# Preservation theorems for transducer outputs

Valérie Berthé, Herman Goulet-Ouellet, Toghrul Karimov,
Dominique Perrin, **Mihir Vahanwala**

**Research School-Conference of Combinatorics, Automata, and Number Theory 2025**
**CIRM, Marseille**

# Combinatorial virtues feat. the Fibonacci word

$$0100101001001\cdots$$

Morphic: $x_{\mathsf{fib}} = \sigma_{\mathsf{fib}}(x_{\mathsf{fib}})$, where $\sigma_{\mathsf{fib}}(0) = 01$, $\sigma_{\mathsf{fib}}(1) = 0$

Recurrent: every factor occurs infinitely often

Furthermore, it is linearly, and hence uniformly recurrent

(linearly) bounded gaps between occurrences of *u*

Well-defined factor frequencies

For every $u$, $\lim_{N\to\infty} \frac{1}{N} |\{i : i < N, x[i,\infty) \in u\Sigma^{\omega}\}|$ is well-defined

Suppose a word $x \in \Sigma^\omega$ is

Morphic

(Linearly, Uniformly) Recurrent

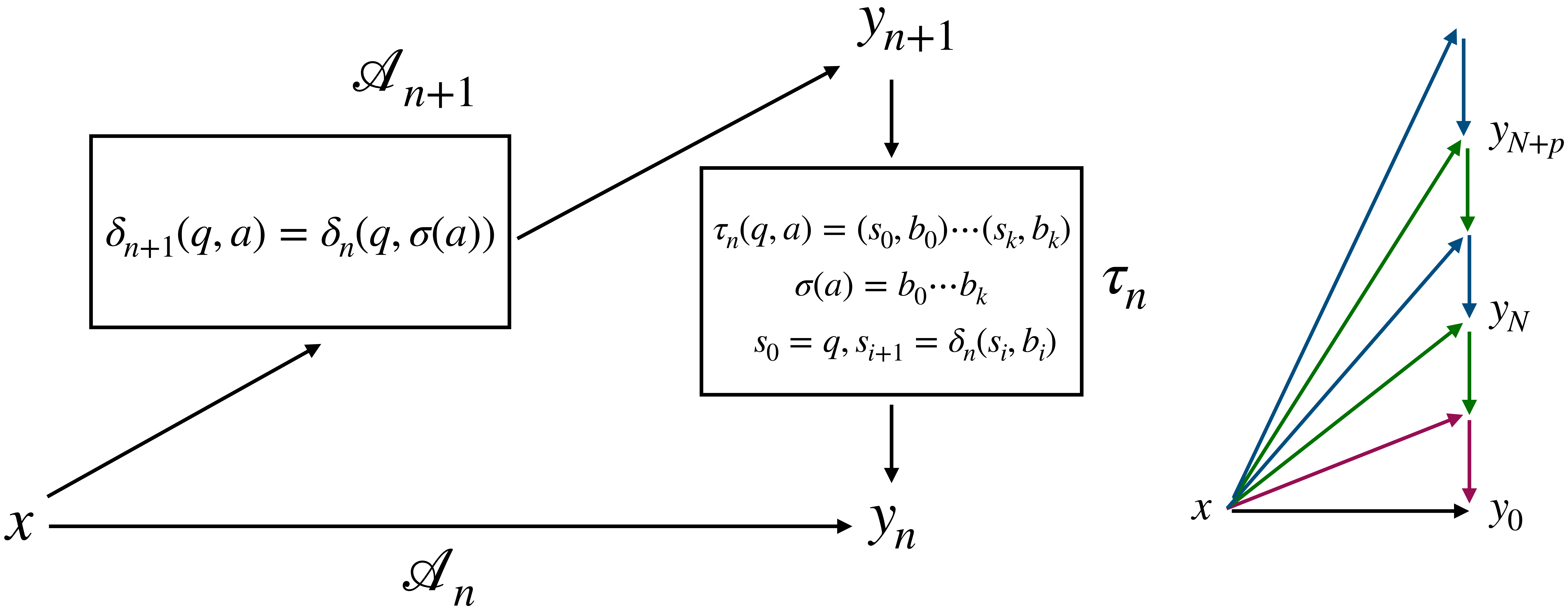
Well-defined factor frequencies

We feed $x$ to an automaton $\mathscr{A} = (Q, q_{\text{init}}, \Sigma, \delta)$ to get run $y \in (Q \times \Sigma)^\omega$

**Are these properties preserved?**

# The run of a substitutive word is morphic

$x \in \Sigma^\omega,\ x = \sigma(x)$ $\qquad$ $(Q, q_{\text{init}}, \Sigma, \delta_0) = \mathscr{A}_0$




$$y_{N+p} = y_N, y_0 = \tau_0 \cdots \tau_{N-1}(y_N), y_N = \tau_N \cdots \tau_{N+p-1}(y_N + p)$$

Suppose a word $x \in \Sigma^\omega$ is

Morphic ✓

(Linearly, Uniformly) Recurrent

Well-defined factor frequencies

Does the run $y \in (Q \times \Sigma)^\omega$ of $\mathscr{A}$ on $x$ have these properties?

**Theorem (Semënov et al., Pritykin).**

If $x$ is uniformly (resp. linearly) recurrent, then
$y$ has a suffix that is uniformly (resp. linearly) recurrent

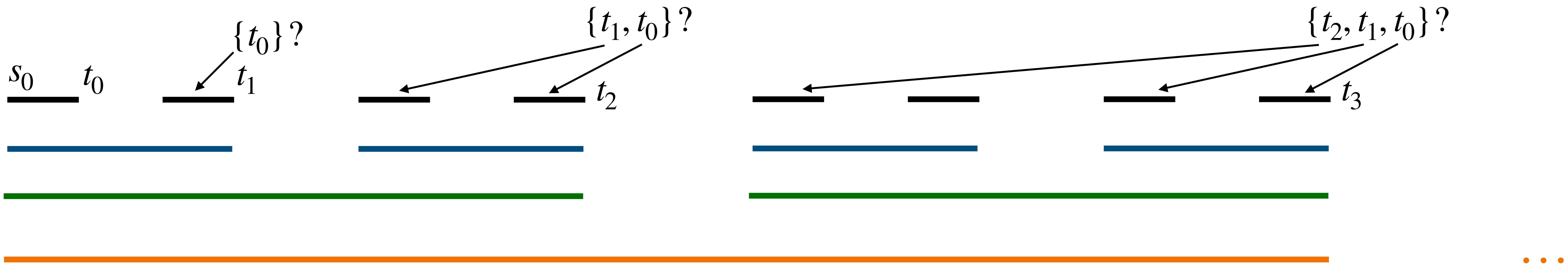
**We can further prove:**

- The corresponding statement holds for recurrent words
- Furthermore, we can effectively identify the suffix

# A flavour of the Semënov argument

Assume that $\mathscr{A}$ is a permutation automaton, i.e.,
each letter induces a permutation of $Q$ (the transitions are reversible)

If $x$ is recurrent, so is the run $y$: any occurrence of $v$ in $y$ has another to its right
Suppose $v$ is the run on $u_0$ starting in $s_0$ and ending in $t_0$




There are finitely many states, this process saturates in at most $|Q| + 1$ iterations

We can also argue that uniform and linear recurrence are preserved ■

# The other special case: no reversibility

Let $\mathscr{A}$ be a reset automaton, i.e.,
the transition function induced by each letter is either the identity or constant

do nothing reset to a state

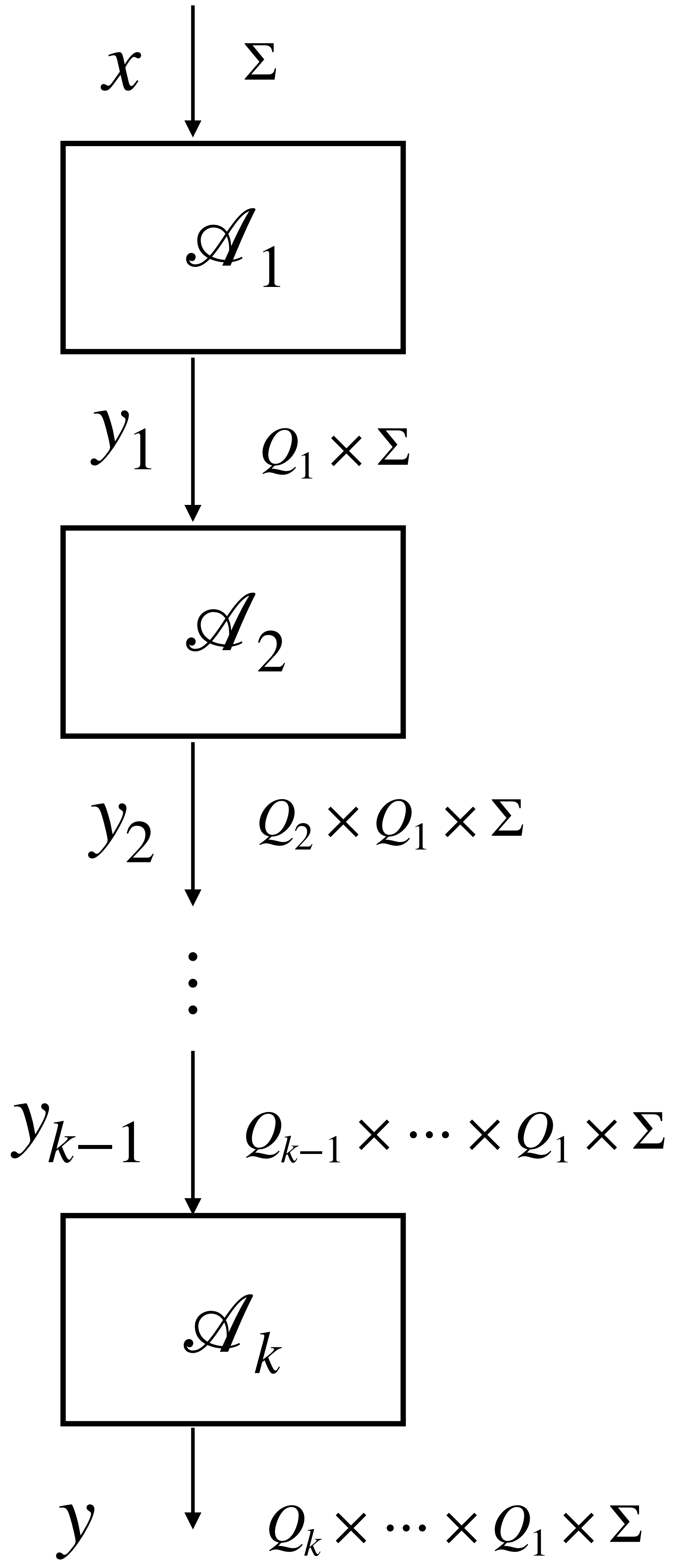
Let $x$ be recurrent and let the first reset letter occur at position $N$

The run $y$ of $\mathscr{A}$ on $x$ is recurrent starting at position $N+1$


The linear and uniform recurrence of $y$, as applicable, can readily be argued ■

# Preservation of recurrence holds for cascades of special automata

permutation automata, reset automata




$$Q = Q_k \times \cdots \times Q_1$$

$$\delta((q^k, \ldots, q^1), a)$$

$$=$$

$$\Big( \delta_k\big(q^k, \ (q^{k-1}, \ldots, q^1, a) \ \big),$$

$$\ldots,$$

$$\delta_j\big(q^j, \ (q^{j-1}, \ldots, q^1, a) \ \big),$$

$$\ldots,$$

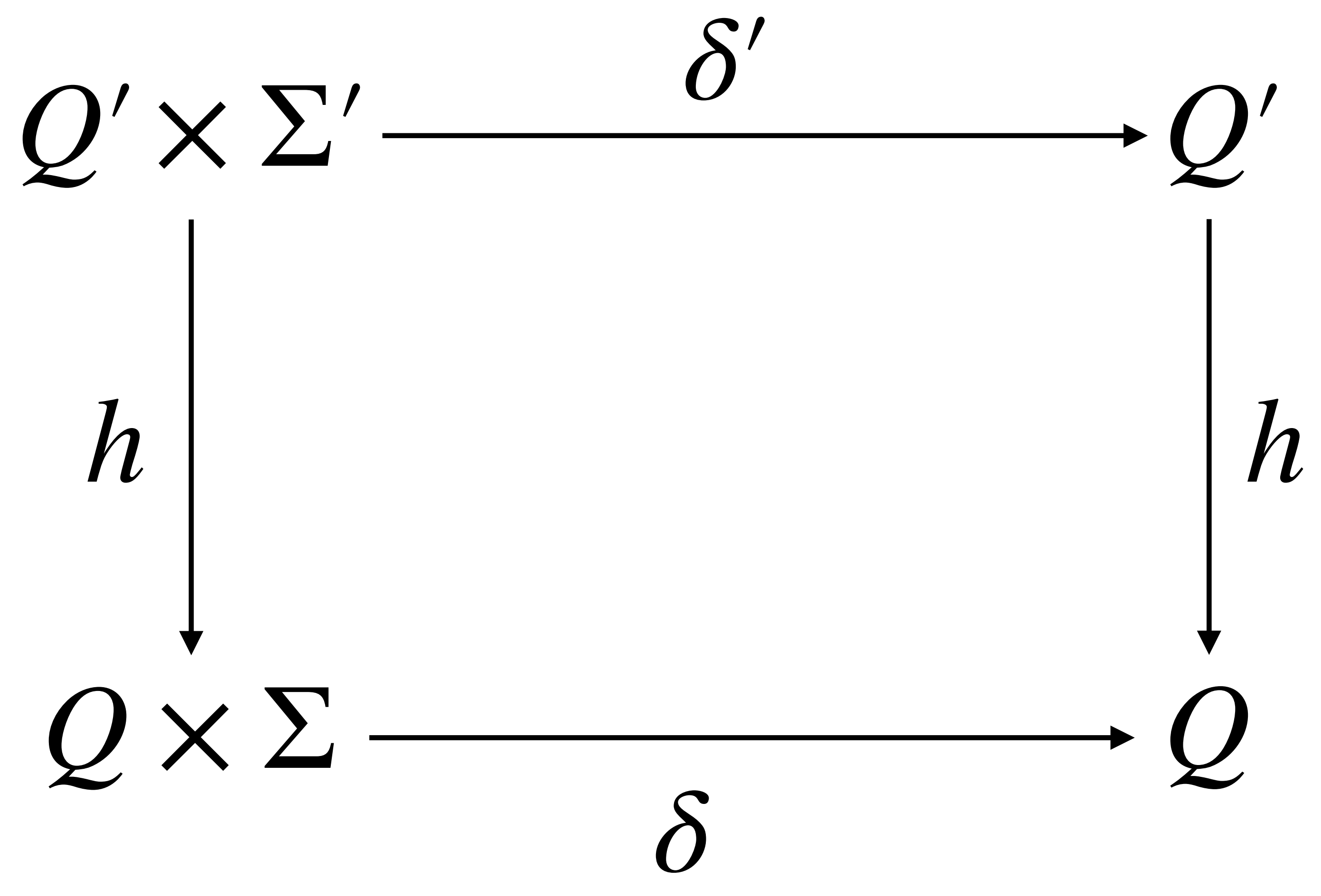
$$\delta_1\big(q^1, a\big) \ \Big)$$

# Preservation of recurrence holds for cascades of special automata

permutation automata, reset automata

## Theorem (Krohn-Rhodes).

Every automaton can be simulated by a cascade of special automata




commutative diagram of primed automaton simulating unprimed automaton through morphism $h$

## Preservation of recurrence holds for all automata ■

Suppose a word $x \in \Sigma^\omega$ has some of the following properties

Morphic ✓

(Linearly, Uniformly) Recurrent ✓

Well-defined factor frequencies

Yes for reset, and hence counter-free automata[1]

Does the run $y \in (Q \times \Sigma)^\omega$ of $\mathscr{A}$ on $x$ have these properties?

[1] A corollary of the Krohn-Rhodes theorem is that any counter-free automaton can be simulated by a cascade of resets.

# The other special case: no reversibility

Let $\mathscr{A}$ be a reset automaton, i.e.,
the transition function induced by each letter is either the identity or constant
do nothing reset to a state

Let $x$ be recurrent and let the first reset letter occur at position $N$

The run $y$ of $\mathscr{A}$ on $x$ is recurrent starting at position $N+1$


The linear and uniform recurrence of $y$, as applicable, can readily be argued ■

Suppose a word $x \in \Sigma^\omega$ has some of the following properties

Morphic ✓

(Linearly, Uniformly) Recurrent ✓

Well-defined factor frequencies

Yes for reset, and hence counter-free automata[1]

Not necessarily for permutation automata

Yes, if Boshernitzan's condition holds

E.g., primitive morphic words, linearly recurrent words, Sturmian words, almost all IETs, a large class of 1D-toric words, Arnoux-Rauzy words with bounded partial quotients, ...

Does the run $y \in (Q \times \Sigma)^\omega$ of $\mathscr{A}$ on $x$ have these properties?

Primitive morphic ⟶ Linearly recurrent ⟶ Boshernitzan ⟶ Factor frequencies

[1] A corollary of the Krohn-Rhodes theorem is that any counter-free automaton can be simulated by a cascade of resets.

$$0100101001001\cdots$$

The first 13 letters of the Fibonacci word

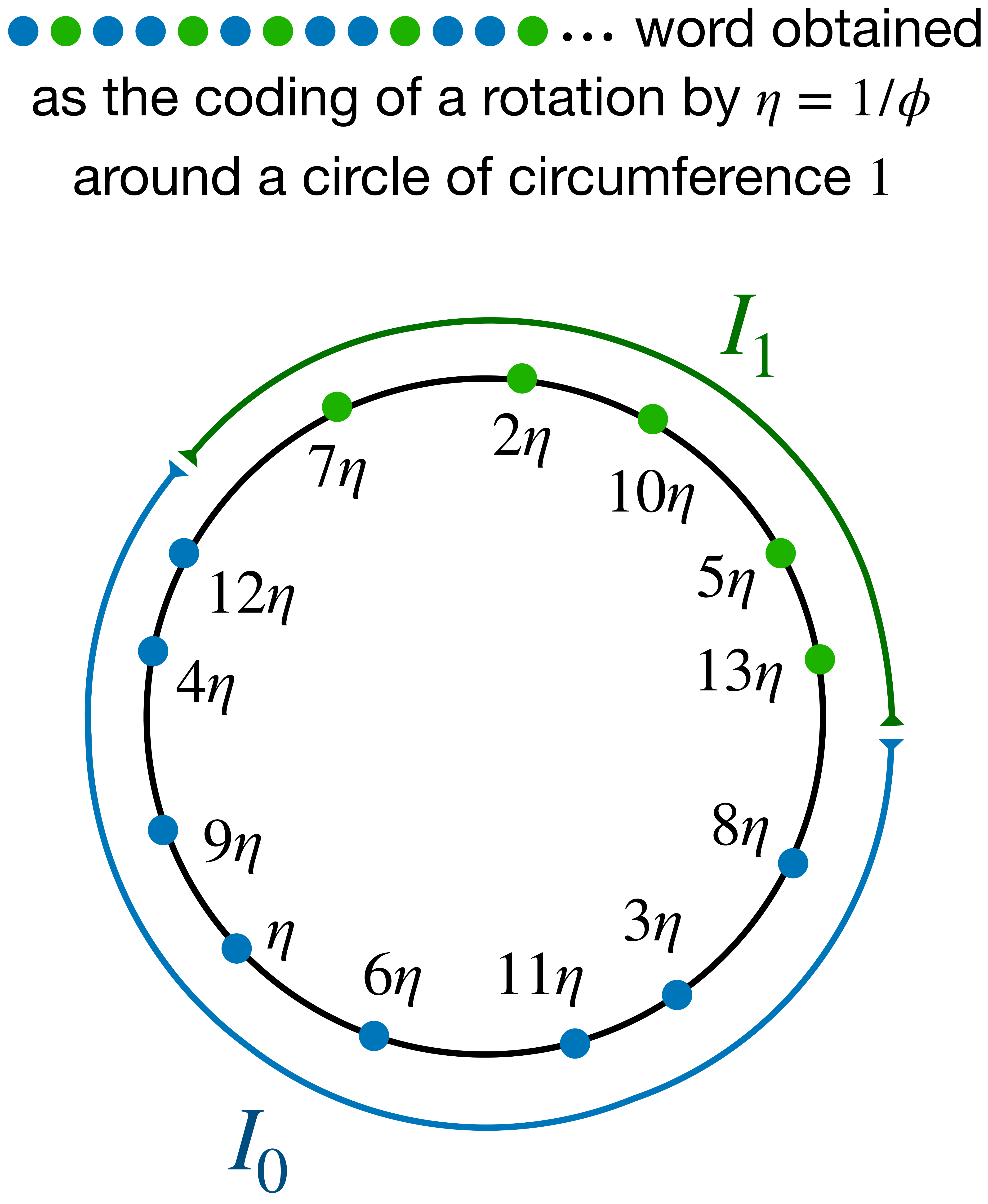
$0:8$ occurrences, frequency $0.615$

$1:5$ occurrences, frequency $0.385$

In the limit,

$0$ occurs with frequency $1/\phi \approx 0.618$

$1$ occurs with frequency $1/\phi^2 \approx 0.382$




$$I_{au} = I_a \cap (I_u - \eta)$$

frequency of factor $u$

=

length of arc $I_u$

**Trace**
Fibonacci word

generated by

**Topological Dynamical System**
Rotation

**Observation**
factor $u$

attributed to

**Causing Region**
arc $I_u$

The punchline, when ergodic theory applies, is:

**temporal frequency of observation**
**equals**
**spatial measure of causing region**

**temporal frequency of observation**
**equals**
**spatial measure of causing region**

# The "default" topological dynamical system

trace $x \in \Sigma^\omega$

endow $\Sigma^\omega$ with the product topology

shift operator $T : \Sigma^\omega \to \Sigma^\omega$, $T(a_0 a_1 a_2 \cdots) = a_1 a_2 \cdots$

Shift (space) $X \subseteq \Sigma^\omega$ : closure of orbit $(x, Tx, T^2 x, \ldots)$

**When and how does this realise the ergodic dream?**

**Trace**

$x \in \Sigma^\omega$

induces →

← generates

**Topological Dynamical System**

shift $X \subseteq \Sigma^\omega$, operator $T$

**Observation**

factor $u$

attributed to

**Causing Region**

cylinder $[u] \subseteq X$

$[u] = X \cap u\Sigma^\omega$

**Temporal frequency**

$\lim_{N\to\infty} \frac{1}{N} \left| \{ i : i < N, x[i, \infty) \in u\Sigma^\omega \} \right|$

equals

**Spatial measure**

$\mu([u])$

We want sensible, invariant measures, i.e., $\mu(L) = \mu(T^{-1}L)$

We consider shifts $X$ that have a unique invariant measure
In such shifts, all words have well-defined factor frequencies
(Oxtoby's theorem)

Suppose a word $x \in \Sigma^\omega$ has some of the following properties

Morphic ✓

(Linearly, Uniformly) Recurrent ✓

Well-defined factor frequencies

Yes for reset, and hence counter-free automata

Not necessarily for permutation automata

Yes, if Boshernitzan's condition holds

E.g., primitive morphic words, linearly recurrent words, Sturmian words, almost all IETs, a large class of 1D-toric words, Arnoux-Rauzy words with bounded partial quotients, ...

Does the run $y \in (Q \times \Sigma)^\omega$ of $\mathscr{A}$ on $x$ have these properties?

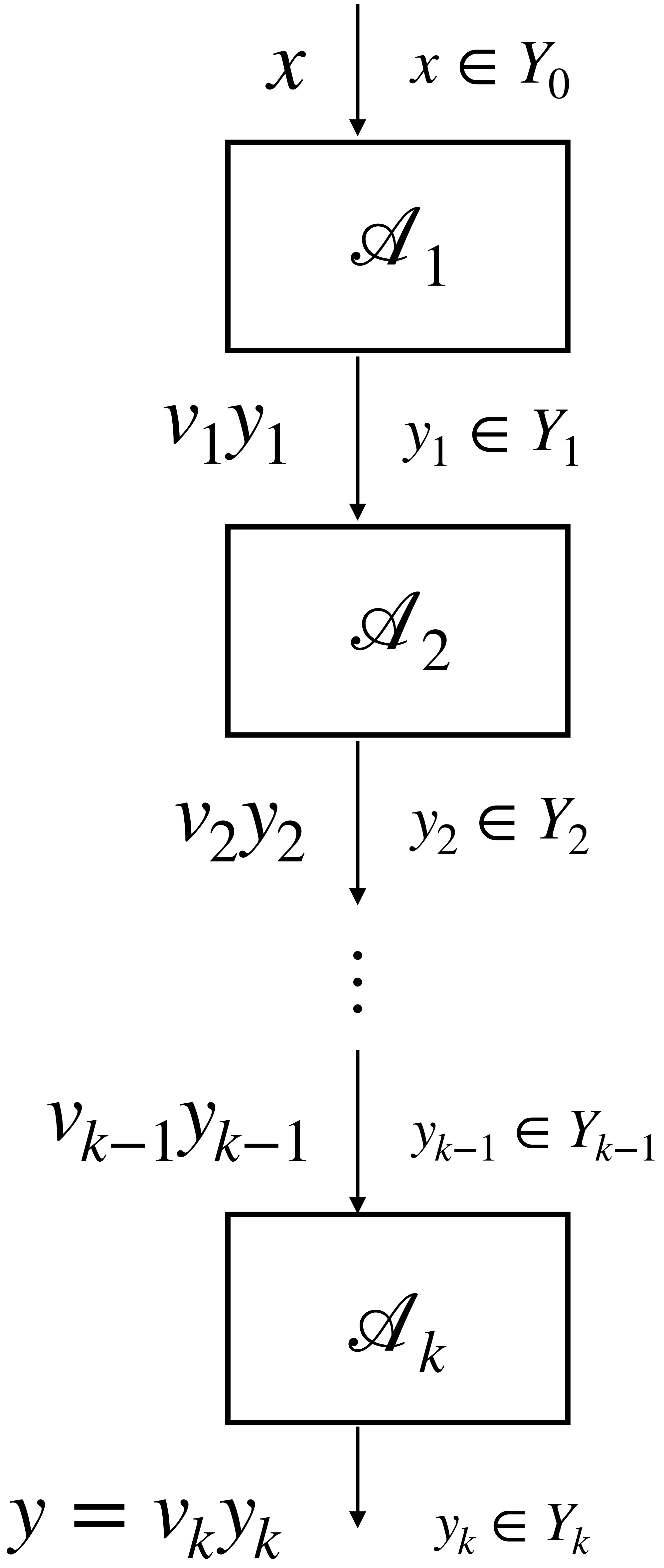
Primitive morphic ⟶ Linearly recurrent ⟶ Boshernitzan ⟶ Factor frequencies

# Boshernitzan's condition

If a minimal shift $X$ has an invariant measure $\mu$ such that

$$\liminf_{n\to\infty}\left(n\cdot\min_{\{u:|u|=n\}}\mu([u])\right)>0$$

then $\mu$ is the unique invariant measure of $X$




**Lemma**

If $Y_i$ satisfies Boshernitzan's condition then so does $Y_{i+1}$

**Corollary**

If $Y_0$ satisfies Boshernitzan's condition then $x, y$ have well-defined factor frequencies

**Bonus**

If $x$ is primitive morphic with computable factor frequencies then so is $y_k$

Suppose a word $x \in \Sigma^\omega$ has some of the following properties

Morphic ✓

(Linearly, Uniformly) Recurrent ✓

Well-defined factor frequencies

Yes for reset, and hence counter-free automata

Not necessarily for permutation automata

Yes, if Boshernitzan's condition holds

E.g., primitive morphic words, linearly recurrent words, Sturmian words, almost all IETs, a large class of 1D-toric words, Arnoux-Rauzy words with bounded partial quotients, ...

Does the run $y \in (Q \times \Sigma)^\omega$ of $\mathscr{A}$ on $x$ have these properties?

**Thank You!**

# Extra technical slides

Consider $x \in X$ and permutation automaton $\mathscr{A}$

The run $y$ is obtained from the the skew product $G \rtimes X$

The group $G$ is the transition group of $\mathscr{A}$, the morphism $\varphi : \Sigma \to G$

The dynamical transition of $G \rtimes X$ is given by $T$

$$T \circ (g, ax) = (g \cdot \varphi(a), x)$$

Let $v$ be the run on $u$ starting in $q$, let $C$ be the coset that maps $q_{\text{init}}$ to $q$

We hope that the frequency of $v$ is the measure of $(C, [u])$

But is $G \rtimes X$ even minimal?

Does its relevant minimal component have a unique invariant measure?

But is $G \rtimes X$ even minimal?

Key construct: cobounding map $\alpha$ from $X$ to set $\mathscr{C}$ of cosets of $H \leq G$

By definition, $\alpha$ is required to be continuous

$X$ is partitioned into finitely many (clopen) cylinders on which $\alpha$ is constant

The argument uses the minimality and compactness of $X$

Define $Y_\alpha = \{(g, x) : g \in \alpha(x)\} \subseteq G \rtimes X$

The point: if $H$ is a minimal subgroup for which a cobounding map exists then $Y_\alpha$ is minimal

Note, however, that there are $|G|/|H|$ choices of $Y_\alpha$, we make the one aligning with the run of $\mathscr{A}$

Does $Y_\alpha$ have a unique invariant measure?

Let $H$ be the minimal subgroup for which cobounding $\alpha$ exists

Let $\mathscr{C}$ be the set of right cosets of $H$

Recall $Y_\alpha = \{(g, x) : g \in \alpha(x)\}$. Does $Y_\alpha$ have a unique invariant measure?

Consider $Z_\alpha = \{(C, x) : \alpha(x) = C\} \subseteq \mathscr{C} \rtimes X$

$Z_\alpha$ is isomorphic to $X$, and in particular has a unique invariant measure

To define the invariant measure on $Z_\alpha$, recall that $X$ is partitioned into cylinders on which $\alpha$ is constant

To determine whether a cylinder exists in $Z_\alpha$ at all, we can use Semënov's theorem

An invariant measure for $Y_\alpha$ : $\mu_Y(g, L) = \dfrac{1}{|H|}\mu_Z(Hg, L)$

Intuitively, this distributes mass equally to all elements of the coset

Could this be *the* invariant measure for $Y$?

An invariant measure for $Y_\alpha : \mu_Y(g, [u]) = \frac{1}{|H|}\mu_Z(Hg, [u]) = \frac{1}{|H|}\mu_X([u])$

For $u$ long enough

Intuitively, this distributes mass equally to all elements of the coset

Could this be *the* invariant measure for $Y$?

Suppose $\mu_X$ is the unique invariant measure by virtue of satisfying

$$\liminf_{n\to\infty} \min_{|u|=n} n\mu_X([u]) > 0,$$

the Boshernitzan condition which is sufficient for unique ergodicity

Then, by inspection, $\mu_Y$ also satisfies the Boshernitzan condition
and is indeed the unique invariant measure of $Y_\alpha$

Let $\mathscr{A}$ be a permutation automaton and let $x$ be such that its shift $X$ has a unique invariant measure that satisfies the Boshernitzan condition

The run $y$ has a shift $Y \cong Y_\alpha$ which also has a unique invariant measure that satisfies the Boshernitzan condition, and hence $y$ has well-defined factor frequencies

We can also show that the Boshernitzan condition is preserved by reset automata

modulo a finite prefix of the run

In the Boshernitzan case, we can then apply the Krohn-Rhodes theorem and establish that the run inherits the property of well-defined factor frequencies