TEMPORO-SPATIAL DIFFERENTIATIONS

Aidan J. Young

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics.

Chapel Hill
2023

Approved by:

Idris Assani

Justin Sawon

Anush Tserunyan

Benjamin Weiss

Mark Williams

ABSTRACT

AIDAN J. YOUNG: Temporo-Spatial Differentiations
(Under the direction of Idris Assani)

Temporo-spatial differentiation problems were first introduced in (Assani and Young, 2022) under the name of spatial-temporal differentiation problems. Given a probability space (X, μ) and measurable map $T : X \to X$, a temporo-spatial differentiation problem is concerned with the limiting behavior of the sequence

$$\frac{1}{\mu(C_k)} \int_{C_k} \frac{1}{k} \sum_{j=0}^{k-1} f \circ T^j \mathrm{d}\mu,$$

where $f \in L^1(X, \mu)$ and $(C_k)_{k=1}^\infty$ is a sequence of measurable subsets of X with positive measure. These problems were then generalized to the setting of non-autonomous dynamical systems in (Assani and Young, 2023).

We will present several of the basic aspects of temporo-spatial differentiation problems, including their connections with the field of ergodic optimization. We also present several positive convergence results for "random temporo-spatial differentiation problems." We then discuss generalizations of temporo-spatial differentiation problems to the setting of actions of groups and semigroups other than $\mathbb{Z}$, as well as the construction of pathological temporo-spatial differentiation problems.

A.M.D.G.

TABLE OF CONTENTS

Introduction

Dynamical systems is classically concerned with the study of some space X equipped with a map $T : X \to X$ that respects some structure of X. Typically, we are especially interested in questions about the "long-term behavior" of these systems, though exactly what long-term behavior means can vary contextually. Measure-theoretic dynamical systems is interested in measurable or measure-preserving maps $T : X \to X$ on a measure space $X = (X, \mu)$, and topological dynamical systems is interested in continuous maps $T : X \to X$ on a compact metrizable topological space X. As a rule, one could consider a field of dynamical systems for every kind of structure: a field of smooth dynamical systems that studies C^r maps $T : X \to X$ on C^r manifolds X, a field of algebraic dynamics that studies endomorphisms $T : X \to X$ of some algebraic object (e.g. abstract groups, topological groups, fields) X, a field of complex dynamics that studies holomorphic maps $T : X \to X$ of a complex manifold X, a field of C*-dynamical systems that studies endomorphisms $T : X \to X$ of C*-algebras X, and so on.

Though there are many separate fields within dynamical systems, in practice we're often interested in the interplay between the dynamics of different structures on a space. For example, we might consider the case where X is a compact metrizable space equipped with a continuous map $T : X \to X$, and μ is a Borel measure on X, making T also measurable; do the measure-theoretic dynamics force certain properties on the topological dynamics, and vice-versa? Given the frequent interplay between measure-theoretic and topological dynamics, it benefits us to not be too rigid about where the boundaries between these fields of dynamics fall. For our purposes, we understand *ergodic theory* to be the study of measurable maps $T : X \to X$ on a probability space (X, μ) -especially their long-term behaviors- as well as of the interplay between measurable dynamics and other fields of dynamics.

More generally, dynamical systems are not just concerned with the dynamics of a single map, but with the dynamics of semigroup actions on spaces. In this broader context, the view of dynamics presented above can be understood as the study of actions of the additive group $\mathbb{N}_0$ of nonnegative integers, where $n \cdot x = T^n x$. At times, we will be interested in the more general case where we allow semigroups other than $\mathbb{N}_0$ to act on

our spaces. However, the study of $\mathbb{N}_0$ actions we've described above is essentially the classical setting for the field. When presenting new concepts, we will often describe them in terms of actions of $\mathbb{N}_0$ if we feel that whatever phenomenon we wish to draw attention to can be seen clearly in this classical setting.

One of the classical areas of study in ergodic theory is the study of ergodic averages: Given a measurable function $f : X \to \mathbb{C}$, what can be said about the limiting behavior of the sequence

$$\left(\frac{1}{k} \sum_{j=0}^{k-1} f \circ T^j \right)_{k=1}^{\infty} ?$$

These averages $k^{-1} \sum_{j=0}^{k-1} f \circ T^j$ are referred to as *ergodic averages*, and admit a physical intuition: If T describes the change in a system X over time, and f is a measurement of X, then $k^{-1} \sum_{j=0}^{k-1} f \circ T^j$ can be understood as the average measurement over a period of time k. In light of this physical description of the ergodic averages, we sometimes call them "temporal averages."

Two foundational results in ergodic theory, namely the von Neumann and Birkhoff Ergodic Theorems, describes the limiting behavior of these ergodic averages in the senses of L^p convergence and pointwise-almost everywhere convergence, respectively.

Theorem 1.0.1 (Von Neumann Ergodic Theorem). *(c.f. (Neumann, 1932)) Consider a probability space $X = (X, \mu)$ along with a μ-preserving map $T : X \to X$. Let $f \in L^2(X, \mu)$. Then*

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f \circ T^j$$

exists in the $L^2(X, \mu)$ norm.

Theorem 1.0.2 (Birkhoff Ergodic Theorem). *(c.f. (Birkhoff, 1931)) Consider a probability space $X = (X, \mu)$ along with a measure-preserving map $T : X \to X$. Let $f \in L^1(X, \mu)$. Then for μ-almost all $x \in X$, the limit*

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f\left(T^j x\right)$$

exists.

We will often write $T^n f := f \circ T^n$, i.e. regard composition with T as an operator which we also write as T.

Another classical area of study within measure theory is the study of spatial derivatives. Given a measure space (X, μ), a sequence $(C_k)_{k=1}^{\infty}$ of measurable subsets of X satisfying $\mu(C_k) \in (0, \infty)$ for all $k \in \mathbb{N}$, and an integrable function $f \in L^1(X, \mu)$, what can be said about the limiting behavior of the sequence

$$\left(\frac{1}{\mu(C_k)} \int_{C_k} f \mathrm{d}\mu \right)_{k=1}^{\infty} ?$$

One classical result along these lines is the Lebesgue Differentiation Theorem (Folland, 1999, Theorem 3.21). A fuller treatment of the study of local spatial differentiation theorems of this kind can be found in (De Guzman, 1976).

This is where we situate the problem that interests us, which we call the temporo-spatial differentiation problem. Consider a probability space (X, μ), and a measurable transformation $T : X \to X$. Consider further a sequence $(C_k)_{k=1}^{\infty}$ of measurable subsets of X such that $\mu(C_k) > 0$ for all $k \in \mathbb{N}$, and a function $f \in L^1(X, \mu)$. What can be said about the limiting behavior of the sequence

$$\left(\frac{1}{\mu(C_k)} \int_{C_k} \frac{1}{k} \sum_{j=0}^{k-1} f \circ T^j \mathrm{d}\mu \right)_{k=1}^{\infty} ?$$

The temporo-spatial differentiation problem was first proposed under the name "spatial-temporal differentiation" in a joint work between the present author and his advisor, I. Assani, in (Assani and Young, 2022). Its name reflects that it is a synthesis of two averaging processes: the "temporal averaging" $f \mapsto \frac{1}{k} \sum_{j=0}^{k-1} f \circ T^j$ and the "spatial averaging" $f \mapsto \frac{1}{\mu(C_k)} \int_{C_k} f \mathrm{d}\mu$. The present author and I. Assani agreed to change the name when we found the older one difficult to pronounce consistently. The nuance of these problems lies in the potential "conflict" between the tendencies of the temporal and spatial averaging processes.

Chapter 2 of this dissertation discusses the temporo-spatial differentiation problem in the context of actions of $\mathbb{Z}$ on a compact metric space X by homeomorphisms that preserve a Borel probability measure on X. We provide some characterizations of when a temporo-spatial differentiation can behave well, and also provide certain convergence results. This chapter is based largely on a joint paper between the present author and I. Assani (Assani and Young, 2022).

Chapter 3 generalizes some of the general results of Chapter 2 to the setting of actions of topological groups. We also discuss several instances of what we call a pointwise reduction heuristic: a temporo-spatial differentiation over a spatial averaging sequence $(C_k)_{k=1}^{\infty}$ of sets containing a fixed point with diameter

going to 0 sufficiently fast, then the temporo-spatial differentiation will have the same limiting behavior as an ergodic average at the point. This heuristic then allows us to convert pointwise convergence results into convergence results for so-called "random temporo-spatial differentiation problems."

Chapter 4 introduces what we call multi-local temporo-spatial differentiations, i.e. temporo-spatial differentiations with respect to finite unions of balls. We demonstrate a multi-local form of the pointwise reduction heuristic from Chapter 3, and then present several constructions of pathological temporo-spatial differentiations.

In Chapter 5, we discuss a generalization of temporo-spatial differentiations to the setting of non-autonomous dynamical systems. We consider random temporo-spatial differentiation problems for systems of endomorphisms of compact abelian metrizable groups, showing that these problems are almost surely well-behaved. We then demonstrate that under mild conditions, the *topologically* generic behavior is that these random temporo-spatial differentiations will diverge. This chapter is based on a joint paper between the present author and I. Assani (Assani and Young, 2023).

In Chapter 6, we extend the notions of ergodic optimization to the setting of so-called "non-commutative topological dynamical systems," i.e. C*-dynamical systems. We consider relative ergodic optimization for C*-dynamical systems, where we attempt to optimize relative to a constrained family of invariant states, and extend several elementary results from the classical setting to the relative non-commutative setting. We also provide alternate proofs of several results using techniques from nonstandard analysis.

Temporo-spatial differentiations for actions of $\mathbb{Z}$

This chapter is based on the article (Assani and Young, 2022), a joint work between the author and his advisor, I. Assani. The only changes made have been as follows:

- The abstract was removed.

- We replaced phrases like "this article" with "this chapter."

- We replaced the phrase "spatial-temporal differentiation" with "temporo-spatial differentiation," for reasons described in Chapter 1.

- In the second paragraph of Section 2.1, we replaced a reference to a secondary source with a reference to the article in which the result originated.

- We revised Remark 2.1.2. In (Assani and Young, 2022), this remark indicated a future line of investigation. However, since that article was accepted, we pursued that line of investigation further, and in fact we pursue that line of investigation further in later chapters of this dissertation. Remark 2.1.2 has been revised to reflect this.

- We added this paragraph to explain the relation between (Assani and Young, 2022) and the current chapter.

All these changes were made so that the contents of the article would make sense in the context of this dissertation. Aside from those changes listed, this chapter is identical in content to (Assani and Young, 2022).

Let $(X, \mathcal{B}, \mu, T)$ be an ergodic topological dynamical system, where X is a compact metric space, $\mathcal{B}$ is the Borel σ-algebra of X, and $T : X \to X$ a homeomorphism that is ergodic with respect to the probability measure μ. We consider *temporo-spatial differentiation problems* of the type

$$\lim_{k \to \infty} \frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \mathrm{d}\mu,$$

5

where $(F_k)_{k=1}^\infty$ is a sequence of measurable sets $F_k \in \mathcal{B}$ with positive measure $\mu(F_k) > 0$; specifically, we consider questions of when this limit exists, and when it exists, what that limit is for $f \in L^\infty(X, \mu)$.

Before proceeding, we pause to distinguish these problems from two other kinds of differentiation problems which we will call temporal and spatial differentiation problems. A temporal differentiation problem typically considers properties of the sequence $\left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right)_{k=1}^\infty$, especially convergence properties, and a spatial differentiation problem might consider convergence properties of sequences of the form $\left(\frac{1}{\mu(F_k)} \int_{F_k} f \, d\mu \right)$, where F_k are sets of positive measure. What we describe here as temporal differentiation is the domain of the classical ergodic theorems, and results exist regarding spatial differentiation problems (e.g. the Lebesgue Differentiation Theorem (Folland, 1999, 3.21), Fundamental Theorem of Calculus). Our problem, however, fits in neither of these bins, except in trivial cases, and these differentiation problems might be called "temporo-spatial" differentiation problems.

A temporo-spatial differentiation problem $\left(\frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) d\mu \right)_{k=1}^\infty$ hinges on three parameters: the dynamical system $(X, \mathcal{B}, \mu, T)$, the sequence $(F_k)_{k=1}^\infty$ of measurable sets, and the function $f \in L^\infty(X, \mu)$. For the most part, the questions we consider in this chapter can be understood as "fixing" two of these parameters and investigating what can be said about the convergence properties of the differentiation when the remaining parameter is allowed to "vary".

The chapter is organized as follows:

1. In Section 2.1, we consider certain functions which behave particularly well with respect to these differentiations, called uniform functions, and analyze them with respect to these temporo-spatial differentiations. We pay special attention to topological dynamical systems and how these temporo-spatial differentiations interact with unique ergodicity and uniformity.

2. In Section 2.2, we consider a non-expansive topological dynamical system, and consider temporo-spatial differentiations along certain random nested sequences of subsets, deriving probabilistic results.

3. In Section 2.3, we consider instead a broader class of Lipschitz maps, and differentiate along randomly chosen sequences of sets; in particular, we derive probabilistic results about temporo-spatial differentiations along random sequences of cylinders in a subshift, as well as find certain pathological counterexamples.

4. In Section 2.4, we turn to study differentiations along random cylinders on Bernoulli shifts, but using a more probabilistic set of tools different from those we employed in the second section. We then use these techniques to consider a different problem of random cylinders, where we allow the cylinders at different steps to have different centers.

2.1 Uniform functions and differentiation theorems

In this section, we consider questions of the following forms: Given an appropriate system $(X, \mathcal{B}, T, \mu)$, are there $f \in L^\infty(X, \mu)$ for which $\left(\frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \mathrm{d}\mu \right)_{k=1}^\infty$ converges for *all* choices of $(F_k)_{k=1}^\infty$? On the other hand, are there restrictions we can place on $(X, \mathcal{B}, \mu, T)$ to ensure that $\left(\frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \mathrm{d}\mu \right)_{k=1}^\infty$ converges for *all* choices of $(F_k)_{k=1}^\infty$ *and* all $f \in C(X)$? The answer to the former question will be centered around the notion of a uniform function (defined below), and the answer to the latter question will be centered around unique ergodicity.

Let X be a compact metrizable space with Borel σ-algebra $\mathcal{B}$, and let $T : X \to X$ be a homeomorphism. Then (X, T) is uniquely ergodic iff the sequence $\left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right)_{k=1}^\infty$ converges in $C(X)$ to a constant function for all $f \in C(X)$ (Oxtoby, 1952, (5.3)), and when this happens, the sequence converges to $\int f \mathrm{d}\mu$, where μ is the unique ergodic T-invariant Borel probability measure. Thus if $(F_k)_{k=1}^\infty$ is *any* sequence of measurable sets of positive measure, then $\frac{1}{\mu(F_k)} \int_{F_k} \frac{1}{k} \sum_{i=0}^{k-1} T^i f \mathrm{d}\mu \to \int f \mathrm{d}\mu$ for all $f \in C(X)$, since $\alpha_{F_k} : f \mapsto \frac{1}{\mu(F_k)} \int_{F_k} f \mathrm{d}\mu$ is a state on $L^\infty(X, \mu)$. Fix $\epsilon > 0$, and choose $K \in \mathbb{N}$ such that

$$ k \geq K \Rightarrow \left\| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \leq \epsilon. $$

Then if $k \geq K$, we have

$$ \left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{F_k} \left(T^i f \right) \right| = \left| \alpha_{F_k} \left(\int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right| $$

$$ \leq \left\| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty $$

$$ \leq \epsilon. $$

More generally, if we have any dynamical system $(Y, \mathcal{A}, \nu, S)$, we can call a function $g \in L^\infty(Y, \nu)$ *uniform* if $\frac{1}{k} \sum_{i=0}^{k-1} S^i g \to \int g \, d\nu$ in L^∞. Let $\mathscr{U}(Y, \mathcal{A}, \nu, S) \subseteq L^\infty(Y, \nu)$ denote the space of all uniform functions on $(Y, \mathcal{A}, \nu, S)$. If g is uniform, then for any sequence $(G_k)_{k=1}^\infty$ of measurable sets of positive measure, we have

$$\frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\nu(G_k)} \int_{G_k} S^i g \, d\nu \to \int g \, d\nu,$$

meaning that essentially any differentiation problem of the type that interests us will behave exceptionally well for that g.

Whenever (X, T) is a uniquely ergodic system, we have $C(X) \subseteq \mathscr{U}(X, \mathcal{B}, \mu, T)$, since

$$\left\| \int f \, d\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \leq \left\| \int f \, d\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_{C(X)}.$$

We collect here a few results about some more general differentiation problems. We first demonstrate a general characterization theorem for uniform functions.

Theorem 2.1.1. *Let $(Y, \mathcal{A}, \nu, S)$ be an ergodic dynamical system, and let $g \in L^\infty(Y, \nu)$. Then g is uniform if and only if for all sequences $(G_k)_{k=1}^\infty$ in $\mathcal{A}$ of measurable sets of positive measure,*

$$\frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\nu(G_k)} \int_{G_k} S^i g \, d\nu \to \int g \, d\nu.$$

Proof. ($\Rightarrow$) If g is uniform, then

$$
\begin{aligned}
\left| \int g \, d\nu - \frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\nu(G_k)} \int_{G_k} S^i g \, d\nu \right| &= \left| \frac{1}{\nu(G_k)} \int_{G_k} \left(\int g \, d\nu - \frac{1}{k} \sum_{i=0}^{k-1} S^i g \right) d\nu \right| \\
&\leq \frac{1}{\nu(G_k)} \int_{G_k} \left\| \int g \, d\nu - \frac{1}{k} \sum_{i=0}^{k-1} S^i g \right\|_\infty \\
&= \left\| \int g \, d\nu - \frac{1}{k} \sum_{i=0}^{k-1} S^i g \right\|_\infty \\
&\to 0.
\end{aligned}
$$

($\Leftarrow$) Suppose that g is not uniform, and set $h_k = \int g \, d\nu - \frac{1}{k} \sum_{i=0}^{k-1} S^i g$. Then $\limsup_{k \to \infty} \|h_k\|_\infty > 0$. Breaking h_k into its real part h_k^{Re} and imaginary part h_k^{Im} tells us that ei-

ther $\limsup_{k\to\infty} \left\|h_k^{\mathrm{Re}}\right\|_\infty > 0$, or $\limsup_{k\to\infty} \left\|h_k^{\mathrm{Im}}\right\|_\infty > 0$. Suppose without loss of generality that $\limsup_{k\to\infty} \left\|h_k^{\mathrm{Re}}\right\|_\infty > 0$. Then at least one of the inequalities

$$\nu\left(\left\{y \in Y : h_k^{\mathrm{Re}}(y) \geq \frac{\epsilon_0}{2}\right\}\right) > 0, \ \nu\left(\left\{y \in Y : h_k^{\mathrm{Re}}(y) \leq -\frac{\epsilon_0}{2}\right\}\right) > 0$$

attains for infinitely many $k \in \mathbb{N}$. Assume without loss of generality that $I = \left\{k \in \mathbb{N} : \nu\left(\left\{y \in Y : h_k^{\mathrm{Re}}(y) \geq \frac{\epsilon_0}{2}\right\}\right) > 0\right\}$ is an infinite set.

Construct a sequence $(G_k)_{k=1}^\infty$ by letting $G_k = \left\{y \in Y : h_k^{\mathrm{Re}}(y) \geq \frac{\epsilon_0}{2}\right\}$ for all $k \in I$, and $G_k = Y$ for $k \in \mathbb{N} \setminus I$. Then if $k \in I$, then

$$\left|\frac{1}{\nu(G_k)} \int_{G_k} h_k \mathrm{d}\nu\right| \geq \left|\frac{1}{\nu(G_k)} \int_{G_k} h_k^{\mathrm{Re}} \mathrm{d}\nu\right|$$

$$= \frac{1}{\nu(G_k)} \int_{G_k} h_k^{\mathrm{Re}} \mathrm{d}\nu$$

$$\geq \frac{1}{\nu(G_k)} \int_{G_k} \frac{\epsilon_0}{2} \mathrm{d}\nu$$

$$= \frac{\epsilon_0}{2}.$$

Therefore, there exist infinitely many $k \in \mathbb{N}$ such that

$$\left|\int g \mathrm{d}\nu - \frac{1}{k}\sum_{i=0}^{k-1} \frac{1}{\nu(G_k)} \int_{G_k} S^i g \mathrm{d}\nu\right| = \left|\frac{1}{\nu(G_k)} \int_{G_k} h_k \mathrm{d}\nu\right| \geq \frac{\epsilon_0}{2},$$

meaning that $\left|\int g \mathrm{d}\nu - \frac{1}{k}\sum_{i=0}^{k-1} \frac{1}{\nu(G_k)} \int_{G_k} S^i g \mathrm{d}\nu\right| \not\to 0$. $\qquad\square$

Remark 2.1.2. In this chapter, we consider averages of the form $\frac{1}{k}\sum_{i=0}^{k-1} \frac{1}{\mu(F_k)} \int_{F_k} T^i f \mathrm{d}\mu$, where $T^{(\cdot)}$ is a probability measure-preserving action of $\mathbb{Z}$ on the probability space $(X, \mathcal{B}, \mu)$, and $F_k \in \mathcal{B}$ is a set of positive measure. We could extend our scope to consider probability measure-preserving actions $T^{(\cdot)}$ of an amenable group G on $(X, \mathcal{B}, \mu)$, and consider averages of the form $\frac{1}{|A_k|}\sum_{g \in A_k} \frac{1}{\mu(F_k)} \int_{F_k} T^g f \mathrm{d}\mu$, where $(A_k)_{k=1}^\infty$ is a Følner sequence for G. This will be the subject of chapters 3 and 4.

Because we will so frequently be considering averages of functions over sets of positive measures, it will benefit us to introduce the following notation.

Notation 2.1.3. Let $(X, \mathcal{B}, \mu)$ be a probability space. When $F \in \mathcal{B}$ is a set of positive measure $\mu(F) > 0$, we denote by α_F the state on $L^\infty(X, \mu)$ given by

$$\alpha_F(f) := \frac{1}{\mu(F)} \int_F f \mathrm{d}\mu.$$

Theorem 2.1.1 hints at why we consider temporo-spatial differentiations of L^∞ functions instead of, for example, differentiations of L^p functions for $p \in [1, \infty)$. One might plausibly propose that if we have a uniquely ergodic dynamical system $(X, \mathcal{B}, \mu, T)$, then we can observe that for all $f \in C(X)$, all temporo-spatial differentiations converge to $\int f \mathrm{d}\mu$. We could then try to extend this convergence to all of $L^1(X, \mu)$, since $C(X)$ is L^1-dense in $L^1(X, \mu)$. However, we know that a uniquely ergodic dynamical system can still have non-uniform L^∞ functions (in fact, any ergodic dynamical system over a non-atomic standard probability space will have them, as seen in Proposition 2.1.14), so this cannot be right. The catch is that for measurable F of nonzero measure, the functional $\alpha_F : f \mapsto \frac{1}{\mu(F)} \int_F f \mathrm{d}\mu$ is of norm 1 with respect to L^∞, but the same can't be said relative to L^p for $p \in [1, \infty)$. As such, the "natural" choice of function for a temporo-spatial differentiation is an L^∞ function.

A similarly plausible but misguided attempt to establish convergence results of temporo-spatial differentiations for all $f \in L^\infty(X, \mu)$ could be through the concept of uniform sets. In (Hansel et al., 1973, Theorem 1), it was established that if $\mathcal{B}$ is separable with respect to the metric $(A, B) \mapsto \mu(A \Delta B)$, then there exists a dense T-invariant subalgebra $\mathcal{B}' \subseteq \mathcal{B}$ of sets such that χ_B is uniform for all $B \in \mathcal{B}'$. Again, one might propose that we could use a density argument to extend convergence results on temporo-spatial differentiations to functions χ_A for all $A \in \mathcal{B} \supseteq \mathcal{B}'$. But again, Theorem 2.1.1 tells us that this would be tantamount to proving that all L^∞ functions are uniform, and we know that there can exist non-uniform L^∞ functions.

Other results are possible regarding topological dynamical systems, as we show below.

Lemma 2.1.4. *Let* $f \in L^\infty(X, \mu)$, *where* $(X, \mathcal{B}, \mu, T)$ *is a dynamical system. Then the sequence* $\left(\left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \right)_{k=1}^\infty$ *is convergent, and*

$$\lim_{k \to \infty} \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty = \inf_{k \in \mathbb{N}} \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty.$$

Proof. Let $a_k = \left\| \sum_{i=0}^{k-1} T^i f \right\|_\infty$. Then the sequence $(a_k)_{k=1}^\infty$ is subadditive. This follows since if $k, \ell \in \mathbb{N}$, then

$$
\begin{aligned}
a_{k+\ell} &= \left\| \sum_{i=0}^{k+\ell-1} T^i f \right\|_\infty \\
&\le \left\| \sum_{i=0}^{k-1} T^i f \right\|_\infty + \left\| \sum_{i=k}^{k+\ell-1} T^i f \right\|_\infty \\
&= \left\| \sum_{i=0}^{k-1} T^i f \right\|_\infty + \left\| T^k \sum_{i=0}^{\ell-1} T^i f \right\|_\infty \\
&= \left\| \sum_{i=0}^{k-1} T^i f \right\|_\infty + \left\| \sum_{i=0}^{\ell-1} T^i f \right\|_\infty \\
&= a_k + a_\ell.
\end{aligned}
$$

The result then follows from the Subadditivity Lemma. $\square$

Definition 2.1.5. For $f \in L^\infty(X, \mu)$, set

$$
\Gamma(f) := \lim_{k \to \infty} \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty.
$$

We call this value $\Gamma(f)$ the *gauge* of f.

This $\Gamma(f)$ satisfies the inequality $\Gamma(f) \ge \int f \, \mathrm{d}\mu$, since

$$
\frac{1}{k} \sum_{i=0}^{k-1} T^i f \le \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty
$$

almost surely, implying that

$$
\int f \, \mathrm{d}\mu = \int \frac{1}{k} \sum_{i=0}^{k-1} T^i f \, \mathrm{d}\mu \le \int \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \mathrm{d}\mu = \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty.
$$

Therefore

$$
\int f \, \mathrm{d}\mu \le \inf_{k \in \mathbb{N}} \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty = \Gamma(f).
$$

Definition 2.1.6. Let X be a compact metric space, and let $C_\mathbb{R}(X)$ denote the (real) space of real-valued continuous functions on X endowed with the uniform norm $\| \cdot \|_{C(X)}$. Let $T : X \to X$ be a continuous

11

homeomorphism, and let $\mathcal{M}_T$ denote the family of all T-invariant Borel probability measures on X. A measure $\mu \in \mathcal{M}_T(X)$ is called *f-maximizing* for some $f \in C_\mathbb{R}(X)$ if $\int f \mathrm{d}\mu = \sup_{\nu \in \mathcal{M}_T} \int f \mathrm{d}\nu$. We denote by $\mathcal{M}_{\max}(f)$ the set of all f-maximizing measures.

The definition of maximizing measures is due to Jenkinson (Jenkinson, 2006a, Definition 2.3). The definition is topological in nature, in the sense that it is defined with reference to a homeomorphism on a compact metric space prior to any other measure that the metric space might possess. A result of Jenkinson (Jenkinson, 2006a, Proposition 2.4) tells us that for every $f \in C_\mathbb{R}(X)$, we have

1. $\mathcal{M}_{\max}(f) \neq \emptyset$,

2. $\mathcal{M}_{\max}(f)$ is a compact metrizable simplex, and

3. the extreme points of $\mathcal{M}_{\max}(f)$ are exactly the ergodic f-maximizing measures. In particular, every $f \in C_\mathbb{R}(X)$ admits an ergodic f-maximizing measure.

For every nonnegative $f \in C_\mathbb{R}(X)$, let μ_f denote an ergodic maximizing measure for f. We claim that $\Gamma(f) \leq \int f \mathrm{d}\mu_f$. To prove this, we note that $\left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \leq \max_{x \in X} \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x)$, where the maximum exists because X is compact and $f \in C_\mathbb{R}(X)$ is continuous. Choose $x_k \in X$ such that $\max_{x \in X} \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) = \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x_k)$. Let δ_{x_k} denote the Borel point-mass probability measure

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

Let $\mu_k = \frac{1}{k} \sum_{i=0}^{k-1} \delta_{T^i x_k}$, so that $\frac{1}{k} \sum_{i=0}^{k-1} T^i f(x_k) = \int f \mathrm{d}\mu_k$.

Since the space of Borel probability measures on X is compact in the weak* topology on $C(X)^*$, there exists a subsequence $(\mu_{k_n})_{n=1}^\infty$ of $(\mu_k)_{k=1}^\infty$ converging to a Borel probability measure μ'. It follows from a classical calculation that μ' is T-invariant.

Therefore, μ' is a T-invariant Borel probability measure on X such that $\int f \mathrm{d}\mu' = \Gamma(f)$. But if μ_f is f-maximal, then

$$\Gamma(f) = \int f \mathrm{d}\mu' \leq \int f \mathrm{d}\mu_f.$$

Under certain conditions, however, we can achieve equality here.

Lemma 2.1.7. *Let $(X, \mathcal{B}, \mu)$ be a probability space, where X is a compact metric space with Borel σ-algebra on X denoted by $\mathcal{B}$. Let $T : X \to X$ be a homeomorphism. If μ is strictly positive, and $f \in C_\mathbb{R}(X)$ is nonnegative, then*

$$\Gamma(f) = \int f \mathrm{d}\mu_f.$$

Proof. If μ is strictly positive, then the L^∞ norm restricted to $C(X)$ agrees with the uniform norm on $C(X)$ (see the discussion after (Folland, 1999, Theorem 6.8)). Therefore $\left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty = \sup_{x \in X} \left| \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) \right|$ for all $k \in \mathbb{N}$. However, we can bound $\int f \mathrm{d}\mu_f$ by

$$
\begin{aligned}
\int f \mathrm{d}\mu_f &= \int \frac{1}{k} \sum_{i=0}^{k-1} T^i f \mathrm{d}\mu \\
&\leq \sup_{x \in X} \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) \\
&= \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \\
\Rightarrow \int f \mathrm{d}\mu_f &\leq \inf_{k \in \mathbb{N}} \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty \\
&= \Gamma(f),
\end{aligned}
$$

establishing the opposite inequality. □

Lemma 2.1.8. *Suppose that $(X, \mathcal{B}, \mu, T)$ consists of a compact metric space X with Borel σ-algebra $\mathcal{B}$ and a strictly positive probability measure μ that is ergodic with respect to a homeomorphism $T : X \to X$. Then the system (X, T) is uniquely ergodic if and only if $\Gamma(f) = \int f \mathrm{d}\mu$ for all nonnegative $f \in C_\mathbb{R}(X)$.*

Proof. ($\Rightarrow$) If (X, T) is uniquely ergodic, then in particular $\mu = \mu_f$ for all nonnegative $f \in C_\mathbb{R}(X)$. Therefore by the previous lemma, we have

$$\int f \mathrm{d}\mu = \int f \mathrm{d}\mu_f = \Gamma(f).$$

($\Leftarrow$) If (X, T) is *not* uniquely ergodic, then we know that $\mathcal{M}_T(X)$ is not a singleton, and thus contains another ergodic measure ν. By a result of Jenkinson (Jenkinson, 2006a, Theorem 3.7), we know that there exists $f \in C(X)$ real-valued such that $\nu = \mu_f$ is the *unique f-maximizing measure*. We may assume without

loss of generality that f is nonnegative, since otherwise we can replace f with $\tilde{f} - \inf_{x \in X} f(x)$. Since we claimed that ν was the unique f-maximizing measure, we can conclude in particular that

$$
\begin{aligned}
\int f \mathrm{d}\mu &< \int f \mathrm{d}\nu \\
&= \int f \mathrm{d}\mu_f \\
&= \Gamma(f).
\end{aligned}
$$

$\square$

Theorem 2.1.9. *Suppose that $(X, \mathcal{B}, \mu, T)$ consists of a compact metric space X with Borel σ-algebra $\mathcal{B}$ and a probability measure μ that is ergodic with respect to a homeomorphism $T : X \to X$. Then the following results are related by the implications (1)$\Rightarrow$(2)$\Rightarrow$(3). Further, if μ is strictly positive, then (3)$\Rightarrow$(1).*

1. *(X, T) is uniquely ergodic.*

2. *For every sequence of Borel-measurable sets $(F_k)_{k=1}^{\infty}$ of positive measure, and for every $f \in C(X)$, the limit $\lim_{k \to \infty} \alpha_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right)$ exists and is equal to $\int f \mathrm{d}\mu$, where $\alpha.$ is as defined in Notation 2.1.3.*

3. *For every sequence of open sets $(U_k)_{k=1}^{\infty}$ of positive measure, and for every $f \in C(X)$, the limit $\lim_{k \to \infty} \alpha_{U_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right)$ exists and is equal to $\int f \mathrm{d}\mu$, where $\alpha.$ is as defined in Notation 2.1.3.*

Proof. (1)$\Rightarrow$(2): If (X, T) is uniquely ergodic, then $\left\| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_{C(X)} \overset{k \to \infty}{\to} 0$, so

$$
\begin{aligned}
\left| \int f \mathrm{d}\mu - \alpha_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right| &= \left| \alpha_{F_k} \left(\int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right| \\
&\leq \left\| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_{C(X)} \\
&\overset{k \to \infty}{\to} 0.
\end{aligned}
$$

(2)$\Rightarrow$(3): Trivial, since an open set is automatically Borel.

$\neg$(1)$\Rightarrow$ $\neg$(3): Suppose (X, T) is *not* uniquely ergodic, and that μ is strictly positive. Then Lemma 2.1.8 tells us that there exists nonnegative $f \in C_{\mathbb{R}}(X)$ for which $\Gamma(f) > \int f \mathrm{d}\mu$. Let L be such that

14

$\int f \mathrm{d}\mu < L < \Gamma(f)$, and consider the open set

$$
U_k = \left\{ x \in X : \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) > L \right\}.
$$

By the proof of Lemma 2.1.7, we know that

$$
\Gamma(f) = \inf_{k \in \mathbb{N}} \max_{x \in X} \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) = \inf_{k \in \mathbb{N}} \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x_k),
$$

where $x_k \in U_k$ for all $k \in \mathbb{N}$. Therefore U_k is a nonempty open set, and since μ is strictly positive, that means $\mu(U_k) > 0$. Therefore

$$
\alpha_{U_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \geq \alpha_{U_k}(L)
$$

$$
= L
$$

$$
\Rightarrow \liminf_{k \to \infty} \alpha_{U_k}(f) \geq L
$$

$$
> \int f \mathrm{d}\mu.
$$

$\square$

Theorem 2.1.10. *Suppose that $(X, \mathcal{B}, \mu, T)$ consists of a compact connected metric space $X = (X, \rho)$ with Borel σ-algebra $\mathcal{B}$ and a probability measure μ that is ergodic with respect to a homeomorphism $T : X \to X$. Suppose further that μ is strictly positive, but (X, T) is not uniquely ergodic. Then there exists a sequence $(U_k)_{k=1}^{\infty}$ of nonempty open subsets of X and a nonnegative continuous function $f \in C(X)$ such that the sequence $\left(\alpha_{U_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right)_{k=1}^{\infty}$ is not Cauchy. Furthermore, if μ is atomless, then we can choose the sequence $(U_k)_{k=1}^{\infty}$ such that $\mu(U_k) \searrow 0$.*

Proof. Lemma 2.1.8 tells us that there exists nonnegative $f \in C_\mathbb{R}(X)$ for which $\Gamma(f) > \int f \mathrm{d}\mu$. Let $L, M \in \mathbb{R}$ such that $\int f \mathrm{d}\mu < L < M < \Gamma(f)$, and consider the open sets

$$V_k = \left\{ x \in X : \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) > M \right\},$$

$$W_k = \left\{ x \in X : \int f \mathrm{d}\mu < \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) < L \right\}.$$

By the proof of Lemma 2.1.7, we know that $V_k \neq \emptyset$, so let $x_k \in V_k$. We also know that there exists z_k in X such that $f(z_k) \leq \int f \mathrm{d}\mu$, since if $f(z) > \int f \mathrm{d}\mu$ for all $z \in X$, then $\int f(z) \mathrm{d}\mu(z) > \int f \mathrm{d}\mu$, a contradiction. By the Intermediate Value Theorem, there then exists $y_k \in W_k$. Construct $(U_k)_{k=1}^\infty$ as

$$U_k = \begin{cases} V_k, & k \text{ odd} \\[2mm] W_k, & k \text{ even.} \end{cases}$$

Then

$$\limsup_{k \to \infty} \alpha_{U_{2k-1}} \left(\frac{1}{2k-1} \sum_{i=0}^{2k-2} T^i f \right) \geq \limsup_{k \to \infty} \alpha_{U_{2k-1}} (M)$$

$$= M,$$

$$\liminf_{k \to \infty} \alpha_{U_{2k}} \left(\frac{1}{2k} \sum_{i=0}^{2k-1} T^i f \right) \leq \liminf_{k \to \infty} \alpha_{U_{2k}} (L)$$

$$= L.$$

Therefore

$$\liminf_{k \to \infty} \alpha_{U_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \leq L < M \leq \limsup_{k \to \infty} \alpha_{U_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right).$$

Moreover, if μ is atomless, then we can choose $(U_k)_{k=1}^\infty$ so that $\mu(U_k) \to 0$ by letting U_k be a ball of sufficiently small radius contained in V_k (if k is odd) or W_k (if k is even). The above calculations can be carried out in the same way. $\qquad\square$

In Theorem 2.3.7, we construct an example of a Bernoulli shift $(X, \mathcal{B}, \mu, T)$ where there exists $(x, f) \in X \times C(X)$ such that the sequence $\left(\alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right)_{k=1}^\infty$ not only does not converge to $\int f \mathrm{d}\mu$, as in

Theorem 2.1.9, but such that it does not converge at all. Theorem 2.1.10 does not encompass that example, since subshifts are a priori totally disconnected.

In the next result, we will be making use of the Jewett-Krieger Theorem in a specific formulation. This is the formulation originally proven by Jewett in (Jewett, 1970) under the assumption that the transformation was weakly mixing; Bellow and Furstenberg later demonstrated in (Bellow and Furstenberg, 1979) that the parts of Jewett's argument which relied on the weakly mixing property could be proven under the weaker assumption of ergodicity. The version of the Jewett-Krieger Theorem we will be using is as follows.

Jewett-Krieger Theorem. *Given an invertible ergodic system $(Y, \mathcal{A}, \nu, S)$ on a standard probability space $(Y, \mathcal{A}, \nu)$, there exists an essential isomorphism $h : (Y, \mathcal{A}, \nu, S) \to (2^\omega, \mathcal{B}, \mu, T)$ (where 2^ω denotes the Cantor space) such that $(2^\omega, T)$ is a strictly ergodic system.*

The following result provides some structure statements about the space $\mathscr{U}(Y, \mathcal{A}, \nu, S)$ of uniform functions.

Theorem 2.1.11. *Let $(Y, \mathcal{A}, \nu)$ be a standard probability space, and $S : Y \to Y$ an ergodic automorphism. Then $\mathscr{U}(Y, \mathcal{A}, \nu, S)$ is a closed S-invariant subspace of $L^\infty(Y, \nu)$ that is closed under complex conjugation, and contains a unital S-invariant C*-subalgebra A which is dense in $L^1(Y, \nu)$. This A is isomorphic as a C*-subalgebra to $C(2^\omega)$.*

Proof. First, we prove that $\mathscr{U}(Y, \mathcal{A}, \nu, S)$ is a closed S-invariant subspace of $L^\infty(Y, \nu)$. The fact it is a subspace of $L^\infty(Y, \nu)$ is clear, so suppose $f \in \mathrm{cl}(\mathscr{U}(Y, \mathcal{A}, \nu, S))$. Then there exists $g \in \mathscr{U}(Y, \mathcal{A}, \nu, S)$ such that $\|f - g\|_\infty \leq \frac{\epsilon}{3}$. Choose $K \in \mathbb{N}$ such that $k \geq K$ implies $\left\| \int g \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i g \right\|_\infty \leq \frac{\epsilon}{3}$. Then

$$
\begin{aligned}
\left\| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f \right\|_\infty &\leq \left\| \int f \mathrm{d}\mu - \int g \mathrm{d}\mu \right\|_\infty \\
&+ \left\| \int g \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i g \right\|_\infty \\
&+ \left\| \frac{1}{k} \sum_{i=0}^{k-1} T^i (g - f) \right\|_\infty \\
&\leq \|f - g\|_\infty + \left\| \int g \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i g \right\|_\infty + \|f - g\|_\infty \\
&\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}
$$

17

Thus $f \in \mathscr{U}(Y, \mathcal{A}, \nu, S)$. Now, we claim that if $f \in \mathscr{U}(Y, \mathcal{A}, \nu, S)$, then $Sf, S^{-1}f \in \mathscr{U}(Y, \mathcal{A}, \nu, S)$. We compute

$$
\begin{aligned}
\left\| \int Sf \mathrm{d}\nu - \frac{1}{k}\sum_{i=0}^{k-1} S^i(Sf) \right\|_\infty &= \left\| \int f \mathrm{d}\nu - \frac{1}{k}\sum_{i=0}^{k-1} S^i(Sf) \right\|_\infty \\
&= \left\| \int f \mathrm{d}\nu - \left(\frac{1}{k}\sum_{i=0}^{k-1} S^i f \right) + \frac{1}{k}\left(f - S^k f \right) \right\|_\infty \\
&\leq \left\| \int f \mathrm{d}\nu - \frac{1}{k}\sum_{i=0}^{k-1} S^i f \right\|_\infty + \frac{2\|f\|}{k} \\
&\overset{k\to\infty}{\Rightarrow} 0.
\end{aligned}
$$

An analogous argument will show that $S^{-1}f \in \mathscr{U}(Y, \mathcal{A}, \nu, S)$. To see that $\mathscr{U}(Y, \mathcal{A}, \nu, S)$ is also closed under complex conjugation, we see that

$$
\begin{aligned}
\left\| \int \overline{f}\mathrm{d}\mu - \frac{1}{k}\sum_{i=0}^{k-1} T^i\overline{f} \right\|_\infty &= \left\| \overline{\int f \mathrm{d}\mu - \frac{1}{k}\sum_{i=0}^{k-1} T^i f} \right\|_\infty \\
&= \left\| \int f \mathrm{d}\mu - \frac{1}{k}\sum_{i=0}^{k-1} T^i f \right\|_\infty.
\end{aligned}
$$

Finally, we prove that $\mathscr{U}(Y, \mathcal{A}, \nu, S)$ contains a unital S-invariant C*-algebra A that's dense in $L^1(Y, \nu)$. By the Jewett-Krieger Theorem, we know there exists an essential isomorphism $\phi : (Y, \mathcal{A}, \nu, S) \to (2^\omega, \mathcal{B}, \mu, T)$, where $(2^\omega, \mathcal{B}, \mu, T)$ is uniquely ergodic. Let $A = \Phi(C(2^\omega))$, where $\Phi : L^\infty(2^\omega, \mu) \to L^\infty(Y, \nu)$ is the pullback of ϕ. Since $C(2^\omega)$ is dense in $L^1(2^\omega, \mu)$, we can infer that $A = \Phi(C(2^\omega))$ is dense in $L^1(Y, \nu)$. Since continuous functions in a uniquely ergodic system are uniform, it follows that the functions of A are uniform.

Because μ is strictly positive, we know that $C(2^\omega)$ is isomorphic as a C*-algebra to its copy in $L^\infty(2^\omega, \mu)$ (see proof of Lemma 2.1.7), so this map Φ is an isomorphism between $C(2^\omega) \subsetneq L^\infty(2^\omega, \mu)$ and $A = \Phi(C(2^\omega))$. $\qquad\square$

Proposition 2.1.12. *Suppose that $(X, \mathcal{B}, \mu, T)$ consists of a compact metric space $X = (X, \rho)$ with Borel σ-algebra $\mathcal{B}$ and a probability measure μ that is ergodic with respect to a homeomorphism $T : X \to X$, where X is connected. Suppose further that $\exists F \in \mathcal{B}$ such that $0 < \mu(F) < 1$. Then there exists $f \in \mathscr{U}(X, \mathcal{B}, \mu, T) \setminus C(X)$.*

Proof. By the Jewett-Krieger Theorem, there exists an essential isomorphism $h : (X, \mathcal{B}, \mu, T) \to (X', \mathcal{B}', \mu', T')$, where $X' = 2^\omega$ and (X', T') is uniquely ergodic. The topological space 2^ω admits a basis $\mathcal{G}$ of clopen sets. We claim that there exists $G \in \mathcal{G}$ such that $0 < \mu'(G) < 1$.

Assume for contradiction that $\mu'(G) \in \{0, 1\}$ for all $G \in \mathcal{G}$. If $(E_k)_{k=1}^\infty$ is some sequence in $\mathcal{B}'$ of sets for which $\mu'(E_k) \in \{0, 1\}$, then

$$\mu'\left(\bigcup_{k=1}^\infty E_k\right) = \max_{k \in \mathbb{N}} \mu'(E_k) \qquad\qquad \in \{0, 1\},$$

$$\mu'\left(\bigcap_{k=1}^\infty E_k\right) = \min_{k \in \mathbb{N}} \mu'(E_k) \qquad\qquad \in \{0, 1\},$$

$$\mu'(X \setminus E_1) = 1 - \mu(E_1) \qquad\qquad \in \{0, 1\}.$$

But since $\mathcal{G}$ generates $\mathcal{B}'$, this would imply that $\mu'(E) \in \{0, 1\}$ for all $E \in \mathcal{B}'$, a contradiction.

Therefore, there exists $G_0 \in \mathcal{B}'$ clopen such that $0 < \mu'(G_0) < 1$. Set $g = \chi_{G_0} \in C(X') \subseteq \mathscr{U}(X', \mathcal{B}', \mu', T')$, and let $f = g \circ h$. Then $f \in \mathscr{U}(X, \mathcal{B}, \mu, T)$. But since f takes values in $\{0, 1\}$, and $\mu(\{x \in X : f(x) = 1\}) \notin \{0, 1\}$, we must conclude that $f \in \mathscr{U}(X', \mathcal{B}', \mu', T') \setminus C(X)$. $\qquad\square$

We conclude this section by remarking that in most situations, we'll have $\mathscr{U}(Y, \mathcal{A}, \nu, S) \neq L^\infty(Y, \nu)$. We cite here a special case of a result of N. Ormes.

Lemma 2.1.13. *Suppose $(Y, \mathcal{A}, \nu)$ is a non-atomic standard probability space, and $S : Y \to Y$ is an ergodic automorphism. Then there exists a minimal homeomorphism $T : 2^\omega \to 2^\omega$ and an affine homeomorphism $p : [0, 1] \to \mathcal{M}_T(2^\omega)$ for which $(2^\omega, \mathcal{B}, p(0), T)$ is essentially isomorphic to $(Y, \mathcal{A}, \nu, S)$, where $\mathcal{B}$ here denotes the Borel σ-algebra on 2^ω.*

Proof. This is a special case of (Ormes, 1997, Corollary 7.4), where we specifically consider the Choquet simplex $[0, 1]$. $\qquad\square$

Since $(2^\omega, T)$ is not uniquely ergodic, it follows that there exists $f_0 \in C(2^\omega)$ such that $\left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f_0\right)_{k=1}^\infty$ does not converge uniformly to the constant $\int f_0 \, d(p(0))$. Since $(2^\omega, T)$ is minimal, and the support of $p(0)$ is a nonempty T-invariant compact subset of 2^ω, it follows that $p(0)$ is strictly positive, and so the uniform

norm on $C(2^\omega)$ coincides with the $L^\infty(2^\omega, p(0))$ norm on $C(2^\omega)$. As such, it follows that

$$\left\| \int f_0 \mathrm{d}(p(0)) - \frac{1}{k}\sum_{i=0}^{k-1} T^i f_0 \right\|_\infty = \sup_{x\in 2^\omega}\left| \int f_0 \mathrm{d}(p(0)) - \frac{1}{k}\sum_{i=0}^{k-1} T^i f_0(x) \right| \overset{k\to\infty}{\not\to} 0.$$

Let $\phi : (Y, \mathcal{A}, \nu, S) \to (2^\omega, \mathcal{B}, p(0), T)$ be an essential isomorphism, and let $\Phi : L^\infty(2^\omega, p(0)) \to L^\infty(Y, \nu)$ be the pullback of ϕ. Then

$$\left\| \int (\Phi f_0)\,\mathrm{d}\nu - \frac{1}{k}\sum_{i=0}^{k-1} S^i(\Phi f_0) \right\|_\infty = \left\| \int f_0 \mathrm{d}(p(0)) - \frac{1}{k}\sum_{i=0}^{k-1} T^i f_0 \right\|_\infty \overset{k\to\infty}{\not\to} 0.$$

Therefore $\Phi f_0 \in L^\infty(Y,\nu) \setminus \mathscr{U}(Y, \mathcal{A}, \nu, S)$.

The following proposition summarizes this discussion.

Proposition 2.1.14. *Suppose $(Y, \mathcal{A}, \nu)$ is a non-atomic standard probability space, and $S : Y \to Y$ is an ergodic automorphism. Then $\mathscr{U}(Y, \mathcal{A}, \nu, S) \neq L^\infty(Y, \nu)$.*

2.2 Non-expansive maps

In this section, as well as in Section 2.3, we investigate for a certain class of dynamical system $(X, \mathcal{B}, \mu, T)$ what can be said about the convergence properties of $\left(\frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k}\sum_{i=0}^{k-1} T^i f \right) \mathrm{d}\mu \right)_{k=1}^\infty$ for $f \in C(X)$ when we consider a "probabilistically generic" sequence $(F_k)_{k=1}^\infty$. In other words, we investigate in some sense a "typical" behavior of $\left(\frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k}\sum_{i=0}^{k-1} T^i f \right) \mathrm{d}\mu \right)_{k=1}^\infty$ for $f \in C(X)$, and find sufficient conditions for this differentiation to converge almost surely to $\int f \mathrm{d}\mu$ for all $f \in C(X)$.

Let $X = (X, \rho)$ be a compact metric space, and $T : X \to X$ a 1-Lipschitz map, i.e. such that $\rho(Tx, Ty) \leq \rho(x, y)$ for all $x, y \in X$. Let $\mathcal{B}$ denote the Borel σ-algebra on X, and μ a T-invariant, ergodic Borel probability measure on X. Then (X, T) has topological entropy 0, and thus $(X, \mathcal{B}, \mu, T)$ is automatically of entropy $0 < \infty$ (Goodman, 1971, Lemma 1). By the Krieger Generator Theorem (Krieger, 1970, 2.1), the ergodic system admits a finite measurable partition $\mathcal{E} = \{E_d\}_{d\in\mathcal{D}}$ of X such that $\{T^i E_d : i \in \mathbb{Z}, d \in \mathcal{D}\}$ generates the σ-algebra $\mathcal{B}$, where $\mathcal{D}$ is a finite indexing set. We call $\mathcal{E}$ a *generator* of $(X, \mathcal{B}, \mu, T)$.

Let $d_i : X \to \mathcal{D}, i \in \mathbb{Z}$ be the measurable random variable uniquely determined by the relation

$$x \in T^{-i} E_{d_i(x)},$$

or equivalently

$$T^i x \in E_{d_i(x)}.$$

Given a word $\mathbf{a} = (a_0, a_1, \dots, a_{\ell-1}) \in \mathcal{D}^\ell$, we define the *cylinder associated to* $\mathbf{a}$ by

$$[a_0, a_1, \dots, a_{\ell-1}] := \bigcap_{i=0}^{\ell-1} T^{-i} E_{a_i}.$$

We also define the *rank-k cylinder associated to* $x \in X$ by

$$C_k(x) := [d_0(x), d_1(x), \dots, d_{k-1}(x)] = \bigcap_{i=0}^{k-1} T^{-i} E_{d_i(x)}.$$

Equivalently, we can define $C_k(x)$ to be the element of $\vee_{i=0}^{k-1} T^{-i} \mathcal{E}$ containing x.

We note here that $\mu(C_k(x)) > 0$ for all $k \in \mathbb{N}$ for almost all $x \in X$, since

$$\{x \in X : \exists k \in \mathbb{N} \text{ s.t. } \mu(C_k(x)) = 0\} = \bigcup_{k \in \mathbb{N}} \{x \in X : \mu(C_k(x)) = 0\}$$

$$= \bigcup_{k \in \mathbb{N}} \left(\bigcup_{\mathbf{d} \in \mathcal{D}^k \text{ s.t. } \mu([\mathbf{d}])=0} [\mathbf{d}] \right)$$

is a countable union of null sets.

Suppose further that $\mathrm{diam}(C_k(x)) \to 0$ for almost all $x \in X$. Our main result for this section is the following.

The following result states that when an ergodic system $(X, \mathcal{B}, \mu, T)$ is equipped with a generating partition satisfying certain topological conditions, then temporo-spatial differentiations of a continuous function f along a nested sequence of cylinders defined with respect to that generating partition will almost surely converge to the expected value of f.

Theorem 2.2.1. *Let $X = (X, \rho)$ be a compact metric space, and $T : X \to X$ a 1-Lipschitz map, i.e. such that $\rho(Tx, Ty) \leq \rho(x, y)$ for all $x, y \in X$. Let $\mathcal{B}$ denote the Borel σ-algebra on X, and μ a T-invariant, ergodic Borel probability measure on X. Let $\mathcal{E} = \{E_d\}_{d \in \mathcal{D}}$ be a finite measurable partition of X which generates $\mathcal{B}$, and let $C_k(x)$ be the element of $\vee_{i=0}^{k-1} T^{-i} \mathcal{E}$ containing x. Suppose further that*

$\mathrm{diam}(C_k(x)) \to 0$ *for almost all $x \in X$. Then the set of $x \in X$ such that*

$$\frac{1}{\mu(C_k(x))} \int_{C_k(x)} \frac{1}{k} \sum_{i=0}^{k-1} T^i f \mathrm{d}\mu \overset{k \to \infty}{\to} \int f \mathrm{d}\mu$$

for all $f \in C(X)$ is of full measure.

Proof. Since X is compact metrizable, we know that $C(X)$ is a separable vector space, so let $\{f_n\}_{n \in \mathbb{N}}$ be a countable set in $C(X)$ such that $\overline{\mathrm{span}}\{f_n\}_{n \in \mathbb{N}} = C(X)$, where the closure is taken in the uniform norm on $C(X)$. Let

$$S_n = \left\{ x \in X : \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f_n \right) \to \int f_n \mathrm{d}\mu \right\}.$$

We claim that $\mu(S_n) = 1$.

Choose $x \in X$ satisfying the following three conditions:

(a) $\mathrm{diam}(C_k(x)) \to 0$,

(b) $\mu(C_k(x)) > 0$ for all $k \in \mathbb{N}$, and

(c) $\frac{1}{k} \sum_{i=0}^{k-1} T^i f_n(x) \to \int f_n \mathrm{d}\mu$.

The set of $x \in X$ satisfying the condition (a) is of full measure by hypothesis, and the set of $x \in X$ satisfying the condition (b) is of full measure by the discussion preceding the statement of this theorem. Finally, the set of $x \in X$ satisfying condition (c) is of full measure by the Birkhoff ergodic theorem (Walters, 2007, Theorem 1.5). Therefore, the set of $x \in X$ satisfying all three conditions is of full measure.

Fix $\epsilon > 0$. Since f_n is uniformly continuous, we know there exists $\delta > 0$ such that $\rho(x_1, x_2) \leq \delta \Rightarrow$ $|f_n(x_1) - f_n(x_2)| \leq \frac{\epsilon}{2}$. Choose $K_1 \in \mathbb{N}$ such that $\mathrm{diam}(C_k(x)) \leq \delta$ for all $k \geq K_1$. Choose $K_2 \in \mathbb{N}$ such

that $k \geq K_2 \Rightarrow \left| \int f_n \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f_n(x) \right| \leq \frac{\epsilon}{2}$. Let $K = \max\{K_1, K_2\}$, and suppose that $k \geq K$. Then

$$
\left| \int f_n \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f_n \right) \right|
$$

$$
\leq \left| \int f_n \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f_n(x) \right| + \left| \frac{1}{k} \sum_{i=0}^{k-1} T^i f_n(x) - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f_n \right) \right|
$$

$$
\leq \frac{\epsilon}{2} + \frac{1}{k} \sum_{i=0}^{k-1} \left| T^i f_n(x) - \frac{1}{\mu(C_k(x))} \int_{C_k(x)} T^i f \right|
$$

$$
= \frac{\epsilon}{2} + \frac{1}{k} \sum_{i=0}^{k-1} \left| \frac{1}{\mu(C_k(x))} \int_{C_k(x)} T^i f_n(x) - T^i f_n \mathrm{d}\mu \right|
$$

$$
= \frac{\epsilon}{2} + \frac{1}{k} \sum_{i=0}^{k-1} \left| \frac{1}{\mu(T^i C_k(x))} \int_{T^i C_k(x)} f_n(x) - f_n \mathrm{d}\mu \right|
$$

$$
\leq \frac{\epsilon}{2} + \frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(T^i C_k(x))} \int_{T^i C_k(x)} |f_n(x) - f_n| \, \mathrm{d}\mu
$$

$$
\leq \frac{\epsilon}{2} + \frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(T^i C_k(x))} \int_{T^i C_k(x)} \frac{\epsilon}{2} \mathrm{d}\mu
$$

$$
= \epsilon,
$$

since $\mathrm{diam}\left(T^i C_k(x) \right) \leq \mathrm{diam}(C_k(x)) \leq \mathrm{diam}(C_K(x)) < \delta$. Thus if $\mu(C_k(x)) > 0$ for all $k \in \mathbb{N}$, if $\mathrm{diam}(C_k(x)) \to 0$, and if $\frac{1}{k} \sum_{i=0}^{k-1} T^i f_n(x) \to \int f_n \mathrm{d}\mu$, then $x \in S_n$. Thus $\mu(S_n) = 1$ for all $n \in \mathbb{N}$, and so $\mu\left(\bigcap_{n \in \mathbb{N}} S_n \right) = 1$.

We claim now that if $x \in S = \bigcap_{n \in \mathbb{N}} S_n$, then $\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \to \int f \mathrm{d}\mu$ for all $f \in C(X)$. Fix $x \in S, f \in C(X), \epsilon > 0$. Then there exist $N \in \mathbb{N}$ and $z_1, \ldots, z_N \in \mathbb{C}$ such that

$$
\left\| f - \sum_{n=1}^{N} z_n f_n \right\|_\infty < \frac{\epsilon}{3}.
$$

Choose $L_1, \ldots, L_N \in \mathbb{N}$ such that

$$
k \geq L_n \Rightarrow \left| \int f_n \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f_n \right) \right| < \frac{\epsilon}{3N \max\{|z_1|, \ldots, |z_N|, 1\}}.
$$

Abbreviate $g = \sum_{n=1}^{N} z_n f_n$, and let $L = \max\{L_1, \ldots, L_N\}$. Then if $k \geq L$, then

$$\left| \int f \, d\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \right|$$

$$\leq \left| \int f \, d\mu - \int g \, d\mu \right| + \left| \int g \, d\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i g \right) \right|$$

$$+ \left| \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i (g - f) \right) \right|$$

$$\leq \|f - g\|_\infty + \sum_{n=1}^{N} |z_n| \left| \int f_n \, d\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f_n \right) \right|$$

$$+ \frac{1}{k} \sum_{i=0}^{k-1} \|g - f\|_\infty$$

$$\leq \epsilon.$$

Thus $x \in S \Rightarrow \lim_{k \to \infty} \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) = \int f \, d\mu$ for all $f \in C(X)$. Since $\mu(S) = 1$, this concludes the proof. $\qquad \square$

Remark 2.2.2. We remark that the cylindrical structure of the $C_k(x)$ was not essential to our proof of Theorem 2.2.1. Rather, the important feature of $(C_k(x))_{k=1}^{\infty}$ was that their diameter went to 0 as $k \to \infty$. To demonstrate this fact, we consider the scenario where we replace the $C_k(x)$ with balls around x of radius decreasing to 0, and note that the technique of proof is remarkably similar to that used to prove Theorem 2.2.1.

Theorem 2.2.3. *Let $X = (X, \rho)$ be a compact metric space, and $T : X \to X$ a 1-Lipschitz map, i.e. such that $\rho(Tx, Ty) \leq \rho(x, y)$ for all $x, y \in X$. Let $\mathcal{B}$ denote the Borel σ-algebra on X, and μ a T-invariant, ergodic Borel probability measure on X. Let $(r_k)_{k=1}^{\infty}$ be a non-increasing sequence of positive numbers $r_k > 0$ such that $\lim_{k \to \infty} r_k = 0$. Let $B_k(x) = \{y \in X : \rho(x, y) < r_k\}$. Then the set of $x \in X$ such that*

$$\frac{1}{\mu(B_k(x))} \int_{B_k(x)} \frac{1}{k} \sum_{i=0}^{k-1} T^i f \, d\mu \xrightarrow{k \to \infty} \int f \, d\mu$$

for all $f \in C(X)$ is of full measure.

Proof. First we will prove that for an arbitrary $f \in C(X)$, the set of all $x \in X$ such that $\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)} \left(T^i f \right) \xrightarrow{k \to \infty} \int f \, d\mu$ is of full measure. Fix $\epsilon > 0$, and choose $\delta > 0$ such that $\rho(x, y) < \delta \Rightarrow$

$|f(x) - f(y)| < \frac{\epsilon}{2}$ (where we invoke the uniform continuity of f). Choose $K_1 \in \mathbb{N}$ such that $r_{K_1} < \delta$.
Then if $k \geq K_1, i \in [0, k-1]$, we have that $y \in B_k(x) \Rightarrow \rho(x,y) < \delta \Rightarrow \rho\left(T^i x, T^i y\right) < \delta$. Let $x \in \operatorname{supp}(\mu), k \geq K_1$. Then

$$
\begin{aligned}
\left|T^i f(x) - \alpha_{B_k(x)}\left(T^i f\right)\right| &= \left|T^i f(x) - \frac{1}{\mu(B_k(x))} \int_{B_k(x)} T^i f \mathrm{d}\mu\right| \\
&= \left|T^i f(x) - \frac{1}{\mu\left(T^i B_k(x)\right)} \int_{T^i B_k(x)} f \mathrm{d}\mu\right| \\
&= \left|\frac{1}{\mu\left(T^i B_k(x)\right)} \int_{T^i B_k(x)} \left(T^i f(x) - f\right) \mathrm{d}\mu\right| \\
&\leq \frac{1}{\mu\left(T^i B_k(x)\right)} \int_{T^i B_k(x)} \left|T^i f(x) - f\right| \mathrm{d}\mu \\
&\leq \frac{\epsilon}{2}.
\end{aligned}
$$

Let $x \in X \cap \operatorname{supp}(\mu)$ such that $\frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) \overset{k \to \infty}{\to} \int f \mathrm{d}\mu$. Choose $K_2 \in \mathbb{N}$ such that

$$
k \geq K_2 \Rightarrow \left|\int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x)\right| < \frac{\epsilon}{2}.
$$

Then if $k \geq \max\{K_1, K_2\}$, then

$$
\begin{aligned}
&\left|\int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)}\left(T^i f\right)\right| \\
&\leq \left|\int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x)\right| + \left|\frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)}\left(T^i\right)\right| \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}
$$

The Birkhoff Ergodic Theorem then tells us that the set $\left\{x \in X : \frac{1}{k} \sum_{i=0}^{k-1} T^i f(x) \overset{k \to \infty}{\to} \int f \mathrm{d}\mu\right\}$ is of full measure, and so we can intersect it with the support of μ to get another set of full measure.

We can now use an argument almost identical to that used in the proof of Theorem 2.2.1 to prove this present theorem. Let $\{f_n\}_{n \in \mathbb{N}}$ be a countable set in $C(X)$ such that $\overline{\operatorname{span}}\{f_n\}_{n \in \mathbb{N}} = C(X)$, where the

closure is taken in the uniform norm on $C(X)$. Let

$$S_n = \left\{ x \in X \cap \mathrm{supp}(\mu) : \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)}(f_n) \to \int f_n \mathrm{d}\mu \right\}.$$

As we have already shown, each S_n is of full measure, and thus so is $\bigcap_{n \in \mathbb{N}} S_n$. From here, appealing to the fact that these $\{f_n\}_{n \in \mathbb{N}}$ generate $C(X)$, we can prove the present theorem. $\qquad \square$

Remark 2.2.4. Assuming that $\mu(\{x\}) = 0$ for all $x \in X$, then $\mu(B_k(x)) \to 0$ for all $x \in X$.

Example 2.2.5. Theorem 2.2.3 ceases to be true if we drop the hypothesis that our system is ergodic. Let $X = (X, \rho)$ be a compact metric space, and let $T : X \to X$ be the identity map $T = \mathrm{id}_X$ on X. Let μ be any non-atomic Borel probability measure μ on X (which is automatically id_X-invariant) that is strictly positive. Fix $x_0 \in X$ and let $f(x) = \rho(x, x_0)$. Let $B_k(x_0) = \{x \in X : \rho(x, x_0) < 1/k\}$.

We claim that $\int f \mathrm{d}\mu > 0$, but

$$\alpha_{B_k(x_0)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \overset{k \to \infty}{\not\to} 0.$$

First, we observe that $\mu(B_k(x_0)) > 0$ for all $k \in \mathbb{N}$, since $x_0 \in \mathrm{supp}(\mu)$. However, since $\bigcap_{k=1}^{\infty} B_k(x_0) = \{x_0\}$, we know that $\mu(B_k(x_0)) \to 0$. Therefore there exists $K \in \mathbb{N}$ such that $0 < \mu(B_K(x_0)) < \mu(B_1(x_0))$. Since f is a nonnegative function, we can then conclude that

$$\int f \mathrm{d}\mu \geq \int_{B_1(x_0) \setminus B_K(x_0)} f \mathrm{d}\mu$$
$$\geq \mu(B_1(x_0) \setminus B_K(x_0)) \frac{1}{K}$$
$$> 0.$$

Then

$$\alpha_{B_k(x_0)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) = \alpha_{B_k(x_0)} \left(\frac{1}{k} \sum_{i=0}^{k-1} f \right)$$
$$= \alpha_{B_k(x_0)}(f)$$
$$= \frac{1}{\mu(B_k(x_0))} \int_{B_k(x_0)} f \mathrm{d}\mu.$$

26

However, we can also say that $\left|\alpha_{B_k(x_0)}(f)\right| \le 1/k$, since

$$\begin{aligned}
\left|\alpha_{B_k(x_0)}(f)\right| &= \left|\frac{1}{\mu(B_k(x_0))} \int_{B_k(x_0)} f \mathrm{d}\mu\right| \\
&\le \frac{1}{\mu(B_k(x_0))} \int_{B_k(x_0)} |f| \mathrm{d}\mu \\
&\le \frac{1}{\mu(B_k(x_0))} \int_{B_k(x_0)} \frac{1}{k} \mathrm{d}\mu \\
&= \frac{1}{k}.
\end{aligned}$$

Thus $\alpha_{B_k(x_0)}\left(\frac{1}{k}\sum_{i=0}^{k-1} T^i f\right) \overset{k\to\infty}{\to} 0 \neq \int f \mathrm{d}\mu$.

Thus for *every* $x_0 \in X$ exists $f_{x_0} \in C(X)$ such that

$$\limsup_{k\to\infty} \left|\int f_{x_0} \mathrm{d}\mu - \alpha_{B_k(x_0)}\left(\frac{1}{k}\sum_{i=0}^{k-1} T^i f_{x_0}\right)\right| > 0.$$

This example highlights how if the system under consideration is not ergodic, then results like Theorem 2.2.3 can fail.

Remark 2.2.6. In this section, as well as in Sections 2.3 and 2.4, we focus on continuous functions $f \in C(X)$. Our reason for this is that we can study these f in relation to the topological properties of (X, T).

2.3 Lipschitz maps and subshifts

Let us consider a compact *pseudo*metric space $X = (X, p)$, and $T : X \to X$ a map that is Lipschitz of constant $L > 1$, i.e. $p(Tx, Ty) \le L \cdot p(x, y)$. Recall that a pseudometric is distinguished from a metric by the fact we do not assume that a pseudometric distinguishes points, i.e. we do not assume that $p(x, y) = 0 \Rightarrow x = y$. Suppose that $(X, \mathcal{B}, \mu, T)$ is of finite entropy, and thus admits a generator $\mathcal{E}$. Suppose further that for almost all $x \in X$ exists a constant $\gamma = \gamma_x \in [1, \infty)$ such that

$$\mathrm{diam}(C_k(x)) \le \gamma \cdot L^{-k} \ (\forall k \in \mathbb{N}).$$

We pause to remark on two points. The first is that our consideration of pseudometric spaces is not generality for generality's sake. As we will see later in this section, this consideration of pseudometric spaces will be useful for studying certain metric spaces. The second is that this class of examples is not

a direct generalization of the class considered in Section 2.2. Though every 1-Lipschitz map is of course Lipschitz for every constant $L > 1$, our condition on $\mathrm{diam}(C_k(x))$ is stronger here, since we ask not just that $\mathrm{diam}(C_k(x))$ go to 0, but that it do so exponentially.

Since we are working in the slightly unorthodox setting of pseudometric spaces rather than metric spaces, we will prove that one of the strong properties of compact metric spaces is also true of compact *pseudo*metric spaces, namely that every continuous function is uniformly continuous. The proof is essentially identical to the "textbook" argument for compact metric spaces. We doubt this is a new result, but we could not find a reference for it, so we prove it here.

Lemma 2.3.1. *Let (X, p) be a compact pseudometric space. Then every continuous function $f : X \to \mathbb{C}$ is uniformly continuous.*

Proof. Fix $\epsilon > 0$. Then for every $x \in X$ exists $\delta_x > 0$ such that $p(x, y) < \delta_x \Rightarrow |f(x) - f(y)| < \epsilon$. Then the family $\mathcal{U} = \left\{ B\left(x, \frac{\delta_x}{2}\right) \right\}_{x \in X}$ is an open cover of X, so there exists a finite subcover $\mathcal{U}' = \left\{ B\left(x_j, \frac{\delta_{x_j}}{2}\right) \right\}_{j=1}^{n}$ of X.

Let $\delta' = \min_{1 \leq j \leq n} \frac{\delta_{x_j}}{2}$, and suppose that $x, y \in X$ such that $p(x, y) < \delta'$. Then there exists $x_j \in X$ such that $p(x, x_j) < \frac{\delta_{x_j}}{2}$, since $\mathcal{U}'$ is a cover of X. Then

$$
\begin{aligned}
p(x_j, y) &\leq p(x_j, x) + p(x, y) \\
&< \frac{\delta_{x_j}}{2} + \frac{\delta_{x_j}}{2} \\
&= \delta_{x_j}.
\end{aligned}
$$

Therefore $p(x_j, x) < \frac{\delta_{x_j}}{2} < \delta_{x_j}, p(x_j, y) < \delta_{x_j}$, so $|f(x) - f(x_j)| < \frac{\epsilon}{2}, |f(y) - f(x_j)| < \frac{\epsilon}{2}$. Thus

$$
\begin{aligned}
|f(x) - f(y)| &\leq |f(x) - f(x_j)| + |f(x_j) - f(y)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}
$$

Therefore $\delta' > 0$ is such that $p(x, y) < \delta' \Rightarrow |f(x) - f(y)| < \epsilon$. Thus we have shown that f is uniformly continuous. $\square$

Now we are able to both state and prove the first main result of this section.

Proposition 2.3.2. *Let (X, p) be a compact pseudometric space, and let $T : X \to X$ be an L-Lipschitz homeomorphism on X with respect to p, where $L > 1$. Suppose μ is a regular Borel probability measure on X such that T is ergodic with respect to μ. Let $\mathcal{E} = \{E_d\}_{d \in \mathcal{D}}$ be a generator of (X, T) such that for almost all $x \in X$ exists $\gamma_x \in \mathbb{R}$ such that $\mathrm{diam}(C_k(x)) \leq \gamma_x \cdot L^{-k}$ for all $k \in \mathbb{N}$. Fix $f \in C(X)$. Then*

$$\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \stackrel{k \to \infty}{\Rightarrow} \int f \mathrm{d}\mu$$

for almost all $x \in X$.

Proof. Our goal is to show that for every $\epsilon > 0$, there exists some $K \in \mathbb{N}$ such that if $k \geq K$, we have

$$\left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \right|$$

$$\leq \left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \left(T^i f \right)(x) \right| + \left| \frac{1}{k} \sum_{i=0}^{k-1} \left(\left(T^i f \right)(x) - \alpha_{C_k(x)} \left(T^i f \right) \right) \right|$$

$$\leq \left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \left(T^i f \right)(x) \right| + \frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f \right)(x) - \alpha_{C_k(x)} \left(T^i f \right) \right|$$

$$\leq \epsilon.$$

We will accomplish this by bounding the terms

$$\left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \left(T^i f \right)(x) \right|, \quad \frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f \right)(x) - \alpha_{C_k(x)} \left(T^i f \right) \right|$$

by ϵ.

We will start with bounding the latter term. We claim that if $x \in X$ such that $\mu(C_k(x)) > 0, \mathrm{diam}(C_k(x)) \leq \gamma_x \cdot L^{-k}$ for all $k \in \mathbb{N}$, then for every $\epsilon > 0$, there exists $K_1 \in \mathbb{N}$ such that

$$k \geq K_1 \Rightarrow \frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f \right)(x) - \alpha_{C_k(x)} \left(T^i f \right) \right| < \frac{\epsilon}{2}.$$

To prove this, choose $\delta > 0$ such that $p(y, z) < \delta \Rightarrow |f(y) - f(z)| < \frac{\epsilon}{4}$. Let $\kappa \in \mathbb{N}$ such that $\gamma_x \cdot L^{-\kappa} < \delta$. Then if $k > \kappa$, then

$$\frac{1}{k} \sum_{i=0}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| = \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right]$$
$$+ \frac{1}{k} \left[\sum_{k-\kappa+1}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right].$$

We will estimate these two terms separately, bounding each by $\frac{\epsilon}{4}$. Beginning with the former, we observe that if $x, y \in C_k(x)$, then

$$p(T^i x, T^i y) \leq L^i p(x, y) \leq L^i \cdot \gamma_x \cdot L^{-k} = \gamma_x \cdot L^{i-k}.$$

In particular, this means that if $i - k \leq -\kappa$, then $\left| (T^i f)(x) - f(z) \right| < \frac{\epsilon}{4}$ for all $z = T^i y \in T^i C_k(x)$, so

$$\frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right]$$
$$= \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| \frac{1}{\mu(C_k(x))} \int_{C_k(x)} ((T^i f)(x)) - T^i f \mathrm{d}\mu \right| \right]$$
$$= \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \frac{1}{\mu(T^i C_k(x))} \int_{T^i C_k(x)} \left| (T^i f)(x) - f \right| \mathrm{d}\mu \right]$$
$$\leq \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \frac{1}{\mu(T^i C_k(x))} \int_{T^i C_k(x)} \frac{\epsilon}{4} \mathrm{d}\mu \right]$$
$$= \frac{k - \kappa + 1}{k} \frac{\epsilon}{4}$$
$$\leq \frac{\epsilon}{4}.$$

On the other hand, we can estimate

$$\frac{1}{k} \left[\sum_{k-\kappa+1}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right] \leq \frac{2\kappa}{k} \|f\|.$$

Choose $K_1 > \kappa$ such that $\frac{2\kappa\|f\|_\infty}{K_1} < \frac{\epsilon}{4}$. Then if $k \geq K_1$, we have

$$\frac{1}{k} \sum_{i=0}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| = \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right]$$
$$+ \frac{1}{k} \left[\sum_{k-\kappa+1}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right]$$
$$\leq \frac{\epsilon}{4} + \frac{\epsilon}{4}$$
$$= \frac{\epsilon}{2}.$$

Now suppose further that $x \in X$ is such that $\frac{1}{k} \sum_{i=0}^{k-1} (T^i f)(x) \overset{k\to\infty}{\to} \int f \mathrm{d}\mu$. Choose $K_2 \in \mathbb{N}$ such that $k \geq K_2 \Rightarrow \left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} (T^i f)(x) \right| < \frac{\epsilon}{2}$. Then if $k \geq \max\{K_1, K_2\}$, then we have

$$\frac{1}{k} \sum_{i=0}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| = \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right]$$
$$+ \frac{1}{k} \left[\sum_{k-\kappa+1}^{k-1} \left| (T^i f)(x) - \alpha_{C_k(x)} (T^i f) \right| \right]$$
$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2}$$
$$= \epsilon.$$

Since the set of $x \in X$ for which this calculation could be performed is of full measure, the proposition follows. $\qquad\square$

From here, we get the following corollary.

Corollary 2.3.3. *Let (X, ρ) be a compact metric space, and let $T : X \to X$ be an L-Lipschitz homeomorphism on X with respect to ρ, where $L > 1$. Suppose μ is a regular Borel probability measure on X such that T is ergodic with respect to μ. Let $\mathcal{E} = \{E_d\}_{d\in\mathcal{D}}$ be a generator of (X, T) such that for almost all $x \in X$ exists $\gamma_x \in \mathbb{R}$ such that $\mathrm{diam}(C_k(x)) \leq \gamma_x \cdot L^{-k}$ for all $k \in \mathbb{N}$. Then the set of $x \in X$ such that*

$$\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} (T^i f) \overset{k\to\infty}{\to} \int f \mathrm{d}\mu$$

for all $f \in C(X)$ is of full measure.

Proof. Let $\{f_n\}_{n\in\mathbb{N}}$ be a countable set in $C(X)$ such that $C(X) = \overline{\text{span}}\{f_n\}_{n\in\mathbb{N}}$. By the previous result, we can extrapolate that the set of $x \in X$ such that $\frac{1}{k}\sum_{i=0}^{k-1}\alpha_{C_k(x)}\left(T^i f_n\right) \overset{k\to\infty}{\to} \int f_n \mathrm{d}\mu$ is of full measure. We can then extend to all of $C(X)$ in the same manner as we did in the proof of Theorem 2.2.1. $\qquad\square$

2.3.1 Two-sided subshifts and systems of finite entropy

This brings us to the matter of (two-sided) subshifts. Let $\mathcal{D}$ be a finite discrete set, and let $T : \mathcal{D}^{\mathbb{Z}} \to \mathcal{D}^{\mathbb{Z}}$ be the map $(Tx)_n = x_{n+1}$, called the *left shift*. We call $X \subseteq \mathcal{D}^{\mathbb{Z}}$ a *subshift* if X is compact and $TX = X$. Assume that μ is a Borel probability measure on X with respect to which T is ergodic.

In a shift space, we will always take our generator to be the family $\mathcal{E} = \{E_d\}_{d\in\mathcal{D}}$ of sets $E_d = \{x \in X : x_0 = d\}, d \in \mathcal{D}$. We claim that for almost all $x \in X$, we have

$$\lim_{k\to\infty}\frac{1}{k}\sum_{i=0}^{k-1}\alpha_{C_k(x)}\left(T^i f\right) = \int f \mathrm{d}\mu$$

for all $f \in C(X)$. First, we want to establish the following lemma. This is no doubt a classical result, but we could not find a reference for it, so we prove it here.

Lemma 2.3.4. *Let $(X, \mathcal{F}, \mu, T)$ be a subshift, where $X \subset \mathcal{D}^{\mathbb{Z}}$. The family*

$$\mathcal{F} = \left\{T^n \chi_{[a_0, a_1, \ldots, a_{\ell-1}]} : (a_0, a_1, \ldots, a_{\ell-1}) \in \mathcal{D}^\ell, \ell \in \mathbb{N}, i \in \mathbb{Z}\right\}$$

generates $C(X)$ in the sense that its span is dense in $C(X)$ with respect to the uniform norm.

Proof. We claim that every $f \in C(X)$ can be approximated uniformly by elements of $\text{span}\,\mathcal{F}$. We will begin by demonstrating the result for real $f \in C(X)$, then extrapolate the result to all complex-valued $f \in C(X)$.

For $\ell \in \mathbb{N}$, set

$$A(a_{-\ell+1}, a_{-\ell+1}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{\ell-1}, a_{\ell-1})$$
$$= \{x \in X : x_j = a_j \; \forall j \in [-\ell+1, \ell-1]\}$$
$$= T^{-\ell+1}[a_{-\ell+1}, a_{-\ell+2}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{\ell-1}, a_{\ell-1}]$$

and let

$$g_\ell = \sum_{\mathbf{a} \in \mathcal{D}^{2\ell-1}} \min\left\{f(y) : y \in A(\mathbf{a})\right\} \chi_{A(\mathbf{a})}$$

$$= \sum_{\mathbf{a} \in \mathcal{D}^{2\ell-1}} \min\left\{f(y) : y \in T^{-\ell+1}[\mathbf{a}]\right\} \chi_{T^{-\ell}[\mathbf{a}]}$$

$$\in \operatorname{span}\mathcal{F}.$$

We claim that $g_\ell \to f$ uniformly. The sequence $\ell \mapsto g_\ell$ is monotonic increasing. Moreover, we claim that it converges pointwise to f. To see this, let $x \in X$, and consider $g_\ell(x)$. Fix $\epsilon > 0$. Then for each $\ell \in \mathbb{N}$ exists $y^{(\ell)} \in X$ such that $g_\ell(x) = f\left(y^{(\ell)}\right)$. However, since $y_j^{(\ell)} = x_j$ for all $j \in [-\ell+1, \ell-1]$, we can conclude that $y^{(\ell)} \to x$, and so by continuity of f, we can conclude that $g_\ell(x) = f\left(y^{(\ell)}\right) \to f(x)$. Thus $g_\ell \nearrow f$ pointwise. Dini's Theorem then gives us uniform convergence. Therefore, if $f \in C(X)$ is real-valued, then $f \in \overline{\operatorname{span}}\mathcal{F}$. On the other hand, any complex-valued function $f \in C(X)$ can be expressed as the sum of its real and imaginary parts, and we can apply this argument to both of those parts separately. $\qquad\square$

Theorem 2.3.5. *Let $X \subseteq \mathcal{D}^{\mathbb{Z}}$ be a subshift, and let μ be a Borel probability measure on X with respect to which the left shift T is ergodic. Then the set of all $x \in X$ such that*

$$\frac{1}{k}\sum_{i=0}^{k-1} \alpha_{C_k(x)}\left(T^i f\right) \to \int f \mathrm{d}\mu$$

for all $f \in C(X)$ is of full measure.

Proof. Our first step is to show that

$$\frac{1}{k}\sum_{i=0}^{k-1} \alpha_{C_k(x)}\left(T^i \chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right) \to \int \chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\mathrm{d}\mu$$

for all finite strings $\mathbf{a} = (a_0, a_1, \ldots, a_{\ell-1}) \in \mathcal{D}^\ell$. Let p be the pseudometric on X given by $p(x,y) = 2^{-\min\{n \geq 0 : x_n \neq y_n\}}$, where $\min(\emptyset) = +\infty$ and $2^{-\infty} = 0$.

We claim that the function $\chi_{[\mathbf{a}]}$ is continuous with respect to the topology of p, and that $(X, \mathcal{B}, \mu, T)$ satisfies the hypotheses of Proposition 2.3.2 for $L = 2$. A straightforward calculation shows that T is

2-Lipschitz and that $\text{diam}\,([\mathbf{a}]) \leq 2 \cdot 2^{-\ell}$ for all $\ell \in \mathbb{N}, \mathbf{a} \in \mathcal{D}^\ell$. Therefore, if

$$R_\mathbf{a} = \left\{ x \in X : \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i \chi_{[a_0,a_1,\ldots,a_{\ell-1}]} \right) \to \int \chi_{[a_0,a_1,\ldots,a_{\ell-1}]} d\mu \right\},$$

then $\mu(R_\mathbf{a}) = 1$ for all $\mathbf{a} \in \bigcup_{\ell=1}^\infty \mathcal{D}^\ell$, and so $R = \bigcap_{\mathbf{a} \in \bigcup_{\ell=1}^\infty \mathcal{D}^\ell} R_\mathbf{a}$ is of full measure. We now claim that if $\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i \chi_{[a_0,a_1,\ldots,a_{\ell-1}]} \right) \to \int \chi_{[a_0,a_1,\ldots,a_{\ell-1}]} d\mu$, then

$$\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i T^n \chi_{[a_0,a_1,\ldots,a_{\ell-1}]} \right) \to \int \chi_{[a_0,a_1,\ldots,a_{\ell-1}]} d\mu$$

for all $n \in \mathbb{Z}$.

It will suffice to prove the result for $n = \pm 1$ and extend to all $n \in \mathbb{Z}$ by induction. To prove the claim for $n = 1$, we observe that

$$\left(\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i(Tf) \right) \right) - \left(\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \right) = \frac{1}{k} \alpha_{C_k(x)} \left(T^k f - f \right)$$

$$\Rightarrow \left| \left(\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i(Tf) \right) \right) - \left(\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \right) \right| \leq \frac{2}{k} \|f\|_\infty$$

$$\to 0.$$

A similar calculation tells us that

$$\left| \left(\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i \left(T^{-1} f \right) \right) \right) - \left(\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \right) \right| \leq \frac{2}{k} \|f\|_\infty \to 0,$$

verifying the claim for $n = -1$. Thus if $\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i f \right) \to \int f d\mu$, then a straightforward induction argument will show that

$$\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{C_k(x)} \left(T^i \left(T^n f \right) \right) \to \int f d\mu = \int T^n f d\mu$$

for all $n \in \mathbb{Z}$.

In particular, this means that if $x \in R$, then $\frac{1}{k}\sum_{i=0}^{k-1}\alpha_{C_k(x)}\left(T^i f\right) \to \int f \mathrm{d}\mu$ for all $f \in \mathcal{F}$. Since the span of $\mathcal{F}$ is dense in $C(X)$, this means that if $x \in R$, then

$$\frac{1}{k}\sum_{i=0}^{k-1}\alpha_{C_k(x)}\left(T^i f\right) \to \int f \mathrm{d}\mu$$

for all $f \in C(X)$. $\qquad\square$

We turn now to apply Theorem 2.3.5 to a slightly broader context. Let $(Y, \mathcal{A}, \nu, S)$ be an invertible ergodic system with finite entropy. Then the system admits a finite generator $\mathcal{E} = \{E_d\}_{d \in \mathcal{D}}$. For each $y \in Y, i \in \mathbb{Z}$, let $e_i(y) \in \mathcal{D}$ be the element of $\mathcal{D}$ such that $y \in S^{-i}E_{e_i(y)}$, or equivalently such that $S^i y \in E_{e_i(y)}$. Define the k-*length cylinder* corresponding to y by

$$F_k(y) = \bigcap_{i=0}^{k-1} S^{-i}E_{e_i(y)}.$$

We denote these cylinders by F_k instead of C_k to indicate that they live in Y, not X.

We define a map $\phi : Y \to \mathcal{D}^{\mathbb{Z}}$ by

$$\phi(y) = (e_i(y))_{i \in \mathbb{Z}}.$$

We call this map ϕ the *itinerary map* on Y induced by $\mathcal{E}$. Let T be the standard left shift on $\mathcal{D}^{\mathbb{Z}}$. The itinerary map commutes with the left shift in the sense that the following diagram commutes:

$$
\begin{array}{ccc}
Y & \xrightarrow{\;\;S\;\;} & Y \\
\downarrow{\scriptstyle\phi} & & \downarrow{\scriptstyle\phi} \\
\mathcal{D}^{\mathbb{Z}} & \xrightarrow{\;\;T\;\;} & \mathcal{D}^{\mathbb{Z}}
\end{array}
$$

We can now state the following corollary.

Corollary 2.3.6. *Let $(Y, \mathcal{A}, \nu, S)$ be an invertible ergodic system with finite entropy and finite generator $\mathcal{E} = \{E_d\}_{d \in \mathcal{D}}$. Let $\mathbb{A} \subseteq L^{\infty}(Y, \nu)$ be the subspace*

$$\mathbb{A} = \overline{\mathrm{span}}\left\{S^n \chi_{\bigcap_{j=0}^{\ell-1} S^{-j}E_{d_j}} : n \in \mathbb{Z}, \ell \in \mathbb{N}, d_j \in \mathcal{D}\right\}.$$

35

Then the set of $y \in Y$ such that

$$\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(F_k(y))}\int_{F_k(y)}S^i g \mathrm{d}\nu \to \int g \mathrm{d}\nu$$

for all $g \in \mathbb{A}$ is of full measure.

Proof. Endow $\mathcal{D}^{\mathbb{Z}}$ with the pushforward measure $\mu(B) = \nu\left(\phi^{-1}B\right)$. Since $\phi^{-1}[d] = E_d \in \mathcal{A}$ for all $d \in \mathcal{D}$, we know that μ is Borel. We also observe that $F_k(y) = \phi^{-1}C_k(\phi(y))$. Consider $f = \chi_{\bigcap_{i=0}^{k-1} E_{d_i}} = \chi_{\phi^{-1}[d_0,d_1,\dots,d_{k-1}]}$. Let $B \subseteq \mathcal{D}^{\mathbb{Z}}$ be the set of all $x \in X$ such that

$$\frac{1}{k}\sum_{i=0}^{k-1}\alpha_{C_k(x)}\left(T^i f\right) \to \int f \mathrm{d}\mu$$

for all $f \in C(X)$, which we know by the previous theorem to be of full measure in X, and let $A = \phi^{-1}B$. Then if $y \in A$, and $d_0, d_1, \dots, d_{\ell-1} \in \mathcal{D}$, then

$$\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\nu(F_k(y))}\int_{F_k(y)}S^i\chi_{\bigcap_{j=0}^{\ell-1}S^{-j}E_{d_j}}\mathrm{d}\nu$$

$$= \frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\nu(F_k(y))}\nu\left(F_k(y)\cap S^{-i}\bigcap_{j=0}^{\ell-1}S^{-j}E_{d_j}\right)$$

$$= \frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\nu\left(\phi^{-1}C_k(\phi(y))\right)}\nu\left(\phi^{-1}\left(C_k(\phi(y))\cap T^{-i}[d_0,d_1,\dots,d_{\ell-1}]\right)\right)$$

$$= \frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(C_k(\phi(y)))}\mu\left(C_k(\phi(y))\cap T^{-i}[d_0,d_1,\dots,d_{\ell-1}]\right)$$

$$= \frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(C_k(\phi(y)))}\int_{C_k(\phi(y)}T^i\chi_{[d_0,d_1,\dots,d_{\ell-1}]}\mathrm{d}\mu$$

$$\to \int\chi_{[d_0,d_1,\dots,d_{\ell-1}]}\mathrm{d}\mu$$

$$= \mu([d_0,d_1,\dots,d_{\ell-1}])$$

$$= \nu\left(\bigcap_{j=0}^{\ell-1}S^{-j}E_{d_j}\right)$$

$$= \int\chi_{\bigcap_{j=0}^{\ell-1}S^{-j}E_{d_j}}\mathrm{d}\nu,$$

since $\chi_{[d_0,d_1,\ldots,d_{\ell-1}]} \in C\left(\mathcal{D}^{\mathbb{Z}}\right)$. By an argument similar to that employed in the proof of Theorem 2.3.5, we can extrapolate that if $y \in A$, then

$$\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(F_k(y))}\int_{F_k(y)}S^iS^n\chi_{\bigcap_{j=0}^{\ell-1}S^{-j}E_{d_j}}\,\mathrm{d}\nu \to \int S^n\chi_{\bigcap_{j=0}^{\ell-1}S^{-j}E_{d_j}}\,\mathrm{d}\nu$$

for $n \in \mathbb{Z}$. By density, it follows that if $g \in \mathbb{A}$, then $\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(F_k(y))}\int_{F_k(y)}S^ig\mathrm{d}\nu \to \int g\mathrm{d}\nu$ for all $y \in A$, and A is a set of full measure. □

2.3.2 Pathological differentiation problems and relations to symbolic distributions

In Theorem 2.3.5, we demonstrated that

$$\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^if\right) \overset{k\to\infty}{\to} \int f\mathrm{d}\mu \; (\forall f \in C(X))$$

for almost all $x \in X$. We take this opportunity to demonstrate that the "almost all" caveat is indispensable, as there can exist $x \in X$ for which $\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^if\right) \not\to \int f\mathrm{d}\mu$ for certain $f \in C(X)$. This is related to the shift not being uniquely ergodic, which we discussed in more detail in Section 2.1. In fact, we even claim the sequence $\left(\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^if\right)\right)_{k=1}^{\infty}$ can fail to be *Cauchy* for certain pairs $(x,f) \in X \times C(X)$.

Theorem 2.3.7. *Let $X = \mathcal{D}^{\mathbb{Z}}$ be a Bernoulli shift with symbol space $\mathcal{D} = \{0,1,\ldots,D-1\}, D \geq 2$, a Borel probability measure μ such that $\mu([d]) \neq 0$ for all $d \in \mathcal{D}$. Let $f = \chi_{[0]}$, and left shift T. Then there exists an uncountable subset $S \subseteq X$ such that $x, y \in S \Rightarrow x_j = y_j \; (\forall j \leq 0)$, and such that the sequence $\left(\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^if\right)\right)_{k=1}^{\infty}$ is not Cauchy for all $x \in S$.*

Proof. We first compute $\alpha_{[x_0,x_1,\ldots,x_{k-1}]}\left(T^if\right)$ for $0 \leq i \leq k-1$ as follows. We see

$$\begin{aligned}
\alpha_{[x_0,x_1,\ldots,x_{k-1}]}\left(T^if\right) &= \frac{1}{\mu([x_0,x_1,\ldots,x_{k-1}])}\int_{[x_0,x_1,\ldots,x_{k-1}]}T^i\chi_{[0]}\mathrm{d}\mu \\
&= \frac{1}{\mu([x_0,x_1,\ldots,x_{k-1}])}\int_{\bigcap_{j=0}^{k-1}T^{-j}[x_j]}\chi_{T^{-i}[0]}\mathrm{d}\mu \\
&= \frac{1}{\mu([x_0,x_1,\ldots,x_{k-1}])}\mu\left(\left(\bigcap_{j=0}^{k-1}T^{-j}[x_j]\right)\cap T^{-i}[0]\right) \\
&= \delta(x_i,0),
\end{aligned}$$

37

where $\delta(\cdot, \cdot)$ refers here to the Kronecker delta. Thus if $x = (x_j)_{j \in \mathbb{Z}} \in X$, then

$$\alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) = \frac{\#\{i \in [0, k-1] : x_i = 0\}}{k} \tag{$\dagger$}$$

The identity ($\dagger$) implies that if there exists $x \in X$ such that $\left(\alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right)_{k=1}^{\infty}$ is not Cauchy, then we can then build our set S. For $x, y \in X$, write $x \sim y$ if $x_j = y_j$ for all $j \leq 0$, and the set $\{j \in \mathbb{N} : x_j \neq y_j\}$ has density 0. This is an equivalence relation. We claim that if $x \sim y$, then $\left| \alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) - \alpha_{C_k(y)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right| \stackrel{k \to \infty}{\to} 0$. By ($\dagger$), we know that

$$\begin{aligned}
\left| \alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) - \alpha_{C_k(y)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right| &= \left| \frac{1}{k} \sum_{i=0}^{k-1} (\delta(x_i, 0) - \delta(y_i, 0)) \right| \\
&\leq \frac{1}{k} \sum_{i=0}^{k-1} |(\delta(x_i, 0) - \delta(y_i, 0))| \\
&\leq \frac{\#\{i \in [0, k-1] : x_i \neq y_i\}}{k} \\
&\stackrel{k \to \infty}{\to} 0.
\end{aligned}$$

Therefore, we can let S be the equivalence class of x under $\sim$. To see that this S is uncountable, let $E \subseteq \mathbb{N}$ be an infinite subset of density 0. Then every subset F of E has density 0. For each $F \subseteq E$, let $x^F \in X$ be a sequence such that $x_j^F = x_j$ for $j \notin F$ and $x_j^F \neq x_j$ for $j \in F$. Since E has uncountably many subsets, and $x \sim x^F$ for all $F \subseteq E$, we have shown that the equivalence class of x by $\sim$ is uncountable. So, assuming that $x \in X$ such that $\left(\alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} f \right) \right)_{k=1}^{\infty}$ is not Cauchy, then we can let $S = \{y \in X : x \sim y\}$.

Our next order of business is to construct some such x. The identity ($\dagger$) also helps us construct an $x \in X$ for which $\left(\alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \right)_{k=1}^{\infty}$ is not Cauchy. Construct $x = (x_j)_{j \in \mathbb{Z}} \in X$ as follows. For brevity,

let $c_n = \sum_{p=1}^n 2^p$. Set

$$
x_j = \begin{cases}
0 & j < 0 \\[4pt]
1 & j = 0 \\[4pt]
0 & 0 < j \le 2 \\[4pt]
1 & 2 < j \le 6 \\[4pt]
0 & 6 < j \le 14 \\[4pt]
1 & 14 < j \le 30 \\[4pt]
\vdots & \\[4pt]
0 & c_{2n} < j \le c_{2n+1} \\[4pt]
1 & c_{2n+1} < j \le c_{2n+2} \\[4pt]
0 & c_{2n+2} < j \le c_{2n+3} \\[4pt]
\vdots &
\end{cases}
$$

In plain language, this sequence begins with 0 for $j < 0$, a 1 at $j = 0$, then 2^1 terms of 0, then 2^2 terms of 1, then 2^3 terms of 0, then 2^4 terms of 1, and so on. We claim that $\liminf_{k \to \infty} \alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \ne \limsup \alpha_{C_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right)$. Sampling along the subsequence $k_n = c_{2n} + 1$, we get

$$
\begin{aligned}
\alpha_{C_{c_{2n}+1}(x)} \left(\frac{1}{c_{2n} + 1} \sum_{i=0}^{c_{2n}} T^i f \right) &= \frac{2 + 8 + 32 + \cdots + 2^{2n-1}}{1 + 2 + 4 + 6 + \cdots + 2^{2n}} & = \frac{\frac{1}{2} \sum_{p=1}^n 4^p}{1 + \sum_{q=1}^{2n} 2^q} \\
&= \frac{1}{3} \cdot \frac{4^n - 1}{4^n - \frac{1}{2}} & \overset{n \to \infty}{\to} \frac{1}{3},
\end{aligned}
$$

where the limit is taken using L'Hospital's Rule. On the other hand, looking at the subsequence $k_n = c_{2n-1}+1$, we get

$$
\begin{aligned}
\alpha_{C_{c_{2n-1}+1}(x)} \left(\frac{1}{c_{2n-1} + 1} \sum_{i=0}^{c_{2n-1}} T^i f \right) &= \frac{2 + 8 + 32 + \cdots + 2^{2n-1}}{1 + 2 + 4 + 6 + \cdots + 2^{2n-1}} & = \frac{\frac{1}{2} \sum_{p=1}^n 4^p}{1 + \sum_{q=1}^{2n-1} 2^q} \\
&= \frac{1}{3} \cdot \frac{4^n - 1}{\frac{1}{2} 4^n - \frac{1}{2}} & \overset{n \to \infty}{\to} \frac{2}{3}.
\end{aligned}
$$

Thus we can say

$$\liminf_{k\to\infty} \alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i f\right) \leq \frac{1}{3} < \frac{2}{3} \leq \limsup_{k\to\infty} \alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i f\right).$$

Therefore the sequence $\left(\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i f\right)\right)_{k=1}^{\infty}$ is divergent, and thus not Cauchy. □

Remark 2.3.8. Theorem 2.3.7 is not encompassed by Theorem 2.1.10, since a subshift is a priori totally disconnected.

This calculation adequately sets up the following result.

Theorem 2.3.9. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic subshift, with $X \subseteq \mathcal{D}^{\mathbb{Z}}$, and let $x \in X$. Then the following statements about $x \in X$ are equivalent.*

1. *For all $f \in C(X)$, the limit*

$$\lim_{k\to\infty} \alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i f\right)$$

   exists and is equal to $\int f\mathrm{d}\mu$.

2. *For all words $(a_0, a_1, \ldots, a_{\ell-1}) \in \bigcup_{\ell=1}^{\infty}\mathcal{D}^{\ell}$, the limit*

$$\lim_{k\to\infty} \alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i \chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right)$$

   exists and is equal to $\mu([a_0, a_1, \ldots, a_{\ell-1}])$.

3. *For all words $(a_0, a_1, \ldots, a_{\ell-1}) \in \bigcup_{\ell=1}^{\infty}\mathcal{D}^{\ell}$, the limit*

$$\lim_{k\to\infty} \frac{\#\{i \in [0, k-\ell] : x_i = a_0, x_{i+1} = a_1, \ldots, x_{i+\ell-1} = a_{\ell-1}\}}{k}$$

   exists and is equal to $\mu([a_0, a_1, \ldots, a_{\ell-1}])$.

4. *For all words $(a_0, a_1, \ldots, a_{\ell-1}) \in \bigcup_{\ell=1}^{\infty}\mathcal{D}^{\ell}$, the limit*

$$\lim_{k\to\infty} \frac{\#\{i \in [0, k-1] : x_i = a_0, x_{i+1} = a_1, \ldots, x_{i+\ell-1} = a_{\ell-1}\}}{k}$$

   exists and is equal to $\mu([a_0, a_1, \ldots, a_{\ell-1}])$.

Proof. Lemma 2.3.4 tells us that (1) $\iff$ (2). That (3) $\iff$ (4) comes from the observation that the absolute difference between the two sequences is at most $\frac{\ell-1}{k}$. To establish (2) $\iff$ (3), we compute $\alpha_{C_k(x)}\left(T^i\chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right)$ for $i \in [0, k-\ell]$ as follows.

$$\alpha_{C_k(x)}\left(T^i\chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right) = \frac{1}{\mu([x_0,x_1,\ldots,x_{k-1}])}\int_{[x_0,x_1,\ldots,x_{k-1}]}\chi_{T^{-i}[a_0,a_1,\ldots,a_{\ell-1}]}\mathrm{d}\mu$$

$$= \begin{cases} 1 & a_i = x_0, a_{i+1} = x_1,\ldots,a_{i+\ell-1} = x_{\ell-1}, \\ 0 & \text{otherwise} \end{cases}.$$

Therefore

$$\frac{1}{k}\sum_{i=0}^{k-\ell}\alpha_{C_k(x)}\left(T^i\chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right)$$

$$= \frac{\#\{i \in [0, k-\ell] : x_i = a_0, x_{i+1} = a_1,\ldots,x_{i+\ell-1} = a_{\ell-1}\}}{k}.$$

Finally, we observe that

$$\left|\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^if\right) - \alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-\ell}T^if\right)\right| = \left|\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=k-\ell+1}^{k-1}T^if\right)\right|$$

$$\leq \frac{\ell-1}{k}\|f\|_\infty.$$

Therefore the end behaviors of $\left(\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i\chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right)\right)_{k=1}^\infty$ and $\left(\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-\ell}T^i\chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right)\right)_{k=1}^\infty$ are identical, i.e. one converges iff the other converges, and if they converge, then they converge to the same value. But then, as has already been established, we know that

$$\left(\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-\ell}T^i\chi_{[a_0,a_1,\ldots,a_{\ell-1}]}\right)\right)_{k=1}^\infty$$

$$= \left(\frac{\#\{i \in [0, k-1] : x_i = a_0, x_{i+1} = a_1,\ldots,x_{i+\ell-1} = a_{\ell-1}\}}{k}\right)_{k=1}^\infty,$$

demonstrating that (2) $\iff$ (3). $\square$

Theorem 2.3.9 gives us an alternate proof of Theorem 2.3.5. Applying the Birkhoff Ergodic Theorem to the functions $\chi_{[\mathbf{a}]}$ tells us that almost all $x \in X$ satisfy $\frac{1}{k} \sum_{i=0}^{k-1} T^i \chi_{[\mathbf{a}]}(x) \overset{k\to\infty}{\to} \mu([\mathbf{a}])$ for all strings $\mathbf{a} \in \bigcup_{\ell=1}^{\infty} \mathcal{D}^{\ell}$. But this is exactly condition (4) from Theorem 2.3.9. Moreover, this result gives us a more concrete characterization of the "set of full measure" that Theorem 2.3.5 alludes to.

Before concluding, we demonstrate that Proposition 2.3.2 does not hinge on the cylinder structure of X.

Theorem 2.3.10. *Let (X, ρ) be a compact metric space, and let $T : X \to X$ be an L-Lipschitz homeomorphism on X with respect to ρ, where $L > 1$. Suppose μ is a regular Borel probability measure on X such that T is ergodic with respect to μ. Let $(r_k)_{k=1}^{\infty}$ be a sequence of positive numbers $r_k > 0$ such that there exists a constant $\gamma \in \mathbb{R}$ such that $r_k \leq \gamma \cdot L^{-k}$ for all $k \in \mathbb{N}$. Fix $f \in C(X)$. Let $B_k(x) = \{y \in X : \rho(x, y) < r_k\}$. Then the set of $x \in X$ such that*

$$\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)} \left(T^i f\right) \overset{k\to\infty}{\to} \int f \mathrm{d}\mu$$

for all $f \in C(X)$ is of full measure.

Proof. Since $C(X)$ is separable, it will suffice to show that given some fixed $f \in C(X)$, we have

$$\frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)} \left(T^i f\right) \overset{k\to\infty}{\to} \int f \mathrm{d}\mu$$

for almost all $x \in X$. Our method of proof will closely resemble our proof of Proposition 2.3.2.

Our goal is to show that for every $\epsilon > 0$ exists some $K \in \mathbb{N}$ such that if $k \geq K(\epsilon)$, we have

$$\left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \alpha_{B_k(x)} \left(T^i f\right) \right|$$

$$\leq \left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \left(T^i f\right)(x) \right| + \left| \frac{1}{k} \sum_{i=0}^{k-1} \left(\left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right)\right) \right|$$

$$\leq \left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \left(T^i f\right)(x) \right| + \frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right|$$

$$\leq \epsilon.$$

We will accomplish this by bounding the terms

$$\left| \int f \mathrm{d}\mu - \frac{1}{k} \sum_{i=0}^{k-1} \left(T^i f\right)(x) \right|, \quad \frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right|$$

by ϵ.

We will start with bounding the latter term. We claim that if $x \in X$ such that $\mu(B_k(x)) > 0, \mathrm{diam}(B_k(x)) \leq \gamma_x \cdot L^{-k}$ for all $k \in \mathbb{N}$, then for every $\epsilon > 0$, there exists $K_1 \in \mathbb{N}$ such that

$$k \geq K_1 \Rightarrow \frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right| < \frac{\epsilon}{2}.$$

To prove this, choose $\delta > 0$ such that $p(y, z) < \delta \Rightarrow |f(y) - f(z)| < \frac{\epsilon}{4}$. Let $\kappa \in \mathbb{N}$ such that $\gamma_x \cdot L^{-\kappa} < \delta$. Then if $k > \kappa$, then

$$\frac{1}{k} \sum_{i=0}^{k-1} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right| = \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right| \right]$$
$$+ \frac{1}{k} \left[\sum_{k-\kappa+1}^{k-1} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right| \right].$$

We will estimate these two terms separately, bounding each by $\frac{\epsilon}{4}$. Beginning with the former, we observe that if $x, y \in B_k(x)$, then

$$p\left(T^i x, T^i y\right) \leq L^i p(x, y) \leq L^i \cdot \gamma_x \cdot L^{-k} = \gamma_x \cdot L^{i-k}.$$

In particular, this means that if $i - k \leq -\kappa$, then $\left| \left(T^i f\right)(x) - f(z) \right| < \frac{\epsilon}{4}$ for all $z = T^i y \in T^i C_k(x)$, so

$$\frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| \left(T^i f\right)(x) - \alpha_{B_k(x)} \left(T^i f\right) \right| \right]$$
$$= \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \left| \frac{1}{\mu(B_k(x))} \int_{B_k(x)} \left(\left(T^i f\right)(x) \right) - T^i f \mathrm{d}\mu \right| \right]$$
$$= \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \frac{1}{\mu\left(T^i B_k(x)\right)} \int_{T^i B_k(x)} \left| \left(T^i f\right)(x) - f \right| \mathrm{d}\mu \right]$$
$$\leq \frac{1}{k} \left[\sum_{i=0}^{k-\kappa} \frac{1}{\mu\left(T^i B_k(x)\right)} \int_{T^i B_k(x)} \frac{\epsilon}{4} \mathrm{d}\mu \right]$$
$$= \frac{k-\kappa+1}{k} \frac{\epsilon}{4}$$
$$\leq \frac{\epsilon}{4}.$$

On the other hand, we can estimate

$$\frac{1}{k}\left[\sum_{k-\kappa+1}^{k-1}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|\right]\leq\frac{2\kappa}{k}\|f\|.$$

Choose $K_1>\kappa$ such that $\frac{2\kappa\|f\|_\infty}{K_1}<\frac{\epsilon}{4}$. Then if $k\geq K_1$, we have

$$\frac{1}{k}\sum_{i=0}^{k-1}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|=\frac{1}{k}\left[\sum_{i=0}^{k-\kappa}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|\right]$$
$$+\frac{1}{k}\left[\sum_{k-\kappa+1}^{k-1}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|\right]$$
$$\leq\frac{\epsilon}{4}+\frac{\epsilon}{4}$$
$$=\frac{\epsilon}{2}.$$

Now suppose further that $x\in X$ is such that $\frac{1}{k}\sum_{i=0}^{k-1}\left(T^i f\right)(x)\overset{k\to\infty}{\to}\int f\mathrm{d}\mu$. Choose $K_2\in\mathbb{N}$ such that $k\geq K_2\Rightarrow\left|\int f\mathrm{d}\mu-\frac{1}{k}\sum_{i=0}^{k-1}\left(T^i f\right)(x)\right|<\frac{\epsilon}{2}$. Then if $k\geq\max\{K_1,K_2\}$, then we have

$$\frac{1}{k}\sum_{i=0}^{k-1}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|=\frac{1}{k}\left[\sum_{i=0}^{k-\kappa}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|\right]$$
$$+\frac{1}{k}\left[\sum_{k-\kappa+1}^{k-1}\left|\left(T^i f\right)(x)-\alpha_{B_k(x)}\left(T^i f\right)\right|\right]$$
$$\leq\frac{\epsilon}{2}+\frac{\epsilon}{2}$$
$$=\epsilon.$$

$\square$

Before looking at a more general family of differentiation problems, we want to take a moment to observe that if $(X,\mathcal{B},T,\mu)$ is an ergodic system, then if the (measure-theoretic) entropy $h(T,\mu)$ of the system is positive, then we automatically have that $\mu(C_k(x))\overset{k\to\infty}{\to}0$: by the Shannon-McMillan-Breiman Theorem (Dajani and Kraaikamp, 2002, Theorem 6.2.1), it follows that for μ-almost every $x\in X$ there exists $K=K_x\in\mathbb{N}$ such that

$$k\geq K\Rightarrow-\frac{1}{k}\log\mu(C_k(x))\geq\frac{h(T,\mu)}{2}.$$

Then if $k \geq K$, we have

$$-\frac{1}{k} \log \mu(C_k(x)) \geq \frac{h(T, \mu)}{2}$$
$$\Rightarrow \log \mu(C_k(x)) \leq -\frac{h(T, \mu)}{2} k \qquad\qquad < 0$$
$$\Rightarrow \mu(C_k(x)) \leq \left(e^{-\frac{h(T, \mu)}{2}} \right)^k \qquad\qquad \overset{k \to \infty}{\to} 0.$$

On the other hand, whether $\mu(B_k(x)) \overset{k \to \infty}{\to} 0$ depends on where $(X, \mathcal{B}, \mu)$ contains atoms. If $\mu(\{x\}) = 0$ for all $x \in X$, then $\mu(B_k(x)) \overset{k \to \infty}{\to} 0$.

2.4   Random cylinders in a Bernoulli shift - a probabilistic approach

In this section, we consider problems similar to those addressed in Sections 2.2 and 2.3, where we take some $(X, \mathcal{B}, \mu, T)$ with specified properties (in this case, we assume the system is Bernoulli), and seek to establish conditions under which for a randomly chosen sequence $(F_k)_{k=1}^\infty$ of sets of positive measure, the sequence $\left(\frac{1}{\mu(F_k)} \int_{F_k} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \mathrm{d}\mu \right)_{k=1}^\infty$ converges almost surely to $\int f \mathrm{d}\mu$ for all $f \in C(X)$.

We provide now an alternate proof of a special case of Theorem 2.3.5. Though the result proved is lesser in scope, we include it for the reason that the proof provided here has a decidedly more probabilistic flavor than the proof provided of Theorem 2.3.5 in Section 2.3. This method of proof also proves slightly more versatile, as it allows us to consider randomly chosen sequences of cylinders which are not necessarily nested.

In this section, $X = \mathcal{D}^{\mathbb{Z}}$ is a Bernoulli shift on a finite alphabet $\mathcal{D}$ with probability vector $\mathbf{p} = (p(d))_{d \in \mathcal{D}}$, and μ is the Borel probability measure on X induced by $\mathbf{p}$. We begin by proving a lemma to which we assign a whimsical title.

Lemma 2.4.1 (The Even Stronger Law of Large Numbers). *Let $(Y, \mathcal{A}, \nu)$ be a probability space, and let $(k_n)_{n=1}^\infty$ be a sequence in $\mathbb{N}$ such that $\sum_{n=1}^\infty k_n^{-2} < \infty$. Let $(\zeta_{i,n})_{0 \leq i \leq k_n - 1, n \in \mathbb{N}}$ be a family of L^∞ real random variables satisfying the following conditions.*

1. *There exists $C \in [1, \infty)$ such that $\|\zeta_{i,n}\|_\infty \leq C$ for all $0 \leq i \leq k_n - 1, n \in \mathbb{N}$.*

2. *$\int \zeta_{i,n} \mathrm{d}\nu = m$ for all $0 \leq i \leq k_n - 1, n \in \mathbb{N}$, where m is a constant.*

3. *For each $n \in \mathbb{N}$, the subfamily $\{\zeta_{i,n}\}_{i=0}^{k_n - 1}$ is mutually independent.*

Then

$$\frac{1}{k_n} \sum_{i=0}^{k_n-1} \zeta_{i,n} \overset{n\to\infty}{\to} m$$

almost surely.

Proof. For the sake of brevity, abbreviate

$$S_n = \sum_{i=0}^{k_n-1} \zeta_{i,n},$$

and assume without loss of generality that $m = 0$ (else, we can just consider $\hat{\zeta}_{i,n} = \zeta_{i,n} - m$). Given $\epsilon > 0$,

set

$$E_{n,\epsilon} = \{y \in Y : |S_n(y)|/k_n \geq \epsilon\} = \{y \in Y : |S_n(y)| \geq k_n\epsilon\}.$$

Then Chebyshev's inequality tells us that

$$\mu(E_{n,\epsilon}) \leq \frac{1}{(k_n\epsilon)^4} \int S_n^4 d\nu.$$

Then

$$\int S_n^4 d\nu = \sum_{r,s,t,u=0}^{k_n-1} \int \zeta_{r,n}\zeta_{s,n}\zeta_{t,n}\zeta_{u,n} d\nu.$$

This sum consists of terms of the forms

1. $\int \zeta_{r,n}^4 d\nu$

2. $\int \zeta_{r,n}^2 \zeta_{s,n}^2 d\nu$

3. $\int \zeta_{r,n}^3 \zeta_{s,n} d\nu$

4. $\int \zeta_{r,n}^2 \zeta_{s,n} \zeta_{t,n} d\nu$

5. $\int \zeta_{r,n}\zeta_{s,n}\zeta_{t,n}\zeta_{u,n} d\nu$

where r, s, t, u are distinct. We assert that the terms of the third, fourth, and fifth forms all vanish by virtue of

independence. This leaves k_n terms of the first form and $3k_n(k_n - 1)$ terms of the second form. Thus there

are $k_n + 3k_n(k_n - 1)$ terms of absolute value $\leq C^4$. Thus

$$\int S_n^4 \mathrm{d}\nu \leq \left(3k_n^2 - 2k_n\right)^2 C^4$$

$$\leq 3k_n^2 C^4$$

$$\Rightarrow \mu(E_{n,\epsilon}) \leq \frac{3k_n^2 C^4}{(k_n \epsilon)^4}$$

$$\Rightarrow \sum_{n=1}^{\infty} \mu(E_{n,\epsilon}) \leq \frac{3C^4}{\epsilon^4} \sum_{n=1}^{\infty} k_n^{-2}$$

$$< \infty.$$

By the Borell-Cantelli Lemma, it follows that $\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} E_{n,\epsilon}\right) = 0$. But

$$\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} E_{n,\epsilon} = \left\{ y \in Y : \limsup_{n \to \infty} \left| \frac{S_n(y)}{k_n} \right| \geq \epsilon \right\},$$

so we can conclude that

$$\mu\left(\left\{ y \in Y : \limsup_{n \to \infty} \left| \frac{S_n(y)}{k_n} \right| > 0 \right\} \right) = \mu\left(\bigcup_{K=1}^{\infty} \left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} E_{n,\frac{1}{K}} \right) \right) = 0.$$

Thus $\frac{S_n}{k_n} \to m$ almost surely. $\qquad\qquad\square$

Now we apply this to estimating

$$\frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(C_k(x))} \int_{C_k(x)} T^i \chi_{[a_0, a_1, \dots, a_{\ell-1}]},$$

where $C_k(x)$ is the rank-k cylinder associated to x (see the discussion near the beginning of Section 2.2). Fix a word $\mathbf{a} = (a_0, a_1, \dots, a_{\ell-1}) \in \mathcal{D}^\ell$. We are going to consider a sequence of families of discrete random variables in X given by

$$\xi_{i,k}^{\mathbf{a}}(x) = \frac{1}{\mu(C_k(x))} \int_{C_k(x)} T^i \chi_{[\mathbf{a}]} \mathrm{d}\mu$$

$$= \frac{\mu\left(C_k(x) \cap T^{-i}[\mathbf{a}]\right)}{\mu(C_k(x))}. \qquad\qquad (0 \leq i \leq k-1)$$

Each random variable is bounded in $L^\infty(X, \mu)$ by 1. We claim that they also have a shared mean $\int \xi_{i,k}^{\mathbf{a}} \mathrm{d}\mu = \mu([\mathbf{a}])$.

$$
\int \xi_{i,k}^{\mathbf{a}} \mathrm{d}\mu
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \left(\prod_{h=0}^{k-1} p(d_h) \right) \alpha_{[d_0, d_1, \ldots, d_{k-1}]} \left(T^i \chi_{[a_0, a_1, \ldots, a_{\ell-1}]} \right)
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \left(\prod_{h=0}^{k-1} p(d_h) \right) \frac{1}{\prod_{h=0}^{k-1} p(d_h)} \int_{[d_0, d_1, \ldots, d_{k-1}]} T^i \chi_{[a_0, a_1, \ldots, a_{\ell-1}]} \mathrm{d}\mu
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \int_{[d_0, d_1, \ldots, d_{k-1}]} T^i \chi_{[a_0, a_1, \ldots, a_{\ell-1}]} \mathrm{d}\mu
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \int_{[d_0, d_1, \ldots, d_{k-1}]} \chi_{T^{-i}[a_0, a_1, \ldots, a_{\ell-1}]} \mathrm{d}\mu
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \mu \left([d_0, d_1, \ldots, d_{k-1}] \cap T^{-i}[a_0, a_1, \ldots, a_{\ell-1}] \right)
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \mu \left([d_0, d_1, \ldots, d_{k-1}] \cap \bigcup_{c_0, c_1, \ldots, c_{i-1}} [c_0, c_1, \ldots, c_{i-1}, a_0, a_1, \ldots, a_{\ell-1}] \right)
$$

$$
= \sum_{\vec{d} \in \mathcal{D}^k} \mu \left([d_0, d_1, \ldots, d_{k-1}] \cap [d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}] \right)
$$

To compute this value, we look at two cases: where $i + \ell \leq k$, and where $i + \ell \geq k$.

If $i + \ell \leq k$, then

$$
[d_0, d_1, \ldots, d_{k-1}] \cap [d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}]
$$

$$
= \begin{cases} [d_0, d_1, \ldots, d_{k-1}] & d_i = a_0, d_{i+1} = a_1, \ldots, d_{i+\ell-1} = a_{\ell-1} \\ \emptyset & \text{otherwise} \end{cases}
$$

This means that $d_0, d_1, \ldots, d_{i-1}$, as well as $d_{i+\ell}, \ldots, d_{k-1}$ are "free". Thus

$$\sum_{\vec{d} \in \mathcal{D}^k} \mu\left([d_0, d_1, \ldots, d_{k-1}] \cap [d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}]\right)$$

$$= \sum_{\vec{d} \in \mathcal{D}^l} \mu([d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}, d_{i+\ell}, \ldots, d_{k-1}])$$

$$= \sum_{\vec{d} \in \mathcal{D}^k} (p(d_0)p(d_1) \cdots p(d_{i-1})) \, (p(a_0)p(a_1) \cdots p(a_{\ell-1})) \, (p(d_{i+\ell}) \cdots p(d_{k-1}))$$

$$= \mu([a_0, a_1, \ldots, a_{\ell-1}]).$$

On the other hand, if $i + \ell \geq k$, then

$$[d_0, d_1, \ldots, d_{k-1}] \cap [d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}]$$

$$= \begin{cases} [d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}] & d_i = a_0, \ldots, d_{k-1} = a_{k-i-1} \\ \emptyset & \text{otherwise} \end{cases}$$

leaving $d_0, d_1, \ldots, d_{i-1}$ "free". Thus

$$\sum_{\vec{d} \in \mathcal{D}^k} \mu\left([d_0, d_1, \ldots, d_{k-1}] \cap [d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}]\right)$$

$$= \sum_{\vec{d} \in \mathcal{D}^k} \mu([d_0, d_1, \ldots, d_{i-1}, a_0, a_1, \ldots, a_{\ell-1}])$$

$$= \sum_{\vec{d} \in \mathcal{D}^k} p(d_0)p(d_1) \cdots p(d_{i-1})p(a_0)p(a_1) \cdots p(a_{\ell-1})$$

$$= \mu([a_0, a_1, \ldots, a_{\ell-1}]).$$

Thus in either case, we have $\int \xi_{i,k}^{\mathbf{a}} \mathrm{d}\mu = \mu([\mathbf{a}])$.

Now, for fixed k, the family $\left\{\xi_{i,k}^{\mathbf{a}}\right\}_{i=0}^{k-1}$ is not necessarily independent, but we can break it up into arithmetic subsequences which are. Consider the families $\left\{\xi_{m\ell+j,k}^{\mathbf{a}}\right\}_{m=0}^{\lfloor k/\ell \rfloor - 1}$ for $j \in \{0, 1, \ldots, \ell - 1\}$. Then these subfamilies are independent, so the Even Stronger Law Of Large Numbers tells us that $\frac{1}{\lfloor k/\ell \rfloor} \sum_{m=0}^{k-1} \xi_{m\ell+j,k}^{\mathbf{a}} \to$

$\mu([\mathbf{a}])$ almost surely. Now we calculate

$$\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(C_k(x))}\int_{C_k(x)}T^i\chi_{[\mathbf{a}]}\mathrm{d}\mu$$

$$=\frac{1}{k}\sum_{i=0}^{k-1}\xi_{i,k}^{\mathbf{a}}(x)$$

$$=\frac{\ell\lfloor k/\ell\rfloor}{k}\left[\frac{1}{\ell\lfloor k/\ell\rfloor}\sum_{i=0}^{k-1}\xi_{i,k}^{\mathbf{a}}(x)\right]$$

$$=\frac{\ell\lfloor k/\ell\rfloor}{k}\left[\frac{1}{\ell}\sum_{j=0}^{\ell-1}\frac{1}{\lfloor k/\ell\rfloor}\sum_{m=0}^{\lfloor k/\ell\rfloor-1}\xi_{m\ell+j,k}^{\mathbf{a}}(x)\right]+\frac{\sum_{i=\ell\lfloor k/\ell\rfloor}^{k-1}\xi_{i,k}^{\mathbf{a}}(x)}{\ell}$$

$$\overset{\text{almost surely}}{\to}(1)\left[\frac{1}{\ell}\sum_{j=0}^{\ell-1}\mu([\mathbf{a}])\right]+0$$

$$=\mu([\mathbf{a}])$$

$$=\int\chi_{[\mathbf{a}]}\mathrm{d}\mu.$$

Taking a countable intersection over $\mathbf{a}\in\bigcup_{\ell=1}^{\infty}\mathcal{D}^\ell$, we can conclude that the set B of all $x\in X$ such that $\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{\mu(C_k(x))}\int_{C_k(x)}T^i\chi_{[\mathbf{a}]}\mathrm{d}\mu\to\int\chi_{[\mathbf{a}]}\mathrm{d}\mu$ for all words $\mathbf{a}$ is of full measure. We can further conclude that if $x\in B$, we have $\frac{1}{\mu(C_k(x))}\int_{C_k(x)}T^iT^n\chi_{[\mathbf{a}]}\mathrm{d}\mu\to\int T^n\chi_{[\mathbf{a}]}$ for all words $\mathbf{a}$ and $n\in\mathbb{Z}$. Since span $\left\{T^n\chi_{[\mathbf{a}]}:\mathbf{a}\in\bigcup_{\ell=1}^{\infty}\mathcal{D}^\ell,n\in\mathbb{Z}\right\}$ is dense in $C(X)$, we can conclude the following special case of Theorem 2.3.5.

Proposition 2.4.2. *Let $X=\mathcal{D}^{\mathbb{Z}}$ be a Bernoulli shift, and let μ be the associated measure. Endow X with the generator $\mathcal{E}=\{E_d\}_{d\in\mathcal{D}}$, where $E_d=\{x\in X:x_0=d\}$. Then the set of all $x\in X$ such that*

$$\alpha_{C_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^if\right)\to\int f\mathrm{d}\mu$$

for all $f\in C(X)$ is of full measure.

However, this technique lends itself to another result that is not encompassed by Theorem 2.3.5. We have looked at temporo-spatial differentiation problems where we are differentiating with respect to the cylinders $C_k(x)$ of a randomly chosen $x\in X$. The next result considers instead the situation where we randomly choose a *sequence* $(x_k)_{k=1}^{\infty}$ in X and differentiating with respect to the sequence $(C_k(x_k))_{k=1}^{\infty}$.

Theorem 2.4.3. *Let $X = \mathcal{D}^{\mathbb{Z}}$ be a Bernoulli shift, and let μ be the associated measure. Endow X with the generator $\mathcal{E} = \{E_d\}_{d \in \mathcal{D}}$, where $E_d = \{x \in X : x_0 = d\}$. Consider the countably infinite product probability space $(X^\infty, \mathcal{B}^\infty, \mu^\infty) = \prod_{k \in \mathbb{N}}(X, \mathcal{B}, \mu)$. Then the set of all $(x_k)_{k=1}^\infty \in X^\infty$ such that*

$$\alpha_{C_k(x_k)}\left(\frac{1}{k}\sum_{i=0}^{k-1}T^i f\right) \to \int f \mathrm{d}\mu$$

for all $f \in C(X)$ is of full μ^∞-measure.

Proof. Our method is very similar to the method used for Proposition 2.4.2. Let $\mathbf{x} = (x_k)_{k=1}^\infty \in X^\infty$ denote a sequence in X.

Fix a word $\mathbf{a} = (a_0, a_1, \ldots, a_{\ell-1}) \in \mathcal{D}^\ell$. We are going to consider a sequence of families of discrete random variables in X given by

$$\begin{aligned}
\zeta_{i,k}^{\mathbf{a}}(\mathbf{x}) &= \frac{1}{\mu(C_k(x_k))}\int_{C_k(x_k)} T^i \chi_{[\mathbf{a}]} \mathrm{d}\mu \\
&= \frac{\mu\left(C_k(x_k) \cap [\mathbf{a}]\right)}{\mu(C_k(x_k))}. \qquad\qquad (0 \le i \le k-1)
\end{aligned}$$

Each random variable $\zeta_{i,k}^{\mathbf{a}}$ is bounded in $L^\infty(X, \mu)$ by 1. By a calculation identical to the one used to prove Proposition 2.4.2, we can conclude that $\int \zeta_{i,k}^{\mathbf{a}} \mathrm{d}\mu = \mu([\mathbf{a}])$.

As before, for fixed k, the family $\left\{\zeta_{i,k}^{\mathbf{a}}\right\}_{i=0}^{k-1}$ is not necessarily independent, but we can break it up into arithmetic subsequences which are. Consider the families $\left\{\zeta_{m\ell+j,k}^{\mathbf{a}}\right\}_{m=0}^{\lfloor k/\ell \rfloor - 1}$ for $j \in \{0, 1, \ldots, \ell - 1\}$. Then these families are independent, and so the Even Stronger Law Of Large Numbers tells us that

$\frac{1}{\lfloor k/\ell \rfloor} \sum_{m=0}^{k-1} \zeta_{m\ell+j,k}^{\mathbf{a}} \to \mu([\mathbf{a}])$ almost surely. Now we calculate

$$\frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(C_k(x_k))} \int_{C_k(x_k)} T^i \chi_{[\mathbf{a}]} \mathrm{d}\mu$$

$$= \frac{1}{k} \sum_{i=0}^{k-1} \zeta_{i,k}^{\mathbf{a}}(\mathbf{x})$$

$$= \frac{\ell \lfloor k/\ell \rfloor}{k} \left[\frac{1}{\ell \lfloor k/\ell \rfloor} \sum_{i=0}^{k-1} \zeta_{i,k}^{\mathbf{a}}(\mathbf{x}) \right]$$

$$= \frac{\ell \lfloor k/\ell \rfloor}{k} \left[\frac{1}{\ell} \sum_{j=0}^{\ell-1} \frac{1}{\lfloor k/\ell \rfloor} \sum_{m=0}^{\lfloor k/\ell \rfloor-1} \zeta_{m\ell+j,k}^{\mathbf{a}}(\mathbf{x}) \right] + \frac{\sum_{i=\ell \lfloor k/\ell \rfloor}^{k-1} \zeta_{i,k}^{\mathbf{a}}(\mathbf{x})}{\ell}$$

$$\xrightarrow{\text{almost surely}} (1) \left[\frac{1}{\ell} \sum_{j=0}^{\ell-1} \mu([\mathbf{a}]) \right] + 0$$

$$= \mu([\mathbf{a}])$$

$$= \int \chi_{[\mathbf{a}]} \mathrm{d}\mu.$$

Again, taking a countable intersection over $\mathbf{a} \in \bigcup_{\ell=1}^{\infty} \mathcal{D}^\ell$, we can conclude that the set B of all $\mathbf{x} \in X^\infty$ such that $\frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(C_k(x_k))} \int_{C_k(x_k)} T^i \chi_{[\mathbf{a}]} \mathrm{d}\mu \to \int \chi_{[\mathbf{a}]} \mathrm{d}\mu$ for all words $\mathbf{a}$ is of full measure. We can further conclude that if $\mathbf{x} \in B$, we have $\frac{1}{\mu(C_k(x_k))} \int_{C_k(x_k)} T^i T^n \chi_{[\mathbf{a}]} \mathrm{d}\mu \to \int T^n \chi_{[\mathbf{a}]}$ for all words $\mathbf{a}$ and $n \in \mathbb{Z}$. Since span $\left\{ T^n \chi_{[\mathbf{a}]} : \mathbf{a} \in \bigcup_{\ell=1}^{\infty} \mathcal{D}^\ell, n \in \mathbb{Z} \right\}$ is dense in $C(X)$, we can conclude that if $\mathbf{x} \in B$, then

$$\alpha_{C_k(x_k)} \left(\frac{1}{k} \sum_{i=0}^{k-1} T^i f \right) \to \int f \mathrm{d}\mu$$

for all $f \in C(X)$. $\qquad \square$

Temporo-spatial differentiations for actions of topological groups and the pointwise reduction heuristic

Our primary goal with this chapter is to extend some of the results in Chapter 2 to the setting of actions of amenable groups. In particular, we demonstrate several special cases of a general heuristic: that a temporo-spatial differentiation relative to a family of sets containing a point x with diameter going to 0 sufficiently fast will be equivalent to a pointwise ergodic average at the point x. This heuristic applies to many kinds of ergodic averages.

In Section 3.1, we provide some general results about temporo-spatial differentiations. In particular, we provide a characterization in terms of ergodic optimization of a kind of "best-case scenario" behavior, where temporo-spatial averages of continuous functions always converge to the integral.

In Section 3.2, we provide convergence theorems for two special cases of temporo-spatial differentiation averages: where the spatial averaging sets have measure going to 1, and where they are constant.

In Section 3.3, we show that in the case where the spatial averaging sets share a common fixed point x and have diameter going to 0 sufficiently fast, then the associated temporo-spatial differentiations can be reduced to a pointwise temporal average at that fixed point x. This then provides us a means to prove convergence results for suitable "random temporo-spatial differentiation problems." In particular, these reduction results can be applied even when the temporal averaging sets are not Følner.

In Section 3.4, we generalize some of the results of Section 3.3 to the setting of weighted temporo-spatial ergodic averages. These include equicontinuous families of continuous weight functions of modulus 1, as well as potentially unbounded weight sequences of complex constants.

3.1   General results and unique ergodicity

Throughout this chapter, by a *topological dynamical system*, we will mean a continuous action T of a locally compact *unimodular* topological group G on a compact metrizable space X, denoted $T : G \curvearrowright X$. We will use m to refer to a left- and right-invariant Haar measure on the group G, hereafter referred to simply as a Haar measure. Our consideration of unimodular groups is primarily to simplify some bookkeeping about

when we are invoking a left-invariant Haar measure and when we are invoking a right-invariant Haar measure. Of course, the class of unimodular groups includes all abelian groups, all discrete groups, and all compact groups, thus encompassing many of the groups ergodic theory classically considers actions of.

Definition 3.1.1. Let $T : G \curvearrowright X$ be a topological dynamical system, where G is an amenable group, and let $f \in C_{\mathbb{R}}(X)$ be a real-valued continuous function on X. The *gauge* of f is the value

$$\Gamma(f) := \sup \left\{ \int f \mathrm{d}\mu : \mu \in \mathcal{M}_T(X) \right\},$$

where $\mathcal{M}_T(X)$ denotes the family of T-invariant Borel probability measures on X. We say that $\mu \in \mathcal{M}_T(X)$ is *f-maximizing* if $\int f \mathrm{d}\mu = \Gamma(f)$, and denote the class of all f-maximizing measures on X by

$$\mathcal{M}_{\max}(f) := \left\{ \mu \in \mathcal{M}_T(X) : \int f \mathrm{d}\mu = \Gamma(f) \right\}.$$

The gauge is well-defined, since if G is amenable, then $\mathcal{M}_T(X)$ is a nonempty Choquet simplex in the weak*-topology. Since $\mathcal{M}_T(X)$ is compact, it follows that $\mathcal{M}_{\max}(f)$ is nonempty.

We now wish to provide an alternative description of the gauge for nonnegative-valued functions.

Lemma 3.1.2. *Let $T : G \curvearrowright X$ be a topological dynamical system, and let $K \subseteq G$ be a compact subset of a locally compact group G. Let $f \in C(X)$. Then the function $x \mapsto \int_K f(T_g x) \mathrm{d}m(g)$ is continuous, where m is a Haar measure on G.*

Proof. We can assume that K is of positive Haar measure, and in particular nonempty, since otherwise this would be trivial.

Fix $\epsilon > 0$, and let ρ be a compatible metric for X. We know a priori that the function $G \times X \to \mathbb{C}$ given by $(g, x) \mapsto f(T_g x)$ is continuous, so for each $g \in K$, choose an open neighborhood $U_g \subseteq G$ of g and a positive number $\delta_g > 0$ such that if $(g', x') \in U_g \times B(x, \delta_g)$, then $\left| f\left(T_{g'} x'\right) - f(T_g x) \right| < \frac{\epsilon}{2m(K)}$. Then $\{U_g\}_{g \in K}$ is an open cover of the compact K, so there exist $g_1, \ldots, g_n \in K$ such that $K \subseteq U_{g_1} \cup \cdots \cup U_{g_n}$. Let $\delta = \min\{\delta_{g_1}, \ldots, \delta_{g_n}\}$. Then if $\rho(x, y) < \delta$, and $g \in K$, then $g \in U_{g_j}$ for some $j \in \{1, \ldots, n\}$. Therefore $(g, x), (g, y) \in U_{g_j} \times B(x, \delta) \subseteq U_{g_j} \times B(x, \delta_j)$, so

$$|f(T_g x) - f(T_g y)| \leq \left| f(T_g x) - f\left(T_{g_j} x\right) \right| + \left| f\left(T_{g_j} x\right) - f(T_g y) \right| < \frac{\epsilon}{2m(K)} + \frac{\epsilon}{2m(K)} = \frac{\epsilon}{m(K)}.$$

Thus there exists $\delta > 0$ such that if $\rho(x, y) < \delta$, then $|f(T_gx) - f(T_gy)| < \frac{\epsilon}{m(K)}$. Therefore, it follows that if $\rho(x, y) < \delta$, then

$$\left| \int_K f(T_gx)\mathrm{d}m(g) - \int_K f(T_gy)\mathrm{d}m(g) \right| = \left| \int_K \left(f(T_gx) - f(T_gy) \right) \mathrm{d}m(g) \right|$$

$$\leq \int_K |f(T_gx) - f(T_gy)|\, \mathrm{d}m(g)$$

$$\leq \int_K \frac{\epsilon}{m(K)} \mathrm{d}m(g)$$

$$= \epsilon.$$

Therefore the function $x \mapsto \int_k f(T_gx)\mathrm{d}m(g)$ is continuous. $\qquad\square$

Notation 3.1.3. (a) Let $T : G \curvearrowright X$ be a continuous action of a locally compact group G with Haar measure m on a topological space X, and let f be a continuous function $X \to \mathbb{C}$. Let K be a compact subset of G with positive Haar measure. We define $\mathrm{Avg}_K f : X \to \mathbb{C}$ to be the continuous function

$$\mathrm{Avg}_K f(x) = \frac{1}{m(K)} \int_K f(T_gx)\mathrm{d}m(g).$$

The continuity of $\mathrm{Avg}_K f$ follows from Lemma 3.1.2. In the event where G is a discrete group, we will also define $\mathrm{Avg}_K f$ for all nonempty compact subsets K of G and $f \in L^1(X, \mu)$ by

$$\mathrm{Avg}_K f = \frac{1}{|K|} \sum_{g \in K} f \circ T_g.$$

(b) Let (X, μ) be a probability space, and let $f \in L^1(X, \mu)$. Let C be a measurable subset of X with $\mu(C) > 0$. We define the functional $\alpha_C : L^1(X, \mu) \to \mathbb{C}$ by

$$\alpha_C(f) := \frac{1}{\mu(C)} \int f\mathrm{d}\mu.$$

Although the functionals α_C are defined here on L^1, we will almost always be interested in their action on L^∞, where they are considerably better behaved.

Definition 3.1.4. Let G be a locally compact topological group. A net $(F_i)_{i \in \mathscr{I}}$ of compact subsets of G is called *Følner* if $m(F_i) > 0$ for all $i \in \mathscr{I}$, and

$$\lim_i \frac{m(gF_i \Delta F_i)}{m(F_i)} = 0 \qquad\qquad (\forall g \in G),$$

where Δ denotes the symmetric difference $A\Delta B = (A \setminus B) \cup (B \setminus A)$.

Theorem 3.1.5. *Let $T : G \curvearrowright X$ be a topological dynamical system, where G is an amenable group with Haar measure m. Let $(F_i)_{i \in \mathscr{I}}$ be a left Følner net for G, and let $f \in C_{\mathbb{R}}(X)$ be a nonnegative-valued continuous function on X. Then the net $\left(\left\| \mathrm{Avg}_{F_i} f \right\|_{C(X)} \right)_{i \in \mathscr{I}}$ converges, and*

$$\Gamma(f) = \lim_i \left\| \mathrm{Avg}_{F_i} f \right\|.$$

Proof. For each $i \in \mathscr{I}$, let σ_i be a Borel probability measure on X such that

$$\int \mathrm{Avg}_{F_i} f \mathrm{d}\sigma_i = \left\| \mathrm{Avg}_{F_i} f \right\|_{C(X)}.$$

For each $i \in \mathscr{I}$, define the Borel probability measure μ_i on X by

$$\int f \mathrm{d}\mu_i = \int_X \left(\mathrm{Avg}_{F_i} f \right) \mathrm{d}\sigma_i = \frac{1}{m(F_i)} \int_{F_i} \left(\int_X T_g f(x) \mathrm{d}\sigma_i(x) \right) \mathrm{d}m(g),$$

where the latter equality follows from Fubini's Theorem.

In order to prove that the net $\left(\int f \mathrm{d}\mu_i \right)_{i \in \mathscr{I}}$ converges to $\Gamma(f)$, it will suffice to prove that for any convergent sub-net $\left(\mu_{i_j} \right)_{j \in \mathscr{J}}$, the sub-net $\left(\int f \mathrm{d}\mu_{i_j} \right)_{j \in \mathscr{J}}$ converges to $\Gamma(f)$, because if $\left(\int f \mathrm{d}\mu_i \right)_{i \in \mathscr{I}}$ didn't converge to $\Gamma(f)$, then we could extract some subnet $\left(\mu_{i_j} \right)_{j \in \mathscr{J}}$ along which $\left(\int f \mathrm{d}\mu_{i_j} \right)_{j \in \mathscr{J}}$ converged to some other point (because the net is contained in a compact subset of $\mathbb{R}$), then take a weak*-convergent subnet of that, yielding a contradiction.

Therefore, we need to show that every weak*-limit point of the net $(\mu_i)_{i \in \mathscr{I}}$ is f-maximizing.

Since the space $\mathcal{M}(X)$ of Borel probability measures on X is weak*-compact, it follows that there exists a weak*-convergent sub-net $\left(\mu_{i_j} \right)_{j \in \mathscr{J}}$, converging to some μ. It follows then that μ is T-invariant, since if

$f_0 \in C(X), g_0 \in G$, then

$$\left| \int T_{g_0} f_0 \mathrm{d}\mu - \int f_0 \mathrm{d}\mu \right| = \left| \int (T_{g_0} f_0 - f_0)\, \mathrm{d}\mu \right|$$

$$= \lim_j \left| \int (T_{g_0} f_0 - f_0)\, \mathrm{d}\mu_{i_j} \right|,$$

where

$$\left| \int_X \frac{1}{m\left(F_{i_j}\right)} \int_{F_{i_j}} (T_{g_0} f_0(x) - f_0(x))\, \mathrm{d}m(g)\mathrm{d}\sigma_{i_j}(x) \right|$$

$$= \left| \int_X \frac{1}{m\left(F_{i_j}\right)} \left(\left(\int_{F_{i_j} g_0 \backslash F_{i_j}} T_g f_0(x)\mathrm{d}m(g) \right) - \left(\int_{F_{i_j} \backslash F_{i_j} g_0} T_g f_0(x)\mathrm{d}m(g) \right) \right) \mathrm{d}\sigma_{i_j}(x) \right|$$

$$\leq \left| \int_X \frac{1}{m\left(F_{i_j}\right)} \int_{g_0 F_{i_j} \backslash F_{i_j}} T_g f_0(x)\mathrm{d}m(g)\mathrm{d}\sigma_{i_j}(x) \right|$$

$$+ \left| \int_X \frac{1}{m\left(F_{i_j}\right)} \int_{F_{i_j} \backslash g_0 F_{i_j}} T_g f_0(x)\mathrm{d}m(g)\mathrm{d}\sigma_{i_j}(x) \right|$$

$$\leq \frac{m\left(g_0 F_{i_j} \backslash F_{i_j}\right) + m\left(F_{i_j} \backslash g_0 F_{i_j}\right)}{m\left(F_{i_j}\right)} \|f_0\|_{C(X)}$$

$$= \frac{m\left(F_{i_j} g_0 \Delta F_{i_j}\right)}{m\left(F_{i_j}\right)} \|f_0\|_{C(X)}$$

$$\overset{j \to \infty}{\to} 0.$$

Therefore $\int T_{g_0} f_0 \mathrm{d}\mu = \int f_0 \mathrm{d}\mu$, meaning μ is T-invariant.

We claim that μ is f-maximizing. On one hand, we know that $\int f \mathrm{d}\mu \leq \Gamma(f)$, because $\mu \in \mathcal{M}_T(X)$. Now, suppose that $\nu \in \mathcal{M}_T(X)$. Then

$$\int f \mathrm{d}\nu = \int_X \frac{1}{m\left(F_{i_j}\right)} T_g f(x)\mathrm{d}m(g)\mathrm{d}\nu(x)$$

$$\leq \left\| \frac{1}{m\left(F_{i_j}\right)} T_g f \mathrm{d}m(g) \right\|_{C(X)}$$

$$= \int f \mathrm{d}\mu_{i_j}$$

$$\Rightarrow \int f \mathrm{d}\nu \leq \lim_j \int f \mathrm{d}\mu_{i_j}$$

$$= \int f \mathrm{d}\mu.$$

Therefore $\int f\mathrm{d}\mu \geq \int f\mathrm{d}\nu$ for all $\nu \in \mathcal{M}_T(X)$, meaning that $\int f\mathrm{d}\mu = \sup_{\nu \in \mathcal{M}_T(X)} \int f\mathrm{d}\nu$, i.e. μ is f-maximizing, so $\int f\mathrm{d}\mu = \Gamma(f)$. $\qquad \square$

From this, we can use the gauge to provide a characterization of uniquely ergodic systems.

Theorem 3.1.6. *Let $T : G \curvearrowright X$ be a topological dynamical system, where G is amenable, and let $\mu \in \mathcal{M}_T(X)$ be a T-invariant Borel probability measure on X that is fully supported on X, i.e. gives positive measure to every nonempty open subset of X. Then $T : G \curvearrowright X$ is uniquely ergodic if and only if $\Gamma(f) = \int f\mathrm{d}\mu$ for all nonnegative-valued $f \in C_{\mathbb{R}}(X)$.*

Proof. Clearly $\int f\mathrm{d}\mu \leq \Gamma(f)$ for all $f \in C_{\mathbb{R}}(X)$.

($\Rightarrow$) If $T : G \curvearrowright X$ is uniquely ergodic, then μ is f-maximizing for all $f \in C_{\mathbb{R}}(X)$, so in particular $\int f\mathrm{d}\mu = \Gamma(f)$ for all nonnegative $f \in C_{\mathbb{R}}(X)$.

($\Leftarrow$) We'll prove the contrapositive. Suppose that $T : G \curvearrowright X$ is *not* uniquely ergodic. Then there exists an ergodic T-invariant Borel probability measure $\nu \neq \mu$. By (Jenkinson, 2006b, Theorem 1), there exists a continuous real-valued function $f \in C_{\mathbb{R}}(X)$ such that $\mathcal{M}_{\max}(f) = \{\nu\}$. By possibly adding a nonnegative constant to f, we can assume that f is nonnegative. But $\Gamma(f) = \int f\mathrm{d}\nu \neq \int f\mathrm{d}\mu$. $\qquad \square$

Finally, we want to provide a connection between unique ergodicity and temporo-spatial differentiation problems. Before stating the main theorem relating these, we prove the following lemma that relates the α_C functionals to the L^∞ norm.

Lemma 3.1.7. *Let (X, μ) be a probability space, and let $f \in L^\infty(X, \mu)$. Then*

$$\frac{1}{2}\|f\|_\infty \leq \sup\left\{|\alpha_C(f)| : C \subseteq X \text{ measurable}, \mu(C) > 0\right\} \leq \|f\|_\infty,$$

and in particular, if f is real-valued, then

$$\sup\left\{|\alpha_C(f)| : C \subseteq X \text{ measurable}, \mu(C) > 0\right\} = \|f\|_\infty.$$

Proof. In either case, it's clear that $|\alpha_C(f)| \leq \|f\|_\infty$ for all $C \subseteq X$ measurable with $\mu(C) > 0$, since α_C is a state on $L^\infty(X, \mu)$

Consider now the case that f is real-valued. If $\|f\|_\infty = 0$, then the equality is immediate, so suppose that $\|f\|_\infty > 0$. Set $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$. Then $\|f\|_\infty = \max\left\{\|f^+\|_\infty, \|f^-\|_\infty\right\}$. Assume

without loss of generality that $\|f\|_\infty = \|f^+\|_\infty$. For each $k \in \mathbb{N}$, set

$$C_k = \left\{ x \in X : f^+(x) > \frac{k}{k+1} \|f^+\|_\infty \right\} = \left\{ x \in X : f(x) > \frac{k}{k+1} \|f\|_\infty \right\}.$$

Then $\mu(C_k) > 0$ and $\alpha_{C_k}(f) \geq \frac{k}{k+1} \|f\|_\infty$ for all $k \in \mathbb{N}$, meaning in particular that

$$\sup \left\{ |\alpha_C(f)| : C \subseteq X \text{ measurable, } \mu(C) > 0 \right\} \geq \|f\|_\infty.$$

Now suppose that f is not necessarily real-valued. Let $h_1, h_2 \in L_\mathbb{R}^\infty(X, \mu)$ be the real and imaginary parts of f, respectively. Then $\|f\|_\infty \leq \|h_1\|_\infty + \|h_2\|_\infty$. Therefore $\max \{\|h_1\|_\infty, \|h_2\|_\infty\} \geq \frac{1}{2}\|f\|_\infty$. Assume without loss of generality that $\|h_1\|_\infty \geq \|h_2\|_\infty$, so $\|h_1\|_\infty \geq \frac{1}{2}\|f\|_\infty$. For each $k \in \mathbb{N}$, choose $C_k' \subseteq X$ measurable such that $\mu(C_k') > 0$, and $\alpha_{C_k'}(h_1) \geq \frac{k}{k+1}\|h_1\|_\infty$, which is possible if we appeal to the real case. Then

$$\begin{aligned}
\left| \alpha_{C_k'}(f) \right| &= \left| \alpha_{C_k'}(h_1) + i\alpha_{C_k'}(h_2) \right| \\
&\geq \left| \alpha_{C_k'}(h_1) \right| \\
&\geq \frac{k}{k+1} \|h_1\|_\infty \\
&\geq \frac{k}{k+1} \left(\frac{1}{2}\|f\|_\infty \right).
\end{aligned}$$

Taking the limit as $k \to \infty$ verifies that

$$\sup \left\{ |\alpha_C(f)| : C \subseteq X \text{ measurable, } \mu(C) > 0 \right\} \geq \frac{1}{2}\|f\|_\infty.$$

$\square$

In the case that X is a compact metrizable space, the measure μ is Borel, and the f is continuous, Lemma 3.1.7 can be sharpened as follows.

Lemma 3.1.8. *Let (X, μ) be a probability space, where X is a compact metrizable space and μ is a Borel probability measure. Let $f \in C(X)$. Then*

$$\frac{1}{2}\|f\|_\infty \leq \sup \left\{ |\alpha_C(f)| : C \subseteq X \text{ open, } \mu(C) > 0 \right\} \leq \|f\|_\infty,$$

and in particular, if f is real-valued, then we have

$$\sup\{|\alpha_C(f)| : C \subseteq X \text{ open}, \mu(C) > 0\} = \|f\|_\infty.$$

Proof. Under these conditions, all the C_k and C_k' in the proof of Lemma 3.1.7 are open. The result follows from the same proof. □

A natural corollary of Lemma 3.1.7 is the following qualitative statement.

Theorem 3.1.9. *Let $T : G \curvearrowright (X, \mu)$ be a measure-preserving action of a discrete (not necessarily amenable) group G on a probability space (X, μ). Let $(F_i)_{i \in \mathscr{I}}$ be a net of compact subsets of G with positive measure. Let $f \in L^\infty(X, \mu)$. Then the following conditions are equivalent.*

  (i) $\mathrm{Avg}_{F_i} f \to \int f \mathrm{d}\mu$ in the norm topology on L^∞.

  (ii) $\alpha_{C_i}(\mathrm{Avg}_{F_i} f) \to \int f \mathrm{d}\mu$ for all nets of measurable subsets C_i of X with positive measure.

Proof. This equivalence follows from the estimates in Lemma 3.1.7. For each $i \in \mathscr{I}$, set

$$f_i = \mathrm{Avg}_{F_i} f - \int f \mathrm{d}\mu.$$

(i)⇒(ii) Suppose that $f_i \to \int f \mathrm{d}\mu$ in L^∞, and let $(C_i)_{i \in \mathscr{I}}$ be a net of measurable subsets C_i of X with positive measure. Then

$$\left|\alpha_{C_i}\left(\mathrm{Avg}_{F_i} f\right) - \int f \mathrm{d}\mu\right| = |\alpha_{C_i}(f_i)| \le \|f_i\|_\infty \overset{i \to \infty}{\to} 0.$$

(ii)⇒(i) We'll prove ¬(i)⇒ ¬(ii). Suppose that $\limsup_i \|f_i\|_\infty > 0$. For each $i \in \mathscr{I}$, choose $C_i \subseteq X$ measurable with positive measure such that

$$|\alpha_{C_i}(f_i)| \ge \frac{1}{2} \sup\{|\alpha_C(f_i)| : C \subseteq X \text{ measurable}, \mu(C) > 0\} \ge \frac{1}{4} \|f_i\|_\infty.$$

Then $\limsup_i \left|\alpha_{C_i}\left(\mathrm{Avg}_{F_i} f\right) - \int f \mathrm{d}\mu\right| = \limsup_i |\alpha_{C_i}(f_i)| \ge \frac{1}{4} \limsup_i \|f_i\|_\infty > 0.$ □

Theorem 3.1.9 gives a qualitative description of the conditions under which we get the "best possible" behavior for a temporo-spatial differentiation problem, i.e. conditions under which $\alpha_{C_i}\left(\mathrm{Avg}_{F_i} f\right) \to \int f \mathrm{d}\mu$

independent of the choice of C_i. However, Lemma 3.1.7 provides a potential avenue for quantitative estimates on the rate of convergence for temporo-spatial averages by "importing" estimates on the rate of L^∞-convergence for $\mathrm{Avg}_{F_i} \to \int f \mathrm{d}\mu$.

In general, the classical ergodic theorems don't give us estimates on the rate of convergence they promise, and this convergence can in fact be very slow. See the "Speed of Convergence" discussion in §1.2 of (Krengel, 2011) for a survey of relevant counterexamples. However, some authors have studied situations where effective estimates on the convergence of certain ergodic averages can be obtained. For a rudimentary example of this type, consider the case where $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is an irrational real, and $T : \mathbb{Z} \curvearrowright (\mathbb{R}/\mathbb{Z})$ is an action of $\mathbb{Z}$ on the circle by $Tx = \alpha + x$, where $\mathbb{R}/\mathbb{Z}$ is endowed with its Haar probability measure μ. Then by appealing to the unique ergodicity of (X, T), we can say that $\frac{1}{k} \sum_{j=0}^{k-1} T^j f \to \int f \mathrm{d}\mu$ in $L^\infty(X, \mu)$ for all $f \in C(X)$. However, if $f(x) = e^{2\pi i n x}$ for some $n \in \mathbb{Z} \setminus \{0\}$, i.e. if f is a nontrivial character on $\mathbb{R}/\mathbb{Z}$, then using a geometric series, we can see that $\left\| \frac{1}{k} \sum_{j=0}^{k-1} T^j f - \int f \mathrm{d}\mu \right\|_\infty \leq A_{\alpha,n} k^{-1}$ for some constant $A_{\alpha,n} \in (0, \infty)$, yielding a quantitative estimate on that convergence rate. In particular, Lemma 3.1.7 tells us that under those circumstances, we'd have that

$$\left| \alpha_{C_k} \left(\frac{1}{k} \sum_{j=0}^{k-1} T^j f \right) \right| \leq k^{-1} A_{\alpha,n}$$

for all choices of $(C_k)_{k=1}^\infty$. In this dissertation, we will say no more on this topic, which is linked to the study of effective equidistribution (see (Einsiedler, 2010)).

Lemma 3.1.10. *Let $T : G \curvearrowright X$ be a topological dynamical system, where G is amenable. Let $(F_i)_{i \in \mathscr{I}}$ be a Følner net for G, and let $f \in C(X), \lambda \in \mathbb{C}$. Then the following conditions are related by the implications (i) $\Longleftrightarrow$ (ii)$\Rightarrow$(iii). If in addition we have that $\mathscr{I} = \mathbb{N}$, i.e. that $(F_i)_{i \in \mathbb{N}}$ is a Følner sequence, then (iii)$\Rightarrow$(i).*

(i) $\int f \mathrm{d}\mu = \lambda$ for all T-invariant Borel probability measures μ on X.

(ii) $\mathrm{Avg}_{F_i} f \to \lambda$ uniformly.

(iii) $\mathrm{Avg}_{F_i} f(x) \to \lambda$ for all $x \in X$.

Proof. (i)$\Rightarrow$(ii): Suppose that $\int f\mathrm{d}\mu = \lambda$ for all T-invariant Borel probabiliy measures μ on X, and let $(x_i)_{i\in\mathscr{I}}$ be a net in X such that

$$\left|\mathrm{Avg}_{F_i} f(x_i) - \lambda\right| = \left\|\mathrm{Avg}_{F_i} f - \lambda\right\|_{C(X)} \qquad (\forall i \in \mathscr{I}).$$

Let $(\mu_i)_{i\in I}$ be the net of Borel probability measures on X given by

$$\int g\mathrm{d}\mu_i = \mathrm{Avg}_{F_i} g(x_i) \qquad (\forall g \in C(X), i \in \mathscr{I}).$$

Appealing to compactness, let $\left(\mu_{i_j}\right)_{j\in\mathscr{J}}$ be a weak*-convergent subnet along which

$$\lim_j \left|\int f\mathrm{d}\mu_{i_j} - \lambda\right| = \limsup_i \left|\int f\mathrm{d}\mu_i - \lambda\right|.$$

Let $\mu = \lim_j \mu_{i_j}$. Since $(F_{i_j})_{j\in\mathscr{J}}$ is Følner, it follows from a classical argument that μ is T-invariant, and

$$\left|\int f\mathrm{d}\mu - \lambda\right| = \limsup_i \left|\int f\mathrm{d}\mu_i - \lambda\right| = \limsup_i \left\|\mathrm{Avg}_{F_i} f - \lambda\right\|_{C(X)}.$$

But $\int f\mathrm{d}\mu = \lambda$ by (i), so it follows that $\limsup_i \left\|\mathrm{Avg}_{F_i} f - \lambda\right\|_{C(X)} = 0$, meaning that $\mathrm{Avg}_{F_i} f \to \lambda$ uniformly.

(ii)$\Rightarrow$(i): Trivial.

(ii)$\Rightarrow$(iii): Trivial.

(iii)$\Rightarrow$(i): $(F_i)_{i\in\mathbb{N}}$ is a Følner *sequence*, and let μ be a Borel probability measure on X. Then $\mathrm{Avg}_{F_i} f \overset{i\to\infty}{\to} \lambda$ pointwise-almost everywhere, and the functions $\mathrm{Avg}_{F_i} f$ are dominated by the constant function $\|f\|_{C(X)}$, so we can appeal to the Dominated Convergence Theorem to say that

$$\int f\mathrm{d}\mu = \int \mathrm{Avg}_{F_i} f\mathrm{d}\mu \overset{i\to\infty}{\to} \int \lambda\mathrm{d}\mu = \lambda.$$

$\square$

Remark 3.1.11. The reason we add the caveat that $\mathscr{I} = \mathbb{N}$ to ensure that (iii)$\Rightarrow$(i) in our proof of Lemma 3.1.10 is that there is in general no Dominated Convergence Theorem for arbitrary nets. For an elementary example, let $\mathscr{I} = \mathcal{P}_F([0,1])$ be the net of finite subsets of $[0,1]$, and define for each $i \in \mathscr{I}$, and for each

$S \in \mathscr{I}$, let $f_i \in C(X)$ be a continuous function such that $f_i|_i \equiv 1$ and $\int f_i \mathrm{d}\mu \leq 1/2$, where μ is the Lebesgue probability measure on $[0, 1]$. Then $\lim_i f_i(x) = 1$ for all $x \in [0, 1]$, but $\limsup_i \int f_i \mathrm{d}\mu \leq 1/2$.

The equivalence (i) $\Longleftrightarrow$ (ii) of Lemma 3.1.10 in the case where $G = \mathbb{Z}$, $F_k = \{0, 1, \ldots, k - 1\}$ can be found in (Herman, 1983, Lemme on pg. 487). This result generalizes the classical result of Oxtoby (Oxtoby, 1952, (5.3)) relating unique ergodicity and uniform convergence of temporal averages, as unique ergodicity is equivalent to $\left\{ \int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X) \right\}$ being singleton for all $f \in C(X)$. Since this property will be important for the remainder of this section, we introduce the following definition.

Definition 3.1.12. Let $T : G \curvearrowright X$ be a topological dynamical system, and let $f \in C(X)$. We say that f is T-*Herman* (or simply *Herman*, when T is clear from context) if $\left\{ \int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X) \right\}$ is singleton.

As Lemma 3.1.10 shows, this Herman property is equivalent to a certain definition of uniformity in terms of uniform convergence of ergodic averages. However, since we used the term "uniform function" in Chapter 2 to refer to a function whose ergodic averages converged in an L^∞ norm, whereas Herman functions converge in the uniform norm, we saw fit to distinguish these two terms, keeping the terminology between chapters more consistent.

The following theorem tells us that the best kind of convergence for temporo-spatial differentiations can be characterized in terms of ergodic optimization.

Theorem 3.1.13. *Let $T : G \curvearrowright X$ be a topological dynamical system, where G is amenable. Let $\mu \in \mathcal{M}_T(X)$ be a T-invariant Borel probability measure on X. Let $(F_i)_{i \in \mathscr{I}}$ be a Følner net for G, and let $f \in C(X)$. Then the following conditions are related by the implications (1)$\Rightarrow$(2)$\Rightarrow$(3), and if μ is fully supported on X, then (3)$\Rightarrow$(1).*

1. *f is Herman.*

2. *For every net $(C_i)_{i \in \mathscr{I}}$ of Borel-measurable sets C_i of positive measure, the net*

$$\left(\alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) \right)_{i \in \mathscr{I}}$$

   converges to $\int f \mathrm{d}\mu$.

3. *For every net $(U_i)_{i \in \mathscr{I}}$ of open sets U_i of positive measure, the net*

$$\left(\alpha_{U_i} \left(\mathrm{Avg}_{F_i} f \right) \right)_{i \in \mathscr{I}}$$

converges to $\int f\mathrm{d}\mu$.

Proof. (1)$\Rightarrow$(2) Suppose that f is Herman, and let $(F_i)_{i\in\mathscr{I}}$ be a Følner net for G. Then by Lemma 3.1.10, the net $\left(\mathrm{Avg}_{F_i} f\right)_{i\in\mathscr{I}}$ converges in $C(X)$-norm to $\int f\mathrm{d}\mu$, and since $\|\cdot\|_\infty \le \|\cdot\|_{C(X)}$, it follows that $\mathrm{Avg}_{F_i} f \to \int f\mathrm{d}\mu$ in $L^\infty(X, \mu)$. Therefore (1)$\Rightarrow$(2) follows from Theorem 3.1.9.

(2)$\Rightarrow$(3) Trivial.

$\neg$(1)$\Rightarrow \neg$(3) Suppose that μ is fully supported. For this direction, we can assume that f is real-valued, since otherwise we can break f into its real and imaginary parts and consider those parts separately. So for the remainder of this proof, we can assume that f is real-valued.

Suppose that f is not Herman, and that μ is strictly positive. Set

$$m_1 = \min\left\{\int f\mathrm{d}\nu : \nu \in \mathcal{M}_T(X)\right\},$$
$$m_2 = \max\left\{\int f\mathrm{d}\nu : \nu \in \mathcal{M}_T(X)\right\}.$$

If $\left\{\int f\mathrm{d}\mu : \mathcal{M}_T(X)\right\}$ is not singleton, then $m_1 < m_2$, and in particular this tells us that at least one of the inequalities $\int f\mathrm{d}\mu < m_2$, $\int f\mathrm{d}\mu > m_1$ is true. We consider two cases:

Case (i): Consider the case where $m_2 > \int f\mathrm{d}\mu$. Set $g = f + \|f\|_{C(X)}$, which is a nonnegative-valued function with

$$\Gamma(g) = m_2 + \|f\|_{C(X)} > \int f\mathrm{d}\mu + \|f\|_{C(X)} = \int g\mathrm{d}\mu.$$

Choose $L \in \left(\int g\mathrm{d}\mu, \Gamma(g)\right)$. For each $i \in \mathscr{I}$, set

$$U_i = \begin{cases} \{x \in X : \mathrm{Avg}_{F_i} g(x) > L\} & \text{if } \left\|\mathrm{Avg}_{F_i} g\right\|_{C(X)} > L, \\ X & \text{if } \left\|\mathrm{Avg}_{F_i} g\right\|_{C(X)} \le L. \end{cases}$$

Because each U_i is a nonempty open set, and μ is fully supported, we know that each U_i has positive measure. By Theorem 3.1.5, we know that

$$\lim_i \left\|\mathrm{Avg}_{F_i} g\right\| = \Gamma(g) > L > \int g\mathrm{d}\mu.$$

Thus $\limsup_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} g\right) \geq L > \int g \mathrm{d}\mu = \int f \mathrm{d}\mu + \|f\|_{C(X)}$, so

$$\limsup_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} f\right) = \limsup_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} g - \|f\|_{C(X)}\right) > \int f \mathrm{d}\mu.$$

Case (ii): Suppose $m_1 < \int f \mathrm{d}\mu$. Consider $g = \|f\|_{C(X)} - f$, a nonnegative-valued function. Then for $\nu \in \mathcal{M}_T(X)$, we have

$$\int g \mathrm{d}\nu = \|f\|_{C(X)} - \int f \mathrm{d}\nu$$
$$\Rightarrow \Gamma(g) = \|f\|_{C(X)} - m_1$$
$$> \|f\|_{C(X)} - \int f \mathrm{d}\mu$$
$$= \int g \mathrm{d}\mu.$$

Choose $L \in (\int g \mathrm{d}\mu, \Gamma(g))$. Construct open subsets U_i of X by

$$U_i = \begin{cases} \{x \in X : \mathrm{Avg}_{F_i} g(x) > L\} & \text{if } \left\|\mathrm{Avg}_{F_i} g\right\| > L, \\ X & \text{if } \left\|\mathrm{Avg}_{F_i} g\right\| \leq L. \end{cases}$$

Then by a similar argument to that used in Case (i), we know that $\limsup_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} g\right) \geq L > \int g \mathrm{d}\mu$. It then follows that

$$\liminf_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} f\right) = \liminf_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} \left(\|f\|_{C(X)} - g\right)\right)$$
$$= \|f\|_{C(X)} - \limsup_i \alpha_{U_i} (\mathrm{Avg}_{F_i} g)$$
$$< \|f\|_{C(X)} - \int g \mathrm{d}\mu$$
$$= \int f \mathrm{d}\mu.$$

$\square$

We now come to a theorem which provides a qualitative connection between unique ergodicity and temporo-spatial differentiation problems.

Theorem 3.1.14. *Let $T : G \curvearrowright X$ be a topological dynamical system, where G is amenable. Let $\mu \in \mathcal{M}_T(X)$ be a T-invariant Borel probability on X. Let $(F_i)_{i \in \mathscr{I}}$ be a Følner net for G. Then the following conditions are related by the implications (1)$\Rightarrow$(2)$\Rightarrow$(3), and if μ is fully supported on X, then (3)$\Rightarrow$(1).*

1. *$T : G \curvearrowright X$ is uniquely ergodic.*

2. *For every net $(C_i)_{i \in \mathscr{I}}$ of Borel-measurable sets C_i of positive measure, the net*

$$\left(\alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) \right)_{i \in \mathscr{I}}$$

   converges to $\int f \mathrm{d}\mu$ for all $f \in C(X)$.

3. *For every net $(U_i)_{i \in \mathscr{I}}$ of open sets U_i of positive measure, the net*

$$\left(\alpha_{U_i} \left(\mathrm{Avg}_{F_i} f \right) \right)_{i \in \mathscr{I}}$$

   converges to $\int f \mathrm{d}\mu$ for all $f \in C(X)$.

Proof. (1)$\Rightarrow$(2) The unique ergodicity of $T : G \curvearrowright X$ is equivalent to every $f \in C(X)$ being Herman. Apply Theorem 3.1.13.

(2)$\Rightarrow$(3) Trivial.

$\neg$(1)$\Rightarrow \neg$(3) Suppose that $T : G \curvearrowright X$ is not uniquely ergodic, and that μ is strictly positive. By Theorem 3.1.6, there exists a nonnegative-valued $f \in C_{\mathbb{R}}(X)$ such that $\int f \mathrm{d}\mu < \Gamma(f) = \lim_i \left\| \mathrm{Avg}_{F_i} f \right\|$. Let $L \in \left(\int f \mathrm{d}\mu, \Gamma(f) \right)$. For each $i \in \mathscr{I}$, set

$$U_i = \begin{cases} \left\{ x \in X : \mathrm{Avg}_{F_i} f(x) > L \right\} & \text{if } \left\| \mathrm{Avg}_{F_i} f \right\| > L, \\ X & \text{if } \left\| \mathrm{Avg}_{F_i} f \right\| \leq L. \end{cases}$$

Because each U_i is a nonempty open set, and μ is fully supported, we know that each U_i has positive measure. Thus $\limsup_i \alpha_{U_i} \left(\mathrm{Avg}_{F_i} f \right) \geq L > \int f \mathrm{d}\mu$. $\qquad \square$

In the event that we're dealing not just with a Følner net, but instead a Følner *sequence*, we can make a stronger claim: that unique ergodicity is equivalent to all the temporo-spatial differentiations of continuous functions by that temporal averaging sequence converging.

Theorem 3.1.15. *Let $T : G \curvearrowright X$ be a topological dynamical system, where G is amenable. Let $\mu \in \mathcal{M}_T(X)$ be a T-invariant Borel probability measure on X, and let $(F_k)_{k=1}^\infty$ be a Følner sequence. Let $f \in C(X)$. Then the following conditions are related by the implications (1)$\Rightarrow$(2)$\Rightarrow$(3)$\Rightarrow$(4), and if μ is fully supported on X, then (4)$\Rightarrow$(1).*

1. *f is Herman.*

2. *For every sequence $(C_k)_{k=1}^\infty$ of Borel-measurable sets C_k of positive measure, the sequence*

$$\left(\alpha_{C_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$$

   converges to $\int f \mathrm{d}\mu$.

3. *For every sequence $(U_k)_{k=1}^\infty$ of open sets U_k of positive measure, the sequence*

$$\left(\alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$$

   converges to $\int f \mathrm{d}\mu$.

4. *For every sequence $(U_k)_{k=1}^\infty$ of open sets U_k of positive measure, the sequence*

$$\left(\alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$$

   converges to some complex number.

Furthermore, if $\left\{\int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X)\right\}$ is not singleton, the measure μ is fully supported, and the space (X, μ) is atomless, then we can choose a sequence $(U_k')_{k=1}^\infty$ of open subsets of X with positive measure and a continuous $f \in C(X)$ such that $\left(\alpha_{U_k'}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$ diverges and $\mu\left(U_k'\right) \searrow 0$.

Proof. That (1)$\Rightarrow$(2)$\Rightarrow$(3) follows immediately from Theorem 3.1.13, and (3)$\Rightarrow$(4) is trivial. Now we'll show that if μ is fully supported, then $\neg$(1)$\Rightarrow \neg$(4). Suppose that f is *not* Herman. We can consider the case where f is real-valued, since otherwise we can break f into its real and imaginary parts. Moreover, we can assume that f is nonnegative-valued, since otherwise we can just replace f with $f + \|f\|_{C(X)}$. So for the remainder of this proof, we assume that f is nonnegative-valued.

Set

$$m_1 = \min\left\{ \int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X) \right\},$$

$$m_2 = \max\left\{ \int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X) \right\}.$$

If $\left\{ \int f \mathrm{d}\mu : \mathcal{M}_T(X) \right\}$ is not singleton, then $m_1 < m_2$, and in particular this tells us that at least one of the inequalities $\int f \mathrm{d}\mu < m_2$, $\int f \mathrm{d}\mu > m_1$ is true.

Case (i): Consider first the case where $m_2 > \int f \mathrm{d}\mu$. Choose $L, M \in \mathbb{R}$ such that $\int f \mathrm{d}\mu < L < M < \Gamma(f)$. Define open sets $V_k, W_k \subseteq X$ for $k \in \mathbb{N}$ by

$$V_k = \left\{ x \in X : \mathrm{Avg}_{F_k} f(x) > M \right\},$$

$$W_k = \left\{ x \in X : \mathrm{Avg}_{F_k} f(x) < L \right\}.$$

Both sets are obviously open, since they're preimages of open subsets of $\mathbb{R}$ under the continuous functions $\mathrm{Avg}_{F_k} f \in C_{\mathbb{R}}(X)$.

First, we know that there exists $K \in \mathbb{N}$ such that $V_k \neq \emptyset$ for all $k \geq K$. This is because we know there exists $K \in \mathbb{N}$ in X such that $\left\| \mathrm{Avg}_{F_k} f \right\|_{C(X)} > M$ for all $k \geq K$, and by the Extreme Value Theorem, we know there exists a sequence $(x_k)_{k=1}^{\infty}$ such that

$$\left\| \mathrm{Avg}_{F_k} f(x) \right\|_{C(X)} = \mathrm{Avg}_{F_k} f(x_k)$$

for all $k \in \mathbb{N}$. In particular, if $k \geq K$, then $x_k \in V_k$. Therefore $V_k \neq \emptyset$ for all $k \geq K$.

Secondly, we claim that W_k is nonempty for all $k \in \mathbb{N}$. To see this, suppose to the contrary that $W_k = \emptyset$ for some $k \in \mathbb{N}$. Then $f(x) \geq L > \int f \mathrm{d}\mu$ for all $x \in X$, meaning that $\int f \mathrm{d}\mu \geq L > \int f \mathrm{d}\mu$, a clear contradiction. So $W_k \neq \emptyset$ for all $k \in \mathbb{N}$.

Now, define a sequence $(U_k)_{k=1}^{\infty}$ of nonempty open subsets of X by

$$U_k = \begin{cases} X & \text{if } k < K, \\ V_k & \text{if } k \geq K \text{ is odd,} \\ W_k & \text{if } k \geq K \text{ is even.} \end{cases}$$

Then

$$
\limsup_{k\to\infty} \alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right) \geq \limsup_{k\to\infty} \alpha_{U_{2k+1}}\left(\mathrm{Avg}_{F_{2k+1}} f\right)
$$

$$
= \limsup_{k\to\infty} \alpha_{V_{2k+1}}\left(\mathrm{Avg}_{F_{2k+1}} f\right)
$$

$$
\geq \limsup_{k\to\infty} \alpha_{V_{2k+1}}(M)
$$

$$
= M,
$$

$$
\liminf_{k\to\infty} \alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right) \leq \liminf_{k\to\infty} \alpha_{U_{2k}}\left(\mathrm{Avg}_{F_{2k}} f\right)
$$

$$
= \liminf_{k\to\infty} \alpha_{W_{2k}}\left(\mathrm{Avg}_{F_{2k}} f\right)
$$

$$
\leq \liminf_{k\to\infty} \alpha_{W_{2k}}(L)
$$

$$
= L.
$$

Therefore

$$
\liminf_{k\to\infty} \alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right) \leq L < M \leq \limsup_{k\to\infty} \alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right),
$$

meaning the sequence diverges.

Case (ii): Consider now the case where $m_1 < \int f \mathrm{d}\mu$. Replacing f with $f' = \|f\|_{C(X)} - f$, another nonnegative-valued continuous function, we see that

$$
\int f' \mathrm{d}\mu = \|f\|_{C(X)} - \int f \mathrm{d}\mu < \|f\|_{C(X)} - m_1 = \max\left\{\int f' \mathrm{d}\nu : \nu \in \mathcal{M}_T(X)\right\}.
$$

We can now carry out the construction from Case (i) on f' instead of f to get a sequence $(U_k)_{k=1}^{\infty}$ of open sets along which $\left(\alpha_{U_k}\left(\mathrm{Avg}_{F_k} f'\right)\right)_{k=1}^{\infty}$ diverges, and thus along which $\left(\alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^{\infty}$ diverges.

Furthermore, if in addition, we assume that (X, μ) is atomless, then we can replace our U_k with subsets U'_k such that $\mu\left(U'_k\right) \searrow 0$. This can be done by recursively constructing a sequence of balls U'_k contained in U_k with sufficiently small radius that $\mu\left(U'_{k+1}\right) \leq \min\left\{\mu\left(U'_k\right), 1/k\right\}$ for all $k \in \mathbb{N}$. This is possible by virtue of the atomlessness of (X, μ), since $0 = \mu(\{y_k\}) = \lim_{n\to\infty} \mu(B(y_k, 1/n))$. The above calculation will proceed the same way with the U_k replaced by U'_k. $\qquad\square$

Theorem 3.1.16. *Let* $T : G \curvearrowright X$ *be a topological dynamical system, where* G *is amenable. Let* $\mu \in \mathcal{M}_T(X)$ *be a* T-*invariant Borel probability on* X, *and let* $(F_k)_{k=1}^\infty$ *be a Følner sequence. Then the following conditions are related by the implications (1)$\Rightarrow$(2)$\Rightarrow$(3)$\Rightarrow$(4), and if* μ *is fully supported on* X, *then (4)$\Rightarrow$(1).*

1. $T : G \curvearrowright X$ *is uniquely ergodic.*

2. *For every sequence* $(C_k)_{k=1}^\infty$ *of Borel-measurable sets* C_k *of positive measure, the sequence*

$$\left(\alpha_{C_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$$

   converges to $\int f \mathrm{d}\mu$ *for all* $f \in C(X)$.

3. *For every sequence* $(U_k)_{k=1}^\infty$ *of open sets* U_k *of positive measure, the sequence*

$$\left(\alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$$

   converges to $\int f \mathrm{d}\mu$ *for all* $f \in C(X)$.

4. *For every sequence* $(U_k)_{k=1}^\infty$ *of open sets* U_k *of positive measure, the sequence*

$$\left(\alpha_{U_k}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$$

   converges to some complex number for all $f \in C(X)$.

Furthermore, if (X, T) *is not uniquely ergodic, the measure* μ *is fully supported, and the space* (X, μ) *is atomless, then we can choose a sequence* $(U_k')_{k=1}^\infty$ *of open subsets of* X *with positive measure and a continuous* $f \in C(X)$ *such that* $\left(\alpha_{U_k'}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty$ *diverges and* $\mu\left(U_k'\right) \searrow 0$.

Proof. That (1)$\Rightarrow$(2)$\Rightarrow$(3) follows immediately from Theorem 3.1.14, and (3)$\Rightarrow$(4) is trivial. Now we'll show that if μ is fully supported, then $\neg$(1)$\Rightarrow \neg$(4). Suppose that $T : G \curvearrowright X$ is *not* uniquely ergodic. By Theorem 3.1.5, there exists a nonnegative-valued $f \in C_{\mathbb{R}}(X)$ such that $\int f \mathrm{d}\mu < \Gamma(f) = \lim_{k\to\infty} \left\|\mathrm{Avg}_{F_k} f\right\|$, i.e. for which $\left\{\int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X)\right\}$ is non-singleton. Appeal to Theorem 3.1.15. $\qquad\square$

Remark 3.1.17. Our Theorem 3.1.15 generalizes Theorem 2.1.10. We thank Benjamin Weiss for pointing out that connectedness was not necessary for that result.

3.2 Special cases of temporo-spatial differentiation problems

We digress here quickly to consider certain special classes of temporo-spatial differentiation problems: where the sequence of spatial averaging sets are constant, and where the spatial averaging sets have measure going to 1.

Proposition 3.2.1. *Let $T : G \curvearrowright (X, \mu)$ be a measure-preserving action of a discrete group G on a probability space (X, μ). Let $f \in L^1(X, \mu)$, and let $(F_k)_{k=1}^\infty$ be a sequence of nonempty finite subsets of G such that the sequence $\left(\mathrm{Avg}_{F_k} f\right)$ converges to a function $f^* \in L^1(X, \mu)$ in the weak topology on $L^1(X, \mu)$. Then for every measurable subset C of X with positive measure, we have*

$$\alpha_C \left(\mathrm{Avg}_{F_k} f\right) \overset{k \to \infty}{\Rightarrow} \alpha_C (f^*).$$

Proof. We know $\mu(C)^{-1} \chi_C \in L^\infty(X, \mu) = \left(L^1(X, \mu)\right)'$, so

$$\alpha_C \left(\mathrm{Avg}_{F_k} f\right) = \int \mu(C)^{-1} \chi_C \, \mathrm{Avg}_{F_k} f \mathrm{d}\mu$$

$$= \left\langle \mathrm{Avg}_{F_k} f, \mu(C)^{-1} \chi_C \right\rangle$$

$$\left[\mathrm{Avg}_{F_k} f \overset{k \to \infty}{\Rightarrow} f^* \text{ in the weak topology on } L^1(X, \mu)\right] \overset{k \to \infty}{\Rightarrow} \left\langle f^*, \mu(C)^{-1} \chi_C \right\rangle$$

$$= \alpha_C (f^*).$$

$\square$

We note that Proposition 3.2.1 is exceptional among all our temporo-spatial convergence results to date, in that it can be applied to a function f which is not L^∞, but merely L^1. It also brings us to the following corollary.

Corollary 3.2.2. *Let $T : G \curvearrowright (X, \mu)$ be a measure-preserving action of a discrete amenable group G on a probability space (X, μ). Let $f \in L^1(X, \mu)$, and let $(F_k)_{k=1}^\infty$ be a Følner sequence for G. Then for every measurable subset C of X of positive measure, we have*

$$\alpha_C \left(\mathrm{Avg}_{F_k} f\right) \overset{k \to \infty}{\Rightarrow} \alpha_C (f^*),$$

where f^* is the projection of f onto the subspace of invariant functions in $L^1(X, \mu)$.

Proof. This is a corollary of Proposition 3.2.1 and the Mean Ergodic Theorem for actions of amenable groups (Kerr and Li, 2016, Theorem 4.23), since the norm topology on L^1 is stronger than the weak topology. □

Proposition 3.2.3. *Let $T : G \curvearrowright (X, \mu)$ be a measure-preserving action of a discrete group G on a probability space (X, μ), and let $(C_k)_{k=1}^\infty$ be a sequence of measurable subsets of X such that $\mu(C_k) \to 1$. Let $(F_k)_{k=1}^\infty$ be a sequence of nonempty finite subsets of G. Then*

$$\lim_{k \to \infty} \alpha_{C_k} \left(\operatorname{Avg}_{F_k} f \right) = \int f \mathrm{d}\mu$$

for all $f \in L^\infty(X, \mu)$.

Proof. Fix $f \in L^\infty(X, \mu)$. Then

$$
\begin{aligned}
&\left| \int f \mathrm{d}\mu - \alpha_{C_k} \left(\operatorname{Avg}_{F_k} f \right) \right| \\
&= \left| \int_X \operatorname{Avg}_{F_k} \mathrm{d}\mu - \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu + \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu - \mu(C_k)^{-1} \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu \right| \\
&\leq \left| \int_X \operatorname{Avg}_{F_k} f \mathrm{d}\mu - \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu \right| + \left| \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu - \mu(C_k)^{-1} \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu \right| \\
&= \left| \int_{X \setminus C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu \right| + \left(1 - \mu(C_k)^{-1} \right) \left| \int_{C_k} \operatorname{Avg}_{F_k} f \mathrm{d}\mu \right| \\
&\leq (1 - \mu(C_k)) \|f\|_\infty + \left(1 - \mu(C_k)^{-1} \right) \mu(C_k) \|f\|_\infty \\
&\xrightarrow{k \to \infty} \infty.
\end{aligned}
$$

□

In light of Proposition 3.2.3, we can see that temporo-spatial differentiation problems become trivial in the case where $\mu(C_k) \to 1$ for $f \in L^\infty(X, \mu)$. The result, however, fails for any unbounded integrable function. Let $f \in L^1(X, \mu) \setminus L^\infty(X, \mu)$, i.e. an unbounded integrable function, and let $E_k = \{x \in X : |f(x)| \geq k\}$ for all $k \in \mathbb{N}$. Then by Chebyshev's inequality, it follows that $0 < \mu(E_k) \leq k^{-1} \|f\|_1$ for all $k \in \mathbb{N}$, and

$\mu(E_k) \searrow 0$. Then if $T : G \curvearrowright (X, \mu)$ is the trivial action, i.e. $T_g = \mathrm{id}_X$ for all $g \in G$, then

$$\alpha_{E_k} \left(\mathrm{Avg}_{F_k} |f| \right) = \alpha_{E_k}(|f|)$$

$$\geq k$$

$$\Rightarrow \alpha_{E_k} \left(\mathrm{Avg}_{F_k} |f| \right) \overset{k \to \infty}{\to} +\infty.$$

Based on this example, we can see that in contrast with Proposition 3.2.1, there's no hope for improving Proposition 3.2.3 to even the case where $f \in L^{\infty-}(X, \mu) = \bigcap_{p \in [1,\infty)} L^p(X, \mu)$.

3.3 Temporo-spatial differentiation theorems around sets of rapidly vanishing diameter

In this section, we'll be concerned with the following general setup and question: Let (X, p) be a compact pseudometric space, and let $T : G \curvearrowright X$ be a continuous action of a locally compact group G on X which preserves a Borel probability measure μ on X. Now fix some point $x_0 \in X$, and consider a net of positive-measure subsets C_i of X containing x_0. When will the temporo-spatial derivative relative to C_i (and some averaging net F_i) resemble the pointwise temporal average at x_0? Theorem 3.3.2 establishes a powerful sufficient condition: If $f : X \to \mathbb{C}$ is uniformly continuous and bounded, and the diameter of the elements of the net C_i go to 0 sufficiently fast, then we'll have that $\left(\mathrm{Avg}_{F_i} f \right)(x_0) \approx \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right)$, where "sufficiently fast" depends upon the (pseudo)metric properties of the continuous action, the averaging net $(F_i)_{i \in \mathscr{I}}$, and the point x_0. In this situation, we can reduce the temporo-spatial problem to a problem of taking a pointwise ergodic average. We then consider cases where narrowing our focus (e.g. considering Hölder actions instead of general continuous actions) allow us to improve the diameter decay rate. We then move on to make statements about the "probabilistically generic" behavior of these temporo-spatial derivatives by appealing to pointwise convergence results from ergodic theory. Finally, we extend this pointwise reduction to the setting of nonconventional ergodic averages with Theorem 3.3.12.

Several results in this section will be quite general in their statement, and as such will sometimes require a number of additional hypotheses that are satisfied automatically in many reasonable situations. We make notes after the proofs of some results to note that certain hypotheses stated explicitly in the results in question are satisfied a priori in certain reasonable cases.

Our first result of this section describes a sufficient condition for the temporo-spatial averages to reduce to pointwise averages.

Lemma 3.3.1. *Let (X, p) be a compact pseudometric space, and let $T : G \curvearrowright X$ be a continuous action of a locally compact topological group G (not necessarily amenable) on X. Let μ be a regular Borel probability measure on X. Fix a point $x_0 \in X$.*

Let $(F_i)_{i \in \mathscr{I}}$ be a net of compact subsets of G with positive Haar measure m. Let $(C_i)_{i \in \mathscr{I}}$ be a net of measurable subsets of X such that $\mu(C_i) > 0$ and $x_0 \in C_i$ for all $i \in \mathscr{I}$. Suppose that for every $\delta > 0$, there exists a net $(A_i)_{i \in \mathscr{I}}$ of measurable subsets of G such that

$$A_i \subseteq \{g \in F_i : \mathrm{diam}(C_i) \leq \delta\}, \qquad\qquad (\forall i \in \mathscr{I})$$

$$\lim_i \frac{m(A_i)}{m(F_i)} = 1.$$

Let $f : X \to \mathbb{C}$ be a continuous function. Then

$$\lim_i \left| \left(\mathrm{Avg}_{F_i} f \right) (x_0) - \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) \right| = 0.$$

Proof. Fix $\epsilon > 0$. Since f is uniformly continuous (by Lemma 2.3.1), there exists $\delta > 0$ such that if $y_1, y_2 \in X$, and $p(y_1, y_2) \leq \delta$, then $|f(y_1) - f(y_2)| \leq \frac{\epsilon}{2\lambda}$. Let $(A_i)_{i \in \mathscr{I}}$ be as in the lemma statement, and set $B_i = F_i \setminus A_i$ for all $i \in \mathscr{I}$, so $\lim_i m(B_i)/m(F_i) = 0$.

Now, we estimate

$$\left| \left(\mathrm{Avg}_{F_i} f \right) (x_0) - \alpha_{C_i}(\mathrm{Avg}_{F_i} f) \right|$$

$$= \left| \alpha_{C_i} \left(\left(\mathrm{Avg}_{F_i} f \right) (x_0) - \mathrm{Avg}_{F_i} f \right) \right|$$

$$= \left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{F_i} \left(f(T_g x_0) - (f \circ T_g) \right) \right) \mathrm{d}m(g) \right|$$

$$\leq \left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{A_i} \left(f(T_g x_0) - (f \circ T_g) \right) \mathrm{d}m(g) \right) \right|$$

$$+ \left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{B_i} \left(f(T_g x_0) - (f \circ T_g) \right) \mathrm{d}m(g) \right) \right|$$

74

Our goal now is to bound both

$$\left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{A_i} (f(T_g x_0) - (f \circ T_g)) \mathrm{d}m(g) \right) \right|,$$

$$\left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{B_i} (f(T_g x_0) - (f \circ T_g)) \mathrm{d}m(g) \right) \right|$$

by $\frac{\epsilon}{2}$.

First, we estimate the term $\left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{A_i} (f(T_g x_0) - (f \circ T_g)) \mathrm{d}m(g) \right) \right|$. We see that if $g \in A_i$, then

$$\left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{A_i} (f(T_g x_0) - (f \circ T_g)) \mathrm{d}m(g) \right) \right|$$

$$= \left| \frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{A_i} (f(T_g x_0) - f(T_g x)) \mathrm{d}m(g) \mathrm{d}\mu(x) \right|$$

$$\leq \frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{A_i} |f(T_g x_0) - f(T_g x)| \, \mathrm{d}m(g) \mathrm{d}\mu(x)$$

But if $y \in T_g C_i$, and $D_{x_0}(g, \mathrm{diam}(C_i)) \leq \delta$, then $x_0, T_{g^{-1}} y \in C_i$, meaning that

$$\rho(T_g x_0, y) = \rho(T_g x_0, T_g(T_{g^{-1}} y)) \leq D_{x_0}(g, T_{g^{-1}} y) \leq D_{x_0}(g, \mathrm{diam}(C_i)) \leq \delta.$$

Therefore

$$\frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{A_i} |f(T_g x_0) - f(T_g x)| \, \mathrm{d}m(g) \mathrm{d}\mu(x) \leq \frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{A_i} \frac{\epsilon}{2} \mathrm{d}m(g) \mathrm{d}\mu(x)$$

$$= \frac{\epsilon}{2}.$$

Now, we bound the term $\left| \alpha_{C_i} \left(\frac{1}{m(F_i)} \int_{B_i} (f(T_g x_0) - (f \circ T_g)) \mathrm{d}m(g) \right) \right|$. By estimates similar to those used to approximate the former term, we have that

$$\left| \frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{B_i} (f(T_g x_0) - f(T_g x)) \mathrm{d}m(g) \mathrm{d}\mu(x) \right|$$

$$\leq \frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{B_i} |f(T_g x_0) - f(T_g x)| \, \mathrm{d}m(g) \mathrm{d}\mu(x)$$

$$\leq \frac{1}{\mu(C_i)} \int_{C_i} \frac{1}{m(F_i)} \int_{B_i} (2\|f\|_u) \, \mathrm{d}m(g) \mathrm{d}\mu(x)$$

Choose $I \in \mathscr{I}$ such that if $i \geq I$, then $\frac{m(B_i)}{m(F_i)} \leq \frac{\epsilon}{4 \max\{1, \|f\|_u\}}$. Then if $i \geq I$, then $\frac{m(B_i)}{m(F_i)} (2\|f\|_u) \leq \frac{\epsilon}{2}$.

Therefore, if $i \geq I$, then $\left|\left(\mathrm{Avg}_{F_i} f\right)(x_0) - \alpha_{C_i}(\mathrm{Avg}_{F_i} f)\right| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. $\qquad\square$

We have stated Lemma 3.3.1 for *pseudo*metric spaces, rather than just metric spaces. In Chapter 2, we found that looking at certain pseudometric spaces helped us to establish convergence results for certain temporo-spatial averages. For example, Proposition 2.3.2 was useful in proving Theorem 2.3.5. For this reason, we state several results of this section in terms of compact pseudometric spaces.

We also observe that Lemma 3.3.1 does not assume that the action $T : G \curvearrowright (X, \mu)$ is measure-preserving, only continuous.

Lemma 3.3.1 as stated is a powerful tool for achieving the kind of reduction to the pointwise setting that we aim for, but we desire still a sufficient condition for the hypotheses of the lemma to attain. The following result states that, under appropriate conditions, we can find an x_0-dependent diameter decay condition on $(C_i)_{i \in \mathscr{I}}$ for this reduction to attain.

Theorem 3.3.2. *Let (X, p) be a compact pseudometric space, and let $T : G \curvearrowright X$ be a continuous action of a locally compact topological group G (not necessarily amenable) on X. Let μ be a regular Borel probability measure on X. Fix a point $x_0 \in X$, and for each $g \in G, r \in (0, \infty)$, let $D_{x_0}(g, r)$ be the value*

$$D_{x_0}(g, r) = \sup \left\{ p(T_g x_0, T_g x) : x \in X, p(x_0, x) \leq r \right\},$$

and assume that $D_{x_0}(\cdot, r) : G \to (0, \infty)$ is measurable for each $r \in (0, \infty)$.

Let $(F_i)_{i \in \mathscr{I}}$ be a net of compact subsets of G with positive Haar measure m. Let $(C_i)_{i \in \mathscr{I}}$ be a net of measurable subsets of X such that $\mu(C_i) > 0$ and $x_0 \in C_i$ for all $i \in \mathscr{I}$. Suppose that for every $\delta > 0$, we have

$$\lim_i \frac{m \left(\{ g \in F_i : D_{x_0}(g, \mathrm{diam}(C_i)) > \delta \} \right)}{m(F_i)} = 0.$$

Let $f : X \to \mathbb{C}$ be a continuous function. Then

$$\lim_i \left| \left(\mathrm{Avg}_{F_i} f \right)(x_0) - \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) \right| = 0.$$

Proof. For each $\delta > 0$, set

$$A_i = \left\{ g \in F_i : D_{x_0}(g, \mathrm{diam}(C_i)) \leq \delta \right\}.$$

The result follows from Lemma 3.3.1. $\qquad\square$

The assumption that $D_{x_0}(\cdot, r) : G \to (0, \infty)$ be a measurable function in G for all $r \in (0, \infty)$, though relevant to make sure the sets A_i in our proof are measurable, is satisfied automatically in the case where G is discrete. Our use of this function D_{x_0} ensures that the condition being imposed is in fact a decay condition on $\mathrm{diam}(C_i)$, in the sense that if $(C_i)_{i \in \mathscr{I}}$ is a net satisfying the condition

$$\lim_i \frac{m\left(\{g \in F_i : D_{x_0}(g, \mathrm{diam}(C_i)) > \delta\}\right)}{m(F_i)} = 0 \qquad (\forall \delta > 0),$$

and $(C_i')_{i \in \mathscr{I}}$ is a net of measurable subsets of X containing x_0 with positive measure, and $\mathrm{diam}\,(C_i') \leq \mathrm{diam}(C_i)$ for all $i \in \mathscr{I}$, then $(C_i')_{i \in \mathscr{I}}$ will also satisfy the condition. However, this decay rate depends on x_0, a shortcoming which can be overcome with some additional conditions on the action T, as will be seen in Theorem 3.3.4.

Definition 3.3.3. Let (X, p) be a pseudometric space (not necessarily compact), and let $T : G \curvearrowright X$ be an action of a group G on X. We call the action *Hölder* if for every $g \in G$ exist $H(g), L(g) \in (0, \infty)$ such that

$$p\left(T_g x, T_g y\right) \leq L(g) \cdot p(x, y)^{H(g)} \qquad (\forall g \in G, x \in X, y \in X).$$

Our next result shows that if we assume that our action is Hölder, and the Hölder parameters of T_g satisfy certain measurability properties as functions of G, then this diameter decay rate can be chosen independent of x_0. We remark now that our statement of the result is quite wordy, with several hypotheses, but as we'll explain shortly, several of these hypotheses are satisfied automatically in many cases.

Theorem 3.3.4. *Let (X, p) be a compact pseudometric space, and let $T : G \curvearrowright X$ be a continuous action of a locally compact topological group G (not necessarily amenable) on X. Let μ be a regular Borel probability measure on X. Assume further that there exist measurable functions $H, L : G \to (0, \infty)$ such that*

$$p\left(T_g x, T_g y\right) \leq L(g) \cdot p(x, y)^{H(g)} \qquad (\forall g \in G, x \in X, y \in X).$$

Let $(F_i)_{i \in \mathscr{I}}$ be a net of compact subsets of G with positive Haar measure m. Let $(C_i)_{i \in \mathscr{I}}$ be a net of measurable subsets of X such that $\mu(C_i) > 0$ and $x \in C_i$ for all $i \in \mathscr{I}$. Suppose that for every $\delta > 0$, we have

$$\lim_i \frac{m\left(\{g \in F_i : L(g) \cdot \mathrm{diam}(C_i)^{H(g)} > \delta\}\right)}{m(F_i)} = 0.$$

Let $x_0 \in X$ be a point in X, and let $f : X \to \mathbb{C}$ be a uniformly bounded continuous function. Then

$$\lim_i \left| \left(\mathrm{Avg}_{F_i} f \right)(x_0) - \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) \right| = 0.$$

Proof. We first observe that if $p(T_g x, T_g y) \le L(g) p(x, y)^{H(g)}$, then $D_{x_0}(g, r) \le L(g) r^{H(g)}$ for all $x_0 \in X$, so $\mathrm{diam}(T_g C_i) \le L(g) \cdot \mathrm{diam}(C_i)^{H(g)}$. Given $\delta > 0$, set

$$A_i = \left\{ g \in F_i : L(g) \cdot \mathrm{diam}(C_i)^{H(g)} \le \delta \right\},$$

and apply Lemma 3.3.1. $\qquad\square$

Remark 3.3.5. • If G is discrete, then the measurability assumptions on H, L are automatically fulfilled.

  • If T_g is Lipschitz for all $g \in G$, then we can take H to be the constant function 1. This is the case in particular if T is an action on a compact Riemannian manifold X by diffeomorphisms.

  • In the special case where $G = \mathbb{Z}$, if both T_1, T_{-1} are Hölder with exponent α_0 and coefficient L_0, then for $n \ge 0$, we can take $H(n) = \alpha_0^{|n|}, L(n) = L_0^{|n|}$.

  • If G acts by isometries, then we can take H, L to both be the constants 1.

Theorem 3.3.4 says that given a Hölder action T of a group G (subject to certain measurability conditions) on a compact pseudometric probability space, and an averaging net $(F_i)_{i \in \mathscr{I}}$, there exists a diameter decay rate such that if $(C_i)_{i \in \mathscr{I}}$ is a net of positive-measure sets containing a fixed point x_0, then the temporo-spatial derivative at C_i will resemble the temporal pointwise average. Notably, this decay rate depends only on the averaging net and the Hölder condition on T, and not on the point x_0 or the function f.

Theorem 3.3.4 cannot be called sharp in the strictest sense, since given any net (C_i) satisfying the hypotheses of Theorem 3.3.4, we could replace all the C_i with $C_i \cup E$, where E is some fixed subset of X with positive diameter but measure 0. A truly sharp Theorem 3.3.4 would -at the very least- have to account for a notion of "essential diameter."

Under an additional assumption on the function being averaged, we can provide quantitative estimates on the approximation in Theorem 3.3.4.

Proposition 3.3.6. *Let (X, p) be a compact pseudometric space, and let $T : G \curvearrowright X$ be a continuous action of a locally compact topological group G (not necessarily amenable) on X. Let μ be a regular Borel*

probability measure on X. Assume further that there exist measurable functions $H, L : G \to (0, \infty)$ such that

$$p\left(T_g x, T_g y\right) \leq L(g) \cdot p(x, y)^{H(g)} \qquad\qquad (\forall g \in G, x \in X, y \in X).$$

Let F be a compact subset of G with positive Haar measure m, and let C be a measurable subset of X such that $\mu(C) > 0$.

Let $x_0 \in X$ be a point in X, and let $f : X \to \mathbb{C}$ be a Hölder function with constants c, β for which

$$|f(x) - f(y)| \leq c \cdot \rho(x, y)^\beta \qquad\qquad (\forall x, y \in X).$$

Then

$$\left|\left(\mathrm{Avg}_F f\right)(x_0) - \alpha_C\left(\mathrm{Avg}_F f\right)\right| \leq \frac{c}{m(F)} \int_F L(g)^\beta \cdot \mathrm{diam}(C)^{\beta H(g)} \mathrm{d}m(g).$$

Proof.

$$
\begin{aligned}
&\left|\left(\mathrm{Avg}_F f\right)(x_0) - \alpha_C\left(\mathrm{Avg}_F f\right)\right| \\
&= \left|\alpha_C\left(\frac{1}{m(F)}\int_F \left(f\left(T_g x_0\right) - (f \circ T_g)\right)\mathrm{d}m(g)\right)\right| \\
&= \left|\frac{1}{m(C)}\int_C \frac{1}{m(F)}\int_F \left(f(T_g x_0) - (f(T_g x))\right)\mathrm{d}m(g)\mathrm{d}\mu(x)\right| \\
&\leq \frac{1}{m(C)}\int_C \frac{1}{m(F)}\int_F \left|f\left(T_g x_0\right) - f\left(T_g x\right)\right|\mathrm{d}m(g)\mathrm{d}\mu(x) \\
&\leq \frac{1}{m(C)}\int_C \frac{1}{m(F)}\int_F c \cdot \rho\left(T_g x_0, T_g x\right)^\beta \mathrm{d}m(g)\mathrm{d}\mu(x) \\
&\leq \frac{1}{m(C)}\int_C \frac{1}{m(F)}\int_F c \cdot \left(L(g) \cdot \rho\left(x_0, x\right)^{H(g)}\right)^\beta \mathrm{d}m(g)\mathrm{d}\mu(x) \\
&\leq \frac{1}{m(C)}\int_C \frac{1}{m(F)}\int_F c \cdot \left(L(g) \cdot \mathrm{diam}(C)^{H(g)}\right)^\beta \mathrm{d}m(g)\mathrm{d}\mu(x) \\
&= c\frac{1}{m(C)}\int_C \frac{1}{m(F)}\int_F L(g)^\beta \cdot \mathrm{diam}(C)^{\beta H(g)}\mathrm{d}m(g)\mathrm{d}\mu(x) \\
&= \frac{c}{m(F)}\int_F L(g)^\beta \cdot \mathrm{diam}(C)^{\beta H(g)}\mathrm{d}m(g)
\end{aligned}
$$

$\square$

Our next result takes us in the direction of a "random temporo-spatial differentiation problem," where we consider a temporo-spatial problem in which the spatial averaging net is considered to be chosen "randomly" according to some scheme or constraints.

Corollary 3.3.7. *Let $T : G \curvearrowright X$ be a continuous action of a locally compact topological group G on a compact pseudometric space $X = (X, p)$, and let μ be a regular Borel probability measure on X, and let $(F_i)_{i \in \mathscr{I}}$ be a net in G. Let $H, L : G \to (0, \infty)$ be measurable functions such that*

$$p\left(T_g x, T_g y\right) \leq L(g) \cdot p(x, y)^{H(g)} \qquad\qquad (\forall g \in G, x \in X, y \in X).$$

Suppose that for each $x \in X$, the net $(C_i(x))_{i \in \mathscr{I}}$ is a net of measurable subsets $C_i(x)$ of X containing the point x such that $\mu(C_i(x)) > 0$ for all $x \in X$, as well as that for almost all $x \in X$, we have

$$\lim_i \frac{m\left(\left\{g \in F_i : L(g) \cdot \operatorname{diam}(C_i(x))^{H(g)} > \delta\right\}\right)}{m(F_i)} = 0$$

for all $\delta > 0$. Let $f : X \to \mathbb{C}$ be a continuous function, and suppose that for almost all $x \in X$, we have that $\lim_i \operatorname{Avg}_{F_i} f(x) = f^*(x)$, where f^* is a measurable function $X \to \mathbb{C}$. Then

$$\lim_i \alpha_{C_i(x)}\left(\operatorname{Avg}_{F_i} f\right) = f^*(x)$$

for almost all $x \in X$.

Remark 3.3.8. Corollary 3.3.7 is a tool that turns almost-sure pointwise convergence results from ergodic theory into almost-sure convergence results for classes of random temporo-spatial differentiations. Corollaries 3.3.10 and 3.3.11, corresponding to the Lindenstrauss pointwise ergodic theorem and Bourgain's theorem on pointwise convergence of averages along polynomials, respectively, are special cases of Corollary 3.3.7. In principle, there is a special case of Corollary 3.3.7 corresponding to any result that ensures the almost-sure pointwise convergence of an ergodic average.

Proof of Corollary 3.3.7. Let

$$A = \bigcap_{k=1}^{\infty} \left\{ x \in X : \lim_i \frac{m\left(\{ g \in F_i : L(g) \cdot \mathrm{diam}(C_i(x))^{H(g)} > 1/k \}\right)}{m(F_i)} = 0 \right\},$$

$$B = \left\{ x \in X : \lim_i \mathrm{Avg}_{F_i} f(x) = f^*(x) \right\}.$$

Both A, B are of full measure by hypothesis, and thus so is $A \cap B$. Let $x \in A \cap B$. Then

$$\left| \alpha_{C_i(x)} \left(\mathrm{Avg}_{F_i} f \right) - f^*(x) \right| \leq \left| \alpha_{C_i(x)} \left(\mathrm{Avg}_{F_i} f \right) - \mathrm{Avg}_{F_i} f(x) \right| + \left| \mathrm{Avg}_{F_i} f(x) - f^*(x) \right|$$

$$\overset{i \to \infty}{\to} 0,$$

where the first summand goes to 0 (by Theorem 3.3.4) because $x \in A$ and the second summand goes to 0 because $x \in B$. $\square$

As a rule, results like Theorem 3.3.4 lead naturally to results like Corollary 3.3.7, and we'll see several other examples of this in this chapter. Theorem 3.3.4 provides a sufficient condition for a spatial averaging net $(C_i)_{i \in \mathscr{I}}$ around a point x to induce a temporo-spatial differentiation problem that's reducible to a pointwise temporal problem at that point x; it then follows that if we have some scheme for associating to every point x a spatial averaging net $(C_i(x))_{i \in \mathscr{I}}$ around x, and we know that $\mathrm{Avg}_{F_i} f(x) \to f^*(x)$ almost surely for $f \in C(X)$ continuous, then we have a convergence result for the "random temporo-spatial differentiation problem" $\left(\alpha_{C_i(x)} \left(\mathrm{Avg}_{F_i} f \right) \right)_{i \in \mathscr{I}}$. There will be several other examples of results like Corollary 3.3.7 in various contexts, taking some temporal pointwise reduction result like Theorem 3.3.4 and extrapolating a statement about random temporo-spatial problems.

It should be noted, however, that the convergence in Corollary 3.3.7 will in general be only for almost every $x \in X$, rather than all $x \in X$. If there exists a point $x \in X$ where $\left(\mathrm{Avg}_{F_i} f(x) \right)_{i \in \mathscr{I}}$ does not converge to $f^*(x)$, then Theorem 3.3.4 tells us that $\left(\alpha_{C_i(x)} \left(\mathrm{Avg}_{F_i} f \right) \right)_{i \in \mathscr{I}}$ won't either.

Corollary 3.3.9. *Let $T : G \curvearrowright X$ be an action of a locally compact topological group G on a compact metric space $X = (X, \rho)$ that preserves a Borel probability measure μ on X, and let $(F_i)_{i \in \mathscr{I}}$ be a net. Let $H, L : G \to (0, \infty)$ be measurable functions such that*

$$p\left(T_g x, T_g y \right) \leq L(g) \cdot p(x, y)^{H(g)} \qquad (\forall g \in G, x \in X, y \in X).$$

81

Suppose that for each $x \in X$, the net $(C_i(x))_{i \in \mathscr{I}}$ is a net of measurable subsets $C_i(x)$ of X containing the point x such that $\mu(C_i(x)) > 0$ for all $x \in X$, and

$$\lim_i \frac{m\left(\{g \in F_i : L(g) \cdot \mathrm{diam}(C_i(x))^{H(g)} > \delta\}\right)}{m(F_i)} = 0$$

for almost all $x \in X$. Suppose that for μ-almost all $x \in X$, we have

$$\lim_i \mathrm{Avg}_{F_i} f(x) = \int f \mathrm{d}\mu \qquad\qquad (\forall f \in C(X))$$

Then for almost all $x \in X$, we have

$$\lim_i \alpha_{C_i(x)} \left(\mathrm{Avg}_{F_i} f\right) = \int f \mathrm{d}\mu.$$

Proof. Since X is compact metrizable, it follows that $C(X)$ is separable, so let $\{f_n : n \in \mathbb{N}\}$ be a subset of $C(X)$ with dense span. For each $n \in \mathbb{N}$, set

$$A_n = \left\{x \in X : \lim_i \alpha_{C_i(x)} \left(\mathrm{Avg}_{F_i} f_n\right) = \int f \mathrm{d}\mu\right\}.$$

Each A_n is of full measure.

Let $f \in C(X)$, and fix $N \in \mathbb{N}$. Choose $J_N \in \mathbb{N}$ and a sequence $z_{1,N}, \ldots, z_{J_N,N} \in \mathbb{C}$ such that

$$\lim_{N \to \infty} \left\| f - \sum_{j=1}^{J_N} z_{j,N} f_j \right\|_{C(X)} \leq \frac{1}{3N}.$$

For convenience, set $\phi_N = \sum_{j=1}^{J_N} z_{j,N} f_j$. Then $\left\| \int f \mathrm{d}\mu - \int \phi_N \mathrm{d}\mu \right\|_{L^\infty(X,\mu)} \leq \|f - \phi_N\|_{C(X)} \leq \frac{\epsilon}{3}$.

Now for $j \in \{1, \ldots, J_N\}$, choose $i_{j,N} \in \mathscr{I}$ such that if $i \geq i_{j,N}$, then

$$\left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f_j(x)\right) - \int f \mathrm{d}\mu \right| \leq \frac{1}{3N^2 \max\{z_{j,N}, 1\}}.$$

Choose $I_N \in \mathscr{I}$ such that $I_N \geq i_{j,N}$ for all $j \in \{1, \ldots, J_N\}$, and let $x \in \bigcap_{n \in \mathbb{N}} A_n$. Then if $i \geq I_N$, we have

$$
\begin{aligned}
\left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) - \int f \mathrm{d}\mu \right| &\leq \left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) - \alpha_{C_i} \left(\mathrm{Avg}_{F_i} \phi_N \right) \right| \\
&+ \left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} \phi_N \right) - \int \phi_N \mathrm{d}\mu \right| \\
&+ \left| \int \phi_N \mathrm{d}\mu - \int f \mathrm{d}\mu \right|.
\end{aligned}
$$

We bound each of the three summands by $\frac{1}{3N}$ in turn. Firstly, we can see that

$$
\left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) - \alpha_{C_i} \left(\mathrm{Avg}_{F_i} \phi_N \right) \right| = \left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} (f - \phi_N) \right) \right| \leq \| f - \phi_N \|_{C(X)} \leq \frac{1}{3N},
$$

which addresses the first summand. For the second summand, we have

$$
\begin{aligned}
\left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} \phi_N \right) - \int \phi_N \mathrm{d}\mu \right| &= \left| \sum_{j=1}^{J_N} z_{j,N} \left(\alpha_{C_i} \left(\mathrm{Avg}_{F_i} f_j \right) - \int f_j \mathrm{d}\mu \right) \right| \\
&\leq \sum_{j=1}^{N} |z_{j,N}| \cdot \left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f_j \right) - \int f_j \mathrm{d}\mu \right| \\
&\leq \sum_{j=1}^{N} |z_{j,N}| \frac{1}{3N^2 \max\{|z_{j,N}|, 1\}} \\
&\leq \frac{1}{3N}.
\end{aligned}
$$

Finally, for the third summand, we have that

$$
\left| \int \phi_N \mathrm{d}\mu - \int f \mathrm{d}\mu \right| \leq \| \phi_N - f \|_{C(X)} \leq \frac{1}{3N}.
$$

Taken together, these tell us that for every $N \in \mathbb{N}$, $x \in \bigcap_{n \in \mathbb{N}} A_n$, there exists $I \in \mathscr{I}$ such that if $i \geq I$, then $\left| \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) - \int f \mathrm{d}\mu \right| \leq \frac{1}{N}$. Therefore $\lim_i \alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) = \int f \mathrm{d}\mu$ for all $x \in \bigcap_{n \in \mathbb{N}} A_n$. Since each A_n is of full measure, it follows that their countable intersection $\bigcap_{n=1}^{\infty} A_n$ is of full measure, yielding our desired almost-sure convergence. $\qquad \square$

This result tells us that if to (almost) every $x \in X$ we assign a net $(C_i(x))$ of sets of positive measure with rapidly decaying diameter, and $\mathrm{Avg}_{F_i} f \to \int f \mathrm{d}\mu$ for all $f \in C(X)$ then the "probabilistically generic" behavior is that $\alpha_{C_i} \left(\mathrm{Avg}_{F_i} f \right) \to \int f \mathrm{d}\mu$.

Corollary 3.3.9 encompasses several results from Chapter 2, including Theorem 2.2.1, Theorem 2.2.3, and Corollary 2.3.3. Proposition 2.3.2 can also be recovered from our Corollary 3.3.7. Corollary 3.3.9 is motivated by the desire to find positive convergence results for temporo-spatial differentiations relative to actions of groups other than $\mathbb{Z}$, as well as to find to find positive convergence results for temporo-spatial differentiations relative to averages over other subsequences of $\mathbb{Z}$. Moreover, Corollary 3.3.9 opens the door to temporo-spatial differentiations along subsequences. We present here a few examples.

Corollary 3.3.10. *Let $T : G \curvearrowright X$ be an action of a locally compact amenable topological group G on a compact pseudometric space $X = (X, p)$ that preserves a regular Borel probability measure μ on X, and let $(F_k)_{k \in \mathbb{N}}$ be a tempered Følner sequence for G. Let $H, L : G \to (0, \infty)$ be measurable functions such that*

$$p\left(T_g x, T_g y\right) \leq L(g) \cdot p(x, y)^{H(g)} \qquad\qquad (\forall g \in G, x \in X, y \in X).$$

Suppose that for each $x \in X$, the sequence $(C_k(x))_{k \in \mathbb{N}}$ is a sequence of measurable subsets $C_k(x)$ of X containing the point x such that $\mu(C_k(x)) > 0$ for all $x \in X$, and

$$\lim_{k \to \infty} \frac{m\left(\left\{g \in F_k : L(g) \cdot \mathrm{diam}(C_k(x))^{H(g)} > \delta\right\}\right)}{m(F_k)} = 0$$

for almost all $x \in X$.

Then given $f \in C(X)$, for almost all $x \in X$, we have

$$\lim_{k \to \infty} \alpha_{C_k(x)} \left(\mathrm{Avg}_{F_k} f \right) = \mathbb{E}f(x),$$

where $\mathbb{E}$ is the projection onto the space of T-invariant functions in $L^\infty(X, \mu)$.

Proof. The Lindenstrauss Ergodic Theorem (Lindenstrauss, 2001, Theorem 3.3) tells us that $\mathrm{Avg}_{F_k} f \to \mathbb{E}f$ almost surely, so we can apply Corollary 3.3.9. $\qquad\square$

Corollary 3.3.11. *Let $P \in \mathbb{R}[t]$ be a polynomial with real coefficients, and let $T : \mathbb{Z} \curvearrowright X$ be an action of $\mathbb{Z}$ on a compact pseudometric space $X = (X, p)$ that preserves a regular Borel probability measure μ on X.*

Let $F_k = \{\lfloor P(1) \rfloor, \lfloor P(2) \rfloor, \ldots, \lfloor P(k) \rfloor\}$ for all $k \in \mathbb{N}$, and let $H, L : G \to (0, \infty)$ be functions such that

$$p\left(T_n x, T_n y\right) \leq L(n) \cdot p(x, y)^{H(n)} \qquad\qquad (\forall n \in \mathbb{Z}, x \in X, y \in X).$$

Suppose that for each $x \in X$, the sequence $(C_k(x))_{k \in \mathbb{N}}$ is a sequence of measurable subsets $C_k(x)$ of X containing the point x such that $\mu(C_k(x)) > 0$ for all $x \in X$, and

$$\lim_{k \to \infty} \frac{m\left(\{n \in F_k : L(n) \cdot \mathrm{diam}(C_k(x))^{H(n)} > \delta\}\right)}{m(F_k)} = 0$$

for almost all $x \in X$. Let $f \in C(X)$. Then there exists a function $f^* \in L^\infty(X, \mu)$ such that for almost all $x \in X$, we have

$$\lim_{k \to \infty} \alpha_{C_k(x)}\left(\mathrm{Avg}_{F_k} f\right) = f^*(x).$$

Proof. By (Bourgain, 1989, Theorem 2), there exists $f^* \in L^\infty(X, \mu)$ such that $\mathrm{Avg}_{F_k} f(x) \to f^*(x)$ almost surely. Apply Corollary 3.3.7. $\qquad\square$

Finally, we remark that a form of the pointwise reduction in Theorem 3.3.2 can be recovered in the context of nonconventional ergodic averages. In order to make the statement of this result a bit more readable, we use slightly different notation for the remainder of this section than we used in previous parts of this article, using T_ℓ to refer to an ℓth homeomorphism, rather than an action of the integer $\ell \in \mathbb{Z}$.

Theorem 3.3.12. *Let (X, p) be a compact pseudometric space, and let $T_1, \ldots, T_L$ be a family of homeomorphisms $T_\ell : X \to X$. Let μ be a regular Borel probability measure on X invariant under each T_ℓ. Let $\left(n_j^{(1)}\right)_{j=1}^\infty, \ldots, \left(n_j^{(L)}\right)_{j=1}^\infty$ be sequences of integers.*

Fix a point $x_0 \in X$, and let $(C_k)_{k=1}^\infty$ be a sequence of measurable subsets of X with positive measure for which $x_0 \in C_k$ and suppose that for each $\ell = 1, \ldots, \ell$, and every $\delta \in (0, \infty)$, we have that

$$\frac{\left|\left\{j \in \{0, 1, \ldots, k-1\} : \mathrm{diam}\left(T_\ell^{n_j^{(\ell)}} C_k\right) \geq \delta\right\}\right|}{k} \to 0.$$

Let $f_0, f_1, \ldots, f_L \in C(X)$. *Then*

$$\lim_{k \to \infty} \left| \left(\frac{1}{k} \sum_{j=0}^{k-1} f_0(x_0) \prod_{\ell=1}^{L} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - \alpha_{C_k} \left(\frac{1}{k} \sum_{j=0}^{k-1} f_0 \prod_{\ell=1}^{L} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right| = 0.$$

Proof. For the sake of making some notation in this proof more concise, we'll write

$$T_0 = \mathrm{id}_X,$$

$$n_j^{(0)} = 1 \qquad\qquad (\forall j \geq 0),$$

meaning that $f_0 \prod_{\ell=1}^{L} T_\ell^{n_j^{(\ell)}} f_\ell = \prod_{\ell=0}^{L} T_\ell^{n_j^{(\ell)}} f_\ell$. We also use $\| \cdot \|_u$ to denote the uniform norm on $C(X)$.

Fix $M = \max \{1, \|f_0\|_u, \|f_1\|_u, \ldots, \|f_L\|_u\}$, and fix $\epsilon > 0$. By appealing to the uniform continuity of the functions $f_0, f_1, \ldots, f_L$, choose $\delta_0, \delta_1, \ldots, \delta_L > 0$ such that

$$\forall x \in X \ \forall y \in X \ \left[(p(x, y) \leq \delta_\ell) \Rightarrow \left(|f_\ell(x) - f_\ell(y)| \leq \frac{\epsilon}{2(L+1)M^L} \right) \right] \qquad (\ell = 0, 1, \ldots, L).$$

Set $\delta = \min \{\delta_0, \delta_1, \ldots, \delta_L\} > 0$, and set

$$A_k^{(\ell)} = \left\{ j \in \{0, 1, \ldots, k-1\} : \mathrm{diam} \left(T_\ell^{n_j^{(\ell)}} C_k \right) < \delta \right\} \qquad (\ell = 1, \ldots, L, k \in \mathbb{N}),$$

$$A_k = \bigcap_{\ell=1}^{L} A_k^{(\ell)}.$$

Then

$$\frac{|\{0, 1, \ldots, k-1\} \setminus A_k|}{k} = \frac{\left| \bigcup_{\ell=1}^{L} \left(\{0, 1, \ldots, k-1\} \setminus A_k^{(L)} \right) \right|}{k}$$

$$\leq \sum_{\ell=1}^{L} \frac{\left| \left(\{0, 1, \ldots, k-1\} \setminus A_k^{(L)} \right) \right|}{k}$$

$$= \sum_{\ell=1}^{L} \frac{\left| \left\{ j \in \{0, 1, \ldots, k-1\} : \mathrm{diam} \left(T_\ell^{n_j^{(\ell)}} C_k \right) \geq \delta \right\} \right|}{k}$$

$$\overset{k \to \infty}{\to} 0.$$

86

We now turn to estimating

$$\left| \left(\sum_{j=0}^{k-1} f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - \alpha_{C_k} \left(\sum_{j=0}^{k-1} f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$\leq \frac{1}{k} \sum_{j=0}^{k-1} \left| \left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - \alpha_{C_k} \left(f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$= \frac{1}{k} \sum_{j=0}^{k-1} \left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$= \frac{1}{k} \left[\sum_{j \in A_k} \left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right| \right]$$

$$+ \frac{1}{k} \left[\sum_{j \in \{0,1,\ldots,k-1\} \setminus A_k} \left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right| \right].$$

In light of this decomposition, we make separate estimates on

$$\left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

based on whether $j \in A_k$ or $j \in \{0, 1, \ldots, k-1\} \setminus A_k$.

If $j \in A_k$, and $x \in A_k$, then $p\left(T_\ell^{n_j^{(\ell)}} x, T_\ell^{n_j^{(\ell)}} x_0 \right) < \delta$. Using an elementary "telescoping" estimate, it follows that

$$\left| \left(f_0(x) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x) \right) - \left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) \right|$$

$$\leq \sum_{h=0}^{L} \left(\prod_{\ell=0}^{h-1} \left| f_\ell \left(T_\ell^{n_j^{(\ell)}} x \right) \right| \right) \left| f_h \left(T_h^{n_j^{(h)}} x \right) - f_h \left(T_h^{n_j^{(h)}} x_0 \right) \right| \left(\left| \prod_{\ell=h+1}^{L} f_\ell \left(T_\ell^{n_j^{(\ell)}} x_0 \right) \right| \right)$$

$$\leq \sum_{h=0}^{L} M^h \frac{\epsilon}{2(L+1)M^L} M^{L-h}$$

$$= \epsilon/2.$$

On the other hand, if $j \in B_k$, then

$$\left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$\leq \left\| \left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right\|$$

$$\leq (2M)^{L+1}$$

$$= 2^{L+1} M^{L+1}.$$

Now, choose $K \in \mathbb{N}$ such that if $k \geq K$, then

$$\frac{|\{0, 1, \ldots, k-1\} \setminus A_k|}{k} \leq \frac{\epsilon}{2^{L+2} M^{L+1}}.$$

Then for all $k \geq K$, we have

$$\left| \left(\sum_{j=0}^{k-1} f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - \alpha_{C_k} \left(\sum_{j=0}^{k-1} f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$\leq \frac{1}{k} \sum_{j=0}^{k-1} \left| \left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - \alpha_{C_k} \left(f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$= \frac{1}{k} \sum_{j=0}^{k-1} \left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right|$$

$$= \frac{1}{k} \left[\sum_{j \in A_k} \left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right| \right]$$

$$+ \frac{1}{k} \left[\sum_{j \in \{0,1,\ldots,k-1\} \setminus A_k} \left| \alpha_{C_k} \left(\left(f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right| \right]$$

$$\leq \frac{|A_k|}{k} \frac{\epsilon}{2} + \frac{|\{0, 1, \ldots, k-1\} \setminus A_k|}{k} 2^{L+1} M^{L+1}.$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2^{L+2} M^{L+1}} 2^{L+1} M^{L+1}$$

$$= \epsilon.$$

Therefore

$$\lim_{k \to \infty} \left| \left(\sum_{j=0}^{k-1} f_0(x_0) \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell(x_0) \right) - \alpha_{C_k} \left(\sum_{j=0}^{k-1} f_0 \prod_{\ell=1}^{\ell} T_\ell^{n_j^{(\ell)}} f_\ell \right) \right| = 0.$$

$\square$

Theorem 3.3.12 can be used to convert pointwise convergence results for nonconventional ergodic averages into convergence results for random temporo-spatial averages, as shown by the following result.

Corollary 3.3.13. *Let (X, p) be a compact pseudometric space, and let $T_1, \ldots, T_L$ be a family of homeomorphisms $T_\ell : X \to X$. Let μ be a regular Borel probability measure on X invariant under each T_ℓ. Let $\left(n_j^{(1)}\right)_{j=1}^\infty, \ldots, \left(n_j^{(L)}\right)_{j=1}^\infty$ be sequences of integers.*

For each point $x \in X$, let $(C_k(x))_{k=1}^\infty$ be a sequence of measurable subsets of X with positive measure for which $x \in C_k(x)$ and suppose that for each $\ell = 1, \ldots, \ell$, and every $\delta \in (0, \infty)$, we have that

$$\frac{\left|\left\{j \in \{0, 1, \ldots, k-1\} : \mathrm{diam}\left(T_\ell^{n_j^{(\ell)}} C_k(x)\right) \geq \delta\right\}\right|}{k} \to 0.$$

Let $f_0, f_1, \ldots, f_L \in C(X)$, and suppose that $f^* \in L^\infty(X, \mu)$ such that

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f_0(x) \prod_{\ell=1}^\ell T_\ell^{n_j^{(\ell)}} f_\ell(x) = f^*(x)$$

for almost all $x \in X$. Then

$$\lim_{k \to \infty} \alpha_{C_k(x)}\left(\frac{1}{k} \sum_{j=0}^{k-1} f_0 \prod_{\ell=1}^\ell T_\ell^{n_j^{(\ell)}} f_\ell\right) = f^*(x)$$

for almost all $x \in X$.

Proof. Set

$$E = \left\{x \in X : \lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f_0(x) \prod_{\ell=1}^\ell T_\ell^{n_j^{(\ell)}} f_\ell(x) = f^*(x)\right\}.$$

If $x \in E$, then Theorem 3.3.12 tells us that

$$\lim_{k \to \infty} \alpha_{C_k(x)}\left(\frac{1}{k} \sum_{j=0}^{k-1} f_0 \prod_{\ell=1}^\ell T_\ell^{n_j^{(\ell)}} f_\ell\right) = f^*(x).$$

$\square$

3.4 Weighed temporo-spatial differentiation theorems

For the duration of this section, we narrow our attention to the case where $G = \mathbb{Z}$, and introduce a generalized form of a temporo-spatial differentiation problem. We also adopt the common notation that the action of the integer $n \in \mathbb{Z}$ be written as T^n. Let (X, μ) consist of a compact pseudometrizable space X endowed with a Borel probability measure μ, and let $T : X \to X$ be a homeomorphism. A *weight* on X is a measurable function $X \to \mathbb{T}$, where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. For convenience, write

$$\operatorname{Avg}_F^\xi f := \frac{1}{|F|} \sum_{j \in F} \xi^j \cdot \left(f \circ T^j \right),$$

where F is a finite nonempty subset of $\mathbb{Z}$. Let $(C_k)_{k=1}^\infty$ be a sequence of measurable subsets of X with $\mu(C_k) > 0$ for all $k \in \mathbb{N}$, and let $f \in L^\infty(X, \mu)$. What can be said of the limiting behavior of the sequence

$$\alpha_{C_k} \left(\operatorname{Avg}_{F_k}^\xi f \right)_{k=1}^\infty ?$$

Moreover, suppose Ξ is some family of measurable functions $X \to \mathbb{T}$. What can be said about the limiting behavior of the sequence

$$\alpha_{C_k} \left(\operatorname{Avg}_{F_k}^\xi f \right)_{k=1}^\infty$$

for all $\xi \in \Xi$?

We consider this problem in analogy with a classical problem of pointwise weighted temporal averages.

Wiener-Wintner pointwise ergodic theorem. *Let (X, μ) be a standard probability space, and let T be an automorphism of the probability space (X, μ). Set $[k] = \{0, 1, \ldots, k-1\} \subseteq \mathbb{Z}$. Then for each $f \in L^1(\mu)$ exists a set $X_f \subseteq X$ of full measure such that for all $x \in X_f$, and all $\theta \in \mathbb{T}$, the sequence*

$$\left(\operatorname{Avg}_{[k]}^\theta f(x) \right)_{k=1}^\infty$$

converges, where we identify the unimodular complex number θ with the constant function $x \mapsto \theta$ on X.

The first alleged proof of the Wiener-Wintner Theorem was presented in (Wiener and Wintner, 1941), but the argument presented was found to be incorrect. However, several proofs of the result have been presented since then. See (Assani, 2003, Chapter 2) for a discussion of several different approaches to the result.

As in Section 3.3, we present a very general result that allows us to reduce certain temporo-spatial problems to certain pointwise temporal problems. Afterwards, we provide specific examples of this reduction. Before we can prove Proposition 3.4.3, we introduce some terminology and prove an elementary technical lemma.

Definition 3.4.1. Let $\xi : (X, p) \to \mathbb{C}$ be a complex-valued function on a pseudometric space (X, p). A *modulus of uniform continuity for* ξ is a function $\Delta : (0, 1) \to (0, \infty)$ such that

$$\forall \epsilon \in (0, 1) \; \forall x_1, x_2 \in X \qquad [(p(x_1, x_2) \leq \Delta(\epsilon)) \Rightarrow |\xi(x_1) - \xi(x_2)| \leq \epsilon].$$

Given a family Ξ of functions $(X, p) \to \mathbb{C}$, we call a function $\Delta : (0, 1) \to (0, \infty)$ a *modulus of uniform equicontinuity* for Ξ if Δ is a modulus of uniform continuity for all $\xi \in \Xi$.

A function ξ is of course uniformly continuous if and only if it admits a modulus of uniform continuity, and a family Ξ is uniformly equicontinuous if and only if it admits a modulus of uniform equicontinuity. Note we make no assumption that a modulus of uniform continuity or modulus of uniform equicontinuity is the "best possible" choice. For example, if $\Xi = \{1\}$ consists solely of the constant function 1, then *any* map $(0, 1) \to (0, 1)$ would be both a modulus of uniform continuity for 1 and a modulus of uniform equicontinuity for Ξ.

Lemma 3.4.2. *Let (X, ρ) be a compact pseudometric space, and let Ξ be a uniformly equicontinuous family of functions $(X, p) \to \mathbb{T}$ with modulus of uniform equicontinuity Δ. Then $\Xi^j = \{\xi^j : \xi \in \Xi\}$ is a uniformly equicontinuous family for all $j \in \mathbb{Z}$, and if $j \neq 0$, then $\epsilon \mapsto \Delta(\epsilon/|j|)$ is a modulus of uniform equicontinuity for Ξ^j.*

Proof. If $j = 0$, then $\Xi^j = \{1\}$, which is trivially uniformly equicontinuous, and in fact any map $(0, 1) \to (0, 1)$ whatsoever will be a modulus of uniform equicontinuity for Ξ^0. Now assume that $j \neq 0$.

We prove this first for $j \in \mathbb{N}$, i.e. $j = |j| > 0$. Let $x_1, x_2 \in X, \xi \in \Xi$. We set up a telescoping sum

$$
\begin{aligned}
\left| \xi^j(x_1) - \xi^j(x_2) \right| &= \left| (\xi(x_1) - \xi(x_2)) \sum_{p=0}^{j-1} \xi^p(x_1) \xi^{j-p-1}(x_2) \right| \\
&= |\xi(x_1) - \xi(x_2)| \cdot \left| \sum_{p=0}^{j-1} \xi^p(x_1) \xi^{j-p-1}(x_2) \right| \\
&\leq |\xi(x_1) - \xi(x_2)| \cdot \sum_{j=0}^{p-1} \left| \xi^p(x_1) \xi^{j-p-1}(x_2) \right| \\
&= |\xi(x_1) - \xi(x_2)| \cdot j.
\end{aligned}
$$

Now, in the case where $j < 0$, i.e. $j = -|j|$, we observe that $\Xi^j = \left(\Xi^{|j|} \right)^{-1} = \left\{ \overline{\zeta} : \zeta \in \Xi^j \right\}$, and conjugation is an isometry. $\qquad \square$

Proposition 3.4.3. *Let (X, p) be a compact pseudometric space, and let $T : X \curvearrowright X$ be a homeomorphism of X. Let μ be a regular Borel probability measure on X. Fix a point $x_0 \in X$, and for each $n \in \mathbb{Z}, r \in (0, \infty)$, let $D_{x_0}(j, r)$ be the value*

$$
D_{x_0}(j, r) = \sup \left\{ p \left(T^j x_0, T^j x \right) : x \in X, p(x_0, x) \leq r \right\}.
$$

Let $(F_k)_{k=1}^{\infty}$ be a sequence of finite nonempty subsets of $\mathbb{Z}$. Let Ξ be a uniformly equicontinuous family of continuous functions $X \to \mathbb{T}$, and for each $j \in \mathbb{Z}$, let Δ^j be a modulus of uniform equicontinuity for Ξ^j. Let $(C_k)_{k=1}^{\infty}$ be a sequence of measurable subsets of X such that $\mu(C_k) > 0$ and $x_0 \in C_k$ for all $k \in \mathbb{N}$. Suppose that for every $\delta > 0, \epsilon > 0$, we have

$$
\lim_{k \to \infty} \frac{|\{j \in F_k : D_{x_0}(g, \operatorname{diam}(C_k)) > \delta\}|}{|F_k|} = 0,
$$
$$
\lim_{k \to \infty} \frac{|\{j \in F_k : \operatorname{diam}(C_k) > \Delta^j(\epsilon)\}|}{|F_k|} = 0.
$$

Let $f \in C(X)$. Finally, suppose there exists a constant $\lambda > 0$ such that $\mu \left(T^j C_k \right) \leq \lambda \mu(C_k)$ for all $j \in \mathbb{N}$. Then for all $\xi \in \Xi$, we have

$$
\lim_{k \to \infty} \left| \left(\operatorname{Avg}_{F_k}^{\xi} f \right) (x_0) - \alpha_{C_k} \left(\operatorname{Avg}_{F_k}^{\xi} f \right) \right| = 0,
$$

and the convergence is uniform in $\xi \in \Xi$.

Proof. Our proof of this result is similar in structure to our proof of Lemma 3.3.2, but with the added wrinkle of accounting for how the weight affects our averages.

Fix $\epsilon > 0$, and let $\xi \in \Xi$. Since f is uniformly continuous, there exists $\delta_1 > 0$ such that if $y_1, y_2 \in X$, and $p(y_1, y_2) \leq \delta$, then $|f(y_1) - f(y_2)| \leq \frac{\epsilon}{4\lambda \max\{1, \|f\|_u\}}$, where $\|\cdot\|_u$ denotes the uniform norm on $C(X)$. Write

$$A_k = \left\{ j \in F_k : D_{x_0}(j, \mathrm{diam}(C_k)) \leq \delta, \ \mathrm{diam}\left(T^j C_k\right) \leq \Delta^j \left(\frac{\epsilon}{4\lambda \max\{1, \|f\|_u\}} \right) \right\},$$

$$B_k = F_k \setminus A_k.$$

Our hypothesis tells us that $|A_k|/|F_k| \to 1, |B_k|/|F_k| \to 0$. We estimate

$$\left| \left(\mathrm{Avg}_{F_k}^\xi f \right)(x_0) - \alpha_{C_k} \left(\mathrm{Avg}_{F_k}^\xi \right) \right|$$

$$= \left| \alpha_{C_k} \left(\left(\mathrm{Avg}_{F_k}^\xi f \right)(x_0) - \mathrm{Avg}_{F_k}^\xi f \right) \right|$$

$$= \left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in F_k} \left[\xi(x_0)^j f\left(T^j x_0\right) - \xi^j \left(f \circ T^j\right) \right] \right) \right|$$

$$\leq \left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in A_k} \left[\xi(x_0)^j f\left(T^j x_0\right) - \xi^j \left(f \circ T^j\right) \right] \right) \right|$$

$$+ \left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in B_k} \left[\xi(x_0)^j f\left(T^j x_0\right) - \xi^j \left(f \circ T^j\right) \right] \right) \right|$$

We estimate these two terms separately, starting with the first.

If $j \in A_k$, then

$$\left| \alpha_{C_k} \left(\xi(x_0)^j f \left(T^j x_0 \right) - \xi^j \left(f \circ T^j \right) \right) \right|$$

$$= \left| \frac{1}{\mu(C_k)} \int_{C_k} \left(\xi^j(x_0) f \left(T^j x_0 \right) - \xi^j(x) f \left(T^j x \right) \right) \mathrm{d}\mu(x) \right|$$

$$\leq \frac{1}{\mu(C_k)} \int_{C_k} \left| \xi^j(x_0) f \left(T^j x_0 \right) - \xi^j(x) f \left(T^j x \right) \right| \mathrm{d}\mu(x)$$

$$= \frac{1}{\mu(C_k)} \int_{T^j C_k} \left| \xi^j (x_0) f \left(T^j x_0 \right) - \xi^j \left(T^{-j} y \right) f (y) \right| \mathrm{d}\mu(y)$$

$$\leq \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j (x_0) f \left(T^j x_0 \right) - \xi^j \left(T^{-j} y \right) f (y) \right| \mathrm{d}\mu(y)$$

$$\leq \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j (x_0) f \left(T^j x_0 \right) - \xi^j (x_0) f (y) \right| \mathrm{d}\mu(y)$$

$$+ \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j (x_0) f (y) - \xi^j \left(T^{-j} y \right) f (y) \right| \mathrm{d}\mu(y)$$

$$= \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j(x_0) \right| \left| f \left(T^j x_0 \right) - f (y) \right| \mathrm{d}\mu(y)$$

$$+ \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j (x_0) - \xi^j \left(T^{-j} y \right) \right| \cdot \left| f(y) \right| \mathrm{d}\mu(y)$$

$$\leq \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| f \left(T^j x_0 \right) - f (y) \right| \mathrm{d}\mu(y)$$

$$+ \frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j (x_0) - \xi^j \left(T^{-j} y \right) \right| \cdot \| f \|_\infty \mathrm{d}\mu(y).$$

First, if $y \in T^j C_k$, and $D_{x_0} \left(j, \mathrm{diam}(C_k) \leq \delta \right)$, then $p \left(T^j x_0, y \right) \leq \delta$, meaning that

$$\frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| f \left(T^j x_0 \right) - f (y) \right| \mathrm{d}\mu(y) \leq \frac{\epsilon}{4},$$

and $y \in T^j C_k \Rightarrow T^{-j} y \in C_k$, meaning that $p \left(x_0, T^{-j} y \right) \leq \mathrm{diam}(C_k) \leq \Delta^j \left(\frac{\epsilon}{4\lambda \max\{1, \|f\|_u\}} \right)$, so

$$\frac{\lambda}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| \xi^j (x_0) - \xi^j \left(T^{-j} y \right) \right| \cdot \| f \|_\infty \mathrm{d}\mu(y) \leq \frac{\epsilon}{4}.$$

Thus

$$\left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in A_k} \left[\xi(x_0)^j f \left(T^j x_0 \right) - \xi^j \left(f \circ T^j \right) \right] \right) \right| \leq \frac{\epsilon}{2}.$$

Suppose now that $j \in B_k$. By a computation similar to the one performed for the case where $j \in A_k$, we get

$$\left|\xi(x_0)^j f\left(T^j x_0\right) - \xi^j\left(f \circ T^j\right)\right|$$

$$\leq \frac{\lambda}{\mu\left(T^j C_k\right)} \int_{T^j C_k} \left|\xi^j\left(x_0\right) f\left(T^j x_0\right) - \xi^j\left(T^{-j} y\right) f\left(y\right)\right| \mathrm{d}\mu(y)$$

$$\leq \frac{\lambda}{\mu\left(T^j C_k\right)} \int_{T^j C_k} \left(\left|\xi^j\left(x_0\right) f\left(T^j x_0\right)\right| + \left|\xi^j\left(T^{-j} y\right) f\left(y\right)\right|\right) \mathrm{d}\mu(y)$$

$$\leq \frac{\lambda}{\mu\left(T^j C_k\right)} \mu\left(T^j C_k\right) \left(2\|f\|_u\right)$$

$$= 2\lambda\|f\|_u.$$

Choose $K \in \mathbb{N}$ such that if $k \geq K$, then $\frac{|B_k|}{|F_k|} \leq \frac{\epsilon}{4\lambda \max\{1, \|f\|_u\}}$. Then if $k \geq K$, we have

$$\left|\left(\mathrm{Avg}_{F_k}^{\xi} f\right)(x_0) - \alpha_{C_k}\left(\mathrm{Avg}_{F_k}^{\xi}\right)\right|$$

$$\leq \left|\alpha_{C_k}\left(\frac{1}{|F_k|} \sum_{j \in A_k} \left[\xi(x_0)^j f\left(T^j x_0\right) - \xi^j\left(f \circ T^j\right)\right]\right)\right|$$

$$+ \left|\alpha_{C_k}\left(\frac{1}{|F_k|} \sum_{j \in B_k} \left[\xi(x_0)^j f\left(T^j x_0\right) - \xi^j\left(f \circ T^j\right)\right]\right)\right|$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2}$$

$$= \epsilon.$$

We note that our estimates on K were independent of our choice of $\xi \in \Xi$, meaning the convergence is uniform in ξ. $\qquad\square$

With this in mind, we can state the following.

Theorem 3.4.4. *Let (X, p) be a compact pseudometric space, and let $T : X \curvearrowright X$ be a homeomorphism of X. Let μ be a regular Borel probability measure on X. Suppose there exist functions $H, L : \mathbb{Z} \to (0, \infty)$ such that*

$$p\left(T^g x, T^g y\right) \leq L(j) \cdot p(x, y)^{H(j)} \qquad\qquad (\forall j \in \mathbb{Z}, x \in X, y \in X).$$

Let $(F_k)_{k=1}^{\infty}$ be a sequence of finite nonempty subsets of $\mathbb{Z}$. Let Ξ be a uniformly equicontinuous family of continuous functions $X \to \mathbb{T}$, and for each $j \in \mathbb{Z}$, let Δ^j be a modulus of uniform equicontinuity for Ξ^j. Let $(C_k)_{k=1}^{\infty}$ be a sequence of measurable subsets of X such that $\mu(C_k) > 0$ and $x_0 \in C_k$ for all $k \in \mathbb{N}$. Suppose that for every $\delta > 0, \epsilon > 0$, we have

$$\lim_{k \to \infty} \frac{\left|\{j \in F_k : L(j) \cdot \operatorname{diam}(C_k)^{H(j)} > \delta\}\right|}{|F_k|} = 0,$$

$$\lim_{k \to \infty} \frac{\left|\{j \in F_k : \operatorname{diam}(C_k) > \Delta^j(\epsilon)\}\right|}{|F_k|} = 0.$$

Let $x_0 \in X$ be a point in X, and let $f : X \to \mathbb{C}$ be a uniformly bounded continuous function Finally, suppose there exists a constant $\lambda > 0$ such that $\mu\left(T^j C_k\right) \leq \lambda\mu(C_k)$ for all $j \in \mathbb{N}$. Then for all $\xi \in \Xi$, we have

$$\lim_{k \to \infty} \left|\left(\operatorname{Avg}_{F_k}^{\xi} f\right)(x_0) - \alpha_{C_k}\left(\operatorname{Avg}_{F_k}^{\xi} f\right)\right| = 0,$$

and the convergence is uniform in $\xi \in \Xi$.

Proof. We have the bound $D_{x_0}(j, r) \leq L(j) \cdot r^{H(j)}$. We can thus apply Lemma 3.4.3. $\qquad \square$

Corollary 3.4.5. *Let* (X, ρ) *be a compact metric space, and let* $T : X \to X$ *be a homeomorphism. Let* μ *be a Borel probability measure on* X *that's* T-*invariant. Let* $f \in C(X)$. *For each* $x \in X$, *let* $C_k(x)$ *be a measurable subset of* X *with positive measure such that*

$$\lim_{k \to \infty} \frac{\left|\{j \in [k] : D_x(j, \operatorname{diam}(C_k(x))) > \delta\}\right|}{k} = 0,$$

$$\lim_{k \to \infty} \frac{\left|\{j \in [k] : \operatorname{diam}(C_k(x)) > \Delta^j(\epsilon)\}\right|}{k} = 0.$$

Then for every $f \in C(X)$ *exists a set* $X_f \subseteq X$ *of full measure such that for all* $x \in X_f$, *and all* $\theta \in \mathbb{T}$, *the sequence*

$$\left(\alpha_{C_k(x)}\left(\operatorname{Avg}_{[k]}^{\theta}\right)\right)_{k=1}^{\infty}$$

converges.

Proof. By the Wiener-Wintner pointwise ergodic theorem, there exists a set $X_f \subseteq X$ of full measure such that for all $x \in X_f$, and all $\theta \in \mathbb{T}$, the sequence

$$\left(\mathrm{Avg}_{[k]}^{\theta} f(x) \right)_{k=1}^{\infty}$$

converges. By Lemma 3.4.3, it follows that $\lim_{k \to \infty} \left| \alpha_{C_k(x)} \left(\mathrm{Avg}_{[k]}^{\theta} f \right) - \left(\mathrm{Avg}_{[k]}^{\theta} f(x) \right) \right| = 0$. Thus the sequence converges. $\qquad\square$

We now consider a different class of weighting sequences, where we choose our weights to be constant functions, but loosen our assumptions about boundedness. Given a sequence $\mathbf{F} = (F_k)_{k=1}^{\infty}$ of finite subsets of $\mathbb{Z}$, set

$$M^{\mathbf{F}} := \left\{ (a_k)_{k \in \mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}} : \sup_{\ell \in \mathbb{N}} \frac{1}{|F_\ell|} \sum_{j \in F_\ell} |a_j| < \infty \right\}.$$

We also introduce the notation

$$\mathrm{Avg}_F^a := \frac{1}{|F|} \sum_{j \in F} a_j f \circ T^j,$$

where F is a finite subset of $\mathbb{Z}$.

Our next result establishes that under a rapidly decaying diameter condition, temporo-spatial differentiations involving weighted ergodic means for continuous functions can be reduced to pointwise temporal averages. The twist here is that the diameter decay condition also hinges on the weighting sequence.

Proposition 3.4.6. *Let (X, p) be a compact pseudometric space, and let $T : X \curvearrowright X$ be a homeomorphism of X. Let μ be a regular Borel probability measure on X. Fix a point $x_0 \in X$, and for each $n \in \mathbb{Z}, r \in (0, \infty)$, let $D_{x_0}(j, r)$ be the value*

$$D_{x_0}(j, r) = \sup \left\{ p \left(T^j x_0, T^j x \right) : x \in X, p(x_0, x) \le r \right\}.$$

Let $\mathbf{F} = (F_k)_{k=1}^{\infty}$ be a sequence of finite nonempty subsets of $\mathbb{Z}$. Let $(a_k)_{k=0}^{\infty} \in M^{\mathbf{F}}$, and let $(C_k)_{k=1}^{\infty}$ be a sequence of measurable subsets of X such that $\mu(C_k) > 0$ and $x_0 \in C_k$ for all $k \in \mathbb{N}$. Suppose that for every $\delta > 0$, we have

$$\lim_{k \to \infty} \frac{1}{|F_k|} \sum_{j \in F_k, \, D_{x_0}(j, \mathrm{diam}(C_k)) > \delta} |a_j| = 0,$$

97

Let $f : X \to \mathbb{C}$ be a continuous function. Finally, suppose there exists a constant $\lambda > 0$ such that $\mu\left(T^j C_k\right) \leq \lambda\mu(C_k)$ for all $j \in \mathbb{N}$. Then we have

$$\lim_{k \to \infty} \left|\left(\operatorname{Avg}_{F_k}^a f\right)(x_0) - \alpha_{C_k}\left(\operatorname{Avg}_{F_k}^a f\right)\right| = 0.$$

Proof. Fix $\epsilon > 0$. Appealing to the uniform continuity and boundedness of f, choose $\delta > 0$ such that

$$p(y_1, y_2) \leq \delta \Rightarrow |f(y_1) - f(y_2)| \leq \epsilon.$$

Set

$$A_k = \left\{j \in F_k : D_{x_0}(j, \operatorname{diam}(C_k)) \leq \delta\right\},$$

$$B_k = \left\{j \in F_k : D_{x_0}(j, \operatorname{diam}(C_k)) > \delta\right\}.$$

Then $\frac{|A_k|}{|F_k|} \leq 1$, and $\frac{1}{|F_k|}\sum_{j \in B_k}|a_j| \to 0$. Using a calculation similar to that used in our proof of Proposition 3.4.3, we get

$$\left|\left(\operatorname{Avg}_{F_k}^a f\right)(x_0) - \alpha_{C_k}\left(\operatorname{Avg}_{F_k}^a\right)\right| \leq \left|\alpha_{C_k}\left(\frac{1}{|F_k|}\sum_{j \in A_k}\left[a_j f\left(T^j x_0\right) - a_j\left(f \circ T^j\right)\right]\right)\right|$$

$$+ \left|\alpha_{C_k}\left(\frac{1}{|F_k|}\sum_{j \in B_k}\left[a_j f\left(T^j x_0\right) - a_j\left(f \circ T^j\right)\right]\right)\right|.$$

As before, we'll estimate these two terms separately. Many of the calculations done here are quite similar to those used in our proof of Proposition 3.4.3, so we will be terser in our presentation here.

First, suppose $j \in A_k$. Then

$$
\begin{aligned}
\left| \alpha_{C_k} \left(a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right) \right| &= \frac{|a_j|}{\mu(C_k)} \left| \int_{C_k} \left(f \left(T^j x_0 \right) - f \left(T^j x \right) \right) \mathrm{d}\mu(x) \right| \\
&\leq \frac{|a_j|}{\mu(C_k)} \int_{C_k} \left| f \left(T^j x_0 \right) - f \left(T^j x \right) \right| \mathrm{d}\mu(x) \\
&\leq \lambda |a_j| \frac{1}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| f \left(T^j x_0 \right) - f \left(y \right) \right| \mathrm{d}\mu(y) \\
&\leq \lambda |a_j| \epsilon.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in A_k} \left[a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right] \right) \right| \\
&\leq \frac{1}{|F_k|} \sum_{j \in A_k} \left| \alpha_{C_k} \left(a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right) \right| \\
&\leq \frac{1}{|F_k|} \sum_{j \in A_k} \lambda |a_j| \epsilon \\
&\leq \frac{1}{|F_k|} \sum_{j \in F_k} |a_j| \lambda \epsilon \\
&\leq \left(\sup_{\ell \in \mathbb{N}} |F_\ell|^{-1} \sum_{j \in F_\ell} |a_j| \right) \lambda \epsilon.
\end{aligned}
$$

Now, consider the case where $j \in B_k$. Then

$$
\begin{aligned}
\left| \alpha_{C_k} \left(a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right) \right| &\leq \lambda |a_j| \frac{1}{\mu \left(T^j C_k \right)} \int_{T^j C_k} \left| f \left(T^j x_0 \right) - f \left(y \right) \right| \mathrm{d}\mu(y) \\
&\leq \lambda |a_j| \left(2 \|f\|_u \right).
\end{aligned}
$$

Choose $K \in \mathbb{N}$ such that if $k \geq K$, then $\frac{1}{|F_k|} \sum_{j \in B_k} |a_j| \leq \epsilon$. Then

$$\left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in A_k} \left[a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right] \right) \right| \leq \frac{1}{|F_k|} \sum_{j \in B_k} \lambda |a_j| (2\|f\|_u)$$

$$\leq 2\lambda \|f\|_u \frac{1}{|F_k|} \sum_{j \in B_k} |a_j|$$

Therefore, if $k \geq K$, we have

$$\left| \left(\mathrm{Avg}_{F_k}^a f \right) (x_0) - \alpha_{C_k} \left(\mathrm{Avg}_{F_k}^a \right) \right|$$

$$\leq \left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in A_k} \left[a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right] \right) \right|$$

$$+ \left| \alpha_{C_k} \left(\frac{1}{|F_k|} \sum_{j \in B_k} \left[a_j f \left(T^j x_0 \right) - a_j \left(f \circ T^j \right) \right] \right) \right|$$

$$\leq \left(\sup_{\ell \in \mathbb{N}} |F_\ell|^{-1} \sum_{j \in F_\ell} |a_j| \right) \lambda \epsilon + 2\lambda \|f\|_u \epsilon$$

$$= \lambda \left(\left(\sup_{\ell \in \mathbb{N}} |F_\ell|^{-1} \sum_{j \in F_\ell} |a_j| \right) + 2\|f\|_u \right) \epsilon.$$

This coefficient on ϵ is independent of our choice of k, so we can conclude that

$$\lim_{k \to \infty} \left| \left(\mathrm{Avg}_{F_k}^a f \right) (x_0) - \alpha_{C_k} \left(\mathrm{Avg}_{F_k}^a f \right) \right| = 0.$$

$\square$

Multi-local temporo-spatial differentiations

In this chapter, we focus on the case where the spatial averaging sequence $(C_k)_{k=1}^\infty$ consists of finite unions of balls, a setting we call "multi-local." We study sufficient conditions for these corresponding temporo-spatial differentiations to converge, as well as the existence and prevalence of pathological multi-local temporo-spatial differentiations.

In Section 4.1, we establish several notations that will be used throughout the chapter, as well as some standing assumptions and conventions.

In Section 4.2, we provide sufficient conditions for multi-local temporo-spatial differentations to converge. We also show how these convergence results can fail if certain assumptions are relaxed.

In Section 4.3, we briefly present the theory of ergodic optimization. In particularly, we characterize the maximum ergodic average in the context of continuous actions of amenable groups.

In Section 4.4, we construct multi-local temporo-spatial differentiations for a given real-valued continuous function f which have a prescribed compact set $\mathcal{K}$ as the set of limit points of the temporo-spatial differentiation.

In Section 4.5, we consider temporo-spatial differentiations as sequences of measures

$$\left(f \mapsto \frac{1}{\mu(C_k)} \int_{\mu(C_k)} \sum_{j=0}^{k-1} T^j f \mathrm{d}\mu \right)_{k=1}^\infty,$$

and consider how to construct sequences $(C_k)_{k=1}^\infty$ for which $\mathrm{LS}\left(\left(f \mapsto \frac{1}{\mu(C_k)} \int_{C_k} \sum_{j=0}^{k-1} T^j f \mathrm{d}\mu \right)_{k=1}^\infty \right)$ is some prescribed subset of the Choquet simplex of T-invariant Borel probability measures on X, where $\mathrm{LS}\left((z_k)_{k=1}^\infty \right)$ denotes the set of all limits of convergent subsequences of $(z_k)_{k=1}^\infty$ (defined in more detail in Section 4.1). In particular, we construct examples of $(C_k)_{k=1}^\infty$ for which $\mathrm{LS}\left(\left(f \mapsto \frac{1}{\mu(C_k)} \int_{C_k} \sum_{j=0}^{k-1} T^j f \mathrm{d}\mu \right)_{k=1}^\infty \right)$ is the entire Choquet simplex of T-invariant measures.

In Section 4.6, we show that for a system (X, T) with a specification-like property that we call the Very Weak Specification Property, there exists a residual set of $x \in X$ that exhibit a strong form of the

maximal Birkhoff average oscillation property. Specifically, there exists a residual set of $x \in X$ such that $\mathrm{LS}\left(\left(\mu_{x,\pi(k)}\right)_{k=1}^{\infty}\right)$ is the entire Choquet simplex of T-invariant measures for all non-constant polynomials $\pi(t) \in \mathbb{Q}[t]$ such that $\pi(\mathbb{N}) \subseteq \mathbb{N}$, where $\mu_{x,k}$ are the the empirical measures $\mu_{x,k} = \frac{1}{k}\sum_{j=0}^{k-1} \delta_x \circ T^j$ for $x \in X$. Consequently, for sequences $(r_k)_{k=1}^{\infty}$ of radii decaying to 0 sufficiently fast, we have for a residual set of $x \in X$ that $\mathrm{LS}\left(\left(f \mapsto \frac{1}{\mu(B(x;r_k))}\int_{B(x;r_k)} \sum_{j=0}^{\pi(k)-1} T^j f \mathrm{d}\mu\right)_{k=1}^{\infty}\right)$ is the entire Choquet simplex of T-invariant measures for all non-constant integer-valued polynomials $\pi(t)$ that send nonnegative integers to nonnegative integers.

4.1   Notations and conventions

Here we identify particular notations and conventions we adopt throughout this chapter. Individual sections might place additional assumptions on some of the objects we define here. We also place more novel definitions in the later sections of the chapter.

We will let (X, ρ) be a compact metric space, and $T : G \curvearrowright X$ will be a continuous monoidal left-action of a discrete monoid G on X by continuous maps $(T_g)_{g \in G}$ (not necessarily invertible). That is to say, the maps $(T_g)_{g \in G}$ will satisfy the laws

$$T_{g_1} \circ T_{g_2} = T_{g_1 g_2} \qquad\qquad (\forall g_1, g_2 \in G),$$
$$T_{1_G} = \mathrm{id}_X,$$

where 1_G denotes the identity element of G. We will use μ to denote a Borel probability measure on X, though we will not in general assume that μ is T-invariant. The support of μ will be denoted $\mathrm{supp}(\mu)$.

Given a finite subset F of G, and a function $f : X \to \mathbb{C}$, we write

$$\mathrm{Avg}_F f := \frac{1}{|F|}\sum_{g \in F} T_g f,$$

where $T_g f := f \circ T_g$. Similarly, if β is a Borel probability measure on X, and $E \subseteq X$ is a Borel subset of X, we will write

$$\left(\beta \circ \mathrm{Avg}_F\right)(E) := \frac{1}{|F|}\sum_{g \in F} \beta\left(T_g^{-1} E\right).$$

These notations are consistent with each other in the sense that if $f \in C(X)$, then

$$\int_X f \mathrm{d} \left(\beta \circ \mathrm{Avg}_F \right) = \int_X \mathrm{Avg}_F \, f \mathrm{d}\beta.$$

If no domain is specified for an integral $\int$, then the integral is assumed to be over X, i.e. $\int := \int_X$.

We will denote the space of all Borel probability measures on X by $\mathcal{M}(X)$. We will always consider $\mathcal{M}(X)$ with the weak*-topology, making $\mathcal{M}_T(X)$ a Choquet simplex. We use $\mathcal{M}_T(X)$ to denote the space of T-invariant Borel probability measures on X, also equipped with the weak*-topology to make $\mathcal{M}_T(X)$ a Choquet simplex.

We use $\partial_e S$ to denote the set of extreme points of a subset S of a real topological vector space, i.e. $\partial_e S$ denotes the set of all points in S which cannot be expressed nontrivially as a convex combination of points in S.

We will use $\mathbb{N}$ to denote the set of positive integers, and $\mathbb{N}_0$ to denote the set of nonnegative integers.

A sequence $(F_k)_{k=1}^{\infty}$ of finite subsets of a group G is called *Følner* if

$$\lim_{k \to \infty} \frac{|F_k \Delta g F_k|}{|F_k|} = 0 \qquad\qquad (\forall g \in G),$$

where $|\cdot|$ denotes cardinality and Δ is the symmetric difference, i.e. $A\Delta B = (A \setminus B) \cup (B \setminus A)$.

Given a sequence $(z_k)_{k=1}^{\infty}$ in a topological space Z, we write

$$\mathrm{LS}\left((z_k)_{k=1}^{\infty}\right) := \left\{ \lim_{\ell \to \infty} z_{k_\ell} : k_1 < k_2 < \cdots , \ \lim_{\ell \to \infty} z_{k_\ell} \text{ exists} \right\}$$

to denote the set of limit points of $(z_k)_{k=1}^{\infty}$, called the *limit set* of $(z_k)_{k=1}^{\infty}$.

4.2   Convergence results and their limitations

Definition 4.2.1. Let $\bar{x} = \left(x^{(1)}, \ldots, x^{(n)}\right) \in X^n, \bar{r} = \left(r^{(1)}, \ldots, r^{(n)}\right) \in (0, \infty)^n$. We write

$$B\left(\bar{x}; \bar{r}\right) := \bigcup_{h=1}^{n} B\left(x^{(h)}, r^{(h)}\right),$$

where $B\left(x; r\right) := \{y \in X : \rho(x, y) < r\}$ is the open ball with center x and radius r. We refer to sets of the form $B(\bar{x}; \bar{r})$ as *multi-balls*.

Lemma 4.2.2. *Every multi-ball $B\left(x^{(1)}, \ldots, x^{(n)}; r^{(1)}, \ldots, r^{(n)}\right)$ can be expressed in the form*

$$B\left(y^{(1)}, \ldots, y^{(m)}; s^{(1)}, \ldots, s^{(m)}\right),$$

where $y^{(1)}, \ldots, y^{(m)}$ are distinct.

Proof. If $x^{(1)}, \ldots, x^{(n)}$ are not already distinct, then we can write $\left\{x^{(1)}, \ldots, x^{(n)}\right\} = \left\{x^{(h_1)}, \ldots, x^{(h_m)}\right\}$, where $h_1, \ldots, h_m \in \{1, \ldots, n\}$, and $x^{(h_1)}, \ldots, x^{(h_m)}$ are distinct. Then

$$B\left(x^{(1)}, \ldots, x^{(n)}; r^{(1)}, \ldots, r^{(n)}\right) = B\left(y^{(1)}, \ldots, y^{(p)}; s^{(1)}, \ldots, s^{(m)}\right),$$

where $y^{(p)} = x^{(h_p)}, s^{(p)} = \max\left\{r^{(h)} : x^{(h)} = y^{(p)}\right\}$. $\qquad\square$

Definition 4.2.3. Let (X, ρ) be a compact metric space, and let $T : G \curvearrowright X$ be an action of a discrete semigroup G by Hölder maps T_g equipped with functions $H, L : G \to (0, \infty)$ such that

$$\rho\left(T_g x, T_g y\right) \leq L(g) \cdot \rho(x, y)^{H(g)} \qquad\qquad (\forall g \in G, x \in X, y \in X).$$

We refer to the pair (H, L) as a *modulus of Hölder continuity* (abbreviated *MoHöC*) for T. Let $\mathbf{F} = (F_k)_{k=1}^{\infty}$ be a sequence of nonempty finite subsets of G. We say that a sequence $(\bar{r}_k)_{k=1}^{\infty}$ of n-tuples $\bar{r}_k = \left(r_k^{(1)}, \ldots, r_k^{(n)}\right)_{k=1}^{\infty}$ of positive numbers *decays $(X, \rho, H, L, \mathbf{F})$-fast* if

$$\lim_{k \to \infty} \frac{\left|\left\{g \in F_k : L(g) \cdot \left(r_k^{(h)}\right)^{H(g)} > \delta\right\}\right|}{|F_k|} = 0 \qquad (\forall \delta \in (0, \infty), h \in \{1, \ldots, n\}),$$

$$\lim_{k \to \infty} r_k^{(h)} = 0 \qquad\qquad (\forall h \in \{1, \ldots, n\}).$$

An immediate observation about this definition is that if $(\bar{r}_k)_{k=1}^{\infty}$ is a sequence of n-tuples of positive numbers that decay $(X, \rho, H, L, \mathbf{F})$-fast, and $(\bar{s}_k)_{k=1}^{\infty}$ is another sequence of n-tuples of positive numbers for which there exists $K \in \mathbb{N}$ such that $s_k^{(h)} \leq r_k^{(h)}$ for all $h \in \{1, \ldots, n\}, k \geq K$, then $(\bar{s}_k)_{k=1}^{\infty}$ decays $(X, \rho, H, L, \mathbf{F})$-fast. So we have in fact described a rapid decay condition. Moreover, any system of Hölder maps with MoHöC (H, L) will admit a sequence $(r_k)_{k=1}^{\infty}$ that decays $(X, \rho, H, L, \mathbf{F})$-fast.

Our assumption that $\lim_{k\to\infty} r_k^{(h)} = 0$ ensures that if $x^{(1)}, \ldots, x^{(h)}$ are distinct points in X, then the balls $\left\{ B\left(x^{(h)}; r_k^{(h)}\right) \right\}_{h=1}^{n}$ are pairwise disjoint for sufficiently large k, since for sufficiently large k we'll have that

$$\max\left\{ r_k^{(1)}, \ldots, r_k^{(n)} \right\} < \frac{1}{2} \min\left\{ \rho\left(x^{(h_1)}, x^{(h_2)}\right) : 1 \leq h_1 < h_2 \leq n \right\}.$$

For the remainder of this section, $T : G \curvearrowright X$ will be an action of a discrete group G on X by Hölder homoeomorphisms with MoHöC (H, L), and $\mathbf{F} = (F_k)_{k=1}^{\infty}$ will be a sequence of nonempty finite subsets of G.

Notation 4.2.4. Let μ be a Borel probability measure on X, and let $E \subseteq X$ be a μ-measurable set such that $\mu(E) > 0$. The functional $\alpha_E : C(X) \to \mathbb{C}$ is defined as

$$\alpha_E(f) := \frac{1}{\mu(E)} \int_E f \mathrm{d}\mu.$$

We will sometimes also treat α_E instead as a Borel probability measure $\alpha_E : A \mapsto \mu(A|E)$. These interpretations are consistent with each other in the sense that $\alpha_E(f) = \int f \mathrm{d}\alpha_E$ for all $f \in C(X)$.

Lemma 4.2.5. *Let* $x \in X$, *and let* $(r_k)_{k=1}^{\infty}$ *be a sequence of positive numbers that decays* $(X, \rho, H, L, \mathbf{F})$*-fast, and suppose* $f \in C(X)$. *Let* μ *be a Borel probability measure on* X, *and let* $x \in \mathrm{supp}(\mu)$. *Then*

$$\lim_{k\to\infty} \left(\alpha_{B(x;r_k)} \left(\mathrm{Avg}_{F_k} f \right) - \mathrm{Avg}_{F_k} f(x) \right) = 0.$$

Moreover, if f *satisfies the Hölder condition*

$$|f(y) - f(z)| \leq c \cdot \rho(y, z)^{\beta} \qquad\qquad (\forall y, z \in X),$$

for some constants $c, \beta \in (0, \infty)$, *then*

$$\left| \alpha_{B(x;r_k)} \left(\mathrm{Avg}_{F_k} f \right) - \mathrm{Avg}_{F_k} f(x) \right| \leq \frac{c}{|F_k|} \sum_{g \in F_k} L(g)^{\beta} \cdot r_k^{\beta H(g)}.$$

Proof. Fix $\epsilon > 0$. Since f is continuous and X is compact, we know that f is *uniformly* continuous, meaning that there exists $\delta > 0$ such that

$$\rho(y, z) \le \delta \Rightarrow |f(y) - f(z)| \le \epsilon.$$

Set

$$A_k = \left\{ g \in F_k : L(g) \cdot r_k^{H(g)} \ge \delta \right\}.$$

By the hypothesis that $(r_k)_{k=1}^\infty$ decays $(X, \rho, H, L, \mathbf{F})$-fast, we know that $\lim_{k \to \infty} \frac{|A_k|}{|F_k|} = 0$.

We estimate

$$\left| \alpha_{B(x;r_k)} \left(\text{Avg}_{F_k} f \right) - \text{Avg}_{F_k} f(x) \right|$$

$$= \left| \frac{1}{\mu(B(x; r_k))} \int_{B(x;r_k)} \frac{1}{|F_k|} \sum_{g \in F_k} (f(T_g y) - f(T_g x)) \, d\mu(y) \right|$$

$$\le \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x;r_k)} |f(T_g y) - f(T_g x)| \, d\mu(y) \tag{$\dagger$}$$

$$= \left(\frac{1}{|F_k|} \sum_{g \in A_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x;r_k)} |f(T_g y) - f(T_g x)| \, d\mu(y) \right)$$

$$+ \left(\frac{1}{|F_k|} \sum_{g \in F_k \setminus A_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x;r_k)} |f(T_g y) - f(T_g x)| \, d\mu(y) \right).$$

We will return to the line marked ($\dagger$) when we compute the estimate for the case where f is Hölder. For now, we estimate these two terms separately.

$$\frac{1}{|F_k|} \sum_{g \in A_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x;r_k)} |f(T_g y) - f(T_g x)| \, d\mu(y)$$

$$\le \frac{1}{|F_k|} \sum_{g \in A_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x;r_k)} \|2f\|_{C(X)} \, d\mu(y)$$

$$= \frac{2|A_k|}{|F_k|} \|f\|_{C(X)}.$$

Choose $K \in \mathbb{N}$ sufficiently large that $\frac{|A_k|}{|F_k|} \le \epsilon$. Then for $k \ge K$, we have that

$$\frac{2|A_k|}{|F_k|} \|f\|_{C(X)} \le 2\|f\|_{C(X)} \epsilon.$$

For the other of second of the two aforementioned terms, we observe that if $g \in F_k \setminus A_k$, then

$$\rho(T_g y, T_g x) \leq L(g) \cdot \rho(x, y)^{H(g)}$$

$$\leq \delta$$

$$\Rightarrow |f(T_g y) - f(T_g x)| \leq \epsilon.$$

Thus

$$\frac{1}{|F_k|} \sum_{g \in F_k \setminus A_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} |f(T_g y) - f(T_g x)| \, \mathrm{d}\mu(y)$$

$$\leq \frac{1}{|F_k|} \sum_{g \in F_k \setminus A_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} \epsilon \mathrm{d}\mu(y)$$

$$= \frac{|F_k| - |A_k|}{|F_k|} \epsilon$$

$$\leq \epsilon.$$

Therefore, if $k \geq K$, then

$$\left| \alpha_{B(x; r_k)} \left(\mathrm{Avg}_{F_k} f \right) - \mathrm{Avg}_{F_k} f(x) \right| \leq \left(2\|f\|_{C(X)} + 1 \right) \epsilon.$$

Finally, in the case where we have the additional hypothesis that f is (c, β)-Hölder, we can instead estimate the earlier (†) as

$$
\begin{aligned}
&\frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} |f(T_g y) - f(T_g x)| \, \mathrm{d}\mu(y) \\
&\leq \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} c \cdot \rho(T_g y, T_g x)^\beta \mathrm{d}\mu(y) \\
&\leq \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} c \cdot \left(L(g) \cdot \rho(x, y)^{H(g)} \right)^\beta \, \mathrm{d}\mu(y) \\
&\leq \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} c \cdot \left(L(g) \cdot r_k^{H(g)} \right)^\beta \, \mathrm{d}\mu(y) \\
&= \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{\mu(B(x; r_k))} \int_{B(x; r_k)} c \cdot \left(L(g)^\beta \cdot r_k^{\beta H(g)} \right) \mathrm{d}\mu(y) \\
&= \frac{c}{|F_k|} \sum_{g \in F_k} L(g)^\beta \cdot r_k^{\beta H(g)}.
\end{aligned}
$$

$\square$

The upshot of Lemma 4.2.5 is that when we consider temepero-spatial differentiations with respect to balls of radius decaying sufficiently fast centered at a fixed point x_0, this temporo-spatial differentiation is equivalent to a pointwise (temporal) ergodic average. On one hand, this means that we can consider certain "random" temepero-spatial differentiations by appealing to pointwise convergence theorems, as in Corollary 4.2.10. On another hand, this means that we can use pathological pointwise ergodic averages to generate pathological temporo-spatial differentiations, as we will see in Section 4.6.

The following lemma lets us describe temporo-spatial averages over multi-balls in terms of temporo-spatial averages over balls, and will be useful going forward.

Lemma 4.2.6. *Let μ be a Borel probability measure on X, and let $x^{(1)}, \ldots, x^{(n)} \in \mathrm{supp}(\mu); r^{(1)}, \ldots, r^{(n)} \in (0, 1)$ such that the balls $\left\{ B\left(x^{(h)}; r^{(h)}\right) \right\}_{h=1}^n$ are pairwise disjoint. Let $f \in L^1(X, \mu)$. Then*

$$
\alpha_{B(\bar{x}, \bar{r})}(f) = \sum_{h=1}^n \frac{\mu\left(B\left(x^{(h)}; r^{(h)}\right)\right)}{\mu\left(B\left(x^{(1)}; r^{(1)}\right)\right) + \cdots + \mu\left(B\left(x^{(n)}; r^{(n)}\right)\right)} \alpha_{B\left(x^{(h)}; r^{(h)}\right)}(f).
$$

Proof.

$$\alpha_{B(\bar{x};,\bar{r})}(f) = \frac{1}{\mu(B(\bar{x};,\bar{r}))} \int_{B(\bar{x};\bar{r})} f \mathrm{d}\mu$$

$$= \sum_{h=1}^{n} \frac{1}{\sum_{u=1}^{n} \mu\left(B\left(x^{(u)};r^{(u)}\right)\right)} \int_{B\left(x^{(h)};r^{(h)}\right)} f \mathrm{d}\mu$$

$$= \sum_{h=1}^{n} \frac{\mu\left(B\left(x^{(h)};r^{(h)}\right)\right)}{\sum_{u=1}^{n} \mu\left(B\left(x^{(u)};r^{(u)}\right)\right)} \frac{1}{\mu\left(B\left(x^{(h)};r^{(h)}\right)\right)} \int_{B\left(x^{(h)};r^{(h)}\right)} f \mathrm{d}\mu$$

$$= \sum_{h=1}^{n} \frac{\mu\left(B\left(x^{(h)};r^{(h)}\right)\right)}{\mu\left(B\left(x^{(1)};r^{(1)}\right)\right) + \cdots + \mu\left(B\left(x^{(n)};r^{(n)}\right)\right)} \alpha_{B\left(x^{(h)};r^{(h)}\right)}(f)$$

$\square$

Theorem 4.2.7. *Let $(\bar{r}_k)_{k=1}^{\infty}$ be a sequence that decays $(X, \rho, H, L, \mathbf{F})$-fast, and let $f \in C(X)$. Suppose $\bar{x} = \left(x^{(1)}, \ldots, x^{(n)}\right)$ is an n-tuple in X such that*

$$\lim_{k \to \infty} \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) = C \qquad\qquad (\forall h \in \{1, \ldots, n\}),$$

where C is independent of h, and let μ be a Borel probability measure on X for which $x^{(1)}, \ldots, x^{(n)} \in \mathrm{supp}(\mu)$. Then

$$\lim_{k \to \infty} \alpha_{B(\bar{x};\bar{r}_k)} \left(\mathrm{Avg}_{F_k} f\right) = C.$$

Proof. By Lemma 4.2.2, we can assume without loss of generality that $x^{(1)}, \ldots, x^{(n)}$ are distinct. Because $r_k^{(h)} \to 0$, we know that for sufficiently large k, we'll have

$$B\left(\bar{x};\bar{r}_k\right) = \bigsqcup_{h=1}^{n} B\left(x^{(h)};r_k^{(h)}\right),$$

where $B(x; r) := \{y \in X : \rho(x, y) < r\}$ is the open ball with center x and radius r, and $\sqcup$ denotes disjoint union. We therefore estimate that

$$\left| \alpha_{B(\bar{x};\bar{r}_k)} \left(\mathrm{Avg}_{F_k} f \right) - C \right|$$

$$= \left| \frac{1}{\mu(B(\bar{x};\bar{r}_k))} \int_{B(\bar{x};\bar{r}_k)} \frac{1}{|F_k|} \sum_{g \in F_k} \left(f(T_g y) - C \right) \mathrm{d}\mu(y) \right|$$

$$= \left| \frac{1}{\mu(B(\bar{x};\bar{r}_k))} \sum_{h=1}^{n} \int_{B\left(x^{(h)};r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \left(f(T_g y) - C \right) \mathrm{d}\mu(y) \right|$$

$$\leq \sum_{h=1}^{n} \left| \frac{1}{\mu(B(\bar{x};\bar{r}_k))} \int_{B\left(x^{(h)};r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \left(f(T_g y) - C \right) \mathrm{d}\mu(y) \right|$$

$$\leq \sum_{h=1}^{n} \left| \frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)} \right) \right)} \int_{B\left(x^{(h)};r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \left(f(T_g y) - C \right) \mathrm{d}\mu(y) \right|$$

$$= \sum_{h=1}^{n} \left| \alpha_{B\left(x^{(h)};r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f - C \right) \right|$$

$$\leq \sum_{h=1}^{n} \left(\left| \alpha_{B\left(x^{(h)};r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f - \mathrm{Avg}_{F_k}(x) \right) \right| + \left| \alpha_{B\left(x^{(h)};r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f(x) - C \right) \right| \right)$$

$$= \sum_{h=1}^{n} \left(\left| \alpha_{B\left(x^{(h)};r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f - \mathrm{Avg}_{F_k} \left(x^{(h)} \right) \right) \right| + \left| \mathrm{Avg}_{F_k} f \left(x^{(h)} \right) - C \right| \right)$$

$$\overset{k \to \infty}{\to} 0,$$

where the limit in the last line follows from Lemma 4.2.5. $\qquad \square$

We recall here the following definition.

Definition 4.2.8. Let $(F_k)_{k=1}^{\infty}$ be a sequence of nonempty finite subsets of a group G. We say that $(F_k)_{k=1}^{\infty}$ is *tempered* if there exists a constant $c > 0$ such that

$$\left| \bigcup_{j=1}^{k-1} F_j^{-1} F_k \right| \leq c |F_k| \qquad\qquad (\forall k \geq 2).$$

Lemma 4.2.9. *Every Følner sequence $(F_k)_{k=1}^{\infty}$ has a tempered subsequence. In particular, every amenable group admits a tempered Følner sequence.*

Proof. (Lindenstrauss, 2001, Proposition 1.4) $\qquad \square$

The existence of tempered subsequences will be relevant to us in later sections.

Corollary 4.2.10. *Suppose G is an amenable group, and $\mathbf{F}$ is a tempered Følner sequence. Suppose further that μ is a Borel probability measure on X that is T-invariant and ergdic. Then for almost all $\bar{x} \in X^n$, we have for all $f \in C(X)$ and all sequences $(\bar{r}_k)_{k=1}^\infty$ that decay $(X, \rho, H, L, \mathbf{F})$-fast that*

$$\lim_{k \to \infty} \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f \right) = \int f \mathrm{d}\mu.$$

Proof. Since X is compact metrizable, it follows that $C(X)$ is separable, so let $\{f_\ell\}_{\ell \in \mathbb{N}}$ be a countable dense subset of $C(X)$. For each $\ell \in \mathbb{N}$, set

$$X_\ell = \left\{ x \in X : \mathrm{Avg}_{F_k} f_\ell(x) = \int f_\ell \mathrm{d}\mu \right\}.$$

By the Lindenstrauss ergodic theorem (Lindenstrauss, 2001, Theorem 3.3), each of these sets X_ℓ has full probability, and so $X' = \bigcap_{\ell \in \mathbb{N}} X_\ell$ also has full probability. Thus $(X')^n$ is of full probability in X^n with respect to the product measure $\underbrace{\mu \times \cdots \times \mu}_{n}$.

Let $\bar{x} \in (X')^n$, and let $(\bar{r}_k)_{k=1}^\infty$ be a sequence of n-tuples of positive numbers that decay $(X, \rho, H, L, \mathbf{F})$-fast. By Theorem 4.2.7, we know that $\lim_{k \to \infty} \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f_\ell \right) = \int f_\ell \mathrm{d}\mu$ for all $\ell \in \mathbb{N}$. Now it remains to prove that this convergence occurs for all $f \in C(X)$.

Let $f \in C(X)$, and fix $\epsilon > 0$. Choose f_ℓ such that $\|f - f_\ell\|_{C(X)} \leq \epsilon$. Then

$$
\begin{aligned}
&\left| \int f \mathrm{d}\mu - \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f \right) \right| \\
&\leq \left| \int f \mathrm{d}\mu - \int f_\ell \mathrm{d}\mu \right| + \left| \int f_\ell \mathrm{d}\mu - \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f_\ell \right) \right| + \left| \alpha_{B(\bar{x}, \bar{r}_k)} \left(\mathrm{Avg}_{F_k} (f_\ell - f) \right) \right| \\
&\leq \|f - f_\ell\|_{C(X)} + \left| \int f_\ell \mathrm{d}\mu - \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f_\ell \right) \right| + \|f - f_\ell\|_{C(X)} \\
&\leq 2\epsilon + \left| \int f_\ell \mathrm{d}\mu - \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f_\ell \right) \right|.
\end{aligned}
$$

Now choose $K \in \mathbb{N}$ such that if $k \geq K$, then $\left| \int f_\ell \mathrm{d}\mu - \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f_\ell \right) \right| \leq \epsilon$. Then for $k \geq K$, we have that

$$\left| \int f \mathrm{d}\mu - \alpha_{B(\bar{x}; \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f \right) \right| \leq 3\epsilon.$$

This demonstrates the convergence. □

Theorem 4.2.7 tells us that if we look at a sequence of concentric multiballs with rapidly vanishing radii, and if the pointwise Birkhoff averages at the centers converge to the same limit, then the temporo-spatial average with respect to these sequences of multiballs will inherit the limiting behavior f the pointwise Birkhoff averages. We might wonder whether Theorem 4.2.7 could be generalized by replacing the assumption that $\lim_{k\to\infty} \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) = C$ with $\limsup_{k\to\infty} \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) = C$, assuming of course that f was real-valued. It turns out this generalization fails, as the next example demonstrates.

Example 4.2.11. Let $X = \{0,1\}^{\mathbb{N}}$, and let μ be the Borel probability measure on X generated by

$$\mu\left([a_1,\ldots,a_\ell]\right) = 2^{-\ell}$$

for all $a_1,\ldots,a_\ell \in \{0,1\}, \ell \in \mathbb{N}$, where $[a_1,\ldots,a_\ell] = \{x \in X : x(1) = a_1,\ldots,x(\ell) = a_\ell\}$. Let $T_j :$ $\mathbb{N}_0 \curvearrowright X$ be the left shift $(Tx)(i) = x(i+j)$, where $\mathbb{N}_0$ denotes the semigroup of nonnegative integers, making (X,μ,T) a one-sided Bernoulli shift. Equip X with the compatible metric

$$\rho(x,y) = \begin{cases} 0 & \text{if } x = y, \\ 2^{-\ell} & \text{if } \ell = \min\{i \in \mathbb{N} : x(i) \neq y(i)\}. \end{cases}$$

Then $B\left(x; 2^{-k}\right) = [x(1),\ldots,x(k)]$, and T_k is 2^k-Lipschitz, i.e. $\rho\left(T_k x, T_k y\right) \leq 2^k \cdot \rho(x,y)$. Set $L(j) = 2^j, H(j) = 1$. We can check that $\left(2^{-k}\right)_{k=1}^\infty$ decays $(X,\rho,H,L,\mathbf{F})$-fast for $\mathbf{F} = (\{0,1,\ldots,k-1\})_{k=1}^\infty$ by observing that for any $\delta > 0$, if $2^{j-k} \geq \delta$ for $\delta \in (0,1), 0 \leq j \leq k-1$, then $j - k \geq \log_2 \delta \iff j \geq k + \log_2 \delta$. Therefore $L(j) \cdot \left(2^{-k}\right)^{H(j)} < \delta$ for all but at most $\lceil |\log_2 \delta| \rceil$ of $j \in \{0,1,\ldots,k-1\}$, so

$$\frac{\left|\left\{j \in F_k : L(j) \cdot \left(2^{-k}\right)^{H(j)} \geq \delta\right\}\right|}{|F_k|} \leq \frac{\lceil |\log_2 \delta| \rceil}{k} \xrightarrow{k\to\infty} 0.$$

Let $(c_n)_{n=1}^\infty$ be a sequence of natural numbers chosen to grow fast enough that

$$\frac{c_n}{c_1 + \cdots + c_n} \geq \frac{n-1}{n} \qquad\qquad (\forall n \in \mathbb{N}).$$

Set $s_n = c_1 + \cdots + c_n$, so our growth condition states that $\frac{c_n}{s_n} \geq \frac{n-1}{n}$. Now construct $x \in X$ by

$$
x(i) = \begin{cases}
0 & 1 \leq i \leq s_1 \\
1 & s_1 < i \leq s_2, \\
0 & s_2 < i \leq s_3, \\
\cdots \\
0 & s_{2n} < i \leq s_{2n+1} \\
1 & s_{2n+1} < i \leq s_{2n+2}. \\
\cdots
\end{cases}
$$

In plain language, this x consists of c_1 terms of 0, then c_2 terms of 1, then c_3 terms of 0, then c_4 terms of 1, etc. We then define $y \in X$ by

$$
y(i) = 1 - x(i) \qquad\qquad (\forall i \in \mathbb{N}),
$$

i.e. replacing all 0's with 1's and vice-versa. Set $f = \chi_{[0]}$. We claim that $\limsup_{k\to\infty} \mathrm{Avg}_{F_k} f(x) = \limsup_{k\to\infty} \mathrm{Avg}_{F_k}(y) = 1$.

Consider the case where we sample along $(s_{2n-1})_{n=1}^{\infty}$. Then

$$
\mathrm{Avg}_{F_{s_{2n-1}}} f(x) = \frac{c_1 + c_3 + c_5 + \cdots + c_{2n-1}}{c_1 + c_2 + c_3 + c_4 + c_5 + c_6 + \cdots + c_{2n-1}} \geq \frac{c_{2n-1}}{s_{2n-1}} \geq \frac{2n-2}{2n-1} \stackrel{n\to\infty}{\to} 1.
$$

But $\mathrm{Avg}_{F_k} f(z) \in [0,1]$ for all $k \in \mathbb{N}, z \in X$, so we can conclude that $\limsup_{k\to\infty} \mathrm{Avg}_{F_k} f(x) = 1$. Likewise, sampling along s_{2n}, we see that

$$
\mathrm{Avg}_{F_{s_{2n}}} f(y) = \frac{c_2 + c_4 + \cdots + c_{2n}}{s_{2n}} \geq \frac{c_{2n}}{s_{2n}} \geq \frac{2n-1}{2n} \stackrel{n\to\infty}{\to} 1.
$$

Thus $\limsup_{k\to\infty} \mathrm{Avg}_{F_k}(y) = 1$.

Computing the temporo-spatial averages, we can see that

$$\alpha_{B\left(x,y;2^{-k},2^{-k}\right)}\left(\mathrm{Avg}_{F_k}\, f\right)$$

$$=\frac{1}{\mu\left(B\left(x,y;2^{-k},2^{-k}\right)\right)}\int_{B\left(x,y;2^{-k},2^{-k}\right)}\frac{1}{k}\sum_{j=0}^{k-1}\chi_{[0]}(T_j z)\mathrm{d}\mu(z)$$

$$=\frac{1}{2^{-k}+2^{-k}}\frac{1}{k}\sum_{j=0}^{k-1}\chi_{[0]}(T_j z)\mathrm{d}\mu(z)$$

$$=2^{k-1}\int_{[x(1),\dots,x(k)]\sqcup[y(1),\dots,y(k)]}\frac{1}{k}\sum_{j=0}^{k-1}\chi_{[0]}(T_j z)\mathrm{d}\mu(z)$$

$$=2^{k-1}\int\frac{1}{k}\sum_{j=0}^{k-1}\left(\chi_{[x(1),\dots,x(k)]}(z)+\chi_{[y(1),\dots,y(k)]}(z)\right)\chi_{[0]}(T_j z)\mathrm{d}\mu(z)$$

$$=2^{k-1}\int\frac{1}{k}\sum_{j=0}^{k-1}\left(\chi_{[x(1),\dots,x(k)]}(z)+\chi_{[y(1),\dots,y(k)]}(z)\right)\chi_{T^{-j}[0]}(z)\mathrm{d}\mu(z)$$

$$=2^{k-1}\int\frac{1}{k}\sum_{j=0}^{k-1}\left(\chi_{[x(1),\dots,x(k)]\cap T^{-j}[0]}(z)+\chi_{[y(1),\dots,y(k)]\cap T^{-j}[0]}(z)\right)\mathrm{d}\mu(z)$$

We know that

$$\chi_{[x(1),\dots,x(k)]\cap T^{-j}[0]}(z)=\begin{cases}1 & \text{if } z(1+j)=x(1+j)=0,\\ 0 & \text{otherwise},\end{cases}$$

$$\chi_{[y(1),\dots,y(k)]\cap T^{-j}[0]}(z)=\begin{cases}1 & \text{if } z(1+j)=y(1+j)=0,\\ 0 & \text{otherwise}.\end{cases}$$

Thus

$$\int\chi_{[x(1),\dots,x(k)]\cap T^{-j}[0]}(z)\mathrm{d}\mu(z)=\begin{cases}2^{-k} & \text{if } x(1+j)=0,\\ 0 & \text{if },\end{cases}$$

$$\int\chi_{[y(1),\dots,y(k)]\cap T^{-j}[0]}(z)=\begin{cases}2^{-k} & \text{if } z(1+j)=y(1+j)=0,\\ 0 & \text{otherwise}.\end{cases}$$

But since $x(1+j) = 0 \iff y(i+j) = 1$, it follows that

$$\int \left(\chi_{[x(1),\dots,x(k)] \cap T^{-j}[0]}(z) + \chi_{[y(1),\dots,y(k)] \cap T^{-j}[0]}(z) \right) \mathrm{d}\mu(z) = 2^{-k}$$

for all $j = 0, 1, \dots, k-1$. Therefore

$$\alpha_{B\left(x,y;2^{-k},2^{-k}\right)} \left(\mathrm{Avg}_{F_k} f \right)$$
$$= 2^{k-1} \int \frac{1}{k} \sum_{j=0}^{k-1} \left(\chi_{[x(1),\dots,x(k)] \cap T^{-j}[0]}(z) + \chi_{[y(1),\dots,y(k)] \cap T^{-j}[0]}(z) \right) \mathrm{d}\mu(z)$$
$$= 2^{k-1} \frac{1}{k} \sum_{j=0}^{k-1} 2^{-k}$$
$$= \frac{1}{2}.$$

So $\limsup_{k\to\infty} \alpha_{B\left(x,y;2^{-k},2^{-k}\right)} \left(\mathrm{Avg}_{F_k} f \right) = 1/2 \neq 1$.

Example 4.2.12. Looking at Theorem 4.2.7, we could also ask whether the result could be generalized to somehow accommodate the case where $\lim_{k\to\infty} \mathrm{Avg}_{F_k} f \left(x^{(h)} \right)$ exists for all $h = 1, \dots, n$, but is allowed to vary with h. However, we can construct examples of points $x, y \in X$, sequences of radii $(r_k)_{k=1}^\infty, (s_k)_{k=1}^\infty \in (0,1)^{\mathbb{N}}$ decaying $(X, \rho, H, L, \mathbf{F})$-fast, and a function $f \in C(X)$ where $\lim_{k\to\infty} \mathrm{Avg}_{F_k} f(x), \lim_{k\to\infty} \mathrm{Avg}_{F_k} f(y)$ both exist, but $\lim_{k\to\infty} \alpha_{B\left(x,y;2^{-k},2^{-k}\right)} \left(\mathrm{Avg}_{F_k} f \right)$ does not. Let $X, \rho, T, \mu, \mathbf{F}$ be as in Example 4.2.11, but choose x, y to be

$$x(i) = \begin{cases} 0 & \text{if } i \text{ is even,} \\ 1 & \text{if } i \text{ is odd,} \end{cases} \qquad y(i) = \begin{cases} 0 & \text{if } i \text{ is divisible by 3,} \\ 1 & \text{otherwise.} \end{cases}$$

Let $f = \chi_{[0]}$. Then $\lim_{k\to\infty} \mathrm{Avg}_{F_k} f(x) = 1/2, \lim_{k\to\infty} \mathrm{Avg}_{F_k} f(y) = 1/3$. Construct sequences of natural numbers $(p_k)_{k=1}^\infty, (q_k)_{k=1}^\infty$ strictly increasing such that

$$\frac{2^{-p_k}}{2^{-p_k} + 2^{-q_k}} \geq \frac{k}{k+1} \qquad \text{(for } k \text{ odd),}$$
$$\frac{2^{-q_k}}{2^{-p_k} + 2^{-q_k}} \geq \frac{k}{k+1} \qquad \text{(for } k \text{ even).}$$

115

Therefore
$$\lim_{n\to\infty} \frac{2^{-p_{2n-1}}}{2^{-p_{2n-1}} + 2^{-q_{2n-1}}} = \frac{2^{-q_{2n}}}{2^{-p_{2n}} + 2^{-q_{2n}}} = 1.$$

Set $r_k = 2^{-p_k}$, $s_k = 2^{-q_k}$. We can see that $(r_k, s_k)_{k=1}^\infty$ decays $(X, \rho, H, L, \mathbf{F})$-fast. By Lemma 4.2.6, we have

$$\alpha_{B(x,y;r_k,s_k)} \left(\mathrm{Avg}_{F_k} f \right) = \frac{2^{-p_k}}{2^{-p_k} + 2^{-q_k}} \alpha_{B(x;r_k)} \left(\mathrm{Avg}_{F_k} f \right) + \frac{2^{-q_k}}{2^{-p_k} + 2^{-q_k}} \alpha_{B(y;s_k)} \left(\mathrm{Avg}_{F_k} f \right).$$

Sampling along even k, we see that

$$\lim_{n\to\infty} \alpha_{B(x,y;r_{2n},s_{2n})} \left(\mathrm{Avg}_{F_{2n}} f \right)$$
$$= \lim_{n\to\infty} \frac{2^{-p_{2n}}}{2^{-p_{2n}} + 2^{-q_{2n}}} \alpha_{B(x;r_{2n})} \left(\mathrm{Avg}_{F_{2n}} f \right) + \frac{2^{-q_{2n}}}{2^{-p_{2n}} + 2^{-q_{2n}}} \alpha_{B(y;s_{2n})} \left(\mathrm{Avg}_{F_{2n}} f \right)$$
$$= 0 \left(\lim_{n\to\infty} \alpha_{B(x;r_{2n})} \left(\mathrm{Avg}_{F_{2n}} f \right) \right) + 1 \left(\lim_{n\to\infty} \alpha_{B(y;s_{2n})} \left(\mathrm{Avg}_{F_{2n}} f \right) \right)$$
$$= \frac{1}{3},$$

where the limits in the last step are taken using Lemma 4.2.5. On the other hand, sampling along odd k, we see that

$$\lim_{n\to\infty} \alpha_{B(x,y;r_{2n-1},s_{2n-1})} \left(\mathrm{Avg}_{F_{2n-1}} f \right)$$
$$= \lim_{n\to\infty} \frac{2^{-p_{2n-1}}}{2^{-p_{2n-1}} + 2^{-q_{2n-1}}} \alpha_{B(x;r_{2n-1})} \left(\mathrm{Avg}_{F_{2n-1}} f \right)$$
$$+ \lim_{n\to\infty} \frac{2^{-q_{2n-1}}}{2^{-p_{2n-1}} + 2^{-q_{2n-1}}} \alpha_{B(y;s_{2n-1})} \left(\mathrm{Avg}_{F_{2n-1}} f \right)$$
$$= 1 \left(\lim_{n\to\infty} \alpha_{B(x;r_{2n-1})} \left(\mathrm{Avg}_{F_{2n-1}} f \right) \right) + 0 \left(\lim_{n\to\infty} \alpha_{B(y;s_{2n-1})} \left(\mathrm{Avg}_{F_{2n-1}} f \right) \right)$$
$$= \frac{1}{2},$$

where we again appeal to Lemma 4.2.5 to take the limits at the end. Thus the sequence

$$\left(\alpha_{B(x,y;r_k,s_k)} \left(\mathrm{Avg}_{F_k} f \right) \right)_{k=1}^\infty$$

is divergent.

The argument employed in Example 4.2.12, where we control the "weight" we give several points at different points in the temporo-spatial differentiation, will have applications in Sections 4.4 and 4.5. However, the following result also demonstrates that absent such tricks, we have predictable convergence behaviors.

Theorem 4.2.13. *Let* $(\bar{r}_k)_{k=1}^{\infty}$ *be a sequence that decays* $(X, \rho, H, L, \mathbf{F})$-*fast, and let* $f \in C(X)$. *Suppose* $\bar{x} = \left(x^{(1)}, \ldots, x^{(n)}\right)$ *is an* n-*tuple in* X *such that*

$$C_h = \lim_{k \to \infty} \mathrm{Avg}_{F_k} f\left(x^{(h)}\right)$$

exists for all $h = 1, \ldots, n$. *Let* μ *be a Borel probability measure on* X *for which* $x^{(1)}, \ldots, x^{(n)} \in \mathrm{supp}(\mu)$. *Suppose further that*

$$D_h = \lim_{k \to \infty} \frac{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)}{\mu\left(B\left(x^{(1)}; r_k^{(1)}\right)\right) + \cdots + \mu\left(B\left(x^{(n)}; r_k^{(n)}\right)\right)}$$

exists for all $h = 1, \ldots, n$. *Then*

$$\lim_{k \to \infty} \alpha_{B(\bar{x}; \bar{r}_k)}\left(\mathrm{Avg}_{F_k} f\right) = \sum_{h=1}^{n} D_h C_h.$$

Proof. This follows immediately from Lemmas 4.2.5, 4.2.6. $\qquad \square$

4.3 Preliminaries from ergodic optimization

Here we prove a generalization of a result of O. Jenkinson (Jenkinson, 2006a, Proposition 2.1) to the setting of actions of amenable topological groups. Our method of proof closely resembles Jenkinson's, but requires that we attend to a few extra details.

Throughout this section, $T : G \curvearrowright X$ will be an action of a discrete amenable group G on a compact metrizable space X by homeomorphisms, and $f \in C_{\mathbb{R}}(X)$ will be a real-valued continuous function on X. Let $\mathbf{F} = (F_k)_{k=1}^{\infty}$ be a Følner sequence for G. Define the set $\mathrm{Reg}(f)$ by

$$\mathrm{Reg}(f) = \left\{ x \in X : \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x) \text{ exists} \right\}.$$

We define the following values:

$$\overline{a}(f) := \sup \left\{ \int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X) \right\},$$

$$\underline{a}(f) := \inf \left\{ \int f \mathrm{d}\nu : \nu \in \mathcal{M}_T(X) \right\},$$

$$\overline{b}(f) := \sup \left\{ \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in \mathrm{Reg}(f) \right\},$$

$$\underline{b}(f) := \inf \left\{ \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in \mathrm{Reg}(f) \right\},$$

$$\overline{c}(f) := \sup \left\{ \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in X \right\},$$

$$\underline{c}(f) := \inf \left\{ \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in X \right\},$$

$$\overline{d}(f) := \lim_{k \to \infty} \left(\sup \left\{ \mathrm{Avg}_{F_k} f(x) : x \in X \right\} \right)$$

$$\underline{d}(f) := \lim_{k \to \infty} \left(\inf \left\{ \mathrm{Avg}_{F_k} f(x) : x \in X \right\} \right).$$

We write $\overline{b}(f) = -\infty, \underline{b}(f) = +\infty$ if $\mathrm{Reg}(f) = \emptyset$. We will show in Theorem 4.3.3 that $\underline{d}(f), \overline{d}(f)$ are well-defined.

The following result is elementary, but will be relevant for much of this chapter, so we state and prove it here.

Lemma 4.3.1. *Let $(F_k)_{k=1}^\infty$ be a Følner sequence for a group G, and let $(\beta_k)_{k=1}^\infty$ be a sequence of Borel probability measures on X. Then if $k_1 < k_2 < \cdots$ is a sequence of natural numbers such that $\nu = \lim_{\ell \to \infty} \beta_{k_\ell} \circ \mathrm{Avg}_{F_{k_\ell}}$ exists, then $\nu \in \mathcal{M}_T(X)$. In particular, if G is amenable, then $\mathcal{M}_T(X) \neq \emptyset$.*

Proof. Assume WLoG that $k_\ell = \ell$ for all $\ell \in \mathbb{N}$. Let $f \in C(X), g \in G$.

$$\left| \int f \mathrm{d}\nu - \int T_g f \mathrm{d}\nu \right| = \lim_{k \to \infty} \left| \left(\frac{1}{|F_k|} \sum_{h \in F_k} \int T_h f \mathrm{d}\beta_k \right) - \left(\frac{1}{|F_k|} \sum_{h' \in gF_k} \int T_{h'} f \mathrm{d}\beta_k \right) \right|$$

$$= \lim_{k \to \infty} \frac{1}{|F_k|} \left| \left(\sum_{h \in F_k \setminus gF_k} T_h f \right) - \left(\sum_{h' \in gF_k \setminus F_k} T_{h'} f \right) \right|$$

$$\leq \limsup_{k \to \infty} \frac{|F_k \Delta g F_k|}{|F_k|} \|f\|_{C(X)}$$

$$= 0.$$

To prove that $\mathcal{M}_T(X) \neq \emptyset$, consider any Borel probability measure β on X, and use the weak*-compactness of $\mathcal{M}(X)$ to extract a convergent subsequence from $\left(\beta \circ \operatorname{Avg}_{F_k}\right)_{k=1}^{\infty}$. The limit of that convergent subsequence will be T-invariant. $\square$

Definition 4.3.2. Let $\nu \in \mathcal{M}_T(X)$, and $f \in C(X)$. A point $x \in X$ is called $(f, \mathbf{F}, \nu)$-*typical* if $\lim_{k \to \infty} \operatorname{Avg}_{F_k} f(x) = \int f \mathrm{d}\nu$.

Theorem 4.3.3. *Suppose $f \in C_\mathbb{R}(X)$. Then the values $\overline{a}(f), \underline{a}(f), \overline{c}(f), \underline{c}(f), \overline{d}(f), \underline{d}(f)$ are all well-defined real numbers, and*

$$\overline{b}(f) \leq \overline{c}(f) = \overline{a}(f) = \overline{d}(f),$$

$$\underline{b}(f) \geq \underline{c}(f) = \underline{a}(f) = \underline{d}(f).$$

Furthermore, if for every ergodic measure $\theta \in \partial_e \mathcal{M}_T(X)$ exists an $(f, \mathbf{F}, \theta)$-typical point, then

$$\overline{a}(f) = \overline{b}(f) = \overline{c}(f) = \overline{d}(f), \ \underline{a}(f) = \underline{b}(f) = \underline{c}(f) = \underline{d}(f).$$

Proof. We will only prove the inequalities and identities for $\overline{a}, \overline{b}, \overline{c}, \overline{d}$, since the analogous relations between $\underline{a}, \underline{b}, \underline{c}, \underline{d}$ can be proven in a parallel fashion.

The well-definedness of $\overline{a}(f)$ follows from the weak*-compactness of $\mathcal{M}_T(X)$. We also know a priori that $\overline{c}(f) \leq \|f\|_{C(X)}$, and thus $\overline{c}(f)$ is well-defined.

It still remains to prove that $\overline{d}(f)$ is well-defined, which we will accomplish by proving that $\overline{d}(f) = \overline{a}(f)$.

For each $k \in \mathbb{N}$, choose $x_k \in X$ such that $\operatorname{Avg}_{F_k} f(x_k) = \sup\left\{\operatorname{Avg}_{F_k} f(x) : x \in X\right\}$. Let μ_k be the Borel probability measure on X defined by

$$\int g \mathrm{d}\mu_k = \operatorname{Avg}_{F_k} f(x_k).$$

Let $(\mu_{k_\ell})_{\ell=1}^{\infty}$ be a weak*-convergent subsequence converging to the measure μ. Then since $\mathbf{F}$ is Følner, it follows from Lemma 4.3.1 that μ is T-invariant. Thus

$$\overline{a}(f) \geq \int f \mathrm{d}\mu = \lim_{\ell \to \infty} \int \operatorname{Avg}_{F_{k_\ell}} f \mathrm{d}\mu_{k_\ell} = \lim_{\ell \to \infty} \left(\sup\left\{\operatorname{Avg}_{F_{k_\ell}} f(x) : x \in X\right\}\right).$$

On the other hand, we know that if $\nu \in \mathcal{M}_T(X)$, then

$$\int f \mathrm{d}\nu = \int \mathrm{Avg}_{F_{k_\ell}} f \mathrm{d}\nu \leq \sup \left\{ \mathrm{Avg}_{F_{k_\ell}} f(x) : x \in X \right\},$$

and thus taking $\ell \to \infty$ tells us that $\int f \mathrm{d}\nu \leq \int f \mathrm{d}\mu$. Therefore this measure μ is f-maximizing, meaning that $\overline{a}(f) = \int f \mathrm{d}\mu = \lim_{\ell \to \infty} \left(\sup \left\{ \mathrm{Avg}_{F_{k_\ell}} f(x) : x \in X \right\} \right)$. Since we know this holds true for *any* weak*-convergent subsequence $(\mu_{k_\ell})_{\ell=1}^\infty$, and $(\mu_k)_{k=1}^\infty$ takes values in the weak*-compact space $\mathcal{M}(X)$, we can conclude that $\overline{d}(f)$ is well-defined and equal to $\overline{a}(f)$.

It follows immediately from the definitions that $\overline{b}(f) \leq \overline{c}(f)$, since

$$\begin{aligned}
\overline{b}(f) &= \sup \left\{ \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in \mathrm{Reg}(f) \right\} \\
&= \sup \left\{ \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in \mathrm{Reg}(f) \right\} \\
&\leq \sup \left\{ \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in X \right\} \qquad\qquad = \overline{c}(f).
\end{aligned}$$

It similarly follows from definitions that $\overline{c}(f) \leq \overline{d}(f)$, since

$$\begin{aligned}
\overline{c}(f) &= \sup \left\{ \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x) : x \in X \right\} \\
&\leq \sup \left\{ \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x_k) : (x_k)_{k=1}^\infty \in X^{\mathbb{N}} \right\} \\
&\leq \sup \left\{ \limsup_{k \to \infty} \left(\sup \left\{ \mathrm{Avg}_{F_k} f(x) : x \in X \right\} \right) : (x_k)_{k=1}^\infty \in X^{\mathbb{N}} \right\} \\
&= \limsup_{k \to \infty} \left(\sup \left\{ \mathrm{Avg}_{F_k} f(x) : x \in X \right\} \right) \qquad\qquad = \overline{d}(f).
\end{aligned}$$

Next we show that $\overline{a}(f) \leq \overline{c}(f)$. Let $k_1 < k_2 < \cdots$ such that $(F_{k_\ell})_{\ell=1}^\infty$ is a *tempered* Følner subsequence, a subsequence which exists by Lemma 4.2.9. Let $\theta \in \partial_e \mathcal{M}_T(X)$. Then by the Lindenstrauss Ergodic Theorem, there exists $x \in X$ such that $\lim_{\ell \to \infty} \mathrm{Avg}_{F_{k_\ell}} f(x) = \int f \mathrm{d}\theta$. Therefore

$$\int f \mathrm{d}\theta = \lim_{\ell \to \infty} \mathrm{Avg}_{F_{k_\ell}} f(x) \leq \limsup_{k \to \infty} \mathrm{Avg}_{F_k} f(x) \leq \overline{c}(f).$$

Suppose $\nu \in \mathcal{M}_T(X)$, and let $(\theta_x)_{x \in X}$ be the ergodic decomposition of $T : G \curvearrowright X$. Then

$$\int f \mathrm{d}\nu = \int \left(\int f \mathrm{d}\theta_x \right) \mathrm{d}\nu(x) \leq \int \overline{c}(f) \mathrm{d}\nu(x) = \overline{c}(f).$$

Taking the supreumum over $\nu \in \mathcal{M}_T(X)$ confirms that $\overline{a}(f) \leq \overline{c}(f)$.

Now assume that for every ergodic measure $\theta \in \partial_e \mathcal{M}_T(X)$ exists $x_\theta \in X$ such that $\int f \mathrm{d}\theta = \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x_\theta)$. We prove that $\overline{a}(f) \leq \overline{b}(f)$. To begin, we'll prove that $\int f \mathrm{d}\theta \leq \overline{b}(f)$ for all ergodic $\theta \in \partial_e \mathcal{M}_T(X)$, and then use the ergodic decomposition to extrapolate to the general case.

First, consider the case where θ is an ergodic measure in $\mathcal{M}_T(X)$. Then there exists $x_\theta \in X$ such that

$$\int f \mathrm{d}\theta = \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x_\theta) \leq \overline{b}(f).$$

Now suppose $\nu \in \mathcal{M}_T(X)$, and let $(\theta_x)_{x \in X}$ be the ergodic decomposition of $T : G \curvearrowright X$. Then

$$\int f \mathrm{d}\nu = \int \left(\int f \mathrm{d}\theta_x \right) \mathrm{d}\nu(x) \leq \int \overline{b}(f) \mathrm{d}\nu(x) = \overline{b}(f).$$

Taking the supremum over $\nu \in \mathcal{M}_T(X)$ confirms that $\overline{a}(f) \leq \overline{b}(f)$. $\square$

What remains unclear to us at this point is whether $\overline{a}(f) \leq \overline{b}(f)$, $\underline{b}(f) \leq \underline{a}(f)$ in general. However, there are several general cases where we know the answer to be yes.

- If $\underline{a}(f) = \overline{a}(f)$, then every $x \in X$ is an $(f, \mathbf{F}, \nu)$-typical point for all $\nu \in \mathcal{M}_T(X)$. In particular, this will occur for all $f \in C_{\mathbb{R}}(X)$ if $T : G \curvearrowright X$ is uniquely ergodic.

- If $\mathbf{F}$ is tempered, then the Lindenstrauss Ergodic Theorem implies that the set of $(f, \mathbf{F}, \theta)$-typical points is of probability 1 with respect to θ for ergodic θ, and a fortiori, that the set is nonempty. This holds in particular if $G = \mathbb{Z}$ and $F_k = \{0, 1, \ldots, k-1\}$ for all $k \in \mathbb{N}$, which is the setting of the classical Birkhoff Ergodic Theorem.

Corollary 4.3.4. *The values $\overline{c}(f), \underline{c}(f), \overline{d}(f), \underline{d}(f)$ are independent of the choice of Følner sequence $\mathbf{F}$, and $\overline{b}(f), \underline{b}(f)$ are independent of the choice of* tempered *Følner sequence.*

Proof. The first claim follows from the fact that $\overline{a}(f), \underline{a}(f)$ are independent of $\mathbf{F}$, combined with Theorem 4.3.3. The second claim follows from the fact that if $\mathbf{F}$ is a *tempered* Følner sequence, then by the Linden-

strauss Ergodic Theorem, every ergodic measure $\theta \in \partial_e \mathcal{M}_T(X)$ admits an $(f, \mathbf{F}, \theta)$-typical point, meaning Theorem 4.3.3 tells us that $\overline{b}(f) = \overline{a}(f), \underline{b}(f) = \underline{a}(f)$. $\qquad\square$

4.4   Pathological multi-local temporo-spatial differentiations of individual functions

This section is motivated by the following question: Given a real-valued function $f \in C_{\mathbb{R}}(X)$, what possible sets $\mathcal{K}$ can be realized as

$$\mathcal{K} = \left\{ \lim_{\ell \to \infty} \alpha_{B\left(x, y; r_{k_\ell}, s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) : k_1 < k_2 < \cdots, \ \lim_{\ell \to \infty} \alpha_{B\left(x, y; r_{k_\ell}, s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) \text{ exists} \right\}$$

through judicious choices of $(\bar{x}; \bar{r}_k)_{k=1}^\infty$? If $\mathcal{K}$ is non-singleton, then the temporo-spatial differentiation will of course be divergent.

Before constructing these pathological temporo-spatial differentiations, we define a measure-theoretic property which will be important to us in this section.

Definition 4.4.1. Let (X, ρ) be a compact metric space, and let μ be a Borel probability measure on X. We say that μ *neglects shells* if

$$\mu \left(\{ y \in X : \rho(x, y) = r \} \right) = 0 \qquad\qquad (\forall x \in X, r \in [0, \infty)).$$

A probability measure which neglects shells is automatically non-atomic, but the converse is false. Consider the case of $X = \left\{ (a, b) \in \mathbb{R}^2 : a^2 + b^2 \leq 2 \right\}$ with the standard Euclidean metric. Let μ be the Borel probability measure

$$\mu(E) = \frac{1}{\mathcal{H}^1(S)} \mathcal{H}^1(S \cap E),$$

where $\mathcal{H}^1$ is the 1-dimensional Hausdorff measure and $S = \left\{ (a, b) \in \mathbb{R}^2 : a^2 + b^2 = 1 \right\}$ is the unit circle in $\mathbb{R}^2$. Then this μ is non-atomic, but does not neglect shells.

Theorem 4.4.2. *The following conditions are equivalent.*

(i) The function $\phi : X \times [0, \infty) \to [0, 1]$ defined by

$$\phi(x, r) = \mu(B(x; r))$$

is continuous.

(ii) μ *neglects shells.*

Proof. (i)$\Rightarrow$(ii): Suppose that ϕ is continuous, and fix $x \in X, r \in [0, \infty)$. Let $r_k = r + 1/k$ for all $k \in \mathbb{N}$. By downward continuity of measures, we know that

$$\lim_{k \to \infty} \phi(x, r_k) = \mu\left(\{y \in X : \rho(x, y) \leq r\}\right) = \phi(x, r) + \mu(\{y \in X : \rho(x, y) = r\}).$$

If $\lim_{k \to \infty} \phi(x, r_k) = \phi(x, r)$, then $\mu(\{y \in X : \rho(x, y) = r\}) = 0$.

(ii)$\Rightarrow$(i): Suppose that μ neglects shells, and let $(x_k, r_k)_{k=1}^{\infty}$ be a sequence in $X \times [0, \infty)$ converging to (x, r). Let $f_k, f \in L^{\infty}(X, \mu)$ be the functions

$$f_k = \chi_{B(x_k; r_k)},$$

$$f = \chi_{B(x; r)}.$$

We claim that $f_k \to f$ pointwise on $\{y \in X : \rho(x, y) \neq r\}$, which under the assumption that μ neglects shells constitutes convergence pointwise almost everywhere. If we can prove that, then we can appeal to the Dominated Convergence Theorem (using the constant function 1 as a dominator) to conclude that $\phi(x_k, r_k) = \int f_k \mathrm{d}\mu \overset{k \to \infty}{\to} \int f \mathrm{d}\mu = \phi(x, r)$, i.e. that ϕ is (sequentially) continuous.

First, consider the case where $\rho(x, y) < r$. Set $\epsilon = r - \rho(x, y)$. Then there exist $K_1, K_2 \in \mathbb{N}$ such that

$$k \geq K_1 \qquad\qquad\qquad \Rightarrow |r_k - r| < \frac{\epsilon}{2},$$

$$k \geq K_2 \qquad\qquad\qquad \Rightarrow \rho(x_k, x) < \frac{\epsilon}{2}.$$

If $k \geq K_1$, then $r_k > r - \frac{\epsilon}{2}$. Set $K = \max\{K_1, K_2\}$, and suppose that $k \geq K$. Then

$$\rho(y, x_k) \leq \rho(y, x) + \rho(x, x_k)$$

$$< \rho(y, x) + \frac{\epsilon}{2}$$

$$= r - \frac{\epsilon}{2}$$

$$< r_k.$$

123

Thus if $k \geq K$, then $f_k(y) = 1 = f(y)$. Therefore $\lim_{k \to \infty} f_k(y) = f(y)$ for $y \in B(x; r)$.

Second, consider the case where $\rho(x, z) > r$. Set $\delta = \min\left\{\rho(x, z) - r, \frac{\rho(x,z)}{2}\right\}$, and choose $L_1, L_2 \in \mathbb{N}$ such that

$$k \geq L_1 \qquad\qquad\qquad \Rightarrow |r_k - r| < \frac{\delta}{2},$$
$$k \geq L_2 \qquad\qquad\qquad \Rightarrow \rho(x_k, x) < \frac{\delta}{2}.$$

Set $L = \max\{L_1, L_2\}$, and consider $k \geq L$. Then

$$\rho(z, x_k) \geq |\rho(z, x) - \rho(x, x_k)|$$
$$= \rho(z, x) - \rho(x, x_k)$$
$$> \rho(z, x) - \frac{\delta}{2}$$
$$> r + \delta - \frac{\delta}{2}$$
$$= r + \frac{\delta}{2}$$
$$> r_k.$$

Thus if $k \geq L$, then $f_k(z) = 0 = f(z)$. Therefore $\lim_{k \to \infty} f_k(z) = f(z)$ for $\rho(z, x) > r$. This completes the proof. $\qquad\qquad\square$

The property of neglecting shells is very important to us in this chapter because of Lemma 4.4.3, which is a valuable tool for several constructions that will follow in this section and the next.

Lemma 4.4.3. *Let μ be a Borel probability measure on X that neglects shells, and let $x^{(1)}, \ldots, x^{(n)} \in \text{supp}(\mu)$. Let $\delta^{(1)}, \ldots, \delta^{(n)} > 0$, and fix $\lambda^{(1)}, \ldots, \lambda^{(n)} \in (0, 1)$ such that $\lambda^{(1)} + \cdots + \lambda^{(n)} = 1$. Then there exist $r^{(1)}, \ldots, r^{(n)} > 0$ such that $0 < r^{(h)} < \delta^{(h)}$, and*

$$\frac{\mu\left(B\left(x^{(h)}; r^{(h)}\right)\right)}{\mu\left(B\left(x^{(1)}; r^{(1)}\right)\right) + \cdots + \mu\left(B\left(x^{(n)}; r^{(n)}\right)\right)} = \lambda^{(h)} \qquad\qquad (h = 1, \ldots, n).$$

Proof. Assume without loss of generality that

$$\delta^{(h)} < \min_{1 \leq i < j \leq n} \rho\left(x^{(i)}, x^{(j)}\right),$$

124

otherwise we can replace each $\delta^{(h)}$ with $\min \left\{ \delta^{(h)}, \frac{1}{4} \min_{1 \leq i < j \leq n} \rho \left(x^{(i)}, x^{(j)} \right) \right\}$.

Choose real numbers $a^{(1)}, \ldots, a^{(h)} \in (0,1)$ such that

$$\frac{a^{(h)}}{a^{(1)} + \cdots + a^{(n)}} = \lambda^{(h)},$$

$$a^{(1)} < \mu \left(B \left(x^{(h)}; \delta^{(h)} \right) \right)$$

for all $h = 1, \ldots, n$. The tuple $\left(a^{(1)}, \ldots, a^{(n)} \right) \in (0,1)^n$ can be found along the line segment $\left\{ \left(t\lambda^{(1)}, \ldots, t\lambda^{(n)} \right) : t \in (0,1) \right\}$. We know that $\mu \left(B \left(x^{(h)}; \delta^{(h)} \right) \right) > 0$ because we assumed that $x^{(h)} \in \mathrm{supp}(\mu)$. Then by Theorem 4.4.2 and the Intermediate Value Theorem, there exist $r^{(h)} \in \left(0, \delta^{(h)} \right)$ such that

$$\mu \left(B \left(x^{(h)}; r^{(h)} \right) \right) = a^{(h)} \qquad\qquad (h = 1, \ldots, n),$$

and therefore

$$\frac{\mu \left(B \left(x^{(h)}; r^{(h)} \right) \right)}{\mu \left(B \left(x^{(1)}; r^{(1)} \right) \right) + \cdots + \mu \left(B \left(x^{(n)}; r^{(n)} \right) \right)} = \lambda^{(h)} \qquad\qquad (h = 1, \ldots, n).$$

$\square$

Theorem 4.4.4. *Let $x, y \in X$ such that*

$$u = \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(x),$$

$$v = \lim_{k \to \infty} \mathrm{Avg}_{F_k} f(y)$$

exist, where $u \leq v$. Suppose $\mathcal{K} \subseteq [u,v]$ is a nonempty compact subset. Let μ be a fully supported Borel probability measure on X that neglects shells. Then there exist sequences $(r_k)_{k=1}^{\infty}, (s_k)_{k=1}^{\infty}$ of positive numbers such that

$$\mathcal{K} = \mathrm{LS} \left(\left(\alpha_{B(x,y;r_k,s_k)} \left(\mathrm{Avg}_{F_k} f \right) \right)_{k=1}^{\infty} \right)$$

Proof. Let $P = \{p_i : i \in I\} \subseteq \mathcal{K}$ be a countable dense subset of $\mathcal{K}$ enumerated by the countable indexing set I, and let $\mathcal{N} = \{\mathcal{N}_i : i \in I\}$ be a partition of $\mathbb{N}$ into countably many infinite subsets, also enumerated by I. For convenience, write $i(k)$ for the $i \in I$ such that $k \in \mathcal{N}_i$.

For each $i \in I$, choose $\lambda_i \in [0, 1]$ such that

$$p_i = \lambda_i u + (1 - \lambda_i)v.$$

For each $k \in \mathbb{N}$, choose $t_k \in (0, 1)$ such that

$$|t_k - \lambda_{i(k)}| \leq 1/k.$$

Using the uniform continuity of $\mathrm{Avg}_{F_k} f$ and Lemma 4.4.3, choose $(r_k, s_k)_{k=1}^{\infty}$ such that

$$\rho(w, z) \leq \max\{r_k, s_k\} \Rightarrow \left|\mathrm{Avg}_{F_k} f(w) - \mathrm{Avg}_{F_k} f(z)\right| \leq 1/k \qquad (\forall w, z \in X),$$
$$\frac{\mu(B(x; r_k))}{\mu(B(x; r_k)) + \mu(B(y; s_k))} = t_k$$

for all $k \in \mathbb{N}$.

For $k \in \mathbb{N}$, we have that

$$\left| p_{i(k)} - \alpha_{B(x,y;r_k,s_k)} \left(\mathrm{Avg}_{F_k} f \right) \right|$$

$$= \left| p_{i(k)} - \left(t_k \alpha_{B(x;r_k)} \left(\mathrm{Avg}_{F_k} f \right) + (1 - t_k) \alpha_{B(y;s_k)} \left(\mathrm{Avg}_{F_k} f \right) \right) \right|$$

$$\leq \left| p_{i(k)} - \left(t_k \left(\mathrm{Avg}_{F_k} f(x) \right) + (1 - t_k) \left(\mathrm{Avg}_{F_k} f(y) \right) \right) \right|$$

$$\quad + t_k \left| \alpha_{B(x;r_k)} \mathrm{Avg}_{F_k} (f(x) - f) \right| + (1 - t_k) \left| \alpha_{B(y;s_k)} \mathrm{Avg}_{F_k} (f(y) - f) \right|$$

$$\leq \left| p_{i(k)} - \left(t_k \left(\mathrm{Avg}_{F_k} f(x) \right) + (1 - t_k) \left(\mathrm{Avg}_{F_k} f(y) \right) \right) \right| + \frac{1}{k}$$

$$= \left| p_{i(k)} - (t_k u + (1 - t_k) v) \right| + t_k \left| u - \mathrm{Avg}_{F_k} f(x) \right| + (1 - t_k) \left| v - \mathrm{Avg}_{F_k} f(y) \right| + \frac{1}{k}$$

$$= \left| \lambda_{i(k)} u + (1 - \lambda_{i(k)}) v - (t_k u + (1 - t_k) v) \right|$$

$$\quad + t_k \left| u - \mathrm{Avg}_{F_k} f(x) \right| + (1 - t_k) \left| v - \mathrm{Avg}_{F_k} f(y) \right| + \frac{1}{k}$$

$$\leq \left| \left(\lambda_{i(k)} - t_k \right) u \right| + \left| \left((1 - \lambda_{i(k)}) - (1 - t_k) \right) v \right|$$

$$\quad + t_k \left| u - \mathrm{Avg}_{F_k} f(x) \right| + (1 - t_k) \left| v - \mathrm{Avg}_{F_k} f(y) \right| + \frac{1}{k}$$

$$= \left| \left(\lambda_{i(k)} - t_k \right) u \right| + \left| \left(\lambda_{i(k)} - t_k \right) v \right|$$

$$\quad + t_k \left| u - \mathrm{Avg}_{F_k} f(x) \right| + (1 - t_k) \left| v - \mathrm{Avg}_{F_k} f(y) \right| + \frac{1}{k}$$

$$\leq \frac{1}{k} |u| + \frac{1}{k} |v| + t_k \left| u - \mathrm{Avg}_{F_k} f(x) \right| + (1 - t_k) \left| v - \mathrm{Avg}_{F_k} f(y) \right| + \frac{1}{k}$$

$$\leq \frac{|u|}{k} + \frac{|v|}{k} + \left| u - \mathrm{Avg}_{F_k} f(x) \right| + \left| v - \mathrm{Avg}_{F_k} f(y) \right| + \frac{1}{k}$$

We now claim that

$$\mathcal{K}$$

$$= \left\{ \lim_{\ell \to \infty} \alpha_{B\left(x,y;r_{k_\ell}, s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) : k_1 < k_2 < \cdots, \ \lim_{\ell \to \infty} \alpha_{B\left(x,y;r_{k_\ell}, s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) \text{ exists} \right\}.$$

We will prove the two sets contain each other, and thus are equal. First, let $q \in \mathcal{K}$, and choose a sequence $(p_{i_\ell})_{\ell=1}^{\infty}$ in S such that $|q - p_{i_\ell}| < 1/\ell$ for all $\ell \in \mathbb{N}$. For each $\ell \in \mathbb{N}$, recursively choose

$k_\ell > \max\{k_1, \dots, k_{\ell-1}\}$ such that

$$\left| u - \mathrm{Avg}_{F_{k_\ell}} f(x) \right| < 1/\ell,$$

$$\left| v - \mathrm{Avg}_{F_{k_\ell}} f(y) \right| < 1/\ell,$$

$$k_\ell \in \mathcal{N}_{i_\ell}.$$

Then

$$
\begin{aligned}
\left| q - \alpha_{B\left(x,y;r_{k_\ell},s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) \right| &\leq \left| q - p_{i_\ell} \right| + \left| p_{i_\ell} - \alpha_{B\left(x,y;r_{k_\ell},s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) \right| \\
&\leq \frac{1}{\ell} + \frac{1}{k_\ell} |u| + \frac{1}{k_\ell} |v| + \left| u - \mathrm{Avg}_{F_{k_\ell}} f(x) \right| \\
&\quad + \left| v - \mathrm{Avg}_{F_{k_\ell}} f(y) \right| + \frac{1}{k_\ell} \\
&\leq \frac{1}{\ell} + \frac{|u|}{\ell} + \frac{|v|}{\ell} + \frac{1}{\ell} + \frac{1}{\ell} + \frac{1}{\ell} \\
&= \frac{4 + |u| + |v|}{\ell} \\
&\overset{\ell \to \infty}{\to} 0.
\end{aligned}
$$

Therefore

$$q \in \mathrm{LS}\left(\left(\alpha_{B(x,y;r_k,s_k)} \left(\mathrm{Avg}_{F_k} f \right) \right)_{k=1}^{\infty} \right).$$

Conversely, let $k_1 < k_2 < \cdots$ be an increasing sequence of natural numbers such that $q = \lim_{\ell \to \infty} \alpha_{B\left(x_{k_\ell}, y_{k_\ell}; r_{k_\ell}, y_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right)$ exists. Fix $\epsilon > 0$, and choose $K \in \mathbb{N}$ sufficiently large that

$$
\begin{aligned}
\Rightarrow \left| u - \mathrm{Avg}_{F_k} f(x) \right| &< \epsilon && (\forall k \geq K), \\
\Rightarrow \left| v - \mathrm{Avg}_{F_k} f(y) \right| &< \epsilon && (\forall k \geq K), \\
\frac{\max\{|u|, |v|, 1\}}{K} &< \epsilon, \\
\left| q - \alpha_{B\left(x,y;r_{k_\ell},s_{k_\ell}\right)} \left(\mathrm{Avg}_{F_{k_\ell}} f \right) \right| &< \epsilon && (\forall \ell \geq K).
\end{aligned}
$$

Then if $\ell \geq K$, we have

$$
\begin{aligned}
\left|p_{i(k_\ell)} - q\right| &\leq \left|p_{i(k_\ell)} - \alpha_{B\left(x,y;r_{k_\ell},s_{k_\ell}\right)}\left(\mathrm{Avg}_{F_{k_\ell}} f\right)\right| + \left|\alpha_{B\left(x,y;r_{k_\ell},s_{k_\ell}\right)}\left(\mathrm{Avg}_{F_{k_\ell}} f\right) - q\right| \\
&\leq \frac{1}{k}|u| + \frac{1}{k}|v| + \left|u - \mathrm{Avg}_{F_k} f(x)\right| + \left|v - \mathrm{Avg}_{F_k} f(y)\right| + \frac{1}{k} + \epsilon \\
&< 6\epsilon.
\end{aligned}
$$

Therefore $\inf_{p \in \mathcal{K}} |p - q| < 6\epsilon$. Since our choice of $\epsilon > 0$ was arbitrary, it follows that $\inf_{p \in \mathcal{K}} |p - q| = 0$, and since $\mathcal{K}$ is compact, this implies that $q \in \mathcal{K}$. $\qquad\square$

Corollary 4.4.5. *Suppose G is an amenable group, and $\mathbf{F} = (F_k)_{k=1}^{\infty}$ is a right Følner sequence for G. Let $f \in C_{\mathbb{R}}(X)$ such that for every ergodic $\theta \in \partial_e \mathcal{M}_T(X)$ exists an $(f, \mathbf{F}, \theta)$-typical point. Let $\mathcal{K}$ be a compact subset of the compact interval*

$$
[\underline{a}(f),\ \overline{a}(f)].
$$

Let μ be a fully supported Borel probability measure on X that neglects shells. Then there exist points $x, y \in X$ and sequences $(r_k)_{k=1}^{\infty}, (s_k)_{k=1}^{\infty}$ of positive numbers such that

$$
\mathcal{K} = \mathrm{LS}\left(\left(\alpha_{B(x,y;r_k,s_k)}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^{\infty}\right).
$$

Proof. By (Jenkinson, 2006a, Proposition 2.4-(iii)), there exist ergodic Borel probability measures θ_1, θ_2 such that

$$
\begin{aligned}
\int f \, \mathrm{d}\theta_1 &= \underline{a}(f), \\
\int f \, \mathrm{d}\theta_2 &= \overline{a}(f).
\end{aligned}
$$

By hypothesis, there exist $x, y \in X$ such that

$$
\begin{aligned}
\lim_{\ell \to \infty} \mathrm{Avg}_{F_{k_\ell}} f(x) &= \int f \, \mathrm{d}\theta_1, \\
\lim_{\ell \to \infty} \mathrm{Avg}_{F_{k_\ell}} f(y) &= \int f \, \mathrm{d}\theta_2.
\end{aligned}
$$

Apply Theorem 4.4.4. $\qquad\square$

Corollary 4.4.6. *Suppose G is an amenable group, and $\mathbf{F} = (F_k)_{k=1}^\infty$ is a right Følner sequence for G. Let $f \in C_{\mathbb{R}}(X)$, and let $\mathcal{K}$ be a compact subset of the compact interval*

$$[\underline{a}(f),\ \overline{a}(f)].$$

Let μ be a fully supported Borel probability measure on X that neglects shells. Then there exist points $x, y \in X$ and sequences $(r_k)_{k=1}^\infty, (s_k)_{k=1}^\infty$ of positive numbers such that

$$\mathcal{K} \subseteq \mathrm{LS}\left(\left(\alpha_{B(x,y;r_k,s_k)}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty\right).$$

Proof. Choose a tempered Følner subsequence $(F_{k_\ell})_{\ell=1}^\infty$ of $\mathbf{F}$. By (Jenkinson, 2006a, Proposition 2.4-(iii)), there exist ergodic Borel probability measures θ_1, θ_2 such that

$$\int f\mathrm{d}\theta_1 = \underline{a}(f),$$
$$\int f\mathrm{d}\theta_2 = \overline{a}(f).$$

By the Lindestrauss Ergodic Theorem, there exist $x, y \in X$ such that

$$\lim_{\ell\to\infty} \mathrm{Avg}_{F_{k_\ell}} f(x) = \int f\mathrm{d}\theta_1,$$
$$\lim_{\ell\to\infty} \mathrm{Avg}_{F_{k_\ell}} f(y) = \int f\mathrm{d}\theta_2.$$

By Theorem 4.4.4, there exist $(r_k)_{k=1}^\infty, (s_k)_{k=1}^\infty \in (0,\infty)^{\mathbb{N}}$ such that

$$\mathcal{K} = \mathrm{LS}\left(\left(\alpha_{B\left(x,y;r_{k_\ell},s_{k_\ell}\right)}\left(\mathrm{Avg}_{F_{k_\ell}} f\right)\right)_{\ell=1}^\infty\right).$$

Then

$$\mathcal{K} \subseteq \mathrm{LS}\left(\left(\alpha_{B(x,y;r_k,s_k)}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^\infty\right).$$

$\square$

Theorem 4.4.7. *Suppose G is an amenable group, and $\mathbf{F} = (F_k)_{k=1}^{\infty}$ is a tempered Følner sequence for G. Let $f \in C_{\mathbb{R}}(X)$, and let $\mathcal{K}$ be a compact subset of the compact interval*

$$[\underline{a}(f),\ \overline{a}(f)]\,.$$

Let μ be a fully supported Borel probability measure on X that neglects shells. Then there exist points $x, y \in X$ and sequences $(r_k)_{k=1}^{\infty}, (s_k)_{k=1}^{\infty}$ of positive numbers such that

$$\mathcal{K} = \mathrm{LS}\left(\left(\alpha_{B(x,y;r_k,s_k)}\left(\mathrm{Avg}_{F_k} f\right)\right)_{k=1}^{\infty}\right).$$

Proof. The Lindenstrauss Ergodic Theorem implies that for every ergodic $\theta \in \partial_e \mathcal{M}_T(X)$ exists an $(f, \mathbf{F}, \theta)$-typical point. Apply Corollary 4.4.5. $\qquad\square$

4.5 Pathological multi-local temporo-spatial differentiations on $C(X)$

In this section, we consider a temporo-spatial differentiation $\left(\alpha_{C_k} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}$ as a sequence in $\mathcal{M}_T(X)$. If G is a discrete amenable group, and $\mathbf{F} = (F_k)_{k=1}^{\infty}$ is a Følner sequence, then Lemma 4.3.1 tells us that

$$\mathrm{LS}\left(\left(\alpha_{C_k} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right) \subseteq \mathcal{M}_T(X)$$

for all sequences $(C_k)_{k=1}^{\infty}$ of measurable subsets of X with positive measure.

We are motivated here by the following question: Consider an action $T : G \curvearrowright X$ of a discrete amenable group G on a compact metrizable space by X, where X is endowed with a Borel probability measure μ. Given a Følner sequence $\mathbf{F} = (F_k)_{k=1}^{\infty}$ for G, can we choose a sequence $(C_k)_{k=1}^{\infty}$ of measurable subsets of X with $\mu(C_k) > 0$ such that

$$\mathrm{LS}\left(\left(\alpha_{C_k} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right) = \mathcal{C},$$

where $\mathcal{C}$ is some prescribed compact subset of $\mathcal{M}_T(X)$? If so, then can the $(C_k)_{k=1}^{\infty}$ be chosen to fit some prescribed constraints?

In this section, we provide positive answers for certain classes of $\mathcal{C}$. Throughout this section, assume that G is a discrete amenable group and $\mathbf{F}$ is a Følner sequence for G. We also assume that $T : G \curvearrowright X$ is a Hölder action with MoHöC (H, L).

Lemma 4.5.1. *Let $\mathcal{L} \subseteq C(X)$ denote the family of functions $f \in C(X)$ for which*

$$|f(x) - f(y)| \leq \rho(x, y) \qquad\qquad (\forall x, y \in X).$$

Then $\mathcal{L}$ has dense span in $C(X)$.

Proof. For $x_0 \in X$, set $\phi_{x_0}(x) = \rho(x, x_0)$. If $x, y \in X$, then by the Reverse Triangle Inequality we know

$$|\phi_{x_0}(x) - \phi_{x_0}(y)| = |\rho(x, x_0) - \rho(y, x_0)| \leq \rho(x, y).$$

Thus the functions ϕ_{x_0} satisfy the prescribed Lipschitz condition, as does the constant function 1. Furthermore, we know that $\{\phi_{x_0} : x_0 \in X\}$ separates points, since if $x, y \in X, x \neq y$, then $0 = \phi_x(x) \neq \phi_x(y)$. Therefore by the Stone-Weierstrass Theorem, we know that $C(X)$ is densely spanned by finite products of elements in $\{\phi_{x_0} : x \in X\} \cup \{1\} \subseteq \mathcal{L}$. We claim, however, that a product of elements in $\mathcal{L}$ is a scalar multiple of an element in $\mathcal{L}$. Let $f_1, f_2 \in \mathcal{L}$. Then

$$
\begin{aligned}
|f_1(x)f_2(x) - f_1(y)f_2(y)| &= |f_1(x)f_2(x) - f_1(x)f_2(y) + f_1(x)f_2(y) - f_1(y)f_2(y)| \\
&\leq |f_1(x)| \cdot |f_2(x) - f_2(y)| + |f_1(x) - f_1(y)| \cdot |f_2(y)| \\
&\leq \|f_1\|_{C(X)} \cdot |f_2(x) - f_2(y)| + |f_1(x) - f_1(y)| \cdot \|f_2\|_{C(X)} \\
&\leq \left(\|f_1\|_{C(X)} + \|f_2\|_{C(X)} \right) \rho(x, y).
\end{aligned}
$$

Let $h = \frac{f_1 f_2}{\|f_1\|_{C(X)} + \|f_2\|_{C(X)} + 1}$. Then $h \in \mathcal{L}$, so $f_1 f_2 = \left(\|f_1\|_{C(X)} + \|f_2\|_{C(X)} + 1 \right) h \in \mathbb{C}\mathcal{L}$. By an inductive argument, we can show that any finite product of elements of $\mathcal{L}$ is an element of $\mathbb{C}\mathcal{L}$. Therefore, the Stone-Weierstrass Theorem tells us that $C(X)$ is densely spanned by $\mathcal{L}$. $\qquad\square$

Theorem 4.5.2. *Let $\theta^{(1)}, \ldots, \theta^{(n)} \in \partial_e \mathcal{M}_T(X)$ be a finite collection of ergodic measures on X, and let $\mathcal{C}$ be a compact subset of the convex hull of $\{\theta^{(1)}, \ldots, \theta^{(n)}\}$. Suppose $\mathbf{F}$ is a tempered Følner sequence, and that μ is a Borel probability measure on X that neglects shells. Then there exist points $x^{(1)}, \ldots, x^{(n)}$ and sequences of radii $\left(r_k^{(1)} \right)_{k=1}^{\infty}, \ldots, \left(r_k^{(n)} \right)_{k=1}^{\infty}$ such that*

$$\mathrm{LS}\left(\left(\alpha_{B\left(x^{(1)}, \ldots, x^{(n)}; r_k^{(1)}, \ldots, r_k^{(n)} \right)} \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty} \right) = \mathcal{C}.$$

Moreover, the set of n-tuples $\left(x^{(1)}, \ldots, x^{(n)}\right) \in X^n$ which admit such sequences $\left(r_k^{(1)}, \ldots, r_k^{(n)}\right)_{k=1}^{\infty}$ is of full probability with respect to the product measure $\theta^{(1)} \times \cdots \times \theta^{(n)}$.

Proof. Assume without loss of generality that $\theta^{(1)}, \ldots, \theta^{(n)}$ are distinct. By the Lindenstrauss Ergodic Theorem, there exist points $x^{(1)}, \ldots, x^{(n)} \in \operatorname{supp}(\mu)$ such that

$$\lim_{k \to \infty} \operatorname{Avg}_{F_k} f\left(x^{(h)}\right) = \int f \, d\theta^{(h)} \qquad\qquad (h = 1, \ldots, n).$$

In fact, the Lindenstrauss Ergodic Theorem tells us that the set of such $\left(x^{(1)}, \ldots, x^{(n)}\right) \in X^n$ is of full measure with respect to $\theta^{(1)} \times \cdots \times \theta^{(n)}$. For the remainder of this proof, let $\bar{x} = \left(x^{(1)}, \ldots, x^{(n)}\right) \in X^n$ be such an n-tuple.

For each $i \in I$, let $\bar{\lambda}_i = \left(\lambda_i^{(1)}, \ldots, \lambda_i^{(n)}\right) \in [0, 1]^n$ be such that $\nu_i = \sum_{h=1}^n \lambda_i^{(h)} \theta^{(h)}$.

Let $\mathscr{N} = \{\mathcal{N}_i : i \in I\}$ be a partition of $\mathbb{N}$ into infinite subsets. For each $k \in \mathbb{N}$, set $i(k) \in I$ such that $k \in \mathcal{N}_{i(k)}$. For each $k \in \mathbb{N}$, choose $\bar{t}_k = \left(t_k^{(1)}, \ldots, t_k^{(n)}\right) \in (0, 1)^n$ such that

$$\sum_{h=1}^n \left| t_k^{(h)} - \lambda_{i(k)}^{(h)} \right| < 1/k,$$

$$\sum_{h=1}^n t_k^{(h)} = 1.$$

For each $k \in \mathbb{N}$, choose $\delta_k > 0$ such that

$$\max_{g \in F_k} \left(L(g) \cdot \delta_k^{H(g)} \right) < 1/k.$$

Now for each $k \in \mathbb{N}$, use Lemma 4.4.3 to choose $\bar{r}_k = \left(r_k^{(1)}, \ldots, r_k^{(n)}\right) \in (0, 1)^n$ such that

$$\frac{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)}{\mu\left(B\left(x^{(1)}; r_k^{(1)}\right)\right) + \cdots + \mu\left(B\left(x^{(n)}; r_k^{(n)}\right)\right)} = t_k^{(h)},$$

$$r_k^{(h)} < \delta_k,$$

$$r_k^{(h)} < \frac{1}{3} \min \left\{ \rho\left(x^{(h_1)}, x^{(h_2)}\right) : 1 \le h_1 < h_2 \le n \right\}.$$

The last condition ensures that the balls $\left\{ B\left(x^{(h)}; r_k^{(h)}\right) : h = 1, \ldots, n \right\}$ are pairwise disjoint. Since the points $x^{(h)}$ each satisfy

$$\lim_{k \to \infty} \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) = \int f \mathrm{d}\theta^{(h)},$$

for all $f \in C(X)$, and the measures $\theta^{(1)}, \ldots, \theta^{(h)}$ are distinct, it follows that the $x^{(1)}, \ldots, x^{(n)}$ are also distinct, meaning that $\min \left\{ \rho\left(x^{(h_1)}, x^{(h_2)}\right) : 1 \le h_1 < h_2 \le n \right\} > 0$.

Let $\mathcal{L} \subseteq C(X)$ denote the family of all continuous functions f on X such that

$$|f(x) - f(y)| \le \rho(x, y) \qquad\qquad (\forall x, y \in X),$$

i.e. the 1-Lipschitz functions $X \to \mathbb{C}$, and let $f \in \mathcal{L}$. Then

$$
\begin{aligned}
&\left| \alpha_{B(\bar{x}, \bar{r}_k)} \left(\mathrm{Avg}_{F_k} f \right) - \int f \mathrm{d}\nu_{i(k)} \right| \\
[\text{Lem. 4.2.6}] =\; & \left| \left[\sum_{h=1}^n \frac{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right) \right)}{\sum_{u=1}^n \mu\left(B\left(x^{(u)}; r^{(u)}\right) \right)} \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) \right] - \int f \mathrm{d}\nu_{i(k)} \right| \\
=\; & \left| \left[\sum_{h=1}^n t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) \right] - \int f \mathrm{d}\nu_{i(k)} \right| \\
=\; & \left| \sum_{h=1}^n \left(t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) - \lambda_{i(k)}^{(h)} \int f \mathrm{d}\theta^{(h)} \right) \right| \\
\le\; & \sum_{h=1}^n \left| t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) - \lambda_{i(k)}^{(h)} \int f \mathrm{d}\theta^{(h)} \right| \\
\le\; & \sum_{h=1}^n \left[\left| t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) - t_k^{(h)} \int f \mathrm{d}\theta^{(h)} \right| + \left| \left(t_k^{(h)} - \lambda_{i(k)}^{(h)} \right) \int f \mathrm{d}\theta^{(h)} \right| \right] \\
\le\; & \left[\sum_{h=1}^n t_k^{(h)} \left| \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) - \int f \mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{k}.
\end{aligned}
$$

We can then estimate

$$
\begin{aligned}
&\left| \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) - \int f \mathrm{d}\theta^{(h)} \right| \\
\le\; & \left| \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f \right) - \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) \right| + \left| \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)} \right|
\end{aligned}
$$

Since $r_k^{(h)} < \delta_k$ for all $k \in \mathbb{N}$, it follows that if $\rho\left(x^{(h)}, y\right) < r_k^{(h)}$, then $\rho\left(T_g x^{(h)}, T_g y\right) < 1/k$ for $g \in F_k$. Since f is 1-Lipschitz, it follows that $\left|f\left(T_g x^{(h)}\right) - f\left(T_g y\right)\right| < 1/k$ for all $g \in F_k$. Thus

$$
\begin{aligned}
&\left|\alpha_{B\left(x^{(h)}; r_k^{(h)}\right)}\left(\mathrm{Avg}_{F_k} f\right) - \mathrm{Avg}_{F_k} f\left(x^{(h)}\right)\right| \\
&= \left|\frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)} \int_{B\left(x^{(h)}; r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k}\left(f(T_g y) - f\left(T_g x^{(h)}\right)\right) \mathrm{d}\mu(y)\right| \\
&\leq \frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)} \int_{B\left(x^{(h)}; r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k}\left|f(T_g y) - f\left(T_g x^{(h)}\right)\right| \mathrm{d}\mu(y) \\
&< \frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)} \int_{B\left(x^{(h)}; r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{k} \mathrm{d}\mu(y) \\
&= \frac{1}{k}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\left|\alpha_{B(\bar{x}, \bar{r}_k)}\left(\mathrm{Avg}_{F_k} f\right) - \int f \mathrm{d}\nu_{i(k)}\right| \\
&\leq \left[\sum_{h=1}^{n} t_k^{(h)}\left|\alpha_{B\left(x^{(h)}; r_k^{(h)}\right)}\left(\mathrm{Avg}_{F_k} f\right) - \int f \mathrm{d}\theta^{(h)}\right|\right] + \frac{\|f\|_{C(X)}}{k} \\
&\leq \left[\sum_{h=1}^{n} t_k^{(h)}\left(\left|\alpha_{B\left(x^{(h)}; r_k^{(h)}\right)}\left(\mathrm{Avg}_{F_k} f\right) - \mathrm{Avg}_{F_k} f\left(x^{(h)}\right)\right| + \left|\mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)}\right|\right)\right] \\
&\quad + \frac{\|f\|_{C(X)}}{k} \\
&= \left[\sum_{h=1}^{n} t_k^{(h)}\left(\frac{1}{k} + \left|\mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)}\right|\right)\right] + \frac{\|f\|_{C(X)}}{k} \\
&= \frac{1}{k} + \left[\sum_{h=1}^{n} t_k^{(h)}\left|\mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)}\right|\right] + \frac{\|f\|_{C(X)}}{k}.
\end{aligned}
$$

Let $\{f_m : m \in \mathbb{N}\}$ be a countable family of functions in $\mathcal{L}$ that densely span $C(X)$, and let dist : $\mathcal{M}(X) \times \mathcal{M}(X) \to [0, 1]$ be the metric

$$
\mathrm{dist}(\beta_1, \beta_2) = \sum_{m=1}^{\infty} 2^{-m} \min\left\{\left|\int f_m \mathrm{d}(\beta_1 - \beta_2)\right|, 1\right\}.
$$

This dist metric is compatible with the weak*-topology on $\mathcal{M}(X)$. We can also say that for all $M \in \mathbb{N}$, we have

$$\text{dist}\left(\alpha_{B(\bar{x},\bar{r}_k)} \circ \text{Avg}_{F_k}, \nu_{i(k)}\right)$$

$$\leq \left[\sum_{m=1}^{M} 2^{-m} \left|\alpha_{B(\bar{x},\bar{r}_k)} \left(\text{Avg}_{F_k} f_m\right) - \int f_m d\nu_{i(k)}\right|\right] + \sum_{m=M+1}^{\infty} 2^{-m}$$

$$\leq \left[\sum_{m=1}^{M} 2^{-m} \left[\frac{1}{k} + \left[\sum_{h=1}^{n} t_k^{(h)} \left|\text{Avg}_{F_k} f_m \left(x^{(h)}\right) - \int f_m d\theta^{(h)}\right|\right] + \frac{\|f_m\|_{C(X)}}{k}\right]\right] + 2^{-M}$$

$$\leq \frac{1 + \max_{1 \leq m \leq M} \|f_m\|_{C(X)}}{k} + 2^{-M} + \max_{1 \leq m \leq M} \max_{1 \leq h \leq n} \left|\text{Avg}_{F_k} f_m \left(x^{(h)}\right) - \int f_m d\theta^{(h)}\right|$$

We claim that $\text{LS}\left(\left(\alpha_{B(\bar{x},\bar{r}_k)} \circ \text{Avg}_{F_k}\right)_{k=1}^{\infty}\right) = \mathcal{C}$.

First, let $\nu \in \mathcal{C}$. Choose a sequence $(\nu_{i_\ell})_{\ell=1}^{\infty}$ such that $\text{dist}\,(\nu, \nu_{i_\ell}) < 1/\ell$. Choose $k_1 < k_2 < \cdots$ such that

$$k \geq k_\ell \Rightarrow \left|\text{Avg}_{F_k} f_m \left(x^{(h)}\right) - \int f_m d\theta^{(h)}\right| \leq \frac{1}{\ell} \qquad (m = 1, \ldots, \ell; h = 1, \ldots, n),$$

$$k_\ell \geq \ell \left(1 + \max_{1 \leq m \leq \ell} \|f_m\|_{C(X)}\right),$$

$$k_\ell \in \mathcal{N}_{i_\ell}$$

for all $\ell \in \mathbb{N}$. Then

$$\text{dist}\left(\alpha_{B\left(\bar{x},\bar{r}_{k_\ell}\right)} \circ \text{Avg}_{F_{k_\ell}}, \nu\right)$$

$$\leq \text{dist}\left(\alpha_{B\left(\bar{x},\bar{r}_{k_\ell}\right)} \circ \text{Avg}_{F_{k_\ell}}, \nu_{i(k_\ell)}\right) + \text{dist}\,(\nu_{i_\ell}, \nu)$$

$$\leq \left[\frac{1 + \max_{1 \leq m \leq \ell} \|f_m\|_{C(X)}}{k_\ell} + 2^{-M} + \max_{1 \leq m \leq \ell} \max_{1 \leq h \leq n} \left|\text{Avg}_{F_{k_\ell}} f_m \left(x^{(h)}\right) - \int f_m d\theta^{(h)}\right|\right]$$

$$+ \frac{1}{\ell}$$

$$\leq \frac{1}{\ell} + 2^{-\ell} + \frac{1}{\ell} + \frac{1}{\ell}$$

$$\overset{\ell \to \infty}{\to} 0.$$

Therefore $\nu \in \text{LS}\left(\left(\alpha_{B(\bar{x},\bar{r}_k)} \circ \text{Avg}_{F_k}\right)_{k=1}^{\infty}\right)$, meaning that $\mathcal{C} \subseteq \text{LS}\left(\left(\alpha_{B(\bar{x},\bar{r}_k)} \circ \text{Avg}_{F_k}\right)_{k=1}^{\infty}\right)$.

To prove the opposite containment, suppose $\gamma \in \mathrm{LS}\left(\left(\alpha_{B(\bar{x},\bar{r}_k)} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right)$, and let $k_1 < k_2 < \cdots$ such that $\gamma = \lim_{\ell \to \infty} \alpha_{B(\bar{x},\bar{r}_{k_\ell})} \circ \mathrm{Avg}_{F_{k_\ell}}$. Fix $f \in \mathcal{L}$. Then

$$
\left| \int f\,\mathrm{d}\gamma - \int f\,\mathrm{d}\nu_{i(k_\ell)} \right|
$$

$$
\leq \left| \int f\,\mathrm{d}\gamma - \alpha_{B(\bar{x},\bar{r}_{k_\ell})}\left(\mathrm{Avg}_{F_{k_\ell}} f\right) \right| + \left| \alpha_{B(\bar{x},\bar{r}_{k_\ell})}\left(\mathrm{Avg}_{F_{k_\ell}} f\right) - \int f\,\mathrm{d}\nu_{i(k_\ell)} \right|
$$

$$
\leq \left| \int f\,\mathrm{d}\gamma - \alpha_{B(\bar{x},\bar{r}_{k_\ell})}\left(\mathrm{Avg}_{F_{k_\ell}} f\right) \right| + \frac{1}{k_\ell}
$$

$$
+ \left[\sum_{h=1}^{n} t_{k_\ell}^{(h)} \left| \mathrm{Avg}_{F_{k_\ell}} f\left(x^{(h)}\right) - \int f\,\mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{k_\ell}
$$

$$
\leq \left| \int f\,\mathrm{d}\gamma - \alpha_{B(\bar{x},\bar{r}_{k_\ell})}\left(\mathrm{Avg}_{F_{k_\ell}} f\right) \right| + \frac{1}{k_\ell}
$$

$$
+ \left[\max_{1 \leq h \leq n} \left| \mathrm{Avg}_{F_{k_\ell}} f\left(x^{(h)}\right) - \int f\,\mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{k_\ell}
$$

$$
\overset{\ell \to \infty}{\to} 0.
$$

Therefore $\gamma = \lim_{\ell \to \infty} \nu_{i(k_\ell)}$, meaning that $\gamma \in \mathcal{C}$. Thus $\mathrm{LS}\left(\left(\alpha_{B(\bar{x},\bar{r}_k)} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right) \subseteq \mathcal{C}$. $\qquad\square$

In Theorem 4.5.2, our assumption that $\mathcal{C}$ live in a finite-dimensional subset of $\mathcal{M}_T(X)$ helps us place an upper bound on $\mathrm{LS}\left(\alpha_{B(\bar{x};\bar{r}_k)} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}$, i.e. show that $\mathrm{LS}\left(\alpha_{B(\bar{x};\bar{r}_k)} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty} \subseteq \mathcal{C}$. In general, it is possible to construct $(C_k)_{k=1}^{\infty}$ for which $\mathrm{LS}\left(\alpha_{B(\bar{x};\bar{r}_k)} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}$ is "maximally large," as the following theorem shows.

Theorem 4.5.3. *Suppose μ is a Borel probability measure on X. Then there exists a sequence $(C_k)_{k=1}^{\infty}$ of multi-balls in X such that*

$$
\mathrm{LS}\left(\left(\alpha_{C_k} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right) = \mathcal{M}_T(X).
$$

Proof. Since $\mathrm{LS}\left(\left(\alpha_{C_k} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right)$ is always a closed subset of $\mathcal{M}_T(X)$, it will suffice to construct $(C_k)_{k=1}^{\infty}$ such that $\mathrm{LS}\left(\left(\alpha_{C_k} \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}\right)$ is dense in $\mathcal{M}_T(X)$.

Let $\mathcal{E} = \left\{\theta^{(h)} : h \in \mathbb{N}\right\} \subseteq \partial_e \mathcal{M}_T(X)$ be a countable dense subset of $\partial_e \mathcal{M}_T(X)$, and set

$$
\mathcal{F} = \left\{ \sum_{h=1}^{n} \lambda^{(h)} \theta^{(h)} : n \in \mathbb{N}, \bar{\lambda} \in [0,1]^n \cap \mathbb{Q}^n, \sum_{h=1}^{n} \lambda^{(h)} = 1 \right\},
$$

i.e. $\mathcal{F}$ is the set of all rational convex combinations of elements of $\mathcal{E}$. Assume that the $\theta^{(h)}, h \in \mathbb{N}$ are distinct. By the Krein-Millman Theorem, the set $\mathcal{F}$ is a countable dense subset of $\mathcal{M}_T(X)$. Let $\{\nu_i : i \in I\}$ be an enumeration of $\mathcal{F}$, where I is some countable indexing set, and let $\mathcal{N} = \{\mathcal{N}_i : i \in I\}$ be a partition of $\mathbb{N}$ into countably infinitely many infinite subsets.

For each $i \in I$, let $(\kappa(i, \ell))_{\ell=1}^{\infty}$ be a strictly increasing sequence such that

$$\kappa(i, \ell) \in \mathcal{N}_i,$$

$$\left(F_{\kappa(i,\ell)}\right)_{\ell=1}^{\infty} \text{ is tempered,}$$

which exists by Lemma 4.2.9.

We are going to construct $(C_k)_{k=1}^{\infty}$ such that $\lim_{\ell \to \infty} \alpha_{C_{\kappa(i,\ell)}} \circ \mathrm{Avg}_{F_{\kappa(i,\ell)}} = \nu_i$ for all $i \in I$. For each $k \in \mathbb{N}$, set $i(k) \in I$ such that $k \in \mathcal{N}_{i(k)}$.

For each $i \in I$, choose $\bar{\lambda}_i \in ([0, 1] \cap \mathbb{Q})^{\mathbb{N}}$ and $n_i \in \mathbb{N}$ such that

$$\sum_{h=1}^{n_i} \lambda_i^{(h)} \theta^{(h)} = \nu_i,$$

$$\sum_{h=1}^{n_i} \lambda_i^{(h)} = 1,$$

$$\lambda_i^{(h)} = 0 \qquad\qquad\qquad\qquad \text{for all } h > n_i.$$

By the Lindenstrauss Ergodic Theorem, there exists for each $\theta^{(h)}$ a point $x^{(h)} \in X$ such that

$$\lim_{\ell \to \infty} \mathrm{Avg}_{F_{\kappa(i,\ell)}} f\left(x^{(h)}\right) = \int f \mathrm{d}\theta^{(h)} \qquad\qquad (\forall f \in C(X), \ \forall i \in I).$$

For each $k \in \mathbb{N}$, choose $\bar{t}_k = \left(t_k^{(1)}, \ldots, t_k^{\left(n_{i(k)}\right)}\right) \in (0, 1)^{n_{i(k)}}$ such that

$$\sum_{h=1}^{n_{i(k)}} \left|t_k^{(h)} - \lambda_{i(k)}^{(h)}\right| < 1/k,$$

$$\sum_{h=1}^{n_{i(k)}} t_k^{(h)} = 1.$$

138

For each $k \in \mathbb{N}$, choose $\delta_k > 0$ such that

$$\max_{g \in F_k} \left(L(g) \cdot \delta_k^{H(g)} \right) < 1/k.$$

Now for each $k \in \mathbb{N}$, use Lemma 4.4.3 to choose $r_k^{(1)}, \ldots, r_k^{\left(n_{i(k)}\right)} \in (0,1)$ such that

$$t_k^{(j)} = \frac{\mu \left(B \left(x^{(h)}; r_k^{(h)} \right) \right)}{\mu \left(B \left(x^{(1)}; r_k^{(1)} \right) \right) + \cdots + \mu \left(B \left(x^{\left(n_{i(k)}\right)}; r_k^{\left(n_{i(k)}\right)} \right) \right)},$$

$$r_k^{(h)} < \delta_k,$$

$$r_k^{(h)} < \frac{1}{3} \min \left\{ \rho \left(x^{(h_1)}, x^{(h_2)} \right) : 1 \le h_1 < h_2 \le n_{i(k)} \right\}.$$

The last condition ensures that the balls $\left\{ B \left(x^{(h)}; r_k^{(h)} \right) : h = 1, \ldots, n_{i(k)} \right\}$ are pairwise disjoint. Since the points $x^{(h)}$ each satisfy

$$\lim_{\ell \to \infty} \mathrm{Avg}_{F_{\kappa(i,\ell)}} f \left(x^{(h)} \right) = \int f \mathrm{d}\theta^{(h)}$$

for all $f \in C(X), i \in I$, and the measures $\theta^{(h)}$ are distinct, it follows that the $x^{(h)}$ are also distinct, meaning that $\min \left\{ \rho \left(x^{(h_1)}, x^{(h_2)} \right) : 1 \le h_1 < h_2 \le n_{i(k)} \right\} > 0$.

For each $k \in \mathbb{N}$, set

$$C_k = B \left(x^{(1)}, \ldots, x^{\left(n_{i(k)}\right)}; r_k^{(1)}, \ldots, r_k^{\left(n_{i(k)}\right)} \right).$$

We now show that

$$\lim_{\ell \to \infty} \alpha_{C_{\kappa(i,\ell)}} \left(\mathrm{Avg}_{F_{\kappa(i,\ell)}} f \right) = \int f \mathrm{d}\nu_i \qquad\qquad (\forall f \in C(X), \ \forall i \in I).$$

In light of Lemma 4.5.1, it will suffice to prove the convergence for $f \in \mathcal{L}$, where

$$\mathcal{L} = \{ \phi \in C(X) : \forall x \in X \ \forall y \in X \ \left(|\phi(x) - \phi(y)| \le \rho(x,y) \right) \}$$

139

is the family of all 1-Lipschitz functions. We see

$$\left| \alpha_{C_k} \left(\mathrm{Avg}_{F_k} f \right) - \int f \mathrm{d}\nu_{i(k)} \right|$$

$$[\text{Lem. 4.2.6}] = \left| \left[\sum_{h=1}^{n_{i(k)}} \frac{\mu\left(B\left(x^{(h)}; r_k^{(h)} \right) \right)}{\sum_{u=1}^{n} \mu\left(B\left(x^{(u)}; r^{(u)} \right) \right)} \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) \right] - \int f \mathrm{d}\nu_{i(k)} \right|$$

$$= \left| \left[\sum_{h=1}^{n_{i(k)}} t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) \right] - \int f \mathrm{d}\nu_{i(k)} \right|$$

$$= \left| \sum_{h=1}^{n_{i(k)}} \left(t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) - \lambda_{i(k)}^{(h)} \int f \mathrm{d}\theta^{(h)} \right) \right|$$

$$\leq \sum_{h=1}^{n_{i(k)}} \left| t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) - \lambda_{i(k)}^{(h)} \int f \mathrm{d}\theta^{(h)} \right|$$

$$\leq \sum_{h=1}^{n_{i(k)}} \left[\left| t_k^{(h)} \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) - t_k^{(h)} \int f \mathrm{d}\theta^{(h)} \right| + \left| \left(t_k^{(h)} - \lambda_{i(k)}^{(h)} \right) \int f \mathrm{d}\theta^{(h)} \right| \right]$$

$$\leq \left[\sum_{h=1}^{n_{i(k)}} t_k^{(h)} \left| \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) - \int f \mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{k}.$$

We can then estimate

$$\left| \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) - \int f \mathrm{d}\theta^{(h)} \right|$$

$$\leq \left| \alpha_{B\left(x^{(h)}; r_k^{(h)} \right)} \left(\mathrm{Avg}_{F_k} f \right) - \mathrm{Avg}_{F_k} f\left(x^{(h)} \right) \right| + \left| \mathrm{Avg}_{F_k} f\left(x^{(h)} \right) - \int f \mathrm{d}\theta^{(h)} \right|$$

Since $r_k^{(h)} < \delta_k$ for all $k \in \mathbb{N}$, it follows that if $\rho\left(x^{(h)}, y\right) < r_k^{(h)}$, then $\rho\left(T_g x^{(h)}, T_g y\right) < 1/k$ for $g \in F_k$.

Since f is 1-Lipschitz, it follows that $\left| f\left(T_g x^{(h)}\right) - f\left(T_g y\right) \right| < 1/k$ for all $g \in F_k$. Thus

$$
\left| \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f\right) - \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) \right|
$$

$$
= \left| \frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)} \int_{B\left(x^{(h)}; r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \left(f(T_g y) - f\left(T_g x^{(h)}\right) \right) \mathrm{d}\mu(y) \right|
$$

$$
\leq \frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)} \int_{B\left(x^{(h)}; r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \left| f(T_g y) - f\left(T_g x^{(h)}\right) \right| \mathrm{d}\mu(y)
$$

$$
< \frac{1}{\mu\left(B\left(x^{(h)}; r_k^{(h)}\right)\right)} \int_{B\left(x^{(h)}; r_k^{(h)}\right)} \frac{1}{|F_k|} \sum_{g \in F_k} \frac{1}{k} \mathrm{d}\mu(y)
$$

$$
= \frac{1}{k}.
$$

Therefore

$$
\left| \alpha_{C_k} \left(\mathrm{Avg}_{F_k} f\right) - \int f \mathrm{d}\nu_{i(k)} \right|
$$

$$
\leq \left[\sum_{h=1}^{n_{i(k)}} t_k^{(h)} \left| \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f\right) - \int f \mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{k}
$$

$$
\leq \left[\sum_{h=1}^{n_{i(k)}} t_k^{(h)} \left(\left| \alpha_{B\left(x^{(h)}; r_k^{(h)}\right)} \left(\mathrm{Avg}_{F_k} f\right) - \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) \right| + \left| \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)} \right| \right) \right]
$$

$$
+ \frac{\|f\|_{C(X)}}{k}
$$

$$
= \left[\sum_{h=1}^{n_{i(k)}} t_k^{(h)} \left(\frac{1}{k} + \left| \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)} \right| \right) \right] + \frac{\|f\|_{C(X)}}{k}
$$

$$
= \frac{1}{k} + \left[\sum_{h=1}^{n_{i(k)}} t_k^{(h)} \left| \mathrm{Avg}_{F_k} f\left(x^{(h)}\right) - \int f \mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{k}.
$$

In particular, this tells us that for fixed $i \in I$, we have

$$
\left| \alpha_{C_{\kappa(i,\ell)}} \left(\operatorname{Avg}_{F_{\kappa(i,\ell)}} f \right) - \int f \mathrm{d}\nu_{i(k)} \right|
$$

$$
= \left| \alpha_{C_{\kappa(i,\ell)}} \left(\operatorname{Avg}_{F_{\kappa(i,\ell)}} f \right) - \int f \mathrm{d}\nu_{i(\kappa(i,\ell))} \right|
$$

$$
\leq \frac{1}{\kappa(i,\ell)} + \left[\sum_{h=1}^{n_i} t^{(h)}_{\kappa(i,\ell)} \left| \operatorname{Avg}_{F_{\kappa(i,\ell)}} f \left(x^{(h)} \right) - \int f \mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{\kappa(i,\ell)}
$$

$$
\leq \frac{1}{\ell} + \left[\max_{1 \leq h \leq n_i} \left| \operatorname{Avg}_{F_{\kappa(i,\ell)}} f \left(x^{(h)} \right) - \int f \mathrm{d}\theta^{(h)} \right| \right] + \frac{\|f\|_{C(X)}}{\ell}
$$

$$
\overset{\ell \to \infty}{\to} 0.
$$

Therefore $\nu_i = \lim_{\ell \to \infty} \alpha_{C_{\kappa(i,\ell)}} \operatorname{Avg}_{F_{\kappa(i,\ell)}}$. Thus LS $\left(\left(\alpha_{C_k} \circ \operatorname{Avg}_{F_k} \right)_{k=1}^{\infty} \right) \supseteq \mathcal{F}$ is dense in $\mathcal{M}_T(X)$, and since LS $\left(\left(\alpha_{C_k} \circ \operatorname{Avg}_{F_k} \right)_{k=1}^{\infty} \right)$ is a closed subset of $\mathcal{M}_T(X)$, it follows that

$$
\text{LS} \left(\left(\alpha_{C_k} \circ \operatorname{Avg}_{F_k} \right)_{k=1}^{\infty} \right) = \mathcal{M}_T(X).
$$

$\square$

We conclude this section by proving a result that does not rely on the measure μ neglecting shells.

Proposition 4.5.4. *There exists a sequence $(x_k)_{k=1}^{\infty}$ of points in X and a sequence $(r_k)_{k=1}^{\infty}$ of radii such that*

$$
\text{LS} \left(\left(\alpha_{B(x_k; r_k)} \circ \operatorname{Avg}_{F_k} \right)_{k=1}^{\infty} \right) \supseteq \partial_e \mathcal{M}_T(X).
$$

Proof. Let $\{\nu_i : i \in I\}$ be a countable dense subset of $\partial_e \mathcal{M}_T(X)$, where I is some countable indexing set, and let $\mathcal{N} = \{\mathcal{N}_i : i \in I\}$ be a partition of $\mathbb{N}$ into countably infinitely many infinite subsets. For each $k \in \mathbb{N}$, set $i(k) \in I$ such that $k \in \mathcal{N}_{i(k)}$.

For each $i \in I$, let $(\kappa(i, \ell))_{\ell=1}^{\infty}$ be a strictly increasing sequence such that

$$
\kappa(i, \ell) \in \mathcal{N}_i,
$$

$$
\left(F_{\kappa(i,\ell)} \right)_{\ell=1}^{\infty} \text{is tempered},
$$

which exists by Lemma 4.2.9. By the Lindenstrauss Ergodic Theorem, for each $i \in I$ exists $y_i \in X$ such that

$$\lim_{\ell \to \infty} \mathrm{Avg}_{F_{\kappa(i,\ell)}} f(y_i) = \int f \mathrm{d}\nu_i \qquad\qquad (\forall f \in C(X)).$$

Set $x_k = y_{i(k)}$.

For each $k \in \mathbb{N}$, choose $\delta_k > 0$ such that

$$\max_{g \in F_k} \left(L(g) \cdot \delta_k^{H(g)} \right) < 1/k,$$

and let $r_k \in (0, \delta_k)$ for all $k \in \mathbb{N}$. If $f \in \mathcal{L}(X)$, then

$$\left| \alpha_{B\left(x_{\kappa(i,\ell)}; r_{\kappa(i,\ell)}\right)} \left(\mathrm{Avg}_{\kappa(i,\ell)} f \right) - \int f \mathrm{d}\nu_i \right|$$

$$\leq \left| \alpha_{B\left(x_{\kappa(i,\ell)}; r_{\kappa(i,\ell)}\right)} \left(\mathrm{Avg}_{\kappa(i,\ell)} f \right) - \mathrm{Avg}_{\kappa(i,\ell)} f \left(x_{\kappa(i,\ell)} \right) \right| + \left| \mathrm{Avg}_{\kappa(i,\ell)} f \left(x_{\kappa(i,\ell)} \right) - \int f \mathrm{d}\nu_i \right|$$

$$= \left| \alpha_{B\left(x_{\kappa(i,\ell)}; r_{\kappa(i,\ell)}\right)} \left(\mathrm{Avg}_{\kappa(i,\ell)} f \right) - \mathrm{Avg}_{\kappa(i,\ell)} f \left(x_{\kappa(i,\ell)} \right) \right| + \left| \mathrm{Avg}_{\kappa(i,\ell)} f (y_i) - \int f \mathrm{d}\nu_i \right|$$

$$\leq \frac{1}{\kappa(i,\ell)} + \left| \mathrm{Avg}_{\kappa(i,\ell)} f (y_i) - \int f \mathrm{d}\nu_i \right|$$

$$\overset{\ell \to \infty}{\to} 0.$$

Therefore $\nu_i \in \mathrm{LS}\left(\left(\alpha_{B(x_k; r_k)} \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty} \right)$ for all $i \in I$. Since $\{\nu_i : i \in I\}$ is dense in $\partial_e \mathcal{M}_T(X)$, and $\mathrm{LS}\left(\left(\alpha_{B(x_k; r_k)} \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty} \right)$ is always closed, it follows that

$$\partial_e \mathcal{M}_T(X) \subseteq \mathrm{LS}\left(\left(\alpha_{B(x_k; r_k)} \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty} \right).$$

$\square$

4.6 Weak specification and maximal oscillation

Specification properties were initially introduced by R. Bowen in (Bowen, 1971) in the course of studying Axiom A diffeomorphisms. In the intervening decades, a considerable amount of effort has been put into the study of other specification-like properties -typically weaker than the Specification Property considered by Bowen- and the connections between them. For a broad overview of these specification-like properties and

the relations between them, we refer the reader to (Kwietniak et al., 2016), whose terminology we will be following.

Throughout this section, let (X, ρ) be a compact metric space, and let $T : \mathbb{N}_0 \curvearrowright X$ be an action of $\mathbb{N}_0$ on X by continuous (not necessarily invertible) maps. For $x \in X, k \in \mathbb{N}$, we define the *kth empirical measure of x* to be the Borel probability measure

$$\mu_{x,k} := \sum_{j=0}^{k-1} \delta_{T_j x},$$

where δ_y denotes the point mass at y, i.e. $\delta_y(A) = \chi_A(y)$. In light of Lemma 4.2.5, the study of local temporo-spatial differentiations is closely tied to the study of pointwise ergodic averages.

A point $x \in X$ is said to have *maximal oscillation* with respect to $T : \mathbb{N}_0 \curvearrowright X$ if

$$\mathrm{LS}\left((\mu_{x,k})_{k=1}^\infty\right) = \mathcal{M}_T(X).$$

This could be understood as the worst possible divergence for the sequence $(\mu_{x,k})_{k=1}^\infty$. M. Denker, C. Grillenberger, and K. Sigmund demonstrated the following prevalence result for points of maximal oscillation. Recall that a subset S of X is called *residual* if S contains a dense G_δ set.

Theorem 4.6.1. *(Denker et al., 2006, Proposition 21.18) If T has the Periodic Specification Property, then the set of points $x \in X$ with maximal oscillation is residual in X.*

Remark 4.6.2. In (Denker et al., 2006), what the authors call the Specification Property (defined there as Definition 21.1) is what (Kwietniak et al., 2016) calls the Periodic Specification Property, which is slightly stronger than what (Kwietniak et al., 2016) -and consequently we- call the Specification Property in Definition 4.6.5.

We introduce here a variation on and strengthening of the definition of maximal oscillation.

Definition 4.6.3. A *sampling family* is a family Π of functions $\mathbb{N} \to \mathbb{N}$ such that $\lim_{k\to\infty} \pi(k) = +\infty$ for all $k \in \mathbb{N}$. Given a sampling family Π, we say that a point $x \in X$ has *maximal oscillation relative to Π* if for every $\pi \in \Pi$, we have that

$$\mathrm{LS}\left((\mu_{x,\pi(k)})_{k=1}^\infty\right) = \mathcal{M}_T(X).$$

Maximal oscillation can then be recovered as the case where $\Pi = \{k \mapsto k\}$ consists solely of the identity function on $\mathbb{N}$.

Maximal oscillation describes the situation where not only does the sequence $(\mu_{x,k})_{k=1}^{\infty}$ diverge, but it diverges to the greatest extent possible. However, because $(\mu_{x,k})_{k=1}^{\infty}$ takes values in the compact space $\mathcal{M}(X)$, we know it will always have convergent subsequences, meaning this divergence will always "disappear" if we restrict our attention to an appropriate subsequence. Our notion of maximal oscillation relative to a sampling family allows us to strengthen the notion of maximal oscillation by prescribing the "worst-case scenario" divergence along a family of subsequences.

We now define a hierarchy of specification-like properties.

Definition 4.6.4. A *specification* is a finite sequence $\xi = \{([a_j, b_j], x_j)\}_{j=1}^{n}$ of finite subintervals $[a_j, b_j]$ of $\mathbb{N}$ and points $x_j \in X$. Given a function $\mathbf{M} : \mathbb{N} \to \mathbb{N}$, we say that the specification $\xi = \{([a_j, b_j], x_j)\}_{j=1}^{n}$ is $\mathbf{M}$-*spaced* if $a_j - b_{j-1} \geq \mathbf{M}(j)$ for all $j = 2, \ldots, n$. If $\mathbf{M}$ is the constant function $N \in \mathbb{N}$, then we say an $\mathbf{M}$-spaced specification is N-*spaced*.

Definition 4.6.5. Let $\xi = \{([a_j, b_j], x_j)\}_{j=1}^{n}$ be a specification, and let $\delta > 0$. We call a point $y \in X$ a δ-*tracing* of ξ if

$$\rho\left(T_i x_j, T_{a_j+i} y\right) < \delta \qquad \qquad (\forall j = 1, \ldots, n; i = 0, 1, \ldots, b_j - a_j).$$

(I) We call a family of functions $(\mathbf{M}_\delta : \mathbb{N} \to \mathbb{N})_{\delta \in (0,1)}$ a *modulus of specification* for (X, T) if every $\mathbf{M}_\delta$-spaced specification ξ admits a δ-tracing, and say that $T : \mathbb{N}_0 \curvearrowright X$ has the *Very Weak Specification Property*.

(II) If T admits a modulus of specification $(\mathbf{M}_\delta)_{\delta \in (0,1)}$ with the additional property that

$$\lim_{n \to \infty} \frac{\mathbf{M}_\delta(n)}{n} = 0 \qquad \qquad (\forall \delta \in (0, 1)),$$

then we say that T has the *Weak Specification Property*.

(III) If T admits a modulus of specification $(\mathbf{M}_\delta)_{\delta \in (0,1)}$ with the additional property that each $\mathbf{M}_\delta$ is a constant function, then we say that T has the *Specification Property*.

Intuitively, these specification-like properties mean that if we have some orbit segments that we want to approximate within δ, then we can find a point whose orbits are close to those segments as long as the segments are spaced far enough apart from each other. Clearly these specification properties are listed in ascending order of strength.

What we call the Weak Specification Property and Specification Property both have precedents in the literature. The Specification Property goes back to R. Bowen's original work (Bowen, 1971), and what we call here the Weak Specification Property can be found in (Marcus, 1980). See (Kwietniak et al., 2016) for a fuller historical discussion. However, to our knowledge, there is no precedent for what we term here the Very Weak Specification Property in the literature. Regardless, our results in this section do not rely on a modulus of specification $(\mathbf{M}_\delta)_{\delta \in (0,1)}$ satisfying the condition that $\mathbf{M}_\delta(n) = o(n)$ for all $\delta \in (0, 1)$, so we see fit to introduce this weaker specification-like property.

Our main theorem of this section is the following.

Theorem 4.6.6. *Let Π be a countable sampling family. Suppose $T : \mathbb{N}_0 \curvearrowright X$ has the Very Weak Specification Property. Then the set*

$$X^\Pi = \left\{ x \in X : \mathrm{LS}\left(\left(\mu_{x,\pi(k)} \right)_{k=1}^\infty \right) = \mathcal{M}_T(X) \text{ for all } \pi \in \Pi \right\}.$$

is residual.

Let $\mathcal{E}$ denote a countable dense subset of $\partial_e \mathcal{M}_T(X)$, and let

$$\mathcal{F} = \left\{ \sum_{i=1}^n \lambda_i \theta_i : n \in \mathbb{N}, \theta_i \in \mathcal{E}, \lambda_i \in \mathbb{Q} \cap [0,1], \sum_{i=1}^n \lambda_i = 1 \right\},$$

i.e. $\mathcal{F}$ is the set of all rational convex combinations of elements of $\mathcal{E}$. Then $\mathcal{F}$ is a countable dense subset of $\mathcal{M}_T(X)$ by the Krein-Millman Theorem. Further, let $\{f_h\}_{h=1}^\infty$ be an enumerated dense subset of $C(X)$.

Lemma 4.6.7. *Let Π be a sampling family. For $\nu \in \mathcal{F}, \epsilon > 0, H \in \mathbb{N}, k_0 \in \mathbb{N}, \pi \in \Pi$, set*

$$E(\nu, \epsilon, H, k_0, \pi) = \bigcap_{h=1}^H \left\{ x \in X : \exists k \geq k_0 \left(\left| \left(\frac{1}{\pi(k)} \sum_{j=0}^{\pi(k)-1} T_j f_h(x) \right) - \int f_h \mathrm{d}\nu \right| < \epsilon \right) \right\}.$$

If T has the Very Weak Specification Property, then $E(\nu, \epsilon, H, k_0, \pi)$ is a dense open subset of X.

Proof. Fix $H \in \mathbb{N}, \nu \in \mathcal{M}_T(X), \epsilon > 0, \pi \in \Pi$. Set

$$A_k = \left\{ x \in X : \left| \left(\frac{1}{\pi(k)} \sum_{j=0}^{\pi(k)-1} T_j f_h(x) \right) - \int f_h d\nu \right| < \epsilon \text{ for } h = 1, \dots, H \right\}.$$

Then $E(\nu, \epsilon, H, k_0) = \bigcup_{k=k_0}^{\infty} A_k$. Clearly $\bigcup_{k=k_0}^{\infty} A_k$ is open, leaving us to show it is dense.

Choose $\theta_0, \theta_1, \dots, \theta_{I-1} \in \mathcal{E}; \lambda_0, \lambda_1, \dots, \lambda_{I-1} \in [0, 1] \cap \mathbb{Q}$ such that

$$\nu = \sum_{i=0}^{I-1} \lambda_i \theta_i,$$

where we can assume without loss of generality that $\lambda_i > 0$ for all $i = 1, \dots, I$. Let $p_0, p_1, \dots, p_{I-1}, q \in \mathbb{N}$ such that

$$\lambda_i = \frac{p_i}{q} \qquad\qquad (i = 0, 1, \dots, I-1).$$

Let $y_0, y_1, \dots, y_{I-1} \in X$ such that $\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} T_j f_h(y_i) = \int f d\theta_i$ for $i = 0, 1, \dots, I-1$, which exist by the Birkhoff Ergodic Theorem. Choose $k_0 \in \mathbb{N}$ such that

$$k \geq k_0 \Rightarrow \left| \left(\frac{1}{k} \sum_{j=0}^{k-1} T_j f_h(y_i) \right) - \int f_h d\theta_i \right| < \epsilon/3 \qquad (i = 0, 1, \dots, I-1; h = 1, \dots, H).$$

Fix $x \in X, \eta > 0$. We will show that there exists $k \geq k_0$ and $y \in A_k$ such that

$$\rho(x, y) \leq \eta.$$

Since $f_1, \dots, f_H$ are uniformly continuous, there exists $\delta > 0$ such that

$$\forall z_1, z_2 \in X \, \forall h \in \{1, \dots, H\} \, (\rho(z_1, z_2) < \delta \Rightarrow |f_h(z_1) - f_h(z_2)| < \epsilon/3).$$

Assume without loss of generality that $\delta < \eta$.

Let $(\mathbf{M}_\delta)_{\delta \in (0,1)}$ be a modulus of specification for $T : \mathbb{N}_0 \curvearrowright X$. Fix

$$N = \max \{\mathbf{M}_\delta(1), \dots, \mathbf{M}_\delta(I+1)\}.$$

147

For $K \in \mathbb{N}$, define a sequence

$$a_{-1}^{(K)} \leq b_{-1}^{(K)} < a_0^{(K)} \leq b_0^{(K)} < a_1^{(K)} \leq b_1^{(K)} < a_2^{(K)} \leq b_2^{(K)} < \cdots < a_{I-1}^{(K)} \leq b_{I-1}^{(K)}$$

by

$$
\begin{aligned}
a_{-1}^{(K)} &= 0, & b_{-1}^{(K)} &= 0, \\
a_0^{(K)} &= N, & b_0^{(K)} &= a_0 + Kp_0 - 1, \\
a_1^{(K)} &= b_0^{(K)} + N, & b_1^{(K)} &= a_1^{(K)} + Kp_1 - 1, \\
a_2^{(K)} &= b_1^{(K)} + N, & b_2^{(K)} &= a_2^{(K)} + Kp_2 - 1, \\
&\;\;\vdots & & \\
a_{I-1}^{(K)} &= b_{I-2}^{(K)} + N, & b_{I-1}^{(K)} &= a_{I-1}^{(K)} + Kp_{I-1} - 1.
\end{aligned}
$$

Written explicitly, we have

$$a_i^{(K)} = (i+1)N + K\sum_{\ell=0}^{i-1} p_\ell,$$

$$b_i^{(K)} = (i+1)N - 1 + K\sum_{\ell=0}^{i} p_\ell.$$

Set

$$x_i = \begin{cases} x & \text{if } i = -1, \\[2mm] y_i & \text{if } 0 \leq i \leq I - 1 \end{cases}$$

Let $\xi^{(K)}$ be the specification

$$\xi^{(K)} = \left\{ \left(\left[a_i^{(K)}, b_i^{(K)} \right], x_i \right) \right\}_{i=-1}^{I-1}.$$

Then $\xi^{(K)}$ is $\mathbf{M}_\delta$-spaced, so by the Weak Specification Property, for each $K \in \mathbb{N}$ exists $y = y^{(K)} \in X$ such that $y^{(K)}$ is a δ-tracing of $\xi^{(K)}$. In particular, since $a_{-1}^{(K)} = 0 = b_{-1}^{(K)}, x_{-1} = x$, this means that $\rho(x, y) < \delta < \eta$. We claim that $y^{(K)} \in E(\nu, \epsilon, H, k_0, \pi)$ for sufficiently large K.

For $k \in \mathbb{N}$, set

$$K = K_k = \left\lfloor \frac{\pi(k) - IN - 1}{q} \right\rfloor,$$

so

$$b^{(K)}_{I-1} + 1 = IN + Kq + 1 \le \pi(k) \le IN + (K+1)q.$$

The following sketch of our argument motivates our definition of $y^{(K)}$. Let $f \in C(X)$. Then

$$\frac{1}{\pi(k)} \sum_{j=0}^{\pi(k)-1} f_h \left(T_j y^{(K)} \right)$$

$$\approx \mathrm{Avg}_{\left[a_0^{(K)},b_0^{(K)}\right] \cup \left[a_1^{(K)},b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)},b_{I-1}^{(K)}\right]} f_h \left(y^{(K)} \right)$$

$$= \frac{1}{Kp_0 + Kp_1 + \cdots Kp_{I-1}} \sum_{i=0}^{I-1} \sum_{j=a_i^{(K)}}^{b_i^{(K)}} f_h \left(T_j y^{(K)} \right)$$

$$= \frac{1}{Kp_0 + Kp_1 + \cdots Kp_{I-1}} \sum_{i=0}^{I-1} \left(b_i^{(K)} - a_i^{(K)} + 1 \right) \mathrm{Avg}_{\left[a_i^{(K)},b_i^{(K)}\right]} f_h \left(y^{(K)} \right)$$

$$= \frac{1}{Kq} \sum_{i=0}^{I-1} Kp_i \, \mathrm{Avg}_{\left[a_i^{(K)},b_i^{(K)}\right]} f_h \left(y^{(K)} \right)$$

$$= \sum_{i=0}^{I-1} \frac{p_i}{q} \, \mathrm{Avg}_{\left[a_i^{(K)},b_i^{(K)}\right]} f_h \left(y^{(K)} \right)$$

$$\approx \sum_{i=0}^{I-1} \frac{p_i}{q} \, \mathrm{Avg}_{\left[0,b_i^{(K)} - a_i^{(K)}\right]} f_h \left(x_i \right)$$

$$\approx \sum_{i=0}^{I-1} \frac{p_i}{q} \int f_h \mathrm{d}\theta_i$$

$$= \int f_h \mathrm{d}\nu,$$

where we write that $s(k) \approx t(k)$ if $|s(k) - t(k)| < \epsilon/3$ for sufficiently large $k \in \mathbb{N}$. So it will suffice to verify those three claims.

Claim (i): We first argue that

$$\left| \left(\frac{1}{\pi(k)} \sum_{j=0}^{\pi(k)-1} f_h \left(T_j y^{(K)} \right) \right) - \mathrm{Avg}_{\left[a_0^{(K)},b_0^{(K)}\right] \cup \left[a_1^{(K)},b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)},b_{I-1}^{(K)}\right]} f_h \left(y^{(K)} \right) \right| \le \frac{\epsilon}{3}$$

for sufficiently large $k \in \mathbb{N}$. We know that

$$
\frac{1}{\pi(k)} \sum_{j=0}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right)
$$

$$
= \frac{1}{\pi(k)} \left(\sum_{i=0}^{I-1} \sum_{j=a_i^{(K)}}^{b_i^{(K)}} f_h\left(T_j y^{(K)}\right) \right) + \frac{1}{\pi(k)} \left(\sum_{j=0}^{a_0^{(K)}} f_h\left(T_j y^{(K)}\right) \right)
$$

$$
+ \frac{1}{\pi(k)} \left(\sum_{i=0}^{I-2} \sum_{j=b_i^{(K)}+1}^{a_{i+1}^{(K)}} f_h\left(T_j y^{(K)}\right) \right) + \frac{1}{\pi(k)} \left(\sum_{j=b_{I-1}^{(K)}+1}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right) \right)
$$

$$
= \frac{Kq}{\pi(k)} \operatorname{Avg}_{\left[a_0^{(K)},b_0^{(K)}\right] \cup \left[a_1^{(K)},b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)},b_{I-1}^{(K)}\right]} f_h\left(y^{(K)}\right) + \frac{1}{\pi(k)} \left(\sum_{j=0}^{a_0^{(K)}} f_h\left(T_j y^{(K)}\right) \right)
$$

$$
+ \frac{1}{\pi(k)} \left(\sum_{i=0}^{I-2} \sum_{j=b_i^{(K)}+1}^{a_{i+1}^{(K)}} f_h\left(T_j y^{(K)}\right) \right) + \frac{1}{\pi(k)} \left(\sum_{j=b_{I-1}^{(K)}+1}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right) \right)
$$

$$
= \operatorname{Avg}_{\left[a_0^{(K)},b_0^{(K)}\right] \cup \left[a_1^{(K)},b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)},b_{I-1}^{(K)}\right]} f_h\left(y^{(K)}\right)
$$

$$
+ \frac{Kq - \pi(k)}{\pi(k)} \operatorname{Avg}_{\left[a_0^{(K)},b_0^{(K)}\right] \cup \left[a_1^{(K)},b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)},b_{I-1}^{(K)}\right]} f_h\left(y^{(K)}\right)
$$

$$
+ \frac{1}{\pi(k)} \left(\sum_{j=0}^{a_0^{(K)}} f_h\left(T_j y^{(K)}\right) \right) + \frac{1}{\pi(k)} \left(\sum_{i=0}^{I-2} \sum_{j=b_i^{(K)}+1}^{a_{i+1}^{(K)}} f_h\left(T_j y^{(K)}\right) \right)
$$

$$
+ \frac{1}{\pi(k)} \left(\sum_{j=b_{I-1}^{(K)}+1}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right) \right).
$$

Therefore

$$
\left| \left(\frac{1}{\pi(k)} \sum_{j=0}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right) \right) - \mathrm{Avg}_{\left[a_0^{(K)}, b_0^{(K)}\right] \cup \left[a_1^{(K)}, b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)}, b_{I-1}^{(K)}\right]} f_h\left(y^{(K)}\right) \right|
$$

$$
\leq \left| \frac{Kq - \pi(k)}{\pi(k)} \mathrm{Avg}_{\left[a_0^{(K)}, b_0^{(K)}\right] \cup \left[a_1^{(K)}, b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)}, b_{I-1}^{(K)}\right]} f_h\left(y^{(K)}\right) \right|
$$

$$
+ \left| \frac{1}{\pi(k)} \left(\sum_{j=0}^{a_0^{(K)}} f_h\left(T_j y^{(K)}\right) \right) \right|
$$

$$
+ \left| \frac{1}{\pi(k)} \left(\sum_{i=0}^{I-2} \sum_{j=b_i^{(K)}+1}^{a_{i+1}^{(K)}} f_h\left(T_j y^{(K)}\right) \right) \right|
$$

$$
+ \left| \frac{1}{\pi(k)} \sum_{j=b_{I-1}^{(K)}+1}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right) \right|
$$

$$
= \frac{\pi(k) - Kq}{\pi(k)} \left| \mathrm{Avg}_{\left[a_0^{(K)}, b_0^{(K)}\right] \cup \left[a_1^{(K)}, b_1^{(K)}\right] \cup \cdots \left[a_{I-1}^{(K)}, b_{I-1}^{(K)}\right]} f_h\left(y^{(K)}\right) \right|
$$

$$
+ \left| \frac{1}{\pi(k)} \left(\sum_{j=0}^{N} f_h\left(T_j y^{(K)}\right) \right) \right|
$$

$$
+ \left| \frac{1}{\pi(k)} \left(\sum_{i=0}^{I-2} \sum_{j=b_i^{(K)}+1}^{b_i^{(K)}+N} f_h\left(T_j y^{(K)}\right) \right) \right| + \left| \frac{1}{\pi(k)} \sum_{j=IN+Kq+1}^{\pi(k)-1} f_h\left(T_j y^{(K)}\right) \right|
$$

$$
\leq \frac{\pi(k) - Kq}{\pi(k)} \|f_h\|_{C(X)} + \frac{N+1}{\pi(k)} \|f_h\|_{C(X)}
$$

$$
+ \frac{(I-1)N}{\pi(k)} \|f_h\|_{C(X)} + \frac{\pi(k) - IN + Kq + 1}{\pi(k)} \|f_h\|_{C(X)}
$$

$$
\leq \left[\frac{IN+1}{\pi(k)} + \frac{N+1}{\pi(k)} + \frac{(I-1)N}{\pi(k)} + \frac{q}{\pi(k)} \right] \cdot \|f_h\|_{C(X)}
$$

$$
\overset{k\to\infty}{\to} 0.
$$

This establishes our estimate for large k.

Claim (ii): We next argue that

$$
\left| \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \mathrm{Avg}_{\left[a_i^{(K)}, b_i^{(K)}\right]} f_h\left(y^{(K)}\right) \right) - \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \mathrm{Avg}_{\left[0, b_i^{(K)} - a_i^{(K)}\right]} f_h\left(x_i\right) \right) \right| < \frac{\epsilon}{3}
$$

for all $k \in \mathbb{N}$. To see this, we can note that

$$\left| \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \operatorname{Avg}_{\left[a_i^{(K)}, b_i^{(K)}\right]} f_h\left(y^{(K)}\right) \right) - \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \operatorname{Avg}_{\left[0, b_i^{(K)} - a_i^{(K)}\right]} f_h\left(x_i\right) \right) \right|$$

$$= \left| \sum_{i=0}^{I-1} \frac{p_i}{q} \frac{1}{b_i^{(K)} - a_i^{(K)} + 1} \sum_{j=0}^{b_i^{(K)} - a_i^{(K)}} \left(f_h\left(T_{j+a_i^{(K)}} y^{(K)}\right) - f_h\left(T_j x_i\right) \right) \right|$$

$$= \left| \sum_{i=0}^{I-1} \frac{p_i}{q} \frac{1}{K p_i} \sum_{j=0}^{K p_i - 1} \left(f_h\left(T_{j+a_i^{(K)}} y^{(K)}\right) - f_h\left(T_j x_i\right) \right) \right|$$

$$\leq \sum_{i=0}^{I-1} \frac{p_i}{q} \frac{1}{K p_i} \sum_{j=0}^{K p_i - 1} \left| f_h\left(T_{j+a_i^{(K)}} y^{(K)}\right) - f_h\left(T_j x_i\right) \right|$$

$$(\dagger) < \sum_{i=0}^{I-1} \frac{p_i}{q} \frac{1}{K p_i} \sum_{j=0}^{K p_i - 1} \frac{\epsilon}{3}$$

$$= \frac{\epsilon}{3},$$

where the estimate ($\dagger$) follows from the fact that y is a δ-tracing of $\xi^{(K)}$.

Claim (iii): Our third step is to show that

$$\left| \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \operatorname{Avg}_{\left[0, b_i^{(K)} - a_i^{(K)}\right]} f_h\left(x_i\right) \right) - \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \int f_h \mathrm{d}\theta_i \right) \right| < \frac{\epsilon}{3}$$

for sufficiently large $k \in \mathbb{N}$. This follows because

$$\left| \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \operatorname{Avg}_{\left[0, b_i^{(K)} - a_i^{(K)}\right]} f_h\left(x_i\right) \right) - \left(\sum_{i=0}^{I-1} \frac{p_i}{q} \int f_h \mathrm{d}\theta_i \right) \right|$$

$$= \left| \sum_{i=0}^{I-1} \frac{p_i}{q} \left(\operatorname{Avg}_{\left[0, b_i^{(K)} - a_i^{(K)}\right]} f_h\left(x_i\right) - \int f_h \mathrm{d}\theta_i \right) \right|$$

$$\leq \sum_{i=0}^{I-1} \frac{p_i}{q} \left| \operatorname{Avg}_{\left[0, b_i^{(K)} - a_i^{(K)}\right]} f_h\left(x_i\right) - \int f_h \mathrm{d}\theta_i \right|$$

$$= \sum_{i=0}^{I-1} \frac{p_i}{q} \left| \left(\frac{1}{K p_i} \sum_{j=0}^{K p_i - 1} f_h\left(T_j x_i\right) \right) - \int f_h \mathrm{d}\theta_i \right|$$

If k is sufficiently large that

$$\left|\left(\frac{1}{Kp_i}\sum_{j=0}^{Kp_i-1}f_h\left(T_jx_i\right)\right)-\int f_h\mathrm{d}\theta_i\right|<\frac{\epsilon}{3}\qquad\text{(for }i=0,1,\ldots,I-1\text{)},$$

then

$$\sum_{i=0}^{I-1}\frac{p_i}{q}\left|\left(\frac{1}{Kp_i}\sum_{j=0}^{Kp_i-1}f_h\left(T_jx_i\right)\right)-\int f_h\mathrm{d}\theta_i\right|<\sum_{i=0}^{I-1}\frac{p_i}{q}\frac{\epsilon}{3}=\frac{\epsilon}{3}.$$

Taking these three claims together, we can say that

$$
\left|\left(\frac{1}{\pi(k)}\sum_{j=0}^{\pi(k)-1}f_h\left(T_jy^{(K)}\right)\right)-\int f\mathrm{d}\nu\right|
$$

$$
\leq\left|\left(\frac{1}{\pi(k)}\sum_{j=0}^{\pi(k)-1}f_h\left(T_jy^{(K)}\right)\right)-\mathrm{Avg}_{\left[a_0^{(K)},b_0^{(K)}\right]\cup\left[a_1^{(K)},b_1^{(K)}\right]\cup\cdots\left[a_{I-1}^{(K)},b_{I-1}^{(K)}\right]}f_h\left(y^{(K)}\right)\right|
$$

$$
+\left|\left(\sum_{i=0}^{I-1}\frac{p_i}{q}\mathrm{Avg}_{\left[a_i^{(K)},b_i^{(K)}\right]}f_h\left(y^{(K)}\right)\right)-\left(\sum_{i=0}^{I-1}\frac{p_i}{q}\mathrm{Avg}_{\left[0,b_i^{(K)}-a_i^{(K)}\right]}f_h\left(x_i\right)\right)\right|
$$

$$
+\left|\left(\sum_{i=0}^{I-1}\frac{p_i}{q}\mathrm{Avg}_{\left[0,b_i^{(K)}-a_i^{(K)}\right]}f_h\left(x_i\right)\right)-\left(\sum_{i=0}^{I-1}\frac{p_i}{q}\int f_h\mathrm{d}\theta_i\right)\right|
$$

$$
<\frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3}
$$

$$
=\epsilon
$$

for sufficiently large $k\in\mathbb{N}$.

For each $h\in\{1,\ldots,H\}$, choose $k_h\in\mathbb{N}$ such that

$$k\geq k_h\Rightarrow\left|\left(\frac{1}{\pi(k)}\sum_{j=0}^{\pi(k)-1}f_h\left(T_jy^{(K)}\right)\right)-\int f\mathrm{d}\nu\right|<\epsilon.$$

Then if $k\geq\max\{k_0,k_1,\ldots,k_H\}$, it follows that $y^{(K)}\in E(\nu,\epsilon,H,k_0,\pi)$. $\qquad\square$

Proof of Theorem 4.6.6. We can metrize $\mathcal{M}(X)$ with the metric $\mathrm{dist}:\mathcal{M}(X)\times\mathcal{M}(X)\to[0,1]$ defined by

$$\mathrm{dist}(\beta_1,\beta_2)=\sum_{h=1}^{\infty}\min 2^{-h}\left\{\left|\int f_h\mathrm{d}(\beta_1-\beta_2)\right|,1\right\}.$$

For $\nu \in \mathcal{F}, k_0 \in \mathbb{N}, n \in \mathbb{N}, \pi \in \Pi$, write

$$B(\nu, n, k_0) = \{x \in X : \exists k \geq k_0 \; (\text{dist}(\mu_{x,k}, \nu) < 1/n)\}.$$

Choose $H_n \in \mathbb{N}$ such that $2^{-H_n} < 1/(2n)$. We claim that

$$B(\nu, n, k_0, \pi) \supseteq E(\nu, 1/(2n), H_n, k_0).$$

If $x \in E(\nu, 1/(2n), H_n, k_0, \pi)$, then there exists $k \geq k_0$ such that

$$\begin{aligned}
\text{dist}(\mu_{x,\pi(k)}, \nu) &= \sum_{h=1}^{\infty} 2^{-h} \min\left\{\left|\int f_h \mathrm{d}\left(\mu_{x,\pi(j)} - \nu\right), 1\right.\right\} \\
&< 2^{-1}\frac{1}{2n} + 2^{-2}\frac{1}{2n} + \cdots + 2^{-H_n}\frac{1}{2n} + \sum_{h=H_\epsilon+1}^{\infty} 2^{-h} \\
&< \frac{1}{2n} + \frac{1}{2n} \\
&= 1/n.
\end{aligned}$$

Thus $x \in B(\nu, n, k_0, \pi)$.

We claim that $X' \supseteq \bigcap_{\pi \in \Pi} \bigcap_{\nu \in \mathcal{F}} \bigcap_{n=1}^{\infty} \bigcap_{k_0=1}^{\infty} B(\nu, n, k_0, \pi)$. Let $x \in \bigcap_{\pi \in \Pi} \bigcap_{\nu \in \mathcal{F}} \bigcap_{n=1}^{\infty} \bigcap_{k_0=1}^{\infty} B(\nu, n, k_0, \pi)$, and consider some $\nu \in \mathcal{M}_T(X)$. Choose a sequence $(\nu_\ell)_{\ell=1}^{\infty}$ in $\mathcal{F}$ such that $\text{dist}(\nu, \nu_\ell) < 1/\ell$ for all $\ell \in \mathbb{N}$. Construct a sequence $(k_\ell)_{\ell=1}^{\infty}$ in $\mathbb{N}$ recursively as follows:

- **Basis step:** Choose $k_1 \in \mathbb{N}$ such that $\text{dist}\left(\mu_{x,\pi(k_1)}, \nu_1\right) < 1$, which exists because $x \in B(\nu_n, n, 1, \pi)$.

- **Recursive step:** Suppose we've chosen $k_1 < k_2 < \cdots < k_\ell$ such that $\text{dist}\left(\mu_{x,\pi(k_n)}, \nu_n\right) < 1/n$ for $n = 1, \ldots, \ell$. Chose $k_{\ell+1} \geq k_\ell + 1$ such that $\text{dist}\left(\mu_{x,\pi(k_{\ell+1})}, \nu_{\ell+1}\right) < 1/(\ell+1)$, which exists because $x \in B(\nu_{\ell+1}, \ell+1, k_\ell + 1, \pi)$.

It follows then that

$$\text{dist}\left(\mu_{x,k_\ell}, \nu\right) \leq \text{dist}\left(\mu_{x,k_\ell}, \nu_\ell\right) + \text{dist}\left(\nu_\ell, \nu\right) < 2/\ell \overset{\ell\to\infty}{\Rightarrow} 0,$$

i.e. $\nu \in \text{LS}\left((\mu_{x,k})_{k=1}^{\infty}\right)$.

But $\bigcap_{\nu \in \mathcal{F}} \bigcap_{n=1}^{\infty} \bigcap_{k_0=1}^{\infty} B(\nu, n, k_0)$ is a countable intersection of residual sets, and thus itself residual.

$\square$

Corollary 4.6.8. *Let* $\mathbf{F} = (\{0, 1, \ldots, k-1\})_{k=1}^{\infty}$, *and suppose that* $T : \mathbb{N}_0 \curvearrowright X$ *is a Hölder action on* X *that has the Very Weak Specification Property. Suppose* Π *is a countable sampling family. Then the set of* $x \in X$ *such that* $\mathrm{LS}\left(\left(\alpha_{B\left(x; r_{\pi(k)}\right)} \circ \mathrm{Avg}_{F_{\pi(k)}}\right)_{k=1}^{\infty}\right) = \mathcal{M}_T(X)$ *for all* $(r_k)_{k=1}^{\infty}$ *that decay* $(X, \rho, H, L, \mathbf{F})$-*fast and* $\pi \in \Pi$ *is a residual subset of* X.

Proof. Lemma 4.2.5 tells us that this is exactly the set considered in Theorem 4.6.6. $\square$

Our Theorem 4.6.6 strengthens the following result of J. Li and M. Wu, since the Specification Property implies the Very Weak Specification Property.

Corollary 4.6.9. *(Li and Wu, 2016, Theorem 1.3) Suppose* $T : \mathbb{N}_0 \curvearrowright X$ *has the Specification Property, and let* $f \in C_{\mathbb{R}}(X)$ *be a real-valued continuous function on* X. *Then the set*

$$\left\{ x \in X : \liminf_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f\left(T_j x\right) = \underline{a}(f), \ \limsup_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f\left(T_j x\right) = \overline{a}(f) \right\}$$

is residual.

Proof. Let $\Pi = \{k \mapsto k\}$ be the sampling family consisting solely of the identity function $\mathbb{N} \to \mathbb{N}$, and consider $x \in X^{\Pi}$. Since the Specification Property implies the Very Weak Specification Property, Theorem 4.6.6 tells us that X^{Π} is residual. Let $\theta_1, \theta_2 \in \partial_e \mathcal{M}_T(X)$ such that

$$\int f \, d\theta_1 = \underline{a}(f),$$
$$\int f \, d\theta_2 = \overline{a}(f).$$

Then there exist $k_1^{(i)} < k_2^{(i)} < k_3^{(i)} \cdots$ for $i = 1, 2$ such that $\lim_{\ell \to \infty} \mu_{x,k_\ell^{(i)}} = \theta_i$. Thus

$$\underline{a}(f) \leq \liminf_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T_j x) \qquad \leq \lim_{\ell \to \infty} \frac{1}{k_\ell^{(1)}} \sum_{j=0}^{k_\ell^{(1)}-1} f(T_j x) = \underline{a}(f)$$

$$\Rightarrow \liminf_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T_j x) = \underline{a}(f),$$

$$\overline{a}(f) \geq \limsup_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T_j x) \qquad \geq \lim_{\ell \to \infty} \frac{1}{k_\ell^{(2)}} \sum_{j=0}^{k_\ell^{(2)}-1} f(T_j x) = \overline{a}(f)$$

$$\Rightarrow \limsup_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T_j x) = \overline{a}(f).$$

Therefore

$$X^{\Pi} \subseteq \left\{ x \in X : \liminf_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T_j x) = \underline{a}(f), \limsup_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T_j x) = \overline{a}(f) \right\},$$

meaning the latter is residual. $\qquad\square$

Non-autonomous temporo-spatial differentiations for group endomorphisms

The remainder of this chapter is a reproduction of the article (Assani and Young, 2023), a joint work between the author and his advisor, I. Assani. It is presented without any changes.

In (Assani and Young, 2022), we introduced the notion of a spatial-temporal differentiation problem. Here, we introduce a generalization of this concept to the setting of non-autonomous dynamical systems, and prove probabilistic and topological results about certain random spatial-temporal differentiations on compact abelian metrizable groups.

This paper is organized as follows:

- In Section 5.1, we provide a definition of non-autonomous dynamical systems for our purposes. We also describe what a spatial-temporal differentiation problem would look like in this non-autonomous setting.

- In Section 5.2, we introduce the notion of uniform distribution, and describe how uniform distribution in a compact group is related to the representation theory of that group. We end in proving a metric result about the uniform distribution of the trajectory of a point under a sequence of group endomorphisms under the hypothesis that the group endomorphisms satisfy a property we call the *Difference Property*.

- In Section 5.3, we consider questions about when the Difference Property makes the group endomorphisms surjective, and whether a sequence with the Difference Property can exist on a given group.

- In Section 5.4, we prove a probabilistic result about non-autonomous spatial-temporal differentiations relative to a sequence of group endomorphisms with the Difference Property and a sequence of concentric balls with rapidly decaying radii, demonstrating that the set of $x \in G$ which generate well-behaved spatial-temporal differentiations is of full measure.

- In Section 5.5, we prove a probabilistic result about uniformly distributed sequences of the form $(T_n \cdots T_1 g_{\Lambda_n})_{n=0}^{\infty}$, where $(g_n)_{n=0}^{\infty} \in G^{\mathbb{N}_0}$ and $(\Lambda_n)_{n=1}^{\infty}$ is an increasing sequence of natural numbers.

- In Section 5.6, we prove a topological counterpoint to Theorem 5.4.2, demonstrating that the set of $x \in G$ which generate pathological spatial-temporal differentiations is comeager.

We thank the referee for their careful reading of this paper.

5.1 Introducing non-autonomous dynamical systems

Our definition of a non-autonomous dynamical system is inspired by the "process formulation" found in (Kloeden and Rasmussen, 2011), though adapted for our ergodic-theoretic purposes. We state the definition in excess generality, because it is more important to us that the definition capture the *concept* of non-autonomy; the study of (autonomous) dynamical systems comes in many diverse flavors, so we would like our definition of non-autonomous dynamical systems to reflect that diversity.

Let $\mathbb{N}_0 = \mathbb{Z} \cap [0, \infty)$ be the semigroup of nonnegative integers (distinguished from the set $\mathbb{N}$ of strictly positive integers), and let X be an object in a category $\mathcal{C}$. Let $\mathrm{Hom}_\mathcal{C}(X, X)$ denote the semigroup of endomorphisms on X in the category $\mathcal{C}$. A *non-autonomous dynamical system* is a pair (X, τ), where τ is a family of maps $\{\tau(s, t) \in \mathrm{Hom}_\mathcal{C}(X, X)\}_{s,t \in \mathbb{N}_0, s \geq t}$ satisfying the following conditions.

1. $\tau(s, s) = \mathrm{id}_X$ for all $s \in \mathbb{N}_0$

2. $\tau(s, u) = \tau(s, t)\tau(t, u)$ for all $s, t, u \in \mathbb{N}_0, s \geq t \geq u$,

where the composition of endomorphisms of X is abbreviated as multiplication. We refer to τ as the *process*.

The essential difference between an autonomous and a non-autonomous system is that the transition map $\tau(s, t)$ is dependent on both the "starting time" t and the "ending time" s. The system would be autonomous if it had the additional property that $\tau(s, t) = \tau(s - t, 0)$ for all $s, t \in \mathbb{N}_0, s \geq t$, indicating that the transition map depends only on the elapsed time between t and s.

Similar to how an autonomous dynamical system can be treated in terms of either an action of $\mathbb{N}_0$ on a phase space, or equivalently in terms of its generating transformation T, a non-autonomous dynamical system as we have formulated it above can be understood in terms of a family of generators $T_t = \tau(t, t - 1), t \in \mathbb{N}$. Likewise, a family of generators $\{T_t \in \mathrm{Hom}_\mathcal{C}(X, X)\}_{t \in \mathbb{N}}$ can be understood as generating a non-autonomous

dynamical system by

$$\tau(s, t) = \tau(s, s-1)\tau(s-1, s-2)\cdots\tau(t+1, t)$$

$$= T_s T_{s-1} \cdots T_{t+1}.$$

The approaches are equivalent, but we will typically be approaching these non-autonomous systems from the perspective of starting with the generators $(T_n)_{n=1}^{\infty}$ and building $\tau(\cdot, \cdot)$ from that sequence.

For our purposes, that category $\mathcal{C}$ will be the category whose objects are compact topological spaces X endowed with Borel probability measures μ, and whose morphisms are continuous maps. We will not in general assume these maps are measure-preserving. Though the assumption that maps are measure-preserving is typically vital in the autonomous setting, we will eventually be considering situations where interesting results are possible without the explicit assumption that the maps in question are measure-preserving. For measurable sets F with $\mu(F) > 0$, set $\alpha_F(f) = \frac{1}{\mu(F)} \int_F f \mathrm{d}\mu$. We are interested in questions of the following forms:

- Let $(F_k)_{k=1}^{\infty}$ be a sequence of Borel subsets of X for which $\mu(F_k) > 0$, and let $f \in L^{\infty}(X, \mu)$. Then what can be said about the limiting behavior of

$$\left(\alpha_{F_k}\left(\frac{1}{k}\sum_{i=0}^{k-1} T_i T_{i-1}\cdots T_1 f\right)\right)_{k=1}^{\infty} ?$$

- Suppose the $F_k = F_k(x)$ are "indexed" by $x \in X$. Then can we make any probabilistic claims about the generic behavior of the sequence $\left(\alpha_{F_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1} T_i T_{i-1}\cdots T_1 f\right)\right)_{k=1}^{\infty}$?

- Under the same conditions, can we make any topological claims about the generic behavior of the sequence $\left(\alpha_{F_k(x)}\left(\frac{1}{k}\sum_{i=0}^{k-1} T_i T_{i-1}\cdots T_1 f\right)\right)_{k=1}^{\infty}$?

Theorem 5.4.2 is a result of the second type, describing the probabilistically generic behavior of a spatial-temporal differentiation along the sequence $B_k(x)$, where $B_k(x)$ is a ball centered at x with radius decaying rapidly to 0. Theorems 5.6.2 and 5.6.7 are of the third type, describing the topologically generic behavior of a spatial-temporal differentiation along the sequence $B_k(x)$.

5.2 Uniform distribution and harmonic analysis

Before proceeding, we define the notion of uniform distribution. The study of uniformly distributed sequences began with Weyl's investigation of "uniform distribution modulo 1", expanding on Kronecker's Theorem in Diophantine approximation (Weyl, 1968). This notion was then extended by Hlawka to apply to compact probability spaces (Hlawka, 1956). The study of uniform distribution in compact groups in particular was first initiated by Eckmann in (Eckmann, 1943); though Eckmann's initial definition of uniform distribution for compact groups contained a significant error, the initial paper still contained several foundational results in the theory, including the Weyl Criterion for Compact Groups (Proposition 5.2.5). For a more through history of the topic, the reader is referred to the note at the end of 4.1 in (Kuipers and Niederreiter, 2012).

Definition 5.2.1. Let X be a compact Hasudorff topological space endowed with a regular Borel probability measure μ. A sequence $(x_n)_{n=0}^\infty$ in X is called *uniformly distributed* with respect to the measure μ if

$$\frac{1}{k}\sum_{i=0}^{k-1} f(x_i) \overset{k\to\infty}{\to} \int f \mathrm{d}\mu$$

for all $f \in C(X)$.

Let G denote a compact topological group with identity element 1 and Haar probability measure μ. Throughout this section, we will be dealing only with the dynamics of compact groups, so G will always denote a compact group, μ will always refer to the Haar probability measure on the compact group G, and $\mathrm{Bo}(X)$ will always refer to the Borel σ-algebra on a topological space X. We will also write $L^p(G) := L^p(G, \mu)$, taking the measure μ to be understood.

When it comes to topological groups, the uniform distribution of topological groups can be characterized in terms of the representation theory of the group. We review here some important concepts from the representation theory of topological groups so that we may state this relation. Our brisk summary of the basic representation theory of compact groups mostly follows (Folland, 2016).

Let $\mathbb{U}(\mathcal{H})$ denote the group of unitary operators on a Hilbert space $\mathcal{H}$, where we endow $\mathbb{U}(\mathcal{H})$ with the strong operator topology. A *unitary representation of G on $\mathcal{H}$* is a continuous group homomorphism $\pi : G \to \mathbb{U}(\mathcal{H})$. Though non-unitary representations exist, we will be dealing here exclusively with unitary representations, so we do not bother to define non-unitary representations.

We call a closed subspace $\mathcal{M}$ of $\mathcal{H}$ an *invariant subspace* for a unitary representation $\pi : G \to \mathbb{U}(\mathcal{H})$ if $\pi(x)\mathcal{M} \subseteq \mathcal{M}$ for all $x \in G$. Since $\pi(x)|_{\mathcal{M}} : \mathcal{M} \to \mathcal{M}$ is unitary on $\mathcal{M}$, we call $\pi^{\mathcal{M}} : x \mapsto \pi(x)|_{\mathcal{M}} \in \mathbb{U}(\mathcal{M})$ the *subrepresentation* of π corresponding to $\mathcal{M}$. A unitary representation $\pi : G \to \mathbb{U}(\mathcal{H})$ is called *irreducible* if its only invariant subspaces are $\mathcal{H}$ and $\{0\}$. Two unitary representations $\pi_1 : G \to \mathbb{U}(\mathcal{H}_1), \pi_2 : G \to \mathbb{U}(\mathcal{H}_2)$ are called *unitarily equivalent* if there exists a unitary map $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $\pi_2(x) = U\pi_1(x)U^{-1}$ for all $x \in G$. We denote by $[\pi]$ the unitary equivalence class of a representation $\pi : G \to \mathbb{U}(\mathcal{H})$.

Fact 5.2.2. *If G is a compact group, then for every irreducible unitary representation $\pi : G \to \mathbb{U}(\mathcal{H})$, the space $\mathcal{H}$ is of finite dimension.*

Proof. (Folland, 2016, Theorem 5.2). $\qquad\square$

Given a unitary representation $\pi : G \to \mathbb{U}(\mathcal{H})$, we define the *dimension* $\dim \pi$ of π to be $\dim(\mathcal{H})$; since a unitary equivalence of representations induces a unitary isometry between the spaces they act on, we can conclude that the dimension of a representation is invariant under unitary equivalence. We use $\hat{G}$ to denote the family of unitary equivalence classes of irreducible unitary representations of G. We note that this notation is consistent with the use of $\hat{G}$ to refer to the Pontryagin dual of a locally compact abelian group G, since when G is locally compact abelian, the irreducible unitary representations are exactly the continuous homomorphisms $G \to \mathbb{S}^1 \subseteq \mathbb{C}$.

In particular, there will always exist at least one irreducible representation of dimension 1, specifically the map $x \mapsto 1 \in \mathbb{C}$. We call this the *trivial representation* of G.

Given an irreducible unitary representation $\pi : G \to \mathbb{U}(\mathcal{H})$, we define the *matrix elements* of π to be the functions $G \to \mathbb{C}$ given by

$$x \mapsto \langle \pi(x)u, v \rangle \qquad\qquad (u, v \in \mathcal{H}).$$

Because π is continuous with respect to the strong operator topology on $\mathcal{H}$, it follows that the matrix elements are continuous. Given an orthonormal basis $\{e_i\}_{i=1}^{n}$ of $\mathcal{H}$, we define the functions $\{\pi_{i,j}\}_{i,j=1}^{n}$ by

$$\pi_{i,j}(x) = \langle \pi(x)e_i, e_j \rangle.$$

These $\pi_{i,j}$ in fact define matrix entries for π in the basis $\{e_i\}_{i=1}^n$.

In particular, the trivial representation will have constant matrix elements.

This is sufficient framework to state the results we will be drawing on.

Fact 5.2.3. *[Peter-Weyl Theorem] Let G be a compact group, and let $V \subseteq C(G)$ be the subspace of $C(G)$ spanned by*

$$\left\{ \pi_{p,q} : p, q = 1, \dots, \dim \pi; [\pi] \in \hat{G} \right\}.$$

Then V is dense in $C(G)$ with respect to the uniform norm. Furthermore,

$$\left\{ \sqrt{\dim \pi}\, \pi_{p,q} : p, q = 1, \dots, \dim \pi; [\pi] \in \hat{G} \right\}$$

is an orthonormal basis for $L^2(G)$.

Proof. (Folland, 2016, Theorem 5.12). $\qquad\square$

Corollary 5.2.4. *If $\pi : G \to \mathbb{U}(\mathcal{H})$ is a nontrivial irreducible unitary representation, then $\int \pi_{p,q} \mathrm{d}\mu = 0$ for all $(p, q) \in \{1, 2, \dots, \dim \pi\}^2$.*

Proof.

$$\int \pi_{p,q} \mathrm{d}\mu = \int \pi_{p,q} \cdot 1 \mathrm{d}\mu = \langle \pi_{p,q}, 1 \rangle_{L^2(G)} = 0,$$

since the constant function 1 is a normalized matrix term of the trivial representation. $\qquad\square$

Finally, we return to the subject of uniform distribution.

Proposition 5.2.5 (Weyl Criterion for Compact Groups)**.** *Let G be a compact group, and $(x_n)_{n=0}^\infty$ a sequence in G. Then the following are equivalent.*

1. *The sequence $(x_n)_{n=0}^\infty$ is uniformly distributed in G.*

2. *For all nontrivial irreducible unitary representations $\pi : G \to \mathbb{U}(\mathcal{H})$, and all $(p, q) \in \{1, 2, \dots, \dim \pi\}^2$ we have*

$$\frac{1}{k} \sum_{i=0}^{k-1} \pi_{p,q}(x_i) \overset{k \to \infty}{\to} 0.$$

3. For all nontrivial irreducible unitary representations $\pi : G \to \mathbb{U}(\mathcal{H})$, we have

$$\frac{1}{k} \sum_{i=0}^{k-1} \pi(x_i) \overset{k \to \infty}{\Rightarrow} \mathbf{0},$$

where $\mathbf{0}$ denotes the zero operator on $\mathcal{H}$, and the convergence is meant in the operator norm.

Proof. (Kuipers and Niederreiter, 2012, Chapter 4, Theorem 1.3). $\qquad \square$

Lemma 5.2.6. *Let* $\varphi : G \twoheadrightarrow G_1$ *be a continuous surjective group homomorphism, and let* $(x_n)_{n=1}^{\infty}$ *be a uniformly distributed sequence in* G. *Then* $(\varphi(x_n))_{n=1}^{\infty}$ *is uniformly distributed in* G_1.

Proof. Let $\pi : G_1 \to \mathbb{U}(\mathcal{H})$ be a nontrivial unitary representation of G_1. Then $\pi \circ \varphi$ is a nontrivial unitary representation of G, since φ is surjective. Therefore

$$\frac{1}{k} \sum_{i=0}^{k-1} \pi(\varphi(x_i)) = \frac{1}{k} \sum_{i=0}^{k-1} (\pi \circ \varphi)(x_i) \qquad\qquad \overset{k \to \infty}{\Rightarrow} \mathbf{0}.$$

We can thus apply Proposition 5.2.5. $\qquad \square$

Remark 5.2.7. Lemma 5.2.6 is listed as Theorem 1.6 in Chapter 4 of (Kuipers and Niederreiter, 2012). We include the proof here for the sake of self-containment. Lemma 5.2.6 will be important when proving Theorem 5.5.9.

We will be interested specifically in the case where the Weyl Criterion only requires us to "check" countably many matrix element functions, or equivalently where $\hat{G}$ is countable. It turns out that this is tantamount to a metrizability assumption.

Lemma 5.2.8. *Let* G *be a compact topological group. Then the following are equivalent.*

1. $C(G)$ *is separable as a vector space with the uniform norm.*

2. $L^2(G)$ *is separable as a Hilbert space.*

3. *The family* $\hat{G}$ *is countable.*

4. G *is metrizable.*

Proof. (1)$\Rightarrow$(2) If $C(G)$ admits a countable set with dense span in $C(G)$, then that same set has dense span in $L^2(G)$. Thus $L^2(G)$ is separable.

(2)$\Rightarrow$(1) If $L^2(G)$ is separable, then every orthonormal basis of $L^2(G)$ is countable, including the family of matrix terms. But these matrix terms have dense span in $C(G)$, so $C(G)$ is separable.

(2) $\iff$ (3) If there are only countably many unitary equivalence classes of irreducible unitary representations of G (which are all necessarily finite-dimensional), then $L^2(G)$ is separable by the Peter-Weyl Theorem. Conversely, if $\hat{G}$ is uncountable, then Peter-Weyl tells us that $L^2(G)$ admits an uncountable orthonormal basis, meaning that $L^2(G)$ is not separable.

(4)$\Rightarrow$(1) If $G = (G, \rho)$ is compact metrizable, then G is separable, admitting a countable dense subset $\{x_j\}_{j \in I}$ Let $f_j = \rho(\cdot, x_j)$. Then $\{f_j\}_{j \in J}$ separates points, since if $f_j(x) = f_j(y)$ for all $j \in J$, then $\rho(x, x_j) = \rho(y, x_j)$ for all $j \in I$. For each $n \in \mathbb{N}$ exists $j_n \in I$ such that $\rho(x, x_{j_n}) \leq \frac{1}{2n}$, since $\{x_j\}_{j \in J}$ is dense in G. Thus $\rho(x, y) \leq \rho(x, x_j) + \rho(x_j, y) \leq \frac{1}{n}$. Thus $x = y$. By Stone-Weierstrass, this implies that

$$\overline{\mathrm{span}}\left\{\prod_{n=1}^{N} f_{j_n} : j_1, \ldots, j_N \in J, N \in \mathbb{N} \cup \{0\}\right\} = C(G),$$

where the empty product is the constant function 1.

(2)$\Rightarrow$(4) Let $\lambda : G \to \mathbb{U}\left(L^2(G)\right)$ be the left regular representation

$$\lambda(x) : (t \mapsto f(t)) \mapsto \left(t \mapsto f\left(x^{-1}t\right)\right).$$

Then λ is a faithful representation of G on $\mathcal{H}$, meaning λ is an embedding of G into $\mathbb{U}\left(L^2(G)\right)$. Therefore $\lambda(G) \cong G$ is a closed subgroup of $\mathbb{U}\left(L^2(G)\right)$. But $\mathbb{U}(\mathcal{H})$ is metrizable when $\mathcal{H}$ is separable, so G is therefore metrizable. $\qquad \square$

When G is compact abelian, the family $\hat{G}$ is exactly the family $\mathrm{Hom}\left(G, \mathbb{S}^1\right)$ of continuous group homomorphisms $G \to \mathbb{S}^1$, where $\mathbb{S}^1 = \{z \in \mathbb{C} : z\overline{z} = 1\}$. Then $\hat{G}$ has the structure of a locally compact abelian topological group under pointwise multiplication (Rudin, 1962, 1.2.6(d)).

The following results describes a condition under which certain sequences will be almost surely uniformly distributed. We remark that the result is a direct generalization of Theorem 4.1 from Chapter 1 of (Kuipers and Niederreiter, 2012), which proves the result (albeit in different language) for the particular case where $G = \mathbb{R}/\mathbb{Z}$, and our method of proof is essentially the same, except expressed in the language of harmonic analysis.

Theorem 5.2.9. *Let G be a compact abelian metrizable group, and let $(\Phi_n)_{n \in \mathbb{N}}$ be a sequence of distinct continuous group endomorphisms of G such that $\Phi_n - \Phi_m$ is a surjection of G onto itself for all $n \neq m$. Then for almost all $x \in G$, the sequence $(\Phi_n x)_{n=0}^{\infty}$ is uniformly distributed.*

Proof. Fix some nontrivial irreducible unitary representation $\gamma \in \hat{G}$. Set

$$S_\gamma(k, x) = \frac{1}{k} \sum_{i=0}^{k-1} \gamma(\Phi_i x).$$

Then

$$|S_\gamma(k, x)|^2 = \frac{1}{k^2} \sum_{i,j=0}^{k-1} \gamma(\Phi_i x) \overline{\gamma(\Phi_j x)}$$

$$= \frac{1}{k^2} \sum_{i,j=0}^{k-1} \gamma((\Phi_i - \Phi_j) x)$$

$$\Rightarrow \int |S_\gamma(k, x)|^2 \mathrm{d}\mu(x) = \frac{1}{k^2} \sum_{i,j=0}^{k-1} \int \gamma((\Phi_i - \Phi_j) x) \mathrm{d}\mu(x)$$

$$= \frac{1}{k}.$$

This cancellation is possible because if $i \neq j$, then $\Phi_i - \Phi_j$ is surjective, meaning that $\gamma \circ (\Phi_i - \Phi_j)$ is a nontrivial character on G. Therefore $\int \gamma((\Phi_i - \Phi_j) x) \mathrm{d}\mu(x) = 0$ for $i \neq j$, meaning only the terms of $i = j$ contribute to the sum.

This tells us that $\sum_{K=1}^{\infty} \int |S_\gamma(K^2, x)|^2 \mathrm{d}\mu(x) = \sum_{K=1}^{\infty} K^{-2} < \infty$, so by Fatou's Lemma we know that $\int \sum_{K=1}^{\infty} |S_\gamma(K^2, x)|^2 \mathrm{d}\mu(x) < \infty$. In particular, this tells us that $\sum_{K=1}^{\infty} |S_\gamma(K^2, x)|^2 < \infty$ for almost all $x \in G$, and so for almost all $x \in G$, we have $S_\gamma(K^2, x) \overset{K \to \infty}{\to} 0$.

Let $x \in G$ such that $S_\gamma(K^2, x) \overset{K \to \infty}{\to} 0$. We want to show that $S_\gamma(k, x) \overset{k \to \infty}{\to} 0$. For any $k \in \mathbb{N}$, we have

$$\lfloor \sqrt{k} \rfloor^2 \leq k \leq \left(\lfloor \sqrt{k} \rfloor + 1 \right)^2.$$

So

$$|S_\gamma(k,x)| = \left| \frac{1}{k} \sum_{i=0}^{k-1} \gamma(\Phi_i x) \right|$$

$$\leq \left| \frac{1}{k} \sum_{i=0}^{\lfloor\sqrt{k}\rfloor^2-1} \gamma(\Phi_i x) \right| + \left| \frac{1}{k} \sum_{j=\lfloor\sqrt{k}\rfloor^2}^{k-1} \gamma(\Phi_j x) \right|$$

$$\leq \left| \frac{1}{\lfloor\sqrt{k}\rfloor^2} \sum_{i=0}^{\lfloor\sqrt{k}\rfloor^2-1} \gamma(\Phi_i x) \right| + \frac{2\lfloor\sqrt{k}\rfloor}{k}$$

$$= \left| S_\gamma\left(\lfloor\sqrt{k}\rfloor^2, x\right) \right| + \frac{2\lfloor\sqrt{k}\rfloor}{k}$$

$$\leq \left| S_\gamma\left(\lfloor\sqrt{k}\rfloor^2, x\right) \right| + \frac{2}{\sqrt{k}}$$

$$\overset{k\to\infty}{\to} 0.$$

Let $E_\gamma = \left\{ x \in G : S_\gamma(k,x) \overset{k\to\infty}{\to} 0 \right\}$, and let $E = \bigcap_{\gamma\in\hat{G}\setminus\{1\}} E_\gamma$. Since $\hat{G}$ is countable, we know $\mu(E) = 1$, proving the theorem. $\qquad\square$

5.3   The Difference Property

We will be interested especially in the situation where the sequence $(\Phi_n)_{n\in\mathbb{N}}$ is generated by a sequence $(T_n)_{n\in\mathbb{N}}$ of continuous group endomorphisms. Motivated by Theorem 5.2.9, we introduce the following definition.

Definition 5.3.1. Let G be a compact abelian metrizable group, and let $(T_n)_{n=1}^\infty$ be a sequence of continuous group endomorphisms of G. Set $\Phi_n = T_n T_{n-1} \cdots T_1$ for $n \in \mathbb{N}$. We say that the sequence $(T_n)_{n=1}^\infty$ has *the Difference Property* if $\Phi_n - \Phi_m$ is surjective for all $n, m \in \mathbb{N}, n \neq m$.

It is not obvious that the Difference Property places any particular restrictions on the individual Φ_n themselves. We have a special interest in when the maps $\{T_n\}_{n=1}^\infty$ are surjective -or perhaps more interestingly, when they are *not* surjective- because a continuous group endomorphism on a compact group is measure-preserving if and only if it is surjective.

Proposition 5.3.2. *Let $T : G \to G$ be a continuous group endomorphism of a compact group G. Then T is surjective if and only if T is measure-preserving.*

Proof. ($\Rightarrow$) Suppose T is surjective. Define a Borel measure ν on G by $\nu(E) = \mu\left(T^{-1}E\right)$. Let $x \in G$, and choose $y \in G$ such that $x = Ty$ Then the measure ν satisfies

$$\nu(xE) = \mu\left(T^{-1}(xE)\right)$$
$$= \mu\left(T^{-1}((Ty)E)\right)$$
$$= \mu\left(yT^{-1}E\right)$$
$$= \mu\left(T^{-1}E\right)$$
$$= \nu(E),$$

meaning that ν is left-invariant. It also satisfies $\nu(G) = \mu\left(T^{-1}G\right) = \mu(G) = 1$, so ν is a probability measure. Therefore ν is a Haar probability measure on G, but by the uniqueness of the Haar measure, this implies that $\nu = \mu$.

($\Leftarrow$) If T is *not* surjective, then there exists $x \in G \setminus TG$. Since G is compact and Hausdorff, the map T is necessarily closed, so TG is closed in G, and a fortiori is measurable. If $\mu(TG) \neq 1$, then we know that $\mu\left(T^{-1}(TG)\right) = \mu(G) \neq \mu(TG)$, so consider the case where $\mu(TG) = 1$. We claim that $T^{-1}(xTG) = \emptyset$, since if $Ty = xTz$ for some $y, z \in G$, then

$$T\left(yz^{-1}\right) = (Ty)\left(Tz^{-1}\right)$$
$$= xTz(Tz)^{-1}$$
$$= x,$$

a contradiction. Thus $\mu\left(T^{-1}(xTG)\right) = 0 \neq 1 = \mu(TG) = \mu(xTG)$. $\qquad\square$

We note that our argument for the forward direction can be found in (Walters, 2007, §1.2). We include it here for the sake of a self-contained treatment.

The possibility that in the non-autonomous case, rich results like Theorem 5.2.9 could be achieved where a nontrivial number of the $\{T_n\}_{n=1}^{\infty}$ are not even measure-preserving intrigues us. The following results show that under certain conditions, the Difference Property imposes surjectivity on the T_n individually.

Proposition 5.3.3. *Let G be a compact abelian metrizable group, and let $(T_n)_{n=1}^\infty$ be a sequence of continuous group endomorphisms on G with the Difference Property such that $T_n T_m = T_m T_n$ for all $m, n \in \mathbb{N}$. Then each T_n is surjective.*

Proof. For each $n \in \mathbb{N}$, we have that

$$\Phi_{n+1} - \Phi_n = T_{n+1}\Phi_n - \Phi_n$$
$$= (T_{n+1} - 1)\Phi_n$$
$$= \Phi_n(T_{n+1} - 1)$$

is surjective, meaning that Φ_n is surjective. Since $\Phi_n = T_n \Phi_{n-1}$ (where we let $\Phi_0 = \mathrm{id}_G$), we can conclude that T_n is surjective. $\square$

This result applies in the autonomous setting, where $T_n = T$ for all $n \in \mathbb{N}$, but not in general. Even in the case of $\mathbb{R}^2/\mathbb{Z}^2$, endomorphisms need not commute. The next example addresses the situation of finite-dimensional tori.

Fact 5.3.4. *If G is a compact connected abelian Lie group, then $G = \mathbb{R}^h/\mathbb{Z}^h$ for some $h \in \mathbb{N}$. In general, if G is a compact abelian Lie group, then $G = G_0 \oplus B$, where G_0 is the identity component of G, and $B \cong A = G/G_0$.*

Proof. For a characterization of the compact connected Lie groups, see (Procesi, 2006, §4.2). We now argue that this implies that any compact abelian Lie group can be expressed in the way described above.

We want to prove that G is isomorphic to a direct sum of G_0 and $A = G/G_0$. The embedding $\iota_1 : G_0 \hookrightarrow G$ is just the canonical embedding, so it remains to find an embedding $\iota_2 : A \hookrightarrow G$.

Since A is finite abelian, there exist $a_1, \ldots, a_m \in A$, as well as $\ell_1, \ldots, \ell_m \geq 2$ such that

$$A = \langle a_1 \rangle_{\ell_1} \oplus \cdots \oplus \langle a_m \rangle_{\ell_m}.$$

Let $\pi : G \to G/G_0$ be the canonical projection. For each $j \in \{1, \ldots, m\}$, choose $y_j \in G$ such that $\pi(y_j) = a_j$. Then $\ell_j y_j \in \ker \pi = G_0$. Since G_0 is divisible, there exists $y_j' \in G_0$ such that $\ell_j y_j' = \ell_j y_j$. Set $b_j = y_j - y_j'$, so that $\ell_j b_j = 0$. However, because $\pi(k b_j) = k\pi(a_j)$ for $k = 1, \ldots, \ell_j - 1$, we know that b_j has order ℓ_j in G.

Let $B = \oplus_{j=1}^m \langle b_j \rangle_{\ell_1}$ denote the subgroup of G generated by $\{b_1, \ldots, b_m\}$. Then $\pi : B \to A$ is an isomorphism. Let $\iota_2 = \pi^{-1} : A \to B \leq G$.

We claim now that $\iota_1 : G_0 \hookrightarrow G, \iota_2 : A \hookrightarrow G$ generate G as a direct sum of G_0 and A. First, we observe that $G_0 \cap B = \{0\}$, since if $x \in G_0 \cap B$, then $\pi(x) = 0$, because $G_0 = \ker \pi$, but also $x = \pi^{-1}(0) = 0$, since $\pi|_B : B \to A$ is an isomorphism. Therefore, the sum $G_0 + B$ is direct.

Now, we have to show that $G = G_0 + B$. If $x \in G$, then choose $b \in B$ such that $\pi(b) = \pi(x)$. Then $x - b \in \ker \pi = G_0$. Thus $x = (x - b) + b$.

Therefore $G = G_0 \oplus B \cong G_0 \oplus A$. $\qquad\square$

Proposition 5.3.5. *Let G be a compact connected abelian Lie group, and let $(T_n)_{n=1}^\infty$ be a sequence of continuous group endomorphisms on G with the Difference Property. Then each T_n is surjective.*

Proof. Let $G = \mathbb{R}^h / \mathbb{Z}^h$. We can express T_n as an $h \times h$ integer matrix. The maps $\Phi_n - \Phi_m$ are surjective if and only if $\det(\Phi_n - \Phi_m) \neq 0$. We can conclude that each T_n is invertible because

$$\Phi_{n+1} - \Phi_n = (T_{n+1} - I)\Phi_n$$

$$= (T_{n+1} - I)T_n T_{n-1} \cdots T_1$$

$$\Rightarrow \det(\Phi_{n+1} - \Phi_n) = \det(T_{n+1} - I) \det(T_n) \det(T_{n-1}) \cdots \det(T_1)$$

$$\neq 0.$$

This implies in particular that $\det(T_n) \neq 0$, meaning that T_n is surjective. $\qquad\square$

We can, however, provide *negative* results on when a sequence with the Difference Property might exist.

Lemma 5.3.6. *Let $G = G_0 \oplus A$ be a compact abelian group, where G_0 is a compact connected abelian group and A is a finite abelian group. Let $\psi : G \to A$ be a continuous group homomorphism. Then*

$$\psi(G) = \psi(0 \oplus A) = \{\psi(0, a) : a \in A\}.$$

Proof. Since G_0 is connected, we know that $\psi|_{G_0} : G_0 \to A$ is the zero map. Clearly $\psi(0 \oplus A) \subseteq \psi(G)$. Now suppose that $a \in \psi(G)$. Then there exist $x_0 \in G_0, a_0 \in A$ such that $\psi(x_0, a_0) = a$. But $\psi(0, a_0) = \psi(x_0, a_0) - \psi(x_0, 0) = \psi(x_0, a_0)$. Thus $a \in \psi(0 \oplus A)$. $\qquad\square$

Proposition 5.3.7. *Let $G = G_0 \oplus A$, where G_0 is a compact connected abelian group, and A a finite group with more than one element. Then there does not exist a sequence $(T_n)_{n=1}^\infty$ on G with the Difference Property. In particular, there does not exist a sequence with the Difference Property on any compact abelian Lie group that is not connected.*

Proof. Assume for contradiction that $(T_n)_{n=1}^\infty$ has the Difference Property. Let $\Delta_{n,m} = \Phi_n - \Phi_m$. Let $\iota : A \hookrightarrow G_0 \oplus A$ be the canonical embedding. Then by the previous lemma, the maps $\Delta_{n,m} \circ \iota = (\Phi_n \circ \iota) - (\Phi_m \circ \iota)$ are surjective. But this is a contradiction, since $n \mapsto \Phi_n \circ \iota$ is a mapping from $\mathbb{N}$ to the finite set A^A, which cannot be injective. Therefore there exist $n, m \in \mathbb{N}, n \neq m$ such that $\Phi_n \circ \iota = \Phi_m \circ \iota$, meaning that $\Delta_{n,m}(G) = \Delta_{n,m} \circ \iota(A) = \{0\} \neq A$, a contradiction.

The special case of compact abelian Lie groups comes from Fact 5.3.4. □

5.4   A random non-autonomous spatial-temporal differentiation problem

Lemma 5.4.1. *Let $G = (G, \rho)$ be a compact metrizable group with metric ρ. Let $(T_n)_{n=1}^\infty$ be a family of Lipschitz-continuous maps $T_n : G \to G$, where T_n is L_n-Lipschitz. Let $\Phi_n = T_n T_{n-1} \cdots T_1, \Phi_0 = \mathrm{id}_G$. Then there exists a sequence $(\eta_k)_{k=1}^\infty$ of positive numbers $\eta_k > 0, \eta_k \overset{k \to \infty}{\to} 0$ such that if $(r_k)_{k=1}^\infty$ is a sequence of positive numbers $0 < r_k \leq \eta_k$, and $B_k(x) = B(x, r_k)$, then for every $f \in C(G)$, and every sequence $(x_n)_{n=1}^\infty \in G^{\mathbb{N}}$, we have*

$$\left| \alpha_{B_k(x_k)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) - \frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f(x_k) \right| \overset{k \to \infty}{\to} 0.$$

Proof. Let $\tilde{L}_k = \max\{1, L_1, L_2, \ldots, L_{k-1}\}$. Set

$$\eta_k = \tilde{L}_k^{-(k-1)} k^{-1}.$$

Let $(r_k)_{k=1}^\infty$ be a sequence such that $0 < r_k \leq \eta_k$. Then each $T_1, \ldots, T_{k-1}$ is $\tilde{L}_k$-Lipschitz, so if $y \in B_k(x)$, i.e. if $\rho(x,y) < r_k$, and if $i \in [0, k-1]$, then

$$
\begin{aligned}
\rho(\Phi_i x, \Phi_i y) &= \rho(T_i T_{i-1} \cdots T_1 x, T_i T_{i-1} \cdots T_1 y) \\
&\leq L_i L_{i-1} \cdots L_1 \rho(x,y) \\
&\leq \left(\tilde{L}_k\right)^i \rho(x,y) \\
&\leq \left(\tilde{L}_k\right)^{i-(k-1)} k^{-1} \\
&\leq k^{-1}.
\end{aligned}
$$

Now let $f \in C(G)$, and fix $\epsilon > 0$. By uniform continuity of f, there exists $K \in \mathbb{N}$ such that $\rho(z,w) \leq \frac{1}{K} \Rightarrow |f(z) - f(w)| \leq \epsilon$. Then if $k \geq K$, then

$$
\begin{aligned}
\left| \alpha_{B_k(x_k)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) - \frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f(x_k) \right| &= \frac{1}{k} \left| \sum_{i=0}^{k-1} \Phi_i f(x_k) - \frac{1}{\mu(B_k(x_k))} \int_{B_k(x_k)} \Phi_i f \mathrm{d}\mu \right| \\
&\leq \frac{1}{k} \sum_{i=0}^{k-1} \left| \Phi_i f(x_k) - \frac{1}{\mu(B_k(x_k))} \int_{B_k(x_k)} \Phi_i f \mathrm{d}\mu \right| \\
&= \frac{1}{k} \sum_{i=0}^{k-1} \left| \frac{1}{\mu(B_k(x_k))} \int_{B_k(x_k)} (\Phi_i f(x_k) - \Phi_i f) \, \mathrm{d}\mu \right| \\
&\leq \frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(B_k(x_k))} \int_{B_k(x_k)} |(\Phi_i f(x_k) - \Phi_i f)| \, \mathrm{d}\mu \\
&\leq \frac{1}{k} \sum_{i=0}^{k-1} \frac{1}{\mu(B_k(x_k))} \int_{B_k(x_k)} \epsilon \mathrm{d}\mu \\
&= \epsilon.
\end{aligned}
$$

Therefore

$$
\left| \alpha_{B_k(x_k)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) - \frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f(x_k) \right| \overset{k \to \infty}{\to} 0.
$$

$\square$

Our assumption that the $\{T_n\}_{n=1}^\infty$ are Lipschitz is not overly restrictive. In particular, if G has the structure of a Lie group with a Riemannian metric, and the maps T_n are C^1, then the maps T_n are automatically Lipschitz in that Riemannian metric.

Theorem 5.4.2. *Let $G = (G, \rho)$ be a compact abelian metrizable group with metric ρ. Let $(T_n)_{n=1}^\infty$ be a family of Lipschitz-continuous group homomorphisms with the Difference Property, where T_n is L_n-Lipschitz. Then there exists a Borel subset E of full measure and a sequence $(\eta_k)_{k=1}^\infty$ of positive numbers $\eta_k > 0, \eta_k \overset{k \to \infty}{\to} 0$ such that if $(r_k)_{k=1}^\infty$ is a sequence of positive numbers $0 < r_k \leq \eta_k$, and $B_k(x) = B(x, r_k)$, then for all $x \in E$, we have*

$$\alpha_{B_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) \overset{k \to \infty}{\to} \int f \mathrm{d}\mu \qquad\qquad (\forall f \in C(G)).$$

Proof. Let $E = \{x \in G : (\Phi_n x)_{n=0}^\infty$ is uniformly distributed in $G\}$. By Theorem 5.2.9, this set is of full measure. Now let $(\eta_k)_{k=1}^\infty$ be as in Lemma 5.4.1. Then if $0 < r_k \leq \eta_k$, and if $x \in E$, then for every $f \in C(G)$ we have

$$\left| \alpha_{B_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) - \int f \mathrm{d}\mu \right|$$

$$\leq \left| \alpha_{B_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) - \frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f(x) \right| + \left| \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f(x) \right) - \int f \mathrm{d}\mu \right|$$

$$\overset{k \to \infty}{\to} 0,$$

where the first term goes to 0 by Lemma 5.4.1, and the second term goes to 0 by the fact that $(\Phi_n x)_{n=0}^\infty$ is uniformly distributed in G. $\qquad\square$

5.5   Further probabilistic results about uniformly distributed sequences

We now consider the distribution properties of randomly chosen sequences $(x_n)_{n=0}^\infty \in G^{\mathbb{N}_0} = \prod_{n=0}^\infty G$ in G.

Lemma 5.5.1. *Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of separable metrizable topological spaces. Then*

$$\mathrm{Bo} \left(\prod_{n \in \mathbb{N}} X_n \right) = \bigotimes_{n \in \mathbb{N}} \mathrm{Bo}(X_n).$$

Proof. (Kallenberg, 2021, Lemma 1.2) $\qquad\square$

Definition 5.5.2. Let X be a nonempty set. We call $\mathcal{A} \subseteq \mathcal{P}(X)$ a *semi-algebra* if

(a) $\emptyset \in \mathcal{A}$

(b) If $A, B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$.

(c) For every $A \in \mathcal{A}$, the set $X \setminus A$ can be written as a disjoint union of finitely many elements of $\mathcal{A}$.

Lemma 5.5.3. *Let $(X_n, \mathcal{B}_n)_{n \in \mathbb{N}}$ be a sequence of measurable spaces, and let*

$$\mathcal{A} = \left\{ B_1 \times \cdots \times B_N \times \prod_{n=N+1}^{\infty} X_n : N \in \mathbb{N}, B_1 \in \mathcal{B}_1, \ldots, B_N \in \mathcal{B}_N \right\}.$$

Then $\mathcal{A}$ is a semi-algebra that generates $\bigotimes_{n \in \mathbb{N}} \mathcal{B}_n$.

Proof. For $n \in \mathbb{N}$, let $\pi_i : \prod_{n \in \mathbb{N}} X_n \twoheadrightarrow X_i$ be the map $\pi_i : (x_n)_{n \in \mathbb{N}} \mapsto x_i$. Then $\mathcal{A}$ can be written as

$$\mathcal{A} = \left\{ \bigcap_{n=1}^{N} \pi_n^{-1}(B_n) : N \in \mathbb{N}, B_1 \in \mathcal{B}_1, \ldots, B_N \in \mathcal{B}_N \right\}.$$

Written this way, it is clear that $\emptyset \in \mathcal{A}$ and that $\mathcal{A}$ is closed under finite intersections. Finally, we will prove that the complement of every set in $\mathcal{A}$ can be expressed as the disjoint union of finitely many elements of $\mathcal{A}$. Let $A = \bigcap_{n=1}^{N} \pi_n^{-1}(B_n)$, where $B_1 \in \mathcal{B}_1, \ldots, B_N \in \mathcal{B}_N$. Then

$$
\begin{aligned}
A^{\mathsf{C}} &= \left(\bigcap_{n=1}^{N} \pi_n^{-1}(B_n) \right)^{\mathsf{C}} \\
&= \bigcup_{n=1}^{N} \pi_n^{-1}\left(B_n^{\mathsf{C}} \right) \\
&= \bigsqcup_{I \in \mathcal{P}(\{1,2,\ldots,N\})} \left[\left(\bigcap_{i \in I} \pi_i^{-1}(B_i)^{\mathsf{C}} \right) \cap \left(\bigcap_{j \in \{1,\ldots,N\} \setminus I} \left(\pi_j^{-1}(B)^{\mathsf{C}} \right)^{\mathsf{C}} \right) \right] \\
&= \bigsqcup_{I \in \mathcal{P}(\{1,2,\ldots,N\})} \left[\left(\bigcap_{i \in I} \pi_i^{-1}\left(B_i^{\mathsf{C}} \right) \right) \cap \left(\bigcap_{j \in \{1,\ldots,N\} \setminus I} \pi_i^{-1}(B_j) \right) \right].
\end{aligned}
$$

Therefore, we have written A^{C} as a disjoint union of elements of $\mathcal{A}$.

Finally, to justify our claim that $\mathcal{A}$ generates $\bigotimes_{n \in \mathbb{N}} \mathcal{B}_n$ as a σ-algebra, we note that $\bigotimes_{n \in \mathbb{N}} \mathcal{B}_n$ is generated as a σ-algebra by $\left\{ \pi_n^{-1}(B_n) : n \in \mathbb{N}, B_n \in \mathcal{B}_n \right\}$, and

$$\left\{ \pi_n^{-1}(B_n) : n \in \mathbb{N}, B_n \in \mathcal{B}_n \right\} \subseteq \mathcal{A} \subseteq \bigotimes_{n \in \mathbb{N}} \mathcal{B}_n.$$

173

□

We go to the trouble of proving Lemma 5.5.3 because when checking whether a map between probability spaces is measure-preserving, it suffices to see how the map behaves on a generating semi-algebra, as in the following result.

Lemma 5.5.4. *Let* $(X_1, \mathcal{B}_1, \mu_1), (X_2, \mathcal{B}_2, \mu_2)$ *be probability spaces, and let* $T : X_1 \to X_2$ *be a map. Let* $\mathcal{A}_2 \subseteq \mathcal{B}_2$ *be a semi-algebra which generates the* σ-*algebra* $\mathcal{B}_2$. *Then* T *is measurable and measure-preserving iff* $T^{-1}A \in \mathcal{B}_1$ *for all* $A \in \mathcal{A}_2$, *and* $\mu_1\left(T^{-1}A\right) = \mu_2(A)$.

Proof. (Dajani and Dirksin, 2008, Theorem 1.2.2) □

With all of this out of the way, we are ready to demonstrate a characterization of the Haar measure of a countable product of compact metrizable groups.

Theorem 5.5.5. *Let* $(H_n)_{n \in \mathbb{N}}$ *be a sequence of compact metrizable groups, and let* ν_n *be the left-invariant (resp. right-invariant) Haar probability measure of* H_n. *Then* $\mu = \prod_{n \in \mathbb{N}} \nu_n$ *is the left-invariant (resp. right-invariant) Haar probability measure on* $G = \prod_{n \in \mathbb{N}} H_n$.

Proof. We will demonstrate the claim for left-invariant Haar measures, since the proof for the right-invariant claim is essentially identical.

We will show that μ is a G-invariant Borel probability measure on G, and then conclude from the uniqueness of the Haar measure that μ must be *the* Haar probability measure on G. By Lemma 5.5.1, we know that $\mathrm{Bo}(G) = \bigotimes_{n \in \mathbb{N}} \mathrm{Bo}(H_n)$. Set

$$\mathcal{A} = \left\{ B_1 \times \cdots \times B_N \times \prod_{n=N+1}^{\infty} H_n : N \in \mathbb{N}, B_1 \in \mathrm{Bo}(H_1), \ldots, B_N \in \mathrm{Bo}(H_N) \right\}.$$

Then by Lemma 5.5.3, this $\mathcal{A}$ is a generating semi-algebra for $\bigotimes_{n \in \mathbb{N}} \mathrm{Bo}(H_n)$.

Now, fix $g = (h_n)_{n \in \mathbb{N}} \in G$. We want to prove that left multiplication on G by g is a μ-preserving transformation. By Lemma 5.5.4, it will suffice to prove that $\mu(gA) = \mu(A)$ for all $A \in \mathcal{A}$. So fix sets

$B_1 \in \mathrm{Bo}(H_1), \ldots, B_N \in \mathrm{Bo}(H_N), N \in \mathbb{N}$. Then

$$\mu\left(g\left(B_1 \times \cdots \times B_N \times \prod_{n=N+1}^{\infty} H_n\right)\right) = \mu\left(h_1 B_1 \times \cdots \times h_N B_N \times \prod_{n=N+1}^{\infty} h_n H_n\right)$$

$$= \mu\left(h_1 B_1 \times \cdots \times h_N B_N \times \prod_{n=N+1}^{\infty} H_n\right)$$

$$= \nu_1(h_1 B_1) \cdots \nu_N(h_N B_N)$$

$$= \nu_1(B_1) \cdots \nu_N(B_N)$$

$$= \mu\left(B_1 \times \cdots \times B_N \times \prod_{n=N+1}^{\infty} H_n\right).$$

We have thus established that μ is G-invariant on $\mathcal{A}$, and so we can infer that μ is G-invariant for all of $\mathrm{Bo}(G)$. $\qquad\square$

For the remainder of this section, let G be a compact abelian metrizable group. Let $S : G^{\mathbb{N}_0} \twoheadrightarrow G^{\mathbb{N}_0}$ be the left shift

$$S(g_n)_{n=0}^{\infty} = (g_{n+1})_{n=0}^{\infty}.$$

Then S is a continuous surjective group endomorphism of $G^{\mathbb{N}_0}$, and given a continuous group homomorphism $T : G \to G$, let $\widehat{T} : G^{\mathbb{N}_0} \to G^{\mathbb{N}_0}$ be the map

$$\widehat{T} : (g_n)_{n=0}^{\infty} \mapsto (T g_n)_{n=0}^{\infty}.$$

We can observe that S and $\widehat{T}$ commute, since

$$S\widehat{T}(g_n)_{n=0}^{\infty} = S\,(Tg_n)_{n=0}^{\infty}$$

$$= (Tg_{n+1})_{n=0}^{\infty}$$

$$= \widehat{T}(g_{n+1})_{n=0}^{\infty}$$

$$= \widehat{T}S(g_n)_{n=0}^{\infty}.$$

Lemma 5.5.6. *Let $T : G \twoheadrightarrow G$ be a continuous surjective group endomorphism of G, and fix $\ell \in \mathbb{N}$. Then* $\mathrm{id}_{G^{\mathbb{N}_0}} - S^\ell \widehat{T}$ *is also surjective.*

Proof. Fix $g = (g_n)_{n=0}^\infty \in G^{\mathbb{N}_0}$, and construct a sequence $g' = (g'_n)_{n=0}^\infty \in G^{\mathbb{N}_0}$ recursively as follows. First, set $g'_n = g_n$ for $n = 0, \ldots, \ell - 1$. Then for $N > \ell - 1$, assuming that $g'_0, g'_1, g'_2, \ldots, g'_N \in G$ have been chosen such that

$$g'_n - Tg'_{n+\ell} = g_n \qquad\qquad (\text{for } n = 0, 1, 2, \ldots, N - \ell),$$

choose $g'_{N+1} \in G$ such that

$$Tg'_{N+1} = g_{N+1-\ell} - g'_{N+1-\ell},$$

which exists because T is surjective. Then

$$g'_{N+1-\ell} - Tg'_{N+1} = g_{N+1-\ell}.$$

Continuing this process gives us a sequence $g' = (g'_n)_{n=0}^\infty$ such that

$$\left(\mathrm{id}_{G^{\mathbb{N}_0}} - S^\ell \widehat{T}\right) g' = g.$$

$\square$

Lemma 5.5.7. *Let $(T_n)_{n=1}^\infty$ be a sequence of surjective group homomorphisms $T_n : G \twoheadrightarrow G$, and let $(\ell_n)_{n=1}^\infty$ be a sequence of natural numbers. Then the sequence $\left(S^{\ell_n} \widehat{T}_n\right)_{n=1}^\infty$ has the Difference Property on $G^{\mathbb{N}_0}$.*

Proof. Set $\Lambda_n = \ell_1 + \cdots + \ell_n$. Let $m, n \in \mathbb{N}, m < n$. Then

$$\left(\left(S^{\ell_m}\widehat{T}_m\right)\cdots\left(S^{\ell_1}\widehat{T}_1\right)\right) - \left(\left(S^{\ell_n}\widehat{T}_n\right)\cdots\left(S^{\ell_1}\widehat{T}_1\right)\right)$$

$$= \left(S^{\Lambda_m}\widehat{T}_m\cdots\widehat{T}_1\right) - \left(S^{\Lambda_n}\widehat{T}_n\cdots\widehat{T}_1\right)$$

$$= \left(\widehat{T}_m\cdots\widehat{T}_1 S^{\Lambda_m}\right) - \left(\widehat{T}_n\cdots\widehat{T}_1 S^{\Lambda_n}\right)$$

$$= \left(\mathrm{id}_{G^{\mathbb{N}_0}} - \widehat{T}_n\widehat{T}_{n-1}\cdots\widehat{T}_{m+1}S^{\Lambda_n-\Lambda_m}\right) S^{\Lambda_m}\widehat{\Phi}_m$$

$$= \left(\mathrm{id}_{G^{\mathbb{N}_0}} - \widehat{\tau(n,m)}S^{\Lambda_n-\Lambda_m}\right) S^{\Lambda_m}\widehat{\Phi}_m$$

$$= \left(\mathrm{id}_{G^{\mathbb{N}_0}} - S^{\Lambda_n-\Lambda_m}\widehat{\tau(n,m)}\right) S^{\Lambda_m}\widehat{\Phi}_m$$

where $\tau(n,m) = T_n T_{n-1} \cdots T_{m+1}$. By Lemma 5.5.6, it follows that $\mathrm{id}_{G^{\mathbb{N}_0}} - \widehat{\tau(n,m)} S^{\Lambda_n - \Lambda_m}$ is surjective. Therefore

$$\left(\left(S^{\ell_m}\widehat{T}_m\right)\cdots\left(S^{\ell_1}\widehat{T}_1\right)\right) - \left(\left(S^{\ell_n}\widehat{T}_n\right)\cdots\left(S^{\ell_1}\widehat{T}_1\right)\right) = \left(\mathrm{id}_{G^{\mathbb{N}_0}} - S^{\Lambda_n - \Lambda_m}\widehat{\tau(n,m)}\right)S^{\Lambda_m}$$

is surjective, since it is a composition of surjections. □

Corollary 5.5.8. *Let $(T_n)_{n=1}^{\infty}$ be a sequence of surjective group homomorphisms $T_n : G \twoheadrightarrow G$, and let $(\ell_n)_{n=1}^{\infty}$ be a sequence of natural numbers. Set $\Lambda_n = \ell_1 + \cdots + \ell_n$. Then for almost every $g \in G$, the sequence $\left(S^{\Lambda_n}\widehat{T}_n \cdots \widehat{T}_1 g\right)_{n=0}^{\infty}$ is uniformly distributed in $G^{\mathbb{N}_0}$.*

Proof. Apply Lemma 5.5.7 and Theorem 5.2.9. □

Theorem 5.5.9. *Let $(T_n)_{n=1}^{\infty}$ be a sequence of surjective group homomorphisms $T_n : G \twoheadrightarrow G$, and let $(\ell_n)_{n=1}^{\infty}$ be a sequence of natural numbers. Set $\Lambda_n = \ell_1 + \cdots + \ell_n$. Then for almost every sequence $(g_n)_{n=0}^{\infty} \in G^{\mathbb{N}_0}$, the sequence $(T_n \cdots T_1 g_{\Lambda_n})_{n=0}^{\infty}$ is uniformly distributed in G.*

Proof. Let $\pi : G^{\mathbb{N}_0} \twoheadrightarrow G$ be the projection onto the first term $\pi : (g_n)_{n=0}^{\infty} \mapsto g_0$. By Corollary 5.5.8, for almost every $(g_n)_{n=0}^{\infty} \in G^{\mathbb{N}_0}$, the sequence $\left(S^{\Lambda_n}\widehat{T}_n \cdots \widehat{T}_1 g\right)_{n=0}^{\infty}$ is uniformly distributed in $G^{\mathbb{N}_0}$.

Now let $(g_n)_{n=0}^{\infty} \in G^{\mathbb{N}_0}$ be such that the sequence $\left(S^{\Lambda_n}\widehat{T}_n \cdots \widehat{T}_1 g\right)_{n=0}^{\infty}$ is uniformly distributed in $G^{\mathbb{N}_0}$. Then by Lemma 5.2.6, the sequence $\left(\pi\left(S^{\Lambda_n}\widehat{T}_n \cdots \widehat{T}_1 g\right)\right)_{n=0}^{\infty}$ is uniformly distributed in G. But

$$\pi\left(S^{\Lambda_n}\widehat{T}_n \cdots \widehat{T}_1 g\right) = T_n \cdots T_1 g_{\Lambda_n}.$$

□

5.6 Topologically generic behaviors of random spatial-temporal differentiation problems

We can interpret Theorem 5.2.9 as saying that if $(T_n)_{n=1}^{\infty}$ is a sequence of continuous group endomorphisms of G with the Difference Property, and $\Phi_n = T_n T_{n-1} \cdots T_1$, then the property of $x \in G$ that $(\Phi_n x)_{n=0}^{\infty}$ is uniformly distributed is "probabilistically generic", in the sense that the set of such x has full measure. In light of Theorem 5.4.2, we can infer that if $B_k(x)$ is a sequence of balls around x with radii

going to 0 sufficiently fast, then it is probabilistically generic that

$$\alpha_{B_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) \xrightarrow{k\to\infty} \int f \mathrm{d}\mu \qquad\qquad (\forall f \in C(G)).$$

In other words, we can see Theorem 5.4.2 as a statement about a probabilistically generic spatial-temporal differentiation of a certain kind.

However, if we try to look at *topologically* generic behaviors, the story changes. Instead, in a sense that we will make precise momentarily, the topologically generic behavior is that the sequence $\left(\alpha_{B_k(x)} \left(\frac{1}{k} \sum_{i=0}^{k-1} \Phi_i f \right) \right)_{k=1}^{\infty}$ is divergent for some $f \in C(G)$.

Definition 5.6.1. Let X be a compact metrizable space, and $S \subseteq X$ a subset. We say that $A \subseteq X$ is *nowhere dense* if for every nonempty open $\mathcal{O} \subseteq X$, there exists a nonempty open subset $W \subseteq \mathcal{O}$ such that $W \cap A = \emptyset$. A subset $A \subseteq X$ is called *meager* if there exists a sequence $(A_n)_{n\in\mathbb{N}}$ of nowhere dense subsets of X such that $A \subseteq \bigcup_{n\in\mathbb{N}} A_n$. We call a subset $B \subseteq X$ *comeager* if $X \setminus B$ is meager. A comeager set is sometimes called *Baire generic*.

Our goal here is to show that the behavior described in Theorem 5.2.9 is -from this topological perspective- exceptional in the sense that the set of such $x \in G$ is meager.

Theorem 5.6.2. *Let $G = (G, \rho)$ be a compact abelian metrizable group with infinitely many elements and metric ρ. Let $(T_n)_{n=1}^{\infty}$ be a sequence of continuous, surjective group endomorphisms $T_n : G \to G$. Let $\Phi_n = T_n T_{n-1} \cdots T_1, \Phi_0 = \mathrm{id}_G$. Suppose that $\bigcup_{m=1}^{\infty} \ker \Phi_m$ is dense in G. Then the set of $x \in G$ such that $(\Phi_n x)_{n=0}^{\infty}$ is uniformly distributed is meager.*

Before we can prove Theorem 5.6.2, we need to prove a few technical lemmas.

Lemma 5.6.3. *Let $G = (G, \rho)$ be a compact metrizable group with infinitely many elements and $\epsilon > 0$. Then there exists $\delta > 0$ such that $\mu(B(0, \delta)) < \epsilon$.*

Proof. Consider the sequence $(B(0, 1/n))_{n=1}^{\infty}$. Then $B(0, 1) \supseteq B(0, 1/2) \supseteq B(0, 1/3) \supseteq \cdots$, and $\bigcap_{n=1}^{\infty} B(0, 1/n) = \{0\}$. But $\mu(\{0\}) = 0$, so by continuity of measure, it follows that $\lim_{n\to\infty} \mu(B(0, 1/n)) = 0$. Thus in particular there exists $N \in \mathbb{N}$ such that $\mu(B(0, 1/N)) < \epsilon$. Let $\delta = 1/N$. $\qquad\square$

Lemma 5.6.4. *Let $G = (G, \rho)$ be a compact abelian metrizable group with infinitely many elements and metric ρ. Let $(T_n)_{n=1}^{\infty}$ be a sequence of continuous, surjective group endomorphisms $T_n : G \to G$. Let*

$\Phi_n = T_n T_{n-1} \cdots T_1, \Phi_0 = \mathrm{id}_G$. *Suppose that* $\bigcup_{m=1}^{\infty} \ker \Phi_m$ *is dense in* G. *Let* $\epsilon, \delta > 0, N_1 \in \mathbb{N}$, *and let* $\mathcal{O} \subseteq G$ *be a nonempty open set. Then there exists a nonempty open set* $W \subseteq \mathcal{O}$ *and* $L \geq N_1$ *such that if* $x \in W$, *then*

$$\frac{\#\{j \leq L - 1 : \Phi_j x \in B(0, \delta/2)\}}{L} \geq 1 - \epsilon.$$

Proof. Choose $m \in \mathbb{N}, a \in \ker \Phi_m$ such that $a \in \mathcal{O}$. Choose $N_2 \in \mathbb{N}$ such that

$$\frac{m}{m + N_2} < \epsilon.$$

Set $L_0 = \max\{N_1, N_2\}$. Let $U \subseteq G$ be the open neighborhood of 0 given by

$$U = \bigcap_{\ell=0}^{L_0 - 1} \tau(m + \ell, m)^{-1} B(0, \delta/2).$$

Finally set

$$W = \left(\Phi_m^{-1} U\right) \cap \mathcal{O}.$$

Then $a \in W$, and $\Phi_{m+\ell} W \subseteq B(0, \delta/2)$ for all $\ell \in \{0, 1, 2, \ldots, L_0 - 1\}$. Let $L = L_0 + m$. Then

$$\frac{\#\{j \leq L - 1 : \Phi_j x \in B(0, \delta/2)\}}{L_0 + m} \geq \frac{L_0}{L_0 + m} = 1 - \frac{m}{m + L_0} \geq 1 - \frac{m}{m + N_2} \geq 1 - \epsilon.$$

$\square$

Proof of Theorem 5.6.2. Let $f : G \to [0, 1]$ be the continuous function

$$f(x) = \begin{cases} 1 & \rho(x, 0) \leq \frac{\delta}{2}, \\ 2 - \frac{2}{\delta}\rho(x, 0) & \frac{\delta}{2} \leq \rho(x, 0) \leq \delta, \\ 0 & \rho(x, 0) \geq \delta, \end{cases}$$

where $\delta > 0$ is chosen such that $\mu(B(0, \delta)) < \frac{1}{2}$, which exists by Lemma 5.6.3. Then $\chi_{B(0, \delta/2)} \leq f \leq \chi_{B(0, \delta)}$, so

$$\mu(B(0, \delta/2)) \leq \int f \mathrm{d}\mu \leq \mu(B(0, \delta)) < \frac{1}{2}.$$

For $K \in \mathbb{N}$, let A_K be the set

$$A_K = \left\{ x \in G : \frac{1}{k} \sum_{i=0}^{k-1} f\left(\Phi_i x\right) < \frac{2}{3} \text{ for all } k \geq K \right\}.$$

We claim the set A_K is nowhere dense. But the set of $x \in G$ such that $(\Phi_n x)_{n=0}^{\infty}$ is uniformly distributed is contained in $\bigcup_{K=1}^{\infty} A_K$, so if we can show that A_K is nowhere dense for all $K \in \mathbb{N}$, then the theorem will be proven. Now fix $K \in \mathbb{N}$.

Let $\mathcal{O} \subseteq G$ be a nonempty open subset of G. By Lemma 5.6.4, there exists a nonempty open subset $W \subseteq \mathcal{O}$ such that

$$\frac{\#\left\{j \leq L-1 : \Phi_j x \in B(0, \delta/2)\right\}}{L} \geq \frac{2}{3}$$

for $x \in W$, where $L \geq K$. So

$$\frac{1}{k} \sum_{i=0}^{k-1} f\left(\Phi_i x\right) \geq \frac{1}{k} \sum_{i=0}^{k-1} \chi_{B(0,\delta/2)}\left(\Phi_i x\right) \geq \frac{2}{3}.$$

Therefore $W \subseteq \mathcal{O} \setminus A_K$. $\qquad \square$

Our proof of Theorem 5.6.2 is based off of (Mance, 2010, Theorem 8.3.1). That theorem can be interpreted as a special case of Theorem 5.6.2 in the case where $G = \mathbb{R}/\mathbb{Z}$, though it is stated there in the language of normality with respect to a Cantor series.

However, this result can be strengthened under some mild additional assumptions. Theorem 5.6.2 states that the family of $x \in G$ for which $\frac{1}{k} \sum_{i=0}^{k-1} f\left(\Phi_i x\right) \overset{k \to \infty}{\to} \int f \mathrm{d}\mu$ for all $f \in C(G)$ is meager. However, it can by shown that under some additional assumptions, there exists $f \in C(G)$ such that the family of $x \in G$ for which $\lim_{k \to \infty} \frac{1}{k} \sum_{i=0}^{k-1} f\left(\Phi_i x\right)$ exists is meager.

Lemma 5.6.5. *Let $f \in C(G)$, and let $(x_n)_{n=0}^{\infty}$ be a sequence in G such that*

$$\frac{1}{k} \sum_{i=0}^{k-1} f(x_i) \overset{k \to \infty}{\to} \lambda$$

for some $\lambda \in \mathbb{C}$. Let $I \subseteq \mathbb{N}_0$ be a subset of density 0, i.e. such that

$$\frac{1}{k} \sum_{i=0}^{k-1} \chi_I(i) \overset{k \to \infty}{\to} 0.$$

Let $(y_n)_{n=0}^{\infty}$ be a sequence in G such that $\{n \in \mathbb{N}_0 : x_n \neq y_n\} \subseteq I$. Then

$$\frac{1}{k} \sum_{i=0}^{k-1} f(y_i) \overset{k \to \infty}{\to} \lambda.$$

In particular, if $(x_n)_{n=0}^{\infty}$ is uniformly distributed, then $(y_n)_{n=0}^{\infty}$ is uniformly distributed.

Proof. First, fix $f \in C(G)$, and suppose that $\frac{1}{k} \sum_{i=0}^{k-1} f(x_i) \overset{k \to \infty}{\to} \lambda$. Then

$$\frac{1}{k} \sum_{i=0}^{k-1} f(y_i) = \left(\frac{1}{k} \sum_{i=0}^{k-1} f(x_i) \right) + \left(\frac{1}{k} \sum_{i=0}^{k-1} \chi_I(i)(f(y_i) - f(x_i)) \right)$$

$$\Rightarrow \left| \left(\frac{1}{k} \sum_{i=0}^{k-1} f(y_i) \right) - \left(\frac{1}{k} \sum_{i=0}^{k-1} f(x_i) \right) \right| \leq \frac{1}{k} \sum_{i=0}^{k-1} \chi_I(i) \, |f(x_i) - f(y_i)|$$

$$\leq \frac{1}{k} \sum_{i=0}^{k-1} \chi_I(i) \, (2 \, \|f\|)$$

$$\overset{k \to \infty}{\to} 0.$$

Now, suppose that $(x_n)_{n=0}^{\infty}$ is uniformly distributed. Then the first part of this lemma tells us that for every $g \in C(G)$, we have that $\lim_{k \to \infty} \frac{1}{k} \sum_{i=0}^{k-1} g(y_i)$ exists and is equal to $\lim_{k \to \infty} \frac{1}{k} \sum_{i=0}^{k-1} g(x_i) = \int g \mathrm{d}\mu$. $\qquad \square$

Lemma 5.6.6. *Let G be a compact abelian metizable group. If the set*

$$F = \{x \in G : (\Phi_n x)_{n=0}^{\infty} \text{ is uniformly distributed}\}$$

is nonempty, and $\bigcup_{m=1}^{\infty} \ker \Phi_m$ is dense in G, then F is dense in G.

Proof. Let $U \subseteq G$ be a nonempty open subset, and let $a_0 \in F$. Then there exists $m \in \mathbb{N}$ and $p \in \bigcup_{m=1}^{\infty} \ker \Phi_m$ such that $y \in U - p$, so $a_0 + p \in U$. Then $\Phi_n a_0 = \Phi_n(a_0 + p)$ for $n \geq m$, so it follows from Lemma 5.6.5 that $(\Phi_n(a_0 + p))_{n=0}^{\infty}$ is also uniformly distributed, i.e. $a_0 + p \in F \cap U$. $\qquad \square$

Theorem 5.6.7. *Let $G = (G, \rho)$ be a compact abelian metrizable group with infinitely many elements and metric ρ. Let $(T_n)_{n=1}^{\infty}$ be a sequence of continuous, surjective group endomorphisms $T_n : G \to G$. Let $\Phi_n = T_n T_{n-1} \cdots T_1, \Phi_0 = \mathrm{id}_G$, and suppose that $\bigcup_{m=1}^{\infty} \ker \Phi_m$ is dense in G. Suppose further that the set*

181

$F = \{x \in G : (\Phi_n x)_{n=0}^{\infty} \text{ is uniformly distributed}\}$ is nonempty. Then there exists $f \in C(G)$ such that the set of $x \in G$ such that $\lim_{k \to \infty} \frac{1}{k} \sum_{i=0}^{k-1} f(\Phi_i x)$ exists is meager.

Proof. Choose $\delta > 0$ such that $\mu(B(0, \delta)) < \frac{1}{8}$, which exists by Lemma 5.6.3. Set

$$f(x) = \begin{cases} 1 & \rho(x, 0) \leq \frac{\delta}{2}, \\ 2 - \frac{2}{\delta}\rho(x, 0) & \frac{\delta}{2} \leq \rho(x, 0) \leq \delta, \\ 0 & \rho(x, 0) \geq \delta, \end{cases}$$

For each $K \in \mathbb{N}$, let A_K be the set

$$A_K = \left\{ x \in G : \left| \left(\frac{1}{k_1} \sum_{i=0}^{k_1-1} f(\Phi_i x) \right) - \left(\frac{1}{k_2} \sum_{i=0}^{k_2-1} f(\Phi_i x) \right) \right| < \frac{1}{4} \text{ for all } k_1, k_2 \geq K \right\}.$$

Since the set of $x \in G$ such that $\lim_{k \to \infty} \frac{1}{k} \sum_{i=0}^{k-1} f(\Phi_i x)$ exists is contained in $\bigcup_{K=1}^{\infty} A_K$, it will suffice to prove that each A_K is nowhere dense. Now fix $K \in \mathbb{N}$.

Let $\mathcal{O}$ be a nonempty open subset of G. By Lemma 5.6.4, there exists $L \geq K$ and a nonempty open subset $W_1 \subseteq \mathcal{O}$ such that $\frac{1}{L} \sum_{i=0}^{L-1} f(\Phi_i x) \geq \frac{7}{8}$. By Lemma 5.6.6, there exists $a \in F \cap W_1$. Since $a \in F$, there exists $N \geq L$ such that

$$k \geq N \Rightarrow \left| \int f d\mu - \frac{1}{k} \sum_{i=0}^{k-1} f(\Phi_i a) \right| \leq \frac{1}{8}.$$

But since $\int f d\mu < \frac{1}{8}$, it follows that

$$\frac{1}{N} \sum_{i=0}^{N-1} f(\Phi_i a) < \frac{1}{4}.$$

Since $\frac{1}{N} \sum_{i=0}^{N-1} f \circ \Phi_i$ is uniformly continuous, it follows that there exists an open neighborhood W_2 of a such that

$$x \in W_2 \Rightarrow \left| \left(\frac{1}{N} \sum_{i=0}^{N-1} f(\Phi_i x) \right) - \left(\frac{1}{N} \sum_{i=0}^{N-1} f(\Phi_i a) \right) \right| < \frac{1}{8}.$$

Then if $x \in W_2$, then

$$\frac{1}{N} \sum_{i=0}^{N-1} f(\Phi_i x) < \left(\frac{1}{N} \sum_{i=0}^{N-1} f(\Phi_i a) \right) + \frac{1}{8} < \frac{1}{4} + \frac{1}{8} = \frac{3}{8}.$$

Therefore, if $x \in W_1 \cap W_2$, then

$$\left(\frac{1}{L} \sum_{i=0}^{L-1} f\left(\Phi_i x\right) \right) - \left(\frac{1}{N} \sum_{i=0}^{N-1} f\left(\Phi_i x\right) \right) > \frac{7}{8} - \frac{3}{8} = \frac{1}{2}.$$

Thus $a \in W_1 \cap W_2 \subseteq \mathcal{O}$, and $(W_1 \cap W_2) \cap A_K = \emptyset$. $\qquad \square$

Throughout this section, we have relied heavily on the assumption that $\bigcup_{m=1}^{\infty} \ker \Phi_m$ is dense in G. This assumption still encompasses a wide class of interesting examples when $G = \mathbb{R}^d/\mathbb{Z}^d, d \in \mathbb{N}$. Endow G with the metric

$$\rho\left((t_1, \ldots, t_d) + \mathbb{Z}^d, (s_1, \ldots, s_d) + \mathbb{Z}^d \right) = \sum_{j=1}^{d} \min_{h \in \mathbb{Z}} |t_j - s_j + h|.$$

Note that the metric ρ is invariant under addition by elements of G.

Lemma 5.6.8. *Let $G = \mathbb{R}^d/\mathbb{Z}^d$, and let $T : G \to G$ be a continuous surjective group endomorphism. Let $A \in \mathbb{Z}^{d \times d}$ be the $d \times d$ integer matrix such that*

$$T : \begin{bmatrix} t_1 \\ \vdots \\ t_d \end{bmatrix} + \mathbb{Z}^d \mapsto A \begin{bmatrix} t_1 \\ \vdots \\ t_d \end{bmatrix} + \mathbb{Z}^d.$$

Then for every $x \in G$ exists $y \in \ker T$ such that $\rho(x, y) \leq d^2 \left\| A^{-1} \right\|_{\text{op}}$, where the operator norm is taken relative to the standard Euclidean norm on $\mathbb{R}^d$.

Proof. Let $\mathbf{e}_1, \ldots, \mathbf{e}_d$ be the standard basis of the real vector space $\mathbb{R}^d$, and let $\mathbf{f}_j = A^{-1} \mathbf{e}_j$ for $j = 1, \ldots, d$. Then

$$\ker T = \mathbb{Z}\mathbf{f}_1 + \cdots + \mathbb{Z}\mathbf{f}_d.$$

Let $x = (t_1, \ldots, t_d) + \mathbb{Z}^d \in G$. Since A is invertible, we know that $\{\mathbf{f}_1, \ldots, \mathbf{f}_d\}$ is a basis for $\mathbb{R}^d$, so there exist $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$ such that $\sum_{j=1}^{d} t_j \mathbf{e}_j = \sum_{j=1}^{d} \lambda_j \mathbf{f}_j$. Let $\ell_j = \lfloor \lambda_j \rfloor$, and let $y = \sum_{j=1}^{d} \ell_j \mathbf{f}_j + \mathbb{Z}^d$. Set

$\kappa_j = \lambda_j - \ell_j \in [0, 1)$. Then $y \in \ker T$, and

$$\rho(x, y) = \rho(0, y - x) = \rho\left(0, \sum_{j=1}^{d} (\lambda_j - \ell_j)\, \mathbf{f}_j + \mathbb{Z}^d\right) = \rho\left(0, \sum_{j=1}^{d} \kappa_j \mathbf{f}_j + \mathbb{Z}^d\right)$$

$$\leq \rho\left(0, \sum_{j=1}^{1} \kappa_j \mathbf{f}_j + \mathbb{Z}^d\right) + \rho\left(\sum_{j=1}^{1} \kappa_j \mathbf{f}_j + \mathbb{Z}^d, \sum_{j=1}^{2} \kappa_j \mathbf{f}_j + \mathbb{Z}^d\right)$$

$$+ \cdots + \rho\left(\sum_{j=1}^{d-1} \kappa_j \mathbf{f}_j + \mathbb{Z}^d, \sum_{j=1}^{d} \kappa_j \mathbf{f}_j + \mathbb{Z}^d\right)$$

$$= \sum_{j=1}^{d} \rho\left(0, \kappa_j \mathbf{f}_j + \mathbb{Z}^d\right).$$

For each $j \in \{1, \ldots, d\}$, let $\mathbf{f}_j = \sum_{i=1}^{d} b_{i,j} \mathbf{e}_j$. Then

$$\sum_{j=1}^{d} \rho\left(0, \kappa_j \mathbf{f}_j + \mathbb{Z}^d\right) = \sum_{j=1}^{d} \rho\left(0, \kappa_j \sum_{i=1}^{d} b_{i,j} \mathbf{e}_i + \mathbb{Z}^d\right)$$

$$= \sum_{j=1}^{d} \sum_{i=1}^{d} \min_{h \in \mathbb{Z}} |h - \kappa_j b_{i,j}|$$

$$\leq \sum_{j=1}^{d} \sum_{i=1}^{d} \kappa_j |b_{i,j}|$$

$$\leq \sum_{j=1}^{d} \sum_{i=1}^{d} |b_{i,j}| \qquad\qquad (0 \leq \kappa_j < 1)$$

$$\leq \sum_{j=1}^{d} \sum_{i=1}^{d} \|A^{-1} \mathbf{e}_j\|$$

$$\leq \sum_{j=1}^{d} \sum_{i=1}^{d} \|A^{-1}\|_{\mathrm{op}}$$

$$= d^2 \|A^{-1}\|_{\mathrm{op}}.$$

$\square$

Proposition 5.6.9. *Let $G = \mathbb{R}^d/\mathbb{Z}^d$, and let $(T_n)_{n=1}^{\infty}$ be a sequence of continuous surjective group endomorphisms of G onto itself. For each $n \in \mathbb{N}$, let $A_n \in \mathbb{Z}^{d \times d}$ be the $d \times d$ integer matrix such that*

$$
T_n : \begin{bmatrix} t_1 \\ \vdots \\ t_d \end{bmatrix} + \mathbb{Z}^d \mapsto A_n \begin{bmatrix} t_1 \\ \vdots \\ t_d \end{bmatrix} + \mathbb{Z}^d.
$$

Then if $\liminf_{n\to\infty} \left\| (A_n A_{n-1} \cdots A_1)^{-1} \right\|_{\mathrm{op}} = 0$, *then* $\bigcup_{m=1}^{\infty} \ker \Phi_m$ *is dense in G.*

Proof. Let $x \in G$, and let $\delta > 0$. Choose $m \in \mathbb{N}$ such that $\left\| (A_m A_{m-1} \cdots A_1)^{-1} \right\|_{\mathrm{op}} < \frac{\delta}{d^2}$. Then Φ_m is implemented by the matrix $A_m A_{m-1} \cdots A_1$, so Lemma 5.6.8 tells us there exists $y \in \ker \Phi_m$ such that $\rho(x,y) \leq d^2 \left\| (A_m A_{m-1} \cdots A_1)^{-1} \right\|_{\mathrm{op}} < \delta$. Therefore $\bigcup_{m=1}^{\infty} \ker \Phi_m$ is dense in G. $\qquad\square$

Noncommutative ergodic optimization

One of the guiding questions of the field of ergodic optimization is the following: Given a topological dynamical system (X, G, U), and a real-valued continuous function $f \in C(X)$, what values can $\int f \mathrm{d}\mu$ take when μ is an invariant Borel probability measure on X, and in particular, what are the extreme values it can take? In a joint work with I. Assani (Assani and Young, 2022, Section 3), we noticed that the field of ergodic optimization was relevant to the study of certain temporo-spatial differentiation problems. Hoping to extend these tools to the study of temporo-spatial differentiation problems in the setting of operator-algebraic dynamical systems, this chapter develops an operator-algebraic formalization of this question of ergodic optimization, re-interpreting it as a question about the values of invariant states on a C*-dynamical system. This framework is then applied to provide a characterization of certain uniquely ergodic C*-dynamical systems with respect to ergodic optimizations.

Section 6.1 develops the theory of ergodic optimization in the context of C*-dynamical systems, where the role of "maximizing measures" is instead played by invariant states on a C*-algebra. The framework we adopt is in fact somewhat more general than the classical framework of maximizing measures, since we consider ergodic optimizations relative to a restricted class of invariant states, which we call relative ergodic optimizations. We also demonstrate that some of the basic results of that classical theory of ergodic optimization extend to the C*-dynamical setting.

In Section 6.2, we define a value called the *gauge* of a singly generated C*-dynamical system, a non-commutative generalization of the functional of the same name defined in (Assani and Young, 2022), and describe its connections to questions of ergodic optimization, as well as the ways in which it can be used to "detect" the unique ergodicity of C*-dynamical systems under certain Choquet-theoretic assumptions.

In Section 6.3, we extend the results of the previous section to the case where the phase group is a countable discrete amenable group. We also provide a characterization of uniquely ergodic C*-dynamical systems of countable discrete amenable groups in terms of various notions of convergence of ergodic averages.

In Section 6.4, we relate the convergence properties of certain ergodic averages to relative ergodic optimizations.

Finally, in Section 6.5, we provide alternate proofs of several results from this chapter using the toolbox of nonstandard analysis.

6.1 Ergodic Optimization in C*-Dynamical Systems

Given a unital C*-algebra $\mathfrak{A}$, let $\mathrm{Aut}(\mathfrak{A})$ denote the family of all *-automorphisms of $\mathfrak{A}$. We endow $\mathrm{Aut}(\mathfrak{A})$ with the *point-norm topology*, i.e. the topology induced by the pseudometrics

$$(\Phi, \Psi) \mapsto \|\Phi(a) - \Psi(a)\| \qquad (a \in \mathfrak{A}).$$

This topology makes $\mathrm{Aut}(\mathfrak{A})$ a topological group (Blackadar, 2006, II.5.5.4).

We define a *C*-dynamical system* to be a triple $(\mathfrak{A}, G, \Theta)$ consisting of a unital C*-algebra $\mathfrak{A}$, a topological group G (called the *phase group*), and a point-continuous left group action $\Theta : G \to \mathrm{Aut}(\mathfrak{A})$.

Notation 6.1.1. Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system, and let $F \subseteq G$ be a nonempty finite subset. We define $\mathrm{Avg}_F : \mathfrak{A} \to \mathfrak{A}$ by

$$\mathrm{Avg}_F x := \frac{1}{|F|} \sum_{g \in F} \Theta_g a.$$

Denote by $\mathcal{S}$ the family of all states on $\mathfrak{A}$ endowed with the weak*-topology, and by $\mathcal{T}$ the subfamily of all tracial states on $\mathfrak{A}$. A state ϕ on $\mathfrak{A}$ is called Θ-*invariant* (or simply *invariant* if the action Θ is understood in context) if $\phi = \phi \circ \Theta_g$ for all $g \in G$. Denote by $\mathcal{S}^G \subseteq \mathcal{S}$ the family of all Θ-invariant states on $\mathfrak{A}$, and by $\mathcal{T}^G \subseteq \mathcal{T}$ the family of all Θ-invariant tracial states on $\mathfrak{A}$. The set $\mathcal{S}^G$ (resp. $\mathcal{T}^G$) is weak*-compact in $\mathcal{S}$ (resp. in $\mathcal{T}$). Unless otherwise stated, whenever we deal with subspaces of $\mathcal{S}$, we consider these subspaces equipped with the weak*-topology.

We will assume for the remainder of this section that $(\mathfrak{A}, G, \Theta)$ is a C*-dynamical system such that $\mathfrak{A}$ is separable, and also that $\mathcal{S}^G \neq \emptyset$. This framework will include every system of the form $(C(Y), G, \Theta)$, where Y is a compact metrizable topological space, the group G is countable, discrete, and amenable, and Θ is of the form $\Theta_g : f \mapsto f \circ U_g$ for all $g \in G$, where $U : G \curvearrowright Y$ is a *right* action of G on Y by homeomorphisms. Because of the correspondence between topological dynamical systems as we've defined them previously in

Section 3.1 and C*-dynamical systems over commutative C*-algebras, it is customary to call a C*-dynamical system a "non-commutative topological dynamical systems."

Before proceeding, we prove the following Krylov–Bogolyubov-type result, which will be useful to establish the Θ-invariance of certain states later.

Lemma 6.1.2. *Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system, and let G be an amenable group. If $(\phi_k)_{k=1}^{\infty}$ is a sequence in $\mathcal{S}$, and $\mathbf{F} = (F_k)_{k=1}^{\infty}$ is a right Følner sequence for G, then any weak*-limit point of the sequence $\left(\phi_k \circ \mathrm{Avg}_{F_k}\right)_{k=1}^{\infty}$ is Θ-invariant. In particular, if K is a nonempty, Θ-invariant, weak*-compact, convex subset of $\mathcal{S}$, then $K \cap \mathcal{S}^G \neq \emptyset$.*

Proof. Let $(\phi_k)_{k=1}^{\infty}$ be a sequence of states, and fix $g_0 \in G, x \in \mathfrak{A}$. Then

$$\left| \phi_k \left(\mathrm{Avg}_{F_k} \Theta_{g_0} x \right) - \phi_k \left(\mathrm{Avg}_{F_k} x \right) \right|$$

$$= \left| \frac{1}{|F_k|} \left[\phi_k \left(\sum_{g \in F_k g_0} \Theta_g x \right) - \phi_k \left(\sum_{g \in F_k} \Theta_g x \right) \right] \right|$$

$$\leq \left| \frac{1}{|F_k|} \phi_k \left(\sum_{g \in F_k g_0 \setminus F_k} \Theta_g x \right) \right| + \left| \frac{1}{|F_k|} \phi_k \left(\sum_{g \in F_k \setminus F_k g_0} \Theta_g x \right) \right|$$

$$\leq \frac{|F_k g_0 \Delta F_k|}{|F_k|} \|x\|$$

$$\overset{k \to \infty}{\to} 0.$$

Therefore, if $k_1 < k_2 < \cdots$ is such that $\psi = \lim_{\ell \to \infty} \phi_{k_\ell} \circ \mathrm{Avg}_{F_{k_\ell}}$ exists, then

$$|\psi(\Theta_g x) - \psi(x)| \leq \limsup_{\ell \to \infty} \frac{|F_{k_\ell} g_0 \Delta F_{k_\ell}|}{|F_{k_\ell}|} \|x\| = 0.$$

Finally, let K be a nonempty, Θ-invariant, weak*-compact, convex subset of $\mathcal{S}$. Let ϕ be any state in K, and consider the sequence $\left(\phi \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty}$. By the convexity and Θ-invariance of K, every term of this sequence is an element of K, and since K is compact, there exists a subsequence of this sequence which converges in K. As has already been shown, that limit must be an element of $\mathcal{S}^G$. $\square$

Remark 6.1.3. Lemma 6.5.1 can be seen as a nonstandard-analytic analogue to Lemma 6.1.2.

Although our manner of proof of Lemma 6.1.2 is scarcely novel, the result as we have stated it here can be used to ensure the existence of invariant states with specific properties that might interest us, as seen

for example in Corollary 6.1.4 and Proposition 6.1.14. Our standing hypothesis that $\mathfrak{A}$ be separable is not necessary for this proof of Lemma 6.1.2.

Corollary 6.1.4. *If $\mathcal{T} \neq \emptyset$, and G is amenable, then $\mathcal{T}^G \neq \emptyset$.*

Proof. Apply Lemma 6.1.2 to the case where $K = \mathcal{T}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Definition 6.1.5. We denote by $\mathfrak{R}$ the real Banach space of all self-adjoint elements of $\mathfrak{A}$, and denote by $\mathfrak{R}^\natural$ the space of all real self-adjoint bounded linear functionals on $\mathfrak{A}$.

Definition 6.1.6. Let V be a locally convex topological real vector space, and let K be a compact subset of V which is contained in a hyperplane that does not contain the origin. We call K a *simplex* if the positive cone $P = \{ck : c \in \mathbb{R}_{\geq 0}, k \in K\}$ defines a lattice ordering on $P - P = \{p_1 - p_2 : p_1, p_2 \in P\} \subseteq V$ with respect to the partial order $a \leq b \iff b - a \in P$.

Remark 6.1.7. In Definition 6.1.6, the assumption that K lives in a hyperplane that does not contain the origin is technically superfluous, but simplifies the theory somewhat (see (Phelps, 2001, Section 10)), and is satisfied by all the simplices that interest us here. Specifically, we know that $\mathcal{S}$ (and by extension $\mathcal{S}^G, \mathcal{T}, \mathcal{T}^G$) lives in the real hyperplane $\{\phi \in \mathfrak{R}^\natural : \phi(1) = 1\}$ defined by the evaluation at 1.

We begin with the following lemma.

Lemma 6.1.8.   *(i) The spaces $\mathcal{S}, \mathcal{S}^G, \mathcal{T}, \mathcal{T}^G$ are compact and metrizable.*

(ii) If $\mathcal{T} \neq \emptyset$, then the space $\mathcal{T}^G$ is a simplex.

Before proving this lemma, we need to introduce some terminology. Let ϕ, ψ be two positive linear functionals on a unital C*-algebra $\mathfrak{A}$. We say that the two positive functionals are *orthogonal*, notated $\phi \perp \psi$, if they satisfy either of the following two equivalent conditions:

(a) $\|\phi + \psi\| = \|\phi\| + \|\psi\|$.

(b) For every $\epsilon > 0$ exists positive $z \in \mathfrak{A}$ of norm ≤ 1 such that $\phi(1 - z) < \epsilon, \psi(z) < \epsilon$.

It is well-know that these conditions are equivalent (Pedersen, 1979, Lemma 3.2.3). For every $\phi \in \mathfrak{R}^\natural$, there exist unique positive linear functionals ϕ^+, ϕ^- such that $\phi = \phi^+ - \phi^-$, and $\phi^+ \perp \phi^-$, called the *Jordan decomposition* of ϕ (Blackadar, 2006, II.6.3.4).

Before proving Lemma 6.1.8, we demonstrate the following property of the Jordan decomposition of a tracial functional.

Lemma 6.1.9. *Let $\mathfrak{A}$ be a unital C*-algebra, and $\phi \in \mathfrak{R}^\natural$. Suppose that $\phi(xy) = \phi(yx)$ for all $x, y \in \mathfrak{A}$. Then $\phi^\pm(xy) = \phi^\pm(yx)$ for all $x, y \in \mathfrak{A}$.*

Proof. Let $\mathcal{U}(\mathfrak{A})$ denote the group of unitary elements in $\mathfrak{A}$. For a unitary element $u \in \mathcal{U}(\mathfrak{A})$, let $\mathrm{Ad}_u \in \mathrm{Aut}(\mathfrak{A})$ denote the inner automorphism

$$\mathrm{Ad}_u\, x = uxu^*.$$

Let $\psi \in \mathfrak{A}'$. We claim that ψ is tracial if and only if $\psi \circ \mathrm{Ad}_u = \psi$ for all unitaries $u \in \mathcal{U}(\mathfrak{A})$.

Let $u \in \mathcal{U}(\mathfrak{A})$ be unitary, and $x \in \mathfrak{A}$ an arbitrary element. Then

$$\phi(ux) = \psi\left(u(xu)u^*\right)$$
$$= \psi(\mathrm{Ad}_u(xu)).$$

So $\psi(ux) = \psi(xu)$ if and only if $\psi(\mathrm{Ad}_u(xu)) = \psi(xu)$.

In one direction, suppose that $\psi = \psi \circ \mathrm{Ad}_u$ for all $u \in \mathcal{U}(\mathfrak{A})$. Fix $x, y \in \mathfrak{A}$. Then we can write $y = \sum_{j=1}^4 c_j u_j$ for some $c_1, \ldots, c_4 \in \mathbb{C}$ and unitaries $u_1, \ldots, u_4 \in \mathcal{U}(\mathfrak{A})$ unitary. Then

$$\psi(xy) = \psi\left(x \sum_{j=1}^4 c_j u_j\right)$$
$$= \sum_{j=1}^4 c_j \psi(xu_j)$$
$$= \sum_{j=1}^4 c_j \psi(\mathrm{Ad}_{u_j}(xu_j))$$
$$= \sum_{j=1}^4 c_j \psi(u_j x)$$
$$= \psi\left(\left(\sum_{j=1}^4 c_j u_j\right) x\right)$$
$$= \psi(yx).$$

Thus ψ is tracial.

In the other direction, suppose there exists $u \in \mathcal{U}(\mathfrak{A})$ such that $\psi \circ \mathrm{Ad}_u \neq \psi$. Let $y \in \mathfrak{A}$ such that $\psi(y) \neq \psi(\mathrm{Ad}_u y)$, and let $x = yu^*$. Then

$$\psi(xu) = \psi(y)$$
$$\neq \psi(\mathrm{Ad}_u y)$$
$$= \psi\left(uyu^*\right)$$
$$= \psi(ux).$$

Therefore ψ is not tracial.

Now, if $\phi \in \mathfrak{R}^\natural$ is tracial, then $\phi \circ \mathrm{Ad}_u = \phi$ for all $u \in \mathcal{U}(\mathfrak{A})$. Then $\phi = \phi \circ \mathrm{Ad}_u = (\phi^+ \circ \mathrm{Ad}_u) - (\phi^- \circ \mathrm{Ad}_u)$. But $\|\phi^\pm \circ \mathrm{Ad}_u\| = \|\phi^\pm\|$, so it follows that $\|\phi\| = \|\phi^+ \circ \mathrm{Ad}_u\| + \|\phi^- \circ \mathrm{Ad}_u\|$. Therefore $\phi = (\phi^+ \circ \mathrm{Ad}_u) - (\phi^- \circ \mathrm{Ad}_u)$ is an orthogonal decomposition of ϕ, and so it is *the* Jordan decomposition. This means that $\phi^\pm = \phi^\pm \circ \mathrm{Ad}_u$. Since this is true for all $u \in \mathcal{U}(\mathfrak{A})$, it follows that $\phi^\pm$ are tracial. $\qquad\square$

Proof of Lemma 6.1.8.   (i) This all follows because $\mathcal{S}$ is a weak*-closed real subspace of the unit ball in the continuous dual of the separable Banach space $\mathfrak{R}$, and the spaces $\mathcal{S}^G, \mathcal{T}, \mathcal{T}^G$ are all closed subspaces of $\mathcal{S}$.

(ii) It is a standard fact that if $\mathcal{T} \neq \emptyset$, then $\mathcal{T}$ is a simplex (Blackadar, 2006, II.6.8.11). Let

$$C^G = \left\{ c\phi : c \in \mathbb{R}_{\geq 0}, \phi \in \mathcal{T}^G \right\}$$

be the positive cone of $\mathcal{T}^G$, and let $\mathfrak{R}^\natural$ denote the (real) space of all bounded self-adjoint tracial linear functionals on $\mathfrak{A}$. Let E^G denote the (real) space of all bounded self-adjoint Θ-invariant linear functionals on $\mathfrak{A}$. We already know that $\mathcal{T}$ lives in a hyperplane of $\mathfrak{R}^\natural$ defined by the evaluation functional $\phi \mapsto \phi(1)$. It will therefore suffice to show that $E^G = C^G - C^G$, and that E^G is a sub-lattice of $\mathfrak{R}^\natural$.

Let $\phi^+, \phi^- \geq 0$ be positive functionals on $\mathfrak{A}$ such that $\phi = \phi^+ - \phi^-$ is tracial, and $\phi^+ \perp \phi^-$. By Lemma 6.1.9, we know that ϕ^+, ϕ^- are tracial. We claim that if $\phi \in E^G$, then $\phi^+, \phi^- \in C^G$. To prove this, let $g \in G$, and consider that $\phi^+ \circ \Theta_g, \phi^- \circ \Theta_g$ are both positive linear functionals such that $\phi = (\phi^+ \circ \Theta_g) - (\phi^- \circ \Theta_g)$.

We claim that $(\phi^+ \circ \Theta_g) \perp (\phi^- \circ \Theta_g)$. Fix $\epsilon > 0$. We know that there exists $z \in \mathfrak{A}$ such that $\|z\| \leq 1, 0 \leq z$, and such that $\phi^+(1-z) < \epsilon, \phi^-(z) < \epsilon$. Then $\Theta_{g^{-1}}(z)$ is a positive element of norm ≤ 1 such that

$$
\begin{aligned}
\phi^+\left(\Theta_g\left(\Theta_{g^{-1}}(1-z)\right)\right) &= \phi^+(1-z) && < \epsilon, \\
\phi^-\left(\Theta_g\left(\Theta_{g^{-1}}(z)\right)\right) &= \phi^-(z) && < \epsilon.
\end{aligned}
$$

Therefore $(\phi^+ \circ \Theta_g) - (\phi^- \circ \Theta_g)$ is a Jordan decomposition of ϕ, and since the Jordan decomposition is unique, it follows that $\phi^+ = \phi^+ \circ \Theta_g, \phi^- = \phi^- \circ \Theta_g$, i.e. that $\phi^+, \phi^- \in C^G$. This means that $E^G = C^G - C^G$.

We now want to show that $E^G = C^G - C^G$ is a sublattice of E, i.e. that it is closed under the lattice operations. Let $\phi, \psi \in E^G$. For this calculation, we draw on the identities listed in (Aliprantis and Burkinshaw, 2006, Theorem 1.3). Then

$$
\begin{aligned}
\phi \vee \psi &= (((\phi - \psi) + \psi) \vee (0 + \psi)) \\
&= ((\phi - \psi) \vee 0) + \psi \\
&= (\phi - \psi)^+ + \psi, \\
\phi \wedge \psi &= ((\phi - \psi) + \psi) \wedge (0 + \psi) \\
&= ((\phi - \psi) \wedge 0) + \psi \\
&= -((-(\phi - \psi)) \vee 0) + \psi \\
&= -(\psi - \phi)^+ + \psi.
\end{aligned}
$$

Therefore, if E^G is a real linear space and is closed under the operations $\phi \mapsto \phi^+, \phi \mapsto \phi^-$, then it is also closed under the lattice operations. Thus E^G is a sublattice of $\mathfrak{R}^\natural$.

Hence, the subset $\mathcal{T}^G$ is a compact metrizable simplex.

$\square$

In order to keep our treatment relatively self-contained, we define here several elementary concepts from Choquet theory that will be relevant in this section.

Definition 6.1.10. Let S_1, S_2 be convex spaces. We call a map $T : S_1 \to S_2$ an *affine map* if for every $v, w \in S_1; \ t \in [0, 1]$, we have

$$T(tv + (1 - t)w) = tT(v) + (1 - t)T(w).$$

In the case where $S_2 \subseteq \mathbb{R}$, we call T an *affine functional*.

Definition 6.1.11. Let K be a convex subset of a locally convex real topological vector space V.

(a) A point $k \in K$ is called an *extreme point* of K if for every pair of points $k_1, k_2 \in K$ and parameter $t \in [0, 1]$ such that $k = tk_1 + (1 - t)k_2$, either $k_1 = k_2$ or $t \in \{0, 1\}$. In other words, we call k extreme if there is no nontrivial way of expressing k as a convex combination of elements of K.

(b) The set of all extreme points of K is denoted $\partial_e K$.

(c) A subset F of K is called a *face* if for every pair $k_1, k_2 \in K, t \in (0, 1)$ such that $tk_1 + (1 - t)k_2 \in F$, we have that $k_1, k_2 \in F$.

(d) A face F of K is called an *exposed face* of K if there exists a continuous affine functional $\ell : K \to \mathbb{R}$ such that $\ell(x) = 0$ for all $x \in F$, and $\ell(y) < 0$ for all $y \in K \setminus F$.

(e) A point $k \in K$ is called an *exposed point* of K if $\{k\}$ is an exposed face of K.

(f) Given a subset $\mathcal{E}$ of K, the closed convex hull of $\mathcal{E}$ is written as $\overline{\mathrm{co}}(\mathcal{E})$.

We now introduce the basic concepts in our treatment of ergodic optimization.

Definition 6.1.12. Let $x \in \mathfrak{R}$ be a self-adjoint element, and let $K \subseteq \mathcal{S}^G$ be a compact convex subset of $\mathcal{S}^G$. Define a value $m(x|K)$ by

$$m(x|K) := \sup_{\psi \in K} \psi(x).$$

We say a state $\phi \in K$ is $(x|K)$-*maximizing* if $\phi(x) = m(x|K)$. Let $K_{\max}(x) \subseteq K$ denote the set of all $(x|K)$-maximizing states. A state $\phi \in K$ is called *uniquely* $(x|K)$-*maximizing* if $K_{\max}(x) = \{\phi\}$.

Remark 6.1.13. We note here a trivial inequality: If $K_1 \subseteq K_2$ are compact convex subsets of $\mathcal{S}^G$, then $m(x|K_1) \leq m(x|K_2)$, and in particular, we will always have $m(x|K_1) \leq m(x|\mathcal{S}^G)$.

We will single out one type of compact convex subset of $\mathcal{S}^G$ which will prove important later. Given a subset $A \subseteq \mathfrak{A}$, set

$$\mathrm{Ann}(A) := \left\{ \phi \in \mathcal{S}^G : A \subseteq \ker \phi \right\}.$$

When $\mathfrak{I} \subseteq \mathfrak{A}$ is a Θ-invariant closed ideal of $\mathfrak{A}$, we have a bijective correspondence between the states in $\mathrm{Ann}(\mathfrak{I})$ and the states on $\mathfrak{A}/\mathfrak{I}$ invariant under the action induced by Θ. We will be referring to this set again in Sections 6.2 and 6.3, when values of the form $m\left(a|\,\mathrm{Ann}(A)\right)$ come up in reference to certain ergodic averages. We observe that $\mathrm{Ann}(\{0\}) = \mathcal{S}^G$, and that $A \subseteq B \subseteq \mathfrak{A} \Rightarrow \mathrm{Ann}(A) \supseteq \mathrm{Ann}(B)$. There is also no a priori guarantee that $\mathrm{Ann}(A) \neq \emptyset$, since for example $\mathrm{Ann}(\{1\}) = \emptyset$. However, Proposition 6.1.14 gives sufficient conditions for $\mathrm{Ann}(A)$ to be nonempty.

Proposition 6.1.14. *Let $A \subseteq \mathfrak{A}$ be such that $\Theta_g A \subseteq A$ for all $g \in G$. Suppose there exists a state on $\mathfrak{A}$ which vanishes on A. Then $\mathrm{Ann}(A) \neq \emptyset$. In particular, if $\mathfrak{I} \subsetneq \mathfrak{A}$ is a proper closed two-sided ideal of $\mathfrak{A}$ for which $\Theta_g \mathfrak{I} = \mathfrak{I}$ for all $g \in G$, then $\mathrm{Ann}(\mathfrak{I}) \neq \emptyset$.*

Proof. Let $K \subseteq \mathcal{S}$ denote the family of all (not necessarily invariant) states on $\mathfrak{A}$ which vanish on A. Then if $\phi \in K$ and $a \in A$, then $\Theta_g a \in A$, so $\phi \circ \Theta_g$ vanishes on A. Therefore $\Theta_g K \subseteq K$ for all $g \in G$. It follows from Lemma 6.1.2 that $K \cap \mathcal{S}^G = \mathrm{Ann}(A) \neq \emptyset$.

Suppose $\mathfrak{I} \subsetneq \mathfrak{A}$ is a proper closed two-sided ideal of $\mathfrak{A}$ for which $\Theta_g \mathfrak{I} = \mathfrak{I}$ for all $g \in G$, and let $\pi : \mathfrak{A} \twoheadrightarrow \mathfrak{A}/\mathfrak{I}$ be the canonical quotient map. Let $\tilde{\Theta} : G \to \mathrm{Aut}(\mathfrak{A}/\mathfrak{I})$ be the induced action of G on $\mathfrak{A}/\mathfrak{I}$ by $\tilde{\Theta}_g(a + \mathfrak{I}) = \Theta_g a + \mathfrak{I}$. Let ψ be a $\tilde{\Theta}$-invariant state on $\mathfrak{A}/\mathfrak{I}$. Then $\psi \circ \pi$ is a Θ-invariant state on $\mathfrak{A}$ which vanishes on $\mathfrak{I}$, i.e. $\psi \circ \pi \in \mathrm{Ann}(\mathfrak{I})$. $\square$

Proposition 6.1.15. *Let $K \subseteq \mathcal{S}^G$ be a nonempty compact convex subset of $\mathcal{S}^G$, and let $x \in \mathfrak{R}$. Then $K_{\max}(x)$ is a nonempty, compact, exposed face of K.*

Proof. To see that $K_{\max}(x)$ is nonempty, for each $n \in \mathbb{N}$, let $\phi_n \in K$ such that $\phi_n(x) \geq m(x|K) - \frac{1}{n}$. Then since K is compact, the sequence $(\phi_n)_{n=1}^{\infty}$ has a convergent subsequence. Let ϕ be the limit of a convergent subsequence of $(\phi_n)_{n=1}^{\infty}$. Then ϕ is $(x|K)$-maximizing.

To see that $K_{\max}(x)$ is compact, consider that

$$K_{\max}(x) = \{\phi \in K : \phi(x) = m(x|K)\},$$

which is a closed subset of K. As for being an exposed face, consider the continuous affine functional $\ell : K \to \mathbb{R}$ given by

$$\ell(\phi) = \phi(x) - m(x|K).$$

Then the functional ℓ exposes $K_{\max}(x|K)$, since it is nonpositive on all of K and vanishes exactly on $K_{\max}(x)$. $\qquad\square$

The following result describes the ways in which some ergodic optimizations interact with equivariant *-homomorphisms of C*-dynamical systems.

Theorem 6.1.16. *Let $(\mathfrak{A}, G, \Theta)$, $\left(\tilde{\mathfrak{A}}, G, \tilde{\Theta}\right)$ be two C*-dynamical systems, and let $\pi : \mathfrak{A} \to \tilde{\mathfrak{A}}$ be a surjective *-homomorphism such that*

$$\tilde{\Theta}_g \circ \pi = \pi \circ \Theta_g \qquad\qquad (\forall g \in G).$$

Let $\tilde{\mathcal{S}}^G$ denote the space of $\tilde{\Theta}$-invariant states on $\tilde{\Theta}$. Then $m\left(\pi(a)|\tilde{\mathcal{S}}^G\right) = m\left(a| \operatorname{Ann}(\ker \pi)\right)$.

Proof. Let $\tilde{\mathcal{S}}^G$ denote the space of $\tilde{\Theta}$-invariant states on $\tilde{\mathfrak{A}}$. We claim that there is a natural bijective correspondence between $\tilde{\mathcal{S}}^G$ and $\operatorname{Ann}(\ker \pi)$. If ϕ is a $\tilde{\Theta}$-invariant state on $\tilde{\mathfrak{A}}$, then we can pull it back to a Θ-invariant state ϕ_0 on $\mathfrak{A}$ by

$$\phi_0 = \phi \circ \pi.$$

This ϕ_0 obviously vanishes on $\ker \pi$, and is Θ-invariant by virtue of the equivariance property of π. Conversely, if we start with a Θ-invariant state ψ on $\mathfrak{A}$ that vanishes on $\ker \pi$, then we can push it to a $\tilde{\Theta}$-invariant state $\tilde{\psi}$ on $\tilde{\mathfrak{A}}$ by

$$\tilde{\psi} \circ \pi = \psi.$$

We claim now that

$$m\left(a| \operatorname{Ann}(\ker \pi)\right) = m\left(\pi(a)|\tilde{\mathcal{S}}^G\right).$$

Let ϕ be a $\left(\pi(a)|\tilde{\mathcal{S}}^G\right)$-maximizing state on $\tilde{\mathfrak{A}}$. Then $\phi \circ \pi \in \operatorname{Ann}(\ker \pi)$, so

$$m\left(\pi(a)|\tilde{\mathcal{S}}^G\right) = \phi(\pi(a)) \leq m\left(a| \operatorname{Ann}(\ker \pi)\right).$$

On the other hand, if $\psi \in \mathrm{Ann}(\ker \pi)$ is $(a|\mathrm{Ann}(\ker \pi))$-maximizing, then let $\tilde{\psi}$ be such that $\tilde{\psi} \circ \pi = \psi$. Then $\tilde{\psi} \in \tilde{\mathcal{S}}^G$, so

$$m\left(a|\mathrm{Ann}(\ker \pi)\right) = \psi(a) = \tilde{\psi}(\pi(a)) \leq m\left(a|\tilde{\mathcal{S}}^G\right).$$

$\square$

The assumption in Theorem 6.1.16 that π is surjective is actually superfluous, as shown in Corollary 6.3.8. We will later provide a proof of this stronger claim that uses the gauge functional, introduced in the context of actions of $\mathbb{Z}$ in Section 6.2 and in the context of actions of amenable groups in Section 6.3.

Moreover, the proof of Theorem 6.1.16 can be extended to establish a correspondence between ergodic optimization over certain compact convex subsets of $\tilde{\mathcal{S}}^G$ and certain compact convex subsets of $\mathrm{Ann}(\ker \pi)$. For example under the same hypotheses, if $\mathcal{T} \neq \emptyset$, then the proof could be modified in a simple manner to establish that $m\left(\pi(a)|\tilde{\mathcal{T}}^G\right) = m\left(a|\mathrm{Ann}(\ker \pi) \cap \mathcal{T}^G\right)$, where $\tilde{\mathcal{T}}^G$ denotes the $\tilde{\Theta}$-invariant tracial states on $\tilde{\mathfrak{A}}$. In lieu of stating Theorem 6.1.16 in greater generality, we content ourselves to state this special case (which we will use in future sections) and remark that the argument can be generalized further.

The following characterization of exposed faces in compact metrizable simplices will prove useful.

Lemma 6.1.17. *Let K be a compact metrizable simplex. Then every closed face of K is exposed.*

Proof. See (Davies, 1967, Theorem 7.4). $\square$

The theorem we are building to in this section is as follows.

Theorem 6.1.18. *Let $K \subseteq \mathcal{S}^G$ be a compact simplex. Then the closed faces of K are exactly the sets of the form $K_{\max}(x)$ for some $x \in \mathfrak{R}$.*

Before we can prove our main theorem of this section, we will need to prove the following result, which gives us a means by which to build an important linear functional.

Theorem 6.1.19. *Let $K \subseteq \mathcal{S}^G$ be a compact simplex, and let $\ell : K \to \mathbb{R}$ be a continuous affine functional. Then there exists a continuous linear functional $\tilde{\ell} : \overline{\mathrm{span}}_{\mathbb{R}}(K) \to \mathbb{R}$ such that $\tilde{\ell}|_K = \ell$.*

To prove this theorem, we break it up into several parts, attaining the extension $\tilde{\ell}$ as the final step of a few subsequent extensions of ℓ.

Lemma 6.1.20. *Let $K \subseteq \mathcal{S}^G$ be a compact metrizable simplex, and let $\ell : K \to \mathbb{R}$ be a continuous affine functional. Let $P = \{c\phi : c \in \mathbb{R}_{\geq 0}, \phi \in K\}$. Then there exists a continuous functional $\ell_1 : P \to \mathbb{R}$ satisfying the following conditions for all $f_1, f_2 \in P; c \in \mathbb{R}_{\geq 0}$:*

(a) $\ell_1(cf_1) = c\ell_1(f_1)$,

(b) $\ell_1(f_1 + f_2) = \ell_1(f_1) + \ell_2(f_2)$,

(c) $\ell_1|_K = \ell$.

Proof. Note that every nonzero element of P can be expressed uniquely as $c\phi$ for some $c \in \mathbb{R}_{\geq 0} \setminus \{0\}, \phi \in K$. As such we define

$$\ell_1(c\phi) = \begin{cases} c\ell(\phi) & c > 0 \\ 0 & c = 0 \end{cases}$$

It is immediately clear that this ℓ_1 satisfies conditions (a) and (c), leaving only (b) to check.

Now, suppose that $f_1 = c_1\phi_1, f_2 = c_2\phi_2$ for some $\phi_1, \phi_2 \in K; c_1, c_2 \in \mathbb{R}_{\geq 0}$. Consider first the case where at least one of c_1, c_2 are nonzero. Then

$$
\begin{aligned}
f_1 + f_2 &= c_1\phi_1 + c_2\phi_2 \\
&= (c_1 + c_2)\left(\frac{c_1}{c_1 + c_2}\phi_1 + \frac{c_2}{c_1 + c_2}\phi_2\right) \\
\Rightarrow \ell_1(f_1 + f_2) &= \ell_1\left(c_1\phi_1 + c_2\phi_2\right) \\
&= (c_1 + c_2)\ell\left(\frac{c_1}{c_1 + c_2}\phi_1 + \frac{c_2}{c_1 + c_2}\phi_2\right) \\
[\text{because } \ell \text{ is affine}] &= (c_1 + c_2)\left(\frac{c_1}{c_1 + c_2}\ell(\phi_1) + \frac{c_2}{c_1 + c_2}\ell(\phi_2)\right) \\
&= c_1\ell(\phi_1) + c_2\ell(\phi_2) \\
&= \ell_1(c_1\phi_1) + \ell_1(c_2\phi_2) \\
&= \ell_1(f_1) + \ell_1(f_2).
\end{aligned}
$$

In the event that $c_1 = c_2 = 0$, then the additivity property attains trivially.

It remains now to show that ℓ_1 is continuous. We will check continuity at nonzero points in P, and then at $0 \in P$. First, consider the case where $c\phi \in P \setminus \{0\}$, and $c \in \mathbb{R}_{\geq 0}, \phi \in K$. Suppose that $(c_n\phi_n)_n$ is a

sequence in P converging in the weak*-topology to $c\phi$. We claim that $c_n \to c$ in $\mathbb{R}$, and $\phi_n \to \phi$ in the weak*-topology.

We first observe that $(c_n\phi_n)(1) = c_n$, so $(c_n)_n$ converges in $\mathbb{R}_{\geq 0}$ to c, meaning in particular that for sufficiently large n, we have that $c_n \in \left[\frac{c}{2}, \frac{3c}{2}\right]$. Now, if $\lambda : \mathfrak{R} \to \mathbb{R}$ is a norm-continuous linear functional, then

$$
\begin{aligned}
\lambda(\phi_n) &= \frac{1}{c_n}\lambda(c_n\phi_n) \\
&\to \frac{1}{c}\lambda(c\phi) \\
&= \lambda(\phi).
\end{aligned}
$$

Therefore $c_n \to c, \phi_n \to \phi$. Thus we can compute

$$
\begin{aligned}
|\ell_1(c\phi) - \ell_1(c_n\phi_n)| &\leq |\ell_1(c\phi) - \ell_1(c_n\phi)| + |\ell_1(c_n\phi) - \ell_1(c_n\phi_n)| \\
&= |c - c_n| \cdot |\ell(\phi)| + |c_n| \cdot |\ell(\phi) - \ell(\phi_n)| \\
&\leq |c - c_n|\left(\sup_{\phi \in K}|\ell(\phi)|\right) + \frac{3c}{2}|\ell(\phi) - \ell(\phi_n)| \\
&\overset{n \to \infty}{\to} 0,
\end{aligned}
$$

where $\sup_{\phi \in K}|\ell(\phi)|$ must be finite because K is weak*-compact, and $|\ell(\phi) - \ell(\phi_n)| \overset{n \to \infty}{\to} 0$ because ℓ is weak*-continuous.

Now, suppose that $(c_n\phi_n)_{n=1}^{\infty}$ converges to 0. Then again we have that $c_n \to 0$ by the same argument used above (i.e. $c_n = (c_n\phi_n)(1)$). Therefore

$$
|\ell_1(c_n\phi_n)| = |c_n| \cdot |\ell(\phi_n)| \leq |c_n|\left(\sup_{\phi \in K}|\ell(\phi)|\right) \overset{n \to \infty}{\to} 0.
$$

We can thus conclude that ℓ_1 is weak*-continuous. $\qquad\square$

Lemma 6.1.21. *Let ℓ_1, P be as in Lemma 6.1.20, and let $V = P - P$. Then there exists a continuous linear functional $\tilde{\ell} : V \to \mathbb{R}$ such that $\tilde{\ell}|_P = \ell_1$.*

Proof. Define $\tilde{\ell} : V \to \mathbb{R}$ by

$$\tilde{\ell}(v) = \ell_1\left(v^+\right) - \ell_1\left(v^-\right),$$

where v^+, v^- are meant in the sense of the lattice structure V possesses by virtue of K being a simplex.

Our first claim is that if $f, g \in P$ such that $v = f - g$, then $\tilde{\ell}(v) = \ell_1(f) - \ell_1(g)$. To see this, we observe that $f + v^- = g + v^+ \in P$. Therefore

$$\ell_1\left(f + v^-\right) = \ell_1\left(g + v^+\right)$$

$$= \ell_1(f) + \ell_1\left(v^-\right) = \ell_1(g) + \ell_1\left(v^+\right)$$

$$\Rightarrow \ell_1(f) - \ell_1(g) = \ell_1\left(v^+\right) - \ell_1\left(v^-\right)$$

$$= \tilde{\ell}(v).$$

This makes linearity fairly straightforward to check. First, to confirm additivity, let $v, w \in V$. Then $v + w = (v^+ + w^+) - (v^- + w^-)$, where $v^+ + w^+, v^- + w^- \in P$. Thus

$$\tilde{\ell}(v + w) = \ell_1\left(v^+ + w^+\right) - \ell_1\left(v^- + w^-\right)$$

$$= \ell_1\left(v^+\right) + \ell_1\left(w^+\right) - \ell_1\left(v^-\right) - \ell_1\left(w^-\right)$$

$$= \ell_1\left(v^+\right) - \ell_1\left(v^-\right) + \ell_1\left(w^+\right) - \ell_1\left(w^-\right)$$

$$= \tilde{\ell}(v) + \tilde{\ell}(w).$$

To check homogeneity, let $c \in \mathbb{R}$. If $c \geq 0$, then $cv^+, cv^- \in P$, and $cv^+ - cv^- = cv$; on the other hand, if $c \leq 0$, then $-cv^-, -cv^+ \in P$, and $cv = -cv^- + cv^+$. In both cases, homogeneity is straightforward to show. This proves that $\tilde{\ell}$ is linear.

It is also quick to show that $\tilde{\ell}|_P = \ell_1$, since if $v \in P$, then $v = v^+$, so $\tilde{\ell}(v) = \ell_1\left(v^+\right) - 0 = \ell_1(v)$.

It remains now to show that $\tilde{\ell}$ is continuous. By (Rudin, 1991, Theorem 1.18), it will suffice to show that $\ker \tilde{\ell}$ is weak*-closed. To prove the kernel is closed, let $(v_n)_{n=1}^{\infty}$ be a sequence in $\ker \tilde{\ell}$ converging in the weak*-topology to $v \in V$. By the Uniform Boundedness Principle, it follows that $\sup_n \|v_n\| < \infty$. By rescaling, we can assume without loss of generality that $\|v_n\| \leq 1$ for all $n \in \mathbb{N}$, and since the unit ball $B \subseteq V$ is weak*-closed by Banach-Alaoglu, we can infer that $\|v\| \leq 1$.

Since the unit ball B is weak*-compact, it follows that the sequences $(v_n^+)_{n=1}^\infty$, $(v_n^-)_{n=1}^\infty$ have convergent subsequences. Let $(n_j)_{j=1}^\infty$ be a subsequence along which $v_{n_j}^+ \to m_1 \in P, v_{n_j}^- \to m_2 \in P$. Then if $x \in \mathfrak{R}$, then

$$
\begin{aligned}
v(x) &= \lim_{n\to\infty} v_n(x) \\
&= \lim_{n\to\infty} \left(v_n^+(x) - v_n^-(x) \right) \\
&= \lim_{j\to\infty} \left(v_{n_j}^+(x) - v_{n_j}^-(x) \right) \\
&= \left(\lim_{j\to\infty} v_{n_j}^+(x) \right) - \left(\lim_{j\to\infty} v_{n_j}^-(x) \right) \\
&= m_1(x) - m_2(x).
\end{aligned}
$$

Therefore $v = m_1 - m_2$, so

$$
\begin{aligned}
\tilde{\ell}(v) &= \tilde{\ell}(m_1) - \tilde{\ell}(m_2) \\
&= \left(\lim_{j\to\infty} \tilde{\ell}\left(v_{n_j}^+ \right) \right) - \left(\lim_{j\to\infty} \tilde{\ell}\left(v_{n_j}^- \right) \right) \\
&= \lim_{j\to\infty} \left(\tilde{\ell}\left(v_{n_j}^+ \right) - \tilde{\ell}\left(v_{n_j}^- \right) \right) \\
&= \lim_{j\to\infty} \tilde{\ell}\left(v_{n_j} \right) \\
&= \lim_{j\to\infty} 0 \\
&= 0.
\end{aligned}
$$

Therefore, we can conclude that $\tilde{\ell}$ is weak*-continuous. $\qquad\square$

Proof of Theorem 6.1.19. This follows from Lemmas 6.1.20 and 6.1.21. $\qquad\square$

Proof of Theorem 6.1.18. Let $F \subseteq K$ be a closed face of K. By Lemma 6.1.17, the face F is exposed, so let $\ell : K \to \mathbb{R}$ be a weak*-continuous affine functional such that

$$
\begin{aligned}
\ell(k) &= 0 & (\forall k \in F), \\
\ell(k) &< 0 & (\forall k \in K \setminus F).
\end{aligned}
$$

Set

$$V = \{c_1\phi_1 - c_2\phi_2 : c_1, c_2 \in \mathbb{R}_{\geq 0}; \phi_1, \phi_2 \in K\},$$

and let $\tilde{\ell} : V \to \mathbb{R}$ be a continuous linear extension of ℓ to V whose existence is promised by Theorem 6.1.19. We can then extend $\tilde{\ell} : V \to \mathbb{R}$ to a weak*-continuous linear functional $\ell' : \mathfrak{R}^\natural \to \mathbb{R}$ (Aliprantis and Burkinshaw, 2006, Theorem 3.6). There thus exists some $x \in \mathbb{R}$ such that $\ell'(\phi) = \phi(x)$ for all $\phi \in \mathfrak{R}^\natural$ (Baggett, 1992, Theorem 5.2). In particular, we have $\ell'(v) = v(x)$ for all $v \in V$. Therefore $F = K_{\max}(x)$.

The converse is contained in Proposition 6.1.15. $\qquad\square$

In particular, we can recover the following corollary.

Corollary 6.1.22. *If $\phi \in \partial_e K$, then there exists $x \in \mathfrak{R}$ such that ϕ is uniquely $(x|K)$-maximizing, i.e. such that $\{\phi\} = K_{\max}(x)$.*

Proof. The singleton $\{\phi\}$ is a closed face, and by Lemma 6.1.17 is therefore an exposed face. Apply Theorem 6.1.18. $\qquad\square$

We have developed the language of ergodic optimization here in a somewhat atypical way, where we speak not of x-maximizing states *simpliciter*, but of a state that is maximizing relative to a compact convex subset K of $\mathcal{S}^G$, especially a compact simplex K. This notion of relative ergodic optimization has precedent in (Zhao, 2016). For our purposes, this relative ergodic optimization means we can consider ergodic optimization problems over different *types* of states. In Section 6.4, we will broaden our scope somewhat to consider ergodic optimization in the noncommutative setting relative to a set of states that aren't necessarily Θ-invariant.

Since Theorem 6.1.18 applies in cases where K is a simplex, we will conclude this section by describing some situations where $\mathcal{S}^G$ is a compact metrizable simplex.

For each $\phi \in \mathcal{S}^G$, let $\pi_\phi : \mathfrak{A} \to \mathscr{B}(\mathscr{H}_\phi)$ be the GNS representation corresponding to ϕ. Define a unitary representation $u_\phi : G \to \mathbb{U}(\mathscr{H}_\phi)$ of G by

$$u_\phi(g)\pi_\phi(a) = \pi_\phi\left(\Theta_{g^{-1}}(a)\right),$$

extending this from $\pi_\phi(\mathfrak{A})$ to $\mathscr{H}_\phi$. Set

$$E_\phi = \{v \in \mathscr{H}_\phi : u_\phi(v) = v \text{ for all } g \in G\}.$$

Let $P_\phi : \mathscr{H}_\phi \twoheadrightarrow E_\phi$ be the orthogonal projection (in the functional-analytic sense) of $\mathscr{H}_\phi$ onto E_ϕ. We call the C*-dynamical system $(\mathfrak{A}, G, \Theta)$ a *G-abelian* system if for every $\phi \in \mathcal{S}^G$, the family of operators $\{P_\phi \pi_\phi(a) P_\phi \in \mathscr{B}(\mathscr{H}_\phi) : a \in \mathfrak{A}\}$ is mutually commutative.

We record here a handful of germane facts about G-abelian systems.

Proposition 6.1.23. *If $(\mathfrak{A}, G, \Theta)$ is G-abelian, then $\mathcal{S}^G$ is a simplex.*

Proof. See (Sakai, 2012, Theorem 3.1.14). □

Definition 6.1.24. We call a system $(\mathfrak{A}, G, \Theta)$ *asymptotically abelian* if there exists a sequence $(g_n)_{n=1}^\infty$ in G such that

$$[\Theta_{g_n} a, b] \overset{n \to \infty}{\Rightarrow} 0$$

for all $a, b \in \mathfrak{A}$, where $[\cdot, \cdot]$ is the Lie bracket $[x, y] = xy - yx$ on $\mathfrak{A}$.

Proposition 6.1.25. *If $(\mathfrak{A}, G, \Theta)$ is asymptotically abelian, then it is also G-abelian.*

Proof. See (Sakai, 2012, Proposition 3.1.16). □

6.2 Unique ergodicity and gauges: the singly generated setting

So far we have spoken about C*-dynamical systems, a noncommutative analog of a topological dynamical systems. But just as classical ergodic theory is often interested in the interplay between topological dynamical systems and the measure-theoretic dynamical systems they can be realized in, we are interested in questions about the interplay between C*-dynamical systems and the non-commutative measure-theoretic dynamical systems they can be realized in. To make this more precise, we introduce the notion of a W*-dynamical system.

A *W*-probability space* is a pair $(\mathfrak{M}, \rho)$ consisting of a von Neumann algebra $\mathfrak{M}$ and a faithful tracial normal state ρ on $\mathfrak{M}$. An *automorphism* of a W*-probability space $(\mathfrak{M}, \rho)$ is a *-automorphism $T : \mathfrak{M} \to \mathfrak{M}$ such that $\rho \circ T = \rho$, i.e. an automorphism of $\mathfrak{M}$ which preserves ρ. A *W*-dynamical system* is a quadruple

$(\mathfrak{M}, \rho, G, \Xi)$, where $(\mathfrak{M}, \rho)$ is a W*-probability space, and $\Xi : G \to \mathrm{Aut}(\mathfrak{M}, \rho)$ is a left action of a discrete topological group G (called the *phase group*) on $\mathfrak{M}$ by ρ-preserving automorphisms of $\mathfrak{M}$, i.e. such that $\rho(\Xi_g x) = \rho(x)$ for all $g \in G, x \in \mathfrak{M}$. Importantly, if $(\mathfrak{M}, \rho, G, \Xi)$ is a W*-dynamical system, then $(\mathfrak{M}, G, \Xi)$ is automatically a W*-dynamical system.

Remark 6.2.1. In the literature, the term "W*-dynamical system" is sometimes used to refer to a more general construction, where the group G is assumed to satisfy some topological conditions, and the action is assumed to be continuous in the strong operator topology, e.g. (Bannon et al., 2018). Other authors use a yet more general definition, e.g. (Blackadar, 2006, III.3.2). Since we are only interested in actions of discrete groups, we adopt a narrower definition.

Definition 6.2.2. Given a W*-probability space, we define $\mathcal{L}^2(\mathfrak{M}, \rho)$ to be the Hilbert space defined by completing $\mathfrak{M}$ with respect to the inner product $\langle x, y \rangle_\rho = \rho(y^* x)$, i.e. the Hilbert space associated with the faithful GNS representation of $\mathfrak{M}$ induced by ρ.

Finally, we introduce the notion of a C*-model, intending to generalize the notion of a topological model from classical ergodic theory to this noncommutative setting.

Definition 6.2.3. Let $(\mathfrak{M}, \rho, G, \Xi)$ be a W*-dynamical system. A *C*-model* of $(\mathfrak{M}, \rho, G, \Xi)$ is a quadruple $(\mathfrak{A}, G, \Theta; \iota)$ consisting of a C*-dynamical system $(\mathfrak{A}, G, \Theta)$ and a *-homomorphism $\iota : \mathfrak{A} \to \mathfrak{M}$ such that

(a) $\iota(\mathfrak{A})$ is dense in the weak operator topology of $\mathfrak{M}$,

(b) $\Xi_g(\iota(\mathfrak{A})) = \iota(\mathfrak{A})$ for all $g \in G$, and

(c) $\Xi_g \circ \iota = \iota \circ \Theta_g$ for all $g \in G$.

We call the C*-model $(\mathfrak{A}, G, \Theta; \iota)$ *faithful* if ι is also injective.

We remark that we can turn any C*-model into a faithful C*-model through a quotienting process. If ι was not injective, then we could instead consider $\tilde{\iota} : \mathfrak{A}/\ker\iota \hookrightarrow \mathfrak{M}$. In the case where $\mathfrak{A}$ is commutative, this quotienting process corresponds (via the Gelfand-Naimark Theorem) to taking a measure-theoretic dynamical system and restricting to the support of the resident probability measure. To see this, let $\mathfrak{A} = C(X)$, where X is a compact metrizable topological space, and let $\mathfrak{M} = L^\infty(X, \mu)$ for some Borel probability measure μ. Let $\iota : C(X) \to L^\infty(X, \mu)$ be the (not necessarily injective) map that maps a continuous function on X to its equivalence class in $L^\infty(X, \mu)$. It can be seen that $f \in \ker\iota$ if and only if the open set $\{x \in X : f(x) \neq 0\}$ is of measure 0, or equivalently if $f|_{\mathrm{supp}(\mu)} = 0$, and in particular that ι is injective if and only if μ is *strictly*

positive (i.e. μ assigns positive measure to all nonempty open sets). As such, we can identify $C(X)/\ker\iota$ with $C(\mathrm{supp}(\mu))$. Let $Y = \mathrm{supp}(\mu)$ denote the support of μ on X, and let $\pi : C(X) \twoheadrightarrow C(Y)$ be the quotient map (which corresponds to a restriction from X to Y, i.e. $\pi f = f|_Y$). Then algebraically, we have a commutative diagram

$$
\begin{array}{ccc}
C(X) & \xrightarrow{\ \pi\ } & C(Y) \\
& \searrow_{\iota} & \ \ \vdots\, \tilde{\iota} \\
& & L^\infty(X,\mu)
\end{array}
$$

So in the commutative case, we can make $\iota : C(X) \to L^\infty(X,\mu)$ injective by looking at $\tilde{\iota} : C(Y) \to L^\infty(Y,\mu) \cong L^\infty(X,\mu)$, i.e. by using the support Y to model $(Y,\mu) \cong (X,\mu)$.

Importantly, so long as $\mathcal{L}^2(\mathfrak{M},\rho)$ is separable, any W*-dynamical system $(\mathfrak{M}, \rho, G, \Xi)$ will admit a faithful separable C*-model. To construct such a C*-model, it suffices to take some separable C*-subalgebra $\mathfrak{B} \subseteq \mathfrak{M}$ which is dense in $\mathfrak{M}$ with respect to the weak operator topology, then let $\mathfrak{A}$ be the norm-closure of the span of $\bigcup_{g \in G} (\Xi_g \mathfrak{B})$. We then define $\Theta_g = \Xi_g|_\mathfrak{A}$ and let $\iota : \mathfrak{A} \hookrightarrow \mathfrak{M}$ be the inclusion map.

One last important concept in this section and the next will be unique ergodicity. A C*-dynamical system $(\mathfrak{A}, G, \Theta)$ is called *uniquely ergodic* if $\mathcal{S}^G$ is a singleton. As in the commutative setting, unique ergodicity can be equivalently characterized in terms of convergence properties of ergodic averages. To our knowledge, the strongest such characterization of unique ergodicity for singly generated C*-dynamical systems can be found in (Abadie and Dykema, 2009, Theorem 3.2), which describes unique ergodicity relative to the fixed point subalgebra. This characterization was then generalized to characterize unique ergodicity relative to the fixed point subalgebra for C*-dynamical systems over amenable phase groups in (Duvenhage and Stroh, 2011, Theorem 5.2); however, in Corollary 6.3.6, we provide a characterization of uniquely ergodic C*-dynamical systems in terms of ergodic averages that is not encompassed by (Duvenhage and Stroh, 2011, Theorem 5.2).

Given a C*-dynamical system $(\mathfrak{A}, \mathbb{Z}, \Theta)$, let $a \in \mathfrak{A}$ be a positive element. We define the *gauge* of a to be

$$
\Gamma(a) := \lim_{k \to \infty} \frac{1}{k} \left\| \sum_{j=0}^{k-1} \Theta_j a \right\|.
$$

To prove this limit exists, it suffices to observe that the sequence $\left(\left\| \sum_{j=0}^{k-1} \Theta_j a \right\| \right)_{k=1}^{\infty}$ is subadditive, since

$$\left\| \sum_{j=0}^{k+\ell-1} \Theta_j a \right\| \leq \left\| \sum_{j=0}^{k-1} \Theta_j a \right\| + \left\| \sum_{j=k}^{k+\ell-1} \Theta_j a \right\|$$

$$= \left\| \sum_{j=0}^{k-1} \Theta_j a \right\| + \left\| \Theta_k \sum_{j=0}^{\ell-1} \Theta_j a \right\|$$

$$= \left\| \sum_{j=0}^{k-1} \Theta_j a \right\| + \left\| \sum_{j=0}^{\ell-1} \Theta_j a \right\|.$$

Therefore, by the Subadditivity Lemma, the sequence $\left(\frac{1}{k} \left\| \sum_{j=0}^{k-1} \Theta_j a \right\| \right)_{k=1}^{\infty}$ converges, and we have the equality

$$\lim_{k \to \infty} \frac{1}{k} \left\| \sum_{j=0}^{k-1} \Theta_j a \right\| = \inf_{k \in \mathbb{N}} \frac{1}{k} \left\| \sum_{j=0}^{k-1} \Theta_j a \right\|.$$

We have the following characterization of Γ in the language of ergodic optimization.

Theorem 6.2.4. *Let* $(\mathfrak{A}, \mathbb{Z}, \Theta)$ *be a C*-dynamical system. Then if* $a \in \mathfrak{A}$ *is a positive element, then* $\Gamma(a) = m \left(a | \mathcal{S}^G \right)$.

Proof. For each $k \in \mathbb{N}$, choose a state σ_k on $\mathfrak{A}$ such that

$$\sigma_k \left(\frac{1}{k} \sum_{j=0}^{k-1} \Theta_j a \right) = \left\| \frac{1}{k} \sum_{j=0}^{k-1} \Theta_j a \right\|.$$

Let $\omega_k = \frac{1}{k} \sum_{j=0}^{k-1} \sigma_k \circ \Theta_j$, so

$$\omega_k(x) = \frac{1}{k} \sum_{j=0}^{k-1} \sigma_k \left(\Theta_j x \right)$$

$$= \sigma_k \left(\frac{1}{k} \sum_{j=0}^{k-1} \Theta_j x \right),$$

$$\omega_k(a) = \sigma_k \left(\frac{1}{k} \sum_{j=0}^{k-1} \Theta_j a \right)$$

$$= \left\| \frac{1}{k} \sum_{j=0}^{k-1} \Theta_j a \right\|.$$

Let $\omega \in \mathcal{S}$ be a weak*-limit point of $(\omega_k : k \in \mathbb{N})$, and let $k_1 < k_2 < \cdots$ be a subsequence such that $\omega_{k_n} \overset{n \to \infty}{\to} \omega$ in the weak*-topology. By Lemma 6.1.2, we know that ω is Θ-invariant. Therefore $\omega(a) = \Gamma(a)$, and ω is a Θ-invariant state on $\mathfrak{A}$, so

$$\Gamma(a) = \omega(a) \leq m\left(a|\mathcal{S}^{\mathbb{Z}}\right).$$

Now, we prove the opposite inequality. Let $\phi \in \mathcal{S}^{\mathbb{Z}}$. Then

$$\phi(a) = \phi\left(\mathrm{Avg}_k \, a\right)$$
$$\leq \|\mathrm{Avg}_k \, a\|$$
$$= \frac{1}{k}\left\|\sum_{j=0}^{k-1} \Theta_j a\right\|$$
$$= \frac{1}{k}\left\|\sum_{j=0}^{k-1} \Theta_j a\right\| \qquad (\forall k \in \mathbb{N})$$
$$\Rightarrow \phi(a) \leq \inf_{k \in \mathbb{N}} \frac{1}{k}\left\|\sum_{j=0}^{k-1} \Theta_j a\right\|$$
$$= \Gamma(a)$$
$$\Rightarrow \sup_{\psi \in \mathcal{S}^{\mathbb{Z}}} \psi(a) \leq \Gamma(a).$$

Therefore

$$m\left(a|\mathcal{S}^{\mathbb{Z}}\right) = \sup_{\psi \in \mathcal{S}^{\mathbb{Z}}} \psi(a) \leq \Gamma(a).$$

This establishes the identity. $\qquad \square$

Corollary 6.2.5. *Let $(\mathfrak{M}, \rho, \mathbb{Z}, \Xi)$ be a W*-dynamical system, and let $(\mathfrak{A}, \mathbb{Z}, \Theta; \iota)$ be a C*-model of $(\mathfrak{M}, \rho, \mathbb{Z}, \Xi)$. If $a \in \mathfrak{A}$ is a positive element, then*

$$\Gamma(\iota(a)) = m\left(a|\operatorname{Ann}(\ker \iota)\right).$$

Proof. Write $\tilde{\mathfrak{A}} = \iota(\mathfrak{A}) \subseteq \mathfrak{M}$, and let $\tilde{\Theta} : \mathbb{Z} \to \operatorname{Aut}\left(\tilde{\mathfrak{A}}\right)$ be the action $\tilde{\Theta}_n = \Xi_n|_{\tilde{\mathfrak{A}}}$ obtained by restricting Ξ to $\tilde{\mathfrak{A}}$. Write $\tilde{\mathcal{S}}^{\mathbb{Z}}$ for the space of $\tilde{\Theta}$-invariant states on $\tilde{\mathfrak{A}}$.

We can write $\Gamma_{\mathfrak{M}}(\iota(a)) = \Gamma_{\tilde{\mathfrak{A}}}(\iota(a))$. By Theorem 6.2.4, we know that $\Gamma_{\tilde{\mathfrak{A}}}(\iota(a)) = m\left(\iota(a)|\tilde{\mathcal{S}}^{\mathbb{Z}}\right)$, and by Theorem 6.1.16, we know that $m\left(\iota(a)|\tilde{\mathcal{S}}^{\mathbb{Z}}\right) = m\left(a|\operatorname{Ann}(\ker \iota)\right)$. $\qquad\square$

Remark 6.2.6. Corollary 6.2.5 can be regarded as an operator-algebraic extension of Lemma 2.3 from (Assani and Young, 2022). The assumption that $(\mathfrak{A}, G, \Theta; \iota)$ is faithful can be understood as analogous to the assumption of strict positivity in that paper.

This Γ value provides an alternative characterization of unique ergodicity, at least under some additional Choquet-theoretic hypotheses.

Theorem 6.2.7. *Let $(\mathfrak{M}, \rho, \mathbb{Z}, \Xi)$ be a W*-dynamical system, and let $(\mathfrak{A}, \mathbb{Z}, \Theta; \iota)$ be a faithful C*-model of $(\mathfrak{M}, \rho, \mathbb{Z}, \Xi)$. Then the following conditions are related by the implications (i) $\Longleftrightarrow$ (ii)$\Rightarrow$(iii).*

(i) The C*-dynamical system $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is uniquely ergodic.

(ii) The C*-dynamical system $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is strictly ergodic.

(iii) $\Gamma(\iota(a)) = \rho(\iota(a))$ for all positive $a \in \mathfrak{A}$.

Further, if $\mathcal{S}^{\mathbb{Z}}$ is a simplex, then (iii)$\Rightarrow$(i).

Proof. (i)$\Rightarrow$(ii) Suppose that $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is uniquely ergodic. Then $\rho \circ \iota$ is an invariant state on $\mathfrak{A}$, so it follows that $\rho \circ \iota$ is the unique invariant state on $\mathfrak{A}$. But $\rho \circ \iota$ is also a faithful state on $\mathfrak{A}$, so it follows that $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is strictly ergodic.

(ii)$\Rightarrow$(i) Trivial.

(i)$\Rightarrow$(iii) Suppose that $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is uniquely ergodic, and let $a \in \mathfrak{A}$ be positive. Let ϕ be a $\mathcal{S}^{\mathbb{Z}}$-maximizing state for a. Then $\phi = \rho \circ \iota$, since both ϕ and $\rho \circ \iota$ are invariant states on $\mathfrak{A}$, and $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is uniquely ergodic. Thus $\phi = \rho \circ \iota$, so $\Gamma(\iota(a)) = \phi(a) = \rho(\iota(a))$.

(iii)$\Rightarrow$(i) Suppose that $\mathcal{S}^{\mathbb{Z}}$ is a simplex, but that $(\mathfrak{A}, \mathbb{Z}, \Theta)$ is *not* uniquely ergodic. By the Krein-Milman Theorem, there exists two distinct extreme points of $\mathcal{S}^{\mathbb{Z}}$, and in particular there exists an extreme point $\phi \in \mathcal{S}^{\mathbb{Z}}$ of $\mathcal{S}^{\mathbb{Z}}$ distinct from $\rho \circ \iota$. Then by Corollary 6.1.22, there exists $a \in \mathfrak{A}$ self-adjoint such that $\{\phi\} = \mathcal{S}^{\mathbb{Z}}_{\max}(a)$. We can assume that a is positive, since otherwise we could replace a with $a + r$ for a sufficiently large positive real number $r > 0$, and $\mathcal{S}^{\mathbb{Z}}_{\max}(a) = \mathcal{S}^{\mathbb{Z}}_{\max}(a + r)$. Then $\Gamma(\iota(a)) = \phi(a)$. But by the assumption that ϕ is uniquely $(a|\mathcal{S}^{\mathbb{Z}})$-maximizing, it follows that $\rho(\iota(a)) < \phi(a)$. Therefore $\Gamma(\iota(a)) \neq \rho(\iota(a))$, meaning that (iii) does not attain. Thus $\neg$(i)$\Rightarrow \neg$(iii). $\qquad\square$

6.3 Unique ergodicity and gauge: the amenable setting

For the duration of this section, we assume that $(\mathfrak{M}, \rho, G, \Xi)$ is a W*-dynamical system with $\mathcal{L}^2(\mathfrak{M}, \rho)$ separable. Assume further that $(\mathfrak{A}, G, \Theta)$ is a C*-dynamical system such that $\mathfrak{A}$ is separable, and that G is amenable. It follows from Corollary 6.1.4 that $\mathcal{S}^G \neq \emptyset$.

In this section, we expand upon some of the ideas presented in Section 6.2, generalizing from the case of actions of $\mathbb{Z}$ to actions of a countable discrete amenable group G. We separate these two sections because our treatment of the more general amenable setting has some additional nuances to it.

Our first result of this section is a generalization of a classical result from ergodic theory regarding unique ergodicity, which is that a (singly generated) topological dynamical system is uniquely ergodic if and only if the averages of the continuous functions converge to a constant. This classical result is well-known, and can be found in many standard texts on ergodic theory, e.g. (Dajani and Dirksin, 2008, Thm 6.2.1), (Eisner et al., 2015, Thm 10.6), (Walters, 2007, Thm 5.17), but the earliest example of a result like this that we could find was (Oxtoby, 1952, 5.3). Theorem 6.3.1 generalizes this classical result not only to the noncommutative setting, but to the setting where the phase group G is amenable.

We define the *weak topology* on a C*-algebra $\mathfrak{A}$ to be the topology generated by the states on $\mathfrak{A}$, i.e.

$$x \mapsto \psi(x) \qquad\qquad (\psi \in \mathcal{S}).$$

In other words, the weak topology is the topology in which a net $(x_i)_{i \in \mathscr{I}}$ converges to x if and only if $(\psi(x_i))_{i \in \mathscr{I}}$ converges to $\psi(x)$ for every state ψ on $\mathfrak{A}$. We say the net $(x_i)_{i \in \mathscr{I}}$ *converges weakly* to x if it converges in the weak topology.

Theorem 6.3.1. *Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system. Then the following conditions are equivalent.*

(i) *$(\mathfrak{A}, G, \Theta)$ is uniquely ergodic.*

(ii) *There exists a right Følner sequences $(F_k)_{k=1}^\infty$ for G and a linear functional $\phi : \mathfrak{A} \to \mathbb{C}$ such that for all $x \in \mathfrak{A}$, the sequence $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^\infty$ converges in norm to $\phi(x)1 \in \mathbb{C}1$.*

(iii) *There exists a left Følner sequences $(F_k)_{k=1}^\infty$ for G and a linear functional $\phi : \mathfrak{A} \to \mathbb{C}$ such that for all $x \in \mathfrak{A}$, the sequence $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^\infty$ converges weakly to $\phi(x)1 \in \mathbb{C}1$.*

(iv) There exists a state ϕ on $\mathfrak{A}$ such that for every right Følner sequence $(F_k)_{k=1}^\infty$ for G, the sequence $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^\infty$ converges in norm to $\phi(x)1 \in \mathbb{C}1$.

(v) There exists a state ϕ on $\mathfrak{A}$ such that for every left Følner sequence $(F_k)_{k=1}^\infty$ for G, the sequence $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^\infty$ converges weakly to $\phi(x)1 \in \mathbb{C}1$.

Proof. Assume throughout that any $x \in \mathfrak{A}$ is nonzero.

(ii)$\Rightarrow$(iii) Follows from the existence of two-sided Følner sequence.

(iv)$\Rightarrow$(v) Follows from the existence of two-sided Følner sequence.

(iv)$\Rightarrow$(ii) Trivial.

(v)$\Rightarrow$(iii) Trivial.

(iii)$\Rightarrow$(i) Suppose that $\mathrm{Avg}_{F_k} x \to \phi(x)1 \in \mathbb{C}1$ weakly for all $x \in \mathfrak{A}$. We claim that ϕ is the unique invariant state of $(\mathfrak{A}, G, \Theta)$. First, we demonstrate that ϕ is Θ-invariant. Fix $g_0 \in G$, and fix $\epsilon > 0$. Choose $K_1, K_2, K_3 \in \mathbb{N}$ such that

$$
\begin{aligned}
k \geq K_1 &\quad\Rightarrow \left|\phi(\phi(x)1) - \phi\left(\mathrm{Avg}_{F_k} x\right)\right| < \frac{\epsilon}{3}, \\
k \geq K_2 &\quad\Rightarrow \left|\phi(\Theta_{g_0}\phi(x)1) - \phi(\Theta_{g_0}\mathrm{Avg}_{F_k} x)\right| < \frac{\epsilon}{3}, \\
k \geq K_3 &\quad\Rightarrow \frac{|g_0 F_k \Delta F_k|}{|F_k|} < \frac{\epsilon}{3\|x\|}.
\end{aligned}
$$

The K_1, K_2 exist because we know that in the weak topology, the functionals $\phi, \phi \circ \Theta_{g_0}$ are both continuous, and K_3 exists by the amenability of G. Let $K = \max\{K_1, K_2, K_3\}$. Then if $k \geq K$, then

$$
\begin{aligned}
|\phi(\Theta_{g_0} x) - \phi(x)| &\leq \left|\phi(\Theta_{g_0} x) - \phi(\Theta_{g_0} \operatorname{Avg}_{F_k} x)\right| \\
&\quad + \left|\phi(\Theta_{g_0} \operatorname{Avg}_{F_k} x) - \phi(\operatorname{Avg}_{F_k} x)\right| + \left|\phi(\operatorname{Avg}_{F_k} x) - \phi(x)\right| \\
&\leq \frac{\epsilon}{3} + \left|\phi(\Theta_{g_0} \operatorname{Avg}_{F_k} x) - \phi(\operatorname{Avg}_{F_k} x)\right| + \frac{\epsilon}{3} \\
&= \frac{2\epsilon}{3} + \left|\phi(\Theta_{g_0} \operatorname{Avg}_{F_k} x) - \phi(\operatorname{Avg}_{F_k} x)\right| \\
&= \frac{2\epsilon}{3} + \left|\phi\left(\frac{1}{|F_k|}\left(\sum_{g \in F_k} \Theta_{g_0 g} x\right) - \left(\frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g x\right)\right)\right| \\
&= \frac{2\epsilon}{3} + \left|\phi\left(\frac{1}{|F_k|}\left(\sum_{g \in g_0 F_k} \Theta_g x\right) - \left(\frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g x\right)\right)\right| \\
&= \frac{2\epsilon}{3} + \left|\phi\left(\frac{1}{|F_k|}\left(\sum_{g \in g_0 F_k \setminus F_k} \Theta_g x\right) - \left(\frac{1}{|F_k|} \sum_{g \in F_k \setminus g_0 F_k} \Theta_g x\right)\right)\right| \\
&\leq \frac{2\epsilon}{3} + \left|\phi\left(\frac{1}{|F_k|} \sum_{g \in g_0 F_k \setminus F_k} \Theta_g x\right)\right| + \left|\phi\left(\frac{1}{|F_k|} \sum_{g \in F_k \setminus g_0 F_k} \Theta_g x\right)\right| \\
&< \frac{2\epsilon}{3} + \frac{|g_0 F_k \Delta F_k|}{|F_k|} \|x\| \\
&= \epsilon.
\end{aligned}
$$

Therefore ϕ is Θ-invariant. To see that it is positive, it suffices to observe that $x \geq 0 \Rightarrow \operatorname{Avg}_{F_k} x \geq 0$, meaning that $\phi(x) = \lim_{k \to \infty} \phi(\operatorname{Avg}_{F_k} x) \geq 0$. To see that $\phi(1) = 1$, we just observe that $\operatorname{Avg}_{F_k} 1 = 1$ for all $k \in \mathbb{N}$.

Now we show that ϕ is the *unique* Θ-invariant state. Let ψ be any invariant state. Then

$$
\begin{aligned}
\psi(x) &= \psi(\operatorname{Avg}_{F_k} x) \\
&\overset{k \to \infty}{\Rightarrow} \psi(\phi(x)1) \\
&= \phi(x)\psi(1) \\
&= \phi(x).
\end{aligned}
$$

Therefore $\psi = \phi$, and so $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic.

(i)$\Rightarrow$(iv) Fix a right Følner sequence $(F_k)_{k=1}^\infty$, and assume for contradiction that $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic with Θ-invariant state ϕ, but that there exists $x \in \mathfrak{A}$ such that $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^\infty$ does *not* converge in norm to a scalar, and in particular does not converge in norm to $\phi(x)1$. Since we can decompose x into its real and imaginary parts, we can assume that $x \in \mathfrak{A}_{\mathrm{sa}}$. Fix $\epsilon_0 > 0$ for which there exists an infinite sequence $k_1 < k_2 < \cdots$ such that $\left\| \mathrm{Avg}_{F_{k_n}} x - \phi(x)1 \right\| \geq \epsilon_0$. Then for each $n \in \mathbb{N}$ exists a state ψ_n on $\mathfrak{A}$ such that $\left| \psi_n \left(\mathrm{Avg}_{F_{k_n}} x - \phi(x)1 \right) \right| = \left\| \mathrm{Avg}_{F_{k_n}} x - \phi(x)1 \right\|$.

Set

$$\omega_n = \psi_n \circ \mathrm{Avg}_{F_{k_n}},$$

so $\omega_n(x - \phi(x)1) = \psi_n \left(\mathrm{Avg}_{F_{k_n}} x - \phi(x)1 \right)$. Then $(\omega_n)_{n=1}^\infty$ has a subsequence, call it $(\omega_{n_j})_{j=1}^\infty$ which converges in the weak*-topology to some ω. This ω is also a state on $\mathfrak{A}$, and by Lemma 6.1.2, we know ω is Θ-invariant. But $\omega \neq \phi$, since

$$
\begin{aligned}
|\omega(x) - \phi(x)| &= \lim_{j \to \infty} \left| \omega_{n_j}(x) - \phi(x) \right| \\
&= \lim_{j \to \infty} \left| \omega_{n_j}(x - \phi(x)1) \right| \\
&= \lim_{j \to \infty} \left| \psi_{n_j}(\mathrm{Avg}_{F_{k_{n_j}}} x - \phi(x)1) \right| \\
&= \lim_{j \to \infty} \left\| \mathrm{Avg}_{F_{k_{n_j}}} x - \phi(x)1 \right\| \\
&\geq \epsilon_0.
\end{aligned}
$$

This contradicts $(\mathfrak{A}, G, \Theta)$ being uniquely ergodic. $\qquad \square$

Remark 6.3.2. Although (Duvenhage and Stroh, 2011, Theorem 5.2) describes conditions under which unique ergodicity of an action of an amenable group on a C*-algebra can be related to the convergence of ergodic averages, that result is not a direct generalization of our Theorem 6.3.1.

In order to develop the gauge machinery from the previous section in the context of actions of amenable groups, we will need to use slightly different techniques, since we do not have access to the Subadditivity Lemma. The main results of the remainder of this section can be summarized as follows.

Main results 6.3.3. *Let* $\mathbf{F} = (F_k)_{k=1}^\infty$ *be a right Følner sequence.*

(a) Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system, and let $\mathbf{F} = (F_k)_{k=1}^\infty$ be a right Følner sequence for G. Then if $a \in \mathfrak{A}$ is a positive element, then the sequence $\left(\left\| \frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right\| \right)_{k=1}^\infty$ converges to $m\left(a | \mathcal{S}^G\right)$.

(b) Let $(\mathfrak{A}, G, \Theta; \iota)$ be a faithful C*-model of $(\mathfrak{M}, \rho, G, \Xi)$. Then the following conditions are related by the implications (i) $\Longleftrightarrow$ (ii)$\Rightarrow$(iii).

   (i) The C*-dynamical system $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic.

   (ii) The C*-dynamical system $(\mathfrak{A}, G, \Theta)$ is strictly ergodic.

   (iii) $\Gamma(\iota(a)) = \rho(\iota(a))$ for all positive $a \in \mathfrak{A}$.

   Further, if $\mathcal{S}^G$ is a simplex, then (iii)$\Rightarrow$(i).

Theorem 6.3.4. Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system, and let $\mathbf{F} = (F_k)_{k=1}^\infty$ be a right Følner sequence for G. Then if $a \in \mathfrak{A}$ is a positive element, then the sequence $\left(\left\| \frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right\| \right)_{k=1}^\infty$ converges to $m\left(a | \mathcal{S}^G\right)$.

Proof. For each $k \in \mathbb{N}$, choose a state σ_k on $\mathfrak{A}$ such that

$$\sigma_k \left(\frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right) = \left\| \frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right\|.$$

Let $\omega_k = \frac{1}{|F_k|} \sum_{g \in F_k} \sigma_k \circ \Theta_g$, so

$$\omega_k(x) = \frac{1}{|F_k|} \sum_{g \in F_k} \sigma_k(\Theta_g x)$$

$$= \sigma_k \left(\frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g x \right),$$

$$\omega_k(a) = \sigma_k \left(\frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right)$$

$$= \left\| \frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right\|.$$

This means that in order to show that $\left(\left\| \frac{1}{|F_k|} \sum_{g \in G} \Theta_g a \right\| \right)_{k=1}^\infty$ converges to $m\left(a | \mathcal{S}^G\right)$, it suffices to show that $\omega_k(a) \overset{k \to \infty}{\to} m\left(a | \mathcal{S}^G\right)$. So for the remainder of this proof, we are going to be looking instead at the sequence $(\omega_k)_{k=1}^\infty$.

Let $k_1 < k_2 < \cdots$ be some sequence such that $(\omega_{k_n})_{n=1}^{\infty}$ converges in the weak*-topology to some ω. It follows from Lemma 6.1.2 that ω is Θ-invariant. To see that $(\omega_k(a))_{k=1}^{\infty} = \left(\left\| \frac{1}{|F_k|} \sum_{g \in G} \Theta_g \iota(a) \right\| \right)_{k=1}^{\infty}$ converges to $m\left(a|\mathcal{S}^G\right)$, it will suffice to show that every limit point ω of $(\omega_k : k \in \mathbb{N})$ satisfies

$$\omega \in \mathcal{S}_{\max}^G(a).$$

This follows because if there existed a subsequence $k_1 < k_2 < \cdots$ of $(\omega_k)_{k=1}^{\infty}$ such that $\omega_{k_n}(a) \overset{n \to \infty}{\to} z \neq m\left(a|\mathcal{S}^G\right)$, then by compactness, that subsequence $(\omega_{k_n} : n \in \mathbb{N})$ would have some subsequence converging to some ω' for which $\omega'(a) = z \neq m\left(a|\mathcal{S}^G\right)$, meaning in particular that $\omega' \notin \mathcal{S}_{\max}^G(a)$.

So let $k_1 < k_2 < \cdots$ be some sequence such that $(\omega_{k_n})_{n=1}^{\infty}$ converges in the weak*-topology to some ω. As has already been remarked, we have that $\omega \in \mathcal{S}^G$, so $\omega(a) \leq m\left(a|\mathcal{S}^G\right)$. We prove the opposite inequality. Let $\phi \in \mathcal{S}^G$. Then

$$\phi(a) = \phi\left(\frac{1}{|F_{k_n}|} \sum_{g \in F_{k_n}} \Theta_g a \right) \qquad (\phi \text{ is } \Theta\text{-invariant})$$

$$\leq \left\| \frac{1}{|F_{k_n}|} \sum_{g \in F_{k_n}} \Theta_g a \right\|$$

$$= \omega_{k_n}(a)$$

$$\Rightarrow \phi(a) \leq \lim_{n \to \infty} \omega_{k_n}(a)$$

$$= \omega(a).$$

Therefore $\omega(a) \geq \sup_{\psi \in \mathcal{S}^G} \psi(a) = m\left(a|\mathcal{S}^G\right)$. This establishes the desired identity. $\qquad \square$

Remark 6.3.5. An alternate proof of Theorem 6.3.4 using nonstandard analysis is presented in Section 6.5.

Corollary 6.3.6. *Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system, and let $\mathbf{F} = (F_k)_{k=1}^{\infty}$ be a right Følner sequence for G. Let $\phi \in \mathcal{S}^G$. Then $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic if and only if for every positive element $a \in \mathfrak{A}$, the sequence $\left(\left\| \frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right\| \right)_{k=1}^{\infty}$ converges to $\phi(a)$.*

Proof. ($\Rightarrow$) Suppose $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic. Then $\phi(a) = m\left(a|\mathcal{S}^G\right)$ for all positive $a \in \mathfrak{A}$, so by Theorem 6.3.4 $\left\| \frac{1}{|F_k|} \sum_{g \in F_k} \Theta_g a \right\| \overset{k \to \infty}{\to} \phi(a)$.

($\Longleftarrow$) We'll prove the contrapositive. Suppose $(\mathfrak{A}, G, \Theta)$ is *not* uniquely ergodic. Then there exists an extreme point ψ of $\mathcal{S}^G$ different from ϕ. By Corollary 6.1.22, there exists $a \in \mathfrak{A}$ self-adjoint such that $\{\phi\} = \mathcal{S}^G_{\max}(a)$. We can assume that a is positive, replacing a by $a + r$ for a sufficiently large positive real number $r > 0$ otherwise. Thus $\lim_{k\to\infty} \left\| \frac{1}{|F_k|} \sum_{g\in F_k} \Theta_g a \right\| = \psi(a) > \phi(a)$. $\qquad\square$

Definition 6.3.7. Given a C*-dynamical system $(\mathfrak{A}, G, \Theta)$, a positive element $a \in \mathfrak{A}$, and a right Følner sequence $\mathbf{F} = (F_k)_{k=1}^\infty$ for G, we define the *gauge* of a to be the limit

$$\Gamma(x) := \lim_{k\to\infty} \left\| \frac{1}{|F_k|} \sum_{g\in F_k} \Theta_g x \right\|.$$

Theorem 6.3.4 shows that the gauge exists, but Theorem 6.3.9 demonstrates the way that the gauge interacts with a W*-dynamical system and a C*-model. Moreover, the gauge is dependent only on $(\mathfrak{A}, G, \Theta)$, and independent of the right Følner sequence $\mathbf{F} = (F_k)_{k=1}^\infty$. As such, even though the gauge as we have described it is computed using a right Følner sequence $\mathbf{F} = (F_k)_{k=1}^\infty$, we do not need to include $\mathbf{F}$ in our notation for Γ.

Corollary 6.3.8. *Let* $(\mathfrak{A}, G, \Theta), \left(\tilde{\mathfrak{A}}, G, \tilde{\Theta}\right)$ *be two C*-dynamical systems, and let* $\pi : \mathfrak{A} \to \tilde{\mathfrak{A}}$ *be a *-homomorphism (not necessarily surjective) such that*

$$\tilde{\Theta}_g \circ \pi = \pi \circ \Theta_g \qquad\qquad (\forall g \in G).$$

Let $\tilde{\mathcal{S}}^G$ *denote the space of* $\tilde{\Theta}$-*invariant states on* $\tilde{\Theta}$. *Then* $m\left(\pi(a)|\tilde{\mathcal{S}}^G\right) = m\left(a|\operatorname{Ann}(\ker \pi)\right)$.

Proof. Let $\mathfrak{B} = \pi(\mathfrak{A})$, and let $H : G \to \operatorname{Aut}(\mathfrak{B})$ be the action $H_g = \tilde{\Theta}_g|_\mathfrak{B}$. Let K denote the space of all H-invariant states on $\mathfrak{B}$. Then

$$
\begin{aligned}
m\left(\pi(a)|\tilde{\mathcal{S}}^G\right) &= \Gamma_{\tilde{\mathfrak{A}}}(\pi(a)) && \text{(Theorem 6.3.4)} \\
&= \Gamma_\mathfrak{B}(\pi(a)) && \\
&= m\left(\pi(a)|K\right) && \text{(Theorem 6.3.4)} \\
&= m\left(a|\operatorname{Ann}(\ker \pi)\right) && \text{(Theorem 6.1.16)}.
\end{aligned}
$$

$\qquad\square$

Corollary 6.3.9. *Let* $(\mathfrak{M}, \rho, G, \Xi)$ *be a W*-dynamical system, and let* $(\mathfrak{A}, G, \Theta; \iota)$ *be a C*-model of* $(\mathfrak{M}, \rho, \mathbb{Z}, \Xi)$. *Then if* $a \in \mathfrak{A}$ *is a positive element, then*

$$\Gamma(\iota(a)) = m\left(a | \operatorname{Ann}(\ker \iota)\right).$$

Proof. Write $\tilde{\mathfrak{A}} = \iota(\mathfrak{A}) \subseteq \mathfrak{M}$, and let $\tilde{\Theta} : G \to \operatorname{Aut}\left(\tilde{\mathfrak{A}}\right)$ be the action $\tilde{\Theta}_g = \Xi_g|_{\tilde{\mathfrak{A}}}$ obtained by restricting Ξ to $\tilde{\mathfrak{A}}$. Write $\tilde{\mathcal{S}}^G$ for the space of $\tilde{\Theta}$-invariant states on $\tilde{\mathfrak{A}}$.

We know $\Gamma_{\mathfrak{M}}(\iota(a)) = \Gamma_{\tilde{\mathfrak{A}}}(\iota(a))$. By Theorem 6.3.4, we know that $\Gamma_{\tilde{\mathfrak{A}}}(\iota(a)) = m\left(\iota(a)|\tilde{\mathcal{S}}^G\right)$, and by Theorem 6.1.16, we know that

$$m\left(\iota(a)|\tilde{\mathcal{S}}^G\right) = m\left(a | \operatorname{Ann}(\ker \iota)\right).$$

$\square$

This brings us to our characterization of unique ergodicity with respect to the gauge for C*-models.

Theorem 6.3.10. *Let* $(\mathfrak{M}, \rho, G, \Xi)$ *be a W*-dynamical system, and let* $(\mathfrak{A}, G, \Theta; \iota)$ *be a faithful C*-model of* $(\mathfrak{M}, \rho, G, \Xi)$. *Then the following conditions are related by the implications (i)* $\Longleftrightarrow$ *(ii)* $\Rightarrow$ *(iii).*

(i) *The C*-dynamical system* $(\mathfrak{A}, G, \Theta)$ *is uniquely ergodic.*

(ii) *The C*-dynamical system* $(\mathfrak{A}, G, \Theta)$ *is strictly ergodic.*

(iii) $\Gamma(a) = \rho(\iota(a))$ *for all positive* $a \in \mathfrak{A}$.

Further, if $\mathcal{S}^G$ *is a simplex, then (iii)* $\Rightarrow$ *(i).*

Proof. (i)$\Rightarrow$(ii) Suppose that $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic. Then $\rho \circ \iota$ is an invariant state on $\mathfrak{A}$, so it follows that $\rho \circ \iota$ is *the* unique invariant state. But $\rho \circ \iota$ is also faithful, so it follows that $(\mathfrak{A}, G, \Theta)$ is strictly ergodic.

(ii)$\Rightarrow$(i) Trivial.

(i)$\Rightarrow$(iii) Suppose that $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic, and let $a \in \mathfrak{A}$ be positive. Let ϕ be an $\left(a|\mathcal{S}^G\right)$-maximizing state on $\mathfrak{A}$. Then $\phi = \rho \circ \iota$, since both are invariant states and $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic. Then $\phi = \rho \circ \iota$, so $\Gamma(a) = \phi(a) = \rho(\iota(a))$.

(iii)$\Rightarrow$(i) Suppose that $\mathcal{S}^G$ is a simplex, but that $(\mathfrak{A}, G, \Theta)$ is *not* uniquely ergodic. Let $\phi \in \mathcal{S}^G$ be an extreme point of $\mathcal{S}^G$ different from $\rho \circ \iota$. Then by Corollary 6.1.22, there exists $a \in \mathfrak{A}$ self-adjoint such

that $\{\phi\} = \mathcal{S}^G_{\max}(a)$. We can assume that a is positive, since otherwise we could replace a with $a + r$ for a sufficiently large positive real number $r > 0$, and $\mathcal{S}^{\mathbb{Z}}_{\max}(a) = \mathcal{S}^{\mathbb{Z}}_{\max}(a + r)$. Then $\Gamma(a) = \phi(a)$. But by the assumption that ϕ is uniquely $(a|\mathcal{S}^G)$-maximizing, it follows that $\rho(\iota(a)) < \phi(a)$. Therefore $\Gamma(a) \neq \rho(\iota(a))$, meaning that (iii) does not attain. Thus ¬(i)⇒ ¬(iii). $\qquad\square$

6.4 A noncommutative Herman ergodic theorem

For the duration of this section, we assume that $(\mathfrak{A}, G, \Theta)$ is a C*-dynamical system such that $\mathfrak{A}$ is separable, and that G is amenable.

Let $\mathbf{F} = (F_k)_{k=1}^\infty$ be a right Følner sequence for G. Write $\mathscr{P}^{\mathbf{F}}(S)$ to denote the set of all limit points of sequences of the form $\left(\phi_k \circ \mathrm{Avg}_{F_k}\right)_{k=1}^\infty$, where $\phi_k \in S$ for all $k \in \mathbb{N}$. Because $\mathbf{F}$ is right Følner, we know from Lemma 6.1.2 that if S is nonempty, then $\mathscr{P}^{\mathbf{F}}(S)$ will be a nonempty compact subset of $\mathcal{S}^G$. In particular, if $S \supseteq \mathcal{S}^G$, then $\mathscr{P}^{\mathbf{F}}(S) = \mathcal{S}^G$ for any choice of $\mathbf{F}$. Moreover, if S is convex and Θ-invariant, then $\mathscr{P}^{\mathbf{F}}(S) = \overline{S}$.

Question 6.4.1. Is $\mathscr{P}^{\mathbf{F}}(S)$ dependent on $\mathbf{F}$ in general?

We now define two quantities.

Notation 6.4.2. Let $\mathbf{F}$ be a right Følner sequence for G, and S a nonempty subset of $\mathcal{S}$. Let $x \in \mathfrak{R}$. Define

$$\overline{a}_{\mathbf{F},S}(x) := \sup\left\{\psi(x) : \psi \in \mathscr{P}^{\mathbf{F}}(x)\right\},$$

$$\underline{a}_{\mathbf{F},S}(x) := \inf\left\{\psi(x) : \psi \in \mathscr{P}^{\mathbf{F}}(x)\right\},$$

$$\overline{d}_{\mathbf{F},S}(x) := \lim_{k\to\infty}\left(\sup\left\{\phi\left(\mathrm{Avg}_{F_k} x\right) : \phi \in S\right\}\right),$$

$$\underline{d}_{\mathbf{F},S}(x) := \lim_{k\to\infty}\left(\inf\left\{\phi\left(\mathrm{Avg}_{F_k} x\right) : \phi \in S\right\}\right).$$

The values $\overline{a}_{\mathbf{F},S}, \overline{d}_{\mathbf{F},S}$ can be compared to the α and δ quantities presented in Section 2 of (Jenkinson, 2006a), respectively. Ergodic optimization is concerned with finding the extrema of sequences of ergodic averages of real-valued functions, but there are several ways we might attempt to formalize what an "extremum" of a sequence of ergodic averages would be. In (Jenkinson, 2006a), O. Jenkinson proposes several different ways we might formalize this notion, then demonstrates that they are equivalent under reasonable conditions (Jenkinson, 2006a, Proposition 2.1). Our Proposition 6.4.3 is an attempt to extend some part of this result to the noncommutative and relative setting.

Proposition 6.4.3. *The quantities $\overline{d}_{\mathbf{F},S}(x), \underline{d}_{\mathbf{F},S}(x)$ are well-defined when $S \subseteq \mathcal{S}^G$ is compact, convex, and Θ-invariant. Moreover, they satisfy*

$$\overline{a}_{\mathbf{F},S}(x) = \overline{d}_{\mathbf{F},S}(x), \qquad\qquad \underline{a}_{\mathbf{F},S}(x) = \underline{d}_{\mathbf{F},S}(x).$$

Proof. We'll prove that $\overline{a}_{\mathbf{F},S}(x) = \overline{d}_{\mathbf{F},S}(x)$, as the proof that $\underline{a}_{\mathbf{F},S}(x) = \underline{d}_{\mathbf{F},S}(x)$ is very similar. We know a priori that $\mathscr{P}^{\mathbf{F}}(S) = \overline{S}$.

Let $(\phi_k)_{k=1}^{\infty}$ be a sequence in S such that for each $k \in \mathbb{N}$, we have

$$\sup \left\{ \phi\left(\mathrm{Avg}_{F_k} x \right) : \phi \in S \right\} - 1/k \leq \phi_k(\mathrm{Avg}_{F_k} x) \leq \sup \left\{ \phi\left(\mathrm{Avg}_{F_k} x \right) : \phi \in S \right\}.$$

We know that any limit point of $\left(\phi_k \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty}$ is in $\overline{S}$. Let $k_1 < k_2 < \cdots$ be chosen such that $\lim_{\ell \to \infty} \phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x \right) = \limsup_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right)$. We can assume that $\left(\phi_{k_\ell} \circ \mathrm{Avg}_{F_{k_\ell}} \right)_{\ell=1}^{\infty}$ is weak*-convergent to a state $\psi \in \overline{S}$, passing to a subsequence if necessary. Then

$$\limsup_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right) = \lim_{\ell \to \infty} \phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x \right) = \psi(x) \leq \overline{a}_{\mathbf{F},S}(x).$$

Assume for contradiction that $\liminf_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right) < \overline{a}_{\mathbf{F},S}(x)$. Let $\psi' \in \mathscr{P}^{\mathbf{F}}(S)$ be such that $\psi'(x) > \liminf_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right)$. Then

$$\psi'(x) = \psi' \left(\mathrm{Avg}_{F_k} x \right) \qquad\qquad \leq \phi_k \left(\mathrm{Avg}_{F_k} x \right) - 1/k.$$

Let $k_1' < k_2' < \cdots$ such that $\left(\phi_{k_\ell'} \left(\mathrm{Avg}_{F_{k_\ell'}} x \right) \right)_{\ell=1}^{\infty}$ converges to $\liminf_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right)$. Then

$$\psi'(x) \leq \lim_{\ell \to \infty} \phi_{k_\ell'} \left(\mathrm{Avg}_{F_{k_\ell'}} x \right) \qquad\qquad = \liminf_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right) < \psi(x),$$

a contradiction. Therefore we conclude that $\liminf_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right) \geq \overline{a}_{\mathbf{F},S}(x)$. Thus

$$\overline{a}_{\mathbf{F},S}(x) \leq \liminf_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right) \leq \limsup_{k \to \infty} \phi_k \left(\mathrm{Avg}_{F_k} x \right) \leq \overline{a}_{\mathbf{F},S}(x).$$

Thus we can conclude that $\overline{d}_{\mathbf{F},S}(x)$ is well-defined and equal to $\overline{d}_{\mathbf{F},S}(x)$. $\qquad \square$

Remark 6.4.4. An alternate proof of Proposition 6.4.3 using nonstandard analysis is presented in Section 6.5.

To our knowledge, the first result like Theorem 6.4.5 is (Herman, 1983, Lemme on pg. 487). Herman's result can be understood as an extension of the classical result that a topological dynamical system is uniquely ergodic if and only if the ergodic averages of all continuous functions converge uniformly to a constant. To our knowledge, the first record of this classical result is (Oxtoby, 1952, (5.3)). If Oxtoby's result can be understood as relating the uniform convergence properties of ergodic averages of *all* continuous functions to the ergodic optimization of all continuous functions, then Herman's result relates the uniform convergence properties of ergodic averages of a *single* continuous function to its ergodic optimization. Our result extends Herman's in a few directions. First, it extends Herman's result to the setting of actions of amenable groups other than $\mathbb{Z}$. Moreover, it extends the result to C*-dynamical systems. Finally, it allows us to relate convergence in certain seminorms to relative ergodic optimizations.

Let $(\mathfrak{A}, G, \Theta)$ be a C*-dynamical system, where G is an amenable group. Given a nonempty subset S of $\mathcal{S}$, define the seminorm $\|\cdot\|_S$ on $\mathfrak{A}$ by

$$\|x\|_S := \sup_{\phi \in S} |\psi(x)|.$$

Theorem 6.4.5. *Let* $\mathbf{F}$ *be a right Følner sequence for* G, *and* $S \subseteq \mathcal{S}$. *Let* $x \in \mathfrak{R}$, *and* $\lambda \in \mathbb{R}$. *Then the following are equivalent.*

(i) $\{\psi(x) : \psi \in \mathscr{P}^{\mathbf{F}}(S)\} = \{\lambda\}$.

(ii) $\lim_{k \to \infty} \left\|\mathrm{Avg}_{F_k} x - \lambda\right\|_S = 0.$

Proof. (i)$\Rightarrow$(ii): We prove the contrapositive. Suppose there exists $\epsilon_0 > 0$ and $k_1 < k_2 < \cdots$ such that

$$\left\|\mathrm{Avg}_{F_{k_\ell}} x - \lambda\right\|_S > \epsilon_0 \qquad\qquad (\forall \ell \in \mathbb{N}).$$

For each $k \in \mathbb{N}$, choose $\phi_k \in S$ such that $\left|\phi_k \left(\mathrm{Avg}_{F_k} x - \lambda\right)\right| \geq \frac{1}{2} \left\|\mathrm{Avg}_{F_k} x - \lambda\right\|_S$. Then in particular we know that

$$\left|\phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x - \lambda\right)\right| > \epsilon_0/2 \qquad\qquad (\forall \ell \in \mathbb{N}).$$

218

By the weak*-compactness of $\mathcal{S}$, there must exist a weak*-convergent subsequence of $\left(\phi_{k_\ell} \circ \mathrm{Avg}_{k_\ell}\right)_{\ell=1}^\infty$. Assume without loss of generality that $\left(\phi_{k_\ell} \circ \mathrm{Avg}_{k_\ell}\right)_{\ell=1}^\infty$ converges in the weak* topology, and write $\psi = \lim_{\ell\to\infty} \phi_{k_\ell} \circ \mathrm{Avg}_{k_\ell}$. Then

$$
\begin{aligned}
|\psi(x - \lambda)| &= \left| \lim_{\ell\to\infty} \phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x - \lambda \right) \right| \\
&= \lim_{\ell\to\infty} \left| \phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x - \lambda \right) \right| \\
&\geq \epsilon_0 / 2.
\end{aligned}
$$

Therefore $\psi(x) \neq \lambda$, meaning that $\left\{ \psi(x) : \psi \in \mathscr{P}^{\mathbf{F}}(S) \right\} \neq \{\lambda\}$.

(ii)$\Rightarrow$(i): Suppose that $\lim_{k\to\infty} \left\| \mathrm{Avg}_{F_k} x - \lambda \right\|_S = 0$. Let $(\phi_k)_{k=1}^\infty$ be a sequence in S, and let $\left(\phi_{k_\ell} \circ \mathrm{Avg}_{F_{k_\ell}}\right)_{\ell=1}^\infty$ be a weak*-convergent subsequence of $\left(\phi_k \circ \mathrm{Avg}_{F_k}\right)_{k=1}^\infty$ with limit ψ. Then

$$
\begin{aligned}
|\psi(x - \lambda)| &= \left| \lim_{\ell\to\infty} \phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x - \lambda \right) \right| \\
&= \lim_{\ell\to\infty} \left| \phi_{k_\ell} \left(\mathrm{Avg}_{F_{k_\ell}} x - \lambda \right) \right| \\
&\leq \limsup_{\ell\to\infty} \left\| \mathrm{Avg}_{F_k} x - \lambda \right\|_S \\
&= 0.
\end{aligned}
$$

Therefore $\left\{ \psi(x) : \psi \in \mathscr{P}^{\mathbf{F}}(S) \right\} = \{\lambda\}$. $\qquad\square$

Remark 6.4.6. An alternate proof of Theorem 6.4.5 using nonstandard analysis is presented in Section 6.5.

Corollary 6.4.7. *Let $\mathbf{F}$ be a right Følner sequence for G. Let $x \in \mathfrak{R}$, and $\lambda \in \mathbb{R}$. Then the following are equivalent.*

(i) $\left\{ \psi(x) : \psi \in \mathcal{S}^G \right\} = \{\lambda\}$.

(ii) $\lim_{k\to\infty} \left\| \mathrm{Avg}_{F_k} x - \lambda \right\| = 0$.

Proof. Apply Theorem 6.4.5 in the case where $S = \mathcal{S}$, implying that $\| \cdot \|_S = \| \cdot \|$ and $\mathscr{P}^{\mathbf{F}}(S) = \mathcal{S}^G$. $\quad\square$

Corollary 6.4.7 strengthens the noncommutative analogue of Oxtoby's characterization of unique ergodicity, as we see below.

Corollary 6.4.8 (A noncommutative extension of Oxtoby's characterization of unique ergodicity). *Let* $\mathbf{F} = (F_k)_{k=1}^{\infty}$ *be a right Følner sequence for* G. *A C*-dynamical system* $(\mathfrak{A}, G, \Theta)$ *is uniquely ergodic if and only if* $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^{\infty}$ *converges in norm to an element of* $\mathbb{C}1 \subseteq \mathfrak{A}$ *for all* $x \in \mathfrak{A}$.

Proof. ($\Rightarrow$): By taking real and imaginary parts, we can reduce to the case where x is self-adjoint. If $(\mathfrak{A}, G, \Theta)$ is uniquely ergodic, then $\left\{\psi(x) : \psi \in \mathcal{S}^G\right\}$ is singleton, so by Corollary 6.4.7 the averages will converge to a scalar.

($\Leftarrow$): Conversely, if $(\mathfrak{A}, G, \Theta)$ is *not* uniquely ergodic, then there exist two states $\psi_1, \psi_2 \in \mathcal{S}^G$ for which there exists $y \in \mathfrak{R}$ such that $\psi_1(y) \neq \psi_2(y)$, implying that $\left\{\psi(y) : \psi \in \mathcal{S}^G\right\}$ is not singleton. Corollary 6.4.7 then tells us that $\left(\mathrm{Avg}_{F_k} x\right)_{k=1}^{\infty}$ doesn't converge in norm. $\square$

6.5   Applications of nonstandard analysis to noncommutative ergodic optimization

The tools of nonstandard analysis can be used to provide alternate proofs of some results in this chapter. In this section, we assume that the reader is familiar with the basic tools and vocabulary of nonstandard analysis. See (Goldblatt, 2012) for references. Since some of the terminology of the field is not entirely universal, we define some of the less universal terms here.

We will assume throughout this section that $(\mathfrak{A}, G, \Theta)$ is a C*-dynamical system, and that $\mathfrak{U}$ is a universe that contains $\mathfrak{A}, G, \mathbb{C}$. Assume that $* : \mathfrak{U} \mapsto \mathfrak{U}'$ is a countably saturated universe embedding. We say that $x \in {}^*\mathbb{C}$ is *unlimited* if $|x| > n$ for all $n \in \mathbb{N}$, and *limited* otherwise. Let $\mathbb{L} = \bigcup_{n \in \mathbb{N}} \{z \in {}^*\mathbb{C} : \|z\| \leq n\}$ denote the external ring of limited elements of ${}^*\mathbb{C}$. For $z, w \in {}^*\mathbb{C}$, we write $z \simeq w$ if $|z - w| < 1/n$ for all $n \in \mathbb{N}$. This $\simeq$ is an equivalence relation on ${}^*\mathbb{C}$. We define the *shadow* $\mathrm{sh} : \mathbb{L} \twoheadrightarrow \mathbb{C}$ to be the $\mathbb{C}$-linear functional mapping $z \in \mathbb{L}$ to the unique (standard) complex number $w \in \mathbb{C}$ for which $z \simeq w$. The shadow is also order-preserving on $\mathbb{L} \cap {}^*\mathbb{R}$. Let ${}^*\mathbb{N}_{\infty} := \{K \in {}^*\mathbb{N} : \forall k \in \mathbb{N} \, (K \geq k)\} = {}^*\mathbb{N} \setminus \mathbb{N}$ denote the unlimited hypernaturals.

We have the following nonstandard analogue of Lemma 6.1.2.

Lemma 6.5.1. *Let* $(\mathfrak{A}, G, \Theta)$ *be a C*-dynamical system, and let* G *be an amenable group. Consider a sequence in* $(\phi_k)_{k=1}^{\infty}$ *in* $\mathcal{S}$, *and a right Følner sequence* $\mathbf{F} = (F_k)_{k=1}^{\infty}$ *for* G. *Let* $K \in {}^*\mathbb{N}_{\infty}$ *be an unlimited hypernatural, and define a state* $\omega : \mathfrak{A} \to \mathbb{C}$ *by*

$$\omega(x) = \mathrm{sh}\left({}^*\phi_K\left(\mathrm{Avg}_{F_K} x\right)\right).$$

Then ω is a well-defined Θ-invariant state, and is a limit point of the sequence $\left(\phi_k \circ \mathrm{Avg}_{F_k}\right)_{k=1}^\infty$.

Proof. First, we take up the well-definedness of ω. If $x \in \mathfrak{A}$, then

$$\forall k \in \mathbb{N} \ \left(\left|\phi_k\left(\mathrm{Avg}_{F_k} x\right)\right| \leq \|x\|\right),$$

and so by the Transfer Principle

$$\forall k \in {}^*\mathbb{N} \ \left(\left|{}^*\phi_k\left(\mathrm{Avg}_{F_k} x\right)\right| \leq \|x\|\right).$$

In particular, it follows that $\left|{}^*\phi_K\left(\mathrm{Avg}_{F_K} x\right)\right| \leq \|x\|$, meaning that ${}^*\phi_K\left(\mathrm{Avg}_{F_K} x\right) \in \mathbb{L}$. Thus $\omega(x)$ is well-defined. We can similarly prove that ω is positive and unital by applying the Transfer Principle to the sentences

$$\forall k \in \mathbb{N} \ \forall x \in \mathfrak{A} \ \left(\phi_k\left(\mathrm{Avg}_{F_k} (x^*x)\right) \geq 0\right),$$

$$\forall k \in \mathbb{N} \ \left(\phi_k\left(\mathrm{Avg}_{F_k} 1\right) = 1\right).$$

To prove the Θ-invariance of ω, we recall from a familiar argument (see proof of Lemma 6.1.2) that if $g_0 \in G, x \in \mathfrak{A}$, then

$$\left|\phi_k\left(\mathrm{Avg}_{F_k} \Theta_{g_0} x\right) - \phi_k\left(\mathrm{Avg}_{F_k} x\right)\right| \leq \frac{|F_k g_0 \Delta F_k|}{|F_k|} \|x\| \overset{k \to \infty}{\to} 0.$$

It follows from a classical result of nonstandard analysis (Goldblatt, 2012, Theorem 6.1.1) that $\left|{}^*\phi_K\left(\mathrm{Avg}_{F_K} \Theta_{g_0} x\right) - {}^*\phi_K\left(\mathrm{Avg}_{F_K} x\right)\right| \leq \frac{|F_K g_0 \Delta F_K|}{|F_K|} \|x\| \simeq 0$, meaning that $\omega(x) = \omega\left(\Theta_{g_0} x\right)$.

Finally, we argue that ω is a limit point of $\left(\phi_k \circ \mathrm{Avg}_{F_k}\right)_{k=1}^\infty$. For $n, \ell, k_0 \in \mathbb{N}; x_1, \ldots, x_\ell \in \mathfrak{A}$, consider the sentence $\sigma_{x_1,\ldots,x_\ell;n,k_0}$ given by

$$\exists k \in \mathbb{N} \ \left[(k \geq k_0) \wedge \left(\min_{1 \leq j \leq \ell} \left|\omega(x_j) - \phi_k\left(\mathrm{Avg}_{F_k} x_j\right)\right| < 1/n\right)\right].$$

Then ${}^*\sigma_{x_1,\ldots,x_\ell;n,k_0}$ is true for all $n, \ell, k_0 \in \mathbb{N}; x_1, \ldots, x_\ell \in \mathfrak{A}$, witnessed by K. Therefore, it follows from the Transfer Principle that $\sigma_{x_1,\ldots,x_\ell;n,k_0}$ is true for all $n, \ell, k_0 \in \mathbb{N}; x_1, \ldots, x_\ell \in \mathfrak{A}$. We know that

$$\left\{ \left\{ \psi \in \mathcal{S} : \min_{1 \leq j \leq \ell} |\omega(x_j) - \psi(x_j)| < 1/n \right\} : n, \ell \in \mathbb{N}; x_1, \ldots, x_\ell \in \mathfrak{A} \right\}$$

is a neighborhood basis for ω in the weak* topology. Thus we have shown that ω is a limit point of the sequence $\left(\phi_k \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty}$. $\qquad\square$

We might ask whether Lemma 6.5.1 is strictly weaker than Lemma 6.1.2, since Lemma 6.5.1 also asserts that the state it describes is a limit point of the sequence that generates it. In fact, the two lemmas are equivalent in the sense that for a sequence $(\phi_k)_{k=1}^{\infty}$ in $\mathcal{S}$, every limit point of the sequence $\left(\phi_k \circ \mathrm{Avg}_{F_k} \right)_{k=1}^{\infty}$ can be written as $\mathrm{sh}\left({}^*\phi_K \left(\mathrm{Avg}_{F_K} x \right) \right)$ for some $K \in {}^*\mathbb{N}_\infty$. To see this, choose $k_1 < k_2 < \cdots$ such that $\psi = \lim_{\ell \to \infty} \phi_{k_\ell} \circ \mathrm{Avg}_{F_{k_\ell}}$ exists. Let $\mathcal{N}$ be a countable neighborhood basis for ψ in the weak*-topology, and for each $U \in \mathcal{N}, k \in \mathbb{N}$, let $S_{U,k}$ be the set

$$S_{U,k} = \left\{ k' \in \mathbb{N} : \left(k' \geq k \right) \wedge \left(\phi_{k'} \circ \mathrm{Avg}_{F_{k'}} \in U \right) \right\}.$$

Then $\{S_{U,k}\}_{U \in \mathcal{N}, k \in \mathbb{N}}$ has the finite intersection property, and so by the countable saturation of our universe embedding, it follows that there exists $K \in {}^*\mathbb{N}$ such that

$$K \in \bigcap_{U \in \mathcal{N}, k \in \mathbb{N}} {}^*S_{U,k},$$

which is necessarily unlimited. Then for any $x \in \mathfrak{A}$, we have that $\left| {}^*\phi_K \left(\mathrm{Avg}_{F_K} x \right) - \psi(x) \right| < 1/n$ for all $n \in \mathbb{N}$, so

$$\mathrm{sh}\left({}^*\phi_K \left(\mathrm{Avg}_{F_K} x \right) \right) = \psi(x).$$

This correspondence can be generalized in the following result.

Proposition 6.5.2. *Let* $\Omega = (\Omega, \tau)$ *be a compact Hausdorff topological space, and let* $* : \mathfrak{U} \to \mathfrak{U}'$ *be a countably saturated extension of a universe* $\mathfrak{U}$ *containing* Ω *and* $\mathbb{N}$. *Let* $\simeq$ *be the binary relation on* ${}^*\Omega$ *defined*

by

$$x \simeq y \qquad \Longleftrightarrow \qquad \forall U \in \tau \ (x \in {}^*U \leftrightarrow y \in {}^*U).$$

Define a map $\mathrm{sh} : {}^*\Omega \to \Omega$ *that sends* $x \in {}^*\Omega$ *to the unique* $y \in \Omega$ *such that* $x \simeq y$, *and let* $(x_k)_{k=1}^\infty$ *be a sequence in* Ω. *Then the map* sh *is well-defined.*

Further, set

$$\mathrm{LS}\left((x_k)_{k=1}^\infty\right)$$
$$= \left\{ \omega \in \Omega : \forall U \in \tau \ \forall k \in \mathbb{N} \ \left[(\omega \in U) \to \left(\exists k' \in \mathbb{N} \ \left((k' \geq K) \wedge (x_{k'} \in U) \right) \right) \right] \right\}.$$

Then

$$\left\{ \mathrm{sh}\left({}^*x_K\right) : K \in {}^*\mathbb{N}_\infty \right\} \subseteq \mathrm{LS}\left((x_k)_{k=1}^\infty\right).$$

In addition, if $*$ *is* κ*-saturated for some uncountable cardinal* $\kappa > |\mathcal{B}|$, *where* $\mathcal{B}$ *is some topological basis* $\mathcal{B}$ *of* τ, *then* $\left\{ \mathrm{sh}\left({}^*x_K\right) : K \in {}^*\mathbb{N}_\infty \right\} = \mathrm{LS}\left((x_k)_{k=1}^\infty\right).$

Proof. The fact that in a compact topological space, for every $x \in {}^*\Omega$ exists exactly one $y \in \Omega$ such that $x \simeq y$ can be found in (Väth, 2006, Corollary 12.41). Let $K \in {}^*\mathbb{N}_\infty$, and consider $y = \mathrm{sh}\left({}^*x_K\right)$. Let $\mathcal{N}_y = \{U \in \mathcal{B} : y \in U\}$, where $\mathcal{B}$ is a topological basis for τ, and consider for $k \in \mathbb{N}, U \in \mathcal{N}_y$ the sentence $\sigma_{U,k}$ defined by

$$\exists k' \in \mathbb{N} \ \left[\left(k' \geq k\right) \wedge (x_k \in U) \right].$$

Then ${}^*\sigma_{k,U}$ is true for all $k \in \mathbb{N}, U \in \mathcal{N}_y$, since ${}^*x_K \in {}^*U$ and $K \geq k$ for all $k \in \mathbb{N}$, so it follows that $\sigma_{k,U}$ is true for all $k \in \mathbb{N}, U \in \mathcal{N}_y$. Since $\mathcal{N}_y$ forms a neighborhood basis for y, it follows that $y \in \mathrm{LS}\left((x_k)_{k=1}^\infty\right)$.

Now suppose that $*$ is κ-saturated for some uncountable cardinal $\kappa > |\mathcal{B}|$, and let $\omega \in \mathrm{LS}\left((x_k)_{k=1}^\infty\right)$. Let $\mathcal{N}_\omega = \{U \in \mathcal{B} : \omega \in U\}$. For $k \in \mathbb{N}, U \in \mathcal{N}_\omega$, consider the set

$$S_{k,U} = \left\{ k' \in \mathbb{N} : \left(k' \geq k\right) \wedge (x_{k'} \in U) \right\}.$$

Then $\{S_{k,U} : k \in \mathbb{N}, U \in \mathcal{N}_\omega\}$ has the finite intersection property, and thus there exists $K \in \bigcap_{k \in \mathbb{N}, U \in \mathcal{N}_\omega} {}^*S_{k,U}$. Thus ${}^*x_K \in {}^*U$ for all $U \in \mathcal{N}_\omega$, and $K \in {}^*\mathbb{N}_\infty$. Thus $\omega = \mathrm{sh}\left({}^*x_K\right)$. $\qquad\square$

Remark 6.5.3. Our definitions of $\simeq$ and sh in the statement of Proposition 6.5.2 is consistent with our definition of $\simeq$ on $\mathbb{L}$ in the following sense. We can write $\mathbb{L} = \bigcup_{n \in \mathbb{N}} {}^*\{z \in \mathbb{C} : |z| \leq n\}$. If $x, y \in \mathbb{L}$, then there exists $n \in \mathbb{N}$ such that $\max\{|x|, |y|\} \leq n$. Then $x, y \in {}^*\{z \in \mathbb{C} : |z| \leq n\}$. The set $\{z \in \mathbb{C} : |z| \leq n\}$ is compact, and the definition of $\simeq$ on that compact space in the sense of Proposition 6.5.2 will agree with our definition of $\simeq$ on $\mathbb{L}$ from the start of this section.

In light of Theorem 6.5.2, several compactness arguments in this chapter can be proven alternatively in the language of nonstandard analysis. Here we provide a few examples.

Proof of Theorem 6.3.4 using nonstandard analysis. For each $k \in \mathbb{N}$, choose a state ϕ_k on $\mathfrak{A}$ such that

$$\phi_k \left(\mathrm{Avg}_{F_k} a \right) = \left\| \mathrm{Avg}_{F_k} a \right\|.$$

Fix $K \in {}^*\mathbb{N}_\infty$, and let $\omega : \mathfrak{A} \to \mathbb{C}$ be the state

$$\omega(x) = \mathrm{sh} \left({}^*\phi_K \left(\mathrm{Avg}_{F_K} x \right) \right).$$

Lemma 6.5.1 tells us that ω is Θ-invariant. We argue now that $\omega(a) = m \left(a | \mathcal{S}^G \right)$. This follows because if $\psi \in \mathcal{S}^G$, then

$$\psi(a) = \psi \left(\mathrm{Avg}_{F_k} a \right) \leq \left\| \mathrm{Avg}_{F_k} a \right\| = \phi_k \left(\mathrm{Avg}_{F_k} a \right)$$

for all $k \in \mathbb{N}$, and thus we can apply the Transfer Principle to the sentence $\forall k \in \mathbb{N} \ \left(\psi(a) \leq \phi_k \left(\mathrm{Avg}_{F_k} a \right) \right)$ to infer

$$\psi(a) \leq {}^*\phi_K \left(\mathrm{Avg}_{F_K} a \right) \Rightarrow \psi(a) \leq \omega(a).$$

Therefore, we've proven that ${}^*\left\| \mathrm{Avg}_{F_K} a \right\| \simeq m \left(a | \mathcal{S}^G \right)$ for all $K \in {}^*\mathbb{N}_\infty$. Therefore by a classical result of nonstandard analysis (Goldblatt, 2012, Theorem 6.1.1), it follows that $\lim_{k \to \infty} \left\| \mathrm{Avg}_{F_k} a \right\| = m \left(a | \mathcal{S}^G \right)$. $\qquad\square$

Proof of Proposition 6.4.3 using nonstandard analysis. We'll prove that $\overline{a}_{\mathbf{F},S}(x) = \overline{d}_{\mathbf{F},S}(x)$, as the proof that $\underline{a}_{\mathbf{F},S}(x) = \underline{d}_{\mathbf{F},S}(x)$ is very similar. We know a priori that $\mathscr{P}^{\mathbf{F}}(S) = \overline{S}$.

Let $(\phi_k)_{k=1}^\infty$ be a sequence in S such that for each $k \in \mathbb{N}$, we have

$$\sup \left\{ \phi \left(\mathrm{Avg}_{F_k} x \right) : \phi \in S \right\} - 1/k \leq \phi_k (\mathrm{Avg}_{F_k} x) \leq \sup \left\{ \phi \left(\mathrm{Avg}_{F_k} x \right) : \phi \in S \right\}.$$

Let $K \in {}^*\mathbb{N}_\infty$, and let $\omega : \mathfrak{A} \to \mathbb{C}$ be the state $\omega(y) = \mathrm{sh}\left({}^*\phi_K\left(\mathrm{Avg}_{F_K} y\right)\right)$. Then $\omega \in \mathscr{P}^{\mathbf{F}}(S)$, so $\omega(x) \leq \bar{a}_{\mathbf{F},S}(x)$.

To prove the opposite inequality, let $\psi \in \mathscr{P}^{\mathbf{F}}(S) = \overline{S}$. Then

$$
\begin{aligned}
\psi(x) &= \psi\left(\mathrm{Avg}_{F_k} x\right) \\
&\leq \sup\left\{\phi\left(\mathrm{Avg}_{F_k} x\right) : \phi \in S\right\} \\
&\leq \phi_k\left(\mathrm{Avg}_{F_k} x\right) + 1/k \qquad\qquad\qquad (\forall k \in \mathbb{N}).
\end{aligned}
$$

Thus the sentence

$$
\forall k \in \mathbb{N} \ \left(\psi(x) \leq \phi_k\left(\mathrm{Avg}_{F_k} x\right) + 1/k\right)
$$

is true. Applying the Transfer Principle then tells us that $\psi(x) \leq {}^*\phi_K\left(\mathrm{Avg}_{F_K} x\right) + 1/K$, implying that $\psi(x) \leq \omega(x)$. Taking a supremum over $\psi \in \overline{S} = \mathscr{P}^{\mathbf{F}}(S)$ tells us that

$$
\bar{a}_{\mathbf{F},S}(x) \leq \omega(x).
$$

Therefore ${}^*\phi_K(x) \simeq \bar{a}_{\mathbf{F},S}(x)$ for all $K \in {}^*\mathbb{N}$. Thus, by a classical result of nonstandard analysis (Goldblatt, 2012, Theorem 6.1.1), it follows that $\lim_{k\to\infty}\left\|\mathrm{Avg}_{F_k} a\right\| = \bar{a}_{\mathbf{F},S}(x)$. $\qquad\square$

Proof of Theorem 6.4.5 using nonstandard analysis. (i)$\Rightarrow$(ii): Suppose $\left\{\psi(x) : \psi \in \mathscr{P}^{\mathbf{F}}(S)\right\} = \{\lambda\}$. For each $k \in \mathbb{N}$, choose $\phi_k \in S$ such that $\left|\phi_k\left(\mathrm{Avg}_{F_k} x - \lambda\right)\right| \geq \frac{1}{2}\left\|\mathrm{Avg}_{F_k} x - \lambda\right\|_S$. Fix $K \in {}^*\mathbb{N}$, and let $\omega : \mathfrak{A} \to \mathbb{C}$ be the state

$$
\omega(y) = \mathrm{sh}\left({}^*\phi_K\left(\mathrm{Avg}_{F_K} y\right)\right).
$$

Lemma 6.5.1 tells us that $\omega \in \mathscr{P}^{\mathbf{F}}(S)$. Thus $\omega(x) = \lambda$. Therefore $\left|{}^*\phi_K\left(\mathrm{Avg}_{F_K} x\right) - \lambda\right| \simeq 0$ for all $K \in {}^*\mathbb{N}_\infty$, meaning a classical result of nonstandard analysis (Goldblatt, 2012, Theorem 6.1.1) tells us that $\lim_{k\to\infty}\left|\phi_k\left(\mathrm{Avg}_{F_k} x\right) - \lambda\right| = 0$. But because $\left\|\mathrm{Avg}_{F_k} x - \lambda\right\|_S \leq 2\left|\phi_k\left(\mathrm{Avg}_{F_k} x\right) - \lambda\right|$ for all $k \in \mathbb{N}$, we can conclude that $\lim_{k\to\infty}\left\|\mathrm{Avg}_{F_k} x - \lambda\right\|_S = 0$.

(ii)$\Rightarrow$(i): Suppose that $\lim_{k\to\infty}\left\|\mathrm{Avg}_{F_k} x - \lambda\right\|_S = 0$. Let $(\phi_k)_{k=1}^\infty$ be a sequence in S, and let $\omega : \mathfrak{A} \to \mathbb{C}$ be the state

$$
\omega(y) = \mathrm{sh}\left({}^*\phi_K\left(\mathrm{Avg}_{F_K} y\right)\right).
$$

Then

$$\left|\omega(x - \lambda)\right| \simeq \left|{}^{*}\phi_K\left(\mathrm{Avg}_{F_K} x - \lambda\right)\right| \leq \left\|{}^{*}\mathrm{Avg}_{F_K} x - \lambda\right\|_S \simeq 0.$$

Therefore $\omega(x) = \lambda$. We can then take a supremum to get

$$\sup_{(\phi_k)_{k=1}^{\infty} \in S^{\mathbb{N}}, K \in {}^{*}\mathbb{N}_{\infty}} \left|\mathrm{sh}\left({}^{*}\phi_K\left(\mathrm{Avg}_{F_K} x\right)\right) - \lambda\right| = 0.$$

But in light of Proposition 6.5.2, we know that

$$\mathscr{P}^{\mathbf{F}}(S) = \left\{y \mapsto \mathrm{sh}\left({}^{*}\phi_K\left(\mathrm{Avg}_{F_K} y\right)\right) : (\phi_k)_{k=1}^{\infty} \in S^{\mathbb{N}}, K \in {}^{*}\mathbb{N}_{\infty}\right\},$$

so this shows that $\psi(x) = \lambda$ for all $\psi \in \mathscr{P}^{\mathbf{F}}(S)$. $\qquad\square$

BIBLIOGRAPHY

Abadie, B. and Dykema, K. (2009). Unique ergodicity of free shifts and some other automorphisms of c*-algebras. *Journal of Operator Theory*, pages 279–294.

Aliprantis, C. D. and Burkinshaw, O. (2006). *Positive operators*, volume 119. Springer Science & Business Media.

Assani, I. (2003). *Wiener Wintner ergodic theorems*. World Scientific Publishing Company.

Assani, I. and Young, A. (2022). Spatial-temporal differentiation theorems. *Acta Mathematica Hungarica*, 168:301–344.

Assani, I. and Young, A. (2023). Non-autonomous spatial-temporal differentiation theorems for group endomorphisms. *Real Analysis Exchange*, pages 397–424.

Baggett, L. W. (1992). *Functional Analysis: a primer*. Dekker.

Bannon, J. P., Cameron, J., and Mukherjee, K. (2018). On noncommutative joinings. *International Mathematics Research Notices*, 2018(15):4734–4779.

Bellow, A. and Furstenberg, H. (1979). An application of number theory to ergodic theory and the construction of uniquely ergodic models. *Israel Journal of Mathematics*, 33(3-4):231–240.

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660.

Blackadar, B. (2006). *Operator algebras: theory of C*-algebras and von Neumann algebras*, volume 122. Springer Science & Business Media.

Bourgain, J. (1989). Pointwise ergodic theorems for arithmetic sets. *Publications Mathématiques de l'IHÉS*, 69:5–41.

Bowen, R. (1971). Periodic points and measures for axiom a diffeomorphisms. *Transactions of the American Mathematical Society*, 154:377–397.

Dajani, K. and Dirksin, S. (2008). A simple introduction to ergodic theory. *University of Utrecht, Lecture notes in Ergodic Theory*.

Dajani, K. and Kraaikamp, C. (2002). *Ergodic theory of numbers*. Number 29. Cambridge University Press.

Davies, E. (1967). A generalized theory of convexity. *Proceedings of the London Mathematical Society*, 3(4):644–652.

De Guzman, M. (1976). *Differentiation of Integrals in Rn*.

Denker, M., Grillenberger, C., and Sigmund, K. (2006). *Ergodic theory on compact spaces*, volume 527. Springer.

Duvenhage, R. and Stroh, A. (2011). Unique ergodicity and disjointness of c*-dynamical systems. *arXiv preprint arXiv:1102.4243*.

Eckmann, B. (1943). Über monothetische gruppen. *Commentarii Mathematici Helvetici*, 16(1):249–263.

Einsiedler, M. (2010). Effective equidistribution and spectral gap. In *European Congress of Mathematics Amsterdam, 14–18 July, 2008*, pages 31–51.

Eisner, T., Farkas, B., Haase, M., and Nagel, R. (2015). *Operator theoretic aspects of ergodic theory*, volume 272. Springer.

Folland, G. B. (1999). *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons.

Folland, G. B. (2016). *A course in abstract harmonic analysis*. CRC Press.

Goldblatt, R. (2012). *Lectures on the hyperreals: an introduction to nonstandard analysis*, volume 188. Springer Science & Business Media.

Goodman, T. N. (1971). Relating topological entropy and measure entropy. *Bulletin of the London Mathematical Society*, 3(2):176–180.

Hansel, G., Raoult, J., and Rosenblatt, M. (1973). Ergodicity, uniformity and unique ergodicity. *Indiana University Mathematics Journal*, 23(3):221–237.

Herman, M. R. (1983). Une méthode pour minorer les exposants de lyapounov et quelques exemples montrant le caractere local d'un théoreme d'arnold et de moser sur le tore de dimension 2. *Commentarii Mathematici Helvetici*, 58(1):453–502.

Hlawka, E. (1956). Folgen auf kompakten räumen. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 20, pages 223–241. Springer.

Jenkinson, O. (2006a). Ergodic optimization. *Discrete & Continuous Dynamical Systems-A*, 15(1):197.

Jenkinson, O. (2006b). Every ergodic measure is uniquely maximizing. *Discrete and Continuous Dynamical Systems*, 16(2):383.

Jewett, R. I. (1970). The prevalence of uniquely ergodic systems. *Journal of Mathematics and Mechanics*, 19(8):717–729.

Kallenberg, O. (2021). *Foundations of Modern Probability*. Springer, Cham, 3 edition.

Kerr, D. and Li, H. (2016). *Ergodic theory*. Springer.

Kloeden, P. E. and Rasmussen, M. (2011). *Nonautonomous dynamical systems*. Number 176. American Mathematical Soc.

Krengel, U. (2011). *Ergodic theorems*, volume 6. Walter de Gruyter.

Krieger, W. (1970). On entropy and generators of measure-preserving transformations. *Transactions of the American Mathematical Society*, 149(2):453–464.

Kuipers, L. and Niederreiter, H. (2012). *Uniform distribution of sequences*. Courier Corporation.

Kwietniak, D., Lacka, M., and Oprocha, P. (2016). A panorama of specification-like properties and their consequences. *Contemporary Mathematics*, 669:155–186.

Li, J. and Wu, M. (2016). Points with maximal birkhoff average oscillation. *Czechoslovak Mathematical Journal*, 66(1):223–241.

Lindenstrauss, E. (2001). Pointwise theorems for amenable groups. *Inventiones mathematicae*, 146(2):259–295.

Mance, B. (2010). *Normal numbers with respect to the Cantor series expansion*. PhD thesis, The Ohio State University.

Marcus, B. (1980). A note on periodic points for ergodic toral automorphisms. *Monatshefte für Mathematik*, 89(2):121–129.

Neumann, J. v. (1932). Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences*, 18(1):70–82.

Ormes, N. S. (1997). Strong orbit realization for minimal homeomorphisms. *Journal d'Analyse Mathématique*, 71(1):103–133.

Oxtoby, J. C. (1952). Ergodic sets. *Bulletin of the American Mathematical Society*, 58(2):116–136.

Pedersen, G. (1979). C*-algebras and their automorphism groups. *Soc. Monogr. Ser. Academic Press, London*.

Phelps, R. R. (2001). *Lectures on Choquet's theorem*. Springer Science & Business Media.

Procesi, C. (2006). *Lie groups: an approach through invariants and representations*. Springer Science & Business Media.

Rudin, W. (1962). *Fourier analysis on groups*, volume 121967. Wiley Online Library.

Rudin, W. (1991). *Functional analysis*. McGraw-Hill, Inc.

Sakai, S. (2012). *C*-algebras and W*-algebras*. Springer Science & Business Media.

Väth, M. A. (2006). *Nonstandard analysis*. Springer Science & Business Media.

Walters, P. (2007). *Ergodic theory—introductory lectures*, volume 458. Springer.

Weyl, H. (1968). Uber ein problem aus dem gebiete der diophantischen, ges. abh. i.

Wiener, N. and Wintner, A. (1941). Harmonic analysis and ergodic theory. *American Journal of Mathematics*, 63(2):415–426.

Zhao, Y. (2016). Maximal integral over observable measures. *Acta Mathematica Sinica. English Series*, 32(5):571–578.