DYNAMICS OF SELF-SIMILAR INTERVAL EXCHANGE TRANSFORMATIONS ON THREE INTERVALS

KELLY YANCEY

Abstract. It is well known that almost all interval exchange transformations are rigid and weakly mixing. Jon Chaika proposed the following question: Is it true that self-similar IETs cannot be both rigid and weakly mixing? The set of self-similar IETs are a set of measure zero and are the ones that are periodic under Rauzy induction. This paper answers Chaika's question in the affirmative for the case where three intervals are exchanged.

1. Introduction

Interval exchange transformations have garnered a lot of attention in the recent years as there has been substantial progress made in regards to an old question of Veech. Given a normalized length vector with d components and a permutation on d letters, a d-interval exchange transformation is an exchange of the d intervals according to the permutation. The question of Veech is the following: Are almost all interval exchange transformations simple, rigid, and weakly mixing? For a fixed permutation, almost all refers to Lebesgue measure on the unit simplex. Veech himself proved that almost all IETs (interval exchange transformations) are rigid [15] and Katok and Stepin proved that almost all 3-IETs are weakly mixing [9]. Later, in 2007, Avila and Forni showed that almost all IETs that are not rotations are weakly mixing [2]. The result of Avila and Forni was a major breakthrough in the theory of IETs. The simplicity part of Veech's question remains open.

Even though almost all IETs are weakly mixing and rigid, these results do not shed any light on the behavior of classes of IETs that have measure zero. For example, a special set of measure zero IETs are the self-similar ones, also referred to as pseudo-Anosov IETs. In the particular case of 3-IETs the results in [9] show that self-similar IETs are weakly mixing. The class of self-similar IETs is periodic under the scheme of Rauzy induction and can be thought of as arising from substitutions. This paper explores the following question of Chaika:

Question 1.1. *Is it true that self-similar IETs cannot be both weakly mixing and rigid?*

The goal is to prove that while almost all IETs are weakly mixing and rigid, these behaviors are incompatible notions in the realm of self-similar IETs.

This paper explores the dynamics of self-similar 3-IETs and answers the question in the affirmative for this case.

Theorem 1.2. *Self-similar* 3*-IETs are not rigid.*

This work was partially supported by NSF Grant DMS 1201534.

To prove this theorem we view the self-similar IET first as a substitution system and then as a cutting and stacking transformation. We then proceed in a combinatorial way to show the absence of rigidity.

A result of Ferenczi, Holton, and Zamboni [6], which states that 3-IETs either have minimal self-joinings or are rigid, allows us to conclude immediately that self-similar 3-IETs have minimal self-joinings. Also, this class of 3-IETs arrise from primitive substitutions and as such are uniquely ergodic. Thus, self-similar 3-IETs satisfy Sarnak's conjecture, that is, this class of transformations is disjoint from the Mobius function. For more information on Sarnak's conjecture see [4] and [13].

The structure of this paper is as follows: In Section 2 we give some background on substitutions, cutting and stacking transformations, and IETs. We prove Theorem 1.2 in Section 3 and in Section 4 we prove the minimal self-joinings property for a specific 3-IET directly.

2. Preliminaries

Suppose (X, β, μ) is a Lebesgue probability space. Let $T : X \to X$ be an invertible measure-preserving transformation. In this case we will call (X, β, μ, T) a *dynamical system.* We will begin with a few standard definitions from ergodic theory.

Definition 2.1. *The transformation T is weakly mixing if $f \in L_2$ and $f \circ T = \lambda f$ for some $\lambda \in \mathbb{C}$ implies that f is constant almost everywhere.*

Definition 2.2. *The transformation T is mildly mixing if there are no nonconstant rigid L_2 functions. (A function $f \in L_2$ is rigid if there exists an increasing sequence of natural numbers (n_m) such that $f \circ T^{n_m} \to f$ in L_2.)*

An equivalent formulation of mildly mixing is that there exists no measurable set A with $0 < \mu(A) < 1$ such that $\liminf_{n\to\infty} \mu(T^n A \triangle A) = 0$.

Definition 2.3. *The transformation T is strongly mixing if*

$$\mu(T^n A \cap B) \to \mu(A)\mu(B)$$

for every $A, B \in \beta$.

Definition 2.4. *The transformation T is rigid if there exists an increasing sequence of natural numbers (n_m) such that*

$$\lim_{m\to\infty} \mu(T^{n_m} A \triangle A) = 0$$

for every set A of positive measure.

With these definitions it is easy to see that rigidity and mild mixing are incompatible notions. Rigidity is also incompatible with strong mixing. Now we will give some definitions from the theory of joinings. For more information on joinings see [7] and [11] and for joinings related to substitutions see [14].

Definition 2.5. *For any integer $k \geq 2$, a k-fold self-joining of (X, β, μ, T) is a measure defined on the Cartesian product X^k that is T^k-invariant and whose marginals are μ.*

Let $J(X)$ denote the set of joinings of (X, β, μ).

Definition 2.6. *A dynamical system (X, β, μ, T) has minimal self-joinings of order k, k-fold MSJ for short, if every ergodic k-fold self-joining of the system $\nu \in J(X)$ is either the product measure μ^k or an off diagonal measure (i.e. $\nu(A_1 \times \cdots \times A_k) = \mu(A_1 \cap T^{i_2} A_2 \cap \cdots \cap T^{i_k} A_k)$).*

Definition 2.7. *A dynamical system has minimal self-joinings, MSJ, if it has minimal self-joinings of every order.*

In Rudolph's book [11] it is shown that transformations with MSJ have trivial centralizer (only the powers commute with the transformation) and are prime (there are no factors). With these observations and the fact that rigid transformations have uncountable centralizer [8], we have that a transformation with MSJ is mildly mixing. Thus if you can show that a transformation has MSJ then it cannot be rigid.

The following theorem is helpful in proving that self-similar IETs have MSJ.

Theorem 2.8. [12] *Suppose T has 2-fold MSJ, is weakly mixing, and not strongly mixing, then T has MSJ.*

2.1. **Substitution Systems.** Suppose $\mathcal{A}$ is a finite *alphabet* of d letters. A *word* is a finite sequence of letters from the alphabet, also referred to as a *block*. The set $\mathcal{A}^*$ refers to the set of all finite words over the alphabet $\mathcal{A}$. The natural metric on $\mathcal{A}^{\mathbb{Z}}$ is defined by $d(x, y) = \sum_{-\infty}^{\infty} \frac{|x_i - y_i|}{2^{|i|}}$. With this metric $\mathcal{A}^{\mathbb{Z}}$ is a compact metric space and the shift map $T : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ defined by $T(x)_i = x_{i+1}$ is a homeomorphism. A set that is closed and shift-invariant is called a *subshift* or *shift space*.

A special class of subshifts that we will study in this paper are substitutions. A map $\theta : \mathcal{A} \to \mathcal{A}^*$ which induces a map from $\mathcal{A}^*$ to $\mathcal{A}^*$ by $\theta(ab) = \theta(a)\theta(b)$ where $a, b \in \mathcal{A}$ is called a *substitution*. A word is *admissible* if it occurs as a subword of $\theta^n(a)$ for some $n \in \mathbb{N}$ and $a \in \mathcal{A}$. The *substitution space* is then defined as the set of all sequences in $\mathcal{A}^{\mathbb{Z}}$ such that every finite word is admissible. We will denote the substitution space associated to the substitution θ by X_θ.

The incidence matrix M associated to a substitution θ has entries m_{ij} where m_{ij} is the number of occurrences of i in $\theta(j)$. The substitution is said to be *primitive* if the incidence matrix is primitive. Substitution spaces arising from primitive substitutions are minimal and uniquely ergodic [3].

2.2. **Cutting and Stacking.** In this paper we will be viewing substitutions as finite rank transformations and in particular will be viewing these transformation through the lens of cutting and stacking. We begin by describing the procedure of cutting and stacking for rank-one maps.

We inductively define a sequence of towers, X_n, each of height h_n. Each X_n is a column of h_n disjoint intervals with equal measure denoted by $\{I_{n,0}, ..., I_{n,h_n-1}\}$. The elements of X_n are called *levels*. We often refer to $I_{n,0}$ as the *bottom level* and I_{n,h_n-1} as the *top level* of X_n. A transformation, T_n, is defined on $\{I_{n,0}, ..., I_{n,h_n-2}\}$ by moving up one level. That is, $T_n(I_{n,i}) = I_{n,(i+1)}$ for all $0 \leq i < h_n - 1$. Note that T_n is not defined on the top level of X_n.

Thus, we must define X_{n+1} by first cutting X_n into q_n subcolumns of equal width. We may then add any number of new levels (called *spacers*) above each subcolumn. Now, we stack every subcolumn of X_n above the subcolumn to its left to form X_{n+1}. Thus, X_{n+1} consists of q_n copies of X_n which may be separated by spacers. Finally, we define $T = \lim_{n\to\infty} T_n$. The transformation T is called a *rank-one* map. For example, Chacon's transformation was the first example of a map that is weakly mixing, but not strongly mixing and is a classic example of a rank-one map.

In this paper we will be using the technique of cutting and stacking with 3 towers instead of 1 tower as in the rank-one case. There will be no spacers, only the subcolumns of each tower will be rearranged. That is, each of the 3 towers of X_n, call them $A^{(n)}, B^{(n)}, C^{(n)}$, will be cut into a given number of subcolumns, and restacked according to a given substitution to form the new columns of X_{n+1}. The transformation acts in the same way as the rank-one case. We will denote by $h_{n,A}$, $h_{n,B}$ and $h_{n,C}$ respectively the heights of the towers $A^{(n)}$, $B^{(n)}$ and $C^{(n)}$.

2.3. **Interval Exchange Transformations and Rauzy Induction.** In this section we will define interval exchange transformations (IETs for short), review the basics of Rauzy induction, and discuss a special class of IETs, the self-similar ones. For more information on IETs see Viana's notes [16].

A d-IET is a piecewise continuous map of the unit interval in which d subintervals are exchanged according to a permutation π.

Definition 2.9. *Suppose $d \geq 2$ and let $\mathcal{A}$ be an alphabet with d letters. Given a partition $\{I_\alpha : \alpha \in \mathcal{A}\}$ of $[0, 1)$ into d subintervals each of length λ_α and a permutation $\pi = (\pi_t, \pi_b)$ of the d letters, define the d-IET $T : [0, 1) \to [0, 1)$ by*

$$T(x) = x + \sum_{\pi_b(\beta)<\pi_b(\alpha)} \lambda_\beta - \sum_{\pi_t(\beta)<\pi_t(\alpha)} \lambda_\beta \qquad \textit{for } x \in I_\alpha.$$

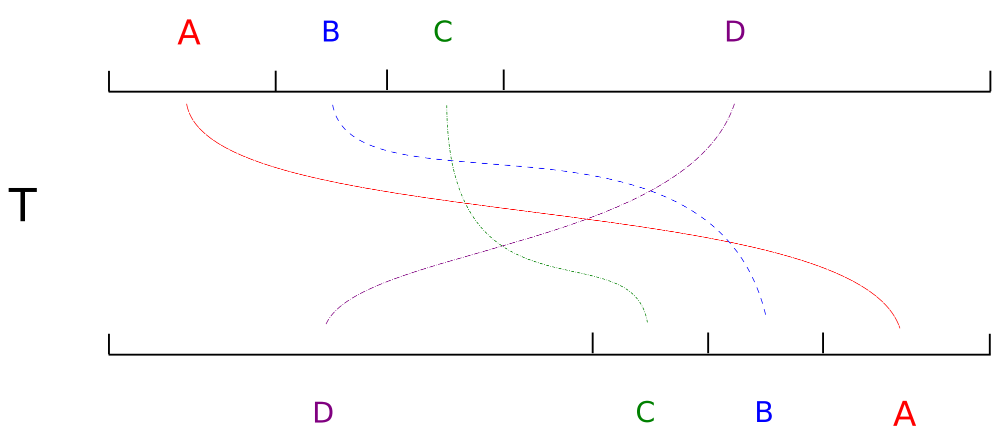

FIGURE 1. 4-IET

An example of a 4-IET with permutation

$$\pi = \begin{pmatrix} \pi_t \\ \pi_b \end{pmatrix} = \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}$$

is given in Figure 1.

Rauzy induction is a specific procedure for inducing on a subinterval of the IET. If the original IET is an exchange of d intervals, the induced map will also be a d-IET. In Rauzy induction there are two types of inducing schemes, top and bottom, depending on the length of the last subinterval. First we will describe the top procedure.

Suppose that $\pi_t(\alpha) = d$ and $\pi_b(\beta) = d$, that is, I_α is the last subinterval of the partition of $[0,1)$ and I_β is the last subinterval of the partition of $T([0,1))$. In the *top produce* we assume that $\lambda_\alpha > \lambda_\beta$. Define the interval J by $J = [0, 1-\lambda_\beta)$ and under these assumptions induce on the interval J to produce a new d-IET, T'. That is, T' is the first return map to the interval J. In this case we call I_α the *winner* and I_β the *loser*. The new IET T' has permutation $\pi' = (\pi'_t, \pi'_b)$ where $\pi'_t = \pi_t$ and in π'_b, α is replaced by α followed by β and β is deleted at the right end (in other terms we perform a cyclic permutation on symbols to the right of the winner symbol α).

For the *bottom procedure* assume that $\lambda_\alpha < \lambda_\beta$. Define the interval J by $J = [0, 1-\lambda_\alpha)$. Produce a new d-IET by inducing on the interval J. In this case we call I_β the winner. The new IET T' has permutation $\pi' = (\pi'_t, \pi'_b)$ where in π'_t, β is replaced by β followed by α and α is deleted at the right end (in other terms we perform a cyclic permutation on symbols to the right of the winner symbol β) and $\pi'_b = \pi_b$.

We will assume that all IETs satisfy the infinite distinct orbit condition (i.d.o.c) which says that the orbits of discontinuities are infinite and distinct. This is also known as the Keane condition. It ensures that Rauzy induction may be repeatedly applied [16].

The following figure displays an example of Rauzy induction applied to a 4-IET. See Example 2.10 for the permutation.

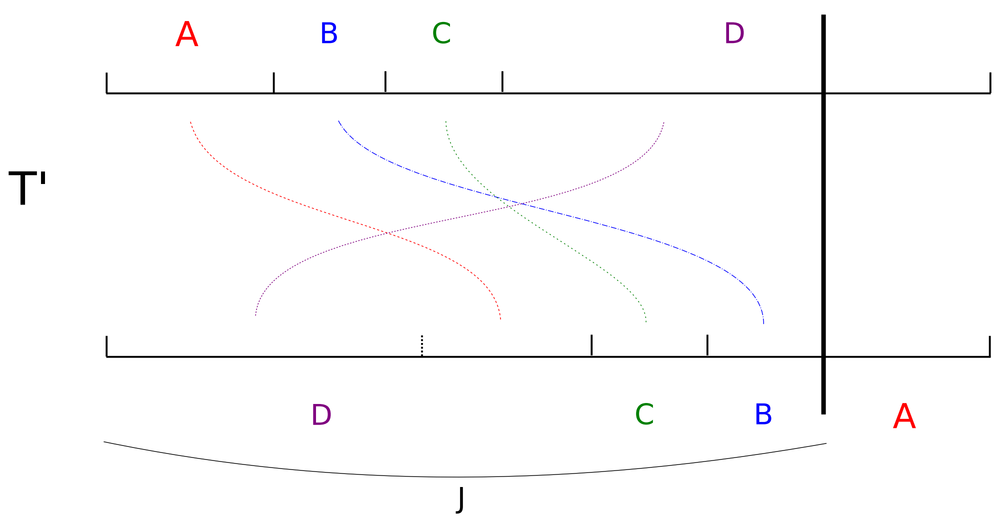

FIGURE 2. Rauzy Induction for Example 2.10

Example 2.10. *Consider the IET from Figure* 1*. As you can see from Figure* 2 *this is a top procedure and* D *is the winner. Thus the new permutation* π' *is given by*

$$\pi = \begin{pmatrix} A & B & C & D \\ D & A & C & B \end{pmatrix}.$$

2.4. **Self-Similar IETs.** The information obtained from Rauzy induction can be put into a diagram called a *Rauzy diagram.* Fix a permutation of d letters, π. Consider the graph whose vertices are permutations of d letters and put an edge from π_1 to π_2 if you can obtain the permutation π_2 from π_1 through Rauzy induction. The edges of this directed graph are labeled by the winner symbols from Rauzy induction. The Rauzy diagram associated to the given permutation π is the connected component of this graph that contains π.

For the rest of this paper we will be considering 3-IETs. The only irreducible 3-IET is given by the permutation $\pi = \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}$. The following figure displays the Rauzy diagram for this permutation.

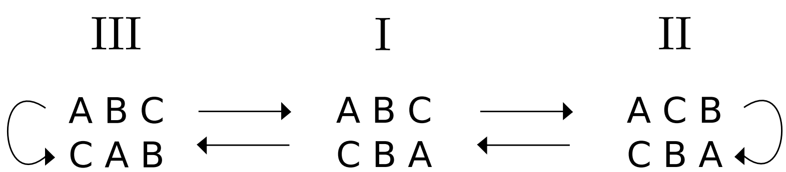

FIGURE 3. Rauzy Diagram on 3 Intervals

Associated to each edge in the Rauzy diagram is a substitution on three letters. Recall that J was the interval that we induced on in the induction scheme. The substitution is obtained by following the images of elements of J in $[0, 1)$ under T until they return to J. For example, the substitution from the top procedure induction in Example 2.10 is the following:

$$A \mapsto AD; B \mapsto B; C \mapsto C; D \mapsto D.$$

The substitutions corresponding to the edges of the Rauzy diagram from Figure 3 are below.

$$\begin{aligned} I \to II &: A \mapsto A; B \mapsto B; C \mapsto AC \\ II \to II &: A \mapsto AB; B \mapsto B; C \mapsto C \\ II \to I &: A \mapsto A; B \mapsto AB; C \mapsto C \\ I \to III &: A \mapsto AC; B \mapsto B; C \mapsto C \\ III \to III &: A \mapsto A; B \mapsto B; C \mapsto BC \\ III \to I &: A \mapsto A; B \mapsto BC; C \mapsto C \end{aligned}$$

Associated to each above substitution is a 3×3 incidence matrix M. To obtain the incidence matrix of the substitution that is derived from moving from permutation I to permutation II and then back to I, you simply multiply the individual incidence matrices from each substitution together. In this case you would obtain

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We are now ready to define the class of self-similar IETs. Intuitively, a *self-similar IET* is one that after inducing a finite number of times (using Rauzy induction) and rescaling so that you have a map of the unit interval, you arrive back to your original IET.

Definition 2.11. *The IET defined by permutation π and length vector λ is self-similar if there exists a loop in the Rauzy diagram beginning at π and an associated Perron-Frobenius matrix M such that $M\lambda = \xi\lambda$ where ξ is the dominant eigenvalue of the matrix M.*

Thus, self-similar IETs correspond to loops in the Rauzy diagram that see every label at least once, where the labels are given by the winners in the induction procedure.

It was shown in [9] that self-similar 3-IETs are weakly mixing. Weak mixing in this case also follows directly from the Veech criterion in [15].

3. 3-IETs are not Rigid

In this section we will show that self-similar 3-IETs are not rigid. One way to accomplish this goal is to view a self-similar 3-IET as a cutting and stacking transformation and proceed in a combinatorial way to obtain the absence of rigidity.

Consider the loop $CBCAA$ in the Rauzy diagram from Figure 3. This loop gives rise to a self-similar 3-IET. To study self-similar IETS we will study such loops in the Rauzy diagram and their corresponding substitutions. The substitutions will be defined over the alphabet $\mathcal{A} = \{A, B, C\}$.

We will always begin our loops in the Rauzy diagram from the permutation labeled I. Starting from this permutation there are two possible directions to follow in the directed graph. We begin by considering loops that first move from permutation I to permutation II.

Consider the path generated by the loop $ABACC$ in the Rauzy diagram. This loop gives rise to a self-similar IET on three intervals determined by the substitution θ defined by

$$\begin{aligned} \theta(A) &= ABAC \\ \theta(B) &= ABBAC \\ \theta(C) &= AC. \end{aligned}$$

Associated to the above substitution is a 3×3 incidence matrix M where

$$M = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

The matrix M has eigenvector $(1, -1, 1)^T$ with eigenvalue 1 in the neutral direction, which in the language of cutting and stacking means that the sum of the heights of the first and third columns is always equal to the height of the second column plus one. That is,

$$h_{n,A} + h_{n,C} = h_{n,B} + 1.$$

Let T denote the transformation associated (via cutting and stacking) to the IET generated by the loop $ABACC$ in the Rauzy diagram. The proposition below says that this transformation is not rigid. To show that T is not rigid, it suffices to find a set of positive measure E such that $\liminf_{n\to\infty} \mu(T^n E \triangle E) > 0$. For our purposes, let E be the bottom level of the A tower at some stage m.

Proposition 3.1. *The self-similar IET generated by the loop $ABACC$ in the Rauzy diagram is not rigid.*

Proof. Suppose for a contradiction, that $\liminf_{n\to\infty} \mu(T^n E \triangle E) = 0$. That means that there exists a subsequence of natural numbers (n_m) such that $\mu(T^{n_m} E \triangle E) \to 0$. Our goal is to follow iterates of the set E and show that if our assumption is true then it must be the case that $TE = E$. To do this we will rely heavily on the fact that $h_{n,A} + h_{n,C} = h_{n,B} + 1$. This will produce a contradiction since E has positive measure strictly less than one and T is ergodic. In [1] Aaronson and Weiss use this approach to show that the classic Chacon's map is not rigid. We are adapting their approach to the setting of finite rank transformations.

Recall that we denote by $h_{n,A}$, $h_{n,B}$ and $h_{n,C}$ respectively the heights of the towers $A^{(n)}$, $B^{(n)}$ and $C^{(n)}$ at stage n.

Let $\epsilon > 0$, and let N be such that $\mu(T^N E \triangle E) < \epsilon$. Fix n such that $N \in [h_{n,A}, h_{n+1,A} - 1]$ and assume that $n \geq m$.

First consider the case when $n = m$. That is, E is the bottom level of the A tower at stage n and $\mu(T^N E \triangle E) < \epsilon$ where $N \in [h_{n,A}, h_{n+1,A} - 1]$. This case will display the main mechanism in the proof.

Before we proceed, we need some notation for how levels in X_n appear in X_{n+2}. This will help us keep track of where the set E travels under iterates of T. It may be helpful to view the substitution after two steps to help visualize how the towers appear.

$$A \mapsto ABACABBACABACAC \tag{1}$$

$$B \mapsto ABACABBACABBACABACAC \tag{2}$$

$$C \mapsto ABACAC. \tag{3}$$

Let us now consider what happens during stage n of the construction to produce the next stage $n+1$. At the n-th stage, the tower $A^{(n)}$ is cut into five pieces, the tower $B^{(n)}$ is cut into three pieces, and the tower $C^{(n)}$ is cut into three pieces before the subcolumns are stacked together. Let A_n be the bottom level of $A^{(n)}$, B_n the bottom level of $B^{(n)}$, and C_n the bottom level of $C^{(n)}$. Thus $A^{(n)}$ is comprised of levels $A_n, TA_n, \ldots, T^{h_{n,A}-1}A_n$, $B^{(n)}$ is comprised of levels $B_n, TB_n, \ldots, T^{h_{n,B}-1}B_n$, and $C^{(n)}$ is comprised of levels $C_n, TC_n, \ldots, T^{h_{n,C}-1}C_n$. As we described before, A_n is cut into five pieces, call them $A_{n,0}, \ldots, A_{n,4}$. Taking into account the way the columns are stacked together to form X_{n+1}, we see that $A_{n+1} = A_{n,0}$, $B_{n+1} = A_{n,2}$, and $C_{n+1} = A_{n,4}$.

Figure 4 displays how X_1 is constructed from X_0.

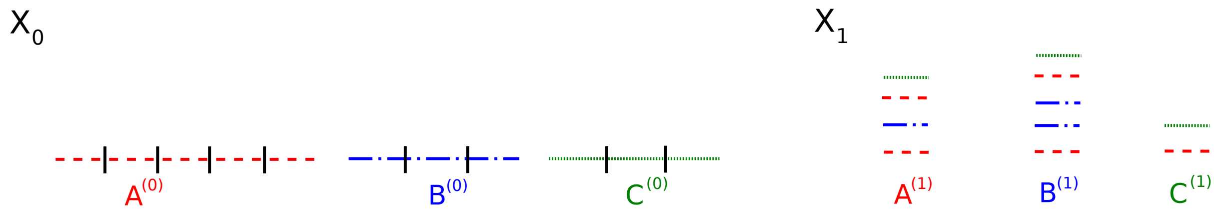

FIGURE 4. Construction of X_1 from X_0

As you can see from the figure, $h_{1,A} + h_{1,C} = 4 + 2 = 6 = 5 + 1 = h_{1,B} + 1$.

At this point it is helpful to define the size of the gaps between consecutive occurrences of $A^{(n)}$ inside the towers of X_{n+2}. To that end define

$$e_j = \begin{cases} h_{n,B}, & \text{if } j = 0, 4, 7, 13, 16 \\ 2h_{n,B}, & \text{if } j = 2, 9, 11 \\ h_{n,C}, & \text{if } j = 1, 3, 5, 8, 10, 12, 14, 17. \end{cases}$$

Notice that e_j is not defined when $j = 6, 15, 18$. Figure 5 displays how the e_j terms were determined for $0 \leq j \leq 5$ which corresponds to gaps between appearances of $A^{(n)}$ inside $A^{(n+2)}$.

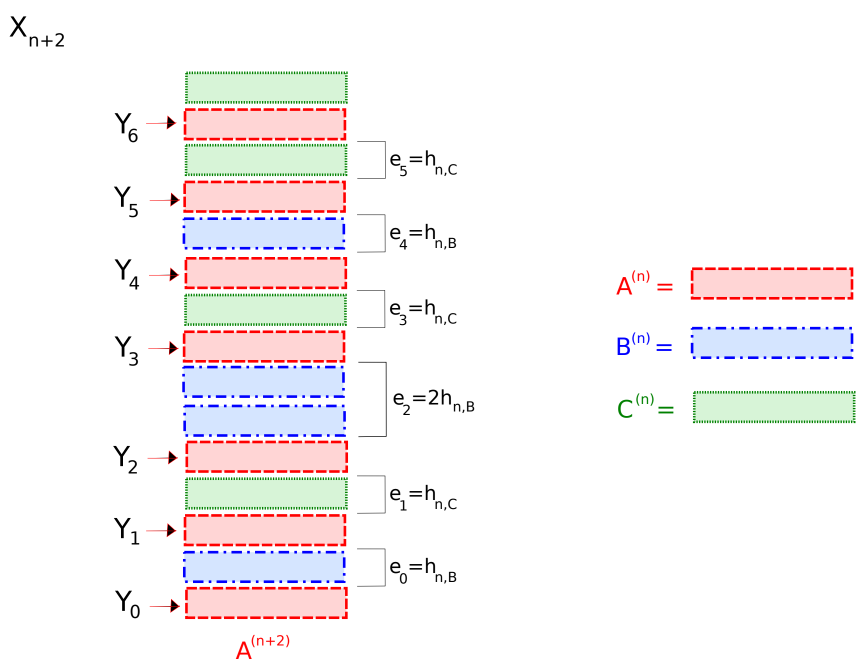

FIGURE 5. Tower $A^{(n+2)}$

Define k_j for $j = 0, 1, \ldots, 18$ by the recursive formula

$$k_j = 0 \quad \text{for } j = 0, 7, 16 \qquad \text{and} \qquad k_{j+1} = k_j + h_{n,A} + e_j \quad \text{otherwise.}$$

These variables allow us to easily describe the appearance of $A^{(n)}$ blocks inside the three towers at stage $n+2$. In that vein, let

$$Y_j = \begin{cases} \bigcup_{i=0}^{h_{n,A}-1} T^{i+k_j} A_{n+2}, & \text{if } j = 0, 1, \ldots, 6 \\ \\ \bigcup_{i=0}^{h_{n,A}-1} T^{i+k_j} B_{n+2}, & \text{if } j = 7, 8, \ldots, 15 \\ \\ \bigcup_{i=0}^{h_{n,A}-1} T^{i+k_j} C_{n+2}, & \text{if } j = 16, 17, 18. \end{cases}$$

You can view Y_j for $0 \le j \le 6$ in Figure 5. Analyzing the Y_j we see that $Y_{j+1} = T^{h_{n,A}+e_j} Y_j$ for all $j \ne 6, 15, 18$.

Recall that $N \in [h_{n,A}, h_{n+1,A} - 1]$. If we write N as $N = a h_{n,A} + b$ then we need to factor $h_{n+1,A}$ in terms of $h_{n,A}$ to determine the potential values for a and b. Now, notice that

$$h_{n+1,A} = 2h_{n,A} + h_{n,B} + h_{n,C} = 2h_{n,A} + (h_{n,A} + h_{n,C} - 1) + h_{n,C} = 3h_{n,A} + 2h_{n,C} - 1.$$

Thus if we write $N = a h_{n,A} + b$ then we have $a = 1, 2, 3$ and $0 \le b < h_{n,A}$ (except when $a = 3$, in which case $b \le 2h_{n,C} - 2$).

For fixed a (determined by N above) let $e_{j,a} = e_j + e_{j+1} + \cdots + e_{j+a-1}$. Recall that our set E belongs to $A^{(n)}$ which has been partitioned into pieces Y_j for $0 \le j \le 18$. To analyze $T^N E$ we can analyze $T^N Y_j$. Observe

$$T^N Y_j = T^{a h_{n,A} + b} Y_j = T^{b - e_{j,a}} Y_{j+a}.$$

Suppose S_1, S_2 are two sets of positive measure. By $S_1 \overset{\epsilon}{\approx} S_2$ we will mean $\mu(S_1 \triangle S_2) < \epsilon$. First we will concentrate on the pieces of E that are inside the tower $A^{(n+2)}$, that is, we will focus on $E \cap Y_j$ where $0 \le j \le 6$. Here we have

$$T^N(E \cap Y_j) = T^N E \cap T^N Y_j \overset{\epsilon}{\approx} E \cap T^{b - e_{j,a}} Y_{j+a}$$

for $0 \le j < 6 - a$. Also notice that

$$T^N(E \cap Y_j) = T^{b - e_{j,a}}(E \cap Y_{j+a})$$

for $0 \le j < 6 - a$. This is true because E appears in the same position inside each Y_j. In particular, E is the bottom level of each Y_j. Putting these together we see that

$$E \cap Y_{j+a} \overset{\epsilon}{\approx} T^{-b + e_{j,a}} E \cap Y_{j+a}.$$

A similar calculation can be carried out for the rest of the j values. Thus $E \overset{\epsilon}{\approx} T^{-b+e_{j,a}} E$.

Now, if we analyze the formulas for e_j we see that for any a we can find j, j' such that $|e_{j,a} - e_{j',a}| = h_{n,B}$. Recall that e_j is determined by the length of the gap between consecutive apprearances of Y_j, which is really determined by the length of the gap between consecutive A blocks in Equations (1) $-$ (3). The difference of $h_{n,B}$ comes from the fact that there is an A block followed by a B block and a B block followed by a B block. Hence, $E \overset{2\epsilon}{\approx} T^{h_{n,B}} E$.

To be more clear, for $a = 1, 2, 3$ choose $j = 0,\ j' = 9$, that is pick j so that Y_j is the first appearance of a Y_i inside $A^{(n+1)}$ in $A^{(n+2)}$ and j' so that $Y_{j'}$ is the first appearance of a Y_i inside $B^{(n+1)}$ in $B^{(n+2)}$.

Now we will use the fact that $h_{n,A} + h_{n,C} = h_{n,B} + 1$. Recall that E is the bottom level of $A^{(m)} = A^{(n)}$ which appears in X_n. Thus E is the bottom level in each Y_j. In particular, E is in the same position inside each Y_j as we noted earlier. Observe the following about Y_1:

$$E \cap Y_2 = T^{h_{n,A}+h_{n,C}}(E \cap Y_1) = T^{h_{n,A}+h_{n,C}}E \cap Y_2 = T^{h_{n,B}+1}E \cap Y_2 \overset{2\epsilon}{\approx} TE \cap Y_2.$$

Since E occupies the same position inside each block Y_j, the same estimate holds for every j. Thus, the above says that

$$E \overset{2\epsilon}{\approx} TE.$$

Next consider the case when $n = m+1$. Thus our set E is the bottom level of $A^{(m)}$ and our rigidity time N is between $h_{m+1,A}$ and $h_{m+2,A}-1$. This time around we want to keep track of how X_m appears inside $X_{n+2} = X_{m+3}$. The substitution is longer, but the same variables can be defined. That is, the set $A^{(m)}$ has been broken up into sets Y_j where the Y_j blocks belong to $X_{m+3} = X_{n+2}$ and e_j represents the length of consecutive gaps between Y_j. This time around if we write $N = ah_{m,A}+b$ (and remember that $N \in [h_{n,A}, h_{n+1,A} - 1] = [h_{m+1,A}, h_{m+2,A} - 1]$) then we need to factor $h_{m+1,A}$ and $h_{m+2,A}$ in terms of $h_{m,A}$. Using the relationship between the heights of the towers we see that the possible values of a and b are $a = 3, 4, \ldots, 13$ and $0 \leq b < h_{m,A}$ (except when $a = 3$, in which case $2h_{m,C} - 1 \leq b < h_{m,A}$ and when $a = 13$, in which case $0 \leq b < 2h_{m,C} + 2h_{m-1,C} - 3$).

Fix a (which depends on N). As before, the key step is finding j, j' such that $|e_{j,a} - e_{j',a}| = h_{m,B}$. The structure of the substitution allows you to find such j, j' (as in the previous case). We will use the previously constructed j and j' to form the new ones for this case. Notice that the useful blocks to analyze from the previous case were Y_0 and Y_9. Those blocks were then subdivided and restacked. We want to consider the restacked sub-blocks that appear inside $A^{(n+2)} = A^{(m+3)}$ and $B^{(n+2)} = B^{(m+3)}$. That means that we will analyze Y_0 (the first appearance of a Y_i inside $A^{(n+1)}$ in $A^{(n+2)}$) and Y_{35} (the third appearance of a Y_i inside $B^{(n+1)}$ in $B^{(n+2)}$). Now, $|e_{j,a} - e_{j',a}| = h_{m,B}$ for any a in the desired range because the difference of $h_{m,B}$ comes from the a values in the previous case ($n = m$) and then the blocks agree after that initial disparity for the desired a values.

The remainder of the proof for this case follows in exactly the same way as when $n = m$. The same analysis can be done for all n where $n \geq m$. In each case the same contradiction is reached and T is not rigid.

□

Above we considered a specific loop in the Rauzy diagram and proved that the corresponding IET is not rigid. Generalized loops in the Rauzy diagram that begin at the same permutation (permutation I) and have A as the first letter in the loop act in a similar manner. Consider the path given by $AB^kACB^\ell C$ where $k, \ell \in \mathbb{Z}$ with at least one of k, ℓ nonzero where the notation B^i means that the letter B is repeated i-times. This path is a loop that covers every letter in the alphabet and generates a self-similar IET. We will consider two separate components in the above path, $CB^\ell C$ and AB^kA. Notice that $CB^\ell C$ corresponds to the following substitution:

$$\begin{aligned}\theta_1(A) &= AC\\ \theta_1(B) &= B^{\ell+1}C\\ \theta_1(C) &= B^\ell C\end{aligned}$$

Likewise, AB^kA corresponds to the following substitution:

$$\begin{aligned}\theta_2(A) &= AB^k\\ \theta_2(B) &= AB^{k+1}\\ \theta_2(C) &= AC.\end{aligned}$$

Putting these two together gives the following substitution, θ, associated to the complete loop:

$$\begin{aligned}\theta(A) &= AB^kAC\\ \theta(B) &= (AB^{k+1})^{\ell+1}AC\\ \theta(C) &= (AB^{k+1})^\ell AC.\end{aligned}$$

For this loop, the incidence matrix associated to the substitution is

$$M = \begin{pmatrix} 2 & \ell+2 & \ell+1\\ k & (k+1)(\ell+1) & (k+1)\ell\\ 1 & 1 & 1\end{pmatrix}$$

and has eigenvector $(1,-1,1)^T$ with eigenvalue 1 in the neutral direction. This allows us to conclude that we have the same relationship between the heights of the towers $A^{(n)}, B^{(n)}, C^{(n)}$ as before. Namely

$$h_{n,A} + h_{n,C} = h_{n,B} + 1.$$

Proposition 3.2. *The self-similar IET generated by the loop $AB^kACB^\ell C$ in the Rauzy diagram is not rigid.*

The proof is the same as the proof of the previous proposition. Instead of supplying all of the details we will describe the main components of the proof. To begin, let E be the bottom level of the A tower in the m-th stage. Let $\epsilon > 0$ and suppose for a contradiction that N is such that $\mu(T^N E \triangle E) < \epsilon$ and $N \in [h_{n,A}, h_{n+1,A} - 1]$ where $n \geq m$. Recall that the first step is to analyze how X_n appears in X_{n+2}. To do that, consider two iterates of the substitution:

$$\begin{aligned}A &\mapsto AB^kAC[(AB^{k+1})^{\ell+1}AC]^kAB^kAC(AB^{k+1})^\ell AC\\ B &\mapsto [AB^kAC[(AB^{k+1})^{\ell+1}AC]^{k+1}]^{\ell+1}AB^kAC(AB^{k+1})^\ell AC\\ C &\mapsto [AB^kAC[(AB^{k+1})^{\ell+1}AC]^{k+1}]^\ell AB^kAC(AB^{k+1})^\ell AC.\end{aligned}$$

The two key components of the previous proof remain present in this more complicated version. Let $N = ah_{m,A} + b$. The first component is that you can always find j, j' such that $|e_{j,a} - e_{j',a}| = h_{m,B}$. This allows you to conclude that $T^{h_{m,B}}E \approx E$. The second component is that the sequence of levels $A^{(n)}C^{(n)}A^{(n)}$ appears in each of the three towers in X_{n+2}. Using this information and the fact that $h_{m,A} + h_{m,C} = h_{m,B} + 1$ you can conclude that $TE \approx E$, which is a contradiction.

Completely generalized loops that start at permutation I and first travel to permutation II are a countable number of combinations of the above loops, and have the same structure and neutral eigenvector. Thus we have the following theorem:

Theorem 3.3. *Any self-similar IET generated by a loop in the Rauzy diagram beginning at permutation I and first traveling to permutation II is not rigid.*

Since the Rauzy diagram is symmetric about permutation I, we have proved Theorem 1.2.

4. Minimal Self-Joinings of 3-IETs

We are now ready to discuss the MSJ property of self-similar 3-IETs. In 2005 Ferenczi, Holton, and Zamboni proved the following theorem:

Theorem 4.1. [6] *Every* 3*-IET either has MSJ or is rigid.*

In light of Theorem 1.2 we obtain a corollary about self-similar 3-IETs.

Corollary 4.2. *Self-similar* 3*-IETs have MSJ.*

As stated in the introduction, these are an important class of 3-IETs, as they satisfy Sarnak's conjecture. Let $\phi(n)$ denote the Mobius function. Sarnak's conjecture says that the Mobius function is disjoint from deterministic dynamical systems. Specifically it states that for any deterministic topological dynamical system (X, T) as $N \to \infty$ we have

$$\sum_{n \leq N} \phi(n) f(T^n x) = O(n)$$

for $x \in X$ and $f \in C(X)$.

In [4] it is stated that any uniquely ergodic topological dynamical system (X, T) such that T^p and T^q are disjoint for $p \neq q$ satisfies Sarnak's conjecture. In [5] del Junco and Rudolph showed that if T is weakly mixing and has minimal self-joinings then T^p and T^q are disjoint for $p \neq q$. Thus self-similar 3-IETs are disjoint from the Mobius function.

Currently there are no known examples of d-IETs where $d > 3$ that have MSJ.

4.1. **Example.**

Proving that a transformation has MSJ is not an easy task. In this section we concentrate on one self-similar IET and show that it has MSJ. We do this directly, by viewing the IET as a substitution system. Unlike the result from [6], our substitution arises from Rauzy induction.

Consider the substitution defined by the loop $ABACC$ in the Rauzy diagram, call this substitution θ. This is the substitution that was discussed in detail in the previous section. Let T denote the shift map on the substitution space X_θ.

Theorem 4.3. *The map T has MSJ.*

Self-similar IETs on 3 intervals are weakly mixing ([9]) and never strongly mixing ([10]). Thus to prove the above theorem, we only need to show 2-fold MSJ. The main tools to prove 2-fold MSJ are the following two lemmas by Rudolph.

Lemma 4.4. *(Lemma 6.14 from [11]) If $(X.\beta,\mu,T)$ is ergodic and $\nu \in J(X)$ is a 2-fold ergodic joining that is $(Id \times T)$ invariant, then $\nu = \mu \times \mu$.*

Lemma 4.5. *(Lemma 6.15 from [11]) Let (X,β,μ,T) be an ergodic dynamical system and $\{P_i\}$ a countable set of cylinders generating β. Let $\overline{\beta} = \{P_\ell \times P_m\}$ be a countable generating algebra of $X \times X$. Assume that*

(1) *$\nu \in J(X)$ is a 2-fold ergodic joining*

(2) *$(x,y) \in X \times X$ satisfies*
$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \chi_A(T^{-n}x, T^{-n}y) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \chi_A(T^{n}x, T^{n}y) = \nu(A)$$
for all $A \in \overline{\beta}$

(3) *there exists intervals $L_k = [i_k, j_k] \subset \mathbb{Z}$, intervals $M_k \subset \mathbb{Z}$, $t_k \in \mathbb{Z}$, and $\gamma > 0$ such that $i_k \leq 0 \leq j_k$, $j_k - i_k \to \infty$, $M_k \subset L_k$, $M_k + t_k \subset L_k$, and $|M_k| \geq \gamma |L_k|$*

(4) *for any cylinder sets P_ℓ and P_m there exists K such that if $k \geq K$ then for all $i \in M_k$*
$$T^i x \in P_\ell \text{ if and only if } T^{i+t_k} x \in P_\ell$$
$$T^i y \in P_m \text{ if and only if } T^{i+t_k+1} y \in P_m.$$

Then, ν is $(Id \times T)$ invariant and hence $\nu = \mu \times \mu$.

The first thing that we need to do is determine the structure of long words. Our substitution θ is defined on three letters, A, B, C, and given by
$$\begin{aligned} \theta(A) &= ABAC \\ \theta(B) &= ABBAC \\ \theta(C) &= AC. \end{aligned}$$
Define D_1 to be the word $\theta(A)$ with the first letter removed and an A appended to the end. Thus, $D_1 = BACA$. Similarly, define D_n to be the word $\theta^n(A)$ with the first letter removed and an A appended to the end. This means that the length of D_n is the length of $\theta^n(A)$. Let $\overline{D_n} = D_n D_{n-1} \cdots D_1$ and define $H_n = |\theta^n(A)| + \cdots + |\theta(A)|$ to be the length of $\overline{D_n}$.

Lemma 4.6. *Iterates of the substitution θ have the following form for $n \geq 2$:*
$$\begin{aligned} \theta^n(A) &= A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \\ \theta^n(B) &= A\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \\ \theta^n(C) &= A\overline{D_{n-1}}C. \end{aligned}$$

Proof. This will be a proof by induction. For the case $n = 2$ consider

$$\begin{aligned} \theta^2(A) &= \theta(ABAC) = ABACABBACABACAC = AD_1BD_1D_1C \\ \theta^2(B) &= \theta(ABBAC) = ABACABBACABBACABACAC = AD_1BD_1BD_1D_1C \\ \theta^2(C) &= \theta(AC) = ABACAC = AD_1C. \end{aligned}$$

Now suppose that the formula holds for an arbitrary n. Thus, for instance $\theta^n(A) = A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C$. From the definition of D_n we see that $D_n = \overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA$. Observe,

$$\begin{aligned} \theta^{n+1}(A) &= \theta^n(ABAC) = \theta^n(A)\theta^n(B)\theta^n(A)\theta^n(C) \\ &= \left(A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C\right)\left(A\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C\right)\left(A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C\right)\left(A\overline{D_{n-1}}C\right) \\ &= A\left(\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\right)\overline{D_{n-1}}B\left(\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\right)\left(\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\right)\overline{D_{n-1}}C \\ &= AD_n\overline{D_{n-1}}BD_nD_n\overline{D_{n-1}}C \\ &= A\overline{D_n}BD_n\overline{D_n}C. \end{aligned}$$

In a similar fashion we have

$$\begin{aligned} \theta^{n+1}(B) &= \theta^n(ABBAC) = \theta^n(A)\theta^n(B)\theta^n(B)\theta^n(A)\theta^n(C) \\ &= A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \\ &\quad A\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}C \\ &= AD_n\overline{D_{n-1}}BD_n\overline{D_{n-1}}BD_nD_n\overline{D_{n-1}}C \\ &= A\overline{D_n}B\overline{D_n}BD_n\overline{D_n}C \end{aligned}$$

and

$$\begin{aligned} \theta^{n+1}(C) &= \theta^n(AC) = \theta^n(A)\theta^n(C) \\ &= A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}C \\ &= AD_n\overline{D_{n-1}}C \\ &= A\overline{D_n}C. \end{aligned}$$

□

A point in X_θ can be written in terms of $\theta^n(A), \theta^n(B)$, and $\theta^n(C)$. The above lemma tells us what form these blocks take when $n \geq 2$. Thus if the length of an admissible word is greater than or equal to 20 then it can be written in terms of the previous lemma.

Proof of Theorem 4.3. Let $\{P_i\}$ be a countable set of cylinders generating the sigma-algebra β. Suppose that $\overline{\beta} = \{P_\ell \times P_m\}$ is a countable set generating the algebra for $X \times X$. Suppose that $\nu \in J(X)$ is an ergodic joining. Let $(x, y) \in X \times X$ be a ν generic point satisfying condition (2) of Lemma 4.5.

If x and y are in the same orbit then there exists $k \in \mathbb{Z}$ such that $T^k x = y$. Then ν is an off diagonal measure, that is, it is the image of μ under the map $z \mapsto (z, T^k z)$.

Now suppose x and y are in different orbits. Write x and y in terms of $\theta^n(A), \theta^n(B)$, and $\theta^n(C)$ blocks. Let $k_n \in \mathbb{Z}$ be such that the block containing the zeroth place of x, denoted by x_0, and the block containing the zeroth place of $T^{k_n}y$, denoted by $(T^{k_n}y)_0$, start at the same location. Then k_n is bounded by half the length of the longest block. Thus, $|k_n| \leq \frac{|\theta^n(B)|}{2}$. Since x and y are in different orbits, there exists a block where they differ, call it the s_nth block. Note that s_n negative means that x and y differ at a block that occurs in the negative direction (i.e. to the left of the zeroth position). Choose s_n such that $|s_n|$ is minimal, that is, s_n should be the first block where they disagree.

Suppose $|s_n| > 20$. In this case, consider a coding of x and $T^{k_n}y$ by $\theta^{n+1}(A)$, $\theta^{n+1}(B)$, and $\theta^{n+1}(C)$ blocks. Since $|s_n| > 20$, the coding in terms of θ^n blocks agree in a large neighborhood of the block that contains x_0. To be precise, the blocks agree in at least 19 places in each direction. Since the longest component of the substitution is $\theta(B)$ (which has length 5), the θ^{n+1} blocks from x and $T^{k_n}y$ agree and are aligned around the block that contains x_0. Thus, there is no need to shift to realign the blocks in this case. Since the block lengths are increasing in length as n increases, we have that $|s_{n+1}| < |s_n|$. Note that we can continue this process until we find n' such that $|s_{n'}| \leq 20$. For notational purposes, we will assume that n satisfies this requirement and that y is shifted by k (instead of k_n). Let ω_{s_n} be the position in x where the s_nth block begins.

The remainder of the proof is split into cases depending on the structure of the first block where x and T^ky differ. Suppose that $s_n > 0$.

Case 1. Suppose that $x_{\omega_{s_n}}$ is the beginning of $\theta^n(A)$ and $(T^ky)_{\omega_{s_n}}$ is the beginning of $\theta^n(B)$. Since every block ends in the same way, namely $\overline{D_{n-1}}C$, and begins in the same way, namely $A\overline{D_{n-1}}$, x and T^ky have the following form around the ω_{s_n}th place:

$$x \ : \ \overline{D_{n-1}}CA\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}} \tag{4}$$

$$T^ky \ : \ \overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}} \tag{5}$$

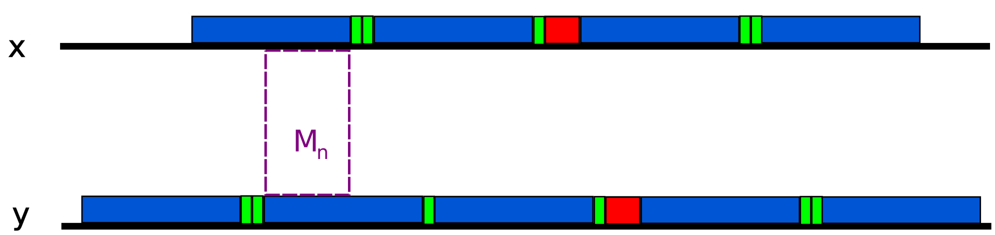

FIGURE 6. The case where $\frac{1}{2}H_{n-1} \leq k \leq \frac{3}{2}H_{n-1}$. In this figure, the long blocks have length H_{n-1}, the medium blocks have length $|\theta^{n-1}(A)|$ and the short blocks have length 1.

If $\frac{1}{2}H_{n-1} \leq k \leq \frac{3}{2}H_{n-1}$ then define M_n to be the interval where the first $\overline{D_{n-1}}$ block from x and the second $\overline{D_{n-1}}$ block from y overlap (see Figure 6). The most you are shifting y is by $\frac{3}{2}H_{n-1}$. In which case the length of M_n is $\frac{1}{2}H_{n-1} + 2$. When y is shifted by $\frac{1}{2}H_{n-1}$ the length of M_n is $\frac{1}{2}H_{n-1} - 2$.

Thus, in this case

$$|M_n| \geq \frac{1}{2}H_{n-1} - 2.$$

Figure 6 shows a shift close to $\frac{1}{2}H_{n-1}$. Define $t_n = H_{n-1} + 1$. Then for cylinders P_ℓ and P_m you can choose N large enough so that for $n \geq N$ we have

$$T^i x \in P_\ell \text{ if and only if } T^{i+t_n+1}x \in P_\ell$$
$$T^i y \in P_m \text{ if and only if } T^{i+t_n}y \in P_m$$

for all $i \in M_n - \max(|P_\ell|, |P_m|)$.

Define the interval L_n as the symmetric interval that contains the first block where x and y differ. Thus, the length of L_n is at most

$$2\left(|\omega_{s_n}| + |\theta^n(B)| + |k|\right) \leq 2\left(20|\theta^n(B)| + |\theta^n(B)| + \frac{1}{2}|\theta^n(B)|\right) \leq 44|\theta^n(B)|.$$

We can use the relationship between the length of the different letters iterated under θ to see that $|\theta^n(B)| \leq 6|\theta^{n-1}(A)|$. Thus,

$$\begin{aligned}
\liminf_{n\to\infty} \frac{|M_n|}{|L_n|} &\geq \liminf_{n\to\infty} \frac{\frac{1}{2}H_{n-1} - 2}{44|\theta^n(B)|} \\
&= \liminf_{n\to\infty} \frac{|\theta(A)| + \cdots |\theta^{n-1}(A)| - 4}{88\theta^n(B)} \\
&\geq \liminf_{n\to\infty} \frac{|\theta^{n-1}(A)|}{88 \cdot 6|\theta^{n-1}(A)|} > 0.
\end{aligned}$$

If $\frac{3}{2}H_{n-1} \leq k \leq \frac{1}{2}|\theta^n(B)|$ we can use the same M_n as before. It is split up for ease of reading. Now, the length of M_n is at least $H_{n-1} - \left(\frac{1}{2}H_{n-1} + \frac{1}{2}\left|\theta^{n-1}(A)\right|\right) = \frac{1}{2}\left(H_{n-1} - |\theta^{n-1}(A)|\right)$. Define $t_n = H_{n-1} + 1$. Then for cylinders P_ℓ and P_m you can choose N large enough so that for $n \geq N$ we have

$$T^i x \in P_\ell \text{ if and only if } T^{i+t_n+1}x \in P_\ell$$
$$T^i y \in P_m \text{ if and only if } T^{i+t_n}y \in P_m$$

for all $i \in M_n - \max(|P_\ell|, |P_m|)$.

We can again use the relationship between the heights to obtain $|\theta^n(B)| \leq 20|\theta^{n-2}(A)|$. Define L_n in the same way as before. Thus,

$$\begin{aligned}
\liminf_{n\to\infty} \frac{|M_n|}{|L_n|} &\geq \liminf_{n\to\infty} \frac{\frac{1}{2}\left(H_{n-1} - |\theta^{n-1}(A)|\right)}{44|\theta^n(B)|} \\
&= \liminf_{n\to\infty} \frac{|\theta(A)| + \cdots |\theta^{n-2}(A)|}{88\theta^n(B)} \\
&\geq \liminf_{n\to\infty} \frac{|\theta^{n-2}(A)|}{88 \cdot 20|\theta^{n-2}(A)|} > 0.
\end{aligned}$$

If $0 \leq k \leq \frac{1}{2}H_{n-1}$ then define M_n to be the interval where the second $\overline{D_{n-1}}$ block from x and the second $\overline{D_{n-1}}$ block from y overlap. Then $|M_n| \geq \frac{1}{2}H_{n-1}$. Define $t_n = H_{n-1} + |\theta^{n-1}(A)| + 2$. The remaining part of the argument from the above paragraphs hold here as well.

Therefore in this case the assumptions of Lemma 4.5 hold and $\nu = \mu \times \mu$.

Case 2. Suppose that $x_{\omega_{s_n}}$ is the beginning of $\theta^n(A)$ and $(T^k y)_{\omega_{s_n}}$ is the beginning of $\theta^n(C)$. The block that must precede $\theta^n(C)$ is $\theta^n(A)$ and the possible blocks that can precede $\theta^n(A)$ are $\theta^n(B)$ or $\theta^n(C)$. Since x and y are the same up until the ω_{s_n}-th place, this case isn't possible.

Case 3. Suppose that $x_{\omega_{s_n}}$ is the beginning of $\theta^n(B)$ and $(T^k y)_{\omega_{s_n}}$ is the beginning of $\theta^n(C)$. Since $\theta^n(A)$ must precede a $\theta^n(C)$ block and x and y are the same up until the ω_{s_n}-th place, the block that appears to the left of $\theta^n(B)$ in x must be $\theta^n(A)$. Also, $\theta^n(A)$ must be followed by a $\theta^n(C)$ block. Thus, x and $T^k y$ have the following form around the ω_{s_n}th place:

$$\begin{aligned} x &: A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \\ T^k y &: A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}CA\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \end{aligned}$$

If $H_{n-1} \leq k \leq \frac{1}{2}|\theta^n(B)|$ then define M_n to be the interval where the first $\overline{D_{n-1}}$ block from x and the second $\overline{D_{n-1}}$ block from y overlap. If y is shifted by H_{n-1} then $|M_n| = H_{n-1} - |\theta^n(A)| - 2$ and if y is shifted by $\frac{1}{2}|\theta^n(B)|$ then $|M_n| = \frac{1}{2}H_{n-1} + \frac{1}{2}|\theta^n(A)| - 1$. Thus, $|M_n| \geq H_{n-1} - |\theta^n(A)| - 2$.

Define $t_n = 3H_{n-1} + |\theta^n(A)| + 4$. Then for cylinders P_ℓ and P_m there is an N large enough so that for $n \geq N$ we have

$$T^i x \in P_\ell \text{ if and only if } T^{i+t_n+1}x \in P_\ell$$

$$T^i y \in P_m \text{ if and only if } T^{i+t_n}y \in P_m$$

for all $i \in M_n - \max(|P_\ell|, |P_m|)$.

Define the interval L_n as the symmetric interval that contains the first block where x and y differ. Thus, the length of L_n is at most $44|\theta^n(B)|$. We can use the relationship between the length of the different letters iterated under θ to see that $|\theta^n(B)| \leq 20|\theta^{n-2}(A)|$. Thus,

$$\begin{aligned} \liminf_{n\to\infty} \frac{|M_n|}{|L_n|} &\geq \liminf_{n\to\infty} \frac{H_{n-1} - |\theta^{n-1}(A)| - 2}{44|\theta^n(B)|} \\ &= \liminf_{n\to\infty} \frac{|\theta(A)| + \cdots |\theta^{n-2}(A)| - 2}{44\theta^n(B)} \\ &\geq \liminf_{n\to\infty} \frac{|\theta^{n-2}(A)| - 2}{44 \cdot 20|\theta^{n-2}(A)|} > 0. \end{aligned}$$

If $0 \leq k \leq \frac{1}{2}H_{n-1}$ then define M_n to be the interval where the third $\overline{D_{n-1}}$ block from x and the third $\overline{D_{n-1}}$ block from y overlap. Then $|M_n| \geq \frac{1}{2}H_{n-1}$. Define $t_n = H_{n-1} + 1$.

For cylinders P_ℓ and P_m there is an N large enough so that for $n \geq N$ we have

$$T^i x \in P_\ell \text{ if and only if } T^{i+t_n+1}x \in P_\ell$$

$$T^i y \in P_m \text{ if and only if } T^{i+t_n}y \in P_m$$

for all $i \in M_n - \max(|P_\ell|, |P_m|)$.

Define the interval L_n as before. Thus,

$$\begin{aligned}\liminf_{n\to\infty}\frac{|M_n|}{|L_n|} &\geq \liminf_{n\to\infty}\frac{\frac{1}{2}H_{n-1}}{44|\theta^n(B)|}\\ &= \liminf_{n\to\infty}\frac{|\theta(A)|+\cdots|\theta^{n-1}(A)|}{88\theta^n(B)}\\ &\geq \liminf_{n\to\infty}\frac{|\theta^{n-1}(A)|}{88\cdot 6|\theta^{n-1}(A)|}>0.\end{aligned}$$

The last interval of k values to consider is $\frac{1}{2}H_{n-1}<k<H_{n-1}$. For these values we need two subcases, depending on what follows $\theta^n(B)$ from x and $\theta^n(A)$ from y.

Subcase 3.1. We still have all of the structure assumptions on x and y as in the beginning of Case 3. Additionally, suppose that $\theta^n(A)$ at the "end" of y is followed by $\theta^n(C)$, which in turn must be followed by the block $A\overline{D_{n-1}}$ since all θ^n blocks begin in that manner. Also, notice that $\theta^n(B)$ in x must be followed by $A\overline{D_{n-1}}$ for a similar reason. Hence, x and T^ky have the following form around the ω_{s_n}th place:

$$\begin{aligned} x &: A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}\\ T^ky &: A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}CA\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}CA\overline{D_{n-1}}\end{aligned}$$

Recall that we are assuming $\frac{1}{2}H_{n-1}<k<H_{n-1}$. Define M_n to be the interval where the first $\overline{D_{n-1}}$ block from x and the second $\overline{D_{n-1}}$ block from y overlap. Thus, $|M_n|\geq\frac{1}{2}H_{n-1}-2$. Define t_n to be $t_n=4H_{n-1}+|\theta^n(A)+6$. Then for cylinders P_ℓ and P_m there is N large enough so that for $n\geq N$ we have

$$T^ix\in P_\ell \text{ if and only if } T^{i+t_n}x\in P_\ell$$

$$T^iy\in P_m \text{ if and only if } T^{i+t_n+1}y\in P_m$$

for all $i\in M_n-\max(|P_\ell|,|P_m|)$. The interval L_n is defined in the same way and the length is at most $2(|\omega_{s_n}|+2|\theta^n(B)|+|k|)\leq 46|\theta^n(B)|$. Thus,

$$\begin{aligned}\liminf_{n\to\infty}\frac{|M_n|}{|L_n|} &\geq \liminf_{n\to\infty}\frac{\frac{1}{2}H_{n-1}-2}{46|\theta^n(B)|}\\ &= \liminf_{n\to\infty}\frac{|\theta(A)|+\cdots|\theta^{n-1}(A)|-2}{92\theta^n(B)}\\ &\geq \liminf_{n\to\infty}\frac{|\theta^{n-1}(A)|-2}{92\cdot 6|\theta^{n-1}(A)|}>0.\end{aligned}$$

Subcase 3.2. In this subcase we suppose that, in addition to the structure of x and y defined at the beginning of Case 3, we have that $\theta^n(A)$ at the "end" of y is followed by $\theta^n(B)$, which in turn must be followed by the block $A\overline{D_{n-1}}$. We also assume that $\theta^n(B)$ in x is followed by another $\theta^n(B)$. Notice that the case where $\theta^n(B)$ from x is followed by $\theta^n(A)$ reduces to

Case 1. Under these assumptions x and $T^k y$ have the following form around the ω_{s_n}th place:

$$\begin{array}{rcl} x & : & A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \\ T^k y & : & A\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}CA\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}CA\overline{D_{n-1}}B\overline{D_{n-1}}BD_{n-1}\overline{D_{n-1}}C \end{array}$$

Define M_n to be the interval where the first $\overline{D_{n-1}}$ block from x and the second $\overline{D_{n-1}}$ block from y overlap. Thus, $|M_n| \geq \frac{1}{2}H_{n-1} - 2$. Define t_n to be $t_n = 6H_{n-1} + 2|\theta^n(A)| + 8$. The rest of the argument is the same as Subcase 3.1.

Therefore in this case the assumptions of Lemma 4.5 hold and $\nu = \mu \times \mu$. □

Acknowledgements. The author would like to thank Giovanni Forni for being a valuable mentor and listening to these ideas. The author would like to thank Matthew Yancey for help with the figures that appear in this paper. Also, she would like to thank the referee for valuable comments that made this paper better.

References

1. Jon Aaronson and Benjamin Weiss, *Remarks on the tightness of cocycles*, Colloq. Math. **84/85** (2000), no. part 2, 363–376, Dedicated to the memory of Anzelm Iwanik. MR 1784202 (2001i:37009)
2. Artur Avila and Giovanni Forni, *Weak mixing for interval exchange transformations and translation flows*, Ann. of Math. (2) **165** (2007), no. 2, 637–664. MR 2299743 (2010c:37002)
3. F. Blanchard, A. Maass, and A. Nogueira (eds.), *Topics in symbolic dynamics and applications*, London Mathematical Society Lecture Note Series, vol. 279, Cambridge University Press, Cambridge, 2000, Papers from the CIMPA-UNESCO Summer School held at Universidad de la Frontera, Temuco, January 6–24, 1997. MR 1776753 (2001e:37003)
4. J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Moebius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83. MR 2986954
5. A. del Junco and D. Rudolph, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory Dynam. Systems **7** (1987), no. 4, 531–557. MR 922364 (89e:28029)
6. Sébastien Ferenczi, Charles Holton, and Luca Q. Zamboni, *Joinings of three-interval exchange transformations*, Ergodic Theory Dynam. Systems **25** (2005), no. 2, 483–502. MR 2129107 (2005m:37004)
7. Eli Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003. MR 1958753 (2004c:37011)
8. A. B. Katok, Ja. G. Sinaĭ, and A. M. Stepin, *The theory of dynamical systems and general transformation groups with invariant measure*, Mathematical analysis, Vol. 13 (Russian), Akad. Nauk SSSR Vsesojuz. Inst. Naučn. i Tehn. Informacii, Moscow, 1975, pp. 129–262. (errata insert). MR 0584389 (58 #28430)
9. A. B. Katok and A. M. Stepin, *Approximations in ergodic theory*, Uspehi Mat. Nauk **22** (1967), no. 5 (137), 81–106. MR 0219697 (36 #2776)
10. Anatole Katok, *Interval exchange transformations and some special flows are not mixing*, Israel J. Math. **35** (1980), no. 4, 301–310. MR 594335 (82e:58060)
11. Daniel J. Rudolph, *Fundamentals of measurable dynamics*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1990, Ergodic theory on Lebesgue spaces. MR 1086631 (92e:28006)
12. V. V. Ryzhikov, *Self-joins of commutative actions with an invariant measure*, Mat. Zametki **83** (2008), no. 5, 792–795. MR 2451369 (2010c:37007)
13. P. Sarnak, *Three lectures on the mobius function randomness and dynamics*, publications/ias.edu/sarnak/ (2011).

14. Younghwan Son, *Substitutions, tiling dynamical systems and minimal self-joinings*, Discrete Contin. Dyn. Syst. **34** (2014), no. 11, 4855–4874. MR 3223833
15. William A. Veech, *The metric theory of interval exchange transformations. I. Generic spectral properties*, Amer. J. Math. **106** (1984), no. 6, 1331–1359. MR 765582 (87j:28024a)
16. M. Viana, *Dynamics of interval exchange transformations and teichmuller flows*, http://w3.impa.br/ viana/out/ietf.pdf.

Kelly B. Yancey, Department of Mathematics, University, Mathematics Building, College Park, MD 20742-4015
E-mail address: kbyancey1@gmail.com