ENTROPY AND SKEW PRODUCTS OVER IRRATIONAL ROTATIONS

NICANOR CARRASCO-VARGAS

1. Introduction

Let $T \curvearrowright (X, \mu)$ be an invertible measure preserving system. One can formalize a system in which T and its inverse T^{-1} are applied randomly and with equal probabilities by means of a skew product dynamical system. More concretely, let S be the shift transformation on $(\{-1,1\}^{\mathbb{Z}}, \nu)$, where ν is the uniform Bernoulli measure. Let $\tau\colon \{-1,1\}^{\mathbb{Z}} \to \{-1,1\}$ be defined by $\tau(y) = y(0)$. Consider the transformation $S \rtimes_\tau T$ acting on $\{-1,1\}^{\mathbb{Z}} \times X$ defined by

$$(y, x) \to (S(y), T^{\tau(y)}(x))$$

This transformation preserves the product measure $\nu \times \mu$. The measure preserving system $S \rtimes_\tau T \curvearrowright (\{-1,1\}^{\mathbb{Z}} \times X, \nu \times \mu)$ is classically called $[T, T^{-1}]$ system. Kalikow [5] proved that when T has positive entropy then the corresponding $[T, T^{-1}]$ system is not loosely Bernoulli, despite being a Kolmogorov automorphism. It is generally regarded as the first natural example of this property. Kalikow's result has been extended in several ways [14, 2], and generalizations of the same construction have continued to provide interesting examples in ergodic theory [3].

Despite the influence of Kalikow's result, it was unknown for a long time whether his result provides a single example, or many non-isomorphic examples, taking different T's. This was settled rather recently by Austin [1]. He proved that there is a highly nontrivial isomorphism invariant I such that $I(S \rtimes_\tau T) = h_\mu(T)$ for every T. A similar result holds [4] for the analogous construction with a one-sided shift $\{-1,1\}^{\mathbb{N}}$.

The objective of this work is investigating analogous results in the zero-entropy regime. We focus on the analogous construction in which the shift $S \curvearrowright \{-1,1\}^{\mathbb{Z}}$ is replaced by an irrational rotation, and the applications of T or T^{-1} are driven by the rotation and a step function on the circle.

Let $\mathbb{T} = [0, 1)$ with the Lebesgue measure m. Let $\tau\colon \mathbb{T} \to \{-1,1\}$ be defined by

$$\tau(x) = 1_{[0,1/2)}(x) - 1_{[1/2,1)}(x)$$

Let $S_\alpha \curvearrowright (\mathbb{T}, m)$ be the irrational rotation by α, and let $T \curvearrowright (X, \mu)$ be an arbitrary invertible measure-preserving system. We consider the transformation $S_\alpha \rtimes_\tau T \curvearrowright (\mathbb{T} \times X, m \times \mu)$ given by

$$(y, x) \mapsto (S_\alpha(y), T^{\tau(y)}(x))$$

Our results are proved under the following Diophantine condition about unbounded partial quotients along odd numbers.

Definition 1.1. *We say that an irrational $\alpha \in [0, 1)$ is good if for all $k \in \mathbb{N}$ we can find $p, q \in \mathbb{N}$ coprime with q odd and such that $|\alpha - p/q| < 1/(kq^2)$.*

The set of irrationals with this property has full Lebesgue measure. This follows from [12, Theorem 3.1] choosing $k_n = 2n + 1$ and $z = 1/k$.

The following result is a direct analogue to those proved in [1, 4].

Theorem 1.1. *Let α be a good irrational. Then there is an isomorphism invariant I such that for every $T \curvearrowright (X,\mu)$ we have*

$$I(S_\alpha \rtimes_\tau T) = h_\mu(T).$$

Instead of the entropy of T, one may ask whether other entropy-type invariants are remembered by the skew product $S_\alpha \rtimes_\tau T$. We prove the following result about slow entropy, which is an entropy-type invariant introduced by Katok and Thouvenot [7].

Theorem 1.2. *Let α be a good irrational, and let $\boldsymbol{a} = \{a_n(t)\}_{n\in\mathbb{N},t>0}$ be a scale. Then there is an isomorphism invariant I such that for every $T \curvearrowright (X,\mu)$ we have*

$$I(S_\alpha \rtimes_\tau T) = \overline{\mathrm{ent}}^{\boldsymbol{a}}_\mu(T).$$

The invariant in the first result is a variation of sequence entropy, and the invariant in the second result is a variation of slow entropy.

Our third main result is a form of rigidity statement under the assumption that T is a K-automorphism. In this case we can prove that not only entropy invariants of a transformation T are remembered by $S_\alpha \rtimes T$, but its complete isomorphism class (up to taking inverses). We remark that no results of this kind were previously known.

Theorem 1.3. *Let α be a good irrational. Let $T_1 \curvearrowright (X_1,\mathcal{B}_1,\mu_1)$ and $T_2 \curvearrowright (X_2,\mathcal{B}_2,\mu_2)$ be invertible K-automorphisms of standard probability spaces. Then $S_\alpha \rtimes_\tau T_1$ is isomorphic to $S_\alpha \rtimes_\tau T_2$ if and only if T_1 is isomorphic to T_2 or the inverse of T_2.*

2. Preliminaries

We write $A \Subset B$ when A is a finite subset of B. The symmetric difference of two sets is written $A \triangle B = (A\cup B) \smallsetminus (A\cap B)$.

Let $(X,\mathcal{B},\mu)$ be a standard probability space and let T be an automorphism. Let ξ be a finite collection of pairwise disjoint measurable subsets of X. Elements in ξ are called atoms. We often assume that ξ is a partition, but not always. We write $\xi \le \eta$ if every atom in ξ can be written as the union of atoms in η. In this situation we say that ξ is subordinate to η and that η refines ξ. We also let $\xi \vee \eta = \{P\cap Q : P\in\xi, Q\in\eta\}$. We write $T^{-i}(\xi) = \{T^{-i}(P) : P \in \xi\}$. Given $F \Subset \mathbb{Z}$ we write

$$\xi^F = \bigvee_{i\in F} T^{-i}(\xi).$$

The Shannon entropy of ξ is defined as $H(\xi) = \sum_{P\in\xi} -\mu(P)\log\mu(P)$.

Proposition 2.1. *Let ξ be a collection of pairwise disjoint atoms (not necessarily a partition) and suppose $\mu(\cup_{P\in\xi}P) = \delta$. Then $H(\xi) \le \delta\log(|\xi|)$.*

Proof. Jensen's inequality shows that $H(\xi)$ is maximized when $\mu(P) = \delta/|\xi|$ for all $P\in\xi$, and for such ξ we have $H(\xi) = \delta\log(|\xi|)$. □

Given $x \in X$ we denote by $\xi(x)$ the atom of ξ containing x (provided it exists).

3. Definitions of invariants

We now define the invariants with which we will work. Let $(X,\mathcal{B},\mu)$ be a standard probability space, and let $T \curvearrowright (X,\mathcal{B},\mu)$ be an automorphism. Let $\mathcal{F} = (F_n)$ be an arbitrary sequence of finite subsets of $\mathbb{Z}$.

3.1. **A variation of sequence entropy.** For a finite and measurable partition ξ of X we define

$$h_\mu^{\mathcal{F}}(T,\xi) = \limsup_{n\to\infty} \frac{1}{|F_n|} H(\xi^{F_n}).$$

Next, we take the supremum over all finite measurable partitions of X,

$$h_\mu^{\mathcal{F}}(T) = \sup_\xi h_\mu^{\mathcal{F}}(T,\xi).$$

This is an isomorphism invariant because we take supremum over partitions. It coincides with measure-theoretic entropy for $F_n = \{0,\dots,n-1\}$ (or any other Følner sequence). In this case we drop the "$\mathcal{F}$" and the corresponding quantities are denoted $h_\mu(T)$ and $h_\mu(T,\xi)$.

Proposition 3.1. *Let ξ and η be finite measurable partitions of X, and let $\rho(\xi,\eta) = H(\xi|\eta) + H(\eta|\xi)$ be their Rokhlin distance. Then*

$$|h_\mu^{\mathcal{F}}(T,\xi) - h_\mu^{\mathcal{F}}(T,\eta)| \le \rho(\xi,\eta)$$

Proof. Same argument as Lemma 1 in [11]. □

Proposition 3.2. *Let Ξ be a collection of finite partitions such that $\{\eta : \eta \le \xi,\ \xi \in \Xi\}$ is dense in the Rokhlin distance. Then*

$$h_\mu^{\mathcal{F}}(T) = \sup_{\xi\in\Xi} h_\mu^{\mathcal{F}}(T,\xi)$$

Proof. Follows from Proposition 3.1. □

3.2. **A variation of slow entropy.** Let ξ be a finite measurable partition of X. Given $F \Subset \mathbb{Z}$ we define the Hamming pseudo-distance $d^H_{\xi,F}$ on X by

$$d^H_{\xi,F}(x,y) = \frac{1}{|F|}|\{i \in F : \xi(T^i(x)) \ne \xi(T^i(y))\}|$$

Recall that $\xi(x)$ the atom in ξ containing x. The corresponding pseudo-ball is denoted $B^H_{\xi,F}(x,r) = \{y \in X : d^H_{\xi,F}(x,y) < r\}$. We define

$$S^H_\xi(T,F,\epsilon,\delta) = \min\{k \in \mathbb{N} : \text{there are } x_1,\dots,x_k \in X, \mu(\bigcup_{i=1}^k B^H_{\xi,F}(x_i,\epsilon)) > 1-\delta\}.$$

By a scale $\mathbf{a} = \{a_n(t)\}_{n\in\mathbb{N},t>0}$ we mean a sequence of monotone functions $a_n\colon (0,\infty) \to (0,\infty)$ such that $a_n(t)$ converges to infinity for every t. We define

$$\overline{\mathrm{ent}}_\mu^{\mathbf{a},\mathcal{F}}(T,\xi) = \lim_{\epsilon,\delta\to 0} \sup\{t \ge 0 : t = 0 \text{ or } \limsup_{n\to\infty} \frac{S^H(T,F_n,\epsilon,\delta)}{a_n(t)} > 0\}$$

$$\overline{\mathrm{ent}}_\mu^{\mathbf{a},\mathcal{F}}(T) = \sup_\xi \overline{\mathrm{ent}}_\mu^{\mathbf{a},\mathcal{F}}(T,\xi)$$

This is an isomorphism invariant because we take supremum over partitions. If $F_n = \{0,\dots,n-1\}$ then this invariant coincides with slow entropy for $\mathbb{Z}$-actions as introduced in [7]. In this case we drop the "$\mathcal{F}$" in the notation. The corresponding quantities are denoted $\overline{\mathrm{ent}}_\mu^{\mathbf{a}}(T,\xi)$ and $\overline{\mathrm{ent}}_\mu^{\mathbf{a}}(T)$.

Proposition 3.3. $S^H_\xi(T,F+n,\epsilon,\delta) = S^H_\xi(T,F,\epsilon,\delta)$ *for* $n \in \mathbb{Z}$.

Proof. From the definition of d^H_ξ we see that that $d^H_{\xi,F}(x,y) = d^H_{\xi,F+n}(T^{-n}(x),T^{-n}(y))$. A direct computation with this relation shows that $T^{-n}(B^H_{\xi,F}(x,\epsilon)) = B^H_{\xi,F+n}(T^{-n}(x),\epsilon)$. Since T preserves μ, it follows that $B^H_{\xi,F}(x,\epsilon)$ has the same measure as $B^H_{\xi,F+n}(T^{-n}(x),\epsilon)$. □

By this equality, in the computation of slow entropy we can also take $F_n = \{1, \dots, n\}$, or any interval in $\mathbb{Z}$ with n elements.

For classical slow entropy a finite generator theorem holds [7, Corollary 1]. Dropping the condition that $(F_n)_{n\in\mathbb{N}}$ is a Følner sequence, we find the following weaker form of generator theorem. Given two ordered partitions $\xi = (P_1, \dots, P_n)$ and $\eta = (Q_1, \dots, Q_n)$ we write

$$\rho_{sym}(\xi, \eta) = \sum_{i=1}^{n} \mu(P_i \triangle Q_i).$$

We extend this to a distance between unordered partitions by taking the minimal value over all orders, and adding empty atoms if necessary so that both have the same cardinality. The distance obtained is still denoted ρ_{sym}.

Proposition 3.4. *Let Ξ be a collection of finite partitions such that $\{\eta : \eta \leq \xi,\ \xi \in \Xi\}$ is dense in the distance ρ_{sym}. Then*

$$\overline{\mathrm{ent}}_{\mu}^{\boldsymbol{a},\mathcal{F}}(T) = \sup_{\xi\in\Xi} \overline{\mathrm{ent}}_{\mu}^{\boldsymbol{a},\mathcal{F}}(T, \xi)$$

Proof. From [7, Equation 1.3] we have that for two finite measurable partitions ξ and η

$$S_{\eta}^{H}(T, F, \epsilon + \rho_{sym}(\xi, \eta)^{1/2}, \delta + \rho_{sym}(\xi, \eta)^{1/2}) \leq S_{\xi}^{H}(T, F, \epsilon, \delta), \quad F \Subset \mathbb{Z}.$$

The claim follows from this inequality with $F = F_n$, and the fact $\overline{\mathrm{ent}}_{\mu}^{\mathbf{a},\mathcal{F}}(T, \eta) \leq \overline{\mathrm{ent}}_{\mu}^{\mathbf{a},\mathcal{F}}(T, \xi)$ when $\eta \leq \xi$. □

4. Preliminary computations on $\mathbb{T}$

We endow $\mathbb{T} = [0, 1)$ with the distance $d_{\mathbb{T}}(x, y) = \min\{d_{\mathbb{R}}(x+n, y+m) : n, m \in \mathbb{Z}\}$, the Borel sigma-algebra $\mathcal{B}_{\mathbb{T}}$, and the Lebesgue measure m. By an arc in $\mathbb{T}$ we mean a set of the form $[a, b)$ or $[0, a) \cup [b, 1)$, $a < b \in \mathbb{T}$. The length of an arc is defined as $b - a$ in the first case, and as $a + (1 - b)$ in the second case.

The fractional part of $x \in \mathbb{R}$ in $[0, 1)$ is denoted $\langle x \rangle$. For $a, b \in \mathbb{T}$ we use the notation $a +_{\mathbb{T}} b := \langle a + b \rangle$.

Let α be a good irrational. In this section we consider the transformation $S_\alpha \curvearrowright (\mathbb{T}, \mathcal{B}_{\mathbb{T}}, m)$ given by $x \mapsto \langle x + \alpha \rangle = x +_{\mathbb{T}} \alpha$. Thus if ξ is a partition of $\mathbb{T}$ and $F \Subset \mathbb{Z}$ then ξ^F denotes $\vee_{i\in F} S_{\alpha}^{-i}(\xi)$.

Proposition 4.1. *Let $\zeta = \{[0, 1/2), [1/2, 1)\}$. Suppose that $|\alpha - p/q|< 1/(kq^2)$, where $p, q \in \mathbb{N}$ are coprime, q is odd, and $k \geq 2$. Then for every $I \in \zeta^{\{0,\dots,q-1\}}$ we have*

$$|\mathrm{length}(I) - \frac{1}{2q}| \leq \frac{2}{kq}. \tag{1}$$

Proof. Let

$$E = \{S_{\alpha}^{-i}(0) : i = 0, \dots, q - 1\}, \quad E' = \{S_{\alpha}^{-i}(1/2) : i = 0, \dots, q - 1\}$$

Writing $\zeta^{\{0,\dots,q-1\}} = \bigvee_{i=0}^{q-1} S_{\alpha}^{-i}(\zeta)$ we see that $\zeta^{\{0,\dots,q-1\}}$ equals the collection of intervals $[a, b)$ whose endpoints a, b are consecutive elements in $E \cup E'$ and $a < b$, plus one atom of the form $[0, a) \cup [b, 1)$ where a, b are the leftmost and rightmost elements in $E \cup E'$. We define

$$C = \{\frac{j}{q} : j = 0, \dots, q - 1\}.$$

We claim that every element in E is at distance at most $1/(kq)$ from a unique element in C. To see this, we multiply the relation $|\alpha - p/q|< 1/(kq^2)$ by $i \in \{0, \dots, q - 1\}$ and find

$$|i\alpha - i\frac{p}{q}| < \frac{i}{kq^2} \leq \frac{1}{kq}$$

Since p and q are coprime, we can find $j \in \{0, \dots, q-1\}$ so that $ip \equiv j \mod q$. Therefore $\langle i\frac{p}{q}\rangle = \frac{j}{q}$. Taking fractional part in the equation above, we see that the distance $d_{\mathbb{T}}$ between $S_\alpha^i(0)$ and j/q is strictly smaller than $1/(kq)$. The uniqueness of j follows from the fact that elements in C can not be too close (if $j' \in \{0, \dots, q-1\}$ is such that $d_{\mathbb{T}}(S_\alpha^i(0), j'/q) < 1/(kq)$ then by triangle inequality we see that $d_{\mathbb{T}}(j/q, j'/q) < \frac{1}{kq} + \frac{1}{kq} < \frac{1}{q}$, which implies $j = j'$ as any two different elements in C have distance $d_{\mathbb{T}}$ at least $1/q$). Let

$$C' = \{\frac{j}{2q} : j = 0, \dots, 2q-1,\ j \text{ odd}\}.$$

Since q is odd, the isometry $x \mapsto \langle x + 1/2\rangle$ sends E to E' and sends C to C'. Therefore every element in E' is at distance at most $1/(kq)$ from a unique element in C'. We obtain a bijection $E \cup E' \to C \cup C'$ which displaces every point by at most $1/(kq)$ in $d_{\mathbb{T}}$. Equation (1) follows from this fact plus the observation that $C \cup C'$ is equal to $\{\frac{j}{2q} : j = 0, \dots, 2q-1\}$. □

Definition 4.1. *Given $x \in \mathbb{T}$ we write $\tau^0(x) = 0$,*

$$\tau^n(x) = \tau(x) + \cdots + \tau(S_\alpha^{n-1}(x)), \quad n \geq 1,$$

$$\tau^F(x) = \{\tau^i(x) : i \in F\}, \quad F \Subset \mathbb{N}.$$

If $x \to \tau^n(x)$ is constant over a set $P \subset \mathbb{T}$, then we define $\tau^n(P) := \tau^n(x)$ for any $x \in P$. We follow a similar convention regarding $\tau^F(P)$.

Definition 4.2. *For $n \geq 1$ choose $p(n), q(n) \in \mathbb{N}$ coprime, with $q(n)$ odd, and such that*

$$|\alpha - \frac{p(n)}{q(n)}| < \frac{1}{2^n q(n)^2}. \tag{2}$$

This is possible by Definition 1.1. We also define

$$F_n = \{q(n), 2q(n), \dots, nq(n)\}$$
$$G_+(n) = \{x \in \mathbb{T} : \tau^{iq(n)}(x) = i,\ \forall i = 1, \dots, n\}$$
$$G_-(n) = \{x \in \mathbb{T} : \tau^{iq(n)}(x) = -i,\ \forall i = 1, \dots, n\}$$
$$G(n) = G_+(n) \cup G_-(n)$$

Proposition 4.2. *For all $n \geq 1$ the sets $G_+(n)$ and $G_-(n)$ have measure at least $1/2 - n/2^{n-1}$.*

Proof. Set $n \geq 1$, $p = p(n)$ and $q = q(n)$ as in Definition 4.2. Thanks to the Denjoy-Koksma inequality we have $|\tau^q(x)| \leq 2$ for all $x \in \mathbb{T}$. Since q is odd, $\tau^q(x)$ is the sum of an odd number of 1's and -1's, and thus $\tau^q(x)$ is also odd. Since the only odd integers between -2 and 2 are -1 and 1, we obtain

$$\forall x \in \mathbb{T} \quad \tau^q(x) \in \{1, -1\}. \tag{3}$$

An elementary induction on $m \geq 1$ shows that the value of τ^m on consecutive arcs of $\zeta^{\{0,\dots,m-1\}}$ differs by 2. Writing $\tau^q(x) = \tau(x) + \cdots + \tau(S_\alpha^{q-1}(x))$ we see that $\tau^q(x)$ is constant on each atom in $\zeta^{\{0,\dots,q-1\}}$. Thus τ^q alternates between -1 and 1 between consecutive arcs in $\zeta^{\{0,\dots,q-1\}}$. There are $2q$ arcs in $\zeta^{\{0,\dots,q-1\}}$, so τ^q takes value 1 on q of them and takes value -1 on the remaining q.

Multiplying the equation $|\alpha - p/q| < 1/(2^n q^2)$ by q and then taking fractional part, we see that $S_\alpha^q(0)$ is at distance at most $1/(2^n q)$ from 0 in $d_{\mathbb{T}}$. Since S_α^q is an isometry we see that S_α^q displaces any point by at most $1/(2^n q)$ in $d_{\mathbb{T}}$.

Take an arc $I = [a, b)$ in $\zeta^{\{0,\dots,q-1\}}$ such that τ^q takes value 1 on I, and let $x \in I$. If the iterates $S_\alpha^{iq}(x)$ fall in the same arc I for all $i = 0, \dots, n-1$ then we have $x \in G_+(n)$. To see this it suffices to write $\tau^{iq}(x) = \tau^q(x) + \tau^q(S_\alpha^q(x)) + \cdots + \tau^q(S_\alpha^{(i-1)q}(x))$ and note that then we have the sum of i 1's. Since each application of S_α^q produces a displacement of at most $1/(2^n q)$, the iterates $S^q(x), S^{2q}(x), \ldots, S^{(n-1)q}(x)$ are all within distance at most $(n-1)/(2^n q)$ of x. This shows that all elements in I at distance at least $(n-1)/(2^n q)$ from the endpoints of the arc belong to $G_+(n)$. Since the length of I is at least $1/2q - 1/(2^{n-1}q)$ (Proposition 4.1), we see that there is a sub-interval of I of length at least $1/2q - 1/(2^{n-1}q) - 2(n-1)/(2^n q) = 1/(2q) - n/(2^{n-1}q)$ contained in $G_+(n)$. The same is true if I is an arc of the form $[0,a) \cup [b,1)$. Since there are exactly q arcs where τ^q takes value 1 we obtain $m(G_+(n)) \geq 1/2 - n/2^{n-1}$.

The claim about $m(G_-(n))$ follows from the same argument with arcs where τ^q equals -1. □

Proposition 4.3. *Let ξ be a partition of $\mathbb{T}$ whose atoms are arcs. There is $n_0 \in \mathbb{N}$ so that for $n \geq n_0$ we can find a pairwise disjoint collection ξ_n of subsets of $\mathbb{T}$ such that:*

1. *For every $Q \in \xi_n$ there is a unique $P \in \xi$ with $Q \subset P$. Furthermore, $m(P \smallsetminus Q) < n/2^{n-1}$.*
2. *Q is a subset of an atom of ξ^{F_n}*
3. *Every $y \in Q$ satisfies $y +_{\mathbb{T}} i\alpha \in P$ for every $i \in F_n$.*

Furthermore $H(\xi^{F_n})/|F_n| \to 0$ as $n \to \infty$.

Proof. By our choice of $q(n)$, $S_\alpha^{q(n)}$ displaces points by distance at most $1/(2^n q(n))$, which is at most $1/2^n$. Thus for $i = 1, \ldots, n$ the endpoints of arcs in $S_\alpha^{-iq(n)}\xi$ are very close to endpoints of arcs in ξ, at distance at most $n/2^n$. Suppose that n is large enough so that $n/2^{n-1}$ is smaller than the length of every arc in ξ. Given an arc $P \in \xi$ we can shrink both endpoints by $n/2^n$ and obtain an arc $Q \subset P$ with $m(Q) \geq m(P) - n/2^{n-1}$. The previous observation shows that then Q is contained in a single atom of ξ^{F_n}. We obtain ξ_n following this procedure for each $P \in \xi$.

The claim that $H(\xi^{F_n})/|F_n| \to 0$ follows from an elementary computation using Proposition 2.1, and the observation that ξ^{F_n} has a subcollection of $|\xi|$ atoms covering most of $\mathbb{T}$ in measure (at least $1 - |\xi| n/2^{n-2}$). □

We now define a class of useful partitions of $\mathbb{T}$ associated to the level sets of $\tau^n \colon \mathbb{T} \to \mathbb{Z}$.

Definition 4.3. *For $n \geq 0$ we define $L(n)$ as the finite and measurable partition of $\mathbb{T}$ such that two elements $x, y \in \mathbb{T}$ are in the same atom if and only if $\tau^n(x) = \tau^n(y)$. Thus atoms in $L(n)$ are equal to $\{x \in \mathbb{T} : \tau^n(x) = k\}$ for some $k \in \mathbb{Z}$. Given $F \Subset \mathbb{Z}$ we define $L(F) = \bigvee_{n \in F} L(n)$.*

If $P \subset \mathbb{T}$ is contained in some atom of $L(n)$ then we define $\tau^n(P) \in \mathbb{Z}$ as $\tau^n(x)$ for some $x \in P$. Similarly, if P is contained in some atom of $L(F)$ then we define $\tau^F(P) \Subset \mathbb{Z}$ as $\tau^F(x)$ for some $x \in P$.

Remark 4.1. *The partition $L(F_n)$ contains both $G_+(n)$ and $G_-(n)$ as atoms.*

Proposition 4.4. *The number of atoms in $L(F_n)$ is at most $2n$. Thus $H(L(F_n))/|F_n| \to 0$.*

Proof. In the next argument we use some ideas from the proof of Proposition 4.2. The atom in $L(F_n)$ containing an element y is determined by the sequence $(\tau^{iq(n)}(y))_{i=1}^n$. Writing $\tau^{iq(n)}(y) = \sum_{j=0}^{i-1} \tau^{q(n)}(S_\alpha^{jq(n)}(y))$ for each $i = 1, \ldots, n$, we see that $(\tau^{iq(n)}(y))_{i=1}^n$ is determined by the sequence $(\tau^{q(n)}(S_\alpha^{jq(n)}(y))_{j=0}^{n-1}$. This is a sequence of 1's and -1. Its values depend on the arc from $\zeta^{\{0,\ldots,q(n)-1\}}$ containing each $S_\alpha^{jq(n)}(y)$, $j = 0, \ldots, n-1$. We observe that $(S_\alpha^{jq(n)}(y))_{j=0}^{n-1}$ changes of arc at most once. This follows by comparing the length of the arcs with the displacement produced by $S^{q(n)}$ (as in the proof of Proposition 4.2). Therefore the sequence $(\tau^{q(n)}(S_\alpha^{jq(n)}(y))_{j=0}^{n-1}$ of 1's and -1's has at most one sign change. There are at most $2n$ such sequences, determined by the starting value and the position of the sign change.

Therefore $H(L(F_n))/|F_n| \leq \log(|L(F_n)|)/n \leq \log(2n)/n \to 0$ as $n \to \infty$. Alternatively, it is easy to derive $H(L(F_n))/|F_n| \to 0$ from the fact that $L(F_n)$ has two atoms with most of the measure, but we shall need the linear bound on the cardinality of $L(F_n)$ later. □

5. Estimates related to $h^{\mathcal{F}}(S_\alpha \rtimes_\tau T)$

In this section we prove Theorem 1.1 and prove further estimates that will be needed for the proof of Theorem 1.3. In this section we set $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$ as the sequence from Definition 4.2.

We fix for this section an arbitrary invertible measure-preserving system $T \curvearrowright (X, \mathcal{B}, \mu)$. We work with the probability space $(\mathbb{T} \times X, \mathcal{B}_{\mathbb{T}} \otimes \mathcal{B}, m \times \mu)$ and the transformation $S_\alpha \rtimes_\tau T$.

Proposition 5.1. *Let ξ be a partition of $\mathbb{T}$. Then $H(\xi) = H(\xi \times \{X\})$.*

Proof. $H(\xi \times \{X\})$ is defined as $\sum_{P\in\xi} -m \times \mu(P \times X) \log(m \times \mu(P \times X))$, which is clearly equal to $\sum_{P\in\xi} -m(P) \log(m(P))$, which equals $H(\xi)$ by definition. □

If β is a partition of $\mathbb{T} \times X$, then β_y denotes the partition of X defined by the condition that $x, x' \in X$ are in the same atom of β_y if (y, x) and (y, x') belong to the same atom of β. In other words, $\beta_y(x) = \beta_y(x') \iff \beta(y, x) = \beta(y, x')$. If $P \subset \mathbb{T}$ is such that $y \to \beta_y$ is constant over $y \in P$, then we also denote $\beta_P = \beta_y$ for some $y \in P$.

Proposition 5.2. *Let ξ be a partition of $\mathbb{T}$ whose atoms are arcs, and let η be a partition of X. Let β be a partition of $\mathbb{T} \times X$ so that $\xi \times \{X\} \leq \beta \leq \xi \times \eta$. It follows that $y \to \beta_y$ is constant over $y \in P$, for each $P \in \xi$. Thus we can write $\beta = \bigcup_{P\in\xi}\{P\} \times \beta_P$. Define γ_n by*

$$\gamma_n = \{P \times Q : P \in L(F_n) \vee \xi_n, Q \in \beta_P^{\tau^{F_n}(P)}\}$$

Here ξ_n is the collection of pairwise disjoint atoms associated to ξ from Proposition 4.3. Thus γ_n is a collection of disjoint atoms in $\mathbb{T} \times X$, and it is not a partition because ξ_n does not cover $\mathbb{T}$. Observe that γ_n is well-defined because given $P \in L(F_n) \vee \xi_n$, both $y \to \beta_y$ and $y \to \tau^{F_n}(y)$ are constant over $y \in P$. Then

$$\lim_{n\to\infty} H(\gamma_n)/|F_n| = \sum_{P\in\xi} m(P) h_\mu(T, \beta_P)$$

Proof. We have

$$H(\gamma_n) = \sum_{P\in L(F_n)\vee\xi_n} \sum_{Q\in\beta_P^{\tau^{F_n}(P)}} -m \times \mu(P \times Q) \log(m \times \mu(P \times Q))$$

By a direct computation this can be rearranged as

$$= \sum_{P\in L(F_n)\vee\xi_n} m(P) H(\beta_P^{\tau^{F_n}(P)})$$

We separate this sum as

$$= \sum_{\substack{P\in L(F_n)\vee\xi_n \\ P\subset G_+(n)}} m(P) H(\beta_P^{\tau^{F_n}(P)}) + \sum_{\substack{P\in L(F_n)\vee\xi_n \\ P\subset G_-(n)}} m(P) H(\beta_P^{\tau^{F_n}(P)}) + \sum_{\substack{P\in L(F_n)\vee\xi_n \\ P\subset \mathbb{T}\setminus G(n)}} m(P) H(\beta_P^{\tau^{F_n}(P)})$$

Recall that $L(F_n)$ has $G_+(n)$ and $G_-(n)$ as atoms, and they cover most of $\mathbb{T}$ in measure. Thus $P \in L(F_n) \vee \xi_n$ satisfies $P \subset G_+(n)$ if and only if P belongs to $\{G_+(n)\} \vee \xi_n$. Furthermore, in this case we have $\tau^{F_n}(P) = \{1, \ldots, n\}$. A similar observation applies to $G_-(n)$. Thus we can continue our computation as

$$= \sum_{P\in\{G_+(n)\}\vee\xi_n} m(P) H(\beta_P^{\{1,\ldots,n\}}) + \sum_{P\in\{G_-(n)\}\vee\xi_n} m(P) H(\beta_P^{\{-n,\ldots,-1\}}) + \sum_{\substack{P\in L(F_n)\vee\xi_n \\ P\subset \mathbb{T}\setminus G(n)}} m(P) H(\beta_P^{\tau^{F_n}(P)})$$

Since $H(\beta_P^{\{-n,\dots,-1\}}) = H(\beta_P^{\{1,\dots,n\}})$, this is equal to

$$= \sum_{P\in\{G_+(n)\}\vee\xi_n} m(P)H(\beta_P^{\{1,\dots,n\}}) + \sum_{P\in\{G_-(n)\}\vee\xi_n} m(P)H(\beta_P^{\{1,\dots,n\}}) + \sum_{\substack{P\in L(F_n)\vee\xi_n\\ P\subset\mathbb{T}\smallsetminus G(n)}} m(P)H(\beta_P^{\tau^{F_n}(P)})$$

The assumption $\xi\times\{X\}\le\beta$ ensures that for any $P\in\xi_n$, we have $\beta_{P\cap G_+(n)} = \beta_{P\cap G_-(n)}$. Thus we can combine the sum of the first two terms and obtain the following

$$= \sum_{P\in\xi_n} m(P\cap G(n))H(\beta_P^{\{1,\dots,n\}}) + \sum_{\substack{P\in L(F_n)\vee\xi_n\\ P\subset\mathbb{T}\smallsetminus G(n)}} m(P)H(\beta_P^{\tau^{F_n}(P)})$$

Recall that ξ_n contains $|\xi|$ atoms, each one contained in one atom of ξ. As n increases, the measure of the symmetric differences decreases to zero. Thus dividing by $n=|F_n|$ and taking the limit we obtain

$$\lim_{n\to\infty}\frac{1}{n}\sum_{P\in\xi_n} m(P\cap G(n))H(\beta_P^{\{1,\dots,n\}}) = \lim_{n\to\infty}\frac{1}{n}\sum_{P\in\xi} m(P)H(\beta_P^{\{1,\dots,n\}}) = \lim_{n\to\infty}\sum_{P\in\xi} m(P)H(\beta_P^{\{1,\dots,n\}})/n$$
$$= \sum_{P\in\xi} m(P)h_\mu(T,\beta_P)$$

To finish the proof we observe that the following term vanishes.

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\substack{P\in L(F_n)\vee\xi_n\\ P\subset\mathbb{T}\smallsetminus G(n)}} m(P)H(\beta_P^{\tau^{F_n}(P)}) \le \lim_{n\to\infty}\frac{1}{n}\sum_{\substack{P\in L(F_n)\vee\xi_n\\ P\subset\mathbb{T}\smallsetminus G(n)}} m(P)n\log|\eta| = \lim_{n\to\infty}(1-m(G(n)))\log|\eta| = 0$$

Here we applied the trivial bounds $H(\beta_P^{\tau^{F_n}(P)}) \le |\tau^{F_n}(P)|\log|\beta_P| \le n\log|\eta|$, which are valid because $|\tau^{F_n}(P)|\le n$ and β_P has at most $|\eta|$ atoms. ☐

Proposition 5.3. *Let ξ be a partition of $\mathbb{T}$ whose elements are arcs, and let η be a partition of X. Let β be a partition of $\mathbb{T}\times X$ such that $\beta\le\xi\times\eta$. Then*

$$h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta) = \int_{\mathbb{T}} h_\mu(T,\beta_y)dm(y)$$

Proof. We first argue that it suffices to prove the claim under the additional assumption that $\xi\times\{X\}\le\beta$, as if this is not the case then we can always replace β by $\beta\vee(\xi\times\{X\})$. First observe that the integral expressions associated to β and $\beta\vee(\xi\times\{X\})$ are equal, as $(\beta\vee(\xi\times\{X\}))_y=\beta_y$ for all $y\in\mathbb{T}$. We also have $h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta) = h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta\vee(\xi\times\{X\}))$. The inequality $\le$ is simply because $\beta\vee(\xi\times\{X\})$ refines β. For the other inequality observe that $(\beta\vee(\xi\times\{X\})^{F_n}$ is equal to $\beta^{F_n}\vee(\xi\times\{X\})^{F_n}$. But $H((\xi\times\{X\})^{F_n}) = H(\xi^{F_n}\times\{X\}) = H(\xi^{F_n})$, and we proved in Proposition 4.3 that $H(\xi^{F_n})/|F_n|\to 0$ as $n\to\infty$.

We now provide a proof under the assumption that $\xi\times\{X\}\le\beta$. This assumption allows us to use Proposition 5.2. We write β as $\beta=\bigcup_{P\in\xi}\xi\times\beta_P$, where $\{\beta_P : P\in\xi\}$ is a collection of partitions of X, each one less fine than η. A direct computation shows that then

$$\int_{\mathbb{T}} h_\mu(T,\beta_y)dm(y) = \sum_{P\in\xi} m(P)h_\mu(T,\beta_P) \tag{4}$$

In what follows we will work with $\beta^{F_n}\vee((L(F_n)\vee\xi_n)\times\{X\})$, where ξ_n is defined in Proposition 4.3. Recall from Proposition 4.4 and Proposition 4.3 that $H(L(F_n))/n$ and $H(\xi_n)/n$ converge to zero as $n\to\infty$. Then the same reasoning used in the previous paragraph shows that

$$\limsup_{n\to\infty} H(\beta^{F_n})/n = \limsup_{n\to\infty} H(\beta^{F_n}\vee((L(F_n)\vee\xi_n)\times\{X\}))/n.$$

By Equation (4) and Proposition 5.2, in order to prove that β has the property in the statement, it is sufficient to prove the equality

$$\limsup_{n\to\infty} H(\gamma_n)/n = \limsup_{n\to\infty} H(\beta^{F_n} \vee ((L(F_n)\vee\xi_n)\times\{X\}))/n. \tag{5}$$

We first prove that

$$\gamma_n \subset ((L(F_n)\vee\xi_n)\times\{X\})\vee\beta^{F_n}.$$

Let $P_0\times Q_0$ be an arbitrary atom in γ_n. By definition of γ_n we have $P_0\in L(F_n)\vee\xi_n$. By definition of ξ_n, we can find a unique $P\in\xi$ containing P_0. By definition of γ_n, Q_0 can be written as

$$Q_0 = \bigcap_{i\in F_n} T^{-\tau^i(P_0)}(Q_i),$$

where $(Q_i)_{i\in F_n}$ are atoms in $\beta_P=\beta_{P_0}$.

Take $(y,x)\in P_0\times Q_0$. Since P_0 is contained in an atom from ξ_n, we have that $\beta_{y+_{\mathbb{T}}i\alpha}=\beta_P$ for $i\in F_n$. It follows that all iterates $(S_\alpha\rtimes_\tau T)^i(y,x)$, $i\in F_n$, fall into elements of β of the form $\{P\}\times\beta_P$, where P is the atom of ξ defined in the previous paragraph. Furthermore, the itinerary described over such atoms from β is fully determined by Q_0. It is also independent of the specific choice of $(y,x)\in P_0\times Q_0$. This shows that we have an inclusion

$$P_0\times Q_0 \subset (P_0\times\{X\})\cap\bigcap_{i\in F_n}(S_\alpha\rtimes_\tau T)^{-i}(P\times Q_i)$$

The same argument, but applied backwards, shows that in fact this is a set-theoretic equality.

$$P_0\times Q_0 = (P_0\times\{X\})\cap\bigcap_{i\in F_n}(S_\alpha\rtimes_\tau T)^{-i}(P\times Q_i)$$

The element at the right is an atom from $((L(F_n)\vee\xi_n)\times\{X\})\vee\beta^{F_n}$. This shows $\gamma_n\subset ((L(F_n)\vee\xi_n)\times\{X\})\vee\beta^{F_n}$. From this it follows inequality $\leq$ in eq. (5).

Let $\gamma_n'\subset((L(F_n)\vee\xi_n)\times\{X\})\vee\beta^{F_n}$ be the collection of atoms not in γ_n. Since γ_n and γ_n' are disjoint,

$$H(((L(F_n)\vee\xi_n)\times\{X\})\vee\beta^{F_n}) = H(\gamma_n)+H(\gamma_n').$$

In order to prove $\geq$ in Equation (5), it is sufficient to show that $H(\gamma_n')/n\to 0$ as $n\to\infty$. The union of atoms in γ_n contains $(\cup_{P\in\xi_n}P)\times\{X\}$, whose measure $m\times\mu$ is equal to $m(\cup_{P\in\xi_n}P)$. Thus the union of the atoms in γ_n' is a set with measure at most $1-m(\cup_{P\in\xi_n}P)$. Then by Proposition 2.1 we have

$$H(\gamma_n')\leq(1-m(\cup_{P\in\xi_n}P))\log|\gamma_n'|$$

We recall from Proposition 4.3 that $(1-m(\cup_{P\in\xi_n}P)\to 0$ as $n\to\infty$. Thus we only need to prove that $\log(|\gamma_n'|)/n$ is bounded. The cardinality of γ_n' is at most $|L(F_n)\vee\xi_n)\times\{X\})\vee\beta^{F_n}|$. By Proposition 4.3 and Proposition 4.4 we can bound this quantity as $|L(F_n)\vee\xi_n)\times\{X\}|\cdot|\beta|^{|F_n|}\leq |L(F_n)|\cdot|\xi_n|\cdot|\beta|\cdot^n\leq 2n|\xi|\cdot|\beta|^n$. Therefore $H(\gamma_n')/n\to 0$ as claimed. ☐

We are finally ready to prove Theorem 1.1.

Proposition 5.4 (Theorem 1.1). *We have* $h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T)=h_\mu(T)$.

Proof. Let ξ be a finite measurable partition of $\mathbb{T}$ whose atoms are arcs, and let η be a finite measurable partition of X. Choosing $\beta=\xi\times\eta$ in Proposition 5.3 we find $h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\xi\times\eta)=h_\mu(T,\eta)$. Taking supremum over this kind of partitions we find by Proposition 3.2 that $h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T)=h_\mu(T)$. ☐

Theorem 5.1. *Let β be an arbitrary finite measurable partition of $\mathbb{T}\times X$. Then*

$$h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta)=\int_{\mathbb{T}} h_\mu(T,\beta_y)dm(y)$$

Proof. Suppose that we have a sequence of partitions $(\beta_n)_{n\in\mathbb{N}}$ with the following properties.

- Each β_n is subordinate to a product partition of $\mathbb{T}\times X$.
- $\lim_{n\to\infty}\rho(\beta,\beta_n)=0$
- For almost every $y\in\mathbb{T}$ we have $\lim_{n\to\infty}\rho(\beta_y,(\beta_n)_y)=0$

Then the statement follows from the following equalities.

$$\begin{aligned} h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta) &= \lim_{n\to\infty} h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta_n) \\ &= \lim_{n\to\infty}\int_{\mathbb{T}} h_\mu(T,(\beta_n)_y)dm(y) \\ &= \int_{\mathbb{T}} h_\mu(T,\beta_y)dm(y) \end{aligned}$$

The first equality is because $|h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta_n)-h^{\mathcal{F}}_{m\times\mu}(S_\alpha\rtimes_\tau T,\beta)|\le\rho(\beta,\beta_n)$ (Proposition 3.1). The second equality is by Proposition 5.3 and the assumption that each β_n is subordinate to a product partition. The third equality is the dominated convergence theorem. Indeed, $|h_\mu(T,\beta_y)-h_\mu(T,(\beta_n)_y)|\le\rho(\beta,\beta_n)\to 0$ for almost every y, the sequence of functions $n\to(y\mapsto h_\mu(T,(\beta_n)_y))$ converges pointwise almost everywhere to $y\mapsto h_\mu(T,\beta_y)$ (Proposition 3.1 again). To see that the sequence of functions we are considering is dominated by an integrable function, note that $|\beta_y|\le|\beta|$ for all y, so $h_\mu(T,\beta_y)\le\log|\beta|$ for all y, and $y\to\log|\beta|$ is integrable over $\mathbb{T}$.

It only remains to construct a sequence $(\beta_n)_{n\in\mathbb{N}}$ with the desired properties. Write $\beta=\{P_1,\dots,P_k\}$. If we endow $\mathcal{B}_{\mathbb{T}}\otimes\mathcal{B}$ with the distance given by the measure of symmetric difference $(P,Q)\to m\times\mu(P\triangle Q)$, then those sets that can be written as unions of rectangles are dense (here a rectangle is a set of the form $P\times Q$, $P\in\mathcal{B}_{\mathbb{T}}$, $Q\in\mathcal{B}$). It follows that for each $n\in\mathbb{N}$ we can find k sets $P_1(n),\dots,P_k(n)$, each of them a unions of rectangles, such that they approximate each atom in β in the sense that

$$m\times\mu(P_1\triangle P_1(n))<2^{-n},\dots,m\times\mu(P_k\triangle P_k(n))<2^{-n}.$$

We define a partition $\beta_n=\{Q_1(n),\dots,Q_k(n)\}$ by setting

$$Q_1(n)=P_1(n),\ Q_2(n)=P_2(n)\smallsetminus P_1(n),\dots,Q_k(n)=P_k(n)\smallsetminus\cup_{i=1}^{k-1}P_i(n).$$

Since $P_1,\dots,P_k$ are pairwise disjoint, each symmetric difference $P_i(n)\triangle Q_i(n)$ is a subset of $\cup_{i=1}^{n}P_i\triangle P_i(n)$, which has measure at most $k2^{-n}$. It follows that $m\times\mu(P_i\triangle Q_i(n))<k2^{-n}$ for each $i=1,\dots,k$. Therefore

$$\sum_{i=1}^{k} m\times\mu(P_i\triangle Q_i(n))<k^2 2^{-n}$$

This shows that $\rho_{sym}(\beta,\beta_n)<k^2 2^{-n}$ and therefore $\rho_{sym}(\beta,\beta_n)\to 0$ as $n\to\infty$.

We emphasize that ρ_{sym} is defined as a minimum value over orderings of the partitions, so we can always take a convenient order and use it as upper bound. Naturally, in this case the convenient order is the one which with β_n was constructed.

If $P\subset\mathbb{T}\times X$ and $y\in\mathbb{T}$, we write $P_y=P\cap\{y\}\times X$. Therefore $\beta_y=\{(P_1)_y,\dots,(P_n)_y\}$, and similarly $(\beta_n)_y=\{(Q_1(n))_y,\dots,(Q_k(n))_y\}$. Choosing the order which $(\beta_n)_y$ inherits from β_n, we see that

$$\rho_{sym}(\beta_y,(\beta_n)_y)\le\sum_{i=1}^{k}\mu((P_i)_y\triangle(Q_i(n))_y)$$

Integrating this inequality over $y \in \mathbb{T}$ we obtain

$$\begin{aligned}\int_{\mathbb{T}} \rho_{sym}(\beta_y, (\beta_n)_y) dm(y) &\leq \int_{\mathbb{T}} \sum_{i=1}^{k} \mu((P_i)_y \triangle (Q_i(n))_y) dm(y) \\ &= \sum_{i=1}^{k} \int_{\mathbb{T}} \mu((P_i)_y \triangle (Q_i(n))_y) dm(y)\end{aligned}$$

Using the relation $(P_i)_y \triangle Q_i(n)_y = (P_i \triangle Q_i(n))_y$, Fubini's Theorem, and the bound from before, we find

$$= \sum_{i=1}^{k} \int_{\mathbb{T}} \mu((P_i \triangle Q_i(n))_y) dm(y) = \sum_{i=1}^{k} m \times \mu(P_i \triangle Q_i(n)) < k^2 2^{-n}$$

It follows that $\int_{\mathbb{T}} \rho_{sym}(\beta_y, (\beta_n)_y) dm(y) \to 0$ as $n \to \infty$. Recall that if a sequence of measurable functions converges in L^1 distance, then along a subsequence we have pointwise convergence almost everywhere [13, Theorem 3.12]. Therefore replacing $(\beta_n)_{n\in\mathbb{N}}$ by a subsequence if needed we have the property $\rho_{sym}(\beta_y, (\beta_n)_y) \to 0$ for almost every $y \in \mathbb{T}$. In this manner we obtain a sequence $(\beta_n)_{n\in\mathbb{N}}$ with the following properties.

- Each β_n is subordinate to a product partition of $\mathbb{T} \times X$.
- $\lim_{n\to\infty} \rho_{sym}(\beta, \beta_n) = 0$
- For almost every $y \in \mathbb{T}$ we have $\lim_{n\to\infty} \rho_{sym}(\beta_y, (\beta_n)_y) = 0$
- Each β_n has the same number of atoms as the initial partition β.

To finish the proof we invoke Proposition 4.3.5 in [8] which states that in the space of partitions of $\mathbb{T} \times X$ with cardinality at most $k = |\beta|$, convergence in the symmetric difference metric ρ_{sym} implies convergence in the conditional entropy (Rokhlin) metric ρ. ☐

6. Proof of Theorem 1.2

We fix for this section an arbitrary invertible measure-preserving system $T \curvearrowright (X, \mu)$. The next result proves Theorem 1.2.

Proposition 6.1. *Let $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$ be the sequence of finite sets defined in Definition 4.2, and let $\boldsymbol{a} = \{a_n(t)\}_{n\in\mathbb{N}, t>0}$ be an arbitrary scale. Then*

$$\overline{\mathrm{ent}}^{\boldsymbol{a},\mathcal{F}}_{m\times\mu}(S_\alpha \rtimes_\tau T) = \overline{\mathrm{ent}}^{\boldsymbol{a}}_{\mu}(T).$$

We divide the proof in several inequalities. Let ξ be a partition of $\mathbb{T}$ whose atoms are arcs, and let η be a partition of X.

Proposition 6.2. *If $(y, x), (y', x') \in G_+(n) \times X$ then we have*

$$d^H_{\eta,\{1,\ldots,n\}}(x, x') \leq d^H_{\xi\times\eta,F_n}((y, x), (y', x')) \leq d^H_{\xi,F_n}(y, y') + d^H_{\eta,\{1,\ldots,n\}}(x, x')$$

Proof. Recall that when $y \in G_+(n)$ we have $\tau^{iq(n)}(y) = i$ for $i = 1, \ldots, n$. Therefore $(S_\alpha \rtimes_\tau T)^{iq(n)}(y, x) = (S_\alpha^{iq(n)}(y), T^{\tau^{iq(n)}(y)}(x)) = (S_\alpha^{iq(n)}(y), T^i(x))$. This shows that the number of $j \in F_n$ for which the images by $(S_\alpha \rtimes_\tau T)^j$ of (y, x) and (y', x') belong to different atoms of $\xi \times \eta$ is at least $|\{j \in \{1, \ldots, n\} : \eta(T^j(x)) \neq \eta(T^j(x'))\}|$, and at most $|\{j \in F_n : \xi(S_\alpha^j(y)) \neq \xi(S_\alpha^j(y'))| + |\{j \in \{1, \ldots, n\} : \eta(T^j(x)) \neq \eta(T^j(x'))\}|$. Dividing by $n = |F_n|$ we obtain the inequalities in the statement. ☐

Proposition 6.3. *For all n large enough (meaning $m(G_+(n)) > 1/3$) we have*

$$S^H_{\xi\times\eta}(S_\alpha \rtimes_\tau T, F_n, \epsilon, \delta) \geq S^H_\eta(T, \{1, \ldots, n\}, 2\epsilon, 3\delta)$$

Proof. Let $(y_1, x_1), \ldots, (y_k, x_k) \in \mathbb{T} \times X$ so that $\bigcup_{i=1}^k B^H_{\xi\times\eta,F_n}((y_i, x_i), \epsilon)$ has measure $m \times \mu$ at least $1-\delta$. This implies that the intersection of $G_+(n)$ with $\bigcup_{i=1}^k B^H_{\xi\times\eta,F_n}((y_i, x_i), \epsilon)$ has measure at least $m(G_+(n)) - \delta$. Increasing the radius of the balls to 2ϵ, we can change their center to $G_+(n) \times X$. That is, for each i such that $B^H_{\xi\times\eta,F_n}((y_i, x_i), \epsilon) \cap G_+(n) \times X$ is nonempty, we pick an element in this intersection. We find a collection of $\ell \leq k$ elements $(u_1, v_1), \ldots, (u_\ell, v_\ell)$ in $G_+(n) \times X$ such that the intersection of $G_+(n)$ with $\bigcup_{i=1}^\ell B^H_{\xi\times\eta,F_n}((u_i, v_i), 2\epsilon)$ has measure at least $m(G_+(n)) - \delta$.

If (y, x) belongs to the intersection of $G_+(n)$ and $\bigcup_{i=1}^\ell B^H_{\xi\times\eta,F_n}((u_i, v_i), 2\epsilon)$, then we must have $d^H_{\eta,\{1,\ldots,n\}}(x, v_i) \leq 2\epsilon$ for some $i = 1, \ldots, \ell$ by Proposition 6.2. Therefore we have

$$G_+(n) \cap \bigcup_{i=1}^{\ell} B^H_{\xi\times\eta,F_n}((u_i, v_i), 2\epsilon) \subset G_+(n) \times \bigcup_{i=1}^{\ell} B^H_{\eta,\{1,\ldots,n\}}(v_i, 2\epsilon)$$

The measure $m \times \mu$ of the set at the right is equal to $m(G_+(n)) \cdot \mu(\bigcup_{i=1}^\ell B^H_{\eta,\{1,\ldots,n\}}(v_i, 2\epsilon))$. By the inclusion above, we know that this value is greater or equal to the measure of the set at the left, which is at least $m(G_+(n)) - \delta$ by previous observations. It follows that

$$m(G_+(n)) - \delta \leq m(G_+(n)) \cdot \mu(\bigcup_{i=1}^{\ell} B^H_{\eta,\{1,\ldots,n\}}(v_i, 2\epsilon)).$$

Dividing this relation by $m(G_+(n)$ we see that the measure μ of $\bigcup_{i=1}^\ell B^H_{\eta,\{1,\ldots,n\}}(v_i, 2\epsilon)$ is at least $1 - \delta/m(G_+(n))$. This number is larger than $1 - 3\delta$ when $m(G_+(n)) \geq 1/3$, which holds for all n large enough by Proposition 4.2. (the same argument works as long as the liminf of $m(G_+(n))$ is positive, up to changing an irrelevant constant in the statement). □

Proposition 6.4. *For all n large enough we have*

$$S^H_{\xi\times\eta}(S_\alpha \rtimes_\tau T, F_n, \epsilon, 2\delta) \leq 2|\xi| S^H_\eta(T, \{1, \ldots, n\}, \epsilon, \delta)$$

Proof. Let n_0 be as defined in Proposition 4.3. Let $n \geq n_0$, and let $\xi_n = \{P_1, \ldots, P_\ell\}$ as defined in Proposition 4.3. Thus $\ell = |\xi|$ and $m(\bigcup_{i=1}^\ell P_i) \geq 1 - |\xi| n/2^{n-1}$.

By Proposition 4.2 we have $m(G_+(n)) \geq 1/2 - n/2^{n-1}$, and thus the intersection of $G_+(n)$ with $\bigcup_{i=1}^\ell P_i$ has measure at least $1/2 - n/2^{n-1} - |\xi|n/2^{n-1} = 1/2 - (1 + |\xi|)n/2^{n-1}$. Next, let $x_1, \ldots, x_k$ be a collection of $k = S^H_\eta(T, \{1, \ldots, n\}, \epsilon, \delta)$ elements in X so that the measure μ of $\bigcup_{j=1}^k B^H_{\eta,\{1,\ldots,n\}}(x_j, \epsilon)$ is larger than $1 - \delta$.

Let

$$A_+(n) = ((G_+(n) \cap \bigcup_{i=1}^{\ell} P_i) \times \bigcup_{j=1}^{k} B^H_{\eta,\{1,\ldots,n\}}(x_j, \epsilon))$$

From the definition of product measure and the observations in the previous paragraph we have

$$m \times \mu(A_+(n)) \geq (1/2 - (1 + |\xi|)n/2^{n-1})(1 - \delta)$$

The value at the right converges to $1/2 - \delta/2$ as n tends to infinity, so it is larger than $1/2 - \delta$ for all n large enough.

Let $\mathcal{I}$ be the set of indices $i = 1, \ldots, \ell$ such that the intersection $G_+(n) \cap P_i$ is nonempty. For each such i we pick an element y_i in this intersection. Given an arbitrary (y, x) in the set $A_+(n)$ we can find $i \in \mathcal{I}$ with $y \in G_+(n) \cap P_i$, and $j \in \{1, \ldots, k\}$ with $x \in B^H_{\eta,\{1,\ldots,n\}}(x_j, \epsilon)$. This is possible simply by the definition of $A_+(n)$. Observe that $d^H_{\xi,F_n}(y, y_i) = 0$ as the distance d^H_{ξ,F_n} between two elements is zero when they belong to the same atom in ξ^{F_n}, which in this case is P_i.

By Proposition 6.2 it follows that then $d^H_{\xi\times\eta,F_n}((y,x),(y_i,x_j)) \le d^H_{\xi,F_n}(y,y_i)+d^H_{\eta,\{1,\dots,n\}}(x,x_j) \le 0+\epsilon$. Thus

$$A_+(n) \subset \bigcup_{i\in\mathcal{I}} \bigcup_{j=1}^{k} B^H_{\xi\times\eta,F_n}((y_i,x_j),\epsilon).$$

We have proved that $A_+(n)$ is contained in the union of at most $\ell\cdot k = |\xi| S^H_\eta(T,\{1,\dots,n\},\epsilon,\delta)$ pseudoballs of radius ϵ in the distance $d^H_{\xi\times\eta,F_n}$.

The next step of the proof is analogue. Recall from Section 3 that $S^H_\eta(T,\{-n,\dots,-1\},\epsilon,\delta) = S^H_\eta(T,\{1,\dots,n\},\epsilon,\delta)$. Consider a collection $z_1,\dots,z_k$ of elements in X with cardinality $k = S^H_\eta(T,\{-n,\dots,-1\},\epsilon,\delta)$, so that $\mu(\bigcup_{j=1}^k B^H_{\eta,\{-n,\dots,-1\}}(z_j,\epsilon)) \ge 1-\delta$, and define $A_-(n)$ analogously. The same argument shows that $A_-(n)$ is contained in the union of at most $\ell\cdot k$ pseudoballs of radius ϵ in $d^H_{\xi\times\eta,F_n}$.

Finally, note that $A_-(n)$ and $A_+(n)$ are disjoint because $G_-(n)$ and $G_+(n)$ are disjoint. Thus $m(A_-(n)\cup A_+(n)) > 1-2\delta$ for n large enough. This finishes the argument. ☐

Proof of Proposition 6.1. Let ξ be a partition of $\mathbb{T}$ whose atoms are arcs, and let η be a partition of X. Reviewing the definitions one immediately sees that by Proposition 6.3 and Proposition 6.4 we have

$$\overline{\mathrm{ent}}^{\mathbf{a},\mathcal{F}}_{m\times\mu}(S_\alpha \rtimes_\tau T, \xi\times\eta) = \overline{\mathrm{ent}}^{\mathbf{a}}_{\mu}(T,\eta). \tag{6}$$

Taking supremum over this kind of product partitions, we obtain $\overline{\mathrm{ent}}^{\mathbf{a},\mathcal{F}}_{m\times\mu}(S_\alpha \rtimes_\tau T) = \overline{\mathrm{ent}}^{\mathbf{a}}_{\mu}(T)$ thanks to Proposition 3.4. ☐

7. The $\mathcal{F}$-Pinsker algebra

Recall that the Pinsker algebra $\Pi(T)$ of a measure-preserving transformation $T \curvearrowright (X,\mathcal{B},\mu)$ is the collection of all $P\in\mathcal{B}$ such that $h_\mu(T,\{P, X\setminus P\})=0$. Also recall that a transformation is called Kolmogorov when $\Pi(T)$ is the collection of sets $P\in\mathcal{B}$ with $\mu(P)\in\{0,1\}$. We shall consider a form of sequential Pinsker algebra.

Definition 7.1. *Let $\mathcal{F}=(F_n)_{n\in\mathbb{N}}$ be a sequence of finite subsets of $\mathbb{Z}$. The $\mathcal{F}$-Pinsker algebra $\Pi^{\mathcal{F}}(T)$ of a measure-preserving transformation $T \curvearrowright (X,\mathcal{B},\mu)$ is the collection of all $P\in\mathcal{B}$ such that $h^{\mathcal{F}}_\mu(T,\{P,X\setminus P\})=0$.*

Our previous results allow us to compute the $\mathcal{F}$-Pinsker algebra of $S_\alpha \rtimes_\tau T$, where $\mathcal{F}$ is the sequence from Definition 4.2.

Proposition 7.1. *Let α be a good irrational, let $\mathcal{F}$ be as defined in Definition 4.2, and let $T \curvearrowright (X,\mathcal{B},\mu)$ be an automorphism of a standard probability space. Then*

$$\Pi^{\mathcal{F}}(S_\alpha \rtimes_\tau T) = \{P\in\mathcal{B}_{\mathbb{T}}\otimes\mathcal{B} : P_y\in\Pi(T) \text{ for almost every } y\in\mathbb{T}\}$$

In particular, if T is a Kolmogorov automorphism,

$$\Pi^{\mathcal{F}}(S_\alpha \rtimes_\tau T) = \{P\in\mathcal{B}_{\mathbb{T}}\otimes\mathcal{B} :\ \mu(P_y) \text{ equals } 0 \text{ or } 1 \text{ for almost every } y\in\mathbb{T}\}$$

Proof. Let P be a measurable subset of $\mathbb{T}\times X$. Theorem 5.1 shows that $h^{\mathcal{F}}_{m\times\mu}(S_\alpha \rtimes_\tau T, \{P, \mathbb{T}\times X\setminus P\})=0$ if and only if for almost every $y\in\mathbb{T}$ we have $h_\mu(T,\{P_y, X\setminus P_y\})=0$. This is equivalent to $P_y\in\Pi(T)$ for almost every $y\in\mathbb{T}$. ☐

In the next result, the intuitive idea is that if ϕ is not independent of its first coordinate, then Φ would map a partition with positive entropy $h^{\mathcal{F}}_{m\times\mu}$ to one with zero entropy (this is clear for product partitions).

Theorem 7.1. *Let α be a good irrational, let $(X, \mathcal{B}, \mu)$ be a standard probability space, and let T_1 and T_2 be K-automorphisms. Suppose that Φ is an isomorphism from $S_\alpha \rtimes_\tau T_1$ to $S_\alpha \rtimes_\tau T_2$, written as*

$$\Phi\colon \mathbb{T} \times X \to \mathbb{T} \times X$$
$$(y, x) \mapsto (\phi(y, x), \psi(y, x)).$$

Then ϕ is independent of the second coordinate in the sense that there is a function $\phi'\colon \mathbb{T} \to \mathbb{T}$ so that $\phi(y, x) = \phi'(y)$ for almost every $(y, x) \in \mathbb{T} \times X$.

Proof. The conclusion of Proposition 7.1 can be re-written (up to null sets) as

$$\Pi^{\mathcal{F}}(S_\alpha \rtimes_\tau T_1) = \Pi^{\mathcal{F}}(S_\alpha \rtimes_\tau T_2) = \{B \times X : B \in \mathcal{B}_{\mathbb{T}}\}.$$

Furthermore, it is clear that an isomorphism of measure-preserving transformations respects the $\mathcal{F}$-Pinsker algebras (up to null sets). It follows that for every $B \in \mathcal{B}_{\mathbb{T}}$ we can find $C \in \mathcal{B}_{\mathbb{T}}$ so that $\Phi^{-1}(B \times \{X\}) = C \times \{X\}$ (up to null sets). Since $\Phi^{-1}(B \times \{X\}) = \{(y, x) : \phi(y, x) \in B\} = \phi^{-1}(B)$, we also have that $\phi^{-1}(B \times \{X\}) = C \times \{X\}$. This shows that $\phi\colon \mathbb{T} \times X \to X$ is a measurable function, where the domain is given the sigma-algebra generated by $\{B \times X : B \in \mathcal{B}_{\mathbb{T}}\}$, and the codomain is given the sigma-algebra $\mathcal{B}_{\mathbb{T}}$. This implies that ϕ depends on its first coordinate almost everywhere by a general result about measurable functions. To see this, let $\pi\colon (\mathbb{T} \times X, \mathcal{B}_{\mathbb{T}} \otimes \mathcal{B}, m \times \mu) \to (X, \mathcal{B}, \mu)$ be the projection to the first coordinate. Let $\pi^{-1}(\mathcal{B})$ be the sigma-algebra generated by $\{\pi^{-1}(B) : B \in \mathcal{B}\}$, which is equal to $\{B \times X : B \in \mathcal{B}_{\mathbb{T}}\}$ (up to null sets). By Lemma 1.14 in [6], the fact that ϕ is measurable with respect to $\pi^{-1}(\mathcal{B})$ implies that $\phi = \pi \circ \upsilon$ (up to null sets) for some measurable function $\upsilon\colon (\mathbb{T}, \mathcal{B}_{\mathbb{T}}) \to (\mathbb{T} \times X, \mathcal{B}_{\mathbb{T}} \otimes \mathcal{B})$. Then $\pi \circ \upsilon$ has the property in the statement. □

8. Isomorphisms between skew products $S_\alpha \rtimes_\tau T$

In this section we prove the following result (observe that there are no assumptions on T_1 or T_2, such as being K).

Theorem 8.1. *Let α be a good irrational. Let $(X, \mathcal{B}, \mu)$ be a standard probability space, and let T_1 and T_2 be two automorphisms. Suppose that $S_\alpha \rtimes_\tau T_1$ is isomorphic to $S_\alpha \rtimes_\tau T_2$ via an isomorphism Φ of the form*

$$\Phi\colon \mathbb{T} \times X \to \mathbb{T} \times X$$
$$(y, x) \mapsto (\phi(y), \psi(y, x)).$$

That is, the first coordinate of Φ depends only on its $\mathbb{T}$-coordinate. Then T_1 is isomorphic to T_2 or the inverse of T_2.

We fix for the rest of this section the objects defined in the statement of this result.

Proposition 8.1. *There exists $c \in \mathbb{T}$ so that $\phi(y) = y + c$ for almost every $y \in \mathbb{T}$.*

Proof. The relation $\Phi \circ (S_\alpha \rtimes_\tau T_1) = (S_\alpha \rtimes_\tau T_2) \circ \Phi$ implies that $S_\alpha \circ \phi = \phi \circ S_\alpha$. Indeed, given $(y, x) \in \mathbb{T} \times X$ we have

$$\Phi((S_\alpha \rtimes_\tau T_1)(y, x)) = \Phi(S_\alpha(y), \dots) = (\phi(S_\alpha(y)), \dots)$$
$$S_\alpha \rtimes_\tau T_2(\Phi(y, x)) = S_\alpha \rtimes_\tau T_2(\phi(y), \dots) = (S_\alpha(\phi(y)), \dots)$$

Let $f\colon \mathbb{T} \to \mathbb{T}$ be given by $f(y) = \phi(y) -_{\mathbb{T}} y$ (here $-$ is interpreted as the inverse operation in the abelian group $(\mathbb{T}, +_{\mathbb{T}})$). The fact that S_α commutes with ϕ implies that $f \circ S_\alpha = f$. Since S_α is ergodic, it follows that f is constant (up to a set with null measure m).

Alternatively, it is easy to see that ϕ preserves m:

$$m(B) = m \times \mu(B \times X) = m \times \mu(\Phi^{-1}(B \times X)) = m \times \mu(\phi^{-1}(B) \times X) = m(\phi^{-1}(B))$$

Therefore ϕ is an element in the automorphism group or centralizer group of the transformation $S_\alpha \curvearrowright (\mathbb{T}, \mathcal{B}_\mathbb{T}, m)$. It is known that this group consists precisely of rotations. $\square$

In what follows we write $\psi(y,x) = \psi_y(x)$. Furthermore, we denote by $\mathrm{Aut}(X, \mathcal{B}, \mu)$ the collection of automorphisms of the probability space $(X, \mathcal{B}, \mu)$, in which two elements are identified if they coincide almost everywhere.

Proposition 8.2. *For almost every y we have $\psi_y \in \mathrm{Aut}(X, \mathcal{B}, \mu)$*

Proof. We first prove that for almost every $y \in \mathbb{T}$ the function $\psi_y : X \to X$ is bijective up to a null set. Since Φ is an isomorphism it admits an inverse $\Phi' : \mathbb{T} \times X \to \mathbb{T} \times X$. Thus $\Phi \circ \Phi' = \Phi' \circ \Phi$ is the identity map on $\mathbb{T} \times X$ up to removing a set with null measure $m \times \mu$. The results we proved for Φ are also valid for Φ', and thus the first coordinate of Φ' depends only on its first coordinate, and is a circle rotation. This rotation must be $y \to y - c$ as it must cancel out with the rotation in the first coordinate of Φ. Thus we can write $\Phi'(y,x) = (y - c, \psi'(y,x))$, for a measurable map $\psi' : \mathbb{T} \times X \to X$. We also write $\psi'(y,x) = \psi'_y(x)$.

For almost every (y,x) we have $(y,x) = \Phi'(\Phi(y,x)) = \Phi'(y + c, \psi_y(x)) = (y, \psi'_{y+c}(\psi_y(x)))$. It follows that for almost every y, $\psi'_{y+c} \circ \psi_y$ is the identity map on X (up to a set with null measure μ). A similar argument shows that for almost every y, $\psi_y \circ \psi'_{y+c}$ is the identity map on X (up to a set with null measure μ). The fact that ψ_y has a measurable inverse implies that it is bijective (up to a set with null measure μ).

Next we prove that for almost every $y \in \mathbb{T}$ the map ψ_y is measure preserving in the sense that $(\psi_y)_*\mu = \mu$. Let $P \in \mathcal{B}_\mathbb{T}$ and $Q \in \mathcal{B}$. Observe that $\Phi^{-1}(P \times Q) = \{(y,x) : y + c \in P, \psi_y(x) \in Q\}$. Therefore

$$m \times \mu(\Phi^{-1}(P \times Q)) = \int_\mathbb{T} 1_P(y + c)\mu(\psi_y^{-1}(Q))dm(y)$$

Applying the change of variables $y \to y - c$ which preserves m we find

$$m \times \mu(\Phi^{-1}(P \times Q)) = \int_\mathbb{T} 1_P(y)\mu(\psi_{y-c}^{-1}(Q))dm(y)$$

On the other hand,

$$m \times \mu(\Phi^{-1}(P \times Q)) = m \times \mu(P \times Q) = \int_\mathbb{T} 1_P(y)\mu(Q)dm(y)$$

We conclude that

$$\int_\mathbb{T} 1_P(y)\mu(Q)dm(y) = \int_\mathbb{T} 1_P(y)\mu(\psi_{y-c}^{-1}(Q))dm(y), \quad P \in \mathcal{B}_\mathbb{T}, Q \in \mathcal{B}.$$

Fix $Q \in \mathcal{B}$. Since the previous equality holds for every $P \in \mathcal{B}$, it follows that $\mu(Q) = \mu(\psi_{y-c}^{-1}(Q))$ for y in a subset of $\mathbb{T}$ with full measure.

Let $\{Q_i : i \in \mathbb{N}\}$ be a countable basis for $\mathcal{B}$. Since a countable intersection of subsets of $\mathbb{T}$ with full measure still has full measure, we can find we find a Borel measurable $K \subset \mathbb{T}$ with $m(K) = 1$ and such that every $y \in K$ satisfies $\mu(Q_i) = \mu(\psi_{y-c}^{-1}(Q_i))$ for all $i \in \mathbb{N}$. Since $\{Q_i : i \in \mathbb{N}\}$ generates $\mathcal{B}$, it follows that $\mu(B) = \mu(\psi_{y-c}^{-1}(B))$, for all $B \in \mathcal{B}$. The set $K + c$ also has full measure, and every $y \in K + c$ verifies $\mu(B) = \mu(\psi_y^{-1}(B))$, for all $B \in \mathcal{B}$. $\square$

The previous proposition ensures that $y \to \psi_y$ is an almost everywhere defined function $\mathbb{T} \to \mathrm{Aut}(X, \mathcal{B}, \mu)$. We will need to prove some statements about it, but first we need to review some facts about $\mathrm{Aut}(X, \mathcal{B}, \mu)$. This is a group with the composition operation $(f, g) \to f \circ g$, $f \circ g(x) = f(g(x))$. We endow $\mathrm{Aut}(X, \mathcal{B}, \mu)$ with the weak topology, which makes it a topological group. In fact, with this topology it is a Polish metrizable space. In this topology we have that

$f_n \to f$ if and only if for every $B \in \mathcal{B}$, we have $\mu(f^{-1}(B) \triangle f_n^{-1}(B)) \to 0$ as $n \to \infty$. The reader is referred to [10, §1] for further details about the weak topology on $\mathrm{Aut}(X, \mathcal{B}, \mu)$.

We endow $\mathrm{Aut}(X, \mathcal{B}, \mu)$ with the Borel sigma-algebra associated to its weak topology. Thus we can speak about the measurability of functions $\mathbb{T} \to \mathrm{Aut}(X, \mathcal{B}, \mu)$.

Proposition 8.3. *The (almost everywhere defined) map $\mathbb{T} \to \mathrm{Aut}(X, \mathcal{B}, \mu)$ given by $y \to \psi_y$ is measurable.*

Proof. In order to prove that $y \to \psi_y$ is measurable, we use the fact that $(y, x) \to \psi(y, x)$ is measurable as a function of both y and x (this seems to be standard, but since we have not found a citable statement, we provide an argument). Recall that Φ is measurable by hypothesis, and therefore the same is true for its projection to the second coordinate $\psi \colon \mathbb{T} \times X \to X$. Also recall that we write $\psi_y(x) = \psi(y, x)$.

Since $\mathrm{Aut}(X, \mathcal{B}, \mu)$ is a topological group, taking inverses is continuous and in particular measurable. Thus it is sufficient to prove that $y \to \psi_y^{-1}$ is measurable. A sub-basis for the weak topology on $\mathrm{Aut}(X, \mathcal{B}, \mu)$ is given by sets of the form

$$U_{a,b,A,B} = \{R \in \mathrm{Aut}(X, \mathcal{B}, \mu) : a < \mu(R(A) \cap B) < b\},$$

where $a, b > 0$ and $A, B \in \mathcal{B}$ (this follows from the definition of $W_{S, A_1, \ldots, A_n, \epsilon}$ given in page 2 of [10]). Fix $a, b > 0$ and A, B in $\mathcal{B}$. To prove that $\{y \in \mathbb{T} : a < \mu(\psi_y^{-1}(A) \cap B) < b\}$ is measurable, it is sufficient to show the measurability of the map

$$f \colon \mathbb{T} \to [0, 1], y \mapsto f(y) = \mu(\psi_y^{-1}(A) \cap B).$$

Observe that we can write

$$f(y) = \int_X 1_{\psi_y^{-1}(A)}(x) \cdot 1_B(x) d\mu(x).$$

Furthermore we have

$$1_{\psi_y^{-1}(A)}(x) = 1 \iff \psi(y, x) \in A \iff 1_{\psi^{-1}(A)}(y, x) = 1$$

Thus we can re-write

$$f(y) = \int_X 1_{\psi^{-1}(A)}(y, x) \cdot 1_B(x) d\mu(x)$$

Since $\psi \colon \mathbb{T} \times X \to X$ is measurable and $A \in \mathcal{B}$, we have that $\psi^{-1}(A)$ is a measurable set. It follows that $(y, x) \to 1_{\psi^{-1}(A)}(y, x) \cdot 1_B(x)$ is measurable as a function of both y and x. Then Tonelli's theorem shows that $y \to f(y)$ is measurable. □

The next result proves Theorem 8.1.

Proposition 8.4 (Theorem 8.1)**.** *There exists $y_* \in \mathbb{T}$ such that $\psi_{y_*} \in \mathrm{Aut}(X, \mathcal{B}, \mu)$ and furthermore $\psi_{y_*} \circ T_1 = T_2 \circ \psi_{y_*}$ or $\psi_{y_*} \circ T_1 = T_2^{-1} \circ \psi_{y_*}$.*

Proof. Recall from Proposition 4.2 that the measure of $G(n)$ tends to 1 as n tends to infinity, that $G(n)$ is the disjoint union of $G_+(n)$ and $G_-(n)$, and that the measure of each of them tends to $1/2$ as n tends to infinity. Furthermore, the map $\mathbb{T} \to \mathbb{T}$, $y \to y - c$ preserves m. Therefore

$$1/2 = \lim_{n \to \infty} m(G_+(n) \cap (G(n) - c)) = \lim_{n \to \infty} m(G_+(n) \cap (G_+(n) - c)) + m(G_+(n) \cap (G_-(n) - c))$$

It follows that

$$1/2 \le \limsup_{n \to \infty} m(G_+(n) \cap (G_+(n) - c)) + \limsup_{n \to \infty} m(G_+(n) \cap (G_-(n) - c))$$

It follows that at least one of both limsups is at least $1/4$. We assume for the rest of the proof that the first of the two limsups is at least $1/4$. Under this assumption, we shall find y as in the statement and with the property $\psi_y \circ T_1 = T_2 \circ \psi_y$ (if the second limsup is at least $1/4$ then one follows a similar argument and finds y_* with $\psi_{y_*} \circ T_1 = T_2^{-1} \circ \psi_{y_*}$). Pick an increasing sequence of natural numbers $(n_k)_{k\in\mathbb{N}}$ such that

$$m(G_+(n_k) \cap (G_+(n_k) - c)) \geq 1/4, \quad k \in \mathbb{N}.$$

By the previous two propositions, there is a Borel measurable set $O \subset \mathbb{T}$ with $m(O) = 1$, and such that for every $y \in O$ we have that $\psi_y \in \mathrm{Aut}(X, \mathcal{B}, \mu)$. By the inner regularity of the Lebesgue measure [13, §2] we have

$$1 = m(O) = \sup\{m(K) : K \subset O \text{ and } K \text{ is compact}\}.$$

Pick a compact set K_0 with $K_0 \subset O$ with $m(K_0) \geq 1 - 1/32$. By Proposition 8.3 the function $K_0 \to \mathrm{Aut}(X, \mathcal{B}, \mu)$ given by $y \to \psi_y$ is measurable. We endow K with the subspace topology, which makes it a Polish metrizable topological space. As we discussed, the weak topology on $\mathrm{Aut}(X, \mathcal{B}, \mu)$ is also Polish metrizable. Then a general form of Lusin's Theorem [9, Theorem 17.12] states that we can find a compact subset of K_0 over which $y \to \psi_y$ is continuous, and with measure m arbitrarily close to $m(K_0)$. Choose a compact set $K \subset K_0$ with $m(K) \geq 1 - 1/16$ and for which the map $K \to \mathrm{Aut}(X, \mathcal{B}, \mu)$, $y \to \psi_y$ is continuous.

A form of Steinhaus's Theorem for the compact group $(\mathbb{T}, +_{\mathbb{T}})$ and the compact set K from before states the following: for every $\epsilon > 0$ we can find a neighborhood I of the identity element of $\mathbb{T}$ such that every $s \in I$ satisfies $|m(K) - m(K \cap (K + s))| < \epsilon$. This follows from the continuity of the map $s \to m(K \cap (K + s))$ (see [15, pp 50]). Thus we can pick an open set I in $\mathbb{T}$ containing the identity element and such that every $s \in I$ satisfies $m(K \cap (K + s)) \geq 1 - 1/8$. Recall that $q(n)\alpha$ converges to the identity element of $\mathbb{T}$ as $n \to \infty$. It follows that there is k_0 so that for all $k \geq k_0$ we have $-q(n_k)\alpha \in I$, and therefore $m(K \cap (K - q(n_k)\alpha)) \geq 1 - 1/8$.

In a probability space, the intersection of a set with measure at least $1 - 1/8$ with another set with measure at least $1/4$ is at least $1/8$. Thus we have

$$m(K \cap (K - q(n_k)\alpha) \cap G_+(n_k) \cap (G_+(n_k) - c)) \geq 1/8, \quad k \geq k_0$$

Thus for every $k \geq k_0$ we can pick y_k in the intersection above. Since K is compact, the sequence $\{y_k : k \geq k_0\}$ has an accumulation point y_* in K. In fact, replacing $(n_k)_{k\in\mathbb{N}}$ by a subsequence if necessary, we simply assume that $(y_k)_{k\in\mathbb{N}}$ converges to y_*. The condition $y_* \in K$ guarantees that $\psi_{y_*} \in \mathrm{Aut}(X, \mathcal{B}, \mu)$.

We will now prove that y_* has the desired property $\psi_{y_*} \circ T_1 = T_2 \circ \psi_{y_*}$. First we note that since Φ intertwines $S_\alpha \rtimes_\tau T_1$ and $S_\alpha \rtimes_\tau T_2$, it also intertwines their powers $(S_\alpha \rtimes_\tau T_1)^m$ and $(S_\alpha \rtimes_\tau T_2)^m$ for $m \in \mathbb{N}$ (this denotes composition m times). Looking at the second coordinate of the relation $\Phi((S_\alpha \rtimes_\tau T_1)^m(y, x)) = (S_\alpha \rtimes_\tau T_2)^m(\Phi(y, x))$ we find the equation

$$\psi_{y+m\alpha} \circ T_1^{\tau^m(y)} = T_2^{\tau^m(y+c)} \circ \psi_y$$

We use this relation for $m = q(n_k)$ and $y = y_k$, $k \in \mathbb{N}$ (here $q(n_k)$ comes from Definition 4.2). Since y_k belongs to $G_+(n_k)$, by definition of $G_+(n_k)$ we have $\tau^{q(n_k)}(y_k) = 1$. Similarly, since y_k belongs to $G_+(n_k) - c$, we also have $\tau^{q(n_k)}(y_k + c) = 1$. Therefore the relation above simplifies to

$$\psi_{y_k+q(n_k)\alpha} \circ T_1 = T_2 \circ \psi_{y_k}, \quad k \in \mathbb{N}$$

Since $q(n_k)\alpha$ converges to 0 as $k \to \infty$, it follows that $y_k + q(n_k)\alpha \to y_*$ as $k \to \infty$. Furthermore, since $y_k \in K \cap (K - q(n_k)\alpha)$, it is also true that $y_k + q(n_k)\alpha$ belongs to K, $k \geq k_0$. Combining these observations with the fact that $y \to \psi_y$ is continuous for $y \in K$, we see that in $\mathrm{Aut}(X, \mathcal{B}, \mu)$ we have the convergences

$$\lim_{k\to\infty} \psi_{y_k+q(n_k)\alpha} = \psi_{y_*} \text{ and } \lim_{k\to\infty} \psi_{y_k} = \psi_{y_*}$$

Since composition and inversion are continuous operations on $\mathrm{Aut}(X, \mathcal{B}, \mu)$, we conclude that

$$\psi_{y_*} \circ T_1 = T_2 \circ \psi_{y_*}$$

This finishes the argument. □

We can now prove Theorem 1.3.

Proof of Theorem 1.3. Since all standard probability spaces are isomorphic, we can assume that T_1 and T_2 act on the same space. Then the statement follows by applying Theorem 7.1 and then Theorem 8.1. □

References

[1] T. Austin. Scenery entropy as an invariant of RWRS processes, May 2014.

[2] K. Ball. Entropy andσ-algebra equivalence of certain random walks on random sceneries. *Israel Journal of Mathematics*, 137(1):35–60, Dec. 2003.

[3] D. Dolgopyat, C. Dong, A. Kanigowski, and P. Nándori. Flexibility of statistical properties for smooth systems satisfying the central limit theorem. *Inventiones mathematicae*, 230(1):31–120, Oct. 2022.

[4] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and Dyadic Equivalence of Random Walks on a Random Scenery. *Advances in Mathematics*, 156(2):157–179, Dec. 2000.

[5] S. A. Kalikow. T, T -1 Transformation is Not Loosely Bernoulli. *The Annals of Mathematics*, 115(2):393, Mar. 1982.

[6] O. Kallenberg. *Foundations of Modern Probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, Cham, 2021.

[7] A. Katok and J.-P. Thouvenot. Slow entropy type invariants and smooth realization of commuting measure-preserving transformations. *Annales de l'I.H.P. Probabilités et statistiques*, 33(3):323–338, 1997.

[8] A. B. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*. Number v. 54 in Encyclopedia of mathematics and its applications. Cambridge University Press, Cambridge ; New York, NY, USA, 1995.

[9] A. Kechris. *Classical Descriptive Set Theory*. New York, 1995.

[10] A. Kechris. *Global Aspects of Ergodic Group Actions*, volume 160 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, Rhode Island, Jan. 2010.

[11] A. G. Kushnirenko. On metric invariants of entropy type. *Russian Mathematical Surveys*, 22(5):53, Oct. 1967.

[12] R. Nair. On metric diophantine approximation and subsequence ergodic theory. *The New York Journal of Mathematics*, 3A:117–124, 1998.

[13] W. Rudin. *Real and complex analysis*. McGraw-Hill, New York, 3rd ed edition, 1987.

[14] D. J. Rudolph. Asymptotically Brownian skew products give non-loosely BernoulliK-automorphisms. *Inventiones Mathematicae*, 91(1):105–128, Feb. 1988.

[15] A. Weil. *L'integration dans les groupes Topologiques et ses applications*. Hermann, Paris, 2. ed edition, 1979. OCLC: 609547847.

Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Email address: `nicanor.vargas@uj.edu.pl`