with $K \subseteq \bigcup_{i\in I} U_i$, there exists a finite subset $F \subseteq I$ such that $K \subseteq \bigcup_{i\in F} U_i$.

Colloquially, we say that K is compact if every open cover of K admits a finite subcover.

Proposition B.3.1. Let (X, τ) be a topological space, and let $K \subseteq X$ be closed. The following are equivalent:

(i) K is compact;

(ii) for any family $\{C_i\}_{i\in I}$ of closed subsets in K with the finite intersection property — i.e., for all finite $F \subseteq I$, $\bigcap_{i\in F} C_i \neq \emptyset$ — it is the case that $\bigcap_{i\in I} C_i \neq \emptyset$;

(iii) every net $(x_\alpha)_\alpha$ in K has a cluster point in K;

(iv) every net $(x_\alpha)_\alpha$ in K admits a convergent subnet.

Fact B.3.9. Let X and Y be topological spaces. If $f\colon X \to Y$ is continuous, then for any compact $K \subseteq X$, $f(K) \subseteq Y$ is compact.

Fact B.3.10. If X and Y are topological spaces with X compact and Y Hausdorff, then any continuous bijection $f\colon X \to Y$ is a homeomorphism.

Fact B.3.11. Let M and N be metric spaces. If $f\colon M \to N$ is continuous, and M is compact, then f is uniformly continuous.

Fact B.3.12. If (M, d) is a metric space, the following are equivalent:

(i) M is compact;

(ii) M is sequentially compact — every sequence in M admits a convergent subsequence;

(iii) M is complete and totally bounded — for any $\varepsilon > 0$, there exists $\{x_1, \dots, x_n\} \subseteq M$ such that $M \subseteq \bigcup_{i=1}^{n} U(x_i, \varepsilon)$.

Definition B.3.17. Let Ω be a compact Hausdorff space. A subset $\mathcal{F} \subseteq C(\Omega)$ is

(a) pointwise bounded if, for all $x \in \Omega$, $\sup_{f\in\mathcal{F}}|f(x)| < \infty$;

(b) equicontinuous at $x \in \Omega$ if for every $\varepsilon > 0$, there is a neighborhood U_x of x such that $|f(y) - f(x)| < \varepsilon$ for all $f \in \mathcal{F}$ and all $y \in U_x$;

(c) equicontinuous if $\mathcal{F}$ is equicontinuous at every $x \in \Omega$.

Theorem B.3.2 (Arzela–Ascoli Theorem)**.** Let Ω be a compact Hausdorff space. If $\mathcal{F} \subseteq C(\Omega)$ is pointwise bounded and equicontinuous, then the uniform closure $\overline{\mathcal{F}}^{\|\cdot\|_u} \subseteq C(\Omega)$ is compact in $C(\Omega)$.

Theorem B.3.3 (Tychonoff's Theorem)**.** Let $(X_i, \tau_i)_{i\in I}$ be a family of compact topological spaces. Then, the product space $\prod_{i\in I} X_i$, equipped with the product topology, is also compact.

Definition B.3.18. A topological space (X, τ) is said to be locally compact if for any $x \in U \in \tau$, there is $V \in \tau$ with compact closure such that $x \in V \subseteq \overline{V} \subseteq U$.

If X is locally compact and Hausdorff, we say it is a LCH space.

Theorem B.3.4 (Urysohn's Lemma for LCH spaces)**.** Let X be a LCH space. Suppose $C, K \subseteq X$ are closed disjoint subsets with K compact. Then, there is a continuous function $f\colon X \to [0, 1]$ with $f|_C = 0$ and $f|_K = 1$.

If $K \subseteq U \subseteq X$ is such that K is compact and U is open, then there is a compactly supported continuous function $f\colon X \to [0, 1]$ such that $f|_K = 1$ and $\operatorname{supp}(f) \subseteq U$.

Definition B.3.19. Let X be a LCH space. A compactification of X is a pair (Z, ι), with

(i) Z compact and Hausdorff;

(ii) $\iota\colon X \to Z$ an embedding;

(iii) $\operatorname{Ran}(\iota) \subseteq Z$ dense.

Theorem B.3.5 (One-Point Compactification)**.** Let X be a noncompact LCH space. There is a compactification (X_α, ι) of X, with

(1) $X_\infty \setminus \iota(X)$ is exactly one point;

(2) for any other compactification of X, (Z, j), where $Z \setminus j(X)$ is one point, Z is homeomorphic to X_∞.

Appendix C

Measure Theory and Integration

In order to properly discuss amenability, we need a strong foundation in measure theory.

C.1 Constructing Measurable Spaces

Fix a set Ω. We let $\mathcal{A} = \{A_i\}_{i \in I}$ be a collection of subsets of Ω.

Definition C.1.1 (Algebra of Subsets)**.** The collection $\mathcal{A} = \{A_i\}_{i \in I}$ is known as an algebra of subsets for Ω if

- $\emptyset, \Omega \in \mathcal{A}$;
- for any $A_i \in \mathcal{A}$, $A_i^c \in \mathcal{A}$;
- for any $A_i, A_j \in \mathcal{A}$, $A_i \cup A_j \in \mathcal{A}$.

We can refine the concept of an algebra of subsets to consider countable unions rather than finite unions. This is known as a σ-algebra.

Definition C.1.2 (σ-Algebra of Subsets)**.** The collection $\mathcal{A} = \{A_i\}_{i \in I}$ is known as a σ-algebra of subsets for Ω if

- $\emptyset, \Omega \in \mathcal{A}$;
- for any $A_i \in \mathcal{A}$, $A_i^c \in \mathcal{A}$;
- for any countable collection $\{A_n\}_{n \geqslant 1} \subseteq \mathcal{A}$, $\bigcup_{n \geqslant 1} A_n \in \mathcal{A}$.

Definition C.1.3 (Measurable Space)**.** A pair $(\Omega, \mathcal{A})$, where Ω is a set and $A \subseteq P(\Omega)$ is a σ-algebra, is called a measurable space. Elements in the measurable space are called $\mathcal{A}$-measurable sets.

Definition C.1.4 (Restriction of a σ-Algebra)**.** For a measurable space $(\Omega, \mathcal{A})$, with $E \in \mathcal{A}$, the family

$$\mathcal{A}_E = \{E \cap A \mid A \in \mathcal{A}\}$$

is a σ-algebra on E, known as the restriction of $\mathcal{A}$ to E.

Definition C.1.5 (Produced σ-Algebra)**.** Let $(\Omega, \mathcal{A})$ be a measurable space, and $f\colon \Omega \to \Lambda$ is a map. The σ-algebra produced by f on Λ is the collection

$$\mathcal{N} = \left\{E \mid E \subseteq \Lambda,\ f^{-1}(E) \in \mathcal{A}\right\}.$$

Definition C.1.6 (Generated σ-Algebra)**.** For a family $\mathcal{E} \subseteq P(\Omega)$, the σ-algebra generated by E is the smallest σ-algebra that contains E.

$$\sigma(\mathcal{E}) = \bigcap \left\{ \mathcal{M} \mid \mathcal{E} \subseteq \mathcal{M}, \mathcal{M} \text{ is a } \sigma\text{-algebra} \right\}.$$

Definition C.1.7 (Borel σ-Algebra)**.** If Ω is a topological space with topology $\tau \subseteq P(\Omega)$, we define

$$\mathcal{B}_\Omega = \sigma(\tau)$$

to be the Borel σ-algebra.

All open, closed, clopen, F_σ, and G_δ subsets of Ω are Borel.

We can now begin examining measurable functions.

Definition C.1.8 (Measurable Functions)**.** Let $(\Omega, \mathcal{M})$ and $(\Lambda, \mathcal{N})$ be measurable spaces.

(1) We say a map $f\colon \Omega \to \Lambda$ is $\mathcal{M}$-$\mathcal{N}$-measurable if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.

(2) We say a map $f\colon \Omega \to \mathbb{R}$ is measurable if it is $\mathcal{M}$-$\mathcal{B}_\mathbb{R}$-measurable.

(3) We say a map $f\colon \Omega \to \mathbb{C}$ is measurable if both $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are measurable.

The set of all measurable functions on $(\Omega, \mathcal{M})$ is denoted $L_0(\Omega, \mathcal{M})$.

The collection of all bounded measurable functions is the set

$$B_\infty(\Omega, \mathcal{M}) = \left\{ f \in L_0(\Omega, \mathcal{M}) \mid \sup_{x \in \Omega} |f(x)| < \infty \right\}.$$

Example C.1.1. If $f\colon \Omega \to \Lambda$ is a continuous map between topological spaces, then f is $\mathcal{B}_\Omega$-$\mathcal{B}_\Lambda$-measurable, since

$$\mathcal{F} = \left\{ E \subseteq \Lambda \mid f^{-1}(E) \in \mathcal{B}_\Omega \right\}$$

is a σ-algebra containing every open set in Λ, so $\mathcal{F}$ contains $\mathcal{B}_\Lambda$.

Example C.1.2. If $(\Omega, \mathcal{M})$ is a measurable space, and $f\colon \Omega \to \Lambda$ is a map, the measurable space $(\Lambda, \mathcal{N})$ produced by f is necessarily $\mathcal{M}$-$\mathcal{N}$-measurable.

Fact C.1.1. If $(\Omega, \mathcal{M})$, $(\Lambda, \mathcal{N})$, and $(\Sigma, \mathcal{L})$ are measurable spaces, with $f\colon \Omega \to \Lambda$ and $g\colon \Lambda \to \Sigma$ measurable, then $g \circ f$ is measurable.

Proposition C.1.1. Let $(\Omega, \mathcal{M})$ be a measurable space. Let $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Suppose $f, g, h_n\colon \Omega \to \mathbb{F}$ are measurable for $n \geq 1$.

(1) If $\alpha \in \mathbb{F}$, then $f + \alpha g$ is measurable.

(2) $\overline{f}$ is measurable.

(3) fg is measurable.

(4) $\frac{f}{g}$ is measurable assuming it is well-defined.

(5) if h_n are $\mathbb{R}$-valued, and $(h_n(x))_n$ is bounded for each $x \in \Omega$, then $\sup h_n$ and $\inf h_n$ are measurable.

(6) If f and g are $\mathbb{R}$ valued, then $\max(f, g)$ and $\min(f, g)$ are measurable. In particular,

$$f_+ = \max(f, 0)$$

$$f_- = \max(0, -f)$$

are measurable.

(7) $|f|$ is measurable.

(8) The pointwise limit of measurable functions is measurable — if $\lim_{n\to\infty} h_n(x)$ exists for all $x \in \Omega$, then $h = \lim_{n\to\infty} h_n$ is measurable.

Definition C.1.9 (Simple Functions)**.** A simple function $s\colon \Omega \to \mathbb{F}$ is a function with finite range. In other words, s is of the form

$$s = \sum_{k=1}^{n} c_k \mathbb{1}_{E_k}$$

for $E_k \subseteq \Omega$ and $c_k \in \mathbb{F}$.

A simple function is measurable if and only if $E_k \in \mathcal{M}$ for each k.

C.2 Constructing Measures

A measure assigns a nonnegative "length" or "volume" to measurable sets.

Definition C.2.1 (Measures on Measurable Spaces)**.** A measure on a measurable space $(\Omega, \mathcal{M})$ is a map $\mu\colon \mathcal{M} \to [0, \infty]$ that satisfies the following.

(i) $\mu(\emptyset) = 0$;

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

The triple $(\Omega, \mathcal{M}, \mu)$ is called a measure space.

A measure μ is finite if $\mu(\Omega) < \infty$

If $\mu(\Omega) = 1$, then μ is called a probability measure.

A measure μ is called finitely additive if $\mu(E \sqcup F) = \mu(E) + \mu(F)$.

A measure μ is called σ-finite if there is a countable family $\{E_n\}_{n\geqslant 1} \subseteq \mathcal{M}$ such that

$$\Omega = \bigcup_{n\geqslant 1} E_n$$

and $\mu(E_n) < \infty$.

A measure μ on $(\Omega, \mathcal{M})$ is called semi-finite if, for every $E \in \mathcal{M}$ with $\mu(E) = \infty$, there exists $F \in \mathcal{M}$ with $F \subseteq E$ and $0 < \mu(F) < \infty$.

Lemma C.2.1. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space.

(1) If $E, F \in \mathcal{M}$ with $F \subseteq E$, then $\mu(F) \subseteq \mu(E)$.

(2) If $(E_n)_n$ is a sequence of measurable sets, then

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

(3) If $(E_n)_{n\geqslant 1}$ is an increasing family of measurable sets, then

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

Definition C.2.2 (Pushforward Measure)**.** Let $(\Omega, \mathcal{M}, \mu)$ be a measure space, and let $(\Lambda, \mathcal{N})$ be a measurable space. Let $f\colon \Omega \to \Lambda$ be measurable. The map

$$f_*\mu\colon \mathcal{N} \to [0, \infty]$$

defined by

$$f_*\mu(E) = \mu\left(f^{-1}(E)\right)$$

defines a measure on $(\Lambda, \mathcal{N})$. This is known as the pushforward measure of μ.

If $\mathcal{N}$ on Λ is produced by f, then the pushforward measure is necessarily defined on $\mathcal{N}$, and that any function $g\colon \Lambda \to \mathbb{F}$ is measurable if and only if $g \circ f$ is measurable.

Definition C.2.3. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space.

A null set is a measurable set $N \in \mathcal{M}$ with $\mu(N) = 0$.

A property which holds for all $x \in \Omega \setminus N$ for some null set N is said to hold μ-almost everywhere, or μ-a.e.

Definition C.2.4. If $(\Omega, \mathcal{M}, \mu)$ is a measure space, we can define an equivalence relation on the set $L_0(\Omega, \mathcal{M}, \mu)$, by

$$f \sim_\mu g \text{ if and only if } \mu(\{x \mid f(x) \neq g(x)\}) = 0.$$

We define the set of all classes of measurable functions by

$$\begin{aligned} L(\Omega, \mu) &= L_0(\Omega, \mathcal{M})/{\sim_\mu} \\ &= \left\{[f]_\mu \mid f \in L_0(\Omega, \mathcal{M})\right\}. \end{aligned}$$

Fact C.2.1. The operations

- $[f]_\mu + [g]_\mu = [f+g]_\mu$;
- $[f]_\mu[g]_\mu = [fg]_\mu$;
- and $\alpha[f]_\mu = [\alpha f]_\mu$

are well-defined.

Definition C.2.5 (Essentially Bounded Functions and Continuous Functions)**.** Let $(\Omega, \mathcal{M}, \mu)$ be a measure space, and $f\colon \Omega \to \mathbb{C}$ be measurable. We say f is μ-essentially bounded if there is $C \geqslant 0$ such that

$$\mu(\{x \in \Omega \mid |f(x)| \geqslant C\}) = 0.$$

We say C is an essential bound for f. The infimum of all essential bounds is the essential supremum, which gives

$$\begin{aligned} \|f\|_\infty &= \operatorname{ess\,sup}(f) \\ &= \inf\{C \geqslant 0 \mid \mu(\{x \in \Omega \mid |f(x)| \geqslant C\}) = 0\}. \end{aligned}$$

The collection of all μ-essentially bounded functions is denoted

$$L_\infty(\Omega, \mu) = \left\{[f]_\mu \in L(\Omega, \mu) \mid \|f\|_\infty < \infty\right\}.$$

Note that $B_\infty(\Omega,\mu) = L_\infty(\Omega,\mu)$ as sets.

If Ω is a topological space, with $\mathcal{B}_\Omega$ the Borel σ-algebra, we have $C(\Omega) \subseteq L_0(\Omega, \mathcal{B}_\Omega)$.

For μ a measure on $(\Omega, \mathcal{B}_\Omega)$, the μ-equivalence classes of continuous functions are

$$C(\Omega,\mu) = \left\{[f]_\mu \mid f \in C(\Omega)\right\}.$$

Members of $L(\Omega,\mu)$ and $L_\infty(\Omega,\mu)$ are equivalence classes of functions (rather than functions themselves), but we use the abuse of notation that $[f]_\mu = f$.

Fact C.2.2. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space, and let $f, g \colon \Omega \to \mathbb{C}$ be measurable, and $\alpha \in \mathbb{C}$. Then, the following are true:

- $\|f+g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$;
- $\|\alpha f\|_\infty = |\alpha|\|f\|_\infty$;
- if $\|f\|_\infty = 0$, then $f = 0$ μ-a.e.;
- $\|f\|_\infty \leqslant \|f\|_u$;
- if f is essentially bounded, then

$$\mu(\{x \mid |f(x)| \geqslant \|f\|_\infty\}) = 0.$$

Definition C.2.6 (Complete Measure Space)**.** A measure space $(\Omega, \mathcal{M}, \mu)$ is said to be complete if all subsets of null sets are measurable (hence null).

C.3 Integration

Definition C.3.1. If $\phi \colon \Omega \to [0,\infty)$ is a positive, simple, and measurable function,

$$\phi = \sum_{k=1}^{n} c_k \mathbb{1}_{E_k},$$

then the integral of ϕ is defined as

$$\int_\Omega \phi \, d\mu = \sum_{k=1}^{n} c_k \mu(E_k),$$

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

The value of this integral is not dependent on the representation of ϕ.

Definition C.3.2. If $f \colon \Omega \to \infty[0,\infty)$ is a positive measurable function, then

$$\int_\Omega f \, d\mu = \sup\left\{\int_\Omega \phi \, d\mu \mid \phi \text{ measurable and simple, } 0 \leqslant \phi \leqslant f\right\}.$$

If $E \subseteq \Omega$ is measurable, we define

$$\int_E f \, d\mu = \int_\Omega f \mathbb{1}_E \, d\mu.$$

Proposition C.3.1. Let $(\Omega, \mathcal{M})$ be a measurable space, and let $f : \Omega \to \mathbb{C}$ be measurable. There is a sequence $(\phi_n)_n$ of simple, measurable functions with $(\phi_n(x))_n \xrightarrow{n\to\infty} f(x)$.

If $f \geqslant 0$, we can take ϕ_n to be positive and pointwise increasing.

If f is bounded, then this convergence is uniform, and $(\phi_n)_n$ can be chosen to be uniformly bounded.

Theorem C.3.1 (Monotone Convergence Theorem)**.** Let $(f_n : \Omega \to [0,\infty))$ be an increasing sequence of positive, measurable functions converging pointwise to $f\colon \Omega \to [0,\infty)$. Then, f is measurable, and

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu.$$

Definition C.3.3. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space.

(1) A measurable function $f\colon \Omega \to [0,\infty)$ is integrable if

$$\int_\Omega f \, d\mu < \infty.$$

(2) A measurable function $f\colon \Omega \to \mathbb{R}$ is integrable if both f_+ and f_- are integrable. We define

$$\int_\Omega f \, d\mu = \int_\Omega f_+ \, d\mu - \int_\Omega f_- \, d\mu.$$

(3) A measurable function $f\colon \Omega \to \mathbb{C}$ is said to be integrable if both $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are integrable. We define

$$\int_\Omega f \, d\mu = \int_\Omega \mathrm{Re}(f) \, d\mu + i \int_\Omega \mathrm{Im}(f) \, d\mu.$$

Fact C.3.1. Let $f, g\colon \Omega \to \mathbb{C}$ be integrable functions, and $\alpha \in \mathbb{C}$. Then,

- $f + \alpha g$ is integrable, and $\displaystyle\int_\Omega (f + \alpha g) \, d\mu = \int_\Omega f \, d\mu + \alpha \int_\Omega g \, d\mu$;
- if f and g are real-valued, and $f \leqslant g$, then $\displaystyle\int_\Omega f \, d\mu \leqslant \int_\Omega g \, d\mu$;
- $\displaystyle\left|\int_\Omega f \, d\mu\right| \leqslant \int_\Omega |f| \, d\mu.$

Fact C.3.2. If $f = g$ μ-a.e., then

$$\int_\Omega f \, d\mu = \int_\Omega g \, d\mu.$$

Fact C.3.3. If $f\colon \Omega \to \mathbb{C}$ is measurable, then $\int_\Omega |f| \, d\mu = 0$ if and only if $f = 0$ μ-a.e.

Fact C.3.4. A measurable function $f\colon \Omega \to \mathbb{C}$ is integrable if and only if $|f|$ is integrable.

Definition C.3.4 (Integrable Functions)**.** Let $(\Omega, \mathcal{M}, \mu)$ be a measure space.

(1) We define the set of (classes of) integrable functions to be

$$L_1(\Omega, \mu) = \left\{ [f]_\mu \in L(\Omega, \mu) \mid f \text{ is integrable} \right\}.$$

(2) We define the set of (classes of) square-integrable functions to be

$$L_2(\Omega, \mu) = \left\{[f]_\mu \in L(\Omega, \mu) \mid |f|^2 \text{ is integrable}\right\}.$$

Definition C.3.5. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space. If f and $(f_n)_n$ are integrable with $\|f - f_n\|_1 \xrightarrow{n\to\infty} 0$, we say $(f_n)_n$ converges in mean to f.

Fact C.3.5. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space.

(1) For $f \in L_1(\Omega, \mu)$, the maps

$$[f]_\mu \longmapsto \int_\Omega f \, d\mu$$
$$[f]_\mu \longmapsto \int_\Omega |f| \, d\mu$$

are well-defined.

(2) For $f \in L_1(\Omega, \mu)$, we define

$$\|f\|_1 = \int_\Omega |f| \, d\mu.$$

This is a well-defined norm.

$$\begin{aligned} \|f + g\|_1 &\leqslant \|f\|_1 + \|g\|_1 \\ \|\alpha f\|_1 &= |\alpha| \|f\|_1 \\ \|f\|_1 = 0 &\Leftrightarrow f = 0 \ \mu\text{-a.e.} \end{aligned}$$

(3)

$$d_1\Big([f]_\mu, [g]_\mu\Big) = \|f - g\|_1$$

is a metric on $L_1(\Omega, \mu)$.

Theorem C.3.2 (Dominated Convergence Theorem)**.** Let $(f_n : \Omega \to \mathbb{C})_n$ be a sequence of measurable functions converging pointwise to a measurable function $f\colon \Omega \to \mathbb{C}$. If there is an integrable $g\colon \Omega \to [0, \infty)$ with $|f_n| \leqslant g$ for all n, then

$$\int_\Omega f_n \, d\mu \xrightarrow{n\to\infty} \int_\Omega f \, d\mu.$$

Corollary C.3.1. If $f\colon \Omega \to \mathbb{C}$ is integrable, then there is a sequence of simple integrable functions $(\phi_n)_n$ with $\|f - \phi_n\|_1 \xrightarrow{n\to\infty} 0$.

Corollary C.3.2. If $f\colon \mathbb{R} \to \mathbb{C}$ is integrable, then there is a sequence $(f_n)_n$ of compactly supported integrable functions such that $\|f - f_n\|_1 \xrightarrow{n\to\infty} 0$.

Theorem C.3.3. If $f\colon \mathbb{R} \to \mathbb{C}$ is integrable, and $\varepsilon > 0$, there is a continuous, compactly supported function g with $\|f - g\|_1 < \varepsilon$.

Proposition C.3.2. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space, and let $(\Lambda, \mathcal{N})$ be a measurable space with $f\colon \Omega \to \Lambda$ a measurable map. Let $f_*\mu$ be the pushforward measure on $(\Lambda, \mathcal{N})$. For a measurable function $g\colon \Lambda \to$

$[0, \infty)$, then

$$\int_{\Lambda} g \, d(f_* \mu) = \int_{\Omega} (g \circ f) \, d\mu.$$

Moreover, if $g \colon \Lambda \to \mathbb{F}$ is integrable with respect to $f_*\mu$, then so too is $g \circ f$ with respect to μ.

C.4 Complex Measures

Example C.4.1. If $(\Omega, \mathcal{M}, \mu)$ is a measure space, then the map $\mu_f(E) = \int_E f \, d\mu$ is a well-defined measure.

Definition C.4.1. Let $(\Omega, \mathcal{M}, \mu)$ be a measurable space.

(1) A complex number on $(\Omega, \mathcal{M}, \mu)$ is a map $\mu \colon \mathcal{M} \to \mathbb{C}$ satisfying the following conditions.

- $\mu(\emptyset) = 0$;
- $\mu\left(\bigsqcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k)$ for $\{E_k\}_{k \geqslant 1} \subseteq \mathcal{M}$.

(2) We write $M(\Omega)$ to be the set of all complex measures on $(\Omega, \mathcal{M})$.

(3) If $\mu \in M(\Omega)$, and $\mu(E) \in \mathbb{R}$ for all $E \in \mathcal{M}$, then we say μ is a real measure on $(\Omega, \mathcal{M})$.

(4) If $\mu \in M(\Omega)$ and $\mu(E) \geqslant 0$ for all $E \in \mathcal{M}$, then we say μ is a positive measure on $(\Omega, \mathcal{M})$.

(5) If μ is a positive measure on $(\Omega, \mathcal{M})$ with $\mu(\Omega) = 1$, we say μ is a probability measure on $(\Omega, \mathcal{M})$. We write $\mathcal{P}(\Omega, \mathcal{M})$ to be the collection of all probability measures on $(\Omega, \mathcal{M})$.

(6) If Ω is a LCH space, we always let $M(\Omega)$ be the set of all complex Borel measures on Ω.

Definition C.4.2. If $(\Omega, \mathcal{M})$ is a measurable space, and $x \in \Omega$, the Dirac measure at x is defined by

$$\delta_x \colon \mathcal{M} \to [0, 1]$$
$$\delta_x(E) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}.$$

If $x_1, \ldots, x_n$ are distinct points in Ω, and $t_1, \ldots, t_n \in [0, 1]$ with $\sum_{j=1}^{n} t_j = 1$, then

$$\mu = \sum_{j=1}^{n} t_j \delta_{x_j}$$

is a probability measure on $(\Omega, \mathcal{M})$.

Fact C.4.1. If μ is a complex measure on $(\Omega, \mathcal{M})$, then $\overline{\mu}$, defined by $\overline{\mu}(E) = \overline{\mu(E)}$ for $E \in \mathcal{M}$, is also a complex measure. Additionally, $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$, defined by

$$\mathrm{Re}(\mu)(E) = \mathrm{Re}(\mu(E))$$
$$\mathrm{Im}(\mu)(E) = \mathrm{Im}(\mu(E))$$

are real measures.

Definition C.4.3. If $\mu \in M(\Omega)$, then the total variation of μ is the quantity

$$|\mu| \colon \mathcal{M} \to [0, \infty]$$

with

$$|\mu|(E) = \sup\left\{ \sum_{j=1}^{\infty} |\mu(E_j)| \,\middle|\, E = \bigsqcup_{j=1}^{\infty} E_j,\ E_j \in \mathcal{M} \right\}.$$

Fact C.4.2. If $\mu \in M(\Omega)$, then $|\mu|$ is a positive, finite measure. Additionally, if $\mu, \nu \in M(\Omega)$ with $\alpha \in \mathbb{C}$, then

(a) $|\mu(E)| \leqslant |\mu|(E)$

(b) $|\mu + \nu|(E) \leqslant |\mu|(E) + |\nu|(E)$

(c) $|\alpha\mu|(E) = |\alpha||\mu|(E)$.

Definition C.4.4 (Absolute Continuity of Measures)**.** Let $(\Omega, \mathcal{M})$ be a measurable space, and let μ and ν be positive measures on this space. If $\mu(A) > 0$ implies $\nu(A) > 0$ for a given $A \in \mathcal{M}$, we say μ is absolutely continuous with respect to ν. We write $\mu \ll \nu$.

Theorem C.4.1 (Radon–Nikodym Theorem)**.** If $\mu \ll \nu$ on $(\Omega, \mathcal{M})$, then there exists a measurable function $f\colon \Omega \to [0, \infty]$ such that

$$\nu(A) = \int_A f \, d\nu$$

for each $A \in \mathcal{M}$.

Remark C.4.1. The Radon–Nikodym theorem extends to signed and complex measures.

Fact C.4.3. Let $(\Omega, \mathcal{M}, \lambda)$ be a measure space, and suppose $f \in L_1(\Omega, \lambda)$. Then, $\mu(E) = \int_E f \, d\lambda$ defines a complex measure. We write $f = \frac{d\mu}{d\lambda}$, which is the Radon–Nikodym derivative of μ with respect to λ.

It is also the case that

$$|\mu|(E) = \int_E |f| \, d\lambda.$$

Fact C.4.4. If $\mu \in M(\Omega)$, there exists a measurable function $f\colon \Omega \to \mathbb{C}$ such that $|f| = 1$ and $\mu(E) = \int_E f \, d|\mu|$ for all $E \in \mathcal{M}$.

Definition C.4.5. Let Ω be a LCH space equipped with the Borel σ-algebra, $\mathcal{B}_\Omega$.

(1) A Borel measure $\mu\colon \mathcal{B}_\Omega \to [0, \infty]$ is called

- inner regular on $E \in \mathcal{B}_\Omega$ if
$$\mu(E) = \sup\{\mu(K) \mid K \subseteq E, K \text{ compact}\};$$
- outer regular on $E \in \mathcal{B}_\Omega$ if
$$\mu(E) = \inf\{\mu(U) \mid U \supseteq E, U \text{ open}\};$$
- regular on E if it is inner regular and outer regular on E;
- regular if it is regular on all $E \in \mathcal{B}_\Omega$;
- Radon if
  - $\mu(K) < \infty$ for all compact $K \subseteq \Omega$;
  - μ is inner regular on all open sets and outer regular on all Borel sets.

(2) A complex Borel measure $\mu\colon \mathcal{B}_\Omega \to \mathbb{C}$ is regular if $|\mu|$ is regular; μ is Radon if $|\mu|$ is Radon.

(3) We write $M_r(\Omega)$ to denote the set of all complex regular measures on $(\Omega, \mathcal{B}_\Omega)$.

Fact C.4.5. Every positive Radon measure is regular. Thus, every complex Borel measure is regular if and only if it is Radon.

Moreover, if Ω is a second countable LCH space, then every complex Borel measure is regular.

Definition C.4.6. Let (Ω, τ) be a topological space, and suppose $\mu \colon \mathcal{B}_\Omega \to [0, \infty]$ is a Borel measure.

(1) The kernel of μ is the set

$$N_\mu = \bigcup \{U \subseteq \Omega \mid U \in \tau,\ \mu(U) = 0\}.$$

(2) The support of μ is the complement of the kernel, $\operatorname{supp}(\mu) = N_\mu^c$.

Fact C.4.6. If μ is a Radon measure on a LCH space Ω, then $\mu(N_\mu) = 0$, meaning $\mu(\Omega) = \mu(\operatorname{supp}(\mu))$.

Theorem C.4.2 (Hahn and Jordan Decomposition)**.** Let $(\Omega, \mathcal{M})$ be a measurable space, and let $\mu \colon \mathcal{M} \to \mathbb{R}$ be a real measure. Then, there is a measurable partition $\Omega = P \sqcup N$ such that for all $E \subseteq P$, $\mu(E) \geqslant 0$, and for all $E \subseteq N$, $\mu(E) \leqslant 0$. This partition is unique up to a μ-null symmetric difference — that is, for any P', N' satisfying the conditions, $\mu(P' \triangle P) = 0$ and $\mu(N' \triangle N) = 0$.

There is a unique decomposition $\mu = \mu_+ - \mu_-$, with $\mu_\pm$ that are positive such that if $E \subseteq P$, then $\mu_-(E) = 0$, and if $E \subseteq N$, $\mu_+(E) = 0$.

Definition C.4.7. Let $(\Omega, \mathcal{M})$ be a measurable space, and let $f \colon \Omega \to \mathbb{C}$ be measurable.

(1) If $\mu \colon \mathcal{M} \to \mathbb{R}$ is a real measure with $\mu = \mu_+ - \mu_-$, we say that f is μ-integrable if it is both μ_+ and μ_--integrable. We define

$$\int_\Omega f \, d\mu = \int_\Omega f \, d\mu_+ - \int_\Omega f \, d\mu_-.$$

(2) If $\mu : \mathcal{M} \to \mathbb{C}$ is a complex measure with $\mu_1 = \operatorname{Re}(\mu)$ and $\mu_2 = \operatorname{Im}(\mu)$, we say f is μ-integrable if it is both μ_1 and μ_2-integrable. We define

$$\int_\Omega f \, d\mu = \int_\Omega f \, d\mu_1 + i \int_\Omega f \, d\mu_2.$$

Theorem C.4.3 (Riesz Representation Theorem on $C_c(\Omega)$)**.** Let Ω be a LCH space. If $\varphi \colon C_c(\Omega) \to \mathbb{C}$ is a positive linear functional, then there is a unique Radon measure μ such that

$$\varphi(f) = \int_\Omega f \, d\mu$$

for all $f \in C_c(\Omega)$. Additionally, for every open $U \subseteq \Omega$, we have

$$\mu(U) = \sup\left\{ \varphi(f) \,\middle|\, f \in C_c(\Omega, [0,1]), \operatorname{supp}(f) \subseteq U \right\},$$

and for every compact $K \subseteq \Omega$, we have

$$\mu(K) = \inf\{ \varphi(f) \mid f \geqslant \mathbb{1}_K \}.$$

Theorem C.4.4 (Riesz Representation Theorem on $C(X)$)**.** Let X be a compact metric space, and let $\varphi \in (C(X))^*$ be a positive linear functional with $\varphi(\mathbb{1}_X) = \|\varphi\| = 1$. Then, for $f \in C(X)$, there is a unique Borel probability measure such that

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

Appendix D

Functional Analysis

D.1 Normed Vector Spaces and Algebras

The fundamental unit of functional analysis is functions — specifically, collections of functions equipped with particular operations and a norm that turn them into vector spaces and algebras. This section will focus on some of the basic facts and theory surrounding normed vector spaces and algebras.

Definition D.1.1 (Seminorms and Norms)**.** Let X be a $\mathbb{F}$-vector space, and let $p\colon X \to [0,\infty)$ be a function. If

- $p(\lambda x) = |\lambda| p(x)$ for all $\lambda \in \mathbb{F}$ and $x \in X$ (homogeneity), and
- $p(x+y) \leqslant p(x) + p(y)$ for all $x, y \in X$ (triangle inequality),

we say p is a seminorm. If p also satisfies

- $p(x) = 0$ if and only if $x = 0$ (positive definite),

then p is a norm. Norms on vector spaces are usually denoted $\|\cdot\|$. Additionally, if X is an algebra, the (semi)norm also has to be sub-multiplicative — i.e.,

$$\|xy\| \leqslant \|x\|\|y\|.$$

Two norms, $\|\cdot\|_a$ and $\|\cdot\|_b$, are said to be equivalent if there exist constants C_1 and C_2 such that

$$\begin{aligned}\|x\|_a &\leqslant C_1\|x\|_b\\ \|x\|_b &\leqslant C_2\|x\|_a\end{aligned}$$

for all $x \in X$.

If X is complete with respect to the metric $d(x,y) = \|x-y\|$, we call X a Banach space. If X is a normed algebra that is complete with respect to its induced metric, then we say X is a Banach algebra.

Theorem D.1.1. Let X be a normed vector space. Then, X is complete if and only if, for every sequence of vectors $(x_k)_k$, if $\sum_{k=1}^{\infty}\|x_k\|$ converges, then $\sum_{k=1}^{\infty} x_k$ converges.

Definition D.1.2 (Open Balls, Closed Balls, Spheres)**.** Let X be a normed vector space.

- We write
$$U(x,\delta) = \{y \in X \mid \|y-x\| < \delta\}$$
to be the open ball of radius δ centered at x. We write $U_X = U(0,1)$.

- We write
$$B(x, \delta) = \{y \in X \mid \|y - x\| \leqslant \delta\}$$
to be the closed ball of radius δ centered at x. We write $B_X = B(0, 1)$.

- We write
$$S(x, \delta) = \{y \in X \mid \|y - x\| = \delta\}$$
to be the sphere of radius δ centered at x. We write $S_X = S(0, 1)$.

Definition D.1.3. Let X be a normed vector space. A subset $A \subseteq X$ is said to be total if its closed linear span is equal to X; $\overline{\text{span}}(A) = X$.

Definition D.1.4. Let $T\colon X \to Y$ be a linear map between normed vector spaces. We say T is bounded if its operator norm, defined by
$$\|T\|_{op} = \sup_{x \in B_X} \|T(x)\|$$
is finite. We write $\mathbb{B}(X, Y)$ for the set of all bounded linear maps between X and Y.

Remark D.1.1. Note that if Y is complete, then $\mathbb{B}(X, Y)$ is a Banach space with pointwise addition and scalar multiplication.

A quick sketch of the proof is as follows: consider a $\|\cdot\|_{op}$-Cauchy sequence $(T_n)_n$ in $\mathbb{B}(X, Y)$, and define T to be the pointwise limit of $(T_n)_n$, which exists as for any $y \in Y$, $(T_n(y))_y$ is Cauchy in Y. Then, it can be shown that defining T in this manner yields convergence in operator norm.

Furthermore, if we define $\mathbb{B}(X) = \mathbb{B}(X, X)$, then this space is a normed algebra with pointwise addition, scalar multiplication, and composition of operators. The algebra $\mathbb{B}(X)$ is complete if X is complete.

Fact D.1.1. The following are equivalent for a linear map T:

- T is continuous at 0;
- T is continuous;
- T is uniformly continuous;
- T is bounded.

Definition D.1.5. Let $T\colon X \to Y$ be a linear map between normed vector spaces.

- We say T is bounded below if there exists $C > 0$ such that $\|T(x)\| \geqslant C\|x\|$ for all $x \in X$.
- If T is bounded and bounded below, we say T is bicontinuous.
- If T is a linear isomorphism that is bicontinuous, we say T is a bicontinuous isomorphism, and say $X \cong Y$ are bicontinuously isomorphic.
- If T is a linear isomorphism and is such that $\|T(x)\| = \|x\|$ for all x, then we say T is an isometric isomorphism.

Definition D.1.6. Let X be a normed vector space. The subset $X^* \subseteq X'$, where X' is the algebraic dual of X, is the set of all continuous linear functionals on X:
$$X^* = \mathbb{B}(X, \mathbb{F}).$$

Definition D.1.7 (Unconditional Summability)**.** If Ω is a set and $f\colon \Omega \to X$ is any function between Ω and suitable vector space X (see Definition D.4.1), we say the unconditional series $\sum_{j\in\Omega} f(j)$ converges to some value $k \in X$ if the net $(s_F)_{F\in\mathcal{F}}$ converges to k, where

$$\mathcal{F} = \{F \mid F \subseteq J, \operatorname{card}(F) < \infty\}$$

is the collection of finite subsets of Ω directed by inclusion.

Definition D.1.8 (Three Fundamental Function Spaces)**.** Let Ω be any set.

- The space $\ell_1(\Omega)$ is the set of all functions $f\colon \Omega \to \mathbb{C}$ such that $\displaystyle\sum_{t\in\Omega} |f(t)| < \infty$.
- The space $\ell_2(\Omega)$ is the set of all functions $f\colon \Omega \to \mathbb{C}$ such that $\displaystyle\sum_{t\in\Omega} |f(t)|^2 < \infty$.
- The space $\ell_\infty(\Omega)$ is the set of all functions $f\colon \Omega \to \mathbb{C}$ such that $\sup_{t\in\Omega}|f(t)| < \infty$.

D.2 The Fundamental Theorems of Banach Spaces

Definition D.2.1. Let X be a topological space. We say X is a Baire space if the intersection of any countable collection of open, dense subsets is also dense.

The Baire category theorem serves as one of the central bridges between functional analysis and topology. Note that the property of being a Baire space is a purely topological definition, while completeness is an analytic concept.

Theorem D.2.1 (Baire Category Theorem)**.** Let X be a complete metric space. Then, X is a Baire space.

The Baire category theorem is used to prove many important theorems in functional analysis, such as the ones that follow. They all fundamentally rely on the completeness of Banach spaces, which is expressed through the Baire category theorem. Proofs for these theorems can be found in functional analysis textbooks such as [Rud73].

Theorem D.2.2 (Open Mapping Theorem)**.** Let $T\colon X \to Y$ be a surjective linear map between Banach spaces. Then, T is an open map — i.e., if $U \subseteq X$ is open, then V is also open.

Corollary D.2.1 (Bounded Inverse)**.** If $T\colon X \to Y$ is a bounded linear map that is bijective, then T^{-1} is also bounded.

Theorem D.2.3 (Closed Graph Theorem)**.** Let $T\colon X \to Y$ be a linear map between Banach spaces. Then, T is bounded if and only if $\operatorname{graph}(T) = \{(x, T(x)) \mid x \in X\} \subseteq X \times Y$ is closed in the product topology.

Theorem D.2.4 (Uniform Boundedness Principle)**.** Let $\{T_i\}_{i\in I}$ be a family of maps in $\mathbb{B}(X, Y)$ such that, for all $x \in X$, $\sup_{i\in I}\|T_i(x)\| < \infty$. Then, $\sup_{i\in I}\|T_i\|_{op} < \infty$.

Now, we turn our attention towards linear functionals. Consider the following problem in algebra: suppose X is a finite-dimensional vector space, and $Y \subseteq X$ is a subspace. If we have a linear functional $\varphi\colon Y \to \mathbb{F}$, can this linear functional be extended to the full space?

The answer is yes — if $\mathcal{B} = \{x_1, \dots, x_m\}$ is a basis for Y, a fundamental result states that this basis can be extended to a basis for X, $\mathcal{C} = \{x_1, \dots, x_m, x_{m+1}, \dots, x_n\}$; if we define $c_i = \varphi(x_i)$ for $1 \leqslant i \leqslant m$, we may define $\varphi(x_i) = 0$ for $m + 1 \leqslant i \leqslant n$. In other words, we may always extend elements of the algebraic dual from a subspace to the full space.[I]

[I]This is part of a more general fact that vector spaces are injective objects. We discuss this a bit in the section on free vector spaces.

However, if X is infinite-dimensional and equipped with a norm (or, more generally, a locally convex topology, see Definition D.4.2), we also care about continuity, norm, and whether these extensions preserve continuity and norm. This is the domain of the Hahn–Banach theorems, which establish extension and separation results in normed vector spaces (and, as detailed later, locally convex topological vector spaces, see Theorem D.4.1).

Definition D.2.2 (Minkowski Functional)**.** We call a map $p\colon X \to [0,\infty)$ a Minkowski functional if

- $p(x+y) \leqslant p(x) + p(y)$ for all $x, y \in X$, and;
- $p(tx) = tp(x)$ for all $x \in X$ and $t > 0$.

Theorem D.2.5 (Hahn–Banach–Minkowski Extension)**.** Let $Y \subseteq X$ be a linear subspace of a normed vector space X, and let $\phi \in Y^*$ and p a Minkowski functional be such that for all $y \in Y$, $\phi(y) \leqslant p(y)$. Then, there is a map $\Phi \in X^*$ such that

- $\Phi|_Y = \phi$, and;
- $\Phi(x) \leqslant p(x)$ for all $x \in X$.

Theorem D.2.6 (Hahn–Banach Continuous Extension)**.** Let X be a normed vector space, and $\phi \in Y^*$, where $Y \subseteq X$ is a linear subspace. Then, there exists a linear functional $\Phi \in X^*$ such that $\|\Phi\|_{op} = \|\phi\|_{op}$, and $\Phi|_Y = \phi$. This extension is not necessarily unique.

The Hahn–Banach extension theorems lend themselves nicely to understanding the separation properties of linear functionals in the continuous dual space. These results allow us to know that there are "enough" linear functionals in the dual space of any normed vector space that allow us to distinguish points from closed subspaces and distinguish points from each other.

Theorem D.2.7 (Hahn–Banach Separation)**.** Let X be a normed vector space.

- For a fixed $x_0 \in X$, there exists a linear functional $\phi \in X^*$ such that $\phi(x_0) = \|x_0\|$.
- For a proper closed subspace $Y \subseteq X$ and some fixed $x_0 \in X \setminus Y$, there is a $\phi \in X^*$ such that $\|\phi\|_{op} \leqslant 1$, $\phi|_Y = 0$, and $\phi(x_0) = \mathrm{dist}_Y(x_0)$.

Corollary D.2.2. Let X be a normed space. For every $x \in X$, we have

$$\sup_{\phi \in B_{X^*}} |\varphi(x)| = \|x\|.$$

D.3 Duality

Here, we discuss a little bit more of the theory of dual spaces.

Definition D.3.1. Let X be a normed vector space. The linear functional $\hat{x}\colon X^* \to \mathbb{C}$, defined by

$$\hat{x}(\varphi) = \varphi(x)$$

is bounded with norm $\|\hat{x}\|_{op} = \|x\|$. We define the embedding $\iota\colon X \hookrightarrow X^{**}$ by

$$\iota(x) = \hat{x}.$$

We call ι the canonical embedding.

Definition D.3.2. Let X be a normed space. A norm completion of X is a pair (Z, j), where Z is a Banach space, $j\colon X \hookrightarrow Z$ is a linear isometry, and $\overline{\mathrm{Ran}}(j) = Z$.

Proposition D.3.1. Let X be a normed space, and set $\widetilde{X} = \overline{\iota_X(X)}^{\|\cdot\|_{op}} \subseteq X^{**}$. Then, $\left(\widetilde{X}, \iota_X\right)$ is a norm completion of X. Additionally, if (Z, j) is any other norm completion of X, then there is an isometric isomorphism $Z \to \widetilde{X}$.

Proposition D.3.2. Let X and Y be normed spaces, and let $T \in \mathbb{B}(X, Y)$. Then, there is a unique $\widetilde{T} \in \mathbb{B}\left(\widetilde{X}, \widetilde{Y}\right)$ such that $\widetilde{T} \circ \iota_X = \iota_Y \circ T$. The diagram below commutes.

$$\begin{tikzcd} \widetilde{X} \arrow[r, "\widetilde{T}"] & \widetilde{Y} \\ X \arrow[u, "\iota_X"] \arrow[r, "T"'] & Y \arrow[u, "\iota_Y"'] \end{tikzcd}$$

Furthermore, we have $\|T\|_{op} = \left\|\widetilde{T}\right\|_{op}$. If T is isometric, then so is $\widetilde{T}$, and if T is an isometric isomorphism, then so is $\widetilde{T}$.

Definition D.3.3. A normed space is called a dual space if there is a normed space Z such that $Z^* \cong X$ are isometrically isomorphic. We call Z the predual of X.

Example D.3.1.

- We have $c_0^* \cong \ell_1$, where c_0 is the space of all sequence vanishing at infinity, and ℓ_1 is the space of all absolutely summable sequences.
- We have $\ell_1^* \cong \ell_\infty$, where ℓ_∞ is space of all bounded sequences.
- If μ is a σ-finite measure on the measurable space $(\Omega, \mathcal{M})$, and $p, q \in (1, \infty)$ are such that $p^{-1} + q^{-1} = 1$, then $L_p(\Omega, \mu)^* \cong L_q(\Omega, \mu)$. Here,

$$L_p(\Omega, \mu) = \left\{f\colon \Omega \to \mathbb{C} \;\middle|\; \int_\Omega |f|^p \, d\mu < \infty\right\}.$$

  Additionally, if μ is semi-finite, then $L_1(\Omega, \mu)^* \cong L_\infty(\Omega, \mu)$.

Theorem D.3.1 (Riesz–Markov Theorem)**.** Let Ω be a LCH space. Then, $M_r(\Omega) \cong C_0(\Omega)^*$ are isometrically isomorphic, where $M_r(\Omega)$ is equipped with the norm $\|\mu\| = |\mu|(\Omega)$ for $\mu \in M_r(\Omega)$.

D.4 Topological Vector Spaces

Earlier, we discussed some of the features of normed vector spaces and Banach spaces. Here, we expand our scope to to examine the analytic properties of vector spaces whose topology is not necessarily induced by a norm.

Definition D.4.1. Let X be a $\mathbb{C}$-vector space, and let τ be a topology on X. We say τ is compatible with the vector space structure of X if

(1) X is T1 (see Definition B.3.9);

(2) scalar multiplication, $(\lambda, x) \mapsto \lambda x$ is continuous, where $\mathbb{C} \times X$ is given the product topology;

(3) vector addition, $(x, y) \mapsto x + y$ is continuous, where $X \times X$ is given the product topology.

If X is equipped with a topology compatible with the vector space structure of X, then (X, τ) is called a topological vector space. We abbreviate it as TVS.

Remark D.4.1. It can be shown that if (X, τ) is a TVS, the topology on X is automatically Hausdorff.

Definition D.4.2. A TVS (X, τ) is called locally convex if X admits a neighborhood base (see Definition B.3.7) consisting of convex sets. We abbreviate as LCTVS.

It can be shown that every LCTVS has a neighborhood base consisting of *balanced* convex sets (see Definition A.2.2).

An important structural result in the theory of topological vector spaces is the fact that every locally convex topology is generated by a separating family of seminorms.

Proposition D.4.1. Let X be a $\mathbb{C}$-vector space, and let $\mathcal{P}$ be a family of seminorms on X. For each $z \in X$ and $p \in \mathcal{P}$, we define $f_{p,z} \colon X \to [0, \infty)$ by

$$f_{p,z}(x) = p(x - z).$$

The topology $\tau_{\mathcal{P}}$ is the initial topology on X induced by the family

$$\mathcal{F}_{\mathcal{P}} = \left\{ f_{p,z} \;\middle|\; p \in \mathcal{P}, z \in X \right\}.$$

If $\mathcal{P}$ is such that for each $x \neq 0$, there is some p such that $p(x) \neq 0$ (i.e., $\mathcal{P}$ separates the points of X), then the family $\mathcal{F}_{\mathcal{P}}$ separates points in X. It is then the case that $(X, \tau_{\mathcal{P}})$ is a LCTVS.

Convergence of nets in the topology $\tau_{\mathcal{P}}$ is defined by $(x_\alpha)_\alpha \xrightarrow{\tau_{\mathcal{P}}} x$ if and only if $p(x_\alpha - x) \to 0$ for all $p \in \mathcal{P}$.

Furthermore, if (X, τ) is any LCTVS, then there is a corresponding family of separating seminorms $\mathcal{P}$ such that $\mathrm{id} \colon (X, \tau) \to (X, \tau_{\mathcal{P}})$ is a homeomorphism.

The Hahn–Banach theorems (such as the extension and separation results) that we established in Theorems D.2.5, D.2.6, and D.2.7 have corresponding results in topological vector spaces. The extension results are relatively straightforward.

Theorem D.4.1 (Hahn–Banach Extension for LCTVS)**.** Let X be a LCTVS, and suppose $E \subseteq X$ is a subspace. If $\varphi \in E^*$, then there is a $\psi \in X^*$ such that $\psi|_E = \varphi$.

Corollary D.4.1. Let X be a LCTVS. Let $\{x_1, \ldots, x_n\} \subseteq X$ be linearly independent, and $\{\alpha_1, \ldots, \alpha_n\} \in \mathbb{C}$. Then, there exists $\varphi \in X^*$ such that $\varphi(x_j) = \alpha_j$ for all j.

To provide some context for the Hahn–Banach separation results, consider two open, disjoint, convex subsets $A, B \subseteq \mathbb{R}^n$. The hyperplane separation theorem from convex optimization (see [BV04, Chapter 2.6]) states that there is a nonzero vector $m \in \mathbb{R}^n$ and some $b \in \mathbb{R}$ such that the map $\varphi \colon \mathbb{R}^n \to \mathbb{R}$, defined by $\varphi(x) = m^T x - b$, is strictly negative for all $x \in A$ and is strictly positive for all $x \in B$. The affine hyperplane defined by $\{x \mid \varphi(x) = b\}$ is known as a separating hyperplane for A and B.

What the Hahn–Banach theorems allow us to do is extend this result beyond $\mathbb{R}^n$ to the case of any TVS — with a special case if the TVS is locally convex.

Theorem D.4.2 (Hahn–Banach Separation for TVS)**.** Let X be a TVS (that may or may not be locally convex) over $\mathbb{C}$. Let A and B be convex and disjoint subsets of X. If A is open, then there exists $\varphi \in X^*$, with $\varphi = u + iv$, and $t \in \mathbb{R}$ such that

$$u(a) < t \leqslant u(b)$$

for all $a \in A$ and $b \in B$.

If A and B are open, then the inequalities can be taken to be strict.

The requirement that A and B be open can be relaxed in the case of a LCTVS, where we can separate closed, disjoint, convex sets, so long as one of the sets is compact. Specifically, we are able to separate the sets by a double hyperplanes if the topology on X is locally convex.

Theorem D.4.3 (Hahn–Banach Separation for LCTVS)**.** Let X be a LCTVS, and suppose $C, K \subseteq X$ are closed, disjoint, convex subsets of X, with K compact. Then, there exists $\varphi \in X^*$, with $\varphi = u + iv$, $t \in \mathbb{R}$, and $\delta > 0$ such that

$$u(x) \leqslant t \leqslant t + \delta \leqslant u(y)$$

for all $x \in C$ and $y \in K$.

Proposition D.4.2. Let $W \subseteq X'$, where X' is the algebraic dual of X. For each $\varphi \in W$, consider the seminorm

$$p_\varphi(x) = |\varphi(x)|.$$

We let $\mathcal{P}_W = \{p_\varphi \mid \varphi \in W\}$. If $\mathcal{P}_W$ separates points, then we may construct the topology $\tau_{\mathcal{P}_W}$ as in Proposition D.4.1.

Alternatively, we may consider the initial topology on X induced by the family W, written $\sigma(X, W)$.

It is the case that $\mathrm{id}\colon (X, \tau_{\mathcal{P}_W}) \to (X, \sigma(X, W))$ is a homeomorphism.

Definition D.4.3 (Norm Topology)**.** Let X be a normed vector space. If $\mathcal{P} = \{\|\cdot\|\}$, then the topology $\tau_{\mathcal{P}}$ is known as the norm topology on X.

Convergence is defined by $(x_n)_n \xrightarrow{\|\cdot\|} x$ if and only if $\|x_n - x\| \to 0$.

Remark D.4.2. Normed vector spaces are metric space, and hence first countable (so sequences are sufficient to define convergence).

Definition D.4.4 (Weak Topology)**.** If X is a normed vector space, we say $\sigma(X, X^*)$ is the weak topology on X.

Convergence is defined by $(x_\alpha)_\alpha \xrightarrow{w} x$ if and only if $(\varphi(x_\alpha))_\alpha \to \varphi(x)$ for all $\varphi \in X^*$.

Definition D.4.5 (Weak* Topology)**.** If X is a normed vector space, we say $\sigma(X^*, \iota(X))$, where ι is the canonical embedding (see D.3.1), is the weak* topology on X^*.

Convergence is defined by $(\varphi_\alpha)_\alpha \xrightarrow{w^*} \varphi$ if and only if $(\varphi_\alpha(x))_\alpha \to \varphi(x)$ for all $x \in X$.

Theorem D.4.4 (Banach–Alaoglu Theorem)**.** Let X be a normed vector space.

(1) The unit ball in the dual space, B_{X^*}, is w^*-compact.

(2) A subset $C \subseteq X^*$ is w^*-compact if and only if C is w^*-closed and norm bounded.

D.5 Hilbert Spaces and Operators

In Chapters 6 and 7, we discuss the relationship between a group Γ and the way the group is represented as an algebra of bounded operators on a Hilbert space. Here, we discuss more exactly what is meant by "algebra of bounded operators on a Hilbert space."

Definition D.5.1. Let X be a vector space. A semi-inner product on X is a map $\langle \cdot, \cdot \rangle\colon X \times X \to \mathbb{C}$ such that

- $\langle \alpha x + y, z \rangle = \alpha \langle x, z \rangle + \langle y, z \rangle$;
- $\langle x, \alpha y + z \rangle = \overline{\alpha} \langle x, y \rangle + \langle x, z \rangle$;

- $\langle x, x\rangle \geqslant 0$.

If $\langle x, x\rangle = 0$ if and only if $x = 0$, then $\langle \cdot, \cdot\rangle$ is an inner product with induced norm $\|x\|^2 = \langle x, x\rangle$. We call X an inner product space if it is equipped with an inner product.

If X is an inner product space that is complete with respect to the induced norm, then we say X is a Hilbert space. We usually denote Hilbert spaces by $\mathcal{H}$.

There are a few important structural results that are established in linear algebra relating to inner product spaces. We detail three that are used extremely often.

Theorem D.5.1 (Polarization Identity)**.** Let X be an inner product space, and let $x, y \in X$. Then,

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=0}^{3}\left\|x + i^k y\right\|^2.$$

Theorem D.5.2 (Cauchy–Schwarz Inequality)**.** Let X be an inner product space, and let $x, y \in X$. Then,

$$|\langle x, y\rangle| \leqslant \|x\|\|y\|.$$

Lemma D.5.1. If X is any inner product space and $F\colon X \times X \to \mathbb{C}$ is a sesquilinear form — i.e., one that satisfies the following properties for all $x, x_1, x_2, y, y_1, y_2 \in X$ and $\alpha \in \mathbb{C}$:

- $F(\alpha x_1 + x_2, y) = \alpha F(x_1, y) + F(x_2, y)$;
- $F(x, y_1 + \alpha y_2) = F(x, y_1) + \overline{\alpha} F(x, y_2)$;

then there exists some $T \in \mathcal{L}(X)$ such that

$$F(x, y) = \langle T(x), y\rangle,$$

where $\langle \cdot, \cdot\rangle$ is the inner product on X.

Note that with these three results, we can show that any two sesquilinear forms F and G are equal along the diagonal — i.e., if $F(x, x) = G(x, x)$ for all x — then they are equal everywhere. Additionally, the Cauchy–Schwarz inequality for sesquilinear forms is expressed as

$$|F(x, y)| \leqslant F(x, x)^{1/2} F(y, y)^{1/2}.$$

These facts are used implicitly throughout the study of Hilbert spaces and operators on them.

Definition D.5.2. Let $\mathcal{H}$ be a Hilbert space. A subset $\{x_i\}_{i\in I}$ is called orthonormal if

$$\begin{aligned}\left\langle x_i, x_j\right\rangle &= \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \\ &= \delta_{ij}.\end{aligned}$$

A maximal orthonormal set in $\mathcal{H}$ is called an orthonormal basis; equivalently, the set $\{x_i\}_{i\in I}$ is an orthonormal basis if and only if $\operatorname{span}(\{x_i\}_{i\in I})$ is dense in $\mathcal{H}$.

Remark D.5.1. Every Hilbert space has an orthonormal basis. This can be found by applying Zorn's lemma on the partially ordered set of all orthonormal subsets of $\mathcal{H}$ ordered by inclusion.

Theorem D.5.3 (Bessel's Inequality and Parseval's Identity)**.** Let $\{e_i\}_{i\in I}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. Then, for any $x \in \mathcal{H}$,

$$\sum_{i\in I}|\langle x, e_i\rangle|^2 \leqslant \|x\|^2.$$

If $\{e_i\}_{i\in I}$ is an orthonormal basis, then

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

Theorem D.5.4. If $M \subseteq \mathcal{H}$ is a closed subspace of a Hilbert space $\mathcal{H}$, then for any $x \in \mathcal{H}$, there is a unique $y_x \in M$ such that $\|x - y_x\|$ is minimal.

The map $P_M : \mathcal{H} \to M$, $x \mapsto y_x$ is known as the orthogonal projection onto M. Additionally, the map P_M has the following properties:

- P_M is linear;
- $P_M^2 = P_M$;
- $\|P_M\|_{op} = 1$ (if $M = \{0\}$, then $\|P_M\|_{op} = 0$).

Setting $M^\perp$ to be the range of $I_{\mathcal{H}} - P_M$, it is also the case that $\mathcal{H}/M \cong M^\perp$, with $\mathcal{H} = M \oplus M^\perp$.

One of the most important structural results on Hilbert spaces relates the continuous dual of a Hilbert space to the inner product.

Theorem D.5.5 (Riesz Representation Theorem for Hilbert Spaces)**.** Let $\mathcal{H}$ be a Hilbert space, and let $\varphi \in \mathcal{H}^*$. Then, there is a unique $f_\varphi \in \mathcal{H}$ such that

$$\varphi(g) = \langle g, f_\varphi\rangle$$

for all $g \in \mathcal{H}$.

Now that we understand the structure of Hilbert spaces and their closed subspaces, we can now begin understanding bounded operators on Hilbert spaces.

Definition D.5.3. Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded operator between Hilbert spaces. We define the adjoint of T to be the unique operator $T^* : \mathcal{H} \to \mathcal{H}$ such that

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle$$

for all $x, y \in \mathcal{H}$. The adjoint satisfies the following properties:

- $(T + \lambda S)^* = T^* + \overline{\lambda} S^*$;
- $T^{**} = T$;
- $(R \circ T)^* = T^* \circ R^*$;
- if T is invertible, then $\left(T^{-1}\right)^* = (T^*)^{-1}$;
- $\|T^*\|_{op} = \|T\|_{op}$;
- $\|T^*T\|_{op} = \|T\|_{op}^2$ (known as the C^*-property).

There are a variety of topologies one can place on the space $\mathbb{B}(\mathcal{H})$. We detail three.

Definition D.5.4. Let $(T_\alpha)_\alpha$ be a net in $\mathbb{B}(\mathcal{H})$.

- We say $(T_\alpha)_\alpha \xrightarrow{\|\cdot\|_{op}} T$ if $\|T_\alpha - T\|_{op} \to 0$. This is the norm topology on $\mathbb{B}(\mathcal{H})$.
- We say $(T_\alpha)_\alpha \xrightarrow{SOT} T$ if, for all $\xi \in \mathcal{H}$, $\|T_\alpha(\xi) - T(\xi)\| \to 0$. This is the strong operator topology (or topology of pointwise convergence) on $\mathbb{B}(\mathcal{H})$.

- We say $(T_\alpha)_\alpha \xrightarrow{\text{WOT}} T$ if, for all $\xi, \eta \in \mathcal{H}$, $\langle T_\alpha(\xi), \eta\rangle \to \langle T(\xi), \eta\rangle$. This is the weak operator topology on $\mathbb{B}(\mathcal{H})$.

Definition D.5.5.

- We say $T \in \mathbb{B}(\mathcal{H})$ is normal if $T^*T = TT^*$.
- We say $T \in \mathbb{B}(\mathcal{H})$ is self-adjoint if $T^* = T$. We write $\mathbb{B}(\mathcal{H})_{\text{s.a.}}$ to refer to the set of all self-adjoint operators in $\mathbb{B}(\mathcal{H})$.
- We say $P \in \mathbb{B}(\mathcal{H})$ is a projection if $P^2 = P^* = P$.
- We say $V \in \mathbb{B}(\mathcal{H})$ is an isometry if $V^*V = I_{\mathcal{H}}$.
- We say $T \in \mathbb{B}(\mathcal{H})$ is a *partial* isometry if $TT^*T = T$.
- We say $U \in \mathbb{B}(\mathcal{H})$ is a unitary if $U^* = U^{-1}$.

We write $\mathcal{U}(\mathcal{H})$ to refer to the set of all unitary operators on $\mathcal{H}$. Two operators $T, S \in \mathbb{B}(\mathcal{H})$ are called unitarily equivalent if there is $U \in \mathcal{U}(\mathcal{H})$ such that $UTU^* = S$.

The space of unitary operators, $\mathcal{U}(\mathcal{H})$, is a group with respect to operator composition.

The set $\mathbb{B}(\mathcal{H})_{\text{s.a.}}$ admits an order structure.

Definition D.5.6. Let $T \in \mathbb{B}(\mathcal{H})_{\text{s.a.}}$. We say T is positive if, for every $\xi \in \mathcal{H}$, we have

$$\langle T(\xi), \xi\rangle \geqslant 0.$$

We write $\mathbb{B}(\mathcal{H})_+$ to refer to the operator norm-closed cone of positive operators in $\mathbb{B}(\mathcal{H})_{\text{s.a.}}$.

If $T, S \in \mathbb{B}(\mathcal{H})_{\text{s.a.}}$, we say $T \geqslant S$ if $T - S \in \mathbb{B}(\mathcal{H})_{\text{s.a.}}$.

Remark D.5.2. It can be shown that an operator $T \in \mathbb{B}(\mathcal{H})_+$ if and only if there is some $S \in \mathbb{B}(\mathcal{H})$ such that $T = S^*S$.

Definition D.5.7. Let $x, y \in \mathcal{H}$. We define the rank-one bounded operator $\theta_{x,y} : \mathcal{H} \to \mathcal{H}$ by

$$\theta_{x,y}(z) = \langle z, y\rangle x.$$

If $T \in \mathbb{B}(\mathcal{H})$ is such that

$$T = \sum_{j=1}^{n} \theta_{x_j, y_j},$$

where $x_j, y_j \in \mathcal{H}$, then T is of finite rank — i.e., $\dim(\mathrm{Ran}(T)) < \infty$. We write $T \in \mathbb{F}(\mathcal{H})$.

A map $T \in \mathbb{B}(\mathcal{H})$ is called compact if T maps bounded sets to sets with compact closure. The space of compact operators is written $\mathbb{K}(\mathcal{H})$.

Theorem D.5.6. The operator norm-closure of the finite rank operators is the compact operators. That is,

$$\overline{\mathbb{F}(\mathcal{H})}^{\|\cdot\|_{\text{op}}} = \mathbb{K}(\mathcal{H}).$$

D.6 Banach Algebras and C^*-Algebras

Earlier, we looked at the definition of an algebra

Definition D.6.1. A Banach $*$-algebra is a complete normed $*$-algebra (see Definition **??**) satisfying $\|a^*\| = a$.

If A is a Banach algebra that satisfies the C*-property, $\|a^* a\| = \|a\|^2$, then A is called a C*-algebra.

There are a variety of distinguished elements in C*-algebras, whose names borrow from their respective names in $\mathbb{B}(\mathcal{H})$.

Bibliography

[Hal66] James D. Halpern. “Bases in vector spaces and the axiom of choice”. In: *Proc. Amer. Math. Soc.* 17 (1966), pp. 670–673. ISSN: 0002-9939,1088-6826. DOI: `10.2307/2035388`. URL: `https://doi.org/10.2307/2035388`.

[Tit72] J Tits. “Free subgroups in linear groups”. In: *Journal of Algebra* 20.2 (1972), pp. 250–270. ISSN: 0021-8693. DOI: `https://doi.org/10.1016/0021-8693(72)90058-0`. URL: `https://www.sciencedirect.com/science/article/pii/0021869372900580`.

[Rud73] Walter Rudin. *Functional analysis*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973, pp. xiii+397.

[Fol84] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). Modern techniques and their applications, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1984, pp. xiv+350. ISBN: 0-471-80958-6.

[Har00] Pierre de la Harpe. *Topics in geometric group theory*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 2000, pp. vi+310. ISBN: 0-226-31719-6.

[Run02] Volker Runde. *Lectures on amenability*. Vol. 1774. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2002, pp. xiv+296. ISBN: 3-540-42852-6. DOI: `10.1007/b82937`. URL: `https://doi.org/10.1007/b82937`.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004, pp. xiv+716. ISBN: 0-521-83378-7. DOI: `10.1017/CBO9780511804441`. URL: `https://doi.org/10.1017/CBO9780511804441`.

[DF04] David S. Dummit and Richard M. Foote. *Abstract algebra*. Third. John Wiley & Sons, Inc., Hoboken, NJ, 2004, pp. xii+932. ISBN: 0-471-43334-9.

[Run05] Volker Runde. *A taste of topology*. Universitext. Springer, New York, 2005, pp. x+176. ISBN: 978-0387-25790-7.

[AB06] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis*. Third. A Hitchhiker's Guide. Springer, Berlin, 2006, pp. xxii+703. ISBN: 978-3-540-32696-0.

[BHV08] Bachir Bekka, Pierre de la Harpe, and Alain Valette. *Kazhdan's property (T)*. Vol. 11. New Mathematical Monographs. Cambridge University Press, Cambridge, 2008, pp. xiv+472. ISBN: 978-0-521-88720-5. DOI: `10.1017/CBO9780511542749`. URL: `https://doi.org/10.1017/CBO9780511542749`.

[Alu09] Paolo Aluffi. *Algebra: Chapter 0*. Vol. 104. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2009, pp. xx+713. ISBN: 978-0-8218-4781-7. DOI: `10.1090/gsm/104`. URL: `https://doi.org/10.1090/gsm/104`.

[Tao09] Terence Tao. *245B, notes 2: Amenability, the ping-pong lemma, and the Banach-Tarski paradox (optional)*. `https://terrytao.wordpress.com/2009/01/08/245b-notes-2-amenability-the-ping-pong-lemma-and-the-banach-tarski-paradox-optional/`. 2009.

[Löh17] Clara Löh. *Geometric group theory*. Universitext. An introduction. Springer, Cham, 2017, pp. xi+389. ISBN: 978-3-319-72253-5. DOI: `10.1007/978-3-319-72254-2`. URL: `https://doi.org/10.1007/978-3-319-72254-2`.

[Run20] Volker Runde. *Amenable Banach algebras*. Springer Monographs in Mathematics. A panorama. Springer-Verlag, New York, 2020, pp. xvii+462. ISBN: 978-1-0716-0351-2. DOI: 10.1007/978-1-0716-0351-2. URL: https://doi.org/10.1007/978-1-0716-0351-2.

[Jus22] Kate Juschenko. *Amenability of discrete groups by examples*. Vol. 266. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2022, pp. xi+165. ISBN: 978-1-4704-7032-6. DOI: 10.1090/surv/266. URL: https://doi.org/10.1090/surv/266.