Generic properties of extensions

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr. rer. nat.)

im Fachgebiet

Mathematik

Vorgelegt

von M.Sc. B.Sc. Michael Schnurr

geboren am 28.09.1991 in Cleveland

Die Annahme der Dissertation wurde empfohlen von:

1. Prof. Dr. Tatjana Eisner, Universität Leipzig
2. Prof. Dr. Christoph Thiele, Universität Bonn

Die Verleihung des akademischen Grades erfolgt mit Bestehen
der Verteidigung am 30.11.2018 mit dem Gesamtprädikat *magna cum laude*.

To my family

You have always been the rocks that I have needed

To my friends

You have helped me pull through my troubled times, and I am forever grateful

Contents

Introduction **iii**

1 Classical Results and Background **1**

1.1 Halmos and Rokhlin Category Theorems 2

1.1.1 Weak Topology 2

1.1.2 Weak Approximation 3

1.1.3 Uniform Topology and Uniform Approximation 7

1.1.4 Category Theorems 11

1.2 Examples of Weakly But Not Strongly Mixing Transformations 15

1.3 Other Category Results 18

2 Genericity for Extensions **22**

2.1 Non-Fixed Factor 22

2.1.1 Notation and Preliminaries 23

2.1.2 Discrete Extensions 25

2.1.3 Weak Approximation Theorem for Extensions on the Unit Square 30

2.1.4 Uniform Approximation 37

2.1.5 Conjugacy Lemma 38

2.1.6 Category Theorem 39

2.1.7 General Case for the Fiber Space 46

2.1.8 Strongly Mixing Extensions 47

2.2 Fixed Factor 49

2.2.1 Notation and Definitions 49

2.2.2 Conjugacy Lemma for Weakly Mixing Extensions 51

2.2.3 Strongly Mixing Extensions 53

2.2.4 Rigid Extensions and Other Relativizations 55

3 Further Questions **60**

Acknowledgments **63**

Selbstständigkeitserklärung **65**

Bibliography **67**

Introduction

Ergodic theory was born in 1931 when von Neumann and Birkhoff presented their ergodic theorems, which lead to a proof of Boltzmann's ergodic hypothesis for ergodic systems. Subsequently, this proof gave a description of ergodic systems in terms of their asymptotic behavior. Ergodic systems are not the only systems which can be described by their asymptotic behavior, though. For example, another desirable property of measure-preserving systems is strongly mixing, which is easy to define, but difficult to check. Between ergodicity and strongly mixing is weakly mixing, which is comparatively more difficult to define, but much easier to check via spectral properties of the system.

From their definitions by their asymptotic behavior, it is immediately apparent that all strongly mixing transformations are weakly mixing, and all weakly mixing transformations are ergodic. On the other hand, it is easy to create an example of a measure-preserving transformation which is ergodic, but not weakly mixing. An irrational rotation on the circle (with the Haar measure) is such an example. However, the case of weakly and strongly mixing transformation proved more difficult. Certainly it was suspected that weakly mixing transformations are not necessarily strongly mixing, but constructing an example of such a transformation proved to be a challenge.

In 1944, Halmos showed in [30] that in some sense "most" measure-preserving transformations are weakly mixing – that being weakly mixing is a generic property of a measure-preserving transformation. Not long after, in 1948, Rokhlin showed in [44] that in that same sense, "most" measure-preserving transformations are *not* strongly mixing – that being strongly mixing is a rare property of measure-preserving transformations. Together, these two results showed with resounding strength that weakly mixing transformations are not necessarily strongly mixing. It is interesting to note, however, that it was not until 1969 that an example of a weakly but not strongly mixing transformation was found, when Chacon constructed an example in [10]. Since Halmos, much research has been done in finding "typical" properties of measure-preserving dynamical systems. For example, in 1970, Katok and Stepin showed in [36] that a generic measure-preserving transformation is rigid. For more examples, see, e.g., [39], [1], [37], [2], [5], [3], [15], [4], [47], [29].

More than thirty years after the work of Halmos and Rokhlin, in completely unrelated efforts, Furstenberg [24] presented his celebrated ergodic theoretic proof of Szemeredi's Theorem on the existence of arbitrarily long arithmetic progressions in large subsets of $\mathbb{N}$. While the original proof by Furstenberg used diagonal measures, the alternative proof by him, Katznelson, and Ornstein [25] building up a tower of so-called compact and weakly

mixing extensions had much greater impact on further development of ergodic theory. This method of finding the characteristic factor has been extended to various ergodic theorems and is an active area of research, see, e.g., [12], [13], [7], [20], [9], [23], [19], [51] [6], [42]. For example, the appropriate characteristic factor for norm convergence of multiple ergodic averages was identified by Host and Kra [32] and has the structure of an inverse limit of nilsystems, see also Ziegler [54].

In 2012, Terence Tao posed the natural question [50]: what do typical extensions look like? More precisely, can we extend the theory on generic properties of measure-preserving transformations to theory on generic properties of measure-preserving extensions? Surprisingly, there seemed to be little to no work done in this direction. In [43], Robertson proved a genericity statement for compact group extensions, but there seemed to be no work on genericity of measure-preserving extensions in general. This work partially fills this gap, and was further filled by Glasner and Weiss in [28].

This thesis is organized as follows. In Chapter 1 we give a survey of the classical theory, including some of the proofs as points of comparison for results that will follow in Chapter 2. First in Section 1.1 we focus on giving a detailed survey of Halmos' (and Rokhlin's) category theorems. Then Section 1.2 discusses some examples of measure-preserving transformations which are weakly, but not strongly mixing. Finally Section 1.3 examines a few other properties of measure-preserving transformations which are generic, namely rigidity, α-rigidity, and κ-weakly mixing. Then in Chapter 2 we begin to work towards genericity in extensions. In an effort to closely follow the classical theory of Halmos, we ultimately show in Section 2.1 that the set of weakly mixing extensions of a non-fixed factor is generic in the set of all extensions in this sense of a non-fixed factor. In the same section we show that the corresponding set of strongly mixing extensions is rare. Then in Section 2.2 we approach the same problem, but this time in the context of extensions of a fixed transformation, which makes the problem more difficult. We generalize the result from [28] that the set of weakly mixing extensions of an antiperiodic transformation is generic, provide a sufficient condition for the set of strongly mixing extensions to be rare, and discuss rigid extensions. Finally in Chapter 3 we articulate some open questions.

Chapter 1

Classical Results and Background

Because the theory we will develop in Chapter 2 is so deeply rooted in the classical theory as well as its techniques, in this chapter we will be taking an extensive look at the classical theory. In particular, we will be presenting many proofs in this chapter as points of comparison for later.

Before beginning, we make a few notational notes that will hold throughout this entire work. Let $(X, \mathcal{F}, m)$ be a measure space. We will be making a few assumptions about this measure space and about measure-preserving transformations on it. First of all, we will consider only measure-preserving transformations which are invertible. Second, we will not distinguish two measure-preserving transformations which disagree on a set of measure 0. That is, we identify T, S if $m(\{x \in X | Tx \neq Sx\}) = 0$. Lastly, $(X, \mathcal{B}, m)$ will be assumed to be the unit interval with the Lebesgue measure.

We make two notes. First, the first two assumptions mean that we are technically considering T as an automorphism of the measure algebra of $\mathcal{F}$. However, we will generally still call T a measure-preserving transformation. Second, though it might seem restrictive to assume X to be a specific measure space, it is indeed not the case. It is known that all standard, non-atomic probability spaces are isomorphic as measure algebras, and the unit interval is the quintessential example (see [30, p. 61]).

As two last notational notes, we will be using χ_A to indicate the characteristic function of the set A, and $\mathbb{1}_X$ for the identity transformation on X.

Finally, we wish to recall a few definitions/characterizations. First, a measure-preserving transformation T on X is ergodic if and only if for all $f, g \in L^2(X)$ (and using the abusive notation of denoting the corresponding Koopman operator by T as well, i.e. $Tf = f \circ T$),

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \langle T^n f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle.$$

Further, T is weakly mixing if and only if for all $f, g \in L^2(X)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle T^n f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| = 0.$$

Lastly T is, by definition, strongly mixing if and only if for all $f, g \in L^2(X)$,

$$\lim_{n\to\infty} |\langle T^n f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| = 0.$$

We denote the set of measure-preserving transformations on X by $\mathcal{G}(X)$, the set of weakly mixing transformations on X by $\mathcal{W}(X)$, and the set of strongly mixing transformations on X by $\mathcal{S}(X)$.

1.1 Halmos and Rokhlin Category Theorems

In this section we will be exploring the proofs of Halmos' and Rokhlin's results, as they are presented in [31]. We give a brief overview of the strategy of Halmos' proof here. First, we define the weak topology on the set of measure-preserving transformations on a given measure space, as well as a metric which induces this topology. Having done that, we will give two results which approximate measure-preserving transformations by other measure-preserving transformations in different ways. We are then able to use those approximation results to show that conjugations of any fixed antiperiodic transformation are dense in the weak topology. With that, we show that the set of weakly mixing transformations form a dense, G_δ set in the weak topology. After, we show that strongly mixing transformations form a first category set in the same topology.

1.1.1 Weak Topology

We wish to define a topology on $\mathcal{G}(X)$. The topology, called the weak topology, corresponds to weak convergence, where a sequence $\{T_n\}$ converges weakly to T if and only if for all measurable sets $E, m(T_n E \triangle TE) \to 0$. We make this explicit by defining a subbasis as follows. Given $T \in \mathcal{G}(X), E \in \mathcal{F}, \varepsilon > 0$, define

$$N_{\varepsilon,E}(T) := \{S \in \mathcal{G}(X) | m(TE \triangle SE) < \varepsilon\}.$$

Then the collection of sets $N_{\varepsilon,E}(T)$ form a subbasis for the weak topology. However, it happens to be superfluous to use all measurable sets E when forming a subbasis for the weak topology.

Definition 1.1. A *dyadic interval of rank* k in X is an interval of the form

$$\left(\frac{j}{2^k}, \frac{j+1}{2^k}\right),$$

where $0 \le j \le 2^k - 1$. A *dyadic set of rank* k is a finite union of dyadic intervals of rank at most k.

An important and well-known fact is that all Lebesgue-measurable sets can be approximated arbitrarily closely by dyadic sets. Thus we see that for a subbasis of the weak topology, it is sufficient to consider only dyadic sets.

Proposition 1.1. *Let $T \in \mathcal{G}(X)$, D be a dyadic set, and $\varepsilon > 0$. Then the collection of sets of the form*

$$\tilde{N}_{\varepsilon,D}(T) := \{S \in \mathcal{G}(X) | m(TD \triangle SD) < \varepsilon\}$$

form a subbasis for the weak topology.

Proof. As every dyadic set D is also a measurable set, we need only show that given a set $N_{\varepsilon,E}(T)$, there exists a dyadic set D and a $\delta > 0$ such that $\tilde{N}_{\delta,D}(T) \subset N_{\varepsilon,E}(T)$. Let D be a dyadic set such that $D \subset E$ and $m(E \backslash D) < \varepsilon/4$. Set $\delta := \varepsilon/2$. Then given $S \in \tilde{N}_{\delta,D}(T)$, we have

$$\begin{aligned} m(TE \triangle SE) &\leq m(TD \triangle SD) + m(T(E \backslash D) \triangle S(E \backslash D)) \\ &< \delta + 2m(E \backslash D) = \varepsilon. \end{aligned}$$

So $S \in N_{\varepsilon,E}(T)$. □

Note that we could have defined the weak topology in terms of the Koopman operators. In this case, we would have that for all measurable sets $E, T_n\chi_E$ converges in $L^2(X)$ to $T\chi_E$. Indeed, $\|T_n\chi_E - T\chi_E\|_2 = m(T_nE \triangle TE) \to 0$. But if $T_n\chi_E \to T\chi_E$ for all measurable sets E, then $T_nf \to Tf$ for all $f \in L^2(X)$. Thus we see that for the Koopman operators, the weak topology coincides with the strong operator topology. Further, since Koopman operators are all unitary, the weak topology also happens to coincide with the weak operator topology.

Our last note about the weak topology is that is indeed metrizable, and a metric is known. Let $\{D_n\}$ be the (in particular countable) collection of dyadic intervals in X. Then we can define the distance between two elements of $\mathcal{G}(X)$ by

$$d(T,S) := \sum_n \frac{m(TD_n \triangle SD_n) + m(T^{-1}D_n \triangle S^{-1}D_n)}{2^n}.$$

While this metric can have its uses, in general we will not be utilizing it unless it becomes necessary, as it is rather cumbersome. We do, however, note that the weights $1/2^n$ are not particularly important. The metric d indicates, more or less, that T and S are "close" if they map all dyadic intervals to similar sets. Thus, we could have defined d by

$$d(T,S) := \sum_n a_n \left(m(TD_n \triangle SD_n) + m(T^{-1}D_n \triangle S^{-1}D_n)\right),$$

where

$$\sum_n a_n < \infty,$$

and we would still have a metric which induces the weak topology.

1.1.2 Weak Approximation

We open this section with the following definition.

Definition 1.2. A *dyadic permutation of rank* k is a measure-preserving transformation on X which maps the dyadic intervals of rank k to each other bijectively, by translation (or more simply, it permutes the dyadic intervals of rank k). A *cyclic dyadic permutation of rank* k is a dyadic permutation of rank k which contains a single cycle in its cycle decomposition. Stated differently, T is a cyclic dyadic permutation of rank k if it is a dyadic permutation (of rank k) such that for all dyadic intervals of rank k, A, B, there exists exactly one n with $0 \leq n \leq 2^n - 1$, such that $T^n A = B$.

A perhaps surprising fact which Halmos showed, and which we will show in this section, is that all measure-preserving transformations can be approximated (in the weak topological sense) by cyclic dyadic permutations. In order to show that, we must take a few steps. The first is a purely measure-theoretic result, and is rather technical, but is crucial.

Lemma 1.1. *Let* $\{E_i : i = 1, \ldots n\}$ *partition the unit interval, and* r_i *be dyadic rationals such that*

$$\sum_{i=1}^{n} r_i = 1,$$

and $|m(E_i) - r_i| < \delta$ *for some* $\delta > 0$ *and for all* i*. Then there exists* $\{F_i : i = 1, \ldots n\}$*, dyadic sets that partition the unit interval, such that* $m(F_i) = r_i$ *and* $m(E_i \triangle F_i) < 2\delta$ *for all* i*.*

Proof. We will construct the partition $\{F_i\}$ in steps. We start by choosing F_i'' merely to be dyadic sets such that for all $i, m(E_i \triangle F_i'') < \gamma$ for some $\gamma > 0$ to be chosen later. We do not require at this time that $m(F_i'') = r_i$, nor that the F_i'' are pairwise disjoint. Note that

$$\begin{aligned} & m(E_i) - m(F_i'') \\ =& (m(E_i \backslash F_i'') + m(E_i \cap F_i'')) - (m(E_i \cap F_i'') + m(F_i'' \backslash E_i)) \\ =& m(E_i \backslash F_i'') - m(F_i'' \backslash E_i). \end{aligned}$$

Thus

$$\begin{aligned} |m(E_i) - m(F_i'')| &= |m(E_i \backslash F_i'') - m(F_i'' \backslash E_i)| \\ \leq |m(E_i \backslash F_i'')| &+ |m(F_i'' \backslash E_i)| = m(E_i \triangle F_i'') < \gamma. \end{aligned}$$

Subsequently,

$$|m(F_i'') - r_i| \leq |m(F_i'') - m(E_i)| + |m(E_i) - r_i| < \gamma + \delta.$$

We will now modify F_i''. First we will make them disjoint. Note that

$$m((E_i \cap E_j) \triangle (F_i'' \cap F_j'')) \leq m(E_i \triangle F_i'') + m(E_j \triangle F_j'') < 2\gamma.$$

However, $\{E_i\}$ form a partition, so for $i \neq j, m(F_i'' \cap F_j'') < 2\gamma$. Thus for fixed i,

$$m\left(F_i'' \cap \bigcup_{j\neq i} F_j''\right) \leq m\left(\bigcup_{j\neq i}(F_i'' \cap F_j'')\right) \leq \sum_{j\neq i} m(F_i'' \cap F_j'') < 2n\gamma.$$

Define

$$F_i' := F_i'' \backslash \left(F_i'' \cap \bigcup_{j\neq i} F_j''\right).$$

Note that as the intersection of dyadic sets is a dyadic set, F_i' are dyadic sets. Now clearly

$$|m(F_i') - r_i| \leq |m(F_i') - m(F_i'')| + |m(F_i'') - r_i| < 2n\gamma + \gamma + \delta.$$

Moreover,

$$m(F_i' \triangle E_i) \leq m(F_i' \triangle F_i'') + m(F_i'' \triangle E_i) < 2n\gamma + \gamma.$$

Now F_i' are disjoint, but their union is not X, nor is their measure necessarily r_i. We will need to make one last modification. First, let

$$\tilde{F} := \bigcup_i \left(F_i'' \cap \bigcup_{j\neq i} F_j''\right).$$

Now for each i such that $m(F_i') > r_i$, let $B_i \subset F_i'$ such that B_i is a dyadic set and $m(F_i') - m(B_i) = r_i$, and define $F_i := F_i' \backslash B_i$. Further, let F be the union of $\tilde{F}$ and all B_i chosen up to this point. Then for F_i' with $m(F_i') < r_i$, let B_i be a dyadic subset of F such that $m(F_i') + m(B_i) = r_i$ (and once a set B_i has been chosen, discard it from F). For these i, define $F_i := F_i' \cup B_i$. Clearly for all $i, m(F_i) = r_i$. Further,

$$m(E_i \triangle F_i) \leq m(E_i \triangle F_i') + m(F_i' \triangle F_i) < 2n\gamma + \gamma + 2n\gamma + \gamma + \delta.$$

Thus if we had chosen γ so that $(4n+2)\gamma < \delta$, then $m(E_i \triangle F_i) < 2\delta$. □

Remark 1.1. The fact that we were using dyadic sets for Lemma 1.1 was not particularly important for the proof. Only certain properties of the class $\mathcal{B}$ of dyadic sets are important. Namely, $\mathcal{B}$ is an algebra, all Lebesgue-measurable sets can be approximated arbitrarily closely by elements of $\mathcal{B}$, and the measure of every element of $\mathcal{B}$ is a dyadic rational. Most importantly, if $B \in \mathcal{B}$ and $0 < r \leq m(B)$, there exists $B_0 \subset B$ such that $B_0 \in \mathcal{B}$ and $m(B_0) = r$. Any collection with these properties will also be satisfactory for Lemma 1.1. For example, if $T \in \mathcal{G}(X)$, then $T\mathcal{B}$ is such a class.

Theorem 1.1. *Dyadic permutations are dense in the weak topology on $\mathcal{G}(X)$.*

Proof. Let $T \in \mathcal{G}(X), \varepsilon > 0$ and let

$$N_\varepsilon(T) := \{S \in \mathcal{G}(X) | m(TD_i \triangle SD_i) < \varepsilon, i = 1, ..., N\}$$

be a dyadic neighborhood of T. Without loss of generality, $N = 2^k$ for some k, and D_i are all (distinct) dyadic intervals of rank k. This can be done because all basic dyadic neighborhoods of T contain a neighborhood of this form. To see this, take any basic

dyadic neighborhood (with δ in place of ε), and let k be common rank of all the dyadic sets in the definition of the neighborhood. Then letting $\varepsilon = \delta/2^k$, we see $N_\varepsilon(T)$, with $N = 2^k$, is contained in this neighborhood.

Consider $\{D_i \cap TD_j | 1 \leq i, j \leq N\}$. As T is invertible, this is a partition of X. Set $D_{ij} := D_i \cap TD_j$. Applying Lemma 1.1, we find a partition of X into dyadic sets $\{E_{ij}\}$ such that $m(D_{ij} \triangle E_{ij}) < \varepsilon/2^{k+1}$. We then apply Lemma 1.1 again (this time in light of Remark 1.1) to find a partition of X into dyadic sets $\{F_{ij}\}$ such that $m(D_{ij} \triangle TF_{ij}) < \varepsilon/2^{k+1}$. When applying Lemma 1.1 the second time, we take the same dyadic rationals r_{ij} that we do in the first application of the lemma, so that $m(E_{ij}) = m(F_{ij})$.

Now let Q be a dyadic permutation which maps $F_{ij} \mapsto E_{ij}$ for all i, j. Fix j. Observe that

$$\begin{aligned} &m\left(D_j \triangle \bigcup_i F_{ij}\right) \\ =&m\left(\bigcup_i T^{-1}D_{ij} \triangle \bigcup_i F_{ij}\right) \\ \leq&\sum_i m(T^{-1}D_{ij} \triangle F_{ij}) < \frac{\varepsilon}{2}. \end{aligned}$$

Subsequently,

$$m\left(QD_j \triangle \bigcup_i E_{ij}\right) < \frac{\varepsilon}{2}.$$

On the other hand,

$$\begin{aligned} m\left(TD_j \triangle \bigcup_i E_{ij}\right) &= m\left(\bigcup_i D_{ij} \triangle \bigcup_i E_{ij}\right) \\ \leq \sum_i m(D_{ij} \triangle E_{ij}) &< \frac{\varepsilon}{2}. \end{aligned}$$

Thus,

$$m(QD_j \triangle TD_j) \leq m\left(QD_j \triangle \bigcup_i E_{ij}\right) + m\left(TD_j \triangle \bigcup_i E_{ij}\right) < \varepsilon.$$

As this holds for all $j, Q \in N_\varepsilon(T)$. □

Finally, we come to the promised result.

Theorem 1.2. *Let $T \in \mathcal{G}(X), \varepsilon > 0$ and $N_\varepsilon(T)$ be a basic dyadic neighborhood of T. Then for all $n \in \mathbb{N}$, there exists $Q \in N_\varepsilon(T)$, a cyclic dyadic permutation of rank $k \geq n$.*

Proof. By Theorem 1.1, we can assume that T is itself a dyadic permutation (we will therefore use P instead of T from here). Let n be fixed, and let N be the number of disjoint cycles in P, and let k_0 be a common rank for all sets PD_i, where D_i are the sets

that define $N_\varepsilon(P)$. Note that P can be seen as a dyadic permutation of rank k_0. Choose $k \geq \max\{n, k_0\}$ such that $N/2^{k-1} < \varepsilon$.

We will construct Q, a cyclic dyadic permutation of rank k. To do this, we begin by choosing E_1 to be any dyadic interval of rank k_0 (which we will henceforth refer to as k_0-interval). Now, if $PE_1 \neq E_1$, we define Q to map the "first" k-interval of E_1 (by first, we mean the left most k-interval within E_1) to the first k-interval of PE_1. Then if $P^2E_1 \neq E_1$, we define Q to map the first k-interval of PE_1 to the first k-interval of P^2E_1.

Eventually we will reach a point where $P^{j_1}E_1 = E_1$. When this happens, we map the first k-interval of $P^{j_1-1}E_1$ to the second k-interval of E_1. We repeat the previous process, replacing "first" with "second", and then third, and so forth. Eventually, we will reach the "last" (rightmost) k-interval of $P^{j_1-1}E_1$. At this point, we choose E_2, a k_0-interval which is not in the same P-cycle as E_1, and we define Q to map the last k-interval of $P^{j_1-1}E_1$ to the first k-interval of E_2. We then repeat this entire process until we have exhausted all independent cycles of P. When we have reached the last k-interval of $P^{j_N-1}E_N$, we define Q to map it back to the first k-interval of E_1 to close the cycle.

Let $F_l, 1 \leq l \leq N$ be the last k-interval in $P^{j_l-1}E_l$ and fix i. Observe that

$$\begin{aligned} & m(PD_i \triangle QD_i) \\ \leq & m\left(\bigcup_l (PF_l \triangle QD_l)\right) \\ \leq & \sum_l m(PF_l \triangle QF_l) \leq \frac{2N}{2^k} < \varepsilon. \end{aligned}$$

Thus $Q \in N_\varepsilon(P)$. □

1.1.3 Uniform Topology and Uniform Approximation

While it is the topology of our primary interest, the weak topology is not the only natural topology on $\mathcal{G}(X)$. There is another topology, which we call the uniform topology. Of particular interest to us are two metrics which induce the uniform topology. First we have

$$\tilde{d}(T, S) := \sup_{E \in \mathcal{B}} m(TE \triangle SE).$$

On the other hand, we also have

$$d'(T, S) := m\{x \in X : Sx \neq Tx\}.$$

Our first goal for this section is showing that $\tilde{d}$ and d' induce the same metrics. First we prove an auxiliary result for later usage.

Proposition 1.2. *The metrics $\tilde{d}, d'$ are invariant under group operations in $\mathcal{G}(X)$. That is, given $R, S, T \in \mathcal{G}(X)$,*

$$\begin{aligned} \tilde{d}(S, T) = \tilde{d}(RS, RT) = \tilde{d}(SR, TR) = \tilde{d}(S^{-1}, T^{-1}), \\ d'(S, T) = d'(RS, RT) = d'(SR, TR) = d'(S^{-1}, T^{-1}). \end{aligned}$$

Proof. First for $\tilde{d}$, we have that

$$\begin{aligned}\tilde{d}(SR, TR) &= \sup_E m(S(RE)\triangle T(RE)) \\ &= \sup_{RE} m(SE\triangle TE) \\ &= \sup_E m(SE\triangle TE) = \tilde{d}(S,T).\end{aligned}$$

On the other hand, as $m(RA\triangle RB) = m(R(A\triangle B)) = m(A\triangle B)$ for all A, B, $\tilde{d}(S,T) = \tilde{d}(RS, RT)$.

Finally, we have

$$\tilde{d}(S^{-1}, T^{-1}) = \tilde{d}(\mathbb{1}_X, T^{-1}S) = \tilde{d}(T,S) = \tilde{d}(S,T).$$

For d', first note that $Sx \neq Tx$ if and only if $RSx \neq RTx$, so $d'(S,T) = d'(RS, RT)$. To show $d'(S^{-1}, T^{-1}) = d'(S,T)$, we show that

$$S\{x|Sx \neq Tx\} = \{x|S^{-1}x \neq T^{-1}x\}.$$

Denote $A := \{x|Sx \neq Tx\}$ and $B := \{x|S^{-1}x \neq T^{-1}x\}$. Let $z \in A$, and define $y := Sz$. Suppose $y \notin B$, i.e. $S^{-1}y = T^{-1}y$. Then

$$Tz = TS^{-1}y = TT^{-1}y = y = Sz,$$

which contradicts $z \in A$. So $y \in B$. On the other hand if we have $y \in B$ and define $z := S^{-1}y$. Suppose $z \notin A$. Then

$$T^{-1}y = T^{-1}Sz = T^{-1}Tz = z = S^{-1}y,$$

which contradicts $y \in B$. So $z \in A$.

Lastly, we have

$$d'(SR, TR) = d'(R^{-1}S^{-1}, R^{-1}T^{-1}) = d'(S^{-1}, T^{-1}) = d'(S,T).$$

□

To prove that $\tilde{d}, d'$ induce the same topology, we need two lemmas. First, recall that T is periodic of period n if $T^n = \mathbb{1}_X$. Further recall that T is antiperiodic if

$$m\left(\bigcup_{n=1}^{\infty} \{x|T^n x = x\}\right) = 0.$$

Lemma 1.2. *Let T be periodic of period n. Then there exists a measurable set E such that $m(E) = 1/n$ and $E, TE, \ldots, T^{n-1}E$ are pairwise disjoint.*

Proof. If $n = 1$ there's nothing to prove, as we can take $X = E$. If $n > 1$ then there exists a set E_1 such that $m(E_1 \triangle TE_1) > 0$. Else, it follows that T is periodic of period 1. Now, we claim that $m(E_1 \backslash TE_1) > 0$. Indeed, if $m(E_1 \backslash TE_1) = 0$, then as $m(E_1 \backslash TE_1) =$

$m(E_1) - m(E_1 \cap TE_1)$, we have that $m(E_1) = m(E_1 \cap TE_1)$. On the other hand, $m(TE_1 \backslash E_1) = m(TE_1) - m(E_1 \cap TE_1)$, and as $m(E_1) = m(TE_1), m(TE_1 \backslash E_1) = 0$. Since $m(E_1 \triangle TE_1) = m(E_1 \backslash TE_1) + m(TE_1 \backslash E_1) = 0$, a contradiction.

Now let $F_1 := E_1 \backslash TE_1$. Clearly F_1, TF_1 are disjoint. If $n = 2$ we stop. Else, we find a set $E_2 \subset F_1$ such that $m(E_2 \triangle T^2 E_2) > 0$. This is again possible as else $T^2 E_2 = E_2$. Again we have $m(E_2 \backslash T^2 E_2) > 0$, so we let $F_2 := E_2 \backslash T^2 E_2$. We claim $F_2, TF_2, T^2 F_2$ are pairwise disjoint. F_2, TF_2 follows as $F_2 \subset F_1$. $F_2, T^2 F_2$ follows from definition of F_2. so suppose $TF_2, T^2 F_2$ is nonempty. If $x \in TF_2 \cap T^2 F_2$, then $x \in TE_2 \cap T^2 E_2$, or $T^{-1} x \in E_2 \cap TE_2$. But this is impossible as $E_2 \cap TE_2$ is empty.

We proceed inductively, repeating this process $n - 1$ times. Let $E^1 := F_{n-1}$. Now if $m(E^1) = 1/n$, let $E := E^1$. In general, we may need to repeat the process on the complement of $\bigcup_{i=0}^{n-1} T^i E^1$, and indeed we may need to proceed transfinitely. The process must end at a countable ordinal β as a disjoint collection of sets of positive measure must be countable. Letting

$$E := \bigcup_{\alpha \leq \beta} E^\alpha$$

will necessarily have that $m(E) = 1/n$. □

The other lemma we need is classical and very well-known, so we will merely state it without proof.

Lemma 1.3 (Rokhlin's Lemma)**.** *Let T be antiperiodic. Then for all $n \in \mathbb{N}$ and $\varepsilon > 0$ there exists a measurable set E such that $E, TE, \ldots, T^{n-1} E$ are pairwise disjoint and*

$$m\left(\bigcup_{i=0}^{n-1} T^i E\right) > 1 - \varepsilon.$$

We do make two notes. First, while Rokhlin's Lemma is usually stated for ergodic T instead of antiperiodic T, the most general case presented here does hold. Note that it is this formulation which appears in [31]. Second, the collection $\{E, TE, \ldots, T^{n-1} E\}$ are known as a Rokhlin Tower of height n (corresponding to ε), and the set E in particular is called the base of the tower.

With this, we can show that the metrics $\tilde{d}, d'$ induce the same topology.

Proposition 1.3. *For all $S, T \in \mathcal{G}(X)$,*

$$\frac{2}{3} d'(S,T) \leq \tilde{d}(S,T) \leq d'(S,T).$$

Remark. As we will not be too interested in using the uniform topology itself, we will only present the proof that $\tilde{d}(S,T) \leq d'(S,T)$, though both are necessary to show $\tilde{d}, d'$ induce the same topology.

Proof. As $\tilde{d}, d'$ are both invariant under left and right multiplication, we need only show that for any $T \in \mathcal{G}(X), \tilde{d}(\mathbb{1}_X, T) \leq d'(\mathbb{1}_X, T)$, because then for any $S, T \in \mathcal{G}(X)$ we have

$$\tilde{d}(S,T) = \tilde{d}(\mathbb{1}_X, S^{-1}T) \leq d'(\mathbb{1}_X, S^{-1}T) = d'(S,T).$$

Let $F := \{x | Tx \neq x\}$. Note $d'(\mathbb{1}_X, T) = m(F)$ Clearly $X \backslash F$ is T-invariant (indeed, so is every subset of $X \backslash F$) and thus F is also T-invariant. Let E be any arbitrary measurable set. Then

$$\begin{aligned} m(E \triangle TE) &\leq m((E \cap F) \triangle (TE \cap F)) + m((E \backslash F) \triangle (TE \backslash F)) \\ &= m((E \cap F) \triangle (TE \cap F)) = m((E \triangle TE) \cap F) \leq m(F). \end{aligned}$$

Thus

$$\tilde{d}(\mathbb{1}_X, T) = \sup_E m(E \triangle TE) \leq m(F) = d'(\mathbb{1}_X, T).$$

□

Our second and final goal of this section is to show, more or less, that periodic transformations are dense in the uniform topology. First, we need one more lemma.

Lemma 1.4. *Let E, F be measurable sets with $m(E) = m(F)$. Then there exists $T \in \mathcal{G}(X)$ such that $m(TE \triangle F) = 0$.*

Proof. Let $S \in \mathcal{G}(X)$ be any ergodic transformation. Then there exists an $n_0 \in \mathbb{N}$ such that $m(E \cap S^{-n_0}F) > 0$ For $x \in E \cap S^{-n_0}F$, define $Tx := S^n x$. Now if $m(E \backslash S^{-n_0}F) = 0$, we're done. If not, then there exists $n_1 \in \mathbb{N}$ such that $m((E \backslash S^{-n_0}F) \cap S^{-n_1}(F \backslash S^{n_0}E))$. Then for $x \in (E \backslash S^{-n_0}F) \cap S^{-n_1}(F \backslash S^{n_0}E)$, define $Tx := S^{n_1}x$. Proceed inductively, possibly transfinitely. Eventually we will reach an ordinal where the process terminates. Afterwards, we apply the same process to $X \backslash E$ and $X \backslash F$ in order to define T on all of X. Note that since $S \in \mathcal{G}(X), T$ will also be invertible. □

Theorem 1.3 (Uniform Approximation Theorem)**.** *Let $T \in \mathcal{G}(X)$ be antiperiodic. Then for all $n \in \mathbb{N}$ and $\varepsilon > 0$, there exists $R \in \mathcal{G}(X)$ which is periodic of period n such that*

$$d'(R,T) \leq \frac{1}{n} + \varepsilon.$$

Proof. Let E be the base of a Rokhlin tower of height n, corresponding to ε. Note that $m(E) \leq 1/n$. Now if $x \in T^i E$ for $0 \leq i \leq n-2$, let $Rx := Tx$. For $x \in T^{n-1}x$, define $Rx := T^{n-1}x$. Clearly if x is in the Rokhlin tower, then x is a point with period n under R, no matter how R is defined on the complement of the tower. Thus no matter how we define R on the complement of the tower (so long as it remains invertible), we have

$$d'(R,T) \leq m(T^{n-1}E) + \varepsilon = m(E) + \varepsilon \leq \frac{1}{n} + \varepsilon.$$

□

1.1.4 Category Theorems

We return to the weak topology now to show that the sets of weakly and strongly mixing transformations are indeed different sets. In fact, we show that $\mathcal{W}(X)$ is much larger than $\mathcal{S}(X)$. We begin by showing $\mathcal{W}(X)$ is a dense, G_δ set in the weak topology. To that end, we will need a crucial lemma which, while now well-known, was quite surprising in its time.

Lemma 1.5 (Conjugacy Lemma)**.** *Let $T_0 \in \mathcal{G}(X)$ be antiperiodic. Then the set*

$$\{S^{-1}T_0S | S \in \mathcal{G}(X)\}$$

is dense in the weak topology.

Proof. Let $T \in \mathcal{G}(X)$, and let $N_\varepsilon(T)$ be a dyadic basic neighborhood of T (let $D_1, \ldots, D_k$ be the dyadic sets which define $N_\varepsilon(T)$). Now consider $N_{\varepsilon/2}(T)$, which is the dyadic neighborhood of T which is defined by the same D_i, but with $\varepsilon/2$ instead of ε. By the Weak Approximation Theorem (Theorem 1.2,) there exists $Q \in N_{\varepsilon/2}(T)$ which is a cyclic dyadic permutation of rank k, where k is greater than the ranks of all the D_i and

$$\frac{1}{2^{k-1}} < \frac{\varepsilon}{2}.$$

Now by the Uniform Approximation Theorem (Theorem 1.3), with 2^k and $1/2^k$ in place of n and ε respectively, there exists $R \in \mathcal{G}(X)$ which is periodic of period 2^k such that

$$d'(R, T_0) \leq \frac{1}{2^k} + \frac{1}{2^k} = \frac{1}{2^{k-1}} < \frac{\varepsilon}{2}.$$

We wish to construct S so that $Q = S^{-1}RS$. To this end, let $q := 2^k$ and let $E_0, E_1, \ldots, E_{q-1}$ be the dyadic intervals of rank k, labeled so that $QE_i = E_{i+1}, i \mod q$. Further, by Lemma 1.2, there exists a set F of measure $1/q$ such that $F, RF, \ldots, R^{q-1}F$ are pairwise disjoint. For the sake of notational convenience, define $F_i := R^iF$. Now define S so that $SE_0 = F_0$. Then for $x \in E_i$, define $Sx := R^iSQ^{-i}x$. This can be summarized by the following diagram:

$$\begin{tikzcd}
E_0 \arrow[r, "Q"] \arrow[d, "S"'] & E_1 \arrow[r, "Q"] \arrow[d, "S"'] & E_2 \arrow[r, "Q"] \arrow[d, "S"'] & \ldots \arrow[r, "Q"] & E_{q-2} \arrow[r, "Q"] \arrow[d, "S"'] & E_{q-1} \arrow[d, "S"'] \\
F_0 \arrow[r, "R"'] & F_1 \arrow[r, "R"'] & F_2 \arrow[r, "R"'] & \ldots \arrow[r, "R"'] & F_{q-2} \arrow[r, "R"'] & F_{q-1}
\end{tikzcd}$$

Commutativity of the diagram shows that $Q = S^{-1}RS$.

We are nearly done. We have

$$d'(Q, S^{-1}T_0S) = d'(S^{-1}RS, S^{-1}T_0S) = d'(R, T_0) < \frac{\varepsilon}{2}.$$

Thus for any D_i we have

$$\begin{aligned} m(TD_i \triangle S^{-1}T_0SD_i) &\leq m(TD_i \triangle QD_i) + m(QD_i \triangle S^{-1}T_0SD_i) \\ &\leq \frac{\varepsilon}{2} + \tilde{d}(Q, S^{-1}T_0S). \end{aligned}$$

But $\tilde{d} \leq d'$, so

$$m(TD_i \triangle S^{-1}T_0SD_i) \leq \frac{\varepsilon}{2} + d'(Q, S^{-1}T_0S) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Lemma 1.5 is incredibly useful because it says, in essence, that for any property which is isomorphism invariant (and is not mutually exclusive with antiperiodic transformations), the set of transformations with that property is dense in $\mathcal{G}(X)$. Indeed, if we have $S, T \in \mathcal{G}(X)$ and we define $R := S^{-1}TS$, then $SR = TS$. As S is measure-preserving, S is an isomorphism between (X, m, T) and (X, m, R). This becomes useful in proving results akin to our next theorem (Theorem 1.4). First, however, we need the following lemma, which is a well-known characterization of weakly mixing transformations.

Lemma 1.6. *Let $T \in \mathcal{G}(X)$. Then $T \in \mathcal{W}(X)$ if and only if there exists a subsequence (n_k) such that for all $f, g \in L^2(X)$,*

$$\lim_{k\to\infty} |\langle T^{n_k} f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| = 0. \tag{1.1}$$

Proof. Suppose $T \in \mathcal{W}(X)$, let $(f_i)_{i=1}^{\infty}$ be a dense subset of $L^2(X)$, and let $l \mapsto (f_i, f_j), 1 \leq l < \infty$ order the set $\{(f_i, f_j) | i, j \in \mathbb{N}\}$ arbitrarily. By the Koopman-von Neumann lemma (see, e.g., [17, p.54]), for each l there exists a subsequence $n_m'^l$ of upper asymptotic density 1 such that

$$\lim_{k\to\infty} \left|\langle T^{n_m'^l} f_i, f_j\rangle - \langle f_i, 1\rangle\langle 1, f_j\rangle\right| = 0.$$

Stated differently, for the pair f_i, f_j corresponding to l, we have the desired convergence along the dense subsequence $(n_m'^l)$. We now define a new density 1 subsequence for each l inductively. Let $(n_m^1) := (n_m'^1)$, and for $l > 1$, define (n_m^l) to be a common density 1 subsequence of (n_m^{l-1}) and $(n_m'^l)$. Now define n_k be a diagonal sequence obtained from (n_m^l), i.e., $n_k := n_k^k$. We now show that for all f_i, f_j, (1.1) holds along (n_k). Indeed, fix $\varepsilon > 0, i, j \in \mathbb{N}$ and let l correspond to i, j. There exists N such that for $m \geq N$,

$$\left|\langle T^{n_m^l} f_i, f_j\rangle - \langle f_i, 1\rangle\langle 1, f_j\rangle\right| < \varepsilon.$$

Now let $K := \max\{l, N\}$. Now for $k \geq K$, there exists m_0 such that $n_k = n_{m_0}^l$ and $m_0 \geq N$. Therefore, for $k \geq K$,

$$|\langle T^{n_k} f_i, f_j\rangle - \langle f_i, 1\rangle\langle 1, f_j\rangle| < \varepsilon.$$

To see n_k satisfies (1.1) for all $h, g \in L^2(X)$, let $h_j \to h, g_j \to g$, with $h_j, g_j \in (f_i)$ for all j (without loss of generality, h_j, g_j, h, g all have norm 1), and let $\varepsilon > 0$. Then

$$\begin{aligned}
|\langle T^{n_k} h, g\rangle - \langle h, 1\rangle\langle 1, g\rangle| \leq\ & |\langle T^{n_k} h, g\rangle - \langle T^{n_k} h_j, g\rangle| \\
&+ |\langle T^{n_k} h_j, g\rangle - \langle T^{n_k} h_j, g_j\rangle| \\
&+ |\langle T^{n_k} h_j, g_j\rangle - \langle h_j, 1\rangle\langle 1, g_j\rangle| \\
&+ |\langle h_j, 1\rangle\langle 1, g_j\rangle - \langle h_j, 1\rangle\langle 1, g\rangle| \\
&+ |\langle h_j, 1\rangle\langle 1, g\rangle - \langle h, 1\rangle\langle 1, g\rangle|.
\end{aligned}$$

For the third line, we can find a K such that for $k \geq K$ (and uniformly in j),

$$|\langle T^{n_k} h_j, g_j\rangle - \langle h_j, 1\rangle\langle 1, g_j\rangle| < \varepsilon/5.$$

For the other four lines we have, for example from the last line, from the Cauchy-Schwarz Inequality and by the fact that $\|g\| = 1$, that

$$|\langle h_j, 1\rangle\langle 1, g\rangle - \langle h, 1\rangle\langle 1, g\rangle| = |\langle h_j - h, 1\rangle\langle 1, g\rangle| \leq \|h_j - h\|.$$

And as $h_j \to h$ in $L^2(X)$, there exists J_1 such that for $j \geq J_1$,

$$|\langle h_j, 1\rangle\langle 1, g\rangle - \langle h, 1\rangle\langle 1, g\rangle| < \frac{\varepsilon}{5}.$$

We can similarly find such J_i for the other 3 lines, and thus for $k \geq K$, we get

$$|\langle T^{n_k} h, g\rangle - \langle h, 1\rangle\langle 1, g\rangle| < \varepsilon.$$

To prove the converse, suppose $T \notin \mathcal{W}(X)$. Then there exists $\lambda \neq 1$ and $f \in L^2(X)$ (f not constant, and without loss of generality, $\|f\| = 1$) such that $Tf = \lambda f$. Note that $\langle f, 1\rangle = 0$ as $f, 1$ are eigenvalues of different eigenfunctions. Then for all n,

$$|\langle T^n f, f\rangle - \langle f, 1\rangle\langle 1, f\rangle| = |\lambda^n \langle f, f\rangle| = 1.$$

Thus for the pair f, f and any subsequence (n_k),

$$\lim_{k\to\infty} |\langle T^{n_k} f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| = 1 \neq 0.$$

□

Theorem 1.4. *The set $\mathcal{W}(X)$ is a dense, G_δ set in the weak topology on $\mathcal{G}(X)$.*

Proof. Let $T \in \mathcal{W}(X)$ and $S \in \mathcal{G}(X)$. Further, by Lemma 1.6, there exists a subsequence (n_k) such that for all $f, g \in L^2(X)$,

$$\lim_{k\to\infty} |\langle T^{n_k} h, g\rangle - \langle h, 1\rangle\langle 1, g\rangle| = 0.$$

Let $\varepsilon > 0$. Then

$$\begin{aligned} &\left|\langle (S^{-1}TS)^{n_k} f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle\right| \\ =&\left|\langle S^{-1}T^{n_k} Sf, g\rangle - \langle f, 1\rangle\langle 1, g\rangle\right| \\ =&|\langle T^{n_k} Sf, Sg\rangle - \langle Sf, 1\rangle\langle 1, Sg\rangle|, \end{aligned}$$

with the last line holding because S is unitary, and $S1 = 1$. Thus for appropriate K we have that for $k \geq K$

$$\left|\langle (S^{-1}TS)^{n_k} f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle\right| < \varepsilon.$$

Thus $\{S^{-1}TS | S \in \mathcal{G}(X)\} \subset \mathcal{W}(X)$. So by Lemma 1.5, $\mathcal{W}(X)$ is dense in $\mathcal{G}(X)$.

To show $\mathcal{W}(X)$ is G_δ, consider the sets

$$A_{i,j,k,n} := \left\{T \in \mathcal{G}(X) \,|\, |\langle T^n f_i, f_j\rangle - \langle f_i, 1\rangle\langle 1, f_j\rangle| < \frac{1}{k}\right\}.$$

We see that from Lemma 1.6,

$$\mathcal{W}(X) = \bigcup_n \bigcap_{i,j,k} A_{i,j,k,n}.$$

As $A_{i,j,k,n}$ is open in the weak operator topology (and thus the weak topology on $\mathcal{G}(X)$), $\mathcal{W}(X)$ is G_δ. □

Having finally shown that weakly mixing transformations are generic, we turn our attention to strongly mixing transformations. By comparison, this is a much more simple task.

Theorem 1.5. *The set $\mathcal{S}(X)$ is a first category set in the weak topology on $\mathcal{G}(X)$.*

Proof. Let $P'_k := \left\{T \in \mathcal{G}(X) | T^k = \mathbb{1}_X\right\}$ and note that for all n,

$$P_n := \bigcup_{k>n} P'_k$$

is dense in $\mathcal{G}(X)$. This holds because dyadic permutations (of sufficiently large ranks) are always contained in P_n and by Theorem 1.1, dyadic permutations are dense in $\mathcal{G}(X)$.

Now let $A := [0, 1/2] \subset X$ and define new sets

$$M'_k := \left\{T \in \mathcal{G}(X) | \left|\langle T^k \chi_A, \chi_A\rangle - \langle \chi_A, 1\rangle\langle 1, \chi_A\rangle\right| \leq \frac{1}{5}\right\}.$$

Further define

$$M_n := \bigcap_{k>n} M'_k$$

and

$$\bigcup_{n=1}^{\infty} M_n.$$

We immediately see that $\mathcal{S}(X) \subset M$. Indeed, if $T \in \mathcal{S}(X)$, given $\varepsilon > 0$ and $f, g \in L^2(X)$ there exists N such for $n > N$,

$$|\langle T^n f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| < \varepsilon,$$

so in particular it must hold for $f = g = \chi_A$ and $\varepsilon = 1/5$. Thus, to show $\mathcal{S}(X)$ is of first category, it is sufficient to show M is of first category. It is therefore sufficient to show that M_n is nowhere dense for all n. Thus it is even further sufficient to show

$$M_n^c = \bigcup_{k>n} (M'_k)^c$$

is dense for all n, as M_n is closed for all n. Finally, it is thus sufficient to show that $P'_k \subset (M'_k)^c$ because, as we previously noted, P_n is dense for all n and if $P'_k \subset (M'_k)^c$, then $P_n \subset M_n^c$ for all n. So let $T \in P'_k$. Then

$$\left|\langle T^k \chi_A, \chi_A\rangle - \langle \chi_A, 1\rangle\langle 1, \chi_A\rangle\right| = \left|\langle \chi_A, \chi_A\rangle - \frac{1}{4}\right| = \left|\frac{1}{2} - \frac{1}{4}\right| > \frac{1}{5}.$$

So $T \in (M'_k)^c$ □

1.2 Examples of Weakly But Not Strongly Mixing Transformations

As Theorems 1.4 and 1.5 show that $\mathcal{W}(X)$ is much larger than $\mathcal{S}(X)$, one might expect that it would be easy to find $T \in \mathcal{W}(X) \backslash \mathcal{S}(X)$. However, this proved to be quite the challenge. In fact, an example wasn't found until 1969 – two decades after Rokhlin proved Theorem 1.5 in [44] – that Chacon constructed an example in [10]. Chacon's method for constructing this transformation is known as cutting-and-stacking. Before we describe the method, we give a definition.

Definition 1.3. An ordered set of (up to endpoints) disjoint subintervals $\mathcal{C} = \{I_1, \ldots, I_N\}$ of $[0, 1]$ is called a *column* if $m(I_i) = m(I_j)$ for all i, j. The number N is called the *height* of the column $\mathcal{C}$. We say $x \in \mathcal{C}$ if there exists I_i such that $x \in I_i$, and we call that particular I_i the *level* of x. We further say $T \in \mathcal{G}(X)$ *maps along the column* $\mathcal{C}$ if for $1 \leq i \leq N-1, TI_i = I_{i+1}$.

The name column is chosen for the visual intuition – we imagine a column of intervals of equal length stacked on top of each other. T maps along the column if it moves each interval up one level until we reach the top, where anything can happen.

If we have two disjoint columns of the same width, then we should be able to create a new column from them.

Definition 1.4. Let $\mathcal{C}_1 = \{I_1, \ldots, I_{N_1}\}, \mathcal{C}_2 = \{I'_1, \ldots, I'_{N_2}\}$ be two columns such that for $1 \leq i \leq N_1, 1 \leq j \leq N_2, m(I_i) = m(I'_j)$ and $m(I_i \cap I'_j) = 0$. We create a new column $\mathcal{C} = \{I''_1, \ldots I''_{N_1+N_2}\}$ where $I''_i = I_i$ for $1 \leq i \leq N_1$ and $I''_{N_1+j} = I'_j$ for $1 \leq j \leq N_2$. We call this process *stacking* $\mathcal{C}_2$ onto $\mathcal{C}_1$.

The basic idea of the cutting-and-stacking method is rather simple: We begin with a column, and then divide (or "cut") the column into some number of equal width subcolumns. We then add "spacer" sets on top of some of the subcolumns. We then stack the subcolumns (including spacer sets) from right to left. We then repeat this process. At each stage of the process we consider a transformation T_n which maps along the column. The limit of this process will yield a measure-preserving transformation T. We now describe the process precisely for the transformation which Chacon constructed.

1. (Base case) Let $\mathcal{C}_0 = (0, 2/3)$ be a column of height 1 (i.e. $h_0 = 1$).

2. (Defining $\mathcal{C}_1$). Cut $\mathcal{C}_0$ into three equal width subcolumns. In particular, we denote $\mathcal{C}_0[0] := \{(0, 2/9)\}, \mathcal{C}_0[1] = \{(2/9, 4/9)\}, \mathcal{C}[2] = \{(4/9, 2/3)\}$. We now add a spacer set on top of $\mathcal{C}[1]$, which is an interval of the same length as the intervals in the subcolumns, which is disjoint from $\mathcal{C}_0$. In this case, we take $S_1 = (2/3, 8/9)$. We then create $\mathcal{C}_1$ by stacking $\mathcal{C}_0[2]$ onto $\mathcal{C}_0[1] \cup S_1$, and that resulting column onto $\mathcal{C}_0[0]$. Note that $\mathcal{C}_1$ is a column of height $4 = 3h_0 + 1$. Let T_1 be a transformation which maps along $\mathcal{C}_1$.

3. (Defining $\mathcal{C}_{n+1}$) Suppose we have defined column $\mathcal{C}_n$. Note that $\mathcal{C}_n$ has a height

$$h_n = 3h_{n-1} + 1 = \sum_{i=0}^{n} 3^i = \frac{3^{n+1} - 1}{2},$$

and the width of the column $\mathcal{C}_n$ is $2/3^{n+1}$. Now cut $\mathcal{C}_n$ into three subcolumns of width $2/3^{n+2}$, which we denote (from left to right, respectively), $\mathcal{C}_n[0], \mathcal{C}_n[1], \mathcal{C}_n[2]$. Let S_{n+1} be the initial interval of length $2/3^{n+2}$ in the complement of $\mathcal{C}_n$, and add it as a spacer set to the top of $\mathcal{C}_n[2]$. Then create $\mathcal{C}_{n+1}$ by stacking $\mathcal{C}_n[2]$ onto $\mathcal{C}_n[1] \cup S_{n+1}$, and then this stacked column onto $\mathcal{C}_n[0]$. Let T_{n+1} map along the column $\mathcal{C}_{n+1}$.

Figure 1.1: Cutting $\mathcal{C}_1$ and adding spacer set S_2 to form $\mathcal{C}_2$

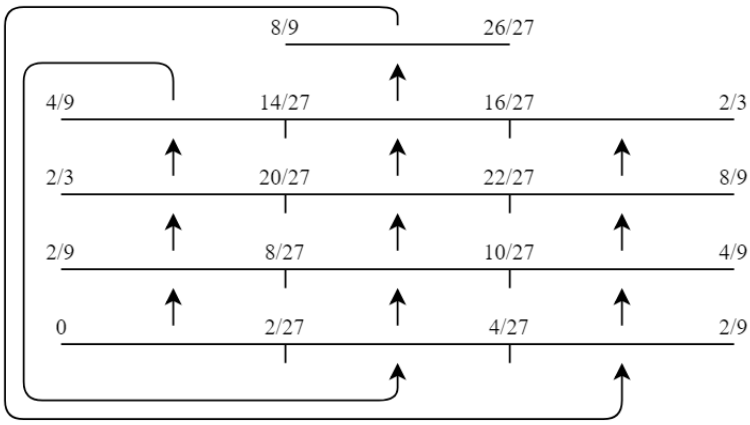


Figure 1.1 shows the process of defining $\mathcal{C}_2$. The 4 intervals of length 2/9 form the column $\mathcal{C}_1$, which is then cut into three equal width subcolumns. The interval $(8/9, 26/27)$ is the spacer set $\mathcal{S}_2$.

Notice that for any n and for any $k \geq n$, if $x \in \mathcal{C}_n \backslash \{I_{h_n}\}$, then $T_k x = T_n x$. Thus the weak limit of T_n exists. Denote it by T. We call T the (canonical) Chacon transformation. We will not fully show that T is weakly but not strongly mixing, but we will give a few notes to give an idea of how it is shown. The key to showing T is not strongly mixing is the following property.

Lemma 1.7. *Let T be the Chacon transformation, let $n \in \mathbb{N}$ be fixed, and let I be a level in $\mathcal{C}_n$. Then for all $k > n$,*

$$m(T^{h_{k-1}} I \cap I) \geq \frac{1}{3} m(I).$$

Lemma 1.7 essentially says that if we take a level of a column, then the level meets itself in a large portion infinitely often. Thus T clearly cannot be strongly mixing.

To show T is weakly mixing, one must first show T is ergodic, a simple task by comparison.[1] We can then take an eigenfunction f for T corresponding to an eigenvalue λ, where $|f| \equiv 1, |\lambda| = 1$. Now we make an observation. If I is a level in $\mathcal{C}_n$, then there are three levels in $\mathcal{C}_{n+1}$ which are all contained in I. Denote these three levels by $I[0], I[1], I[2]$, where $I[0]$ is the lowest level of the three, and $I[2]$ is the highest. We observe that

$$T^{h_n} I[0] = I[1], \quad T^{h_n} I[1] = T^{-1} I[2].$$

[1]In fact, it should be noted that any transformation which is constructed by a cutting-and-stacking method (a so-called *rank one transformations*) can be shown to be ergodic in the same way.

Using this observation, we are able to conclude that there exists x such that

$$\begin{aligned}\left|\lambda^{h_n} f(x) - f(x)\right| &< 2\varepsilon,\\ \left|\lambda^{2h_n+1} f(x) - f(x)\right| &< 2\varepsilon.\end{aligned}$$

But as $|f| \equiv 1$, we see $\lambda = 1$, so T has continuous spectrum.

While the Chacon transformation was the first example of a weakly but not strongly mixing transformation, it is not the only known example. We will briefly discuss some other examples.

Let $\lambda := (\lambda_1, \ldots, \lambda_n) \in (0,1)^n$ be such that $\sum_{i=1}^{n} \lambda_i = 1$. Further, let σ be a permutation on n objects. Further, for each i, define

$$a_i := \sum_{1 \le j < i} \lambda_j, \quad a_i' := \sum_{1 \le j < \sigma(i)} \lambda_{\sigma^{-1}(j)}.$$

We now define $T := T_{\sigma,\lambda}$ as follows: For $x \in (a_i, a_{i+1})$ (for $1 \le i < n$), define $Tx := x - a_i + a_i'$. We call T the *interval exchange transformation* associated with σ, λ.

Intuitively, an interval exchange transformation "permutes" a finite number of subintervals of varying lengths. More accurately, an interval exchange transformation takes a partition of X into intervals and maps the elements of this partition by translation to a new partition, according to the permutation σ.

As an example, suppose we have

$$\lambda = \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right),$$

and $\sigma = (1\ 3\ 2)$ (i.e., $\sigma(1) = 3, \sigma(3) = 2, \sigma(2) = 1$). Then we have

$$\begin{aligned} a_1 &= 0, \quad a_1' = \frac{5}{6},\\ a_2 &= \frac{1}{6}, \quad a_2' = 0\\ a_3 &= \frac{1}{2}, \quad a_3' = \frac{1}{3}. \end{aligned}$$

So we have

$$T\left(0, \frac{1}{6}\right) = \left(\frac{5}{6}, 1\right), \quad T\left(\frac{1}{6}, \frac{1}{2}\right) = \left(0, \frac{1}{3}\right), \quad T\left(\frac{1}{2}, 1\right) = \left(\frac{1}{3}, \frac{5}{6}\right).$$

In [35], Katok showed that interval exchange transformations are not strongly mixing for any choice of λ and σ. It is, however, not the case that $T_{\sigma,\lambda}$ is weakly mixing for all choices of σ, λ. For example, if $\lambda_i = \lambda_j$ for all i, j, then $T_{\sigma,\lambda}$ is a permutation of n intervals of length $1/n$, and is subsequently not even ergodic. However, these cases are somewhat negligible. Katok and Stepin [36] and Veech [52] made some progress before Avila and Forni [8] showed that for almost every $\lambda, T_{\sigma,\lambda}$ is weakly mixing. It should be

noted that interval exchange transformations remain component of active research today. See, e.g., [40], [14], [21], [22] [11], [16], [53], [38].

Our last example is due to Kakutani, though the transformation itseelf was originally constructed by von Neumann. We begin by defining intervals I_n. Let $I_0 := (0, 1/2)$, and inductively define I_n to be the interval of length $1/2^{n+1}$ whose left endpoint is the right endpoint of I_{n-1}. More explicitly, we can define

$$I_n := \left(1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right).$$

We can now define $T \in \mathcal{G}(X)$ as follows. For $x \in I_n$, define

$$Tx := x - \left(1 - \frac{1}{2^n} - \frac{1}{2^{n+1}}\right).$$

Essentially T maps I_n to $1 - I_n$, though it does so by translation rather than reflection. In [34], Kakutani showed that T is both weakly mixing and not strongly mixing.

1.3 Other Category Results

While Theorems 1.4 and 1.5 were the first category results of their kind, they were far from the last. For years after Halmos and Rokhlin published their results, various category results began popping up. In this section we will discuss just a few of these many results.

First, we recall what it means for a transformation to be rigid.

Definition 1.5. We say that $T \in \mathcal{G}(X)$ is *rigid* if there exists a subsequence (n_k) such that T^{n_k} converges to $\mathbb{1}_X$ strongly. That is, if for all $f \in L^2(X)$,

$$\lim_{k\to\infty} \|T^{n_k} f - f\| = 0.$$

The sequence n_k is called a *rigidity sequence* for T

Let $\mathcal{R}(X) \subset \mathcal{G}(X)$ denote the set of rigid transformations on X. Though a triviality, we make note of the following equivalence for later.

Proposition 1.4. *A measure-preserving transformation $T \in \mathcal{G}(X)$ is rigid if and only if there exists a subsequence (n_k) such that for all $f, g \in L^2(X)$,*

$$\lim_{k\to\infty} |\langle T^{n_k} f, g\rangle - \langle f, g\rangle| = 0. \tag{1.2}$$

Proof. Equation (1.2) says that $T^{n_k} \to \mathbb{1}_X$ weakly, so one direction is trivial. For the other direction, suppose there is a subsequence (n_k) such that (1.2) holds. As T is unitary, clearly $\|T^{n_k} f\| \to \|\mathbb{1}_X f\|$ for all f, so $T^{n_k} \to \mathbb{1}_X$ strongly. □

In [36], Katok and Stepin showed that $\mathcal{R}(X)$ is a dense, G_δ set. We provide a proof of this below, as it will be relevant in Chapter 2

Theorem 1.6 (Katok, Stepin)**.** *The set $\mathcal{R}(X)$ is a dense, G_δ subset of $\mathcal{G}(X)$.*

Proof. Note first of all that there exists a rigid transformation which is antiperiodic. Take, for example, a rotation by an irrational number λ. It is ergodic, and hence antiperiodic. On the other hand, there exists a subsequence (n_k) such that $e^{2\pi i n_k \lambda} \to 1$ as $k \to \infty$, so (n_k) is a rigidity sequence for this transformation.

Let $T \in \mathcal{R}(X), S \in \mathcal{G}(X)$, and (n_k) be a rigidity sequence for T. Then for all $f \in L^2(X)$ we have

$$\begin{aligned} &\left\|(S^{-1}TS)^{n_k} f - f\right\| \\ =&\left\|S^{-1}T^{n_k}Sf - f\right\| \\ \leq&\left\|S^{-1}\right\| \left\|T^{n_k}(Sf) - Sf\right\| \\ =&\left\|T^{n_k}(Sf) - Sf\right\|. \end{aligned}$$

As (n_k) is a rigidity sequence for T, this goes to 0 as $k \to \infty$, so $S^{-1}TS$ is rigid. So by Lemma 1.5, $\mathcal{R}(X)$ is dense in $\mathcal{G}(X)$.

Now let (f_i) be a countable, dense subset of $L^2(X)$ and consider the sets

$$R_{i,k,n} := \left\{T \in \mathcal{G}(X) | \left\|T^n f_i - f_i\right\| < \frac{1}{k}\right\}.$$

We see that $R_{i,k,n}$ are open in the strong operator topology. Further, we see by definition that

$$\mathcal{R}(X) = \bigcap_{i,k} \bigcup_{n>k} R_{i,k,n}.$$

So $\mathcal{R}(X)$ is G_δ. □

Theorem 1.6 has a couple of interesting consequences. First, combined with Theorem 1.4, it says that a generic transformation $T \in \mathcal{G}(X)$ is both weakly mixing and rigid. Stated another way, it says that there exist subsequences $(n_k), (m_k)$ such that for all $f, g \in L^2(X)$

$$\lim_{k\to\infty} |\langle T^{n_k} f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| = 0, \text{ and } \lim_{k\to\infty} |\langle T^{m_k} f, g\rangle - \langle f, g\rangle| = 0.$$

Further, since the subsequence (n_k) can be assumed to have upper asymptotic density 1 by the Koopman-von Neumann lemma, the rigidity sequence (m_k) must have density zero.

Second, as $\mathcal{R}(X) \subset \mathcal{S}(X)^c$, Theorem 1.5 is a corollary to Theorem 1.6.

Definition 1.6. Let α be an element of the unit circle in $\mathbb{C}$. We say that $T \in \mathcal{G}(X)$ is *α-rigid* if there exists a subsequence (n_k) such that $T^{n_k} \to \alpha \mathbb{1}_X$ strongly. That is, for all $f \in L^2(X)$,

$$\lim_{k\to\infty} \|T^{n_k} f - \alpha f\| = 0.$$

Denote the set of α-rigid transformations by $\mathcal{R}^\alpha(X)$. It is clear from the definition that $\mathcal{R}^1(X) = \mathcal{R}(X)$. However, more can be said. Indeed, for all $\alpha, \mathcal{R}^\alpha(X) \subset \mathcal{R}(X)$. This can be shown in two steps. First, if T is α-rigid, then T is α^n-rigid for all n. This holds because if $T^{n_k} \to \alpha \mathbb{1}_X$, then $T^{n \cdot n_k} \to \alpha^n \mathbb{1}_X$. If α is rational, then we are done. If α is irrational, then there exists a subsequence n_j such that $\alpha^{n_j} \to 1$, and then taking a diagonal subsequence of the α^n-rigidity sequences will produce a rigidity sequence for T. Indeed, the same proof shows that if T is α-rigid for α irrational, then T is β-rigid for all β in the unit circle of $\mathbb{C}$.

The following result can be found in [39].

Proposition 1.5. *Let α be an element of the unit circle of $\mathbb{C}$. Then $\mathcal{R}^\alpha(X)$ is a dense, G_δ subset of $\mathcal{G}(X)$.*

Proof. We first note that there exists $T \in \mathcal{R}^\alpha(X)$ which is antiperiodic. Consider T a rotation on the circle by some irrational λ. It is ergodic, and hence antiperiodic. On the other hand, there exists a subsequence (n_k) such that $\lambda^{n_k} \to \alpha$ as $k \to \infty$, so (n_k) is an α-rigidity sequence for this transformation.

Let $T \in \mathcal{R}^\alpha(X), S \in \mathcal{G}(X)$, and (n_k) be an α-rigidity sequence for T. Then for all $f \in L^2(X)$,

$$\begin{aligned} &\left\|(S^{-1}TS)^{n_k} f - \alpha f\right\| \\ =& \left\|S^{-1}T^{n_k}Sf - \alpha f\right\| \\ \leq& \left\|S^{-1}\right\| \left\|T^{n_k}(Sf) - S(\alpha f)\right\| \\ =& \left\|T^{n_k}(Sf) - \alpha Sf\right\|. \end{aligned}$$

As (n_k) is an α-rigidity sequence for T, this goes to 0 as $k \to \infty$, so $S^{-1}TS$ is α-rigid. So by Lemma 1.5, $\mathcal{R}^\alpha(X)$ is dense in $\mathcal{G}(X)$.

Now let (f_i) be a countable, dense subset of $L^2(X)$ and consider the sets

$$R_{i,k,n} := \left\{T \in \mathcal{G}(X) | \left\|T^n f_i - \alpha f_i\right\| < \frac{1}{k}\right\}.$$

We see that $R_{i,k,n}$ are open in the strong operator topology. Further, we see by definition that

$$\mathcal{R}^\alpha(X) = \bigcap_{i,k} \bigcup_{n>k} R_{i,k,n}.$$

So $\mathcal{R}^\alpha(X)$ is G_δ. □

The final type of transformation we will discuss is a combination of weakly mixing and rigid transformations.

Definition 1.7 (Stepin)**.** Let $\kappa \in [0,1]$. We say that $T \in \mathcal{G}(X)$ is *κ-weakly mixing* if there exists a subsequence (n_k) such that for all $f, g \in L^2(X)$,

$$\lim_{k\to\infty} |\langle T^{n_k} f, g\rangle - (\kappa\langle f, 1\rangle\langle 1, g\rangle + (1-\kappa)\langle f, g\rangle)| = 0.$$

Denote the set of κ-weakly mixing transformations by $\mathcal{W}^\kappa(X)$. We immediately see that $\mathcal{W}^1(X) = \mathcal{W}(X)$ and $\mathcal{W}^0(X) = \mathcal{R}(X)$. It is further easy to see that κ-weakly mixing transformations can be viewed as an interpolation of weakly mixing and rigid transformations. Interestingly, it turns out that for $\kappa \neq 0, \mathcal{W}^\kappa(X) \subset \mathcal{W}(X)$. We will not show this, but the proof can be found in [48]. However, in that same paper, Stepin showed that $\mathcal{W}^\kappa(X)$ is residual.

Proposition 1.6 (Stepin)**.** *For all $\kappa \in [0,1]$, the set $\mathcal{W}^\kappa(X)$ is a dense, G_δ subset of $\mathcal{G}(X)$.*

Proof. That $\mathcal{W}^\kappa(X)$ is nonempty can be found in [48]. Further, for $\kappa \neq 0, \mathcal{W}^\kappa(X) \subset \mathcal{W}(X)$, so every element of such $\mathcal{W}^\kappa(X)$ is, in particular, antiperiodic. For $\kappa = 0, \mathcal{W}^\kappa(X) = \mathcal{R}(X)$.

Now let κ be fixed, $T \in \mathcal{W}^\kappa(X), S \in \mathcal{G}(X)$, and (n_k) be a subsequence such that for all $f, g \in L^2(X)$,

$$\lim_{k\to\infty} |\langle T^{n_k} f, g\rangle - (\kappa\langle f, 1\rangle\langle 1, g\rangle + (1-\kappa)\langle f, g\rangle)| = 0.$$

Then we have

$$\begin{aligned}
&\left|\langle (S^{-1}TS)^{n_k} f, g\rangle - (\kappa\langle f, 1\rangle\langle 1, g\rangle + (1-\kappa)\langle f, g\rangle)\right| \\
=&\left|\langle S^{-1}T^{n_k}Sf, g\rangle - (\kappa\langle f, 1\rangle\langle 1, g\rangle + (1-\kappa)\langle f, g\rangle)\right| \\
=&|\langle T^{n_k}(Sf), Sg\rangle - (\kappa\langle Sf, S1\rangle\langle S1, Sg\rangle + (1-\kappa)\langle Sf, Sg\rangle)| \\
=&|\langle T^{n_k}(Sf), Sg\rangle - (\kappa\langle Sf, 1\rangle\langle 1, Sg\rangle + (1-\kappa)\langle Sf, Sg\rangle)|
\end{aligned}$$

But by assumption, this goes to 0 as $k \to \infty$. Thus $S^{-1}TS \in \mathcal{W}^\kappa(X)$, and thus $\mathcal{W}^\kappa(X)$ is dense in $\mathcal{G}(X)$.

Now let (f_i) be a countable, dense subset of $L^2(X)$ and consider the sets

$$W^\kappa_{i,j,k,n} = \left\{T \in \mathcal{G}(X) \,|\, |\langle T^n f_i, f_j\rangle - (\kappa\langle f_i, 1\rangle\langle 1, f_j\rangle + (1-\kappa)\langle f_i, f_j\rangle)| < \frac{1}{k}\right\}.$$

We see that $W_{i,j,k,n}$ is open in the weak operator topology. We further see that

$$\mathcal{W}^\kappa(X) = \bigcap_{i,j,k} \bigcup_{n>k} W^\kappa_{i,j,k,n}$$

so $\mathcal{W}^\kappa(X)$ is G_δ. □

Before ending this section, we make a few notes. First, there are many more examples of such category results. For other examples, see, e.g., [39], [1], [37], [2], [5], [3], [15], [4], [47], [29], [18].

Chapter 2

Genericity for Extensions

2.1 Non-Fixed Factor

In 2014, Host, Kra, and Maass attempted to find examples of systems which have a non-trivial nilfactor [33]. To their surprise, they found it difficult to find such examples, finding that for most known classes of systems, the only nilfactor was the Kronecker factor, and found a sufficient condition for this property to occur. This lead Tao [50] to pose the question: is a typical extension of a circle rotation (for example) a weakly mixing extension? If true, then the dichotomy between weakly mixing and isometric extensions would explain the findings of Host, Kra, and Maass. Moreover, Tao's framing of the problem beckons the classical theory of Halmos, which provided a strategy for its solution.

However, the initial attempts at the problem ran into some issues. Certain aspects of Halmos' techniques could not be easily replicated. Most notably was Halmos' use of dyadic permutations (see Section 1.1.2). The natural generalization only made sense if the factor transformation was itself a dyadic permutation. As dyadic permutations represent a small set of all possible factor transformations, this posed a problem.

The idea then arose: what if one considers the set of extensions of a non-fixed factor? To make sense of this, recall Rokhlin's skew product representation theorem (see, for example, [27, p.69]), which says that if (Z, μ, T) is an extension of (X, m, S), then there exists a measure space (Y, ν) and a measurable map $\phi : X \to \mathcal{G}(Y)$ such that (Z, μ, T) is isomorphic to $(X \times Y, m \times \nu, \hat{T})$, where $\hat{T}(x, y) = (Tx, \phi(x)y)$. In other words, all extensions in the measure-preserving case can be represented as skew products over their factors. The idea is then as follows: fix a product space, and consider all measure-preserving transformations on this space which are skew products over a system on the first coordinate. This solves the issue of dyadic permutations, as now all dyadic permutations (which are themselves extensions) can be considered and utilized. This section works in this setting, culminating in Theorem 2.5, which says that the set of weakly mixing extensions is a dense, G_δ set.

The work appearing in this section will appear in *Ergodic Theory and Dynamical Systems*. See [46].

2.1.1 Notation and Preliminaries

Throughout Section 2.1 we will be working with extensions on product spaces through the natural projection. To be more precise, we let (X, m) be a non-atomic standard probability space, (Y, ν) be a probability space, $(Z, \mu) = (X \times Y, m \times \nu)$, and T, T' be measure-preserving transformations on $(Z, \mu), (X, m)$ respectively, such that (Z, μ, T) is an extension of (X, m, T') through the natural projection map $\pi : Z \to X$ onto the first coordinate. We will assume throughout that T, T' are invertible, and will identify two transformations if they differ on a set of measure zero. We will say "T is an extension of T'" or "T extends T'" if T and T' satisfy all conditions stated above. We will assume without loss of generality that X is the unit interval and m is the Lebesgue measure. We can assume this because all non-atomic standard probability spaces are isomorphic (see [30, p. 61]).

Let $\mathcal{G}(Z)$ denote the set of all invertible, measure-preserving transformations on (Z, μ) and let $\mathcal{G}_X = \{T \in \mathcal{G}(Z) : \exists\, T' \in \mathcal{G}(X) \text{ s.t. } T \text{ extends } T'\}$. Note that if we say $T \in \mathcal{G}_X$, we assume that the transformation on the factor will be notated by T'. Further note that we will also write $\mathcal{G}_X$ to denote the corresponding set of Koopman operators. Note that another way we could define $\mathcal{G}_X$ is given a probability space $(Z, \mathcal{B}, \mu)$ and a fixed sub-algebra $\mathcal{F} \subset \mathcal{B}$, we consider all invertible measure-preserving transformations on $(Z, \mathcal{B}, \mu), T$, for which $\mathcal{F}$ is T-invariant, and we denote by T' the action of T on $\mathcal{F}$.

The weak topology on $\mathcal{G}_X$ is defined as the subspace topology of the weak topology on $\mathcal{G}(Z)$. That is, it is given by the subbasic neighborhoods

$$N_\varepsilon(T; E) = \{S \in \mathcal{G}_X : \mu(TE \triangle SE) < \varepsilon\},$$

where $\varepsilon > 0$ and E is some measurable subset of Z. Note that if Z is, say, the unit square with the Lebesgue measure, then it is sufficient for a subbasis to consider only dyadic sets (i.e., a finite union of dyadic squares). Recall from Section 1.1.1 that the weak topology happens to coincide with the weak (and strong) operator topology for the corresponding Koopman operators.

Further recall from Section 1.1.3, the two metrics on $\mathcal{G}(Z)$ defined by

$$\tilde{d}(S, T) := \sup_E \mu(SE \triangle TE)$$
$$d'(S, T) := \mu\left(\{z \in Z : Sz \neq Tz\}\right)$$

where the $\sup$ in the first definition is taken over all measurable sets E. For the important properties of d, d', see Section 1.1.3.

Let $L^2(Z|X)$ denote the Hilbert module over $L^\infty(X)$. More precisely, for $f \in L^2(Z)$,

$$f \in L^2(Z|X) \text{ if and only if } \mathbb{E}(|f|^2 \,|X)^{1/2} \in L^\infty(X),$$

where $\mathbb{E}(f|X)$ is the conditional expectation of f with respect to X. More specifically, it is the conditional expectation with respect to $\mathcal{A} := \{\pi^{-1}(A) : A \in \mathcal{L}\}$, where $\mathcal{L}$ is the Lebesgue sigma algebra on X. Let

$$\|f\|_{L^2(Z|X)} := \mathbb{E}(|f|^2 \,|X)^{1/2}$$

and

$$\langle f, g\rangle_{L^2(Z|X)} := \mathbb{E}(f \cdot \overline{g}|X).$$

For more on $L^2(Z|X)$, see [49]. One important result of $L^2(Z|X)$ that we do wish to emphasize for later is the Cauchy-Schwarz Inequality.

Proposition 2.1. *Let $f, g \in L^2(Z|X)$. Then*

$$\left|\langle f, g\rangle_{L^2(Z|X)}\right| \leq \|f\|_{L^2(Z|X)} \|g\|_{L^2(Z|X)} \text{ a.e.}$$

Next we give a definition for weakly mixing extensions, cf. [49].

Definition 2.1. An extension T is said to be a *weakly mixing extension of* T' or T is *weakly mixing relative to* T' if for all $f, g \in L^2(Z|X)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \|\mathbb{E}(T^n f \cdot \overline{g}|X) - (T')^n \mathbb{E}(f|X) \cdot \mathbb{E}(g|X)\|_{L^2(X)} = 0.$$

For other possible (equivalent) definitions of weakly mixing extensions, see [27, p.192]. We denote by $\mathcal{W}_X \subset \mathcal{G}_X$ the set of weakly mixing extensions on Z.

To close this section, we prove that the Baire Category Theorem is applicable to $\mathcal{G}_X$, and further that $\mathcal{G}_X$ is topologically a small subset of $\mathcal{G}(Z)$.

Proposition 2.2. *Suppose Y has more than one point. Then $\mathcal{G}_X$ is closed and nowhere dense in $\mathcal{G}(Z)$.*

Proof. We first prove that $\mathcal{G}_X$ is closed. Let $T \in \mathcal{G}(Z)\backslash\mathcal{G}_X$. We wish to find a neighborhod of T that is disjoint from $\mathcal{G}_X$. To this end, let $E \subset Z$ be a cylinder set (to be precise, E is of the form $D \times Y$ for some measurable $D \subset X$), such that TE is not a cylinder set, even up to measure zero. Define $M := \mu(E)$. Then for all cylinder sets C with $\mu(C) = M, \mu(TE\triangle C) > 0$.

We claim that indeed $\inf_C \mu(TE\triangle C) > 0$, where the inf is taken over all cylinder sets C with measure exactly M. Suppose to the contrary that $\inf_C \mu(TE\triangle C) = 0$. Let C_n be a sequence of such cylinder sets such that not only $\mu(TE\triangle C_n) \to 0$, but further such that

$$\sum_{n=1}^{\infty} \mu(TE\backslash C_n) < \infty.$$

Define

$$\hat{C} := \bigcup_{n=1}^{\infty} \bigcap_{m\geq n} C_m.$$

We claim that $\mu(TE\triangle\hat{C}) = 0$. As $\hat{C}$ is clearly a cylinder set, we will arrive at a contradiction.

First consider

$$\hat{C}\backslash TE = \left(\bigcup_{n=1}^{\infty} \bigcap_{m\geq n} C_m\right)\backslash TE = \bigcup_{n=1}^{\infty} \bigcap_{m\geq n} (C_m\backslash TE).$$

Now because $\mu(C_m \triangle TE) \to 0, \mu\left(\bigcap_{m\geq n}(C_m \backslash TE)\right) = 0$ for all n. But then

$$\mu(\hat{C}\backslash TE) \leq \sum_{n=1}^{\infty} \mu\left(\bigcap_{m\geq n}(C_m\backslash TE)\right) = 0.$$

On the other hand,

$$TE\backslash\hat{C} = TE\backslash\left(\bigcup_{n=1}^{\infty}\bigcap_{m\geq n} C_m\right) = \bigcap_{n=1}^{\infty}\bigcup_{m\geq n}(TE\backslash C_m).$$

But by assumption, $\sum_{n=1}^{\infty}\mu(TE\backslash C_n) < \infty$, so by the Borel-Cantelli lemma, $\mu(TE\backslash\hat{C}) = 0$.

Now, let $\varepsilon := \inf_C \mu(TE\triangle C)$. We claim that for any $S \in \mathcal{G}_X, S \notin N_\varepsilon(T; E)$. Indeed, SE is (up to a null set) a cylinder set, and $\mu(SE) = M$, so $\mu(TE\triangle SE) \geq \varepsilon$ by definition of ε.

Now because $\mathcal{G}_X$ is closed, in order to prove that it is nowhere dense, it is sufficient to show that $\mathcal{G}(Z)\backslash\mathcal{G}_X$ is dense. Fix $T \in \mathcal{G}_X$, let $\varepsilon > 0$ and let

$$N_\varepsilon(T) = \{S \in \mathcal{G}(Z) : \mu(TE_i\triangle SE_i) < \varepsilon, i = 1, \ldots, n\},$$

where E_i are measurable sets. Now let $A \subset Z$ be a measurable set such that $0 < \mu(A) < \varepsilon$, and A is not a cylinder set. Further let $B \subset Z$ be a cylinder set such that $\mu(A \cap B) = \mu(A \cap B^c)$, and define $A_1 := A \cap B, A_2 := A \cap B^c$.

We now take $S \in \mathcal{G}(Z)$ with the following properties: if $z \in Z\backslash A, Sz := Tz, SA_1 = TA_2$, and $SA_2 = TA_1$. Note that because T is an extension and A is not a cylinder set, $S \notin \mathcal{G}_X$. Further note that $\{z \in Z : Sz \neq Tz\} = A$. Therefore,

$$\sup_E \mu(TE\triangle SE) = d(T, S) \leq d'(T, S) = \mu(A) < \varepsilon.$$

So $S \in N_\varepsilon(T)$. □

By Proposition 2.2, $\mathcal{G}_X$ is a closed subset of a Baire space, so $\mathcal{G}_X$ is itself a Baire space. Further, because $\mathcal{G}_X$ is nowhere dense, the classical Halmos and Rokhlin results can provide no information about $\mathcal{G}_X$.

2.1.2 Discrete Extensions

As stated in Section 2.1.1, throughout this work we will let (X, m) be the unit interval with Lebesgue measure. For this section, let $Z = X \times \{1, \ldots, L\}$, with $L \geq 2$, w be a probability measure on $\{1, \ldots, L\}$, and $w_i := w(i)$ (without loss of generality, $w_i \neq 0$ for all i). Let μ be the product measure of m and w on Z.

In this section we will be exploring some results regarding these discrete extension measure spaces. We begin by showing that such systems can never be weakly mixing extensions.

Proposition 2.3. *Let $(Z,\mu),(X,m)$ be as above. Then $\mathcal{W}_X=\emptyset$.*

Proof. Fix $T\in\mathcal{G}_X$. It suffices to show that there exists an $f\in L^2(Z|X)$ with relative mean zero (that is, $\mathbb{E}(f|X)=0$ $m-$almost everywhere) such that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left(\left|\mathbb{E}(T^nf\cdot\overline{f}|X)\right|^2\right)^{1/2}\neq 0.$$

In particular, we will construct f such that $\mathbb{E}(T^nf\cdot\overline{f}|X)(x)$ can take only a finite number of possible values, none of which are 0. Thus $\left|\mathbb{E}(T^nf\cdot\overline{f}|X)\right|^2(x)$ is always positive,

$$\begin{aligned}&\inf_{x\in X}\frac{1}{N}\sum_{n=0}^{N-1}\left(\left|\mathbb{E}(T^nf\cdot\overline{f}|X)\right|^2\right)^{1/2}(x)\\ \geq&\min\left|\mathbb{E}(T^nf\cdot\overline{f}|X)\right|(x)>0\end{aligned}$$

for all N, and

$$\frac{1}{N}\sum_{n=0}^{N-1}\left(\left|\mathbb{E}(T^nf\cdot\overline{f}|X)\right|^2\right)^{1/2}$$

cannot converge (in $L^2(X)$) to the zero function on X.

Consider $f(x,y)$, where $f(x,i)=1$ for all $x\in X, i=2,\dots,L$, and

$$f(x,1)=\frac{-\sum_{i=2}^{L}w_i}{w_1}$$

for all $x\in X$. We first see that f has relative mean zero when $L\geq 2$. Indeed, for all $x\in X$ we have that,

$$\begin{aligned}\int_Y f(x,y)dw&=w_1\left(\frac{-\sum_{i=2}^{L}w_i}{w_1}\right)+\sum_{i=2}^{L}w_i\\ &=-\sum_{i=2}^{L}w_i+\sum_{i=2}^{L}w_i=0.\end{aligned}$$

Let $\sigma_{n,x}(i)$ be such that $T^n(x,i)=((T')^nx,\sigma_{n,x}(i))$ for all $(x,i)\in Z$. Now,

$$\begin{aligned}&\mathbb{E}(T^nf\cdot\overline{f}|X)(x)\\ =&\sum_{j=1}^{L}w_jf(x,j)f((T')^nx,\sigma_{n,x}(j))\\ =&\sum_{j=1}^{L}w_jf(x,j)f(x,\sigma_{n,x}(j)),\end{aligned}$$

where the last equality is because f is constant on any given level. Thus we see that because T is invertible, $\sigma_{n,x}$ is a permutation on an L element set, and the value of $\mathbb{E}(T^n f \cdot \overline{f}|X)(x)$ is completely determined by the specific permutation $\sigma_{n,x}$. As there are $L!$ permutations of the L levels, there are finitely many possible values of $\mathbb{E}(T^n f \cdot \overline{f}|X)(x)$.

To see $\mathbb{E}(T^n f \cdot \overline{f}|X)(x) \neq 0$, consider 2 cases. In the first case, we have $\sigma_{n,x}(1) = 1$. In this case we have that for $j = 1$,

$$w_j f(x,j) f(x, \sigma_{n,x}(j)) = w_1 f(x,1) f(x,1) > 0,$$

and for all other j,

$$w_j f(x,j) f(x, \sigma_{n,x}(j)) = w_j.$$

Thus we have that every summand of

$$\sum_{j=1}^{L} w_j f(x,j) f(x, \sigma_{n,x}(j))$$

is positive, and thus the sum is positive (in particular, nonzero). So now suppose $\sigma_{n,x}(i) = 1, i \neq 1$. In this case we have

$$\sum_{j=1}^{L} w_j f(x,j) f(x, \sigma_{n,x}(j)) = f(x,1)(w_1 + w_i) + \left(\sum_{j=2}^{L} w_j \right) - w_i.$$

Consider

$$f(x,1)(w_1 + w_i) = \frac{-\sum_{j=2}^{L} w_j}{w_1}(w_1 + w_i) = \left(-\sum_{j=2}^{L} w_j \right) \left(1 + \frac{w_i}{w_1} \right).$$

Note that

$$\left(-\sum_{j=2}^{L} w_j \right) \left(1 + \frac{w_i}{w_1} \right) \leq -\sum_{j=2}^{L} w_j.$$

Thus,

$$\begin{aligned} & f(x,1)(w_1 + w_i) + \left(\sum_{j=2}^{L} w_j \right) - w_i \\ \leq & \left(-\sum_{j=2}^{L} w_j \right) + \left(\sum_{j=2}^{L} w_j \right) - w_i = -w_i < 0. \end{aligned}$$

So $\mathbb{E}(T^n f \cdot \overline{f}|X)(x)$ is always nonzero, and $\left|\mathbb{E}(T^n f \cdot \overline{f}|X)\right|^2 (x)$ is always positive, as desired. $\square$

We make two notes here. First, the proof of Proposition 2.3 never used any assumptions on the factor, (X, m, T'), and thus it will hold when the factor is any probability space, with any measure-preserving transformation on that space. Second, the proof is still valid in the case that Z has countably many levels instead of finitely many.

Corollary 2.1. *Let (Y, w) be a measure space where Y is a countable set and let $(Z, \mu) = (X \times Ym \times w)$. Then $\mathcal{W}_X$ is empty.*

Proof. Note that for almost all z, if z, Tz are on levels k_1, k_2 respectively (where the level of a point $z \in Z$ is defined to be k if $z = (x, k)$ for some $x \in X$), then $w_{k_1} = w_{k_2}$. Indeed, suppose we have a set A such that both A and TA are contained to a single level. That is, there exists $k_1, k_2 \in Y$ such that for all $z \in A$, there exists $x \in X$ such that $z = (x, k_1)$, and similarly for all $z \in TA$, there exists x such that $z = (x, k_2)$. Thus we can write A as $B \times \{k_1\}$ for some $B \subset X$ and $k_1 \in Y$ (note, in particular, that $m(B) > 0$). Subsequently, $TA = T'B \times \{k_2\}$ for some $k_2 \in Y$. But then $\mu(A) = \mu(B \times k_1) = m(B)w_{k_1}$ and $\mu(TA) = \mu(T'B \times k_2) = m(T'B)w_{k_2} = m(B)w_{k_2}$ as T' is measure-preserving. But as T is measure-preserving, we get that $m(B)w_{k_1} = m(B)w_{k_2}$ and thus either $\mu(A) = 0$ or $w_{k_1} = w_{k_2}$.

Thus (Z, μ, T) can be decomposed into countably many invariant subsystems to which we can apply Proposition 2.3. □

Though $\mathcal{W}_X$ is empty on these discrete extension spaces, they are still worth exploring. But before we can proceed, we will henceforth suppose that the probability measure w is the normalized counting measure. That is, $w_i = \frac{1}{L}$ for all i. With this assumption, we extend the notion of dyadic sets and permutations on X to dyadic sets and permutations on Z.

Definition 2.2. If D is a dyadic interval of rank k in X, then a dyadic square of rank k in Z is a set of the form $D \times \{i\}$. A dyadic set in Z is a union of dyadic squares. A dyadic permutation of rank k on Z is a permutation of the dyadic squares of rank k. A *column-preserving (dyadic) permutation* (of rank k) on Z is a dyadic permutation on Z which is an extension of a dyadic permutation on X.

We wish to generalize the fact that dyadic permutations are dense in $\mathcal{G}(X)$ to density of column-preserving permutations in $\mathcal{G}_X$. First we introduce the following notation: we write $A \subset i$ if there exists $A' \subset X$ such that $A = A' \times \{i\}$.

We now move to the main result of this section, which is a generalization of Theorem 1.1.

Theorem 2.1 (Density of column-preserving permutations)**.** *Column-preserving permutations are dense in $\mathcal{G}_X$. More precisely, let $T \in \mathcal{G}_X$. Given $N_\varepsilon(T)$, a dyadic neighborhood of T, there exists $Q \in N_\varepsilon(T)$, a column-preserving permutation.*

Proof. Without loss of generality, assume

$$N_\varepsilon(T) = \{S \in \mathcal{G}_X : \mu(TD_l \triangle SD_l) < \varepsilon, l = 1, \ldots, L(2^n)\},$$

where D_l are every dyadic square of some fixed rank n (note that D_{l_1}, D_{l_2} are disjoint up to boundary points). We can assume this as every dyadic neighborhood contains such a neighborhood.

Let $k \in \{1, \ldots, L\}$, and let $P_k := \{D_i \cap TD_j | D_i \subset k, j = 1, \ldots, L(2^n)\}$. Note that P_k partitions level k. If π is the natural projection onto X, then let $P'_k := \pi P_k = \{\pi E | E \in P_k\}$. P'_k is a partition of X. Let $P' = \{\hat{A}_\lambda | \lambda \in \Lambda\}$ be a common refinement of P'_k for $k = 1, \ldots, L$, and let $P = \{\hat{A}_{\lambda,k} | \lambda \in \Lambda, k = 1, \ldots, L\}$ be a partition of Z obtained by lifting every element of P' to every level ($\hat{A}_{\lambda,k} \subset k$).

Applying a weaker version of Lemma 1.1 (one where we do not care about the value of $|m(E_i) - r_i|$ in the formulation of the lemma) to the partition P', we obtain a partition $\{A_\lambda\}$ of X into dyadic sets so that

$$m(\hat{A}_\lambda \triangle A_\lambda) < \frac{L\varepsilon}{2\,|\Lambda|}.$$

Applying Lemma 1.1 again, we get a partition of X into dyadic sets B_λ so that

$$m((T')^{-1}\hat{A}_\lambda \triangle B_\lambda) < \frac{L\varepsilon}{2\,|\Lambda|}.$$

Note the full strength of Lemma 1.1 guarantees we can select this partition so that $m(A_\lambda) = m(B_\lambda)$ (as $m(\hat{A}_\lambda) = m((T')^{-1}\hat{A}_\lambda)$). We can now lift A_λ, B_λ to sets $A_{\lambda,k}, B_{\lambda,k}$ so that $A_{\lambda,k}, B_{\lambda,k} \subset k$. Note that

$$\mu(\hat{A}_{\lambda,k} \triangle A_{\lambda,k}) < \frac{\varepsilon}{2\,|\Lambda|}, \text{ and } \mu(T^{-1}\hat{A}_{\lambda,k_2} \triangle B_{\lambda,k_1}) < \frac{\varepsilon}{2\,|\Lambda|}, \tag{2.1}$$

where k_1, k_2 are such that if i, j are such that $\hat{A}_{\lambda,k_2} \subset D_i \cap TD_j$, then $D_i \subset k_2, D_j \subset k_1$.

We will now define Q of some rank $r \in \mathbb{N}$ where r is at least as large as the ranks of D_i for every i, and A_λ, B_λ for every λ. We first define Q' a dyadic permutation on X as any dyadic permutation which maps B_λ to A_λ for every λ. Next we define Q, a column preserving permutation of rank r. First let k_1, k_2 be as before: if i, j are such that $\hat{A}_{\lambda,k_2} \subset D_i \cap TD_j$, then $D_i \subset k_2, D_j \subset k_1$. Then Q will be the extension of Q' such that $B_{\lambda,k_1} \mapsto A_{\lambda,k_2}$. Note that for all λ, k, we have that level $Q^{-1}A_{\lambda,k} =$ level $T^{-1}\hat{A}_{\lambda,k}$, where level $A := k$ if and only if $A \subset k$.

We now show that $\mu(TD_j \triangle QD_j) < \varepsilon$ for all j. Fix $j \in 1, \ldots, L(2^n)$, and define k so that $D_j \subset k$. Let $\Lambda_j := \{\lambda \in \Lambda | (T')^{-1}\hat{A}_\lambda \subset \pi D_j\}$. For $\lambda \in \Lambda_j$, let $i_{\lambda,j}$ be such that $T^{-1}\hat{A}_{\lambda,i_{\lambda,j}} \subset D_j$. Then

$$D_j = \bigcup_{\lambda \in \Lambda_j} T^{-1}\hat{A}_{\lambda,i_{\lambda,j}}.$$

Further, by the definitions of Q and Λ_j, as well as the previous note, we have that $Q^{-1}A_{\lambda,i_{\lambda,j}} = B_{\lambda,k}$.

Note all unions and sums will be taken over $\lambda \in \Lambda_j$. We have

$$\mu\left(D_j \triangle \bigcup B_{\lambda,k}\right) = \mu\left(\bigcup T^{-1}\hat{A}_{\lambda,i_{\lambda,j}} \triangle \bigcup B_{\lambda,k}\right) \leq \sum \mu(T^{-1}\hat{A}_{\lambda,i_{\lambda,j}} \triangle B_{\lambda,k}). \tag{2.2}$$

But by (2.1), $\mu(T^{-1}\hat{A}_{\lambda,i_{\lambda,j}} \triangle B_{\lambda,k}) < \dfrac{\varepsilon}{2\,|\Lambda|}$ so (2.2) $< \sum \dfrac{\varepsilon}{2\,|\Lambda|} = \dfrac{\varepsilon}{2}$. Therefore

$$\mu\left(QD_j \triangle \bigcup A_{\lambda,i_{\lambda,j}}\right) = \mu\left(D_j \triangle \bigcup B_{\lambda,k}\right) < \frac{\varepsilon}{2}.$$

On the other hand,

$$\mu\left(\bigcup A_{\lambda,i_{\lambda,j}} \triangle TD_j\right) = \mu\left(\bigcup A_{\lambda,i_{\lambda,j}} \triangle \bigcup \hat{A}_{\lambda,i_{\lambda,j}}\right) \leq \sum (A_{\lambda,i_{\lambda,j}} \triangle \hat{A}_{\lambda,i_{\lambda,j}}). \tag{2.3}$$

Again by (2.1), $u(A_{\lambda,i_{\lambda,j}} \triangle \hat{A}_{\lambda,i_{\lambda,j}}) < \dfrac{\varepsilon}{2\,|\Lambda|}$ so (2.3) $< \sum \dfrac{\varepsilon}{2\,|\Lambda|} = \dfrac{\varepsilon}{2}$. Finally,

$$\mu(TD_j \triangle QD_j) \leq \mu\left(TD_j \triangle \bigcup A_{\lambda,i_{\lambda,j}}\right) + \left(\bigcup A_{\lambda,i_{\lambda,j}} \triangle QD_j\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

And because this holds for all j, we have that $Q \in N_\varepsilon(T)$. □

2.1.3 Weak Approximation Theorem for Extensions on the Unit Square

Now we let (Z, m_2) be $X \times X$ with the Lebesgue measure. If we need further clarity, we will write the Lebesgue measure on X as m_1, but in general we will denote both Lebesgue measures by m.

We begin by drawing some connections to Section 2.1.2. First, however, we need some more notation. For $L \in \mathbb{N}$, define

$$Z_L := \bigcup_{j=0}^{L-1} \left(X \times \left\{\frac{j}{L}\right\}\right) \subset Z.$$

Also define μ_L, a measure on Z_L, to be the product of the Lebesgue measure with a normalized counting measure on L points. Further, $\pi_L : Z \to Z_L$ be the natural projection onto Z_L. That is, if $z = \left(x, \dfrac{j}{L} + \gamma\right)$ for $\gamma \in \left[0, \dfrac{1}{L}\right)$, then $\pi_L(z) = \left(x, \dfrac{j}{L}\right)$.

Definition 2.3. Let $T \in \mathcal{G}(Z)$. We say that T is *discrete equivalent* if there exists L and $T_L \in \mathcal{G}(Z_L)$, such that (Z, m, T) is an extension of (Z_L, μ_L, T_L) through the factor map π_L. Further, we say that T is *simply discrete equivalent* if T is an identity extension. That is, if we write Z as $Z_L \times \left[0, \dfrac{1}{L}\right)$, then $T = T_L \times \mathbb{1}_Y$. If we wish to emphasize the number of levels, L, we will say T is *L-(simply) discrete equivalent*.

Definition 2.3 is fairly easy to visualize. We take the square and divide it into L equal measure horizontal pieces. Then T is discrete equivalent if T moves fibers on each small piece to other such fibers, and is simply discrete equivalent if it does not move any points within the fiber. Note that in general, a discrete equivalent T need not be in $\mathcal{G}_X$. However, if $T \in \mathcal{G}_X$, then the corresponding T_L is also an extension of T'.

Our goal for this section is to provide a version of Halmos' Weak Approximation Theorem (Theorem 1.2) when restricted to $\mathcal{G}_X$. Mostly this will mean proving a result equivalent to Theorem 2.1. However we first need to lay some ground work. Definitions of dyadic squares, sets, and permutations are all standard in this case, so we do not re-define them. Column-preserving permutations are defined just as they are in Definition 2.2.

Before moving on, we make a few remarks.

Remark 2.1. Lemma 1.1 holds on (Z, m), because (Z, m), like (X, m), is a non-atomic standard probability space, and Lemma 1.1 holds for all such spaces (replacing "dyadic sets" in the statement of Lemma 1.1 with a class $\mathcal{B}$ which is isomorphic to the class of dyadic sets). Alternatively, one can simply prove Lemma 1.1 again in the context of the square. No part of the proof relies on the the fact that we were working on the unit interval, so nothing changes in the proof.

Remark 2.2. We will make the following notational convenience. If $T \in \mathcal{G}(Z)$ and $S' \in \mathcal{G}(X)$, then we will write $S'T$ in place of $(S' \times \mathbb{1}_Y)T$. As S' can be viewed as a transformation on Z itself in this way, there is nothing wrong with doing this, though it may well look strange at first.

Remark 2.3. If $Q \in \mathcal{G}(Z)$ is a dyadic permutation of rank K, then Q is L-simply discrete equivalent with $L = 2^K$. Further if $S' \in \mathcal{G}(X)$, then $S'Q$ is also L-simply discrete equivalent. If Q is further an extension of $Q' \in \mathcal{G}(X)$, then $S'Q$ is an extension of $S'Q'$.

We have some work to do before we can prove our generalization of Theorem 1.1. The idea is as follows: suppose we are applying the same strategy as in the proof of Theorem 1.1. We have found dyadic partitions $\{E_{ij}\}$ and $\{F_{ij}\}$ such that E_{ij} is close to D_{ij}, and F_{ij} is close to $T^{-1}D_{ij}$. Now suppose for a particular i, j we have that $E_{ij} = X \times [0, 1/4]$ and $F_{ij} = [0, 1/4] \times Y$. Clearly there is no way to map $F_{ij} \to E_{ij}$ that result in a column-preserving dyadic permutation. We need to better control the dyadic partitions in order to create a column-preserving permutation.

Lemma 2.1. *Let $\{E_1, \dots E_N\}$ be a finite partition of Z, $\varepsilon > 0$, and suppose $\{\tilde{F}_1, \dots, \tilde{F}_N\}$ is another partition of Z, where $\tilde{F}_i$ are all dyadic sets, and $m(E_i \triangle \tilde{F}_i) < \varepsilon$. Let $K :=$ max rank $\tilde{F}_i$ and let $E_{ij} := E_i \cap \pi^{-1}C_j, j \in \{1, \dots, 2^K\}$, where C_j is a dyadic interval of rank K. Let r_{ij} be dyadic rationals (possibly zero) such that*

$$\sum_{i=1}^{N} r_{ij} = \frac{1}{2^K}$$

for all j and

$$|m(E_{ij}) - r_{ij}| < \frac{\varepsilon}{2^K}$$

for all i, j. Then there exists $\{F_1, \dots, F_N\}$ a partition of Z such that F_i is dyadic set for all $i, m(F_{ij}) = r_{ij}$ (with F_{ij} similarly defined as E_{ij}) for all i, j, and $m(E_i \triangle F_i) < 3\varepsilon$ for all i.

Proof. Similar to the definition of E_{ij}, define $\tilde{F}_{ij} := \tilde{F}_i \cap \pi^{-1}C_j$. Note that by choice of $K, \tilde{F}_{ij}$ is of the product of C_j and a dyadic set for all i, j. Now fix j, and for all $\tilde{F}_{ij}$ with $m(\tilde{F}_{ij}) > r_{ij}$, let $A_{ij} \subset \tilde{F}_{ij}$ of the form $A_{ij} = C_j \times B_{ij}$ with B_{ij} a dyadic set, and $m(A_{ij}) = m(\tilde{F}_{ij}) - r_{ij}$. Define $F_{ij} := \tilde{F}_{ij} \backslash A_{ij}$. Let A be the union of all A_{ij} chosen up to this point. Now for $\tilde{F}_{ij}$ with $m(\tilde{F}_{ij}) < r_{ij}$, let $A_{ij} \subset A$ of the same form as above, this time with $m(A_{ij}) = r_{ij} - m(\tilde{F}_{ij})$. In this case, define $F_{ij} := \tilde{F}_{ij} \cup A_{ij}$. Now repeat this process for all j, and then let

$$F_i := \bigcup_{j=1}^{2^K} F_{ij}.$$

Note that some F_{ij} may be empty. In particular, $F_{ij} = \emptyset$ if and only if $r_{ij} = 0$.

Note that by definition, $m(F_{ij}) = r_{ij}$ and note further that $\tilde{F}_{ij} \triangle F_{ij} = A_{ij}$. We claim

$$\sum_{j=1}^{2^K} m(A_{ij}) < 2\varepsilon$$

for all i. Let i be fixed, and consider

$$\begin{aligned} \sum_j m(A_{ij}) &= \sum_j \left| m(\tilde{F}_{ij}) - r_{ij} \right| \\ &\leq \sum_j |m(E_{ij}) - r_{ij}| + \sum_j \left| m(E_{ij}) - m(\tilde{F}_{ij}) \right|. \end{aligned}$$

We have $|m(E_{ij}) - r_{ij}| < \dfrac{\varepsilon}{2^K}$ so $\sum_j |m(E_{ij}) - r_{ij}| < \varepsilon$. On the other hand,

$$\sum_j \left| m(E_{ij}) - m(\tilde{F}_{ij}) \right| \leq \sum_j m(E_{ij} \triangle \tilde{F}_{ij}) = m\left(\bigcup_j (E_{ij} \triangle \tilde{F}_{ij}) \right).$$

Now, because $m(E_{ij_1} \cap \tilde{F}_{ij_2}) = 0$ if $j_1 \neq j_2$, we have that

$$m\left(\bigcup_j (E_{ij} \triangle \tilde{F}_{ij}) \right) = m\left(\bigcup_j E_{ij} \triangle \bigcup_j \tilde{F}_{ij} \right) = m(E_i \triangle \tilde{F}_i) < \varepsilon.$$

Therefore, $\sum_j m(A_{ij}) < 2\varepsilon$.

We will now show $m(E_i \triangle F_i) < 3\varepsilon$. Firstly, we have $m(E_i \triangle F_i) \leq m(E_i \triangle \tilde{F}_i) + m(\tilde{F}_i \triangle F_i)$. But $m(E_i \triangle \tilde{F}_i) < \varepsilon$. Further,

$$\begin{aligned} m(\tilde{F}_i \triangle F_i) &= m\left(\bigcup_j \tilde{F}_{ij} \triangle \bigcup_j F_{ij} \right) \\ &\leq m\left(\bigcup_j (\tilde{F}_{ij} \triangle F_{ij}) \right) = \sum_j m(\tilde{F}_{ij} \triangle F_{ij}). \end{aligned}$$

But as previously noted, $\tilde{F}_{ij} \triangle F_{ij} = A_{ij}$, and we already showed $\sum_j m(A_{ij}) < 2\varepsilon$. Thus, $m(E_i \triangle F_i) < 3\varepsilon$ as desired. □

With Lemma 2.1, we can now prove the equivalent version of Theorem 2.1 for the unit square, which will be the core result for proving our version of the Weak Approximation Theorem.

Theorem 2.2 (Density of column-preserving permutations)**.** *Column-preserving permutations are dense in $\mathcal{G}_X$. More precisely, let $T \in \mathcal{G}_X$. Given $N_\varepsilon(T)$, a dyadic neighborhood of T, there exists $Q \in N_\varepsilon(T)$, a column-preserving permutation.*

Proof. We may assume without loss of generality that

$$N_\varepsilon(T) = \{S : \mu(TD_i \triangle SD_i) < \varepsilon, l = 1, \ldots, (2^{2N})\},$$

where D_i are dyadic squares of some fixed rank N. We start with the case where $T' = \mathbb{1}_X$.

Let $D_{ij} := D_i \cap TD_j$. Note that $\{D_{ij}\}$ partitions Z. By Lemma 1.1, there exists a partition of Z into dyadic sets, $\{\tilde{E}_{ij}\}$, such that

$$m(D_{ij} \triangle \tilde{E}_{ij}) < \frac{\varepsilon}{6M},$$

where $M := 2^{2N}$. Further by Lemma 1.1, we can find a dyadic partition of Z into sets $\{\tilde{F}_{ij}\}$ where

$$m(T^{-1}D_{ij} \triangle \tilde{F}_{ij}) < \frac{\varepsilon}{6M}.$$

Note that because $m(D_{ij}) = m(T^{-1}D_{ij})$, we can assume that $m(\tilde{E}_{ij}) = m(\tilde{F}_{ij})$ Let $K = \max \text{rank}\{\tilde{E}_{ij}, \tilde{F}_{ij}\}$. We can now apply Lemma 2.1 to both $\tilde{E}_{ij}$ and $\tilde{F}_{ij}$ to get dyadic partitions $\{E_{ij}\}$ and $\{F_{ij}\}$ such that

$$m(D_{ij} \triangle E_{ij}) < \frac{\varepsilon}{2M} \text{ and } m(T^{-1}D_{ij} \triangle F_{ij}) < \frac{\varepsilon}{2M}. \tag{2.4}$$

Recall that if C_k is a dyadic interval of rank K, then in the notation of Lemma 2.1, $E_{ijk} := E_{ij} \cap \pi^{-1}C_k$ and $F_{ijk} := F_{ij} \cap \pi^{-1}C_k$. Note that not only do we have $m(D_{ij}) = m(T^{-1}D_{ij})$, but because T is an extension of identity,

$$m(T^{-1}D_{ij} \cap \pi^{-1}C_k) = m(T^{-1}(D_{ij} \cap \pi^{-1}C_k)) = m(D_{ij} \cap \pi^{-1}C_k).$$

Thus we are able to choose the same dyadic rationals in both applications of Lemma 2.1, and subsequently have that $m(E_{ijk}) = m(F_{ijk})$ for $i, j = 1, \ldots 2^{2N}, k = 1, \ldots, 2^K$.

We now define Q as the permutation which maps F_{ijk} to E_{ijk}. Note that in particular, Q will map F_{ij} to E_{ij}. Further note that Q will be an extension of the identity.

Let j be fixed. We will now show $m(QD_j \triangle TD_j) < \varepsilon$. Recall $D_{ij} = D_i \cap TD_j$, so $T^{-1}D_{ij} = T^{-1}D_i \cap D_j$ and $D_j = \bigcup_i T^{-1}D_{ij}$. We have

$$m\left(D_j \triangle \bigcup_i F_{ij}\right) = m\left(\bigcup_i T^{-1}D_{ij} \triangle \bigcup_i F_{ij}\right) \leq \sum_i m(T^{-1}D_{ij} \triangle F_{ij}). \tag{2.5}$$

But per (2.4), $m(T^{-1}D_{ij} \triangle F_{ij}) < \frac{\varepsilon}{2M}$, so $(2.5) < \sum_i \frac{\varepsilon}{2M} \leq \frac{\varepsilon}{2}$. Therefore

$$m\left(QD_j \triangle \bigcup_i E_{ij}\right) = m\left(D_j \triangle \bigcup_i F_{ij}\right) < \frac{\varepsilon}{2}.$$

On the other hand,

$$m\left(TD_j \triangle \bigcup_i E_{ij}\right) = m\left(\bigcup_i D_{ij} \triangle \bigcup_i E_{ij}\right) \leq \sum_i m(D_{ij} \triangle E_{ij}). \tag{2.6}$$

Again, per (2.4), $m(D_{ij} \triangle E_{ij}) < \frac{\varepsilon}{2M}$, so (2.6) $< \sum_i \frac{\varepsilon}{2M} = \frac{\varepsilon}{2}$. Therefore,

$$\begin{aligned} & m(TD_j \triangle QD_j) \\ \leq & m\left(TD_j \triangle \bigcup_i E_{ij}\right) + m\left(\bigcup_i E_{ij} \triangle QD_j\right) \\ < & \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

As this holds for all j, we have that $Q \in N_\varepsilon(T)$.

Now suppose T is an extension of some invertible T'. Define $\tilde{T} := (T')^{-1}T$. Then $\tilde{T}$ is an extension of the identity, so there exists a column-preserving permutation $\tilde{Q} \in N_{\varepsilon/2}(\tilde{T})$. But then $T'\tilde{Q} \in N_{\varepsilon/2}(T)$ as

$$m(T'\tilde{Q}D_i \triangle TD_i) = m(\tilde{Q}D_i \triangle \tilde{T}D_i) < \frac{\varepsilon}{2}.$$

By Remark 2.3, $T'\tilde{Q}$ is L-simply discrete equivalent, with $L = 2^{\text{rank}\,\tilde{Q}}$. If we let $G_i := \pi_L D_i$ and let

$$\tilde{N}_{\varepsilon/2}(\pi_L(T'\tilde{Q})) := \{S_L : \mu_L(\pi_L(T'\tilde{Q})G_i \triangle S_L G_i) < \frac{\varepsilon}{2} \forall i\},$$

then by Theorem 2.1 there exists a column-preserving dyadic permutation $\hat{Q} \in \tilde{N}_{\varepsilon/2}(\pi_L(T'\tilde{Q}))$. Now we define Q to be the simply discrete equivalent extension of $\hat{Q}$. Note that because L was dyadic, Q is a (column-preserving) dyadic permutation. Further, $Q \in N_{\varepsilon/2}(T'\tilde{Q})$ as

$$m(QD_i \triangle T'\tilde{Q}D_i) = \mu_L(\hat{Q}G_i \triangle \pi_L(T'\tilde{Q})G_i) < \frac{\varepsilon}{2}.$$

So

$$m(QD_i \triangle TD_i) \leq m(QD_i \triangle T'\tilde{Q}D_i) + m(T'\tilde{Q}D_i \triangle TD_i) < \varepsilon$$

for all i, and thus $Q \in N_\varepsilon(T)$. □

We close this section with the promised version of Halmos' Weak Approximation Theorem for extensions.

Theorem 2.3 (Weak Approximation Theorem for Extensions)**.** *Let $T \in \mathcal{G}_X$, and let $N_\varepsilon(T)$ be a dyadic neighborhood of T. Then for any $k_0 \in \mathbb{N}$, there exists $k \geq k_0$ and $Q \in \mathcal{G}_X$ such that the following hold:*

- *Q, Q' are dyadic permutations of rank k on Z, X respectively,*
- *Q' is cyclic (of period 2^k),*
- *Q is periodic with period 2^k everywhere,*
- *$Q \in N_\varepsilon(T)$.*

Proof. Because Theorem 2.2 tells us that $N_{\varepsilon/2}(T)$ will contain $P \in \mathcal{G}_X$, (P a column-preserving permutation) we need only prove the case where T is a permutation itself (we will therefore proceed using P, P' in place of T, T').

Fix $k_0 \in \mathbb{N}$. Because P is a permutation and D_i is a dyadic set, PD_i is also a dyadic set. Let M be the maximum rank of PD_i (so that P is a permutation of rank M), K be the number of disjoint cycles in P', and k be chosen to be greater than both M and k_0, and such that $\dfrac{K}{2^{k-1}} < \varepsilon$.

We will now construct Q of rank k. Note that following the proof there will be an example of this construction. To start, let E_1 be any dyadic square of rank M. If πE_1 is not a fixed point of P', we have Q map the "first" rank k dyadic square (which we will henceforth refer to as a k-square) of E_1 to the "first" k-square in PE_1. By "first" k-square, we mean the top left k-square. Now, if $(P')^2\pi E_1 \neq \pi E_1$, we continue to map to the first k-square in $(P')^2\pi E_1$. Eventually, however, we reach a point where $(P')^l\pi E_1 = \pi E_1$. From where we are in $(P)^{l-1}E_1$, we continue to map to the "second" k-square in P^lE_1 (by "second" we mean the one to the right of the first). Note that $P^lE_1 \neq E_1$ in general.

We now repeat the entire process, replacing "first" for "second," eventually "third" and so on, as well as replacing E_1 with P^lE_1. Eventually we will arrive at a k-square whose projection is at the far right of $(P')^{l-1}\pi E_1$. At this point, we choose an M-square E_2 such that πE_2 is not in the P' cycle of πE_1 (assuming such an E_2 exists). Then from our current position, we map to the first k-square of E_2, and repeat the process.

We continue on like this until we have exhausted every P' cycle (including fixed points), at which point we return to the the first k-square of E_1. Note that we have visited every k-column exactly once. We are not quite done yet, though. We now choose a k-square on the same column as the first k-square in E_1, and we repeat the entire process. Now shifting to rows within the M-squares that correspond to our new choice of starting point. That is, in the original process, we were in the top row of every M-square, because our original k-square was in the top row. If our new k-square is in the 3rd row within its M-square, say, all our choices will be in the 3rd row of the respective M-squares. Note that while this precision is unnecessary, it can be useful for ensuring no mistakes are made. Repeating this process, we eventually define Q for all k-squares.

We now find a bound for $m(PD_i \triangle QD_i)$. Note that by our construction the only points that can be in $PD_i \triangle QD_i$ come from k-squares in D_i whose projections are in the last k-interval in each P' cycle. Let E_j be such a k-square. Then

$$m(PE_j \triangle QE_j) \leq \frac{2}{2^{2k}} = \frac{1}{2^{2k-1}}.$$

There are 2^k such E_j per k-column, and there are K such k-columns. Thus,

$$m(PD_i \triangle QD_i) \leq \bigcup_j m(PE_j \triangle QE_j) \leq \frac{K2^k}{2^{2k-1}} = \frac{K}{2^{k-1}} < \varepsilon.$$

□

The construction in the proof of Theorem 2.3 can be difficult to follow closely, so we provide an example of the construction. We first provide a P which, in this case, will be of

rank 2. See Figure 2.1 for reference on how we label the 2-squares. Note that we will define P using cycle decomposition notation. That is, if we write $R = (1\ 2\ 3)$, then we mean that the image under R of the square labeled 1 is the square labeled 2. Similarly the image of "2" is "3" and the image of "3" is "1". Any squares not written explicitly in the decomposition are fixed points. Now, we let $P := (1\ 11\ 5\ 3)(13\ 15)(9\ 7)(2\ 6\ 14)(4\ 16\ 12\ 8)$. Note that P extends $P' := (1\ 3)$ on X.

Figure 2.1

1	2	3	4
1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Figure 2.2

(1,1)	(5,3)	(1,5)	(5,6)	(5,2)	(1,4)	(5,7)	(7,8)
(3,1)	(6,3)	(6,5)	(4,6)	(6,2)	(3,4)	(3,7)	(4,8)
(5,1)	(1,3)	(2,5)	(1,6)	(2,2)	(2,4)	(7,7)	(2,8)
(6,1)	(3,3)	(7,5)	(6,6)	(8,2)	(8,4)	(4,7)	(8,8)
(2,1)	(2,3)	(8,5)	(8,6)	(1,2)	(5,4)	(2,7)	(1,8)
(8,1)	(8,3)	(3,5)	(3,6)	(3,2)	(6,4)	(8,7)	(6,8)
(7,1)	(7,3)	(5,5)	(2,6)	(7,2)	(7,4)	(1,7)	(5,8)
(4,1)	(4,3)	(4,5)	(7,6)	(4,2)	(4,4)	(6,7)	(3,8)

Suppose we were to construct Q to be a rank 3 permutation. Rather than write the entire cycle decomposition of Q (as it would involve writing all 64 3-squares), we label Figure 2.2 to define Q. Here we have labeled the 3-squares such that for a square labeled (n, k), we have that $Q(n, k) = (n, k + 1), (k \mod 8)$ (for consistency, here we have $8 \mod 8 := 8$ instead of 0 as it typically would be). Further, if $n_1 \neq n_2$, then $(n_1, k_1), (n_2, k_2)$ are in independent cycles. It is easy to see with this notation that Q is an extension of a cyclic permutation Q' on X. We also note that the Q we constructed is

not the only possible Q we could have constructed, as we have many free choices in the construction.

To close this section, we note that a very simple modification of the proof of Theorem 2.3 would yield a column preserving permutation Q such that not only Q' is cyclic, but Q is cyclic as well. In our example seen in Figure 2.2, this modification would be accomplished by changing the definition of Q slightly so that $Q(n, 8) = (n + 1, 1), (n \bmod 8)$. This formulation is more akin to the classical theorem. However, we choose the formulation given in Theorem 2.3 as it is this formulation we need for further results.

2.1.4 Uniform Approximation

Our goal in this section is to prove results that are generalizations of those needed for Halmos' classical Conjugacy Lemma (Lemma 1.5), and whose proofs quickly follow from the classical results and their proofs.

Lemma 2.2. *Let $T \in \mathcal{G}_X$ where T' is periodic of period n (almost) everywhere. Then there exists a set E such that $E = \pi^{-1}E'$ for some $E' \subset X$, and $\{E, TE, \ldots, T^{n-1}E\}$ partition Z.*

Proof. Because T' is has period n everywhere, there exists E' such that $\{E', T'E', \ldots, (T')^{n-1}E'\}$ partitions X. Setting $E := \pi^{-1}E'$ we have

$$\{E, TE, \ldots, T^{n-1}E\}$$

are pairwise disjoint because T extends T'. Further, because $m(E) = m(E') = \dfrac{1}{n}$, we have

$$m\left(\bigcup_{i=0}^{n-1} T^i E\right) = \sum_{i=0}^{n-1} m(T^i E) = 1,$$

or $\displaystyle\bigcup_{i=0}^{n-1} T^i E = X.$ □

Next we move to a version of Rokhlin's lemma (Lemma 1.3).

Lemma 2.3. *Let $T \in \mathcal{G}_X$ where T' is antiperiodic. Then for every $n \in \mathbb{N}$ and $\varepsilon > 0$ there exists E such that $E = \pi^{-1}E'$ for some E', $\{E, TE, \ldots, T^{n-1}E\}$ are pairwise disjoint, and $m\left(\displaystyle\bigcup_{i=0}^{n-1} T^i E\right) > 1 - \varepsilon$.*

Proof. Let $n \in \mathbb{N}$ and $\varepsilon > 0$. Because T' is antiperiodic, there exists $E' \subset X$ such that $\{E', T'E', \ldots, (T')^{n-1}E'\}$ are pairwise disjoint and

$$m\left(\bigcup_{i=0}^{n-1} (T')^i E'\right) > 1 - \varepsilon.$$

Let $E := \pi^{-1}E'$. Because T extends T', $\{E, TE, \ldots, T^{n-1}E\}$ are pairwise disjoint. Further,

$$m\left(\bigcup_{i=0}^{n-1} T^i E\right) = \sum_{i=0}^{n-1} m(T^i E)$$
$$= \sum_{i=0}^{n-1} m((T')^i E') = m\left(\bigcup_{i=0}^{n-1} (T')^i E'\right) > 1 - \varepsilon.$$

□

We conclude this section with a version of Halmos' Uniform Approximation Theorem (Theorem 1.3).

Theorem 2.4 (Uniform Approximation Theorem for Extensions)**.** *Let $T \in \mathcal{G}_X$ where T' is antiperiodic. Then for every $n \in \mathbb{N}$ and $\varepsilon > 0$ there exists $R \in \mathcal{G}_X$, such that both R and R' are periodic with period n almost everywhere, and $d'(R, T) \leq \frac{1}{n} + \varepsilon$.*

Proof. By Lemma 2.3, there exists E a cylinder set, such that $\{E, TE, \ldots, T^{n-1}E\}$ are pairwise disjoint, and

$$m\left(\bigcup_{i=0}^{n-1} T^i E\right) > 1 - \varepsilon.$$

If $z \in \bigcup_{i=0}^{n-2} T^i E$, define $Rz := Tz$, and if $z \in T^{n-1}E$, define $Rz := T^{-(n-1)}z$, thus making R have period n for all points on which we have thus far defined it. Further, because T extends T', R is also an extension. And for any definition of R on the remainder of Z, we have $d'(R, T) \leq m(T^{n-1}E) + \varepsilon \leq \frac{1}{n} + \varepsilon$.

All that remains is to define R on the remainder of Z so that R is an extension, and R, R' have period n. Since the remainder is a cylinder set, this can be done by defining R' on the projection of the remainder, as in the classical case (see the proof of Theorem 1.3), and then letting $R = R' \times \mathbb{1}_Y$ on this set of measure ε. □

2.1.5 Conjugacy Lemma

We now prove a generalization of Halmos' Conjugacy Lemma (Lemma 1.5), using the same techniques as Halmos' original proof.

Lemma 2.4 (Conjugacy Lemma for Extensions)**.** *Let $T \in \mathcal{G}_X, T_0 \in \mathcal{G}_X$ such that T_0' is antiperiodic, and let $N_\varepsilon(T) = \{V \in \mathcal{G}_X : m(VD_i \triangle TD_i) < \varepsilon, i = 1, \ldots, N\}$ be a dyadic neighborhood of T. Then there exists $S \in \mathcal{G}_X$ such that $S^{-1}T_0S \in N_\varepsilon(T)$.*

Proof. Let $k_0 \in \mathbb{N}$ be greater than the ranks of all D_i and $\frac{1}{2^{k_0-2}} < \varepsilon$. Further, let $Q \in N_{\varepsilon/2}(T)$, a dyadic permutations of rank $k \geq k_0$ with all properties guaranteed by the Weak

Approximation Theorem for Extensions, Theorem 2.3 (Q' is cyclic, Q is 2^k periodic). Applying the Uniform Approximation Theorem for Extensions, Theorem 2.4, with 2^k in place of n and $\frac{1}{2^k}$ in place of ε, there exists $R \in \mathcal{G}_X$ such that R, R' have period 2^k almost everywhere, and $d'(R, T_0) \leq \frac{1}{2^k} + \frac{1}{2^k} < \frac{\varepsilon}{2}$.

We will show Q and R are conjugate by some $S \in \mathcal{G}_X$. Let $q = 2^k$ and $E_0, \ldots, E_{q-1}$ be cylinder sets of dyadic intervals of rank k in X, arranged so that $QE_i = E_{i+1}$ ($i \mod q$). Note that $m(E_i) = \frac{1}{q}$ By Lemma 2.2 there exists F_0 a cylinder set, such that $m(F_0) = \frac{1}{q}$ and $F_0, RF_0, \ldots, R^{q-1}F_0$ partition Z. Let $F_i := R^i F_0$. Let S be any measure preserving transformation which maps E_0 to F_0 as an extension of some S'. Then for $z \in E_i$, let $Sz := R^i S Q^{-i} z$. This can be seen in the following diagram:

$$\begin{tikzcd}
E_0 \arrow[r, "Q"] \arrow[d, "S"'] & E_1 \arrow[r, "Q"] \arrow[d, "S"'] & E_2 \arrow[r, "Q"] \arrow[d, "S"'] & \ldots \arrow[r, "Q"] & E_{q-2} \arrow[r, "Q"] \arrow[d, "S"'] & E_{q-1} \arrow[d, "S"'] \\
F_0 \arrow[r, "R"'] & F_1 \arrow[r, "R"'] & F_2 \arrow[r, "R"'] & \ldots \arrow[r, "R"'] & F_{q-2} \arrow[r, "R"'] & F_{q-1}
\end{tikzcd}$$

Commutation of the diagram shows that $Q = S^{-1}RS$. Further, because $Q, R, S \restriction_{E_0}$ are extensions of $Q', R', S' \restriction_{\pi E_0}$, S is an extension of S'.

Now, because d' is invariant under group operations, we have

$$d'(Q, S^{-1}T_0S) \leq d'(S^{-1}RS, S^{-1}T_0S) = d'(R, T_0) < \frac{\varepsilon}{2}.$$

Thus, for any D_i we have:

$$\begin{aligned} m(TD_i \triangle S^{-1}T_0SD_i) &\leq m(TD_i \triangle QD_i) + m(QD_i \triangle S^{-1}T_0SD_i) \\ &\leq \frac{\varepsilon}{2} + \tilde{d}(Q, S^{-1}T_0S). \end{aligned}$$

But $\tilde{d} \leq d'$, so

$$m(TD_i \triangle S^{-1}T_0SD_i) \leq \frac{\varepsilon}{2} + d'(Q, S^{-1}T_0S) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

This lemma is so important because as we will see in our main result in Theorem 2.5, the conjugacy class of $\mathcal{W}_X$ is $\mathcal{W}_X$ itself. Thus, in proving Lemma 2.4, we have indeed proven half of Theorem 2.5.

2.1.6 Category Theorem

We are fast approaching our main goal: that $\mathcal{W}_X$ is a dense, G_δ subset of $\mathcal{G}_X$. Before we can prove it, we need to prove a few technical results. First we have a quick consequence of the Cauchy-Schwarz Inequality, Proposition 2.1, but it will be important enough to make a special note of it.

Proposition 2.4. *Let $f, g \in L^2(Z|X)$. Then*

$$\|\mathbb{E}(f \cdot \overline{g}|X)\|_{L^2(X)} \leq \left\| \|f\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|g\|_{L^2(Z)} .$$

Proof. By the Cauchy-Schwarz Inequality for $L^2(Z|X)$, we have

$$|\mathbb{E}(f \cdot \overline{g}|X)| \leq \mathbb{E}(|f|^2 |X)^{1/2} \mathbb{E}(|g|^2 |X)^{1/2}$$

pointwise. Now, by definition of $L^2(Z|X), \mathbb{E}(|f|^2 |X)^{1/2} \in L^\infty(X)$. Letting

$$M := \left\| \|f\|_{L^2(Z|X)} \right\|_{L^\infty(X)},$$

we have

$$|\mathbb{E}(f \cdot \overline{g}|X)| \leq M \mathbb{E}(|g|^2 |X)^{1/2}. \tag{2.7}$$

Now note that by Fubini,

$$\begin{aligned} \left\| \mathbb{E}(|g|^2 |X)^{1/2} \right\|_{L^2(X)} &= \left(\int_X \left(\left(\int_Y |g|^2 (x,y) d\mu_Y \right)^{1/2} \right)^2 d\mu_X \right)^{1/2} \\ &= \left(\int_Z |g|^2 \, dm_2 \right)^{1/2} = \|g\|_{L^2(Z)} , \end{aligned}$$

where $Y = X, \mu_X = \mu_Y = m_1$ (the notation here was changed to clarify what integrals were intended). And so taking $L^2(X)$ norm on both sides of (2.7), we have

$$\|\mathbb{E}(f \cdot \overline{g}|X)\|_{L^2(X)} \leq M \left\| \mathbb{E}(|g|^2 |X)^{1/2} \right\|_{L^2(X)} = \left\| \|f\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|g\|_{L^2(Z)}$$

as desired. □

Our immediate goal is to prove that $T \in \mathcal{W}_X$ is equivalent to the existence of a subsequence n_k such that for all $f, g \in L^2(Z|X)$,

$$\lim_{k \to \infty} \|\mathbb{E}(T^{n_k} f \cdot \overline{g}|X) - (T')^{n_k} \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0. \tag{2.8}$$

To this end, we first show that if (2.8) holds for an $L^2(Z)$-dense subset of $L^2(Z|X)$, it holds for all of $L^2(Z|X)$.

Lemma 2.5. *Let $T \in \mathcal{G}_X$, and let $D \subset L^2(Z|X)$, with*

$$\left\| \|\overline{f}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \leq 1$$

for all $f \in D$, such that D is dense in the unit ball of $L^2(Z|X)$, with respect to the $L^2(Z)$ norm topology. Further suppose that there exists a subsequence (n_k) such that for all $f_i, f_j \in D$,

$$\lim_{k \to \infty} \left\| \mathbb{E}(T^{n_k} f_i \cdot \overline{f_j}|X) - (T')^{n_k} \mathbb{E}(f_i|X) \cdot \mathbb{E}(\overline{f_j}|X) \right\|_{L^2(X)} = 0.$$

Then for all $h, g \in L^2(Z|X)$

$$\lim_{k \to \infty} \|\mathbb{E}(T^{n_k} h \cdot \overline{g}|X) - (T')^{n_k} \mathbb{E}(h|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0.$$

Proof. Fix $h, g \in L^2(Z|X)$ and let $(h_j), (g_j) \subset D$ such that

$$h_j \xrightarrow{L^2(Z)} h, g_j \xrightarrow{L^2(Z)} g.$$

We first claim that

$$\lim_{j\to\infty} \|\mathbb{E}(T^n h_j \cdot \overline{g_j}|X) - \mathbb{E}(T^n h \cdot \overline{g}|X)\|_{L^2(X)} = 0,$$

uniformly with respect to n. Indeed, we have

$$\begin{aligned} & \|\mathbb{E}(T^n h_j \cdot \overline{g_j}|X) - \mathbb{E}(T^n h \cdot \overline{g}|X)\|_{L^2(X)} \\ \leq\ & \|\mathbb{E}(T^n h_j \cdot \overline{g_j} - T^n h_j \cdot \overline{g}|X)\|_{L^2(X)} + \|\mathbb{E}(T^n h_j \cdot \overline{g} - T^n h \cdot \overline{g}|X)\|_{L^2(X)} \\ =\ & \|\mathbb{E}(T^n h_j \cdot (\overline{g_j} - \overline{g})|X)\|_{L^2(X)} + \|\mathbb{E}(\overline{g} \cdot (T^n h_j - T^n h)|X)\|_{L^2(X)}. \end{aligned}$$

Now by Proposition 2.4

$$\begin{aligned} \|\mathbb{E}(T^n h_j \cdot (\overline{g_j} - \overline{g})|X)\|_{L^2(X)} &\leq \left\| \|T^n h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(Z)}, \\ \|\mathbb{E}(\overline{g} \cdot (T^n h_j - T^n h)|X)\|_{L^2(X)} &\leq \left\| \|\overline{g}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|T^n(h - h_j)\|_{L^2(Z)}. \end{aligned}$$

In turn, as $\left\| \|T^n h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} = \left\| \|h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)}$ and T is an isometry, we get

$$\begin{aligned} \left\| \|T^n h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(Z)} &\leq \left\| \|h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(Z)}, \\ \left\| \|\overline{g}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|T^n(h - h_j)\|_{L^2(Z)} &\leq \left\| \|\overline{g}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|h - h_j\|_{L^2(Z)}. \end{aligned}$$

Because $h_j \to h, g_j \to g$ in $L^2(Z)$, we have the desired result. Note that a similar argument using Proposition 2.4 will show $\mathbb{E}(\overline{g_j}|X) \to \mathbb{E}(\overline{g}|X), (T')^n\mathbb{E}(h_j|X) \to (T')^n\mathbb{E}(h|X))$ in $L^2(X)$. For the reader's convenience, we show the argument again. In the first case we have

$$\begin{aligned} & \|\mathbb{E}(\overline{g_j}|X) - \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = \|\mathbb{E}(\overline{g_j} - \overline{g}|X)\| \\ \leq & \left\| \|1\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(X)} = \|\overline{g_j} - \overline{g}\|_{L^2(X)}, \end{aligned}$$

and in the latter case we have

$$\begin{aligned} & \|(T')^n\mathbb{E}(h_j|X) - (T')^n\mathbb{E}(h|X)\|_{L^2(X)} = \|\mathbb{E}(h_j - h|X)\| \\ \leq & \left\| \|1\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|h_j - h\|_{L^2(X)} = \|h_j - h\|_{L^2(X)}. \end{aligned}$$

Now,

$$\begin{aligned} & \|\mathbb{E}(T^n h \cdot \overline{g}|X) - (T')^n\mathbb{E}(h|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} \\ \leq\ & \|\mathbb{E}(T^n h \cdot \overline{g}|X) - \mathbb{E}(T^n h_j \cdot \overline{g_j}|X)\|_{L^2(X)} \\ +\ & \|\mathbb{E}(T^n h_j \cdot \overline{g_j}|X) - (T')^n\mathbb{E}(h_j|X) \cdot \mathbb{E}(\overline{g_j}|X)\|_{L^2(X)} \\ +\ & \|(T')^n\mathbb{E}(h_j|X) \cdot \mathbb{E}(\overline{g_j}|X) - (T')^n\mathbb{E}(h|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)}. \end{aligned}$$

By hypothesis, there is a subsequence n_k (independent of j) such that the middle term converges to 0. Further, the first and third terms converge to 0 as $j \to \infty$ uniformly in n, so

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k}h \cdot \overline{g}|X) - (T')^{n_k}\mathbb{E}(h|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0$$

as desired. □

Next, recall that a function $f \in L^2(Z|X)$ is called *generalized eigenfunction* for a given $T \in \mathcal{G}_X$ (see [27, p.179]) if the $L^\infty(X)$-module spanned by $\{T^n f : n \in \mathbb{N}\}$ has finite rank. In other words, there exists $g_1, \ldots, g_l \in L^2(Z|X)$ such that for all n, there exists $c_j^n \in L^\infty(X), 1 \leq j \leq l$ such that

$$T^n f(x,y) = \sum_j c_j^n(x) g_j(x,y).$$

Lemma 2.6. *Let $T \in \mathcal{G}_X$. Then $T \in \mathcal{W}_X$ if and only if there exists a subsequence n_k such that for all $f, g \in L^2(Z|X)$*

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k}f \cdot \overline{g}|X) - (T')^{n_k}\mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0. \tag{2.9}$$

Proof. Let $T \in \mathcal{W}_X$. Lemma 2.5 tells us that we need only show that there exists a subsequence such that (2.9) holds for an $L^2(Z)$-dense subset of $L^2(Z|X)$. Let $(f_i)_{i=1}^\infty$ be an enumeration of that dense subset, and let $l \mapsto (f_i, f_j), 1 \leq l < \infty$ order the set $\{(f_i, f_j)|i, j \in \mathbb{N}\}$ arbitrarily. Now, by the definition of a weakly mixing extension and the Koopman-von Neumann Lemma (see, e.g., [17, p.54]), for each l there exists a subsequence $n_m'^l$ of upper asymptotic density 1 such that

$$\lim_{m\to\infty} \left\|\mathbb{E}(T^{n_m'^l}f_i \cdot \overline{f_j}|X) - (T')^{n_m'^l}\mathbb{E}(f_i|X) \cdot \mathbb{E}(\overline{f_j}|X)\right\|_{L^2(X)} = 0.$$

Stated differently, for the pair f_i, f_j corresponding to l, we have the desired convergence along the dense subsequence $(n_m'^l)$. We now define a new density 1 subsequence for each l inductively. Let $(n_m^1) := (n_m'^1)$, and for $l > 1$, define n_m^l to be a common density 1 subsequence of (n_m^{l-1}) and $(n_m'^l)$. Now define (n_k) be a diagonal sequence obtained from (n_m^l), i.e., $n_k := n_k^k$. We now show that for all f_i, f_j, (2.9) holds along (n_k). Indeed, fix $\varepsilon > 0, i, j \in \mathbb{N}$ and let l correspond to i, j. There exists N such that for $m \geq N$,

$$\left\|\mathbb{E}(T^{n_m^l}f_i \cdot \overline{f_j}|X) - (T')^{n_m^l}\mathbb{E}(f_i|X) \cdot \mathbb{E}(\overline{f_j}|X)\right\|_{L^2(X)} < \varepsilon.$$

Now let $K := \max\{l, N\}$. Now for $k \geq K$, there exists m_0 such that $n_k = n_{m_0}^l$ and $m_0 \geq N$. Therefore, for $k \geq K$,

$$\|\mathbb{E}(T^{n_k}f \cdot \overline{g}|X) - (T')^{n_k}\mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} < \varepsilon.$$

To prove the converse, suppose $T \in \mathcal{G}_X \backslash \mathcal{W}_X$. Then there exists $f \in L^2(Z|X) \backslash L^\infty(X)$ that is a generalized eigenfunction for T, see [27, p.192]. Without loss of generality, $\|f\|_{L^1(Z)} = 1$. Let $g_1, \ldots, g_l \in L^2(Z|X)$ be a basis for the module spanned by $T^n f$. We want g_i to be "relatively orthonormal". That is, we want $\mathbb{E}(g_i \cdot \overline{g}_j|X) = 0$ $a.e.$ when

$i \neq j$ and $\mathbb{E}(|g_j|^2 |X) = 1$ $a.e.$ This can be accomplished with a relative Gram-Schmidt process. We start by defining $h_1' := g_1$. For any x such that $\mathbb{E}(|h_1'|^2 |X)(x) = 0$, we have that $h_1'(x,y) := h_{1,x}'(y) \equiv 0$. Thus by setting the corresponding $c_1^n(x) = 0$ for all n we can define $h_{1,x}(y)$ arbitrarily, so long as it is not identically 0. For all other x, define $h_{1,x}(y) := h_{1,x}'(y)$. Now, having defined $h_j, 1 \leq j \leq i-1$, we define h_i' by

$$h_i' := g_i - \sum_{j=1}^{i-1} \frac{\mathbb{E}(g_i \cdot \overline{h}_j|X)}{\mathbb{E}(|h_j|^2 |X)} h_j.$$

Similar to the above, if there are any x such that $\mathbb{E}(|h_i'|^2 |X)(x) = 0$, we define $h_{i,x}(y) \not\equiv 0$ (again changing $c_i^n(x)$ to 0 for all n), but now with the additional condition that $h_{i,x}$ is orthogonal to $h_{j,x}$ for $1 \leq j \leq i-1$. For all other $x, h_{i,x}(y) := h_{i,x}'(y)$. Notice that if we have $1 \leq k < i$, then for x such that we did not need to redefine $h_{i,x}$, we have:

$$\begin{aligned}
\mathbb{E}(h_i \cdot \overline{h}_k|X)(x) &= \mathbb{E}\left(\left(g_i - \sum_{j=1}^{i-1} \frac{\mathbb{E}(g_i \cdot \overline{h}_j|X)}{\mathbb{E}(|h_j|^2 |X)} h_j\right) \cdot \overline{h}_k|X\right)(x) \\
&= \mathbb{E}(g_i \cdot \overline{h}_k|X)(x) - \mathbb{E}\left(\left(\sum_{j=1}^{i-1} \frac{\mathbb{E}(g_i \cdot \overline{h}_j|X)}{\mathbb{E}(|h_j|^2 |X)} h_j\right) \cdot \overline{h}_k|X\right)(x) \\
&= \mathbb{E}(g_i \cdot \overline{h}_k|X)(x) - \mathbb{E}\left(\frac{\mathbb{E}(g_i \cdot \overline{h}_k|X)}{\mathbb{E}(|h_k|^2 |X)} h_k \cdot \overline{h}_k|X\right)(x) \\
&= \mathbb{E}(g_i \cdot \overline{h}_k|X)(x) - \frac{\mathbb{E}(g_i \cdot \overline{h}_k|X)}{\mathbb{E}(|h_k|^2 |X)} \mathbb{E}(h_k \cdot \overline{h}_k|X)(x) = 0.
\end{aligned}$$

Note that we get from the second line to the third as we assume we have already checked relative orthoganlity of h_j, h_k for $1 \leq j, k \leq i-1$. Further note that for x for which we did redefine $h_{i,x}$, we have that $\mathbb{E}(h_i \cdot \overline{h}_k|X)(x) = 0$ by choice of $h_{i,x}$. Finally we normalize and redefine g_j so that

$$g_j(x,y) := \frac{h_j(x,y)}{\mathbb{E}(|h_j|^2 |X)(x)^{1/2}}.$$

Now define a function $j : \mathbb{N} \to \{1, \ldots, l\}$ such that $\left\|c_{j(n)}^n\right\|_{L^2(Z)} \geq \|c_i^n\|_{L^2(Z)}, 1 \leq i \leq l$. Note that for each n, $\left\|c_{j(n)}^n\right\|_{L^2(Z)} \geq 1/l$ as else

$$\|T^n f\|_{L^1(Z)} \leq \sum_i \|c_i^n g_i\|_1 \leq \sum_i \|c_i^n\|_2 \|g_i\|_2 < 1.$$

Fix n and suppose for now that $\mathbb{E}(f|X) \equiv 0$ almost everywhere. Note that this is guaranteed to be possible because if $f = f_0 + \mathbb{E}(f|X)$ is a generalized eigenfunction with basis $\{g_1, \ldots g_l\}$, then f_0 is a generalized eigenfunction with spanning set $\{g_1, \ldots, g_l, \mathbb{E}(f|X)\}$,

and $\mathbb{E}(f_0|X) \equiv 0$ by design. Now, by relative orthonormality we have

$$\begin{aligned}
&\left\|\mathbb{E}(T^n f \cdot \overline{g}_{j(n)}|X) - (T')^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}_{j(n)}|X)\right\|_{L^2(X)} \\
&= \left\|\mathbb{E}(T^n f \cdot \overline{g}_{j(n)}|X)\right\|_{L^2(X)} \\
&= \left\|\mathbb{E}\left(\left(\sum_{i=1}^{l} c_i^n g_i\right) \cdot \overline{g}_{j(n)}|X\right)\right\|_{L^2(X)} \\
&= \left\|\mathbb{E}(c_{j(n)}^n g_{j(n)} \cdot \overline{g}_{j(n)}|X)\right\|_{L^2(X)} \\
&= \left\|c_j^n \mathbb{E}(\left|g_{j(n)}\right|^2 |X)\right\|_{L^2(X)} \\
&= \left\|c_j^n\right\|_{L^2(Z)} \geq \frac{1}{l}.
\end{aligned}$$

Define $B_i := j^{-1}(i)$. By the above work, if $n \in B_i$,

$$\|\mathbb{E}(T^n f \cdot \overline{g}_i|X)\|_{L^2(X)} \geq \frac{1}{l}.$$

Now given any subsequence (n_k) there will be at least one $i \in \{1, \ldots, l\}$ such that (n_k) intersects B_i infinitely often. Thus, $\|\mathbb{E}(T^{n_k} f \cdot \overline{g_i}|X)\|$ does not converge to 0 as $k \to \infty$.

If $\mathbb{E}(f|X) \neq 0$, we write $f = f_0 + h$ where $\mathbb{E}(f_0|X) = 0$ and $h := \mathbb{E}(f|X)$. Then

$$\begin{aligned}
&\mathbb{E}(T^n f \cdot \overline{g}|X) - (T')^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X) \\
&= \mathbb{E}(T^n (f_0 + h) \cdot \overline{g}|X) - (T')^n \mathbb{E}(f_0 + h|X) \cdot \mathbb{E}(\overline{g}|X).
\end{aligned}$$

By linearity of the conditional expectation, this is the same as

$$\mathbb{E}(T^n f_0 \cdot \overline{g}|X) - (T')^n \mathbb{E}(f_0|X) \cdot \mathbb{E}(\overline{g}|X) + \mathbb{E}(T^n h \cdot \overline{g}|X) - (T')^n \mathbb{E}(h|X) \cdot \mathbb{E}(\overline{g}|X).$$

The second term is 0 as $\mathbb{E}(f_0|X) \equiv 0$. Further, because $h \in L^\infty(X), \mathbb{E}(T^n h \cdot \overline{g}|X) = (T')^n h \cdot \mathbb{E}(\overline{g}|X) = (T')^n \mathbb{E}(h|X) \cdot \mathbb{E}(\overline{g}|X)$, which cancels with the fourth term. We are left with $\mathbb{E}(T^n f_0 \overline{g}|X)$ and have reduced this to the previous case. □

Finally, we arrive at our goal.

Theorem 2.5 (Weakly Mixing Extensions are Residual)**.** *$\mathcal{W}_X$ is a dense, G_δ subset of $\mathcal{G}_X$.*

Proof. We begin by proving that if $T \in \mathcal{W}_X$ and $S \in \mathcal{G}_X$, then $S^{-1}TS \in \mathcal{W}_X$. With the note that there are weakly mixing extensions of antiperiodic factors, we then use Lemma 2.4 to conclude $\mathcal{W}_X$ is dense.

We need to prove that there exists a subsequence of

$$\begin{aligned}
&\left\|\mathbb{E}((S^{-1}TS)^n f \cdot \overline{g}|X) - ((S')^{-1}T'S')^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\right\|_{L^2(X)} \\
&= \left\|\mathbb{E}(S^{-1}T^n S f \cdot \overline{g}|X) - (S')^{-1}(T')^n S' \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\right\|_{L^2(X)}
\end{aligned}$$

which converges to 0. Indeed, the above is equal to

$$\begin{aligned}
&\left\|\mathbb{E}(S^{-1}T^nSf\cdot(S^{-1}S)\overline{g}|X)-(S')^{-1}(T')^nS'\mathbb{E}(f|X)\cdot((S')^{-1}S')\mathbb{E}(\overline{g}|X)\right\|_{L^2(X)}\\
=&\left\|(S')^{-1}\mathbb{E}(T^n(Sf)\cdot(S\overline{g})|X)-(S')^{-1}(T')^n\mathbb{E}(Sf|X)\cdot\mathbb{E}(S\overline{g}|X)\right\|\\
\leq&\left\|(S')^{-1}\right\|\left\|\mathbb{E}(T^n(Sf)\cdot(S\overline{g})|X)-(T')^n\mathbb{E}(Sf|X)\cdot\mathbb{E}(S\overline{g}|X)\right\|\\
=&\left\|\mathbb{E}(T^n(Sf)\cdot(S\overline{g})|X)-(T')^n\mathbb{E}(Sf|X)\cdot\mathbb{E}(S\overline{g}|X)\right\|.
\end{aligned}$$

But T is a weakly mixing extension of T' so there exists a subsequence for which the above converges to 0.

To prove that $\mathcal{W}_X$ is G_δ, let $\{f_i\}\subset L^2(Z|X)$ be dense with respect to $L^2(Z)$ and for $i,j,k,n\in\mathbb{N}$, consider the sets

$$A_{i,j,k,n}:=\left\{S\in\mathcal{G}_X:\left\|\mathbb{E}(S^nf_i\cdot\overline{f_j}|X)-(S')^n\mathbb{E}(f_i|X)\cdot\mathbb{E}(\overline{f_j}|X)\right\|_{L^2(X)}<\frac{1}{k}\right\}.$$

Due to Lemmas 2.5, 2.6, we see that $\bigcap_{i,j,k}\bigcup_{n\geq k}A_{i,j,k,n}=\mathcal{W}_X$.

Thus it is sufficient to prove that each $A_{i,j,k,n}$ is open. For this, we show that for fixed $n\in\mathbb{N}$, $f,g\in L^2(Z|X)$ and $\varepsilon>0$, the set

$$\{S\in\mathcal{G}_X:\|\mathbb{E}(S^nf\cdot\overline{g}|X)-(S')^n\mathbb{E}(f|X)\cdot\mathbb{E}(\overline{g}|X)\|<\varepsilon\}$$

is open in the weak topology. To this end, we show that the complement

$$V(n,f,g,\varepsilon):=\{S\in\mathcal{G}_X:\|\mathbb{E}(S^nf\cdot\overline{g}|X)-(S')^n\mathbb{E}(f|X)\cdot\mathbb{E}(\overline{g}|X)\|\geq\varepsilon\}$$

is closed. Let $(S_m)\subset V(n,f,g,\varepsilon)$ be a sequence of Koopman operators with (S_m) converging weakly to a Koopman operator S. Note that this implies that $S_m\to S$ strongly (as Koopman operators are all isometries).

First note that in general, if we have functions $g,h,h_1,h_2,\ldots\in L^2(Z|X)$, and $h_m\to h$ in $L^2(Z)$, then $\mathbb{E}(h_m\cdot g|X)\to\mathbb{E}(h\cdot g|X)$ in $L^2(X)$. Indeed, by Proposition 2.4,

$$\begin{aligned}
\|\mathbb{E}(g\cdot h_m|X)-\mathbb{E}(g\cdot h|X)\|_{L^2(X)}&=\|\mathbb{E}(g\cdot(h-h_m)|X)\|_{L^2(X)}\\
&\leq\left\|\|g\|_{L^2(Z|X)}\right\|_{L^\infty(X)}\|h_m-h\|_{L^2(Z)}.
\end{aligned}$$

Second, note that $S_m\to S$ strongly implies $(S_m)'\to S'$ strongly. To see this, let $h\in L^2(X)$ and let $\hat{h}\in L^2(Z)$ be defined so that $\hat{h}(x,y):=h(x)$ (that is, $\hat{h}$ is constant on fibers). Then by Fubini, $\|(S_m)'h-S'h\|_{L^2(X)}=\left\|S_m\hat{h}-S\hat{h}\right\|_{L^2(Z)}$.

Lastly note that if $S_m\to S$ strongly, then $S_m^n\to S^n$ strongly. This is a specific case of the fact that if $T_n\to T,S_n\to S$ strongly (With T_n,S_n uniformly bounded), then $T_nS_n\to TS$ strongly . Indeed, for $f\in L^2(Z)$,

$$\begin{aligned}
\|(T_nS_n-TS)f\|&=\|(T_nS_n-T_nS+T_nS-TS)f\|\\
&\leq\|T_n(S_n-S)f\|+\|(T_n-T)Sf\|
\end{aligned}$$

Letting $\varepsilon > 0, N_1 \in \mathbb{N}$ be such that for $n > N_1, \|(S_n - S)f\| \leq \frac{\varepsilon}{2}$ (as $\|T_n\| = 1$ for all n), N_2 such that for $n > N_2, \|(T_n - T)Sf\| < \frac{\varepsilon}{2}$, we get $\|(T_nS_n - TS)f\| < \varepsilon$ for $n > \max N_1, N_2$.

With these facts, we see that $\mathbb{E}(S_m^n f \cdot \overline{g}|X) - (S'_m)^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)$ converges to $\mathbb{E}(S^n f \cdot \overline{g}|X) - (S')^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)$ strongly (consider $h_m := S_m^n f$ in our above observation). Thus, $S \in V(n, f, g, \varepsilon)$, and so $V(n, f, g, \varepsilon)$ is closed. □

Remark 2.4. Our assumption that (X, m) was the Lebesgue measure on the unit interval was only important for the proof of density, where we need a non-atomic probability space to have an antiperiodic factor. Proving $\mathcal{W}_X$ is G_δ never required anything of X, and so the proof will hold for (X, m) being replaced by any probability space.

2.1.7 General Case for the Fiber Space

There is still a case left to consider. Namely the case where the "vertical" measure is neither purely non-atomic nor purely discrete. Let (X, m) be as before, and let (Y, η) be a probability space with $Y = A \dot{\cup} B$, where B is an at most countable set, each of which is an atom of η, and $\eta \restriction_A$ is non-atomic (in particular, $0 < \eta(A) < 1$). Let $(Z, \mu, T) = (X \times Y, m \times \eta, T)$, where $T \in \mathcal{G}_X$ is an extension of T'. Let $C := X \times A$ and $D := X \times B$

We first show that T cannot mix points on the discrete and non-discrete parts of Y.

Proposition 2.5. *Let $C, D \subset Z$ be as defined above. Then up to a set of measure zero, $TC \subset C$ and $TD \subset D$.*

Proof. First, suppose there exists $D' \subset D$ such that $\mu(D') > 0$ and $TD' \subset C$. We can assume without loss of generality that there exists k, a level of D, such that $D' \subset k$ (see Section 2.1.2 for an explanation of this notation). We claim that $m(TD') = 0$, so that T is not measure-preserving. Note that because T is an extension of invertible T' and D' is contained to a single level of D, then for each fixed $x \in X, TD' \cap \pi^{-1}(x)$ contains at most one point. Therefore by Fubini

$$\mu(TD') = \int_Z \chi_{TD'} d\mu = \int_X \left(\int_Y \chi_{TD'}(x, y) d\eta \right) dm.$$

But by our previous note and the fact that $(A, \eta \restriction_A)$ is non-atomic, $\int_Y \chi_{TD'}(x, y) d\eta = 0$ for all x, and thus $\mu(TD') = 0$.

Now suppose there exists $C' \subset C$ such that $\mu(C') > 0$ and $TC' \subset D$. Note that there exists $x_0 \in X$ such that $\pi^{-1}(x_0) \cap C'$ is uncountable (else $\mu(C') = 0$ with a similar argument as above). But $T(\pi^{-1}(x_0) \cap C') \subset T'x_0 \times B$ is a countable set. This contradicts the invertibility of T. □

A quick consequence of Proposition 2.5 is that there are no weakly mixing extensions on Z.

Corollary 2.2. *Let $(X, m, T'), (Y, \eta), (Z, \mu, T), C$, and D be as defined above, with A and B, continuous and discrete parts of Y, respectively, nonempty. Then $T \notin \mathcal{W}_X$.*

Proof. Note that we can assume without loss of generality that $TC \subset C$ and $TD \subset D$ (that is, not up to a set of measure 0, but rather everywhere). Define $f(z)$ as

$$f(z) := \begin{cases} \dfrac{1}{\eta(A)} & \text{if } z \in C \\ \dfrac{-1}{\eta(B)} & \text{if } z \in D \end{cases}.$$

Clearly $f \in L^2(Z|X)$ and simple calculation shows that f has relative mean zero. Further, by Proposition 2.5, f is T-invariant, and so

$$\mathbb{E}(T^n f \cdot \overline{f}|X) = \mathbb{E}(f \cdot \overline{f}|X) = \mathbb{E}(|f|^2 |X) > 0.$$

Therefore, T is not a weakly mixing extension of T'. □

Remark 2.5. As with the G_δ part of Theorem 2.5, Proposition 2.5 and Corollary 2.2 hold when (X, m) is replaced with any standard probability space.

2.1.8 Strongly Mixing Extensions

In this section we first extend the notion of strongly mixing transformations to extensions, just as the notions of ergodic and weakly mixing transformations were extended to extensions. Afterwards we will show that the set of strongly mixing extensions form a set of first category in $\mathcal{G}_X$.

Definition 2.4. Let $(X, \nu), (Z, \mu)$ be probability spaces. We say that $T \in \mathcal{G}_X$ is a *(strongly) mixing extension* of T' or T is *(strongly) mixing relative to* T' if for all $f, g \in L^2(Z|X)$,

$$\lim_{n\to\infty} \|\mathbb{E}(T^n f \cdot \overline{g}|X) - (T')^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0.$$

Let $\mathcal{S}_X \subset \mathcal{G}_X$ denote the set of strongly mixing extensions.

Definition 2.4 yields some of the properties one would hope to have from the extension of the notion of strongly mixing transformations. For example, $\mathcal{S}_X$ is in general not empty. Indeed, any direct product transformation where the second component is strongly mixing will be a strongly mixing extension. We also have that if X is a single point, then the definition coincides with classical strongly mixing transformations. Further, it is clear that $\mathcal{S}_X \subset \mathcal{W}_X$.

We once again return to the case where (X, m) is the unit interval with the Lebesgue measure, and (Z, m) is the unit square with the Lebesgue measure. Analogous to Rokhlin's result and its proof, we now show that $\mathcal{S}_X$ is a first category subset of $\mathcal{G}_X$.

Theorem 2.6 (Strongly Mixing Extensions are of First Category)**.** *$\mathcal{S}_X \subset \mathcal{G}_X$ is of first category.*

Proof. For $k \in \mathbb{N}$, let $P_k := \{T \in \mathcal{G}_X | T^k = \mathbb{1}_Z\}$. Note in particular that if $T \in P_k$ then $(T')^k = \mathbb{1}_X$. For $n \in \mathbb{N}$, let $\hat{P}_n := \bigcup_{k>n} P_k$. Note that the Weak Approximation Theorem for Extensions (Theorem 2.3) implies that $\hat{P}_n$ is dense in $\mathcal{G}_X$.

Let $A := [0,1] \times [0,1/2]$ (the bottom half of Z). Note that $\mathbb{E}(\chi_A|X) = 1/2$ for all $x \in X$. We now define new sets,

$$M_k := \left\{ T \in \mathcal{G}_X \,\middle|\, \left\| \mathbb{E}(T^k \chi_A \cdot \chi_A | X) - (T')^k \mathbb{E}(\chi_A|X) \cdot \mathbb{E}(\chi_A|X) \right\|_{L^2(X)} \leq \frac{1}{5} \right\}.$$

Using the same arguments as used in the proof of Theorem 2.5 for the sets $V(n, f, g, \varepsilon)$, we see that M_k is closed for all k. Now let

$$M := \bigcup_{n=1}^{\infty} \bigcap_{k>n} M_k.$$

It is easy to see that $\mathcal{S}_X \subset M$. Indeed, if $T \in \mathcal{S}_X$, then for $f = g = \chi_A$ (in particular), for all $\varepsilon > 0$ there exists n such that for $k > n$,

$$\left\| \mathbb{E}(T^k f \cdot g | X) - (T')^k \mathbb{E}(f|X) \cdot \mathbb{E}(g|X) \right\|_{L^2(X)} \leq \varepsilon.$$

In particular, this holds for $\varepsilon = \dfrac{1}{5}$.

Thus it is sufficient to show that M is of first category. It is in turn sufficient to show that $\bigcap_{k>n} M_k$ is nowhere dense for all n, and further, given that $\bigcap_{k>n} M_k$ is closed for all n, it is thus sufficient to show that $\mathcal{G}_X \backslash \bigcap_{k>n} M_k$ is dense. Lastly, as

$$\mathcal{G}_X \backslash \bigcap_{k>n} M_k = \bigcup_{k>n} (\mathcal{G}_X \backslash M_k),$$

it will suffice to show that $P_k \subset (\mathcal{G}_X \backslash M_k)$ for all k, as then

$$\hat{P}_n = \bigcup_{k>n} P_k \subset \bigcup_{k>n} (\mathcal{G}_X \backslash M_k)$$

and $\hat{P}_n$ is dense.

Now, if $T \in P_k$, then $T^k = \mathbb{1}_Z, (T')^k = \mathbb{1}_X$, so

$$\begin{aligned} &\left\| \mathbb{E}(T^k \chi_A \cdot \chi_A | X) - (T')^k \mathbb{E}(\chi_A|X) \cdot \mathbb{E}(\chi_A|X) \right\|_{L^2(X)} \\ &= \left\| \mathbb{E}(\chi_A|X) - \mathbb{E}(\chi_A|X)^2 \right\| = \left\| \frac{1}{2} - \frac{1}{4} \right\| = \frac{1}{4} > \frac{1}{5}. \end{aligned}$$

Thus, $T \notin M_k$. □

Corollary 2.3. *$\mathcal{S}_X$ is a proper subset of $\mathcal{W}_X$.*

2.2 Fixed Factor

While Theorem 2.5 is a step forward, it does not answer Tao's question. For that, we need to return to the situation of a fixed factor transformation. As it turns out, though, it did not take long to bridge this gap. Shortly after the work in Section 2.1 was made publicly available, Glasner and Weiss, having seen the work, were able to show that a generic extension of an ergodic transformation is weakly mixing [28]. However, their proof actually showed a more general statement of density: they showed that the set of conjugations of an extension T of an ergodic transformation T_0 by extensions of the identity is dense in the set of all extensions of T_0. Indeed, Lemma 2.8 is an analogue of Lemmas 1.5 and 2.4. This section serves to present Glasner and Weiss' result, as well as discuss other considerations, such as strongly mixing and rigid extensions.

2.2.1 Notation and Definitions

As in Section 2.1, we let (X, m) be the unit interval with the Lebesgue measure, and $\mathcal{G}(X)$ denote the set of invertible measure-preserving transformations on X. We will also have (Y, ν) denote the unit interval with the Lebesgue measure, and let $(Z, \mu) := (X \times Y, m \times \nu)$. For $T_0 \in \mathcal{G}(X)$, denote by $\mathcal{G}_{T_0}$ the subset of $\mathcal{G}(Z)$ which are extensions of T_0 (through the natural projection onto X). Note that we will also use T_0 to represent the transformation $T_0 \times \mathbb{1}_Y$. The topology on $\mathcal{G}(X)$ and $\mathcal{G}_{T_0}$ will always be assumed to be the weak topology. For $T \in \mathcal{G}_{T_0}$, let $\{T_x\}_{x \in X} \subset \mathcal{G}(Y)$ be such that $T(x, y) = (T_0 x, T_x y)$. Note that we have changed the notation for representing the factor from T' in Section 2.1 to T_0 in this section to emphasize the difference between having a non-fixed and fixed factor transformation, respectively.

For $p \in \mathbb{N}, p > 1$, a p-adic interval of rank k in Y is an interval of the form

$$\left(\frac{j}{p^k}, \frac{j+1}{p^k}\right), 0 \leq j < p^k,$$

and a p-adic set is a union of p-adic intervals. A p-adic permutation (of rank k) is a permutatiion of the p-adic intervals of rank k.

Lemma 2.7. *For all $p > 1$, the p-adic permuations are dense in $\mathcal{G}(Y)$.*

The proof is nearly identical to that of Theorem 1.1, but by replacing each instance of "dyadic" with "p-adic" in the proof (one must also do the same thing for Lemma 1.1 and Proposition 1.1).

Let d denote a metric on $\mathcal{G}(Y)$ which induces the weak topology on $\mathcal{G}(Y)$. In particular, for $T, S \in \mathcal{G}(Y)$, we will have d be of the form

$$d(T, S) = \sum_{n=1}^{\infty} \frac{a_n}{2^{k_n}} \left(\nu(TD_n \triangle SD_n) + \nu(T^{-1}D_n \triangle S^{-1}D_n)\right),$$

where D_n are all dyadic intervals of rank at least 1, k_n is the rank of D_n and

$$\sum_{n=1}^{\infty} a_n < \infty.$$

Recall that in the classical theory, the weights $a_n/2^{k_n}$ are simply $1/2^n$ (see Section 1.1.1). Now note that d induces a metric on $\mathcal{G}_{T_0}, d^*$. For $T, S \in \mathcal{G}_{T_0}, d^*$ is given by

$$d^*(T,S) := \int_X d(T_x, S_x)dm.$$

Proposition 2.6. *The metric d^* induces the weak topology on $\mathcal{G}_{T_0}$.*

Proof. Note that for all dyadic intervals $D_j \subset Y$,

$$\int_X \nu(T_x D_j \triangle S_x D_j)dm = \sum_{i=1}^{2^{k_n}} \left(\mu(TD_{i,j} \triangle SD_{i,j}) + \mu(T^{-1}D_{i,j} \triangle S^{-1}D_{i,j})\right),$$

where $D_{i,j}$ are dyadic squares in Z of rank k_j whose projection onto Y are D_j. We get then

$$\int_X d(T_x, S_x)dm = \sum_{j=1}^{\infty}\sum_{i=1}^{2^{k_j}} \frac{a_j}{2^{k_j}} \left(\mu(TD_{i,j} \triangle SD_{i,j}) + \mu(T^{-1}D_{i,j} \triangle S^{-1}D_{i,j})\right),$$

with

$$\sum_{j=1}^{\infty}\sum_{i=1}^{2^{k_j}} \frac{a_j}{2^{k_j}} = \sum_{n=1}^{\infty} a_n < \infty.$$

Let $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be a bijection. If $n = f(i,j)$, set $b_n := a_j/2^{k_j}$. Then

$$d^*(T,S) = \sum_{n=1}^{\infty} b_n \left(\mu(TD_n \triangle SD_n) + \mu(T^{-1}D_n \triangle S^{-1}D_n)\right),$$

with

$$\sum_{n=1}^{\infty} b_n < \infty,$$

i.e., d^* induces the weak topology. □

We next recall the definition of a strongly mixing extension from Section 2.4.

Definition. We say that $T \in \mathcal{G}_{T_0}$ is a *(strongly) mixing extension* of T_0 or T is *(strongly) mixing relative to* T_0 if for all $f, g \in L^2(Z|X)$,

$$\lim_{n\to\infty} \|\mathbb{E}(T^n f \cdot \overline{g}|X) - T_0^n \mathbb{E}(f|X)\mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0.$$

Let $\mathcal{S}_{T_0} \subset \mathcal{G}_{T_0}$ denote the set of strongly mixing extensions of T_0. Lastly for definition reminders, we have the following, from [28].

Definition 2.5. Let $T \in \mathcal{G}_{T_0}$. We say that T is *a piecewise constant skew product* over T_0 if there exists a partition of $X, \{A_1, \ldots, A_N\}$, and $\{R_1, \ldots, R_N\} \subset \mathcal{G}(Y)$ such that for $1 \le j \le N$ and for all $x \in A_j, T_x = R_j$.

Of particular note is that the set of piecewise constanst skew products over T_0 is dense in $\mathcal{G}_{T_0}$. For the reader's convenience, we provide the proof below, cf. [28].

Proposition 2.7 (Glasner, Weiss)**.** *Let $T_0 \in \mathcal{G}(X)$ be fixed. Then the set of piecewise constant skew products over T_0 is dense in $\mathcal{G}_{T_0}$.*

Proof. Let $\rho : X \to \mathcal{G}(Y)$ be a measurable map such that $\rho(x) = T_x$ and let $\rho_*(m) = m \circ \rho$ denote the pushforward measure. As $\rho_*(m)$ is a regular measure on $\mathcal{G}(Y)$, for a fixed $\varepsilon > 0$, there exists a compact set $K \subset \mathcal{G}(Y)$ with $\rho_*(m)(K) > 1 - \varepsilon$.

Let $\{C_i | 1 \leq i \leq N\}$ be a finite partition of K and let C_0 be the complement of K. For $1 \leq i \leq N$, let $A_i := \rho^{-1}(C_i)$, and let $A_0 := \rho^{-1}(C_0)$.

We now define S, a piecewise constant skew product over T_0, as follows: if $x \in A_0, S_x := \mathbb{1}_Y$, and if $x \in A_i, S_x := R_i$ where $R_i \in C_i$ is fixed for each i. Then

$$\begin{aligned} d^*(T,S) &= \int_X d(T_x, S_x) dm \\ &= \int_{A_0} d(T_x, S_x) dm + \sum_{i=1}^{N} \int_{A_i} d(T_x, S_x) dm \\ &\leq \varepsilon + \sum_{i=1}^{N} \int_{A_i} \varepsilon \, dm \leq \varepsilon + \varepsilon \rho_*(m)(K) < 2\varepsilon. \end{aligned}$$

□

2.2.2 Conjugacy Lemma for Weakly Mixing Extensions

We begin exploring category theorems in $\mathcal{G}_{T_0}$ with the following result of Glasner and Weiss [28]. While the result and its proof (including much of their notation) can be found in their note, we present it with slightly more generality.

Lemma 2.8. *Let $T_0 \in \mathcal{G}(X)$ be antiperiodic, and $\hat{T} \in \mathcal{G}_{T_0}$. Then $\{S^{-1}\hat{T}S | S \in \mathcal{G}_{\mathbb{1}_X}\}$ is dense in $\mathcal{G}_{T_0}$.*

Proof. Let $T \in \mathcal{G}_{T_0}$ and let $\varepsilon > 0$, and $N_\varepsilon(T)$ a dyadic neighborhood of T. We wish to find $S \in \mathcal{G}_{\mathbb{1}_X}$ such that $S^{-1}\hat{T}S \in N_\varepsilon(T)$. Without loss of generality, we can assume that T is a piecewise constant skew product over T_0. Let $\{A_1, \ldots, A_k\}$ be a partition of X and $\{R_1, \ldots, R_k\} \subset \mathcal{G}(Y)$ such that for all j, and all $x \in A_j, T_x = R_j$.

Now let $\{B, T_0B, \ldots, T_0^{n-1}B\}$ be a Rokhlin tower with respect to $\varepsilon/2$, and $1/n < \varepsilon/2$, that is, $B, T_0B, \ldots, T_0^{n-1}B$ are pairwise disjoint, and

$$m\left(\bigcup_{i=0}^{n-1} T_0^i B\right) > 1 - \varepsilon.$$

We now refine the tower with respect to the partition of $X, \{A_j | 1 \leq j \leq k\}$. This will create a partition of B into sets $\{B_l | 1 \leq l \leq L\}$ such that for all i, l, there exists $\alpha(l,i)$ between 1 and k such that $T_0^i B_l \subset A_{\alpha(l,i)}$.

We will now define $S \in \mathcal{G}_{\mathbb{1}_X}$, a piecewise constant skew product. First, for all $x \in B$ define $S_x := \mathbb{1}_Y$. Then, inductively, having defined S for $0 \leq i < n-1$ for all l, define S_x for $x \in T^{i+1}B_l$ (for all l) to be $S_x := \hat{T}_{T_0^{-1}x} S_{T_0^{-1}x} R^{-1}_{\alpha(l,i)}$. Note that this implies $S_x^{-1}\hat{T}_{T_0^{-1}x} S_{T_0^{-1}x} = R_{\alpha(l,i)}$. This defines S on the fibers corresponding to the Rokhlin tower. Finally, if $x \notin \bigcup_{i=0}^{n-1} T_0^i B$, define $S_x := \mathbb{1}_Y$.

We see that by how we constructed S, if $x \in \bigcup_{i=0}^{n-2} T_0^i B$ then $T_x := R_x = (S^{-1}\hat{T}S)_x$. Thus it is easy to see that for any measurable set $E \in Z$,

$$\mu(TE \triangle S^{-1}\hat{T}SE) \leq m\left(X \setminus \bigcup_{i=0}^{n-1} T_0^i B\right) + m(T^{n-1}B) < \frac{\varepsilon}{2} + \frac{1}{n} < \varepsilon,$$

so certainly $S^{-1}\hat{T}S \in N_\varepsilon(T)$. □

We note the similarities between Lemmas 1.5, 2.4, and 2.8. Just as Lemma 1.5 was a useful tool for proving category results in the case of measure-preserving transformations, Lemma 2.8 is equally useful in proving category results for extensions. The first of these results, as was the case classically, is for weakly mixing extensions. The result can also be attributed to Glasner and Weiss in [28], but the proof is slightly different.

Theorem 2.7. *Let T_0 be antiperiodic and let $\mathcal{W}_{T_0}$ denote the set of weakly mixing extensions of T_0. Then $\mathcal{W}_{T_0}$ is a dense, G_δ subset of $\mathcal{G}_{T_0}$.*

Proof. Let $T \in \mathcal{W}_{T_0}$ and $S \in \mathcal{G}_{\mathbb{1}_X}$. We wish to prove that $S^{-1}TS \in \mathcal{W}_{T_0}$ as then by Lemma 2.8, $\mathcal{W}_{T_0}$ is dense in $\mathcal{G}_{T_0}$. Thus, by Lemma 2.6, we need to show that there exists a subsequence n_k such that for all $f, g \in L^2(Z|X)$,

$$\lim_{k\to\infty} \left\|\mathbb{E}((S^{-1}TS)^{n_k} f \cdot \overline{g}|X) - T_0^{n_k}\mathbb{E}(f|X)\cdot\mathbb{E}(\overline{g}|X)\right\|_{L^2(X)} = 0. \tag{2.10}$$

Note that

$$\begin{aligned}
&\left\|\mathbb{E}((S^{-1}TS)^n f \cdot \overline{g}|X) - T_0^n\mathbb{E}(f|X)\cdot\mathbb{E}(\overline{g}|X)\right\|_{L^2(X)} \\
&= \left\|\mathbb{E}(S^{-1}T^n Sf \cdot \overline{g}|X) - T_0^n\mathbb{E}(f|X)\cdot\mathbb{E}(\overline{g}|X)\right\| \\
&= \left\|\mathbb{E}(S^{-1}T^n Sf \cdot (S^{-1}S)\overline{g}|X) - T_0^n\mathbb{E}(Sf|X)\cdot\mathbb{E}(S\overline{g}|X)\right\| \\
&= \left\|\mathbb{E}(S^{-1}(T^n Sf \cdot S\overline{g})|X) - T_0^n\mathbb{E}(Sf|X)\cdot\mathbb{E}(S\overline{g}|X)\right\| \\
&= \left\|\mathbb{E}(T^n(Sf)\cdot(S\overline{g})|X) - T_0^n\mathbb{E}(Sf|X)\cdot\mathbb{E}(S\overline{g}|X)\right\|
\end{aligned}$$

But as $T \in \mathcal{W}_{T_0}$, we know that this converges to 0 along the subsequence n_k which corresponds to T. Thus, $S^{-1}TS \in \mathcal{W}_{T_0}$.

Now let $\{f_i\}$ be an $L^2(Z)$-dense subset of $L^2(Z|X)$ and define

$$U_{i,j,k,n} := \left\{T \in \mathcal{G}_{T_0} \,\middle|\, \left\|\mathbb{E}(T^n f_i \cdot \overline{f_j}|X) - T_0^n\mathbb{E}(f_i|X)\cdot\mathbb{E}(\overline{f_j}|X)\right\|_{L^2(X)} < \frac{1}{k}\right\}.$$

Lemmas 2.5 and 2.6 show that

$$\mathcal{W}_{T_0} = \bigcap_{i,j,k} \bigcup_{n>k} U_{i,j,k,n},$$

so we need only show $U_{i,j,k,n}$ are open. To this end, we instead show that

$$V(n, f, g, \varepsilon) := \{S \in \mathcal{G}_{T_0} : \|\mathbb{E}(S^n f \cdot \overline{g}|X) - T_0^n \mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\| \geq \varepsilon\}$$

is closed. Let $(S_m) \subset V(n, f, g, \varepsilon)$ be a sequence of Koopman operators with (S_m) converging weakly to a Koopman operator S. Note that this implies that $S_m \to S$ strongly (as Koopman operators are all isometries).

We claim that if we have functions $g, h, h_1, h_2, \ldots \in L^2(Z|X)$, and $h_m \to h$ in $L^2(Z)$, then $\mathbb{E}(h_m g|X) \to \mathbb{E}(hg|X)$ in $L^2(X)$. Indeed,

$$\begin{aligned} \|\mathbb{E}(g \cdot h_m|X) - \mathbb{E}(g \cdot h|X)\|_{L^2(X)} &= \|\mathbb{E}(g \cdot (h - h_m)|X)\|_{L^2(X)} \\ &\leq \left\| \|g\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|h_m - h\|_{L^2(Z)}. \end{aligned}$$

Lastly note that if $S_m \to S$ strongly, then $S_m^n \to S^n$ strongly.

With these facts, we see that $\mathbb{E}(S_m^n f \cdot \overline{g}|X) - \mathbb{E}(T_0^n f \cdot \overline{g}|X)$ converges to $\mathbb{E}(S^n f \cdot \overline{g}|X) - \mathbb{E}(T_0^n f \cdot \overline{g}|X)$ strongly (consider $h_m := S_m^n f$ in our above observation). Thus, $S \in V(n, f, g, \varepsilon)$, and so $V(n, f, g, \varepsilon)$ is closed. □

2.2.3 Strongly Mixing Extensions

To fit the classical theory, we would like to prove that for all $T_0, \mathcal{S}_{T_0}$ is a first category subset of $\mathcal{G}_{T_0}$. Unfortunately, this goal proves to be more difficult than one would expect. We will instead be able to show that for a set of T_0 which is generic in $\mathcal{G}(X), \mathcal{S}_{T_0}$ is of first category. Indeed, we begin with the following sufficient condition.

Lemma 2.9. *Let $T_0 \in \mathcal{G}(X)$ be fixed. If $\mathcal{R}(Z) \cap \mathcal{G}_{T_0}$ is dense in $\mathcal{G}_{T_0}$, then $\mathcal{S}_{T_0}$ is a first category subset of $\mathcal{G}_{T_0}$.*

Proof. Let $A := X \times [0, 1/2]$ and for all $k \in \mathbb{N}$ define

$$P_k' := \left\{ T \in \mathcal{G}_{T_0} | \mu(T^k A \cap A) > \frac{9}{20} \right\}.$$

Note that for all n, $P_n := \bigcup_{k>n} P_k'$ is dense in $\mathcal{G}_{T_0}$ as $\mathcal{R}(Z) \cap \mathcal{G}_{T_0} \subset P_n$ for all n.

Now define

$$M_k' := \left\{ T \in \mathcal{G}_{T_0} | \left\| \mathbb{E}(T^k \chi_A \cdot \chi_A|X) - \frac{1}{4} \right\|_{L^2(X)} \leq \frac{1}{5} \right\},$$

and $M_n := \bigcap_{k>n} M'_k$. Note that as $\mathbb{E}(\chi_A|X) \equiv 1/2, \mathcal{S}_{T_0} \subset \bigcup_n M_n$. Further, as M'_k is closed, to prove $\mathcal{S}_{T_0}$ is of first category, it is sufficient to prove M_n is dense for all n. But this is true as $P'_k \subset (M'_k)^c$ for all k. Indeed, if $T \in P'_k$, then

$$\begin{aligned}&\left\|\mathbb{E}(T^k\chi_A \cdot \chi_A|X) - 1/4\right\|_{L^2(X)} \geq \left\|\mathbb{E}(T^k\chi_A \cdot \chi_A|X) - 1/4\right\|_{L^1(X)}\\ &= \left\|T^k\chi_A \cdot \chi_A\right\|_{L^1(Z)} - 1/4 > 9/20 - 1/4 = 1/5.\end{aligned}$$

□

While Lemma 2.9 is a nice sufficient condition, it certainly cannot be applied for all factors T_0. Indeed, if T_0 is strongly mixing, if $T \in \mathcal{G}_{T_0}$ and B is any X-measurable set, then for no subsequence n_k does $T^{n_k}\chi_B \to \chi_B$. All the same, it will be an invaluable tool, as we will see in the following two results.

Theorem 2.8. *If T_0 is periodic, then $\mathcal{S}_{T_0}$ is a first category subset of $\mathcal{G}_{T_0}$.*

Proof. By Lemma 2.9, we need only show that $\mathcal{R}(Z) \cap \mathcal{G}_{T_0}$ is dense in $\mathcal{G}_{T_0}$. Let p be the smallest integer greater than 1 such that $T_0^p = \mathbb{1}_X$. Further fix $T \in \mathcal{G}_{T_0}, \varepsilon > 0$, and let $N_\varepsilon(T)$ be a p-adic neighborhood in the weak topology. That is,

$$N_\varepsilon(T) := \{S \in \mathcal{G}_{T_0}|\mu(TD_{ij}\triangle SD_{ij}) < \varepsilon, 1 \leq i, j \leq p^K\}$$

where D_{ij} are (without loss of generality) every p-adic square of rank $K \in \mathbb{N}$. Let $\{E_i\}, \{F_i\}$ be dyadic intervals of rank K in X and Y respectively such that $D_{ij} = E_i \times F_j$.

By Proposition 2.7, there exists $S \in N_{\varepsilon/2}(T)$ that is a piecewise constant skew product over T_0. Let $\{A_1, \dots A_N\}$ partition X and $\{R_1, \dots R_N\} \subset \mathcal{G}(Y)$ such that for all $x \in A_k, S_x = R_k$.

Now for $V \in \mathcal{G}(Y)$, let

$$N'_\varepsilon(V) := \{U \in \mathcal{G}(Y)|\nu(VF_i\triangle UF_i) < \varepsilon, 1 \leq i \leq p^K\},$$

and for each R_k, let $P_k \in N'_{\varepsilon/(2Nm(A_k))}(R_k)$ be a p-adic permutation. Define a new piecewise constant extension of T_0, Q such that for all $x \in A_k, Q_x = P_k$. Fix D_{ij}. Then by Fubini,

$$\begin{aligned}&\mu(SD_{ij}\triangle QD_{ij}) = \int_{E_i} \nu(S_xF_j\triangle Q_xF_j)dx\\ &= \sum_{k=1}^N \nu(R_kF_j\triangle P_kF_j)m(A_k \cap E_i) < \sum_{k=1}^N \frac{\varepsilon}{2Nm(A_k)}m(A_k \cap E_i) \leq \frac{\varepsilon}{2}.\end{aligned}$$

Thus,

$$\mu(TD_{ij}\triangle QD_{ij}) \leq \mu(TD_{ij}\triangle SD_{ij}) + \mu(SD_{ij}\triangle QD_{ij}) < \varepsilon.$$

As this holds for all $i, j, Q \in N_\varepsilon(T)$. Let M be the maximum of the ranks of all P_k. Then $Q^{p^{M+1}} = \mathbb{1}_Z$, so $Q \in \mathcal{R}(Z) \cap \mathcal{G}_{T_0}$. □

Theorem 2.9. *Let $T_0 \in \mathcal{G}(X)$ be antiperiodic and rigid. Then $\mathcal{S}_{T_0}$ is a first category subset of $\mathcal{G}_{T_0}$.*

Proof. Let $T \in \mathcal{R}(Z) \cap \mathcal{G}_{T_0}$. We know this set is nonempty as the rigidity of T_0 on X implies the rigidity of $T_0 \times \mathbb{1}_Y$ on Z. Now consider the set

$$\{S^{-1}TS | S \in \mathcal{G}_{\mathbb{1}_X}\}.$$

By Lemma 2.8, this set is dense in $\mathcal{G}_{T_0}$. Further, as S can be viewed as an element of $\mathcal{G}(Z)$, we know that for each $S, S^{-1}TS \in \mathcal{R}(Z)$. Thus by Lemma 2.9, $\mathcal{S}_{T_0}$ is of first category. □

To close this section we show that, though it is not completely clear for exactly which set of transformations T_0 we have $\mathcal{S}_{T_0}$ is of first category, we are able to show a step in that direction.

Proposition 2.8. *Let $T_0 \in \mathcal{G}(X)$ be antiperiodic. Then $\mathcal{S}^c_{T_0}$ is dense in $\mathcal{G}_{T_0}$.*

Proof. Let $\mathcal{A}$ denote the measurable subsets of Z such that $A \in \mathcal{A}$ if and only if $\mathbb{E}(\chi_A|X) \equiv 1/2$. Further, define

$$\mathcal{A}_{T_0} := \left\{ T \in \mathcal{G}_{T_0} | \ \exists\, (n_k), A \in \mathcal{A} \ s.t. \ \left\| \mathbb{E}(T^{n_k}\chi_A \cdot \chi_A|X) - \frac{1}{4} \right\| \geq \frac{1}{5} \ \forall k \right\}.$$

Note that for all $T_0, \mathcal{A}_{T_0}$ is nonempty, as for $T = T_0 \times \mathbb{1}_Y$ we can take A as in the proof of Lemma 2.9. Further note that $\mathcal{A}_{T_0} \subset \mathcal{S}^c_{T_0}$. Now fix $T \in \mathcal{A}_{T_0}$, as well as $A \in \mathcal{A}$, and (n_k) corresponding to T. Now fix $S \in \mathcal{G}_{\mathbb{1}_X}$ and define $f := S^{-1}\chi_A$. Then we have

$$\mathbb{E}((S^{-1}TS)^{n_k} f \cdot f|X) = \mathbb{E}(T^{n_k}(Sf) \cdot (Sf)|X) = \mathbb{E}(T^{n_k}\chi_A \cdot \chi_A|X).$$

Thus we see that $S^{-1}TS \in \mathcal{A}_{T_0}$. By Theorem 2.8, $\mathcal{A}_{T_0}$ is dense in $\mathcal{G}_{T_0}$. □

2.2.4 Rigid Extensions and Other Relativizations

Recall from Chapter 1 that in [36], Katok and Stepin showed a generalization of Rokhlin's result that strongly mixing transformations are of first category, namely that rigid transformations are residual. Indeed, if a transformation is rigid, then it is not strongly mixing. One might think to try to show that the set of rigid extensions of T_0 are a dense, G_δ set. However, first one must know what it means for an extension to be rigid.

Definition 2.6. Let $T \in \mathcal{G}_{T_0}$. We say that T is *rigid relative to* T_0, or T is a *rigid extension* of T_0 if there exists a subsequence (n_k) such that for all $f, g \in L^2(Z|X)$,

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k} f \cdot \overline{g}|X) - \mathbb{E}(T_0^{n_k} f \cdot \overline{g}|X)\|_{L^2(X)} = 0.$$

Let $\mathcal{R}_{T_0} \subset \mathcal{G}_{T_0}$ denote the set of rigid extensions of T_0. Note that if X were a single point, then Proposition 1.4 shows that relative rigidity coincides with classical rigidity.

Proposition 2.9. *Let X be a single point. Then $\mathcal{R}_{T_0} = \mathcal{R}(Z)$.*

Proof. First of all note that if X is a single point then the only element of $\mathcal{G}(X)$ is $\mathbb{1}_X$ so we must have that $T_0 = \mathbb{1}_X$. Further, $L^2(Z|X) = L^2(Z)$ as $\mathbb{E}(|f|^2 |X)^{1/2} \in L^\infty(X)$ if and only if

$$\left(\int_Y |f|^2 \, dm\right)^{1/2} = \|f\|_{L^2(Z)}$$

is finite. Thus, for $f, g \in L^2(Z)$,

$$\|\mathbb{E}(T^{n_k} f \cdot \overline{g}|X) - \mathbb{E}(T_0^{n_k} f \cdot \overline{g}|X)\|_{L^2(X)} = |\langle T^{n_k} f, g\rangle - \langle f, g\rangle| \, .$$

So $\mathcal{R}_{T_0} = \mathcal{R}(Z)$. □

Just as rigid transformations are mutually exclusive with strongly mixing transformations, we see that the same is true of rigid and strongly mixing extensions.

Lemma 2.10. *For all $T_0 \in \mathcal{G}(X), \mathcal{S}_{T_0} \subset \mathcal{R}_{T_0}^c$.*

Proof. Let $T \in \mathcal{S}_{T_0}$ and let $A = X \times [0, 1/2]$. Note then that $\mathbb{E}(\chi_A|X) \equiv 1/2$ and for all $n, T_0^n \mathbb{E}(\chi_A|X)\mathbb{E}(\chi_A|X) \equiv 1/4$, so,

$$\lim_{n\to\infty} \left\| \mathbb{E}(T^n \chi_A \cdot \chi_A|X) - \frac{1}{4} \right\|_{L^2(X)} = 0.$$

On the other hand, all $n, \mathbb{E}(T_0^n \chi_A \cdot \chi_A|X) \equiv 1/2$. Subsequently for all subsequence (n_k),

$$\lim_{k\to\infty} \left\| \mathbb{E}(T^{n_k} \chi_A \cdot \chi_A|X) - \frac{1}{2} \right\|_{L^2(X)} \geq \frac{1}{4} > 0.$$

□

In fact, we can say something even stronger. Letting $\mathcal{A}_{T_0}$ be as in the proof of Proposition 2.8, then it's easy to see that $\mathcal{R}_{T_0} \subset \mathcal{A}_{T_0}$. Thus we have $\mathcal{S}_{T_0} \subset \mathcal{A}_{T_0}^c \subset \mathcal{R}_{T_0}^c$.

With the additional note that $\mathcal{R}_{T_0}$ is always nonempty by taking $T_0 \times \mathbb{I}_Y$, it seems that $\mathcal{R}_{T_0}$ has all the properties we would expect from a relativization of rigidity. Unfortunately, one property may fail. Consider $S \in \mathcal{G}_{\mathbb{1}_X}$. It may well fail to be the case that for $T \in \mathcal{R}_{T_0}, S^{-1}TS \in \mathcal{R}_{T_0}$. Because of this, we cannot apply Lemma 2.8, and are unable to show $\mathcal{R}_{T_0}$ is dense. However, although we cannot show $\mathcal{R}_{T_0}$ is a dense, G_δ, we can still show it is a G_δ. First, we prove the following technical lemma.

Lemma 2.11. *Let $T \in \mathcal{G}_{T_0}$, and let $D \subset L^2(Z|X)$, with*

$$\left\| \|f\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \leq 1 \textit{ for all } f \in D,$$

such that D is dense in the unit ball of $L^2(Z|X)$, with respect to the $L^2(Z)$ norm topology. Further suppose that there exists a subsequence (n_k) such that for all $f_i, f_j \in D$,

$$\lim_{k\to\infty} \left\| \mathbb{E}(T^{n_k} f_i \cdot \overline{f_j}|X) - \mathbb{E}(T_0^{n_k} f_i \cdot \overline{f_j}|X) \right\|_{L^2(X)} = 0.$$

Then $T \in \mathcal{R}_{T_0}$.

Proof. Fix h, g in the unit ball in $L^2(Z|X)$ and let $\{h_j\}, \{g_j\} \subset D$ such that

$$h_j \xrightarrow{L^2(Z)} h, g_j \xrightarrow{L^2(Z)} g.$$

The proof of Lemma 2.5 shows that (uniformly in n) $\mathbb{E}(T^n h_j \cdot \overline{g_j}|X) \to \mathbb{E}(T^n h \cdot \overline{g}|X)$, and $\mathbb{E}(T_0^n h_j \cdot \overline{g_j}|X) \to \mathbb{E}(T_0^n h \cdot \overline{g}|X)$ in $L^2(X)$. For the reader's convenience, we reproduce the proof below. First for $\mathbb{E}(T^n h_j \overline{g_j}|X)$, we have

$$\begin{aligned} & \|\mathbb{E}(T^n h_j \cdot \overline{g_j}|X) - \mathbb{E}(T^n h \cdot \overline{g}|X)\|_{L^2(X)} \\ \leq\ & \|\mathbb{E}(T^n h_j \cdot \overline{g_j} - T^n h_j \cdot \overline{g}|X)\|_{L^2(X)} + \|\mathbb{E}(T^n h_j \cdot \overline{g} - T^n h \cdot \overline{g}|X)\|_{L^2(X)} \\ =\ & \|\mathbb{E}(T^n h_j \cdot (\overline{g_j} - \overline{g})|X)\|_{L^2(X)} + \|\mathbb{E}(\overline{g} \cdot (T^n h_j - T^n h)|X)\|_{L^2(X)}. \end{aligned}$$

Now by Proposition 2.4

$$\begin{aligned} \|\mathbb{E}(T^n h_j \cdot (\overline{g_j} - \overline{g})|X)\|_{L^2(X)} &\leq \left\| \|T^n h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(Z)}, \\ \|\mathbb{E}(\overline{g} \cdot (T^n h_j - T^n h)|X)\|_{L^2(X)} &\leq \left\| \|\overline{g}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|T^n(h - h_j)\|_{L^2(Z)}. \end{aligned}$$

In turn, as $\left\| \|T^n h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} = \left\| \|h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)}$ and T is an isometry, we get

$$\begin{aligned} \left\| \|T^n h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(Z)} &\leq \left\| \|h_j\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|\overline{g_j} - \overline{g}\|_{L^2(Z)}, \\ \left\| \|\overline{g}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|T^n(h - h_j)\|_{L^2(Z)} &\leq \left\| \|\overline{g}\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|h - h_j\|_{L^2(Z)}. \end{aligned}$$

Because $h_j \to h, g_j \to g$ in $L^2(Z)$, we have the desired result. Note that the proof that $\mathbb{E}(T_0^n h_j \overline{g_j}|X) \to \mathbb{E}(T_0^n h \overline{g}|X)$ is exactly the same, replacing T with T_0. Now,

$$\begin{aligned} & \|\mathbb{E}(T^n h \cdot \overline{g}|X) - \mathbb{E}(T_0^n h \cdot \overline{g}|X)\|_{L^2(X)} \\ \leq\ & \|\mathbb{E}(T^n h \cdot \overline{g}|X) - \mathbb{E}(T^n h_j \cdot \overline{g_j}|X)\|_{L^2(X)} \\ +\ & \|\mathbb{E}(T^n h_j \cdot \overline{g_j}|X) - \mathbb{E}(T_0^n h_j \cdot \overline{g_j}|X)\|_{L^2(X)} \\ +\ & \|\mathbb{E}(T_0^n h_j \cdot \overline{g_j}|X) - \mathbb{E}(T_0^n h \cdot \overline{g}|X)\|_{L^2(X)}. \end{aligned}$$

By hypothesis, there is a subsequence (n_k) (independent of j) such that the middle term converges to 0. Further, the first and third terms converge to 0 as $j \to \infty$ uniformly in n, so

$$\lim_{k \to \infty} \|\mathbb{E}(T^{n_k} h \cdot \overline{g}|X) - \mathbb{E}(T_0^{n_k} h \cdot \overline{g}|X)\|_{L^2(X)} = 0.$$

By scaling this will also hold for all $h, g \in L^2(Z|X)$, so $T \in \mathcal{R}_{T_0}$. □

Proposition 2.10. *Let $T_0 \in \mathcal{G}(X)$ be fixed. Then $\mathcal{R}_{T_0}$ is a G_δ set.*

Proof. Let $\{f_i\}$ be an $L^2(Z)$-dense subset of $L^2(Z|X)$ and define

$$U_{i,j,k,n} := \left\{ T \in \mathcal{G}_{T_0} \,|\, \left\| \mathbb{E}(T^n f_i \cdot \overline{f_j}|X) - \mathbb{E}(T_0^n f_i \cdot \overline{f_j}|X) \right\|_{L^2(X)} < \frac{1}{k} \right\}.$$

Lemma 2.11 shows that

$$\mathcal{R}_{T_0} = \bigcap_{i,j,k} \bigcup_{n>k} U_{i,j,k,n},$$

so we need only show $U_{i,j,k,n}$ are open. To this end, we instead show that

$$V(n, f, g, \varepsilon) := \{S \in \mathcal{G}_{T_0} : \|\mathbb{E}(S^n f \cdot \overline{g}|X) - \mathbb{E}(T_0^n f \cdot \overline{g}|X)\| \geq \varepsilon\}$$

is closed. Let $(S_m) \subset V(n, f, g, \varepsilon)$ be a sequence of Koopman operators with (S_m) converging weakly to a Koopman operator S. Note that this implies that $S_m \to S$ strongly (as Koopman operators are all isometries).

We claim that if we have functions $g, h, h_1, h_2, \ldots \in L^2(Z|X)$, and $h_m \to h$ in $L^2(Z)$, then $\mathbb{E}(h_m g|X) \to \mathbb{E}(hg|X)$ in $L^2(X)$. Indeed, we have

$$\begin{aligned}\|\mathbb{E}(g \cdot h_m|X) - \mathbb{E}(g \cdot h|X)\|_{L^2(X)} &= \|\mathbb{E}(g \cdot (h - h_m)|X)\|_{L^2(X)} \\ &\leq \left\| \|g\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \|h_m - h\|_{L^2(Z)}.\end{aligned}$$

Lastly note that if $S_m \to S$ strongly, then $S_m^n \to S^n$ strongly. This is a specific case of the fact that if $T_n \to T, S_n \to S$ strongly (With T_n, S_n uniformly bounded), then $T_n S_n \to TS$ strongly . Indeed, for $f \in L^2(Z)$,

$$\begin{aligned}\|(T_n S_n - TS)f\| &= \|(T_n S_n - T_n S + T_n S - TS)f\| \\ &\leq \|T_n(S_n - S)f\| + \|(T_n - T)Sf\|\end{aligned}$$

With these facts, we see that $\mathbb{E}(S_m^n f \cdot \overline{g}|X) - \mathbb{E}(T_0^n f \cdot \overline{g}|X)$ converges to $\mathbb{E}(S^n f \cdot \overline{g}|X) - \mathbb{E}(T_0^n f \cdot \overline{g}|X)$ strongly (consider $h_m := S_m^n f$ in our above observation). Thus, $S \in V(n, f, g, \varepsilon)$, and so $V(n, f, g, \varepsilon)$ is closed. □

Rigidity is not the only property we can relativize. Indeed, having already relativized rigidity immediately allows us to define relativiations of α-rigidity and κ-weak mixing.

Definition 2.7. Let $\alpha \in \mathbb{C}, |\alpha| = 1$, and $T \in \mathcal{G}_{T_0}$. We say that T is *α-rigid relative to* T_0, or T is an *α-rigid extension* of T_0, if there exists a subsequence (n_k) such that for all $f, g \in L^2(Z|X)$,

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k} f \cdot \overline{g}|X) - \alpha\mathbb{E}(T_0^{n_k} f \cdot \overline{g}|X)\|_{L^2(X)} = 0.$$

We denote the set of α-rigid extensions of T_0 by $\mathcal{R}_{T_0}^{\alpha}$.

Definition 2.8 (Kunde, Schnurr)**.** Let $\kappa \in [0, 1]$, and $T \in \mathcal{G}_{T_0}$. We say that T is *κ-weakly mixing* relative to T_0 or T is a *κ-weakly mixing extension* of T_0, if there exists a subsequence (n_k) such that for all $f, g \in L^2(Z|X)$,

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k} f \cdot \overline{g}|X) - (\kappa T_0^{n_k}\mathbb{E}(f|X) \cdot \mathbb{E}(g|X) + (1 - \kappa)\mathbb{E}(T_0^{n_k} f \cdot \overline{g}|X))\|_{L^2(X)} = 0.$$

We denote the set of κ-weakly mixing extensions of T_0 by $\mathcal{W}_{T_0}^{\kappa}$. Definition 2.7 and 2.8 suffer from the same problem as Definition 2.6: they are not closed under conjugation by elements of $\mathcal{G}_{\mathbb{1}_X}$. However, similar just as with rigid extensions, we can show that both are G_δ, which is our next goals. First, however, a couple of lemmas.

Lemma 2.12. *Let $\alpha \in \mathbb{C}, |\alpha| = 1, T \in \mathcal{G}_{T_0}$, and let $D \subset L^2(Z|X)$, with*

$$\left\| \|f\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \leq 1 \textit{ for all } f \in D,$$

such that D is dense in the unit ball of $L^2(Z|X)$, with respect to the $L^2(Z)$ norm topology. Further suppose that there exists a subsequence (n_k) such that for all $f_i, f_j \in D$,

$$\lim_{k\to\infty} \left\| \mathbb{E}(T^{n_k} f_i \cdot \overline{f_j}|X) - \alpha \mathbb{E}(T_0^{n_k} f_i \cdot \overline{f_j}|X) \right\|_{L^2(X)} = 0.$$

Then $T \in \mathcal{R}_{T_0}^{\alpha}$.

Analagously, one has the following.

Lemma 2.13. *Let $\kappa \in [0,1], T \in \mathcal{G}_{T_0}$, and let $D \subset L^2(Z|X)$, with*

$$\left\| \|f\|_{L^2(Z|X)} \right\|_{L^\infty(X)} \leq 1 \textit{ for all } f \in D,$$

such that D is dense in the unit ball of $L^2(Z|X)$, with respect to the $L^2(Z)$ norm topology. Further suppose that there exists a subsequence (n_k) such that for all $f_i, f_j \in D$,

$$\lim_{k\to\infty} \left\| \mathbb{E}(T^{n_k} f_i \cdot \overline{f_j}|X) - \left(\kappa T_0^{n_k} \mathbb{E}(f_i|X) \cdot \mathbb{E}(f_j|X) + (1-\kappa)\mathbb{E}(T_0^{n_k} f_i \cdot \overline{f_j}|X)\right) \right\|_{L^2(X)} = 0.$$

Then $T \in \mathcal{W}_{T_0}^{\kappa}$.

See Lemma 2.12 for details on both proofs.

Proposition 2.11. *Let $T_0 \in \mathcal{G}(X)$ be fixed and $\alpha \in \mathbb{C}, |\alpha| = 1$. Then $\mathcal{R}_{T_0}^{\alpha}$ is G_δ.*

Proof. Let $\{f_i\}$ be an $L^2(Z)$-dense subset of $L^2(Z|X)$ and define

$$U_{i,j,k,n} := \left\{ T \in \mathcal{G}_{T_0} | \left\| \mathbb{E}(T^n f_i \cdot \overline{f_j}|X) - \alpha \mathbb{E}(T_0^n f_i \cdot \overline{f_j}|X) \right\|_{L^2(X)} < \frac{1}{k} \right\}.$$

Lemma 2.12 shows that

$$\mathcal{R}_{T_0}^{\alpha} = \bigcap_{i,j,k} \bigcup_{n>k} U_{i,j,k,n},$$

The rest follows as in Proposition 2.10 □

Proposition 2.12. *Let $T_0 \in \mathcal{G}(X)$ be fixed and $\kappa \in [0,1]$. Then $\mathcal{W}_{T_0}^{\kappa}$ is G_δ.*

Proof. Let $\{f_i\}$ be an $L^2(Z)$-dense subset of $L^2(Z|X)$ and define

$$U_{i,j,k,n} := \left\{ T \in \mathcal{G}_{T_0} | \left\| \mathbb{E}(T^n f_i \cdot \overline{f_j}|X) - \left(\kappa T_0^n \mathbb{E}(f_i|X) \cdot \mathbb{E}(\overline{f_j}|X) + (1-\kappa)\mathbb{E}(T_0^n f_i \cdot \overline{f_j}|X)\right) \right\|_{L^2(X)} < \frac{1}{k} \right\}.$$

Lemma 2.13 shows that

$$\mathcal{W}_{T_0}^{\kappa} = \bigcap_{i,j,k} \bigcup_{n>k} U_{i,j,k,n},$$

The rest follows as in Proposition 2.10 □

Chapter 3

Further Questions

In closing, we would like to formulate some open questions that remain.

Question 1. Is $\mathcal{S}_{T_0}$ of first category for all base transformations T_0? If not, is the condition in Lemma 2.9 a necessary one?

There is reason to believe that the rarity of strongly mixing extensions actually depends on the transformation T_0. In fact, in the course of many discussions with Jean-Paul Thouvenot, the following conjecture has been made.

Conjecture 1 (Thouvenot). Let T_0 be a Bernoulli shift. Then $\mathcal{S}_{T_0}$ contains a dense, G_δ set.

If Conjecture 1 is true, it marks an interesting divergence in this theory from its classical counterpart.

Question 2. What is the correct definition of rigid extensions?

A property being closed under conjugation by extensions of identity means that property is a relative isomorphism invariant (essentially, a relative isomorphism is an isomorphism which preserves a certain factor algebra: in this case, the factor algebra generated by X). While it is not strictly a problem that relative rigidity as defined in Section 2.2.4 is not a relative isomorphism invariant, it is desirable for a given property to be a relative isomorphism invariant, and so if there is a definition of relative rigidity which is a relative isomorphism invariant, it is likely a good one. It should be noted however that if Conjecture 1 is true, then any definition of relative rigidity which is a relative isomorphism invariant must not be G_δ for all base transformations T_0. The reason for this is clear: any proper definition of a rigid extension must be mutually exclusive with strongly mixing extensions and if relative rigidity were both a relative isomorphism invariant and G_δ, it would be generic, and hence strongly mixing extensions would be rare for all factor transformations.

Further note that we can ask Question 2 about α-rigid and κ-weakly mixing extensions as well.

Question 3. What other properties can be relativized to produce category results?

While most think of mildly mixing transformations as those which have no rigid factor, one can also equivalently define mildly mixing by IP-convergence, see, for example,

[26]. As mildly mixing transformations are known to be rare, perhaps the natural corresponding definition of a mildly mixing extension can also be shown to be rare. While Conjecture 1 would answer this negatively for Bernoulli base transformations, perhaps it can still be shown for other possible factor transformations.

It is necessarily the case that for all $T_0, \mathcal{S}_{T_0} \subsetneq \mathcal{W}_{T_0}$, as we can take $S \in \mathcal{G}(Y)$ to be any transformation which is weakly but not strongly mixing (for example, the Chacon transformation), and then $T := T_0 \times S$ will be a weakly but not strongly mixing extension of T_0.

Proposition 3.1. *Let $T_0 \in \mathcal{G}(X)$ be any transformation, and $S \in \mathcal{G}(Y)$ be any transformation which is weakly mixing but not strongly mixing. Then $T := T_0 \times S$ is a weakly, but not strongly mixing extension of T_0.*

Proof. We wish to show that for all $f, g \in L^2(Z|X)$, there exists a subsequence n_k such that

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k} f \cdot \overline{g}|X) - T_0^{n_k}\mathbb{E}(f|X) \cdot \mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0.$$

However, it suffices to take f and g to be the characteristic functions of product sets. Standard linearity and approximation arguments will handle all other cases. Let $A, C \subset X, B, D \subset Y$ be measurable sets, and let (n_k) be a weakly mixing subsequence for S. Then

$$\begin{aligned}
&\|\mathbb{E}(T^{n_k}\chi_{A\times B} \cdot \chi_{C\times D}|X) - T_0^{n_k}\mathbb{E}(\chi_{A\times B}|X) \cdot \mathbb{E}(\chi_{C\times D}|X)\|_{L^2(X)} \\
&= \left\|\mathbb{E}(\chi_{T_0^{-n_k}A\times S^{-n_k}B} \cdot \chi_{C\times D}|X) - (\chi_{T_0^{-n_k}A}\nu(B))(\chi_C\nu(D))\right\|_{L^2(X)} \\
&= \left\|\mathbb{E}(\chi_{(T_0^{-n_k}A\cap C)\times(S^{-n_k}B\times D)}|X) - \chi_{T_0^{-n_k}A}\chi_C\nu(B)\nu(D)\right\|_{L^2(X)} \\
&= \left\|\chi_{T_0^{-n_k}A\cap C}\nu(S^{-n_k}B\cap D) - \chi_{T_0^{-n_k}A\cap C}\nu(B)\nu(D)\right\|_{L^2(X)} \\
&\leq \left\|\chi_{T_0^{-n_k}A\cap C}\right\|_{L^2(X)} \left|\nu(S^{-n_k}B\cap D) - \nu(B)\nu(D)\right|.
\end{aligned}$$

But as (n_k) is a weakly mixing subsequence for S,

$$\lim_{k\to\infty} \left|\nu(S^{-n_k}B\cap D) - \nu(B)\nu(D)\right| = 0.$$

Further, since $\left\|\chi_{T_0^{-n_k}A\cap C}\right\|_{L^2(X)} \leq 1, T$ is a weakly mixing extension of T_0.

To see that T is not a strongly mixing extension of T_0, let B be a measurable set in Y and define $A := X \times B$. Then we have $\mathbb{E}(T^n\chi_A \cdot \chi_A|X) = \int S^n\chi_B \cdot \chi_B d\nu$ and $T_0^n\mathbb{E}(\chi_A|X) = \mathbb{E}(\chi_A|X) = \int \chi_B d\nu$. Thus

$$\begin{aligned}
&\|\mathbb{E}(T^n\chi_A \cdot \chi_A|X) - T_0^n\mathbb{E}(\chi_A|X) \cdot \mathbb{E}(\chi_A|X)\|_{L^2(X)} \\
&= \left|\int S^n\chi_B \cdot \chi_B d\nu - \left(\int \chi_B d\nu\right)^2\right|.
\end{aligned}$$

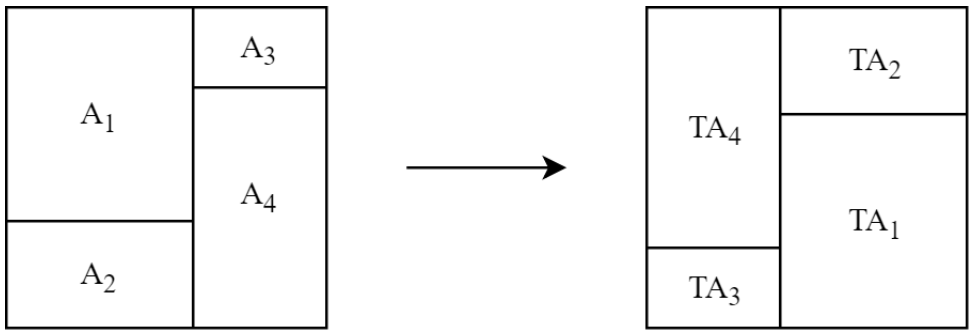
By assumption, S is not strongly mixing, so T is not a strongly mixing extension of T_0. $\square$

However, such examples are trivial. It would be interesting to find a nontrivial example of such an extension – one which is not a direct product, nor obtained by conjugating a direct product. Thus we have the following question.

Question 4. For a given T_0, can one construct a nontrivial extension of T_0 which is a weakly, but not strongly mixing extension?

One possible idea for how one might construct such an example is a two dimensional interval exchange transformation. Figure 3.1 shows a simple example of such a transformation. Of course, such transformations could only exist when the factor is itself an interval exchange transformation. Still, it may be a good first step for finding examples.

Figure 3.1



Acknowledgements

First and foremost I wish to thank my advisor, Tanja Eisner. Nothing I have accomplished in this work would have been achievable without her support and guidance. She was the one who even brought ergodic theory as a field to my attention, let alone the specific problems I tackled. I would have been lost without her, and I will be forever grateful for the kindness and support she has showed me over the past three years.

Secondly, I want to thank everyone else who has helped me in the course of my research. Terence Tao asked Tanja the question which turned into this work. Jean-Paul Thouvenot helped to give me a new direction once I had completed the main problem of my focus. Vitaly Bergelson assisted me achieving clarity about my research. And many thanks to everyone else who I had discussions with and bounced ideas off of, including Philipp Kunde, Claudia Raithel, Thilo Simon, and Ben Stanley.

Thirdly, I wish to show my appreciation for some of the teachers in my life. My high school calculus teacher, Jim Stevens, sparked my interest in math. My undergraduate advisor, Todd Eisworth, kindled that spark as I learned what math was truly about. My unoffical Master's advisor, Jeff Connor, helped me navigate my first forays into research. Lastly, many thanks to my ninth grade English teacher, Mary Duiker, for being the first teacher to truly push me and make me wish to do the best work I possibly could.

I also give my eternal and undying gratitude to my family. I know it was difficult for them – especially my mother – when I chose to move to Germany to study, but they have never shown anything but love and support for me.

Lastly, I cannot thank my friends enough. Over the years they have all been an invaluable and indispensable part of my life. From listening to my problems, to trusting me with their own, and even to simply being there to help me mentally decompress, I do not know what I would do without them.

Mike

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialen oder erbrachten Dienstleistungen als solche gekennzeichnet.

. . . . . .
(Ort, Datum)

. . . . . .
(Unterschrift)

Bibliography

[1] O.N. Ageev. The generic automorphism of a Lebesgue space conjugate to a G-extension for any finite abelian group G. *Doklady Akademii Nauk*, 374:439–442, 2000.

[2] O.N. Ageev. On the multiplicity function of generic group extensions with continuous spectrum. *Ergodic Theory and Dynamical Systems*, 21:321–338, 2001.

[3] O.N. Ageev. On the genericity of some nonasymptotic dynamic properties. *Uspekhi Matematicheskikh Nauk*, 58:177–178, 2003.

[4] O.N. Ageev. The homogeneous spectrum problem in ergodic theory. *Inventiones mathematicae*, 160:417–446, 2005.

[5] Steve Alpern and V.S. Prasad. Properties generic for Lebesgue space automorphisms are generic for measure-preserving manifold homeomorphisms. *Ergodic Theory and Dynamical Systems*, 22:1587–1620, 2002.

[6] Idris Assani, David Duncan, and Ryo Moore. Pointwise characteristic factors for Wiener-Wintner double recurrence theorem. *Ergodic Theory and Dynamical Systems*, 36:1037–1066, 2016.

[7] Idris Assani and Kimberly Presser. Pointwise characteristic factors for the multiterm return times theorem. *Ergodic Theory and Dynamical Systems*, 32:341–360, 2012.

[8] A Avila and G Forni. Weak Mixing for Interval Exchange Transformations and Translation Flows. *Annals of Mathematics*, 165(2):637–664, 2007.

[9] Vitaly Bergelson, Terence Tao, and Tamar Ziegler. Multiple recurrence and convergence results associated to F_P^ω-actions. *Journal d'Analyse Mathématique*, 127:329–378, 2015.

[10] R.V. Chacon. Weakly mixing transformations which are not strongly mixing. *Proceedings of the American Mathematical Society*, 22:559–562, 1969.

[11] Jon Chaika and Jon Fickenscher. Topological mixing for some residual sets of interval exchange transformations. *Communications in Mathematical Physics*, 333:483–503, 2015.

[12] Qing Chu. Convergence of weighted polynomial multiple ergodic averages. *Proceedings of the American Mathematical Society*, 137:1363–1369, 2009.

[13] Qing Chu, Nikos Frantzinakis, and Bernard Host. Ergodic averages of commuting transformations with distinct degree polynomial iterates. *Proceedings of the London Mathematical Society*, 102:801–842, 2011.

[14] Francois Dahmani, Koji Fujiwara, and Vincent Guirardel. Free groups of interval exchange transformations are rare. *Groups, Geometry, and Dyanmics*, 7:883–910, 2013.

[15] Thierry de la Rue and Jose de Sam Lazaro. Une transformation générique peut être insérée dans un flot. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 39:121–134, 2003.

[16] Ivan Dynnikov and Alexandra Skripchenko. Minimality of interval exchange transformations with restrictions. *Journal of Modern Dynamics*, 11:219–248, 2017.

[17] Manfred Einsiedler and Thomas Ward. *Ergodic Theory with a view towards Number Theory*. Springer, 2011.

[18] Tanja Eisner. Rigidity of contractions on Hilbert Spaces. In *Stability of Operators and Operator Semigroups*. Springer Verlag, 2010.

[19] Tanja Eisner and Ben Krause. (Uniform) convergence of twisted ergodic averages. *Ergodic Theory and Dynamical Systems*, 36:2172–2202, 2016.

[20] Tanja Eisner and Pavel Zorin-Kranich. Uniformity in the Wiener-Wintner theorem for nilsequences. *Discrete and Continuous Dynamical Systems*, 33:3497–3516, 2013.

[21] Sébastien Ferenczi. A generalization of the self-dual induction to every interval exchange transformation. *Annales de l'Institut Fourier*, 64:1947–2002, 2014.

[22] Giovanni Forni and Carlos Matheus. Introduction to Teichmüller theory and its applications to dynamics of interval exchange transformations, flows on surfaces and billiards. *Journal of Modern Dynamics*, 8:271–436, 2014.

[23] Nikos Frantzinakis and Pavel Zorin-Kranich. Multiple recurrence for non-commuting transformations along rationally independent polynomials. *Ergodic Theory and Dynamical Systems*, 35:403–411, 2015.

[24] H Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *Journal d'Analyse Mathématique*, 31:204–256, 1977.

[25] H. Furstenberg, Y. Katznelson, and D. Ornstein. The Ergodic Theoretical Proof of Szemeredi's Theorem. *Bulletin of the American Mathematical Society*, 7(3):527–552, 1982.

[26] Harry Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, 1981.

[27] Eli Glasner. *Ergodic Theory via Joinings*. American Mathematical Society, 2003.

[28] Eli Glasner and Benjamin Weiss. Relative Weak Mixing is Generic. 2017. https://arxiv.org/abs/1707.06425.

[29] Pierre-Antoine Guihéneuf. Dynamical properties of spatial discretizations of a generic homeomorphism. *Ergodic Theory and Dynamical Systems*, 35:1474–1523, 2015.

[30] Paul Halmos. In General a Measure Preserving Transformation is Mixing. *Annals of Mathematics*, 45(4):786–792, 1944.

[31] Paul Halmos. *Lectures on Ergodic Theory*. Chelsea Publishing Company, 1956.

[32] Bernard Host and Bryna Kra. Nonconventional ergodic averages and nilmanifolds. *Annals of Mathematics*, 161:397–488, 2005.

[33] Bernard Host, Bryna Kra, and Alejandro Maass. Complexity of nilsystems and systems lacking nilfactors. *Journal d'Analyse Mathématique*, 124:261–295, 2014.

[34] Shizuo Kakutani. *Examples of ergodic measure preserving transformations which are weakly mixing but not strongly mixing*, chapter Recent Advances in Topological Dynamics, pages 143–149. Springer-Verlag, 1973.

[35] A.B. Katok. Interval exchange transformations and some special flows are not mixing. *Israel Journal of Mathematics*, 35:301–310, 1980.

[36] A.B. Katok and A.M. Stepin. Metric properties of homeomorphisms that preserve measure. *Uspekhi Matematicheskikh Nauk*, 25:193–220, 1970.

[37] JLF King. The generic transformation has roots of all orders. *Colloquium Mathematicae*, 84/85:521–547, 2000.

[38] Luis-Miguel Lopez and Philippe Narbel. Infinite interval exchange transformations from shifts. *Ergodic Theory and Dynamical Systems*, 37:1935–1965, 2017.

[39] M.G. Nadkarni. *Spectral theory of dynamical systems*, chapter 8. Baire Category Theorems of Ergodic Theory, pages 51–61. Birkhäuser, 1998.

[40] Christopher Novak. Continuous interval exchange actions. *Algebraic & Geometric Topology*, 10:1609–1625, 2010.

[41] Karl Peterson. *Ergodic Theory*. Cambridge University Press, 1989.

[42] Donald Robertson. Characteristic factors for commuting actions of amenable groups. *Journal d'Analyse Mathématique*, 129:165–196, 2016.

[43] E. Arthur Robinson. The maximal abelian subextension determines weak mixing for group extensions. *Proceedings of the American Mathematical Soceity*, 114:443–450, 1992.

[44] V. Rokhlin. A "general" measure-preserving transformation is not mixing. *Doklady Akademii Nauk SSSR*, 60:349–351, 1948.

[45] Mike Schnurr. A Note on Strongly Mixing Extensions. 2017. https://arxiv.org/abs/1712.06192.

[46] Mike Schnurr. Generic properties of extensions. *Ergodic Theory and Dynamical Systems*, 2018. Awaiting a volume. Currently published online: https://doi.org/10.1017/etds.2017.145.

[47] Slawomir Solecki. Closed subgroups generated by generic measure automorphisms. *Ergodic Theory and Dynamical Systems*, 34:1011–1017, 2014.

[48] A.M. Stepin. Spectral properties of generic dynamical systems. *Mathematics of the USSR-Izvestiya*, 29:159–192, 1987.

[49] Terence Tao. *Poincare's Legacies, Part II: pages from year two of a mathematical blog*, chapter 2. Ergodic Theory, pages 161–355. American Mathematical Society, 2009.

[50] Terence Tao, 2012. Personal communication with Tanja Eisner.

[51] Terence Tao and Tamar Ziegler. Concatenation theorems for anti-Gowers-uniform functions and Host-Kra characteristic factors. *Discrete Analysis*, 2016.

[52] W.A. Veech. The Metric Theory of Interval Exchange Transformations I. Generic Spectral Properties. *American Journal of Mathematics*, 106(6):1331–1359, 1984.

[53] Yue Wu. Whirly 3-interval exchange transformations. *Ergodic Theory and Dyanimcal Systems*, 37:995–1008, 2017.

[54] Tamar Ziegler. Universal characteristic factors and Furstenberg averages. *Journal of the American Mathematical Society*, 20:53–97, 2007.