Approximation and classification in the ergodic theory of nonamenable groups

Thesis by
Peter Burton

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2017
Defended May 15, 2017

Peter Burton
ORCID: 0000-0002-0348-8385

ACKNOWLEDGEMENTS

I thank my advisor, Alexander Kechris, for his guidance during my time in graduate school. I thank my coauthors for their contributions to the work presented herein.

I would also like to thank my family and friends for helping me become who I am today.

This thesis is dedicated to the State of California.

ABSTRACT

This thesis is a contribution to the theory of measurable actions of discrete groups on standard probability spaces. The focus is on nonamenable acting groups. It is organized into two parts. The first part deals with a notion called weak equivalence, which describes a sense in which such actions can approximate each other. The second part deals with the concept of entropy for measure preserving actions of sofic groups.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Tim Austin and Peter Burton. Uniform mixing and completely positive sofic entropy. *To appear in Journal d'Analyse Mathematique*, 2016. Available at http://arxiv.org/abs/1603.09026.
Both authors contributed equally to this project.

[2] Peter Burton. Topology and convexity in the space of actions modulo weak equivalence. *To appear in Ergodic Theory and Dynamical Systems*, 2015. Available at http://arxiv.org/abs/1501.04079.
This paper is entirely my own work.

[3] Peter Burton. Naive entropy of dynamical systems. *To appear in Israel Journal of Mathematics*, 2015. Available at http://arxiv.org/abs/1503.06360.
This paper is entirely my own work.

[4] Peter Burton. A topological semigroup structure on the space of actions modulo weak equivalence. Preprint. Available at http://arxiv.org/abs/1501.04373, 2015.
This paper is entirely my own work.

[5] Peter Burton and Alexander S. Kechris. Invariant random subgroups and action versus representation maximality. *To appear in Proceedings of the American Mathematical Society*, 2016. Available at http://arxiv.org/abs/1608.01331.
Both authors contributed equally to this project.

[6] Peter Burton, Martino Lupini, and Omer Tamuz. Weak equivalence of stationary actions and the entropy realization problem. Submitted. Available at http://arxiv.org/abs/1603.05013, 2016.
All authors contributed equally to this project.

Chapter 0

INTRODUCTION

We introduce the various topics in the thesis in the order they appear.

0.1 Part I: Weak containment of measure preserving group actions

Fix a standard probability space (X, μ). We will denote by $\mathrm{Aut}(X, \mu)$ the group of all measure preserving transformations of (X, μ). In [45], P.R. Halmos defined two topologies on this group, called the weak topology and the uniform topology. The uniform topology strictly refines the weak topology. With these topologies the space $\mathrm{Aut}(X, \mu)$ provides a framework to develop a global theory of $\mathbb{Z}$-systems, allowing one to formulate questions about approximation, genericity, and classification. Similarly, one can consider the space $\mathrm{A}(G, X, \mu)$ of measure preserving actions of an arbitrary countable discrete group G. This space carries two topologies corresponding to the two topologies on $\mathrm{Aut}(X, \mu)$, and thus opens the door to analyzing global aspects of G-systems. It turns out that there is a rich interplay between properties of G and the structure of $\mathrm{A}(G, X, \mu)$. This is the subject of the book [53].

An important tool in analyzing $\mathbb{Z}$-systems is the Rokhlin Lemma, which asserts that for any $\mathbb{Z}$-system (X, μ, T) and any $\epsilon > 0$ there exists $n \in \mathbb{N}$ and a measurable set $A \subseteq X$ such that the shifted sets $A, TA, T^2A, \ldots, T^nA$ are pairwise disjoint and $\mu(A \cup TA \cup \cdots \cup T^nA) > 1 - \epsilon$. More abstractly, this asserts that any measure preserving transformation can be approximated by periodic transformations arbitrary well in the uniform topology on $\mathrm{Aut}(X, \mu)$. Since it is easy to see that any two periodic transformations with the same period are conjugate, it follows that the conjugacy class of any aperiodic transformation is dense in the set of aperiodic transformations with respect to the uniform topology.

For any G, the group $\mathrm{Aut}(X, \mu)$ acts on $\mathrm{A}(G, X, \mu)$ by conjugation: for $T \in \mathrm{Aut}(X, \mu)$ let $(T \cdot \mathbf{a})(g) = T\mathbf{a}(g)T^{-1}$, where $\mathbf{a}(g)$ is the transformation corresponding to $g \in G$ under the action $\mathbf{a} \in \mathrm{A}(G, X, \mu)$. We refer to the orbits of this action as conjugacy classes. Analogs of the Rokhlin Lemma were developed for actions of amenable groups by Ornstein and Weiss in [67] and in [36] M. Foreman and Weiss used them to show that when G is amenable the conjugacy class of every free action is dense in the set of free actions with respect to the uniform topology on $\mathrm{A}(G, X, \mu)$. Moreover, by the results of [55] this condition characterizes amenability.

In [53] Kechris defined a notion of 'weak containment' for measure preserving G-systems: an action $\mathbf{a}$ is said to be weakly contained in an action $\mathbf{b}$ if $\mathbf{a}$ lies in the weak closure of the conjugacy class of $\mathbf{b}$. We denote this by $\mathbf{a} \preceq \mathbf{b}$. If $\mathbf{a} \preceq \mathbf{b}$ and $\mathbf{b} \preceq \mathbf{a}$ we say that $\mathbf{a}$ is weakly equivalent to $\mathbf{b}$. Then the Rokhlin Lemma can be interpreted as saying that all free actions of an amenable group are weakly equivalent.

Chapter 1: Invariant random subgroups and action versus representation maximality

Associated to a measure-preserving action $\mathbf{a} \in \mathrm{A}(G, X, \mu)$, one has the Koopman representation $\kappa^{\mathbf{a}}$ of G on $L^2(X, \mu)$. It is more natural to consider the restriction $\kappa_0^{\mathbf{a}}$ of this representation to the orthogonal complement of the constant functions. There is a notion of weak containment for unitary representations analagous to weak containment of actions (see Appendix F of [8] for the definition). It is obvious that weak containment of actions implies weak containment of the corresponding Koopman representations, and it is not too hard to construct examples where the converse fails. However, these easy counterexamples come from non-ergodic actions and it remained an open problem to find ergodic examples. In this chapter, we prove the following theorem, showing in a strong way that weak containment of free ergodic actions is different from weak containment of the corresponding Koopman representations in the case of $\mathbf{F}_\infty$, the free group on infinitely many generators.

Theorem 0.1.1 (Burton-Kechris, [24])**.** *There exists a free ergodic action* $\mathbf{a}$ *of* $\mathbf{F}_\infty$ *which is not maximal in the order of weak containment of actions such that the corresponding Koopman representation* $\kappa_0^{\mathbf{a}}$ *is maximal in the order of weak containment of representations.*

The proof of Theorem 0.1.1 is probabilistic, relying on the construction of a particular invariant random subgroup of $\mathbf{F}_\infty$. If G is a discrete group, an invariant random subgroup (IRS) of G is a conjugation-invariant probability measure on the space of subgroups of G. The notion of an IRS was introduced by Abért, Y. Glasner and B. Virág in [4] as a stochastic generalization of normal subgroups. The IRS we build to prove Theorem 0.1.1 is supported on the subgroups H of G such that the corresponding generalized Bernoulli shift action of G on $[0, 1]^{G/H}$ is maximal in the order of weak containment (for both actions and representations).

This chapter is joint work with with Alexander Kechris.

Chapter 2: Topology and convexity in the space of actions modulo weak equivalence

In this chapter, we analyse the structure of the quotient of the space of actions by the relation of weak equivalence. In [3] M. Abért and G. Elek introduced a compact Polish topology on the set of weak equivalence classes of G-systems. We will denote this space by $\mathrm{A}_{\sim}(G, X, \mu)$. Freeness is an invariant of weak equivalence, and we denote the subspace of free weak equivalences classes by $\mathrm{FR}_{\sim}(G, X, \mu)$. Thus $\mathrm{FR}_{\sim}(G, X, \mu)$ represents the extent to which the Rokhlin Lemma fails for G, and studying its structure provides an approach to understanding measure preserving actions of nonamenable groups. In [74], R.D. Tucker-Drob introduced a slightly modified notion called 'stable weak containment' which avoids certain minor technical pathologies of weak containment. We denote the space of stable weak equivalence classes by $\mathrm{A}^{s}_{\sim}(G, X, \mu)$ and the subspace of free stable weak equivalence classes by $\mathrm{FR}^{s}_{\sim}(G, X, \mu)$.

$\mathrm{A}_{\sim}(\Gamma, X, \mu)$ carries a natural operation of convex combination. We introduce a variant of an abstract construction of Fritz which encapsulates the convex combination operation on $\mathrm{A}_{\sim}(\Gamma, X, \mu)$. This formalism allows us to define the geometric notion of an extreme point. We also discuss a topology on $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ due to Abert and Elek in which it is Polish and compact, and show that this topology is equivalent others defined in the literature. We show that the convex structure of $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ is compatible with the topology, and as a consequence deduce that $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ is path connected. Using ideas of Tucker-Drob we are able to give a complete description of the topological and convex structure of $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ for amenable Γ by identifying it with the simplex of invariant random subgroups. In particular we conclude that $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ can be represented as a compact convex subset of a Banach space if and only if Γ is amenable. In the case of general Γ we prove a Krein-Milman type theorem asserting that finite convex combinations of the extreme points of $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ are dense in this space. We also consider the space $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ of stable weak equivalence classes and show that it can always be represented as a compact convex subset of a Banach space. In the case of a free group $\mathbb{F}_N$, we show that if one restricts to the compact convex set $\mathrm{FR}_{\sim_s}(\mathbb{F}_N, X, \mu) \subseteq \mathrm{A}_{\sim_s}(\mathbb{F}_N, X, \mu)$ consisting of the stable weak equivalence classes of free actions, then the extreme points are dense in $\mathrm{FR}_{\sim_s}(\mathbb{F}_N, X, \mu)$.

Chapter 3: A topological semigroup structure on the space of actions modulo weak equivalence.

In this chapter, we introduce a topology on the space of actions modulo weak equivalence finer than the one previously studied in the literature. We show that the product of actions is a continuous operation with respect to this topology, so that the space of actions modulo weak equivalence becomes a topological semigroup.

Chapter 4: Weak equivalence of stationary actions and the entropy realization problem

In this chapter, we introduce the notion of weak containment for stationary actions of a countable group and define a natural topology on the space of weak equivalence classes. We prove that Furstenberg entropy is an invariant of weak equivalence, and moreover that it descends to a continuous function on the space of weak equivalence classes.

This chapter is joint work with Martino Lupini and Omer Tamuz.

0.2 Part II: sofic entropy

In this part, we study entropy theory for actions of nonamenable groups. Given a standard probability space (K, κ), the Bernoulli shift of G over the base space (K, κ) is the action of G on (K^G, κ^G) given by shifting indices. If K is countable we may refer to it as the 'alphabet'. One of the first major problems in ergodic theory was to determine whether the Bernoulli shift of $\mathbb{Z}$ over a two-point space with uniform measure is isomorphic to the Bernoulli shift of $\mathbb{Z}$ over a three-point space with uniform measure. This problem was answered in the negative by A.N. Kolmogorov in [64] through the introduction of an isomorphism invariant for $\mathbb{Z}$-systems with a finite generating partition. Known as entropy, the invariant was extended to arbitrary $\mathbb{Z}$-systems by Y.G. Sinaĭ in [73]. The entropy of the shift of $\mathbb{Z}$ over two points is $\log 2$, while the entropy of the shift of $\mathbb{Z}$ over three points is $\log 3$. As such these systems are not isomorphic. In the subsequent decades the entropy theory of $\mathbb{Z}$-systems developed into a vast panoply of mathematics, with the outstanding achievement being D.S. Ornstein's proof in [66] that two Bernoulli shifts of $\mathbb{Z}$ are isomorphic if and only if their base spaces have the same Shannon entropy. For an account of the entropy theory of $\mathbb{Z}$-systems, see [33] or Part 2 of [42].

The core of entropy theory was extended from $\mathbb{Z}$-systems to actions of amenable groups by Ornstein and B. Weiss in [67]. However, their approach is based on taking limits over Følner sequences, and so until recently it was not at all clear how to define entropy for more actions of general groups. In his groundbreaking papers [14] and [22], L. Bowen introduced a notion of entropy for G-systems when G is a free group, and then a family of entropy notions for G-systems when G is a sofic group and the G-system admits a finite generating partition. He used these invariants to prove that Bernoulli shifts of sofic groups over finite alphabets are classified by the Shannon entropy of their base. The class of sofic groups includes all amenable groups and all residually finite groups. It is a major open problem to determine whether every group is sofic. Informally, a group G is sofic if it admits sequence of approximate actions on finite sets which are, asymptotically, good replicas of the translation action of G on itself. Such a sequence is called a 'sofic approximation'. A precise statement appears as Definition 1 in [14]. The book [26] and the survey [68] provide more abstract perspectives.

Bowen's sofic entropy is constructed relative to a choice of a sofic approximation and while it is known that the entropy of a system can depend on this choice, the extent and nature of this dependence is poorly understood. However, in [17] Bowen showed that when G is amenable, sofic entropy relative to any approximation always agrees with classical Komogorov-Sinaĭ entropy. Sofic entropy was defined for arbitrary G-systems by D. Kerr in [58]. In [61] Kerr and H. Li defined topological entropy for actions of sofic groups by homeomorphisms of compact metric spaces and proved a variational principle relating it to measure-theoretic entropy. The theory of sofic entropy has proved very fruitful, with papers on the subject including [6], [7], [15], [16], [17], [18], [20], [27], [39], [48], [47], [49], [50], [51], [59], [60], [62], [63] and [78]. The article [77] provides a survey of the area.

Chapter 5: Naive entropy

In this chapter, we study an invariant of dynamical systems called naive entropy, which is defined for both measurable and topological actions of any countable group. We focus on nonamenable groups, in which case the invariant is two-valued, with every system having naive entropy either zero or infinity. Bowen has conjectured that when the acting group is sofic, zero naive entropy implies sofic entropy at most zero for both types of systems. We prove the topological version of this conjecture by showing that for every action of a sofic group by homeomorphisms of a compact metric space, zero naive entropy implies sofic entropy at most zero. This result and the simple definition of naive entropy allow us to show that the generic action of a free group on the Cantor set has sofic entropy at most zero. We observe that a distal Γ-system has zero naive entropy in both senses, if Γ has an element of infinite order. We also show that the naive entropy of a topological system is greater than or equal to the naive measure entropy of the same system with respect to any invariant measure.

Chapter 6: Uniform mixing and completely positive sofic entropy

Let G be a countable discrete sofic group. In this chapter, we define a concept of uniform mixing for measure-preserving G-actions and show that it implies completely positive sofic entropy. When G contains an element of infinite order, we use this to produce an uncountable family of pairwise nonisomorphic G-actions with completely positive sofic entropy. None of our examples is a factor of a Bernoulli shift.

This chapter is joint work with Tim Austin.

Part I

Weak equivalence of measurable group actions

Chapter 1

INVARIANT RANDOM SUBGROUPS AND ACTION VERSUS REPRESENTATION MAXIMALITY

Peter Burton and Alexander S. Kechris [1]

1.1 Introduction

Let G be a countably infinite group and (X, μ) a standard non-atomic probability space. We denote by $A(G, X, \mu)$ the space of measure preserving actions of G on (X, μ) with the weak topology. If $\mathbf{a}, \mathbf{b} \in A(G, X, \mu)$, we say that $\mathbf{a}$ is **weakly contained** in $\mathbf{b}$, in symbols $\mathbf{a} \preceq \mathbf{b}$, if $\mathbf{a}$ is in the closure of the set of isomorphic copies of $\mathbf{b}$ (i.e., it is in the closure of the orbit of $\mathbf{b}$ under the action of the automorphism group of (X, μ) on $A(G, X, \mu)$; see [53]). We say that $\mathbf{a} \in A(G, X, \mu)$ is **action-maximal** if for all $\mathbf{b} \in A(G, X, \mu)$ we have $\mathbf{b} \preceq \mathbf{a}$. Such $\mathbf{a}$ exist by a result of Glasner-Thouvenot-Weiss, Hjorth; see [53, Theorem 10.7]).

Now let H be a separable, infinite-dimensional Hilbert space and denote by $\mathrm{Rep}(G, H)$ the space of unitary representations of G on H with the weak topology (see [53, Appendix H]). For $\pi, \rho \in \mathrm{Rep}(G, H)$ we denote by $\pi \preceq \rho$ the usual relation of **weak containment** of representations (see [8], [53, Appendix H]). We say that $\pi \in \mathrm{Rep}(G, H)$ is **representation-maximal** if for all $\rho \in \mathrm{Rep}(G, H)$ we have $\rho \preceq \pi$. It is easy to check that such π exist.

For any action $\mathbf{a} \in A(G, X, \mu)$, let $\kappa^{\mathbf{a}}$ be the associated representation on $L^2(X, \mu)$, called the **Koopman representation**, and by $\kappa_0^{\mathbf{a}}$ its restriction to the orthogonal of the constant functions (see [53, page 66]). Then we have

$$\mathbf{a} \preceq \mathbf{b} \implies \kappa_0^{\mathbf{a}} \preceq \kappa_0^{\mathbf{b}}$$

but the converse fails; see [53, pages 66 and 68] and also [28, page 155] for examples. However in all these examples the actions $\mathbf{a}, \mathbf{b}$ were not both ergodic and this led to the following question.

[1] Research partially supported by NSF Grant DMS-1464475

Problem 1.1.1. *If* $\mathbf{a}, \mathbf{b} \in A(G, X, \mu)$ *are free, ergodic, does* $\kappa_0^{\mathbf{a}} \preceq \kappa_0^{\mathbf{b}}$ *imply* $\mathbf{a} \preceq \mathbf{b}$*?*

We provide a negative answer below. The proof is based on a result about invariant random subgroups of $G = \mathbf{F}_\infty$, the free group on a countably infinite set of generators, which might be of independent interest.

If I is a countable set and α is an action of a countable group G on I, we will write $\mathbf{s}_\alpha$ for the corresponding **generalized shift action** on 2^I with the usual product measure, given by $(\mathbf{s}_\alpha(g) \cdot f)(i) = f(\alpha(g)^{-1} \cdot i)$. If $I = G/H$, for some $H \leq G$, we will write $\tau_{G/H}$ for the left-translation action of G on G/H and $\mathbf{s}_{G/H}$ instead of $\mathbf{s}_{\tau_{G/H}}$. If H is trivial, we write $\mathbf{s}_G$ instead of $\mathbf{s}_{G/H}$.

We also let λ_α be the representation on $\ell^2(I)$ given by $(\lambda_\alpha(g) \cdot f)(i) = f(\alpha(g)^{-1} \cdot i)$. Note that $\lambda_{\tau_{G/H}}$ is the usual **quasi-regular representation** of G on $\ell^2(G/H)$, which we will denote by $\lambda_{G/H}$.

We call a subgroup $H \leq G$ with $[G : H] = \infty$ **action-maximal** if $\mathbf{s}_{G/H}$ is action-maximal and **representation-maximal** if $\lambda_{G/H}$ is representation-maximal. It was shown in [54] that there are H which are action-maximal and also H which are representation-maximal, for any non-abelian free group G.

An **invariant random subgroup (IRS)** of G is a probability Borel measure on $\mathrm{Sub}(G)$, the compact space of subgroups of G, which is invariant under the (continuous) action of G on $\mathrm{Sub}(G)$ by conjugation. Denote by $\mathcal{M}_G \subseteq \mathrm{Sub}(G)$ the set of all $H \leq G$ that are both action-maximal and representation-maximal. We show the following:

Theorem 1.1.1. *Let* $G = \mathbf{F}_\infty$*. Then there exists an IRS of* G *which is supported by* $\mathcal{M}_G$*.*

Using this and the result of Dudko-Grigorchuk [34, Proposition 8], we then prove the following:

Theorem 1.1.2. *Let* $G = \mathbf{F}_\infty$*. Then there exists a free, ergodic* $\mathbf{a} \in A(G, X, \mu)$ *such that* $\mathbf{a}$ *is not action-maximal but* $\kappa_0^{\mathbf{a}}$ *is representation-maximal.*

Let $\mathbf{a}$ be as in Theorem 1.1.2. Since $G = \mathbf{F}_\infty$ does not have property (T), the free, ergodic actions $\mathbf{b} \in A(G, X, \mu)$ are dense in $A(G, X, \mu)$ (see [53, Theorems 12.2 and 10.8]), so there is a free, ergodic $\mathbf{b} \in A(G, X, \mu)$ such that $\mathbf{b} \npreceq \mathbf{a}$. On the other hand $\kappa_0^{\mathbf{b}} \preceq \kappa_0^{\mathbf{a}}$, and thus we have a negative answer to Problem 1.1.1.

We employ below the following notation:

If α is an action of G on I and $S \subseteq G$, we write $\alpha(S) = \{\alpha(g) : g \in S\} \subseteq \mathrm{Sym}(I)$. For $G = \mathbf{F}_\infty$, we let $g_0, g_1, \dots$ be free generators of G and let $G_n = \langle g_0, g_1, \dots, g_n \rangle \leq G$.

If x is a real number, we write $\lfloor x \rfloor$ for the largest integer less than or equal to x. If x, y are real numbers and $\epsilon > 0$, we write $x \approx_\epsilon y$ to mean $|x - y| < \epsilon$. Finally, $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{N}^+ = \{1, 2, 3, \dots\}$

For the rest of the paper, $G = \mathbf{F}_\infty$.

1.2 Proof of Theorem 1.1.1

The structure of the proof is as follows. In Subsection 1.2 we state three lemmas. Temporarily assuming these lemmas, in Subsection 1.2 we give the main argument establishing Theorem 1.1.1. Then in Subsection 1.2 we prove the lemmas from Subsection 1.2.

Recall that for $\mathbf{a} \in A(G, X, \mu)$, we have $\mathbf{a} \preceq \mathbf{b}$ if and only if $\mathbf{a}$ lies in the closure of the isomorphic copies of $\mathbf{b}$. In particular, $\mathbf{b}$ is action-maximal if and only if the isomorphic copies of $\mathbf{b}$ are dense in $A(G, X, \mu)$. We will use these equivalences without comment several times in the sequel.

Statements of lemmas

The first lemma provides a general method for constructing invariant random subgroups.

Lemma 1.2.1. *Let α be an action of G on a countably infinite set I. Suppose there is an increasing sequence of non-empty finite subsets $(F_n)_{n=0}^\infty$ of I such that $\bigcup_{n=0}^\infty F_n = I$ and F_n is $\alpha(G_n)$-invariant. Let θ_n be the probability measure on* $\mathrm{Sub}(G)$ *given by the pushfoward of the uniform measure on F_n under the map $v \mapsto \mathrm{stab}_\alpha(v)$ (where $\mathrm{stab}_\alpha(v)$ is the stabilizer of v in α). Let θ be any weak-star limit point of the θ_n. Then θ is an invariant random subgroup of G.*

In order to state the second lemma, we need the following definition.

Definition 1.2.1. *Let α be an action of G on a finite set V and let n be such that all $\alpha(g_k), k > n$, act trivially. Let β be an action of G on a countably infinite set I. Let $Q \subseteq I$ be a finite set. We will say that α (relative to n)* ***appears in β within*** *Q if there is a $\beta(G_n)$-invariant set $W \subseteq Q$ and a bijection $\phi : V \to W$ such that $\phi(\alpha(g) \cdot v) = \beta(g) \cdot \phi(v)$ for all $v \in V$ and $g \in G_n$. We will say that α* ***appears in*** *β if it appears within some finite subset of I.*

Note that if α appears in β as above, then $\mathbf{s}_{\alpha\restriction G_n}$ is a factor of $\mathbf{s}_{\beta\restriction G_n}$.

Lemma 1.2.2. *There exists a sequence of finite sets $(V_n)_{n=1}^{\infty}$, with $|V_n| \to \infty$, and actions $(\alpha_n)_{n=1}^{\infty}$ of G, where α_n acts transitively on V_n so that all $g_k, k > n$, act trivially in α_n, such that if β is a transitive action of G on a countably infinite set and α_n (relative to n) appears in β for each n, then $\mathbf{s}_\beta$ is action-maximal and λ_β is representation-maximal.*

Fix a sequence of finite sets V_n and actions α_n of G on V_n, $n \geq 1$, as in Lemma 1.2.2. Given $f : \mathbb{N} \to \mathbb{N}^+$, $m > 0$, write $C_m(f) = \sum_{n=0}^{m-1}(|V_{f(n)}| + 1)$. We will need a function f with the following properties.

Lemma 1.2.3. *There exists a function $f : \mathbb{N} \to \mathbb{N}^+$ such that:*

(i) for every $n \geq 1$ there exists positive integer $K = K_n$ such that for all j there is l with $\left\lfloor \frac{j}{K} \right\rfloor = \left\lfloor \frac{l}{K} \right\rfloor$ and $f(l) = n$,

(ii) for every $\epsilon > 0$, there exists $t > 0$, such that for all $m > 0$ we have

$$\frac{1}{C_m(f)} \sum_{n=1}^{t} (|V_n| + 1) \cdot \left|\left\{j \in \{0, \ldots, m-1\} : f(j) = n\right\}\right| > 1 - \epsilon.$$

Main argument

Let, for $n \geq 1$, α_n and V_n be as in Lemma 1.2.2 and let f be as in Lemma 1.2.3. Choose a pairwise disjoint sequence of finite sets W_n, $n \geq 0$, such that $|W_n| = |V_{f(n)}|$. Define an action of α of G on $\bigcup_{n=0}^{\infty} W_n$ by identifying W_n with $V_{f(n)}$ and letting G act on W_n according to $\alpha_{f(n)}$. Let $\{u_n\}_{n=0}^{\infty}$ be an enumeration of a countably infinite set disjoint from the W_n. We now modify α to obtain a new action β of G on $I = \left(\bigcup_{n=0}^{\infty} W_n\right) \cup \{u_n\}_{n=0}^{\infty}$. We will have that $\beta(g_k)$ agrees with $\alpha(g_k)$ on W_n when $k \in \{0, \ldots, f(n)\}$.

For each n, choose a point $w_n \in W_n$ and let $\beta(g_{f(n)+1})$ transpose w_n with u_n. Let $(l_n)_{n=0}^{\infty}$ be a strictly increasing sequence of indices such that $\max(n, f(0), \ldots, f(n+1)) + 1 < l_n$. Let $\beta(g_{l_n})$ transpose w_n and w_{n+1}.

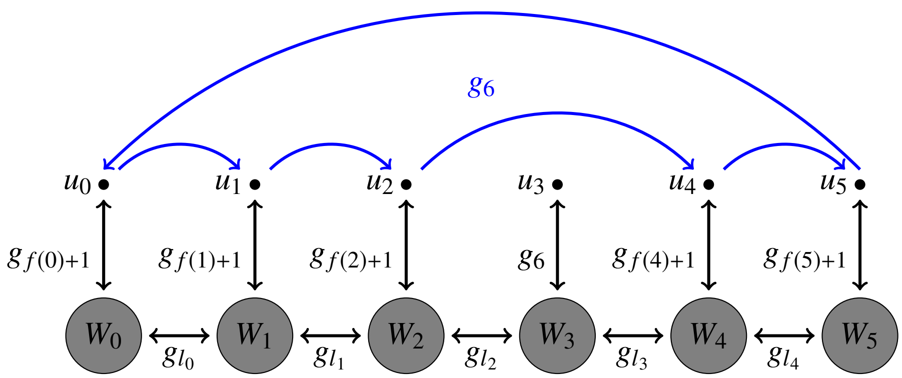
Fix $n \geq 1$. We now define how $\beta(g_n)$ acts on $\{u_j\}_{j=0}^{\infty}$. For $k \in \mathbb{N}$, consider the discrete interval

$$D_{n,k} = \{k \cdot n, \ldots, (k+1) \cdot n - 1\}.$$

We would like to have $\beta(g_n)$ make a cycle out $\{u_j, j \in D_{n,k}\}$ for each k. Unfortunately, we cannot achieve that exactly since there may by some $j \in D_{n,k}$ for which $f(j) + 1 = n$, and in this case we will have already used g_n to link W_j with u_j. Thus for each k, we will let $\beta(g_n)$ make a cycle out of the set

$$\{u_j : j \in D_{n,k} \text{ and } f(j) + 1 \neq n\},$$

making no modification to the action of $\beta(g_n)$ on those u_j for which $f(j) + 1 = n$. We will call these cycles the top cycles of $\beta(g_n)$. We have the following picture of β, where $n = f(3) + 1 = 6$ and we consider the interval $D_{6,0}$.



Finally β is defined trivially for all other points. Clearly β acts transitively. Write for $m > 0$, $\left(\bigcup_{k=0}^{m-1} W_k\right) \cup \{u_0, \ldots, u_{m-1}\} = T_m$ and for $m \geq 0$, $T_{m!} = F_m$. Thus F_m is invariant under $\beta(G_m)$. For each m, define a measure θ_m on $\mathrm{Sub}(G)$ be letting θ_m be the pushforward of the uniform measure on F_m under the map $v \mapsto \mathrm{stab}_\beta(v)$. Let θ be a weak-star limit point of θ_m. By Lemma 1.2.1, θ is an invariant random subgroup of G.

We claim that θ is supported on $\mathcal{M}_G$. Let $(Q_k)_{k=0}^{\infty}$ be an increasing sequence of finite subsets of G with $\bigcup_{k=0}^{\infty} Q_k = G$. For $H \leq G$, let $Q_k/H = \{gH : g \in Q_k\}$. Write, for $n \geq 1$, $k \in \mathbb{N}$,

$$A_{n,k} = \left\{H \leq G : \alpha_n \text{ appears in } \tau_{G/H} \text{ within } Q_k/H\right\}.$$

By definition, if $H \in \bigcup_{k=0}^{\infty} A_{n,k}$, then α_n appears in $\tau_{G/H}$. Therefore by Lemma 1.2.2, we have

$$\bigcap_{n=1}^{\infty}\bigcup_{k=0}^{\infty} A_{n,k} \subseteq \mathcal{M}_G.$$

Thus it suffices to show that for each $n \geq 1$ we have $\sup_{k<\infty} \theta(A_{n,k}) = 1$. Fix n and $\epsilon > 0$. Since the set $A_{n,k}$ is clopen for each k, it is enough to show the following:

Claim 1.2.1. *There is some $k \in \mathbb{N}$, such that for all $m > 0$, we have $\theta_m(A_{n,k}) > 1-\epsilon$.*

Let t be large enough that Lemma 1.2.3(ii) holds for our chosen ϵ. We now define five finite subsets of G.

- Let $S_1 \subseteq G$ consist of $\{1_G\}$ together with every word in the generators $g_0, \dots, g_t$ with length at most $\max_{1\leq j\leq t} |V_j|$. If $f(j) \leq t$, this choice will allow us to pass between points in W_j using an element of S_1.
- Let $S_2 = \{1_G, g_0, \dots, g_{t+1}\}$. If $f(j) \leq t$, this choice will allow us to pass to u_j from some point in W_j using an element of S_2.
- Let S_3 consist of all words in the generators g_K, g_{2K}, g_{3K} of length at most $3K$, where $K = K_n$ is the number provided by Lemma 1.2.3(i) for our fixed n. We will explain this choice later.
- Let $S_4 = \{g_{n+1}\}$. If $f(l) = n$, we will use g_{n+1} to pass from u_l to some point in W_l.
- Let S_5 consists of all words in the generators $g_1, \dots, g_n$ of length at most $|V_n|$. If $f(l) = n$, this choice will allow us to pass between any two points of W_l using an element of S_5.

Let k be large enough that Q_k contains $S_5 \cdot S_4 \cdot S_3 \cdot S_2 \cdot S_1$. We assert that the following implies Claim 1.2.1.

Claim 1.2.2. *If $v \in W_j \cup \{u_j\}$ and $f(j) \leq t$, then α_n appears in $\tau_{G/\mathrm{stab}_\beta(v)}$ within $Q_k/\mathrm{stab}_\beta(v)$.*

Indeed, suppose Claim 1.2.2 holds and let $m > 0$. Note that $C_{m!}(f)$ defined as in Lemma 1.2.3 is exactly $|T_{m!}|$. Thus we have

$$\begin{aligned}
\theta_m(A_{n,k}) &= \frac{1}{|T_{m!}|} \cdot \left|\left\{v \in T_{m!} \colon \operatorname{stab}_\beta(v) \in A_{n,k}\right\}\right| && (1.1)\\
&\geq \frac{1}{|T_{m!}|} \cdot \left|\left\{v \in T_{m!} : v \in W_j \cup \{u_j\} \text{ and } f(j) \leq t\right\}\right| && (1.2)\\
&= \frac{1}{|T_{m!}|} \sum_{n=1}^{t} (|V_n| + 1) \cdot \left|\left\{j \in \{0, \ldots, m! - 1\} : f(j) = n\right\}\right| && (1.3)\\
&> 1 - \epsilon, && (1.4)
\end{aligned}$$

where

- (1.1) follows from the definition of θ_m,
- (1.2) follows from (1.1) by Claim 1.2.2,
- (1.3) follows from (1.2) since $|W_j| = |V_{f(j)}|$,
- (1.4) follows from (1.3) by Lemma 1.2.1(ii).

Thus it remains to establish Claim 1.2.2.

Fix j with $f(j) \leq t$. By our choice of K, there is some l such that $\lfloor j/K \rfloor = \lfloor l/K \rfloor$ and $f(l) = n$. Fix $v \in W_j \cup \{u_j\}$. Write $H = \operatorname{stab}_\beta(v)$ and let $P = \{gH : \beta(g) \cdot v \in W_l\}$. Since $\beta(G_n)$ acts on W_l according to α_n, it follows that α_n appears in $\tau_{G/H}$ within P. Therefore it is enough to show that $P \subseteq Q_k/H$, or equivalently $W_l \subseteq \beta(Q_k) \cdot v$. The idea is that we have chosen k large enough that we can reach any point in W_l from v using the β action of a word from Q_k.

By our choice of S_1, if $v \in W_j$ there is an element $\gamma \in S_1$ such that $\beta(\gamma) \cdot v = w_j$ where w_j is the point in W_j connected to u_j. The connection between w_j and u_j is made by $\beta(g_{f(j)+1})$. We have $g_{f(j)+1} \in S_2$ since $f(j) \leq t$. Thus $u_j = \beta(\gamma) \cdot v$, where $\gamma \in S_2 \cdot S_1$.

Note that our assumption on l guarantees that l lies between the same pair of multiples of K as j does. We would like to say that this allows us to pass from u_j to u_l using $\beta(g_K)^i$ for some $i \in [-K, K]$. However, there is the minor issue of the points u_d which are skipped the top cycles of $\beta(g_K)$. We can easily overcome this obstacle by noting that for any d, at most one of $\beta(g_K), \beta(g_{2K})$, and $\beta(g_{3K})$ skips over u_d, and therefore there is a word γ' in g_K, g_{2K}, g_{3K} of length at most $3K$ such that $\beta(\gamma') \cdot u_j = u_l$. We have $\gamma' \in S_3$.

Since $f(l) = n$, we see that u_l is connected to W_l by $\beta(g_{f(l)+1}) = \beta(g_{n+1})$. Therefore $\beta(g_{n+1}\gamma'\gamma) \cdot v \in W_l$. Since $W_l \subseteq \beta(S_5) \cdot \beta(g_{n+1}\gamma'\gamma) \cdot v$, we have that $W_l \subseteq \beta(Q_k) \cdot v$ and we are done.

Proofs of lemmas

Proof of Lemma 1.2.1. Let $h_1, \ldots, h_l, k_1, \ldots, k_{l'}, g \in G$ and let $\epsilon > 0$. Let m be large enough that $h_1, \ldots, h_l$, $k_1, \ldots, k_{l'}, g$ are words in the generators $\{g_0, \ldots, g_m\}$. Write

$$C = \{H \leq G : h_1, \ldots, h_l \in H \text{ and } k_1, \ldots, k_{l'} \notin H\}.$$

Note that C is a clopen set and therefore there is some $n \geq m$ such that

$$\theta(C) \approx_\epsilon \theta_n(C) \text{ and } \theta(gCg^{-1}) \approx_\epsilon \theta_n(gCg^{-1}). \tag{1.5}$$

Noting that F_n is $\alpha(\langle g, h_1, \ldots, h_l, k_1, \ldots, k_{l'}\rangle)$ invariant we have

$$\begin{aligned}
\theta_n(gCg^{-1}) = {} & \frac{1}{|F_n|} \cdot \big|\{v \in F_n : \alpha(gh_jg^{-1}) \cdot v = v \text{ for all } j \in \{1, \ldots, l\} \\
& \text{and } \alpha(gk_jg^{-1}) \cdot v \neq v \text{ for all } j \in \{1, \ldots, l'\}\}\big| \\
= {} & \frac{1}{|F_n|} \cdot \big|\{v \in F_n : \alpha(h_j)\alpha(g^{-1}) \cdot v = \alpha(g^{-1}) \cdot v \text{ for all } j \in \{1, \ldots, l\} \\
& \text{and } \alpha(k_j)\alpha(g^{-1}) \cdot v \neq \alpha(g^{-1}) \cdot v \text{ for all } j \in \{1, \ldots, l'\}\}\big| \\
= {} & \frac{1}{|F_n|} \cdot \big|\{w \in F_n : \alpha(h_j) \cdot w = w \text{ for all } j \in \{1, \ldots, l\} \\
& \text{and } \alpha(k_j) \cdot w \neq w \text{ for all } j \in \{1, \ldots, l'\}\}\big| \\
= {} & \theta_n(C).
\end{aligned}$$

Then from (1.5) we have $\theta(C) \approx_{2\epsilon} \theta(gCg^{-1})$. □

Proof of Lemma 1.2.2. It is clearly enough to find such V_n, α_n such that for any β as in that lemma, $\mathbf{s}_\beta$ is action-maximal and another sequence, also denoted below by V_n, α_n, such that for any β as in that lemma, λ_β is representation-maximal. Then by interlacing these two sequences, we have a sequence that achieves both goals.

Case 1: We first find the sequence for which the appropriate $\mathbf{s}_\beta$ is action-maximal. By [54, Theorem 5.1], there is a countably infinite set J and a transitive action α of

G on J such that $\mathbf{s}_\alpha$ is action-maximal. Identify (X, μ) with 2^J carrying the usual product measure. For a finite set $T \subseteq J$ and $\rho \in 2^T$, write

$$N_\rho = \left\{x \in 2^J : x(v) = \rho(v) \text{ for all } v \in T\right\}.$$

For $n \geq 1$, $\epsilon > 0$ and a finite set $T \subseteq J$, let $U_{n,\epsilon,T}$ be the set of all $\mathbf{c} \in A(G, X, \mu)$ such that

$$\mu\big(\mathbf{s}_\alpha(g_k) \cdot N_\rho \cap N_\sigma\big) \approx_\epsilon \mu\big(\mathbf{c}(g_k) \cdot N_\rho \cap N_\sigma\big), \forall \sigma, \rho \in 2^T,\ k \in \{0, \ldots, n-1\}.$$

Observe that the collection of all $U_{n,\epsilon,T}$ is a neighborhood basis at $\mathbf{s}_\alpha \in A(G, X, \mu)$. Let $(T_n)_{n=1}^\infty$ be an increasing sequence of finite subsets of J with $\bigcup_{n=1}^\infty T_n = J$. Write $U_n = U_{n,2^{-n-|T_n|},T_n}$. Then the sets U_n form a neighborhood basis at $\mathbf{s}_\alpha$. Note that for each $n \geq 1$ and each $k \in \{0, \ldots, n-1\}$, we can extend $\alpha(g_k) \upharpoonright \left(T_n \cup \bigcup_{j=0}^{n-1} \alpha(g_j) \cdot T_n\right)$ to a permutation of J which is trivial on the complement of a finite set containing $T_n \cup \bigcup_{j=0}^{n-1} \alpha(g_j) \cdot T_n$. Hence for each $n \geq 1$, we can find an action $\widehat{\alpha}_n$ of G on J with the following properties:

(I) $\widehat{\alpha}_n(g_k) \cdot v = \alpha(g_k) \cdot v$, if $k \in \{0, \ldots, n-1\}$ and $v \in T_n$.

(II) $\widehat{\alpha}_n(g_k)$ acts trivially if $k > n$.

(III) There is a $\widehat{\alpha}_n$-invariant finite set $V_n \subseteq J$ such that $\widehat{\alpha}_n \upharpoonright (J \setminus V_n)$ is trivial and $\widehat{\alpha}_n \upharpoonright V_n$ is transitive.

By (I) we see that $\mathbf{s}_{\widehat{\alpha}_n}(g_k) \cdot N_\rho = \mathbf{s}_\alpha(g_k) \cdot N_\rho$ for all $\rho \in 2^{T_n}$ and $k \in \{0, \ldots, n-1\}$. Therefore $\mathbf{s}_{\widehat{\alpha}_n} \in U_n$. Write $\alpha_n = \widehat{\alpha}_n \upharpoonright V_n$. By (II) all $g_k, k > n$, act trivially in α_n. Observe that (III) implies that $\mathbf{s}_{\widehat{\alpha}_n} \cong \mathbf{s}_{\alpha_n} \times \boldsymbol{\iota}$, where $\boldsymbol{\iota}$ is the trivial action of G on a nonatomic standard probability space. Thus for each $n \geq 1$ there is an isomorphic copy of $\mathbf{s}_{\alpha_n} \times \boldsymbol{\iota}$ in U_n.

Suppose β is a transitive action of G on a countably infinite set such that α_n appears in β for each $n \geq 1$. Note that s_β is ergodic (see, e.g., [55, 2.1]). Then $\mathbf{s}_{\alpha_n \upharpoonright G_n}$ is a factor of $\mathbf{s}_{\beta \upharpoonright G_n}$ and hence $\mathbf{s}_{\alpha_n \upharpoonright G_n} \times (\boldsymbol{\iota} \upharpoonright G_n)$ is a factor of $\mathbf{s}_{\beta \upharpoonright G_n} \times (\boldsymbol{\iota} \upharpoonright G_n)$. Using the fact that the definition of U_n depends only on G_n, this implies that for each $n \geq 1$ there is an isomorphic copy of $\mathbf{s}_\beta \times \boldsymbol{\iota}$ in U_n. Therefore there is a sequence of isomorphic copies of $\mathbf{s}_\beta \times \boldsymbol{\iota}$ in $A(G, X, \mu)$ which converges to $\mathbf{s}_\alpha$. Since the isomorphic copies of $\mathbf{s}_\alpha$ are dense in $A(G, X, \mu)$, this implies that the isomorphic copies of $\mathbf{s}_\beta \times \boldsymbol{\iota}$ are dense in $A(G, X, \mu)$.

By [74, Theorem 3.11], we see that any ergodic action $\mathbf{d}$ of G is weakly contained in almost every ergodic component of $\mathbf{s}_\beta \times \boldsymbol{\iota}$. In particular, any ergodic action $\mathbf{d}$ of G is weakly contained in $\mathbf{s}_\beta$ and therefore the isomorphic copies of $\mathbf{s}_\beta$ are dense in the ergodic actions. Since G does not have Property (T), [53, Theorem 12.2] implies that the isomorphic copies of $\mathbf{s}_\beta$ are dense in $\mathrm{A}(G, X, \mu)$.

Case 2: We next find a sequence V_n, α_n, for which the appropriate λ_β is representation-maximal. We start with a transitive action α of G on a countably infinite set J such that λ_α is representation-maximal (see [54, Theorem 5.5]. Then proceed as in the proof of Case 1 to find V_n, α_n such that for some isomorphic copy σ_n of $\lambda_{\alpha_n} \oplus \infty 1_G$, (σ_n) converges to λ_α, where 1_G is the trivial one-dimensional representation of G and $\infty 1_G$ the direct sum of countably many copies of 1_G, i.e., the trivial representation on a separable, infinite-dimensional Hilbert space. Let now β be as above. Then the isomorphic copies of $\lambda_\beta \oplus \infty 1_G$ converge to λ_α. By a result of Hjorth, see [53, H.7], the irreducible representations are dense in $\mathrm{Rep}(G, H)$. Every irreducible representation π is $\preceq \lambda_\alpha$ and thus $\preceq_Z \lambda_\alpha \preceq_Z \lambda_\beta \oplus \infty 1_G$, where $\preceq_Z$ is weak containment in the sense of Zimmer. Recall that $\sigma \preceq_Z \rho$ iff σ is in the closure of the isomorphic copies of ρ. Also $\sigma \preceq_Z \rho \implies \sigma \preceq \rho$ and for σ irreducible, $\sigma \preceq_Z \rho \iff \sigma \preceq \rho$ (see [8, page 397] and [53, page 209]). Then by [1, Proposition 3.5] π is a subrepresentation of an ultrapower of $\lambda_\beta \oplus \infty 1_G$, which is of course of the form $\lambda^*_\beta \oplus \eta^*$, where λ^* is an ultrapower of λ_β and η^* a trivial representation of G on a Hilbert space H^*. Let H_1 be the space on which this subrepresentation acts, which is a G-invariant subspace of the direct sum of the space of λ^*_β and H^*. Then if $v \in H^*$ and v_1 is its projection on H_1, v_1 is G-invariant, so as π is irreducible, $v_1 = 0$, i.e., $H^* \perp H_1$. Thus H_1 is contained in the space of λ^*_β, i.e., π is a subrepresentation of λ^*_β, so $\pi \preceq_Z \lambda_\beta$. Thus the isomorphic copies of λ_β are dense in $\mathrm{Rep}(G, H)$, i.e., λ_β is representation-maximal. □

Proof of Lemma 1.2.3. Note that letting for $n \geq 1$, $A_n = f^{-1}(\{n\})$ the statement of the lemma is equivalent to the existence of a partition $\mathbb{N} = \bigsqcup_{n \geq 1} A_n$ with the following properties:

(i) For each $n \geq 1$ there is positive integer K_n such that A_n intersects each interval $I^n_i = [iK_n, (i+1)K_n), i = 0, 1, 2, \dots$.

(ii) Let $g : \mathbb{N}^+ \to \mathbb{N}^+$ be defined by $g(n) = |V_n| + 1$, where V_n is as in Lemma 1.2.2. Then we have that for each $\epsilon > 0$, there is $t > 0$, such that for all $m > 0$:

$$\frac{\sum_{n>t}(|A_n \cap m| \cdot g(n))}{\sum_n(|A_n \cap m| \cdot g(n))} < \epsilon,$$

where we identify here m with $\{0, 1, \ldots, m-1\}$.

To construct A_n, K_n, first chose $a_2 < a_3 < \ldots$ to be large enough so that a_n is divisible by 3 and

$$\sum_{n=2}^{\infty} \frac{1}{a_2 \cdots a_n} < \frac{1}{3} \text{ and } \frac{a_n}{3} > \frac{g(n)2^n}{g(n-1)}.$$

We let $A'_1 = \{2i : i \in \mathbb{N}\}$ and also put $K_1 = 2, K_n = 2a_2 \cdots a_n$ for $n \geq 2$. We will then inductively define pairwise disjoint $A_2, A_3, \ldots,$ which are also disjoint from A'_1, to satisfy (ii) above and so that for $n \geq 2$, A_n has exactly one member in each interval I^n_i as above, and finally we put $A_1 = \mathbb{N} \setminus \bigcup_{n=2}^{\infty} A_n$.

So assume that $A'_1, A_2, \ldots, A_{n-1}$ have been constructed (this is just A'_1, if $n = 2$). To find A_n, so that (i) above is satisfied, it is enough to have for each i,

$$K_n > \frac{3}{2} \left|(A'_1 \cup A_2 \cup \cdots A_{n-1}) \cap I^n_i\right|.$$

But

$$\left|(A'_1 \cup A_2 \cup \cdots A_{n-1}) \cap I^n_i\right| = a_2 \cdots a_n + a_3 \cdots a_n + \cdots + a_{n-1}a_n + a_n,$$

so this follows from $\sum_{n=2}^{\infty} \frac{1}{a_2 \cdots a_n} < \frac{1}{3}$. Also for $i = 0$, we can choose the element of A_n in $[0, K_n)$ to be $\geq \frac{K_n}{3}$.

We finally check that (ii) is satisfied. Fix $\epsilon > 0$ and choose $t > 1$ so that $\sum_{n>t}^{\infty} 2^{-n} < \epsilon$. Consider now any $m > 0$ and $n > t$.

Case 1. $m \geq K_n$. Then for some $s > 1$, we have that $m \in I^n_{s-1}$ and $|A_n \cap m| \leq s$, while

$$\sum_n |A_n \cap m| \cdot g(n) \geq |A_{n-1} \cap m| \cdot g(n-1) \geq (s-1)a_n \cdot g(n-1)$$

so

$$\frac{|A_n \cap m| \cdot g(n)}{\sum_n |A_n \cap m| \cdot g(n)} \leq \frac{s \cdot g(n)}{(s-1) \cdot g(n-1)} \cdot \frac{1}{a_n} < 2^{-n}.$$

Case 2. $m < K_n$. Then either $m \leq \frac{K_n}{3}$ and $|A_n \cap m| = 0$ or $m > \frac{K_n}{3}$ and $|A_n \cap m| \leq 1$, in which case also

$$|A_{n-1} \cap m| \geq \frac{a_n}{3}.$$

So for any $m < K_n$,

$$\frac{|A_n \cap m| \cdot g(n)}{\sum_n |A_n \cap m| \cdot g(n)} \le \frac{g(n)}{(\frac{a_n}{3})g(n-1)} < 2^{-n}.$$

Thus for any $n > t$, we have

$$\frac{|A_n \cap m| \cdot g(n)}{\sum_n |A_n \cap m| \cdot g(n)} < 2^{-n}$$

and so

$$\frac{\sum_{n>t}(|A_n \cap m| \cdot g(n))}{\sum_n(|A_n \cap m| \cdot g(n))} < \epsilon$$

□

1.3 Proof of Theorem 1.1.2

We note that if $\lambda_{G/H}$ is representation-maximal, then H is not amenable. This is because $1_G \le \lambda_{G/H}$ implies $\tau_{G/H}$ is amenable (see [55, Theorem 1.1]).

We will use the notion of a random Bernoulli shift over an invariant random subgroup; we refer the reader to [74, Section 5.3] and [4, Proposition 45] for details. Let θ be the invariant random subgroup constructed in Theorem 1.1.1 and let $\mathbf{s}_\theta$ be the θ-random Bernoulli shift. Note that for almost every ergodic component $\mathbf{b}$ of $\mathbf{s}_\theta$, almost all stabilizers of $\mathbf{b}$ lie in $\mathcal{M}_G$ and hence the type of $\mathbf{b}$ is supported on $\mathcal{M}_G$. Fix such an action $\mathbf{b}$. Let (Y, ν) be the underlying space of $\mathbf{b}$.

For $y \in Y$ write $H_y = \mathrm{stab}_{\mathbf{b}}(y)$. By [34, Proposition 8] we have $\lambda_{G/H_y} \le \kappa_0^{\mathbf{b}}$ for ν-almost every $y \in Y$. Since the type of $\mathbf{b}$ is supported on $\mathcal{M}_G$, for ν-almost every y we have that λ_{G/H_y} is representation-maximal and so $\kappa_0^{\mathbf{b}}$ is representation-maximal. Let $\mathbf{a} = \mathbf{b} \times \mathbf{s}_G$. Then $\mathbf{a}$ is free and ergodic and $\kappa_0^{\mathbf{a}}$ is representation-maximal. Suppose, toward a contradiction, that $\mathbf{a}$ were action-maximal.

Let $S \subseteq G^2$ be the collection of all pairs (g, h) such that $\langle g, h\rangle$ is nonamenable. Since λ_{G/H_y} is representation-maximal for ν-almost every $y \in Y$, and so H_y is not amenable, we see that $S \cap H_y^2$ is nonempty for ν-almost every y. Let $\phi : \mathbb{N} \to S$ be an enumeration of S. For $y \in Y$ let $\phi_y = \min\{n : \phi(n) \in H_y^2\}$. Then there is some $k \in \mathbb{N}$ such that $\nu(\{y : \phi_y = k\}) > 0$. Write $A = \{y : \phi_y = k\}$ and let N be the subgroup of G generated by the coordinates of $\phi(k)$. Note that for $y \in A$, we have $N \subseteq H_y$, and so $\mathbf{b} \upharpoonright N$ is trivial on A. By [53, Page 74], since $\mathbf{a}$ is action-maximal for G, we have that $\mathbf{a} \upharpoonright N$ is action-maximal for N. Observe that

$$\mathbf{a} \upharpoonright N = (\mathbf{b} \upharpoonright N) \times (\mathbf{s}_G \upharpoonright N) \cong (\mathbf{b} \upharpoonright N) \times (\mathbf{s}_N)^{\mathbb{N}} \cong (\mathbf{b} \upharpoonright N) \times \mathbf{s}_N.$$

So writing $\mathbf{c} = (\mathbf{b} \upharpoonright N) \times \mathbf{s}_N$, we have that $\mathbf{c}$ is action-maximal for N.

By [74, Theorem 3.11], this implies that any ergodic action $\mathbf{d}$ of N is weakly contained in almost every ergodic component of $\mathbf{c}$. Note that if $y \in A$, then $\boldsymbol{\iota}_{\{y\}} \times \mathbf{s}_N \cong \mathbf{s}_N$ is an ergodic component of $\mathbf{c}$, where by $\boldsymbol{\iota}_{\{y\}}$ we mean the trivial action of N on the one-point space $\{y\}$. Therefore $\mathbf{d} \preceq \mathbf{s}_N$. Since N does not have property (T), the ergodic actions of N are dense in $A(N, X, \mu)$ (see [53, 12.2]), so the isomorphic copies of $\mathbf{s}_N$ are dense in $A(N, X, \mu)$. But by [53, Proposition 13.2] this contradicts the fact that N is nonamenable.

Remark 1.3.1. *For $G = \mathbf{F}_\infty$, let $\mathbf{a}$ be as in Theorem 1.1.2. Then for any irreducible π we have $\pi \preceq \kappa_0^{\mathbf{a}}$, so $\pi \preceq_Z \kappa_0^{\mathbf{a}}$. Thus, as the irreducible representations are dense, $\pi \preceq_Z \kappa_0^{\mathbf{a}}$, for all π. Thus there is a free ergodic action $\mathbf{b}$ such that $\kappa_0^{\mathbf{b}} \preceq_Z \kappa_0^{\mathbf{a}}$ but $\mathbf{b} \not\preceq \mathbf{a}$, which is a somewhat stronger negative answer to Problem 1.1.1.*

Remark 1.3.2. *It is possible that one could use the techniques developed in this paper to show that Theorem 1.1.2 also holds for the free groups with finitely many generators $n > 1$ but we have not verified that.*

Chapter 2

TOPOLOGY AND CONVEXITY IN THE SPACE OF ACTIONS MODULO WEAK EQUIVALENCE

Peter Burton[1]

2.1 Introduction

By a probability space we mean a standard Borel space Y with a Borel probability measure ν. If ν is nonatomic, we say the pair (Y, ν) is a standard probability space. If ν is nonatomic then Y must be uncountable and thus by Theorem 17.41 in [56] every standard probability space is isomorphic to the unit interval with Lebesgue measure. Let Γ be a countable discrete group. By a measure-preserving action of Γ on (Y, ν) we mean a Borel action $a : \Gamma \times Y \to Y$ which preserves the measure ν. We write $\Gamma \curvearrowright^a (Y, \nu)$. In accordance with the standard conventions of ergodic theory, we identify two actions which agree almost everywhere. Thus a measure-preserving action of Γ on (Y, ν) is equivalently a homomorphism from Γ into the group $\mathrm{Aut}(Y, \nu)$ of measure-preserving automorphisms of (Y, ν), where again two such automorphisms are identified if they agree almost everywhere.

We fix a standard probability space (X, μ) throughout the remainder of the paper. As in [53] we can define the Polish space $\mathrm{A}(\Gamma, X, \mu)$ of measure-preserving actions of Γ. Kechris defines the following relation of weak containment among measure-preserving actions of Γ, by analogy with the standard notion of weak containment for representations.

Definition 2.1.1. *[53] If $\Gamma \curvearrowright^a (X, \mu)$ and $\Gamma \curvearrowright^b (Y, \nu)$ are measure-preserving actions of Γ on probability spaces, we say a is **weakly contained** in b and write $a \prec b$ if for any finite sequence $A_1, \ldots, A_n$ of measurable subsets of X, finite $F \subseteq \Gamma$ and $\epsilon > 0$ there exist measurable subsets $B_1, \ldots, B_n$ of Y such that for all $\gamma \in F$ and all $i, j \leq n$ we have*

$$|\mu(\gamma^a A_i \cap A_j) - \nu(\gamma^b B_i \cap B_j)| < \epsilon.$$

[1]Research partially supported by NSF grant DMS-0968710

*We say a is **weakly equivalent** to b and write $a \sim b$ if $a \prec b$ and $b \prec a$.*

We may assume in this definition that $A_1, \ldots, A_n$ form a partition of X. Note that we do not require (X, μ) and (Y, ν) to be standard, that is to say we include the case where they might be countable. The relation of weak containment is G_δ, so the quotient $\mathrm{A}_\sim(\Gamma, X, \mu)$ of $\mathrm{A}(\Gamma, X, \mu)$ by weak equivalence is well-behaved.

We also consider a generalization of weak containment, due to Tucker-Drob. For probability spaces $(Y_i, \nu_i), 1 \leq i \leq m$ and positive real numbers $\alpha_i, 1 \leq i \leq m$ with $\sum_{i=1}^m \alpha_i = 1$ we let $\bigsqcup_{i=1}^m \alpha_i Y_i$ be the probability space formed by endowing the disjoint union of the Y_i with the measure $\sum_{i=1}^m \alpha_i \nu_i$ given by $\left(\sum_{i=1}^m \alpha_i \nu_i\right)(A) = \sum_{i=1}^m \alpha_i \nu_i(A \cap Y_i)$. If $\Gamma \curvearrowright^{a_i} (Y_i, \nu_i)$ are measure-preserving actions, then $\sum_{i=1}^m \alpha_i a_i$ is the action on $\bigsqcup_{i=1}^m \alpha_i Y_i$ given by letting Γ act like a_i on Y_i.

Definition 2.1.2. *[74] If $\Gamma \curvearrowright^a (X, \mu)$ and $\Gamma \curvearrowright^b (Y, \nu)$ are measure-preserving actions, we say a is **stably weakly contained** in b if for all $A_1, \ldots, A_k \in \mathrm{MALG}_\mu$, all finite $F \subseteq \Gamma$ and all $\epsilon > 0$ there exist $\alpha_1, \ldots, \alpha_m$ such that $\sum_{i=1}^m \alpha_1 = 1$ and sets $B_1, \ldots, B_k \subseteq \bigsqcup_{i=1}^m \alpha_i Y_i$ such that*

$$\left| \mu(\gamma^a A_i \cap A_j) - \sum_{i=1}^m \alpha_i \nu \left(\gamma^{\sum_{i=1}^m \alpha_i b} B_i \cap B_j \right) \right| < \epsilon.$$

We write $a \prec_s b$ if a is stably weakly contained in b and $a \sim_s b$ for $a \prec_s b$ and $b \prec_s a$.

When we wish to distinguish between an action and its equivalence class, we write $[a]$ for the weak equivalence class of a and $[a]_s$ for the stable weak equivalence class. The quotient of $\mathrm{A}(\Gamma, X, \mu)$ by the relation of stable weak containment is denoted $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$. The goal of this paper is to analyze the topological and geometric structure of $\mathrm{A}_\sim(\Gamma, X, \mu)$ and $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$.

More specifically, unlike $\mathrm{A}(\Gamma, X, \mu)$, the spaces $\mathrm{A}_\sim(\Gamma, X, \mu)$ and $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ carry a well-defined operation of convex combination. This is inherited from the operation of endowing the disjoint union of two probability spaces with a convex combination of their respective measures. In Section 2 we introduce a variation of a construction of Fritz [37] which abstracts the idea of convex combinations. Fritz's objects are referred to as 'convex spaces'; we weaken the definition in order to encompass the convex structure on $\mathrm{A}_\sim(\Gamma, X, \mu)$, obtaining the notion of 'weak convex space'. We show that this relates naturally to other ideas of convexity, define a notion of convex function, and generalize the important geometric notions of 'convex hull', 'extreme point', and 'face' from the classical situation of vector spaces to this abstract framework. We also define 'topological weak convex spaces' as weak convex structures which are appropriately compatible with an underlying topology.

In Section 3 we consider methods of topologizing $\mathrm{A}_{\sim}(\Gamma, X, \mu)$. The first topology defined on this space was in [1], and a second formulation was given in [74]. These are equivalent, Polish, compact and finer than the quotient of the weak topology on $\mathrm{A}(\Gamma, X, \mu)$. We discuss a third topology, implicit in [1] and pointed out to us by Kechris. This is shown to be equivalent to the previous two. We also consider a natural topology on $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$.

In Section 4 we describe how to endow $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ with the structure of a weak convex space and show that it is in fact a topological weak convex space. Furthermore, we show that the metric giving $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ its Polish topology is compatible with the convex structure in the sense that the distance function to any compact convex set is a convex function.

In Section 5 we analyze the structure of $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ for amenable Γ. The main tool is the following idea. Let $\mathrm{Sub}(\Gamma)$ be the space of subgroups of Γ, regarded as a subspace of $\{0, 1\}^{\Gamma}$ with the product topology. $\mathrm{Sub}(\Gamma)$ is then a compact metric space on which Γ acts by conjugation.

Definition 2.1.3. *An **invariant random subgroup** of* Γ *is a conjugation-invariant Borel probability measure on* $\mathrm{Sub}(\Gamma)$.

Invariant random subgroups have been studied in numerous recent papers, including [4], [11], [13] and [35]. If $\Gamma \curvearrowright^{a} (X, \mu)$ is a measure-preserving action, then the pushforward measure $(\mathrm{stab}_a)_* \mu$ is an invariant random subgroup of Γ called the type of a. We extend ideas of Tucker-Drob from [74] to show the following.

Theorem 2.1.1. *If* Γ *is amenable, then* $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ *is isomorphic to* $\mathrm{IRS}(\Gamma)$ *as a topological convex space. In particular, if* Γ *is amenable then* $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ *is isomorphic to a compact convex subset of a Banach space.*

In Section 6 we consider the structure of $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ for general Γ. If Γ is non-amenable, the existence of strongly ergodic actions of Γ implies that the convex structure on this space has the pathology that the convex combination of a point x with itself might be different from x. This is why we need to consider weak convex spaces instead of just convex spaces. The main result of this section is the following Krein-Milman type theorem.

Theorem 2.1.2. $\mathrm{A}_\sim(\Gamma, X, \mu)$ *is equal to the closed convex hull of its extreme points. In other words, finite convex combinations of the extreme points of* $\mathrm{A}_\sim(\Gamma, X, \mu)$ *are dense in* $\mathrm{A}_\sim(\Gamma, X, \mu)$.

Given this result, it seems interesting to describe the extreme points of $\mathrm{A}_\sim(\Gamma, X, \mu)$. In the amenable case, the identification with $\mathrm{IRS}(\Gamma)$ provides a complete such description, since the extreme points of $\mathrm{IRS}(\Gamma)$ are known to be the ergodic measures and consequently the extreme points of $\mathrm{A}_\sim(\Gamma, X, \mu)$ for amenable Γ are exactly those actions with ergodic type. In the nonamenable case this description does not suffice. It is clear that any strongly ergodic action is an extreme point. We are able to show the following.

Theorem 2.1.3. *Suppose* $[a] \in \mathrm{A}_\sim(\Gamma, X, \mu)$ *is an extreme point. Let* $a = \int_Z a_z d\eta(z)$ *be the ergodic decomposition of a. Then there is a measure-preserving action b of* Γ *such that for* η*-almost all* $z \in Z$ *we have* $[a_z] = [b]$.

Let $\mathrm{FR}_\sim(\Gamma, X, \mu)$ denote the subspace of $\mathrm{A}_\sim(\Gamma, X, \mu)$ consisting of the weak equivalence classes of free actions. We prove:

Theorem 2.1.4. *Let* $\mathbb{F}_N$ *be a free group of finite or countably infinite rank. Then the weak equivalence classes containing a free ergodic action are dense in* $\mathrm{FR}_\sim(\mathbb{F}_N, X, \mu)$.

In Section 7 we use a characterization of convex subsets of Banach spaces from [25] to show the following.

Theorem 2.1.5. *For any* Γ*, the space* $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ *is isomorphic to a compact convex subset of a Banach space.*

We characterize the extreme points of $A_{\sim_s}(\Gamma, X, \mu)$ as precisely those stable weak equivalence classes which contain an ergodic action. This result was obtained by Tucker-Drob and Bowen independently of the author. Tucker-Drob and Bowen have also shown that $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ is a simplex, and the set $\mathrm{FR}_{\sim_s}(\Gamma, X, \mu)$ of stable weak equivalence classes of free actions is a subsimplex. Recall that a Poulsen simplex is a simplex such that the extreme points are dense. Thus from Theorem 2.1.4 we have:

Corollary 2.1.1. *Let $\mathbb{F}_N$ be a free group of finite or countably infinite rank. Then* $\mathrm{FR}_{\sim_s}(\mathbb{F}_N, X, \mu)$ *is a Poulsen simplex.*

Acknowledgements
We would like to thank Alexander Kechris for introducing us to this topic and for many helpful discussions. We also thank Robin Tucker-Drob for informing us of his result with Bowen that the space of stable weak equivalence classes forms a simplex, and for raising the question of when it forms a Poulsen simplex. We thank the anonymous referee for useful comments.

2.2 Weak convex spaces

We first describe the formalism realized by $\mathrm{A}_\sim(\Gamma, X, \mu)$.

Convex spaces and weak convex spaces.
Convex spaces were introduced in [37] and further developed in [25] as an abstract setting to study the notion of convex combination.

Definition 2.2.1. *[37] A **convex space** is a set X together with a family $\mathcal{V}$ of binary operations cc_t for each $t \in [0,1]$ such that for all $x, y, z \in X$ and all $s, t \in [0,1]$*

(1) $cc_0(x, y) = x$,

(2) $cc_t(x, x) = x$,

(3) $cc_t(x, y) = cc_{1-t}(y, x)$,

(4) $cc_t(cc_s(x, y), z) = cc_{st}\left(x, cc_{\frac{t(1-s)}{1-st}}(y, z)\right)$.

We will usually write $tx +_{\mathcal{V}} (1-t)y$ for $cc_t(x, y)$, omitting the subscript $\mathcal{V}$ when the convex structure being considered is clear. Note that (4) allows us to unambiguously define $\sum_{i=1}^n \lambda_i x_i$ for $(x_i)_{i=1}^n \subseteq X$ and $(\lambda_i)_{i=1}^n \subseteq [0,1]$ such that $\sum_{i=1}^n \lambda_i = 1$. We will need to weaken the definition of a convex space to cover the situation where a convex combination of a point x with itself could be different from x.

Definition 2.2.2. *An **weak convex space** is a set X with a family cc_t of binary operations for $t \in [0,1]$, satisfying* (1), (3)*, and* (4) *of Definition 2.2.1.*

Definition 2.2.3. *A **topological (weak) convex space** is a topological space X carrying a (weak) convex structure such that the ternary operation $cc : [0,1]\times X^2 \to X$ given by $cc(t,x,y) = cc_t(x,y)$ is continuous.*

Extreme points and faces

We can define extreme points in a weak convex space in exactly the same way as in a vector space.

Definition 2.2.4. *If A is a convex set in a weak convex space, we say $x \in A$ is an **extreme point** if $x = ty + (1-t)z$ for $0 < t < 1$ and some $y, z \in A$ implies $y = z = x$. Write $\mathrm{ex}(A)$ for the set of extreme points of A. If A is a compact convex subset of a topological weak convex space, we say a **face** of A is a nonempty closed subset $F \subseteq A$ such that if $x, y \in A$, $0 < t < 1$ and $tx + (1-t)y \in F$ then $x, y \in F$.*

2.3 Topology on the space of weak equivalence classes

Let Γ be a countable group and $\mathrm{A}_\sim(\Gamma, X, \mu)$ be its space of actions modulo weak equivalence. We consider a metric on $\mathrm{A}_\sim(\Gamma, X, \mu)$ which is implicit in [1].

Fix an enumeration $(\gamma_i)_{i=0}^\infty$ of Γ. If $\mathcal{A} = \{A_1, \ldots, A_k\}$ is a partition of X into k pieces, $a \in \mathrm{A}(\Gamma, X, \mu)$ and $n \in \mathbb{N}$, let $M_{n,k}^{\mathcal{A}}(a) \in [0,1]^{n\times k\times k}$ be the point whose p, q, r coordinate is $\mu(\gamma_p^a A_q \cap A_r)$, where $p \leq n$ and $q, r \leq k$. Let $C_{n,k}(a) = \overline{\{M_{n,k}^{\mathcal{A}}(a) : \mathcal{A} \text{ is a partition of } X \text{ into } k \text{ pieces.}\}}$ Then we can define a pseudometric d on $\mathrm{A}(\Gamma, X, \mu)$ by the formula

$$d(a,b) = \sum_{n,k=1}^{\infty} \frac{1}{2^{n+k}} d_H(C_{n,k}(a), C_{n,k}(b)),$$

where d_H is the Hausdorff distance in the hyperspace of compact subsets of $[0,1]^{n\times k\times k}$. It is easy to see that $a \sim b$ if and only if $d(a,b) = 0$, so d descends to a metric on $\mathrm{A}_\sim(\Gamma, X, \mu)$, which we also denote by d. Let τ_1 be the topology induced by d. We note that this definition extends to actions on countable spaces. We will write $A_\sim^*(\Gamma)$ for the space of all actions of Γ on probability spaces.

We now describe a different construction of the topology on $\mathrm{A}_\sim(\Gamma, X, \mu)$ due to Tucker-Drob [74] in order to show it agrees with the one we have just introduced. (Tucker-Drob shows in [74] that his formulation agrees with the one from [1]).

Let S be a compact Polish space, and consider S^Γ, which is also a compact Polish space. Γ acts on S^Γ by the shift action s given by $(\gamma^s f)(\delta) = f(\gamma^{-1}\delta)$. Let $M_s(S^\Gamma)$ be the compact Polish space of shift-invariant probability measures on S^Γ and let $\mathcal{K}_S = \mathcal{K}(M_s(S^\Gamma))$ be the hyperspace of compact subsets of $M_s(S^\Gamma)$ equipped with the Hausdorff topology. Then $\mathcal{K}_S$ is again compact and Polish. Now consider an S-valued random variable $\phi \in L(X, \mu, S)$ on X, that is to say a measurable map $\phi : X \to S$. For each measure-preserving action $a \in \mathrm{A}(\Gamma, X, \mu)$ we get a map $\Phi_S^{\phi,a} : X \to S^\Gamma$ by letting $\Phi_S^{\phi,a}(x)(\gamma) = \phi((\gamma^{-1})^a x)$ and consequently a shift-invariant measure $(\Phi_S^{\phi,a})_*\mu$ on S^Γ. Then define a subset $E(a, S)$ of $M_s(S^\Gamma)$ by

$$E(a, S) = \{(\Phi_S^{\phi,a})_*\mu : \phi : X \to S \text{ is measurable}\}.$$

Let $\Phi_S : \mathrm{A}(\Gamma, X, \mu) \to \mathcal{K}_S$ be given by $\Phi_S(a) = \overline{E(a, S)}$. When $S = K$ is the Cantor set, we omit the subscript S on the notations just introduced. By Proposition 3.5 in [74], we have $a \sim b$ if and only if $\Phi(a) = \Phi(b)$ so we can consider the initial topology on $\mathrm{A}_\sim(\Gamma, X, \mu)$ induced by Φ. Call this τ_2. We now work towards showing τ_1 agrees with τ_2. There will be a series of preliminary steps. This entire argument can be regarded as a 'perturbed' version of Proposition 3.5 in [74].

We first fix a compatible metric on $M_s(K^\Gamma)$. Let $\mathcal{A}_K$ be the collection of clopen subsets of K^Γ of the form $\pi_F^{-1}\left(\prod_{\gamma\in F} A_\gamma\right)$ where $A_\gamma \subseteq K$ an element of some fixed countable clopen basis for K, $F \subseteq \Gamma$ is finite and $\pi : K^\Gamma \to K^F$ is the projection onto the F-coordinates. Since the elements of $\mathcal{A}_K$ generate the Borel σ-algebra of K^Γ, for $(\nu_n)_{n=1}^\infty \subseteq M_s(K^\Gamma)$ we have $\nu_n \to \nu$ in $M_s(K^\Gamma)$ if and only if $\nu_n(A) \to \nu(A)$ for every $A \in \mathcal{A}_K$. So, enumerating the elements of $\mathcal{A}_K$ as $(A_i^K)_{i=1}^\infty$, δ_K given by

$$\delta_K(\nu, \rho) = \sum_{i=1}^\infty \frac{1}{2^i}|\nu(A_i^K) - \rho(A_i^K)|$$

is a compatible metric on $M_s(K^\Gamma)$.

Lemma 2.3.1. *For any $\epsilon > 0$ there exists $k \in \mathbb{N}$ such that every a and every $\phi \in L(X, \mu, K)$ there is $\psi \in L(X, \mu, K)$ with $\delta_K((\Phi^{\phi,a})_*\mu, (\Phi^{\psi,a})_*\mu) < \epsilon$ such that the range of ψ has size $\leq k$. Note that k depends only on ϵ, not on a or ϕ.*

Proof. Fix ϵ. Choose N large enough that $\sum_{i=N}^\infty \frac{1}{2^i} < \epsilon$. For each $i \leq N$, write $A_i = \pi_{F_i}^{-1}\left(\prod_{\gamma\in F_i} A_\gamma^i\right)$ for $A_\gamma^i \subseteq K$ clopen and $F_i \subseteq \Gamma$ finite. We have for all

$a \in \mathrm{A}(\Gamma, X, \mu)$ and $\phi, \psi \in L(X, \mu, K)$,

$$\begin{aligned} |\Phi^{\phi,a}(A_i) - \Phi^{\psi,a}(A_i)| &= \left|\Phi^{\phi,a}\left(\pi_{F_i}^{-1}\left(\prod_{\gamma \in F_i} A_\gamma^i\right)\right) - \Phi^{\psi,a}\left(\pi_{F_i}^{-1}\left(\prod_{\gamma \in F_i} A_\gamma^i\right)\right)\right| \\ &= |\mu(\{x : \Phi^{\phi,a}(x)(\gamma) \in A_\gamma^i \text{ for all } \gamma \in F_i\}) \\ &\qquad - \mu(\{x : \Phi^{\psi,a}(x)(\gamma) \in A_\gamma^i \text{ for all } \gamma \in F_i\})| \\ &= |\mu(\{x : \phi((\gamma^{-1})^a x) \in A_\gamma^i \text{ for all } \gamma \in F_i\}) \\ &\qquad - \mu(\{x : \psi((\gamma^{-1})^a x) \in A_\gamma^i \text{ for all } \gamma \in F_i\})| \\ &= \left|\mu\left(\bigcap_{\gamma \in F_i} \gamma^a \phi^{-1}(A_\gamma^i)\right) - \left(\bigcap_{\gamma \in F_i} \gamma^a \psi^{-1}(A_\gamma^i)\right)\right|. \end{aligned} \tag{2.1}$$

Now, fix $\phi \in L(X, \mu, K)$. Let $(B_j)_{j=1}^k$ be the finite partition of K given by the atoms of the Boolean algebra generated by $(A_\gamma^i)_{i \leq N, \gamma \in F_i}$. Note that k depends only on ϵ. For each $j \leq k$, let y_j be any point in B_j. Define a map $\psi : X \to K$ by letting $\psi(x) = y_j$ for the unique j such that $x \in \phi^{-1}(B_j)$. Then $\psi^{-1}(B_j) = \phi^{-1}(B_j)$ for each j, and hence $\phi^{-1}(A_\gamma^i) = \psi^{-1}(A_\gamma^i)$ for each $i \leq N$ and $\gamma \in F_i$. Therefore the value of the expression (1) is 0 and $\delta_K((\Phi^{\phi,a})_*\mu, (\Phi^{\psi,a})_*\mu) < \epsilon$. □

Lemma 2.3.2. *If $\overline{E(a_n, L)} \to \overline{E(a, L)}$ in $\mathcal{K}(M_s(L^\Gamma))$ for every finite set L then $\overline{E(a_n, K)} \to \overline{E(a, K)}$ in $\mathcal{K}(M_s(K^\Gamma))$.*

Proof. Fix $\epsilon > 0$ in order to show that eventually $d_{\mathcal{K}}\left(\overline{E(a_n, K)}, \overline{E(a, K)}\right) < \epsilon$, where $d_{\mathcal{K}}$ is the Hausdorff distance in $\mathcal{K}(M_s(K^\Gamma))$ constructed from δ_K. For $k \in \mathbb{N}$ and $b \in \mathrm{A}(\Gamma, X, \mu)$ let

$$E_k(b, K) = \{(\Phi^{\phi,a})_*\mu : \phi : X \to K \text{ is measurable and the range of } \phi \text{ has size } \leq k\}.$$

By Lemma 2.3.1 we can choose $k \in \mathbb{N}$ such that $E(b, K) \subseteq B_{\frac{\epsilon}{4}}(E_k(b, K))$ for every $b \in \mathrm{A}(\Gamma, X, \mu)$ where $B_r(A) = \{\nu \in M_s(K^\Gamma) : \delta_K(\nu, \rho) < r \text{ for some } \rho \in A\}$. Notice that $E_k(b, K) = \bigcup_{\substack{L \subseteq K, \\ |L| = k}} E(b, L)$. Fix a set L of size k and choose N large enough such that if $n \geq N$ then $d_{\mathcal{K}_L}\left(\overline{E(a_n, L)}, \overline{E(a, L)}\right) < \frac{\epsilon}{4}$ where $d_{\mathcal{K}_L}$ is the Hausdorff distance in $\mathcal{K}(M_s(L^\Gamma))$. Since the construction is independent of the set chosen to realize L, we have in fact $d_{\mathcal{K}_L}\left(\overline{E(a_n, L)}, \overline{E(a, L)}\right) < \frac{\epsilon}{4}$ for every finite set L of size k. For a fixed finite $L \subseteq K$ let $E_L(b, K) = \{(\Phi^{b,\phi})_*\mu : \phi : X \to K \text{ measurable}, \phi(X) \subseteq L\}$. Then for any $b, c \in \mathrm{A}(\Gamma, X, \mu)$ we have

$$d_{\mathcal{K}}\left(\overline{E_L(b, K)}, \overline{E_L(c, K)}\right) = d_{\mathcal{K}_L}\left(\overline{E(b, L)}, \overline{E(c, L)}\right),$$

so that when $n \geq N$,

$$d_{\mathcal{K}}\left(\overline{E_k(a_n, K)}, \overline{E_k(a, K)}\right) = d_{\mathcal{K}}\left(\bigcup_{\substack{L \subseteq K \\ |L|=k}} \overline{E(a_n, L)}, \bigcup_{\substack{L \subseteq K \\ |L|=k}} \overline{E(a, L)}\right)$$
$$\leq \sup_{\substack{L \subseteq K \\ |L|=k}} d_{\mathcal{K}_L}\left(\overline{E(a_n, L)}, \overline{E(a, L)}\right) < \frac{\epsilon}{4}.$$

Therefore when $n \geq N$,

$$\begin{aligned} d_{\mathcal{K}}\left(\overline{E(a_n, K)}, \overline{E(a, K)}\right) &\leq d_{\mathcal{K}}\left(\overline{E(a_n, K)}, \overline{E_k(a_n, K)}\right) + d_{\mathcal{K}}\left(\overline{E_k(a_n, K)}, \overline{E_k(a, K)}\right) \\ &\quad + d_{\mathcal{K}}\left(\overline{E_k(a, K)}, \overline{E(a, K)}\right) \\ &< \frac{3\epsilon}{4}. \end{aligned}$$

□

Lemma 2.3.3. *Let L be a finite set of size k. Then for each finite set $(A_p)_{p=1}^q$ of basic clopen sets $A_p \subseteq L^\Gamma$ and $\epsilon > 0$ there is $\delta > 0$ such that if $d(a, b) < \delta$ then for all $\phi \in L(X, \mu, L)$ there exists $\psi \in L(X, \mu, L)$ such that $|(\Phi_L^{a,\phi})_*\mu(A_p) - (\Phi_L^{b,\psi})_*\mu(A_p)| < \epsilon$ for all $p \leq q$.*

Proof. Write $A_p = \bigcap_{\gamma \in F_p} \pi_\gamma^{-1}(\jmath_p(\gamma))$ for some $F_p \subseteq \Gamma$ finite, $\jmath : F_p \to k$ and fix $\epsilon > 0$. Choose a finite $F \subseteq \Gamma$ with $(F_p)^2 \subseteq F$ for all $p \leq q$. We may assume the identity $e \in F$. Suppose $d(a, b) < \frac{\delta}{2^{|F|+k^{|F|}}}$; we will specify a value for δ later. Now fix $\phi : X \to k$ and let $B_i = \phi^{-1}(i)$. Given $\eta : F \to k$, let $B_\eta = \bigcap_{\gamma \in F} \gamma^a B_{\eta(\gamma)}$. We can then find a partition $\{D_\eta\}_{\eta \in k^F}$ such that

$$|\mu(\gamma^a B_{\eta_1} \cap B_{\eta_2}) - \mu(\gamma^b D_{\eta_1} \cap D_{\eta_2})| < \delta$$

for all $\eta_1, \eta_2 \in k^F$ and $\gamma \in F$. Define $\psi : X \to k$, by $\psi(y) = l$ if $y \in D_\eta$ for some η with $\eta(e) = l$. Furthermore, for each $l \leq k$ let $D_l = \bigsqcup\{D_\eta : \eta \in k^F$ and $\eta(e) = l\} = \psi^{-1}(l)$. For each $J \subseteq F$ and $\sigma \in k^J$ let $D_\sigma = \bigsqcup\{D_\eta : \eta \in k^F$ and $\sigma \sqsubseteq \eta\}$, where $\sigma \sqsubseteq \eta$ means η extends σ and let $\tilde{D}_\sigma = \bigcap_{\gamma \in J} \gamma^b D_{\sigma(\gamma)}$. Furthermore if $\gamma \in \Gamma$, $J \subseteq \Gamma$ and $\sigma \in k^J$ let $\gamma \cdot \sigma \in k^{\gamma J}$ be given by $(\gamma \cdot \sigma)(\delta) = \sigma(\gamma^{-1}\delta)$. For $\sigma \in K^{F_p}$ and $\gamma \in F_p$ we have

$$\begin{aligned} &|\mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma}) - \mu(\gamma^a B_\sigma \cap B_{\gamma\cdot\sigma})| \\ &\qquad \leq \sum_{(\eta \in k^F : \sigma \sqsubseteq \eta)} \sum_{(\eta' \in k^F : \gamma\cdot\sigma \sqsubseteq \eta')} |\mu(\gamma^b D_\eta \cap D_{\eta'}) - \mu(\gamma^a B_\eta \cap B_{\eta'})| \\ &\qquad \leq \delta(k^{|F|})^2. \end{aligned}$$

In particular, setting $\gamma = e$ we see $|\mu(B_\sigma) - \mu(D_\sigma)| < \delta k^{2|F|}$ for every $\sigma : F_p \to k$. Since $\gamma^a B_\sigma = B_{\gamma\cdot\sigma} = \gamma^a B_\sigma \cap B_{\gamma\cdot\sigma}$ we have

$$\begin{aligned}|\mu(D_\sigma) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| &\le |\mu(D_\sigma) - \mu(\gamma^a B_\sigma)| \\ &\qquad + |\mu(\gamma^a B_\sigma \cap B_{\gamma\cdot\sigma}) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| \\ &= |\mu(D_\sigma) - \mu(B_\sigma)| \\ &\qquad + |\mu(\gamma^a B_\sigma \cap B_{\gamma\cdot\sigma}) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| \\ &< 2\delta k^{2|F|}\end{aligned}$$

and also

$$\begin{aligned}|\mu(D_{\gamma\cdot\sigma}) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| &\le |\mu(D_{\gamma\cdot\sigma}) - \mu(B_{\gamma\cdot\sigma})| \\ &\qquad + |\mu(\gamma^a B_\sigma \cap B_{\gamma\cdot\sigma}) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| \\ &< 2\delta k^{2|F|}.\end{aligned}$$

Therefore

$$\begin{aligned}\mu((\gamma^b D_\sigma)\triangle(D_{\gamma\cdot\sigma})) &= \mu(\gamma^b D_\sigma) + \mu(D_{\gamma\cdot\sigma}) - 2\mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma}) \\ &\le |\mu(D_{\gamma\cdot\sigma}) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| + |\mu(D_{\gamma\cdot\sigma}) - \mu(\gamma^b D_\sigma \cap D_{\gamma\cdot\sigma})| \\ &< 4\delta k^{2|F|}. \end{aligned} \tag{2.2}$$

Since $(D_\eta)_{\eta\in k^F}$ is a partition of X and $(F_p)^2 \subseteq F$, we have

$$D_{J_p} = \bigsqcup_{\substack{\eta\in k^F \\ J_p\sqsubseteq\eta}} D_\eta = \bigcap_{\gamma\in F_p} \bigsqcup_{\substack{\sigma\in k^{\gamma F_p} \\ \sigma(\gamma)=J_p(\gamma)}} D_\sigma = \bigcap_{\gamma\in F_p} \bigsqcup_{\substack{\sigma\in k^{F_p} \\ \sigma(e)=J_p(\gamma)}} D_{\gamma\cdot\sigma}.$$

Now, by (2),

$$\mu\left(\left(\bigcap_{\gamma\in F_p} \bigsqcup_{\substack{\sigma\in k^{F_p} \\ \sigma(e)=J(\gamma)}} D_{\gamma\cdot\sigma}\right) \triangle \left(\bigcap_{\gamma\in F_p} \bigsqcup_{\substack{\sigma\in k^{F_p} \\ \sigma(e)=J_p(\gamma)}} \gamma^b D_\sigma\right)\right) < (|F_p|k^{|F_p|})(4\delta k^{2|F|}). \tag{2.3}$$

Note that

$$\bigcap_{\gamma\in F_p} \bigsqcup_{\substack{\sigma\in k^{F_p} \\ \sigma(e)=J_p(\gamma)}} \gamma^b D_\sigma = \bigcap_{\gamma\in F_p} \gamma^b D_{J_p(\gamma)} = \tilde{D}_{J_p},$$

so (3) reads

$$|\mu(D_{J_p}) - \mu(\tilde{D}_{J_p})| < (|F_p|k^{|F_p|})(4\delta k^{2|F|}).$$

Moreover,

$$\begin{aligned}
(\Phi_L^{b,\psi})_*\mu(A_p) &= \mu(\{x : \Phi_L^{b,\psi}(x) \in A_p\}) \\
&= \mu(\{x : \Phi_L^{b,\psi}(x)(\gamma) = \jmath_p(\gamma) \text{ for all } \gamma \in F_p\}) \\
&= \mu(\{x : \psi((\gamma^{-1})^b x) = \jmath_p(\gamma) \text{ for all } \gamma \in F_p\}) \\
&= \mu(\{x : x \in \gamma^b \psi^{-1}(\jmath_p(\gamma)) \text{ for all } \gamma \in F_p\}) \\
&= \mu\left(\bigcap_{\gamma \in F_p} \gamma^b D_{\jmath_p(\gamma)}\right) \\
&= \mu(\tilde{D}_{\jmath_p}).
\end{aligned}$$

Similarly, $(\Phi_L^{a,\phi})_*\mu(A_p) = \mu(B_{\jmath_p})$. So we finally have

$$\begin{aligned}
|(\Phi_L^{b,\psi})_*\mu(A_p) - (\Phi_L^{a,\phi})_*\mu(A_p)| &= |\mu(\tilde{D}_{\jmath_p}) - \mu(B_{\jmath_p})| \\
&\leq |\mu(\tilde{D}_{\jmath_p}) - \mu(D_{\jmath_p})| + |\mu(D_{\jmath_p}) - \mu(B_{\jmath_p})| \\
&< (|F_p| k^{|F_p|})(4\delta k^{2|F|}) + 2\delta k^{2|F|}.
\end{aligned}$$

Since k is fixed in advance, $|F_p| \leq |F|$ and F depends only on $(A_p)_{p=1}^q$, it is clear that δ can be chosen so $(|F_p| k^{|F_p|})(4\delta k^{2|F|}) + 2\delta k^{2|F|} < \epsilon$ for all $p \leq q$. □

We can now prove the main result of this section.

Theorem 2.3.1. $\tau_1 = \tau_2$.

Proof. Suppose that $a_n \to a$ in τ_1. We need to prove $\Phi(a_n) \to \Phi(a)$ in $\mathcal{K}(M_s(K^\Gamma))$. By Lemma 2.3.2 it suffices to fix a finite set L and show $\overline{E(a_n, L)} \to \overline{E(a, L)}$ in $\mathcal{K}(M_s(L^\Gamma))$. Let $k = |L|$. Write $E_n = E(a_n, L)$ and $E = E(a, L)$. As before, if we let $\mathcal{A}_L = (A_i^L)_{i=1}^\infty$ be the collection of clopen subsets of L^Γ of the form $\bigcap_{\gamma \in F} \pi_\gamma^{-1}(j_\gamma)$ for a finite $F \subseteq \Gamma$ and $j_\gamma \leq k$, then

$$\delta_L(\nu, \rho) = \sum_{i=1}^{\infty} \frac{1}{2^i} |\nu(A_i^L) - \rho(A_i^L)|$$

is a compatible metric on $M_s(L^\Gamma)$. Fix $\epsilon > 0$ in order to show that eventually $d_L(\overline{E_n}, \overline{E}) < \epsilon$, where d_L is the Hausdorff distance in $\mathcal{K}(M_s(L^\Gamma))$ constructed from δ_L. Choose N sufficiently large that $\sum_{i=N}^\infty \frac{1}{2^i} < \frac{\epsilon}{2}$. By Lemma 2.3.3 there is $\delta > 0$ such that if $d(a, b) < \delta$ then for each $i \leq N$ and all $\phi \in L(X, \mu, L)$ there exists $\psi \in L(X, \mu, L)$ such that $|(\Phi_L^{a,\phi})_*\mu(A_i^L) - (\Phi_L^{b,\psi})_*\mu(A_i^L)| < \frac{\epsilon}{2}$. Thus if M is large enough that $d(a_n, a) < \delta$ for $n \geq M$, we have $d_L(\overline{E_n}, \overline{E}) < \epsilon$.

Now suppose $\Phi(a_n) \to \Phi(a)$ in $\mathcal{K}(M_s(K^\Gamma))$. Fix r, q and $\epsilon > 0$ in order to show that eventually $d_H(C_{r,q}(a_n), C_{r,q}(a)) < \epsilon$. Choose q distinct points $(x_p)_{p=1}^q \in K$ and let $(D_p)_{p=1}^q$ be a family of disjoint clopen subsets of K with $x_p \in D_p$. Now let M be large enough that all sets of the form $\pi_{\gamma_s}^{-1}(D_p) \cap \pi_e^{-1}(D_t)$ for $s \leq r$ and $p, t \leq q$ appear as some A_i^K for $i \leq M$ in our previously chosen clopen basis $\mathcal{A}_K$. Then choose N large enough that when $n \geq N$, $d_{\mathcal{K}}(\Phi(a_n), \Phi(a)) < \frac{\epsilon}{2^M}$. Then for each $\phi \in L(X, \mu, K)$ we have $\psi \in L(X, \mu, K)$ such that $\delta_K((\Phi^{a_n,\phi})_*\mu, (\Phi^{a,\psi})_*\mu) < \frac{\epsilon}{2^M}$. So in particular, if $n \geq N$ then for each $\phi \in L(X, \mu, K)$ there exists $\psi \in L(X, \mu, K)$ such that

$$|(\Phi^{a_n,\phi})_*\mu(\pi_{\gamma_s}^{-1}(D_p) \cap \pi_e^{-1}(D_t)) - (\Phi^{a,\psi})_*\mu(\pi_{\gamma_s}^{-1}(D_p) \cap \pi_e^{-1}(D_t))| < \epsilon$$

for all $p, t \leq q$ and $s \leq r$.

Now suppose $n \geq N$ and let $(B_p)_{p=1}^q$ be a partition of X. Define $\phi : X \to K$ by taking $\phi(x) = x_p$ for the unique $p \leq q$ with $x \in B_p$ so by the previous paragraph we have a corresponding ψ. Observe that for all $\gamma \in \Gamma$ we have

$$\begin{aligned}
\mu(\gamma^{a_n} B_p \cap B_t) &= \mu(\gamma^{a_n} \phi^{-1}(D_p) \cap \phi^{-1}(D_t)) \\
&= \mu(\{x : \phi((\gamma^{a_n})^{-1}x) \in D_p \text{ and } \phi(x) \in D_t\}) \\
&= \mu(\{x : \Phi^{\phi,a_n}(x)(\gamma) \in D_p \text{ and } \Phi^{\phi,a_n}(x)(e) \in D_t\}) \\
&= \mu(\{x : \Phi^{\phi,a_n}(x) \in \pi_\gamma^{-1}(D_p) \text{ and } \Phi^{\phi,a_n}(x) \in \pi_e^{-1}(D_t)\}) \\
&= \mu(\{x : \Phi^{\phi,a_n}(x) \in \pi_\gamma^{-1}(D_p) \cap \pi_1^{-1}(D_t)\}) \\
&= (\Phi^{\phi,a_n})_*\mu(\pi_\gamma^{-1}(D_p) \cap \pi_1^{-1}(D_t)).
\end{aligned}$$

Similarly letting $H_p = \psi^{-1}(D_p)$ we have

$$\mu(\gamma^a H_p \cap H_t) = (\Phi^{\psi,a_n})_*\mu(\pi_\gamma^{-1}(D_p) \cap \pi_1^{-1}(D_t)).$$

Thus for all $p, t \leq q$ and $s \leq r$,

$$\begin{aligned}
&|\mu(\gamma_s^{a_n} B_p \cap B_t) - \mu(\gamma_s^a H_p \cap H_t)| \\
&\qquad = |(\Phi^{\phi,a_n})_*\mu(\pi_{\gamma+s}^{-1}(D_p) \cap \pi_e^{-1}(D_t)) - (\Phi^{\psi,a_n})_*\mu(\pi_{\gamma_s}^{-1}(D_p) \cap \pi_e^{-1}(D_t))| \\
&\qquad < \epsilon.
\end{aligned}$$

We have shown that when $n \geq N$, $C_{r.q}(a_n) \subseteq B_\epsilon(C_{r,q}(a))$. The argument that eventually $C_{r,q}(a) \subseteq B_\epsilon(C_{r,q}(a_n))$ is identical. $\square$

Topology on the space of stable weak equivalence classes

Let $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ be the space of stable weak equivalence classes and let ι be the trivial action of Γ on an standard probability space. By Lemma 3.7 in [74], we have $a \prec_s b$ if and only if $a \prec \iota \times b$. Moreover, Theorem 1.1 in [74] says that $\overline{E(a \times \iota, K)} = \mathrm{cch}(E(a, K))$, where $M_s(K^\Gamma)$ carries its natural topological convex structure as a compact convex subset of a Banach space. Letting $\Psi : \mathrm{A}(\Gamma, X, \mu) \to \mathcal{K}(M_s(K^\Gamma))$ be the map $a \mapsto \mathrm{cch}(E(a, K))$ we have $\Psi(a) = \Psi(b)$ if and only if $a \sim_s b$. Tucker-Drob gives $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ the initial topology induced by Ψ, in which it is a compact Polish space. Thus we have $a_n \to a$ in the topology of $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ if and only if $a_n \times \iota \to a \times \iota$ in the topology of $\mathrm{A}_{\sim}(\Gamma, X, \mu)$. Therefore we can introduce a metric d_s on $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ by setting $d_s(a, b) = d(a \times \iota, b \times \iota)$.

2.4 The space of weak equivalence classes as a weak convex space

We now describe how to give $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ the structure of a weak convex space. Given $t \in [0, 1]$ and $a, b \in \mathrm{A}_{\sim}(\Gamma, X, \mu)$ we let $c \in A(\Gamma, X_1 \sqcup X_2, t\mu_1 + (1-t)\mu_2)$ be the disjoint sum of representative actions a and b on the disjoint union of two copies X_1 and X_2 of X with the first copy carrying a copy of the measure μ weighted by t and the second copy carrying a copy of μ weighted by $(1-t)$. To get an action in $\mathrm{A}(\Gamma, X, \mu)$ we need to choose an isomorphism of (X, μ) with $(X_1 \sqcup X_2, t\mu_1 + (1-t)\mu_2)$, but the weak equivalence class of c does not depend on this or on the representatives we chose. So we have a well-defined binary operation $\mathrm{A}_{\sim}(\Gamma, X, \mu)^2 \to \mathrm{A}_{\sim}(\Gamma, X, \mu)$. Call this cc_t. It is clear that (1), (3) and (4) of Definition 2.2.1 are satisfied, so $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ is a weak convex space. Moreover, we have the following.

Proposition 2.4.1. $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ *is a topological weak convex space.*

Proof. We must show that cc is continuous. Suppose that $t_j \to t$ in $[0, 1]$ and $a_j \to a$ and $b_j \to b$ in the topology of $A_{\sim}(\Gamma, Y, \mu)$. Write $c_j = t_j a_j + (1-t_j)b_j$ and $c = ta + (1-t)b$. Fixing $l, m \in \mathbb{N}$ write $C(d)$ for $C_{l,m}(d)$. We need to prove that for every $\epsilon > 0$ there is J so that if $j > J$ then we have $d_H(C(c_j), C(c)) < \epsilon$, where d_H is the Hausdorff distance in $[0, 1]^{l \times m^2}$.

First we must show that for sufficiently large j, for every partition $B_1, \ldots, B_l$ of Y there is a partition $D_1, \ldots, D_l$ of Y depending on j such that for all $s, t \leq l$ and $p \leq m$,

$$|\mu(\gamma_p^{c_j} D_s \cap D_t) - \mu(\gamma_p^{c} B_s \cap B_t)| < \epsilon.$$

Choose J_1 so that if $j > J_1$ then $|t_j - t| < \frac{\epsilon}{6}$. Choose $J_2 > J_1$ so if $j > J_2$ then $d_H(C_{a_j}, C_a) < \frac{\epsilon}{6}$ and $d_H(C_{b_j}, C_b) < \frac{\epsilon}{6}$. Fix $j > J_2$. Writing θ for the isomorphism from $(Y_1 \sqcup Y_2, t\mu + (1-t)\mu)$ to (Y, μ) and θ_j for the isomorphism from $(Y_1 \sqcup Y_2, t_j\mu + (1-t_j)\mu)$ to (Y, μ) we have a partition $(B_{s,i})_{s=1}^l$ of Y_i given by $B_{s,i} = \theta^{-1}(B_s) \cap Y_i$. So we can find a partition $(D_{s,i})_{s=1}^l$ of Y_i such that for all $p \le m$ and all $s, t \le l$ we have

$$|\mu(\gamma_p^{a_j} D_{s,1} \cap D_{t,1}) - \mu(\gamma_p^a B_{s,1} \cap B_{t,1})| < \frac{\epsilon}{6}$$

and

$$|\mu(\gamma_p^{b_j} D_{s,2} \cap D_{t,2}) - \mu(\gamma_p^b B_{s,2} \cap B_{t,2})| < \frac{\epsilon}{6}.$$

Now, let $D_s = \theta_j(D_{s,1} \sqcup D_{s,2})$. Note that since each $\theta_j(Y_i)$ is c_j invariant,

$$\begin{aligned}
\mu(\gamma_p^{c_j} D_s \cap D_t) &= \mu(\gamma_p^{c_j}\theta_j(D_{s,1}) \cap \theta_j(D_{t,1})) + \mu(\gamma_p^{c_j}\theta_j(D_{s,2}) \cap \theta_j(D_{t,2})) \\
&= \mu(\theta_j(\gamma_p^{a_j} D_{s,1} \cap D_{t,1})) + \mu(\theta_j(\gamma_p^{b_j} D_{s,2} \cap D_{t,2})) \\
&= t_j\mu(\gamma_p^{a_j} D_{s,1} \cap D_{t,1}) + (1-t_j)\mu(\gamma_p^{b_j} D_{s,2} \cap D_{t,2}).
\end{aligned}$$

Similarly since $\theta(Y_i)$ is c-invariant we have

$$\begin{aligned}
\mu(\gamma_p^c B_s \cap B_t) &= \mu(\gamma_p^c\theta(B_{s,1}) \cap \theta(B_{t,1})) + \mu(\gamma_p^c\theta(B_{s,2}) \cap \theta(B_{t,2})) \\
&= \mu(\theta(\gamma_p^a B_{s,1} \cap B_{t,1})) + \mu(\theta(\gamma_p^b B_{s,2} \cap B_{t,2})) \\
&= t\mu(\gamma_p^a \cap B_{s,1} \cap B_{t,1}) + (1-t)\mu(\gamma_p^b \cap B_{s,2} \cap B_{t,2}).
\end{aligned}$$

Note that if $|x_1 - x_2| < \delta$ and $|y_1 - y_2| < \delta$ then $|x_1y_1 - x_2y_2| < 3\delta$. So our assumptions guarantee that we have

$$|t_j\mu(\gamma_p^{a_j} D_{s,1} \cap D_{t,1}) - t\mu(\gamma_p^a \cap B_{s,1} \cap B_{t,1})| < \frac{\epsilon}{2}$$

and

$$|(1-t_j)\mu(\gamma_p^{b_j} D_{s,2} \cap D_{t,2}) - (1-t)\mu(\gamma_p^b \cap B_{s,2} \cap B_{t,2})| < \frac{\epsilon}{2},$$

and hence

$$|\mu(\gamma_p^{c_j} D_s \cap D_t) - \mu(\gamma_p^c B_s \cap B_t)| < \epsilon$$

as claimed.

Now we must show that for sufficiently large j, every partition $B_1, \ldots, B_l$ of Y there is a partition $D_1, \ldots, D_l$ of Y depending on j such that for all $s, t \le l$ and $p \le m$ we have

$$|\mu(\gamma_p^c D_s \cap D_t) - \mu(\gamma_p^{c_j} B_s \cap B_t)| < \epsilon.$$

The argument is similar to the previous step, so we omit it. □

Corollary 2.4.1. *$A_\sim(\Gamma, Y, \mu)$ is path connected.*

Corollary 2.4.2. *$A_\sim(\Gamma, Y, \mu)$ is uncountable.*

We now record a lemma which will be useful later, guaranteeing that the metric on $\mathrm{A}_\sim(\Gamma, X, \mu)$ behaves nicely with respect to the convex structure.

Lemma 2.4.1. *For any convex set $K \subseteq \mathrm{A}_\sim(\Gamma, X, \mu)$ the function $d(\cdot, K) = \inf_{b\in K} d(\cdot, b)$ is convex.*

Proof. Let $x, y \in \mathrm{A}_\sim(\Gamma, X, \mu)$ and consider $tx + (1-t)y$. Fix n, k and write $C(a)$ for $C_{n,k}(a)$. It suffices to show that

$$\inf_{b\in K} d_H(C(tx + (1-t)y), C(b)) \leq t(\inf_{b\in K} d_H(C(x), C(b)) + (1-t)(\inf_{b\in K} d_H(C(y), C(b))),$$

where d_H is the Hausdorff distance in the space $[0, 1]^{n\times k^2}$. Fix $\epsilon > 0$. It suffices to find $a \in K$ with

$$d_H(C(tx+(1-t)y), C(a)) \leq t(\inf_{b\in K} d_H(C(x), C(b))+\epsilon)+(1-t)(\inf_{b\in K} d_H(C(y), C(b))+\epsilon). \tag{2.4}$$

Choose $c \in K$ with $d_H(C(x), C(c)) < \inf_{b\in K} d_H(C(x), C(b)) + \epsilon$ and choose $d \in K$ with $d_H(C(x), C(d)) < \inf_{b\in K} d_H(C(y), C(b)) + \epsilon$. Note that since K is convex, $tc + (1-t)d \in K$. We claim

$$d_H(C(tx + (1-t)y), C(tc + (1-t)d)) \leq td_H(C(x), C(c)) + (1-t)d_H(C(y), C(d)),$$

which implies (4). Let $\delta > 0$, it then suffices to show

$$d_H(C(tx+(1-t)y), C(tc+(1-t)d)) \leq t(d_H(C(x), C(c))+\delta)+(1-t)(d_H(C(y), C(d))+\delta). \tag{2.5}$$

Let X_1 and X_2 be two copies of X and ν be the measure on $X_1 \sqcup X_2$ given by $t(\mu \upharpoonright X_1)+(1-t)(\mu \upharpoonright X_2)$. Let $\mathcal{P} = (P_i)_{i=1}^k$ be a partition of $X_1 \sqcup X_2$. This induces a partition $\mathcal{P}_1 = (P_i^1)_{i=1}^k$ of X_1 given by $P_i^1 = P_i \cap X_1$ and similarly we have a partition $\mathcal{P}_2 = (P_i^2)_{i=1}^k$ of X_2. We can find a partition $\mathcal{Q}_1 = (Q_i^1)_{i=1}^k$ of X_1 such that for $m \leq n$ and $i, j \leq k$ we have

$$|\mu(\gamma_m^x P_i^1 \cap P_j^1) - \mu(\gamma_m^c Q_i^1 \cap Q_j^1)| < d_H(C(x), C(c)) + \delta$$

and similarly we can find a partition $Q_2 = (Q_i^2)_{i=1}^k$ of X_2 such that for $m \leq n$ and $i, j \leq k$ we have

$$|\mu(\gamma_m^y P_i^2 \cap P_j^2) - \mu(\gamma_m^d Q_i^2 \cap Q_j^2)| < d_H(C(y), C(d)) + \delta.$$

Let $Q = (Q_i)_{i=1}^k$ be the partition of $X_1 \sqcup X_2$ given by $Q_i = Q_i^1 \sqcup Q_i^2$. Write $t(d_H(C(x), C(c)) + \delta) + (1-t)(d_H(C(y), C(d)) + \delta) = r$. Then for all $m \leq n$ and $i, j \leq k$ we have

$$\begin{aligned}
&|\nu(\gamma_m^{tx+(1-t)y} P_i \cap P_j) - \nu(\gamma_m^{tc+(1-t)d} Q_i \cap Q_j)| \\
&\qquad \leq |t\mu(\gamma_m^x P_i^1 \cap P_j^1) - t\mu(\gamma_m^c Q_i^1 \cap Q_j^1)| \\
&\qquad\quad + |(1-t)\mu(\gamma_m^y P_i^2 \cap P_j^2) - (1-t)\mu(\gamma_m^d Q_i^2 \cap Q_j^2)| \\
&\qquad \leq r.
\end{aligned}$$

We have shown that $C(tx + (1-t)y) \subseteq B_r(C(tc + (1-t)d))$. The argument that $C(tc + (1-t)d) \subseteq B_r(C(tx + (1-t)y))$ is identical, so we omit it. Thus we conclude $d_H(C(tx + (1-t)y), C(tc + (1-t)d)) \leq r$ and (5) holds. □

We note that $\mathrm{A}_\sim(\Gamma, X, \mu)$ in fact has additional structure in that it admits convex combinations of infinitely many elements. We first consider the case of a countable convex combination. If $\lambda_i \in [0, 1]$ are such that $\sum_{i=1}^\infty \lambda_i = 1$ and $a_i \in \mathrm{A}_\sim(\Gamma, X, \mu)$ then we can naturally define an action $\sum_{i=1}^\infty \lambda_i a_i$ on the disjoint sum $\bigsqcup_{i=1}^\infty X_i$ with the i copy of X weighted by λ_i. It remains to check that this is independent of the choice of representatives a_i.

Proposition 2.4.2. *If $a_i \prec b_i$ for all i, then $\sum_{i=1}^\infty \lambda_i a_i \prec \sum_{i=1}^\infty \lambda_i b_i$.*

Proof. Let $A_1, \ldots, A_k \subseteq \bigsqcup_{m=1}^\infty X_m$, $\epsilon > 0$ and $F \subseteq \Gamma$ finite be given. Choose N such that $\sum_{m=N}^\infty \lambda_m < \frac{\epsilon}{2}$. For each $m < N$, consider the partition $A_1^m, \ldots, A_k^m$ of X_m given by $A_i^m = A_i \cap X_m$. We can find for each $m < N$ a partition $B_1^m, \ldots, B_k^m$ such that for all $\gamma \in F$ and $i, j \leq k$ we have

$$|\mu(\gamma^{a_i} A_i^m \cap A_j^m) - \mu(\gamma^{b_i} B_i^m \cap B_j^m)| < \frac{\epsilon}{2}.$$

Let $B_i = \bigsqcup_{m=1}^{\infty} B_i^m$. Then

$$\begin{aligned}
&\left|\mu\left(\gamma^{\sum_{m=1}^{\infty}\lambda_m a_m} A_i \cap A_j\right) - \mu\left(\gamma^{\sum_{m=1}^{\infty}\lambda_m b_m} B_i \cap B_j\right)\right| \\
&\qquad \leq |\sum_{m=1}^{N} \lambda_m \mu(\gamma^{a_m} A_i^m \cap A_j^m) - \sum_{m=1}^{N} \lambda_m \mu(\gamma^{b_m} B_i^m \cap B_j^m)| \\
&\qquad\quad + |\sum_{m=M}^{\infty} \lambda_m \mu(\gamma^{a_m} A_i^m \cap A_j^m) - \sum_{m=M}^{\infty} \lambda_m \mu(\gamma^{b_m} B_i^m \cap B_j^m)| \\
&\qquad \leq \sum_{m=1}^{N} \lambda_m |\mu(\gamma^{a_i} A_i^m \cap A_j^m) - \mu(\gamma^{b_i} B_i^m \cap B_j^m)| + \frac{\epsilon}{2} \\
&\qquad \leq \frac{\epsilon}{2}\left(\sum_{m=1}^{N} \lambda_m\right) + \frac{\epsilon}{2} \leq \epsilon.
\end{aligned}$$

□

It is in fact possible to define integrals of weak equivalence classes of actions over a probability measure. Let (Z, η) be a probability space and suppose that for each z we have a probability space (Y_z, ν_z) and a measure-preserving action $\Gamma \curvearrowright^{a_z} (Y_z, \nu_z)$ such that the map $z \mapsto [a_z]$ from (Z, η) to $A_{\sim}^*(\Gamma)$ is measurable, where $[a_z]$ is the weak equivalence class of a_z. Note that we do not require (X_z, ν_z) or (Z, η) to be standard. Let $Y = \bigsqcup_{z \in Z} Y_z$ and put a measure ν on Y by taking $\nu(A) = \int_Z \nu_z(A \cap Y_Z) d\eta(z)$. Y will be a standard probability space isomorphic to (X, μ) if (Z, η) is standard or η-almost all (Y_z, ν_z) are standard. Let $\Gamma \curvearrowright^a (Y, \nu)$ be given by letting Γ act like a_z on Y_z. We write $a = \int_Z a_z d\eta(z)$. We then have a map $\phi : Y \to Z$ given by letting $\phi(y)$ be the unique z such that $y \in Y_z$. This is clearly a factor map from a to $\iota_{Z,\eta}$ and $\nu = \int_Z \nu_z d\eta(z)$ is the disintegration of ν over η via ϕ. Thus Theorem 3.12 in [74] guarantees that if b_z are actions of Γ on (Y_z, ν_z) with $b_z \sim a_z$ then if $b = \int_Z b_z d\eta(z)$ we have $a \sim b$. Therefore this construction gives a well-defined weak equivalence class of actions of γ. If we restrict (Y_z, ν_z) to be standard, then we in fact have a mapping from the space $M(\mathrm{A}_{\sim}(\Gamma, X, \mu))$ of probability measures on $\mathrm{A}_{\sim}(\Gamma, X, \mu)$ to $\mathrm{A}_{\sim}(\Gamma, X, \mu)$.

Lemma 2.4.2. *For any n, k, and (Z, η) and measurable assignment $z \mapsto a_z$, we have $C_{n,k}\left(\int_Z a_z d\eta(z)\right) \subseteq \mathrm{cch}\left(\bigcup_{z \in Z} C_{n,k}(a_z)\right)$.*

Proof. Fix n, k and let $a = \int_Z a_z d\eta(z)$. Let (X_z, μ_z) by the underlying measure space of a_z. Let $\mathcal{L}$ be a countable dense subset of MALG $\left(\bigsqcup_{z \in Z} X_z, \int_Z \mu_z d\eta(z)\right)$, so that

$\mathcal{L}^k$ is dense in the space of k-partitions of $\bigsqcup_{z\in Z} X_z$. Then $\{M_{\mathcal{A}}(a)\}_{\mathcal{A}\in\mathcal{L}^n}$ is dense in $C_{n,k}\left(\int_Z a_z d\eta(z)\right)$, so it suffices to show that each $M_{\mathcal{A}}(a) \in \mathrm{cch}\left(\bigcup_{z\in Z} C(a_z)\right)$. For each $\mathcal{A}$, the function $f_{\mathcal{A}} : Z \to \mathbb{R}^{n\times k\times k}$ given by $z \mapsto M_{\mathcal{A}_z}(a_z)$ is a Borel function, where $\mathcal{A}_z$ is the partition of X_z given by $(A\cap X_z)_{A\in\mathcal{A}}$. Thus $M_{\mathcal{A}}(a) = \int_Z f_{\mathcal{A}}(a)d\eta(z)$. We may assume that Z carries a Polish topology such that $f_{\mathcal{A}}$ is continuous for all $\mathcal{A} \in \mathcal{L}^n$. Choose a sequence of measures $(\nu_i)_{i=1}^\infty$ such that ν_i has finite support and $\nu_i \to \eta$ in the topology of $M(Z)$, the space of all Borel probability measures on Z. If we write $\nu_i = \sum_{j=1}^{j(i)} \alpha_j \delta_{z_j}$ then

$$\int_Z f_{\mathcal{A}}(z)d\nu_i(z) = \sum_{j=1}^{j(i)} \alpha_j f_{\mathcal{A}}(z_j) \in \mathrm{ch}\left(\bigcup_{z\in A} C(a_z)\right).$$

Since $\nu_i \to \eta$, we have

$$\int_Z f_{\mathcal{A}}(z)d\nu_i(z) \to \int_Z f_{\mathcal{A}}(z)d\eta(z)$$

which proves the lemma. □

2.5 The structure of the space of weak equivalence classes for amenable acting groups

When Γ is amenable, the structure of $\mathrm{A}_\sim(\Gamma, X, \mu)$ can be completely described using the notion of an invariant random subgroup. We begin with the following, the following extends Theorem 1.8 in [74]. Recall that if $\Gamma \curvearrowright^a (X,\mu)$ is a measure-preserving action, we have a map $\mathrm{stab}_a : X \to \mathrm{Sub}(\Gamma)$ given by $x \mapsto \mathrm{stab}_a(x)$. The type of a is the invariant random subgroup of Γ given by $(\mathrm{stab}_a)_*\mu$.

Proposition 2.5.1. *If Γ is amenable and $a, b \in \mathrm{A}(\Gamma, X, \mu)$ then* $\mathrm{type}(a) = \mathrm{type}(b)$ *if and only if $a \sim b$.*

Proof. By [1] type is an invariant of weak equivalence so suppose $\mathrm{type}(a) = \mathrm{type}(b)$.

Let $X_\infty^a = \{x \in X : [\Gamma : \mathrm{stab}_a(x) = \infty]\}$ and $X_\infty^b = \{x \in X : [\Gamma : \mathrm{stab}_b(x)] = \infty\}$. Notice that X_∞^a is a-invariant and X_∞^b is b-invariant and since $\mathrm{type}(a) = \mathrm{type}(b)$, $\mu(X_\infty^a) = \mu(X_\infty^b)$. Suppose that $\mu(X_\infty^a) > 0$ and let $a_\infty = a \upharpoonright X_\infty^a$ with normalized measure $\frac{\mu\upharpoonright X_\infty^a}{\mu(X_\infty^a)}$ and define b_∞ similarly. Then $\mathrm{type}(a_\infty) = \mathrm{type}(b_\infty)$ and these are concentrated on the infinite index subgroups of Γ, therefore $a_\infty \sim b_\infty$ by Theorem 1.8 (2) in [74]. Thus to prove the proposition it suffices to show the following. Note that for this we do not require Γ to be amenable.

Lemma 2.5.1. *Suppose $a, b \in \mathrm{A}(\Gamma, X, \mu)$ are actions such that* $\mathrm{type}(a) = \mathrm{type}(b)$ *and these are concentrated on the finite-index subgroups of* Γ. *Then $a \sim b$.*

Proof. We may assume that $\theta = \mathrm{type}(a) = \mathrm{type}(b)$ is concentrated on the subgroups of index n for some fixed n. Consider an a-orbit C. For each linear ordering $<^i_C$ of C, we get a homomorphism $\psi^i_C : \Gamma \to \mathrm{Sym}(n)$, where $\mathrm{Sym}(n)$ is the symmetric group on n letters. Place a Borel linear order $\sqsubset$ on $\mathrm{Sym}(n)^\Gamma$. Let then $<^a_C = <^{i_0}_C$ be the linear order such that $\psi^{i_0}_C$ is $\sqsubset$-least among all the ψ^i_C. Write ϕ^a_C for $\psi^{i_0}_C$. Use this same construction to choose homomophisms ϕ^b_D for each b-orbit D. Write ϕ^a_x for $\phi^a_{[x]_{E_a}}$ and similarly ϕ^b_x for $\phi^b_{[x]_{E_b}}$.

For a homomorphism $\phi : \Gamma \to \mathrm{Sym}(n)$ let j_ϕ be the corresponding action of Γ on $\{1, \ldots, n\}$. Say ϕ is transitive if j_ϕ is transitive. Each transitive homomorphism $\phi : \Gamma \to \mathrm{Sym}(n)$ determines a conjugacy class $\mathcal{H}_\phi$ of index n subgroups of Γ as the stabilizers of j_ϕ. For each a-orbit $[x]_{E_a}$ the stabilizers of the action of Γ on $[x]_{E_a}$ also determine a conjugacy class $\mathcal{H}^a_x$ of index n subgroups of Γ. Let c be the action of $\mathrm{Sym}(n)$ on $\mathrm{Sym}(n)^\Gamma$ by $(f \cdot \phi)(\gamma)(k) = f\phi(\gamma)f^{-1}(k)$. Then $[\phi^a_x]_{E_c} = \left\{\psi^i_{[x]_{E_a}} : <^i_{[x]_{E_a}} \text{ is a linear ordering of } [x]_{E_a}\right\}$. Let $\mathcal{L}$ be the set of all transitive homomorphisms $\phi : \Gamma \to \mathrm{Sym}(n)$ such that ϕ is $\sqsubset$-least in $[\phi]_{E_c}$. It is clear that for $\phi \in \mathcal{L}$, $\phi^a_x = \phi$ if and only if $\mathcal{H}^a_x = \mathcal{H}_\phi$. Similarly $\phi^b_x = \phi$ if and only if $\mathcal{H}^b_x = \mathcal{H}_\phi$. Thus for any $A \subseteq \mathcal{L}$, we have

$$\begin{aligned}
\mu(\{x : \phi^a_x \in A\}) &= \mu(\{x : \mathcal{H}^a_x = \mathcal{H}_\phi \text{ for some } \phi \in A\}) \\
&= \mu(\{x : \mathrm{stab}_a(x) \text{ is conjugate to an element of} \\
&\qquad\qquad \mathcal{H}_\phi \text{ for some } \phi \in A\}) \\
&= \theta(\{H \in \mathrm{Sub}(\Gamma) : H \text{ is conjugate to an element of} \\
&\qquad\qquad \mathcal{H}_\phi \text{ for some } \phi \in A\}) \\
&= \mu(\{x : \mathrm{stab}_b(x) \text{ is conjugate to an element of} \\
&\qquad\qquad \mathcal{H}_\phi \text{ for some } \phi \in A\}) \\
&= \mu(\{x : \phi^b_x \in A\}).
\end{aligned}$$

Now, fix a finite set $F \subseteq \Gamma$ and a partition $A_1, \ldots, A_m$ of X. For each map $\omega : F \to \mathrm{Sym}(n)$ let $X^a_\omega = \{x \in X : \phi^a_x \upharpoonright F = \omega\}$ and similarly $X^b_\omega = \{x \in X : \phi^b_x \upharpoonright F = \omega\}$. Then $(X^a_\omega)_{\omega \in \mathrm{Sym}(n)^F}$ and $(X^b_\omega)_{\omega \in \mathrm{Sym}(n)^F}$ are finite decompositions of X with $\mu(X^a_\omega) = \mu(X^b_\omega)$. For $k \leq n$ let

$$X^a_{\omega,k} = \left\{x \in X^a_\omega : x \text{ is in the } k\text{-position with respect to } <^a_{[x]_{E_a}}\right\}$$

and define $X^b_{\omega,k}$ similarly. We claim that for each k there is a measure-preserving bijection $S^a_{\omega,k}$ of $X^a_{\omega,k}$ with $X^a_{\omega,1}$. Let $\sqsubset_1$ be a wellordering of Γ. For each $\gamma \in \Gamma$ let

$$X^a_{\omega,k,\gamma} = \left\{ x \in X^a_{\omega,k} : \text{ the } \sqsubset_1 \text{ - least } \delta \in \Gamma \text{ with } \delta^a x \in X^a_{\omega,1} \text{ is equal to } \gamma \right\}.$$

Put then $S^a_{\omega,k} \restriction X^a_{\omega,k,\gamma} = \gamma^a$. In particular, this shows that $\mu(X^a_{\omega,k}) = \frac{\mu(X^a_\omega)}{n}$. We can perform the same construction for b and we see that $\mu(X^b_{\omega,k}) = \frac{\mu(X^b_\omega)}{n}$. So $\mu(X^a_{\omega,1}) = \mu(X^b_{\omega,1})$ and hence there is a measure-preserving bijection $T_{\omega,1}$ of each $X^a_{\omega,1}$ with $X^b_{\omega,1}$. Define a measure-preserving bijection T_ω of X^a_ω with X^b_ω by letting $T_\omega(x) = (S^b_{\omega,k})^{-1} T S^a_{\omega,k}(x)$ for $x \in X^a_{\omega,k}$. Let then $T = \bigcup_{\omega \in \mathrm{Sym}(n)^F} T_\omega$ so $T \in \mathrm{Aut}(X, \mu)$.

We claim that for all $\gamma \in F$ and all $x \in X$, we have $T(\gamma^a x) = \gamma^b T(x)$. Indeed, suppose $x \in X^a_{\omega,k}$ so that x is in the k-position with respect to $<^a_{[x]_{E_a}}$. Then $\gamma^a x$ is in the $\phi^a_x(\gamma)(k) = \omega(k)$ position with respect to $<^a_{[x]_{E_a}}$ so $T(\gamma^a_x)$ is in the $\omega(k)$ position of the E_b-class D such that $T_{\omega,1} S^a_{\omega,k}(x) \in D$, where D has the canonical order $<^b_D$. On the other hand, $T(x) = T_\omega(x)$ is in the k-position of D with respect to $<^b_D$. Hence $\gamma^b T(x)$ is in the $\phi^b_{T(x)}(\gamma)(k) = \omega(k)$ position of D and we have the claim. Now, for $i \leq m$ putting $B_i = T(A_i)$ we have for any γ in F and $i, j \leq m$,

$$\begin{aligned} \mu(\gamma^b B_i \cap B_j) &= \mu(\gamma^b T(A_i) \cap T(A_j)) \\ &= \mu(T(\gamma^a A_i) \cap T(A_j)) \\ &= \mu(T(\gamma^a A_i \cap A_j)) \\ &= \mu(\gamma^a A_i \cap A_j) \end{aligned}$$

and therefore $a \sim b$. $\square$

$\square$

In [74], Tucker-Drob shows that for amenable Γ, the space $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ of stable weak equivalence classes is homeomorphic to the space $\mathrm{IRS}(\Gamma)$ of invariant random subgroups of Γ. Indeed, $\mathrm{type}(a) = \mathrm{type}(b)$ if and only if $a \sim_s b$ and the map $A_{\sim_s}(\Gamma, X, \mu) \longrightarrow \mathrm{IRS}(\Gamma)$ given by $a \mapsto \mathrm{type}(a)$ is a homeomorphism. So we have the following.

Corollary 2.5.1. *For amenable Γ, $a \sim_s b$ if and only if $a \sim b$.*

Moreover, let $x \in X$, $t \in [0,1]$ and $a, b \in \mathrm{A}(\Gamma, X, \mu)$ and consider the action $ta + (1-t)b$ on $tX_1 \sqcup (1-t)X_2$. We have $\mathrm{stab}_{ta+(1-t)b} = \mathrm{stab}_a(x)$ if $x \in X_1$ and $\mathrm{stab}_b(x)$ if $x \in X_2$. Thus for any $H \leq \Gamma$, $\{x : \mathrm{stab}_{ta+(1-t)b}(x) = H\} = \{x \in X_1 : \mathrm{stab}_a(x) = H\} \sqcup \{x \in X_2 : \mathrm{stab}_b(x) = H\}$ so for any $A \subseteq \mathrm{Sub}(\Gamma)$ we have

$$\begin{aligned}
(t\mu_1 + (1-t)\mu_2)(\{x : \mathrm{stab}_{ta+(1-t)b}(x) \in A\}) \\
&= (t\mu_1 + (1-t)\mu_2)(\{x \in X_1 : \mathrm{stab}_a(x) \in A\} \\
&\sqcup \{x \in X_2 : \mathrm{stab}_b(x) \in A\}) \\
&= t\mu(\{x : \mathrm{stab}_a(x) \in H\}) + (1-t)\mu(\{x : \mathrm{stab}_b(x) \in A\}).
\end{aligned}$$

Therefore $\mathrm{type}(ta + (1-t)b) = t(\mathrm{type}(a)) + (1-t)(\mathrm{type}(b))$ and Theorem 2.1.1 follows. Note in particular that if Γ is amenable then $ta + (1-t)a \sim a$, so for amenable groups $\mathrm{A}_\sim(\Gamma, X, \mu)$ is actually a convex space, not just a weak convex space.

It is known (see for example [35]) that $\mathrm{IRS}(\Gamma)$ is a simplex in $C(\mathrm{Sub}(\Gamma))^*$, the dual of the Banach space $C(\mathrm{Sub}(\Gamma))$ of continuous functions on $\mathrm{Sub}(\Gamma)$. So by the classical Krein-Milman theorem we have that for amenable Γ, $\mathrm{cch}(\mathrm{ex}(\mathrm{A}_\sim(\Gamma, X, \mu))) = \mathrm{A}_\sim(\Gamma, X, \mu)$. We will prove an analogous result for general Γ using other means. Moreover, $\mathrm{ex}(\mathrm{IRS}(\Gamma))$ is precisely the ergodic measures in $\mathrm{IRS}(\Gamma)$ so when Γ is amenable, $\mathrm{ex}(\mathrm{A}_\sim(\Gamma, X, \mu))$ is the set of actions with ergodic type.

2.6 The structure of the space of weak equivalence classes for general acting groups

Recall from [53] that E_0 is the equivalence relation given by eventual equality on $2^{\mathbb{N}}$ and if E is an equivalence relation on X and F is an equivalence relation on Y then a Borel homomorphism from E to F is a Borel map $f : X \to Y$ such that $x_1 E x_2$ implies $f(x_1) F f(x_2)$. A equivalence relation E on a measure space is said to be strongly ergodic (or E_0-ergodic) if for any homomorphism from E to E_0, the preimage of some E_0-class is conull. By Proposition 5.6 in [29] if a is strongly ergodic then every b with $b \sim a$ is ergodic. In particular, $\frac{1}{2}a + \frac{1}{2}a$ is not ergodic, so $\frac{1}{2}a + \frac{1}{2}a$ is not weakly equivalent to a when a is strongly ergodic. By Theorem 1.2 in [55], the Bernoulli shift $\Gamma \curvearrowright ([0,1]^\Gamma, \lambda^\Gamma)$ with λ Lebesgue measure on $[0,1]$ is strongly ergodic when Γ is nonamenable. Thus when Γ is nonamenable, $\mathrm{A}_\sim(\Gamma, X, \mu)$ is not a convex space, only a weak convex space. We now prove Theorem 2.1.2

Proof. **(of Theorem 2.1.2)** Write $A = \mathrm{A}_\sim(\Gamma, X, \mu)$. Let $B = \mathrm{cch}(\mathrm{ex}(A))$ and suppose toward a contradiction that there exists $x \in A \setminus B$. Since B is compact, $d(x, B) > 0$.

Let $\alpha = \sup_{y\in A} d(y, B)$ and let $C = \{y \in A : d(y, B) = \alpha\}$. Then C is nonempty, disjoint from B and C is a face of A.

Let $\mathcal{F}$ be the family of faces of C, ordered by reverse inclusion. Suppose $\{F_i\}_{i\in I}$ is a linearly ordered subset of $\mathcal{F}$ and consider $\bigcap_{i\in I} F_i$. If $x, y \in C$ and $0 < t < 1$ are such that $tx + (1-t)y \in \bigcap_{i\in I} F_i$, then $x, y \in F_i$ for each i since each F_i is a face. Hence $\bigcap_{i\in I} F_i$ is a face. It is nonempty by compactness. So Zorn's Lemma guarantees there exist minimal elements of $\mathcal{F}$. Let F be such a minimal element.

Choose $y \in F$ and suppose toward a contradiction that there exists $y' \in F$ with $y' \notin \mathrm{cch}(\{y\})$. Then $\mathrm{cch}(\{y\})$ is a compact convex set, so letting $G = \left\{z \in F : d(z, \mathrm{cch}(\{y\})) = \sup_{w\in F} d(w, \mathrm{cch}(\{y\}))\right\}$, G is a nonempty face of F disjoint from $\mathrm{cch}(\{y\})$, contradiction the minimality of F. So for all $y \in F$ we have $F \subseteq \mathrm{cch}(\{y\})$. Fix such a y. Note that $\mathrm{cch}(\{y\}) = \mathrm{ch}(\{y\})$. We claim that y is an extreme point of C. Assuming this, since C is a face of A we have that y is an extreme point of A and we have a contradiction to the hypothesis that $C \cap B = \emptyset$.

Suppose first that there do not exist $a, b \in C$ and $0 < t < 1$ such that $y = ta + (1-t)b$. Then y is an extreme point of C be definition. So let $a, b \in C$ and $0 < t < 1$ be such that $y = ta + (1-t)b$. We must show that $y \sim a \sim b$. Since F is a face of C, we have $a, b \in F$. Thus we can write $a = \sum_{i=1}^{n} s_i y$ and $b = \sum_{i=1}^{k} r_i y$ for $s_i, r_i \in [0, 1]$. By Proposition 2.4.2 and associativity we have $y \sim \left(\sum_{i=1}^{n} ts_i y + \sum_{i=1}^{k} (1-t) r_i y\right)$. Since $0 < t < 1$, iterating this argument we find that for any $\delta > 0$, there is $m \in \mathbb{N}$ and $(\lambda_i)_{i=1}^{m} \subseteq [0, 1]$ such that $\lambda_i \leq \delta$ for all i and $y \sim \sum_{i=1}^{m} \lambda_i y$.

We claim that this implies $y \sim \kappa y + (1-\kappa)y$ for all $\kappa \in [0, 1]$. Note that $\kappa y + (1-\kappa)y$ is isomorphic to $\iota_{\kappa,1-\kappa} \times y$, where $\iota_{\kappa,1-\kappa}$ is the trivial action of Γ on $(\{0, 1\}, m_\kappa)$ where $m_\kappa(\{0\}) = \kappa$ and $m_\kappa(\{1\}) = 1 - \kappa$. Hence y is a factor of $\kappa y + (1-\kappa)y$ and it thus suffices to show $\kappa y + (1-\kappa)y \prec y$.

Let X_1, X_2 be two copies of X, let $n, k \in \mathbb{N}$, $\epsilon > 0$ and a partition $\mathcal{P} = (P_i)_{i=1}^{k}$ of $X_1 \sqcup X_2$ be given. As before, we get a partition $\mathcal{P}_1 = (P_i^1)_{i=1}^{k}$ with $P_i^1 = P_i \cap X_1$ of X_1 and similarly a partition $\mathcal{P}_2 = (P_i^2)_{i=1}^{k}$ with $P_i^2 = P_i \cap X_2$ of X_2. Now, choose $\delta < \frac{\epsilon}{2}$. Then we can find m and $(\lambda_p)_{p=1}^{m}$ such that $y \sim \sum_{p=1}^{m} \lambda_p y$ and for some $l \leq m$ we have $\kappa - \frac{\epsilon}{2} \leq \sum_{p=1}^{l} \lambda_p \leq \kappa$. Let now X'_p be a copy of X for each $p \leq m$, and for $q \in \{0, 1\}$ let $P_{i,p}^{q}$ be the corresponding copy of P_i^q sitting in X'_p. Let $\mathcal{Q} = (Q_i)_{i=1}^{k}$ be the partition of $\bigsqcup_{p=1}^{m} X'_p$ given by $Q_i = \left(\bigsqcup_{p=1}^{l} P_{i,p}^1\right) \sqcup \left(\bigsqcup_{p=l+1}^{m} P_{i,p}^2\right)$. Then for $s \leq n$ and $i, j \leq k$ we have

$$\begin{aligned}
&\left|(\kappa\mu + (1-\kappa)\mu)(\gamma_s^{\kappa y+(1-\kappa)y} P_i \cap P_j) - \left(\sum_{p=1}^{m} \lambda_p \mu\right)\left(\gamma_s^{\sum_{p=1}^{m} \lambda_p y} Q_i \cap Q_j\right)\right| \\
&\leq \left|\kappa\mu(\gamma_s^y P_i^1 \cap P_j^1) - \left(\sum_{p=1}^{l} \lambda_p \mu(\gamma_s^y P_{i,p}^1 \cap P_{j,p}^1)\right)\right| \\
&\qquad + \left|(1-\kappa)\mu(\gamma_s^y P_i^2 \cap P_j^2) - \left(\sum_{p=l+1}^{m} \lambda_p \mu(\gamma_s^y P_{i,p}^2 \cap P_{j,p}^2)\right)\right| \\
&= \left|\kappa\mu(\gamma_s^y P_i^1 \cap P_j^1) - \left(\sum_{p=1}^{l} \lambda_p\right)\mu(\gamma_s^y P_i^1 \cap P_j^1)\right| \\
&\qquad + \left|(1-\kappa)\mu(\gamma_s^y P_i^2 \cap P_j^2) - \left(\sum_{p=l+1}^{m} \lambda\right)\mu(\gamma_s^y P_i^2 \cap P_j^2)\right| \\
&= \left|\left(\kappa - \sum_{p=1}^{l} \lambda_p\right)\mu(\gamma_s^y P_i^1 \cap P_j^1)\right| \\
&\qquad + \left|\left((1-\kappa) - \sum_{p=l+1}^{m} \lambda_p\right)\mu(\gamma_s^y P_i^2 \cap P_j^2)\right| \\
&\leq \left|\left(\kappa - \sum_{p=1}^{l} \lambda_p\right)\right| + \left|\left((1-\kappa) - \sum_{p=l+1}^{m} \lambda_p\right)\right| \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Since $y \sim \sum_{p=1}^{m} \lambda_p y$, $\kappa y + (1-\kappa)y \prec y$ and we are done. $\square$

We note that a metrizable topological vector space V is locally convex if and only if its topology is induced by a countable family of seminorms $\left(|\cdot|_n^V\right)_{n=1}^{\infty}$. Then $p(v, w) = \sum_{n=1}^{\infty} \frac{1}{2^n}|v - w|_n^V$ is a compatible metric on V, which is easily seen to obey Lemma 2.4.1. Thus the technique used to prove Theorem 2.1.2 works to prove the metrizable case of the classical Krein-Milman theorem using only the convex and metric structure of V, not the vector space structure in the form of linear functionals.

Before proving Theorem 2.1.3, we briefly discuss the ergodic decomposition in the context of weak equivalence classes. Suppose $a \in \mathrm{A}(\Gamma, X, \mu)$ and $a = \int_Z a_z d\eta(z)$ is the ergodic decomposition of a, that is to say we have a factor map $\pi : (X, \mu) \to (Z, \eta)$ such that if $\mu = \int_Z \mu_z d\eta(z)$ is the disintegration of μ over (Z, η) via π then $\mu_z(\pi^{-1}(z)) = 1$ and $\Gamma \curvearrowright^a (\pi^{-1}(z), \mu_z)$ is isomorphic to a_z. Furthermore, the assignment $z \mapsto \mu_z$ from $(Z, \eta) \to M_a(X)$ is Borel, where $M_a(X)$ is the space of a-invariant probablity measures on X (we may assume here that X is a Polish space). Recall that $A^*_\sim(\Gamma)$ is the space of weak equivalence classes of all measure-preserving actions of Γ, including those actions on finite space. $A^*_\sim(\Gamma)$ is topologized using the exact same metric as we use to topologize $\mathrm{A}_\sim(\Gamma, X, \mu)$. We would like to conclude that the assignment $z \mapsto [a_z]$ is measurable from (Z, η) to $A^*_\sim(\Gamma)$, where $[a_z]$ is the weak equivalence class of a_z. This is a consequence of the following lemma.

Lemma 2.6.1. *Let $\Gamma \curvearrowright^a Y$ be a Borel action of Γ on a Polish space Y. Then the map Θ from $M_a(Y)$ to $A^*_\sim(\Gamma)$ given by $\nu \mapsto [a_\nu]$ is Borel, where $[a_\nu]$ is the weak equivalence class of the measure preserving action $a_\nu = \Gamma \curvearrowright^a (Y, \nu)$.*

Proof. Fix a measure $\nu \in M_a(Y)$ and consider $\Theta^{-1}(U)$, where

$$U = \{[a] \in A^*_\sim(\Gamma) : d_H(C_{n,k}(a_\nu), C_{n,k}(a)) < \epsilon \text{ for all } n, k \leq N\}$$

for some $N \in \mathbb{N}$ and $\epsilon > 0$, so U is a basic open neighborhood of $\Theta(\nu) = a_\nu$. Since

$$U = \bigcup_{m=1}^{\infty} \bigcap_{n,k=1}^{N} \left\{[b] \in A^*_\sim(\Gamma) : d_H(C_{n,k}(a_\nu), C_{n,k}(b)) \leq \epsilon - \frac{1}{m}\right\},$$

it suffices to show $\Theta^{-1}(V)$ is Borel for a set V of the form

$$V = \{[b] \in A^*_\sim(\Gamma) : d_H(C_{n,k}(a_\nu), C_{n,k}(b)) \leq r\}.$$

Fixing n and k we write $C(b)$ for $C_{n,k}(b)$. Now, let K and L be compact subsets of a compact Polish space W with metric p, let D_K be dense in K and D_L be dense in L. We have

$$\begin{aligned} d_H(K, L) \leq r \iff & \max_{x \in K} \inf_{y \in L} p(x, y) \leq r \text{ and } \max_{y \in L} \inf_{x \in K} p(y, x) \leq r \\ \iff & (\forall x \in K)(\forall \delta > 0)(\exists y \in L)(p(x, y) < r + \delta) \\ & \wedge (\forall y \in L)(\forall \delta > 0)(\exists x \in K)(p(y, x) < r + \delta) \\ \iff & (\forall x \in D_K)(\forall \delta > 0)(\exists y \in D_L)(p(x, y) < r + \delta) \\ & \wedge (\forall y \in D_L)(\forall \delta > 0)(\exists y \in D_L)(p(y, x) < r + \delta). \end{aligned}$$

If $\mathbf{L}$ is a countable algebra generating the Borel σ-algebra $\mathbf{B}(Y)$ of Y, then $\mathbf{L}$ is dense in $\mathrm{MALG}(Y,\rho)$ for any Borel probability measure ρ on Y. Regarding a partition of Y into k pieces as an element of $\mathbf{B}(Y)^N$ and considering $\mathbf{L}^k$, we see that there exists a fixed countable family $(\mathcal{A}_m)_{m=1}^\infty$ of partitions of Y such that for any Borel probability measure ρ on Y, $(\mathcal{A}_m)_{m=1}^\infty$ is dense in the set of k-partitions of X with topology inherited from $\mathrm{MALG}(Y,\rho)$. We may further assume that each element of each $\mathcal{A}_m$ is clopen. This implies that the set $(M^{\mathcal{A}_m}(a_\rho))_{m=1}^\infty$ is dense in $C(a_\rho)$ for any Borel probability measure ρ. Therefore we have

$$V = \left(\bigcap_{m=1}^\infty\bigcap_{l=1}^\infty\bigcup_{i=1}^\infty\left\{b \in A^*_\sim(\Gamma) : \left|M^{\mathcal{A}_m}(a_\nu) - M^{\mathcal{A}_i}(b)\right| < r + \frac{1}{l}\right\}\right)$$
$$\cap\left(\bigcap_{m=1}^\infty\bigcap_{l=1}^\infty\bigcup_{i=1}^\infty\left\{b \in A^*_\sim(\Gamma) : \left|M^{\mathcal{A}_i}(a_\nu) - M^{\mathcal{A}_m}(b)\right| < r + \frac{1}{l}\right\}\right).$$

Now, $|M^{\mathcal{A}_i}(a_\nu) - M^{\mathcal{A}_m}(a_\rho)| < s$ if and only if $|\nu(\gamma^a A_i^j \cap A_i^t) - \rho(\gamma^a A_m^i \cap A_m^t)| < s$ for all $A_i^j, A_i^t \in \mathcal{A}_i$ and $A_m^i, A_m^t \in \mathcal{A}_m$. Since for any pair $J_1, J_2 \subseteq Y$ the set $\{\rho : |\nu(J_1) - \nu(J_2)| < s\}$ is Borel, we see

$$\Theta^{-1}\left(\left\{b \in A^*_\sim(\Gamma) : \left|M^{\mathcal{A}_i}(a_\nu) - M^{\mathcal{A}_m}(b)\right| < r + \frac{1}{l}\right\}\right)$$

is Borel and consequently $\Theta^{-1}(V)$ is Borel. □

We now prove Theorem 2.1.3

Proof. **(of Theorem 2.1.3)** Let $\Theta : Z \to A^*_\sim(\Gamma)$ be the map sending each point in z to the weak equivalence class $[a_z]$, so Θ is measurable by Lemma 2.6.1. Suppose towards a contradiction that the theorem fails. Then for every set $Z' \subseteq Z$ with $\eta(Z') = 1$, there is more than one weak equivalence class in the set $\{[a_z] : z \in Z'\}$. Equivalently, the measure $\Theta_*\eta$ on $A^*_\sim(\Gamma)$ is not supported on a single point. We can thus split $A^*_\sim(\Gamma)$ into two disjoint sets Y_1, Y_2 such that $0 < \Theta_*\eta(Y_1), \Theta_*\eta(Y_2) < 1$. Letting $A_i = \Theta^{-1}(Y_i)$, we get disjoint measurable sets $A_1, A_2 \subseteq Z$ such that $0 < \eta(A_1), \eta(A_2) < 1$ and for all $z \in A_1$ and all $w \in A_2$ we have that $z \nsim w$.

Recall that for a measure-preserving action b of Γ and $n, k \in \mathbb{N}$ the set $C_{n,k}(a) \subseteq [0,1]^{n\times k\times k}$ was defined in Section 2.3.

Lemma 2.6.2. *For any action b of Γ on a probability space (Y,ν), we have* $\mathrm{cch}(C_{n,k}(b)) \subseteq C_{n,k}(\iota \times b)$.

Proof. Write $C_{n,k}(b) = C(b)$. Suppose $x \in \mathrm{cch}(C(b))$. Then we can find points $(x_i)_{i=1}^{\infty}$ such that $\lim_{i\to\infty} x_i = x$ and each x_i has the form $x_i = \sum_{j=1}^{j(i)} \alpha_i^j x_i^j$ for $(x_i^j)_{j=1}^{j(i)} \subseteq C(b)$ and $(\alpha_i^j)_{j=1}^{j(i)} \subseteq [0,1]$ with $\sum_{j=1}^{j(i)} \alpha_j^i = 1$ for each i. Without loss of generality we may assume that each x_i^j has the form $M^{\mathcal{A}_i^j}(b)$ for a partition $\mathcal{A}_i^j = (A_{i,l}^j)_{l=1}^k$ of Y into k pieces. Fixing i consider the action $\sum_{j=1}^{j(i)} \alpha_i^j b$ on the space $\left(\bigsqcup_{j=1}^{j(i)} Y_j, \sum_{j=1}^{j(i)} \alpha_i^j \nu_j\right)$, where each (Y_j, ν_j) is a copy of (Y, ν). Let $\mathcal{B} = (B_l)_{l=1}^k$ be the partition of $\bigsqcup_{j=1}^{j(i)} Y_j$ given by letting $B_l = \bigsqcup_{j=1}^{j(i)} A_{i,l}^j$, where $A_{i,l}^j$ sits inside the j copy of Y. For any $p \le n$ and $l, m \le k$ and $x \in [0,1]^{n\times k\times k}$ let $(x)_{p,l,m}$ be the p, l, m coordinate of x. We then have

$$\begin{aligned}
\left(M^{\mathcal{B}}\left(\sum_{j=1}^{j(i)} \alpha_i^j b\right)\right)_{p,l,m} &= \left(\sum_{j=1}^{j(i)} \alpha_i^j \nu_j\right)\left(\gamma_p^{\sum_{j=1}^{j(i)} \alpha_i^j b} B_l \cap B_m\right) \\
&= \sum_{j=1}^{j(i)} \left(\alpha_i^j \nu_j(\gamma_p^b A_{i,l}^j \cap A_{i,m}^j)\right) \\
&= \sum_{j=1}^{j(i)} \alpha_i^j \left(M^{\mathcal{A}_i^j}(b)\right)_{p,l,m}.
\end{aligned}$$

Therefore

$$M^{\mathcal{B}}\left(\sum_{j=1}^{j(i)} \alpha_i^j b\right) = \sum_{j=1}^{j(i)} \alpha_i^j \left(M^{\mathcal{A}_i^j}(b)\right) = x_i.$$

We have shown that $x_i \in C\left(\sum_{j=1}^{j(i)} \alpha_i^j b\right)$. Since $\sum_{j=1}^{j(i)} \alpha_i^j b$ is a factor of $b \times \iota$, we have $x_i \in C(b \times \iota)$. Since $\lim_{i\to\infty} x_i = x$ and $C(b\times\iota)$ is closed, the lemma follows. □

It is clear that for any two measure-preserving actions b, c we have $b \prec c$ if and only if $C_{n,k}(b) \subseteq C_{n,k}(c)$ for all n, k. We claim that there are disjoint subsets $A_3, A_4 \subseteq Z$ of positive measure such that for some pair n_0, k_0, every $z \in A_3$ and every $w \in A_4$ we have $C_{n_0,k_0}(a_z) \nsubseteq \mathrm{cch}(C_{n_0,k_0}(a_w))$. For $z \in A_3$ let $R_z = \{w \in A_2 : a_z \nprec a_w\}$. Since a_z is ergodic, $a_z \prec a_w \times \iota$ implies $a_z \prec a_w$. Therefore $R_z = \{w \in A_2 : a_z \nprec a_w \times \iota\}$.

Assume first that there is a set $D_3 \subseteq A_1$ with $\eta(D_3) > 0$ such that for each $z \in D_3$ we have $\eta(R_z) > 0$. Write $\hat{K}$ for $\mathrm{cch}(K)$. By Lemma 2.6.2 we can write $R_z = \bigcup_{n,k=1}^{\infty} R_z^{n,k}$ where $R_z^{n,k} = \left\{w \in A_2 : C_{n,k}(a_z) \nsubseteq \widehat{C_{n,k}(a_w)}\right\}$. Thus for each z there is a lexicographically least pair (n_z, k_z) such that $\eta(R_z^{n_z,k_z}) > 0$. Therefore there is a pair n_0, k_0 and a set $D_4 \subseteq D_3$ such that $\eta(D_4) > 0$ and for all $z \in D_4$ we have $\eta(R_z^{n_0,k_0}) > 0$. Fixing n_0 and k_0 we write $C(b)$ for $C_{n_0,k_0}(b)$. Let $(w_j)_{j=1}^{\infty} \subseteq A_2$ be a sequence of points such that the family $\left(\widehat{C(a_{w_j})}\right)_{j=1}^{\infty}$ is dense in the space $\left\{\widehat{C(a_w)} : w \in A_2\right\}$ with respect to the Hausdorff metric d_H on the space on compact subsets of $[0,1]^{n_0 \times k_0 \times k_0}$. Let then $F_{j,l} = \left\{w \in A_2 : d_H\left(\widehat{C(a_w)}, \widehat{C(a_{w_j})}\right) < \frac{1}{l}\right\}$.

Fix $z \in D_4$ and choose $w \in R_z^{n_0,k_0}$. By hypothesis there is $\epsilon > 0$ such that $C(a_z) \nsubseteq B_\epsilon\left(\widehat{C(a_w)}\right)$, where $B_\epsilon(K)$ denotes the ball of radius ϵ around K. Then if we choose j so that $d_H\left(\widehat{C(a_{w_j})}, \widehat{C(w)}\right) < \frac{\epsilon}{2}$ and l so that $\frac{1}{l} < \frac{\epsilon}{2}$ we have $w \in F_{j,l} \subseteq R_z^{n_0,k_0}$. Hence there is a subset $\mathcal{J} \subseteq \mathbb{N}^2$ such that $R_z^{n_0,k_0} = \bigcup_{(j,l)\in\mathcal{J}} F_{j,l}$. So for each z we can choose a lexicographically least pair (j_z, l_z) such that $\eta(F_{j_z,l_z}) > 0$ and $F_{j_z,l_z} \subseteq R_z^{n_0,k_0}$. There is then a pair (j_0, l_0) and a set $E_3 \subseteq D_3$ with $\eta(E_3) > 0$ such that $\eta(F_{j_0,l_0}) > 0$ and for all $z \in E_3$ and all $w \in F_{j_0,l_0}$ we have $C(a_z) \nsubseteq \widehat{C(a_w)}$. So take $A_3 = E_3$ and $A_4 = F_{j_0,l_0}$. Thus we are left with the case $\eta(R_z) = 0$ for almost all $z \in A_1$. Then for almost all $w \in A_2$ and almost all $z \in A_1$ we must have $a_w \not\prec a_z$, so a symmetric argument gives the claim.

Given a (real) topological vector space V, we say a hyperplane in V is a set of the form $H_{\ell,\alpha} = \{v \in V : \ell(v) = \alpha\}$ for some continuous linear functional ℓ and $\alpha \in \mathbb{R}$. Given disjoint compact subsets $W_1, W_2 \subseteq V$ we say that $H_{\ell,\alpha}$ separates W_1 from W_2 if $W_1 \subseteq \{v \in V : \ell(v) < \alpha\}$ and $W_2 \subseteq \{v \in V : \ell(v) > \alpha\}$.

Lemma 2.6.3. *Let $S \subseteq \mathbb{R}^n$ be compact. Then there is a countable family $(H_i)_{i=1}^{\infty}$ of hyperplanes such that for any $x \in S$ and any compact convex $W \subseteq S$ there is i so H_i separates $\{x\}$ from W.*

Proof. Let $(\ell_j)_{j=1}^{\infty}$ be a countable set of linear functionals which is dense in the sup norm on S. Enumerate $\mathbb{Q}$ as $(q_m)_{m=1}^{\infty}$ and let $H_{j,m} = \{s \in S : \ell_j(s) = q_m\}$. Given x and W, by Hahn-Banach find a linear functional ℓ and $\alpha \in \mathbb{R}$ so that $H = H_{\ell,\alpha}$ separates x from W. Let $r = \min\left(\inf_{h\in H} ||x - h||, \inf_{\substack{h\in H,\\ w\in W}} ||h - w||\right)$ so $r > 0$. Then choose m so $|q_m - \alpha| < \frac{r}{2}$ and j so $\sup_{s\in S} |\ell(s) - \ell_j(s)| < \frac{r}{2}$. Then $H_{j,m}$ separates x from W. □

Now take $S = [0,1]^{n_0 \times k_0 \times k_0}$ and fix a family $(H_i)_{i=1}^{\infty}$ of hyperplanes as in the lemma. Since $\widehat{C(a_w)}$ is compact convex for each $w \in A_4$ and for all $z \in A_3$ we have $C(a_z) \nsubseteq \widehat{C(a_w)}$, for each pair $(z, w) \in A_3 \times A_4$ there is an index $i(z, w)$ and a point

$x_{z,w} \in C(a_z)$ such that $H_{i(z,w)}$ separates $x_{z,w}$ from $\widehat{C(a_w)}$. Fix $z \in A_3$. Taking $(w_j)_{j=1}^{\infty}$ as before, for $(j,l) \in \mathbb{N}^2$ let $G_{j,l} = \left\{w \in A_4 : d_H\left(\widehat{C(a_w)}, \widehat{C(a_{w_j})}\right) < \frac{1}{l}\right\}$. Choosing $w \in A_4$, let $\epsilon = d_H\left(\widehat{C(w)}, H_{i(z,w)}\right)$ so $\epsilon > 0$. Finding $j_{z,w}$ so $d_H\left(\widehat{C(w_j)}, \widehat{C(w)}\right) < \frac{\epsilon}{2}$ and $l_{z,w}$ so $\frac{1}{l} < \frac{\epsilon}{2}$ we have $w \in G_{j_{z,w},l_{z,w}}$ and $H_{i(z,w)}$ separates $x_{z,w}$ from $\widehat{C(u)}$ for all $u \in G_{j_{z,w},l_{z,w}}$. Then we have $A_4 = \bigcup_{\substack{(j_{z,w},l_{z,w}):\\ w \in A_4}} G_{j_{z,w},l_{z,w}}$ so we can find w_0 so that $\eta\left(G_{j_{z,w_0},l_{z,w_0}}\right) > 0$. Let then $G_z = G_{j_{z,w_0},l_{z,w_0}}$, $x_z = x_{z,w_0}$ and $i(z) = i(z,w_0)$ so that $H_{i(z)}$ separates x_z from $\widehat{C(u)}$ for all $u \in G_z$. Since the G_z were chosen from a countable family, we can find a set $A_5 \subseteq A_3$ of positive measure such that $G_z = G$ is the same for all $z \in A_5$. We can then find an index i and a set $A_6 \subseteq A_5$ of positive measure such that for all $z \in K$, $H_i = H$ separates x_z from $\widehat{C(u)}$ for all $u \in G$. H splits $[0,1]^{n\times k\times k}$ into two closed convex sets H_+ and H_-, where H_+ contains the x_z and H_- contains the $C(u)$.

For $S \subseteq Z$ with $\eta(S) > 0$ let $\eta_S = \frac{\eta\restriction S}{\eta(S)}$ be normalized measure on S. By Lemma 2.4.2 we have $C\left(\int_G a_u d\eta_G(u)\right) \subseteq \operatorname{cch}\left(\bigcup_{u\in G} C(u)\right) \subseteq H_-$. Write $A_6 = \bigcup_{p=1}^{\infty} A_6^p$, where $A_6^p = \left\{z \in A_6 : d_H(x_z, H) \geq \frac{1}{p}\right\}$ and find p so $\eta(A_6^p) > 0$. Letting $K = A_6^p$, for all $z \in K$, x_z is an element of the closed convex set $H_+^p = \{y \in H_+ : d_H(y,H) \geq \frac{1}{p}\}$ and H_+^p is disjoint from H_-. We have $\int_K x_z d\eta_K(z) \in C\left(\int_K a_z d\eta_K(z)\right)$ and $\int_K x_z d\eta_K(z) \in H_+^p$. Since $C\left(\int_G a_u d\eta_G(u)\right) \subseteq H_-$ we see that $C\left(\int_K a_z d\eta_K(z)\right) \nsubseteq C\left(\int_G a_u d\eta_G(u)\right)$ and it follows that $\int_K a_z d\eta_K(z) \nsim \int_G a_u d\eta_G(u)$. Let $L_1 = K, L_2 = G$ then there is $i \in \{1,2\}$ with $\int_{L_i} a_z d\eta_{L_i}(z) \nsim a$. Since $0 < \eta(L_i) < 1$, we can write

$$a = \eta(L_i)\left(\int_{L_i} a_z d\eta_{L_i}(z)\right) + \eta(Z \setminus L_i)\left(\int_{Z\setminus L_i} a_z d\eta_{Z\setminus L_i}(z)\right),$$

which contradicts our assumption that a is an extreme point. □

We now prove Theorem 2.1.4. Recall that the uniform topology on $\operatorname{Aut}(X,\mu)$ is given by the metric $d_u(T,S) = \mu(\{x : Tx \neq Sx\})$. If $\mathcal{P} = \{P_1, \ldots, P_p\}$ is a partition of a space on which $\mathbb{F}_N$ acts by an action a, $J \subseteq \mathbb{F}_N$ is finite and $\tau : J \to p$ let $P_\tau^a = \bigcap_{\gamma\in J} \gamma^a P_{\tau(\gamma)}$.

Proof. **(of Theorem 2.1.4)** Let a be a free action of $\mathbb{F}_N$. By replacing a with $a \times \iota$ if necessary, we may assume that for each n, k the set $C_{n,k}(a)$ is closed and convex. Fix integers n_0 and k_0 and $\epsilon > 0$. It is enough to find a free ergodic action b of

$\mathbb{F}_N$ such that for all $n \leq n_0$ and $k \leq k_0$ we have $d_H(C_{n,k}(a), C_{n,k}(b)) < \epsilon$. Let $\{\gamma_1, \ldots, \gamma_{n_0}\} = F_0$ be the finite subset of $\mathbb{F}_N$ under consideration. Let $s = s_{\mathbb{F}_N}$ be the Bernoulli shift of $\mathbb{F}_N$ acting on $\left(2^{\mathbb{F}_N}, \nu\right)$ where ν is the product measure. For any action c of $\mathbb{F}_N$ on (X, μ) and $\gamma \in \mathbb{F}_N$ we have

$$\{(x, y) \in X \times 2^{\mathbb{F}_N} : \gamma^{c\times s}(x, y) \neq \gamma^{a\times x}(x, y)\} = \{x \in X : \gamma^c x \neq \gamma^a x\} \times Y$$

and hence

$$(\mu \times \nu)(\{(x, y) \in X \times 2^{\mathbb{F}_N} : \gamma^{c\times s}(x, y) \neq \gamma^{a\times x}(x, y)\}) = \mu(\{x \in X : \gamma^c x \neq \gamma^a x\}).$$

Assume $d_u(\gamma^a, \gamma^c) < \frac{\epsilon}{16}$ for all $\gamma \in F_0$. Then for any measurable partition $\mathcal{A} = A_1, \ldots, A_k$ of $X \times 2^{\mathbb{F}_N}$, all $\gamma \in F_0$ and all $i, j \leq k$ we have

$$|(\mu \times \nu)(\gamma^{a\times s} A_i \cap A_j) - (\mu \times \nu)(\gamma^{c\times s} A_i \cap A_j)| < \frac{\epsilon}{16}$$

for all $\gamma \in F_0$. In the notation of Section 2.3, $\rho\left(M^{\mathcal{A}}_{n,k}(a \times s), M^{\mathcal{A}}_{n,k}(c \times s)\right) < \frac{\epsilon}{16}$ where ρ is the supremum metric on $[0, 1]^{n\times k\times k}$. Choose a finite collection $\mathcal{L}$ of measurable subsets of $X \times 2^{\mathbb{F}_N}$ such that for every measurable partition $\mathcal{A}$ of $X \times 2^{\mathbb{F}_N}$ there is a partition $\mathcal{B} \subseteq \mathcal{L}$ such that $\rho\left(M^{\mathcal{A}}_{n,k}(a \times s), M^{\mathcal{B}}_{n,k}(a \times s)\right) < \frac{\epsilon}{16}$. Then for every such $\mathcal{A}$ there exists $\mathcal{B} \subseteq \mathcal{L}$ such that $\rho\left(M^{\mathcal{A}}_{n,k}(c \times s), M^{\mathcal{B}}_{n,k}(c \times s)\right) < \frac{3\epsilon}{16}$.

For $\gamma \in \mathbb{F}_N$ let $\pi_\gamma : 2^{\mathbb{F}_N} \to 2$ be projection onto the γ coordinate. For $i \in \{0, 1\}$ let $S_i = \pi_e^{-1}(\{i\})$ and put $\mathcal{S} = \{S_1, S_2\}$. Choose now a finite partition $\mathcal{R} = \{R_1, \ldots, R_r\}$ of X and a finite subset $F \subseteq \mathbb{F}_N$ containing F_0 such that for every $A \in \mathcal{L}$ there are sets R_j with $1 \leq j \leq r$ and a family of functions $(\tau_j)_{j=1}^t$ with $\tau_j : F \to 2$ such that

$$\mu\left(\left(\bigsqcup_{j=1}^{t} R_j \times S^s_{\tau_j}\right) \triangle A\right) < \frac{\epsilon}{16}.$$

Write $\mathcal{P} = \mathcal{R} \times \mathcal{S}$. We can identify a function $\theta : F \to r \times 2$ with a pair (σ, τ) where $\sigma : F \to r$ and $\tau : F \to 2$ so

$$P^{c\times s}_\theta = \bigcap_{\gamma\in F} \gamma^b P^{c\times s}_{\theta(\gamma)} = \left(\bigcap_{\gamma\in F} \gamma^c R_{\sigma(\gamma)}\right) \times \left(\bigcap_{\gamma\in F} \gamma^s S_{\tau(\gamma)}\right) = R^c_\sigma \times S^s_\tau.$$

Note that for any $j \leq r$, $R_j \times S^s_\tau$ is a finite disjoint union of sets of the form $R^c_\sigma \times S^s_\tau$, hence any $A \in \mathcal{L}$ is within $\frac{\epsilon}{16}$ of finite disjoint union of sets of the form $P^{c\times s}_\theta$ for $\theta : F \to r \times 2$.

Let $\delta = \frac{\epsilon}{4(2r)^{2|F|}}$. Fix an ergodic action c of $\mathbb{F}_N$ such that $d_u(\gamma^a, \gamma^c) < \frac{\delta^2}{32|F|^2(2r)^{|F|^2}}$ for all $\gamma \in F$. (For example use the fact that the ergodic automorphisms are uniformly dense in $\mathrm{Aut}(X, \mu)$ to move one of the generators γ of $\mathbb{F}_N$ so it acts ergodically but is still sufficiently close to γ^a). Then clearly $d_H(C_{n,k}(a), C_{n,k}(c)) < \frac{\epsilon}{2}$ for all $n \leq n_0$ and $k \leq k_0$. Let $b = c \times s$. Since c is ergodic and s is free and mixing, b is free and ergodic. Thus it is sufficient to show $d_H(C_{n,k}(c), C_{n,k}(b)) < \frac{\epsilon}{2}$ for all $n \leq n_0$, $k \leq k_0$. Since $c \prec b$, it is sufficient to show that for every partition $\mathcal{A}$ of $X \times 2^{\mathbb{F}_N}$ there is a partition C of X such that $\rho\left(M^{\mathcal{A}}_{n,k}(b), M^{C}_{n,k}(c)\right) < \frac{\epsilon}{2}$. By our previous reasoning, for each partition $\mathcal{A} = (A_1, \ldots, A_k)$ of $X \times 2^{\mathbb{F}_N}$ there is a partition $\mathcal{B}$ whose pieces are disjoint unions of sets of the form P^b_θ for $\theta : F \to r \times 2$ such that $\rho\left(M^{\mathcal{A}}_{n,k}(b), M^{\mathcal{B}}_{n,k}(b)\right) < \frac{\epsilon}{4}$.

Claim 2.6.1. *There is a partition Q of X indexed by $r \times 2$ such that for every $\theta : J \to r \times 2$ with $J \subseteq F_0 F$ we have $|(\mu \times \nu)(P^b_\theta) - \mu(Q^c_\theta)| < \delta$.*

Suppose the claim holds. Regard $\mathbb{F}_N$ as acting on $\bigcup_{J \subseteq \mathbb{F}_N} \{\theta : J \to 2 \times r\}$ by shift, $\gamma \cdot \theta(\gamma') = \theta(\gamma^{-1}\gamma')$. Thus the domain $\mathrm{dom}(\gamma \cdot \theta) = \gamma \mathrm{dom}(\theta)$. Then for any $\theta, \kappa : F \to 2 \times r$ and $\gamma \in F_0$ we have

$$\gamma^b P^b_\theta \cap P^b_\kappa = \begin{cases} P^b_{\gamma \cdot \theta \cup \kappa} & \text{if } \gamma \cdot \theta \text{ and } \kappa \text{ are compatible,} \\ \emptyset & \text{if not.} \end{cases}$$

and similarly

$$\gamma^c Q^c_\theta \cap Q^c_\kappa = \begin{cases} Q^c_{\gamma \cdot \theta \cup \kappa} & \text{if } \gamma \cdot \theta \text{ and } \kappa \text{ are compatible,} \\ \emptyset & \text{if not.} \end{cases}$$

Therefore the claim gives $|(\mu \times \nu)(\gamma^b P^b_\theta \cap P^b_\kappa) - \mu(\gamma^c Q^c_\theta \cap Q^c_\kappa)| < \delta$ for all $\theta, \kappa : F \to r \times 2$. So if $\mathcal{B} = \{B_1, \ldots, B_k\}$ is a partition such that $B_i = \bigsqcup_{s=1}^t P^b_{\theta_i(s)}$ for functions $\theta_i(s) : F \to r \times 2$ and we let $C_i = \bigsqcup_{s=1}^t Q^c_{\theta_i(s)}$, then we have

$$\begin{aligned} |(\mu \times \nu)(\gamma^b B_i \cap B_j) - \mu(\gamma^c C_i \cap C_j)| &= \left|(\mu \times \nu)\left(\bigsqcup_{s,s'=1}^{t} \gamma^b P^b_{\theta_i(s)} \cap P^b_{\theta_j(s')}\right)\right. \\ &\qquad \left. - \mu\left(\bigsqcup_{s,s'=1}^{t} \gamma^c Q^c_{\theta_i(s)} \cap Q^c_{\theta_j(s')}\right)\right| \\ &\leq t^2 \delta \leq (2r)^{2|F|}\delta < \frac{\epsilon}{4}, \end{aligned}$$

since $t \le (2r)^{|F|}$. Taking $C = (C_i)_{i=1}^k$ we get $\rho\left(M_{n,k}^{\mathcal{B}}(b), M_{n,k}^{C}(c)\right) < \frac{\epsilon}{4}$, which implies the theorem.

It remains to show Claim 2.6.1. This part of the argument follows the proof of Theorem 1 in [2] and the extensions of these ideas developed in [74]. Let $G = F_0F$. Assume without loss of generality that G is closed under taking inverses. Note that it suffices to prove the claim for θ defined on all of G. In order to find Q we will find a partition $\mathcal{T} = \{T_1, T_2\}$ and set $Q_{i,j} = R_i \cap T_j$ for $1 \le i \le r$, $1 \le j \le 2$. Thus we are looking for $\mathcal{T} = \{T_1, T_2\}$ such that for all (τ, σ) with $\sigma : G \to r$ and $\tau : G \to 2$ we have

$$|(\mu \times \nu)(R_\sigma^c \times S_\tau^s) - \mu(R_\sigma^c \cap T_\tau^c)| < \delta.$$

Note that $\nu(S_\tau^s) = 2^{-|G|}$ for any such τ so we are looking for $\mathcal{T}$ such that

$$\left|2^{-|G|}\mu(R_\sigma^c) - \mu(R_\sigma^c \cap T_\tau^c)\right| < \delta.$$

The idea is that a random $\mathcal{T}$ should have this property.

Without loss of generality we may assume X is a compact metric space with a compatible metric p. For $\eta > 0$ let

$$D_\eta = \{x \in X : \text{ for all } \gamma, \gamma' \in G, \gamma_1 \neq \gamma_2 \text{ implies } p(\gamma_1^c x, \gamma_2^c x) > \eta\}$$

and

$$E_\eta = \{(x, x') \in D_\eta^2 : \text{ for all } \gamma_1, \gamma_2 \in G, p(\gamma_1^c x, \gamma_2^c x') > \eta\}.$$

Lemma 2.6.4. *There is $\eta > 0$ such that $\mu(D_\eta) > 1 - \frac{\delta^2}{16(2r)^{|F|^2}}$ and $\mu^2(X^2 \setminus E_\eta) < \frac{\delta^2}{16(2r)^{2|F|}}$.*

Proof. Clearly if $\eta_1 < \eta_2$ then $D_{\eta_2} \subseteq D_{\eta_1}$. We have $X \setminus \bigcup_{\eta>0} D_\eta = \{x \in X :$ for some $\gamma_1 \neq \gamma_2 \in G$, $\gamma_1^c x = \gamma_2^c x\}$. Now since a is free, if $\gamma_1^c x = \gamma_2^c x$ then we must have $\gamma_i^c x \neq \gamma_i^a x$ for some $i \in \{1, 2\}$. Each $\gamma \in G$ is a product $f_1 f_2$ for $f_1 \in F_0$ and $f_2 \in F$, thus for any $\gamma \in G$ we have

$$d_u(\gamma^c, \gamma^a) < d_u(f_1^a, f_1^c) + d_u(f_2^a, f_2^c) < \frac{\delta^2}{16|F|^2(2r)^{|F|^2}}$$

since $f_i \in F$. Therefore

$$\mu(\{x : \text{ for some } \gamma \in G, \gamma^c x \neq \gamma^a x\}) < |G| \frac{\delta^2}{16|F|^2(2r)^{|F|^2}} < \frac{\delta^2}{16(2r)^{2|F|}}$$

and hence $\mu\left(X \setminus \bigcup_{\eta>0} D_\eta\right) < \frac{\delta^2}{16(2r)^{|F|^2}}$. So we can find $\eta = \eta_0$ such that D_{η_0} satisfies the lemma. Now for any $\eta > 0$,

$$\begin{aligned}
D_{\eta_0}^2 \setminus \bigcup_{\eta>0} E_\eta = \{(x,x') \in D_{\eta_0}^2 :\ &\text{for all } \eta > 0 \\
&\text{there exist } \gamma_1, \gamma_2 \in G \text{ such that } p(\gamma_1 x, \gamma_2 x') < \eta\} \\
= \{(x,x') \in D_{\eta_0}^2 :\ &\text{there exist } \gamma_1, \gamma_2 \in G \text{ such that } \gamma_1 x = \gamma_2 x'\}.
\end{aligned}$$

For a fixed x, $\{(x,x') \in D_{\eta_0}^2 :$ there exist $\gamma_1, \gamma_2 \in G$ such that $\gamma_1 x = \gamma_2 x'\}$ is finite so $\mu\left(D_{\eta_0}^2 \setminus \bigcup_{\eta>0} E_\eta\right)$ has measure 0 by Fubini and hence we have the lemma for E_η. $\square$

Let $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$ be a partition of X into pieces with diameter $< \frac{\eta}{4}$. For $x \in X$ let $Y(x)$ be the unique $l \leq m$ such that $x \in Y_i$. Let κ be the uniform (= product) probability measure on 2^m and for each $\omega \in 2^m$ define a partition $Z(\omega) = \{Z_1^\omega, Z_2^\omega\}$ by letting $x \in Z_i^\omega$ if and only if $\omega(Y(x)) = i$. Thus we have a random variable $Z : (2^m, \kappa) \to \mathrm{MALG}(X,\mu)^2$ given by $\omega \mapsto Z(\omega)$. Fix now $\tau : G \to 2$ and an arbitrary subset $A \subseteq X$. We compute the expected value of $\mu(Z(\omega)_\tau \cap A)$. Let χ_B be the characteristic function of B.

$$\begin{aligned}
\mathbb{E}[\mu(Z_\tau \cap A)] &= \int_{2^m} \mu(Z(\omega)_\tau \cap A) d\kappa(\omega) \\
&= \int_{2^m} \int_X \chi_{Z(\omega)_\tau \cap A}(x) d\mu(x) d\kappa^m(\omega) \\
&= \int_A \int_{2^m} \chi_{Z(\omega)_\tau}(x) d\kappa(\omega) d\mu(x) \\
&= \int_{D_\eta \cap A} \int_{2^m} \chi_{Z(\omega)_\tau}(x) d\kappa(\omega) d\mu(x) + \int_{A \setminus D_\eta} \int_{2^m} \chi_{Z(\omega)_\tau}(x) d\kappa(\omega) d\mu(x).
\end{aligned} \tag{2.6}$$

Now if $x \in D_\eta$ then for all $\gamma_1 \neq \gamma_2 \in G$ we have $p(\gamma_1^c x, \gamma_2^c x) \geq \eta$ so that $Y(\gamma_1^c x) \neq Y(\gamma_2^c x)$ and hence the events $\omega(Y(\gamma_1^c x)) = i$ and $\omega(Y(\gamma_2^c x)) = j$ are independent. We have $x \in \gamma^c Z(\omega)_{\tau(\gamma)}$ if and only if $\omega(Y((\gamma^{-1})^c x)) = \tau(\gamma)$, so if $x \in D_\eta$ and $\gamma_1 \neq \gamma_2 \in G$ the events $x \in \gamma^c Z(\omega)_{\tau(\gamma_1)}$ and $x \in \gamma^c Z(\omega)_{\tau(\gamma_2)}$ are independent. So for $x \in D_\eta$,

$$\int_{2^m} \chi_{Z(\omega)_\tau}(x)d\kappa(\omega) = \kappa(\{\omega : x \in \gamma^c Z(\omega)_{\tau(\gamma)} \text{ for all } \gamma \in G\})$$
$$= \prod_{\gamma\in G} \kappa\left(\left\{\omega : \omega(Y((\gamma^{-1})^c x)) = \tau(\gamma)\right\}\right) = 2^{-|G|}. \tag{2.7}$$

Since $\mu(X \setminus D_\eta) < \frac{\delta^2}{16(2r)^{|F|^2}}$, we have $2^{-|G|}\left(\mu(A) - \frac{\delta^2}{16(2r)^{|F|^2}}\right) \leq (6) \leq 2^{-|G|}\mu(A) + \frac{\delta^2}{16(2r)^{|F|^2}}$ and thus $\left|\mathbb{E}[\mu(Z_\tau \cap A)] - \mu(A)2^{-|G|}\right| < \frac{\delta^2}{16(2r)^{|F|^2}}$. We now compute the second moment of $\mu(Z_\tau \cap A)$, in order to estimate its variance.

$$\begin{aligned}
\mathbb{E}\left[\mu(Z_\tau \cap A)^2\right] &= \int_{2^m} \mu(Z_\tau(\omega) \cap A)^2 d\kappa(\omega) \\
&= \int_{2^m} \left(\int_A \chi_{Z_\tau(\omega)}(x)d\mu(x)\right)^2 d\kappa(\omega) \\
&= \int_{2^m}\int_{A^2} \chi_{Z_\tau(\omega)}(x_1)\chi_{Z_\tau(\omega)}(x_2)d\mu^2(x_1,x_2)d\kappa(\omega) \\
&= \int_{A^2}\int_{2^m} \chi_{Z_\tau(\omega)}(x_1)\chi_{Z_\tau(\omega)}(x_2)d\kappa(\omega)d\mu^2(x_1,x_2) \\
&= \int_{A^2\cap E_\eta}\int_{2^m} \chi_{Z_\tau(\omega)}(x_1)\chi_{Z_\tau(\omega)}(x_2)d\kappa(\omega)d\mu^2(x_1,x_2) \\
&\qquad + \int_{A^2\setminus E_\eta}\int_{2^m} \chi_{Z_\tau(\omega)}(x_1)\chi_{Z_\tau(\omega)}(x_2)d\kappa(\omega)d\mu^2(x_1,x_2).
\end{aligned} \tag{2.8}$$

Now if $(x_1, x_2) \in E_\eta$ then for any pair $\gamma_1, \gamma_2 \in G$ we have $p(\gamma_1^c x_1, \gamma_2^c x_2) > \eta$ so that $Y(\gamma_1^c x_1) \neq Y(\gamma_2^c x_2)$ and thus for a fixed pair (x_1, x_2) the events $\omega(Y(\gamma^{-1})^c x_1) = \tau(\gamma)$ for all $\gamma \in G$ and $\omega(Y(\gamma^{-1})^c x_2) = \tau(\gamma)$ for all $\gamma \in G$ are independent. Hence for a fixed $(x_1, x_2) \in E_\eta$ we have

$$\begin{aligned}
&\int_{2^m} \chi_{Z_\tau(\omega)}(x_1)\chi_{Z_\tau(\omega)}(x_2)d\kappa(\omega) \\
&\quad = \kappa(\{\omega : x_1 \in \gamma^c Z(\omega)_{\tau(\gamma)} \text{ and } x_2 \in \gamma^c Z(\omega)_{\tau(\gamma)} \text{ for all } \gamma \in G\}) \\
&\quad = \kappa(\{\omega : \omega(Y((\gamma^{-1})^c x_1) = \tau(\gamma) \text{ and } \omega(Y((\gamma^{-1})^c)x_2) = \tau(\gamma) \text{ for all } \gamma \in G\}) \\
&\quad = \kappa\left(\left\{\omega : \omega(Y((\gamma^{-1})^c x_1)) = \tau(\gamma) \text{ for all } \gamma \in G\right\}\right) \\
&\qquad\qquad \cdot \kappa\left(\left\{\omega : \omega(Y((\gamma^{-1})^c x_2)) = \tau(\gamma) \text{ for all } \gamma \in G\right\}\right) \\
&\quad = 2^{-2|G|}
\end{aligned}$$

by (7) and the fact that $E_\eta \subseteq D^2_\eta$. Since $\mu^2(A \setminus E_\eta) < \frac{\delta^2}{16(2r)^{|F|^2}}$ we see

$$\left(\mu(A)^2 - \frac{\delta^2}{16(2r)^{|F|^2}}\right) 2^{-2|G|} \leq (8) \leq 2^{-2|G|}\mu(A)^2 + \frac{\delta^2}{16(2r)^{|F|^2}}$$

and hence $\left|\mathbb{E}[\mu(Z_\tau \cap A)^2] - \mu(A)^2 2^{-2|G|}\right| < \frac{\delta^2}{16(2r)^{|F|^2}}$. Therefore

$$\begin{aligned}
\mathrm{Var}(\mu(Z_\tau \cap A)) &= \mathbb{E}[\mu(Z_\tau \cap A)^2] - \mathbb{E}[\mu(Z_\tau \cap A)]^2 \\
&\leq \left|\mathbb{E}[\mu(Z_\tau \cap A)^2] - \mu(A)^2 2^{-2|G|}\right| + \mu(A)^2 2^{-2|G|} \\
&\qquad - \left(-\left|\mathbb{E}[\mu(Z_\tau \cap A)] - \mu(A)2^{-|G|}\right| + \mu(A)2^{-|G|}\right)^2 \\
&\leq \frac{\delta^2}{16(2r)^{|F|^2}} + \mu(A)^2 2^{-2|G|} - \left(-\frac{\delta^2}{16(2r)^{|F|^2}} + \mu(A)2^{-|G|}\right)^2 \\
&= \frac{\delta^2}{16(2r)^{|F|^2}} - \frac{\delta^4}{(16(2r)^{|F|^2})^2} + 2\mu(A)2^{-|G|}\frac{\delta^2}{16(2r)^{|F|^2}} \leq \frac{\delta^2}{8(2r)^{|F|^2}}.
\end{aligned}$$

Therefore Chebyshev's inequality for $\mu(Z_\tau \cap A)$ gives

$$\begin{aligned}
\kappa\left(\left\{\omega : |\mu(Z_\tau(\omega) \cap A) - \mathbb{E}[\mu(Z_\tau \cap A)]| \geq \frac{\delta}{2}\right\}\right) &\leq \frac{\mathrm{Var}(\mu(Z_\tau \cap A))}{\left(\frac{\delta}{2}\right)^2} \\
&\leq \frac{1}{2(2r)^{|F|^2}}.
\end{aligned}$$

Now since $\left|\mathbb{E}[\mu(Z_\tau \cap A)] - \mu(A)2^{-|G|}\right| < \frac{\delta}{2}$ we have

$$\kappa\left(\left\{\omega : \left|\mu(Z_\tau(\omega) \cap A) - \mu(A)2^{-|G|}\right| \geq \delta\right\}\right) \leq \frac{1}{2(2r)^{|F|^2}}.$$

Since this is true for each $\tau \in 2^G$ we have

$$\begin{aligned}
&\kappa\left(\left\{\omega : \left|\mu(Z_\tau(\omega) \cap A) - \mu(A)2^{-|G|}\right|| \geq \delta \text{ for some } \tau : G \to 2\right\}\right) \\
&\leq \frac{1}{2r^{|F|^2}}.
\end{aligned}$$

Finally, letting A range over the sets R_σ for $\sigma \in r^G$ we get

$$\kappa\left(\left\{\omega : \left|\mu(Z_\tau(\omega) \cap R^c_\sigma) - \mu(R^c_\sigma)2^{-|G|}\right| \geq \delta \text{ for some } \tau : G \to 2 \text{ and } \sigma : G \to r\right\}\right) \leq \frac{1}{2}.$$

Then any member of the nonempty complement of

$$\left\{\omega : \left|\mu(Z_\tau(\omega) \cap R_\sigma^c) - \mu(R_\sigma^c)2^{-|G|}\right| \geq \delta \text{ for some } \tau : G \to 2 \text{ and } \sigma : G \to r\right\}$$

works as $\mathcal{T}$. This completes the proof of Theorem 2.1.4.

□

We note that the proof of Theorem 2.1.4 goes through for any group Γ such that an arbitrary free action can be approximated in the uniform topology by ergodic actions - for example any group of the form $\mathbb{Z} * H$. Such an approximation is impossible if Γ has property (T), and in this case the extreme points of $\mathrm{FR}_{\sim_s}(\Gamma, X, \mu)$ are closed. Therefore the following question is natural.

Question 2.6.1. *Let* Γ *be a group without property* (T). *Can every free action of* Γ *be approximated in the uniform topology of* $A(\Gamma, X, \mu)$ *by ergodic actions?*

2.7 The space of stable weak equivalence classes

$\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ can be given the structure of a weak convex space in exactly the same way as $\mathrm{A}_\sim(\Gamma, X, \mu)$. Moreover, it is clear that for any $a \in \mathrm{A}(\Gamma, X, \mu)$ and $t \in [0, 1]$ we have $a \sim_s ta + (1 - t)a$, so $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ is in fact a convex space. Recall that the metric d_s on $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ is defined by $d_s(a, b) = d(a \times \iota, b \times \iota)$ where d is the metric on $\mathrm{A}_\sim(\Gamma, X, \mu)$.

Proposition 2.7.1. *For any* $a, b, c \in \mathrm{A}(\Gamma, X, \mu)$ *and* $t \in [0, 1]$, *we have* $d_s(ta + (1 - t)c, tb + (1 - t)c) \leq td_s(a, b)$.

It is clear that $(ta + (1 - t)c) \times \iota \sim t(a \times \iota) + (1 - t)(c \times \iota)$, so it suffices to show the following.

Proposition 2.7.2. *For any* $a, b, c \in \mathrm{A}(\Gamma, X, \mu)$ *and* $t \in [0, 1]$ *we have* $d(ta + (1 - t)c, tb + (1 - t)c) \leq td(a, b)$.

Proof. Fix n, k and write $C(a) = C_{n,k}(a)$ in order to show that $d_H(C(ta + (1 - t)c), C(tb + (1 - t)c)) \leq td_H(C(a), C(b))$. Fix $\epsilon > 0$. Let $\mathcal{P} = (P_i)_{i=1}^n$ be a partition of $X_1 \sqcup X_2$ where X_1 and X_2 are disjoint copies of X. Let $P_i^l = P_i \cap X_l$ for $l \in \{1, 2\}$. Find a partition $Q = (Q_i)_{i=1}^n$ such that for $i, j \leq n$ and $p \leq k$ we have

$$|\mu(\gamma_p^a P_i^1 \cap P_j^1) - \mu(\gamma_p^b Q_i \cap Q_j)| < d_H(C(a), C(b)) + \epsilon.$$

Then if we take $Q'_i = Q_i \sqcup P_i^2$ for all $i, j \leq n$,

$$\begin{aligned}
&|(t\mu + (1-t)\mu)(\gamma_p^{ta+(1-t)c} P_i \cap P_j) - (t\mu + (1-t)\mu)(\gamma_p^{tb+(1-t)c} Q'_i \cap Q'_j)| \\
&\qquad = |t\mu(\gamma_p^a P_i^1 \cap P_j^1) + (1-t)\mu(\gamma_p^c P_i^2 \cap P_j^2) \\
&\qquad\qquad - t\mu(\gamma_p^b Q_i \cap Q_j) - (1-t)\mu(\gamma_p^c P_i^2 \cap P_j^2)| \\
&\qquad = |t\mu(\gamma_p^a P_i^1 \cap P_j^1) + t\mu(\gamma_p^b Q_i \cap Q_j)| \leq t(d_H(C(a), C(b)) + \epsilon).
\end{aligned}$$

□

Theorem 2.1.5 now follows from Proposition 2.7.1 and Corollary 12 in [25]. Tucker-Drob and Bowen have obtained the next result independently of the author.

Proposition 2.7.3. *The extreme points of* $\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ *are precisely those stable weak equivalence classes which contain an ergodic action.*

Proof. Suppose that a is ergodic and we have $a \sim_s tb + (1-t)c$ for $t \in (0,1)$. Therefore $a \prec \iota \times (tb + (1-t)c) \sim t(b \times \iota) + (1-t)(c \times \iota)$. Since a is ergodic, Theorem 3.11 in [74] implies that $a \prec b$ and $a \prec c$. Suppose toward a contradiction that $b \not\prec_s c$, so that for some n, k we have $C_{n,k}(b) \not\subseteq \mathrm{cch}(C_{n,k}(c))$. Fixing n, k write $C(d)$ for $C_{n,k}(d)$. Let $\alpha = \sup_{x \in C(b)} p(x, \mathrm{cch}(C(c)))$ where p is the metric on $[0,1]^{n \times k \times k}$. Choose $x_0 \in C(b)$ so that $p(x_0, \mathrm{cch}(C(c))) = \alpha$. Choose $y_0 \in \mathrm{cch}(C(c))$ so that $p(x_0, y_0) = \alpha$. Consider the point $tx_0 + (1-t)y_0 \in \mathrm{cch}(C(tb + (1-t)c))$. It is easy to see that

$$p(tx + (1-t)z, ty + (1-t)z) \leq tp(y, z)$$

for any x, y, z so we have

$$\begin{aligned}
p(tx_0 + (1-t)y_0, x_0) &= p(tx_0 + (1-t)y_0, tx_0 + (1-t)x_0) \\
&\leq (1-t)p(x_0, y_0) < \alpha
\end{aligned}$$

since $0 < t$. Since $\alpha = \inf_{y \in \mathrm{cch}(C(c))} p(x_0, y)$ we see that $tx_0 + (1-t)y_0 \notin \mathrm{cch}(C(c))$ and hence $\mathrm{cch}(C(tb + (1-t)c) \not\subseteq \mathrm{cch}(C(c))$. Since for any two actions d, e we have $d \prec_s e$ if and only if $\mathrm{cch}(C_{n,k}(d)) \subseteq \mathrm{cch}(C_{n,k}(e))$ for all n, k this implies that $tb + (1-t)c \not\prec_s c$. But $tb + (1-t)c \prec_s a \prec c$ by hypothesis, so we have a contradiction and we conclude $b \prec_s c$. A symmetric argument shows $c \prec_s b$, so $b \sim_s c$. Since

$\mathrm{A}_{\sim_s}(\Gamma, X, \mu)$ obeys (2) of Definition 2.2.1, we get that $a \sim_s b \sim_s c$. Therefore if a stable weak equivalence class contains an ergodic action, it is an extreme point of $A_{\sim_s}(\Gamma, X, \mu)$. On the other hand, an argument identical to the proof of Theorem 2.1.3 shows that if the stable weak equivalence class of an action a is an extreme point of $A_{\sim_s}(\Gamma, X, \mu)$ then if we write $a = \int_Z a_z d\eta(z)$ then there is an ergodic action b such that $a_z \sim_s b$ for all $z \in Z$. Thus $a \sim_s b \times \iota \sim_s b$ and we see that a is stably weakly equivalent to an ergodic action. $\square$

Chapter 3

A TOPOLOGICAL SEMIGROUP STRUCTURE ON THE SPACE OF ACTIONS MODULO WEAK EQUIVALENCE.

Peter Burton

3.1 Introduction.

Let Γ be a countable group and let (X, μ) be a standard probability space. All partitions considered in this chapter will be assumed to be measurable. If a is a measure-preserving action of Γ on (X, μ) and $\gamma \in \Gamma$ we write γ^a for the element of $\mathrm{Aut}(X, \mu)$ corresponding to γ under a. Let $\mathrm{A}(\Gamma, X, \mu)$ be the space of measure-preserving actions of Γ on (X, μ). We have the following basic definition, due to Kechris.

Definition 3.1.1. *For actions $a, b \in \mathrm{A}(\Gamma, X, \mu)$ we say that a is* ***weakly contained*** *in b if for every partition $(A_i)_{i=1}^n$ of (X, μ), finite set $F \subseteq \Gamma$ and $\epsilon > 0$ there is a partition $(B_i)_{i=1}^n$ of (X, μ) such that*

$$\left|\mu\left(\gamma^a A_i \cap A_j\right) - \mu\left(\gamma^b B_i \cap B_j\right)\right| < \epsilon$$

for all $i, j \leq n$ and all $\gamma \in F$. We write $a \prec b$ to mean that a is weakly contained in b. We say a is ***weakly equivalent*** *to b and write $a \sim b$ if we have both $a \prec b$ and $b \prec a$. $\sim$ is an equivalence relation and we write $[a]$ for the weak equivalence class of a.*

For more information on the space of actions and the relation of weak equivalence, we refer the reader to [53]. Let $\mathrm{A}_\sim(\Gamma, X, \mu) = \mathrm{A}(\Gamma, X, \mu)/\sim$ be the set of weak equivalence classes of actions. Freeness is invariant under weak equivalence, so the set $\mathrm{FR}_\sim(\Gamma, X, \mu)$ of weak equivalence classes of free actions is a subset of $\mathrm{A}_\sim(\Gamma, X, \mu)$.

Given $[a], [b] \in \mathrm{A}_\sim(\Gamma, X, \mu)$ with representatives a and b consider the action $a \times b$ on $\left(X^2, \mu^2\right)$. We can choose an isomorphism of $\left(X^2, \mu^2\right)$ with (X, μ) and thereby regard $a \times b$ as an action on (X, μ). The weak equivalence class of the resulting action on (X, μ) does not depend on our choice of isomorphism, nor on the choice of representatives. So we have a well-defined binary operation $\times$ on $\mathrm{A}_\sim(\Gamma, X, \mu)$. This is clearly associative and commutative. In Section 3.2 we introduce a new topology on $A_\sim(\Gamma, X, \mu)$ which is finer than the one studied in [1], [23] and [74]. We call this the fine topology. The goal of this note is to prove the following result.

Theorem 3.1.1. $\times$ *is continuous with respect to the fine topology, so that in this topology* $(A_\sim(\Gamma, X, \mu), \times)$ *is a commutative topological semigroup.*

In [74], Tucker-Drob shows that for any free action a we have $a \times s_\Gamma \sim a$, where s_Γ is the Bernoulli shift on $\left([0,1]^\Gamma, \lambda^\Gamma\right)$ with λ being Lebesgue measure. Thus if we restrict attention to the free actions there is additional algebraic structure.

Corollary 3.1.1. *With the fine topology,* $(FR_\sim(\Gamma, X, \mu), \times)$ *is a commutative topological monoid.*

Acknowledgements.
We would like to thank Alexander Kechris for introducing us to this topic and posing the question of whether the product is continuous.

3.2 Definition of the fine topology.

Fix an enumeration $\Gamma = (\gamma_s)_{s=1}^\infty$ of Γ. Given $a \in A(\Gamma, X, \mu)$, $t, k \in \mathbb{N}$ and a partition $\mathcal{A} = (A_i)_{i=1}^k$ of X into k pieces let $M_{t,k}^{\mathcal{A}}(a)$ be the point in $[0,1]^{t\times k\times k}$ whose s, l, m coordinate is $\mu\left(\gamma_s^a A_l \cap A_m\right)$. Endow $[0,1]^{t\times k\times k}$ with the metric given by the sum of the distances between coordinates and let d_H be the corresponding Hausdorff metric on the space of compact subsets of $[0,1]^{t\times k\times k}$. Let $C_{t,k}(a)$ be the closure of the set

$$\left\{M_{t,k}^{\mathcal{A}}(a) : \mathcal{A} \text{ is a partition of } X \text{ into } k \text{ pieces}\right\}.$$

We have $a \sim b$ if and only if $C_{t,k}(a) = C_{t,k}(b)$ for all t, k. Define a metric d_f on $A_\sim(\Gamma, X, \mu)$ by

$$d_f([a],[b]) = \sum_{t=1}^\infty \frac{1}{2^t}\left(\sup_k d_H\left(C_{t,k}(a), C_{t,k}(b)\right)\right).$$

This is clearly finer than the topology on $A_\sim(\Gamma, X, \mu)$ discussed in the references.

Definition 3.2.1. *The topology induced by d_f is called the the* ***fine topology***.

We have $[a_n] \to [a]$ in the fine topology if and only if for every finite set $F \subseteq \Gamma$ and $\epsilon > 0$ there is N so that when $n \geq N$, for every $k \in \mathbb{N}$ and every partition $(A_l)_{l=1}^k$ of (X, μ) there is a partition $(B_l)_{l=1}^k$ so that

$$\sum_{l,m=1}^k \left|\mu\left(\gamma^{a_n} A_l \cap A_m\right) - \mu\left(\gamma^a B_l \cap B_m\right)\right| < \epsilon$$

for all $\gamma \in F$ and $l, m \leq k$.

3.3 Proof of the theorem.

We begin by showing a simple arithmetic lemma.

Lemma 3.3.1. *Suppose I and J are finite sets and $(a_i)_{i\in I}, (b_i)_{i\in I}, (c_j)_{j\in J}, (d_j)_{j\in J}$ are sequences of elements of $[0,1]$ with $\sum_{i\in I} a_i = 1$, $\sum_{j\in J} d_j = 1$, $\sum_{i\in I} |a_i - b_i| < \delta$ and $\sum_{j\in J} |c_j - d_j| < \delta$. Then $\sum_{(i,j)\in I\times J} |a_i c_j - b_i d_j| < 2\delta$.*

Proof. Fix i. We have

$$\begin{aligned}\sum_{j\in J} |a_i c_j - b_i d_j| &\leq \sum_{j\in J} (|a_i c_j - a_i d_j| + |d_j a_i + d_j b_i|) \\ &= \sum_{j\in J} (a_i |c_j - d_j| + d_j |a_i - b_i|) \\ &\leq \delta a_i + |a_i - b_i|.\end{aligned}$$

Therefore

$$\sum_{(i,j)\in I\times J} |a_i c_j - b_i d_j| \leq \sum_{i\in I} (a_i \delta + |a_i - b_i|) \leq 2\delta.$$

□

We now give the main argument.

Proof of Theorem 3.1.1. Suppose $[a_n] \to [a]$ and $[b_n] \to [b]$ in the fine topology. Fix $\epsilon > 0$ and $t \in \mathbb{N}$. Let N be large enough so that when $n \geq N$ we have

$$\max\left(\sup_k d_H\left(C_{t,k}(a_n), C_{t,k}(a)\right), \sup_k d_H\left(C_{t,k}(b_n), C_{t,k}(b)\right)\right) < \frac{\epsilon}{4}. \tag{3.1}$$

Fix $n \geq N$. Let $k \in \mathbb{N}$ be arbitrary and consider a partition $\mathcal{A} = (A_l)_{l=1}^k$ of X^2 into k pieces. Find partitions $\left(D_i^1\right)_{i=1}^p$ and $\left(D_i^2\right)_{i=1}^q$ of X such that for each $l \leq k$ there are pairwise disjoint sets $I_l \subseteq p \times q$ such that if we write $D_l = \bigcup_{(i,j)\in I_l} D_i^1 \times D_j^2$ then

$$\mu^2\left(D_l \triangle A_l\right) < \frac{\epsilon}{4k^2}. \tag{3.2}$$

Write $(\gamma_s)_{s=1}^t = F$. By (3.1) we can find a partition $\left(E_i^1\right)_{i=1}^p$ of X such that for all $\gamma \in F$ we have

$$\sum_{i,j=1}^p \left|\mu\left(\gamma^a D_i^1 \cap D_j^1\right) - \mu\left(\gamma^{a_n} E_i^1 \cap E_j^1\right)\right| < \frac{\epsilon}{4} \tag{3.3}$$

and a partition $\left(E_i^2\right)_{i=1}^q$ of X such that for all $\gamma \in F$ we have

$$\sum_{i,j=1}^{q} \left| \mu\left(\gamma^b D_i^2 \cap D_j^2\right) - \mu\left(\gamma^{b_n} E_i^2 \cap E_j^2\right) \right| < \frac{\epsilon}{4}. \tag{3.4}$$

Define a partition $\mathcal{B} = (B_l)_{l=1}^k$ of X^2 by setting $B_l = \bigcup_{(i,j)\in I_l} E_i^1 \times E_j^2$. For $\gamma \in F$ we now have

$$
\begin{aligned}
&\sum_{l,m=1}^{k} \left|\mu^2(\gamma^{a\times b} D_l \cap D_m) - \mu^2(\gamma^{a_n\times b_n} B_l \cap B_m)\right| \\
&= \sum_{l,m=1}^{k} \left|\mu^2\left(\gamma^{a\times b}\left(\bigcup_{(i_1,j_1)\in I_l} D^1_{i_1}\times D^2_{j_1}\right)\cap\left(\bigcup_{(i_2,j_2)\in I_m} D^1_{i_2}\times D^2_{j_2}\right)\right)\right. \\
&\qquad\qquad \left. - \mu^2\left(\gamma^{a_n\times b_n}\left(\bigcup_{(i_1,j_1)\in I_l} E^1_{i_1}\times E^2_{j_1}\right)\cap\left(\bigcup_{(i_2,j_2)\in I_m} E^1_{i_2}\times E^2_{j_2}\right)\right)\right| \\
&= \sum_{l,m=1}^{k} \left|\mu^2\left(\left(\bigcup_{(i_1,j_1)\in I_l} \gamma^a D^1_{i_1}\times \gamma^b D^2_{j_1}\right)\cap\left(\bigcup_{(i_2,j_2)\in I_m} D^1_{i_2}\times D^2_{j_2}\right)\right)\right. \\
&\qquad\qquad \left. - \mu^2\left(\left(\bigcup_{(i_1,j_1)\in I_l} \gamma^{a_n} E^1_{i_1}\times \gamma^{b_n} E^2_{j_1}\right)\cap\left(\bigcup_{(i_2,j_2)\in I_m} E^1_{i_2}\times E^2_{j_2}\right)\right)\right| \\
&= \sum_{l,m=1}^{k} \left|\mu^2\left(\bigcup_{\substack{(i_1,j_1,i_2,j_2)\\ \in I_l\times I_m}} \left(\gamma^a D^1_{i_1}\times \gamma^b D^2_{j_1}\right)\cap\left(D^1_{i_2}\times D^2_{j_2}\right)\right)\right. \\
&\qquad\qquad \left. - \mu^2\left(\bigcup_{\substack{(i_1,j_1,i_2,j_2)\\ \in I_l\times I_m}} \left(\gamma^{a_n} E^1_{i_1}\times E^2_{j_1}\right)\cap\left(\gamma^{b_n} E^1_{i_2}\times E^2_{j_2}\right)\right)\right| \\
&= \sum_{l,m=1}^{k} \left|\mu^2\left(\bigcup_{\substack{(i_1,j_1,i_2,j_2)\\ \in I_l\times I_m}} \left(\gamma^a D^1_{i_1}\cap D^1_{i_2}\right)\times\left(\gamma^b D^2_{j_1}\cap D^2_{j_2}\right)\right)\right. \\
&\qquad\qquad \left. - \mu^2\left(\bigcup_{\substack{(i_1,j_1,i_2,j_2)\\ \in I_l\times I_m}} \left(\gamma^{a_n} E^1_{i_1}\cap E^1_{i_2}\right)\times\left(\gamma^{b_n} E^2_{j_1}\cap E^2_{j_2}\right)\right)\right| \\
&\le \sum_{l,m=1}^{k} \sum_{\substack{(i_1,j_1,i_2,j_2)\\ \in I_l\times I_m}} \left|\mu\left(\gamma^a D^1_{i_1}\cap D^1_{i_2}\right)\mu\left(\gamma^b D^2_{j_1}\cap D^2_{j_2}\right)\right. \\
&\qquad\qquad \left. - \mu\left(\gamma^{a_n} E^1_{i_1}\cap E^1_{i_2}\right)\mu\left(\gamma^{b_n} E^2_{j_1}\cap E^2_{j_2}\right)\right| \\
&\le \sum_{\substack{(i_1,j_1,i_2,j_2)\\ \in p\times q\times p\times q}} \left|\mu\left(\gamma^a D^1_{i_1}\cap D^1_{i_2}\right)\mu\left(\gamma^b D^2_{j_1}\cap D^2_{j_2}\right) - \mu\left(\gamma^{a_n} E^1_{i_1}\cap E^1_{i_2}\right)\mu\left(\gamma^{b_n} E^2_{j_1}\cap E^2_{j_2}\right)\right| \\
&= \sum_{\substack{(i_1,i_2,j_1,j_2)\\ \in p^2\times q^2}} \left|\mu\left(\gamma^a D^1_{i_1}\cap D^1_{i_2}\right)\mu\left(\gamma^b D^2_{j_1}\cap D^2_{j_2}\right) - \mu\left(\gamma^{a_n} E^1_{i_1}\cap E^1_{i_2}\right)\mu\left(\gamma^{b_n} E^2_{j_1}\cap E^2_{j_2}\right)\right|.
\end{aligned}
\tag{3.5}
$$

Now (3.3) and (3.4) let us apply Lemma 3.3.1 with $I = p^2, J = q^2$ and $\delta = \frac{\epsilon}{4}$ to conclude that (3.5) $\leq \frac{\epsilon}{2}$. Note that for any three subsets S_1, S_2, S_3 of a probability space (Y, ν) we have

$$
\begin{aligned}
|\nu(S_1 \cap S_3) - \nu(S_2 \cap S_3)| &= |\nu(S_1 \cap S_2 \cap S_3) \\
&\quad + \nu((S_1 \setminus S_2) \cap S_3) - \nu(S_1 \cap S_2 \cap S_3) - \nu((S_2 \setminus S_1) \cap S_3)| \\
&\leq \nu(S_1 \triangle S_2),
\end{aligned}
$$

and hence for any $l, m \leq k$ and any action $c \in \mathrm{A}\left(\Gamma, X^2, \mu^2\right)$ we have

$$
\begin{aligned}
&\left|\mu^2(\gamma^c A_l \cap A_m) - \mu^2(\gamma^c D_l \cap D_m)\right| \\
&\leq \left|\mu^2(\gamma^c A_l \cap A_m) - \mu^2(\gamma^c D_l \cap A_m)\right| + \left|\mu^2\left(\gamma^c D_l \cap A_m\right) - \mu^2\left(\gamma^c D_l \cap D_m\right)\right| \\
&\leq \mu^2\left(\gamma^c A_l \triangle \gamma^c D_l\right) + \mu^2(A_m \triangle D_m) \leq \frac{\epsilon}{2k^2},
\end{aligned}
$$

where the last inequality follows from (3.2). Hence for all $\gamma \in F$,

$$
\begin{aligned}
&\sum_{l,m=1}^{k} \left|\mu^2(\gamma^{a\times b} A_l \cap A_m) - \mu^2(\gamma^{a_n \times b_n} B_l \cap B_m)\right| \\
&\leq \sum_{l,m=1}^{k} (\left|\mu^2(\gamma^a A_l \cap A_m) - \mu^2(\gamma^a D_l \cap D_m)\right| \\
&\qquad + \left|\mu^2(\gamma^{a\times b} D_l \cap D_m) - \mu^2(\gamma^{a_n \times b_n} B_l \cap B_m)\right|) \\
&\leq \sum_{l,m=1}^{k} \left(\frac{\epsilon}{2k^2} + \left|\mu^2(\gamma^{a\times b} D_l \cap D_m) - \mu^2(\gamma^{a_n \times b_n} B_l \cap B_m)\right|\right) \\
&\leq \frac{\epsilon}{2} + (3.5) \leq \epsilon.
\end{aligned}
$$

Therefore $M^{\mathcal{A}}_{t,k}(a \times b)$ is within ϵ of $M^{\mathcal{B}}_{t,k}(a_n \times b_n)$ and we have shown that for all k, $C_{t,k}(a \times b)$ is contained in the ball of radius ϵ around $C_{t,k}(a_n \times b_n)$. A symmetric argument shows that if $n \geq N$ then for all k, $C_{t,k}(a_n \times b_n)$ is contained in the ball of radius ϵ around $C_{t,k}(a \times b)$ and thus the theorem is proved. $\square$

Chapter 4

WEAK EQUIVALENCE OF STATIONARY ACTIONS AND THE ENTROPY REALIZATION PROBLEM

Peter Burton, Martino Lupini and Omer Tamuz

4.1 Introduction

Let G be a countable discrete group and let m be a probability measure on G. Let also (X, μ) be a standard probability space. A measurable action of G on (X, μ) is said to be m-stationary if the corresponding convolution of m with μ is equal to μ. More explicitly, this means $\sum_{g \in G} m(g) \cdot \mu(gA) = \mu(A)$ for all measurable subsets A of X. Stationary actions are automatically nonsingular, and form a natural intermediate class between measure-preserving actions and general nonsingular actions. We will write $\mathrm{Stat}(G, m, X, \mu)$ for the set of m-stationary actions of G on (X, μ). Given an action $a \in \mathrm{Stat}(G, m, X, \mu)$ we will write g^a for the nonsingular transformation of (X, μ) corresponding to g.

In [53], Kechris defined a notion of weak containment for measure-preserving actions of countable groups analogous to the standard notion of weak containment for unitary representations. The same definition can be given for stationary actions.

Definition 4.1.1. *Let $a, b \in \mathrm{Stat}(G, m, X, \mu)$. We say that a is weakly contained in b, in symbols $a \preceq b$, if the following condition holds. For every $\epsilon > 0$, every finite $F \subseteq G$ and every finite collection $A_1, \ldots, A_n$ of measurable subsets of X, there are measurable subsets $B_1, \ldots, B_n$ of X such that*

$$|\mu(g^a A_i \cap A_j) - \mu(g^b B_i \cap B_j)| < \epsilon$$

for all $g \in F$ and all $i, j \in \{1, \ldots, n\}$. We say that a is weakly equivalent to b, in symbols $a \sim b$, if $a \preceq b$ and $b \preceq a$.

Thus a is weakly contained in b if the statistics of a on finite partitions can be simulated arbitrary well in the action b. Weak equivalence is a much coarser relation than isomorphism; for example in [36] it is shown that all free measure-preserving actions of an amenable group are weakly equivalent. It is also better behaved from the perspective of descriptive set theory: there is in general no standard Borel structure on the set of isomorphism classes of m-stationary actions, whereas in Section 4.3 we will define a natural Polish topology on the set of weak equivalence classes of m-stationary actions for any pair (G, m).

In [38], Furstenberg introduced an invariant $h_m(X, \mu, a)$ which quantifies how far an m-stationary action a is from being measure-preserving. Later termed Furstenberg entropy, this is defined by

$$h_m(X, \mu, a) = -\sum_{g \in G} m(g) \cdot \int_X \log \frac{\mathrm{d}g^a \mu}{\mathrm{d}\mu}(x) \mathrm{d}\mu(x).$$

By Jensen's inequality, we have that $h_m(X, \mu, a)$ is nonnegative, and it is zero if and only if a is measure-preserving. The following problem has been studied in articles such as [19], [21], [31], [46], [52] and [65].

Problem 4.1.1 (Furstenberg entropy realization problem)**.** *For a fixed pair (G, m), describe the possible values of Furstenberg entropy on ergodic ν-stationary systems.*

The goal of this paper is to establish the following theorem, which shows that the above problem can be regarded as a problem about the structure of the space of weak equivalence classes.

Theorem 4.1.1. *Furstenberg entropy is an invariant of weak equivalence and descends to a continuous function on the space of weak equivalence classes.*

4.2 A characterization of weak containment

In this section we verify that one obtains an equivalent notion if one alters the definition of weak containment to allow shifts on both sides of the intersections.

Proposition 4.2.1. *Let $a, b \in \mathrm{Stat}(G, m, X, \mu)$. Then the following are equivalent.*

(i) a is weakly contained b.

(ii) For any finite subset F of G, $\epsilon > 0$, and measurable subsets $A_1, \ldots, A_n$ of X, there exist measurable subsets $B_1, \ldots, B_n$ of X such that

$$\left|\mu(g^a A_i \cap h^a A_j) - \mu(g^b B_i \cap h^b B_j)\right| < \epsilon \tag{4.1}$$

for all $g, h \in F$ and $i, j \in \{1, \ldots, n\}$.

Proof. Taking $h = 1_G$ it is clear that (ii) implies (i). We now show (i) implies (ii). Suppose that $F = \{g_0, \ldots, g_m\}$ is a finite subset of G, n is a natural number, and $A_0, \ldots, A_n$ are measurable subsets of X. Without loss of generality, we can assume that $n = m$, $g_0 = 1_G$ and $A_0 = X$. Fix $\epsilon > 0$ and choose $0 < \delta < \epsilon/7$. Set $A_{i,j} = g_j^a A_i$ for $i, j \in \{1, \ldots, n\}$. In particular we have $A_{i,0} = A_i$ and $A_{i,j} = g_j^a A_{i,0}$ for $i, j \in \{1, \ldots, n\}$. By assumption there exist measurable subsets $B_{i,j}$ of X such that

$$|\mu(A_{i,j} \cap g_m^a A_{l,k}) - \mu(B_{i,j} \cap g_m^b B_{l,k})| < \delta$$

for all $i, j, k, l, m \in \{1, \ldots, n\}$. Since $A_{0,0} = X$ and $g_0 = 1_G$, we have that $\mu(B_{0,0}) > 1 - \delta$. It follows that

$$|\mu(g_m^a A_{l,k}) - \mu(g_m^b B_{l,k})| < 2\delta$$

for $m, l, k \in \{1, \ldots, n\}$. Therefore

$$\begin{aligned}
\mu(B_{j,m} \triangle g_m^b B_{j,0}) &= \mu(B_{j,m}) + \mu(g_m^b B_{j,0}) - 2\mu(B_{j,m} \cap g_m^b B_{j,0}) \\
&\leq 6\delta + \mu(A_{j,m}) + \mu(g_m^a A_{j,0}) - 2\mu(A_{j,m} \cap g_m^a A_{j,0}) \\
&= 6\delta + \mu(A_{j,m} \triangle g_m^a A_{j,0}) = 6\delta.
\end{aligned}$$

In conclusion

$$|\mu(g_k^a A_i \cap g_m^a A_j) - \mu(g_k^b B_{i,0} \cap g_m^b B_{j,0})| \leq |\mu(g_k^a A_{i,0} \cap A_{j,m}) - \mu(g_k^b B_{i,0} \cap B_{j,m})| + 6\delta \leq 7\delta < \epsilon.$$

for every $i, j, k, m \in \{1, \ldots, n\}$. Thus we can take $B_i = B_{i,0}$ to obtain (4.1). □

4.3 The space of weak equivalence classes

For $a \in \mathrm{Stat}(G, m, X, \mu)$ we will write $\tilde{a}$ for the weak equivalence class of a. Let $(g_k)_{k=1}^{\infty}$ be an enumeration of G. For a natural number m and an ordered finite partition $\overline{A} = \{A_1, \ldots, A_n\}$ of X, we will write $M_{m,\overline{A}}(a)$ for the point in $[0,1]^{m \times n \times n}$ whose (k, i, j)-coordinate is $\mu(g_k^a A_i \cap A_j)$. Let then $C_{m,n}(a)$ be the closure in $[0,1]^{m \times n \times n}$ of the set

$$\left\{ M_{m,\overline{A}} : \overline{A} \text{ is a partition of } X \text{ into } n \text{ pieces.} \right\}.$$

Clearly we have $a \preceq b$ if and only if $C_{m,n}(a) \subseteq C_{m,n}(b)$ for all natural numbers m, n. Let

$$\delta(a, b) = \sum_{m,n=1}^{\infty} \frac{1}{2^{m+n}} \cdot d_H\big(C_{m,n}(a), C_{m,n}(b)\big),$$

where d_H is the Hausdorff distance on the space of compact subsets of $[0,1]^{m\times n\times n}$. Then for any $a, b, c, d \in \mathrm{Stat}(G, m, X, \mu)$ with $a \sim c$ and $b \sim d$ we have $\delta(a, b) = \delta(c, d)$. Thus the quantity $\tilde{\delta}(\tilde{a}, \tilde{b}) = \delta(a, b)$ is a well-defined metric on the space of weak equivalence classes. The corresponding topology is easily seen to be Polish. We denote this space by $\widetilde{\mathrm{Stat}}(G, m, X, \mu)$. As in the measure-preserving case, an ultraproduct construction shows that $\widetilde{\mathrm{Stat}}(G, m, X, \mu)$ is compact.

In addition to its topology, $\widetilde{\mathrm{Stat}}(G, m, X, \mu)$ carries a convex structure. Given $a, b \in \mathrm{Stat}(G, m, X, \mu)$, and $t \in (0, 1)$ one can realize a as an action on $[0, t)$ and realize b as an action on $[t, 1]$. One then defines $ta + (1-t)b$ to be the action on $[0, 1]$ which agrees with a on $[0, t)$ and b on $[t, 1]$. It is easy to see that this procedure gives a well-defined operation on $\widetilde{\mathrm{Stat}}(G, m, X, \mu)$. As in the measure-preserving case discussed in [23], the convex structure is better behaved if one instead considers the relation $\preceq_s$ of stable weak containment. This is defined by letting $a \preceq_s b$ if and only if $a \preceq b \times \iota$, where ι is the trivial action of G on a standard probability space. Write $\widetilde{\mathrm{Stat}}_s(G, m, X, \mu)$ for the space of stable weak equivalence classes. $\tilde{\delta}$ gives a Polish topology on $\widetilde{\mathrm{Stat}}_s(G, m, X, \mu)$ and since $h_m(X, \mu, a \times \iota) = h_m(X, \mu, a)$, Theorem 4.1.1 continues to hold if we replace weak equivalence by stable weak equivalence. The arguments from [23] carry over to show that $\widetilde{\mathrm{Stat}}_s(G, m, X, \mu)$ is isomorphic to a compact convex subset of a Banach space, and that its extreme points are exactly those stable weak equivalence classes containing an ergodic action. Moreover, the map $a \mapsto h_m(X, \mu, a)$ respects the convex combination operation. Thus understanding the convex structure of $\widetilde{\mathrm{Stat}}_s(G, m, X, \mu)$ could give new understanding of Problem 4.1.1.

4.4 Proof of Theorem 4.1.1

For each n, let $a_n \in \mathrm{Stat}(G, m, X, \mu)$; let also $a \in \mathrm{Stat}(G, m, X, \mu)$. Assume that $\tilde{a}_n$ converges to $\tilde{a}$ in $\widetilde{\mathrm{Stat}}(G, m, X, \mu)$. Fixing $g \in G$, it is enough to show the following: for any $c \geq 0$ we have

$$\lim_{n\to\infty} \mu\left(\left\{x \in X : \frac{\mathrm{d}g^{a_n}\mu}{\mathrm{d}\mu}(x) > c\right\}\right) = \mu\left(\left\{x \in X : \frac{\mathrm{d}g^{a}\mu}{\mathrm{d}\mu}(x) > c\right\}\right).$$

Let M be a positive constant such that $\frac{\mathrm{d}g^b\mu}{\mathrm{d}\mu} \leq M$ for any m-stationary action a. Let $\omega_n = \frac{\mathrm{d}g^a\mu}{\mathrm{d}\mu}$ and $\omega_n = \frac{\mathrm{d}g^{a_n}\mu}{\mathrm{d}\mu}$. Write $C = \{x \in X : \omega(x) > c\}$, and $C_n = \{x \in X : \omega_n(x) > c\}$. We will prove that $\mu(C) \leq \liminf_n \mu(C_n)$. The proof that $\mu(C) \geq \limsup_n \mu(C_n)$ is analogous. Suppose by contradiction $\mu(C) >$

$\liminf_n \mu(C_n)$. Thus, after passing to a subsequence, we can assume that there is $\delta > 0$ such that $\mu(C_n) \leq \mu(C) - \delta$ for every $n \in \mathbb{N}$. Identify X with $[0, 1]$, so that we have a Borel linear order on X. Define the Borel linear order $\sqsubseteq$ on X by letting $t \sqsubseteq s$ iff $\omega(t) < \omega(s)$ or $\omega(t) = \omega(s)$ and $t < s$. Similarly define $\sqsubseteq_n$ in terms of ω_n. Note that if D is a terminal segment of $\sqsubseteq$ then we have $\mu(g^a D) \geq \mu(g^a E)$ for any E with $\mu(E) = \mu(D)$. For $n \in \mathbb{N}$ write D_n for the terminal segment of $\sqsubseteq$ such that $\mu(D_n) = \mu_n(C_n)$ and write E_n for the terminal segment of $\sqsubseteq_n$ such that $\mu(C) = \mu_n(E_n)$. Let also F_n be the terminal segment of $\sqsubseteq$ such that $\mu(F_n) = \mu(C_n) + \delta$ and let K_n be the terminal segment of $\sqsubseteq_n$ such that $\mu_n(K_n) = \mu(C_n) + \delta$. Clearly $D_n \subseteq F_n \subseteq C$ and $C_n \subseteq K_n \subseteq E_n$. We have

$$\mu(F_n \setminus D_n) = \mu(F_n) - \mu(D_n) = \delta = \mu_n(K_n) - \mu_n(C_n) = \mu_n(K_n \setminus C_n) \tag{4.2}$$

and similarly

$$\mu(C \setminus F_n) = \mu_n(E_n \setminus K_n). \tag{4.3}$$

Note that since $\omega(x) > c \geq \omega_n(y)$ if $x \in C$ but $y \in X \setminus C_n$, (4.3) implies

$$\mu\left(g^a(C \setminus F_n)\right) \geq \mu_n\left(g^{a_n}(E_n \setminus K_n)\right). \tag{4.4}$$

Let H be the terminal segment of $\sqsubseteq$ such that $\mu(H) = \mu(C) - \delta$ so that by (4.2) we have $\delta = \mu(C \setminus H) = \mu(F_n \setminus D_n)$. Since $F_n \setminus D_n \subseteq C$ and $C \setminus H$ has the lowest Radon-Nikodym derivative of any subset of C with measure δ this implies

$$\mu\left(g^a(C \setminus H)\right) \leq \mu\left(g^a(F_n \setminus D_n)\right). \tag{4.5}$$

For $n \in \ltimes$ from (4.2), (4.4) and (4.5) we have

$$\begin{aligned}
&\mu\left(g^a(C \setminus D_n)\right) - \mu\left(g^{a_n}(E_n \setminus C_n)\right) \\
&= \mu\left(g^a(C \setminus F_n)\right) + \mu\left(g^a(F_n \setminus D_n)\right) - \mu\left(g^{a_n}(E_n \setminus K_n)\right) - \mu\left(g^{a_n}(K_n \setminus C_n)\right) \\
&\geq \mu\left(g^a(F_n \setminus D_n)\right) - \mu\left(g^{a_n}(K_n \setminus C_n)\right) \geq \mu\left(g^a(F_n \setminus D_n)\right) - c \cdot \mu\left(K_n \setminus C_n\right) \\
&= \mu\left(g^a(F_n \setminus D_n)\right) - c\delta \geq \mu\left(g^a(C \setminus H)\right) - c\delta.
\end{aligned} \tag{4.6}$$

For $x \in C$ we have $\omega(x) > c$ so the last quantity is strictly positive. Choose

$$0 < \varepsilon < \frac{1}{2(4+M)} \cdot \left(\mu\left(g^a(C \setminus H)\right) - c\delta\right). \tag{4.7}$$

Since $\tilde{a}_n \to \tilde{a}$, for every Borel partition $A_1, \ldots, A_k$ of X there is a partition $B_1, \ldots, B_k$ of X such that $|\mu(A_i) - \mu(B_i)| < \varepsilon$ and $\left|\mu(g^a A_i \cap A_j) - \mu(g^{a_n} B_i \cap B_j)\right| < \varepsilon$ for all

$i, j \in \{1, \ldots, k\}$. Fixing n, write $C' = C_n$, $D = D_n$, $E = E_n$ and $\sqsubseteq' = \sqsubseteq_n$. Note that from (4.6) and (4.7) we have

$$2(4+M)\varepsilon < \mu\left(g^{a}(C \setminus D)\right) - \mu\left(g^{a_n}(E \setminus C')\right). \tag{4.8}$$

Let $A_1 = X \setminus C$ and $A_2 = C$. Find $B_1, B_2 \subseteq X$ such that $|\mu(A_i) - \mu(B_i)| < \varepsilon$ and

$$\left|\mu(g^{a}A_i \cap A_j) - \mu(g^{a_n}B_i \cap B_j)\right| < \varepsilon$$

for each $i, j \in \{1, 2\}$. Note that

$$\mu\left(X \setminus (B_1 \cup B_2)\right) \leq 2\varepsilon.$$

We have

$$\begin{aligned}
\mu(g^{a}A_1) &= \mu(g^{a}A_1 \cap A_1) + \mu(g^{a}A_1 \cap A_2) \\
&\geq \mu(g^{a_n}B_1 \cap B_1) + \mu(g^{a_n}B_1 \cap B_2) - 2\varepsilon \\
&\geq \mu(g^{a_n}B_1 \cap B_1) + \mu(g^{a_n}B_1 \cap B_2) + \mu\left(g^{a_n}B_1 \setminus (B_1 \cup B_2)\right) - 4\varepsilon \\
&\geq \mu(g^{a_n}B_1) - 4\varepsilon. \qquad (4.9)
\end{aligned}$$

Note that

$$\mu(B_1) \geq \mu(A_1) - \varepsilon = \mu(X \setminus C) - \varepsilon.$$

Write L for the initial segment of $\sqsubseteq'$ such that $\mu(L) = \mu(X \setminus C) - \varepsilon$. Note that $\mu(X \setminus E) = \mu(X \setminus C)$ and so $\mu\left(X \setminus (E \cup L)\right) = \varepsilon$. We have

$$\mu\left(g^{a_n}(X \setminus E)\right) = \mu(g^{a_n}L) + \mu\left(g^{a_n}\left(X \setminus (E \cup L)\right)\right)$$

and therefore

$$\mu(g^{a_n}L) \geq \mu\left(g^{a_n}(X \setminus E)\right) - M\varepsilon. \tag{4.10}$$

Since $\mu(B_1) \geq \mu(L)$ and $\mu(g^{a_n}L) \leq \mu(g^{a_n}J)$ for any $J \subseteq X$ with $\mu(J) \geq \mu(L)$ from (4.10) we see

$$\mu(g^{a_n}B_1) \geq \mu\left(g^{a_n}(X \setminus E)\right) - M\varepsilon.$$

From (4.9) we have

$$\mu\left(g^{a}(X \setminus C)\right) \geq \mu\left(g^{a_n}(X \setminus E)\right) - (4+M)\varepsilon. \tag{4.11}$$

Now write $A_1 = C'$ and $A_2 = X \setminus C'$. Find $B_1, B_2 \subseteq X$ such that

$$\left|\mu(A_i \cap A_j) - \mu(B_i \cap B_j)\right| < \varepsilon$$

and

$$\left|\mu(g^{a_n}A_i \cap A_j) - \mu(g^a B_i \cap B_j)\right| < \varepsilon$$

for each $i, j \in \{1, 2\}$. Arguing as before we have $\mu(g^{a_n}A_1) \leq \mu(g^a B_1) + 4\varepsilon$ and $\mu(g^a B_1) \leq \mu(g^a D) + M\varepsilon$ so that

$$\mu(g^{a_n}C') \leq \mu(g^a D) + (4 + M)\varepsilon. \tag{4.12}$$

From (4.11) and (4.12) we have

$$\mu\left(g^a(X \setminus C) \cup D)\right) \geq \mu\left(g^{a_n}\left((X \setminus E) \cup C'\right)\right) - 2(4 + M)\varepsilon. \tag{4.13}$$

Note that

$$D \sqcup (C \setminus D) \sqcup (X \setminus C) = X$$

and

$$C' \sqcup (E \setminus C') \sqcup (X \setminus E) = X.$$

Thus from (4.8) and (4.13) we have

$$\begin{aligned}
1 &= \mu\left(g^a\left(D \cup (X \setminus C)\right)\right) + \mu\left(g^a(C \setminus D)\right) \\
&\geq \mu\left(g^{a_n}\left(C' \cup (X \setminus E)\right)\right) - 2(4 + M)\varepsilon + \mu\left(g^a(C \setminus D)\right) \\
&> \mu\left(g^{a_n}\left(C' \cup (X \setminus E)\right)\right) + \mu\left(g^{a_n}(E \setminus C')\right) = 1,
\end{aligned}$$

which is the desired contradiction. This concludes the proof of Theorem 4.1.1.

Part II

Sofic entropy

Chapter 5

NAIVE ENTROPY OF DYNAMICAL SYSTEMS

Peter Burton

5.1 Introduction.

A fundamental aspect of the theory of dynamical systems is the invariant known as entropy. Defined for both measurable and topological systems, this is a nonnegative real number which quantifies how random the given dynamics are. Entropy was introduced for measurable $\mathbb{Z}$-systems by Kolmogorov in [64] and Sinai in [73] and for topological $\mathbb{Z}$-systems by Adler, Konheim and McAndrew in [5]. In [67], Ornstein and Weiss extended much of entropy theory from $\mathbb{Z}$-systems to Γ-systems for amenable groups Γ. More recently, there has been significant progress in creating ideas of entropy for systems where the acting group is nonamenable. The most significant aspect of this new work is Bowen's theory of sofic entropy, initially developed by him for measurable systems in the papers [14], [15], [18] and [12], and further developed for both types of systems by Kerr and Li in [61], [63] and [62] and by Kerr in [58] and [59]. Another thread is the concept of Rokhlin entropy, developed for measurable systems by Seward in [70], [71] and [72]. In this paper we begin to study a third notion of entropy for general systems, called naive entropy. This idea was suggested by Bowen in [12] as the most direct way of generalizing the definition for $\mathbb{Z}$-systems. While he considered only the measurable context, a similar definition can be made for topological systems.

Following an observation of Bowen, we show that if Γ is a nonamenable countable group then any topological or measurable Γ-system has naive entropy either 0 or ∞. Thus for nonamenable groups naive entropy is best understood as a dichotomy rather than an invariant. A natural question is to what extent the dichotomy between zero and infinite naive entropy corresponds to the dichotomy between zero and positive sofic entropy. Bowen has conjectured in [12] that zero naive entropy implies sofic entropy at most zero. In Section 5.4 we prove the following topological version of this conjecture. Here $h^{\mathrm{tp}}_{\mathrm{nv}}$ is the naive topological entropy and h^{tp}_{Σ} is the sofic entropy with respect to a sofic approximation Σ.

Theorem 5.1.1. *Let Γ be a sofic group, let $\Gamma \curvearrowright X$ be a topological Γ-system and let Σ be a sofic approximation to Γ. If $h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright X) = 0$ then $h^{\mathrm{tp}}_{\Sigma}(\Gamma \curvearrowright X) \leq 0$.*

One advantage of naive entropy is that in many cases it is easy to see that a system has zero naive entropy. For example in Section 5.2 we observe that if Γ has an element of infinite order, then any distal Γ-system has zero naive entropy in both senses. This gives a partial answer to a question of Bowen. Furthermore, in Section 5.2 we are able show that if Γ is a free group, then the generic Γ-system with phase space the Cantor set has zero naive topological entropy. More precisely, if X is a compact metric space and Γ a countable group, let $\mathrm{A}_{\mathrm{top}}(\Gamma, X)$ denote the Polish space of topological Γ-systems with phase space X. We say a sequence $(\Gamma \curvearrowright^{a_n} X)_{n=1}^{\infty} \subseteq \mathrm{A}_{\mathrm{top}}(\Gamma, X)$ of Γ-systems converges to a system $\Gamma \curvearrowright^{a} X$ if for every $\gamma \in \Gamma$ the sequence of homeomorphisms corresponding to γ in a_n converges uniformly to the homeomorphism corresponding to γ in a.

Theorem 5.1.2. *Let $2^{\mathbb{N}}$ denote the Cantor set and let $\mathbb{F}$ be any countable free group. The set of topological $\mathbb{F}$-systems with zero naive entropy is comeager in $\mathrm{A}_{\mathrm{top}}\left(\mathbb{F}, 2^{\mathbb{N}}\right)$.*

Combining Theorems 5.1.1 and 5.1.2 we have the following corollary.

Corollary 5.1.1. *If $\mathbb{F}$ is a countable free group, then the set of $\mathbb{F}$-systems with sofic entropy at most 0 is comeager in $\mathrm{A}_{\mathrm{top}}\left(\mathbb{F}, 2^{\mathbb{N}}\right)$.*

Another natural question to ask is whether there is a relation between naive measure entropy and naive topological entropy. In Section 5.2 we show half of such a variational principle. Let h_{nv} denote the naive measure entropy.

Theorem 5.1.3. *If $\Gamma \curvearrowright X$ is a topological Γ-system and μ is an invariant measure for $\Gamma \curvearrowright X$ then*

$$h_{\mathrm{nv}}(\Gamma \curvearrowright (X, \mu)) \leq h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright X).$$

Notational preliminaries.

Throughout the paper Γ will denote a countable discrete group. A measurable Γ-system $\Gamma \curvearrowright^{a} (X, \mu)$ consists of a standard probability space (X, μ) and measure-preserving action on Γ on (X, μ), equivalently a homomorphism $a : \Gamma \to \mathrm{Aut}(X, \mu)$, where $\mathrm{Aut}(X, \mu)$ is the group of measure-preserving bijections from (X, μ) to itself.

We use Kechris's convention from [53] and write γ^a instead of $a(\gamma)$ for $\gamma \in \Gamma$. We identify two measure-preserving bijections if they agree almost everywhere, and thus identify two Γ-systems $\Gamma \curvearrowright^a (X, \mu)$ and $\Gamma \curvearrowright^b (X, \mu)$ if $\gamma^a = \gamma^b$ almost everywhere for each $\gamma \in \Gamma$.

A topological Γ-system $\Gamma \curvearrowright^a X$ consists of a compact metrizable space X and a homomorphism $a : \Gamma \to \mathrm{Homeo}(X)$, where $\mathrm{Homeo}(X)$ is the group of homeomorphisms of X. As in the measurable case, we write γ^a instead of $a(\gamma)$. If $\Gamma = \mathbb{Z}$ we use the standard notation and write $a(1) = T$, denoting the system by (X, T) or (X, μ, T).

For $n \in \mathbb{N}$, we let $[n]$ denote the set $\{1, \ldots, n\}$.

Acknowledgments.

We thank Alexander Kechris for introducing us to this topic, and Lewis Bowen for allowing us to read his preprint [12]. We also thank the anonymous referee for numerous helpful comments. This research was partially supported by NSF grant DMS-0968710.

Additional note.

After communicating our results to Brandon Seward, he informed us that the measurable case of Bowen's naive entropy conjecture has been proved independently by a number of researchers including Miklos Abert, Tim Austin, Seward himself, and Benjamin Weiss. This together with our Theorem 5.1.3, the variational principle for sofic entropy and the fact that a topological system with no invariant measure has sofic entropy $-\infty$ give an alternate, indirect proof of our Theorem 5.1.1. Our work was done independently of the (as yet unpublished) work of these authors on the measurable case.

5.2 Naive entropy.

Naive measure entropy.

In this section we introduce the naive measure entropy of a dynamical system. Fix a measurable Γ-system $\Gamma \curvearrowright^a (X, \mu)$. All partitions considered will be assumed to be measurable. If $\alpha = (A_1, \ldots, A_n)$ is a finite partition of (X, μ) the Shannon entropy

$H_\mu(\alpha)$ of α is defined by

$$H_\mu(\alpha) = -\sum_{i=1}^{n} \mu(A_i)\log(\mu(A_i)).$$

If α and β are partitions of (X, μ), the join $\alpha \vee \beta$ is the partition consisting of all intersections $A \cap B$ where $A \in \alpha$ and $B \in \beta$. We make a similar definition for the join $\bigvee_{i=1}^{n} \alpha_i$ of a finite family $(\alpha_i)_{i=1}^{n}$ of partitions. If α is partition and $\gamma \in \Gamma$ we let $\gamma^a\alpha$ be the partition $\{\gamma^a A : A \in \alpha\}$. For a finite set $F \subseteq \Gamma$ let α^F denote the partition $\bigvee_{\gamma \in F} \gamma^a\alpha$. If (X, μ, T) is a $\mathbb{Z}$-system and $F = [0, n]$ we write α_0^n for α^F. Recall the classical definition of entropy for $\mathbb{Z}$-systems.

Definition 5.2.1. *Let (X, μ, T) be a measurable $\mathbb{Z}$-system. The dynamical entropy $h_\mu(\alpha)$ of a finite partition α is defined by*

$$h_\mu(\alpha) = \inf_{n \in \mathbb{N}} \frac{1}{n} H_\mu\left(\alpha_0^n\right).$$

The ***measure entropy*** *$h(X, \mu, T)$ of the system is defined by*

$$h(X, \mu, T) = \sup\{h_\mu(\alpha) : \alpha \textit{ is a finite partition of } X.\}$$

See Chapter 14 of [42] for more information on the entropy of $\mathbb{Z}$-systems. In [12], L. Bowen has introduced the following analog of Definition 5.2.1.

Definition 5.2.2. *Let $\Gamma \curvearrowright (X, \mu)$ be a measurable Γ-system. The dynamical entropy $h_\mu(\alpha)$ of a finite partition α is defined by*

$$h_\mu(\alpha) = \inf_{F} \frac{1}{|F|} H_\mu\left(\alpha^F\right),$$

where the infimum is over all nonempty finite subsets F of Γ. The ***naive measure entropy*** *$h_{\mathrm{nv}}(\Gamma \curvearrowright (X, \mu))$ of the system is defined by*

$$h_{\mathrm{nv}}(\Gamma \curvearrowright (X, \mu)) = \sup\{h_\mu(\alpha) : \alpha \textit{ is a finite partition of } X\}.$$

In the case of $\mathbb{Z}$, Definition 5.2.2 agrees with Definition 5.2.1. The next fact was proven by Bowen in [12].

Theorem 5.2.1. *If Γ is nonamenable then for any measurable Γ-system $\Gamma \curvearrowright (X, \mu)$ we have $h_{\mathrm{nv}}(\Gamma \curvearrowright (X, \mu)) \in \{0, \infty\}$.*

Proof. Suppose there is a finite partition α with $h_\mu(\alpha) = c > 0$. Choose $r \in \mathbb{R}$. Since Γ is nonamenable, there is a finite set $W \subseteq \Gamma$ such that

$$\inf_F \frac{|WF|}{|F|} \geq \frac{r}{c},$$

where the infimum is over all nonempty finite subsets of Γ. Then we have

$$\begin{aligned} h_\mu\left(\alpha^W\right) &= \inf_F \frac{1}{|F|} H_\mu\left(\alpha^{WF}\right) \\ &= \inf_F \frac{|WF|}{|F|}\left(\frac{1}{|WF|} H_\mu\left(\alpha^{WF}\right)\right) \\ &\geq \inf_F \frac{|WF|}{|F|} h_\mu(\alpha) \\ &\geq r. \end{aligned}$$

□

Naive topological entropy.

In this section we introduce the naive topological entropy of a dynamical system. Fix a topological Γ-system $\Gamma \curvearrowright^a X$. If $\mathcal{U}$ is an open cover of a compact metric space X, let $N(\mathcal{U})$ denote the minimal cardinality of a subcover of $\mathcal{U}$. If $\mathcal{U}$ and $\mathcal{V}$ are open covers of X, the join $\mathcal{U} \vee \mathcal{V}$ is the open cover consisting of all intersections $U \cap V$ where $U \in \mathcal{U}$ and $V \in \mathcal{V}$. We make a similar definition for the join $\bigvee_{i=1}^n \mathcal{U}_i$ of a finite family $(\mathcal{U}_i)_{i=1}^n$ of open covers. If $\mathcal{U}$ is an open cover and $\gamma \in \Gamma$ we let $\gamma^a \mathcal{U}$ be the open cover $\{\gamma^a U : U \in \mathcal{U}\}$. For a finite set $F \subseteq \Gamma$, write $\mathcal{U}^F$ to refer to $\bigvee_{\gamma \in F} \gamma^a \mathcal{U}$. If (X, T) is a $\mathbb{Z}$-system and $F = [0, n]$ we write $\mathcal{U}_0^n$ for $\mathcal{U}^F$. Again we recall the definition of entropy for $\mathbb{Z}$-systems.

Definition 5.2.3. *Let (X, T) be a topological $\mathbb{Z}$-system. The entropy $h^{\mathrm{tp}}(\mathcal{U})$ of a finite open cover $\mathcal{U}$ is defined by*

$$h^{\mathrm{tp}}(\mathcal{U}) = \inf_{n \in \mathbb{N}} \frac{1}{n} \log\left(N\left(\mathcal{U}_0^n\right)\right),$$

*and the **topological entropy** $h^{\mathrm{tp}}(X, T)$ of the system is defined by*

$$h^{\mathrm{tp}}(\mathbb{Z} \curvearrowright X) = \sup\{h^{\mathrm{tp}}(\mathcal{U}) : \mathcal{U} \textit{ is a finite open cover of } X\}.$$

Following Definition 5.2.2 we make the following definition.

Definition 5.2.4. *Let $\Gamma \curvearrowright X$ be a topological Γ-system. Given a finite open cover $\mathcal{U}$ of X we define the entropy $h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{U})$ of $\mathcal{U}$ by*

$$h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{U}) = \inf_F \frac{1}{|F|} \log\left(N\left(\mathcal{U}^F\right)\right),$$

where the infimum is over all nonempty finite subsets of Γ*. We define the* ***naive topological entropy*** $h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright X)$ *of* $\Gamma \curvearrowright X$ *by*

$$h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright X) = \sup\{h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{U}) : \mathcal{U} \text{ is a finite open cover of } X\}.$$

A similar concept has been studied in [9], [10] and [40] and is discussed the text [33]. If Γ has a finite generating set S, these authors define the entropy of an open cover $\mathcal{U}$ by the formula

$$\limsup_{n\to\infty} \frac{1}{n} \log\left(N\left(\mathcal{U}^{S^n}\right)\right)$$

and the entropy of the system by taking the supremum over finite open covers. Clearly a system with zero entropy in this sense has $h^{\mathrm{tp}}_{\mathrm{nv}}$ equal to zero. Hence we work with $h^{\mathrm{tp}}_{\mathrm{nv}}$ in order to get the strongest form of Theorem 5.1.1. An identical argument to the proof of Theorem 5.2.1 shows that if Γ is nonamenable then any topological Γ-system has naive topological entropy either 0 or ∞.

We record the following observation, which is immediate from the definition.

Proposition 5.2.1. *If* $h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright^a X) > 0$ *then for every* $\gamma \in \Gamma$ *with infinite order we have* $h^{\mathrm{tp}}(X, \gamma^a) > 0$*, where we regard* (X, γ^a) *as a* $\mathbb{Z}$*-system.*

Equivalent definitions of naive topological entropy.

We now introduce two standard reformulations of the definition of naive topological entropy, due originally in the case of $\mathbb{Z}$ to R. Bowen. For a metric space (X, d) and $\epsilon > 0$ say a set $S \subseteq X$ is ϵ-separated if for each distinct pair $x_1, x_2 \in S$ we have $d(x_1, x_2) \geq \epsilon$. Say that S is ϵ-spanning if for every $x \in X$ there is $x_0 \in S$ with $d(x, x_0) \leq \epsilon$. Define $\mathrm{Sep}(X, \epsilon, d)$ to be the maximal cardinality of an ϵ-separated subset of X, and $\mathrm{Span}(X, \epsilon, d)$ to be the minimal cardinality of an ϵ-spanning subset of X. It is clear that

$$\mathrm{Span}(X, \epsilon, d) \leq \mathrm{Sep}(X, \epsilon, d) \leq \mathrm{Span}\left(X, \frac{\epsilon}{2}, d\right). \tag{5.1}$$

Now, fix a Γ-system $\Gamma \curvearrowright^a X$ and a compatible metric d on X. For a nonempty finite subset $F \subseteq \Gamma$ define a metric d_F on X by letting $d_F(x_1, x_2) = \max_{\gamma\in F} d\left(\gamma^a x_1, \gamma^a x_2\right)$. The proof of the following is an immediate generalization of the corresponding statement for $\mathbb{Z}$-systems, which can be found as Proposition 14.11 in [42].

Proposition 5.2.2. *Letting* F *range over the nonempty finite subsets of* Γ *we have*

$$h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright^a X) = \sup_{\epsilon>0} \inf_F \frac{1}{|F|} \log(\mathrm{Sep}(X, \epsilon, d_F)) = \sup_{\epsilon>0} \inf_F \frac{1}{|F|} \log(\mathrm{Span}(X, \epsilon, d_F)).$$

Proof. Fix $\epsilon > 0$ and $F \subseteq \Gamma$ finite. Write F^{-1} for $\{\gamma^{-1} : \gamma \in F\}$. Let $\mathcal{U}$ be an open cover of X with Lebesgue number ϵ. Let $S \subseteq X$ be an ϵ-spanning set of minimal cardinality with respect to $d_{F^{-1}}$. For every $x \in X$ there is $s \in S$ with $d(\gamma^a x, \gamma^a s) \leq \epsilon$ for all $\gamma \in F^{-1}$. Write $B_\epsilon(s)$ for the ball of radius ϵ around s with respect to d. We have $\gamma^a x \in B_\epsilon(\gamma^a s)$ or equivalently $x \in \left(\gamma^{-1}\right)^a B_\epsilon(\gamma^a s)$ for all $\gamma \in F^{-1}$. Therefore $x \in \bigcap_{\gamma \in F^{-1}} \left(\gamma^{-1}\right)^a B_\epsilon(\gamma^a s)$ and so $\bigcup_{s \in S} \bigcap_{\gamma \in F^{-1}} \left(\gamma^{-1}\right)^a B_\epsilon(\gamma^a s)$ is an open cover of X. Now, for every $s \in S$ and $\gamma \in F^{-1}$ we have that $B_\epsilon(\gamma^a s)$ is contained in some element of $\mathcal{U}$ and hence $\bigcap_{\gamma \in F^{-1}} \left(\gamma^{-1}\right)^a B_\epsilon(\gamma^a s)$ is contained in an element of $\mathcal{U}^F$. It follows that

$$N\left(\mathcal{U}^F\right) \leq |S| = \operatorname{Span}\left(X, \epsilon, d_{F^{-1}}\right). \tag{5.2}$$

If $\mathcal{V}$ is an open cover of X, let $\operatorname{diam}(\mathcal{V})$ denote the supremum of the diameters of elements of $\mathcal{V}$. Let $\mathcal{V}$ be an open cover of X with $\operatorname{diam}(\mathcal{V}) \leq \epsilon$. Let R be an ϵ-separated set of maximal cardinality with respect to d_F. An element of $\mathcal{V}^F$ contains at most one point of R, and hence

$$\operatorname{Sep}\left(X, \epsilon, d_F\right) \leq N\left(\mathcal{V}^F\right). \tag{5.3}$$

By (5.1), (5.2) and (5.3) if $\mathcal{U}$ has Lebesgue number ϵ and $\operatorname{diam}(\mathcal{V}) \leq \epsilon$ we have for all finite $F \subseteq \Gamma$:

$$\begin{aligned}
h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{U}) &= \inf_F \frac{1}{|F|} \log\left(N\left(\mathcal{U}^F\right)\right) \\
&\leq \inf_F \frac{1}{|F|} \log\left(\operatorname{Span}\left(X, \epsilon, d_F\right)\right) \\
&\leq \inf_F \frac{1}{|F|} \log\left(\operatorname{Sep}\left(X, \epsilon, d_F\right)\right) \\
&\leq \inf_F \frac{1}{|F|} \log\left(N\left(\mathcal{V}^F\right)\right) \\
&= h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{V}) \\
&\leq h^{\mathrm{tp}}_{\mathrm{nv}}\left(\Gamma \curvearrowright^a X\right).
\end{aligned} \tag{5.4}$$

Assume $h^{\mathrm{tp}}_{\mathrm{nv}}\left(\Gamma \curvearrowright^a X\right) < \infty$. Given $\kappa > 0$ find an open cover $\mathcal{U}$ so that $h^{\mathrm{tp}}_{\mathrm{nv}}\left(\Gamma \curvearrowright^a X\right) - \kappa \leq h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{U})$. Then if ϵ is less than the Lebesgue number of $\mathcal{U}$, (5.4) implies that

$$\begin{aligned}
h^{\mathrm{tp}}_{\mathrm{nv}}\left(\Gamma \curvearrowright^a X\right) - \kappa &\leq \inf_F \frac{1}{|F|} \log\left(\operatorname{Span}\left(X, \epsilon, d_F\right)\right) \\
&\leq \inf_F \frac{1}{|F|} \log\left(\operatorname{Sep}\left(X, \epsilon, d_F\right)\right) \\
&\leq h^{\mathrm{tp}}_{\mathrm{nv}}\left(\Gamma \curvearrowright^a X\right).
\end{aligned}$$

Assume $h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright^a X) = \infty$. Given $r \in \mathbb{R}$ find an open cover $\mathcal{U}$ so that $r \leq h^{\mathrm{tp}}_{\mathrm{nv}}(\mathcal{U})$. Then if ϵ is less than the Lebesgue number of $\mathcal{U}$, we have again by (5.4) that

$$r \leq \inf_F \frac{1}{|F|} \log\left(\mathrm{Span}\left(X, \epsilon, d_F\right)\right) \leq \inf_F \frac{1}{|F|} \log\left(\mathrm{Sep}\left(X, \epsilon, d_F\right)\right).$$

□

In particular we see from Proposition 5.2.2 that the quantities

$$\sup_{\epsilon>0} \inf_F \frac{1}{|F|} \log(\mathrm{Sep}(X, \epsilon, d_F))$$

and

$$\sup_{\epsilon>0} \inf_F \frac{1}{|F|} \log(\mathrm{Span}(X, \epsilon, d_F))$$

are independent of the choice of compatible metric d.

Proof of Theorem 5.1.3.

Recall that if $\alpha = (A_1, \ldots, A_k)$ and $\beta = (B_1, \ldots, B_m)$ are finite partitions of (X, μ), the conditional Shannon entropy $H(\alpha|\beta)$ of α given β is defined by

$$H(\alpha|\beta) = -\sum_{i=1}^{k} \sum_{j=1}^{m} \mu(A_i \cap B_j) \log\left(\frac{\mu(A_i \cap B_j)}{\mu(B_j)}\right).$$

We will use the following well-known facts about Shannon entropy, which appear in [42] as Propositions 14.16, 14.18.2 and 14.18.4 respectively.

Proposition 5.2.3. **(1)** *$H(\alpha_1 \vee \alpha_2) = H(\alpha_1) + H(\alpha_2|\alpha_1)$, in particular $H(\alpha_1 \vee \alpha_2) \geq H(\alpha_1)$,*

(2) *If β_2 refines β_1 then $H(\alpha|\beta_2) \leq H(\alpha|\beta_1)$,*

(3) *$H(\alpha_1 \vee \alpha_2|\beta) \leq H(\alpha_1|\beta) + H(\alpha_2|\beta)$.*

The following argument is a straightforward generalization of the corresponding proof for $\mathbb{Z}$-systems given as Part I of Theorem 17.1 in [42].

Proof of Theorem 5.1.3. Let μ be an invariant measure for the topological Γ-system $\Gamma \curvearrowright^a X$. Let $\alpha = (A_i)_{i=1}^k$ be a measurable partition of (X, μ). Choose closed sets $B_i \subseteq A_i$ such that $\mu(A_i \triangle B_i)$ is small enough so $H(\alpha|\beta) \leq 1$, where β is the partition

$(B_i)_{i=1}^{k+1}$ and $B_{k+1} = X - \bigcup_{i=1}^{k} B_i$. Then for any finite set $F \subseteq \Gamma$ by (2) and (3) of Proposition 5.2.3 we have

$$\begin{aligned} H_\mu\left(\alpha^F \middle| \beta^F\right) &\leq \sum_{\gamma \in F} H_\mu\left(\gamma^a \alpha \middle| \beta^F\right) \\ &\leq \sum_{\gamma \in F} H_\mu(\gamma^a \alpha | \gamma^a \beta) \\ &= |F| \cdot H_\mu(\alpha|\beta) \\ &\leq |F|. \end{aligned}$$

Hence by (1) of Proposition 5.2.3 we have

$$\begin{aligned} H_\mu\left(\alpha^F\right) &\leq H_\mu\left(\alpha^F \vee \beta^F\right) \\ &= H_\mu\left(\beta^F\right) + H_\mu\left(\alpha^F \middle| \beta^F\right) \\ &\leq H_\mu\left(\beta^F\right) + |F| \end{aligned}$$

and consequently

$$\begin{aligned} h_\mu(\alpha) &= \inf_F \frac{1}{|F|} H_\mu\left(\alpha^F\right) \\ &\leq \inf_F \frac{1}{|F|}\left(H_\mu\left(\beta^F\right) + |F|\right) \\ &= h_\mu(\beta) + 1. \end{aligned} \tag{5.5}$$

Now let $U_i = B_i \cup B_{k+1}$. Then $X - U_i = \bigcup_{\substack{1 \leq j \leq k, \\ j \neq i}} B_j$ so U_i is open and $\mathcal{U} = (U_i)_{i=1}^k$ is an open cover of X. Note that the only elements of β meeting U_i are B_i and B_{k+1}. Let $\mathcal{V}(F)$ be an open subcover of $\mathcal{U}^F$ with minimal cardinality. We claim that each element of $\mathcal{V}(F)$ meets at most $2^{|F|}$ elements of β^F. Indeed suppose $\phi : F \to [k]$ is a function such that $\bigcap_{\gamma \in F} \gamma^a U_{\phi(\gamma)} \in \mathcal{V}(F)$ and let $x \in \bigcap_{\gamma \in F} \gamma^a U_{\phi(\gamma)}$. Then if $\psi : F \to [k+1]$ is any function so that $x \in \bigcap_{\gamma \in F} \gamma^a B_{\psi(\gamma)} \in \beta^F$ we must have $B_{\psi(\gamma)} \cap U_{\phi(\gamma)} \neq \emptyset$ and hence $\psi(\gamma) \in \{\phi(\gamma), k+1\}$ for all $\gamma \in F$. Therefore

$$\left|\beta^F\right| \leq 2^{|F|} \left|\mathcal{V}(F)\right|.$$

It follows that

$$\begin{aligned} H_\mu\left(\beta^F\right) &\leq \log\left(\left|\beta^F\right|\right) \\ &\leq \log\left(2^{|F|} \cdot |\mathcal{V}(F)|\right) \\ &\leq |F| \log 2 + \log\left(|\mathcal{V}(F)|\right) \\ &= |F| \log 2 + \log\left(N\left(\mathcal{U}^F\right)\right) \end{aligned} \tag{5.6}$$

and hence by (5.5) and (5.6) we have

$$\begin{aligned}
h_\mu(\alpha) &\leq h_\mu(\beta) + 1 \\
&= \left(\inf_F \frac{1}{|F|} H_\mu\left(\beta^F\right)\right) + 1 \\
&\leq \left(\inf_F \frac{1}{|F|} \left(|F| \log 2 + \log\left(N\left(\mathcal{U}^F\right)\right)\right)\right) + 1 \\
&= h_{\mathrm{nv}}^{\mathrm{tp}}(\mathcal{U}) + 1 + \log 2.
\end{aligned}$$

Therefore

$$h_{\mathrm{nv}}\left(\Gamma \curvearrowright (X,\mu)\right) \leq h_{\mathrm{nv}}^{\mathrm{tp}}\left(\Gamma \curvearrowright X\right) + 1 + \log 2.$$

Now observe that the measure μ^n on X^n is invariant for the n^{th} Cartesian power of the system $\Gamma \curvearrowright X$. Therefore the same argument shows

$$h_{\mathrm{nv}}\left(\Gamma \curvearrowright (X^n,\mu^n)\right) \leq h_{\mathrm{nv}}^{\mathrm{top}}\left(\Gamma \curvearrowright X^n\right) + 1 + \log 2. \tag{5.7}$$

Immediate generalizations of the proofs of Theorems 14.14 and 14.31 in [42] show that both forms of naive entropy are additive under direct products. Thus (5.7) implies

$$n \cdot h_{\mathrm{nv}}\left(\Gamma \curvearrowright (X,\mu)\right) \leq n \cdot h_{\mathrm{nv}}^{\mathrm{top}}\left(\Gamma \curvearrowright X\right) + 1 + \log 2$$

for all $n \geq 1$ and therefore we must have

$$h_{\mathrm{nv}}\left(\Gamma \curvearrowright (X,\mu)\right) \leq h_{\mathrm{nv}}^{\mathrm{top}}\left(\Gamma \curvearrowright X\right).$$

□

Examples.

Example 5.2.1. *Let (Y,ν) be a standard probability space. Assume ν is not supported on a single point. Consider the Bernoulli shift $\Gamma \curvearrowright (X,\mu)$ where $X = Y^\Gamma$ and $\mu = \nu^\Gamma$. Let $\alpha = (A_1, A_2)$ be a partition of (Y,ν) with positive entropy and $\hat{\alpha} = \left(\hat{A}_1, \hat{A}_2\right)$ be the partition of (X,μ) given by*

$$\hat{A}_i = \{\omega \in X : \omega\left(e_\Gamma\right) \in A_i\},$$

where e_Γ is the identity of Γ. Then as in the case of a $\mathbb{Z}$-system distinct shifts of $\hat{\alpha}$ are independent and so we have $H_\mu\left(\hat{\alpha}^F\right) = |F| \cdot H_\mu\left(\hat{\alpha}\right)$. Thus

$$h_\mu\left(\hat{\alpha}\right) = H_\mu\left(\hat{\alpha}\right) = H_\nu(\alpha) > 0.$$

By Theorem 5.2.1 we see that if Γ *is nonamenable then* $h_{\mathrm{nv}}(\Gamma \curvearrowright (X,\mu)) = \infty$*. Thus Theorem 5.1.3 implies that the corresponding topological system* $\Gamma \curvearrowright X$ *has infinite naive entropy.*

Example 5.2.2. *Let* $\Gamma \curvearrowright^a X$ *be a topological system and* d *a compatible metric on* X*. Recall that* $\Gamma \curvearrowright^a X$ *is said to be distal if for every pair* x_1, x_2 *of distinct points in* X *we have* $\inf_{\gamma\in\Gamma} d(\gamma^a x_1, \gamma^a x_2) > 0$*. In particular, an isometric system such as a circle rotation is distal.*

Now, suppose that $\Gamma \curvearrowright^a X$ *is distal and* Γ *has an element* γ *of infinite order. Then* (X, γ^a) *is a distal* $\mathbb{Z}$*-system. Theorem* 18.19 *in [42] implies that distal* $\mathbb{Z}$*-systems have zero entropy. Thus Proposition 5.2.1 guarantees that* $h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright^a X) = 0$*. By Theorem 5.1.3,* $h_{\mathrm{nv}}(\Gamma \curvearrowright^a (X,\mu)) = 0$ *for any invariant measure* μ*. It is likely that a distal* Γ*-system has zero naive topological entropy for an arbitrary* Γ*, but we were unable to prove this despite significant effort.*

Proof of Theorem 5.1.2

We first show three preliminary lemmas.

Lemma 5.2.1. *Let* $\mathcal{U}$ *be a finite open cover of a compact metrizable space* X*. Fix a finite set* $F \subseteq \Gamma$ *and* $k \in \mathbb{N}$*. Then*

$$Z(\mathcal{U}, F, k) = \left\{ (\Gamma \curvearrowright^a X) \in \mathrm{A_{top}}(\Gamma, X) : N\left(\bigvee_{\gamma\in F} \gamma^a \mathcal{U} \right) \le k \right\}$$

is open.

Proof. Write $\mathcal{U} = (U_i)_{i=1}^n$. Let $(\Gamma \curvearrowright^a X) \in Z(\mathcal{U}, F, k)$ and let $\mathcal{V}$ be a subcover of $\bigvee_{\gamma\in F} \gamma^a \mathcal{U}$ with cardinality $\le k$. Let d be a compatible metric on X and let d_u be the metric

$$d_u(f,g) = \sup_{x\in X} d(f(x), g(x)).$$

Note that to obtain the uniform topology on Homeo(X) we must use the metric

$$d'_u(f,g) = d_u(f,g) + d_n(f^{-1}, g^{-1}).$$

However the topology induced by d_u on $\mathrm{A_{top}}(\Gamma, X)$ is the same as the one induced by d'_u so we will continue to work with the former.

Let ϵ be a Lebesgue number for $\mathcal{V}$ with respect to d. Let $(\phi_j)_{j=1}^k$ be a sequence of functions from F to $[n]$ so that

$$\mathcal{V} = \left(\bigcap_{\gamma\in F} \gamma^a U_{\phi_j(\gamma)}\right)_{j=1}^k .$$

Let $\delta > 0$ be small enough that for all $\gamma \in F$ and $x_1, x_2 \in X$, $d(x_1, x_2) < \delta$ implies $d(\gamma^a x_1, \gamma^a x_2) < \epsilon$. Then for any $x \in X$, $\left(\gamma^{-1}\right)^a B_\epsilon(x)$ contains $B_\delta\left(\left(\gamma^{-1}\right)^a x\right)$. Suppose $d_u\left(\left(\gamma^{-1}\right)^a, \left(\gamma^{-1}\right)^b\right) < \delta$ for all $\gamma \in F$. We claim

$$\left(\bigcap_{\gamma\in F} \gamma^b U_{\phi_j(\gamma)}\right)_{j=1}^k$$

is a cover of X. Let $x \in X$. Then there is $j \le k$ so that $B_\epsilon(x) \subseteq \bigcap_{\gamma\in F} \gamma^a U_{\phi_j(\gamma)}$, equivalently $\left(\gamma^{-1}\right)^a B_\epsilon(x) \subseteq U_{\phi_j(\gamma)}$ for all $\gamma \in F$. Since $d\left(\left(\gamma^{-1}\right)^a x, \left(\gamma^{-1}\right)^b x\right) < \delta$, we see that $\left(\gamma^{-1}\right)^b x \in U_{\phi_j(\gamma)}$. Therefore $x \in \gamma^b U_{\phi_j(\gamma)}$ for all $\gamma \in F$. □

Lemma 5.2.2. *For any system $\Gamma \curvearrowright X$, if $(\mathcal{U}_n)_{n=1}^\infty$ is a sequence of finite open covers such that* $\lim_{n\to\infty} \operatorname{diam}(\mathcal{U}_n) = 0$*, then* $\lim_{n\to\infty} h^{\mathrm{tp}}(\mathcal{U}_n) = h^{\mathrm{tp}}_{\mathrm{nv}}(\Gamma \curvearrowright X)$*.*

Proof. It is clear that if $\mathcal{U}$ refines $\mathcal{V}$ then $h^{\mathrm{tp}}(\mathcal{V}) \le h^{\mathrm{tp}}(\mathcal{U})$. Thus if $\mathcal{V}$ is an arbitrary open cover of X, by choosing n so that $\operatorname{diam}(\mathcal{U}_n)$ is less than the Lebegsue number of $\mathcal{V}$ we have $h^{\mathrm{tp}}(\mathcal{V}) \le h^{\mathrm{tp}}(\mathcal{U}_n)$. □

Lemma 5.2.3. *For any countable group Γ and compact metrizable space X, the set of systems with zero naive topological entropy is G_δ in* $\mathrm{A}_{\mathrm{top}}(\Gamma, X)$*.*

Proof. If $\mathcal{U}$ is an open cover of X, $F \subseteq \Gamma$ is finite and $\epsilon > 0$ set

$$\tilde{Z}(\mathcal{U}, F, \epsilon) = \left\{(\Gamma \curvearrowright^a X) \in \mathrm{A}_{\mathrm{top}}(\Gamma, X) : \frac{1}{|F|} \log\left(N\left(\bigvee_{\gamma\in F} \gamma^a \mathcal{U}\right)\right) < \epsilon\right\}.$$

Note that in the notation of Lemma 5.2.1, we have

$$\tilde{Z}(\mathcal{U}, F, \epsilon) = Z\left(\mathcal{U}, F, \lfloor \exp(\epsilon|F|)\rfloor\right),$$

and hence $\tilde{Z}(\mathcal{U}, F, \epsilon)$ is open. If $(\mathcal{U}_n)_{n=1}^\infty$ is a sequence of finite open covers with $\lim_{n\to\infty} \operatorname{diam}(\mathcal{U}_n) = 0$ then by Lemma 5.2.2, the set of systems with zero naive topological entropy is equal to the G_δ set

$$\bigcap_{n=1}^\infty \bigcap_{k=1}^\infty \bigcup_F \tilde{Z}\left(\mathcal{U}_n, \frac{1}{k}, F\right),$$

where the union is over all nonempty finite subsets of Γ. □

Proof of Theorem 5.1.2. By Lemma 5.2.3, it suffices to show the set of systems with zero entropy is dense in $\mathrm{A}_{\mathrm{top}}\left(\Gamma, 2^{\mathbb{N}}\right)$. By Corollary 2.5 in [41], the set of homeomorphisms with zero entropy is uniformly dense in Homeo $\left(2^{\mathbb{N}}\right)$. Therefore the set of systems in $\mathrm{A}_{\mathrm{top}}\left(\Gamma, 2^{\mathbb{N}}\right)$ for which the first generator of Γ acts with zero entropy is dense. The theorem follows from this fact and Proposition 5.2.1. □

5.3 Sofic groups and sofic entropy.

Sofic groups.

Sofic groups were introduced by Gromov in [44] and Weiss in [75]. Let Sym(n) denote the symmetric group on n letters. Let u_n denote the uniform probability measure on $[n]$ so that $u_n(A) = \frac{|A|}{n}$. In keeping with our convention for dynamical systems, if σ is a function from Γ to Sym(n) we write $\gamma^{\sigma} m$ for $\sigma(\gamma)(m)$.

Definition 5.3.1. *Let Γ be a countable discrete group. Let $\Sigma = (\sigma_i)_{i=1}^{\infty}$ be a sequence of functions $\sigma_i : \Gamma \to \mathrm{Sym}(n_i)$ such that $n_i \to \infty$. Note that the σ_i are not assumed to be homomorphisms. We say Σ is a* ***sofic approximation*** *to Γ if for every pair $\gamma_1, \gamma_2 \in \Gamma$ we have*

$$\lim_{i\to\infty} u_{n_i}(\{m \in [n_i] : (\gamma_1\gamma_2)^{\sigma_i} m = \gamma_1^{\sigma_i}\gamma_2^{\sigma_i} m\}) = 1,$$

and for every pair $\gamma_1 \neq \gamma_2$ we have

$$\lim_{i\to\infty} u_{n_i}(\{m \in [n_i] : \gamma_1^{\sigma_i} m \neq \gamma_2^{\sigma_i} m\}) = 1.$$

We say Γ is ***sofic*** *if there exists a sofic approximation to Γ.*

Thus the first condition guarantees that the σ_i are asymptotically homomorphisms, and the second condition guarantees that the corresponding approximate actions on $[n_i]$ are asymptotically free. The standard examples of sofic groups are residually finite groups and amenable groups. It is unknown whether every countable group is sofic.

Topological sofic entropy.

In [57] and [62], Kerr and Li developed a topological counterpart to Bowen's theory of sofic entropy, based initially on operator-algebraic considerations. We will use the 'spatial' formulation of these ideas. Fix a group Γ and a topological Γ-system $\Gamma \curvearrowright^{a} X$. Fix a compatible metric d for X. Define the metrics d^2 and d^{∞} on the set of maps from $[n]$ to X by

$$d^2(\phi, \psi) = \left(\frac{1}{n}\sum_{m=1}^{n} d\left(\phi(m), \psi(m)\right)^2\right)^{\frac{1}{2}}$$

and

$$d^\infty(\phi, \psi) = \max_{m \in [n]} d(\phi(m), \psi(m)).$$

Definition 5.3.2. *Let $F \subseteq \Gamma$ be finite, $\delta > 0$ and $\sigma : \Gamma \to \mathrm{Sym}(n)$. Define* $\mathrm{Map}(\sigma, F, \delta)$ *to be the collection of functions $\phi : [n] \to X$ such that $d^2(\phi \circ \gamma^\sigma, \gamma^a \circ \phi) \leq \delta$ for all $\gamma \in F$.*

Definition 5.3.3. *Let $\Sigma = (\sigma_i)_{i=1}^\infty$ be a sofic approximation to Γ with $\sigma_i \in \mathrm{Sym}(n_i)^\Gamma$. Define the* ***topological sofic entropy*** *$h_\Sigma^{\mathrm{tp}}(\Gamma \curvearrowright^a X)$ of $\Gamma \curvearrowright^a X$ with respect to Σ as follows. Letting F range over the nonempty finite subsets of Γ and $\delta, \epsilon > 0$ define*

$$\begin{aligned}
h_\Sigma^{\mathrm{tp}}(\delta, F, \epsilon) &= \limsup_{i \to \infty} \frac{1}{n_i} \log(\mathrm{Sep}(\mathrm{Map}(\sigma_i, F, \delta), \epsilon, d^\infty)),\\
h_\Sigma^{\mathrm{tp}}(F, \epsilon) &= \inf_{\delta > 0} h_\Sigma^{\mathrm{tp}}(\delta, F, \epsilon),\\
h_\Sigma^{\mathrm{tp}}(\epsilon) &= \inf_{F} h_\Sigma^{\mathrm{tp}}(F, \epsilon),\\
h_\Sigma^{\mathrm{tp}}(\Gamma \curvearrowright^a X) &= \sup_{\epsilon > 0} h_\Sigma^{\mathrm{tp}}(\epsilon).
\end{aligned}$$

5.4 Proof of Theorem 5.1.1

This argument builds on the framework used to prove Lemma 5.1 in [62].

Choosing parameters

In this subsection we set the values of some initial parameters for our construction. Let $\Sigma = (\sigma_n)_{n=1}^\infty$ be a sofic approximation to Γ, where $\sigma_n : \Gamma \to \mathrm{Sym}(n)$. The case where σ_n is a function from Γ to $[k_n]$ for some $k_n \neq n$ can be handled with trivial modifications. Choose κ with $0 < \kappa < 1$. It suffices to show that $h_\Sigma^{\mathrm{tp}}(\Gamma \curvearrowright^a X) \leq \kappa$. Choose $\epsilon > 0$, so that it suffices to show that $h_\Sigma^{\mathrm{tp}}(\epsilon) \leq \kappa$. Let

$$\eta = \frac{\kappa}{4 \log \left(\mathrm{Sep}\left(X, \frac{\epsilon}{4}, d\right)\right)} \tag{5.8}$$

and choose $k \in \mathbb{N}$ such that

$$\frac{1}{k} \leq \frac{\eta}{2}. \tag{5.9}$$

By our assumption that $h_{\mathrm{nv}}^{\mathrm{tp}}(\Gamma \curvearrowright^a X) = 0$, we can choose a finite set $F \subseteq \Gamma$ such that

$$\frac{1}{|F|} \log \left(\mathrm{Sep}\left(X, \frac{\epsilon}{4}, d_F\right)\right) \leq \frac{\kappa}{4k}. \tag{5.10}$$

Lemma 5.4.1. *Let $F' \subseteq F$ be such that $|F'| \geq \frac{|F|}{k}$. Then*

$$\mathrm{Sep}\left(X, \frac{\epsilon}{4}, d_{F'}\right) \leq \exp\left(\frac{\kappa |F'|}{4}\right).$$

Proof of Lemma 5.4.1. Since

$$\operatorname{Sep}\left(X, \frac{\epsilon}{4}, d_{F'}\right) \leq \operatorname{Sep}\left(X, \frac{\epsilon}{4}, d_F\right),$$

we have

$$\begin{aligned}
\frac{1}{|F'|} \log\left(\operatorname{Sep}\left(X, \frac{\epsilon}{4}, d_{F'}\right)\right) &\leq \frac{1}{|F'|} \log\left(\operatorname{Sep}\left(X, \frac{\epsilon}{4}, d_F\right)\right) \\
&\leq k\left(\frac{1}{|F|} \log\left(\operatorname{Sep}\left(X, \frac{\epsilon}{4}, d_F\right)\right)\right) \\
&\leq \frac{\kappa}{4},
\end{aligned}$$

where the last inequality follows from (5.10). □

Write $s = |F|$. Let $\delta > 0$ be small enough that

$$\delta \leq \left(\frac{\epsilon}{8}\right)^2, \tag{5.11}$$

$$\delta \leq \frac{\eta}{4s^3} \tag{5.12}$$

(so in particular $s\delta < 1$) and finally

$$-(s\delta \log(s\delta) + (1 - s\delta)\log(1 - s\delta)) \leq \frac{\kappa}{4}. \tag{5.13}$$

For a finite $S \subseteq \Gamma$ let

$$\begin{aligned}
Q(S)_n = \{m \in [n] : (\gamma_1\gamma_2)^{\sigma_n} m = \gamma_1^{\sigma_n}\gamma_2^{\sigma_n} m \text{ for all } \gamma_1, \gamma_2 \in S\} \\
\cap \{m \in [n] : \gamma_1^{\sigma_n} m \neq \gamma_2^{\sigma_n} m \text{ for all } \gamma_1 \neq \gamma_2 \in S\}
\end{aligned}$$

Write $\hat{F}$ for the symmetrization of F. Since Σ is a sofic approximation, we can find N so that if $n \geq N$ then

$$\left|Q(\hat{F})_n\right| \geq \left(1 - \frac{\eta}{4s^2}\right) n. \tag{5.14}$$

Choosing a separated subset

In this subsection we find a large ϵ-separated subset V of $\operatorname{Map}(\sigma, F, \delta)$ such that every element of V is approximately equivariant on a fixed large subset of $[n]$. Fix $n \geq N$ and write $\sigma = \sigma_n$. Let D be an ϵ-separated subset of $\operatorname{Map}(\sigma, F, \delta)$ with respect to d^∞ of maximal cardinality. For every $\phi \in \operatorname{Map}(\sigma, F, \delta)$ by definition we have $d^2(\phi \circ \gamma^\sigma, \gamma^a \circ \phi) \leq \delta$ for all $\gamma \in F$. Explicitly,

$$\left(\frac{1}{n}\sum_{m=1}^{n} d\big(\phi(\gamma^\sigma m), \gamma^a \phi(m)\big)^2\right)^{\frac{1}{2}} \leq \delta.$$

Hence for each fixed $\gamma \in F$ at least $(1-\delta)n$ elements m of $[n]$ have

$$d\left(\phi\left(\gamma^{\sigma} m\right), \gamma^{a} \phi(m)\right) \leq \sqrt{\delta}.$$

Hence the set Θ_{ϕ} of all $m \in [n]$ such that

$$d\left(\phi\left(\gamma^{\sigma} m\right), \gamma^{a} \phi(m)\right) \leq \sqrt{\delta}$$

for all $\gamma \in F$ has size at least $(1-s\delta)n$.

By a standard estimate from information theory (see for example Lemma 16.19 in [30]) the number of subsets of $[n]$ of size at most $s\delta n$ is at most

$$\exp\bigl(-n(s\delta \log(s\delta) + (1-s\delta)\log(1-s\delta))\bigr)$$

and by (5.13) this is bounded above by $\exp\left(\frac{\kappa n}{4}\right)$. Hence there at at most $\exp\left(\frac{\kappa n}{4}\right)$ possible choices for the sets $\{\Theta_{\phi} : \phi \in D\}$ and thus there are at least $\exp\left(-\frac{\kappa n}{4}\right)|D|$ elements of D for which Θ_{ϕ} is the same. So we can find $V \subseteq D$ and $\Theta \subseteq [n]$ such that

$$|D| \leq \exp\left(\frac{\kappa n}{4}\right)|V| \tag{5.15}$$

and for all $\phi \in V$ we have $\Theta_{\phi} = \Theta$. Note that since $|\Theta| \geq (1-s\delta)n$, (5.12) implies that

$$|[n] - \Theta| \leq \frac{\eta n}{4s^2}. \tag{5.16}$$

Furthermore, by (5.11) and the definition of Θ, for all $\phi \in V$ and all $m \in \Theta$ we have

$$d\left(\phi\left(\gamma^{\sigma} m\right), \gamma^{a} \phi(m)\right) \leq \frac{\epsilon}{8}. \tag{5.17}$$

Disjoint subsets of the sofic graph

Endow $[n]$ with the structure of the graph G_{σ} corresponding to σ, where m_1 is connected to m_2 if and only if there is $\gamma \in F$ such that $(\gamma)^{\sigma} m_1 = m_2$ or $\left(\gamma^{-1}\right)^{\sigma} m_1 = m_2$. In this section we find a maximal collection of disjoint subsets of G_{σ} which resemble a nontrivial part of F.

By (5.14) and (5.16),

$$|G_{\sigma} - (Q(\hat{F})_n \cap \Theta)| \leq \frac{\eta n}{2s^2}.$$

Let J be the collection of points c in G_{σ} such that the ball of radius 1 around c in G_{σ} is contained in $Q(\hat{F})_n \cap \Theta$, and let I be the collection of points c in J such that the ball of radius 1 around c is contained in J. Then

$$|G_{\sigma} - J| \leq s \cdot |G_{\sigma} - (Q(\hat{F})_n \cap \Theta)| \leq \frac{\eta n}{2s}$$

and

$$|G_\sigma - I| \leq s \cdot |G_\sigma - J| \leq \frac{\eta n}{2}. \tag{5.18}$$

If $c \in J$ then the mapping from F to G_σ given by $\gamma \mapsto \gamma^\sigma c$ is injective. We now begin an inductive procedure. Choose $c_1 \in J$ and take $F_1 = F$. Suppose we have chosen $c_1, \ldots, c_j \in J$ and $F_1, \ldots, F_j \subseteq F$ such that the sets $\left(F_i^\sigma c_i\right)_{i=1}^j$ are pairwise disjoint and $\frac{|F|}{k} \leq |F_i|$ for all $i \in \{1, \ldots, j\}$. Write $F_i^\sigma c_i = B_i$

Assume we cannot extend this process further, so that there do not exist c_{j+1} and F_{j+1} satisfying the two conditions. Write $W = \bigcup_{i=1}^j B_i$. Our assumption implies that for every $c \in J$, at least $\left(1 - \frac{1}{k}\right)|F|$ of the points in $F^\sigma c$ lie in W. Suppose toward a contradiction that $\frac{|J|}{k} < |I - W|$. For each point b in I, there are exactly $|F|$ points $c \in J$ such that $b \in F^\sigma c$, in symbols $|\{c \in J : b \in F^\sigma c\}| = |F|$. Indeed $b \in F^\sigma c$ if and only if $b = \gamma^\sigma c$ for some $c \in F$. Since $b, c \in Q(F)_n$, this is equivalent to $\left(\gamma^{-1}\right)^\sigma b = c$. Since $b \in Q(F^{-1})_n$, the map $\gamma^{-1} \mapsto \left(\gamma^{-1}\right)^\sigma b$ is injective. Therefore

$$\begin{aligned}
|\{c \in J : b \in F^\sigma c\}| &= \left|\left\{c \in J : c \in \left(F^{-1}\right)^\sigma b\right\}\right| \\
&= |F^{-1}| \\
&= |F|.
\end{aligned}$$

So we have

$$\sum_{b \in I - W} |\{c \in J : b \in F^\sigma c\}| = |F| \cdot |I - W| > \frac{|F| \cdot |J|}{k}.$$

We can write

$$\sum_{b \in I - W} |\{c \in J : b \in F^\sigma c\}| = \sum_{b \in I - W} \sum_{c \in J} \mathbf{1}_{F^\sigma c}(b),$$

where $\mathbf{1}_Y$ is the characteristic function of Y. So we have

$$\sum_{c \in J} \sum_{b \in I - W} \mathbf{1}_{F^\sigma c}(b) > \frac{|F| \cdot |J|}{k}.$$

Since there are $|J|$ terms in the outer sum, there must be some $c_0 \in J$ with

$$\sum_{b \in I - W} \mathbf{1}_{F^\sigma c_0}(b) > \frac{|F|}{k},$$

or equivalently $|(I - W) \cap F^\sigma c_0| > \frac{|F|}{k}$. Thus $|W \cap F^\sigma c_0| < \left(1 - \frac{1}{k}\right)|F|$, which contradicts our assumption. It follows that for a maximal pair of sequences $(c_i)_{i=1}^j$ and $(F_i)_{i=1}^j$ satisfying the relevant conditions, we have

$$|I - W| \leq \frac{|J|}{k}. \tag{5.19}$$

Fix such a maximal pair $(c_i)_{i=1}^j$ and $(F_i)_{i=1}^j$. Note that by our choice of k in (5.9) we have

$$\frac{|J|}{k} \leq \frac{n}{k} \leq \frac{\eta n}{2}. \tag{5.20}$$

Therefore if we put $P = G_\sigma - W$ then by (5.18), (5.19) and (5.20) we have

$$\begin{aligned} |P| &\leq |G_\sigma - I| + |I - W| \\ &\leq \frac{\eta n}{2} + \frac{\eta n}{2} = \eta n. \end{aligned} \tag{5.21}$$

Controlling sofic entropy by naive entropy

In this subsection we use the data previously constructed to bound the size of an appropriately separated subset of $\mathrm{Map}(\sigma, F, \delta)$ in terms of the separation numbers used to compute naive entropy. For $B \subseteq [n]$, let d_B^∞ be the pseudometric on the collection of maps from $[n]$ to X given by $d_B^\infty(\phi, \psi) = \max_{m \in B} d(\phi(m), \psi(m))$. Let $i \leq j$ and take an $\frac{\epsilon}{2}$-spanning set V_i of V of minimal cardinality with respect to the pseudometric $d_{B_i}^\infty$. We claim

$$|V_i| \leq \exp\left(\frac{\kappa |F_i|}{4}\right).$$

To see this, let U be a maximal $\frac{\epsilon}{2}$-separated subset of V with respect to $d_{B_i}^\infty$. Then U is also $\frac{\epsilon}{2}$-spanning with respect to $d_{B_i}^\infty$ and hence $|V_i| \leq |U|$. For any two elements ϕ and ψ of V we have $c_i \in J \subseteq \Theta = \Theta_\psi = \Theta_\phi$. Since $F_i \subseteq F$ it follows from (5.17) that $d\left(\gamma^a \phi(c_i), \phi\left(\gamma^\sigma c_i\right)\right) \leq \frac{\epsilon}{8}$ for all $\gamma \in F_i$, and similarly for ψ. So for all $\gamma \in F_i$ we have

$$\begin{aligned} d\left(\gamma^a \phi(c_i), \gamma^a \psi(c_i)\right) &\geq d\left(\phi\left(\gamma^\sigma c_i\right), \psi\left(\gamma^\sigma c_i\right)\right) \\ &\qquad - d\left(\gamma^a \phi(c_i), \phi\left(\gamma^\sigma c_i\right)\right) - d\left(\gamma^a \psi(c_i), \psi\left(\gamma^\sigma c_i\right)\right) \\ &\geq d\left(\phi\left(\gamma^\sigma c_i\right), \psi\left(\gamma^\sigma c_i\right)\right) - \frac{\epsilon}{4}. \end{aligned} \tag{5.22}$$

Now, since U is $\frac{\epsilon}{2}$-separated with respect to $d_{B_i}^\infty$, for any $\phi, \psi \in U$ we have

$$d_{B_i}^\infty(\phi, \psi) = \max_{b \in B_i} d(\phi(b), \psi(b)) = \max_{\gamma \in F_i} d\left(\phi\left(\gamma^\sigma c_i\right), \psi\left(\gamma^\sigma c_i\right)\right) \geq \frac{\epsilon}{2}. \tag{5.23}$$

By (5.22) and (5.23),

$$\begin{aligned} d_{F_i}(\phi(c_i),\psi(c_i)) &= \max_{\gamma\in F_i} d\left(\gamma^a\phi(c_i),\gamma^a\psi(c_i)\right) \\ &\geq \max_{\gamma\in F_i}\left(d\left(\phi\left(\gamma^\sigma c_i\right),\psi\left(\gamma^\sigma c_i\right)\right)-\frac{\epsilon}{4}\right) \\ &= \left(\max_{\gamma\in F_i} d\left(\phi\left(\gamma^\sigma c_i\right),\psi\left(\gamma^\sigma c_i\right)\right)\right)-\frac{\epsilon}{4} \\ &\geq \frac{\epsilon}{2}-\frac{\epsilon}{4}=\frac{\epsilon}{4}. \end{aligned}$$

It follows that $\{\phi(c_i) : \phi \in U\}$ is an $\frac{\epsilon}{4}$-separated subset of X with respect to d_{F_i} of size $|U|$ and hence by Lemma 5.4.1 we have

$$|U| \leq \mathrm{Sep}\left(X,\frac{\epsilon}{4},d_{F_i}\right) \leq \exp\left(\frac{\kappa|F_i|}{4}\right),$$

and consequently

$$|V_i| \leq \exp\left(\frac{\kappa|F_i|}{4}\right). \tag{5.24}$$

Now, take an $\frac{\epsilon}{2}$-spanning subset V_P of V of minimal cardinality with respect to d_P^∞. Since a maximal $\frac{\epsilon}{2}$-separated subset is also $\frac{\epsilon}{2}$-spanning, we have

$$|V_P| \leq \mathrm{Sep}\left(V,\frac{\epsilon}{2},d_P^\infty\right). \tag{5.25}$$

For a compact pseudometric space (Z,ρ) and $r > 0$ write $\mathrm{Cov}(Z,r,\rho)$ for the minimal cardinality of a family of ρ-balls of radius r which covers Z. It is easy to see that for any r we have

$$\mathrm{Cov}(Z,r,\rho) \leq \mathrm{Sep}(Z,r,\rho) \leq \mathrm{Cov}\left(Z,\frac{r}{2},\rho\right).$$

Now, let $\{B_1,\dots,B_j\}$ be a cover of X by balls of radius $\frac{\epsilon}{4}$. We can construct a cover of $X^{[n]}$ by considering the collection of all sets of the form $\prod_{p=1}^n Y_p$ where Y_p is equal to some B_i if $p\in P$ and equal to X if $p\notin P$. Each of these sets is a d_P^∞-ball of radius $\frac{\epsilon}{4}$ and so we see that

$$\begin{aligned} \mathrm{Sep}\left(V,\frac{\epsilon}{2},d_P^\infty\right) &\leq \mathrm{Cov}\left(V,\frac{\epsilon}{4},d_P^\infty\right) \leq \mathrm{Cov}\left(X^{[n]},\frac{\epsilon}{4},d_P^\infty\right) \\ &\leq \mathrm{Cov}\left(X,\frac{\epsilon}{4},d\right)^{|P|} \leq \mathrm{Sep}\left(X,\frac{\epsilon}{4},d\right)^{|P|}. \end{aligned} \tag{5.26}$$

(5.21), (5.25), and (5.26) imply

$$|V_P| \leq \mathrm{Sep}\left(X,\frac{\epsilon}{4},d\right)^{\eta n}$$

and hence

$$|V_P| \leq \exp\left(\frac{\kappa n}{4}\right) \tag{5.27}$$

by our choice of η in (5.8).

Conclusion

Let Z be the set of all maps $\phi : [n] \to X$ such that $\phi \restriction P = \psi \restriction P$ for some $\psi \in V_P$ and for each $i \leq j$ we have $\phi \restriction B_i = \psi_i \restriction B_i$ for some $\psi_i \in V_i$. Note that since we chose the sets $B_i = F_i^\sigma c_i$ to be pairwise disjoint, and the maps $\gamma \mapsto \gamma^\sigma c_i$ for $\gamma \in F_i$ are bijective, we have $\sum_{i=1}^{j} |F_i| \leq n$. Thus by (5.24) and (5.27) we have

$$\begin{aligned}
|Z| &\leq |V_P| \left(\prod_{i=1}^{j} |V_i|\right) \\
&\leq \exp\left(\frac{\kappa n}{4}\right) \left(\prod_{i=1}^{j} \exp\left(\frac{\kappa |F_i|}{4}\right)\right) \\
&= \exp\left(\frac{\kappa n}{4} + \frac{\kappa}{4}\left(\sum_{i=1}^{j} |F_i|\right)\right) \\
&\leq \exp\left(\frac{\kappa n}{2}\right). \qquad (5.28)
\end{aligned}$$

Note that if $\phi \in V$, then by the hypothesis that V_i is $\frac{\epsilon}{2}$-spanning for V with respect to the metric $d^\infty_{B_i}$ we have that $\max_{b \in B_i} d(\phi(b), \psi_i(b)) \leq \frac{\epsilon}{2}$ for some element ψ_i of V_i, and similarly for P and V_P. Hence every element of V is within d^∞ distance $\frac{\epsilon}{2}$ of some element of Z. Define a map $f : V \to Z$ by letting $f(\phi)$ be any element of Z within d^∞ distance $\frac{\epsilon}{2}$ of ϕ. Since V is a subset of D and we assumed that D was ϵ-separated with respect to d^∞, it follows that f is injective. Therefore we have $|V| \leq |Z|$. Then it follows from (5.15) and (5.28) that if $n \geq N$ then

$$\begin{aligned}
\mathrm{Sep}(\mathrm{Map}(F, \delta, \sigma_n), \epsilon, d^\infty) &= |D| \\
&\leq \exp\left(\frac{\kappa n}{2}\right) |V| \\
&\leq \exp\left(\frac{\kappa n}{2}\right) |Z| \\
&\leq \exp\left(\frac{\kappa n}{2}\right) \exp\left(\frac{\kappa n}{2}\right) \\
&= \exp(\kappa n).
\end{aligned}$$

This concludes the proof of Theorem 5.1.1.

Chapter 6

UNIFORM MIXING AND COMPLETELY POSITIVE SOFIC ENTROPY

Tim Austin and Peter Burton

6.1 Introduction

Let G be a countable discrete sofic group, (X, μ) a standard probability space and $T : G \curvearrowright X$ a measurable G-action preserving μ. In [14], Lewis Bowen defined the sofic entropy of (X, μ, T) relative to a sofic approximation under the hypothesis that the action admits a finite generating partition. The definition was extended to general (X, μ, T) by Kerr and Li in [61] and Kerr gave a more elementary approach in [58]. In [17] Bowen showed that when G is amenable, sofic entropy relative to any sofic approximation agrees with the standard Kolmogorov-Sinai entropy. Despite some notable successes such as the proof in [14] that Bernoulli shifts with distinct base-entropies are nonisomorphic, many aspects of the theory of sofic entropy are still relatively undeveloped.

Rather than work with abstract measure-preserving G-actions, we will use the formalism of G-processes. If G is a countable group and A is a standard Borel space, we will endow A^G with the right-shift action given by $(g \cdot a)(h) = a(hg)$ for $g, h \in G$ and $a \in A^G$. A G-process over A is a Borel probability measure μ on A^G which is invariant under this action. Any measure-preserving action of G on a standard probability space is measure-theoretically isomorphic to a G-process over some standard Borel space A. We will assume the state space A is finite, which corresponds to the case of measure-preserving actions which admit a finite generating partition. Note that by results of Seward from [71] and [72], the last condition is equivalent to an action admitting a countable generating partition with finite Shannon entropy.

In [7], the first author introduced a modified invariant called model-measure sofic entropy which is a lower bound for Bowen's sofic entropy. Let $\Sigma = (\sigma_n : G \to$

$\mathrm{Sym}(V_n)$) be a sofic approximation to G. Model-measure sofic entropy is constructed in terms of sequences $(\mu_n)_{n=1}^{\infty}$ where μ_n is a probability measure on A^{V_n}. If these measures replicate the process (A^G, μ) in an appropriate sense then we say that $(\mu_n)_{n=1}^{\infty}$ locally and empirically converges to μ. We refer the reader to [7] for the precise definitions. We have substituted the phrase 'local and empirical convergence' for the phrase 'quenched convergence' which appeared in [7]. This has been done to avoid confusion with an alternative use of the word 'quenched' in the physics literature. A process is said to have completely positive model-measure sofic entropy if every nontrivial factor has positive model-measure sofic entropy. The goal of this paper is the to prove the following theorem, which generalizes the main theorem of [32].

Theorem 6.1.1. *Let G be a countable sofic group containing an element of infinite order. Then there exists an uncountable family of pairwise nonisomorphic G-processes each of which has completely positive model-measure sofic entropy (and hence completely positive sofic entropy) with respect to any sofic approximation to G. None of these processes is a factor of a Bernoulli shift.*

In order to prove Theorem 6.1.1 we introduce a concept of uniform mixing for sequences of model-measures. This uniform model-mixing will be defined formally in Section 6.3. It implies completely positive model-measure sofic entropy.

Theorem 6.1.2. *Let G be a countable sofic group and let (A^G, μ) be a G-process with finite state space A. Suppose that for some sofic approximation Σ to G, there is a uniformly model-mixing sequence $(\mu_n)_{n=1}^{\infty}$ which locally and empirically converges to μ over Σ. Then (A^G, μ) has completely positive lower model-measure sofic entropy with respect to Σ.*

As in [32], the examples we exhibit to establish Theorem 6.1.1 are produced via a coinduction method for lifting H-processes to G-processes when $H \leq G$. If (A^H, ν) is an H-process then we can construct a corresponding G-process (A^G, μ) as follows. Let T be a transversal for the right cosets of H in G. Identify G as a set with $H \times T$ and thereby identify A^G with $(A^H)^T$. Set $\mu = \nu^T$. We call (A^G, μ) the coinduced process and denote it by $\mathrm{CInd}_H^G(\nu)$. (See page 72 of [53] for more details on this construction.) When $H \cong \mathbb{Z}$ this procedure preserves uniform mixing.

Theorem 6.1.3. *Let G be a countable sofic group and let $(A^{\mathbb{Z}}, \nu)$ be a uniformly mixing $\mathbb{Z}$-process with finite state space A. Let $H \leq G$ be a subgroup isomorphic* *to $\mathbb{Z}$ and identify $A^{\mathbb{Z}}$ with A^H. Then for any sofic approximation Σ to G, there is a uniformly model-mixing sequence of measures which locally and empirically converges to* $\mathrm{CInd}_H^G(\nu)$ *over* Σ.

We remark that it is easy to see that if (A^G, μ) is a Bernoulli shift (that is to say, μ is a product measure), then there is a uniformly model-mixing sequence which locally and empirically converges to μ. Indeed, if $\mu = \eta^G$ for a measure η on A then the measures η^{V_n} on A^{V_n} are uniformly model-mixing and locally and empirically converge to μ. Thus Theorem 6.1.2 shows that Bernoulli shifts with finite state space have completely positive sofic entropy, giving another proof of this case of the main theorem from [59]. We believe that completely positive sofic entropy for general Bernoulli shifts can be deduced along the same lines, requiring only a few additional estimates, but do not pursue the details here.

Acknowledgements

The first author's research was partially supported by the Simons Collaboration on Algorithms and Geometry. The second author's research was partially supported by NSF grants DMS-0968710 and DMS-1464475.

6.2 Preliminaries

Notation

The notation we use closely follows that in [7]; we refer the reader to that reference for further discussion. Let A be a finite set. For any pair of sets $W \subseteq S$ we let $\pi_W : A^S \to A^W$ be projection onto the W-coordinates (thus our notation leaves the larger set S implicit). Let G be a countable group and let (A^G, μ) be a G-process. For $F \subseteq G$ we will write $\mu_F = \pi_{F*}\mu \in \mathrm{Prob}(A^F)$ for the F-marginal of μ.

Let B be another finite set and let $\phi : A^G \to B$ be a measurable function. If $F \subseteq G$ we will say that ϕ is F-local if it factors through π_F. We will say ϕ is local if it is F-local for some finite F. Let $\phi^G : A^G \to B^G$ be given by $\phi^G(a)(g) = \phi(g \cdot a)$ and note that ϕ^G is equivariant between the right-shift on A^G and the right-shift on B^G.

Let V be a finite set and let σ be a map from G to $\mathrm{Sym}(V)$. For $g \in G$ and $v \in V$ we write $\sigma^g \cdot v$ instead of $\sigma(g)(v)$. For $F \subseteq G$ and $S \subseteq V$ we define

$$\sigma^F(S) = \{\sigma^g \cdot s : g \in F, s \in S\}.$$

For $v \in V$ we write $\sigma^F(v)$ for $\sigma^F(\{v\})$. We write $\Pi^\sigma_{v,F}$ for the map from A^V to A^F given by $\Pi^\sigma_{v,F}(\overline{a})(g) = \overline{a}(\sigma^g \cdot v)$ for $\overline{a} \in A^V$ and $g \in F$. We write Π^σ_v for $\Pi^\sigma_{v,G}$. With $\phi : A^G \to B$ as before, we write ϕ^σ for the map from A^V to B^V given by $\phi^\sigma(\overline{a})(v) = \phi\big(\Pi^\sigma_v(\overline{a})\big)$.

If D is a finite set and η is a probability measure on D then $\mathrm{H}(\eta)$ denotes the Shannon entropy of η, and for $\epsilon > 0$ we define

$$\mathrm{cov}_\epsilon(\eta) = \min\big\{|F| : F \subseteq D \text{ is such that } \eta(F) > 1 - \epsilon\big\}.$$

If $\phi : D \to E$ is a map to another finite set then we may write $\mathrm{H}_\mu(\phi)$ in place of $\mathrm{H}(\phi_*\mu)$. For $p \in [0, 1]$ we let $\mathrm{H}(p) = -p \log p - (1 - p) \log(1 - p)$.

We use the $o(\cdot)$ and $\lesssim$ asymptotic notations with respect to the limit $n \to \infty$. Given two functions f and g on $\mathbb{N}$, the notation $f \lesssim g$ means that there is a positive constant C such that $f(n) \le Cg(n)$ for all n.

An information theoretic estimate

Lemma 6.2.1. *Let A be a finite set and let $(V_n)_{n=1}^\infty$ be a sequence of finite sets such that $|V_n|$ increases to infinity. Let μ_n be a probability measure on A^{V_n}. We have*

$$\liminf_{n\to\infty} \frac{\mathrm{H}(\mu_n)}{|V_n|} \le \sup_{\epsilon>0} \liminf_{n\to\infty} \frac{1}{|V_n|} \log \mathrm{cov}_\epsilon(\mu_n).$$

Proof. Let μ be a probability measure on a finite set F and let $E \subseteq F$. By conditioning on the partition $\{E, F \setminus E\}$ and then recalling that entropy is maximized by uniform distributions we obtain

$$\begin{aligned} \mathrm{H}(\mu) &= \mu(E) \cdot \mathrm{H}(\mu(\cdot \mid E)) + \mu(F \setminus E) \cdot \mathrm{H}(\mu(\cdot \mid F \setminus E)) + \mathrm{H}(\mu(E)) \\ &\le \mu(E) \cdot \log(|E|) + (1 - \mu(E)) \cdot \log(|F \setminus E|) + \mathrm{H}(\mu(E)). \end{aligned} \tag{6.1}$$

Now let μ_n and V_n be as in the statement of the lemma. Let $\epsilon > 0$ and let $S_n \subseteq A^{V_n}$ be a sequence of sets with $\mu_n(S_n) > 1 - \epsilon$ and $|S_n| = \mathrm{cov}_\epsilon(\mu_n)$. By applying (6.1) with $F = A^{V_n}$ and $E = S_n$ we have

$$\begin{aligned}\liminf_{n\to\infty} \frac{\mathrm{H}(\mu_n)}{|V_n|} &\leq \liminf_{n\to\infty} \frac{1}{|V_n|}\big(\mu(S_n)\cdot\log(|S_n|) \\ &\qquad\qquad + (1-\mu(S_n))\cdot\log(|A^{V_n}\setminus S_n|) + \mathrm{H}(\mu(S_n))\big) \\ &\leq \liminf_{n\to\infty} \frac{1}{|V_n|}\big(\log(|S_n|) + \epsilon\cdot\log(|A^{V_n}|) + \mathrm{H}(\epsilon)\big) \\ &\leq \left(\liminf_{n\to\infty} \frac{1}{|V_n|}\log\operatorname{cov}_\epsilon(\mu_n)\right) + \epsilon\cdot\log(|A|).\end{aligned}$$

Now let ϵ tend to zero to obtain the lemma. □

6.3 Metrics on sofic approximations and uniform model-mixing

Let us fix a proper right-invariant metric ρ on G: for instance, if G is finitely generated then ρ can be a word metric, and more generally we may let $w : G \to [0,\infty)$ be any proper weight function and define ρ to be the resulting weighted word metric. Again let V be a finite set and let σ be a map from G to $\mathrm{Sym}(V)$. Let H_σ be the graph on V with an edge from v to w if and only if $\sigma^g\cdot v = w$ or $\sigma^g\cdot w = v$ for some $g\in G$. Define a weight function W on the edges of H_σ by setting

$$W(v,w) = \min\big\{\rho(g,1_G) : \sigma^g\cdot v = w \text{ or } \sigma^g\cdot w = v\big\}.$$

If v and w are in the same connected component of H_σ let ρ_σ be the W-weighted graph distance between v and w, that is

$$\begin{aligned}\rho_\sigma(v,w) = \min\Bigg\{\sum_{i=0}^{k-1} W(p_i,p_{i+1}) :&(v = p_0, p_1, \ldots, p_{k-1}, p_k = w) \\ &\text{is an } H_\sigma\text{-path from } v \text{ to } w\Bigg\}.\end{aligned}$$

Having defined ρ_σ on the connected components of H_σ, choose some number M much larger than the ρ_σ-distance between any two points in the same connected component. Set $\rho_\sigma(v,w) = M$ for any pair v,w of vertices in distinct connected components of H_σ. Note that if $(\sigma_n : G\to\mathrm{Sym}(V_n))$ is a sofic approximation to G then for any fixed $r<\infty$ once n is large enough the map $g\mapsto\sigma_n^g\cdot v$ restricts to an isometry from $B_\rho(1_G,r)$ to $B_{\rho_{\sigma_n}}(v,r)$ for most $v\in V_n$.

In the sequel the sofic approximation will be fixed, and we will abbreviate ρ_{σ_n} to ρ_n. We can now state the main definition of this paper.

Definition 6.3.1. *Let $(V_n)_{n=1}^{\infty}$ be a sequence of finite sets with $|V_n| \to \infty$ and for each n let σ_n be a map from G to $\mathrm{Sym}(V_n)$. Let A be a finite set. For each $n \in \mathbb{N}$ let μ_n be a probability measure on A^{V_n}. We say the sequence $(\mu_n)_{n=1}^{\infty}$ is* ***uniformly model-mixing*** *if the following holds. For every finite $F \subseteq G$ and every $\epsilon > 0$ there is some $r < \infty$ and a sequence of subsets $W_n \subseteq V_n$ such that*

$$|W_n| = (1 - o(1))|V_n|$$

and if $S \subseteq W_n$ is r-separated according the metric ρ_n then

$$\mathrm{H}\left(\pi_{\sigma_n^F(S)*}\mu_n\right) \geq |S| \cdot (\mathrm{H}(\mu_F) - \epsilon).$$

This definition is motivated by Weiss' notion of uniform mixing from the special case when G is amenable: see [76] and also Section 4 of [32]. Let us quickly recall that notion in the setting of a G-process (A^G, μ). First, if $K \subseteq G$ is finite and $S \subseteq G$ is another subset, then S is K**-spread** if any distinct elements $s_1, s_2 \in S$ satisfy $s_1 s_2^{-1} \notin K$. The process (A^G, μ) is **uniformly mixing** if, for any finite-valued measurable function $\phi : A^G \to B$ and any $\epsilon > 0$, there exists a finite subset $K \subseteq G$ with the following property: if $S \subseteq G$ is another finite subset which is K-spread, then

$$\mathrm{H}\big((\phi_*^G \mu)_S\big) \geq |S| \cdot (\mathrm{H}_\mu(\phi) - \epsilon).$$

Beware that we have reversed the order of multiplying s_1 and s_2^{-1} in the definition of 'K-spread' compared with [32]. This is because we work in terms of observables such as ϕ rather than finite partitions of A^G, and shifting an observable by the action of $g \in G$ corresponds to shifting the partition that it generates by g^{-1}.

The principal result of [69] is that completely positive entropy implies uniform mixing. The reverse implication also holds: see [43] or Theorem 4.2 in [32]. Thus, uniform mixing is an equivalent characterization of completely positive entropy.

The definition of uniform mixing may be rephrased in terms of our proper metric ρ on G as follows. The process (A^G, μ) is uniformly mixing if and only if, for any finite-valued measurable function $\phi : A^G \to B$ and any $\epsilon > 0$, there exists an $r < \infty$ with the following property: if $S \subseteq G$ is r-separated according to ρ, then

$$\mathrm{H}\big((\phi_*^G \mu)_S\big) \geq |S| \cdot (\mathrm{H}_\mu(\phi) - \epsilon).$$

This is equivalent to the previous definition because a subset $S \subseteq G$ is r-separated according to ρ if and only if it is $B_\rho(1_G, r)$-spread. The balls $B_\rho(1_G, r)$ are finite, because ρ is proper, and any other finite subset $K \subseteq G$ is contained in $B_\rho(1_G, r)$ for all sufficiently large r.

This is the point of view on uniform mixing which motivates Definition 6.3.1. We use the right-invariant metric ρ rather than the general definition of 'K-spread' sets because it is more convenient later.

Definition 6.3.1 is directly compatible with uniform mixing in the following sense. If G is amenable and $(F_n)_{n=1}^{\infty}$ is a Følner sequence for G, then the sets F_n may be regarded as a sofic approximation to G: an element $g \in G$ acts on F_n by translation wherever this stays inside F_n and arbitrarily at points which are too close to the boundary of F_n. If (A^G, μ) is an ergodic G-process, then it follows easily that the sequence of marginals μ_{F_n} locally and empirically converge to μ over this Følner-set sofic approximation. If (A^G, μ) is uniformly mixing, then this sequence of marginals is clearly uniformly model-mixing in the sense of Definition 6.3.1.

On the other hand, suppose that (A^G, μ) admits a sofic approximation and a locally and empirically convergent sequence of measures over that sofic approximation which is uniformly model-mixing. Then our Theorem 6.1.2 shows that (A^G, μ) has completely positive sofic entropy. If G is amenable then sofic entropy always agrees with Kolmogorov-Sinai entropy [17], and this implies that (A^G, μ) has completely positive entropy and hence is uniformly mixing, by the result of [69].

Thus if G is amenable then completely positive entropy and uniform mixing are both equivalent to the existence of a sofic approximation and a locally and empirically convergent sequence of measures over it which is uniformly model-mixing. If these conditions hold, then we expect that one can actually find a locally and empirically convergent and uniformly model-mixing sequence of measures over *any* sofic approximation to G. This should follow using a similar kind of decomposition of the sofic approximants into Følner sets as in Bowen's proof in [17]. However, we have not explored this argument in detail.

Definition 6.3.1 applies to a shift-system with a finite state space. It can be transferred to an abstract measure-preserving G-action on (X, μ) by fixing a choice of finite measurable partition of X. However, in order to study actions which do not admit a finite generating partition, it might be worth looking for an extension of

Definition 6.3.1 to G-processes with arbitrary compact metric state spaces, similarly to the setting in [7]. We also do not pursue this generalization here.

6.4 Proof of Theorem 6.1.2

We will use basic facts about the Shannon entropy of observables (i.e. random variables with finite range), for which we refer the reader to Chapter 2 of [30]. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$, (A^G, μ) and $(\mu_n)_{n=1}^{\infty}$ be as in the statement of Theorem 6.1.2. The following is the 'finitary' model-measure analog of Lemma 5.1 in [32].

Lemma 6.4.1. *Let $F \subseteq G$ be finite. Let B be a finite set and let $\phi : A^G \to B$ be an F-local observable. Let $S_n \subseteq V_n$ be a sequence of sets such that $|S_n| \gtrsim |V_n|$. Then we have*

$$\mathrm{H}(\mu_F) - \frac{1}{|S_n|}\mathrm{H}(\pi_{\sigma_n^F(S_n)*}\mu_n) \geq \mathrm{H}_\mu(\phi) - \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n*}\phi_*^{\sigma_n}\mu_n\big) - o(1).$$

Proof of Lemma 6.4.1. Let $\theta : A^F \to B$ be a map with $\theta \circ \pi_F = \phi$. Fix $n \in \mathbb{N}$ and $S \subseteq V_n$. Let $\alpha = \pi_{\sigma_n^F(S)} : A^{V_n} \to A^{\sigma_n^F(S)}$ and let $\beta = \pi_S \circ \phi^{\sigma_n} : A^{V_n} \to B^S$. For $s \in S$ let $\alpha_s = \Pi_{s,F}^{\sigma_n} : A^{V_n} \to A^F$ and let $\beta_s = \theta \circ \Pi_{s,F}^{\sigma_n} : A^{V_n} \to B$. Then we have $\alpha = (\alpha_s)_{s\in S}$ and $\beta = (\beta_s)_{s\in S}$. Enumerate $S = (s_k)_{k=1}^m$ and write $\alpha_{s_k} = \alpha_k$. All entropies in the following display are computed with respect to μ_n. We have

$$\begin{aligned}
\mathrm{H}(\alpha) &= \mathrm{H}(\alpha_1, \ldots, \alpha_m)\\
&= \mathrm{H}(\alpha_1) + \sum_{k=1}^{m-1} \mathrm{H}(\alpha_{k+1}|\alpha_1, \ldots, \alpha_k)\\
&= \mathrm{H}(\alpha_1, \beta_1) + \sum_{k=1}^{m-1} \mathrm{H}(\alpha_{k+1}, \beta_{k+1}|\alpha_1, \ldots, \alpha_k)\\
&= \mathrm{H}(\beta_1) + \mathrm{H}(\alpha_1|\beta_1) + \sum_{k=1}^{m-1} \mathrm{H}(\beta_{k+1}|\alpha_1, \ldots, \alpha_k) + \sum_{k=1}^{m-1} \mathrm{H}(\alpha_{k+1}|\beta_{k+1}, \alpha_1, \ldots, \alpha_k)\\
&\leq \mathrm{H}(\beta_1) + \sum_{k=1}^{m-1} \mathrm{H}(\beta_{k+1}|\beta_1, \ldots, \beta_k) + \sum_{k=1}^{m} \mathrm{H}(\alpha_k|\beta_k)\\
&= \mathrm{H}(\beta) + \sum_{k=1}^{m} \mathrm{H}(\alpha_k|\beta_k).
\end{aligned}$$

Let ι be the identity map on A^F. Then

$$\begin{aligned}
|S| \cdot \mathrm{H}(\mu_F) - \mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) &= |S| \cdot \mathrm{H}_{\mu_F}(\iota) - \mathrm{H}_{\mu_n}(\alpha) \\
&\geq |S| \cdot \mathrm{H}_{\mu_F}(\theta) + |S| \cdot \mathrm{H}_{\mu_F}(\iota|\theta) \\
&\qquad - \mathrm{H}_{\mu_n}(\beta) - \sum_{s\in S} \mathrm{H}_{\mu_n}(\alpha_s|\beta_s) \\
&= |S| \cdot \mathrm{H}_{\mu}(\phi) - \mathrm{H}\big(\pi_{S*}\phi_*^{\sigma_n}\mu_n\big) \\
&\qquad + |S| \cdot \mathrm{H}_{\mu_F}(\iota|\theta) - \sum_{s\in S} \mathrm{H}_{\mu_n}(\alpha_s|\beta_s). \qquad (6.2)
\end{aligned}$$

Now allowing n to vary, let $S_n \subseteq V_n$ be a sequence of sets such that $|S_n| \gtrsim |V_n|$. Write $\nu_n = \pi_{\sigma_n^F(S_n)*}\mu_n$. Let $s \in S_n$ be such that the obvious map from F to $\sigma_n^F(s)$ is injective. Then the function $\overline{a} \mapsto \Pi_{s,F}^{\sigma_n}(\overline{a})$ provides an identification of $A^{\sigma_n^F(s)}$ with A^F. This identification sends α_s to ι and β_s to θ. When n is large the $\sigma_n^F(s)$ marginal of μ_n will resemble μ_F for most $s \in S_n$. Since α_s and β_s are $\pi_{\sigma_n^F(s)}$ measurable this implies that $\mathrm{H}_{\mu_F}(\iota|\theta) \approx \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)$ for most s. More precisely, we can find a sequence of sets $C_n \subseteq S_n$ with

$$|C_n| = (1 - o(1))|S_n|$$

such that

$$\max_{s\in C_n} \big|\mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\big| = o(1).$$

Thus

$$\begin{aligned}
\left| |S_n| \cdot \mathrm{H}_{\mu_F}(\iota|\theta) - \sum_{s\in S_n} \mathrm{H}_{\nu_n}(\alpha_s|\beta_s) \right| &\leq \sum_{s\in C_n} \big|\mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\big| \\
&\qquad + \sum_{s\in S_n\setminus C_n} \big|\mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\big| \\
&= o(|S_n|).
\end{aligned}$$

The lemma then follows from (6.2) and the above. □

Recall that for a measure space (X, μ) and two observables α and β on X the Rokhlin distance between α and β is defined by

$$d_\mu^{\mathrm{Rok}}(\alpha, \beta) = \mathrm{H}_\mu(\alpha|\beta) + \mathrm{H}_\mu(\beta|\alpha).$$

This is a pseudometric on the space of observables on X. An easy computation shows that if $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$ are two families of observables on X then

$$d_\mu^{\mathrm{Rok}}((\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n)) \leq \sum_{k=1}^{n} d_\mu^{\mathrm{Rok}}(\alpha_k, \beta_k).$$

Lemma 6.4.2. *Let $\phi, \psi : A^G \to B$ be two local observables. Let $S_n \subseteq V_n$ be a sequence of sets with $|S_n| \gtrsim |V_n|$. Then we have*

$$\frac{1}{|S_n|}\left|\mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_* \mu_n) - \mathrm{H}(\pi_{S_n *}\psi^{\sigma_n}_* \mu_n)\right| \leq d^{\mathrm{Rok}}_\mu(\phi, \psi) + o(1).$$

Proof. Let $\alpha_n = \pi_{S_n} \circ \phi^{\sigma_n} : A^{V_n} \to B^{S_n}$ and let $\beta_n = \pi_{S_n} \circ \psi^{\sigma_n} : A^{V_n} \to B^{S_n}$. Let F be a finite subset of G such that both ϕ and ψ are F-local. Let $\theta : A^F \to B$ be a map such that $\theta \circ \pi_F = \phi$ and let $\kappa : A^F \to B$ be a map such that $\kappa \circ \pi_F = \psi$. For $s \in S_n$ let $\alpha_{n,s} = \theta \circ \Pi^{\sigma_n}_{s,F} : A^{V_n} \to B$ so that $\alpha_n = (\alpha_{n,s})_{s \in S_n}$. Also let $\beta_{n,s} = \kappa \circ \Pi^{\sigma_n}_{s,F} : A^{V_n} \to B$. Then we have

$$\begin{aligned}
\frac{1}{|S_n|}\left|\mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_* \mu_n) - \mathrm{H}(\pi_{S_n *}\psi^{\sigma_n}_* \mu_n)\right| &= \frac{1}{|S_n|}\left|\mathrm{H}_{\mu_n}(\alpha_n) - \mathrm{H}_{\mu_n}(\beta_n)\right| \\
&\leq \frac{1}{|S_n|} \cdot d^{\mathrm{Rok}}_{\mu_n}(\alpha_n, \beta_n) \\
&= \frac{1}{|S_n|} \cdot d^{\mathrm{Rok}}_{\mu_n}\big((\alpha_{n,s})_{s \in S_n}, (\beta_{n,s})_{s \in S_n}\big) \\
&\leq \frac{1}{|S_n|} \sum_{s \in S_n} d^{\mathrm{Rok}}_{\mu_n}(\alpha_{n,s}, \beta_{n,s}). \qquad (6.3)
\end{aligned}$$

If the map $g \mapsto \sigma_n^g \cdot s$ is injective on F, we can identify $A^{\sigma_n^F(s)}$ with A^F and thereby identify $\alpha_{n,s}$ with θ and $\beta_{n,s}$ with κ. Note that

$$d^{\mathrm{Rok}}_{\mu_F}(\theta, \kappa) = d^{\mathrm{Rok}}_\mu(\phi, \psi).$$

It follows that for any $\epsilon > 0$ we can find a weak star neighborhood O of μ such that if $s \in S_n$ is such that $(\Pi^{\sigma_n}_s)_* \mu_n \in O$ then

$$\left|d^{\mathrm{Rok}}_{\mu_n}(\alpha_{n,s}, \beta_{n,s}) - d^{\mathrm{Rok}}_\mu(\phi, \psi)\right| < \epsilon.$$

Thus, since μ_n locally and empirically converges to μ, there are sets $C_n \subseteq S_n$ with $|C_n| = (1 - o(1))|S_n|$ such that

$$\max_{s \in C_n}\left|d^{\mathrm{Rok}}_{\mu_n}(\alpha_{n,s}, \beta_{n,s}) - d^{\mathrm{Rok}}_\mu(\phi, \psi)\right| = o(1). \qquad (6.4)$$

The lemma now follows from (6.3) and (6.4). □

Corollary 6.4.1. *Let $\big(\phi_m : A^G \to B\big)_{m=1}^\infty$ be a sequence of local observables and let $\phi : A^G \to B$ be a local observable. Let $S_n \subseteq V_n$ be a sequence of sets with $|S_n| \gtrsim |V_n|$. Then if $(m_n)_{n=1}^\infty$ increases to infinity at a slow enough rate we have*

$$\frac{1}{|S_n|}\left|\mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_* \mu_n) - \mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_{m_n *} \mu_n)\right| \leq d^{\mathrm{Rok}}_\mu(\phi, \phi_{m_n}) + o(1).$$

Proof of Theorem 6.1.2. Let B be a finite set and let $\psi : A^G \to B$ be an observable with $\mathrm{H}_\mu(\psi) > 0$. Let $(\phi_m)_{m=1}^\infty$ be an AL approximating sequence for ψ rel μ (see Definition 4.4 in [7]). Then the sequence ϕ_m converges to ψ in d_μ^{Rok}. In particular, ϕ_m is a Cauchy sequence and so we can find $M \in \mathbb{N}$ so that for all $m \geq M$ we have

$$d_\mu^{\mathrm{Rok}}(\phi_m, \phi_M) \leq \frac{\mathrm{H}_\mu(\psi)}{8}. \tag{6.5}$$

We will also assume M is large enough that

$$\mathrm{H}_\mu(\phi_M) \geq \frac{\mathrm{H}_\mu(\psi)}{2}. \tag{6.6}$$

Let F be a finite subset of G such that ϕ_M is F-local. Then Definition 6.3.1 provides an $r < \infty$ and a sequence of subsets $W_n \subseteq V_n$ such that $|W_n| = (1 - o(1))|V_n|$ and if $S \subseteq W_n$ is r-separated then

$$\mathrm{H}(\mu_F) - \frac{1}{|S|}\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \leq \frac{\mathrm{H}_\mu(\phi_M)}{2}. \tag{6.7}$$

Let $K = |B_\rho(1_G, r)|$. Since σ_n is a sofic approximation there are sets $W_n' \subseteq V_n$ with $|W_n'| = (1 - o(1))|V_n|$ such that if $w \in W_n'$ then the ρ_n ball of radius r around w has cardinality at most K. Write $Y_n = W_n \cap W_n'$ and note that we have $|Y_n| = (1 - o(1))|V_n|$. For each n let S_n be an r-separated subset of Y_n with maximal cardinality. Then $Y_n \subseteq \bigcup_{s \in S_n} B_{\rho_n}(s, r)$ so that

$$|S_n| \geq \frac{|Y_n|}{K} = (1 - o(1))\frac{|V_n|}{K}. \tag{6.8}$$

By Lemma 6.4.1 and (6.7) we have

$$\mathrm{H}_\mu(\phi_M) - \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n*}\phi^{\sigma_n}_{M*}\mu_n\big) - o(1) \leq \frac{\mathrm{H}_\mu(\phi_M)}{2}$$

so that from (6.6) we have

$$\frac{\mathrm{H}_\mu(\psi)}{4} - o(1) \leq \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n*}\phi^{\sigma_n}_{M*}\mu_n\big). \tag{6.9}$$

By Proposition 5.15 in [7] if $(m_n)_{n=1}^\infty$ increases to infinity at a slow enough rate then $(\phi^{\sigma_n}_{m_n})_*\mu_n$ will locally and empirically converge to $\psi^G_*\mu$. Since A is finite, by the same argument as for Proposition 8.1 in [7] we have

$$\begin{aligned}\underline{\mathrm{h}}^{\mathrm{q}}_\Sigma\big(\psi^G_*\mu\big) &\geq \sup_{\epsilon>0}\liminf_{n\to\infty}\frac{1}{|V_n|}\log\mathrm{cov}_\epsilon\big((\phi^{\sigma_n}_{m_n})_*\mu_n\big) \\ &\geq \liminf_{n\to\infty}\frac{1}{|V_n|}\,\mathrm{H}\big((\phi^{\sigma_n}_{m_n})_*\mu_n\big),\end{aligned} \tag{6.10}$$

where the second inequality follows from Lemma 6.2.1. We also assume that $(m_n)_{n=1}^\infty$ increases slowly enough for Corollary 6.4.1 to hold. By (6.5) we have

$$\left|\frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}\phi_{M*}^{\sigma_n}\mu_n\big) - \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}(\phi_{m_n}^{\sigma_n})_*\mu_n\big)\right| \leq \frac{\mathrm{H}_\mu(\psi)}{8} + o(1).$$

Combining this with (6.9) we see that

$$\frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}(\phi_{m_n}^{\sigma_n})_*\mu_n\big) \geq \frac{\mathrm{H}_\mu(\psi)}{8} - o(1).$$

By the above and (6.8) we have that for all sufficiently large n,

$$\mathrm{H}\big((\phi_{m_n}^{\sigma_n})_*\mu_n\big) \geq \frac{\mathrm{H}_\mu(\psi)}{8K+1}|V_n|. \tag{6.11}$$

Theorem 6.1.2 now follows from (6.10) and (6.11). □

6.5 Proof of Theorem 6.1.3

Let $(A^{\mathbb{Z}}, \nu)$ be a uniformly mixing $\mathbb{Z}$-process, and for each positive integer l let ν_l be the marginal of ν on A^l. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$ be an arbitrary sofic approximation to G. Let $h \in G$ have infinite order and write $H = \langle h \rangle \cong \mathbb{Z}$. We construct a measure μ_n on A^{V_n} for each $n \in \mathbb{N}$. We will later show that the sequence $(\mu_n)_{n=1}^\infty$ is uniformly model-mixing and locally and empirically converges to μ over Σ.

We first construct a measure μ_n^l on A^{V_n} for each pair (n, l) with l much smaller than n. For a given n, the single permutation σ_n^h partitions V_n into a disjoint union of cycles. Since h has infinite order and Σ is a sofic approximation, once n is large most points will be in very long cycles. In particular we assume that most points are in cycles with length much larger than l. Partition the cycles into disjoint paths so that as many of the paths have length l as possible, and let $\mathcal{P}_n^l = (P_{n,1}^l, \dots, P_{n,k_n}^l)$ be the collection of all length-l paths that result (so $\mathcal{P}_n^l$ is not a partition of the whole of V_n, but covers most of it). Fix any element $\overline{a}_0 \in A^{V_n}$ and define a random element $\overline{a} \in A^{V_n}$ by choosing each restriction $\overline{a} \restriction_{P_{n,i}^l}$ independently with the distribution of ν_l and extending to the rest of V_n according to $\overline{a}_0$. Let μ_n^l be the law of this $\overline{a}$.

Now let $(l_n)_{n=1}^\infty$ increase to infinity at a slow enough rate that the following two conditions are satisfied:

(a) The number of points of V_n that lie in some member of the family $\mathcal{P}_n^{l_n}$ is $(1 - o(1))|V_n|$.

(b) Whenever $g, g' \in G$ lie in distinct right cosets of H, so that $g^{-1}h^p g' \neq 1_G$ for all $p \in \mathbb{Z}$, we have

$$|\{v \in V_n : (\sigma_n^g)^{-1}(\sigma_n^h)^p \sigma_n^{g'} \cdot v = v \text{ for some } p \in \{-l_n, \ldots, l_n\}\}| = o(|V_n|)$$

Set $\mu_n = \mu_n^{l_n}$. We separate the proof that $(\mu_n)_{n=1}^{\infty}$ has the required properties into two lemmas.

Lemma 6.5.1. $(\mu_n)_{n=1}^{\infty}$ *locally and empirically converges to* μ *over* Σ.

Proof of Lemma 6.5.1. Since (A^G, μ) is ergodic, by Corollary 5.6 in [7] it suffices to show that μ_n locally weak star converges to μ. For a set $I \subseteq \mathbb{Z}$ write $h^I = \{h^i : i \in I\}$. Fix a finite set $F \subseteq G$. By enlarging F if necessary we can assume there is an interval $I \subseteq \mathbb{Z}$ such that $F = \bigcup_{k=1}^m h^I t_k$ for $t_1, \ldots, t_m$ in some transversal for the right cosets of H in G. For each $g \in F$ let j_g be a fixed element of A. Let $B \subseteq A^G$ be defined by

$$B = \left\{a \in A^G : a(g) = j_g \text{ for all } g \in F\right\}$$

and let $\epsilon > 0$. Then sets such as

$$O = \left\{\eta \in \mathrm{Prob}(A^G) : \eta(B) \approx_\epsilon \mu(B)\right\}$$

form a subbasis of neighborhoods around μ. It therefore suffices to show that when n is large we have $(\Pi_v^{\sigma_n})_* \mu_n \in O$ with high probability in the choice of $v \in V_n$.

For $k \in \{1, \ldots, m\}$ let

$$B_k = \left\{x \in A^{\mathbb{Z}} : x(i) = j_{h^i t_k} \text{ for all } i \in I\right\}.$$

Note that μ is defined in such a way that $\mu(B) = \prod_{i=1}^k \nu(B_k)$. Now, let W_n be the set of all points $v \in V_n$ such that the following conditions hold.

(i) The map $g \mapsto \sigma_n^g \cdot v$ is injective on F.

(ii) $\sigma_n^{h^i t_k} \cdot v = (\sigma_n^h)^i \sigma_n^{t_k} \cdot v$ for all $i \in I$ and $k \in \{1, \ldots, m\}$.

(iii) For all pairs $g, g' \in F$, $\sigma_n^g \cdot v$ is in the same path as $\sigma_n^{g'} \cdot v$ if and only if g and g' lie in the same right coset of H. In particular, each of the images $\sigma_n^g \cdot v$ for $g \in F$ is contained in some member of $\mathcal{P}_n^{l_n}$.

We claim that $|W_n| = (1-o(1))|V_n|$. Clearly Conditions (i) and (ii) are satisfied with high probability in v, and so is the last part of Condition (iii), by Condition (a) in the choice of $(l_n)_{n=1}^{\infty}$.

Fix $g, g' \in F$ and suppose that g and g' are in the same coset of H, so that we have $g = h^i t_k$ and $g' = h^{i'} t_k$ for some $k \in \{1, \ldots, m\}$ and $i, i' \in I$. If v satisfies Condition (ii) then we have

$$(\sigma_n^h)^{i'-i}\sigma_n^g \cdot v = (\sigma_n^h)^{i'-i}(\sigma_n^h)^i\sigma_n^{t_k} \cdot v = (\sigma_n^h)^{i'}\sigma_n^{t_k} \cdot v = \sigma_n^{g'} \cdot v$$

so that $\sigma_n^g \cdot v$ and $\sigma_n^{g'} \cdot v$ will lie in the same path assuming that $\sigma_n^{t_k} \cdot v$ is not one of the first or last $|I|$ elements of its path. Note that for any $v \in V_n$ we have

$$\left|\left\{w : \sigma_n^{t_k} \cdot w = v \text{ for some } k \in \{1, \ldots, m\}\right\}\right| \leq m.$$

It follows that the number of points $v \in V_n$ such that $\sigma_n^{t_k} \cdot v$ is one of the first or last $|I|$ elements of a path is at most $2mp_n|I| + o(|V_n|)$ where p_n is the total number of paths in V_n. By Condition (a) in the choice of $(l_n)_{n=1}^{\infty}$, most of V_n is covered by paths whose lengths increase to infinity. Since also $p_n = o(V_n)$, it follows that $\sigma_n^g \cdot v$ lies in the same path as $\sigma_n^{g'} \cdot v$ with high probability in v.

On the other hand, suppose that g and g' are in distinct cosets of H. Assume that $\sigma_n^g \cdot v$ and $\sigma_n^{g'} \cdot v$ are in the same path. Then there is $p \in \{-l_n, \ldots, l_n\}$ with $\sigma_n^g \cdot v = (\sigma_n^h)^p \sigma_n^{g'} \cdot v$, and hence $(\sigma_n^g)^{-1}(\sigma_n^h)^p\sigma_n^{g'} \cdot v = v$. By Condition (b) in the choice of $(l_n)_{n=1}^{\infty}$ there are only $o(|V_n|)$ vertices v for which this holds. Thus we have established the claim.

Now let $v \in W_n$. We have

$$(\Pi_v^{\sigma_n})_*\mu_n(B) = \mu_n\left(\left\{\overline{a} \in A^{V_n} : \overline{a}(\sigma_n^g \cdot v) = j_g \text{ for all } g \in F\right\}\right).$$

For each $k \in \{1, \ldots, m\}$ the set $\{(\sigma_n^h)^i\sigma_n^{t_k} \cdot v : i \in I\}$ is contained in a single path. Since the marginal of μ_n on each path is ν_{l_n} the probability that

$$\overline{a}\left((\sigma_n^h)^i\sigma_n^{t_k} \cdot v\right) = j_{h^i t_k}$$

for all $i \in I$ is equal to $\nu_{l_n}(B_k) = \nu(B_k)$. On the other hand, the marginals of μ_n on distinct paths are independent, so the probability that $\overline{a}(\sigma_n^g \cdot v) = j_g$ for all $g \in F$ is actually equal to $\prod_{i=1}^{k} \nu(B_k)$. $\square$

If $(A^{\mathbb{Z}}, \nu)$ is weakly mixing, then so is the co-induced G-action. In particular, this certainly holds if $(A^{\mathbb{Z}}, \nu)$ is uniformly mixing. Therefore we may immediately promote Lemma 6.5.1 to the fact that $(\mu_n)_{n=1}^{\infty}$ locally and doubly empirically converges to μ over Σ, by Lemma 5.15 of [7]. In fact, we suspect that local and double empirical convergence holds here whenever $(A^{\mathbb{Z}}, \nu)$ is ergodic.

Lemma 6.5.2. $(\mu_n)_{n=1}^{\infty}$ *is uniformly model-mixing.*

Proof of Lemma 6.5.2. Let $F \subseteq G$ be finite and let $\epsilon > 0$. Again decompose $F = \bigcup_{k=1}^{m} h^I t_k$ for some interval $I \subseteq \mathbb{Z}$ and elements $t_k \in T$. Note that the restriction of the metric ρ to H is a proper right invariant metric on $H \cong \mathbb{Z}$, even though it might be different from the usual metric on $\mathbb{Z}$. Thus since ν is uniformly mixing we can find some $r_0 < \infty$ such that if $(I_j)_{j=1}^{q}$ is a family of intervals in $\mathbb{Z}$ which are each of length $|I|$ and are pairwise at distance at least r_0 then writing $K = \bigcup_{j=1}^{q} I_j$ we have

$$\mathrm{H}(\nu_K) \geq q \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right). \tag{6.12}$$

Let $r < \infty$ be large enough that for all $g, g' \in G$ if $\rho(g, g') \geq r$ then $\rho(fg, f'g') \geq r_0$ for all $f, f' \in F$. Such a choice of r is possible since by right-invariance of ρ we have $\rho(fg, g) = \rho(f, 1_G)$ and $\rho(f'g', g') = \rho(f', 1_G)$. Let W_n be as in the proof of Lemma 6.5.1 and recall that $|W_n| = (1 - o(1))|V_n|$. Let $S \subseteq W_n$ be r-separated according to ρ_n.

Fix a path $P \in \mathcal{P}_n^{l_n}$ and let S_P be the set of points $v \in S$ such that $\sigma_n^{t_{k(v)}} \cdot v \in P$ for some $k(v) \in \{1, \ldots, m\}$. Since $S \subseteq W_n$, Condition (iii) from the previous proof implies that

$$\sigma_n^F(S) \cap P = \bigcup_{v \in S_P} \{(\sigma_n^h)^i \sigma_n^{t_{k(v)}} \cdot v : i \in I\}.$$

Each of the sets in the latter union is an interval of length $|I|$ in P and by our choice of r these are pairwise at distance r_0 in ρ_n restricted to P. Since the marginal of μ_n on P is equal to ν_{n_l}, (6.12) implies that

$$\mathrm{H}(\pi_{(\sigma_n^F(S) \cap P)*}\mu_n) \geq |S_P| \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right).$$

Since the marginals of μ_n on distinct paths are independent, this implies that

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq \left(\sum_{P \in \mathcal{P}_n^{l_n}} |S_P|\right) \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right). \tag{6.13}$$

By Condition (iii) in the definition of W_n, each $v \in S$ appears in S_P for exactly m paths P. Therefore

$$\sum_{P \in \mathcal{P}_n^{l_n}} |S_P| = m \cdot |S|. \tag{6.14}$$

Now $\mathrm{H}(\mu_F) = m \cdot \mathrm{H}(\nu_I)$ so from (6.13) and (6.14) we have

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq |S| \cdot (\mathrm{H}(\mu_F) - \epsilon)$$

as required. □

Proof of Theorem 6.1.3. Theorem 6.1.3 now follows from Theorem 6.1.2 and Lemmas 6.5.1 and 6.5.2. □

6.6 Proof of Theorem 6.1.1

Proof of Theorem 6.1.1. This part of the argument is essentially the same as the corresponding part of [32]. Consider the family of uniformly mixing $\mathbb{Z}$-processes $\{(4^{\mathbb{Z}}, \nu_\omega) : \omega \in 2^{\mathbb{N}}\}$ constructed in Section 6 of [32]. Fix an isomorphic copy H of $\mathbb{Z}$ in G and let $\mu_\omega = \mathrm{CInd}_H^G(\nu_\omega)$. By Theorems 6.1.2 and 6.1.3 the process $(4^G, \mu_\omega)$ has completely positive model-measure sofic entropy. Note that the restriction of the G-action to H is a permuted power of the original $\mathbb{Z}$-process in the sense of Definition 6.5 from [32]. Thus by Proposition 6.6 in that reference, the processes $\{(4^G, \mu_\omega) : \omega \in 2^{\mathbb{N}}\}$ are pairwise nonisomorphic.

Suppose toward a contradiction that for some ω, $(4^G, \mu_\omega)$ is a factor of a Bernoulli shift (Z^G, η^G) over some standard probability space (Z, η). Let $\psi : Z^G \to 4^G$ be an equivariant measurable map with $\psi_* \eta^G = \mu_\omega$. Note that the restricted right-shift action $H \curvearrowright (Z^G, \eta^G)$ is still isomorphic to a Bernoulli shift and ψ is still a factor map from this process to the restricted action $H \curvearrowright (4^G, \mu_\omega)$. Thus the latter $\mathbb{Z}$-process is isomorphic to a Bernoulli shift and so is its factor $(4^{\mathbb{Z}}, \nu_\omega)$. This contradicts Corollary 6.4 in [32]. □

BIBLIOGRAPHY

[1] M. Abert and G. Elek. The space of actions, partition metric and combinatorial rigidity. Preprint. Available at http://arxiv.org/abs/1108.2147, 2011.

[2] M. Abért and B. Weiss. Bernoulli actions are weakly contained in any free action. *Ergodic Theory and Dynam. Systems*, 33(2):323–333, 2013.

[3] Miklós Abért and Gábor Elek. Dynamical properties of profinite actions. *Ergodic Theory Dynam. Systems*, 32(6):1805–1835, 2012. ISSN 0143-3857. doi: 10.1017/S0143385711000654. URL `http://dx.doi.org/10.1017/S0143385711000654`.

[4] Miklós Abért, Yair Glasner, and Bálint Virág. Kesten's theorem for invariant random subgroups. *Duke Math. J.*, 163(3):465–488, 2014. ISSN 0012-7094. doi: 10.1215/00127094-2410064. URL `http://dx.doi.org/10.1215/00127094-2410064`.

[5] R. Adler, A. Konheim, and M. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.*, 57:309–319, 1965.

[6] A. Alpeev. The entropy of Gibbs measures on sofic groups. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 436(Teoriya Predstavlenii, Dinamicheskie Sistemy, Kombinatornye Metody. XXV):34–48, 2015. ISSN 0373-2703.

[7] Tim Austin. Additivity properties of sofic entropy and measures on model spaces. *To appear in Forum of Mathematics, Sigma*, 2015. Available at http://arxiv.org/abs/1510.02392.

[8] Bachir Bekka, Pierre de la Harpe, and Alain Valette. *Kazhdan's property (T)*, volume 11 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2008. ISBN 978-0-521-88720-5. doi: 10.1017/CBO9780511542749. URL `http://dx.doi.org/10.1017/CBO9780511542749`.

[9] A. Biś. Entropies of a semigroup of maps. *Discrete Contin. Dyn. Syst.*, 11 (2-3):639–648, 2004.

[10] A. Biś and M. Urbański. Some remarks on the topological entropy of a semigroup of continuous maps. *Cubo*, 2(2):63–71, 2006.

[11] L. Bowen. Invariant random subgroups of the free group. http://arxiv.org/abs/1204.5939, 2013.

[12] L. Bowen. Examples in sofic entropy theory. preprint, 2014.

[13] L. Bowen, R. Grigorchuk, and R. Kravchenko. Invariant random subgroups of lamplighter groups. preprint, http://arxiv.org/abs/1206.6780, 2013.

[14] Lewis Bowen. Measure conjugacy invariants for actions of countable sofic groups. *J. Amer. Math. Soc.*, 23(1):217–245, 2010. ISSN 0894-0347. doi: 10.1090/S0894-0347-09-00637-7. URL `http://dx.doi.org/10.1090/S0894-0347-09-00637-7`.

[15] Lewis Bowen. The ergodic theory of free group actions: entropy and the f-invariant. *Groups Geom. Dyn.*, 4(3):419–432, 2010. ISSN 1661-7207. doi: 10.4171/GGD/89. URL `http://dx.doi.org/10.4171/GGD/89`.

[16] Lewis Bowen. Entropy for expansive algebraic actions of residually finite groups. *Ergodic Theory Dynam. Systems*, 31(3):703–718, 2011. ISSN 0143-3857. doi: 10.1017/S0143385710000179. URL `http://dx.doi.org/10.1017/S0143385710000179`.

[17] Lewis Bowen. Sofic entropy and amenable groups. *Ergodic Theory Dynam. Systems*, 32(2):427–466, 2012. ISSN 0143-3857. doi: 10.1017/S0143385711000253. URL `http://dx.doi.org/10.1017/S0143385711000253`.

[18] Lewis Bowen. Entropy theory for sofic groupoids I: The foundations. *J. Anal. Math.*, 124:149–233, 2014. ISSN 0021-7670. doi: 10.1007/s11854-014-0030-9. URL `http://dx.doi.org/10.1007/s11854-014-0030-9`.

[19] Lewis Bowen. Random walks on random coset spaces with applications to Furstenberg entropy. *Invent. Math.*, 196(2):485–510, 2014. ISSN 0020-9910. doi: 10.1007/s00222-013-0473-0. URL `http://dx.doi.org/10.1007/s00222-013-0473-0`.

[20] Lewis Bowen and Yonatan Gutman. A Juzvinskii addition theorem for finitely generated free group actions. *Ergodic Theory Dynam. Systems*, 34(1):95–109, 2014. ISSN 0143-3857. doi: 10.1017/etds.2012.126. URL `http://dx.doi.org/10.1017/etds.2012.126`.

[21] Lewis Bowen, Yair Hartman, and Omer Tamuz. Property (T) and the Furstenberg entropy of nonsingular actions. *Proc. Amer. Math. Soc.*, 144(1):31–39, 2016. ISSN 0002-9939. doi: 10.1090/proc/12685. URL `http://dx.doi.org/10.1090/proc/12685`.

[22] Lewis Phylip Bowen. A measure-conjugacy invariant for free group actions. *Ann. of Math. (2)*, 171(2):1387–1400, 2010. ISSN 0003-486X. doi: 10.4007/annals.2010.171.1387. URL `http://dx.doi.org/10.4007/annals.2010.171.1387`.

[23] Peter Burton. Topology and convexity in the space of actions modulo weak equivalence. *To appear in Ergodic Theory and Dynamical Systems*, 2015.

Available at http://arxiv.org/abs/1501.04079.
This paper is entirely my own work.

[24] Peter Burton and Alexander S. Kechris. Invariant random subgroups and action versus representation maximality. *To appear in Proceedings of the American Mathematical Society*, 2016. Available at http://arxiv.org/abs/1608.01331. Both authors contributed equally to this project.

[25] V. Capraro and T. Fritz. On the axiomatization of convex subsets of Banach spaces. http://arxiv.org/abs/1105.1270, 2011.

[26] Valerio Capraro and Martino Lupini. *Introduction to sofic and hyperlinear groups and Connes' embedding conjecture*, volume 2136 of *Lecture Notes in Mathematics*. Springer, Cham, 2015. ISBN 978-3-319-19332-8; 978-3-319-19333-5. doi: 10.1007/978-3-319-19333-5. URL `http://dx.doi.org/10.1007/978-3-319-19333-5`. With an appendix by Vladimir Pestov.

[27] Nhan-Phu Chung. Topological pressure and the variational principle for actions of sofic groups. *Ergodic Theory Dynam. Systems*, 33(5):1363–1390, 2013. ISSN 0143-3857. doi: 10.1017/S0143385712000429. URL `http://dx.doi.org/10.1017/S0143385712000429`.

[28] C.T. Conley and A.S.. Kechris. Measurable chromatic and independence numbers for ergodic graphs and group actions. *Groups, Geometry and Dynamics*, 7:127–180, 2013.

[29] C.T. Conley, A.S. Kechris, and R.D. Tucker-Drob. Ultraproducts of measure preserving actions and graph combinatorics. *Ergodic Theory and Dynamical Systems*, 33:334–374, 2013.

[30] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006. ISBN 978-0-471-24195-9; 0-471-24195-4.

[31] Alexandre Danilenko. Furstenberg entropy values for nonsingular actions of groups without property (T),. Preprint. Available at arXiv:1512.05139, 2015.

[32] A. Dooley, V. Golodets, D. Rudolph, and S. Sinel'shchikov. Non-Bernoulli systems with completely positive entropy. *Ergodic Theory and Dynamical Systems*, 28:87–124, 2007.

[33] Tomasz Downarowicz. *Entropy in dynamical systems*, volume 18 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2011. ISBN 978-0-521-88885-1. doi: 10.1017/CBO9780511976155. URL `http://dx.doi.org/10.1017/CBO9780511976155`.

[34] A. Dudko and R. Grigorchuk. On spectra of Koopman, groupoid and quasi-regular representations. preprint.

[35] A. Eisenmann and Y. Glasner. Generic IRS in free groups, after Bowen. preprint, http://arxiv.org/abs/1406.1261, 2014.

[36] Matthew Foreman and Benjamin Weiss. An anti-classification theorem for ergodic measure preserving transformations. *J. Eur. Math. Soc. (JEMS)*, 6(3): 277–292, 2004. ISSN 1435-9855. URL `http://link.springer.de/cgi/linkref?issn=1435-9855&year=04&volume=6&page=277`.

[37] T. Fritz. Convex spaces I: definition and examples. preprint, http://arxiv.org/pdf/0903.5522.pdf, 2009.

[38] Harry Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963. ISSN 0002-9947.

[39] Damien Gaboriau and Brandon Seward. Cost, ℓ^2-Betti numbers and the sofic entropy of some algebraic actions. Submitted. Available at http://arxiv.org/abs/1509.02482, 2015.

[40] E. Ghys, R. Langevin, and P. Walczak. Entropie géométrique des feuilletague. *Acta Math.*, 160(1-2):105–142, 1988.

[41] E. Glasner and B. Weiss. The topological Rokhlin property and topological entropy. *Amer. J. Math.*, 123(6):917–935, 2001.

[42] Eli Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003. ISBN 0-8218-3372-3. doi: 10.1090/surv/101. URL `http://dx.doi.org/10.1090/surv/101`.

[43] V. Golodets and S. Sinel'shchikov. Completely positive entropy and non-Bernoullicity for transformation groups. *Colloq. Math.*, 84/85:421–429, 2000.

[44] M. Gromov. Endomorphisms of symbolic algebraic varieties. *J. Eur. Math. Soc.*, 2:109–197, 1999.

[45] Paul R. Halmos. *Lectures on ergodic theory*. Publications of the Mathematical Society of Japan, no. 3. The Mathematical Society of Japan, 1956.

[46] Yair Hartman and Omer Tamuz. Furstenberg entropy realizations for virtually free groups and lamplighter groups. *J. Anal. Math.*, 126:227–257, 2015. ISSN 0021-7670. doi: 10.1007/s11854-015-0016-2. URL `http://dx.doi.org/10.1007/s11854-015-0016-2`.

[47] Ben Hayes. Sofic entropy of Gaussian actions. *To appear in Ergodic Theory and Dynamical Systems*, 2015. Available at http://arxiv.org/abs/1509.07835.

[48] Ben Hayes. Polish models and sofic entropy. *To appear in Journal of the Institute of Mathematics Jussieu*, 2015. Available at arXiv:1411.1510.

[49] Ben Hayes. Fuglede–Kadison Determinants and Sofic Entropy. *Geom. Funct. Anal.*, 26(2):520–606, 2016. ISSN 1016-443X. doi: 10.1007/s00039-016-0370-y. URL http://dx.doi.org/10.1007/s00039-016-0370-y.

[50] Ben Hayes. Mixing and spectral gap relative to Pinsker factors for sofic groups. *To appear in the Proceedings in honor of Vaughan F. R. Jones' 60th birthday conferences*, 2016. Available at http://arxiv.org/abs/1509.07839.

[51] Ben Hayes. Relative entropy and the Pinsker product formula for sofic groups. Preprint. Available at https://arxiv.org/abs/1605.01747, 2016.

[52] V. A. Kaĭmanovich and A. M. Vershik. Random walks on discrete groups: boundary and entropy. *Ann. Probab.*, 11(3):457–490, 1983. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(198308)11:3<457:RWODGB>2.0.CO;2-5&origin=MSN.

[53] Alexander S. Kechris. *Global aspects of ergodic group actions*, volume 160 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2010. ISBN 978-0-8218-4894-4. doi: 10.1090/surv/160. URL http://dx.doi.org/10.1090/surv/160.

[54] Alexander S. Kechris. Weak containment in the space of actions of a free group. *Israel J. Math.*, 189:461–507, 2012. ISSN 0021-2172. doi: 10.1007/s11856-011-0182-6. URL http://dx.doi.org/10.1007/s11856-011-0182-6.

[55] Alexander S. Kechris and Todor Tsankov. Amenable actions and almost invariant sets. *Proc. Amer. Math. Soc.*, 136(2):687–697 (electronic), 2008. ISSN 0002-9939. doi: 10.1090/S0002-9939-07-09116-2. URL http://dx.doi.org/10.1090/S0002-9939-07-09116-2.

[56] A.S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer, 1995.

[57] D. Kerr and H. Li. Bernoulli actions and infinite entropy. *Groups Geom. Dyn.*, 5:663–672, 2011.

[58] David Kerr. Sofic measure entropy via finite partitions. *Groups Geom. Dyn.*, 7(3):617–632, 2013. ISSN 1661-7207. doi: 10.4171/GGD/200. URL http://dx.doi.org/10.4171/GGD/200.

[59] David Kerr. Bernoulli actions of sofic groups have completely positive entropy. *Israel J. Math.*, 202(1):461–474, 2014. ISSN 0021-2172. doi: 10.1007/s11856-014-1077-0. URL http://dx.doi.org/10.1007/s11856-014-1077-0.

[60] David Kerr and Hanfeng Li. Bernoulli actions and infinite entropy. *Groups Geom. Dyn.*, 5(3):663–672, 2011. ISSN 1661-7207. doi: 10.4171/GGD/142. URL http://dx.doi.org/10.4171/GGD/142.

[61] David Kerr and Hanfeng Li. Entropy and the variational principle for actions of sofic groups. *Invent. Math.*, 186(3):501–558, 2011. ISSN 0020-9910. doi: 10.1007/s00222-011-0324-9. URL `http://dx.doi.org/10.1007/s00222-011-0324-9`.

[62] David Kerr and Hanfeng Li. Combinatorial independence and sofic entropy. *Commun. Math. Stat.*, 1(2):213–257, 2013. ISSN 2194-6701.

[63] David Kerr and Hanfeng Li. Soficity, amenability, and dynamical entropy. *Amer. J. Math.*, 135(3):721–761, 2013. ISSN 0002-9327. doi: 10.1353/ajm.2013.0024. URL `http://dx.doi.org/10.1353/ajm.2013.0024`.

[64] A. N. Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk SSSR*, 124:754–755, 1959. ISSN 0002-3264.

[65] Amos Nevo. The spectral theory of amenable actions and invariants of discrete groups. *Geom. Dedicata*, 100:187–218, 2003. ISSN 0046-5755. doi: 10.1023/A:1025839828396. URL `http://dx.doi.org/10.1023/A:1025839828396`.

[66] Donald Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4:337–352, 1970. ISSN 0001-8708.

[67] Donald S. Ornstein and Benjamin Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.*, 48:1–141, 1987. ISSN 0021-7670. doi: 10.1007/BF02790325. URL `http://dx.doi.org/10.1007/BF02790325`.

[68] Vladimir G. Pestov. Hyperlinear and sofic groups: a brief guide. *Bull. Symbolic Logic*, 14(4):449–480, 2008. ISSN 1079-8986. doi: 10.2178/bsl/1231081461. URL `http://dx.doi.org/10.2178/bsl/1231081461`.

[69] Daniel J. Rudolph and Benjamin Weiss. Entropy and mixing for amenable group actions. *Ann. of Math. (2)*, 151(3):1119–1150, 2000. ISSN 0003-486X. doi: 10.2307/121130. URL `http://dx.doi.org/10.2307/121130`.

[70] B. Seward. Ergodic actions of countable groups and finite generating partitions. preprint, http://arxiv.org/pdf/1206.6005.pdf, 2014.

[71] B. Seward. Krieger's finite generator theorem for ergodic actions of countable groups I. preprint, http://arxiv.org/abs/1405.3604, 2014.

[72] B. Seward. Krieger's finite generator theorem for ergodic actions of countable groups II. preprint, http://arxiv.org/abs/1501.03367, 2014.

[73] Ja. Sinaĭ. On the concept of entropy for a dynamic system. *Dokl. Akad. Nauk SSSR*, 124:768–771, 1959. ISSN 0002-3264.

[74] Robin D. Tucker-Drob. Weak equivalence and non-classifiability of measure preserving actions. *Ergodic Theory Dynam. Systems*, 35(1):293–336, 2015. ISSN 0143-3857. doi: 10.1017/etds.2013.40. URL `http://dx.doi.org/10.1017/etds.2013.40`.

[75] B. Weiss. Sofic groups and dynamical systems. *Ergodic theory and harmonic analysis*, 62:350–359, 2000.

[76] Benjamin Weiss. Actions of amenable groups. In *Topics in dynamics and ergodic theory*, volume 310 of *London Math. Soc. Lecture Note Ser.*, pages 226–262. Cambridge Univ. Press, Cambridge, 2003. doi: 10.1017/CBO9780511546716.012. URL `http://dx.doi.org/10.1017/CBO9780511546716.012`.

[77] Benjamin Weiss. Entropy and actions of sofic groups. *Discrete Contin. Dyn. Syst. Ser. B*, 20(10):3375–3383, 2015. ISSN 1531-3492. doi: 10.3934/dcdsb.2015.20.3375. URL `http://dx.doi.org/10.3934/dcdsb.2015.20.3375`.

[78] Guohua Zhang. Local variational principle concerning entropy of sofic group action. *J. Funct. Anal.*, 262(4):1954–1985, 2012. ISSN 0022-1236. doi: 10.1016/j.jfa.2011.11.029. URL `http://dx.doi.org/10.1016/j.jfa.2011.11.029`.