ON THE DYNAMICS OF ACTIONS ON COMPACT METRIZABLE SPACES

by

Jade Raymond

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2025

Approved by:

Dr. Kevin McGoff

Dr. Will Brian

Dr. Michael Grabchak

Dr. Tony Brizendine

ABSTRACT

JADE RAYMOND. On the dynamics of actions on compact metrizable spaces.
(Under the direction of DR. KEVIN MCGOFF)

The field of dynamical systems is generally interested in the study of collections of transformations acting upon some state space. In this work, we are primarily interested in actions on compact metrizable spaces.

In Chapter 1, we focus our attention on a particular kind of dynamical system, where for a finite set of symbols $\mathcal{A}$ and group G, we consider the space $\mathcal{A}^G$ of labellings of G by symbols in $\mathcal{A}$, and take a canonical action of G on this space by shifting these labellings. A shift is any topological subsystem of this system, and in general, there is a wide variety of such subsystems. Within these shifts there are the subcategories of SFTs, sofic shifts, and strongly irreducible shifts, and in general, these subcategories do not coincide. For locally finite groups, we show that every sofic shift is an SFT, every SFT is strongly irreducible, every strongly irreducible shift is an SFT, every SFT is entropy minimal, and every SFT has a unique measure of maximal entropy, among some other properties which are more complex to state. Furthermore, we show that if any of these properties hold for a group, then the group must be locally finite. These results are collected into two main theorems which characterize the local finiteness of groups by purely dynamical properties. In pursuit of these results we present a formal construction of free extension shifts on a group G, which takes a shift on a subgroup H of G and naturally extends it to a shift on all of G.

In Chapter 2, we turn our attention on studying dynamical systems more generally. For a compact metrizable space X endowed with the Borel σ-algebra, and a collection $\mathcal{T}$ of Borel measurable transformations from X to itself, the pair $(X, \mathcal{T})$ is an arbitrary dynamical system. A principle goal in ergodic theory is to characterize the invariant measures of dynamical systems, which are probability measures

μ on the Borel sets of X such that for any Borel set $E \subset X$ and every $T \in \mathcal{T}$, we have $\mu(E) = \mu(T^{-1}(E))$. We develop the notion of the completion $(X, \mathcal{T}^*)$ of a dynamical system $(X, \mathcal{T})$, which consists of all transformations which preserve every invariant measure of the original system $(X, \mathcal{T})$. We demonstrate that the collection $\mathcal{T}^*$ is a monoid under composition, contains all inverses of bijections, is stable under wobbling, and is closed in a novel topology on the set of all Borel measurable transformations from X to itself. Using this completion, we define Birkhoff systems, for which a version of the pointwise ergodic theorem holds, and show that most classically studied systems are Birkhoff, despite the definition being stronger than how usual pointwise ergodic theorems are stated. Additionally, we show that a dynamical system is Birkhoff if and only if its completion is Birkhoff, which makes it possible to transfer a known pointwise ergodic theorem from one dynamical system to another. For Birkhoff systems, we define the notion of the dynamical independence of two sets A and B, and prove that an invariant measure is ergodic if and only if whenever A is dynamically independent of B, then A must be probabilistically independent of B. This result enables the identification of independence structures within ergodic measures from purely dynamical concepts, and is useful in characterizing the ergodic invariant measures. Finally, we use these concepts in conjunction to study permutation systems, semicontractible systems, products of systems, joinings of systems, as well as power systems. These tools also enable a proof and extension of De Finetti's Theorem which characterizes the semicontractible systems for which the conclusion of De Finetti's Theorem holds. This is further extended to semicontractible power systems, which gives a broad class of generalizations of De Finetti's Theorem where further restrictions on the distributions are considered.

DEDICATION

This dissertation is dedicated to my dear husband Milo, whose love, support, encouragement, and delicious food was vital to making it possible for me to commit so much time and energy towards this work.

ACKNOWLEDGEMENTS

I would like to acknowledge, first and foremost, my advisor Dr. Kevin McGoff for his mentorship throughout the years. His guidance and support has been vital in my growth as a professional, and the opportunities which I have been afforded as his student has enabled me to be highly technical with a constant mind towards application. I would also like to thank the remaining members of my committee, Dr. William Brian, Dr. Michael Grabchak, and Dr. Tony Brizendine, for their helpful comments and direction. I also acknowledge the University of North Carolina at Charlotte for 5 years of GASP tuition and healthcare support, as well as the Department of Mathematics and Statistics for providing me with a Graduate Teaching Assistantship for the entirety of my tenure at UNC Charlotte, which has permitted me to focus on my degree.

TABLE OF CONTENTS

PREFACE

This document is organized into two chapters, each of which is intended to be treated at two separate works. There are relationships between the two, however they each use their own notation and terminology, which is presented individually in each chapter. Chapter 1 is work that has been published (as a solo author) in the journal Ergodic Theory and Dynamical Systems [44]. Chapter 2 is unpublished work which has been in development since the work of Chapter 1 was completed.

CHAPTER 1: Shifts of Finite Type on Locally Finite Groups

For a finite set of symbols $\mathcal{A}$ and a group G, the field of symbolic dynamics studies the action of G by translations on the set $\mathcal{A}^G$, called *full G-shift with alphabet* $\mathcal{A}$, and the subsystems within. Equipped with the product topology (with the discrete topology on $\mathcal{A}$), a closed, translation invariant subset of $\mathcal{A}^G$ is called a *G-shift*, and understanding what properties such subsystems can exhibit is central to symbolic dynamics. In its conception, the primary group of interest was $\mathbb{Z}$, the group of integers under addition. Even in this case, complex behavior arises, though much is known in general about shifts on $\mathbb{Z}$ [36]. A natural extension of this case is the group $\mathbb{Z}^d$ for some natural number d, the study of which has been called multi-dimensional symbolic dynamics. More recently, interest in shifts on $\mathbb{Z}^d$ has grown, though this case already adds much complexity [46, 38, 28], and less is known about $\mathbb{Z}^d$-shifts in general. Interest in the general group case is even more recent, and as may be expected, is even less tractable than the case of $\mathbb{Z}^d$, though a recent result about tilings of amenable groups [17] has made a few results about shifts on amenable groups possible [19, 11, 10].

The class of *G-shifts of finite type*, or G-SFTs, are of particular interest, as they are characterized by a finite amount of information. More precisely, a G-SFT X is a G-shift for which there is a finite collection of *patterns* (an element of $\mathcal{A}^F$ for a finite $F \subset G$) so that X is the collection of all *configurations* in $\mathcal{A}^G$ for which these patterns never appear. The finite nature of G-SFTs makes them amenable to analysis using finitary and combinatorial methods, and in general G-SFTs are well behaved in comparison to general shifts. Furthermore, every shift on a group can be represented as an intersection of SFTs, so in this sense, SFTs are plentiful and are good approximations for shifts in general. Formal definitions of G-shifts and G-SFTs can be found in Section 1.1.2.

Understanding what properties are possible for SFTs on groups is at the core of symbolic dynamics. One such property is the *entropy* (Definition 1.1.14) of an SFT on a countable amenable group G, or in particular, the set of entropies which are attainable by SFTs on G, which is denoted $\mathcal{E}(G)$. $\mathcal{E}(\mathbb{Z})$ was classified by Lind [35], and more recently, $\mathcal{E}(\mathbb{Z}^d)$ for $d \geq 2$ was classified by Hochman and Meyerovitch [29]. Recent results by Barbieri [5] classifies $\mathcal{E}(G)$ as $\mathcal{E}(\mathbb{Z}^d)$ for a certain class of amenable groups. Currently to the knowledge of the author, there are no known finitely generated groups G for which $\mathcal{E}(G)$ does not coincide with either $\mathcal{E}(\mathbb{Z})$ or $\mathcal{E}(\mathbb{Z}^2)$, and further classifying $\mathcal{E}(G)$ for other groups and classes of groups is an open goal in symbolic dynamics. Another property is *strong irreducibility* (Definition 1.1.9), which loosely gives that a G-shift is large, and contains a large variety of configurations. In general, a G-SFT need not be strongly irreducible, and a strongly irreducible G-shift need not be a G-SFT. The additional structure which strong irreducibility imposes on a shift has been useful in proving results about shifts [14, 41, 10]. We also explore several other properties of shifts, which are outlined in Section 1.1, and discussed informally after the statement of our two main theorems below.

Our motivation for studying locally finite groups comes from the following example. Let $G = \bigoplus_{n\in\mathbb{N}} \mathbb{Z}/2\mathbb{Z}$, the countable direct sum of the 2 element group. Elements of G are infinite sequences of 0s and 1s which only contain finitely many 1s, and the group operation is component-wise addition modulo 2. Using elementary methods for computing the entropy on shifts, it is possible to show directly that

$$\mathcal{E}(G) = \left\{ \frac{\log(n)}{2^{m-1}} : n, m \in \mathbb{N} \right\} \subsetneq \mathcal{E}(\mathbb{Z}),$$

providing an example of an infinitely generated group for which $\mathcal{E}(G)$ does not coincide with $\mathcal{E}(\mathbb{Z})$ or $\mathcal{E}(\mathbb{Z}^2)$. In general, classifying the entropies which are attainable by SFTs for a group G is quite difficult, however the process is made tractable for this group by the fact that

$$H_n = \left(\bigoplus_{k=1}^{n} \mathbb{Z}/2\mathbb{Z}\right) \oplus \left(\bigoplus_{k=n+1}^{\infty} \{0\}\right)$$

is a sequence of finite subgroups of G such that $H_n \leq H_{n+1}$ and $G = \bigcup_{n\in\mathbb{N}} H_n$, which makes $\{H_n\}$ a Følner sequence for G. As it turns out, a countable group with such a sequence $\{H_n\}$ of finite subgroups is necessarily *locally finite*. A group is locally finite if every finitely generated subgroup is finite. In fact, any countable locally finite group must have such a sequence of subgroups, and so this property coincides exactly with the property of being locally finite when the group is countable. Locally finite groups naturally extend finite groups in a way that allows for finitary methods to be used when analyzing the groups, despite being possibly infinite. As a result, one may suspect SFTs on locally finite groups are highly structured and have many nice dynamical properties.

The main results of this chapter confirm that SFTs on locally finite groups have very strong dynamical properties. Furthermore, we show that locally finite groups are the only groups for which all SFTs exhibit these properties. These results are grouped in two, one in the case where G is an arbitrary group, and the second where G is a countable amenable group. The first is given below, and followed by a brief explanation of each statement in the result, though formal definitions for every term below can be found in Section 1.1.

Theorem I. *Let G be a group. Then the following are equivalent.*

(a) G is locally finite.

(b) Every G-SFT is the free extension of some SFT on a finite subgroup of G.

(c) Every G-SFT is strongly irreducible.

(d) Every strongly irreducible G-shift is a G-SFT.

(e) Every sofic G-shift is a G-SFT.

(f) For every G-SFT X, Aut(X) *is locally finite.*

Statement I(b) is not a typical dynamical property, but involves a specific type of shift defined in Section 1.2 called a *free extension* shift. Free extensions shifts are by no means a new concept and have been used in the past [29, 5], however we present a formal construction and derive many useful properties of free extensions, some of which are new to the knowledge of the author. The equivalence between statement I(b) and I(a) is at the core of nearly every argument involved in proving this theorem and the next. Free extension shifts are defined for general groups in Section 1.2, and may be useful in studying shifts on groups in general, beyond the study of shifts on locally finite groups. Statement I(e) involves sofic shifts, which are the image of SFTs under continuous, shift invariant factor maps. Along with SFTs, sofic shifts are a noteworthy class of shifts which are defined by a finite amount of information. Every SFT is necessarily sofic, however the converse does not hold in general, and Theorem I gives that the converse holds only in the case that the group is locally finite. The definition of factor maps and sofic shifts can be found in Section 1.1.2.3. Statement I(c) gives that every SFT on a locally finte group is strongly irreducibile. A formal definition is given by Definition 1.1.9, but informally, strong irreducibility is a property which guarantees that for any two elements in the shift, there exists an element of the shift which is equal to one of the elements on a finite subset, and equal to the other on any sufficiently separated finite subset. In this sense, strongly irreducible shifts are rich with configurations. Statement I(d) is the converse of the previous statement, and is independently equivalent to the group being locally finite. These two statements in combination give that the set of G-SFTs and the set of strongly irreducible G-shifts coincides exactly when G is locally finite, but that neither is contained in the other when G is not locally finite. Statement I(f) involves $\mathrm{Aut}(X)$, the automorphism group of an SFT X. This group consists of homeomorphisms from X to itself which preserve the action of G, and is formally defined in Section 1.1.2.3.

For the second result, we restrict to the case that G is a countable amenable group, which permits the development of topological entropy, and each of the statements in the result involves this entropy. A brief discussion of each statement follows the statement of the result, and the formal definitions of every term can be found in Section 1.1. In statement II(d), we use the non-standard notation $H \ll G$ to denote that H is a *finite* subgroup of G.

Theorem II. *Let G be a countable amenable group. Then the following are equivalent.*

(a) G is locally finite

(b) If X is a nonempty G-SFT with $h(X) = 0$, then $X = \{x\}$, where x is a fixed point.

(c) Every G-SFT is entropy minimal.

(d) G is locally non-torsion, and

$$\mathcal{E}(G) = \left\{ \frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N} \right\} \subset \mathbb{Q}^{+}_{\log} = \left\{ \frac{\log(n)}{m} : n, m \in \mathbb{N} \right\}.$$

(e) Every G-SFT has a unique measure of maximal entropy.

We remark that while we restrict the results to countable amenable groups, entropy can also be extended to the more general class of countable sofic groups [12], however we will not need this more general definition, since any countable locally finite group is necessarily amenable. The definition of entropy can be found in Section 1.1.2.4.

Statement II(b) is about what sorts of zero topological entropy SFTs can exist, and in the case of locally finite groups, there is a single zero-entropy SFT (up to conjugacy). This result indirectly answers a question of Barbieri in the affirmative:

Question 1.0.1 (3.19 [5])**.** Does there exist an amenable group G and a G-SFT which does not contain a zero-entropy G-SFT?

Since the only 0 entropy SFTs on locally finite groups are single fixed points, it suffices to construct an SFT which contains no fixed points, which is trivial to do using free extensions. There is further discussion about this construction in Section 1.4.

Statement II(c) involves entropy minimality, which is the property that a shift has no proper subshift with the same entropy as the entire shift. A formal definition is given by Definition 1.1.16. Statement II(d) consists of two parts. The first, is that G is locally non-torsion, and means that every finitely generated subgroup of G is either finite, or contains an element of infinite order. The need for this requirement in this statement is discussed further in Sections 1.3.2.3 and 1.4. The second part of the statement classifies the set of entropies attainable on any locally finite group, and gives the following corollary which may be of independent interest to the remainder of the Theorem.

Corollary. *Let G be a countable locally finite group. Then*

$$\mathcal{E}(G) = \left\{ \frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N} \right\}$$

Finally, the last statement in the theorem, Statement II(e), involves measure theoretic entropy and measures of maximal entropy, which are invariant measures on the SFT that have a measure theoretic entropy equal to the topological entropy of the system. Formal definitions for these can be found starting at Definition 1.1.20.

1.0.1 Overview

In Section 1.1, we present the relevant background and notation used in the remainder of the chapter. In Section 1.2, we define free extension shifts generally for groups, and then prove some properties of these shifts. In Section 1.3 we prove Theorems I and II, which is broken down into several individual lemmas. Finally, in Section 1.4, we discuss some general consequences of Theorems I and II and properties of free extensions, and indicate possible directions for future work.

1.1 Definitions and Notation

We begin with defining all necessary background terms and notation. The section is broken up into subsections based on what is being defined.

1.1.1 Sets and Groups

For any set A, let $B \Subset A$ denote that B is a *finite* subset of A. The set difference of two sets A and B is denoted by $A \setminus B$. The disjoint union of two sets A and B is denoted by $A \sqcup B$. The symmetric difference of two sets A and B is denoted by $A \triangle B$

Given two sets A and B, the set A^B refers to the collection of all functions $f : B \to A$. If A is endowed with a topology, then A^B is endowed with the product topology.

For a group G, we denote that a subset $H \subset G$ is a subgroup of G by $H \leq G$, and to additionally specify that H is a finite subgroup of G, we use the notation $H \lll G$. For $F \subset G$, the subgroup of G generated by F is denoted as $\langle F \rangle$. A group is *periodic* if all of its elements have finite order, and is *torsion* if it is periodic and infinite. This definition of a torsion group is non-standard, as typically the terms torsion and periodic are equivalent, however we require the distinction between arbitrary periodic groups and infinite periodic groups. If P is a property which a group can posses, then a group G is said to be *locally* P if $\forall F \Subset G$, the subgroup $\langle F \rangle$ has property P. A group G is then *locally finite* if $\forall F \Subset G$ we have $\langle F \rangle \lll G$, that is, every finitely generated subgroup of G is finite. G is *locally non-torsion* if $\forall F \Subset G$, the subgroup $\langle F \rangle$ is non-torsion, or in other words, either finite or not periodic. In addition, we use the terminology *non-locally finite* to mean that a group is not locally finite (and similarly for non-locally non-torsion).

Given a group G and subgroup $H \leq G$, we denote the set of *right* cosets of H in G by $H \backslash G$. This notation is similar to the one used for set difference (though the spacing is different), however which is being referred to is generally clear from context.

A countable group G is *amenable* if there exists a sequence $\{F_n\}_{n=1}^{\infty}$ such that $F_n \Subset G$, $\{F_n\}$ exhausts G so that $G = \bigcup_n F_n$, and $\forall g \in G$,

$$\lim_{n\to\infty} \frac{|gF_n \triangle F_n|}{|F_n|} = 0.$$

Such a sequence is called a *left Følner sequence*, and similarly, *right Følner sequences* exist for amenable groups, which satisfy

$$\lim_{n\to\infty} \frac{|F_n g \triangle F_n|}{|F_n|} = 0.$$

1.1.2 G-shifts

For the remainder of the section, $\mathcal{A}$ is a finite *alphabet* (set), endowed with the discrete topology.

Definition 1.1.1. Let G be a group. $\mathcal{A}^G$ is endowed with the product topology, which makes it a compact Hausdorff space. When G is countable $\mathcal{A}^G$ is metrizable, and we will take this fact to be evident when G is countable. For any $g \in G$, define $\sigma^g : \mathcal{A}^G \to \mathcal{A}^G$ by

$$(\sigma^g x)(h) = x(hg)$$

for any $h \in G$. Each σ^g is a homeomorphism from $\mathcal{A}^G$ to $\mathcal{A}^G$, and $\sigma^g \circ \sigma^h = \sigma^{gh}$ for all $g, h \in G$. Also, $\sigma^e x = x$ for each $x \in X$, and therefore the collection $\sigma = \{\sigma^g : g \in G\}$ is a continuous action of G on $\mathcal{A}^G$. The pair $(\mathcal{A}^G, \sigma)$ is called the *full G-shift with alphabet* $\mathcal{A}$, or simply the *full G-shift* when the alphabet $\mathcal{A}$ is clear, which is typically the case. The elements of $\mathcal{A}^G$ are referred to as *configurations*.

Though the full G-shift is interesting in its own right, we are primarily interested in subsystems of the full G-shift, which are called *G-shifts*.

Definition 1.1.2. Let G be a group. A subset $X \subset \mathcal{A}^G$ is said to be *G-invariant*, or merely *shift invariant* when the group G is clear from context, if for every $x \in X$, and $g \in G$, $\sigma^g x \in X$. A closed, G-invariant subset $X \subset \mathcal{A}^G$, along with the action of G on $\mathcal{A}^G$ restricted to X, is called a *G-shift of $\mathcal{A}^G$*, or just *G-shift* when the full shift is clear from context.

1.1.2.1 Patterns

Although an element of $\mathcal{A}^G$ is known as a configuration, the term *pattern* is used when considering elements of $\mathcal{A}^F$ for some $F \subset G$. In addition, we define a few operations on patterns that are quite useful when working with shifts.

Definition 1.1.3. For any group G and $F \subset G$, an element $w \in \mathcal{A}^F$ is called an *$\mathcal{A}$-pattern on F*, or just a *pattern* if $\mathcal{A}$ is clear. The *shape* of a pattern $w \in \mathcal{A}^F$ is the set F itself.

For $E, F \subset G$ and patterns $w \in \mathcal{A}^E$ and $v \in \mathcal{A}^F$, we say w and v are disjoint if E and F are disjoint. Similarly, $w \in \mathcal{A}^F$ is said to be finite if F is finite, and infinite if F is infinite.

For any $E \subset F \subset G$ (including $E = F = G$), the restriction of a pattern $w \in \mathcal{A}^F$ to E, which is denoted $w|_E$ and contained in $\mathcal{A}^E$, is defined as $w|_E(g) = w(g)$ for every $g \in E$. Conversely, for some $w \in \mathcal{A}^E$, the set of *F-extensions of w* is defined as

$$[w]_F = \left\{ v \in \mathcal{A}^F : v|_E = w \right\}.$$

In the case that $F \leq G$, then $[w]_F$ is known as a *cylinder set*. In the case that $F = G$, then $[w]$ is used instead of $[w]_G$, unless clarity is necessary.

Patterns are very useful in describing the structure of G-shifts. For any G-shift X (including the full G-shift), the set

$$\mathfrak{B} = \left\{ [w]_G \cap X : F \Subset G, w \in \mathcal{A}^F \right\}$$

is a basis for the subspace topology on X as a subspace of the full G-shift. Note that $[w]_G \cap X$ may be empty or nonempty, and we define the following sets in order to distinguish when this is or is not the case.

Definition 1.1.4. For any G-shift X, and any $F \subset G$, let $\mathcal{L}_F(X)$ denote the F-*language* of X, which is defined as

$$\mathcal{L}_F(X) = \{x|_F : x \in X\} \subset \mathcal{A}^F.$$

We then let $\mathcal{L}(X)$ be the *language* of X, which is defined as

$$\mathcal{L}(X) = \bigsqcup_{F \Subset G} \mathcal{L}_F(X).$$

By this definition, note that $w \in \mathcal{L}(X)$ if and only if $[w]_G \cap X \neq \varnothing$. In addition, let $\mathcal{L}^\infty(X)$ denote the set

$$\mathcal{L}^\infty(X) = \bigsqcup_{F \subset G} \mathcal{L}_F(X).$$

The main difference between this and $\mathcal{L}(X)$ is that $\mathcal{L}^\infty(X)$ also contains infinite patterns.

We also let $\mathcal{F}_F(X) = \mathcal{A}^F \setminus \mathcal{L}_F(X)$, and

$$\mathcal{F}(X) = \bigsqcup_{F \Subset G} \mathcal{F}_F(X).$$

These sets are known as the *forbidden F-patterns* of X and the *forbidden patterns* of X, respectively.

In constructions which appear in Section 1.2, we utilize an extension of the shift action σ to $\mathcal{L}^\infty(X)$, as well as a joining operation which allows taking two disjoint patterns and combining them into one pattern. These are defined next.

Definition 1.1.5. Let G be a group, and X be a G-shift. Let $g \in G$. Then for any

$F \subset G$, define $\sigma_F^g : \mathcal{L}_F(X) \to \mathcal{L}_{Fg^{-1}}(X)$ by

$$(\sigma_F^g w)(h) = w(hg), \quad \forall h \in Fg^{-1}.$$

Note that in the case $F = G$, this covers the typical shift maps. We then define $\sigma^g : \mathcal{L}^\infty(X) \to \mathcal{L}^\infty(X)$ for any $F \subset G$ and pattern $w \in \mathcal{A}^F$ as

$$\sigma^g w = \sigma_F^g w.$$

Restricting patterns to subshapes and shifting behave well in relation to each other. Let $E \subset F \subset G$, and let $g \in G$. Then for any $w \in \mathcal{L}_{Eg}(X)$, the pattern $\sigma^g w$ has shape $Egg^{-1} = E$, and for any $h \in E$,

$$\sigma^g(w|_{Eg})(h) = (w|_{Eg})(hg) = w(hg) = (\sigma^g w)(h).$$

Since this holds for any $h \in E$, it follows that

$$\sigma^g(w|_{Eg}) = (\sigma^g w)|_E.$$

This rule is used in many proofs without reference.

Similar interplay exists between the shifts and extension sets. Let $E \subset F \subset G$, and $g \in G$. Then $Eg \subset Fg$, and for any $w \in \mathcal{L}_{Eg}(X)$,

$$\sigma^g[w]_{Fg} = [\sigma^g w]_F.$$

This is also used in many proofs without reference.

Along with this natural notion of shifting patterns, there is a natural way to define joining two disjoint patterns.

Definition 1.1.6. Let G be a group, and X be a G-shift. For any disjoint $u, v \in \mathcal{L}^\infty(X)$, with shapes F_u and F_v respectively (so that $F_u \cap F_v = \varnothing$), we define the *join* of u and v, denoted by $u \vee v$, as follows. Let $w = u \vee v$ be defined as

$$w(g) = \begin{cases} u(g), & g \in F_u \\ v(g), & g \in F_v \end{cases},$$

which is a pattern with shape $F_u \sqcup F_v$. Since F_u and F_v must be disjoint to take a join, it is clear that $\vee$ is commutative.

Additionally, the shift action distributes over $\vee$. For any disjoint $u, v \in \mathcal{L}^\infty(X)$ and $g \in G$, it is always the case that $\sigma^g(u \vee v) = (\sigma^g u) \vee (\sigma^g v)$.

Furthermore, for any infinite collection of mutually disjoint patterns, all of these patterns can be joined together into one (possibly infinite) pattern, and by this commutativity, the order of the infinite join is irrelevant. Also, the shifts commute with infinite joins for similar reasons. Infinite joins and the commutativity of the shifts with infinite joins are an integral part of several proofs in Section 1.2.

1.1.2.2 Properties of G-shifts

Each G-shift X defines a set of forbidden patterns, however it is also possible to define a G-shift from a set of forbidden patterns.

Definition 1.1.7. Let G be a group, $\mathcal{A}$ be a finite alphabet, and let $\mathbf{F} \subset \mathcal{L}(\mathcal{A}^G)$ be a set of forbidden patterns. Define

$$\mathcal{X}^G[\mathbf{F}] = \left\{ x \in \mathcal{A}^G : \forall g \in G, \forall F \Subset G,\ (\sigma^g x)|_F \notin \mathbf{F} \right\}$$

It is an elementary exercise to show that $\mathcal{X}^G[\mathbf{F}]$ is a G-shift (though possibly empty), so $\mathcal{X}^G[\mathbf{F}]$ is called the *G-shift defined by* $\mathbf{F}$. $\mathcal{X}[\mathbf{F}]$ is used whenever G is clear from the context.

Another elementary result is that $X = \mathcal{X}[\mathcal{F}(X)]$ for any G-shift X, and therefore every G-shift is generated by some set of finite forbidden patterns.

While $\mathcal{F}(X)$ is always a set of forbidden patterns which defines the G-shift X, there may be much smaller sets of forbidden patterns which also define X. In some cases, there may be a finite set of forbidden patterns which defines a G-shift X, in which case the G-shift is called a *shift of finite type*.

Definition 1.1.8. Let G be a group, and X a G-shift. Then X is called a *G-shift of finite type*, or typically a *G-SFT*, if there exists a finite $\mathbf{F} \Subset \mathcal{L}(X)$ such that $X = \mathcal{X}[\mathbf{F}]$.

For any G-SFT X there always exists some $F \Subset G$ such that $X = \mathcal{X}[\mathcal{F}_F(X)]$. Such a shape F is called a *forbidden shape* for X. Additionally, given some forbidden shape F, any $H \Subset G$ with $F \subset H$ is also a forbidden shape, meaning $X = \mathcal{X}[\mathcal{F}_F(X)] = \mathcal{X}[\mathcal{F}_H(X)]$. This property is used in many results without reference.

The finitary nature of G-SFTs makes them amenable to analysis using more combinatorial methods, and they are generally well behaved in many regards. Another strong property a G-shift can possess is *strong irreducibility*, which is a strong mixing type property that is of general interest in the literature.

Definition 1.1.9. Let G be a group, and X be a G-shift. Then X is *strongly irreducible* if there exists a finite $K \Subset G$ with the following property. For any $u, v \in \mathcal{L}(X)$ with shapes F_u and F_v, if $F_u \cap KF_v = \varnothing$, then there exists $x \in X$ such that $x|_{F_u} = u$ and $x|_{F_v} = v$.

This definition differs from typical definitions of strong irreducibility of shifts on finitely generated groups [19]. In the case that G is finitely generated, this definition is equivalent to more typical definitions using the distance induced by a word metric, and is merely an extension of the more typical definition to (possibly) infinitely generated groups.

1.1.2.3 Factors and Sofic shifts

We begin with the definition of factor maps on shift spaces.

Definition 1.1.10. Let $\mathcal{A}$ and $\mathcal{B}$ be two finite alphabets, G be a group, X be a G-shift of $\mathcal{A}^G$ and Y be a G-shift of $\mathcal{B}^G$. Then a map $\phi : X \to Y$ is a *factor map* if

- ϕ is continuous,
- ϕ is surjective, and
- for every $g \in G$, $\sigma^g \circ \phi = \phi \circ \sigma^g$.

In the case that a factor map ϕ is a homeomorphism, then ϕ is called a *conjugacy*, and X and Y are said to be *conjugate*. The collection of conjugacies from a G-shift X to itself forms a group under composition denoted $\mathrm{Aut}(X)$.

This definition of a factor map applies more generally between actions of a group on two topological spaces, however in the context of G-shifts however, factor maps have a very specific structure. We begin by defining a specific kind of factor map which can be constructed between two G-shifts.

Definition 1.1.11. Let $\mathcal{A}$ and $\mathcal{B}$ be two finite alphabets, let G be a group, and let X be a G-shift of $\mathcal{A}^G$. For some $F \Subset G$, let $\beta : \mathcal{L}_F(X) \to \mathcal{B}$ be any function, called a *block map*. Then β induces a map $\phi_\beta^G : X \to \mathcal{B}^G$ called a *block code* by

$$\big(\phi_\beta^G(x)\big)(g) = \beta\big((\sigma^g x)|_F\big),$$

and $Y = \phi_\beta^G(X)$ is a G-shift of $\mathcal{B}^G$. Rather than $\mathcal{B}^G$ however, we consider the co-domain of ϕ_β^G to be Y, which makes ϕ_β^G surjective and therefore a factor map from X to Y.

Block codes are generally easy to work with, due to the finitary nature of the block map that generates them. Surprisingly, any factor map between G-shifts (on possibly different alphabets) is a block code generated by some block map, and this fact is given by the following theorem.

Theorem 1.1.12 (Curtis-Hedlund-Lyndon)**.** *Let G be a group, $\mathcal{A}$ and $\mathcal{B}$ be finite alphabets, X be a G-shift of $\mathcal{A}^G$ and Y a G-shift of $\mathcal{B}^G$, and let $\phi : X \to Y$. Then ϕ is a factor map if and only if there exists $F \Subset G$ and block map $\beta : \mathcal{L}_F(X) \to \mathcal{B}$ such that $\phi = \phi_\beta^G$.*

A proof of the theorem at this level of generality can be found in [51, Corollary 6]. Informally, the theorem gives that factor maps for G-shifts are defined by a finite amount of information. A broader class of G-shifts which are defined by a finite amount of information, which contains all SFTs but generally includes more shifts, is the class of sofic G-shifts.

Definition 1.1.13. A G-shift Y is called a *sofic G-shift* if there exists a G-SFT X such that Y is a factor of X.

Weiss noted when first introducing sofic $\mathbb{Z}$-shifts that "the finite type subshifts are flawed by not being closed under the simplest operation, namely that of taking [factors]" [53]. The collection of all sofic shifts is clearly closed under taking factors, and this is one of the many reasons the class of sofic shifts is of interest in symbolic dynamics.

1.1.2.4 Entropy

Another important aspect of shifts which is studied in dynamics is entropy (both topological and measure theoretic), though this theory is generally restricted to countable amenable groups, as computing averages for shifts on non-amenable groups is not possible in general. Notions of entropy do exist for the broader class of countable sofic groups [12], however certain undesirable properties arise from such definitions, such as the potential for the entropy of a factor of a system being higher than the entropy of the system itself [54]. Formal treatment of topological and measure theoretic entropy for G-shifts (and more generally continuous actions of groups on metric spaces), as well as results about these notions of entropy, can be found in [33].

Definition 1.1.14. Let G be a countable amenable group. Then the *(topological) entropy* of a nonempty G-shift X is defined as

$$h(X) = \inf_n \frac{\log\big(|\mathcal{L}_{F_n}(X)|\big)}{|F_n|} = \lim_{n\to\infty} \frac{\log\big(|\mathcal{L}_{F_n}(X)|\big)}{|F_n|},$$

where $\{F_n\}$ is some Følner sequence for G. This limit always exists and is equal to the infimum above [33, Section 9.9]. The entropy of X is also independent of the choice of Følner sequence.

Furthermore, some results pertain to the set of real numbers which are attained as the (topological) entropies for SFTs on a particular group.

Definition 1.1.15. Let G be a countable amenable group. Then let

$$\mathcal{E}(G) = \big\{h(X) : X \text{ a nonempty } G\text{-SFT}\big\}$$

Note that $\mathcal{E}(G)$ is a countable subset of $[0, \infty)$, since there are only countably many G-SFTs for any countable group G. Determining exactly what the set $\mathcal{E}(G)$ is for a given group G is in general quite difficult. A classic result of Lind [35] precisely classifies $\mathcal{E}(\mathbb{Z})$ as non-negative rational multiples of logarithms of Perron numbers. More recently, Hochman and Meyerovich determined that $\mathcal{E}(\mathbb{Z}^d)$ is the set of non-negative upper semi-computable real numbers [29]. For finitely generated amenable groups G with decidable word problem which admit a translation like action by $\mathbb{Z}^2$, recent work by Barbieri [5] has classified $\mathcal{E}(G)$ as the set of non-negative upper semi-computable real numbers.

With entropy, we may also define the following notion of minimality.

Definition 1.1.16. Let G be a countable amenable group, and X a G-shift. Then X is *entropy minimal* if for each subshift $Y \subsetneq X$, we have $h(Y) < h(X)$.

A weaker but related notion of minimality is SFT-entropy minimality.

Definition 1.1.17. Let G be a countable amenable group, and X a G-shift. Then X is *SFT-entropy minimal* if for each SFT $Y \subsetneq X$, we have $h(Y) < h(X)$.

Although in general SFT-entropy minimality is weaker than entropy minimality, they are in fact equivalent if the shift in question is an SFT. Proving this is a fairly standard argument involving approximating subshifts by SFTs, so we omit its proof. This fact is quite useful for proving that an SFT is entropy minimal, as it significantly reduces the amount of shifts to consider when proving entropy minimality.

Along with topological entropy, measure-theoretic entropy can be defined if the shift X is additionally endowed with a Borel probability measure (that is always Radon, since $\mathcal{A}^G$ is metrizable when G is countable) that behaves nicely with the shift action of G.

Definition 1.1.18. Let G be a countable amenable group, and let X be a G-shift. A measure μ on X is *G-invariant* if for any $g \in G$ and measurable $E \subset X$, it is the case that $\mu(\sigma^{g^{-1}} E) = \mu(E)$.

Let $\mathcal{M}(X)$ denote the set of all G-invariant Borel probability measures μ on X.

For a G-shift X and $w \in \mathcal{L}(X)$, $\mu[w]$ is used as a shorthand for $\mu([w] \cap X)$. To define the μ-entropy of X, first an associated partition entropy must be defined.

Definition 1.1.19. Let G be a countable amenable group, X be a G-shift, and $\mu \in \mathcal{M}(X)$. Then, for any $F \Subset G$, the (F, μ)-*entropy* of X is defined as

$$H_\mu(X, F) = -\sum_{w \in \mathcal{L}_F(X)} \mu[w] \log(\mu[w]),$$

where $0 \cdot \log(0)$ is taken to be 0 by convention. The maximum of $H_\mu(X, F)$ over $\mathcal{M}(X)$ is $\log(|\mathcal{L}_F(X)|)$, and is attained only by any $\mu \in \mathcal{M}(X)$ for which $\mu[w] = \frac{1}{|\mathcal{L}_F(X)|}$ for all $w \in \mathcal{L}_F(X)$ [52, Corollary 4.2.1].

With this, the measure theoretic entropy can be defined.

Definition 1.1.20. Let G be a countable amenable group, X be a G-shift, and $\mu \in \mathcal{M}(X)$. Then for any Følner sequence $\{F_n\}_{n=1}^{\infty}$ for G, the *μ-entropy* of X is defined as

$$h_\mu(X) = \inf_n \frac{H_\mu(X, F_n)}{|F_n|} = \lim_{n \to \infty} \frac{H_\mu(X, F_n)}{|F_n|}.$$

As with topological entropy, this limit always exists, is equal to this infimum, and is independent of the choice of Føner sequence [33, Section 9.3]. Furthermore, the Variational Principle [33, Theorem 9.43] gives that

$$h(X) = \sup_{\mu \in \mathcal{M}(X)} h_\mu(X).$$

A measure $\mu \in \mathcal{M}(X)$ satisfying $h(X) = h_\mu(X)$ is called a *measure of maximal entropy*, and for G-shifts, there always exists at least one measure of maximal entropy, since shift actions are expansive and the entropy map $\mu \to h_\mu(X)$ is upper semi-continuous in this case [54].

1.2 Free Extension Shifts

Though the primary purpose of this chapter is to prove that locally finite groups are precisely the groups which exhibit strong dynamical properties for all SFTs, proving many of these properties directly is somewhat tedious. Instead, we develop a general theory of *free extension* shifts, which simplifies (and even trivializes) many of the results for locally finite groups. Essentially all of the primary results in this chapter use properties of free extensions, which are constructed in this section.

The notion of a free extension shift is not new however. Hochman and Meyerovich [29] used them (though not explicitly by name) in their landmark paper characterizing the possible entropies of $\mathbb{Z}^d$ SFTs. The term free extension and some associated notation used were coined by Barbieri [5], with free extensions appearing as a special case of a far more general method of constructing "extensions" of shifts. The definition given here is far less general than Barbieri's construction, but it is perhaps more amenable to specifically analyzing free extensions.

1.2.1 Definition of Free Extensions

Though there are a few equivalent ways of defining free extensions, we use the following as the primary definition, and prove its equivalence to other definitions.

Definition 1.2.1. Let G be a group, $H \leq G$, and Y be an H-shift with alphabet $\mathcal{A}$. Then the *free G-extension* of Y, which is denoted $Y^{\uparrow G}$, is defined as

$$Y^{\uparrow G} = \left\{ x \in \mathcal{A}^G : \forall g \in G, (\sigma^g x)|_H \in Y \right\}.$$

Given a free extension $Y^{\uparrow G}$, we call Y the *base* shift.

While it is clear from this definition that $Y^{\uparrow G}$ is always G-invariant, it is less clear that $Y^{\uparrow G}$ is necessarily closed, and therefore a G-shift. Rather than proving this directly, we use the following lemma to deduce that $Y^{\uparrow G}$ is a G-shift.

Lemma 1.2.2. *Let G be a group, $H \leq G$. Then for any $\mathbf{F} \subset \mathcal{L}(\mathcal{A}^H)$, we have $(\mathcal{X}^H[\mathbf{F}])^{\uparrow G} = \mathcal{X}^G[\mathbf{F}]$.*

Proof. First, we show that $(\mathcal{X}^H[\mathbf{F}])^{\uparrow G} \subset \mathcal{X}^G[\mathbf{F}]$. To do so, let $x \in (\mathcal{X}^H[\mathbf{F}])^{\uparrow G}$, $g \in G$, and $F \Subset G$, and since $\mathbf{F} \subset \mathcal{L}(\mathcal{A}^H)$, we may consider only when $F \Subset H$. By definition, we have that $y = (\sigma^g x)|_H \in \mathcal{X}^H[\mathbf{F}]$, which gives that $y|_F = (\sigma^e y)|_F \notin \mathbf{F}$. With $F \Subset H$, we have $y|_H = \big((\sigma^g x)|_H\big)|_F = (\sigma^g x)|_F \notin \mathbf{F}$. Since x, g, and F were arbitrary, we obtain the desired inclusion.

To show that $\mathcal{X}^G[\mathbf{F}] \subset (\mathcal{X}^H[\mathbf{F}])^{\uparrow G}$, let $x \in \mathcal{X}^G[\mathcal{F}(Y)]$ and $g \in G$, and we must show that $(\sigma^g x)|_H \in \mathcal{X}^H[\mathbf{F}]$. Let $h \in H$ and $F \Subset H$. Then

$$\big(\sigma^h(\sigma^g x)|_H\big)|_F = \big((\sigma^{hg} x)|_H\big)|_F = (\sigma^{hg} x)|_F \notin \mathbf{F},$$

by the fact that $x \in \mathcal{X}^G[\mathcal{F}(Y)]$. As such, $(\sigma^g x)|_H \in \mathcal{X}^H[\mathbf{F}]$, so we have shown the desired result. □

With this lemma, we can easily see that $Y^{\uparrow G}$ is a G-shift.

Corollary 1.2.3. *Let G be a group, $H \leq Y$, and Y an H-shift. Then $Y^{\uparrow G} = \mathcal{X}^G[\mathcal{F}(Y)]$, and in particular $Y^{\uparrow G}$ is a G-shift.*

Proof. Taking $\mathbf{F} = \mathcal{F}(Y)$ in Lemma 1.2.2 gives the desired result, along with noting that $\mathcal{X}^G[\mathcal{F}(Y)]$ is a G-shift. □

In addition to proving that free extensions are shifts, the previous corollary gives an alternative characterization of free extensions, namely the free extension of an H-shift Y is the G-shift defined by the set of finite forbidden patterns for Y. We now provide another characterization of free extensions which is useful in constructing elements of the free extension of a shift. We begin with defining the following function.

Definition 1.2.4. With $H\backslash G$ denoting the set of right cosets of H in G, let $\mathscr{C}(H\backslash G)$ denote the set of all choice sets for $H\backslash G$, whose existence is given by the axiom of choice. In particular, an element $C \in \mathscr{C}(H\backslash G)$ is a subset of G such that $\{Hc\}_{c\in C}$ is an enumeration of the right cosets of H in G.

For any group G, $H \leq G$, $C \in \mathscr{C}(H\backslash G)$, and alphabet $\mathcal{A}$, define a map $\kappa_C^{\mathcal{A}} : (\mathcal{A}^H)^C \to \mathcal{A}^G$ by

$$\kappa_C^{\mathcal{A}}(\{w_c\}_{c\in C}) = \bigvee_{c\in C} \sigma^{c^{-1}} w_c.$$

Such a map is called a *construction* function. When $\mathcal{A}$ is clear from context, κ_C is used instead.

Note that with $\{w_c\}_{c\in C} \in (\mathcal{A}^H)^C$, each w_c has shape H, which makes $\sigma^{c^{-1}} w_c$ have shape Hc. Since $\{Hc\}_{c\in C}$ is an enumeration of all right cosets of H in G, this makes all $\sigma^{c^{-1}} w_c$ disjoint, and so we may join them together to form a configuration on G, giving that κ_C is well defined.

We now show that every construction function is bijective.

Lemma 1.2.5. *Let G be a group, $H \leq G$, and $C \in \mathscr{C}(H\backslash G)$. Then κ_C is a bijection, and $\kappa_C^{-1}(x) = \{(\sigma^c x)|_H\}_{c\in C}$.*

Proof. First, we show that κ_C is injective. Let $\{w_c\}_{c\in C}, \{v_c\}_{c\in C} \in (\mathcal{A}^H)^C$ be such that $\kappa_C(\{w_c\}) = \kappa_C(\{v_c\})$. For any $d \in C$, it must be that $\kappa_C(\{w_c\})|_{Hd} = \kappa_C(\{v_c\})|_{Hd}$. Since

$$\kappa_C(\{w_c\})|_{Hd} = (\bigvee_{c\in C} \sigma^{c^{-1}} w_c)|_{Hd},$$

and each $\sigma^{c^{-1}} w_c$ has shape Hc, this restriction must result in $\sigma^{d^{-1}} w_d$. Similarly, $\kappa_C(\{v_c\})|_{Hd} = \sigma^{d^{-1}} v_d$, and so $\sigma^{d^{-1}} w_d = \sigma^{d^{-1}} v_d$. Applying σ^d to both sides gives that $w_d = v_d$. Since $d \in C$ was arbitrary, we have that $\{w_c\} = \{v_c\}$, and therefore κ_C is injective.

Next, let us show that κ_C is surjective. Let $x \in \mathcal{A}^G$. Then

$$x = \bigvee_{c \in C} x|_{Hc} = \bigvee_{c \in C} \sigma^{c^{-1}}\big(\sigma^c(x|_{Hc})\big)$$
$$= \bigvee_{c \in C} \sigma^{c^{-1}}\big((\sigma^c x)|_H\big) = \kappa_C\big(\{(\sigma^c x)|_H\}_{c \in C}\big),$$

and so κ_C is surjective.

This makes κ_C a bijection, and is therefore invertible, and the previous display gives the exact rule for κ_C^{-1}. □

In fact, with $(\mathcal{A}^H)^C$ endowed with the product topology, κ_C is a homeomorphism from $(\mathcal{A}^H)^C$ to $\mathcal{A}^G$, however we do not use this. For the remainder of the chapter, it will be taken as a given that κ_C is a bijection. With κ_C, we may now give our last characterization of free extensions.

Lemma 1.2.6. *Let G be a group, $H \leq G$, and Y be an H-shift. Then for any $C \in \mathscr{C}(H\backslash G)$, we have $Y^{\uparrow G} = \kappa_C(Y^C)$.*

Proof. First, let $x \in Y^{\uparrow G}$. Then $\kappa_C^{-1}(x) = \{(\sigma^c x)|_H\}_{c \in C}$, and by definition of $Y^{\uparrow G}$, we have $(\sigma^c x)|_H \in Y$ for each $c \in C$, and therefore $\{(\sigma^c x)|_H\}_{c \in C} \in Y^C$, so $x \in \kappa_C(Y^C)$. As such, $Y^{\uparrow G} \subset \kappa_C(Y^C)$.

Now let $x \in \kappa_C(Y^C)$. Then by definition, we have that $\kappa_C^{-1}(x) = \{(\sigma^c x)|_H\}_{c \in C} \in Y^C$, and so $(\sigma^c x)|_H \in Y$ for each $c \in C$. Let $g \in G$. Then there exists a unique $c \in C$ and $h \in H$ such that $g = hc$. Since $(\sigma^c x)|_H \in Y$, by shift invariance we also have $\sigma^h(\sigma^c x)|_H \in Y$, and therefore we have

$$(\sigma^g x)|_H = (\sigma^{hc} x)|_H = \sigma^h(\sigma^c x)|_H \in Y.$$

As $g \in G$ was arbitrary, this gives that $x \in Y^{\uparrow G}$, and therefore $\kappa_C(Y^C) \subset Y^{\uparrow G}$. □

Each of these characterizations provides a useful perspective on the structure of free extensions, and are useful in proving different properties of free extensions. The definition used here indicates that a free extension shift is locally a shift on a subgroup, which is broadly useful in ensuring restrictions of configurations in the free extension are an element of the shift on the subgroup. The second characterization by forbidden patterns clearly makes free extensions a type of shift, and provides an implicit connection between free extensions and SFTs. The final characterization with construction functions gives free extensions a natural strong mixing condition, in the sense that for any $C \in \mathscr{C}(H\backslash G)$ and collection $\{x_c\}_{c\in C} \subset Y^{\uparrow G}$, there exists an element of $Y^{\uparrow G}$ which is equal to x_c on the coset Hc. This strong mixing condition is at the core of the utility of studying free extensions in general.

1.2.2 Properties of Free Extensions and their Base Shifts

Now that we have established that free extensions are well defined, we now prove that many useful properties can be transferred from a base shift to its extension, and vice-versa.

First, we observe that the topological entropy of a free extension is the same as its base shift.

Proposition 1.2.7 (Proposition 5.2, [5])**.** *Let G be a countable amenable group, $H \leq G$, and X an H-shift. Then*

$$h(X^{\uparrow G}) = h(X).$$

Second, if there are 3 groups $K \leq H \leq G$, then taking a K-shift and extending it to G produces the same thing as first extending to H, and then to G.

Lemma 1.2.8. *Let G be a group, $K \leq H \leq G$, and Y a K-shift. Then $Y^{\uparrow G} = (Y^{\uparrow H})^{\uparrow G}$.*

Proof. By Corollary 1.2.3, we have $Y^{\uparrow G} = \mathcal{X}^G[\mathcal{F}(Y)]$ and $Y^{\uparrow H} = \mathcal{X}^H[\mathcal{F}(Y)]$. Then, by Lemma 1.2.2, we have

$$(Y^{\uparrow H})^{\uparrow G} = (\mathcal{X}^H[\mathcal{F}(Y)])^{\uparrow G} = \mathcal{X}^G[\mathcal{F}(Y)] = Y^{\uparrow G}.$$

□

Next, we prove a stability result for free extensions, namely that the intersection of free extensions is the free extension of an intersection of shifts.

Lemma 1.2.9. *Let G be a group, $H \leq G$, and $\{Y_i\}_{i\in I}$ be a collection of H-shifts. Then $\bigcap_{i\in I} Y_i^{\uparrow G} = \left(\bigcap_{i\in I} Y_i\right)^{\uparrow G}$.*

Proof. Let $C \in \mathscr{C}(H\backslash G)$. Then by Lemma 1.2.6 and the fact that κ_C is a bijection,

$$\bigcap_{i\in I} Y_i^{\uparrow G} = \bigcap_{i\in I} \kappa_C(Y_i^C) = \kappa_C\left(\bigcap_{i\in I} Y_i^C\right) = \kappa_C\left(\left(\bigcap_{i\in I} Y_i\right)^C\right) = \left(\bigcap_{i\in I} Y_i\right)^{\uparrow G}.$$

□

Lemma 1.2.2 readily gives that the free extension of an SFT remains an SFT. Perhaps surprisingly, the converse also holds; if the free extension of a shift is an SFT, then the base shift must have been an SFT to begin with.

Lemma 1.2.10. *Let G be a group, $H \leq G$, and Y be an H-shift. Then Y is an H-SFT if and only if $Y^{\uparrow G}$ is a G-SFT.*

Proof. First, suppose Y is an SFT. Let $F \Subset H$ be a forbidden shape for Y, and $\mathbf{F} \Subset \mathcal{A}^F$ be a set of forbidden F-patterns so that $Y = \mathcal{X}^H[\mathbf{F}]$. By Lemma 1.2.2, $Y^{\uparrow G} = (\mathcal{X}^H[\mathbf{F}])^{\uparrow G} = \mathcal{X}^G[\mathbf{F}]$, which is clearly an SFT.

Now suppose that $Y^{\uparrow G}$ is an SFT. Let $C \in \mathscr{C}(H\backslash G)$, and let $d \in C$ such that $H = Hd$. Let $F \Subset G$ be a forbidden shape for $Y^{\uparrow G}$ so that $Y^{\uparrow G} = \mathcal{X}^G[\mathcal{F}_F(Y^{\uparrow G})]$. Now, define $C_0 \subset C$ to be the set of all $c \in C$ for which $F \cap Hc \neq \varnothing$, and for $c \in C_0$, let $F_c = F \cap Hc$. This partitions F into the finitely many disjoint subsets F_c, which are each contained within a separate coset of H. Then define

$$E = \bigcup_{c \in C_0} F_c c^{-1} \subset H,$$

and

$$\hat{F} = \bigcup_{c \in C_0} Ec.$$

Note that for each $c \in C_0$ we have $F_c c^{-1} \subset E$, and therefore $F_c = (F_c c^{-1})c \subset Ec \subset \hat{F}$, so $F \subset \hat{F}$. As such, $Y^{\uparrow G} = \mathcal{X}[\mathcal{F}_{\hat{F}}(Y^{\uparrow G})]$. For any $w \in \mathcal{A}^E$, let

$$P(w) = \bigcup_{c \in C_0} [\sigma^{c^{-1}} w]_{\hat{F}} \subset \mathcal{A}^{\hat{F}},$$

and define

$$\mathbf{F} = \left\{ w \in \mathcal{A}^E : P(w) \subset \mathcal{F}_{\hat{F}}(Y^{\uparrow G}) \right\}.$$

Let us now show that $Y = \mathcal{X}^H[\mathbf{F}]$, which clearly shows Y is a H-SFT, since $\mathbf{F}$ is finite.

First, let $x \in Y^{\uparrow G} = \mathcal{X}^G[\mathcal{F}_{\hat{F}}(Y^{\uparrow G})]$. Let $g \in G$, and pick any $c \in C_0$. Then, from the definition of $P(w)$,

$$(\sigma^{c^{-1}g} x)|_{\hat{F}} \in [(\sigma^{c^{-1}g} x)|_{Ec}]_{\hat{F}} = [\sigma^{c^{-1}} (\sigma^g x)|_E]_{\hat{F}} \subset P\big((\sigma^g x)|_E\big).$$

Since $x \in \mathcal{X}^G[\mathcal{F}_{\hat{F}}(Y^{\uparrow G})]$, it follows that $(\sigma^{c^{-1}g} x)|_{\hat{F}} \notin \mathcal{F}_{\hat{F}}(Y^{\uparrow G})$, and therefore, $P\big((\sigma^g x)|_E\big) \not\subset \mathcal{F}_{\hat{F}}(Y^{\uparrow G})$. This gives that $(\sigma^g x)|_E \notin \mathbf{F}$. Since g was arbitrary, this implies $x \in \mathcal{X}^G[\mathbf{F}]$, and therefore $Y^{\uparrow G} \subset \mathcal{X}^G[\mathbf{F}]$.

Now let $x \in \mathcal{X}^G[\mathbf{F}]$. By definition, for each $g \in G$, it must be that $(\sigma^g x)|_E \notin \mathbf{F}$, and therefore for every $c \in C_0$, we have $P\big((\sigma^{cg} x)|_E\big) \not\subset \mathcal{F}_{\hat{F}}(Y^{\uparrow G})$. As such, for each $c \in C_0$, there exists $w_c \in P\big((\sigma^{cg} x)|_E\big) \setminus \mathcal{F}_{\hat{F}}(Y^{\uparrow G})$. Note that this means for each

$c \in C_0$, there exists $d_c \in C_0$ such that $(\sigma^{d_c} w_c)|_E = (\sigma^{cg} x)|_E$. Let $x_c \in Y^{\uparrow G}$ be such that $x_c|_{\hat{F}} = w_c$. This must be possible since $w_c \notin \mathcal{F}_{\hat{F}}(Y^{\uparrow G})$, and $Y^{\uparrow G}$ is shift invariant. Let $\{a^c_d\}_{d \in C} = \kappa_C^{-1}(x_c)$, and note that $a^c_d \in Y$ for each $c \in C_0$ and $d \in C$ by Lemma 1.2.6. Furthermore, for each $c \in C_0$, we have $a^c_{d_c}|_E = (\sigma^{cg} x)|_E$, which gives that

$$(\sigma^{c^{-1}} a^c_{d_c})|_{Ec} = \sigma^{c^{-1}}(a^c_{d_c}|_E) = \sigma^{c^{-1}}\big((\sigma^{cg} x)|_E\big) = (\sigma^g x)|_{Ec}.$$

Define $\{y_d\}_{d \in C} \in Y^C$ as follows. For each $c \in C_0$, let $y_c = a^c_{d_c}$, and for $d \in C \setminus C_0$, let $y_d \in Y$ (it does not matter how these are chosen). Then $y = \kappa_C(\{y_d\}) \in Y^{\uparrow G}$, so for each $c \in C_0$, it is the case that $y|_{Ec} = (\sigma^{c^{-1}} a_{c,d_c})|_{Ec} = (\sigma^g x)|_{Ec}$, which gives that $y|_{\hat{F}} = (\sigma^g x)|_{\hat{F}}$. Since $y \in Y^{\uparrow G}$, this implies $(\sigma^g x)|_{\hat{F}} \notin \mathcal{F}_{\hat{F}}(Y^{\uparrow G})$. As this is true for all $g \in G$, this implies $x \in \mathcal{X}[\mathcal{F}_{\hat{F}}(Y^{\uparrow G})] = Y^{\uparrow G}$, and therefore $\mathcal{X}^G[\mathbf{F}] \subset Y^{\uparrow G}$.

The results of the two previous paragraphs gives that $Y^{\uparrow G} = \mathcal{X}^G[\mathbf{F}]$. By Lemma 1.2.2, since $\mathbf{F} \subset \mathcal{A}^E$, and $E \subset H$, we have $Y^{\uparrow G} = \mathcal{X}^G[\mathbf{F}] = (\mathcal{X}^H[\mathbf{F}])^{\uparrow G}$, and so $\kappa_C(Y^C) = \kappa_C\big((\mathcal{X}^H[\mathbf{F}])^C\big)$. Since κ_C is a bijection, it must be that $Y^C = (\mathcal{X}^H[\mathbf{F}])^C$, and therefore $\mathcal{X}^H[\mathbf{F}] = Y$. $\square$

Next, we show a similar result to the previous one, replacing the property of being a G-SFT with being strongly irreducible.

Lemma 1.2.11. *Let G be a group, $H \leq G$, and Y be an H-shift. Then Y is strongly irreducible if and only if $Y^{\uparrow G}$ is strongly irreducible.*

Proof. First, suppose that Y is strongly irreducible. Then there exists $K \Subset H$ such that for any $u, v \in \mathcal{L}(Y)$ with shapes F_u and F_v such that $F_u \cap KF_v = \varnothing$, then there exists $x \in Y^{\uparrow G}$ such that $x|_{F_u} = u$ and $x|_{F_v} = v$. We will now show that $X = Y^{\uparrow G}$ is strongly irreducible with the same K. Let $u, v \in \mathcal{L}(X)$ with shapes F_u and F_v such that $F_u \cap KF_v = \varnothing$. Since $u, v \in \mathcal{L}(X)$, let $x_u, x_v \in X$ such that $x_u|_{F_u} = u$ and $x_v|_{F_v} = v$. Let $C \in \mathscr{C}(H \backslash G)$. Since $x_u, x_v \in X = Y^{\uparrow G}$, we may take $\{y_c\}_{c \in C} = \kappa_C^{-1}(x_u)$ and $\{z_c\}_{c \in C} = \kappa_C^{-1}(x_v)$ with $y_c, z_c \in Y$ for all $c \in C$ by Lemma

1.2.6. Let $c \in C$, and define $U_c = F_u c^{-1} \cap H$ and $V_c = F_v c^{-1} \cap H$. With U_c and V_c being subsets of H, we have that $y_c|_{U_c} \in \mathcal{L}(Y)$ and $z_c|_{V_c} \in \mathcal{L}(Y)$. Furthermore, we have

$$U_c \cap KV_c \subset F_u c^{-1} \cap KF_v c^{-1} = (F_u \cap KF_v)c^{-1} = \varnothing.$$

As such, the strong irreducibility of Y gives that there exists $w_c \in Y$ such that $w_c|_{U_c} = y_c|_{U_c}$ and $w_c|_{V_c} = z_c|_{V_c}$. Let $x = \kappa_C(\{w_c\}_{c\in C}) \in X$. We now show that $x|_{F_u} = u$ and $x|_{F_v} = v$. Indeed, as

$$F_u = \bigsqcup_{c\in C} F_u \cap Hc = \bigsqcup_{c\in C} (F_u c^{-1} \cap H)c = \bigsqcup_{c\in C} U_c c,$$

and similarly for F_v with V_c in place of U_c, we may obtain that $x|_{F_u} = u$ by checking $x|_{U_c c} = u|_{U_c c}$ for every $c \in C$ (and similarly for $x|_{F_v} = v$). For $c \in C$, we have $U_c c \subset Hc$, and so

$$\begin{aligned} x|_{U_c c} = (x|_{Hc})|_{U_c c} = (\sigma^{c^{-1}} x_c)|_{U_c c} = \sigma^{c^{-1}}(x_c|_{U_c}) = \sigma^{c^{-1}}(y_c|_{U_c}) \\ = \sigma^{c^{-1}}\big((\sigma^c x_u)|_{U_c}\big) = \sigma^{c^{-1}}\sigma^c(x_u|_{U_c c}) = u|_{U_c c}, \end{aligned}$$

where the final equality follows from the fact that $x_u|_{F_u} = u$. By the same argument, replacing U_c with V_c, y_c with z_c, x_u with x_v, and u with v, we obtain $x|_{V_c c} = v|_{V_c c}$, and therefore $x|_{F_u} = u$ and $x|_{F_v} = v$. As such, $Y^{\uparrow G}$ is strongly irreducible.

Now, suppose that $X = Y^{\uparrow G}$ is strongly irreducible. Let $L \Subset G$ be such that if $u, v \in \mathcal{L}(X)$ with shapes E_u and E_v satisfy $E_u \cap LE_v = \varnothing$, then there exists $x \in X$ such that $x|_{E_u} = u$ and $x|_{E_v} = v$. Let $K = L \cap H$, and we will now show that Y is strongly irreducible with this K. Note that K must be nonempty because L must contain an element of H (otherwise, if E_u and E_v are finite subsets of H which intersect, then $E_u \cap LE_v = \varnothing$, but if u and v disagree on some element of $E_u \cap E_v$, there clearly cannot be an element in X that contains both u and v). Let $u, v \in \mathcal{L}(Y)$

with shapes F_u and F_v such that $F_u \cap KF_v = \varnothing$. Since $u, v \in \mathcal{L}(Y)$, we clearly have that $u, v \in \mathcal{L}(X)$. We now show that $F_u \cap LF_v = \varnothing$. Indeed, since $K = L \cap H$, $F_u \cap KF_v = \varnothing$, and $F_u, F_v \subset H$,

$$
\begin{aligned}
F_u \cap LF_v &= \big(F_u \cap (L \cap H)F_v\big) \cup \big(F_u \cap (L \setminus H)F_v\big) \\
&\subset \big(F_u \cap KF_v\big) \cup \big(H \cap (L \setminus H)H\big) \\
&= \bigcup_{l \in L \setminus H} H \cap lH
\end{aligned}
$$

Since H is a subgroup of G, for any $l \in L \setminus H$ we have $H \cap lH = \varnothing$, since lH is a proper left coset of H. As such, $F_u \cap LF_v = \varnothing$, and so by the strong irreducibility of X, there exists $x \in X$ such that $x|_{F_u} = u$ and $x|_{F_v} = v$. Let $y = x|_H \in Y$, and so $y|_{F_u} = u$ and $y|_{F_v} = v$. Therefore, Y is strongly irreducible. □

Lastly, factor maps which are defined by block maps on a base shifts can be extended to a factor map of the free extension of the base shift in a natural manner. This does not apply to arbitrary factor maps from a free extension shifts however, and only applies to factor maps whose block maps are defined on a subset of the group for the base shift. In the result, for a function $\phi : X \to X$, we denote by $(\phi)^C$ the product function on X^C.

Lemma 1.2.12. *Let G be a group, $F \Subset H \leq G$ and $C \in \mathscr{C}(H \backslash G)$. For finite alphabets $\mathcal{A}$ and $\mathcal{B}$, let Y be an H-shift of $\mathcal{A}^H$, and let $\beta : \mathcal{L}_F(Y) \to \mathcal{B}$ be a block map. Then for any $x \in Y^{\uparrow G}$,*

$$\phi^G_\beta(x) = (\kappa^{\mathcal{B}}_C \circ (\phi^H_\beta)^C \circ (\kappa^{\mathcal{A}}_C)^{-1})(x).$$

Proof. Let $x \in Y^{\uparrow G}$. Then for any $g \in G$, let $h \in H$ and $d \in C$ such that $g = hd$.

Expanding the definition of $\kappa_C^{\mathcal{B}}$, we have

$$\begin{aligned}\big((\kappa_C^{\mathcal{B}} \circ (\phi_\beta^H)^{H\backslash G} \circ (\kappa_C^{\mathcal{A}})^{-1})(x)\big)(g) &= \left(\bigvee_{c\in C} \sigma^{c^{-1}}\big(\phi_\beta^H((\kappa_C^{\mathcal{A}})^{-1}(x)_c)\big)\right)(g) \\ &= \Big(\sigma^{d^{-1}}\big(\phi_\beta^H((\kappa_C^{\mathcal{A}})^{-1}(x)_d)\big)\Big)(hd),\end{aligned}$$

where $g \in Hd$ implies we only need to observe the pattern in the join on Hd. Applying the shift $\sigma^{d^{-1}}$, and expanding the definition of ϕ_β^H at h, we obtain

$$\Big(\sigma^{d^{-1}}\big(\phi_\beta^H((\kappa_C^{\mathcal{A}})^{-1}(x)_d)\big)\Big)(hd) = \phi_\beta^H\big((\sigma^d x)|_H\big)(h) = \beta\Big(\big(\sigma^h\big((\sigma^d x)|_H\big)\big)|_F\Big).$$

Simplifying, we have

$$\beta\Big(\big(\sigma^h\big((\sigma^d x)|_H\big)\big)|_F\Big) = \beta\big(((\sigma^{hd}x)|_{Hh^{-1}})|_F\big) = \beta\big((\sigma^g x)|_F\big) = \big(\phi_\beta^G(x)\big)(g)$$

by the definition of ϕ_β^G at g. With this, we have shown the desired result. □

A direct consequence of the previous result is that certain factors of a free extension are equal to the free extension of a factor of the base shift. This is not the case for all factors, however the property is quite useful nevertheless.

Corollary 1.2.13. *Let G be a group, $H \leq G$, $\mathcal{A}$ and $\mathcal{B}$ be finite alphabets, and Y be an H-shift of $\mathcal{A}^H$. Let $F \Subset H$ and $\beta : \mathcal{L}_F(Y) \to \mathcal{B}$ be a block map. Then*

$$\phi_\beta^G(Y^{\uparrow G}) = \phi_\beta^H(Y)^{\uparrow G}$$

Proof. Let $C \in \mathscr{C}(H\backslash G)$. By Lemma 1.2.12, for any $x \in Y^{\uparrow G}$, we have $\phi_\beta^G(x) =$

$\big(\kappa_C^{\mathcal{B}} \circ (\phi_\beta^H)^C \circ (\kappa_C^{\mathcal{A}})^{-1}\big)(x)$, and therefore, using Lemma 1.2.6,

$$\begin{aligned}
\phi_\beta^G(x) &= \big(\kappa_C^{\mathcal{B}} \circ (\phi_\beta^H)^C \circ (\kappa_C^{\mathcal{A}})^{-1}\big)(x) \\
&= \big(\kappa_C^{\mathcal{B}} \circ (\phi_\beta^H)^C\big)\big((\kappa_C^{\mathcal{A}})^{-1}(x)\big) \\
&\in \kappa_C^{\mathcal{B}}\Big((\phi_\beta^H)^C\big(Y^C\big)\Big) \\
&= \kappa_C^{\mathcal{B}}\big(\phi_\beta^H(Y)^C\big) \\
&= \phi_\beta^H(Y)^{\uparrow G}.
\end{aligned}$$

Similarly, for any $y \in \phi_\beta^H(Y)^{\uparrow G}$, we have $(\kappa_C^{\mathcal{B}})^{-1}(y) \in \phi_\beta^H(Y)^C = (\phi_\beta^H)^C(Y^C) = (\phi_\beta^H)^C\big((\kappa_C^{\mathcal{A}})^{-1}(Y^{\uparrow G})\big)$. Applying $\kappa_C^{\mathcal{B}}$ to both sides, we obtain by Lemma 1.2.12 that

$$y \in \big(\kappa_C^{\mathcal{B}} \circ (\phi_\beta^H)^C \circ (\kappa_C^{\mathcal{A}})^{-1}\big)(Y^{\uparrow G}) = \phi_\beta^G(Y^{\uparrow G}).$$

□

1.2.3 Applications of free extensions to shifts on groups

Using free extensions, it is possible to analyze shifts on arbitrary groups, though only to an extent. First, we can look at SFTs on arbitrary groups. We use the following result extensively in the study of SFTs on locally finite groups in particular, however it applies in full generality to all groups.

Lemma 1.2.14. *Let G be a group, and X a G-SFT. Then there exists $F \Subset G$ and $\langle F\rangle$-SFT Y such that $X = Y^{\uparrow G}$. In other words, every SFT on a group G is the free extension of an SFT on a finitely generated subgroup of G.*

Proof. Since X is an SFT, let $F \Subset G$ be a forbidden shape for X, so $X = \mathcal{X}^G[\mathcal{F}_F(X)]$. Let $H = \langle F\rangle \leq G$, which makes H finitely generated. Additionally, since $F \Subset H$, the H-shift $Y = \mathcal{X}^H[\mathcal{F}_F(X)]$ is an H-SFT. By Lemma 1.2.2,

$$Y^{\uparrow G} = \mathcal{X}^H[\mathcal{F}_F(X)]^{\uparrow G} = \mathcal{X}^G[\mathcal{F}_F(X)] = X,$$

which proves the desired result. ☐

In the case that G itself is finitely generated, it may be that $F \Subset G$ is such that $\langle F \rangle = G$, and so $X = Y = Y^{\uparrow G}$, which is a trivial result. In the case that G is infinitely generated however, this can never be the case, and leads to interesting results such as the following.

Corollary 1.2.15. *Let G be an infinitely generated amenable group. Then*

$$\mathcal{E}(G) = \bigcup_{F \Subset G} \mathcal{E}(\langle F \rangle).$$

Proof. This follows immediately from the previous lemma and Proposition 1.2.7. ☐

In addition to SFTs, which are defined by a finite forbidden shape, sofic shifts are defined by an SFT and a finite block map, and using a similar technique as the lemma above, we can show that any sofic shift on a group is the free extension of a sofic shift on a finitely generated subgroup.

Lemma 1.2.16. *Let G be a group and X be a sofic G-shift. Then there exists $F \Subset G$ and sofic $\langle F \rangle$-shift Y such that $X = Y^{\uparrow G}$.*

Proof. Since X is sofic, let Z be a G-SFT, and $\beta : \mathcal{L}_E(Z) \to \mathcal{A}$ be a block map such that $X = \phi_\beta^G(Z)$ with $E \Subset G$. Since Z is a G-SFT, by Lemma 1.2.14 that there exists $F \Subset G$ and $\langle F \rangle$-SFT W such that $W^{\uparrow G} = Z$. Let $H = \langle F \cup E \rangle$, and we have that $U = W^{\uparrow H}$ is an H-SFT by Lemma 1.2.10, and that $U^{\uparrow G} = (W^{\uparrow H})^{\uparrow G} = W^{\uparrow G} = Z$ by Lemma 1.2.8. As such, Lemma 1.2.12 gives us that

$$X = \phi_\beta^G(Z) = \phi_\beta^G(U^{\uparrow G}) = (\phi_\beta^H(U))^{\uparrow G},$$

since $E \Subset H$. Since U is an H-SFT, we have that $Y = \phi_\beta^H(U)$ is a sofic H-shift, and clearly H is finitely generated with $X = Y^{\uparrow G}$. ☐

Furthermore, the finite nature of the strong irreducibility condition (namely the finiteness of K) allows us to prove the same result as for SFTs and sofic shifts.

Lemma 1.2.17. *Let G be a group and X be a strongly irreducible G-shift. Then there exists $F \Subset G$ and strongly irreducible $\langle F \rangle$-shift Y such that $X = Y^{\uparrow G}$.*

Proof. Since X is strongly irreducible, let $K \Subset G$ be such that for any $u, v \in \mathcal{L}(X)$ with shapes F_u and F_v respectively such that $F_u \cap KF_v = \varnothing$, then there exists $x \in X$ such that $x|_{F_u} = u$ and $x|_{F_v} = v$. We will now show that K is an option for the finite set F in the statement of the lemma. For any $F \Subset H$ let $\mathbf{F}_F = \mathcal{F}_F(X)$, and define $Y_F = \mathcal{X}^H[\mathbf{F}_F]$. Then since F is finite, Y_F is an H-SFT for each F. Furthermore, by Lemma 1.2.2, we have that $Y_F^{\uparrow G} = (\mathcal{X}^H[\mathbf{F}_F])^{\uparrow G} = \mathcal{X}^G[\mathbf{F}_F]$, and so clearly $X \subset Y_F^{\uparrow G}$. As such, $X \subset \bigcap_{F \Subset H} Y_F^{\uparrow G}$, and by Lemma 1.2.9, we have that

$$\bigcap_{F \Subset H} Y_F^{\uparrow G} = \left(\bigcap_{F \Subset H} Y_F \right)^{\uparrow G}.$$

Let $Y = \bigcap_{F \Subset H} Y_F$, which is an H-shift, so we have $X \subset Y^{\uparrow G}$. We now show that $Y^{\uparrow G} \subset X$.

Let $C \in \mathscr{C}(H \backslash G)$, let $z \in Y^{\uparrow G}$, and let $g \in G$ and $F \Subset H$. By the construction of $Y^{\uparrow G}$, we have $z \in Y_F^{\uparrow G} = \mathcal{X}^G[\mathbf{F}_F]$, and so $z|_F \notin \mathbf{F}_F = \mathcal{F}_F(X)$. Therefore, there exists $x_F \in X$ such that $z|_F = x_F|_F$. In particular, this shows that the set $E_F = [z|_F] \cap X$ is nonempty and closed. Additionally, since for each $g \in G$ we have $\sigma^g z \in Y^{\uparrow G}$, we also have that $[(\sigma^g z)|_F] \cap X = \sigma^g([z|_{Fg}] \cap X)$ is nonempty and closed. Since σ^g is a homeomorphism on $\mathcal{A}^G$ and X, we have that $E_{Fg} = [z|_{Fg}] \cap X$ is a nonempty closed subset of X. As such,

$$\mathscr{G} = \{E_{Fg} : F \Subset H, g \in G\}$$

is a collection of nonempty closed subsets of X. We now show that $\mathscr{G}$ has the finite intersection property.

Let $E_{F_1g_1}, E_{F_2g_2}, \dots, E_{F_ng_n} \in \mathscr{G}$. Note that since $F_i \Subset H$ and $g_i \in G$, we have that $F_ig_i \Subset Hc_i$ for some unique $c_i \in C$. If we have that $F_jg_j \Subset Hc_i$ for some $j \neq i$, then $F_ig_i \cup F_jg_j \Subset Hc_i$, giving $(F_ig_i \cup F_jg_j)c_i^{-1} \Subset H$, and so we have

$$E_{F_ig_i} \cap E_{F_jg_j} = [z|_{F_ig_i}] \cap [z|_{F_jg_j}] \cap X = [z|_{(F_ig_i \cup F_jg_j)c_i^{-1}c_i}] \cap X = E_{F_ig_i \cup F_jg_j},$$

which is an element of $\mathscr{G}$, and so we may assume without loss of generality that $c_i \neq c_j$ for $i \neq j$. For finite induction, we have that $E_{F_1g_1}$ is nonempty, so suppose that we have shown $\bigcap_{i=1}^k E_{F_ig_i}$ is nonempty for some $k < n$. As such, there exists an element $x \in X$ such that $u = x|_{\bigcup_{i=1}^k F_ig_i} = z|_{\bigcup_{i=1}^k F_ig_i}$, meaning that $u \in \mathcal{L}(X)$. Let $v = z|_{F_{k+1}g_{k+1}}$, and note that $v \in \mathcal{L}(X)$, as $E_{F_{k+1}g_{k+1}}$ is nonempty. Now, since $F_ig_i \subset Hc_i$ for each i, and $K \subset H$ by definition of H, we have that

$$\left(\bigcup_{i=1}^k F_ig_i\right) \cap K(F_{k+1}g_{k+1}) \subset \left(\bigcup_{i=1}^k Hc_i\right) \cap H(Hc_{k+1}) = \left(\bigcup_{i=1}^k Hc_i\right) \cap Hc_{k+1}.$$

Since $c_i \neq c_j$ for $i \neq j$, we have in the rightmost set an intersection of a right coset with a union of distinct right cosets, which is necessarily empty, and so we have

$$\left(\bigcup_{i=1}^k F_ig_i\right) \cap K(F_{k+1}g_{k+1}) = \varnothing.$$

By the strong irreducibility of X, there exists $x \in X$ such that $x|_{\bigcup_{i=1}^k F_ig_i} = u$ and $x|_{F_{k+1}g_{k+1}} = v$. This gives that $x \in \bigcap_{i=1}^{k+1} E_{F_ig_i}$, so this set is nonempty. By inducing until n, we obtain that $\bigcap_{i=1}^n E_{F_ig_i}$ is nonempty. As such, $\mathscr{G}$ has the finite intersection property.

Since X is a closed subset of $\mathcal{A}^G$, which is compact, we have that X is compact. As such, since $\mathscr{G}$ is a collection of closed subsets of X with the finite intersection property, $\bigcap \mathscr{G}$ is also nonempty, in particular there exists $x \in X$ such that

$$x \in \bigcap_{g \in G} \bigcap_{F \Subset H} E_{Fg}.$$

With $\{e\} \Subset H$, this gives that for each $g \in G$, we have $x \in E_{\{g\}} = [z|_{\{g\}}] \cap X$, which gives that $x(g) = z(g)$ and therefore $x = z \in X$. Since $z \in Y^{\uparrow G}$ was arbitrary, we have shown that $Y^{\uparrow G} \subset X$, and therefore $X = Y^{\uparrow G}$.

Finally, by Lemma 1.2.11, since $X = Y^{\uparrow G}$ is strongly irreducible, we have that Y is strongly irreducible. $\square$

As is the case with SFTs, the two previous results may be trivial in the case that G is finitely generated, however this is not the case when G is infinitely generated. Further study into the properties of free extensions and which properties translate from a free extension to its base shift and vice-versa, may prove to show that the study of SFTs on arbitrary groups may be reducible to studying SFTs on finitely generated groups. While we do not require any further properties for the results of this chapter, it may be fruitful to explore other such properties in the context of free extensions.

1.3 Locally Finite Groups

With the theory of free extensions sufficiently developed, we may move on to proving properties of SFTs on locally finite groups. This section contains all parts of the proofs of Theorems I and II.

We first begin by introducing the following construction, which applies to any group G which is not locally finite, and which will be referenced throughout the remainder of the section.

Definition 1.3.1. Let G be a non-locally finite group, and $\mathcal{A} = \{0, 1\}$. Since G is non-locally finite, there exists an infinite, finitely generated group $H \leq G$. Let $S \Subset H$ be such that $e \in S$ and $\langle S \rangle = H$. Then taking $\mathbf{F} = \mathcal{A}^S \setminus \{0^S, 1^S\}$, where 0^S and 1^S are the constant 0 and 1 patterns, let $\mathcal{2}_H = \mathcal{X}^H[\mathbf{F}]$.

$\mathcal{2}_H$ is clearly an SFT from this construction, and in particular contains exactly 2 points, the constant 0 and 1 patterns on H, which will be denoted 0^H and 1^H respectively. By Lemma 1.2.10, $\mathcal{2}_H^{\uparrow G}$ is also an SFT.

1.3.1 Proof of Theorem I

We now have everything needed to prove Theorem I. Each of the results in this section which contributes to the Theorem will be marked with the implication that it provides. For instance, the following result is marked as (I(a) $\Longrightarrow$ I(b)) to indicate that it provides the implication that if G is locally finite, then every G-SFT is the free extension of an SFT on a finite subgroup of G. Many of these results follow readily from the properties of free extensions developed in the previous section. Theorem I is restated below for convenience.

Theorem I. *Let G be a group. Then the following are equivalent.fallow*

(a) G is locally finite.

(b) Every G-SFT is the free extension of some SFT on a finite subgroup of G.

(c) Every G-SFT is strongly irreducible.

(d) Every strongly irreducible G-shift is a G-SFT.

(e) Every sofic G-shift is a G-SFT.

(f) For every G-SFT X, $\mathrm{Aut}(X)$ is locally finite.

We begin by proving the following chain of implications:

$$\mathrm{I(a)} \Longrightarrow \mathrm{I(b)} \Longrightarrow \mathrm{I(c)} \Longrightarrow \mathrm{I(a)}$$

The first of these implications follows directly from Lemma 1.2.14.

Proposition 1.3.2 (I(a) $\Longrightarrow$ I(b))**.** *Let G be a locally finite group, and X a G-SFT. Then there exists $H \ll G$ and an H-SFT Y such that $X = Y^{\uparrow G}$.*

Proof. By Lemma 1.2.14, there exists $F \Subset G$ and $\langle F \rangle$-SFT Y such that $X = Y^{\uparrow G}$. But G is locally finite, so $H = \langle F \rangle$ is finite, which gives the desired result. □

Next we show that if G is a group for which every G-SFT is the free extension of a shift on a finite subgroup of H, then every G-SFT is strongly irreducible. In fact, we can show the following result, which is stronger; if X is a G-SFT for which there exists $H \ll G$ and H-SFT Y such that $X = Y^{\uparrow G}$, then X is strongly irreducible.

Lemma 1.3.3 (I(b) $\Longrightarrow$ I(c))**.** *Let G be a group and X be a G-SFT such that there exists $H \ll G$ and H-SFT Y such that $X = Y^{\uparrow G}$. Then X is strongly irreducible.*

Proof. Since H is finite, Y is vacuously strongly irreducible with $K = H$. By Lemma 1.2.11, $X = Y^{\uparrow G}$ is strongly irreducible. □

Lastly, we prove the final implication by contrapositive, where we use the SFT $2_H^{\uparrow G}$ as an example of an SFT on non-locally finite groups which is not strongly irreducible.

Lemma 1.3.4 (I(c) $\Longrightarrow$ I(a))**.** *Let G be a non-locally finite group. Then there exists a G-SFT which is not strongly irreducible.*

Proof. Let $H \leq G$ be an infinite, finitely generated subgroup of G, which must exist because G is not locally finite.

To show that $2_H^{\uparrow G}$ is not strongly irreducible, it is necessary to show that for all $K \Subset G$, there exist patterns $u, v \in \mathcal{L}(2_H^{\uparrow G})$ with shapes F_u and F_v respectively such that $F_u \cap KF_v = \varnothing$, but there is no $x \in X$ with $x|_{F_u} = u$ and $x|_{F_v} = v$.

Let $K \Subset G$. Since K is finite, it must be that $H \setminus K$ is nonempty, so let $h \in H \setminus K$. Let $u = 0^{\{h\}}$ and $v = 1^{\{e\}}$, which are trivially in $\mathcal{L}(2_H^{\uparrow G})$. Then $F_u = \{h\}$ and $F_v = \{e\}$, and clearly since $h \notin K$, we have $F_u \cap KF_v = \{h\} \cap K = \varnothing$. But, for any $x \in X$, it must be that $x|_H \in \{0^H, 1^H\}$, and therefore $x|_{\{h\}} = x|_{\{e\}}$, so it cannot be that $x|_{F_u} = u$ and $x|_{F_v} = v$ simultaneously.

Therefore, $2_H^{\uparrow G}$ is not strongly irreducible. □

Next, we shall prove that I(a) $\Longrightarrow$ I(d), and prove the converse direction in the subsection immediately following, as we will need an example introduced then.

Lemma 1.3.5 (I(a) $\Longrightarrow$ I(d))**.** *Let G be a locally finite group, and X a strongly irreducible G-shift. Then X is a G-SFT.*

Proof. By Lemma 1.2.17, there exists $F \Subset G$ and strongly irreducible $\langle F \rangle$-shift Y such that $X = Y^{\uparrow G}$. Since G is locally finite and F is finite, $H = \langle F \rangle$ is finite, and therefore Y is an H-SFT. By Lemma 1.2.10, $Y^{\uparrow G} = X$ is a G-SFT. □

1.3.1.1 Sofic shifts on locally finite groups

Next, we prove the following implication involving the statement that every sofic G-shift is a G-SFT.

$$\text{I(a)} \Longleftrightarrow \text{I(e)}$$

First, we show directly that all sofic G-shifts on locally finite groups are G-SFTs.

Lemma 1.3.6 (I(a) $\Longrightarrow$ I(e))**.** *Let G be a locally finite group, X be an SFT, and ϕ be a factor map. Then $\phi(X)$ is an SFT.*

Proof. By Proposition 1.3.2, there exists $H \ll G$ and H-SFT Y such that $X = Y^{\uparrow G}$. Let $F \Subset G$ and let $\beta : \mathcal{L}_F(X) \to \mathcal{B}$ be a block map such that $\phi = \phi_\beta^G$. Let $K = \langle H \cup F \rangle$, which is finite because G is locally finite, and let $Z = Y^{\uparrow K}$. By Lemma 1.2.8, we have $X = Y^{\uparrow G} = (Y^{\uparrow K})^{\uparrow G} = Z^{\uparrow G}$. Since $F \subset K$, Lemma 1.2.13 gives that

$$\phi(X) = \phi_\beta^G(Z^{\uparrow G}) = \phi_\beta^K(Z)^{\uparrow G}.$$

Since K is finite, the K-shift $\phi_\beta^K(Z) \subset \mathcal{A}^K$ is an SFT, and by Lemma 1.2.10, we obtain that $\phi_\beta^K(Z)^{\uparrow G} = \phi(X)$ is an SFT. □

For the converse result, we will prove the contrapositive by constructing for any non-locally finite group, a sofic shift which is not an SFT. We begin with the construction.

Definition 1.3.7 (Example 1.11 [34])**.** Let H be an infinite, finitely generated group, and let $S \Subset H$ such that $S = S^{-1}$, $e \notin S$, and $H = \langle S \rangle$. The *even H-shift* $S_{\text{even}} \subset \{0,1\}^H$ is the set of all configurations x such that any maximal finite connected component of $x^{-1}(1) \subset H$ in the Cayley graph $\Gamma(H, S)$ has even size. Said otherwise, each finite connected component of ones has even size.

Proposition 1.10 of [34] gives that S_{even} is a sofic H-shift, but not an H-SFT. Using this, we can prove the converse result.

Lemma 1.3.8 (I(e) $\Longrightarrow$ I(a))**.** *Let G be a non-locally finite group. Then there exists a sofic G-shift which is not a G-SFT.*

Proof. Let $H \leq G$ be infinite and finitely generated. Then S_{even} as defined above is a sofic H-shift, but not an H-SFT. Let X be an H-SFT and $\phi : X \to S_{\text{even}}$ be a factor map, which must exist by the soficity of S_{even}. Then there exists $F \Subset H$ and a block map $\beta : \mathcal{L}_F(X) \to \{0,1\}$ such that $\phi = \phi_\beta^H$. Then by Lemma 1.2.10, it follows $X^{\uparrow G}$ is a G-SFT, and by Lemma 1.2.13, we obtain

$$S_{\text{even}}^{\uparrow G} = \phi_\beta^H(X)^{\uparrow G} = \phi_\beta^G(X^{\uparrow G}),$$

and therefore $S_{\text{even}}^{\uparrow G}$ is sofic. By the contrapositive of Lemma 1.2.10 however, $S_{\text{even}}^{\uparrow G}$ is not an SFT. □

In addition to S_{even} being a sofic H-shift, we also have that it is strongly irreducible. With $K = (S \cup \{e\})^2$, and two patterns $u, v \in S_{\text{even}}$ with shapes F_u and F_v such that $F_u \cap KF_v = \varnothing$, we may extend u to a pattern on $(S \cup \{e\})F_u$ by using the symbol 0 or 1 in a manner that ensures this extension has an even number of 1s in any connected component of 1s. The same can be done for v. By the definition of K, we have that these extensions are disjoint, so let $x \in \{0,1\}^H$ which matches these extensions of u and v, and is 0 elsewhere. Since the extensions of u and v have connected components of 1s of even size, x only has connected components of 1s of even size, even if a connected component in the extension of u is connected with a connected component of the extension of v, since both individually have even size. As such, we have the following result.

Lemma 1.3.9 (I(d) $\Longrightarrow$ I(a))**.** *Let G be a non-locally finite group. Then there exists a strongly irreducible G-shift which is not a G-SFT.*

Proof. Let $H \leq G$ be infinite and finitely generated. Then S_{even} as defined above is a strongly irreducible H-shift, but not an H-SFT. By Lemma 1.2.10, $S_{\text{even}}^{\uparrow G}$ is not a G-SFT, and by Lemma 1.2.11, $S_{\text{even}}^{\uparrow G}$ is strongly irreducible. □

1.3.1.2 Automorphism groups for locally finite SFTs

Finally, we prove the last implications for Theorem I in the following manner.

$$\text{I(a)} \Longleftrightarrow \text{I(f)}$$

First, we show that the automorphism group for any SFT on a locally finite groups is locally finite.

Lemma 1.3.10 (I(a) $\Longrightarrow$ I(f))**.** *Let G be locally finite and X a G-SFT. Then* $\text{Aut}(X)$ *is locally finite.*

Proof. Let $F \Subset G$ be a forbidden shape for X so that $X = \mathcal{X}^G[\mathcal{F}_F(X)]$. Let $E = \{\phi_1, \phi_2, \ldots, \phi_n\} \subset \mathrm{Aut}(X)$ be a finite set of autoconjugacies, and let $K = \langle E \rangle$. Without loss of generality, E may be assumed to be symmetric. Then for each ϕ_i, there exists $F_i \Subset G$ and block maps $\beta_i : \mathcal{L}_{F_i}(X) \to \mathcal{A}$ such that $\phi_i = \phi^G_{\beta_i}$. Now, let

$$H = \left\langle F \cup \bigcup_{i=1}^{n} F_i \right\rangle.$$

H must be finite, since G is locally finite. Then, since $F \subset H$, it is the case that $X = \mathcal{X}^G[\mathcal{F}_H(X)]$, and by Lemma 1.2.2, we have $X = \mathcal{X}^G[\mathcal{F}_H(X)] = \mathcal{X}^H[\mathcal{F}_H(X)]^{\uparrow G}$. For simplicity, let $Y = \mathcal{X}^H[\mathcal{F}_H(X)]$. Additionally, by Corollary 1.2.13, for each i, we obtain

$$Y^{\uparrow G} = \phi_i(Y^{\uparrow G}) = \phi^H_{\beta_i}(Y)^{\uparrow G},$$

and therefore $Y = \phi^H_{\beta_i}(Y)$, which gives $\phi^H_{\beta_i} \in \mathrm{Aut}(Y)$. Let $C \in \mathscr{C}(H \backslash G)$. Then for each i and j and using Lemma 1.2.12, we have for every $x \in Y^{\uparrow G}$ that

$$(\phi_i \circ \phi_j)(x) = \left(\kappa_C \circ (\phi^H_{\beta_i})^C \circ \kappa_C^{-1} \circ \kappa_C \circ (\phi^H_{\beta_j})^C \circ \kappa_C^{-1}\right)(x) = \left(\kappa_C \circ (\phi^H_{\beta_i} \circ \phi^H_{\beta_j})^C \circ \kappa_C^{-1}\right)(x),$$

with $\phi^H_{\beta_i} \circ \phi^H_{\beta_j} \in \mathrm{Aut}(Y)$. As such, the behavior of each $\phi \in K$ on $Y^{\uparrow G}$ is entirely determined by an element in $\mathrm{Aut}(Y)$. Since H is finite, Y is finite, and therefore $\mathrm{Aut}(Y) \subset Y^Y$ is also finite, which gives that K must be finite. Since E was arbitrary, this gives that $\mathrm{Aut}(X)$ is locally finite. □

Lastly, we show that if the automorphism group of an SFT is locally finite, then the group on which the SFT is defined must be locally finite.

Lemma 1.3.11 (I(f) $\Longrightarrow$ I(a))**.** *Let G be a group. If for every G-SFT X,* $\mathrm{Aut}(X)$ *is locally finite, then G is locally finite.*

Proof. If every G-SFT X satisfies $\mathrm{Aut}(X)$ is locally finite, in particular this is true of the full G-shift Σ (on at least 2 symbols). Clearly, the map $\psi : G \to \mathrm{Aut}(\Sigma)$

defined by $\psi(g) = \sigma^g$ is an injective homomorphism, since for any $h \neq g$, we have $\sigma^h \neq \sigma^g$ on Σ, since it is possible to describe a configuration which gets sent to different configurations under σ^h and σ^g. As such, $\psi(G) \leq \mathrm{Aut}(\Sigma)$. Since $\mathrm{Aut}(\Sigma)$ is locally finite, $\psi(G)$ is locally finite. But $\psi(G)$ and G are isomorphic, and therefore G is locally finite. □

1.3.2 Proof of Theorem II

Next, we will prove Theorem II. As with the previous section, results pertaining to certain implications in Theorem II are marked. The main additional assumption we will need is that G is a countable amenable group, rather than any group. Most of these results also depend heavily on the properties of free extensions developed in the previous section. We restate Theorem II below for convenience.

Theorem II. *Let G be a countable amenable group. Then the following are equivalent.*

(a) G is locally finite

(b) If X is a nonempty G-SFT with $h(X) = 0$, then $X = \{x\}$, where x is a fixed point.

(c) Every G-SFT is entropy minimal.

(d) G is locally non-torsion, and

$$\mathcal{E}(G) = \left\{ \frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N} \right\} \subset \mathbb{Q}^+_{\log}.$$

(e) Every G-SFT has a unique measure of maximal entropy.

Each of the equivalences in the theorem will be shown individually to be equivalent to II(a). We begin by showing its equivalence to II(b). Additionally, note that all countable locally finite groups are amenable, so we omit amenable as an assumption for a few of the results.

1.3.2.1 Zero entropy SFTs on locally finite groups

We begin by showing that zero entropy SFTs on locally finite groups consist of single fixed points.

Lemma 1.3.12 (II(a) $\implies$ II(b))**.** *Let G be a countable locally finite group. Then if X is a non-empty G-SFT with $h(X) = 0$, then $X = \{x\}$ for some fixed point x.*

Proof. Let X be a G-SFT with $h(X) = 0$. Then by assumption, $X = Y^{\uparrow G}$ for some $H \ll G$ and H-shift Y. By Proposition 1.2.7, we have $h(X) = h(Y) = 0$. Since H is finite,

$$0 = h(Y) = \frac{1}{|H|}\log(|Y|),$$

which implies that $|Y| = 1$. Then for any $C \in \mathscr{C}(H\backslash G)$ we have $|Y^C| = 1$, and therefore $|X| = |\kappa_C(Y^C)| = 1$ so $X = \{x\}$ for the only $x \in X$. Since X is shift invariant, it must be that x is a fixed point. □

To show the converse, recall the definition of the SFT $\mathcal{2}_H$ from the beginning of Section 1.3.

Lemma 1.3.13 (II(b) $\implies$ II(a))**.** *Let G be a countable amenable non-locally finite group. Then there exists a G-SFT X with 0 topological entropy, however $|X| > 1$.*

Proof. Let $H \leq G$ be an infinite, finitely generated subgroup. By Proposition 1.2.7, we have $h(\mathcal{2}_H^{\uparrow G}) = h(\mathcal{2}_H)$. Since G is countable and amenable, and $H \leq G$, it is the case that H is also countable and amenable, so let $\{F_n\}_{n=1}^{\infty}$ be a Følner sequence for H. Since $\mathcal{2}_H$ contains exactly 2 points, 0^H and 1^H, it is clear to see that $\mathcal{L}_{F_n}(\mathcal{2}_H) = \{0^{F_n}, 1^{F_n}\}$, and therefore $|\mathcal{L}_{F_n}(\mathcal{2}_H)| = 2$. Additionally, it must be that $\lim_{n\to\infty} |F_n| =$

∞, because H is an infinite subgroup. Then

$$h(2_H^{\uparrow G}) = h(2_H) = \lim_{n\to\infty} |F_n|^{-1} \log\bigl(|\mathcal{L}_{F_n}(2_H)|\bigr) = \log(2) \lim_{n\to\infty} |F_n|^{-1} = 0.$$

Also, since $|2_H| = 2$, $|2_H^{\uparrow G}| > 1$, which gives the desired result. □

1.3.2.2 Entropy minimality of SFTs on locally finite groups

Recall that a G-SFT X is entropy minimal if for every G-shift $Y \subsetneq X$, we have $h(Y) < h(X)$. The following result shows that for a countable locally finite group, every SFT on the group is entropy minimal.

Lemma 1.3.14 (II(a) $\Longrightarrow$ II(c)). *Let G be a countable locally finite group, and X be a G-SFT. Then X is entropy minimal.*

Proof. Since X is an SFT, let $F \Subset G$ be such that $X = \mathcal{X}^G[\mathcal{F}_F(X)]$. Let $Y \subsetneq X$ also be an SFT, and let $E \Subset G$ be such that $Y = \mathcal{X}^G[\mathcal{F}_E(Y)]$. Let $H = \langle F \cup E\rangle$, which is finite because E and F are finite and G is locally finite. Also we have $E, F \subset H$, and therefore $X = \mathcal{X}^G[\mathcal{F}_H(X)]$ and $Y = \mathcal{X}^G[\mathcal{F}_H(Y)]$. By Lemma 1.2.2, we obtain $X = (\mathcal{X}^H[\mathcal{F}_H(X)])^{\uparrow G}$ and $Y = (\mathcal{X}^H[\mathcal{F}_H(Y)])^{\uparrow G}$. Given that $Y \subsetneq X$, it must then be that $\mathcal{X}^H[\mathcal{F}_H(Y)] \subsetneq \mathcal{X}^H[\mathcal{F}_H(X)]$. With Proposition 1.2.10, this gives

$$\begin{aligned} h(Y) &= h(\mathcal{X}^H[\mathcal{F}_H(Y)]) = |H|^{-1}\log(|\mathcal{X}^H[\mathcal{F}_H(Y)]|) \\ &< |H|^{-1}\log(|\mathcal{X}^H[\mathcal{F}_H(X)]|) = h(\mathcal{X}^H[\mathcal{F}_H(X)]) = h(X), \end{aligned}$$

where the strict inequality follows from the fact that log is strictly increasing, and this implies that X is SFT-entropy minimal. Since X is an SFT, and SFT-entropy minimality and entropy minimality are equivalent for SFTs, we have that X is entropy minimal. □

For the converse result about entropy minimality, we again use the SFT 2_H.

Lemma 1.3.15 (II(c) $\implies$ II(a))**.** *Let G be a countable amenable non-locally finite group. Then there exists a G-SFT X which is not entropy minimal.*

Proof. Let $H \leq G$ be an infinite, finitely generated subgroup. We have that $2_H^{\uparrow G}$ is a G-SFT, and $h(2_H^{\uparrow G}) = 0$, an argument for which can be found in Lemma 1.3.13. It is clear that $\{0^G\} \subset 2_H^{\uparrow G}$, and $\{0^G\}$ is clearly a G-shift, as it is conjugate to the full G-shift on 1 symbol. Additionally, $h(\{0^G\}) = 0$, and therefore $2_H^{\uparrow G}$ is not entropy minimal. □

1.3.2.3 The set of SFT entropies for locally finite groups

Next, we establish the set of all entropies that SFTs can obtain for locally finite groups. The following result shows that II(a) implies II(d). The first part of the implication is trivial; if G is locally finite, then every finitely generated subgroup is finite, and therefore G is locally non-torsion. The second part of the implication is given below.

Lemma 1.3.16 (II(a) $\implies$ II(d))**.** *Let G be a countable locally finite group. Then*

$$\mathcal{E}(G) = \left\{ \frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N} \right\} \subset \mathbb{Q}^+_{\log}.$$

Proof. First, consider the case when G is finite. Let X be a G-SFT. Then $h(X) = \frac{\log(|X|)}{|G|} \in \mathbb{Q}^+_{\log}$, and so

$$\mathcal{E}(G) \subset \left\{ \frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N} \right\},$$

since $G \ll G$. Now let $H \ll G$ and $n \in \mathbb{N}$. Since G and H are finite, let $m = \frac{|G|}{|H|} \in \mathbb{N}$. Let $\mathcal{A}$ be a finite alphabet with $|\mathcal{A}| = n^m$. Then, let $X = \{a^G : a \in \mathcal{A}\}$, which is a G-SFT, and $|X| = n^m$. Then

$$h(X) = \frac{\log(|X|)}{|G|} = \frac{\log(n^m)}{|G|} = \frac{m\log(n)}{|G|} = \frac{\log(n)}{|H|},$$

and therefore,

$$\left\{\frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N}\right\} \subset \mathcal{E}(G),$$

which gives the desired result.

If G is infinite, then G must be infinitely generated, and so by Corollary 1.2.15,

$$\mathcal{E}(G) = \bigcup_{F \Subset G} \mathcal{E}(\langle F \rangle).$$

Since G is locally finite, $H \ll G$ if and only if H is finitely generated, which gives

$$\begin{aligned}\mathcal{E}(G) = \bigcup_{H \ll G} \mathcal{E}(H) &= \bigcup_{H \ll G} \left\{\frac{\log(n)}{|K|} : K \ll H, n \in \mathbb{N}\right\} \\ &= \left\{\frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N}\right\}.\end{aligned}$$

□

Many locally finite groups do not satisfy $\mathcal{E}(G) = \mathbb{Q}^+_{\log}$, due to the lack of subgroups of certain orders. For example, $\bigoplus_{n\in\mathbb{N}} \mathbb{Z}/2\mathbb{Z}$ is locally finite, but only has subgroups of order 2^n. There are locally finite groups which do attain $\mathcal{E}(G) = \mathbb{Q}^+_{\log}$ however, with the most prominent example likely being Hall's universal group $\mathbb{U}$ [24], which has the property that every countable locally finite group can be embedded within it, which includes all finite groups. As such, it has finite subgroups of every order, and so $\mathcal{E}(\mathbb{U}) = \mathbb{Q}^+_{\log}$.

A direct converse of the previous lemma has been elusive to the author, which is the reason for the additional statement that G is locally non-torsion in II(d). The following lemma gives the most general form of a converse that has been found by the author.

Lemma 1.3.17. *Let G be a countable amenable group such that $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$. Then G is periodic.*

Proof. We proceed by the contrapositive. Let G be a group which is not periodic, meaning there exists $h \in G$ whose order is infinite. Let $H = \langle h \rangle$ so that H is isomorphic to $\mathbb{Z}$, and define $\mathbf{F} = \{1^{\{e,h\}}\} \subset \{0,1\}^{\{e,h\}}$, and let $X = \mathcal{X}^H[\mathbf{F}]$. Since G is amenable, and $H \leq G$, it must be that H is amenable. Then X is conjugate to the well known golden mean shift on $\mathbb{Z}$, so $h(X) = \log(\varphi)$, where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. X is also clearly an SFT, so by Lemma 1.2.10 the G-shift $X^{\uparrow G}$ is an SFT, and by Proposition 1.2.7 we have $h(X^{\uparrow G}) = h(X) = \log(\varphi)$. It is an elementary number theory exercise to show that φ^n is irrational for all $n \in \mathbb{N}$, and so it must be that for any $n, m \in \mathbb{N}$, we have $\varphi^m \neq n$. Therefore $\forall n, m \in \mathbb{N}$, it is the case that $\log(\varphi) \neq \frac{\log(n)}{m}$, so $\log(\varphi) \notin \mathbb{Q}^+_{\log}$. But $\log(\varphi) \in \mathcal{E}(G)$, and therefore $\mathcal{E}(G) \not\subset \mathbb{Q}^+_{\log}$. □

It remains to show that periodic but not locally finite groups have SFTs with entropy outside of $\mathbb{Q}^+_{\log}$, however it is in general quite difficult to construct SFTs on such groups in a manner conducive to computing its topological entropy. As a result, we instead add the statement that G is locally non-torsion, which removes the need to consider such groups.

Lemma 1.3.18 (II(d) $\implies$ II(a))**.** *Let G be a countable amenable group which is locally non-torsion, and $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$. Then G is locally finite.*

Proof. By Lemma 1.3.17, G is periodic. Let $F \Subset G$, and consider $H = \langle F \rangle$. Since G is periodic, H is periodic. Since G is locally non-torsion, H is finite or not periodic, and therefore H must be finite. Since $F \Subset G$ was arbitrary, G is locally finite. □

The author suspects that if $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$, then $\mathcal{E}(G)$ must be locally finite. This would allow for II(d) to have the locally non-torsion assumption removed, and only leave $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$.

1.3.2.4 Measures of maximal entropy for SFTs on locally finite groups

Finally, we show that every SFT on a countable locally finite group has a unique measure of maximal entropy, and that if every SFT on a countable amenable group has a unique measure of maximal entropy, then the group must be locally finite. First, we require a simple but powerful result about the topological structure of SFTs on countable locally finite groups.

Lemma 1.3.19. *Let G be a countable locally finite group, and let X be a G-SFT. Then there exists a sequence $\{H_n\}_{n=1}^{\infty}$ with $H_n \leq H_{n+1} \ll G$ for all n, such that $G = \bigcup_{n\in\mathbb{N}} H_n$, and there exist H_n-SFTs Y_n such that $X = Y_n^{\uparrow G}$ for all n. Furthermore, the set*

$$\mathfrak{B}[\{Y_n\}] = \{[y] \cap X : n \in \mathbb{N}, y \in Y_n\}$$

is a basis for the subspace topology on X.

Proof. First, since X is a G-SFT, by Proposition 1.3.2, there exists $H_1 \ll G$ and H_1-SFT Y_1 such that $X = Y_1^{\uparrow G}$. Then, since G is countable, let $G = \{g_n : n \in \mathbb{N}\}$ be an enumeration of G. Define for $n \geq 2$,

$$H_n = \langle H_1 \cup \{g_i : i < n\} \rangle.$$

Since H_1 and $\{g_i : i < n\}$ are both finite, H_n is finitely generated, and therefore finite. Furthermore, for any $g \in G$, there is some $n \in \mathbb{N}$ for which $g = g_n$, and clearly $g_n \in H_{n+1}$. Also, $H_n \leq H_{n+1}$.

Now, for each $n \geq 2$, let $Y_n = Y_1^{\uparrow H_n}$. By Lemma 1.2.8, we obtain $X = Y_1^{\uparrow G} = (Y_1^{\uparrow H_n})^{\uparrow G} = Y_n^{\uparrow G}$.

Finally, let $\mathfrak{B}$ be the standard basis of all cylinder sets for X. To show that $\mathfrak{B}[\{Y_n\}]$ is a basis for the topology on X, first note that $\mathfrak{B}[\{Y_n\}] \subset \mathfrak{B}$, and therefore it suffices to show that any set in $\mathfrak{B}$ can be constructed by sets in $\mathfrak{B}[\{Y_n\}]$. Let $w \in \mathcal{L}(X)$ so that $[w] \cap X$ is nonempty, and let F be the shape of w. Since $G = \bigcup_{n\in\mathbb{N}} H_n$ and

$H_n \leq H_{n+1}$, it follows there exists $N \in \mathbb{N}$ such that $F \subset H_N$. Then, it is clear that

$$[w] \cap X = \bigcup_{z \in [w]_{H_N} \cap Y_N} [z] \cap X,$$

which implies that $\tau(\mathfrak{B})$, the topology generated by $\mathfrak{B}$, is contained in $\tau(\mathfrak{B}[\{Y_n\}])$, so $\mathfrak{B}[\{Y_n\}]$ is a basis for the topology on X. □

Lemma 1.3.20 (II(a) $\Longrightarrow$ II(e))**.** *Let G be a countable locally finite group. Then for any G-SFT X, there exists a unique measure of maximal entropy.*

Proof. Since shift actions of countable amenable groups are expansive, the map $\mu \mapsto h_\mu(X)$ is upper semi-continuous [15, Theorem 2.1], and so X has a measure of maximal entropy $\mu \in \mathcal{M}(X)$ such that $h_\mu(X) = h(X)$.

By Lemma 1.3.19, there exists $\{H_n\}_{n=1}^\infty$ and H_n-SFTs Y_n such that $\mathfrak{B}[\{Y_n\}]$ is a basis for the topology on X, and therefore also generates the Borel σ-algebra on X. Furthermore, since $X = Y_n^{\uparrow G}$, Lemma 1.2.7 gives that

$$h(X) = h(Y_1) = h(Y_n) = \frac{\log(|Y_1|)}{|H_1|} = \frac{\log(|Y_n|)}{|H_n|}$$

for all $n \in \mathbb{N}$. Also note that $\{H_n\}$ is a Følner sequence for G, and therefore

$$h_\mu(X) = \inf_n \frac{H_\mu(X, H_n)}{|H_n|}.$$

As such, we obtain $h_\mu(X) \leq \frac{H_\mu(X,H_n)}{|H_n|}$ for all $n \in \mathbb{N}$. But

$$H_\nu(X, H_n) \leq \log(|\mathcal{L}_{H_n}(X)|) = \log(|Y_n|)$$

for any $\nu \in \mathcal{M}(X)$, and therefore

$$\frac{\log(|Y_n|)}{|H_n|} = h(X) = h_\mu(X) \leq \frac{H_\mu(X, H_n)}{|H_n|} \leq \frac{\log(|Y_n|)}{|H_n|},$$

so all of these quantities must be equal, which further implies that for every n and $y \in Y_n$, we have $\mu[y] = \frac{1}{|Y_n|}$. This is true for any $n \in \mathbb{N}$, and therefore any measure of maximal entropy must take these specific values for every element of $\mathfrak{B}[\{Y_n\}]$. By the Carathéodory Extension Theorem, there exists a unique Borel probability measure with these properties, and therefore there exists only one measure of maximal entropy. $\square$

Though the previous proof does not explicitly mention how to construct the measure of maximal entropy, its construction is fairly simple. For a countable locally finite group G and G-SFT X, take some $H \ll G$ and H-SFT Y such that $X = Y^{\uparrow G}$. Let ν be a measure on Y defined by $\nu(y) = \frac{1}{|Y|}$ for all $y \in Y$. Then for any $C \in \mathscr{C}(H\backslash G)$, the pushforward measure $\mu = (\nu)^C \circ \kappa_C^{-1}$ is an invariant measure of maximal entropy for X. Informally, μ is the uniform measure on X, which is obtained as the push forward of a product measure under a construction function. It can also be shown that μ is independent of choice of H and Y for which $X = Y^{\uparrow G}$.

For the converse result, we give an SFT on any non-locally finite group which has multiple measures of maximal entropy.

Lemma 1.3.21 ((e) $\Longrightarrow$ (a))**.** *Let G be a countable amenable non-locally finite group. Then there exists a G-SFT X which has multiple measures of maximal entropy.*

Proof. Since $h(2_H^{\uparrow G}) = 0$, the Variational Principle gives that for all $\mu \in \mathcal{M}(X)$, we have $0 \leq h_\mu(2_H^{\uparrow G}) \leq h(2_H^{\uparrow G}) = 0$, and so $h_\mu(2_H^{\uparrow G}) = h(2_H^{\uparrow G})$. This means every measure $\mu \in \mathcal{M}(X)$ is a measure of maximal entropy.

Since 0^G and 1^G are both elements of $2_H^{\uparrow G}$, the two Dirac measures δ_{0^G} and δ_{1^G} are distinct, and since both 0^G and 1^G are fixed points they are both invariant, and therefore contained within $\mathcal{M}(2_H^{\uparrow G})$. As such, $2_H^{\uparrow G}$ has at least 2 measures of maximal entropy. $\square$

1.4 Final Remarks

The main results of this chapter gives that the class of locally finite groups presents interesting dynamical behaviors that are unexpected in general. This combined with the converse results, which show these interesting behaviors are unique to locally finite groups, gives insights into the types of groups where interesting behavior is possible. As mentioned in the introduction of this chapter, Theorem II(b) gives that the only groups for which there are only trivial zero-entropy dynamics are precisely the locally finite groups, and this property indirectly answers in the affirmative Question 3.19 of Barbieri [5]: "Does there exist an amenable group G and a G-SFT which does not contain a zero-entropy G-SFT?" For any countable locally finite group G, take any finite $H \ll G$ with $|H| > 1$, and pick any H-SFT Y which does not contain any fixed points. Then $X = Y^{\uparrow G}$ also does not contain a fixed point, and therefore contains no zero-entropy SFTs. This answer to the question leads to the following refinement of the question, as infinite locally finite groups are necessarily infinitely generated:

Question 1.4.1. Does there exist an infinite, *finitely generated* amenable group G and a G-SFT which does not contain a zero-entropy G-SFT?

Theorem II(d) also aids in the overall classification of the possible sets which are attainable as the set of entropies of SFTs on a specific amenable group. In the case that G is locally finite, $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$ (and in particular, an exact form for $\mathcal{E}(G)$ is known). In the case that G is not periodic, then it contains an element of infinite order (and therefore a subgroup isomorphic to $\mathbb{Z}$), and thus by Lemma 1.2.14, $\mathcal{E}(\mathbb{Z}) \subset \mathcal{E}(G)$. Though more research is needed to classify $\mathcal{E}(G)$ exactly for these types of groups (such as the work of Barbieri [5]), at least it is known that $\mathbb{Z}$-SFT entropies are attainable. The remaining class of groups are the finitely generated amenable torsion groups. We have shown in Lemma 1.3.17 that $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$ does imply that the G is periodic, however it is unclear whether the following can be answered in the affirmative:

Question 1.4.2. If G is a countable amenable group such that $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$, then must it be the case that G is locally finite? If not, is

$$\mathcal{E}(G) = \left\{ \frac{\log(n)}{|H|} : H \ll G, n \in \mathbb{N} \right\}$$

sufficient to conclude that G is locally finite?

Answering either of these questions in the affirmative would permit the locally non-torsion statement in II(d) to be dropped, leading to a strictly stronger result. Following the method used in proving that $\mathcal{E}(G) \subset \mathbb{Q}^+_{\log}$ implies periodicity, it would suffice to produce for any finitely generated, amenable, torsion group H, an H-SFT with entropy outside of $\mathbb{Q}^+_{\log}$. Then, for any amenable torsion group G which is not locally finite, it must contain a finitely generated torsion subgroup H (potentially the whole group), and this SFT can be defined on H, and then freely extended to G with the same entropy. Defining SFTs is not difficult in general; the primary difficulty is in computing their entropy, especially when arbitrary finitely generated groups are considered.

Including strengthening statement II(d), there are likely other statements which could be added to Theorems I and II. The types of dynamical properties explored in this work are by no means exhaustive, so future work may be able to add to these theorems, and any such work will likely use free extensions extensively as they have been used here. In addition to extending these theorems, expanding the theory of free extensions may be fruitful in the study of shifts on groups. For instance, while the forward direction of Lemma 1.2.10 is true in greater generality using the more general embeddings of Barbieri [5], the reverse direction for the specific case of free extensions is a new result to the knowledge of the author. Lemmas 1.2.14, 1.2.16, and 1.2.17 indicate that the study of SFTs, sofic shifts, and strongly irreducible shifts may be reduced to the study of such shifts on finitely generated groups. Furthermore, it is possible to take a minimal such finitely generated subgroup, so that the shift may not be further reduced from the perspective of free extensions. Such shifts may be considered *intrinsic* to the group, in the sense that they do not arise as the free extension of any shift on a proper subgroup.

Given that Lemmas 1.2.14, 1.2.16, and 1.2.17 give strong connections between free extensions, and the finite type, sofic, and strongly irreducible properties, along with Lemmas 1.2.10 and 1.2.11 giving that the finite type and strongly irreducible properties transfer readily between a free extension and its base shift, there is some indication that a similar result may exist for sofic shifts. Lemmas 1.2.12 and 1.2.10 readily give that if a shift is sofic, then any free extension of it is also sofic, however the converse result is not so simple. Jeandel first posed whether the free $\mathbb{Z}^2$-extension of a $\mathbb{Z}$-shift X being sofic implies that X is sofic, which has remained open since at least 2011 [42]. We may say a group G *has property S* if for any subgroup $H \leq G$ and H-shift Y, the G-shift $Y^{\uparrow G}$ being sofic implies that Y is sofic. Jeandel's question may then be posed more generally for all groups as:

Question 1.4.3. Which groups have property S?

By Theorem I and Lemma 1.2.10, we have that any locally finite group has property S, and so there are groups with this property. However, not all groups have property S, as Barbieri, Sablik, and Salo have shown that a certain class of non-amenable G do not have property S [6]. It remains to be seen whether groups such as $\mathbb{Z}$ and $\mathbb{Z}^2$ have property S, and perhaps whether amenable or sofic groups have property S.

Lastly, the mere existence of the two main theorems suggests that it may be possible to classify other dynamical properties by properties of the group, such as property S. To the knowledge of the author, these results may be the only results in symbolic dynamics that gives implications about the group only from dynamical properties of the group, let alone a complete characterization of the group by dynamical properties.

By assuming additional structure on the group, it may be possible to characterize other dynamical properties by this structure, and derive similar theorems as Theorems I and II for other classes of groups.

CHAPTER 2: The Completion of a Dynamical System and its Uses

In dynamical system, one of the core goals of classical theory is to understand the long term behavior of a transformation T from some state space X to itself. Whether it be to make predictions about the upcoming weather [27], or analyzing the spread of disease within a population [22], understanding how systems evolve over time allows us to better understand and predict the future. In the realm of ergodic theory, we seek to do this by finding the invariant distributions of systems, which in some sense captures the long term statistical behavior. Such distributions encapsulate an equilibrium, as these distributions do not change as the system evolves [48]. Additionally, if we pick a single state of the system and observe its empirical distribution as the state evolves over time, in the limit we generally expect this empirical distribution to converge to an equilibrium state. This idea is formalized in the various ergodic theorems, with Birkhoff's pointwise ergodic theorem being one of the most celebrated results [9].

Beyond the study of systems which evolve with respect to a single transformation, one might instead consider a continuum of transformations $\{T_r\}_{r\in\mathbb{R}}$ to represent continuous evolution over time. Techniques for analyzing such systems are fairly similar to discrete systems, since any such system can be discretized at an arbitrarily small time scale. Ergodic theorems for these systems also largely as a consequence of ergodic theorems for discrete transformations using Riemann sums ((iii), page 35 [52]). One might also consider an arbitrary group of transformations, although in this case, treating these transformations as encoding evolution over time becomes difficult or impossible. Instead, these transformations are seen as a collection of possible transitions or merely a set of invariance constraints. In general, studying arbitrary groups of transformations can lead to seemingly paradoxical results such as the Banach-Tarski paradox [4], so group dynamics are generally restricted to the amenable case in order to avoid this. In this case, many techniques for analysis have been developed [33], and Lindenstrauss has given a pointwise ergodic theorem for amenable groups, including non-discrete ones [37].

In general however, when the focus is on characterizing the invariant measures, there is not much of a difference between these different circumstances. At their core, we have a space X along with some collection of transformations $\mathcal{T}$ which we see as imposing invariance conditions on the space. At this level of generality, many standard techniques cease to hold due to even the most minor assumptions made on $\mathcal{T}$ which do not necessarily hold. One notable exception is Choquet Theory, which does tackle the set of invariant measures for arbitrary collections of transformations (see Chapter 12 [43]). In this chapter, we seek to expand on the theory of general dynamical systems in order to develop tools to aid in the characterization of the invariant measures of systems.

Formally, we let X be a compact metrizable topological space endowed with the Borel σ-algebra, and let $\mathcal{T}$ be any collection of Borel measurable transformations $T : X \to X$. The pair $(X, \mathcal{T})$ is called a dynamical system. We say that a Borel probability measure μ on X is $\mathcal{T}$-invariant if for every Borel set $E \subset X$ and every $T \in \mathcal{T}$, we have that $\mu(E) = \mu(T^{-1}(E))$. Our principal goal is to give some sort of characterization of the set of invariant measures. The exact form of such a characterization will depend heavily on the exact nature of the system in consideration, so our goal is to develop tools which may be helpful in doing so, rather than providing some universal characterization theorem, which may be impossible, given the complexity which can arise for systems [13].

An immediate observation that can be made about a dynamical system $(X, \mathcal{T})$ is that the more transformations $\mathcal{T}$ contains, the more restricted the set of invariant measures becomes. In the extreme case, where $\mathcal{T}$ is the set of all Borel measurable transformations from X to itself, a simple characterization is possible. If $|X| = 1$, then the set of invariant measures is just the singleton point mass δ_x, which is the only Borel probability measure on this space. Of course, this case is trivial, and we are generally uninterested it in. However, if $|X| > 1$, then the set of invariant measures is empty, because no probability measure can be invariant with respect to the constant transformations $T(x) = x_0$ and $T(x) = x_1$ where $x_0 \neq x_1$ are two distinct points in X. At the other extreme, if $\mathcal{T}$ is empty or consists only of the identity transformation on the space X, then every probability measure on X is invariant. As such, if we start with the set of transformations $\mathcal{T}$, where all probability measures are invariant, and begin to add in transformations, eventually we add in enough constraints in order for there to be no invariant measures. This begs the question, how exactly does this occur? Is it the case that every time we add in a transformation, the set of invariant measures shrinks? Immediately, the answer to this is no, because if a measure is invariant for T, it is also automatically invariant for $T^2 = T \circ T$. This means that sometimes, when we add in a transformation, it does not change the set of invariant measures. Intuitively however, the more transformations we have, the more invariance constraints we place on measures, and so it is more likely that we can find a precise characterization of the invariant measures. So, this leads us to the following question.

Question 2.0.1. Given a dynamical system $(X, \mathcal{T})$, how can we ascertain whether or not a transformation $T \notin \mathcal{T}$ can be added to $\mathcal{T}$ without changing the set of invariant measures? Is there a maximal collection of transformations such that, no matter which $T \notin \mathcal{T}$ we pick, adding T to this maximal collection would change the set of invariant measures?

This question is addressed in Section 2.3 using the notion of the completion $(X, \mathcal{T}^*)$ of a dynamical system $(X, \mathcal{T})$, which answers the second question above in the affirmative. This completion is always nonempty, is closed under composition, is closed under taking the inverses of bijections, is stable under wobbling (see Subsection 2.3.2), and is closed in a novel topology on the set of measurable transformations from X to itself (see Section 2.2), given by Theorem 2.3.26.

Interestingly, this completion also seems to appear anywhere that invariant measures are involved. For instance, given some collection of measures $\mathcal{M}$, we say that a set E is $(\mathcal{M}, \mathcal{T})$-invariant if for every $\mu \in \mathcal{M}$ and $T \in \mathcal{T}$, we have $\mu(E \triangle T^{-1}(E)) = 0$, where $\triangle$ denotes the symmetric difference. It turns out that with $\mathcal{M}$ the collection of invariant measures, under the assumption that it is closed, the set of $(\mathcal{M}, \mathcal{T})$-invariant sets coincides exactly with the set of $(\mathcal{M}, \mathcal{T}^*)$-invariant sets (Proposition 2.3.32), despite the fact that the latter is a much stricter requirement.

Going beyond more basic constructs such as invariant sets, one would of course consider pointwise ergodic theorems, and what dynamical systems have such theorems. We present in Section 2.4 the novel notion of a Birkhoff system, for which a version of the pointwise ergodic theorem holds. This definition makes heavy use of the completion of a dynamical system, and holds when there is some method of taking ergodic averages which converges to an object which is in some sense invariant with respect to the completion of a dynamical system, rather than only the original system as most pointwise ergodic theorems are stated. Despite this seemingly stronger definition, it is possible to use existing pointwise ergodic theorems to prove that most classically studied systems are Birkhoff systems, given by Theorems 2.4.11 and 2.4.13. In addition, the definition of a Birkhoff systems makes it almost immediate that if a system is Birkhoff, so is its completion. As a result, since the completion of dynamical systems are typically rather large in comparison to typical systems, large enough to no longer fall under the scope of more classical theory. As such, it is possible to transfer pointwise ergodic theorems from well studied systems to other systems using the completion. An example of this being done is given by Lemma 2.5.29.

Using Birkhoff systems, we also use completions to develop the notion of dynamical independence, which we use to provide a new characterization of ergodicity. While the precise definition of dynamical independence is somewhat technical, there are many ways to show that two sets are dynamically independent which do not involve identifying any invariant measures (Propositions 2.4.24 and 2.4.26). The characterization of ergogodicity is given by Theorem 2.4.22, and is stated rather simply; for Birkhoff systems, an invariant measure is ergodic if and only every pair of dynamically independent sets is probabilistically independent. As a result, it is possible to start with a dynamical system, identify dynamically independent pairs of sets within the system, and then every ergodic measure must be defined in a way that makes these pairs probabilistically independent. Being able to identify such structures can be very helpful in the characterization of the invariant measures of a system, as is demonstrated by Lemma 2.5.29.

We then turn our attention to applying these concepts to a broad class of dynamical systems which is of general interest in statistics. Using the completion, Birkhoff systems, and dynamical independence, we prove many results about dynamical systems on sequence spaces which act on the index sets. A notable and celebrated result of this variety is De Finetti's Theorem, for which we prove an extension. Classically, De Finetti's Theorem can be interpreted as a characterization of the invariant measures of a particular system, and our extension applies to a broad family of particular systems. Furthermore, Theorem 2.5.43 gives a further extension of De Finetti's Theorem which gives a characterization of the invariant measures for an even broader class of systems. This theorem enables us to easily give a wide variety of restricted De Finetti type results with little to no work, as is given by Example 2.5.47.

2.0.1 Outline

We begin with Section 2.1, which states and proves some standard results in functional analysis and the theory of probability measures on compact metrizable spaces. This section also outlines a lot of important notation for the remainder of the chapter. In Section 2.2, we develop a novel topology on the space of measurable transformations between a compact metrizable space X and a compact metrizable space Y, which naturally arises in the study of completions. In Section 2.3, we define dynamical systems and give their properties, and also define and develop basic theory around the completion of a dynamical system. In Section 2.4, we define Birkhoff systems, prove that many systems studied in classical contexts are Birkhoff under this definition, and then define the notion of dynamical independence and prove its relation to ergodic measures. Then, in Section 2.5 we turn our attention to studying the invariant measures of system on sequence spaces, as well as product systems, joinings of systems, and power systems, and we use the completion, Birkhoff systems, and dynamical independence to this extent. In Section 2.5.4, we prove our extension of De Finetti's theorem to a broad class of systems on sequence spaces. Finally, in Section 2.6, we summarize the results of this chapter and indicate possible directions for future work.

2.1 Preliminaries

Within this section, we define and state many well known results that do not necessarily appear all in one place, and so we provide proofs for the majority of these results. We also use some non-standard notation for many of these concepts, as this notation will be easier to use than more standard notation in future sections, so it is simpler to re-prove many of these results using the notation of this chapter. We begin with the following simple definition.

Definition 2.1.1. For a set S, let $\mathscr{F}(S)$ denote the set of finite subsets of S.

Next, we define the notion of a space used throughout this chapter. Briefly, a space is a compact metrizable topological space endowed with the Borel σ-algebra.

Definition 2.1.2. A *space* is a triple $(X, \tau, \mathscr{A})$ where X is a set, τ is a compact metrizable topology on X, and $\mathscr{A}$ is the σ-algebra generated by τ, otherwise known as the Borel σ-algebra. For conciseness, we refer to spaces only by the set X on which they are defined, and refer to the topology on X by τ_X and the σ-algebra on X by $\mathscr{A}_X$.

Given such a space, we define some functional spaces from this space into the reals. These constructions are well known, and we use fairly standard notation.

Definition 2.1.3. For a topological space X let $\mathbb{R}^X$ denote the set of all functions from X to $\mathbb{R}$. Let $B(X) \subset \mathbb{R}^X$ the set of bounded $\mathscr{A}_X$-measurable functions from X to $\mathbb{R}$ with its usual σ-algebra, and $C(X) \subset B(X)$ denote the set of bounded τ_X-continuous functions from X to $\mathbb{R}$ with its usual topology. We endow $C(X)$ and $B(X)$ with the supremum norm

$$\|f\| = \sup_{x \in X} |f(x)|,$$

making each a Banach space [47].

Additionally, since spaces are endowed with a σ-algebra, we may define measures on spaces, in particular we will solely be interested in probability measures. As such, we will use the terms probability measure and measure interchangeably.

Definition 2.1.4. For any space X, let $\mathcal{P}X$ denote the set of probability measures on the Borel σ-algebra of X. Additionally, for any $f \in B(X)$, let $\mathcal{L}f : \mathcal{P}X \to \mathbb{R}$ be the function defined by

$$\mathcal{L}f(\mu) = \int_X f \, \mathrm{d}\mu.$$

We endow $\mathcal{P}X$ with the weakest topology $\tau_{\mathcal{P}X}$ such that for each $f \in C(X)$, the map $\mathcal{L}f$ is continuous. In other words, we endow $\mathcal{P}X$ with the weakest topology such that $\mathcal{L}C(X) = \{\mathcal{L}f : f \in C(X)\} \subset C(\mathcal{P}X)$. This is the well known topology on $\mathcal{P}X$ inherited as a subset of $C(X)^*$ endowed with the weak* topology, which is compact and metrizable (Section 21.5 of [47]). We also endow $\mathcal{P}X$ with the Borel σ-algebra $\mathscr{A}_{\mathcal{P}X}$ which is generated by $\tau_{\mathcal{P}X}$, and this makes $(\mathcal{P}X, \tau_{\mathcal{P}X}, \mathscr{A}_{\mathcal{P}X})$ a space. As such, $\mathcal{P}^2X = \mathcal{P}(\mathcal{P}X)$ is well defined, and we may take $\tau_{\mathcal{P}X}$ to be the weakest topology such that for each $\psi \in C(\mathcal{P}X)$, the map $\mathcal{L}\psi$ is continuous. This continues for any finite n, obtaining a space $(\mathcal{P}^nX, \tau_{\mathcal{P}^nX}, \mathscr{A}_{\mathcal{P}^nX})$ for every $n \in \mathbb{N}$.

As a consequence of identifying $\mathcal{P}X$ as a subset of a vector space (in particular, the intersection of a cone and a hyperplane), it is convex, meaning that for any $\mu, \nu \in \mathcal{P}X$ and $t \in (0, 1)$, we have that $t\mu + (1 - t)\nu \in \mathcal{P}X$.

For any element x of a space X, we may take the point mass δ_x which is a probability measure on X that is concentrated at the point x. Viewing δ as a map from X to $\mathcal{P}X$, this transformation demonstrates the strong connection between the topology on $\mathcal{P}X$ and the original topology on X.

Proposition 2.1.5. *Let X be a space. Then the map $\delta : X \to \mathcal{P}X$ defined by $\delta(x) = \delta_x$ (the measure in $\mathcal{P}X$ for which $\delta_x(\{x\}) = 1$) is a topological embedding of X into $\mathcal{P}X$.*

Proof. First, clearly the map δ is injective, since if $\delta_x = \delta_y$, then $\delta_x(\{x\}) = 1 = \delta_y(\{y\}) = \delta_x(\{y\})$, so if $x \neq y$, then $\delta_x(\{x, y\}) = 2$ by disjoint additivity.

Now, suppose that $\{x_n\}_{n\in\mathbb{N}} \subset X$ is a sequence which converges to x in X. For $f \in C(X)$, we then have by the continuity of f that

$$\lim_{n\to\infty} \mathcal{L}f(\delta_{x_n}) = \lim_{n\to\infty} \int_X f \, \mathrm{d}\delta_{x_n} = \lim_{n\to\infty} f(x_n) = f(x) = \mathcal{L}f(\delta_x),$$

and as this holds for every $f \in C(X)$, we therefore have that $\{\delta_{x_n}\}_{n\in\mathbb{N}}$ converges to δ_x in $\mathcal{P}X$. This shows that δ is continuous, since X is metrizable, so sequential continuity and continuity coincide.

Finally, since X is compact metrizable, any closed subset $F \subset X$ is also compact, and so $\delta(F)$ is compact in $\mathcal{P}X$. Since $\mathcal{P}X$ is compact metrizable, it follows that $\delta(F)$ is closed, and therefore δ is a closed map. As such, the restricted map $\delta : X \to \delta(X)$ is a homeomorphism, which makes δ a topological embedding of X into $\mathcal{P}X$. □

Additionally, for spaces and countable products of spaces, the notion of taking product measures is well defined and even continuous.

Proposition 2.1.6. *Let I be a countable set and X_i be a space for every $i \in I$. Then the map $\otimes : \prod_{i\in I} \mathcal{P}X_i \to \mathcal{P}\prod_{i\in I} X_i$ defined by*

$$\otimes(\{\mu_i\}) = \bigotimes_{i\in I} \mu_i$$

is a topological embedding of $\prod_{i\in I} \mathcal{P}X_i$ into $\mathcal{P}\prod_{i\in I} X_i$.

Proof. By Theorem 2 of [45], the map $\otimes$ is continuous and injective (noting that every Borel Probability measure in a compact metrizable space is Radon, and therefore τ-smooth, and that the product Borel σ-algbra and the Borel σ-algebra of the product space coincide since both are compact metrizable and therefore second countable). Since $\prod_{i\in I} \mathcal{P}X_i$ is compact, the image of any closed and therefore compact set under $\otimes$ is then compact and therefore closed in $\mathcal{P}\prod_{i\in I} X_i$. As such, $\otimes$ is closed, so the restriction of $\otimes$ onto its image is a homeomorphism, making it a topological embedding. □

2.1.1 Properties of $\mathcal{L}$

We now turn our attention to proving many properties of the map $\mathcal{L}$ which takes a function $f \in B(X)$ and produces a mapping from $\mathcal{P}X$ to $\mathbb{R}$. We begin with the following definition.

Definition 2.1.7. Let C be a convex set, and V a vector space. A function $f : C \to V$ is *affine* if for every $x, y \in C$ and $t \in (0,1)$, we have

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y).$$

We now show that the image of f under $\mathcal{L}$ is always affine.

Lemma 2.1.8. *Let X be a space, and $f \in B(X)$. Then $\mathcal{L}f$ is affine.*

Proof. Let $\mu, \nu \in \mathcal{P}X$, and let $t \in (0,1)$ so that $p\mu + (1-p)\nu \in \mathcal{P}X$. Then

$$\begin{aligned}\mathcal{L}f(p\mu + (1-p)\nu) &= \int_X f \,\mathrm{d}(p\mu + (1-p)\nu) = p\int_X f \,\mathrm{d}\mu + (1-p)\int_X f \,\mathrm{d}\nu \\ &= p\mathcal{L}f(\mu) + (1-p)\mathcal{L}f(\nu),\end{aligned}$$

and thus $\mathcal{L}f$ is affine. The middle equality follows from the fact that $\mathcal{P}X$ is identified with a convex subset of the vector space $C(X)^*$. □

Next, we may show that $\mathcal{L}$, as an operator, is linear. We further strengthen this result below.

Lemma 2.1.9. *Let X be a space. Then $\mathcal{L} : B(X) \to \mathbb{R}^{\mathcal{P}X}$ is linear.*

Proof. Let $a \in \mathbb{R}$ and $f, g \in B(X)$. Then for any $\mu \in \mathcal{P}X$, by the linearity of integrals, we have

$$\mathcal{L}[af+g](\mu) = \int_X af + g \,\mathrm{d}\mu = a\int_X f \,\mathrm{d}\mu + \int_X g \,\mathrm{d}\mu = a\mathcal{L}f(\mu) + \mathcal{L}g(\mu),$$

and therefore $\mathcal{L}$ is linear. □

On top of being linear, for $f \in B(X)$ (which is bounded), we have that $\mathcal{L}f$ is also bounded, and the bound for $\mathcal{L}f$ is precisely that for f.

Lemma 2.1.10. *Let X be a space and $f \in B(X)$. Then $\mathcal{L}f$ is bounded, and*

$$\sup_{\mu \in \mathcal{P}X} |\mathcal{L}f(\mu)| = \|f\|.$$

Proof. First, we have for every $\mu \in \mathcal{P}X$ that

$$|\mathcal{L}f(\mu)| = \left|\int_X f \,\mathrm{d}\mu\right| \leq \int_X |f| \,\mathrm{d}\mu \leq \int_X \|f\| \,\mathrm{d}\mu = \|f\|,$$

and therefore $\mathcal{L}f$ is bounded, and furthermore

$$\sup_{\mu \in \mathcal{P}X} |\mathcal{L}f(\mu)| \leq \|f\|.$$

Next, for $\epsilon > 0$, there exists $x \in X$ such that $|f(x)| > \|f\| - \epsilon$ by definition of $\|f\|$. As such,

$$\sup_{\mu \in \mathcal{P}X} |\mathcal{L}f(\mu)| \geq |\mathcal{L}f(\delta_x)| = \left|\int_X f \,\mathrm{d}\delta_x\right| = |f(x)| > \|f\| - \epsilon,$$

and so $\sup_{\mu \in \mathcal{P}X} |\mathcal{L}f(\mu)| \geq \|f\|$ since $\epsilon > 0$ is arbitrary. This gives the desired conclusion. □

Before we outright show that $\mathcal{L}$ is a continuous operator, we first need the following result which follows readily from the Bounded Convergence Theorem.

Lemma 2.1.11. *Let X be a space. If $\{f_n\}_{n\in\mathbb{N}} \subset B(X)$ is a uniformly bounded sequence which converges pointwise to $f \in \mathbb{R}^X$, then $f \in B(X)$ and*

$$\lim_{n\to\infty} \mathcal{L}f_n = \mathcal{L}f$$

pointwise for all $\mu \in \mathcal{P}X$.

Proof. Let $\mu \in \mathcal{P}X$. Since the pointwise limit of measurable functions is measurable, $f \in B(X)$ and by the Bounded Convergence Theorem,

$$\lim_{n\to\infty} \mathcal{L}f_n(\mu) = \lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu = \mathcal{L}f(\mu).$$

□

Next, we have the following technical result which gives equivalent statements about the measurability of certain maps from $\mathcal{P}X$ for a space X. We use this result immediately, and again for the main theorem characterizing the standard σ-algebra $\mathcal{P}X$ is endowed with as a space.

Lemma 2.1.12. *Let X be a space, and let $\mathscr{A}$ be any σ-algebra on $\mathcal{P}X$. Then the following are equivalent.*

(a) For every $f \in C(X)$, $\mathcal{L}f$ is $\mathscr{A}$-measurable.

(b) For every compact $K \subset X$, $\mathcal{L}\chi_K$ is $\mathscr{A}$-measurable.

(c) For every $E \in \mathscr{A}_X$, $\mathcal{L}\chi_E$ is $\mathscr{A}$-measurable.

(d) For every $f \in B(X)$, $\mathcal{L}f$ is $\mathscr{A}$-measurable.

Proof. First suppose (a), let d be a metric for X, let $K \subset X$ be a compact, and for each n, define the function $f_n : X \to \mathbb{R}$ by

$$f_n(x) = \max\{1 - nd(x, K), 0\},$$

where $d(x,K) = \inf\{d(x,k) : k \in K\}$. Since K is closed, $d(x,K) = 0$ if and only if $x \in K$. Then each f_n is continuous as $d(\cdot, K)$ is continuous, and the remaining operations preserve continuity. Furthermore, if $x \in K$, then $d(x,K) = 0$ and

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \max\{1 - nd(x,K), 0\} = \lim_{n\to\infty} \max\{1, 0\} = 1.$$

Additionally, for $x \notin K$, let $\epsilon = d(x,K)$, and choose N such that $\frac{1}{N} < \epsilon$. As such, $N\epsilon > 1$, and for each $n \geq N$,

$$f_n(x) = \max\{1 - nd(x,K), 0\} = \max\{1 - n\epsilon, 0\},$$

however $1 - n\epsilon \leq 1 - N\epsilon < 1 - 1 = 0$, and therefore $\max\{1 - n\epsilon, 0\} = 0$. Then $f_n(x) = 0$ for $n \geq N$, and therefore

$$\lim_{n\to\infty} f_n(x) = 0.$$

As such, the pointwise limit of the f_n is exactly χ_K, the characteristic function of K. Furthermore, each $f_n \leq 1$, so by Lemma 2.1.11,

$$\lim_{n\to\infty} \mathcal{L} f_n = \mathcal{L}\chi_K$$

pointwise. By (a), each f_n is continuous, and therefore each $\mathcal{L} f_n$ is measurable. With $\mathcal{L}\chi_K$ the pointwise limit of these functions, we have that $\mathcal{L}\chi_K$ must also be measurable. This proves (b).

Now suppose that (b) holds, and let

$$\mathscr{L} = \{E \in \mathscr{A}_X : \mathcal{L}\chi_E \text{ is measurable}\},$$

and note that $\mathscr{L}$ contains every compact subset of X, including both $\varnothing$ and X. Now, for any $E \in \mathscr{L}$, $\mathcal{L}\chi_E$ is measurable, and therefore

$$\mathcal{L}\chi_{X\setminus E}(\mu) = \int_X \chi_{X\setminus E}\,\mathrm{d}\mu = \mu(X\setminus E) = 1-\mu(E) = 1-\int_X \chi_E\,\mathrm{d}\mu = 1-\mathcal{L}\chi_E(\mu),$$

so $\mathcal{L}\chi_{X\setminus E}$ is measurable. As such, $X\setminus E \in \mathscr{L}$. Lastly, let $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{G}$ be mutually disjoint and define $E = \bigcup_{n=1}^{\infty} E_n$. Then for every $\mu \in \mathcal{P}X$, by countable additivity, we have

$$\begin{aligned}\mathcal{L}\chi_E(\mu) = \int_X \chi_E\,\mathrm{d}\mu = \mu(E) = \mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty}\mu(E_n)\\ = \sum_{n=1}^{\infty}\int_X \chi_{E_n}\,\mathrm{d}\mu = \sum_{n=1}^{\infty}\mathcal{L}\chi_{E_n}(\mu),\end{aligned}$$

and so $\mathcal{L}\chi_E = \sum_{i=1}^{\infty}\mathcal{L}\chi_{E_n}$. Since finite sums of measurable functions are measurable, and pointwise limits of measurable functions are measurable, this gives that $\mathcal{L}\chi_E$ is measurable. Therefore $E \in \mathscr{L}$, and this makes $\mathscr{L}$ a λ-system.

Finally, since the set of compact subsets of X forms a π-system $\mathscr{P}$, and $\mathscr{P} \subset \mathscr{L}$, we have by the $\pi-\lambda$ Theorem that the σ-algebra generated by the compact subsets of X is contained within $\mathscr{L}$. But the σ-algebra generated by all compact subsets of X is precisely $\mathscr{A}_X$, the Borel σ-algebra on X. As such, $\mathscr{A}_X \subset \mathscr{L}$, but by definition $\mathscr{L} \subset \mathscr{A}_X$, and so $\mathscr{L} = \mathscr{A}_X$. Therefore, for any measurable set $E \subset X$, the function $\mathcal{L}\chi_E$ is measurable. This proves (c).

Next suppose (c) holds. For any simple function $f = \sum_{i=1}^{k} a_i\chi_{E_i}$ we have by Lemma 2.1.9 that

$$\mathcal{L}f = \sum_{i=1}^{k} a_i\mathcal{L}\chi_{E_i}.$$

By (c), each $\mathcal{L}\chi_{E_i}$ is measurable, and since the linear combination of measurable functions is measurable, we have that $\mathcal{L}f$ is measurable. Lastly, for any arbitrary $f \in B(X)$, there is an increasing sequence $\{f_n\}$ of simple functions such that f is the pointwise limit of the f_n, so by Lemma 2.1.11,

$$\lim_{n\to\infty} \mathcal{L}f_n = \mathcal{L}f$$

With each f_n simple, each $\mathcal{L}f_n$ is measurable, and therefore, being the pointwise limit of measurable functions, $\mathcal{L}f$ is measurable. This proves that (d) holds.

Finally, (d) immediately implies (a) as $C(X) \subset B(X)$. □

With this result, we prove the main theorem about $\mathcal{L}$, which is that it is an isometric linear operator from the Banach space $B(X)$ to $B(\mathcal{P}X)$, and from $C(X)$ to $C(\mathcal{P}X)$.

Theorem 2.1.13. *Let X be a space. Then $\mathcal{L} : B(X) \to B(\mathcal{P}X)$ is an isometric linear embedding of $B(X)$ into $B(\mathcal{P}X)$, and of $C(X)$ into $C(\mathcal{P}X)$.*

Proof. First, note that the map is well defined, because by definition of the topology on $\mathcal{P}X$, we have $\mathcal{L}C(X) \subset C(\mathcal{P}X)$. Since $C(\mathcal{P}X) \subset B(\mathcal{P}X)$, for each $f \in C(X)$ we have $\mathcal{L}f$ is $\mathscr{A}_{\mathcal{P}X}$-measurable, and therefore by Lemma 2.1.12 for each $f \in B(X)$ we have $\mathcal{L}f$ is $\mathscr{A}_{\mathcal{P}X}$-measurable, or in other words $\mathcal{L}f \in B(\mathcal{P}X)$. As such, $\mathcal{L}B(X) \subset B(\mathcal{P}X)$, so this map is well defined between its domain and co-domain.

Next, we have $\mathcal{L}$ is linear by Lemma 2.1.9.

Finally, for any $f \in B(X)$, we have $\mathcal{L}f \in B(\mathcal{P}X)$ and by Lemma 2.1.10 we obtain

$$\|\mathcal{L}f\| = \sup_{\mu\in\mathcal{P}X} |\mathcal{L}f(\mu)| = \|f\|,$$

and therefore $\mathcal{L}$ is an isometry. □

It follows from this result that for $f \in B(X)$, we have $\mathcal{L}f \in B(\mathcal{P}X)$, and thus we have $\mathcal{L}^2 f = \mathcal{L}(\mathcal{L}f) \in B(\mathcal{P}^2 X)$, and so on, obtaining that for $f \in B(X)$, we have $\mathcal{L}^n f \in B(\mathcal{P}^n X)$ for every $n \in \mathbb{N}$. Similarly, for $f \in C(X)$, we have $\mathcal{L}f \in C(\mathcal{P}X)$, and so $\mathcal{L}^2 f \in C(\mathcal{P}^2 X)$, and so on, giving $\mathcal{L}^n f \in C(\mathcal{P}^n X)$ for every $n \in \mathbb{N}$.

Next, we have a map which is similar in nature to the maps δ and $\otimes$, however we require Theorem 2.1.13 in order to prove its properties. This map is called the barycenter map, and is essentially the expected value of a random measure, otherwise known as the intensity of a random measure (see [31]).

Proposition 2.1.14. *Let X be a space. Then the map $\beta : \mathcal{P}^2 X \to \mathcal{P}X$ defined by*

$$\beta[m](E) = \int_{\mathcal{P}X} \mu(E)\, m(\mathrm{d}\mu)$$

for $m \in \mathcal{P}^2 X$ and $E \in \mathscr{A}_X$, is a continuous affine surjection. Furthermore, for every $f \in B(X)$, we have that

$$\mathcal{L}f \circ \beta = \mathcal{L}^2 f.$$

Proof. First, for any $m \in \mathcal{P}^2 X$, it is necessary to prove that $\beta(m) \in \mathcal{P}X$. Indeed, we have

$$\beta[m](\varnothing) = \int_{\mathcal{P}X} \mu(\varnothing)\, m(\mathrm{d}\mu) = \int_{\mathcal{P}X} 0\, m(\mathrm{d}\mu) = 0,$$

and

$$\beta[m](X) = \int_{\mathcal{P}X} \mu(X)\, m(\mathrm{d}\mu) = \int_{\mathcal{P}X} 1\, m(\mathrm{d}\mu) = 1.$$

Now, for $\{E_n\}_{n\in\mathbb{N}} \subset \mathscr{A}_x$ disjoint, for each $\mu \in \mathcal{P}X$, note that by countable additivity, we have

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mathcal{L}\chi_{E_n}(\mu),$$

where this limit is of a monotonically increasing sequence of real numbers, as each $\mathcal{L}\chi_{E_n}$ is non-negative. Furthermore, by Theorem 2.1.13, we have that each $\mathcal{L}\chi_{E_n}$ is in $B(\mathcal{P}X)$, and so each is $\mathscr{A}_{\mathcal{P}X}$-measurable. Thus, as the finite sum of measurable functions is measurable, each $\sum_{n=1}^{N} \mathcal{L}\chi_{E_n}$ is, and so we have a monotonically increasing pointwise limit of measurable functions. By the Bounded Convergence Theorem and linearity of integrals, we then have that

$$\begin{aligned}
\beta[m]\left(\bigcup_{n=1}^{\infty} E_n\right) &= \int_{\mathcal{P}X} \mu\left(\bigcup_{n=1}^{\infty} E_n\right) m(\mathrm{d}\mu) \\
&= \int_{\mathcal{P}X} \lim_{N\to\infty} \sum_{n=1}^{N} \mathcal{L}\chi_{E_n}(\mu)\, m(\mathrm{d}\mu) \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\mathcal{P}X} \mathcal{L}\chi_{E_n}(\mu)\, m(\mathrm{d}\mu) \\
&= \sum_{n=1}^{\infty} \int_{\mathcal{P}X} \mu(E_n)\, m(\mathrm{d}\mu) \\
&= \sum_{n=1}^{\infty} \beta[m](E_n),
\end{aligned}$$

and therefore $\beta(m)$ has countable disjoint additivity, which shows $\beta(m) \in \mathcal{P}X$.

Next, it is clear that β is a surjection, as for any $\mu \in \mathcal{P}X$, we have that $\delta_\mu \in \mathcal{P}^2X$, and for any $E \in \mathscr{A}_X$,

$$\beta[\delta_\mu](E) = \int_{\mathcal{P}X} \nu(E)\, \delta_\mu(\mathrm{d}\nu) = \mu(E),$$

and thus $\beta(\delta_\mu) = \mu$.

Then, for $m_1, m_2 \in \mathcal{P}^2X$ and $t \in (0,1)$, we have that $tm_1 + (1-t)m_2 \in \mathcal{P}^2X$, and

$$\begin{aligned}
\beta[tm_1 + (1-t)m_2](E) &= \int_{\mathcal{P}X} \mu(E)\, (tm_1 + (1-t)m_2)(\mathrm{d}\mu) \\
&= t\int_{\mathcal{P}X} \mu(E)\, m_1(\mathrm{d}\mu) + (1-t)\int_{\mathcal{P}X} \mu(E)\, m_2(\mathrm{d}\mu) \\
&= t\beta[m_1](E) + (1-t)\beta[m_2](E),
\end{aligned}$$

so β is affine.

We now prove that for any $f \in B(X)$, we have that $\mathcal{L}f \circ \beta = \mathcal{L}^2 f$. Let $m \in \mathcal{P}^2X$ and $E \in \mathscr{A}_X$. Noting that $\beta[m](E) = \int_X \chi_E \, \mathrm{d}\beta(m) = \mathcal{L}\chi_E(\beta(m)) = [\mathcal{L}\chi_E \circ \beta](m)$, we have

$$\begin{aligned}[\mathcal{L}\chi_E \circ \beta](m) &= \int_{\mathcal{P}X} \mu(E)\, m(\mathrm{d}\mu) = \int_{\mathcal{P}X} \int_X \chi_E \,\mathrm{d}\mu\, m(\mathrm{d}\mu) \\ &= \int_{\mathcal{P}X} \mathcal{L}\chi_E(\mu)\, m(\mathrm{d}\mu) = \mathcal{L}[\mathcal{L}\chi_E)](m) = \mathcal{L}^2\chi_E(m).\end{aligned}$$

As $m \in \mathcal{P}^2X$ was arbitrary, for every $E \in \mathscr{A}_X$, we have $\mathcal{L}\chi_E \circ \beta = \mathcal{L}^2\chi_E$. Now, for any simple function $f \in B(X)$, we have that $f = \sum_{i=1}^n a_i\chi_{E_i}$ for some $E_i \in \mathscr{A}_X$ and $a_i \in \mathbb{R}$, and by Theorem 2.1.13 and this previously proven fact that

$$\begin{aligned}\mathcal{L}f \circ \beta = \mathcal{L}\left(\sum_{i=1}^n a_i\chi_{E_i}\right) \circ \beta &= \left(\sum_{i=1}^n a_i\mathcal{L}\chi_{E_i}\right) \circ \beta \\ &= \sum_{i=1}^n a_i(\mathcal{L}\chi_{E_i} \circ \beta) = \sum_{i=1}^n a_i\mathcal{L}^2\chi_{E_i} = \mathcal{L}^2\left(\sum_{i=1}^n a_i\chi_{E_i}\right) \\ &= \mathcal{L}^2 f.\end{aligned}$$

Finally, for every $f \in B(X)$, there exists a monotonically increasing sequence of simple functions $\{f_n\}_{n\in\mathbb{N}} \subset B(x)$ whose pointwise limit is f. By Lemma 2.1.11, we have

$$\mathcal{L}f \circ \beta = \left(\lim_{n\to\infty} \mathcal{L}f_n\right) \circ \beta = \lim_{n\to\infty} \mathcal{L}f_n \circ \beta = \lim_{n\to\infty} \mathcal{L}^2 f_n = \mathcal{L}^2 f.$$

Finally, we show that β is continuous. To do so, by definition of the topology on $\mathcal{P}X$, it will suffice to show that for every $f \in C(X)$, that $\mathcal{L}f \circ \beta$ is continuous. But by the result of the previous paragraph, we have that $\mathcal{L}f \circ \beta = \mathcal{L}^2 f$, and since we have $f \in C(X)$, this means $\mathcal{L}f \in C(\mathcal{P}X)$, and therefore $\mathcal{L}^2 f = \mathcal{L}(\mathcal{L}f) \in C(\mathcal{P}^2X)$, which is clearly continuous. Therefore, $\mathcal{L}f \circ \beta$ is continuous for every $f \in C(X)$, which proves that β is continuous. $\square$

Finally, we require another technical result similar to Lemma 2.1.12, but for topologies on $\mathcal{P}X$.

Lemma 2.1.15. *Let X be a space, and let $D \subset C(X)$ be dense. Let τ be a topology on $\mathcal{P}X$. Then the following are equivalent.*

1. *For every $f \in D$, $\mathcal{L}f$ is τ-continuous.*

2. *For every $f \in C(X)$, $\mathcal{L}f$ is τ-continuous.*

Proof. Clearly (2) implies (1), and so suppose that for every $g \in D$, $\mathcal{L}g$ is τ-continuous, and let $f \in C(X)$. Since D is dense in $C(X)$, there is a sequence $\{f_n\}_{n\in\mathbb{N}} \subset D$ such that $\lim_{n\to\infty} \|f - f_n\| = 0$. By Lemmas 2.1.9 and 2.1.10, we have

$$\sup_{\mu\in\mathcal{P}X} |\mathcal{L}f(\mu) - \mathcal{L}f_n(\mu)| = \sup_{\mu\in\mathcal{P}X} |\mathcal{L}[f - f_n](\mu)| = \|f - f_n\|,$$

and thus $\{\mathcal{L}f_n\}_{n\in\mathbb{N}}$ converges uniformly to $\mathcal{L}f$. Since each $\mathcal{L}f_n$ is τ-continuous, the Uniform Limit Theorem gives that $\mathcal{L}f$ is τ-continuous. □

Using this Lemma and Lemma 2.1.12, we now give equivalent characterizations for $\mathscr{A}_{\mathcal{P}X}$ apart from its original definition as the Borel σ-algebra of the topological space $\mathcal{P}X$. We use this characterization in the following section in some proofs of the measurability of maps into $\mathcal{P}X$.

Theorem 2.1.16. *Let X be a space. Then $\mathscr{A}_{\mathcal{P}X}$, the σ-algebra generated by the topology on $\mathcal{P}X$, is equal to each of the following.*

1. *The smallest σ-algebra $\mathscr{A}_1$ on $\mathcal{P}X$ such that for each $f \in C(X)$, $\mathcal{L}f$ is measurable.*

2. *The smallest σ-algebra $\mathscr{A}_2$ on $\mathcal{P}X$ such that for each $E \in \mathscr{A}_X$, $\mathcal{L}\chi_E$ is measurable.*

3. *The smallest σ-algebra $\mathscr{A}_3$ on $\mathcal{P}X$ such that for each $f \in B(X)$, $\mathcal{L}f$ is measurable.*

Proof. First, note that Lemma 2.1.12 gives that $\mathscr{A}_1 = \mathscr{A}_2 = \mathscr{A}_3$, as any σ-algebra $\mathscr{A}$ for which $\mathcal{L}f$ is measurable for each $f \in C(X)$ automatically implies that $\mathcal{L}\chi_E$ is measurable for each measurable E, and $\mathcal{L}f$ is measurable for each $f \in B(X)$. As such $\mathscr{A}_2, \mathscr{A}_3 \subset \mathscr{A}_1$. The other containments are all proven in a similar manner, all that remains is to prove that $\mathscr{A}_{\mathcal{P}X}$ is also equal to these σ-algebras.

Next, since the topology on $\mathcal{P}X$ is the weakest which satisfies $\mathcal{L}C(X) \subset C(\mathcal{P}X)$, and $C(\mathcal{P}X) \subset B(\mathcal{P}X)$, we have that for each $f \in C(X)$, $\mathcal{L}f$ is measurable. This implies that $\mathscr{A}_1 \subset \mathscr{A}_{\mathcal{P}X}$, and so it only remains to show the reverse inclusion.

Since X is compact, $C(X)$ is separable, so let $D \subset C(X)$ be countable and dense. By Lemma 2.1.15, $\tau_{\mathcal{P}X}$ is the weakest topology such that for each $f \in D$, $\mathcal{L}f \in C(\mathcal{P}X)$. Furthermore, since $\mathbb{R}$ is second countable, it has a countable base $\{U_n\}_{n\in\mathbb{N}}$ of open subsets of $\mathbb{R}$, and $\tau_{\mathcal{P}X}$ is the smallest topology containing all sets of the form $(\mathcal{L}f)^{-1}(U_n)$ for $f \in D$ and $n \in \mathbb{N}$. Letting

$$\mathscr{B} = \left\{ \bigcap_{i=1}^{k} (\mathcal{L}f_i)^{-1}(U_{n_i}) : f_1, \ldots, f_k \in D, n_1, \ldots, n_k \in \mathbb{N} \right\},$$

this gives that $\mathscr{B}$ is a base for $\tau_{\mathcal{P}X}$. Since each U_n is measurable in $\mathbb{R}$, for any $f \in D \subset C(X)$ we have $(\mathcal{L}f)^{-1}(U_n)$ is in $\mathscr{A}_1$ by definition, and therefore $\mathscr{B} \subset \mathscr{A}_1$. Additionally, $\mathscr{B}$ is clearly countable, and so any $U \in \tau_{\mathcal{P}X}$ is the countable union of sets in $\mathscr{B}$, which implies that $\tau_{\mathcal{P}X} \subset \mathscr{A}_1$. Therefore $\mathscr{A}_{\mathcal{P}X} \subset \mathscr{A}_1$, which completes the proof. $\square$

2.2 Transformations between Spaces

In this section, we define sets of functions between spaces and endow these sets with a topology which is novel to the knowledge of the author. We develop some properties of this topology, and prove that while it is similar to existing topologies, it is not equal. This topology appears naturally in the study of dynamical systems, as we demonstrate in the next section. We begin by defining these sets of transformations.

Definition 2.2.1. Let X and Y be a spaces. Let Y^X denote the set of functions from X to Y. Let $\mathcal{B}(X,Y) \subset Y^X$ be the set of $\mathscr{A}_X/\mathscr{A}_Y$-measurable functions from X to Y, and let $\mathcal{C}(X,Y) \subset \mathcal{B}(X,Y)$ be the set of τ_X/τ_Y-continuous functions from X to Y. For simplicity, we use the notation $\mathcal{B}(X) = \mathcal{B}(X,X)$ and $\mathcal{C}(X) = \mathcal{C}(X,X)$ when pertinent.

Next, for any transformation $T \in \mathcal{B}(X,Y)$, we may define the usual pushforward map T_* which maps from $\mathcal{P}X$ to $\mathcal{P}Y$, however we use the notation $\mathcal{P}T$ for this map instead. We discuss the reason for this following the definition.

Definition 2.2.2. Let X and Y be spaces. For $T \in \mathcal{B}(X,Y)$, let $\mathcal{P}T : \mathcal{P}X \to \mathcal{P}Y$ be defined by

$$\mathcal{P}T[\mu](E) = \mu\big(T^{-1}(E)\big),$$

for $\mu \in \mathcal{P}X$ and $E \in \mathscr{A}_Y$.

We remark here that the double usage of $\mathcal{P}$ is deliberate, as a nod to the functorial nature of $\mathcal{P}$. The class of compact metrizable topological spaces along with either continuous or measurable transformations as morphisms between these spaces forms a category for which $\mathcal{P}$ is a functor. Given spaces X and Y, and a transformation $T : X \to Y$ (again, continuous or measurable), we have that $\mathcal{P}X$ and $\mathcal{P}Y$ are spaces, and that $\mathcal{P}T$ is a (continuous or measurable, as is shown later) transformation between $\mathcal{P}X$ and $\mathcal{P}Y$. In fact, $\mathcal{P}$ is not only a categorical functor, but a monad, when accompanied by the maps δ and β, or more precisely the natural transformations these maps induce on every space X. See [3] for more about monads and probabilistic monads in particular, including examples such as the Giry and Kantorovich Monads (in Section 4.2 of the same paper). In fact, the monad presented here sits somewhere between the Giry and Kantorovich Monads, which lie on the categories of measurable spaces and complete metric spaces respectively. To the knowledge of the author, the study of the probabilistic monad on the particular category of compact metrizable spaces is new, however there is nothing particularly surprising about the nature of this monad. In any case, we continue to show relevant properties of $\mathcal{P}T$, including many that are outside of this categorical scope.

2.2.1 Properties of $\mathcal{P}$

As we have now defined a map $\mathcal{P}$, we prove core properties of this map, just as we have done so for $\mathcal{L}$ in the previous section. The first is that $\mathcal{P}T$ is affine for every $T \in \mathcal{B}(X, Y)$.

Lemma 2.2.3. *Let X and Y be spaces, and let $T \in \mathcal{B}(X, Y)$. Then $\mathcal{P}T$ is affine.*

Proof. Let $\mu, \nu \in \mathcal{P}X$ and $t \in [0, 1]$ so that $t\mu + (1 - t)\nu \in \mathcal{P}X$. Then for $E \in \mathscr{A}_Y$, we have

$$\begin{aligned}
\mathcal{P}T[t\mu + (1-t)\nu](E) &= [t\mu + (1-t)\nu](T^{-1}E) \\
&= t\mu(T^{-1}E) + (1-t)\nu(T^{-1}E) \\
&= t\mathcal{P}T[\mu](E) + (1-t)\mathcal{P}T[\nu](E) \\
&= [t\mathcal{P}T(\mu) + (1-t)\mathcal{P}T(\nu)](E),
\end{aligned}$$

and so $\mathcal{P}T(t\mu + (1-t)\nu) = t\mathcal{P}T(\mu) + (1-t)\mathcal{P}T(\nu)$. □

Next, we have that $\mathcal{P}$ is a sort of homomorphism of composition.

Lemma 2.2.4. *Let X, Y, and Z be spaces. For $T \in \mathcal{B}(Y, Z)$ and $S \in \mathcal{B}(X, Y)$, we have that $T \circ S \in \mathcal{B}(X, Z)$, and*

$$\mathcal{P}(T \circ S) = \mathcal{P}T \circ \mathcal{P}S.$$

Proof. For $\mu \in \mathcal{P}X$ and $E \in \mathscr{A}_Z$, we have

$$\begin{aligned}\mathcal{P}[T \circ S][\mu](E) &= \mu\big((T \circ S)^{-1}(E)\big) = \mu\big(S^{-1}(T^{-1}(E))\big) \\ &= \mathcal{P}S[\mu]\big(T^{-1}(E)\big) = \mathcal{P}T[\mathcal{P}S(\mu)](E) \\ &= [\mathcal{P}T \circ \mathcal{P}S][\mu](E),\end{aligned}$$

and thus $\mathcal{P}(T \circ S) = \mathcal{P}T \circ \mathcal{P}S$. □

Now, we have that $\mathcal{P}$ preserves bijectivity.

Lemma 2.2.5. *Let X and Y be spaces, and let $T \in \mathcal{B}(X, Y)$. If T is bijective, then so is $\mathcal{P}T$, and $(\mathcal{P}T)^{-1} = \mathcal{P}T^{-1}$.*

Proof. If T is a bijection, then since X and Y are standard Borel spaces and T is Borel-measurable, we have by Corollary 15.2 of [32] that T^{-1} is also Borel measurable, and therefore $T^{-1} \in \mathcal{B}(Y, X)$. For a space Z, let I_Z denote the identity on Z. Also, note that for $\mu \in \mathcal{P}Z$ and $E \in \mathscr{A}_Z$, we have

$$\mathcal{P}I_Z[\mu](E) = \mu(I_Z^{-1}(E)) = \mu(E) = I_{\mathcal{P}Z}[\mu](E),$$

and so $\mathcal{P}I_Z = I_{\mathcal{P}Z}$. Then for $\mu \in \mathcal{P}Y$, and $E \in \mathscr{A}_Y$, we have by Lemma 2.2.4 that

$$\mathcal{P}T \circ \mathcal{P}T^{-1} = \mathcal{P}[T \circ T^{-1}] = \mathcal{P}I_Y = I_{\mathcal{P}Y},$$

and for $\mu \in \mathcal{P}X$ and $E \in \mathscr{A}_X$, we have by Lemma 2.2.4 that

$$\mathcal{P}T^{-1} \circ \mathcal{P}T = \mathcal{P}[T^{-1} \circ T] = \mathcal{P}I_X = I_{\mathcal{P}X}.$$

As such, $(\mathcal{P}T)^{-1} = \mathcal{P}T^{-1}$, which shows that $\mathcal{P}T$ is invertible, and so a bijection. $\square$

The next property of $\mathcal{P}$ is its interoperability with $\mathcal{L}$, and the following result is essentially a translation of the change of variables formula in terms of $\mathcal{L}$ and $\mathcal{P}$.

Lemma 2.2.6. *Let X and Y be spaces. For $T \in \mathcal{B}(X, Y)$ and $f \in B(Y)$, we have that $f \circ T \in B(X)$, and*

$$\mathcal{L}(f \circ T) = \mathcal{L}f \circ \mathcal{P}T.$$

Proof. For $\mu \in \mathcal{P}X$, we have by a change of variables (Theorem 16.3 [8]) that

$$[\mathcal{L}f \circ \mathcal{P}T](\mu) = \mathcal{L}f(\mathcal{P}T(\mu)) = \int_Y f \, \mathrm{d}\mathcal{P}T(\mu) = \int_X f \circ T \, \mathrm{d}\mu = \mathcal{L}[f \circ T](\mu),$$

and thus $\mathcal{L}(f \circ T) = \mathcal{L}f \circ \mathcal{P}T$. $\square$

Finally, it turns out that $\mathcal{P}$ preserves the measurability and continuity of maps, which means $\mathcal{P}$ is in some sense an operator.

Lemma 2.2.7. *Let X and Y be spaces. For $T \in \mathcal{B}(X, Y)$, we have $\mathcal{P}T \in \mathcal{B}(\mathcal{P}X, \mathcal{P}Y)$, and for $T \in \mathcal{C}(X, Y)$, we have $\mathcal{P}T \in \mathcal{C}(\mathcal{P}X, \mathcal{P}Y)$.*

Proof. First let $T \in \mathcal{B}(X, Y)$. By Theorem 2.1.16, $\mathscr{A}_{\mathcal{P}Y}$ is the weakest σ-algebra making $\mathcal{L}\chi_E$ measurable for every $E \in \mathscr{A}_Y$, and so $\mathcal{P}T \in \mathcal{B}(\mathcal{P}X, \mathcal{P}Y)$ if and only if for every $E \in \mathscr{A}_Y$, the map from $\mathcal{P}X$ to $\mathbb{R}$ given by $(\mathcal{L}\chi_E) \circ \mathcal{P}T$ is measurable. Also, for $E \in \mathscr{A}_Y$, let $\theta_E^Y = \mathcal{L}\chi_E$ for conciseness, and note that

$$\theta_E^Y(\mu) = \mathcal{L}\chi_E(\mu) = \int_X \chi_E \, \mathrm{d}\mu = \mu(E)$$

for every $\mu \in \mathcal{P}Y$. Also, for $E \in \mathscr{A}_X$, let $\theta_E^X = \mathcal{L}\chi_E$, and similarly for $\mu \in \mathcal{P}X$, $\theta_E^X(\mu) = \mu(E)$.

Now, let $E \in \mathscr{A}_Y$, and note that

$$\left[\theta_E^Y \circ \mathcal{P}T\right](\mu) = \mathcal{P}T[\mu](E) = \mu\big(T^{-1}(E)\big) = \theta_{T^{-1}(E)}^X(\mu),$$

and thus $\theta_E^Y \circ \mathcal{P}T = \theta_{T^{-1}(E)}^X$. Since $T \in \mathcal{B}(X, Y)$, we have that $T^{-1}(E) \in \mathscr{A}_X$, so by Theorem 2.1.16 we have that $\theta_{T^{-1}(E)}^X$ is measurable, and therefore so is $\theta_E^Y \circ \mathcal{P}T$. As this holds for every $E \in \mathscr{A}_Y$, it must be that $\mathcal{P}T \in \mathcal{B}(\mathcal{P}X, \mathcal{P}Y)$.

Next, let $T \in \mathcal{C}(X, Y)$. By definition, $\tau_{\mathcal{P}Y}$ is the weakest topology making $\mathcal{L}f$ continuous for each $f \in C(Y)$, and therefore $\mathcal{P}T \in \mathcal{C}(\mathcal{P}X, \mathcal{P}Y)$ if and only if for every $f \in C(Y)$ we have $\mathcal{L}f \circ \mathcal{P}T$ is continuous. Indeed, we have that $\mathcal{L}f \circ \mathcal{P}T = \mathcal{L}(f \circ T)$ by Lemma 2.2.6. Since T is continuous from X to Y, and f is continuous from Y to $\mathbb{R}$, $f \circ T \in C(X)$, and therefore we have that $\mathcal{L}(f \circ T)$ is continuous by definition of $\tau_{\mathcal{P}X}$. Thus, $\mathcal{L}f \circ \mathcal{P}T$ is continuous for each $f \in C(Y)$, which gives that $\mathcal{P}T \in \mathcal{C}(\mathcal{P}X, \mathcal{P}Y)$. □

As a result of this lemma, we have for $T \in \mathcal{B}(X, Y)$ that $\mathcal{P}T \in \mathcal{B}(\mathcal{P}X, \mathcal{P}Y)$, and thus we may take $\mathcal{P}^2T = \mathcal{P}(\mathcal{P}T) \in \mathcal{B}(\mathcal{P}^2X, \mathcal{P}^2Y)$ and so on, to obtain that for every $n \in \mathbb{N}$ we have $\mathcal{P}^nT \in \mathcal{B}(\mathcal{P}^nX, \mathcal{P}^nY)$. Similarly, for $T \in \mathcal{C}(X, Y)$, we have for every $n \in \mathbb{N}$ that $\mathcal{P}^nT \in \mathcal{C}(\mathcal{P}^nX, \mathcal{P}^nY)$. With this, we demonstrate the connection between $\mathcal{P}T$ and the barycenter map β defined in the previous section.

Lemma 2.2.8. *Let X and Y be spaces, and let $T \in \mathcal{B}(X, Y)$. Then with the barycenter maps $\beta_X : \mathcal{P}^2X \to \mathcal{P}X$ and $\beta_Y : \mathcal{P}^2Y \to \mathcal{P}Y$, we have*

$$\mathcal{P}T \circ \beta_Y = \beta_X \circ \mathcal{P}^2T,$$

where $\mathcal{P}^2T = \mathcal{P}(\mathcal{P}T)$ is a map from $\mathcal{P}^2X$ to $\mathcal{P}^2Y$.

Proof. Let $m \in \mathcal{P}^2 X$ and $E \in \mathscr{A}_Y$. Then by a change of variables, we have

$$\begin{aligned}
[\mathcal{P}T \circ \beta_Y][m](E) &= \beta_Y[m](T^{-1}(E)) \\
&= \int_{\mathcal{P}X} \mu(T^{-1}(E))\, m(\mathrm{d}\mu) \\
&= \int_{\mathcal{P}X} \mathcal{P}T[\mu](E)\, m(\mathrm{d}\mu) \\
&= \int_{\mathcal{P}Y} \nu(E)\, \mathcal{P}^2 T(m)(\mathrm{d}\nu) \\
&= [\beta_Y \circ \mathcal{P}^2 T][m](E).
\end{aligned}$$

As this holds for all $m \in \mathcal{P}^2 X$ and $E \in \mathscr{A}_Y$, we have shown the desired result. □

Lastly, we have the following Proposition which largely follows from many of the lemmas in this section, but since we use it extensively in Section 2.5, it is useful to have it stated as its own result.

Proposition 2.2.9. *Let X and Y be spaces, and let $T \in \mathcal{C}(X, Y)$ be a homeomorphism. Then $\mathcal{P}T \in \mathcal{C}(\mathcal{P}X, \mathcal{P}Y)$ is an affine homeomorphism.*

Proof. By Lemmas 2.2.3, 2.2.5, and 2.2.7, we have that $\mathcal{P}T$ is an affine continuous bijection between $\mathcal{P}X$ and $\mathcal{P}Y$. Since $\mathcal{P}X$ is compact, and $\mathcal{P}Y$ is Hausdorff, we have that $\mathcal{P}T$ is an affine homeomorphism. □

2.2.2 A topology on $\mathcal{B}(X, Y)$

We now move to endow $\mathcal{B}(X, Y)$ (and as a result $\mathcal{C}(X, Y)$) with a topology that is in some sense compatible with both $\mathcal{P}$ and the topology on $\mathcal{P}Y$. It is defined as follows.

Definition 2.2.10. Let X and Y be spaces. We endow $\mathcal{B}(X, Y)$ with the weakest topology so that for each $\mu \in \mathcal{P}X$, the function $\rho_\mu : \mathcal{B}(X, Y) \to \mathcal{P}Y$ defined by

$$\rho_\mu(T) = \mathcal{P}T(\mu)$$

is continuous. $\mathcal{C}(X, Y) \subset \mathcal{B}(X, Y)$ is endowed with the corresponding subspace topology.

In order to understand the behavior of this topology, we prove some results about the continuity of composition with respect to this topology. When both T and S are only measurable, we have the following result about the continuity of $T \circ S$.

Lemma 2.2.11. *Let X, Y, and Z be spaces. Then the composition operator $\circ$: $\mathcal{B}(Y, Z) \times \mathcal{B}(X, Y) \to \mathcal{B}(X, Z)$ is continuous in its left component.*

Proof. To show that $\circ$ is continuous in its left component, we need to show that for each $S \in \mathcal{B}(X, Y)$ the map $\kappa_S : \mathcal{B}(Y, Z) \to \mathcal{B}(X, Z)$ defined by $\kappa_S(T) = T \circ S$ is continuous. Since the topology on $\mathcal{B}(X, Z)$ is the weakest such that for every $\mu \in \mathcal{P}X$, the function ρ_μ is continuous, it suffices to show that $\rho_\mu \circ \kappa_S$ is continuous for each $\mu \in \mathcal{P}X$ in order to show that κ_S is continuous.

Indeed, let $\mu \in \mathcal{P}X$ and we have by Lemma 2.2.4 that

$$\begin{aligned}[]
[\rho_\mu \circ \kappa_S](T) &= \rho_\mu(T \circ S) = \mathcal{P}[T \circ S](\mu) = [\mathcal{P}T \circ \mathcal{P}S](\mu) \\
&= \mathcal{P}T\big(\mathcal{P}S(\mu)\big) = \rho_{\mathcal{P}S(\mu)}(T),
\end{aligned}$$

thus $\rho_\mu \circ \kappa_S = \rho_{\mathcal{P}S(\mu)}$. As $\mathcal{P}S(\mu) \in \mathcal{P}Y$, $\rho_{\mathcal{P}S(\mu)}$ is continuous by definition of the topology on $\mathcal{B}(Y, Z)$, so $\rho_\mu \circ \kappa_S$ is continuous, completing the proof. □

For continuity in the right component, we need some additional assumption on the maps. In particular, if $T \circ S$ is to be continuous in S, it must be that $\mathcal{P}T$ is continuous, which happens precisely when T is, giving the following result.

Lemma 2.2.12. *Let X, Y, and Z be spaces. Then the composition operator $\circ$: $\mathcal{C}(Y, Z) \times \mathcal{B}(X, Y) \to \mathcal{B}(X, Z)$ is separately continuous.*

Proof. By Lemma 2.2.11 $\circ$ is continuous in its left component, so it suffices to show that $\circ$ is continuous in its right component in order to prove that it is separately continuous. To do so, we need to show that for each $T \in \mathcal{C}(Y, Z)$ the map $\lambda_T : \mathcal{B}(X, Y) \to \mathcal{B}(X, Z)$ defined by $\lambda_T(S) = T \circ S$ is continuous. Since the topology on $\mathcal{B}(X, Z)$ is the weakest such that for every $\mu \in \mathcal{P}X$, the function ρ_μ is continuous, it suffices to show that $\rho_\mu \circ \lambda_T$ is continuous for each $\mu \in \mathcal{P}X$ in order to show that λ_T is continuous.

Indeed, let $\mu \in \mathcal{P}X$ and we have by Lemma 2.2.4 that

$$\begin{aligned}[][\rho_\mu \circ \lambda_T](S) &= \rho_\mu(T \circ S) = \mathcal{P}[T \circ S](\mu) = [\mathcal{P}T \circ \mathcal{P}S](\mu) \\ &= \mathcal{P}T\big(\mathcal{P}S(\mu)\big) = \mathcal{P}T\big(\rho_\mu(S)\big) = [\mathcal{P}T \circ \rho_\mu](S),\end{aligned}$$

and therefore $\rho_\mu \circ \lambda_T = \mathcal{P}T \circ \rho_\mu$. By Lemma 2.2.7, since $T \in \mathcal{C}(Y, Z)$, we have that $\mathcal{P}T \in \mathcal{C}(\mathcal{P}Y, \mathcal{P}Z)$, so $\mathcal{P}T$ is a continuous map. By definition of the topology on $\mathcal{B}(X, Z)$, we also have that ρ_μ is continuous, and therefore $\mathcal{P}T \circ \rho_\mu$ is continuous, which gives that $\rho_\mu \circ \lambda_T$ is continuous, completing the proof. □

Beyond the continuity of composition, we move to comparing this topologies to other topologies on $\mathcal{B}(X, Y)$. Immediately, we have two usual notions of convergence for functions in $\mathcal{B}(X, Y)$; pointwise convergence, which is represented by the subspace topology on $\mathcal{B}(X, Y)$ inherited from the product topology on Y^X, and the uniform convergence topology on $\mathcal{B}(X, Y)$. The latter is metrizable by choosing a bounded metric d on Y, and defining a metric m on $\mathcal{B}(X, Y)$ by $m(T, S) = \sup_{x \in X} d(T(x), S(x))$. First, we have that the topology on $\mathcal{B}(X, Y)$ is at least as fine as the product topology.

Lemma 2.2.13. *Let X and Y be spaces. Then the topology on $\mathcal{B}(X, Y)$ is finer than the subspace topology inherited from Y^X with the product topology.*

Proof. The product topology on Y^X is generated by sets of the form

$$S(x, U) = \{T \in Y^X : T(x) \in U\}$$

for every $x \in X$ and $U \subset Y$ open, so we need only show that each $S(x, U)$ contains a basic open subset of $\mathcal{B}(X, Y)$.

For $x \in X$ and $U \subset Y$ open, we have that $\delta_x \in \mathcal{P}X$. Furthermore, since δ embeds Y as a closed subspace of $\mathcal{P}Y$, we have that the topology on Y is equal to the topology on $\delta(Y)$ as a subspace of $\mathcal{P}Y$ when Y and δY are identified. As such, for any open $U \subset Y$, there exists an open set $V \subset \mathcal{P}Y$ such that $\delta(U) = V \cap \delta(Y)$.

Now, consider the basic open subset $\rho_{\delta_x}^{-1}(V)$ of $\mathcal{B}(X, Y)$. First, for any $T \in \rho_{\delta_x}^{-1}(V)$, we have that $\mathcal{P}T(\delta_x) \in V$, but $\mathcal{P}T(\delta_x) = \delta_{T(x)}$, and so in fact $\mathcal{P}T(\delta_x) \in V \cap \delta(Y) = \delta(U)$. As such, we have that $\delta_{T(x)} \in \delta(U)$, which directly implies that $T(x) \in U$, or alternatively, that $T \in S(x, U)$. As this holds for every $T \in \rho_{\delta_x}^{-1}(V)$, we have that $\rho_{\delta_x}^{-1}(V) \subset S(x, U)$, which completes the proof. □

Next, we have that the topology on $\mathcal{B}(X, Y)$ is no finer that the uniform convergence topology.

Lemma 2.2.14. *Let X and Y be spaces. Then the topology on $\mathcal{B}(X, Y)$ is coarser than the uniform convergence topology on $\mathcal{B}(X, Y)$.*

Proof. Let $I : \mathcal{B}(X, Y) \to \mathcal{B}(X, Y)$ be the identity map, where the domain is endowed with the uniform convergence topology, and the co-domain is endowed with the topology on $\mathcal{B}(X, Y)$. Showing that I is continuous will give the desired result. Since the uniform convergence topology is metrizable, we have that I is continuous if and only if it is sequentially continuous [7], so let $\{T_n\}_{n \in \mathbb{N}} \subset \mathcal{B}(X, Y)$ be a sequence which converges uniformly to T. Now for $f \in C(Y)$, since Y is compact f is uniformly continuous, and therefore $\{f \circ T_n\}_{n \in \mathbb{N}}$ converges uniformly to $f \circ T$, which directly implies that this convergence is in $B(X)$. By Lemma 2.1.13, we have that $\mathcal{L}(f \circ T_n)$ converges to $\mathcal{L}(f \circ T)$ (in $B(\mathcal{P}X)$, and therefore also pointwise), and by Lemma 2.2.6, we have that $\mathcal{L}(f \circ T_n) = \mathcal{L}f \circ \mathcal{P}T_n$ and $\mathcal{L}(f \circ T) = \mathcal{L}f \circ \mathcal{P}T$, and thus for every

$\mu \in \mathcal{P}X$ we have

$$\lim_{n\to\infty} \mathcal{L}f(\rho_\mu(T_n)) = \lim_{n\to\infty} \mathcal{L}f(\mathcal{P}T_n(\mu)) = \lim_{n\to\infty} \mathcal{L}[f \circ T_n](\mu)$$
$$= \mathcal{L}[f \circ T](\mu) = \mathcal{L}f(\mathcal{P}T(\mu)) = \mathcal{L}f(\rho_\mu(T)).$$

As this holds for every $f \in C(Y)$, we have that $\{\rho_\mu(T_n)\}_{n\in\mathbb{N}}$ is a sequence which converges to $\rho_\mu(T)$ in $\mathcal{P}Y$. As $\mu \in \mathcal{P}X$ was arbitrary, it follows that $\{T_n\}_{n\in\mathbb{N}}$ converges to T in the topology on $\mathcal{B}(X, Y)$ by definition. Therefore, I is sequentially continuous, which is the desired result. □

As such, the previous two lemmas gives us that this topology sits somewhere in between the pointwise convergence topology and the uniform convergence topology. The following result gives us that this topology is far closer to the former.

Lemma 2.2.15. *Let X and Y be spaces. If $\{T_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathcal{B}(X, Y)$ that converges pointwise to a function $T \in \mathcal{B}(X, Y)$, then $\{T_n\}_{n\in\mathbb{N}}$ converges to T in $\mathcal{B}(X, Y)$.*

Proof. Let $f \in C(Y)$. Since f is continuous, we have that

$$\lim_{n\to\infty} f(T_n(x)) = f\left(\lim_{n\to\infty} T_n(x)\right) = f(T(x))$$

for every $x \in X$, and thus $f \circ T_n$ converges pointwise to $f \circ T$. Then, for any $\mu \in \mathcal{P}X$, we have by the Dominated Convergence Theorem that

$$\lim_{n\to\infty} \int_X f \circ T_n \, \mathrm{d}\mu = \int_X f \circ T \, \mathrm{d}\mu.$$

By definition, we have $\int_X f \circ T_n \, \mathrm{d}\mu = \mathcal{L}[f \circ T_n](\mu)$, and by Lemma 2.2.6, we further have that $\mathcal{L}[f \circ T_n](\mu) = [\mathcal{L}f \circ \mathcal{P}T_n](\mu)$. As $\mathcal{P}T_n(\mu) = \rho_\mu(T_n)$, we have that

$$\lim_{n\to\infty} \mathcal{L}f(\rho_\mu(T_n)) = \mathcal{L}f(\rho_\mu(T)).$$

As $f \in C(Y)$ was arbitrary, this gives us that $\rho_\mu(T_n)$ converges to $\rho_\mu(T)$ in $\mathcal{P}Y$. As this holds for every $\mu \in \mathcal{P}X$, we have that T_n converges to T in $\mathcal{B}(X, Y)$. □

Additionally, since it is obviously the case that a pointwise convergent sequence does not automatically convergence uniformly, the previous lemma shows that the topology on $\mathcal{B}(X, Y)$ is strictly coarser than the uniform convergence topology. While this result would seem to indicate that the pointwise convergence topology coincides with the topology on $\mathcal{B}(X, Y)$, and this is true in some more trivial cases. However, in most interesting cases, it is not. The following result gives an exact characterization for when the topology on $\mathcal{B}(X, Y)$ is strictly finer than the pointwise convergence topology.

Theorem 2.2.16. *Let X and Y be spaces with Y having at least 2 points. Then the topology on $\mathcal{B}(X, Y)$ is strictly finer than the subspace topology inherited from Y^X with the product topology if and only if X is uncountable.*

Proof. First, we proceed by the contrapositive and assume that X is countable. As we have already shown in Lemma 2.2.13 that the topology on $\mathcal{B}(X, Y)$ is finer than the subspace topology inherited from Y^X with the product topology, it suffices to show that this subspace topology is finer than the topology on $\mathcal{B}(X, Y)$. We do this by showing that the identity map $I : \mathcal{B}(X, Y) \to \mathcal{B}(X, Y)$, where the domain is endowed with the subspace topology and co-domain is endowed with the topology on $\mathcal{B}(X, Y)$, is continuous. As X is countable, the space Y^X is therefore a countable product of metrizable spaces and is therefore also metrizable, so we have that the subspace topology on $\mathcal{B}(X, Y)$ is metrizable. As such, I is continuous if and only if it is sequentially continuous [7], and the sequential continuity of I is given by Lemma

2.2.15, therefore I is continuous and the subspace topology and topology on $\mathcal{B}(X, Y)$ coincide.

For the converse, suppose that X is uncountable. We now show that there exists a net $\{T_\alpha\} \subset \mathcal{B}(X, Y)$ and $T \in \mathcal{B}(X, Y)$ such that $T_\alpha \to T$ in the subspace topology, but $T_\alpha \not\to T$ in the topology on $\mathcal{B}(X, Y)$, which proves that the topology on $\mathcal{B}(X, Y)$ is strictly finer than the subspace topology. By assumption, there exists at least two distinct points y_0 and y_1 in Y, which we may choose arbitrarily. By ordering $\mathscr{F}(X)$, the set of finite subsets of X, by inclusion, $\mathscr{F}(X)$ is a directed set. For each $\alpha \in \mathscr{F}(X)$, define $T_\alpha \in \mathcal{B}(X, Y)$ by

$$T_\alpha(x) = \begin{cases} y_1 & x \in \alpha \\ y_0 & x \in X \setminus \alpha \end{cases}.$$

Note that each T_α is clearly measurable, since the preimage of any Borel set is either $\varnothing$, α, $X \setminus \alpha$, or X, with α being a finite, and therefore Borel, set. Define $T \in \mathcal{B}(X, Y)$ by $T(x) = y_1$. Now, the product topology on Y^X is generated by open sets of the form

$$\prod_{x \in X} U_x,$$

where each U_x is an open set in Y, and $U_x = Y$ for all but finitely many $x \in X$. For any open set U of this form, let $\operatorname{supp}(U) = \{x \in X : U_x \neq Y\}$, and note that $\operatorname{supp}(U)$ is necessarily finite, so $\operatorname{supp}(U) \in \mathscr{F}(X)$.

Now, let U be any basic open set (a set of the form displayed above) in Y^X containing T. Let $\alpha = \operatorname{supp}(U)$, and note that for each $x \in \alpha$, we have that $T(x) = y_1 \in U_x$. Then, for any $\alpha \subset \beta \in \mathscr{F}(X)$, we have that for every $x \in \alpha \subset \beta$ that $T_\beta(x) = y_1 \in U_x$, and therefore $T_\beta \in U$. Since U was an arbitrary basic open set, we have shown that $T_\alpha \to T$ with respect to the subspace topology on $\mathcal{B}(X, Y)$.

Finally, since X is an uncountable compact metric space, there exists some nonatomic measure $\mu \in \mathcal{P}X$ (Chapter II, Theorem 8.1, [40]), so we have for every $E \in \mathscr{F}(X)$ that $\mu(E) = 0$ by additivity. Additionally, since Y is metric it is completely Hausdorff, and so there exists a continuous function $f : Y \to \mathbb{R}$ such that $f(y_0) = 0$ and $f(y_1) = 1$. Then for every $\alpha \in \mathscr{F}(X)$, we have

$$\begin{aligned}
\mathcal{L}f(\rho_\mu(T_\alpha)) &= \int_X (f \circ T_\alpha)(x)\, \mu(\mathrm{d}x) \\
&= \int_\alpha f(T_\alpha(x))\, \mu(\mathrm{d}x) + \int_{X \setminus \alpha} f(T_\alpha(x))\, \mu(\mathrm{d}x) \\
&= \int_\alpha f(y_1)\, \mathrm{d}\mu + \int_{X \setminus \alpha} f(y_0)\, \mathrm{d}\mu \\
&= 1\mu(\alpha) + 0\mu(X \setminus \alpha) \\
&= 0.
\end{aligned}$$

On the other hand, we have

$$\mathcal{L}f(\rho_\mu(T)) = \int_X (f \circ T)\, \mathrm{d}\mu = \int_X f(y_1)\, \mathrm{d}\mu = 1,$$

and therefore we do not have that $\mathcal{L}f(\rho_\mu(T_\alpha)) \to \mathcal{L}f(\rho_\mu(T))$ in $\mathbb{R}$, which implies that we do not have $\rho_\mu(T_\alpha) \to \rho_\mu(T)$ in $\mathcal{P}Y$. Therefore, the net $\{T_\alpha\}_{\alpha \in \mathscr{F}(X)}$ does not converge to T in $\mathcal{B}(X, Y)$, so this topology must be strictly finer than the subspace topology on $\mathcal{B}(X, Y)$ from the product topology on Y^X. $\square$

As a result, for most interesting choices of the space X and Y, it will be the case that this topology on $\mathcal{B}(X, Y)$ is stronger than pointwise convergence, but weaker than uniform convergence. Despite this, when observing sequences only, the topology on $\mathcal{B}(X, Y)$ is essentially pointwise convergence. This topology emerges naturally in the study of dynamical systems (defined in the next section), so it is important to understand its structure, which we have done by placing it between two well known topologies.

2.3 Dynamical Systems and their Completions

We now turn our attention to defining the notion of a dynamical system which we are primarily interested in studying. These systems are fairly general, and a large majority of systems classically studied are dynamical systems in this sense. The purpose of defining dynamical systems so generally will become clear throughout this section.

Definition 2.3.1. A *dynamical system* is a pair $(X, \mathcal{T})$ of a space X and a (possibly empty) set of transformations $\mathcal{T} \subset \mathcal{B}(X)$. When $\mathcal{T}$ consists of a single transformation T, we use the shorthand (X, T) to denote a dynamical system. A *measure system* is a pair $(X, \mathcal{M})$ of a space X and a (possibly empty) set of probability measures $\mathcal{M} \subset \mathcal{P}X$. When $\mathcal{M}$ consists of a single measure μ, we use the shorthand (X, μ).

Measure systems are not a common notion in the realm of ergodic theory, but the following definition gives us a way of transferring between these two types of systems.

Definition 2.3.2. Let X be a space. For a measure $\mu \in \mathcal{P}X$ and a transformation $T \in \mathcal{B}(X)$, we say that μ is *T-invariant*, or that T is *a μ-preserving function* if $\mathcal{P}T(\mu) = \mu$. Furthermore, let

$$\mathcal{I}_X(T) = \{\mu \in \mathcal{P}X : \mathcal{P}T(\mu) = \mu\} \subset \mathcal{P}X$$

be the set of all T-invariant measures on X, which may be nonempty.

$$\mathcal{F}_X(\mu) = \{T \in \mathcal{B}(X) : \mathcal{P}T(\mu) = \mu\} \subset \mathcal{B}(X)$$

the set of all μ-preserving functions on X, which is always nonempty, as the identity map I_X on X is μ-preserving for all measures $\mu \in \mathcal{P}X$. If $\mathcal{T} \subset \mathcal{B}(X)$ is a collection of transformations, we say that a measure $\mu \in \mathcal{P}X$ is *$\mathcal{T}$-invariant* if for every $T \in \mathcal{T}$,

μ is T-invariant, and we define

$$\mathcal{I}_X(\mathcal{T}) = \bigcap_{T \in \mathcal{T}} \mathcal{I}_X(T)$$

to be the set of $\mathcal{T}$-invariant measures, with the convention that $\mathcal{I}_X(\varnothing) = \mathcal{P}X$. Also, if $\mathcal{M} \subset \mathcal{P}X$ is a collection of measures, we say that a transformation T is $\mathcal{M}$-*preserving* if for every $\mu \in \mathcal{M}$, T is μ-preserving, and we define

$$\mathcal{F}_X(\mathcal{M}) = \bigcap_{\mu \in \mathcal{M}} \mathcal{F}_X(\mu)$$

to be the set of $\mathcal{M}$-preserving functions, with the convention that $\mathcal{F}_X(\varnothing) = \mathcal{B}(X)$. When the space X is clear from the context, we write $\mathcal{I}(\mathcal{T}) = \mathcal{I}_X(\mathcal{T})$ or $\mathcal{F}(\mathcal{M}) = \mathcal{F}_X(\mathcal{M})$.

From these definitions, we can make the following associations.

Definition 2.3.3. For a dynamical system $(X, \mathcal{T})$, we have that $\mathcal{I}_X(\mathcal{T})$ is a collection of measures on X, and thus $(X, \mathcal{I}_X(\mathcal{T}))$ is a measure system, referred to as the *associated measure system to* $(X, \mathcal{T})$. Similarly, for a measure system $(X, \mathcal{M})$, we have that $\mathcal{F}_X(\mathcal{M})$ is a collection of transformations on X, and thus $(X, \mathcal{F}_X(\mathcal{M}))$ is a dynamical system, referred to as the *associated dynamical system to* $(X, \mathcal{M})$.

As such, dynamical systems and measure systems are in some sense dual notions. Classically, $\mathcal{I}_X(\mathcal{T})$ is denoted $\mathcal{P}_{\mathcal{T}}(X)$, however for the purposes of this chapter, having $\mathcal{T}$ in a subscript would become cumbersome and may conflict with other notation, so we use this notation instead. A principal goal in Ergodic theory is to be able to take an arbitrary dynamical system and produce a description of its associated measure system. Tools for doing so generally involve using the dynamical system to identify structures within the associated measure system directly to produce a description of the invariant measures. The goal of the rest of this section is to develop new tools in the pursuit of characterizing the associated measure system of a dynamical system.

2.3.1 Properties of $\mathcal{I}_X(\mathcal{T})$

In general, the structure of $\mathcal{I}_X(\mathcal{T})$ for a dynamical system is rather well studied, and many of the properties we discuss here are fairly well known, however we do introduce some concepts which are new to the knowledge of the author. We indicate when this is the case, but we begin with some basic properties that are simple to prove. First, it is not necessarily the case that $\mathcal{I}_X(\mathcal{T})$ is nonempty, except in some specific scenarios.

Theorem 2.3.4 (Krylov-Bogolyubov)**.** *If $(X, \mathcal{T})$ is a dynamical system such that $\mathcal{T} \subset \mathcal{C}(X)$, and either $\mathcal{T} = \{T\}$ or $\mathcal{T}$ is a countable amenable group under composition, then $\mathcal{I}(\mathcal{T})$ is nonempty.*

Next, we have that for a subcollection of transformations, the set of invariant measures with respect to this subset is a superset of the invariant measures of the original system.

Lemma 2.3.5. *Let $(X, \mathcal{T})$ be a dynamical system, and $\mathcal{S} \subset \mathcal{T}$. Then $\mathcal{I}(\mathcal{T}) \subset \mathcal{I}(\mathcal{S})$.*

Proof. If

$$\mu \in \mathcal{I}(\mathcal{T}) = \bigcap_{T \in \mathcal{T}} \mathcal{I}(T),$$

we have that $\mu \in \mathcal{I}(T)$ for every $T \in \mathcal{T}$. As $\mathcal{S} \subset \mathcal{T}$ by assumption, it follows that for every $S \in \mathcal{S} \subset \mathcal{T}$ that $\mu \in \mathcal{I}(S)$, and thus

$$\mu \in \bigcap_{S \in \mathcal{S}} \mathcal{I}(S) = \mathcal{I}(\mathcal{S}).$$

Since this holds for every $\mu \in \mathcal{I}(\mathcal{T})$, we obtain that $\mathcal{I}(\mathcal{T}) \subset \mathcal{I}(\mathcal{S})$. $\square$

Using this result, we have the following simple, but generally nonstandard result which we use in Section 2.5.

Lemma 2.3.6. *Let X be a space and $\mathcal{T}, \mathcal{S} \subset \mathcal{B}(X)$ be such that $\mathrm{id}_X \in \mathcal{T}$ and $\mathrm{id}_X \in \mathcal{S}$. Then with*

$$\mathcal{T} \circ \mathcal{S} = \{T \circ S : T \in \mathcal{T}, S \in \mathcal{S}\},$$

we have that

$$\mathcal{I}(\mathcal{T} \cup \mathcal{S}) = \mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S}) = \mathcal{I}(\mathcal{T} \circ \mathcal{S})$$

Proof. First, note that

$$\mathcal{I}(\mathcal{T} \cup \mathcal{S}) = \bigcap_{T \in \mathcal{T} \cup \mathcal{S}} \mathcal{I}(T) = \bigcap_{T \in \mathcal{T}} \mathcal{I}(T) \cap \bigcap_{S \in \mathcal{S}} \mathcal{I}(S) = \mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S}).$$

Now, since $\mathrm{id}_X \in \mathcal{T}$, we have that $\mathcal{S} = \mathrm{id}_X \circ \mathcal{S} \subset \mathcal{T} \circ \mathcal{S}$, and so we have that $\mathcal{I}(\mathcal{T} \circ \mathcal{S}) \subset \mathcal{I}(\mathcal{S})$ by Lemma 2.3.5. Similarly, since $\mathrm{id}_X \in \mathcal{S}$, we have that $\mathcal{T} = \mathcal{T} \circ \mathrm{id}_X \subset \mathcal{T} \circ \mathcal{S}$, and so we have $\mathcal{I}(\mathcal{T} \circ \mathcal{S}) \subset \mathcal{I}(\mathcal{T})$ by Lemma 2.3.5. As such, we have $\mathcal{I}(\mathcal{T} \circ \mathcal{S}) \subset \mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S})$.

Finally, for $\mu \in \mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S})$, we have for every $T \in \mathcal{T}$ that $\mathcal{P}T(\mu) = \mu$, and for every $S \in \mathcal{S}$ that $\mathcal{P}T(\mu) = \mu$. As such, we have for every $T \in \mathcal{T}$ and $S \in \mathcal{S}$ by Lemma 2.2.4 that

$$\mathcal{P}[T \circ S](\mu) = (\mathcal{P}T \circ \mathcal{P}S)(\mu) = \mathcal{P}T(\mathcal{P}S(\mu)) = \mathcal{P}T(\mu) = \mu,$$

and thus we have that $\mu \in \mathcal{I}(T \circ S)$. Since that holds for every $T \in \mathcal{T}$ and $S \in \mathcal{S}$, we have that $\mu \in \mathcal{I}(\mathcal{T} \circ \mathcal{S})$. Since $\mu \in \mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S})$ was arbitrary, we have shown $\mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S}) \subset \mathcal{I}(\mathcal{T} \circ \mathcal{S})$. Combined with the result of the previous paragraph gives $\mathcal{I}(\mathcal{T}) \cap \mathcal{I}(\mathcal{S}) = \mathcal{I}(\mathcal{T} \circ \mathcal{S})$ as desired. □

While it is not true in general that the set of invariant measures for a dynamical system will be closed in $\mathcal{P}X$, there is a particular class of dynamical systems for which it is.

Lemma 2.3.7. *Let $(X, \mathcal{T})$ be a dynamical system such that $\mathcal{T} \subset \mathcal{C}(X)$. Then $\mathcal{I}(\mathcal{T})$ is closed in $\mathcal{P}X$.*

Proof. Since $\mathcal{I}(\mathcal{T}) = \bigcap_{T\in\mathcal{T}} \mathscr{I}(T)$, it will suffice to prove that $\mathscr{I}(T)$ is closed for every $T \in \mathcal{C}(X)$, at which point $\mathcal{I}(\mathcal{T})$ is an intersection of closed sets and therefore closed. Furthermore, since $\mathcal{P}X$ is metrizable, it suffices to show that $\mathcal{I}(T)$ is sequentially closed. Indeed, let $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathcal{I}(T)$ be a sequence which converges to $\mu \in \mathcal{P}X$, so we have that $\mathcal{P}T(\mu_n) = \mu_n$ for every $n \in \mathbb{N}$. Then, since $T \in C(X)$, we have $\mathcal{P}T \in \mathcal{C}(\mathcal{P}X)$, and so $\{\mathcal{P}T(\mu_n)\}_{n\in\mathbb{N}}$ converges to $\mathcal{P}T(\mu)$ in $\mathcal{P}X$. But $\mathcal{P}T(\mu_n) = \mu_n$, and therefore we have that $\{\mu_n\}_{n\in\mathbb{N}}$ converges to $\mathcal{P}T(\mu)$ as well as μ in $\mathcal{P}X$. Since $\mathcal{P}X$ is Hausdorff, it must be that $\mathcal{P}T(\mu) = \mu$, and therefore $\mu \in \mathcal{I}(T)$. This shows that $\mathcal{I}(T)$ is closed, which is the desired result. □

In light of this result, we give the following definition.

Definition 2.3.8. A measure system $(X, \mathcal{M})$ is *closed* if $\mathcal{M}$ is closed (or compact) in $\mathcal{P}X$.

Interestingly, for a dynamically system $(X, \mathcal{T})$, while continuity of every element in $\mathcal{T}$ is sufficient to guarantee that $(X, \mathcal{I}(\mathcal{T}))$ is closed, it turns out that in the majority of the cases here, it suffices to just assume that $(X, \mathcal{I}(\mathcal{T}))$ is closed (or $\mathcal{I}(\mathcal{T})$ is closed) in order to derive important properties

2.3.1.1 Invariant sets and ergodicity

In order to develop more properties of $\mathcal{I}_X(\mathcal{T})$, we need to define the notion of invariant sets, and ultimately the notion of $\mathcal{T}$-ergodic measures. We begin with the following definition.

Definition 2.3.9. Let X be a space, let $\mu \in \mathcal{P}X$, and let $T \in \mathcal{B}(X)$. Then a set $E \in \mathscr{A}_X$ is (μ, T)-invariant if $\mu(E \triangle T^{-1}(E)) = 0$, where $\triangle$ denotes the symmetric difference. We define

$$\mathscr{I}_X(\mu, T) = \{E \in \mathscr{A}_X : \mu(E \triangle T^{-1}(E)) = 0\}$$

to be the set of all (μ, T)-invariant sets. Furthermore, for $\mathcal{M} \subset \mathcal{P}X$ and $\mathcal{T} \subset \mathcal{B}(X)$, a set $E \in \mathscr{A}_X$ is $(\mathcal{M}, \mathcal{T})$-invariant if for every $\mu \in \mathcal{M}$ and $T \in \mathcal{T}$, it is the case that E is (μ, T)-invariant. We define

$$\mathscr{I}_X(\mathcal{M}, \mathcal{T}) = \bigcap_{\mu \in \mathcal{M}} \bigcap_{T \in \mathcal{T}} \mathscr{I}_X(\mu, T)$$

to be the set of all $(\mathcal{M}, \mathcal{T})$-invariant sets. In the case that $\mathcal{M} = \varnothing$ or $\mathcal{T} = \varnothing$, then $\mathscr{I}_X(\mathcal{M}, \mathcal{T}) = \mathscr{A}_X$ vacuously. Additionally, when X is clear from context, we omit X in the notation, so $\mathscr{I}(\mathcal{M}, \mathcal{T}) = \mathscr{I}_X(\mathcal{M}, \mathcal{T})$. Furthermore, if $\mathcal{M} = \{\mu\}$, we will write $\mathscr{I}(\mu, \mathcal{T}) = \mathscr{I}(\mathcal{M}, \mathcal{T})$ and if $\mathcal{T} = \{T\}$, we will write $\mathscr{I}(\mathcal{M}, T) = \mathscr{I}(\mathcal{M}, \mathcal{T})$.

This notion of an invariant set does not necessarily align with more classical definitions of invariant sets, which are defined next, however as can be seen in Chapter 12 of [43], this notion of an invariant set is the correct notion to use in this context where $\mathcal{T}$ may have a rather large number of transformations. This reference however only defines $\mathscr{I}_X(\mu, \mathcal{T})$ (with different notation), and does consider $\mathscr{I}_X(\mathcal{M}, \mathcal{T})$ for finite $\mathcal{M}$, however does not make any mentions to considering when $\mathcal{M}$ is infinite. To the knowledge of the author, this notion of an invariant set with respect to multiple measures and multiple transformations has not been previously studied. Now, we define the more classic notion of an invariant set.

Definition 2.3.10. Let X be a space, and let $T \in \mathcal{B}(X)$. Then a set $E \in \mathscr{A}_X$ is T-invariant if $T^{-1}(E) = E$. We define

$$\mathscr{I}_X(T) = \{E \in \mathscr{A}_X : T^{-1}(E) = E\}$$

to be the set of all T-invariant sets. Furthermore, for $\mathcal{T} \subset \mathcal{B}(X)$, a set $E \in \mathscr{A}_X$ is $\mathcal{T}$-invariant if for every $T \in \mathcal{T}$, we have that E is T-invariant. We define

$$\mathscr{I}_X(\mathcal{T}) = \bigcap_{T \in \mathcal{T}} \mathscr{I}_X(T)$$

to be the set of all $\mathcal{T}$-invariant sets. In the case that $\mathcal{T} = \varnothing$, then $\mathscr{I}_X(\mathcal{T}) = \mathscr{A}_X$ vacuously. When X is clear from the context, we omit X in the notation, so $\mathscr{I}(\mathcal{T}) = \mathscr{I}_X(\mathcal{T})$.

Next, we show some basic properties of $\mathscr{I}(\mathcal{M}, \mathcal{T})$ and $\mathscr{I}(\mathcal{T})$, and how they are immediately connected.

Theorem 2.3.11. *Let X be a space, $\mathcal{M} \subset \mathcal{P}X$, and $\mathcal{T} \subset \mathcal{B}(X)$. Then*

(a) if $\mathcal{N} \subset \mathcal{M}$, then $\mathscr{I}(\mathcal{M}, \mathcal{T}) \subset \mathscr{I}(\mathcal{N}, \mathcal{T})$,

(b) if $\mathcal{S} \subset \mathcal{T}$, then $\mathscr{I}(\mathcal{M}, \mathcal{T}) \subset \mathscr{I}(\mathcal{M}, \mathcal{S})$,

(c) if $\mathcal{S} \subset \mathcal{T}$, then $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{S})$,

(d) $\mathscr{I}(\mathcal{M}, \mathcal{T})$ is a sub σ-algebra of $\mathscr{A}_X$,

(e) $\mathscr{I}(\mathcal{T})$ is a sub σ-algebra of $\mathscr{A}_X$,

(f) $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{M}, \mathcal{T})$, and

(g) $\mathscr{I}(\mathcal{T}) = \mathscr{I}(\mathcal{P}X, \mathcal{T})$.

Proof. For (a), as $\mathcal{N} \subset \mathcal{M}$, if $E \in \mathscr{I}(\mathcal{M}, \mathcal{T})$, then for every $\mu \in \mathcal{M}$, we have $E \in \mathscr{I}(\mu, \mathcal{T})$. As $\mathcal{N} \subset \mathcal{M}$, it follows that for every $\nu \in \mathcal{N} \subset \mathcal{M}$, we have $E \in \mathscr{I}(\nu, \mathcal{T})$, and therefore $E \in \mathscr{I}(\mathcal{N}, \mathcal{T})$. With E arbitrary, this gives $\mathscr{I}(\mathcal{M}, \mathcal{T}) \subset \mathscr{I}(\mathcal{N}, \mathcal{T})$.

For (b) and (c), the proof is nearly identical as for (a).

Next, for (d), we first prove that for $\mu \in \mathcal{P}X$ and $T \in \mathcal{T}$, that $\mathscr{I}(\mu, T)$ is a σ-algebra. Indeed, we have that $\mu(\varnothing \triangle T^{-1}(\varnothing)) = \mu(\varnothing \triangle \varnothing) = \mu(\varnothing) = 0$, and so

$\varnothing \in \mathscr{I}(\mu, T)$. Additionally, we have that $\mu(X \triangle T^{-1}(X)) = \mu(X \triangle X) = \mu(\varnothing) = 0$, and so $X \in \mathscr{I}(\mu, T)$. Now, suppose that $E \in \mathscr{I}(\mu, T)$. Then $(X \setminus E) \triangle T^{-1}(X \setminus E) = (X \setminus E) \triangle X \setminus T^{-1}(E) = E \triangle T^{-1}(E)$ always holds for the symmetric difference, and therefore

$$\mu((X \setminus E) \triangle T^{-1}(X \setminus E)) = \mu(E \triangle T^{-1}(E)) = 0,$$

because $E \in \mathscr{I}(\mu, T)$, and thus $X \setminus E \in \mathscr{I}(\mu, T)$. Finally, let $\{E_n\}_{n \in \mathbb{N}} \subset \mathscr{I}(\mu, T)$. Then, we have that

$$\begin{aligned} \left(\bigcup_{n \in \mathbb{N}} E_n \right) \triangle T^{-1} \left(\bigcup_{n \in \mathbb{N}} E_n \right) &= \left(\bigcup_{n \in \mathbb{N}} E_n \right) \triangle \left(\bigcup_{n \in \mathbb{N}} T^{-1}(E_n) \right) \\ &\subset \bigcup_{n \in \mathbb{N}} E_n \triangle T^{-1}(E_n), \end{aligned}$$

and so by the monotonicity and subadditivity of μ, we have

$$\begin{aligned} \mu \left(\left(\bigcup_{n \in \mathbb{N}} E_n \right) \triangle T^{-1} \left(\bigcup_{n \in \mathbb{N}} E_n \right) \right) &\leq \mu \left(\bigcup_{n \in \mathbb{N}} E_n \triangle T^{-1}(E_n) \right) \\ &\leq \sum_{n \in \mathbb{N}} \mu(E_n \triangle T^{-1}(E_n)) \\ &= \sum_{n \in \mathbb{N}} 0 = 0, \end{aligned}$$

as $\mu(E_n \triangle T^{-1}(E_n)) = 0$ because $E_n \in \mathscr{I}(\mu, T)$. Therefore, $\bigcup_{n \in \mathbb{N}} E_n \in \mathscr{I}(\mu, T)$, and so $\mathscr{I}(\mu, T)$ is a σ-algebra. For arbitrary $\mathcal{M} \subset \mathcal{P}X$ and $\mathcal{T} \subset \mathcal{B}(X)$, if $\mathcal{M} = \varnothing$ or $\mathcal{T} = \varnothing$, then $\mathscr{I}(\mathcal{M}, \mathcal{T})$ is a σ-algebra. Otherwise, if both $\mathcal{M}$ and $\mathcal{T}$ are nonempty, then $\mathscr{I}(\mathcal{M}, \mathcal{T})$ is an intersection of σ-algebras (as $\mathscr{I}(\mu, T)$ is a σ-algebra), and is therefore a σ-algebra.

Then, for (e), we first prove that for $T \in \mathcal{T}$, that $\mathscr{I}(T)$ is a σ-algebra. Indeed, we have that $T^{-1}(\varnothing) = \varnothing$, and that $T^{-1}(X) = X$, and therefore $\varnothing, X \in \mathscr{I}(T)$. Then, if $E \in \mathscr{I}(T)$, we have that $T^{-1}(E) = E$, and so $T^{-1}(X \setminus E) = X \setminus T^{-1}(E) = X \setminus E$, so this gives that $X \setminus E \in \mathscr{I}(T)$. Finally, let $\{E_n\}_{n \in \mathbb{N}} \subset \mathscr{I}(T)$. Then $T^{-1}(E_n) = E_n$

for every $n \in \mathbb{N}$, and so

$$T^{-1}\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \bigcup_{n\in\mathbb{N}} T^{-1}(E_n) = \bigcup_{n\in\mathbb{N}} E_n,$$

which means that $\bigcup_{n\in\mathbb{N}} E_n \in \mathscr{I}(T)$, and therefore $\mathscr{I}(T)$ is a σ-algebra. Lastly, $\mathscr{I}(\mathcal{T})$ is an intersection of sets of the form $\mathscr{I}(T)$, which are all σ-algebras, and therefore $\mathscr{I}(\mathcal{T})$ is a σ-algebra (unless $\mathcal{T} = \varnothing$, in which case $\mathscr{I}(\mathcal{T}) = \mathscr{A}_X$ is clearly a σ-algebra).

Now, for (f), let $E \in \mathscr{I}(\mathcal{T})$. Then $T^{-1}(E) = E$, and so $E \triangle T^{-1}(E) = \varnothing$, and therefore for any $\mu \in \mathcal{M}$, we have that $\mu(E \triangle T^{-1}(E)) = \mu(\varnothing) = 0$, which gives that $E \in \mathscr{I}(\mathcal{M}, \mathcal{T})$. As $E \in \mathscr{I}(\mathcal{T})$ was arbitrary, this gives $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{M}, \mathcal{T})$ as desired.

Finally, for (g), let $E \in \mathscr{I}(\mathcal{P}X, \mathcal{T})$. Then for every $x \in X$ and $T \in \mathcal{T}$, we have that $\delta_x(E \triangle T^{-1}(E)) = 0$, and therefore $x \notin E \triangle T^{-1}(E)$. As this is true for every $x \in X$, we must have that $E \triangle T^{-1}(E) = \varnothing$, which implies that $T^{-1}(E) = E$. As such, $E \in \mathscr{I}(\mathcal{T})$, and since $E \in \mathscr{I}(\mathcal{P}X, \mathcal{T})$ was arbitrary, this gives $\mathscr{I}(\mathcal{P}X, \mathcal{T}) \subset \mathscr{I}(\mathcal{T})$. By (f), we also have the reverse inclusion, and therefore $\mathscr{I}(\mathcal{T}) = \mathscr{I}(\mathcal{P}X, \mathcal{T})$. □

The invariant sets of a dynamical system are those which are trivial in some core sense of the dynamical system, in that they are essentially unchanged by the dynamics. For measures systems, we have the following notion.

Definition 2.3.12. Let X be a space, and $\mu \in \mathcal{P}X$. A set $E \in \mathscr{A}_X$ is called *μ-trivial* if $\mu(E) \in \{0, 1\}$, so either E is a null set or a full set. Let $\mathscr{T}_X(\mu) \subset \mathscr{A}_X$ denote the collection of μ-trivial sets. Furthermore, for a measure system $(X, \mathcal{M})$, a set $E \in \mathscr{A}_X$ is *$\mathcal{M}$-trivial* if for every $\mu \in \mathcal{M}$, we have that E is μ-trivial. Define

$$\mathscr{T}_X(\mathcal{M}) = \bigcap_{\mu\in\mathcal{M}} \mathscr{T}_X(\mu)$$

to be the set of all $\mathcal{M}$-trivial sets. As a convention, we define $\mathscr{T}_X(\varnothing) = \mathscr{A}_X$. Also, a set $E \in \mathscr{A}_X$ is called *μ-null* if $\mu(E) = 0$. Let $\mathscr{N}_X(\mu) \subset \mathscr{A}_X$ denote the collection of

μ-trivial sets. Furthermore, for a measure system $(X, \mathcal{M})$, a set $E \in \mathscr{A}_X$ is $\mathcal{M}$*-null* if for every $\mu \in \mathcal{M}$, we have that E is μ-null. Define

$$\mathscr{N}_X(\mathcal{M}) = \bigcap_{\mu \in \mathcal{M}} \mathscr{N}_X(\mu)$$

to be the set of all $\mathcal{M}$-null sets. As a convention, we define $\mathscr{N}_X(\varnothing) = \mathscr{A}_X$. When X is clear from the context, we will omit it and use $\mathscr{T}(\mathcal{M}) = \mathscr{T}_X(\mathcal{M})$ and $\mathscr{N}(\mathcal{M}) = \mathscr{N}_X(\mathcal{M})$.

Informally, $\mathcal{M}$-trivial sets contain either "almost everything" or "almost nothing" with respect to every measure in $\mathcal{M}$ (though, different measures may disagree on whether or not a particular set is one or the other). The null sets are those which are "almost nothing" with respect to every measure in the system. At the intersection of invariant sets and trivial sets are the ergodic measures.

Definition 2.3.13. Let $(X, \mathcal{T})$ be a dynamical system. A measure $\mu \in \mathcal{I}(\mathcal{T})$ is $\mathcal{T}$*-ergodic* if $\mathscr{I}(\mu, \mathcal{T}) \subset \mathscr{T}(\mu)$, or in other words that for $E \in \mathscr{I}(\mu, \mathcal{T})$, we have $\mu(E) \in \{0, 1\}$. Let $\mathcal{E}_X(\mathcal{T}) \subset \mathcal{I}(\mathcal{T})$ denote the set of all $\mathcal{T}$-ergodic measures. We omit X in the notation when X is clear from the context, so $\mathcal{E}(\mathcal{T}) = \mathcal{E}_X(\mathcal{T})$.

Following this definition, we can strengthen the characterization of ergodicity.

Lemma 2.3.14. *For a dynamical system $(X, \mathcal{T})$, we have that $\mu \in \mathcal{I}(\mathcal{T})$ is $\mathcal{T}$-ergodic if and only if*

$$\mathscr{I}(\mu, \mathcal{T}) = \mathscr{T}(\mu).$$

Furthermore, for any $\mathcal{M} \subset \mathcal{E}(\mathcal{T})$, we have that

$$\mathscr{I}(\mathcal{M}, \mathcal{T}) = \mathscr{T}(\mathcal{M}).$$

Proof. For the reverse direction, if it holds for $\mu \in \mathcal{I}(\mathcal{T})$ that $\mathscr{I}(\mu, \mathcal{T}) = \mathscr{T}(\mu)$, then

$\mathscr{I}(\mu,\mathcal{T}) \subset \mathscr{T}(\mu)$, which gives that μ is $\mathcal{T}$-ergodic by definition, so we now only need to prove that $\mathscr{T}(\mu) = \mathscr{I}(\mu,\mathcal{T})$ if μ is $\mathcal{T}$-ergodic.

Now, if $\mu \in \mathcal{I}(\mathcal{T})$ is $\mathcal{T}$-ergodic, then by definition $\mathscr{I}(\mu,\mathcal{T}) \subset \mathscr{T}(\mu)$, so we must show the reverse inclusion for the desired result. Now, for $E \in \mathscr{T}(\mu)$, we either have that $\mu(E) = 0$ or $\mu(E) = 1$. Additionally, since $\mu \in \mathcal{I}(\mathcal{T})$, we have for every $T \in \mathcal{T}$ that $\mathcal{P}T(\mu) = \mu$. If $\mu(E) = 0$, then

$$\begin{aligned} \mu(E \triangle T^{-1}(E)) &\leq \mu(E \cup T^{-1}(E)) \leq \mu(E) + \mu(T^{-1}(E)) \\ &= 0 + [\mathcal{P}T(\mu)](E) = \mu(E) = 0, \end{aligned}$$

and therefore $E \in \mathscr{I}(\mu, T)$. On the other hand, if $\mu(E) = 1$, then $\mu(X \setminus E) = 0$, and so $\mu((X \setminus E) \triangle T^{-1}(X \setminus E)) = 0$ by the result above. But $T^{-1}(X \setminus E) = X \setminus T^{-1}(E)$, and $A \triangle B = (X \setminus A) \triangle (X \setminus B)$ for any $A, B \subset X$, and therefore

$$0 = \mu((X \setminus E) \triangle T^{-1}(X \setminus E)) = \mu((X \setminus E) \triangle X \setminus T^{-1}(E)) = \mu(E \triangle T^{-1}(E)),$$

so $E \in \mathscr{I}(\mu, T)$. As such, for every $E \in \mathscr{T}(\mu)$, we have that $E \in \mathscr{I}(\mu, T)$ for every $T \in \mathcal{T}$, and so $E \in \mathscr{I}(\mu,\mathcal{T})$. With $E \in \mathscr{T}(\mu)$ arbitrary, this gives that $\mathscr{T}(\mu) \subset \mathscr{I}(\mu,\mathcal{T})$ as desired.

Finally, by definition and the result above we have

$$\mathscr{I}(\mathcal{M},\mathcal{T}) = \bigcap_{\mu \in \mathcal{M}} \mathscr{I}(\mu,\mathcal{T}) = \bigcap_{\mu \in \mathcal{M}} \mathscr{T}(\mu) = \mathscr{T}(\mathcal{M}).$$

□

As such, the ergodic measures are those for which the notions of triviality and invariance are identical. Interestingly, there is also a geometric classification of these measures within $\mathcal{I}(\mathcal{T})$.

Proposition 2.3.15. *Let $(X, \mathcal{T})$ be a dynamical system, and $(X, \mathcal{I}(\mathcal{T}))$ its associated measure system. Then $\mathcal{I}(\mathcal{T})$ is convex (more accurately, a simplex), and the extreme points of $\mathcal{I}(\mathcal{T})$ are exactly $\mathcal{E}(\mathcal{T})$.*

Proof. This result is a summarization of the results of Chapter 12 of [43], however we will give the proof of the fact that $\mathcal{I}(\mathcal{T})$ is convex for the sake of completion.

Let $\mu, \nu \in \mathcal{I}(\mathcal{T})$, and let $t \in [0, 1]$. Then for any $T \in \mathcal{T}$, we have that $\mathcal{P}T(\mu) = \mu$ and $\mathcal{P}T(\nu) = \nu$, and therefore using Lemma 2.2.3, we have

$$\mathcal{P}T(t\mu + (1-t)\nu) = t\mathcal{P}T(\mu) + (1-t)\mathcal{P}T(\nu) = t\mu + (1-t)\nu,$$

and therefore $t\mu + (1-t)\nu \in \mathcal{I}(T)$. As this is true for every $T \in \mathcal{T}$, we have that $t\mu + (1-t)\nu \in \mathcal{I}(\mathcal{T})$, so $\mathcal{I}(\mathcal{T})$ is convex. □

Without further assumptions on $\mathcal{I}(\mathcal{T})$, we cannot say much more about its structure. Under the assumption that $\mathcal{I}(\mathcal{T})$ is closed, which is the case when $\mathcal{T}$ consists of continuous functions (Lemma 2.3.7), then these extreme points in some sense ecode enough information about the entirety of $\mathcal{I}(\mathcal{T})$.

Proposition 2.3.16. *Let $(X, \mathcal{T})$ be a dynamical system, and suppose its associated measure system $(X, \mathcal{I}(\mathcal{T}))$ is closed. Then*

(a) $\mathcal{E}(\mathcal{T})$ is a G_δ subset of $\mathcal{P}X$,

(b) for every $m \in \mathcal{P}^2X$ with $m(\mathcal{I}(\mathcal{T})) = 1$, we have $\beta(m) \in \mathcal{I}(\mathcal{T})$,

(c) for every $\mu \in \mathcal{I}(\mathcal{T})$, there exists a unique $m \in \mathcal{P}^2X$ with $m(\mathcal{E}(\mathcal{T})) = 1$ and $\beta(m) = \mu$, and

(d) if $\mathcal{E}(\mathcal{T})$ is compact (closed in $\mathcal{P}X$), then β is an affine homeomorphism of $\mathcal{P}\mathcal{E}(\mathcal{T})$ and $\mathcal{I}(\mathcal{T})$.

Proof. The first statement is given by the fact that $\mathcal{E}(\mathcal{T})$ are the extreme points of $\mathcal{I}(\mathcal{T})$ (Proposition 2.3.15), and Proposition 1.3 of [43].

For the second statement, let $T \in \mathcal{T}$. Since $m(\mathcal{I}(\mathcal{T})) = 1$, we have for every $E \in \mathscr{A}_X$ that

$$\begin{aligned}\mathcal{P}T[\beta(m)](E) = \beta[m](T^{-1}(E)) &= \int_{\mathcal{P}X} \mu(T^{-1}(E))\, m(\mathrm{d}\mu) \\ &= \int_{\mathcal{I}(\mathcal{T})} \mathcal{P}T[\mu](E)\, m(\mathrm{d}\mu) \\ &= \int_{\mathcal{I}(\mathcal{T})} \mu(E)\, m(\mathrm{d}\mu) = \int_{\mathcal{P}X} \mu(E)\, m(\mathrm{d}\mu) = \beta[m](E),\end{aligned}$$

and so $\mathcal{P}T[\beta(m)] = \beta(m)$. Since $T \in \mathcal{T}$ was arbitrary, we have that $\beta(m) \in \mathcal{I}(\mathcal{T})$.

The third statement comes from the fact that $\mathcal{I}(\mathcal{T})$ is a closed simplex in $\mathcal{P}X$ (Proposition 2.3.15), and the main Theorem of Chapter 10 of [43] (one of Choquet's Theorem).

For the fourth statement, Proposition 2.1.14 gives that $\beta : \mathcal{P}^2X \to \mathcal{P}X$ is a continuous surjection. Since $\mathcal{E}(\mathcal{T})$ is a closed subset of $\mathcal{P}X$, we have that $\mathcal{E}(\mathcal{T})$ is compact and so a space. This gives that $\mathcal{P}\mathcal{E}(\mathcal{T})$ is a space. By the second statement however, since $m \in \mathcal{P}\mathcal{E}(\mathcal{T}) \subset \mathcal{P}\mathcal{I}(\mathcal{T})$, we have that $m(\mathcal{I}(\mathcal{T})) \geq m(\mathcal{E}(\mathcal{T})) = 1$ and so $\beta(m) \in \mathcal{I}(\mathcal{T})$. Additionally, the third statement is that β is an injection, and so β is a continuous bijection from $\mathcal{P}\mathcal{E}(\mathcal{T})$ to $\mathcal{I}(\mathcal{T})$. Since $\mathcal{P}\mathcal{E}(\mathcal{T})$ is compact and $\mathcal{I}(\mathcal{T})$ is Hausdorff, β is a homeomorphism between $\mathcal{P}\mathcal{E}(\mathcal{T})$ and $\mathcal{I}(\mathcal{T})$. □

As a result, when it is known that $\mathcal{I}_X(\mathcal{T})$ is closed, then providing a characterization of $\mathcal{E}_X(\mathcal{T})$ automatically gives a characterization of $\mathcal{I}_X(\mathcal{T})$ by this proposition.

2.3.2 The wobbling extension

Before we move on to proving properties of $\mathcal{F}_X(\mathcal{M})$ for a measure system $(X, \mathcal{M})$ we first need to define the notion of the wobbling extension of a collection of transformations. We begin with its definition, and then briefly discuss it.

Definition 2.3.17. Let $(X, \mathcal{T})$ be a dynamical system, and let $\mathcal{S} = \{S_n\}_{n\in\mathbb{N}} \subset \mathcal{T}$ be a countable, indexed subset of $\mathcal{T}$. Let $\alpha = \{\alpha_n\}_{n\in\mathbb{N}} \subset \mathscr{A}_X$ be a countable, measurable partition of X (possibly consisting of empty sets), so $\bigcup_{n\in\mathbb{N}} \alpha_n = X$, and for every $n \neq m \in \mathbb{N}$, $\alpha_n \cap \alpha_m = \varnothing$. If the collection $\beta = \{S_n^{-1}(\alpha_n)\}_{n\in\mathbb{N}}$ is a countable, measurable partition of X, then $(\mathcal{S}, \alpha)$ is an $(X, \mathcal{T})$-*wobbling pair*. For any $(X, \mathcal{T})$-wobbling pair $(\mathcal{S}, \alpha)$, we may define the *α-wobble of $\mathcal{S}$*, denoted $W_\alpha^{\mathcal{S}} : X \to X$, by defining for each $x \in \beta_n = S_n^{-1}(\alpha_n)$, that $W_\alpha^{\mathcal{S}}(x) = S_n(x)$. This defines a function from X to itself, as β is a partition of X. Then the *wobbling extension of $\mathcal{T}$* is defined as

$$W_X(\mathcal{T}) = \{W_\alpha^{\mathcal{S}} : (\mathcal{S}, \alpha) \text{ is an } (X, \mathcal{T})\text{-wobbling pair}\}.$$

When clear from the context we will omit the X and simply write $W(\mathcal{T}) = W_X(\mathcal{T})$.

This definition of the wobbling extension is inspired by the definition of the wobbling group of a group given in Definition 5.9 of [50]. The definition has been modified to handle objects which are not bijections, while also restricting to generally measurable objects (and as is shown below, the wobbling extension itself only contains measurable functions). The primary motivation for this particular definition is that it allows for an analogous version of Lemma 5.13 of the same paper to be proven for this new notion of the wobbling extension (done in the following section), which states that any invariant mean on the group will be invariant under the entire wobbling group. In principal, the wobbling extension of a collection of transformations is quite large in comparison to the original collection, so it is perhaps surprising that invariance should be maintained. In order to further understand the behavior of the wobbling extension, we first need the following result.

Lemma 2.3.18. *Let $(X, \mathcal{T})$ be a dynamical system, and let $(\mathcal{S}, \alpha)$ be an $(X, \mathcal{T})$-*

wobbling pair. Then for each $n \in \mathbb{N}$ and $E \subset X$, we have

$$[W_\alpha^{\mathcal{S}}]^{-1}(E \cap \alpha_n) = S_n^{-1}(E \cap \alpha_n).$$

Proof. First, suppose that $x \in [W_\alpha^{\mathcal{S}}]^{-1}(E \cap \alpha_n)$. Since we are taking a preimage of an intersection, we have $x \in [W_\alpha^{\mathcal{S}}]^{-1}(E) \cap [W_\alpha^{\mathcal{S}}]^{-1}(\alpha_n)$, and so we have $x \in [W_\alpha^{\mathcal{S}}]^{-1}(E)$ and $x \in [W_\alpha^{\mathcal{S}}]^{-1}(\alpha_n)$. Since $x \in [W_\alpha^{\mathcal{S}}]^{-1}(\alpha_n)$, and $\{S_n^{-1}(\alpha_n)\}_{n\in\mathbb{N}}$ is a partition of X, we must have that $x \in S_m^{-1}(\alpha_m)$ for some $m \in \mathbb{N}$. But then, this implies $W_\alpha^{\mathcal{S}}(x) = S_m(x) \in \alpha_m$, and thus we have

$$x \in [W_\alpha^{\mathcal{S}}]^{-1}(\alpha_n) \cap [W_\alpha^{\mathcal{S}}]^{-1}(\alpha_m) = [W_\alpha^{\mathcal{S}}]^{-1}(\alpha_n \cap \alpha_m).$$

As $\alpha_n \cap \alpha_m = \varnothing$ for $n \neq m$, it must be that $n = m$, and so we have that $x \in S_n^{-1}(\alpha_n)$, thus $S_n(x) \in \alpha_n$. Additionally, since $x \in [W_\alpha^{\mathcal{S}}]^{-1}(E)$, which gives that $S_n(x) = W_\alpha^{\mathcal{S}}(x) \in E$, and therefore $S_n(x) \in E \cap \alpha_n$ and $x \in S_n^{-1}(E \cap \alpha_n)$. As $x \in [W_E^{\mathcal{S}}]^{-1}(E \cap \alpha_n)$ was arbitrary, we have $[W_E^{\mathcal{S}}]^{-1}(E \cap \alpha_n) \subset S_n^{-1}(E \cap \alpha_n)$.

Now, suppose that $x \in S_n^{-1}(E \cap \alpha_n) = S_n^{-1}(E) \cap S_n^{-1}(\alpha_n)$, so we have $x \in S_n^{-1}(E)$ and $x \in S_n^{-1}(\alpha_n)$. With $x \in S_n^{-1}(\alpha_n)$, we have by definition that $W_\alpha^{\mathcal{S}}(x) = S_n(x) \in \alpha_n$, and so $W_\alpha^{\mathcal{S}}(x) \in \alpha_n$. Additionally $x \in S_n^{-1}(E)$ gives that $S_n(x) \in E$, and thus $W_\alpha^{\mathcal{S}}(x) = S_n(x) \in E$. Putting these together gives that $W_\alpha^{\mathcal{S}}(x) \in E \cap \alpha_n$, and so $x \in [W_\alpha^{\mathcal{S}}]^{-1}(E \cap \alpha_n)$. As $x \in S_n^{-1}(E \cap \alpha_n)$ was arbitrary, we have that $S_n^{-1}(E \cap \alpha_n) \subset [W_\alpha^{\mathcal{S}}]^{-1}(E \cap \alpha_n)$.

Putting the results of the two previous paragraphs together, we obtain the desired result. □

With this lemma, we may now prove the following main properties of the wobbling extension of a collection of measurable transformations.

Theorem 2.3.19. *Let X be a space. Then the wobbling extension operator $W(\cdot)$ on* *the subsets of $\mathcal{B}(X)$ is a closure operator, that is, for any dynamical system $(X, \mathcal{T})$, we have*

(a) $W(\mathcal{T}) \subset \mathcal{B}(X)$,

(b) $\mathcal{T} \subset W(\mathcal{T})$,

(c) If $\mathcal{U} \subset \mathcal{T}$*, then* $W(\mathcal{U}) \subset W(\mathcal{T})$*, and*

(d) $W(W(\mathcal{T})) = W(\mathcal{T})$.

Proof. First, we must show that $W(\mathcal{T}) \subset \mathcal{B}(X)$, which requires showing that $W_\alpha^{\mathcal{S}}$ is measurable for any $(X, \mathcal{T})$-wobbling pair $(\mathcal{S}, \alpha)$. Indeed, let $(\mathcal{S}, \alpha)$ be an $(X, \mathcal{T})$-wobbling pair, and let $E \in \mathscr{A}_X$. Then the collection $\{E \cap \alpha_n\}_{n \in \mathbb{N}}$ is a partition of E into measurable subsets, and thus by Lemma 2.3.18,

$$[W_\alpha^{\mathcal{S}}]^{-1}(E) = [W_\alpha^{\mathcal{S}}]^{-1}\left(\bigsqcup_{n \in \mathbb{N}} E \cap \alpha_n\right) = \bigsqcup_{n \in \mathbb{N}} [W_\alpha^{\mathcal{S}}]^{-1}(E \cap \alpha_n) = \bigsqcup_{n \in \mathbb{N}} S_n^{-1}(E \cap \alpha_n).$$

With each $S_n \in \mathcal{T} \subset \mathcal{B}(X)$, each S_n is measureable, and since E and α_n are measurable, so is $E \cap \alpha_n$, and therefore $S_n^{-1}(E \cap \alpha_n)$ is as well. As such, $[W_\alpha^{\mathcal{S}}]^{-1}(E)$ is a countable (disjoint) union of measurable sets, and is therefore measurable. As $E \in \mathscr{A}_X$ was arbitrary, this gives that $W_\alpha^{\mathcal{S}} \in \mathcal{B}(X)$, and so $W(\mathcal{T}) \subset \mathcal{B}(X)$.

Next, to show that $\mathcal{T} \subset W(\mathcal{T})$, let $T \in \mathcal{T}$, and it suffices to find a $(X, \mathcal{T})$-wobbling pair $(\mathcal{S}, \alpha)$ such that $W_\alpha^{\mathcal{S}} = T$. Let $S_n = T$ for all n, and let $\alpha_1 = X$, and $\alpha_n = \varnothing$ for $n \geq 2$. Then for $x \in S_1^{-1}(\alpha_1) = S_1^{-1}(X) = X$, we have $W_\alpha^{\mathcal{S}}(x) = S_1(x) = T(x)$, and thus $W_\alpha^{\mathcal{S}} = T$.

Now, if $\mathcal{U} \subset \mathcal{T}$, and $(\mathcal{S}, \alpha)$ is an $(X, \mathcal{U})$-wobbling pair, then it is clearly also an $(X, \mathcal{T})$-wobbling pair, and so for any $W_\alpha^{\mathcal{S}} \in W(\mathcal{U})$, it is automatically the case that $W_\alpha^{\mathcal{S}} \in W(\mathcal{T})$.

Finally, we show that $W(W(\mathcal{T})) = W(\mathcal{T})$. For an element $V \in W(W(\mathcal{T}))$, let $(\mathcal{S}, \alpha)$ be its $(X, W(\mathcal{T}))$-wobbling pair. For each $S_n \in \mathcal{S}$, we have that $S_n \in W(\mathcal{T})$, so let $(\mathcal{U}^n, \beta^n)$ be its associated $(X, \mathcal{T})$-wobbling pair. We now create an $(X, \mathcal{T})$-wobbling pair $(\mathcal{R}, \gamma)$ such that the γ-wobble of $\mathcal{R}$ is V. Indeed, let $\mathcal{R} = \{R_{n,m}\}_{n,m\in\mathbb{N}}$ be defined by $R_{n,m} = U^n_m \in \mathcal{U}^n$, and let $\gamma = \{\gamma_{n,m}\}_{n,m\in\mathbb{N}}$ be defined by $\gamma_{n,m} = \alpha_n \cap \beta^n_m$. For simplicity, it is easier to define these with double indices, however they are both countable nevertheless. It is clear that $U^n_m \in \mathcal{U}^n \subset \mathcal{T}$, and so $\mathcal{R} \subset \mathcal{T}$. Furthermore, let $x \in X$. Since α is a partition of X, there exists exactly one n such that $x \in \alpha_n$. Furthermore, β^n is also a partition of X, so there exists exactly one m such that $x \in \beta^n_m$. As such, $x \in \alpha_n \cap \beta^n_m = \gamma_{n,m}$, so γ covers X. Then, for γ_{n_1,m_1} and γ_{n_2,m_2}, where $(n_1, m_1) \neq (n_2, m_2)$ (or in other words, $n_1 \neq n_2$ or $m_1 \neq m_2$), if $n_1 \neq n_2$, then

$$\gamma_{n_1,m_1} \cap \gamma_{n_2,m_2} \subset \alpha_{n_1} \cap \alpha_{n_2} = \varnothing,$$

as α is a partition of X. Otherwise, $n_1 = n_2$ but $m_1 \neq m_2$, and

$$\gamma_{n_1,m_1} \cap \gamma_{n_2,m_2} \subset \beta^{n_1}_{m_1} \cap \beta^{n_2}_{m_2} = \beta^{n_1}_{m_1} \cap \beta^{n_1}_{m_2} = \varnothing,$$

as β^{n_1} is a partition of X. As such, γ is a partition of X. Next, we must show that $\{R^{-1}_{n,m}(\gamma_{n,m})\}_{n,m\in\mathbb{N}}$ is a partition of X. Since $(\mathcal{S}, \alpha)$ is a $(X, W(\mathcal{T}))$-wobbling pair, we have that $\{S^{-1}_n(\alpha_n)\}_{n\in\mathbb{N}}$ is a partition of X, and it will suffice to show that $\{R^{-1}_{n,m}(\gamma_{n,m})\}_{m\in\mathbb{N}}$ is a partition of $S^{-1}_n(\alpha_n)$. Indeed, since β^n is a partition of X, so is $\{S^{-1}_n(\beta^n_m)\}_{m\in\mathbb{N}}$, and thus

$$S^{-1}_n(\alpha_n) = \bigsqcup_{m\in\mathbb{N}} S^{-1}_n(\alpha_n) \cap S^{-1}_n(\beta^n_m) = \bigsqcup_{m\in\mathbb{N}} S^{-1}_n(\alpha_n \cap \beta^n_m).$$

With S_n a wobbling itself, with $(X, \mathcal{T})$-wobbling pair $(\mathcal{U}^n, \beta^n)$, we have by Lemma 2.3.18 with $E = \alpha_n$ that

$$S^{-1}_n(\alpha_n \cap \beta^n_m) = [U^n_m]^{-1}(\alpha_n \cap \beta^n_m) = R^{-1}_{n,m}(\alpha_n \cap \beta^n_m).$$

As such, this gives us

$$S_n^{-1}(\alpha_n) = \bigsqcup_{m\in\mathbb{N}} R_{n,m}^{-1}(\alpha_n \cap \beta_m^n),$$

or in other words that $\{R_{n,m}^{-1}(\alpha_n \cap \beta_m^n)\}_{m\in\mathbb{N}}$ is a partition of $S_n^{-1}(\alpha_n)$. As such, $(\mathcal{R}, \gamma)$ is a wobbling pair, so let Z be its corresponding γ-wobble of $\mathcal{R}$. Lastly, recalling that V is the α-wobble of $\mathcal{S}$, we must show that $V(x) = Z(x)$ for all $x \in X$, so let $x \in X$. By definition, there exists exactly one $n \in \mathbb{N}$ such that $x \in \alpha_n$, and then $V(x) = S_n(x)$. Additionally, since S_n is the β^n-wobble of $\mathcal{U}^n$, we have that there exists exactly one $m \in \mathbb{N}$ such that $x \in \beta_m^n$, and $S_n(x) = U_m^n(x) = R_{n,m}(x)$. Also, this means that $x \in \alpha_n \cap \beta_m^n = \gamma_{n,m}$, and thus we have $Z(x) = R_{n,m}(x) = S_n(x) = V(x)$. Therefore, $V = Z \in W(\mathcal{T})$, and so we have shown that $V \in W(W(\mathcal{T}))$ is also in $W(\mathcal{T})$. As $V \in W(W(\mathcal{T}))$ was arbitrary, we have shown that $W(W(\mathcal{T})) \subset W(\mathcal{T})$. With part (b), we have that $W(\mathcal{T}) \subset W(W(\mathcal{T}))$, and thus we have that $W(W(\mathcal{T})) = W(\mathcal{T})$. □

As $W(\cdot)$ is a closure operator, any set of the form $W(\mathcal{T})$ is "closed" with respect to the operator, and an alternative characterization for such a "closed" set $\mathcal{S}$ is that $W(\mathcal{S}) = \mathcal{S}$. In order to give a concise term for such sets, we define the following terminology.

Definition 2.3.20. Let $(X, \mathcal{T})$ be a dynamical system. We say that $\mathcal{T}$ is *stable under wobbling*, or merely *stable*, if $W(\mathcal{T}) = \mathcal{T}$, or in other words that $\mathcal{T}$ is closed with respect to the closure operator $W(\cdot)$.

2.3.3 Properties of $\mathcal{F}_X(\mathcal{M})$

We can now turn our attention to the dual notion of the set of invariant measures for a dynamical system, the set of measure-preserving functions of a measure system. Somewhat surprisingly, there is little to no mention of this set in the literature. There are some results which do in some sense deal with this set [50], however in general it is not an object that is studied. As we demonstrate in future sections, this object plays a rather important role in the study of dynamical systems, so it is useful to understand its properties. To begin, we have the following simple result, analogous to Lemma 2.3.5.

Lemma 2.3.21. *Let $(X, \mathcal{M})$ be a measure system, and $\mathcal{N} \subset \mathcal{M}$. Then $\mathcal{F}(\mathcal{M}) \subset \mathcal{F}(\mathcal{N})$.*

Proof. If

$$T \in \mathcal{F}(\mathcal{M}) = \bigcap_{\mu \in \mathcal{M}} \mathcal{F}(\mu),$$

we have that $T \in \mathcal{F}(\mu)$ for every $\mu \in \mathcal{M}$. As $\mathcal{N} \subset \mathcal{M}$ by assumption, it follows that for every $\nu \in \mathcal{N} \subset \mathcal{M}$ that $T \in \mathcal{F}(\nu)$, and thus

$$T \in \bigcap_{\nu \in \mathcal{N}} \mathcal{F}(\nu) = \mathcal{F}(\mathcal{N}).$$

Since this holds for every $T \in \mathcal{F}(\mathcal{M})$, we obtain that $\mathcal{F}(\mathcal{M}) \subset \mathcal{F}(\mathcal{N})$. □

We now turn our attention to the main structural theorem for $\mathcal{F}(\mathcal{M})$. The first three of these properties are generally well known properties (although not usually stated in this manner), however the final two are novel to the knowledge of the author.

Theorem 2.3.22. *Let $(X, \mathcal{M})$ be a measure system, and $(X, \mathcal{F}(\mathcal{M}))$ its associated dynamical system. Then*

(a) $\mathcal{F}(\mathcal{M})$ contains the identity map, so is nonempty,

(b) $\mathcal{F}(\mathcal{M})$ is closed under composition, so is a monoid,

(c) If $T \in \mathcal{F}(\mathcal{M})$ is a bijection, then $T^{-1} \in \mathcal{F}(\mathcal{M})$,

(d) $\mathcal{F}(\mathcal{M})$ is stable under wobbling, and

(e) $\mathcal{F}(\mathcal{M})$ is (topologically) closed in $\mathcal{B}(X)$.

Proof. For (a), with $I : X \to X$ the identity function on X, which is clearly measurable and so $I \in \mathcal{B}(X)$. Then, we have for every $\mu \in \mathcal{M}$ and $E \in \mathscr{A}_X$ that

$$[\mathcal{P}I(\mu)](E) = \mu(I^{-1}E) = \mu(E),$$

as $I^{-1}(E) = E$, and thus $\mathcal{P}I(\mu) = \mu$. This gives that $I \in \mathcal{F}(\mu)$, and as this applies to every $\mu \in \mathcal{M}$, we have that $I \in \mathcal{F}(\mathcal{M})$, which also makes $\mathcal{F}(\mathcal{M})$ nonempty.

For (b), if $S, T \in \mathcal{F}(\mathcal{M})$ we have by Lemma 2.2.4 that $\mathcal{P}(T \circ S) = \mathcal{P}T \circ \mathcal{P}S$, and thus for any $\mu \in \mathcal{M}$, we have

$$\mathcal{P}[T \circ S](\mu) = [\mathcal{P}T \circ \mathcal{P}S](\mu) = \mathcal{P}T(\mathcal{P}S(\mu)).$$

With $S, T \in \mathcal{F}(\mathcal{M})$ and $\mu \in \mathcal{M}$, we have that $\mathcal{P}S(\mu) = \mu$ and $\mathcal{P}T(\mu) = \mu$, and so

$$\mathcal{P}[T \circ S](\mu) = \mathcal{P}T(\mathcal{P}S(\mu)) = \mathcal{P}T(\mu) = \mu,$$

and therefore $T \circ S \in \mathcal{F}(\mu)$. This is true for every $\mu \in \mathcal{M}$ which gives that $T \circ S \in \mathcal{F}(\mathcal{M})$. With $\mathcal{F}(\mathcal{M})$ containing the identity function, which is the identity for the operation of composition, and composition is always associative, this shows $\mathcal{F}(\mathcal{M})$ is a monoid.

For (c), let $T \in \mathcal{F}(\mathcal{M})$ such that T is a bijection. By Lemma 2.2.5, we have that $\mathcal{P}T$ is a bijection, and $(\mathcal{P}T)^{-1} = \mathcal{P}T^{-1}$. As such, for $\mu \in \mathcal{M}$, we have by definition that $\mathcal{P}T(\mu) = \mu$, and since $\mathcal{P}T$ is invertible, we have $\mu = (\mathcal{P}T)^{-1}(\mu) = \mathcal{P}T^{-1}(\mu)$, and thus $T^{-1} \in \mathcal{F}(\mu)$. Since this holds for every $\mu \in \mathcal{M}$, we have that $T^{-1} \in \mathcal{F}(\mathcal{M})$.

Next, for (d), it will suffice to show that $W(\mathcal{F}(\mathcal{M})) \subset \mathcal{F}(\mathcal{M})$, as Theorem 2.3.19 gives that $\mathcal{F}(\mathcal{M}) \subset W(\mathcal{F}(\mathcal{M}))$, and thus $W(\mathcal{F}(\mathcal{M})) = \mathcal{F}(\mathcal{M})$, showing that $\mathcal{F}(\mathcal{M})$ is stable. Let $T \in W(\mathcal{F}(\mathcal{M}))$, and $(\mathcal{S}, \alpha)$ the corresponding $(X, \mathcal{F}(\mathcal{M}))$-wobbling pair for T. Let $\mu \in \mathcal{M}$ and $E \in \mathscr{A}_X$. Then with α a partition of X, we also have that $\{T^{-1}(\alpha_n)\}_{n\in\mathbb{N}}$ is a partition of X, and thus with Lemma 2.3.18,

$$T^{-1}(E) = \bigsqcup_{n\in\mathbb{N}} T^{-1}(E) \cap T^{-1}(\alpha_n) = \bigsqcup_{n\in\mathbb{N}} T^{-1}(E \cap \alpha_n) = \bigsqcup_{n\in\mathbb{N}} S_n^{-1}(E \cap \alpha_n).$$

As this union is countable and disjoint, by the countable additivity of μ, we have

$$\begin{aligned}[\mathcal{P}T(\mu)](E) = \mu(T^{-1}(E)) &= \mu\left(\bigsqcup_{n\in\mathbb{N}} S_n^{-1}(E \cap \alpha_n)\right) \\ &= \sum_{n\in\mathbb{N}} \mu(S_n^{-1}(E \cap \alpha_n)) = \sum_{n\in\mathbb{N}} [\mathcal{P}S_n(\mu)](E \cap \alpha_n).\end{aligned}$$

As $S_n \in \mathcal{F}(\mathcal{M})$, and $\mu \in \mathcal{M}$, we have that $\mathcal{P}S_n(\mu) = \mu$. Furthermore, with α a partition of X, we have that $E = \bigsqcup_{n\in\mathbb{N}} E \cap \alpha_n$ and thus by the countable additivity of μ, we have

$$[\mathcal{P}T(\mu)](E) = \sum_{n\in\mathbb{N}} \mu(E \cap \alpha_n) = \mu\left(\bigsqcup_{n\in\mathbb{N}} E \cap \alpha_n\right) = \mu(E).$$

As this is true for every $E \in \mathscr{A}_X$, it follows that $\mathcal{P}T(\mu) = \mu$, and thus $T \in \mathcal{F}(\mu)$. As this holds for every $\mu \in \mathcal{M}$, this gives that $T \in \mathcal{F}(\mathcal{M})$. With $T \in W(\mathcal{F}(\mathcal{M}))$ arbitrary, this gives that $W(\mathcal{F}(\mathcal{M})) \subset \mathcal{F}(\mathcal{M})$, which is the desired result.

Finally, for (e), we have by definition that for any $\mu \in \mathcal{P}X$, we have $\mathcal{F}(\mu) = \{T \in \mathcal{B}(X) : \mathcal{P}T(\mu) = \mu\}$ and $\mathcal{P}T(\mu) = \rho_\mu(T)$ for any $T \in \mathcal{B}(X)$. Therefore, for every $\mu \in \mathcal{P}X$, it is the case that

$$\mathcal{F}(\mu) = \{T \in \mathcal{B}(X) : \rho_\mu(T) = \mu\} = \rho_\mu^{-1}(\{\mu\}).$$

By definition then,

$$\mathcal{F}(\mathcal{M}) = \bigcap_{\mu\in\mathcal{M}} \mathcal{F}(\mu) = \bigcap_{\mu\in\mathcal{M}} \rho_\mu^{-1}(\{\mu\}).$$

Now since $\mathcal{P}X$ is endowed with a Hausdorff topology, singletons are closed, and therefore $\{\mu\}$ is a closed subset of $\mathcal{P}X$. Additionally, ρ_μ is continuous by the definition of the topology on $\mathcal{B}(X)$, and therefore $\rho_\mu^{-1}(\{\mu\})$ is a closed subset of $\mathcal{B}(X)$ for every $\mu \in \mathcal{P}X$. As $\mathcal{F}(\mathcal{M})$ is written above as an intersection of such sets, $\mathcal{F}(\mathcal{M})$ is closed in $\mathcal{B}(X)$. □

2.3.4 The measure and dynamical completions

To summarize the results of this section so far, we have defined the notion of a measure system associated to a dynamical system, and also the notion of a dynamical system associated to a measure system, and have proven many properties of these associated systems. Now, note that we can do the following. Start with a dynamical system $(X, \mathcal{T})$, obtain its associated measure system $(X, \mathcal{I}(\mathcal{T}))$, and then obtain the associated dynamical system $(X, \mathcal{F}(\mathcal{I}(\mathcal{T})))$ to this measure system. This associated dynamical system encodes some information about the original associated measure system, and may prove to be useful in obtaining a description of the invariant measures. For instance, if it is possible to directly deduce elements of $\mathcal{F}(\mathcal{I}(\mathcal{T}))$ solely from $\mathcal{T}$ without identifying any invariant measures beforehand, these additional elements may be helpful in characterizing the invariant measures. Similarly, if we start with a measure system $(X, \mathcal{M})$, we may consider its associated dynamical system $(X, \mathcal{F}(\mathcal{M}))$, and then the associated measure system $(X, \mathcal{I}(\mathcal{F}(\mathcal{M})))$ to this dynamical one. This may also prove fruitful to identifying dynamical systems with particular measures as invariant measures, and while we prove some properties of this, our primary focus will be the first example here we have given. To begin developing this theory, we provide some terminology.

Definition 2.3.23. For a dynamical system $(X, \mathcal{T})$, define $\mathcal{T}^* = \mathcal{F}_X(\mathcal{I}_X(\mathcal{T}))$, and for a measure system $(X, \mathcal{M})$, defined $\mathcal{M}^\star = \mathcal{I}_X(\mathcal{F}_X(\mathcal{M}))$.

Following this definition, we have the following basic properties.

Lemma 2.3.24. *Let* $(X, \mathcal{T})$ *and* $(X, \mathcal{S})$ *be dynamical systems, and* $(X, \mathcal{M})$ *and* $(X, \mathcal{N})$ *be measure systems. Then*

(a) $\mathcal{T} \subset \mathcal{T}^*$*,*

(b) $\mathcal{M} \subset \mathcal{M}^\star$*,*

(c) If $\mathcal{T} \subset \mathcal{S}$*, then* $\mathcal{T}^* \subset \mathcal{S}^*$*, and*

(d) If $\mathcal{M} \subset \mathcal{N}$*, then* $\mathcal{M}^\star \subset \mathcal{N}^\star$*.*

Proof. For (a), let $S \in \mathcal{T}$. With

$$\mathcal{I}(\mathcal{T}) = \bigcap_{T \in \mathcal{T}} \mathcal{I}(T) = \bigcap_{T \in \mathcal{T}} \{\mu \in \mathcal{P}X : \mathcal{P}T(\mu) = \mu\},$$

it follows that we have for every $\mu \in \mathcal{I}(\mathcal{T})$ that $\mathcal{P}S(\mu) = \mu$. As such, $S \in \{T \in \mathcal{B}(X) : \mathcal{P}T(\mu) = \mu\}$ for every $\mu \in \mathcal{I}(\mathcal{T})$, and so

$$S \in \bigcap_{\mu \in \mathcal{I}(\mathcal{T})} \{T \in \mathcal{B}(X) : \mathcal{P}T(\mu) = \mu\} = \bigcap_{\mu \in \mathcal{I}(\mathcal{T})} \mathcal{F}(\mu) = \mathcal{F}(\mathcal{I}(\mathcal{T})) = \mathcal{T}^*.$$

Since this holds for every $S \in \mathcal{T}$, we have $\mathcal{T} \subset \mathcal{T}^*$.

For (b), let $\nu \in \mathcal{M}$. With

$$\mathcal{F}(\mathcal{M}) = \bigcap_{\mu \in \mathcal{M}} \mathcal{F}(\mu) = \bigcap_{\mu \in \mathcal{M}} \{T \in \mathcal{B}(X) : \mathcal{P}T(\mu) = \mu\},$$

it follows that we have for every $T \in \mathcal{F}(\mathcal{M})$ that $\mathcal{P}T(\nu) = \nu$. As such, $\nu \in \{\mu \in \mathcal{P}X : \mathcal{P}T(\mu) = \mu\}$ for every $T \in \mathcal{F}(\mathcal{M})$, and so

$$\nu \in \bigcap_{T \in \mathcal{F}(\mathcal{M})} \{\mu \in \mathcal{P}X : \mathcal{P}T(\mu) = \mu\} = \bigcap_{T \in \mathcal{F}(\mathcal{M})} \mathcal{I}(T) = \mathcal{I}(\mathcal{F}(\mathcal{M})) = \mathcal{M}^\star.$$

Since this holds for every $\nu \in \mathcal{M}$, we have $\mathcal{M} \subset \mathcal{M}^\star$.

For (c), if $\mathcal{T} \subset \mathcal{S}$, then by Lemma 2.3.5, we have that $\mathcal{I}(\mathcal{S}) \subset \mathcal{I}(\mathcal{T})$. Using $\mathcal{M} = \mathcal{I}(\mathcal{S})$ and $\mathcal{N} = \mathcal{I}(\mathcal{T})$, we have $\mathcal{M} \subset \mathcal{N}$, so it then follows by Lemma 2.3.21 that $\mathcal{F}(\mathcal{N}) \subset \mathcal{F}(\mathcal{M})$. With $\mathcal{F}(\mathcal{N}) = \mathcal{F}(\mathcal{I}(\mathcal{T})) = \mathcal{T}^*$ and $\mathcal{F}(\mathcal{M}) = \mathcal{F}(\mathcal{I}(\mathcal{S})) = \mathcal{S}^*$, we obtain $\mathcal{T}^* \subset \mathcal{S}^*$.

For (d), if $\mathcal{M} \subset \mathcal{N}$, then by Lemma 2.3.21, we have that $\mathcal{F}(\mathcal{N}) \subset \mathcal{F}(\mathcal{M})$. Using $\mathcal{T} = \mathcal{F}(\mathcal{N})$ and $\mathcal{S} = \mathcal{F}(\mathcal{M})$, we have $\mathcal{T} \subset \mathcal{S}$, so it then follows by Lemma 2.3.5 that $\mathcal{I}(\mathcal{S}) \subset \mathcal{I}(\mathcal{T})$. With $\mathcal{I}(\mathcal{S}) = \mathcal{I}(\mathcal{F}(\mathcal{M})) = \mathcal{M}^\star$ and $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{F}(\mathcal{N})) = \mathcal{N}^\star$, we obtain $\mathcal{M}^\star \subset \mathcal{N}^\star$. □

From these inequalities, we can prove the following equalities.

Lemma 2.3.25. *Let $(X, \mathcal{T})$ be a dynamical system and let $(X, \mathcal{M})$ be a measure system. Then*

(a) $\mathcal{I}(\mathcal{T})^\star = \mathcal{I}(\mathcal{F}(\mathcal{I}(\mathcal{T}))) = \mathcal{I}(\mathcal{T}^*) = \mathcal{I}(\mathcal{T})$,

(b) $\mathcal{F}(\mathcal{M})^* = \mathcal{F}(\mathcal{I}(\mathcal{F}(\mathcal{M}))) = \mathcal{F}(\mathcal{M}^\star) = \mathcal{F}(\mathcal{M})$,

(c) $(\mathcal{T}^*)^* = \mathcal{T}^*$, *and*

(d) $(\mathcal{M}^\star)^\star = \mathcal{M}^\star$.

Proof. For (a), the first two equalities are evident from the definitions of completions. Next, by Lemma 2.3.24(a), we have that $\mathcal{T} \subset \mathcal{T}^*$, and by Lemma 2.3.5, that $\mathcal{I}(\mathcal{T}^*) \subset \mathcal{I}(\mathcal{T})$. Finally, by Lemma 2.3.24(b), we have that $\mathcal{I}(\mathcal{T}) \subset \mathcal{I}(\mathcal{T})^\star$. Connecting these containments together, we have

$$\mathcal{I}(\mathcal{T})^\star = \mathcal{I}(\mathcal{T}^*) \subset \mathcal{I}(\mathcal{T}) \subset \mathcal{I}(\mathcal{T})^\star,$$

and thus all of these sets are equal, and $\mathcal{I}(\mathcal{T})^\star = \mathcal{I}(\mathcal{T})$.

For (b), the first two equalities are evident from the definitions of completions. Next, by Lemma 2.3.24(b), we have that $\mathcal{M} \subset \mathcal{M}^\star$, and by Lemma 2.3.21, that

$\mathcal{F}(\mathcal{M}^\star) \subset \mathcal{F}(\mathcal{M})$. Finally, by Lemma 2.3.24(a), we have that $\mathcal{F}(\mathcal{M}) \subset \mathcal{F}(\mathcal{M})^*$. Connecting these containments together, we have

$$\mathcal{F}(\mathcal{M})^* = \mathcal{F}(\mathcal{M}^\star) \subset \mathcal{F}(\mathcal{M}) \subset \mathcal{F}(\mathcal{M})^*,$$

and thus all of these sets are equal, and $\mathcal{F}(\mathcal{M})^* = \mathcal{F}(\mathcal{M})$.

For (c), we have by part (b) with $\mathcal{M} = \mathcal{I}(\mathcal{T})$ that

$$(\mathcal{T}^*)^* = \mathcal{F}(\mathcal{I}(\mathcal{T}))^* = \mathcal{F}(\mathcal{I}(\mathcal{T})) = \mathcal{T}^*.$$

For (d), we have by part (a) with $\mathcal{T} = \mathcal{F}(\mathcal{M})$ that

$$(\mathcal{M}^\star)^\star = \mathcal{I}(\mathcal{F}(\mathcal{M}))^\star = \mathcal{I}(\mathcal{F}(\mathcal{M})) = \mathcal{M}^\star.$$

□

Combining the results of the two previous lemmas, we obtain the following theorem which summarizes all of the core properties of taking the completion of an object.

Theorem 2.3.26. *The operator $\cdot^*$ is a closure operator on the subsets of $\mathcal{B}(X)$. This means that for $\mathcal{T} \subset \mathcal{B}(X)$, we have*

(a) $\mathcal{T} \subset \mathcal{T}^*$*,*

(b) if $\mathcal{S} \subset \mathcal{T}$*, then* $\mathcal{S}^* \subset \mathcal{T}^*$*, and*

(c) $(\mathcal{T}^*)^* = \mathcal{T}^*$*.*

The operator $\cdot^\star$ is a closure operator on the subsets of $\mathcal{P}X$. This means that for $\mathcal{M} \subset \mathcal{B}(X)$, we have

(a) $\mathcal{M} \subset \mathcal{M}^\star$*,*

(b) if $\mathcal{N} \subset \mathcal{M}$*, then* $\mathcal{N}^* \subset \mathcal{M}^*$*, and*

(c) $(\mathcal{M}^\star)^\star = \mathcal{M}^\star$.

Proof. This is just a combination of Lemmas 2.3.24 and 2.3.25. □

As both $\cdot^*$ and $\cdot^\star$ are closure operators, we give the following definitions for the closed sets under these operators.

Definition 2.3.27. For a dynamical system $(X, \mathcal{T})$, if it holds that $\mathcal{T} = \mathcal{T}^*$, then $\mathcal{T}$ is said to be *measure-complete*. For a measure system $(X, \mathcal{M})$, if it holds that $\mathcal{M} = \mathcal{M}^\star$, then $\mathcal{M}$ is said to be *dynamically complete*.

Following this, we have that the following systems are complete.

Proposition 2.3.28. *Let $(X, \mathcal{T})$ be a dynamical system and $(X, \mathcal{M})$ a measure system. Then*

(a) $(X, \mathcal{F}(\mathcal{M}))$ *is measure-complete,*

(b) $(X, \mathcal{I}(\mathcal{T}))$ *is dynamically-complete,*

(c) $(X, \mathcal{T}^*)$ *is measure-complete, and*

(d) $(X, \mathcal{M}^\star)$ *is dynamically-complete.*

Proof. By Lemma 2.3.25, we have that $(\mathcal{F}(\mathcal{M}))^* = \mathcal{F}(\mathcal{M})$, and therefore $(X, \mathcal{F}(\mathcal{M}))$ is measure-complete. Also by Lemma 2.3.25, we have that $(\mathcal{I}(\mathcal{T}))^\star = \mathcal{I}(\mathcal{T})$, and therefore $(X, \mathcal{I}(\mathcal{T}))$ is dynamically-complete. By Theorem 2.3.26, we have that $(\mathcal{T}^*)^* = \mathcal{T}^*$, and therefore $(X, \mathcal{T}^*)$ is measure-complete. Finally, by Theorem 2.3.26, we have that $(\mathcal{M}^\star)^\star = \mathcal{M}$, and therefore $(X, \mathcal{M}^\star)$ is dynamically-complete. □

In light of this proposition, we give the following definitions.

Definition 2.3.29. The *measure-completion* of the dynamical system $(X, \mathcal{T})$ is the dynamical system $(X, \mathcal{T}^*)$, and the *dynamical-completion* of the measure system $(X, \mathcal{M})$ is the measure system $(X, \mathcal{M}^\star)$.

2.3.5 Properties of measure-complete systems

With the measure-completions of dynamical systems well defined, we now move to prove the properties of the completion. Dynamical-completions of measure systems are largely uninteresting in the study of dynamical systems, so we do not give a concise statement of their properties. In any case, many of these properties require that the measure system arise from a dynamical system, in order to speak of the ergodic measures, which are rather simply characterized as the extreme points of a dynamically-complete measure system. As a result, we only focus on giving properties of the measure-completion, which is the primary object of interest for this chapter anyways.

Proposition 2.3.30. *Let $(X, \mathcal{T})$ be a dynamical system, and $(X, \mathcal{T}^*)$ its measure completion. Then*

(a) $\mathcal{T}^*$ *contains the identity map, so is nonempty,*

(b) $\mathcal{T}^*$ *is closed under composition, so is a monoid,*

(c) If $T \in \mathcal{T}^*$ *is a bijection, then* $T^{-1} \in \mathcal{T}^*$*,*

(d) $\mathcal{T}^*$ *is stable under wobbling,*

(e) $\mathcal{T}^*$ *is (topologically) closed in* $\mathcal{B}(X)$*,*

(f) $\mathcal{T}^*$ *is closed under taking pointwise limits of sequences,*

(g) $\mathcal{I}(\mathcal{T}^*) = \mathcal{I}(\mathcal{T})$*, and*

(h) $\mathcal{E}(\mathcal{T}^*) = \mathcal{E}(\mathcal{T})$*.*

Proof. By definition, we have that $\mathcal{T}^* = \mathcal{F}(\mathcal{I}(\mathcal{T}))$, and therefore $\mathcal{T}^* = \mathcal{F}(\mathcal{M})$ for some measure system $(X, \mathcal{M})$. By Theorem 2.3.22, the first five properties hold. For (f), if $\{T_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathcal{T}^*$ which converges to some function T, since each

$T_n \in \mathcal{B}(X)$ is measurable, so is T, and therefore $T \in \mathcal{B}(X)$. Then, by Lemma 2.2.15, we have that $\{T_n\}_{n\in\mathbb{N}}$ converges to T in $\mathcal{B}(X)$, and since $\mathcal{T}^*$ is closed, it must be that $T \in \mathcal{T}^*$. Property (g) is given by Lemma 2.3.25. For (h), we have by (g) that $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}^*)$, and by Proposition 2.3.15, the extreme points of $\mathcal{I}(\mathcal{T})$ are exactly $\mathcal{E}(\mathcal{T})$, and the extreme points of $\mathcal{I}(\mathcal{T}^*)$ are exactly $\mathcal{E}(\mathcal{T}^*)$. But $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}^*)$, and the extreme points of this set is not dependent on $\mathcal{T}$, and so we must have that $\mathcal{E}(\mathcal{T}) = \mathcal{E}(\mathcal{T}^*)$. □

Beyond these core properties of the completion of a dynamical system, the invariants sets for a dynamical and its completion behave well in relation to each other.

Proposition 2.3.31. *Let $(X, \mathcal{T})$ be a dynamical system. Then*

(a) for $\mu \in \mathcal{E}(\mathcal{T})$, we have $\mathscr{I}(\mu, \mathcal{T}) = \mathscr{I}(\mu, \mathcal{T}^*)$,

(b) for $\mu \in \mathcal{I}(\mathcal{T})$, we have that $\mu \in \mathcal{E}(\mathcal{T})$ if and only if $\mathscr{I}(\mu, \mathcal{T}^*) \subset \mathscr{T}(\mu)$, and

(c) for every $\mathcal{M} \subset \mathcal{E}(\mathcal{T})$, we have $\mathscr{I}(\mathcal{M}, \mathcal{T}) = \mathscr{I}(\mathcal{M}, \mathcal{T}^*)$.

Proof. First, for (a), let $\mu \in \mathcal{E}(\mathcal{T})$, and since $\mathcal{E}(\mathcal{T}) = \mathcal{E}(\mathcal{T}^*)$ by Proposition 2.3.30, we have $\mu \in \mathcal{E}(\mathcal{T}^*)$ as well. As a result, we have by Lemma 2.3.14 applied to both the systems $(X, \mathcal{T})$ and $(X, \mathcal{T}^*)$ that

$$\mathscr{I}(\mu, \mathcal{T}) = \mathscr{T}(\mu) = \mathscr{I}(\mu, \mathcal{T}^*)$$

as desired.

Now, for (b), let $\mu \in \mathcal{E}(\mathcal{T})$. Then by definition we have that $\mathscr{I}(\mu, \mathcal{T}) \subset \mathscr{T}(\mu)$, and by (a), we have that $\mathscr{I}(\mu, \mathcal{T}^*) \subset \mathscr{T}(\mu)$. Now suppose that $\mu \in \mathcal{I}(\mathcal{T})$ and that $\mathscr{I}(\mu, \mathcal{T}^*) \subset \mathscr{T}(\mu)$. By Proposition 2.3.30, we have that $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}^*)$, and therefore by definition we have that $\mu \in \mathcal{E}(\mathcal{T}^*)$. But by Proposition 2.3.30, we have $\mathcal{E}(\mathcal{T}^*) = \mathcal{E}(\mathcal{T})$, so it follows that $\mu \in \mathcal{E}(\mathcal{T})$, which proves the desired result.

Finally for (c), by (a) we have

$$\mathscr{I}(\mathcal{M},\mathcal{T}) = \bigcap_{\mu\in\mathcal{M}} \mathscr{I}(\mu,\mathcal{T}) = \bigcap_{\mu\in\mathcal{M}} \mathscr{I}(\mu,\mathcal{T}^*) = \mathscr{I}(\mathcal{M},\mathcal{T}^*).$$

□

In the case that we assume $\mathcal{I}(\mathcal{T})$ to be closed, we can strengthen this result even further.

Proposition 2.3.32. *Let $(X,\mathcal{T})$ be a dynamical system, and suppose its associated measure system $(X,\mathcal{I}(\mathcal{T}))$ is closed. Then*

$$\mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}^*) = \mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}) = \mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}^*) = \mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}) = \mathscr{T}(\mathcal{E}(\mathcal{T})).$$

Proof. First, we have by Lemma 2.3.14 that $\mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}) = \mathscr{T}(\mathcal{E}(\mathcal{T}))$, and by Proposition 2.3.31 that $\mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}^*) = \mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T})$. Then, note we have $\mathcal{E}(\mathcal{T}) \subset \mathcal{I}(\mathcal{T})$ by definition and that $\mathcal{T} \subset \mathcal{T}^*$ by Theorem 2.3.26, and so Theorem 2.3.11 gives that

$$\mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}^*) \subset \mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}) \subset \mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}) = \mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}^*). \tag{2.1}$$

Now, let $E \in \mathscr{I}(\mathcal{E}(\mathcal{T}),\mathcal{T}^*)$, and let $\mu \in \mathcal{I}(\mathcal{T})$ and $T \in \mathcal{T}^*$. By Proposition 2.3.16 there exists $m \in \mathcal{P}^2(X)$ with $m(\mathcal{E}(\mathcal{T})) = 1$ and for which $\beta(m) = \mu$. Then by Proposition 2.1.14, we have

$$\mu(E\triangle T^{-1}(E)) = [\beta(m)](E\triangle T^{-1}(E)) = \int_{\mathcal{P}X} \nu(E\triangle T^{-1}(E))\, m(\mathrm{d}\nu).$$

Now, since $m(\mathcal{E}(\mathcal{T})) = 1$, we have that

$$\int_{\mathcal{P}X} \nu(E\triangle T^{-1}(E))\, m(\mathrm{d}\nu) = \int_{\mathcal{E}(\mathcal{T})} \nu(E\triangle T^{-1}(E))\, m(\mathrm{d}\nu),$$

and for each $\nu \in \mathcal{E}(\mathcal{T})$, we have that $E \in \mathscr{I}(\mathcal{E}(\mathcal{T}), \mathcal{T}^*) \subset \mathscr{I}(\nu, \mathcal{T}^*)$, and therefore $\nu(E \triangle T^{-1}(E)) = 0$. As such, we have

$$\mu(E \triangle T^{-1}(E)) = \int_{\mathcal{E}(\mathcal{T})} \nu(E \triangle T^{-1}(E))\, m(\mathrm{d}\nu) = \int_{\mathcal{E}(\mathcal{T})} 0 \,\mathrm{d}m = 0,$$

and therefore $E \in \mathscr{I}(\mu, T)$. As this holds for every $\mu \in \mathcal{I}(\mathcal{T})$ and $T \in \mathcal{T}^*$, we have that $E \in \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$. Finally, as $E \in \mathscr{I}(\mathcal{E}(\mathcal{T}), \mathcal{T}^*)$ was arbitrary, this shows that $\mathscr{I}(\mathcal{E}(\mathcal{T}), \mathcal{T}^*) \subset \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$, and thus the containments in (2.1) must be equalities, which is the desired result. □

This proposition is surprising in that it says that for a set $E \in \mathscr{A}_X$, the property of being $\mathcal{E}(\mathcal{T})$-trivial, being $(\mathcal{E}(\mathcal{T}), \mathcal{T})$-invariant, being $(\mathcal{E}(\mathcal{T}^*), \mathcal{T}^*)$-invariant, being $(\mathcal{I}(\mathcal{T}), \mathcal{T})$-invariant, being $(\mathcal{I}(\mathcal{T}^*), \mathcal{T}^*)$-invariant, and being $\mathcal{E}(\mathcal{T})$-trivial are all equivalent, so as long as the set of invariant measures for the system is closed. This last one is rather interesting, as it means if we know $\mathcal{M}$ is the closed set of invariant measures for some dynamical system, the collection of trivial sets for the extreme points of $\mathcal{M}$ must be the set of $(\mathcal{M}, \mathcal{T})$-invariant sets for any collection $\mathcal{T}$ of transformations for which $\mathcal{M} = \mathcal{I}(\mathcal{T})$. In this sense, the exact choice of $\mathcal{T}$ only matters in identifying $\mathcal{I}(\mathcal{T})$, and it is possible to deduce what sets must be invariant for this system, when all invariant measures (or just the ergodic measures) are taken into account. Additionally, because we have $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T})$ this proposition also gives that $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$, which is also surprising, as it shows that any $\mathcal{T}$-invariant set is always $(\mathcal{I}(\mathcal{T}^*), \mathcal{T}^*)$-invariant.

2.4 Birkhoff Systems and Dynamical Independence

A rather important result in Ergodic Theory is the Birkhoff ergodic theorem, though at least in its original form, only applies to a single transformation T (although it also applies to more general measure spaces). Other pointwise ergodic theorems have been shown to hold in other contexts, however none hold generally for the dynamical systems presented here. As such, we develop a notion of a Birkhoff System, for which a version of the pointwise ergodic theorem holds. Interestingly, this notion of a Birkhoff System transfers readily to the completion of a system. Furthermore, we shall demonstrate that many classically studied systems (where there are known pointwise ergodic theorems) are Birkhoff Systems. Then, for Birkhoff systems we will be able to define a notion of dynamical independence of two sets within a system, and ultimately show that for an invariant measure, the notion of ergodicity and the property that dynamically independent sets are always probabilistically independent, are equivalent. We begin by defining Birkhoff systems.

2.4.1 Birkhoff systems

In order to define the notion of a dynamical system, we first give a definition of the usual notion of a conditional expectation.

Definition 2.4.1. Let X be a space, let $f \in B(X)$, let $\mu \in \mathcal{P}X$, and let $\mathscr{B} \subset \mathscr{A}_X$ be a σ-algebra. Then $g \in B(X)$ is a version of the $(\mu, \mathscr{B})$-*expectation of* f if

(a) g is $\mathscr{B}$-measurable, and

(b) for every $E \in \mathscr{B}$, it holds that

$$\int_E g \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu.$$

We denote that g is a version of this expectation of f with the notation $g = \mathbb{E}_\mu[f|\mathscr{B}]$.

Next, we have the fairly standard fact that two versions of the conditional function are always equal almost everywhere.

Lemma 2.4.2. *Let X be a space, let $f \in B(X)$, let $\mu \in \mathcal{P}X$, let $\mathscr{B} \subset \mathscr{A}_X$ be a σ-algebra, and let $g = \mathbb{E}_\mu[f|\mathscr{B}]$. Then for $h \in B(X)$ that is $\mathscr{B}$-measurable, $h = \mathbb{E}_\mu[f|\mathscr{B}]$ if and only if there exists a set $E \in \mathscr{N}(\mu)$ such that for all $x \in X \setminus E$, $h(x) = g(x)$.*

Proof. First, suppose that $h = \mathbb{E}_\mu[f|\mathscr{B}]$, and let $E_1 = (g-h)^{-1}((0,\infty))$ and $E_2 = (h-g)^{-1}((0,\infty))$, and note that because both g and h are $\mathscr{B}$-measurable, so are $g-h$ and $h-g$, and $(0,\infty) \subset \mathbb{R}$ is a measurable subset of $\mathbb{R}$, so we have that $E_1, E_2 \in \mathscr{B}$. As such, since g and h are versions of $\mathbb{E}_\mu[f|\mathscr{B}]$, we have

$$\int_{E_1} g \, \mathrm{d}\mu = \int_{E_1} f \, \mathrm{d}\mu = \int_{E_1} h \, \mathrm{d}\mu,$$

and therefore

$$0 = \int_{E_1} g \, \mathrm{d}\mu - \int_{E_1} h \, \mathrm{d}\mu = \int_{E_1} g - h \, \mathrm{d}\mu.$$

Since $g - h > 0$ on E_1, it must be that $\mu(E_1) = 0$. Similarly, we have

$$0 = \int_{E_2} h \, \mathrm{d}\mu - \int_{E_2} g \, \mathrm{d}\mu = \int_{E_2} h - g \, \mathrm{d}\mu,$$

and since $h - g > 0$ on E_2, it must be that $\mu(E_2) = 0$. Therefore $E = \{x \in X : g(x) \neq h(x)\} = E_1 \cup E_2$ satisfies $\mu(E) = 0$, which gives that $E \in \mathscr{N}(\mu)$.

Now suppose that there is some $E \in \mathscr{N}(\mu)$ such that for all $x \in X \setminus E$, $h(x) = g(x)$. Now, for $F \in \mathscr{B}$, we have

$$\int_F f \, \mathrm{d}\mu = \int_F g \, \mathrm{d}\mu = \int_{F \setminus E} g \, \mathrm{d}\mu = \int_{F \setminus E} h \, \mathrm{d}\mu = \int_F h \, \mathrm{d}\mu,$$

and since h is $\mathscr{B}$-measurable, we have $h = \mathbb{E}_\mu[f|\mathscr{B}]$. □

We now introduce a new notion of conditional expectation with respect to a dynamical system. This is heavily related to the classical notion of conditional expectation, but whose core property involves multiple measures, as opposed to just a single one.

Definition 2.4.3. Let $(X, \mathcal{T})$ be a dynamical system and $f \in B(X)$. Then $g \in B(X)$ is a version of the $(X, \mathcal{T})$*-expectation of* f, referred to as a *dynamical expectation of* f, if

(a) g is $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$-measurable, and

(b) for every $\mu \in \mathcal{I}(\mathcal{T})$ and $E \in \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$, it holds that

$$\int_E g \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu.$$

We denote that g is a version of this expectation of f with the notation $g = \mathbb{E}_{\mathcal{T}}[f]$. Additionally, note that it follows immediately from the definitions that $g = \mathbb{E}_{\mathcal{T}}[f]$ if and only if $g = \mathbb{E}_{\mu}[f | \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)]$ for every $\mu \in \mathcal{I}(\mathcal{T})$.

Unlike the conditional expectation, whose existence is guaranteed by the existence of the Radon-Nikodym derivative [8], it is not necessarily the case that the dynamical expectation of any function should exist. As such, when we say "let $g = \mathbb{E}_{\mathcal{T}}[f]$", we are also making the assertion that such a function g exists. In any case, it follows immediately from this definition that for $f \in B(X)$, the $(X, \mathcal{T})$-dynamical expectation of a function f is the same as the $(X, \mathcal{T}^*)$-dynamical expectation of f assuming that either of the two exists.

Proposition 2.4.4. *Let* $(X, \mathcal{T})$ *be a dynamical system and* $f \in B(X)$*. Then* $g = \mathbb{E}_{\mathcal{T}}[f]$ *if and only if* $g = \mathbb{E}_{\mathcal{T}^*}[f]$*.*

Proof. First, if $g = \mathbb{E}_{\mathcal{T}}[f]$, then g is $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$-measurable. By Proposition 2.3.30, we have that $\mathcal{I}(\mathcal{T}^*) = \mathcal{I}(\mathcal{T})$, and Theorem 2.3.26 that $(\mathcal{T}^*)^* = \mathcal{T}^*$, and so $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*) = \mathscr{I}(\mathcal{I}(\mathcal{T}^*), (\mathcal{T}^*)^*)$, which means g is $\mathscr{I}(\mathcal{I}(\mathcal{T}^*), (\mathcal{T}^*)^*)$-measurable.

Furthermore, for $\mu \in \mathcal{I}(\mathcal{T}^*) = \mathcal{I}(\mathcal{T})$ and $E \in \mathscr{I}(\mathcal{I}(\mathcal{T}^*), (\mathcal{T}^*)^*) = \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$, we have

$$\int_E g \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu,$$

and therefore $g = \mathbb{E}_{\mathcal{T}^*}[f]$.

If $g = \mathbb{E}_{\mathcal{T}^*}[f]$, then we have $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*) = \mathscr{I}(\mathcal{I}(\mathcal{T}^*), (\mathcal{T}^*)^*)$ and $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}^*)$ as above, which gives that $g = \mathbb{E}_{\mathcal{T}}[f]$. □

Next, we have that the dynamical expectation of f is always invariant almost everywhere for every transformation in the completion.

Lemma 2.4.5. *Let $(X, \mathcal{T})$ be a dynamical system and $f \in B(X)$, and suppose $g = \mathbb{E}_{\mathcal{T}}[f]$. Then for any $T \in \mathcal{T}^*$, we have that $g \circ T = \mathbb{E}_{\mathcal{T}}[f]$ and that there exists $E \in \mathscr{N}(\mathcal{I}(\mathcal{T}))$ such that for $x \in X \setminus E$, we have $g(T(x)) = g(x)$.*

Proof. Since g is $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$-measurable, and $T \in \mathcal{B}(X)$, we have that $g \circ T$ is $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$-measurable. Additionally, for each $\mu \in \mathcal{I}(\mathcal{T})$, we have that $\mathcal{P}T(\mu) = \mu$ by definition, and also for $E \in \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$, it is automatically the case that $\mu(E \triangle T^{-1}(E)) = 0$. Using these facts and a change of variables, we have

$$\int_E g \circ T \, \mathrm{d}\mu = \int_{T^{-1}(E)} g \circ T \, \mathrm{d}\mu = \int_E g \, \mathrm{d}\mathcal{P}T(\mu) = \int_E g \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu.$$

Since $E \in \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$ and $\mu \in \mathcal{I}(\mathcal{T})$ were arbitrary, we have shown that $g \circ T = \mathcal{E}_{\mathcal{T}}[f]$.

Finally, note that $g = \mathcal{E}_{\mathcal{T}}[f]$ occurs if and only if $g = \mathbb{E}_\mu[f | \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)]$ for every $\mu \in \mathcal{I}(\mathcal{T})$, and similarly $g \circ T = \mathcal{E}_{\mathcal{T}}[f]$ occurs if and only if $g \circ T = \mathbb{E}_\mu[f | \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)]$ for every $\mu \in \mathcal{I}(\mathcal{T})$. As such, for each $\mu \in \mathcal{I}(\mathcal{T})$, we have by Lemma 2.4.2 that there exists $E_\mu \in \mathscr{N}(\mu)$ such that for $x \in X \setminus E_\mu$, we have $g(T(x)) = g(x)$. Let $E = \{x \in X : g(T(x)) \neq g(x)\}$, which is a measurable set since both $g(x)$ and $g(T(x))$ are measurable. Additionally, we have that $E \subset E_\mu$ for every $\mu \in \mathcal{I}(\mathcal{T})$, and therefore by monotonicity, we have that $\mu(E) \leq \mu(E_\mu) = 0$, which gives that $E \in \mathscr{N}(\mu)$. As this holds for every $\mu \in \mathcal{I}(\mathcal{T})$, we have that $E \in \mathscr{N}(\mathcal{I}(\mathcal{T}))$, and clearly for every $x \in X \setminus E$, we have $g(T(x)) = g(x)$, as desired. □

Furthermore, we have that the $(X, \mathcal{T})$-dynamical expectation of f is constant almost everywhere for every $\mu \in \mathcal{E}(\mathcal{T})$, as should be expected from existing results in Ergodic Theory (Theorem 1.6 [52]).

Proposition 2.4.6. *Let $(X, \mathcal{T})$ be a dynamical system, let $f \in B(X)$, let $\mu \in \mathcal{E}(\mathcal{T})$, and let $g = \mathbb{E}_{\mathcal{T}}[f]$. Then there exists $E \in \mathscr{N}(\mu)$ such that for $x \in X \setminus E$, we have*

$$g(x) = \int_X f \, \mathrm{d}\mu.$$

Proof. For $z \in \mathbb{Z}$ and $k \in \mathbb{N}$, define

$$E_z^k = \left[\frac{z}{2^k}, \frac{z+1}{2^k}\right),$$

and note that for each $k \in \mathbb{N}$, we have $\mathbb{R} = \bigsqcup_{z \in \mathbb{Z}} E_z^k$. Furthermore, each E_z^k is nonempty and measurable. As such, we have $g^{-1}(E_z^k) \in \mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*)$. By Theorem 2.3.11 we have that $\mathscr{I}(\mathcal{I}(\mathcal{T}), \mathcal{T}^*) \subset \mathscr{I}(\mu, \mathcal{T})$, and by Lemma 2.3.14, since $\mu \in \mathcal{E}(\mathcal{T})$, we have that $\mathscr{I}(\mu, \mathcal{T}) = \mathscr{T}(\mu)$, and therefore $g^{-1}(E_z^k) \in \mathscr{T}(\mu)$, so $\mu(g^{-1}(E_z^k)) \in \{0, 1\}$.

For a given $k \in \mathbb{N}$, we have by countable disjoint additivity that

$$1 = \mu(X) = \mu\big(g^{-1}(\mathbb{R})\big) = \mu\left(\bigsqcup_{z \in \mathbb{Z}} g^{-1}(E_z^k)\right) = \sum_{z \in \mathbb{Z}} \mu(g^{-1}(E_z^k)).$$

Since $\mu(g^{-1}(E_z^k)) \in \{0, 1\}$ for each $z \in \mathbb{Z}$, there exists exactly one $z_k \in \mathbb{Z}$ for which

$$\mu(g^{-1}(E_{z_k}^k)) = 1.$$

Let $F_k = \overline{E^k_{z_k}}$ so that each F_k is compact and $\mu(g^{-1}(F_k)) = 1$. It is also clear that we must have $F_{k+1} \subset F_k$ for every k, otherwise F_{k+1} and F_k would be disjoint (based on their definition) making $\mu(X) > 1$. Also, $\operatorname{diam}(F_k) = \sup\{|x - y| : x, y \in F_k\} = \frac{1}{2^k}$, and so $\{F_k\}$ is a contracting sequence of nonempty compact subsets of $\mathbb{R}$. By a version of the Cantor Intersection Theorem (see Section 9.4 of [47]), there exists $r \in \mathbb{R}$ such that $\bigcap_{k\in\mathbb{N}} F_k = \{r\}$. Additionally, by the continuity of μ, it must be that

$$\mu(g^{-1}(\{r\})) = \mu\left(g^{-1}\left(\bigcap_{k\in\mathbb{N}} F_k\right)\right) = \mu\left(\bigcap_{k\in\mathbb{N}} g^{-1}(F_k)\right) = \lim_{k\to\infty} \mu(g^{-1}(F_k)) = 1.$$

Let $F = g^{-1}(\{r\})$, and since $\mu(F) = 1$, we have $\mu(X \setminus F) = 0$, and therefore using that $g = \mathbb{E}_{\mathcal{T}}[f]$, we have for $x \in F$ that

$$g(x) = r = \int_F g \, \mathrm{d}\mu = \int_X g \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu,$$

and so taking $E = X \setminus F$ gives the desired result. □

Next, we define the usual notion of an ergodic average, but for an arbitrary finite subset of transformations.

Definition 2.4.7. Let X be a space, and let $x \in X$ and $f \in B(X)$. For a set $F \in \mathscr{F}(\mathcal{B}(X))$, define

$$A_F[f] = \frac{1}{|F|} \sum_{T\in F} f \circ T$$

to be the *F-average of f*.

With the ergodic theorem, we are interested in looking at the behavior of ergodic averages along some sequence of finite sets of transformations. For simplicity, we give a concise term for such sequences.

Definition 2.4.8. Let I be a set. An *averaging sequence in I* is a sequence $\mathcal{A} = \{\mathcal{A}_n\}_{n\in\mathbb{N}}$ of finite subsets of I, so that $\mathcal{A} \in \mathscr{F}(I)^{\mathbb{N}}$.

With these notions defined, we may give the definition of a Birkhoff System.

Definition 2.4.9. Let $(X, \mathcal{T})$ be a dynamical system. An averaging sequence $\mathcal{A}$ in $\mathcal{T}^*$ is called a *Birkhoff sequence for* $(X, \mathcal{T})$ if for every $f \in B(X)$, there exists $g = \mathbb{E}_{\mathcal{T}}[f](x)$ such that for some $E \in \mathscr{N}(\mathcal{I}(\mathcal{T}))$ we have for $x \in X \setminus E$ that

$$\lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = g(x).$$

We say that $(X, \mathcal{T})$ is a *Birkhoff system* if it has a Birkhoff sequence.

The structure of this statement follows any of the well known pointwise ergodic theorems (Theorem 1.14 [52], Theorem 1.2 [37], Theorem 2.3 [49], Theorem 3.41 [23]), however there are a few notable differences. First, the sequence of subsets along which we take the ergodic averages is only required to be contained within the completion of $\mathcal{T}$, rather than being directly related to it. Additionally, rather than the limit being a conditional expectation with respect to a single measure, or just some invariant function, we are taking the dynamical expectation. As previously discussed, it is not clear that the dynamical expectation of a system should always exist, so a necessary condition for a system to be Birkhoff is for the dynamical expectation of every function to exist. As with the dynamical expectation, it is almost immediate from the definition that if a system is Birkhoff, so is its completion.

Proposition 2.4.10. *Let* $(X, \mathcal{T})$ *be a dynamical system. Then* $\mathcal{A}$ *is a Birkhoff sequence for* $(X, \mathcal{T})$ *if and only if* $\mathcal{A}$ *is a Birkhoff sequence for* $(X, \mathcal{T}^*)$*. As such,* $(X, \mathcal{T})$ *is a Birkhoff system if and only if* $(X, \mathcal{T}^*)$ *is a Birkhoff system.*

Proof. Let $\mathcal{A}$ be a Birkhoff sequence for $(X, \mathcal{T})$. Then for $f \in B(X)$, there exists $g = \mathbb{E}_{\mathcal{T}}[f]$ such that for some $E \in \mathscr{N}(\mathcal{I}(\mathcal{T}))$ we have for $x \in X \setminus E$ that

$$\lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = g(x).$$

Then $\mathcal{A}$ is an averaging sequence in $\mathcal{T}^*$, and by Proposition 2.4.4, we have $g = \mathbb{E}_{\mathcal{T}^*}[f]$, and by Proposition 2.3.30 we have $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}^*)$, and so there exists $E \in \mathscr{N}(\mathcal{I}(\mathcal{T})) = \mathscr{N}(\mathcal{I}(\mathcal{T}^*))$ such that for $x \in X \setminus E$, we have

$$\lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = g(x),$$

which shows that $\mathcal{A}$ is a Birkhoff sequence for $(X, \mathcal{T}^*)$.

Using essentially the same argument, if $\mathcal{A}$ is a Birkhoff sequence for $(X, \mathcal{T}^*)$, it will also be a Birkhoff sequence for $(X, \mathcal{T})$. □

While this notion of a Birkhoff sequence and Birkhoff systems plays nicely with a dynamical system and its measure-completion, it would be rather useless unless well known systems were in fact Birkhoff systems. The two following theorems prove that many of the classically studied types of dynamical systems are in fact Birkhoff systems. We begin with doing so for classical systems consisting of a single transformation.

Theorem 2.4.11. *Let (X, T) be a dynamical system consisting of a single transformation, where $\mathcal{I}(T)$ is closed. Then the sequence $\mathcal{A}$ defined by*

$$\mathcal{A}_n = \{I, T, T^2, \dots, T^{n-1}\}$$

is a Birkhoff sequence for (X, T), and therefore (X, T) is a Birkhoff system.

Proof. First, let

$$g(x) = \limsup_{n\to\infty} A_{\mathcal{A}_n}[f](x),$$

which is a limsup of linear combinations of functions in $B(X)$, and is therefore in $B(X)$. We now show that $g = g \circ T$. Indeed, for $x \in X$, we have

$$g(T(x)) = \limsup_{n\to\infty} A_{\mathcal{A}_n}[f](T(x)),$$

but at the same time,

$$\begin{aligned}
&\left|\left(\limsup_{n\to\infty} A_{\mathcal{A}_n}[f](T(x))\right) - g(x)\right| \\
&\qquad \leq \left|\left(\limsup_{n\to\infty} A_{\mathcal{A}_n}[f](T(x)) - A_{\mathcal{A}_n}[f](x)\right) + \left(\limsup_{n\to\infty} A_{\mathcal{A}_n}[f](x)\right) - g(x)\right| \\
&\qquad \leq \limsup_{n\to\infty} \left|\frac{1}{n}\sum_{k=0}^{n-1} f(T^k(T(x))) - \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))\right| \\
&\qquad = \limsup_{n\to\infty} \left|\frac{1}{n}(f(T^n(x)) - f(x))\right| \\
&\qquad \leq \limsup_{n\to\infty} \frac{2\|f\|}{n} = 0.
\end{aligned}$$

As such, we must have that $g(T(x)) = g(x)$ for every $x \in X$, and so $g = g \circ T$. Then, for every measurable $E \subset \mathbb{R}$, we have

$$g^{-1}(E) = (g \circ T)^{-1}(E) = T^{-1}(g^{-1}(E)),$$

which gives that g is $\mathscr{I}(T)$-measurable. Next, for $\mu \in \mathcal{I}(T)$, we have by Theorem 2.3 of [49] that for $h = \mathbb{E}_\mu[f|\mathscr{I}(T)]$, there exists $E \in \mathscr{N}(\mathcal{I}(T))$ such that for each $x \in X \setminus E$,

$$\lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = h(x).$$

In order for this to occur, the limit on the left hand side must exist for $x \in X \setminus E$, and therefore we have for $x \in X \setminus E$ that

$$g(x) = \limsup_{n\to\infty} A_{\mathcal{A}_n}[f](x) = \lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = h(x),$$

and since g is $\mathscr{I}(T)$-measurable, we have by Proposition 2.4.2 that $g = \mathbb{E}_\mu[f|\mathscr{I}(T)]$. By Theorem 2.3.11(f), we have $\mathscr{I}(T) \subset \mathscr{I}(\mathcal{I}(T), T)$, and therefore g is $\mathscr{I}(\mathcal{I}(T), T)$-measurable. Then, since $\mathcal{I}(T)$ is closed, Proposition 2.3.32 gives that $\mathscr{I}(\mathcal{I}(T), T) = \mathscr{I}(\mathcal{I}(T), \{T\}^*)$, and so g is $\mathscr{I}(\mathcal{I}(T), \{T\}^*)$-measurable. Now, let $E \in \mathscr{I}(\mathcal{I}(T), \{T\}^*)$, and define

$$E_\infty = \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}(E).$$

From the proof of (i) $\implies$ (ii) of Theorem 1.5 of [52], we have that $T^{-1}(E_\infty) = E_\infty$, and thus $E_\infty \in \mathscr{I}(T)$. Furthermore, for every $\mu \in \mathcal{I}(T)$, we have that $\mu(E \triangle E_\infty) = 0$, and therefore using that $g = \mathbb{E}_\mu[f|\mathscr{I}(T)]$, we have

$$\int_E g \,\mathrm{d}\mu = \int_{E_\infty} g \,\mathrm{d}\mu = \int_{E_\infty} f \,\mathrm{d}\mu = \int_E f \,\mathrm{d}\mu,$$

and therefore $g = \mathbb{E}_T[f]$.

Finally, for every $\mu \in \mathcal{I}(T)$, we have by Theorem 2.3 of [49] that with $g = \mathbb{E}_\mu[f|\mathscr{I}(T)]$, there exists $E_\mu \in \mathscr{N}(\mu)$ such that for $x \in X \setminus E_\mu$,

$$\lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = g(x).$$

Let F be the set of all x for which this limit exists and equals $g(x)$. This is clearly a measurable set, as each function in the limit is measurable, and g is measurable, and so $E = X \setminus F$ is also measurable. It must then be that for each $\mu \in \mathcal{I}(T)$ that $X \setminus E_\mu \subset F$, and therefore $E = X \setminus F \subset E_\mu$. This gives that $\mu(E) \leq \mu(E_\mu) = 0$, and so $E \in \mathscr{N}(\mu)$ for every $\mu \in \mathcal{I}(T)$, whence $E \in \mathscr{N}(\mathcal{I}(\mathcal{T}))$. Noting that $g = \mathbb{E}_T[f]$, we have shown that $\mathcal{A}$ is a Birkhoff sequence as desired. $\square$

Before proving that the next type of classical system is a Birkhoff system, let us recall the definition of a Følner sequence, and define the notion of a tempered Følner sequence.

Definition 2.4.12. Let G be a countable group. A sequence $\{F_n\} \subset \mathscr{F}(G)$ is a *Følner sequence* if for every $g \in G$ we have that

$$\lim_{n\to\infty} \frac{|gF_n \triangle F_n|}{|F_n|} = 0.$$

A group G has a Følner sequence if and only if it is amenable [33]. A Følner squence $\{F_n\}$ for G is said to be *tempered* if there exists some $C > 0$ such that for all $n \in \mathbb{N}$, we have

$$\left|\bigcup_{k<n} F_k^{-1} F_n\right| \leq C|F_n|.$$

Next, utilizing a version of Lindenstrauss's pointwise ergodic theorem for amenable groups [37], we can prove that any countable amenable group acting measurably on a compact metrizable space (for which the set of invariant measures is closed) is a Birkhoff system.

Theorem 2.4.13. *Let $(X, \mathcal{G})$ be a dynamical system where $\mathcal{G}$ is a countable amenable group and $\mathcal{I}(\mathcal{G})$ is closed. Then any tempered Følner sequence $\mathcal{A}_\mathbb{N} = \{\mathcal{A}_n\}_{n\in\mathbb{N}}$ in $\mathcal{G}$ is a Birkhoff sequence for $(X, \mathcal{G})$, so $(X, \mathcal{G})$ is a Birkhoff system.*

Proof. First, by Theorem 4.10 of [33], there exists a two-sided Følner sequence in $\mathcal{G}$, and by Proposition 1.4 of [37], this Følner sequence has a tempered subsequence $\mathcal{D} = \{\mathcal{D}_n\}_{n\in\mathbb{N}}$. Let

$$g(x) = \limsup_{n\to\infty} A_{\mathcal{D}_n}[f](x),$$

which is a limsup of linear combinations of functions in $B(X)$, and is therefore in $B(X)$. We now show that for each $T \in \mathcal{G}$, that $g = g \circ T$. Indeed, for $x \in X$ we have

$$g(T(x)) = \limsup_{n\to\infty} A_{\mathcal{D}_n}[f](T(x)),$$

but at the same time,

$$\begin{aligned}
&\left|\left(\limsup_{n\to\infty} A_{\mathcal{D}_n}[f](T(x))\right) - g(x)\right| \\
&\qquad \le \left|\left(\limsup_{n\to\infty} A_{\mathcal{D}_n}[f](T(x)) - A_{\mathcal{D}_n}[f](x)\right) + \left(\limsup_{n\to\infty} A_{\mathcal{D}_n}[f](x)\right) - g(x)\right| \\
&\qquad \le \limsup_{n\to\infty}\left|\frac{1}{|\mathcal{D}_n|}\sum_{S\in\mathcal{D}_n} f(S(T(x))) - \frac{1}{|\mathcal{D}_n|}\sum_{S\in\mathcal{D}_n} f(S(x))\right| \\
&\qquad = \limsup_{n\to\infty}\left|\frac{1}{|\mathcal{D}_n|}\sum_{S\in\mathcal{D}_n\circ T\setminus\mathcal{D}_n} f(S(x)) - \frac{1}{|\mathcal{D}_n|}\sum_{S\in\mathcal{D}_n\setminus\mathcal{D}_n\circ T} f(S(x))\right| \\
&\qquad \le \limsup_{n\to\infty}\frac{|\mathcal{D}_n\circ T\triangle\mathcal{D}_n|}{|\mathcal{D}_n|}\|f\| = 0,
\end{aligned}$$

which is true because $\mathcal{D}$ is a right Følner sequence for $\mathcal{T}$. As such, we have that $g(T(x)) = g(x)$ for every $x \in X$, and so $g = g \circ T$. Then, for every measurable $E \subset \mathbb{R}$, we have

$$g^{-1}(E) = (g \circ T)^{-1}(E) = T^{-1}(g^{-1}(E)),$$

which gives that g is $\mathscr{I}(\mathcal{G})$-measurable. Next, for $\mu \in \mathcal{I}(\mathcal{G})$, we have by Theorem 2.3.11(f) that $\mathscr{I}(\mathcal{G}) \subset \mathscr{I}(\mu,\mathcal{G})$, and therefore g is $\mathscr{I}(\mu,\mathcal{G})$-measurable. We also have by Theorem 4.28 of [33] that for $h = \mathbb{E}_\mu[f|\mathscr{I}(\mu,\mathcal{G})]$ there exists $E \in \mathscr{N}(\mu)$ such that for each $x \in X \setminus E$,

$$\lim_{n\to\infty} A_{\mathcal{D}_n}[f](x) = h(x).$$

In order for this to occur, the limit on the left hand side must exist for $x \in X \setminus E$, and therefore we have for $x \in X \setminus E$ that

$$g(x) = \limsup_{n\to\infty} A_{\mathcal{D}_n}[f](x) = \lim_{n\to\infty} A_{\mathcal{D}_n}[f](x) = h(x),$$

and since g is $\mathscr{I}(\mu,\mathcal{G})$-measurable, we have by Lemma 2.4.2 that $g = \mathbb{E}_\mu[f|\mathscr{I}(\mu,\mathcal{G})]$. Since g is $\mathscr{I}(\mu,\mathcal{G})$-measurable for each $\mu \in \mathcal{I}(\mathcal{G})$, it follows that g is $\mathscr{I}(\mathcal{I}(\mathcal{G}),\mathcal{G})$-

measurable, and since $\mathcal{I}(\mathcal{G})$ is closed, we have by Proposition 2.3.32 that $\mathscr{I}(\mathcal{I}(\mathcal{G}),\mathcal{G}) = \mathscr{I}(\mathcal{I}(\mathcal{G}),\mathcal{G}^*)$, and therefore g is $\mathscr{I}(\mathcal{I}(\mathcal{G}),\mathcal{G}^*)$-measurable. Furthermore, for every $\mu \in \mathcal{I}(\mathcal{G})$ and $E \in \mathscr{I}(\mathcal{I}(\mathcal{G}),\mathcal{G}^*) \subset \mathscr{I}(\mu,\mathcal{G})$, since $g = \mathbb{E}_\mu[f|\mathscr{I}(\mu,\mathcal{G})]$, we have that

$$\int_E g \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu,$$

and therefore $g = \mathbb{E}_{\mathcal{G}}[f]$.

Finally, considering now an arbitrary tempered Følner sequence $\mathcal{A} = \{\mathcal{A}_n\}_{n\in\mathbb{N}}$, for $\mu \in \mathcal{I}(\mathcal{G})$, we have by Theorem 4.28 of [33] that for $g = \mathbb{E}_\mu[f|\mathscr{I}(\mu,\mathcal{G})]$ there exists $E_\mu \in \mathscr{N}(\mu)$ such that for each $x \in X \setminus E_\mu$,

$$\lim_{n\to\infty} A_{\mathcal{A}_n}[f](x) = g(x).$$

Let F be the set of all x for which this limit exists and equals $g(x)$. This is clearly a measurable set, as each function in the limit is measurable, and g is measurable, and so $E = X \setminus F$ is also measurable. It must then be that for each $\mu \in \mathcal{I}(\mathcal{G})$, $X \setminus E_\mu \subset F$, and therefore $E = X \setminus F \subset E_\mu$. This gives that $\mu(E) \leq \mu(E_\mu) = 0$, and so $E \in \mathscr{N}(\mu)$ for every $\mu \in \mathcal{I}(\mathcal{G})$, whence $E \in \mathscr{N}(\mathcal{I}(\mathcal{G}))$. Noting that $g = \mathbb{E}_{\mathcal{G}}[f]$, we have shown that $\mathcal{A}$ is a Birkhoff sequence as desired. $\square$

The requirement that $\mathcal{I}(\mathcal{T})$ be closed in the two previous Theorems is only used in one particular place in the argument, ultimately to show that $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}^*)$ through the fact that $\mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}) = \mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}^*)$ when $\mathcal{I}(\mathcal{T})$ is closed. While it is certainly not true that $\mathcal{I}(\mathcal{T})$ is always closed, it may always be the case that $\mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}) = \mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}^*)$, or even more likely that $\mathscr{I}(\mathcal{T}) \subset \mathscr{I}(\mathcal{I}(\mathcal{T}),\mathcal{T}^*)$ regardless of whether or not $\mathcal{I}(\mathcal{T})$ is closed, although a proof or disproof of either remains elusive to the author. In that case, the additional requirement that $\mathcal{I}(\mathcal{T})$ is closed could be dropped, leaving that all classically studied systems consisting of a single measurable transformation and measurable actions of countable amenable groups to all be Birkhoff systems.

Additionally, the structure of these two proofs is rather similar, and demonstrates that showing a system is Birkhoff follows almost entirely from existing statements of pointwise ergodic theorems for dynamical systems, giving credence to this definition of a Birkhoff system as a broad generalization to describe systems for which a version of the pointwise ergodic theorem holds. Theorem 2.4.11 highlights an important distinction about the definition of a Birkhoff sequence, which is that the sequence need not actually be contained solely within the original collection of transformations $\mathcal{T}$ in order to be Birkhoff, we only require it to be contained within $\mathcal{T}^*$. In this case, the Birkhoff sequence is very related to $\mathcal{T}$, however a particular use case of Birkhoff systems which appears in Section 2.5 demonstrates that a Birkhoff net may be very different from $\mathcal{T}$ (in particular, in the proof of Theorem 2.5.22). Furthermore, since a Birkhoff sequence $\mathcal{A}$ for $(X, \mathcal{T})$ is also a Birkhoff sequence for $(X, \mathcal{T}^*)$, Birkhoff sequence need not "exhaust" or in some sense encapsulate the entirety of the system, as is the case with taking iterates of T or Følner sequences for countable amenable groups.

We now prove the following fact about ergodic measures for Birkhoff systems.

Proposition 2.4.14. *Let $\mathcal{A}$ be a Birkhoff sequence for the Birkhoff system $(X, \mathcal{T})$, and let $\mu \in \mathcal{I}(\mathcal{T})$. Then $\mu \in \mathcal{E}(\mathcal{T})$ if and only if for every $A, B \in \mathscr{A}_X$,*

$$\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T\in\mathcal{A}_n} \mu(T^{-1}(A) \cap B) = \mu(A)\mu(B).$$

Proof. First, suppose that $\mu \in \mathcal{E}(\mathcal{T})$. For $\chi_A \in B(X)$, since $\mathcal{A}$ is a Birkhoff sequence, we have for $g = \mathbb{E}_{\mathcal{T}}[\chi_A]$ and some $E_1 \in \mathscr{N}(\mathcal{I}(\mathcal{T}))$ that for $x \in X \setminus E_1$,

$$\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T\in\mathcal{A}_n} \chi_A(T(x)) = g(x).$$

Additionally, since $\mu \in \mathcal{E}(\mathcal{T})$, by Proposition 2.4.6, there is $E_2 \in \mathscr{N}(\mu)$ so that for $x \in X \setminus E_2$,

$$g(x) = \int_X \chi_A \, \mathrm{d}\mu = \mu(A).$$

Also, $\chi_A(T(x)) = \chi_{T^{-1}(A))}(x)$. As such, with $E = E_1 \cup E_2$ where $\mu(E) = 0$, for $x \in X \setminus E$ we have

$$\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T \in \mathcal{A}_n} \chi_{T^{-1}(A)}(x) = \mu(A).$$

Now by the continuity of multiplication, we have for $x \in X \setminus E$ that

$$\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T \in \mathcal{A}_n} \chi_{T^{-1}(A)}(x)\chi_B(x) = \mu(A)\chi_B(x).$$

Note as well that $\chi_{T^{-1}(A)}(x)\chi_B(x) = \chi_{T^{-1}(A)\cap B}(x)$, and by the Dominated Covergence Theorem (as $\mu(E) = 0$), we have

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T \in \mathcal{A}_n} \mu(T^{-1}(A) \cap B) &= \lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T \in \mathcal{A}_n} \int_X \chi_{T^{-1}(A)\cap B} \, \mathrm{d}\mu \\ &= \int_X \mu(A)\chi_B \, \mathrm{d}\mu = \mu(A)\mu(B)\end{aligned}$$

as desired.

Finally, suppose that for every $A, B \in \mathscr{A}_X$, that

$$\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T \in \mathcal{A}_n} \mu(T^{-1}(A) \cap B) = \mu(A)\mu(B).$$

By Proposition 2.3.31, we have that $\mu \in \mathcal{E}(\mathcal{T})$ if and only if $\mathscr{I}(\mu, \mathcal{T}^*) \subset \mathscr{T}(\mu)$, so let $A \in \mathscr{I}(\mu, \mathcal{T}^*)$. Then for every $T \in \mathcal{T}^*$, we have $\mu(A \triangle T^{-1}A) = 0$, and therefore we also have that $\mu(A \setminus T^{-1}(A)) = 0$ by monotonicity, which gives that

$$\mu(T^{-1}(A) \cap A) = \mu(T^{-1}(A) \cap A) + \mu(A \setminus T^{-1}(A)) = \mu(A).$$

By assumption, we then have that

$$\mu(A)^2 = \lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T\in\mathcal{A}_n} \mu(T^{-1}(A)\cap A) = \lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T\in\mathcal{A}_n} \mu(A) = \mu(A),$$

since $\mathcal{A}_n \subset \mathcal{T}^*$ for every $n \in \mathbb{N}$. As such, we have $\mu(A)^2 = \mu(A)$, which means $\mu(A) \in \{0,1\}$. Therefore, $A \in \mathscr{T}(\mu)$, and since $A \in \mathscr{I}(\mu, \mathcal{T}^*)$ was arbitrary, we have shown that $\mathscr{I}(\mu, \mathcal{T}^*) \subset \mathscr{T}(\mu)$, and therefore $\mu \in \mathcal{E}(\mathcal{T})$. □

2.4.2 Dynamical Independence

With the development of Birkhoff systems and their Birkhoff sequences, it is now possible to describe a notion of the dynamical independence of two sets, a property of which will be used to characterize ergodicity for Birkhoff systems. Before we can get to this, we describe (in greater generality than needed) a notion of weak convergence associated with taking averages, as well as a Theorem which will be of importance for describing dynamical independence.

2.4.2.1 $\mathcal{A}$-weak convergence

We begin with the following definition.

Definition 2.4.15. Let I be a set, V be a normed vector space over the reals, and D be a directed set. For a set $F \in \mathscr{F}(I)$ and $x \in V^I$, let

$$A_F(x) = \frac{1}{|F|} \sum_{i\in F} x_i,$$

called the *F-average of x*. A net $\mathcal{A} \in \mathscr{F}(I)^D$ is called an *averaging net in I*, and to any such net we may associate the net $A_{\mathcal{A}}(x) = \{A_{\mathcal{A}_\alpha}(x)\}_{\alpha\in D}$, called the *$\mathcal{A}$-averages of x*. We say that x is *$\mathcal{A}$-weakly Cauchy* if the net $A_{\mathcal{A}}(x)$ is a Cauchy net in V, and that x is *$\mathcal{A}$-weakly convergent* (to $y \in V$) if $A_{\mathcal{A}}(x)$ is convergent (to y) in V. We also say that x is *bounded along $\mathcal{A}$* if there exists a constant $M \geq 0$ such that for every $\alpha \in D$ and every $i \in \mathcal{A}_\alpha$, we have $\|x_i\| \leq M$.

In addition to this notion of weak convergence, along an averaging net we make look at the density of some subset of I.

Definition 2.4.16. For a set I and subset $B \subset I$, we may take an averaging net $\mathcal{A}$ in I to define the $\mathcal{A}$-*density of* B as

$$\overline{d}_{\mathcal{A}}(B) = \sup\big\{\delta \in [0,1] : \forall \alpha \in D, \exists \beta \geq \alpha, |\mathcal{A}_\beta \cap B| \geq \delta|\mathcal{A}_\beta|\big\}.$$

This notion of density follows fairly standard notions of density, and this particular notion of density mimics that used in [26]. It is not hard to see that this notion of density is monotonically increasing in set containment, but we prove this nevertheless.

Lemma 2.4.17. *Let I be a set, and $C \subset B \subset I$. Then for any averaging net $\mathcal{A}$ in I, we have that $\overline{d}_{\mathcal{A}}(C) \leq \overline{d}_{\mathcal{A}}(B)$.*

Proof. Suppose that $\delta \in [0,1]$ is such that for every $\alpha \in D$, there exists $\beta \geq \alpha$ such that $|\mathcal{A}_\beta \cap C| \geq \delta|\mathcal{A}_\beta|$. Then since $C \subset B$, we clearly have that $\mathcal{A}_\beta \cap C \subset \mathcal{A}_\beta \cap B$, and therefore we have

$$|\mathcal{A}_\beta \cap B| \geq |\mathcal{A}_\beta \cap C| \geq \delta|\mathcal{A}_\beta|.$$

As such, for every $\alpha \in D$, there exists $\beta \geq \alpha$ such that $|\mathcal{A}_\beta \cap B| \geq \delta|\mathcal{A}|$. Therefore, we have

$$\begin{aligned}\big\{\delta \in [0,1] :& \forall \alpha \in D, \exists \beta \geq \alpha, |\mathcal{A}_\beta \cap C| \geq \delta|\mathcal{A}_\beta|\big\} \\ &\subset \big\{\delta \in [0,1] : \forall \alpha \in D, \exists \beta \geq \alpha, |\mathcal{A}_\beta \cap B| \geq \delta|\mathcal{A}_\beta|\big\},\end{aligned}$$

so taking the suprememum of both of these sets gives the desired result. □

The primary reason for discussing these concepts is the following theorem, which roughly states that for $x \in V^I$, if x is "mostly" constant, then x will converge $\mathcal{A}$-weakly to that constant.

Theorem 2.4.18. *Let I be a set, $\mathcal{A}$ an averaging net in I, and for a normed vector space (over the reals) V, let $x \in V^I$ be bounded along $\mathcal{A}$ and $\mathcal{A}$-weakly Cauchy. If there exists $y \in V$ such that the set $B_y = \{i \in I : x_i = y\}$ satisfies $\overline{d}_{\mathcal{A}}(B_y) = 1$, then x converges $\mathcal{A}$-weakly to y.*

Proof. Let $\epsilon > 0$. Since x is $\mathcal{A}$-weakly Cauchy, let $\alpha \in D$ such that for every $\beta, \gamma \geq \alpha$, we have that $\|A_{\mathcal{A}_\beta}(x) - A_{\mathcal{A}_\gamma}(x)\| < \frac{\epsilon}{2}$. As such, for any $\beta, \gamma \geq \alpha$, we have

$$\|A_{\mathcal{A}_\beta}(x) - y\| \leq \|A_{\mathcal{A}_\beta}(x)_\beta - A_{\mathcal{A}_\gamma}(x)\| + \|A_{\mathcal{A}_\gamma}(x) - y\| < \frac{\epsilon}{2} + \|A_{\mathcal{A}_\gamma}(x) - y\|.$$

Also, since x is bounded along $\mathcal{A}$, let $M \geq 0$ be such that for every $\alpha \in D$ and $i \in \mathcal{A}_\alpha$ we have $\|x_i\| \leq M$. With $\delta = 1 - \frac{\epsilon}{2(\|y\|+M)} < 1 = \overline{d}_{\mathcal{A}}(B_y)$, we have that for every $\beta \geq \alpha$, there exists $\gamma \geq \beta \geq \alpha$, so that $|\mathcal{A}_\gamma \cap B_y| \geq \delta|\mathcal{A}_\gamma|$, or alternatively, that $\delta \leq \frac{|\mathcal{A}_\gamma \cap B_y|}{|\mathcal{A}_\gamma|}$. With this γ, we have

$$\begin{aligned}
\|A_{\mathcal{A}_\gamma}(x) - y\| &= \left\| \frac{1}{|\mathcal{A}_\gamma|} \sum_{i \in \mathcal{A}_\gamma} x_i - y \right\| \\
&= \left\| \frac{1}{|\mathcal{A}_\gamma|} \sum_{i \in \mathcal{A}_\gamma \cap B_y} x_i - y + \frac{1}{|\mathcal{A}_\gamma|} \sum_{i \in \mathcal{A}_\gamma \setminus B_y} x_i \right\| \\
&\leq \left(1 - \frac{|\mathcal{A}_\gamma \cap B_y|}{|\mathcal{A}_\gamma|}\right) \|y\| + \frac{1}{|\mathcal{A}_\gamma|} \sum_{i \in \mathcal{A}_\gamma \setminus B_y} \|x_i\| \\
&\leq \left(1 - \frac{|\mathcal{A}_\gamma \cap B_y|}{|\mathcal{A}_\gamma|}\right) \|y\| + \left(1 - \frac{|\mathcal{A}_\gamma \cap B_y|}{|\mathcal{A}_\gamma|}\right) M \\
&\leq (1 - \delta)(\|y\| + M) \\
&= \frac{\epsilon}{2},
\end{aligned}$$

and thus for this same value of $\gamma \geq \beta$,

$$\|A_{\mathcal{A}_\beta}(x) - y\| < \frac{\epsilon}{2} + \|A_{\mathcal{A}_\gamma}(x) - y\| < \epsilon.$$

As this holds for every $\beta \geq \alpha$ (with appropriate choice of γ), we have that $A_{\mathcal{A}}(x)$ converges to y and so x converges $\mathcal{A}$-weakly to y. □

2.4.2.2 Dynamical Independence

Now that we have established the above definitions and theorem, we may state the definition of dynamical independence, and prove that it is a natural notion to consider when looking at dynamical systems.

Definition 2.4.19. Let $(X, \mathcal{T})$ be a Birkhoff system, let $\mu \in \mathcal{I}(\mathcal{T})$, and let $A, B \in \mathscr{A}_X$. Then we say that A *is* $(\mu, \mathcal{T})$*-dynamically independent of* B if for the set

$$I^{\mu}_{\mathcal{T}}(A, B) = \{T \in \mathcal{T}^* : \mu(T^{-1}(A) \cap B) = \mu(A \cap B)\},$$

there exists a Birkhoff sequence $\mathcal{A}$ for which $\overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}}(A, B)) = 1$.

While this definition may not be the easiest to work with, we show that there are simpler versions which are easier show hold, which also imply dynamical independence. The first of these is that dynamical independence readily transfers to the completion of a system.

Proposition 2.4.20. *Let $(X, \mathcal{T})$ be a Birkhoff system, let $\mu \in \mathcal{I}(\mathcal{T})$, and let $A, B \in \mathscr{A}_X$. Then A is $(\mu, \mathcal{T})$-dynamically independent of B if and only if A is $(\mu, \mathcal{T}^*)$-dynamically independent of B.*

Proof. First, by Proposition 2.3.30, we have that $\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}^*)$, and so $\mu \in \mathcal{I}(\mathcal{T}^*)$. Then, Theorem 2.3.26 gives that $(\mathcal{T}^*)^* = \mathcal{T}^*$, and thus

$$I^{\mu}_{\mathcal{T}^*}(A, B) = \{T \in (\mathcal{T}^*)^* = \mathcal{T}^* : \mu(T^{-1}(A) \cap B) = \mu(A \cap B)\} = I^{\mu}_{\mathcal{T}}(A, B).$$

As such, there exists a Birkhoff sequence $\mathcal{A}$ for which $\overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}}(A, B)) = 1$ if and only if there exists a Birkhoff sequence $\mathcal{A}$ for which $\overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}^*}(A, B)) = 1$. □

Let us now recall the notion of probabilistic independence.

Definition 2.4.21. Let X be a space and $\mu \in \mathcal{P}X$. Then two sets $A, B \in \mathcal{A}_X$ are *μ-independent* if

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

We may now prove our main theorem regarding dynamical independence, which is that dynamical independence implies probabilistic independence precisely for all ergodic measures.

Theorem 2.4.22. *Let $(X, \mathcal{T})$ be a Birkhoff system. Then for $\mu \in \mathcal{I}(\mathcal{T})$, the following are equivalent.*

(a) $\mu \in \mathcal{E}(\mathcal{T})$, and

(b) whenever A is $(\mu, \mathcal{T})$-dynamically independent of B, then A and B are μ-independent.

Proof. First, suppose that $\mu \in \mathcal{E}(\mathcal{T})$, and suppose that A is $(\mu, \mathcal{T})$-dynamically independent of B, so let $\mathcal{A}$ be a Birkhoff sequence for which $\overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}}(A, B)) = 1$. Then by Proposition 2.4.14, we have that

$$\lim_{n\to\infty} \frac{1}{|\mathcal{A}_n|} \sum_{T \in \mathcal{A}_n} \mu(T^{-1}(A) \cap B) = \mu(A)\mu(B).$$

By letting $I = \mathcal{T}^*$, $V = \mathbb{R}$, and $D = \mathbb{N}$, we may define $x_T = \mu(T^{-1}(A) \cap B)$, and thus x is $\mathcal{A}$-weakly convergent to $\mu(A)\mu(B)$, which clearly means x is $\mathcal{A}$-weakly Cauchy. But since $\overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}}(A, B)) = 1$ and by definition

$$I^{\mu}_{\mathcal{T}}(A, B) = \{T \in \mathcal{T}^* : x_T = \mu(A \cap B)\},$$

it follows from Theorem 2.4.18 that x also converges $\mathcal{A}$-weakly to $\mu(A \cap B)$. Therefore, it must be that $\mu(A \cap B) = \mu(A)\mu(B)$, so A and B are μ-independent.

Now, suppose that whenever A is $(\mu, \mathcal{T})$-dynamically independent of B, then A and B are μ-independent, and let $E \in \mathscr{I}(\mu, \mathcal{T}^*)$. Then for $T \in \mathcal{T}^*$, we have that $\mu(T^{-1}(E) \triangle E) = 0$, which implies $\mu(E \setminus T^{-1}(E)) = 0$ by monotonicity, and so

$$\mu(T^{-1}(E) \cap E) = \mu(T^{-1}(E) \cap E) + \mu(E \setminus T^{-1}(E)) = \mu(E) = \mu(E \cap E),$$

and therefore $T \in I^{\mu}_{\mathcal{T}}(E, E)$. With $T \in \mathcal{T}^*$ arbitrary, this shows that $I^{\mu}_{\mathcal{T}}(E, E) = \mathcal{T}^*$. As such, since $(X, \mathcal{T})$ is a Birkhoff system, there exists a Birkhoff sequence $\mathcal{A}$, and clearly for every $n \in \mathbb{N}$, we have

$$|\mathcal{A}_n \cap I^{\mu}_{\mathcal{T}}(E, E)| = |\mathcal{A}_n \cap \mathcal{T}^*| = 1|\mathcal{A}_n|,$$

and thus $\overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}}(E, E)) = 1$, which shows that E is $(\mu, \mathcal{T})$-dynamically independent of itself. By assumption then, E is μ-independent of itself, so

$$\mu(E) = \mu(E \cap E) = \mu(E)^2,$$

which gives that $\mu(E) \in \{0, 1\}$. As such $\mathscr{I}(\mu, \mathcal{T}^*) \subset \mathscr{T}(\mu)$, which by Proposition 2.3.31 implies that $\mu \in \mathcal{E}(\mathcal{T})$. □

In a sense, this theorem is a strengthening of the usual definition of ergodicity, as any $(\mu, \mathcal{T})$-invariant set is automatically $(\mu, \mathcal{T})$-dynamically independent of itself. The usual definition of ergodicity gives restrictions for where ergodic measures may be concentrated, however this notion of dynamical independence allows for the identification of independence structures which must be present in ergodic measures, further restricting their structure. We now turn our attention to more restricted notions of dynamical independence which will ultimately imply the most general definition above, but which will be easier to verify in practice. The original definition involves a particular choice of invariant measure μ in order to determine if a set is dynamically independent from another, which may in some sense defeat the utility of the definition, as it may be primarily useful in classifying which measures are invariant (and ergodic) to begin with. As such, the following definitions will be more useful in these circumstances.

Definition 2.4.23. Let $(X, \mathcal{T})$ be a Birkhoff system, and let $A, B \in \mathscr{A}_X$. Then we say that A *is* $\mathcal{T}$*-dynamically independent of* B if for the set

$$I_{\mathcal{T}}(A, B) = \{T \in \mathcal{T}^* : \exists S \in \mathcal{T}^*, T^{-1}(A) \cap B = S^{-1}(A \cap B)\},$$

there exists a Birkhoff sequence $\mathcal{A}$ for which $\overline{d}_{\mathcal{A}}(I_{\mathcal{T}}(A, B)) = 1$.

It is not too hard to see that this notion of dynamical independence implies the first notion (which is proven in the proposition below), however without reliance on a particular invariant measure μ in the definition of the set $I_{\mathcal{T}}(A, B)$, is it possible to identify that a set is dynamically independent of another without first needing to identify an invariant measure.

Proposition 2.4.24. *Let* $(X, \mathcal{T})$ *be a Birkhoff system, and let* $A, B \in \mathscr{A}_X$*. Then for any* $\mu \in \mathcal{I}(\mathcal{T})$*, we have that* $I_{\mathcal{T}}(A, B) \subset I^{\mu}_{\mathcal{T}}(A, B)$*, and so if* A *is* $\mathcal{T}$*-dynamically independent of* B*, then for any* $\mu \in \mathcal{I}(\mathcal{T})$*,* A *is* $(\mu, \mathcal{T})$*-dynamically independent of* B*.*

Proof. First, note that for any $\mu \in \mathcal{I}(\mathcal{T})$, we have that if $T \in I_{\mathcal{T}}(A, B)$, then there exists $S \in \mathcal{T}^*$ such that $T^{-1}(A) \cap B = S^{-1}(A \cap B)$. As such, we have $\mathcal{P}S(\mu) = \mu$, and so

$$\mu(T^{-1}(A) \cap B) = \mu(S^{-1}(A \cap B)) = \mathcal{P}S(\mu)(A \cap B) = \mu(A \cap B),$$

and therefore $T \in I^{\mu}_{\mathcal{T}}(A, B)$. As $T \in \mathcal{I}_{\mathcal{T}}(A, B)$ was arbitrary, this gives the containment $I_{\mathcal{T}}(A, B) \subset I^{\mu}_{\mathcal{T}}(A, B)$.

Therefore, if A is $\mathcal{T}$-dynamically independent of B, there is some Birkhoff sequence $\mathcal{A}$ for $(X, \mathcal{T})$ such that $\overline{d}_{\mathcal{A}}(I_{\mathcal{T}}(A, B)) = 1$, and Lemma 2.4.17 gives us that since $I_{\mathcal{T}}(A, B) \subset I^{\mu}_{\mathcal{T}}(A, B)$, then

$$1 = \overline{d}_{\mathcal{A}}(I_{\mathcal{T}}(A, B)) \leq \overline{d}_{\mathcal{A}}(I^{\mu}_{\mathcal{T}}(A, B)) \leq 1,$$

and so A is $(\mu, \mathcal{T})$-dynamically independent of B. □

The next somewhat cumbersome detail surrounding the notion of dynamical independence is the existence of some Birkhoff sequence along which the density of a particular set is 1. Identifying Birkhoff sequences is not an easy task, and in general, without further assumption on the structure of Birkhoff sequences, it is unclear that the existence of a single Birkhoff sequence implies the existence of many more. However, in the specific case of dynamical systems $(X, \mathcal{T})$ where $\mathcal{T}$ is a countable amenable group and $\mathcal{I}(\mathcal{T})$ is closed, any Følner sequence for $\mathcal{T}$ has a Birkhoff subsequence (by taking a tempered subsequence, as per Proposition 1.4 of [37] and Theorem 2.4.13), and so Birkhoff sequences are plentiful, and as long as one remains in the realm of Følner sequences, it is easy to transform Følner sequences into other Følner sequences. All of these properties come together for the following definition, specialized to this case.

Definition 2.4.25. Let $(X, \mathcal{G})$ be a dynamical system such that $\mathcal{G}$ is a countable amenable group and $\mathcal{I}(\mathcal{G})$ is closed. A subset $I \subset \mathcal{G}$ is said to be *thick* if for every finite set $F \subset \mathcal{G}$, there exists $T \in \mathcal{G}$ such that $F \circ T \subset I$. For $A, B \in \mathscr{A}_X$, define the set

$$I(A, B) = \{T \in \mathcal{G} : \exists S \in \mathcal{G}, T^{-1}(A) \cap B = S^{-1}(A \cap B)\}.$$

We say that A is dynamically independent of B if $I(A, B)$ is thick in $\mathcal{G}$.

The parallel between the set $I(A, B)$ and $I_{\mathcal{G}}(A, B)$ are rather clear, however it is not so immediately clear that requiring $I(A, B)$ to be thick in $\mathcal{G}$ will be sufficient to demonstrate the existence of a Birkhoff sequence for which $I(A,B)$ will have density 1.

Proposition 2.4.26. *Let $(X,\mathcal{G})$ be a dynamical system such that $\mathcal{G}$ is a countable amenable group and $\mathcal{I}(\mathcal{G})$ is closed, and let $A,B \in \mathscr{A}_X$. If A is dynamically independent of B, then A is $(\mu,\mathcal{G})$-dynamically independent of B for any $\mu \in \mathcal{I}(\mathcal{T})$.*

Proof. Since A is dynamically independent of B, it follows that $I(A,B)$ is thick in $\mathcal{G}$. Let $\mathcal{A}$ be any Følner sequence for $\mathcal{G}$. By the thickness of $I(A,B)$ in $\mathcal{G}$, for every $\mathcal{A}_n$, which is a finite subset of $\mathcal{G}$, let $T_n \in \mathcal{G}$ be such that $\mathcal{A}_n \circ T_n \subset I(A,B)$. Now for any $S \in \mathcal{G}$, we have that

$$\frac{|S \circ \mathcal{A}_n \circ T_n \triangle \mathcal{A}_n \circ T_n|}{|\mathcal{A}_n \circ T_n|} = \frac{|(S \circ \mathcal{A}_n \triangle \mathcal{A}_n) \circ T_n|}{|\mathcal{A}_n \circ T_n|} = \frac{|S \circ \mathcal{A}_n \triangle \mathcal{A}_n|}{|\mathcal{A}_n|},$$

and so it is clear that the sequence $\mathcal{A}_n \circ T_n$ is a Følner sequence. By Proposition 1.4 of [37], this sequence has a tempered subsequence $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$, and so by Theorem 2.4.13, $\mathcal{D}$ is a Birkhoff sequence. Furthermore, by construction we have $\mathcal{D}_n \subset I(A,B)$ for every $n \in \mathbb{N}$, and so

$$|\mathcal{D}_n \cap I(A,B)| = 1|\mathcal{D}_n|,$$

which gives that $\overline{d}_{\mathcal{D}}(I(A,B)) = 1$. It is further clear that $I(A,B) \subset I_{\mathcal{G}}(A,B)$, and therefore we must have that $\overline{d}_{\mathcal{D}}(I_{\mathcal{G}}(A,B)) = 1$. Therefore, A is $\mathcal{G}$-dynamically independent of B, and so by Proposition 2.4.24, this implies that A is $(\mu,\mathcal{G})$-dynamically independent of B for any $\mu \in \mathcal{I}(\mathcal{T})$. $\square$

2.5 Permutation and semicontractible systems

In 1985, Aldous [2] gave the following definition for a general process, reworded and narrowed slightly to fit the context of this chapter.

Definition 2.5.1. Let X be a space and let I be a countable (discrete) set. Then X^I endowed with the product topology is also a space. Let $\Theta(I)$ denote the set of bijections from any $J \subset I$ to I, and let $\Pi(I)$ denote the set of all bijections from I to I, or in other words the set of permutations of I. For $g \in \Theta(I)$, we may define a map $T_g : X^I \to X^I$ by

$$[T_g(x)]_i = x_{g^{-1}(i)}.$$

It is clear then that the inverse image of any open cylinder set in X^I under T_g will again be an open cylinder set, and thus $T_g \in \mathcal{C}(X^I)$. Furthermore, T_g is surjective for every $g \in \Theta(I)$, and a bijection for every $g \in \Pi(I)$ with $(T_g)^{-1} = T_{g^{-1}}$. For any subgroup $G \leq \Pi(I)$, define

$$\mathcal{G}_G = \{T_g : g \in G\},$$

which gives that $(X^I, \mathcal{G}_G)$ is a dynamical system. Any dynamical system $(X^I, \mathcal{G})$ which arises in such a manner is called a *permutation system*, and any $\mu \in \mathcal{I}(\mathcal{G})$ is called a *partially exchangeable process* with respect to $(X^I, \mathcal{G})$.

Aldous poses The Characterization Problem, which is simply stated as classifying all partially exchangeable processes for every permutation system $(X^I, \mathcal{G})$. As $\mathcal{G} \subset \mathcal{C}(X^I)$ for every permutation system $(X^I, \mathcal{G})$, it follows by Lemma 2.3.7 that $\mathcal{I}_{X^I}(\mathcal{G})$ is always closed, so characterizing these invariant measures is equivalent to characterizing $\mathcal{E}_{X^I}(\mathcal{G})$ by Proposition 2.3.16. We may study more general forms of processes where $G \subset \Pi(I)$ is not necessarily a subgroup, but the invariant measures for this system will be the same as those for the process generated by the group generated by G, as a result of Proposition 2.3.30. While Aldous used the term partially exchangeable process for the invariant measures of permutation systems, we will refer to them as the latter term to remain consistent with the rest of this chapter. For the remainder of this section we discuss useful tools in the characterization of these measures, including many proofs which hold in greater generality than just permutation systems including products, joinings, and power systems. We also give a new proof and extension of De Finetti's Theorem, which characterizes the invariant measures of exchangeable systems. Additionally, we give even further extensions of De Finetti's Theorem with a direction towards generalizing the Aldous-Hoover Theorem.

2.5.1 Product spaces

We begin with some basic facts about countable products of spaces which will be used throughout the section. The first tool we will need is the notion of cylinder sets.

Definition 2.5.2. Let I be a countable set and X_i be a space for each $i \in I$, so that $X_I = \prod_{i \in I} X_i$ endowed with the product topology is a space. Let $\pi_i : X_I \to X_i$ be defined as $\pi_i(x) = x_i$, which are each continuous and open (image of open sets is open) by the definition of X_I. The openness of π_i also guarantees that for any $E \in \mathscr{A}_{X_I}$ that $\pi_i(E) \in \mathscr{A}_{X_i}$. Also, for $J \subset I$, let $X_J = \prod_{j \in J} X_j$ and define $\pi_J : X_I \to X_J$ be defined as $[\pi_J(x)]_j = x_j$ (which are also all continuous). For any $J \subset I$ and collection of sets $E_j \in \mathscr{A}_{X_j}$ for every $j \in J$, the set

$$E = \bigcap_{j \in J} \pi_j^{-1}(E_j)$$

is the set of all $x \in X_I$ such that for every $j \in J$, we have that $x_j \in E_j$. Such a set is called a *cylinder set*. In the case that J is finite, we call it a *finite cylinder*, and if J is finite and each E_j is open, we call such a set an *open cylinder* (any such cylinder set is open in X_I, though only when J is finite). Open cylinders for a basis for the topology on X_I, and the set of all finite cylinders forms a π-system (closed under finite intersections) which generates $\mathscr{A}_{X_I}$.

Next, we define marginal measures, which will provide useful language for discussing the structure of the invariant measures for permutation systems.

Definition 2.5.3. Let I be a countable set and for each $i \in I$ let X_i be a space so that $X_I = \prod_{i\in I} X_i$ endowed with the product topology is a space. Then for $\mu \in \mathcal{P}X$ and $J \subset I$, the *J-marginal of* μ, denoted μ_J, is the probability measure $\mathcal{P}\pi_J(\mu)$. This means that for $E \in X_J$, we have

$$\mu_J(E) = \mathcal{P}\pi_J[\mu](E) = \mu(\pi_J^{-1}(E)) = \mu(E \times X_{I\setminus J}).$$

In the case that $J = \{i\}$, we denote the marginal by μ_i.

2.5.2 Product systems and joinings

Before shifting our focus to permutation systems, we develop some theory of product systems and joinings. These concepts will be useful in characterizing the invariant measures of permutation systems, however they hold in greater generality. We start with the definitions of these systems.

Definition 2.5.4. Given a countable set I and a dynamical system $(X_i, \mathcal{S}_i)$ for each $i \in I$, their *product* with $X_I = \prod_{i\in I} X_i$ (which is a space) and $\mathcal{T}_I = \prod_{i\in I} \mathcal{T}_i$ is $(X_I, \mathcal{T}_I)$, where an element of $\mathcal{T}_I$ is

$$T = T_1 \times T_2 \times \cdots ,$$

with $T_i \in \mathcal{T}_i$, so that $\pi_i \circ T = T_i \circ i$ for every $i \in I$.

The invariant measures of a product of dynamical systems have a rather simple characterization in fairly general cases (Proposition 2.5.7), however there is a certain class of subsystems of product systems which will be of particular interest in the study of permutation systems.

Definition 2.5.5. Given a countable set I and a dynamical system $(X_i, \mathcal{S}_i)$ for each

$i \in I$. A *joining* of these systems is a dynamical system $(X_I, \mathcal{T})$ where $\mathcal{T} \subset \mathcal{S}_I$, and for every $i \in I$ and $S \in \mathcal{S}_i$, there exists $T \in \mathcal{T}$ such that $\pi_i \circ T = S \circ \pi_i$.

All permutation systems are shown to be joinings of permutation systems which are in some sense indecomposable, and so it is important to describe the structure of the invariant measures of joinings in terms of the invariant measures of the systems joined together. This is given by the following result.

Proposition 2.5.6. *Let $(X_I, \mathcal{T})$ be a joining of the dynamical systems $(X_i, \mathcal{S}_i)$ for i in a countable set I. Then*

(a) if $\mu \in \mathcal{I}_{X_I}(\mathcal{T})$, then $\mu_i \in \mathcal{I}_{X_i}(\mathcal{S}_i)$ for every $i \in I$,

(b) if $\mu \in \mathcal{E}_{X_I}(\mathcal{T})$, then $\mu_i \in \mathcal{E}_{X_i}(\mathcal{S}_i)$ for every $i \in I$,

(c) if for each $i \in I$ we have some $\mu_i \in \mathcal{I}_{X_i}(\mathcal{S}_i)$, then $\bigotimes_{i \in I} \mu_i \in \mathcal{I}_{X_I}(\mathcal{T})$, and

(d) assuming $\mathcal{I}_{X_I}(\mathcal{T})$ is closed, if for each $i \in I$, we have $\mu_i \in \mathcal{E}_{X_i}(\mathcal{S}_i)$, then there exists $\nu \in \mathcal{E}_{X_I}(\mathcal{T})$ such that $\nu_i = \mu_i$ for every i.

Proof. For (a), let $\mu \in \mathcal{I}_{X_I}(\mathcal{T})$. For $i \in I$, we have $\mu_i = \mathcal{P}\pi_i(\mu)$, and for $S \in \mathcal{S}_i$, we have by assumption that there exists $T \in \mathcal{T}$ such that $\pi_i \circ T = S \circ \pi_i$, and since $\mu \in \mathcal{I}_{X_I}(\mathcal{T})$, we have that $\mathcal{P}T(\mu) = \mu$. So by Lemma 2.2.4 it must be that

$$\mathcal{P}S(\mu_i) = \mathcal{P}S(\mathcal{P}\pi_i(\mu)) = \mathcal{P}[S \circ \pi_i](\mu) = \mathcal{P}[\pi_i \circ T](\mu) = \mathcal{P}\pi_i(\mathcal{P}T(\mu)) = \mu_i,$$

whence $\mu_i \in \mathcal{I}_{X_i}(\mathcal{S}_i)$ as $S \in \mathcal{S}_i$ was arbitrary.

For (b), let $\mu \in \mathcal{E}_{X_I}(\mathcal{T})$. For $i \in I$, let $E \in \mathscr{I}(\mu_i, \mathcal{S}_i)$, so that for $S \in \mathcal{S}_i$, we have $\mu_i(E \triangle S^{-1}(E)) = 0$. By assumption, for $T \in \mathcal{T} \subset \mathcal{S}_I$, there is some $S \in \mathcal{S}_i$ so that $\pi_i \circ T = S \circ \pi_i$, which gives that

$$T^{-1}(\pi_i^{-1}(E)) = (\pi_i \circ T)^{-1}(E) = (S \circ \pi_i)^{-1}(E) = \pi_i^{-1}(S^{-1}(E)),$$

and so

$$\begin{aligned}\mu\big(\pi_i^{-1}(E)\triangle T^{-1}(\pi_i^{-1}(E))\big) &= \mu\big(\pi_i^{-1}(E)\triangle\pi_i^{-1}(S^{-1}(E))\big)\\ &= \mu(\pi_i^{-1}(E\triangle S^{-1}(E)))\\ &= \mathcal{P}\pi_i[\mu](E\triangle S^{-1}(E))\\ &= \mu_i(E\triangle S^{-1}(E)) = 0.\end{aligned}$$

Since $T \in \mathcal{T}$ was arbitrary, this demonstrates that $\pi_i^{-1}(E) \in \mathscr{I}(\mu, \mathcal{T})$, and since $\mu \in \mathcal{E}_{X_I}(\mathcal{T})$, this gives that $\mu_i(E) = [\mathcal{P}\pi_i(\mu)](E) = \mu(\pi_i^{-1}(E)) \in \{0,1\}$. With $E \in \mathscr{I}(\mu_i, \mathcal{S}_i)$ arbitrary, we have shown that $\mu_i \in \mathcal{E}_{X_i}(\mathcal{S}_i)$.

For (c), let $\mu_i \in \mathcal{I}_{X_i}(\mathcal{S}_i)$ for each $i \in I$, and define $\nu = \bigoplus_{i\in I}\mu_i$ to be the product measure on X_I (which is well defined by Proposition 2.1.6). Let $T \in \mathcal{T}$, and let $\mathscr{L} = \{E \in \mathscr{A}_{X_I} : \mathcal{P}T[\nu](E) = \nu(E)\}$. We have that $\varnothing, X \in \mathscr{L}$ trivially, as $T^{-1}(X) = X$ and $T^{-1}(\varnothing) = \varnothing$. Next, if $E \in \mathscr{L}$, we have that $\mathcal{P}T[\nu](E) = \nu(E)$, and thus

$$\mathcal{P}T[\nu](X \setminus E) = 1 - \mathcal{P}T[\nu](E) = 1 - \nu(E) = \nu(X \setminus E),$$

and so $X\setminus E \in \mathscr{L}$. Finally, let $\{E_n\}_{n\in\mathbb{N}} \subset \mathscr{L}$ be mutually disjoint. Then by countable additivity, we have

$$\mathcal{P}T[\nu]\left(\bigsqcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mathcal{P}T[\nu](E_n) = \sum_{n\in\mathbb{N}} \nu(E_n) = \nu\left(\bigsqcup_{n\in\mathbb{N}} E_n\right),$$

and thus $\bigsqcup_{n\in\mathbb{N}} E_n \in \mathscr{L}$. This shows that $\mathscr{L}$ is a λ-system. Now, let $J \subset I$ be finite, and let $E_j \in \mathscr{A}_{X_j}$ for each $j \in J$. For $T \in \mathcal{T}$, we have for each $j \in J$ that there exists $S_j \in \mathcal{S}_j$ such that $\pi_j \circ T = S_j \circ \pi_j$, and note that $\mathcal{P}S_j(\nu_j) = \mathcal{P}S_j(\mu_j) = \mu_j = \nu_j$. With $\bigcap_{j\in J}\pi_j^{-1}(E_j) \in \mathscr{A}_{X_I}$ an arbitrary finite cylinder, we have by the definition of ν as a product measure that

$$\begin{aligned}
\mathcal{P}T[\nu]\left(\bigcap_{j\in J}\pi_j^{-1}(E_j)\right) &= \nu\left(T^{-1}\left(\bigcap_{j\in J}\pi_j^{-1}(E_j)\right)\right) \\
&= \nu\left(\bigcap_{j\in J}T^{-1}(\pi_j^{-1}(E_j))\right) = \nu\left(\bigcap_{j\in J}\pi_j^{-1}(S_j^{-1}(E_j))\right) \\
&= \prod_{j\in J}\nu((S_j\circ\pi_j)^{-1}(E_j)) \\
&= \prod_{j\in J}\mathcal{P}[S_j\circ\pi_j](\nu)(E_j) = \prod_{j\in J}\mathcal{P}S_j[\nu_j](E_j) \\
&= \prod_{j\in J}\nu_j(E) = \prod_{j\in J}\nu(\pi_j^{-1}(E_j)) = \nu\left(\bigcap_{j\in J}\pi_j^{-1}(E_j)\right),
\end{aligned}$$

and thus the collection of all finite cylinders $\mathscr{P}$ is contained in $\mathscr{L}$. Furthermore, $\mathscr{P}$ is a π-system, and thus by the $\pi-\lambda$ Theorem, the σ-algebra generated by $\mathscr{P}$ (which is $\mathscr{A}_{X_I}$), must be contained in $\mathscr{L}\subset\mathscr{A}_{X_I}$, which gives that $\mathscr{L}=\mathscr{A}_{X_I}$. This proves that $\mathcal{P}T(\eta)=\eta$. As $T\in\mathcal{T}$ was arbitrary, it follows that $\nu\in\mathcal{I}_{X_I}(\mathcal{T})$.

Finally, for (d), let $\mu_i\in\mathcal{E}_{X_i}(\mathcal{S}_i)$ for each $i\in I$. Note that for each $i\in I$, we have $\pi_i\in\mathcal{C}(X_I,X_i)$ so by Lemma 2.2.7, we have $\mathcal{P}\pi_i\in\mathcal{C}(\mathcal{P}X_I,\mathcal{P}X_i)$ for each $i\in I$. Thus $[\mathcal{P}\pi_i]^{-1}(\{\mu\})$ is a closed subset of $\mathcal{P}X_I$ for each $i\in I$. Therefore,

$$C(\{\mu_i\}_{i\in I})=\{\nu\in\mathcal{P}X_I:\forall i\in I,\nu_i=\mu_i\}=\bigcap_{i\in I}[\mathcal{P}\pi_i]^{-1}(\{\mu\})$$

is a closed subset of $\mathcal{P}X_I$. By the assumption that $\mathcal{I}_{X_I}(\mathcal{T})$ is closed, this gives that $J(\{\mu_i\}_{i\in I})=C(\{\mu_i\}_{i\in I})\cap\mathcal{I}_{X_I}(\mathcal{T})$ is closed. Furthermore, both sets are convex, and we know by (c) that $\nu=\bigotimes_{i\in I}\mu_i\in\mathcal{I}_{X_I}(\mathcal{T})$ and clearly $\nu\in C(\{\mu_i\}_{i\in I})$, and therefore $J(\{\mu_i\}_{i\in I})$ is a nonempty compact convex set. By the Choquet Representation Theorem, it has an extreme point $\eta\in J(\{\mu_i\}_{i\in I})$. Now suppose that $\eta^1,\eta^2\in\mathcal{I}_{X_I}(\mathcal{T})$ and

$t \in (0,1)$ such that $\eta = t\eta^1 + (1-t)\eta^2$. Then for $i \in I$ and $E \in \mathscr{A}_{X_i}$,

$$\begin{aligned}\mu_i(E) &= \eta_i(E) = \eta(\pi_i^{-1}(E)) \\ &= t\eta^1(\pi_i^{-1}(E)) + (1-t)\eta^2(\pi_i^{-1}(E)) = t\eta_i^1(E) + (1-t)\eta_i^2(E).\end{aligned}$$

However, by Proposition 2.3.15 since $\mu_i \in \mathcal{E}_{X_i}(\mathcal{S}_i)$, it is also an extreme point of $\mathcal{I}_{X_i}(\mathcal{S}_i)$. With (a) giving that $\eta_i^1, \eta_i^2 \in \mathcal{I}_{X_i}(\mathcal{S}_i)$, the only way this is possible is for $\mu_i = \eta_i^1 = \eta_i^2$. As this holds for every $i \in I$, it follows that $\eta^1, \eta^2 \in J(\{\mu_i\}_{i\in I})$. But η was an extreme point of $J(\{\mu_i\}_{i\in I})$, and so for $\eta = t\eta^1 + (1-t)\eta^2$ with $t \in (0,1)$, it must be that $\eta = \eta_1 = \eta_2$. As such, η is an extreme point of $\mathcal{I}_{X_I}(\mathcal{T})$, so must be in $\mathcal{E}_{X_I}(\mathcal{T})$. □

The requirement that $\mathcal{I}_{X_I}(\mathcal{T})$ is closed in the proposition above for (d) is only so that there is an extreme point of the set of invariant measures with the appropriate marginals. This could be replaced with a sufficiently general ergodic decomposition theorem such at [18] once adequately translated to the context presented here. In the well behaved circumstance of the product of systems, it is possibly to precisely describe the ergodic measures in fairly general circumstances.

Proposition 2.5.7. *Let $(X_I, \mathcal{T}_I)$ be the product of the dynamical systems $(X_i, \mathcal{T}_i)$ for i in some countable set I where each $\mathcal{T}_i$ is a countable amenable group, and suppose that $\mathcal{I}_{X_I}(\mathcal{T}_I)$ is closed. Then*

$$\mathcal{E}_{X_I}(\mathcal{T}_I) = \bigotimes_{i\in I} \mathcal{E}_{X_i}(\mathcal{T}_i) = \left\{ \bigotimes_{i\in I} \mu_i : \forall i \in I, \mu_i \in \mathcal{E}_{X_i}(\mathcal{T}_i) \right\}.$$

Proof. First, note that the countable direct sum of countable amenable groups is again countable and amenable, since finite products of countable amenable groups is countable and amenable, and the direct union of a set of groups is countable and amenable. As such, $\mathcal{T}_I^0 = \bigoplus_{i\in I} \mathcal{T}_i$ (which is the direct sum of each $\mathcal{T}_i$) is a countable amenable group. Furthermore, since $(\mathcal{T}_I^0)^*$ is closed in $\mathcal{B}(X_I)$ by Proposition 2.3.30 gives that for a sequence $\{S_n\}_{n\in\mathbb{N}}$ in $(\mathcal{T}_I^0)^*$ which converges pointwise to T, it must be that $T \in (\mathcal{T}_I^0)^*$. For $T \in \mathcal{T}_I$, we have that

$$T = T_1 \times T_2 \times \cdots \times T_n \times \cdots$$

for $T_i \in \mathcal{T}_i$. Let

$$S_n = T_1 \times T_2 \times \cdots \times T_n \times e_{n+1} \times e_{n+2} \times \cdots ,$$

where e_i is the identity of $\mathcal{T}_i$. Then clearly $S_n \in \mathcal{T}_I^0$, and S_n converges pointwise to T, and therefore $T \in (\mathcal{T}_I^0)^*$. As such, we have $\mathcal{T}_I^0 \subset \mathcal{T}_I \subset (\mathcal{T}_I^0)^*$, and thus by Theorem 2.3.26 we have that $(\mathcal{T}_I)^* = (\mathcal{T}_I^0)^*$, which further gives by Proposition 2.3.30 that

$$\mathcal{I}_{X_I}(\mathcal{T}_I) = \mathcal{I}_{X_I}((\mathcal{T}_I)^*) = \mathcal{I}_{X_I}((\mathcal{T}_I^0)^*) = \mathcal{I}_{X_I}(\mathcal{T}_I^0),$$

which is closed by assumption. As such, by Proposition 2.4.10 and Theorem 2.4.13 it follows that a tempered Følner sequence for $\mathcal{T}_I^0$ is a Birkhoff sequence for $(X_I, \mathcal{T}_I^0)$, which is also a Birkhoff sequence for $(X_I, (\mathcal{T}_I^0)^* = (\mathcal{T}_I)^*)$, which is also a Birkhoff sequence for $(X_I, \mathcal{T}_I)$.

Next, let $\mu \in \mathcal{E}_{X_I}(\mathcal{T}_I)$, and we proceed by induction. For the base case, let $J = \{j\} \subset I$ and $E_j \in \mathscr{A}_{X_j}$, and note that

$$\mu(\pi_j^{-1}(E_j)) = \mathcal{P}\pi_j[\mu](E_j) = \mu_j(E_j).$$

Now, suppose that for some $k \in \mathbb{N}$, that for any $J \subset I$ with $|J| = k$ and any $E_j \in \mathscr{A}_{X_j}$

for $j \in J$, we have that

$$\mu\left(\bigcap_{j\in J} \pi_j^{-1}(E_j)\right) = \prod_{j\in J} \mu_j(E_j).$$

Let $J \subset I$ with $|J| = k+1$ and $E_j \in \mathscr{A}_{X_j}$ for $j \in J$, and chose some $i \in J$. Let $E = \bigcap_{j\in J\setminus\{i\}} \pi_j^{-1}(E_j)$ and $F = \pi_i^{-1}(E_i)$. Then for $T \in \mathcal{T}_I$, we have for each $j \in J \setminus \{i\}$ that there is some $T_j \in \mathcal{T}_j$ so that $\pi_j \circ T = T_j \circ \pi_j$. Let $S \in \mathcal{T}_I$ be the map such that $\pi_j \circ S = T_j \circ \pi_j$ for every $j \in J \setminus \{i\}$ and $\pi_k \circ S = e_k \circ \pi_k$ for all other $k \in I \setminus (J \setminus \{i\})$. Then we have

$$\begin{aligned} T^{-1}(E) \cap F = T^{-1}\left(\bigcap_{j\in J} \pi_j^{-1}(E_j)\right) \cap F &= \left(\bigcap_{j\in J} T^{-1}(\pi_j^{-1}(E_j))\right) \cap F \\ &= \left(\bigcap_{j\in J} \pi_j^{-1}(T_j^{-1}(E_j))\right) \cap F. \end{aligned}$$

On the other hand, since $i \in I \setminus (J \setminus \{i\})$, we have

$$\begin{aligned} S^{-1}(E \cap F) &= S^{-1}\left(\bigcap_{j\in J} \pi_j^{-1}(E_j) \cap \pi_i^{-1}(E_i)\right) \\ &= \bigcap_{j\in J} S^{-1}(\pi_j^{-1}(E_j)) \cap S^{-1}(\pi_i^{-1}(E_i)) \\ &= \bigcap_{j\in J} \pi_j^{-1}(T_j^{-1}(E_j)) \cap \pi_i^{-1}(e_i^{-1}(E_i)) \\ &= \bigcap_{j\in J} \pi_j^{-1}(T_j^{-1}(E_j)) \cap F, \end{aligned}$$

and so $T^{-1}(E) \cap F = S^{-1}(E \cap F)$. As such, we have that $\mathcal{T}_I \subset I_{\mathcal{T}_I}(E, F)$. Given a tempered Følner sequence $\mathcal{A}$ for $\mathcal{T}_I^0$, we then have that $\mathcal{A}$ is a Birkhoff sequence for $(X_I, \mathcal{T}_I)$, and with $\mathcal{T}_I^0 \subset \mathcal{T}_I \subset I_{\mathcal{T}_I}(E, F)$, it follows that

$$|I_{\mathcal{T}_I}(E, F) \cap \mathcal{A}_n| = 1|\mathcal{A}_n|$$

for every $n \in \mathbb{N}$, and therefore $\overline{d}_{\mathcal{A}}(I_{\mathcal{T}_I}(E, F)) = 1$. As such, Proposition 2.4.24 gives us that E is $(\mu, \mathcal{T}_I)$-dynamically independent of B. Since $\mu \in \mathcal{E}_{X_I}(\mathcal{T}_I)$, we have by Theorem 2.4.22 that $\mu(E \cap F) = \mu(E)\mu(F)$. By the inductive hypothesis, since $|J \setminus \{i\}| = k$, we have that

$$\mu(E) = \mu\left(\bigcap_{j \in J\setminus\{i\}} \pi_j^{-1}(E_j)\right) = \prod_{j \in J\setminus\{i\}} \mu_j(E_j),$$

and also that $\mu(F) = \mu(\pi_i^{-1}(E_i)) = \mathcal{P}\pi_i[\mu](E_i) = \mu_i(E_i)$. So with $\mu(E \cap F) = \mu(E)\mu(F)$, we have that

$$\begin{aligned}\mu\left(\bigcap_{j \in J} \pi_j^{-1}(E_j)\right) &= \mu\left(\bigcap_{j \in J\setminus\{i\}} \pi_j^{-1}(E_j) \cap \pi_i^{-1}(E_i)\right) = \mu(E \cap F) \\ &= \mu(E)\mu(F) = \mu_i(E_i) \prod_{J \in J\setminus\{i\}} \mu_j(E_j) = \prod_{j \in J} \mu_j(E_j).\end{aligned}$$

By induction, it then follows that this holds for every $J \subset I$ finite. Since the Kolmogorov Extension Theorem gives that any such measure μ is unique, and the measure $\bigotimes_{i \in I} \mu_i$ also has the property of the display above, it must be that $\mu = \bigotimes_{i \in I} \mu_i$. Since $\mu \in \mathcal{E}_{X_I}(\mathcal{T}_I)$ was arbitrary, it follows that $\mathcal{E}_{X_I}(\mathcal{T}_I) \subset \bigotimes_{i \in I} \mathcal{E}_{X_i}(\mathcal{T}_i)$.

Finally, for each $i \in I$, let $\mu_i \in \mathcal{E}_{X_i}(\mathcal{T}_i)$. Since $\mathcal{I}_{X_I}(\mathcal{T}_I)$ is closed, we have by Proposition 2.5.6(d) that there exists $\nu \in \mathcal{E}_{X_I}(\mathcal{T}_I)$ such that $\nu_i = \mu_i$ for all $i \in I$. But by the argument above, we have that

$$\nu = \bigotimes_{i \in I} \nu_i = \bigotimes_{i \in I} \mu_i,$$

and thus $\mu = \bigotimes_{i \in I} \mu_i$ is in $\mathcal{E}_{X_I}(\mathcal{T}_I)$. Since $\mu \in \bigotimes_{i \in I} \mathcal{E}_{X_i}(\mathcal{T}_i)$ was arbitrary, this gives us that $\bigotimes_{i \in I} \mathcal{E}_{X_i}(\mathcal{T}_i) \subset \mathcal{E}_{X_I}(\mathcal{T}_I)$, which when combined with the result of the previous paragraph gives the desired result. $\square$

Beyond this result, it is difficult to classify the invariant measures for an arbitrary joining of dynamical systems. For products of permutation systems, there are clearly many ergodic measures, in particular if we take the product of a given permutation system $(X^I, \mathcal{G})$ over a countable set J, we have that $(X^I)^J = X^{I\times J}$

$$\mathcal{G}_J = \prod_{j\in J} \mathcal{G} = \mathcal{G}^J,$$

so for the product permutation system $(X^{I\times J}, \mathcal{G}^J)$, the Theorem above gives that

$$\mathcal{E}_{X^{I\times J}}(\mathcal{G}^J) = \prod_{j\in J} \mathcal{E}_{X^I}(\mathcal{G}),$$

so every possible combination of a countable product of ergodic measures will be ergodic. On the other hand, with $(X^{I\times J}, \mathcal{T})$ the joining of the permutation system $(X^I, \mathcal{G})$ for j in a countable set J, such that $\mathcal{T}$ is the diagonal of $\mathcal{G}^J$, meaning that $\mathcal{T}$ is comprised of transformations of the form $T = \prod_{j\in J} S$ for $S \in \mathcal{G}$. Then $\mathcal{E}_{X^{I\times J}}(\mathcal{T})$ consists of a far smaller variety of invariant measures. As such, the sorts of invariant measures which can be obtained from joinings can be rather varied. Nevertheless, with the proposition above, we use the theory of completions in order to extend this result to a broad class of joinings.

Theorem 2.5.8. *Let $(X_I, \mathcal{T})$ be a joining of the dynamical systems $(X_i, \mathcal{S}_i)$ for i in some countable set I where each $\mathcal{S}_i$ is a countable amenable group, $\mathcal{I}_{X_i}(\mathcal{T})$ is closed, and $\mathcal{S}_I \subset \mathcal{T}^*$. Then*

$$\mathcal{E}_{X_I}(\mathcal{T}) = \bigotimes_{i\in I} \mathcal{E}_{X_i}(\mathcal{S}_i).$$

Proof. By assumption, we have that $\mathcal{T} \subset \mathcal{S}_I \subset \mathcal{T}^*$, and therefore by Theorem 2.3.26, we have that $(\mathcal{S}_I)^* = \mathcal{T}^*$. Therefore, by Proposition 2.3.30 and Proposition 2.5.7 we have

$$\mathcal{E}_{X_I}(\mathcal{T}) = \mathcal{E}_{X_I}(\mathcal{T}^*) = \mathcal{E}_{X_I}((\mathcal{S}_I)^*) = \mathcal{E}_{X_I}(\mathcal{S}_I) = \bigotimes_{i\in I} \mathcal{E}_{X_i}(\mathcal{S}_i).$$

□

In light of this result, we give the following definition.

Definition 2.5.9. Let $(X_I, \mathcal{T})$ be a joining of the dynamical systems $(X_i, \mathcal{S}_i)$ for i in some countable set I where each $\mathcal{S}_i$ is a monoid. Then the joining is said to be *independent* if $\mathcal{S}_I \subset \mathcal{T}^*$.

The restriction that $\mathcal{S}_i$ is a monoid is rather mild, as for any collection $\mathcal{U}_i$ of transformation, the monoid $\mathcal{S}_i$ generated by it (by adding in the identity map if necessary) will be contained in $(\mathcal{S}_i)^*$ by Proposition 2.3.30, and so the set of invariant measures for $(X, \mathcal{U}_i)$ will be identical to those for $(X, \mathcal{S}_i)$. While we have not classified in full generality the invariant measures for independent joinings of dynamical systems, the results shown lead to the following conjecture.

Conjecture 2.5.10. *Let $(X_I, \mathcal{T})$ be an independent joining of the dynamical systems $(X_i, \mathcal{S}_i)$ for i in some countable set I where each $\mathcal{S}_i$ contains id_{X_i}. Then*

$$\mathcal{E}_{X_I}(\mathcal{T}) = \bigotimes_{i \in I} \mathcal{E}_{X_i}(\mathcal{S}_i).$$

Note that the requirement that each $\mathcal{S}_i$ contains some well behaved transformation (such as id_{X_i}) is a necessary condition for this conjecture to hold at this level of generality. Indeed, if each $\mathcal{S}_i$ was just a single non-identity transformation S_i, then the only joining of these systems is the product system, which would of course be independent. This joining however would be a joining in the classical sense, and it is known in general that if μ is ergodic for (X, T) and ν is ergodic for (Y, S) that $\mu \otimes \nu$ is the only ergodic measure for $(X \times Y, T \times S)$ if and only if (X, T, μ) and (Y, S, μ) are disjoint (in fact, this is the definition, see Chapter 6 Section 6 of [23]). But there are plenty of non-disjoint systems, and so we must have some additional requirement on the $\mathcal{S}_i$ in order for the conjecture above to hold. It may be sufficient for each $\mathcal{S}_i$

to only be a semigroup however, although it seems more likely that $\mathcal{S}_i$ will need to also have an identity based on the proof of Proposition 2.5.7 in the construction of the element S. In any case, this conjecture would follow from the one below similarly to how Theorem 2.5.8 follows Proposition 2.5.7, so perhaps it is the one that should be proven.

Conjecture 2.5.11. *Let $(X_I, \mathcal{T}_I)$ be the product of the dynamical systems $(X_i, \mathcal{S}_i)$ for i in some countable set I where each $\mathcal{S}_i$ is a monoid. Then*

$$\mathcal{E}_{X_I}(\mathcal{T}_I) = \bigotimes_{i \in I} \mathcal{E}_{X_i}(\mathcal{T}_i).$$

While Theorem 2.5.8 is simple to prove in light of Proposition 2.5.7, it is a rather powerful tool in the analysis of permutation systems. We demonstrate this with the following example.

Example 2.5.12. Let P be the set of positive prime numbers. For each $p \in P$ let $I_p = \{0, 1, \ldots, p-1\}$ and let $g_p : I_p \to I_p$ be defined by

$$g_p(i) = i + 1 \mod p,$$

and let $G_p \leq \Pi(I_p)$ be the cyclic group of order p generated by g_p. Let $I = \bigcup_{p \in P} I_p$, where we take this union to be disjoint (so for $0_p \in I_p$ and $0_q \in I_q$, $0_p \neq 0_q$, etc.). Define $g : I \to I$ by $g|_{I_p} = g_p$, and let $G \leq \Pi(I)$ be the cyclic group generated by g, which has infinite order, so G is isomorphic to $\mathbb{Z}$. For a space X, let us take the permutation system $(X^I, \mathcal{G}_G)$. Then

$$\mathcal{E}_{X^I}(\mathcal{G}_G) = \bigotimes_{p \in P} \mathcal{E}_{X^{I_p}}(\mathcal{G}_{G_p}).$$

Proof. First, note that the space X^I may be modeled as the product space $\prod_{p \in P} X^{I_p}$

and let us denote $\pi_p : X^I \to X^{I_p}$ the corresponding projection map. Also, note that

$$T_g = \prod_{p \in P} T_{g_p},$$

so we have that $T_g \in \prod_{p\in P} \mathcal{G}_{G_p}$, and subsequently that $\mathcal{G}_G \subset \mathcal{G}_P = \prod_{p\in P} \mathcal{G}_{G_p}$ since $\mathcal{G}_G$ is the cyclic group generated by T_g. Furthermore, for $T_h \in \mathcal{G}_{G_p}$, we have that $h = g_p^k$ for some $k \in I_p$, and thus

$$\pi_p \circ T_{g^k} = \pi_p \circ (T_g)^k = (T_{g_p})^k \circ \pi_p = T_{g_p^k} \circ \pi_p = T_h \circ \pi_p,$$

which gives that $(X^I, \mathcal{G}_G)$ is a joining of the systems $(X^{I_p}, \mathcal{G}_{G_p})$ for $p \in P$. Also note that $\mathcal{G}_{G_p}$ is countable and amenable for each p (it is finite), and that $\mathcal{I}_{\mathcal{G}_G}(X^I)$ is closed because $(X^I, \mathcal{G}_G)$ is a permutation system.

Now, let $T \in \mathcal{G}_P$, so $T = \prod_{p\in P} T_{h_p}$ for $T_{h_p} \in \mathcal{G}_{G_p}$ for every $p \in P$. Let $k_p \in I_p$ be such that $h_p = g_p^{k_p}$. Now for each n, let $P_n \subset P$ be the set of the n smallest prime numbers, so that we have a finite system of congruences $x = k_p \mod p$ for $p \in P_n$. By the Chinese Remainder Theorem with $m_n = \prod_{p\in P_n} p$, this system has a unique solution $x = l_n \mod m_n$. Define $T_n = (T_g)^{l_n} \in \mathcal{G}$, which is a sequence. By definition, we have that

$$T_n = (T_g)^{l_n} = \prod_{p\in P} (T_{g_p})^{l_n} = \prod_{p\in P} T_{g_p^{l_n}}.$$

For $p \in P_n$, since g_p has order p, and we have that $l_n = k_p \mod p$, so it follows that

$$T_{g_p^{l_n}} = T_{g_p^{k_p}} = T_{h_p},$$

which gives that

$$T_n = (T_g)^{l_n} = \prod_{p\in P_n} T_{h_p} \times \prod_{p\in P\setminus P_n} (T_{g_p})^{l_n}.$$

Taking a limit in n, using that $P = \bigcup_{n\in\mathbb{N}} P_n$, it follows that

$$\lim_{n\to\infty} T_n = \lim_{n\to\infty} (T_g)^{l_n} = \prod_{p\in P} T_{h_p} = T$$

pointwise for every $x \in X^I$. Since we have that $T_n \in \mathcal{G}_G \subset (\mathcal{G}_G)^*$ by Theorem 2.3.26 we therefore have by Proposition 2.3.30 that $T \in (\mathcal{G}_G)^*$. Since $T \in \mathcal{G}_P$ was arbitrary, we have shown that $\mathcal{G}_P \subset (\mathcal{G}_G)^*$. By Theorem 2.5.8, we have that

$$\mathcal{E}_{X^I}(\mathcal{G}_G) = \bigotimes_{p\in P} \mathcal{E}_{X^{I_p}}(\mathcal{G}_{G_p}).$$

□

This example has two notable features. The first is that, up until the final sentence, measures are not referenced a single time in the entire proof. This highlights why working with the completion can simplify many arguments and focus primarily on the dynamics and the restrictions they place on the invariant measures. The second is that this permutation system is a rather classical dynamical system given by a homeomorphism of some compact metrizable space. From this classical perspective, it would be rather tedious to show that the ergodic measures decompose in this fashion given access to the single homeomorphism T_g. In the classical theory of joinings, the system above is the product topological joining of the action of T_{g_p} on X^{I_p}, however these systems are far from disjoint in the sense of topological joinings, and even if this was the case, this does not imply measure disjointness (page 141, [23]). It is really through the connection between this $\mathbb{Z}$ action and the action of the rather unruly group $\prod_{p\in P} G_p$ that the proof above is viable in light of the completion of the dynamical system.

2.5.3 General results on Permutation systems

We now turn our attention to permutation systems and prove some general results about them. In fact, we can prove these results for a more broad class of systems called semi-contractible systems.

Definition 2.5.13. Let X be a space and I a countable (discrete) set. Then X^I endowed with the product topology is also a space. For monoid $M \subset \Theta(I)$, define

$$\mathcal{S}_M = \{T_g : g \in M\}.$$

This gives a dynamical system $(X^I, \mathcal{S}_M)$, and any dynamical system $(X^I, \mathcal{S})$ which arises in such a manner is called a *semicontractible system*. Of course, every permutation system is a semi-contractible system.

The first of these is that it is rather easy to describe the behavior of cylinder sets when transformed by some function in the system.

Lemma 2.5.14. *Let X be a space, and I be a countable set so that X^I endowed with the product topology is a space. Then for any cylinder set*

$$E = \bigcap_{j \in J} \pi_j^{-1}(E_j),$$

we have for $g \in \Theta(I)$ that

$$T_g^{-1}(E) = \bigcap_{j \in J} \pi_{g^{-1}(j)}^{-1}(E_j).$$

Proof. Let $F = \bigcap_{j \in J} \pi_{g^{-1}(j)}^{-1}(E_j)$. First, suppose that $x \in T_g^{-1}(E)$, so we have for every $j \in J$ that $[T_g(x)]_j \in E_j$. Then for every $j \in J$, we have that $[T_g(x)]_j = x_{g^{-1}(j)} \in E_j$, and therefore $x_{g^{-1}(j)} \in E_j$. As such, we must have that $x \in F$, and since $x \in T_g^{-1}(E)$ was arbitrary, we have shown that $T_g^{-1}(E) \subset F$.

Now, let $x \in F$. Then we have for every $j \in J$ that $x_{g^{-1}(j)} \in E_j$, and so $[T_g(x)]_j = x_{g^{-1}(j)} \in E_j$, which gives that $T_g(x) \in E$, and therefore $x \in T_g^{-1}(E)$. Since $x \in F$ was arbitrary, we have shown that $F \subset T_g^{-1}(E)$, which when combined with the result of the previous paragraph is the desired result. □

Due to the simplicity of this result, we use it without reference, and only state it to ensure that it is clear. With it, we may prove that permutation systems indeed have many invariant and ergodic measures.

Lemma 2.5.15. *Let $(X^I, \mathcal{S})$ be a semicontractible system, and let $\mu \in \mathcal{P}X$. Then $\nu = \bigotimes_{i \in I} \mu$ is in $\mathcal{I}_{X^I}(\mathcal{S})$.*

Proof. Let $M \subset \Theta(I)$ be such that $\mathcal{S} = \mathcal{S}_M$, and let $T_g \in \mathcal{S}$ for $g \in M$. Let $\mathscr{L} = \{E \in \mathscr{A}_{X^I} : \mathcal{P}T_g[\nu](E) = \nu(E)\}$, and note that

$$\mathcal{P}T_g[\nu](X) = \nu(T_g^{-1}(X)) = \nu(X) \quad \text{and} \quad \mathcal{P}T_g[\nu](\varnothing) = \nu(T_g^{-1}(\varnothing)) = \nu(\varnothing),$$

and so we have $\varnothing, X \in \mathscr{L}$. Now, for $E \in \mathscr{L}$, we have $\mathcal{P}T_g[\nu](E) = \nu(E)$, and so

$$\begin{aligned}\mathcal{P}T_g[\nu](X \setminus E) &= \nu(T_g^{-1}(X \setminus E)) = \nu(X \setminus T_g^{-1}(E)) = 1 - \nu(T_g^{-1}(E)) \\ &= 1 - \mathcal{P}T_g[\nu](E) = 1 - \nu(E) = \nu(X \setminus E),\end{aligned}$$

and so $X \setminus E \in \mathscr{L}$. Finally, let $\{E_n\}_{n \in \mathbb{N}} \subset \mathscr{L}$ be disjoint, so we have $\mathcal{P}T_g[\nu](E_n) = \nu(E_n)$ for every $n \in \mathbb{N}$. Then by the countable additivity of ν and $\mathcal{P}T_g(\nu)$, we have

$$\mathcal{P}T_g[\nu]\left(\bigsqcup_{n \in \mathbb{N}} E_n\right) = \sum_{n \in \mathbb{N}} \mathcal{P}T_g[\nu](E_n) = \sum_{n \in \mathbb{N}} \nu(E_n) = \nu\left(\bigsqcup_{n \in \mathbb{N}} E_n\right),$$

and so $\bigsqcup_{n \in \mathbb{N}} E_n \in \mathscr{L}$. As such, $\mathscr{L}$ is a λ-system.

Now, let $\mathscr{P}$ be the set of all finite cylinders, and let $E \in \mathscr{P}$ so that for $J \subset I$ finite and $E_j \in \mathscr{A}_X$ for each $j \in J$, we have

$$E = \bigcap_{j \in J} \pi_j^{-1}(E_j).$$

Then, by definition of ν as a product measure, we have

$$\nu(E) = \nu\left(\bigcap_{j \in J} \pi_j^{-1}(E_j)\right) = \prod_{j \in J} \nu_j(E_j) = \prod_{j \in J} \mu(E_j).$$

Additionally, we have

$$T_g^{-1}(E) = T_g^{-1}\left(\bigcap_{j \in J} \pi_j^{-1}(E_j)\right) = \bigcap_{j \in J} \pi_{g^{-1}(j)}(E_j),$$

and therefore

$$\mathcal{P}T_g[\nu](E) = \nu(T_g^{-1}(E)) = \nu\left(\bigcap_{j \in J} \pi_{g^{-1}(j)}(E_j)\right) = \prod_{j \in J} \nu_{g^{-1}(j)}(E_j) = \prod_{j \in J} \mu(E_j).$$

This gives that $\mathcal{P}T_g[\nu](E) = \nu(E)$, and so $E \in \mathscr{L}$. As $E \in \mathscr{P}$ was arbitrary, we have shown $\mathscr{P} \subset \mathscr{L}$. By the $\pi - \lambda$ Theorem, the σ-algebra generated by $\mathscr{P}$ (which is $\mathscr{A}_{X^I}$) must be contained in $\mathscr{L} \subset \mathscr{A}_{X^I}$, and so $\mathscr{L} = \mathscr{A}_{X^I}$. This proves that $\mathcal{P}T_g[\nu] = \nu$, and since $T_g \in \mathcal{S}$ was arbitrary, we have shown that $\nu \in \mathcal{I}_{X^I}(\mathcal{S})$. □

Not much more can be said about what sorts of invariant measures exist for semi-contractible systems without some additional assumptions. We begin with the following relation.

Definition 2.5.16. Let I be a countable set and $M \subset \Theta(I)$ a monoid. Define a relation $\sim_M$ on I by $i \sim_M j$ if there exists $g \in M$ such that $g(i) = j$ or $g(j) = i$. This relation is clearly both reflexive (since $\mathrm{id}_I \in M$) and symmetric. Let $\approx_M$ denote the transitive closure of $\sim_M$, which is then clearly an equivalence relation.

This equivalence relation is a well known orbit equivalence relation, where $i \approx_M j$

if i and j are in the same orbit (in some sense, see Proposition A.3 of [39]). We may now use this equivalence relation to define an important property of semicontractible systems.

Definition 2.5.17. Let $(X^I, \mathcal{S}_M)$ be a semicontractible system. Then $(X^I, \mathcal{S}_M)$ is *transitive* if for every $i, j \in I$, it holds that $i \approx_M j$.

Next, we prove the core property of invariant measures for transitive semicontractible systems.

Lemma 2.5.18. *Let $(X^I, \mathcal{S}_M)$ be a semicontractible system, and let $\mu \in \mathcal{I}(\mathcal{S})$. If $i \sim_M j$ or $i \approx_M j$, we have that $\mu_i = \mu_j$. As such, if $(X^I, \mathcal{S})$ is transitive, then for every $i, j \in I$, we have $\mu_i = \mu_j$.*

Proof. Let $M \subset \Theta(I)$ be such that $\mathcal{S} = \mathcal{S}_M$. First, suppose that $i \sim_M j$ so there exists $g \in G$ such that $g(i) = j$ or $g(j) = i$. Without loss of generality, suppose that $g(i) = j$ by swapping the roles of i and j if necessary. For $E \in \mathscr{A}_X$, and $k \in I$ we have

$$\mu_k(E) = \mathcal{P}\pi_j[\mu](E) = \mu(\pi_k^{-1}(E)),$$

and note that $\pi_k^{-1}(E)$ is a cylinder set. Since $T_g \in \mathcal{G}$, we have that $\mu = \mathcal{P}T_g(\mu)$, and thus

$$\mu(\pi_j^{-1}(E)) = \mathcal{P}T_g[\mu](\pi_j^{-1}(E)) = \mu(T_g^{-1}(\pi_j^{-1}(E))).$$

Note then that $g^{-1}(j) = i$, and so

$$T_g^{-1}(\pi_j^{-1}(E)) = \pi_{g^{-1}(j)}^{-1}(E) = \pi_i^{-1}(E),$$

and so

$$\mu_j(E) = \mu(\pi_i^{-1}(E)) = \mathcal{P}\pi_i[\mu](E) = \mu_i(E).$$

As $E \in \mathscr{A}_X$ was arbitrary, we have that $\mu_i = \mu_j$.

Now, the map $\phi(i) = \mu_i$ is a surjective homomorphism from the relation $(I, \sim_M)$ onto $(\{\mu_i\}_{i\in I}, =)$. Thus, we have that $(\phi(I), \phi(\sim_M))$ is a relation contained in the equivalence relation $(\{\mu_i\}_{i\in I}, =)$, and thus its transitive closure must also be contained within it. As such, it must be that ϕ is also a surjective homomorphism from the relation $(I, \approx_M)$ onto $(\{\mu_i\}_{i\in I}, =)$, and as such if $i \approx_M j$, then $\mu_i = \mu_j$.

Finally, if $(X^I, \mathcal{S})$ is transitive, then for every $i, j \in I$ we have $i \approx_M j$, and therefore $\mu_i = \mu_j$. □

Beyond stating that the invariant measures for transitive semicontractible systems are identically distributed, it is difficult to describe in any generality what sort of dependence structures may exist between indices. Despite this, it is still possible to in some sense decompose semicontractible systems into transitive semicontractible systems. We first define the equivalence classes of equivalence relation $\approx_M$.

Definition 2.5.19. Let I be a countable set, and $M \subset \Theta(I)$ a monoid. Let $O_M(I)$ denote the collection of equivalence classes of $\approx_M$.

Note that $O_M(I)$ is always countable, since I is countable. Using these equivalence classes, we may describe a method for decomposing semicontractible systems into transitive ones.

Theorem 2.5.20. *Let $(X^I, \mathcal{S}_M)$ be a semicontractible system. Then $(X^I, \mathcal{S}_M)$ is a joining of the transitive semicontractible systems $(X^J, \mathcal{S}_{M^J})$ for $J \in O_M(I)$, where*

$$M^J = \{g|_J : g \in M\}.$$

Furthermore, if $\mathcal{S}$ is a countable amenable group, then so is each $\mathcal{S}_{M^J}$.

Proof. First, since $O_M(I)$ is a partition of I, we have that

$$X^I = \prod_{i\in I} X = \prod_{J\in O_M(I)} X^J.$$

Next, let $\phi : M \to \Theta(J)$ be defined by $\phi(g) = g|_J$. This map is well defined, because for $g \in M$ and $j \in J$, we have that $j \approx_M g(j)$ since $g(j) = g(j)$, and therefore g maps J into itself since J is an equivalence class, so $g|_J \in \Theta(J)$. Additionally, for $g, h \in M$, we have that $g \circ h \in M$ by the monoidal structure of M, and so

$$\phi(g \circ h) = (g \circ h)|_J = g|_J \circ h|_J = \phi(g) \circ \phi(h),$$

and so ϕ is a monoidal homomorphism. As such, $\phi(M)$ is a monoid in $\Theta(I)$. If on the other hand M is a group, by the first isomorphism theorem for groups, the set $N = \{g \in M : \phi(M) = \mathrm{id}_J \in \Pi(J)\}$ is a normal supgroup of M, that

$$M^J = \phi(M) = \{\phi(g) : g \in M\} = \{g|_J : g \in M\}$$

is a subgroup of $\Pi(J)$, and that M^J is isomorphic as a group to the quotient group M/N. This last property gives that M^J is amenable, since $\mathcal{S}$ is isomorphic to M, so M is amenable, and the quotient of an amenable group is amenable. Otherwise, in general we have that M^J is a monoid in $\Theta(J)$, and so $(X^J, \mathcal{S}_{M^J})$ is a semicontractible system. Also, since $J \in O_M(I)$, we have for all $j, k \in J$ that $j \approx_M k$, which gives that $(X^J, \mathcal{S}_{M^J})$ is transitive.

Finally, all that remains is to show that $(X^I, \mathcal{S})$ is a joining of these systems. Indeed, for $T_g \in \mathcal{S}_M$, we have that $g|_J \in \Theta(J)$ for every $J \in O_M(I)$, so with $T_{g|_J}$ a map on X^J for each $J \in O_M(I)$, we have that

$$T_g = \prod_{J \in O_M(I)} T_{g|_J},$$

and thus

$$\mathcal{S} \subset \prod_{J \in O_M(I)} \mathcal{S}_{M^J}.$$

Then, for $J \in O_M(I)$ and $T_h \in \mathcal{S}_{M^J}$, there exists $g \in M$ such that $g|_J = h$, and so with $\pi_J : X^I \to X^J$ the canonical projection, we have

$$\pi_J \circ T_g = T_{g|_J} \circ \pi_J = T_h \circ \pi_J,$$

and therefore $(X^I, \mathcal{S})$ is a joining of the systems $(X^J, \mathcal{S}_{M^J})$. □

With this theorem, the task of characterizing the invariant measures for semicontractible systems reduces to characterizing the measures of those which are transitive, and then finding ways to join them together as the marginals of a measure on the joining of these systems. For a generic transitive permutation system (let alone transitive semicontractible systems), there may be an incredibly wide variety of ergodic measures. For instance, in the case that $I = \mathbb{Z}$ and $M \leq \Pi(\mathbb{Z})$ is the map generated by $g(z) = z + 1$, the transitive permutation system $(X^I, \mathcal{S}_M)$ has the property that $\mathcal{E}_{X^I}(\mathcal{S}_M)$ is a residual set in $\mathcal{I}_{X^I}(\mathcal{S}_M)$ (A dense countable intersection of dense open sets) [13]. As such there are far too many ergodic measures to fully characterize, at least for general systems. On the other hand, there are certain transitive permutation systems where the set of ergodic measures can be concisely described, which we present in the following subsections.

2.5.4 An extension of De Finetti's Theorem

Although initially proven by De Finetti almost a century ago using probabilistic tools and very much not with an eye on describing the invariant measures of semicontractible systems, De Finetti's Theorem is fundamentally such a characterization [16]. Several generalizations, extensions, and proofs of the theorem exists in the literature, such as Hewitt and Savage's result [25], for which there are many proofs such as that given in [23] (which is the most similar to the one given here), as well as category theoretical proofs of the result [20, 18]. Aldous also discusses variations of the theorem for more general contexts, and also proposed the characterization problem for partially exchangeable processes [2]. In this section, we prove an extension of De Finetti's Theorem which gives a characterization of exchangeability in terms of both completions and dynamical independence, giving a complete characterization of all semicontractible systems for which the conclusion of De Finetti's Theorem holds. In order to state this theorem, we need the following definition.

Definition 2.5.21. Let I be a countable set. For $g \in \Pi(I)$, define the support of g to be

$$\operatorname{supp}(g) = \{i \in I : g(i) \neq i\},$$

and define $\pi(I) \leq \Pi(I)$ as

$$\pi(I) = \{g \in \Pi(I) : |\operatorname{supp}(g)| < \infty\}.$$

$\pi(I)$ is clearly a group as $\operatorname{supp}(e) = \varnothing$, $\operatorname{supp}(g) = \operatorname{supp}(g^{-1})$ and $\operatorname{supp}(g \circ h) \subset \operatorname{supp}(g) \cup \operatorname{supp}(h)$.

We now state the theorem, then demonstrate how the original De Finetti's Theorem is a corollary, and proceed to prove it in several lemmas afterwards.

Theorem 2.5.22. *Let $(X^I, \mathcal{S})$ be a semicontractible system with I infinite. Then the following are equivalent.*

(a) $\mathcal{S}_{\pi(I)} \subset \mathcal{S}^*$,

(b) $\mathcal{S}_{\Pi(I)} \subset \mathcal{S}^*$,

(c) $\mathcal{S}_{\Theta(I)} \subset \mathcal{S}^*$,

(d) $\mathcal{S}^* = (\mathcal{S}_{\pi(I)})^*$, *and $(X^I, \mathcal{S})$ is a Birkhoff system such that for every pair of finite cylinder sets $E = \bigcap_{j \in J} \pi_j^{-1}(E_j)$ and $F = \bigcap_{k \in K} \pi_k^{-1}(F_k)$ ($E_j, F_k \in \mathscr{A}_X$) where J and K are disjoint, E is $(\mu, \mathcal{S})$-dynamically independent of F,*

(e) the map $\Psi : \mathcal{P}X \to \mathcal{E}_{X^I}(\mathcal{S})$ *defined by*

$$\Psi(\mu) = \bigoplus_{i \in I} \mu = \mu^I$$

is a well defined homeomorphism, and

(f) the map $\beta \circ \mathcal{P}\Psi : \mathcal{P}^2 X \to \mathcal{I}_{X^I}(\mathcal{S})$ *is a well defined affine homeomorphism.*

This theorem of course, does not necessarily claim that there are any permutation systems which satisfy any of these constraints, however there is one permutation system which rather trivially satisfies the first statement, and this is precisely the permutation system for which De Finetti's original theorem applies.

Corollary 2.5.23 (De Finetti's Theorem)**.** *For the permutation system* $(X^I, \mathcal{G}_{\pi(I)})$*, the map* $\beta \circ \mathcal{P}\Psi$ *is an affine homeomorphism from* $\mathcal{P}^2 X$ *to* $\mathcal{I}_{X^I}(\mathcal{G}_{\pi(I)})$*.*

Proof. By Theorem 2.3.26 we have $\mathcal{G}_{\pi(I)} \subset (\mathcal{G}_{\pi(I)})^*$, and so Theorem 2.5.22 gives us the desired result. □

This statement in particular is a modern restatement of De Finetti's Theorem as can be found in [23]. In reference to the classical statement of De Finetti's Theorem, we give the following definition.

Definition 2.5.24. Any permutation system $(X^I, \mathcal{G})$ which satisfies any of the equivalent conditions of Theorem 2.5.22 (as well as I being infinite) is called an *exchangeable system*. Any semicontractible system which satisfies any of the equivalent conditions of Theorem 2.5.22 (as well as I being infinite) is called a *contractible system*.

The second term defined above refers to contractible sequences, which is equivalent to exchangeability (Theorem 6.1A [2]). We do not explore contractible sequences directly here, since they have already shown to be equivalent to exchangeable ones. Before proving Theorem 2.5.22, we prove the following fact about contractible systems.

Lemma 2.5.25. *Let $(X^I, \mathcal{S})$ be a contractible system where $|X| > 1$. Then $(X^I, \mathcal{S})$ is transitive.*

Proof. By Theorem 2.5.22 we have that $(\mathcal{S}_M)^* = (\mathcal{S}_{\pi(I)})^*$, and that $\beta \circ \mathcal{P}\Psi : \mathcal{P}^2 X \to \mathcal{I}_{X^I}(\mathcal{S}_M)$ is a well defined affine homeomorphism. By Theorem 2.5.20, $(X^I, \mathcal{S}_M)$ is a joining of the transitive semicontractible systems $(X^J, \mathcal{S}_{M^J})$ for $J \in O_M(I)$. Choose some $J \in O_M(I)$, and suppose we have some $K \in O_M(I) \setminus \{J\}$. Since X has at least two points, $\mathcal{P}X$ does as well, and so let $\mu_1 \neq \mu_2 \in \mathcal{P}X$. By Lemma 2.5.15 we have that $\nu_J = \prod_{j \in J} \mu_1$ is in $\mathcal{I}_{X^J}(\mathcal{S}_{M^J})$, and for all $K \in O_M(I) \setminus \{J\}$, let $\nu_K = \prod_{k \in K} \mu_2$. Then, by Proposition 2.5.6, we have that $\nu = \bigotimes_{J \in O_M(I)} \nu_J$ is in $(X^I, \mathcal{S}_M)$. This however contradicts that $\beta \circ \mathcal{P}\Psi$ is a homeomorphism, since ν is not in the image of Ψ because $\mu_1 \neq \mu_2$. As such, it must be that there is not some $K \in O_M(I) \setminus \{J\}$, meaning $O_M(I) = \{J\}$. As such, $(X^I, \mathcal{S}_M)$ is transitive. $\square$

We now turn our attention to proving Theorem 2.5.22 by proving that

$$\text{(a)} \implies \text{(c)} \implies \text{(b)} \implies \text{(a)},$$

and then that

$$\text{(c)} \implies \text{(d)} \implies \text{(e)} \implies \text{(f)} \implies \text{(a)},$$

which completes the proof. For the first of these chains, it is clear that (c) $\implies$ (b) $\implies$ (a), as $\pi(I) \subset \Pi(I) \subset \Theta(I)$, and therefore $\mathcal{S}_{\pi(I)} \subset \mathcal{S}_{\Pi(I)} \subset \mathcal{S}_{\Theta(I)} \subset \mathcal{S}^*$, so it suffices to prove that (a) $\implies$ (c) below.

Lemma 2.5.26 ((a) $\implies$ (c))**.** *Let $(X^I, \mathcal{S})$ be a semicontractible system with I infinite. If $\mathcal{S}_{\pi(I)} \subset \mathcal{S}^*$, then $\mathcal{S}_{\Theta(I)} \subset \mathcal{S}^*$.*

Proof. Let us identify I with $\mathbb{N}$ by some enumeration of I, and let $g \in \Theta(\mathbb{N})$ be a bijection $g : J \to \mathbb{N}$ for some $J \subset \mathbb{N}$. For $n \in \mathbb{N}$, define g_n so that $g_n(i) = g^{-1}(i)$ for every $i \in [1, n]$. Define $K = [1, n] \cup g^{-1}([1, n])$, and let $K_0 = K \setminus [1, n]$ and

$K_1 = K \setminus g^{-1}([1,n])$. For $i \in \mathbb{N} \setminus K$, define $g_n(i) = i$. By definition, we have that g^{-1} is a bijection from $[1,n]$ to $g^{-1}([1,n])$, so these sets have the same cardinality, and therefore we must have that $|K_0| = |K_1|$. As such, there exists a bijection h from K_0 to K_1, and define $g_n(i) = h(i)$ for every $i \in K_0$. Thus, g_n is a bijection from $[1,n]$ to $g^{-1}([1,n])$, and from K_0 to K_1, and since $[1,n] \cap K_0 = \varnothing$ and $K_0 \cup [1,n] = K$, and also $g^{-1}([1,n]) \cap K_1 = \varnothing$ and $g^{-1}([1,n]) \cup K_1 = K$, we have that g_n is a bijection from K to K. Furthermore, $|\operatorname{supp}(g_n)| \leq |K| < 2n$, and therefore $g_n \in \pi(I)$. Also, since $\pi(I)$ is a group, we have $g_n^{-1} \in \pi(I)$ Additionally, for any $i \in \mathbb{N}$ and all $n \geq i$, we have by definition that $g_n(i) = g^{-1}(i)$, and therefore

$$\lim_{n\to\infty} g_n(i) = g^{-1}(i).$$

As a result, for any $x \in X^{\mathbb{N}}$ and $i \in \mathbb{N}$, we have

$$\lim_{n\to\infty} [T_{g_n^{-1}}(x)]_i = \lim_{n\to\infty} x_{g_n(i)} = x_{g^{-1}(i)} = [T_g(x)]_i,$$

and therefore $\{T_{g_n^{-1}}\}_{n\in\mathbb{N}}$ converges pointwise to T_g. Since $g_n^{-1} \in \pi(I)$ for each $n \in \mathbb{N}$, we have that $T_{g_n^{-1}} \in \mathcal{S}_{\pi(I)} \subset \mathcal{S}^*$. By Proposition 2.3.30, we must then have that $T_g \in \mathcal{S}^*$. Since $g \in \Theta(I)$ was arbitrary, we have shown that $\mathcal{S}_{\Theta(I)} \subset \mathcal{S}^*$. □

Now, we move on to prove the next chain of implications. We first need the following two lemmas.

Lemma 2.5.27. *Let I be a countable set. Then $\pi(I)$ is countable, locally finite, and amenable.*

Proof. For $J \subset I$ finite, it is clear that the set $\{g \in \pi(I) : \operatorname{supp}(g) = J\}$ is finite, and since there are only countably many such J, it follows that $\pi(I)$ is countable.

Next, let $g_1, g_2, \ldots, g_n \in \pi(I)$. Let $J = \bigcup_{k=1}^n \operatorname{supp}(g_k)$. Since $\operatorname{supp}(g_i)$ is a finite set for every i, it follows that J is also finite. It is also then clear than for any $i \in I \setminus J$

that $g_k(i) = i$, and so the group generated by $g_1, g_2, \ldots, g_n$ has elements with support contained in J, making $\pi(I)$ locally finite.

Finally, every countable locally finite group is amenable. $\square$

Lemma 2.5.28. *The permutation system $(X^I, \mathcal{S}_{\pi(I)})$ is transitive.*

Proof. For any $i, j \in I$, let $g \in \Pi(I)$ be defined by $g(i) = j$, $g(j) = i$, and for every $k \in I \setminus \{i, j\}$, $g(k) = k$. Then $\operatorname{supp}(g) = \{i, j\}$, so $g \in \pi(I)$. As such, there exists $g \in \pi(I)$ such that $g(i) = j$, and so $i \sim_{\pi(I)} j$, so $(X^I, \mathcal{G})$ is transitive. $\square$

With this, we can now prove the first of the implications in the second chain. This result is a rather strong example of the uses of completions of dynamical systems and using them to transfer pointwise ergodic theorems from one system to another. In some sense, this Lemma is the core of the proof of De Finetti's Theorem, and it barely mentions measures except at the very end.

Lemma 2.5.29 ((c) $\Longrightarrow$ (d))**.** *Let $(X^I, \mathcal{S})$ be a semicontractible system with I infinite. If $\mathcal{S}_{\Theta(I)} \subset \mathcal{S}^*$, then $\mathcal{S}^* = (\mathcal{S}_{\pi(I)})^*$, and $(X^I, \mathcal{S})$ is a Birkhoff system such that for every pair of finite cylinder sets $E = \bigcap_{j \in J} \pi_j^{-1}(E_j)$ and $F = \bigcap_{k \in K} \pi_k^{-1}(F_k)$ ($E_j, F_k \in \mathscr{A}_X$) where J and K are disjoint, E is $(\mu, \mathcal{S})$-dynamically independent of F.*

Proof. First, note that by Theorem 2.3.26 we have $\mathcal{S}_{\pi(I)} \subset (\mathcal{S}_{\pi(I)})^*$, and thus by Lemma 2.5.26, it holds that $\mathcal{S}_{\Theta(I)} \subset (\mathcal{S}_{\pi(I)})^*$. Also, since $\pi(I) \subset \Theta(I)$, we have that $\mathcal{S}_{\pi(I)} \subset \mathcal{S}_{\Theta(I)} \subset (\mathcal{S}_{\pi(I)})^*$. Next, $(X^I, \mathcal{S})$ is semicontractible, there exists a monoid $M \subset \Theta(I)$ for which $\mathcal{S} = \mathcal{S}_M$. As such, we have $\mathcal{S}_M \subset \mathcal{S}_{\Theta(I)} \subset (\mathcal{S}_M)^*$. Using Theorem 2.3.26, we then have that

$$(\mathcal{S}_{\pi(I)})^* \subset (\mathcal{S}_{\Theta(I)})^* \subset ((\mathcal{S}_{\pi(I)})^*)^* = (\mathcal{S}_{\pi(I)})^*,$$

and also that

$$\mathcal{S}^* = (\mathcal{S}_M)^* \subset (\mathcal{S}_{\Theta(I)})^* \subset ((\mathcal{S}_M)^*)^* = (\mathcal{S}_M)^* = \mathcal{S}^*,$$

and therefore

$$\mathcal{S}^* = (\mathcal{S}_{\Theta(I)})^* = (\mathcal{S}_{\pi(I)})^*.$$

As such, Proposition 2.3.30 also gives that

$$\mathcal{I}_{X^I}(\mathcal{S}) = \mathcal{I}_{X^I}(\mathcal{S}^*) = \mathcal{I}_{X^I}((\mathcal{S}_{\pi(I)})^*) = \mathcal{I}_{X^I}(\mathcal{S}_{\pi(I)}).$$

By Lemma 2.3.7 we have that $\mathcal{I}_{X^I}(\mathcal{S}_{\pi(I)})$ is closed, and since $\pi(I)$ is countable and amenable by Lemma 2.5.27, so is $\mathcal{S}_{\pi(I)}$, and therefore by Theorem 2.4.13, any tempered Følner sequence $\mathcal{A}$ for $\mathcal{S}_{\pi(I)}$ is a Birkhoff sequence for $(X^I, \mathcal{S}_{\pi(I)})$. By Proposition 2.4.10, $\mathcal{A}$ is also a Birkhoff sequence for $(X^I, (\mathcal{S}_{\pi(I)})^*) = (X^I, \mathcal{S}^*)$, and hence is also a Birkhoff sequence for $(X^I, \mathcal{S})$, which proves that $(X^I, \mathcal{S})$ is a Birkhoff system.

Let J and K be disjoint finite subsets of I, and let $E_j, F_k \in \mathscr{A}_X$ for every $j \in J$ and $k \in K$. Then

$$E = \bigcap_{j\in J} \pi_j^{-1}(E_j) \quad \text{and} \quad F = \bigcap_{k\in K} \pi_k^{-1}(E_k)$$

are finite cylinders such that J and K are disjoint, as in the statement of this lemma.

Let us now define the set $H = \{g \in \pi(I) : g^{-1}(J) \cap K = \varnothing\}$, and let $D \subset \pi(I)$ be finite. Define $L = J \cup \bigcup_{h\in D} \operatorname{supp}(h)$, which is finite because D is finite and each $\operatorname{supp}(h)$ is finite. Enumerate $K = \{k_1, k_2, \ldots, k_m\}$, which is finite, and $I \setminus L = \{i_1, i_2, \ldots\}$ (which is infinite since L is finite), and define $g(k_l) = i_l$ for $l \in [1, m]$. Also, since $J \cap K = \varnothing$, and $J \subset L$, so $J \cap (I \setminus L) = \varnothing$, we may define $g(j) = j$ for $j \in J$. As such, g as defined is a bijection from $J \cup K$ to $J \cup h(K)$. We may then extend this to a bijection from $J \cup K \cup h(K)$ to itself by arbitrary mapping the remaining elements to each other. By construction, we then have that $g(K) \cap L = \varnothing$, but $g(j) = j$ for all $j \in J$. Then for $h \in D$, we have for $j \in J$ that $[h \circ g]^{-1}(j) = g^{-1}(h^{-1}(j))$. We either have that $h^{-1}(j) \in J$ which gives that $h^{-1}(j) \in L$, or that $h^{-1}(j) \neq j$, in which case $h(h^{-1}(j)) = j \neq h^{-1}(j)$, and therefore $h^{-1}(j) \in \text{supp}(h) \subset L$. As such, we have that $g^{-1}(h^{-1}(j)) \in g^{-1}(L)$. But since $g(K) \cap L = \varnothing$, we have $K \cap g^{-1}(L) = g^{-1}(\varnothing) = \varnothing$, and therefore $g^{-1}(h^{-1}(j)) \notin K$. As this holds for all $j \in J$, we have shown that $[h \circ g]^{-1}(J) \cap K = \varnothing$, and so $h \circ g \in H$. Since $h \in D$ was arbitrary, we have shown that $D \circ g \subset H$, and since $D \subset \pi(I)$ was an arbitrary finite set, we have shown that H is thick.

Then, for $g \in H$, we have

$$T_g^{-1}(E) \cap F = T_{g^{-1}} \left(\bigcap_{j \in J} \pi_j^{-1}(E_j) \right) \cap F = \bigcap_{j \in J} \pi_{g^{-1}(j)}^{-1}(E_j) \cap F.$$

Define $h \in \pi(I)$ by $h(g^{-1}(j)) = j$, and hence $h^{-1}(j) = g^{-1}(j)$, for every $j \in J$. Since g^{-1} is a bijection, h is a bijection from $g^{-1}(J)$ to J. Extend h to a bijection from $g^{-1}(J) \cup J$ to itself by arbitrary assigning the remaining value, and set h to be the identity on all $i \in I \setminus (g^{-1}(J) \cup J)$. Also, since $g \in H$, we have $g^{-1}(J) \cap K = \varnothing$, and since $J \cap K = \varnothing$ by assumption, we have that $K \subset I \setminus (g^{-1}(J) \cup J)$, and so h fixes every point of K. We then have

$$\begin{aligned} T_h^{-1}(E \cap F) &= T_{h^{-1}} \left(\bigcap_{j \in J} \pi_j^{-1}(E_j) \cap \bigcap_{k \in K} \pi_k^{-1}(F_k) \right) \\ &= \bigcap_{j \in J} \pi_{h^{-1}(j)}(E_j) \cap \bigcap_{k \in K} \pi_{h^{-1}(k)}(F_k) \\ &= \bigcap_{j \in J} \pi_{g^{-1}(j)}(E_j) \cap \bigcap_{k \in K} \pi_k(F_k) \\ &= \bigcap_{j \in J} \pi_{g^{-1}(j)}^{-1}(E_j) \cap F, \end{aligned}$$

and thus $T_g^{-1}(E) \cap E = T_h^{-1}(E \cap F)$, so $T_g \in I(E, F)$. As $T_g \in \mathcal{S}_H$ was arbitrary, this shows that $\mathcal{G}_H \subset I(E,F)$, so H being thick in $\pi(I)$ implies that $\mathcal{G}_H$ is thick in $\mathcal{S}_{\pi(I)}$. As such, E is dynamically independent of F.

Then, for $\mu \in \mathcal{I}_{X^I}(\mathcal{S})$, Lemma 2.3.25 gives us that

$$\mathcal{I}_{X^I}(\mathcal{S}) = \mathcal{I}_{X^I}(\mathcal{S}^*) = \mathcal{I}_{X^I}((\mathcal{S}_{\pi(I)})^*) = \mathcal{I}_{X^I}(\mathcal{S}_{\pi(I)}),$$

and so is E is $(\mu, \mathcal{S}_{\pi(I)})$-dynamically independent of F by Proposition 2.4.26. By Proposition 2.4.20, this implies E is $(\mu, (\mathcal{S}_{\pi(I)})^*) = (\mu, \mathcal{S}^*)$-dynamically independent of F, and therefore E is $(\mu, \mathcal{S})$-dynamically independent of F. □

Next, we may move directly to prove the next implication.

Lemma 2.5.30 ((d) $\implies$ (e))**.** *Let $(X^I, \mathcal{S})$ be a semicontractible system with I infinite. Also, suppose that $\mathcal{S}^* = (\mathcal{S}_{\pi(I)})^*$, and that $(X^I, \mathcal{S})$ is a Birkhoff system such that for every pair of finite cylinder sets $E = \bigcap_{j \in J} \pi_j^{-1}(E_j)$ and $F = \bigcap_{k \in K} \pi_k^{-1}(F_k)$ $(E_j, F_k \in \mathscr{A}_X)$ where J and K are disjoint, E is $(\mu, \mathcal{S})$-dynamically independent of F.. Then the map $\Psi : \mathcal{P}X \to \mathcal{E}_{X^I}(\mathcal{S})$ defined by*

$$\Psi(\mu) = \bigoplus_{i \in I} \mu = \mu^I$$

is a well defined homeomorphism.

Proof. First, we show that Ψ is well defined, by showing that for $\mu \in \mathcal{P}X$ we have that $\mu^I \in \mathcal{E}_{X^I}(\mathcal{S})$. Since Proposition 2.3.31 gives us that

$$\mathcal{E}_{X^I}(\mathcal{S}) = \mathcal{E}_{X^I}(\mathcal{S}^*) = \mathcal{E}_{X^I}((\mathcal{S}_{\pi(I)})^*) = \mathcal{E}_{X^I}(\mathcal{S}_{\pi(I)}),$$

and so it will suffice to show that $\mu^I \in \mathcal{E}_{X^I}(\mathcal{S}_{\pi(I)})$. The proof of this follows the proof of De Finetti's Theorem in [23]. First, note that $\mu^I \in \mathcal{I}_{X^I}(\mathcal{S}_{\pi(I)})$ by Lemma 2.5.15, and take any enumeration $I = \{i_n : n \in \mathbb{N}\}$ of I and for each $n \in \mathbb{N}$, let $I_n = \{i_1, i_2, \ldots, i_n\}$. With $\pi_{I_n} : X^I \to X^{I_n}$ the canonical projection map, let $E_n = \pi_{I_n}^{-1}(\pi_{I_n}(E)) \in \mathscr{A}_{X^I}$ so that $E \subset E_n$ for every n, and additionally, $E_{n+1} \subset E_n$. Furthermore, if $x \in \bigcap_{n\in\mathbb{N}} E_n$, then $\pi_{I_n}(x) \in \pi_{I_n}(E)$ for all $n \in \mathbb{N}$, and so it must be that $x \in E$. As such, $E = \bigcap_{n\in\mathbb{N}} E_n$, where the intersection is descending. By the continuity of μ^I, we have

$$\lim_{n\to\infty} \mu^I(E_n \setminus E) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mu^I(E_n) = \mu^I(E).$$

Next, for $n \in \mathbb{N}$, let define $g_n \in \pi(I)$ by $g_n(i_k) = i_{k+n}$ and $g_n(i_{k+n}) = i_k$ for every $i_k \in I_n$, and $g_n(i) = i$ for all $i \in I \setminus I_{2n}$. Then clearly $|\operatorname{supp}(g_n)| < \infty$, and so $g_n \in \pi(I)$, which means that $T_n = T_{g_n} \in \mathcal{S}_{\pi(I)}$. Furthermore, $g_n^{-1} = g_n$, and so $T_n^{-1} = T_n$. By the invariance of μ^I, we have

$$\begin{aligned} 0 &= \lim_{n\to\infty} \mu^I(E_n \setminus E) = \lim_{n\to\infty} \mathcal{P}T_n[\mu^I](E_n \setminus E) = \lim_{n\to\infty} \mu^I(T_n^{-1}(E_n \setminus E)) \\ &= \lim_{n\to\infty} \mu^I(T_n(E_n) \setminus T_n(E)) = \lim_{n\to\infty} \mu^I(T_n(E_n) \triangle T_n(E)). \end{aligned}$$

Also, with $T_n = T_n^{-1} \in \mathcal{S}_{\pi(I)}$ and $E \in \mathscr{I}(\mu^I, \mathcal{S}_{\pi(I)})$, we have that

$$\mu^I(E \triangle T_n(E)) = \mu^I(E \triangle T_n^{-1}(E)) = 0,$$

and so this gives by the triangle inequality (since $\mu(A\triangle B)$ is a pseudometric on $\mathscr{A}_{X^I}$) that

$$\begin{aligned} 0 &\le \lim_{n\to\infty} \mu^I(T_n(E_n) \triangle E_n) \\ &\le \lim_{n\to\infty} \mu^I(T_n(E_n) \triangle T_n(E)) + \mu^I(T_n(E) \triangle E) + \mu^I(E \triangle E_n) \\ &= 0. \end{aligned}$$

As such, we have

$$\lim_{n\to\infty} \mu^I(T_n(E_n) \cap E_n) = \lim_{n\to\infty} \mu(E_n) = \mu(E).$$

On the other hand, by the construction of g_n, we have that $\pi_{I\setminus I_n}(E_n) = X^{I\setminus I_n}$, and that $\pi_{X\setminus(I_{2n}\setminus I_n)}(T_n(E_n)) = X^{I\setminus I_{2n}\setminus I_n}$, and so E_n and $T_n(E_n)$ are supported on disjoint subsets of I. As such, by the definition of μ^I, we have

$$\mu^I(T_n(E_n) \cap E_n) = \mu^I(T_n(E_n))\mu^I(E_n) = \mathcal{P}T_n^{-1}[\mu^I](E_n)\mu^I(E_n) = [\mu^I(E_n)]^2,$$

and so

$$\lim_{n\to\infty} \mu^I(T_n(E_n) \cap E_n) = \lim_{n\to\infty} [\mu^I(E_n)]^2 = [\mu^I(E)]^2.$$

Therefore $\mu^I(E) = [\mu^I(E)]^2$, which implies that $\mu^I(E) \in \{0, 1\}$ and that $E \in \mathscr{T}(\mu^I)$. Since $E \in \mathscr{I}(\mu^I, \mathcal{S}_{\pi(I)})$ was arbitrary, we have shown that $\mu^I \in \mathcal{E}_{X^I}(\mathcal{S}_{\pi(I)}) = \mathcal{E}_{X^I}(\mathcal{S})$.

Next, we show that Ψ is a surjection, so we show that for every $\nu \in \mathcal{E}_{X^I}(\mathcal{S})$ there exists $\mu \in \mathcal{P}X$ such that $\Psi(\mu) = \nu$. Since $\nu_i = \nu_j$ for every $i, j \in I$, let $\mu = \nu_i$. Now, for $j \in I$ and $E_j \in \mathscr{A}_X$, we have

$$\nu(\pi_j^{-1}(E_j)) = \mathcal{P}\pi_j[\nu](E_j) = \nu_j(E_j) = \mu(E_j).$$

For induction, suppose there exists $n \in \mathbb{N}$ such that whenever $J \subset I$ satisfies $|J| = n$, and we have $E_j \in \mathscr{A}_X$ for each $j \in J$, then

$$\nu\left(\bigcap_{j\in J} \pi_j^{-1}(E_j)\right) = \prod_{j\in J} \mu(E_j).$$

Let $K \subset I$ with $|K| = n + 1$, and let $E_k \in \mathscr{A}_X$ for every $k \in K$. Choose any $i \in K$, and define

$$E = \bigcap_{k \in K \setminus \{i\}} \pi_k^{-1}(E_k)$$

and $F = \pi_i^{-1}(E_i)$. Then clearly E and F are finite cylinders, and $K \setminus \{i\} \cap \{i\} = \varnothing$. Therefore, E is $(\nu, \mathcal{S})$-dynamically independent of F by assumption, and so by Theorem 2.4.22, we have that $\nu(E \cap F) = \nu(E)\nu(F)$. Since we have that $|K \setminus \{i\}| = n$, we have by assumption that

$$\nu(E) = \nu\left(\bigcap_{k \in K \setminus \{i\}} \pi_k^{-1}(E_k)\right) = \prod_{k \in K \setminus \{i\}} \mu(E_k),$$

and from the first display, we have that $\nu(F) = \mu(E_i)$, and so

$$\begin{aligned}\nu\left(\bigcap_{k \in K} \pi_k^{-1}(E_k)\right) &= \nu\left(\bigcap_{k \in K \setminus \{i\}} \pi_k^{-1}(E_k) \cap \pi_i^{-1}(E_i)\right) \\ &= \nu(E \cap F) = \nu(E)\nu(F) \\ &= \mu(E_i) \prod_{k \in K \setminus \{i\}} \mu(E_k) \\ &= \prod_{k \in K} \mu(E_k).\end{aligned}$$

As such, by induction, we have for every finite $J \subset I$ and $E_j \in \mathscr{A}_X$ for every $j \in J$ that

$$\nu\left(\bigcap_{j \in J} \pi_j^{-1}(E_j)\right) = \prod_{j \in J} \mu(E_j).$$

As such, it must be that $\nu = \mu^I = \Psi(\mu)$, so Ψ is a surjection.

Finally, Proposition 2.1.6 gives that Ψ is also a continuous injection, and is therefore a continuous bijection from $\mathcal{P}X$, which is compact, to $\mathcal{E}_{X^I}(\mathcal{S})$, which is Hausdorff, and therefore Ψ is a homeomorphism. □

The last two lemmas are rather short, so we prove them one after the other.

Lemma 2.5.31 ((e) $\implies$ (f))**.** *Let $(X^I, \mathcal{S})$ be a semicontractible system with I infinite. If the map $\Psi : \mathcal{P}X \to \mathcal{E}_{X^I}(\mathcal{S})$ defined by*

$$\Psi(\mu) = \bigoplus_{i \in I} \mu = \mu^I$$

is a well defined homeomorphism, then the map $\beta \circ \mathcal{P}\Psi : \mathcal{P}^2 X \to \mathcal{I}_{X^I}(\mathcal{S})$ is a well defined affine homeomorphism.

Proof. Since Ψ is a homeomorphism from $\mathcal{P}X$ to $\mathcal{E}_{X^I}(\mathcal{G})$, this gives us that $\Psi \in \mathcal{C}(\mathcal{P}X, \mathcal{E}_{X^I}(\mathcal{G}))$. By Proposition 2.2.9, $\mathcal{P}\Psi$ is an affine homeomorphism between $\mathcal{P}^2 X$ and $\mathcal{P}\mathcal{E}_{X^I}(\mathcal{G})$. By Proposition 2.3.16, β is an affine homeomorphism of $\mathcal{P}\mathcal{E}_{X^I}(\mathcal{G})$ to $\mathcal{I}_{X^I}(\mathcal{G})$, and so $\beta \circ \mathcal{P}\Psi$ is an affine homeomorphism from $\mathcal{P}^2 X$ to $\mathcal{I}_{X^I}(\mathcal{G})$. □

Lemma 2.5.32 ((f) $\implies$ (a))**.** *Let $(X^I, \mathcal{S})$ be a semicontractible system with I infinite. If the map $\beta \circ \mathcal{P}\Psi : \mathcal{P}^2 X \to \mathcal{I}_{X^I}(\mathcal{S})$ is a well defined homeomorphism, then $\mathcal{S}_{\pi(I)} \subset \mathcal{S}^*$.*

Proof. Let $\mu \in \mathcal{I}_{X^I}(\mathcal{S})$, and let $m \in \mathcal{P}^2 X$ such that $\mu = [\beta \circ \mathcal{P}\Psi](m)$. Then for $T_g \in \mathcal{S}_{\pi(i)}$, let $\mathscr{L} = \{E \in \mathscr{A}_{X^I} : \mathcal{P}T_g[\mu](E) = E\}$, and note that

$$\mathcal{P}T_g[\mu](X) = \mu(T_g^{-1}(X)) = \mu(X) \quad \text{and} \quad \mathcal{P}T_g[\mu](\varnothing) = \mu(T_g^{-1}(\varnothing)) = \mu(\varnothing),$$

and so we have $\varnothing, X \in \mathscr{L}$. Now, for $E \in \mathscr{L}$, we have $\mathcal{P}T_g[\mu](E) = \mu(E)$, and so

$$\begin{aligned} \mathcal{P}T_g[\mu](X \setminus E) &= \mu(T_g^{-1}(X \setminus E)) = \mu(X \setminus T_g^{-1}(E)) = 1 - \mu(T_g^{-1}(E)) \\ &= 1 - \mathcal{P}T_g[\mu](E) = 1 - \mu(E) = \mu(X \setminus E), \end{aligned}$$

and so $X \setminus E \in \mathscr{L}$. Finally, let $\{E_n\}_{n \in \mathbb{N}} \subset \mathscr{L}$ be disjoint, so we have $\mathcal{P}T_g[\mu](E_n) = \mu(E_n)$ for every $n \in \mathbb{N}$. Then by the countable additivity of μ and $\mathcal{P}T_g(\mu)$, we have

$$\mathcal{P}T_g[\mu]\left(\bigsqcup_{n \in \mathbb{N}} E_n\right) = \sum_{n \in \mathbb{N}} \mathcal{P}T_g[\mu](E_n) = \sum_{n \in \mathbb{N}} \mu(E_n) = \mu\left(\bigsqcup_{n \in \mathbb{N}} E_n\right),$$

and so $\bigsqcup_{n\in\mathbb{N}} E_n \in \mathscr{L}$. As such, $\mathscr{L}$ is a λ-system.

Now, let $\mathscr{P}$ be the set of all finite cylinders, and let $E \in \mathscr{P}$ so that for $J \subset I$ finite and $E_j \in \mathscr{A}_X$ for each $j \in J$, we have

$$E = \bigcap_{j\in J} \pi_j^{-1}(E_j).$$

Then, by definition of μ, we have,

$$\mu(E) = [\beta \circ \mathcal{P}\Psi][m](E) = \int_{\mathcal{P}X} \nu^I(E)\, m(\mathrm{d}\nu),$$

and since ν^I is a product measure for each $\nu \in \mathcal{P}X$, we have that

$$\nu^I(E) = \nu^I\left(\bigcap_{j\in J} \pi_j^{-1}(E_j)\right) = \prod_{j\in J} \nu(E_j),$$

which gives

$$\mu(E) = \int_{\mathcal{P}X} \prod_{j\in J} \nu(E_j)\, m(\mathrm{d}\nu).$$

Additionally, we have for $T_g \in \mathcal{S}_{\pi(I)}$ that

$$T_g^{-1}(E) = T_{g^{-1}}(E) = T_{g^{-1}}\left(\bigcap_{j\in J} \pi_j^{-1}(E_j)\right) = \bigcap_{j\in J} \pi_{g^{-1}(j)}(E_j),$$

and we have for $\nu \in \mathcal{P}X$ that

$$\nu^I(T_g^{-1}(E)) = \nu^I\left(\bigcap_{j\in J} \pi_{g^{-1}(j)}(E_j)\right) = \prod_{j\in J} \nu(E_j),$$

and therefore

$$\begin{aligned}\mathcal{P}T_g[\mu](E) &= \mu(T_g^{-1}(E)) = [\beta \circ \mathcal{P}\Psi][m](T_g^{-1}E)\\ &= \int_{\mathcal{P}X} \nu^I(T_g^{-1}(E))\, m(\mathrm{d}\mu) = \int_{\mathcal{P}X} \prod_{j\in J} \nu(E_j)\, m(\mathrm{d}\nu).\end{aligned}$$

This gives that $\mathcal{P}T_g[\mu](E) = \mu(E)$, and so $E \in \mathscr{L}$. As $E \in \mathscr{P}$ was arbitrary, we have shown $\mathscr{P} \subset \mathscr{L}$. By the $\pi - \lambda$ Theorem, the σ-algebra generated by $\mathscr{P}$ (which is $\mathscr{A}_{X^I}$) must be contained in $\mathscr{L} \subset \mathscr{A}_{X^I}$, and so $\mathscr{L} = \mathscr{A}_{X^I}$. This proves that $\mathcal{P}T_g[\nu] = \nu$, and so we have shown that $T_g \in \mathcal{F}(\mu)$. Since $T_g \in \mathcal{S}_{\pi(I)}$ was arbitrary, we have shown $\mathcal{S}_{\pi(I)} \subset \mathcal{F}_{X^I}(\mu)$, and since $\mu \in \mathcal{I}_{X^I}(\mathcal{S})$ was arbitrary, we have shown that $\mathcal{S}_{\pi(I)} \subset \mathcal{F}_{X^I}(\mathcal{I}_{X^I}(\mathcal{S})) = \mathcal{S}^*$. □

With this lemma, we have proven Theorem 2.5.22. Before looking at variations of exchangeability, we first given a few examples of contractible systems outside of $(X^I, \mathcal{S}_{\pi(I)})$. The main purpose of this Proposition is not so much to develop a useful collection of contractible systems, but rather to demonstrate the power of Theorem 2.5.22. We discuss the result more after we prove it.

Proposition 2.5.33. *Let I be a countably inifinite set and X a space. For the following sets $M \subset \Theta(I)$, $(X^I, \mathcal{S}_M)$ is a contractible system.*

(a) $M = \Pi(I)$,

(b) $M = \Theta(I)$,

(c) $M = \theta(I) \subset \Theta(I)$, *the set of elements in $\Theta(I)$ with $\{i \in I : g^{-1}(i) \neq i\}$ finite,*

(d) $M = \alpha(I) \subset \pi(I)$, *the group of alternating permutations with finite support,*

(e) For $I = \{i_1, i_2, \dots\}$, and any monotonically increasing $f : \mathbb{N} \to \mathbb{N}$, define

$$C = \{g \in \Theta(I) : \exists N \in \mathbb{N}, \forall n \geq N, \exists m > f(n), g^{-1}(i_n) = i_m\}$$

to be the set of permutations for which the tail contracts towards i_1 at a minimum rate specified by f, and let M be the monoid generated by C.

Proof. For $M = \Pi(I)$ and $C = \Theta(I)$, we have by Theorem 2.3.26 that

$$\mathcal{S}_{\Pi(I)} \subset (\mathcal{S}_{\Pi(I)})^* \quad \text{and} \quad \mathcal{S}_{\Theta(I)} \subset (\mathcal{S}_{\Theta(I)})^*,$$

and so clearly $(X^I, \mathcal{S}_M)$ is contractible.

For $M = \theta(I)$, we have that $\pi(I) \subset \theta(I)$, and therefore we have by Theorem 2.3.26 that

$$\mathcal{S}_{\pi(I)} \subset (\mathcal{S}_{\pi(I)})^* \subset (\mathcal{S}_{\theta(I)})^*,$$

and so $(X^I, \mathcal{S}_{\theta(I)})$ is contractible.

For $M = \alpha(I)$, let $g \in \pi(I)$. Of course, if $g \in \alpha(I)$, then we have $T_g \in \mathcal{S}_{\alpha(I)} \subset (\mathcal{S}_{\alpha(I)})^*$ by Theorem 2.3.26, so suppose that $g \notin \alpha(I)$. Thus, g decomposes into a product of an odd number of 2-cycles. Let $I = \{i_1, i_2, \dots\}$ be an enumeration of I, and let $J = \mathrm{supp}(g)$. Choose $N \in \mathbb{N}$ such that for all $j \in J$, there is $m < N$ with $j = i_m$. For $n \geq N$, define $h_n(i_n) = i_{n+1}$ and $h_n(i_{n+1}) = i_n$ to be a 2-cycle (so it is the identity elsewhere), and note that $\mathrm{supp}(h_n)$ is always disjoint from $\mathrm{supp}(g)$. Then $g_n = g^{-1} \circ h_n$ decomposes into a product of an even number of 2-cycles, and therefore $g_n \in \alpha(I)$ for every $n \geq N$. Now, for $m < n$, we have $[g^{-1} \circ h_n](i_m) = g^{-1}(i_m)$ because the support of h_n is $\{i_n, i_{n+1}\}$. As such, for every $i \in I$ we have

$$\lim_{n \to \infty} g_n(i) = g^{-1}(i).$$

Now, since $g_n \in \alpha(I)$, we also have $g_n^{-1} \in \alpha(I)$, and so for every $n \geq N$, we have $T_{g_n^{-1}} \in \mathcal{S}_{\alpha(I)}$, and for any $x \in X^I$ and $i \in I$, we have

$$\lim_{n \to \infty} [T_{g_n^{-1}}(x)]_i = \lim_{n \to \infty} x_{g_n(i)} = x_{g^{-1}(i)} = [T_g(x)]_i,$$

and thus $\{T_{g_n^{-1}}\}_{n\geq N}$ is a sequence in $\mathcal{S}_{\alpha(I)} \subset (\mathcal{S}_{\alpha(I)})^*$ (Theorem 2.3.26) which converges pointwise to T_g. By Proposition 2.3.30, we have that $T_g \in (\mathcal{S}_{\alpha(I)})^*$. Since $g \in \pi(I)$ was arbitrary, we have shown that $\mathcal{S}_{\pi(I)} \subset (\mathcal{S}_{\alpha(I)})^*$, and so $(X^I, \mathcal{S}_{\alpha(I)})$ is a contractible system.

Finally, let $I = \{i_1, i_2, \dots\}$ be an enumeration of I, let $f : \mathbb{N} \to \mathbb{N}$ be monotonically increasing, and let

$$C = \{g \in \Theta(I) : \exists N \in \mathbb{N}, \forall n \geq N, \exists m > f(n), g^{-1}(i_n) = i_m\},$$

and let M be the monoid generated by C. Let $g \in \pi(I)$, and let $J = \mathrm{supp}(g)$. Choose $N \in \mathbb{N}$ such that for all $j \in J$, there is $m < N$ with $j = i_m$. For $n \geq N$, define $g_n^{-1}(i_m) = g^{-1}(i_m)$ for every $m < n$, and note that this is a permutation of the set $\{i_1, \dots, i_{n-1}\}$. Now, for each $m \geq n$, let $k_m = f(m) + 1$, and define $g_n^{-1}(i_m) = i_{k_m}$. With f monotonically increasing, it must be that g_n^{-1} is an injection from I to $g_n^{-1}(I)$, so letting $J = g_n^{-1}(I)$, we have that $g_n : J \to I$ is a bijection. Also, for all $m \geq n$, we have that $g_n^{-1}(i_m) = i_{k_m}$ where $k_m > f(m)$, and thus we have that $g_n \in C$. Finally, for every $i_m \in I$, there exists $N > m$, and for every $n \geq N$, we have

$$g_n^{-1}(i_m) = g^{-1}(i_m),$$

and so we have that

$$\lim_{n\to\infty} g_n^{-1}(i_m) = g^{-1}(i_m).$$

Now, since $g_n \in C$ for every $n \geq N$, we have $T_{g_n} \in \mathcal{S}_C$, and for any $x \in X^I$ and $i \in I$, we have

$$\lim_{n\to\infty} [T_{g_n}(x)]_i = \lim_{n\to\infty} x_{g_n^{-1}(i)} = x_{g^{-1}(i)} = [T_g(x)]_i,$$

and thus $\{T_{g_n}\}_{n\geq N}$ is a sequence in $\mathcal{S}_C \subset (\mathcal{S}_C)^*$ (Theorem 2.3.26) which converges pointwise to T_g. By Proposition 2.3.30, we have that $T_g \in (\mathcal{S}_C)^*$. Since $g \in \pi(I)$

was arbitrary, we have shown that $\mathcal{S}_{\pi(I)} \subset (\mathcal{S}_C)^*$. Finally, we have that $\mathcal{S}_C \subset \mathcal{S}_M$ by definition, and so Theorem 2.3.26 gives that $\mathcal{S}_{\pi(I)} \subset (\mathcal{S}_C)^* \subset (\mathcal{S}_M)^*$, and so $(X^I, \mathcal{S}_M)$ is a contractible system. □

The first three sets have the main purpose of demonstrating that it is possible to classify the invariant measures of rather unruly semi-contractible systems. First, we have that $\Pi(I)$ is not a discrete amenable group, as it contains as a subgroup the free group of any finite number of generators, and is also uncountable. $\theta(I)$ is also uncountable and $\Theta(I)$ is a superset. The fourth example is notable because $\alpha(I)$ is a strict subgroup of $\pi(I)$, which would generally imply that the set of invariant measures with respect to $\alpha(I)$ is larger than those for $\pi(I)$, however they end up coinciding. The last class of examples is one where the tail behavior of elements is rather unruly, and is not even a monoid (though we take the monoid generated by this set C). Additionally, we prove that $\mathcal{S}_{\pi(I)} \subset (\mathcal{S}_C)^*$, however we have that $\pi(I) \cap C = \varnothing$, and so this example also demonstrates that C has no tangible relation to $\pi(I)$.

2.5.5 Variations of exchangeable and contractible systems

We now turn our attention to using Theorem 2.5.22 to characterize the invariant measures of systems which are related to contractible systems. We begin with the following slight variation, which is an analog of Proposition 3.8 of [2]. We discuss the connection between it and the statement of Proposition 3.8 following the proof.

Proposition 2.5.34. *Let Y be space, and let $(X^I, \mathcal{G})$ be a countable amenable exchangeable system. Then $Y \times X^I$ is a space, and let*

$$\mathcal{T} = \{\mathrm{id}_Y \times T : T \in \mathcal{G}\}$$

This gives a dynamical system $(Y \times X^I, \mathcal{T})$. Then $\beta \circ \mathcal{P}(\otimes \circ (\delta \times \Psi))$ is an affine homeomorphism from $\mathcal{P}(Y \times \mathcal{P}X)$ to $\mathcal{I}_{Y \times X^I}(\mathcal{T})$.

Proof. First, note that by definition, the system $(Y \times X^I, \mathcal{T})$ is the product of the systems (Y, id_Y) with id_Y the identity on Y, and $(X^I, \mathcal{G})$. Clearly we have $\mathcal{I}_Y(\mathrm{id}_Y) = \mathcal{P}Y$, and since $\delta(Y)$ is the set of extreme points of $\mathcal{P}Y$, we have that $\mathcal{E}_Y(\mathrm{id}_Y) = \delta(Y)$ by Proposition 2.3.15. Note as well that $\{\mathrm{id}_Y\}$ is clearly a countable amenable group and so is $\mathcal{G}$ by assumption, and since each $\mathrm{id}_Y \times T$ for $T \in \mathcal{G}$ is continuous on $Y \times X^I$, we have that $\mathcal{I}_{Y\times X^I}(\mathcal{T})$ is closed by Lemma 2.3.7. By Proposition 2.5.7, we have that

$$\mathcal{E}_{Y\times X^I}(\mathcal{T}) = \mathcal{E}_Y(\mathrm{id}_Y) \otimes \mathcal{E}_{X^I}(\mathcal{G}) = \delta(Y) \otimes \mathcal{E}_{X^I}(\mathcal{G}).$$

By Proposition 2.1.5, we have that $\delta : Y \to \delta(Y)$ is a homeomorphism and by Theorem 2.5.22, we have that $\Psi : \mathcal{P}(X) \to \mathcal{E}_{X^I}(\mathcal{G})$ is a homeomorphism, so $\delta \times \Psi$ is a homeomorphism from $Y \times \mathcal{P}X$ to $\delta(Y) \times \mathcal{E}_{X^I}(\mathcal{G}_{\pi(I)})$. Additionally, Proposition 2.1.6 gives that the map $\otimes : \delta(Y) \times \mathcal{E}_{X^I}(\mathcal{G}) \to \delta(Y) \otimes \mathcal{E}_{X^I}(\mathcal{G})$ is a homeomorphism, and therefore $\otimes \circ (\delta \times \Psi)$ is a homeomorphism from $Y \times \mathcal{P}X$ to $\mathcal{E}_{Y\times X^I}(\mathcal{T})$. By Proposition 2.2.9, we have that $\mathcal{P}(\otimes \circ (\delta \times \Psi))$ is an affine homeomorphism from $\mathcal{P}(Y \times \mathcal{P}(X))$ to $\mathcal{P}(\mathcal{E}_{Y\times X^I}(\mathcal{T}))$. Finally, since $\mathcal{E}_{Y\times X^I}(\mathcal{T})$ is homeomorphic to $Y \times \mathcal{P}X$, it is compact, and so by Proposition 2.3.16, β is an affine homeomorphism from $\mathcal{P}(\mathcal{E}_{Y\times X^I}(\mathcal{T}))$ to $\mathcal{I}_{Y\times X^I}(\mathcal{T})$, whence

$$\beta \circ \mathcal{P}(\otimes \circ (\delta \times \Psi))$$

gives the desired affine homeomorphism. □

First, the requirement that $(X^I, \mathcal{G})$ be an amenable exchangeable system as opposed to an arbitrary contractible system $(X^I, \mathcal{S})$ is due to the lack of a more general characterization of the ergodic measures of product systems, and this will also be the case for the subsequent result. In any case, the statement of Proposition 3.8 of [2] is that for a random variable V and a sequence of random variables $Z_1, Z_2, \ldots,$ if for every $g \in \pi(\mathbb{N})$ it holds that

$$(V, Z_1, Z_2, \dots) \equiv (V, Z_{g(1)}, Z_{g(2)}, \dots),$$

that (Z_i) is conditionally i.i.d. (independent and identically distributed) given (V, α), where α is the directing random measure for (Z_i), and that Z and V are conditionally independent given α. We may interpret the proposition above to be stating the same thing. In the case of the proposition, V is a random variable which takes values in Y, and (Z_i) takes values in X. A measure $m \in \mathcal{P}(Y \times \mathcal{P}X)$ is in some sense the joint distribution of (V, α), so if we know exactly what value $(y, \mu) \in Y \times \mathcal{P}X$ the pair (V, α) takes, then μ^I is the residual distribution on (Z_i), so it is i.i.d.. On the other hand, if we are given α, which is the same as being told the value of $m_{\mathcal{P}X}$, we obtain the conditional distribution $m(E|\alpha = \mu)$, which is merely a distribution on Y and is clearly independent of the resulting distribution μ^I on (Z_i). Beyond this however, Proposition 3.8 says nothing in the way of uniqueness of representation, which the proposition above does.

Next, we prove a generalization of Corollary 3.9 of [2] which states that for a finite collection of separately exchangeable sequences, meaning that the joint distribution of all sequences is invariant when a single one of the sequences is transformed by a permutation of finite support, then each sequence is i.i.d. conditioned on some directing measure. We extend this result to countably many sequences.

Proposition 2.5.35. *Let J be a countable set and let I be a countably infinite set, and for each $j \in J$ let X_j be a space. Also, for every $j \in J$, let $C_j \subset \Pi(I)$ be such that $(X_j^I, \mathcal{G}_{C_j})$ is a countable amenable exchangeable system. Define a dynamical system on $X_J^I = \prod_{j \in J} X_j^I$ by taking $\mathcal{T}$ to be the collection of all $T \in \prod_{j \in J} \mathcal{G}_{C_j}$ such that there exists $j \in J$ for which $\pi_j \circ T = T_j \circ \pi_j$ for some $T_j \in \mathcal{G}_{C_j}$, and $\pi_k \circ T = \mathrm{id}_{X_k^I} \circ \pi_k$ for* *all others $k \in J \setminus J$. Then $\beta \circ \mathcal{P}(\otimes \circ (\prod_{j\in J} \Psi_j))$ is an affine homeomorphism from*

$$\mathcal{P}\left(\prod_{j\in J} \mathcal{P}X_j\right) \quad \textit{to} \quad \mathcal{I}_{X_J^I}(\mathcal{T}),$$

where Ψ_j is the homeomorphism $\Psi_j : \mathcal{P}X_j \to \mathcal{E}_{X_j^I}(\mathcal{G}_{C_j})$ given by Theorem 2.5.22.

Proof. First, note that we clearly have that $\mathcal{T}$ is a joining of the systems $(X_j^I, \mathcal{G}_{C_j})$ for each $j \in J$, and this joining is actually independent. To see this, we have by Theorem 2.3.22 that $\mathcal{T}^*$ is a monoid, and with $\mathcal{T} \subset \mathcal{T}^*$ by Theorem 2.3.26, thus the group generated by $\mathcal{T}$, which is clearly $\prod_{j\in J} \mathcal{G}_{C_j}$ must be contained in $\mathcal{T}^*$, so the joining is independent. Also, each $T \in \mathcal{T}$ is continuous, so Lemma 2.3.7 gives that $\mathcal{I}_{X_J^I}(\mathcal{T})$ is closed. As such, by Theorem 2.5.8 we have that

$$\mathcal{E}_{X_J^I}(\mathcal{T}) = \bigotimes_{j\in J} \mathcal{E}_{X_j^I}(\mathcal{G}_{C_j}).$$

Proposition 2.1.6 thus gives that $\otimes : \prod_{j\in J} \mathcal{E}_{X_j^I}(\mathcal{G}_{C_j}) \to \mathcal{E}_{X_J^I}(\mathcal{T})$ is a homeomorphism. Additionally, since each $(X_j^I, \mathcal{G}_{C_j})$ is exchangeable, we have by Theorem 2.5.22 that $\Psi_j : \mathcal{P}X_j \to \mathcal{E}_{X_j^I}(\mathcal{G}_{C_j})$ is a homeomorphism, and thus $\prod_{j\in J} \Psi_j$ is a homeomorphism from $\prod_{j\in J} \mathcal{P}X_j$ to $\prod_{j\in J} \mathcal{E}_{X_j^I}(\mathcal{G}_{C_j})$. Thus, $\otimes \circ (\prod_{j\in J} \Psi_j)$ is a homeomorphism from $\prod_{j\in J} \mathcal{P}X_j$ to $\mathcal{E}_{X_J^I}(\mathcal{T})$, which proves that $\mathcal{E}_{X_J^I}(\mathcal{T})$ is compact, and Proposition 2.2.9 gives us that $\mathcal{P}(\otimes \circ (\prod_{j\in J} \Psi_j))$ is an affine homeomorphism from $\mathcal{P}(\prod_{j\in J} \mathcal{P}X_j)$ to $\mathcal{P}\mathcal{E}_{X_J^I}(\mathcal{T})$. Finally, Proposition 2.3.16 gives that $\beta : \mathcal{P}\mathcal{E}_{X_J^I}(\mathcal{T}) \to \mathcal{I}_{X_J^I}(\mathcal{T})$ is an affine homeomorphism, so $\beta \circ \mathcal{P}(\otimes \circ (\prod_{j\in J} \Psi_j))$ is as desired. □

In line with this theorem, there is also an analogous theorem for characterizing the jointly exchangeable sequences, where rather than applying a permutation of finite support to a single sequence, it is applied to all sequences simultaneously. This result will follow from a result of the proceeding section, and so we wait until then to prove it.

2.5.6 Power systems

For this subsection, the main object of concern will be the power system of a dynamical system with respect to an index system. Without further ado, we give the definition of index systems and power systems.

Definition 2.5.36. For a countable set I and a monoid $M \subset \Theta(I)$, the pair (I, M) is an *index system*. Given a dynamical system $(X, \mathcal{T})$ with $\mathrm{id}_X \in \mathcal{T}$, we may form the *power system* $(X, \mathcal{T})^{(I,M)}$ which is a dynamical system $(X^I, \mathcal{T}^M)$, where an element $T^g \in \mathcal{T}^M$ for $T \in \mathcal{T}$ and $g \in M$ is defined as

$$[T^g(x)]_i = T(x_{g^{-1}(i)}).$$

Note in this instance $\mathcal{T}^M$ is not the typical M-fold product of $\mathcal{T}$ with itself, but rather just the notation we use. In reality, the transformations in $\mathcal{T}^M$ are in bijection with $\mathcal{T} \times M$, since we have T^g for every $T \in \mathcal{T}$ and $g \in M$. Alternatively, if we let $\widehat{T} = \prod_{i \in I} T$ denote the I-fold product of T with itself, which is a map in $\mathcal{B}(X^I)$, we have that $T^g = \hat{T} \circ T_g$, where T_g is defined as usual for semicontractible systems.

Power systems are generalizations of semicontractible systems, as taking the dynamical system (X, id_X), we exactly recover the semicontractible system $(X^I, \mathcal{S}_M)$. Also, taking (I, id_I) as the index system, the power system is $(X^I, \widehat{\mathcal{T}})$, which is an I-fold self join of the system $(X, \mathcal{T})$ with itself, where we only take the diagonal transformations. It is for this reason we require both M and $\mathcal{T}$ to contain their respective identities, otherwise it may be rather difficult in general to describe the behavior of power systems. With previous results, it is rather easy to characterize the invariant measures of power systems.

Lemma 2.5.37. *Let $(X, \mathcal{T})^{(I,M)} = (X^I, \mathcal{T}^M)$ be a power system. Then*

$$\mathcal{I}_{X^I}(\mathcal{T}^M) = \mathcal{I}_{X^I}(\widehat{\mathcal{T}}) \cap \mathcal{I}_{X^I}(\mathcal{S}_M).$$

Proof. Since by definition, we have that $\mathcal{T}^M = \widehat{\mathcal{T}} \circ \mathcal{S}_M$, Lemma 2.3.6 gives the desired result. $\square$

In order to provide a more precise characterization, we will need the notion of a transitive power system, which we transfer over from the notion of a transitive semicontractible system.

Definition 2.5.38. An index system (I, M) is *transitive* if the semicontractible system $(X^I, \mathcal{S}_M)$ is transitive. An invertible power system $(X, \mathcal{T})^{(I,M)}$ is said to be *transitive* if (I, M) is transitive.

Also, with semicontractible systems, we can decompose them into a joining of transitive semicontractible systems by Theorem 2.5.20. Similarly, any power system will decompose as a joining of transitive power systems over the index systems on each of the orbits $O_M(I)$.

Proposition 2.5.39. *Let $(X, \mathcal{T})^{(I,M)}$ be a power system. Then $(X, \mathcal{T})^{(I,M)}$ is a joining of the transitive power systems $(X, \mathcal{T})^{(J,M^J)}$ for $J \in O_M(I)$, where*

$$M^J = \{g|_J : g \in M\}.$$

Furthermore, if $\mathcal{S}_M$ is a countable amenable group, then so is each $\mathcal{S}_{M^J}$.

Proof. Let $\widehat{\mathcal{T}} = \{\widehat{T} : T \in \mathcal{T}\} \subset \mathcal{B}(X^I)$, and note that for the power system $(X^I, \mathcal{T}^M)$, we have by definition that $\mathcal{T}^M = \widehat{\mathcal{T}} \circ \mathcal{S}_M$ (where the composition here is pairwise composition of transformations in the individual sets). We then have by Theorem 2.5.20 that $(X^I, \mathcal{S}_M)$ is a joining of the transitive semicontractible systems $(X^J, \mathcal{S}_{M^J})$ for $J \in O_M(I)$, with M^J as defined above, and if $\mathcal{S}_M$ is an amenable group, so is each $\mathcal{S}_{M^J}$. Since $\widehat{\mathcal{T}}$ has no effect on I, let $\widehat{\mathcal{T}}_J$ be the set of all $\prod_{j \in J} T$ for $T \in \mathcal{T}$, and we have that

$$(X, \mathcal{T})^{(I,M)} = (X^I, \widehat{\mathcal{T}} \circ \mathcal{S}_M)$$

is a joining of the transitive power systems

$$(X,\mathcal{T})^{(J,M^J)} = (X^J, \widehat{\mathcal{T}}_J \circ \mathcal{S}_{M^J}).$$

□

As a result, as is the case for semicontractible systems, we may turn our attention to characterizing the invariant measures for transitive power systems. As it turns out, such a characterization follows largely from Lemma 2.5.37 and previous results.

Proposition 2.5.40. *Let $(X,\mathcal{T})^{(I,M)} = (X^I, \mathcal{T}^M)$ be a transitive power system. Choose any $i_0 \in I$ and for any $\mu \in \mathcal{I}_{X^I}(\mathcal{T}^M)$, let $\mu_0 = \mu_{i_0}$ denote the i_0-marginal of μ. Then*

(a) if $\mu \in \mathcal{I}_{X^I}(\mathcal{T}^M)$, we have $\mu_0 = \mu_i$ for every $i \in I$,

(b) if $\mu \in \mathcal{I}_{X^I}(\mathcal{T}^M)$, we have $\mu_0 \in \mathcal{I}_X(\mathcal{T})$, and

(c) for $\mu \in \mathcal{I}_X(\mathcal{T})$, then $\prod_{i\in I} \mu \in \mathcal{I}_{X^I}(T^M)$.

Proof. First, by Lemma 2.5.37, we have that

$$\mathcal{I}_{X^I}(\mathcal{T}^M) = \mathcal{I}_{X^I}(\widehat{\mathcal{T}}) \cap \mathcal{I}_{X^I}(\mathcal{S}_M)$$

where $(X^I, \widehat{\mathcal{T}})$ is an I-fold joining of $(X,\mathcal{T})$ with itself, and $(X^I, \mathcal{S}_M)$ is a transitive semicontractible system.

For (a), let $\mu \in \mathcal{I}_{X^I}(\mathcal{T}^M) \subset \mathcal{I}_{X^I}(\mathcal{S}_M)$. By Lemma 2.5.18, we have that for all $i,j \in I$ that $\mu_i = \mu_j$, and so in particular for $i_0 \in I$ we have $\mu_0 = \mu_{i_0} = \mu_i$ for all $i \in I$.

For (b), let $\mu \in \mathcal{I}_{X^I}(\mathcal{T}^M) \subset \mathcal{I}_{X^I}(\widehat{\mathcal{T}})$. Then by Proposition 2.5.6(a), we have that $\mu_i \in \mathcal{I}_X(\mathcal{T})$ for every $i \in I$, and (a) implies that $\mu_0 \in \mathcal{I}_X(\mathcal{T})$.

For (c), let $\mu \in \mathcal{I}_X(\mathcal{T})$. By Proposition (c), we have that $\nu = \prod_{i \in I} \mu \in \mathcal{I}_{X^I}(\widehat{\mathcal{T}})$, and by Lemma 2.5.15, we have that $\nu \in \mathcal{I}_{X^I}(\mathcal{S}_M)$, as such, we have

$$\nu \in \mathcal{I}_{X^I}(\widehat{\mathcal{T}}) \cap \mathcal{I}_{X^I(\mathcal{S}_M)} = \mathcal{I}_{X^I}(\mathcal{T}^M).$$

□

Beyond these statements, it is not possible to give statements about the ergodic measures as Proposition 2.5.6 does for joinings. The Theorem below, which only applies to a certain kind of power system we define next, gives an indication as to why it is difficult in general to describe the ergodic measures of such systems.

Definition 2.5.41. A power system $(X, \mathcal{T})^{(I,M)}$ is said to be *contractible* if the semi-contractible system $(X^I, \mathcal{S}_M)$ is contractible.

We now pove a lemma which will be vital to characterizatiing the invariant measures of contractible power systems.

Lemma 2.5.42. *Let X be a space, I be a countably infinite set, and let $T \in \mathcal{B}(X)$. Then with the map $\Psi : \mathcal{P}X \to \mathcal{P}X^I$ defined by $\Psi(\mu) = \mu^I$, we have that $\mathcal{P}\widehat{T} \circ \Psi = \Psi \circ \mathcal{P}T$.*

Proof. Define $\mathscr{L} = \{E \in \mathscr{A}_{X^I} : [\mathcal{P}\widehat{T} \circ \Psi][\mu](E) = [\Psi \circ \mathcal{P}T][\mu](E)\}$. Then

$$[\mathcal{P}\widehat{T} \circ \Psi][\mu](X) = \mu^I(\widehat{T}^{-1}(X)) = \mu^I(X) = 1 = [\mathcal{P}T(\mu)]^I(X) = [\Psi \circ \mathcal{P}T][\mu](X),$$

and

$$[\mathcal{P}\widehat{T} \circ \Psi][\mu](\varnothing) = \mu^I(\widehat{T}^{-1}(\varnothing)) = \mu^I(\varnothing) = 0 = [\mathcal{P}T(\mu)]^I(\varnothing) = [\Psi \circ \mathcal{P}T][\mu](\varnothing),$$

which gives that $\varnothing, X \in \mathscr{L}$. Now, for $E \in \mathscr{L}$, we have that $[\mathcal{P}\widehat{T} \circ \Psi][\mu](E) =$

$[\Psi \circ \mathcal{P}T][\mu](E)$, and since both $[\mathcal{P}\widehat{T} \circ \Psi](\mu)$ and $[\Psi \circ \mathcal{P}T](\mu)$ are measures, we have

$$[\mathcal{P}\widehat{T} \circ \Psi][\mu](X \setminus E) = 1 - [\mathcal{P}\widehat{T} \circ \Psi][\mu](E) = 1 - [\Psi \circ \mathcal{P}T][\mu](E) = [\Psi \circ \mathcal{P}T][\mu](X \setminus E),$$

and so $X \setminus E \in \mathscr{L}$. Finally, let $\{E_n\}_{n\in\mathbb{N}} \subset \mathscr{L}$ be a countable disjoint collection of sets such that for each $n \in \mathbb{N}$, we have $[\mathcal{P}\widehat{T} \circ \Psi][\mu](E_n) = [\Psi \circ \mathcal{P}T][\mu](E_n)$. Again, since both $[\mathcal{P}\widehat{T} \circ \Psi](\mu)$ and $[\Psi \circ \mathcal{P}T](\mu)$ are measures, we have by countable additivity that

$$\begin{aligned} [\mathcal{P}\widehat{T} \circ \Psi][\mu]\left(\bigsqcup_{n\in\mathbb{N}} E_n\right) &= \sum_{n\in\mathbb{N}} [\mathcal{P}\widehat{T} \circ \Psi][\mu](E_n) \\ &= \sum_{n\in\mathbb{N}} [\Psi \circ \mathcal{P}T][\mu](E_n) \\ &= [\Psi \circ \mathcal{P}T][\mu]\left(\bigsqcup_{n\in\mathbb{N}} E_n\right), \end{aligned}$$

and therefore $\bigcup_{n\in\mathbb{N}} E_n \in \mathscr{L}$. As such, $\mathscr{L}$ is a λ-system.

Now, let $\mathscr{P}$ be the collection of all finite cylinders for X^I, and note that $\mathscr{P}$ is closed under finite intersections so that it is a π-system. Let $J \subset I$ be finite with $E_j \in \mathscr{A}_X$ for each $j \in J$ so that $E = \bigcap_{j\in J} \pi_j^{-1}(E_j) \in \mathscr{P}$ is an arbitrary finite cylinder. Then we have for every $i \in I$ that $\pi_i \circ \widehat{T} = T \circ \pi_i$, and so

$$\begin{aligned}
[\mathcal{P}\widehat{T} \circ \Psi][\mu](E) &= \mu^I \left(\widehat{T}^{-1} \bigcap_{j \in J} \pi_j^{-1}(E_j) \right) \\
&= \mu^I \left(\bigcap_{j \in J} \widehat{T}^{-1}(\pi_j^{-1}(E_j) \right) \\
&= \mu^I \left(\bigcap_{j \in J} \pi_j^{-1}(T^{-1}(E_j)) \right) \\
&= \prod_{j \in J} \mu(T^{-1}(E_j)) \\
&= \prod_{j \in J} \mathcal{P}T[\mu](E_j) \\
&= [\mathcal{P}T(\mu)]^I \left(\bigcap_{j \in J} \pi_j^{-1}(E_j) \right) \\
&= [\Psi \circ \mathcal{P}T][\mu](E)
\end{aligned}$$

As such, we have that $E \in \mathscr{L}$, and since $E \in \mathscr{P}$ was arbitrary, we have shown that $\mathscr{P} \subset \mathscr{L}$. By the $\pi - \lambda$ Theorem, the σ-algebra generated by $\mathscr{P}$ (which is $\mathscr{A}_{X^I}$) is contained in $\mathscr{L} \subset \mathscr{A}_{X^I}$, and therefore we have that $[\mathcal{P}\widehat{T} \circ \Psi](\mu) = [\Psi \circ \mathcal{P}T](\mu)$. Since $\mu \in \mathcal{P}X$ was arbitrary, we have shown that $\mathcal{P}\widehat{T} \circ \Psi = \Psi \circ \mathcal{P}T$. $\square$

This lemma allows us to give the following surprising result about the invariant measures of contractible power systems.

Theorem 2.5.43. *Let $(X, \mathcal{T})^{(I,M)}$ be a contractible power system. Then with $\mathcal{P}\mathcal{T} = \{\mathcal{P}T : T \in \mathcal{P}\}$, we have $\beta \circ \mathcal{P}\Psi : \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T}) \to \mathcal{I}_{X^I}(\mathcal{T}^M)$ is an affine homeomorphism.*

Proof. First, by Lemma 2.5.37, we have that $\mathcal{I}_{X^I}(\mathcal{T}^M) = \mathcal{I}_{X^I}(\widehat{\mathcal{T}}) \cap \mathcal{I}_{X^I}(\mathcal{S}_M)$, and since $(X^I, \mathcal{T}^M)$ is contractible, we have that $(X^I, \mathcal{S}_M)$ is contractible, and so by Theorem 2.5.22, we have that $\beta \circ \mathcal{P}\Psi$ is an affine homeomorphism from $\mathcal{P}^2 X$ to $\mathcal{I}_{X^I}(\mathcal{S}_M)$. Now, let us note that by Lemmas 2.2.8, 2.2.4, and 2.5.42 that we have for every $T \in \mathcal{T}$ that

$$\begin{aligned}\mathcal{P}\widehat{T} \circ \beta \circ \mathcal{P}\Psi &= \beta \circ \mathcal{P}^2\widehat{T} \circ \mathcal{P}\Psi = \beta \circ \mathcal{P}[\mathcal{P}\widehat{T} \circ \Psi] \\ &= \beta \circ \mathcal{P}[\Psi \circ \mathcal{P}T] = \beta \circ \mathcal{P}\Psi \circ \mathcal{P}^2 T.\end{aligned}$$

Then for $\mu \in \mathcal{I}_{X^I}(\mathcal{T}^M) \subset \mathcal{I}_{X^I}(\mathcal{S}_M)$, we have that there is a unique $m \in \mathcal{P}^2 X$ for which $\mu = [\beta \circ \mathcal{P}\Psi](m)$. We also have that $\mu \in \mathcal{I}_{X^I}(\widehat{\mathcal{T}})$, and so for $T \in \mathcal{T}$ (which gives an arbitrary $\widehat{T} \in \widehat{\mathcal{T}}$), we have $\mathcal{P}\widehat{T}(\mu) = \mu$. Putting these two facts together, along with the result of the previous display gives the following result.

$$\mu = \mathcal{P}\widehat{T}(\mu) = [\mathcal{P}\widehat{T} \circ \beta \circ \mathcal{P}\Psi](m) = [\beta \circ \mathcal{P}\Psi \circ \mathcal{P}^2 T](m) = [\beta \circ \mathcal{P}\Psi](\mathcal{P}^2 T(m)).$$

But, by assumption, m is the unique measure for which $[\beta \circ \mathcal{P}\Psi](m) = \mu$, and therefore it must be that $m = \mathcal{P}^2 T(m)$. This however is precisely what it means for $m \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}T)$, and since this holds for every $T \in \mathcal{T}$, we have that $m \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$. Additionally, for $m \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$, we have for every $T \in \mathcal{T}$ that $\mathcal{P}^2 T(m) = m$, and therefore we have for every $T \in \mathcal{T}$ that

$$\begin{aligned}\mu &= [\beta \circ \mathcal{P}\Psi](m) = [\beta \circ \mathcal{P}\Psi](\mathcal{P}^2 T(m)) = [\beta \circ \mathcal{P}\Psi \circ \mathcal{P}^2 T](m) \\ &= [\mathcal{P}\widehat{T} \circ \beta \circ \mathcal{P}\Psi](m) = \mathcal{P}\widehat{T}([\beta \circ \mathcal{P}\Psi](m)) = \mathcal{P}\widehat{T}(\mu),\end{aligned}$$

and therefore $\mathcal{P}\widehat{T}(\mu) = \mu$ for every $T \in \mathcal{T}$, which gives that $\mu \in \mathcal{I}_{X^I}(\widehat{\mathcal{T}})$. As such, for $\mu \in \mathcal{I}_{X^I}(\mathcal{S}_M)$, we have that $\mu \in \mathcal{I}_{X^I}(\widehat{\mathcal{T}})$ if and only if $m = [\beta \circ \mathcal{P}\Psi]^{-1}(\mu) \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$, and since $\mathcal{I}_{X^I}(\widehat{\mathcal{T}}) \cap \mathcal{I}_{X^I}(\mathcal{S}_M) = \mathcal{I}_{X^I}(\mathcal{T}^M)$, it follows that $[\beta \circ \mathcal{P}\Psi]^{-1}(\mathcal{I}_{X^I}(\mathcal{T}^M)) = \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$, which is the desired result. □

Rather surprisingly, the invariant measures for every contractible power system $(X, \mathcal{T})^{(I,M)}$ are in exact correspondence with the set of invariant measures of the system $(\mathcal{P}X, \mathcal{P}\mathcal{T})$. In general, these invariant measures are far more complicated than the invariant measures for the original system, however the elements of $\mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$ have a characteristic property in terms of $\mathcal{I}_X(\mathcal{T})$ using the barycenter map.

Proposition 2.5.44. *Let $(X,\mathcal{T})$ be a dynamical system. Then $\beta(\mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})) = \mathcal{I}_X(\mathcal{T})$.*

Proof. First, suppose that $m \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$, and let $T \in \mathcal{T}$, so we have $\mathcal{P}T \in \mathcal{P}\mathcal{T}$, and therefore $\mathcal{P}^2T(m) = m$. Then we have by Lemma 2.2.8 that

$$\mathcal{P}T(\beta(m)) = [\mathcal{P}T \circ \beta](m) = [\beta \circ \mathcal{P}^2T](m) = \beta(\mathcal{P}^2T(m)) = \beta(m),$$

and therefore we have that $\beta(m) \in \mathcal{I}_X(T)$. Since this holds for every $T \in \mathcal{T}$, we have that $\beta(m) \in \mathcal{I}_X(\mathcal{T})$. As $m \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$ was arbitrary, we have that $\beta(\mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})) \subset \mathcal{I}_X(\mathcal{T})$.

Now, let $\mu \in \mathcal{I}_X(\mathcal{T})$, and let $m = \delta_\mu \in \mathcal{P}^2X$. Then for $T \in \mathcal{T}$, we have $\mathcal{P}T(\mu) = \mu$, and so

$$\mathcal{P}^2T(\delta_\mu) = \delta_{\mathcal{P}T(\mu)} = \delta_\mu,$$

and therefore $\delta_\mu \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}T)$. Since this holds for every $T \in \mathcal{T}$, we have that $\delta_\mu \in \mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$. Also, note that for $E \in \mathscr{A}_X$, we have

$$\beta[\delta_\mu](E) = \int_{\mathcal{P}X} \nu(E)\, \delta_\mu(\mathrm{d}\nu) = \int_{\{\mu\}} \nu(E)\, \delta_\mu(\mathrm{d}\nu) = \mu(E),$$

and therefore $\beta(\delta_\mu) = \mu$, which gives that $\mu \in \beta(\mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T}))$. Since $\mu \in \mathcal{I}_X(\mathcal{T})$ was arbitrary, we have shown $\mathcal{I}_X(\mathcal{T}) \subset \beta(\mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T}))$, which when combined with the result of the previous paragraph gives the desired result. ☐

Informally, the barycenter of any invariant measure for $(\mathcal{P}X, \mathcal{P}\mathcal{T})$ must be an invariant measure for $(X,\mathcal{T})$, and every invariant measure for $(X,\mathcal{T})$ is the barycenter of some invariant measure for $(\mathcal{P}X, \mathcal{P}\mathcal{T})$. It is not the case however that any measure on $\mathcal{P}^2X$ whose barycenter is an invariant measure for $(X, \mathcal{T})$ will be an invariant measure for $(\mathcal{P}X, \mathcal{P}\mathcal{T})$, as is demonstrated with the following example.

Example 2.5.45. Let $X = \{0, 1, 2\}$, and let $T : X \to X$ be defined by $T(x) = x + 1 \mod 3$. Then clearly we have $\mu = \frac{1}{3}(\delta_0 + \delta_1 + \delta_2)$ is the only element of $\mathcal{I}_X(T)$. Now, define

$$m = \frac{1}{2}\delta\left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1\right) + \frac{1}{4}\delta\left(\frac{1}{3}\delta_0 + \frac{2}{3}\delta_2\right) + \frac{1}{4}\delta\left(\frac{1}{3}\delta_1 + \frac{2}{3}\delta_2\right)$$

to be an element of $\mathcal{P}^2X$. Then we have

$$\begin{aligned}\beta(m) &= \left(\frac{1}{2}\frac{1}{2} + \frac{1}{4}\frac{1}{3}\right)\delta_0 + \left(\frac{1}{2}\frac{1}{2} + \frac{1}{4}\frac{1}{3}\right)\delta_1 + \left(\frac{1}{4}\frac{2}{3} + \frac{1}{4}\frac{2}{3}\right)\delta_2 \\ &= \frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_2 = \mu,\end{aligned}$$

however

$$\mathcal{P}^2T(m) = \frac{1}{2}\delta\left(\frac{1}{2}\delta_1 + \frac{1}{2}\delta_2\right) + \frac{1}{4}\delta\left(\frac{1}{3}\delta_1 + \frac{2}{3}\delta_0\right) + \frac{1}{4}\delta\left(\frac{1}{3}\delta_2 + \frac{2}{3}\delta_0\right),$$

which is clearly not equal to m.

It is however still a useful heuristic for determining whether or not an element $m \in \mathcal{P}^2X$ is invariant for $(\mathcal{P}X, \mathcal{P}\mathcal{T})$. Now, to demonstrate Theorem 2.5.43 in action, we prove an analog to Proposition 2.5.35, but where we apply the same permutation to all sequences simultaneously.

Proposition 2.5.46. *Let J be a countable set, and I be a countably infinite set, and for each $j \in J$ let X_j be a space. Choose some monoid $M \subset \Theta(I)$ such that for each $j \in J$, $(X_j^I, \mathcal{S}_M)$ is a contractible system. Define a dynamical system $(X_J^I, \mathcal{T})$ on $X_J^I = \prod_{j\in J} X_j^I$ by taking $\mathcal{T}$ to be the collection of transformations of the form*

$$\tilde{T}_g = \prod_{j\in J} T_g$$

for every $g \in M$. With $X_J = \prod_{j\in J} X_j$, then $\beta \circ \mathcal{P}\Psi$ is an affine homeomorphism from $\mathcal{P}^2 X_J$ to $\mathcal{I}_{X_J^I}(\mathcal{T})$.

Proof. Note that the system $(X_J^I, \mathcal{T})$ is identical to the contractible power system $(X_J, \mathrm{id}_{X_J})^{(I,G)}$, by realizing that rather than treating this system as a collection of J sequences on X_j^I, which we are all transforming by $g \in M$ simultaneously, we are instead treating this system as a single sequence on $(X_J)^I$, which we are transforming by $g \in M$. Additionally, we do nothing else to X_J, hence why the base system only has the identity as a transformation. By Theorem 2.5.43, we have that $\beta \circ \mathcal{P}\Psi$ is an affine homeomorphism from $\mathcal{I}_{\mathcal{P}X_J}(\mathcal{P}\,\mathrm{id}_{X_J})$ to $\mathcal{I}_{X_J^I}(\mathcal{T})$. Since $\mathcal{P}\,\mathrm{id}_{X_J} = \mathrm{id}_{\mathcal{P}X_J}$, we clearly have that $\mathcal{I}_{\mathcal{P}X_J}(\mathrm{id}_{\mathcal{P}X_J}) = \mathcal{P}^2 X_J$, which is the desired result. □

This example is a rather simple form of power system, but it still is one, and it is interesting to see it applied to this case. Interpreting this result, it states that if we have J sequences which we simultaneously transform by an element $g \in M$ (for some contractible system $(X^I, \mathcal{S}_M)$), then there may be correlation between the sequences, but this correlation can only happen within one specific index, similar to the usual statement of De Finetti's Theorem. Next, we give a more complex example of using Theorem 2.5.43.

Proposition 2.5.47. *Let $X = \{0,1\}$, and let $T : X \to X$ be defined as $T(0) = 1$ and $T(1) = 0$. Then $\mathcal{I}_{\mathcal{P}X}(\mathcal{P}T)$ is affinely homeomorphic to the set of probability measures on $[0, \frac{1}{2}]$.*

Proof. First note that

$$\mathcal{P}X = \{t\delta_0 + (1-t)\delta_1 : t \in [0,1]\},$$

and so $\mathcal{P}X$ is affinely homeomorphic to $[0,1]$. Now, note that for $\mu_t = t\delta_0 + (1-t)\delta_1$, we have that

$$\mathcal{P}T(\mu_t) = t\mathcal{P}T(\delta_0) + (1-t)\mathcal{P}T(\delta_1) = t\delta_{T(0)} + (1-t)\delta_{T(1)} = (1-t)\delta_0 + t\delta_1 = \mu_{1-t}.$$

Using the affine heomorphism to $[0,1]$, the map $\mathcal{P}T$ becomes $S(x) = 1-x$, and thus $\mathcal{I}_{\mathcal{P}X}(\mathcal{P}T)$ is affinely homeomorphic to $\mathcal{I}_{[0,1]}(S)$. As such, any $m \in \mathcal{I}_{[0,1]}(S)$ must have probabilities symmetric about $\frac{1}{2}$, giving that $\mathcal{P}S(m|_{[0,\frac{1}{2})}) = m|_{(\frac{1}{2},1]}$, and with $1 = m([0,1]) = m([0,\frac{1}{2})) + m(\{\frac{1}{2}\}) + m((\frac{1}{2},1])$. Any such measure is uniquely constructed by choosing a measure n on $[0,\frac{1}{2}]$, and taking $m = \frac{1}{2}n + \frac{1}{2}\mathcal{P}S(n)$, so the invariant measures on $\mathcal{I}_{\mathcal{P}X}(\mathcal{P}\mathcal{T})$ are affinely homeomorphic to $\mathcal{P}[0,\frac{1}{2}]$ by the map

$$m = \int_{[0,\frac{1}{2}]} \frac{1}{2}\delta(t\delta_0 + (1-t)\delta_1) + \frac{1}{2}\delta((1-t)\delta_0 + t\delta_1)\, n(\mathrm{d}t).$$

Additionally, letting $\mathcal{T} = \{\mathrm{id}_X, T\}$ and (I, M) be an index set such that $(X,\mathcal{T})^{(I,M)}$ is a contractible power system, then using Theorem 2.5.43, we have for every $\mu \in \mathcal{I}_{X^I}(T^M)$ there exists a unique $n \in \mathcal{P}[0,\frac{1}{2}]$ such that

$$\mu(E) = \int_{[0,\frac{1}{2}]} \frac{1}{2}(t\delta_0 + (1-t)\delta_1)^I(E) + \frac{1}{2}((1-t)\delta_0 + (1-t)\delta_1)^I(E)\, n(\mathrm{d}t).$$

□

With $I = \mathbb{N}$, this is an analog for De Finetti's Theorem on 0-1 exchangeable sequences which has the additional requirement that the probability of seeing some sequence $x_1x_2\cdots x_n$ is the same as seeing $\bar{x}_1\bar{x}_2\cdots\bar{x}_n$ where $\bar{x}_i = 1 - x_i$. For example, it is equally likely to observe 100101 as it is to observe 011010, and this holds for any arbitrary finite sequence of values.

2.5.6.1 Connections to the Aldous-Hoover Theorem

Re-framed to the language of dynamical systems as presented here, the Aldous-Hoover Theorem is classically a characterization of the invariant measures of the following system.

Definition 2.5.48. Let X be a space, and let I and J be countably infinite sets, and define $X^{I\times J}$, which is a space. Then, for $(g,h)\in\pi(I)\times\pi(J)$, define $T_{g,h}: X^{I\times J}\to X^{I\times J}$ by

$$[T_{(g,h)}(x)]_{(i,j)} = x_{(g^{-1}(i),h^{-1}(j))},$$

and let $\mathcal{T}=\{T_{(g,h)}:(g,h)\in\pi(I)\times\pi(J)\}$. This gives a system $(X^{I\times J},\mathcal{T})$.

Classically, the set of invariant measures for this system is classified as follows.

Theorem 2.5.49 (Aldous-Hoover Theorem [1, 30])**.** *A measure $\mu\in\mathcal{P}X^{I\times J}$ is in $\mathcal{I}_{X^{I\times J}}(\mathcal{T})$ if and only if there exists some function $f:[0,1]^4\to X$ and random variables $\alpha,\eta_i,\xi_j,\zeta_{i,j}$ for all $i\in I$ and $j\in J$ with $\zeta_{i,j}=\eta_{j_i}$ which are uniform i.i.d. such that $\mu_{i,j}$ is the distribution of the random variable $f(\alpha,\eta_i,\xi_j,\zeta_{i,j})$ for every $i\in I$ and $j\in J$.*

This characterization of the invariant measure is very much one of probabilistic origin (in fact its statement and proof is entirely in the language of probability theory). The result also extends to an arbitrary number of dimensions, and is not only restricted to two dimensions. It is also worth noting that different choices of f above may result in the same measure μ, and while Hoover has given a characterization of when two such f yield the same invariant distribution, but it would be preferable to have a bijective correspondence. As with the purely dynamical characterization of De Finetti's Theorem would indicate, there should be a dynamical characterization of the Aldous-Hoover Theorem, and this characterization should be bijective. As it turns out, the system defined above is a contractible power system, once adequately re-framed.

Lemma 2.5.50. *The contractible power system $(X^I,\mathcal{S}_{\pi(I)})^{(J,\pi(J))}$ is the same as the system $(X^{I\times J},\mathcal{T})$.*

Proof. Indeed for $T_{(g,h)} \in \mathcal{T}$ with $(g,h) \in \pi(I) \times \pi(J)$, we have for every $i \in I$ and $j \in J$ that

$$[T_{(g,h)}(x)]_{(i,j)} = x_{(g^{-1}(i),h^{-1}(j))}.$$

Now, for $T_g \in \mathcal{S}_{\pi(I)}$ and $h \in \pi(J)$, we have by definition of power systems that for $x \in (X^I)^J$ that

$$[T_g^h(x)]_j = T_g(x_{h^{-1}(j)}),$$

where $[T_g^h(x)]_j \in X^I$, and so at index i of this sequence, we have

$$\big[[T_g^h(x)]_j\big]_i = [T_g(x_{h^{-1}(j)})]_i = (x_{h^{-1}(j)})_{g^{-1}(i)}.$$

Identifying points in $(X^I)^J$ with $X^{I\times J}$, we have that $T_g^h = T_{g,h}$, and so these systems are identical. □

As a result, we may give the following version of the Aldous-Hoover in the context of dynamical systems.

Proposition 2.5.51. *Let $(X^I, \mathcal{S})$ be a contractible system, and let $(X^I, \mathcal{S})^{(J,M)}$ be any contractible power system with $(X^I, \mathcal{S})$ as the base. Then the invariant measures for $(\mathcal{P}X^I, \mathcal{P}\mathcal{S})$ are affinely homeomorphic to the invariant measures of $(X^I, \mathcal{S})^{(J,\mathcal{T})} = ((X^I)^J, \mathcal{S}^M)$ by the affine homeomorphism $\beta \circ \mathcal{P}\Psi$.*

Proof. This follows immediately from applying Theorem 2.5.43. □

Now, this characterization is in some sense incomplete, as it is not entirely clear what $\mathcal{I}_{\mathcal{P}X^I}(\mathcal{P}\mathcal{S})$ would be, as this is a rather complex system. However, Theorem 2.5.43 has reduced characterizing the invariant measures of a two dimensional system to characterizing the invariant measures of system which is effectively one dimensional, and furthermore this restriction is in exact bijection with the invariant measures of the original system. If a nice characterization of $\mathcal{I}_{\mathcal{P}X^I}(\mathcal{P}\mathcal{S})$ could be found, then this would provide a characterization of the invariant measures of separately exchangeable arrays just as the Aldous-Hoover Theorem does, however this characterization would be in perfect bijection with the invariant measures. Additionally, if we continue to take further contractible powers of the system, an analog of the Aldous-Hoover Theorem for higher dimensions may be recoverable.

2.6 Final Remarks

This chapter, at its core, introduces the concept of completions of dynamical systems, and demonstrates their relevance to dynamical systems of all varieties when the objects of interest are the invariant measures. From invariant sets, to pointwise ergodic theorems, to characterizing the invariant measures, it is clear that the completion of a dynamical system is vital to take into consideration in ergodic theory. Regardless of the structure of an original system $(X, \mathcal{T})$, the completion $\mathcal{T}^*$ always has rather nice properties, given by Theorem 2.3.26. Completions enable stating the pointwise ergodic theorem in a more general form, and also transferring the results of pointwise ergodic theorems to systems for which it was not previously known to hold. Dynamical independence as a concept is also a very useful tool in the characterization of the invariant measures of dynamical systems when it is suspected that the dynamics induce the independence of sets. These properties come together to extend De Finetti's Theorem to a far broader class of systems, and also indicates how a similar extension is possible for other characterizations of the invariant measures of a system, such as the Aldous-Hoover Theorem.

Open problems such as the Furstenberg $\times 2 \times 3$ conjecture on the invariant measures of the system $(S^1, \{\times 2, \times 3\})$ [21], with $S^1 = \mathbb{R}/\mathbb{Z}$, and the maps $\times 2$ and $\times 3$ multiply by 2 and 3 modulo 1, may benefit from considering the completion of this system. By either identifying other maps which must be in the completion of this system, or by characterizing the complete dynamical system which would yield the conjectured invariant measures, it may be possible to find an approach to the theorem outside of the scope of existing methods.

In terms of future work, connecting the notion of the completion of a dynamical system to any other concept within ergodic theory may prove to be useful, as well as exploring other properties that the completion might posses. In terms of particular questions for future work, the first major question is whether the condition that

$\mathcal{I}(\mathcal{T})$ is closed is actually necessary whenever it is used. The only reason for this assumption anywhere it is used is to apply Choquet's Theorem, however a suitably general ergodic decomposition could replace this in many contexts. The results of [18] are rather promising, and with a careful translate to the typical language of dynamical systems, would allow for the replacement of Choquet's Theorem wherever it appears, whether directly or indirectly. Notably, this would remove the requirements for Theorems 2.4.11 and 2.4.13, making the proofs apply to all dynamical systems of a single transformation, and every dynamical system over a countable amenable group, solidifying the definitions of the dynamical expectation of Birkhoff systems as an appropriate generalization of the concepts. Additionally, there are many results in Section 2.5 which have additional constraints that are necessary for the proofs given, but for which there is no clear reason why (or if) the constraint is necessary. For example, for Proposition 2.5.7 about the invariant measures of product systems, it seems as though the result should still hold when each $\mathcal{T}_i$ is not necessarily a countable amenable group. While it may still be necessary for there to be some assumption on it (such as assuming that $\mathcal{T}_i$ contains the identity), however the countability and amenability only really appear so that dynamical independence can be invoked. Additionally, with the invariant measures of contractible power systems $(X, \mathcal{T})^{(I,M)}$ coinciding with the invariant measures of $(\mathcal{P}X, \mathcal{P}\mathcal{T})$, it would be fruitful to develop tools to classify these invariant measures in terms of the invariant measures of $(X, \mathcal{T})$, if possible. In particular, characterizing the invariant measures of $(\mathcal{P}X^I, \mathcal{P}\mathcal{S}_{\pi(I)})$ would enable a stronger form of the Aldous-Hoover Theorem.

REFERENCES

[1] David J. Aldous, *Representations for partially exchangeable arrays of random variables*, Journal of Multivariate Analysis **11** (1981), no. 4, 581–598.

[2] David J. Aldous, *Exchangeability and related topics*, École d'Été de Probabilités de Saint-Flour XIII — 1983 (Berlin, Heidelberg) (P. L. Hennequin, ed.), Springer Berlin Heidelberg, 1985, pp. 1–198.

[3] Julian Asilis, *Probability monads*, Ph.D. thesis, PhD thesis, Harvard University Cambridge, Massachusetts, 2020.

[4] Stefan Banach and Alfred Tarski, *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fundamenta Mathematicae **6** (1924), no. 1, 244–277.

[5] Sebastián Barbieri, *On the entropies of subshifts of finite type on countable amenable groups*, Groups, Geometry, and Dynamics **15** (2021), no. 2, 607–638.

[6] SebastiÃ¡n Barbieri, Mathieu Sablik, and Ville Salo, *Soficity of free extensions of effective subshifts*, Discrete and Continuous Dynamical Systems **45** (2025), no. 4, 1117–1149.

[7] S. Baron and Solomon Leader, *5299*, The American Mathematical Monthly **73** (1966), no. 6, 677–678.

[8] Patrick Billingsley, *Probability and measure*, 3rd ed., John Wiley & Sons, 1995.

[9] George D Birkhoff, *Proof of the ergodic theorem*, Proceedings of the National Academy of Sciences **17** (1931), no. 12, 656–660.

[10] Robert Bland, *An embedding theorem for subshifts over amenable groups with the comparison property*, Ergodic Theory and Dynamical Systems (2024), 1–31.

[11] Robert Bland, Kevin McGoff, and Ronnie Pavlov, *Subsystem entropies of shifts of finite type and sofic shifts on countable amenable groups*, Ergodic Theory and Dynamical Systems (2022), 1–34.

[12] Lewis Bowen, *Measure conjugacy invariants for actions of countable sofic groups*, Journal of the American Mathematical Society **23** (2010), no. 1, 217–245.

[13] Silas Carvalho and Alexander Paul Condori Huaman, *Generic properties of invariant measures of full-shift systems over perfect polish metric spaces*, Stochastics and Dynamics **21** (2021), 2150040.

[14] Tullio Ceccherini-Silberstein and Michel Coornaert, *On the density of periodic configurations in strongly irreducible subshifts*, Nonlinearity **25** (2012), no. 7, 2119.

[15] Nhan-Phu Chung and Guohua Zhang, *Weak expansiveness for actions of sofic groups*, Journal of Functional Analysis **268** (2015), no. 11, 3534–3565.

[16] Bruno De Finetti, *Funzione caratteristica di un fenomeno aleatorio*, Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928, 1929, pp. 179–190.

[17] Tomasz Downarowicz, Dawid Huczek, and Guohua Zhang, *Tilings of amenable groups*, Journal für die reine und angewandte Mathematik (Crelles Journal) **2019** (2019), no. 747, 277–298.

[18] Noé Ensarguet and Paolo Perrone, *Categorical probability spaces, ergodic decompositions, and transitions to equilibrium*, 2024.

[19] Joshua Frisch and Omer Tamuz, *Symbolic dynamics on amenable groups: the entropy of generic shifts*, Ergodic Theory and Dynamical Systems **37** (2017), no. 4, 1187–1210.

[20] Tobias Fritz, Tomáš Gonda, and Paolo Perrone, *De finetti's theorem in categorical probability*, Journal of Stochastic Analysis **2** (2021), no. 2, 1–26.

[21] Harry Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation*, Mathematical systems theory **1** (1967), no. 1, 1–49.

[22] Winston Garira and Bothwell Maregere, *The transmission mechanism theory of disease dynamics: Its aims, assumptions and limitations*, Infectious Disease Modelling **8** (2022), 122–144.

[23] Eli Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, no. 101, American Mathematical Soc., 2003.

[24] Philip Hall, *Some constructions for locally finite groups*, Journal of the London Mathematical Society **1** (1959), no. 3, 305–319.

[25] Edwin Hewitt and Leonard J Savage, *Symmetric measures on cartesian products*, Transactions of the American Mathematical Society **80** (1955), no. 2, 470–501.

[26] Neil Hindman and Dona Strauss, *Density in arbitrary semigroups*, Semigroup Forum, vol. 73, Springer, 2006, pp. 273–300.

[27] Assaf Hochman, Pinhas Alpert, Tzvi Harpaz, Hadas Saaroni, and Gabriele Messori, *A new dynamical systems perspective on atmospheric predictability: Eastern mediterranean weather regimes as a case study*, Science Advances **5** (2019), no. 6, eaau0936.

[28] Michael Hochman, *On the automorphism groups of multidimensional shifts of finite type*, Ergodic Theory and Dynamical Systems **30** (2010), no. 3, 809–840.

[29] Michael Hochman and Tom Meyerovitch, *A characterization of the entropies of multidimensional shifts of finite type*, Annals of Mathematics (2010), 2011–2038.

[30] Douglas N Hoover, *Relations on probability spaces and arrays of random variables*, Institute for Advanced Study (1979), 1–63.

[31] Olav Kallenberg et al., *Random measures, theory and applications*, vol. 1, Springer, 2017.

[32] Alexander Kechris, *Classical descriptive set theory*, vol. 156, Springer Science & Business Media, 2012.

[33] David Kerr and Hanfeng Li, *Ergodic theory: Independence and dichotomies*, Springer, 2016.

[34] Sebastián Andrés Barbieri Lemp, *Shift spaces on groups: computability and dynamics*, Ph.D. thesis, Université de Lyon, 2017.

[35] Douglas A Lind, *The entropies of topological markov shifts and a related class of algebraic integers*, Ergodic Theory and Dynamical Systems **4** (1984), no. 2, 283–300.

[36] Douglas Alan Lind and Brian Marcus, *An introduction to symbolic dynamics and coding*, 2nd ed., Cambridge University Press, 2021.

[37] Elon Lindenstrauss, *Pointwise theorems for amenable groups*, Inventiones mathematicae **146** (2001), no. 2, 259–295.

[38] Kevin McGoff and Ronnie Pavlov, *Random $\mathbb{Z}^d$-shifts of finite type*, Journal of Modern Dynamics **10** (2016), no. 2, 287–330.

[39] Sean Moss and Paolo Perrone, *A category-theoretic proof of the ergodic decomposition theorem*, Ergodic Theory and Dynamical Systems **43** (2023), no. 12, 4166–4192.

[40] Kalyanapuram Rangachari Parthasarathy, *Probability measures on metric spaces*, vol. 352, American Mathematical Soc., 2005.

[41] Ronnie Pavlov, *Perturbations of multidimensional shifts of finite type*, Ergodic Theory and Dynamical Systems **31** (2011), no. 2, 483–526.

[42] ______, *A class of nonsofic multidimensional shift spaces*, Proceedings of the American Mathematical Society **141** (2013), no. 3, 987–996.

[43] Robert R Phelps, *Lectures on choquet's theorem*, Springer, 2001.

[44] Jade Raymond, *Shifts of finite type on locally finite groups*, Ergodic Theory and Dynamical Systems **44** (2024), no. 12, 3565–3598.

[45] Paul Ressel, *Some continuity and measurability results on spaces of measures*, Mathematica Scandinavica **40** (1977), no. 1, 69–78.

[46] Raphael M. Robinson, *Undecidability and nonperiodicity for tilings of the plane.*, Inventiones mathematicae **12** (1971), 177–209.

[47] Halsey Royden and Patrick Michael Fitzpatrick, *Real analysis*, China Machine Press, 2010.

[48] David Ruelle, *Thermodynamic formalism: the mathematical structure of equilibrium statistical mechanics*, Cambridge University Press, 2004.

[49] Omri Sarig, *Lecture notes on ergodic theory*, Lecture Notes, Penn. State University (2009), 1–145.

[50] Friedrich Martin Schneider and Andreas Thom, *On følner sets in topological groups*, Compositio Mathematica **154** (2018), no. 7, 1333–1361.

[51] Simon Wacker, *Cellular automata on group sets and the uniform curtis–hedlund–lyndon theorem*, Natural Computing **18** (2019), no. 3, 459–487.

[52] Peter Walters, *An introduction to ergodic theory*, vol. 79, Springer Science & Business Media, 2000.

[53] Benjamin Weiss, *Subshifts of finite type and sofic systems*, Monatshefte für Mathematik **77** (1973), no. 5, 462–474.

[54] ______, *Entropy and actions of sofic groups*, Discrete Contin. Dyn. Syst. Ser. B **20** (2015), no. 10, 3375–3383.