AN ERGODIC SYSTEM IS DOMINANT EXACTLY WHEN IT HAS POSITIVE ENTROPY

TIM AUSTIN, ELI GLASNER, JEAN-PAUL THOUVENOT, AND BENJAMIN WEISS

Abstract. An ergodic dynamical system $X$ is called dominant if it is isomorphic to a generic extension of itself. It was shown in [8] that Bernoulli systems with finite entropy are dominant. In this work we show first that every ergodic system with positive entropy is dominant, and then that if $X$ has zero entropy then it is not dominant.

Contents

Introduction 1
1. Relative Bernoulli 3
2. Positive entropy systems are dominant 4
3. Zero entropy systems are not dominant 7
4. The positive entropy theorem for amenable groups 11
References 12

Introduction

We say that an ergodic system $X = (X, X, \mu, T)$ is dominant if a generic extension $\hat{T}$ of $T$ is isomorphic to $T$. We obtain the surprising result that every ergodic positive entropy system of an amenable group has the property that its generic extension is isomorphic to it. For $\mathbb{Z}$ systems we show that conversely, when an ergodic system has zero entropy then it is not dominant. Our first result for $\mathbb{Z}$ actions follows from an extension of a result from [8] according to which a generic extension of a Bernoulli system is Bernoulli with the same entropy (and hence is isomorphic to it by Ornstein’s fundamental result) to the relative situation - together with Austin’s weak Pinsker theorem [3]. The extension to all countable amenable groups relies on the results in [17], [20] and [5]. For the result that zero entropy is not dominant for $\mathbb{Z}$ actions we use an idea from the slow entropy developed in [12].

To make the definition of dominance more precise, as in [9] and [8], we present a convenient way of parametrising the space of extensions of $T$ as follows: Let $X = (X, X, \mu, T)$ be an ergodic system. We will assume throughout this work (excepting the last section, where we will comment about the infinite entropy case) that it is infinite and has finite entropy, which for convenience we assume is equal 1. Let
Let $\mathcal{R} \subset \mathcal{X}$ be a finite generating partition. Let $\mathcal{S}$ be the collection of Rokhlin cocycles with values in the Polish group of measure preserving automorphisms of the unit interval $\text{MPT}(I, \mathcal{C}, \lambda)$, where $\lambda$ is the normalized Lebesgue measure and $\mathcal{C}$ is the Borel $\sigma$-algebra on $I = [0, 1]$. Thus an element $S \in \mathcal{S}$ is a measurable map $x \mapsto S_x \in \text{MPT}(I, \lambda)$, and we associate to it the skew product transformation

$$\hat{S}(x, u) = (Tx, S_x u), \quad (x \in X, u \in I),$$

on the measure space $(X \times I, \mathcal{X} \times \mathcal{C}, \mu \times \lambda)$.

We recall that, by Rokhlin's theorem, every ergodic extension $Y \to X$ either has this form or it is $n$ to $1$ a.e for some $n \in \mathbb{N}$ (see e.g. [7, Theorem 3.18]). Thus the collection $\mathcal{S}$ parametrises the ergodic extensions of $X$ with infinite fibers. This defines a Polish topology on $\mathcal{S}$ which is inherited from the Polish group $\text{MPT}(X \times I, \mu \times \lambda)$ of all the measure preserving transformations.

In [8] we have shown that for a fixed ergodic finite entropy $T$ with property $\mathbf{A}$, a generic extension $\hat{T}$ of $T$ also has the property $\mathbf{A}$, where $\mathbf{A}$ stands for each of the following properties: (i) having the same entropy as $T$, (ii) Bernoulli, (iii) $K$, and (iv) loosely Bernoulli.

Now with these notations at hand the definition above becomes:

0.1. Definition. An ergodic system $X = (X, \mathcal{X}, \mu, T)$ is dominant if there is a dense $G_\delta$ subset $S_0 \subset \mathcal{S}$ such that for each $S \in S_0$ we have $\hat{S} \cong T$.

From [8, Theorems 4.1 and 5.1], if $B$ is a Bernoulli system with finite entropy, then its generic extension is again Bernoulli having the same entropy. By Ornstein's theorem [16] such an extension is isomorphic to $B$. This proves the following.

0.2. Proposition. Every Bernoulli system with finite entropy is dominant.

We recall that an ergodic system $X$ is coalescent if every endomorphism $E$ of $X$ is an automorphism. Note that when an extension $\hat{S}$ as above with $\hat{S} \cong T$ exists then the system $X$ is not coalescent. In fact, if $\pi : \hat{S} \to T$ is the (infinite to one) extension, and $\theta : T \to \hat{S}$ is an isomorphism then $E = \pi \circ \theta$ is an endomorphism of $X$ which is not an automorphism. Thus we have the following:

0.3. Proposition. A dominant system is not coalescent.

Hahn and Parry [10] showed that totally ergodic automorphisms with quasi-discrete spectrum are coalescent. In [15] Dan Newton says:

A question put to me by Parry in conversation is the following: if $T$ has positive entropy does it follow that $T$ is not coalescent?

Using theorems of Ornstein [16] and Austin [3], and results from [8] we can now prove the following theorem.

0.4. Theorem. An ergodic system with positive entropy is not coalescent.

Proof. We first observe that a Bernoulli system is never coalescent (if $B$ is Bernoulli and $B' \to B$ is an isometric extension which is again Bernoulli then, by Ornstein's theorem, $B' \cong B$). Now let $X = (X, \mathcal{X}, \mu, T)$ be an ergodic system with positive entropy. By Austin's weak Pinsker theorem [3] we can write $X$ as a product system $B \times \mathbb{Z}$ with $B$ a Bernoulli system of finite entropy. Finally, as noted in [15, Proposition
1] if \( T = T_1 \times T_2 \), where \( T_1 \) is not coalescent, then \( T \) is not coalescent. In fact, given an endomorphism \( E \) of \( T_1 \) which is not an automorphism, the map \( E \times \text{Id} \), where \( \text{Id} \) denotes the identity automorphism on the second coordinate, is an endomorphism of \( T \) which is not an automorphism. Applying this observation to \( T \) denotes the identity automorphism on the second coordinate, is an endomorphism of \( 1 \) if \( T \in MPT(\mathcal{I}, \lambda) \), 

Thus by Proposition 0.3 we conclude that: the set of non-dominant automorphisms is comeager in \( MPT(\mathcal{I}, \lambda) \), hence also in the dense \( G_\delta \) subset of \( MPT(\mathcal{I}, \lambda) \) comprising the zero entropy automorphisms. However, as we will show in section 3 using a slow entropy argument, the answer is affirmative for every ergodic system with zero entropy.

1. Relative Bernoulli

1.1. Definition. Let \( X = (X, \mathcal{X}, \mu, T) \) be an ergodic system and \( \mathcal{X}_0 \subset \mathcal{X} \) a \( T \)-invariant \( \sigma \)-subalgebra. Let \( X_0 = (X_0, \mathcal{X}_0, \mu_0, T_0) \) be the corresponding factor system and let \( \pi : X \to X_0 \) denote the factor map. We say that \( X \) is relatively Bernoulli over \( X_0 \) if there is a \( T \)-invariant \( \sigma \)-algebra \( \mathcal{X}_1 \subset \mathcal{X} \) independent of \( \mathcal{X}_0 \) such that \( \mathcal{X} = \mathcal{X}_0 \vee \mathcal{X}_1 \), and there is a \( \mathcal{X}_1 \)-generating finite partition \( \mathcal{K} \subset \mathcal{X}_1 \) such that the partitions \( \{T^i\mathcal{K}\}_{i \in \mathbb{Z}} \) are independent; in other words, the corresponding system \( X_1 = (X_1, \mathcal{X}_1, \mu_1, T_1) \) is Bernoulli and \( X \cong X_0 \times X_1 \).

If \( R_0 \) is a finite generating partition for \( \mathcal{X}_0 \) and \( R \) is a finite generating partition for \( \mathcal{X} \) then J.-P. Thouvenot showed that there is a condition called relatively weak Bernoulli, which is equivalent to the extension being relatively Bernoulli, see [23] and also [14]. This condition is as follows:

1.2. Definition. The partition \((R, T)\) is relatively Bernoulli over \((R_0, T)\) if for every \( \epsilon > 0 \) there is \( N \) such that for a collection \( \mathcal{G} \) of atoms \( A \) of the partition \( \vee_{i=-\infty}^{1} T^{-i} R \), and a collection \( \mathcal{G}_0 \) of atoms \( B \) of the partition \( \vee_{i=-\infty}^{1} T^{-i} R_0 \), we have

\[
\mu \left( \bigcup \{ A \cap B : A \in \mathcal{G}, B \in \mathcal{G}_0 \} \right) > 1 - \epsilon,
\]

for all such \( A \) and \( B \).

Since \( \vee_{i=-k}^{1} T^{-i} R \nRightarrow \vee_{i=-\infty}^{1} T^{-i} R \) and \( \vee_{i=-k}^{k} T^{-i} R_0 \nRightarrow \vee_{i=-\infty}^{1} T^{-i} R_0 \), this can be formulated in finite terms as: for every \( \epsilon > 0 \) there is \( N \) and \( \exists k_0 \) such that for all \( k > k_0 \) there is a collection \( \mathcal{G} \) of atoms \( A \) of \( \vee_{i=-k}^{1} T^{-i} R \) and a collection \( \mathcal{G}_0 \) of atoms \( B \) of \( \vee_{i=-k}^{k} T^{-i} R_0 \) such that

\[
\mu \left( \bigcup \{ A \cap B : A \in \mathcal{G}, B \in \mathcal{G}_0 \} \right) > 1 - \epsilon,
\]

for all such \( A \) and \( B \).
One last change — instead of (2b) we can also require that for \( A, A' \in \mathcal{G} \), \( B \in \mathcal{G}_0 \)
\[
\bar{d}_N (\text{dist}(\bigvee_{i=0}^{N-1} T^{-i} \mathcal{R} \upharpoonright A \cap B), \text{dist}(\bigvee_{i=0}^{N-1} T^{-i} \mathcal{R} \upharpoonright A' \cap B)) < \epsilon.
\]
That (2b) implies (3) with \( 2\epsilon \) is immediate.

For the converse implication observe first that the distribution \( \text{dist}(\bigvee_{i=0}^{N-1} T^{-i} \mathcal{R} \upharpoonright B) \) is the average of \( \text{dist}(\bigvee_{i=0}^{N-1} T^{-i} \mathcal{R} \upharpoonright A \cap B) \) over all \( A \in \mathcal{G} \) and for all \( \mathcal{R} \). We will first show that the collection of the elements \( \hat{S} \) of \( T \) is relatively Bernoulli over \( \mathbf{X}_0 \).

Proof. For convenience we assume that the relative entropy is 1.

As in [8] let \( \mathcal{R} \subset \mathcal{X} \) be a finite relatively generating partition for \( \mathbf{X} \) over \( \mathbf{X}_0 \) with entropy 1 (so that \( \mathcal{R} \) is a Bernoulli partition independent of \( \mathbf{X}_0 \)), and let \( \mathcal{R}_0 \subset \mathcal{X}_0 \) be a finite generator for \( \mathbf{X}_0 \). Let \( \mathcal{S} \) be the collection of Rokhlin cocycles with values in \( \text{MPT}(I, \lambda) \), where \( \lambda \) is the normalized Lebesgue measure on the unit interval \( I = [0, 1] \).

Thus an element \( S \in \mathcal{S} \) is a measurable map \( x \mapsto S_x \in \text{MPT}(I, \lambda) \), and we associate to it the skew product transformation
\[
\hat{S}(x, u) = (Tx, S_x u), \quad (x \in X, u \in I).
\]

Let \( Y = X \times I \) and set \( \mathbf{Y} = (Y, \mathbf{y}, \mu \times \lambda) \), with \( \mathbf{y} = X \otimes \mathcal{C} \).

Part I: By Theorem 4.1 of [8] there is a dense \( G_\delta \) subset \( \mathcal{S}_0 \subset \mathcal{S} \) with \( h(\hat{S}) = 1 \) for every \( S \in \mathcal{S}_0 \). We will first show that the collection of the elements \( S \in \mathcal{S}_0 \) for which the corresponding \( \hat{S} \) is relatively Bernoulli over \( \mathbf{X}_0 \) forms a \( G_\delta \) set.

As the inverse limit of relatively Bernoulli systems is relatively Bernoulli, see [22, Proposition 7], to show that a transformation \( T \) on \( (X, \mathcal{X}, \mu) \) is relatively Bernoulli over \( \mathbf{X}_0 \) it suffices to show that for a refining sequence of partitions
\[
P_1 \prec \cdots \prec P_n \prec P_{n+1} \prec \cdots
\]
such that the corresponding algebras \( \hat{P}_n \) satisfy \( \bigvee_{n \in \mathbb{N}} \hat{P}_n = \mathcal{X} \), for each \( n \), the process \( (T, P_n) \) is relatively VWB relative to \( (T, \mathcal{R}_0) \).

For each \( n \in \mathbb{N} \) let \( \mathcal{Q}_n \) denote the dyadic partition of \([0, 1]\) into intervals of size \( 1/2^n \), and let
\[
P_n = \mathcal{R} \times \mathcal{Q}_n.
\]

For any \( S \in \mathcal{S}_0 \) the relative entropy of \( \mathbf{Y} = X \times [0, 1] \) over \( \mathbf{X}_0 \) is also 1. Thus for all \( n \) we have
\[
H(P_n \mid (\bigvee_{i=-\infty}^{i=n} \hat{S}^{-i} P_n) \lor (\bigvee_{i=-\infty}^{i=n} \hat{S}^{-i} \mathcal{R}_0)) = 1,
\]
and for all \( N \geq 1 \)
\[
H(\bigvee_{i=-N}^{i=0} \hat{S}^{-i} P_n \mid (\bigvee_{i=-\infty}^{i=1} \hat{S}^{-i} P_n) \lor (\bigvee_{i=-\infty}^{i=\infty} \hat{S}^{-i} \mathcal{R}_0)) = N.
\]
Therefore, we can find a suitably small $\delta > 0$ such that for $k_0$ large enough
\[ H(\bigvee_{i=0}^{N-1} \hat{S}^{-i}p_n | (\bigvee_{i=k_0}^{1} \hat{S}^{-i}p_n) \lor (\bigvee_{i=k_0}^{0} \hat{S}^{-i}R_0)) < N + \delta. \]

Now, conditioned on the partition
\[ (\bigvee_{i=k_0}^{1} \hat{S}^{-i}p_n) \lor (\bigvee_{i=k_0}^{0} \hat{S}^{-i}R_0) \]
the partition $\bigvee_{i=0}^{N-1} \hat{S}^{-i}p_n$ will be $\eta$-independent of
\[ (\bigvee_{i=k_0-1}^{1} \hat{S}^{-i}p_n) \lor (\bigvee_{i=k_0+1}^{0} \hat{S}^{-i}R_0) \]
for all $k \geq k_0$ for $\eta$ small enough (see Definition 5.1 in [8] and the following discussion), so that the inequality (3) in Section 1 (with $\mathcal{P}_n$ replacing $\mathcal{R}$) for $k = k_0$ will imply (3) with $2\epsilon$, for all $k > k_0$.

Define the set $U(n, N_1, N_2, \epsilon, \delta)$ to consist of those $S \in \mathcal{S}_0$ that satisfy:

1. $H(\bigvee_{i=0}^{N_1-1} \hat{S}^{-i}p_n | (\bigvee_{i=-N_2}^{1} \hat{S}^{-i}p_n) \lor (\bigvee_{i=-N_2}^{0} \hat{S}^{-i}R_0)) < N_1 + \delta$,

2. $d_{N_1} \left( (\bigvee_{i=0}^{N_1-1} \hat{S}^{-i}p_n \upharpoonright A \cap B, \bigvee_{i=0}^{N_1-1} \hat{S}^{-i}p_n \upharpoonright A' \cap B) \right) < \epsilon,$

for a set of atoms $A, A' \in \mathcal{G}$, $B \in \mathcal{G}_0$,

where $\mathcal{G} \subset \bigvee_{i=-N_2}^{1} \hat{S}^{-i}p_n$, $\mathcal{G}_0 \subset \bigvee_{i=-N_2}^{0} \hat{S}^{-i}R_0$ and $(\mu \times \lambda) \left( \bigcup \{A \cap B : A \in \mathcal{G}, B \in \mathcal{G}_0\} \right) > 1 - \epsilon.$

Now the sets $U(n, N_1, N_2, \epsilon, \delta)$ are open (easy to check) and that the $G_\delta$ set
\[ S_1 = \bigcap_{n,k,l} U(n, N_1, N_2, 1/k, 1/l) \]
comprises exactly the elements $S \in \mathcal{S}_0$ for which the corresponding $\hat{S}$ is relatively Bernoulli over $\mathbf{X}_0$. Thus, if $S \in \mathcal{S}_0$ is such that $\hat{S}$ is relatively Bernoulli, then for every $n, \epsilon, \delta$, there are $N_1, N_2$ such that $S \in U(n, N_1, N_2, \epsilon, \delta)$, and conversely, for every relatively Bernoulli $\hat{S}$ the corresponding $S$ is in $\mathcal{S}_1$.

**Part II:** The collection $S_1$ is nonempty. To see this we first note that the Bernoulli system $\mathbf{X}_1$ admits a proper extension $\mathbf{X}_1 \to \mathbf{X}_1$ which is also Bernoulli and has the same entropy. This follows e.g. by a deep result of Rudolph [18, 19] who showed that every weakly mixing group extension of $\mathbf{X}_1$ is again a Bernoulli system. An explicit example of such an extension of the 2-shift is given by Adler and Shields, [2]. Since $\mathbf{X}_1$ is weakly mixing the product system $\hat{X} = \mathbf{X}_0 \times \mathbf{X}_1$ is ergodic and $\hat{X} \to \mathbf{X}_0$ is an element of $S_1$.

Now apply the relative Halmos theorem [9, Proposition 2.3], to deduce that the $G_\delta$ subset $S_1$ is dense in $\mathcal{S}$, as claimed. \qed

We can now deduce the positive entropy part of our main result.

2.2. Theorem. Every ergodic system $\mathbf{X} = (X, \mathcal{X}, \mu, T)$ of positive finite entropy is dominant.
By Austin’s weak Pinsker theorem [3] we can present \( X \) as a product system \( X = B \times Z \), where \( B \) is a Bernoulli system with finite entropy. Thus \( X \) is relatively Bernoulli over \( Z \), and by Theorem 2.1 it follows that a generic extension \( \hat{S} \) of \( X \) is relatively Bernoulli over \( Z \). Therefore, for such \( \hat{S} \) the system \( Y = (X \times I, \mathcal{X} \times \mathcal{E}, \mu \times \lambda, \hat{S}) \) is again of the form \( Y = B' \times Z \) with \( B' \) a Bernoulli system with the same entropy as that of \( B \). By Ornstein’s theorem [16] \( B \sim B' \), whence also \( X \sim Y \) and our proof is complete. \( \square \)

2.3. Remark. With notations as in the proofs of Theorems 2.1 and 2.2 observe that for every \( S \in \mathcal{S} \) the system \((Y, \mu \times \lambda, \hat{S})\) admits \( Z = (Z, \mathcal{Z}, \mu, T) \) (with \( Z \) considered as a subalgebra of \( X \)) as a factor:

\[(Y, \mu \times \lambda, \hat{S}) \to X \to Z.\]

In the Polish group \( G = \text{MPT}(Y, \mu \times \lambda) \) consider the closed subgroup \( G_Z = \{g \in G : gA = A, \forall A \in \mathcal{Z}\} \). We now observe that the residual set \( S_1 \subset S_0 \), of those \( S \in S_0 \) for which \( \hat{S} \) is Bernoulli over \( Z \) with the same relative entropy over \( X \) is a single orbit for the action of \( G_Z \) under conjugation.

In the last section (Section 4) we will show that the positive entropy theorem holds for any countable amenable group.

In [8, Theorem 6.4] it was shown that the generic extension of a K-automorphism is a mixing extension. We will next prove an analogous theorem for a general ergodic system with positive entropy. We first prove the following relatively Bernoulli analogue of Theorem 6.2 in [8].

2.4. Theorem. Let \( X = (X, \mathcal{X}, \mu, T) \) be a relatively Bernoulli system over \( X_0 \), and \( S \) a Rokhlin cocycle with values in \( \text{MPT}(I, \lambda) \), where \( I = [0, 1] \) and \( \lambda \) is Lebesgue measure on \( I \). We denote by \( \hat{S} \) the transformation

\[\hat{S}(x, u) = (Tx, S_x u),\]

on \( Y = X \times I \), and let

\[\hat{S}(x, u, v) = (Tx, S_x u, S_x v), \quad (x, u, v) \in W = X \times I \times I,\]

be the relative independent product of \( Y \) with itself over \( X \). Then for a generic \( S \in \mathcal{S} \) the transformation \( \hat{S} \) is relatively Bernoulli over \( X_0 \).

Proof. For the \( G_\delta \) part we follow, almost verbatim, the proof of Theorem 2.1, where we now let \( Q_n \) denote the product dyadic partition of \( I \times I \) into squares of size \( \frac{1}{2^n} \times \frac{1}{2^n} \) and, with notations as in the proof of Theorem 2.1, we let \( P_n = R \times Q_n \).

Thus it only remains to show that the \( G_\delta \) set \( S_1 \), comprising those \( S \in S_0 \) for which \( \hat{S} \) is relatively Bernoulli on \( W = X \times I \times I \) relative to \( X_0 \), is non empty. Now examples of skew products over a Bernoulli system with such properties are provided by Hoffman in [11]. The base Bernoulli transformation that Hoffman constructs for his example can be arranged to have arbitrarily small entropy by an appropriate choice of the parameters used in the construction in section 3 (the skew product example is in section 4 and the proof of Bernoullicity is in section 5). Using such construction on \( X \) (where the cocycle is measurable with respect to the Bernoulli direct component of \( X \)) we obtain our required extension of \( X \). This completes our proof. \( \square \)
We also recall the following criterion [8, Lemma 6.5]:

2.5. Lemma. Let $X$ be ergodic and $Y$ be a factor of $X$. Then the following are equivalent:

(1) $X$ is a relatively mixing extension of $Y$.
(2) In the relatively independent product $X \times X$, the Koopman operator restricted to $L^2(Y)^\perp$ is mixing.

2.6. Theorem. Let $X = (X, X, \mu, T)$ be an ergodic system with positive entropy, then the generic extension of $X$ is relatively mixing over $X$.

Proof. By the weak Pinsker theorem [3] we can present $X$ as a product system $X = Z \times B$, where $B$ is a Bernoulli system with finite entropy. Thus $X$ is relatively Bernoulli over $Z$, and by Theorem 2.4 it follows that a generic extension $\tilde{S}$ of $X$ to $X \times I \times I$ is still relatively Bernoulli over $Z$. Thus the extended system $W$ on $W = X \times I \times I$ with $\tilde{S}$ action, has the form $W = Z \times B'$ with $B'$ again a Bernoulli system.

Now for the system $Y$, defined on $Y = X \times I$ by

$$\tilde{S}(x, u) = (Tx, S_x u),$$

we have that the corresponding relative product system $Y \times Y$ is isomorphic to $W_X$, which is a Bernoulli extension of $Z$ and therefore, by Lemma 2.5, a relatively mixing extension of $Z$. A fortiori $Y \times Y$ is a relatively mixing extension of $X$ and our proof is complete. $\Box$

3. Zero entropy systems are not dominant

3.1. Definition.

- For $\omega, \omega' \in \{0, 1\}^n$ the Hamming (or $d$-distance) is defined by

$$d(\omega, \omega') = \frac{1}{n} \# \{ 0 \leq i < n : \omega_i \neq \omega'_i \}.$$

- For two measurable partitions $Q = \{A_i\}_{i=1}^n$, $\hat{Q} = \{B_i\}_{i=1}^n$ of a measured space $(X, \mu)$, the distance $d(Q, \hat{Q})$ is defined by

$$d(Q, \hat{Q}) = \frac{1}{2} \sum_{i=1}^n \mu(A_i \triangle B_i).$$

3.2. Theorem. Every ergodic system $X$ with zero entropy is not dominant.

3.3. Remark. Recently Terrence Adams [1] has proved a somewhat analogous result in the setting of MPT, the group of all measure preserving transformations of the unit interval with Lebesgue measure. It is well known that generically a $T$ in MPT has zero entropy. What Adams shows is that for any preassigned growth rate for slow entropy, the generic transformation has a complexity which exceeds that rate. In our proof of theorem 3.2 we don’t introduce a formal definition of slow entropy but its definition lies behind our lemma 3.4.
Proof. We first choose a strictly ergodic model $X = (X, X, \mu_0, T)$ for our system which is a subshift of $\{0, 1\}^\mathbb{Z}$. By the variational principle this model will have zero topological entropy. (To see that such a model exists, see for example [6], where this fact can be deduced from property (b) on page 281 and Theorem 29.2 on page 301.) Denote by $a_n$ the number of $n$-blocks in $X$, so that $a_n$ is sub-exponential.

For $x_0 \in X$ and $Q = \{Q_0, Q_1\}$ a partition of $X$ let

$$B_n(x_0, \epsilon) = \{x \in X : d_n(Q_n(x), Q_n(x_0)) < \epsilon\},$$

where for a point $x \in X$ and $n \geq 1$ we write

$$Q_n(x) = \omega_0\omega_1\omega_2 \ldots \omega_{n-1}, \text{ when } x \in \bigcap_{i=0}^{n-1} T^{-i}(Q_{\omega_i}).$$

3.4 Lemma. For $\epsilon < \frac{1}{100}$ and $\delta < \frac{1}{100}$ there is an $N$ such that for all $n \geq N$, if $m$ is the minimal number such that there are points $x_1, x_2, \ldots, x_m$ with

$$\mu_0(\bigcup_{i=1}^m B_n(x_i, \epsilon)) > 1 - \delta,$$

then $m \leq a_{2n}$.

Proof. Denote by $\mathcal{P} = \{P_1, P_2\}$ the partition of $X$ according to the 0-th coordinate. Given $\epsilon > 0$ there is some $k_0$ and a partition $\hat{Q}$ measurable with respect to $\bigvee_{i=-k_0}^{k_0} T^i\mathcal{P}$ such that

$$d(\mathcal{Q}, \hat{Q}) < \frac{\epsilon}{2}.$$

By ergodicity there exists an $N$ such that for $n \geq N$ there is a set $A \subset X$ with $\mu_0(A) > 1 - \delta$ with

$$d_n(Q_n(x), \hat{Q}_n(x)) < \epsilon, \quad \forall x \in A.$$

Let $\{\alpha_i\}_{i=1}^\ell$ be those atoms of $\bigvee_{i=-k_0}^{n+k_0} T^i\mathcal{P}$ such that $\alpha_i \cap A \neq \emptyset$, so that $\ell \leq a_{n+2k_0+1}$.

Choose $x_i \in \alpha_i \cap A$, $1 \leq i \leq \ell$. We claim that

$$A \subset \bigcup_{i=1}^\ell B_n(x_i, \epsilon).$$

For $x \in \bigcup_{i=1}^\ell \alpha_i$ we denote by $i(x)$ that index such that $x \in \alpha_{i(x)}$. Now since $x$ and $x_{i(x)}$ are in $A$ we have

$$\hat{d}_n(Q_n(x), \hat{Q}_n(x)) < \epsilon \quad \text{and} \quad d_n(Q_n(x), \hat{Q}_n(x)) < \epsilon.$$

Since $x \in \alpha_{i(x)}$, $\hat{Q}_n(x) = Q_n(x)$. Therefore

$$\hat{d}_n(Q_n(x), Q_n(x_{i(x)})) < 2\epsilon,$$

whence $x \in B_n(x_{i(x)}, \epsilon)$. This proves our claim and we conclude that $m \leq \ell \leq a_{n+2k_0+1}$. Thus for sufficiently large $n$ we indeed get $m \leq a_{2n}$.

We will show that a generic extension of $T$ to $(Y, \mu) = (X \times [0, 1], \mu_0 \times \lambda)$, with $\lambda$ Lebesgue measure on $[0, 1]$, is not isomorphic to $X$. To do this we will show that for a generic extension $\hat{S}$ the partition $\mathcal{Q}$ of $Y$, defined by splitting $X \times [0, 1]$ into $\{Q_0, Q_1\} = \{X \times [0, \frac{1}{2}], X \times [\frac{1}{2}, 1]\}$, will not satisfy the conclusion of this lemma.

Notations:

- $S$ is the Polish space comprising the measurable Rohklin cocycles $x \mapsto S_x \in \mathrm{MPT}([0, 1], \lambda)$.
- For $S \in S$ let $\hat{S}(x, u) = (T_x, S_x u)$. 

• $Q^\hat{S}_n(y) = \omega_0\omega_1\omega_2\ldots\omega_{n-1}$, where $y \in \cap_{i=0}^{n-1} \hat{S}^{-i}(Q_{\omega_i})$.
• $C(\hat{S}, n, \epsilon, \delta) = \min \{ k : \exists y_1, y_2, \ldots, y_k \in Y, \text{ such that } \mu(\cup_{i=1}^{n} B^\hat{S}_{i}(y_i, \epsilon)) > 1 - \delta \}$.

Define now

$$\mathcal{U}(N, \epsilon, \delta) = \{ S \in \mathcal{S} : \exists n \geq N \text{ such that } C(\hat{S}, n, \epsilon, \delta) > 2a_{2n} \}.$$  

This is an open subset of $\mathcal{S}$ (see e.g. [8] for similar claims). We will show that, for sufficiently small $\epsilon$ and $\delta$, it is dense in $\mathcal{S}$.

First consider the case $S_0 = \text{id}$. Let $\eta > 0$ be given and choose $M$ so that $1/M < \eta$. Now build a Rohklin tower for $T$, with base $B_0$ and heights $mM > N$ and $mM + 1$ for a suitable $m$, filling all of $X$. Let $B = B_0 \times [0,1]$ be the base of the corresponding tower in $(Y, \mu, \hat{S})$. We modify $S_0 = \text{id}$ only on the levels $T^{jM-1}B_0$ for $1 \leq j \leq m$, so that the new $S$ will be within $\eta$ of $S_0$. The $Q$-names of the points in $T^{jM-1}B$ are constant for all $0 \leq j < m$. We modify $S_0$ on the levels $T^{jM-1}B$ so that we see all possible 0-1 names for the $M$-blocks as we move up the tower with equal measure. A similar procedure is described as independent cutting and stacking and is explained in details in section I.10.d in Paul Shields book [21].

3.5. **Lemma.** Any $B_{mM}(y, \epsilon)$ ball has measure at most $2^m(\frac{1}{2} + H(2\epsilon,1-2\epsilon))$.

**Proof.** The $Q_{mM}$-names of points $y \in B$ are constant on blocks of length $M$ and all sequences of zeros and ones have equal probability by construction. So by a well known estimation (using Stirling’s formula), in $\{0,1\}^m$ with uniform measure, the measure of an $\epsilon$-ball in normalized Hamming metric is $\leq 2^m(\frac{1}{2} + H(2\epsilon,1-2\epsilon))$.

For points in the lower half of the tower over $B$ we have a similar estimate with $m$ replaced by some $\ell > \frac{1}{2}m$ and $\epsilon$ replaced by $\frac{2}{\ell}\epsilon < 2\epsilon$. For points in the upper half of the tower, for some $\ell < \frac{1}{2}m$ we have that $\hat{S}^{\ell}y \in B$ and then we get an estimate with $m - \ell > \frac{1}{2}m$. This proves the lemma.

From this lemma it follows that in order to achieve even $\frac{1}{2}$ as $\mu(\cup_{i=1}^{L} B_{mM}(y_i, \epsilon))$ we must have $L \cdot 2^m(\frac{1}{2} + H(2\epsilon,1-2\epsilon)) > \frac{1}{2}$, hence

$$L \geq \frac{1}{2} \cdot 2^m(\frac{1}{2} - H(2\epsilon,1-2\epsilon)).$$

Since $a_n$ is sub-exponential this lower bound certainly exceeds $a_{2mM}$ if $m$ is sufficiently large. This shows that this modified $S$ is an element of $\mathcal{U}(N, \epsilon, \delta)$.

A similar construction can be carried out for any $S \in \mathcal{S}$. The main point that needs to be checked is that for small $\epsilon$, no $B_{M}^\hat{S}(y, \epsilon)$-ball can have measure greater than $\frac{1}{2} + 2\epsilon$.

3.6. **Lemma.** For any $\hat{S}$ and all $y_0$

$$\mu(B_M^\hat{S}(y_0, \epsilon)) \leq \frac{1}{2} + 2\epsilon.$$  

**Proof.** Let $Q_M^\hat{S}(y_0) = \omega_0\omega_1\ldots\omega_{M-1}$. Then

$$\tilde{d}_M(Q_M^\hat{S}(y), Q_M^\hat{S}(y_0)) = \frac{1}{M} \sum_{i=0}^{M-1} 1_{Q_{\omega_i}}(\hat{S}^i y_0)(1 - 1_{Q_{\omega_i}}(\hat{S}^i y)).$$
and
\[ \int_Y \tilde{d}_M(Q_M^S(y), Q_M^S(y_0)) \, d\mu = \frac{1}{2}. \]

Since \( \tilde{d}_M \leq 1 \), the measure of the set where \( \tilde{d}(Q_M^S(y), Q_M^S(y_0)) \leq \epsilon \) cannot exceed \( \frac{1}{2} + 2\epsilon \).

This lemma, which is formulated for the measure \( \mu \) on the entire space \( Y \), in fact holds as well for any level \( L_j = \hat{S}^i_j B \) in the tower, when we replace \( \mu \) by the measure \( \mu \) restricted to \( L_j \). This is so because the partition \( \{Q_0, Q_1\} \) intersects each level of the tower in relative measure \( \frac{1}{2} \) and \( \hat{S} \) is measure preserving.

We now mimic the proof outlined for \( S_0 = \text{id} \) and, given \( S \in \mathcal{S} \), using an independent cutting and stacking we change \( \hat{S} \) as follows. For the level \( L_j = \hat{S}^i_j B \) consider the partition
\[ \mathcal{R}_j = \bigvee_{i=0}^{M-1} \hat{S}^{-i}(Q \cap \hat{S}^{j+i} B). \]

We change the transformation \( \hat{S} \) at the transition from level \( jM - 1 \) to level \( jM \), so that these partitions \( \mathcal{R}_j \) will become independent.

We want to estimate the size of an \( mM \)-\( \epsilon \) ball around a point \( y_0 \in B \). If \( y \in B \) belongs to this ball there is a set \( A \subset \{0, 1, \ldots, nM - 1\} \) with \( |A| \leq \epsilon mM \) where the \( mM \)-names of \( y \) and \( y_0 \) differ. We need now a simple lemma:

**3.7. Lemma.** Let \( A \subset \{0, 1, \ldots, mM - 1\} \) such that \( |A| \leq \epsilon mM \). Denote \( I_j = \{jM, jM + 1, \ldots, jM + M - 1\}, 0 \leq j < m - 1 \). Let \( J \subset \{0, 1, \ldots, m - 1\} \) be the set of \( \ell \) such that
\[ |I_\ell \cap A| < \sqrt{\epsilon}M. \]
Then \( |J| > (1 - \sqrt{\epsilon})m \).

**Proof.** Let \( K = \{0, 1, \ldots, mM - 1\} \setminus J \). Then
\[ \epsilon mM \geq |\bigcup_{k \in K} I_k \cap A| \geq M\sqrt{\epsilon}|K|. \]
Thus \( |K| \leq \sqrt{\epsilon}m \), whence \( |J| > (1 - \sqrt{\epsilon})m \). \( \square \)

Next using Lemma 3.6 for each level of the form \( T_j^{JM} B_0 \), we will estimate the size of an \( mM \)-\( \epsilon \) ball. So fix a point \( y_0 \in B \). If \( y \in B_{mM}(y_0, \epsilon) \) then by Lemma 3.7 there is a set of indices \( J_y \subset \{1, 2, \ldots, m\} \) such that
1. \( |J_y| \geq (1 - \sqrt{\epsilon})m \),
2. For each \( j \in J_y \), \( \hat{S}^j y \in B_M(\hat{S}^i y_0, \sqrt{\epsilon}) \).

The number of possible sets that satisfy (1) is bounded by \( 2^{mH(\sqrt{\epsilon}, 1 - \sqrt{\epsilon})} \). By Lemma 3.6 and by the independence, for a fixed such \( J_y \) the measure of the set of points that satisfy (2) is at most
\[ \left( \frac{1}{2} + 2\sqrt{\epsilon} \right)^{m(1 - \sqrt{\epsilon})}. \]
Write \( \left( \frac{1}{2} + 2\sqrt{\epsilon} \right)^{1 - \sqrt{\epsilon}} = 2^{-c} \), where \( c \geq c_0 > 0 \) for all sufficiently small \( \epsilon \). Then
\[ 2^{-cm} \cdot 2^{mH(2\epsilon, 1 - 2\epsilon)} = 2^{m(-c + H(2\epsilon, 1 - 2\epsilon))} \leq 2^{-\frac{m}{2}c_0}, \]
for \( H(2\epsilon, 1 - 2\epsilon) \leq \frac{1}{2}c_0 \). We now see that the measure of the ball \( B_{mM}(y_0, \epsilon) \) is bounded by \( 2^{-\frac{m}{2}c_0} \).
AN ERGODIC SYSTEM IS DOMINANT EXACTLY WHEN IT HAS POSITIVE ENTROPY

This was done for \( y_0 \in B \) and as in the proof of Lemma 3.5 we obtain the suitable estimations for any \( y \) in the tower over \( B \). We conclude the argument as in the case \( S = \text{id} \) and again it follows that the resultant modified \( S \) is an element of \( \mathcal{U}(N, \epsilon, \delta) \).

Finally for fixed sufficiently small \( \epsilon \) and \( \delta \) setting
\[
\mathcal{E} = \bigcap_{N=1}^{\infty} \mathcal{U}(N, \epsilon, \delta)
\]
we obtain the required dense \( G_\delta \) subset of \( S \), where for each \( S \in \mathcal{E} \) the corresponding \( \hat{S} \) is not isomorphic to \( T \). In fact, if \( \hat{S} \) would be isomorphic to \( T \) then the isomorphism would take the partition \( \mathcal{Q} \) of \( Y \) to a partition \( \tilde{\mathcal{Q}} \) of \( X \). Applying Lemma 3.4 to \( \tilde{\mathcal{Q}} \) we see that there is some \( N \) such that for all \( n \geq N \) the conclusion of the lemma holds. But since \( S \in \mathcal{E} \) this is a contradiction. \( \square \)

4. THE POSITIVE ENTROPY THEOREM FOR AMENABLE GROUPS

We fix an arbitrary infinite countable amenable group \( G \). We let \( \mathbb{A}(G, \mu) \) denote the Polish space of measure preserving actions \( \{T_g\}_{g \in G} \) of \( G \) on the Lebesgue space \( (X, \mathcal{X}, \mu) \). (For a description of the topology on \( \mathbb{A}(G, \mu) \) we refer e.g. to [13].)

As in the proof of Theorem 2.1 let \( \mathcal{S} \) be the collection of Rokhlin cocycles from \( X \) with values in \( \text{MPT}(I, \lambda) \), that is, \( \mathcal{S} \) is a family \( \{S^g\}_{g \in G} \), where each element \( S^g \) is a collection of measurable maps \( x \mapsto S^g_x \in \text{MPT}(I, \lambda) \), such that for \( g, h \in G \) and \( x \in X \) we have
\[
S^{gh}(x) = S^g(T_h x)S^h(x), \quad \mu \text{ a.e.}
\]
We associate to \( S \in \mathcal{S} \) the skew product transformation
\[
\hat{S}^g(x, u) = (T_g x, S^g_x u), \quad (x \in X, u \in I).
\]
Let \( Y = X \times I \) and set \( Y = (Y, \mathcal{Y}, \mu \times \lambda) \), with \( \mathcal{Y} = \mathcal{X} \otimes \mathcal{C} \).

A free \( G \)-action \( X \) defines an equivalence relation \( R \subset X \times X \), where \( (x, x') \in R \) iff \( \exists g \in G, \; x' = gx \), and a cocycle \( S \in \mathcal{S} \) defines uniquely a cocycle \( \alpha \) on \( R^1 \):
\[
\alpha(x, x') = S^g_x.
\]
This map is one-to-one and onto from the set of cocycles on \( X \) to the set of cocycles on \( R \). For more details on this correspondence see [13, Section 20, C].

Let now
\[
X = (X, \mathcal{X}, \mu, \{T_g\}_{g \in G}) \rightarrow X_0 = (X_0, \mathcal{X}_0, \mu_0, \{(T_0)_g\}_{g \in G})
\]
be a \( G \)-Bernoulli extension, where this notion is defined exactly as in Definition 1.1.

4.1. Definition. If \( G \) and \( H \) are two countable groups acting as measure preserving transformations \( \{T_g\}_{g \in G}, \{S_h\}_{h \in H} \) on the measure space \( (Z, \nu) \) we say that the actions are orbit equivalent if for \( \nu \)-a.e. \( z \in Z \), \( Gz = Hz \).

---

\(^{1}\) A cocycle \( \alpha \) on \( R \) is a function from \( R \) to \( \text{MPT}(I, \lambda) \) which satisfies the cocycle equation
\[
\alpha(x, z) = \alpha(y, z)\alpha(x, y).
\]
In [17] and [4] it is shown that any ergodic measure preserving action of an amenable group is orbit equivalent to an action of $\mathbb{Z}$.

We can now state and prove an extension of Theorem 2.1 to free actions of $G$, and moreover we can also get rid of the finite entropy assumption on $X$.

4.2. **Theorem.** Let $X = (X,\mathcal{X},\mu,\{T_g\}_{g \in G})$ be an ergodic $G$-system which is relative Bernoulli over a free system $X_0$ with finite relative entropy, so that $X = X_0 \times X_1$. Then, the generic extension $\hat{S}$ of $\{T_g\}_{g \in G}$ is relatively Bernoulli over $X_0$.

**Proof.** By [17] there is $T_0 : X_0 \to X_0$ such that orbits of $T_0$ coincide with $G$-orbits on $X_0$, and such that $T_0$ has zero entropy. The $G$-factor map $X = X_0 \times X_1 \to X_0$ is given by a constant cocycle whose constant value is the Bernoulli action on the Bernoulli factor $X_1$. We use this cocycle, now viewed as a cocycle on the equivalence relation defined by $T_0$, to define an extension $T : X \to X$. By [20] the relative entropy of a generic such $T$ over $T_0$ is the same as that of the $G$-action $X$ over $X_0$. By [5] the extension of $\mathbb{Z}$-systems $\pi : T \to T_0$ is again relatively Bernoulli. Applying Theorem 2.1 to $\pi$ we conclude that a dense $G_\delta$ subset $S_1(\mathbb{Z})$ of extensions of $T$ is such that each $\hat{S} \in S_1(\mathbb{Z})$ is relatively Bernoulli over $T_0$. Finally applying [5] in the other direction we conclude that the corresponding set of extensions $S_1(G)$ is again a dense $G_\delta$ subset of $S(G)$ and that for each $S \in S_1(G)$, the corresponding $G$-system is relatively Bernoulli over $X_0$. \hfill \Box

As in the case of $\mathbb{Z}$-actions, with the same proof, we now obtain the following theorem.

4.3. **Theorem.** Every ergodic free $G$-system $X$ of positive entropy is dominant.

In view of Theorem 3.2 the following question naturally arises:

4.4. **Problem.** Can Theorem 3.2 be extended to all infinite countable amenable groups?

---

**References**

[1] Terrence Adams, *Genericity and rigidity for slow entropy transformations*, New York J. of Mathematics, New York J. Math. **27**, (2021), 393–416.

[2] R. L. Adler and P. C. Shields, *Skew products of Bernoulli shifts with rotations*, Israel J. Math. **12**, (1972), 215–222.

[3] Tim Austin, *Measure concentration and the weak Pinsker property*, Publ. Math. IHES **128**, (2018), 1–119.

[4] A. Connes, J. Feldman and B. Weiss, *An amenable equivalence relation is generated by a single transformation*, Ergodic Theory Dynamical Systems, **1**, (1981), no. 4, 431–450 (1982).

[5] Alexandre I. Danilenko and Kyewon K. Park, *Generators and Bernoullean factors for amenable actions and cocycles on their orbits*, Ergodic Theory Dynamical Systems, **22**, (2002), no. 6, 1715–1745.

[6] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic theory on compact spaces*, Lecture Notes in Math. **527**, Springer-Verlag, 1976.

[7] Eli Glasner, *Ergodic Theory via joinings*, Math. Surveys and Monographs, AMS, **101**, 2003.
AN ERGODIC SYSTEM IS DOMINANT EXACTLY WHEN IT HAS POSITIVE ENTROPY

[8] Eli Glasner, Jean-Paul Thouvenot and Benjamin Weiss, On some generic classes of ergodic measure preserving transformations, Tr. Mosk. Mat. Obs., 82, no. 1, MCCME, M., 2021, (Trans. Moscow Math. Soc., 82:1 (2021)), arXiv:2009.06901.

[9] Eli Glasner and Benjamin Weiss, Relative weak mixing is generic, Sci. China Math. 62, (2019), no. 1, 69–72.

[10] F. J. Hahn and W. Parry, Some characteristic properties of dynamical systems with quasi-discrete spectra, Math. Systems Theory, 2, (1968), 179–190.

[11] Christopher Hoffman, The behavior of Bernoulli shifts relative to their factors, Ergod. Th. & Dynam. Sys. (1999), 19, 1255–1280

[12] Anatole Katok and Jean-Paul Thouvenot, Slow entropy type invariants and smooth realization of commuting measure-preserving transformations, Ann. Inst. H. Poincaré Probab. Statist. 33, (1997), no. 3, 323–338.

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

[14] John C. Kieffer, A simple development of the Thouvenot relative isomorphism theory, Ann. Probab. 12, (1984), no. 1, 204–211.

[15] Dan Newton, Coalescence and spectrum of automorphisms of a Lebesgue space, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 19, (1971), 117–122.

[16] D. S. Ornstein, Bernoulli shifts with the same entropy are isomorphic, Advances in Math. 4, (1970), 337-352.

[17] Donald S. Ornstein and Benjamin Weiss, Ergodic theory of amenable group actions. I. The Rohlin lemma, Bull. Amer. Math. Soc. (N.S.), 2 (1980), no. 1, 161–164.

[18] D. J. Rudolph, Classifying the isometric extensions of a Bernoulli shift, J. Analyse Math. 34, (1978), 36–60 (1979).

[19] D. J. Rudolph, k-fold mixing lifts to weakly mixing isometric extensions, Ergodic Theory Dynam. Systems, 5, (1985), no. 3, 445–447.

[20] Daniel J. Rudolph and Benjamin Weiss, Entropy and mixing for amenable group actions, Ann. of Math. (2) 151, (2000), no. 3, 1119–1150.

[21] P. Shields, The ergodic theory of discrete sample paths, Graduate studies in mathematics 13, Amer. Math. Soc. 1996.

[22] J.-P. Thouvenot, Quelques propriétés des systéms dynamiques qui se décomposent en un produit de deux systèmes dont l’un est un schéma de Bernoulli, Israel J. of Math. 21, (1975), Nos. 2-3, 177–207.

[23] J.-P. Thouvenot, Remarques sur les systèmes dynamiques donnés avec plusieurs facteurs, Israel J. of Math. 21, (1975), no. 2-3, 215–232.

[24] W. A. Veech, A criterion for a process to be prime, Monatsh. Math. 94, (1982), 335–341.