

Automatic sequences are orthogonal to aperiodic multiplicative functions

Mariusz Lemańczyk* Clemens Müllner†

November 5, 2018

Abstract

Given a finite alphabet \mathbb{A} and a primitive substitution $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ (of constant length λ), let (X_θ, S) denote the corresponding dynamical system, where X_θ is the closure of the orbit via the left shift S of a fixed point of the natural extension of θ to a self-map of $\mathbb{A}^{\mathbb{Z}}$. The main result of the paper is that all continuous observables in X_θ are orthogonal to any bounded, aperiodic, multiplicative function $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$, i.e.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(S^n x) \mathbf{u}(n) = 0$$

for all $f \in C(X_\theta)$ and $x \in X_\theta$. In particular, each primitive automatic sequence, that is, a sequence read by a primitive finite automaton, is orthogonal to any bounded, aperiodic, multiplicative function.

Introduction

Throughout the paper, by an *automatic* sequence $(a_n)_{n \geq 0} \subset \mathbb{C}$, we mean a continuous observable in a substitutional system (X_θ, S) , i.e. $a_n = f(S^n x)$, $n \geq 0$, for some $f \in C(X_\theta)$ and $x \in X_\theta$.¹ Here, we assume that $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ is a substitution of constant length λ (over the alphabet \mathbb{A}), and we let (X_θ, S) denote the corresponding subshift of the full shift $(\mathbb{A}^{\mathbb{Z}}, S)$ (see Section 2 for details).

By $\boldsymbol{\mu} : \mathbb{N} \rightarrow \{-1, 0, 1\}$ we denote the classical Möbius function: $\boldsymbol{\mu}(p_1 \dots p_k) = (-1)^k$ for different primes p_1, \dots, p_k , $\boldsymbol{\mu}(1) = 1$ and $\boldsymbol{\mu}(n) = 0$ for all non square-free numbers $n \in \mathbb{N}$. In connection with the celebrated Sarnak's conjecture [37] on Möbius orthogonality of zero entropy systems, i.e.

$$(1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(S^n x) \boldsymbol{\mu}(n) = 0$$

*Research supported by a Narodowe Centrum Nauki grant.

†Research supported by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No 648132, by the project F55-02 of the Austrian Science Fund FWF which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications" and by Project I1751 (FWF), called MUDERA (Multiplicativity, Determinism, and Randomness).

¹This is a slight extension of the classical notion of automatic sequence meant as a sequence obtained by a finite code of a fixed point of a substitution, see e.g. [3], [14], [36] to see also a relationship with sequences read by finite automata.

for each zero entropy dynamical system (X, S) , all $f \in C(X)$ and $x \in X$ (for more information on the subject, see the survey article [12]), it has been proved in [35] that all automatic sequences (a_n) are orthogonal to the classical Möbius function $\boldsymbol{\mu}$, i.e. $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_n \boldsymbol{\mu}(n) = 0$. This triggers the question whether (1) remains true if we replace $\boldsymbol{\mu}$ by another arithmetic function. The Möbius function is an example of an arithmetic function which is multiplicative ($\boldsymbol{\mu}(mn) = \boldsymbol{\mu}(m)\boldsymbol{\mu}(n)$ whenever m, n are coprime), hence, it is natural to ask whether automatic sequences are orthogonal to each zero mean,² bounded, multiplicative function \mathbf{u} . Said that, one realizes immediately that the answer to such a question is negative as periodic functions are automatic sequences and we have many periodic, multiplicative functions.³ But even if we consider the non-periodic case, still, there are non-periodic automatic sequences which are (completely) multiplicative, zero mean functions [2], [38], [41]. On the other hand, it has been proved in [13] that many automatic sequences given by so called bijective substitutions are orthogonal to all zero mean, bounded, multiplicative functions (in [9], it is proved that they are orthogonal to the Möbius function).

A stronger requirement than zero mean of \mathbf{u} which one can consider is that of *aperiodicity*, that is

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mathbf{u}(an + b) = 0$$

for each $a, b \in \mathbb{N}$, i.e. \mathbf{u} has a mean, equal to zero, along each arithmetic progression. Many classical multiplicative functions are aperiodic, e.g. $\boldsymbol{\mu}$ or the Liouville function $\boldsymbol{\lambda}$.

The aim of the present paper is to prove the following ($h(\theta)$ and $c(\theta)$ stand, respectively, for the height and the column number of the substitution θ , see Section 2):

Theorem 0.1. *Let θ be a primitive substitution of constant length λ . Then, each automatic sequence $a_n = f(S^n x)$, $n \geq 0$, in (X_θ, S) is orthogonal to any bounded, aperiodic, multiplicative function $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$, i.e.*

$$(2) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_n \mathbf{u}(n) = 0.$$

More precisely:

(i) *If $c(\theta) = h(\theta)$ then each automatic sequence $(f(S^n x))_{n \geq 0}$ is orthogonal to any bounded, aperiodic, arithmetic function \mathbf{u} .⁴*

(ii) *If $c(\theta) > h(\theta)$ then: the automatic sequences $(f(S^n x))_{n \geq 0}$ for which the spectral measure of f is continuous are orthogonal to all bounded, multiplicative functions. If the spectral measure is discrete then (i) applies. All other automatic sequences in (X_θ, S) are orthogonal to all bounded, aperiodic, multiplicative functions.*

²Recall that a sequence $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$ has zero mean if $M(\mathbf{u}) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mathbf{u}(n)$ exists and equals zero.

³Indeed, examples of periodic multiplicative functions are given by: Dirichlet characters, or $n \mapsto (-1)^{n+1}$.

⁴We emphasize that no multiplicativity on \mathbf{u} is required.

We have already mentioned that examples of automatic sequences which are multiplicative functions are known, but they are quite special. For example, in [38], it is proved that completely multiplicative, never vanishing functions “produced” by finite automata are limits in the upper density of periodic sequences, that is, they are Besicovitch rationally almost periodic. In fact, we can strengthen this result by showing that they are even Weyl rationally almost periodic,⁵ which is a consequence of Theorem 0.1:

Corollary 0.2. All multiplicative and automatic sequences produced by primitive automata⁶ are Weyl rationally almost periodic. Furthermore, the automaton/substitution that produces such an automatic sequence can be chosen to have equal column number and height.

The main problem in this paper belongs to number theory and combinatorics, and the proof of Möbius orthogonality for automatic sequences in [35] relied on combinatorial properties of sequences produced by automata and an application of the (number theoretic) method of Mauduit and Rivat [31], [32]. However, the problem of orthogonality of sequences is deeply related to classical ergodic theory, namely, to Furstenberg’s disjointness of dynamical systems, see [12] for an exhaustive presentation of this approach. A use of ergodic theory tools is the main approach in the present paper: we make use of some old results on the centralizer of substitutional systems [22] and, more surprisingly, we find an ergodic interpretation of the combinatorial approach from [35] in terms of joinings of substitutional systems. Finally, we relativize some of the arguments from [13] to reach the goal. For example, the reason behind the orthogonality of automatic sequences to all bounded, multiplicative functions (whenever the spectral measure of an automatic sequence is continuous) in (ii) of Theorem 0.1 is that the essential centralizer of substitutional systems is finite [22]. This “small size” of the essential centralizer puts serious restrictions on possible joinings between (sufficiently large) different prime powers S^p , S^q of S and allows one to use the numerical KBSZ criterion on the orthogonality of numerical sequences with bounded, multiplicative functions (see Section 1.5).

The ergodic theory approach (together with number-theoretic tools) turn out to be very effective in some attempts to prove Sarnak’s conjecture. The recent results of Frantzikinakis and Host [16] show: if instead of orthogonality (1), we ask for its weaker, namely, logarithmic version:

$$(3) \quad \lim_{N \rightarrow \infty} \frac{1}{\log N} \sum_{n \leq N} \frac{1}{n} f(S^n x) \boldsymbol{\mu}(n) = 0$$

then this orthogonality indeed holds for all zero entropy (X, S) systems whose set of ergodic (invariant) measures is countable, in particular, it applies to substitutional systems. Moreover, in (3), we can replace $\boldsymbol{\mu}$ by many other so called non-pretentious multiplicative functions [17]. However, so far, the approach

⁵A sequence (b_n) taking values in a finite set \mathbb{B} is called *Weyl rationally almost periodic* if it can be approximated by periodic sequences (with values in \mathbb{B}) in the pseudo-metric

$$d_W(x, y) = \limsup_{N \rightarrow \infty} \sup_{\ell \geq 1} \frac{1}{N} |\{\ell \leq n < \ell + N : x(n) \neq y(n)\}|.$$

⁶Sequences produced by finite automata take only finitely many values and are, therefore, bounded.

through logarithmic averaging does not seem to answer the main conjecture (1) (see also [20] and [39]).

The ergodic theory approach allows one for a further extension of the notion of being automatic sequence which triggers one more natural question. Namely, a uniquely ergodic topological dynamical system (Y, T) (with a unique invariant measure ν) is called *MT-substitutional* if there is a primitive substitution $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ such that the measure-theoretic dynamical systems (Y, ν, T) and $(X_\theta, \mu_\theta, S)$ are measure-theoretically isomorphic. Then, any sequence $a_n = g(T^n y)$, $n \geq 0$ (with $g \in C(Y)$ and $y \in Y$), can be called *MT-automatic*.⁷ To cope with the MT-automatic case we need to control a behaviour of \mathbf{u} on so called short intervals:

$$(4) \quad \lim_{K \rightarrow \infty} \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} \mathbf{u}(n) \right| = 0$$

(for each (b_k) satisfying $b_{k+1} - b_k \rightarrow \infty$) which is much stronger than the requirement that $M(\mathbf{u}) = 0$. It turns out however that (4) is satisfied for many so called non-pretentious multiplicative functions, see [29], [30]; in particular, it is satisfied for the Möbius function μ .

Corollary 0.3. Any MT-automatic sequence (a_n) satisfies (2) for any multiplicative, bounded, aperiodic $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$ satisfying (4). For each MT-substitutional system (Y, T) and any $g \in C(Y)$, we have

$$\frac{1}{N} \sum_{n \leq N} g(T^n y) \mathbf{u}(n) \rightarrow 0 \text{ when } N \rightarrow \infty$$

uniformly in $y \in Y$. In particular, the uniform convergence takes place in (X_θ, S) for any primitive substitution θ .

In the course of the proof of Theorem 0.1, given a substitutional system (X_θ, S) , we will build successive continuous extensions of it, which are also given by substitutions, where for the largest extension, Theorem 0.1 will be easier to handle. As a byproduct of this procedure, we will clear up Remark 9.1 from [36], p. 229, about the form of a cocycle in a skew product representation of $(X_\theta, \mu_\theta, S)$ over the Kronecker factor.

Added in June 2018: In the recent paper [24], it is proved that all q -multiplicative sequences are either almost periodic which roughly would correspond to (i) of Theorem 0.1 or are orthogonal to all bounded, multiplicative functions. This makes some overlap (some automatic sequences are q -multiplicative) with our main result but the classes considered in [24] and in the present paper are essentially different.

1 Ergodic theory necessities

1.1 Joinings and disjointness

By a (measure-theoretic) *dynamical system* we mean (X, \mathcal{B}, μ, S) , where (X, \mathcal{B}, μ) is a probability standard Borel space and $S : X \rightarrow X$ is an a.e. bijection which

⁷For example, sequences given by pieces of automatic sequences in (X_θ, S) : $g(S^n x_k)$ for $b_k \leq n < b_{k+1}$, $k \geq 1$, where $b_1 = 1$ and $b_{k+1} - b_k \rightarrow \infty$, for $g \in C(X_\theta)$ and $(x_k) \subset X_\theta$, are MT-automatic, see Section 8.

is bimeasurable and measure-preserving. If no confusion arises, we will speak about S itself and call it an *automorphism*.⁸

Remark 1.1. Each homeomorphism S of a compact metric space X determines many (measure-theoretic) dynamical systems $(X, \mathcal{B}(X), \mu, S)$ with $\mu \in M(X, S)$, where $M(X, S)$ stands for the set of probability Borel measures on X ($\mathcal{B}(X)$ stands for the σ -algebra of Borel sets). Recall that by the Krylov-Bogolyubov theorem, $M(X, S) \neq \emptyset$, and moreover, $M(X, S)$ endowed with the weak-* topology becomes a compact metrizable space. The set $M(X, S)$ has a natural structure of a convex set (in fact, it is a Choquet simplex) and its extremal points are precisely the ergodic measures. We say that the topological system (X, S) is *uniquely ergodic* if it has only one invariant measure (which must be ergodic). The system (X, S) is called *minimal* if it does not contain a proper subsystem (equivalently, the orbit of each point is dense).

Remark 1.2. Basic systems considered in the paper are *subshifts* whose definition we now recall. Let \mathbb{A} be a finite, nonempty set (alphabet). By a *block* (or *word*) over \mathbb{A} , we mean $B \in \mathbb{A}^n$ (for some $n \geq 0$) and $n = |B|$ is the *length* of B . Hence $B = (a_{i_0}, a_{i_1}, \dots, a_{i_{n-1}})$ with $a_{i_k} \in \mathbb{A}$ for $k = 0, \dots, n-1$. We will also use the following notation: $B = B_0 B_1 \dots B_{n-1}$, where $B_j = a_{i_j}$ for $j = 0, \dots, n-1$. If $0 \leq i \leq j < n$ then we write $B[i, j]$ for $B_i B_{i+1} \dots B_j$, in particular, notationally, $B[i] = B_i$. We say that the (sub)block $B[i, j]$ *appears* in B . The notation we have just presented has its natural extension to infinite sequences.

Given $\eta \in \mathbb{A}^{\mathbb{N}}$, we can define

$$X_\eta := \{x \in \mathbb{A}^{\mathbb{Z}} : \text{each block appearing in } x \text{ appears in } \eta\}.$$

It is not hard to see that X_η is closed and S -invariant, where $S : \mathbb{A}^{\mathbb{Z}} \rightarrow \mathbb{A}^{\mathbb{Z}}$ is the left *shift*, i.e.

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

Then the dynamical system (X_η, S) is called a *subshift* (given by η). If η has the property that each block appearing in it reappears infinitely often (and such are substitutional systems which are considered in the paper) then there exists $\bar{\eta} \in \mathbb{A}^{\mathbb{Z}}$ satisfying: $\bar{\eta}_k = \eta_k$ for each $k \geq 0$ and $X_\eta = \{S^m \bar{\eta} : m \in \mathbb{Z}\}$.

Given another system (Y, \mathcal{C}, ν, T) , we may consider the set $J(S, T)$ of *joinings* of automorphisms S and T . Namely, $\kappa \in J(S, T)$ if κ is an $S \times T$ -invariant probability measure on $\mathcal{B} \otimes \mathcal{C}$ with the projections μ and ν on X and Y , respectively.⁹ Note that the projections maps $p_X : X \times Y \rightarrow X$, $p_Y : X \times Y \rightarrow Y$ settle factor maps between the dynamical systems

$$(X \times Y, \mathcal{B} \otimes \mathcal{C}, \kappa, S \times T) \text{ and } (X, \mathcal{B}, \mu, S), (Y, \mathcal{C}, \nu, T), \text{ respectively.}$$

The automorphisms S and T are called *disjoint* if the only joining of S and T is product measure $\mu \otimes \nu$, i.e. $J(S, T) = \{\mu \otimes \nu\}$. We will then write $S \perp T$. Note

⁸In what follows we will also use notation $S \in \text{Aut}(X, \mathcal{B}, \mu)$, where $\text{Aut}(X, \mathcal{B}, \mu)$ stands for the Polish group of all automorphisms of (X, \mathcal{B}, μ) . The topology is given by the strong operator topology of the corresponding unitary operators $U_S, U_S f := f \circ S$ on $L^2(X, \mathcal{B}, \mu)$.

⁹We can also speak about (topological) joinings in the context of topological dynamics. Indeed, if $(X, S), (Y, T)$ are two topological systems then each closed subset $M \subset X \times Y$ invariant under $S \times T$ and with full projections is called a (topological) joining.

that if $S \perp T$ then at least one of these automorphisms must be ergodic. If both are ergodic, the subset $J^e(S, T)$ of ergodic joinings (i.e. of those $\rho \in J(S, T)$ for which the system $(X \times Y, \rho, S \times T)$ is ergodic) is non-empty; in fact, the ergodic decomposition of a joining consists (a.e.) of ergodic joinings.

If (X, \mathcal{B}, μ, S) and (Y, \mathcal{C}, ν, T) are isomorphic, i.e. for some (invertible¹⁰) $W : (X, \mathcal{B}, \mu) \rightarrow (Y, \mathcal{C}, \nu)$, we have equivariance $W \circ S = T \circ W$, then W yields the corresponding *graph joining* $\mu_W \in J(S, T)$ determined by

$$\mu_W(B \times C) = \mu(B \cap W^{-1}C).$$

Then, $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \mu_W, S \times T)$ is isomorphic to S (hence μ_W is ergodic if S was). When $S = T$, we speak about *self-joinings* of S .

We recall that an automorphism $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$ has *discrete spectrum* if $L^2(Z, \mathcal{D}, \kappa)$ is spanned by the eigenfunctions of the unitary operator $U_R: f \mapsto f \circ R$ on $L^2(Z, \mathcal{D}, \kappa)$. Assuming ergodicity, we have:

(5) If R has discrete spectrum then each its ergodic self-joining is graphic.

Each automorphism $S \in \text{Aut}(X, \mathcal{B}, \mu)$ has the largest factor which has discrete spectrum. It is called the *Kronecker factor* of S .

Joinings are also considered in topological dynamics. If S_i is a homeomorphisms of a compact metric space X_i , $i = 1, 2$ then each $S_1 \times S_2$ -invariant, closed subset $M \subset X_1 \times X_2$ with the full natural projections, is called a *topological joining* of S_1 and S_2 .

1.2 Discrete suspensions

Given $S \in \text{Aut}(X, \mathcal{B}, \mu)$ and $h \in \mathbb{N}$, consider the probability space $(\tilde{X}, \tilde{\mu})$, where $\tilde{X} = X \times \{0, 1, \dots, h-1\}$ and $\tilde{\mu} = \mu \otimes \rho_h$, where ρ_h is the normalized uniform measure on $\mathbb{Z}/h\mathbb{Z} \cong \{0, 1, \dots, h-1\}$. We define now \tilde{S} on $(\tilde{X}, \tilde{\rho})$ by setting

$$(6) \quad \tilde{S}(x, j) = (x, j+1) \text{ whenever } x \in X, j = 0, 1, \dots, h-2$$

and

$$(7) \quad \tilde{S}(x, h-1) = (Sx, 0) \text{ for } x \in X.$$

It is not hard to see that $\tilde{S} \in \text{Aut}(\tilde{X}, \tilde{\mu})$ and it is called the *h-discrete suspension* of S (it is ergodic if and only if so is S).

Note that the map $(x, j) \mapsto j$ for $(x, j) \in \tilde{X}$ yields a factor map between \tilde{S} and the rotation $\tau_h : x \mapsto x+1$ on $\mathbb{Z}/h\mathbb{Z}$. Note also that $\tilde{S}^h(x, 0) = (Sx, 0)$, so in fact we can view \tilde{S} as the *h-discrete suspension* of $\tilde{S}^h|_{X \times \{0\}}$. As a matter of fact, an automorphism $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$ is an *h-discrete suspension* if and only if

$$(8) \quad R \text{ has the rotation } \tau_h \text{ as a factor.}$$

Indeed, if π settles a factor map between R and τ_h , then set $A_0 := \pi^{-1}(\{0\})$, note that $R^h(A_0) = A_0$ and show that R is isomorphic to the *h-discrete suspension* of $R^h|_{A_0}$.

¹⁰Without non-invertibility, we say that T is a *factor* of S .

Finally, consider $(m, h) = 1$. Then, it is easy to see that:

- (9) The h -discrete suspension $\widetilde{\tau}_m$ is isomorphic to direct product $\tau_m \times \tau_h$.

We refer the reader to [19] for more information on ergodic theory, in particular, on the theory of joinings.

1.3 Group and isometric extensions

Given (X, \mathcal{B}, μ, S) an ergodic dynamical system, consider a measurable $\varphi : X \rightarrow G$ (i.e. φ is a *cocycle*), where G is a compact metric group. Let m_G denote Haar measure of G . The automorphism $S_\varphi : X \times G \rightarrow X \times G$, given by

$$S_\varphi(x, g) = (Sx, \varphi(x)g),$$

is called a *compact group extension* of S . We obtain the dynamical system $(X \times G, \mathcal{B} \otimes \mathcal{B}(G), \mu \otimes m_G, S_\varphi)$ which need not be ergodic. For example, it is not ergodic when $\varphi(x) = \xi(Sx)^{-1}\xi(x)$ for a measurable $\xi : X \rightarrow G$, i.e. when φ is a *coboundary* (indeed, the map $(x, g) \mapsto (x, \xi(x)g)$ settles an isomorphism between S_φ and $S \times Id_G$). Note that $(S_\varphi)^m(x, g) = (S^m x, \varphi^{(m)}(x)g)$, where $\varphi^{(m)}(x) = \varphi(S^{m-1}x) \dots \varphi(Sx)\varphi(x)$ for $m \geq 1$.

Proposition 1.3 ([18]). *Assume that S and T are ergodic automorphisms on (X, \mathcal{B}, μ) and (Y, \mathcal{C}, ν) , respectively. Assume moreover that $S \perp T$ and S_φ, T_ψ are ergodic group extensions of S and T , respectively ($\psi : Y \rightarrow H$). If the product measure $(\mu \otimes m_G) \otimes (\nu \otimes m_H)$ is ergodic¹¹ then $S_\varphi \perp T_\psi$.*

The following results are also classical.

Lemma 1.4. *Assume that S_φ and T_ψ are ergodic and let $\tilde{\rho} \in J^e(S_\varphi, T_\psi)$. Then (up to a natural permutation of coordinates) $(X \times G \times Y \times H, \mathcal{B} \otimes \mathcal{B}(G) \otimes \mathcal{C} \otimes \mathcal{B}(H), \tilde{\rho}, S_\varphi \times T_\psi)$ is a compact group extension of $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \rho, S \times T)$, where $\rho := \tilde{\rho}|_{X \times Y}$.*

Moreover, if the relatively independent extension¹² $\hat{\rho}$ of ρ is ergodic then $\tilde{\rho} = \hat{\rho}$.

A group extension is a special case of so called *skew products*. Assume that we have a measurable map $\Sigma : X \rightarrow \text{Aut}(Z, \mathcal{D}, \rho)$. Then we can consider S_Σ :

$$S_\Sigma(x, z) := (Sx, \Sigma(x)(z))$$

which is an automorphism of $(X \times Z, \mathcal{B} \otimes \mathcal{D}, \mu \otimes \rho)$. If additionally, Z is a compact metric space and $\Sigma(x)$, $x \in X$, are isometric then we call S_Σ an *isometric extension*. Since the group $\text{Iso}(Z)$ of isometries of Z considered with the uniform topology is a compact metric group (by Arzela-Ascoli theorem), it is not hard to see that each isometric extension is a factor of a group extension (especially, if

¹¹It means that the only situation in which we loose disjointness is when the cocycles φ and ψ “add” a common non-trivial eigenvalue for U_{S_φ} and U_{T_ψ} .

¹²This measure is defined by

$$\hat{\rho}(\tilde{A} \times \tilde{B}) = \int_X \mathbb{E}(\tilde{A}|X)(x) \mathbb{E}(\tilde{B}|X)(y) d\rho(x, y) \text{ for Borel subsets } \tilde{A}, \tilde{B} \subset X \times G.$$

we assume that the isometric extension is uniquely ergodic).¹³ If the isometric extension is ergodic, one can choose the group extension also ergodic.

Remark 1.5. If S_φ is a group extension, then for each closed subgroup $F \subset G$, the automorphism $S_{\varphi,F}$, given by

$$S_{\varphi,F}(x, gF) := (Sx, \varphi(x)gF),$$

is an isometric extension of S ; it acts on the space $X \times G/F$ considered with $\mu \otimes m_{G/F}$, where the measure $m_{G/F}$ on the homogenous space G/F is the natural image of Haar measure m_G .

Remark 1.6. Each finite extension is isometric. It is a factor of a group extension by G , where G is finite.

1.4 Odometers

Odometers are given by the inverse limits of cyclic groups: we have $n_{t-1} | n_t$ for each $t \geq 1$ and

$$X = H_{(n_t/n_{t-1})} := \liminv \mathbb{Z}/n_t\mathbb{Z}$$

with the rotation R by 1 on each coordinate. If for each t , $n_{t+1}/n_t = \lambda \geq 2$, then we speak about λ -*odometer* and denote it by $H_\lambda (= H_{(\lambda)})$.

Each odometer (X, R) is uniquely ergodic (with the unique measure being Haar measure m_X of X). Then (X, m_X, R) has discrete spectrum with the group of eigenvalues given by all roots of unity of degree n_t , $t \geq 1$. Furthermore, R^r is ergodic (uniquely ergodic) iff $(r, n_t) = 1$ for each $t \geq 1$. In this case R^r and R are isomorphic as both are ergodic and their spectra are the same, so the claim follows by the Halmos-von Neumann theorem (e.g. [19]). It easily follows that whenever $p, q \in \mathcal{P}$ are different prime numbers (by \mathcal{P} we denote the set of prime numbers) not dividing any n_t then each $\rho \in J^e(R^p, R^q)$ is a graph joining (of an isomorphism between R^p and R^q), see e.g. [28].

1.4.1 h -discrete suspensions of odometers

Assume that $(h, n_t) = 1$ for each $t \geq 1$. Let \tilde{R} denote the h -discrete suspension of R . Then (cf. (9)), we obtain that

$$(10) \quad \tilde{R} \text{ is isomorphic to } R \times \tau_h.$$

Indeed, both automorphisms have discrete spectrum (and are ergodic). The group of eigenvalues of $U_{\tilde{R}}$ is equal to $\{e^{2\pi ij/(hn_t)} : j \in \mathbb{Z}, t \geq 0\}$, while the group of eigenvalues of $U_{R \times \tau_h}$ is generated by $\{e^{2\pi ij/n_t} : j \in \mathbb{Z}, t \geq 0\}$ and the group of h -roots of unity. It follows that $U_{R \times \tau_h}$ and $U_{\tilde{R}}$ have the same group of eigenvalues, hence again by the Halmos-von Neumann theorem, they are isomorphic.

¹³A group extension can be defined on $X \times \text{Iso}(Z)$, acting by the formula $(x, I) \mapsto (Sx, \Sigma(x) \circ I)$; to obtain a factor map, fix a point (transitive for $\text{Iso}(Z)$) $z_0 \in Z$ and consider $(x, I) \mapsto (x, I(z_0))$.

1.4.2 Group extensions of odometers - special assumptions

We will assume that R is an odometer with “small spectrum”, that is, the set $\{p \in \mathcal{P} : p|n_t \text{ for some } t \geq 1\}$ is finite (this will always be the case for λ -odometers). Moreover, we will assume that R_φ (more precisely, U_{R_φ}) has continuous spectrum on the space $L^2(X \times G, m_X \otimes m_G) \ominus L^2(X, m_X)$.¹⁴ It follows that if $p \in \mathcal{P}$ is sufficiently large then $(R_\varphi)^p$ is ergodic (uniquely ergodic if φ is continuous). Furthermore, we assume that if $p, q \in \mathcal{P}$ are different and sufficiently large then the only ergodic joinings between $(R_\varphi)^p$ and $(R_\varphi)^q$ are relatively independent extensions of isomorphisms between R^p and R^q .¹⁵ By all these assumptions $(X \times G \times X \times G, \widehat{(m_X)}_W, (R_\varphi)^p \times (R_\varphi)^q)$ is ergodic and has the same eigenvalues as the odometer R .

Lemma 1.7. *Assume that φ is continuous. Under the above assumptions for each (real-valued) $F \in C(X \times G)$ such that $F \perp L^2(X, m_X)$, we have*

$$\frac{1}{N} \sum_{n \leq N} F((R_\varphi)^{pn}(x, g)) F((R_\varphi)^{qn}(x, g)) \rightarrow 0$$

for each $(x, g) \in X \times G$ and different primes p, q sufficiently large.

Proof. First notice that any accumulation point ρ of $(\frac{1}{N} \sum_{n \leq N} \delta_{(R^p \times R^q)^n(x, x)})$ is ergodic (cf. e.g. [25]), hence is graphic. It follows that any accumulation point $\tilde{\rho}$ of $(\frac{1}{N} \sum_{n \leq N} \delta_{((R_\varphi)^p \times (R_\varphi)^q)^n((x, g), (x, g))})$ will be the relatively independent extension of an isomorphism between R^p and R^q in view of the second part of Lemma 1.4. By the definition of the relatively independent extension, $\int F \otimes F d\tilde{\rho} = 0$ and the result follows. \square

1.5 Application - KBSZ criterion

Recall the following result (to which we refer as the KBSZ criterion) about the orthogonality of numerical sequences with bounded, multiplicative functions:

Theorem 1.8 (Kátai [23], Bourgain, Sarnak and Ziegler [6]). *Let a bounded sequence $(a_n) \subset \mathbb{C}$ satisfy*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_{pn} \bar{a}_{qn} = 0$$

for each $p \neq q$ sufficiently large prime numbers. Then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_n \mathbf{u}(n) = 0$$

for each bounded, multiplicative function $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$.

¹⁴If V is a unitary operator on a Hilbert space \mathcal{H} then it is said to have *continuous spectrum* if for all $x \in \mathcal{H}$ the spectral measure σ_x is continuous; the latter measure (on \mathbb{S}^1) is determined by its Fourier transform: $\hat{\sigma}_x(m) := \int_{\mathbb{S}^1} z^m d\sigma_x(z) = \langle V^m x, x \rangle$ for all $m \in \mathbb{Z}$.

¹⁵We recall that if $W : X \rightarrow X$ settles an isomorphism of R^p and R^q , then it yields an ergodic joining $(m_X)_W(A \times B) := m_X(A \cap W^{-1}B)$; its relatively independent extension $\widehat{(m_X)}_W$ is defined by

$$\widehat{(m_X)}_W(\tilde{A} \times \tilde{B}) = \int_X \mathbb{E}(\tilde{A}|X)(x) \mathbb{E}(\tilde{B}|X)(Wx) dm_X(x) \text{ for Borel subsets } \tilde{A}, \tilde{B} \subset X \times G.$$

In the context of topological dynamical systems, that is, given (X, S) , we use this result with $a_n = f(S^n x)$ with $f \in C(X)$ and $x \in X$. It is not hard to see that this criterion applies for any uniquely ergodic (X, S) with the property that $S^p \perp S^q$ (disjointness is meant if we consider the unique invariant measure μ) and we consider $f \in C(X)$ with $\int_X f d\mu = 0$ and arbitrary $x \in X$. However, even if we do not have disjointness of sufficiently large (prime) powers, we can apply this criterion for **particular** continuous functions if we control the limit joinings, as, for example, in Lemma 1.7.

2 Basics on substitutions of constant length

Let \mathbb{A} be an alphabet (a non-empty finite set). Denote $\mathbb{A}^* := \bigcup_{m \geq 0} \mathbb{A}^m$, where \mathbb{A}^m stands for the set of words $w = a_0 a_1 \dots a_{m-1}$ over \mathbb{A} of length $|w|$ equal to m (\mathbb{A}^0 consists only of the empty word). Fix $\mathbb{N} \ni \lambda \geq 2$. By a *substitution of (constant) length λ* we mean a map

$$\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$$

which we also write as $\theta(a) = \theta(a)_0 \theta(a)_1 \dots \theta(a)_{\lambda-1}$ for $a \in \mathbb{A}$. Via the concatenation of words, there is a natural extension of θ to a map from \mathbb{A}^m to $\mathbb{A}^{m\lambda}$ (for each $m \geq 1$) or from \mathbb{A}^* to itself, or even from $\mathbb{A}^{\mathbb{Z}}$ to itself. In particular, we can iterate θ k times:

$$\mathbb{A} \xrightarrow{\theta} \mathbb{A}^\lambda \xrightarrow{\theta} \dots \xrightarrow{\theta} \mathbb{A}^{\lambda^k}$$

which can be viewed as the substitution θ^k (the k th-iterate of θ) of length λ^k :

$$\theta^k : \mathbb{A} \rightarrow \mathbb{A}^{\lambda^k}.$$

The following formula is well-known and follows directly by definition:

$$(11) \quad \theta^{k+\ell}(a)_{j'\lambda^k+j} = \theta^k(\theta^\ell(a)_{j'})_j$$

for each $a \in \mathbb{A}$, $j' < \lambda^\ell$, $j < \lambda^k$. Indeed, $|\theta^\ell(a)| = \lambda^\ell$ and consider $\theta^\ell(a)_{j'}$, i.e. the j' -th letter in the word $\theta^\ell(a)$. We now let act θ^k on $\theta^\ell(a)$ which transforms letters in the word $\theta^\ell(a)$ into blocks of length λ^k , in particular the j' -th letter of $\theta^\ell(a)$ becomes the j' -th block of length λ^k . So counting j -th letter in this block is the same as counting $(j'\lambda^k + j)$ -th letter in $\theta^{k+\ell}(a)$.

The *subshift* $X_\theta \subset \mathbb{A}^{\mathbb{Z}}$ is determined by all words that appear in $\theta^k(a)$ for some $k \geq 1$ and $a \in \mathbb{A}$:

$$(12) \quad X_\theta := \left\{ x \in \mathbb{A}^{\mathbb{Z}} : \begin{array}{l} \text{for each } r < s, \text{ we have} \\ x[r, s] = \theta^k(a)[i, i + s - r - 1] \\ \text{for some } k \geq 1, a \in \mathbb{A} \text{ and } 0 \leq i < \lambda^k - (s - r) \end{array} \right\}.$$

That is, X_θ is closed and invariant for the left shift S acting on $\mathbb{A}^{\mathbb{Z}}$. Then, clearly, we have

$$(13) \quad X_{\theta^k} \subset X_\theta \text{ for each } k \geq 1.$$

Note also that for each $a \in \mathbb{A}$, there exist $k, \ell \geq 1$ such that

$$\theta^k(a)_0 = \theta^{k+\ell}(a)_0$$

from which we deduce that there are a letter $a \in \mathbb{A}$ and $\ell \geq 1$ such that

$$(14) \quad \theta^\ell(a)_0 = a_0.$$

Hence, by iterating the substitution θ^ℓ , we obtain a fixed point $u \in \mathbb{A}^{\mathbb{N}}$ for the map $\theta^\ell : \mathbb{A}^{\mathbb{N}} \rightarrow \mathbb{A}^{\mathbb{N}}$. This, similarly to the RHS of (12) defines a subshift $X_u \subset \mathbb{A}^{\mathbb{Z}}$ for which we have $X_u \subset X_\theta$, cf. Remark 1.2. In general, we do not have equalities in (13), and (the more) X_u is a proper subshift of X_{θ^ℓ} above. The situation changes if we assume that θ is *primitive*, that is, when there exists $k \geq 1$ such that

$$(15) \quad \text{for each } a \in \mathbb{A} \text{ the word } \theta^k(a) \text{ contains all letters from } \mathbb{A}.$$

Then, we have equalities in (13), and moreover, for each $\ell \geq 1$ if $u \in \mathbb{A}^{\mathbb{N}}$ satisfies $\theta^\ell(u) = u$ then $X_u = X_\theta$. Moreover, we have

Proposition 2.1. *Assume that θ is a primitive substitution of constant length λ . Then (X_θ, S) is minimal. Moreover, there exists exactly one invariant measure μ_θ on X_θ .*

In what follows, we assume that $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ is primitive. Under this assumption, it follows directly by the Perron-Frobenius theorem that, for any letter $a \in \mathbb{A}$, the density

$$\delta_a := \lim_{n \rightarrow \infty} \frac{\#\{j < \lambda^n : \theta^n(a')_j = a\}}{\lambda^n}$$

exists and is independent of $a' \in \mathbb{A}$. More precisely, first, denote by $M(\theta) \in \mathbb{Z}^{|\mathbb{A}| \times |\mathbb{A}|}$ the incidence matrix of θ , i.e. $M_{a,a'}(\theta) := |\theta(a')|_a :=$ the number of occurrences of a in $\theta(a')$. The vector $(\delta_a)_{a \in \mathbb{A}}$ is then the unique right eigenvector for the (maximal) eigenvalue λ , i.e.

$$(16) \quad \lambda \cdot \delta_a = \sum_{a' \in \mathbb{A}} M(\theta)_{a,a'} \cdot \delta_{a'}$$

holds for all $a \in \mathbb{A}$ and uniquely defines $(\delta_a)_{a \in \mathbb{A}}$, [36].

As θ is primitive, we can also assume that for some $a_0 \in \mathbb{A}$, we have

$$(17) \quad \theta(a_0)_0 = a_0.$$

By iterating θ at a_0 we obtain $u \in \mathbb{A}^{\mathbb{N}}$ such that $\theta(u) = u$ (and $X_u = X_\theta$).

We now recall the definition of the *height* $h(\theta)$ of θ following [36]. Therefore, for $k \geq 0$, set

$$(18) \quad S_k = S_k(\theta) := \{r \geq 1 : u[k+r] = u[k]\}$$

and

$$(19) \quad g_k := \gcd S_k.$$

This allows us to define

$$(20) \quad h(\theta) := \max\{m \geq 1 : (m, \lambda) = 1, m|g_0\}.$$

We list now some basic properties of $h = h(\theta)$.

1. $1 \leq h \leq |\mathbb{A}|$.
2. $h = \max\{m \geq 1 : (m, \lambda) = 1, m|g_k\}$ for every $k \geq 0$.
3. If, for $j = 0, \dots, h-1$, we consider the set

$$C_j := \{u[j + rh] : r \geq 0\},$$

then, by 2., the letter $u[j]$ can be equal to some $u[s]$ only if $s \in C_j$. Hence, the sets C_j form a partition of \mathbb{A} . If we identify, in u , the letters in the same set C_j , $j = 0, \dots, h-1$, we thus obtain a periodic (of period h) sequence (cf. (8)), and h is the largest integer $\leq |\mathbb{A}|$, coprime to λ , with this property.

Moreover, if by $\mathbb{A}^{(h)}$ we denote the set $\{u[mh(\theta), mh(\theta) + h(\theta) - 1] : m \geq 0\} \subseteq \mathbb{A}^h$ (which is of course finite) then we can define $\eta : \mathbb{A}^{(h)} \rightarrow (\mathbb{A}^{(h)})^\lambda$ by setting

$$\eta(w) = \theta(w)[0, h(\theta) - 1]\theta(w)[h(\theta), 2h(\theta) - 1] \dots \theta(w)[(\lambda - 1)h(\theta), \lambda h(\theta) - 1].$$

Then η turns out to be a primitive substitution and $h(\eta) = 1$. Moreover, in view of 3. above, as a (measure-theoretic) dynamical system, $(X_\theta, \mu_\theta, S)$ is isomorphic to the $h(\theta)$ -discrete suspension \tilde{S} of (X_η, μ_η, S) . In what follows we will denote η by $\theta^{(h)}$.

If θ is primitive then so is θ^k (for each $k \geq 1$). Take now u a fixed point for θ . This is also a fixed point of θ^k . Moreover, $h = h(\theta) = h(\theta^k)$. It follows easily that

$$(21) \quad (\theta^{(h)})^k = (\theta^k)^{(h)}.$$

Note also that if $(a_0, \dots, a_{h-1}) \in \mathbb{A}^{(h)}$ then

$$\theta^{(h)}(a_0, \dots, a_{h-1}) = \theta(a_0)\theta(a_1) \dots \theta(a_{h-1}),$$

where the block on the RHS is divided into blocks of length h which are elements of $\mathbb{A}^{(h)}$. The j -th such element ($j < \lambda$) has the beginning at the position jh which is of the form

$$jh = i\lambda + y, \quad 0 \leq y < \lambda$$

with $i = \lfloor jh/\lambda \rfloor$ (and $y = jh - i\lambda$). It follows that if

$$\theta^{(h)}(a_0, \dots, a_{h-1})_j = (b_0, \dots, b_{h-1})$$

then $b_0 = \theta(a_i)_{jh - i\lambda}$. Replacing in this reasoning θ by θ^k and using (21), we obtain the following: If $(a_0, \dots, a_{h-1}) \in \mathbb{A}^{(h)}$, $k \geq 1$, $j < \lambda^k$ and

$$(\theta^{(h)})^k(a_0, \dots, a_{h-1})_j = (b_0, \dots, b_{h-1})$$

then

$$(22) \quad b_0 = \theta^k(a_i)_{jh - i\lambda^k}, \quad \text{where } i = \lfloor jh/\lambda^k \rfloor.$$

Following [8], a substitution θ is called *pure* if $\theta = \theta^{(h)}$, i.e. if $h(\theta) = 1$.

Denote by $c = c(\theta)$ the *column number* of θ . Recall its definition: we consider iterations $\theta^k : \mathbb{A} \rightarrow \mathbb{A}^{\lambda^k}$ and each time we consider sets $\{\theta^k(a)_j : a \in \mathbb{A}\}$ for $j = 0, \dots, \lambda^k - 1$, where $\theta^k(a) = \theta^k(a)_0\theta^k(a)_1 \dots \theta^k(a)_{\lambda^k - 1}$. Then $c(\theta)$ is defined as the minimal cardinality (we run over $k \geq 1$ and $0 \leq j \leq \lambda^k - 1$) of such sets.

In what follows we will make use of the following observation.

Lemma 2.2. *Let $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ be a primitive substitution with $c(\theta) = 1$, then θ is pure, i.e. we have $h(\theta) = 1$.*

Proof. By assumption, there exist $a \in \mathbb{A}$, $k \in \mathbb{N}$ and $0 \leq j < \lambda^k$ such that $\theta^k(x)_j = a$ for all $x \in \mathbb{A}$. Recalling that the fixpoint of θ is u we find that $u[j] = u[j + \lambda^k] = a$. This shows that $\lambda^k \in S_j(\theta)$ and thus, $g_j \mid \lambda^k$. As $h(\theta) \mid g_j$ and is coprime with λ , we have $h(\theta) = 1$. \square

In fact, more is true:

Lemma 2.3. *Let $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ be a primitive substitution. Then*

$$c(\theta) = h(\theta) \cdot c(\theta^{(h)}).$$

In particular, $h(\theta) \mid c(\theta)$.

This result seems to be a folklore but we could not find a proof of it in the literature. We postpone the proof of Lemma 2.3 to Section 6.4.

The following result is well known for pure substitutions but a use of Lemma 2.3 yields the following:

Proposition 2.4. *Assume that $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ is primitive. Then, as the measure-theoretic dynamical system, $(X_\theta, \mu_\theta, S)$ is isomorphic to a $c(\theta)$ -point extension (cf. Remark 1.6) of the λ -odometer $(H_\lambda, m_{H_\lambda}, R)$.*

Proof. We have an isomorphism of X_θ (we omit automorphisms and measures) with $\widetilde{X_{\theta^{(h)}}}$. If $\pi : X_{\theta^{(h)}} \rightarrow H_\lambda$ denotes the factor map which is $c(\theta^{(h)})$ to 1 (a.e.), then $\pi \times Id_{\mathbb{Z}/h\mathbb{Z}}$ yields a factor map between the h -discrete suspensions $\widetilde{X_{\theta^{(h)}}}$ and $\widetilde{H_\lambda}$. It is again $c(\theta^{(h)})$ to 1 (a.e.). However, in view of (10), the suspension $\widetilde{H_\lambda}$ is isomorphic to the direct product $H_\lambda \times \mathbb{Z}/h\mathbb{Z}$.¹⁶ This makes H_λ a factor of X_θ with (a.e.) fiber of cardinality $c(\theta^{(h)}) \cdot h$. The result follows from Lemma 2.3. \square

As a matter of fact, whenever θ is primitive and $h(\theta) = 1$, the system (H_λ, R) represents the so called maximal equicontinuous factor of (X_θ, S) . The corresponding factor map is usually seen in the following way: Each point $x \in X_\theta$ has a unique λ^t -skeleton structure. By that, one means a sequence $(j_t)_{t \geq 1}$ with $0 \leq j_t < \lambda^t - 1$ for which, for each $t \geq 1$, we have

$$x[-j_t + s\lambda^t, -j_t + (s+1)\lambda^t - 1] = \theta^t(c_s)$$

for each $s \in \mathbb{Z}$ and some $c_s \in \mathbb{A}$. Now, the map $x \mapsto (j_t)$ yields the factor map which we seek.

Notice also that if $\mathcal{R} \subset \mathbb{A} \times \mathbb{A}$ is an equivalence relation which is θ -consistent, that is:

$$(23) \quad (a, b) \in \mathcal{R} \Rightarrow (\theta(a)_j, \theta(b)_j) \in \mathcal{R} \text{ for each } j = 0, \dots, \lambda - 1,$$

¹⁶The factor H_λ is represented in $H_\lambda \times \mathbb{Z}/h\mathbb{Z}$ as the first coordinate σ -algebra and it is a factor of $\widetilde{H_\lambda}$ represented by an invariant σ -algebra. However, because of ergodicity and the fact that H_λ has discrete spectrum, there is only one invariant σ -algebra in $\widetilde{H_\lambda}$ representing H_λ . The same argument shows that H_λ is a factor of X_θ in a canonical way.

then we have the quotient substitution $\theta_{\mathcal{R}} : \mathbb{A}/\mathcal{R} \rightarrow (\mathbb{A}/\mathcal{R})^\lambda$ given by

$$\theta_{\mathcal{R}}([a]) := [\theta(a)_0] \dots [\theta(a)_{\lambda-1}]$$

is correctly defined and the dynamical system $(X_{\theta_{\mathcal{R}}}, S)$ is a (topological) factor of (X_θ, S) . Let alone $(X_{\theta_{\mathcal{R}}}, \mu_{\theta_{\mathcal{R}}}, S)$ is a measure-theoretic factor of $(X_\theta, \mu_\theta, S)$ (clearly, $\theta_{\mathcal{R}}$ is primitive). A particular instance of a θ -consistent equivalence relation \mathcal{R} given by

$$(a, b) \in \mathcal{R} \text{ if and only if } \theta(a) = \theta(b).$$

In this situation, the quotient dynamical system $(X_{\theta_{\mathcal{R}}}, S)$ is in fact (topologically) isomorphic to (X_θ, S) . Therefore, no harm arises if we assume that

$$(24) \quad \theta \text{ is 1-1 on letters, i.e., } \theta(a) \neq \theta(b) \text{ whenever } a \neq b.$$

Finally, we assume that

$$(25) \quad \theta \text{ is aperiodic,}$$

that is, there is a non-periodic element $x \in X_\theta$. In fact, since we assume that θ is primitive, θ is aperiodic if and only if X_θ is infinite.

From now on, we consider only substitutions θ satisfying (15), (17), (24) and (25). Often, we will additionally assume that θ is pure.

3 Proof of Theorem 0.1 (i)

This and Section 5 will be devoted to prove Theorem 0.1 in some particular cases. We assume that $c(\theta) = 1$, so according to Proposition 2.4 we deal (from ergodic theory point of view) with discrete spectrum case.¹⁷ In fact, we will prove even a stronger property.

So we have $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$, and we assume that for some $a \in \mathbb{A}$, $\theta(a)_0 = a$. Moreover, by replacing θ by its iterate if necessary, we can assume that

$$|\{0 \leq j < \lambda : |\{\theta(b)_j : b \in \mathbb{A}\}| = 1\}| := \ell_1 \geq 1.$$

Hence, since $u = \theta(u)$ is a concatenation of words $\theta(b)$, $b \in \mathbb{A}$, in the interval $[0, \lambda - 1]$ the number of positions j , for which $u[j + s\lambda] = u[j]$ for each $s \geq 1$ is ℓ_1 . Let us pass to θ^2 . We are interested in

$$\ell_2 := |\{0 \leq j < \lambda^2 - 1 : |\{\theta^2(b)_j : b \in \mathbb{A}\}| = 1\}|.$$

We have $\ell_2 \geq (\lambda - \ell_1)\ell_1 + \ell_1\lambda$. Inductively, if

$$\ell_{k-1} := |\{0 \leq j < \lambda^{k-1} - 1 : |\{\theta^{k-1}(b)_j : b \in \mathbb{A}\}| = 1\}|$$

then

$$(26) \quad \ell_k \geq (\lambda - \ell_1)\ell_{k-1} + \ell_1\lambda^{k-1}.$$

¹⁷A prominent example in this class is the Baum-Sweet sequence given by the substitution $a \mapsto ab, b \mapsto cb, c \mapsto bd$ and $d \mapsto dd$. It is not hard to see $\{\theta^3(y)_5 : y \in \{a, b, c, d\}\} = \{d\}$.

Moreover, by definition,

$$(27) \quad \text{the sequence } \ell_k/\lambda^k, k \geq 1, \text{ is increasing.}$$

In view of (26), we have

$$\frac{\ell_k}{\lambda^k} \geq \frac{\lambda - \ell_1}{\lambda} \frac{\ell_{k-1}}{\lambda^{k-1}} + \frac{\ell_1}{\lambda},$$

and since, by (27), the sequence ℓ_k/λ^k is convergent, the above recurrence formula implies

$$(28) \quad \lim_{k \rightarrow \infty} \frac{\ell_k}{\lambda^k} = 1.$$

Note that an interpretation of ℓ_k is that it is the number of coordinates j for $j = 0, 1, \dots, \lambda^k - 1$ such that $u[j + s\lambda^k] = u[j]$ for each $s \geq 1$.

Recall now the notion of Weyl rationally almost periodic sequences (WRAP). A sequence $(x[n]) \in \mathbb{A}^{\mathbb{N}}$ is WRAP if it is the limit of periodic sequences in the Weyl pseudo-metric d_W :

$$d_W(z, z') = \limsup_{N \rightarrow \infty} \sup_{m \geq 0} \frac{1}{N} |\{m \leq n \leq m + N - 1 : z[n] \neq z'[n]\}|.$$

Now, (28) implies immediately that:

Proposition 3.1. *If $c(\theta) = 1$ then the fixed point $u = \theta(u)$ is WRAP, where the periodic sequences can be chosen with period λ^k .*

From the point of view of Möbius disjointness, dynamical systems (X_x, S) given by WRAP sequences x have already been studied in [4] and it is proved there that all continuous observables $(f(S^n y))$ (for $f \in C(X_x)$) are orthogonal to the Möbius function μ . But a rapid look at the proof in [4] shows that the only property of μ used in it was the aperiodicity of μ (what is essential in the proof is that all points $y \in X_x$ are also WRAP).

Corollary 3.2 ([4]). *If $x \in \mathbb{A}^{\mathbb{N}}$ is WRAP then for each bounded $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$ which is aperiodic, we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(S^n y) \mathbf{u}(n) = 0$$

for all $f \in C(X_x)$ and $y \in X_x$. In particular, the above assertion holds for substitutional dynamical systems with $c(\theta) = 1$.¹⁸

Remark 3.3. Theorem 0.1 (i) also follows from [10] but because the relations between synchronized automata and substitutions with $c(\theta) = 1$ do not seem to be explained explicitly in literature, we gave a more general and direct argument.

¹⁸The automatic sequences given by observables in case $c(\theta) = 1$ represent so called *synchronized* case in [35].

4 More ergodic theory prerequisites

4.1 Essential centralizer

Assume that T is an ergodic¹⁹ automorphism of (Y, \mathcal{C}, ν) . By the *centralizer* $C(T)$ of T we mean the group of all invertible automorphisms $V \in \text{Aut}(Y, \mathcal{C}, \nu)$ commuting with T . Clearly, $\{T^n : n \in \mathbb{Z}\}$ is a normal subgroup of $C(T)$ and the group $EC(T) := C(T)/\{T^n : n \in \mathbb{Z}\}$ is called the *essential centralizer* of T .

Lemma 4.1. *Assume that $EC(T)$ is finite and $C(T) = C(T^p)$ for all sufficiently large $p \in \mathcal{P}$. Then, for all sufficiently large $p, q \in \mathcal{P}$, $p \neq q$, the automorphisms T^p and T^q are not isomorphic.*

Proof. Since $EC(T)$ is finite, we have

$$C(T) = \{T^n V_i : n \in \mathbb{Z}, i = 0, 1, \dots, K\},$$

where $V_0 = Id$ and $V_i \in C(T) \setminus \{T^j : j \in \mathbb{Z}\}$ for $i = 1, \dots, K$. For $i = 1, \dots, K$, let $m_i \geq 1$ be the smallest natural number such that $V_i^{m_i} \in \{T^j : j \in \mathbb{Z}\}$ (of course m_i is precisely the order of the coset given by V_i in $C(T)/\{T^j : j \in \mathbb{Z}\}$). In what follows, we consider only prime numbers p, q which are dividing no m_i for $i = 1, \dots, K$. Suppose that for $p \neq q$ (sufficiently large) we have an isomorphism of T^p and T^q . As obviously T^q has a q -root, it follows that there is a root of degree q of T^p , i.e. there exists $W \in \text{Aut}(Y, \mathcal{C}, \nu)$ such that $W^q = T^p$.²⁰ Now, $W \in C(T^p)$, hence (by assumption) $W \in C(T)$. It follows that, for some $n \in \mathbb{Z}$ and $0 \leq i \leq K$, $W = T^n \circ V_i$, whence $W^q = (T^n \circ V_i)^q$ and we obtain

$$V_i^q = T^{p-nq}$$

which, if $i \neq 0$, is impossible as q is coprime to m_i and if $i = 0$, $T^p = T^{nq}$ which yields $q|p$ (or $p = 0$ if $n = 0$), a contradiction. \square

Remark 4.2. Notice that in the above proof, in fact we proved that T^p cannot be isomorphic to U^q with $U \in \text{Aut}(Y, \mathcal{C}, \nu)$, in other words, we proved that T^p cannot have a q -root.

4.1.1 Centralizer of h -discrete suspensions

We assume that $T \in \text{Aut}(Y, \mathcal{C}, \nu)$ and let \tilde{T} denote its h -discrete suspension, see (6) and (7). Note that whenever $V \in C(T)$, the formula $\bar{V}(y, j) := (Vy, j)$ for $(y, j) \in \tilde{Y}$ defines an element of the centralizer of \tilde{T} . In fact, we have the following:

Proposition 4.3 (Cor. 1.4 in [7]). *If $T \in \text{Aut}(Y, \mathcal{C}, \nu)$ is ergodic then $C(\tilde{T}) = \{\bar{V} \circ \tilde{T}^m : V \in C(T), m \in \mathbb{Z}\}$.*

It follows immediately from Proposition 4.3 that:

Corollary 4.4. *If $T \in \text{Aut}(Y, \mathcal{C}, \nu)$ is ergodic and $EC(T)$ is finite, also $EC(\tilde{T})$ is finite.*

¹⁹We assume also that T is *aperiodic*: for ν -a.e. $y \in Y$, the map $m \mapsto T^m y$ is one-to-one on \mathbb{Z} .

²⁰Note that if τ is an automorphism and $\tau = \sigma^q$ for another automorphism σ , then $\sigma \in C(\tau)$ as $\sigma \circ \tau = \sigma \circ \sigma^q = \sigma^q \circ \sigma = \tau \circ \sigma$.

4.1.2 Essential centralizer of substitutions of constant length

The result below has been proved by Host and Parreau in [22] for pure substitutions of constant length. However, taking into account Corollary 4.4 and the fact that each substitutional system is an h -discrete suspension of its pure basis (which is also given by a substitution of constant length), see Section 2, we obtain the following²¹ result.

Theorem 4.5 ([22]). *Let θ be a primitive substitution of constant length. Then (remembering that (X_θ, S) is uniquely ergodic), $EC(S)$ is finite if and only if $c(\theta) > h(\theta)$.*

Remark 4.6. If $c(\theta) = h(\theta)$ (in particular, if $c(\theta) = 1$) then we are in the synchronizing case. Thus, the spectrum of the corresponding dynamical system is discrete. Therefore, the essential centralizer is uncountable. In fact, S^p is isomorphic to S^q for all sufficiently large $p, q \in \mathcal{P}$.

Remark 4.7. Assume that $V : \mathbb{A} \rightarrow \mathbb{A}$ is a bijection “commuting” with θ :

$$(29) \quad \theta(V(a))_j = V(\theta(a))_j$$

for each $j = 0, 1, \dots, \lambda - 1$. Then, we claim that V has a natural extension to a homeomorphism $V : X_\theta \rightarrow X_\theta$ so that $S \circ V = V \circ S$. Indeed, treat V as a 1-code map on $\mathbb{A}^{\mathbb{Z}}$ (so obviously, it commutes with the shift). Clearly, we only need to show that

$$y \in X_\theta \Rightarrow Vy \in X_\theta.$$

Now, for each $i \leq j$, $y[i, j] = \theta^k(a)[m, m + (j - i)]$ for some $a \in \mathbb{A}$, $k \geq 1$ and $m \geq 0$. It follows from (29) that

$$(Vy)[i, j] = V\theta^k(a)[m, m + (j - i)] = \theta^k(Va)[m, m + (j - i)]$$

and the claim follows. Note that if θ is primitive, then $V : X_\theta \rightarrow X_\theta$ preserves μ_θ , so $V \in C(S)$.

4.2 Joinings of powers of finite extensions of odometers

Lemma 4.8 ([13]). *Assume that T acting on (Y, \mathcal{C}, ν) is ergodic (aperiodic) and has discrete spectrum. Let $\varphi : Y \rightarrow G$ be a cocycle with G finite. Assume that T_φ is ergodic. Moreover, assume that for $p \in \mathcal{P}$ large enough the corresponding group extension $(T_\varphi)^p$ is also ergodic.²² Then for $p \in \mathcal{P}$ large enough, we have $C(T_\varphi) = C((T_\varphi)^p)$.*

Using Lemmas 4.1 and 4.8, we obtain the following.

Proposition 4.9. *Assume that T_φ is an ergodic G -extension (G a finite group) $T \in \text{Aut}(Y, \mathcal{C}, \nu)$ with discrete spectrum, so that $EC(T_\varphi)$ is finite. Assume moreover that $(T_\varphi)^p$ is ergodic for all $p \in \mathcal{P}$ sufficiently large. Then $(T_\varphi)^p$ and $(T_\varphi)^q$ are not isomorphic for all $p, q \in \mathcal{P}$ which are different and large enough.*

²¹For the particular instance of bijective substitutions, the result was proved to hold slightly earlier in [27].

²²This assumption requires small spectrum of U_T .

In case of odometers, we can prove more. Assume that R is the odometer given by λ (that is, $R \in \text{Aut}(H_\lambda, m_{H_\lambda})$ is ergodic, has discrete spectrum and the eigenvalues are roots of unity of degree λ^k , $k \geq 1$). Consider $p, q \in \mathcal{P}$ relatively prime with λ . Then R is isomorphic to R^p (also to R^q), so R^p is isomorphic to R^q . Recall also that the only **ergodic** joinings between R^p and R^q are given by the graphs of isomorphisms between R^p and R^q .

Proposition 4.10. *Assume that R is the λ -odometer. Let $\varphi : H_\lambda \rightarrow G$ be a cocycle with G finite, so that R_φ is ergodic, and $(R_\varphi)^p$ is also ergodic for $p \in \mathcal{P}$ large enough. Assume moreover that $EC(R_{\varphi F})$ is finite for each proper subgroup $F \subset G$.²³ Then, for each different $p, q \in \mathcal{P}$ large enough, the **only** joinings²⁴ between $(R_\varphi)^p$ and $(R_\varphi)^q$ that project onto ergodic joinings of R^p and R^q are relatively independent extensions of the projections.*

Proof. A general theory of groups extensions (see e.g. [28], [34]) tells us that if

$$\rho \in J^e((R_\varphi)^p, (R_\varphi)^q) = J^e(R_{\varphi^{(p)}}^p, R_{\varphi^{(q)}}^q)$$

then there exist $F_1, F_2 \subset G$ normal subgroups of G such that an isomorphism W between R^p and R^q (we have assumed that $\rho|_{H_\lambda \times H_\lambda}$ is the graph of W) lifts to an isomorphism \widetilde{W} of the factors given by $H_\lambda \times G/F_1$ and $H_\lambda \times G/F_2$. But $R_{\varphi^{(p)}F_1}^p = (R_{\varphi F_1})^p$. Moreover, F_1, F_2 depend on p, q but altogether we have only finitely many possibilities for F_1, F_2 . Assume that F_1 is a proper subgroup of G . We use our assumption and Lemma 4.1 together with Remark 4.2 to get a contradiction. It follows that we can obtain only relatively independent extensions of graphs as joinings between $(R_\varphi)^p$ and $(R_\varphi)^q$ whenever $p \neq q$ are large enough. \square

Remark 4.11. By Mentzen's theorem [33] on (partly continuous spectrum) factors of substitutions of constant length, when R_φ is given by a substitution (non-synchronizing, primitive) then the natural factors $R_{\varphi F}$ for F proper subgroup of G are also (up to measure-theoretic isomorphism) substitutions (cf. Footnote 23), so the assumptions of Proposition 4.10 will be satisfied.

4.3 Strategy of the proof of the main result

To prove Theorem 0.1, we first show that each substitutional system (X_θ, S) is a **topological** factor of another substitutional system $(X_{\widehat{\Theta}}, S)$, where, from the measure-theoretic point of view, the system $(X_{\widehat{\Theta}}, S)$, which is uniquely ergodic, is isomorphic to R_φ satisfying the assumption of Proposition 4.10, see Remark 4.11. Moreover, we will show that the odometer of the original substitutional system is R .

Let (H_λ, R) denote the λ -odometer associated with $(X_{\widehat{\Theta}}, S)$. If the height of $\widehat{\Theta}$ is one, the λ -odometer is the maximal equicontinuous factor²⁵ of $(X_{\widehat{\Theta}}, S)$. Hence, there is a continuous equivariant map $\pi : X_{\widehat{\Theta}} \rightarrow H_\lambda$. Moreover, the following observation we already used in the proof of Lemma 1.7:

²³ We recall that the factor, as automorphism, is given by $R_{\varphi F}(x, gF) = (Rx, \varphi(x)gF)$. Note that, when applied to substitutions, this assumption requires H_λ to represent the Kronecker factor, cf. Remark 4.6. However, if we replace H_λ by $H_\lambda \times \mathbb{Z}/h\mathbb{Z}$, this assumption is satisfied and the proof of Proposition 4.10 remain the same.

²⁴ Note that we drop the assumption of ergodicity, cf. the second part of Lemma 1.4.

²⁵ If $h(\widehat{\Theta}) > 1$ then we must replace H_λ with $H_\lambda \times \mathbb{Z}/h\mathbb{Z}$.

Lemma 4.12. *Take different $p, q \in \mathcal{P}$ sufficiently large. Then, for each $y \in H_\lambda$, the point (y, y) is generic for an ergodic $R^p \times R^q$ invariant measure κ , i.e. for each $F \in C(H_\lambda \times H_\lambda)$, we have*

$$\frac{1}{N} \sum_{n \leq N} F(R^{pn}y, R^{qn}y) \rightarrow \int_{H_\lambda \times H_\lambda} F d\kappa.$$

It follows that (y, y) is generic for $\kappa = (m_{H_\lambda})_W$, the graph joining given by an isomorphism W between R^p and R^q .

Take any $x \in X_{\hat{\Theta}}$. We study now the sequence ($p \neq q$ sufficiently large)

$$\frac{1}{N} \sum_{n \leq N} \delta_{(S^p \times S^q)^n(x, x)}, \quad N \geq 1.$$

Any of its limit points yields a joining $\rho \in J((R_\varphi)^p, (R_\varphi)^q)$. Now, $\rho|_{H_\lambda \times H_\lambda}$ is precisely obtained as the limit of

$$\frac{1}{N_k} \sum_{n \leq N_k} \delta_{(R^p \times R^q)^n(\pi(x), \pi(x))},$$

for a relevant subsequence (N_k) . But we have already noticed that any such limit must be an **ergodic** joining of R^p and R^q , hence it is a graph. Now, we use our results to obtain that the limit of $\frac{1}{N_k} \sum_{n \leq N_k} \delta_{(S^p \times S^q)^n(x, x)}$ also exists and it is the relatively independent extension of the underlying graph joining between R^p and R^q . We have proved the following.

Proposition 4.13. *If different $p, q \in \mathcal{P}$ are large enough, then the set*

$$\{\rho \in J(S^p, S^q) : \rho = \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \delta_{(S^p \times S^q)^n(x, x)} \text{ for some } N_k \rightarrow \infty\}$$

*is contained in the set of relative products over the graphs of isomorphisms between R^p and R^q .*²⁶

Assume now that

$$(30) \quad F \in C(X_{\hat{\Theta}}) \text{ and } F \perp L^2(\pi^{-1}(\mathcal{B}(H_\lambda))).$$

Again, fix $x \in X_{\hat{\Theta}}$. We have

$$\frac{1}{N_k} \sum_{n \leq N_k} F(S^{pn}x) \overline{F(S^{qn}x)} \rightarrow \int_{X_{\hat{\Theta}} \times X_{\hat{\Theta}}} F \otimes \overline{F} d\rho =$$

$$\int_{H_\lambda \times H_\lambda} \mathbb{E}(F \otimes \overline{F} | H_\lambda \times H_\lambda) d\rho|_{H_\lambda \times H_\lambda} = \int_{H_\lambda} \mathbb{E}(F | H_\lambda) \cdot \overline{\mathbb{E}(F | H_\lambda) \circ W} d\rho = 0.$$

by the definition of the relative product.

The only thing which is missing now is to be sure that we have sufficiently many functions F satisfying (30). In fact, we aim at showing that $(X_{\hat{\Theta}}, S)$ has a

²⁶In fact, by Lemma 4.12, it follows that each point (x, x) is generic.

topological factor $(X_{\theta'}, S)$, which measure-theoretically is equal to (H_λ, R) and for each $F \in C(X_{\hat{\Theta}})$, we have

$$(31) \quad F = F' + F'',$$

where $F' \in C(X_{\theta'})$ (with some abuse of notation), and $F'' \perp L^2(\pi^{-1}(\mathcal{B}(H_\lambda)))$ (of course, F'' is also continuous). This will allow us to conclude the proof of Theorem 0.1 using Theorem 1.8 for F'' and dealing separately with F' .

In the bijective case, the existence of a “good” $\hat{\Theta}$ is known, see Section 5 below: in $(X_{\hat{\Theta}}, S)$ we have many “good” continuous function, as the Kronecker factor of $(X_{\hat{\Theta}}, S)$ has a “topological” realization (see [13]). However, in the general case such an approach seems to be unknown, and in Section 6, we will show a new general construction of an extension of a substitutional system in which we will see sufficiently many continuous functions satisfying (30).

5 Proof of Theorem 0.1 in the bijective and quasi-bijective case

A substitution $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ is called *quasi-bijective*²⁷ if for all $n \geq n_0$ and all $j \in \{0, \dots, \lambda^n - 1\}$, we have

$$c(\theta) = |\{\theta^n(a)_j : a \in \mathbb{A}\}|.$$

If, additionally, $c(\theta) = |\mathbb{A}|$, then we speak about a *bijective*²⁸ substitution (sometimes, such a substitution is also called *invertible*). The proof of Theorem 0.1 in the bijective case is provided in [13]. The proof is also provided in [13] for the Rudin-Shapiro substitution and the proof can be repeated for some other quasi-bijective substitutions. The proof of Theorem 0.1 covers the general case.

6 Substitutions of constant length - one more point of view

6.1 Substitution joinings of substitutional systems

Assume that we have two substitutions $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ and $\zeta : \mathbb{B} \rightarrow \mathbb{B}^\lambda$. Assume that

$$(32) \quad \mathcal{A} \subset \mathbb{A} \times \mathbb{B}, \quad p_{\mathbb{A}}(\mathcal{A}) = \mathbb{A}, \quad p_{\mathbb{B}}(\mathcal{A}) = \mathbb{B},$$

where $p_{\mathbb{A}}, p_{\mathbb{B}}$ stand for the projections on \mathbb{A} and \mathbb{B} , respectively. Moreover, we assume that

$$(33) \quad (a, b) \in \mathcal{A} \Rightarrow (\theta(a)_j, \zeta(b)_j) \in \mathcal{A}$$

for each $j = 0, \dots, \lambda - 1$. Then it is easy to see that the formula

$$(34) \quad \Sigma(a, b) = (\theta(a)_0, \zeta(a)_0)(\theta(a)_1, \zeta(a)_1) \dots (\theta(a)_{\lambda-1}, \zeta(a)_{\lambda-1})$$

²⁷The Rudin-Shapiro substitution is a prominent example of a quasi-bijective substitution which is not bijective.

²⁸Even for a bijective substitution, the height can be > 1 , see Example 3 in [33]. Also, substitution $\hat{\Theta}$ in Example 6.4 enjoys the same properties.

defines a substitution $\Sigma : \mathcal{A} \rightarrow \mathcal{A}^\lambda$ of length λ . We can also use notation:

$$\Sigma = \theta \vee \zeta.$$

Remark 6.1. We note that in general the above Σ need not be primitive, that is, it does not necessarily satisfy (15). Indeed, consider for example $\zeta = \theta$ and then take for \mathcal{A} the product set $\mathbb{A} \times \mathbb{A}$. On the other hand, the “diagonal” $\mathcal{A} := \{(a, a) : a \in \mathbb{A}\}$ yields a primitive substitution (clearly isomorphic to θ).

In order to see a less trivial primitive example, consider $\mathbb{A} = \{0, 1, 2\}$ and $\zeta = \theta$, where (see the last example in Section 2 [8])

$$0 \mapsto 010, 1 \mapsto 102, 2 \mapsto 201$$

with $\mathcal{A} = \{(0, 1), (1, 0), (0, 2), (2, 0)\}$, which gives a primitive substitution different from the “diagonal” one.²⁹ If we slightly change the definition of θ , namely:

$$0 \mapsto 010, 1 \mapsto 201, 2 \mapsto 102$$

(which is still primitive with $c = h = 2$) then Σ is primitive but (17) is not satisfied.

Remark 6.2. Note that X_Σ is a topological joining of X_θ and X_ζ . Indeed, up to a natural rearrangement of coordinates, $X_{\theta \vee \zeta} \subset X_\theta \times X_\zeta$. Then, for every $(x, y) \in X_\Sigma$, the orbits of x and y are dense in X_θ and X_ζ , respectively.³⁰ Now, the image of the natural projection $(x, y) \mapsto x$ is contained in X_θ and if $\theta(u) = u$, $\zeta(v) = v$ then $(\theta \vee \zeta)(u, v) = (u, v)$ (after a rearrangement of coordinates). So u is in the image of $p_{X_\theta}(x, y) = x$ and therefore

$$(35) \quad p_{X_\theta} : X_{\theta \vee \zeta} \rightarrow X_\theta, p_{X_\zeta}(X_{\theta \vee \zeta}) = X_\zeta.$$

Since p_{X_θ} is continuous and equivariant, it settles a topological factor map between the relevant substitutional systems.

Definition 6.1. We call Σ a *substitution joining* of θ and ζ if Σ satisfies (15) and (17) (note that (24) and (25) are satisfied automatically as θ, ζ are substitutions).

6.2 Joining with the synchronizing part

In general, when dealing with the dynamical system given by a substitution of constant length, we would like to see its Kronecker factor as a topological factor realized “in the same category of objects”, that is, realized by another substitution. This is not always possible, even in the class of bijective substitutions, see Herning’s example [21]. For the purpose of orthogonality with an arithmetic function \mathbf{u} , we need however only an extension of the original substitution which is given by another substitution (of the same length) and require that in the extended system we have a “good” realization of the Kronecker factor. This is done by a joining of θ with its synchronizing part.

²⁹It is not hard to see that in this example, θ is primitive and $c(\theta) = 2$. If by u we denote the fixed point of θ obtained by the iterations of 0, then $S_0 = 2\mathbb{N}$, so $g_0 = 2$ and hence $h(\theta) = 2 = c(\theta)$. This means that the dynamical system $(X_\theta, \mu_\theta, S)$ has discrete spectrum and hence since Σ is also primitive, as measure-theoretic dynamical systems, $(X_\theta, \mu_\theta, S)$ and $(X_\Sigma, \mu_\Sigma, S)$ are isomorphic, cf. (5).

³⁰This follows by the minimality of (X_θ, S) and (X_ζ, S) , respectively.

6.2.1 Synchronizing part of θ

Given a substitution $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$, set

$$\mathcal{X}(\theta) = \mathcal{X} := \{M \subset \mathbb{A} : M \text{ realizes the column number of } \theta\},$$

i.e. there exist $k_M \geq 1$, $j_M \in \{0, \dots, \lambda^{k_M} - 1\}$ such that

$$M = \{\theta^{k_M}(a)_{j_M} : a \in \mathbb{A}\}, |M| = c(\theta).$$

We can assume without loss of generality that $k_M = k_{M'}$ for all $M, M' \in \mathcal{X}$. Note that when $M \in \mathcal{X}$ then for each $j = 0, \dots, \lambda - 1$, we have

$$|\theta(M)_j| = c(\theta),$$

where by $\theta(M)_j$ we denote the set $\{\theta(a)_j : a \in M\}$. It follows that the formula

$$(36) \quad \tilde{\theta}(M) = \theta(M)_0 \theta(M)_1 \dots \theta(M)_{\lambda-1}$$

defines a substitution $\tilde{\theta} : \mathcal{X} \rightarrow \mathcal{X}^\lambda$.

Proposition 6.3. *Substitution $\tilde{\theta}$ has the following basic properties:*

- (i) $c(\tilde{\theta}) = 1$.
- (ii) $\tilde{\theta}$ is primitive.
- (iii) $h(\tilde{\theta}) = 1$.

Proof. For (i), let $M \in \mathcal{X}$ and suppose that $\theta^{k_M}(\mathbb{A})_{j_M} = M$. Then, for each $M' \in \mathcal{X}$, we have $\theta^{k_M}(M')_{j_M} = M$. The validity of (ii) follows by the same argument. Finally, (iii) follows from Lemma 2.2. \square

Definition 6.2. We call $\tilde{\theta}$ the *synchronizing part* of θ .

Also, note that

$$(37) \quad \bigcup_{M \in \mathcal{X}} M = \mathbb{A}.$$

Indeed, fix $M \in \mathcal{X}$ and let $a \in \mathbb{A}$. Take any $x \in M$ and (by primitivity) choose $k \geq 1$ so that $\theta^k(x)_j = a$ for some $0 \leq j < \lambda^k$. Then $a \in \tilde{\theta}^k(M)_j$ and $\tilde{\theta}^k(M)_j \in \mathcal{X}$.

Remark 6.4. If the union in (37) is additionally a partition, then we obtain an equivalence relation on \mathbb{A} which is θ -consistent (cf. (23)) and the dynamical system $(X_{\tilde{\theta}}, S)$ is a topological factor of (X_θ, S) . However, in general, there is no reason for $(X_{\tilde{\theta}}, S)$ to be a topological factor of (X_θ, S) .

Note that, in general, the union in (37) is not a partition as the following example shows.

Example 6.1. Consider the following substitution θ :

$$(38) \quad \begin{aligned} \theta(a) &= ab \\ \theta(b) &= ca \\ \theta(c) &= ba. \end{aligned}$$

By looking at θ^3 , we see that θ is primitive. Moreover, straightforward computations give that $c(\theta) = 2$ and $h(\theta) = 1$. In fact, $\mathcal{X} = \{\{a, b\}, \{a, c\}\}$. Therefore, the union in (37) is not a partition.

Note that the first column of (38) and all its iterates yields set $\{a, b, c\}$, so θ is not quasi-bijective.

Furthermore by replacing $\{a, b\}$ by 0 and $\{a, c\}$ by 1, we compute $\tilde{\theta}$:

$$\begin{aligned}\tilde{\theta}(0) &= 10 \\ \tilde{\theta}(1) &= 00.\end{aligned}$$

We also have the following result:

Proposition 6.5. *Let $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ be a substitution. Then:*

(39) *θ is bijective if and only if $\tilde{\theta}$ is trivial.*

(40) *θ is quasi-bijective if and only if $X_{\tilde{\theta}}$ is finite, i.e. a fixed point \tilde{u} of $\tilde{\theta}$ is periodic.*

Proof. First of all, (39) follows directly from (37). Let us pass to the proof of (40).

\Rightarrow : We assume that θ is quasi-bijective, i.e. for all $n \geq n_0$ and all $j < \lambda^n$, we have

$$|\theta^n(\mathbb{A})_j| = c(\theta).$$

We define $M_j := \theta^{n_0}(\mathbb{A})_j$ and it follows immediately that $\tilde{\theta}^{n_0}(M)_j = M_j$ for all $M \in \mathcal{X}$. Let \tilde{u} be a fixed point of $\tilde{\theta}$. Now, for $i \geq 0$ and $0 \leq j < \lambda^{n_0}$, we find that (use (11) letting $\ell \rightarrow \infty$),

$$\tilde{u}[j + i\lambda^{n_0}] = \tilde{\theta}^{n_0}(\tilde{u})[j + i\lambda^{n_0}] = \tilde{\theta}^{n_0}(\tilde{u}[i])_j = M_j,$$

which shows that \tilde{u} is periodic (in fact, λ^{n_0} is a period of \tilde{u}).

\Leftarrow : Let \tilde{u} be a periodic fixed point of $\tilde{\theta}$ with period p . First, we claim that \tilde{u} is also periodic with period λ^{n_0} for some $n_0 \geq 0$. Indeed, by Proposition 3.1 (and Proposition 6.3), it follows that, in the Weyl pseudo-metric d_W , we can approximate \tilde{u} by periodic sequences that have period λ^n . Thus, there exists a λ^{n_0} -periodic sequence v such that

$$d_W(\tilde{u}, v) < \frac{1}{2p}.$$

By basic properties of d_W , we obtain

$$d_W(S^{\lambda^{n_0}}(\tilde{u}), S^{\lambda^{n_0}}(v)) \leq d_W(\tilde{u}, v) < \frac{1}{2p}.$$

Hence, the triangle inequality implies

$$d_W(\tilde{u}, S^{\lambda^{n_0}}(\tilde{u})) < \frac{1}{p}.$$

Since \tilde{u} and $S^{\lambda^{n_0}}(\tilde{u})$ are both p -periodic functions with distance less than $1/p$, they must coincide and the claim follows.

As, by Proposition 6.3, $\tilde{\theta}$ is primitive, there exists $n \geq 0$ such that for each $M \in \mathcal{X}$, we can find $j_1 < \lambda^n$ such that $\tilde{\theta}^n(\tilde{u}[0])_{j_1} = M$. Moreover, for all $j_0 < \lambda^{n_0}$ and all $j < \lambda^n$ (use (11) and $j_0 + j\lambda^{n_0} < \lambda^{n+n_0}$), we have

$$\begin{aligned}\tilde{u}[j_0] &= \tilde{u}[j_0 + j\lambda^{n_0}] = \tilde{\theta}^{n+n_0}(\tilde{u})[j_0 + j\lambda^{n_0}] = \\ &= \tilde{\theta}^{n+n_0}(\tilde{u}[0])[j_0 + j\lambda^{n_0}] = \tilde{\theta}^{n_0}(\tilde{\theta}^n(\tilde{u}[0]))_{j_0}.\end{aligned}$$

Letting $j = j_1$, this shows that $\tilde{\theta}^{n_0}(M)_{j_0} = \tilde{u}[j_0]$ for all $M \in \mathcal{X}, j_0 < \lambda^{n_0}$. As \mathcal{X} covers \mathbb{A} in view of (37), this shows that for all $j_0 < \lambda^{n_0}$, we have $\tilde{\theta}^{n_0}(\mathbb{A})_{j_0} = \tilde{u}[j_0]$ and, therefore,

$$|\tilde{\theta}^{n_0}(\mathbb{A})_{j_0}| = c(\tilde{\theta}),$$

which concludes the proof. \square

Remark 6.6. It follows from Propositions 6.3 and 6.5 that once θ is **not** quasi-bijective, $(X_{\tilde{\theta}}, S)$ is a “realization” of (H_λ, R) .

6.2.2 Joining θ with its synchronizing part

Define

$$\overline{\mathcal{X}} := \{(a, M) : a \in M, M \in \mathcal{X}\} \subset \mathbb{A} \times \mathcal{X}.$$

Now, in view of (37), the first compatibility condition (32) immediately follows. If $(a, M) \in \overline{\mathcal{X}}$ then $a \in M$, so for each $j = 0, \dots, \lambda - 1$, we have $\theta(a)_j \in \theta(M)_j$, and in view of (36), $\theta(a)_j \in \tilde{\theta}(M)_j$. Therefore, the second compatibility condition (33) also follows.

Set

$$\begin{aligned}\Theta &:= \theta \vee \tilde{\theta} : \overline{\mathcal{X}} \rightarrow \overline{\mathcal{X}}^\lambda, \\ \Theta(a, M) &= (\theta(a)_0, \tilde{\theta}(M)_0)(\theta(a)_1, \tilde{\theta}(M)_1) \dots (\theta(a)_{\lambda-1}, \tilde{\theta}(M)_{\lambda-1}).\end{aligned}$$

Remark 6.7. Although, by primitivity, we can assume that there are $a \in \mathbb{A}$ and $M \in \mathcal{X}$ such that $\theta(a)_0 = a$ and $\tilde{\theta}(M)_0 = M$, it may happen that (17) is not satisfied for Θ . However, as it does for any substitution of constant length, there exists some $k \in \mathbb{N}$ such that $\Theta^k = \theta^k \vee \tilde{\theta}^k$ ³¹ satisfies (17). Since the following proposition shows that Θ satisfies (15), which assures us that $X_\Theta = X_{\Theta^k}$, we can assume without loss of generality that Θ satisfies (17).

Proposition 6.8. *Substitution Θ is primitive and $h(\Theta) = h(\theta)$.*

Proof. We first show that Θ is primitive. Let $(a_1, M_1), (a_2, M_2) \in \overline{\mathcal{X}}$. By the definition of $\tilde{\theta}$, it follows that there are some $k \in \mathbb{N}, j < \lambda^k$ such that $\theta^k(\mathbb{A})_j = M_2$. Hence, $\tilde{\theta}^k(M)_j = M_2$ for every $M \in \mathcal{X}$. This ensures that there exists $a \in \mathbb{A}$ such that $\theta^k(a)_j = a_2$. Using the primitivity of θ , choose $k' \in \mathbb{N}$ and $j' < \lambda^{k'}$ so that $\theta^{k'}(a_1)_{j'} = a$. We define $M := \tilde{\theta}^{k'}(M_1)_{j'}$ and using (11), we obtain

$$\begin{aligned}\Theta^{k+k'}((a_1, M_1))_{j'\lambda^k+j} &= \Theta^k(\Theta^{k'}((a_1, M_1))_{j'})_j \\ &= \Theta^k((a, M))_j = (a_2, M_2)\end{aligned}$$

³¹ We always have $\Theta^k = (\theta \vee \tilde{\theta})^k = \theta^k \vee \tilde{\theta}^k = \theta^k \vee \tilde{\theta}^k$.

which shows that Θ is primitive.

Let us denote by u a fixed point of θ , by v a fixed point of Θ , where $u[0] = a_0, v[0] = (a_0, M_0)$. According to (18), if we set

$$S_k(\theta) = \{r \geq 1 : u[k+r] = u[k]\},$$

then

$$h(\theta) = \max\{m \geq 1 : (m, \lambda) = 1, m \mid \gcd S_k\},$$

for all $k \geq 0$.

We recall also that in the first part of the proof we showed that for $(a_1, M_1) = (a_2, M_2) = (a_0, M_0)$ (and using (11)), we obtain that

$$\Theta^{k+k'}((a_0, M_0))_{j'\lambda^k+j} = (a_0, M_0)$$

for any $k' \in \mathbb{N}, j' < \lambda^{k'}$ such that $\theta^{k'}(a_0)_{j'} = a$. Since, $u[0, \lambda^{k'} - 1] = \theta^{k'}(a_0)$, the above means that there exists $k \in \mathbb{N}, j < \lambda^k$ such that for all j' with $u[j'] = a$, we have $v[j'\lambda^k + j] = (a_0, M_0)$. We denote by j'_0 the smallest such j' . Thus, we find

$$\lambda^k \cdot S_{j'_0}(\theta) \subset S_{j'_0\lambda^k+j}(\Theta).$$

This implies that $\gcd S_{j'_0\lambda^k+j}(\Theta) \mid \lambda^k \gcd S_{j'_0}(\theta)$, and, therefore, $h(\Theta) \mid h(\theta)$. Moreover, by Lemma 6.18 (below), we obtain that $h(\theta) \mid h(\Theta)$, which completes the proof.³² \square

Remark 6.9. Note that from the measure-theoretic point of view, the substitutional system $(X_{\theta \vee \tilde{\theta}}, \mu_{\theta \vee \tilde{\theta}}, S)$ (which is an extension of $(X_\theta, \mu_\theta, S)$) is isomorphic to $(X_\theta, \mu_\theta, S)$. Indeed, by the unique ergodicity of $(X_{\theta \vee \tilde{\theta}}, S)$, it represents an ergodic joining of $(X_\theta, \mu_\theta, S)$ and $(X_{\tilde{\theta}}, \mu_{\tilde{\theta}}, S)$ and the latter system has discrete spectrum contained in the discrete spectrum of $(X_\theta, \mu_\theta, S)$. By ergodicity, the only such a joining is graphic. We will detail on this soon.

Hence, using Proposition 6.8, the above remark and the ergodic interpretation of the column number, cf. Proposition 2.4, we obtain the following (which is also an immediate consequence of Lemma 6.18³³):

Corollary 6.10. For a substitution $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$, we have $c(\Theta) = c(\theta)$.

Remark 6.11. Note also that $(X_{\tilde{\theta}}, S)$ is a topological factor of $(X_{\theta \vee \tilde{\theta}}, S)$ via the map:

$$(x_n, M_n)_{n \in \mathbb{Z}} \mapsto (M_n)_{n \in \mathbb{Z}}.$$

Assume that $M \in \mathcal{X}$. As, for each $j = 0, \dots, \lambda - 1$, the set $\{\theta(a)_j : a \in M\}$ has $c(\theta)$ elements, we obtain the following:

Lemma 6.12. *Substitution $\theta \vee \tilde{\theta}$ has the following property: for each $M \in \mathcal{X}$, we have $(\theta \vee \tilde{\theta})(\cdot, M)_j$ is bijective for each $j = 0, \dots, \lambda - 1$.*³⁴

Remark 6.13. In this way, using joinings, we explained one of the strategies in [35] which consists in representation of each automatic sequence as a “combination” of the synchronized and relatively invertible parts.

³²Alternatively, $h(\theta) \mid h(\Theta)$ follows from the fact that (X_θ, S) is a topological factor of (X_Θ, S) .

³³Indeed, as $c(\tilde{\theta}) = 1$, Lemma 6.18 shows that $c(\theta) \leq c(\Theta) \leq c(\theta)$.

³⁴More precisely, it is bijective on a set whose cardinality is $c(\theta)$, i.e., a bijection from M to $\tilde{\theta}(M)_j$.

6.2.3 Description of the subshift $(X_{\theta \vee \tilde{\theta}}, S)$

Each point of the space $X_{\theta \vee \tilde{\theta}}$ is of the form $(x_n, M_n)_{n \in \mathbb{Z}}$, where $\overline{M} := (M_n) \in X_{\tilde{\theta}}$ with $M_n \in \mathcal{X}$, and $x = (x_n) \in X_{\theta}$, so for each $n \in \mathbb{Z}$, $x_n \in M_n \subset \mathbb{A}$. More than that, like every substitutional system, such a point must have its (unique) λ^t -skeleton structure $(j_t)_{t \geq 1}$ on which we will detail more. First of all, by the definition of $\theta \vee \tilde{\theta}$, (j_t) must be the λ^t -skeleton structure of both \overline{M} and x (we strongly use here that the skeleton structure is unique; for that, we need that the substitution $\tilde{\theta}$ is non-trivial, in other words the argument makes no sense if θ is quasi-bijective).

So $(j_t)_{t \geq 1}$ is the λ^t -skeleton structure of $\overline{M} \in X_{\tilde{\theta}}$: for each $t \geq 1$, we have

$$\overline{M}[-j_t + s\lambda^t, -j_t + (s+1)\lambda^t - 1] = \tilde{\theta}^t(R_s)$$

for each $s \in \mathbb{Z}$ and some $R_s \in \mathcal{X}$. Fix $t \geq 1$ and consider $s = 0$. We have

$$\overline{M}[-j_t, -j_t + \lambda^t - 1] = \tilde{\theta}^t(R_0)$$

and $\mathcal{X} \ni R_0 = \{r_0, \dots, r_{c-1}\}$, where $c = c(\theta)$ and $r_j \in \mathbb{A}$. Hence

$$(41) \quad \tilde{\theta}^t(R_0) = \begin{array}{cccc} \theta^t(r_0)_0 & \dots & \theta^t(r_0)_{j_t} & \dots & \theta^t(r_0)_{\lambda^t-1} \\ \theta^t(r_1)_0 & \dots & \theta^t(r_1)_{j_t} & \dots & \theta^t(r_1)_{\lambda^t-1} \\ \dots & \dots & \dots & \dots & \dots \\ \theta^t(r_{c-1})_0 & \dots & \theta^t(r_{c-1})_{j_t} & \dots & \theta^t(r_{c-1})_{\lambda^t-1}, \end{array}$$

where the columns “represent” sets in \mathcal{X} . We look at $\tilde{\theta}^t(R_0)_{j_t}$ the j_t -th element of $\tilde{\theta}^t(R_0)$ which is “represented” by the j_t -th column of the matrix above and is equal to M_0 (the zero coordinate of \overline{M}). Now, $x_0 \in M_0$, hence $x_0 = \theta^t(r_\ell)_{j_t}$ for a unique $0 \leq \ell \leq c-1$. Moreover, $x[-j_t, -j_t + \lambda^t - 1] = \theta^t(r_\ell)$. Hence, unless $j_t = \lambda^t - 1$, x_0 determines x_1 .

Note that, we can reverse this reasoning in the case:

$$(42) \quad \min(j_t, \lambda^t - j_t - 1) \rightarrow \infty.$$

Indeed, in this case, given \overline{M} , we can choose $x_0 \in M_0$ arbitrarily, and then successively fill in the second coordinate by placing there the ℓ -th row in the matrix (41), where $x_0 = \theta^t(r_\ell)_{j_t}$. This shows that over all $\overline{M} \in X_{\tilde{\theta}}$ we have c points $(x, \overline{M}) \in X_{\theta \vee \tilde{\theta}}$ whenever the λ^t -skeleton of \overline{M} satisfies (42). What remains are points for which

$$\text{either } j_{t+1} = j_t \text{ or } \lambda^{t+1} - j_{t+1} - 1 = \lambda_t - j_t - 1 \text{ for } t \geq T.$$

When projected down to the maximal equicontinuous factor, this condition defines a countable set, in particular, of (Haar) measure zero. For the unique measure $\mu_{\tilde{\theta}}$ we have hence a.e. a c -point extension. For \overline{M} which do not satisfy (42), we need (at most) two coordinates for x to determine all other, so the fibers have at most c^2 elements.

Example 6.2. Let $\mathbb{A} = \{a, b, c, d\}$. We consider the following substitution $\theta : \mathbb{A} \rightarrow \mathbb{A}^4$:

$$\begin{aligned} \theta(a) &= adda, \quad \theta(b) = bccb \\ \theta(c) &= abbc, \quad \theta(d) = baad. \end{aligned}$$

It is not hard to see that θ is pure and $c = 2$, $\mathcal{X} = \{M_0, M_1\}$, where $M_0 = \{a, b\}$ and $M_1 = \{c, d\}$. Moreover,

$$\begin{aligned}\tilde{\theta}(M_0) &= M_0 M_1 M_1 M_0 \\ \tilde{\theta}(M_1) &= M_0 M_0 M_0 M_1.\end{aligned}$$

Let us now consider the block $\overline{M}^{(0)} = M_0 M_0 \in \mathcal{X}^2$, which we write as $\overline{M}_{-1}^{(1)} \overline{M}_0^{(1)}$. Acting on it by $\tilde{\theta}$, we obtain

$$\overline{M}^{(1)} = \tilde{\theta}(M_0) \tilde{\theta}(M_0) = \overline{M}_{-4}^{(1)} \dots \overline{M}_{-1}^{(1)} \overline{M}_0^{(1)} \dots \overline{M}_3^{(1)}.$$

By iterating this procedure and passing to the limit, we obtain a two-sided sequence which is a fixed point for θ and which we denote by $\overline{M} = \tilde{\theta}^\infty(M_0) \cdot \tilde{\theta}^\infty(M_0)$, the “dot” indicating the zero position of the sequence. A similar procedure can be made on the θ -side, starting with $a.a$, $a.b$, $b.a$ and $b.b$. It is not hard to see that, up to a natural rearrangement of coordinates, the following four points $(\theta^\infty(a) \cdot \theta^\infty(a), \overline{M})$, $(\theta^\infty(a) \cdot \theta^\infty(b), \overline{M})$, $(\theta^\infty(b) \cdot \theta^\infty(a), \overline{M})$, $(\theta^\infty(b) \cdot \theta^\infty(b), \overline{M})$ are members of $X_{\theta \vee \tilde{\theta}}$. It follows that the fiber over \overline{M} has four points.

6.3 Toward a skew-product measure-theoretic representation - making the relative invertibility clearer

In this part we want to rename the alphabet of Θ to get a new substitution, which makes the invertible part easier to handle. Namely, our new alphabet will be the set $\{0, \dots, c-1\} \times \mathcal{X}$, where $c = c(\theta)$. The only thing we need is to give a “good” identification of $\{0, \dots, c-1\} \times M$ with $\{(a, M) : a \in M\}$ (for $M \in \mathcal{X}$). We start by giving a classical example.

Example 6.3 (Rudin-Shapiro sequence). We consider the Rudin-Shapiro substitution θ defined by $\mathbb{A} = \{a, b, c, d\}$ and

$$\begin{aligned}a &\mapsto ab, & d &\mapsto dc \\ b &\mapsto ac, & c &\mapsto db.\end{aligned}$$

We find that $\mathcal{X} = \{\{a, d\}, \{b, c\}\}$ and consequently for Θ , we have:

$$\begin{aligned}(a, \{a, d\}) &\mapsto (a, \{a, d\})(b, \{b, c\}), & (d, \{a, d\}) &\mapsto (d, \{a, d\})(c, \{b, c\}) \\ (b, \{b, c\}) &\mapsto (a, \{a, d\})(c, \{b, c\}), & (c, \{b, c\}) &\mapsto (d, \{a, d\})(b, \{b, c\}).\end{aligned}$$

Let us now fix a partial ordering on \mathbb{A} that is complete on every $M \in \mathcal{X}$, e.g., $a < b, c < d$. Thus, we can identify $(a, \{a, d\}) \cong (0, \{a, d\})$ and $(d, \{a, d\}) \cong (1, \{a, d\})$ as a is the smallest element of $\{a, d\}$ and d is the second smallest element. This gives the new substitution $\tilde{\Theta} : \{0, \dots, c-1\} \times \mathcal{X} \rightarrow (\{0, \dots, c-1\} \times \mathcal{X})^\lambda$: which in our example is given by:

$$\begin{aligned}(0, \{a, d\}) &\mapsto (0, \{a, d\})(0, \{b, c\}), & (1, \{a, d\}) &\mapsto (1, \{a, d\})(1, \{b, c\}) \\ (0, \{b, c\}) &\mapsto (0, \{a, d\})(1, \{b, c\}), & (1, \{b, c\}) &\mapsto (1, \{a, d\})(0, \{b, c\}).\end{aligned}$$

Note that $\tilde{\theta}$ is very simple in this example, as identifying $\{a, d\}$ with M_0 and $\{b, c\}$ with M_1 , gives $\tilde{\theta}(M_0) = M_0 M_1$ and $\tilde{\theta}(M_1) = M_0 M_1$, so we obtain

a periodic sequence. We obtain the classical Rudin-Shapiro sequence by considering the fixpoint of $\tilde{\Theta}$ and then applying the projection $(i, M) \mapsto i$. (We emphasize that the classical Rudin-Shapiro sequence itself is not a fixpoint of a substitution.)

We will be dealing with θ which is not quasi-bijective (which, via Propositions 6.3 and 6.5, guarantees that $(X_{\tilde{\theta}}, S)$ “captures” the whole discrete spectrum of $(H_\lambda, m_{H_\lambda}, R)$). We repeat the above construction: fix a partial order on \mathbb{A} that is complete on every $M \in \mathcal{X}$ and identify (a, M) with (j, M) if the ordering of the elements in M according to the ordering on \mathbb{A} restricted to M yields a as the $(j + 1)$ -th smallest element and rewrite instructions using the alphabet. This yields the substitution

$$\tilde{\Theta} : \{0, \dots, c - 1\} \times \mathcal{X} \rightarrow (\{0, \dots, c - 1\} \times \mathcal{X})^\lambda.$$

The second coordinate is still $\tilde{\theta}$ which corresponds to the synchronizing part and can be defined independently of the first coordinate. The first coordinate (which now depends on the second coordinate), gives the “invertible part”.

Lemma 6.14. $\tilde{\Theta}$ is primitive, $h(\tilde{\Theta}) = h(\theta)$ and $c(\tilde{\Theta}) = c(\theta)$.

Proof. As $\tilde{\Theta}$ is obtained by renaming the alphabet of Θ , this follows immediately from Proposition 6.8 and Corollary 6.10. \square

Remark 6.15. A point in the space $X_{\tilde{\Theta}}$ is of the form $(y_n, M_n)_{n \in \mathbb{Z}}$ with $y_n \in \{0, \dots, c - 1\}$ and there is an equivariant map between $X_{\tilde{\Theta}}$ and $X_{\theta \vee \tilde{\theta}}$ given by

$$(43) \quad (y_n, M_n) \mapsto (z_n, M_n),$$

where $z_n \in M_n$ is the $(y_n + 1)$ -th letter in the ordering on M_n . Composing this map with the projection on the first coordinate yields the topological factor (X_θ, S) .

The projection of (43) on the second coordinate yields the factor $(X_{\tilde{\theta}}, S)$. According to Subsection 6.2.3, we can now represent $(X_{\tilde{\Theta}}, \mu_{\tilde{\Theta}}, S)$ as a skew product over $(X_{\tilde{\theta}}, \mu_{\tilde{\theta}}, S)$ using the following: first

$$\text{identifying } (i_n, M_n)_{n \in \mathbb{Z}} \text{ with } ((M_n)_{n \in \mathbb{Z}}, i_0) \in X_{\tilde{\theta}} \times \{0, \dots, c - 1\}$$

(which is possible for all $\overline{M} = (M_n)_{n \in \mathbb{Z}}$ satisfying (42), cf. Subsection 6.2.3) and then setting

$$(44) \quad (\overline{M}, i_0) \mapsto (S\overline{M}, \sigma_{\overline{M}}(i_0)),$$

where $\sigma_{\overline{M}}$ (in fact, it is σ_{M_0}) is a permutation of $\{0, \dots, c - 1\}$ determined by M_0 : we come back to the matrix (41) in which M_0 is represented as the j_t -th column. The next column is representing M_1 , and if we have z_0 (which belongs to M_0 and corresponds to i_0) in the i_0 -th row then z_1 is just the next element in the same row. Now, the elements in the j_t -th column have their names in $\{0, \dots, c - 1\}$ and this is the same for $(j_t + 1)$ -st column. Hence $\sigma_{\overline{M}}$ is a permutation of $\{0, \dots, c - 1\}$ sending an element in the j_t -th column to the neighboring one in the next column.

Remark 6.16. As from the measure-theoretic point of view, $(X_{\tilde{\Theta}}, S)$ is still (X_θ, S) (and $(X_{\tilde{\theta}}, S)$ is the same as (H_λ, R)), the above formula (44) clears up Remark 9.1 in [36], p. 229, about the form of a cocycle representing $(X_\theta, \mu_\theta, S)$ as a skew product over $(H_\lambda, m_{H_\lambda}, R)$.

Coming back to our main problem, note that $(X_{\tilde{\Theta}}, S)$ has (X_θ, S) as its **topological** factor and it has also $(X_{\tilde{\theta}}, S)$ “representing” measure-theoretically (H_λ, R) as its **topological** factor, but **still** we do not know whether we have decomposition (31) for each $F \in C(X_{\tilde{\Theta}})$. To assure it, we will need another extension.

At the end of this section we want to discuss one particular ordering that has useful properties.

We start by taking an arbitrary complete order $<_{M_0}$ on M_0 such that a_0 is the minimum. By Proposition 6.3, $\tilde{\theta}$ is synchronizing and primitive, thus we find k'_0, j'_0 such that $\tilde{\theta}^{k'_0}(M)_{j'_0} = M_0$ for all $M \in \mathcal{X}$. As $\theta^{k'_0}_{j'_0}$ permutes the elements of M_0 , we find by iterating $\theta^{k'_0}_{j'_0}$, some $k_0, j_0 < \lambda^{k_0}$ such that $\theta^{k_0}(a)_{j_0} = a$ for all $a \in M_0$ and $\tilde{\theta}^{k_0}(M)_{j_0} = M_0$ for all $M \in \mathcal{X}$.

Using (37), we extend now the ordering on M_0 to a partial (but not complete in general, as below, $a \neq b$ remain incomparable if $\theta^{k_0}(a)_{j_0} = \theta^{k_0}(b)_{j_0}$) ordering on \mathbb{A} by

$$a < b \text{ if and only if } \theta^{k_0}(a)_{j_0} <_{M_0} \theta^{k_0}(b)_{j_0}.$$

This gives a complete ordering on any $M \in \mathcal{X}$ and, thus, allows us to define $\tilde{\Theta}$. With this ordering we obtain directly that for all $M \in \mathcal{X}$ and $i = 0, 1, \dots, c-1$, we have

$$(45) \quad \tilde{\Theta}^{k_0}(i, M)_{j_0} = (i, M_0)$$

which will be useful later.

Remark 6.17. As the order on each $M \in \mathcal{X}$ is fixed, we have $\tilde{\Theta}^k = \tilde{\Theta}^k$ (cf. footnote 31).

6.4 Column number, height and pure base – revisited

In this part we come back to connections between the column number and the height of primitive substitutions. We start with simple observations about substitution joinings.

Lemma 6.18. *Let $\theta : \mathbb{A} \rightarrow \mathbb{A}^\lambda$ and $\zeta : \mathbb{B} \rightarrow \mathbb{B}^\lambda$ be substitutions of length λ fulfilling (15), (17), (24) and (25), but are not necessarily pure. For some $\mathcal{A} \subset \mathbb{A} \times \mathbb{B}$, let $\Sigma = \theta \vee \zeta : \mathcal{A} \rightarrow \mathcal{A}^\lambda$ be a substitutional joining (in particular, we assume that Σ is primitive). Then,*

$$(46) \quad \text{lcm}(h(\theta), h(\zeta)) \mid h(\Sigma),$$

which also yields a lower bound for $h(\Sigma)$. Furthermore,

$$\max(c(\theta), c(\zeta)) \leq c(\Sigma) \leq c(\theta) \cdot c(\zeta),$$

which, in particular, yields lower and upper bounds for the column number of Σ .

Proof. One sees directly that $S_k(\Sigma) = S_k(\theta) \cap S_k(\zeta)$ and, therefore, $g_k(\theta) \mid g_k(\Sigma)$ and $g_k(\zeta) \mid g_k(\Sigma)$ which implies (46).³⁵

As the projections of \mathcal{A} on \mathbb{A} and \mathbb{B} are full, the lower bound for $c(\Sigma)$ is obvious. Furthermore, if $|\theta^k(\mathbb{A})_j| = c(\theta)$, $|\zeta^{k'}(\mathbb{B})_{j'}| = c(\zeta)$ then (using (11)), we have

$$\begin{aligned} \Sigma^{k'+k}(\mathcal{A})_{j\lambda^{k'}+j'} &\subseteq \Sigma^{k'+k}(\mathbb{A} \times \mathbb{B})_{j\lambda^{k'}+j'} = \Sigma^{k'}(\Sigma^k(\mathbb{A} \times \mathbb{B}))_{j'} \\ &\subseteq \Sigma^{k'}(\theta^k(\mathbb{A})_j \times \mathbb{B})_{j'} \subseteq \theta^{k'}(\theta^k(\mathbb{A}))_{j'} \times \zeta^{k'}(\mathbb{B})_{j'}. \end{aligned}$$

This shows directly that $|\Sigma^{k'+k}(\mathcal{A})_{j\lambda^{k'}+j'}| \leq c(\theta) \cdot c(\zeta)$ and completes the proof. \square

Let us now pass to the proof of Lemma 2.3:

Proof of Lemma 2.3 (We assume that the ordering on \mathbb{A} is as the one described at the end of Subsection 6.3.) We have already seen that $c := c(\theta) = c(\tilde{\Theta})$ and $h := h(\theta) = h(\tilde{\Theta})$, so we consider without loss of generality the substitution $\tilde{\Theta} : (\{0, \dots, c-1\} \times \mathcal{X}) \rightarrow (\{0, \dots, c-1\} \times \mathcal{X})^\lambda$. It follows directly from the basic properties of the height that we discussed at the beginning of the paper that there exists a 1-coding $f : \mathcal{A} := (\{0, \dots, c-1\} \times \mathcal{X}) \rightarrow \{0, \dots, h-1\}$ such that the fixed point $u = \tilde{\Theta}(u)$ is mapped to the periodic sequence

$$01 \dots (h-1)01 \dots (h-1) \dots$$

It follows by a simple computation (cf. (11)) that

$$u[j'\lambda^k + j] = \tilde{\Theta}^k(u[j'])_j$$

whenever $j' \geq 0$ and $j < \lambda^k$. As $f(u[\ell]) \equiv \ell \pmod{h}$ for each $\ell \geq 0$, it follows that (letting j' be so that $(i, M) = u[j']$ and $\ell = j'\lambda^k + j$)

$$f(\tilde{\Theta}^k(i, M)_j) = f(\tilde{\Theta}^k(u[j'])_j) = f(u[j'\lambda^k + j]) \equiv j'\lambda^k + j \pmod{h}.$$

By the same token, $f(i, M) \equiv j' \pmod{h}$, so finally

$$(47) \quad f(\tilde{\Theta}^k(i, M)_j) \equiv f(i, M)\lambda^k + j \pmod{h}$$

for each $k \geq 0$ and $j < \lambda^k$.

Furthermore, we know by the construction of $\tilde{\Theta}$ that there exist k_0, j_0 such that $\tilde{\Theta}^{k_0}(i, M)_{j_0} = (i, M_0)$ for every $i \in \{0, \dots, c-1\}$ and $M \in \mathcal{X}$, as noticed in (45). Thus, in view of (47) (for $j = j_0$ and $k = k_0$), for all $i \in \{0, \dots, c-1\}$ and $M \in \mathcal{X}$, we obtain that

$$(f(i, M)\lambda^{k_0} + j_0 \pmod{h}) = f(\tilde{\Theta}^{k_0}(i, M)_{j_0}) = f(i, M_0).$$

This implies that $f(i, M)$ does not depend on M (as λ is a multiplicatively invertible element in the ring $\mathbb{Z}/h\mathbb{Z}$), so $f(i, M) = f(i, M_0)$ and f only depends on the first coordinate. Therefore, we denote $f'(i) := f(i, M_0)$.

³⁵One can also see this result dynamically, as both (X_θ, S) and (X_ζ, S) are topological factors of (X_Σ, S) .

As the second coordinate in $\tilde{\Theta}$ equals $\tilde{\theta}$ is independent of the first coordinate and $\tilde{\Theta}(\cdot, M)_j$ is a bijection from $\{0, \dots, c-1\}$ to itself, we obtain the following formula for the incidence matrix $M(\tilde{\Theta})$,

$$(48) \quad \sum_{i' < c} M(\tilde{\Theta})_{(i,M),(i',M')} = M(\tilde{\theta})_{M,M'}.$$

We now claim that

$$(49) \quad \delta_{(i,M)} = \delta_M \cdot \frac{1}{c},$$

i.e. the density of (i, M) for $\tilde{\Theta}$ is equal to $e_{i,M} := \frac{1}{c}\delta_M$. Indeed, using the definition of $e_{i,M}$, (48) and (16) (for $\tilde{\theta}$), we obtain

$$\begin{aligned} \sum_{(i',M') \in \{0, \dots, c-1\} \times \mathcal{X}} e_{(i',M')} \cdot M(\tilde{\Theta})_{(i,M),(i',M')} &= \sum_{M' \in \mathcal{X}} \frac{1}{c} \delta_{M'} \cdot \sum_{i' < c} M(\tilde{\Theta})_{(i,M),(i',M')} \\ &= \sum_{M' \in \mathcal{X}} \frac{1}{c} \delta_{M'} \cdot M(\tilde{\theta})_{M,M'} \\ &= \lambda \cdot \frac{1}{c} \delta_M = \lambda \cdot e_{i,M}. \end{aligned}$$

But (by the Perron-Frobenius theorem) there is a unique solution to the above, given by $\delta_{i,M}$, so $e_{i,M} = \delta_{i,M}$ and the claim (49) follows.

Clearly, each letter $j \in \{0, \dots, h-1\}$ appears in $01 \dots (h-1)01 \dots$ with density $\frac{1}{h}$ and, recall that, this periodic sequence is the image of u under f . Using this and (49), we obtain

$$(50) \quad 1/h = \sum_{(i,M): f(i,M)=j} \delta_{(i,M)} = \sum_{i: f'(i)=j} \sum_M \frac{1}{c} \delta_M = \frac{1}{c} |\{i < c : f'(i) = j\}|.$$

Equation (50) shows directly that $h|c$ and $\#\{i < c : f'(i) = j\}$ is constant and equals c/h . It remains to show that the column number of the pure base of θ equals c/h . For this aim, let us first notice that we can also view f as well defined on \mathbb{A} and that for all $M \in \mathcal{X}$ and $j < h$, we have

$$(51) \quad \#\{a \in M : f(a) = j\} = c/h.$$

Recall also the construction of the pure basis $\theta^{(h)}$ of θ . It is a substitution defined over the alphabet $\mathbb{A}^{(h)}$ which is the set $\{u[mh(\theta), mh(\theta) + h - 1] : m \geq 0\}$ (which is of course finite). Then, we define $\theta^{(h)} : \mathbb{A}^{(h)} \rightarrow (\mathbb{A}^{(h)})^\lambda$ by setting

$$\theta^{(h)}(w)_j = \theta(w)[jh(\theta), (j+1)h(\theta) - 1]$$

for $j = 0, \dots, \lambda - 1$.

If $w = (w_0, \dots, w_{h-1}) \in \mathbb{A}^{(h)}$, then there exists some m such that $a_i = u[mh + i]$ for $i = 0, \dots, h-1$. It follows that

$$(52) \quad f(w_i) = f(u[mh + i]) \equiv mh + i \equiv i \pmod{h}$$

which gives $f(w_i) = i$. For each $k \geq 1$ and $j < \lambda^k$, we have

$$(53) \quad (\theta^{(h)})^k(w)_j = \theta^k(w)[jh(\theta), (j+1)h(\theta) - 1].$$

Choose $k > 1$ such that $\lambda^k \equiv 1 \pmod{h}$ and $\lambda^k > (j_0 + 1)h$. Now, $j_0 \frac{\lambda^k - 1}{h} + j_0 < \lambda^{k+k_0}$. In view of (53) (for $k + k_0$ and $j_0(\lambda^k - 1)/h + j_0$) and then (11) (which we may use as $(j_0 + 1)(h - 1) < \lambda^k$), we obtain that

$$\begin{aligned} & (\theta^{(h)})^{k+k_0}(w_0, \dots, w_{h-1})_{j_0(\lambda^k-1)/h+j_0} \\ &= \theta^{k+k_0}(w_0, \dots, w_{h-1})[j_0\lambda^k + j_0(h-1), j_0\lambda^k + (j_0 + 1)(h-1)] \\ &= \theta^k(\theta^{k_0}(w_0)_{j_0})[j_0(h-1), (j_0 + 1)(h-1)]. \end{aligned}$$

It follows from this calculation that the column number of $\theta^{(h)}$ is not larger than the number of $\theta^{k_0}(a_0)_{j_0}$, when we run over all possible $(w_0, \dots, w_{h-1}) \in \mathbb{A}^{(h)}$. But, as before (cf. (47)), $f(\theta^k(a)_j) = f(a)\lambda^k + j \pmod{h}$ (for each $a \in \mathbb{A}$). It follows that $f(\theta^{k_0}(w_0)_{j_0}) \equiv j_0 \pmod{h}$ (as $f(w_0) = 0$, see (52)). Moreover, by (45), $\theta^{k_0}(a)_{j_0} \in M_0$ (for each $a \in \mathbb{A}$). This shows that $c(\theta^{(h)})$ is not larger than the number of $b \in M_0$ such that $f(b) \equiv j_0 \pmod{h}$ and hence, in view of (51), $c(\theta^{(h)}) \leq c(\theta)/h(\theta)$.

Next, we want to give a lower bound for $c(\theta^{(h)})$. We fix any $k > 0$ and $j < \lambda^k$. Denote $i = \lfloor \frac{jh}{\lambda^k} \rfloor$. In view of (22), it follows that we have for any $(w_0, \dots, w_{h-1}) \in \mathbb{A}^{(h)}$,

$$(\theta^{(h)})^k(w_0, \dots, w_{h-1})_j = (v_0, \dots, v_{h-1}),$$

where

$$v_0 = \theta^k(w_i)_{jh-i\lambda^k}.$$

Moreover, $f(w_i) = i$ (see (52)). Thus, we find

$$\#(\theta^{(h)})^k(\mathbb{A}^{(h)})_j \geq \#\theta^k(\mathbb{A}_i)_{jh-i\lambda^k},$$

where $\mathbb{A}_i := \{a \in \mathbb{A} : f(a) = i\}$. Finally, we recall the relative bijectivity: for each $M \in \mathcal{X}$ and any different $a, a' \in M$, we have $\theta^k(a)_{j'} \neq \theta^k(a')_{j'}$ for each $j' < \lambda^k$, see the last phrase before Lemma 6.12. This shows that

$$c(\theta^{(h)}) \geq \#\{a \in M_0 : f(a) = i\} = \frac{c(\theta)}{h(\theta)},$$

which completes the proof. \square

Corollary 6.19. The dynamical system $(X_\theta, \mu_\theta, S)$ given by (primitive) θ has purely discrete spectrum if and only if $h(\theta) = c(\theta)$.

6.5 Toward a group extension

We will now aim at an extension of the substitution $\tilde{\Theta}$ to a substitution $\hat{\Theta}$ (of length λ) over the alphabet $G \times \mathcal{X}$, where G is a subgroup of the group \mathcal{S}_c of bijections on $\{0, \dots, c-1\}$. We would like to note that in case of bijective substitutions a similar idea was used in [36] to study the spectrum of the associated system.

Recall that for each $j = 0, \dots, \lambda - 1$ and $M \in \mathcal{X}$, $\tilde{\Theta}(\cdot, M)_j$ is bijective in view of Lemma 6.12 and the definition of $\tilde{\Theta}$. We define the corresponding

permutation $\sigma_{M,j} \in \mathcal{S}_c$ by $\sigma_{M,j}(m) = n$ iff $\tilde{\Theta}(n, M)_j = (m, \tilde{\theta}(M)_j)$. This allows us to define the aforementioned group G as the group generated by the $\sigma_{M,j}$, i.e.

$$(54) \quad G := \langle \sigma_{M,j} : M \in \mathcal{X}, j \in \{0, \dots, \lambda - 1\} \rangle.$$

Finally, we define $\hat{\Theta} : G \times \mathcal{X} \rightarrow (G \times \mathcal{X})^\lambda$ by setting

$$(55) \quad \hat{\Theta}(\sigma, M)_j := (\sigma \circ \sigma_{M,j}, \tilde{\theta}(M)_j),$$

for $j = 0, \dots, \lambda - 1$.

Repeating the same with θ replaced with θ^k (cf. Remark 6.17), we define the permutations $\sigma_{M,j}^{(k)}$ of $\{0, \dots, c - 1\}$ so that

$$(56) \quad \sigma_{M,j}^{(k)}(m) = n \text{ if and only if } \tilde{\Theta}^k(n, M)_j = (m, \tilde{\theta}^k(M)_j).$$

This is a very important formula, as it highlights the connection between $\tilde{\Theta}$ and $\hat{\Theta}$. Furthermore, analogously to (55), we find that

$$(57) \quad \hat{\Theta}^k(\sigma, M)_j = (\sigma \circ \sigma_{M,j}^{(k)}, \tilde{\theta}^k(M)_j),$$

for $j = 0, \dots, \lambda^k - 1$.

It follows directly that $\sigma_{M,j}^{(1)} = \sigma_{M,j}$ and from (11) (applied to $\hat{\Theta}$) that

$$(58) \quad \sigma_{M,j_1 \lambda^{k_2} + j_2}^{(k_1 + k_2)} = \sigma_{M,j_1}^{(k_1)} \circ \sigma_{\tilde{\theta}^{k_1}(M)_{j_1}, j_2}^{(k_2)}$$

for $j_1 < \lambda^{k_1}$ and $j_2 < \lambda^{k_2}$.

If $\tilde{\theta}(M_0)_0 = M_0$, then clearly $\tilde{\theta}^k(M_0)_0 = M_0$ and applying (58) for (σ, M_0) , with $j_1 = j_2 = 0$, we obtain $\sigma_{M_0,0}^{(k)} = (\sigma_{M_0,0})^k$. Hence, by possibly passing to an iterate of $\hat{\Theta}$, we can assume that

$$(59) \quad \hat{\Theta}(\sigma, M_0)_0 = (\sigma, M_0)$$

for each $\sigma \in G$.

We recall (45), which shows the existence of some $k_0 \in \mathbb{N}$ and $j_0 < \lambda^{k_0}$ such that $\tilde{\Theta}^{k_0}(i, M)_{j_0} = (i, M_0)$ for all $i < c$ and $M \in \mathcal{X}$. This gives directly, by (56), that $\sigma_{M,j_0}^{(k_0)} = id$, or in other words

$$(60) \quad \hat{\Theta}^{k_0}(\sigma, M)_{j_0} = (\sigma, M_0)$$

holds for all $\sigma \in G, M \in \mathcal{X}$.

The next few lemmas will be used to show that $\hat{\Theta}$ is primitive. We start by giving some kind of a dual statement to (60).

Lemma 6.20. *For every $M \in \mathcal{X}$ there exist $k_M \in \mathbb{N}$ and $j_M < \lambda^{k_M}$ such that $\hat{\Theta}^{k_M}(\sigma, M_0)_{j_M} = (\sigma, M)$ for each $\sigma \in G$.*

Proof. We fix $M \in \mathcal{X}$. As $\tilde{\theta}$ is primitive, we can find $k \in \mathbb{N}, j < \lambda^k$ such that $\tilde{\theta}^k(M_0)_j = M$, which gives $\hat{\Theta}^k(\sigma, M_0)_j = (\sigma \circ \sigma_{M,j}^{(k)}, M)$ for all $\sigma \in G$. In view of (11), we obtain that

$$(61) \quad \begin{aligned} \hat{\Theta}^{k_0+k}(\sigma, M_0)_{j \lambda^{k_0} + j_0} &= \hat{\Theta}^{k_0}(\hat{\Theta}^k(\sigma, M_0)_j)_{j_0} \\ &= \hat{\Theta}^{k_0}(\sigma \circ \sigma_{M,j}^{(k)}, M)_{j_0} = (\sigma \circ \sigma_{M,j}^{(k)}, M_0). \end{aligned}$$

Set $\ell := \text{ord}(\sigma_{M,j}^{(k)})$. Now, by iterating (61), using (11), we obtain that

$$\widehat{\Theta}^{k'}(\sigma, M_0)_{j'} = (\sigma \circ (\sigma_{M,j}^{(k)})^{\ell-1}, M_0),$$

where $k' = (\ell - 1)(k_0 + k)$ and $j' = \lambda^{(\ell-2)(k_0+k)}(j\lambda^{k_0} + j_0) + \dots + \lambda^{k_0+k}(j\lambda^{k_0} + j_0) + j\lambda^{k_0} + j_0$. It follows that

$$\begin{aligned} \widehat{\Theta}^{k+k'}(\sigma, M_0)_{j+\lambda^k j'} &= \widehat{\Theta}^k(\widehat{\Theta}^{k'}(\sigma, M_0)_{j'})_j = \widehat{\Theta}^k(\sigma \circ (\sigma_{M,j}^{(k)})^{\ell-1}, M_0)_j \\ &= (\sigma \circ (\sigma_{M,j}^{(k)})^\ell, M) = (\sigma, M) \end{aligned}$$

which completes the proof. \square

Next, we find a better description for G .

Lemma 6.21. *We have*

$$G = \langle \sigma_{M_0,j}^{(k)} : k \in \mathbb{N}, j < \lambda^k, \widetilde{\theta}^k(M_0)_j = M_0 \rangle.$$

Proof. By (11), it is clear that any $\sigma_{M_0,j}^{(k)} \in G$. Thus, it remains to write any $\sigma_{M,j}$ as $\sigma_{M_0,j'}^{(k)}$, where $\widetilde{\theta}^k(M_0)_{j'} = M_0$. Using (11), Lemma 6.20, (55) and finally (60), we have

$$\begin{aligned} \widehat{\Theta}^{k_0+1+k_M}(\sigma, M_0)_{j_0+\lambda^{k_0}j+\lambda^{k_0+1}j_M} &= \widehat{\Theta}^{k_0}(\widehat{\Theta}(\widehat{\Theta}^{k_M}(\sigma, M_0)_{j_M})_j)_{j_0} \\ &= \widehat{\Theta}^{k_0}(\widehat{\Theta}(\sigma, M)_j)_{j_0} \\ &= \widehat{\Theta}^{k_0}(\sigma \circ \sigma_{M,j}, \widetilde{\theta}(M)_j)_{j_0} \\ &= (\sigma \circ \sigma_{M,j}, M_0), \end{aligned}$$

which, by (57), ends the proof. \square

This allows us to give the final description for G .

Lemma 6.22. *We have*

$$G = \{\sigma_{M_0,j}^{(k)} : k \in \mathbb{N}, j < \lambda^k, \widetilde{\theta}^k(M_0)_j = M_0\}.$$

Proof. It just remains to show that

$$\{\sigma_{M_0,j}^{(k)} : k \in \mathbb{N}, j < \lambda^k, \widetilde{\theta}^k(M_0)_j = M_0\}$$

is indeed a (finite) group. As each element in our group G is of finite order, all we need to show is that the multiplication of two elements $\sigma_{M_0,j_1}^{(k_1)}, \sigma_{M_0,j_2}^{(k_2)}$, where $\widetilde{\theta}^{k_i}(M_0)_{j_i} = M_0$, is again an element of the set. A simple computation yields

$$\begin{aligned} \widehat{\Theta}^{k_2+k_1}(\sigma, M_0)_{j_2+\lambda^{k_2}j_1} &= \widehat{\Theta}^{k_2}(\widehat{\Theta}^{k_1}(\sigma, M_0)_{j_1})_{j_2} = \widehat{\Theta}^{k_2}(\sigma \circ \sigma_{M_0,j_1}^{(k_1)}, M_0)_{j_2} \\ &= (\sigma \circ (\sigma_{M_0,j_1}^{(k_1)} \circ \sigma_{M_0,j_2}^{(k_2)}), M_0), \end{aligned}$$

so the result follows. \square

Thus, for any $g \in G$, we have some $k_g \geq 1$ and $j_g < \lambda^{k_g}$ such that $g = \sigma_{M_0,j_g}^{(k_g)}$ and $\widetilde{\theta}^{k_g}(M_0)_{j_g} = M_0$. Now, we can finally show the following result.

Proposition 6.23. $\widehat{\Theta}$ is primitive.

Proof. As we have $\widehat{\Theta}(\sigma, M_0)_0 = (\sigma, M_0)$, it is sufficient to show that for all $g, h \in G$ and $M, M' \in \mathcal{X}$ there exist $k_1 \geq 1$, $j_1 < \lambda^{k_1}$ such that $\widehat{\Theta}^{k_1}(id, M_0)_{j_1} = (g, M)$ and $k_2 \geq 1$, $j_2 < \lambda^{k_2}$ such that $\widehat{\Theta}^{k_2}(h, M')_{j_2} = (id, M_0)$. Indeed, for all $k \geq k_1 + k_2$, we obtain that

$$\begin{aligned} \widehat{\Theta}^k(h, M')_{j_2 \lambda^{k-k_2} + j_1} &= \widehat{\Theta}^{k_1}(\widehat{\Theta}^{k-k_1-k_2}(\widehat{\Theta}^{k_2}(h, M')_{j_2})_0)_{j_1} \\ &= \widehat{\Theta}^{k_1}(\widehat{\Theta}^{k-k_1-k_2}(id, M_0)_0)_{j_1} = \widehat{\Theta}^{k_1}(id, M_0)_{j_1} = (g, M). \end{aligned}$$

As there are only finitely many $g, h \in G$ and $M, M' \in \mathcal{X}$, we find that there exists some (large) k such that for all $g, h \in G$ and $M, M' \in \mathcal{X}$ there exists j such that $\widehat{\Theta}^k(h, M')_j = (g, M)$.

Fix now $g \in G$ and $M \in \mathcal{X}$. Since $g = \sigma_{M_0, j_g}^{(k_g)}$, where $\tilde{\theta}^{k_g}(M_0)_{j_g} = M_0$, we have (using Lemma 6.20)

$$\widehat{\Theta}^{k_M+k_g}(id, M_0)_{j_M + \lambda^{k_M} j_g} = \widehat{\Theta}^{k_M}(\widehat{\Theta}^{k_g}(id, M_0)_{j_g})_{j_M} = \widehat{\Theta}^{k_M}(g, M_0)_{j_M} = (g, M).$$

Furthermore, we have (using (59))

$$\begin{aligned} \widehat{\Theta}^{k_{h^{-1}}+k_0}(h, M')_{j_{h^{-1}} + \lambda^{k_{h^{-1}}} j_0} &= \widehat{\Theta}^{k_{h^{-1}}}(\widehat{\Theta}^{k_0}(h, M')_{j_0})_{j_{h^{-1}}} = \widehat{\Theta}^{k_{h^{-1}}}(h, M_0)_{j_{h^{-1}}} \\ &= (h \circ h^{-1}, M_0) = (id, M_0), \end{aligned}$$

which completes the proof. \square

Proposition 6.24. $(X_{\widehat{\Theta}}, S)$ is a topological factor of (X_{Θ}, S) .

Proof. We note that a point in the space $X_{\widehat{\Theta}}$ is of the form $(\sigma_n, M_n)_{n \in \mathbb{Z}}$ and define the equivariant map

$$(62) \quad (\sigma_n, M_n) \mapsto (i_n, M_n),$$

where i_n is defined by $\sigma_n(i_n) = 0$. We claim that the fixed point of $\widehat{\Theta}$ that starts with (id, M_0) is mapped to the fixed point of $\tilde{\Theta}$ that starts with $(0, M_0)$: Indeed, in view of (57), we have for all $k \geq 0$, $n < \lambda^k$ that $\widehat{\Theta}^k(id, M_0)_n = (\sigma_{M_0, n}^{(k)}, \tilde{\theta}^k(M_0)_n)$ and, by (56), we have that $\sigma_{M_0, n}^{(k)}(i) = 0$ if and only if $\tilde{\Theta}^k(0, M_0)_n = (i, \tilde{\theta}^k(M_0)_n)$. Thus, the map (62), sends the fixpoint of $\widehat{\Theta}$ starting with (id, M_0) to the fixpoint of $\tilde{\Theta}$ starting with $(0, M_0)$. Now, the proof follows by the minimality of the dynamical systems under consideration. \square

As $(X_{\widehat{\Theta}}, S)$ is topologically isomorphic to (X_{Θ}, S) , using Remark 6.11, we obtain the following.

Corollary 6.25. (X_{Θ}, S) is a topological factor of $(X_{\widehat{\Theta}}, S)$.

We have seen that $h(\theta) = h(\Theta) = h(\tilde{\Theta})$, and $c(\theta) = c(\Theta) = c(\tilde{\Theta})$. These equalities do not carry over to $\widehat{\Theta}$.

Lemma 6.26. We have $c(\widehat{\Theta}) = |G|$.

Proof. We recall that, by (45), there are k_0, j_0 such that

$$\widehat{\Theta}^{k_0}(\sigma, M)_{j_0} = (\sigma, M_0),$$

which shows that $c(\widehat{\Theta}) \leq |\{(\sigma, M_0) : \sigma \in G\}| = |G|$.

Furthermore, it follows from the definition of $\widehat{\Theta}$ that for all $M \in \mathcal{X}$, we have

$$\widehat{\Theta}^k(\{(g, M) : g \in G\})_j = \{(g \cdot \sigma_{M,j}^{(k)}, \tilde{\theta}^k(M)_j) : g \in G\} = \{(g', \tilde{\theta}^k(M)_j) : g' \in G\}.$$

This shows directly that

$$\widehat{\Theta}^k(G \times \mathcal{X})_j = G \times \tilde{\theta}^k(\mathcal{X})_j,$$

which implies that $c(\widehat{\Theta}) \geq |G|$. \square

Even though for the height one could expect equality, it is not the case in general. The following example shows that $h(\widehat{\Theta})$ and $h(\theta)$ can be different.

Example 6.4. We consider the following bijective (whence $c(\theta) = 3$) substitution θ :

$$\begin{aligned}\theta(a) &= aab \\ \theta(b) &= bcc \\ \theta(c) &= cba.\end{aligned}$$

Since $a \mapsto aab$, obviously, $h(\theta) = 1$. Of course $\mathcal{X} = \{a, b, c\}$, so in fact for $\widehat{\Theta}$ we are interested only in the first coordinate. The bijections given by the columns yield three permutations of $\{0, 1, 2\}$: $\sigma_0 = Id$, $\sigma_1 = (12)$ and $\sigma_2 = (012)$. As they generate \mathcal{S}_3 , we have $G = \mathcal{S}_3$. We also have

$$\widehat{\Theta}(\sigma) = (\sigma)(\sigma \circ (12))(\sigma \circ (021)),$$

for all $\sigma \in \mathcal{S}_3$.

We consider the fixpoint u that starts with id and find that $u[4] = (12)(12) = id$ and $u[26] = (021)(021)(021) = id$. Thus, see (19), we have $g_0 = \gcd(S_0) \leq 2$. Furthermore, we find that for $n = \sum_{i=0}^r n_i 3^i$,

$$u[n] = \widehat{\Theta}^{r+1}(id)_n = \sigma_{n_r} \circ \dots \circ \sigma_{n_0}.$$

We consider the sign of a permutation and find that

$$\text{sgn}(\sigma_0) = \text{sgn}(\sigma_2) = 1, \text{sgn}(\sigma_1) = -1$$

which gives

$$\text{sgn}(u[n]) = \prod_{i=0}^r \text{sgn}(\sigma_{n_i}) = \prod_{i=0}^r (-1)^{n_i} = \prod_{i=0}^r (-1)^{n_i 3^i} = (-1)^{\sum n_i 3^i} = (-1)^n.$$

This shows that $h(\widehat{\Theta}) = 2$.

This shows that $\tilde{\Theta}$ is much closer to θ than $\widehat{\Theta}$, but $\widehat{\Theta}$ has a structural advantage, as it relies on a group G . Indeed, we are able to find a representation of the height within the group G .

Lemma 6.27. *Let $\widehat{\Theta} : G \times \mathcal{X} \rightarrow (G \times \mathcal{X})^\lambda$, be as described above. Then, there exists $G_0 \leq G$ such that $G/G_0 \cong \mathbb{Z}/h(\widehat{\Theta})\mathbb{Z}$. Furthermore, if we define $\zeta : G/G_0 \times \mathcal{X} \rightarrow (G/G_0 \times \mathcal{X})^\lambda$ by*

$$\zeta(G_0g, M)_j = (G_0g\sigma_{M,j}, \widetilde{\theta}(M)_j),$$

then ζ is the joining of a periodic substitution p (of period h) and $\widetilde{\theta}$, i.e. $\zeta = p \vee \widetilde{\theta}$.

Proof. Since $\widehat{\Theta}$ has height $h = h(\widehat{\Theta})$, there exists $f : G \times \mathcal{X} \rightarrow \{0, \dots, h-1\}$ (which we also treat as a 1-code) such that the fixed point u of $\widehat{\Theta}$ obtained by iterations of $\widehat{\Theta}$ at (id, M_0) is mapped via f to $01 \dots (h-1)01 \dots$, i.e. $f(u[n]) = n \bmod h$. Furthermore, we can assume without loss of generality that $\lambda \equiv 1 \bmod h$, as we can always replace $\widehat{\Theta}$ by $\widehat{\Theta}^\ell$, if necessary. We find similarly to (47) that

$$f(\widehat{\Theta}^k(\sigma, M)_j) \equiv f((\sigma, M))\lambda^k + j \bmod h$$

for each $k \geq 0$ and $j < \lambda^k$. Using additionally (60), this gives

$$f((\sigma, M_0)) = f(\widehat{\Theta}^{k_0}(\sigma, M)_{j_0}) \equiv f((\sigma, M))\lambda^{k_0} + j_0 \equiv f((\sigma, M)) + j_0 \bmod h,$$

which shows that $f((\sigma, M))$ only depends on σ . Since (for each k) u begins with $\widehat{\Theta}^k(id, M_0)$, by the definition of f , we have $f(\sigma_{M_0, j}^{(k)}) = j \bmod h$. Due to Lemma 6.22, we can write each $g_1, g_2 \in G$ as $g_i = \sigma_{M_0, j_i}^{(k_i)}$. Thus, we find

$$f(g_1g_2) = f(\widehat{\Theta}^{k_2}(\widehat{\Theta}^{k_1}(id, M_0)_{j_1})_{j_2}) \equiv f(\widehat{\Theta}^{k_1}(id, M_0)_{j_1}) + j_2 \equiv j_1 + j_2 \bmod h.$$

This shows, that for all $g_1, g_2 \in G$ we have

$$f(g_1g_2) \equiv f(g_1) + f(g_2) \bmod h.$$

It follows that f is a group homomorphism and therefore $G_0 := f^{-1}(0)$ is a normal subgroup. We can identify Gg by $f(g)$ and the last statement follows immediately for $p : \{0, \dots, h-1\} \rightarrow \{0, \dots, h-1\}^\lambda$ defined by $p(i)_j = \lambda i + j \bmod h$. \square

Remark 6.28. Lemma 6.27 allows us to find a representation of the maximal equicontinuous factor of $\widehat{\Theta}$ in the case that $\widetilde{\theta}$ is not periodic. In general, we need to rely on a more complicated construction.

First, we would like to note that we can find a representation of G in the centralizer of $(X_{\widehat{\Theta}}, S)$. Indeed, given $\tau \in G$, consider $V_\tau : G \times \mathcal{X} \rightarrow G \times \mathcal{X}$ defined by

$$(63) \quad V_\tau(\sigma, M) = (\tau \circ \sigma, M).$$

Note that in view of (55), $\widehat{\Theta}$ “commutes” with V_τ . It follows from Remark 4.7 that V_τ (uniquely) extends to a homeomorphism V_τ of $X_{\widehat{\Theta}}$ and commutes with S :

$$(64) \quad V_\tau((\sigma_n, M_n)_{n \in \mathbb{Z}}) = (\tau \circ \sigma_n, M_n)_{n \in \mathbb{Z}}$$

for each $(\sigma_n, M_n)_{n \in \mathbb{Z}} \in X_{\hat{\Theta}}$. It follows that we have a (finite) group $\mathcal{V} := \{V_\tau : \tau \in G\}$ of homeomorphisms of $X_{\hat{\Theta}}$ commuting with the shift.

Next, we want to determine the dynamical system obtained by factoring $(X_{\hat{\Theta}}, S)$ by \mathcal{V} . For this aim, consider the map $g : (G \times \mathcal{X})^{\mathbb{Z}} \rightarrow (G \times \mathcal{X} \times \mathcal{X})^{\mathbb{Z}}$ defined by

$$g((\sigma_n, M_n)_{n \in \mathbb{Z}}) = (\sigma_{n+1}^{-1} \circ \sigma_n, M_n, M_{n+1})_{n \in \mathbb{Z}},$$

which means that g is a sliding block code with (right-) radius 1. We see directly that $g((\sigma_n, M_n)_{n \in \mathbb{Z}}) = g((\sigma'_n, M'_n)_{n \in \mathbb{Z}})$ if and only if $(\sigma'_n, M'_n)_{n \in \mathbb{Z}} = V_\tau((\sigma_n, M_n)_{n \in \mathbb{Z}})$ for some $\tau \in G$. We want to describe $g(X_{\hat{\Theta}})$. As g is a sliding block code, it follows that $g(X_{\hat{\Theta}})$ is a (minimal) subshift. We want to show that it is actually a substitutional dynamical system, where the substitution has column number 1, i.e. it is synchronizing.

To this end, let us define a substitution $\eta : G \times \mathcal{X} \times \mathcal{X} \rightarrow (G \times \mathcal{X} \times \mathcal{X})^\lambda$ as follows

$$\eta(\sigma, M, M')_j = \begin{cases} (\sigma_{M,j+1}^{-1} \circ \sigma_{M,j}, \tilde{\theta}(M)_j, \tilde{\theta}(M)_{j+1}), & \text{for } j < \lambda - 1 \\ (\sigma_{M',0}^{-1} \circ \sigma \circ \sigma_{M,\lambda-1}, \tilde{\theta}(M)_{\lambda-1}, \tilde{\theta}(M')_0), & \text{for } j = \lambda - 1. \end{cases}$$

A simple computation gives

$$\eta^k(\sigma, M, M')_j = \begin{cases} ((\sigma_{M,j+1}^{(k)})^{-1} \circ \sigma_{M,j}^{(k)}, \tilde{\theta}^k(M)_j, \tilde{\theta}^k(M)_{j+1}), & \text{for } j < \lambda^k - 1 \\ ((\sigma_{M',0}^{(k)})^{-1} \circ \sigma \circ \sigma_{M,\lambda-1}^{(k)}, \tilde{\theta}^k(M)_{\lambda-1}, \tilde{\theta}^k(M')_0), & \text{for } j = \lambda^k - 1. \end{cases}$$

We denote the fixpoint of $\hat{\Theta}$ that starts with (id, M_0) by u and denote $u[n] = (\sigma_n, M_n)$. This allows us to reduce the alphabet of the substitution η from $G \times \mathcal{X} \times \mathcal{X}$ to \mathbb{B} , where

$$\mathbb{B} := \{(\sigma_{n+1}^{-1} \circ \sigma_n, M_n, M_{n+1}) : n \geq 0\}.$$

The formula above shows that g maps the fixed point of $\hat{\Theta}$ starting with $(\sigma_0, M_0)(\sigma_1, M_1) \dots$ to the fixed point of η starting with $(\sigma_1^{-1} \circ \sigma_0, M_0, M_1)$. Furthermore, we see directly that the restriction of $f : X_{\hat{\Theta}} \rightarrow \mathbb{B}^{\mathbb{Z}}$ is still well-defined on the fixedpoint of $\hat{\Theta}$ and therefore on $X_{\hat{\Theta}}$. This restriction is necessary to ensure that η is primitive.

Thus we showed that $g(X_{\hat{\Theta}}) = X_\eta$.

It remains to show that η has column number 1. As $\tilde{\theta}$ has column number 1, we find by (28) some integers $k \geq 1$, $j < \lambda^k - 1$ and $M' \in \mathcal{X}$ such that $\tilde{\theta}^k(M)_j = M'$ holds for all $M \in \mathcal{X}$. It follows that for any $\sigma \in G$ and $M_1, M_2 \in \mathcal{X}$, we have

$$\begin{aligned} \eta^{2k}(\sigma, M_1, M_2)_{j\lambda^k+j} &= \eta^k(\eta^k(\sigma, M_1, M_2)_j)_j \\ &= \eta^k((\sigma_{M_1,j+1}^{(k)})^{-1} \circ \sigma_{M_1,j}^{(k)}, M', \theta^k(M_1)_{j+1})_j \\ &= ((\sigma_{M',j+1}^{(k)})^{-1} \circ \sigma_{M',j}^{(k)}, M', \theta^k(M')_{j+1}) \end{aligned}$$

which shows that $c(\eta) = 1$.

Example 6.5. We give the explicit construction of η for the Rudin-Shapiro sequence, which was already introduced in Example 6.3. We recall that,

$$\begin{aligned}\tilde{\Theta}(0, M_0) &= (0, M_0)(0, M_1), & \tilde{\Theta}(1, M_0) &= (1, M_0)(1, M_1), \\ \tilde{\Theta}(0, M_1) &= (0, M_0)(1, M_1), & \tilde{\Theta}(1, M_1) &= (1, M_0)(0, M_1).\end{aligned}$$

This gives directly

$$\begin{aligned}\hat{\Theta}(\sigma, M_0) &= (\sigma, M_0)(\sigma, M_1), \\ \hat{\Theta}(\sigma, M_1) &= (\sigma, M_0)(\sigma \circ (01), M_1).\end{aligned}$$

We see that $\mathbb{B} = \{(\sigma, M, M') \in G \times \mathcal{X} \times \mathcal{X} : M \neq M'\}$ as the only fixpoint of $\tilde{\theta}$ is given by $M_0 M_1 M_0 M_1 \dots$. Hence, η is defined as follows:

$$\begin{aligned}\eta(\sigma, M_0, M_1) &= (id, M_0, M_1)(\sigma, M_1, M_0), \\ \eta(\sigma, M_1, M_0) &= ((01), M_0, M_1)(\sigma \circ (01), M_1, M_0).\end{aligned}$$

A simple computation yields $\eta^2(\sigma, M, M')_0 = (id, M_0, M_1)$ for all $(\sigma, M, M') \in \mathbb{B}$ which shows that η has column number 1.

This procedure works well to find a representation of the maximal equicontinuous factor when $h(\hat{\Theta}) = 1$. However, in general we are factoring out too much and need to take a subgroup of \mathcal{V} , namely $\mathcal{V}_1 := \{V_\tau : \tau \in G_0\}$. Therefore we define $g_h : (G \times \mathcal{X})^{\mathbb{Z}} \rightarrow (G \times \mathcal{X} \times \mathcal{X} \times \{0, \dots, h-1\})^{\mathbb{Z}}$ as an "extension" of the previously mentioned function $g : (G \times \mathcal{X})^{\mathbb{Z}} \rightarrow (G \times \mathcal{X} \times \mathcal{X})^{\mathbb{Z}}$ in the following way (see the proof of Lemma 6.27 for f and its properties):

$$g_h((\sigma_n, M_n)_{n \in \mathbb{Z}}) = (\sigma_{n+1}^{-1} \circ \sigma_n, M_n, M_{n+1}, f(\sigma_n))_{n \in \mathbb{Z}}.$$

We see directly that $g((\sigma_n, M_n)_{n \in \mathbb{Z}}) = g((\sigma'_n, M'_n)_{n \in \mathbb{Z}})$ if and only if $(\sigma'_n, M'_n)_{n \in \mathbb{Z}} = V_\tau((\sigma_n, M_n)_{n \in \mathbb{Z}})$ for some $\tau \in G_0$, as $f(\tau\sigma) \equiv f(\tau) + f(\sigma) \pmod{h}$.

Thus, we consider now $\eta_h : G \times \mathcal{X} \times \mathcal{X} \times \{0, \dots, h-1\} \rightarrow (G \times \mathcal{X} \times \mathcal{X} \times \{0, \dots, h-1\})^\lambda$, where the first three coordinates coincide with η and the last coordinate can be seen as another substitution $p : \{0, \dots, h-1\} \rightarrow \{0, \dots, h-1\}^\lambda$ where $p(i)_j = \lambda i + j \pmod{h}$.

This gives that $\eta_h = \eta \vee p$. As we have seen that X_η is measure-theoretically isomorphic to $(H_\lambda, m_{H_\lambda}, R)$ and $(X_p, S) \cong (\mathbb{Z}/h\mathbb{Z}, m_{\mathbb{Z}/h\mathbb{Z}}, \tau_h)$, we have found a factor of $X_{\hat{\Theta}}$ that is itself given by X_{η_h} and is measure-theoretically isomorphic to the maximal equicontinuous factor of $X_{\hat{\Theta}}$.

7 Proof of Theorem 0.1, Corollary 0.2 and some questions

Proof of Theorem 0.1: We have already shown (i) of Theorem 0.1. We now need to handle the case $c(\theta) > h(\theta)$.

To this end, we have defined $\theta \vee \tilde{\theta}$ extending θ (and being primitive), where $(X_{\tilde{\theta}}, S)$ "represents" (measure-theoretically) either the $(H_\lambda, m_{H_\lambda}, R)$ factor of (X_θ, S) if θ is not quasi-bijective or it is finite. Via an isomorphic copy $\tilde{\Theta}$ of $\theta \vee \tilde{\theta}$, we finally have got the substitution $\hat{\Theta}$ which is primitive by Proposition 6.23.

Moreover, by Corollary 6.25, $(X_{\hat{\Theta}}, S)$ has (X_θ, S) as its topological factor, and therefore $c(\hat{\Theta}) > h(\hat{\Theta})$ (it cannot be synchronized).³⁶ We only need to prove (ii) of Theorem 0.1 for $(X_{\hat{\Theta}}, S)$.

Via (64), we have shown that there is a compact group, a copy of the group G , namely $\mathcal{V} = \{V_\tau : \tau \in G\}$, in the centralizer of S which consists of homeomorphisms. Because of Lemma 6.26 and Proposition 2.4, the factor σ -algebra $\mathcal{B}(\mathcal{V})$ of subsets which are fixed by all elements of \mathcal{V} “represents” (measure-theoretically) $(H_\lambda, m_{H_\lambda}, R)$. Of course, the L^2 -space for $(X_{\hat{\Theta}}, \mu_{\hat{\Theta}}, S)$ is spanned by continuous functions. Now, the conditional expectation of $F \in C(X_{\hat{\Theta}})$ with respect to $\mathcal{B}(\mathcal{V})$ is given by $\frac{1}{|G|} \sum_{\tau \in G} F \circ V_\tau$, so it is still a continuous function. It follows that each continuous function F is represented as $F_1 + F_2$, where F_1 is continuous and orthogonal to the $L^2(\mathcal{B}(\mathcal{V}))$, and F_2 is represented by a continuous function that belongs to $L^2(\mathcal{B}(\mathcal{V}))$. If we knew that H_λ represents the Kronecker factor of $(X_{\hat{\Theta}}, \mu_{\hat{\Theta}}, S)$, via Proposition 4.10, we could apply Section 4.3 to conclude the proof. But, as we have already noticed, we cannot control $h = h(\hat{\Theta})$ and the Kronecker factor \mathcal{K} of $(X_{\hat{\Theta}}, \mu_{\hat{\Theta}}, S)$ is given by $(H_\lambda \times \mathbb{Z}/h\mathbb{Z}, m_{H_\lambda} \otimes m_{\mathbb{Z}/h\mathbb{Z}}, R \times \tau_h)$. But $\mathcal{B}(\mathcal{V}) \subset \mathcal{K}$, so by a result of Veech [40], there is a compact subgroup $\mathcal{V}_1 \subset \mathcal{V}$ such that \mathcal{K} is the σ -algebra of subsets fixed by all elements of \mathcal{V}_1 . (We are actually able to give a concrete description of \mathcal{V}_1 as $\{V_\tau : \tau \in G_0\}$, where G_0 was defined in Lemma 6.27.) So, we can now repeat all the above arguments (cf. Footnote 23) with \mathcal{V} replaced by \mathcal{V}_1 to complete the proof. \square

Proof of Corollary 0.2: We begin with a 1-code $F : X_\theta \rightarrow \mathbb{C}$ (that is, $F(y)$ depends only on the 0-coordinate $y[0]$ of $y \in X_\theta$). We assume that for some fixed point $y \in X_\theta$, the sequence $(F(S^n y))_{n \in \mathbb{N}}$ is multiplicative. Such an F is a continuous function and (X_θ, S) is a topological factor of $(X_{\hat{\Theta}}, S)$, say $\pi : X_{\hat{\Theta}} \rightarrow X_\theta$ settles a factor map. Now, $F \circ \pi \in C(X_{\hat{\Theta}})$ and (as in the proof of Theorem 0.1) we have $F \circ \pi = F_1 + F_2$, where both F_1, F_2 are continuous, $F_1 \in L^2(\mathcal{B}(\mathcal{V}))$ and $F_2 \perp L^2(\mathcal{B}(\mathcal{V}))$. Take $x \in \pi^{-1}(y)$ that can be chosen as a fixed point of $\hat{\Theta}$. By Theorem 0.1,

$$\frac{1}{N} \sum_{n \leq N} \overline{F_2(S^n x)} F(S^n(\pi(x))) \rightarrow 0$$

since $(F(S^n \pi(x)))$ is multiplicative. On the other hand, by unique ergodicity,

$$\frac{1}{N} \sum_{n \leq N} \overline{F_2(S^n x)} F(S^n(\pi(x))) \rightarrow \int_{X_{\hat{\Theta}}} \overline{F_2} \cdot F \circ \pi \, d\mu_{\hat{\Theta}}.$$

It follows that $F \circ \pi \perp F_2$, which implies $F_2 = 0$. It follows that the spectral measure of $F \circ \pi = F_1$ is discrete.

Let us first assume that $h(\hat{\Theta}) = 1$. First we note that π is actually a 1-code, which shows that $F \circ \pi$ is also a 1-code. Furthermore, we have an explicit formula for F_1 :

$$(65) \quad F_1(x) = \frac{1}{|G|} \sum_{\tau \in G} (F \circ \pi) \circ V_\tau(x).$$

³⁶It has also $(X_{\hat{\theta}}, S)$ as its topological factor.

As $F \circ \pi$ is a 1-code and V_τ is also a 1-code, it follows that F_1 is a 1-code as well. Moreover, by (65), we obtain that

$$F_1(\sigma, M) = \frac{1}{|G|} \sum_{\tau \in G} (F \circ \pi)(\tau\sigma, M) = \frac{1}{|G|} \sum_{\sigma' \in G} (F \circ \pi)(\sigma', M).$$

Thus, $F_1(\sigma, M)$ only depends on M and we see that

$$F(S^n y) = F_1(S^n \overline{M}) = F_1(M_n).$$

As we have shown that $F_1 \in C(X_{\tilde{\theta}})$ is a 1-code and $\tilde{\theta}$ is synchronizing by Proposition 6.3 (i) and the result follows.

Consider now the case $h(\hat{\Theta}) > 1$. We recall that there exists some $f : G \times \mathcal{X} \rightarrow \{0, \dots, h-1\}$ such that $f(y[n]) \equiv n \pmod{h}$. As we have seen at the end of Section 6.5, we can identify \mathcal{V}_1 as $\{V_\tau : \tau \in G_0\}$. Now, we find that $F_1(\sigma, M)$ depends on M and σG_0 , or equivalently on M and $f(\sigma, M) = f(\sigma G_0, M)$. This gives that

$$F(S^n y) = F_1((n \bmod h), M_n).$$

As $(M_n)_{n \in \mathbb{N}}$ and $(n \bmod h)_{n \in \mathbb{N}}$ are both Weyl-rationally almost periodic, we see directly that $(F_1((n \bmod h), M_n))_{n \in \mathbb{N}}$ is also Weyl-rationally almost periodic. Furthermore, we define a substitution $p : \{0, \dots, h-1\} \rightarrow \{0, \dots, h-1\}^\lambda$, by setting:

$$p(i)_j = i\lambda + j \pmod{h},$$

for all $j < \lambda$ and $i \in \{0, \dots, h-1\}$. We immediately compute that $c(p) = h(p) = h$. Moreover, we find that $((n \bmod h), M_n)_{n \in \mathbb{N}}$ is the fixpoint of $p \vee \tilde{\theta}$. We find by Lemma 6.18 that

$$\max(c(p), c(\tilde{\theta})) \leq c(p \vee \tilde{\theta}) \leq c(p)c(\tilde{\theta})$$

and since $c(\tilde{\theta}) = 1$ this gives $c(p \vee \tilde{\theta}) = h$. Furthermore, we have

$$\text{lcm}(h(p), h(\tilde{\theta})) | h(p \vee \tilde{\theta})$$

and as $h(\tilde{\theta}) = 1$ this gives that $h \leq h(p \vee \tilde{\theta})$. However, Lemma 2.3 implies that $h(p \vee \tilde{\theta}) \leq c(p \vee \tilde{\theta})$ and therefore $h(p \vee \tilde{\theta}) = h$. It follows that $(F(S^n(y)))_{n \in \mathbb{N}}$ is given as a 1-code of a fixpoint of a substitution of constant length for which the column number equals the height, or in other words, it has purely discrete spectrum. \square

Remark 7.1. If we want to obtain in the assertion a weaker conclusion,³⁷ namely that such functions are Besicovitch rationally almost periodic, another proof of Corollary 0.2 can be obtained in the following way.

We first recall that if

$$\mathcal{M} := \{\mathbf{v} : \mathbb{N} \rightarrow \mathbb{C} : |\mathbf{v}| \leq 1 \text{ and } \mathbf{v} \text{ is multiplicative}\}$$

³⁷Note that μ^2 is Besicovitch rationally almost periodic but is not Weyl rationally almost periodic, e.g. [4].

then to check that \mathbf{v} has a mean along an arbitrary arithmetic progression, it is enough to show that $\mathbf{v} \cdot \chi$ has a mean for each Dirichlet character χ , see e.g. Proposition 3.1 in [5]. Now, the argument used in the proof of Lemma 2 in [38] works well and it shows that a function $\mathbf{v} \in \mathcal{M}$ taking finitely many values has a mean along any arithmetic progression. Furthermore, Theorem 1.3 in [5] says that \mathbf{v} is either Besicovitch rationally almost periodic or it is uniform. But uniformity is equivalent to aperiodicity by [15], and we have obtained an aperiodic multiplicative automatic sequence, which is in conflict with Theorem 0.1.

8 Proof of Corollary 0.3

Let $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$ be a bounded multiplicative function. Following [1], we say that a topological dynamical system (Y, T) satisfies the strong \mathbf{u} -MOMO property³⁸ if for each increasing sequence $(b_k)_{k \geq 1} \subset \mathbb{N}$, $b_1 = 1$, $b_{k+1} - b_k \rightarrow \infty$, each sequence $(y_k) \subset Y$ and each $f \in C(Y)$, we have

$$(66) \quad \lim_{K \rightarrow \infty} \frac{1}{b_K} \sum_{k \leq K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^n y_k) \mathbf{u}(n) \right| = 0$$

Substituting $f = 1$ above, we see that \mathbf{u} has to satisfy (4).

Clearly, given (Y, T) , (66) implies (1). In fact (for $\mathbf{u} = \boldsymbol{\mu}$), in the class of zero entropy systems, they are equivalent: it is proved in [1] that Sarnak's conjecture holds if and only if all zero entropy systems enjoy the strong MOMO property.³⁹

Lemma 8.1. *For each primitive substitution θ , the system (X_θ, S) has the strong \mathbf{u} -MOMO property for each bounded, aperiodic, multiplicative $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$ satisfying (4).*

Proof. We only need to prove the result for $(X_{\hat{\theta}}, S)$. Then, the result follows immediately from Theorem 25 in [1] or, more precisely, from its proof. Indeed, all ergodic rotations satisfy the strong \mathbf{u} -MOMO property (Corollary 28 in [1]) and the only ergodic joinings between S^p and S^q whenever $p \neq q$ are large enough, are relative products over isomorphisms of Kronecker factors. Moreover, the structure of the space of continuous functions allows us to represent each $F \in C(X_{\hat{\theta}})$ as $F = F_1 + F_2$, with both F_i continuous, F_1 measurable with respect to the Kronecker factor, and F_2 orthogonal to L^2 of that factor. Finally, we apply the reasoning from the proof of Theorem 25 to F_2 . \square

Proof of Corollary (0.3) It follows from Lemma 8.1 and the equivalence of Properties 1 and 3 in Main Theorem in [1] that all MT-substitutional systems satisfy the strong \mathbf{u} -MOMO property for each bounded, aperiodic, multiplicative \mathbf{u} satisfying (4). The uniform convergence follows now from Theorem 7 in [1]. \square

Note that among MT-substitutional systems there are uniquely ergodic models which are topologically mixing [26]. In such models the maximal equicontin-

³⁸Instead of the strong $\boldsymbol{\mu}$ -MOMO property, we speak about the strong MOMO property. The acronym MOMO stands for *Möbius Orthogonality of Moving Orbits*.

³⁹Moreover, no positive entropy systems has the strong MOMO property. We recall (see [11]) that there are positive entropy systems (Y, T) satisfying (1) for all $f \in C(Y)$.

uous factor must be trivial, that is, there is no topological “realization” of the Kronecker factor.

Theorem 0.1 tell us that for each primitive substitution θ , each bounded, multiplicative, aperiodic \mathbf{u} , (1) holds for each $f \in C(X_\theta)$ and arbitrary $x \in X_\theta$. We do not know however whether this convergence is uniform in x . Note that Corollary 0.3 gives the positive answer if \mathbf{u} satisfies additionally (4).

References

- [1] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, T. de la Rue, *Möbius disjointness for models of an ergodic system and beyond*, Israel J. Math., published online (2018), arXiv:1704.03506.
- [2] J.-P. Allouche and Goldmakher, *Mock Characters and the Kronecker Symbol*, arXiv 16080395/v2
- [3] J.-P. Allouche, J. Shallit, *Automatic Sequences, Theory, Applications, Generalizations*, Cambridge Univ. Press, 2003.
- [4] V. Bergelson, J. Kułaga-Przymus, M. Lemańczyk, F. Richter, *Rational sets, polynomial recurrence and symbolic dynamics*, to appear in Ergodic Theory Dynam. Systems, arXiv: 1611.08352.
- [5] V. Bergelson, J. Kułaga-Przymus, M. Lemańczyk, F. Richter, *A structure theorem for level sets of multiplicative functions and applications*, to appear in International Math. Research Notices, arXiv:1708.02613.
- [6] J. Bourgain, P. Sarnak, T. Ziegler, *Disjointness of Möbius from horocycle flows*, in From Fourier analysis and number theory to Radon transforms and geometry, vol. 28 of Dev. Math., Springer, New York, 2013, pp. 67-83.
- [7] A. Danilenko, M. Lemańczyk, *Spectral multiplicities for ergodic flows*, Discrete Continuous Dynam. Systems **33** (2013), 4271-4289.
- [8] F.M. Dekking, *The spectrum of dynamical systems arising from substitutions of constant length*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **41** (1977/78), no. 3, 221–239.
- [9] M. Drmota, *Subsequences of automatic sequences and uniform distribution*, in *Uniform Distribution and Quasi-Monte Carlo Methods*, Radon Series in Computational and Applied Mathematics, vol. 15 (de Gruyter, Berlin, 2014), 87-104.
- [10] J.-M. Deshouillers, M. Drmota, C. Müllner, *Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture*, Studia Mathematica **231** (2015), 83-95.
- [11] T. Downarowicz, J. Serafin, *Almost full entropy subshifts uncorrelated to the Möbius function*, published online (2017), International Math. Research Notices.

- [12] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, *Sarnak's Conjecture – what's new*, in: Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics, CIRM Jean-Morlet Chair, Fall 2016, Editors: S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, Lecture Notes in Mathematics 2213, Springer International Publishing, pp. 418
- [13] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, C. Mauduit, *Substitutions and Möbius disjointness*, in Proceedings of the Oxtoby Centennial Conference, AMS Contemporary Mathematics Series 678 (2016), 151-174.
- [14] N.P. Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lecture Notes in Mathematics 1794 (Springer, Berlin, 2002) Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.
- [15] N. Frantzikinakis, B. Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc. **30** (2017), 67-157.
- [16] N. Frantzikinakis, B. Host, *The logarithmic Sarnak conjecture for ergodic weights*, Annals of Math. **187** (2018), 869–931.
- [17] N. Frantzikinakis, B. Host, *Furstenberg systems of bounded multiplicative functions and applications*, arXiv 1804.08556
- [18] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1-49.
- [19] E. Glasner, *Ergodic theory via joinings*, vol. 101 of Mathematical Surveys and Monographs, American Mathematical Society, Providence 2003.
- [20] A. Górnika, D. Kwietniak, M. Lemańczyk, *Sarnak's conjecture implies the Chowla conjecture along a subsequence*, in: Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics, CIRM Jean-Morlet Chair, Fall 2016, Editors: S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, Lecture Notes in Mathematics 2213, Springer International Publishing,
- [21] J.L. Herning, *Spectrum and Factors of Substitution Dynamical Systems*, PdD dissertation, George Washington University, 2013.
- [22] B. Host, F. Parreau, *Homomorphismes entre systemes dynamiques définis par substitutions* (French) [Homomorphisms between dynamical systems defined by substitutions] Ergodic Theory Dynam. Systems **9** (1989), 469–477.
- [23] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47**(1-2) (1986), 223-225.
- [24] J. Konieczny, *On Sarnak's conjecture of q -multiplicative sequences*, arXiv:1808.06196
- [25] J. Kułaga-Przymus, M. Lemańczyk, *The Möbius function and continuous extensions of rotations*, Monatshefte Math. **178** (2015), 553-582.
- [26] E. Lehrer, *Topological mixing and uniquely ergodic models*, Israel J. Math. **57** (1987), 239-255.

- [27] M. Lemańczyk, M.K. Mentzen, *On metric properties of substitutions*, Compositio Math. **65** (1988), 241-263.
- [28] M. Lemańczyk, M.K. Mentzen, *Compact subgroups in the centralizers of natural factors of an ergodic group extension of a rotation determine all factors*, Ergodic Theory Dynam. Systems **10** (1990), 763-776.
- [29] K. MATOMÄKI AND M. RADZIWIŁŁ, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056.
- [30] K. MATOMÄKI, M. RADZIWIŁŁ, AND T. TAO, *An averaged form of Chowla’s conjecture*, Algebra & Number Theory **9** (2015), 2167–2196.
- [31] C. Mauduit, J. Rivat, *Sur un problème de Gelfond : la somme des chiffres des nombres premiers*, Ann. of Math. (2), **171** (3) (2010), 1591–1646.
- [32] C. Mauduit, J. Rivat, *Prime numbers along Rudin-Shapiro sequences*, J. Eur. Math. Soc., **17** (10) (2015), 2595–2642.
- [33] M.K. Mentzen, *Invariant sub- σ -algebras for substitutions of constant length*, Studia Math. **92** (1989), 257–273.
- [34] M.K. Mentzen, *Ergodic properties of group extensions of dynamical systems with discrete spectra*, Studia Math. **101** (1991), 19–31.
- [35] C. Müllner, *Automatic sequences fulfill the Sarnak conjecture*, Duke Math. J. **166** (2017), 3219–3290.
- [36] M. Queffélec, *Substitution Dynamical Systems - Spectral Analysis*, second edition, Lecture Notes in Mathematics, vol. 1294, Springer-Verlag, Berlin Heidelberg 2010.
- [37] P. Sarnak, *Three lectures on the Möbius function, randomness and dynamics*, <http://publications.ias.edu/sarnak/>.
- [38] J.-C. Schlage-Puchta, *Completely multiplicative automatic functions*, Integers **11** (2011), A31, 8pp.
- [39] T. Tao, <https://terrytao.wordpress.com/2017/10/20/the-logarithmically-averaged-and-non-logarithmically-averaged-chowla-conjectures/>
- [40] W.A. Veech, *A criterion for a process to be prime*, Monatsh. Math. **94** (1982), 335–341.
- [41] S. Yazdani, *Multiplicative functions and k -automatic sequences*, J. Théor. Nombres Bordeaux **13** (2001), 651–658.

Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopin street 12/18, 87-100 Toruń, Poland: mlem@mat.umk.pl

Institut für Diskrete Mathematik und Geometrie, TU Wien, Wiedner Hauptstr. 8-10, 1040 Wien, Austria: clemens.muellner@tuwien.ac.at