

Preservation Theorems for Transducer Outputs

Anonymous Author(s)

ABSTRACT

Suppose we have a deterministic finite-state transducer \mathcal{A} and an infinite word x , and run \mathcal{A} on x to obtain an infinite word $\mathcal{A}(x)$. Which properties of x are guaranteed to also hold for $\mathcal{A}(x)$? In this paper, we study this preservation question for various well-known combinatorial properties, e.g., recurrence, being morphic, and having factor frequencies. The celebrated Krohn-Rhodes theorem provides the framework for proving our preservation results, and our techniques are based on the ergodic theory of symbolic dynamical systems, i.e., shift spaces.

CCS CONCEPTS

• Theory of computation \rightarrow Logic and verification.

KEYWORDS

Krohn-Rhodes Theorem, Transducers, Ergodic Theory, Word Combinatorics, Morphic Words

ACM Reference Format:

Anonymous Author(s). 2026. Preservation Theorems for Transducer Outputs. In . ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

Words with elegant combinatorial properties often accurately capture the behaviour of symbolic dynamical systems and natural phenomena. Classes of self-similar words like *morphic words* [36, 37], and in particular, *automatic sequences* [2] have profound connections with logic and automata theory [39]. Such words have been of interest for over a century: e.g., the ubiquitous Thue-Morse word (as qualified in [1]) was defined to study a problem concerning geodesics on surfaces [36, Chap. 5]. It is moreover an example of a *uniformly recurrent* word, i.e., every factor occurs infinitely often (this first condition makes the word *recurrent*) and there is a bound on the gaps between its consecutive occurrences (this additional condition makes the word *uniformly recurrent*). This *almost periodic* nature of uniformly recurrent words has been used to model quasicrystals in the setting of aperiodic order [3]: one may consult [36] and its bibliography for such applications in physics.

A perennially relevant class of uniformly recurrent words is that of *toric words* [7]. They have traditionally attracted interest in the study of dynamical systems and ergodic theory (see [41, Chap. 1] and [42]). More recently, toric words have been identified to accurately capture the behaviour of *linear loops* in program verification

[28]. They also play a central role in the decidability of the monadic second-order (MSO) theory of the structure $\langle \mathbb{N}; <, a_1^{\mathbb{N}}, \dots, a_k^{\mathbb{N}} \rangle$ as established by [6].

Intuitively, if the infinite word x is a trace of a system (e.g., the coding of an orbit of a point under the action of a dynamical system), and \mathcal{A} is a deterministic finite-state transducer, then $\mathcal{A}(x)$ is obtained by augmenting x with annotations with respect to some specification provided by the transducer. It is natural to ask: are combinatorial properties of x preserved even after adding such annotations, and to what extent? More precisely, this paper considers the following concrete questions.

- (1) If x is recurrent, then is $\mathcal{A}(x)$ also recurrent? Can the *recurrence function* of $\mathcal{A}(x)$ be described and computed?
- (2) If x admits *factor frequencies*, then does $\mathcal{A}(x)$? If these frequencies are effective for x , are they effective for $\mathcal{A}(x)$?

In the above, (i) the recurrence function $R_x: \mathbb{N} \rightarrow \mathbb{N} \cup \{\infty\}$ of a recurrent infinite word x returns for each n , the maximum gap between two consecutive occurrences of a factor u of length n ; (ii) an infinite word x admits factor frequencies if for every finite word u , we have that $\lim_{N \rightarrow \infty} \frac{1}{N} \cdot |\{i \mid 0 \leq i < N, x(i, i + |u|) = u\}|$ exists.

A classical result regarding the preservation of factor frequencies by transducers is Agafonov's theorem (see e.g., [5]), which states that if \mathcal{A} is a deterministic *oblivious subsequence-selecting* transducer, then x is *normal* if and only if $\mathcal{A}(x)$ is normal. A normal word $x \in \Sigma^\omega$ admits factor frequencies such that for any $u \in \Sigma^+$, the frequency of u is $1/|\Sigma|^{|u|}$. Normality is a prerequisite for *randomness*, and its preservation by transducers has been studied in [11, 12]. This paper complements the above results by showing that transducers can also preserve the existence of factor frequencies by virtue of preserving combinatorial structure.

We identify that the Krohn-Rhodes theorem [33] provides a convenient framework to answer our preservation-related questions. This is because it guarantees that $\mathcal{A}(x)$ can be obtained as a cascade $\sigma \circ \mathcal{B}_k \circ \dots \circ \mathcal{B}_1(x)$, where $\mathcal{B}_1, \dots, \mathcal{B}_k$ are “simple” letter-to-letter transducers whose underlying automata are either *permutation automata* or *reset automata*, and σ is a letter-to-word substitution. In other words, it suffices to answer the above questions for substitutions (done in Lem. 4.1), permutation automata, and reset automata.

To answer the first question regarding recurrence, we use the Krohn-Rhodes framework to revisit the work of Pritykin [35], who used a result of Semënov et al. [34] to show that if x has a uniformly recurrent suffix, then so does $\mathcal{A}(x)$. The advantages of using the Krohn-Rhodes theorem are as follows. (i) We can prove a slightly stronger result, i.e., if x has a recurrent suffix, then so does $\mathcal{A}(x)$. (ii) We can already give an effective answer in the affirmative to the second question regarding frequencies in the special case of uniformly recurrent words and *counter-free* transducers. (iii) In the original case of uniform recurrence and arbitrary transducers, we can explicitly describe and compute the starting index of the uniformly recurrent suffix of $\mathcal{A}(x)$, and also the recurrence function of this suffix. In particular, if x is *linearly recurrent* (i.e., the recurrence

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://doi.org/10.1145/nnnnnnn).
Conference'17, July 2017, Washington, DC, USA

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
<https://doi.org/10.1145/nnnnnnn>

function of x grows linearly), then the suffix of $\mathcal{A}(x)$ is also linearly recurrent; likewise if x is polynomially recurrent, then so is the suffix of $\mathcal{A}(x)$. Moreover, if x is (linearly, polynomially, uniformly) recurrent and the underlying automaton of \mathcal{A} is a permutation automaton, then $\mathcal{A}(x)$ itself possesses the corresponding recurrence property. These preservation results are stated and proved in Sec. 5, and will be useful to treat the second question regarding factor frequencies in greater generality.

What about preservation of having factor frequencies? The astute reader might already provide the counterexample of the (non-recurrent automatic) word $x \in \{0, 1\}^\omega$ such that $x(n) = 1$ if and only if n is a power of 3, and the two-state transducer \mathcal{A} that permutes the states upon reading the letter 1 and prints the current state. The word x admits factor frequencies but $\mathcal{A}(x)$ does not. Applying Lem. 6.3, we deduce that $\mathcal{A}(x)$ is also automatic, and thus automatic (and hence morphic) words do not necessarily admit factor frequencies.

The question of what structure a word needs in order to admit factor frequencies is indeed profound. We note in Lem. 3.1 the intrinsic connection to symbolic dynamics and ergodic theory: a word x admits factor frequencies if and only if it is *generic* for an *invariant measure* on the *shift space* X that it generates. Normal words, for instance, are by definition generic for the uniform Borel probability measure on the full shift Σ^ω . We therefore seek to understand when we can guarantee that $\mathcal{A}(x)$ is also a generic point for some invariant measure on the shift space Y it generates.

The most reliable way of doing so is to ensure that the shift Y generated by $\mathcal{A}(x)$ is *uniquely ergodic*, i.e., it admits a single invariant measure; by Oxtoby's theorem, all words in Y would then be generic for this measure. It does not suffice to merely require that x generate a uniquely ergodic shift: this condition is met by the counterexample above. Nevertheless, if a shift space is uniquely ergodic, then the support of its invariant measure will be a *minimal* shift space, i.e., a shift space generated by a uniformly recurrent word. In the above counterexample, this support is the singleton shift space generated by 0^ω . We further prove that all words in a minimal uniquely ergodic shift space admit computable factor frequencies (Lem. 3.2).

It is tempting to try to establish that if x generates a minimal shift space X that is uniquely ergodic, then $\mathcal{A}(x)$ generates a shift space Y that is uniquely ergodic. Unfortunately, as the symbolic traces of ingenious counterexamples shown by Veech [40], Sataev [38], Chaika [13], and Guenais & Parreau [26] demonstrate, this is not the case. Intuitively, the introduction of “annotations” gives rise to words whose factor frequencies “oscillate” between different limits, or, in technical terms, a generic point in X need not lift to a generic point in Y .

What additional structure must we impose on x in order to ensure that $\mathcal{A}(x)$ generates a uniquely ergodic shift space? The following is an ideal starting point. A *primitive morphic* word x is a uniformly recurrent word that is morphic, i.e., obtained by applying a substitution τ to the fixed point of a non-trivial substitution σ . Primitive morphic words are in fact linearly recurrent, and are classic examples of words that generate uniquely ergodic shifts. We show that if x has a primitive morphic suffix, then so does $\mathcal{A}(x)$, and this result is fully effective, i.e., we can compute the factor frequencies of $\mathcal{A}(x)$; see Thm. 6.2. We mention that we incidentally

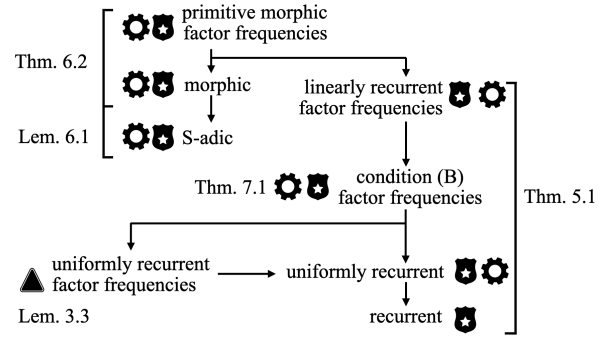


Figure 1: Summary of main results. Arrow from block P to block Q means that the properties in P imply those in Q . Shield besides P means that if x has a suffix with the properties in P , then so does $\mathcal{A}(x)$. Gear besides the shield means that the relevant suffix y of $\mathcal{A}(x)$ is effectively identifiable, and the properties in P are furthermore effective for y if they are for x . Uniformly recurrent x admitting factor frequencies does not guarantee that $\mathcal{A}(x)$ does.

prove that more general notions of self-similarity are preserved while obtaining $\mathcal{A}(x)$ from x (see Lem. 6.1).

Our main contribution lies in proving that, in order to guarantee preservation of having factor frequencies, it suffices to impose far less structure on x than that implied by being primitive morphic. More precisely, we consider an ergodic-theoretic property known as *Boshernitzan's condition*, or *Condition (B)* (Def. 3.6) on *minimal* shifts. If a minimal shift X satisfies Boshernitzan's condition, then it is uniquely ergodic. Boshernitzan's condition is rather relaxed: it is known that all linearly recurrent words, as well as all *Sturmian* words, generate shifts that satisfy it. We show that if x is uniformly recurrent and generates a shift that satisfies Boshernitzan's condition, then the uniformly recurrent suffix of $\mathcal{A}(x)$ also generates a shift that satisfies Boshernitzan's condition, and $\mathcal{A}(x)$ thus admits factor frequencies. This main result is Thm. 7.1 in the paper. As a special case, we show that the factor frequencies of $\mathcal{A}(x)$ are computable when x is a computable Sturmian word and the underlying automaton of \mathcal{A} is a permutation automaton (Cor. 8.2).

Our preservation theorems are summarised in Fig. 1, which depicts the relations between combinatorial properties, whether $\mathcal{A}(x)$ inherits them from x , and whether the corresponding preservation theorem is effective.

2 PRELIMINARIES

2.1 Notation and Terminology

Throughout this paper, we shall work with words over finite non-empty alphabets, which will usually be denoted by Σ and sometimes by Γ . The set of infinite words over Σ (indexed by \mathbb{N}) is denoted by Σ^ω , the set of finite words is denoted by Σ^* , the empty word is denoted by ε , and the set of finite non-empty words is denoted by Σ^+ . For a word x , the notation $x(i)$ denotes the letter in the i -th position of x , the notation $x(i, j)$ denotes the finite subword, or *factor* of x starting in position i and ending at position $j - 1$ (i.e.,

the length of the factor is $j - i$). When $u = x(i, j)$, we say that u has an *occurrence* in x at index i . If x is an infinite word, the notation $x(i, \infty)$ denotes the suffix of x starting at position i . We shall denote the length of a finite word u as $|u|$. The factor complexity of a word x is the function P_x that counts for each n , the number of distinct length- n factors of x . We denote by $\mathcal{L}(x)$ the *factor language* of x , i.e., the set of finite factors u of x . We shall sometimes use $\mathcal{L}_n(x)$ as a convenient shorthand for $\mathcal{L}(x) \cap \Sigma^n$.

A word $x \in \Sigma^\omega$ is said to be *recurrent* if every factor u of x occurs infinitely often. In this case, we can define the set of *return words* to a factor u of x : these are words r such that ru is a factor of x , and ru has exactly two occurrences of u , once as a prefix, and once as a suffix. A recurrent word x is said to be *uniformly recurrent* if for every factor u , its set of return words $\mathcal{R}_x(u)$ is finite. For a uniformly recurrent word x , we can define a *return-time* (or *recurrence*) function $R_x: \mathbb{N} \rightarrow \mathbb{N}$, such that $R_x(n)$ gives the length of the longest return word to a length- n factor of x . A uniformly recurrent word x is said to be *linearly recurrent* if $R_x(n) \in O(n)$. A helpful perspective is as follows: in a uniformly recurrent word x , for any factor u , the gaps between its consecutive occurrences are bounded. If x is linearly recurrent, the bound is linear in $|u|$. If for a uniformly (respectively, linearly) recurrent word $x \in \Sigma^\omega$, given any n , we can compute $x(n)$ and $R_x(n)$, we say that x is *effectively* uniformly (respectively, linearly) recurrent.

Remark 2.1. If $x(n)$ and $R_x(n)$ are computable, we can compute $\mathcal{L}_n(x)$ as the set of length- n factors of $x(0, n + R_x(n))$. This is because if the occurrence of u at index i is its first occurrence, then $v = x(0, i)$ must be a suffix of some return word r to u , since indeed, vu is a recurrent factor of x . In particular, R_x gives a bound on the index of the first occurrence of a factor. Conversely, if x is uniformly recurrent and $x(n)$ and $\mathcal{L}_n(x)$ are computable given any n , then we can compute $R_x(n)$ by enumerating $\mathcal{L}(x)$ till we find an N such that all length- N factors contain at least two occurrences of each length- n factor.

Let u be a factor of an infinite word x . The *frequency* of u in x , denoted $f_x(u)$, is given by $\lim_{N \rightarrow \infty} \frac{1}{N} |\{i \mid x(i, i + |u|) = u, i < N\}|$ (i.e., the limiting fraction of indices which mark an occurrence of u), and is defined if the limit exists. We say that a word $x \in \Sigma^\omega$ *admits factor frequencies* if for every recurrent factor u of x , the above limit exists. If furthermore, given any u , this limit can be computed, then we say that x admits *computable* factor frequencies. Formally, there is a Turing machine which, given factor u and tolerance δ as input, computes a rational number ρ such that $|f_x(u) - \rho| < \delta$. Moreover, for any given u , if the convergence of the sequence $\frac{1}{N} |\{i \mid x(i, i + |u|) = u, k \leq i < N + k\}|$ toward $f_x(u)$ holds uniformly in k , then x is said to admit *uniform frequencies* (see e.g. [36, 37]).

2.2 Shift Spaces and Measures

We shall use ergodic theory to study which subsets of uniformly recurrent words that admit factor frequencies are closed under transduction. Shift spaces will serve as our underlying dynamical systems in order to do so. We also remark that combinatorial properties (e.g., recurrence) are often studied for shift spaces, but it is usually straightforward to obtain analogous results for words.

The shift operator T maps a word $x \in \Sigma^\omega$ to $x(1, \infty)$, i.e., it deletes the starting letter. We write $T^n x$ for the n -fold application of T to $x \in \Sigma^\omega$. We endow Σ^ω with the usual product topology, thus making it a compact metric space. A *shift system* (also called a *shift*) is the dynamical system (X, T) where $X \subseteq \Sigma^\omega$ is closed and satisfies $TX \subseteq X$. When the dynamics is clear from the context, we also refer to X as a shift.

Given a word x , the *shift generated by x* is the the topological closure $\{T^n x \mid n \in \mathbb{N}\}$ of the orbit of x under T . For a shift X , we define $\mathcal{L}(X) = \{u \mid u \in \mathcal{L}(x'), x' \in X\}$, and as before, we declare $\mathcal{L}_n(X)$ to be shorthand for $\mathcal{L}(X) \cap \Sigma^n$. We have that if X is the shift generated by x , then $\mathcal{L}(X) = \mathcal{L}(x)$, and $X = \{x' \mid \mathcal{L}(x') \subseteq \mathcal{L}(x)\}$ [37, Prop. 4.6]. A shift X is said to be *minimal* if it does not have a non-empty proper closed subset Y such that $TY \subseteq Y$. Therefore, the shift generated by any x' in a minimal shift X must be X itself. We have that a word x generates a minimal shift if and only if it is uniformly recurrent [37, Prop. 4.7].

A cylinder of $X \subseteq \Sigma^\omega$ is a set of the form $\{ux' \mid x' \in \Sigma^\omega\} \cap X$ where $u \in \Sigma^*$. The cylinder generated by u will be denoted by $[u]_X$; we omit the superscript X when it is clear from the context. The cylinders are clopen (i.e., both closed and open, and in particular are the open balls of X), and a subset of X is clopen if and only if it is a finite union of cylinders. Since there are countably many cylinders, any open set is a countable union of cylinders.

We refer the reader to [22, App. B.5] or a standard text such as [4] for the basic concepts of measure spaces. Given a topological space X , its collection of *Borel-measurable sets*, or simply *Borel sets*, is the the σ -algebra generated by the open subsets of X . If X is a shift space, the σ -algebra can equivalently be generated by the cylinders of X .

A (Borel probability) measure μ on X maps each Borel set B to $\mu(B) \in [0, 1]$, satisfies $\mu(X) = 1$, and is countably additive, i.e., for any countable collection $(B_n)_{n=0}^\infty$ of disjoint Borel sets we have $\mu(\bigcup_n B_n) = \sum_n \mu(B_n)$.

A function $\pi: Y \rightarrow X$ is called *Borel measurable* if for every Borel set $B \subseteq X$, the pre-image $\pi^{-1}(B)$ is a Borel set of Y . In particular, if π is continuous, then it is Borel measurable. Indeed, if B is open, then by continuity $\pi^{-1}(B)$ is open; and we have that $\pi^{-1}(X \setminus B) = Y \setminus \pi^{-1}(B)$ and $\pi^{-1}(\bigcup_i B_i) = \bigcup_i \pi^{-1}(B_i)$. In other words, the family of sets B for which $\pi^{-1}(B)$ is Borel contains the open sets of X and forms a σ -algebra; in particular it must contain the σ -algebra generated by the family of open sets of X , i.e., the Borel sets of X . A measurable function $\pi: Y \rightarrow X$ *projects* a Borel measure ν on Y to a Borel measure μ on X as $\mu = \nu \circ \pi^{-1}$.

An *invariant measure* μ of a compact dynamical system (X, T) is a (Borel probability) measure μ such that $\mu(T^{-1}B) = \mu(B)$ for every measurable $B \subseteq X$. By the Krylov-Bogolyubov theorem [22, 3.8.4], every (X, T) has an invariant measure.

In the case of shift spaces, we can obtain an invariant measure μ by defining a probability pre-measure μ_0 on the Boolean algebra of cylinders, and then invoking the Carathéodory extension theorem [22, Thm. B.5.1] to uniquely extend it to a Borel measure μ , since the σ -algebras generated by cylinders and open sets of a shift space coincide. We require that the pre-measure μ_0 evaluate to 1 on $X =$

$[\varepsilon]_X$, and satisfy the compatibility conditions

$$\mu_0([u]_X) = \sum_a \mu_0([ua]_X) = \sum_a \mu_0([au]_X). \quad (1)$$

Observe in particular that $\mu_0 \circ T^{-1}([u]_X) = \mu_0([u]_X)$, and more generally, for any set U in the Boolean algebra of cylinders, $\mu_0(U) = \mu_0(T^{-1}U)$. The Borel measure μ is defined as

$$\mu(B) = \inf_{(U_n)_n} \sum_n \mu_0(U_n),$$

where $(U_n)_n$ ranges over sequences of (disjoint) sets in the Boolean algebra of cylinders such that $B \subseteq \bigcup_n U_n$.

Since T is continuous, $\mu \circ T^{-1}$ will also be a Borel measure ν . The corresponding pre-measure ν_0 obtained by restricting ν to the Boolean algebra of cylinders, however, is identical to μ_0 . Thus by the uniqueness of the extension, $\nu = \mu \circ T^{-1} = \mu$, i.e., μ is indeed an invariant measure.

We say that $x \in X$ is a *generic point* of an invariant measure μ (or simply, x is generic for μ) if for every continuous $h \in C(X, \mathbb{R})$ (where $C(X, \mathbb{R})$ denotes the set of all continuous functions from X to \mathbb{R}),

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} h(T^i x) = \int h d\mu.$$

We refer the reader to [22, App. B.5.2] for the precise formal definition of the integral.

2.3 Morphic Words and Primitivity

A *substitution* is a map $\sigma: \Sigma \rightarrow \Gamma^*$, and is extended to Σ^* and Σ^ω in the obvious way, i.e., concatenation. A substitution $\sigma: \Sigma \rightarrow \Sigma^*$ is said to be *primitive* if there exists a positive integer m such that for every letter $a \in \Sigma$, all letters occur in $\sigma^m(a)$. A substitution σ is said to be *non-erasing* if there does not exist a letter a such that $\sigma(a) = \varepsilon$.

A word $x \in \Sigma^\omega$ (we assume x contains all the letters in Σ) is called *substitutive* if there exists a substitution σ which is non-trivial (i.e., not the identity) over the letters of x , such that $\sigma(x) = x$. If σ is primitive, then x is said to be a *primitive substitutive word*. As an example, the Fibonacci word $x = 0100101001001 \dots$ is primitive substitutive, and is the fixed point of the substitution that maps 0 to 01 and 1 to 0.

A word $y \in \Gamma^\omega$ is called *morphic* if there exists a substitutive word $x \in \Sigma^\omega$ and a substitution $\tau: \Sigma \rightarrow \Gamma^*$ such that $y = \tau(x)$. If x is primitive substitutive, then y is said to be *primitive morphic*. Morphic words are presented as σ, τ , a prefix u such that $\sigma(u) = u$, and a letter a such that $\sigma(a)$ prolongs a (i.e., begins with a and is more than one letter long). We can convert any representation into one such that the defining substitutions are non-erasing by restricting the alphabet of σ .

Remark 2.2. Some authors require τ to be a coding, i.e., $\tau: \Sigma \rightarrow \Gamma$ in the above definition. A result of Cobham observes that this is not a restriction. Indeed, these definitions can (effectively) be used interchangeably, including in the case of primitive morphic words. The idea to prove the equivalence (presented in [20, Prop. 17]) is to define a substitution $\hat{\sigma}$ over an auxiliary alphabet $\hat{\Sigma}$, constructed by taking $|\tau(a)|$ copies of each letter $a \in \Sigma$.

There are several other sources the reader can consult for a detailed technical exposition of this remark, see e.g., [2, Cor. 7.7.5], [19, Thm. 3.8] (Thm. 9 in the arXiv version), or [27]. For completeness, we sketch the idea of [20, Prop. 17] here. For $0 \leq i < |\tau(a)| - 1$, the substitution $\hat{\sigma}$ maps the i -th copy $a^{(i)}$ of a to the concatenation $b_i^{(0)} b_i^{(1)} \dots$ of all the copies of the i -th letter b_i of $\sigma(a)$ (if $|\sigma(a)| \leq i$, then the image is the empty word). The last copy of a is mapped similarly to the copies of the remaining letters of $\sigma(a)$. We then get \hat{x} , a “stuttering” version of x , as the fixed point of $\hat{\sigma}$ iterated on $c^{(0)}$, the foremost copy of the starting letter of x . Finally, y is obtained as $\hat{\tau}(\hat{x})$, where $\hat{\tau}$ maps $a^{(i)}$ to the i -th letter of $\tau(a)$.

Remark 2.3. It is straightforward to give an effective proof of the fact that any suffix of a morphic word is morphic [2, Thm. 7.6.1]. It is also known how to decide whether a morphic word (presented as the image under τ of the fixed point of σ) is uniformly recurrent; should the decision be yes, the procedure computes letter-to-letter τ' and primitive σ' such that the input is the image under τ' of the fixed point of σ' [21, Thm. 1, Thm. 3, Sec. 4, Sec. 5] (see also [20, Lem. 4] and the statement of [23, Thm. 24(1)]). It is well known that primitive substitutive words are, in fact, *linearly* recurrent [17, Prop. 25]. By Lem. 2.4 below, the linear recurrence function and factor frequencies are effective. These properties extend to primitive morphic words by Lem. 4.1. The proof of effectiveness is implicit in classical textbook, we provide it in App. A for completeness.

LEMMA 2.4. *Let $x \in \Sigma^\omega$ be a substitutive word obtained as the fixed point of a primitive substitution σ , and let $y = \tau(x)$ be a primitive morphic word. We have that x, y are effectively linearly recurrent and admit computable factor frequencies.*

Morphic words, by their definition, can be regarded as being self-similar. This notion of self-similarity can be generalised as follows.

Definition 2.5. Let S be a set of substitutions. A word $x \in \Sigma^\omega$ is said to be *S-adic* if there exists a sequence $\sigma = (\sigma_n)_{n=0}^\infty$ of substitutions from S , and words $x^{(0)} = x, x^{(1)}, x^{(2)}, \dots$ such that for all n , $x^{(n)} = \sigma_n(x^{(n+1)})$. The sequence σ is called the *directive sequence* of the word x , and we say that x is directed by σ .

Note that substitutive and morphic words are special cases of *S-adic* words where the directive sequence is periodic and eventually periodic, respectively. The most common example of a class of *S-adic* words is the class of Sturmian words (see e.g. [8, 36]): these are words over the binary alphabet with a factor complexity of $n+1$. Let $S = \{\lambda_0, \lambda_1, \rho_0, \rho_1\}$, where $\lambda_i(i) = i, \lambda_i(j) = ij, \rho_i(i) = i, \rho_i(j) = ji$. It is well known that any Sturmian word is *S-adic* with S as above, and an analogous statement can also be made for the generalisation to Arnoux-Rauzy words [24]. As an example, the Fibonacci word is a Sturmian word, and the sequence of substitutions is $(\lambda_0 \lambda_1)^\omega$.

3 FREQUENCIES AND ERGODIC THEORY

In this section, we shall prove the following, and use it to motivate the conditions we impose on a word x (i.e., Boshernitzan’s condition) in order to ensure that $\mathcal{A}(x)$ admits factor frequencies.

LEMMA 3.1. *A word x admits factor frequencies if and only if x is generic for an invariant measure μ on the shift X that it generates. Should this be the case, we have $F_X(u) = \mu([u]_X)$.*

PROOF. The “if” implication is obvious: we simply choose the continuous function h to be the indicator 1_u which evaluates to 1 on the cylinder $[u]_X$ and 0 elsewhere.

Conversely, suppose x admits factor frequencies and generates the shift X . Recall that $\mathcal{L}(X) = \mathcal{L}(x)$, and hence the factor frequencies of x defines a pre-measure μ_0 on each cylinder in a way that satisfies the compatibility condition (1). The pre-measure μ_0 is thus invariant, and extends to an invariant measure μ . We now need to prove that x is generic for μ , i.e., for every continuous h , we have $\int h d\mu = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} h(T^i x)$.

We have a special case where the Stone-Weierstraß theorem applies. Observe that since X is compact, the function h is uniformly continuous, admits a modulus of continuity, and we can define the following sequence $(g_M)_M$ of functions that converges uniformly to h (i.e., for every δ there exists M_δ such that for all $M \geq M_\delta$ we have $\max_X |g_M - h| \leq \max_{u \in \mathcal{L}_M(X)} (\max_{[u]_X} h - \min_{[u]_X} h) < \delta$):

$$g_M = \sum_{u \in \mathcal{L}_M(X)} \max_{[u]_X} h \cdot 1_u.$$

By the dominated convergence theorem [22, Thm. B.5.3], we also have that $\int h d\mu = \lim_{M \rightarrow \infty} \int g_M d\mu$.

In particular, for every $\delta > 0$, we can choose M large enough such that

$$\left| \int h d\mu - \int g_M d\mu \right| < \delta/3,$$

and for all N , we have by uniform convergence that

$$\left| \frac{1}{N} \sum_{i=0}^N g_M(T^i x) - \frac{1}{N} \sum_{i=0}^N h(T^i x) \right| < \delta/3.$$

Since x admits factor frequencies, it is also clear that for each M , we have $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} g_M(T^i x) = \int g_M d\mu$, and in particular for our chosen M , for every δ there exists N_δ such that for all $N \geq N_\delta$, we have

$$\left| \int g_M d\mu - \frac{1}{N} \sum_{i=0}^{N-1} g_M(T^i x) \right| < \delta/3.$$

Adding our inequalities together and applying the triangle inequality, we get that for every δ there exists N_δ such that for all $N \geq N_\delta$, we have $\left| \int h d\mu - \frac{1}{N} \sum_{i=0}^{N-1} h(T^i x) \right| < \delta$, or in other words, $\int h d\mu = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} h(T^i x)$, as desired. \square

We now describe measures for which “almost” all points are generic. An invariant measure μ is called *ergodic* when the following condition is met: if a Borel set B is such that $T^{-1}B = B$, then $\mu(B)$ is either 0 or 1. Since X is a compact metric space, (X, T) is guaranteed to have an ergodic measure [22, Prop. 3.8.9]. The Birkhoff ergodic theorem [22, Thm. 3.8.5] states that if μ is an ergodic measure, then for every integrable function h , the set of points x for which $\frac{1}{N} \sum_{i=0}^{N-1} h(T^i x) \neq \int h d\mu$ has μ -measure 0. Thus, for each $u \in \mathcal{L}(X)$, the set $B_u = \{x \mid f_x(u) \neq \mu([u]_X)\}$ has μ -measure 0. By Lem. 3.1, the complement of the countable union of all such B_u is precisely the set of points that are generic for μ . Thus, the set G of generic points for μ has μ -measure 1. If (X, T) is minimal, we additionally have that G is dense (in the topological sense) in X .

¹We overload inequality to also include the case where the limit does not exist.

The exception of non-generic points does not arise in dynamical systems (X, T) that are *uniquely ergodic*, i.e., there exists only one invariant measure μ , which is guaranteed to be ergodic [22, Cor. 3.8.10]. In this case, Oxtoby’s theorem [37, Thm. 4.3] tells us that all points $x \in X$ are generic for μ . Moreover, unique ergodicity is even equivalent to each word in the shift having uniform factor frequencies [22, Prop. 3.8.14]. A word x admits uniform factor frequencies if for every $v \in \mathcal{L}(x)$ and every δ , there exists M such that for every $m \geq M$ and every $u \in \mathcal{L}_m(x)$, we have $||u|_v|/|u| - f_x(v)| < \delta$, where $|u|_v$ denotes the number of occurrences of v in u . Note that for technical convenience, we may replace $|u|_v|/|u|$ in the above by $|u|_v/(|u| - |v| + 1)$, which we denote by $f_u(v)$. This alternate definition is also natural because the denominator is the number of indices of u at which an occurrence of v is possible.

Note that if (X, T) is uniquely ergodic, then the support X' of the invariant measure μ is a minimal shift. Indeed, if it were to contain a shift Y , then Y itself would admit an invariant measure ν [22, Prop. 3.8.4] which, when extended to X' , would be distinct from μ : a contradiction. We henceforth focus on minimal shifts.

LEMMA 3.2. *If an effectively uniformly recurrent word x generates a uniquely ergodic shift X with invariant measure μ , then x admits computable factor frequencies.*

PROOF. Given a factor v of length l and a tolerance δ , we shall show that we can compute L such that for all $u \in \mathcal{L}_{L+l-1}(x)$, we have $|f_x(v) - f_u(v)| < \delta$. It then remains to count the occurrences of v in any appropriate u , e.g., $u = x(0, L+l-1)$. For notational convenience, in this proof we abbreviate \mathcal{L}_{L+l-1} as simply \mathcal{L} .

It suffices to show that for every $L \geq 1$, we have $\min_{u \in \mathcal{L}} f_u(v) \leq f_x(v) \leq \max_{u \in \mathcal{L}} f_u(v)$. Indeed, by the definition of uniform frequencies, the difference between the bounds is less than δ for some L which can be found by brute enumeration (made possible by the fact that x is effectively uniformly recurrent).

We prove the above claim by showing that $f_x(v)$ is a convex combination of the values of $f_u(v)$; more precisely, we have $f_x(v) = \sum_{u \in \mathcal{L}} f_x(u) f_u(v)$. To that end, we define the continuous function χ , which maps x to $f_{x(0, L+l-1)}(v)$. Clearly by the unique ergodicity of X , we have $\int \chi d\mu$ is equal to the summation of the right. We shall conclude by showing that $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \chi(T^i x) = f_x(v)$.

Observe that by definition, $\chi(x) = \frac{1}{L} \sum_{j=0}^{L-1} 1_v(T^j x)$. Thus, we can express $\sum_{i=0}^{N-1} \chi(T^i x) = \frac{1}{L} \sum_{i=0}^{N-1} \sum_{j=i}^{i+L-1} 1_v(T^j x)$. We can further rearrange the summation to obtain

$$\frac{1}{L} \left(L \sum_{i=L-1}^{N-L} 1_v(T^i x) + \sum_{i=0}^{L-2} (i+1)(1_v(T^i x) + 1_v(T^{N-1-i} x)) \right).$$

Thus, $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \chi(T^i x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=L-1}^{N-L} 1_v(T^i x)$. The latter sequence is $(N-2L)/N$ times the average of $N-2L$ terms, and indeed converges to the limit of the usual Birkhoff averages, i.e., to $f_x(v)$. \square

The existence of factor frequencies is very susceptible to failure when words are drawn from minimal shift spaces that are not uniquely ergodic. Note that the set of all invariant measures of (X, T) form a convex polytope whose extremal points are the

ergodic measures [22, Prop. 3.8.10]. In particular, if (X, T) is not uniquely ergodic, it will have at least two distinct ergodic measures.

LEMMA 3.3 (THM. 1.1 IN [16]). *Let X be a minimal shift space which is not uniquely ergodic. There exists a dense set $B \subseteq X$ of words that are not generic for any invariant measure, and hence do not admit factor frequencies.*

Remark 3.4. To summarise the discussion on the existence of factor frequencies, for a minimal shift X , we have the following. If X is uniquely ergodic, then all $x \in X$ admit (uniform and computable) factor frequencies, and moreover agree on the frequencies of each factor. On the other hand, if X is not uniquely ergodic, then it will have a dense subset of words which do not admit factor frequencies, and a dense (by virtue of including the orbit of any generic point) subset of generic points for its ergodic measures. These generic points do admit factor frequencies, but they do not agree on the frequencies of all factors because they are generic for different measures.

Let us now explain how to obtain the counterexamples to the preservation of unique ergodicity alluded to in the Introduction.

LEMMA 3.5. *There exists a uniformly recurrent word $x \in \{0, 1\}^\omega$ that admits factor frequencies, but $\mathcal{A}(x)$ does not, where \mathcal{A} is the two-state transducer whose underlying automaton changes the state on reading 1, stays in the same state on reading 0, and prints (q, b) on reading b , where q is the current state.*

PROOF. We shall choose x to be the symbolic trace of a carefully constructed topological dynamical system. Chaika [13] constructs a minimal and uniquely ergodic dynamical system (X', T') as follows: the space X' is the unit circle identified by the interval $[0, 1)$, and T' maps ζ to $\zeta + \alpha$, where $\alpha < 1/3$ is an irrational number. The invariant measure μ' is the usual Lebesgue measure. There is a distinguished set I which is the union of the intervals $[0, \gamma)$ and $[\alpha + \beta, \alpha + \beta + \gamma)$, where β, γ are constants depending on α . The dynamics of the skew product $Y' = (\mathbb{Z}_2 \rtimes X', T')$ maps (b, ζ) to $(b + 1_I(\zeta), \zeta + \alpha)$, where $1_I(\zeta)$ is 1 if $\zeta \in I$, and 0 otherwise. This skew product is minimal, but admits precisely two distinct ergodic measures ν'_0, ν'_1 .

For each ζ , we define its symbolic trace $x = \text{tr}(\zeta)$ as $x(n) = 1_I(\zeta + n\alpha)$. We claim that $X = \{\text{tr}(\zeta) \mid \zeta \in X'\}$ is a minimal shift. To that end, for each $u \in \{0, 1\}^+$, we shall inductively define sets I_u with the property that u occurs at index n in $\text{tr}(\zeta)$ if and only if $\zeta + n\alpha \in I_u$. Clearly, $I_1 = I$, and $I_0 = [0, 1) \setminus I$, and $I_{bu} = I_b \cap (-\alpha + I_u)$. By examining the orientation of the intervals, we have that each I_u is either empty, or contains a non-empty open set. Thus, $\text{tr}(\zeta)$ is uniformly recurrent, and furthermore $\mathcal{L}(\text{tr}(\zeta))$ is the set of factors u such that I_u is non-empty (we take I_ϵ to be $[0, 1)$). It follows that every $\text{tr}(\zeta)$ generates the same shift X , and that $f_{\text{tr}(\zeta)}(u) = \mu'(I_u)$. The unique ergodicity of X follows readily from Lem. A.2 in the appendix.

We now define the shift Y to be generated by the run of \mathcal{A} on some $\text{tr}(\zeta)$. We claim that Y is minimal, projects to X , and in fact the projection map is two-to-one. Minimality can be seen either by alternately defining Y via Y' similarly as above, or directly from the uniform recurrence of Y as given by Lem. 5.2. Now, Y is (isomorphic to) a closed subset of $\mathbb{Z}_2 \times X$. Since \mathbb{Z}_2 is compact, projection is a

closed map, and hence the projection of Y onto X is a closed set which is moreover shift-invariant. It follows that this projection must be X itself. Since Y includes the set of symbolic traces of the minimal skew product $Y' = \mathbb{Z}_2 \rtimes X'$, it must in fact be isomorphic to $\mathbb{Z}_2 \times X$.

Finally, we show that Y admits two distinct invariant measures ν_0, ν_1 . For this, observe that $\text{tr}: Y' \rightarrow Y$ is measurable. Indeed, the subsets B of Y for which $\text{tr}^{-1}(B)$ is Lebesgue-measurable form a σ -algebra which includes the cylinders of Y as they are images of measurable sets $\{b\} \times I_u$. Thus, the ergodic measures ν'_0, ν'_1 of Y' project to invariant measures $\nu_b = \nu'_b \circ \text{tr}^{-1}$.

To complete the proof, we record that ν_0, ν_1 are indeed distinct because ν'_0, ν'_1 are such that $\nu'_b(b \times B) = \mu'(B)$, i.e., each measure concentrates its mass in one copy of X' [13, Prop. 3].

We thus have that Y is a shift that is not uniquely ergodic, and contains the runs of \mathcal{A} on all $x \in X$. In particular, by Lem. 3.3, there exists some counterexample $y = \mathcal{A}(x)$ that does not admit factor frequencies despite x doing so. \square

The natural questions that follow are: can one identify when a shift is uniquely ergodic? Is there a stronger condition on X which precludes the misbehaviour of Y described above? Boshernitzan [10] gave a sufficient (but not necessary) condition in answer to the former. In this paper, we shall show that this condition answers the latter in the affirmative.

Definition 3.6 (Condition (B)). Let $X \subseteq \Sigma^\omega$ be a minimal shift space. For an invariant measure μ , define $\kappa: \mathbb{N} \rightarrow \mathbb{R}$ as

$$\kappa(n) = \min_{u \in \mathcal{L}_n(X)} (n \cdot \mu([u]_X)).$$

The minimal shift space $X \subseteq \Sigma^\omega$ satisfies Boshernitzan's condition (also referred to as Condition (B)) if it admits an invariant measure μ such that the corresponding κ satisfies $\limsup_{n \rightarrow \infty} \kappa(n) > 0$. A uniformly recurrent word $x \in \Sigma^\omega$ satisfies Boshernitzan's condition if it generates a shift space X that satisfies Boshernitzan's condition.

LEMMA 3.7 (BOSHERNITZAN [10]). *A minimal shift space $X \subseteq \Sigma^\omega$ that satisfies Boshernitzan's condition is uniquely ergodic. If a word $x \in \Sigma^\omega$ satisfies Boshernitzan's condition, then the frequency of every factor u of x is well defined and uniform.*

Remark 3.8. If a word x is linearly recurrent, then it satisfies Boshernitzan's condition [18, Thm. 15]. And even, linear recurrence for a minimal shift space is equivalent to $\liminf \kappa(n) > 0$ by a result due to Boshernitzan (see [36, Exercise 174]). By Rmk. 2.3 below, Boshernitzan's condition is also satisfied by primitive morphic words. The work of Damanik and Lenz [14] is important from the perspective of applicability: they show that Sturmian words satisfy Boshernitzan's condition, as do almost all words coding the orbits of interval-exchange transformations [15, Thm. 5], a large class of 1-dimensional toric words [15, Sec. 5], and almost all Arnoux-Rauzy words [15, Thm. 12].

4 THE KROHN-RHODES FRAMEWORK

In this paper, we shall work with deterministic automata and transducers. Henceforth, when we refer to automata and transducers, we mean the deterministic models of computation described below.

A transducer \mathcal{A} is given by $(Q, q_{\text{init}}, \Sigma, \Gamma, \delta^o, \delta^t)$, where Q is a finite non-empty set of states, $q_{\text{init}} \in Q$ is the initial state, Σ is the input alphabet, Γ is the output alphabet, $\delta^t: Q \times \Sigma \rightarrow Q$ is the transition function, and $\delta^o: Q \times \Sigma \rightarrow \Gamma^*$ is the output function. For convenience, we shall extend the domain of δ^t, δ^o to $Q \times \Sigma^*$ in the obvious way: $\delta^t(q, \varepsilon) = q$ and $\delta^o(q, \varepsilon) = \varepsilon$ for all q , while $\delta^t(q, ua) = \delta^t(\delta^t(q, u), a)$ and $\delta^o(q, ua) = \delta^o(q, u) \cdot \delta^o(\delta^t(q, u), a)$. A *uniform* transducer is one where we have that for all q, a , we have $|\delta^o(q, a)|$ is the same: if the common length is l , we have an l -uniform transducer. A *non-erasing* transducer is one where for all $q \in Q, a \in \Sigma$ we have $\delta^o(q, a) \neq \varepsilon$.

Automata are special cases of 1-uniform transducers where $\Gamma = Q \times \Sigma$, and δ^o is the identity function, and we therefore present them simply as $(Q, q_{\text{init}}, \Sigma, \delta)$. We shall denote the output of a transducer \mathcal{A} upon reading a word x (finite or infinite) as $\mathcal{A}(x)$. When \mathcal{A} is an automaton, we call the sequence $\mathcal{A}(x)$ of state-letter pairs the *run* of \mathcal{A} on x , because indeed if we denote $\mathcal{A}(x)(n) = (q_n, a_n)$, then for every n we have $\delta(q_n, a_n) = q_{n+1}$.

Given a transducer \mathcal{A} , we shall denote its *underlying automaton* (obtained as described above) by \mathcal{A}_0 . We observe that we can always define a substitution $\sigma: (Q \times \Sigma) \rightarrow \Gamma^*$ such that for all x , we have that $\mathcal{A}(x) = \sigma(\mathcal{A}_0(x))$. The following lemma assures us that it suffices to focus only on automata to prove our results.

LEMMA 4.1. *Let $\sigma: \Sigma \rightarrow \Gamma^*$ be a substitution, and let $x \in \Sigma^\omega$ be an infinite word. If x has one of the following properties, then $\sigma(x)$ also has that property, provided it is an infinite word.*

- (1) *primitive morphic*
- (2) *recurrent*
- (3) *(effectively) uniformly recurrent*
- (4) *(effectively) linearly recurrent*
- (5) *uniformly recurrent and admits factor frequencies*
- (6) *satisfies Boshernitzan's condition*

In item (5), if x is effectively uniformly recurrent and admits computable factor frequencies, then $\sigma(x)$ also has these properties.

PROOF. Item (1) follows immediately by definition (see Rmk. 2.2).

Items (2) through (4) have similar proofs. To formalise the proof, consider the *attribution function* α from indices of $\sigma(x)$ to indices of x . We define $\alpha(j)$ to be the smallest index i such that $j < |\sigma(x(0, i+1))|$, i.e., the letter at index j of $\sigma(x)$ is attributed to the image of $x(i)$. We define $\alpha(J) = \bigcup_{j \in J} \alpha(j)$. In particular, when v is a factor of $\sigma(x)$ at index j , then we set $J = [j, j + |v|)$, obtain $[i, i')$ as the smallest interval that contains $\alpha(J)$, and say that the above occurrence of v is attributed to an occurrence of $u = x(i, i')$ in x at index i .

Clearly, if an occurrence of v is attributed to an occurrence of u , then the recurrence of u implies that of v . This proves item (2).

In case x is uniformly recurrent, an occurrence of v will always be attributed to a factor u that is at most $R_x(1) \cdot |v|$ in length. This is because the gaps between occurrences of non-erasing letters (letters a for which $\sigma(a)$ is not the empty word) are bounded by $R_x(1)$. The gaps between the occurrences of u are themselves bounded by $R_x(R_x(1) \cdot |v|)$. Finally, the gaps between the occurrences of v must be bounded by $(\max_{a \in \Sigma} |\sigma(a)|) \cdot R_x(R_x(1) \cdot |v|)$. This proves item (3), and also item (4): if $R_x(n)$ is linear, then the bound on the gaps between occurrences of v will also be linear in $|v|$.

We prove item (5) in two steps. As in Rmk. 2.2, we decompose σ as $\tau \circ \hat{\sigma}$, where τ is a letter-to-letter substitution, and $\hat{\sigma}$ maps each letter a to $|\sigma(a)|$ copies of a , i.e., $\hat{\sigma}(a) = (a, 0)(a, 1) \cdots (a, |\sigma(a)| - 1)$. We denote $\hat{\sigma}(x)$ by z , and $\tau(z)$, which is the same as $\sigma(x)$, by y . Observe by the preceding discussion that z, y are uniformly recurrent. We shall show that z admits factor frequencies, and in turn so does y .

By uniform recurrence, we have that any factor w of z can be attributed to one of finitely many factors $\{u_1, \dots, u_d\}$ of x , and this factor will have length at most $R_x(1) \cdot |w|$. By the definition of attribution, there is no pair u_i, u_j such that u_i is a factor of u_j (in particular, the cylinders $[u_1]_X, \dots, [u_d]_X$ are disjoint); moreover by the positional-encoding nature of the construction, each u_i , upon restricting to non-erasing letters, will give a word u such that $\hat{\sigma}(u)$ contains exactly one occurrence of w (and conversely, the set of factors which begin and end with a non-erasing letter, and give u upon restricting to non-erasing letters is $\{u_1, \dots, u_d\}$). We thus obtain $f_z(w) = f_z(\hat{\sigma}(u))$. We claim that $f_z(\hat{\sigma}(u)) = F/L$, where $F = \sum_{i=1}^d f_x(u_i)$, and L is the non-zero constant $\sum_{a \in \Sigma} f_x(a) |\hat{\sigma}(a)|$. Indeed, $L \geq 1/R_x(1)$, and is hence non-zero.

We note that $F = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=1}^d \mathbf{1}_{u_j}(T^i x)$ by Lem. 3.1. We also have that $L = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} |\sigma(x(i))|$ (consider the function that maps x to $|\sigma(x(0))|$). We rephrase these observations by taking appropriate subsequences of Birkhoff averages as follows. Let $(\beta_N)_{N \in \mathbb{N}}$ denote the strictly increasing sequence whose N -th element β_N is the index (in x) of the N -th occurrence of an element of $\{u_1, \dots, u_d\}$. We have in particular that $\lim_{N \rightarrow \infty} \frac{N}{\beta_N} = F$. Similarly, we define the strictly increasing sequence $(\gamma_N)_{N \in \mathbb{N}}$ of indices of occurrences of $\hat{\sigma}(u)$ in z as $\gamma_N = \sum_{i=0}^{\beta_N-1} |\sigma(x(i))|$, and observe that $\lim_{N \rightarrow \infty} \frac{\gamma_N}{\beta_N} = L$. Since $L \neq 0$, we obtain that $\lim_{N \rightarrow \infty} \frac{N}{\gamma_N} = \frac{F}{L}$. Denoting $U = \{\gamma_N \mid N \in \mathbb{N}\}$, it remains to observe (by elementary analytic means) that $\lim_{M \rightarrow \infty} \frac{|U \cap [0, M)|}{M} = \lim_{N \rightarrow \infty} \frac{N}{\gamma_N} = \frac{F}{L}$.

Indeed, the sequences agree whenever $M = \gamma_N$. The sequence on the left decreases in the interim, and only increases when M steps from $\gamma_N - 1$ to γ_N . A simple calculation shows that the increment is bounded by $1/(\gamma_N - 1)$, i.e., asymptotically vanishes. It then follows that the sequences converge to the same limiting value F/L .

The proof that y inherits the property of admitting factor frequencies from z is simpler because τ is letter-to-letter. For any factor v , there are finitely many factors w_1, \dots, w_d of the same length such that $\tau(w_i) = v$. The frequency $f_y(v)$ is simply $\sum_i f_z(w_i)$.

To prove item (6)², observe that

$$\min_{v \in \mathcal{L}_n(y)} f_y(v) \geq \frac{1}{L} \min_{u \in \mathcal{L}_{n \cdot R}(x)} f_x(u),$$

where $R = R_x(1)$. Writing $\kappa'(n)$ as shorthand for $\min_{u \in \mathcal{L}} f(u)$, using X, Y to denote the shifts generated by x, y , and defining μ_X, μ_Y using f_x, f_y , we get $L(R+1)n \cdot \kappa'_Y(n) \geq \max_{0 \leq r < R} (nR+r) \cdot \kappa'_X(nR+r)$. This allows us to conclude that $\limsup_{n \rightarrow \infty} \kappa_Y(n) > 0$, given $\limsup_{n \rightarrow \infty} \kappa_X(n) > 0$, thus establishing the preservation of Boshernitzan's condition. \square

We record an interesting corollary of the above proof.

²The proof of [14, Thm. 8(a)], which states item (6), implicitly assumes the substitution is non-erasing.

COROLLARY 4.2. *Let σ be a non-erasing substitution. If x admits (computable) factor frequencies, then so does $\sigma(x)$.*

Our main tool to prove preservation results for automata will be the Krohn-Rhodes theorem, which decomposes automata into simple, “well-behaved” components that are connected “in series”. The strategy is then to prove preservation results for these simpler automata. The Krohn-Rhodes theorem is often stated for semigroups [29]; see [33, Thm. 1] for a formulation in terms of automata (see also [31, Thm. 3]). We now introduce the terminology required to state the theorem.

Given an automaton $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \delta)$, for every $u \in \Sigma^*$, we define $\delta_u: Q \rightarrow Q$ to be the transition function induced by u , i.e., $\delta_u(q) = \delta(q, u)$. An automaton \mathcal{A} is called a *reset automaton* if for every $a \in \Sigma$, the function δ_a is either the identity function, or constant-valued. Intuitively, a reset automaton is a flip-flop whose memory records the last reset. An automaton \mathcal{A} is called a *permutation automaton* (also a *group automaton*) if for every $a \in \Sigma$, δ_a defines a permutation of Q , i.e., is bijective. The automaton \mathcal{A} is called *counter-free* if there does not exist $u \in \Sigma^+$, $Q' \subseteq Q$ such that δ_u induces a non-trivial permutation on Q' . For the sake of brevity, we declare that permutation, reset, and counter-free transducers are respectively those whose underlying automata are permutation, reset, and counter-free automata. We remark that counter-free automata are of particular interest because of their equivalence with first-order-definable (aperiodic) languages, as established by the McNaughton–Papert theorem [32].

We say that an automaton $\mathcal{A}' = (Q', q'_{\text{init}}, \Sigma, \delta')$ covers an automaton $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \delta)$ if there exists a map φ from Q' to Q that respects the initial states and commutes with the transition relations, i.e., $\varphi(q'_{\text{init}}) = q_{\text{init}}$ and for every $q' \in Q'$, $a \in \Sigma$, we have that $\varphi(\delta'(q', a)) = \delta(\varphi(q'), a)$. This implies, in particular that for a word u over Σ , its runs v, v' over $\mathcal{A}, \mathcal{A}'$ respectively are related by $v = \sigma'(v')$ where σ' is a letter-to-letter substitution.

If $\mathcal{A}_1 = (Q_1, \Sigma, \delta_1)$, and $\mathcal{A}_2 = (Q_2, Q_1 \times \Sigma, \delta_2)$, then the *cascade* $\mathcal{A}_2 \circ \mathcal{A}_1$ is an automaton over Σ with states $Q_2 \times Q_1$ defined by the following property. The run of a word u is a word v' over $(Q_2 \times Q_1 \times \Sigma)$: its projection v onto $Q_1 \times \Sigma$ is the run of \mathcal{A}_1 on u , and v' is the run of \mathcal{A}_2 on v . Formally, the transition function is given by

$$\delta((q_2, q_1), a) = (\delta_1(q_1, a), \delta_2(q_2, (q_1, a))).$$

The cascade $\mathcal{A}_k \circ \dots \circ \mathcal{A}_1$ is implemented by performing the rightmost cascade first, akin to function composition.

THEOREM 4.3 (KROHN-RHODES [33]). *For every automaton \mathcal{A} , we can compute a cascade $\mathcal{A}' = \mathcal{B}_k \circ \dots \circ \mathcal{B}_1$ such that:*

- (1) \mathcal{A}' covers \mathcal{A} .
- (2) Each \mathcal{B}_i is either a permutation automaton or a two-state reset automaton.
- (3) If \mathcal{B}_i is a permutation automaton, then its transition group is homomorphic to a subgroup of the transition monoid of \mathcal{A} .

In particular, if \mathcal{A} is a counter-free automaton, then each \mathcal{B}_i is a reset automaton.

Remark 4.4. Our proofs of preservation theorems for transducer outputs will follow the following template: (i) prove the theorem for reset automata and permutation automata; (ii) deduce that the

theorem holds for any cascade of these special automata; (iii) apply Lem. 4.1 (we only need the easy case of σ being a coding here) and deduce that the theorem holds for any automaton covered by such a cascade; (iv) apply the Krohn-Rhodes theorem and deduce that the theorem indeed holds for all automata; (v) apply Lem. 4.1 to deduce the theorem holds for all transducers. In the above, only Step (i) will require work.

5 PRESERVATION OF RECURRENCE

Semënov showed that if a word $x \in \Sigma^\omega$ is *effectively almost-periodic* (i.e., the word x is effective, and given any $u \in \Sigma^+$, it can be decided whether u occurs infinitely often in x , and if it is the case, we can compute $\mathcal{R}_x(u)$), then for any transducer \mathcal{A} , the word $\mathcal{A}(x)$, if infinite, is also effectively almost-periodic (see [34, Sec. 3] for an exposition of that result in English). Using this result and techniques from the proof, Pritykin showed that if a word $y \in \Sigma^\omega$ has a uniformly recurrent suffix x , then for any transducer \mathcal{A} , the word $\mathcal{A}(y)$, if infinite, also has a uniformly recurrent suffix [35]. We use the Krohn-Rhodes theorem to prove (in an arguably simpler and more insightful manner) the following strengthening. We remark that the Krohn-Rhodes theorem is crucially needed in the case where we do not assume uniform recurrence.

THEOREM 5.1. *Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \Gamma, \delta^o, \delta^t)$ be a transducer. Consider a word $x \in \Sigma^\omega$. If x has one of the following properties, then $\mathcal{A}(x)$ also has the same property, provided that it is an infinite word.*

- (1) *recurrent suffix*
- (2) *(effectively) uniformly recurrent suffix (whose starting index is computable)*
- (3) *(effectively) linearly recurrent suffix (whose starting index is computable)*

By Rmk. 4.4, the preservation theorem above would follow from preservation lemmas for permutation automata and reset automata. Observe that in the proofs, we can assume without losing generality that the input word x is recurrent by simply considering the run starting at the beginning of the recurrent suffix. The following distills the core idea of Semënov’s argument.

LEMMA 5.2. *Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \delta)$ be a permutation automaton. Consider a word $y \in \Sigma^\omega$. If x has one of the following properties, then $\mathcal{A}(x)$ also has that property.*

- (1) *recurrent*
- (2) *(effectively) uniformly recurrent*
- (3) *(effectively) linearly recurrent*

PROOF. We prove the preservation of recurrence, and obtain preservation of the other properties as corollaries of the proof. We shall consider an arbitrary occurrence of a factor v of $\mathcal{A}(x)$, and prove that v occurs infinitely often by showing that we can find another occurrence to the right of the one under consideration.

To that end, let us establish some notation. We denote the projection of $v = v_0$ onto Σ by u_0 , and denote by i_0 the index of the occurrence of v_0 in $\mathcal{A}(x)$ (which is the same as that of the corresponding occurrence of u_0 in x). If the first letter of v_0 is $(q, u_0(0))$ and $\delta_{u_0}(q) = q_0$, then we use (u_0, q_0) as shorthand for v_0 . The shorthand indeed defines v_0 because \mathcal{A} is a permutation automaton, and hence for every word u , δ_u admits an inverse, which we shall denote

by δ_u^{-1} . In other words, the transition function is “reversible”: if we know the state after reading u , we can determine all the states along the run of u .

We inductively define sequences $(u_k)_k$ of factors of x , $(v_k)_k$ of factors of $\mathcal{A}(x)$, $(i_k)_k$ of indices, and $(q_k)_k$ of states. We shall maintain that the projection of each v_k onto Σ is the corresponding u_k . We always have that u_k has infinitely many occurrences in x , and we will maintain that u_k occurs in x at index i_0 . We find its next occurrence at index $i_{k+1} > i_0$, and define $u_{k+1} = x(i_0, i_{k+1} + |u_k|)$ to be the *extended return word* that spans consecutive occurrences of u_k . Correspondingly, we define $v_{k+1} = \mathcal{A}(x)(i_0, i_{k+1} + |u_k|)$. This word v_{k+1} will be denoted by (u_{k+1}, q_{k+1}) . In particular, the word u_k is a prefix of u_{k+1} .

By the pigeonhole principle, there exist distinct l, k with $0 \leq l < k \leq |Q|$ such that $q_l = q_k$. By construction, we also have that u_l is both a prefix and a suffix of u_k . This means that we have found two occurrences of $v_l = (u_l, q_l)$, one at index i_0 , and the other at index $i_{k+1} + |u_k| - |u_l|$. Our construction also ensures in particular that u_0 is a prefix of u_l , and since the transition function is reversible, we have found an occurrence of v_0 at index $i_{k+1} + |u_k| - |u_l|$, which is to the right of i_0 by at most $|u_{k+1}|$.

We have thus far proven that $\mathcal{A}(x)$ is recurrent. In the case of (effective) uniform recurrence, we observe that $i_{k+1} - i_0 \leq R_x(|u_k|)$, and thus $|u_{k+1}| \leq R_x(|u_k|) + |u_k|$. For convenience, we define the auxiliary function W_x (read as “window function”) as $W_x(n) = R_x(n) + n$. From the discussion above, we get that $R_{\mathcal{A}(x)}(n) \leq W_x^{(|Q|+1)}(n)$, where $W_x^{(k)}$ denotes the k -fold composition of W_x . Clearly, if R_x is computable (respectively, linear), then so is $R_{\mathcal{A}(x)}$.

Finally, we check that in the case of uniform recurrence, the index of the first occurrence of v is bounded by $R_{\mathcal{A}(x)}(|v|)$. This is indeed the case, because if i_0 exceeds the above bound, we can apply the above argument *mutatis mutandis* to find an occurrence of v to the left of i_0 . This completes the proof. \square

LEMMA 5.3. *Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \delta)$ be a reset automaton. Consider a word $x \in \Sigma^\omega$. If x has one of the following properties, then $\mathcal{A}(x)$ also has that property.*

- (1) *recurrent suffix*
- (2) *(effectively) uniformly recurrent suffix (whose starting index is computable)*
- (3) *(effectively) linearly recurrent suffix (whose starting index is computable)*
- (4) *uniformly recurrent suffix and admits factor frequencies*

In item (4), if the suffix of x is effectively uniformly recurrent, its starting index is computable, and x admits computable factor frequencies, then $\mathcal{A}(x)$ also has these properties.

PROOF. We shall assume without losing generality that x is recurrent by simply considering the run starting at the beginning of its recurrent suffix. The lemma is trivial if x does not contain any recurrent reset letters, i.e., letters a such that δ_a is a constant-valued function. We shall therefore assume that x has a recurrent reset letter a , whose first occurrence is at index N , which can be computed in the case of effective uniform recurrence. We claim that the suffix of $\mathcal{A}(x)$ starting at index $N + 1$ is recurrent.

The key will be to focus on return words to the reset letter a . By definition, these words begin with a , and have exactly a single

occurrence of a because a is a single-letter word. We have that $x(N, \infty) = r_0 r_1 r_2 \dots$, a concatenation of return words.

As before, we shall show that for any occurrence of an arbitrary factor v at index $i_0 > N$, we can find another occurrence at an index to the right of i_0 . Let u be the projection of v onto Σ^+ , i.e., $x(i_0, i_0 + |v|) = u$. We observe by the above factorisation of $x(N, \infty)$ into return words that there exists an index i with $N \leq i < i_0$ such that $x(i) = a$. We take i to be maximal, and in particular we have $i_0 - i < R_x(1)$. Let u' denote the factor $x(i, i_0 + |v|)$. By recurrence, u' will have another occurrence at an index $i' > i$, and this will lead to an occurrence of v in $\mathcal{A}(x)$ at index $i' - i + i_0$. This proves that $y = \mathcal{A}(x)(N + 1, \infty)$ is recurrent, establishing item (1).

If x is uniformly recurrent, then $|i' - i| \leq R_x(|v| + R_x(1))$, and hence $R_y(n) \leq R_x(n + R_x(1))$. Clearly R_y is effective (respectively, linear) if R_x is. We also check that if $i_0 - N$ exceeds $R_x(n + R_x(1))$, then we can use the same argument as above to find an occurrence of v in $\mathcal{A}(x)$ to the left of i_0 . This proves items (2) and (3).

To study factor frequencies of v , we partition the occurrences of u , and account precisely for which partitions result in an occurrence of v . Formally, we construct a tree whose root is u , vertices are factors u' of the form wu , leaves are factors u' that have u as a proper suffix and begin with a reset letter, and the successors of the internal nodes u' are of the form au' . The depth of the tree is thus at most $R_x(1)$. The frequency $f_y(v)$ is simply $\sum_{u'} f_x(u')$, where the summation ranges over leaves u' such that the run of the reset automaton on the word u' has v as its suffix. \square

Remark 5.4. In the above proof, consider the case x is uniformly recurrent, and let y be the recurrent suffix of $\mathcal{A}(x)$. Observe that $\min_{v \in \mathcal{L}_n(y)} f_y(v) \geq \min_{u \in \mathcal{L}_{n+R_x(1)}(x)} f_x(u)$.

Lem. 5.2, Lem. 5.3, and Rmk. 4.4 prove Thm. 5.1. In fact, item (4) of Lem. 5.3, along with the counter-free case of Thm. 4.3, gives the following result.

THEOREM 5.5. *Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \Gamma, \delta^0, \delta^t)$ be a counter-free transducer. Let $x \in \Sigma^\omega$ be uniformly recurrent and admit factor frequencies. We have that if $\mathcal{A}(x)$ is an infinite word, then it has a suffix y which is uniformly recurrent and admits factor frequencies. When x is effectively uniformly recurrent and admits computable factor frequencies, the suffix y also has these properties, and its starting index in $\mathcal{A}(x)$ can be computed.*

6 PRESERVATION OF SELF-SIMILARITY

In this section, we show that transducers preserve properties that indicate self-similarity³.

LEMMA 6.1. *Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \Gamma, \delta^0, \delta^t)$ be a transducer, let S be a set of substitutions, let $x \in \Sigma^\omega$ be an S -adic word directed by the sequence $(\sigma_n)_{n=0}^\infty$. We have that the word $\mathcal{A}(x)$, if infinite, is \hat{S} -adic for a set \hat{S} of substitutions defined using only S , and the directive sequence $(\hat{\sigma}_n)_{n=0}^\infty$ can be defined using only \mathcal{A} and $(\sigma_n)_{n=0}^\infty$.*

³Every word $x = a_0 a_1 a_2 \dots$ is vacuously S -adic for $S = \{\sigma_1, \dots, \sigma_{|\Sigma|}, \tau_1, \dots, \tau_{|\Sigma|}\}$ over the alphabet $\Sigma \cup \{b\}$, where σ_i replaces the distinguished letter b with $a_i \in \Sigma$ and is identical elsewhere, and τ_i replaces b with ba_i and is identical elsewhere. We then see that x is directed by $\tau_{a_0} \sigma_{a_1} \sigma_{a_2} \dots$. The point of directive sequences is to encode more information about self-similar structure, and the lemma conveys the sense in which this structure is preserved. In particular, it allows us to prove that if x is morphic then so is $\mathcal{A}(x)$.

PROOF. We can assume without losing generality that $\mathcal{A} = \mathcal{A}_0$ is an automaton (Lem. 4.1). Let $x_0 = x, x_1, \dots$ be the sequence of words such that for all n , $x_n \in \Sigma_n^\omega$ and $x_n = \sigma_n(x_{n+1})$. To show that the run $y_0 = \mathcal{A}_0(x_0)$ is δ -adic, we shall construct a sequence $(\mathcal{A}_n)_{n=0}^\infty$ of automata, where $\mathcal{A}_n = (Q, q_{\text{init}}, \Sigma_n, \delta_n)$, then let $y_n = \mathcal{A}_n(x_n)$, and finally define a sequence of substitutions $(\hat{\sigma}_n)_{n=0}^\infty$ such that for all n , $y_n = \hat{\sigma}_n(y_{n+1})$.

The key idea is that x_{n+1} can be intuited as a “compression” of x_n using σ_n . Hence, in \mathcal{A}_{n+1} , we define δ_{n+1} such that for all $a \in \Sigma_{n+1}$, $\delta_{n+1}(q, a) = \delta_n(q, \sigma_n(a))$. Thus, the run y_{n+1} can be regarded as a “fast-forwarded” version of the run y_n . We are therefore motivated to define $\hat{\sigma}_n$ as the dual “slow-motion” operator, i.e., $\hat{\sigma}_n((q, a))$ gives the run of \mathcal{A}_n on $\sigma_n(a)$ starting in state q . It is now easy to check that $\mathcal{A}_n(x_n) = \hat{\sigma}_n(\mathcal{A}_{n+1}(x_{n+1}))$ for all n . \square

THEOREM 6.2. Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \Gamma, \delta^0, \delta^t)$ be a transducer, and let $x \in \Sigma^\omega$ be a primitive morphic word given as the image under τ of a particular fixed point of a primitive substitution σ . When $\mathcal{A}(x)$ is an infinite word:

- (1) The word $\mathcal{A}(x)$ has a primitive morphic suffix y . We can compute the index at which y begins, and compute substitutions $\hat{\sigma}, \hat{\tau}$ such that $\hat{\sigma}$ is primitive and y is the image under $\hat{\tau}$ of a fixed point of $\hat{\sigma}$.
- (2) The word $\mathcal{A}(x)$ admits computable factor frequencies.

PROOF. We note that item (2) follows from item (1) by applying Lem. 2.4. We therefore focus on proving item (1). We remark that it suffices to give a proof assuming \mathcal{A} is an automaton.

We first prove that $\mathcal{A}(x)$ is morphic, and we will follow the construction given in the proof of Lem. 6.1 in order to do so. The directive sequence of x is $\tau\sigma^\omega$, and in particular, for all $n \geq 1$, $\Sigma_n = \Sigma_1$ and $x_n = x_1$. Observe that this implies that the set $\{\mathcal{A}_n \mid n \in \mathbb{N}\}$ is finite and can be effectively enumerated. Furthermore, when $n \geq 1$, \mathcal{A}_{n+1} and δ_n will depend only on \mathcal{A}_n and σ . By the pigeonhole principle, there exist computable m, n such that $m < n$ and $\mathcal{A}_m = \mathcal{A}_n$. We have that $\mathcal{A}_m(x)$ is a fixed point of $\hat{\sigma}_m \cdots \hat{\sigma}_{n-1}$, and its image under $\hat{\sigma}_0 \cdots \hat{\sigma}_{m-1}$ gives $\mathcal{A}(x)$, which is hence morphic.

Now, since x is primitive morphic, it is effectively linearly recurrent (Lem. 2.4). The run $\mathcal{A}(x)$ therefore has a suffix y that is effectively linearly recurrent, and the starting index N of this suffix can be computed (Thm. 5.1). This suffix y of the morphic word $\mathcal{A}(x)$ is again morphic, and the substitutions describing it can be computed (Rmk. 2.3). Having proved that y is an effectively linearly recurrent morphic word, we invoke Rmk. 2.3 again, this time to deduce that y is indeed primitive morphic, and we can compute substitutions $\hat{\sigma}, \hat{\tau}$ such that $\hat{\sigma}$ is primitive, and y is the image under $\hat{\tau}$ of a fixed point of $\hat{\sigma}$. \square

For the sake of completeness, we comment on the case of *automatic* words. Recall that a word $x \in \Gamma^\omega$ is automatic [2, Chap. 5] if for some $k \geq 2$, there exists an automaton that computes $x(n)$ when given the k -ary representation of n as input: we then say that x is k -automatic. Clearly, for $a \geq 1$, we have x is k^a -automatic if and only if x is k -automatic. Equivalently, by a result due to Cobham [2, Thm. 6.3.2], a word x is k -automatic if and only if it is the image under a coding of a fixed point of a k -uniform morphism σ (for every $a \in \Sigma$, $|\sigma(a)| = k$). Note that the definition can be relaxed

(analogously to Rmk. 2.2) to replace the coding by a k' -uniform morphism [2, Cor. 6.8.3]. Our proof of Lem. 6.1 can thus be adapted to deduce the main result of [2, Sec. 6.9]:

LEMMA 6.3. If \mathcal{A} is a uniform transducer and x is k -automatic, then $\mathcal{A}(x)$ is k -automatic.

7 PRESERVATION OF CONDITION (B)

In this section, we prove the following main result using the Krohn-Rhodes theorem. We give a general statement, but observe that by Rmk. 4.4, it suffices to prove the result separately for reset automata and permutation automata.

THEOREM 7.1. Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \Gamma, \delta^0, \delta^t)$ be a transducer, and let $x \in \Sigma^\omega$ be a/an (effectively) uniformly recurrent word that satisfies Boshernitzan’s condition. We have that the word $\mathcal{A}(x)$, if infinite, has a/an (effectively) uniformly recurrent suffix y that satisfies Boshernitzan’s condition, and hence admits (computable) factor frequencies.

We already have argued the existence of a/an (effectively) uniformly recurrent suffix y in Thm. 5.1. We shall prove that this suffix satisfies Boshernitzan’s condition. The fact that y admits (computable) factor frequencies would then follow immediately from (Lem. 3.2 and) Lem. 3.7.

LEMMA 7.2. Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \delta)$ be a reset automaton and let $x \in \Sigma^\omega$ be a uniformly recurrent word that satisfies Boshernitzan’s condition. We have that the word $\mathcal{A}(x)$ has a uniformly recurrent suffix y that satisfies Boshernitzan’s condition.

PROOF. Recall from Lem. 5.3 that $\mathcal{A}(x)$ has a uniformly recurrent suffix y that admits factor frequencies. We can define an invariant measure μ on the minimal shift Y defined by y , as $\mu([v]_Y) = f_y(v)$ (the invariant measure on the shift defined by x is similarly defined). Let us now use Rmk. 5.4 to argue that Y satisfies Boshernitzan’s condition (Def. 3.6) by virtue of admitting the above invariant measure. We have from Rmk. 5.4 that for n large enough, $\min_{v \in \mathcal{L}_n(y)} f_y(v)$ is lower bounded by $\min_{u \in \mathcal{L}_{2n}(x)} f_x(u)$ as well as $\min_{u \in \mathcal{L}_{2n+1}(x)} f_x(u)$, or in other words

$$3 \cdot \kappa_Y(n) \geq \max(\kappa_X(2n), \kappa_X(2n+1)).$$

This implies that $\limsup_{n \rightarrow \infty} \kappa_Y(n) > 0$, and the word y indeed satisfies Boshernitzan’s condition. \square

The case where \mathcal{A} is a permutation automaton is much more involved, and we need to invoke arguments from cohomology and topological dynamics inspired by [9]. We devote the rest of this section to the proof of the following lemma.

LEMMA 7.3. Let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \delta)$ be a permutation automaton. If x is a uniformly recurrent word that satisfies Boshernitzan’s condition, then so is $\mathcal{A}(x)$.

The group G generated by the transitions of the permutation automaton, and the natural onto morphism $\varphi: \Sigma^* \rightarrow G$ (the morphism maps a word u to the element corresponding to δ_u) will play a key role in the proof. Let X be the minimal shift defined by x , let μ_X be its unique invariant measure, and recall that by Def. 3.6, $\limsup_{n \rightarrow \infty} \kappa_X(n) > 0$. We define the *skew product* topological

dynamical system⁴ $(G \rtimes X, T)$ with the compact metric state space $\{(g, x) \mid g \in G, x \in X\}$, and update $T(g, x) = (g \cdot \varphi(x(0)), x(1, \infty))$. Observe that $T^{-1} \circ (g, x) = \{(ga^{-1}, ax) \mid ax \in X\}$, where a^{-1} is shorthand for $\varphi(a)^{-1}$.

We shall equivalently (in fact, interchangeably since we deal with a permutation automaton) regard the skew product as a shift space Z over the alphabet $(G \times \Sigma)$. We map (g, x) to y (and also denote y by (g, x)), where $y(n) = (g \cdot \varphi(x(0, n)), x(n))$. This perspective is isomorphic and each element $y \in Z$ can naturally be viewed as the run of an automaton with states G on its projection $x \in X$. Note that $\mathcal{A}(x)$ can be obtained by simply applying a coding (letter-to-letter substitution) on y . By Lem. 5.2, each $y \in Z$ is uniformly recurrent (however, Z is not necessarily minimal). Given a word x whose run $\mathcal{A}(x) = (e, x) = y$ we are to study, let Y be the minimal shift defined by y : note that Y is contained in (but not necessarily equal to) Z . We use π to denote the projection of Y onto X , and π^{-1} to denote its pre-image.

Our goal is to prove that if X satisfies Boshernitzan's condition (Def. 3.6), then so does Y . We shall do so⁵ by starting with an arbitrary invariant measure ν on Y , and using it to construct an invariant measure μ_Y on Y that satisfies $\kappa_Y(n) = c \cdot \kappa_X(n)$ for all large enough n , where c is a constant.

We shall study the symmetry of Y in order to understand how much structure Y inherits from X . Consider the group action of G on Z , where $h \in G$ maps (g, x) to (hg, x) . By the associativity of the binary group operation, we have that the action of h commutes with the shift operation T_Z and its inverse. Define the subgroup H as $\{h \in G \mid hY = Y\}$. We shall show the following connection between invariant measures on Y and X .

LEMMA 7.4. *Let ν be an arbitrary invariant measure on Y .*

- (1) *The map $\nu \circ h$ is an invariant measure on Y for each $h \in H$.*
- (2) *The map $\nu \circ \pi^{-1}$ on X is the same as the unique invariant measure μ_X .*

In particular, the invariant measure μ_Y defined as $\frac{1}{|H|} \sum_{h \in H} \nu \circ h$ satisfies $\mu_Y \circ \pi^{-1} = \mu_X$.

The following lemma, which is inspired by [30, Prop. 2.1], is instrumental in making π^{-1} explicit: it tells us that we can express $\pi^{-1}(x) = \{(g, x) \mid g \in \alpha(x)\}$.

LEMMA 7.5. *The map α given by $\alpha(x) = \{g \mid (g, x) \in Y\}$ is a well defined function from X to $C = H \backslash G$, satisfies $\alpha(ax) = \alpha(x) \cdot a^{-1}$, and is continuous. In particular, X can be expressed as a finite union of cylinders on which α is constant.*

PROOF. Note that $\pi(Y)$ must be a shift contained in X because the projection π commutes with T and is a closed map as G is compact; since X is a minimal shift, $\pi(Y) = X$. This assures us that α maps each $x \in X$ to a non-empty set.

We first prove that $\alpha(x)$ is indeed a coset. By the defining property of the subgroup H , we deduce that if $(g, x) \in Y$, then $Hg \subseteq \alpha(x)$. We now need to prove that $\alpha(x)$ cannot span more than one coset, i.e., if $(g, x) \in Y$ and $(g', x) \in Y$, then $g' = hg$ for some $h \in H$.

⁴We acknowledge that the notation T is overloaded, but nevertheless use it for clarity because, as we shall see, it performs the same shift operation in spirit. If the shift operator is invoked on multiple shift spaces within the same context, we shall use a distinguishing subscript.

⁵Some proof ideas are inspired by [9], especially Prop. 3.8 *ibidem*.

Let $Y' = g'g^{-1}Y$, and observe that Y' is also a minimal shift (recall that the group action commutes with the shift operation). We have that (g, x) is contained in the intersection of the minimal shifts Y, Y' : this is only possible if $Y = Y'$. In other words, $h = g'g^{-1} \in H$. We henceforth refer to α as a function from X to C ; in technical parlance it is a *(minimal) cobounding map modulo H* .

The next requirement follows readily. Let $x, ax \in X$. Since $\alpha(x)$ and $\alpha(ax)$ are both cosets, they have the same cardinality, and it suffices to observe from the construction of Y that $\alpha(ax) \cdot \varphi(a) \subseteq \alpha(x)$ in order to establish that $\alpha(ax) = \alpha(x) \cdot a^{-1}$.

Thirdly, we show that α is continuous. Suppose for the sake of deriving a contradiction that α is not continuous at some $x \in X$, i.e., $\alpha(x) = C$, and for every index n , there exists $x_n \in X$ such that $x(0, n) = x_n(0, n)$ but $\alpha(x_n) = C_n \neq C$. By the pigeonhole principle, we deduce that there exist a coset C' , infinitely many indices n , and words x_n such that $x_n(0, n) = x(0, n)$, but $\alpha(x_n) = C' \neq C$. Consider an element $g' \in C$, and observe that $((g', x_n))_n$ is an infinite sequence of elements in the compact space Y , and this sequence converges to (g', x) . However, $g' \notin \alpha(x)$, i.e., the limit does not exist in Y : a contradiction, as desired.

It remains to show that X is a finite union of cylinders on which α is constant. From the preceding argument for continuity, it follows that for any $x \in X$, we can obtain cylinders $[u_j]_X$ on which α is constant and $x(j, \infty) \in [u_j]_X$. We show by contradiction that these cylinders cover X : if they leave some x' uncovered, we can use continuity to identify a cylinder $[u']_X$ on which α is constant and $x' \in [u']_X$. However, by the minimality of X , u' must have been a prefix of some $x(j, \infty)$: a contradiction. Finally, since X is compact, the open cover $([u_j]_{j=0}^\infty)$ of cylinders with the desired property admits a finite subcover $([u_j]_{j=0}^N)$, which we use to establish the last claim of the lemma. \square

PROOF OF LEM. 7.4. Lem. 7.5 gives the framework to deduce that: (i) the operators T_Y, T_Y^{-1} on Y give the same evaluations as their counterparts T_Z, T_Z^{-1} ; (ii) the inverse image π^{-1} commutes with T and T^{-1} . These will respectively imply the two parts of the claim.

Observe that $T_Y^{-1} \circ (g, x) \subseteq T_Z^{-1} \circ (g, x) = \{(ga^{-1}, ax) \mid ax \in X\}$. On the other hand, $\alpha(ax) = \alpha(x) \cdot a^{-1}$: since $g \in \alpha(x)$, this proves by the definition of α that the set inclusion is actually an equality, i.e., $T_Y^{-1} = T_Z^{-1}$ (that $T_Y = T_Z$ on Y is obvious by the definition of Y). This proves (i). It then immediately follows that $h \circ T_Y^{-1} = T_Y^{-1} \circ h$, and hence $\nu \circ h$ is an invariant measure on Y for each $h \in H$. Indeed, $\nu \circ h \circ T_Y^{-1} = \nu \circ T_Y^{-1} \circ h = \nu \circ h$.

To show (ii), observe furthermore that $\pi^{-1} \circ T_X^{-1}(x) = \{(g', ax) \mid g' \in \alpha(ax), ax \in X\} = \{(ga^{-1}, ax) \mid g \in \alpha(x), ax \in X\} = T_Y^{-1} \circ \pi^{-1}(x)$. This implies that for any invariant measure ν , we have $\nu \circ \pi^{-1} \circ T_X^{-1} = \nu \circ T_Y^{-1} \circ \pi^{-1} = \nu \circ \pi^{-1}$. It remains to observe that $\nu \circ \pi^{-1}$ is indeed a Borel measure on X because the function π is Borel measurable by virtue of being continuous. Thus, $\nu \circ \pi^{-1}$ is an invariant measure on X ; since (X, T_X) is uniquely ergodic, it can only be the same as μ_X . \square

We are now ready to complete the proof that Y inherits the property of admitting an invariant measure that satisfies Boshernitzan's condition from X .

Let us examine what μ_Y evaluates to on cylinders. For convenience, we shall denote a finite factor v as (g, u) , where $v(n) = (g \cdot \varphi(u(0, n)), u(n))$ for all n . We have

$$\mu_Y([(g, u)]_Y) = \frac{1}{|H|} \sum_{h \in H} \nu \circ h([(g, u)]_Y) = \frac{1}{|H|} \sum_{g' \in Hg} \nu([(g', u)]_Y),$$

and by symmetry, $\mu_Y([(g, u)]_Y) = (1/|H|) \cdot \sum_{g' \in Hg} \mu_Y([(g', u)]_Y)$.

Now, by Lem. 7.5, for all long enough factors $u \in \mathcal{L}(X)$, we have that α is constant-valued on $[u]_X$. This implies that for u long enough, $\pi^{-1}([u]_X) = \bigcup_{g' \in \alpha([u]_X)} [(g', u)]_Y$. In other words, $\mu_Y([(g, u)]_Y) = (1/|H|) \cdot \mu_Y \circ \pi^{-1}([u]_X)$. Since $\mu_Y \circ \pi^{-1} = \mu_X$, we have thus proven that for all long enough finite factors, we have $\mu_Y([(g, u)]_Y) = (1/|H|) \cdot \mu_X([u]_X)$, and it follows that $\kappa_Y(n) = (1/|H|)\kappa_X(n)$. We have thus proved that the run y of an automaton with states G and initial state e (where e is the group identity) defines a minimal shift Y that satisfies Boshernitzan's condition (Def. 3.6).

We obtain $\mathcal{A}(x)$ by a simple coding that substitutes (g, a) with (q, a) , where q is obtained by the action of g on q_{init} . By Lem. 4.1, $\mathcal{A}(x)$ satisfies Boshernitzan's condition (Def. 3.6), and hence by Lem. 3.7, admits factor frequencies. This proves Lem. 7.3. \square

8 DISCUSSION

An open question is to find an example of a word x (or prove one does not exist) that does not satisfy Boshernitzan's condition, but $\mathcal{A}(x)$ generates uniquely ergodic shifts for every \mathcal{A} .

An obvious direction for future work is to generalise the following beyond the case of primitive morphic words: if the factor frequencies of x are given by *effective closed-form* expressions, then so are those of $\mathcal{A}(x)$. Concretely, if we can decide whether $f_x(u) = r$ where r is a given constant, our current proof of Thm. 7.1 via Lem. 3.2 does not imply that $\mathcal{A}(x)$ inherits this property.

The main technical obstacle to the above is obtaining an effective version of Lem. 7.5, i.e., the computation of *minimal cobounding maps* on the shift X defined by the input word. This is closely related to the behaviour of *return groups*. Recall that $\mathcal{R}_X(u)$ denotes the set of return words to a factor $u \in \mathcal{L}(X)$. Let G be a group, and φ be a morphism from Σ^* into G . The return group of u (with respect to G, φ) is generated by the return words to u , and is given by $\langle \varphi(\mathcal{R}_X(u)) \rangle$. If u' is a prefix of u , then the return group of u is a subgroup of that of u' [23, Lem. 3]. We say that the return groups *stabilise* at u if for every factor u'' such that u is a prefix of u'' , the return groups of u and u'' are the same. As [9, Prop. 7.8] observes, determining minimal cobounding maps is equivalent to determining when return groups stabilise.

Our effective results for primitive morphic words indicate that the stabilisation of return groups can be determined in this case, and it has indeed been proven [23, Prop. 27]. Nevertheless, this remains open in the general case, to the best of our knowledge. A possible approach is to generalise the derivation-based techniques of Durand applied in both [23] and our paper.

Another potential approach leverages recent advances in word combinatorics, in particular the study of so-called *suffix-connected* words and their generalisations [25]. An important property of suffix-connected words (which include Sturmian words and Arnoux-Rauzy words) is that all factors are *stable* [25, Cor. 1.2], i.e., for

any factor u , the return group (with respect to G, φ) is G . While this property is ideal to prove an effective version of Lem. 7.3, it remains difficult to determine when an automaton preserves suffix-connectedness, and hence obtain a preservation result for an entire cascade of automata. As a positive case in point, we give the following application.

PROPOSITION 8.1. *Let $\mathcal{A} = (Q', q_{\text{init}}, \Sigma, \delta)$ be a permutation automaton, and let $x \in \Sigma^\omega$ be a Sturmian word. Let G be the group of permutations induced by the transitions, and let Q be the orbit of q_{init} under G . The run of \mathcal{A} on x visits each state $q \in Q$ with frequency $1/|Q|$.*

PROOF. Let X be the minimal shift defined by x . Since the word x is stable by virtue of being Sturmian and hence suffix-connected, we have by [9, Prop. 7.8] that the trivial cobounding map on X is minimal. This implies that the shift Y corresponding to the entire skew product $G \rtimes X$ is minimal. Furthermore, since x is a Sturmian word, it satisfies Boshernitzan's condition (Rmk. 3.8), and by Lem. 7.3, the shift Y also satisfies Boshernitzan's condition. From the proof of Lem. 7.3, we have that the unique invariant measure on Y is given by $\mu_Y([(g, u)]_Y) = \frac{1}{|G|} \mu_X([u]_X)$.

Consider $q \in Q$. The set $H_q = \{g \mid q_{\text{init}} \cdot g = q\}$ is a coset of the subgroup $H = \{h \mid q_{\text{init}} \cdot h = q_{\text{init}}\}$, and hence $|Q| \cdot |H| = |G|$. Now by Oxtoby's theorem, $f_{\mathcal{A}(x)}(q, a) = \sum_{g \in H_q} \mu_Y([(g, a)]_Y)$, which by the definition of μ_Y and the above observation, simplifies to $\frac{1}{|Q|} \mu_X([a]_X) = \frac{1}{|Q|} f_x(a)$. The frequency with which a run visits a state $q \in Q$ is then simply $\frac{1}{|Q|} \sum_a f_x(a) = \frac{1}{|Q|}$. \square

We conclude by recording an interesting corollary of Thm. 7.1, Rmk. 3.8, and the proof of Prop. 8.1.

COROLLARY 8.2. *Let $\Sigma = \{0, 1\}$, let $\mathcal{A} = (Q, q_{\text{init}}, \Sigma, \Gamma, \delta^o, \delta^t)$ be a transducer, and let $x \in \Sigma^\omega$ be a Sturmian word whose factor frequencies have effective closed form expressions. If \mathcal{A} is a permutation transducer, then the factor frequencies in $\mathcal{A}(x)$ have effective closed form expressions.*

REFERENCES

- [1] Jean-Paul Allouche and Jeffrey Shallit. The ubiquitous Prouhet-Thue-Morse sequence. In *Sequences and their applications (Singapore, 1998)*, Springer Ser. Discrete Math. Theor. Comput. Sci., pages 1–16. Springer, London, 1999.
- [2] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003.
- [3] Micael Baake and Uwe Grimm. *Aperiodic order. Vol. 1*, volume 149 of *Encyclopedia of Mathematics and its Application*. Cambridge University Press, 2013.
- [4] Heinz Bauer. *Measure and Integration Theory*. De Gruyter, Berlin, New York, 2001.
- [5] Verónica Becher and Pablo Ariel Heiber. Normal numbers and finite automata. *Theoretical Computer Science*, 477:109–116, 2013.
- [6] Valérie Berthé, Toghru Karimov, Joris Nieuwveld, Joël Ouaknine, Mihir Vahanwala, and James Worrell. On the decidability of monadic second-order logic with arithmetic predicates. In *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 1–14, 2024.
- [7] Valérie Berthé, Toghru Karimov, Joris Nieuwveld, Joël Ouaknine, Mihir Vahanwala, and James Worrell. The monadic theory of toric words. *Theoretical Computer Science*, 1025:114959, 2025.
- [8] Valérie Berthé and Vincent Delecroix. Beyond substitutive dynamical systems: S-adic expansions, 2017.
- [9] Valérie Berthé, Herman Goulet-Ouellet, Carl-Fredrik Nyberg-Brodda, Dominique Perrin, and Karl Petersen. Density of group languages in shift spaces, 2024.
- [10] Michael D. Boshernitzan. A condition for unique ergodicity of minimal symbolic flows. *Ergodic Theory and Dynamical Systems*, 12(3):425–428, 1992.

- [11] Olivier Carton. Preservation of normality by unambiguous transducers. In *Developments in language theory*, volume 13257 of *Lecture Notes in Comput. Sci.*, pages 90–101. Springer, Cham, [2022] ©2022.
- [12] Olivier Carton and Elisa Orduna. Preservation of normality by transducers. *Inform. and Comput.*, 282:Paper No. 104650, 9, 2022.
- [13] Jon Chaika. Skew products over rotations with exotic properties. *Geometriae Dedicata*, 168(1):369–385, 2014.
- [14] David Damanik and Daniel Lenz. A condition of Boshernitzan and uniform convergence in the multiplicative ergodic theorem. *Duke Mathematical Journal*, 133(1):95 – 123, 2006.
- [15] David Damanik and Daniel Lenz. Zero-measure cantor spectrum for schrödinger operators with low-complexity potentials. *Journal de Mathématiques Pures et Appliquées*, 85(5):671–686, 2006.
- [16] Tomasz Downarowicz and Benjamin Weiss. Lifting generic points. *Ergodic Theory and Dynamical Systems*, 44(9):2565–2580, 2024.
- [17] F. Durand, B. Host, and C. Skau. Substitutional dynamical systems, bratteli diagrams and dimension groups. *Ergodic Theory and Dynamical Systems*, 19(4):953–993, 1999.
- [18] Fabien Durand. Linearly recurrent subshifts have a finite number of non-periodic subshift factors. *Ergodic Theory and Dynamical Systems*, 20, 08 2008.
- [19] Fabien Durand. Decidability of the HDOL ultimate periodicity problem. *RAIRO - Theoretical Informatics and Applications*, 47, 11 2011.
- [20] Fabien Durand. HDOL ω -equivalence and periodicity problems in the primitive case. *Unif. Distrib. Theory*, 7(1):199–215, 2012.
- [21] Fabien Durand. Decidability of uniform recurrence of morphic sequences. *International Journal of Foundations of Computer Science*, 24(01):123–146, 2013.
- [22] Fabien Durand and Dominique Perrin. *Dimension Groups and Dynamical Systems: Substitutions, Bratteli Diagrams and Cantor Systems*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2022.
- [23] France Gheeraert, Herman Goulet-Ouellet, Julien Leroy, and Pierre Stas. Stability properties for subgroups generated by return words. *European Journal of Combinatorics*, 130:104224, 2025.
- [24] Amy Glen and Jacques Justin. Episturmian words: a survey. *RAIRO - Theoretical Informatics and Applications*, 43(3):403–442, 3 2009.
- [25] Herman Goulet-Ouellet. Suffix-connected languages. *Theoretical Computer Science*, 923:126–143, 2022.
- [26] M Guénais and F Parreau. Valeurs propres de transformations liées aux rotations irrationnelles et aux fonctions en escalier (eigenvalues of transformations arising from irrational rotations and step functions) (2006). *arXiv preprint math/0605250*.
- [27] Juha Honkala. On the simplification of infinite morphic words. *Theoretical Computer Science*, 410(8):997–1000, 2009.
- [28] Toghrul Karimov. Algorithmic verification of linear dynamical systems, 2023.
- [29] Kenneth Krohn and John Rhodes. Algebraic theory of machines. i. prime decomposition theorem for finite semigroups and machines. *Transactions of The American Mathematical Society - TRANS AMER MATH SOC*, 116, 04 1965.
- [30] Mariusz Lemańczyk and Mieczysław K. Mentzen. Topological ergodicity of real cocycles over minimal rotations. *Monatshefte für Mathematik*, 134(3):227–246, 2002.
- [31] Oded Maler. *On the Krohn-Rhodes Cascaded Decomposition Theorem*, pages 260–278. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [32] Robert McNaughton and Seymour A. Papert. *Counter-Free Automata (M.I.T. research monograph no. 65)*. The MIT Press, 1971.
- [33] A.R. Meyer and C. Thompson. Remarks on algebraic decomposition of automata. *Mathematical Systems Theory*, 3:110–118, 1969.
- [34] A. Muchnik, A. Semenov, and M. Ushakov. Almost periodic sequences. *Theoretical Computer Science*, 304(1):1–33, 2003.
- [35] Yu. L. Pritykin. Finite-automaton transformations of strictly almost-periodic sequences. *Mathematical Notes*, 80(5):710–714, 2006.
- [36] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.
- [37] Martine Queffélec. *Substitution Dynamical Systems - Spectral Analysis*, volume 1294 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 2nd edition, 2010.
- [38] Evgueni A Sataev. On the number of invariant measures for flows on orientable surfaces. *Mathematics of the USSR-Izvestiya*, 9(4):813, 1975.
- [39] Jeffrey Shallit. *The logical approach to automatic sequences—exploring combinatorics on words with Walnut*, volume 482 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2023.
- [40] William A Veech. Strict ergodicity in zero dimensional dynamical systems and the Kronecker-Weyl theorem mod. 2. *Transactions of the American Mathematical Society*, 140:1–33, 1969.
- [41] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.
- [42] Hermann Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Mathematische Annalen*, 77(3):313–352, 1916.

A OMITTED PROOFS

LEMMA A.1. *Let $x \in \Sigma^\omega$ be a substitutive word obtained as the fixed point of a primitive substitution σ , and let $y = \tau(x)$ be a primitive morphic word. We have that x, y are effectively linearly recurrent and admit computable factor frequencies.*

PROOF. We prove the lemma for the primitive substitutive x , and defer the proof for the primitive morphic y to Lem. 4.1, whence it follows immediately. In the proof, we shall associate to the substitution σ an adjacency matrix M , whose (j, i) -th entry records the number of times the j -th letter a_j appears in the image $\sigma(a_i)$ of the i -th letter a_i . Because σ is primitive, we can assume up to replacing σ with some σ^k that M has strictly positive entries, which allows us to use Perron–Frobenius theory.

We first show that x is effectively linearly recurrent, i.e., $R_x(n) \leq Ln$ for some computable L . The proof of linear recurrence in [17, Prop. 25] shows that $L = C \cdot \max_{a \in \Sigma} |\sigma(a)| \cdot R_x(2)$, where C is such that for all $n \geq 1$, $\max_{a \in \Sigma} |\sigma^n(a)| \leq C \cdot \min_{a \in \Sigma} |\sigma^n(a)|$.

It suffices to take $C = (\max_{i,j} (v_i/v_j))^2$, where $(v_i)_{i=1}^{|\Sigma|}$ are the strictly positive entries of a left eigenvector v^\top of M corresponding to the Perron-Frobenius eigenvalue ρ (i.e., the positive one with maximum absolute value). Indeed, for the i -th letter a , $|\sigma^n(a)| = \sum_j (M^n)_{j,i}$. The expression on the right is the sum of entries in the i -th column of M^n , and is lower bounded by $\sum_j (v_j/\max_j v_j) \cdot (M^n)_{j,i}$, and upper bounded by $\sum_j (v_j/\min v_j) \cdot (M^n)_{j,i}$. These bounds respectively evaluate to the i -th entries of $(v^\top M^n)/(\max_j v_j)$ and $(v^\top M^n)/(\min_j v_j)$. Since v^\top is an eigenvector, we get that

$$\rho^j (\min_j v_j) / (\max_j v_j) \leq |\sigma^n(a)| \leq \rho^j (\max_j v_j) / (\min_j v_j).$$

We now show how to compute (an upper bound on) $R_x(2)$. First, we can compute $\mathcal{L}_x(2)$ by starting with $x(0, 2)$, and recording the length-2 factors that are produced upon repeatedly applying σ . The set of factors will saturate within $|\Sigma|^2$ iterations. In particular, we can find k large enough such that for all $a \in \Sigma$, $\sigma^k(a)$ contains every factor in $\mathcal{L}_x(2)$. Clearly, $R_x(2) \leq 2 \max_a |\sigma^k(a)|$. This proves effective linear recurrence.

We note that the textbook discussions in [37, Chap. 5.4] and [22, Chap. 3.8.5] prove x has computable factor frequencies. We present the techniques here for clarity in exposition. Recall the adjacency matrix M of σ . It follows that the frequency of the i -th letter a in x is the i -th entry of the eigenvector corresponding to the Frobenius eigenvalue.

This same idea is extrapolated to describe factor frequencies as follows. We define an alphabet Σ_n , whose letters correspond to words in $\mathcal{L}_n(x)$ (see Rmk. 2.1 for how to compute this alphabet), and the word $x_n \in \Sigma_n^\omega$, where $x_n(i)$ corresponds to the factor $x(i, i+n)$. We define the substitution σ_n such that $\sigma_n(u)$ is the ordered list of the first $|\sigma(u(0))|$ length- n factors of $\sigma(u)$. Crucially, this substitution on Σ_n is primitive [37, Lem. 5.3], and we can again use Perron-Frobenius theory to obtain the frequency of factor $u \in \Sigma_n$, which is a letter in Σ_n . \square

LEMMA A.2. *Consider the partition of the unit circle into finitely many semi-open intervals $I_j = [\beta_j, \gamma_j)$, and let α be irrational. The shift X generated by the codings (with respect to the intervals I_j) of the rotation by α is uniquely ergodic.*

PROOF. Denote by \tilde{I}_j the interval $(\beta_j, \gamma_j]$. By symmetry, the set of codings with respect to the partition of “dual” intervals is also included in X . We shall prove that X comprises precisely of the codings with respect to the original partition, and codings with respect to the dual partition.

For each $u \in \mathcal{L}_n(X)$, we define I_u and \tilde{I}_u as before, and also define the non-empty closed interval \bar{I}_u as their union. If $x \in X$ is a coding of the rotation starting at ζ , then for each prefix u of x , we must have $\zeta \in \bar{I}_u$. Conversely, if ζ is in each \bar{I}_u , then x is its coding because by the orientation of the intervals, ζ is either in each I_u , or

in each \tilde{I}_u (if ζ was on the left endpoint of I_u and right endpoint of \tilde{I}_{uv} , it would imply that I_{uv} is empty, a contradiction).

Thus, each $x \in X$ defines an infinite sequence of nested closed intervals, and there necessarily exists a point ζ in their countable intersection. By the preceding discussion x codes a rotation of this ζ with respect to either the original partition, or the dual partition. This proves that each $x \in X$ admits factor frequencies, with $f_x(u)$ given by the Lebesgue measure of I_u . We use Oxtoby’s theorem to conclude that x is uniquely ergodic. \square