Ergod. Th. & Dynam. Sys., (2026), **46**, 2645–2701 © The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

doi:10.1017/etds.2026.10318

Density of group languages in shift spaces

VALÉRIE BERTHÉ†, HERMAN GOULET-OUELLET‡, CARL-FREDRIK NYBERG BRODDA§, DOMINIQUE PERRIN¶ and KARL PETERSEN‖

† *IRIF, Université Paris Cité, Paris, France*
(e-mail: berthe@irif.fr)
‡ *Département de mathématiques et de statistique, Université de Moncton, Moncton, Canada*
(e-mail: herman.goulet-ouellet@umoncton.ca)
§ *School of Mathematics, Korea Institute for Advanced Study, Seoul, Republic of Korea*
(e-mail: cfnb@kias.re.kr)
¶ *Laboratoire d'Informatique Gaspard Monge, Université Paris Est, Paris, France*
(e-mail: dominique.perrin@esiee.fr)
‖ *Mathematics Department, University of North Carolina, Chapel Hill, NC, USA*
(e-mail: petersen@math.unc.edu)

(*Received* 29 *August* 2024 *UTC and accepted in revised form* 5 *May* 2026 *UTC*)

Abstract. The density of a rational language can be understood as the frequency of some pattern in the shift space, for example, a pattern like 'words with an even number of a given letter'. We study the density of group languages, that is, rational languages recognized by morphisms onto finite groups, inside shift spaces. We show that the density with respect to any given ergodic measure on a shift space exists for every group language, because it can be computed by using any ergodic lift of the given measure to a skew product between the shift space and the recognizing group. We then further study densities in shifts of finite type (with a suitable notion of irreducibility) and then in minimal shifts. In the latter case, we obtain a closed formula for the density under the condition that the aforementioned skew product has minimal closed invariant subsets that are ergodic under the product of the original measure and the uniform probability measure on the group. The formula is derived in part from a characterization of minimal closed invariant subsets for skew products between shifts and finite groups relying on notions of cocycles and coboundaries. In the case where the whole skew product is ergodic under the product measure, then the density is just the cardinality of the subset of the group that defines the language divided by the cardinality of the group. Moreover, we provide sufficient conditions for the skew product to have minimal closed invariant subsets that are ergodic under the product

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

measure. Finally, we investigate the link between minimal closed invariant subsets, return words, and bifix codes.

Key words: symbolic dynamics, ergodic theory, automata theory
2020 Mathematics Subject Classification: 37B10 (Primary); 37A20, 68Q45 (Secondary)

Contents

1 Introduction 2646
2 Symbolic dynamics 2650
3 Group languages and skew products 2653
4 The case of Bernoulli measures 2666
5 Densities in SFTs 2669
6 A characterization of minimality via return words 2677
7 Minimal subsets and modular cobounding maps 2684
8 Ergodicity for primitive substitutions 2692
9 Skew products based on Sturmian shifts 2696
Acknowledgements 2699
References 2699

1. *Introduction*

The study of language density can be traced back to the work of Schützenberger in the 1960s [**53**], Berstel in the 1970s [**6**], and Hansel and Perrin in the 1980s [**32**]. The idea also appears in Eilenberg's monograph [**26**] and in the monograph by Berstel, Perrin, and Reutenauer [**8**]. These earlier works, motivated mainly by automata theory and the theory of codes, focused on density with respect to *Bernoulli measures* (see §4).

In the present paper, we draw our motivation from symbolic dynamics and turn to *ergodic measures on shift spaces*. Within this setting, the density of a given language can be understood as the frequency (with respect to some given ergodic measure) of some pattern in the shift space—for example, a pattern like "words with an even number of a given letter". More precisely, given a shift X on the alphabet A endowed with a shift-invariant probability measure μ, and $L \subseteq A^*$ a language, the *density* $\delta_\mu(L)$ *of* L *under the measure* μ is defined as the Cesàro limit of $\mu(\{u \in L : |u| = n\})$ as $n \to \infty$, whenever it exists.

We restrict our attention to *group languages*, i.e., languages recognized by morphisms onto finite groups: by fixing a morphism $\varphi\colon A^* \to G$ onto a finite group G, we consider a language of the form $L = \varphi^{-1}(K)$, $K \subseteq G$. The density of the languages of this form can be understood using a *skew product* (sometimes called a *group extension*), which, briefly put, is a dynamical system built from a group G, together with a dynamical system acting on G. In our case, we consider a first dynamical system, a shift X over A that supports a measure μ, which is then enriched through an action on a finite group G defined via a morphism φ. This yields a further dynamical system, the skew product, which is governed both by the dynamics of the original system X and by the algebraic structure of the group G. A precise definition of the skew products used in this paper may be found at the

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

beginning of §3. The key idea in this paper is that the density can be expressed in terms of limits of *ergodic sums in a skew product* (cf. Theorem 1.1). This allows us to show that the density always exists and, under suitable conditions, to calculate it. This approach is strongly related with foundational work by Furstenberg, Veech, Schmidt, and Zimmer (among others) on ergodic properties of skew products [**29**, **52**, **56**, **57**, **59**]. Veech's work, in particular, is concerned with a special case of the very same notion of density studied here and it played a key role in guiding our investigation.

Given a shift space X with an ergodic measure μ, one of our main results relates the density $\delta_\mu(L)$ with ergodic measures on the skew product between X and G, denoted $G \rtimes_\varphi X$, where the skewing function is the *cocycle* determined by φ (a skewing function which depends only on the symbol at the zero coordinate of elements of the shift). We say that a measure $\overline{\mu}$ on the skew product *projects to* μ if $\overline{\mu}(G \times B) = \mu(B)$ for every measurable $B \subseteq X$.

THEOREM 1.1. *Let X be a shift space on a finite alphabet A with a shift-invariant probability measure μ and let $\varphi\colon A^* \to G$ be a morphism onto a finite group G. For every group language $L = \varphi^{-1}(K)$, where $K \subseteq G$, the density $\delta_\mu(L)$ exists. Moreover, if μ is ergodic, the density is given by*

$$\delta_\mu(L) = \sum_{g \in G} \overline{\mu}(\{g\} \times X)\, \overline{\mu}(gK \times X),$$

where $\overline{\mu}$ is any ergodic measure on the skew product $G \rtimes_\varphi X$ that projects to μ.

The result is proved in §3. The proof uses the known fact that the skew product always admits an ergodic measure $\overline{\mu}$ that projects to μ (see Lemma 3.4). A natural candidate for the measure $\overline{\mu}$ is the product measure $\nu \times \mu$, where ν is the uniform probability measure on G. When $\nu \times \mu$ happens to be ergodic, then the above formula takes on the following very simple form (see Corollary 3.6):

$$\delta_\mu(L) = \sum_{g \in G} (\nu \times \mu)(\{g\} \times X)\ (\nu \times \mu)(gK \times X) = \sum_{g \in G} (1/|G|)(|gK|/|G|) = |K|/|G|.$$

However, this is not always the case: see for instance Examples 3.7, 5.19, 7.12, and 7.13. We also introduce in §3.3 a pointwise version of density (Definition 3.10). This leads to a further proof of the existence of the density of group languages, expressed as an integral over pointwise densities.

In the rest of the paper, we apply Theorem 1.1 to two cases: when μ is a Markov measure on a shift of finite type (SFT) and when μ is an ergodic measure on a minimal shift space. In the former case, we present a characterization of ergodicity of the product measure $\nu \times \mu$ when μ is an *r-step Markov* measure, $r \geq 1$, meaning that

$$\mu(x_m = a \mid x_{[0,m)} = w) = \mu(x_m = a \mid x_{[m-r,m)} = w_{[m-r,m)})$$

for all words w of length $m \geq r$ such that $\mu(x_m = a \mid x_{[0,m)} = w) \neq 0$. We define for this purpose a suitable notion of irreducibility, called *φ-irreducibility* (Definition 5.1), which leads to the following theorem.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

THEOREM 1.2. *Let μ be an r-step Markov measure on $A^{\mathbb{Z}}$, $r \geq 1$, and let X be the support of μ. Let $\varphi\colon A^* \to G$ be a morphism onto a finite group with uniform probability measure ν. If X is φ-irreducible, then the product measure $\nu \times \mu$ is ergodic on the skew product $G \rtimes_\varphi X$.*

In particular, for every $K \subseteq G$, the language $L = \varphi^{-1}(K)$ has density $\delta_\mu(L) = |K|/|G|$.

When μ is a Bernoulli measure, the above result also follows from the work of Hansel and Perrin [32, Theorem 3], which uses a completely different approach based on the theory of codes. We provide a survey of this approach in §4.

In the case of minimal shift spaces, we specialize Theorem 1.1 into a formula that holds under the condition that $\nu \times \mu$ is ergodic on the minimal closed invariant subsets of $G \rtimes_\varphi X$. This generalizes the simple formula found when the product measure $\nu \times \mu$ is ergodic.

Let us briefly present the notion, inspired by [38, Proposition 2.1], which lies at the core of this more general formula. Given a subgroup $H \leq G$, a map α from X to the right coset space $H\backslash G$ is called a *cobounding map* mod H if it satisfies the following coboundary type equation, where S denotes the shift map:

$$\alpha(Sx) = \alpha(x)\varphi(x_0).$$

The term *cobounding* appears for instance in a paper by Baggett *et al* [3]; the terms *transfer function* and *intertwining* have also been used with similar meanings in [3, 4, 51]. As the name suggests, this notion is related with *cocycles* and *coboundaries*, and more broadly to the long history of cohomology in ergodic theory. In addition to [38] and the previously mentioned work of Furstenberg, Veech, Schmidt, and Zimmer, we drew inspiration also from [2, 20]. The following theorem is our third main result; its statement uses a natural partial order on cobounding maps, which is introduced in §7.2.

THEOREM 1.3. *Let X be a minimal shift space on A with an ergodic measure μ and $\varphi\colon A^* \to G$ a morphism onto a finite group G with uniform probability measure ν. Suppose that $\nu \times \mu$ is ergodic on each of the minimal closed invariant subsets of $G \rtimes_\varphi X$. Then, for every group language $L = \varphi^{-1}(K)$, where $K \subseteq G$, the density $\delta_\mu(L)$ is given by the following formula, where $\alpha\colon X \to H\backslash G$ is any minimal cobounding map:*

$$\delta_\mu(L) = \frac{1}{|H|} \sum_{k \in K} \sum_{Hg \in H\backslash G} \mu(\alpha^{-1}(Hg))\mu(\alpha^{-1}(Hgk)).$$

Behind this last result is a bijection between the minimal closed invariant subsets of $G \rtimes_\varphi X$ and the cobounding maps that are minimal under the aforementioned partial order (Proposition 7.4). This entails, among other things, that whenever X is minimal, $\nu \times \mu$ can be ergodic only when $G \rtimes_\varphi X$ itself is minimal (Corollary 7.2). The formula in Theorem 1.3 is derived from Theorem 1.1, where the role of the ergodic measure $\overline{\mu}$ is played by an appropriate rescaling of the measure $\nu \times \mu$, with ν being the uniform probability measure on G. This rescaling explains the factor $1/|H|$ in the formula.

Theorem 1.3 motivates further study of the structure of the skew products between minimal shifts and finite groups, which includes a characterization of minimality of such

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

skew products in terms of *return words* (Theorem 6.1). We first give a combinatorial proof of this characterization using the theory of bifix codes, inspired by ideas from earlier works [**7**, **12**]. Later, we show that the same conclusion can be reached using cobounding maps (Proposition 7.8).

Having in mind Theorem 1.3, we also provide sufficient conditions for ergodicity of $\nu \times \mu$ on minimal closed invariant subsets of $G \rtimes_\varphi X$. A first condition is given in Corollary 7.11 combined with Proposition 7.10, and a second one, restricted to the case of shift spaces generated by *primitive substitutions*, is given in Proposition 8.3. We deduce as a corollary of the second condition that *substitutive dendric shifts* (definitions are recalled in §8) have ergodic skew products with all finite groups and, consequently, the density of $L = \varphi^{-1}(K)$ is simply $|K|/|G|$, $K \subseteq G$ (see Corollary 3.6).

Here is an example of what these results tell us about specific systems. Consider the Fibonacci shift on the two-letter alphabet $\{a, b\}$, whose definition is recalled in §3.4. For each word w in the language of the shift space, denote by $|w|_a$ the number of occurrences of a in w. Then, we can apply Theorem 1.1 to deduce that the average probability over sufficiently long words w that $|w|_a$ is congruent to r mod m is approximately $1/m$ for each $r = 0, \ldots, m-1$. However, the convergence of these probabilities holds only in the sense of Cesàro mean. This specific example with $m = 2$ is developed in detail in §3.4.

There have been previous results of a similar nature concerning equidistribution modulo m. For instance, Veech [**56**, **57**] studied the parity of the number of visits to an interval by the orbit of a point under an irrational rotation (see Example 3.11), and Jager and Liardet [**34**] studied the congruence classes of the matrices in $\mathrm{GL}(2, \mathbb{Z})$ associated with continued fraction expansions of real numbers (see Example 3.16). We also revisit these results in §9 in the particular case of Sturmian shifts. In each case, it is ergodicity of a relevant skew product that implies equidistribution among cosets. Our approach provides still more examples. These examples include the Thue–Morse shift (explored in Examples 2.1 and 7.12) and the case of substitutive dendric shifts (presented in §8.2), which includes substitutive Sturmian shifts and substitutive codings of interval exchanges.

Let us give a brief overview of the paper's structure. Section 2 gives some preliminaries on symbolic dynamics. Section 3 recalls the definition of skew products and gives an elementary proof of our first main result, Theorem 1.1, which shows in particular that the density always exists. We also provide an alternative proof of the existence using a pointwise version of density. We then survey the original approach to density under Bernoulli measures in §4, relying on notions from algebraic theory of formal languages. The study of the density for SFTs is then handled in Theorem 1.2 together with a discussion on various notions of irreducibility. Section 6 contains the material on bifix codes and a characterization of minimal skew products in terms of return words. Our third main result, Theorem 1.3, is presented in §7, where cobounding maps are studied. The section also contains a simple sufficient condition for ergodicity of the product measure $\nu \times \mu$ on minimal closed invariant subsets. In §8, we take a closer look at shifts generated by primitive substitutions and we consider particular examples of skew products based on Sturmian shifts in §9.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

2. *Symbolic dynamics*

Let A be a finite alphabet. Let ε stand for the empty word of the free monoid A^* and $A^+ = A^* \setminus \{\varepsilon\}$. We denote by $A^{\mathbb{Z}}$ the set of two-sided infinite words on A. For any word w in the free monoid A^* (endowed with concatenation), $|w|$ denotes the length of w and $|w|_a$ stands for the number of occurrences of the letter a in the word w. We start indexing finite words with 0, so that a word $w \in A^*$ has the form $w = w_0 w_1 \cdots w_{n-1}$, where $n = |w|$. Given $0 \le i \le j < n$, we let

$$w_{[i,j)} = w_i w_{i+1} \cdots w_j,$$

and we extend a similar notation for infinite words.

A *factor* of a (finite or infinite) word w is defined as a finite concatenation of consecutive letters occurring in w, that is, a word u is a factor of w if there exist indices $i \le j$ such that $u = w_{[i,j)}$. If w is a finite word, then u is a factor of w precisely when there are words p and s such that $w = pus$. When $p = \varepsilon$ (respectively, $s = \varepsilon$), we say that u is a *prefix* (respectively, *suffix*) of w.

An infinite word $x = (x_n)_{n\in\mathbb{Z}}$ is *uniformly recurrent* if every word occurring in x occurs in an infinite number of positions with bounded gaps; in other words, for every factor w of x, there exists a positive integer m such that for every n, w is a factor of $x_{[n,n+ms)}$.

We view closed shift-invariant sets of two-sided infinite words as dynamical systems under the map S, called the *shift map*, defined by

$$S((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}.$$

A *shift space* (also shortened to *shift*) is a pair (X, S), where X is a topologically closed and shift-invariant subset of $A^{\mathbb{Z}}$ for some finite alphabet A. We usually shorten (X, S) as X when we refer to the system (X, S). The *language* of X is defined as the set $\mathcal{L}(X)$ of factors of elements of X

$$\mathcal{L}(X) = \{x_{[i,j)} \mid x \in X,\ i, j \in \mathbb{Z},\ i \le j\}.$$

When a shift X is said to be defined on the alphabet A, we assume that $A \subseteq \mathcal{L}(X)$.

A shift space is said to be *minimal* if it admits no proper non-empty closed and shift-invariant subset; equivalently, the S-orbit of every element of X is dense. Note that a shift space X is minimal if and only if every infinite word $x \in X$ is uniformly recurrent. However, a shift space is called *irreducible* if *there exists* an element $x \in X$ with dense S-orbit. This is equivalent to the following property of $\mathcal{L}(X)$: for every $u, v \in \mathcal{L}(X)$, there exists $w \in A^*$ such that $uwv \in \mathcal{L}(X)$.

Let X be a shift space on A and fix $w \in \mathcal{L}(X)$. We denote by $\mathcal{R}_X(w)$ the set of (right) *return words* to w. It is, by definition, the set of words r such that rw is in $\mathcal{L}(X)$ and has exactly two factors equal to w, one as a prefix and the other one as a suffix; that is,

$$\mathcal{R}_X(w) = \{r \in A^* \mid rw \in \mathcal{L}(X) \cap wA^* \setminus A^+wA^+\}.$$

Let $X \subseteq A^{\mathbb{Z}}$ be a shift space equipped with a Borel probability measure μ. For $w \in \mathcal{L}(X)$, we denote

$$[w]_X = \{x \in X \mid x_{[0,n)} = w\}$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

the right cylinder defined by w (*right* refers here to the fact this definition involves only non-negative indices). When X is clear from context, we often drop the subscript X and write simply $[w]$. Note that we have $\mu([\varepsilon]) = 1$ and

$$\sum_{a\in A} \mu([wa]) = \mu([w]).$$

The measure μ is *invariant* if $\mu(S^{-1}U) = \mu(U)$ for every Borel set $U \subseteq X$. Note that μ is invariant if and only if, for every $w \in \mathcal{L}(X)$,

$$\sum_{a\in A} \mu([aw]) = \mu([w]).$$

Recall that an invariant measure μ is *ergodic* if every Borel set U which is invariant (that is, $S^{-1}U = U$) has measure 0 or 1. A well-known equivalent condition is that μ is ergodic if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(U \cap S^{-i}V) = \mu(U)\mu(V), \tag{2.1}$$

for every pair U, V of Borel sets [48, p. 56, Exercise 4(a)]. If instead of converging on average (that is, in Cesàro's sense) the sequence $(\mu(U \cap S^{-n}V))_{n\in\mathbb{N}}$ converges directly to $\mu(U)\mu(V)$ for all pairs of Borel sets U, V, then the measure is said to be *mixing*.

The shift X is said to be *uniquely ergodic* if there is a unique invariant probability measure on X, in which case, this unique measure is necessarily ergodic. An important class of uniquely ergodic shifts is given by *primitive substitution shifts*, whose definition may be recalled in [25, §1.4] for instance. By a theorem of Michel, every primitive substitution shift is uniquely ergodic [43]. Other important sufficient conditions for unique ergodicity of shift spaces are due to Boshernitzan [16, 17]. More details on such results may be found in [25, 50].

Example 2.1. The Thue–Morse shift $X = X(\sigma)$ with $\sigma : a \mapsto ab, b \mapsto ba$ is uniquely ergodic by Michel's theorem. Its unique ergodic measure μ is depicted in Figure 1.

The aim of this paper is to study the density of languages under a given measure μ, defined as follows.

Definition 2.2. Let L be a language on an alphabet A and μ a measure on $A^{\mathbb{Z}}$. The *density* of L under the measure μ is the following limit whenever it exists:

$$\delta_\mu(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu([L \cap A^i]).$$

In other words, $\delta_\mu(L)$ is the Cesàro limit of $\mu([L \cap A^n])$ as $n \to \infty$. Since $\mu(w) = 0$ when $w \notin \mathcal{L}(X)$, we have of course that $\delta_\mu(L) = \delta_\mu(L \cap \mathcal{L}(X))$ and $\delta_\mu(\mathcal{L}(X)) = 1$. Moreover, it follows from the definition of density that, whenever L and L' are disjoint,

$$\delta_\mu(L \cup L') = \delta_\mu(L) + \delta_\mu(L').$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

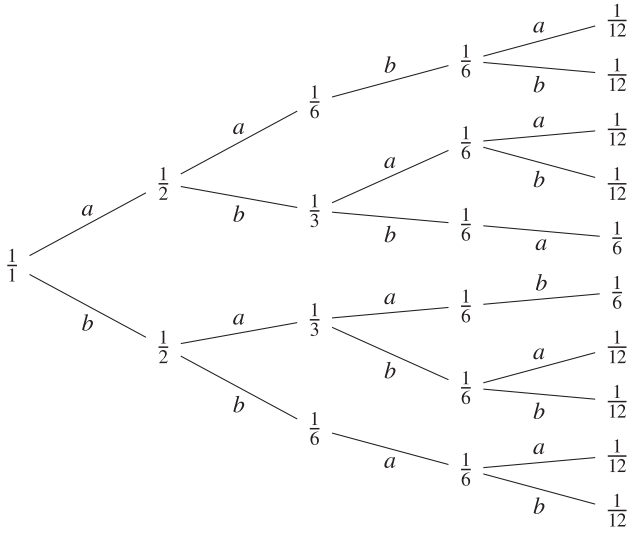


FIGURE 1. The invariant probability measure on the Thue–Morse shift.

There is also a dual rule for the intersection: when $L \cup L' = A^*$, the density of the intersection is given by $\delta_\mu(L \cap L') = \delta_\mu(L) + \delta_\mu(L') - 1$.

Note however that the density of a general intersection $L \cap L'$ might not exist, even when the density of both L and L' exist. For instance, if we consider the two languages

$$L = \{w \in A^* : |w| \equiv 1 \bmod 2\}, \quad L' = \{w \in A^* : |w| \equiv \lfloor \log_2(|w|) \rfloor \bmod 2\}$$

over some fixed finite alphabet A, then it is not hard to see that $\delta_\mu(L) = \delta_\mu(L') = \frac{1}{2}$, no matter the probability measure μ. However, the sequence of Cesàro sums $s_n = (1/n) \sum_{i=1}^{n} \mu([L \cap L' \cap A^i])$ has two subsequences $(s_{2^{2n}})_{n\in\mathbb{N}}$ and $(s_{2^{2n+1}})_{n\in\mathbb{N}}$, which converge to $1/3$ and $1/6$, respectively. Therefore, the sequence of Cesàro sums does not converge and $L \cap L'$ does not have a density.

An easy observation is that a finite language has zero density for every measure. This observation can be generalized to a larger class of languages called *thin languages*, see §4.

Next, we introduce a language that serves as a guiding example throughout the paper.

Example 2.3. Consider integers m, d, r with $m, d \geq 2$ and $0 \leq r \leq m - 1$. Define the language

$$L_r := \{w \in A^* : |w|_1 \equiv r \bmod m\}$$

on the alphabet $A = \{1, 2, \ldots, d\}$, as the set of finite words having a number of 1s congruent to r modulo m. This is a group language, as defined in the next section (see also Example 3.2). One of our motivations for the present study is to establish the existence of the density of $\delta_\mu(L_r)$ and obtain conditions for *equidistribution*, that is, $\delta_\mu(L_r) = 1/m$.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

3. *Group languages and skew products*

We first recall basic definitions on group languages and skew products in §3.1. We then prove Theorem 1.1 in §3.2. The proof uses any ergodic lift to the skew product of the given ergodic measure on the shift space. We also consider in Corollary 3.6 the special case where the product between the original measure and the uniform probability measure on the group is ergodic. If the shift space is uniquely ergodic, then this is even equivalent to unique ergodicity of the skew product, cf. Proposition 3.9. We then consider in §3.3 a pointwise version of density. Lastly, we focus on the case of the Fibonacci shift in §3.4, where we prove that the density considered as a Cesàro mean converges, whereas the sequence of measures $(\mu([L \cap A^n]))_{n\in\mathbb{N}}$ does not converge in the classical sense.

3.1. *First definitions.*

A *rational language* is a language recognized by a finite automaton. Equivalently, a language is rational if and only if it is of the form $\varphi^{-1}(N)$, where $\varphi\colon A^* \to M$ is a morphism onto a finite monoid M and $N \subseteq M$ is a subset.

A *group language* is a set of the form $L = \varphi^{-1}(K)$, where $\varphi\colon A^* \to G$ is a morphism onto a finite group G and $K \subseteq G$. Note that such languages are in particular rational, being recognized by finite groups. For more on the topic, see [8].

Let X be a minimal shift space equipped with an invariant measure μ. We consider the *skew product* of G and X with respect to φ, denoted $G \rtimes_\varphi X$: it is the dynamical system $G \rtimes_\varphi X = (G \times X, T_\varphi)$, where T_φ is the transformation defined by

$$T_\varphi(g, x) = (g\varphi(x_0), Sx). \tag{3.1}$$

More generally, T_φ satisfies, for every $n \in \mathbb{Z}$,

$$T_\varphi^n(g, x) = (g\varphi^{(n)}(x), S^n x),$$

where $\varphi^{(n)}$ is defined by

$$\varphi^{(n)}(x) = \begin{cases} \varphi(x_{[0,n)}) & \text{if } n \geq 0, \\ \varphi(x_{[n,0)})^{-1} & \text{if } n < 0. \end{cases} \tag{3.2}$$

The map $(n, x) \mapsto \varphi^{(n)}(x)$ is the *cocycle* defined by φ. When the morphism φ is clear from context, we may simply write T and $G \rtimes X$. Skew products constitute one of the basic extensions of dynamical systems; see e.g. [21, 48].

LEMMA 3.1. *The skew product $G \rtimes_\varphi X$ is topologically conjugate to a shift space on $G \times A$ via the map $\Psi\colon G \rtimes_\varphi X \to (G \times A)^{\mathbb{Z}}$ defined by*

$$\Psi(g, x)_n = (g\varphi^{(n)}(x), x_n), \quad n \in \mathbb{Z}.$$

Note that the lemma does not say that Ψ is surjective on $(G \times A)^{\mathbb{Z}}$, but simply that Ψ is a topological conjugacy from $G \rtimes_\varphi X$ to the subshift $\Psi(G \rtimes_\varphi X)$.

Proof. The map Ψ is continuous and injective; hence, it is an homeomorphism onto its image. Moreover, it follows from the definitions that $\Psi \circ T_\varphi = S \circ \Psi$. □

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

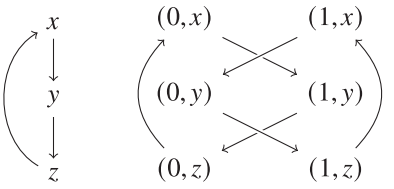


FIGURE 2. The finite shift X generated by $x = (abc)^\infty$ and its skew product with $\mathbb{Z}/2\mathbb{Z}$ from Example 3.3.

Example 3.2. Let $L_r = \{w \in A^* : |w|_1 \equiv r \bmod m\}$ be the language from Example 2.3. Let X be a shift on $\{1, \dots, d\}$. We consider the skew product $\mathbb{Z}/m\mathbb{Z} \rtimes_\varphi X$, where $\mathbb{Z}/m\mathbb{Z}$ is the m-element additive group, and φ is the morphism $\{1, \dots, d\}^* \to \mathbb{Z}/m\mathbb{Z}$ given by $1 \mapsto 1$ and $i \mapsto 0$ for $2 \le i \le d$. The language L_r is indeed a group language, since $L_r = \varphi^{-1}(\{r\})$. In particular, one has $\varphi^{(n)}(x) = |x_0 \cdots x_{n-1}|_1$ modulo m for $n \ge 0$. Understanding the equidistribution properties of the sequence $(\varphi^{(n)}(x))_n$ in $G = \mathbb{Z}/m\mathbb{Z}$ allows one to see how often factors have a given congruence modulo m of occurrences of the letter 1; see §3.4 as an illustration.

Example 3.3. Let X be the three-element shift defined as the finite orbit of the periodic word $x = (abc)^\infty$, so $X = \{x, y, z\}$ with $y = Sx$, $z = Sy$. Let $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$ be the morphism

$$\varphi(a) = \varphi(b) = 1, \quad \varphi(c) = 0.$$

The skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ has six elements. Viewed as a shift on the alphabet $\mathbb{Z}/2\mathbb{Z} \times A$, it is the disjoint union of the orbits of two periodic words,

$$\Psi(0, x) = ((0, a)(1, b)(0, c))^\infty \quad \text{and} \quad \Psi(1, x) = ((1, a)(0, b)(1, c))^\infty.$$

Despite its apparent simplicity, this example proves to be more profound than it seems, by offering a clear way to illustrate ergodicity (see Examples 3.7 and 3.14), irreducibility (see Examples 5.6 and 5.8), as well as minimality properties (see Example 5.6). It can also be generalized as detailed in Example 3.8. The shift X and the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ are depicted in Figure 2.

3.2. *A formula for the density.* The aim of this section is to prove our first main result, Theorem 1.1. The proof uses the existence of an ergodic measure $\overline{\mu}$ on the skew product which projects to μ (see Lemma 3.4). This known fact follows from classical results in ergodic theory (see for example [**23**, **30**, **37**, **46**]); we provide a proof for the sake of completeness. Let us recall briefly that a *factor map* between two dynamical systems (Z, R) and (Y, T) is a map $\pi\colon Z \to Y$ that is onto and satisfies $\pi \circ R = T \circ \pi$.

LEMMA 3.4. *Let (Z, R) and (Y, T) be two compact dynamical systems with a factor map $\pi\colon Z \to Y$. Then, for each ergodic measure μ on (Y, T), there is an ergodic measure $\overline{\mu}$ on (Z, R) such that $\overline{\mu} \circ \pi^{-1} = \mu$.*

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Proof. Let $\mathcal{M}$ be the set of invariant measures on (Z, R) and $\mathcal{M}_\mu$ be the subset of those $\zeta \in \mathcal{M}$ such that $\zeta \circ \pi^{-1} = \mu$. First, we show that $\mathcal{M}_\mu \neq \emptyset$. Start by choosing a point $y \in Y$ satisfying

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i y) = \int_Y f \, d\mu$$

for every continuous function on X; the point y is called a *generic point* for μ. It is well known that ergodicity of μ implies that μ-almost all points of Y have this property (see for instance [30, Proposition 3.7]); thus, such a point y exists. Take a preimage $z \in \pi^{-1}(y)$ and let ζ_i be the point mass measure concentrated on $R^i z$. Consider the sequence of Cesàro averages $(1/n)\sum_{i=0}^{n-1} \zeta_i$. By the Banach–Alaoglu theorem, we may choose a subsequence that converges for the weak-$*$ topology in the space of all probability measures on (Z, R), say

$$\zeta = \lim_{k\to\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} \zeta_i.$$

Observe that $\zeta_i \circ R^{-1} = \zeta_{i+1}$ and, thus, $\zeta - \zeta \circ R^{-1} = \lim_{k\to\infty}(1/n_k)(\zeta_0 - \zeta_{n_k} + 1) = 0$, that is, $\zeta \in \mathcal{M}$. Moreover, for every measurable subset $B \subseteq Y$,

$$\zeta \circ \pi^{-1}(B) = \lim_{k\to\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} \zeta_i(\pi^{-1}(B)) = \lim_{k\to\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} 1_B(T^i y) = \mu(B).$$

Thus, $\zeta \in \mathcal{M}_\mu$.

Observe that $\mathcal{M}_\mu$ is a closed convex subspace of $\mathcal{M}$ and, since it is non-empty, we can apply the Krein–Milman theorem to conclude that it contains an extreme point. Since the ergodic measures on (Z, R) are precisely the extreme points of $\mathcal{M}$, it remains only to show that any extreme point $\overline{\mu} \in \mathcal{M}_\mu$ is also extreme in $\mathcal{M}$. Assume that $\overline{\mu} = s\zeta' + (1-s)\zeta''$, where $0 < s < 1$ and $\zeta', \zeta'' \in \mathcal{M}$. Let us prove that $\overline{\mu}$ cannot be expressed as a non-trivial convex combination of *distinct* elements of $\mathcal{M}$. Then, we have

$$\mu = \overline{\mu} \circ \pi^{-1} = s\zeta' \circ \pi^{-1} + (1-s)\zeta'' \circ \pi^{-1}.$$

Since μ, being ergodic, is an extreme point in the space of invariant measures of (Y, T), we conclude that $\zeta' \circ \pi^{-1} = \zeta'' \circ \pi^{-1} = \mu$. Therefore, $\zeta', \zeta'' \in \mathcal{M}_\mu$ and since $\overline{\mu}$ is an extreme point in $\mathcal{M}_\mu$, it follows that $\zeta' = \zeta''$. Thus, $\overline{\mu}$ is an extreme point in $\mathcal{M}$. □

Let us recall now our first main result, which expresses the density of a group language in a shift space in terms of an ergodic measure on the skew product $G \rtimes X$. This is specialized later for SFTs (§5) and then for minimal shifts (§7).

THEOREM 1.1. *Let X be a shift space on a finite alphabet A with a shift-invariant probability measure μ and let $\varphi\colon A^* \to G$ be a morphism onto a finite group G. For every group language $L = \varphi^{-1}(K)$, where $K \subseteq G$, the density $\delta_\mu(L)$ exists. Moreover, if μ is ergodic, the density is given by*

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

$$\delta_\mu(L) = \sum_{g\in G} \overline{\mu}(\{g\} \times X)\, \overline{\mu}(gK \times X),$$

where $\overline{\mu}$ is any ergodic measure on the skew product $G \rtimes_\varphi X$ that projects to μ.

Proof. We first assume that μ is ergodic.

For $C \subseteq G$, let $U_C = C \times X$; in the case where $C = \{g\}$, $g \in G$, we write simply U_g. Then, we find

$$\{g\} \times [L \cap A^i]_X = (\{g\} \times X) \cap T_\varphi^{-i}(\{gK\} \times X) = U_g \cap T_\varphi^{-i} U_{gK}.$$

Since the projection $G \times X \to X$ is a factor map, by Lemma 3.4, we may take an ergodic measure $\overline{\mu}$ on $G \rtimes_\varphi X$ that projects to μ. In terms of $\overline{\mu}$, we have

$$\mu([L \cap A^i]) = \overline{\mu}(G \times [L \cap A^i]) = \sum_{g\in G} \overline{\mu}(\{g\} \cap [L \cap A^i]) = \sum_{g\in G} \overline{\mu}(U_g \cap T_\varphi^{-i} U_{gK}).$$

Since $\overline{\mu}$ is ergodic, we can use (2.1) to obtain

$$\begin{aligned}\delta_\mu(L) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu([L \cap A^i]) \\ &= \sum_{g\in G} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \overline{\mu}(U_g \cap T_\varphi^{-i} U_{gK}) \\ &= \sum_{g\in G} \overline{\mu}(U_g)\, \overline{\mu}(U_{gK}).\end{aligned}$$

We now only assume that μ is a shift-invariant probability measure on X. Let $\mathcal{E} = \mathcal{E}(X, S)$ be the set of ergodic invariant probability measures on (X, S). By e.g. [**49**, Ch. 12], there exists a Borel probability measure τ on $\mathcal{E}$ such that

$$\mu = \int_{\mathcal{E}} \nu \, d\tau(\nu).$$

Then, by the dominated convergence theorem

$$\delta_\mu(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu([L \cap A^i]) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathcal{E}} \nu([L \cap A^i])\, d\tau(\nu) = \int_{\mathcal{E}} \delta_\nu(L)\, d\tau(\nu),$$

where the densities $\delta_\nu(L)$ exist by the first part of the proof. □

Remark 3.5. Under the stronger assumption that $\overline{\mu}$ is mixing, then the density converges in a stronger sense, that is,

$$\delta_\mu(L) = \lim_{n\to\infty} \mu([L \cap A^n]).$$

See for instance Example 5.11.

In the case where $\nu \times \mu$ is ergodic on the skew product, the above theorem yields the following simple formula for the density.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

COROLLARY 3.6. *Let X be a shift space on a finite alphabet A with an ergodic measure μ and let $\varphi\colon A^* \to G$ be a morphism onto a finite group G with uniform probability measure ν. If the product measure $\nu \times \mu$ is ergodic on $G \times X$, then for every group language $L = \varphi^{-1}(K)$ with $K \subseteq G$, $\delta_\mu(L) = |K|/|G|$.*

Proof. Since $(\nu \times \mu)(U_C) = |C|/|G|$ for all $C \subseteq G$, the above theorem with $\overline{\mu} = \nu \times \mu$ yields $\delta_\mu(L) = \sum_{g \in G} |K|/|G|^2 = |K|/|G|$. □

There are however many examples where $\nu \times \mu$ is not ergodic on the skew product, such as the following simple one.

Example 3.7. Let X be the three-element shift, φ be as in Example 3.3, and $L = \varphi^{-1}(0)$. The shift X is uniquely ergodic, with ergodic measure μ given by the uniform probability measure. Moreover,

$$L \cap \mathcal{L}(X) = (abc)^*\{\varepsilon, ab\} \cup (bca)^*\{\varepsilon\} \cup (cab)^*\{\varepsilon, c\},$$

and, thus,

$$\mu([L \cap A^i]) = \begin{cases} 1 & \text{if } i \equiv 0 \bmod 3, \\ 1/3 & \text{otherwise.} \end{cases}$$

It follows that $\delta_\mu(L)$ is given by $\delta_\mu(L) = (1 + 1/3 + 1/3)/3 = 5/9$. In contrast, if $\nu \times \mu$ would be ergodic, then we should find instead $1/2$. The fact that $\nu \times \mu$ is not ergodic can also be observed directly by noting that $\mathbb{Z}/2\mathbb{Z} \curlywedge_\varphi X$ has two invariant subsets, each of $\nu \times \mu$-measure $1/2$ (which are, in fact, minimal invariant subsets).

We can also compute $\delta_\mu(L) = 5/9$ using the formula from Theorem 1.1. Indeed, there are two ergodic measures $\overline{\mu}_1$ and $\overline{\mu}_2$ that project to μ, each supported on one of the two invariant subsets of $\mathbb{Z}/2\mathbb{Z} \curlywedge_\varphi X$. For each i, $\overline{\mu}_i(\{h\} \times X)$ is either $1/3$ or $2/3$, depending on whether $h = 0$ or 1. Therefore, for $i = 1$ or 2, we find $\delta_\mu(L) = \overline{\mu}_i(\{0\} \times X)^2 + \overline{\mu}_i(\{1\} \times X)^2 = 5/9$.

Example 3.8. We can, in fact, generalize the above example as follows. Take $d \geq 3$ and consider an alphabet $A = \{a_0, \ldots, a_{d-1}\}$ of size d. Let X be the d-element shift space generated by the periodic word $(a_0 \cdots a_{d-1})^\infty$ and $\varphi\colon A^* \to \mathbb{Z}/d\mathbb{Z}$ mapping a_0 to 0, and a_i to 1 for $0 < i < d$. Then, the density of $\varphi^{-1}(0)$ is $(2n-1)/d^2$. Thus, the actual value of the density tends to be twice what ergodicity of $\nu \times \mu$ would yield (namely $1/d$), in the sense that $d\delta_\mu(L) \to 2$ as $d \to \infty$.

Next, we show that if X is uniquely ergodic, then ergodicity of $\nu \times \mu$ implies unique ergodicity of $G \curlywedge_\varphi X$. This is used, e.g., in Example 3.15. It was proved by Veech [**56**] for the case where X is a binary coding of an irrational rotation and $G = \mathbb{Z}/2\mathbb{Z}$. We show below how Veech's argument can be adapted in a straightforward way to the case where X is any uniquely ergodic shift and G is a finite group. The result is true more generally when G is a compact group with normalized Haar measure ν, cf. [**30**, Proposition 3.10].

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

PROPOSITION 3.9. *Let X be a uniquely ergodic shift on A with ergodic measure μ and $\varphi\colon A^* \to G$ be a morphism onto a finite group G with uniform probability measure ν. If $\nu \times \mu$ is an ergodic measure on $G \rtimes_\varphi X$, then $G \rtimes_\varphi X$ is uniquely ergodic.*

Proof. Let ζ be an invariant measure on $G \rtimes_\varphi X$. Observe that the measure $\zeta \circ P_X^{-1}$, where $P_X\colon G \rtimes_\varphi X \to X$ denotes the projection on the second component, is an invariant measure on X; thus, it must be equal to μ as X is uniquely ergodic. In other words,

$$\zeta(G \times E) = \mu(E) \quad \text{for every measurable set } E \subseteq X.$$

Let $g \in G$ act on the left of $G \rtimes_\varphi X$ by $g(h, x) = (gh, x)$. This action is T_φ-commuting, as well as measure-preserving as is easily checked on rectangular sets (that is, sets of the form $F \times E$, where $F \subseteq G$, and $E \subseteq X$ is measurable). Thus, the measure $g\zeta$ defined by $g\zeta(F) = \zeta(gF)$ is also an invariant measure. We claim that the average measure $\overline{\zeta} = (\sum_{g\in G} g\zeta)/|G|$ is equal to $\nu \times \mu$. Indeed, for every measurable set $E \subseteq X$ and $h \in G$, we have

$$\overline{\zeta}(\{h\} \times E) = \frac{1}{|G|} \sum_{g\in G} \zeta(\{gh\} \times E) = \frac{1}{|G|}\zeta(G \times E) = \frac{1}{|G|}\mu(E).$$

Thus, we conclude that $\nu \times \mu = \overline{\zeta}$. If $\nu \times \mu$ is ergodic, then it is an extremal point in the convex set of invariant measures of $G \rtimes_\varphi X$. Since $\nu \times \mu = \overline{\zeta}$ is a uniform convex combination of the measures $g\zeta$, it follows that $g\zeta = \nu \times \mu$ for all $g \in G$. □

3.3. *Pointwise densities.* We show in this section that the concept of density of a language can be adapted to Cesàro averages along orbits of individual points (see Theorem 3.12). This leads to another proof of the fact that the density always exists for group languages (as stated in Theorem 1.1) and an expression for the density as an integral of pointwise densities. For the following definition of the pointwise density, we recall the notation $\varphi^{(i)}(x) = \varphi(x_{[0,i)})$ from (3.2).

Definition 3.10. Let X be a shift space on a finite alphabet A with a shift-invariant measure μ and let $\varphi\colon A^* \to G$ be a morphism onto a finite group G. For every group language $L = \varphi^{-1}(K)$, where $K \subseteq G$, we define the *pointwise density* of L along the orbit of a point $x \in X$ as the following limit whenever it exists:

$$\dot{\delta}_L(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_K(\varphi^{(i)}(x)),$$

where 1_K stands for the indicator function of the set K.

The main result of the section, namely Theorem 3.12, states the almost everywhere existence of pointwise densities for all group languages under all invariant measures.

Example 3.11. We continue Example 3.2. Recall that $A = \{1, \dots, d\}$ and $L = \{w \in A^* : |w|_1 \equiv r \bmod m\}$, $0 \le r < m$. Alternatively, $L = \varphi^{-1}(\{r\})$, where $\varphi\colon A^* \to \mathbb{Z}/m\mathbb{Z}$ is the morphism that sends $1 \in A$ to $1 \in \mathbb{Z}/m\mathbb{Z}$ and all other letters of A to 0. The pointwise density $\dot{\delta}_L(x)$, when it exists, is the frequency of occurrences of prefixes in x having a

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

number of 1s congruent to r mod m. In particular, Theorem 3.12 states that this frequency exists for μ-almost every x, for all invariant probability measures μ on $A^{\mathbb{Z}}$.

We now turn to Theorem 3.12. First, let us recall part of Birkhoff's ergodic theorem; for a complete statement, see [**48**, Theorem 2.3]. Let (X, T) be a dynamical system with an invariant probability measure μ. Given an L^1 map $f\colon X \to \mathbb{R}$, let us define the *time average* of f as

$$\overline{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x),$$

whenever it exists. The ergodic theorem states that $\overline{f}$ exists μ-almost everywhere and, moreover,

$$\int_X \overline{f}\, d\mu = \int_X f\, d\mu.$$

THEOREM 3.12. *Let X be a shift space on a finite alphabet A with a shift-invariant probability measure μ and a morphism $\varphi\colon A^* \to G$ onto a finite group G. For every group language $L = \varphi^{-1}(K)$, the pointwise density $\dot{\delta}_L(x)$ exists for μ-almost every $x \in X$. Moreover, the density $\delta_\mu(L)$ exists and is given by*

$$\delta_\mu(L) = \int_X \dot{\delta}_L\, d\mu.$$

Proof. Let $\overline{\mu}$ be any T_φ-invariant measure on the skew product $G \rtimes_\varphi X$ that projects to μ (for instance, the product measure $\nu \times \mu$, where ν is the uniform probability measure on G). For $g \in G$, let f_g be the indicator function of $\{g\} \times X$. By the ergodic theorem applied to $G \rtimes_\varphi X$, the time average function $\overline{f}_g$ exists on some subset $E_g \subseteq G \times X$ with $\overline{\mu}(E_g) = 1$.

Next, note that $\varphi^{(i)}(x) = g$ if and only if $T_\varphi^i(1_G, x) \in \{g\} \times X$, so

$$\dot{\delta}_L(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_K(\varphi^{(i)}(x)) = \sum_{g\in K} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f_g(T_\varphi^i(1_G, x)) = \sum_{g\in K} \overline{f}_g(1_G, x).$$

Let $E = \bigcap_{g\in G} E_g$ and $F = \{x \in X : (1_G, x) \in E\}$. Then, $\dot{\delta}_L(x)$ certainly exists for all $x \in F$ by the above, so it suffices to show that $\mu(F) = 1$. However, observe that $T_\varphi^i(h, x) \in \{g\} \times X$ is equivalent to $T_\varphi^i(1_G, x) \in \{h^{-1}g\} \times X$, so that

$$\overline{f}_g(h, x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f_g(T_\varphi^i(h, x)) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f_{h^{-1}g}(T_\varphi^i(1_G, x)) = \overline{f}_{h^{-1}g}(1_G, x).$$

It follows that if $(h, x) \in E$, then $\overline{f}_g(1_G, x) = \overline{f}_{hg}(h, x)$ for all $g \in G$ and, thus, $x \in F$. Therefore, we have $E \subseteq G \times F$. Since $\overline{\mu}$ projects to μ, we conclude that $\mu(F) = \overline{\mu}(G \times F) \geq \overline{\mu}(E)$. However, E is a finite intersection of sets of $\overline{\mu}$-measure 1; thus, $\overline{\mu}(E) = 1$ and $\mu(F) = 1$.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

For the last part of the statement, we first note that for $w \in A^i$, we have $\varphi^{(i)}(x) \in K$ if and only if $x \in [w]$ for some $w \in L \cap A^i$, and thus,

$$\int_X 1_K(\varphi^{(i)}(x))\, d\mu = \sum_{w \in L \cap A^i} \int_X 1_{[w]}(x)\, d\mu = \mu([L \cap A^i]).$$

Hence, we have

$$\begin{aligned}\int_X \dot{\delta}_L(x)\, d\mu &= \int_X \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_K(\varphi^{(i)}(x))\, d\mu \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_X 1_K(\varphi^{(i)}(x))\, d\mu \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu([L \cap A^i]) = \delta_\mu(L),\end{aligned}$$

where the integral and limit can be swapped by the dominated convergence theorem. □

Remark 3.13. We can also recover the formula for $\delta_\mu(L)$ in Theorem 1.1 using pointwise densities. Indeed, suppose that μ is an ergodic measure on X and let $\overline{\mu}$ be an ergodic measure on $G \rtimes_\varphi X$ that projects to μ (which exists by Lemma 3.4). We can then use the trick from the proof of Theorem 1.1 involving the sets $U_C = C \times X$ as follows:

$$\begin{aligned}\int_X \dot{\delta}_L\, d\mu &= \int_X \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_K(\varphi^{(i)}(x))\, d\mu(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_X 1_K(\varphi^{(i)}(x))\, d\mu(x) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(\{x \in X : \varphi^{(i)}(x) \in K\}) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \overline{\mu}(G \times \{x \in X : \varphi^{(i)}(x) \in K\}) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \sum_{g \in G} \overline{\mu}(U_g \cap T_\varphi^{-i} U_{gK}) = \sum_{g \in G} \overline{\mu}(U_g)\overline{\mu}(U_{gK}).\end{aligned}$$

Example 3.14. Consider the three-point example introduced in Example 3.3, where X is the orbit closure of the two-sided sequence $x = (abc)^\infty$, the group is $\mathbb{Z}/2\mathbb{Z}$ (with its additive structure), and the morphism $\varphi\colon \{a, b, c\}^* \to \mathbb{Z}/2\mathbb{Z}$ maps a, b to 1 and c to 0. The sequences $\varphi^{(i)}(y)$ for $y = x, Sx, S^2x$ are in turn,

$$\begin{array}{ccc} 100 & 100 & 100\ldots \\ 110 & 110 & 110\ldots \\ 010 & 010 & 010\ldots. \end{array}$$

For $L = \varphi^{-1}(0)$, the pointwise densities $\dot{\delta}_L(y)$ are the average frequencies of 0 in each of these sequences, which are in turn $2/3$, $1/3$, $2/3$. As X is a finite space with a uniform probability measure, the integral of $\dot{\delta}_L$ is simply the average $(2/3 + 1/3 + 2/3)/3 = 5/9$.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Thus, $\delta_\mu(L) = 5/9$ by Theorem 3.12, which agrees with the value found in Example 3.7. Note that the two ergodic measures on the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ are not fully supported and that the function $\dot{\delta}_L(x)$ is not constant.

We now give two examples that show how our equidistribution statement, namely Corollary 3.6, together with the unique ergodicity result from Proposition 3.9, can be applied in this pointwise setting. Both examples are pursued in §9 for Sturmian shifts (see Examples 9.1 and 9.2).

Example 3.15. This example is inspired by the equidistribution results from Veech [**56**, **57**] for irrational rotations in continuation of Example 3.2. Let L_r and φ be as in Example 3.2. Assume that X is a uniquely ergodic shift whose ergodic measure μ is such that the product measure $\nu \times \mu$ is ergodic on $G \rtimes_\varphi X$; or equivalently, $G \rtimes_\varphi X$ is uniquely ergodic (Proposition 3.9). Then, Corollary 3.6 yields that $\delta_\mu(L_r)$ exists and equals $1/m$.

We now revisit this result in terms of elements of X. Recall that $1_{\{r\}}$ stands for the indicator function of the set $\{r\}$ in $\mathbb{Z}/m\mathbb{Z}$. Applying unique ergodicity to the continuous function $1_{\{r\}} \times 1$ on $G \rtimes_\varphi X$ yields indeed that for every $x \in X$,

$$\begin{aligned}\lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} (1_{\{r\}} \times 1) T_\varphi^i(0, x) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} (1_{\{r\}} \times 1)((\varphi^{(i)}(x), S^i x)) \\ &= \lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{\{r\}}(\varphi^{(i)}(x)) = \int 1_{\{r\}} \times 1 = 1/m.\end{aligned}$$

In other words, for every $x \in X$ and for every r, one has

$$\frac{1}{N} \operatorname{Card}\{n : 0 \le n \le N-1,\ |x_0 \cdots x_{n-1}|_1 \equiv r \bmod m\} \to \frac{1}{m}.$$

In the next example, we consider a skew product with a non-Abelian skewing group, namely $G(2) = \mathrm{GL}(2, \mathbb{Z}/2\mathbb{Z})$, that is, the group of 2×2 matrices with entries in $\mathbb{Z}/2\mathbb{Z}$ and determinant 1. The example could also be carried out with $G(m)$ for arbitrary $m \ge 2$, but we treat only the case $m = 2$ for simplicity. It is inspired by the work of Jager and Liardet [**34**] (see also [**15**, **44**, **55**] for related works), and is motivated by distribution properties for convergents in continued fraction expansions.

Example 3.16. Let X be a shift on the alphabet $A = \{1, 2\}$ and consider the morphism

$$\varphi\colon A^* \to G(2), \quad k \mapsto \begin{pmatrix} 0 & 1 \\ 1 & \overline{k} \end{pmatrix},$$

where $\overline{k}$ stands for the congruence class of the integer k modulo 2. Note that this morphism is onto. Let $x = (x_n)_{n\in\mathbb{Z}} \in X$. Consider the real number in the unit interval $[0, 1]$ that admits $(x_n)_{n\in\mathbb{N}}$ as its sequence of partial quotients and let $(p_n(x)/q_n(x))_{n\in\mathbb{N}}$ stand for the associated sequence of rational approximations. Recall that one has $q_{-1}(x) = 0$, $p_{-1}(x) = 1$, $q_0(x) = 1$, $p_0(x) = 0$, and for all positive n,

$$\begin{aligned} q_{n+1}(x) &= x_{n+1} q_n(x) + q_{n-1}(x), \\ p_{n+1}(x) &= x_{n+1} p_n(x) + p_{n-1}(x) \end{aligned}$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

[33, Theorem 149] and, moreover,

$$\varphi^{(n)}(x) = \begin{pmatrix} \overline{\frac{p_{n-1}(x)}{q_{n-1}(x)}} & \overline{\frac{p_n(x)}{q_n(x)}} \end{pmatrix}.$$

In the case where the skew product $G(2) \rtimes_\varphi X$ turns out to be uniquely ergodic, the following distribution results for the sequence $(\varphi^{(n)}(x))_{n\in\mathbb{N}}$ in the group $G(2)$ can be deduced: for every $k = 1, 2$ and for every $x \in X$,

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Card}\{n : 1 \le n \le N,\ q_n(x) \equiv 0 \bmod 2\} = \frac{1}{3},$$
$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Card}\{n : 1 \le n \le N,\ q_n(x) \equiv 1 \bmod 2\} = \frac{2}{3}.$$

The numbers $1/3$ and $2/3$ come from the counting measure on $G(2)$. Consider indeed the entries of index $(2, 2)$ in the matrices of $G(2)$. There are twice as many matrices in $G(2)$ for which this entry is odd than matrices for which it is even. We show in Example 9.2 that this analysis applies, in particular, to the Fibonacci shift, because there, unique ergodicity of the skew product follows from Theorem 8.11.

3.4. *An example in the substitutive Fibonacci shift.* We now proceed to illustrate Theorem 1.1 (and, more specifically, Corollary 3.6, see also Example 3.11) in the substitutive Fibonacci shift, whose precise definition is recalled below. We consider the language $L = \varphi^{-1}(0)$, where $\varphi\colon \{a, b\}^* \to \mathbb{Z}/2\mathbb{Z}$ is the morphism defined by $\varphi(a) = 1$ and $\varphi(b) = 0$; in other words,

$$L = \{w \in \{a, b\}^* : |w|_a \equiv 0 \bmod 2\}.$$

We show that $\delta_\mu(L) = 1/2$ for μ the unique ergodic measure on the substitutive Fibonacci shift (Proposition 3.17), but also that the sequence $(\mu([L \cap A^n]))_{n\in\mathbb{N}}$ does not converge in the classical sense (Proposition 3.18). In particular, the measure $\nu \times \mu$ on the skew product is ergodic, but not mixing, cf. Remark 3.5.

Consider the substitution $\sigma : a \mapsto ab, b \mapsto a$, called the *Fibonacci substitution*. The substitution σ is primitive; thus, the shift space $X = X(\sigma)$ generated by σ, called the *substitutive Fibonacci shift*, or *Fibonacci shift* for short, is uniquely ergodic by Michel's theorem (this alternatively follows from Boshernitzan's criterion [16]). Its unique ergodic measure μ viewed as a map on $\mathcal{L}(X)$ is depicted in Figure 3.

We next prove that the skew product of the Fibonacci shift with $\mathbb{Z}/2\mathbb{Z}$ for the skewing function determined by φ is also uniquely ergodic, as an application of Michel's theorem. The argument is a special case of the general method described in §8.

PROPOSITION 3.17. *The skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ is uniquely ergodic.*

Proof. Take the substitution $\overline{\sigma}$ defined as follows on the alphabet $\mathbb{Z}/2\mathbb{Z} \times A$, whose letters are denoted a_0, a_1, b_0, b_1 for conciseness,

$$\overline{\sigma}\colon \begin{array}{ll} a_0 \mapsto a_0b_1a_0a_1b_1, & a_1 \mapsto a_1b_0a_0a_1b_0, \\ b_0 \mapsto a_0b_1a_1, & b_1 \mapsto a_1b_0a_0. \end{array}$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

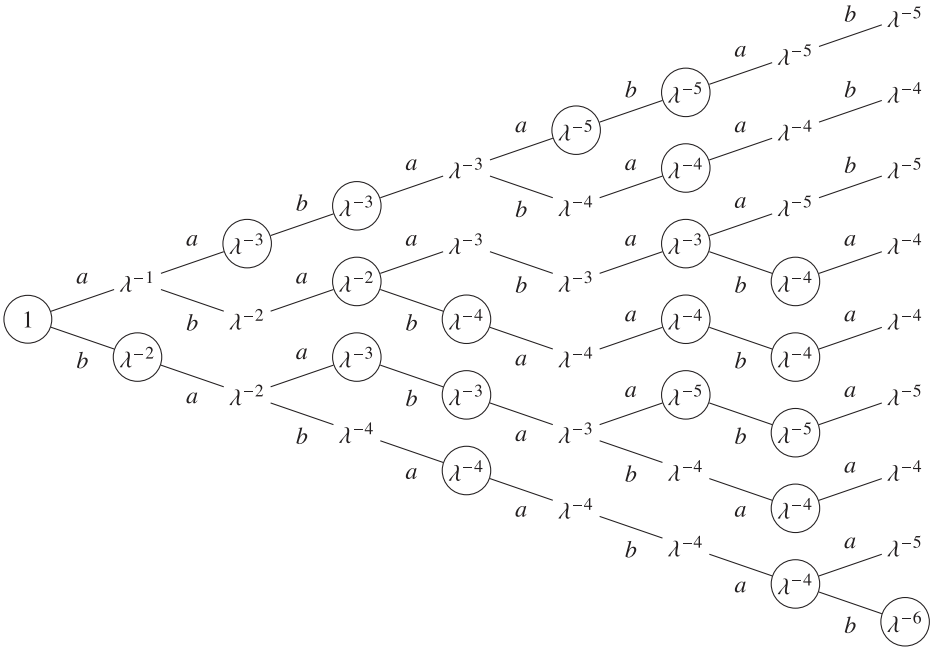


FIGURE 3. The invariant probability measure on the Fibonacci shift ($\lambda =$ the golden ratio). Circled nodes represent elements from the language $L = \{w \in \{a, b\}^* : |w|_a \equiv 0 \bmod 2\}$.

This substitution is primitive and satisfies $\pi \circ \overline{\sigma} = \sigma^3$, where $\pi : (\mathbb{Z}/2\mathbb{Z} \times A)^* \to A^*$ is the natural projection (mapping a_i to a and b_i to b). Moreover, $\varphi \circ \sigma^3 = \varphi$ and $\overline{\sigma}(\mathcal{L}(Y)) \subseteq \mathcal{L}(Y)$, where $Y = \Psi(G \rtimes_\varphi X)$ is the skew product viewed as a shift space on $\mathbb{Z}/2\mathbb{Z} \times A$ via the map Ψ of Lemma 3.1. It follows that Y is the shift space generated by $\overline{\sigma}$. Therefore, by Michel's theorem, Y is uniquely ergodic; hence, so is $G \rtimes_\varphi X$. □

As an immediate consequence of Corollary 3.6 (see also Example 3.11), we conclude that the density $\delta_\mu(L)$, where $L = \{w \in \{a, b\}^* : |w|_a \equiv 0 \bmod 2\}$, exists and is $1/2$. In other words, the sequence $(\mu([L \cap A^n]))_{n\in\mathbb{N}}$ converges to $1/2$ in Cesàro's sense, even though, as we next show, the sequence itself does not converge (this is in contrast with Example 5.11, where $\nu \times \mu$ is mixing). The rest of the section is devoted to the proof of the following result, where $F = (F(n))_{n\in\mathbb{N}}$ is the Fibonacci number sequence (starting with $F(0) = 0$, $F(1) = 1$).

PROPOSITION 3.18. *Let $L = \{w \in \{a, b\}^* : |w|_a \equiv 0 \bmod 2\}$. Then, the sequence of measures $(\mu([L \cap A^n]))_{n\in\mathbb{N}}$ does not have a limit, as*

$$\lim_{n\to\infty} \mu([L \cap A^{F(4n)}]) = 1, \quad \lim_{n\to\infty} \mu([L \cap A^{F(4n+2)}]) = 0.$$

The proof relies on a number of key facts about the structure of the language of the Fibonacci shift. An important tool is the notion of *special word*. Recall that, for a general shift space $X \subseteq A^{\mathbb{Z}}$, a word $w \in \mathcal{L}(X)$ is called *left special* if $aw, bw \in \mathcal{L}(X)$ for some $a, b \in A$, $a \neq b$; it is called *right special* if instead $wa, wb \in \mathcal{L}(X)$ for some $a, b \in A$,

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

$a \neq b$; and it is called *bispecial* if it is both left and right special. Since it is Sturmian, the Fibonacci shift has the property that for every $n \in \mathbb{N}$, $\mathcal{L}(X) \cap A^n$ contains *exactly* one left and one right special factor (which might or might not coincide). We next establish two lemmas which are central in the proof of Proposition 3.18.

LEMMA 3.19. *Let u be a left special factor and v be a right special factor in the Fibonacci shift X.*

(i) *The word $u' = \sigma^2(ub)$ is also left special.*

(ii) *The word $v' = \sigma^2(va)$ is also right special.*

Proof. (i) As u is left special, then au and bu are in $\mathcal{L}(X)$. Let c and d be right extensions of respectively au and bu; let c' and d' be such that $\sigma^2(c) = abc'$ and $\sigma^2(d) = abd'$ (c' and d' are either a or ε). Then, the following words also belong to $\mathcal{L}(X)$:

$$\sigma^2(auc) = aba\sigma^2(u)abc' = abau'c',$$
$$\sigma^2(bud) = ab\sigma^2(u)abd' = abu'd'.$$

In particular, au' and bu' belong to $\mathcal{L}(X)$.

(ii) The proof of the second part follows similar lines. Let c and d be right extensions of respectively va, vb (in fact, $d = a$ since bb does not occur in X); let c' be such that $\sigma^2(c) = ac'$. Then, we find that the following words belong to $\mathcal{L}(X)$:

$$\sigma^2(vac) = v'ac', \quad \sigma^2(vba) = \sigma^2(v)ababa = v'ba.$$

Hence, both $v'a, v'b \in \mathcal{L}(X)$. □

LEMMA 3.20. *Let u_n and v_n denote respectively the left and right special factors of length n in the Fibonacci shift X. Then, whenever $n = F(2k+2) - 1$ for some $k \geq 0$, the equality $bu_n = v_n b$ holds and, moreover,*

$$\varphi(u_n) = \varphi(v_n) \equiv k \mod 2,$$

where $\varphi\colon \{a, b\}^* \to \mathbb{Z}/2\mathbb{Z}$ is the morphism defined by $\varphi(a) = 1$ and $\varphi(b) = 0$.

Proof. Let us first prove that $bu_n = v_n b$ when $n = F(2k+2) - 1$, $k \geq 0$. Consider the following recursively defined sequence of words:

$$w_0 = \varepsilon, \quad w_{k+1} = \sigma^2(w_k b).$$

In other words, this is the sequence of words starting with

$$\varepsilon, \quad \sigma^2(b), \quad \sigma^4(b)\sigma^2(b), \quad \sigma^6(b)\sigma^4(b)\sigma^2(b), \quad \ldots.$$

It is clear from Lemma 3.19 that this is a sequence of left special factors of X. We claim that $|w_k| = F(2k+2) - 1$ for every $k \geq 0$. Indeed, observe that, for every $i \in \mathbb{N}$, $|\sigma^{2i}(b)| = F(2i+1)$ (a fact easily established by induction) and thus,

$$1 + |w_k| = 1 + \sum_{i=1}^{k} |\sigma^{2i}(b)| = \sum_{i=0}^{k} F(2i+1) = F(2k+2).$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Thus, we are reduced to show that $bw_k = v_n b$, where $n = F(2k+2) - 1$; or in other words that removing the last letter from bw_k yields a right special factor. We do so by induction on k. The basis $k = 0$ is trivial since $w_0 = \varepsilon$. Assume that the equality $bw_k = v_n b$ holds for some $k \geq 0$. Then, we have

$$abw_{k+1}a = ab\sigma^2(w_k b)a = \sigma^2(bw_k b)a = \sigma^2(v_n bb)a = \sigma^2(v_n a)ba.$$

By Lemma 3.19, the word $\sigma^2(v_n a)$ is right special and, thus, so is the word $a^{-1}\sigma^2(v_n a)$ obtained by removing its leading letter a. Finally, we observe that $a^{-1}\sigma^2(v_n a)b = bw_{k+1}$ and, thus, $bw_{k+1} = v_m b$, where $m = F(2k+4) - 1$.

Next, let us fix $k \geq 0$ and let $n = F(2k+2) - 1$. Then, by the first part of the proof, we know that $w_k = u_n$, $bu_n = v_n b$, and then

$$\varphi(u_n) = \varphi(bu_n) = \varphi(v_n b) = \varphi(v_n).$$

We need to show that $\varphi(w_k) \equiv k \bmod 2$. We do so by induction on k. Since $w_0 = \varepsilon$ and $w_1 = \sigma^2(b) = ab$, the result is clear for $k = 0, 1$. Assume then that the result holds for some $k \geq 0$. By definition, $w_{k+2} = \sigma^4(w_k b)\sigma^2(b) = \sigma^4(w_k b)ab$. However, observe that

$$\sigma^4(a) = abaababa \quad \text{and} \quad \sigma^4(b) = abaab.$$

Thus, we have the relation $\varphi(\sigma^4(x)) \equiv |x| \bmod 2$ for all $x \in A^*$. By induction, we then have

$$\varphi(w_{k+2}) = \varphi(\sigma^4(w_k b)ab) = \varphi(\sigma^4(w_k b)) + \varphi(ab) \equiv |w_k b| + 1 \equiv |w_k| + 2 \equiv k + 2 \bmod 2,$$

which concludes the proof. □

Proof of Proposition 3.18. Recall that $L = \{w \in \{a, b\}^* : |w|_a \equiv 0 \bmod 2\} = \varphi^{-1}(0)$, where $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$ is the morphism defined by $\varphi(a) = 1$, $\varphi(b) = 0$. Let $\leq_{\text{lex}}$ denote the lexicographic order on words. It follows immediately from the description of $\leq_{\text{lex}}$ on $\mathcal{L}(X) \cap A^n$ by Perrin and Restivo [47, Theorem 2] that for every $w \in \mathcal{L}(X)$ with $|w| = n + 1$,

$$\varphi(w) = \begin{cases} \varphi(v_n a) & \text{if } w \leq_{\text{lex}} v_n a, \\ \varphi(v_n b) & \text{if } w \geq_{\text{lex}} v_n b. \end{cases}$$

Moreover, [47, Proposition 2] states that the lexicographically maximal element of $\mathcal{L}(X) \cap A^{n+1}$ is bu_n. In particular, if $bu_n = v_n b$, then $v_n b$ is maximal in $\mathcal{L}(X) \cap A^{n+1}$ and for such values of n,

$$[L \cap A^{n+1}] = \begin{cases} [v_n b] & \text{if } \varphi(v_n b) = 0, \\ X \setminus [v_n b] & \text{if } \varphi(v_n b) = 1. \end{cases}$$

However, by Lemma 3.20, we know that $\varphi(v_n b) = \varphi(v_n) \equiv k \bmod 2$, where $n = F(2k+2) - 1$, and

$$\mu([L \cap A^{n+1}]) = \begin{cases} \mu([v_n b]) & \text{if } k \text{ is even}, \\ 1 - \mu([v_n b]) & \text{if } k \text{ is odd}. \end{cases}$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Finally, k is even precisely when $n+1 = F(4\ell+2)$ for some ℓ and k is odd precisely when $n+1 = F(4\ell)$ for some ℓ. We conclude the proof by observing that $\lim_{n\to\infty} \mu([v_n]) = 0$, which is a straightforward consequence of [24, Proposition 13]. □

4. *The case of Bernoulli measures*

This section gives a brief account of the original approach to densities under Bernoulli measures due to Schützenberger [53], Berstel [6], and Hansel and Perrin [32] with an approach based on the algebraic theory of formal languages. Details may be found in the monograph [8].

A probability measure μ on the full shift $A^{\mathbb{Z}}$ is a *Bernoulli measure* when the map $A^* \to [0, 1]$ induced by evaluating μ on cylinders is a morphism. More explicitly,

$$\mu([uv]) = \mu([u])\mu([v]) \quad \text{and} \quad \sum_{a\in A} \mu(a) = 1.$$

A Bernoulli measure is called *positive* if $\mu([a]) > 0$ for all $a \in A$. The simplest example of a positive Bernoulli measure is that of the uniform distribution given by $\mu([a]) = 1/|A|$.

A first result on densities under Bernoulli measures is the following, which generalizes the simple observation that finite languages have zero density. We say that a language $L \subseteq A^*$ is *thin* if there is a word $w \in A^*$ that does not appear in any word of L as a factor.

PROPOSITION 4.1. [8, Proposition 13.2.3] *Let μ be a Bernoulli measure on $A^{\mathbb{Z}}$ and $L \subseteq A^*$. If L is thin, then one has $\delta_\mu(L) = 0$.*

Recall that a *rational language* is a language recognized by a finite automaton, which may, without loss of generality, be assumed to be complete. Formally, this means that there is a triple $\mathcal{A} = (Q, i, F)$, where Q is a finite set of states on which A^* acts on the right, $i \in Q$ is the initial state, and $F \subseteq Q$ is the set of final states, and such that $L = \{w \in A^* \mid i \cdot w \in F\}$.

Example 4.2. The language $L = \{aa, ab, b\}^*$ is rational. An automaton recognizing L is shown in Figure 4; it is, in fact, the *minimal* automaton of L.

For rational languages, we have the following converse of Proposition 4.1. We provide a sketch of proof using notions from the algebraic theory of formal languages, which may be recalled in [8].

PROPOSITION 4.3. *Let μ be a positive Bernoulli measure on $A^{\mathbb{Z}}$ and $L \subseteq A^*$ be a rational language. If $\delta_\mu(L) = 0$, then L is thin.*

Proof. Let $\varphi\colon A^* \to M$ be a morphism onto a finite monoid M such that $L = \varphi^{-1}(\varphi(L))$ and let K be the minimal ideal of M. For $q \in M$, we denote by L_q the language $\varphi^{-1}(q)$. Observe that every element $m \in K$ satisfies $MmM = K$ and, in particular, L_q is thin whenever $q \in M \setminus K$. As a result, we deduce that $\delta_\mu(L) = \delta_\mu(L \cap L_K)$ and $\delta_\mu(L_K) = 1$, where $L_K = \varphi^{-1}(K)$.

We claim that $\delta_\mu(L_m) > 0$ for all $m \in K$. First, observe that there must be at least one $n \in K$ with $\delta_\mu(L_n) > 0$. Next, fix $m \in K$ and let $r, s \in M$ such that $rns = m$; let

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

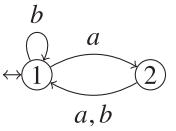


FIGURE 4. An automaton recognizing the language $L = \{aa, ab, b\}^*$ with the state 1 being the initial state and the only final state.

$u, v \in A^*$ such that $\varphi(u) = r$ and $\varphi(v) = s$. It follows that $uL_nv \subseteq L_m$. Then, the fact that μ is a Bernoulli measure implies that

$$\delta_\mu(uL_nv) = \mu([u])\mu([v])\delta_\mu(L_n).$$

Since μ is positive, we get $\mu([L_m]) \geq \mu([u])\mu([v])\delta_\mu(L_n) > 0$. It now follows that if $\delta_\mu(L) = 0$, then L must be a union of L_q with $q \in M \setminus K$. This concludes the proof since finite unions of thin codes are thin, cf. [8, Proposition 2.5.8]. □

Note however that the previous proposition fails for non-rational languages: take, for instance,

$$L = \{a^n w \mid n \geq 0, w \in \{a, b\}^n\}.$$

Then, one has $\mu([L \cap A^{2n}]) = 1/\mu([a])^n$, which implies $\delta_\mu(L) = 0$. However, L is clearly not thin, since every word appears as a suffix.

Next, we give a proof of the fact that the density of a rational language under a Bernoulli measure always exists.

PROPOSITION 4.4. *Let μ be a Bernoulli measure on $A^{\mathbb{Z}}$ and $L \subseteq A^*$ be a rational language. Then, the density $\delta_\mu(L)$ exists and if $\mu([A]) \subseteq \mathbb{Q}$, it is a rational number.*

Proof. Let $\mathcal{A} = (Q, i, F)$ be a complete deterministic automaton recognizing L. For every terminal state $t \in F$, let L_t be the language recognized by (Q, i, t). Since $L = \bigcup_{t \in F} L_t$, it is enough to prove that every L_t has a density. Thus, we may assume that $F = \{t\}$. Let M be the $Q \times Q$-matrix defined by

$$M_{p,q} = \mu(\{x \in A^{\mathbb{Z}} \mid p \cdot x_0 = q\}).$$

Observe that M is a stochastic matrix and since the sets $\{x \in A^{\mathbb{Z}} \mid p \cdot x_0 = q\}$ are disjoint unions of cylinders $[a]$ for $a \in A$, it has rational entries under our hypothesis. Moreover, this matrix has the property that

$$\mu([L \cap A^n]) = M^n_{i,t}.$$

Without loss of generality, we may assume that every $q \in Q$ is on a path from i to t (otherwise, q can be removed). Let U be the set of labels of paths $i \to t$ and which pass by t exactly once; let V be the set of labels of simple loops $t \to t$. Then, $L = UV^*$, $0 \leq \mu([U]) < \infty$, and $\delta_\mu(UV^*) = \mu([U])\delta_\mu(V^*)$ exists whenever $\delta_\mu(V^*)$ exists [8, Proposition 13.2.5]. Hence, we may assume moving forward that there is a path from t to i. In particular, it means that the matrix M is *irreducible*, that is, for all p, q in Q, there exists an integer k such that $M^k_{p,q} > 0$ or else $(M + I)^{|Q|-1} > 0$, where I is the

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

$Q \times Q$ identity matrix. By the Perron–Frobenius theorem, the numbers $M_{i,t}^n$ converge in average and, thus, the density of L exists. Moreover, the matrices M^n converge as $n \to \infty$ to a matrix with all rows equal to a stochastic eigenvector v of M of eigenvalue 1. However, since M has rational entries, the eigenspace of eigenvalue 1 (which has dimension one by the Perron–Frobenius theorem) has a spanning vector u with only rational entries. Therefore, the aforementioned vector v is equal to $\pm u/\|u\|_1$, where $\|u\|_1$ denotes the 1-norm of u and, thus, also has rational entries. □

Example 4.5. Let L and $\mathcal{A}$ be the language and automaton from Example 4.2. Set $\mu([a]) = p$ and $\mu([b]) = 1 - p$. The matrix M, which is irreducible, is

$$M = \begin{bmatrix} 1-p & p \\ 1 & 0 \end{bmatrix}.$$

The normalized left eigenvector is

$$v = \begin{bmatrix} 1/(1+p) & p/(1+p) \end{bmatrix}.$$

Thus, $\delta_\mu(L) = 1/(1+p)$ (by considering here $M_{1,1}$ since $i = t = 1$).

Remark 4.6. The proof of the above theorem uses a Markov chain associated to an automaton. It can also be understood as a skew product $Q \rtimes X$ of the set Q with the shift $X = A^{\mathbb{Z}}$, in the same way as we did with groups. Indeed, if $\mathcal{A} = (Q, i, F)$ is a deterministic automaton, then the set $Q \times X$ may be turned into a dynamical system using the transformation

$$T(q, x) = (q \cdot x_0, Sx).$$

Let ρ be the unique probability measure on Q such that

$$\rho(q) = \sum_{p \cdot a = q} \rho(p)\mu([a]).$$

Then, $\rho \times \mu$ is an invariant probability measure on the skew product $Q \rtimes X$, which is ergodic whenever the matrix of the automaton is irreducible. Moreover, it is mixing as soon as $\mathcal{A}$ is aperiodic, that is, when the gcd of the lengths of cycles in $\mathcal{A}$ is 1; or equivalently, the matrix M is primitive. In this case, we recover the conclusion of Remark 3.5. Also, by ergodicity of Bernoulli measures, the existence of the density in Proposition 4.4 is a special case of Theorem 1.1.

We end with a discussion on prefix codes, a notion which also appears later in §6. A *prefix code* is a subset $U \subseteq A^*$, where no word is a strict prefix of another. Any prefix code must satisfy $\mu([U]) \leq 1$ for any Bernoulli measure μ, cf. [**8**, Proposition 3.7.1].

If U is a prefix code such that $\mu([U]) = 1$, then its *average length* (relatively to μ) is defined as

$$\ell(U) = \sum_{p \in U} |p|\, \mu([p]).$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

It is well known that the average length satisfies

$$\ell(U) = \mu([P]), \tag{4.1}$$

where P is the set of proper prefixes of the words in U [8, Proposition 3.7.11]. The following is closely related to our second main result, Theorem 1.2. Its proof is based on relations between associated generating functions.

PROPOSITION 4.7. [8, Theorem 13.2.11] *Let μ be a positive Bernoulli measure on $A^{\mathbb{Z}}$. Let U be a prefix code such that $\mu([U]) = 1$ and $\ell(U) < \infty$. Then, $\delta_\mu(U^*) = 1/\ell(U)$.*

Note that a *rational* prefix code satisfying $\mu([U]) = 1$ must satisfy the assumption $\ell(U) < \infty$. Indeed, any rational code is thin by [8, Proposition 2.5.20]; hence, the set P of proper prefixes of U is also thin, which implies $\mu([P]) < \infty$ by [8, Proposition 2.5.12].

Let $\varphi\colon A^* \to G$ be a morphism from A^* onto a finite group G. If H is a subgroup of G, then the submonoid $M = \varphi^{-1}(H)$ is generated by the prefix code U consisting of the non-empty words in M with no non-trivial prefix in M; we say that U is a *group code*. Observe in fact that U is also equal to the non-empty words in M with no non-trivial suffix in M and, as a result, has the dual property of being a *suffix code* (no element of U is the suffix of another). Note that sets which are both prefix and suffix codes are called *bifix codes*; bifix codes are at the heart of §6.1.

PROPOSITION 4.8. *Let μ be a positive Bernoulli measure on $A^{\mathbb{Z}}$. Let $\varphi\colon A^* \to G$ be a morphism from A^* onto a finite group G and let H be a subgroup of G. Let U be the group code defined by $\varphi\colon A^* \to G$ and $U^* = \varphi^{-1}(H)$. Then, $\ell(U) = d$ and $\delta_\mu(U^*) = 1/d$ with $d = [G : H] = |H|/|G|$.*

Proof. The proof that $\ell(U) = d$ uses the fact that the set P of proper prefixes of U is a disjoint union of d suffix codes U_i such that $\mu([U_i]) = 1$. Then, using (4.1), we obtain $\ell(U) = \sum_{i=1}^{d} \mu([U_i]) = d$. For further details, see [8, Corollary 6.13.16, Theorem 13.2.9], noting that the code U is a maximal bifix code [8, pp. 64, 65], which is moreover rational and, thus, also thin [8, Proposition 2.5.20]. □

Example 4.9. Let $A = \{a, b\}$. Set $p = \mu([a]), q = \mu([b])$. Let $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$ be defined by $\varphi(a) = 0$, $\varphi(b) = 1$. Then, $\varphi^{-1}(1) = U^*$ with $U = \{a\} \cup ba^*b$. The set P of proper prefixes of U is $P = \{\varepsilon\} \cup ba^*$ and we have $\ell(U) = \mu([\{\varepsilon\} \cup ba^*]) = 1 + q/(1-p) = 2$.

5. *Densities in SFTs*

The aim of this section is to apply our main density formula (Theorem 1.1) within the setting of *SFTs*. We provide a simple condition which guarantees topological transitivity of the skew product with a finite group; this, in turn, implies ergodicity for the product of the uniform probability measure on the group and a Markov measure on the shift. This also includes the case of Bernoulli measures, which we already discussed in §4.

The question of ergodicity for skew products over Bernoulli measures was studied by Kakutani [35] and the more general Markov case was studied by Bufetov [18]. The latter introduced a condition that he called 'strongly connected' to characterize ergodicity of certain skew products involving Markov measures [18, Theorem 4]. Restating Bufetov's

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

criterion in our setting, for consistency of terminology, we rename this condition *strong irreducibility*; his result was also recently extended in a paper by Lummerzheim *et al* [**40**, Theorem 4.3].

In §5.1, we define φ-irreducibility of a subshift with respect to a morphism onto a finite group, which characterizes topological transitivity in skew products. Section 5.2 treats skew products of shift spaces over Markov measures, using the fact that irreducibility implies ergodicity. In §5.3, we discuss strong irreducibility, which provides topological transitivity simultaneously for all skew products.

5.1. *SFTs and φ-irreducibility.* Recall that a shift X is an r-step SFT ($r \geq 1$) if there is a list $F \subseteq A^{r+1}$ of *forbidden factors* of length $r+1$, with the property that an infinite word $x \in A^{\mathbb{Z}}$ belongs to X precisely when none of its factors of length $r+1$ are in F. Recall also that a shift X is *topologically transitive* if for every pair (U, V) of non-empty open sets in X, there is $n > 0$ for which $S^n U \cap V \neq \varnothing$. This is equivalent with the *irreducibility* of X, that is, for every $u, v \in \mathcal{L}(X)$, there is $w \in A^*$ such that $uwv \in \mathcal{L}(X)$. For more on SFTs, see e.g. [**39**, **48**]. Topological transitivity is used to prove ergodicity for r-step Markov measures fully supported on r-step SFTs in §5.2.

Definition 5.1. Let X be a shift on A, G a finite group, and $\varphi\colon A^* \to G$ be a morphism onto G. We say that X is *φ-irreducible* if, for all $u, v \in \mathcal{L}(X)$, there exists $w \in A^*$ such that $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = 1_G$.

Clearly, φ-irreducibility always implies irreducibility. This is a special case of the following remark.

Remark 5.2. For a morphism $\varphi\colon A^* \to G$ as above, let

$$\ker(\varphi) = \{(u, v) \in A^* \times A^* \mid \varphi(u) = \varphi(v)\}.$$

Observe that if $\psi\colon A^* \to H$ is another morphism onto a finite group H such that $\ker(\varphi) \subseteq \ker(\psi)$, then there exists a unique morphism $\beta\colon G \to H$ such that $\beta \circ \varphi = \psi$ and, as a result, φ-irreducibility implies ψ-irreducibility.

We prove the following result, which shows that φ-irreducibility is precisely the notion needed for topological transitivity of the original shift to propagate to skew products with finite groups.

THEOREM 5.3. *Let X be an r-step SFT on A and let $\varphi\colon A^* \to G$ be a morphism onto a finite group. Then, X is φ-irreducible if and only if the skew product $G \rtimes_\varphi X$ is topologically transitive.*

Before doing so, we establish an intermediate result involving the following notion.

Definition 5.4. We say that X is *fiber ergodic* with respect to φ if, for every $g, h \in G$, there exists $w \in \mathcal{L}(X)$ such that $g\varphi(w) = h$; or, equivalently, the restriction of φ to $\mathcal{L}(X)$ is onto.

Fiber ergodicity follows from φ-irreducibility for SFTs, as shown next.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

LEMMA 5.5. *Let X be a SFT on A and let $\varphi\colon A^* \to G$ be a morphism onto a finite group. If X is φ-irreducible, then it is fiber ergodic with respect to φ.*

Proof. Let $r \geq 1$ such that X is an r-step SFT. Take $g \in G$. Since φ is onto, we may find letters $a_1, \ldots, a_k \in A$ such that $\varphi(a_1 \cdots a_k) = g$. For $i = 1, \ldots, k$, let u_i be a word of length $r+1$ in $\mathcal{L}(X)$ starting with a_i. Let t_i and v_i be the suffix and prefix of length r of u_i. By assumption, there exists, for each $i = 1, \ldots k-1$, a word w_i such that $t_i w_i v_{i+1} \in \mathcal{L}(X)$ and $\varphi(t_i w_i) = 1_G$. Since X is an r-step shift, it follows that the word $z = u_1 w_1 \cdots w_{k-1} u_k w_k$ belongs to $\mathcal{L}(X)$ and

$$\varphi(z) = \varphi(a_1)\varphi(t_1 w_1) \cdots \varphi(a_k)\varphi(t_k w_k) = \varphi(a_1 \cdots a_k) = g,$$

which proves that X is fiber ergodic. □

Proof of Theorem 5.3. For $g \in G$, define a relation $\prec_g$ on $\mathcal{L}(X)$ by $u \prec_g v$ if there exists $w \in \mathcal{L}(X)$ such that $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = g$. Observe that $u \prec_g v$ precisely when $T^m(\{1_G\} \times [u]_X)$ intersects $\{g\} \times [v]_X$ for some $m \geq |u|$. Therefore, $G \rtimes_\varphi X$ is topologically transitive precisely when all relations $\prec_g$, $g \in G$, are total. In particular, whenever this is the case, $\prec_{1_G}$ must contain all pairs $u, v \in \mathcal{L}(X)$, which is precisely the definition of φ-irreducibility. Thus, topological transitivity of $G \rtimes_\varphi X$ implies φ-irreducibility of X. It remains to prove the converse.

Assume that X is φ-irreducible, so that $\prec = \prec_{1_G}$ contains all pairs of words in $\mathcal{L}(X)$. Take $u, v \in \mathcal{L}(X)$ and $g \in G$; we need to show that $u \prec_g v$.

By fiber ergodicity (which holds thanks to Lemma 5.5), there is $u' \in \mathcal{L}(X)$ such that $\varphi(u') = g$. Since $u \prec u'$, there is $z \in \mathcal{L}(X)$ such that $uzu' \in \mathcal{L}(X)$ and $\varphi(uz) = 1_G$. Then, $w_0 = zu'$ satisfies $uw_0 \in \mathcal{L}(X)$ and $\varphi(uw_0) = g$.

Extend uw_0 to a word $wu_0v' \in \mathcal{L}(X)$ with $|v'| \geq r+1$. Since $v' \prec v$, there is a word v'' such that $v'v''v \in \mathcal{L}(X)$ and $\varphi(v'v'') = 1_G$. Then, all subwords of length $r+1$ of $uw_0v'v''v$ are in $\mathcal{L}(X)$ and, hence, $uw_0v'v''v \in \mathcal{L}(X)$. Finally, letting $w = w_0v'v''$, we have $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = g$. This shows that $u \prec_g v$, which concludes the proof. □

It is not hard to see directly that topological transitivity of $G \rtimes_\varphi X$ implies fiber ergodicity (as also follows from Theorem 5.3); but the converse is false, as shown by the following example.

Example 5.6. Take again the skew product from Example 3.3, based on the three-element shift generated by the periodic infinite word $(abc)^\infty$ taken with respect to the morphism $\varphi\colon \{a, b, c\}^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = \varphi(b) = 1$, $\varphi(c) = 0$. Then, X is fiber ergodic, but $G \rtimes_\varphi X$ is not topologically transitive.

Finally, we prove the following useful property of φ-irreducibility in SFTs. It shows that for SFTs, the task of verifying the condition in Definition 5.1 can be reduced to a finite set of words.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

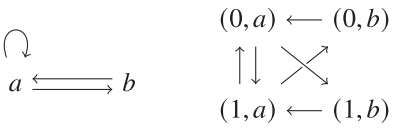


FIGURE 5. The SFT and skew product from Example 5.9. The arrows represent the two respective dynamics, namely the shift map and the skew product map.

PROPOSITION 5.7. *Let X be an r-step SFT on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group. Then, for X to be φ-irreducible, it suffices that, for all $u, v \in \mathcal{L}(X)$ with $|u| = |v| = r$, there exists $w \in A^*$ such that $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = 1_G$.*

Proof. Let $u, v \in \mathcal{L}(X)$. If $|v| < r$, then we may simply replace v by one of its extensions in $\mathcal{L}(X)$ of length r. If $|v| > r$, then we may likewise replace it by its prefix of length r. Thus, we may assume moving forward that $|v| = r$.

First, we assume that $|u| > r$. Let $u = pq = p'q'$, where $|q| = |p'| = r$. By assumption, there are words w, w' such that $qw'p', qwv \in \mathcal{L}(X)$ and $\varphi(qw) = \varphi(qw') = 1_G$. Observe that, for every $n \in \mathbb{N}$, $u(w'u)^n wv$ has all of its factors of length r in $\mathcal{L}(X)$. Thus, the word $z_n = (w'u)^n w$ is such that $uz_n v \in \mathcal{L}(X)$, while $\varphi(uz_n) = \varphi(p)^n$. Taking $n = |G|$, we get $\varphi(uz_n) = 1_G$, as needed.

It remains to handle the case where $|u| < r$. Take a word p such that $|p| = r - |u|$ and $pu \in \mathcal{L}(X)$. Then, we may find words w, w' such that $puw'uwv \in \mathcal{L}(X)$ with $\varphi(puw') = 1_G$ and $\varphi(puw'uw) = 1_G$; thus, $\varphi(uw) = 1_G$. □

Example 5.8. Consider once again the three-element shift X from Example 3.3 generated by the periodic word $(abc)^\infty$, which is an irreducible 1-step SFT. Let $\varphi\colon \{a, b, c\}^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = \varphi(b) = 1$ and $\varphi(c) = 0$.

Observe that every word w such that $awb \in \mathcal{L}(X)$ is of the form $(bca)^n$ for some $n \geq 0$. In particular, it follows that for every such word w, $\varphi(aw) = 1$; thus, X is not φ-irreducible (though note that it is fiber ergodic). It is however φ-irreducible if φ is similarly defined, but takes values instead in $\mathbb{Z}/3\mathbb{Z}$.

Example 5.9. Let X be the *golden mean* shift, that is, the 1-step SFT formed by sequences in $\{a, b\}^{\mathbb{Z}}$ avoiding the factor bb. Take the morphism $\varphi\colon \{a, b\}^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = 1$, $\varphi(b) = 0$. As seen before, this choice of a morphism allows one to count the parity of occurrences of as. Then, X is φ-irreducible, as evidenced by the fact that $aab, aaa, ba, baab \in \mathcal{L}(X)$, according to Proposition 5.7. It is also not hard to verify directly that the skew product viewed as a SFT under the topological conjugacy Ψ from Lemma 3.1 is indeed irreducible. The shift X and the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ are depicted in Figure 5. Note that the same example works with $G = \mathbb{Z}/m\mathbb{Z}$ by considering longer words.

5.2. *Markov measures.* Using the topological transitivity condition from Theorem 5.3, we establish ergodicity for skew products involving *invariant Markov measures*. Recall

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

that a measure μ on $A^{\mathbb{Z}}$ is an r-step Markov measure, $r \geq 1$, when for every word $w \in A^*$ of length $m \geq r$ such that $\mu(x_{[0,m)} = w) \neq 0$ and every letter $a \in A$,

$$\mu(x_m = a \mid x_{[0,m)} = w) = \mu(x_m = a \mid x_{[m-r,m)} = w_{[m-r,m)}).$$

Invariant Markov measures are also called *stationary*. The support of an r-step invariant Markov measure μ is an r-step SFT. We say that an r-step Markov measure μ is *irreducible* when for all $u, v \in A^r$, there exists $m > 0$ such that

$$\mu(x_{[m,m+r)} = v \mid x_{[0,r)} = u) > 0.$$

Note that this is equivalent to irreducibility of the SFT supporting μ. It is well known that a Markov measure is ergodic if and only if it is irreducible. The 1-step case may be found in [**48**, pp. 51–53], while the general r-step case can be reduced to $r = 1$ by passing to the higher block shift (whose definition is recalled after Example 5.11). We now prove our second main result, which we recall.

THEOREM 1.2. *Let μ be an r-step Markov measure on $A^{\mathbb{Z}}$, $r \geq 1$, and let X be the support of μ. Let $\varphi\colon A^* \to G$ be a morphism onto a finite group with uniform probability measure ν. If X is φ-irreducible, then the product measure $\nu \times \mu$ is ergodic on the skew product $G \rtimes_\varphi X$.*

In particular, for every $K \subseteq G$, the language $L = \varphi^{-1}(K)$ has density $\delta_\mu(L) = |K|/|G|$.

Recall that if X is φ-irreducible, then it is irreducible and so, under the assumptions of the above theorem, μ itself must be ergodic. Recall also, from Lemma 3.1, that for every morphism $\varphi\colon A^* \to G$ onto a finite group G and every shift X, the skew product $X \rtimes_\varphi G$ is topologically conjugate to a subshift of $(G \times A)^{\mathbb{Z}}$ via the map

$$\Psi(g, x)_n = (g\varphi^{(n)}(x), x_n), \quad n \in \mathbb{Z}.$$

LEMMA 5.10. *Let X be an r-step SFT, $r \geq 1$, and $\varphi\colon A^* \to G$ be a morphism onto a finite group G with uniform probability measure ν.*

(i) *The set $\Psi(G \rtimes_\varphi X)$ is an r-step SFT.*

(ii) *For every r-step Markov measure μ fully supported on X, the measure $(\nu \times \mu) \circ \Psi^{-1}$ is a Markov measure fully supported on $\Psi(G \rtimes_\varphi X)$.*

Proof. (i) Take a word $\omega \in (G \times A)^*$ with $\omega_i = (g_i, w_i)$ and $m = |\omega| \geq r + 1$. Observe that ω belongs to the language of $\Psi(G \rtimes_\varphi X)$ as long as $w = w_0 \cdots w_{m-1}$ belongs to $\mathcal{L}(X)$ and $g_i\varphi(w_i) = g_{i+1}$. Both of those conditions only need to be verified on factors of length $r + 1$ of ω; hence, $\Psi(G \rtimes_\varphi X)$ is indeed an r-step SFT.

(ii) Fix an r-step Markov measure on X and let $\pi = (\nu \times \mu) \circ \Psi^{-1}$. Fix a word $\omega \in (G \times A)^*$, with $\omega_i = (g_i, w_i)$, $m = |\omega| \geq r$, and $w = w_0 \cdots w_{m-1}$. Assume that ω belongs to the language of $\Psi(G \rtimes_\varphi X)$, which means that $w \in \mathcal{L}(X)$ and $g_{i+1} = g_i\varphi(w_i)$, $i = 0, \ldots, m-2$. Take a letter $\alpha = (g, a) \in G \times A$. In case $g \neq g_{m-1}\varphi(a)$, then it is clear that

$$\pi(\xi_m = \alpha \mid \xi_{[0,m)} = \omega) = \pi(\xi_m = \alpha \mid \xi_{[m-r,m)} = \omega_{[m-r,m)}) = 0.$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Thus, we may suppose from now on that $g = g_{m-1}\varphi(a)$. Then, we have

$$\begin{aligned}\pi(\xi_m = \alpha \mid \xi_{[0,m)} = \omega) &= \frac{\pi(\xi_m = \alpha,\ \xi_{[0,m)} = \omega)}{\pi(\xi_{[0,m)} = \omega)} \\ &= \frac{(\nu \times \mu)(\{g_0\} \times [wa])}{(\nu \times \mu)(\{g_0\} \times [w])} \\ &= \frac{\mu(x_m = a,\ x_{[0,m)} = w_{[0,m)})}{\mu(x_{[0,m)} = w)} \\ &= \mu(x_m = a \mid x_{[0,m)} = w).\end{aligned}$$

Using a similar argument together with the invariance of π, we find

$$\pi(\xi_m = \alpha \mid \xi_{[m-r,m)} = \omega_{[m-r,m)}) = \mu(x_m = a \mid x_{[m-r,m)} = w_{[m-r,m)}).$$

The fact that π is r-step Markov now follows directly from the fact that μ is. □

Proof of Theorem 1.2. In light of Lemma 5.10, the measure $(\nu \times \mu) \circ \Psi^{-1}$ is an r-step Markov measure fully supported on a SFT, which is topologically transitive by Theorem 5.3. Therefore, $(\nu \times \mu) \circ \Psi^{-1}$ must be ergodic and since Ψ is a topological conjugacy, we deduce that the measure $\nu \times \mu$ is ergodic as well.

The last part of the statement is then an immediate consequence of Corollary 3.6. □

Example 5.11. Consider the full shift on $A = \{a, b\}$, equipped with a Bernoulli measure $\mu\colon a \mapsto p$, $b \mapsto 1 - p$, where $0 < p < 1$. Let $\varphi \to \mathbb{Z}/2\mathbb{Z}$ be the morphism defined by $a \mapsto 1$, $b \mapsto 0$, and ν be the uniform measure probability measure on $\mathbb{Z}/2\mathbb{Z}$. Since, clearly, the full shift is φ-irreducible, the product measure $\nu \times \mu$ is ergodic on the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi A^{\mathbb{Z}}$. However, looking directly at the transition diagram of the skew product, we can see that it is not only irreducible, but also aperiodic. Hence, we deduce that the skew product is strongly mixing, and by Remark 3.5,

$$\delta_\mu(L) = \lim_{n\to\infty} \mu([L \cap A^n]) = 1/2.$$

Next, let us briefly describe how ergodicity for r-step Markov measures can also be reduced to the 1-step case by passing to the higher block shift. Let μ be an invariant r-step Markov measure, $r \geq 1$, with support a SFT X. Let $A_X^{[r]} = \mathcal{L}(X) \cap A^r$, which we view as an alphabet. We define the map $\beta_r\colon A^{\mathbb{Z}} \to (A_X^{[r]})^{\mathbb{Z}}$ by

$$\beta_r(x)_i = x_{[i,i+r)}.$$

The image $\beta_r(X)$ forms a shift space on $A_X^{[r]}$ denoted $X^{[r]}$, which is called the *higher block shift*. The image measure of μ under β_r, denoted $\mu^{[r]}$, is an invariant 1-step Markov measure on this shift space. Moreover, μ is ergodic exactly when $\mu^{[r]}$ is.

The density of a group language also carries over to the higher block shift, as follows. Given a morphism $\varphi\colon A^* \to G$ onto a finite group, let $\varphi^{[r]}\colon A_X^{[r]} \to G$ be the morphism defined by $\varphi(w) = \varphi(w_0)$ for $w = w_0 \cdots w_{r-1} \in A_X^{[r]}$. To the group language $L = \varphi^{-1}(K)$, $K \subseteq G$, corresponds the group language $L^{[r]} = (\varphi^{[r]})^{-1}(K)$. It is straightforward to check that $\mu^{[r]}(L^{[r]} \cap (A_X^{[r]})^i) = \mu([L \cap A^i])$ and, as a result,

$$\delta_\mu(L) = \delta_{\mu^{[r]}}(L^{[r]}).$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

5.3. *Strong irreducibility.* We now relate φ-irreducibility to a condition introduced by Bufetov [18], originally under the name *strongly connected*. The term *strictly irreducible* was used by Lummerzheim *et al* [40]. We make a compromise between the two and use the term *strongly irreducible*. We show below that if it holds, then *all* skew products by morphisms onto finite groups are topologically transitive; however, when it fails, such skew products can behave in a variety of manners.

Definition 5.12. Let X be a shift on an alphabet A. For $r \geq 1$, we define the relation $\sim_r$ on $\mathcal{L}(X)$ by $u \sim_r v$ if there exists $w \in A^r$ such that $wu, wv \in \mathcal{L}(X)$. Since $\sim_r$ is symmetric, its transitive closure, which we denote by $\simeq_r$, is an equivalence relation.

An r-step SFT is called *strongly irreducible* if it is irreducible and the relation $\simeq_r$ is total, meaning that it has a single equivalence class.

Remark 5.13. It may happen that $\simeq_r$ has a single equivalence class even in the absence of irreducibility. This is the case in the 1-step SFT X on the alphabet $A = \{a, b\}$ consisting of sequences avoiding the factor ba.

Recall the notation $\prec = \prec_{1_G}$, used in the proof of Theorem 5.3 for the relation defined by $u \prec v$ if and only if there exists $w \in \mathcal{L}(X)$ such that $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = 1_G$.

PROPOSITION 5.14. *In an irreducible r-step SFT, $r \geq 1$, the following implication holds:*

$$u \simeq_r v \implies u \prec v.$$

Proof. Let X be an irreducible r-step SFT. We start by showing that given $u, v \in \mathcal{L}(X)$ with $u \sim_r v$, there is a word w such that $|w| \geq r$, $uwv \in \mathcal{L}(X)$, and $\varphi(uw) = 1_G$. By definition of $\sim_r$, there is a word z such that $|z| = r$ and $zu, zv \in \mathcal{L}(X)$. Let $t \in A^*$ be a word such that $zut \in \mathcal{L}(X)$ and $|ut| \geq r$. Since X is irreducible, there exists a word $w_0 \in A^*$ such that $utw_0zut \in \mathcal{L}(X)$. Then, $(utw_0z)^n u, z(utw_0z)^n v \in \mathcal{L}(X)$ for all $n \geq 1$, since every factor of length r in either of those words is a factor of either zv, zut, or utw_0z, all of which are in $\mathcal{L}(X)$. Taking $n = |G|$, we find $\varphi((utw_0z)^n) = 1_G$. The word $w = (tw_0zu)^{n-1}tw_0z$ then satisfies $uw = (tw_0zu)^n$, and, thus, we get $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = 1_G$, as required.

Suppose next $u = u_0 \sim_r u_1 \sim_r \cdots \sim_r u_{n-1} \sim_r u_n = v$. We want to find a word w such that $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = 1_G$. Using the above claim, there are words w_i such that $u_i w_i u_{i+1} \in \mathcal{L}(X)$, $\varphi(u_i w_i) = 1_G$, and $|w_i| \geq r$. Then, defining $w = w_0 u_1 w_1 \cdots u_{n-1} w_{n-1}$ produces the requisite word such that $uwv \in \mathcal{L}(X)$ and $\varphi(uw) = 1_G$. □

We then deduce the following. It is our version of Bufetov's theorem [18, Theorem 5] and its generalization by Lummerzheim *et al* [40, Theorem 4.3], which we here specialize to the case of skew products with finite groups and morphisms, but generalize to the case of higher step shifts.

THEOREM 5.15. *Let X be an irreducible r-step SFT on A, $r \geq 1$. If X is strongly irreducible, then it is φ-irreducible for every morphism $\varphi\colon A^* \to G$ onto a finite group G. When $r = 1$, the converse holds.*

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Proof. For the first part of the statement, note that when X is strongly irreducible, then by Proposition 5.14, the relation $\prec = \prec_{1_G}$ must be total for every morphism $\varphi\colon A^* \to G$ onto a finite group G, which is precisely the definition of X being φ-irreducible.

We now suppose that $r = 1$. We prove the contrapositive: assuming that X is not strongly irreducible, then it is not φ-irreducible for some φ. The construction presented is essentially the same as that used in the proof of [**40**, Theorem 4.3].

Fix an equivalence class C of $\simeq_1$ restricted to $A \times A$. For each $a \in A$, observe that the set of right extensions of a in X,

$$\mathrm{R}(a) = \{b \in A \mid ab \in \mathcal{L}(X)\},$$

is contained in a class of $\simeq_1$; thus, either $\mathrm{R}(a) \subseteq C$ or $\mathrm{R}(a) \subseteq A \setminus C$. Let

$$B = \{a \in C \mid \mathrm{R}(a) \subseteq A \setminus C\} \cup \{a \in A \setminus C \mid \mathrm{R}(a) \subseteq C\}.$$

Define a morphism $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$ by $\varphi(a) = 1$ if $a \in B$ and $\varphi(a) = 0$ otherwise. Let $a \in C$ and $b \in A$; take a word w such that $awb \in \mathcal{L}(X)$. We claim that $\varphi(aw) = 0$ if and only if $b \in C$. Observe that this claim finishes the proof, as it implies that $a \not\prec b$ whenever $b \in A \setminus C$.

To establish the claim, we argue by induction on $|w|$. We first consider the case $|w| = 0$. We have to check that $\varphi(a) = 0$ if and only if $b \in C$. However, $\varphi(a) = 0$ if and only if $a \notin B$, which is in turn equivalent to $\mathrm{R}(a) \subseteq C$, that is, $b \in C$. For the induction step, suppose that $w = w'c$, $c \in A$, with $aw'b \in \mathcal{L}(X)$, and that the claim holds for w'. First, assume that $\varphi(aw) = 0$. If $c \in C$, then by induction, $\varphi(aw') = 0$ and $\varphi(aw) = \varphi(c) = 0$, and as $b \in \mathrm{R}(c)$, we get $b \in C$. If, however, $c \in A \setminus C$, then $\varphi(aw') = 1$, since the claim holds for w' and $c \in A \setminus C$; hence, $\varphi(aw) = 1 + \varphi(c) = 0$, and so $\varphi(c) = 1$ and $c \notin B$. As $b \in \mathrm{R}(c)$, we again follows that $b \in C$. Conversely, assume that $b \in C$. If $c \in C$, then as $b \in \mathrm{R}(c)$, we get $c \notin B$ and $\varphi(c) = 0$, while by induction, $\varphi(aw') = 0$; hence, $\varphi(w) = 0 + 0 = 0$. Likewise, if $c \in A \setminus C$, then $c \in B$, and $\varphi(aw') = 1 = \varphi(c)$ and $\varphi(aw) = 1 + 1 = 0$. □

At this time, we are unsure whether the converse holds when $r > 1$. Nonetheless, the first part of the above combined with Theorem 5.3 yields the following immediate corollary.

COROLLARY 5.16. *Let X be a strongly irreducible (hence, irreducible) r-step SFT on A, $r \geq 1$. For every morphism $\varphi\colon A^* \to G$, assumed to be onto the finite group G, the skew product $G \rtimes_\varphi X$ is topologically transitive.*

To end this section, we give a few examples of SFTs which are or are not strongly irreducible. However, first we make the following simple observation, similar to Proposition 5.7.

PROPOSITION 5.17. *Let X be an irreducible r-step SFT on A, $r \geq 1$. Then, for X to be strongly irreducible, it suffices that $u \simeq_r v$ for all pairs $u, v \in \mathcal{L}(X)$ with $|u| = |v| = r$.*

Proof. Let u' and v' be words in $\mathcal{L}(X)$ of arbitrary lengths. Choose some words u, v such that $|u| = |v| = r$, u is a prefix of u' or *vice versa*, and v is a prefix of v' or *vice versa*.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

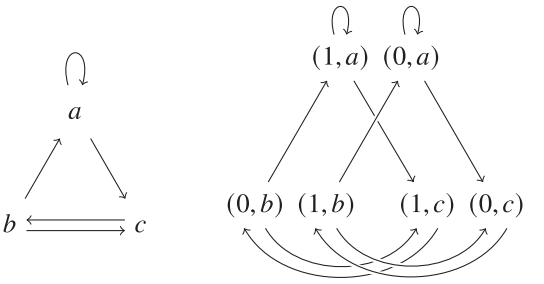


FIGURE 6. The SFT and skew product from Example 5.19.

By assumption, we may find n words $w_0, \ldots, w_{n-1}$ in A^* and $n+1$ words $t_0, \ldots, t_n$, also all of length r, such that $t_0 = u$, $t_n = v$, and

$$w_i t_i,\, w_i t_{i+1} \in \mathcal{L}(X), \quad 0 \le i \le n-1.$$

Then, note that $w_0 u'$ is also in $\mathcal{L}(X)$; indeed, this is obvious when u' is a prefix of u and otherwise, it follows from the fact that X is an r-step SFT. Likewise, $w_{n-1} v' \in \mathcal{L}(X)$, which shows that $u' \simeq_r v'$. □

Example 5.18. Continuing from Example 5.9, we claim that the golden mean shift is, in fact, strongly irreducible. Indeed, it suffices to note that in this irreducible 1-step SFT, for any letters u and v in A, one has au and av in $\mathcal{L}(X)$; hence, $u \sim_1 v$ and, by Proposition 5.17, X is indeed strongly irreducible. As a result, by Corollary 5.16, all skew products of the golden mean shift with skewing functions given by morphisms onto finite groups are topologically transitive, including the one in Example 5.9.

Example 5.19. Let X be the 1-step SFT on $A = \{a, b, c\}$ formed by the sequences avoiding the words ca, ab, bb, cc. Then, X is irreducible, but not strongly irreducible. Indeed, the relation $\simeq_1$ has two classes in $A \times A$, namely $\{b\}$ and $\{a, c\}$. Taking for instance $C = \{b\}$, the set B in the second part of the proof of Theorem 5.15 equals $\{b, c\}$. Consider now the morphism $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = 0$, $\varphi(b) = \varphi(c) = 1$. One has that X is irreducible, whereas $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ is not. If $w \in A^*$ and $bwa \in \mathcal{L}(X)$, we must have $\varphi(bw) = 1$ and, therefore, X is not φ-irreducible. The shift X and the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ are depicted in Figure 6.

In particular, it follows from Theorem 5.3 that no Markov measure fully supported on $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ (viewed as a SFT as per Lemma 5.10) can be ergodic. This includes every measure of the form $\nu \times \mu$, where ν is the uniform probability measure on $\mathbb{Z}/2\mathbb{Z}$ and μ is a Markov measure fully supported on X.

6. *A characterization of minimality via return words*

We now shift our focus away from SFTs and towards *minimal shift spaces*. We start by giving a first characterization of minimality for skew products (Theorem 6.1), stated in terms of return words (§2). Its proof, given in §6.2, strongly relies on the deep links between the skew products under consideration and bifix codes, recalled in §6.1.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

THEOREM 6.1. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. The following conditions are equivalent.*

(i) *For every $n > 0$ and $x \in X$, $\{\varphi(x_{[0,m)}) \mid m \in \mathbb{N}, x_{[m,m+n)} = x_{[0,n)}\} = G$.*
(ii) *The skew product $G \rtimes_\varphi X$ is minimal.*
(iii) *For every $u \in \mathcal{L}(X)$, the restriction of φ to $\mathcal{R}_X(u)^*$ is surjective.*

We provide later a further characterization of minimality in Theorem 7.5, stated in terms of cobounding maps, in the flavor of Anzai's theorem on ergodicity of skew products [2]. We also provide an alternate proof of the equivalence between conditions (ii) and (iii) (Remark 7.9).

Remark 6.2. Condition (i) is reminiscent of the welldoc property (which stands for *well-distributed occurrences*) studied by Balková *et al* [5] in the context of pseudorandom number generators. For convenience, we recall that an infinite word X is said to have the welldoc property when it satisfies the following condition where Ab denotes the Parikh Abelianization map $A^* \to \mathbb{Z}^{|A|}$:

$$\text{for all } k \geq 2, \text{ for all } n \geq 0, \quad \{\mathrm{Ab}(x_{[0,m)}) \bmod k \mid m \geq 0, x_{[m,m+n)} = x_{[0,n)}\} = (\mathbb{Z}/k\mathbb{Z})^{|A|}.$$

It follows from the above theorem that the welldoc property is equivalent to all skew products of the form $(\mathbb{Z}/k\mathbb{Z})^d \rtimes_\varphi X$ being ergodic.

6.1. *More on the theory of codes.* We survey some basic notions from the theory of codes, some of which already appeared in §4. Again, we refer to [8] for additional details.

Recall that a *prefix code* is a subset $U \subseteq A^*$ where no word is a strict prefix of another and, likewise, a *suffix code* is a subset of A^* where no word is a strict suffix of another. Let X be a shift space. A prefix code $U \subseteq \mathcal{L}(X)$ is called *X-complete* if every word $w \in \mathcal{L}(X)$ is either a prefix of a word in U or has a prefix which is an element of U; replacing *prefix* by *suffix* in this definition, we obtain the corresponding notion of *X-complete suffix code*.

Example 6.3. Let X be the Fibonacci shift considered in §3.4. Then, the set $U = \{a, ba\}$ is an X-complete prefix code, though clearly not a suffix code.

We recall that a set U is a *bifix code* if it is both a prefix and a suffix code. It is X-complete if it is both an X-complete prefix code and an X-complete suffix code. When $X = A^{\mathbb{Z}}$, we simply say *complete* instead of *$A^{\mathbb{Z}}$-complete*, be it for prefix, suffix, or bifix codes.

Let $\varphi\colon A^* \to G$ be a morphism onto a finite group G. If $H \leq G$ is a subgroup, then the submonoid $M = \varphi^{-1}(H)$ of A^* is uniquely generated by a bifix code Z (an argument is given in [8, p. 64]), which is also a group code (as introduced in §4). It is, in fact, a complete bifix code.

Let Z be a group code. Let P (respectively S) be the set of proper prefixes (respectively suffixes) of the words in Z. Note that P (respectively S) is also the set of words which have

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

no prefix (respectively suffix) in Z. The *Z-degree* $d(u)$ of a word $u \in A^*$ is then defined as any of the following numbers, which all coincide [8, Proposition 6.1.6]:

(i) the number of suffixes of u that are in P;
(ii) the number of prefixes of u that are in S;
(iii) the number of *Z-parses* of u, that is, the number of triples (s, z, p) such that $u = szp$ with $s \in S$, $z \in Z^*$ and $p \in P$.

It follows from the third definition that for every $u, v, w \in A^*$, the Z-degrees of u and uvw satisfy

$$d(v) \leq d(uvw). \tag{6.1}$$

Indeed, if (s, z, p) is a Z-parse of v, let $us = s'z'$ with $s' \in S$ and $z' \in Z^*$, and $pw = z''p'$ with $z'' \in Z^*$ and $p' \in P$. Then, $(s'z'zz'', p')$ is a Z-parse of uvw which extends (s, z, p). This shows that every parse of v extends to a parse of uvw.

PROPOSITION 6.4. *Let $\varphi\colon A^* \to G$ be a morphism onto a finite group G and H be a subgroup of index d in G. Let Z be the group code such that $\varphi^{-1}(H) = Z^*$ and let S be the set of proper suffixes of elements of Z. For every $p, q \in S$ such that q is a proper prefix of p, $\varphi(p)H \neq \varphi(q)H$. In particular, every word has Z-degree at most d.*

Proof. Let $p = qx$ and assume by contradiction that $\varphi(p)H = \varphi(q)H$. Then, it follows that x is in $\varphi^{-1}(H)$; hence, it is a non-empty word of Z^* and, thus, it has a suffix in Z. Set $x = sr$ with $r \in Z$. Let $z \in Z$ be such that $z = tp$. Then,

$$z = tp = tqx = tqsr,$$

so $r \in Z$ is a proper suffix of $z \in Z$, which contradicts the fact that Z is bifix. □

Let X be a minimal shift space and let $U = Z \cap \mathcal{L}(X)$. The *$X$-degree* of U, denoted $d_X(U)$, is the maximal value of the Z-degrees of all words in $\mathcal{L}(X)$. The following is essentially a reformulation of [7, Theorem 4.2.11]; we include a proof for the convenience of the reader.

THEOREM 6.5. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. Let H be a subgroup of index d in G and Z be the group code such that $\varphi^{-1}(H) = Z^*$. The set $U = Z \cap \mathcal{L}(X)$ is a finite X-complete bifix code of X-degree $d_X(U) \leq d$.*

The proof makes use of the following simple observation.

LEMMA 6.6. *Let v be a word with maximal Z-degree. Then, no word of U can be of the form uvw with u or w non-empty.*

Proof. Assume the contrary. Then, we would have $d(uvw) > d(v)$ because every Z-parse of v extends to a Z-parse of uvw, while the latter has the additional Z-parse $(\varepsilon, uvw, \varepsilon)$. □

Proof of Theorem 6.5. Since Z is a bifix code, the same is true for U. Consider a word $v \in \mathcal{L}(X)$ of maximal Z-degree. By Lemma 6.6, there is no word in U containing v as a

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

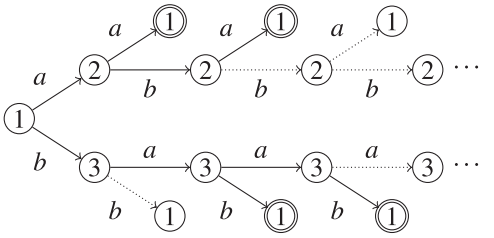


FIGURE 7. Representation of the X-complete bifix code U of Example 6.7. Nodes are labeled by the image of 1 under the permutation given by the label of the path. Elements of U correspond to paths ending in double-circled nodes.

strict factor; thus, U must be finite. Indeed, since X is minimal, every long enough element of $\mathcal{L}(X)$ is of the form uvw with u, w non-empty and, thus, cannot be a factor of a word in U; hence, the length of the words in U is bounded.

Next, we show that U is an X-complete prefix code. Consider a non-empty word $u \in \mathcal{L}(X)$. Since X is minimal, there is a word w such that $uwv \in \mathcal{L}(X)$. Set $u = au'$, where a is a letter. Since $d(au'wv) \geq d(u'wv) \geq d(v)$ and since $d(v)$ is maximal, we have $d(au'wv) = d(u'wv)$. This forces the word $au'wv$ to have a prefix in Z and, hence, in U. Thus, either u has a prefix which is an element of U or it is a prefix of an element of U. Hence, U is an X-complete prefix code. The proof that U is an X-complete suffix code is similar. □

Example 6.7. Let X be the Fibonacci shift on $A = \{a, b\}$ (as in §3.4) and let $\varphi\colon A^* \to S_3$ be the morphism onto the symmetric group S_3 defined by $\varphi(a) = (1\ 2)$, $\varphi(b) = (1\ 3)$, where permutations are written in usual cycle notation. Note that S_3 is isomorphic to the group of matrices $G(2)$ considered in Example 3.16. Let H be the subgroup of G formed of the permutations fixing 1. Let Z be the group code such that $\varphi^{-1}(H) = Z^*$; it is given by $Z = ab^*a \cup ba^*b$. The elements of the infinite set Z are represented in the tree found in Figure 7 as the labels of paths from the root to the leaves. The X-complete bifix code $U = Z \cap \mathcal{L}(X)$, equal to $\{aa, aba, baab, bab\}$, is depicted in Figure 7.

The next result uses ideas found in the proof of the main result of [**12**].

PROPOSITION 6.8. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. Let $H \leq G$ be a subgroup of G, Z be the group code such that $\varphi^{-1}(H) = Z^*$, and $U = Z \cap \mathcal{L}(X)$. Assume that there exists some word $u \in \mathcal{L}(X)$ with maximal Z-degree such that $\varphi(\mathcal{R}_X(u)^*)H = G$. Then, the X-degree of U is $[G : H]$.*

Proof. Let F_A be the free group A and let K be the subgroup of F_A generated by U. Let $u \in \mathcal{L}(X)$ be of maximal X-degree. Let Q be the set of prefixes of u which are suffixes of some element of U (so $|Q|$ is the X-degree of U). Observe that, for distinct elements p and $q \in Q$, the cosets pK and qK are distinct by Proposition 6.4; indeed, since $\varphi(K) \leq H$, Proposition 6.4 implies that φ maps pK and qK inside disjoint right cosets of H in G.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Let us define

$$V = \{v \in F_A \mid vQ \subseteq QK\}.$$

For every $v \in V$, the map $\pi(v)\colon p \mapsto q$ defined by $vp \in qK$ is a permutation. Indeed, let $vp, vp' \in qK$ for some $q \in Q$. Then, $v^{-1}q$ is in $pK \cap p'K$ and, thus, $p = p'$.

We claim that V is a subgroup of F_A. First, let $v \in V$. Then, for any $q \in Q$, since $\pi(v)$ is a permutation of Q, there is a $p \in Q$ such that $vp \in qK$. Then, $v^{-1}q \in pK$. This shows that $v^{-1} \in V$. Next, if v and $w \in V$, then $vwQ \subseteq vQK \subseteq QK$ and, thus, $vw \in V$. Since it is clear that V contains the identity element, this proves the claim.

Next, V contains $\mathcal{R}_X(u)$. Indeed, let $q \in Q$ and $y \in \mathcal{R}_X(u)$. Since q is a prefix of u, yq is a prefix of yu, and since yu is in $\mathcal{L}(X)$ (by definition of $\mathcal{R}_X(u)$), yq is also in $\mathcal{L}(X)$. Since, by Theorem 6.5, U is an X-complete bifix code, it is an X-complete suffix code. This implies that yq is a suffix of a word in U^* and, thus, there is a suffix r of U such that $yq \in rU^*$. We verify that the word r is a suffix of u. Since $y \in \mathcal{R}_X(u)$, there is a word y' such that $yu = uy'$. Consequently, r is a prefix of uy' and, in fact, the word r is a prefix of u. Indeed, one has $|r| \leq |u|$, since otherwise, u would be in the set of internal factors of U and this cannot be the case by Lemma 6.6 (recall that u has maximal Z-degree). Thus, we have $r \in Q$. Since $U^* \subseteq K$ and $r \in Q$, we have $yq \in QK$; hence, $y \in V$.

We finish the proof by showing that $|Q| = [V : K \cap V] \geq [G : H]$. The fact that $|Q| = [V : K \cap V]$ follows by noting that $\pi(v)$ is the identity on Q if and only if v is in K. For the remaining inequality, observe first that, as $K \leq \varphi^{-1}(H)$, the map $V/(K \cap V) \to G/H$, $v(K \cap V) \mapsto \varphi(v)H$ is well defined. Moreover, the fact that V contains $\mathcal{R}_X(u)$ implies that $\varphi(V)H = G$; hence, this map is surjective, thereby showing that $[V : K \cap V] \geq [G : H]$. Since the X-degree of U is at most $[G : H]$ (by Theorem 6.5) and equal to $|Q|$, this concludes the proof. □

The next example shows that there can be elements of maximal degree which fail the condition $\varphi(\mathcal{R}_X(u)^*)H = G$. There are in fact even infinitely many of them.

Example 6.9. Take $A = \{a, b\}$, $\varphi\colon A \to \mathbb{Z}/2\mathbb{Z}$ defined by $\varphi(a) = 1, \varphi(b) = 0$, and X equal to the Thue–Morse shift, considered in Example 2.1. The bifix code $U = \varphi^{-1}(0) \cap \mathcal{L}(X)$ is equal to $\{b, aa, aba, abba\}$. It has X-degree 2, but $\mathcal{R}_X(u)$ is contained in $\varphi^{-1}(0)$ for every u sufficiently long. This last claim can be checked by finding one word u with this property and then using the return preservation property [13]. Nonetheless, the word a has maximal degree and $\mathcal{R}_X(a) = \{a, ab, abb\}$ satisfies $\varphi(\mathcal{R}_X(a)^*) = \mathbb{Z}/2\mathbb{Z}$.

Here is another example with the Thue–Morse shift where $d_X(U) < [G : H]$. Let $\varphi\colon \{a, b\}^* \to \mathbb{Z}/n\mathbb{Z}$ be a morphism with $\varphi(a) = 1 = -\varphi(b)$. Set $Z^* = \varphi^{-1}(0)$ and $U = Z \cap \mathcal{L}(X)$. For every $n \geq 3$, we have $U = \{ab, ba, aabb, bbaa, aababb, bbabaa\}$, while $d_X(U) = 3 < n$ as soon as $n > 3$.

6.2. *Proof of the characterization of minimality.* We are now ready to give the proof of the main theorem of this section.

Proof of Theorem 6.1. (i) implies (ii). Let $Y \subseteq G \rtimes_\varphi X$ be a minimal closed invariant subset. Fix $g \in G$ and $x \in X$; let us show that $(g, x) \in Y$.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

First, note that the projection $P_X(Y) = X$, since X is minimal. Thus, we may find h such that $(h, x) \in Y$. By condition (i), we can find a sequence of positive integers $(m_n)_{n\in\mathbb{N}}$ such that $\varphi(x_{[0,m_n)}) = h^{-1}g$ and $x_{[0,n)} = x_{[m_n,m_n+n)}$. It follows that

$$T_\varphi^{m_n}(h, x) \to (g, x) \quad \text{as } n \to \infty,$$

showing that $(g, x) \in Y$.

(ii) implies (iii). Fix a word $u \in \mathcal{L}(X)$, an element $g \in G$, and consider the two clopen subsets $\{1_G\} \times [u]$ and $\{g\} \times [u]$. By minimality of $G \rtimes_\varphi X$, there exists $n \in \mathbb{Z}$ such that

$$(\{1_G\} \times [u]) \cap T_\varphi^{-n}(\{g\} \times [u]) \neq \varnothing.$$

Choose a point $(1_G, x)$ in that intersection. Let $w = x_{[0,n)}$ if $n \geq 0$, and $w = x_{[n,0)}$ otherwise. Then, we have $T_\varphi^n(1_G, x) = (\varphi(w), S^n x) \in \{g\} \times [u]$ if $n \geq 0$, and $T_\varphi^n(1_G, x) = (\varphi(w)^{-1}, S^n x) \in \{g\} \times [u]$ otherwise. Thus, $w \in \mathcal{R}_X(u)^*$ and $g^{\pm 1} \in \varphi(\mathcal{R}_X(u)^*)$.

(iii) implies (i). Fix $x \in X$, $n \geq 0$ and $g \in G$. We want to prove that we can find $m \geq 0$ with $x_{[0,m)} \in \varphi^{-1}(g)$ and $x_{[0,n)} = x_{[m,m+n)}$.

Let $u = x_{[0,n)}$ and $Y = \mathcal{D}_u(X)$ be the derivative shift of X with respect to u (whose definition may be found in [25, p. 291]). Thus, Y is a shift space on an alphabet $B = B_u$ with a bijection $\theta_u : B \to \mathcal{R}_X(u)$ such that $\theta_u(Y) = X$. Let $y \in Y$ be such that $\theta_u(y) = x$. Consider moreover the morphism $\psi = \varphi \circ \theta_u$, guaranteed to be onto by condition (iii). Let Z be the group code on B such that $Z^* = \psi^{-1}(1_G)$ and $U = Z \cap \mathcal{L}(Y)$. By Theorem 6.5, the set U is a Y-complete bifix code and, thus, it is, in particular, non-empty.

Next, observe that the morphism ψ also satisfies the condition of Proposition 6.8: the restriction of ψ to every $\mathcal{R}_Y(v)^*$ is onto. Indeed, for every $v \in \mathcal{L}(Y)$, the morphism θ_v coding the return words to v satisfies $\theta_u \circ \theta_v = \theta_w$ with $w = \phi_u(v)u$. Thus, we may apply Proposition 6.8 to conclude that the Y-degree of U is equal to $|G|$. Since U is a Y-complete prefix code, y has arbitrary long prefixes in U^*. If such a prefix v is long enough, it has Z-degree equal to the Y-degree of U, that is, $|G|$.

We claim that v has a prefix p such that $\psi(p) = g$. Indeed, since the Z-degree of v is $|G|$, it has $|G|$ prefixes which belong to the set of proper suffixes of the elements of Z. By Lemma 6.6, all these prefixes have distinct images by ψ and, thus, one such prefix p is mapped to g by ψ. Letting $m = |\theta_u(p)|$, we find that $x_{[0,n)} = u = x_{[m,m+n)}$ and, moreover, $\varphi(x_{[0,m)}) = \psi(p) = g$. □

Remark 6.10. We remark that there is also a direct proof that (ii) implies (i), which sheds some light on these equivalences. Assume minimality of the skew product $G \rtimes_\varphi X$ and fix $x \in X$, $n > 0$. Let $w_n = x_{[0,n)}$ and $g \in G$ be given. We need to find a factor v_n such that $x \in [w_n v_n w_n]_X$ and $\varphi(w_n v_n) = g$. By minimality of $G \rtimes_\varphi X$, there is a sequence $(n_j)_{j\in\mathbb{N}}$ such that $T_\varphi^{n_j}(1_G, x) = (\varphi^{(n_j)}(x), S^{n_j}x) \to (g, x)$. Thus, for large enough j, the sequences $S^{n_j}x$ and x agree on arbitrarily long prefixes p_j, and $x \in [p_j q_j p_j]_X$ for some factor q_j such that $|p_j q_j| = n_j$. If j is large enough, then the given prefix w_n of x is a prefix of p_j and (since G is finite) $\varphi(p_j q_j) = \varphi^{(n_j)}(x) = g$.

Example 6.11. Consider once more the Fibonacci substitution $\sigma : a \mapsto ab$, $b \mapsto a$ and the Fibonacci shift X. Let $\varphi : A^* \to \mathbb{Z}/2\mathbb{Z}$ be the morphism $\varphi : a \mapsto 1$, $b \mapsto 0$. We saw in §3.4

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

that the skew product $G \rtimes_\varphi X$ is minimal and uniquely ergodic. We provide here another argument for this which uses bifix codes.

First, let Z be the bifix code such that $Z^* = \varphi^{-1}(0)$ and $U = Z \cap \mathcal{L}(X)$. We find $U = \{aa, aba, b\}$. Consider an alphabet $B = \{u, v, w\}$ and define a morphism $\phi\colon B^* \to A^*$ by

$$\phi\colon u \mapsto aa, v \mapsto aba, w \mapsto b.$$

By construction, the image of ϕ has the same intersection with $\mathcal{L}(X)$ as the submonoid $\varphi^{-1}(0) = Z^*$. Moreover, we have $\sigma^3 \circ \phi = \phi \circ \tau$, where τ is the morphism

$$\tau\colon u \mapsto vvwuw, v \mapsto vvwuwuw, w \mapsto v.$$

Letting Y be the shift space generated by τ, the skew product $G \rtimes_\varphi X$ can be identified with the tower $\widehat{Y}$ relative to the function $f(x) = |\phi(x_0)|$ (see [**25**, §1.1.3] for details). Since τ is primitive, Y is minimal and uniquely ergodic by Michel's theorem, and so is $\widehat{Y}$.

6.3. *Relation with average length.* We finish the section by coming back to the notion of average length discussed in §4. Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. Let Z be the group code such that $\varphi^{-1}(1_G) = Z^*$, and $U = Z \cap \mathcal{L}(X)$ the prefix code considered in Proposition 6.8. Recall the formula for the average length of U relative to μ:

$$\ell(U) = \sum_{u \in U} |u|\, \mu([u]).$$

Example 6.12. The bifix code U in Example 6.11 has average length

$$\ell(U) = \mu([\varepsilon]) + \mu([a]) + \mu([ab]) = 1 + \frac{1}{\lambda} + \frac{1}{\lambda^2} = 2,$$

where λ is the golden ratio (cf. Figure 10).

By (4.1), one has also

$$\ell(U) = \sum_{w \in P} \mu([w]),$$

where P is the set of proper prefixes of some element of U. Moreover, under the hypotheses of Theorem 1.1, that is, ergodicity of the appropriate product measure on the skew product $G \rtimes_\varphi X$, where $L = Z^* = \varphi^{-1}(1_G)$, the following equalities hold, as an extension of Proposition 4.8:

$$\delta_\mu(L) = \frac{1}{|G|} = \frac{1}{\ell(U)}.$$

Indeed, we show that ergodicity of the skew product entails minimality (Corollary 7.2), which implies that the restriction of φ on the sets $\mathcal{R}_X(u)^*$ is onto by Theorem 6.1. Then, by Proposition 6.8, the degree $d_X(U)$ of U is $|G|$. We thus are in the scope of [**7**, Corollary 4.3.8], which states that $d_X(U) = \ell(U)$ and, so, $\ell(U) = |G|$.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

7. *Minimal subsets and modular cobounding maps*

We now generalize some results from the previous sections, in particular, Theorem 1.1, which allows a simple expression of the density under the assumption of ergodicity, and Theorem 6.1, which characterizes minimality (referring here to the action of the skew product T_φ as defined in (3.1), since the action of the shift map S on X is already assumed to be minimal). We rely on the key notion of *modular cobounding map* (Definition 7.3), closely related with the notion of coboundary. It allows, in particular, to obtain a further characterization of minimality (Theorem 7.5) inspired by Anzai's theorem on ergodicity of skew products [**2**].

We prove first that the minimal skew products under consideration (that is, skew products $G \rtimes_\varphi X$ with φ being a morphism onto a finite group G) are finite disjoint unions of their minimal closed invariant subsets, all of which have the same measure (Proposition 7.1). We conclude that ergodicity implies minimality (Corollary 7.2). Moreover, Proposition 7.10 together with Corollary 7.11 provide sufficient conditions for ergodicity on every minimal closed invariant subset of $G \rtimes_\varphi X$. We prove our third main result, Theorem 1.3, in §7.4 and conclude with examples in §7.5.

7.1. *Modular coboundaries.*

We start by examining the general structure of skew products in terms of minimal closed invariant subsets.

PROPOSITION 7.1. *Let X be a minimal shift space on A with an invariant measure μ and $\varphi\colon A^* \to G$ be a morphism onto a finite group G with uniform probability measure ν. The skew product $G \rtimes_\varphi X$ is a disjoint union of its minimal closed invariant subsets, which are finite in number and equal in $\nu \times \mu$-measure.*

Proof. Let G act on the left of $G \rtimes_\varphi X$ by $g(h, x) = (gh, x)$. This action commutes with the transformation T_φ of the skew product (defined in (3.1)) and is continuous, so G acts by automorphisms. In particular, it permutes the set of minimal closed invariant subsets.

Let us fix a pair $(g, x) \in G \times X$ and show that it is contained in some minimal closed invariant subset of $G \rtimes_\varphi X$. By Zorn's lemma, there exists at least one minimal closed invariant subset, say $Y \subseteq G \rtimes_\varphi X$. Its projection on X is also a minimal closed invariant subset; thus, by minimality of X, it must be X itself. It follows that $(h, x) \in Y$ for some $h \in G$, and then $gh^{-1}Y$ is a minimal closed invariant subset containing (g, x).

This also shows that G acts transitively on the set of all minimal closed invariant subsets; thus, it must be finite and with cardinality dividing $|G|$. Moreover, this action by G is measure-preserving, as is easily checked on rectangular sets: indeed, for every measurable sets $E \subseteq G$ and $F \subseteq X$, we have

$$(\nu \times \mu)(g(E \times F)) = (\nu \times \mu)(gE \times F) = (\nu \times \mu)(E \times F).$$

Thus, all minimal closed invariant subsets must have the same measure. □

The above result has the following straightforward consequences.

COROLLARY 7.2. *Let X be a minimal shift space on A with an invariant measure μ and $\varphi\colon A^* \to G$ be a morphism onto a finite group G with uniform probability measure ν.*

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

(i) *If $\nu \times \mu$ restricts to an ergodic measure on a closed invariant subset Y, then Y must be minimal.*
(ii) *If $\nu \times \mu$ restricts to an ergodic measure on any minimal closed invariant subset, then it must restrict to an ergodic measure on each of them.*
(iii) *If $\nu \times \mu$ is ergodic, then $G \rtimes_\varphi X$ must be minimal.*

We now proceed to describe the minimal closed invariant subsets of $G \rtimes_\varphi X$, using the following key notion inspired by [38].

Definition 7.3. Let (X, S) be a shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. A *cobounding map* for φ on X is a continuous map $\alpha\colon X \to H\backslash G$ to the set of right cosets of a subgroup $H \leq G$ such that, for all $x \in X$, the following relates the images of α on x and on its shifted image Sx:

$$\alpha(Sx) = \alpha(x)\varphi(x_0).$$

We also say that α is a cobounding map *mod* H. Observe that if α is a cobounding map, then

$$\alpha(S^n x) = \alpha(x)\varphi^{(n)}(x)$$

for all $n \in \mathbb{Z}$, with $\varphi^{(n)}$ as in (3.2).

As the name suggests, this definition is related to the cohomological equations developed in ergodic theory. These equations have a rich history, as evidenced, for instance, by [2, 20, 38, 52, 56, 57, 59]. In particular, we may view a cobounding map as a 'certificate' that (the cocycle defined by) φ is a *coboundary* mod H. The next proposition clarifies the link between cobounding maps and closed invariant subsets. It is a special case of a result of Lemańczyk and Mentzen [38, Proposition 2.1].

PROPOSITION 7.4. *Let X be a minimal shift space on A with an invariant measure μ and $\varphi\colon A^* \to G$ be a morphism onto a finite group G with uniform probability measure ν. Let α be a cobounding map mod H for φ and $Y_\alpha = \{(g, x) \mid \alpha(x) = Hg\}$.*
(i) *The set Y_α is a closed invariant subset of $G \rtimes_\varphi X$ of $\nu \times \mu$-measure $1/[G : H]$.*
(ii) *If Y is a minimal closed invariant subset of $G \rtimes_\varphi X$, then $Y = Y_\alpha$ for some cobounding map α.*

Proof. (i) Note that $Y_\alpha = \bigcup_{Hg \in H\backslash G} Hg \times \alpha^{-1}(Hg)$. Since α is continuous, each $\alpha^{-1}(Hg)$ is closed and so is Y_α. Fix a pair $(g, x) \in Y_\alpha$, which means that $g \in Hg = \alpha(x)$. Hence, $g\varphi^{(n)}(x) \in \alpha(x)\varphi^{(n)}(x) = \alpha(S^n x)$ and $T_\varphi^n(g, x) = (g\varphi^{(n)}(x), S^n x) \in Y_\alpha$. Thus, Y_α is invariant. It has measure

$$\begin{aligned}(\nu \times \mu)(Y_\alpha) &= \sum_{Hg \in H\backslash G} (\nu \times \mu)(Hg \times \alpha^{-1}(Hg)) \\ &= \frac{|H|}{|G|} \sum_{Hg \in H\backslash G} \mu(\alpha^{-1}(Hg)) = \frac{1}{[G : H]}.\end{aligned}$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

(ii) Let Y be a minimal closed invariant subset of $G \rtimes_\varphi X$. Consider the subgroup $H = \{h \in G \mid hY = Y\}$ and for each $x \in X$, let

$$\alpha(x) = \{h \in H \mid (h, x) \in Y\}.$$

Since X is minimal, the restriction to Y of the projection $P_X \colon G \times X \to X$ is surjective and, therefore, α is defined everywhere on X. Fix an element $g \in \alpha(x)$. We claim that $\alpha(x) = Hg$. On the one hand, it is clear that for $h \in H$, $(hg, x) \in hY = Y$; thus, $hg \in \alpha(x)$. This shows that $Hg \subseteq \alpha(x)$. On the other hand, for $k \in \alpha(x)$, we have that $(g, x) = (gk^{-1}k, x) \in Y \cap gk^{-1}Y$. However, note that $gk^{-1}Y$ is also a closed invariant subset of $G \rtimes_\varphi X$, since Y is and T_φ commutes with the left action of G on $G \rtimes_\varphi X$. This shows that $Y \cap gk^{-1}Y$ is a non-empty closed invariant subset of Y and, thus, $Y \cap gk^{-1}Y = Y$ by minimality of Y.

For every $(h, y) \in Y$, we conclude that $h = kg^{-1}h'$ for some $h' \in G$ such that $(h', y) \in Y$, which implies that $kg^{-1}(h, y) = (h', y) \in Y$. This means that $kg^{-1} \in H$ and, thus, $k = kg^{-1}g \in Hg$, as claimed. In particular, this shows that α is a map $X \to H\backslash G$ and that $Y = Y_\alpha$. It remains to show that α is a cobounding map.

To establish continuity, it suffices to show that $\alpha^{-1}(Hg)$ is closed for each $g \in G$. However, observe that $\alpha^{-1}(Hg)$ can be written in terms of the two component projections $P_G \colon G \rtimes_\varphi X \to G$ and $P_X \colon G \rtimes_\varphi X \to X$:

$$\alpha^{-1}(Hg) = \{x \in X \mid g \in \alpha(x)\} = \{x \in X \mid (g, x) \in Y\} = P_X(P_G^{-1}(g) \cap Y).$$

Since Y is a closed subspace, P_G is continuous, and P_X is a closed map, we conclude that $\alpha^{-1}(Hg)$ is indeed closed.

We end the proof by showing that $\alpha(x)\varphi^{(n)}(x) = \alpha(S^n x)$ for every $x \in X$ and $n \in \mathbb{Z}$. Take first $h \in \alpha(x)$, so $(h, x) \in Y$. Since Y is invariant,

$$(h\varphi^{(n)}(x), S^n x) = T_\varphi^n(h, x) \in Y$$

and, therefore, $h\varphi^{(n)}(x) \in \alpha(S^n x)$. This shows that $\alpha(x)\varphi^{(n)}(x) \subseteq \alpha(S^n x)$. As this inclusion holds for all $x \in X$ and $n \in \mathbb{Z}$, it also holds for x replaced by $S^n x$ and n replaced by $-n$, which yields

$$\alpha(S^n x)\varphi^{(-n)}(S^n x) \subseteq \alpha(S^{-n}S^n x) = \alpha(x).$$

Since $\varphi^{(-n)}(S^n x) = \varphi((S^n x)_{[-n,0)})^{-1} = (\varphi^{(n)}(x))^{-1}$ by (3.2), this also proves the other inclusion. □

7.2. *A characterization of minimality via cobounding maps.* There is always at least one cobounding map, namely the constant map $X \to G\backslash G$, which we call the *trivial cobounding map*. The corresponding closed invariant subset is then the whole skew product. It is immediately apparent that the existence of a non-trivial cobounding map thus forbids minimality of the skew product. In fact, we have the following consequence of Proposition 7.4, which is reminiscent of Anzai's theorem on ergodicity of skew products [**2**] (see also [**48**, Ch. 2, Theorem 4.8]).

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

THEOREM 7.5. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. The skew product $X \rtimes_\varphi G$ is minimal if and only if there exists no non-trivial cobounding maps for φ.*

Proof. We prove the contrapositive implications. If there is a non-trivial cobounding map $\alpha\colon X \to H\backslash G$, then the subset Y_α from Proposition 7.4 is a proper, non-empty, closed invariant subset of $G \rtimes_\varphi X$. Conversely, if $G \rtimes_\varphi X$ is not minimal, then it has (by Zorn's lemma) a proper minimal closed invariant subset Y, which must then be of the form $Y = Y_\alpha$ for some cobounding map $\alpha\colon X \to H\backslash G$, by condition (ii) of Proposition 7.4. Since Y_α is proper, α must be non-trivial. □

Given two cobounding maps α and β, respectively mod H and K, write $\alpha \le \beta$ if $H \le K$ and $\alpha^{-1}(Hg) \subseteq \beta^{-1}(Kg)$ for all $g \in G$. This gives a partial order on cobounding maps, which corresponds directly to the ordering of the corresponding closed invariant subsets under inclusion.

PROPOSITION 7.6. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. For any two cobounding maps α and β, $Y_\alpha \subseteq Y_\beta$ if and only if $\alpha \le \beta$. Therefore, the minimal closed invariant subsets of $G \rtimes_\varphi X$ correspond to the minimal cobounding maps under $\le$.*

Proof. Assume $\alpha\colon X \to H\backslash G$ and $\beta\colon X \to K\backslash G$. If $\alpha \le \beta$, then $Hg \subseteq Kg$ and $\alpha^{-1}(Hg) \subseteq \beta^{-1}(Kg)$ for every $g \in G$; thus,

$$Y_\alpha = \bigcup_{Hg \in H\backslash G} Hg \times \alpha^{-1}(Hg) \subseteq \bigcup_{Kg \in K\backslash G} Kg \times \beta^{-1}(Kg) = Y_\beta.$$

Conversely, assume that $Y_\alpha \subseteq Y_\beta$ and fix $Hg \in H\backslash G$. Then, $Hg \times \alpha^{-1}(Hg) \subseteq Kg' \times \beta^{-1}(Kg')$ for some $g' \in G$. In particular, $g \in Kg'$ so we may assume $g = g'$. We then deduce that $H \subseteq K$ and $\alpha^{-1}(Hg) \subseteq \beta^{-1}(Kg)$. □

The left action of G on $G \rtimes_\varphi X$ corresponds to the left action on cobounding maps given by $(g\alpha)(x) = g(\alpha(x))$, where $g\alpha$ is viewed as a cobounding map mod $H^g = gHg^{-1}$. This cobounding map is such that $g\alpha(x) = H^g h \iff \alpha(x) = Hg^{-1}h$; hence, $X_{g\alpha} = gX_\alpha$. This shows that the passage from minimal closed invariant subsets to minimal cobounding maps preserves the left action of G. In particular, G acts transitively on the set of minimal cobounding maps.

COROLLARY 7.7. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. Let H be a subgroup of G such that there exists a minimal cobounding map $\alpha\colon X \to H\backslash G$. Then, for any other minimal cobounding map $\beta\colon X \to K\backslash G$, the group K is a conjugate of H. Consequently, the number of minimal closed invariant subsets of the skew product $G \rtimes_\varphi X$ equals $[G : H]$.*

7.3. *Cobouding maps and return words.* We have established in Theorem 6.1 a characterization of minimality in terms of return words. Without surprise, we also find

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

links between cobounding maps and return words. This makes the relationship between conditions (ii) and (iii) in Theorem 6.1 more transparent (see Remark 7.9).

PROPOSITION 7.8. *Let X be a minimal shift space on A and $\varphi\colon A^* \to G$ be a morphism onto a finite group G. Let H be a subgroup of G.*

(i) *If a cobounding map $\alpha\colon X \to H\backslash G$ takes constant value Hg on a cylinder $[u]_X$, then $\varphi(\mathcal{R}_X(u)) \subseteq g^{-1}Hg$.*
(ii) *If a word $u \in \mathcal{L}(X)$ satisfies $\varphi(\mathcal{R}_X(u)) \subseteq H$, then there exists a cobounding map $\alpha\colon X \to H\backslash G$ that takes constant value H on $[u]_X$.*
(iii) *A cobounding map $\alpha\colon X \to H\backslash G$ is minimal if and only if $\varphi(\mathcal{R}_X(u))$ generates $g^{-1}Hg$ whenever u is such that α takes constant value Hg on $[u]_X$.*

Proof. (i) Consider $w \in \mathcal{R}_X(u)^*$. Taking $x \in [wu]_X \subseteq [u]_X$, we find

$$\alpha(x) = \alpha(S^{|w|}x) = \alpha(x)\varphi(w) = Hg\varphi(w),$$

while the stabilizer of Hg under the right action of G on $H\backslash G$ is exactly $g^{-1}Hg$.

(ii) Let $u \in \mathcal{L}(X)$ be such that $\varphi(\mathcal{R}_X(u)) \leq H$. For $x \in X$, let

$$C_x = \{j \geq 0 \mid x_{[j,j+|u|)} = u\}.$$

Observe that under the assumption that $\varphi(\mathcal{R}_X(u)) \subseteq H$, the value of $H\varphi(x_{[0,j)})^{-1}$ is identical for every $j \in C_x$. Hence, we may define a map $\alpha\colon X \to H\backslash G$ by

$$\alpha(x) = H\varphi(x_{[0,j)})^{-1} \quad \text{where } j \in C_x.$$

Since X is a minimal shift space, there exists a constant $m > 0$ independent of x such that $\min(C_x) < m$; hence, α is continuous, since its value is determined by the first m letters. Moreover, fixing $x \in X$ and $j \in C_x$ with $j \geq 1$, we find that

$$\alpha(Sx) = H\varphi(x_{[1,j)})^{-1} = H\varphi(x_{[0,j)})^{-1}\varphi(x_0) = \alpha(x)\varphi(x_0).$$

(iii) Assume first that α is minimal. Fix $h \in H$ and suppose that α takes constant value Hg on some cylinder $[u]$. We need to show that $g^{-1}hg$ belongs to $\langle\varphi(\mathcal{R}_X(u))\rangle$. Since $\{g\} \times [u]$ and $\{hg\} \times [u]$ are two non-empty clopen subsets of the minimal closed invariant subset Y_α, we may find $k \in \mathbb{Z}$ with $(\{g\} \times [u]) \cap T_\varphi^{-k}(\{hg\} \times [u]) \neq \varnothing$. Choose x in that intersection, and let $w = x_{[0,k)}$ if $k \geq 0$ and $w = x_{[k,0)}$ otherwise. It follows that

$$(hg, x) = T_\varphi^k(g, x) = \begin{cases} (g\varphi(w), S^k x) & \text{if } k \geq 0, \\ (g\varphi(w)^{-1}, S^k x) & \text{otherwise;} \end{cases}$$

hence, $\varphi(w) = g^{-1}h^{\pm 1}g$ (according to whether $k \geq 0$ or $k < 0$). However, w is a concatenation of elements of $\mathcal{R}_X(u)$; thus, $g^{-1}hg$ belongs to the subgroup of G generated by $\varphi(\mathcal{R}_X(u))$.

To prove the converse, we consider a cobounding map $\beta\colon X \to K\backslash G$ such that $\beta \leq \alpha$. Let u be a word such that both α and β are constant on $[u]$, say with $\alpha([u]) = Hg$ and $\beta([u]) = Kh$. Note that $h \in Kh \subseteq Hg$, so we may assume that $h = g$. By part (i) of the statement, $\varphi(\mathcal{R}_X(u))$ generates a subgroup of $g^{-1}Kg$, while it also generates $g^{-1}Hg$ by

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

our assumption. Thus, $H = K$ and $\beta = \alpha$ on $[u]$. As we may partition X into a union of such cylinders, we get $\alpha = \beta$, and thus show that α is minimal. □

Remark 7.9. From this result combined with Theorem 7.5, we deduce an alternate proof for the equivalence between conditions (ii) and (iii) in Theorem 6.1. Indeed, on the one hand, if the skew product is not minimal, then Theorem 7.5 states that there exists a non-trivial cobounding map; hence, it follows from condition (i) of Proposition 7.8 that for a sufficiently long word u, all return words in $\mathcal{R}_X(u)$ are mapped inside some proper subgroup of G. On the other hand, if the image of some return set $\mathcal{R}_X(u)$ fails to generate G, then by condition (ii) of Proposition 7.8, there exists a non-trivial cobounding map; hence, the skew product cannot be minimal by Theorem 7.5.

Roughly speaking, the smaller the subgroup, the more restrictive the coboundary condition and, hence, the most stringent cobounding maps are the cobounding maps mod 1, where the notation mod 1 refers to working modulo the trivial subgroup. One important fact is that having such cobounding maps turns out to be sufficient for the measure $\nu \times \mu$ to be ergodic on each of the minimal closed invariant subsets of $G \rtimes_\varphi X$.

PROPOSITION 7.10. *Let X be a minimal shift space on A with an ergodic measure μ and $\varphi\colon A^* \to G$ a morphism onto a finite group G with uniform probability measure ν. If there exists a cobounding map $\alpha\colon X \to G$ mod 1, then the product measure $\nu \times \mu$ is ergodic on Y_α.*

Proof. The map $\gamma\colon X \to Y_\alpha, \gamma(x) = (\alpha(x), x)$, is a homeomorphism which intertwines S and T_φ, and satisfies $\mu(E)/|G| = (\nu \times \mu)(\gamma(E))$ for every measurable set $E \subseteq X$. Thus, $(Y_\alpha, T_\varphi, \nu \times \mu)$ is measure-theoretically isomorphic to (X, S, μ), and since the latter is ergodic, so is the former. □

We moreover observe that cobounding maps mod 1 are minimal by Proposition 7.8(iii). In this special case, Proposition 7.8 also yields the following corollary.

COROLLARY 7.11. *Let X be a minimal shift space on A with and $\varphi\colon A^* \to G$ a morphism onto a finite group G. The following conditions are equivalent:*

(i) *φ has a cobounding map mod 1 on X;*
(ii) *$\varphi(\mathcal{R}_X(u)) = 1$ for every long enough $u \in \mathcal{L}(X)$;*
(iii) *$\varphi(\mathcal{R}_X(u)) = 1$ for some word $u \in \mathcal{L}(X)$.*

Proof. That (i) implies (ii) follows from Proposition 7.8(i); (ii) implies (iii) is trivial; (iii) implies (i) follows from Proposition 7.8(ii). □

7.4. *A formula for density in terms of cobounding maps.* The next theorem is our third main result. It gives a simple closed form for the density in terms of any minimal cobounding map under suitable assumptions. It generalizes Corollary 3.6 within the setting of minimal shift spaces.

THEOREM 1.3. *Let X be a minimal shift space on A with an ergodic measure μ and $\varphi\colon A^* \to G$ a morphism onto a finite group G with uniform probability measure ν.*

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Suppose that $\nu \times \mu$ is ergodic on each of the minimal closed invariant subsets of $G \rtimes_\varphi X$. Then, for every group language $L = \varphi^{-1}(K)$, where $K \subseteq G$, the density $\delta_\mu(L)$ is given by the following formula, where $\alpha\colon X \to H\backslash G$ is any minimal cobounding map:

$$\delta_\mu(L) = \frac{1}{|H|} \sum_{k\in K} \sum_{Hg\in H\backslash G} \mu(\alpha^{-1}(Hg))\mu(\alpha^{-1}(Hgk)).$$

Note that when $G \rtimes_\varphi X$ is ergodic, the trivial cobounding map $\alpha\colon X \to G\backslash G$ is minimal and we recover the formula from Corollary 3.6:

$$\delta_\mu(L) = \frac{1}{|G|} \sum_{k\in K} \sum_{Gg\in G\backslash G} \mu(\alpha^{-1}(Gg))\mu(\alpha^{-1}(Ggk)) = \frac{|K|}{|G|}.$$

Proof. Without loss of generality, we may assume that $L = \varphi^{-1}(k)$ for some fixed $k \in G$. Fix a minimal cobounding map $\alpha\colon X \to H\backslash G$ and let $\overline{\mu}$ be the measure on $G \rtimes_\varphi X$ defined by

$$\overline{\mu}(B) = [G : H](\nu \times \mu)(B \cap Y_\alpha).$$

Note that $\overline{\mu}$ is ergodic by assumption. For $g \in G$, let $U_g = \{g\} \times X$. We claim that $\nu \times \mu$ projects to μ. Indeed, for every measurable subset $B \subseteq X$,

$$\begin{aligned}
\overline{\mu}(G \times B) &= \sum_{Hg\in H\backslash G} \overline{\mu}((Hg \times B) \cap Y_\alpha) \\
&= [G : H] \sum_{Hg\in H\backslash G} (\nu \times \mu)(Hg \times (B \cap \alpha^{-1}(Hg))) \\
&= \sum_{Hg\in H\backslash G} \mu(B \cap \alpha^{-1}(Hg)) \\
&= \mu(B).
\end{aligned}$$

Applying Theorem 1.1 with $\{g\} \times X = U_g$, we obtain

$$\begin{aligned}
\delta_\mu(\varphi^{-1}(k)) &= \sum_{g\in G} \overline{\mu}(U_g)\overline{\mu}(U_{gk}) \\
&= [G : H]^2 \sum_{g\in G} (\nu \times \mu)(U_g \cap Y_\alpha)\ (\nu \times \mu)(U_{gk} \cap Y_\alpha) \\
&= [G : H]^2 \sum_{g\in G} \frac{1}{|G|^2} \mu(\alpha^{-1}(Hg))\mu(\alpha^{-1}(Hgk)) \\
&= \frac{1}{|H|^2} \sum_{g\in G} \mu(\alpha^{-1}(Hg))\mu(\alpha^{-1}(Hgk)),
\end{aligned}$$

which after regrouping like terms,

$$= \frac{1}{|H|} \sum_{Hg\in H\backslash G} \mu(\alpha^{-1}(Hg))\mu(\alpha^{-1}(Hgk)),$$

which concludes the proof. □

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

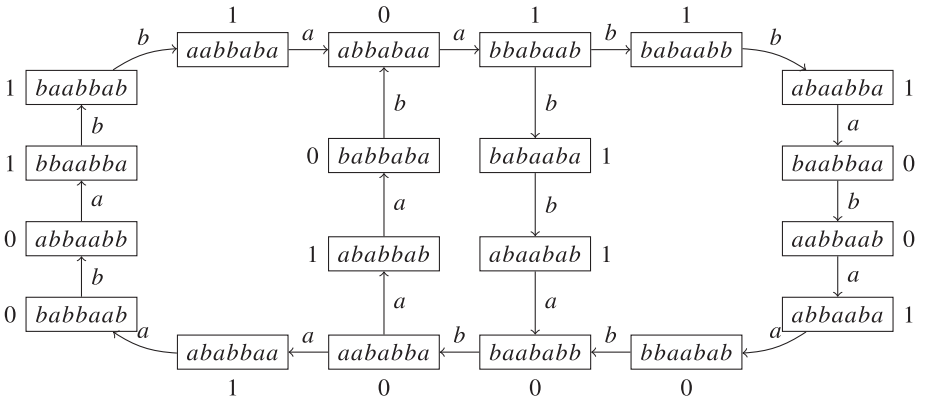


FIGURE 8. One of the two cobounding maps $X \to \mathbb{Z}/2\mathbb{Z}$ on the Thue–Morse shift for the morphism $\varphi\colon \{a, b\}^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = 1$, $\varphi(b) = 0$. The map is constant on cylinders of length 7.

Recall from Corollary 7.2 that the ergodicity assumption from the above theorem is equivalent to $\nu \times \mu$ being ergodic on *any* closed invariant subset (which is then necessarily minimal).

7.5. *Examples.* We finish the section with two examples that illustrate various aspects of Theorem 1.3.

Example 7.12. Let X be the Thue–Morse shift with its unique invariant measure μ (see Example 2.1) and let $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = 1$, $\varphi(b) = 0$.

The morphism φ has two cobounding maps mod 1 on X; hence, by Corollary 7.11, $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ has two minimal closed invariant subsets on which the measure $\nu \times \mu$ is ergodic. The cobounding maps take constant values on the cylinders of length 7; one of them is depicted in Figure 8. Note that the cobounding maps are not constant on cylinders of length 6: for instance, the map of Figure 8 satisfies $\alpha([baabbab]) = 1$ and $\alpha([baabbaa]) = 0$, but $[baabbaa]$ and $[baabbaa]$ are both contained in the cylinder $[baabba]$ of length 6.

Observe that the cobounding maps are *fair*, in the sense that the preimages $\alpha^{-1}(g)$ have the same μ-measure for all elements of the group. Therefore, even though the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ is *not* ergodic, the shift space X is still evenly distributed with respect to φ, in the sense that

$$\delta_\mu(\varphi^{-1}(1)) = \delta_\mu(\varphi^{-1}(0)) = 1/2.$$

Example 7.13. Consider the following substitution and morphism onto S_3 defined on the alphabet $A = \{a, b, c\}$:

$$\sigma : a \mapsto aab,\ b \mapsto acb,\ c \mapsto ba, \quad \varphi(a) = \varphi(c) = (1\,2\,3), \quad \varphi(b) = (1\,2).$$

Let X be the shift space generated by the primitive substitution σ. We claim that the skew product $S_3 \rtimes_\varphi X$ has *three* minimal closed invariant subsets. These minimal subsets correspond to three cobounding maps

$$\alpha_{12}\colon X \to \langle(1\,2)\rangle\backslash S_3, \quad \alpha_{13}\colon X \to \langle(1\,3)\rangle\backslash S_3, \quad \alpha_{23}\colon X \to \langle(2\,3)\rangle\backslash S_3.$$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

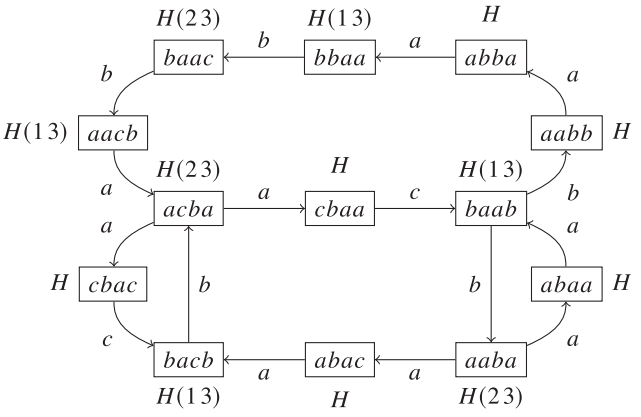


FIGURE 9. Cobounding map mod $H = \langle(1\,2)\rangle$ on the shift of Example 7.13 for the morphism $\varphi\colon \{a, b, c\}^* \to S_3$, $\varphi(a) = \varphi(c) = (1\,2\,3)$, $\varphi(b) = (1\,2)$. The map is constant on cylinders of length 4.

The cobounding map α_{12} is depicted in Figure 9 and the other two can be deduced from α_{12} using the natural left action of S_3 on cobounding maps. Note that in all cases, the subgroup involved has index 3 in S_3, in accordance with Corollary 7.7.

We now briefly sketch a proof of the fact that the above cobounding maps are indeed minimal. In the present case, this is equivalent to showing that there are no cobounding maps mod 1. Recall that by Corollary 7.11, there is a cobounding map mod 1 if and only if $\mathcal{R}_X(u)$ has trivial image under φ for all sufficiently long words u.

By the first main result from [**13**], since σ is a bifix encoding, there exists a constant $K > 0$ such that for all $u \in \mathcal{L}(\sigma)$ with $|u| \geq K$,

$$\mathcal{R}_X(\sigma(u)) = \sigma(\mathcal{R}_X(u)).$$

Using the formula provided in [**13**], we find the upper bound $K \leq 6$, but direct computations show that we can take $K = 2$; in fact, the one-letter word $u = c$ is the *only* word which fails the above equality. Now, take the sequence of words $u_n = \sigma^n(a)$; we claim that $(2\,3) \in \mathcal{R}_X(u_n)$ for infinitely many n. Indeed, observe that the following equalities hold:

$$\varphi \circ \sigma = \varphi \circ \sigma^8, \quad \varphi \circ \sigma(a) = \varphi \circ \sigma(b) = \varphi \circ \sigma(c) = (2\,3).$$

As $\mathcal{R}_X(u_0) = \{a, ba, bba, cba\}$, it follows that $(2\,3) = \varphi(\sigma^{7k+1}(a))$ belongs to $\mathcal{R}_X(u_{7k+1})$ for all $k \geq 0$. This shows that $\mathcal{R}_X(u_n)$ have non-trivial images for infinitely many n; thus, φ has no cobounding map mod 1 on X. This confirms that the above cobounding maps are minimal.

8. *Ergodicity for primitive substitutions*

In this section, we focus on the special case of shift spaces defined by primitive substitutions. Our main result is a sufficient condition for the minimal closed invariant subsets of skew products to be uniquely ergodic (Proposition 8.3). As a corollary, we deduce that substitutive dendric shifts have ergodic skew products with all finite groups (Theorem 8.11). Note that the family of dendric shifts, studied in [**7**, **9**–**12**], encompasses

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

several classical families of shifts, such as Sturmian shifts, codings of interval exchanges, and Arnoux–Rauzy shifts.

8.1. *Skew products based on primitive shifts.* Let us fix a primitive substitution σ over a finite alphabet A and let $X = X(\sigma)$ be the shift space defined by σ. The shift X is a minimal shift space and we recall that it is uniquely ergodic by Michel's theorem.

Definition 8.1. Let $\varphi\colon A^* \to G$ be a morphism onto a finite group G. We say that the primitive substitution σ is *invertible under* φ if:

$$\text{there exists } n \geq 1, \quad \varphi \circ \sigma^n = \varphi. \tag{8.1}$$

Example 8.2. Let $\sigma\colon a \mapsto ab, b \mapsto a$ be the Fibonacci substitution and consider the morphism $\varphi\colon A^* \to \mathbb{Z}/2\mathbb{Z}$, $\varphi(a) = 1, \varphi(b) = 0$. One checks that

$$\sigma^3(a) = abaab, \quad \sigma^3(b) = aba;$$

hence, $\varphi \circ \sigma^3 = \varphi$, so σ is invertible under φ. This property has already been used in the proof of Proposition 3.17. In fact, σ has the much stronger property of *being invertible under every homomorphism onto a finite group*, as we show later (Lemma 8.5).

PROPOSITION 8.3. *Let σ be a primitive substitution, X be its shift space, and $\varphi\colon X \to G$ be a morphism onto a finite group G. If σ is invertible under φ, then the minimal closed invariant subsets of $G \rtimes_\varphi X$ are uniquely ergodic.*

Proof. Up to replacing σ by some power, we may assume without loss of generality (since this does not change the shift space) that $\varphi \circ \sigma = \varphi$ and that σ has a fixed point $y \in X$.

Let Ψ be the topological conjugacy from Lemma 3.1. The fact that $\varphi = \varphi \circ \sigma$ entails the existence of a substitution $\overline{\sigma}$ on $(G \times A)^*$ such that $\Psi(g, \sigma(x)) = \overline{\sigma}(\Psi(g, x))$, namely, when $\sigma(a) = b_0 \cdots b_{n-1}$,

$$\overline{\sigma}(g, a) = (g_0, b_0)(g_1, b_1) \cdots (g_{n-1}, b_{n-1}) \quad \text{where } g_0 = g,\ g_{i+1} = g_i \varphi(b_i).$$

Observe that, for every $g \in G$, the infinite word $z = \Psi(g, y)$ is a uniformly recurrent fixed point of $\overline{\sigma}$, since

$$\overline{\sigma}(z) = \overline{\sigma}(\Psi(g, y)) = \Psi(g, \sigma(y)) = \Psi(g, y) = z.$$

Let B be the subset of letters in $G \times A$ appearing in z; it follows that $\overline{\sigma}$ restricts to a substitution over B. Moreover, z belongs to a minimal subset of $(G \times A)^{\mathbb{Z}}$; hence, it must be uniformly recurrent. Since $\overline{\sigma}$ is a growing substitution fixing a uniformly recurrent word, it must be primitive and have for shift space the closed orbit of z, which is $\Psi(Y)$. In particular, $\Psi(Y)$ is uniquely ergodic by Michel's theorem and so is Y. □

Observe that Example 7.13 fails both the invertibility property (8.1) *and* the property of Corollary 7.11. At this time, we do not know whether the product measure is ergodic on the minimal closed invariant subsets of the skew product. In contrast, here is an example which satisfies both (8.3) and Corollary 7.11.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Example 8.4. Consider the following substitution and morphism onto $\mathbb{Z}/2\mathbb{Z}$ defined on the alphabet $A = \{a, b, c, d\}$:

$$\sigma : a \mapsto baa,\ b \mapsto adc,\ c \mapsto cdc,\ d \mapsto ad, \quad \varphi : a, b \mapsto 0,\ c, d \mapsto 1.$$

Then, $\varphi \circ \sigma = \varphi$, and hence, the minimal closed invariant subsets of $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$ are uniquely ergodic. The substitution $\overline{\sigma}$ satisfying $\Psi(g, \sigma(x)) = \overline{\sigma}(\Psi(g, x))$ is defined as follows on the alphabet $\mathbb{Z}/2\mathbb{Z} \times A$, written a_0, a_1, b_0, b_1, etc. for convenience,

$$\begin{aligned} &a_0 \mapsto b_0a_0a_0, \quad b_0 \mapsto a_0d_0c_1, \quad c_0 \mapsto c_0d_1c_0, \quad d_0 \mapsto a_0d_0, \\ &a_1 \mapsto b_1a_1a_1, \quad b_1 \mapsto a_1d_1c_0, \quad c_1 \mapsto c_1d_0c_1, \quad d_1 \mapsto a_1d_1. \end{aligned}$$

This substitution splits in two the primitive substitutions defined respectively on the alphabets $B = \{a_0, b_0, c_1, d_0\}$ and $C = \{a_1, b_1, c_0, d_1\}$. Each corresponds to one of the two minimal closed invariant subsets of the skew product $\mathbb{Z}/2\mathbb{Z} \rtimes_\varphi X$.

In what follows, we say that a substitution $\sigma : A^* \to A^*$ is *invertible* if its extension to an endomorphism on the free group F_A is an automorphism. We now state two further properties when the substitution is assumed to be invertible.

LEMMA 8.5. *Let σ be a primitive substitution. If σ is invertible, then it is invertible under every morphism $\varphi : A^* \to G$ onto a finite group G.*

Proof. Let $\mathrm{Aut}(F_A)$ be the automorphism group of F_A and $\mathrm{Hom}(F_A, G)$ be the set of morphisms $F_A \to G$. For $\tau \in \mathrm{Aut}(F_A)$, denote by τ_* the self-map of $\mathrm{Hom}(F_A, G)$ defined by

$$\tau_*(\varphi) = \varphi \circ \tau.$$

Since $\rho_* \circ \tau_* = (\tau \circ \rho)_*$, the map $\tau \mapsto \tau_*$ is a morphism from $\mathrm{Aut}(F_A)$ to the symmetric group on $\mathrm{Hom}(F_A, G)$. Moreover, observe that $\mathrm{Hom}(F_A, G)$ is a finite set, being in bijection with the set of maps $A \to G$. As a result, for every automorphism τ of F_A, τ_* is a permutation of $\mathrm{Hom}(F_A, G)$ with finite order; in other words, there exists $n \geq 0$ such that $\tau_*^n = \mathrm{id}$, that is, $\varphi \circ \tau^n = \varphi$ for every morphism $F_A \to G$. Applying this to the extension of σ to an automorphism of F_A yields the result. □

In what follows, we say that a substitution is *aperiodic* if it generates a shift space that contains no finite orbit. We next establish the following lemma that is used in the next section. Observe that the generation property expressed below implies the one stated in Theorem 6.1.

LEMMA 8.6. *Let σ be a primitive aperiodic substitution and X the shift generated by σ. If $\mathcal{R}_X(u)$ generates F_A for every $u \in \mathcal{L}(X)$, then σ is invertible.*

Proof. First, fix a point $x \in X$ which is periodic under σ, meaning $\sigma^k(x) = x$ for some $k > 0$ [**25**, Proposition 1.4.8]. Let w be a word of the form $w = x_{[-n,n)}$ and let $u = x_{[0,n)}$ be its suffix of length n, for some $n \geq 0$. As Almeida and Costa observed in the proof of [**1**, Proposition 5.5], it follows from Mossé's recognizability theorem [**45**, Theorem 3.1bis] that when n is large enough, $u\mathcal{R}_X(w)u^{-1} \subseteq \mathrm{im}(\sigma^k)$. Since by assumption $\mathcal{R}_X(w)$

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

generates F_A, we conclude that σ^k is a surjective endomorphism of F_A. Since finitely generated free groups have the Hopfian property [41, Proposition 3.5], it follows that σ^k is invertible and, hence, so is σ. □

For the sake of completeness, we give an example showing that the converse of the above lemma fails, that is, invertibility does not guarantee that all $\mathcal{R}_X(u)$ generate F_A.

Example 8.7. Let σ be the primitive substitution from Example 8.4 defined on the four-letter alphabet $A = \{a, b, c, d\}$ by

$$\sigma : a \mapsto baa,\ b \mapsto adc,\ c \mapsto cdc,\ d \mapsto ad.$$

This is an invertible substitution, with σ^{-1} given by

$$a \mapsto bc^{-1}d^{-1}b,\ b \mapsto ab^{-1}dcb^{-2}dcb^{-1},\ c \mapsto d^{-1}b,\ d \mapsto b^{-1}dcb^{-1}d.$$

Nonetheless, the following is not a generating set of F_A:

$$\mathcal{R}_X(a) = \{a, ba, dcba, dcdca, dcdcba\}.$$

In fact, this set of return words generate the rank 3 subgroup of F_A with basis $\{a, b, dc\}$.

8.2. *Skew products based on dendric shifts.* We turn now to the *dendric* case. First, we recall the definition. Let X be a shift space. For $w \in \mathcal{L}(X)$, let $\mathrm{L}(w) = \{a \in A \mid aw \in \mathcal{L}(X)\}$ and $\mathrm{R}(w) = \{a \in A \mid wa \in \mathcal{L}(X)\}$. We denote by $\mathrm{E}(w)$ the graph with vertices the disjoint union of $\mathrm{R}(w)$ and $\mathrm{L}(w)$, and edges the pairs $(a, b) \in A \times A$ such that $awb \in \mathcal{L}(X)$; it is called the *extension graph* of w. A shift space X is *dendric* if for every $w \in \mathcal{L}(X)$, the extension graph $\mathrm{E}(w)$ is a tree. For instance, every Sturmian shift is dendric [11].

An important result concerning dendric shifts is the so-called *return theorem* of Berthé *et al*, which we quote next.

THEOREM 8.8. [11, Theorem 4.5] *Let X be a dendric shift on an alphabet A. For every $w \in \mathcal{L}(X)$, the set $\mathcal{R}_X(w)$ is a basis of the free group on A.*

Therefore, it follows from Theorem 6.1 that every skew product of a dendric shift and a finite group is minimal.

Example 8.9. The Fibonacci shift (see §3.4) is Sturmian and therefore dendric. There, we have $\mathrm{R}(a) = \mathrm{L}(a) = \{a, b\}$ and the graph $\mathrm{E}(a)$ is shown in Figure 10. Moreover, $\mathcal{R}_X(a) = \{a, ab\}$, which is obviously a basis of the free group on $\{a, b\}$.

We also give an example of a shift space which is not dendric.

Example 8.10. Let X be the Thue–Morse shift from Example 2.1, generated by the two-letter substitution $\sigma : a \mapsto ab, b \mapsto ba$. For the word $w = aba$, we find $\mathrm{L}(w) = \mathrm{R}(w) = \{a, b\}$ and the graph $\mathrm{E}(w)$, depicted in Figure 11, is not connected.

We recover, as a consequence of the following general result, the unique ergodicity of the skew product of §3.4.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

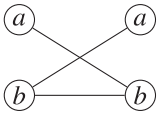


FIGURE 10. The extension graph E(a) in the Fibonacci shift.

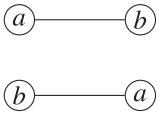


FIGURE 11. The extension graph E(w), $w = aba$, in the Thue–Morse shift.

THEOREM 8.11. *Let X be a dendric shift generated by a primitive substitution. Then, the skew product $G \rtimes_\varphi X$ is uniquely ergodic for every morphism $\varphi\colon A^* \to G$ onto a finite group G.*

The proof uses Lemmas 8.5 and 8.6. Observe that dendric shift spaces, in particular, fall under the scope of Lemma 8.6, thanks to the return theorem of Berthé *et al*, stated above as Theorem 8.8.

Proof. Assume that X is generated by the primitive morphism σ. Observe that σ must be aperiodic, as dendric spaces cannot contain finite orbits. Moreover, by the return theorem, $\langle \mathcal{R}_X(w)\rangle = F_A$ for all $w \in \mathcal{L}(X)$. Therefore, we may apply Lemmas 8.5 and 8.6 to conclude that σ is invertible and, as a result, invertible under every morphism $\varphi\colon A^* \to G$ onto a finite group G. Applying Proposition 8.3, it follows that the skew product $G \rtimes_\varphi X$ has uniquely ergodic minimal closed invariant subsets. As the skew product is also minimal by the return theorem (Theorem 8.8) and Theorem 6.1, this completes the proof. □

We thus obtain, as a direct application of Corollary 3.6, the following result about the density of group languages in substitutive dendric shifts.

COROLLARY 8.12. *Let X be a dendric shift generated by a primitive substitution and let μ be its unique ergodic measure. For every morphism $\varphi\colon A^* \to G$ onto a finite group G and every language $L = \varphi^{-1}(K)$, $K \subseteq G$, the density $\delta_\mu(L)$ exists and is equal to $|K|/|G|$.*

9. *Skew products based on Sturmian shifts*

We end with a discussion about earlier related works on skew products based on irrational rotations, mostly by Veech [**56**, **57**] and Jager and Liardet [**34**]. Due to the nature of the examples involved and to stay consistent with the relevant literature, it is convenient here to use alphabets consisting of natural numbers, such as {0, 1} and {1, 2}.

Among the first classical examples of skew products, skew translations (that is, skew products with base an irrational rotation on the unit circle), and their ergodic and spectral properties have been widely investigated; see e.g. [**27**, **31**, **42**, **54**, **56**, **57**] and the classical references [**21**, **48**]. In particular, they have been used to produce examples of interval exchanges that are not uniquely ergodic [**36**]. Such examples are based on skew products

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

of irrational rotations associated with the group $\mathbb{Z}/2\mathbb{Z}$ that are minimal and not uniquely ergodic, with the skewing function being the indicator function of an interval [56, 57].

More precisely, let α be an irrational number in $[0, 1]$. We consider the rotation $R_\alpha : x \mapsto x + \alpha$ modulo 1 defined on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Let $I = [0, \beta)$ be a semi-open interval of $\mathbb{T}$. Let $1_I : \mathbb{T} \to \{0, 1\}$ be the indicator function of I, that is, $1_I(x) = 1$ if and only if $x \in I$. Let $m \geq 2$ and let $G = \mathbb{Z}/m\mathbb{Z}$. Let $\varphi : \{0, 1\}^* \to G$ be the morphism defined by $0 \mapsto 0$, $1 \mapsto 1$. We then consider the skew product $G \rtimes_I \mathbb{T}$ of R_α defined as

$$(k, x) \mapsto (k + \varphi \circ 1_I(x), x + \alpha) = (k + \varphi \circ 1_I(x), R_\alpha(x)).$$

Such skew products over rotations are closely related to symbolic skew products such as those considered in the present paper and, more precisely, to skew products over shifts obtained as binary codings of rotations. Let I^c stand for the complement of I in $\mathbb{T}$. We use the notation $I_0 = I$ and $I_1 = I^c$. Let x be the infinite word in $\{0, 1\}^{\mathbb{Z}}$ obtained by coding the orbit of 0 under R_α with respect to the partition $\mathcal{I} = \{I, I^c\}$, that is, for any $n \in \mathbb{Z}$, $x_n = 1$ if and only if $R_\alpha^n(0) \in I$, or in other words, $x_n = 1_I(R_\alpha^n(0))$ for all $n \in \mathbb{Z}$. Let (X, S) be the shift space generated by x. Then, this shift is minimal and uniquely ergodic since α is irrational and I is semi-open. Indeed, we associate in a bijective way words $w = w_1 \cdots w_n$ in $\{0, 1\}^n$ with sets $I_w := I_{w_1} \cap R_\alpha^{-1} I_{w_2} \cap \cdots \cap R_\alpha^{-n+1} I_{w_n}$ as follows: w occurs at index k in x if and only if $k\alpha \in I_w$. It is important to stress the fact that I is assumed to be semi-open. As an illustration of the relevance of this hypothesis, consider the case $I = [0, \alpha]$, with $\alpha < 1/2$; then, 00 occurs only once in the orbit of 0 and minimality fails. The fact that I is semi-open guarantees that if I_w is not empty, then there are *infinitely many* k such that $k\alpha \in I_w$, by density of the sequence $(k\alpha)_k$; hence, the uniform recurrence of x and, thus, the minimality of X. When the length of I equals α or $1 - \alpha$, such binary codings of rotations are Sturmian.

Not all intervals I lead to ergodic skew products. In fact, by [56], if α is not badly approximable, then there exists an interval I such that the skew product $G \rtimes_I \mathbb{T}$ of R_α is not ergodic, a result which inspired the elegant characterization of ergodicity from [31] stated in terms of Ostrowski's numeration. Moreover, as proved in [57], skewing a badly approximable rotation over a finite number of intervals with rational endpoints still provides a uniquely ergodic skew product. However, [19] provides a $\mathbb{Z}/2\mathbb{Z}$ skew product of a badly approximable rotation that is minimal and not uniquely ergodic, by skewing over the indicator function of a union of two intervals.

In most examples considered in [56, 57], intervals have lengths that do not belong to $\mathbb{Z}\alpha + \mathbb{Z}$. We consider here the complementary case of Sturmian shifts. We can then apply Theorem 8.11 and Corollary 8.12 when they are furthermore assumed to be generated by a substitution, such as exemplified below. Note also that substitutive Sturmian shifts have been characterized in [22, 58] (in particular, the parameter α of the underlying rotation is quadratic and, thus, badly approximable).

Example 9.1. Let X be the Fibonacci shift on the alphabet $\{0, 1\}$. We consider the skew product $\mathbb{Z}/m\mathbb{Z} \rtimes_\varphi X$, where φ is the morphism $\{0, 1\}^* \to \mathbb{Z}/m\mathbb{Z}$ given by $0 \mapsto 0$, $1 \mapsto 1$. In particular, one has $\varphi^{(n)}(x) = |x_0 \cdots x_{n-1}|_1$ modulo m for $n \geq 0$, as explained in Example 3.2. By Theorem 8.11, $\mathbb{Z}/m\mathbb{Z} \rtimes_\varphi X$ is uniquely ergodic, which yields

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

equidistribution results on the congruence of the number of visits of R_α to the interval $[0, \alpha)$, where $\alpha = (\sqrt{5}-1)/2$. In other words, for every $x \in X$, $r \in \mathbb{Z}/m\mathbb{Z}$ and $a \in \{0, 1\}$, one has

$$\frac{1}{N} \operatorname{Card}\{n : 0 \le n \le N-1,\ |x_0 \cdots x_{n-1}|_a \equiv r \bmod m\} \to \frac{1}{m},$$

or, in other words,

$$\frac{1}{N} \operatorname{Card}\{n : 0 \le n \le N-1,\ \operatorname{Card}\{i : 0 \le i < n,\ i\alpha \in [0, \alpha)\} \equiv r \bmod m\} \to \frac{1}{m}.$$

Finally, we consider in the next example, inspired by the work of Jager and Liardet [34], equidistribution properties for convergents in continued fraction expansions.

Example 9.2. Let X be the Fibonacci shift on the alphabet $\{1, 2\}$. We continue Example 3.16 with the skew product $G(2) \rtimes_\varphi X$ with the non-Abelian skewing group $G(2) = \mathrm{GL}(2, \mathbb{Z}/2\mathbb{Z})$ in relation to distribution properties modulo 2 for convergents of continued fraction expansions, inspired by the work [34], which handles the case of a random real number.

Let $x = (x_n)_{n\in\mathbb{Z}} \in X$. Consider the real number in the unit interval $[0, 1]$ that admits $(x_n)_{n\ge1}$ as its sequence of partial quotients and let $(p_n(x)/q_n(x))_{n\in\mathbb{N}}$ stand for the associated sequence of rational approximations. By Theorem 8.11, the skew product $G(2) \rtimes_\varphi X$ is uniquely ergodic. We can thus deduce, as detailed in Example 3.16, the following distribution results in the group $G(2)$, as a counterpart of [34, Theorem 3.11], which holds for almost every real number in $[0, 1]$ (see also [15, 44, 55] for related works). For every $k = 1, 2$ and for every $x \in X$,

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Card}\{n : 1 \le n \le N,\ q_n(x) \equiv 0 \bmod 2\} = \frac{1}{3},$$
$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Card}\{n : 1 \le n \le N,\ q_n(x) \equiv 1 \bmod 2\} = \frac{2}{3}.$$

In fact, the distribution in the group $G(2)$ of the sequence of continued fraction convergents $(p_n/q_n)_{n\in\mathbb{N}}$, whose sequence of partial quotients is given by elements of the Fibonacci shift X, behaves like that of a random irrational number. Note that we recover the well-known fact that certain residue classes are attained more frequently than others (as already noted at the end of Example 3.16, the entries of equidistributed matrices in $G(2)$ are not equidistributed modulo 2).

This statement can be considered as a modulo m counterpart of Lévy's theorem stating that $\lim_{n\to\infty}((\log q_n)/n) = \pi^2/12 \log 2$ almost everywhere. It is therefore natural to ask how the convergents behave without taking them modulo m, in terms of convergence of $(\log q_n)/n$. Consider the cocyle map $\psi\colon \{0, 1\}^* \to \mathrm{GL}(2, \mathbb{R})$ defined similarly as φ, but now with the matrix $\psi(k)$ being considered in $\mathrm{GL}(2, \mathbb{R})$ (without reduction modulo 2). By [28, Theorem 3], there exists some constant $\Lambda_X > 0$ such that $\lim_{n\to\infty}(1/n) \log \|\psi^{(n)}(x)\| = \Lambda_X$ for all $x \in X$, which yields the existence of $\lim_{n\to\infty}(1/n) \log q_n(x) = \Lambda_X$ for all $x \in X$. In other words, the cocycle ψ is uniform, with the terminology of [28]. We use here the fact that X is minimal, uniquely ergodic, and the cocycle map ψ is such that the entries of $\psi^{(2)}$ are positive.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

Lastly, observe that similar results can be obtained for higher-dimensional continued fractions via skew products defined with primitive dendric shifts on larger alphabets, such as codings of interval exchanges; consider for instance the Jacobi–Perron algorithm whose equidistribution properties modulo m are studied in [14] for random numbers.

Acknowledgements. We would first like to thank the anonymous referees for carefully reading through our paper and for making very helpful comments and suggestions. We warmly thank Samuel Petite for leading us to a mistake in an early version of this paper and France Gheeraert for suggesting a simplification in the proof of Theorem 6.1. We also thank Olivier Carton and Vincent Delecroix for insightful conversations about the notion of pointwise densities.

V.B. is supported by Agence Nationale de la Recherche through the project SymDynAr (ANR-23-CE40-0024) and 2024 ERC Synergy Project DynAMiCs (101167561). H.G.-O. was partially supported by the CTU Global Postdoc Fellowship program. C.-F.N.-B. is supported by the KIAS Individual Grant (MG094701) at Korea Institute for Advanced Study. V.B. and D.P. were supported by the Agence Nationale de la Recherche through the project 'IZES' (ANR-22-CE40-0011).

References

[1] J. Almeida and A. Costa. Presentations of Schützenberger groups of minimal subshifts. *Israel J. Math.* **196**(1) (2013), 1–31.
[2] H. Anzai. Ergodic skew product transformations on the torus. *Osaka J. Math.* **3** (1951), 83–99.
[3] L. Baggett, H. Medina and K. Merrill. Cohomology of polynomials under an irrational rotation. *Proc. Amer. Math. Soc.* **126**(10) (1998), 2909–2918.
[4] L. Baggett and K. Merrill. Representations of the Mautner group and cocycles of an irrational rotation. *Michigan Math. J.* **33**(2) (1986), 221–229.
[5] L'. Balková, M. Bucci, A. De Luca, J. Hladký and S. Puzynina. Aperiodic pseudorandom number generators based on infinite words. *Theor. Comput. Sci.* **647** (2016), 85–100.
[6] J. Berstel. Sur la densité asymptotique de langages formels. *Automata Languages and Programming*. Ed. M. Nivat. North-Holland, Amsterdam, 1972, pp. 345–358.
[7] J. Berstel, C. De Felice, D. Perrin, C. Reutenauer and G. Rindone. Bifix codes and Sturmian words. *J. Algebra* **369** (2012), 146–202.
[8] J. Berstel, D. Perrin and C. Reutenauer. *Codes and Automata*. Cambridge University Press, Cambridge, 2009.
[9] V. Berthé, C. De Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer and G. Rindone. Bifix codes and interval exchanges. *J. Pure Appl. Algebra* **219** (2015), 2781–2798.
[10] V. Berthé, C. De Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer and G. Rindone. Maximal bifix decoding. *Discrete Math.* **338** (2015), 725–742.
[11] V. Berthé, C. De Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer and G. Rindone. Acyclic, connected and tree sets. *Monatsh. Math.* **176**(4) (2015), 521–550.
[12] V. Berthé, C. De Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer and G. Rindone. The finite index basis property. *J. Pure Appl. Algebra* **219**(7) (2015), 2521–2537.
[13] V. Berthé and H. Goulet-Ouellet. On substitutions preserving their return sets. *Combinatorics on Words (Lecture Notes in Computer Science, 13899)*. Ed. A. Frid and R. Mercas. Springer, Cham, 2023, pp. 77–90.
[14] V. Berthé, H. Nakada and R. Natsui. Arithmetic distributions of convergents arising from Jacobi–Perron algorithm. *Indag. Math. (N.S.)* **17**(2) (2006), 169–185.
[15] B. Borda. Equidistribution of continued fraction convergents in $\mathrm{SL}(2,\mathbb{Z}_m)$ with an application to local discrepancy. *J. Mod. Dyn.* **21** (2025), 327–359.
[16] M. Boshernitzan. A unique ergodicity of minimal symbolic flows with linear block growth. *J. Anal. Math.* **44**(1) (1984), 77–96.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

[17] M. Boshernitzan. A condition for unique ergodicity of minimal symbolic flows. *Ergod. Th. & Dynam. Sys.* **12**(3) (1992), 425–428.
[18] A. I. Bufetov. Skew products and ergodic theorems for group actions. *J. Math. Sci. (N.Y.)* **113**(4) (2003), 548–557.
[19] J. Chaika. Skew products over rotations with exotic properties. *Geom. Dedicata* **168** (2014), 369–385.
[20] J. P. Conze. Équirépartition et ergodicité de transformations cylindriques. *Publ. Séminaire Math. Info. Renne* **2** (1976), Talk no. 2.
[21] I. P. Cornfeld, S. V. Fomin, and Y. G. Sinaĭ. *Ergodic Theory* (*Grundlehren der Mathematischen Wissenschaften, 245)*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ.
[22] D. Crisp, W. Moran, A. Pollington and P. Shiue. Substitution invariant cutting sequences. *J. Théor. Nombres Bordeau* **5**(1) (1993), 123–137.
[23] M. Denker, C. Grillenberger and K. Sigmund. *Ergodic Theory on Compact Spaces*. Springer, Berlin, 1976.
[24] F. Durand. Linearly recurrent subshifts have a finite number of non-periodic subshift factors. *Ergod. Th. & Dynam. Sys.* **20** (2000), 1061–1078.
[25] F. Durand and D. Perrin. *Dimension Groups and Dynamical Systems*. Cambridge University Press, Cambridge, 2022.
[26] S. Eilenberg. *Automata, Languages and Machines (Pure and Applied Mathematics, A)*. Academic Press, New York, 1974.
[27] S. Ferenczi and P. Hubert. Minimality and unique ergodicity of Veech 1969 type interval exchange transformations. *Geom. Dedicata* **218**(3) (2024), 77.
[28] A. Furman. On the multiplicative ergodic theorem for uniquely ergodic systems. *Ann. Inst. H. Poincaré Probab. Statist.* **33**(6) (1997), 797–815.
[29] H. Furstenberg. Strict ergodicity and transformation of the torus. *Am. J. Math.* **83**(4) (1961), 573–601.
[30] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, Princeton, NJ, 1981.
[31] M. Guenais and F. Parreau. Eigenvalues of transformations arising from irrational rotations and step functions. *Preprint*, 2006, arXiv:0605250.
[32] G. Hansel and D. Perrin. Codes and Bernoulli partitions. *Math. Syst. Theory* **16**(1) (1983), 133–157.
[33] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers (Oxford Mathematics)*, 6th edn. Oxford University Press, Oxford, 2008.
[34] H. Jager and P. Liardet. Distributions arithmétiques des dénominateurs de convergents de fractions continues. *Indag. Math. (N.S.)* **50**(2) (1988), 181–197.
[35] S. Kakutani. Random ergodic theorems and Markoff processes with a stable distribution. *Proc. 2nd Berkeley Symp. on Mathematical Statistics and Probability*. Ed. J. Neyman. University of California Press, Berkeley, CA, 1951, pp. 247–262.
[36] H. B. Keynes and D. Newton. A 'minimal', non-uniquely ergodic interval exchange transformation. *Math. Z.* **148**(2) (1976), 101–105.
[37] N. Kryloff and N. Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Ann. of Math. (2)* **38**(1) (1937), 65.
[38] M. Lemańczyk and M. K. Mentzen. Topological ergodicity of real cocycles over minimal rotations. *Monatsh. Math.* **134** (2002), 227–246.
[39] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.
[40] P. Lummerzheim, F. Pogorzelski and E. Zimmermann. Strict irreducibility of Markov chains and ergodicity of skew products. *Math. Z.* **311**(4) (2025).
[41] R. C. Lyndon and P. E. Schupp. *Combinatorial Group Theory*. Springer, Berlin, 2001.
[42] K. D. Merrill. Cohomology of step functions under irrational rotations. *Israel J. Math.* **52**(4) (1985), 320–340.
[43] P. Michel. Stricte ergodicité d'ensembles minimaux de substitution. *C. R. Math. Acad. Sci. Paris* **278** (1974), 811–813.
[44] R. Moeckel. Geodesics on modular surfaces and continued fractions. *Ergod. Th. & Dynam. Sys.* **2**(1) (1982), 69–83.
[45] B. Mossé. Puissance de mots et reconnaissabilité des points fixes d'une substitution. *Theor. Comput. Sci.* **99**(2) (1992), 327–334.
[46] J. C. Oxtoby. Ergodic sets. *Bull. Amer. Math. Soc. (N.S.)* **58**(2) (1952), 116–136.
[47] D. Perrin and A. Restivo. A note on Sturmian words. *Theoret. Comput. Sci.* **429** (2012), 265–272.
[48] K. Petersen. *Ergodic Theory*. Cambridge University Press, Cambridge, 1983.
[49] R. R. Phelps. *Lectures on Choquet's Theorem (Lecture Notes in Mathematics, 1757)*. Springer, Berlin, 2001.
[50] M. Queffélec. *Substitution Dynamical Systems: Spectral Analysis*. Springer, Berlin, 2010.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press

[51] A. Ramsay. Nontransitive quasi-orbits in Mackey's analysis of group extensions. *Acta Math.* **137** (1976), 17–48.
[52] K. Schmidt. *Cocycles on Ergodic Transformation Groups*. MacMillan, Delhi, 1977.
[53] M. P. Schützenberger. Sur certains sous-monoïdes libres. *Bull. Soc. Math. France* **93** (1965), 209–223.
[54] M. Stewart. Irregularities of uniform distribution. *Acta Math.* **37**(1-3) (1981), 185–221.
[55] P. Szüsz. Über die metrische Theorie der diophantischen approximation. II. *Acta Math.* **8** (1962/63), 225–241.
[56] W. A. Veech. Strict ergodicity in zero dimensional dynamical systems and the Kronecker–Weyl theorem mod 2. *Trans. Amer. Math. Soc.* **140** (1969), 1.
[57] W. A. Veech. Finite group extensions of irrational rotations. *Israel J. Math.* **21**(2) (1975), 240–259.
[58] S.-I. Yasutomi. On Sturmian sequences which are invariant under some substitutions. *Number Theory and its Applications, Kyoto 1997 (Developments in Mathematics, 2)*. Ed. S. Kanemitsu and K. Gyory. Springer, New York, 1999, pp. 347–373.
[59] R. J. Zimmer. Extensions of ergodic group actions. *Illinois J. Math.* **20** (1976), 373–409.

https://doi.org/10.1017/etds.2026.10318 Published online by Cambridge University Press