Ergod. Th. & Dynam. Sys. (2019), **39**, 3144–3168 © *Cambridge University Press, 2018*
doi:10.1017/etds.2017.145

Generic properties of extensions

MIKE SCHNURR

Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany
(e-mail: schnurr@mis.mpg.de)

(*Received* 22 *April* 2017 *and accepted in revised form* 23 *October* 2017)

Abstract. Motivated by the classical results by Halmos and Rokhlin on the genericity of weakly but not strongly mixing transformations and the Furstenberg tower construction, we show that weakly but not strongly mixing extensions on a fixed product space with both measures non-atomic are generic. In particular, a generic extension does not have an intermediate nilfactor.

1. *Introduction*

The classical results by Halmos [**22**] and Rokhlin [**31**] state that a 'typical' measure-preserving transformation on a probability space (X, μ) is weakly but not strongly mixing. More precisely, the set of weakly mixing transformations is a dense, G_δ (hence residual) set for the weak topology and the set of strongly mixing transformations is of first category. The combination of these two results proved the existence of a weakly but not strongly mixing transformation, without providing a concrete example (for such an example, see [**9**]). Since then, much research has been done in finding 'typical' properties of measure-preserving dynamical systems. See, e.g., [**1**–**5**, **12**, **21**, **26**–**28**, **32**].

More than 30 years later, in completely unrelated efforts, Furstenberg [**17**] 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 [**18**] building up a tower of so-called compact and weakly mixing extensions had a much greater impact on the 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., [**6**–**8**, **10**, **11**, **14**–**16**, **29**, **34**]. For example, the correct characteristic factor for norm convergence of multiple ergodic averages was identified by Host and Kra [**24**] and has the structure of an inverse limit of nilsystems, see also Ziegler [**35**].

The purpose of this paper is to prove analogs of the Halmos and Rokhlin category theorems for extensions (see Theorems 5 and 6), extending a result of Robertson on compact group extensions [**30**]. Inspired by Rokhlin's skew product representation

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

theorem (see, e.g., [**19**, p. 69]), we consider extensions defined on product spaces with the natural projection as the factor map. We show that for a fixed product space, where both measures are non-atomic, a 'typical' extension is weakly but not strongly mixing. Here, by an extension we mean an invertible extension of some (non-fixed), invertible, measure-preserving transformation on the factor. Note that the set of extensions is a closed, nowhere dense set in all invertible transformations on the product space (see Proposition 2), so the classical Halmos and Rokhlin results cannot be applied. The proof for weakly mixing extensions is a non-trivial adaptation of the original construction by Halmos. In particular, a 'typical' extension does not have an intermediate nilfactor. For examples of systems lacking non-trivial nilfactors, see [**25**].

The paper is organized as follows. After discussing some preliminaries in §2, we consider the case of discrete extensions in §3, in particular showing that there are no weakly mixing extensions on these spaces (Proposition 3), but that permutations are dense (Theorem 1). We then prove the weak approximation theorem for extensions on the unit square (Theorem 3) in §4 using the density result for discrete extensions mentioned above. In §5 we generalize a few results, including Halmos' uniform approximation theorem, which are necessary to prove our conjugacy lemma for extensions (Lemma 5) in §6. Section 7 is devoted to the proof that weakly mixing extensions on the unit square are residual and §8 addresses the case of general vertical measure. In §9 we define strongly mixing extensions, and show that such extensions are of first category (Theorem 6). Finally in §10 we formulate some open questions. After this paper was made public, the first part of Question 2 was answered by Glasner and Weiss, see [**20**].

2. *Preliminaries*

As explained in the introduction, in this paper 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, η) be a probability space, $(Z, \mu) = (X \times Y, m \times \eta)$, and T, T' be measure-preserving transformations on (Z, μ), (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. Throughout this paper, 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 [**22**, 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.

The weak topology on $\mathcal{G}(Z)$ is the topology defined by the subbasic neighborhoods

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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). See [**23**] for discussions of this topology. It is helpful to note that the weak topology happens to coincide with the weak (and strong) operator topology for the corresponding Koopman operators. Further, in this paper we will be interested in the weak topology on $\mathcal{G}_X$, by which we mean the subspace topology inherited by the weak topology.

We will need the following two metrics on $\mathcal{G}(Z)$ defined by

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

where the sup in the first definition is taken over all measurable sets E. These metrics were used by Halmos in his proof of the category theorem, see [**23**]. We note that both metrics induce the same topology on $\mathcal{G}(Z)$, but that topology is not the weak topology. Moreover, they satisfy $d(S, T) \leq d'(S, T)$ for all $S, T \in \mathcal{G}(Z)$. The last important note is that d' is invariant under multiplication by transformations. That is to say, for all $R, S, T \in \mathcal{G}(Z)$,

$$d'(RS, RT) = d'(S, T) = d'(SR, TR).$$

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

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

For more on $L^2(Z|X)$, see [**33**]. One important result of $L^2(Z|X)$ that we do wish to emphasize for later is the Cauchy–Schwarz inequality.

PROPOSITION 1. *Let $f, g \in L^2(Z|X)$. Then*

$$|\langle f, g\rangle_{L^2(Z|X)}| \leq \|f\|_{L^2(Z|X)}\|g\|_{L^2(Z|X)} \text{ almost everywhere.}$$

Next we give a definition for weakly mixing extensions, cf. [**33**].

Definition 1. An extension T of T' is said to be *weakly mixing* if for all $f, g \in L^2(Z|X)$,

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

For other possible (equivalent) definitions of weakly mixing extensions, see [**19**, p. 192].

We denote by $\mathcal{W}_X \subset \mathcal{G}_X$ the set of weakly mixing extensions on Z. We finally prove that the Baire category theorem is applicable to $\mathcal{G}_X$, and further that $\mathcal{G}_X$ is topologically a small subset of $\mathcal{G}(Z)$.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

PROPOSITION 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 neighborhood 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 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

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

Now because $\mu(C_m\triangle TE) \to 0$, $\mu(\bigcap_{m\geq n}(C_m\backslash TE)) = 0$ 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,

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

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 $\mu(A) < \varepsilon$, 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 $A_1 := A \cap B$, $A_2 := A \cap B^c$.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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, $\mathcal{G}_X$ 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 $\mathcal{G}_X$.

3. *Discrete extensions*

As stated in §2, throughout this paper 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 3. *Let (Z, μ), (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

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

In particular, we will construct f such that $\mathbb{E}(T^n f \overline{f}|X)(x)$ can take only a finite number of possible values, none of which are 0. Thus $|\mathbb{E}(T^n f \overline{f}|X)|^2(x)$ is always positive, $(1/N) \sum_{n=0}^{N-1} |\mathbb{E}(T^n f \overline{f}|X)|^2(x)$ is bounded away from 0, and $(1/N) \sum_{n=0}^{N-1} (\int_X |\mathbb{E}(T^n f \overline{f}|X)|^2 \, dm)^{1/2}$ cannot converge to the zero function on X.

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

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

for all $x \in X$. It is easy to see that f has relative mean zero when $L \geq 2$, which is why we made this assumption at the beginning of the section.

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

$$\mathbb{E}(T^n f \overline{f}|X)(x) = \sum_{j=1}^{L} w_j f(x, j) f((T')^n x, \sigma_{n,x}(j)) = \sum_{j=1}^{L} w_j f(x, j) f(x, \sigma_{n,x}(j)),$$

with the last equality 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 $\mathbb{E}(T^n f \overline{f}|X)(x)$

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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 $\mathbb{E}(T^n f\overline{f}|X)(x)$.

To see $\mathbb{E}(T^n f\overline{f}|X)(x) \neq 0$, consider two cases. In the first case, we have $\sigma_{n,x}(1) = 1$. In this case it is easy to see 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, non-zero). 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 $(-\sum_{j=2}^{L} w_j)(1 + w_i/w_1) \leq -\sum_{j=2}^{L} w_j$. Thus,

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

So $\mathbb{E}(T^n f\overline{f}|X)(x)$ is always non-zero, and $|\mathbb{E}(T^n f\overline{f}|X)|^2(x)$ is always positive, as desired. □

We make two notes here. First, the proof of Proposition 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. The key observation is that for almost all z, if z, Tz are on levels k_1, k_2 respectively, then $w(k_1) = w(k_2)$. As for any fixed $\alpha \in (0, 1)$, there can be only finitely many levels k with $w(k) = \alpha$, T decomposes into invariant subsystems, to each of which we can apply Proposition 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, $w_i = 1/L$ 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. 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 $\mathcal{G}_X$. To this end, we make a couple of notes. First we introduce the following notation: we write $A \subset i$ if there exists $A' \subset X$ such that $A = A' \times \{i\}$.

Second, we will require the use of the following lemma by Halmos (for proof, see [**23**, p. 67]).

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

LEMMA 1. *Let $\{E_i : i = 1, \dots n\}$ 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 $\{F_i : i = 1, \dots n\}$, 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.*

We now move to the main result of this section.

THEOREM 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

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

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).

Let $k \in \{1, \dots, L\}$, and let $P_k := \{D_i \cap TD_j | D_i \subset k, j = 1, \dots, L(2^n)\}$. 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 $k = 1, \dots, L$, and let $P = \{\hat{A}_{\lambda,k} | \lambda \in \Lambda, k = 1, \dots, L\}$ 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 (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 $m(\hat{A}_\lambda \triangle A_\lambda) < L\varepsilon/2|\Lambda|$. Applying Lemma 1 again, we get a partition of X into dyadic sets B_λ so that $m((T')^{-1}\hat{A}_\lambda \triangle B_\lambda) < L\varepsilon/2|\Lambda|$. Note the full strength of Lemma 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

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

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 $j \in 1, \dots, L(2^n)$, 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

$$\mu\Big(D_j \triangle \bigcup B_{\lambda,k}\Big) = \mu\Big(\bigcup T^{-1}\hat{A}_{\lambda,i_{\lambda,j}} \triangle \bigcup B_{\lambda,k}\Big) \leq \sum \mu(T^{-1}\hat{A}_{\lambda,i_{\lambda,j}} \triangle B_{\lambda,k}). \tag{2}$$

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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

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

On the other hand,

$$\mu\Big(\bigcup A_{\lambda,i_{\lambda,j}} \triangle TD_j\Big) = \mu\Big(\bigcup A_{\lambda,i_{\lambda,j}} \triangle \bigcup \hat{A}_{\lambda,i_{\lambda,j}}\Big) \leq \sum (A_{\lambda,i_{\lambda,j}} \triangle \hat{A}_{\lambda,i_{\lambda,j}}). \tag{3}$$

Again by (1), $u(A_{\lambda,i_{\lambda,j}} \triangle \hat{A}_{\lambda,i_{\lambda,j}}) < \varepsilon/2|\Lambda|$ so (3) $< \sum (\varepsilon/2|\Lambda|) = \varepsilon/2$. Finally,

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

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

4. *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 §3. First, however, we need some more notation. For $L \in \mathbb{N}$, define $Z_L := \bigcup_{j=0}^{L-1} (X \times \{j/L\}) \subset Z$, μ_L, a measure on Z_L, to be the product of the Lebesgue measure with a normalized counting measure on L points. Further, $\pi_L : Z \to Z_L$ is the natural projection onto Z_L. That is, if $z = (x,\ j/L + \gamma)$ for $\gamma \in [0,\ 1/L)$, then $\pi_L(z) = (x,\ j/L)$.

Definition 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 $Z_L \times [0,\ 1/L)$, then $T = T_L \times I$. If we wish to emphasize the number of levels, L, we will say T is *L-(simply) discrete equivalent*.

Definition 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 T_L is also an extension of T'.

Our goal for this section is to provide a version of Halmos' weak approximation theorem (see [**23**, p. 65]) when restricted to $\mathcal{G}_X$. Mostly this will mean proving a result equivalent to Theorem 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 redefine them. Column-preserving permutations are defined just as they are in Definition 2.

Before moving on, we make a few remarks.

Remark 1. Lemma 1 holds on (Z, m), because (Z, m), like (X, m), is a non-atomic standard probability space, and Lemma 1 holds for all such spaces (replacing 'dyadic sets' in the statement of Lemma 1 with a class $\mathcal{B}$ which is isomorphic to the class of dyadic sets). Alternatively, one can simply prove Lemma 1 again in the context of the square. No part of the proof relies on the fact that we were on the unit interval, so nothing changes in the proof.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

Remark 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 $(S' \times I)T$.

Remark 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 $S'Q'$.

The key to our goal is the following strengthening of Lemma 1.

LEMMA 2. *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 $\sum_{i=1}^N r_{ij} = 1/2^K$ for all j and $|m(E_{ij}) - r_{ij}| < \varepsilon/2^K$ for all i, j. Then there exists $\{F_1, \dots, F_N\}$ a partition of Z such that F_i is a 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, 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 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

$$\sum_j m(A_{ij}) = \sum_j |m(\tilde{F}_{ij}) - r_{ij}| \le \sum_j |m(E_{ij}) - r_{ij}| + \sum_j |m(E_{ij}) - m(\tilde{F}_{ij})|.$$

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

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

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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

$$m(\tilde{F}_i \triangle F_i) = m\left(\bigcup_j \tilde{F}_{ij} \triangle \bigcup_j F_{ij}\right) \le m\left(\bigcup_j (\tilde{F}_{ij} \triangle F_{ij})\right) = \sum_j m(\tilde{F}_{ij} \triangle F_{ij}).$$

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, we can now prove the equivalent version of Theorem 1 for the unit square, which will be the core result for proving our version of the weak approximation theorem.

THEOREM 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 $T' = I_X$.

Let $D_{ij} := D_i \cap TD_j$. Note that $\{D_{ij}\}$ partitions Z. By Lemma 1, there exists a partition of Z into dyadic sets, $\{\tilde{E}_{ij}\}$, such that $m(D_{ij} \triangle \tilde{E}_{ij}) < \varepsilon/6M$, where $M := 2^{2N}$. Further by Lemma 1, we can find a dyadic partition of Z into sets $\{\tilde{F}_{ij}\}$ where $m(T^{-1}D_{ij} \triangle \tilde{F}_{ij}) < \varepsilon/6M$. 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 to both $\tilde{E}_{ij}$ and $\tilde{F}_{ij}$ to get dyadic partitions $\{E_{ij}\}$ and $\{F_{ij}\}$ such that

$$m(D_{ij} \triangle E_{ij}) < \frac{\varepsilon}{2M} \quad \text{and} \quad m(T^{-1}D_{ij} \triangle F_{ij}) < \frac{\varepsilon}{2M}. \tag{4}$$

Recall that if C_k is a dyadic interval of rank K, then in the notation of Lemma 2, $E_{ijk} := E_{ij} \cap \pi^{-1}C_k$ 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, 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

$$m\left(D_j \triangle \bigcup_i F_{ij}\right) = m\left(\bigcup_i T^{-1}D_{ij} \triangle \bigcup_i F_{ij}\right) \le \sum_i m(T^{-1}D_{ij} \triangle F_{ij}). \tag{5}$$

But per (4), $m(T^{-1}D_{ij} \triangle F_{ij}) < \varepsilon/2M$, so (5) $< \sum_i \varepsilon/2M \le \varepsilon/2$. Therefore

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

On the other hand,

$$m\left(TD_j \triangle \bigcup_i E_{ij}\right) = m\left(\bigcup_i D_{ij} \triangle \bigcup_i E_{ij}\right) \leq \sum_i m(D_{ij} \triangle E_{ij}). \tag{6}$$

Again, per (4), $m(D_{ij} \triangle E_{ij}) < \varepsilon/2M$, so (6) $< \sum_i (\varepsilon/2M) = \varepsilon/2$. Therefore,

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

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 $m(T'\tilde{Q}D_i \triangle TD_i) = m(\tilde{Q}D_i \triangle \tilde{T}D_i) < \varepsilon/2$. By Remark 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 $\tilde{N}_{\varepsilon/2}(\pi_L(T'\tilde{Q})) := \{S_L : \mu_L(\pi_L(T'\tilde{Q})G_i \triangle S_L G_i) < \varepsilon/2 \forall i\}$, then by Theorem 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 $m(QD_i \triangle T'\tilde{Q}D_i) = \mu_L(\hat{Q}G_i \triangle \pi_L(T'\tilde{Q})G_i) < \varepsilon/2$. 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 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;*
- *Q is periodic with period 2^k everywhere;*
- *$Q \in N_\varepsilon(T)$.*

Proof. Because Theorem 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 $K/2^{k-1} < \varepsilon$.

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

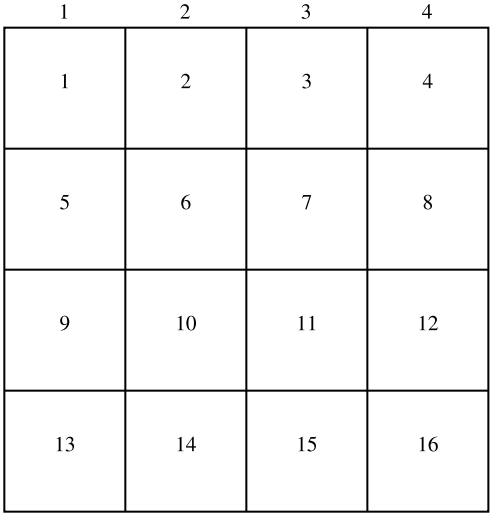


FIGURE 1. Labeling of 2-squares for the definition of P.

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 $P^l E_1$ (by 'second' we mean the one to the right of the first). Note that $P^l E_1 \neq E_1$ in general.

We now repeat the entire process, replacing 'first' for 'second,' eventually 'third' and so on, as well as replacing E_1 with $P^l E_1$. 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 third row within its M-square, say, all our choices will be in the third row of the respective M-squares. 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 $m(PE_j \triangle QE_j) \leq 2/2^{2k} = 1/2^{2k-1}$. There are 2^k such E_j per k-column, and there are K such k-columns. Thus, $m(PD_i \triangle QD_i) \leq \bigcup_j m(PE_j \triangle QE_j) \leq K2^k/2^{2k-1} = K/2^{k-1} < \varepsilon$. □

The construction in the proof of Theorem 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 two. See Figure 1 for reference on how we label the 2-squares.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

(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)

FIGURE 2. Labeling of the 3-squares which define Q.

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.

Suppose we were to construct Q to be a rank-three permutation. Rather than write the entire cycle decomposition of Q (as it would involve writing all 64 3-squares), we label Figure 2 to define Q. Here we have labeled the 3-squares such that for a square labeled (n, k), we have that $Q(n, k) = (n, k + 1)$, $(k \bmod 8)$ (for consistency, here we have $8 \bmod 8 := 8$ 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 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, 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 3 as it is this formulation we need for further results.

5. *Uniform approximation*

Our goal in this section is to prove results that are generalizations of those needed for Halmos' classical conjugacy lemma (the key lemma for proving that weakly mixing transformations on X are dense in $\mathcal{G}(X)$), and whose proofs quickly follow from the classical results and their proofs.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

LEMMA 3. *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' 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 $m(E) = m(E') = 1/n$, we have $m(\bigcup_{i=0}^{n-1} T^iE) = \sum_{i=0}^{n-1} m(T^iE) = 1$, or $\bigcup_{i=0}^{n-1} T^iE = X$. □

Next we move to a version of Rokhlin's lemma (see, e.g., [**23**, p. 71]).

LEMMA 4. *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 $m(\bigcup_{i=0}^{n-1} T^iE) > 1 - \varepsilon$.*

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 $m(\bigcup_{i=0}^{n-1}(T')^iE') > 1 - \varepsilon$. Let $E := \pi^{-1}E'$. Because T extends T', $\{E, TE, \ldots, T^{n-1}E\}$ are pairwise disjoint. Further $m(\bigcup_{i=0}^{n-1} T^iE) = \sum_{i=0}^{n-1} m(T^iE) = \sum_{i=0}^{n-1} m((T')^iE') = m(\bigcup_{i=0}^{n-1}(T')^iE') > 1 - \varepsilon$. □

We conclude this section with a version of Halmos' uniform approximation theorem (see [**23**, p. 75]).

THEOREM 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 $d'(R, T) \leq 1/n + \varepsilon$.*

Proof. By Lemma 4, there exists E a cylinder set, such that $\{E, TE, \ldots, T^{n-1}E\}$ are pairwise disjoint, and $m(\bigcup_{i=0}^{n-1} T^iE) > 1 - \varepsilon$. If $z \in \bigcup_{i=0}^{n-2} T^iE$, 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 $d'(R, T) \leq m(T^{n-1}E) + \varepsilon \leq 1/n + \varepsilon$.

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 you would in the classical case, and then letting $R = R' \times I$ on this set of measure ε. □

6. *Conjugacy lemma*

We now prove a generalization of Halmos' conjugacy lemma (see [**23**, p. 77]), using the same techniques as Halmos' original proof.

LEMMA 5. (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 $S^{-1}T_0S \in N_\varepsilon(T)$.*

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

Proof. Let $k_0 \in \mathbb{N}$ be greater than the ranks of all D_i and $1/2^{k_0-2} < \varepsilon$. 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 3 (Q' is cyclic, Q is 2^k periodic). Applying the uniform approximation theorem for extensions, Theorem 4, with 2^k in place of n and $1/2^k$ in place of ε, there exists $R \in \mathcal{G}_X$ such that R, R' have period 2^k almost everywhere, and $d'(R, T_0) \leq 1/2^k + 1/2^k < \varepsilon/2$.

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 $QE_i = E_{i+1}$ ($i \bmod q$). Note that $m(E_i) = 1/q$. By Lemma 3, there exists F_0, a cylinder set, such that $m(F_0) = 1/q$ 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:

$$\begin{array}{ccccccccccc}
E_0 & \xrightarrow{Q} & E_1 & \xrightarrow{Q} & E_2 & \xrightarrow{Q} & \cdots & \xrightarrow{Q} & E_{q-2} & \xrightarrow{Q} & E_{q-1} \\
\downarrow{\scriptstyle S} & & \downarrow{\scriptstyle S} & & \downarrow{\scriptstyle S} & & & & \downarrow{\scriptstyle S} & & \downarrow{\scriptstyle S} \\
F_0 & \xrightarrow[R]{} & F_1 & \xrightarrow[R]{} & F_2 & \xrightarrow[R]{} & \cdots & \xrightarrow[R]{} & F_{q-2} & \xrightarrow[R]{} & F_{q-1}
\end{array}$$

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

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

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

This lemma is so important because as we will see in our main result in Theorem 5, the conjugacy class of $\mathcal{W}_X$ is $\mathcal{W}_X$ itself. Thus, in proving Lemma 5, we have indeed proven half of Theorem 5.

7. *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 1, but it will be important enough to make a special note of it.

PROPOSITION 4. *Let $f, g \in L^2(Z|X)$. Then*

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

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

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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

$$|\mathbb{E}(fg|X)| \le M\mathbb{E}(|g|^2|X)^{1/2}. \tag{7}$$

Now note that by Fubini,

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

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 (7), we arrive at the desired inequality. □

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

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k} f\overline{g}|X) - (T')^{n_k}\mathbb{E}(f|X)\mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0. \tag{8}$$

To this end, we first show that if (8) holds for an $L^2(Z)$-dense subset of $L^2(Z|X)$, it holds for all of $L^2(Z|X)$.

LEMMA 6. *Let $T \in \mathcal{G}_X$, and let $D \subset L^2(Z|X)$, with $\|\|\overline{f}\|_{L^2(Z|X)}\|_{L^\infty(X)} \le 1$ for all $f \in D$, such that D is dense in the unit ball of $L^2(Z|X)$, but with respect to the $L^2(Z)$ norm topology. Further suppose that there exists a subsequence n_k such that for all f_i, $f_j \in D$,*

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

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

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

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

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

uniformly with respect to n. Indeed, we have

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

Now by Proposition 4

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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

$$\|\|T^n h_j\|_{L^2(Z|X)}\|_{L^\infty(X)}\|\overline{g_j} - \overline{g}\|_{L^2(Z)} \le \|\|h_j\|_{L^2(Z|X)}\|_{L^\infty(X)}\|\overline{g_j} - \overline{g}\|_{L^2(Z)},$$
$$\|\|\overline{g}\|_{L^2(Z|X)}\|_{L^\infty(X)}\|T^n(h - h_j)\|_{L^2(Z)} \le \|\|\overline{g}\|_{L^2(Z|X)}\|_{L^\infty(X)}\|h - h_j\|_{L^2(Z)}.$$

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 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 $L^2(X)$.

Now,

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

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

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

as desired. □

Next, recall that a function $f \in L^2(Z|X)$ is called a *generalized eigenfunction* for a given $T \in \mathcal{G}_X$ (see [**19**, 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 $c_j^n \in L^\infty(X)$, $1 \le j \le l$ such that

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

LEMMA 7. *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 $f, g \in L^2(Z|X)$*

$$\lim_{k\to\infty} \|\mathbb{E}(T^{n_k} f\overline{g}|X) - (T')^{n_k}\mathbb{E}(f|X)\mathbb{E}(\overline{g}|X)\|_{L^2(X)} = 0. \tag{9}$$

Proof. Let $T \in \mathcal{W}_X$. Lemma 6 tells us that we need only show that there exists a subsequence such that (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 $l \mapsto (f_i, f_j)$, $1 \le l < \infty$ order the set $\{(f_i, f_j)\}$ arbitrarily. Now, by the definition of a weakly mixing extension and the Koopman–von Neumann lemma (see, e.g., [**13**, p. 54]), for each l there exists a subsequence $n_m'^l$ of upper asymptotic density 1 such that

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

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 as a diagonal sequence obtained from n_m^l, i.e. $n_k := n_k^k$. It is easy to see that for all f_i, f_j, (9) holds.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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 [**19**, 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 $\mathbb{E}(g_i \overline{g}_j|X) = 0$ almost everywhere when $i \neq j$ and $\mathbb{E}(|g_j|^2|X) = 1$ almost everywhere. 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

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

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) and for all other x, $h_{i,x}(y) := h_{i,x}'(y)$. 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 $\|c_{j(n)}^n\|_{L^2(Z)} \geq \|c_i^n\|_{L^2(Z)}$, $1 \leq i \leq l$. Note that for each n, $\|c_{j(n)}^n\|_{L^2(Z)} \geq 1/l$ 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

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

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

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

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, $\|\mathbb{E}(T^{n_k} f \overline{g_i}|X)\|$ does not converge to 0 as $k \to \infty$.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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

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

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

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

The second term is 0 as $\mathbb{E}(f_0|X) \equiv 0$. Further, because $h \in L^\infty(X)$, $\mathbb{E}(T^n h\overline{g}|X) = (T')^n h\mathbb{E}(\overline{g}|X) = (T')^n\mathbb{E}(h|X)\mathbb{E}(\overline{g}|X)$, 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 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 5 to conclude $\mathcal{W}_X$ is dense.

We need to prove that there exists a subsequence of

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

which converges to 0. Indeed, the above equals

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

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

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

Due to Lemmas 6, 7, 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

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

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

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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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 $\mathbb{E}(h_m g|X) \to \mathbb{E}(hg|X)$ in $L^2(X)$. Indeed, by Proposition 4,

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

Second, note that $S_m \to S$ strongly implies $(S_m)' \to S'$. 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, $\|(S_m)'h - S'h\|_{L^2(X)} = \|S_m\hat{h} - S\hat{h}\|_{L^2(Z)}$.

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

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

Remark 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.

8. *General case*

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, T') be as before with T' invertible, and let (Y, η) be a probability space where $Y = A \dot{\cup} B$, 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 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 §3 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

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

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 5 is that there are no weakly mixing extensions on Z.

COROLLARY 1. *Let (X, m, T'), (Y, η), (Z, μ, T), C, and D be as defined above, with A and B, continuous and discrete parts of Y, respectively, non-empty. 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 5, f is T-invariant, and so

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

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

Remark 5. As with the G_δ part of Theorem 5, Proposition 5 and Corollary 1 hold when (X, m) is replaced with any standard probability space.

9. *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 4. Let (X, ν), (Z, μ) 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 $f, g \in L^2(Z|X)$,

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

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

Definition 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 $\mathcal{G}_X$.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

THEOREM 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 $P_k := \{T \in \mathcal{G}_X | T^k = I_Z\}$. Note in particular that if $T \in P_k$ then $(T')^k = I_X$. For $n \in \mathbb{N}$, let $\hat{P}_n := \bigcup_{k>n} P_k$. Note that the weak approximation theorem for extensions (Theorem 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,

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

Using the same arguments as used in the proof of Theorem 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 $\mathcal{S}_X \subset M$. 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 $\mathcal{G}_X \setminus \bigcap_{k>n} M_k$ is dense. Lastly, as $\mathcal{G}_X \setminus \bigcap_{k>n} M_k = \bigcup_{k>n} (\mathcal{G}_X \setminus M_k)$, it will suffice to show that $P_k \subset (\mathcal{G}_X \setminus M_k)$ for all k as then $\hat{P}_n = \bigcup_{k>n} P_k \subset \bigcup_{k>n} (\mathcal{G}_X \setminus M_k)$ and $\hat{P}_n$ is dense.

Now, if $T \in P_k$, then $T^k = I_Z$, $(T')^k = I_X$, so

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

Thus, $T \notin M_k$. □

COROLLARY 2. $\mathcal{S}_X$ *is a proper subset of* $\mathcal{W}_X$.

10. *Further questions*

To conclude this paper, we formulate some open questions.

Question 1. Let (X, ν) be any probability space, (Y, η) be a non-atomic probability space and $(Z, \mu) = (X \times Y, \nu \times \eta)$. Is $\mathcal{W}_X$ a dense, G_δ subset of $\mathcal{G}_X$?

It would be sufficient for Question 1 to consider the case where (X, ν) is purely atomic. We cannot use the conjugacy lemma in this case because there are no antiperiodic transformations on a discrete set.

Question 2. Let (X, m) be a (potentially non-atomic) probability space, (Y, η) be a non-atomic probability space and $(Z, \mu) = (X \times Y, \nu \times \eta)$. Let $R \in \mathcal{G}(X)$ be fixed. Define

$$\mathcal{G}_R := \{T \in \mathcal{G}_X : T \text{ extends } R\}$$

and

$$\mathcal{W}_R := \{T \in \mathcal{G}_R : T \text{ is a weakly mixing extension of } R\}.$$

Is $\mathcal{W}_R$ a dense, G_δ subset of $\mathcal{G}_R$?

Similarly let

$$\mathcal{S}_R := \{T \in \mathcal{G}_R : T \text{ is a strongly mixing extension of } R\}.$$

Is $\mathcal{S}_R$ a first category subset of $\mathcal{G}_R$?

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

The main difficulty as of now in answering Question 2 for weakly mixing extensions is proving the density. The freedom of having a non-fixed factor allowed a generalization of Halmos' conjugacy lemma. One cannot conjugate in $\mathcal{G}_R$ unless R is the identity on X, as $\mathcal{G}_R$ is not closed under inverses for any other invertible R. That is not even to mention lack of dyadic permutations or other tools which we had at our disposal throughout this paper. Even strongly mixing extensions at first present some trouble, as the proof relied on periodic T, and if R is not periodic, then neither is T.

We note that in the case of compact group extensions of a fixed weakly mixing R, Robinson gave a positive answer to the question of genericity of weakly mixing extensions [**30**].

Further, we see from the following proposition that Question 2, cannot be derived from Theorems 5 and 6. Recall that (X, m) is the unit interval with the Lebesgue measure.

PROPOSITION 6. *For all $T' \in \mathcal{G}(X)$, $\mathcal{G}_{T'}$ is a closed, nowhere dense subset of $\mathcal{G}_X$.*

Proof. Fix $T' \in \mathcal{G}(X)$. We first show that $\mathcal{G}_{T'}$ is closed. Let $S \in \mathcal{G}_X \backslash \mathcal{G}_{T'}$. As $S' \neq T'$, there exists $E \subset X$ such that $m(S'E \triangle T'E) > 0$. Let $\varepsilon := m(S'E \triangle T'E)/2$.

Consider

$$N_\varepsilon(S) := \{R \in \mathcal{G}_X : \mu(R\pi^{-1}E \triangle S\pi^{-1}E) < \varepsilon\}.$$

For any $T \in \mathcal{G}_{T'}$, $\mu(T\pi^{-1}E \triangle S\pi^{-1}E) = m(T'E \triangle S'E) = 2\varepsilon > \varepsilon$, so $T \notin N_\varepsilon(S)$, and thus $\mathcal{G}_{T'}$ is closed.

Now, to show $\mathcal{G}_{T'}$ is nowhere dense, we show $\mathcal{G}_X \backslash \mathcal{G}_{T'}$ is dense. Fix $T \in \mathcal{G}_{T'}$, $\varepsilon > 0$ and let

$$N_\varepsilon(T) := \{S \in \mathcal{G}_X : \mu(TD_i \triangle SD_i) < \varepsilon, i = 1, \ldots 2^{2N}\}$$

where D_i are all dyadic squares of some rank N. Fix i and let $E \subset \pi D_i$ with $m(E) < \varepsilon$. Let E_1, E_2 be disjoint sets whose union is E and $m(E_1) = m(E_2)$.

Select $R \in \mathcal{G}(X)$ with the following properties: for $x \in X \backslash E$, $Rx = x$, $RE_1 = E_2$, $RE_2 = E_1$. Define $\tilde{T} := T(R \times I)$. Note that for j such that $\pi D_j \neq \pi D_i$, $(R \times I)D_j = D_j$, so for such j, $\mu(\tilde{T}D_j \triangle TD_j) = 0$. On the other hand, if $\pi D_j = \pi D_i$, then $\mu(\tilde{T}D_j \triangle TD_j) < m(E) = \varepsilon$. Thus $\tilde{T} \in N_\varepsilon(T)$, but $\tilde{T} \notin \mathcal{G}_{T'}$. □

Question 3. Can one find an extension analog of Katok's result (see, e.g., [**28**]) that rigid transformations form a residual set?

Note that is not completely clear how one should define a rigid extension. The author currently has no notion of how it should be defined.

Acknowledgements. The question of 'typical' behavior of extensions was asked by Terence Tao for a fixed factor, motivated by [**25**], cf. Question 2 and note in particular the following discussion on the difficulties with fixed factors. The author is very grateful to him for the inspiration. The author also thanks Tanja Eisner for introducing him to the problem and for many helpful discussions. The author is further grateful to Ben Stanley for being available to exchange ideas, the referee of this paper for careful reading and valuable

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

comments which improved the paper, and to Bryna Kra, Michael Lin, Philipp Kunde, and Yonatan Gutman for helpful remarks. Lastly, the support of the Max Planck Institute is greatly acknowledged.

References

[**1**] O. N. Ageev. The generic automorphism of a Lebesgue space conjugate to a G-extension for any finite abelian group G. *Dokl. Akad. Nauk* **374** (2000), 439–442.
[**2**] O. N. Ageev. On the multiplicity function of generic group extensions with continuous spectrum. *Ergod. Th. & Dynam. Sys.* **21** (2001), 321–338.
[**3**] O. N. Ageev. On the genericity of some nonasymptotic dynamic properties. *Uspekhi Mat. Nauk* **58** (2003), 177–178.
[**4**] O. N. Ageev. The homogeneous spectrum problem in ergodic theory. *Invent. Math.* **160** (2005), 417–446.
[**5**] S. Alpern and V. S. Prasad. Properties generic for Lebesgue space automorphisms are generic for measure-preserving manifold homeomorphisms. *Ergod. Th. & Dynam. Sys.* **22** (2002), 1587–1620.
[**6**] I. Assani, D. Duncan and R. Moore. Pointwise characteristic factors for Wiener–Wintner double recurrence theorem. *Ergod. Th. & Dynam. Sys.* **36** (2016), 1037–1066.
[**7**] I. Assani and K. Presser. Pointwise characteristic factors for the multiterm return times theorem. *Ergod. Th. & Dynam. Sys.* **32** (2012), 341–360.
[**8**] V. Bergelson, T. Tao and T. Ziegler. Multiple recurrence and convergence results associated to F_P^{ω}-actions. *J. Anal. Math.* **127** (2015), 329–378.
[**9**] R. V. Chacon. Weakly mixing transformations which are not strongly mixing. *Proc. Amer. Math. Soc.* **22** (1969), 559–562.
[**10**] Q. Chu. Convergence of weighted polynomial multiple ergodic averages. *Proc. Amer. Math. Soc.* **137** (2009), 1363–1369.
[**11**] Q. Chu, N. Frantzinakis and B. Host. Ergodic averages of commuting transformations with distinct degree polynomial iterates. *Proc. Lond. Math. Soc. (3)* **102** (2011), 801–842.
[**12**] T. de la Rue and J. de Sam Lazaro. Une transformation générique peut être insérée dans un flot. *Ann. Inst. Henri Poincaré Probab. Stat.* **39** (2003), 121–134.
[**13**] M. Einsiedler and T. Ward. *Ergodic Theory with a View Towards Number Theory*. Springer, New York, 2011.
[**14**] T. Eisner and B. Krause. (Uniform) convergence of twisted ergodic averages. *Ergod. Th. & Dynam. Sys.* **36** (2016), 2172–2202.
[**15**] T. Eisner and P. Zorin-Kranich. Uniformity in the Wiener–Wintner theorem for nilsequences. *Discrete Contin. Dyn. Syst.* **33** (2013), 3497–3516.
[**16**] N. Frantzinakis and P. Zorin-Kranich. Multiple recurrence for non-commuting transformations along rationally independent polynomials. *Ergod. Th. & Dynam. Sys.* **35** (2015), 403–411.
[**17**] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerdi on arithmetic progressions. *J. Anal. Math.* **31** (1977), 204–256.
[**18**] H. Furstenberg, Y. Katznelson and D. Ornstein. The ergodic theoretical proof of Szemeredi's theorem. *Bull. Amer. Math. Soc. (N.S.)* **7**(3) (1982), 527–552.
[**19**] E. Glasner. *Ergodic Theory via Joinings*. American Mathematical Society, Providence, RI, 2003.
[**20**] E. Glasner and B. Weiss. Relative weak mixing is generic. *Preprint*, 2017, arXiv:1707.06425.
[**21**] P.-A. Guihneuf. Dynamical properties of spatial discretizations of a generic homeomorphism. *Ergod. Th. & Dynam. Sys.* **35** (2015), 1474–1523.
[**22**] P. Halmos. In general a measure preserving transformation is mixing. *Ann. of Math. (2)* **45**(4) (1944), 786–792.
[**23**] P. Halmos. *Lectures on Ergodic Theory*. Chelsea Publishing Company, New York, 1956.
[**24**] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)* **161** (2005), 397–488.
[**25**] B. Host, B. Kra and A. Maass. Complexity of nilsystems and systems lacking nilfactors. *J. Anal. Math.* **124** (2014), 261–295.
[**26**] A. B. Katok and A. M. Stepin. Metric properties of homeomorphisms that preserve measure. *Uspekhi Mat. Nauk* **25** (1970), 193–220.
[**27**] J. L. F. King. The generic transformation has roots of all orders. *Colloq. Math.* **84/85** (2000), 521–547.
[**28**] M. G. Nadkarni. Spectral theory of dynamical systems. *Baire Category Theorems of Ergodic Theory*. Birkhäuser, 1998, Ch. 8, pp. 51–61.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press

[**29**] D. Robertson. Characteristic factors for commuting actions of amenable groups. *J. Anal. Math.* **129** (2016), 165–196.
[**30**] E. A. Robinson. The maximal abelian subextension determines weak mixing for group extensions. *Proc. Amer. Math. Soc.* **114** (1992), 443–450.
[**31**] V. Rokhlin. A 'general' measure-preserving transformation is not mixing. *Dokl. Akad. Nauk SSSR* **60** (1948), 349–351.
[**32**] S. Solecki. Closed subgroups generated by generic measure automorphisms. *Ergod. Th. & Dynam. Sys.* **34** (2014), 1011–1017.
[**33**] T. Tao. Poincare's legacies, Part II: pages from year two of a mathematical blog. *Ergodic Theory*. American Mathematical Society, Providence, RI, 2009, Ch. 2, pp. 161–355.
[**34**] T. Tao and T. Ziegler. Concatenation theorems for anti-Gowers-uniform functions and Host–Kra characteristic factors. *Discrete Anal.* (2016).
[**35**] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20** (2007), 53–97.

https://doi.org/10.1017/etds.2017.145 Published online by Cambridge University Press