Balancedness, natural codings, and recognisability

Wolfgang Steiner

IRIF, CNRS, Université de Paris

Dyadisc 6: Dynamics on the Cantor set and applications

Amiens, July 5–7, 2023

(joint work with Marie-Pierre Béal, Valérie Berthé, Olivier Carton, Nicolas Chevallier, Dominique Perrin, Léo Poirier, Antonio Restivo, Jörg Thuswaldner, and Reem Yassawi)

Chairman assignment problem (Tijdeman '80)

"Suppose d states form a union and every year a union chairman has to be selected in such a way that at any time the accumulated number of chairmen from each state is proportional to its weight."

How to get in an effective way an assignment with small discrepancy?

(letter) discrepancy of $u = u_0 u_1 \cdots \in A^{\mathbb{N}}$, $A = \{1, 2, \ldots, d\}$:

$$\Delta(u) = \sup_{n \geq 0} \left\| \boldsymbol{\ell}(u_{[0,n)}) - n\,\boldsymbol{\alpha} \right\|_\infty$$

$$u_{[k,n)} = u_k u_{k+1} \cdots u_{n-1}$$

$\boldsymbol{\ell}(w) = (|w|_1, \ldots, |w|_d)$, $\quad |w|_i$ number of occurrences of i in w

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d), \quad \alpha_i = \lim_{n \to \infty} \frac{|u_{[0,n)}|_i}{n} \quad \text{(if limit exists)}$$

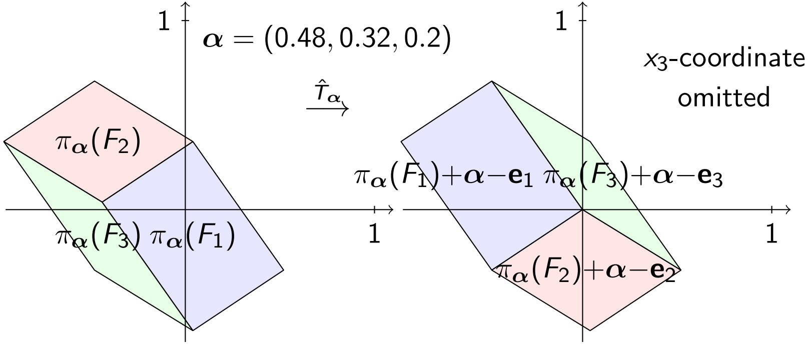
Codings of hypercubic billiards

$$\boldsymbol{\alpha} \in (0,1)^d,\ \sum_{i=1}^d \alpha_i = 1, \quad \mathbf{1}^\perp = \{(x_1,\ldots,x_d) \in \mathbb{R}^d : \sum_{i=1}^d x_i = 0\}$$

$$\hat{T}_{\boldsymbol{\alpha}} : \mathbf{1}^\perp \to \mathbf{1}^\perp, \quad \mathbf{x} \mapsto \mathbf{x} + \boldsymbol{\alpha} - \mathbf{e}_j \quad \text{if } \frac{1-x_j}{\alpha_j} = \min_{1\le i\le d} \frac{1-x_i}{\alpha_i}$$

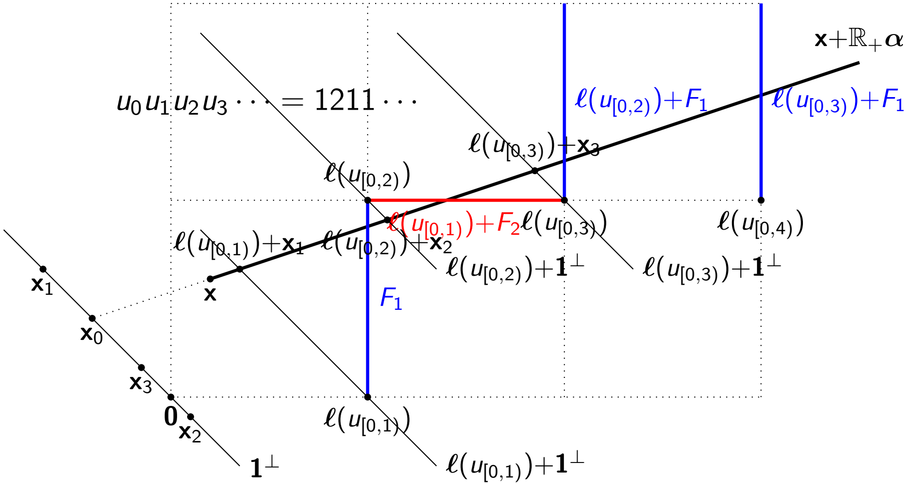
$$\hat{T}_{\boldsymbol{\alpha}}(E_{\boldsymbol{\alpha}}) = E_{\boldsymbol{\alpha}},\ E_{\boldsymbol{\alpha}} = \pi_{\boldsymbol{\alpha}}([0,1]^d) = \bigcup_{i=1}^d \pi_{\boldsymbol{\alpha}}(F_i),\ F_i = \{\mathbf{x} \in [0,1]^d : x_i = 1\}$$

$\pi_{\boldsymbol{\alpha}}$ projection along $\mathbb{R}\boldsymbol{\alpha}$ on $\mathbf{1}^\perp$, i.e., $\pi_{\boldsymbol{\alpha}}(\mathbf{e}_i) = \mathbf{e}_i - \boldsymbol{\alpha}$



$$\mathbf{x}_n = \hat{T}_{\boldsymbol{\alpha}}^n(\mathbf{x}_0) = \mathbf{x}_0 + n\,\boldsymbol{\alpha} - \ell(u_{[0,n)}),\ \text{with } \hat{T}_{\boldsymbol{\alpha}}^n(\mathbf{x}_0) \in \pi_{\boldsymbol{\alpha}}(F_{u_n})$$

Hypercubic billiards, $\mathbf{x}, \boldsymbol{\alpha} \in [0,1]^d$, $\sum_{i=1}^{d} \alpha_i = 1$



$$\mathbf{x}_n = \mathbf{x}_0 + n\,\boldsymbol{\alpha} - \ell(u_{[0,n)}), \quad \text{with } \{\mathbf{x}_0\} = (\mathbf{x}+\mathbb{R}\boldsymbol{\alpha}) \cap \mathbf{1}^{\perp}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \boldsymbol{\alpha} - \ell(u_n), \quad (\mathbf{x}+\mathbb{R}\boldsymbol{\alpha}) \cap (\ell(u_{[0,n)})+F_{u_n}) \neq \emptyset$$

Billiards as natural codings of toral translations

$\hat{T}_{\boldsymbol{\alpha}}$ exchange of domains on $E_{\boldsymbol{\alpha}} = \bigcup_{i=1}^{d} \pi_{\boldsymbol{\alpha}}(F_i)$

$E_{\boldsymbol{\alpha}}$ fundamental domain of $\mathbf{1}^{\perp}/(\mathbb{Z}^d \cap \mathbf{1}^{\perp})$

$\boldsymbol{\alpha} - \mathbf{e}_i \equiv \boldsymbol{\alpha} - \mathbf{e}_j \bmod \mathbb{Z}^d \cap \mathbf{1}^{\perp}$

$\Rightarrow$ induced map of $\hat{T}_{\boldsymbol{\alpha}}$ on $\mathbf{1}^{\perp}/(\mathbb{Z}^d \cap \mathbf{1}^{\perp})$ translation by $\boldsymbol{\alpha} - \mathbf{e}_1$

$u_0 u_1 \cdots$ s.t. $\hat{T}_{\boldsymbol{\alpha}}^n(\mathbf{x}_0) \in \pi_{\boldsymbol{\alpha}}(F_{u_n})$ for some $\mathbf{x}_0$ and all $n \geq 0$

natural coding of this induced map

$$\iota : \mathbf{1}^{\perp} \to \mathbb{R}^{d-1}, \ (y_1, \ldots, y_d) \mapsto (y_1, \ldots, y_{d-1}) \quad \text{bijection}$$

$\iota(E_{\boldsymbol{\alpha}})$ fundamental domain of $\mathbb{T}^{d-1} = \mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$

$\iota(E_{\boldsymbol{\alpha}}) = \bigcup_{i=1}^{d} \iota(\pi_{\alpha}(F_i))$ (topological) partition

$u_0 u_1 \cdots$ s.t. $\iota(\hat{T}_{\boldsymbol{\alpha}}^n(\mathbf{x}_0)) \in \iota(\pi_{\alpha}(F_{u_n}))$ for some $\mathbf{x}_0$ and all $n \geq 0$

natural coding of the toral translation

$$T_{\boldsymbol{\alpha}} : \mathbb{T}^{d-1} \to \mathbb{T}^{d-1}, \ \mathbf{x} \to \mathbf{x} + \iota(\boldsymbol{\alpha})$$

subshift defined by $u_0 u_1 \cdots$ has *purely discrete spectrum*

Properties of hypercubic billiard sequences

Coordinates of $\boldsymbol{\alpha}$ linearly independent over $\mathbb{Q}$, i.e., $\iota(\boldsymbol{\alpha})$ totally irrational

$\Rightarrow$ $T_{\boldsymbol{\alpha}}$ minimal, i.e., $\{T_{\boldsymbol{\alpha}}^n(\iota(\mathbf{x}_0)) : n \geq 0\}$ dense in $\mathbb{T}^{d-1}$ for all $\mathbf{x}_0$

$\Rightarrow$ $\{\hat{T}_{\boldsymbol{\alpha}}^n(\mathbf{x}_0) : n \geq 0\}$ dense in $E_{\boldsymbol{\alpha}}$

$$\Delta(u) = \sup_{n\geq 0} \|n\,\boldsymbol{\alpha} - \boldsymbol{\ell}(u_{[0,n)})\|_\infty = \sup_{n\geq 0} \|\mathbf{x}_n - \mathbf{x}_0\|_\infty = \max_{\mathbf{y}\in E_{\boldsymbol{\alpha}}} \|\mathbf{y} - \mathbf{x}_0\|_\infty$$

$\Rightarrow$ $\Delta(u) \in [\mathrm{diam}(E_{\boldsymbol{\alpha}})/2, \mathrm{diam}(E_{\boldsymbol{\alpha}})]$

$$\mathrm{diam}(E_{\boldsymbol{\alpha}}) = 1 + (d-2)\|\boldsymbol{\alpha}\|_\infty \in \left(2 - \tfrac{2}{d},\, d-1\right)$$

- $(d-1)$-balanced (Vuillon '03, see also Andrieu–Vivion '23)
- not factor-balanced for $d \geq 3$ (Bédaride–Berthé?)
- factor complexity $\Theta(n^{d-1})$ (Arnoux–Mauduit–Shiokawa–Tamura'94, Baryshnikov'95, Andrieu)
- natural coding of a rotation on $\mathbb{T}^{d-1}$ $\Rightarrow$ purely discrete spectrum

Letter balancedness and discrepancy

$\mathcal{L} \subset A^*$ *letter-C-balanced*:

$$\|\boldsymbol{\ell}(w) - \boldsymbol{\ell}(w')\|_\infty \le C \qquad \text{for all } w, w' \in \mathcal{L} \text{ with } |w| = |w'|$$

($|w|$ denotes the length of w)

letter-balanced: letter-C-balanced for some $C \ge 0$

Lemma 1

$\mathcal{L} \subset A^*$, $\boldsymbol{\alpha} \in [0,1]^d$, $\|\boldsymbol{\ell}(w) - |w|\boldsymbol{\alpha}\|_\infty \le C$ for all $w \in \mathcal{L}$
$\Rightarrow$ $\mathcal{L}$ letter-$(2C)$-balanced

in particular, $\Delta(u) \le C \Leftrightarrow \|\boldsymbol{\ell}(u_{[0,n)}) - n\boldsymbol{\alpha}\|_\infty \le C$ for all $n \ge 0$
$\Rightarrow \|\boldsymbol{\ell}(u_{[k,n)}) - (n-k)\boldsymbol{\alpha}\|_\infty \le 2C$ for all $n \ge k \ge 0$
$\Rightarrow$ language of u letter-$(4C)$-balanced

Lemma 2 (e.g. Poirier–St)

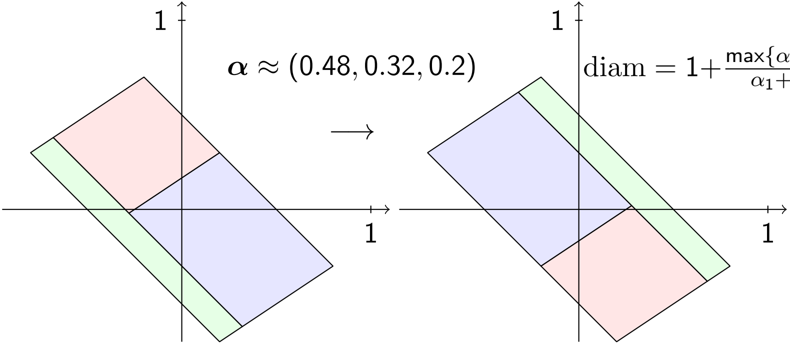
$\mathcal{L} \subset A^*$ infinite letter-C-balanced factorial language
$\Rightarrow \exists\, \boldsymbol{\alpha} \in [0,1]^d$: $\|\boldsymbol{\ell}(w) - |w|\boldsymbol{\alpha}\|_\infty \le C$ for all $w \in \mathcal{L}$

(factorial: if $w \in \mathcal{L}$, then all factors of w are in $\mathcal{L}$)

in particular, language of u letter-C-balanced $\Rightarrow \Delta(u) \le C$

2-balanced sequences with frequency vector $\boldsymbol{\alpha}$, $d \in \{3,4\}$ (Dvořáková–Mašek–Pelantová '23)

$\mathbf{u} \in \{1,3\}^{\mathbb{N}}$ Sturmian with slope $\alpha_1+\alpha_2$
$\mathbf{v} \in \{1,2\}^{\mathbb{N}}$ Sturmian with slope $\frac{\alpha_1}{\alpha_1+\alpha_2}$
$\mathbf{w} \in \{3,4\}^{\mathbb{N}}$ Sturmian with slope $\frac{\alpha_3}{\alpha_3+\alpha_4}$ (if $d=4$)
$d=3$: replace sequence of 1s in $\mathbf{u}$ by $\mathbf{v}$
$\Rightarrow$ 2-balanced sequence with frequency vector $\boldsymbol{\alpha}$



$\boldsymbol{\alpha} \approx (0.48, 0.32, 0.2)$ $\longrightarrow$ $\mathrm{diam} = 1 + \frac{\max\{\alpha_1,\alpha_2\}}{\alpha_1+\alpha_2} \in \left(\frac{3}{2}, 2\right)$

$d=4$: replace sequence of 1s in $\mathbf{u}$ by $\mathbf{v}$, 3s by $\mathbf{w}$
$\Rightarrow$ 2-balanced sequence, $\mathrm{diam} = 1 + \max\left\{\frac{\alpha_1}{\alpha_1+\alpha_2}, \frac{\alpha_2}{\alpha_1+\alpha_2}, \frac{\alpha_3}{\alpha_3+\alpha_4}, \frac{\alpha_4}{\alpha_3+\alpha_4}\right\}$

Tijdeman sequences

$C \geq 1 - \frac{1}{d}$,

$\hat{T}_{\boldsymbol{\alpha},C,C'} : \mathbf{1}^\perp \to \mathbf{1}^\perp \cap [-C,\infty)^d, \quad \mathbf{x} \mapsto \mathbf{x} + \boldsymbol{\alpha} - \mathbf{e}_j$

$$\text{if } x_j + \alpha_j - 1 \geq -C \text{ and } \frac{C' - x_j}{\alpha_j} = \min_{1 \leq i \leq d:\, x_i + \alpha_i - 1 \geq -C} \frac{C' - x_i}{\alpha_i}$$

(such j exists because $\sum_{i=1}^{d} (x_i + \alpha_i + C - 1) \geq 0$)

hypercubic billiards: $\hat{T}_{\boldsymbol{\alpha}} = \hat{T}_{\boldsymbol{\alpha},\infty,1}$

Theorem 3 (Tijdeman '80, Berthé–Carton–Chevallier–St–Yassawi)

$C, C' \geq 1 - \frac{1}{d}$, $C + C' \geq 2 - \frac{1+\min_{1 \leq i \leq d} \alpha_i}{d-1}$

$$\Rightarrow \quad \hat{T}^n_{\boldsymbol{\alpha},C,C'}(\mathbf{0}) \in [-C, C']^d \quad \textit{for all } n \geq 0$$

$(\boldsymbol{\alpha}, C, C', \mathbf{x}_0)$ *Tijdeman parameters* if $\hat{T}^n_{\boldsymbol{\alpha},C,C'}(\mathbf{x}_0) \in [-C, C']^d$ for all $n \geq 0$; we call natural codings of $\hat{T}_{\boldsymbol{\alpha},C,C'}$ *Tijdeman sequences*

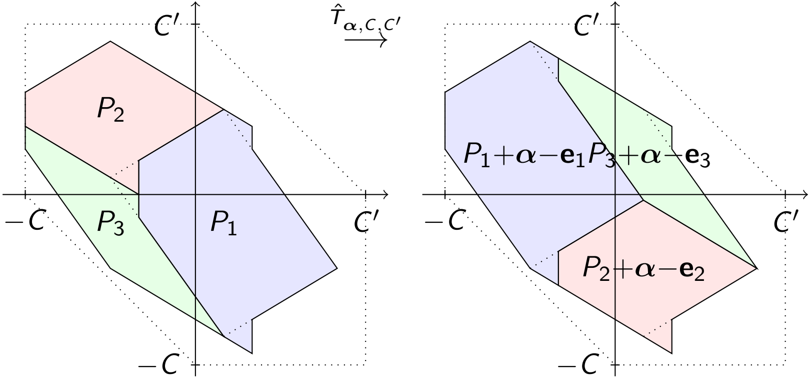
$C + C' < 2 \Rightarrow$ 3-balanced, $C + C' < 3/2$ $(d = 3) \Rightarrow$ 2-balanced

Theorem 4

For Tijdeman parameters $(\boldsymbol{\alpha}, C, C', \mathbf{0})$ *with* $C, C' < 1$,

$$P_{\boldsymbol{\alpha},C,C',i} = \overline{\{\hat{T}^n_{\boldsymbol{\alpha},C,C'}(\mathbf{0}) : n \geq 0\}}$$

are unions of convex polytopes,
$\mathbb{T}^{d-1} = \bigcup_{i=1}^{d} \iota(P_{\boldsymbol{\alpha},C,C',i})$ *(topological) partition*



$P_i = P_{\boldsymbol{\alpha},C,C',i}$ for $d = 3$, $\boldsymbol{\alpha} \approx (0.5, 0.3, 0.2)$, $C = C' = 3/4$

Factor-balancedness (or total balancedness)

language $\mathcal{L}$ over a finite alphabet A is

- *C-balanced w.r.t. $v \in A^*$*:

$$|w|_v - |w'|_v \leq C \qquad \text{for all } w, w' \in \mathcal{L} \text{ with } |w| = |w'|$$

- *balanced w.r.t. $v \in A^*$*: C-balanced w.r.t. $v \in A^*$ for some $C \geq 0$
- *(C-)balanced for length n*: (C-)balanced for all $v \in A^n$
- *letter-(C-)balanced*: (C-)balanced for length 1
- *factor-(C-)balanced*: (C-)balanced for all lengths $n \geq 1$

$$\text{factor-balanced} \iff \text{balanced w.r.t. all } v \in A^*$$
$$\not\Longleftrightarrow \text{factor-}C\text{-balanced for some } C \geq 1$$

factor-balancedness is also called *total balancedness* (Sadun '16)

Examples of (factor-)balanced languages

Morse–Hedlund '40:
language of a Sturmian word letter-1-balanced

Fagnot–Vuillon '02:
language of a Sturmian word factor-balanced, more precisely
$|v|$-balanced w.r.t. all v, factor-C-balanced for some $C \geq 1$ if
(and only if) slope has bounded partial quotients

language of Thue-Morse word letter-2-balanced
(if $|w| = 2n+1$, then $(|w|_0, |w|_1) \in \{(n, n{+}1), (n{+}1, n)\}$;
if $|w| = 2n$, then $(|w|_0, |w|_1) \in \{(n, n), (n{-}1, n{+}1), (n{+}1, n{-}1)\}$)

Berthé–Cecchi Bernales '19:
language of Thue-Morse word NOT balanced for length 2

Berthé–Cecchi Bernales–Durand–Leroy–Perrin–Petite '21:
For $\mathcal{S}$-adic languages defined by sequences of (left or right) proper morphisms with unimodular incidence matrices,
factor-balancedness is equivalent to letter-balancedness

Morphisms preserve balancedness

Lemma 5

$\mathcal{L} \subset A^*$ letter-balanced factorial language, $\sigma : A^* \to B^*$ morphism $\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ letter-balanced ($\mathcal{F}(w) = \{$*set of factors of* $w\}$)

Lemma 6

$\mathcal{L} \subset A^*$ factorial language, $\sigma : A^* \to B^*$ morphism with invertible incidence matrix M_σ, $\mathcal{F}(\sigma(\mathcal{L}))$ letter-balanced $\Rightarrow \mathcal{L}$ letter-balanced

(*incidence matrix* $M_\sigma = (|\sigma(b)|_a)_{a \in A, b \in B}$)

Proposition 7

$\mathcal{L} \subset A^*$ factorial language that is balanced for length n, $\sigma : A^* \to B^*$ morphism, $u \in B^*$ (possibly empty) prefix of $\sigma(a)u$ for all $a \in A$ (or suffix of $u\sigma(a)$ for all $a \in A$)
$\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ balanced for length $\min_{w \in A^{n-1} \cap \mathcal{L}} |\sigma(w)| + |u| + 1$

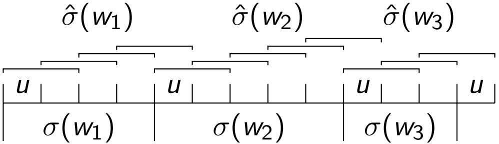
- *σ non-erasing $\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ balanced for length n*
- *σ left or right proper $\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ balanced for length $n+1$*

(σ is *left* (resp. *right*) *proper* if $\sigma(a)$ starts (resp. ends) with the same letter for all $a \in A$)

Proof: case $n = 1$, $\sigma(a)$ starts with $u \in B$ for all $a \in A$:

$$\text{morphism } \hat{\sigma} : A^* \to (B^2)^*,\ w \mapsto (\sigma(w)u)^{(2)}$$

m-th higher block code $(a_1 a_2 \cdots a_N)^{(m)} =$
$(a_1 \cdots a_m)(a_2 \cdots a_{m+1}) \cdots (a_{N-m+1} \cdots a_N) \in (A^m)^{N-m+1})$



$\mathcal{L}$ letter-balanced $\Rightarrow \mathcal{F}(\hat{\sigma}(\mathcal{L}))$ letter-balanced
$\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ balanced for length 2

$m \leq \min_{w \in A^{n-1} \cap \mathcal{L}} |\sigma(w)| + |u| + 1$, morphism
$\hat{\sigma} : (A^n \cap \mathcal{L})^* \to (B^m)^*,\ a_1 a_2 \cdots a_n \mapsto \big(\sigma(a_1)\mathrm{pref}_{m-1}(\sigma(a_2 \cdots a_n)u)\big)^{(m)}$
$\mathcal{L}$ balanced for length $n \Leftrightarrow \mathcal{L}^{(n)}$ letter-balanced
$\Rightarrow \mathcal{F}(\hat{\sigma}(\mathcal{L}^{(n)}))$ letter-balanced $\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ balanced for length m

Theorem 8

$\mathcal{L} \subset A^*$ factor-balanced factorial language, $\sigma : A^* \to B^*$ morphism
$\Rightarrow \mathcal{F}(\sigma(\mathcal{L}))$ factor-balanced

$\mathcal{S}$-adic languages and shifts

$\boldsymbol{\sigma} = (\sigma_n)_{n\geq 0}$ sequence of morphisms (or substitutions) $\sigma_n : A_{n+1}^* \to A_n^*$

$$A_0^* \xleftarrow{\sigma_0} A_1^* \xleftarrow{\sigma_1} A_2^* \xleftarrow{\sigma_2} \cdots \xleftarrow{\sigma_{n-1}} A_n^* \xleftarrow{\sigma_n} A_{n+1}^* \xleftarrow{\sigma_{n+1}} \cdots$$

alphabet rank $\mathrm{AR}(\boldsymbol{\sigma}) = \liminf_{n\to\infty} \# A_n$

$$\mathcal{L}_{\boldsymbol{\sigma}} = \{w \in A_0^* : w \text{ factor of } \sigma_{[0,n)}(A_n) \text{ for infinitely many } n \geq 0\}$$

$$\sigma_{[m,n)} = \sigma_m \circ \sigma_{m+1} \circ \cdots \circ \sigma_{n-1}$$

$$X_{\boldsymbol{\sigma}} = \{u \in A_0^{\mathbb{Z}} : \text{each factor of } u \text{ is in } \mathcal{L}_{\boldsymbol{\sigma}}\}$$

$\mathcal{L}_{\boldsymbol{\sigma}}^{(k)} = \{w \in A_k^* : w \text{ factor of } \sigma_{[k,n)}(A_n) \text{ for infinitely many } n \geq k\}$

$X_{\boldsymbol{\sigma}}^{(k)} = \{u \in A_k^{\mathbb{Z}} : \text{each factor of } u \text{ is in } \mathcal{L}_{\boldsymbol{\sigma}}^{(k)}\}$

$(\sigma_k)_{k\geq 0}$ is left (resp. right) *proper* if $\forall k \geq 0\ \exists n > k$ such that $\sigma_{[k,n)}$ is left (resp. right) proper

$(\sigma_k)_{k\geq 0}$ is *everywhere growing* if $\lim_{k\to\infty} \min_{a\in A_k} |\sigma_{[0,k)}(a)| = \infty$

Factor-balanced $\mathcal{S}$-adic and substitutive languages

Theorem 9

- *$\boldsymbol{\sigma}$ left or right proper sequence of morphisms, $\mathcal{L}_{\boldsymbol{\sigma}}^{(k)}$ letter-balanced for infinitely many $k \Rightarrow \mathcal{L}_{\boldsymbol{\sigma}}$ factor-balanced*
- *$\boldsymbol{\sigma} = (\sigma_k)_{k\geq 0}$ left or right proper, $\mathcal{L}_{\boldsymbol{\sigma}}$ letter-balanced, incidence matrices M_{σ_k} invertible $\Rightarrow \mathcal{L}_{\boldsymbol{\sigma}}$ factor-balanced*
- *$\boldsymbol{\sigma}$ everywhere growing, $\mathcal{L}_{\boldsymbol{\sigma}}^{(k)}$ is balanced for length 2 for infinitely many $k \Rightarrow \mathcal{L}_{\boldsymbol{\sigma}}$ factor-balanced*
- *$\sigma : A^* \to A^*$, σ^k left or right proper for some $k \geq 1$, $\mathcal{L}_{\sigma^\infty}$ letter-balanced $\Rightarrow \mathcal{L}_{\sigma^\infty}$ factor-balanced*
- *$\sigma : A^* \to A^*$ everywhere growing, $\mathcal{L}_{\sigma^\infty}$ balanced for length 2 $\Rightarrow \mathcal{L}_{\sigma^\infty}$ factor-balanced (cf. Queffélec '87, Adamczewski '03, '04)*

Letter-balancedness of (primitive) $\mathcal{S}$-adic languages

Necessary conditions:

$$\bigcap_{n\geq 0} M_{\sigma_{[0,n)}} \mathbb{R}_+^{\# A_n} = \mathbb{R}_+ \boldsymbol{\alpha}$$

$$\left\{\boldsymbol{\ell}(\sigma_{[0,n)}(i)) - |\sigma_{[0,n)}(i)|\, \boldsymbol{\alpha} \,:\, i \in A_n, n \geq 0\right\} \quad \text{bounded}$$

not always sufficient: see Cassaigne–Ferenczi–Zamboni '00 and Cassaigne–Ferenczi–Messaoudi '08 for Arnoux–Rauzy words, Delecroix–Hejda–St '13 for Brun words, Andrieu '18 for Cassaigne words

Factor complexity

Theorem 10 (Berthé–Carton–Chevallier–St-Yassawi)
natural coding of totally irrational translation on $\mathbb{T}^{d-1}$
w.r.t. unions of convex polytopes $\Rightarrow$ *factor complexity* $\Theta(n^{d-1})$

Theorem 11 (Espinoza '23, "S-adic conjecture")
Factor complexity of minimal subshift X *is* $O(n) \Leftrightarrow X = X_{\boldsymbol{\sigma}}$ *s.t.*

1. $\forall n \geq 1\ \exists\, W_n \subset A_n^*$ *with bounded* $\# W_n$*:*
   $\forall\, i \in A_n$*:* $\sigma_{[0,n)}(i) = w^k$ *for some* $w \in W_n$*,* $k \geq 1$
2. $\frac{|\sigma_{[0,n)}(i)|}{|\sigma_{[0,n)}(j)|}$ *bounded for* $i, j \in A_n$*,* $n \geq 0$
3. $|\sigma_n(i)|$ *bounded for* $i \in A_{n+1}$*,* $n \geq 0$

Recognisability

morphism $\sigma : A^* \to B^*$, left shift map S
$(u, k) \in A^{\mathbb{Z}} \times \mathbb{Z}$ *centered σ-representation* of $v \in B^{\mathbb{Z}}$ if

$$v = S^k \sigma(u), \quad 0 \le k < |\sigma(u_0)|$$

σ *fully recognisable* if each $v \in B^{\mathbb{Z}}$ has at most one centered σ-repr., *recognisable in $X \subseteq A^{\mathbb{Z}}$* if each $v \in B^{\mathbb{Z}}$ has at most one centered σ-representation (u, k) with $u \in X$, *for aperiodic points* if each $v \in B^{\mathbb{Z}}$ has at most one centered σ-repr. (u, k) with aperiodic u

$\sigma : A^* \to B^*$ *elementary* if
$\sigma = \alpha \circ \beta$, $\beta : A^* \to C^*$, $\alpha : C^* \to B^* \Rightarrow \#C \ge \#A$

Theorem 12 (Karhumäki–Manúch–Plandowski '03, see also Béal–Perrin–Restivo '23)

Elementary morphisms are fully recognisable for aperiodic points.

Theorem 13 (Mossé '92, Bezuglyi–Kwiatkowski–Medynets '09, Berthé–St–Thuswaldner–Yassawi '19, Béal–Perrin–Restivo '23)

Each $\sigma : A^* \to A^*$ is recognisable for aperiodic points in the substitutive shift X_{σ^∞}.

Representability and recognisability of $\mathcal{S}$-adic shifts

$\boldsymbol{\sigma} = (\sigma_n)_{n\geq 0}$ *recognisable at level n* if σ_n recognisable in $X_{\boldsymbol{\sigma}}^{(n+1)}$,
$\boldsymbol{\sigma}$ *recognisable* if recognisable at each level,
eventually recognisable if recognisable at all sufficiently large levels

$\boldsymbol{\sigma} = (\sigma_n)_{n\geq 0}$ *representable at level n* if $X_{\boldsymbol{\sigma}}^{(n)} = \bigcup_{k\in\mathbb{Z}} S^k \sigma_n\big(X_{\boldsymbol{\sigma}}^{(n+1)}\big)$
Holds when σ_n is non-erasing but not necessarily otherwise, e.g.:

$$\sigma_0 : \ a \mapsto a, \ b \mapsto \varepsilon, \ c \mapsto \varepsilon$$
$$\sigma_n : \ a \mapsto a, \ b \mapsto bb, \ c \mapsto cab \quad \text{for all } n \geq 1.$$

$$\sigma_{[1,n)}(a) = a, \ \sigma_{[1,n)}(b) = b^{2^{n-1}}, \ \sigma_{[1,n)}(c) = cabab^2 \cdots ab^{2^{n-2}}$$

$$X_{\boldsymbol{\sigma}}^{(1)} = \{{}^\infty b^\infty, {}^\infty bab^\infty\}, \quad X_{\boldsymbol{\sigma}}^{(0)} = \{{}^\infty a^\infty\} \neq \bigcup_{k\in\mathbb{Z}} S^k \sigma_0\big(X_{\boldsymbol{\sigma}}^{(1)}\big)$$

$\boldsymbol{\sigma}$ *representable* if representable at each level, *eventually representable* if $\boldsymbol{\sigma}$ representable for all sufficiently large levels

Representability and recognisability of $\mathcal{S}$-adic shifts

$\boldsymbol{\sigma} = (\sigma_n)_{n\geq 0}$ *recognisable at level n* if σ_n recognisable in $X_{\boldsymbol{\sigma}}^{(n+1)}$,
$\boldsymbol{\sigma}$ *recognisable* if recognisable at each level,
eventually recognisable if recognisable at all sufficiently large levels
$\boldsymbol{\sigma} = (\sigma_n)_{n\geq 0}$ *representable at level n* if $X_{\boldsymbol{\sigma}}^{(n)} = \bigcup_{k\in\mathbb{Z}} S^k \sigma_n\big(X_{\boldsymbol{\sigma}}^{(n+1)}\big)$

Theorem 14 (Béal–Perrin-Restivo–St)

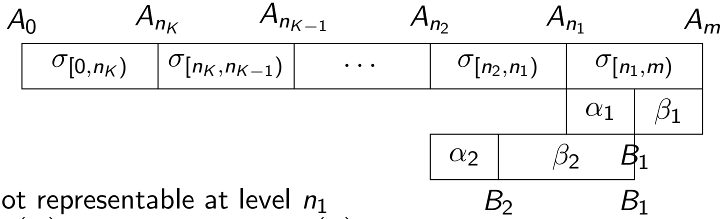
finite alphabet rank $\Rightarrow$ $\boldsymbol{\sigma}$ eventually recognisable for aperiodic points and eventually representable

number of levels at which $\boldsymbol{\sigma}$ not recognisable for aperiodic points at most $\mathrm{AR}(\boldsymbol{\sigma})-2$, number of levels at which $\boldsymbol{\sigma}$ not representable at most $\mathrm{AR}(\boldsymbol{\sigma})-1$

in particular, σ^∞ recognisable for aperiodic points and representable

Proof of recognisability and representability

$0 \le n_K < \cdots < n_2 < n_1 < m$ s.t. $\boldsymbol{\sigma}$ not representable or not rec. for aper. points at level n_k, $1 \le k \le K$, $\#A_m = \mathrm{AR}(\boldsymbol{\sigma})$



$\boldsymbol{\sigma}$ not representable at level n_1

$\Rightarrow X_{\boldsymbol{\sigma}}^{(n_1)} \neq \bigcup_{h\in\mathbb{Z}} S^h \sigma_{[n_1,m)}\big(X_{\boldsymbol{\sigma}}^{(m)}\big) \Rightarrow \sigma_{[n_1,m)}$ erasing $\Rightarrow \sigma_{[n_1,m)}$ not elem.

$\boldsymbol{\sigma}$ not recognisable f.a.p. at $n_1 \Rightarrow \sigma_{[n_1,m)}$ not recognisable f.a.p. in $X_{\boldsymbol{\sigma}}^{(m)}$ or $X_{\boldsymbol{\sigma}}^{(n_1+1)} \neq \bigcup_{h\in\mathbb{Z}} S^h \sigma_{[n_1+1,m)}\big(X_{\boldsymbol{\sigma}}^{(m)}\big)$ (thus $\sigma_{[n_1+1,m)}$ erasing)

all cases $\Rightarrow \sigma_{[n_1,m)}$ not elementary $\Rightarrow \sigma_{[n_1,m)} = \alpha_1 \circ \beta_1$, $\#B_1 < \#A_m$

$\boldsymbol{\sigma}$ not representable or not recognisable f.a.p. at n_2

$\Rightarrow \sigma_{[n_2,n_1)} \circ \alpha_1$ not elementary $\Rightarrow \sigma_{[n_2,n_1)} \circ \alpha_1 = \alpha_2 \circ \beta_2$, $\#B_2 < \#B_1$

$\cdots \Rightarrow \sigma_{[n_K,n_{K-1})} \circ \alpha_{K-1}$ not elementary $\Rightarrow \sigma_{[n_K,n_{K-1})} \circ \alpha_{K-1} = \alpha_K \circ \beta_K$

$\#B_K < \#B_{K-1} < \cdots < \#B_2 < \#B_1 < \#A_m = \mathrm{AR}(\boldsymbol{\sigma})$

Examples of non-recognisability

$K = \mathrm{AR}(\boldsymbol{\sigma}) - 2$ levels of non-recognisability f.a.p.,
non-recognisability f.a.p. at all levels with $\#A_n = n+2$ ($K = \infty$):
$A_n = \{a_0, \ldots, a_{n+1}\}$ for $0 \le n < K$, $A_n = \{a_0, \ldots, a_{K+1}\}$ for $n \ge K$,

$$\sigma_n : a_i \mapsto a_0 a_i a_0,\ 0 \le i \le n+1,\ a_{n+2} \mapsto a_{n+1}, \quad \text{if } 0 \le n < K$$

$$\sigma_n : a_i \mapsto a_0 a_i a_0,\ 0 \le i \le K+1, \quad \text{if } n \ge K$$

$$\sigma_n({}^\infty a_0 a_{n+1} a_0^\infty) = {}^\infty a_0 a_{n+1} a_0^\infty = \sigma_n({}^\infty a_0 a_{n+2} a_0^\infty), \quad 0 \le n < K$$

primitive, injective, proper example of non-recognisability:

$$\sigma_0 : a \mapsto aa,\ b \mapsto ab,\ c \mapsto ba$$

$$\sigma_n : a \mapsto acaba,\ b \mapsto acaaa\ c \mapsto aaaba, \quad n \ge 1$$

$$\tau_n : a \mapsto aabaa,\ b \mapsto aaaaa, \quad n \ge 0$$

$$\sigma_0 \circ \sigma_1 = \tau_0 \circ \sigma_0 \ \Rightarrow\ \sigma_{[0,n+1)} = \tau_{[0,n)} \circ \sigma_0$$

σ_0 is of constant length 2, σ_1 is of odd constant length and proper
$\Rightarrow \sigma_0(X_{\boldsymbol{\sigma}}^{(1)}) = S\sigma_0(X_{\boldsymbol{\sigma}}^{(1)}) = X_{\boldsymbol{\sigma}}^{(0)}$
can be extended to non-recognisability at all levels with exponentional alphabet growth

other examples of non-recognisability: Bedaride–Hilion–Lustig '22

Theorem 15 (Donoso–Durand–Maass–Petite '21)

all $X_{\boldsymbol{\sigma}}^{(n)}$ minimal, all σ_n injective (on letters)
$\Rightarrow$ at most $\log_2 \mathrm{AR}(\boldsymbol{\sigma})$ levels of non-recognisability f.a.p.

Theorem 16 (Donoso–Durand–Maass–Petite '21)

An expansive finite topological rank minimal Cantor system is topologically conjugate to a primitive recognisable $\mathcal{S}$-adic subshift with finite alphabet rank. Moreover, the alphabet rank of the corresponding directive sequence is bounded by the rank of the minimal Cantor system.
Any primitive $\mathcal{S}$-adic subshift (so minimal) with bounded alphabet rank is a topological factor of a finite topological rank minimal Cantor system whose rank is bounded by the square of the alphabet rank of the subshift. Moreover, when the $\mathcal{S}$-adic subshift is recognisable then the factor map is invertible and the subshift has finite topological rank.

(recognisable, left, right proper $\Rightarrow$ conjugate to Bratteli–Vershik system)

Non-unimodular $\mathcal{S}$-adic Rauzy fractals

$\boldsymbol{\sigma} = (\sigma_n)_{n\ge 0}$ primitive, $\sigma_n : A^* \to A^*$, $A = \{1,2,\ldots,d\}$
M_{σ_n} invertible, $\bigcap_{n\ge 0} M_{\sigma_{[0,n)}} \mathbb{R}_+^d = \mathbb{R}_+ \boldsymbol{\alpha}$

$$\mathcal{R}_{\boldsymbol{\sigma}}(i) = \overline{\{\pi_{\boldsymbol{\sigma}}(\ell(w)) : wi \text{ prefix of } \sigma_{[0,n)}(i) \text{ for infinitely many } n, i\}}$$

$$\pi_{\boldsymbol{\sigma}} : \mathbb{Z}^d \to \mathbf{1}^\perp \times \mathbb{Z}_{\boldsymbol{\sigma}}, \quad \mathbf{x} \mapsto (\pi_{\boldsymbol{\alpha}}(\mathbf{x}), (\mathbf{x} + M_{\sigma_{[0,n)}}\mathbb{Z}^d)_{n\ge 0})$$

$$\mathbb{Z}_{\boldsymbol{\sigma}} = \varprojlim \mathbb{Z}^d / M_{\sigma_{[0,n)}}\mathbb{Z}^d$$

$$= \{(\mathbf{x}_n + M_{\sigma_{[0,n)}}\mathbb{Z}^d)_{n\ge 0} : \mathbf{x}_{n+1} \in \mathbf{x}_n + M_{\sigma_{[0,n)}}\mathbb{Z}^d \text{ for all } n \ge 0\}$$

$$\simeq \left\{ \sum_{n=0}^{\infty} M_{\sigma_{[0,n)}} \mathbf{d}_n : \mathbf{d}_n \in D_n \right\}, \quad \text{with } M_{\sigma_{[0,n)}}\mathbb{Z}^d \to \{\mathbf{0}\}$$

D_n complete residue system of $\mathbb{Z}^d / M_{\sigma_n}\mathbb{Z}^d$, i.e., $\#D_n = |\det M_{\sigma_n}|$

under similar conditions as for unimodular substitutions
(see Berthé–St-Thuswaldner '19,'22, Pytheas Fogg–Noûs '20):
$\bigcup_{i=1}^d \mathcal{R}_{\boldsymbol{\sigma}}(i)$ partition of $(\mathbf{1}^\perp \times \mathbb{Z}_{\boldsymbol{\sigma}})/\pi_{\boldsymbol{\sigma}}(\mathbb{Z}^d \cap \mathbf{1}^\perp)$
$X_{\boldsymbol{\sigma}}$ natural coding of translation by $\pi_{\boldsymbol{\sigma}}(\mathbf{e}_i)$ w.r.t. this partition

N-continued fractions, $N \geq 1$ (Burger–Gell-Redman–Kravitz–Walton–Yates '08)

$$T_N : [0,1) \to [0,1), \quad x \mapsto \frac{N}{x} - \left\lfloor \frac{N}{x} \right\rfloor, \text{ i.e., } x = \frac{N}{\lfloor N/x \rfloor + T_N(x)}$$

$$\begin{pmatrix} 1 \\ x \end{pmatrix} = \frac{1}{\lfloor N/x \rfloor + T_N(x)} \begin{pmatrix} \lfloor N/x \rfloor & 1 \\ N & 0 \end{pmatrix} \begin{pmatrix} 1 \\ T_N(x) \end{pmatrix}$$

$$\tilde{T}_N : \left[0, \frac{1}{N}\right) \to \left[0, \frac{1}{N}\right), \ x \mapsto \frac{1}{N}\left(\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor\right), \text{ i.e., } x = \frac{1}{\lfloor 1/x \rfloor + N \tilde{T}_N(x)}$$

$$\tilde{T}_N\left(\frac{x}{N}\right) = \frac{1}{N} T_N(x)$$

$$\begin{pmatrix} 1 \\ x \end{pmatrix} = \frac{1}{\lfloor 1/x \rfloor + N \tilde{T}_N(x)} \begin{pmatrix} \lfloor 1/x \rfloor & N \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \tilde{T}_N(x) \end{pmatrix}$$

$$\sigma_n : 1 \mapsto 1^{\lfloor 1/\tilde{T}_N^n(x) \rfloor} 2 \qquad \lfloor 1/\tilde{T}_N^n(x) \rfloor \geq N$$

$$2 \mapsto 1^N$$

Generalised continued fractions

$$x = \cfrac{1}{a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{a_2 + \ddots}}} \qquad \sigma_n: \begin{array}{l} 1 \mapsto 1^{a_n} 2 \\ 2 \mapsto 1^{b_n} \end{array} \qquad \bigcap_{n \ge 0} M_{\sigma_{[0,n)}} \mathbb{R}_+^2 = \mathbb{R}_+ \begin{pmatrix} 1 \\ x \end{pmatrix}$$

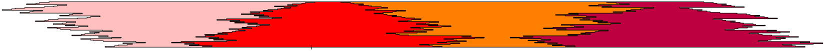
converges if $a_{n+1} \ge b_n$, σ_n is a β-substitution, $\beta^2 = a_n \beta + b_n$, if $a_n \ge b_n$

Theorem 17 (Langeveld–Rossi–Thuswaldner '23)

$X_{\boldsymbol{\sigma}}$ is letter-$\frac{2K-N}{K+1-N}$-balanced if $a_n \ge K \ge N = b_n$ for all $n \ge 0$
factor complexity $\le 2n$

Theorem 18 (Minervino–St '14)

$a_n \ge b_n$ constant: $\mathcal{R}_{\boldsymbol{\sigma}}(1) \cup \mathcal{R}_{\boldsymbol{\sigma}}(2)$ partition of $(\mathbf{1}^\perp \times \mathbb{Z}_{\boldsymbol{\sigma}})/\pi_{\boldsymbol{\sigma}}(\mathbb{Z}^d \cap \mathbf{1}^\perp)$



picture of $\pi_{\boldsymbol{\sigma}}(\mathbf{x}) + (\mathcal{R}_{\boldsymbol{\sigma}}(1) \cup \mathcal{R}_{\boldsymbol{\sigma}}(2))$, $\mathbf{x} \in \mathbb{Z}^2 \cap \mathbf{1}^\perp$, $a_n = b_n = 2$

can be generalised to arbitrary $(a_n)_{n \ge 0}$, $(b_n)_{n \ge 0}$, $a_{n+1} \ge b_n$