

ISOMORPHISM AND EMBEDDING OF BOREL SYSTEMS ON FULL SETS

MICHAEL HOCHMAN

ABSTRACT. A Borel system consists of a measurable automorphism of a standard Borel space. We consider Borel embeddings and isomorphisms between such systems modulo null sets, i.e. sets which have measure zero for every invariant probability measure. For every $t > 0$ we show that in this category there exists a unique free Borel system (Y, S) which is strictly t -universal in the sense that all invariant measures on Y have entropy $< t$, and if (X, T) is another free system obeying the same entropy condition then X embeds into Y off a null set. One gets a strictly t -universal system from mixing shifts of finite type of entropy $\geq t$ by removing the periodic points and “restricting” to the part of the system of entropy $< t$.

As a consequence, after removing their periodic points the systems in the following classes of are completely classified by entropy up to Borel isomorphism off null sets: mixing shifts of finite type, mixing positive-recurrent countable state Markov chains, mixing sofic shifts, beta shifts, synchronized subshifts, and axiom-A diffeomorphisms. In particular any two equal-entropy systems from these classes are entropy conjugate in the sense of Buzzi, answering a question of Boyle, Buzzi and Gomez.

1. INTRODUCTION

1.1. Background. Motivated by the conjugacy problem for symbolic covers of partially hyperbolic systems, in the late 1990’s Buzzi [7, 6] introduced the notion of an entropy conjugacy between topological dynamical systems: X, Y are entropy conjugate if there is a measurable, shift-commuting bijection $X \setminus X_0 \rightarrow Y \setminus Y_0$, where $X_0 \subseteq X, Y_0 \subseteq Y$ are invariant Borel sets which have measure zero for every invariant probability measure of sufficiently large entropy.¹ Entropy conjugacy for mixing shifts of finite type (SFTs) of the same entropy follows from classical work of Adler and Marcus on almost topological conjugacy [1], and the so-called subshifts of quasi-finite-type shifts (QFTs) were classified up to entropy conjugacy by Buzzi [7]; this class is of interest because it includes symbolic covers of many examples, e.g piecewise entropy-expanding maps and multidimensional beta-shifts [7]. Finally, Boyle, Buzzi and Gomez classified the strong positive recurrent (SPR) countable

Research supported by NSF grant 0901534.

¹A notion with the same name and similar (but not identical) definition was introduced earlier by R. Bowen for shifts of finite type [3]

state Markov shifts up to entropy conjugacy, showing that entropy and period are complete invariants. Their proof shows that the entropy conjugacy can be realized as an almost isomorphism [4].

The purpose of this note is to show that on rather general grounds, a stronger equivalence than entropy conjugacy holds for quite a wide class of systems.

1.2. Borel systems and full sets. Our setting is that of Borel dynamics, whose basic definitions we summarize in the next few paragraphs; for further detail see [9, 15]. Recall that a *standard Borel space* (X, \mathcal{F}) is a set X and σ -algebra \mathcal{F} which arises as the σ -algebra of Borel sets for some complete, separable metric on X . Up to isomorphism (i.e. measurable bijections) there are only three standard Borel spaces: those which are finite, those which are countable, and those which are isomorphic to $[0, 1]$ with the Borel σ -algebra. We shall always assume the last of these. We suppress \mathcal{F} in our notation and assume implicitly that all sets are measurable unless otherwise stated. Similarly, all our measures are defined on the given σ -algebra.

A *Borel system* (X, T) consists of an uncountable standard Borel space X and a bijection $T : X \rightarrow X$ such that T, T^{-1} are measurable. T is *free* if it contains no periodic points, i.e. $T^n x \neq x$ for all $x \in X$ and all $n \neq 0$. By [15], the set of periodic points is a Borel set, and if the complement of the periodic points is uncountable then restricting T to it gives rise to a Borel system which we call the *free part* of (X, T) .

Two Borel systems $(X, T), (Y, S)$ are *isomorphic* if there is a Borel isomorphism $\varphi : X \rightarrow Y$ such that $\varphi T = S \varphi$. If instead φ is only a Borel injection and $\varphi T = S \varphi$ we call it a *Borel embedding*, and say that (X, T) embeds into (Y, S) . We note that the image of a Borel injection $\varphi : X \rightarrow Y$ is a Borel subset of Y , and the partially defined inverse $\varphi^{-1}|_{\varphi(X)}$ is measurable.

For a Borel system (X, T) , let

$$\mathcal{E}(X, T) = \{\text{ergodic } T\text{-invariant probability measures on } X\}$$

We say that a Borel set $X_0 \subseteq X$ is *universally null* if $\mu(X_0) = 0$ for all $\mu \in \mathcal{E}(X, T)$. The set $\mathcal{E}(X, T)$ may be empty, but if it is not then the universally null sets form a σ -ideal. A set is *full* if its complement is universally null. We say that Borel systems $(X, T), (Y, S)$ are Borel isomorphic on full sets if there are invariant, full Borel subsets of each on which the restricted maps are isomorphic. We say that (X, T) embeds into (Y, S) on a full set if there is a full, invariant set in X which embeds, with the restricted actions, into Y .

The Gurevich entropy [11] of a Borel system is defined by

$$h(X, T) = \sup_{\mu \in \mathcal{E}(X, T)} h(T, \mu)$$

or $-\infty$ if $\mathcal{E}(X, T) = \emptyset$. When the transformation is clear from the context, we abbreviate the notation to $\mathcal{E}(X)$, $h(X)$, $h(\mu)$ etc.

1.3. Embedding and universality. There are two obvious obstructions to embedding on full sets. One is the entropy of invariant measures: clearly one cannot hope to embed (X, T) into (Y, S) if $h(X) > h(Y)$. The other obstruction comes as usual from periodic points. Every full set includes all periodic points, since a periodic point has positive mass with respect to the unique invariant probability measure on its orbit. Hence an embedding of X into Y requires that for each k the number of periodic points of period k in Y be no less than the corresponding count in X .

Since the periodic points in a Borel system form an invariant Borel set [15] one may deal with this set separately; indeed, this part of a system is classified up to Borel isomorphism by the cardinality of the set of points of period each k . For this reason we shall mostly be interested in free actions, or the free parts of non-free actions.

Definition 1.1. A Borel system (X, T) is *t-universal* if any free Borel system (Y, S) of entropy $< t$ embeds into X on a full set.

According to this definition (X, T) may have measures of entropy $\geq t$ and periodic points. A stronger property is given in the following definition.

Definition 1.2. A *t-universal* system which is free and has no ergodic measures of entropy $\geq t$ is *strictly t-universal*.

It is convenient to introduce one more definition:

Definition 1.3. A *t-slice* of a Borel system X is a Borel subset $X_t \subseteq X$ such that for $\mu \in \mathcal{E}(X, T)$,

$$\mu(X_t) = \begin{cases} 1 & h(T, \mu) < t \\ 0 & \text{otherwise} \end{cases}$$

A *free t-slice* is a *t-slice* of the free part of X .

Note that, up to universally null sets, a Borel system has a unique *t-slice*, and we therefore refer to it as *the t-slice*. Similarly the free *t-slice* is unique up to null sets in the free part of X . We show that *t-slices* exist in section 4.1.

Elementary considerations and Cantor-Bernstein type arguments now establish the following basic properties of universal and strictly universal systems.

Proposition 1.4. *Let $t > 0$.*

- (1) *Any two strictly t-universal systems are isomorphic on full sets.*
- (2) *The free t-slice of a t-universal system is strictly t-universal.*

- (3) *If (X, T) is free, supports no ergodic measures of entropy $\geq t$, and contains an s -universal subsystem for every $s < t$, then (X, T) is strictly t -universal.*

1.4. Main results. We next turn to our main results. The next theorem shows that t -universal (and hence strictly t -universal) systems exist. See section for the definition of mixing shifts of finite type 2, and section 3 for discussion of generic points.

Theorem 1.5. *A mixing shift of finite type (X, T) is $h(X)$ -universal, and for every $t \leq h(X)$, the free t -slice of X is strictly t -universal.*

Note that the Krieger generator theorem states that for any non-atomic invariant probability measure μ on a Borel system Y , if X is a mixing SFT and $h(\mu) < h(X)$ then there is a set of full measure for μ which embeds into X . In fact this is a Borel theorem in the sense that, given a free Borel system (Y, S) with $h(Y) < h(X)$, the proof of the generator theorem produces a Borel map from a full set in Y into X , which, for every $\mu \in \mathcal{E}(Y, S)$, is an injection on a set of μ -measure one (a similar observation is made in [9] concerning the Jewet-Krieger theorem). The main innovation in the proof of Theorem 1.5 is a coding argument which ensures that for distinct ergodic measures μ, ν on Y the images of μ, ν under the embedding are disjoint.

An immediate consequence is:

Theorem 1.6. *For X in any of the following classes and $t_0 \leq h(X, T)$ then the t_0 -slice of X is t_0 -universal:*

- *Mixing sofic shifts,*
- *Mixing countable state Markov shifts,*
- *Axiom-A diffeomorphisms,*
- *Mixing synchronized subshifts,*
- *The natural extensions of β -shifts,*
- *Intrinsically ergodic mixing shifts of quasi-finite type.*

Indeed, members of these classes contain embedded mixing SFTs of arbitrarily large entropy; the theorem then follows from the previous one and Proposition 1.4 (3).

If a free t -universal system has a unique measure of maximal entropy then the space can be partitioned into a set supporting only this measure and no other invariant probability measures, and the remainder, which is strictly t -universal. Thus in this case isomorphism of the original system on a full set is determined by the isomorphism type (in the sense of ergodic theory) of the measure of maximal entropy. In particular, when two such systems have measures of maximal entropy which are Bernoulli they are isomorphic on full sets. Using the fact that systems

in the classes above, with one exception, support a unique measure of maximal entropy which is Bernoulli, we have:

Theorem 1.7. *Let X, Y be any two systems drawn from the classes listed mentioned in Theorem 1.6, excluding Markov shifts which are not positive recurrent. If $h(X) = h(Y)$ then the free parts of X, Y are isomorphic on full sets, and in particular X, Y are entropy conjugate.*

Similarly, positive recurrent Markov shifts are classified up to isomorphism on full sets by entropy and period; and similarly the class of recurrent non-positive recurrent Markov shifts.

This answers Question 6.5 in [4], showing that the SPR property of Markov shifts is not an invariant of entropy-conjugacy, and whether positive recurrent Markov shifts of equal entropy and period be entropy-conjugate.

Quasi-finite-type (QFTs) systems were introduced in [7], where they were shown to be Borel isomorphic to the union of finitely many strong positive recurrent Markov shifts together with a Borel system of lower entropy [7]. Let us say that a dynamical system is entropy-mixing if it has a mixing measure of maximal entropy.

Theorem 1.8. *The free part of entropy-mixing quasi-finite-type subshifts are classified up to isomorphism on full subsets by the number and periods of their measures of maximal entropy.*

The hypothesis of entropy-mixing is necessary: Consider a positive entropy SFT X of period 2, and let X' be the union of X with a mixing SFT of lower entropy. Then X, X' are QFTs whose (unique) maximal measures are isomorphic, but X' supports a mixing measure while X does not. Therefore they cannot be isomorphic on full sets.

1.5. Open problems. We end this introduction with several questions which arise from this work. A natural question is whether the isomorphisms we construct can be given any amount of continuity when the original systems carry a topology. Let us suggest the following concrete problem, which we have been unable to resolve:

Problem 1.9. Let X, Y be mixing SFTs on finite alphabets, and $h(X) = h(Y)$. Let X', Y' denote the sets obtained by removing all periodic points from X, Y . Is there a topological conjugacy between the (non-compact) systems X' and Y' ?

We do not know the answer even if we require only that X', Y' be full sets. Note that by the Adler-Marcus theorem on almost topological conjugacy [1], one can find a topological isomorphism between subsets of X'', Y'' which are dense G_δ 's, contain all points with dense orbits, and have measure one for every invariant measure of sufficiently large entropy. Also, any closed subshift of X without periodic points

can be topologically embedded in Y , and visa versa. However the known methods do not seem to extend all the way “down” to the very low complexity part of the systems.

In another direction, one may ask for a purely Borel formulation. While our methods produce Borel maps, the sets on which they are defined are “large” only from an ergodic-theory point of view. A more natural notion in the Borel category is the following one, introduced by Shelah and Weiss [14]. Let (X, T) be a Borel system. A Borel set $W \subseteq X$ is wandering if its iterates $T^n W$ are all disjoint. Let \mathcal{W} denote the set of countable unions of wandering sets, and all Borel subsets of such unions. If $X \in \mathcal{W}$ we say that X is purely wandering. If this is not the case then \mathcal{W} is a σ -ideal, and it is natural to say that a set is Borel-null if it belongs to \mathcal{W} , and Borel-full if its complement is in \mathcal{W} . Thus we arrive at the following question:

Problem 1.10. Let (X, T) be a Borel system, Y a mixing SFT, and suppose that $h(T, \mu) < h(Y)$ for all $\mu \in \mathcal{E}(X, T)$. Is there an invariant Borel-full subset $X' \subseteq X$ such that (X', T) can be embedded by a Borel map into (Y, S) ?

Like the universally null sets that we have been using, Borel-null sets have a characterization in terms of measures: $A \in \mathcal{W}$ if and only if $\mu(A) = 0$ for every conservative invariant σ -finite measure μ (this is proved in [14] with “non-singular” in place of “invariant”, but from the methods in [14] the version given here follows easily). It is also well known that if μ is a σ -finite infinite invariant measure on (X, T) and Y is a mixing SFT then there is a set of μ -measure 1 which embeds in a Borel way into Y . Thus one might hope that Problem 1.10 can be answered by treating each measure individually. However, unlike invariant probability measures, for infinite measures it is not possible to partition X into disjoint sets each of which supports a single σ -finite measure. Thus it appears that new ideas will be needed to answer Problem 1.10.

This problem is essentially equivalent to one posed by Benjamin Weiss in [15], asking whether a Borel action without invariant probability measures always has a 2-set generator modulo wandering sets. Indeed, given X, Y as in the problem, we may embed a full set in X using Theorem 1.5; the complement of this full set has no invariant probability measure, and an answer to Weiss’ question would probably allow one to embed it in Y .

A positive answer to Problem 1.10 would lead to a Borel isomorphism, modulo wandering sets, between mixing SFTs (and, more generally, SPR Markov shift) of the same entropy. Following this work, Boyle, Buzzi and Gomez have shown that this conclusion at least is true [5].

In the next two sections we present some background and define generic points in our context. In Section 4 we establish some properties of t -universal systems and show how to deduce our results from Theorem 1.5. In Section 5 we prove Theorem 1.5.

Acknowledgement. My thanks to Mike Boyle for some very interesting discussions and many useful suggestions on the presentation of this paper. This work was supported by NSF grant 0901534.

2. SHIFTS OF FINITE TYPE

We briefly recall the definitions of SFTs needed for Theorem 1.6. Let $G = (V, E)$ be a directed graph on finite vertex set V . Endow $E^{\mathbb{Z}}$ with the product topology, which is compact and metrizable. The shift of finite type (SFT) defined by G is the subshift $X_G \subseteq E^{\mathbb{Z}}$ consisting of bi-infinite paths through the graph, that is

$$X_G = \{x \in E^{\mathbb{Z}} : \text{initial vertex of } x_{i+1} = \text{terminal vertex of } x_i\}$$

The shift transformation $S : X_G \rightarrow X_G$ is defined by

$$(Sx)_i = x_{i+1}$$

If the graph G is strongly irreducible, i.e. there is a directed path between every pair of vertices, then X_G is said to be irreducible; this is equivalent to the existence of a dense orbit under the shift. If G is also aperiodic, i.e. there is some N so that for every $v \in V$ and $n > N$ there is a path from v to itself of length n , then X_G is aperiodic in the sense that there is no factor map from X_G to a periodic orbit. An SFT is topologically mixing if and only if it is irreducible and aperiodic.

Shifts of finite type have a well developed theory. We mention the following:

Theorem. *Let X be a mixing SFT X . Then*

- *X has a unique invariant probability measure of maximal entropy which is isomorphic (in the ergodic category) to a Bernoulli process.*
- *For any ergodic system (Y, T, ν) of entropy less than $h(X)$ there is a shift-invariant measure on X isomorphic to (Y, T, σ) .*

The only fact we shall use about the systems in Theorem 1.6 is that they have SFT subsystems of entropy arbitrarily close to the entropy of the system, and in most cases (except non positive recurrent countable state Markov shifts) a unique measure of maximal entropy which is ergodic-theoretically Bernoulli. For further information see [12] (countable state Markov shifts); [13] (sofic shifts and synchronized systems); [3] (axiom-A diffeomorphisms); [2] (beta shifts); [7] (subshifts of quasi finite type).

3. SYMBOLIC REPRESENTATION AND GENERIC POINTS

As is often the case in Borel dynamics, it is convenient to introduce a topology to our systems. Most convenient for us is the possibility of modeling our systems as shift-invariant subsets of shift spaces over countable alphabets. Endow $\mathbb{N}^{\mathbb{Z}}$ with the product topology and let T denote the shift map $(Tx)_i = x_{i+1}$.

Theorem (Weiss [15]). *Any free Borel system can be embedded on a full set into $(\mathbb{N}^{\mathbb{Z}}, T)$, where T is the shift.*

Next, it will be useful to partition our systems into sets which support a unique invariant probability measure. More precisely, we shall describe the set G of generic points of $\mathbb{N}^{\mathbb{Z}}$ and the corresponding partition of G according to ergodic probability measures. These definitions are standard in topological dynamics but require a little care since the space $\mathbb{N}^{\mathbb{Z}}$ in our setting is not compact.

Let \mathcal{C} be the algebra of sets in $\mathbb{N}^{\mathbb{Z}}$ generated by the cylinder sets. We say that $x \in \mathbb{N}^{\mathbb{Z}}$ has limiting frequencies if every $C \in \mathcal{C}$, the limit

$$(3.1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_C(S^n x)$$

exists. The set of points with limiting frequencies is Borel.

If $\mu \in \mathcal{E}(\mathbb{N}^{\mathbb{Z}})$ then for μ -a.e. x the limit in (3.1) exists and equals $\mu(C)$. We denote by G_μ the set of points which satisfy this condition for every set $C \in \mathcal{C}$. Thus $\mu(G_\mu) = 1$ and the sets $G_\mu, \mu \in \mathcal{E}(\mathbb{N}^{\mathbb{Z}})$ are pairwise disjoint.

We claim that $G = \bigcup_{\mu \in \mathcal{E}(\mathbb{N}^{\mathbb{Z}})} G_\mu$ is Borel. To this end, let $Z_N = \{1, \dots, N+1\}^{\mathbb{Z}}$ and let $\pi_N : \mathbb{N}^{\mathbb{Z}} \rightarrow Z_N$ denote the factor map defined componentwise by mapping the symbol $i \in \mathbb{N}$ to $\min\{i, N+1\}$. Clearly if x has limiting frequencies then for every N , the point $\pi_N(x) \in Z_N$ satisfies the same condition for every closed and open set $C \subseteq Z_N$. Then, since Z_N is compact, there is a unique invariant probability measure μ_N on Z_N such that limit (3.1) is equal to $\mu_N(C)$.

We say that $x \in \mathbb{N}^{\mathbb{Z}}$ is a *generic point* if it has limiting frequencies, each μ_N is ergodic, and

$$\lim_{N \rightarrow \infty} \mu_N([N+1]) = 0$$

Here $[N+1]$ is the cylinder set $\{z \in Z_N : z_0 = N+1\}$. The set of generic points is seen to be Borel. We conclude by showing that the set of generic points is precisely G . Indeed, clearly $\bigcup G_\mu$ consists only of generic points. Conversely, if x is generic then $x \in G_\mu$ for the measure μ which is the inverse limit of the measures μ_N .

4. EMBEDDING AND UNIVERSALITY

4.1. Generalities. This section fills in some details which were omitted from Section 1.3. We first show below that t -slices exist:

Lemma 4.1. *Let (X, T) be a Borel system. Then (X, T) has a t -slice.*

Proof. We may assume that $X \subseteq \mathbb{N}^{\mathbb{Z}}$ and T is the shift. Let G be the Borel set of generic points, so that $G \cap X$ is a full set in X . For $x \in G \cap X$ one obtains a measure μ_x in a Borel way so that $x \in G_{\mu_x}$. Since the entropy $h(T, \mu_x)$ is a Borel function of μ_x , we may form the Borel set $X_t = \{t \in G : h(T, \mu_x) < t\}$. Clearly this is a t -slice.

We shall need the following to prove Proposition 1.4: □

Lemma 4.2. *Let (X, T) and (Y, S) be Borel systems. Suppose that for each $s < t$ there is an embedding $\varphi_s : X_s \hookrightarrow Y$ of the s -slice $X_s \subseteq X$ into Y . Then there is an embedding $\varphi : X_t \hookrightarrow Y$.*

Proof. Let $s_n \nearrow h(X, T)$, $s_n = 1, 2, 3, \dots$. Set $X^{(1)} = X_{s_1}$ and $X^{(n)} = X_{s_n} \setminus X^{(n-1)}$, so $X^{(n)}$ are disjoint. Clearly $X_t = \bigcup X^{(n)}$ and, defining $\varphi|_{X^{(n)}} = \varphi_{s_n}|_{X^{(n)}}$, we have $\varphi : X_t \hookrightarrow Y$ as desired. □

An immediate consequence is:

Corollary 4.3. *If (X, T) is t -universal and (Y, S) is a free system whose invariant measures have entropy $< t$, then Y embeds into X on a full set.*

We can now prove Proposition 1.4:

Proof. (of Proposition 1.4) (1) Let (X, T) and (Y, S) be strictly t -universal. Hence by the previous lemma there are full invariant sets $X_0 \subseteq X$ and $Y_0 \subseteq Y$ and Borel embeddings $\varphi : X_0 \hookrightarrow Y$ and $\psi : Y_0 \hookrightarrow X$. We note that φ preserves universally null sets between X_0 and $\varphi(X_0)$ and similarly for ψ , and henceforth we shall work modulo universally null sets.

We proceed precisely as in the Cantor-Bernstein theorem in the category of sets. Form the map $\tau = \varphi\psi$ and let $Y_1 = Y_0 \setminus \varphi(X_0)$ (if this is a universally null set we are done, since φ is the desired isomorphism). Consider the sets $Y_n = \tau^n Y_1$. One easily verifies that they are pairwise disjoint, as are $X_n = \psi Y_n$. Let $Y_\infty = Y_0 \setminus \bigcup_{n=1}^{\infty} Y_n$ and $X_\infty = X_0 \setminus \bigcup_{n=1}^{\infty} X_n$. Then $\varphi : X_\infty \rightarrow Y_\infty$ is bijective and $\psi^{-1} : \bigcup_{n=1}^{\infty} X_n \rightarrow \bigcup_{n=1}^{\infty} Y_n$ is bijective. These sets are complementary up to a universally null set, and one verifies that all sets involved are invariant; hence this is the desired isomorphism.

(2) is trivial.

(3) follows immediately from Lemma 4.2. □

4.2. Some proofs. In the next section we prove Theorem 1.5, showing that a mixing SFT Y is $h(Y)$ -universal. Assuming this we have:

Proof. (of Theorem 1.6). Suppose that (X, T) is a Borel system and that for $t < t_0$ there is a mixing SFT of entropy at least t embedded in X . Since mixing SFTs are universal at their entropy, it follows that X is t -universal for every $t < t_0$. By Lemma 4.2, this implies that X is t_0 -universal, and hence the t_0 -slice of the free part of X is t_0 -universal. \square

Proof. (of Theorem 1.7). Let $(X, T), (Y, S)$ be two systems of entropy h from the list in Theorem 1.7, excluding the countable Markov chains, and assume we have already removed the periodic points from them. Let μ, ν denote, respectively, the unique measures of maximal entropy on X, Y and let $X' = G_\mu \subseteq X$ and $Y' = G_\nu \subseteq Y$ be the points generic for μ, ν . Clearly X' and the h -slice $X_h \subseteq X$ are disjoint (up to null sets), and similarly Y' and Y_h . Also μ, ν are the only invariant probability measures on X', Y' , they have the same entropy as X, Y and therefore as each other, and in the classes in question μ, ν are Bernoulli. Thus from the Ornstein theory there is an isomorphism $X' \rightarrow Y'$ defined on a full set (i.e. from a set of μ -measure one to a set of ν -measure one). Also, by Theorem 1.6, X_h, Y_h are Borel isomorphic on full sets. Now the required isomorphism is defined by combining the two given isomorphisms $X_h \rightarrow Y_h$ and $X' \rightarrow Y'$, and noting that $X_h \cup X'$ and $Y_h \cup Y'$ are full.

For the class of positive recurrent Markov shifts we proceed as follows. For mixing positive-recurrent chains the proof is exactly as above, since there is a unique measure of maximal entropy and it is Bernoulli. In the case that $(X, T), (Y, S)$ are PR and both have period $p > 0$, one considers the p -th power of the shift map S . Then X, Y split, respectively, into p disjoint T^p - and S^p -invariant aperiodic PR Markov shifts of the same entropy which are permuted cyclically by T, S . Applying the above to one pair X_0, Y_0 we obtain an isomorphism $f : (X'_0, T^p) \rightarrow (Y'_0, S^p)$ defined on full sets $X'_0 \subseteq X_0$ and $Y'_0 \subseteq Y_0$. Extend f to $T^i X'_0$ by $f(T^i x) = S^i f(x)$. This is an isomorphism between the full sets $X' = \cup_{i=0}^{p-1} T^i X'_0$ and $Y' = \cup_{i=0}^{p-1} S^i Y'_0$.

The case where X, Y are recurrent but not positive recurrent is simpler, since, in the mixing case, absent a measure of maximal entropy an h -slice is already a full set, and so the result follows directly from Theorem 1.5 and proposition 1.4. When there is periodicity we pass to an appropriate power of the shift and continue as above. \square

Proof. (of Theorem 1.8). This follows from [7, Corollary 1], which implies that any entropy-mixing QFT shift of entropy h is h -universal and hence the free h -slices of any two such systems are isomorphic on full sets. As above, the isomorphism can be extended to a full set by defining it separately on sets which support the measures of maximal entropy, which are products of a Bernoulli and a periodic measure by [7, Theorem 1(1)]. \square

5. UNIVERSALITY OF MIXING SFTs

5.1. Assumptions. If (X, T) is a Borel system then by a theorem of Weiss [15], (X, T) may be embedded on a full set into $\mathbb{N}^{\mathbb{Z}}$ as a shift-invariant subset. Thus to prove Theorem 1.5 it suffices, for $t > 0$, to prove it for the free part of the t -slice of $\mathbb{N}^{\mathbb{Z}}$. We may focus our efforts on this case.

For the remainder of this paper we fix the following notation. (Y, S) is a mixing SFT on a finite alphabet Λ and $t < h(Y)$. Let and $X = X_t$ is the free part of the t -slice of $\mathbb{N}^{\mathbb{Z}}$. We wish to show that we can embed a full set of X into Y . Let $G_t = G \cap X_t$, where G is the set of generic points in $\mathbb{N}^{\mathbb{Z}}$ (see Section 3). For simplicity of notation we suppress the index t and write $G = G_t$. No ambiguity should arise since the set of all generic points will not be used again.

5.2. The Krieger generator theorem. The Krieger generator theorem (see e.g. [8]) asserts that for every $\mu \in \mathcal{E}(X)$ one can find a set of μ -measure 1 and an embedding of this set into Y (recall that according to our assumptions, $h(\mu) < t < h(Y)$ and there are no periodic points in X). There are two facts about this theorem which we shall need. One is that it is a Borel theorem, in the sense that the embedding obtained is a Borel function of the data; see e.g. [9]. More importantly, for each $\mu \in \mathcal{E}(X)$ the inverse map of the embedding is finitary (see below). This feature of the theorem follows easily from most of the standard proofs of the generator theorem; we provide a precise statement below.

Recall that, given a shift-invariant probability measure ν on $\mathbb{N}^{\mathbb{Z}}$, a finitary factor map π into Y is a factor map defined on a set of ν -measure 1 with the property that for ν -a.e. x , the symbol $\pi(x)_0$ is determined by a finite block $x|_{[-n, n]}$, where $n = n(x)$ is a measurable function of x . Such a map can be represented as a countable sequence of pairs (a, b) , with $a \in \bigcup_n \Lambda^{[-n, n]}$ and $b \in \Delta$, consisting of all pairs such that $\pi(x)_0 = b$ if $x|_{[-n, n]} = a$. The space of such sequences of pairs (whether they give a well-defined factor map of ν or not) is a Borel space with the obvious structure. We can further represent such a sequence of pairs as a sequence of 0's and 1's in an appropriate coding.

We are now ready to state the version of the Krieger generator theorem which we shall need.

Theorem 5.1. *Let t , $X_t \subseteq \mathbb{N}^{\mathbb{Z}}$ and Y be as above. There exist Borel maps $\varphi : X_t \rightarrow Y$ and $\psi : X_t \rightarrow \{0, 1\}^{\mathbb{N}}$ such that, for every $\mu \in \mathcal{E}(X_t)$, the map φ is an injective factor map $X_t \cap G_\mu \rightarrow Y$, and its inverse is μ -a.s. given by the finitary map encoded by $\psi(x)$.*

This folklore version apparently has not been stated in the literature but it follows from the standard proofs (e.g. [8, Theorem 28.1]), except that one needs to

construct Rohlin towers simultaneously for all measures in a measurable way. This can be achieved using Proposition 7.9 of [10].

It follows that the map $x \mapsto (\varphi(x), \psi(x))$ is injective on a full set in $X = X_t$. Our task, therefore, will be to construct an embedding $\varphi' : X \rightarrow Y$ which encodes both $\varphi(x)$ and $\psi(x)$ into $\varphi'(x)$, that is, φ' is shift commuting, defined on a full set, and from $\varphi'(x)$ we can recover $\varphi(x)$ and $\psi(x)$. We accomplish this by partitioning the output sequence $y = \varphi(x)$ into two subsequences, one a low density subsequence which encodes $\psi(x)$, and the complementary subsequence which encodes $\varphi(x)$. This is where we use the fact that $h(T, \mu) < t < h(Y, S)$ for all $\mu \in \mathcal{E}(X)$: this means that encoding x requires less than t bits per symbol of input, while in Y we have more than t bits available per symbol. Thus we can aside a small positive-density subsequence and still be able to encode $\varphi(x)$ in the complement.

5.3. Gaps and markers. Let Λ be the alphabet on which the subshift Y is defined. For a block $w \in \bigcup_n \Lambda^n$, let Y_w denote the set obtained from Y by removing all points containing w , which is again an SFT.

A *marker* in Y is a block w which is admissible for Y (that is, appears as a sub-block in some $y \in Y$) and such that no two occurrences of w overlap. As is well known, for every $\varepsilon > 0$ one can find arbitrarily long markers in Y with the property that $h(Y_w) > h(Y) - \varepsilon$, and such that Y_w is a mixing SFT.

Relying on this, we proceed as follows:

- Choose a marker w in Y such that the SFT

$$Y' = Y_w$$

is mixing, and

$$h' = h(Y_w) > t$$

Also, we assume that w is long enough that every symbol in Λ appears in Y' ; this is guaranteed if w is chosen so that $h(Y_w) > h(Y_\sigma)$ for all $\sigma \in \Lambda$.

Write

$$t' = \frac{1}{2}(h' + t)$$

Since Y' is a mixing SFT there exists a transition length $M \geq 1$ such that for all admissible blocks a, b in Y' and all $m \geq M$ there is a block $v \in \Lambda^m$ such that avb is admissible. Since $h(Y') > 0$, clearly Y' consists of more than one point, so we may

- Choose two distinct blocks u, v of length M in Y' .

5.4. The domain of the map $x \mapsto y$. In what follows we will be given a point $x \in G$, and we shall define a point $y = \varphi'(x)$. Our choices will depend on this point

in a Borel way. For $C \in \mathcal{C}$ we denote by $\mu(C)$ the value of the limit

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_C(S^n x)$$

and by definition of G , the function μ is (the restriction to \mathcal{C} of) an ergodic invariant measure on X . We denote this measure also by μ .

Let $*$ be a symbol which does not belong to Λ ; to construct y , we start out with $y = \dots * * * * * \dots$, and replace the $*$'s in several steps until eventually all are removed.

5.5. Marker structure. We first insert a sequence of w 's in y whose purpose is to partition y into identifiable blocks consisting of the gaps between the w 's.

- Choose a block $a \in \mathbb{N}^*$ such that $\mu([a]) > 0$ and such that the minimal distance between appearances of a in x is at least N , where

$$N > \frac{h'}{h' - t} \cdot (\ell(w) + 4M + 1)$$

Such a block a exists because x is not a periodic point. To make the choice Borel and dependent only on μ , select a to be the first such block with respect to the lexicographical ordering on \mathbb{N}^* .

- Let

$$I_1 = \{i \in \mathbb{Z} : a \text{ appears in } x \text{ at index } i\}$$

- For $i \in I_1$, set $y_i \dots y_{i+\ell(w)-1} = w$.

Since $N > \ell(w)$, the blocks from i to $i + \ell(w) - 1$ do not overlap, so the last step produces a well defined sequence of w 's in y .

Note that the block a was chosen in a manner which depends only on the the set $\{C \in \mathcal{C} : \mu(C) > 0\}$. Therefore if x' is any other point giving rise to the same function $\mu(\cdot)$, then the same block a will be chosen, so the choice is consistent on G_μ .

5.6. Coding x . Recall that, given a set $A \subseteq \mathbb{N}^{\mathbb{Z}}$, for $z \in A$ the first return time of z to A is

$$r_A(z) = \min\{n > 0 : S^n z \in A\}$$

with $r_A(z) = \infty$ if $S^n z \notin A$ for all $n > 0$. The induced map $S_A : A \rightarrow A$ is defined by

$$S_A z = S^{r_A(z)} z$$

whenever $r_A(z) < \infty$. If $\mu(A) > 0$ then by the Poincaré recurrence theorem, all powers (positive and negative) of S_A are defined at μ -a.e. point in A , and S_A preserves the measure

$$\mu_A(\cdot) = \frac{1}{\mu(A)} \mu(\cdot \cap A)$$

The entropy of (A, μ_A, S_A) is given by Abramov's formula,

$$h(S_A, \mu) = \frac{1}{\mu(A)} h(S, \mu)$$

Let a be the block chosen above, let

$$A = \mathbb{N}^{\mathbb{Z}} \setminus \bigcup_{n=-M}^{\ell(w)+3M} S^{-n}[a]$$

and let S_A denote the corresponding induced map.

Let P denote the partition on $\mathbb{N}^{\mathbb{Z}}$ according to the symbols at coordinates $0, \dots, \ell(w) + 4M$ symbols of a sequence, and order the elements of P in some fixed way as $P = \{P_1, P_2, \dots\}$. It is easy to see that P is a generating partition for S_A , i.e. for $x \in A$ the sequence $n(i)$, defined by $S_A^{i}x \in P_{n(i)}$, determines x . This sequence $(n_i) = (n_i(x))$ is called the itinerary of x .

Let $x' \in \mathbb{N}^{\mathbb{Z}}$ denote the itinerary of x under S_A with respect to the partition P . Let ν denote the measure on $\mathbb{N}^{\mathbb{Z}}$ obtained from μ_A by pushing it forward through $x \mapsto x'$, so that $(\mathbb{N}^{\mathbb{Z}}, S, \nu) \cong (\mathbb{N}^{\mathbb{Z}}, \mu|_A, S_A)$, and note that the map $x \mapsto x'$ is μ_A -a.e. an injection (and by definition measure-preserving w.r.t. ν).

The block a was chosen in such a way that the frequency of occurrences of a (i.e. the density of I_1) does not exceed $1/N$ and therefore $\mu([a]) \leq 1/N$. Therefore,

$$\mu(A) \geq 1 - \frac{\ell(w) + 4M}{N}$$

so by Abramov's formula,

$$h(S, \nu) \leq \frac{1}{\mu(A)} h(S, \mu) \leq \frac{1}{1 - (\ell(w) + 4M)/N} t < h'$$

Let φ, ψ be the maps in Theorem 5.1 with respect to h' and the SFT Y' , that is, a factor map of the free part of the h' -slice of $\mathbb{N}^{\mathbb{Z}}$ into Y' .

- Let $y' = \varphi(x')$.
- Let

$$I_2 = \{i \in \mathbb{Z} : S^i x \in A\}$$

and let $n(i) \in \mathbb{Z}$ be the unique integer such that $S_A^{n(i)} x = S^i x$ (in other words, $n(i)$ is the cardinality of the set $I_2 \cap (0, i)$ or $I_2 \cap [i, 0)$ (depending on the sign of i).

- For $i \in I_2$ let $y_i = (y')_{n(i)}$.

It is not hard to verify that, since φ commutes with S , this definition of y commutes with S . Also, note that A was defined so that if $i \in I_1$ and $j \in I_2$ then either $j < i - M$ or $j > i + \ell(w) + 3M$. Therefore the definition above does not conflict with the previous one.

5.7. Coding $\psi(x')$. Note that we have determined y except for a block of M symbols preceding each w , and a block of $3M$ symbols following each w . We next wish to encode $\psi(x')$ into the remaining space, specifically into the block of length M starting M symbols to the right of each w . Recall that $\psi(x')$ is an infinite sequence $(\sigma_n)_{n=1}^{\infty}$, $\sigma_n \in \{0, 1\}$. Note that ψ is a shift-invariant function.

We encode (σ_n) in y by controlling the frequency of a certain block (this is one of many possible methods). We rely on the following elementary fact: If $|r_n - 2^{-n^2}| < 2^{-n^2}$ and $f = \sum_{n=1}^{\infty} \sigma_n r_n$ is given, then we can recover σ_n from f (instead of 2^{-n^2} we could choose any other rapidly decreasing sequence). Recall that u, v were chosen above to be blocks of length M in Y' .

- For $n = 1, 2, 3, \dots$ choose pairwise disjoint $U_n \in \mathcal{C}$ with $U_n \subseteq [a]$ and such that

$$\left| \frac{\mu(U_n)}{\mu([a])} - 2^{-n^2} \right| < 2^{-n^2}$$

Such a family exists as long as μ arises from a non-atomic and regular measure. To make the choice Borel, for each n choose U_n to be the first element of \mathcal{C} such that $U_n \subseteq [a] \setminus \cup_{k < n} U_k$ and which satisfies the inequality above.

- For $i \in I_1$,
 - If $T^i x \in \cup_{n=1}^{\infty} U_n$ and $\sigma_n = 1$ then set the block in y starting at $i + \ell(w) + 2M$ to u ,
 - Otherwise, set this block to v .

We have now determined y except at blocks of $*$ of length M which occur immediately before and after a w , and another such block beginning at offset $2M$ to the right of each w . From the definition of M , we can now eliminate the $*$'s:

- Fill in each block of $*$'s with an admissible block from Y' chosen in a manner which depends only on the symbols adjacent to the block.

Notice that, since w does not appear in Y' and u, v are admissible for Y' , the blocks in y which make up the complement of the original w -blocks do not contain any w 's. Since w is a marker, it follows that the only occurrences of w in y are those which we created intentionally, i.e. those starting at indices $i \in I_1$

5.8. Decoding. The procedure which produced y is Borel, shift-invariant and defined μ -a.e. for every $\mu \in \mathcal{E}^c(X, T)$. It remains to explain why it is invertible and why the images of different ergodic measures do not intersect.

Suppose y is given. First, note that w, u, v were chosen independently of μ and x .

Since w is a marker we can identify occurrences of it in y , and from these occurrences are precisely those at indices $i \in I_1$, i.e. those whose starting index

occurs at times when x visits $[a]$. We thus recover I_1 , and its density, which is $\mu([a])$.

Having found I_1 we now look at the density of u 's which begin at indices $i + \ell(w) + M$, $i \in I_1$ (note that there may also be u 's which appear elsewhere). Let p be this density. Then by construction we have

$$\frac{p}{\mu([a])} = \sum \sigma_n \frac{\mu(U_n)}{\mu([a])}$$

so from $p/\mu([a])$ we can recover (σ_n) , and hence obtain $\psi(x')$.

Finally, having found I_1 we can recover I_2 because

$$I_2 = \mathbb{Z} \setminus \bigcup_{i \in I_1} [i - M, i + \ell(w) + 3M]$$

From $y|_{I_2}$ we reconstruct $y' = \varphi(x')$. Since we know $\psi(x')$, we apply it and recover x' , and from x' and I_2 we recover x .

REFERENCES

- [1] Roy L. Adler and Brian Marcus. Topological entropy and equivalence of dynamical systems. *Mem. Amer. Math. Soc.*, 20(219):iv+84, 1979.
- [2] F. Blanchard. β -expansions and symbolic dynamics. *Theoret. Comput. Sci.*, 65(2):131–141, 1989.
- [3] Rufus Bowen. Topological entropy for noncompact sets. *Trans. Amer. Math. Soc.*, 184:125–136, 1973.
- [4] Mike Boyle, Jerome Buzzi, and Ricardo Gómez. Almost isomorphism for countable state Markov shifts. *J. Reine Angew. Math.*, 592:23–47, 2006.
- [5] Mike Boyle, Jerome Buzzi, and Ricardo Gómez. Preliminary manuscript. 2010.
- [6] Jérôme Buzzi. Intrinsic ergodicity of smooth interval maps. *Israel J. Math.*, 100:125–161, 1997.
- [7] Jérôme Buzzi. Subshifts of quasi-finite type. *Invent. Math.*, 159(2):369–406, 2005.
- [8] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic theory on compact spaces*. Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin, 1976.
- [9] Matthew Foreman. A descriptive view of ergodic theory. In *Descriptive set theory and dynamical systems (Marseille-Luminy, 1996)*, volume 277 of *London Math. Soc. Lecture Note Ser.*, pages 87–171. Cambridge Univ. Press, Cambridge, 2000.
- [10] E. Glasner and B. Weiss. On the interplay between measurable and topological dynamics. In *Handbook of dynamical systems. Vol. 1B*, pages 597–648. Elsevier B. V., Amsterdam, 2006.
- [11] B. M. Gurevič. Shift entropy and Markov measures in the space of paths of a countable graph. *Dokl. Akad. Nauk SSSR*, 192:963–965, 1970.
- [12] Bruce P. Kitchens. *Symbolic dynamics*. Universitext. Springer-Verlag, Berlin, 1998. One-sided, two-sided and countable state Markov shifts.
- [13] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.
- [14] Saharon Shelah and Benjamin Weiss. Measurable recurrence and quasi-invariant measures. *Israel J. Math.*, 43(2):154–160, 1982.

- [15] Benjamin Weiss. Measurable dynamics. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 395–421. Amer. Math. Soc., Providence, RI, 1984.

Current address: Fine Hall, Washington Road, Princeton University, Princeton, NJ 08544

E-mail address: `hochman@math.princeton.edu`