On equivariant triangularization of matrix cocycles

by

Joseph Anthony Horan
B.Math, University of Waterloo, 2013

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

On equivariant triangularization of matrix cocycles

by

Joseph Anthony Horan
B.Math, University of Waterloo, 2013

Supervisory Committee

Dr. Anthony Quas, Co-supervisor
(Department of Mathematics and Statistics)

Dr. Christopher Bose, Co-supervisor
(Department of Mathematics and Statistics)

Supervisory Committee

Dr. Anthony Quas, Co-supervisor
(Department of Mathematics and Statistics)

Dr. Christopher Bose, Co-supervisor
(Department of Mathematics and Statistics)

ABSTRACT

The Multiplicative Ergodic Theorem is a powerful tool for studying certain types of dynamical systems, involving real matrix cocycles. It gives a block diagonalization of these cocycles, according to the Lyapunov exponents. We ask if it is always possible to refine the diagonalization to a block upper-triangularization, and if not over the real numbers, then over the complex numbers. After building up to the posing of the question, we prove that there are counterexamples to this statement, and give concrete examples of matrix cocycles which cannot be block upper-triangularized.

Contents

ACKNOWLEDGEMENTS

I would like to thank:

my fellow grads, who keep me both laughing and thinking;

my supervisors, who convinced me that I belong;

Tuesday Tea participants, who think I'm a decent baker and poet;

my housemate, who kept me company through everything;

my friends, both near and far, who stay with me regardless of distance;

my family, who have always been there for me.

Though I name no names, know that I keep each and every one of you in mind.

DEDICATION

To all of those people who ever thought I was worth caring about;
I would never have made it this far without your love.
I hope this serves as at least partial justification.

Chapter 1

Introduction

For as long as humanity has looked up to the sky and wondered about the underlying mechanisms of the cosmos, we have really just studied dynamical systems, if in a very applied sense. Of course, given the development of mathematics, the way in which we study dynamical systems has changed drastically. Questions such as "How fast is this thing expanding?" and "Does this system have an equilibrium?", are now more rigorously formulated as the concepts of growth rates, stability, and chaos. To study these concepts, we look at the Lyapunov exponents for the system, which determine rates of expansion or contraction for elements of the system, and by find subspaces of the system which grow asymptotically according to those rates.

As a simple example, consider a differentiable system, such as a system of ordinary differential equations. Around an equilibrium point of the system, the dynamics are governed by a single matrix: the derivative. The Lyapunov exponents can be found by considering the eigenvalues of the matrix, and the subspaces corresponding to each growth rate are the eigenspaces. This information is very easily obtained when the matrix is in Jordan Normal form, which is upper-triangular.

If we wish to obtain information for the system away from equilibrium points, then we must consider how the derivative changes over time. This is modelled by a matrix cocycle, which is a product of matrices changing over time. The Multiplicative Ergodic Theorem (abbreviated as MET), originally proved by Oseledets in 1968 [16], asserts the existence of Lyapunov exponents and corresponding subspaces for these cocycles, which can be seen as a block diagonalization of the cocycle. This theorem was incredibly influential in the study of dynamical systems, and it spawned an entirely new branch of research in the area.

The original proof of the MET involves considering a matrix cocycle, and constructing an equivalent cocycle comprised solely of upper-triangular matrices, over a carefully chosen auxiliary space. It is natural and important to ask if we may upper-triangularize over the original space instead. Moreover, given a single real matrix, while it may not be triangularizable over the real numbers, it is always triangularizable over the complex numbers. Given this, it is reasonable to ask this question: is every matrix cocycle upper-triangularizable? If not over the real numbers, then over the complex numbers?

In this thesis, we develop the background to rigorously talk about matrix cocycles and upper-triangularization, and then work towards answering the question posed above. In Chapter 2, we state some measure-theoretic preliminaries, and work with some important recurring examples, before providing a development of basic dynamical systems theory, and introducing ergodicity. We specify two particular types of systems, the induced transformation and the skew product, which will be extremely useful. Following that, the MET is stated, and accompanied by examples and historical discussion. In Chapter 3, we formalize the question we wish to ask, and find its context in the mathematical literature. The question is then answered, via abstract theorem, and illustrated by three distinct examples. We conclude by briefly examining possible avenues for further research in this direction. For convenience and interest, proofs of some useful results may be found in the Appendix.

Chapter 2

Preliminaries

The study of measurable cocycles is, in the grand scheme of mathematics, a fairly recent development; however, the cocycle is a natural object to study, because it models the concept of accumulating data along an orbit in a system. As a branch of analysis, it builds on the foundations established in measure theory. We shall present here a brief overview of the definitions necessary to present a coherent development of measurable cocycles, the Multiplicative Ergodic Theorem, and equivariant triangularization of matrix cocycles. Interspersed with the definitions are important examples which will be both useful for placing the definitions into context and for saving computational work later on, and some results which build upon the definitions and lay the solid foundation for the remainder of the thesis.

2.1 Measure theory and background analysis and geometry

In this thesis, we will assume a standard amount of measure theory and general analysis, as seen in a solid pure mathematics undergraduate degree. That is to say, basic definitions such as measure spaces and standard theorems such as Carathéodory's

Extension Theorem will be assumed. However, we will provide a refresher on some of the specific results which find use in the work, so as to keep the thesis as self-contained as possible. See [12] or similar measure theory reference for a much more detailed presentation.

We will be dealing solely with positive measure spaces, so 'measure space' is implicitly a positive measure space. In general, $\mathbb{N} = \{0, 1, 2, \dots\}$, and $\mathbb{N}_+ = \{1, 2, \dots\}$.

Definition 2.1. Let X be a set. A *π-system* is a collection $\mathcal{P}$ of subset of X such that for any $A, B \in \mathcal{P}$, we have $A \cap B \in \mathcal{P}$.

Definition 2.2. Let X be a set. A *semi-algebra* is a collection $\mathcal{P}$ of subsets of X such that:

1. If $A, B \in \mathcal{P}$, then $A \cap B \in \mathcal{P}$.
2. If $A \in \mathcal{P}$, then there are n pairwise disjoint $C_i, \dots, C_n \in \mathcal{P}$ such that $X \setminus A = C_1 \cup \dots \cup C_n$.

Definition 2.3. Let X be a set. An *algebra* is a collection $\mathcal{A}$ of subsets of X such that:

1. $X \in \mathcal{A}$.
2. If $A \in \mathcal{A}$, then $X \setminus A \in \mathcal{A}$.
3. If $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.

A *σ-algebra* is an algebra which is also closed under countable unions: if $A_i \in \mathcal{B}$ for all $i \geq 1$, then

$$\bigcup_{i \geq 1} A_i \in \mathcal{B}$$

.

Lemma 2.4. *Let X be a set, and let $\mathcal{S}$ be a collection of subsets of X. Then there is a σ-algebra which contains $\mathcal{S}$ and which is smallest out of all such σ-algebras; we call this the σ-algebra* generated by $\mathcal{S}$, *and we denote this $\sigma(\mathcal{S})$.*

Lemma 2.5. *Let $\mathcal{P}$ be a semi-algebra of subsets of X. The algebra generated by $\mathcal{P}$, denoted $\mathcal{A}(\mathcal{P})$, is*

$$\mathcal{A}(\mathcal{P}) = \{A \ : \ A = C_1 \cup \cdots \cup C_n,\ C_i \in \mathcal{P}, i = 1 \ldots n, C_i \cap C_j = \emptyset \forall i \neq j\}.$$

Theorem 2.6 (Monotone Class Theorem)**.** *Let X be a set, and $\mathcal{A}$ be an algebra of subsets of X. Let $\mathcal{M}(\mathcal{A})$ be the monotone class generated by $\mathcal{A}$; that is, $\mathcal{M}(\mathcal{A})$ is the smallest monotone class (ie. is closed under unions of increasing sequences of sets and under intersections of decreasing sequences of sets) containing $\mathcal{A}$. Then $\mathcal{M}(\mathcal{A})$ is a σ-algebra, and is precisely equal to $\sigma(\mathcal{A})$.*

Example 2.7. If X and Y are sets with semi-algebras $\mathcal{P}$ and $\mathcal{Q}$ respectively, then the collection of measurable rectangles

$$\mathcal{R} = \{A \times B \ : \ A \in \mathcal{P}, B \in \mathcal{Q}\}$$

is a semi-algebra of subsets of $X \times Y$. For $A, C \in \mathcal{P}$ and $B, D \in \mathcal{Q}$:

1. $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$, and $A \cap C \in \mathcal{P}$, $B \cap D \in \mathcal{Q}$ Hence we have $(A \times B) \cap (C \times D) \in \mathcal{R}$.

2. First, write
$$X \setminus A = \bigcup_{i=1}^{n} C_i, \quad Y \setminus B = \bigcup_{i=1}^{n} D_i,$$

for $C_i \in \mathcal{P}$, $D_j \in \mathcal{Q}$, all pairwise disjoint. Then we obtain:

$$\begin{aligned}(X \times Y) \setminus (A \times B) &= \{(x, y) \ : \ x \in X \setminus A \text{ or } y \in Y \setminus B\} \\ &= ((X \setminus A) \times B) \cup (A \times (Y \setminus B)) \cup ((X \times A) \times (Y \setminus B)) \\ &= \left(\bigcup_{i=1}^{n} C_1 \times B\right) \cup \left(\bigcup_{j=1}^{m} A \times D_j\right) \cup \left(\bigcup_{i=1}^{n} \bigcup_{j=1}^{m} C_i \times D_j\right),\end{aligned}$$

which is a disjoint union of things in $\mathcal{R}$.

Hence $\mathcal{R}$ is a semi-algebra. Just looking at part (1) of that computation, the same statement would be true if instead of semi-algebras, $\mathcal{P}$ and $\mathcal{Q}$ were π-systems. This will be useful later on.

If $\mathcal{A} = \sigma(\mathcal{P})$ and $\mathcal{B} = \sigma(\mathcal{Q}$, then the *product* σ-algebra is the σ-algebra generated by $\mathcal{R}$, which we denote by $\mathcal{A} \otimes \mathcal{B}$; this is true for π-systems or for semi-algebras. As is reasonably clear, it is *not* the Cartesian product of $\mathcal{A}$ and $\mathcal{B}$. If we have measures on $\mathcal{A}$ and $\mathcal{B}$, then the product measure can be defined on $\mathcal{A} \otimes \mathcal{B}$, utilizing the Monotone Class Theorem and Carathéodory's Extension Theorem.

In more generality, if we have a family of measurable spaces $(X_i, \mathcal{B}_i)_{i \in \mathbb{Z}}$, we may define the product σ-algebra as the σ-algebra generated by the semi-algebra $\mathcal{R}$, given by:

$$\mathcal{R} = \left\{ \prod_{i \in \mathbb{Z}} B_i \ : \ B_i = X_i \text{ for all but finitely many } i \right\}.$$

This is equivalent to finding the smallest σ-algebra such that all of the projection maps $\pi_i : \prod_{i \in \mathbb{Z}} \to X_i$ are measurable. We denote the product σ-algebra by $\bigotimes_{i \in \mathbb{Z}} \mathcal{B}_i$.

Of course, in the previous lemma, our semi-algebras could be the full σ-algebras on each space.

Theorem 2.8 (Fubini's Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \times \nu)$ be the product measure space. If $f : X \times Y \to \mathbb{R}$*

is either integrable or both measurable and non-negative, then we have that the sections

$$x \mapsto \int_Y f(x,y)\ d\nu(y), \quad y \mapsto \int_X f(x,y)\ d\mu(x),$$

are measurable functions of X *and* Y*, respectively, and*

$$\int_{X\times Y} f\ d\mu\times\nu = \int_X\int_Y f(x,y)\ d\nu(y)d\mu(x) = \int_Y\int_X f(x,y)\ d\mu(x)d\nu(y).$$

Definition 2.9. Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be measure spaces, and let $\phi : X \to Y$ be measurable. We may define a set function on $\mathcal{B}$, $\phi_*\mu$, by setting

$$\phi_*\mu(B) = \mu(\phi^{-1}(B)),$$

for $B \in \mathcal{B}$. This is well-defined, as ϕ is measurable, and because pullbacks under any function distribute over unions and preserve disjointness of sets, we may see that $\phi_*\mu$ is a measure on $(Y,\mathcal{B})$:

1. $\phi_*\mu(\emptyset) = \mu(\phi^{-1}(\emptyset)) = \mu(\emptyset) = 0$;

2. $\phi_*\mu(A) = \mu(\phi^{-1}(A)) \geq 0$;

3. If $\{A_i\}_{i=1}^{\infty}$ are disjoint sets, then

$$\begin{aligned}\phi_*\mu\left(\bigcup_{i=1}^{\infty} A_i\right) &= \mu\left(\phi^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)\right) = \mu\left(\bigcup_{i=1}^{\infty}\phi^{-1}(A_i)\right) = \sum_{i=1}^{\infty}\mu(\phi^{-1}(A_i)) \\ &= \sum_{i=1}^{\infty}\phi_*\mu(A_i).\end{aligned}$$

We call $\phi_*\mu$ the *pushforward measure* of μ by ϕ. If $\phi_*\mu = \nu$, then we say that ϕ is *measure-preserving.*

Theorem 2.10 (A specific case of the Kolmogorov Extension theorem)**.** *For each $i \in \mathbb{Z}$, let $(X_i, \mathcal{B}_i, \mu_i)$ be a probability space, and let*

$$X = \prod_{i=-\infty}^{\infty} X_i, \ \mathcal{B} = \bigotimes_{i=-\infty}^{\infty} \mathcal{B}_i.$$

Then there exists a unique probability measure μ on $\mathcal{B}$ such that for any finite set $F \subset \mathbb{Z}$ with associated measurable projection map

$$\pi_F : X \to \prod_{i \in F} X_i, \ \pi_F(x) = (x_i)_{i \in F},$$

we have

$$(\mu_F)_* \mu = \prod_{i \in F} \mu_i.$$

Example 2.11 (Bernoulli Shift)**.** Let A be a countable set, called the *alphabet.* Let $\mathcal{B} = \mathcal{P}(A)$, and for each $a \in A$, assign a weight p_a to a, such that

$$\sum_{a \in A} p_a = 1.$$

It is easy to check that the set function $\nu : \mathcal{B} \to \mathbb{R}$ defined by

$$\nu(B) = \sum_{b \in B} p_b$$

is a probability measure on $(A, \mathcal{P}(A))$.

For each $n \in \mathbb{Z}$, let $(X_n, \mathcal{A}_n, \mu_n) = (A, \mathcal{B}, \nu)$. Then we may apply the case of the Kolmogorov Extension Theorem (Theorem 2.10) to obtain a probability space $(X, \mathcal{A}, \mu)$ with

$$X = \prod_{n \in \mathbb{Z}} A, \ \mathcal{A} = \prod_{n \in \mathbb{Z}} \mathcal{P}(A), \ \mu = \prod_{n \in \mathbb{Z}} \nu.$$

This is called the *bilateral shift space* over the alphabet A. Elements of X are sequences $x = (\dots x_{-1} \cdot x_0 x_1 \dots)$, and a generating semi-algebra for $\mathcal{A}$ is the collection

$$\mathcal{C} = \left\{ \pi_F^{-1}\left(\prod_{i \in F} B_i\right) \ : \ F \subset \mathbb{Z} \text{ finite}, \ B_i \subset A \right\}.$$

We shall call these sets *cylinder sets.* By the projection property of this measure, if $F_1, F_2 \subset \mathbb{Z}$ are disjoint, the measure of the intersection of sets $C_1, C_2 \in \mathcal{C}$ fixing those components is given by:

$$\mu(C_1 \cap C_2) = \nu_{F_1,F_2}(\pi_{F_1,F_2}(C_1 \cap C_2)) = \nu_{F_1}(\pi_{F_1}(C_1))\nu_{F_2}(\pi_{F_2}(C_2)) = \mu(C_1)\mu(C_2).$$

In particular, we will deal with sets of the form

$$C = \bigcap_{i \in F} \pi_i^{-1}\{b_i\}$$

where $b_i \in A$ and $F \subset \mathbb{Z}$ is finite, so that some finite number of symbols are fixed. The measure of such a set is

$$\mu(C) = \prod_{i \in F} p_i.$$

At times, we will denote

$$\bigcap_{i=j}^{n} \pi_i^{-1}\{a_i\} = C(x_j \dots x_n = a_j \dots a_n).$$

We shall call these *contiguous string* cylinder sets.

We may find a generating π-system for $\mathcal{A}$, which contains only contiguous string cylinder sets. This will allow us to significantly simplify later proofs.

Lemma 2.12. *If $(X, \mathcal{A})$ is the bilateral shift space over a countable alphabet A with σ-algebra generated by the cylinder sets as above, then the set of contiguous string* *cylinder sets which span the* 0 *index,*

$$\mathcal{D} = \left\{ C(x_{-t} \dots x_{r-1} = a_{-t} \dots a_{r-1}) \ : \ t \geq 0, r \geq 1, \ a_i \in A, i = -t, \dots r-1 \right\},$$

is a generating π*-system for* $\mathcal{A}$*.*

Proof. To see that $\mathcal{D}$ generates $\mathcal{A}$, first note that $\mathcal{D} \subset \mathcal{C}$, so that $\sigma(\mathcal{D}) \subset \sigma(\mathcal{C}) = \mathcal{A}$. Then note that if $C \in \mathcal{C}$, then we have, for some finite $F \subset \mathbb{Z}$ and $B_i \subset A$ for each $i \in F$:

$$C = \bigcap_{i \in F} \pi_i^{-1}(B_i) = \bigcap_{i \in F} \bigcup_{a \in B_i} \pi_i^{-1}\{a\} = \bigcup_{(a_i)_{i \in F} \in \prod_{i \in F} B_i} \left(\bigcap_{j \in F} \pi_i^{-1}\{a_i\} \right).$$

Let $f_1 = \min(\{f \in F\} \cup \{0\}$, $f_2 = \max(\{f \in F\} \cup \{1\}$, and denote $E = \{f_1, f_2 - 1\} \setminus F$. We may then observe that:

$$C = \bigcup_{(a_i)_{i \in E} \in \prod_{i \in E} A} \ \bigcup_{(a_i)_{i \in F} \in \prod_{i \in F} B_i} C(x_{f_1} \dots x_{f_2 - 1} = a_{f_1} \dots a_{f_2 - 1}),$$

where this is a countable disjoint union. Hence, we see that $\mathcal{C} \subset \sigma(\mathcal{D})$, and so $\mathcal{A} = \sigma(\mathcal{C}) \subset \sigma(\mathcal{D})$, which shows that $\mathcal{A} = \sigma(\mathcal{D})$. Therefore $\mathcal{D}$ generates $\mathcal{A}$. To show that $\mathcal{D}$ is a π-system, let

$$C_1 = C(x_{-t_1} \dots x_{r_1 - 1} = a_{-t_1} \dots a_{r_1 - 1}), \quad D_2 = C(x_{-t_2} \dots x_{r_2 - 1} = b_{-t_2} \dots b_{r_2 - 1}),$$

where $t_1, t_2 \geq 0$, and $r_1, r_2 \geq 1$. Denote

$$t^* = \max\{t_1, t_2\}, \ t_* = \min\{t_1, t_2\}, \ r^* = \max\{r_1, r_2\}, \ r_* = \min\{r_1, r_2\}.$$

If $a_i \neq b_i$ for some $i \in \{-t_*, \dots, r_* - 1\}$, then $C_1 \cap C_2 = \emptyset$. Otherwise, we have

$$C_1 \cap C_2 = C(x_{-t^*} \dots x_{r^*-1} = a_{-t^*} \dots a_{r_*-1} b_{r_*} \dots b_{r^*-1}).$$

Hence $C_1 \cap C_2 \in \mathcal{D}$, so that $\mathcal{D}$ is a π-system, which concludes the proof. □

Definition 2.13. Let (G, τ) be a locally compact Hausdorff topological group, with $\mathcal{B}$ the resulting Borel σ-algebra. The σ-finite left-translation-invariant regular measure m on $(G, \mathcal{B})$ which is unique up to a positive scaling is called the *Haar* measure. If G is compact, we may choose the constant so that m is a probability measure.

Example 2.14. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, with τ being the quotient topology on $\mathbb{T}$. Then $(\mathbb{T}, \tau)$ is a compact Hausdorff topological group, under addition; this is by the quotient group and quotient topology constructions. Let $\mathcal{B}$ be the Borel σ-algebra on $\mathbb{T}$; the normalized Haar measure on $(\mathbb{T}, \mathcal{B})$ is the Lebesgue measure λ on the half-open unit interval $[0, 1)$; it is easy to see that λ is normalized, regular, and translation-invariant.

Example 2.15. Let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, with $\tau = \mathcal{B} = \mathcal{P}(\mathbb{Z}_n)$. Then $(\mathbb{Z}_n, \mathcal{B})$ is a compact Hausdorff topological group, again under addition. Define the measure c on $\mathcal{B}$ by $c\{k\} = \frac{1}{n}$; this is the counting measure on $\mathbb{Z}_n$. We can see that c is normalized and regular; since each point in $\mathbb{Z}_n$ has the same measure and translation is a bijection, c is seen to be translation-invariant, so that c is the normalized Haar measure.

We have need to talk about the situation where two measure spaces are ‘the same’, in some fundamental way. The strongest concept is that of ‘isomorphism of measure spaces’, which we now define. There are other related concepts, such as ‘conjugacy of measure algebras’, which are useful, but not relevant to the current work. For more on these ideas, see [31].

Definition 2.16. We say that measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are *isomorphic* (in the *measure-theoretic* sense) when there exists $X_1 \subset X$, $Y_1 \subset Y$ with $\mu(X \setminus X_1) =$

$\nu(Y \setminus Y_1) = 0$, and a bijective measure-preserving transformation $\phi : X_1 \to Y_1$ whose inverse ϕ^{-1} is also measure-preserving. We call ϕ an *isomorphism of measure spaces* or sometimes a *measure-theoretic isomorphism*.

Definition 2.17. Let $k, n \in \mathbb{N}$, with $1 \leq k < n$, and let $\mathbb{F}$ be a field. We denote the set of the k-dimensional subspaces of $\mathbb{F}^n$ by $\mathrm{Gr}_k(\mathbb{F}^n)$; this is the *Grassmannian* of dimension k subspaces of $\mathbb{F}^n$. When $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$, the Grassmannian is a compact metric space, with distance defined by, for k-d subspaces V_1 and V_2 of $\mathbb{F}^n$,

$$d(V_1, V_2) = d_H(V_1 \cap B, V_2 \cap B),$$

where B is the closed unit ball in $\mathbb{F}^n$ and d_H is the Hausdorff distance.

The following proposition allows us to describe $\mathrm{Gr}_1(\mathbb{C}^2)$ in terms of a more well-known object. A proof for this may be found in the appendix, in section A.1.

Proposition 2.18. *If* $\{v_1, v_2\}$ *are given by*

$$v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

then any subspace of $Gr_1(\mathbb{C}^2)$ *may be written as either* $\mathrm{span}_{\mathbb{C}}\{v_1 + zv_2\}$*, for some* $z \in \mathbb{C}$*, or as* $\mathrm{span}_{\mathbb{C}}\{v_2\}$*, so that the map* $\psi : Gr_1(\mathbb{C}^2) \to \bar{\mathbb{C}}$ *given by*

$$\psi(\mathrm{span}_{\mathbb{C}}\{v_1 + zv_2\}) = z, \quad \psi(\mathrm{span}_{\mathbb{C}}\{v_2\}) = \infty,$$

is a bijection between $Gr_1(\mathbb{C}^2)$ *and* $\bar{C}$*. Moreover,* ψ *is continuous with continuous inverse, and hence measurable with measurable inverse.*

Of course, our particular choice of basis is not unique; we could have picked any basis. However, this particular basis is a very appropriate choice later on, when we deal with real orthogonal matrices.

2.2 Ergodic theory and dynamical systems

This thesis is primarily concerned with a theorem stated in the setting of *discrete-time* dynamical systems. Dynamical systems usually involve an action of a group or semi-group (for example, $\mathbb{R}, \mathbb{R}^+, \mathbb{Z}, \mathbb{N}$) on a topological space or a measure space. When the action is over $\mathbb{R}$ or $\mathbb{R}^+$, we typically call it a *flow*; these tend to be actions on manifolds or other such constructs. In this work, we shall be focusing entirely on iterations of measurable and measure-preserving maps, which are actions of $\mathbb{Z}$ or $\mathbb{N}$ on a space.

Definition 2.19. Let $(X, \mathcal{B})$ be a measurable space, and let $T : X \to X$ be measurable. There is a natural $\mathbb{Z}$-action on X, given by $(n, x) \mapsto T^n(x)$. The tuple $(X, \mathcal{B}, T)$ is called a *discrete-time measurable dynamical system*, or just *dynamical system*.

We may be more specific about which spaces we will consider; in particular, adding a measure to the space allows for a rich theory of dynamical systems on measure spaces.

Definition 2.20. Let $(X, \mathcal{B}, \mu, T)$ be a measure space with a measurable map $T : X \to X$. When T is measure-preserving as in Definition 2.9 (that is, when $T_*\mu = \mu$), we say that T *preserves* μ and that μ is T-*invariant*, and we call the tuple $(X, \mathcal{B}, \mu, T)$ a *measure-preserving system*.

It is important to know if the underlying measure space for a measure-preserving system is finite or infinite; finiteness of the measure yields a much more manageable situation for the dynamics. Of course, because a finite measure may be normalized by dividing by the measure of the space, the study of dynamics preserving finite measures reduces to the study of dynamics preserving probability measures. Where possible, we will still attempt to state results for both finite and infinite measures.

Showing that a map is measure-preserving would be challenging if not for the following result for finite measure spaces [31, p. 20].

Proposition 2.21. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system with $\mu(X) < \infty$, and let $\mathcal{P}$ be a semi-algebra which generates $\mathcal{B}$. If T preserves the measure of every set in $\mathcal{P}$, then T preserves μ.*

Proof. We shall show that

$$\mathcal{M} = \left\{B \in \mathcal{B} \ : \ \mu(T^{-1}(B)) = \mu(B)\right\} = \mathcal{B}.$$

First, let $\mathcal{A}(\mathcal{P})$ denote the algebra generated by $\mathcal{P}$. Since $\mathcal{P}$ is a semi-algebra, we have a very particular structure for $\mathcal{A}(\mathcal{P})$; we have that any set in $\mathcal{A}(\mathcal{P})$ can be written as a finite disjoint union of sets in $\mathcal{P}$ (see Lemma 2.5). So for $E \in \mathcal{A}(\mathcal{P})$, we have:

$$\begin{aligned}\mu(T^{-1}(E)) = \mu(T^{-1}(C_1 \cup \cdots \cup C_n)) &= \mu(T^{-1}(C_1) \ldots T^{-1}(C_n)) \\ = \sum_{i=1}^{n} \mu(T^{-1}(C_i)) &= \sum_{i=1}^{n} \mu(C_i) = \mu(E),\end{aligned}$$

since T preserves disjointness of unions, hence $\mathcal{A}(\mathcal{P}) \subset \mathcal{M}$. Finally, let $\{M_i\}_{i=1}^{\infty}$ be an increasing sequence of sets inside $\mathcal{M}$. Then we have:

$$\begin{aligned}\mu\left(T^{-1}\left(\bigcup_{i=1}^{\infty} M_i\right)\right) = \mu\left(\bigcup_{i=1}^{\infty} T^{-1}(M_i)\right) &= \lim_{i\to\infty} \mu(T^{-1}(M_i)) \\ &= \lim_{i\to\infty} \mu(M_i) = \mu\left(\bigcup_{i=1}^{\infty} M_i\right),\end{aligned}$$

because pulling back under T preserves order by inclusion. So we obtain $\bigcup_{i=1}^{\infty} M_i \in \mathcal{M}$. Similarly, if $\{M_i\}_{i=1}^{\infty}$ is a decreasing sequence of sets inside of $\mathcal{M}$, then since each of the sets has finite measure, we have:

$$\begin{aligned}\mu\left(T^{-1}\left(\bigcap_{i=1}^{\infty} M_i\right)\right) &= \mu\left(\bigcap_{i=1}^{\infty} T^{-1}(M_i)\right) = \lim_{i\to\infty} \mu(T^{-1}(M_i)) \\ &= \lim_{i\to\infty} \mu(M_i) = \mu\left(\bigcap_{i=1}^{\infty} M_i\right),\end{aligned}$$

again by continuity along chains of the measure, so we get $\bigcap_{i=1}^{\infty} M_i \in \mathcal{M}$. Thus, $\mathcal{M}$ is a monotone class. Then the monotone class generated by $\mathcal{A}(\mathcal{P})$, denoted $\mathcal{M}(\mathcal{A}(\mathcal{P}))$, is contained by $\mathcal{M}$, since $\mathcal{A}(\mathcal{P}) \subset \mathcal{M}$. But by the Monotone Class Theorem (Theorem 2.6), we have that $\mathcal{M}(\mathcal{A}(\mathcal{P})) = \sigma(\mathcal{A}(\mathcal{P})) = \mathcal{B}$, and so $\mathcal{B} \subset \mathcal{M} \subset \mathcal{B}$, forcing $\mathcal{B} = \mathcal{M}$, as desired. □

Let us look at some examples which will arise later.

Example 2.22. Let $X = \mathbb{T} = \mathbb{R}/\mathbb{Z}$, $\mathcal{B}$ be the Borel sets, and λ be normalized Lebesgue measure. Let $\eta \in \mathbb{T}$, and let $T(x) = x + \eta \mod 1$. T is typically called a rotation, as applying T is rotating the unit interval like a circle. We may show that T is measure-preserving by showing that any T preserves the measure of any interval, because the set of intervals form a semi-algebra which generates $\mathcal{B}$. Note that on the circle, λ may be described for intervals as

$$\lambda(a,b) = \begin{cases} b-a & b > a \\ 1-(a-b) & b < a \end{cases}.$$

Then we have:

$$\begin{aligned}
\lambda(T^{-1}(a,b)) &= \lambda(a-\eta, b-\eta) \\
&= \begin{cases} b-\eta-(a-\eta) & b-\eta > a-\eta \\ 1-(a-\eta-(b-\eta)) & b-\eta < a-\eta \end{cases} \\
&= \begin{cases} b-a & b > a \\ 1-(a-b) & b < a \end{cases} \\
&= \lambda(a,b).
\end{aligned}$$

Thus T is measure-preserving by Proposition 2.21, and so $(\mathbb{T}, \mathcal{B}, \lambda, T)$ is a measure-preserving system.

Whether η is rational or irrational determines many other properties of the system. We shall only have need for the case that η is irrational. We do, however, require a slight generalization of this system.

Example 2.23. Consider the same measure space as the previous example, but this time define the map $F(x) = -x$ (or $1-x$, depending on how explicit we wish to be).

This is clearly measurable; if (a, b) is an interval in $\mathbb{T}$, we have:

$$\begin{aligned}
\lambda(F^{-1}(a,b)) &= \lambda(-b,-a) \\
&= \begin{cases} -a-(-b) & -a > -b \\ 1-(-b-(-a)) & -a < -b \end{cases} \\
&= \begin{cases} b-a & b > a \\ 1-(a-b) & b < a \end{cases} \\
&= \lambda(a,b).
\end{aligned}$$

Hence F preserves Lebesgue measure on $\mathbb{T}$. This in itself is not useful, but given a rotation $T(x) = x + \eta$ on $\mathbb{T}$, it allows us to say that the map $T \circ F : \mathbb{T} \to \mathbb{T}$, given explicitly by $T(F(x)) = \eta - x$, is measure-preserving (it is trivial to see that the composition of transformations preserving a single measure on a space still preserves that measure).

The example of a rotation on the unit interval is a specific example of the following more general result which applies to compact groups.

Example 2.24. Let G be a compact topological group, with $\mathcal{B}$ the Borel σ-algebra, m the normalized Haar measure, and let $g \in G$. Since m is invariant under left-translation, the map $R_g(x) = gx$ preserves m, and therefore the system $(G, \mathcal{B}, m, R_g)$ is measure-preserving.

A more exotic example is the following.

Example 2.25 (Bernoulli Shift)**.** Consider the space $(X, \mathcal{A}, \mu)$ defined in Example 2.11. We define the map $L : X \to X$, by

$$L(x) = L(\dots x_{-1} \cdot x_0 x_1 \dots) = (\dots x_0 \cdot x_1 x_2 \dots),$$

so that $L(x)_n = x_{n+1}$. L is the *left shift* on X. L is clearly invertible on X, with inverse $L^{-1} = R$, where R is the analogously defined right shift, with $R(x)_n = x_{n-1}$.

To show that $(X, \mathcal{A}, \mu, L)$ is a measure-preserving system, we must show that L is measurable and preserves μ; it suffices to prove both facts for cylinder sets. Let $F \subset \mathbb{Z}$, $B_i \subset A$ for each $i \in F$, and let

$$C = \bigcap_{i \in F} \pi_i^{-1}(B_i).$$

For $m \in \mathbb{Z}$, denote $F + m = \{n \in \mathbb{Z} \ : \ n = f + m, f \in F\}$. Then we have:

$$\begin{aligned}
L^{-1}(C) &= \{x \in X \ : \ L(x) \in C\} = \{x \in X \ : \ L(x)_i \in B_i, \ i \in F\} \\
&= \{x \in X \ : \ x_{i+1} \in B_i, \ i \in F\} = \{x \in X \ : \ x_j \in B_{j-1}, j \in F+1\} \\
&= \bigcap_{j \in F+1} \pi_j^{-1}(B_{j-1}),
\end{aligned}$$

which shows that L is measurable. To show that it preserves measure, we see that:

$$\begin{aligned}
\mu(L^{-1}(C)) = \mu\left(\bigcap_{j \in F+1} \pi_j^{-1}(B_{j-1})\right) &= \prod_{j \in F+1} \nu(B_{j-1}) \\
= \prod_{j \in F+1} \nu(B_{j-1}) &= \prod_{i \in F} \nu(B_i) = \mu(C).
\end{aligned}$$

Hence L preserves μ.

Finally, we can prove that given two measure-preserving systems, the product of the two maps over the product measure space is again a measure-preserving system (it's actually a product in the category of measure-preserving systems, even, but we will not discuss this in detail in this thesis).

Example 2.26 (Product Measure-Preserving System)**.** Consider measure-preserving systems $(X, \mathcal{A}, \mu, S)$ and $(Y, \mathcal{B}, \nu, T)$, and let $S \times T$ be defined on the product measure space $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \times \nu)$. Let $A \in \mathcal{A}, B \in \mathcal{B}$, and observe that:

$$\begin{aligned}(S \times T)^{-1}(A \times B) &= \{(x, y) \ : \ S(x) \in A, \ T(y) \in B\} \\ &= \left\{(x, y) \ : \ x \in S^{-1}(A), \ y \in T^{-1}(B)\right\} \\ &= S^{-1}(A) \times T^{-1}(B),\end{aligned}$$

since the components are independent. Then we have:

$$\begin{aligned}\mu \times \nu((S \times T)^{-1}(A \times B)) = \mu \times \nu(S^{-1}(A) \times T^{-1}(B)) &= \mu(S^{-1}(A))\nu(T^{-1}(B)) \\ = \mu(A)\nu(B) &= \mu \times \nu(A \times B).\end{aligned}$$

Since measurable rectangles are a generating semi-algebra for $\mathcal{A} \otimes \mathcal{B}$, we have that $S \times T$ preserves $\mu \times \nu$, so that $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \times \nu, S \times T)$ is a measure-preserving system.

We may build upon our notion of measure-theoretic isomorphism to obtain a notion of when two measure-preserving systems are 'the same'. This notion will allow us to make so-called 'coordinate' changes later on, and help us to understand more complicated maps.

Definition 2.27. We say that the two measure-preserving systems $(X, \mathcal{A}, \mu, S)$ and $(Y, \mathcal{B}, \nu, T)$ are *isomorphic* if there exists a measure-theoretic isomorphism h from X to Y such that h satisfies $h \circ S = T \circ h$ (and such that the sets X_1 and Y_1 in Definition 2.16 are S- and T-invariant, respectively). h is then called an *isomorphism of measure-preserving systems* or a *measure-theoretic dynamical isomorphism*.

Example 2.28. Let $\beta \in \mathbb{T}$ be irrational, and consider the irrational rotation systems $(\mathbb{T}, \mathcal{B}, \lambda, T_\beta)$ and $(\mathbb{T}, \mathcal{B}, \lambda, T_{1-\beta})$, where $T_{1-\beta}$ is the rotation by $1 - \beta$. These measure-preserving systems are essentially the same:

Lemma 2.29. *The map $F(x) = 1 - x$ is a measure-theoretic isomorphism between the systems $(\mathbb{T}, \mathcal{B}, \lambda, T_\beta)$ and $(\mathbb{T}, \mathcal{B}, \lambda, T_{1-\beta})$.*

Proof. Since $F^{-1} = F$, we simply compute $F \circ T_\beta \circ F$:

$$\begin{aligned} F \circ T_\beta \circ F(x) &= F \circ T_\beta(1-x) \\ &= \begin{cases} F(1-x+\beta) & 1-x < 1-\beta \\ F(\beta - x) & 1-x \geq 1-\beta \end{cases} \\ &= \begin{cases} 1-(1-x+\beta) & x > \beta \\ 1-(\beta - x) & x \leq \beta \end{cases} \\ &= \begin{cases} (x-\beta)+1 & x > \beta \\ x+1-\beta & x \leq \beta \end{cases} \\ &= x+1-\beta = T_{1-\beta}(x). \end{aligned}$$

It is easy to see that F is an measure-preserving transformation on $(\mathbb{T}, \lambda)$, so as it is invertible, it is a measure isomorphism on $(\mathbb{T}, \lambda)$. Thus we are done. □

We also can talk about mapping one dynamical system to another, in a way which commutes with the respective dynamics, but isn't necessarily an isomorphism.

Definition 2.30. Let $(X, \mathcal{A}, S)$ be a dynamical system. We say that $(Y, \mathcal{B}, T)$ is a *factor* of $(X, \mathcal{A}, \mu, S)$ when there exists a measurable map $h : X \to Y$ (not necessarily invertible) such that $h \circ T = S \circ h$. In the case of two measure-preserving systems on probability spaces, we may also require h to be measure-preserving, so that $h_*\mu = \nu$.

The following important definitions allows us to delve much further into the study of dynamical systems.

Definition 2.31. Let $(X, \mathcal{B}, T)$ be a dynamical system. We say that $A \in \mathcal{B}$ is a T-*invariant* set if $T^{-1}(A) = A$.

Definition 2.32. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, and let A be a measurable set, with $\mu(A) > 0$. Recall that the collection of subsets $\mathcal{B}_A = \{B \cap A \ : \ B \in \mathcal{B}\}$ is a σ-algebra on A. We may define a set function $\mu_A : \mathcal{B}_A \to \mathbb{R}$ by setting $\mu_A(B \cap A) = \mu(B \cap A)$. It is easy to see that μ_A is a measure on $\mathcal{B}_A$; we call it the *restriction of* μ *to* A. If $\mu(A) < \infty$, then we may normalize μ_A to get $\mu_A(B \cap A) = \dfrac{\mu(B \cap A)}{\mu(A)}$, which is a probability measure. Note that if $B \subset A$ then $\mu_A(B) = \mu_A(B \cap A)$, so we will abuse notation and use the former on occasion.

Moreover, if A is T-invariant, we see that for $B \cap A \in \mathcal{B}_A$:

$$\begin{aligned}\mu_A(T^{-1}(B \cap A)) &= \mu(T^{-1}(B \cap A) \cap A) = \mu(T^{-1}(B) \cap T^{-1}(A) \cap T^{-1}(A)) \\ &= \mu(T^{-1}(B \cap A) = \mu(B \cap A) = \mu_A(B \cap A).\end{aligned}$$

Thus T preserves μ_A, and $(A, \mathcal{B}_A, \mu_A, T)$ is a measure-preserving system (the normalized version is immediately seen to be preserved also).

Definition 2.33. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system. The measure and map pair (μ, T) is *ergodic* if for any T-invariant set $A \in \mathcal{B}$, either $\mu(A) = 0$ or $\mu(X \setminus A) = 0$.

Given a measure-preserving system and an invariant set A for the dynamics, note that the complement of A, $X \setminus A$, is also T-invariant:

$$T^{-1}(X \setminus A) = T^{-1}(X) \setminus T^{-1}(A) = X \setminus A.$$

If the measure of A is neither zero nor full, then neither is the measure of $X \setminus A$, and we obtain two separate measure-preserving systems which sit inside of X, as in

Definition 2.32. The dynamics on X may then be studied simply by studying the separate dynamics on A and $X \setminus A$; in particular, if X is a probability space, then for any measurable set B of X, we have

$$\mu(B) = \mu(B \cap A) + \mu(B \cap (X \setminus A)) = \mu(A)\mu_A(B \cap A) + \mu(X \setminus A)\mu_{X\setminus A}(B \cap A),$$

so that μ decomposes as a convex combination of μ_A and $\mu_{X\setminus A}$.

The statement that this cannot happen is exactly that T is ergodic with respect to μ, and is a natural next step up from measure-preserving systems. That said, ergodicity is the weakest formulation of 'mixing' for measure-preserving transformations [11]; we will see a much stronger formulation later. As an interesting aside, it should be noted that given a fixed map T on a 'nice enough' space, one may develop an 'ergodic decomposition' for T, which represents any preserved measure μ for T as an 'weighted average' of the ergodic measures for T (of course, this average may be an integral, in the case when there are uncountably many ergodic measures), because ergodic measures are extreme points of the convex set of T-invariant measures (see [18]).

Among the first major results in ergodic theory are the so-called 'Ergodic Theorems' proved by Birkhoff and von Neumann in the 1930s [4, 30]. We will state Birkhoff's theorem here, as it is used to prove a characterization of ergodicity soon to be presented, and some of the ideas of the theorem arise later in a somewhat different context, which provides a useful analogy.

Theorem 2.34 (Birkhoff)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system on a σ-finite measure space, and let f be an integrable function. Then*

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

converges for almost every $x \in X$*, and the resulting function* $\tilde{f}$ *is in* L^1*, with* $\tilde{f} \circ T = \tilde{f}$ *almost everywhere and* $\|\tilde{f}\|_1 \leq \|f\|_1$*. If* $\mu(X) < \infty$*, then*

$$\int_X \tilde{f} \, d\mu = \int_X f \, d\mu.$$

Corollary 2.35. *If* $(X, \mathcal{B}, \mu, T)$ *is ergodic, then* $\tilde{f}$ *as obtained above is almost everywhere constant, and if* $\mu(X) < \infty$*, then*

$$\tilde{f} = \frac{1}{\mu(X)} \int_X f \, d\mu.$$

The resulting function $\tilde{f}$ is the asymptotic *time average* of the function f on orbits starting at the point x. The corollary states that for ergodic maps on finite measure spaces, the time average equals the space average almost everywhere. In general, it can be shown that $\tilde{f}$ is the conditional expectation of f with respect to the σ-algebra of T-invariant measurable subsets of X, which is constant for ergodic systems (as any T-invariant sets are either null sets or are the whole space).

We would like to know when a system is ergodic; there are many characterisations of ergodicity, some applying only in certain situations, and the ones which we shall use here are as follows. We leave the proof to the appendix. We do need to state another definition, first.

Definition 2.36. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, $A \in \mathcal{B}$ is *almost-T-invariant* if $\mu(T^{-1}(A) \Delta A) = 0$.

Theorem 2.37. *Let* $(X, \mathcal{B}, \mu, T)$ *be a measure-preserving system, with possibly infinite measure. The following are equivalent:*

1. (μ, T) *is ergodic.*

2. *If* A *is almost* T*-invariant, then* $\mu(A) = 0$ *or* $\mu(X \setminus A) = 0$*.*

3. If $f \in L^\infty$, and $f \circ T = f$ almost everywhere, then f is constant almost everywhere.

If moreover $\mu(X) = 1$, then the above are also equivalent to the following:

4. If $f \in L^2(X)$, and $f \circ T = f$ almost everywhere, then f is constant almost everywhere.

5. If $A, B \in \mathcal{P}$, where $\mathcal{P}$ is a π-system which generates the σ-algebra $\mathcal{B}$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B).$$

It is worthwhile to have an intuitive grasp of these equivalent conditions. Condition (1) is the definition, while condition (2) states that ergodicity is really a modulo 0 concept, in that we can safely ignore sets of zero measure. Condition (3) is a functional definition of ergodicity, which is useful when it is more convenient to consider functions on the space rather than the space itself. Condition (3) may be modified, to require f to simply be measurable.

Condition (4) says that when X has finite measure, we may restrict ourselves to L^2 functions, which works very well when there exists an orthonormal basis for L^2 which behaves nicely with the map on the space. Condition (5) is a useful theoretical characterization for doing abstract computations, but which can be used in a concrete setting as well, as shall be shown.

A stronger formulation of the idea of 'mixing' is the following:

Definition 2.38. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system on a probability space. We say that T is *strongly mixing* with respect to μ if for any $A, B \in \mathcal{B}$, we have that

$$\mu(T^{-k}(A) \cap B) \underset{k\to\infty}{\longrightarrow} \mu(A)\mu(B).$$

There is an obvious strengthening of this definition, in that the defining property may be shown only on a π-system generating the σ-algebra, analogous to the one for ergodicity.

Proposition 2.39. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, and let $\mathcal{P}$ be a generating π-system for $\mathcal{B}$. T is strongly mixing with respect to μ if and only if for every $A, B \in \mathcal{P}$, we have*

$$\mu(T^{-k}(A) \cap B) \underset{k\to\infty}{\longrightarrow} \mu(A)\mu(B).$$

The proof is similar, and may be found in the appendix (in A.2). One may compare this to condition (5) above; we can say that T is ergodic if two sets mix on the average. For concreteness, we may show the following proposition:

Proposition 2.40. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system on a probability space. If T is strongly mixing with respect to μ, then T is also ergodic.*

Proof. The idea for this proof comes from [25]. Let E be a T-invariant measurable set, so $E = T^{-1}(E)$. Iterating we obtain $E = T^{-k}(E)$, and so we obtain:

$$\mu(E) = \mu(E \cap E) = \mu(E \cap T^{-k}(E)) \underset{k\to\infty}{\longrightarrow} \mu(E)\mu(E) = \mu(E)^2,$$

using the mixing property, and so $\mu(E) \in \{0, 1\}$, thereby proving that T is ergodic. □

This shows that ergodicity is a potentially weaker property than strong mixing. However, ergodicity is still sufficient for many applications, as we will see. It must also be noted that ergodicity and strong mixing are invariants of measure-theoretic dynamical isomorphisms:

Lemma 2.41. *Let* $(X, \mathcal{A}, \mu, S)$ *and* $(Y, \mathcal{B}, \nu, T)$ *be isomorphic measure-preserving systems. If* $(X, \mathcal{B}, \mu, T)$ *is ergodic, so is* $(Y, \mathcal{A}, \nu, S)$*. If* $\mu(X), \nu(Y) < \infty$*, then if* $(X, \mathcal{A}, \mu, S)$ *is strongly mixing, so is* $(Y, \mathcal{B}, \nu, T)$*.*

Proof. For the first claim, suppose that (μ, S) is ergodic, and let $B \in \mathcal{B}$ be T-invariant. After dropping down to sets of full measure, let $h : X \to Y$ be the dynamical isomorphism. Observe that $h^{-1}(B)$ is S-invariant:

$$T^{-1}(h^{-1}(B) = (h \circ T)^{-1}(B) = (S \circ h)^{-1}(B) = h^{-1}(S^{-1}(B)) = h^{-1}(B).$$

Thus either

$$\nu(B) = \mu(h^{-1}(B)) = 0$$

or

$$\nu(Y \setminus B) = \mu(h^{-1}(Y \setminus B)) = \mu(X \setminus h^{-1}(B)) = 0,$$

so that (ν, T) is ergodic, as desired.

For the second claim, suppose that (μ, S) is strongly mixing, and let $B, D \in \mathcal{B}$. Note that $h^{-1}(T^{-k}(B)) = S^{-k}(h^{-1}(B))$, by induction. Then we have:

$$\begin{aligned}\nu(T^{-k}(B) \cap D) &= \mu(h^{-1}(T^{-k}(B) \cap D)) = \mu(h^{-1}(T^{-k}(B)) \cap h^{-1}(D)) \\ &= \mu(S^{-k}(h^{-1}(B)) \cap h^{-1}(D)) \underset{k\to\infty}{\longrightarrow} \mu(h^{-1}(B))\mu(h^{-1}(D)) = \nu(B)\nu(D).\end{aligned}$$

Hence (ν, T) is strongly mixing, as desired. ☐

We now give some concrete examples of ergodicity and strong mixing.

Example 2.42. We will show that the irrational rotation on the circle by $\eta \in Q^c$, as introduced in Example 2.22, is an ergodic system, but not strongly mixing. To show that (T, μ) is ergodic, we will use characterization (4) above. Let $f : \mathbb{T} \to \mathbb{C}$ be

T-invariant. Using the Fourier series for f, we obtain, for almost every $x \in \mathbb{T}$:

$$\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x} = f(x) = f(T(x)) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n T(x)} = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n \eta} e^{2\pi i n x}.$$

Hence we obtain a relation in the coefficients, $c_n = c_n e^{2\pi i n\eta},\ n \in \mathbb{Z}$. This forces $c_n = 0$ for $n \neq 0$, as η is not rational, thus $e^{2\pi i n\eta}$ cannot be equal to 1. Thus f is constant almost everywhere, and we see that T is ergodic.

To show that (T, μ) is not strongly mixing, let A be the interval $(0, \frac{1}{3})$, so its length is $\frac{1}{3}$. We will show that there are infinitely many positive integers n such that $T^{-n}(A) \cap A$ is empty. Observe that

$$T^{-k}(A) = (-k\eta, \frac{1}{3} - k\eta),$$

and note that $\{-k\eta \ :\ k \in \mathbb{N}\}$ is dense in the unit interval when taken modulo 1. Then $-k\eta \in [\frac{1}{3}, \frac{2}{3})$ infinitely often, in which case $T^{-k}(A) \subset [\frac{1}{3}, 1)$, which implies that $T^{-k}(A) \cap A = \emptyset$.

Then we obtain

$$\liminf_{k\to\infty} \mu(T^{-k}(A) \cap A) = 0,$$

which implies that

$$\lim_{k\to\infty} \mu(T^{-k}(A) \cap A) \neq \mu(A)^2 = \frac{1}{9},$$

if the limit even exists, and so (T, μ) is not strongly mixing.

For certain groups equipped with Haar measure, we have a sufficient condition for the ergodicity of a rotation. (See Appendix A.3 for a brief recap of character groups.)

Proposition 2.43. *Let $(G, \mathcal{B}, m)$ be a compact topological Abelian group (written multiplicatively) equipped with normalized Haar measure, with character group $\hat{G}$. Let $a \in G$, and define the measure-preserving transformation R_a on G by $R_a(x) = ax$.*

If for every non-trivial character $\gamma \in \hat{G}$ we have $\gamma(a) \neq 1$, then R_a is ergodic with respect to m.

Proof. For any character γ and $x \in G$, we have

$$\gamma(R_a(x)) = \gamma(ax) = \gamma(a)\gamma(x).$$

We know that $\hat{G}$ is an orthonormal basis for $L^2(m)$; we will do a Fourier series computation to obtain our result. Let $f \in L^2(m)$ be R_a-invariant, so that $f \circ R_a = f$. Then we obtain:

$$\sum_{\gamma \in \hat{G}} b_\gamma \gamma(a)\gamma(x) = \sum_{\gamma \in \hat{G}} b_\gamma \gamma(ax) = f \circ R_a(x) = f(x) = \sum_{\gamma \in \hat{G}} b_\gamma \gamma(x).$$

Since the characters are orthonormal, we obtain $b_\gamma \gamma(a) = b_\gamma$ for all $\gamma \in \hat{G}$. This is equivalent to $b_\gamma(\gamma(a) - 1) = 0$; by hypothesis, for all non-trivial γ, we have $\gamma(a) \neq 1$. This implies $b_\gamma = 0$ for all non-trivial γ, and so f must be constant. Applying condition (4) from Theorem 2.37, we get that R_a is ergodic. $\square$

Example 2.44. Let $(X, \mathcal{B}, \mu, L)$ be the invertible left shift on a shift space with a countable alphabet A. We will show that this is a strongly mixing system. Recall from the end of Example 2.11 that a π-system generating $\mathcal{B}$ is the collection $\mathcal{D}$ of contiguous string cylinder sets spanning the 0 index. If we can show that the mixing property in Definition 2.38 holds for two of those sets, we are finished.

Let $t_1, t_2 \geq 0$ and $r_1, r_2 \geq 1$, and define

$$B = C(x_{-t_1} \dots x_{r_1 - 1} = a_{-t_1} \dots a_{r_1 - 1}), \quad D = C(x_{-t_2} \dots x_{r_2 - 1} = b_{-t_2} \dots b_{r_2 - 1}).$$

For any $k \in \mathbb{Z}$, we have:

$$\begin{aligned} L^{-k}(B) &= C(L^k(x)_{-t_1} \dots L^k(x)_{r_1-1} = a_{-t_1} \dots a_{r_1-1}) \\ &= C(x_{-t_1+k} \dots x_{r_1-1+k} = a_{-t_1} \dots a_{r_1-1}). \end{aligned}$$

So for $k \geq r_2 + t_1$, we have $-t_1 + k \geq r_2 > r_2 - 1$, and so

$$\{-t_2, \dots, r_2 - 1\} \cap \{-t_1 + k, \dots, r_1 - 1 + k\} = \emptyset.$$

By the projection property of μ as in Example 2.11, we obtain:

$$\lim_{k\to\infty} \mu(L^{-k}(B) \cap D) = \lim_{k\to\infty} \mu(L^{-k}(B))\mu(D) = \lim_{k\to\infty} \mu(B)\mu(D) = \mu(B)\mu(D),$$

and L is strongly mixing.

The above argument will return in a slightly different form, later on in the work, thanks to the usefulness of the π-system $\mathcal{D}$.

There are a number of important constructions in ergodic theory. We shall deal with two of them, in particular: the induced transformation, and the skew product. These will prove to be useful for both conceptual and computational purposes. See [5] for a related treatment of each of these constructions.

First, we build towards the induced transformation. A result by Poincaré tells us about how often orbits of points starting in a set of positive measure return to that set, in a measure-preserving system on a probability space.

Theorem 2.45 (Poincaré Recurrence)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system with $\mu(X) = 1$, and let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then for $\mu-a.e.$ $x \in A$, x returns infinitely often to A during its orbit under T.*

Proof. Let

$$\begin{aligned} B &= \{x \in A \;:\; x \text{ never returns to } A\} \\ &= \left\{x \in A \;:\; T^k(x) \notin A,\ \forall k \geq 1\right\} \\ &= A \cap \bigcap_{k=1}^{\infty} T^{-k}(X \setminus A). \end{aligned}$$

Consider pre-images of B under T^i; these are the points in X which return to at A in i steps (or fewer) and then never return to A after that. If $x \in T^{-i}(B) \cap T^{-j}(B)$ with $0 \leq i < j$, then $T^i(x) \in B$ and $T^j(x) = T^{j-i}(T^i(x)) \in A$, which means that $T^i(x)$ is simultaneously in B and not in B. Hence the pre-images of B are disjoint, and so we have:

$$1 \geq \mu\left(\bigcup_{i=0}^{\infty} T^{-i}(B)\right) = \sum_{i=0}^{\infty} \mu(T^{-i}(B)) = \sum_{i=0}^{\infty} \mu(B),$$

which forces $\mu(B) = 0$. Finally, the set of all points in A which return only finitely many times to A is

$$A \cap \bigcup_{i=0}^{\infty} T^{-i}(B) = \bigcup_{i=0}^{\infty} A \cap T^{-i}(B),$$

and we have

$$\mu\left(\bigcup_{i=0}^{\infty} A \cap T^{-i}(B)\right) \leq \sum_{i=0}^{\infty} \mu(A \cap T^{-i}(B)) \leq \sum_{i=0}^{\infty} \mu(T^{-i}(B)) = \sum_{i=0}^{\infty} \mu(B) = 0,$$

proving the claim. □

If T is not just measure-preserving with respect to μ, but also ergodic, then we see that almost every point in X reaches a set of non-zero measure in finite time, under iteration by T.

Lemma 2.46. *Let (T, μ) be an ergodic map and measure on a probability space, and* *let* $\mu(A) > 0$. *Then*

$$\mu\left(\bigcup_{n=0}^{\infty} T^{-n}(A)\right) = 1.$$

Proof. Let

$$B = \bigcup_{n=0}^{\infty} T^{-n}(A)$$

be the set of points which map to A in finite time. Note that

$$T^{-1}(B) = T^{-1}\left(\bigcup_{n=0}^{\infty} T^{-n}(A)\right) = \bigcup_{n=0}^{\infty} T^{-(n+1)}(A) = \bigcup_{n=1}^{\infty} T^{-n}(A) \subset B.$$

Then we have

$$\begin{aligned}\mu(T^{-1}(B)\Delta B) &= \mu(T^{-1}(B) \setminus B) + \mu(B \setminus T^{-1}(B)) \\ &= \mu(\emptyset) + \mu(B) - \mu(T^{-1}(B)) = 0,\end{aligned}$$

since T preserves μ. Hence B is almost-T-invariant; by condition (2) of Theorem 2.37, we see that B has either full or zero measure, but $A \subset B$, so $\mu(B) \geq \mu(A) > 0$, and so $\mu(B) = 1$, as desired. □

Recall that by Definition 2.32, if A is a measurable subset of X with $0 < \mu(A) < \infty$, then we may form a probability space $(A, \mathcal{A}, \mu_A)$ together with the restricted σ-algebra $\mathcal{A}$ on A and the restricted measure μ_A, where for $B \in \mathcal{A}$, $\mu_A(B) = \dfrac{\mu(B)}{\mu(A)}$, forms a probability space $(A, \mathcal{A}, \mu_A)$. If A is not T-invariant, then points in A may leave A by iteration of T.

Definition 2.47. For a subset A of a measure-preserving system $(X, \mathcal{B}, \mu, T)$, define the *first return time* $n_{A,T} : A \to \mathbb{N}_+ \cup \{\infty\}$ to be:

$$n_{A,T}(x) = \min\left\{n \geq 1 \ : \ T^n(x) \in A\right\},$$

which may be infinite if x never returns to A. The Poincaré Recurrence Theorem (Theorem 2.45) shows that $n_{A,T}$ is finite almost everywhere on A, and moreover that the set of points where $n_{A,T}$ is infinite is T-invariant. So after removing this invariant null set from A (without changing the notation), we may define the *induced map* of T on A by

$$T_A : A \to A,\ T_A(x) = T^{n_{A,T}(x)}(x).$$

In the simple case when A is T-invariant, then $n_{A,T}$ is identically 1, so that $T_A = T$ and the restricted system is still measure-preserving, as in Definition 2.32. We can prove that this holds in general.

Proposition 2.48. *If* $(X, \mathcal{B}, \mu, T)$ *is a measure-preserving space with* $\mu(X) = 1$ *and* A *is a measurable set with* $0 < \mu(A) < \infty$*, then* $(A, \mathcal{A}, \mu_A, T_A)$ *is a measure-preserving system on a probability space.*

Proof. First, we show that $n_{A,T}$ is measurable, from $(A, \mathcal{A})$ to $(\mathbb{N}_+, \mathcal{P}(\mathbb{N}_+))$. We have, for $k \geq 1$,

$$\begin{aligned} n_{A,T}^{-1}\{k\} &= \big\{x \in A \ :\ T^k(x) \in A, T^i(x) \notin A\ \forall i = 1, \ldots, k-1\big\} \\ &= A \cap T^{-k}(A) \setminus \bigcup_{i=1}^{k-1} T^{-i}(A), \end{aligned}$$

and T is measurable, so $n_{A,T}$ is measurable. For any $B \in \mathcal{A}$, we have:

$$\begin{aligned} T_A^{-1}(B) &= \big\{x \in A \ :\ T^{n_{A,T}(x)}(x) \in B\big\} \\ &= \bigcup_{k=1}^{\infty} (n_{A,T}^{-1}\{k\} \cap T^{-k}(B)), \end{aligned}$$

so T_A is measurable. To show that T_A preserves the measure μ_A, we compute, for

$B \in \mathcal{A}$:

$$\begin{aligned}
\mu(B) &= \mu(T^{-1}(B)) = \mu(T^{-1}(B) \cap A) + \mu(T^{-1}(B) \setminus A) \\
&= \mu(T^{-1}(B) \cap A) + \mu(T^{-2}(B) \setminus T^{-1}(A)) \\
&= \mu(T^{-1}(B) \cap A) + \mu((T^{-2}(B) \setminus T^{-1}(A)) \cap A) + \mu((T^{-2}(B) \setminus T^{-1}(A)) \setminus A) \\
&= \mu(T^{-1}(B) \cap A) + \mu((T^{-2}(B) \setminus T^{-1}(A)) \cap A) + \mu(T^{-2}(B) \setminus (T^{-1}(A) \cup A)) \\
&= \dots \\
&= \sum_{k=1}^{n} \mu\left(\left(T^{-k}(B) \setminus \bigcup_{i=1}^{k-1} T^{-i}(A)\right) \cap A\right) + \mu\left(T^{-n}(B) \setminus \bigcup_{i=0}^{n-1} T^{-i}(A)\right).
\end{aligned}$$

Now, letting $\bar{A} = \bigcup\limits_{i=0}^{\infty} T^{-i}(A)$, we see that $T^{-n}(B) \subset T^{-n}(A) \subset \bar{A}$, so that:

$$\begin{aligned}
0 &\leq \limsup_{n\to\infty} \mu\left(T^{-n}(B) \setminus \bigcup_{i=0}^{n-1} T^{-i}(A)\right) \leq \limsup_{n\to\infty} \mu\left(\bar{A} \setminus \bigcup_{i=0}^{n-1} T^{-i}(A)\right) \\
&= \limsup_{n\to\infty} \left(\mu(\bar{A}) - \mu\left(\bigcup_{i=0}^{n-1} T^{-i}(A)\right)\right) \\
&= \mu(\bar{A}) - \lim_{n\to\infty} \mu\left(\bigcup_{i=0}^{n-1} T^{-i}(A)\right) \\
&= \mu(\bar{A}) - \mu(\bar{A}) = 0.
\end{aligned}$$

So we obtain

$$\sum_{k=1}^{\infty} \mu\left(\left(T^{-k}(B) \setminus \bigcup_{i=1}^{k-1} T^{-i}(A)\right) \cap A\right) = \mu(B),$$

which allows us to conclude:

$$\begin{aligned}\mu_A(T_A^{-1}(B)) &= \frac{1}{\mu(A)}\mu\left(\bigcup_{k=1}^{\infty} n_{A,T}^{-1}\{k\} \cap T^{-k}(B)\right)\\ &= \frac{1}{\mu(A)}\sum_{k=1}^{\infty}\mu\left(T^{-k}(B)\cap A\cap T^{-k}(A)\setminus\bigcup_{i=1}^{k-1}T^{-i}(A)\right)\\ &= \frac{1}{\mu(A)}\sum_{k=1}^{\infty}\mu\left(A\cap\left(T^{-k}(B)\setminus\bigcup_{i=1}^{k-1}T^{-i}(A)\right)\right)\\ &= \frac{1}{\mu(A)}\mu(B) = \mu_A(B).\end{aligned}$$

□

We shall be utilizing the following example and variations thereof in multiple situations throughout the rest of the paper.

Example 2.49. Let $\eta \in \mathbb{T}$ be irrational, and let $(\mathbb{T}, \mathcal{B}, \lambda, T)$ be the irrational rotation measure-preserving system, as in Example 2.22. Let $A = [1-\eta, 1)$; we shall explicitly compute the induced map T_A. We first need to compute the first return time to A. To do so, note that there exists $k \geq 1$ such that $k\eta < 1 < (k+1)\eta$, since η is irrational and in $\mathbb{T}$. Let $q = 2-(k+1)\eta$. Then for $x \in [1-\eta, q)$, we see that

$$\begin{aligned}x + k\eta &\in [1-\eta+k\eta, q+k\eta) = [2-\eta+(k\eta-1), 2-\eta),\\ x+(k+1)\eta &\in [1-\eta+(k+1)\eta, q+(k+1)\eta) = [2-\eta+((k+1)\eta-1), 2),\end{aligned}$$

hence $n_{A,T}(x) = k+1$, since $1 \leq 1+(k-1)\eta$ and $2-\eta = 1-\eta$ modulo 1. Similarly for $x \in [q, 1)$, we have:

$$\begin{aligned}x + k\eta &\in [q+k\eta, 1+k\eta) = [2-\eta, 2+(k\eta-1)),\\ x+(k+1)\eta &\in [q+(k+1)\eta, 1+(k+1)\eta) = [2, 2+((k+1)\eta-1)),\end{aligned}$$

which gives us $n_{A,T}(x) = k$, as $2 + (k\eta - 1) < 2$. So we obtain:

$$T_A(x) = \begin{cases} T^{k+1}(x) & x \in [1-\eta, q) \\ T^k(x) & x \in [q, 1) \end{cases}$$
$$= \begin{cases} x + (k+1)\eta - 1 & x \in [1-\eta, q) \\ x + k\eta - 1 & x \in [q, 1). \end{cases}$$

We write the map with a '-1' to indicate that this is exact, not just modulo 1. The induced measure, λ_A, is simply Lebesgue measure on $[1-\eta, 1)$ divided by η. As seen in Proposition 2.48, λ_A is preserved by T_A. Now, we could leave this system like this, but we can rewrite it in a much more intuitive form, using a change of coordinates. Let

$$\phi : [1-\eta, 1) \to \mathbb{T},\ \phi(x) = \frac{1-x}{\eta};$$

ϕ is then a homeomorphism between $[1-\eta, 1)$ and $\mathbb{T}$, with inverse given by

$$\phi^{-1} : \mathbb{T} \to [1-\eta, 1),\ \phi^{-1}(x) = 1 - \eta x.$$

ϕ acts to switch the interval around and expand it, and ϕ^{-1} acts to switch the interval around and compress it. Note then that the measure for the coordinate-changed dynamics is $\phi_* \lambda_A$. Since any interval (a, b) with $a > b$ may be decomposed as

$$(a, b) = (a, 1) \cup [0, b),$$

we may do all of our computations with intervals of the form $a < b$, by additivity of the measure and distribution of preimages over disjoint unions. So for intervals (a, b)

where $a < b$, we have:

$$\begin{aligned}\phi_*\lambda_A(a,b) &= \lambda_A(\phi^{-1}(a,b)) = \lambda_A(1-\eta b, 1-\eta a) \\ &= \frac{1-\eta a-(1-\eta b)}{\eta} \\ &= \frac{\eta}{\eta}(b-a) \\ &= \lambda(a,b).\end{aligned}$$

Thus ϕ maps $[1-\eta, 1)$ onto $\mathbb{T}$ with normalized Lebesgue measure. Now, we may compute the action of T_A as a map on $\mathbb{T}$, by calculating $S = \phi \circ T_A \circ \phi^{-1}$. Recall that $k\eta < 1 < (k+1)\eta$; this yields $k < \frac{1}{\eta} < k+1$, and we find that the fractional part of $\frac{1}{\eta}$ is $\{\frac{1}{\eta}\} = \frac{1}{\eta} - k$, which we shall denote β. Note that

$$\phi(q) = \frac{1-(2-(k+1)\eta)}{\eta} = k+1-\frac{1}{\eta} = 1-\beta.$$

Then we have:

$$
\begin{aligned}
\phi \circ T_A \circ \phi^{-1}(x) &= \phi \circ T_A(1-\eta x) \\
&= \begin{cases} \phi(1-\eta x + (k+1)\eta - 1) & \phi^{-1}(x) \in [1-\eta, q) \\ \phi(1-\eta x + k\eta - 1) & \phi^{-1}(x) \in [q, 1) \end{cases} \\
&= \begin{cases} \dfrac{1-(k+1-x)\eta}{\eta} & x \in \phi[1-\eta, q) = [1-\beta, 1) \\ \dfrac{1-(k-x)\eta}{\eta} & x \in \phi[q, 1) = [0, 1-\beta) \end{cases} \\
&= \begin{cases} x + \dfrac{1}{\eta} - k - 1 & x \in \phi[1-\eta, q) = [1-\beta, 1) \\ x + \dfrac{1}{\eta} - k & x \in \phi[q, 1) = [0, 1-\beta) \end{cases} \\
&= \begin{cases} x + \beta - 1 & x \in [1-\beta, 1) \\ x + \beta & x \in [0, 1-\beta) \end{cases}
\end{aligned}
$$

Computing this modulo 1, we see that $S(x) = x + \beta$, where $\beta = \left\{ \dfrac{1}{\eta} \right\}$ is irrational. Hence inducing an irrational rotation on an interval of length matching the rotation yields another irrational rotation.

In general, an irrational rotation induces an *interval exchange* map on three intervals; it is only in the special case where the rotation and the interval length are the same that we obtain a rotation. See [20, 29] for a much more indepth look at this type of system.

Example 2.50. Let $(X, \mathcal{A}, \mu, L)$ be the invertible left shift on the shift space over the two symbol alphabet $\{0, 1\}$. Let

$$A = C(x_0 = 1) = \pi_0^{-1}\{1\},$$

and let us compute the induced map L_A. First, we must compute the first-return time, $n_{A,L}$. We are interested in starting in A (so $x_0 = 1$), then counting how many steps it takes to return back to A. This is easily interpreted by saying that $n(x) = k$ when $x_1, \ldots, x_{k-1} = 0$ and $x_k = 1$, or if $k = 1$, then simply $x_1 = 1$. So $n_{A,L}(x) = k$ when $x \in C(x_0 x_1 \ldots x_{k-1} x_k = 1\underbrace{0\ldots0}_{k-1}1)$, and $n_{A,L}(x) = \infty$ when $x_k = 0$ for all $k > 0$. Similar considerations hold for the right shift $R = L^{-1}$ and $n_{A,R}$. Let B_+ be the set of points in A which have finitely many 1's in the positive direction, and B_- be the set of points in A which have finite many 1's in the negative direction; then we have:

$$B_+ = \bigcup_{i=0}^{\infty} L^{-i}(n_{A,L}^{-1}\{\infty\}), \quad B_- = \bigcup_{i=0}^{\infty} R^{-i}(n_{A,R}^{-1}\{\infty\}).$$

By subadditivity of the measure μ and the Poincaré Recurrence Theorem, both B_+ and B_- have measure 0, and it is clear that $B_+ \cup B_-$ is L-invariant (and R-invariant). Hence we may remove $B_+ \cup B_-$ from A to obtain $\tilde{A}$, which we shall still call A.

We now compute L_A. We have:

$$\begin{aligned} L_A(x) &= L^{n_{A,L}}(x) \\ &= \begin{cases} L^k(x), \quad x \in C(x_0 x_1 \ldots x_{k-1} x_k = 1\underbrace{0\ldots0}_{k-1}1) \end{cases} \\ &= \begin{cases} (\ldots 1\underbrace{0\ldots0}_{k-1}\cdot 1 x_{k+1} x_{k+2} \ldots), \quad x \in C(x_0 x_1 \ldots x_{k-1} x_k = 1\underbrace{0\ldots0}_{k-1}1) \end{cases} \end{aligned}$$

Essentially, we are skipping the next string of zeroes, and returning the string centered at the next one. This yields a measure-preserving system $(A, \mathcal{M}_A, m_A, L_A)$, where $\mathcal{M}_A$ is the restriction of $\mathcal{M}$ to A.

Here, we make a clever change of coordinates. Let $(Z, \mathcal{M}, \nu, S)$ be the invertible left Bernoulli shift space over the alphabet $\mathbb{N}$, where the weights on the symbols are $p_k = \dfrac{1}{2^{k+1}}$, corresponding to the symbol $k \geq 0$ (as in Example 2.11). Define the map

$\phi : A \to Z$, by the following. Note that elements x of A can be written as

$$(\dots 1\underbrace{0\dots0}_{k_{-1}}\cdot 1\underbrace{0\dots0}_{k_0}1\underbrace{0\dots0}_{k_1}\dots),$$

where $k_i \geq 0$, so we define

$$\phi(x) = (\dots k_{-1} \cdot k_0 k_1 \dots) \in Z.$$

To show that ϕ is measurable, let $B \subset \mathbb{N}$, and compute:

$$\begin{aligned}\phi^{-1}(\pi_n^{-1}(B)) &= \phi^{-1}(\{z \in Z \ : \ z_n \in B\}) \\ &= \big\{x \in A \ : \ (n+1)^{\text{th}} \text{ string of zeroes is length } j, j \in B\big\} \\ &= \bigcup_{j\in B} \ \bigcup_{k_0,\dots,k_{n-1}\in\mathbb{N}} C(x_0 \dots x_{n+\sum_i k_i + j+1} = 1\underbrace{0\dots0}_{k_0}1\dots1\underbrace{0\dots0}_{k_{n-1}}1\underbrace{0\dots0}_{j}1),\end{aligned}$$

which lies in $\mathcal{B}$, since the unions are both countable, and general cylinder sets are pulled back to intersections of these sets. ϕ is thus measurable, by definition of $\mathcal{M}_A$. Hence we may push m_A forward with ϕ, to obtain a measure $\phi_* m_A$ on Z. We now show that $\phi_* m_A = \nu$, by showing that $\phi_* m_A$ pushes forward to the product measure ν_F on any finite product Z_F. So let $F \subset \mathbb{Z}$ be finite, and consider the product measure ν_F on the product Z_F. We compute, for $(k_i)_{i\in F} \in Z_F$:

$$\begin{aligned}(\pi_F)_*\phi_* m_A(\{(k_i)_{i\in F}\}) &= m_A(\phi^{-1}(\pi_F^{-1}(\{(k_i)_{i\in F}\}))) \\ &= m_A(\big\{x \in A \ : \ b^{\text{th}} \text{ string of zeroes is of length } k_b, \ b \in F\big\}) \\ &= \prod_{b\in F} \frac{1}{2^{k_b+1}} \qquad \text{(by independence of the } \{0,1\} \text{ shift)} \\ &= \nu_F(\{(k_1, \dots, k_{|F|})\}).\end{aligned}$$

Any (measurable) set $C \subset Z_F$ is a disjoint union of points $(k_i)_{i\in F}$, so the above computation holds for C due to countable additivity. Hence $\phi_* m_A = \nu$, by uniqueness of the Kolmogorov extension measure. Therefore ϕ is a measure-theoretic isomorphism between $(A, \mathcal{A}, m_A)$ and $(Z, \mathcal{M}, \nu)$. From the definitions of ϕ and L_A, it is easy to observe that $\phi \circ L_A \circ \phi^{-1}$ acts as the left shift S on Z. Hence, ϕ gives a dynamical isomorphism between $(A, \mathcal{M}_A, m_A, L_A)$ and $(Z, \mathcal{M}, \nu, S)$.

The following proposition describes the relationship between ergodicity of the original map T and the induced map T_A.

Proposition 2.51. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system with $\mu(X) = 1$, and let $A \subset X$ with $\mu(A) > 0$. Let T_A be the induced map of T on A, as above. Then we have the following:*

1. *If T is ergodic, then T_A is ergodic.*
2. *If T_A is ergodic, and*
$$\mu\left(\bigcup_{n=0}^{\infty} T^{-n}(A)\right) = 1,$$
then T is ergodic.

Proof. For the first statement, assume that (T, μ) is ergodic. Let $B \subset A$ be measurable and T_A-invariant, and suppose that $\mu_A(B) = \dfrac{\mu(B)}{\mu(A)} > 0$, so $\mu(B) > 0$ also. Let

$$\bar{B} = \bigcup_{i=0}^{\infty} T^{-i}(B);$$

then by Lemma 2.46, $\bar{B}$ has measure 1 in X. We shall now show that $B = \bar{B} \cap A$. First, note that

$$\bar{B} \cap A = \bigcup_{i=0}^{\infty} T^{-i}(B) \cap A = \left\{x \in A \ :\ \exists\, i \geq 0 \text{ such that } T^i(x) \in B\right\};$$

the inclusion of B in $\bar{B} \cap A$ is trivial, since the first term of the union is $B \cap A = B$. To show the reverse inclusion, note that if $x \in \bar{B} \cap A$, then there exists $i \geq 0$ such that $T^i(x) \in B \subset A$. If $i = 0$, then we are done, so we assume that $i > 0$. If x makes $k \geq 0$ returns to A prior to the i^{th} iteration of T in the orbit, then we have that $T^i(x) = T_A^{k+1}(x)$. By the T_A-invariance of B, we know that $B = T_A^{-(k+1)}(B)$; we therefore see that

$$x \in T_A^{-(k+1)}(B) = B,$$

and so $\bar{B} \cap A \subset B$. Hence $B = \bar{B} \cap A$, so since $\mu(\bar{B}) = 1$, we have

$$\mu_A(B) = \frac{\mu(\bar{B} \cap A \cap A)}{\mu(A)} = \frac{\mu(A)}{\mu(A)} = 1,$$

and therefore (T_A, μ_A) is ergodic.

For the second statement, assume that (T_A, μ_A) is ergodic, and let B be a measurable T-invariant subset of X, with $\mu(B) > 0$. We have also assumed that

$$\bar{A} = \bigcup_{i=0}^{\infty} T^{-i}(A)$$

has full measure in X, which means that $\exists i \geq 0$ such that

$$\mu(B \cap A) = \mu(T^{-i}(B \cap A)) = \mu(T^{-i}(B) \cap T^{-i}(A)) = \mu(B \cap T^{-i}(A)) > 0.$$

Now, $B \cap A$ has positive measure in A; we shall also show that it is T_A-invariant. We have:

$$\begin{aligned} T_A^{-1}(B \cap A) &= \bigcup_{k=1}^{\infty} n_{A,T}^{-1}\{k\} \cap T^{-k}(B) \cap T^{-k}(A) = \bigcup_{k=1}^{\infty} n_{A,T}^{-1}\{k\} \cap T^{-k}(A) \cap B \\ &= B \cap \bigcup_{k=1}^{\infty} n_{A,T}^{-1}\{k\} \cap T^{-k}(A) = B \cap T_A^{-1}(A) = B \cap A. \end{aligned}$$

Hence by ergodicity of T_A, $\mu_A(B \cap A) = 1$, so $\mu(B \cap A) = \mu(A)$, so $B \cap A = A \setminus N$, where $\mu(N) = 0$. We then have:

$$\begin{aligned}
\mu(B) = \mu(B \cap \bar{A}) = \mu\left(B \cap \bigcup_{i=0}^{\infty} T^{-i}(A)\right) &= \mu\left(\bigcup_{i=0}^{\infty} T^{-i}(B \cap A)\right) \\
= \mu\left(\bigcup_{i=0}^{\infty} T^{-i}(A \setminus N)\right) &\geq \mu\left(\left(\bigcup_{i=0}^{\infty} T^{-i}(A)\right) \setminus \bigcup_{i=0}^{\infty} T^{-i}(N)\right) \\
= \mu\left(\bigcup_{i=0}^{\infty} T^{-i}(A)\right) &= 1.
\end{aligned}$$

Hence (T, μ) is ergodic, and we are done. □

In particular, we shall use statement (2) quite effectively, as it can happen that the induced map can give rise to a much simpler system, or at least one more well-known.

The second of our two important transformations, the skew product, will be the main type of map we consider in the work.

Definition 2.52. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system on a probability space, and let $(Y, \mathcal{B}, S_x)_{x \in X}$ be a family of measurable spaces equipped with measurable maps on Y, such that $S : X \times Y \to Y$ defined by $S(x, y) = S_x(y)$ is measurable with respect to $\mathcal{A} \otimes \mathcal{B}$. Define the map $R : X \times Y \to X \times Y$, $R(x, y) = (T(x), S_x(y)) = (T(x), S(x, y))$. Then we call R a *skew product* on $X \times Y$.

Observe that R is indeed measurable. To see this, let $T_1 = T \circ \pi_X$, so that T_1 and S are measurable from $X \times Y$ to X and Y, respectively. For any measurable rectangle $A \times B$, we have:

$$T_1^{-1}(A) = \{(x, y) \ : \ T_1(x, y) = T(x) \in A\} = T^{-1}(A) \times Y,$$

$$S^{-1}(B) = \{(x, y) \ : \ S_x(y) \in B\},$$

so that we obtain

$$\begin{aligned} R^{-1}(A \times B) &= \{(x,y) \ : \ T(x) \in A, S_x(y) \in B\} \\ &= \big\{(x,y) \ : \ x \in T^{-1}(A), (x,y) \in S^{-1}(B)\big\} \\ &= T_1^{-1}(A) \cap S^{-1}(B) \in \mathcal{A} \otimes \mathcal{B}. \end{aligned}$$

Moreover, for computational purposes, we have:

$$S^{-1}(B) = \bigcup_{x \in X} \{(x,y) \ : \ S_x(y) \in B\} = \bigcup_{x \in X} \{x\} \times S_x^{-1}(B),$$

$$R^{-1}(A \times B) = T^{-1}(A) \times Y \cap \bigcup_{x \in X} \{x\} \times S_x^{-1}(B) = \bigcup_{x \in T^{-1}(A)} \{x\} \times S_x^{-1}(B).$$

In general, R may not preserve any measure on the product $X \times Y$. Given some conditions on a family of measures ν_x and a measure M on $X \times Y$, it is possible to give abstract necessary and sufficient conditions for when the measure is preserved by R, using disintegration of the measure (see [1]). For our purposes, we shall only consider the case where $\nu_x = \nu$ for all $x \in X$, and where $(S_x)_* \nu = \nu$. That is, when each of the fibre maps preserves the fixed measure ν, where ν may be σ-finite. We have the following lemma.

Lemma 2.53. *In the above situation, R preserves the product measure $M = \mu \times \nu$.*

Proof. The collection of measurable rectangles in $X \times Y$ form a semi-algebra which generates $\mathcal{A} \otimes \mathcal{B}$, so it suffices to show that R preserves the measure of measurable rectangles in $\mathcal{A} \otimes \mathcal{B}$. Hence, let $A \in \mathcal{A}, B \in \mathcal{B}$, and compute:

$$\begin{aligned}
M(R^{-1}(A \times B)) &= \int_{X \times Y} \mathbb{1}_{R^{-1}(A \times B)}(x, y)\, dM(x, y) \\
&= \int_X \int_Y \mathbb{1}_{T_1^{-1}(A)}(x) \mathbb{1}_{S_x^{-1}(B)}(y)\, d\nu(y) d\mu(x) \\
&= \int_X \mathbb{1}_{T_1^{-1}(A)}(x) \left(\int_Y \mathbb{1}_{S_x^{-1}(B)}(y)\, d\nu(y) \right) d\mu(x) \\
&= \int_X \mathbb{1}_{T_1^{-1}(A)}(x) \nu(S_x^{-1}(B))\, d\mu(x) \\
&= \int_X \mathbb{1}_{T_1^{-1}(A)}(x) \nu(B)\, d\mu(x) \\
&= \nu(B) \int_X \mathbb{1}_{T_1^{-1}(A)}(x)\, d\mu(x) \\
&= \nu(B) \mu(T^{-1}(A)) = \nu(B)\mu(A) = M(A \times B).
\end{aligned}$$

Hence we are done. □

Skew products can have an interpretation as a *random dynamical system*. The base dynamics (X, μ, T) may be considered a random or probabilistic process, which decides a particular map to be applied to (Y, ν, S). In this way, the dynamics on Y can be considered random. A treatment of this very broad topic can be found in [1].

We now discuss a specific example of a skew product, which we shall study for the remainder of the thesis.

Definition 2.54. Let $(X, \mathcal{A}, T)$ be a dynamical system, let $(\mathbb{Z}, \mathcal{M})$ be the integers with the discrete σ-algebra, and let M be a topological monoid (that is, a topological space equipped with a continuous binary operation and an identity element e). A map $f : \mathbb{N} \times X \to M$ satisfying

1. $\forall n, m \in \mathbb{N},\ x \in X,\ f(n+m, x) = f(n, T^m(x)) f(m, x)$, and

2. $f(0, x) = e$,

is called a *cocycle* for or over T. If M is given the Borel σ-algebra and f is measurable, then f is a measurable cocycle. If T is invertible and f takes values in a topological group, then we insist that f must satisfy $f(-n,x) = f(n,T^{-n}(x))^{-1}$, which makes condition (1) above hold for all $n, m \in \mathbb{Z}$.

Note that because

$$f(n,x) = f(1,T^{n-1}(x))f(1,T^{n-2}(x))\dots f(1,T(x))f(1,x),$$

any cocycle is generated by the function $f(1,\cdot) : X \to M$. Moreover, if we have some measurable function $g : X \to M$, we may construct a cocycle G over T by specifying $G(1,x) = g(x)$ and requiring G to respect the cocycle properties. In general, we shall abuse notation and refer to either the cocycle f or the function $f(1,\cdot)$ on X as f.

One interpretation of cocycles is that they form some sort of 'running statistic' along orbits in a dynamical system. That is, if we wish to study the cumulative value of some observable quantity of a system (X,T) along an orbit, the cocycle is the natural object to use, because for a fixed n and x, the cocycle's value at (n,x) depends in a cumulative way on the points along the orbit of x for n time steps.

An important way to study cocycles is to create a skew product of a measure-preserving system and a topological group. Given a measure-preserving system $(X,\mathcal{B},\mu,T)$, a topological group $(G,\mathcal{A},m)$ with Borel σ-algebra $\mathcal{A}$ and Haar measure m, and a cocycle $f : \mathbb{Z} \times X \to G$, we may construct a skew product on $X \times G$ by defining $R : X \times G \to X \times G$, with $R(x,g) = (T(x), f(x)g)$. Note that $g \mapsto f(x)g$ is a rotation on a group, and group rotations preserve Haar measure. Hence, by Lemma 2.53, R preserves the product measure on $X \times G$, so we obtain a skew product measure-preserving system. This type of system will be utilized in the remainder of this work; here, we shall give a slightly involved example for illustration of an atypical method of showing ergodicity of a system.

Example 2.55. Recall the irrational rotation system $(\mathbb{T}, \mathcal{B}, \lambda, \sigma)$, where $\sigma(x) = x+\eta$, with η an irrational. Let $f : \mathbb{Z} \times \mathbb{T} \to \mathbb{T}$ be a cocycle generated by $f(1, x) = x$. Define the skew product R on $\mathbb{T}^2$ by $T(x, y) = (\sigma(x), f(1, x)+y) = (x+\eta, y+x)$, and denote the product measure by $\lambda \times \lambda$. We have the following theorem, considered (in more generality) by Furstenberg in [9].

Theorem 2.56. *T is an ergodic map, with respect to the product Lebesgue measure.*

Here, we shall give a geometric proof, using a particular characterization of ergodicity (condition (5) in Theorem 2.37), rather than the usual Fourier series computation, similar to the computation in Example 2.42.

Proof. Recall that by condition (5) of Theorem 2.37, T is ergodic if and only if for any A, B in a generating π-system for $\mathcal{B}$, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} \lambda \times \lambda(T^{-k}(A) \cap B) \underset{n\to\infty}{\longrightarrow} \lambda(A)\lambda(B).$$

Note that T is invertible; its inverse is given by $T^{-1}(x, y) = (x - \eta, y - x + \eta)$, as we can see:

$$T(T^{-1}(x, y)) = T(x - \eta, y - x + \eta) = (x - \eta + \eta, y - x + \eta + x - \eta) = (x, y)$$
$$T^{-1}(T(x, y)) = T^{-1}(x + \eta, y + x) = (x + \eta - \eta, y + x - (x + \eta) + \eta) = (x, y).$$

Both T and T^{-1} are measure-preserving. Then, recall that T is ergodic if and only if T^{-1} is ergodic; to see this, note that

$$(T^{-1})^{-1}(A) = T(A) = A \iff T^{-1}(T(A)) = A = T^{-1}(A),$$

and so our claim is easily verified. Thus the characterization of ergodicity of T is equivalent to saying that for all A, B in a generating semi-algebra for $\mathcal{B}$:

$$\frac{1}{n}\sum_{k=0}^{n-1} \lambda \times \lambda(T^k(A) \cap B) \underset{n\to\infty}{\longrightarrow} \lambda(A)\lambda(B).$$

We shall show that this holds for any half-open rectangles

$$A = [a, b) \times [c, d),\ B = [e, f) \times [g, h),$$

since we know that the half-open rectangles generate the Borel σ-algebra on $\mathbb{T}$ and form a π-system, hence by work in Example 2.7, the products form a generating π-system for the Borel σ-algebra on $\mathbb{T}^2$. Assume that

$$0 \le a < b \le 1,\ 0 \le c < d \le 1,\ 0 \le e < f \le 1,\ 0 \le g < h \le 1,$$

since any other rectangle can be formed as a finite disjoint union of these rectangles. Our job, then, is to estimate what $\lambda \times \lambda(T^k(A) \cap B)$ is.

First, we determine what $T^k(A)$ looks like. We compute T^k essentially by inspection, though we could do an induction argument:

$$T^k(x, y) = \left(x + k\eta, y + kx + \frac{(k-1)k}{2}\eta\right).$$

This seems mildly daunting, but it is less so if we consider it first as a map on $\mathbb{R}^2$. We'll look at $T(A)$ to start. A is the rectangle bounded by the vertical line segments

$$L_1 = \{(a, y)\ :\ y \in (c, d)\},\ L_2 = \{(b, y)\ :\ y \in (c, d)\},$$

and the horizontal line segments

$$L_3 = \{(x,c)\ :\ x \in (a,b)\},\ L_4 = \{(x,d)\ :\ x \in (a,b)\}.$$

Applying T to these lines gives:

$$T(a,y) = (a+\eta, y+a),\ \ T(b,y) = (b+\eta, y+b),$$
$$T(x,c) = (x+\eta, c+x),\ \ T(x,d) = (x+\eta, d+x).$$

In particular, T shifts the rectangle to the right by η, and then *shears* the rectangle into a parallelogram, by raising the right edge of the rectangle $b-a$ more than the left edge. As the second line of computation shows, the horizontal line segments are mapped to line segments of slope 1, with appropriate endpoints. Similarly, we obtain that:

$$\begin{aligned} T^k(a,y) &= \left(a + k\eta, y + ka + \frac{(k-1)k}{2}\eta\right), \\ T^k(b,y) &= \left(b + k\eta, y + kb + \frac{(k-1)k}{2}\eta\right), \\ T^k(x,c) &= \left(x + k\eta, c + kx + \frac{(k-1)k}{2}\eta\right), \\ T^k(x,d) &= \left(x + k\eta, d + kx + \frac{(k-1)k}{2}\eta\right). \end{aligned}$$

So after k iterations, the rectangle is mapped to a heavily vertically-sheared parallelogram, where the base and top of the parallelogram have slope k. This means that while it takes up the same width in the x-axis, it stretches over more of the y-axis. The exact value of that stretch is the height of the vertical sides plus the height attributable to the shear:

$$H_k = (d-c) + y + kb + \frac{(k-1)k}{2}\eta - \left(y + ka + \frac{(k-1)k}{2}\eta\right) = (d-c) + k(b-a).$$

We also wish to know how wide the parallelogram is (that is, the width of the horizontal segment of the resulting set), parallel to the x-axis, while it is sloping upwards. We can figure this out by some trigonometry; observe that the width W_k lies on a right-angled triangle opposite the angle θ, where the hypotenuse has slope k and the other edge of the triangle is of length $d-c$. That yields

$$\frac{W_k}{d-c} = \tan(\theta) = \tan\left(\arctan\left(\frac{1}{k}\right)\right) = \frac{1}{k},$$

which tells us that $W_k = \frac{d-c}{k}$.

This describes the parallelogram, ignoring the wrap-around that actually occurs while on the torus. Taking this into account, we see that the (not-necessarily-integer-valued) number of times $T^k(A)$ wraps completely around the vertical direction of the torus is given by H_k. Note that we ignore the vertical shifting, because it is negligible for our purposes, considering the extreme wraparound.

We would like to estimate the intersection of $T^k(A)$ and B. The former is a collection of stripes, of width W_k. Any full intersection of a stripe with B is of height $h-g$. We estimate the number of stripes by considering the intersection of $\sigma^k(a,b)$ and (e,f), where $\sigma(x) = x + \eta$. The distance between left end-points of stripes is exactly $\frac{1}{k}$, because the parallelograms have slope k, and the torus has vertical height 1. Then the number of stripes in the intersection is approximated by:

$$\frac{\lambda(\sigma^k(a,b) \cap (e,f))}{\frac{1}{k}},$$

with error $\mathcal{O}(1)$, because we may be missing part of a stripe depending on vertical location of $T^k(a,b)$. Then we may approximate the area of intersection of $T^k(A)$ and B by:

$$(h-g)\frac{(d-c)}{k}\frac{\lambda(\sigma^k(a,b)\cap(e,f))}{\frac{1}{k}} = (h-g)(d-c)\lambda(\sigma^k(a,b)\cap(e,f)),$$

with error $\mathcal{O}(\frac{1}{k})$, because the leftover area from the extra partial stripe would be at most $(h-g)\dfrac{(d-c)}{k}$. The ergodic average of $\lambda(\sigma^k(a,b)\cap(e,f))$ converges to $(b-a)(f-e)$, since σ^{-1} is ergodic (by condition (5) of Theorem 2.37); this allows to conclude:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\lambda(T^k(A)\cap B) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left((h-g)(d-c)\lambda(\sigma^k(a,b)\cap(e,f))+\mathcal{O}\left(\frac{1}{k}\right)\right)\\
&= (h-g)(d-c)\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\lambda(\sigma^k(a,b)\cap(e,f))+\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mathcal{O}\left(\frac{1}{k}\right)\\
&= (h-g)(d-c)(b-a)(f-e)+\lim_{n\to\infty}\frac{1}{n}\mathcal{O}(\log(n))\\
&= (b-a)(d-c)(f-e)(h-g).
\end{aligned}$$

Therefore T is ergodic. □

It should be noted that proving ergodicity of a concrete, hands-on system is in general a non-trivial task. If one cannot utilize equivalent characterizations to reduce the problem to some sort of algebraic or analytic trickery, such as an easy Fourier series computation, or if one cannot prove something stronger which implies ergodicity, then it ends up being very difficult. Indeed, for certain systems, it is surprisingly troublesome to prove ergodicity, as we shall see later. We do have the following proposition, which is useful for product systems:

Lemma 2.57. *Let* $(X, \mathcal{A}, \mu, S)$ *be a measure-preserving system where* S *is strongly mixing and* $\mu(X) = 1$*, and let* $(Y, \mathcal{B}, \nu, T)$ *be a measure-preserving system where* T *is ergodic and* $\nu(Y) = 1$*. Then the product map* $S \times T$ *on* $X \times Y$ *is ergodic with respect to* $\mu \times \nu$*.*

Proof. We will show that condition (5) in Theorem 2.37 is satisfied by $S \times T$. The set of measurable rectangles $A \times B$ is a semi-algebra generating $\mathcal{A} \otimes \mathcal{B}$, so it suffices to consider the condition applied to such sets. So let $A, C \in \mathcal{A}$ and $B, D \in \mathcal{B}$. For $k \in \mathbb{N}$, we have

$$\begin{aligned}(S \times T)^{-k}(A \times B) &= \big\{(x, y) \ : \ S^k(x) \in A, T^k(y) \in B\big\} \\ &= \big\{(x, y) \ : \ x \in S^{-k}(A), y \in T^{-k}(B)\big\} \\ &= S^{-k}(A) \times T^{-k}(B),\end{aligned}$$

so that we obtain:

$$\begin{aligned}\mu \times \nu((S \times T)^{-k}(A \times B) \cap (C \times D)) &= \mu \times \nu((S^{-k}(A) \times T^{-k}(B)) \cap (C \times D)) \\ &= \mu \times \nu((S^{-k}(A) \cap C) \times (T^{-k}(B) \cap D)) \\ &= \mu(S^{-k}(A) \cap C)\nu(T^{-k}(B) \cap D).\end{aligned}$$

Let $\epsilon > 0$. Because S is strongly mixing with respect to μ, we may find $K \in \mathbb{N}$ such that for all $k \geq K$,

$$\big|\mu(S^{-k}(A) \cap C) - \mu(A)\mu(C)\big| < \frac{\epsilon}{3}.$$

Because T is ergodic with respect to ν, we have that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \nu(T^{-k}(B) \cap D) = \nu(B)\nu(D).$$

We may pick $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, the absolute value of the difference of the two above quantities is less than $\frac{\epsilon}{3}$. Moreover, if we neglect a finite number of initial terms, the result is still true; we have:

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \sum_{k=K}^{n-1} \nu(T^{-k}(B) \cap D) &= \lim_{n\to\infty} \left(\frac{1}{n} \sum_{k=0}^{n-1} \nu(T^{-k}(B) \cap D) - \frac{1}{n} \sum_{k=0}^{K-1} \nu(T^{-k}(B) \cap D) \right) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=K}^{n-1} \nu(T^{-k}(B) \cap D) - \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{K-1} \nu(T^{-k}(B) \cap D) \\
&= \nu(B)\nu(D) - 0 = \nu(B)\nu(D).
\end{aligned}$$

We may pick N_2 such that the finite number of terms over n is less than $\frac{\epsilon}{3}$ for all

$n \geq N_2$. Hence, we have, for $n \geq \max\{N_1, N_2\}$:

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{k=0}^{n-1}\mu \times \nu((S \times T)^{-k}(A \times B) \cap (C \times D)) - \mu \times \nu(A \times B)\mu \times \nu(C \times D)\right| \\
&= \left|\frac{1}{n}\sum_{k=0}^{n-1}\mu(S^{-k}(A) \cap C)\nu(T^{-k}(B) \cap D) - \mu(A)\nu(B)\mu(C)\nu(D)\right| \\
&\leq \left|\frac{1}{n}\sum_{k=0}^{n-1}\mu(S^{-k}(A) \cap C)\nu(T^{-k}(B) \cap D) - \frac{1}{n}\sum_{k=K}^{n-1}\mu(S^{-k}(A) \cap C)\nu(T^{-k}(B) \cap D)\right| \\
&\qquad + \left|\frac{1}{n}\sum_{k=K}^{n-1}\mu(S^{-k}(A) \cap C)\nu(T^{-k}(B) \cap D) - \frac{1}{n}\sum_{k=K}^{n-1}\mu(A)\mu(C)\nu(T^{-k}(B) \cap D)\right| \\
&\qquad + \left|\frac{1}{n}\sum_{k=K}^{n-1}\mu(A)\mu(C)\nu(T^{-k}(B) \cap D) - \mu(A)\nu(B)\mu(C)\nu(D)\right| \\
&= \left|\frac{1}{n}\sum_{k=0}^{K-1}\mu(S^{-k}(A) \cap C)\nu(T^{-k}(B) \cap D)\right| \\
&\qquad + \left|\frac{1}{n}\sum_{k=K}^{n-1}\left(\mu(S^{-k}(A) \cap C) - \mu(A)\mu(C)\right)\nu(T^{-k}(B) \cap D)\right| \\
&\qquad + \left|\frac{\mu(A)\mu(C)}{n}\left(\sum_{k=K}^{n-1}\nu(T^{-k}(B) \cap D) - \nu(B)\nu(B)\right)\right| \\
&< \frac{\epsilon}{3} + \frac{1}{n}\sum_{k=K}^{n-1}\frac{\epsilon}{3}\nu(T^{-k}(B) \cap D) + \mu(A)\mu(C)\frac{\epsilon}{3} \\
&< \frac{\epsilon}{3} \cdot 3 = \epsilon.
\end{aligned}
$$

Therefore condition (5) is satisfied, and $S \times T$ is ergodic. □

Another method for proving ergodicity of certain types of skew products is to induce the skew product on a set of positive measure, and then use Proposition 2.51; this method will be utilized later in the thesis. To this end, we have the following useful lemma.

Lemma 2.58. *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic measure-preserving system on a probability space, let $(Y, \mathcal{B}, \nu, S_x)_{x \in X}$ be a family of measure-preserving systems on probability* *spaces, and let* $R : X \times Y \to X \times Y$ *be the skew product of* $(S_x)_x$ *over* T*, so that* $R(x, y) = (T(x), S_x(y))$*. Let* $C = A \times Y$*, for* $A \in \mathcal{A}$ *with* $\mu(A) > 0$*. Then we have:*

$$\mu \times \nu \left(\bigcup_{i=0}^{\infty} R^{-i}(C) \right) = 1.$$

Proof. First, we compute the inverse image of C under R:

$$\begin{aligned} R^{-1}(C) &= \{(x, y) \ : \ T(x) \in A, \ S_x(y) \in Y\} \\ &= \big\{(x, y) \ : \ x \in T^{-1}(A), \ y \in Y\big\} = T^{-1}(A) \times Y. \end{aligned}$$

By iteration, we obtain $R^{-i}(C) = T^{-i}(A) \times Y$. We then have:

$$\bigcup_{i=0}^{\infty} R^{-i}(C) = \bigcup_{i=0}^{\infty} T^{-i}(A) \times Y = \left(\bigcup_{i=0}^{\infty} T^{-i}(A) \right) \times Y.$$

Then by Lemma 2.46, since the base measure-preserving system is ergodic on a probability space, we may conclude:

$$\mu \times \nu \left(\bigcup_{i=0}^{\infty} R^{-i}(C) \right) = \mu \left(\bigcup_{i=0}^{\infty} T^{-i}(A) \right) \nu(Y) = 1 \cdot 1 = 1,$$

and so we are done. □

There is a very important class of cocycles, which will become the foundation of the forthcoming chapter.

Example 2.59. Consider a general dynamical system $(X, \mathcal{A}, T)$, and let $d \in \mathbb{N}_+$. Let $A(1, \cdot) : X \to M_d(\mathbb{R})$ be measurable, and let $A(1, \cdot)$ generate a *matrix* cocycle, where the monoid operation is matrix multiplication on the left. This is the main object of study for the remainder of the thesis, where we shall consider a slightly more specific instance, where A takes values in $GL_d(\mathbb{R})$.

A ‘real life’ example of one of these structures is the composition of derivative matrices of dynamics on a smooth manifold. That is, let $(M_{\mathcal{B}}, \mu, \sigma)$ be a measure-preserving system on a probability space where M has a real smooth manifold structure of dimension d, and σ be smooth. The derivative of σ at a point p is a linear function $D_\sigma(p) : T_p(M) \to T_{\sigma(p)}$, which maps vectors in the tangent space at p to vectors in the tangent space at $\sigma(p)$. The Chain Rule applies to the iterates of σ:

$$D_{\sigma^k}(p) = D_\sigma(\sigma^{k-1}(p)) \circ \cdots \circ D_\sigma(\sigma(p)) \circ D_\sigma(p).$$

If we choose to represent $D_\sigma(p)$ as acting on the standard basis in $\mathbb{R}^d$, it becomes a matrix, and the map $D : \mathbb{N} \times M \to M_d(\mathbb{R})$ given by $D(k,p) = D_{\sigma^k}(p)$ is a matrix cocycle, where the cocycle property is verified by inspecting the Chain Rule computation above.

2.3 The MET

The foundation of this work is the Multiplicative Ergodic Theorem (MET), for invertible matrix cocycles over an invertible base measure-preserving system. This must be specified, because there are many variations of this theorem, either for different hypotheses on the matrices and the underlying dynamics [8], or for more general objects like operators on arbitrary Banach spaces [10]. The fact that there are so many variations is indicative of the similar nature of these objects, and proving different versions is of great importance, for usage in many physical problems [7].

We shall concern ourselves with perhaps the most specific version of the MET, proven in 1968 by Oseledets [16], and examine how far we can push the theorem in a particular direction. In doing so, we shall see that the MET is in some sense optimal.

Theorem 2.60 (Multiplicative Ergodic Theorem for Matrix Cocycles, Invertible Version). *Let* $(X, \mathcal{B}, \mu, T)$ *be an invertible and ergodic measure-preserving system on a probability space, and let* $A(1,\cdot) : X \to GL_d(\mathbb{R})$ *be the generator of a measurable matrix cocycle* A*. Suppose that* A *satisfies*

$$\int_X \log(\|A(1,x)\|)\, d\mu < \infty, \quad \int_X \log(\|A(1,x)^{-1}\|)\, d\mu < \infty.$$

Then there exists $k \in \mathbb{N}_+$*,* $\lambda_1 > \lambda_2 > \cdots > \lambda_k > -\infty$ *called* Lyapunov exponents*, positive integers* $m_1, m_2, \ldots, m_k$ *such that* $m_1 + \cdots + m_k = d$*, a measurable decomposition* $V_1(x) \oplus \cdots \oplus V_k(x)$ *with dimensions* $(m_1, \cdots, m_k)$*, and a* T*-invariant set of full measure,* $\tilde{X} \subset X$ *with the following properties:*

1. ***Equivariance:*** *For all* $x \in \tilde{X}$*, and* $i = 1, \cdots, k$*, we have:*

$$A(1,x)V_i(x) = V_i(T(x)).$$

2. ***Bilateral Growth:*** *Given* $x \in \tilde{X}$ *and* $i \in \{1, \ldots, k\}$*, for all* $v \in V_i(x) \setminus \{0\}$*, we have:*

$$\frac{1}{n}\log(\|A(n,x)v\|) \underset{n\to\infty}{\longrightarrow} \lambda_i, \ \frac{1}{n}\log(\|A(-n,x)v\|) \underset{n\to\infty}{\longrightarrow} -\lambda_i.$$

The growth condition in the conclusion informally says that for for $v \in V_i$, $\|A(n,x)v\| \sim e^{n\lambda_i}$ and $\|A(-n,x)v\| \sim e^{-n\lambda_i}$. That is, the λ_i describes the exponential growth and decay rates of vectors in the i^{th} subspace. The equivariance condition is needed for this to make sense, because vectors in $V_i(x)$ must move to $V_i(T(x))$ and therefore their images will have the same growth rate.

Example 2.61. Let $X = \{1\}$ be a one-element set, $\mathcal{B} = \{\emptyset, X\}$ be the only σ-algebra on X, μ a set function on X given by $\mu(X) = 1$ and $\mu(\emptyset) = 0$, and $T(1) = 1$ be the identity map on X. This is a trivially ergodic and invertible measure-preserving system. Let $A \in GL_d(\mathbb{R})$ be diagonalizable, and conflate the matrix A with the map $A : X \to GL_d(\mathbb{R})$. A then generates a cocycle over T, with $A(n,1) = A^n$. Applying the MET in this case gives us a subspace decomposition of $\mathbb{R}^d$ which has the properties above. If μ is an eigenvalue of A with corresponding eigenvector w, then we have:

$$\begin{aligned}\frac{1}{n}\log\|A(n,1)w\| = \frac{1}{n}\log\|A^n w\| &= \frac{1}{n}\log|(|\,\mu^n\,\|w\|) \\ &= \log|\mu| + \frac{1}{n}\log\|w\| \underset{n\to\infty}{\longrightarrow} \log|\mu|\,.\end{aligned}$$

So the Lyapunov exponent corresponding to the V_i containing w is the logarithm of the absolute value of the eigenvalue. Since A is diagonalizable, there is a basis of eigenvectors $\{w_i\}_{i=1}^n$ with k eigenvalues $\{\mu_1, \ldots, \mu_k\}$. This gives us k distinct eigenspaces $V_1, \ldots, V_k$ corresponding to Lyapunov exponents $\lambda_i = \log|\mu_i|$. In this way, we recover the diagonalization of A.

If A is not diagonalizable but has real eigenvalues, it still has a collection of generalized eigenspaces corresponding to each eigenvalue, and the Jordan normal form specifies basis vectors such that the asymptotic expansion rate is the same as in the diagonalized case. The MET recovers the subspaces in the Jordan form for A, via Lyapunov exponents; of course, it doesn't necessarily reveal an obvious way to arrange the vectors such that the matrix takes an upper-triangular form, like in the Jordan form. It is only block diagonal.

Example 2.62. Let $(X, \mathcal{B}, \mu, T)$ be an invertible and ergodic measure-preserving system, with $\mu(X) = 1$. If $d = 1$, then real invertible matrices are simply non-zero real numbers, which act on $\mathbb{R}$ by multiplication. Let $A : \mathbb{Z} \times X \to GL_1(\mathbb{R}) = \mathbb{R} \setminus \{0\}$ be a measurable cocycle satisfying the log-integrability conditions in both directions, and observe that for non-zero $v \in \mathbb{R}$, we have:

$$\begin{aligned}\frac{1}{n}\log\|A(n,x)v\| &= \frac{1}{n}\log\left(\left|A(1,T^{n-1}(x))\right|\ldots|A(1,x)|\,|v|\right)\\ &= \frac{1}{n}\sum_{i=0}^{n-1}\log\left|A(a,T^{i}(x))\right| + \frac{1}{n}\,|v| \underset{n\to\infty}{\longrightarrow} \int_X \log|A(1,x)|\;d\mu(x),\end{aligned}$$

by Birkhoff's theorem (Theorem 2.34), since the log-integrability conditions mean exactly that $\log|A(1,x)|$ is an L^1 function. Hence in this case, the one Lyapunov exponent for the system is given by the time average of $\log|A(1,\cdot)|$ (which equals the space average).

For an integrable function $f : X \to \mathbb{R}$, we note that $\log\left|e^{f(x)}\right| = \log(e^{f(x)}) = f(x)$ is integrable, so letting $A(1,x) = e^{f(x)}$ we obtain both a 'matrix' cocycle, generated by $A(1,x)$, and an Abelian group cocycle, generated by $f(x)$. Letting $v = 1$, the MET then gives the almost-everywhere convergence of the time averages of f, which is exactly the statement of Birkhoff's theorem. Hence the MET generalizes Birkhoff's theorem. This makes some intuitive sense, as the theorems have a very similar flavour, when unpacked.

In the case of non-invertible T and matrices $A(1,x)$, the MET yields a weaker statement. Instead of a subspace decomposition, there exists a measurable filtration of subspaces on which the Lyapunov exponents increase; see Raghunathan, [19]. If at least the map T is invertible, then there is an in-between result, which restores the subspace decomposition in a slightly weaker form [8].

There are also Multiplicative Ergodic Theorems proven for more exotic spaces. In 1982, Ruelle proved an MET for bounded operators on a Hilbert space, which was the first extension of these ideas into infinite dimensional space [24]. Shortly thereafter, in 1983 Mañé proved a version of the MET for compact operators on Banach spaces [15], and there have been other METs proven in cases of bounded linear operators on

Banach spaces with certain conditions (for instance, see Lian and Lu [14], González-Tokman and Quas [10]).

Along with the many variations of the MET, there are many proofs thereof, and different proofs can give us intuition and insight towards these objects. The proof by Raghunathan in [19], in the case where both the base dynamics and the matrix cocycle are not assumed to be invertible, works hard to construct the subspaces while making sure that they are measurable. The proof by Walters in [32] utilizes a result concerning maps which take the form of skew products over the the product of the base dynamics space and a compact metric space, which can be applied to the product of the base space and a Grassmannian. The proof given by Barreira and Pesin in [3] follows the lines that Oseledets used in the original paper in the area, [16]. The main idea is that it is much easier to prove the result for triangular cocycles; given a matrix cocycle on a particular space X, they extend the space by a copy of $SO_d(\mathbb{R})$ and construct a new matrix cocycle over the new space, and show that there exists a measurable decomposition for the cocycle on X if and only if there is one for the cocycle over the extended space. This new cocycle is triangular, so the proof becomes simpler.

The usefulness of this triangular cocycle over an extended space leads us consider the idea of upper triangularizing the original cocycle, similar to how upper triangularizing matrices can be very informative and computationally beneficial. We saw in Example 2.61 that the MET yields subspaces, but we don't necessarily obtain finer structure. We shall make precise the mathematics and examine this concept in the next chapter.

Chapter 3

Equivariant triangularization

3.1 Setup

Consider again the invertible version of the MET. We may consider this equivariant decomposition of $\mathbb{R}^d$ into subspaces as a block diagonalization of the matrix cocycle A, in the following formulation:

Theorem 3.1 (Equivalent Formulation of MET for Invertible Matrix Cocycles)**.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible and ergodic measure-preserving system on a probability space, and $A : X \times \mathbb{Z} \to GL_d(\mathbb{R})$ be a measurable cocycle. Suppose that A satisfies*

$$\int_X \log^+(\|A(1,x)\|)\; d\mu < \infty, \quad \int_X \log^+(\left\|A(1,x)^{-1}\right\|)\; d\mu < \infty.$$

Then there exists $k \in \mathbb{N}_+$, $\lambda_1 > \lambda_2 > \cdots > \lambda_k \geq -\infty$, positive integers $m_1, m_2, \ldots, m_k$ such that $m_1 + \cdots + m_k = d$, a measurable function $C : X \to GL_d(\mathbb{R})$, and a T-invariant set of full measure, $\tilde{X} \subset X$ with the following properties:

1. ***Equivariance:*** *For all $x \in \tilde{X}$, we have that*

$$C(T(x))^{-1}A(1,x)C(x)$$

is block diagonal with block sizes $(m_1, \cdots, m_k)$*;*

2. ***Bilateral Growth:*** *Given* $x \in \tilde{X}$ *and* $i \in \{1, \ldots, k\}$*, for all non-zero* v *in the columnspace of the* i^{th} *block, we have:*

$$\frac{1}{n}\log(\|A(n,x)v\|) \underset{n\to\infty}{\longrightarrow} \lambda_i, \ \frac{1}{n}\log(\|A(-n,x)v\|) \underset{n\to\infty}{\longrightarrow} -\lambda_i.$$

Given the subspaces $V_i(x),\ i = 1, \ldots, k$ in the classical MET, one may measurably choose basis vectors for each of them, which Walters showed in [32]; call them v_i^j, for $j = 1, \ldots, m_i$. Then construct the matrix $C(x)$ whose columns are the v_i^j, in order; we see that the matrix $C(T(x))A(1,x)C(x)^{-1}$ is block diagonal. Conversely, given the matrices $C(x)$, we may measurably construct the equivariant subspaces by taking $V_i(x) = \text{span}_{\mathbb{R}}\{C(x)e_{M_i+1}, \ldots, C(x)e_{M_i+m_i}\}$, where $i = 1, \ldots, k$ and $M_i = m_1 + \cdots + m_{i-1}$, with $M_1 = 0$.

We see that this restatement of the MET is a result about block diagonalizing matrices in a well-behaved way. However, aside from the overall structure for each block, we learn nothing about the internal structure of the blocks. Is it possible that there may always be finer structure to the blocks, which is unspecified by the theorem? Namely, may we always upper triangularize these blocks? If not over $\mathbb{R}$, over $\mathbb{C}$?

Some authors have investigated this topic. Thieullen, in 1997, investigated the case of 2-by-2 real invertible matrices, and showed that there were four distinct cases for real normal forms [28]. His approach involved complex conformal mapping theory, which foreshadows work later in this thesis, and he proves the results for determinant one cocycles before extending the result to all invertible matrix cocycles.

In 1999, Arnold, Nguyen, and Oseledets proved a so-called Jordan normal form for matrix cocycles, without assuming a log-integrability condition [2]; they obtain an upper-triangular form for the whole matrix, with blocks on the diagonal. In particular, this form does not necessarily have triangular blocks; it has blocks along the diagonal, with zeroes below, but nothing specific above. If the hypotheses of the MET are also satisfied, their results allow for a block triangular form, which is not necessarily any more refined than the blocks already obtained.

Both of these results deal with the situation of matrices over the real numbers, however. In Example 2.61 we analyzed a cocycle which was simply powers of a single matrix A. It is well-known that if we allow for conjugation by complex matrices (that is, letting our change of basis matrix be complex-valued), we may always find a basis (corresponding to a matrix U) in which A is upper triangular. First, we may apply the MET in the real situation to obtain a block-diagonal form for A. Then for each block, we may conjugate by a complex matrix U_i to obtain a complex upper-triangular block, yielding a a block-upper-triangular form for the matrix (and hence, the cocycle).

This example leads us to the following definition and question:

Definition 3.2. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and let A be a cocycle over T of real n-by-n matrices, which together satisfy the hypotheses of the MET; then A may be block-diagonalized over $\mathbb{R}$, with block sizes $m_1, \ldots, m_k$. We say that A is *block upper-triangularizable over* $\mathbb{F}$ if there is a measurable cocycle $C : X \to GL_n(\mathbb{F})$ such that

$$C(T(x))^{-1}A(1, x)C(x)$$

is block upper-triangular, with entries in $\mathbb{F}$ and block sizes $m_1, \ldots, m_k$.

Question. *Given a cocycle A satisfying the MET (Theorem 3.1), can we necessarily always block upper-triangularize A, possibly over $\mathbb{C}$?*

Answer. *We answer this question in the negative. We shall state and prove a sufficient condition to ensure that a matrix cocycle cannot be block upper-triangularized over $\mathbb{R}$ or over $\mathbb{C}$. We shall then present three different explicit matrix cocycles, showing how the condition may (or may not) be satisfied.*

3.2 General framework

Let us introduce some notation, after which we will state the aforementioned condition. Denote the complex unit circle $\left\{z \in \bar{\mathbb{C}} \,:\, |z| = 1\right\}$ by $\mathbb{S}$, and denote the real orthogonal group of 2-by-2 matrices by $O_2(\mathbb{R})$. Proposition 2.18 says that there is a homeomorphism ψ from $\mathrm{Gr}_1(\mathbb{C}^2)$ to $\bar{\mathbb{C}}$. Define maps $\iota : \bar{\mathbb{C}} \setminus \mathbb{S} \to \mathbb{Z}_2$ and $\tau : \bar{\mathbb{C}} \setminus \{0, \infty\} \to \mathbb{T}$ by:

$$\iota(z) = \begin{cases} 0 & |z| < 1 \\ 1 & |z| > 1 \text{ or } z = \infty \end{cases}, \qquad \tau(z) = \frac{1}{2\pi}\arg(z).$$

It is easy to see that these are measurable.

Let $(X, \mathcal{B}, \mu, T)$ be an invertible and ergodic measure-preserving system on a probability space, and let $A : \mathbb{Z} \times X \to O_2(\mathbb{R})$ be a measurable matrix cocycle over T. For each $x \in X$, $A(1, x)$ is an orthogonal, hence invertible, matrix, and it defines a map on $\mathrm{Gr}_1(\mathbb{C}^2)$, as in Section 3.1. Using the map ψ, we obtain a measurable map $M(x) = \psi \circ A(1, x) \circ \psi^{-1}$ which acts on $\bar{\mathbb{C}}$, and we get a skew product $N : X \times \bar{\mathbb{C}} \to X \times \bar{\mathbb{C}}$ with $N(x, z) = (T(x), M(x)z)$.

Since $A(1, x) \in O_2(\mathbb{R})$, each map either rotates by some angle α_x, or flips in the line with angle β_x; let $X_r \subset X$ be where $A(1, x)$ is a rotation, and let $X_f \subset X$ be where $A(1,x)$ is a flip. Define the maps $f_x : \mathbb{T} \to \mathbb{T}$ and $g_x : \mathbb{Z}_2 \to \mathbb{Z}_2$ by:

$$f_x(y) = \begin{cases} y + \dfrac{\alpha_x}{\pi} & x \in X_r, \\ \dfrac{2\beta_x}{\pi} - y & x \in X_f, \end{cases} \qquad g_x(a) = \begin{cases} a & x \in X_r, \\ a+1 & x \in X_f. \end{cases}$$

From these maps, define skew products $R : X \times \mathbb{Z}_2 \to X \times \mathbb{Z}_2$ given by $R(x,a) = (T(x), g_x(a))$, and $S : X \times \mathbb{T} \to X \times \mathbb{T}$ given by $S(x,y) = (T(x), f_x(y))$. The maps R and S will be shown to be factors of N, restricted appropriately. The theorem can now be stated.

Theorem 3.3. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible and ergodic measure-preserving system on a probability space, and let $A : Z \times X \to O_2(\mathbb{R})$ be a measurable cocycle of orthogonal 2-by-2 real matrices over T. Let $R : X \times \mathbb{Z}_2 \to X \times \mathbb{Z}_2$ and $S : X \times \mathbb{T} \to X \times \mathbb{T}$ be as described above. If S is ergodic, then the cocycle A cannot be block upper-triangularized over $\mathbb{R}$. If both R and S are ergodic on their respective spaces, then the cocycle A cannot be block upper-triangularized over $\mathbb{C}$.*

The proof will proceed in roughly four steps:

1. Make rigorous and prove the statements in the above discussion and setup.

2. Assume, for contradiction, the existence of an upper-triangularization for A, and obtain an equivariant family of 1-D complex subspaces.

3. Translate the equivariant subspaces into an invariant graph on $\mathbb{C}$ under a skew product involving M, and then translate into invariant graphs for each of the factor maps f and g above.

4. Utilize ergodicity of R and S to show that the existence of invariant graphs is a contradiction.

Step 1: First, we must understand the action of $A(x)$ on $\mathrm{Gr}_1(\mathbb{C}^2)$, to obtain the action of $A(x)$ on $\bar{\mathbb{C}}$. We recall that $O_2(\mathbb{R})$ is composed of rigid symmetries of the unit circle in $\mathbb{R}^2$, which are rotations and reflections. So for each $x \in X$, $A(1,x)$ is either a rotation by some angle $\alpha_x \in [0, 2\pi)$ (so $A(1,x) = \mathrm{rot}_{\alpha_x}$) or is a reflection over the line with angle $\beta_x \in [0, \pi)$ (so $A(1,x) = \mathrm{refl}_{\beta_x}$). Let X_r and X_f be the sets on which $A(1,x)$ is a rotation and a flip, respectively. We prove the following lemma:

Lemma 3.4. *The map $M(x) = \psi \circ A(1,x) \circ \psi^{-1}$ is measurable and invertible, and has the following form. For $x \in X_r$ with $A(1,x) = rot_{\alpha_x}$, we have*

$$M(x)z = \begin{cases} e^{2i\alpha_x} z & z \neq \infty, \\ \infty & z = \infty. \end{cases}$$

For $x \in X_f$ with $A(1,x) = A(1,x) = refl_{\beta_x}$, we have

$$M(x)z = \begin{cases} \dfrac{e^{4i\beta_x}}{z} & z \notin \{0, \infty\}, \\ \infty & z = 0, \\ 0 & z = \infty. \end{cases}$$

Proof. The measurability of $M(x)$ is clear, as all three of the maps in the composition are continuous, hence measurable. In order to describe the action of $A(1,x)$ on $\mathrm{Gr}_1(\mathbb{C}^2)$, we use Proposition 2.18. As there, let

$$v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix};$$

we may write every subspace in $\mathrm{Gr}_1(\mathbb{C}^2)$ as either $\mathrm{span}_{\mathbb{C}}\{v_1 + zv_2\}$ for some $z \in \mathbb{C}$, or $\mathrm{span}_{\mathbb{C}}\{v_2\}$. We can compute the action of $A(1,x)$ on $\mathrm{Gr}_1(\mathbb{C}^2)$ by applying them to

$v_1 + zv_2$, from above, and then taking the span.

If $x \in X_r$, then $A(1,x) = \mathrm{rot}_{\alpha_x}$. We have:

$$\begin{aligned} A(1,x)(v_1 + zv_2) &= \begin{bmatrix} \cos(\alpha_x) & -\sin(\alpha_x) \\ \sin(\alpha_x) & \cos(\alpha_x) \end{bmatrix} \left(\begin{bmatrix} 1 \\ i \end{bmatrix} + z \begin{bmatrix} 1 \\ -i \end{bmatrix} \right) \\ &= \begin{bmatrix} \cos(\alpha_x) - i\sin(\alpha_x) \\ \sin(\alpha_x) + i\cos(\alpha_x) \end{bmatrix} + z \begin{bmatrix} \cos(\alpha_x) + i\sin(\alpha_x) \\ \sin(\alpha_x) - i\cos(\alpha_x) \end{bmatrix} \\ &= \begin{bmatrix} e^{-i\alpha_x} \\ e^{-i\alpha_x} i \end{bmatrix} + z \begin{bmatrix} e^{i\alpha_x} \\ -e^{i\alpha_x} i \end{bmatrix} \\ &= e^{-i\alpha_x}(v_1 + e^{2i\alpha_x} zv_2). \end{aligned}$$

Therefore, we see that

$$A(1,x)\,\mathrm{span}_{\mathbb{C}}\{v_1 + zv_2\} = \mathrm{span}_{\mathbb{C}}\{v_1 + e^{2i\alpha_x} zv_2\},$$

when $A(1,x)$ is a rotation. We also have:

$$A(1,x)v_2 = \begin{bmatrix} \cos(\alpha_x) & -\sin(\alpha_x) \\ \sin(\alpha_x) & \cos(\alpha_x) \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = e^{i\alpha_x} \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

which yields

$$A(1,x)\,\mathrm{span}_{\mathbb{C}}\{v_2\} = \mathrm{span}_{\mathbb{C}}\{v_2\}.$$

Similarly, if $x \in X_f$, then $A(1,x) = \text{refl}_{\beta_x}$. We have, for $z \in \mathbb{C} \setminus \{0\}$:

$$\begin{aligned}
A(1,x)(v_1 + zv_2) &= \begin{bmatrix} \cos(2\beta_x) & \sin(2\beta_x) \\ \sin(2\beta_x) & -\cos(2\beta_x) \end{bmatrix} \left(\begin{bmatrix} 1 \\ i \end{bmatrix} + z \begin{bmatrix} 1 \\ -i \end{bmatrix} \right) \\
&= \begin{bmatrix} \cos(2\beta_x) + i\sin(2\beta_x) \\ \sin(2\beta_x) - i\cos(2\beta_x) \end{bmatrix} + z \begin{bmatrix} \cos(2\beta_x) - i\sin(2\beta_x) \\ \sin(2\beta_x) + i\cos(2\beta_x) \end{bmatrix} \\
&= \begin{bmatrix} e^{2i\beta_x} \\ e^{2i\beta_x}(-i) \end{bmatrix} + z \begin{bmatrix} e^{-2i\beta_x} \\ e^{-2i\beta_x}(i) \end{bmatrix} \\
&= e^{2i\beta_x} v_2 + e^{-2i\beta_x} z v_1 \\
&= e^{-2i\beta_x} z \left(v_1 + \frac{e^{4i\beta_x}}{z} v_2 \right).
\end{aligned}$$

Therefore, we see that

$$A(1,x)\,\text{span}_{\mathbb{C}}\{v_1 + zv_2\} = \text{span}_{\mathbb{C}} \left\{ v_1 + \frac{e^{4i\beta_x}}{z} v_2 \right\},$$

when $A(1,x)$ is a reflection. We also have:

$$\begin{aligned}
A(1,x)v_1 &= \begin{bmatrix} \cos(2\beta_x) & \sin(2\beta_x) \\ \sin(2\beta_x) & -\cos(2\beta_x) \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = e^{2i\beta_x} \begin{bmatrix} 1 \\ -i \end{bmatrix}, \\
A(1,x)v_2 &= \begin{bmatrix} \cos(2\beta_x) & \sin(2\beta_x) \\ \sin(2\beta_x) & -\cos(2\beta_x) \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = e^{-2i\beta_x} \begin{bmatrix} 1 \\ i \end{bmatrix},
\end{aligned}$$

which implies that

$$\begin{aligned}
A(1,x)\,\text{span}_{\mathbb{C}}\{v_1\} &= \text{span}_{\mathbb{C}}\{v_2\}, \\
A(1,x)\,\text{span}_{\mathbb{C}}\{v_2\} &= \text{span}_{\mathbb{C}}\{v_1\}.
\end{aligned}$$

We have fully described the action of $A(1,x)$ on $\mathrm{Gr}_1(\mathbb{C}^2)$. Proposition 2.18 also tells us that $\mathrm{Gr}_1(\mathbb{C}^2)$ is homeomorphic to $\bar{\mathbb{C}}$, where the homeomorphism is given by $\psi : \mathrm{Gr}_1(\mathbb{C}^2) \to \bar{\mathbb{C}}$, defined by

$$\psi(\mathrm{span}_{\mathbb{C}}\{v_1 + zv_2\}) = z, \quad \psi(\mathrm{span}_{\mathbb{C}}\{v_2\}) = \infty.$$

From this, we may easily read off the family of maps $M(x) = \psi \circ A(1,x) \circ \psi^{-1} : \bar{\mathbb{C}} \to \bar{\mathbb{C}}$ to be the desired maps. The invertibility of $M(x)$ is straightforward to see. $\square$

We now define the skew product $N : X \times \bar{\mathbb{C}} \to X \times \bar{\mathbb{C}}$, given by $N(x,z) = (T(x), M(x)z)$. This represents the action of A on the Grassmannian $\mathrm{Gr}_1(\mathbb{C}^2)$ as a cocycle over the base dynamics T. We shall elaborate on some properties of N, and then construct the factor maps as mentioned above.

Lemma 3.5. *Let N be as defined above. Denote $P_C = \{z \ : \ |z| = C\} \cup \left\{z \ : \ |z| = \frac{1}{C}\right\}$ for $C \in [0,\infty]$ (where we define $P_0 = P_\infty = \{0,\infty\}$). Then sets of the form $X \times P_C$ are N-invariant; in particular, $X \times \mathbb{S}$ and $X \times \{0,\infty\}$ (and their complements) are N-invariant.*

Proof. Observe that N is invertible; the map $N^{-1}(x,z) = (T^{-1}(x), M(T^{-1}(x))^{-1}(z))$

is its inverse, which we can check:

$$\begin{aligned}
N^{-1}(N(x,z)) &= N^{-1}(T(x), M(x)z) \\
&= (T^{-1}(T(x)), M(T^{-1}(T(x)))^{-1}M(x)z) \\
&= (x, M(x)^{-1}M(x)z) \\
&= (x, z), \\
N(N^{-1}(x,z)) &= N(T^{-1}(x), M(T^{-1}(x))^{-1}z) \\
&= (T(T^{-1}(x)), M(T^{-1}(x))M(T^{-1}(x))^{-1}z) \\
&= (x, z).
\end{aligned}$$

It then suffices to show that if $(x,z) \in X \times P_C$, then $N(x,z) = (T(x), M(x)z) \in X \times P_C$. Clearly $T(x) \in X$, so consider that $|z| \in \{C, \frac{1}{C}\}$. Then

$$|M(x)z| \in \left\{|z|, \frac{1}{|z|}\right\} = \left\{C, \frac{1}{C}\right\},$$

so that $M(x)z \in P_C$ and so $N(x,z) \in X \times P_C$. Thus $X \times P_C$ is N-invariant, for any $C \in [0,\infty]$. In particular, since $P_1 = \mathbb{S}$, and $P_0 = \{0,\infty\}$, these sets are N-invariant, as are their complements. □

Note how P_1 is actually just $\mathbb{S}$, and is *not* a union of two disjoint circles. If we were to merely project onto $\mathbb{S}$, we would not be able to isolate the flipping action of $M(x)$. We will get around this by obtaining *two* factors of N, when N is restricted to either of the N-invariant sets $X \times (\bar{\mathbb{C}} \setminus \{0,\infty\})$ or $X \times (\bar{\mathbb{C}} \setminus \mathbb{S})$.

Lemma 3.6. *For each $x \in X$, define the maps $f_x : \mathbb{T} \to \mathbb{T}$ and $g_x : \mathbb{Z}_2 \to \mathbb{Z}_2$, given* *by:*

$$f_x(y) = \begin{cases} y + \dfrac{\alpha_x}{\pi} & x \in X_r, \\ \dfrac{2\beta_x}{\pi} - y & x \in X_f, \end{cases} \qquad g_x(a) = \begin{cases} a & x \in X_r, \\ a+1 & x \in X_f. \end{cases}$$

Define skew products $S : X \times \mathbb{T} \to X \times \mathbb{T}$ *and* $R : X \times \mathbb{Z}_2 \to X \times \mathbb{Z}_2$ *by:*

$$S(x,y) = (T(x), f_x(y)), \qquad R(x,a) = (T(x), g_x(a)).$$

Then S *is a factor of* $N\big|_{X\times(\bar{\mathbb{C}}\setminus\{0,\infty\})}$*, by* $\mathrm{id}\times\tau$*, and* R *is a factor of* $N\big|_{X\times(\bar{\mathbb{C}}\setminus\mathbb{S})}$*, by* $\mathrm{id}\times\iota$*.*

Proof. Clearly all of the maps are measurable. First, we deal with S; recall that the argument function arg satisfies $\arg(z_1z_2) = \arg(z_1) + \arg(z_2)$ for non-zero $z_1, z_2 \in \mathbb{C}$. For $x \in X$ and $z \in \bar{C} \setminus \{0,\infty\}$, we have:

$$\begin{aligned} (\mathrm{id}\times\tau)\circ N(x,z) &= \begin{cases} (T(x), \tau(e^{2i\alpha_x} z)) & x \in X_r \\ \left(T(x), \tau\left(\dfrac{e^{4i\beta_x}}{z}\right)\right) & x \in X_f \end{cases} \\ &= \begin{cases} \left(T(x), \tau(z) + \dfrac{\alpha_x}{\pi}\right) & x \in X_r \\ \left(T(x), \dfrac{2\beta_x}{\pi} - \tau(z)\right) & x \in X_f \end{cases} \\ &= (T(x), f_x(\tau(z))) = S\circ(\mathrm{id}\times\tau)(x,z). \end{aligned}$$

Hence S is a factor of N restricted to $X \times (\bar{\mathbb{C}} \setminus \{0,\infty\})$. Next, we deal with R; let

$z \in \bar{\mathbb{C}} \setminus \mathbb{S}$. We then have:

$$(\mathrm{id} \times \iota) \circ N(x,z) = \begin{cases} (T(x), \iota(e^{2i\alpha_x} z)) & x \in X_r \\ \left(T(x), \iota\left(\dfrac{e^{4i\beta_x}}{z}\right)\right) & x \in X_f \end{cases}$$

$$= \begin{cases} (T(x), \iota(z)) & x \in X_r \\ (T(x), \iota(z) + 1) & x \in X_f \end{cases}$$

$$= (T(x), g_x(\iota(z))) = R \circ (\mathrm{id} \times \iota)(x,z),$$

where we implicitly deal with the cases where $z = 0$ or $z = \infty$, since the computation is identical. Hence R is a factor of N restricted to $X \times (\bar{\mathbb{C}} \setminus \mathbb{S})$. □

It is easy to see that both S and R are invertible. A very important thing to note, now, is that there are measures which are preserved by S or by R. On $X \times \mathbb{T}$, we have the product measure $\mu \times \lambda$, where λ is the normalized Lebesgue measure, and on $X \times \mathbb{Z}$, we have the product measure $\mu \times c$, where c is the normalized counting measure on $\mathbb{Z}_2$ (which assigns weight $\frac{1}{2}$ to each set $\{0\}$ and $\{1\}$, and happens to be the normalized Haar measure on $\mathbb{Z}_2$). When $x \in X_r$, f_x is measure-preserving by Example 2.22 (or, of course, 2.24), and g_x is the identity map and hence obviously measure-preserving. When $x \in X_f$, f_x is measure-preserving by Example 2.23, and g_x is measure-preserving by Example 2.24. By Lemma 2.53, S preserves $\mu \times \lambda$ and R preserves $\mu \times c$. We are now in position to begin the main argument in the proof.

Step 2: For contradiction, we assume that there is an upper-triangularization for A. So we can find $C : X \to GL_2(\mathbb{C})$ such that $C(T(x))^{-1}A(1,x)C(x)$ is upper-triangular. The following lemma is true in more general settings than the one we currently have.

Lemma 3.7. *Let* $(X, \mathcal{B}, T)$ *be a dynamical system, and let* $A : \mathbb{Z} \times X \to GL_d(\mathbb{R})$ *be a measurable matrix cocycle over* T*. Suppose there is a complex block upper-triangularization for* A*, by* $C : X \to GL_d(\mathbb{C})$*, so that* $C(T(x))^{-1}A(1,x)C(x)$ *is block upper-triangular over* $\mathbb{C}$*. Then we may find a measurable equivariant family* $V(x)$ *of* 1*-D complex subspaces of* $\mathbb{C}^d$ *for the cocycle* A*.*

Proof. Let δ_x be the top-left entry of $C(T(x))^{-1}A(1,x)C(x)$. δ_x is non-zero for all x, since all three of these matrices are invertible. Let

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{C}^d$$

and let $V(x)$ be the family of 1-D subspaces given by

$$V(x) = \operatorname{span}_{\mathbb{C}}\{C(x)e_1\} = C(x)\operatorname{span}_{\mathbb{C}}\{e_1\}.$$

We apply $A(1,x)$ to $V(x)$:

$$A(1,x)C(x)e_1 = C(T(x))C(T(x))^{-1}A(1,x)C(x)e_1 = \delta_x C(T(x))e_1.$$

So $A(1,x)V(x) = V(T(x))$, ie. V is equivariant.

For the measurability of $V(x)$, we note that $x \mapsto C(x)$ is assumed to be measurable. $C(x)$ is continuous as a map on $\mathbb{C}^d$, and the induced action of $C(x)$ on the Grassmannian $\mathrm{Gr}_1(\mathbb{C}^d)$ is continuous because of that, hence measurable. Hence

$$x \mapsto V(x) = C(x)\operatorname{span}_{\mathbb{C}}\{e_1\}$$

is measurable, and the lemma is proved. □

We apply Lemma 3.7 in the case of our invertible and ergodic base dynamics $(X, \mathcal{B}, \mu, T)$ and orthogonal 2-by-2 real matrix cocycle A. We may compose the map $V : X \to \mathrm{Gr}_1(\mathbb{C}^2)$ with ψ, the homeomorphism from the Grassmannian to $\bar{\mathbb{C}}$, so we obtain $w : X \to \bar{\mathbb{C}}$, $w(x) = \psi \circ V(x)$. The equivariance property of V with respect to A carries over, in that we have:

$$M(x)(w(x)) = M(x) \circ \psi \circ V(x) = \psi \circ A(1, x)V(x) = \psi(V(T(x))) = w(T(x)). \quad (*)$$

Hence w is equivariant with respect to M. We shall almost exclusively work with w and M instead of V and A.

Step 3: Recall that equivariance is only an almost-everywhere statement, with respect to the measure μ on X. Let

$$B = \{x \in X \ : \ M(x)w(x) \neq w(T(x))\}\,;$$

by the definition of equivariance, we have $\mu(T^{-k}(B)) = 0$, for any $k \in \mathbb{Z}$, so that

$$C = \bigcup_{k \in \mathbb{Z}} T^{-k}(B) = \{x \in X \ : \ \text{equivariance of } w \text{ under } M \text{ fails during the orbit of } x\}$$

has measure zero and is T-invariant. We remove this set from X, so that our equivariance condition is now satisfied by every point in $X \setminus C$, which we shall call X for ease of notation. Also recall that $N(x, z) = (T(x), M(x)z)$ is the skew product of M over T. This leads us to an important lemma, regarding the graph of w (the lemma holds more generally, for any skew product of an equivariant map over invertible base dynamics).

Lemma 3.8. *Let w and N be given as above. The graph of w, denoted*

$$\Gamma_w = \{(x, w(x)) \ : \ x \in X\},$$

is N-invariant.

Proof. We have already shown that N is invertible. It then suffices to show that for any $x \in X$, $N(x, w(x)) \in \Gamma_w$, because T is invertible on X. We compute:

$$N(x, w(x)) = (T(x), M(x)w(x)) = (T(x), w(T(x))) \in \Gamma_w,$$

by $(*)$ (that is, the equivariance of w with respect to M). Thus Γ_w is N-invariant. □

We see that equivariance of w has yielded invariance of its graph under related skew product dynamics. It turns out that N has many invariant sets. Namely, by Lemma 3.5, any set of the form $X \times P_C$ is invariant for N. In particular, the graph of w must be contained in one such pair:

Lemma 3.9. *Consider the above situation, with $(X, \mathcal{B}, \mu, T)$, N, and w. There exists $C \in [0, \infty]$ such that for almost every $x \in X$, $w(x) \in P_C$.*

Proof. Define $k : X \to \mathbb{R}$ by

$$k(x) = \min \left\{ |w(x)|, \frac{1}{|w(x)|} \right\},$$

where we say that $\frac{1}{0} = \infty$ and vice versa, and observe that k is a T-invariant map on

X:

$$\begin{aligned} k(T(x)) &= \min\left\{|w(T(x))|\,, \frac{1}{|w(T(x))|}\right\} \\ &= \min\left\{|M(x)w(x)|\,, \frac{1}{|M(x)w(x)|}\right\} \\ &= \min\left\{|w(x)|\,, \frac{1}{|w(x)|}\right\} \qquad \text{since } M(x) \text{ either fixes } |z| \text{ or flips it} \\ &= k(x). \end{aligned}$$

Since T is ergodic, k must be constant almost everywhere, with constant $C \geq 0$. Then for almost every $x \in X$, we have $|w(x)| = C$ or $|w(x)| = \frac{1}{C}$, so that $w(x) \in P_C$, as desired. □

Let $D \subset X$ be the set of points for which $w(x) \notin P_C$, which has measure zero and is T-invariant. Remove this set from X, hence leaving $X \setminus D$, (which we shall still call X). Then Γ_w (when restricted to this new set of full measure) sits inside of P_C.

We have shown that the image of w sits inside of either a pair of distinct circles (including the pair of degenerate circles, $\{0, \infty\}$) lying outside the unit circle, one outside and one inside, or sits inside the unit circle itself (ie, when $C = 1$ in the lemma above, so $P_1 = \{z \ : \ |z| = 1\} = \mathbb{S}$). In each case, we push the set forward by either ι or τ, depending on which function contains the set in its domain. We may then make a statement about the graphs of $\tau \circ w$ and $\iota \circ w$. Recall that $S(x, y) = (T(x), f_x(y))$ and $R(x, a) = (T(x), g_x(a))$ are skew products on $X \times \mathbb{T}$ and $X \times \mathbb{Z}_2$, respectively, and that the graphs of $\tau \circ w$ and $\iota \circ w$ sit inside of those sets.

Lemma 3.10. *In the above situation, the set* $\mathrm{id} \times \tau(\Gamma_w) = \Gamma_{\tau \circ w}$ *is* S*-invariant, and the set* $\mathrm{id} \times \iota(\Gamma_w) = \Gamma_{\iota \circ w}$ *is* R*-invariant.*

Proof. The equality of sets above is by observing that

$$\text{id} \times \tau(x, w(x)) = (x, \tau(w(x))),\ \text{id} \times \iota(x, w(x)) = (x, \iota(w(x))).$$

To show invariance, in the first case we apply S:

$$S(x, \tau(w(x))) = (T(x), f_x(\tau(w(x)))) = (T(x), \tau(M(x)w(x))) = (T(x), \tau(w(T(x)))),$$

by $(*)$, so that $S(\Gamma_{\tau \circ w}) = \Gamma_{\tau \circ w}$. In the second case, we apply R:

$$R(x, \iota(w(x))) = (T(x), g_x(\iota(w(x)))) = (T(x), \iota(M(x)w(x))) = (T(x), \iota(w(T(x)))),$$

also by $(*)$, so that $R(\Gamma_{\iota \circ w}) = \Gamma_{\iota \circ w}$. $\square$

These invariant sets lie in measure-preserving systems, not just arbitrary measurable dynamical systems, so we may attempt to use classical ergodic theory. We could have tried to work with the map N on $X \times \bar{\mathbb{C}}$, but N had a huge number of invariant sets, and there is no obviously useful measure for our purposes (ie. to derive a contradiction via the existence of w); chances are, N would not be ergodic unless the measure was pathological. However, we are now working in much nicer spaces, so we have much less trouble with which to deal, in this regard.

Step 4: From these invariant graphs, we obtain obstructions to the ergodicity of the dynamics R and S.

Lemma 3.11. *Again considering the above situation, R and S cannot be ergodic, assuming the existence of the invariant graphs from Lemma 3.10.*

Proof. Due to the fact that the two transformations are on different spaces, we will deal with each case separately.

First, consider the two-disjoint-circles case, where $(X \times \mathbb{Z}_2, \mu \times c, R)$ is the measure-preserving system being studied. We know that $\Gamma_{\iota \circ w}$ is an R-invariant set. We may compute the measure of this set, using Fubini's theorem. The key point to notice is that $\iota \circ w$ takes exactly one of two values in Z_2, for every point in X. Hence, we obtain:

$$\begin{aligned}\mu \times c(\Gamma_{\iota \circ w}) &= \int_{X \times \mathbb{Z}_2} \mathbb{1}_{\Gamma_{\iota \circ w}}(x, a)\ d\mu \times c = \int_X \left(\frac{1}{2} \mathbb{1}_{\Gamma_{\iota \circ w}}(x, 0) + \frac{1}{2} \mathbb{1}_{\Gamma_{\iota \circ w}}(x, 1) \right) d\mu \\ &= \int_X \frac{1}{2}\ d\mu = \frac{1}{2}.\end{aligned}$$

Hence $\Gamma_{\iota \circ w}$ is an R-invariant set of non-full positive measure, which means that R cannot be ergodic.

We now deal with the circle case. We know that $\Gamma_{\tau \circ w}$, as above, is an S-invariant set; however, it is clear that the previous argument does not immediately work, because this set has measure zero. Instead, we'll find a non-constant S-invariant function.

To do this, first rewrite S using the explicit formula for f_x. Denoting $\tilde{\alpha}_x = \dfrac{\alpha_x}{\pi}$ and $\tilde{\beta}_x = \dfrac{2\beta}{\pi}$, we obtain:

$$S(x, y) = \begin{cases} (T(x), y + \tilde{\alpha}_x) & x \in X_r, \\ (T(x), \tilde{\beta}_x - y) & x \in X_f. \end{cases}$$

Next, denote $|\cdot, \cdot|_c : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ as the wrap-around distance on the unit interval, defined by

$$|x, y|_c = \min\{|x - y|, 1 - |x - y|\};$$

so $|.5, .7|_c = .2$ but $|.1, .9| = .2$, not $.8$. Finally, define $h(x, y) = |y - \tau \circ w(x)|_c$. Note the distance $|\cdot, \cdot|_c$ is continuous, hence measurable, and $\mathrm{id} \times \tau \circ w$ is measurable, so

$h = |\cdot,\cdot|_c \circ (\mathrm{id} \times \tau \circ c)$ is also measurable. Then we observe that:

$$\begin{aligned}
h(S(x,y)) &= \begin{cases} |y + \tilde{\alpha}_x - \tau \circ w(T(x))|_c & x \in X_r, \\ \left|\tilde{\beta}_x - y - \tau \circ w(T(x))\right|_c & x \in X_f; \end{cases} \\
&= \begin{cases} |y + \tilde{\alpha}_x - \tau \circ M(x)w(x)|_c & x \in X_r, \\ \left|\tilde{\beta}_x - y - \tau \circ M(x)w(x)\right|_c & x \in X_f; \end{cases} \\
&= \begin{cases} |y + \tilde{\alpha}_x - (\tau \circ w(x) + \tilde{\alpha}_x)|_c & x \in X_r, \\ \left|\tilde{\beta}_x - y - (\tilde{\beta}_x - \tau \circ w(x))\right|_c & x \in X_f; \end{cases} \\
&= \begin{cases} |y - \tau \circ w(x)|_c & x \in X_r, \\ |\tau \circ w(x) - y|_c & x \in X_f; \end{cases} \\
&= h(x,y).
\end{aligned}$$

Hence h is S-invariant. However, h is certainly not constant on $X \times \mathbb{T}$, because it is exactly the vertical distance of a point in $X \times \mathbb{T}$ away from the graph of V. Hence S cannot be ergodic, by Theorem 2.37, since h is measurable. □

Proof of Theorem 3.3. We have assumed that R is ergodic with respect to $\mu \times \lambda$, and that S is ergodic with respect to $\mu \times \lambda$. Assume, for contradiction, that the cocycle A is block-upper-triangular. Then by Lemma 3.7, there is an equivariant family of 1-dimensional subspaces, which by Lemma 3.8 yields an invariant graph on $\bar{\mathbb{C}}$, which lies either on the unit circle or on two disjoint circles related by inversion, by Lemma 3.9. By Lemma 3.10, we know that the projection of the graph down to either $X \times \mathbb{Z}_2$ or $X \times \mathbb{T}$ is invariant under the dynamics, and Lemma 3.11 says that because of this, R and S cannot be ergodic. This is a contradiction, which means that the assumption about the existence of an upper-triangularization of A, which is the only other assumption we have made, must be incorrect, and therefore A cannot be upper-triangularized over $\mathbb{C}$. $\square$

We list two remarks about the proof, the latter addressing the fact that the statement regarding block upper-triangularization over $\mathbb{R}$ was not explicitly proven.

1. An alternate proof for showing that the ergodicity of S leads to a contradiction involves computing the measure of the set $h^{-1}[0, \frac{1}{4}]$. This set is the set of all points in $X \times \mathbb{T}$ which are at most distance $\frac{1}{4}$ in the vertical direction from the graph of $\tau \circ w$. Utilizing Fubini's theorem essentially returns the same integration as in the two circles case. This argument is (lovingly) called the 'invariant ponytail' argument, as $h^{-1}[0, \frac{1}{4}]$ looks like a fattened version of $\Gamma_{\tau \circ w}$, which looks like a ponytail if the graph is drawn smoothly enough.

2. It should be noted that if one merely wished to determine if a real cocycle of 2-by-2 matrices, satisfying all of the same hypotheses as above, was measurably equivariantly block upper-triangularizable over the *real* numbers, one would note that $\mathrm{Gr}_1(\mathbb{R}^2)$ is homeomorphic to the interval $[0, \pi)$ by considering the angle from an axis (usually the horizontal axis in the plane), and thus to $\mathbb{T}$. The dynamics arising from this construction are exactly the dynamics derived in the one circle case. Thus this is exactly what we could do to show that A is not equivariantly upper-triangularizable over the real numbers. It must be noted that this occurred due to the particular structure of the action of M on $\bar{\mathbb{C}}$; the graph of w remained away from 0 and ∞.

Now that we have a sufficient condition for a matrix cocycle A to not be block upper-triangularizable, we will explore three examples. The first example will be a cocycle which is not block upper-triangularizable over the real numbers, but may be diagonalized over the complex numbers. The example is a good illustration of how both conditions are explicitly needed. The second and third examples are cocycles which are not block upper-triangularizable over $\mathbb{R}$ or $\mathbb{C}$; they differ in that they are over different base dynamics spaces, and that the technical details of proving ergodicity are considerably different.

3.3 Example 1: rotation cocycle over a rotation

Let $(\mathbb{T}, \mathcal{B}, \lambda, T)$ be the irrational rotation by η over the unit interval with normalized Lebesgue measure, and consider the matrix cocycle A generated by

$$A(1,x) = \begin{bmatrix} \cos(\pi x) & -\sin(\pi x) \\ \sin(\pi x) & \cos(\pi x) \end{bmatrix} = \mathrm{rot}_{\pi x}.$$

Theorem 3.12. *The cocycle A is block upper-triangularizable over $\mathbb{C}$, but not over $\mathbb{R}$.*

Proof. The matrix $A(1,x)$ is a rotation for all $x \in \mathbb{T}$, with rotation angle $\alpha_x = \pi x$, and $X_r = \mathbb{T}$, using the notation in 3.4. Then considering $M(x) = \psi \circ A(1,x) \circ \psi^{-1}$, we see that $M(x)$ acts on $\bar{\mathbb{C}}$ by

$$M(x)z = e^{2\pi i x} z,$$

for all $x \in \mathbb{T}$. From here, we construct the restricted map $M_\circ(x)$, and the factor map $f_x : \mathbb{T} \to \mathbb{T}$; we see that $f_x(y) = x + y$ for all $x \in \mathbb{T}$. Then the skew product $S : \mathbb{T} \times \mathbb{T} \to \mathbb{T} \times \mathbb{T}$ is given by

$$S(x,y) = (T(x), f_x(y)) = (x + \eta, x + y).$$

This map was shown to be ergodic in Example 2.55, and so we see that by Theorem

3.3, A cannot be block upper-triangularized over the real numbers.

However, observe that each rotation matrix $A(1,x)$ is diagonalized by the same matrix

$$C(x) = C = [v_1\ v_2] = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix},$$

where the notation for v_1 and v_2 comes from Lemma 3.4. In particular, we have

$$\begin{aligned} C(T(x))^{-1}A(1,x)C(x) &= C^{-1}A(1,x)C \\ &= \frac{-1}{2i}\begin{bmatrix} -i & -1 \\ -i & 1 \end{bmatrix}\begin{bmatrix} \cos(\pi x) & -\sin(\pi x) \\ \sin(\pi x) & \cos(\pi x) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \\ &= \frac{-1}{2i}\begin{bmatrix} -i & -1 \\ -i & 1 \end{bmatrix}\begin{bmatrix} e^{-\pi i x} & e^{\pi i x} \\ ie^{-\pi i x} & -ie^{\pi i x} \end{bmatrix} \\ &= \frac{-1}{2i}\begin{bmatrix} -2ie^{-\pi i x} & 0 \\ 0 & -2ie^{\pi i x} \end{bmatrix} \\ &= \begin{bmatrix} e^{-\pi i x} & 0 \\ 0 & e^{\pi i x} \end{bmatrix}. \end{aligned}$$

Hence we see that $C(T(x))^{-1}A(1,x)C(x)$ is diagonal in $\mathbb{C}$, and so upper-triangular. □

There are *two* equivariant subspaces, here, which are independent of x:

$$V_1(x) = \text{span}_{\mathbb{C}}\{Ce_1\} = \text{span}_{\mathbb{C}}\{v_1\}, \quad V_2(x) = \text{span}_{\mathbb{C}}\{Ce_2\} = \text{span}_{\mathbb{C}}\{v_2\}.$$

In the dynamical picture, these are the points 0 and ∞ in $\bar{\mathbb{C}}$, and we have seen that rotations leave these points fixed. This gives rise to R-invariant sets in $\mathbb{T} \times \mathbb{Z}_2$ which have positive measure, and there is nothing to prevent this, as the map R is given by

$R(x,a) = (T(x), a)$. Indeed, $\mathbb{T} \times \{0\}$ and $\mathbb{T} \times \{1\}$ are R-invariant sets of positive (but less than full) measure.

3.4 Example 2: rotation and flip cocycle over a rotation

Let the base dynamics space be the same as the previous example: $(\mathbb{T}, \mathcal{B}, \lambda, T)$, where T is rotation on the unit interval by an irrational η, with normalized Lebesgue measure. This time, we define a matrix cocycle A over T, with

$$A(1,x) = \begin{cases} \begin{bmatrix} \cos(\pi\alpha) & -\sin(\pi\alpha) \\ \sin(\pi\alpha) & \cos(\pi\alpha) \end{bmatrix} & x \in [0, 1-\eta), \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & x \in [1-\eta, 1). \end{cases}$$

Theorem 3.13. *The cocycle A over $(\mathbb{T}, \mathcal{B}, \lambda, T)$ as defined above is not block upper-triangularizable over $\mathbb{C}$.*

To do the proof, we fit the situation into the general framework. $A(1,x)$ is a rotation by $\pi\alpha$ for $x \in [0, 1-\eta)$, and a reflection in the horizontal axis for $x \in [1-\eta, 1)$. In the notation, $X_r = [0, 1-\eta)$, with a fixed rotation angle $\alpha_x = \pi\alpha$, and $X_f = [1-\eta, 1)$, with fixed reflection axis $\beta_x = 0$. Computing $M(x)$ yields

$$M(x)z = \begin{cases} e^{2\pi i\alpha} z & x \in [0, 1-\eta), \\ \dfrac{1}{z} & x \in [1-\eta, 1). \end{cases}$$

We compute both the two-point and the circle extensions of $\mathbb{T}$, beginning with S.

The map $f_x : \mathbb{T} \to \mathbb{T}$ is given by:

$$f_x(y) = \begin{cases} y + \alpha & x \in [0, 1-\eta), \\ 1 - y & x \in [1-\eta, 1), \end{cases}$$

where we may write '$1 - y$' in place of '$-y$' to emphasize the modulo 1 aspect of $\mathbb{T}$ (we ignore doing this in the rotation case for notational simplicity, for the moment). The resulting skew product $S : \mathbb{T} \times \mathbb{T} \to \mathbb{T} \times \mathbb{T}$ is:

$$S(x, y) = \begin{cases} (x + \eta, y + \alpha) & x \in [0, 1-\eta), \\ (x + \eta, 1 - y) & x \in [1-\eta, 1). \end{cases}$$

The measure-preserving system here is $(\mathbb{T} \times \mathbb{T}, \lambda \times \lambda, S)$. Next, the map $g_x : \mathbb{Z}_2 \to \mathbb{Z}_2$ is given by:

$$g_x(a) = \begin{cases} a & x \in [0, 1-\eta), \\ a + 1 & x \in [1-\eta, 1), \end{cases}$$

which gives us the skew product $R : \mathbb{T} \times \mathbb{Z}_2 \to \mathbb{T} \times \mathbb{Z}_2$, with

$$R(x, a) = \begin{cases} (x + \eta, a) & x \in [0, 1-\eta), \\ (x + \eta, a + 1) & x \in [1-\eta, 1). \end{cases}$$

The measure-preserving system is $(\mathbb{T} \times \mathbb{Z}_2, \lambda \times c, R)$. We will now show that both R and S are ergodic, which by Theorem 3.3 will show that A cannot be upper-triangularized (over $\mathbb{R}$ or over $\mathbb{C}$).

Proposition 3.14. *R, in the above situation, is ergodic with respect to $\lambda \times c$.*

The flip makes this non-trivial, because as the map iterates, we do not know exactly when the Z_2 component is flipped. We can attempt to get rid of the problem with the flip by inducing the map R on $B = X_f \times \mathbb{Z}_2 = [1-\eta) \times \mathbb{Z}_2$, to get the induced map R_B; this can be motivated by recalling Proposition 2.51.

Lemma 3.15. *For B as defined above, the induced map R_B on $(B, (\lambda \times c)_B)$ is measure-theoretically isomorphic to the map $\tilde{R}_B : \mathbb{T} \times \mathbb{Z}_2 \to \mathbb{T} \times \mathbb{Z}_2$ on $(\mathbb{T} \times \mathbb{Z}_2, \lambda \times c)$, given by*

$$\tilde{R}_B(x, a) = (x + \beta, a + 1),$$

where $\beta = \left\{\frac{1}{\eta}\right\}$.

Proof. Intuitively, this map acts on points (x, a) in B by applying R once in the flip case, because the $\mathbb{T}$ component x is in $[1-\eta, 1)$, and then applying R in the rotation case until x returns to $[1-\eta, 1)$; this means that in the $\mathbb{Z}_2$ component, the cumulative action is to flip a to $a+1$. More explicitly, we may compute R_B. The return time is dependent only on the $\mathbb{T}$ component (since B is the product of X_f and all of $\mathbb{Z}_2$), so we have $n_{R,B} = n_{T,X_f}$, and by our work in Example 2.49, letting k be the unique positive integer such that $k\eta < 1(k+1)\eta$ and letting $q = 2 - (k+1)\eta$, we compute:

$$\begin{aligned}
R_B(x, a) &= R^{n_{R,B}(x,a)}(x, a) = R^{n_{T,X_f}(x)}(x, a) \\
&= \begin{cases} R^{k+1}(x, a) & x \in [1-\eta, q) \\ R^k(x, a) & x \in [q, 1) \end{cases} \\
&= \begin{cases} (x + (k+1)\eta, a+1) & x \in [1-\eta, q) \\ (x + k\eta, a+1) & x \in [1-\eta, q) \end{cases}.
\end{aligned}$$

As well, we may make the same coordinate change as in Example 2.49, which maps $[1-\eta, 1)$ to $\mathbb{T}$ via $\phi(x) = \dfrac{1-x}{\eta}$. Then $\phi \times \mathrm{id}$ is a coordinate change mapping B to $\mathbb{T} \times \mathbb{Z}_2$, and is a measure-theoretic isomorphism between the two spaces by our work in Example 2.49. We have

$$\tilde{R}_B(x,a) = (x+\beta, a+1) = (x,a) + (\beta, 1),$$

and we are done proving the lemma. □

Lemma 3.16. *The map $\tilde{R}_B$ as defined above is ergodic, with respect to $\lambda \times c$ on $\mathbb{T} \times \mathbb{Z}_2$.*

Proof. We observe that $\tilde{R}_B$ is a rotation by $(\beta, 1)$ on the compact Abelian group $\mathbb{T} \times \mathbb{Z}_2$; this puts us in the situation to use Proposition 2.43, which says that $\tilde{R}_B$ will be ergodic if the only character satisfying $\gamma(\beta, 1) = 1$ is the trivial character $\gamma \equiv 1$. By Lemma A.5, the characters on $\mathbb{T} \times \mathbb{Z}_2$ have the form $(x,a) \mapsto e^{2\pi i n x}(-1)^{ab}$, where $n \in \mathbb{Z}$ and $b \in \mathbb{Z}_2$. Let $\gamma_{n,b}$ be the character corresponding to n and b; we apply $\gamma_{n,b}$ to $(\beta, 1)$ and set it equal to 1:

$$\gamma_{n,b}(\beta, 1) = e^{2\pi i n \beta}(-1)^{1 \cdot b} = (e^{2\pi i \beta})^n (-1)^b = 1.$$

Since β is irrational, $e^{2\pi i \beta}$ is never a root of unity and hence never a root of 1 or of -1; this shows that $n = 0$, which implies that $b = 0$. Therefore the only character which satisfies that equation is $\gamma_{0,0} \equiv 1$, so $\tilde{R}_B$ is ergodic. □

Proof of Proposition 3.14. We apply the lemmas in reverse order to prove Proposition 3.14. $\tilde{R}_B$ is ergodic by Lemma 3.16, and so R_B is ergodic because those two maps are isomorphic by Lemma 3.15. Then Proposition 2.51 applies, thanks to Lemma 2.58, and so R is ergodic. □

We now tackle the circle extension case. This case requires slightly more work than the two-point extension.

Proposition 3.17. *S, as defined before, is ergodic with respect to $\lambda \times \lambda$.*

As in the two-point extension case, the flip makes proving ergodicity non-trivial. The same strategy we tried before will serve us well, however, alongside a tweak or two. Let $B = X_f \times \mathbb{T}$; we shall compute the induced map. We keep the same notation from the previous case when speaking about the base map T_η, so that k is still the unique positive integer such that $k\eta < 1 < (k+1)\eta$, $q = 2-(k+1)\eta$, and $\beta = \left\{\frac{1}{\eta}\right\}$. Note that

$$\frac{1-q}{\eta} = \frac{1-(2-(k+1)\eta)}{\eta} = k+1-\frac{1}{\eta} = 1-\beta,$$

which also came up in Example 2.49.

Lemma 3.18. *For B as above, the induced map S_B is measure-theoretically isomorphic to $\tilde{S}_B$ acting on $\mathbb{T}\times\mathbb{T}$, given by*

$$\tilde{S}_B(x,y) = \begin{cases} (x+\beta, k\alpha - y) & x \in [1-\beta, 1), \\ (x+\beta, (k-1)\alpha - y) & x \in [0, 1-\beta). \end{cases}$$

Proof. As before, the return time for S to B is only dependent on the return time for T to X_f, ie. $n_{S,B} = n_{T,X_f}$. Applying S_B is then the same as applying S in the flip case exactly once, then applying S in the rotation case until returning to B. We have:

$$\begin{aligned} S_B(x,y) &= S^{n_{S,B}(x,y)}(x,y) = S^{n_{T,X_f}(x)}(x,y) \\ &= \begin{cases} (x+(k+1)\eta, k\eta - y) & x \in [1-\beta,1), \\ (x+k\eta, (k-1)\eta - y) & x \in [0,1-\beta). \end{cases} \end{aligned}$$

We then apply the map $\phi\times\mathrm{id}$ and obtain a map on $\mathbb{T}\times\mathbb{T}$ isomorphic in the measure-theoretic sense to S_B, again by our work in Example 2.49. This map, $\tilde{S}_B$, is given by

$$\tilde{S}_B(x,y) = \begin{cases} (x+\beta, k\eta - y) & x \in [1-\beta, 1), \\ (x+\beta, (k-1)\eta - y) & x \in [0, 1-\beta). \end{cases}$$

Hence we are done. □

Here, we note that if we were to square $\tilde{S}_B$, we would eliminate the flip in the variable y, and be left with something only involving rotations, which would possibly be much less awkward. This strategy is justified by the following abstract lemma.

Lemma 3.19. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system. If T^2 is ergodic, then so is T.*

Proof. Let $A \subset X$ be a T-invariant set, so that $T^{-1}(A) = A$. Then we have:

$$(T^2)^{-1}(A) = T^{-2}(A) = T^{-1}(T^{-1}(A)) = T^{-1}(A) = A,$$

that is, A is T^2-invariant, and since T^2 is ergodic, A has either zero or full measure. This holds for all T-invariant sets, and so T is ergodic. □

Denote $P = (\tilde{S}_B)^2$, and let T_β be rotation by β. To compute P, we need to consider how the base moves with respect to the partition $\{[0, 1-\beta), [1-\beta, 1)\}$, so that the second application of the map may be determined. This is because there are many possible scenarios, depending on where $x \in \mathbb{T}$ starts. As well, since we have two applications of T_β, we care about how big 2β is. If $\beta > \frac{1}{2}$, then $2\beta > 1$; by Lemmas 2.29 and 2.41, without loss of generality we assume that $\beta < \frac{1}{2}$, for notational convenience.

Lemma 3.20. *Let $\tilde{S}_B$ be as above. If $\beta < \frac{1}{2}$, then the map $P = (\tilde{S}_B)^2$ is given by:*

$$P(x,y) == \begin{cases} (x+2\beta, y) & x \in [0, 1-2\beta) \\ (x+2\beta, y+\alpha) & x \in [1-2\beta, 1-\beta) \\ (x+2\beta, y-\alpha) & x \in [1-\beta, 1) \end{cases} \cdot$$

Proof. For $\beta < \frac{1}{2}$, we have $\beta < 1-\beta$. Then we get:

$$T_\beta[0, 1-\beta) = [\beta, 1) = [\beta, 1-\beta) \cup [1-\beta, 1),$$
$$T_\beta[1-\beta, 1) = [0, \beta).$$

We see that we know how to correctly choose branches of $\tilde{S}_B$ depending in which set $T_\beta(x)$ ends up: one of $[0,\beta)$, $[\beta, 1-\beta)$, or $[1-\beta, 1)$. We may then take inverse images of those sets by T_β:

$$T_\beta^{-1}[0, \beta) = [1-\beta, 1),$$
$$T_\beta^{-1}[\beta, 1-\beta) = [0, 1-2\beta),$$
$$T_\beta^{-1}[1-\beta, 1) = [1-2\beta, 1-\beta).$$

We also have $0 < 1-2\beta < 1-\beta < 1$, so that

$$\big\{[0, 1-2\beta), [1-2\beta, 1-\beta), [1-\beta, 1)\big\}$$

is a partition of $\mathbb{T}$. Each of these sets lies strictly inside one of $\{[0, 1-\beta), [1-\beta, 1)\}$, and their images lies strictly inside of one of those sets, also.

We may now compute $P = (\tilde{S}_B)^2$ directly:

$$P(x,y) = \begin{cases} \tilde{S}_B(x+\beta, (k-1)\alpha - y) & x \in [0, 1-2\beta) \quad (x+\beta \in [\beta, 1-\beta)) \\ \tilde{S}_B(x+\beta, (k-1)\alpha - y) & x \in [1-2\beta, 1-\beta) \quad (x+\beta \in [1-\beta, 1)) \\ \tilde{S}_B(x+\beta, k\alpha - y) & x \in [1-\beta, 1) \quad (x+\beta \in [0, \beta)) \end{cases}$$

$$= \begin{cases} (x+2\beta, (k-1)\alpha - ((k-1)\alpha - y)) & x \in [0, 1-2\beta) \\ (x+2\beta, k\alpha - ((k-1)\alpha - y)) & x \in [1-2\beta, 1-\beta) \\ (x+2\beta, (k-1)\alpha - (k\alpha - y)) & x \in [1-\beta, 1) \end{cases}$$

$$= \begin{cases} (x+2\beta, y) & x \in [0, 1-2\beta) \\ (x+2\beta, y+\alpha) & x \in [1-2\beta, 1-\beta) \\ (x+2\beta, y-\alpha) & x \in [1-\beta, 1) \end{cases} \cdot$$

□

This seems no better than before, especially since there are now three branches to consider. However, the base rotation is 2β, and the interval $[1-2\beta, 1)$ has the same length as the rotation, and removing the portion of the space where y doesn't change will simplify things. Moreover, for $B = [1-2\beta, 1) \times \mathbb{T}$, Lemma 2.58 says that we have

$$\lambda \times \lambda \left(\bigcup_{i=0}^{\infty} P^{-i}(B) \right) = 1.$$

Thus we consider the induced map P_B. We are exactly in the same situation as Example 2.49, so we may also make the same coordinate change to get back to a map on $\mathbb{T} \times \mathbb{T}$. Hence, we let $\zeta = \left\{ \frac{1}{2\beta} \right\}$ and use the coordinate change $\xi(x) = \frac{1-x}{2\beta}$ to obtain a map Q which is measure-theoretically isomorphic to P_B. We have

$$\xi(1-\beta) = \frac{1-(1-\beta)}{2\beta} = \frac{1}{2},$$

and so Q is given by:

$$Q(x,y) = \begin{cases} (x+\zeta, y-\alpha) & x \in [0, \frac{1}{2}) \\ (x+\zeta, y+\alpha) & x \in [\frac{1}{2}, 1). \end{cases}$$

We would like to prove that Q is ergodic. To do this, we may use the following two results:

Proposition 3.21. *Suppose $\sigma : \mathbb{T} \to \mathbb{T}$ is measure-preserving and ergodic with respect to Lebesgue measure, and let $f : \mathbb{T} \to \mathbb{R}$ be a measurable function, with range $f(\mathbb{T}) \subset \alpha\mathbb{Z}$, where α is irrational. Let $\mathbb{T}^2$ have the usual Lebesgue product measure and Borel sets, and let $T_f : \mathbb{T}^2 \to \mathbb{T}^2$ be the skew product extension of σ and f to $\mathbb{T}^2$, so that:*

$$T_f(x,y) = (\sigma(x), y + f(x)).$$

Let $\tilde{T}_f : \mathbb{T} \times \alpha\mathbb{Z} \to \mathbb{T} \times \alpha\mathbb{Z}$ be the skew product extension of σ and f to $\mathbb{T} \times \alpha\mathbb{Z}$ with the product measure $\lambda \times c$ (Lebesgue and counting, with the discrete σ-algebra for the counting measure), so that:

$$\tilde{T}_f(x, n\alpha) = (\sigma(x), n\alpha + f(x)).$$

Then if $\tilde{T}_f$ is ergodic, so is T_f.

Proof. Let $h : \mathbb{T}^2 \to \mathbb{R}$ be a bounded measurable function invariant under T_f, so $h \circ T_f = h$. We shall show that h must be a.e. constant; this implies that T_f is ergodic, by condition (3) of Theorem 2.37. For $y \in \mathbb{T}$, define the measurable function

$$\pi_y : \mathbb{T} \times \alpha\mathbb{Z} \to \mathbb{T}^2,\ \pi_y(x, n\alpha) = (x, y + n\alpha).$$

Then we see that $T_f \circ \pi_y = \pi_y \circ \tilde{T}_f$. In addition, define $\tilde{h}_y = h \circ \pi_y$, so that $\tilde{h}_y$ is measurable. Since π_y intertwines the dynamics on the two spaces, we get the following:

$$\tilde{h}_y \circ \tilde{T}_y = h \circ \pi_y \circ \tilde{T}_f = h \circ T_f \circ \pi_y = h \circ \pi_y = \tilde{h}_y.$$

Thus $\tilde{h}_y$ is invariant under $\tilde{T}_f$, and so is constant a.e. with respect to the product measure $\lambda \times c$, since $\tilde{T}_f$ is ergodic.

We wish to use the fact that $\tilde{h}_y$ is a.e. constant for each $y \in \mathbb{T}$ to show that h is constant a.e. To do this, we make an intermediate step. Define

$$I : \mathbb{T} \to \mathbb{R},\ I(y) = \int_0^1 h(x, y)\ dx = \int_0^1 \tilde{h}_y(x, 0)\ dx.$$

Because h is bounded, I is not infinite, hence well-defined, and by Fubini's theorem, I is measurable. Moreover, we have the following, since $\tilde{h}_y$ is a.e. constant on $\mathbb{T} \times \alpha\mathbb{Z}$:

$$\begin{aligned} I(y + \alpha) &= \int_0^1 h(x, y + \alpha)\ dx = \int_0^1 \tilde{h}_y(x, \alpha)\ dx = \int_0^1 \tilde{h}_y(x, 0)\ dx \\ &= \int_0^1 h(x, y)\ dx = I(y). \end{aligned}$$

$y \mapsto y + \alpha$ is an ergodic map on $\mathbb{T}$, thus we see that I is a.e. constant on $\mathbb{T}$; write $I(y) = C$ for a.e. $y \in \mathbb{T}$. Note that for all y, $\tilde{h}_y$ is a.e. constant on $\mathbb{T} \times \alpha\mathbb{Z}$, so we see that for a.e. $x \in \mathbb{T}$, $\tilde{h}_y(x, 0) = I(y)$. Denote $Y_G = \{y \in \mathbb{T}\ :\ I(y) = C\}$; this set has full measure in $\mathbb{T}$. If $y \in Y_G$, then for a.e. x, $h(x, y) = \tilde{h}_y(x, 0) = C$. Computing the measure of the set of points where $h \neq C$ via Fubini's Theorem yields the final statement: $h = C$ almost everywhere. Hence T_f is ergodic. □

Proposition 3.22 (Schmidt, [26]). *Consider the space $\mathbb{T} \times \alpha\mathbb{Z}$ as defined in the previous proposition. Define the map $\tilde{Q}$ on $\mathbb{T} \times \alpha\mathbb{Z}$ by*

$$\tilde{Q}(x, n\alpha) = \begin{cases} (x + \zeta, (n+1)\alpha) & x \in [0, \frac{1}{2}), \\ (x + \zeta, (n-1)\alpha) & x \in [\frac{1}{2}, 1). \end{cases}$$

Then $\tilde{Q}$ is an ergodic measure-preserving transformation.

Proof. The result follows from Theorem 3.9, Corollary 5.4, and Theorem 12.8 of [26], upon the relabelling of $\mathbb{Z}$ to $\alpha\mathbb{Z}$. □

Lemma 3.23. *The map Q as defined above is ergodic, with respect to $\lambda \times \lambda$.*

Proof. The map $\tilde{Q}$ in Proposition 3.22 is related to the map Q in exactly the way outlined in Proposition 3.21, so since $\tilde{Q}$ is ergodic, Q must be ergodic also. □

Proof of Proposition 3.17. To conclude the proof that the original map S is ergodic, Q is ergodic by Lemma 3.23. Thus, so are P_B, P, and $\tilde{S}_A$, by the isomorphism, Proposition 2.51 (with Lemma 2.58) and Lemma 3.19, therefore S is ergodic, again by Proposition 2.51 and Lemma 2.58. Hence we are done. □

Proof of Theorem 3.13. By Propositions 3.14 and 3.17, R and S are ergodic, which by Theorem 3.3 allows us to conclude that the original matrix cocycle A is not block upper-triangular over $\mathbb{C}$. □

It should be noted that the proof that $\tilde{Q}$ is ergodic, as done by Schmidt in [26] and for a specific case in [27], relies on an object associated to a measure-preserving transformation involving a cocycle on $\mathbb{Z}$ called the set of *essential values* of the transformation. The mechanisms behind this object are rather technical and somewhat elaborate, and out of the scope of this thesis. In lieu of going through an extensive proof of this material, we shall state another example, which uses more well-known mathematical techniques and avoids using a black-box theorem.

3.5 Example 3: rotation and flip cocycle over a Bernoulli shift

Let the base dynamics space be the left Bernoulli shift on two symbols each with weight $\frac{1}{2}$, denoted $(X, \mathcal{B}, \mu, T)$, which was studied in Example 2.25. Let A be a matrix cocycle over T, generated by the map $A(1, \cdot) : X \to O_2(\mathbb{R})$, given by:

$$A(1,x) = \begin{cases} \begin{bmatrix} \cos(\pi\alpha) & -\sin(\pi\alpha) \\ \sin(\pi\alpha) & \cos(\pi\alpha) \end{bmatrix} & x_0 = 0, \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & x_0 = 1. \end{cases}$$

Theorem 3.24. *The cocycle A over $(X, \mathcal{B}, \mu, T)$ as defined above is not block upper-triangularizable over $\mathbb{C}$.*

As in the previous example, we shall fit this system into the general framework, and then work towards proving that the resulting maps are ergodic. Recall that we use the notation $\pi_0^{-1}\{a\} = C(x_0 = a)$, where π_0 is the projection of X onto the 0^{th} coordinate, so that this set is the collection of $x \in X$ with $x_0 = a$. We see that $A(1, x)$ is a rotation by $\pi\alpha$ for $x \in C(x_0 = 0)$, so $X_r = C(x_0 = 0)$ with $\alpha_x = \pi\alpha$, and $A(1, x)$ is a reflection in the horizontal axis for $x \in C(x_0 = 1)$, so $X_f = C(x_0 = 1)$ with $\beta_x = 0$. The cocycle is, in flavour, identical to the cocycle in the previous example, except that the underlying space is different.

The next step is to compute the action of $A(1,x)$ on $\bar{\mathbb{C}}$, which was denoted $M(x)$. We have:

$$M(x)z = \begin{cases} e^{2\pi i\alpha}z & x \in C(x_0 = 0), \\ \dfrac{1}{z} & x \in C(x_0 = 1). \end{cases}$$

We can then compute the circle extension and the two-point extension dynamics. First, in the circle extension case, we have

$$f_x(y) = \begin{cases} y + \alpha & x \in C(x_0 = 0), \\ 1 - y & x \in C(x_0 = 1), \end{cases}$$

so that the map $S : X \times \mathbb{T} \to X \times \mathbb{T}$ is given by

$$S(x,y) = \begin{cases} (T(x), y + \alpha) & x \in C(x_0 = 0) \\ (T(x), 1 - y) & x \in C(x_0 = 1). \end{cases}$$

Then, in the two-point extension case, we have $g_x : \mathbb{Z}_2 \to \mathbb{Z}_2$, given by

$$g_x(a) = \begin{cases} a & x \in C(x_0 = 0) \\ a + 1 & x \in C(x_0 = 1), \end{cases}$$

so that we obtain the map $R : X \times \mathbb{Z}_2 \to Z \times \mathbb{Z}_2$, with formula

$$R(x,a) = \begin{cases} (T(x), a) & x \in C(x_0 = 0) \\ (T(x), a + 1) & x \in C(x_0 = 1). \end{cases}$$

We shall now prove that R and S are ergodic maps on their respective spaces, starting with R.

Proposition 3.25. *R as above is ergodic on $X \times \mathbb{Z}_2$, with respect to $\mu \times c$.*

We utilize the same strategy as before: namely, we shall induce upon $B = X_f \times \mathbb{Z}_2 = C(x_0 = 1) \times \mathbb{Z}_2$ and show that the resulting map is ergodic, via Proposition 2.51.

Lemma 3.26. *Let (Z, ν, L) be the left Bernoulli shift on the bilateral shift space over the alphabet $\mathbb{N}$, where each symbol $k \in \mathbb{N}$ is given the weight $\frac{1}{2^{k+1}}$ (as seen in Example 2.50). The induced map R_B on $(B, (\mu \times c)_B)$ is measure-theoretically isomorphic to the map $\tilde{R}_B$ on $(Z \times \mathbb{Z}_2, \nu \times c)$, with*

$$\tilde{R}_B(z, a) = (L(z), a + 1).$$

Proof. As in Example 2.50, let B_+ be the set of points in A which have finitely many 1's in the positive direction, and B_- be the set of points in A which have finite many 1's in the negative direction; then $B_+ \cup B_-$ has measure 0 in A and is T-invariant. We remove $B_+ \cup B_-$ from A and continue to write $A \setminus (B_+ \cup B_-)$ as A. Then define the map $\phi : A \to Z$, by

$$\phi(x) = (\dots k_{-1} \cdot k_0 k_1 \dots) \in Z,$$

since elements x of A can be written as

$$(\dots 1 \underbrace{0 \dots 0}_{k_{-1}} \cdot 1 \underbrace{0 \dots 0}_{k_0} 1 \underbrace{0 \dots 0}_{k_1} \dots).$$

This defines a measure isomorphism of the two spaces.

Also in Example 2.50, the induced map T_{X_f} was computed to be the map which shifted the sequence left until the next 1 was encountered. In symbols, it is:

$$\begin{aligned} T_{X_f}(x) &= T^{n_{X_f,T}}(x) \\ &= \begin{cases} T^k(x), & x \in C(x_0x_1\dots x_{k-1}x_k = 1\underbrace{0\dots0}_{k-1}1) \end{cases} \\ &= \begin{cases} (\dots 1\underbrace{0\dots0}_{k-1}\cdot 1x_{k+1}x_{k+2}\dots), & x \in C(x_0x_1\dots x_{k-1}x_k = 1\underbrace{0\dots0}_{k-1}1) \end{cases} \end{aligned}$$

It is immediately clear that ϕ intertwines the dynamics, so that it is also a dynamical isomorphism. Then $\phi \times \mathrm{id}$ is a dynamical isomorphism between R_B and $\tilde{R}_B$, since the translation by 1 in the $\mathbb{Z}_2$ component happens only when $x_0 = 1$ in A, which happens exactly once per symbol in Z. □

Lemma 3.27. *$\tilde{R}_B$ as above is ergodic, with respect to $\nu \times c$.*

Proof. Observe that $\tilde{R}_B$ is the product of the left shift L on Z and the translation by 1 on $\mathbb{Z}_2$. By Example 2.44, we see that L is strongly mixing. For the translation, it is easily seen that $\emptyset$ and $\mathbb{Z}_2$ are invariant sets, whereas $\{1\}$ and $\{0\}$ are not; this is a complete listing of all the subsets of $\mathbb{Z}_2$, so we clearly see that translation by 1 is ergodic. Applying Proposition 2.57 allows us to conclude that $\tilde{R}_B$ is ergodic. □

Proof of Proposition 3.25. Since $\tilde{R}_B$ is ergodic by Lemma 3.27, R_B must also be ergodic, because they are isomorphic by Lemma 3.26. Hence by Lemma 2.58 and Proposition 2.51, R is ergodic, so we are done. □

This handles one of the cases, as in the previous example. We now deal with the circle extension.

Proposition 3.28. *S as defined above is ergodic, with respect to $\mu \times \lambda$.*

The proof will follow along the same lines as the proof for Proposition 3.17. We shall induce, calculate an isomorphic map, square it, and find another isomorphic map which will be easier to prove to be ergodic.

Lemma 3.29. *The induced map S_B is measure-theoretically isomorphic to $\tilde{S}_B$ acting on $Z \times \mathbb{T}$, given by*

$$\tilde{S}_B(z,y) = \begin{cases}(L(z), k\alpha - y) & z \in C(z_0 = k).\end{cases}$$

Proof. The map $\phi \times \mathrm{id} : B \times \mathbb{T} \to Z \times \mathbb{T}$ is still a dynamical isomorphism, similarly to 3.26; the $\mathbb{T}$-component of the action arises from the flip given by the 1 in the base, followed by k rotations by α, one for each 0 encountered. □

The flip in the $\mathbb{T}$ component is still getting in the way, so we square the map to eliminate the flip. Let $P = (\tilde{S}_B)^2$. Unlike in the rotation case, this is much less cumbersome to describe:

$$\begin{aligned} P(z,y) &= \begin{cases}\tilde{S}_B(L(z), k\alpha - y) & z \in C(z_0 = k),\end{cases} \\ &= \begin{cases}(L^2(z), j\alpha - (k\alpha - y)) & z \in C(z_0 z_1 = kj),\end{cases} \\ &= \begin{cases}(L^2(z), y + (j-k)\alpha) & z \in C(z_0 z_1 = kj).\end{cases} \end{aligned}$$

So P acts by applying the left shift twice in the Z component, and then rotating around the circle by a multiple of α which depends on both of the symbols z_0, z_1. We change coordinates again, using the next lemma.

Lemma 3.30. *Let (Z, ν, L) be the left Bernoulli shift over a countable alphabet A with weights p_a for $a \in A$, and let (W, ρ, N) be the left Bernoulli shift over the alphabet $A \times A$, meaning that each symbol is a pair $(a_1 a_2) := (a_1, a_2) \in A \times A$, with weights given to each symbol equal to $p_{a_1 a_2} = p_{a_1} p_{a_2}$. Then (Z, ν, L^2) is dynamically isomorphic to the space (W, ρ, N).*

Proof. Define the map $\psi : Z \to W$ by

$$\psi(\dots z_{-2}z_{-1} \cdot z_0 z_1 \dots) = (\dots(z_{-2}z_{-1}) \cdot (z_0 z_1) \dots);$$

it is clear that ψ is invertible, with ψ^{-1} reversing the process. Moreover, ψ is measurable, since we have

$$\psi^{-1}(C(w_i = kj)) = \{z \in Z \ : \ z_{2i}z_{2i+1} = kj\} = C(z_{2i}z_{2i+1} = kj)$$

and every cylinder set in W is a finite intersection of these; ψ is also measure-preserving, because

$$\nu(\psi^{-1}(C(w_i = kj))) = \nu(C(z_{2i}z_{2i+1} = kj)) = p_k p_j = \rho(C(w_i = kj)).$$

We must also check if ψ^{-1} is measurable and measure-preserving. We have:

$$\begin{aligned}
(\psi^{-1})^{-1}(C(z_i = k)) &= \psi(C(z_i = k)) \\
&= \begin{cases} \left\{w \in w \ : \ w_{\lfloor \frac{i}{2} \rfloor} = kj, j \in A\right\} & i \text{ is even} \\ \left\{w \in w \ : \ w_{\lfloor \frac{i}{2} \rfloor} = jk, j \in A\right\} & i \text{ is odd} \end{cases} \\
&= \begin{cases} \bigcup_{j \in A} C(w_{\lfloor \frac{i}{2} \rfloor} = kj) & i \text{ is even} \\ \bigcup_{j \in A} C(w_{\lfloor \frac{i}{2} \rfloor} = jk) & i \text{ is odd} \end{cases},
\end{aligned}$$

so that ψ^{-1} is measurable, and we have:

$$\rho((\psi^{-1})^{-1}(C(z_i = k))) = \begin{cases} \rho\left(\bigcup_{j\in A} C(w_{\lfloor \frac{i}{2} \rfloor} = kj)\right) & i \text{ is even} \\ \rho\left(\bigcup_{j\in A} C(w_{\lfloor \frac{i}{2} \rfloor} = jk)\right) & i \text{ is odd} \end{cases}$$

$$= \sum_{j\in A} p_k p_j = p_k = \nu(C(z_i = k)),$$

since the sum of the weights of elements in A is equal to 1 and each of those sets are disjoint. Hence ψ is a measure space isomorphism. Finally, we may compute:

$$\begin{aligned} \psi \circ L^2(z) &= \psi \circ L^2(\ldots z_{-2}z_{-1} \cdot z_0 z_1 \ldots) \\ &= \psi(\ldots z_0 z_1 \cdot z_2 z_3 \ldots) \\ &= (\ldots (z_0 z_1) \cdot (z_2 z_3) \ldots) \\ &= N(\ldots (z_{-2}z_{-1}) \cdot (z_0 z_1) \ldots) \\ &= N \circ \psi(z). \end{aligned}$$

Therefore $\psi \circ L^2 = N \circ \psi$ and ψ is a dynamical isomorphism of the two spaces. □

From here, we have that $\psi \times \mathrm{id}$ is a dynamical isomorphism between P acting on $(Z \times \mathbb{T}, \nu \times \lambda)$ and $Q = (\psi \times \mathrm{id}) \circ P \circ (\psi^{-1} \times \mathrm{id})$ acting on $(W \times \mathbb{T}, \rho \times \lambda)$. Denote $Y : W \to \mathbb{T}$ as the function

$$Y(w) = Y(w_0) = \begin{cases} (j-k)\alpha & w_0 = kj; \end{cases}$$

it is easy to see that Y is measurable. Then we may write

$$Q(w, y) = (N(w), y + Y(w)).$$

Lemma 3.31. *The map Q as obtained above is strongly mixing with respect to $\rho \times \lambda$.*

As a corollary of this lemma, a result of Rudolph, in [23], shows that the map Q is actually more than strongly mixing; it is *Bernoulli*, that is, isomorphic to a Bernoulli shift over some alphabet.

To prove the lemma, it suffices to show that the mixing property holds for measurable rectangles in $W \times \mathbb{T}$ which are products of contiguous string cylinder sets $A, B \subset W$ with intervals $[a, b], [c, d] \subset \mathbb{T}$ (we may use closed intervals because closed intervals also generate the Borel σ-algebra on $\mathbb{T}$), since this collection of sets forms a generating π-system for the product σ-algebra. Namely, we are done if we show, for such sets,

$$\lim_{n\to\infty} \rho \times \lambda(Q^{-n}(A \times [a, b]) \cap (B \times [c, d])) = \rho(A)\rho(B)(b - a)(d - c).$$

After some preliminary lemmas, the proof will then proceed as follows: we shall rewrite the measure of the intersection as something fixed multiplied by an expectation of a $\mathbb{T}$-valued function with respect to a particular measure, and then show that the expectation converges to what we need.

Lemma 3.32. *For $m \in \mathbb{Z}$, we have $q_m = \rho(Y^{-1}\{m\alpha\}) = \dfrac{1}{3 \cdot 2^{|m|}}$.*

Proof. We have, for $m \geq 0$:

$$\begin{aligned}
q_m = \rho(Y^{-1}\{m\alpha\}) &= \rho(\{w \ : \ w_0 = kj, j - k = m\}) \\
&= \rho(\{w \ : \ w_0 = kj, j = m + k\}) \\
&= \rho(\{w \ : \ w_0 = k(m + k)\}) \\
&= \sum_{k=0}^{\infty} \frac{1}{2^{m+2k+2}} = \frac{1}{2^{m+2}} \sum_{k=0}^{\infty} \frac{1}{4^k} \\
&= \frac{1}{4 \cdot 2^m} \frac{1}{1 - \frac{3}{4}} = \frac{1}{4 \cdot 2^m} \frac{4}{4 - 1} = \frac{1}{3 \cdot 2^m}.
\end{aligned}$$

When $m < 0$, we have:

$$\begin{aligned} q_m &= \rho(Y^{-1}\{m\alpha\}) = \rho(\{w \ : \ w_0 = kj, j-k = m\}) \\ &= \rho(\{w \ : \ w_0 = kj, j-k = -|m|\}) \\ &= \rho(\{w \ : \ w_0 = kj, k-j = |m|\}) \\ &= q_{|m|} = \frac{1}{3 \cdot 2^{|m|}}, \end{aligned}$$

by the symmetry of the symbol space. □

In particular, all of the sets $Y^{-1}\{m\alpha\}$ are disjoint, and depend solely on w_0. We may thus partition the space using this coarser structure; note that each length one cylinder set $C(w_i = kj)$ lies inside $(Y \circ N^i)^{-1}\{(j-k)\alpha\}$. In particular, since these sets are disjoint, we may write $(Y \circ N^i)^{-1}\{m\alpha\}$ as a countable disjoint union of length one cylinder sets. The lemma also says that:

$$1 = \mu(X) = \mu\left(\bigcup_{k\in\mathbb{Z}} Y^{-1}\{k\alpha\}\right) = \sum_{k\in\mathbb{Z}} \mu(Y^{-1}\{k\alpha\}) = \sum_{k\in\mathbb{Z}} q_k,$$

which will be used later.

It turns out that Y is more important than just being a shorthand for the cocycle value. It will prove useful to know the distribution of Y, as well as a sum of multiple copies of Y (which we could see as arising from iterating the map Q). The distribution of Y is given by the measure η on $\mathbb{T}$, defined by

$$\eta(\{m\alpha\}) = \rho(Y^{-1}\{m\alpha\}) = q_m,$$

so that

$$\eta = \sum_{m\in\mathbb{Z}} q_m \delta_{\{m\alpha\}}.$$

We know that the distribution of a finite sum of n independent copies of Y is the n-fold convolution of η, which we denote by $\eta^{\otimes n}$ (see [21]).

We will use the following lemma, which has roots in Fourier analysis on groups (again, see [21]):

Lemma 3.33. $\eta^{\otimes n} \underset{n\to\infty}{\longrightarrow} \lambda$ *in the weak sense.*

Proof. The Fourier transform of η is $\hat{\eta}$, given by

$$\hat{\eta}(m) = \int_{\mathbb{T}} e^{2\pi imx}\ d\eta(x) = \sum_{k\in\mathbb{Z}} q_k e^{2\pi imk\alpha}.$$

We know that the Fourier transform of a convolution becomes the product of Fourier transforms, so that

$$\widehat{\eta^{\otimes n}}(m) = \prod_{j=1}^{n} \hat{\eta}(m) = \left(\sum_{k\in\mathbb{Z}} q_k e^{2\pi imk\alpha}\right)^n.$$

Recalling that $\mathbb{C}$ is a strictly convex space (that is, any line through two points on the unit circle intersects the unit circle only at those points), for any two points $x \neq y$ on the unit circle and $\lambda \in (0,1)$ we have that

$$|\lambda x + (1-\lambda)y| < \lambda\,|x| + (1-\lambda)\,|y| \leq \lambda + 1 - \lambda = 1.$$

This implies that, when $m \neq 0$:

$$\begin{aligned}\left|\sum_{k\in\mathbb{Z}} q_k e^{2\pi imk\alpha}\right| &= \left|\frac{1}{3\cdot 2}e^{2\pi im\alpha} + \frac{1}{3\cdot 2}e^{-2\pi im\alpha} + \sum_{k\neq\pm 1} q_k e^{2\pi imk\alpha}\right| \\ &\leq \frac{1}{3}\left|\frac{1}{2}e^{2\pi im\alpha} + \frac{1}{2}e^{-2\pi im\alpha}\right| + \left|\sum_{k\neq\pm 1} q_k e^{2\pi imk\alpha}\right| \\ &< \frac{1}{3}\cdot 1 + \sum_{k\neq\pm 1} q_k \cdot 1 \\ &= \frac{1}{3} + 1 - \frac{1}{3\cdot 2} - \frac{1}{3\cdot 2} = 1.\end{aligned}$$

Thus we see that

$$\widehat{\eta^{\otimes n}}(m) \underset{n\to\infty}{\longrightarrow} 0$$

for $m \neq 0$, and

$$\widehat{\eta^{\otimes n}}(0) = \left(\sum_{k\in\mathbb{Z}} q_k\right)^n = (1)^n = 1 \underset{n\to\infty}{\longrightarrow} 1.$$

The Fourier transform of the Lebesgue measure is, using periodicity and symmetry of $x \mapsto e^{2\pi imx}$,

$$\hat{\lambda}(m) = \begin{cases} 1 & m = 0, \\ 0 & m \neq 0. \end{cases}$$

Thus $\widehat{\eta^{\otimes n}} \underset{n\to\infty}{\longrightarrow} \hat{\lambda}$ pointwise; uniqueness of the Fourier transform forces $\eta^{\otimes n} \underset{n\to\infty}{\longrightarrow} \lambda$ weakly, as desired (see [13, Theorem I.7.2]). □

In particular, this lemma means that for any continuous function f on $\mathbb{T}$, we have

$$\int_{\mathbb{T}} f \; d\eta^{\otimes n} \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{T}} f \; d\lambda.$$

This will be used to prove the next lemma, which is similar to the discussion in the first chapter of [17].

Lemma 3.34. *For $a < b$ in $\mathbb{T}$,*

$$\int_{\mathbb{T}} \mathbb{1}_{[a,b]}\ d\eta^{\otimes n} \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{T}} \mathbb{1}_{[a,b]}\ d\lambda = b - a.$$

If $a > b$, then the result is still true, by splitting the interval up into $[0, b]$ and $[a, 1]$.

Proof. Let $\epsilon > 0$. Let $f, g : \mathbb{T} \to \mathbb{R}$ be given by:

$$f(x) = \begin{cases} 1 & x \in [a, b] \\ \frac{3}{\epsilon}(x - a) + 1 & x \in \left(a - \frac{\epsilon}{3}, a\right) \\ \frac{-3}{\epsilon}\left(x - b - \frac{\epsilon}{3}\right) & x \in \left(b, b + \frac{\epsilon}{3}\right) \\ 0 & \text{otherwise,} \end{cases}$$

$$g(x) = \begin{cases} 1 & x \in \left[a + \frac{\epsilon}{3}, b - \frac{\epsilon}{3}\right] \\ \frac{3}{\epsilon}\left(x - a - \frac{\epsilon}{3}\right) + 1 & x \in \left(a, a + \frac{\epsilon}{3}\right) \\ \frac{-3}{\epsilon}(x - b) & x \in \left(b - \frac{\epsilon}{3}, b\right) \\ 0 & \text{otherwise.} \end{cases}$$

Both f and g are continuous, and satisfy $0 \le g \le \mathbb{1}_{[a,b]} \le f \le 1$; observe further that

$$\int_{\mathbb{T}} f - \mathbb{1}_{[a,b]}\ d\lambda = \int_{\mathbb{T}} \mathbb{1}_{[a,b]} - g\ d\lambda = 2 \cdot \frac{\epsilon}{6} = \frac{\epsilon}{3}.$$

Finally, by Lemma 3.33, we have:

$$\begin{aligned} \liminf_{n\to\infty} \int_{\mathbb{T}} g\ d\eta^{\otimes n} &= \lim_{n\to\infty} \int_{\mathbb{T}} g\ d\eta^{\otimes n} = \int_{\mathbb{T}} g\ d\lambda, \\ \limsup_{n\to\infty} \int_{\mathbb{T}} f\ d\eta^{\otimes n} &= \lim_{n\to\infty} \int_{\mathbb{T}} f\ d\eta^{\otimes n} = \int_{\mathbb{T}} f\ d\lambda. \end{aligned}$$

For each $n \in \mathbb{N}_+$, denote

$$I_n = \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\eta^{\otimes n} - \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda.$$

Then we have, using the equalities and inequalities involving f,g, and $\mathbb{1}_{[a,b]}$ above:

$$\liminf_{n\to\infty} I_n \geq \liminf_{n\to\infty} \left(\int_{\mathbb{T}} g \ d\eta^{\otimes n} - \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda \right) = \int_{\mathbb{T}} g - \mathbb{1}_{[a,b]} \ d\lambda = \frac{-\epsilon}{3},$$
$$\limsup_{n\to\infty} I_n \leq \limsup_{n\to\infty} \left(\int_{\mathbb{T}} f \ d\eta^{\otimes n} - \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda \right) = \int_{\mathbb{T}} f - \mathbb{1}_{[a,b]} \ d\lambda = \frac{\epsilon}{3}.$$

Combining these inequalities, we get:

$$0 \leq \limsup_{n\to\infty} I_n - \liminf_{n\to\infty} I_n \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon.$$

Since ϵ was arbitrary, we conclude that $\limsup_{n\to\infty} I_n = \liminf_{n\to\infty} I_n$, and so $\lim_{n\to\infty} I_n$ exists. To calculate the limit, observe that we also have:

$$\begin{aligned} \frac{-\epsilon}{3} &= \int_{\mathbb{T}} g - \mathbb{1}_{[a,b]} \ d\lambda \\ &\leq \lim_{n\to\infty} I_n \\ &\leq \int_{\mathbb{T}} f - \mathbb{1}_{[a,b]} \ d\lambda = \frac{-\epsilon}{3}, \end{aligned}$$

so that $\lim_{n\to\infty} I_n = 0$, since ϵ was arbitrary. Finally, we have:

$$\begin{aligned} \lim_{n\to\infty} \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\eta^{\otimes n} &= \lim_{n\to\infty} \left(\int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\eta^{\otimes n} - \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda + \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda \right) \\ &= \left(\lim_{n\to\infty} I_n \right) + \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda \\ &= 0 + \int_{\mathbb{T}} \mathbb{1}_{[a,b]} \ d\lambda = b - a, \end{aligned}$$

as desired. ☐

For notational purposes, define $X_n : W \to \mathbb{T}$ by

$$X_n(w) = \sum_{i=0}^{n-1} Y(N^i(w)),$$

so that $Q^n(w, y) = (N^n(w), y + X_n(w))$. In particular, the definition of Y says that X_n depends on $w_0, \ldots, w_{n-1}$, and that X_n is measurable. As well, we see that

$$\begin{aligned} X_n^{-1}\{m\alpha\} &= \left\{ w \ : \ w_j \in (Y \circ N^j)^{-1}\{m_j\alpha\}, j = 0 \ldots n-1, \text{ where} \sum_{j=0}^{n-1} m_j = m \right\} \\ &= \bigcup_{\sum_{j=0}^{n-1} m_j = m} \bigcap_{j=0}^{n-1} (Y \circ N^j)^{-1}\{m_j\alpha\}. \end{aligned}$$

Since each of the sets $(Y \circ N^j)^{-1}\{m_j\alpha\}$ may be written as a countable disjoint union of length one cylinder sets, each fixing a symbol, we see that a finite intersection of n of them is a countable disjoint union of length n contiguous string cylinder sets, so that the entire set is a countable disjoint union of such sets. In particular, we also have:

$$\begin{aligned} &X_n^{-1}\{m\alpha\} \cap C(w_0 \ldots w_{n-1} = (k_0 j_0) \ldots (k_{n-1} j_{n-1})) \\ &\qquad = \begin{cases} \emptyset & \sum_{l=0}^{n-1}(j_l - k_l) \neq m, \\ C(w_0 \ldots w_{n-1} = (k_0 j_0) \ldots (k_{n-1} j_{n-1})) & \text{else.} \end{cases} \end{aligned}$$

We shall now directly work towards showing that Q is strongly mixing. By Lemma 2.12, part of Example 2.7, and Proposition 2.39, it suffices to show the mixing property for Q for products of a contiguous string cylinder set and a closed interval. This will be done in two steps; first we show that the measure of the given set is equal to an expectation, and then we show that the expectation converges in the appropriate way.

Lemma 3.35. *Let $t_1, t_2 \geq 0$ and $r_1, r_2 \geq 1$. Let A be a contiguous string cylinder set starting at the index $-t_1$ and ending at the index $r_1 - 1$, and let B be a contiguous string cylinder set starting at the index $-t_2$ and ending at the index $r_2 - 1$. Let $a, b, c, d \in \mathbb{T}$, with $a < b$ and $c < d$ (otherwise, split the intervals up into ones where this assumption holds). Then we have that, for $k \geq t_1 + r_2 = k_*$:*

$$\begin{aligned}\rho \times \lambda(Q^{-k}(A \times [a,b]) \cap (B \cap [c,d]))&\\ = \rho(A)\rho(B)\ \mathbb{E}_{\eta^{\otimes k-k^*}}&\left[\lambda\left(\left[\tilde{a} - \sum_{l=0}^{k-k^*-1} Y, \tilde{b} - \sum_{l=0}^{k-k^*-1} Y\right] \cap [c,d]\right)\right],\end{aligned}$$

where $\tilde{a} = a - C$ and $\tilde{b} = b - C$ are translates of a and b by the same fixed constant, independent of k.

Proof. First, we shall compute the set $Q^{-t_1}(A \times [a,b])$, using the computation with X_n above. Observe that $N^{-t_1}(A)$ is a contiguous string cylinder set from 0 to $t_1 + r_1 - 1$, and that for any $m \in \mathbb{Z}$, $X_{t_1}^{-1}\{m\alpha\}$ is a disjoint union of contiguous string cylinder sets from 0 to $t_1 - 1 < t_1 + r_1 - 1$. This means that X_{t_1} is constant on $N^{-t_1}(A)$, with value $m_A\alpha$. Thus we have:

$$N^{-t_1}(A) \cap X_{t_1}^{-1}\{m\alpha\} = \begin{cases} N^{-t_1}(A) & m = m_A\alpha, \\ \emptyset & \text{otherwise.} \end{cases}$$

We have:

$$\begin{aligned}Q^{-t_1}(A \times [a,b]) &= \bigcup_{m\in\mathbb{Z}} (X_{-t_1}^{-1}\{m\alpha\} \cap N^{-t_1}(A)) \times [a - m\alpha, b - m\alpha] \\ &= N^{-t_1}(A) \times [a - m_A\alpha, b - m_A\alpha].\end{aligned}$$

We wish to take further preimages of $A \times [a,b]$ under Q, and intersect them with

$B \times [c,d]$. Since B fixes symbols from $-t_2$ to $r_2 - 1$, X_{r_2} is going to be constant, with value $m_B\alpha$, by the same argument as before. Thus we have:

$$\begin{aligned}
&Q^{-r_2}(Q^{-t_1}(A \times [a,b])) \cap (B \times [c,d]) \\
&\quad = Q^{-r_2}(N^{-t_1}(A) \times [a - m_A\alpha, b - m_A\alpha]) \cap (B \times [c,d]) \\
&\quad = \bigcup_{m \in \mathbb{Z}} \Big((X_{-r_2}^{-1}\{m\alpha\} \cap N^{-t_1-r_2}(A) \\
&\qquad\qquad \times [a - (m_A + m)\alpha, b - (m_A + m)\alpha]) \cap (B \times [c,d]) \Big) \\
&\quad = (N^{-(t_1+r_2)}(A) \cap B) \times ([a - (m_A + m_B)\alpha, b - (m_A + m_B)\alpha] \cap [c,d]).
\end{aligned}$$

Let $C = (m_A + m_B)\alpha$ and $k^* = t_1 + r_2$. Then, for any $k > k^*$, X_k depends on the symbols with indices $0, \ldots, k-1$, but only the symbols with indices $r_2, \ldots, k - t_1 - 1$ are not fixed by either A or B. Hence, we obtain:

$$\begin{aligned}
&Q^{-k}(A \times [a,b]) \cap (B \times [c,d]) \\
&\quad = Q^{-(k-t_1)}(N^{-t_1}(A) \times [a - m_A\alpha, b - m_A\alpha]) \cap (B \times [c,d]) \\
&\quad = \bigcup_{\sum_{l=r_2}^{k-t_1-1} m_l = m \in \mathbb{Z}} \left(\left(\bigcap_{l=r_2}^{k-t_1-1} (Y \circ N^l)^{-1}\{m_l\alpha\} \cap N^{-k}(A) \cap B \right) \right. \\
&\qquad\qquad \left. \times [a - C - m\alpha, b - C - m\alpha] \cap [c,d] \right).
\end{aligned}$$

This is a disjoint union of measurable rectangles of $W \times \mathbb{T}$, and the W-component of each rectangle is the intersection of three sets in W which fix symbols on disjoint indices. This means we may certainly compute the measure of $Q^{-k}(A \times [a,b]) \cap (B \times$

$[c,d])$. Let $\tilde{a} = a - C, \tilde{b} = b - C$; we then obtain:

$$
\begin{aligned}
&\rho \times \lambda(Q^{-k}(A \times [a,b]) \cap (B \times [c,d])) \\
&\quad = \sum_{\sum_{l=r_2}^{k-t_1-1} m_l = m \in \mathbb{Z}} \left(\rho\left(\bigcap_{l=r_2}^{k-t_1-1} (Y \circ N^l)^{-1}\{m_l\alpha\} \cap N^{-k}(A) \cap B \right) \right. \\
&\qquad\qquad \left. \cdot \lambda\big([\tilde{a} - m\alpha, \tilde{b} - m\alpha] \cap [c,d]\big) \right) \\
&\quad = \sum_{\sum_{l=r_2}^{k-t_1-1} m_l = m \in \mathbb{Z}} \left(\left(\prod_{l=r_2}^{k-t_1-1} q_{m_l} \right) \rho(N^{-k}(A))\rho(B)\lambda\big([\tilde{a} - m\alpha, \tilde{b} - m\alpha] \cap [c,d]\big) \right) \\
&\quad = \rho(A)\rho(B) \sum_{\sum_{l=r_2}^{k-t_1-1} m_l = m \in \mathbb{Z}} \left(\prod_{l=r_2}^{k-t_1-1} q_{m_l} \right) \lambda\big([\tilde{a} - m\alpha, \tilde{b} - m\alpha] \cap [c,d]\big)
\end{aligned}
$$

The last observation to make is that because (W, ρ, N) is a Bernoulli shift, when symbols are not fixed, they are independent of each other. That means the choice of the m_l in the large sum is an *independent* choice. In particular, we know that the values of $Y \circ N^i$ and $Y \circ N^j$ are independent, if there is no other knowledge. Then, we note that

$$
m\alpha = \sum_{l=r_2}^{k-t_1-1} Y \circ N^l(w), \;\; w \in \bigcap_{l=r_2}^{k-t_1-1} (Y \circ N^l)^{-1}\{m_l\alpha\},
$$

and see that indeed, because none of the symbols between r_2 and $k - t_1 - 1$ are fixed by A or B, $\sum_{l=r_2}^{k-t_1-1} Y \circ N^l$ is a sum of independent random variables, each having the same distribution as Y. There are $k - t_1 - 1 - r_2 + 1 = k - k^*$ copies of Y in the sum, so that the sum of the $k - k^*$ copies of Y has distribution $\eta^{\otimes k - k^*}$, the $(k - k^*)$-fold convolution of η. Hence we may write, finally:

$$\begin{aligned}
&\rho \times \lambda(Q^{-k}(A \times [a,b]) \cap (B \times [c,d])) \\
&\quad = \rho(A)\rho(B) \sum_{\sum_{l=r_2}^{k-t_1-1} m_l = m \in \mathbb{Z}} \left(\prod_{l=r_2}^{k-t_1-1} q_{m_l} \right) \lambda\big([\tilde{a} - m\alpha, \tilde{b} - m\alpha] \cap [c,d]\big) \\
&\quad = \rho(A)\rho(B)\ \mathbb{E}_{\eta^{\otimes k-k^*}} \left[\lambda \left(\left[\tilde{a} - \sum_{l=0}^{k-k^*-1} Y, \tilde{b} - \sum_{l=0}^{k-k^*-1} Y \right] \cap [c,d] \right) \right],
\end{aligned}$$

where we have reindexed the sum of copies of Y to start at $l = 0$. We are done with the lemma. □

The limit of the expectation as k goes to infinity is the same as the limit as $k - k^*$ goes to infinity, so we may compute the limit as k goes to infinity for notational simplicity.

Lemma 3.36. *Let $a < b$ and $c < d$ in $\mathbb{T}$ and let $\tilde{a}, \tilde{b}$ be translations of a and b by the same constant. Let Y be defined as above. Then we have:*

$$\lim_{k \to \infty} \mathbb{E}_{\eta^{\otimes k}} \left[\lambda \left(\left[\tilde{a} - \sum_{l=0}^{k-1} Y, \tilde{b} - \sum_{l=0}^{k-1} Y \right] \cap [c,d] \right) \right] = (b-a)(d-c).$$

Proof. First, we rewrite the expectation as an integral involving $\eta^{\otimes k}$:

$$\begin{aligned}
&\mathbb{E}_{\eta^{\otimes k}} \left[\lambda \left(\left[\tilde{a} - \sum_{l=0}^{k-1} Y, \tilde{b} - \sum_{l=0}^{k-1} Y \right] \cap [c,d] \right) \right] \\
&\quad = \int_{\mathbb{T}} \lambda([\tilde{a} - x, \tilde{b} - x] \cap [c,d])\ d\eta^{\otimes k}(x) \\
&\quad = \int_{\mathbb{T}} \int_{\mathbb{T}} \mathbb{1}_{[\tilde{a}-x, \tilde{b}-x]}(y) \mathbb{1}_{[c,d]}(y) d\lambda(y) d\eta^{\otimes k}(x) \\
\text{(Fubini)} &\quad = \int_{\mathbb{T}} \mathbb{1}_{[c,d]}(y) \int_{\mathbb{T}} \mathbb{1}_{[\tilde{a},\tilde{b}]}(y+x) d\eta^{\otimes k}(x) d\lambda(y) \\
&\quad = \int_{\mathbb{T}} \mathbb{1}_{[c,d]}(y) \int_{\mathbb{T}} \mathbb{1}_{[\tilde{a}-y, \tilde{b}-y]}(x) d\eta^{\otimes k}(x) d\lambda(y).
\end{aligned}$$

By Lemma 3.34, the inside integral converges, for any $y \in \mathbb{T}$:

$$\int_{\mathbb{T}} \mathbb{1}_{[\tilde{a}-y,\tilde{b}-y]}(x) d\eta^{\otimes k}(x) \underset{k\to\infty}{\longrightarrow} \int_{\mathbb{T}} \mathbb{1}_{[\tilde{a}-y,\tilde{b}-y]}(x)\ d\lambda(x)$$
$$= \tilde{b} - y - (\tilde{a} - y) = \tilde{b} - \tilde{a} = b - m\alpha - (a - m\alpha) = b - a,$$

regardless of y. Each of the integrals is bounded above by 1 and is a measurable function of y (by Fubini), so we may use the Lebesgue Dominated Convergence Theorem to conclude:

$$\mathbb{E}_{\eta^{\otimes k}}\left[\lambda\left(\left[\tilde{a} - \sum_{l=0}^{k-1} Y, \tilde{b} - \sum_{l=0}^{k-1} Y\right] \cap [c,d]\right)\right]$$
$$\underset{n\to\infty}{\longrightarrow} \int_{\mathbb{T}} \mathbb{1}_{[c,d]}(y)(b-a)d\lambda(y) = (b-a)(d-c).$$

This proves the lemma. □

Lemma 3.37. *Let $A \times [a,b]$ and $B \times [c,d]$ be as in the hypotheses of Lemma 3.35. Then the mixing property for Q holds for A and B.*

Proof. Apply Lemmas 3.35 and 3.36. □

Proof of Lemma 3.31. Thanks to Lemma 3.37, we see that Q is mixing for all sets in a generating π-system for the shift on W. Hence by Proposition 2.39, we see that Q is strongly mixing. □

Proof of Proposition 3.28. Since Q is strongly mixing, it is also ergodic by Proposition 2.40. This shows that P is ergodic, by the isomorphism, so that $\tilde{S}_B$ is ergodic by Lemma 3.19. Hence S is ergodic by Proposition 2.51 using Lemma 2.58. □

Proof of Theorem 3.24. Now that we know that R and S are ergodic on their respective spaces by Propositions 3.17 and 3.28, we see that A is not diagonalizable over $\mathbb{C}$ by 3.3, as desired. Hence we are finished the proof. □

Chapter 4

Conclusion

We have seen that not every real matrix cocycle is block upper-triangularizable, even over the complex numbers. To do this, we looked no further than cocycles of 2-by-2 real orthogonal matrices, and found our counterexamples there; the particular structure of orthogonal matrices in two dimensions allowed us to achieve concrete criteria for existence of a counterexample. From here, there are a few directions in which to proceed.

One option is to consider non-orthogonal matrices, even just in two dimensions. Is there a theorem or a class of counterexamples lurking in this setting? We could also increase the dimension of the matrices, while keeping them orthogonal, and see what we can find, or jump straight to arbitrary invertible n-by-n matrices. Finally, an option is to ask if is a generic property of matrix cocycles (if wanted, just 2-by-2 orthogonal matrices) to be non-block-upper-triangularizable, where we mean generic in the technical sense: does the set of such cocycles contain a dense G_δ set in an appropriate topology, so that 'most' cocycles have this property?

As per usual, finding the answer to one question opens the door to many other questions. It is our responsibility as mathematicians to not only ask those questions, but see where they take us.

Chapter A

Appendix

A.1 The Grassmannian

We provide a proof of Proposition 2.18.

Proof. $\{v_1, v_2\}$ is easily seen to be an orthogonal set with respect to the standard complex inner product, hence is also linearly independent and therefore is a basis for $\mathbb{C}^2$. If V is a 1-dimensional subspace of $\mathbb{C}^2$, then for some $a, b \in \mathbb{C}$, we have

$$V = \text{span}_{\mathbb{C}}\{av_1 + bv_2\} = \text{span}_{\mathbb{C}}\{v_1 + \frac{b}{a}v_2\},$$

when $a \neq 0$, and if $a = 0$, then $V = \text{span}_{\mathbb{C}}\{v_2\}$. It is clear that if $a \neq 0$, then

$$\text{span}_{\mathbb{C}}\{av_1 + bv_2\} \neq \text{span}_{\mathbb{C}}\{v_2\},$$

and if we have $z, w \in \mathbb{C}$ with

$$\text{span}_{\mathbb{C}}\{v_1 + zv_2\} = \text{span}_{\mathbb{C}}\{v_1 + wv_2\},$$

then there is $c \in \mathbb{C}$ such that

$$v_1 + zv_2 = c(v_1 + wv_2) = cv_1 + cwv_2,$$

which forces $c = 1$ and $z = w$, since $\{v_1, v_2\}$ is a basis. Hence, ψ is injective. If $z \in \bar{\mathbb{C}} \setminus \{\infty\}$, then $\psi(\text{span}_{\mathbb{C}}\{v_1 + zv_2\}) = z$, and $\psi(\text{span}_{\mathbb{C}}\{v_2\}) = \infty$, so ψ is surjective. Thus ψ is a bijection.

Let $V_n \underset{n\to\infty}{\longrightarrow} V$ in $\text{Gr}_1(\mathbb{C}^2)$, where $\psi(V_n) = z_n$ and $\psi(V) = z \in \bar{\mathbb{C}} \setminus \{\infty\}$. Then unit vectors sitting inside V_n and V are

$$\frac{v_1 + (z_n)v_2}{\sqrt{2}(1 + |z_n|^2)^{\frac{1}{2}}}, \quad \frac{v_1 + zv_2}{\sqrt{2}(1 + |z|^2)^{\frac{1}{2}}}.$$

If the subspaces converge, then these vectors must get arbitrarily close also, by definition of the metric on the Grassmannian (we may choose to fix angles so that the coefficient of v_1 is always real). But then we have

$$\left| \frac{(1 + |z_n|^2)^{\frac{1}{2}} z_n}{(1 + |z_n|^2)^{\frac{1}{2}}} - \frac{(1 + |z|^2)^{\frac{1}{2}} z}{(1 + |z|^2)^{\frac{1}{2}}} \right| = |z_n - z|,$$

and we know the former gets small. Hence indeed, z_n converges to z.

To deal with the case that $\psi(V) = \infty$, we suppose that the vector

$$\frac{v_1 + (z_n)v_2}{\sqrt{2}(1 + |z_n|^2)^{\frac{1}{2}}}$$

approaches the vector $\dfrac{v_2}{\sqrt{2}}$. Then

$$\frac{1}{\sqrt{2}(1 + |z_n|^2)^{\frac{1}{2}}} \to 0,$$

which can only happen if $|z_n|$ gets arbitrarily large. Thus $z_n \underset{n\to\infty}{\longrightarrow} \infty$, and we've shown that ψ is continuous.

Finally, $\mathrm{Gr}_1(\mathbb{C}^2)$ is compact and $\bar{C}$ is Hausdorff, so since ψ is a continuous bijection, we obtain the fact that ϕ^{-1} is continuous for free. We are done. □

A.2 Ergodic equivalences

Here, we present the proof of Theorem 2.37. First we need a quick lemma.

Lemma A.1. *Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $A, B, C \in \mathcal{B}$. Then we have that μ satisfies a 'triangle inequality' (see [11]):*

$$\mu(A\Delta B) \leq \mu(A\Delta C) + \mu(C\Delta B).$$

Proof. Note that for any set C,

$$A\Delta B \subset (A\Delta C) \cup (C\Delta B),$$

by inspecting an Euler diagram, so because μ is subadditive, we obtain

$$\mu(A\Delta B) \leq \mu(A\Delta C) + \mu(C\Delta B).$$

□

We also recall the definition of a λ-system, together with a useful theorem involving it.

Definition A.2. Let X be a set. A collection of subsets $\mathcal{L}$ is a λ-*system* for X if we have:

1. $X \in \mathcal{L}$,

2. If $A \in \mathcal{L}$, then $X \setminus A = A^c \in \mathcal{L}$,

3. If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{L}$ with $A_i \cap A_j = \emptyset$ for $i \neq j$, then

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{L}.$$

Theorem A.3. *Let X be a set, let $\mathcal{P}$ be a π-system in X and let $\mathcal{L}$ be a λ-system in X containing $\mathcal{P}$. Then $\sigma(\mathcal{P}) \subset \mathcal{L}$, that is, the σ-algebra generated by $\mathcal{P}$ is contained inside of $\mathcal{L}$.*

Proof of Theorem 2.37. ($1 \implies 2$): (due to Sarig: [25]) Let A be almost T-invariant. We shall find a set $B \in \mathcal{B}$ such that B is strictly T-invariant and has the same measure as A. Define the set B by:

$$\begin{aligned} B &= \big\{x \in X \ : \ T^k(x) \text{ returns to A infinitely often } \big\} \\ &= \big\{x \in X \ : \ \forall n \geq 1, \exists k \geq n \text{ such that } T^k(x) \in A\big\} \\ &= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} T^{-k}(A). \end{aligned}$$

B is thus measurable, by the last equality. We also have:

$$T^{-1}(B) = \big\{x \in X \ : \ \forall n \geq 1, \exists k \geq 1 \text{ such that } T^k(T(x)) = T^{k+1}(x) \in A\big\} = B,$$

so indeed B is T-invariant, and hence either $\mu(B) = 0$ or $\mu(X \setminus B) = 0$. To show that B has the same measure as A, observe that if $x \in B \setminus A$, then there is k such that $x \in T^{-k}(A) \setminus A$, and if $x \in A \setminus B$, then x returns to A finitely many times and so there is k such that $x \in A \setminus T^{-k}(A)$, so we obtain:

$$A \Delta B \subset \bigcup_{k=1}^{\infty} A \Delta T^{-k}(A).$$

Using this and the symmetric difference triangle inequality, we get:

$$\begin{aligned}
\mu(A\Delta B) &\leq \mu\left(\bigcup_{k=1}^{\infty} A\Delta T^{-k}(A)\right) \leq \sum_{k=1}^{\infty} \mu(A\Delta T^{-k}(A)) \\
&\leq \sum_{k=1}^{\infty}\sum_{i=1}^{k} \mu(T^{-i}(A)\Delta T^{-(i-1)}(A)) = \sum_{k=1}^{\infty}\sum_{i=1}^{k} \mu(T^{-(i-1)}((A)\Delta T^{-1}(A))) \\
&= \sum_{k=1}^{\infty}\sum_{i=1}^{k} \mu(A\Delta T^{-1}(A)) = \sum_{k=1}^{\infty} k\mu(A\Delta T^{-1}(A)) = \sum_{k=1}^{\infty} 0 = 0,
\end{aligned}$$

utilizing the fact that A is almost T-invariant. Therefore $\mu(A\Delta B) = 0$, which yields $\mu(A \setminus B) = \mu(B \setminus A) = 0$. Finally, we compute:

$$\mu(A) = \mu(A \setminus B) + \mu(A \cap B) = \mu(A \cap B) = \mu(B \cap A) + \mu(B \setminus A) = \mu(B),$$

and so either $\mu(A) = 0$ or $\mu(X \setminus A) = 0$, as desired.

(2 $\implies$ 1): Let $A = T^{-1}(A)$. Then

$$\mu(A\Delta T^{-1}(A)) = \mu(A\Delta A) = \mu(\emptyset) = 0,$$

and so $\mu(A) = 0$ or $\mu(X \setminus A) = 0$. Hence (μ, T) is ergodic.

(2 $\implies$ 3): Let f be measurable, with $f \circ T = f$ almost everywhere. Then we have, for any $c \in \mathbb{R}$:

$$T^{-1}f^{-1}(-\infty, c) = (f \circ T)^{-1}(-\infty, c) = f^{-1}(-\infty, c),$$

up to a set of measure zero. Then $\mu(f^{-1}(-\infty, c))$ has either zero measure or full measure, ie. either $f(x) \geq c$ or $f(x) < c$ almost everywhere. Let

$$s = \sup\left\{c \ : \ \mu(f^{-1}(-\infty, c)) = 0\right\}$$

. Almost everywhere, we have that for any $c < s$, $f(x) \geq c$, and for any $c > s$, $f(x) < c$. Letting $a_n \underset{n\to\infty}{\longrightarrow} s$ monotonically from below and $b_n \underset{n\to\infty}{\longrightarrow} s$ monotonically from above, we have that:

$$\begin{aligned}
\mu(X \setminus f^{-1}\{s\}) &= \mu(f^{-1}(\mathbb{R}) \setminus f^{-1}(\bigcap_{n=1}^{\infty}[a_n, b_n))) \\
&= \mu(f^{-1}(\bigcup_{n=1}^{\infty}(-\infty, a_n) \cup [b_n, \infty))) \\
&\leq \sum_{n=1}^{\infty}(\mu(f^{-1}(-\infty, a_n)) + \mu(f^{-1}[b_n, \infty)) \\
&= \sum_{n=1}^{\infty} 0 = 0.
\end{aligned}$$

Thus $f^{-1}\{s\}$ has full measure, so that $f(x) = s$ almost everywhere.

($3 \implies 2$): Let A be almost T-invariant. Then $\mathbb{1}_A(x) = \mathbb{1}_{T^{-1}(A)}(x) = \mathbb{1}_A(T(x))$ for almost every x, and characteristic functions of measurable sets are measurable functions, thus $\mathbb{1}_A$ is constant almost everywhere. Hence $\mu(A) = 0$ or $\mu(X \setminus A) = 0$.

Note that we may interchange (3) with the following statement (3m): *If f is measurable, and $f \circ T = f$ almost everywhere, then f is constant almost everywhere.*

($3m \implies 3$): Since every L^∞ function is measurable, the implication clearly holds.

($3 \implies 3m$): Let $g : \mathbb{R} \to \mathbb{R}$ be given by $g(x) = \arctan(x)$; g is continuous hence measurable. Then if $f : \mathbb{R} \to \mathbb{R}$ is measurable and $f \circ T = f$ almost everywhere, $g \circ f$ is measurable and bounded, hence in L^∞. We also have that $g \circ f \circ T = g \circ f$

almost everywhere, thus $g \circ f$ is constant almost everywhere. But g is a bijection between $\mathbb{R}$ and $(-\frac{\pi}{2}, \frac{\pi}{2})$, with $g^{-1}(x) = \tan(x)$, so applying g^{-1} to both sides, we see that $f \circ T = f$ almost everywhere, as desired.

We now specify that $\mu(X) = 1$. Note that for any $A \in \mathcal{B}$, we have:

$$\int_X \mathbb{1}_A^2 \, d\mu = \int_X \mathbb{1}_A \, d\mu = \mu(A) \leq \mu(X) = 1 < \infty,$$

so that $\mathbb{1}_A \in L^2(X)$.

$(4 \implies 1)$: Let $A = T^{-1}(A)$. Then we have $\mathbb{1}_A = \mathbb{1}_A \circ T$ as before, hence constant almost everywhere, so $\mu(A) \in \{0, 1\}$.

$(3 \implies 4)$: Let $f \in L^2(X)$ be T-invariant. Define $f_n : X \to \mathbb{R}$ by the following:

$$f_n(x) = \begin{cases} f(x) & |f(x)| < n, \\ n & f(x) \geq n, \\ -n & f(x) \leq -n. \end{cases}$$

Then f_n is bounded and measurable, hence in L^∞. Finally, since

$$T^{-1} f^{-1}[-n, n] = (f \circ T)^{-1}[-n, n] = f^{-1}[-n, n],$$

we easily see that $f_n(T(x)) = f_n(x)$ for almost every x, so f_n is almost everywhere constant with constant C_n, for each n. We may compute the norm of each f_n, to be

$$\|f_n\| = (\int_X f_n^2 \, d\mu)^{\frac{1}{2}} = |C_n| \, .$$

Now, we see that $|f_1| \leq |f_2| \leq \cdots \leq |f|$, which implies $\|f_1\| \leq \|f_2\| \leq \cdots \leq \|f\| < \infty$. Then the sequence $\{|C_n|\}$ is bounded above and monotonically increasing, hence it converges to a limit $|C|$; moreover, if $|C_n| < |C_{n+1}|$, then by definition of f_n we must have $|C_{n+1}| = |C_n| + 1$, because if f_n doesn't take the values n or $-n$, f_{n+1} will equal f_n. Hence eventually, $|C_n| = |C_{n+1}| = \cdots = |C|$, and because the $\operatorname{sgn}(f_i) = \operatorname{sgn}(f_j)$ for any $i \neq j$, we get $f_n = f_{n+1} = \cdots = C$ almost everywhere. Then since f is the pointwise limit of the sequence of f_n, we get

$$f(x) = \lim_{n\to\infty} f_n(x) = C,$$

almost everywhere, as desired.

($1 \implies 5$): Let $A, B \in \mathcal{P} \subset \mathcal{B}$. Since $\mu(X) = 1$, $\mathbb{1}_A$ and $\mathbb{1}_B$ are both integrable. Since (μ, T) is ergodic, we see that for almost every $x \in X$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_A(T^k(x)) = \int_X \mathbb{1}_A \, d\mu = \mu(A),$$

by the corollary to Birkhoff's theorem (which we have not yet used in the proof, so there is no circular argument). Then we are done, by the Lebesgue Dominated Convergence Theorem (using the constant function 1 as an integrable majorant):

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-1}(A) \cap B) &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_X \mathbb{1}_{T^{-k}(A)\cap B} \, d\mu \\
&= \lim_{n\to\infty} \int_X \mathbb{1}_B \left(\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_A(T^k(x)) \right) \, d\mu \\
\text{(LDCT)} \qquad &= \int_X \mathbb{1}_B \mu(A) \, d\mu \\
&= \mu(A)\mu(B).
\end{aligned}$$

(5 $\implies$ 1): For $B \in \mathcal{P}$, let

$$\mathcal{M}_B = \left\{ A \in \mathcal{B} \ : \ \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B) \right\}.$$

By hypothesis, we know that this contains $\mathcal{P}$; we shall show that this is a λ-system. First, we assume that this is true. Then the π-λ Theorem tells us that $\mathcal{B} = \sigma(\mathcal{P}) \subset \mathcal{M}_B \subset \mathcal{B}$, and thus that $\mathcal{M}_B = \mathcal{B}$. Note that we could have swapped A and B in the definition of $\mathcal{M}_B$, and the argument would not have changed. Then, for any $A \in \mathcal{B}$, we know that $\mathcal{M}_A$ is a λ-system containing $\mathcal{P}$, hence by the π-λ Theorem again we see that $\mathcal{M}_A = \mathcal{B}$, and since this holds for any $A \in \mathcal{B}$, we are done.

It thus remains to show that $\mathcal{M}_B$ is a λ-system.

- We see that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(X) \cap B) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(X \cap B) = \frac{n}{n}\mu(B) = \mu(X)\mu(B),$$

  so indeed, $X \in \mathcal{M}_B$.

- Let $A \in \mathcal{M}_B$. Then we have $\mu(T^{-k}(X \setminus A) \cap B) = \mu((T^{-k}(X) \setminus T^{-k}(A)) \cap B) = \mu(X \cap B) - \mu(T^{-k}(A) \cap B) = \mu(B) - \mu(T^{-k}(A) \cap B)$, so that:

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(X \setminus A) \cap B) \\
&\quad = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\mu(B) - \mu(T^{-k}(A) \cap B)) \\
&\quad = \lim_{n\to\infty} \left(\mu(B) - \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) \right) \\
&\quad = \mu(B) - \mu(B)\mu(A) = (1 - \mu(A))\mu(B) = \mu(X \setminus A)\mu(B).
\end{aligned}$$

  Therefore $X \setminus A \in \mathcal{M}_B$.

- Finally, let $\{A_i\}_{i=1}^{\infty} \subset \mathcal{M}_B$, with $A_i \cap A_j = \emptyset$ when $i \neq j$, and $A = \bigcup_{i=1}^{\infty} A_i$. Let $\epsilon > 0$. Since

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i) \leq \mu(X) = 1,$$

we may find $N > 0$ such that for

$$\sum_{i=N+1}^{\infty} \mu(A_i) < \frac{\epsilon}{3}.$$

For each $i = 1, \ldots, N$, we can find n_i large enough that for all $n > n_i$,

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A_i) \cap B) - \mu(A_i)\mu(B) \right| < \frac{\epsilon}{3N};$$

let $n^* = \max n_i$. Then for $n > n^*$, noting that

$$\mu(T^{-k}(A_i) \cap B) \leq \mu(T^{-k}(A_i) = \mu(A_i),$$

we obtain:

$$\begin{aligned}
&\left| \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) - \mu(A)\mu(B) \right| \\
&= \left| \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{i=1}^{N} \mu(T^{-k}(A_i) \cap B) + \sum_{i=N+1}^{\infty} \mu(T^{-k}(A_i) \cap B) \right) \right. \\
&\qquad\qquad \left. - \mu(B) \sum_{i=1}^{N} \mu(A_i) - \mu(B) \sum_{i=N+1}^{\infty} \mu(A_i) \right| \\
&\leq \left| \sum_{i=1}^{N} \left(\frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A_i) \cap B) - \mu(A_i)\mu(B) \right) \right| \\
&\qquad + \left| \frac{1}{n} \sum_{k=0}^{n-1} \sum_{i=N+1}^{\infty} \mu(T^{-k}(A_i) \cap B) \right| + \left| \mu(B) \sum_{i=N+1}^{\infty} \mu(A_i) \right| \\
&< N \cdot \frac{\epsilon}{3N} + \frac{1}{n} \cdot \frac{n\epsilon}{3} + 1 \cdot \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Hence $A \in \mathcal{M}_B$, and so $\mathcal{M}_B$ is a λ-system, as desired.

□

We also present the proof of Proposition 2.39.

Proof. The forwards direction is immediate. For the backwards direction, we perform a similar computation to that in the proof of Theorem 2.37 above, in the ($5 \implies 1$) direction. For $B \in \mathcal{P}$, let

$$\mathcal{M}_B = \left\{ A \in \mathcal{B} \ : \ \lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B) \right\}.$$

By hypothesis, we know that this contains $\mathcal{P}$; we shall show that this is a λ-system. First, we assume that this is true. Then the π-λ Theorem tells us that $\mathcal{B} = \sigma(\mathcal{P}) \subset \mathcal{M}_B \subset \mathcal{B}$, and thus that $\mathcal{M}_B = \mathcal{B}$. Note that we could have swapped A and B in the definition of $\mathcal{M}_B$, and the argument would not have changed. Then, for any $A \in \mathcal{B}$, we know that $\mathcal{M}_A$ is a λ-system containing $\mathcal{P}$, hence by the π-λ Theorem again we see that $\mathcal{M}_A = \mathcal{B}$, and since this holds for any $A \in \mathcal{B}$, we are done.

It thus remains to show that $\mathcal{M}_B$ is a λ-system.

- We see that

$$\lim_{n\to\infty} \mu(T^{-n}(X) \cap B) = \lim_{n\to\infty} \mu(X \cap B) = \mu(B) = \mu(X)\mu(B),$$

  so indeed, $X \in \mathcal{M}_B$.

- Let $A \in \mathcal{M}_B$. Then we have

$$\begin{aligned}\mu(T^{-n}(X \setminus A) \cap B) &= \mu((T^{-n}(X) \setminus T^{-n}(A)) \cap B) \\ &= \mu(X \cap B) - \mu(T^{-n}(A) \cap B) = \mu(B) - \mu(T^{-n}(A) \cap B),\end{aligned}$$

so that:

$$\lim_{n\to\infty} \mu(T^{-n}(X \setminus A) \cap B) = \lim_{n\to\infty} (\mu(B) - \mu(T^{-n}(A) \cap B))$$
$$= \mu(B) - \mu(B)\mu(A) = (1 - \mu(A))\mu(B) = \mu(X \setminus A)\mu(B).$$

Therefore $X \setminus A \in \mathcal{M}_B$.

- Finally, let $\{A_i\}_{i=1}^{\infty} \subset \mathcal{M}_B$, with $A_i \cap A_j = \emptyset$ when $i \neq j$, and $A = \bigcup_{i=1}^{\infty} A_i$. Let $\epsilon > 0$. Since

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i) \leq \mu(X) = 1,$$

we may find $N > 0$ such that for

$$\sum_{i=N+1}^{\infty} \mu(A_i) < \frac{\epsilon}{3}.$$

For each $i = 1, \ldots, N$, we can find n_i large enough that for all $n > n_i$,

$$\left|\mu(T^{-n}(A_i)) \cap B) - \mu(A_i)\mu(B)\right| < \frac{\epsilon}{3N};$$

let $n^* = \max n_i$. Then for $n > n^*$, noting that

$$\mu(T^{-k}(A_i) \cap B) \leq \mu(T^{-k}(A_i) = \mu(A_i),$$

we obtain:

$$
\begin{aligned}
&\left|\mu(T^{-n}(A)\cap B)-\mu(A)\mu(B)\right| \\
&=\left|\sum_{i=1}^{N}\mu(T^{-n}(A_i)\cap B)+\sum_{i=N+1}^{\infty}\mu(T^{-n}(A_i)\cap B)\right. \\
&\qquad\qquad\qquad \left. -\mu(B)\sum_{i=1}^{N}\mu(A_i)-\mu(B)\sum_{i=N+1}^{\infty}\mu(A_i)\right| \\
&\leq\left|\sum_{i=1}^{N}\left(\mu(T^{-n}(A_i)\cap B)-\mu(A_i)\mu(B)\right)\right| \\
&\qquad\qquad +\left|\sum_{i=N+1}^{\infty}\mu(T^{-n}(A_i))\cap B)\right|+\left|\mu(B)\sum_{i=N+1}^{\infty}\mu(A_i)\right| \\
&< N\cdot\frac{\epsilon}{3N}+\frac{\epsilon}{3}+1\cdot\frac{\epsilon}{3}=\epsilon.
\end{aligned}
$$

Hence $A \in \mathcal{M}_B$, and so $\mathcal{M}_B$ is a λ-system, as desired.

□

A.3 Character Theory

We present some results from character theory which are utilized in the thesis. See [6] and [21] for more detailed treatments.

Definition A.4. Let G be a locally compact Abelian topological group, equipped with Haar measure m on the Borel σ-algebra $\mathcal{B}$. The *character group* of G, denoted $\hat{G}$, is the set of continuous group homomorphisms from G into the complex unit circle, where the group operation is pointwise multiplication.

Lemma A.5. *All characters of* $\mathbb{T} \times \mathbb{Z}_2$ *are of the form* $(x, b) \mapsto e^{2\pi i n x}(-1)^{ab}$*, where* $n \in \mathbb{Z}$*, and* $a \in \mathbb{Z}_2$*.*

Proof. For Abelian groups G, H with characters $\hat{G}, \hat{H}$, let γ be a character of $G \times H$. Observe that γ_G defined by $\gamma_G(g) = \gamma(g, e_H)$ is an element of $\hat{G}$; similarly, γ_H, defined by $\gamma_H(h) = \gamma(e_G, h)$, is an element of $\hat{H}$. However, we have

$$\gamma(g,h) = \gamma((g,e_h)\cdot(e_G,h)) = \gamma(g,e_H)\gamma(e_G,h) = \gamma_G(g)\gamma_H(h),$$

and so $\gamma \in \hat{G} \times \hat{H}$. The reverse inclusion is trivial to prove, and so all characters of $G \times H$ are products of a character of G and a character of H.

Next, we prove that the characters of Z_n are functions $k \mapsto e^{\frac{2\pi imk}{n}}$, for each $m \in Z_n$. Observe that if γ is a character of Z_n, then

$$1 = \gamma(0) = \gamma(\underbrace{1 + \cdots + 1}_{n}) = \gamma(1)^n,$$

so that $\gamma(1)$ is an n^{th} root of unity, of which there are n distinct ones, each taking the form $e^{\frac{2\pi im}{n}}$. The statement is proved by observing $\gamma(k) = \gamma(1)^k$. In the case of $n = 2$, the square roots of 1 are $+1$ and -1.

Finally, we determine the characters of $\mathbb{T}$. Let $\gamma \in \hat{\mathbb{T}}$. Since γ is assumed to take values on the unit circle and is continuous, we may find $\delta \in (0,1)$ such that

$$\int_0^\delta \gamma(x)\ dx = C \neq 0.$$

Using the homomorphism property, multiply by $\gamma(t)$ to obtain:

$$C\gamma(t) = \int_0^\delta \gamma(t)\gamma(x)\ dx = \int_0^\delta \gamma(x+t)\ dx = \int_t^{t+\delta} \gamma(x)\ dx.$$

Since γ is continuous, we see that this expression is differentiable, hence γ is differentiable. We differentiate the expression $\gamma(x+t) = \gamma(x)\gamma(t)$ with respect to t to get

$$\gamma'(x+t) = \gamma(x)\gamma'(t)$$

and set $t = 0$ to obtain:

$$\gamma'(x) = \gamma'(0)\gamma(x),$$

which is a first-order differential equation on the real line (with initial condition $\gamma(0) = 1$), which has solution

$$\gamma(x) = e^{2\pi i \gamma'(0) x}.$$

Indeed, γ is periodic, but we also require that $\gamma(0) = \gamma(1)$ to fulfil continuity requirements, which forces $\gamma'(0) = n \in \mathbb{Z}$. Hence $\gamma = \gamma_n$, with $\gamma_n(x) = e^{2\pi i n x}$. □

Bibliography

[1] L. Arnold. *Random dynamical systems.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[2] L. Arnold, D. C. Nguyen, and V. I. Oseledets. Jordan normal form for linear cocycles. *Random Operators and Stochastic Equations*, 7, 1999.

[3] L. Barreira and Y. Pesin. *Introduction to Smooth Ergodic Theory.* American Mathematical Society, 2013.

[4] G. D. Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 1931.

[5] A. Boyarsky and P. Gòra. *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension.* Birkhäuser, 1997.

[6] T. Bröcker and T. tom Dieck. *Representations of Compact Lie Groups.* Springer-Verlag, 1985.

[7] M. Dellnitz, G. Froyland, C. Horenkamp, and K. Padberg. On the approximation of transport phenomena—a dynamical systems approach. *GAMM-Mitteilungen*, 32, 2009.

[8] G. Froyland, S. Lloyd, and A. Quas. A semi-invertible Oseledets theorem with applications to transfer operator cocycles. *Discrete and Continuous Dynamical Systems*, 33, 2013.

[9] H. Furstenberg. Strict ergodicity and transformation of the torus. *American Journal of Mathematics*, 83, 1961.

[10] C. González-Tokman and A. Quas. A concise proof of the multiplicative ergodic theorem on Banach spaces. *Preprint*, 2014.

[11] P. R. Halmos. Approximation theories for measure preserving transformations. *Transactions of the American Mathematical Society*, 55, 1944.

[12] P. R. Halmos. *Measure theory*. D. Van Nostrand Company, Inc., New York, N. Y., 1950.

[13] Y. Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004.

[14] Z. Lian and K. Lu. Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space. *Memoirs of the American Mathematical Society*, 206, 2010.

[15] R. Mañé. Lyapounov exponents and stable manifolds for compact transformations. In *Geometric dynamics (Rio de Janeiro, 1981)*, volume 1007 of *Lecture Notes in Math.* Springer, Berlin, 1983.

[16] V. I. Oseledets. A multiplicative ergodic theorem. Characteristic Lyapunov exponents of dynamical systems. *Trudy Moskovskogo Matematičeskogo Obščestva*, 19, 1968.

[17] W. Parry. *Topics in Ergodic Theory*. Cambridge University Press, 1981.

[18] R. R. Phelps. *Lectures on Choquet's theorem*, volume 1757 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001.

[19] M. S. Raghunathan. A proof of Oseledec's multiplicative ergodic theorem. *Israel Journal of Mathematics*, 32, 1979.

[20] G. Rauzy. Échanges d'intervalles et transformations induites. *Polska Akademia Nauk. Instytut Matematyczny. Acta Arithmetica*, 34, 1979.

[21] W. Rudin. *Fourier Analysis on Groups*. Interscience Publishers (a division of John Wiley and Sons), 1962.

[22] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Company, New York, third edition, 1987.

[23] D. J. Rudolph. Classifying the isometric extensions of a Bernoulli shift. *Journal d'Analyse Mathématique*, 34, 1978.

[24] D. Ruelle. Characteristic exponents and invariant manifolds in Hilbert space. *Annals of Mathematics. Second Series*, 115, 1982.

[25] O. Sarig. Lecture notes on ergodic theory, 2008. Last accessed Feb. 13, 2015. http://www.math.psu.edu/sarig/506/ErgodicNotes.pdf.

[26] K. Schmidt. *Lectures on Cocycles of Ergodic Transformation Groups*. Macmillan Company of India, Ltd, 1977.

[27] K. Schmidt. A cylinder flow arising from irregularity of distribution. *Compositio Mathematica*, 36, 1978.

[28] P. Thieullen. Ergodic reduction of random products of two-by-two matrices. *Journal d'Analyse Mathématique*, 73, 1997.

[29] W. A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Annals of Mathematics. Second Series*, 115, 1982.

[30] J. von Neumann. Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 1932.

[31] P. Walters. *Introduction to Ergodic Theory.* Springer-Verlag, 1982.

[32] P. Walters. A dynamical proof of the multiplicative ergodic theorem. *Transactions of the American Mathematical Society*, 335, 1993.