CHARACTERISTIC MEASURES FOR LANGUAGE STABLE SUBSHIFTS

VAN CYR AND BRYNA KRA

Abstract. We consider the problem of when a symbolic dynamical system supports a Borel probability measure that is invariant under every element of its automorphism group. It follows readily from a classical result of Parry that the full shift on finitely many symbols, and more generally any mixing subshift of finite type, supports such a measure. Frisch and Tamuz recently dubbed such measures characteristic, and further showed that every zero entropy subshift has a characteristic measure. While it remains open if every subshift over a finite alphabet has a characteristic measure, we define a new class of shifts, which we call language stable subshifts, and show that these shifts have characteristic measures. This is a large class that is generic in several senses and contains numerous positive entropy examples.

1. Introduction

Suppose X is a compact metric space and $T\colon X \to X$ is a homeomorphism. The Krylov-Bogolioubov Theorem says that there exists a Borel probability measure μ that is supported on X and is invariant under T, meaning that $T_*\mu = \mu$. Using tools of ergodic theory to study the measure-preserving system (X, T, μ), we can often obtain information about the topological dynamical system (X, T). While the homeomorphism σ determines a $\mathbb{Z}$-action on X, there is another (often larger) action on X determined by the automorphism group

$$\mathrm{Aut}(X) := \{\psi \in \mathrm{Homeo}(X)\colon \psi T = T\psi\}$$

of all self-conjugacies of (X, T). This is a natural approach for studying when two topological systems (X, T) and (Y, S) are topologically conjugate: if (X, T) and (Y, S) are conjugate as $\mathbb{Z}$-systems, then $\mathrm{Aut}(X) \cong \mathrm{Aut}(Y)$ as groups, and $(X, \mathrm{Aut}(X))$ and $(Y, \mathrm{Aut}(Y))$ are topologically conjugate as $\mathrm{Aut}(X)$-systems. Systems that can be distinguished as $\mathrm{Aut}(X)$-systems can therefore also be distinguished as $\mathbb{Z}$-systems, and the action of the larger group sometimes makes it easier to distinguish non-conjugate systems.

With this in mind, it is natural to ask if there is an analog of the Krylov-Bogolioubov Theorem for the action of the automorphism group of a dynamical system. Using terminology introduced by Frisch and Tamuz [14], if (X, T) is a topological dynamical system and $\mathrm{Aut}(X)$ denotes its automorphism group, a Borel probability measure μ is *characteristic* for (X, σ) if $\psi_*\mu = \mu$ for all $\psi \in \mathrm{Aut}(X)$. They [14, Question 1.3] ask: does every subshift (over a finite alphabet) have a characteristic measure?

2010 *Mathematics Subject Classification.* 37B10, 68R15.

Key words and phrases. subshift, block complexity.

The second author was partially supported by NSF grant DMS-1800544.

For a subshift, the classical theorem of Curtis, Hedlund, and Lyndon states that every automorphism is given as a sliding block code, and in particular it follows that there are only countably many automorphisms. However, the group of automorphisms can be quite complicated: for example (see [16, 5]), the automorphism group of any mixing shift of finite type contains isomorphic copies of any finite group, the free group on any number of generators, along with copies of many other known groups. A natural way to attempt to answer Frisch and Tamuz's question is to find all of the automorphisms of a subshift and all of the ergodic measures, and then check if any of these measures is a characteristic measure. Unfortunately, this is not a practical method, as many easy to state questions even about full shifts remain open. For example, we do not have enough information about the automorphism group of a full shift to determine whether the automorphism group of the full shift on two symbols is isomorphic to that on three symbols. While this type of strategy is doomed to failure with current tools, there are other methods for showing the existence of a characteristic measure without knowing either what all of the automorphisms are, or much about the simplex of invariant measures, $\mathcal{M}(X)$.

We defer the precise definitions and explanations of these results until Section 2, but give a quick summary of four currently known methods for proving that a subshift (X, σ) has a characteristic measure:

(1) If $\mathrm{Aut}(X)$ is amenable, this follows from the Krylov-Bogolioubov Theorem.
(2) When the topological entropy $h_{\mathrm{top}}(X)$ is 0, this is the main result in Frisch and Tamuz [14].
(3) If there exists a σ-invariant probability measure μ supported on X such that
$$\{\nu \in \mathcal{M}(X) \colon (X, \sigma, \mu) \text{ is measurably isomorphic to } (X, \sigma, \nu)\}$$
is finite, then X has a characteristic measure.
(4) If there exists a closed subshift $Y \subseteq X$ that supports an $\mathrm{Aut}(Y)$-characteristic measure and there are only finitely many $Z \subseteq X$ such that (Y, σ) is topologically conjugate to (Z, σ), then X has a characteristic measure.

We discuss the first two methods further in Section 2; properties (3) and (4) are implicit in the literature and are described in Lemmas 3.1 and 3.3. In light of Frisch and Tamuz's theorem, listed as the method (2), determining whether every symbolic system has a characteristic measure comes down to developing techniques for finding characteristic measures in symbolic systems of positive entropy.

Our main result gives a large class of symbolic systems, containing numerous systems of positive entropy, that admit characteristic measures. To define this class, we introduce a new condition called language stability (see Section 4). This condition is defined by considering the words that do not appear in the language of the system, and placing restrictions on the frequency with which such "forbidden words" arise. We show language stability guarantees the existence of a characteristic measure:

Theorem 1.1. *Every language stable subshift supports a characteristic measure.*

In fact we deduce (see Corollary 4.2) that the characteristic measure we find is a measure of maximal entropy for the subshift. Furthermore, in Section 5 we show that the set of language stable subshifts is generic, in a strong sense. We show that the language stable shifts are a dense G_δ, with respect to the Hausdorff topology, in both the space of all subshifts with a fixed alphabet and in the subspace of positive entropy subshifts with a fixed alphabet.

To further indicate that our theorem identifies characteristic measures not previously known to exist, in Section 6 we build an example of a language stable subshift to which none of the four previously known methods of proving the existence of characteristic measures seem to apply. More precisely, the automorphism group of this system is not amenable, the system has positive entropy, and every measure supported on the system is measurably isomorphic to infinitely many other measures on the system, ensuring that none of the first three conditions is satisfied. The fourth condition is not as easily ruled out by our construction, and it is possible that it is satisfied. However, the application of Theorem 1.1 to our (language stable) example produces a characteristic measure that is a measure of maximal entropy, and even if our example did satisfy the fourth criterion, we show that method could not produce a measure of maximal entropy in our system. Therefore our result guarantees the existence of a characteristic measure of maximal entropy that could not have been seen to exist by any of the four previously described methods (and it is possible that none of those methods could be used to show the existence of any characteristic measure for our system).

Our example is a product of three systems, and the properties of these systems may be of independent interest. Further, we emphasize that our theorem shows the existence of a characteristic measure in our system without our having to either determine the algebraic structure of its group of automorphisms or describe all of its invariant measures. This is highly advantageous: even if we could describe the automorphism group of each of the three systems, the automorphism group of their product may be significantly more complicated than just the product of their individual automorphism groups. By way of example, we mention a recent theorem of Salo and Schraudner [25] that if $X \subseteq \{0,1\}^{\mathbb{Z}}$ is the "sunny side up shift" consisting of all elements of $\{0,1\}^{\mathbb{Z}}$ that have at most one 1, then $\mathrm{Aut}(X) \cong \mathbb{Z}$ while $\mathrm{Aut}(X \times X) \cong (\mathbb{Z}^{\infty} \rtimes S_{\infty}) \rtimes (\mathbb{Z}^{\infty} \rtimes S_{\infty})$.

We conclude this introduction with an application to further motivate the study of characteristic measures. Beyond existence of a characteristic measure for a subshift, it is a natural question whether knowledge of such a measure gives practical information about the subshift. There has been significant interest in determining the algebraic properties of $\mathrm{Aut}(X)$ for different subshifts X. Most of these advances begin with assumptions on X, such as a constraint on some type of complexity of the shift, like the growth rate of the block complexity function or the visiting complexity for recurrence, or structure in the dynamics of the shift, such as being a shift of finite type or being a Toeplitz shift. One then uses these constraints to show that $\mathrm{Aut}(X)$ either must have or cannot have certain algebraic properties. Another, less explored, way to study $\mathrm{Aut}(X)$ on some explicitly given shift is to find small range block codes that define automorphisms and study the relations in the subgroup of $\mathrm{Aut}(X)$ that they generate. This approach seems fruitful both because of its computational nature and because there are many natural questions about $\mathrm{Aut}(X)$ that, so far, have resisted being answered with previously developed approaches. As a specific example, it is unknown whether there exists a subshift whose automorphism group contains a finitely generated, nonabelian, infinite, nilpotent group. One could imagine a computational approach to this problem where a specific subshift (X, σ) is selected, the block codes of small range are checked to see which ones determine automorphisms, and it is checked whether, among them, there exist $\varphi, \psi \in \mathrm{Aut}(X)$ of infinite order such that

$$\theta = [\varphi, \psi] = \varphi\psi\varphi^{-1}\psi^{-1} \neq \mathrm{Id}, \quad \varphi\theta = \theta\varphi, \quad \text{and} \quad \psi\theta = \theta\psi.$$

Then the group generated by these relations would be an infinite, nonabelian, quotient of the discrete Heisenberg group. For explicitly given block codes, we note that these relations can be checked easily simply by finding the block code each of these group elements defines and checking to see if it is the identity (a different approach would be needed to check if φ and ψ have infinite order). A challenge to investigating the structure of $\mathrm{Aut}(X)$ in this way, however, is that we must find block codes that determine automorphisms of X. This leads to the question:

Question 1.2. *Given a subshift X and a block code φ defined on X, determine whether φ determines an automorphism of X.*

In Section 7, we explain how a characteristic measure gives rise to a set of nontrivial conditions on a block code that are necessary for it to be invertible. We give an example where they can be used to show a certain block code does not have an inverse. Although these conditions are not sufficient for a block code to have an inverse, they do allow one to eliminate many block codes as candidates for being automorphisms.

Acknowledgment. We thank Anthony Quas for helpful conversations during the preparation of this article.

2. Preliminaries and Notation

2.1. **Upper Banach density.** If $S \subseteq \mathbb{N}$, then the *upper Banach density* $d^*(S)$ of S is defined to be

$$d^*(S) := \limsup_{n\to\infty} \max_k \frac{|S \cap \{k, k+1, k+2, \dots, k+n-1\}|}{n}.$$

A set S has upper Banach density 1 if and only if there are arbitrarily long runs of consecutive integers that are all elements of S.

2.2. **Symbolic systems.** Let $\mathcal{A}$ be a finite set and let $\mathcal{A}^{\mathbb{Z}}$ be the set of functions $x \colon \mathbb{Z} \to \mathcal{A}$, and denote $x \in \mathcal{A}^{\mathbb{Z}}$ as $x = (x_i)_{i\in\mathbb{Z}}$. The space $\mathcal{A}^{Z}$ is a compact metric space when endowed with the metric

$$d\big((x_i), (y_i)\big) := 2^{-\inf\{|i| \colon x_i \neq y_i\}}.$$

The *left-shift* map $\sigma \colon \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ defined by $(\sigma x)_i := x_{i+1}$ for all $i \in \mathbb{Z}$ is a homeomorphism. For each $w = (w_0, \dots, w_{n-1}) \in \mathcal{A}^n$ the *cylinder set* is

$$[w]_0^+ := \{x \in \mathcal{A}^{\mathbb{Z}} \colon x_i = w_i \text{ for all } 0 \leq i < n\}$$

and the collection of sets $\{\sigma^i[w]_0^+ \colon w \in \mathcal{A}^*,\ i \in \mathbb{Z}\}$ gives a basis for the topology of $\mathcal{A}^{\mathbb{Z}}$. If $X \subset \mathcal{A}^{\mathbb{Z}}$ is closed and σ-invariant, then (X, σ) is a *subshift*, and when the shift σ is understood from the context, we omit the transformation from the notation and refer to $X \subset \mathcal{A}^{\mathbb{Z}}$ as a subshift.

The *language* of a subshift (X, σ) is

$$\mathcal{L}(X) := \{w \in \mathcal{A}^* \colon [w]_0^+ \cap X = \emptyset\}$$

and any $w \in \mathcal{L}(X)$ is called a *word* in the language ($w \in \mathcal{L}(X)$ is sometimes referred to as a *factor* in the literature). The *complexity* $P_X(n)$ counts the number of words of length n in the language $\mathcal{L}(X)$.

Homeo(X) forms a group under composition and Aut(X) is the centralizer of σ in Homeo(X). A map $\psi\colon X \to \mathcal{A}^{\mathbb{Z}}$ is called a *sliding block code* if there exists $R \in \mathbb{N}$ and a map $\Psi\colon \mathcal{L}_{2R+1}(X) \to \mathcal{A}$ such that for all $x \in X$ and $i \in \mathbb{Z}$, we have

$$(\psi x)_i = \Psi(x_{i-R}, \ldots, x_i, \ldots, x_{i+R}).$$

In this case the number R is called a *range* for ψ. A classical result characterizes the automorphisms of a subshift:

Theorem 2.1 (Curtis-Hedlund-Lyndon Theorem [16])**.** *Every element of* Aut(X) *is a sliding block code. Conversely, any sliding block code from X to X that has a sliding block code inverse is an element of* Aut(X)*.*

Following Frisch and Tamuz [14], we define the central object of study:

Definition 2.2. *Let (X, T) be a topological dynamical system and let*

$$\mathrm{Aut}(X) := \{\psi \in \mathit{Homeo}(X) \colon \psi T = T\psi\}$$

be its automorphism group. A Borel probability measure μ is characteristic *for (X, T) if $\psi_*\mu = \mu$ for all $\psi \in \mathrm{Aut}(X)$.*

We study this notion for subshift (X, σ).

2.3. **Coding of orbits.** Assume (X, T) is an invertible topological dynamical system and let $\mathcal{P}$ be a partition of the space X into finitely many sets, meaning that $\mathcal{P} = \{P_1, \ldots, P_n\}$ for some sets $P_i \subset X$ satisfying $\bigcup_{i=1}^n P_i = X$. Note that we make no assumption that the sets P_i are open and require that their union cover all of X. If $x \in X$, then the *coding of the orbit of* x is the sequence $(x_j)_{j\in\mathbb{Z}}$ defined by $x_j = i$ if and only if $T^j x \in P_i$. The *coding of the system* (X, T) is the symbolic system obtained by taking the closure of the codings of all $x \in X$, and it is easy to check that this is a subshift of $\{1, \ldots, n\}^{\mathbb{Z}}$.

2.4. **The forbidden word construction.** We recall how to construct subshifts by specifying a list of forbidden words. Let $\mathcal{F} \subseteq \mathcal{A}^*$ be a (finite or infinite) set of words. Define

$$Y_{\mathcal{F}} := \{y \in \mathcal{A}^{\mathbb{Z}} \colon \sigma^i y \notin [w]_0^+ \text{ for all } w \in \mathcal{F} \text{ and all } i \in \mathbb{Z}\}.$$

It is immediate that $Y_{\mathcal{F}}$ is a subshift and we say that $\mathcal{F}$ is a list of *forbidden words* in $Y_{\mathcal{F}}$. If X is a subshift, define

$$\mathcal{F}(X) := \{w \in \mathcal{A}^* \colon w \notin \mathcal{L}(X)\}.$$

It follows immediately from the definitions that $X \subseteq Y_{\mathcal{F}(X)}$ and it follows from the compactness of X that $Y_{\mathcal{F}(X)} \subseteq X$. Thus $X = Y_{\mathcal{F}(X)}$. In other words, every subshift can be obtained by specifying an appropriate set of forbidden words.

On the other hand, if $\mathcal{F}$ is a set of forbidden words and

$$\mathcal{M} = \{w \in \mathcal{F} \colon \text{ no proper subword of } w \text{ lies in } \mathcal{F}\}$$

then $Y_{\mathcal{M}} = Y_{\mathcal{F}}$. Thus different sets can present the same subshift through the forbidden words construction. But we can obtain a canonical set $\mathcal{M}$ of forbidden words, the minimal ones, if for all $w \in \mathcal{M}$, no proper subword of w also lies in $\mathcal{M}$. In other words, every set of forbidden words has a canonical subset of minimal forbidden words that presents the same subshift via the forbidden words construction. Thus for a subshift X, we define the *minimal forbidden words* $\mathcal{M}(X)$ by setting

$$\mathcal{M}(X) := \{w \in \mathcal{F}(X) \colon \text{ no proper subword of } w \text{ is in } \mathcal{F}(X)\}.$$

For each $n \in \mathbb{N}$ we define $\mathcal{M}_n(X) := \mathcal{M}(X) \cap \mathcal{A}^n$ to be the set of minimal forbidden words of length n in X. The growth rate of the minimal forbidden words was shown to be a conjugacy invariant in Béal, Mignosi, and Restivo [7], and properties of the subshift that can be deduced from its presentation via the minimal forbidden words is further studied in [6, 22].

2.5. **Subshifts of finite type and the cover of a subshift.** A subshift X is called a *subshift of finite type* (SFT) if $\mathcal{M}(X)$ is finite. Bowen [8] gave an equivalent formulation (see [20, Theorem 2.1.8] for a proof):

Proposition 2.3. *The shift (X, σ) is a subshift of finite type if and only if there exists $g \geq 0$ such that whenever $uw, wv \in \mathcal{L}(X)$ and $|w| \geq g$ then we also have $uwv \in \mathcal{L}(X)$.*

For any subshift X there is a well-known way to write it as the intersection of a descending chain of subshifts of finite type: for each $n \in \mathbb{N}$, define

$$X_n := Y_{\bigcup_{k=1}^n \mathcal{M}_k(X)}$$

to be the subshift of finite type using the forbidden words up to length n in X. Then

$$X_1 \supseteq X_2 \supseteq X_3 \supseteq \cdots \supseteq X_n \supseteq \dots$$

and $X = \bigcap_{n=1}^\infty X_n$. This sequence $\{X_n\}_{n=1}^\infty$ is called the *SFT cover* of X.

A subshift X is *topologically transitive* if there exists some $x \in X$ such that

$$X = \overline{\{\sigma^i x \colon i \in \mathbb{Z}\}}.$$

It is *forward transitive* if there exists some $x \in X$ such that

$$X = \overline{\{\sigma^i x \colon i \in \mathbb{N}\}}.$$

Parry [23] showed that any forward transitive subshift of finite type is *intrinsically ergodic*, meaning there is a unique σ-invariant measure μ, supported on X, such that $h_\mu(X) = h_{\text{top}}(X)$ (in other words, the system has a unique measure of maximal entropy). For general (not necessarily transitive) subshifts of finite type, we record the following elementary lemma:

Lemma 2.4. *Every subshift of finite type supports at most finitely many ergodic measures of maximal entropy.*

Proof. Let X be an subshift of finite type whose forbidden words all have length at most n. For any $u, v \in \mathcal{L}_n(X)$, say that $u \sim v$ if there exist $m_1, m_2 > 0$ such that $[u] \cap \sigma^{m_1}[v] \neq \emptyset$ and $[v] \cap \sigma^{m_2}[u] \neq \emptyset$. Let

$$\mathcal{W} := \{w \in \mathcal{L}_n(X) \colon w \sim w\}.$$

Since X is an subshift of finite type whose forbidden words have length at most n, note that $\sim$ is an equivalence relation on the set $\mathcal{W}$ (it would not necessarily be a transitive relation for more general shifts).

Let μ be an ergodic measure of maximal entropy. For any $u, v \in \mathcal{L}_n(X)$, by ergodicity it follows that for μ-almost every $x \in X$

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=0}^{m-1} \mathbf{1}_{[u]}(\sigma^k x) = \mu([u]) \quad \text{and} \quad \lim_{m\to\infty} \frac{1}{m} \sum_{k=0}^{m-1} \mathbf{1}_{[v]}(\sigma^k x) = \mu([v])$$

and also

$$\lim_{m\to\infty} \frac{1}{2m+1} \sum_{k=-m+1}^{m-1} \mathbf{1}_{[v]}(\sigma^k x) = \mu([v]).$$

It follows from the first two equalities that either $\mu([u]) \cdot \mu([v]) = 0$ or $u \sim v$. Using the third equality, it follows that if $\mu([v]) > 0$ then for μ-almost every $x \in X$ and for every $m \in \mathbb{Z}$ we have $v \sim (x_m, x_{m+1}, \dots, x_{m+n-1})$. There must be at least one $v \in \mathcal{L}_n(X)$ for which $\mu([v]) > 0$, and for this v we have

$$\mu\left(\bigcup_{u\sim v} [u]\right) = 1.$$

Furthermore, the measure μ is supported on the sub-subshift of finite type $X(v)$, defined by forbidding all words forbidden in X and forbidding all words of length n that are not equivalent to v. Note that $X(v)$ is a forward transitive subshift of finite type and so by Parry's Theorem, $X(v)$ has a unique measure of maximal entropy. But μ is an example of such a measure, since $h_\mu(X(v)) \leq h_{\text{top}}(X(v)) \leq h_{\text{top}}(X) = h_\mu(X) = h_\mu(X(v))$ (since μ is supported on $X(v)$ and is a measure of maximal entropy on X) and so all of the inequalities are actually equalities.

Thus every ergodic measure of maximal entropy on X is supported on (and is a measure of maximal entropy on) a forward transitive subshift of finite type obtained by forbidding words of length n in X. Since $\mathcal{L}_n(X)$ is finite, there can only be finitely many such forward transitive subshifts of finite type that arise from this construction, each of which supports a unique measure of maximal entropy. Thus X supports only finitely many measures of maximal entropy. □

3. Previously known criteria implying the existence of characteristic measures

3.1. **Finitely many measurably isomorphic systems.** Suppose (X, σ) is a subshift and μ is a σ-invariant measure supported on X. A key tool that underlies most of the cases where it is known how to prove that characteristic measures exist, is the following: if $\psi \in \mathrm{Aut}(X)$, then (X, σ, μ) and $(X, \sigma, \psi_*\mu)$ are conjugate as measure-preserving systems. This gives the following criterion (which is used implicitly in the literature, for example in [14]) for establishing the existence of characteristic measures:

Lemma 3.1. *Let (X, σ) be a subshift and suppose there exists a σ-invariant measure μ supported on X for which the set*

$$\mathcal{I}(\mu) := \{\nu \colon \nu \text{ is a } \sigma\text{-invariant measure supported on } X \text{ and } (X,\sigma,\mu) \cong (X,\sigma,\nu)\}$$

is finite. Then X supports a characteristic measure.

Proof. Define

$$\xi := \frac{1}{|\mathcal{I}(\mu)|} \sum_{\nu \in \mathcal{I}(\mu)} \nu.$$

Then ξ is a σ-invariant measure supported on X. For any $\psi \in \mathrm{Aut}(X)$, note that ψ_* induces a permutation on $\mathcal{I}(\mu)$ by finiteness of $\mathcal{I}(\mu)$ and the fact that $(X, \sigma, \nu) \cong (X, \sigma, \varphi_* \nu)$ for any $\nu \in \mathcal{I}(\mu)$. Thus $\psi_* \xi = \xi$. □

This gives the existence of a characteristic measure for any full shift, and more generally any shift of finite type:

Corollary 3.2. *Every subshift of finite type has a characteristic measure.*

Proof. Let X be a subshift of finite type. The entropy map, $\mu \mapsto h_\mu(\sigma)$, is upper semi-continuous (see [27, Theorem 8.2]), and by [27, Theorem 8.7, Part (v)], the set of measures of maximal entropy is nonempty, and using Part (iii) of the same theorem, there is an ergodic measure of maximal entropy. By Lemma 2.4, there are only finitely many ergodic measures of maximal entropy. Since the properties of being ergodic and being a measure of maximal entropy are preserved under measure theoretic isomorphism, Lemma 3.1 guarantees that X supports a characteristic measure. □

The special case that the subshift of finite type is forward transitive is not a new result (we did not find the non-transitive case in the literature, but expect Corollary 3.2 to be no surprise to experts). Coven and Paul [9] showed that any automorphism preserves entropy and Parry [23] showed that a forward transitive subshift of finite type has a unique measure of maximal entropy. It follows that this unique measure (of maximal entropy) is preserved under the full automorphism group. However we need further use of some of the tools used to prove Corollary 3.2 as well as its result in our arguments.

3.2. **Finitely many topologically conjugate systems.** We also make use of a topological analog of Lemma 3.1:

Lemma 3.3. *Let (X, σ) be a subshift and suppose there exists a closed subshift $Y \subseteq X$ that supports an* $\mathrm{Aut}(Y)$*-characteristic measure and for which the set*

$$\mathcal{J}(Y) := \{Z \subseteq X \colon (Y, \sigma) \text{ is topologically conjugate to } (Z, \sigma)\}$$

is finite. Then X supports a characteristic measure.

Proof. If $Z \in \mathcal{J}(Y)$, then using the topological conjugacy between (Y, σ) and (Z, σ) and pushing forward the characteristic measure on $\mathrm{Aut}(Y)$, it follows that Z supports an $\mathrm{Aut}(Z)$-characteristic measure (in fact $\mathrm{Aut}(Y) \cong \mathrm{Aut}(Z)$ in this case). Let ν be an $\mathrm{Aut}(Y)$-characteristic measure supported on Y.

Every element of $\mathrm{Aut}(X)$ induces a permutation on $\mathcal{J}(Y)$. Thus we can define the group

$$\Pi := \{\pi \in \mathrm{Sym}(\mathcal{J}(Y)) \colon \psi \text{ induces the permutation } \pi \text{ for some } \psi \in \mathrm{Aut}(X)\}.$$

For each $\pi \in \Pi$, choose an automorphism $\phi(\pi) \in \mathrm{Aut}(X)$ that induces the permutation π on $\mathcal{J}(Y)$. Finally define

$$\xi := \frac{1}{|\Pi|} \sum_{\pi \in \Pi} \phi(\pi)_* \nu.$$

If $\psi \in \mathrm{Aut}(X)$ and α is the permutation on $\mathcal{J}(Y)$ induced by ψ,

$$\begin{aligned}\psi_*\xi &= \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\psi_*\phi(\pi)_*\nu = \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\phi(\alpha)_*\phi(\pi)_*\nu \\ &= \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\phi(\alpha\pi)_*\nu = \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\phi(\pi)_*\nu = \xi,\end{aligned}$$

where the penultimate equality holds because $\alpha\pi$ runs over every element of Π exactly once as π runs over Π. □

This means, for example, that any subshift that contains periodic points has a characteristic measure, a fact already pointed out in Fricsch and Tamuz [14]. A version of this fact about periodic points also appears implicitly in Boyle and Krieger [4], in their defining and study of the gyration function. But by making use of the main result in [14], this also means that if X has only finitely many minimal subsystems of topological entropy zero, then X has a characteristic measure.

3.3. **Invariant measures for amenable actions.** One final tool for proving the existence of characteristic measures, as a classical result in the literature, is the Krylov-Bogolioubov Theorem for actions of amenable groups. An immediate application of it yields:

Theorem 3.4 (Krylov-Bogolioubov Theorem; see [19, 3, 1])**.** *Let X be a subshift and suppose the countable, discrete group* $\mathrm{Aut}(X)$ *is amenable. Then X supports a characteristic measure.*

4. Language stable shifts

4.1. **Defining language stable shifts.** Recall that if X is a subshift and $n \in \mathbb{N}$, then $\mathcal{M}_n(X)$ denotes the set of minimal forbidden words of length n in $\mathcal{L}(X)$. We make the following new definition.

Definition 4.1. *A subshift (X,σ) is* language stable *if the set*

$$\{n \in \mathbb{N} \colon \mathcal{M}_n(X) = 0\}$$

has upper Banach density 1.

Note that any subshift of finite type is language stable, because in this case $\mathcal{M}_n(X) = 0$ for all but finitely many $n \in \mathbb{N}$. Moreover, any subshift X can be approximated arbitrarily well by unforbidding enough of its forbidden words to make it satisfy Definition 4.1 (see Section 5 for discussion of the metric on subshifts). More precisely, let X be any subshift and let $S \subseteq \mathbb{N}$ be any set with upper Banach density 1. Define

$$\mathcal{M} := \bigcup_{n\notin S}\mathcal{M}_n(X).$$

Then $Y_{\mathcal{M}}$ is language stable, X is a subshift of $Y_{\mathcal{M}}$, and by choosing S to be very sparse we can ensure that $\mathcal{M}(Y) \setminus \mathcal{M}(X)$ is also very sparse.

4.2. **Language stable shifts support a characteristic measure.** We use the tools of the last two sections to prove our main theorem:

Proof of Theorem 1.1. Assume that X be a language stable subshift. If X is a subshift of finite type, then the result follows from Corollary 3.2. Thus it suffices to assume that X is not a subshift of finite type.

Let $\{X_n\}_{n=1}^{\infty}$ be the SFT cover of X. By Lemma 2.4, each X_n supports at most finitely many measures of maximal entropy. By Lemma 3.1, each X_n carries an $\mathrm{Aut}(X_n)$-characteristic measure ξ_n. Thus for each $n \in \mathbb{N}$, we have that ξ_n is a σ-invariant measure supported on (in general, a proper subset of) $\mathcal{A}^{\mathbb{Z}}$. By assumption, the set

$$S := \{n \in \mathbb{N} \colon \mathcal{M}_n(X) = 0\}$$

has upper Banach density 1. Let $r \colon S \to \mathbb{N}$ be the function

$$r(s) = -s + \min\{t \notin S \colon t > s\}$$

defined to be the longest run of consecutive integers that lie in S, starting from s (since X is not an subshift of finite type, this run is well-defined). By assumption, the set S has upper Banach density 1, and so we can choose a subset $S' \subseteq S$ along which $r(s)$ is strictly increasing.

Since $\mathcal{A}^{\mathbb{Z}}$ is a compact metric space, it follows from the Banach-Alaoglu Theorem that the set $\{\xi_{s'+r(s')-1}\}_{s \in S'}$ has weak* accumulation points in the set of all Borel measures on $\mathcal{A}^{\mathbb{Z}}$. Let ξ be one such accumulation point. We claim that ξ is a characteristic measure for X. Passing from S' to S'' if necessary, we can assume the weak* limit of the sequence $\{\xi_{s'+r(s')-1}\}_{s' \in S'}$ exists.

Let $\psi \in \mathrm{Aut}(X)$ be fixed. By the Curtis-Hedlund-Lyndon Theorem (Theorem 2.1), ψ is a block code of some range $R \geq 0$. Let $\Psi \colon \mathcal{L}_{2R+1}(X) \to \mathcal{A}$ be such that for all $x \in X$ and all $i \in \mathbb{Z}$ we have

$$(\psi x)_i = \Psi(x_{i-R}, \ldots, x_i, \ldots, x_{i+R}).$$

If $R' \geq R$, then ψ is also a block code of range R' and so we may assume (increasing the value of R if necessary) that R is the range for both ψ and ψ^{-1}. From hereon, we fix such R. By abusing notation, we now extend the domain of ψ to include any element of $\mathcal{A}^{\mathbb{Z}}$ to which the range R block code defining ψ can be applied. Since the function $r(s)$ is strictly increasing on S', there exists some M such that $r(s') \geq 2R+1$ for all $s' \in S'$ satisfying $s' > M$; let some such s' be fixed. Notice that

$$X_{s'} = X_{s'+1} = \cdots = X_{s'+r(s')-1} \tag{1}$$

since these are all subshifts of finite type with identical sets of minimal forbidden words. Furthermore, since $X \subseteq X_{s'+r(s')-1}$ and there is no word of length at most $s' + r(s') - 1$ forbidden in X that was not also forbidden in $X_{s'+r(s')-1}$, it follows that

$$\mathcal{L}_{s'+r(s')-1}(X) = \mathcal{L}_{s'+r(s')-1}(X_{s'+r(s')-1}).$$

For each $w \in \mathcal{L}_{s'+r(s')-1}(X)$, applying the sliding block code Ψ determines a word $\Psi(w) \in \mathcal{L}_{s'+r(s')-1-2R}(X)$. This implies that

$$\mathcal{L}_{s'+r(s')-1-2R}(\psi(X_{s'+r(s')-1})) \subseteq \mathcal{L}_{s'+r(s')-1-2R}(X_{s'+r(s')-1-2R})$$

and so $\psi(X_{s'+r(s')-1}) \subseteq X_{s'+r(s')-1-2R}$. Since $r(s') \geq 2R+1$, it follows from this and (1) that $\psi(X_{s'+r(s')-1}) \subseteq X_{s'+r(s')-1}$.

Applying the analogous argument with ψ^{-1} instead of ψ shows that $\psi^{-1}(X_{s'+r(s')-1}) \subseteq X_{s'+r(s')-1}$, and so

$$\psi(X_{s'+r(s')-1}) = X_{s'+r(s')-1}.$$

In other words, $\psi \in \mathrm{Aut}(X_{s'+r(s')-1})$. But by definition, $\xi_{s'+r(s')-1}$ is a $\mathrm{Aut}(X_{s'+r(s')-1})$-characteristic measure and so

$$\psi_*\xi_{s'+r(s')-1} = \xi_{s'+r(s')-1}. \tag{2}$$

Since (2) holds for any $s' \in S'$ such that $s' > M$, it follows that

$$\psi_*\xi = \psi_*\left(\lim_{s'\in S'} \xi_{s'+r(s')-1}\right) = \lim_{s'\in S'} \psi_*\xi_{s'+r(s')-1} = \lim_{s'\in S'} \xi_{s'+r(s')-1} = \xi.$$

Since this holds for any $\psi \in \mathrm{Aut}(X)$, we have that ξ is a characteristic measure for X. □

Note that in the proof, the measure ξ produced is a weak-* limit of the measures ξ_n that are measures of maximal entropy on the shifts of finite type X_n. For each $n \in \mathbb{N}$ we have $h_{top}(X_n) \geq h_{top}(X)$ (since $X \subseteq X_n$) and so $h_{\xi_n}(\sigma) \geq h_{top}(X)$. By upper semi-continuity of the entropy map (see [27, Theorem 8.2]), it follows that $h_\xi(\sigma) \geq h_{top}(X)$. Since ξ is supported on X we have equality, so it is a measure of maximal entropy on X. Therefore our proof shows that:

Corollary 4.2. *Any language stable shift has a characteristic measure that is a measure of maximal entropy.*

5. Genericity of language stable shifts

We show that the set of language stable shifts is a dense G_δ, with respect to the Hausdorff topology, in both the space of all subshifts with a fixed alphabet and in the subspace of positive entropy subshifts with a fixed alphabet.

We start by defining the distance between two subshifts over the same alphabet: if $\mathcal{A}$ is a finite alphabet and $X, Y \subseteq \mathcal{A}^\mathbb{Z}$ are two subshifts, define

$$d(X,Y) := 2^{-\inf\{n\colon \mathcal{L}_n(X)\neq\mathcal{L}_n(Y)\}}. \tag{3}$$

This metric is equivalent to the usual Hausdorff metric but is more convenient for our purposes, and two shifts are close if their sequences agree on large (finite) subsets of the integers. Endowed with the metric (3), the space of all subshifts of $\mathcal{A}^\mathbb{Z}$ is a compact metric space (recall that by definition, a subshift X is a closed subset of $\mathcal{A}^\mathbb{Z}$).

Fixing a finite alphabet $\mathcal{A}$, a property of subshifts is said to be *generic* if it holds for a dense G_δ subset of the space of all subshifts, in this topology. A property is *generic among shifts of positive entropy* if it defines a dense G_δ (in the induced topology) in the subspace of subshifts of positive entropy. Let S denote the set of all subshifts of $\mathcal{A}^\mathbb{Z}$ and let S^+ denote the set of all subshifts of $\mathcal{A}^\mathbb{Z}$ with positive topological entropy. For any $c > 0$, let $S^+_{\geq c}$ denote the set of all subshifts of $\mathcal{A}^\mathbb{Z}$ with topological entropy greater than or equal to c.

Frisch and Tamuz [14] show that subshifts with zero entropy are generic in the space of all shifts, and that more generally subshifts with entropy c are generic in the space of all shifts with entropy at least c. Along these lines, we prove that thee language stable subshifts are generic. We start by showing that these shifts are a G_δ:

Theorem 5.1. *The set of language stable subshifts in S is a G_δ subset of S.*

Proof. A subshift X is language stable if and only if for all $k \in \mathbb{N}$ there exists $n_k \in \mathbb{N}$ such that

$$X_{n_k} = X_{n_k+1} = X_{n_k+2} = \cdots = X_{n_k+k-1},$$

where $\{X_n\}_{n=1}^\infty$ denotes the SFT cover of X, and we show that the set of subshifts in S having this property is G_δ.

For each fixed $n \in \mathbb{N}$, note that

$$\mathcal{W}_n := \{\mathcal{L}_n(X) \colon X \in S\}$$

is finite, as $\mathcal{L}_n(X) \subseteq \mathcal{L}_n(\mathcal{A}^\mathbb{Z})$ for any $X \in S$. Enumerate the elements of $\mathcal{W}_n$ as $L_1^n, L_2^n, \dots, L_{|\mathcal{W}_n|}^n$. For each $1 \leq i \leq |\mathcal{W}_n|$, let $\mathcal{X}(i,n)$ denote the subshift of finite type whose set of forbidden words is $\mathcal{L}_n(\mathcal{A}^\mathbb{Z}) \setminus L_i^n$. For each k, let

$$\mathcal{B}(i,n,k) := \{Y \in S \colon d(\mathcal{X}(i,n), Y) < 2^{-n-k+1}\}.$$

In other words, $Y \in \mathcal{B}(i,n,k)$ if and only if $\mathcal{L}_{n+k-1}(Y) = \mathcal{L}_{n+k-1}(\mathcal{X}(i,n))$. For any such Y, it follows that $\mathcal{L}_j(Y) = \mathcal{L}_j(\mathcal{X}(i,n))$ for any $1 \leq j \leq n+k-1$. Fix such a Y and let $\{Y_m\}_{m=1}^\infty$ be the SFT cover of Y. Then, since $\mathcal{X}(i,n)$ is a subshift of finite type whose forbidden words all have length at most n, we have

$$Y_n = Y_{n+1} = Y_{n+2} = \cdots = Y_{n+k-1}.$$

Since $\mathcal{B}(i,n,k)$ is open, the set

$$\mathcal{U}_k := \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{|\mathcal{W}_n|} \mathcal{B}(i,n,k)$$

is open. Therefore the set

$$\mathrm{LS} := \bigcap_{k=1}^{\infty} \mathcal{U}_k$$

is G_δ.

We claim LS is precisely the set of language stable shifts. A subshift $Y \in \mathrm{LS}$ if and only if for all $k \geq 1$ there exists n and $1 \leq i \leq |\mathcal{W}_n|$ such that $Y \in \mathcal{B}(i,n,k)$. We have already seen that the statement $Y \in \mathcal{B}(i,n,k)$ implies that

$$Y_n = Y_{n+1} = \cdots = Y_{n+k-1},$$

where $\{Y_n\}_{n=1}^\infty$ is the SFT cover of Y. So if $Y \in \mathrm{LS}$ then for all k there exists some n_k such that

$$Y_{n_k} = Y_{n_k+1} = \cdots = Y_{n_k+k-1},$$

and this is equivalent to being language stable. Thus LS is contained in the set of language stable shifts. Conversely, for any language stable shift Y and any $k \geq 1$, there exists n_k such that

$$Y_{n_k} = Y_{n_k+1} = \cdots = Y_{n_k+k-1}.$$

Let $i_k \in \{1, \dots, |\mathcal{W}_{n_k}|\}$ be the index for which $\mathcal{L}_{n_k}(Y) = L_i^{n_k}$. Then $Y \in \mathcal{B}(i_k, n_k, k) \subseteq \mathcal{U}_k$. Since this holds for any k, $Y \in \bigcap_k \mathcal{U}_k$, meaning $Y \in \mathrm{LS}$. This proves the claim that the the G_δ set LS is equal to the set of language stable shifts. □

Corollary 5.2. *The set of language stable subshifts in S^+ is a G_δ subset of S^+. More generally, for any $c \geq 0$, the set of language stable subshifts in S_c^+ is a G_δ subset of S_c^+.*

Proof. By definition of the induced topology, the intersection of S^+ (resp. $S^+_{\geq c}$) with any G_δ subset of S is a G_δ subset of S^+ (resp. $S^+_{\geq c}$), so both the statements follow from Theorem 5.1. □

Theorem 5.3. *For any $c \geq 0$, the set of language stable subshifts in $S^+_{\geq c}$ is dense in $S^+_{\geq c}$. Analogously, the set of language stable subshifts in S is dense in S and the set of language stable subshifts in S^+ is dense in S^+.*

Proof. Let $\mathcal{U}$ be an arbitrary nonempty, open subset of S^+ and fix some $X \in \mathcal{U}$. Let $\{X_n\}_{n=1}^{\infty}$ be the SFT cover of X. Then, by definition of the metric, there exists N such that

$$\{Y \in S^+ \colon \mathcal{L}_N(Y) = \mathcal{L}_N(X)\} \subseteq \mathcal{U}.$$

For such an N notice that X_N is language stable, that $\mathcal{L}_N(X) = \mathcal{L}_N(X_N)$, and that

$$h_{\text{top}}(X_N) \geq h_{\text{top}}(X) \geq c, \tag{4}$$

since $X \subseteq X_N$. Therefore $X_N \in \mathcal{U}$. Since $\mathcal{U}$ was arbitrary, the set of language stable subshifts in $S^+_{\geq c}$ is dense in $S^+_{\geq c}$ for any $c \geq 0$.

For the analogous results for S and S^+, the only small modification is that in inequality (4) we only have $h_{\text{top}}(X) \geq 0$ in S and $h_{\text{top}}(X) > 0$ in S^+. □

Combining the results of this section, we get:

Corollary 5.4. *The set of language stable subshifts of S is generic in S, the set of language stable subshifts of S^+ is generic in S^+, and for any $c \geq 0$ the set of language stable subshifts of $S^+_{\geq c}$ is generic in $S^+_{\geq c}$.*

6. The example

6.1. **Language stability gives rise to new shifts with characteristic measures.** In this section, we construct a language stable subshift that carries a characteristic measure, of maximal entropy, that cannot be seen to exist for any of the four reasons given in the introduction. In particular we build a language stable subshift (W, σ) with the following properties:

(1) The automorphism group of the subshift W is nonamenable.
(2) The subshift W and all of its nonempty subsystems have positive topological entropy.
(3) Each characteristic measure supported on W is measurably isomorphic to infinitely many other measures supported on W.
(4) Each closed subshift $W' \subseteq W$ either does not support an $\text{Aut}(W)$-characteristic measure or has strictly lower topological entropy than W (meaning it does not support an $\text{Aut}(W)$-characteristic measure that is a measure of maximal entropy). The problem of determining whether W' has an $\text{Aut}(W')$-characteristic measure, to which Lemma 3.3 could be applied, does not appear to be significantly easier than determining whether W has an $\text{Aut}(W)$-characteristic measure. However, even if such a W' did support an $\text{Aut}(W')$-characteristic measure, the measure produced by Lemma 3.3 would still not be a measure of maximal entropy on W and so would not be the measure guaranteed to exist by Corollary 4.2.

Our method is to build the shift W as a product $W = X\times Y\times Z$ of three subshifts. The shift X, described in Section 6.2, is a language stable, positive entropy, minimal subshift. The shift Y is the full shift on 2 symbols and its properties are described in Section 6.3. The shift Z, described in Section 6.4, has countably many ergodic measures, all of which are isomorphic to each other, and in Section 6.5 we check that the product system W is language stable.

Before giving our example, we reiterate that W is language stable and so has a characteristic measure by Corollary 4.2. One might wonder if there is an easier way to see this and so we indicate why what may seem like a natural approach to finding a characteristic measure on W is not viable without significant extra information. It is tempting to try to find a characteristic measure for W by analyzing X, Y, and Z separately, finding an $\mathrm{Aut}(X)$-characteric measure α, an $\mathrm{Aut}(Y)$-characteristic measure β, and an $\mathrm{Aut}(Z)$-characteristic measure γ, and guessing that $\alpha\times\beta\times\gamma$ may be an $\mathrm{Aut}(W)$-characteristic measure. However, a recent result of Salo and Schraudner [25] shows that the automorphism group of a product of systems may be much larger than the product of their individual automorphism groups. Specifically, they show that if $X\subseteq\{0,1\}^{\mathbb{Z}}$ is the *sunny side up shift*, consisting of all configurations with at most one 1, then $\mathrm{Aut}(X)\cong\mathbb{Z}$ (in fact it is generated by the shift). On the other hand, they showed that $\mathrm{Aut}(X\times X)\cong(\mathbb{Z}^\infty\rtimes S_\infty)\rtimes(\mathbb{Z}^\infty\rtimes S_\infty)$. Thus, returning to our example, although $\alpha\times\beta\times\gamma$ is certainly invariant under the subgroup $\mathrm{Aut}(X)\times\mathrm{Aut}(Y)\times\mathrm{Aut}(Z)\subseteq\mathrm{Aut}(X\times Y\times Z)$, there is no reason to believe it is $\mathrm{Aut}(W)$-invariant without fully describing the algebraic structure of $\mathrm{Aut}(W)$, which may be significantly larger.

6.2. **There exists a language stable, positive entropy, minimal subshift.**

Lemma 6.1. *Let S be a topologically transitive subshift of finite type whose minimal forbidden words all have length at most N and let $\varepsilon>0$. For any $m\in\mathbb{N}$, there exists $d\in\mathbb{N}$ such that*

(5) $S':=\{x\in S\colon$ *every subword of x of length d contains every word in* $\mathcal{L}_m(S)\}$

is a topologically transitive subshift of finite type and $h_{\mathrm{top}}(S')\geq h_{\mathrm{top}}(S)-\varepsilon$.

Proof. If S contains only a single periodic point, then the lemma is trivially true. Thus without loss, we can assume that S contains at least two distinct periodic points.

For a fixed $m\in\mathbb{N}$, S' is the subshift of finite type obtained by forbidding all the minimal forbidden words in S, as well as all words in $\mathcal{L}(S)$ of length d that omit at least one element of $\mathcal{L}_m(S)$. Since S is a subshift of finite type, by Proposition 2.3 there exists $g\in\mathbb{N}$ such that if $u,v,w\in\mathcal{L}(S)$ are such that $uv,vw\in\mathcal{L}(S)$ and $|v|\geq g$, then $uvw\in\mathcal{L}(S)$. For the remainder of the proof, we assume that $m\geq g$.

If $d\geq m$ is fixed and the shift S' is defined by (5), then this shift is nonempty for sufficiently large d. Fixing such d, for any $a,b\in\mathcal{L}(S')$ we can find words $a',b'\in\mathcal{L}(S')$ such that: $|a'|\geq|a|+d$, $|b'|\geq|b|+d$, a is the leftmost subword of a', b is the rightmost subword of b', and the rightmost subword of length m in a' coincides with the leftmost subword of length m in b'. Note that we can do this because all words in $\mathcal{L}(S')$ can be legally extended arbitrarily far to the right and left (in at least one way) and all words of length d in S' contain a copy of every word in $\mathcal{L}_m(S)$. Therefore, setting b'' to be the word obtained by omitting the leftmost $m\leq d$ letters of b', we have that $a'b''\in\mathcal{L}(S')$ and $a'b''=acb$ for some

$c \in \mathcal{L}(S')$. Thus for any two words $a, b \in \mathcal{L}(S')$, there exists $c \in \mathcal{L}(S')$ such that $acb \in \mathcal{L}(S')$ and so S' is topologically transitive.

We are left with showing that given $\varepsilon > 0$, we can choose $d \in \mathbb{N}$ sufficiently large such that $h_{\mathrm{top}}(S') \geq h_{\mathrm{top}}(S) - \varepsilon$. We carry this out in two stages.

Step 1: we build a high entropy shift missing only one word. Let y be a periodic point in S of minimal period p and choose $w \in \mathcal{L}_p(X)$ such that $y = \dots wwww \dots$. Let μ be the Parry measure on S and μ is an ergodic and nonatomic measure satisfying $h_\mu(\sigma) = h_{\mathrm{top}}(S)$ (see [23]). By the Shannon-McMillan-Breiman Theorem, for μ-almost every $z \in S$ we have

$$\lim_{n\to\infty} -\frac{1}{2n+1} \log \mu([z_{-n} \dots z_0 \dots z_n]) = h_\mu(\sigma) = h_{\mathrm{top}}(S).$$

Thus for any $\delta > 0$, there exists $N \in \mathbb{N}$ and a set $Q \subseteq S$ such that $\mu(Q) > 1 - \delta/2$ and

$$|\mu([z_{-n} \dots z_0 \dots z_n]) - h_\mu(\sigma)| < \delta$$

for all $z \in Q$ and all $n \geq N$. For any fixed $k \geq 1$, let $w(k) := \underbrace{ww\dots w}_{k \text{ times}}$. By the Pointwise Ergodic Theorem we have

$$\lim_{n\to\infty} \frac{1}{2n+1} \sum_{i=-n}^{n} 1_{[w(k)]}(\sigma^i z) = \mu\left([w(k)]\right)$$

for μ-almost every $z \in S$. Therefore there exists $M \in \mathbb{N}$ and a set $R \subseteq S$ such that $\mu(R) > 1 - \delta/2$

$$\left| \frac{1}{2n+1} \sum_{i=-n}^{n} 1_{[w(k)]}(\sigma^i z) - \mu\left([w(k)]\right) \right| < \delta$$

for all $z \in R$ and all $n \geq M$. Taking the maximum of N and M, without loss we can assume that $N = M$. Then $\mu(Q \cap R) > 1 - \delta$ and for any $z \in Q \cap R$, we have that for all $n \geq N$,

$$\begin{aligned} \exp\big(-(2n+1)\cdot(h_{\mathrm{top}}(S)+\delta)\big) &< \mu([z_{-n} \dots z_0 \dots z_n]) \\ &< \exp\big(-(2n+1)\cdot(h_{\mathrm{top}}(S)-\delta)\big) \end{aligned} \tag{6}$$

and

$$\mu([w(k)]) - \delta < \frac{1}{2n+1} \sum_{i=-n}^{n} 1_{[w(k)]}(\sigma^i z) < \mu([w(k)]) + \delta.$$

Let $\mathcal{W}(n,k) \subseteq \mathcal{L}_{2n+1}(S)$ denote the set of words of the form $z_{-n} \dots z_0 \dots z_n$ for some $z \in Q \cap R$ and fix some $n \geq N$. Since $\mu(Q \cap R) > 1 - \delta$, using the upper bound for $\mu([z_{-n} \dots z_0 \dots z_n])$ given in (6), for each $z \in Q \cap R$ there are at least

$$\frac{1-\delta}{\exp\big(-(2n+1)\cdot(h_{\mathrm{top}}(S)-\delta)\big)}$$

cylinder sets based on words of length $2n+1$ in $\mathcal{L}(S)$ for which the number of occurrences of $w(k)$ as a subword of them lies between $(2n+1)(\mu([w(k)]) - \delta)$ and $(2n+1)(\mu([w(k)]) + \delta)$.

Recall that the word $w(k)$ is a periodic self-concatenation of the word w. Since S is a topologically transitive shift of finite type that contains at least two distinct periodic points, we can find a periodic word q whose length is a multiple of $|w|$ that is not just a self-concatenation of w but still starts with the first letter of w. Define the modified word $\tilde{w}(k)$ to be the word $w(k)$ with the leftmost $|q|$ symbols changed from $ww\dots w$ to q. Note that in any element of S and any occurrence of $w(k)$ in that element, the modified sequence that replaces this occurrence of $w(k)$ with $\tilde{w}(k)$ also determines an element of S. Therefore, for any fixed k and any element of S we can produce an element of S that does not contain $w(k)$ as a subword, by finding each occurrence of $w(k)$ and modifying it to $\tilde{w}(k)$. Let $Y_k \subseteq S$ be the subshift obtained by forbidding the word $w(k)$. We claim that $h_{\rm top}(Y_k)$ can be made arbitrarily close to $h_{\rm top}(S)$ by taking k sufficiently large. To see this note that for any $n \geq N$, the number of words in $\mathcal{L}_{2n+1}(Y_k)$ is at least

$$\frac{\left(\frac{1-\delta}{\exp(-(2n+1)\cdot(h_{\rm top}(S)-\delta))}\right)}{2^{(2n+1)\mu([\tilde{w}(k)]}}$$

since the map taking a word in $\mathcal{L}(S)$ to a word in $\mathcal{L}(Y_k)$ is at most $2^{(2n+1)\mu([\tilde{w}(k)])}$-to-one on the words arising from elements of $Q \cap R$. By taking k sufficiently large we can guarantee that $\mu([w(k-|p|/|w|)])$ is as close as we choose to zero, hence $\mu([\tilde{w}(k)])$ is also (since it contains $w(k-|p|/|w|)$ as its rightmost subword). Therefore the exponential growth rate of $|\mathcal{L}_{2n+1}(Y_k)|$ is at least $h_{\rm top}(S)-2\delta$ for any sufficiently large k.

Step 2: we use Y_k to define S'. Choose a word $w \in \mathcal{L}(S)$ that contains every element of $\mathcal{L}_m(S)$ as a subword. For each $k \geq 1$, we use this word w in the construction of Y_k. Fixing $\varepsilon > 0$, we can choose k sufficiently large such that $h_{\rm top}(Y_k) > h_{\rm top}(S) - \varepsilon/4$. We pick N such that for all $n \geq N$, we have $\log P_X(n) > n(h_{\rm top}(S) - \varepsilon/2)$. Since S is a topologically transitive shift of finite type, by Proposition 2.3 there exists a bound B such that for any $v \in \mathcal{L}(S)$ there are words $g_{w,v}, h_{w,v} \in \mathcal{L}(S)$ of length $|g_{w,v}|, |h_{w,v}| \leq B$ and such that $wg_{w,v}vh_{w,v}w \in \mathcal{L}(S)$. Choose one such $g_{w,v}$ and $h_{w,v}$ for each $v \in \mathcal{L}(X)$. Then there exist $b_1, b_2 \leq B$ such that

$$\lim_{n\to\infty} \frac{1}{n} \left|\{v \in \mathcal{L}_n(X) \colon |g_{w,v}| = b_1 \text{ and } |h_{w,v}| = b_2\}\right| > h_{\rm top}(S) - \varepsilon/4.$$

Define

$$\mathcal{W}_n := \{wg_{w,v}vh_{w,v} \in \mathcal{L}_n(Y_k) \colon |g_{w,v}| = b_1 \text{ and } |h_{w,v}| = b_2\}.$$

Without loss of generality, increasing N if necessary, we can assume that $|\mathcal{W}_n| > n(h_{\rm top}(S) - \varepsilon/2)$ for all $n \geq N$. By Proposition 2.3, the elements of $\mathcal{W}_n$ can be freely concatenated.

For each fixed n, note that the shift S' defined using $d := b_1 + b_2 + n$ contains all elements of S that can be written as concatenations of elements of $\mathcal{W}_n$. Define the shift Z_n to be the shift by only allowing elements of S that can be written as bi-infinite concatenations of elements of $\mathcal{W}_n$. Then the topological entropy of S' is at least as large as the shift Z_n and we now estimate the topological entropy of Z_n.

For $k, n \geq N$, the number of words in $\mathcal{L}_{kn}(Z_n)$ is at least $|\mathcal{W}_n|^k > (h_{\rm top}(S) - \varepsilon/2)^{kn}$; namely, all words in $\mathcal{W}_n$ have the same length and therefore, all ways of concatenating k elements of $\mathcal{W}_n$ result in distinct words of length $k(n+b_1+b_2)$ in $\mathcal{L}(S)$. For any $\delta > 0$, we can take n sufficiently large such that $k(n+b_1+b_2) < kn(1+\delta)$ and so $P_{Z_n}(n+b_1+b_2) > (h_{\rm top}(S)-\varepsilon/2)^{k(n+b_1+b_2)/(1+\delta)}$. For δ sufficiently small, this is larger than $(h_{\rm top}(S)-\varepsilon)^{k(n+b_1+b_2)}$. For a fixed such n, this holds for all k and so

$$\liminf_{k\to\infty} \frac{1}{n} \log P_{Z_n}(kn) > h_{\rm top}(S) - \varepsilon.$$

Therefore $h_{\text{top}}(S') \geq h_{\text{top}}(Z_n) > h_{\text{top}}(S) - \varepsilon$. □

Lemma 6.2. *There exists a language stable, positive entropy, minimal subshift. Moreover, given a sequence of intervals of consecutive integers:*

$$\{i_1, i_1+1, \ldots, i_1+j_1\}, \{i_2, i_2+1, \ldots, i_2+j_2\}, \ldots, \{i_n, i_n+1, \ldots, i_n+j_n\}, \ldots$$

there exists a language stable, positive entropy, minimal subshift X which has the property that $X_{i_n} = X_{i_n+j_n}$ for infinitely many n (where X_k is the k^{th} term in the SFT cover of X).

Proof. We inductively construct a sequence of subshifts

$$X_1 \supseteq X_2 \supseteq X_3 \supseteq \cdots \supseteq X_n \supseteq X_{n+1} \supseteq \cdots$$

and show that $X := \bigcap_{i=1}^{\infty} X_i$ is minimal, language stable, and has positive topological entropy.

Fix a sequence $\{\varepsilon_i\}_{i=1}^{\infty}$ of elements of $(0,1)$ such that $\sum_i \varepsilon_i < \log 2$. Let $X_1 := \{0,1\}^{\mathbb{Z}}$ and so $h_{\text{top}}(X_1) = \log 2$. Fix the parameter $m_1 := 1$. Taking $m := m_1$ in Lemma 6.1, choose $d_1 > m_1$ such that

$$X_2 := \{x \in X_1 \colon \text{every subword of length } d_1 \text{ contains every word in } \mathcal{L}_{m_1}(X_1)\}$$

is a topologically transitive subshift of finite type satisfying $h_{\text{top}}(X_2) \geq h_{\text{top}}(X_1) - \varepsilon_1$. Suppose we have constructed a nested sequence of topologically transitive subshifts of finite type $X_1 \supseteq X_2 \supseteq X_3 \supseteq \cdots \supseteq X_n$ as well as parameters m_i, d_i for all $i = 1, \ldots, n-1$ satisfying

$$m_{i+1} \geq \max\{d_i + i, i_k + j_k\}$$

where $i_k = \min\{i_\ell \colon i_\ell > d_i\}$ and $1 \leq i < n-1$. Suppose further that for all $x \in X_n$ and all $1 \leq i < n$, every subword of x of length d_i contains every element of $\mathcal{L}_{m_i}(X_i)$ as a subword. Set $k' = \min\{i_\ell \colon i_\ell > d_{n-1}\}$, define $m_n := \max\{d_{n-1} + n, i_{k'} + j_{k'}\}$, and by Lemma 6.1, find $d_n > m_n$ such that

$$X_{n+1} := \{x \in X_n \colon \text{every subword of length } d_n \text{ contains every word in } \mathcal{L}_{m_n}(X_n)\}$$

is a topologically transitive subshift of finite type satisfying $h_{\text{top}}(X_{n+1}) \geq h_{\text{top}}(X_n) - \varepsilon_n$. Inductively, this defines the shift X_i for all $i \geq 1$ and we set $X := \bigcap_i X_i$. We claim that X is minimal, language stable, and has positive entropy.

First we show X is minimal. Let $w \in \mathcal{L}(X)$ and pick i such that $m_i > |w|$. Since $\mathcal{L}_{m_i}(X) \subseteq \mathcal{L}_{m_i}(X_i)$, it follows by construction that for any $x \in X_{i+1}$ the word w occurs in every subword of x of length d_i. But $X \subseteq X_{i+1}$ and so w occurs syndetically, with gap at most d_i, in every element of X. Since this holds for any $w \in \mathcal{L}(X)$, X is minimal.

Next we show that X is language stable. Fix $i \in \mathbb{N}$ and recall that $m_{i+1} \geq \max\{d_i + i, i_k + j_k\}$ where $i_k = \min\{i_\ell \colon i_\ell > d_i\}$. The shift X_{i+1} is a topologically transitive subshift of finite type whose minimal forbidden words all have length at most d_i. By construction, every word in $\mathcal{L}_{m_{i+1}}(X_{i+1})$ occurs in every element of X_{i+2} and hence in every element of X. Therefore $\mathcal{L}_{m_{i+1}}(X) = \mathcal{L}_{m_{i+1}}(X_{i+1})$. It follows that $\mathcal{L}_k(X) = \mathcal{L}_k(X_{i+1})$ for all $1 \leq k \leq m_{i+1}$. Since there are no minimal forbidden words in X_{i+1} of length greater than d_i and since $m_{i+1} \geq d_i + i$, it follows that there are no minimal forbidden words in X of lengths $d_i + 1, \ldots, d_i + i$, and moreover there are no forbidden words of any length in the interval $\{i_k, i_k + 1, \ldots, i_k + j_k\}$. Since this holds for any i, it follows that the set

$$\{k \colon X \text{ has no minimal forbidden words of length } k\}$$

has upper Banach density 1. In other words, X is language stable. Furthermore, there are infinitely many n for which $X_{i_n} = X_{i_n+j_n}$.

Finally we show that $h_{\mathrm{top}}(X) > 0$. Since X_i is a topologically transitive subshift of finite type, it follows from Parry [23] that X_i supports a unique measure of maximal entropy μ_i. Passing to a subsequence if necessary, we assume that the sequence $\{\mu_i\}_{i=1}^{\infty}$ converges to a weak* limit μ. Note that μ is supported on $X = \bigcap_i X_i$ and by upper semi-continuity of the entropy map (see [27, Theorem 8.2]) for subshifts, we have that

$$h_\mu(\sigma) \geq \limsup_{i\to\infty} h_{\mu_i}(\sigma) \geq h_{\mathrm{top}}(X_1) - \sum_{i=1}^{\infty} \varepsilon_i = \log(2) - \sum_{i=1}^{\infty} \varepsilon_i > 0.$$

By the Variational Principle (see for example [27, Theorem 8.6]) we have $h_{\mathrm{top}}(X) \geq h_\mu(\sigma) > 0$. □

6.3. **The characteristic measures on a full shift.** Boyle, Lind, and Rudolph [5, Corollary 10.2] show that for any topologically mixing subshift of finite type, the measure of maximal entropy is the unique characteristic measure of positive entropy, and all characteristic measures of entropy zero are countable convex combinations of purely atomic measures supported on unions of periodic orbits. In particular, for the full shift on two symbols, their result says the following:

Lemma 6.3. *For the full* 2*-shift* $(\{0,1\}^{\mathbb{Z}}, \sigma)$*, every characteristic measure is a convex combination of the Bernoulli measure (assigning measure* $(1/2)^n$ *to each cylinder set determined by a word of length* n*) and atomic measures supported on unions of periodic orbits.*

6.4. **There exists a language stable shift with countably many ergodic measures, all of whose ergodic measures are isomorphic to each other.**

Lemma 6.4. *Let* $\alpha, \beta \in (0,1)$ *and* $\alpha \notin \mathbb{Q}$*. Let* $R_\alpha \colon [0,1) \to [0,1)$ *be the map* $R_\alpha(x) := x + \alpha \pmod 1$*. Let* $\mathcal{P} := \{[0,\beta), [\beta,1)\}$ *and let* X_β *be the coding of the system* $([0,1), R_\alpha)$ *by the partition* $\mathcal{P}$*. For* $n \geq 2$*, let*

$$\mathcal{P}_n := \bigvee_{i=1}^{n-1} R_\alpha^{-i}\mathcal{P}.$$

Finally let $\mathcal{S}_0 = \{0, \beta\}$ *and* $\mathcal{S}_n := R_\alpha^{-n}\mathcal{S}_0 = \{-n\alpha, \beta - n\alpha\}$*. If* X_β *has a minimal forbidden word of length* $n+1$*, then at least one element of* $\mathcal{R}_0$ *lies in the same cell of* $\mathcal{P}_n$ *as an element of* $\mathcal{S}_{n+1}$*.*

Proof. Suppose X_β has a minimal forbidden word w of length $n+1$. Writing $w := (a_0, a_1, \ldots, a_{n-1}, a_n) \in \mathcal{A}^{n+1}$, then since w is minimal, we have that

$$(a_0, a_1, \ldots, a_{n-1}), (a_1, a_2, \ldots, a_n) \in \mathcal{L}_n(X_\beta).$$

Let $u := (a_1, a_2, \ldots, a_{n-1}) \in \mathcal{L}_{n-1}(X_\beta)$ denote the "interior" of w.

Say there is a unique $b \in \mathcal{A}$ such that $ub \in \mathcal{L}_n(X_\beta)$. Then since $(a_1, a_2, \ldots, a_n) \in \mathcal{L}_n(X_\beta)$, it follows that $b = a_n$. But then since $a_0u = (a_0, a_1, \ldots, a_{n-1}) \in \mathcal{L}_n(X_\beta)$, it follows that $w = a_0ua_n \in \mathcal{L}_{n+1}(X_\beta)$, as

$$\{x \in X_\beta \colon x_i = a_i \text{ for all } 0 \leq i < n\} \neq \emptyset,$$

and so the only possibility is that $x_n = a_n$ because $(x_1, \ldots, x_{n-1}) = u$. But this contradicts the assumption that w is a forbidden word, and so we conclude that there is no unique $b \in \mathcal{A}$ such that $ub \in \mathcal{L}_n(X_\beta)$. Similarly there cannot be a unique $c \in \mathcal{A}$ such that $cu \in \mathcal{L}_n(X_\beta)$.

Recall that when $([0,1), R_\alpha)$ is coded by the partition $\mathcal{P}$, the cells of $\mathcal{P}_n$ are in one-to-one correspondence with cylinder sets of the form:

$$\sigma^{-1}[v]_0^+ := \{x \in X_\beta \colon x_i = v_i \text{ for all } 1 \leq i \leq n-1\}$$

where $v \in \mathcal{L}_{n-1}(X_\beta)$. We maintain the same notation for w being a minimal forbidden word of length $n+1$ and u denoting its interior. Since there is no unique $b \in \mathcal{A}$ such that $ub \in \mathcal{L}(X_\beta)$, this means that the cell of $\mathcal{P}_n$ that corresponds to u is subdivided into at least two different cells in the refined partition $\bigvee_{i=1}^{n} R_\alpha^{-i}\mathcal{P} = \mathcal{P}_n \vee R_\alpha^{-n}\mathcal{P}$. The only two cells that are subdivided in this way are the cells containing the elements of $\mathcal{R}_n$. Similarly, since there is no unique $c \in \mathcal{A}$ such that $cu \in \mathcal{L}(X_\beta)$, it follows that the cell of $\mathcal{P}_n$ corresponding to u is subdivided into at least two different cells in the refined partition $\bigvee_{i=0}^{n-1} R_\alpha^{-i}\mathcal{P} = \mathcal{P} \vee \mathcal{P}_n$. The only two cells that are subdivided in this way are the cells containing the elements of $\mathcal{R}_0$. Therefore, the cell of $\mathcal{P}_n$ corresponding to u contains at least one element of $\mathcal{R}_0$ and at least one element of $\mathcal{R}_n$. □

In preparation for our next lemma, we define a second partition $\mathcal{Q} := \{[0,\alpha), [\alpha, 1)\}$ and for $n > 2$, define $\mathcal{Q}_n := \bigvee_{i=2}^{n-1} R_\alpha^{-i}\mathcal{Q}$. We make use of an auxiliary result.

Lemma 6.5. *For fixed $n > 2$, every cell of $\mathcal{Q}_n$ can be written as a union of cells from $\mathcal{P}_n$.*

Proof. Observe that $\mathcal{Q}_n$ is the partition of $[0,1)$ into intervals whose endpoints come from the set

$$\{-\alpha, -2\alpha, -3\alpha, \dots, -(n-1)\alpha\}.$$

Notice that $\mathcal{P}_n$ is the partition of $[0,1)$ into intervals whose endpoints come from the set

$$\{-\alpha, -2\alpha, 3-\alpha, \dots, -(n-1)\alpha\} \cup \{\beta-\alpha, \beta-2\alpha, \beta-3\alpha, \dots, \beta-(n-1)\alpha\}.$$

Therefore each interval in $\mathcal{Q}_n$ can be written as a union of intervals from $\mathcal{P}_n$. □

The intervals that comprise $\mathcal{Q}_n$ have a natural adjacency relation in $\mathbb{R}/\mathbb{Z}$: we say two cells of $\mathcal{Q}_n$ are *adjacent* if they share an endpoint in $\mathbb{R}/\mathbb{Z}$ and are *twice adjacent* if there is a third cell that is adjacent to both of them.

Lemma 6.6. *Maintaining the notation of Lemma 6.4, if X_β has a minimal forbidden word of length $n+1$ then at least one of the following holds:*

(1) 0 *and* $-n\alpha$ *lie in the same, adjacent, or twice adjacent cells of* $\mathcal{Q}_n$*;*
(2) β *and* $-n\alpha$ *lie in the same cell of* $\mathcal{Q}_n$*;*
(3) $-\beta$ *and* $-n\alpha$ *lie in the same, adjacent, or twice adjacent cells of* $\mathcal{Q}_n$*.*

Proof. By Lemma 6.4, at least one element of $\{0,\beta\}$ must lie in the same cell of $\mathcal{P}_n$ as an element of $\{-n\alpha, \beta-n\alpha\}$. If $-n\alpha$ lies in the same cell of $\mathcal{P}_n$ as an element of $\{0,\beta\}$, then one of (1) or (2) occurs since the cells of $\mathcal{Q}_n$ can be written as unions of cells in $\mathcal{P}_n$. Otherwise, one of the following holds:

(1) β and $\beta - n\alpha$ lie in the same cell of $\mathcal{P}_n$ (hence also the same cell of $\mathcal{Q}_n$);
(2) 0 and $\beta - n\alpha$ lie in the same cell of $\mathcal{P}_n$ (hence also the same cell of $\mathcal{Q}_n$).

Let d denote the metric on $[0, 1)$ inherited from the Euclidean metric on $\mathbb{R}/\mathbb{Z}$. Note that $d(\beta, \beta - n\alpha) = d(0, -n\alpha)$, and so if β and $\beta - n\alpha$ lie in the same cell of $\mathcal{P}_n$ then $d(0, -n\alpha)$ is at most the length, L, of that cell. By the Three Lengths Theorem [26], for any $n > 2$ the intervals comprising $\mathcal{Q}_n$ have at least two and at most three distinct lengths. Moreover, for any n where the intervals have three distinct lengths, the longest of the lengths is the sum of the shorter two and the sum of the lengths of any two consecutive cells is at least the longest length. Therefore, since $d(0, -n\alpha) = L$ is at most the longest length of any interval in $\mathcal{Q}_n$, 0 and $-n\alpha$ can be in the same cell of $\mathcal{Q}_n$, adjacent cells of $\mathcal{Q}_n$, or twice adjacent cells of $\mathcal{Q}_n$. In particular, if β and $\beta - n\alpha$ lie in the same cell of $\mathcal{P}_n$ then (1) holds. Finally, if 0 and $\beta - n\alpha$ lie in the same cell of $\mathcal{P}_n$, then $d(-\beta, -n\alpha) = d(0, \beta - n\alpha)$ and similarly (3) holds. □

The interest in Lemma 6.6 is that we have removed the dependence on $\mathcal{P}_n$ (a partition which depends on β) and replaced it with $\mathcal{Q}_n$ (a partition which depends only on α).

We recall some facts that follow from the Three Lengths Theorem of Sos [26]. First, if n is such that $\mathcal{Q}_n$ has only two distinct lengths, then a new length is created in $\mathcal{Q}_{n+1}$ by subdividing one of the intervals from $\mathcal{Q}_n$ with the longest length. Further, there is a simple formula for the number of intervals of each length in $\mathcal{Q}_n$:

Theorem 6.7 (See for example [2, Theorem 2.6.1])**.** *Let $\alpha \notin \mathbb{Q}$ and let $\alpha = [0, a_1, a_2, a_3, \dots]$ be its continued fraction expansion. Let $p_k/q_k = [0, a_1, a_2, \dots, a_k]$ be its k^{th} convergent. Then the sequence $\{q_k\}_{k=1}^{\infty}$ is nondecreasing, tends to infinity, and for every integer $n \geq 1$ there exists a unique k such that there are numbers $1 \leq m \leq a_{k+1}$ and $0 \leq r < q_k$ satisfying $n = mq_k + q_{k-1} + r$. The partition $\mathcal{Q}_n$ has:*

1. *$r + 1$ intervals of length $\eta_{k-1} - m\eta_k$ (Type 1);*
2. *$n + 1 - q_k$ intervals of length η_k (Type 2);*
3. *$q_k - r - 1$ intervals of length $\eta_{k-1} - (m - 1)\eta_k$ (Type 3),*

where $\eta_k := (-1)^k(q_k\alpha - p_k)$.

An immediate corollary is the following.

Corollary 6.8. *No interval in $\mathcal{Q}_{q_k+q_{k-1}}$ is divided into more than two subintervals in $\mathcal{Q}_{(a_{k+1}+1)q_k+q_{k-1}-1}$. In particular, the orbit segment*

$$\{-n\alpha \colon q_k + q_{k-1} \leq n < (a_{k+1} + 1)q_k + q_{k-1}\}$$

does not visit any cell in $\mathcal{Q}_{q_k+q_{k-1}}$ more than once.

We now construct our language stable shift. Choose $0 < \beta_1 < \alpha$. Sppose we have chosen numbers $\beta_1 < \beta_2 < \cdots < \beta_i < \alpha$ and integers $k_1 < k_2 < \cdots < k_i$ such that for each $1 \leq j \leq i$ we have

1. $(a_{k_j+1} + 1)q_{k_j} > 7j^2$;
2. the numbers $\beta_j, \beta_{j+1}, \beta_{j+2}, \dots, \beta_i$ all lie in the same cell of $\mathcal{Q}_{(a_{k_j+1}+1)q_{k_j}+q_{k_j-1}-1}$ as α;
3. the numbers $-\beta_j, -\beta_{j+1}, -\beta_{j+2}, \dots, -\beta_i$ also all lie in the same cell of $\mathcal{Q}_{(a_{k_j+1}+1)q_{k_j}+q_{k_j-1}-1}$ as α.

Then for any $q_k + q_{k-1} \leq n < (a_{k+1}+1)q_k + q_{k-1}$ the numbers $\beta_j, \beta_{j+1}, \ldots, \beta_i$ all lie in the same cell of $\mathcal{Q}_n$ as each other and, similarly, the numbers $-\beta_j, -\beta_{j+1}, \ldots, -\beta_i$ also all lie in the same cell of $\mathcal{Q}_n$ as each other. By Corollary 6.8 and Lemma 6.6, for any $j_1, j_2 \in \{1, 2, \ldots, i\}$ there are at most three values of n in the interval $q_{k_{j_1}} + q_{k_{j_1}-1} \leq n < (a_{k_{j_1}}+1)q_{k_{j_1}} + q_{k_{j_1}-1}$ for which $X_{\beta_{j_2}}$ has a minimal forbidden word of length $n+1$. But combining these results with conditions (2) and (3), we also have that for any $1 \leq j_1 \leq i$, the set of n in the interval $q_{k_{j_1}} + q_{k_{j_1}-1} \leq n < (a_{k_{j_1}}+1)q_{k_{j_1}} + q_{k_{j_1}-1}$ for which there exists $1 \leq j_2 \leq i$ such that $X_{\beta_{j_2}}$ has a minimal forbidden word of length $n+1$ is at most $7j_1$. By condition (1), the interval $q_{k_{j_1}} + q_{k_{j_1}-1} \leq n < (a_{k_{j_1}}+1)q_{k_{j_1}} + q_{k_{j_1}-1}$ has length at least $7j_1^2$ and so there must be a subinterval of length j_1 on which none of the shifts $X_{\beta_1}, X_{\beta_2}, \ldots, X_{\beta_i}$ have any minimal forbidden words. Since the sequence $\{q_k\}_{k=1}^{\infty}$ is nondecreasing and tends to infinity, we can choose k_{i+1} sufficiently large such that $(a_{k_{i+1}+1}+1)q_{k_{i+1}} > 7(i+1)^2$ and choose $\beta_{i+1} \in (\beta_i, \alpha)$ such that β_{i+1} lies in the same cell of $\mathcal{Q}_{(a_{k_{i+1}+1}+1)q_{k_{i+1}}+q_{k_{i+1}-1}-1}$ as α, and $-\beta_{i+1}$ lies in the same cell of $\mathcal{Q}_{(a_{k_{i+1}+1}+1)q_{k_{i+1}}+q_{k_{i+1}-1}-1}$ as $-\alpha$. Since the partitions $\mathcal{Q}_k$ refine each other as k increases, it follows that for any $1 \leq j_1 \leq i+1$ that β_{i+1} lies in the same cell of $\mathcal{Q}_{(a_{k_{j_1}+1}+1)q_{k_{j_1}}+q_{k_{j_1}-1}-1}$ as α, and $-\beta_{i+1}$ lies in the same cell of $\mathcal{Q}_{(a_{k_{j_1}+1}+1)q_{k_{j_1}}+q_{k_{j_1}-1}-1}$ as $-\alpha$. By induction, we construct a sequence

$$0 < \beta_1 < \beta_2 < \beta_3 < \cdots < \beta_i < \cdots < \alpha$$

which satisfies conditions (1), (2), and (3) for all i, j. Therefore for any j there is an interval of integers (between $q_{k_j} + q_{k_j-1}$ and $(a_{k_j+1}+1)q_{k_j} + q_{k_j-1}$) of length j such that X_{β_i} has no minimal forbidden words of any lengths in that interval, for any $i = 1, 2, 3, \ldots$ Consider the shift

$$Z := \overline{\bigcup_{i=1}^{\infty} X_{\beta_i}}. \tag{7}$$

A word $w \in \{0,1\}^*$ is in the language of Z if and only if there exists i such that $w \in \mathcal{L}(X_{\beta_i})$. It follows that Z is language stable. Moreover we claim that

$$Z \setminus \bigcup_{i=1}^{\infty} X_{\beta_i}$$

is the shift X_α. We show this in two steps. First, suppose $z \in Z$ but $z \notin X_{\beta_i}$ for any i. Then there is a sequence $\{m_i\}$ tending to infinity, and points $x_i \in X_{\beta_{m_i}}$ such that $\lim_i x_i = z$. In other words, for each fixed N we have $z_j = (x_i)_j$ for all $|j| \leq N$. Now let $s_i \in S^1$ be a point whose coding with respect to the partition $\{[0, \beta_{m_i}), [\beta_{m_i}, 1)\}$ agrees with $(x_i)_j$ for all $|j| \leq N$. Passing to a subsequence if necessary, we can assume there exists $y \in S^1$ such that $\lim s_i = y$. Note that $R_\alpha^j s_i$ only codes differently with respect to the partitions $\{[0, \beta_{m_i}), [\beta_{m_i}, 1)\}$ and $\{[0, \alpha), [\alpha, 1)\}$ if $R_\alpha^j s_i \in [\beta_{m_i}, \alpha)$. Since $\lim_i s_i = y$ and $\lim_i \beta_{m_i} = \alpha$, for fixed N we have, for all sufficiently large i, that $R_\alpha^j s_i$ codes in the same way with respect to both of these partitions for all $|j| \leq N$ unless $R_\alpha^j y = \alpha$ for some $|j| \leq N$. Assuming this holds, note that $R_\alpha^j s_i$ and $R_\alpha^j y$ code in the same way with respect to both $\{[0, \beta_{m_i}), [\beta_{m_i}, 1)\}$ and $\{[0, \alpha), [\alpha, 1)\}$ for all $|j| \leq N$ and all sufficiently large i. Therefore the coding of $R_\alpha^j y$ is z_j for all $|j| \leq N$. This holds for all N, so $z \in X_\alpha$. The other possibility is that $R_\alpha^j y = \alpha$ for some $|j| \leq N$. In this case, note that for fixed N, for any sufficiently small $|\gamma| > 0$ the points $R_\alpha^j(s_i + \gamma)$ and $R_\alpha^j s_i$ code in the same way with respect to $\{[0,\beta_{m_i}),[\beta_{m_i},1)\}$ for all $|j|\le N$ and all sufficiently large i (since $\lim s_i = y$) and so we can reduce to the previous case to see that $z\in X_\alpha$. It follows from minimality of X_α that either

$$Z\setminus\bigcup_{i=1}^{\infty}X_{\beta_i}=X_\alpha \quad \text{or} \quad Z\setminus\bigcup_{i=1}^{\infty}X_{\beta_i}=\varnothing.$$

To show it is the former fix $y\in S^1$ to be a point whose R_α-orbit does not include $\alpha\in S^1$. For each fixed N, note that $R_\alpha^j y$ codes in the same way with respect to $\{[0,\beta_i),[\beta_i,1)\}$ and $\{[0,\alpha),[\alpha,1)\}$ for all $|j|\le N$ and all sufficiently large i. The coding of the R_α-orbit of y is an element of X_{β_i} and therefore this is a sequence of points in $\bigcup_i X_{\beta_i}$ whose limit is the coding of y with respect to $\{[0,\alpha),[\alpha,1)\}$, meaning there is an element of X_α in Z.

Furthermore, note that in our construction, we could always choose that $\beta_i\in\{n\alpha\colon n=1,2,\dots\}$, and henceforth we insist on this. Then since $[0,\beta_i)$ can be written as union of intervals in $\bigvee_{i=0}^{m}R_\alpha^{-i}\{[0,\alpha),[\alpha,1)\}$ for sufficiently large m, there is a block code $\varphi_i\colon X_\alpha\to X_{\beta_i}$. By a result of Durand [11, Corollary 12], such a block code must be invertible and so X_α is topologically conjugate to X_{β_i} for all i. Since X_α is uniquely ergodic, it follows that Z is the union of countably many uniquely ergodic, topologically conjugate subshifts. In particular it has only countably many ergodic measures and they are all (measurably) isomorphic to each other.

6.5. **The example and its properties.** The example is the shift $X\times Y\times Z$.

Lemma 6.9. *The shift (W,σ), where $W=X\times Y\times Z$ and X is defined as in Lemma 6.2, Y is the full 2-shift, and Z is defined in (7), is language stable.*

Proof. For any $n\in\mathbb{N}$, we have that $\mathcal{L}_n(X\times Y\times Z)=\mathcal{L}_n(X)\times\mathcal{L}_n(Y)\times\mathcal{L}_n(Z)$. Therefore $(w_X,w_Y,w_Z)\in\{0,1\}^n\times\{0,1\}^n\times\{0,1\}^n$ is forbidden if and only if at least one of the component words is forbidden, meaning that at least one of the following holds: w_X is forbidden in X, w_Y is forbidden in Y, or w_Z is forbidden in Z. Similarly, the word (w_X,w_Y,w_Z) is minimal and forbidden if and only if at least one of w_X, w_Y, and w_Z is a minimal forbidden word in its respective shift. In particular, if each of X, Y, and Z have no minimal forbidden words of length n, then W also has no minimal forbidden words of length n. We claim that there are arbitrarily long intervals of the form $\{N,N+1,N+2,\dots,N+k-1\}$ for which $W=X\times Y\times Z$ has no minimal forbidden words of any length in the interval.

Since Y is the full 2-shift, it does not introduce any forbidden words. Let $\{Z_n\}_{n=1}^{\infty}$ be the SFT cover of the shift Z defined by (7) and let

$$\{i_1,i_1+1,\dots,i_1+j_1\},\{i_2,i_2+1,\dots,i_2+j_2\},\dots,\{i_n,i_n+1,\dots,i_n+j_n\},\dots$$

be a sequence of intervals of consecutive integers for which the sequence $(j_n)_{n\in\mathbb{N}}$ is strictly increasing and such that $Z_{i_n}=Z_{i_n+j_n}$ for all n (this is possible since Z is language stable). By Lemma 6.2 we can construct X such that for infinitely many n we also have $X_{i_n}=X_{i_n+j_n}$. Therefore there are infinitely many n such that none of X, Y, or Z has a minimal forbidden word of any length in the interval $\{i_n,i_n+1,\dots,i_n+j_n\}$, and so $X\times Y\times Z$ also has no minimal forbidden word in any such interval. Since the sequence $(j_n)_{n\in\mathbb{N}}$ is strictly increasing, $W:=X\times Y\times Z$ is language stable. □

Combining this with Corollary 4.2, we have that the existence of a characteristic measure on this system:

Corollary 6.10. *The system (W, σ) has a characteristic measure. Moreover it has a characteristic measure that is a measure of maximal entropy.*

We next check that the existence of this characteristic measure for the system (W, σ) does not follow from results already previously in the literature.

Lemma 6.11. *If μ is any characteristic measure on (W, σ), then the set*

$$\{\nu \in \mathcal{M}(X)\colon (X, \sigma, \nu) \text{ is measurably isomorphic to } (X, \sigma, \mu)\}$$

is infinite. In particular, Lemma 3.1 does not apply to $W = X \times Y \times Z$.

Proof. Let μ be a characteristic measure on (W, σ). Let μ_{YZ} be the marginal measure obtained by projecting μ onto $Y \times Z$, meaning that for any measurable $A \subseteq Y \times Z$ we have $\mu_{YZ}(A) := \mu(X \times A)$. Note that μ_{YZ} is a shift invariant probability measure on $Y \times Z$. Next set μ_Y to be the marginal of μ_{YZ} projected onto Y and set μ_Z to be the marginal of μ_{YZ} projected onto Z; thus for measurable $B \subseteq Y$ and $C \subseteq Z$, we have $\mu_Y(B) := \mu_{YZ}(B \times Z)$ and $\mu_Z(C) := \mu_{YZ}(Y \times C)$. Then μ_Y is an invariant measure on Y and μ_Z is an invariant measure on Z.

If $\varphi \in \mathrm{Aut}(Y)$ is an automorphism of Y and if $B \subseteq Y$ is a measurable set, then (letting Id denote the identity) we have

$$\begin{aligned}
\mu_Y(\varphi^{-1}B) &= \mu_{YZ}((\varphi \times \mathrm{Id})^{-1}(B \times Z)) \\
&= \mu((\mathrm{Id} \times \varphi \times \mathrm{Id})^{-1}(X \times B \times Z)) = \mu(X \times B \times Z) \\
&= \mu_{YZ}(B \times Z) = \mu_Y(B),
\end{aligned}$$

where the third equality holds because $\mathrm{Id} \times \varphi \times \mathrm{Id} \in \mathrm{Aut}(X \times Y \times Z)$ and μ is characteristic. Therefore, μ_Y is an $\mathrm{Aut}(Y)$-characteristic measure on the full shift Y. Thus by Lemma 6.3, we can decompose the measure μ_Y into a convex combination of the symmetric Bernoulli measure on Y and atomic measures supported on unions of periodic orbits in Y, writing

$$\mu_Y = c_0\mu_B + \sum_{i=1}^{\infty} c_i \mu_{p_i} \tag{8}$$

where $0 \le c_i \le 1$ for all i, $\sum_{i=0}^{\infty} c_i = 1$, μ_B is the symmetric Bernoulli measure on Y, p_i is a collection of pairwise disjoint unions of periodic orbits, and μ_{p_i} is a characteristic measure supported on p_i.

Furthermore, the measure μ_{YZ} is a joining of the measures μ_Y and μ_Z. But $h_{\mathrm{top}}(Z) = 0$ and $h(\mu_Z) = 0$. Therefore, the measure μ_B (in the decomposition (8) of μ_Y) is disjoint from μ_Z. Moreover, μ_Z is an invariant measure on Z and so it is an at most countable convex combination of ergodic measures that are all isomorphic to the same irrational circle rotation. It follows that μ_Z is disjoint from every finite rotation, and in particular from μ_{p_i} for all i. Combining these two observations, it follows that μ_Y and μ_Z are disjoint, and so we have that $\mu_{YZ} = \mu_Y \times \mu_Z$.

We write the decomposition of μ_Z as

$$\mu_Z = \sum_{i=1}^{\infty} d_i \mu_{Z,i},$$

where $\mu_{Z,i}$ is an enumeration of the countably many ergodic measures supported on Z with weights $0 \le d_i \le 1$ satisfying $\sum_{i=1}^{\infty} d_i = 1$. For each integer $k \ge 0$, the measure μ_Z^k defined by

$$\mu_Z^k := \sum_{i=1}^{\infty} d_{i+k}\mu_{Z,i}$$

is measurably isomorphic to μ_Z, and the resulting measures $\{\mu_Z^k\}_{k=0}^{\infty}$ are pairwise distinct. Therefore, for each $k \geq 0$, the measure $\mu_{YZ} = \mu_Y \times \mu_Z$ is measurably isomorphic to each of the pairwise distinct measures $\mu_Y \times \mu_Z^k$.

Finally, set μ_X to be the marginal obtained by projecting μ onto X, and so for measurable $A \subseteq X$ we have $\mu_X(A) := \mu(A \times Y \times Z)$. Then μ is a joining of μ_X with μ_{YZ}. Since μ_{YZ} is isomorphic to $\mu_Y \times \mu_Z^k$, there is a joining of μ_X with $\mu_Y \times \mu_Z^k$ that is isomorphic to μ (and distinct from it since it has a different marginal onto $Y \times Z$). These joinings are pairwise distinct for each $k \geq 0$, so μ is measurably isomorphic to infinitely many measures supported on (W, σ).

Note that if μ were the measure produced by Lemma 3.1 then μ would only be measurably isomorphic to finitely many other measures on (W, σ). Therefore μ does not result from applying Lemma 3.1 to any measure on (W, σ). Since μ is an arbitrary characteristic measure on (W, σ), Lemma 3.1 cannot be applied to this system. □

Finally we check that the automorphism group of (W, σ) is not amenable, meaning that we can not apply the Krylov-Bogolioubov Theorem to produce a characteristic measure for this system.

Lemma 6.12. *The automorphism group of* (W, σ) *is not amenable (as a countable discrete group).*

Proof. The automorphism group of Y is nonamenable, since Y is a full-shift on at least two symbols (see [5]). For each $\varphi \in \mathrm{Aut}(Y)$, the map $\varphi \mapsto \mathrm{Id} \times \varphi \times \mathrm{Id}$ gives an embedding of $\mathrm{Aut}(Y)$ into $\mathrm{Aut}(W)$. Since any subgroup of an amenable, countable discrete group is also amenable, it follows that $\mathrm{Aut}(W)$ is also nonamenable. □

7. An application

7.1. **When does a block code define an automorphism?** Suppose $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is a subshift and $\varphi\colon \mathcal{L}_{2R+1}(X) \to \mathcal{A}$ is a block code. A natural question is whether φ defines an automorphism of X, and answering this question in general is challenging. We note, however, that if μ is a characteristic measure for (X, σ), then μ gives rise to a family of necessary conditions that must be satisfied if φ is an automorphism. Making this precise, we have:

Lemma 7.1. *Let* $X \subseteq \mathcal{A}^{\mathbb{Z}}$ *be a subshift and let* $\varphi\colon \mathcal{L}_{2R+1}(X) \to \mathcal{A}$ *be a range* R *block code with* $\varphi(X) \subseteq X$*. For each* $w \in \mathcal{L}(X)$ *define*

$$\varphi^{-1}(w) := \{v \in \mathcal{L}_{2R+|w|}(X)\colon \varphi(v) = w\}.$$

Suppose μ *is a characteristic measure for* (X, σ)*. If there exists* $w \in \mathcal{L}(X)$ *such that* $\mu([w]) \neq \mu([\varphi^{-1}(w)])$*, then* φ *does not define an automorphism of* X*.*

In other words, in shifts where one can estimate the measure of cylinder sets for a characteristic measure, one can quickly eliminate many block codes for consideration as automorphisms of the system. Here we note that Lemma 7.1 generalizes a theorem of Hedlund in the special case of full shifts [16, Theorem 5.4]. We demonstrate how to use Lemma 7.1, in the case of the well-known Fibonacci shift.

7.2. **Use of Lemma 7.1 for the Fibonacci shift.** The *Fibonacci shift* is the subshift of finite type whose only forbidden word is 11; it is called the Fibonacci shift because the size of $\mathcal{L}_n(X)$ grows according to the Fibonacci sequence (note this should not be confused with the Fibonacci substitution system, also known as the Fibonacci Sturmian shift). It is not hard to show that it is topologically mixing and so by Parry's theorem [23], it has a unique measure of maximal entropy μ, called its *Parry measure*. As already noted, μ is a characteristic measure for X and so any automorphism on X preserves μ. The Parry measure of a subshift of finite type can be written down explicitly and a computation following [23] (see also [15, Section 4.4]) gives the measures:

$$\mu([00]) = \frac{b}{b+2} \approx 0.44721;$$

$$\mu([10]) = \mu([01]) = \frac{1}{b+2} \approx 0.27639;$$

$$\mu([0000]) = \mu([0010]) = \mu([0100]) = \frac{a}{b+2} \approx 0.17082;$$

$$\mu([0001]) = \mu([0101]) = \mu([1000]) = \mu([1010]) = \frac{a^2}{b+2} \approx 0.10557;$$

$$\mu([1001]) = \frac{a^3}{b+2} \approx 0.06525,$$

where $a := 2/(\sqrt{5}+1)$ and $b := 2/(\sqrt{5}-1)$. As a particular example, let φ denote the block code with range 1 given by:

$$\varphi(000) = \varphi(001) = 0;$$
$$\varphi(010) = \varphi(100) = \varphi(101) = 1.$$

To determine if φ defines an automorphism of X, we first check that $\varphi(X) \subseteq X$ by checking that the image of any element of $\mathcal{L}_4(X)$ is an element of $\mathcal{L}_2(X)$ (this guarantees that the forbidden word 11 does not occur in any element of $\varphi(X)$). Then we check if the necessary conditions provided by Lemma 7.1 are satisfied. In this case, we conclude that φ does *not* have an inverse block code because

$$\mu([\varphi^{-1}(00)]) = \mu([0000] \cup [0001]) = \mu([0000]) + \mu([0001]) < \mu([00]).$$

Of course for subshifts that are not shifts of finite type, the problem of determining whether $\varphi(X) \subseteq X$ is more challenging. Further, even in cases when it is known to exist, finding a characteristic measure and explicitly writing down the measure of small cylinder sets is significantly more difficult. However, we mention that the characteristic measure we construct on language stable shifts is a weak* limit of Parry measures on shifts of finite type (at least in the case when the terms in the SFT cover are themselves topologically mixing) and this allows one to approximate the measure our characteristic measure would give to a (small) cylinder set by studying the measure it gets in the terms of the SFT cover.

References

[1] D. V. Anosov. On N. N. Bogolyubov's contribution to the theory of dynamical systems. *Uspekhi Mat. Nauk* **49** (1994), no. 5 (299), 5–20; translation in *Russian Math. Surveys* **49** (1994), no. 5, 1–18.

[2] J.-P. Allouche and J. Shallit. Automatic sequences. Theory, applications, generalizations. Cambridge University Press, Cambridge, 2003. xvi+571 pp.

[3] N. N. Bogolyubov. On some ergodic properties of continuous transformation groups. *Nauch. Zap. Kiev Univ. Phys.-Mat. Sb.* **4**:3 (1939), 45–53.

[4] M. Boyle and W. Krieger. Periodic points and automorphisms of the shift. *Trans. Amer. Math. Soc.* **302** (1987), no. 1, 125–149.

[5] M. Boyle, D. Lind, and D. Rudolph. The automorphism group of a shift of finite type. *Trans. Amer. Math. Soc.* **306** (1988), no. 1, 71–114.

[6] M. P. Béal, F. Mignosi, and A. Restivo. Minimal forbidden words and symbolic dynamics. *STACS 96 (Grenoble, 1996)*, 555–566, Lecture Notes in Comput. Sci., **1046**, Springer, Berlin, 1996.

[7] M. P. Béal, F. Mignosi, A. Restivo, and M. Sciortino. Forbidden words in symbolic dynamics. *Adv. in Appl. Math.* **25** (2000), no. 2, 163–193.

[8] R. Bowen. Topological entropy and axiom A. 1970 Global Analysis (*Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif.,* 1968) pp. 23–41 Amer. Math. Soc., Providence, R.I.

[9] E. M. Coven and M. E.. Paul. Endomorphisms of irreducible subshifts of finite type. *Math. Systems Theory* **8** (1974/75), no. 2, 167–175.

[10] E. M. Coven and J. Smítal. Entropy-minimality. *Acta Math. Univ. Comenian. (N.S.)* **62** (1993), no. 1, 117–121.

[11] F. Durand. Linearly recurrent subshifts have a finite number of non-periodic subshift factors. *Ergodic Theory Dynam. Systems* **20** (2000), no. 4, 1061–1078.

[12] S. Fischler. Palindromic prefixes and episturmian words. *J. Combin. Theory Ser. A* **113** (2006), no. 7, 1281–1304.

[13] J. Frish and O. Tamuz. Symbolic dynamics on amenable groups: the entropy of generic shifts. *Ergodic Theory Dynam. Systems* **37** (2017), no. 4, 1187–1210.

[14] J. Frish and O. Tamuz. Characteristic measures of symbolic dynamical systems. `arXiv:1908.02930`.

[15] B. Hasselblatt and A. Katok. Introduction to the modern theory of dynamical systems. Cambridge University Press, Cambridge, 1995.

[16] G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory.* **3** (1969), 320–375.

[17] M. Hochman. Genericity in topological dynamics. *Ergodic Theory Dynam. Systems* **28** (2008), no. 1, 125–165.

[18] A. Johnson. Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers. *Israel J. Math.* **77** (1992) no. 1-2, 211–240.

[19] N. Kryloff and N. Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des sysèmes dynamiques de la mécanique non linéaire. *Ann. of Math.* (2) **38** (1937), no. 1, 65–113.

[20] D. Lind and B. Marcus. An introduction to symbolic dynamics and coding. Cambridge University Press, Cambridge, 1995.

[21] E. Lindenstrauss. Pointwise theorems for amenable groups. *Inv. Math.* **146** (2001), no. 2, 259–295.

[22] F. Mignosi, A. Restivo, and M. Sciortino. Words and forbidden factors. WORDS (Rouen, 1999). *Theoret. Comput. Sci.* **273** (2002), no. 1-2, 99–117.

[23] W. Parry. Intrinsic Markov chains. *Trans. Amer. Math. Soc.* **112** (1964), 55–66.

[24] D. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergodic Theory Dynam. Systems* **10** (1990), no.2, 395–406.

[25] V. Salo and M. Schraudner. Automorphism groups of subshifts through group extensions. Preprint.

[26] V. Sos. On the distribution mod 1 of the sequence $n\alpha$. *Ann. Univ. Sci. Budapest, Eötvös Sect. Math.*, **1** (1958), 127–134.

[27] P. Walters. An introduction to ergodic theory. Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.

Bucknell University, Lewisburg, PA 17837 USA
Email address: van.cyr@bucknell.edu

Northwestern University, Evanston, IL 60208 USA
Email address: kra@math.northwestern.edu