# $S$-ADIC EXPANSIONS, INVARIANT MEASURES AND SUBSTITUTIVE SUBSHIFTS

HANDOUT FOR MY TALK AT THE "WORDS" MEETING IN NANCY, JULY 2025

MARTIN LUSTIG (MARSEILLE)

## 1. Goals of my talk

There are several purposes for this talk:

(1) Explain and advertise the methods the author has developed in collaboration with Nicolas Bédaride and Arnaud Hilion over a series of papers in the past 10 years.
(2) Present 2 non-evident results (about the existence of "many" invariant measures) which have been proved by these methods.
(3) Give a bit of "moral" input from somebody outside the symbolic dynamics main stream.
(4) [if time permits] Give a glimpse about a very new result about the characterization of substitutive subshifts.

## 2. Standard and not so standard basics

### 2.1. Subshifts and their languages.

We always denote by $\mathcal{A} = \{a_1, \dots, a_d\}$ (or by $\mathcal{A}_n = \{a_1, \dots, a_{d(n)}\}$ or $\mathcal{B} = \{b_1, \dots, b_{d(\mathcal{B})}\}$) a finite *alphabet*, and by $\mathcal{A}^*$ the free monoid over $\mathcal{A}$. Let $\mathcal{A}^\mathbb{Z}$ be the set of biinfinite words $\mathbf{x} = \dots x_{-1}x_0x_1 \dots$ with $x_k \in \mathcal{A}$, equipped with the shift operator $T : \mathcal{A}^\mathbb{Z} \to \mathcal{A}^\mathbb{Z}$, where $T(\mathbf{x})$ arises from $\mathbf{x}$ by subtracting 1 from the index $k$ of any of the letters $x_k$. The set $\mathcal{O}(\mathbf{x}) = \{T^n(\mathbf{x}) \mid n \in \mathbb{Z}\} \subseteq \mathcal{A}^\mathbb{Z}$ is the *shift-orbit* of $\mathbf{x}$.

A *subshift $X$ over $\mathcal{A}$* is a subset $X \subseteq \mathcal{A}^\mathbb{Z}$ which is non-empty, $T$-invariant and closed (in the product topology of $\mathcal{A}^\mathbb{Z}$ with respect to the discrete topology on $\mathcal{A}$). We denote by $\Sigma(\mathcal{A})$ the set of all subshifts $X \subseteq \mathcal{A}^\mathbb{Z}$.

A subshift $X \subseteq \mathcal{A}^\mathbb{Z}$ is *minimal* if it is a minimal element of $\Sigma(\mathcal{A})$ with respect to the inclusion; equivalently, for any $\mathbf{x} \in X$ one recovers $X$ as closure of $\mathcal{O}(\mathbf{x})$ in $\mathcal{A}^\mathbb{Z}$.

For every subshift $X$ we denote by $\mathcal{L}(X) \subseteq \mathcal{A}^*$ the *language* of $X$, where $w = w_1 \dots w_m \in \mathcal{L}(X)$ iff for some $\mathbf{x} \in X$ one has $x_1 = w_1, \dots, x_m = w_m$. Conversely, by virtually the same rule we see that any infinite set $\mathcal{L} \subseteq \mathcal{A}^*$ determines a subshift $X(\mathcal{L})$, with $X(\mathcal{L}(X)) = X$ and $\mathcal{L}(X(\mathcal{L})) \subseteq \mathcal{L}$, where equality holds if $\mathcal{L}$ is subshift language.

Hence subshifts and subshift languages determine each other vice versa. BUT: The union of infinitely many subshifts may not be closed and hence not be a subshift, while the union of infinitely many subshift languages is always a subshift language. Conversely, the intersection of infinitely many subshifts, if non-empty, is always a subshift, while the intersection of subshift languages, even if assumed to be infinite, may well not be a subshift language.

The *topological entropy* $h_X$ of a subshift $X \subseteq \mathcal{A}^\mathbb{Z}$ is defined via

$$h_X = \lim_{n \to \infty} \frac{\log \#\{w \in \mathcal{L}(X) \mid |w| = n\}}{n},$$

where $|w|$ denotes the length (= the number of letters) of the word $w \in \mathcal{A}^*$.

## 2.2. **Morphisms and the "image subshift".**

Let $\mathcal{A}$ and $\mathcal{B}$ be finite alphabets, and let $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ be a monoid morphism. All of our monoid morphisms are assumed to be *non-erasing*, by which we mean that no $a_i \in \mathcal{A}$ may be mapped to the empty word $\varepsilon \in \mathcal{B}^*$.

The morphism $\sigma$ determines an *incidence matrix*

$$M(\sigma) \;=\; (|\sigma(a_j)|_{b_i})_{b_i \in \mathcal{B},\, a_j \in \mathcal{A}}$$

where for any $v, w \in \mathcal{B}^*$ we denote by $|w|_v$ the number of occurrences of the word $v$ as factor (= subword) in $w$.

Via biinfinite prolongation the morphism $\sigma$ induces a map $\sigma^{\mathbb{Z}} : \mathcal{A}^{\mathbb{Z}} \to \mathcal{B}^{\mathbb{Z}}$, but the set $\sigma^{\mathbb{Z}}(\mathcal{A}^{\mathbb{Z}}) \subseteq \mathcal{B}^{\mathbb{Z}}$ will in general not be closed under the shift operator $T$ on $\mathcal{B}^{\mathbb{Z}}$. Nevertheless, for any shift-orbit $\mathcal{O}(\mathbf{x}) \subseteq \mathcal{A}^{\mathbb{Z}}$ there is a well defined *image orbit* $\sigma^{\mathcal{O}}(\mathcal{O}(\mathbf{x})) := \mathcal{O}(\sigma^{\mathbb{Z}}(\mathbf{x})) \subseteq \mathcal{B}^{\mathbb{Z}}$.

Any monoid morphism $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ (always assumed to be non-erasing !) induces a subshift map

$$\sigma^{\Sigma} : \Sigma(\mathcal{A}) \to \Sigma(\mathcal{B})\,,\; X \mapsto \sigma(X)\,,$$

where for every subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ the *image subshift* $\sigma(X) := \sigma^{\Sigma}(X)$ is defined as the union of the image orbits of the orbits of $X$ (which is always closed !). Equivalently, we can set $\sigma(X) = X(\sigma(\mathcal{L}(X)))$. The subshift $\sigma(X)$ is also the smallest subshift which contains the set $\sigma^{\mathbb{Z}}(X)$, which is in general not a subshift and may not be confused with the image subshift $\sigma(X)$.

## 2.3. **Recognizability.**

A monoid morphism $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ is *recognizable* in a subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ if, roughly speaking, every biinfinite word $\mathbf{y}$ of the image subshift $Y = \sigma(X) \subseteq \mathcal{B}^{\mathbb{Z}}$ can be lifted ("desubstituted") in precisely one way to a preimage word $\mathbf{x} \in X$. The precise definition is technical and a bit tedious, but one can show:

**Lemma 2.1** ([5])**.** *(1) If $X$ is aperiodic ($\Leftrightarrow$ $X$ does not contain any $\ldots www \ldots$), then $\sigma$ is recognizable in $X$ if and only if the induced map $\sigma^{\mathcal{O}}$ is injective on the set of shift-orbits of $X$.*

*(2) If some $\ldots www \ldots$ is contained in $X$, then one has to require in addition that, if $\sigma(w)$ is a proper power (i.e. $\sigma(w) = v^m$ for some $m \geqslant 2$), then so is $w$.*

## 2.4. **Invariant measures and the measure transfer.**

An *invariant measure* $\mu$ on a subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is a Borel measure on $\mathcal{A}^{\mathbb{Z}}$ which is invariant under the shift operator $T$ and has support in $X$. The set of such invariant measures is denoted by $\mathcal{M}(X)$, which is a cone in that it possesses a canonical $\mathbb{R}_{\geqslant 0}$-linear structure. We say that $\mu \in \mathcal{M}(X)$ is *probability* if $\mu(X) = 1$. The support of any non-zero $\mu \in \mathcal{M}(X)$ is a subshift $X' \subseteq X$.

The set $\mathcal{M}(X)$ embeds canonically into the infinite dimensional vector space $\mathbb{R}^{\mathcal{A}^*}$ via $\mu \mapsto (\mu([w]))_{w \in \mathcal{A}^*}$, where the *cylinder* $[w] \subseteq \mathcal{A}^{\mathbb{Z}}$ denotes the set of all biinfinite words $\mathbf{x} \in \mathcal{A}^{\mathbb{Z}}$ with mono-infinite positive half-word that starts with $w$. From the embedding $\mathcal{M}(X) \subseteq \mathbb{R}^{\mathcal{A}^*}$ the set $\mathcal{M}(X)$ inherits the product topology, which agrees with the more generally known (but less practical) *weak*-topology.*

A measure $\mu \in \mathcal{M}(X)$ is *ergodic* if it can not be written as non-trivial linear combination within $\mathcal{M}(X)$ (i.e. $\mu$ is an extremal point of $\mathcal{M}(X)$). Ergodic probability measures are always linearly independent, so that the existence of $m \geqslant 1$ distinct ergodic probability measures on $X$ assures us that $\dim \mathcal{M}(X) \geqslant m$ (and conversely).

For any subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ we consider the map

$$\zeta_X : \mathcal{M}(X) \to \mathbb{R}^{\mathcal{A}}_{\geqslant 0}\;,\;\; \mu \mapsto \vec{v}^{\,\mu}$$

where the *letter frequency vector* $\vec{v}^{\,\mu}$ for any invariant measure $\mu$ is given by $\vec{v}^{\,\mu} := (\mu([a_1]), \ldots, \mu([a_d]))$. We denote by $C(X) := \zeta_X(\mathcal{M}(X)) \subseteq \mathbb{R}^{\mathcal{A}}$ the $\mathbb{R}_{\geqslant 0}$-linear cone which is image of the map $\zeta_X$. Note that for any $\mu \in \mathcal{M}(X)$ we have $\mu(X) = \sum_{a_i \in \mathcal{A}} \mu([a_i])$.

For any monoid morphism $\sigma : \mathcal{A}^* \to \mathcal{B}^*$, any subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ and any measure $\mu \in \mathcal{M}(X)$ there is a well defined *transferred measure* $\mu^\sigma \in \mathcal{M}(\sigma(X))$, and the issuing map

$$\sigma_X^{\mathcal{M}} : \mathcal{M}(X) \to \mathcal{M}(\sigma(X))\,,\; \mu \mapsto \mu^\sigma$$

has the following properties, for any $\mu \in \mathcal{M}(X)$ (see [5], [4]):

(1) $\sigma_X^{\mathcal{M}}$ is $\mathbb{R}_{\geqslant 0}$-linear.
(2) $\sigma_X^{\mathcal{M}}$ is continuous.
(3) $\sigma_X^{\mathcal{M}}$ is surjective.
(4) $\sigma_X^{\mathcal{M}}$ is injective, if $\sigma$ is recognizable in $X$. (Indeed, it suffices that $\sigma^{\mathcal{O}}$ is injective in $X$.)
(5) The support of $\mu^\sigma$ is equal to the image subshift of the support of $\mu$.
(6) The letter frequency vector $\vec{v}^{\,\mu^\sigma}$ of the transferred measure $\mu^\sigma$ is derived from the letter frequency vector $\vec{v}^{\,\mu}$ of $\mu$ by the linear map
$$\vec{v}^{\,\mu} \quad\mapsto\quad \vec{v}^{\,\mu^\sigma} = M(\sigma) \cdot \vec{v}^{\,\mu}\,.$$
(7) If $\mu$ is ergodic, then so is $\mu^\sigma$.
(8) For any $w \in \mathcal{A}^*$ the cylinder measures satisfy $\mu^\sigma([\sigma(w)]) \;\geqslant\; \mu([w])$.
(9) For any $v \in \mathcal{B}^*$ the value of $\mu^\sigma([v])$ can be computed (by hand) via a fairly elegant formula from the values of the $\mu([w])$ of any $w \in \mathcal{A}^*$ with $|w| \leqslant |v| + 2$.
(10) For any non-empty $w \in \mathcal{A}^*$ the atomic measure $\mu_w$ (which is zero outside the orbit of $\ldots www \ldots$ and satisfies $\mu_w(\mathcal{A}^{\mathbb{Z}}) = |w|$) is mapped by $\sigma_X^{\mathcal{M}}$ to the atomic measure $\mu_{\sigma(w)}$.
(11) In particular, for a probability measure $\mu$ the transferred measure $\mu^\sigma$ will in general *not* be probability.

### 2.5. $S$-adic expansions and vector towers.

A backwards infinite (*"directive"*) sequence of monoid morphisms $\overleftarrow{\sigma} = (\sigma_n : \mathcal{A}_{n+1}^* \to \mathcal{A}_n^*)_{n\geqslant 0}$ is said to be *everywhere growing* if the value $\beta_-(n) := \min\{|\sigma_0 \circ \ldots \circ \sigma_{n-1}(a_i)| \mid a_i \in \mathcal{A}_n\}$ (= the *minimal level letter length*) tends to $\infty$ for $n \to \infty$. With this (throughout the sequel assumed) hypothesis the sequence $\overleftarrow{\sigma}$ defines the subshift $X_{\overleftarrow{\sigma}} := X(\mathcal{L}(\overleftarrow{\sigma})) \subseteq \mathcal{A}_0^{\mathbb{Z}}$ as generated by the language $\mathcal{L}(\overleftarrow{\sigma}) := \bigcup\limits_{n\geqslant 0} \{\sigma_0 \circ \ldots \circ \sigma_{n-1}(a_i) \mid a_i \in \mathcal{A}_n\}$. The directive sequence $\overleftarrow{\sigma}$ is an *$S$-adic expansion* (or *$S$-adic development*) of a subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ if one has $\mathcal{A} = \mathcal{A}_0$ and $X = X_{\overleftarrow{\sigma}}$.

For any $n_0 \geqslant 0$ the *truncated* directive sequence $\overleftarrow{\sigma}\dagger_{n_0} := (\sigma_n)_{n\geqslant n_0}$ generates the *level $n_0$ subshift* $X_{n_0} := X_{\overleftarrow{\sigma}\dagger_{n_0}}$, and one has $\sigma_n(X_{n+1}) = X_n$ for all $n \geqslant 0$. We thus obtain the infinite commutative diagram

$$\begin{array}{ccccccccccc}
\ldots & \xrightarrow{\sigma_{n+1}^{\mathcal{M}}} & \mathcal{M}(X_{n+1}) & \xrightarrow{\sigma_n^{\mathcal{M}}} & \mathcal{M}(X_n) & \xrightarrow{\sigma_{n-1}^{\mathcal{M}}} & \ldots & \xrightarrow{\sigma_1^{\mathcal{M}}} & \mathcal{M}(X_1) & \xrightarrow{\sigma_0^{\mathcal{M}}} & \mathcal{M}(X) \\
 & & \downarrow \zeta_{X_{n+1}} & & \downarrow \zeta_{X_n} & & & & \downarrow \zeta_{X_1} & & \downarrow \zeta_X \\
\ldots & \xrightarrow{M(\sigma_{n+1})} & \mathcal{C}(X_{n+1}) & \xrightarrow{M(\sigma_n)} & \mathcal{C}(X_n) & \xrightarrow{M(\sigma_{n-1})} & \ldots & \xrightarrow{M(\sigma_1)} & \mathcal{C}(X_1) & \xrightarrow{M(\sigma_0)} & \mathcal{C}(X)
\end{array}$$

where $(\sigma_n)_{X_{n+1}}^{\mathcal{M}}$ has been abbreviated to $\sigma_n^{\mathcal{M}}$.

From the above surjectivity of the measure transfer map we know that for every $\mu \in \mathcal{M}(X_{\overleftarrow{\sigma}})$ there exists for any $n \geqslant 0$ a measure $\mu_n \in \mathcal{M}(X_n)$ (with $\mu_0 = \mu$) such that together they form a *measure tower* $(\mu_n)_{n\geqslant 0}$ *on* $\overleftarrow{\sigma}$ in that $\mu_n = \mu_{n+1}^{\sigma_n}$ for all $n \geqslant 0$. Via the above maps $\zeta_{X_n}$ we deduce a *vector tower* $(\vec{v}_n)_{n\geqslant 0}$ *on* $\overleftarrow{\sigma}$ (i.e. $\vec{v}_n = M(\sigma_n) \cdot \vec{v}_{n+1}$ for all $n \geqslant 0$) by setting $\vec{v}_n = \vec{v}^{\,\mu_n}$ to be the letter frequency vector of $\mu_n$ (as defined above).

If the sequence $\overleftarrow{\sigma}$ is *totally recognizable* (i.e. every $\sigma_n$ is recognizable in $X_{n+1}$) then $\mu$ determines uniquely the measure tower $(\mu_n)_{n\geqslant 0}$ and hence the vector tower $(\vec{v}_n)_{n\geqslant 0}$. Without that hypothesis we still have:

**Theorem 2.2** ([2],[4])**.** *For any everywhere growing directive sequence $\overleftarrow{\sigma}$ every vector tower $(\vec{v}_n)_{n\geqslant 0}$ on $\overleftarrow{\sigma}$ is derived (as letter frequency vectors) from some measure tower $(\mu_n)_{n\geqslant 0}$ on $\overleftarrow{\sigma}$. In particular, every vector tower $\overleftarrow{v} = (\vec{v}_n)_{n\geqslant 0}$ on $\overleftarrow{\sigma}$ determines an invariant measure $\mu^{\overleftarrow{v}} := \mu_0$ on the subshift $X_{\overleftarrow{\sigma}}$. The issuing map $\mathfrak{m}_{\overleftarrow{\sigma}} : \overleftarrow{v} \mapsto \mu^{\overleftarrow{v}}$ from the set of vector towers on $\overleftarrow{\sigma}$ to $\mathcal{M}(X)$ is $\mathbb{R}_{\geqslant 0}$-linear and surjective.*

As a consequence, we observe that the family of incidence matrices for the level maps of any $S$-adic development of a subshift $X$ determines the measure cone $\mathcal{M}(X)$, up to $\mathbb{R}_{\geqslant 0}$-linear isomorphisms.

To finish this subsection we quote two well known criteria (see [6]) for subshifts $X$ by means of any $S$-adic expansion $(\sigma_n : \mathcal{A}_{n+1}^* \to \mathcal{A}_n^*)_{n\geqslant 0}$ of $X$, both stated here only for the case needed below:

**Proposition 2.3.** *If for infinitely many integers $n \geqslant 0$ the incidence matrix $M(\sigma_n)$ is positive, then $X$ is minimal.*

**Proposition 2.4.** *If the (above defined) minimal level letter length $\beta_-(n)$ satifies $\lim_{n\to\infty} \frac{\log(\mathrm{card}(\mathcal{A}_n))}{\beta_-(n)} = 0$, then $X$ has topological entropy $h_X = 0$.*

## 3. Subshifts with many ergodic measures

For any integer $d \geqslant 1$ we now consider for an arbitrary parameter $\ell \in \mathbb{R}_{\geqslant 0}$ the $d \times d$-matrix $M_d(\ell) := \ell \cdot \mathrm{Id}_{d\times d} + \mathbf{1}_{d\times d}$, where $\mathbf{1}_{d\times d}$ is the "Attila matrix" with all entries equal to 1. We observe that the $d$-dimensional *center vector* $\vec{c}_d$ (also with all entries equal to 1) satisfies $M_d(\ell)\vec{c}_d = (\ell+d)\vec{c}_d$, while for the $k$-th coordinate unit vector $\vec{e}_k$ we obtain $M_d(\ell)\vec{e}_k = \ell\vec{e}_k + \vec{c}_d$. Hence the 2-dimensional $\mathbb{R}_{\geqslant 0}$-subcone spanned by $\vec{c}_d$ and $\vec{e}_k$ is mapped into itself, where the extremal direction of $\vec{c}_d$ is fixed, while the other extremal direction, given by $\vec{e}_k$, is mapped arbitrarily close to itself by choosing the parameter $\ell$ sufficiently large. It follows directly that for any sufficiently fast growing sequence of parameters $\ell(n) \geqslant 0$ the nested infinite intersection $\ldots \subseteq C^2 \subseteq C^1 \subseteq C^0$ of the cones $C^n := M_d(\ell(0)) \cdot \ldots \cdot M_d(\ell(n)) \cdot \mathbb{R}^d_{\geqslant 0}$ has dimension $d$. This shows (invoking also Proposition 2.4):

**Theorem 3.1** ([3])**.** *For any integer $d \geqslant 1$ with alphabet $\mathcal{A}_d = \{a_1, \ldots, a_d\}$, and for any sufficiently fast growing family of positive integers $(\ell(n))_{n\geqslant 0}$, the subshift $X \subseteq \mathcal{A}_d^{\mathbb{Z}}$ generated by the directive sequence $(\sigma_d(\ell(n)) : \mathcal{A}_d^* \to \mathcal{A}_d^*)_{n\geqslant 0}$, with $\sigma_d(\ell(n)) : a_k \mapsto a_k^{\ell(n)} a_1 a_2 \ldots a_d$ for all $k = 1, \ldots, d$, is minimal and possesses d distinct invariant ergodic probability measures.*

*[In [3] it is also shown that the above sequence of the morphisms $\sigma_d(\ell(n))$ can be refined to a sequence of morphisms $\sigma_n : \mathcal{A}_d^* \to \mathcal{A}_d^*$ where each $\sigma_n$ belongs to an a priori given set $S$ of finitely many (indeed 4 suffice) substitutions.]*

We now use the same basic method to produce a large family of subshifts $X$ with a bit of an exotic combination of properties, as $X$ is "small" in that $X$ is minimal and has entropy $h_X = 0$, while simultaneously $X$ is "large" since it carries infinitely many invariant ergodic probability measures. The first such families have recently been established by V. Cyr and B. Kra in [9].

For this purpose we define for every $d \geqslant 2$ the morphism $\tau_d : \mathcal{A}_{d+1}^* \to \mathcal{A}_d^*$ via $a_k \mapsto a_k^2$ if $k = 1, \ldots, d$ and $a_{d+1} \mapsto a_1 a_2 \ldots a_d$. The exponent 2 is used in showing that $\tau_d$ is recognizable in any subshift $X \subseteq \mathcal{A}_{d+1}^{\mathbb{Z}}$ which does not contain any of the periodic words $\ldots a_k a_k a_k \ldots$. Similarly, the morphisms $\sigma_d(\ell)$ from Theorem 3.1 are recognizable in the full shift $\mathcal{A}_d^{\mathbb{Z}}$, if $\ell \geqslant 2$. It follows that the composition $\sigma_d(\ell) \circ \tau_d$ induces a measure transfer map $(\sigma_d(\ell) \circ \tau_d)_X^{\mathcal{M}}$ which is injective for any aperiodic subshift $X \subseteq \mathcal{A}_{d+1}^{\mathbb{Z}}$.

We now observe that, again, for any $d \geqslant 2$ the $d \times (d+1)$-matrix $M'_d(\ell) := M(\sigma_d(\ell) \circ \tau_d)$ maps $\vec{c}_{d+1}$ to a multiple of $\vec{c}_d$ , and $\vec{e}_k$ (for $1 \leqslant k \leqslant d$) to some $\lambda \vec{e}_k + \lambda' \vec{c}_d$, with $\lambda / \lambda'$ large for large $\ell$. Hence the same arguments as before show that, for any fixed $d \geqslant 2$ and varying large $n \geqslant d$, the subcone of $\mathbb{R}^{n+1}_{\geqslant 0}$ spanned by $\vec{e}_1, \ldots, \vec{e}_d$ has for the products $M'_d(\ell_d(d)) \cdot M'_{d+1}(\ell_d(d+1)) \cdot \ldots \cdot M'_n(\ell_d(n))$ nested image-cone intersection "from infinity" of dimension $d$, provided that the parameters $\ell_d(n)$ grow fast enough for $n \to \infty$. Setting $\ell(n) := \max\{\ell_2(n), \ell_3(n), \ldots, \ell_n(n)\}$ we obtain from the above injectivity of the measure transfer and from the criteria in Propositions 2.3 and 2.4 the desired conclusion:

**Theorem 3.2** ([4], Thm. 7.4)**.** *For any sufficiently fast growing family of positive integers $(\ell(n))_{n \geqslant 2}$ the subshift $X \subseteq \mathcal{A}_2^{\mathbb{Z}}$ generated by the directive sequence $(\sigma_n(\ell(n)) \circ \tau_n)_{n \geqslant 2}$ possesses infinitely many distinct invariant ergodic probability measures. Furthermore $X$ is minimal and has entropy $h_X = 0$.*

The method presented here admits many more applications, both with respect to showing unique ergodicity as well as showing the existence of infinitely many ergodic probability measures. Preliminary calculations of the speaker indicate the possibility of such applications for instance to the Avila-Damanik-Zhang counter-examples to the Simon conjecture in [1], to the "Grillenberger-type" subshifts exhibited by Cassaigne-Nicolas in §4.4.3 of [7], to the results of Méla-Petersen [11] for the Pascal-adic subshift, and also towards unique ergodicity results "à la Boshernitzan" for subshifts of infinite alphabet rank.

## 4. What really is the true nature of a subshift ?

The free monoid $\mathcal{A}^*$ over the alphabet $\mathcal{A} = \{a_1, \ldots, a_d\}$ embeds canonically into the free group $F(\mathcal{A})$ over $\mathcal{A}$. But contrary to $\mathcal{A}^*$, where the minimal generating system $\mathcal{A}$ is uniquely determined by $\mathcal{A}^*$, in $F(\mathcal{A})$ there are infinitely many sets $\mathcal{B} = \{w_1, \ldots, w_d\} \subseteq F(\mathcal{A})$ with canonical isomorphisms $F(\mathcal{B}) \cong F(\mathcal{A}) \cong F_d$, and none of these *bases* for the free group $F_d$ of *rank* $d \geqslant 2$ is preferred in any way. Any subshift $X$ over $\mathcal{A}$ gives canonically rise (by passing to the language $\mathcal{L}(X)$) to a "subshift with inverses" over $\mathcal{B}$, which has led to the basis-free notion of an *algebraic lamination* in $F_d$ (see [8]), together with a canonical embedding $\Sigma(\mathcal{A}) \to \Lambda(F(\mathcal{A}))$ of the space of subshifts $\Sigma(\mathcal{A})$ into the space of algebraic laminations $\Lambda(F(\mathcal{A}))$. Similarly, the space of invariant measures $\mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ embeds canonically into the space of *currents* $\mathcal{M}(F(\mathcal{A}))$.

ASIDE: Symbolic dynamists feel traditionally uneasy about the behavior of inverses under morphisms, but this is mainly due to the fact that the notion of *train track maps* has not yet dissipated into the symbolic dynamics community. With this tool the whole $S$-adic machinery as well as most other symbolic dynamics methods and results could (and should) be carried over from symbolic dynamics to geometric group theory.

There is also a "response" from geometric group theory towards symbolic dynamics, namely that any property of a subshift which is not invariant under change of basis is not accepted as "intrinsic" property, just like properties of matrix groups are not intrinsic properties of the group in question if they are not invariant under group isomorphisms, or similarly for properties of topological objects, if they depend on the embedding of the object in an ambient space (like the well known "2-sided coloring" criterion for a surface to be orientable or not).

To stay within symbolic dynamics terminology I'd like to make this a bit more precise:

**Definition 4.1.** A property of a subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is said to be *intrinsic* if, first, for any monoid morphisms $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ which is recognizable in $X$, the property must also hold for the subshift $\sigma(X) \subseteq \mathcal{B}^{\mathbb{Z}}$. Second, for any monoid morphism $\sigma' : \mathcal{C}^* \to \mathcal{A}^*$ and any subshift $Y \subseteq \mathcal{C}^{\mathbb{Z}}$ with $\sigma'(Y) = X$, if $\sigma'$ is recognizable in $Y$, then the property must also hold for $Y$.

Examples of intrinsic properties are minimality, unique ergodicity, the number $e(X)$ of ergodic probability measures, the statements $h_X = 0$ or $h_X > 0$, and the growth-type of the complexity function $p_X(\cdot)$ . The value of $h_X > 0$ however is *not* intrinsic, and neither is the complexity function itself (and not even its equivalence class $\Theta(p_X)$, see [10]).

## 5. Classification of substitutive subshifts

The speaker has very recently discovered (in the context of investigating a certain type of free group automorphisms) a new computable invariant for any minimal substitutive subshift, which consists of a cyclic sequence of finite graphs and graph maps between them. This invariant appears to be (work in progress) a characterizing invariant of the given subshift, up to recognizable morphisms as in Definition 4.1.

The technicalities of the graphs in question are not yet matured enough to be presented here (other than via the examples given in the Annex below), but the main idea ought to be conveyed anyway:

For any $n \geqslant 0$ the *level* $2n$ *Rauzy graph* $R_{2n}(X)$ of a subshift $X \subseteq \mathcal{A}$ can be reinterpreted as obtained in the following way: One first realizes $X$ graphically as a (typically infinite) collection of lines $\gamma(\mathbf{x})$, one for every $\mathbf{x} \in X$, subdivided as biinfinite edge path, with edges labeled by letters from $\mathcal{A}$ according to the letters $x_k$ on $\mathbf{x} = \ldots x_{-1}x_0x_1\ldots$. In a second step we identify any two vertices $P \in \gamma(\mathbf{x})$ and $Q \in \gamma(\mathbf{x}')$ iff the finite sub-edge-paths of length $2n$ on $\mathbf{x}$ and on $\mathbf{x}'$, centered at $P$ and $Q$ respectively, read off the same word. Finally, to get the finite graph $R_{2n}(X)$ we need to identify any two edges with same endpoints and same label. The subshift $X$ can then be read off from suitable edge paths in $\mathbb{R}_{2n}(X)$, and $X$ is characterized by the fact that this "read-off property" holds for any $n \geqslant 1$. Another pay-off of this alternative construction are canonical label-preserving graph morphisms $R_{2m}(X) \to R_{2n}(X)$ for any $m \geqslant n \geqslant 0$, which define in turn a canonical $S$-adic *Rauzy development* of $X$ which is always totally recognizable. Locally, each of these label-preserving graph morphisms is a composition of vertex-identifications and edge-foldings.

In the special case where $X$ is the subshift generated by a primitive substitution $\sigma$, we can use the incidence matrix $M(\sigma)$ and one of its (left) row PF-eigenvectors $\vec{v}^{\,*}$, in order to define from the coefficients of $\vec{v}^{\,*}$ a length function $L$ on the letters of $\mathcal{A}$ (and thus by summation on all of $\mathcal{A}^*$) which satisfies $L(\sigma(a_i)) = \lambda L(a_i)$ for each $a_i \in \mathcal{A}$, where $\lambda > 1$ is the PF-eigenvalue of $M(\sigma)$.

We can now repeat the above definition of the Rauzy graph $R_{2n}(X)$, but replace the combinatorial length used there (when considering for the vertex identification the $2n$-length sub-edge-paths) by the length $L$, and apply the identification device not just to vertices but also to points in the interior of edges (with "read-off equality of sub-edge-paths" refined by passing to $\sigma$-iterates of those paths).

This gives a continuity of graphs $R(\vec{v}^{\,*})$, one for any positive eigenvector $\vec{v}^{\,*}$ within the uniquely determined PF-eigen-direction of $M(\sigma)$, and for $\vec{v}_2^{\,*} = \lambda \vec{v}_1^{\,*}$ the corresponding graphs are related by a graph isomorphism $R(\vec{v}_1^{\,*}) \to R(\vec{v}_2^{\,*})$ which stretches every edge by the factor $\lambda$. The composition $R(\vec{v}_1^{\,*}) \to R(\vec{v}_2^{\,*}) \to R(\vec{v}_1^{\,*})$ of this homothetic graph isomorphism with the above "vertex-identification & edge-folding" map $R(\vec{v}_2^{\,*}) \to R(\vec{v}_1^{\,*})$ is a topological realization of some substitution $\sigma'$ which generates $X$.

To get to the desired finite cycle of graphs we have to discretize the just obtained continuous "loop of graphs" in a canonical way. There are several possibilities for this canonical discretization (all built on periodic points of the composed maps $R(\vec{v}_1^{\,*}) \to R(\vec{v}_2^{\,*}) \to R(\vec{v}_1^{\,*})$, see the examples in the Annex below), and the choice of the most natural among them is one of the problems still on my desk.

## References


[1] A. Avila, D. Damanik and Z. Zhang, *Singular Density of States Measure for Subshift and Quasi-Periodic Schrödinger Operators.* Communications in Mathematical Physics **330** (2014), 469–498

[2] N. Bédaride, A. Hilion and M. Lustig, *Graph towers, laminations and their invariant measures.* J. London Math. Soc. (2) **101** (2020), 1112–1172

[3] N. Bédaride, A. Hilion and M. Lustig, *Tower power for $S$-adics.* Math. Z. **297** (2021), 1853–1875

[4] N. Bédaride, A. Hilion and M. Lustig, *Measure transfer and $S$-adic developments for subshifts.* Ergodic Theory Dynam. Systems **44** (2024), 3120–3154

[5] N. Bédaride, A. Hilion and M. Lustig, *The measure transfer for subshifts induced by a morphism of free monoids.* Nonlinearity **38** (2025), 025001

[6] V Berthé and V. Delecroix, *Beyond substitutive dynamical systems: $S$-adic expansions.* RIMS Kôkyûroku Bessatsu **B46** (2014), 81–123

[7] V. Berthé and M. Rigo (eds.), *Combinatorics, automata and number theory.* Encyclopedia Math. Appl. **135**. Cambridge Univ. Press, Cambridge, 2010.

[8] T. Coulbois, A. Hilion and M. Lustig, *$\mathbb{R}$-trees and laminations for free groups I: Algebraic laminations.* J. Lond. Math. Soc. (2) **78** (2008), 723–736

[9] V. Cyr and B. Kra, *Realizing ergodic properties in zero entropy subshifts.* Isr. J. Math. **240** (2020), 119–148

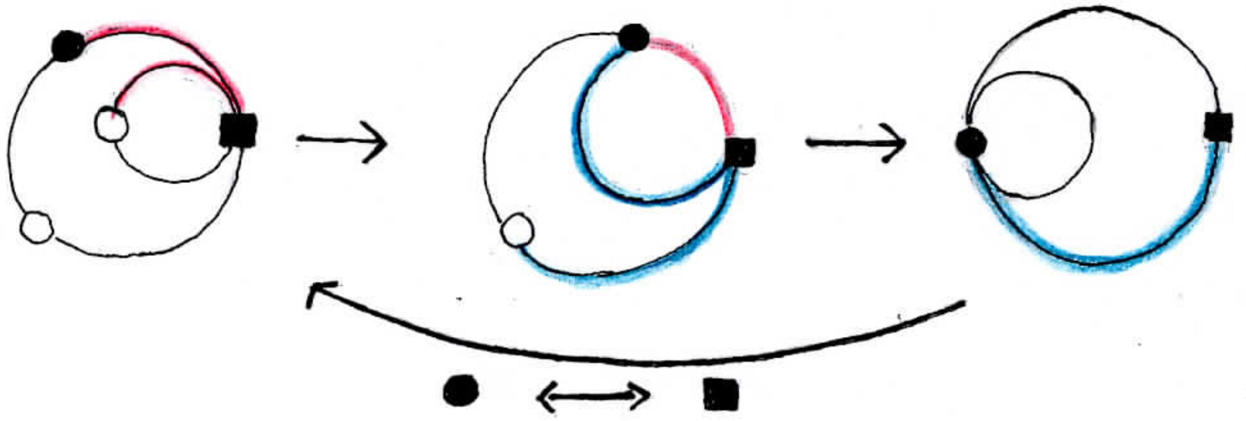
[10] M. Lustig, *How do topological entropy and factor complexity behave under monoid morphisms and free group basis changes ?* arXiv:2204.00816

[11] X. Méla and K. Petersen. *Dynamical properties of the Pascal adic transformation.* Ergodic Theory Dynam. Systems **25** (2005), 227–256

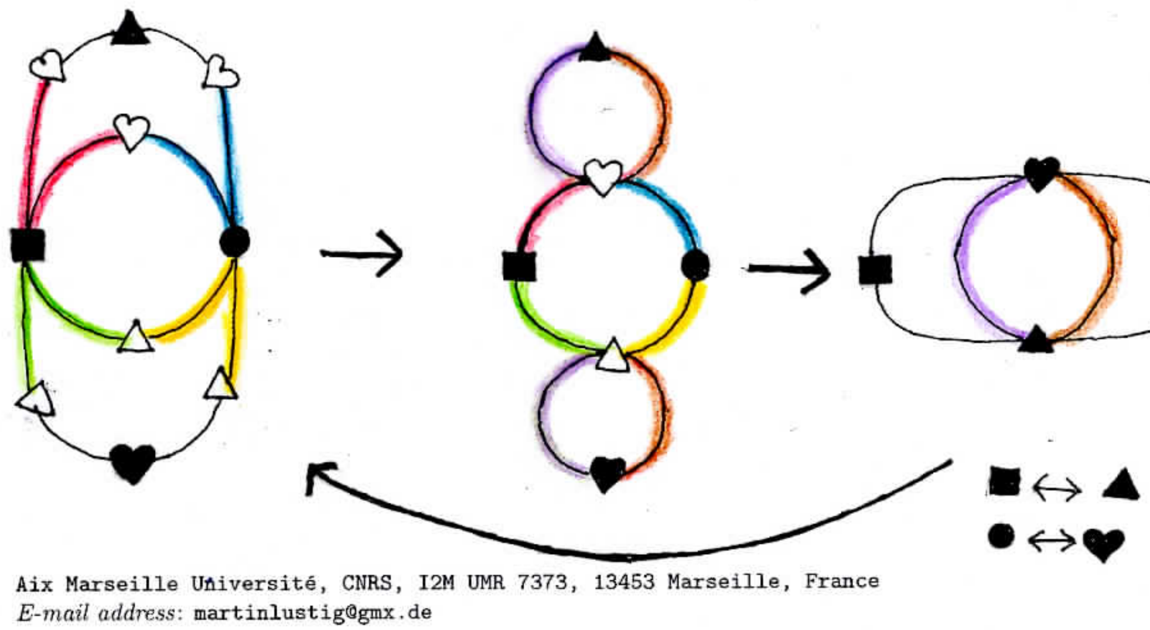
## 6. Annex

Below we draw the characteristic graph morphism cycles for (a) the Fibonacci subshift and (b) the Thue-Morse subshift. Each of the two cycles consists of 3 graphs and 3 graph morphism, where two of the morphisms drawn from left to right, and the last one, from the right-most graph back to the left-most graph, is a graph isomorphism. The vertices are highlighted by symbols (square, heart, ...) which are preserved by the "from-left-to-right" maps and permuted (as indicated) by the "from-right-to-left" isomorphisms. The bold symbols correspond to periodic vertices, while the outlined ones are only pre-periodic. Edges are always mapped to edges, and if two edges are mapped to the same edge this is indicated by colors.

(a) Fibonacci



(b) Tue-Morse



Aix Marseille Université, CNRS, I2M UMR 7373, 13453 Marseille, France
*E-mail address*: martinlustig@gmx.de