Title
Relative slow entropy and applications

Author
Lott, Adam

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Relative slow entropy and applications

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Adam Matthew Lott

2023

ABSTRACT OF THE DISSERTATION

Relative slow entropy and applications

by

Adam Matthew Lott
Doctor of Philosophy in Mathematics
University of California, Los Angeles, 2023
Professor Timothy Derek Austin, Chair

Originally developed in the 1950s, Kolmogorov–Sinai entropy is a very powerful isomorphism invariant for measure-preserving systems that measures the "complexity" or "randomness" of a system. There is also a relative version of this notion which measures the *additional* complexity of a system when compared to a fixed reference system. In 1997, Katok–Thouvenot and Ferenczi independently introduced a notion of "slow entropy" as a way to quantitatively compare measure-preserving systems with zero Kolmogorov–Sinai entropy. The goal of this thesis is to develop a relative version of this theory and apply it to several other natural dynamical questions.

Chapters 1 and 2 lay out the definition and basic properties of relative slow entropy. Our definition inherits many desirable properties that make it a natural generalization of both the Katok–Thouvenot/Ferenczi theory and the classical conditional Kolmogorov–Sinai entropy. In Chapter 3, we address the question: under what conditions is a generic extension of a system also isomorphic to the base system? Using relative slow entropy as a tool to prove systems non-isomorphic, we show that a generic extension is *not* isomorphic to the base system whenever the base has zero Kolmogorov–Sinai entropy. Chapter 4 concerns the well-studied notions of isometric and weakly mixing extensions. We give necessary and sufficient entropy-theoretic conditions for extensions to be isometric or weakly mixing. Finally, Chapter 5 investigates the notion of rigidity. Although there is a well known definition for what it means for a single system to be rigid, there is no standard definition for the notion of a rigid extension. We provide a new candidate definition and investigate its consequences. We show that rigid extensions are generic and give an entropy-theoretic sufficient condition for an extension to be rigid. As a consequence, we obtain a new entropy-theoretic characterization of rigid systems which may be of independent interest.

The dissertation of Adam Matthew Lott is approved.

Mario Bonk

Georg Menz

Terence Chi-Shen Tao

Timothy Derek Austin, Committee Chair

University of California, Los Angeles

2023

TABLE OF CONTENTS

1 Introduction . . . **1**

1.1 Historical background . . . 1

1.1.1 Slow entropy . . . 1

1.1.2 Relativization . . . 3

1.2 Notation and definitions . . . 4

1.3 Outline of results . . . 6

2 Basic properties of relative slow entropy . . . **13**

2.1 Monotonicity and isomorphism invariance . . . 13

2.1.1 Upward monotonicity . . . 13

2.1.2 Downward monotonicity . . . 15

2.2 Relatively generating partitions . . . 17

2.3 Relationship to Kolmogorov-Sinai entropy . . . 21

2.4 Open questions . . . 25

3 Non-dominance of zero entropy actions . . . **26**

3.1 Cocycles and generic extensions . . . 27

3.2 Generic extensions have large relative slow entropy . . . 28

3.2.1 Preliminaries . . . 29

3.2.2 Setup for proof of Proposition 3.5 . . . 33

3.2.3 Construction of the perturbed cocycle 35
3.2.4 Estimating the relative covering number 38
3.2.5 Measurability of the perturbed cocycle 41
3.3 Completing the proof of Theorem 3.1 47
3.3.1 Controlling slow entropy of the base system 47
3.3.2 Conclusion of the proof 49
3.4 Open questions 49
4 Isometric extensions 51
4.1 Isometric implies bounded 52
4.2 Background on conditional weak mixing 55
4.3 Bounded implies isometric 56
5 Rigid extensions 67
5.1 Definitions 67
5.2 Genericity 70
5.3 Relationship between rigidity and slow entropy 74
5.4 A necessary condition for rigidity 82
5.5 Open questions 84
References 86

LIST OF TABLES

1.1 Commonly used notation 12

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Tim Austin. Everything I've learned about probability and ergodic theory has been shaped in some form by his always insightful perspective. Most of the results in this thesis originated from projects that he suggested. He has always been patient with me and has never failed to provide a helpful suggestion whenever I get stuck (which is most of the time). Even in non-mathematical matters, he has always gone above and beyond to support me throughout my graduate school experience. It is a huge understatement to say that this thesis would not exist without Tim, and I am extremely grateful to have gotten the chance to be advised by him.

The contents of this thesis were greatly improved by helpful interactions with several other mathematicians. I am grateful to Bryna Kra for introducing me to the work [Fer97], which was the inspiration for Chapter 4, Benjy Weiss for several helpful suggestions regarding the contents of Chapter 3, and Karl Petersen for introducing me to the works [Blu95, Blu97, Blu98].

I have been fortunate to have many amazing math teachers throughout my life. Without their influence, I would never have made the decision to pursue a career in math, and for that I thank (in roughly chronological order) Leslie Wagner, John Giles, Will Rose, and David Stein. Also deserving of a special acknowledgment is Alex Iosevich, my undergraduate advisor at the Univeristy of Rochester. Alex was extremely influential in the development of my mathematical maturity and in my decision to go to graduate school.

There are many factors that positively impacted my time at UCLA. First, my fellow graduate students – I thank Redmond McNamara, Chris Shriver, and James Leng for many interesting conversations about dynamics, and Will Swartworth, Matthew Stone, and Ryan Wallace for all of the time spent preparing for qualifying exams. Second, many thanks are due to all of the math department staff, especially Martha Contreras and Brenda Buenrostro, for all of their tireless work behind the scenes. Special thanks also go to Aryeh and Sharona Kaplan, Naftali and Chayi Hanfling, and Dovid and Elisa Gurevich for keeping me well fed throughout my first few years at UCLA.

Most importantly, I am thankful for my family. I thank my parents for always supporting me and encouraging me to achieve my goals. I am immensely grateful to my wife Maya for agreeing to move to California and constantly supporting me through all of the craziness that comes with being a Ph.D. student. Finally, a huge thank you goes to my son Lior for being very cute.

VITA

2017	B.S. with Honors (Mathematics), University of Rochester
2017	B.A. (Physics), University of Rochester
2019	M.A. (Mathematics), UCLA

CHAPTER 1

Introduction

1.1 Historical background

1.1.1 Slow entropy

In the study of measure-preserving actions of countable amenable groups, one of the most powerful and classical isomorphism invariants is the Kolmogorov–Sinai entropy rate. It was first introduced for actions of $\mathbb{Z}$ by Kolmogorov and Sinai in order to answer the question of whether all Bernoulli shifts are isomorphic [Kol58, Kol59, Sin59a, Sin59b]. The theory was eventually extended to actions of any amenable group by Kieffer, Ornstein, Weiss, and others (see e.g. [KW72, Kie75, OW80, Mou85]), and has been applied to many other problems in ergodic theory. For example, it is known that a system has the K-property if and only if every partition has strictly positive entropy.

One limitation of Kolmogorov–Sinai entropy is that it is not a complete invariant. That is, it is possible for non-isomorphic systems to have the same entropy. In particular, there are many non-isomorphic systems of interest that all have zero entropy, so there has been considerable interest in developing more refined isomorphism invariants specifically designed to distinguish between systems with zero Kolmogorov–Sinai entropy.

The earliest example of this, called "sequence entropy", was developed by Kušnirenko in the 1960s [Ku67]. Many dynamical properties have been successfully characterized using sequence entropy [Ku67, Pic69, Hul82, Zha92, Sal77]. Some examples of other, more specialized entropy-type invariants include "Fried average entropy" [Fri83, KKR14], "scaled entropy" [ZP15, ZP16], "entropy dimension" [Car97, FP07, DHP19], and "entropy convergence rate" [Blu95, Blu97, Blu98]. We will give more details about some of these invariants in Section 2.4.

However, the invariant that has attracted the most attention recently, and the subject of this thesis, is called **slow entropy**. It was developed in 1997 independently and simultaneously by Ferenczi [Fer97] and Katok–Thouvenot [KT97]. One advantage of slow entropy over the other invariants mentioned above is that it can be easily defined for actions of any amenable group.

To motivate the definition of slow entropy, we briefly describe here a non-standard but equivalent way of defining Kolmogorov–Sinai entropy for the special case of an ergodic shift system $(\{0,1\}^{\mathbb{Z}}, \mu)$. Let $\mu_n \in \mathrm{Prob}(\{0,1\}^n)$ denote the marginal distribution of μ on the coordinates $(0, 1, \ldots, n-1)$. Given $\epsilon > 0$, let $C(\mu, n, \epsilon)$ be the smallest number of elements of $\{0,1\}^n$ required to cover a set of μ_n-measure at least $1 - \epsilon$. Then the Kolmogorov–Sinai entropy of μ is given by

$$h_{\mathrm{KS}}(\mu) \;=\; \lim_{\epsilon \to 0} \; \limsup_{n \to \infty} \; \frac{1}{n} \log C(\mu, n, \epsilon).$$

The equivalence of this definition and the standard definition is an easy consequence of the Shannon–McMillan theorem.

The goal of entropy is to quantify how "complex" or "random" the measure μ is. Informally, the quantity $C(\mu, n, \epsilon)$ represents the number of paths of length n

that are likely to occur under μ. As there are 2^n total possible paths, one typically expects this number to grow exponentially in n, so the Kolmogorov–Sinai entropy simply records the exponential growth rate of this quantity. If μ is very random, then there will be a large number of paths which are all roughly equally likely, and the entropy will be large. On the other hand, if μ has low complexity, then there will be a small number of paths that are overwhelmingly more likely to occur than all of the others, and the entropy will be small. If $C(\mu, n, \epsilon)$ grows sub-exponentially, then the entropy is zero. It is natural to hope that a suitable notion of slow entropy can be obtained by simply looking at the growth rate of $C(\mu, n, \epsilon)$ with respect to a different (sub-exponential) rate function. The crucial observation of Ferenczi and Katok–Thouvenot is that in order to make this work, instead of counting the number of *points* required to cover μ_n, one must count the number of *Hamming balls of small radius* required to cover μ_n. This small modification leads to the definition of slow entropy; this will be addressed further in Section 1.2.

1.1.2 Relativization

Another general theme in abstract ergodic theory is the idea of *relativization*, i.e. asking questions about the *relative* behavior of a system with respect to a fixed "base" system, rather than its absolute behavior. A relevant example here is the notion of relative Kolmogorov–Sinai entropy [AR62, WZ92, RW00]. While the absolute entropy measures how much randomness is contained in a system, the relative entropy effectively measures how much *additional* randomness is present in a system that is not already contained in the base system.

There have been many instances in which relative versions of well known theories have been used for new, interesting applications. For example, a relativization of the classical dichotomy between compactness and weak mixing was developed by Furstenberg as a key ingredient in his proof of Szemerédi's theorem [Fur77]. Another example is Thouvenot's development [Tho75] of a relative version of the famous Ornstein theory for detecting Benoullicity [Orn70]. More recently, this relative Ornstein theory was used by Tim Austin to prove the weak Pinsker conjecture [Aus18].

1.2 Notation and definitions

Let G be a countable discrete amenable group. A (left) **Følner sequence** for G is a sequence of finite sets $F_n \subseteq G$ satisfying

$$\lim_{n\to\infty} \frac{|gF_n \cap F_n|}{|F_n|} \;=\; 1$$

for any fixed $g \in G$. Unless otherwise specified, all Følner sequences will be assumed to be left Følner sequences.

Let T be a measure preserving action of G on the standard measure space $(X, \mathcal{B}_X, \mu)$ and write $\mathbf{X} = (X, \mathcal{B}_X, \mu, T)$. For $g \in G$, write $T^g x$ for the action of g on the point $x \in X$, and for a subset $F \subseteq G$, write $T^F x = \left\{T^f x : f \in F\right\}$. For a finite partition $P = \{P_0, \ldots, P_{r-1}\}$ of X, define $P(x)$ to be the unique $i \in \{0, 1, \ldots, r-1\}$ such that $x \in P_i$. Also, for a finite subset $F \subseteq G$, let P^F be the partition $\bigvee_{f\in F} T^{f^{-1}} P$ and let $P^F(x)$ be the $(\mathbf{P}, \mathbf{F})$**-name of x**, i.e. the word

$$(P(T^f x))_{f\in F} \;\in\; \{0, 1, \ldots, r-1\}^F.$$

For any partition P and finite subset $F \subseteq G$, define the pseudo-metric $d_{P,F}$ on X by

$$d_{P,F}(x, x') \;=\; \frac{1}{|F|} \sum_{f\in F} 1_{P(T^f x) \neq P(T^f x')}.$$

This is just the normalized Hamming distance between the two names $P^F(x)$ and $P^F(x')$. For $E \subseteq X$, let $\mathrm{diam}_{P,F}(E)$ be the diameter of E with respect to $d_{P,F}$.

Definition 1.1. Let F be a finite subset of G, λ any probability measure on X, P any partition of X, and $\epsilon > 0$. We denote by

$$\mathrm{cov}(\lambda, P, F, \epsilon)$$

the **Hamming ϵ-covering number** – the smallest M such that there exist sets $E_1, \ldots, E_M \subseteq X$ satisfying $\mathrm{diam}_{P,F}(E_i) \leq \epsilon$ for all i and $\lambda\left(\bigcup E_i\right) \geq 1 - \epsilon$.

We remark that this number is always finite because X is totally bounded when equipped with any of the pseudo-metrics $d_{P,F}$ (in fact, for any P and any F, X is the union of finitely many sets of diameter 0 according to $d_{P,F}$).

Definition 1.2. A **rate function** is an increasing function $U : \mathbb{N} \to (0, \infty)$ such that $U(n) \to \infty$ as $n \to \infty$.

In [KT97] and [Fer97], slow entropy is defined as follows. Let U be a rate function and (F_n) be a Følner sequence for G. Given a partition P, one defines

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}, P) \;=\; \sup_{\epsilon>0} \; \limsup_{n\to\infty} \; \frac{\mathrm{cov}(\mu, P, F_n, \epsilon)}{U(|F_n|)}.$$

Then the slow entropy of $\mathbf{X}$ with respect to the rate function U and Følner sequence (F_n) is defined to be

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}) \;=\; \sup_{P} \; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}, P),$$

where the supremum is taken over all finite partitions of X into measurable sets.

We now relativize this definition to an extension $\pi : \mathbf{X} \to \mathbf{Y} := (Y, \mathcal{B}_Y, \nu, S)$. Let $\mu = \int \mu_y \, d\nu(y)$ be the disintegration of μ over π (see for example [Bil95, Theorem 33.3]).

Definition 1.3. Given a partition P, a finite set $F \subseteq G$, and $\epsilon > 0$, we define the **relative Hamming ϵ-covering number** $\operatorname{cov}(\mu, P, F, \epsilon \,|\, \pi)$ to be the smallest M with the following property: there exists a set $S \in \mathcal{B}_Y$ with $\nu(S) \geq 1 - \epsilon$ such that for any $y \in S$, $\operatorname{cov}(\mu_y, P, F, \epsilon) \leq M$. It can equivalently be defined as

$$\operatorname{cov}(\mu, P, F, \epsilon \,|\, \pi) \;=\; \inf_{S \in \mathcal{B}_Y \,:\, \nu(S) \geq 1-\epsilon} \;\sup_{y \in S}\; \operatorname{cov}(\mu_y, P, F, \epsilon).$$

Definition 1.4. We now define the **relative slow entropy** of π by the analogous formulas

$$\begin{aligned} h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P \,|\, \pi) \;&=\; \sup_{\epsilon>0} \; \limsup_{n\to\infty} \; \frac{\operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi)}{U(|F_n|)} \\ h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) \;&=\; \sup_{P} \; h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P \,|\, \pi). \end{aligned}$$

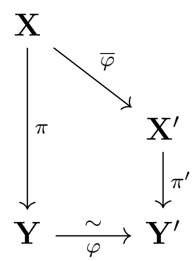
For convenient reference, much of the notation introduced here and later is summarized in Table 1.1.

1.3 Outline of results

The goal of this thesis is to develop a relative version of slow entropy and provide several applications of the theory. In this outline, many theorems will be stated informally or in less than full generality for the sake of readability. For full, precise statements and definitions, the reader should consult the relevant chapter. Most of the work in Chapters 2, 4 and 5 appears in [Lot22a], while Chapter 3 is based on [Lot22b]. Several chapters conclude with an "open questions" section to discuss avenues for further study.

In Chapter 2, the basic properties of relative slow entropy are established. Many of these properties are either direct relativizations of analogous properties of slow entropy or "slowifications" of properties of relative Kolmogorov–Sinai entropy. The first essential property it is that it is monotone in both the top system and bottom system.

Theorem 2.3. *Consider a commutative diagram of the form*

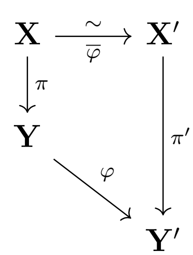


where $\overline{\varphi}$ *is a factor map and* φ *is an isomorphism. Then*

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}' \,|\, \pi') \;\leq\; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi)$$

for any rate function U *and any Følner sequence* (F_n)*.*

Theorem 2.6. *Consider a commutative diagram of the form*



where $\overline{\varphi}$ *is an isomorphism and* φ *is a factor map. Then*

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) \;\leq\; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}' \,|\, \pi')$$

for any rate function U *and any Følner sequence* (F_n)*.*

We also note that it follows immediately from either one of these two results that if $\overline{\varphi}$ and φ are both isomorphisms, then $h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \mid \pi) = h_{\text{slow}}^{U,(F_n)}(\mathbf{X}' \mid \pi')$ for any choice of U and (F_n). We refer to this fact throughout as the "isomorphism invariance of relative slow entropy".

Remark 1.5. Let $\pi : \mathbf{X} \to \mathbf{Y}$ and $\pi' : \mathbf{X}' \to \mathbf{Y}'$ be two extensions. If U is any rate function and (F_n) is any Følner sequence, then in order to show

$$h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \mid \pi) \;\geq\; h_{\text{slow}}^{U,(F_n)}(\mathbf{X}' \mid \pi')$$

it is sufficient to show that for any $\epsilon' > 0$ and partition P' of X', there exist an $\epsilon > 0$ and partition P of X so that

$$\operatorname{cov}(\mu, P, F_n, \epsilon \mid \pi) \;\geq\; \operatorname{cov}(\mu', P', F_n, \epsilon' \mid \pi')$$

for all n sufficiently large.

One of the most important properties of the classical Kolmogorov-Sinai relative entropy is that it can be computed via a sequence of relatively generating partitions (see for example [ELW21, Theorem 2.20]). We show that relative slow entropy also has this property.

Theorem 2.8. *Let $(P_m)_{m=1}^{\infty}$ be a refining sequence of partitions that is generating for $\mathbf{X}$ relative to π. Then*

$$h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \mid \pi) \;=\; \lim_{m\to\infty} h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P_m \mid \pi) \;=\; \sup_m \; h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P_m \mid \pi)$$

for any rate function U and any Følner sequence (F_n).

In the non-relative setting for slow entropy, it is known (see [Fer97, Proposition 2] or [Kat80, Theorem 1.1]) that using an exponential rate function recovers the classical entropy rate. We show the same fact for relative slow entropy.

Theorem 2.14. *Assume that* $\mathbf{X}$ *is ergodic. For* $t > 0$*, let* $U_t(n) = \exp(t \cdot n)$*. Let* (F_n) *be any Følner sequence. Then we have*

$$h_{\mathrm{slow}}^{U_t,(F_n)}(\mathbf{X}, P \,|\, \pi) \;=\; \begin{cases} \infty & \text{if } \; t < h_{\mathrm{KS}}(\mathbf{X}, P \,|\, \pi) \\ 0 & \text{if } \; t > h_{\mathrm{KS}}(\mathbf{X}, P \,|\, \pi) \end{cases}$$

for any partition P*. Equivalently, we have*

$$\sup_{\epsilon > 0} \limsup_{n \to \infty} \frac{1}{|F_n|} \log \operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \;=\; h_{\mathrm{KS}}(\mathbf{X}, P \,|\, \pi)$$

for any partition P*.*

The remainder of the thesis (Chapters 3 to 5) is devoted to several different applications of relative slow entropy. Each of these chapters can be read independently of each other, but they all depend on Chapter 2.

In Chapter 3, we investigate the following question. Given an ergodic base system $\mathbf{Y}$, when is a generic extension $\mathbf{X}$ of $\mathbf{Y}$ also isomorphic to $\mathbf{Y}$? In the special case $G = \mathbb{Z}$, this question was already answered in [AGT21], and the answer is "if and only if $\mathbf{Y}$ has positive Kolmogorov–Sinai entropy". In the same paper, it was also established that for general G, having positive entropy is a sufficient condition. The main result of Chapter 3 is to complete the picture by proving the following.

Theorem 3.1. *Let* G *be any countable amenable group, and let* $\mathbf{Y}$ *be any free ergodic action with zero Kolmogorov–Sinai entropy. Then a generic extension* $\pi : \mathbf{X} \to \mathbf{Y}$ *is not an isomorphism.*

We obtain this as a consequence of the following result about the relative slow entropy of a generic extension, which can be viewed as a relative version of [Ada21, Theorem 1].

Theorem 3.2. *For any Følner sequence (F_n) and any sub-exponential rate function U, the generic extension $\pi : \mathbf{X} \to \mathbf{Y}$ satisfies $h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) = \infty$.*

In Chapter 4, we characterize isometric and weakly mixing extensions in terms of their relative slow entropy. These results are both relativizations and generalizations to all amenable groups of Ferenczi's result [Fer97, Proposition 3].

Theorem 4.4. *Suppose that $\mathbf{X}$ is ergodic. Then the following are equivalent.*

(1) π is an isometric extension.

(2) There exists a Følner sequence (F_n) such that $h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) = 0$ for all rate functions U.

(3) For any Følner sequence (F_n), $h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) = 0$ for all rate functions U.

Corollary 4.17. *Suppose that $\mathbf{X}$ is ergodic. Then π is a weakly mixing extension if and only if for every partition P that is not $\mathbf{Y}$-measurable, there exists a rate function U and a Følner sequence (F_n) such that $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P \,|\, \pi) > 0$.*

Finally, in Chapter 5, we specialize to $G = \mathbb{Z}$ and explore the notion of relative rigidity. Rigidity is a classical dynamical property that has been well studied, but there is no standardized relative version of the theory. We propose a new definition of what it means for an extension π to be rigid and investigate some of its consequences. First, we show that rigid extensions are generic.

Theorem 5.5. *Let $\mathbf{Y}$ be ergodic. Then the generic extension $\pi : \mathbf{X} \to \mathbf{Y}$ is rigid.*

We also obtain a sufficient condition in terms of relative slow entropy.

Theorem 5.10. *Let* $\mathbf{Y}$ *be ergodic. Suppose that there exists a Følner sequence* (F_n) *for* $\mathbb{N}$ *such that* $h_{\text{slow}}^{L,(F_n)}(\mathbf{X} \mid \pi) = 0$*, where* $L(n) = \log n$*. Then* π *is a rigid extension.*

In the non-relative setting, we are also able to give full characterizations of rigidity and mild mixing in terms of slow entropy, which may be of independent interest.

Theorem 5.13. *The following are equivalent.*

(1) $\mathbf{X}$ *is rigid.*

(2) For every rate function U*, there exists a Følner sequence* (F_n) *for* $\mathbb{N}$ *such that* $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}) = 0$.

(3) For the rate function $L(n) = \log n$*, there exists a Følner sequence* (F_n) *for* $\mathbb{N}$ *such that* $h_{\text{slow}}^{L,(F_n)}(\mathbf{X}) = 0$.

Corollary 5.16. *A measure preserving system* $\mathbf{X}$ *is mildly mixing if and only if for all partitions* P *of* X *and all Følner sequences* (F_n) *for* $\mathbb{N}$*, we have* $h_{\text{slow}}^{L,(F_n)}(\mathbf{X}, P) > 0$*, where* $L(n) = \log n$.

G	A countable amenable group
(F_n)	A left Følner sequence for G
$\mathbf{X} = (X, \mathcal{B}_X, \mu, T)$, $\mathbf{Y} = (Y, \mathcal{B}_Y, \nu, S)$	Standard measure-preserving actions of G
π	A factor map from $\mathbf{X}$ to $\mathbf{Y}$
$\mu = \int \mu_y \, d\nu(y)$	The disintegration of μ over π
$P^F(x)$	The (P, F)-name of a point $x \in X$
$d_{P,F}(x, x')$	The normalized Hamming distance between $P^F(x)$ and $P^F(x')$
$\mathrm{diam}_{P,F}(E)$	The diameter of a set E in the pseudo-metric $d_{P,F}$
U	A rate function
P, Q	Partitions of X
$\mathrm{dist}_\mu(P, Q)$	The distance between P and Q according to μ
$\mathrm{dist}_y(P, Q)$	The distance between P and Q according to μ_y
$\mathrm{cov}(\mu, P, F, \epsilon)$	The Hamming ϵ-covering number
$\mathrm{cov}(\mu, P, F, \epsilon \,\vert\, \pi)$	The relative Hamming ϵ-covering number
α, β	Cocycles on Y
$\mathcal{D}_k$	The depth-k dyadic partition of $[0, 1]$

Table 1.1: Commonly used notation

CHAPTER 2

Basic properties of relative slow entropy

2.1 Monotonicity and isomorphism invariance

2.1.1 Upward monotonicity

Definition 2.1. Let $\pi : \mathbf{X} \to \mathbf{Y}$ and $\pi' : \mathbf{X}' \to \mathbf{Y}'$ be extensions. We say that π is an **upward extension** of π' if there is a commutative diagram

$$\begin{array}{ccc} \mathbf{X} & \xrightarrow{\overline{\varphi}} & \mathbf{X}' \\ \big\downarrow{\scriptstyle \pi} & & \big\downarrow{\scriptstyle \pi'} \\ \mathbf{Y} & \xrightarrow[\varphi]{\sim} & \mathbf{Y}' \end{array}$$

where $\overline{\varphi}$ is a factor map and φ is an isomorphism.

Lemma 2.2. *Let π be an upward extension of π' as in Definition 2.1. Then*

$$\overline{\varphi}_* \left(\mu_{\varphi^{-1}y'}\right) \;=\; \mu'_{y'}$$

for ν'-a.e. $y' \in Y'$.

Proof. By the essential uniqueness of disintegrations, it suffices to show the two properties

(1) $\mu' = \int \overline{\varphi}_* \left(\mu_{\varphi^{-1}y'}\right) \, d\nu'(y')$, and

(2) $\overline{\varphi}_* \left(\mu_{\varphi^{-1}y'}\right) ((\pi')^{-1}y') = 1$ for ν'-a.e. y'.

Property (1) is immediate from the definition of a factor map. To show (2), it is sufficient to show that $\int \overline{\varphi}_* \left(\mu_{\varphi^{-1}y'}\right) ((\pi')^{-1}y') \, d\nu'(y') = 1$. This is also immediate from the fact that the diagram in Definition 2.1 commutes. □

Theorem 2.3. *Let π be an upward extension of π'. Then*

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}' \,|\, \pi') \;\leq\; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi)$$

for any rate function U and any Følner sequence (F_n).

Proof. Let $\varphi : \mathbf{Y} \to \mathbf{Y}'$ and $\overline{\varphi} : \mathbf{X} \to \mathbf{X}'$ be the maps as in Definition 2.1. Let $\epsilon > 0$ and let P' be any finite partition of X'. Let P be the partition $\overline{\varphi}^{-1}P'$ of X. Suppose that $\mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi) = L$. Let $S \subseteq Y$ be such that $\nu(Y) \geq 1 - \epsilon$ and $\mathrm{cov}(\mu_y, P, F_n, \epsilon) \leq L$ for all $y \in S$. Fix such a y. We will show that $\mathrm{cov}(\mu'_{\varphi y}, P', F_n, \epsilon) \leq L$ as well.

Let $B_1, \ldots, B_L \subseteq X$ be such that $\mathrm{diam}_{P,F_n}(B_i) \leq \epsilon$ and $\mu_y \left(\bigcup B_i\right) \geq 1 - \epsilon$.

We want to claim that the sets $\overline{\varphi}B_1, \ldots, \overline{\varphi}B_L$ satisfy the analogous properties in X', but because $\overline{\varphi}$ is only a factor map and not necessarily an isomorphism, these sets need not be measurable. So we define B_i' to be the union of all $(P')^{F_n}$-cells that meet $\overline{\varphi}B_i$, and we show that the sets $B_1', \ldots, B_L'$ have the right properties.

To check the diameter condition, note that by construction, $\mathrm{diam}_{P',F_n} B_i' = \mathrm{diam}_{P',F_n}(\overline{\varphi}B_i)$, so it suffices to estimate the latter. Observe that if $\overline{\varphi}x, \overline{\varphi}z \in \overline{\varphi}B_i$, then

$$
\begin{aligned}
d_{P',F_n}(\overline{\varphi}x, \overline{\varphi}z) &= \frac{1}{|F_n|} \sum_{f \in F_n} 1_{P'(T'^f \overline{\varphi}x) \neq P'(T'^f \overline{\varphi}z)} = \frac{1}{|F_n|} \sum_{f \in F_n} 1_{P'(\overline{\varphi}T^f x) \neq P'(\overline{\varphi}T^f z)} \\
&= \frac{1}{|F_n|} \sum_{f \in F_n} 1_{P(T^f x) \neq P(T^f z)} = d_{P,F_n}(x, z) \le \epsilon,
\end{aligned}
$$

so $\mathrm{diam}_{P',F_n}(\overline{\varphi}B_i) \le \epsilon$.

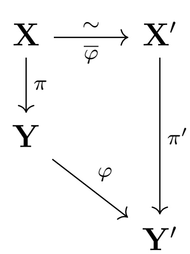
Now to check the measure condition, apply Lemma 2.2 to write

$$
\mu'_{\varphi y}\left(\bigcup B'_i\right) = \mu_y\left(\overline{\varphi}^{-1} \bigcup B'_i\right) = \mu_y\left(\bigcup \overline{\varphi}^{-1} B'_i\right) \ge \mu_y\left(\bigcup B_i\right) \ge 1 - \epsilon.
$$

This shows that $\mathrm{cov}(\mu'_{\varphi y}, P', F_n, \epsilon) \le L$ for all $y \in S$. Taking $S' = \varphi(S)$, we have $\nu'(S') \ge 1 - \epsilon$ as well, so $\mathrm{cov}(\mu', P', F_n, \epsilon \,|\, \pi) \le L = \mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi)$ for all n. Since this holds for any P' and any ϵ, the desired result follows. □

2.1.2 Downward monotonicity

Definition 2.4. Let $\pi : \mathbf{X} \to \mathbf{Y}$ and $\pi' : X' \to \mathbf{Y}'$ be extensions. We say that π' is a **downward extension** of π if there is a commutative diagram of the form



where $\overline{\varphi}$ is an isomorphism and φ is a factor map.

Lemma 2.5. *Let π' be a downward extension of π as in Definition 2.4. Write $\nu = \int \nu_{y'}\, d\nu'(y')$ for the disintegration of ν over φ. Then we have*

$$\mu'_{y'} \;=\; \overline{\varphi}_* \int \mu_y \, d\nu_{y'}(y)$$

for ν'-a.e. $y' \in Y'$.

Proof. Denote the right hand side of the above by $\lambda_{y'}$. By the essential uniqueness of disintegration, it suffices to show the two properties

(1) $\mu' = \int \lambda_{y'}\, d\nu'(y')$, and

(2) for ν'-a.e. y', $\lambda_{y'}$ is supported on the fiber $(\pi')^{-1}y'$.

Property (1) follows from the fact that $\iint \mu_y \, d\nu_{y'}(y)\, d\nu(y) = \int \mu_y \, d\nu(y) = \mu$. To see property (2), note that $\int \mu_y \, d\nu_{y'}(y)$ is a mixture of measures that are all supported on $\pi^{-1}\varphi^{-1}y' = \overline{\varphi}^{-1}(\pi')^{-1}y'$. Therefore $\lambda_{y'}$ is supported on $(\pi')^{-1}y'$ as desired. □

Theorem 2.6. *Suppose that π' is a downward extension of π. Then*

$$h^{U,(F_n)}_{\text{slow}}(\mathbf{X} \,|\, \pi) \;\leq\; h^{U,(F_n)}_{\text{slow}}(\mathbf{X}' \,|\, \pi')$$

for any rate function U and any Følner sequence (F_n).

Proof. Let $\epsilon > 0$ and let P be any partition of X, and set $P' = \overline{\varphi}P$. Suppose that $\operatorname{cov}(\mu', P', F_n, \epsilon^2/4 \,|\, \pi') = L$. It suffices to show that $\operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \leq L$ as well.

Let $S' \subseteq Y'$ be the set of $y' \in Y'$ that satisfy $\operatorname{cov}(\mu'_{y'}, P', F_n, \epsilon^2/4) \leq L$ and note that by definition, $\nu'(S') \geq 1 - \epsilon^2/4$. Fix $y' \in S'$. Let $B'_1, \ldots, B'_L$ be subsets of X' satisfying $\operatorname{diam}_{P',F_n}(B'_i) \leq \epsilon^2/4$ and $\mu'_{y'}(\bigcup B'_i) \geq 1 - \epsilon^2/4$. Set $B_i = \overline{\varphi}^{-1}B'_i$. First, note that it follows immediately from the fact that $\overline{\varphi}$ is an isomorphism that $\operatorname{diam}_{P,F_n}(B_i) = \operatorname{diam}_{P',F_n}(B'_i) \leq \epsilon^2/4$.

We now show that for most y, the sets B_i cover most of μ_y. By Lemma 2.5, we have

$$1-\epsilon^2/4 \;\le\; \mu'_{y'}\left(\bigcup B'_i\right) \;=\; \int \mu_y\left(\overline{\varphi}^{-1}\bigcup B'_i\right)\, d\nu_{y'}(y) \;=\; \int \mu_y\left(\bigcup B_i\right)\, d\nu_{y'}(y),$$

so Markov's inequality implies that the set

$$S(y') \;:=\; \left\{y \in \varphi^{-1}y' : \mu_y\left(\bigcup B_i\right) \ge 1-\epsilon/2\right\}$$

satisfies $\nu_{y'}(S(y')) \ge 1-\epsilon/2$. This shows that any $y \in S(y')$ satisfies

$$\mathrm{cov}(\mu_y, P, F_n, \epsilon) \;\le\; L.$$

Finally, this construction was valid for any $y' \in S'$. Therefore, let

$$S \;:=\; \{y \in Y : \mathrm{cov}(\mu_y, P, F_n, \epsilon) \le L\}$$

and observe that S contains $S(y')$ for all $y' \in S'$. So we conclude that

$$\begin{aligned}\nu(S) \;&=\; \int \nu_{y'}(S)\, d\nu'(y') \;\ge\; \int_{y'\in S'} \nu_{y'}(S)\, d\nu'(y') \;\ge\; \int_{y'\in S'} \nu_{y'}(S(y'))\, d\nu'(y')\\ &\ge\; (1-\epsilon/2)^2 \;\ge\; 1-\epsilon,\end{aligned}$$

implying that $\mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \le L$ as desired. □

2.2 Relatively generating partitions

Definition 2.7. A partition P of X is said to be **generating for X relative to** π if

$$P^G \vee \pi^{-1}\mathcal{B}_Y \;=\; \mathcal{B}_X \mod \mu.$$

Here we have identified in the natural way the partition P^G with the σ-algebra that it induces. Similarly, a sequence of partitions $(P_m)_{m=1}^{\infty}$ is said to be generating for $\mathbf{X}$ relative to π if

$$\left(\bigvee_{m=1}^{\infty} P_m^G\right) \vee \pi^{-1}\mathcal{B}_Y \;=\; \mathcal{B}_X \mod \mu.$$

The sequence is also said to be **refining** if P_{m+1} refines P_m for every m.

One of the most important properties of the classical Kolmogorov-Sinai relative entropy is that it can be computed via a sequence of relatively generating partitions [ELW21, Theorem 2.20]. In this section, we show an analogous result for relative slow entropy.

Theorem 2.8. *Let $(P_m)_{m=1}^{\infty}$ be a refining sequence of partitions that is generating for $\mathbf{X}$ relative to π. Then*

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}\,|\,\pi) \;=\; \lim_{m\to\infty} h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X},P_m\,|\,\pi) \;=\; \sup_m \; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X},P_m\,|\,\pi)$$

for any rate function U and any Følner sequence (F_n).

Remark 2.9. In the case where $\mathbf{Y}$ is trivial, this reduces to the non-relative version of the same result (see [KT97, Proposition 1] or [Fer97, Lemma 1]).

Definition 2.10. For finite partitions $P = \{P_1,\dots,P_r\}$ and $Q = \{Q_1,\dots,Q_r\}$ of X and a probability measure $\lambda \in \mathrm{Prob}(X)$, the partition distance with respect to λ is defined as

$$\mathrm{dist}_\lambda(P,Q) \;=\; \lambda\{x\in X : P(x)\neq Q(x)\} \;=\; \frac{1}{2}\sum_{i=1}^{r}\lambda(P_i \triangle Q_i).$$

Lemma 2.11. *For any partition Q of X and any $0<\epsilon<1$, there exists $\epsilon'>0$ such that if Q' is any other partition of X satisfying $\mathrm{dist}_\mu(Q,Q')\le\epsilon'$, then*

$$\mathrm{cov}(\mu,Q,F_n,\epsilon\,|\,\pi) \;\le\; \mathrm{cov}(\mu,Q',F_n,\epsilon'\,|\,\pi) \qquad \textit{for all } n \textit{ sufficiently large.}$$

Proof. Set $\epsilon' = (\epsilon/2)^4$. Suppose that Q' satisfies $\mathrm{dist}_\mu(Q, Q') = \mu\{x : Q(x) \neq Q'(x)\} \leq \epsilon'$ and define

$$X' := \left\{ x : \frac{1}{|F_n|} \sum_{f \in F_n} 1_{Q(T^f x) \neq Q'(T^f x)} \leq \sqrt{\epsilon'} \right\}.$$

By Markov's inequality, we have

$$\mu(X \setminus X') \leq \frac{1}{\sqrt{\epsilon'}} \int \frac{1}{|F_n|} \sum_{f \in F_n} 1_{\{z : Q(z) \neq Q'(z)\}}(T^f x)\, d\mu(x) \leq \frac{1}{\sqrt{\epsilon'}} \mathrm{dist}_\mu(Q, Q') \leq \sqrt{\epsilon'}.$$

Applying Markov's inequality again, we obtain a set $S_1 \subseteq Y$ such that $\nu(S_1) \geq 1 - (\epsilon')^{1/4} = 1 - \epsilon/2$ such that $\mu_y(X') \geq 1 - (\epsilon')^{1/4} = 1 - \epsilon/2$ for all $y \in S_1$.

Now let $L = \mathrm{cov}(\mu, Q', F_n, \epsilon' \,|\, \pi)$ and let $S_2 \subseteq Y$ be such that $\nu(S_2) \geq 1 - \epsilon'$ and

$$\mathrm{cov}(\mu_y, Q', F_n, \epsilon') \leq L$$

for $y \in S_2$. We claim that $\mathrm{cov}(\mu_y, Q, F_n, \epsilon) \leq L$ for $y \in S := S_1 \cap S_2$. Because $\nu(S) \geq 1 - \epsilon' - \epsilon/2 \geq 1 - \epsilon$, this is enough. Fix $y \in S$ and let $B_1, \ldots, B_L \subset X$ be such that $\mathrm{diam}_{Q', F_n}(B_i) \leq \epsilon'$ and $\mu_y\left(\bigcup B_i\right) \geq 1 - \epsilon'$. Let $B_i' = B_i \cap X'$. Notice that we have $\mu_y\left(\bigcup B_i'\right) \geq 1 - \epsilon' - \epsilon/2 \geq 1 - \epsilon$, so we just need to estimate $\mathrm{diam}_{Q, F_n}(B_i')$. If $x, z \in B_i'$, then

$$\begin{aligned} d_{Q,F_n}(x, z) &= \frac{1}{|F_n|} \sum_{f \in F_n} 1_{Q(T^f x) \neq Q(T^f z)} \\ &\leq \frac{1}{|F_n|} \sum_{f \in F_n} 1_{Q(T^f x) \neq Q'(T^f x)} + \frac{1}{|F_n|} \sum_{f \in F_n} 1_{Q'(T^f x) \neq Q'(T^f z)} \\ &\qquad + \frac{1}{|F_n|} \sum_{f \in F_n} 1_{Q'(T^f z) \neq Q(T^f z)} \\ &\leq \sqrt{\epsilon'} + \epsilon' + \sqrt{\epsilon'}. \end{aligned}$$

The estimates on the first and third terms hold because $x, z \in X'$, and the estimate on the second term holds because $\mathrm{diam}_{Q', F_n}(B_i') \leq \mathrm{diam}_{Q', F_n}(B_i) \leq \epsilon'$. Therefore we have $\mathrm{diam}_{Q, F_n}(B_i') \leq 3\sqrt{\epsilon'} = 3\epsilon^2/4 \leq \epsilon$ as desired. □

Lemma 2.12. *Let P be any partition of X, and let Q' be another partition which is measurable with respect to $P^{F_k} \vee \pi^{-1}\mathcal{B}_Y$ for some $k \in \mathbb{N}$. Then for any $\epsilon > 0$, there exists ϵ' such that*

$$\operatorname{cov}(\mu, Q', F_n, \epsilon \,|\, \pi) \;\leq\; \operatorname{cov}(\mu, P, F_n, \epsilon' \,|\, \pi) \qquad \text{for all } n \text{ sufficiently large.}$$

Proof. Set $\epsilon' = \epsilon/(2|F_k|)$. Let $\operatorname{cov}(\mu, P, F_n, \epsilon' \,|\, \pi) = L$ and let $S \subseteq Y$ be such that $\nu(S) \geq 1 - \epsilon'$ and $\operatorname{cov}(\mu_y, P, F_n, \epsilon') \leq L$ for all $y \in S$. We claim that also $\operatorname{cov}(\mu_y, Q', F_n, \epsilon) \leq L$ for $y \in S$. Fix $y \in S$ and let $B_1, \ldots, B_L$ be such that $\operatorname{diam}_{P,F_n}(B_i) \leq \epsilon'$ and $\mu_y\left(\bigcup B_i\right) \geq 1 - \epsilon'$. We may assume without loss of generality that each B_i is contained in the fiber $\pi^{-1}y$ because doing so can only decrease their diameters and does not change their measure according to μ_y.

We will be done as soon as we show that $\operatorname{diam}_{Q',F_n}(B_i) \leq \epsilon$. To do that, let $x, z \in B_i$ and observe that because x and z lie in the same fiber of π, $Q'(x) \neq Q'(z)$ implies that x and z must be in different cells of P^{F_k}. Therefore we can estimate

$$\begin{aligned}
d_{Q',F_n}(x,z) &= \frac{1}{|F_n|}\sum_{f\in F_n} 1_{Q'(T^f x)\neq Q'(T^f z)} \;\leq\; \frac{1}{|F_n|}\sum_{f\in F_n} 1_{P^{F_k}(T^f x)\neq P^{F_k}(T^f z)} \\
&\leq \frac{1}{|F_n|}\sum_{f\in F_n}\sum_{g\in F_k} 1_{P(T^{gf}x)\neq P(T^{gf}z)} \\
&= \frac{1}{|F_n|}\sum_{h\in F_kF_n} \#\{(g,f)\in F_k\times F_n : gf = h\}\cdot 1_{P(T^h x)\neq P(T^h z)} \\
&\leq \frac{1}{|F_n|}\sum_{h\in F_kF_n} |F_k|\cdot 1_{P(T^h x)\neq P(T^h z)} \\
&= \frac{1}{|F_n|}\sum_{h\in F_n} |F_k|\cdot 1_{P(T^h x)\neq P(T^h z)} + \frac{1}{|F_n|}\sum_{h\in F_kF_n\setminus F_n} |F_k|\cdot 1_{P(T^h x)\neq P(T^h z)} \\
&\leq |F_k|\cdot d_{P,F_n}(x,z) + \frac{|F_k|\cdot|F_kF_n\setminus F_n|}{|F_n|}
\end{aligned}$$

So as soon as n is sufficiently large, because (F_n) is a Følner sequence, we have $\mathrm{diam}_{Q',F_n}(B_i) \le 2|F_k|\epsilon' = \epsilon$ as desired. □

Proof of Theorem 2.8. To prove this theorem, it is clearly sufficient to prove the following finitary version: for any partition Q of X and any $\epsilon > 0$, there exist $m \in \mathbb{N}$ and $\epsilon' > 0$ such that

$$\mathrm{cov}(\mu, Q, F_n, \epsilon \,|\, \pi) \;\le\; \mathrm{cov}(\mu, P_m, F_n, \epsilon' \,|\, \pi)$$

for all sufficiently large n.

Fix a partition Q and $\epsilon > 0$. Let ϵ' be as in the statement of Lemma 2.11. Then, by the definition of relatively generating sequence of partitions, we can find $m, k \in \mathbb{N}$ and a partition Q' refined by $P_m^{F_k} \vee \pi^{-1}\mathcal{B}_Y$ such that $\mathrm{dist}_\mu(Q, Q') \le \epsilon'$. Apply Lemma 2.11 to conclude that

$$\mathrm{cov}(\mu, Q, F_n, \epsilon \,|\, \pi) \;\le\; \mathrm{cov}(\mu, Q', F_n, \epsilon' \,|\, \pi) \tag{2.1}$$

for n sufficiently large. Then apply Lemma 2.12 with ϵ' in place of ϵ to produce an ϵ'' satisfying

$$\mathrm{cov}(\mu, Q', F_n, \epsilon' \,|\, \pi) \;\le\; \mathrm{cov}(\mu, P_m, F_n, \epsilon'' \,|\, \pi) \tag{2.2}$$

for sufficiently large n. Combining (2.1) and (2.2) gives the desired result. □

2.3 Relationship to Kolmogorov-Sinai entropy

In this section, we show that with an exponential rate function, relative slow entropy recovers the classical relative Kolomogorov-Sinai entropy. Given an extension $\pi : \mathbf{X} \to \mathbf{Y}$ and a partition P of X, let $h_{KS}(\mathbf{X}, P \,|\, \pi)$ denote the relative Kolmogorov-Sinai entropy rate (see e.g. [ELW21, Definition 2.18] for the definition).

Definition 2.13. Define

$$h'(\mathbf{X}, P \,|\, \pi) \;:=\; \sup_{\epsilon>0} \limsup_{n\to\infty} \frac{1}{|F_n|} \log \operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi).$$

Note that *a priori* this quantity depends on the choice of Følner sequence, but the next theorem shows that it actually does not.

Theorem 2.14. *Assume that* $\mathbf{X}$ *is ergodic. For any partition* P *of* X*, we have* $h'(\mathbf{X}, P \,|\, \pi) = h_{\mathrm{KS}}(\mathbf{X}, P \,|\, \pi)$.

The non-relative version of this result appears in [Fer97, Proposition 2]

Proof. For this proof, we will abuse notation and write $P^{F_n}(x)$ to mean the cell of the partition P^{F_n} that contains x, rather than the name of the cell. Abbreviate $h = h_{\mathrm{KS}}(\mathbf{X}, P \,|\, \pi)$ and $h' = h'(\mathbf{X}, P \,|\, \pi)$.

Step 1: setup. Fix an arbitrary $\gamma > 0$. For $y \in Y$, define

$$\mathcal{G}_{y,n} \;:=\; \left\{ x \in X : \exp(|F_n|(-h-\gamma)) < \mu_y \left(P^{F_n}(x)\right) < \exp(|F_n|(-h+\gamma)) \right\}.$$

Also define

$$\mathcal{G}_{n} \;:=\; \left\{ x \in X : \exp(|F_n|(-h-\gamma)) < \mu_{\pi x} \left(P^{F_n}(x)\right) < \exp(|F_n|(-h+\gamma)) \right\}.$$

By the amenable version of the relative Shannon-McMillan theorem (see for example [WZ92, Theorem 3.2] or [RW00, Corollary 4.6]), $\mu(\mathcal{G}_n) \to 1$ as $n \to \infty$. Because $\mu = \int \mu_y \, d\nu(y)$, this implies also $\mu_y(\mathcal{G}_n) \to 1$ for ν-a.e. y. Finally, since μ_y is supported only on $\pi^{-1}y$, we have $\mu_y(\mathcal{G}_{y,n}) = \mu_y(\mathcal{G}_n) \to 1$ for ν-a.e. y.

Step 2: $\mathbf{h' \le h}$. Let $\epsilon > 0$. Because $\mu_y(\mathcal{G}_{y,n}) \to 1$ for ν-a.e. y, for all n sufficiently large we can find a set $S \subseteq Y$ such that $\nu(S) \geq 1-\epsilon$ and $\mu_y(\mathcal{G}_{y,n}) \geq 1-\epsilon$

for all $y \in S$. We now estimate $\mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi)$. Because $\nu(S) \geq 1-\epsilon$, it suffices to estimate $\mathrm{cov}(\mu_y, P, F_n, \epsilon)$ for $y \in S$. Fix such a y; we claim that $\mathrm{cov}(\mu_y, P, F_n, \epsilon) \leq \exp(|F_n|(h+\gamma))$. Let $C_1, \ldots, C_L$ be all of the cells of P^{F_n} that meet $\mathcal{G}_{y,n}$. Then each $\mathrm{diam}_{P,F_n}(C_i) = 0$ and

$$\mu_y\left(\bigcup C_i\right) \;=\; \mu_y(\mathcal{G}_{y,n}) \;\geq\; 1-\epsilon,$$

so $\mathrm{cov}(\mu_y, P, F_n, \epsilon) \leq L$. But by definition of $\mathcal{G}_{y,n}$, each C_i has μ_y−measure at least

$$\exp(|F_n|(-h-\gamma)),$$

so $L \leq \exp(|F_n|(h+\gamma))$ as claimed. Since this holds for all $y \in S$, this shows that

$$\mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \;\leq\; \exp(|F_n|(h+\gamma))$$

for sufficiently large n. Therefore we can take $n \to \infty$ and then $\epsilon \to 0$ to conclude $h' \leq h + \gamma$. But since γ is arbitrary we get $h' \leq h$ as desired.

Step 3: $\mathbf{h' \geq h}$. Again fix $0 < \epsilon < 1/4$ and let n and S be as in step 2. Also let us enumerate $P = \{P_0, \ldots, P_{r-1}\}$. This time we will estimate a lower bound for $\mathrm{cov}(\mu_y, P, F_n, \epsilon)$ for any $y \in S$. Suppose $E_1, \ldots, E_M$ are sets in X satisfying $\mathrm{diam}_{P,F_n}(E_i) \leq \epsilon$ and $\mu_y(\bigcup E_i) \geq 1-\epsilon$. Without loss of generality, we may assume that each E_i is a union of P^{F_n}-cells because replacing each E_i by $\bigcup_{x \in E_i} P^{F_n}(x)$ makes each E_i larger without changing its diameter according to d_{P,F_n}. Because $y \in S$, we then have

$$\mu_y\left(\bigcup E_i \cap \mathcal{G}_{y,n}\right) \;\geq\; 1-2\epsilon,$$

and we can estimate

$$1/2 \;\leq\; 1-2\epsilon \;\leq\; \mu_y\left(\bigcup E_i \cap \mathcal{G}_{y,n}\right) \;\leq\; \sum_{i=1}^{M} \mu_y(E_i \cap \mathcal{G}_{y,n}).$$

Each E_i is a union of P^{F_n}-cells, and any P^{F_n}-cell meeting $\mathcal{G}_{y,n}$ has μ_y-measure at most

$$\exp(|F_n|(-h+\gamma))$$

by construction. Therefore the above sum is at most

$$\sum_{i=1}^{M} \exp(|F_n|(-h+\gamma)) \cdot (\# \text{ of } P^{F_n}\text{-cells contained in } E_i).$$

Since E_i has diameter $\leq \epsilon$ according to d_{P,F_n}, the elements of $\{0,1,\dots,r-1\}^{F_n}$ corresponding to the P^{F_n}-cells contained in E_i all fit inside a fixed ball of radius ϵ in the normalized Hamming metric on $\{0,1,\dots,r-1\}^{F_n}$. It is well known (for example, it is an easy consequence of [Gra11, Lemma 3.6]) that the number of words in any such Hamming ball is at most

$$\exp\Big(|F_n| \cdot (\epsilon \log(r-1) + H(\epsilon, 1-\epsilon))\Big),$$

where H is the Shannon entropy function $H(t,1-t) = -t\log t - (1-t)\log(1-t)$. Therefore the above sum is bounded by

$$M \cdot \exp\Big(|F_n|(-h+\gamma+\epsilon\log(r-1) + H(\epsilon,1-\epsilon))\Big),$$

implying that

$$M \;\geq\; (1/2)\exp\Big(|F_n|(h-\gamma-\epsilon\log(r-1)+H(\epsilon,1-\epsilon))\Big) \;=:\; Z.$$

Since we started with an arbitrary covering set this implies that $\mathrm{cov}(\mu_y, P, F_n, \epsilon) \geq Z$ for every $y \in S$. Since $\nu(S) > 1-\epsilon$, it has positive ν-measure intersection with any other set of ν-measure $\geq 1-\epsilon$, so it is impossible to find a different set S' with $\mu(S') > 1-\epsilon$ and $\mathrm{cov}(\mu_y, P, F_n, \epsilon) \leq Z$ for every $y \in S'$. Therefore $\mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \geq Z$. Taking $n \to \infty$ gives

$$\limsup_{n\to\infty} \frac{1}{|F_n|} \log \mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \;\geq\; h - \gamma - \log(r) \cdot H(\epsilon, 1-\epsilon),$$

then taking $\epsilon \to 0$ gives $h' \geq h - \gamma$, and again γ is arbitrary so we conclude $h' \geq h$. $\square$

2.4 Open questions

In this section, we assume $G = \mathbb{Z}$.

In [Blu95,Blu97], Blume defines a notion of "entropy convergence rate" as follows. Let $a(n)$ be a sub-linear rate function, i.e. a rate function satisfying $\lim_{n\to\infty} a(n)/n = 0$. Then, given a partition P of X, one defines

$$h_a(\mathbf{X}, P) \;:=\; \limsup_{n\to\infty} \frac{1}{a(n)} H\big(\mu, P^{[0,n)}\big) \tag{2.3}$$

where H denotes the Shannon entropy function. Then, as usual, the definition is completed by letting

$$h_a(\mathbf{X}) \;=\; \sup_P \; h_a(\mathbf{X}, P).$$

Note that using the rate function $a(n) = n$ yields the standard definition of Kolmogorov–Sinai entropy, so this definition is very similar in sprit to that of slow entropy. However, Adams shows in [Ada21] that they actually behave very differently. Specifically, given any sub-linear rate a and any sub-exponential rate U, he gives an explicit construction of a system $\mathbf{X}$ that satisfies $h_{\mathrm{slow}}^{U,([0,n))}(\mathbf{X}) = 0$ and $h_a(\mathbf{X}) = \infty$.

One can easily define a relative version of entropy convergence rate by substituting the conditional Shannon entropy into (2.3). Then it is natural to ask if Adams's result also holds in the relative setting.

Question 2.15. Given a base system $\mathbf{Y}$, a sub-linear rate function a, and a sub-exponential rate function U, is there always an example of an extension $\pi : \mathbf{X} \to \mathbf{Y}$ that satisfies $h_{\mathrm{slow}}^{U,([0,n))}(\mathbf{X} \,|\, \pi) = 0$ and $h_a(\mathbf{X} \,|\, \pi) = \infty$?

CHAPTER 3

Non-dominance of zero entropy actions

Let **Y** be free and ergodic. It is natural to ask what properties of **Y** are preserved by a generic extension **X** (a precise definition of "generic extension" is discussed in section 3.1). For example, it was shown in [GTW21] that a generic **X** has the same Kolmogorov–Sinai entropy as **Y** and that if **Y** is a nontrivial Bernoulli shift, then a generic **X** is also Bernoulli. The system **Y** is said to be **dominant** if a generic extension **X** is actually *isomorphic* to **Y**. So, for example, the aforementioned results from [GTW21] together with Ornstein's famous isomorphism theorem [Orn70] imply that all nontrivial Bernoulli shifts are dominant. More generally, it has been shown in [AGT21] that

(1) if $G = \mathbb{Z}$, then **Y** is dominant if and only if it has positive Kolmogorov–Sinai entropy, and

(2) for any G, if **Y** has positive Kolmogorov–Sinai entropy, then it is dominant.

In this chapter, we complete the picture by proving the following result.

Theorem 3.1. *Let G be any discrete amenable group, and let* **Y** *be any free ergodic action with zero Kolmogorov–Sinai entropy. Then* **Y** *is not dominant.*

3.1 Cocycles and generic extensions

Let I be the unit interval $[0,1]$, and let m be Lebesgue measure on I. Denote by $\mathrm{Aut}(I,m)$ the group of invertible m-preserving transformations of I, modulo the equivalence relation of m-a.e. agreement. A **Rokhlin cocycle** on Y is a family of measurable maps $\alpha_g : Y \to \mathrm{Aut}(I,m)$ indexed by $g \in G$ that satisfies the **cocycle condition**: for every $g, h \in G$ and ν-a.e. y, $\alpha_{hg}(y) = \alpha_h(S^g y) \circ \alpha_g(x)$. From now on, we will always simply write "cocyle" to mean a Rokhlin cocycle. A cocycle can equivalently be thought of as a measurable map $\alpha : R \to \mathrm{Aut}(I,m)$, where $R \subseteq Y \times Y$ is the orbit equivalence relation induced by S (i.e. $(y,y') \in R$ if and only if $y' = S^g y$ for some $g \in G$). With this perspective, the cocycle condition takes the form $\alpha(y_1,y_3) = \alpha(y_2,y_3) \circ \alpha(y_1,y_2)$. A cocycle α induces the skew product extension T_α on the larger space $X := Y \times I$ defined by

$$T^g_\alpha(y,t) \ := \ (S^g y, \alpha_g(y)(t)).$$

This action preserves the measure $\mu := \nu \times m$ and is an extension of the original action $\mathbf{Y}$. We denote the skew product system (X, T_α, μ) by $\mathbf{X}_\alpha$.

By a classical theorem of Rokhlin (see for example [Gla03, Theorem 3.18]), any infinite-to-one ergodic extension of $\mathbf{Y}$ is isomorphic to $\mathbf{X}_\alpha$ for some cocycle α. Therefore, by topologizing the space of all cocycles on Y we can capture the notion of a "generic" extension – a property is said to hold for a generic extension if it holds for a dense G_δ set of cocycles. Denote the space of all cocycles on Y by $\mathrm{Co}(Y)$. Topologizing $\mathrm{Co}(Y)$ is done in a few stages.

(1) On $\mathrm{Aut}(I,m)$, consider the weak topology defined by the property that a sequence (φ_n) converges to φ if and only if $m(\varphi_n^{-1}E \triangle \varphi^{-1}E) \to 0$ for all Borel

sets $E \subseteq I$.

(2) With this topology, $\mathrm{Aut}(I, m)$ becomes a Polish space. Let d_A be a compatible complete metric such that $(\mathrm{Aut}(I, m), d_A)$ has diameter 1 (see [Kec10, Section 1.1] for details).

(3) If α_0, β_0 are measurable maps $Y \to \mathrm{Aut}(I, m)$, then define

$$\mathrm{dist}(\alpha_0, \beta_0) \;:=\; \int d_A(\alpha_0(y), \beta_0(y))\, d\nu(y).$$

(4) The metric defined in the previous step induces a topology on $\mathcal{A} :=$ the set of all measurable maps $Y \to \mathrm{Aut}(I, m)$. Therefore, because $\mathrm{Co}(Y)$ is just a certain (closed) subset of $\mathcal{A}^G$, it just inherits the product topology.

To summarize, given a cocycle α, a basic open neighborhood of α is specified by two parameters: a finite subset $F \subseteq G$ and $\eta > 0$. The (F, η)-neighborhood of α is $\{\beta \in \mathrm{Co}(Y) : \mathrm{dist}(\alpha_g, \beta_g) < \eta \text{ for all } g \in F\}$. In practice, we will always arrange things so that there is a set of y of measure $\geq 1 - \eta$ on which $\alpha_g(y) = \beta_g(y)$ for all $g \in F$, which is sufficient to guarantee that β is in the (F, η)-neighborhood of α.

3.2 Generic extensions have large relative slow entropy

Let (F_n) be a Følner sequence for G and let $\mathbf{Y}$ be a system with zero Kolmogorov–Sinai entropy. Given a cocycle α, let $\pi : \mathbf{X}_\alpha \to \mathbf{Y}$ be the canonical factor map.

We will say that a rate function U is **sub-exponential** if

$$\lim_{n\to\infty} \frac{1}{n} \log U(n) \;=\; 0.$$

The purpose of this section is to prove the following.

Theorem 3.2. *For any sub-exponential rate function U, there exists a dense G_δ set $\mathscr{C}$ of cocycles α such that $h^{U,(F_n)}_{\mathrm{slow}}(\mathbf{X}_\alpha \,|\, \pi) = \infty$ for all $\alpha \in \mathscr{C}$.*

We will actually prove a formally weaker version of the above in which "$= \infty$" is replaced by "≥ 2". These are clearly equivalent because one may always replace a sub-exponential rate function U by a slightly faster rate function U' that satisfies

$$\lim_{n\to\infty} \frac{U'(n)}{U(n)} \;=\; \infty$$

but is still sub-exponential.

3.2.1 Preliminaries

Let U be a fixed sub-exponential rate function and let P be the fixed partition

$$\{Y \times [0, 1/2), Y \times [1/2, 1]\}$$

of X. To emphasize the dependence on the cocycle α, in this section we replace the notation $\mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi)$ by $\mathrm{cov}(\mu, \alpha, P, F_n, \epsilon \,|\, \pi)$. Define the sets

$$\mathcal{U}_{N,\epsilon} \;:=\; \{\alpha \in \mathrm{Co}(Y) : \mathrm{cov}(\mu, \alpha, P, F_n, \epsilon \,|\, \pi) > 2 \cdot U(|F_n|) \text{ for some } n > N\}$$

and let

$$\mathscr{C} \;:=\; \bigcap_{m \geq 100} \; \bigcap_{N \geq 1} \mathcal{U}_{N, \epsilon = 1/m}.$$

It is clear that every $\alpha \in \mathscr{C}$ satisfies

$$h^{U,(F_n)}_{\mathrm{slow}}(\mathbf{X}_\alpha \,|\, \pi) \;\geq\; h^{U,(F_n)}_{\mathrm{slow}}(\mathbf{X}_\alpha, P \,|\, \pi) \;\geq\; 2.$$

So, by the Baire category theorem, in order to prove Theorem 3.2 it suffices to show that each $\mathcal{U}_{N,\epsilon}$ is both open and dense in $\mathrm{Co}(Y)$.

We first establish the openness. Again, to emphasize the dependence on the cocycle α, we use the notation $P_\alpha^F(y,t)$ to denote the (P,F)-name of the point $(y,t) \in X$ under the action T_α. Let $\mathcal{D}$ denote the partition $\{[0,1/2),[1/2,1]\}$ of I.

Lemma 3.3. *Let $\beta^{(n)}$ be a sequence of cocycles converging to α and let F be a finite subset of G. Then we have*

$$\mu\left\{(y,t) : P_{\beta^{(n)}}^F(y,t) \;=\; P_\alpha^F(y,t)\right\} \;\to\; 1 \qquad \text{as } n \to \infty.$$

Proof. For the names $P_{\beta^{(n)}}^F(y,t)$ and $P_\alpha^F(y,t)$ to be the same means that for every $g \in F$,

$$P\left(S^g y, \beta_g^{(n)}(y)t\right) \;=\; P\left(S^g y, \alpha_g(y)t\right),$$

which is equivalent to

$$\mathcal{D}\left(\beta_g^{(n)}(y)t\right) \;=\; \mathcal{D}\left(\alpha_g(y)t\right). \tag{3.1}$$

The idea is the following. For fixed g and y, if $\alpha_g(y)$ and $\beta_g^{(n)}(y)$ are close in d_A, then (3.1) fails for only a small measure set of t. And if $\beta^{(n)}$ is very close to α in the cocycle topology, then $\beta_g^{(n)}(y)$ and $\alpha_g(y)$ are close for all $g \in F$ and most $y \in Y$. Then, by Fubini's theorem, we will get that the measure of the set of (y,t) failing (3.1) is small.

Here are the details. Fix $\rho > 0$; we will show that the measure of the desired set is at least $1-\rho$ for n sufficiently large. First, let σ be so small that for any $\varphi, \psi \in \mathrm{Aut}(I,m)$,

$$d_A(\varphi,\psi) < \sigma \quad \text{implies} \quad m\{t : \mathcal{D}(\varphi t) = \mathcal{D}(\psi t)\} > 1 - \rho/2.$$

This is possible because

$$\{t : \mathcal{D}(\varphi t) \neq \mathcal{D}(\psi t)\} \;\subseteq\; (\varphi^{-1}[0,1/2) \,\triangle\, \psi^{-1}[0,1/2)) \cup (\varphi^{-1}[1/2,1] \,\triangle\, \psi^{-1}[1/2,1]).$$

Then, from the definition of the cocycle topology, we have

$$\nu\left\{y \in Y : d_A\left(\beta_g^{(n)}(y), \alpha_g(y)\right) < \sigma \quad \text{for all } g \in F\right\} \ \to\ 1 \qquad \text{as } n \to \infty.$$

Let n be large enough so that the above is larger than $1 - \rho/2$. Then, by Fubini's theorem, we have

$$\begin{aligned} &\mu\left\{(y,t) : P^F_{\beta^{(n)}}(y,t) \ = \ P^F_\alpha(y,t)\right\} \\ &= \int m\left\{t : \mathcal{D}\left(\beta_g^{(n)}(y)t\right) = \mathcal{D}\left(\alpha_g(y)t\right) \quad \text{for all } g \in F\right\} d\nu(y). \end{aligned}$$

We have arranged things so that the integrand above is $> 1 - \rho/2$ on a set of y of ν-measure $> 1 - \rho/2$, so the integral is at least $(1-\rho/2)(1-\rho/2) > 1-\rho$ as desired. □

Lemma 3.4. *For any finite $F \subseteq G$, $\epsilon > 0$, and $L > 0$, the set*

$$\{\alpha \in \mathrm{Co}(Y) : \mathrm{cov}(\mu, \alpha, P, F, \epsilon \,|\, \pi) > L\}$$

is open in $\mathrm{Co}(Y)$.

Proof. For this proof, let us overload our notation slightly and consider P^F_α to be a map from X to $\{0,1\}^F$ that sends a point in X to its (P,F)-name according to T_α. Suppose $\beta^{(n)}$ is a sequence of cocycles converging to α and satisfying

$$\mathrm{cov}(\mu, \beta^{(n)}, P, F, \epsilon \,|\, \pi) \ \le\ L$$

for all n. We will show that $\mathrm{cov}(\mu, \alpha, P, F, \epsilon \,|\, \pi) \le L$ as well.

By Lemma 3.3, we have

$$\mu\left\{x : P^F_{\beta^{(n)}}(y,t) \ = \ P^F_\alpha(y,t)\right\} \ = \ \int \mu_y\left\{x : P^F_{\beta^{(n)}}(x) \ = \ P^F_\alpha(x)\right\} d\nu(y) \ \to\ 1$$

as $n \to \infty$. Therefore, there is a subsequence (n_j) that satisfies

$$\mu_y \left\{ t : P^F_{\beta^{(n_j)}}(x) \;=\; P^F_\alpha(x) \right\} \;\to\; 1$$

for ν-a.e. y. Let us immediately pass to this subsequence and relabel, so we assume that ν-a.e. y satisfies

$$\mu_y \left\{ x : P^F_{\beta^{(n)}}(x) \;=\; P^F_\alpha(x) \right\} \;\to\; 1 \qquad \text{as } n \to \infty. \tag{3.2}$$

Let Y^* denote the (full measure) set of y for which this convergence holds.

For each n, let $Y_n \subseteq Y$ be the set of y that satisfy $\mathrm{cov}(\mu_y, \beta^{(n)}, P, F, \epsilon) \leq L$. We have $\nu(Y_n) \geq 1 - \epsilon$ for each n, so it follows that

$$\overline{Y} \;:=\; Y^* \;\cap\; \bigcap_{N \geq 1} \bigcup_{n \geq N} Y_n$$

also satisfies $\nu\left(\overline{Y}\right) \geq 1 - \epsilon$. Therefore it suffices to show that $\mathrm{cov}(\mu_y, \alpha, P, F, \epsilon) \leq L$ for all $y \in \overline{Y}$.

Fix $y \in \overline{Y}$. By definition, this means that $\mathrm{cov}(\mu_y, \beta^{(n)}, P, F, \epsilon) \leq L$ for infinitely many n. By passing to a further subsequence (which depends on y, but y is now fixed for the remainder of the proof) and again relabeling, we may assume that this is true for all n. The covering number $\mathrm{cov}(\mu_y, \beta^{(n)}, P, F, \epsilon)$ is a quantity which really depends only on the measure

$$\left(P^F_{\beta^{(n)}}\right)_* \mu_y \;\in\; \mathrm{Prob}\left(\{0,1\}^F\right),$$

which we now call λ_n for short. The assumption that

$$\mathrm{cov}(\mu_y, \beta^{(n)}, P, F, \epsilon) \;\leq\; L \qquad \text{for all } n$$

says that for each n, there is a collection of L names $w^{(n)}_1, \ldots, w^{(n)}_L \in \{0,1\}^F$ such that the Hamming balls of radius ϵ centered at these names cover a set of λ_n-measure at least $1-\epsilon$. Since $\{0,1\}^F$ is a finite set, there are only finitely many possibilities for the collection $(w_1^{(n)},\ldots,w_L^{(n)})$. Therefore by passing to a further subsequence and relabeling again we may assume that there is a fixed collection of words $w_1,\ldots,w_L$ with the property that if we let B_i be the Hamming ball of radius ϵ centered at w_i, then $\lambda_n\left(\bigcup_{i=1}^L B_i\right)\geq 1-\epsilon$ for every n.

Now, because we chose $y\in Y^*$, we know by (3.2) that

$$\mu_y\left\{x : P^F_{\beta^{(n_j)}}(x) \;=\; P^F_\alpha(x)\right\} \;\to\; 1.$$

This implies that the sequence of measures λ_n converges in the total variation norm on $\mathrm{Prob}\left(\{0,1\}^F\right)$ to $\lambda := \left(P^F_\alpha\right)_*\mu_y$. Since $\lambda_n\left(\bigcup_{i=1}^L B_i\right)\geq 1-\epsilon$ for every n, we conclude that $\lambda\left(\bigcup_{i=1}^L B_i\right)\geq 1-\epsilon$ also, which implies that $\mathrm{cov}(\mu_y,\alpha,P,F,\epsilon)\leq L$ as desired. □

Finally, because

$$\mathcal{U}_{N,\epsilon} \;=\; \bigcup_{n>N}\{\alpha\in\mathrm{Co}(Y) : \mathrm{cov}(\mu,\alpha,P,F_n,\epsilon\,|\,\pi) > 2\cdot U(|F_n|)\},$$

we conclude that each $\mathcal{U}_{N,\epsilon}$ is open as claimed. So to complete the proof of Theorem 3.2, we only need to establish the density.

Proposition 3.5. *For each $N\geq 1$ and $0<\epsilon<1/100$, $\mathcal{U}_{N,\epsilon}$ is dense in* $\mathrm{Co}(Y)$.

This result is the main technical result of this chapter, and the remainder of this section is devoted to proving it.

3.2.2 Setup for proof of Proposition 3.5

Let $N\geq 1$ and $0<\epsilon<1/100$ be fixed and let α_0 be an arbitrary cocycle. Consider a neighborhood of α_0 determined by a finite set $F\subseteq G$ and $\eta>0$. We can assume without loss of generality that $\eta \ll \epsilon$. We will produce a new cocycle $\alpha \in \mathcal{U}_{N,\epsilon}$ such that there is a set Y' of measure $\geq 1-\eta$ on which $\alpha_f(y) = (\alpha_0)_f(y)$ for all $f \in F$, implying that α is in the (F,η)-neighborhood of α_0. The construction of such an α is based on the fact that the orbit equivalence relation R is hyperfinite.

Theorem 3.6 ([CFW81, Theorem 10])**.** *There is an increasing sequence of equivalence relations $R_n \subseteq Y \times Y$ such that*

- *each R_n is measurable as a subset of $Y \times Y$,*
- *every cell of every R_n is finite, and*
- *$\bigcup_n R_n$ agrees μ-a.e. with R.*

Fix such a sequence (R_n) and for $y \in Y$, write $R_n(y)$ to denote the cell of R_n that contains y.

Lemma 3.7. *There exists an m_1 such that $\nu\{y \in Y : S^F y \subseteq R_{m_1}(y)\} > 1-\eta$.*

Proof. Almost every y satisfies $S^G y = \bigcup_m R_m(y)$, so in particular, for ν-a.e. y, there is an m_y such that $S^F y \subseteq R_m(y)$ for all $m \geq m_y$. Letting $Y_\ell = \{y \in Y : m_y \leq \ell\}$, we see that the sets Y_ℓ are increasing and exhaust almost all of Y. Therefore we can pick m_1 so that $\nu(Y_{m_1}) > 1-\eta$. □

Now we drop $R_1, \ldots, R_{m_1-1}$ from the sequence and assume that $m_1 = 1$.

Lemma 3.8. *There exists a K such that $\nu\{y : |R_1(y)| \leq K\} > 1-\eta$.*

Proof. Every R_1-cell is finite, so if we define $Y_k = \{y \in Y : |R_1(y)| \leq k\}$, then the Y_k are increasing and exhaust all of Y. So we pick K so that $\nu(Y_K) > 1-\eta$. □

Continue to use the notation $Y_K = \{y \in Y : |R_1(y)| \leq K\}$.

Lemma 3.9. *For all n sufficiently large, $\nu\{y \in Y : \frac{|(S^{F_n}y) \cap Y_K|}{|F_n|} > 1 - 2\eta\} > 1 - \eta$.*

Proof. We have $|(S^{F_n}y) \cap Y_K| = \sum_{f \in F_n} 1_{Y_K}(S^f y)$. By the mean ergodic theorem [Gla03, Theorem 3.33], we get

$$\frac{|(T^{F_n}y) \cap Y_K|}{|F_n|} \;\to\; \nu(Y_K) \;>\; 1 - \eta \qquad \text{in probability as } n \to \infty.$$

Therefore, in particular, $\nu\left\{y \in Y : \frac{|(T^{F_n}y) \cap Y_K|}{|F_n|} > 1 - 2\eta\right\} \to 1$ as $n \to \infty$, so this measure is $> 1 - \eta$ for all n sufficiently large. $\square$

From now on, let n be a fixed number that is large enough so that the above lemma holds, $n > N$, and $\frac{1}{2}\exp\left(\frac{1}{8K^2} \cdot |F_n|\right) > 2 \cdot U(|F_n|)$. This is possible because U is sub-exponential. The relevance of the final condition will appear at the end.

Lemma 3.10. *There is an m_2 such that $\nu\{y \in Y : T^{F_n}y \subseteq R_{m_2}(y)\} > 1 - \eta$.*

Proof. Same proof as Lemma 3.7. $\square$

Again, drop $R_2, \ldots, R_{m_2 - 1}$ from the sequence of equivalence relations and assume $m_2 = 2$.

3.2.3 Construction of the perturbed cocycle

Let (R_n) be the relabeled sequence of equivalence relations from the previous section. The following measure theoretic fact is well known. Recall that two partitions P and P' of I are said to be **independent** with respect to m if $m(E \cap E') = m(E)m(E')$ for any $E \in P$, $E' \in P'$.

Lemma 3.11. *Let P and P' be two finite partitions of I. Then there exists a $\varphi \in \mathrm{Aut}(I, m)$ such that P and $\varphi^{-1}P'$ are independent with respect to m.*

Proposition 3.12. *For any $\alpha_0 \in \mathrm{Co}(Y)$, there is an $\alpha \in \mathrm{Co}(Y)$ such that*

(1) $\alpha_g(y) = (\alpha_0)_g(y)$ whenever $(y, T^g y) \in R_1$, and

(2) for ν-a.e. y, the following holds. If C is an R_1-cell contained in $R_2(y)$, consider the map $Z_C : t \mapsto P_\alpha^{\{g:T^g y \in C\}}(y, t)$ as a random variable on the underlying space (I, m). Then as C ranges over all such R_1-cells, the random variables Z_C are independent.

Proof. The proof we give here is almost complete, but we do not verify the measurability of the α that we construct. We leave that detail to Section 3.2.5. It is more convenient to adopt the perspective of a cocycle as a map $\alpha : R \to \mathrm{Aut}(I, m)$ satisfying the condition $\alpha(y_1, y_3) = \alpha(y_2, y_3) \circ \alpha(y_1, y_2)$.

Step 1. For $(y_1, y_2) \in R_1$, let $\alpha(y_1, y_2) = \alpha_0(y_1, y_2)$.

Step 2. Fix an R_2-cell $\overline{C}$. Enumerate by $\{C_1, \ldots, C_k\}$ all of the R_1-cells contained in $\overline{C}$ and choose from each a representative $y_i \in C_i$.

Step 3. Recall that $\mathcal{D}$ denotes the partition $\{[0, 1/2), [1/2, 1]\}$ of I. Define $\alpha(y_1, y_2)$ to be an element of $\mathrm{Aut}(I, m)$ such that

$$\bigvee_{y' \in C_1} \alpha(y_1, y')^{-1}\mathcal{D} \qquad \text{and} \qquad \alpha(y_1, y_2)^{-1} \left(\bigvee_{y' \in C_2} \alpha(y_2, y')^{-1}\mathcal{D} \right)$$

are independent. These expressions are well defined because α has already been defined on R_1 and we use Lemma 3.11 to guarantee that such an element of $\mathrm{Aut}(I, m)$ exists.

Step 4. There is now a unique way to extend the definition of α to $(C_1 \cup C_2) \times (C_1 \cup C_2)$ that is consistent with the cocycle condititon. For arbitrary $z_1 \in C_1, z_2 \in C_2$, define

$$\begin{aligned} \alpha(z_1, z_2) &= \alpha(y_2, z_2) \circ \alpha(y_1, y_2) \circ \alpha(z_1, y_1) \qquad \text{and} \\ \alpha(z_2, z_1) &= \alpha(z_1, z_2)^{-1}. \end{aligned}$$

The middle term in the first equation was defined in the previous step and the outer two terms were defined in step 1.

Step 5. Extend the definition of α to the rest of the C_i inductively, making each cell independent of all the previous ones. Suppose α has been defined on $(C_1 \cup \cdots \cup C_j) \times (C_1 \cup \cdots \cup C_j)$. Using Lemma 3.11 again, define $\alpha(y_1, y_{j+1})$ to be an element of $\text{Aut}(I, m)$ such that

$$\bigvee_{y' \in C_1 \cup \cdots \cup C_j} \alpha(y_1, y')^{-1} \mathcal{D} \qquad \text{and} \qquad \alpha(y_1, y_{j+1})^{-1} \left(\bigvee_{y' \in C_{j+1}} \alpha(y_{j+1}, y')^{-1} \mathcal{D} \right)$$

are independent. Then, just as in step 4, there is a unique way to extend the definition of α to all of $(C_1 \cup \cdots \cup C_{j+1}) \times (C_1 \cup \cdots \cup C_{j+1})$. At the end of this process, α has been defined on all of $\overline{C} \times \overline{C}$. This was done for an arbitrary R_2-cell $\overline{C}$, so now α is defined on R_2.

Step 6. For each $N \geq 2$, extend the definition of α from R_N to R_{N+1} with the same procedure, but there is no need to set up any independence. Instead, every time there is a choice for how to define α between two of the cell representatives, just take it to be the identity. This defines α on $\bigcup_{N \geq 1} R_N$, which is equal mod ν to the full orbit equivalence relation, so α is a well defined cocycle.

Now we verify the two claimed properties of α. Property (1) is immediate from step 1 of the construction. To check property (2), fix y and let C_j be any of the

R_1-cells contained in $R_2(y)$. Note that the name $P_\alpha^{\{g:S^g y\in C_j\}}(y,t)$ records the data $P(T_\alpha^g(y,t)) = P(S^g y, \alpha_g(y)t) = \mathcal{D}(\alpha_g(y)t)$ for all g such that $S^g y \in C_j$, which, by switching to the other notation is the same data as $\mathcal{D}(\alpha(y,z)t)$ for $z \in C_j$. So, the set of t for which $P_\alpha^{\{g:S^g y\in C_j\}}(y,t)$ takes a particular value is given by a corresponding particular cell of the partition $\bigvee_{y'\in C_j} \alpha(y,y')^{-1}\mathcal{D} = \alpha(y,y_1)^{-1}\left(\bigvee_{y'\in C_j} \alpha(y_1,y')^{-1}\mathcal{D}\right)$. The construction of α was defined exactly so that the partitions $\bigvee_{y'\in C_j} \alpha(y_1,y')^{-1}\mathcal{D}$ are all independent and the names $P_\alpha^{\{g:S^g y\in C_j\}}(y,t)$ are determined by these independent partitions pulled back by the fixed m-preserving map $\alpha(y,y_1)$, so they are also independent. □

Letting $\widetilde{Y} = \{y \in Y : S^F y \subseteq R_1(y)\}$, this construction guarantees that $\alpha_f(y) = (\alpha_0)_f(y)$ for all $f \in F, y \in \widetilde{Y}$. By Lemma 3.7, $\nu(\widetilde{Y}) > 1-\eta$, so this shows that α is in the $(F,\eta)-$neighborhood of α_0.

3.2.4 Estimating the relative covering number

Let α be the cocycle constructed in the previous section. Now, for the rest of this section, whenever we talk about about (P,F)-names, we will always mean with respect to the action T_α determined by this fixed α. We will obtain a lower bound for $\operatorname{cov}(\mu,\alpha,P,F_n,\epsilon\,|\,\pi)$ by showing that any set of small d_{P,F_n}-diameter must have small μ_y-measure for most $y \in Y$. The following formulation of Hoeffding's inequality will be quite useful [Ver18, Theorem 2.2.6].

Theorem 3.13. *Let $Z_1,\ldots,Z_\ell$ be independent random variables such that each $Z_i \in [0,K]$ almost surely. Let $a = \mathbb{E}\left[\sum Z_i\right]$. Then for any $u > 0$,*

$$\mathbb{P}\left(\sum_{i=1}^{\ell} Z_i < a-u\right) \;\leq\; \exp\left(-\frac{2u^2}{K^2\ell}\right).$$

Let $Y_0 = \left\{ y \in Y : \frac{|(S^{F_n}y) \cap Y_K|}{|F_n|} > 1 - 2\eta \text{ and } S^{F_n}y \subseteq R_2(y) \right\}$. By Lemmas 3.9 and 3.10, $\nu(Y_0) > 1 - 2\eta$. Also note that because $\mu = \nu \times m$, the disintegration of μ over π is given explicitly by $\mu_y = \delta_y \times m$.

Proposition 3.14. *Let $y \in Y_0$. If $B \subseteq X$ is any set satisfying* $\operatorname{diam}_{P,F_n}(B) \leq \epsilon$*, then*

$$\mu_y(B) \;\leq\; \exp\left(-\frac{1}{8K^2} \cdot |F_n| \right).$$

Proof. Because we only care about the μ_y-measure of B, we may assume without loss of generality that B is contained in the fiber above y. Furthermore, because any set of diameter at most ϵ is contained in a ball of radius ϵ, it suffices to assume that

$$B \;=\; \{y\} \times \{t \in I : d_{P,F_n}((y,t_0),(y,t)) \leq \epsilon\}$$

for some fixed $t_0 \in I$. Therefore,

$$\mu_y(B) \;=\; m\,\{t \in I : d_{P,F_n}((y,t_0),(y,t)) \leq \epsilon\}\,. \tag{3.3}$$

Let $\mathcal{C}$ be the collection of R_1-cells C that meet $S^{F_n}y$ and satisfy $|C| \leq K$. For each $C \in \mathcal{C}$, let $F_C = \{f \in F_n : S^f y \in C\}$. Define

$$Z(t) \;=\; |F_n| \cdot d_{P,F_n}\left((y,t_0),(y,t)\right) \;=\; \sum_{f \in F_n} 1_{P(T^f_\alpha(y,t_0)) \neq P(T^f_\alpha(y,t))},$$

and for each $C \in \mathcal{C}$, define

$$Z_C(t') \;=\; \sum_{f \in F_C} 1_{P(T^f_\alpha(y,t_0)) \neq P(T^f_\alpha(y,t))}.$$

Then we have

$$Z(t) \;\geq\; \sum_{C \in \mathcal{C}} Z_C(t),$$

so to get an upper bound for $\mu_y(B)$, it is sufficient to control

$$m\left\{t \in I : \sum_{C\in\mathcal{C}} Z_C(t) < \epsilon|F_n|\right\}.$$

View each $Z_C(t)$ as a random variable on the underlying probability space (I, m). Our construction of the cocycle α guarantees that the collection of names $P_\alpha^{F_C}(y,t)$ as C ranges over all of the R_1-cells contained in $R_2(y)$ is an *independent* collection. Therefore, in particular, the Z_C for $C \in \mathcal{C}$ are independent (the assumption that $y \in Y_0$ guarantees that all $C \in \mathcal{C}$ are contained in $R_2(y)$). We also have that each $Z_C \in [0, K]$ and the expectation of the sum is

$$\begin{aligned} a \;:=\; \sum_{C\in\mathcal{C}} \int Z_C(t)\,dm(t) \;&=\; \sum_{C\in\mathcal{C}}\sum_{f\in F_C} \int 1_{P(T_\alpha^f(y,t_0))\neq P(T_\alpha^f(y,t))}\,dm(t) \;=\; \sum_{C\in\mathcal{C}} \frac{1}{2}|F_C| \\ &=\; \frac{1}{2}\sum_{C\in\mathcal{C}} |C\cap(S^{F_n}y)| \;>\; \frac{1}{2}(1-2\eta)|F_n|, \end{aligned}$$

where the final inequality is true because $y \in Y_0$. So, we can apply Theorem 3.13 with $u = a - \epsilon|F_n|$ to conclude

$$\begin{aligned} m\left\{t : \sum_{C\in\mathcal{C}} Z_C(t) < \epsilon|F_n|\right\} \;&\le\; \exp\left(\frac{-2u^2}{K^2|\mathcal{C}|}\right) \;\le\; \exp\left(\frac{-2(1/2-\eta-\epsilon)^2|F_n|^2}{K^2|F_n|}\right) \\ &\le\; \exp\left(-\frac{1}{8K^2}\cdot|F_n|\right). \end{aligned}$$

The final inequality holds because $\epsilon < 1/100$ and $\eta \ll \epsilon$ is small enough so that $1/2 - \eta - \epsilon > 1/4$. □

Corollary 3.15. *We have* $\mathrm{cov}(\mu, \alpha, P, F_n, \epsilon \,|\, \pi) \ge \frac{1}{2}\exp\left(\frac{1}{8K^2}\cdot|F_n|\right)$.

Proof. For any $y \in Y_0$, Proposition 3.14 implies that if $\mathrm{diam}_{P,F_n}(B) \le \epsilon$, then

$$\mu_y(B) \;\le\; \exp\left(-\frac{1}{8K^2}\cdot|F_n|\right).$$

Therefore, for any such y, it must require at least

$$\frac{1-\epsilon}{\exp\left(-\frac{1}{8K^2}\cdot|F_n|\right)} \;\geq\; \frac{1}{2}\exp\left(\frac{1}{8K^2}\cdot|F_n|\right)$$

many sets of diameter less than ϵ to cover at least $1-\epsilon$ of μ_y, implying that

$$\mathrm{cov}(\mu_y, P, F_n, \epsilon) \;\geq\; \frac{1}{2}\exp\left(\frac{1}{8K^2}\cdot|F_n|\right).$$

Finally, because this holds for every $y \in Y_0$ and $\nu(Y_0) \geq 1-2\eta > 1-\epsilon > 1/2$, it follows that

$$\mathrm{cov}(\mu, \alpha, P, F_n, \epsilon \,|\, \pi) \;\geq\; \frac{1}{2}\exp\left(\frac{1}{8K^2}\cdot|F_n|\right)$$

as desired. □

Finally, recall that n was chosen fixed so that $n > N$ and

$$\frac{1}{2}\exp\left(\frac{1}{8K^2}\cdot|F_n|\right) \;>\; 2\cdot U(|F_n|).$$

Therefore, the above result shows that $\alpha \in \mathcal{U}_{N,\epsilon}$. Because α was constructed to lie arbitrarily close to any cocycle $\alpha_0 \in \mathrm{Co}(Y)$, this shows that $\mathcal{U}_{N,\epsilon}$ is dense in $\mathrm{Co}(Y)$. This completes the proof of Proposition 3.5, which in turn completes the proof of Theorem 3.2.

3.2.5 Measurability of the perturbed cocycle

In this section, we give a more careful proof of Proposition 3.12 that addresses the issue of measurability. We will need to use at some point the following measurable selector theorem [Fre06, Proposition 433F].

Theorem 3.16. *Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be standard Borel spaces. Let $\mathbb{P}$ be a probability measure on $(\Omega_1, \mathcal{F}_1)$ and suppose that $f : \Omega_2 \to \Omega_1$ is measurable and* *surjective. Then there exists a measurable selector $g : \Omega_1 \to \Omega_2$ which is defined $\mathbb{P}$-a.e. (meaning $g(\omega) \in f^{-1}(\omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega_1$).*

Given $y \in Y$, there is a natural bijection between $S^G y$ and G because S is a free action. We can also identify subsets – if $E \subseteq S^G y$, then we will write $\widetilde{E} := \{g \in G : S^g y \in E\}$. Note that this set depends on the "base point" y. If y and z are two points in the same G-orbit, then the set $\widetilde{E}$ based at y is a translate of the same set based at z. It will always be clear from context what the intended base point is.

Definition 3.17. A **pattern** in G is a pair $(H, \mathscr{P})$, where H is a finite subset of G and $\mathscr{P}$ is a partition of H.

Definition 3.18. For $y \in Y$, define $\text{pat}_n(y)$ to be the pattern $(H, \mathscr{P})$, where $H = \widetilde{R_n(y)}$ and $\mathscr{P}$ is the partition of H into the sets $\widetilde{C}$ where C ranges over all of the R_{n-1}-cells contained in $R_n(y)$.

Lemma 3.19. $\text{pat}_n(y)$ *is a measurable function of* y.

Proof. Because there are only countably many possible patterns, it is enough to fix a pattern $(H, \mathscr{P})$ and show that $\{y : \text{pat}_n(y) = (H, \mathscr{P})\}$ is measurable. Enumerate $\mathscr{P} = \{C_1, \ldots, C_k\}$. Saying that $\text{pat}_n(y) = (H, \mathscr{P})$ is the same as saying that $S^H y = R_n(y)$ and each $S^{C_i} y$ is a cell of R_{n-1}. We can express the set of y satisfying this as

$$\left(\bigcap_{i=1}^{k} \bigcap_{g,h \in C_i} \{y : (S^g y, S^h y) \in R_{n-1}\} \quad \cap \quad \bigcap_{(g,h) \in G^2 \setminus \bigcup (C_i \times C_i)} \{y : (S^g y, S^h y) \notin R_{n-1}\} \right)$$
$$\cap \quad \left(\bigcap_{g \in H} \{y : (y, S^g y) \in R_n\} \quad \cap \quad \bigcap_{g \notin H} \{y : (y, S^g y) \notin R_n\} \right).$$

Because each R_n is a measurable set and each S^g is a measurable map, this whole thing is measurable. □

For each pattern $(H, \mathscr{P})$, let $Y^{(n)}_{H,\mathscr{P}} = \{y \in Y : \mathrm{pat}_n(y) = (H, \mathscr{P})\}$. We will define our cocycle α inductively on the equivalence relations R_n. For each n, the sets $Y^{(n)}_{H,\mathscr{P}}$ partition Y into countably many measurable sets, so it will be enough to define α measurably on each $Y^{(n)}_{H,\mathscr{P}}$. At this point, fix a pattern $(H, \mathscr{P})$, fix $n = 2$, and write $Y_{H,\mathscr{P}}$ instead of $Y^{(2)}_{H,\mathscr{P}}$. Define

$$\begin{aligned}
\Omega_2^{H,\mathscr{P}} &= \Big\{\psi : H \times H \to \mathrm{Aut}(I, m) : \psi(h_1, h_3) = \psi(h_2, h_3) \circ \psi(h_1, h_2) \\
&\qquad\qquad \text{for all } h_1, h_2, h_3 \in H\Big\}, \\
\Omega_1^{H,\mathscr{P}} &= \Big\{\sigma : \bigcup_{C \in \mathscr{P}} C \times C \to \mathrm{Aut}(I, m) : \sigma(g_1, g_3) = \sigma(g_2, g_3) \circ \sigma(g_1, g_2) \\
&\qquad\qquad \text{whenever } g_1, g_2, g_3 \text{ all lie in the same cell } C\Big\}, \\
\Omega_2^{H,\mathscr{P},\mathrm{ind}} &= \Big\{\psi \in \Omega_2^{H,\mathscr{P}} : \psi \text{ is } (H, \mathscr{P})\text{-independent}\Big\},
\end{aligned}$$

where $\psi \in \Omega_2^{H,\mathscr{P}}$ is said to be **$(\mathbf{H}, \mathscr{P})$-independent** if for any fixed $h_0 \in H$, the partitions

$$\bigvee_{h \in C} \psi(h_0, h)^{-1} \mathcal{D}$$

as C ranges over $\mathscr{P}$ are independent with respect to m.

Proposition 3.20. *For every $\sigma \in \Omega_1^{H,\mathscr{P}}$, there is some $\psi \in \Omega_2^{H,\mathscr{P},ind}$ that extends σ.*

Proof. The idea is exactly the same as the construction described in steps 3-5 in the sketched proof of Proposition 3.12, but we write it out here also for completeness.

Enumerate $\mathscr{P} = \{C_1, \ldots, C_k\}$ and for each i fix an element $g_i \in C_i$. First, obviously we will define $\psi = \sigma$ on each $C_i \times C_i$. Next, define $\psi(g_1, g_2)$ to be an element of $\mathrm{Aut}(I, m)$ such that

$$\bigvee_{g \in C_1} \sigma(g_1, g)^{-1} \mathcal{D} \qquad \text{and} \qquad \psi(g_1, g_2)^{-1} \left(\bigvee_{g \in C_2} \sigma(g_2, g)^{-1} \mathcal{D} \right)$$

are independent. Then, define ψ on all of $(C_1 \cup C_2) \times (C_1 \cup C_2)$ by setting

$$\begin{aligned} \psi(h_1, h_2) &= \sigma(g_2, h_2) \circ \psi(g_1, g_2) \circ \sigma(h_1, g_1) \qquad \text{and} \\ \psi(h_2, h_1) &= \psi(h_1, h_2)^{-1} \end{aligned}$$

for any $h_1 \in C_1, h_2 \in C_2$. Continue this definition inductively, making each new step independent of all the steps that came before it. If ψ has been defined on $(C_1 \cup \cdots \cup C_j) \times (C_1 \cup \cdots \cup C_j)$, then define $\psi(g_1, g_{j+1})$ to be an element of $\mathrm{Aut}(I, m)$ such that

$$\bigvee_{g \in C_1 \cup \cdots \cup C_j} \psi(g_1, g)^{-1}\mathcal{D} \qquad \text{and} \qquad \psi(g_1, g_{j+1})^{-1} \left(\bigvee_{g' \in C_{j+1}} \sigma(g_{j+1}, g')^{-1}\mathcal{D} \right)$$

are independent. Then extend the definition of ψ to all of $(C_1 \cup \cdots \cup C_{j+1}) \times (C_1 \cup \cdots \cup C_{j+1})$ in the exact same way.

At the end of this process, ψ has been defined on $(C_1 \cup \cdots \cup C_k) \times (C_1 \cup \cdots \cup C_k) = H \times H$, and it satisfies the cocycle condition by construction. To verify that it also satisfies the independence condition, notice that the construction has guaranteed that

$$\bigvee_{h \in C} \psi(g_1, h)^{-1}\mathcal{D}$$

are independent partitions as C ranges over $\mathscr{P}$. To get the same conclusion for an arbitrary base point h_0, pull everything back by the fixed map $\psi(h_0, g_1)$. Because this map is measure preserving, pulling back all of the partitions by it preserves their independence. □

Now we would like to take this information about cocycles defined on patterns and use it to produce cocycles defined on the actual space Y. Define the map $\sigma^{H,\mathscr{P}}$:

$Y_{H,\mathscr{P}} \to \Omega_1^{H,\mathscr{P}}$ by $\sigma_y^{H,\mathscr{P}}(g_1, g_2) := \alpha_0(S^{g_1}y, S^{g_2}y)$. Note that this is a measurable map because α_0 is a measurable cocycle.

By Theorem 3.16 applied to the measure

$$\mathbb{P} \ := \ (\sigma^{H,\mathscr{P}})_*(\nu(\cdot \,|\, Y_{H,\mathscr{P}})) \in \mathrm{Prob}(\Omega_1^{H,\mathscr{P}}),$$

we get a *measurable* map $E^{H,\mathscr{P}} : \Omega_1^{H,\mathscr{P}} \to \Omega_2^{H,\mathscr{P},\mathrm{ind}}$ defined $\mathbb{P}$-a.e. such that $E^{H,\mathscr{P}}(\sigma)$ extends σ. Denote the composition $E^{H,\mathscr{P}} \circ \sigma^{H,\mathscr{P}}$ by $\psi^{H,\mathscr{P}}$ and write the image of y under this map as $\psi_y^{H,\mathscr{P}}$. To summarize, for every pattern $(H, \mathscr{P})$, there is a measurable map $\psi^{H,\mathscr{P}} : Y_{H,\mathscr{P}} \to \Omega_2^{H,\mathscr{P},\mathrm{ind}}$ defined ν-a.e. with the property that $\psi_y^{H,\mathscr{P}}$ extends $\sigma_y^{H,\mathscr{P}}$.

It is now natural to define our desired cocycle α on the equivalence relation R_2 by the formula $\alpha(y, S^g y) := \psi_y^{\mathrm{pat}_2(y)}(e, g)$. It is then immediate to verify the two properties of α claimed in the statement of Proposition 3.12. The fact that α agrees with α_0 on R_1 follows from the fact that $\psi_{H,\mathscr{P}}$ extends $\sigma^{H,\mathscr{P}}$ and the claimed independence property of α translates directly from the independence property that the $\psi_y^{H,\mathscr{P}}$ were constructed to have (see also the discussion after step 6 in the sketched proof of Proposition 3.12). Also, α is measurable because for each fixed g, the map $y \mapsto \alpha(y, T^g y)$ is simply a composition of other maps already determined to be measurable. The only problem is that α, when defined in this way, need not satisfy the cocycle condition. To see why, observe that the cocycle condition $\alpha(y, S^h y) = \alpha(S^g y, S^h y) \circ \alpha(y, S^g y)$ is equivalent to the condition

$$\psi_y^{\mathrm{pat}_2(y)}(e, h) \ = \ \psi_{S^g y}^{\mathrm{pat}_2(S^g y)}(e, hg^{-1}) \circ \psi_y^{\mathrm{pat}_2(y)}(e, g). \tag{3.4}$$

But in defining the maps $\psi^{H,\mathscr{P}}$, we have simply applied Theorem 3.16 arbitrarily to each pattern separately, so $\psi^{\mathrm{pat}_2(y)}$ and $\psi^{\mathrm{pat}_2(S^g y)}$ have nothing to do with each other.

However, we can fix this problem with a little extra work, and once we do, we will have defined $\alpha : R_2 \to \mathrm{Aut}(I, m)$ with all of the desired properties.

Start by declaring two patterns equivalent if they are translates of each other, and fix a choice of one pattern from each equivalence class. Since there are only countably many patterns in total, there is no need to worry about how to make this choice. For each representative pattern $(H_0, \mathscr{P}_0)$, apply Theorem 3.16 arbitrarily to get a map $\psi^{H_0,\mathscr{P}_0}$. This does not cause any problems because two patterns that are not translates of each other can not appear in the same orbit (this follows from the easy fact that $\mathrm{pat}_2(S^g y) = g^{-1} \cdot \mathrm{pat}_2(y)$), so it doesn't matter that their ψ maps are not coordinated with each other. For convenience, let us denote the representative of the equivalence class of $\mathrm{pat}_2(y)$ by $\mathrm{rp}(y)$. Now for every $y \in Y$, let $g^*(y)$ be the unique element of G with the property that $\mathrm{pat}_2(S^{g^*(y)} y) = \mathrm{rp}(y)$. Notice that the maps g^* and rp are both constant on each subset $Y_{H,\mathscr{P}}$ and are therefore measurable.

Now for an arbitrary pattern $(H, \mathscr{P})$ and $y \in Y_{H,\mathscr{P}}$, we define the map $\psi^{H,\mathscr{P}}$ by

$$\psi_y^{H,\mathscr{P}}(g, h) \;:=\; \psi^{\mathrm{rp}(y)}_{S^{g^*(y)}y}(g \cdot g^*(y)^{-1}, h \cdot g^*(y)^{-1}).$$

Notice that this is still just a composition of measurable functions, so $\psi^{H,\mathscr{P}}$ is measurable. All that remains is to verify that this definition satisfies (3.4). The right hand side of (3.4) is

$$\begin{aligned}
&\psi^{\mathrm{rp}(S^g y)}_{S^{g^*(S^g y)} S^g y}(eg^*(S^g y)^{-1}, hg^{-1} g^*(S^g y)^{-1}) \;\circ\; \psi^{\mathrm{rp}(y)}_{S^{g^*(y)}y}(eg^*(y)^{-1}, gg^*(y)^{-1}) \\
&= \psi^{\mathrm{rp}(y)}_{S^{g^*(y)g^{-1}} S^g y}((g^*(y)g^{-1})^{-1}, hg^{-1}(g^*(y)g^{-1})^{-1}) \;\circ\; \psi^{\mathrm{rp}(y)}_{S^{g^*(y)}y}(g^*(y)^{-1}, gg^*(y)^{-1}) \\
&= \psi^{\mathrm{rp}(y)}_{S^{g^*(y)}y}(gg^*(y)^{-1}, hg^*(y)^{-1}) \;\circ\; \psi^{\mathrm{rp}(y)}_{S^{g^*(y)}y}(g^*(y)^{-1}, gg^*(y)^{-1}) \\
&= \psi^{\mathrm{rp}(y)}_{S^{g^*(y)}y}(g^*(y)^{-1}, hg^*(y)^{-1}),
\end{aligned}$$

which is by definition equal to the left hand side of (3.4) as desired.

This, together with the discussion surrounding (3.4), shows that if we construct the maps $\psi^{H,\mathscr{P}}$ in this way, then making the definition $\alpha(y, S^g y) = \psi_y^{\mathrm{pat}_2(y)}(e, g)$ gives us a true measurable cocycle with all of the desired properties. Finally, to extend the definition of α to R_n with $n \geq 3$, repeat the exact same process, except it is even easier because there is no need to force any independence. The maps $\psi^{H,\mathscr{P}}$ only need to be measurable selections into the space $\Omega_2^{H,\mathscr{P}}$, and then everything else proceeds in exactly the same way.

3.3 Completing the proof of Theorem 3.1

In this section, we deduce non-dominance from the fact that generic extensions have arbitrarily large relative slow entropy.

3.3.1 Controlling slow entropy of the base system

The first step is to find a fixed sub-exponential rate function that controls the slow entropy of the base system.

Lemma 3.21. *Let $b(m, n) \geq 0$ be real numbers satisfying*

- $\lim_{n\to\infty} b(m, n) = 0$ *for each fixed m, and*
- $b(m + 1, n) \geq b(m, n)$ *for all m, n.*

Then there exists a sequence (a_n) such that $a_n \to 0$ and for each fixed m, $b(m, n) \leq a_n$ for n sufficiently large (depending on m).

Proof. For each m, let N_m be such that $b(m, n) < 1/m$ for all $n > N_m$. Without loss of generality, we may assume that $N_m < N_{m+1}$. Then we define the sequence (a_n)

by $a_n = b(1,n)$ for $n \leq N_2$ and $a_n = b(m,n)$ for $N_m < n \leq N_{m+1}$. We have $a_n \to 0$ because $a_n < 1/m$ for all $n > N_m$. Finally, the fact that $b(m+1,n) \geq b(m,n)$ implies that for every fixed m, $a_n \geq b(m,n)$ as soon as $n > N_m$. □

Proposition 3.22. *Suppose* $\mathbf{Y}$ *has zero Kolmogorov–Sinai entropy and let* (F_n) *be a Følner sequence for* G*. Then there exists a sub-exponential rate function* U *such that*

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{Y}) \;=\; 0.$$

As was the case with Theorem 3.2, it is sufficient to prove this result with "$= 0$" replaced by "≤ 1".

Proof. By Krieger's theorem, let P be a generating partition for $\mathbf{Y}$. Because $\mathbf{Y}$ has zero Kolmogorov–Sinai entropy, we have

$$\lim_{n\to\infty} \frac{1}{|F_n|} \log \operatorname{cov}(\nu, P, F_n, \epsilon) \;=\; 0$$

for each $\epsilon > 0$. Let $b(m,n) = |F_n|^{-1} \log \operatorname{cov}(\nu, P, F_n, \epsilon = 1/m)$. Then these numbers satisfy the hypotheses of Lemma 3.21, so let (a_n) be the sequence guaranteed by that lemma.

Finally, define U by first setting $U(|F_n|) = \exp(|F_n| \cdot a_n)$ for each n and then interpolating by letting $U(k) = U(|F_n|)$ for all $|F_n| \leq k < |F_{n+1}|$. Because $a_n \to 0$, it follows that U is sub-exponential. Also, since $b(m,n)$ is eventually bounded by a_n for each fixed m, we have

$$\limsup_{n\to\infty} \frac{\operatorname{cov}(\nu, P, F_n, \epsilon = 1/m)}{U(|F_n|)} \;=\; \limsup_{n\to\infty} \frac{\exp(|F_n| \cdot b(m,n))}{\exp(|F_n| \cdot a_n)} \;\leq\; 1$$

for all m, which implies that

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{Y}) \;=\; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{Y}, P) \;\leq\; 1$$

as desired. $\square$

3.3.2 Conclusion of the proof

Let $\mathbf{Y}$ be a fixed base system with zero Kolmogorov–Sinai entropy. By Proposition 3.22, let U be a sub-exponential rate function such that $h_{\text{slow}}^{U,(F_n)}(\mathbf{Y}) = 0$. Now by Theorem 3.2, there is a dense G_δ set $\mathscr{C} \subseteq \mathrm{Co}(Y)$ such that $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}_\alpha \,|\, \pi) = \infty$ for all $\alpha \in \mathscr{C}$. Finally, Theorem 2.6 implies that $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}_\alpha) \geq h_{\text{slow}}^{U,(F_n)}(\mathbf{X}_\alpha \,|\, \pi)$, so $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}_\alpha) > h_{\text{slow}}^{U,(F_n)}(\mathbf{Y})$, implying that $\mathbf{Y}$ is not isomorphic to $\mathbf{X}_\alpha$ for any $\alpha \in \mathscr{C}$. Therefore $\mathbf{Y}$ is not dominant.

3.4 Open questions

In the special case $G = \mathbb{Z}$, there is another notion of equivalence between measure preserving systems, called **Kakutani equivalence**, that is weaker than measure theoretic isomorphism. Two systems $\mathbf{Y}$ and $\mathbf{Y}'$ are said to be Kakutani equivalent if there are subsets $E \subseteq Y$, $E' \subseteq Y'$ such that the induced return-time systems on E and E' are measure theoretically isomorphic.

Now that the question of dominance has been completely resolved, it is natural to ask the same question with the notion of isomorphism replaced by the weaker notion of Kakutani equivalence. Specifically, say that a system $\mathbf{Y}$ is **Kakutani dominant** if a generic extension $\mathbf{X}$ of $\mathbf{Y}$ is Kakutani equivalent to $\mathbf{Y}$. Because Kakutani equivalence is weaker than measure theoretic isomorphism, the work of [AGT21] already shows that all system $\mathbf{Y}$ with positive Kolmorogov–Sinai entropy are Kakutani dominant. However, nothing is known about the zero entropy case.

In [Rat81], Ratner introduced an invariant for Kakutani equivalence. The definition of Ratner's invariant is almost identical to the definition of slow entropy. The only difference is that for slow entropy, one measures the distance between (P, F)-names using the normalized Hamming metric, while for Ratner's invariant, one uses a different metric known as the f-metric, which is more flexible than the Hamming metric.

The fact that Ratner's invariant is so similar to slow entropy suggests that similar methods may be used to approach the question of Kakutani dominance. However, there are significant additional technical difficulties. One challenge is that because the f-metric is not as rigid as the Hamming metric, it is difficult to estimate the f-distance between two long paths constructed via the process in Proposition 3.12, so it would be difficult to obtain an analogue of Proposition 3.14 Another challenge is presented by the existence of *zero entropy loosely Bernoulli* (or *loosely Kronecker*) systems, which have the property that the ϵ-covering number with respect to the f-metric is equal to one. This means that there is only one (P, F)-name which carries almost all of the measure, which also presents an obstruction to performing the construction described in Proposition 3.12. This phenomenon is completely different from anything that arises in the usual Hamming metric setting. Therefore, the following question has been proposed by Tim Austin.

Question 3.23. Is it true that a zero entropy system $\mathbf{Y}$ is Kakutani dominant if and only if it is loosely Kronecker?

CHAPTER 4

Isometric extensions

In this chapter, we consider a slightly different notion of cocycle than that of Chapter 3. Let H be a compact group. In this chapter, a **cocycle** from Y to H is a measurable map $\alpha : G \times Y \to H$ satisfying the cocycle condition $\alpha(g'g, y) = \alpha(g', S^g y)\alpha(g, y)$ for every $g, g' \in G$ and ν-a.e. y. Given a cocycle α from Y to H and a closed subgroup $K \subseteq H$, we denote by $\mathbf{Y} \times_\alpha \mathbf{H}/\mathbf{K}$ the system $(Y \times H/K, \nu \times m_{H/K}, T_\alpha)$, where

- $m_{H/K}$ is the image of the Haar measure m_H under the quotient map $H \to H/K$, and
- $T_\alpha^g(y, hK) := (S^g y, \alpha(g, y)hK)$.

Note that $\mathbf{Y} \times_\alpha \mathbf{H}/\mathbf{K}$ is an extension of $\mathbf{Y}$ via the projection map onto the first coordinate. Such an extension is also called a **homogeneous skew product** over $\mathbf{Y}$.

Definition 4.1. An extension $\pi : \mathbf{X} \to \mathbf{Y}$ is said to be **isometric** if π is isomorphic to the projection map $\mathbf{Y} \times_\alpha \mathbf{H}/\mathbf{K} \to \mathbf{Y}$, i.e. there is a commutative diagram

$$\begin{array}{ccc} \mathbf{X} & \xrightarrow{\sim} & \mathbf{Y} \times_\alpha \mathbf{H}/\mathbf{K} \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \text{proj}} \\ \mathbf{Y} & \xrightarrow{\text{id}} & \mathbf{Y} \end{array} .$$

Remark 4.2. There are several different equivalent definitions in the literature for what it means for π to be an isometric extension, but the above is the most convenient choice for our purposes. Also, the terms "isometric extension" and "compact extension" are often used interchangeably. The two notions are formally different, but they are known to be equivalent. See [Zor, Definition 4, Definition 15, Theorem 22] and [Gla03, Definition 9.10, Theorem 9.14] for more details.

Definition 4.3. We say $\pi : \mathbf{X} \to \mathbf{Y}$ has **bounded complexity** with respect to the Følner sequence (F_n) if $h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) = 0$ for every rate function U. Equivalently, π has bounded complexity if for every $\epsilon > 0$ and every partition P of X, $\limsup_{n\to\infty} \mathrm{cov}(\mu, P, F_n, \epsilon \,|\, \pi) < \infty$.

The purpose of this chapter is to prove the following characterization of isometric extensions.

Theorem 4.4. *Suppose* $\mathbf{X}$ *is ergodic and let* $\pi : \mathbf{X} \to \mathbf{Y}$ *be an extension. Then the following are equivalent.*

(1) π *is isometric.*

(2) π *has bounded complexity with respect to some Følner sequence.*

(3) π *has bounded complexity with respect to every Følner sequence.*

4.1 Isometric implies bounded

Proposition 4.5. *If* π *is isometric, then it has bounded complexity with respect to any Følner sequence.*

Every homogeneous skew product as in Definition 4.1 is a factor of a group rotation skew product on $Y \times H$ (i.e. a homogeneous skew product with K the trivial subgroup). So, by Theorem 2.3, it is sufficient to assume that $X = Y \times H$, $\mu = \nu \times m_H$, and $T^g(y,h) = (S^g y, \alpha(g,y)h)$ for some cocycle α from Y to H. Let ρ be a translation-invariant metric on H, and let (F_n) be any choice of Følner sequence for G.

Proposition 4.6. *Let $Q = \{Q_1, \ldots, Q_k\}$ be a partition of H. Let P be the partition $\{Y \times Q_1, \ldots, Y \times Q_k\}$ of X. Then for any $\epsilon > 0$, there exists $L = L(Q, \epsilon)$ such that*

$$\operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \;\leq\; L$$

for all n sufficiently large.

Proof. For each i, let Q_i' be a compact subset of Q_i such that

$$m_H(Q_i') \;>\; (1 - \epsilon/4) \cdot m_H(Q_i).$$

Let $E = H \setminus \bigcup Q_i'$, so $m_H(E) \leq \epsilon/4$. Let $\overline{E} = Y \times E$, so $\mu(\overline{E}) \leq \epsilon/4$. Now because the Q_i' are pairwise disjoint compact sets, there is some $\delta = \delta(Q,\epsilon)$ such that $\rho(Q_i', Q_j') \geq \delta$ for all $i \neq j$. Let L be the smallest number of balls of ρ-radius at most $\delta/2$ required to cover H, and let $B_1, \ldots, B_L \subseteq H$ be a collection of such balls. Let $\overline{B_i} = Y \times B_i$.

We claim that $\operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi) \leq L$ for all n sufficiently large. By the mean ergodic theorem and the fact that $\mu(\overline{E}) \leq \epsilon/4$, we have

$$\mu \left\{ (y,h) : \frac{1}{|F_n|} \sum_{g \in F_n} 1_{\overline{E}}(T^g(y,h)) < \epsilon/2 \right\} \;\to\; 1$$

as $n \to \infty$. Call this set X', and let n be sufficiently large so that $\mu(X') \geq 1 - \epsilon^2$. By Markov's inequality, we have a set $S \subseteq Y$ with $\nu(S) \geq 1 - \epsilon$ such that $\mu_y(X') \geq 1 - \epsilon$

for all $y \in S$. Fix $y \in S$; we now estimate $\mathrm{cov}(\mu_y, P, F_n, \epsilon)$. Let $B_i' = \overline{B_i} \cap X' \cap \pi^{-1}y$. Because the $\overline{B_i}$ cover all of X, we have

$$\mu_y\left(\bigcup B_i'\right) \; = \; \mu_y(X') \; \geq \; 1 - \epsilon.$$

Therefore we just need to estimate $\mathrm{diam}_{P,F_n}(B_i')$.

Suppose $(y,h),(y,h') \in B_i'$. We have

$$\begin{aligned} d_{P,F_n}((y,h),(y,h')) \; &= \; \frac{1}{|F_n|}\sum_{g\in F_n} 1_{P(T^g(y,h))\neq P(T^g(y,h'))} \\ &= \; \frac{1}{|F_n|}\sum_{g\in F_n} 1_{P(S^g y,\alpha(g,y)h)\neq P(S^g y,\alpha(g,y)h')} \\ &= \; \frac{1}{|F_n|}\sum_{g\in F_n} 1_{Q(\alpha(g,y)h)\neq Q(\alpha(g,y)h')}. \end{aligned}$$

By definition of the B_i, we have $\rho(h,h') \leq \delta/2$, and because ρ is translation invariant, we also have $\rho(\alpha(g,y)h, \alpha(g,y)h') \leq \delta/2$ for all g. Therefore if $\alpha(g,y)h$ and $\alpha(g,y)h'$ are in different cells of Q, it must be the case that either $\alpha(g,y)h \in E$ or $\alpha(g,y)h' \in E$. So the above becomes

$$\begin{aligned} d_{P,F_n}((y,h),(y,h')) \; &\leq \; \frac{1}{|F_n|}\sum_{g\in F_n} 1_E(\alpha(g,y)h) + 1_E(\alpha(g,y)h') \\ &= \; \frac{1}{|F_n|}\sum_{g\in F_n} 1_{\overline{E}}(T^g(y,h)) + 1_{\overline{E}}(T^g(y,h')) \\ &\leq \; \epsilon \end{aligned}$$

because $(y,h),(y,h') \in X'$. So we have $\mathrm{diam}_{P,F_n}(B_i') \leq \epsilon$. This shows that

$$\mathrm{cov}(\mu_y, P, F_n, \epsilon) \; \leq \; L$$

for all $y \in S$ and n sufficiently large as desired. □

Proof of Proposition 4.5. For each m, let Q_m be a partition of H into sets of diameter at most $1/m$ and let $P_m = \{Y \times C : C \in Q_m\}$ as in Proposition 4.6. It is clear that the sequence $(P_m)_{m=1}^{\infty}$ is generating for $\mathbf{X}$ relative to π. By Proposition 4.6, we have $\limsup_{n\to\infty} \mathrm{cov}(\mu, P_m, F_n, \epsilon \,|\, \pi) < \infty$ for every m and every $\epsilon > 0$. Then by Theorem 2.8, for any partition R of X, we have

$$h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}, R \,|\, \pi) \;\leq\; h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X} \,|\, \pi) \;=\; \lim_{m\to\infty} h_{\mathrm{slow}}^{U,(F_n)}(\mathbf{X}, P_m \,|\, \pi) \;=\; 0$$

for any rate function U, as desired. □

4.2 Background on conditional weak mixing

The second half of the proof of Theorem 4.4 requires the theory of compact and weakly mixing extensions originally developed in [Fur77]. All of the necessary background material presented here can be found in [KL16, Chapter 3 and Appendix D].

Definition 4.7. We say that a subset $\Gamma \subseteq G$ has **absolute density** 1 if

$$\lim_{n\to\infty} \frac{|\Gamma \cap F_n|}{|F_n|} \;=\; 1$$

for *any* Følner sequence (F_n).

Definition 4.8. For $y \in Y$ and $f, g \in L^2(\mathbf{X})$, define

$$\langle f, g \rangle_y \;:=\; \int f\overline{g}\, d\mu_y.$$

Let $L^2(\mathbf{X} \,|\, \pi)$ denote the space of $f \in L^2(\mathbf{X})$ such that $y \mapsto \langle f, f \rangle_y \in L^\infty(\mathbf{Y})$. We also say that $f, g \in L^2(\mathbf{X} \,|\, \pi)$ are **conditionally orthogonal given** π if $\langle f, g \rangle_y = 0$ for ν-a.e. $y \in Y$.

In this section, we identify the action T with its Koopman representation on $L^2(\mathbf{X})$. So, for $g \in G$ and $f \in L^2(\mathbf{X})$, we write $T^g f$ to mean $f \circ T^g$.

Definition 4.9. A function $f \in L^2(\mathbf{X}\,|\,\pi)$ is said to be **conditionally weakly mixing given** π if for any Følner sequence (F_n) and any $g \in L^2(\mathbf{X}\,|\,\pi)$, we have

$$\lim_{n\to\infty} \frac{1}{|F_n|} \sum_{s \in F_n} \int \left| \langle T^s f, g \rangle_y \right| \, d\nu(y) \;=\; 0.$$

Equivalently, for any $g \in L^2(\mathbf{X}\,|\,\pi)$ and any $\epsilon > 0$, the set

$$\Gamma_{f,g,\epsilon} \;:=\; \left\{ s \in G : \int \left| \langle T^s f, g \rangle_y \right| \, d\nu(y) < \epsilon \right\}$$

has absolute density 1. The set of all conditionally weakly mixing functions is denoted $W(\mathbf{X}\,|\,\pi)$.

The main fact we will need to use is the following characterization of the maximal intermediate isometric extension (essentially [KL16, Proposition 3.9 and Lemma 3.11]).

Theorem 4.10. *Let* $\mathbf{X}$ *be ergodic and* $\pi : \mathbf{X} \to \mathbf{Y}$ *be an extension. Then there exists an intermediate extension* $\mathbf{X} \to \mathbf{Z} \to \mathbf{Y}$ *such that*

- $\mathbf{Z}$ *is the maximal isometric extension of* $\mathbf{Y}$ *in* $\mathbf{X}$*, and*
- *For* $f \in L^2(\mathbf{X}\,|\,\pi)$*,* $f \in W(\mathbf{X}\,|\,\pi)$ *if and only if* f *is conditionally orthogonal to every* $\mathbf{Z}$*-measurable* $h \in L^2(\mathbf{X}\,|\,\pi)$*.*

4.3 Bounded implies isometric

Proposition 4.11. *If* π *has bounded complexity with respect to some Følner sequence* (F_n)*, then it is isometric.*

Suppose for contradiction that π is not isometric but does have bounded complexity with respect to some Følner sequence (F_n). Let $\mathbf{Z}$ be the maximal intermediate isometric extension as in Theorem 4.10. Because of the assumption that π is not isometric, we know that $\mathbf{Z}$ is a strict factor of $\mathbf{X}$, so we can choose a partition $P = \{P_0, P_1\}$ of X satisfying

(1) P is independent of $\mathbf{Z}$

(2) $\mu_y(P_0), \mu_y(P_1) \geq 1/3$ for ν-a.e. y. [1]

Fix this partition for the rest of this section. Also let $0 < \epsilon < 10^{-6}$ be fixed. Finally, using the notation of Definition 2.10, for $y \in Y$ we abbreviate $\mathrm{dist}_y := \mathrm{dist}_{\mu_y}$.

The outline of the proof of Proposition 4.11 is as follows. First, using the assumption that π has bounded complexity, we will find a "positive density" set of pairs of times $(s,t) \in G^2$ such that $T^{s^{-1}}P$ and $T^{t^{-1}}P$ are very close to each other (in most of the fibers of π). This is Lemma 4.12. Then, using the independence conditions built in to the definition of P, we show essentially that the partition P is conditionally weakly mixing given π, which allows us to find a "density one" set of pairs of times $(s,t) \in G^2$ such that $T^{s^{-1}}P$ and $T^{t^{-1}}P$ are approximately independent of each other (in most of the fibers of π). This is Lemma 4.16. Therefore we can find a pair of times (s,t) for which $T^{s^{-1}}P$ and $T^{t^{-1}}P$ are both close together and approximately independent of each other. But it is impossible for two nontrivial partitions to satisfy this, so we will get a contradiction.

Lemma 4.12. *For $y \in Y$, define*

$$\mathscr{C}_y \ := \ \left\{ (s,t) \in G^2 : \mathrm{dist}_y(T^{s^{-1}}P, T^{t^{-1}}P) \ < 5\sqrt{\epsilon} \right\}.$$

[1] If the factor map $\mathbf{X} \to \mathbf{Z}$ is infinite-to-one, then 1/3 may be replaced by 1/2.

Then there is a constant $c = c(\epsilon) > 0$ such that the following holds. For every n, there is a set $Y_n \subseteq Y$ satisfying $\nu(Y_n) \geq 1 - \epsilon$ and

$$\frac{|\mathscr{C}_y \cap F_n^2|}{|F_n^2|} \;\geq\; c(\epsilon)$$

for all $y \in Y_n$.

Proof. Let $L = L(\epsilon) = \sup_n \operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi)$ and let n be arbitrary. Let Y_n be the set of $y \in Y$ such that $\operatorname{cov}(\mu_y, P, F_n, \epsilon) \leq L$. We have $\nu(Y_n) \geq 1 - \epsilon$ by definition. For the rest of this proof, let $y \in Y_n$ be fixed. We seek to bound $|\mathscr{C}_y \cap F_n^2|/|F_n^2|$ from below by a quantity depending only on ϵ.

Let $B_1, \ldots, B_L$ be subsets of X such that each B_i has d_{P,F_n}-diameter at most ϵ and $\mu_y \left(\bigcup B_i\right) \geq 1 - \epsilon$. Let $X' = \bigcup B_i$. Without loss of generality, we may assume that the B_i are disjoint. For each i, fix a point $x_i \in B_i$. Then, for $x \in X'$, define $r(x)$ to be the unique x_i such that $x \in B_i$.

By construction, we know that for each $x \in X'$, $T^s x$ and $T^s r(x)$ lie in the same P-cell for most $s \in F_n$, but the set of good "times" s changes as x varies. We now apply a form of Markov's inequality to upgrade this to the statement that for most $s \in F_n$, μ_y-most x satisfy $P(T^s x) = P(T^s r(x))$. Define

$$A \;=\; \Big\{ s \in F_n : \mu_y\{x \in X' : P(T^s x) = P(T^s r(x))\} \geq 1 - \sqrt{2\epsilon} \Big\}.$$

We have

$$\begin{aligned}
\sum_{s\in F_n} \mu_y\{x \in X' : P(T^s x) = P(T^s r(x))\} &= \sum_{s\in F_n}\sum_{i=1}^{L} \mu_y\{x \in B_i : P(T^s x) = P(T^s x_i)\} \\
&= \sum_{s\in F_n}\sum_{i=1}^{L} \int_{B_i} 1_{P(T^s x)=P(T^s x_i)}\, d\mu_y(x) \\
&= \sum_{i=1}^{L} \int_{B_i} \sum_{s\in F_n} 1_{P(T^s x)=P(T^s x_i)}\, d\mu_y(x) \\
&= \sum_{i=1}^{L} \int_{B_i} |F_n|(1 - d_{P,F_n}(x, x_i))\, d\mu_y(x) \\
&\geq \sum_{i=1}^{L} \mu_y(B_i)|F_n|(1-\epsilon) \\
&\geq |F_n|(1-\epsilon)^2 \;\geq\; |F_n|(1-2\epsilon).
\end{aligned}$$

But the original sum above also satisfies

$$\begin{aligned}
\sum_{s\in F_n} \mu_y\{x \in X' : P(T^s x) = P(T^s r(x))\} &= \sum_{s\in A} \mu_y\{x \in X' : P(T^s x) = P(T^s r(x))\} + \\
&\quad \sum_{s\notin A} \mu_y\{x \in X' : P(T^s x) = P(T^s r(x))\} \\
&\leq |A| + (|F_n| - |A|)(1-\sqrt{2\epsilon}) \\
&= |F_n|(1-\sqrt{2\epsilon}) + |A|\cdot\sqrt{2\epsilon}.
\end{aligned}$$

Combining these two inequalities shows that

$$|A| \;\geq\; \frac{|F_n|(1-2\epsilon-(1-\sqrt{2\epsilon}))}{\sqrt{2\epsilon}} \;=\; |F_n|(1-\sqrt{2\epsilon}). \tag{4.1}$$

The set A decomposes as

$$A \;=\; \bigcup_{w\in\{0,1\}^L} \{s \in A : (P(T^s x_i))_{i=1}^{L} = w\}.$$

By (4.1) and the pigeonhole principle, there is some $w \in \{0,1\}^L$ such that

$$\left|\{s \in A : (P(T^s x_i))_{i=1}^L = w\}\right| \ \geq \ |F_n|(1-\sqrt{2\epsilon}) \cdot 2^{-L}. \tag{4.2}$$

Call this set $\mathcal{E}$. For $s, t \in \mathcal{E}$, say that x is (s,t)-*good* if $P(T^s x) = P(T^s r(x))$ and $P(T^t x) = P(T^t r(x))$. By definition of A, the set of x that are not (s,t)-good has μ_y-measure at most $2\sqrt{2\epsilon}$. Now, for $s, t \in \mathcal{E}$, we can estimate

$$\begin{aligned} \operatorname{dist}_y(T^{s^{-1}}P, T^{t^{-1}}P) \ &= \ \int 1_{P(T^s x) \neq P(T^t x)} \, d\mu_y(x) \ \leq \ \epsilon + \int_{X'} 1_{P(T^s x) \neq P(T^t x)} \, d\mu_y(x) \\ &\leq \ \epsilon + 2\sqrt{2\epsilon} + \int_{\{x \in X' : \ x \text{ is } (s,t)\text{-good}\}} 1_{P(T^s x) \neq P(T^t x)} \, d\mu_y(x) \\ &= \ \epsilon + 2\sqrt{2\epsilon} + \sum_{i=1}^{L} \mu_y(B_i) 1_{P(T^s x_i) \neq P(T^t x_i)}. \end{aligned}$$

By definition of $\mathcal{E}$, $P(T^s x_i) = w_i = P(T^t x_i)$ for all i, so this final sum vanishes and we conclude $\operatorname{dist}_y(T^{s^{-1}}P, T^{t^{-1}}P) \leq \epsilon + 2\sqrt{2\epsilon} \leq 5\sqrt{\epsilon}$ whenever $s, t \in \mathcal{E}$.

Finally, observe that $\mathscr{C}_y$ contains $\mathcal{E} \times \mathcal{E}$. Therefore, by (4.2), we have

$$\frac{|\mathscr{C}_y \cap F_n^2|}{|F_n^2|} \ \geq \ \left((1-2\sqrt{\epsilon}) \cdot 2^{-L(\epsilon)}\right)^2 \ > \ 0$$

as claimed. □

This finishes the first half of our outline. For convenience, we now introduce some new definitions before starting the second half.

Definition 4.13. Given $y \in Y$ and two sets $A, B \subseteq X$, we define the **dependence score** with respect to μ_y to be

$$\mathscr{D}_y(A,B) \ := \ |\mu_y(A \cap B) - \mu_y(A)\mu_y(B)|.$$

We also define the **averaged dependence score**

$$\mathscr{D}(A,B) \ = \ \int \mathscr{D}_y(A,B) \, d\nu(y).$$

Finally, if Q and Q' are two finite partitions of X, then the averaged dependence score between Q and Q' is defined to be

$$\mathscr{D}(Q, Q') \;=\; \max_{i,j} \mathscr{D}(Q_i, Q'_j).$$

Lemma 4.14. *Let (F_n) be a Følner sequence for G and let (g_n) be an arbitrary sequence of elements of G. Then $(F_n g_n)$ is also a Følner sequence for G.*

Proof. For any $h \in G$, we have

$$\frac{|hF_n g_n \cap F_n g_n|}{|F_n g_n|} \;=\; \frac{|(hF_n \cap F_n)g_n|}{|F_n g_n|} \;=\; \frac{|hF_n \cap F_n|}{|F_n|} \;\to\; 1$$

as $n \to \infty$. □

Lemma 4.15. *Let $\Gamma \subseteq G$ be a subset of absolute density 1. Define $\Gamma' = \{(s,t) \in G^2 : ts^{-1} \in \Gamma\}$. Then if (F_n) is any Følner sequence for G, we have*

$$\lim_{n\to\infty} \frac{|\Gamma' \cap F_n^2|}{|F_n^2|} \;=\; 1.$$

Proof. Let (F_n) be a left Følner sequence for G. We calculate

$$\begin{aligned}
|\Gamma' \cap F_n^2| \;&=\; \#\{(s,t) \in F_n^2 : ts^{-1} \in \Gamma\} \\
&=\; \sum_{s \in F_n} \#\{t \in F_n : ts^{-1} \in \Gamma\} \\
&=\; \sum_{s \in F_n} |F_n \cap \Gamma s| \\
&=\; \sum_{s \in F_n} |F_n s^{-1} \cap \Gamma| \\
&\geq\; |F_n| \cdot |F_n s_n^{-1} \cap \Gamma|,
\end{aligned}$$

where $s_n \in F_n$ is defined to be the element of F_n that minimizes $|F_n s^{-1} \cap \Gamma|$ over all $s \in F_n$. By Lemma 4.14, $(F_n s_n^{-1})$ is also a Følner sequence, so because Γ has absolute density 1 we get

$$\begin{aligned}\lim_{n\to\infty} \frac{|\Gamma' \cap F_n^2|}{|F_n^2|} &\geq \lim_{n\to\infty} \frac{|F_n| \cdot |F_n s_n^{-1} \cap \Gamma|}{|F_n|^2} \\ &= \lim_{n\to\infty} \frac{|F_n s_n^{-1} \cap \Gamma|}{|F_n|} = \lim_{n\to\infty} \frac{|F_n s_n^{-1} \cap \Gamma|}{|F_n s_n^{-1}|} = 1.\end{aligned}$$

□

Now fix another parameter $0 < \eta \ll \epsilon$ which is small enough so that $3\eta^{1/4} < c(\epsilon)/2$, where $c(\epsilon)$ is the quantity from Lemma 4.12.

Lemma 4.16. *For $y \in Y$, define*

$$\mathcal{I}_y := \left\{ (s,t) \in G^2 : \mathscr{D}_y(T^{s^{-1}}P, T^{t^{-1}}P) \leq \sqrt{\eta} \right\}.$$

Then, for all sufficiently large n, there is a set $Y_n^\dagger \subseteq Y$ such that $\nu(Y_n^\dagger) \geq 1 - 3\eta^{1/4}$ and

$$\frac{|\mathcal{I}_y \cap F_n^2|}{|F_n^2|} \geq 1 - 3\eta^{1/4}$$

for all $y \in Y_n^\dagger$.

Proof. Property (1) in the definition of the partition P implies that if f is any P-measurable function satisfying $\int f \, d\mu_y = 0$ for ν-a.e. y, and h is any $\mathbf{Z}$-measurable function, then also $\langle f, h \rangle_y = 0$ for ν-a.e. y. By the second bullet point of Theorem 4.10, this implies that any such f is conditionally weak mixing given π.

Let $f_0 = 1_{P_0} - \mu(P_0)$ and $f_1 = 1_{P_1} - \mu(P_1)$. Clearly these are both P-measurable, and because P is independent of $\mathbf{Z}$ and therefore also independent of $\mathbf{Y}$, we also have $\int f_0\, d\mu_y = \int f_1\, d\mu_y = 0$ for ν-a.e. y. Therefore f_0 and f_1 are both conditionally weakly mixing. Observe that

$$\begin{aligned}
\left|\langle f_0, T^s f_1\rangle_y\right| &= \left|\int f_0 \cdot T^s f_1\, d\mu_y\right| = \left|\int (1_{P_0} - \mu(P_0))(1_{T^{s^{-1}}P_1} - \mu(P_1))\, d\mu_y\right| \\
&= \left|\mu_y(P_0 \cap T^{s^{-1}}P_1) - \mu(P_0)\mu_y(T^{s^{-1}}P_1) - \mu(P_1)\mu_y(P_0) + \mu(P_0)\mu(P_1)\right| \\
&= \left|\mu_y(P_0 \cap T^{s^{-1}}P_1) - \mu_y(P_0)\mu_y(T^{s^{-1}}P_1)\right| \\
&= \mathscr{D}_y(P_0, T^{s^{-1}}P_1).
\end{aligned}$$

In the second to last line we again used the fact that P is independent of $\mathbf{Y}$. So by the discussion in Definition 4.9, the set

$$\Gamma_{0,1} := \left\{s \in G : \mathscr{D}(P_0, T^{s^{-1}}P_1) < \eta\right\}$$

has absolute density 1.

Applying the same analysis to $\langle f_0, T^s f_0\rangle_y$, $\langle f_1, T^s f_0\rangle_y$, and $\langle f_1, T^s f_1\rangle_y$ gives the same conclusion for each of the sets

$$\Gamma_{i,j} := \left\{s \in G : \mathscr{D}(P_i, T^{s^{-1}}P_j) < \eta\right\}.$$

It follows that the set

$$\Gamma := \bigcap_{0\le i,j\le 1} \Gamma_{i,j} = \left\{s \in G : \mathscr{D}(P, T^{s^{-1}}P) < \eta\right\}$$

also has absolute density 1.

As in Lemma 4.15, we now define the set of pairs

$$\Gamma' := \left\{(s,t) \in G^2 : \mathscr{D}(T^{s^{-1}}P, T^{t^{-1}}P) < \eta\right\} = \left\{(s,t) \in G^2 : ts^{-1} \in \Gamma\right\}.$$

Fix any $(s,t) \in \Gamma'$. For each i, j, Markov's inequality implies that there is a subset of Y of measure at least $1 - \sqrt{\eta}$ on which $\mathscr{D}_y(T^{s^{-1}}P_i, T^{t^{-1}}P_j) < \sqrt{\eta}$. Let $Y_{s,t}$

be the intersection of those sets over $0 \le i, j \le 1$; then we have $\nu(Y_{s,t}) \ge 1 - 4\sqrt{\eta}$ and $\mathscr{D}_y(T^{s^{-1}}P, T^{t^{-1}}P) < \sqrt{\eta}$ for all $y \in Y_{s,t}$.

Finally, we estimate the size of $\mathcal{I}_y$ for most y (recall the statement of Lemma 4.16 for the definition of $\mathcal{I}_y$). We have

$$\begin{aligned}
\int \left|\mathcal{I}_y \cap F_n^2\right| \, d\nu(y) &= \int \sum_{(s,t)\in F_n^2} 1_{\mathscr{D}_y(T^{s^{-1}}P, T^{t^{-1}}P) \le \sqrt{\eta}} \quad d\nu(y) \\
&= \sum_{(s,t)\in F_n^2} \nu\left\{ y \in Y : \mathscr{D}_y(T^{s^{-1}}P, T^{t^{-1}}P) \le \sqrt{\eta} \right\} \\
&\ge \sum_{(s,t)\in \Gamma' \cap F_n^2} \nu(Y_{s,t}) \\
&\ge |\Gamma' \cap F_n^2| \cdot (1 - 4\sqrt{\eta}).
\end{aligned}$$

By Lemma 4.15, Γ' has density 1 with respect to (F_n^2), so for n sufficiently large we have

$$\int \frac{|\mathcal{I}_y \cap F_n^2|}{|F_n^2|} \, d\nu(y) \;\ge\; \frac{|\Gamma' \cap F_n^2| \cdot (1 - 4\sqrt{\eta})}{|F_n^2|} \;\ge\; 1 - 5\sqrt{\eta}.$$

It follows by Markov's inequality that there is a set $Y_n^\dagger \subseteq Y$ with $\nu(Y_n^\dagger) \ge 1 - \sqrt{5\sqrt{\eta}} \ge 1 - 3\eta^{1/4}$ such that

$$\frac{|\mathcal{I}_y \cap F_n^2|}{|F_n^2|} \;\ge\; 1 - \sqrt{5\sqrt{\eta}} \;\ge\; 1 - 3\eta^{1/4}$$

for all $y \in Y_n^\dagger$, as claimed. □

Proof of Proposition 4.11. We show that there exists $y \in Y$ such that $\mathcal{C}_y \cap \mathcal{I}_y \neq \emptyset$. This is sufficient because $(s,t) \in \mathcal{C}_y$ implies that $\mathrm{dist}_y(T^{s^{-1}}P, T^{t^{-1}}P) \le \epsilon$, while $(s,t) \in \mathcal{I}_y$ implies that $\mathscr{D}_y(T^{s^{-1}}P, T^{t^{-1}}P) \le \sqrt{\eta}$. But because $\mu_y(P_0), \mu_y(P_1) \ge 1/3$ for all y and $\eta < \epsilon < 10^{-6}$, these two conditions contradict each other. Indeed, $\mathrm{dist}_y(T^{s^{-1}}P, T^{t^{-1}}P) \le \epsilon$ implies in particular that

$$\mu_y(T^{s^{-1}}P_0 \cap T^{t^{-1}}P_1) < \epsilon. \tag{4.3}$$

But the dependence score condition implies that

$$\mu_y(T^{s^{-1}}P_0 \cap T^{t^{-1}}P_1) \;>\; \mu_y(T^{s^{-1}}P_0)\mu_y(T^{t^{-1}}P_1) - \sqrt{\eta} \;>\; \frac{1}{9} - \sqrt{\eta},$$

which contradicts (4.3).

To find such a y, first choose n large enough to satisfy the hypothesis of Lemma 4.16. Then, let Y_n be the set guaranteed by Lemma 4.12 and $Y_n^\dagger$ be the set guaranteed by Lemma 4.16. We have chosen ϵ and η small enough to ensure $\nu(Y_n \cap Y_n^\dagger) > 0$, so choose $y \in Y_n \cap Y_n^\dagger$. Then $|\mathcal{C}_y \cap F_n^2| \geq c(\epsilon) \cdot |F_n^2|$ and $|\mathcal{I}_y \cap F_n^2| \geq (1 - 3\eta^{1/4}) \cdot |F_n^2|$, so by our choice of η, we are guaranteed that $\mathcal{C}_y \cap \mathcal{I}_y \neq \emptyset$. □

As a corollary, we also get a characterization of weakly mixing extensions in terms of relative slow entropy. Recall that $\pi : \mathbf{X} \to \mathbf{Y}$ is said to be **weakly mixing** if there are no intermediate isometric extensions except for the trivial one $\mathbf{Y} \to \mathbf{Y}$.

Corollary 4.17. *Suppose* $\mathbf{X}$ *is ergodic. Then* $\pi : \mathbf{X} \to \mathbf{Y}$ *is weakly mixing if and only if for every partition* P *which is not* $\mathbf{Y}$*-measurable, there exists a rate function* U *and a Følner sequence* (F_n) *such that* $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P \,|\, \pi) > 0$.

Proof. First suppose π is not weakly mixing. Then there is a nontrivial isometric extension $\mathbf{Z} \to \mathbf{Y}$. Then if P is any $\mathbf{Z}$-measurable partition, Theorem 4.4 implies that $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P \,|\, \pi) = 0$ for every rate function U and every Følner sequence (F_n). Because $\mathbf{Z}$ strictly extends $\mathbf{Y}$, we can choose this P to not be $\mathbf{Y}$-measurable.

Conversely, suppose there is a partition P, not measurable with respect to $\mathbf{Y}$, satisfying

$$h_{\text{slow}}^{U,(F_n)}(\mathbf{X}, P \,|\, \pi) \;=\; 0$$

for every U and every (F_n). Then the T-invariant σ-algebra $\pi^{-1}\mathcal{B}_Y \vee \bigvee_{s \in G} T^{s^{-1}}P$ corresponds to an intermediate extension $\mathbf{Z} \to \mathbf{Y}$. Because P is not $\mathbf{Y}$-measurable, this is a nontrivial extension. So because P is relatively generating for $\mathbf{Z}$ with respect to $\mathbf{Y}$, this implies that $h^{U,(F_n)}_{\mathrm{slow}}(\mathbf{Z}\,|\,\pi) = 0$ for all U and (F_n). Therefore Theorem 4.4 implies that $\mathbf{Z} \to \mathbf{Y}$ is isometric, so π is not weakly mixing. $\square$

CHAPTER 5

Rigid extensions

5.1 Definitions

Let $G = \mathbb{Z}$. Recall that a system $\mathbf{X}$ is said to be **rigid** if there exists a sequence $0 = n_0 < n_1 < n_2 < \dots$ such that

$$\lim_{k\to\infty} \mu(T^{-n_k} A \triangle A) \ = \ 0$$

for all measurable $A \subseteq X$.

While there have been some attempts to relativize this notion and define what it means for an extension $\pi : \mathbf{X} \to \mathbf{Y}$ to be rigid (see for example [Sch18, Definition 4]), thus far no definition has been completely satisfactory. In this chapter, we will give a new definition of rigid extension and demonstrate some of its properties.

Definition 5.1. Let $\mathrm{Aut}(I, m)$ denote the space of Lebesgue measure-preserving automorphisms of the unit interval I, modulo the equivalence relation of m-a.e. agreement. This space is a Polish topological group when endowed with the weak topology defined by the property that a sequence (φ_n) converges to φ if and only if $m(\varphi_n^{-1} E \triangle \varphi^{-1} E) \to 0$ for all measurable $E \subseteq I$. We can define a metric that generates this topology as follows. For $k \geq 1$, let $\mathcal{D}_k$ be the partition of I into intervals of length 2^{-k}. Then, for φ, $\psi \in \mathrm{Aut}(I, m)$, define

$$d_A(\varphi, \psi) \;=\; \sum_{k \geq 1} 2^{-k} \, \mathrm{dist}_m(\varphi^{-1}\mathcal{D}_k, \psi^{-1}\mathcal{D}_k).$$

Note in particular that a sequence (φ_n) converges to $\mathrm{id} \in \mathrm{Aut}(I, m)$ if and only if

$$\lim_{n \to \infty} \mathrm{dist}_m \left(\mathcal{D}_j, \varphi_n^{-1} \mathcal{D}_j \right) \;=\; 0$$

for any fixed j.

In this chapter, we use the same definition of cocycle as in Chapter 3. However, because we only consider the special case $G = \mathbb{Z}$, the cocycle condition implies that the full cocycle $(\alpha_n)_{n \in \mathbb{Z}}$ is completely determined by α_0. Therefore, we will abbreviate and simply say that a **cocycle** on Y is any measurable map $\alpha : Y \to \mathrm{Aut}(I, m)$. Recall that a cocycle α on Y induces the skew product system

$$\mathbf{X}_\alpha \;=\; (Y \times I, \nu \times m, T_\alpha)$$

where $T_\alpha(y, t) := (Sy, \alpha(y)t)$. For $n \in \mathbb{N}$, define

$$\alpha_n(y) \;:=\; \alpha(S^{n-1}y) \circ \cdots \circ \alpha(Sy) \circ \alpha(y),$$

so that $T_\alpha^n(y, t) = (S^n y, \alpha_n(y)t)$.

Definition 5.2. We say that $\mathbf{X}_\alpha$ is a **rigid extension** of $\mathbf{Y}$ if for ν-a.e. $y \in Y$, there is a subsequence (n_k) such that $\alpha_{n_k}(y) \to \mathrm{id}$ as $k \to \infty$. Such a sequence (n_k) is called a **rigidity sequence** for y. We will also use the terminology that α is a **rigid cocycle**.

Remark 5.3. There are a few things to note about this definition.

(1) The rigidity sequence (n_k) is allowed to depend on the base point y. This is the main difference between our definition and previous definitions and it is crucial to everything we are able to prove about rigid extensions.

(2) If $\mathbf{Y}$ is trivial, this definition reduces to the usual definition of a rigid system.

(3) By Rokhlin's skew product theorem, any infinite-to-one ergodic extension of $\mathbf{Y}$ is isomorphic to a skew product of the above form. However, we are not able to show that this definition of rigid extension is isomorphism invariant, so for now we are limited to making this definition only for skew product systems.

For future convenience, we record here a simple condition that implies the rigidity of a cocycle α.

Lemma 5.4. *Let α be a cocycle on Y. Let $\mathcal{D}_k$ be the depth-k dyadic partition of I and let $P_k = \{Y \times E : E \in \mathcal{D}_k\}$. In order to show that α is a rigid cocycle, it is sufficient to show that for every $k \geq 1$ and every $\epsilon > 0$, we have*

$$\nu\left\{y \in Y : \mathrm{dist}_y(P_k, T_\alpha^{-n} P_k) < \epsilon \text{ for infinitely many } n\right\} \ = \ 1.$$

Proof. For each k, define

$$\begin{aligned} \mathcal{R}_k \ &:= \ \left\{y \in Y : \mathrm{dist}_y(P_k, T_\alpha^{-n} P_k) < 1/k \text{ for infinitely many } n\right\} \quad \text{and} \\ \mathcal{R} \ &:= \ \bigcap_{k \geq 1} \mathcal{R}_k. \end{aligned}$$

By assumption, we have $\nu(\mathcal{R}) = 1$. We show that any $y \in \mathcal{R}$ has a rigidity sequence.

Fix $y \in \mathcal{R}$. For each k, pick a sequence of times $n_{k,1} < n_{k,2} < \ldots$ such that $\mathrm{dist}_y(P_k, T_\alpha^{-n_{k,\ell}} P_k) < 1/k$ for every k and every ℓ. Such a sequence exists by the definition of the sets $\mathcal{R}_k$. By simply deleting finitely many times if necessary, we may also assume that $n_{k+1,k+1} > n_{k,k}$ for every k.

Now we claim that the sequence of times $(n_{k,k})$ is a rigidity sequence for y. By the discussion in Definition 5.1, it suffices to show that for any fixed j,

$$\mathrm{dist}_m(\mathcal{D}_j, \alpha_{n_{k,k}}(y)^{-1}\mathcal{D}_j) \to 0 \qquad \text{as } k \to \infty.$$

Observe that for any $s \in \mathbb{N}$,

$$\begin{aligned}
\mathrm{dist}_m(\mathcal{D}_j, \alpha_s(y)^{-1}\mathcal{D}_j) &= m\{t : \mathcal{D}_j(t) \neq \mathcal{D}_j(\alpha_s(y)t)\} \\
&= \mu_y\{(y,t) : \mathcal{D}_j(t) \neq \mathcal{D}_j(\alpha_s(y)t)\} \\
&= \mu_y\{(y,t) : P_j(y,t) \neq P_j(T_\alpha^s(y,t))\} \\
&= \mathrm{dist}_y(P_j, T_\alpha^{-s} P_j).
\end{aligned}$$

Because P_j is refined by P_k for all $k \geq j$, we get

$$\begin{aligned}
\lim_{k\to\infty} \mathrm{dist}_m(\mathcal{D}_j, \alpha_{n_{k,k}}(y)^{-1}\mathcal{D}_j) &= \lim_{k\to\infty} \mathrm{dist}_y(P_j, T_\alpha^{-n_{k,k}} P_j) \\
&\leq \lim_{k\to\infty} \mathrm{dist}_y(P_k, T_\alpha^{-n_{k,k}} P_k) \;=\; 0
\end{aligned}$$

as desired. □

5.2 Genericity

First, we show that generic extensions of an ergodic system are rigid. First, let us recall some basic definitions. We denote by $\mathrm{Co}(Y)$ the set of all cocycles on Y, i.e. the set of all measurable maps $\alpha : Y \to \mathrm{Aut}(I, m)$. Each $\alpha \in \mathrm{Co}(Y)$ induces an extension $\mathbf{X}_\alpha = (Y \times I, \nu \times m, T_\alpha)$ of $\mathbf{Y}$ via the skew product transformation $T_\alpha(y,t) = (Sy, \alpha(y)t)$. By identifying each $\alpha \in \mathrm{Co}(Y)$ with the skew product $T_\alpha \in$

$\mathrm{Aut}(Y \times I, \nu \times m)$, we endow $\mathrm{Co}(Y)$ with the topology it inherits as a subspace of the weak topology on $\mathrm{Aut}(Y \times I, \nu \times m)$.

Theorem 5.5. *Let* $\mathbf{Y}$ *be ergodic. Then the set of rigid cocycles* $\alpha \in \mathrm{Co}(Y)$ *is a dense* G_δ *set.*

Remark 5.6. In [Sch18], the author uses a different definition of rigid extension and shows that the set of rigid cocycles forms a G_δ set, but is not able to show that it is dense. Here, because we allow the rigidity sequence to depend on the base point, we are able to establish density as well.

As in the previous section, let $\mathcal{D}_k$ be the level-k dyadic partition of I and let

$$P_k = \{Y \times E : E \in \mathcal{D}_k\}.$$

Given $k, N \in \mathbb{N}$, $\epsilon > 0$, and $\alpha \in \mathrm{Co}(Y)$, define the set

$$\mathcal{R}_{k,N,\epsilon}(\alpha) := \left\{y \in Y : \text{ there exists an } n > N \text{ such that } \mathrm{dist}_y(P_k, T_\alpha^{-n} P_k) < \epsilon\right\}.$$

Given another parameter $\eta > 0$, also define

$$\mathcal{U}_{k,N,\epsilon,\eta} := \{\alpha \in \mathrm{Co}(Y) : \nu(\mathcal{R}_{k,N,\epsilon}(\alpha)) > 1 - \eta\}.$$

Lemma 5.7. *The set of rigid cocycles* $\alpha \in \mathrm{Co}(Y)$ *is given by*

$$\bigcap_{k \geq 1} \bigcap_{\epsilon \searrow 0} \bigcap_{\eta \searrow 0} \bigcap_{N \geq 1} \mathcal{U}_{k,N,\epsilon,\eta},$$

where the intersections over η *and* ϵ *should be interpreted as intersections over countable sequences tending to* 0.

Proof. It's clear that every rigid cocycle α satisfies $\nu(\mathcal{R}_{k,N,\epsilon}(\alpha)) = 1$ for all k, N, ϵ, η, so therefore α is an element of every $\mathcal{U}_{k,N,\epsilon,\eta}$.

Conversely, suppose α is an element of every $\mathcal{U}_{k,N,\epsilon,\eta}$. This implies that

$$\nu(\mathcal{R}_{k,N,\epsilon}(\alpha)) \;>\; 1-\eta$$

for all $\eta > 0$, so $\nu(\mathcal{R}_{k,N,\epsilon}(\alpha)) = 1$. This holds for every N, so

$$\nu\left(\bigcap_{N\geq 1} \mathcal{R}_{k,N,\epsilon}(\alpha)\right) \;=\; \nu\{y \in Y : \operatorname{dist}_y(P_k, T_\alpha^{-n} P_k) < \epsilon \text{ for infinitely many } n\} \;=\; 1$$

as well. Finally, this holds for every k and every $\epsilon > 0$, so by Lemma 5.4, we conclude that α is a rigid cocycle. □

Lemma 5.8. *Each* $\mathcal{U}_{k,N,\epsilon,\eta}$ *is dense in* $\operatorname{Co}(Y)$.

Proof. Recall that a **dyadic permutation of rank** M is an element $\varphi \in \operatorname{Aut}(I, m)$ that permutes the cells of $\mathcal{D}_M$ and acts as a translation on each cell. By [GW19, Lemma 1.2], the set of piecewise constant cocycles is dense in $\operatorname{Co}(Y)$. By Halmos's Weak Approximation Theorem [Hal56, page 65], the set of dyadic permutations is dense in $\operatorname{Aut}(I, m)$. Therefore, we consider the dense set $\mathscr{D}$ of cocycles α such that $\{\alpha(y) : y \in Y\}$ is a finite set of dyadic permutations. We show that each of the sets $\mathcal{U}_{k,N,\epsilon,\eta}$ contains $\mathscr{D}$. To do this, it is clearly sufficient to show that each $\alpha \in \mathscr{D}$ is a rigid cocycle.

Fix $\alpha \in \mathscr{D}$. Because α takes only finitely many values, there is some M such that each $\alpha(y)$ is a dyadic permutation of rank M. So we may consider α to be a map from Y into Sym_M (the subgroup of $\operatorname{Aut}(I, m)$ consisting of dyadic permutations of rank M, isomorphic to the symmetric group on M elements). Define

$$\mathcal{R} \;=\; \{y \in Y : \alpha_n(y) = \operatorname{id} \text{ for infinitely many } n\}.$$

We want to show that $\nu(\mathcal{R}) = 1$. To do this, fix $y \in Y$ and let

$$\Sigma_y \;=\; \{\sigma \in \operatorname{Sym}_M : \alpha_n(y) = \sigma \text{ for infinitely many } n\}.$$

Observe that $\{n \in \mathbb{N} : \alpha_n(y) \in \Sigma_y\}$ must be co-finite. Now we claim that if $\alpha_n(y) \in \Sigma_y$, then $S^n y \in \mathcal{R}$. This is because if $\alpha_n(y) \in \Sigma_y$, then there are infinitely many $m > n$ satisfying $\alpha_m(y) = \alpha_n(y)$. For all such m, we have $\alpha_{m-n}(S^n y)\alpha_n(y) = \alpha_m(y)$, which implies that $\alpha_{m-n}(S^n y) = \text{id}$ for infinitely many m, so $S^n y \in \mathcal{R}$.

Therefore we have shown that for every $y \in Y$, the set of n such that $S^n y \in \mathcal{R}$ is co-finite. By ergodicity, this implies that $\nu(\mathcal{R}) = 1$ as desired. □

Lemma 5.9. *Each $\mathcal{U}_{k,N,\epsilon,\eta}$ is open in* $\text{Co}(Y)$.

Proof. Fix $\alpha \in \mathcal{U}_{k,N,\epsilon,\eta}$. We may write the set $\mathcal{R}_{k,N,\epsilon}(\alpha)$ as

$$\mathcal{R}_{k,N,\epsilon}(\alpha) = \bigcup_{M>N} \bigcup_{\epsilon'<\epsilon} \{y \in Y : \text{ there exists some } n \in (N, M] \text{ such that } \text{dist}_y(P_k, T_\alpha^{-n} P_k) < \epsilon'\}.$$

Since $\nu(\mathcal{R}_{k,N,\epsilon}) > 1 - \eta$, it follows that there exist $M > N$, $\epsilon' < \epsilon$, and $\eta' < \eta$ such that

$$\nu(\mathcal{R}') := \nu\{y \in Y : \text{ there exists some } n \in (N, M] \text{ such that } \text{dist}_y(P_k, T_\alpha^{-n} P_k) < \epsilon'\} = 1 - \eta'.$$

Let $\sigma > 0$ be a parameter that is so small that $\sigma < \epsilon - \epsilon'$ and $(M - N)\sigma < \eta - \eta'$. Let $\mathcal{O}$ be an open neighborhood of α that is so small that for any $\beta \in \mathcal{O}$, we have

$$\text{dist}_\mu(T_\alpha^{-n} P_k, T_\beta^{-n} P_k) \ < \ \sigma^2$$

for all $n \in (N, M]$. This is possible because

- for every n, the map $T_\beta \mapsto T_\beta^n$ is a continuous map from $\text{Aut}(Y \times I, \nu \times m)$ to itself because $\text{Aut}(Y \times I, \nu \times m)$ is a topological group, and

- for any fixed $\gamma \in \mathrm{Co}(Y)$, the map $\beta \mapsto \mathrm{dist}_\mu(T_\gamma^{-1} P_k, T_\beta^{-1} P_k)$ is a continuous map $\mathrm{Co}(Y) \to [0,1]$ by definition of the weak topology on $\mathrm{Aut}(Y \times I, \nu \times m)$.

We show that any $\beta \in \mathcal{O}$ is also in $\mathcal{U}_{k,N,\epsilon,\eta}$.

Because $\mathrm{dist}_\mu(P,Q) = \int \mathrm{dist}_y(P,Q)\, d\nu(y)$, we apply Markov's inequality to conclude that for each $n \in (N, M]$, there is a set of measure $> 1 - \sigma$ on which $\mathrm{dist}_y(T_\alpha^{-n} P_k, T_\beta^{-n} P_k) < \sigma$. Now define

$$\widetilde{Y} \ := \ \{y \in Y : \mathrm{dist}_y(T_\alpha^{-n} P_k, T_\beta^{-n} P_k) < \sigma \text{ for all } n \in (N, M]\}$$

and note that $\nu(\widetilde{Y}) > 1 - (M - N)\sigma$.

Now consider some $y \in \widetilde{Y} \cap \mathcal{R}'$. Because $y \in \mathcal{R}'$, there is $n \in (N, M]$ such that $\mathrm{dist}_y(P_k, T_\alpha^{-n} P_k) < \epsilon'$. Then, for that same n, we get the estimate

$$\mathrm{dist}_y(P_k, T_\beta^{-n} P_k) \ \leq \ \mathrm{dist}_y(P_k, T_\alpha^{-n} P_k) + \mathrm{dist}_y(T_\alpha^{-n} P_k, T_\beta^{-n} P_k) \ < \ \epsilon' + \sigma \ < \ \epsilon,$$

where the second inequality holds because $y \in \widetilde{Y}$.

This shows that for all $y \in \widetilde{Y} \cap \mathcal{R}'$, there exists an $n \in (N, M]$ such that $\mathrm{dist}_y(P_k, T_\beta^{-n} P_k) < \epsilon$, showing that $\mathcal{R}_{k,N,\epsilon}(\beta) \supseteq \widetilde{Y} \cap \mathcal{R}'$. Since $\nu(\widetilde{Y} \cap \mathcal{R}') > 1 - (M - N)\sigma - \eta' > 1 - \eta$, it follows that $\beta \in \mathcal{U}_{k,N,\epsilon,\eta}$ as desired. □

Proof of Theorem 5.5. Follows immediately from Lemmas 5.7 to 5.9 and the Baire category theorem. □

5.3 Relationship between rigidity and slow entropy

Throughout this section, let L denote the rate function $L(n) = \log n$. First we give a sufficient condition for an extension to be rigid.

Theorem 5.10. *Assume that* $\mathbf{Y}$ *is ergodic. Let* α *be a cocycle on* Y *and let* $\pi : \mathbf{X}_\alpha \to \mathbf{Y}$ *denote projection onto the first coordinate. Suppose that there exists a Følner sequence* (F_n) *for* $\mathbb{N}$ *such that* $h_{\mathrm{slow}}^{L,(F_n)}(\mathbf{X}_\alpha \,|\, \pi) = 0$*. Then* $\mathbf{X}_\alpha$ *is a rigid extension of* $\mathbf{Y}$*.*

Proof. For a partition P, $\epsilon > 0$, and $m \in \mathbb{N}$, define

$$\mathcal{R}_{P,\epsilon,m} \;=\; \left\{ y \in Y : \text{ there exists } k > m \text{ such that } \operatorname{dist}_y(P, T_\alpha^{-k} P) < 5\sqrt{\epsilon} \right\}.$$

The first step is to show that $\nu(\mathcal{R}_{P,\epsilon,m}) \geq 1 - 4\sqrt{\epsilon}$ for every P, ϵ, m.

Let (F_n) be the Følner sequence given by the hypothesis of Theorem 5.10. By the assumption that $h_{\mathrm{slow}}^{L,(F_n)}(\mathbf{X} \,|\, \pi) = 0$, for all n sufficiently large we have

$$|P|^{\operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi)} \;\leq\; \frac{\epsilon}{m} \cdot |F_n|. \tag{5.1}$$

Also, by the mean ergodic theorem, for all sufficiently large n we have

$$\nu \left\{ y \in Y : \frac{|\{t \in F_n : S^t y \in \mathcal{R}_{P,\epsilon,m}\}|}{|F_n|} \;<\; \nu(\mathcal{R}_{P,\epsilon,m}) + \epsilon \right\} \;\geq\; 1 - \epsilon. \tag{5.2}$$

So fix an n which is large enough so that (5.1) and (5.2) both hold.

Let

$$C \;=\; C(n) = \operatorname{cov}(\mu, P, F_n, \epsilon \,|\, \pi)$$

and let $Y_n \subseteq Y$ be the set of y satisfying $\operatorname{cov}(\mu_y, P, F_n, \epsilon) \leq C$. By definition we have $\nu(Y_n) \geq 1 - \epsilon$, so by (5.2), we may fix a point y that is an element of both Y_n and the set appearing in (5.2).

We now repeat the construction from the proof of Lemma 4.12, which we partially reproduce here for convenience. Let $B_1, \ldots, B_L$ be subsets of X such that each B_i has d_{P,F_n}-diameter at most ϵ and $\mu_y\left(\bigcup B_i\right) \geq 1 - \epsilon$. Let $X' = \bigcup B_i$. Without loss of generality, we may assume that the B_i are disjoint. For each i, fix a point $x_i \in B_i$. Then, for $x \in X'$, define $r(x)$ to be the unique x_i such that $x \in B_i$.

Define

$$A \;=\; \Big\{ s \in F_n : \mu_y\{x \in X' : P(T_\alpha^s x) = P(T_\alpha^s r(x))\} \geq 1 - \sqrt{2\epsilon} \Big\}.$$

In the proof of Lemma 4.12, we proved the estimate

$$|A| \;\geq\; |F_n|(1 - \sqrt{2\epsilon}).$$

Now decompose the set A as

$$A \;=\; \bigcup_{w \in \{0,1,\ldots,|P|-1\}^C} \{s \in A : (P(T_\alpha^s x_i))_{i=1}^L = w\} \;=:\; \bigcup_{w \in \{0,1,\ldots,|P|-1\}^C} A_w.$$

In the proof of Lemma 4.12, we also showed that

$$\mathrm{dist}_y(T_\alpha^{-s}P, T_\alpha^{-t}P) \;\leq\; 5\sqrt{\epsilon} \qquad \text{whenever } s, t \text{ lie in the same } A_w. \tag{5.3}$$

Using this decomposition of A into the sets A_w, we can show that $S^t y \in \mathcal{R}_{P,\epsilon,m}$ for most $t \in F_n$. By (5.3), we conclude that $\{t \in F_n : S^t y \in \mathcal{R}_{P,\epsilon,m}\}$ contains all of the elements of A, except for possibly the m largest elements of each A_w. Therefore, we can use (5.1) to estimate

$$\begin{aligned} \#\{t \in F_n : S^t y \in \mathcal{R}_{P,\epsilon,m}\} \;&\geq\; \sum_{w \in \{0,1,\ldots,|P|-1\}^C} (|A_w| - m) \;=\; |A| - m \cdot |P|^C \\ &\geq\; |F_n|(1 - \sqrt{2\epsilon}) - m \cdot \frac{\epsilon}{m} \cdot |F_n| \\ &\geq\; |F_n|(1 - 3\sqrt{\epsilon}). \end{aligned}$$

Combining this estimate with (5.2), we conclude that

$$\nu(\mathcal{R}_{P,\epsilon,m}) \;\geq\; 1 - 3\sqrt{\epsilon} - \epsilon \;\geq\; 1 - 4\sqrt{\epsilon}$$

as desired.

Now let

$$\mathcal{R}_{P,\epsilon} \;=\; \bigcap_{M\geq 1}\bigcup_{m\geq M} \mathcal{R}_{P,\epsilon,m} \;=\; \left\{y \in Y : \mathrm{dist}_y(P, T_\alpha^{-n}P) < 5\sqrt{\epsilon} \text{ for infinitely many } n\right\}.$$

Because each $\nu(\mathcal{R}_{P,\epsilon,m}) \geq 1 - 4\sqrt{\epsilon}$, we have $\nu(\mathcal{R}_{P,\epsilon}) \geq 1 - 4\sqrt{\epsilon}$ as well.

Finally, let P_k be the partition $\{Y \times E : E \in \mathcal{D}_k\}$, let $\epsilon_k = 1/k$, and let $\mathcal{R}_k = \mathcal{R}_{P_k,\epsilon_k}$. We have $\nu(\mathcal{R}_k) \geq 1 - 4/\sqrt{k}$. Now let $\overline{\mathcal{R}} = \bigcap_{K\geq 1}\bigcup_{k\geq K}\mathcal{R}_k$ and note that

$$\nu\left(\overline{\mathcal{R}}\right) \;=\; \lim_{K\to\infty} \nu\left(\bigcup_{k\geq K}\mathcal{R}_k\right) \;\geq\; \lim_{K\to\infty}\nu(\mathcal{R}_K) \;=\; 1.$$

We claim that every $y \in \overline{\mathcal{R}}$ has a rigidity sequence.

Fix $y \in \overline{\mathcal{R}}$. Then, by construction, there are infinitely many k that satisfy

$$\mathrm{dist}_y(P_k, T_\alpha^{-n}P_k) < 5/\sqrt{k} \quad \text{for infinitely many } n.$$

So, by repeating the diagonalization argument from Lemma 5.4, we again are able to conclude that y has a rigidity sequence. □

Corollary 5.11. *Isometric extensions are rigid.*

Proof. Suppose $\pi : \mathbf{X}_\alpha \to \mathbf{Y}$ is an isometric extension. Then by Theorem 4.4, for any Følner sequence and any rate function U, π has zero relative slow entropy. So in particular, there is a Følner sequence for which π has zero relative slow entropy with respect to the rate function $L(n) = \log n$. By Theorem 5.10, this implies π is rigid. □

Remark 5.12. In the non-relative setting, this result can be proven directly from the definitions using the fact that any orbit of a compact group rotation is dense in some closed subgroup. In the relative setting, it can proven in a similar but more complicated way by appealing to the theory of the Mackey group. It is interesting to note that we are able to provide another proof of this result using entropy methods.

In the non-relative setting, we are also able to prove a converse and obtain necessary and sufficient conditions for rigidity in terms of slow entropy. For this part, we use interchangeably the notations $f(m) \ll g(m)$ and $f(m) = o(g(m))$ to mean that $f(m)/g(m) \to 0$ as $m \to \infty$.

Theorem 5.13. *The following are equivalent.*

(1) $\mathbf{X}$ *is rigid.*

(2) For every rate function U*, there exists a Følner sequence* (F_n) *for* $\mathbb{N}$ *such that* $h_{\text{slow}}^{U,(F_n)}(\mathbf{X}) = 0$.

(3) There exists a Følner sequence (F_n) *for* $\mathbb{N}$ *such that* $h_{\text{slow}}^{L,(F_n)}(\mathbf{X}) = 0$.

Proof. The implication (2) $\implies$ (3) is trivial and the implication (3) $\implies$ (1) is the special case of Theorem 5.10 where $\mathbf{Y}$ is trivial, so we just need to show that (1) $\implies$ (2). Assume that $\mathbf{X}$ is rigid and let (n_k) be a rigidity sequence. Let P be a finite generating partition for $\mathbf{X}$. Such a partition must exist by Krieger's theorem [Kri70] because rigid systems have zero entropy. Applying the definition of rigidity to each of the finitely many cells of P, it follows that

$$\lim_{k\to\infty} \operatorname{dist}_\mu(T^{-n_k}P, P) \;=\; 0.$$

Now replace the rigidity sequence (n_k) with a sufficiently thin subsequence so that we may assume that

$$\operatorname{dist}_\mu(T^{-n_k}P, P) \;<\; 2^{-k}. \tag{5.4}$$

Also assume that the rigidity sequence is sparse enough so that $n_{k+1} - n_k > k$ for every k.

Now let U be an arbitrary rate function and assume without loss of generality that $U(m) \ll \exp(m)$. Let V be another rate function satisfying $V(m) \ll \log U(m)$. We define our Følner sequence (F_m) by the formula

$$F_m \ := \ [0, V(m)) \cup [n_1, n_1 + V(m)) \cup \cdots \cup [n_{m-1}, n_{m-1} + V(m)).$$

It's clear that this is a Følner sequence for $\mathbb{N}$.

It will be useful later to have a good estimate for $|F_m|$. Clearly $|F_m| \leq m \cdot V(m)$, but we can also show that it is not much smaller than this. Observe that

$$\begin{aligned} |F_m| \ &\geq \ V(m) \cdot \#\{k \leq m : [n_k, n_k + V(m)) \cap [n_{k+1}, n_{k+1} + V(m)) = \emptyset\} \\ &\geq \ V(m) \cdot \#\{k \leq m : n_{k+1} - n_k > V(m)\} \\ &\geq \ V(m) \cdot \#\{k \leq m : k > V(m)\} \\ &= \ V(m) \cdot \max(m - V(m), 0) \\ &= \ V(m) \cdot (m - o(m)) \end{aligned} \tag{5.5}$$

Our goal is to show that $h_{\text{slow}}^{U,(F_m)}(\mathbf{X}) = 0$. Since P is a generating partition, it suffices to show that $h_{\text{slow}}^{U,(F_m)}(\mathbf{X}, P) = 0$. To do this, let $\epsilon > 0$. We seek to estimate $\operatorname{cov}(\mu, P, F_m, \epsilon)$ for m sufficiently large. Let $C = \operatorname{cov}(\mu, P, [0, V(m)), \epsilon)$ and let $B_1, \ldots, B_C$ be subsets of X satisfying $\mu\left(\bigcup B_i\right) \geq 1 - \epsilon$ and

$$\operatorname{diam}_{P,[0,V(m))}(B_i) \ \leq \ \epsilon.$$

We now show that we can restrict the B_i to a large subset of X such that after the restsriction, $\operatorname{diam}_{P,F_m}(B_i)$ is also small.

Let $k_0 = k_0(m)$ be the smallest integer that satisfies

$$\sum_{k \geq k_0} \mathrm{dist}_\mu(T^{-n_k} P, P) \ < \ \frac{\epsilon}{V(m)}.$$

Because of the condition that $\mathrm{dist}_\mu(T^{-n_k} P, P) < 2^{-k}$, it follows that

$$k_0(m) \ \leq \ \log_2 \left(\frac{V(m)}{\epsilon} \right) \ \ll \ \log \log U(m) \ \ll \ m. \tag{5.6}$$

For $0 \leq i < V(m)$, define the "good sets"

$$\mathcal{G}_i \ := \ \{ x \in X : P(T^i x) = P(T^{n_k + i} x) \text{ for all } k \geq k_0 \} \quad \text{and} \tag{5.7}$$

$$\mathcal{G} \ := \ \bigcap_{i=0}^{V(m)} \mathcal{G}_i. \tag{5.8}$$

By the definition of k_0 and the T-invariance of μ, we have

$$\mu(\mathcal{G}^c) \ \leq \ \sum_{i=0}^{V(m)-1} \mu(\mathcal{G}_i^c) \ \leq \ \sum_{i=0}^{V(m)-1} \sum_{k \geq k_0} \mathrm{dist}_\mu(T^{-(n_k+i)} P, T^{-i} P) \ \leq \ \epsilon. \tag{5.9}$$

Now replace each B_i by $B_i' = B_i \cap \mathcal{G}$, so we still have $\mu\left(\bigcup B_i'\right) \geq 1 - 2\epsilon$. It remains to show that each B_i' has small diameter according to d_{P,F_m}.

If $x, y \in B_i'$, then

$$\begin{aligned}
|F_m| \cdot d_{P,F_m}(x,y) \ &\leq \ \sum_{k=0}^{m-1} \sum_{i=0}^{V(m)-1} 1_{P(T^{n_k+i}x) \neq P(T^{n_k+i}y)} \\
&\leq \ k_0 \cdot V(m) + \sum_{k=k_0}^{m-1} \sum_{i=0}^{V(m)-1} 1_{P(T^{n_k+i}x) \neq P(T^{n_k+i}y)} \\
&= \ k_0 \cdot V(m) + (m - k_0) \cdot \sum_{i=0}^{V(m)-1} 1_{P(T^i x) \neq P(T^i y)} \\
&\leq \ k_0 \cdot V(m) + m \cdot V(m) \cdot d_{P,[0,V(m))}(x,y).
\end{aligned}$$

Therefore, by (5.5) and (5.6), we have

$$\begin{aligned}\operatorname{diam}_{P,F_m}(B_i') &\le \frac{k_0 \cdot V(m)}{|F_m|} + \frac{m \cdot V(m) \cdot \operatorname{diam}_{P,[0,V(m))}(B_i')}{|F_m|} \\ &\le \frac{o(m)}{m - o(m)} + \frac{\epsilon \cdot m}{m - o(m)} \\ &\le 3\epsilon\end{aligned}$$

for sufficiently large m.

Thus we have shown that

$$\operatorname{cov}(\mu, P, F_m, 3\epsilon) \le \operatorname{cov}(\mu, P, [0, V(m)), \epsilon) \le |P|^{V(m)} \ll U(m) \ll U(|F_m|)$$

for any $\epsilon > 0$, and the desired conclusion follows. □

Remark 5.14. The part of this proof that breaks down in the relative setting is the estimate (5.9). Here we have used the T-invariance of μ critically to deduce that if the partitions P and $T^{-n_k}P$ are close with respect to μ, then so are $T^{-i}P$ and $T^{-(n_k+i)}P$. In the relative setting this breaks down because if P and $T_\alpha^{-n_k}P$ are close with respect to μ_y, then $T_\alpha^{-i}P$ and $T_\alpha^{-(n_k+i)}P$ are only close with respect to $\mu_{S^i y}$.

Remark 5.15. In [Ada21, Theorem 1], the author shows that for the Følner sequence $F_n = [0, n)$ and any sub-exponential rate function U, there is a dense G_δ set of systems $\mathbf{I} = ([0, 1], T, m)$ that are both rigid and satisfy $h_{\text{slow}}^{U,(F_n)}(\mathbf{I}) = \infty$. Combined with Theorem 5.13, this shows that generically, the slow entropy of a system depends quite strongly on the choice of Følner sequence. This is in contrast with Kolmogorov–Sinai entropy, which is independent of the choice of Følner sequence.

As a corollary of Theorem 5.13, we get a similar condition that characterizes mild mixing systems in terms of slow entropy. Recall that a system is said to be **mildly mixing** if it has no nontrivial rigid factors [FW78].

Corollary 5.16. *The system* $\mathbf{X}$ *is mildly mixing if and only if for all partitions* P *of* X *and all Følner sequences* (F_n) *for* $\mathbb{N}$*, we have* $h_{\text{slow}}^{L,(F_n)}(\mathbf{X}, P) > 0$.

Proof. Suppose $\mathbf{X}$ is not mildly mixing. Then there is a nontrivial rigid system $\mathbf{Y}$ and a factor map $\pi : \mathbf{X} \to \mathbf{Y}$. Let Q be any partition of Y and let $P = \pi^{-1}Q$. The rigidity of $\mathbf{Y}$ implies that we can find a Følner sequence (F_n) such that $h_{\text{slow}}^{L,(F_n)}(\mathbf{Y}, Q) = 0$. Then, using the definition of factor map, it immediately follows that $h_{\text{slow}}^{L,(F_n)}(\mathbf{X}, P) = 0$ as well.

Conversely, suppose that there exist a partition P and a Følner sequence (F_n) so that

$$h_{\text{slow}}^{L,(F_n)}(\mathbf{X}, P) \;=\; 0.$$

Then consider the factor $\mathbf{Y}$ corresponding to the T-invariant σ-algebra $\bigvee_{n\in\mathbb{Z}} T^{-n}P$. Because P is a generating partition for this factor, it follows that $h_{\text{slow}}^{L,(F_n)}(\mathbf{Y}) = 0$, which implies that $\mathbf{Y}$ is rigid, so $\mathbf{X}$ is not mildly mixing. □

5.4 A necessary condition for rigidity

In light of Theorem 5.5 and our failure to prove the converse of Theorem 5.10, one may wonder whether or not *every* extension is rigid. In this section, we show that this is not the case by exhibiting a natural non-empty class of extensions that can not be rigid.

Definition 5.17. Given a system $\mathbf{Y}$, a cocycle α on Y is said to be **strongly mixing** if for ν-a.e. y, we have

$$m\left(E \cap \alpha_n(y)^{-1}E\right) \;\to\; m(E)^2$$

for all measurable $E \subseteq I$. Also, we will say that a skew product extension $\mathbf{X}_\alpha \to \mathbf{Y}$ is a **strongly mixing extension** if α is a strongly mixing cocycle.

This definition is similar to the definition of strongly mixing extension given in [Sch18], but here we have phrased it to be more analogous to our definition of rigidity. It is unknown whether or not the definitions of rigidity and strong mixing presented here are equivalent to the definitions given in [Sch18].

Proposition 5.18. *The set of rigid cocycles and the set of strongly mixing cocycles are disjoint.*

Proof. Let $E \subseteq I$ be any subset of measure $1/2$. If α is a strongly mixing cocycle, then for a.e. y,

$$m\left(E \cap \alpha_n(y)^{-1}E\right) \ \to \ 1/4$$

as $n \to \infty$. But if α were also a rigid cocycle, then there would have to be a subsequence (n_k) along which

$$m\left(E \cap \alpha_{n_k}(y)^{-1}E\right) \ \to \ 1/2$$

as $k \to \infty$, a contradiction. □

Finally, let us remark that the set of strongly mixing cocycles is non-empty. Indeed, if $\mathbf{I} = (I, m, \alpha_0)$ is any strongly mixing system and α is the constant cocycle $\alpha(y) = \alpha_0$, then α is clearly a strongly mixing cocycle. These cocycles correspond to direct product transformations on $Y \times I$ where the transformation in the I coordinate is strongly mixing.

5.5 Open questions

In this section, we outline some unanswered questions about our new notion of rigid extension. First, in light of Theorems 5.10 and 5.13, the most obvious question to ask is

Question 5.19. Does the converse of Theorem 5.10 hold?

Next, as was briefly mentioned earlier, we do not know if this notion is isomorphism invariant. More specifically, recall that the cocycle given by Rokhlin's skew product is unique up to *cohomology* – two cocycles $\alpha, \beta \in \mathrm{Co}(Y)$ are said to be *cohomologous* if there is another map $u : Y \to \mathrm{Aut}(I)$ such that $\beta(y) = u(Sy)^{-1} \circ \alpha(y) \circ u(y)$ for ν-a.e. y. This leads to the following natural question.

Question 5.20. If $\alpha \in \mathrm{Co}(Y)$ is a rigid cocycle and β is cohomologous to α, is β also a rigid cocycle?

What makes this question difficult? If α is a rigid cocycle, then for almost every y, $\alpha_n(y)$ is close to the identity for infinitely many n. So, if β is cohomologous to α, then

$$\beta_n(y) \;=\; u(S^n y)^{-1} \circ \alpha_n(y) \circ u(y) \;\approx\; u(S^n y)^{-1} \circ u(y)$$

for all such n. But because u is arbitrary, there is no guarantee that $u(S^n y)$ will be close to $u(y)$, so we can not conclude anything about $\beta_n(y)$ being close to the identity.

An affirmative answer to this question would allow rigidity to be well defined for an abstract extension $\pi : \mathbf{X} \to \mathbf{Y}$ by using *any* cocyle given by Rokhlin's theorem. Also note that if we were able to prove the converse of Theorem 5.10, then the answer to this question would be "yes" by the isomorphism invariance of relative slow entropy.

One may also ask about explicit constructions. In [Ada21], Adams uses cutting and stacking procedures to construct an explicit example of a transformations that is weakly mixing, rigid, and has large slow entropy. One can ask about relative versions of this construction.

Question 5.21. Given a base system $\mathbf{Y}$, a sub-exponential rate function U, and the Følner sequence $F_n = [0, n)$, is it possible to give an explicit construction of an extension $\pi : \mathbf{X} \to \mathbf{Y}$ which is weakly mixing, rigid, and satisfies $h_{\text{slow}}^{U,(F_n)}(\mathbf{X} \mid \pi) = \infty$?

REFERENCES

[Ada21] Terrence Adams. "Genericity and rigidity for slow entropy transformations." *New York J. Math.*, **27**:393–416, 2021.

[AGT21] Tim Austin, Eli Glasner, Jean-Paul Thouvenot, and Benjamin Weiss. "An ergodic system is dominant exactly when it has positive entropy." arXiv:2112.03800, 2021.

[AR62] L. M. Abramov and V. A. Rohlin. "Entropy of a skew product of mappings with invariant measure." *Vestnik Leningrad. Univ.*, **17**(7):5–13, 1962.

[Aus18] Tim Austin. "Measure concentration and the weak Pinsker property." *Publ. Math. Inst. Hautes Études Sci.*, **128**:1–119, 2018.

[Bil95] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[Blu95] Frank Blume. *The rate of entropy convergence*. ProQuest LLC, Ann Arbor, MI, 1995. Thesis (Ph.D.)–The University of North Carolina at Chapel Hill.

[Blu97] Frank Blume. "Possible rates of entropy convergence." *Ergodic Theory Dynam. Systems*, **17**(1):45–70, 1997.

[Blu98] Frank Blume. "Minimal rates of entropy convergence for completely ergodic systems." *Israel J. Math.*, **108**:1–12, 1998.

[Car97] Maria de Carvalho. "Entropy dimension of dynamical systems." *Portugal. Math.*, **54**(1):19–40, 1997.

[CFW81] Alain Connes, Jacob Feldman, and Benjamin Weiss. "An amenable equivalence relation is generated by a single transformation." *Ergodic Theory Dynam. Systems*, **1**(4):431–450, 1981.

[DHP19] Dou Dou, Wen Huang, and Kyewon Koh Park. "Entropy dimension of measure preserving systems." *Trans. Amer. Math. Soc.*, **371**(10):7029–7065, 2019.

[ELW21] Manfred Einsiedler, Elon Lindenstrauss, and Thomas Ward. *Entropy in ergodic theory and topological dynamics*. Book draft, available online at `https://tbward0.wixsite.com/books/entropy`, 2021.

[Fer97] Sébastien Ferenczi. "Measure-theoretic complexity of ergodic systems." *Israel J. Math.*, **100**:189–207, 1997.

[FP07] Sébastien Ferenczi and Kyewon Koh Park. "Entropy dimensions and a class of constructive examples." *Discrete Contin. Dyn. Syst.*, **17**(1):133–141, 2007.

[Fre06] D. H. Fremlin. *Measure theory. Vol. 4.* Torres Fremlin, Colchester, 2006. Topological measure spaces. Part I, II, Corrected second printing of the 2003 original.

[Fri83] David Fried. "Entropy for smooth abelian actions." *Proc. Amer. Math. Soc.*, **87**(1):111–116, 1983.

[Fur77] Harry Furstenberg. "Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions." *J. Analyse Math.*, **31**:204–256, 1977.

[FW78] Hillel Furstenberg and Benjamin Weiss. "The finite multipliers of infinite ergodic transformations." In *The structure of attractors in dynamical systems (Proc. Conf., North Dakota State Univ., Fargo, N.D., 1977)*, volume 668 of *Lecture Notes in Math.*, pp. 127–132. Springer, Berlin, 1978.

[Gla03] Eli Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.

[Gra11] Robert M. Gray. *Entropy and information theory*. Springer, New York, second edition, 2011.

[GTW21] E. Glasner, J.-P. Thouvenot, and B. Weiss. "On some generic classes of ergodic measure preserving transformations." *Trans. Moscow Math. Soc.*, **82**:15–36, 2021.

[GW19] Eli Glasner and Benjamin Weiss. "Relative weak mixing is generic." *Sci. China Math.*, **62**(1):69–72, 2019.

[Hal56] Paul R. Halmos. *Lectures on ergodic theory*, volume 3 of *Publications of the Mathematical Society of Japan*. Mathematical Society of Japan, Tokyo, 1956.

[Hul82] Paul Hulse. "Sequence entropy and subsequence generators." *J. London Math. Soc. (2)*, **26**(3):441–450, 1982.

[Kat80] A. Katok. "Lyapunov exponents, entropy and periodic orbits for diffeomorphisms." *Inst. Hautes Études Sci. Publ. Math.*, (51):137–173, 1980.

[Kec10] Alexander S. Kechris. *Global aspects of ergodic group actions*, volume 160 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2010.

[Kie75] J. C. Kieffer. "A generalized Shannon-McMillan theorem for the action of an amenable group on a probability space." *Ann. Probability*, **3**(6):1031–1037, 1975.

[KKR14] Anatole Katok, Svetlana Katok, and Federico Rodriguez Hertz. "The Fried average entropy and slow entropy for actions of higher rank abelian groups." *Geom. Funct. Anal.*, **24**(4):1204–1228, 2014.

[KL16] David Kerr and Hanfeng Li. *Ergodic theory*. Springer Monographs in Mathematics. Springer, Cham, 2016. Independence and dichotomies.

[Kol58] A. N. Kolmogorov. "A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces." *Dokl. Akad. Nauk SSSR (N.S.)*, **119**:861–864, 1958.

[Kol59] A. N. Kolmogorov. "Entropy per unit time as a metric invariant of automorphisms." *Dokl. Akad. Nauk SSSR*, **124**:754–755, 1959.

[Kri70] Wolfgang Krieger. "On entropy and generators of measure-preserving transformations." *Trans. Amer. Math. Soc.*, **149**:453–464, 1970.

[KT97] Anatole Katok and Jean-Paul Thouvenot. "Slow entropy type invariants and smooth realization of commuting measure-preserving transformations." *Ann. Inst. H. Poincaré Probab. Statist.*, **33**(3):323–338, 1997.

[Ku67] A. G. Kušnirenko. "Metric invariants of entropy type." *Uspehi Mat. Nauk*, **22**(5 (137)):57–65, 1967.

[KW72] Yitzhak Katznelson and Benjamin Weiss. "Commuting measure-preserving transformations." *Israel J. Math.*, **12**:161–173, 1972.

[Lot22a] Adam Lott. "Relative slow entropy." arXiv:2210.05054, 2022.

[Lot22b] Adam Lott. "Zero entropy actions of amenable groups are not dominant." arXiv:2204.11459, 2022.

[Mou85] Jean Moulin Ollagnier. *Ergodic theory and statistical mechanics*, volume 1115 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985.

[Orn70] Donald Ornstein. "Bernoulli shifts with the same entropy are isomorphic." *Advances in Math.*, **4**:337–352, 1970.

[OW80] Donald S. Ornstein and Benjamin Weiss. "Ergodic theory of amenable group actions. I. The Rohlin lemma." *Bull. Amer. Math. Soc. (N.S.)*, **2**(1):161–164, 1980.

[Pic69] B. S. Pickel' "Certain properties of A-entropy." *Mat. Zametki*, **5**:327–334, 1969.

[Rat81] Marina Ratner. "Some invariants of Kakutani equivalence." *Israel J. Math.*, **38**(3):231–240, 1981.

[RW00] Daniel J. Rudolph and Benjamin Weiss. "Entropy and mixing for amenable group actions." *Ann. of Math. (2)*, **151**(3):1119–1150, 2000.

[Sal77] Alan Saleski. "Sequence entropy and mixing." *J. Math. Anal. Appl.*, **60**(1):58–66, 1977.

[Sch18] Mike Schnurr. "A note on strongly mixing extensions." arXiv:1712.06192, 2018.

[Sin59a] Ja. Sinaĭ. "Flows with finite entropy." *Dokl. Akad. Nauk SSSR*, **125**:1200–1202, 1959.

[Sin59b] Ja. Sinaĭ. "On the concept of entropy for a dynamic system." *Dokl. Akad. Nauk SSSR*, **124**:768–771, 1959.

[Tho75] Jean-Paul Thouvenot. "Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli." *Israel J. Math.*, **21**(2-3):177–207, 1975.

[Ver18] Roman Vershynin. *High-dimensional probability*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2018. An introduction with applications in data science, With a foreword by Sara van de Geer.

[WZ92] Thomas Ward and Qing Zhang. "The Abramov-Rokhlin entropy addition formula for amenable group actions." *Monatsh. Math.*, **114**(3-4):317–329, 1992.

[Zha92] Qing Zhang. "Sequence entropy and mild mixing." *Canad. J. Math.*, **44**(1):215–224, 1992.

[Zor] Pavel Zorin-Kranich. "Compact extensions are isometric." https://www.math.uni-bonn.de/∼pzorin/notes/compact-isometric.pdf.

[ZP15] Yun Zhao and Yakov Pesin. "Scaled entropy for dynamical systems." *J. Stat. Phys.*, **158**(2):447–475, 2015.

[ZP16] Yun Zhao and Yakov Pesin. "Erratum to: Scaled entropy for dynamical systems [MR3299885]." *J. Stat. Phys.*, **162**(6):1654–1660, 2016.