Discrete Derivatives of Recognizable Series

SDA2 2026, Lyon, France

Nicolas Ollinger

LIFO, Université d'Orléans

nicolas.ollinger@univ-orleans.fr

Today's question

When does an $\mathcal{S}$-regular sequence have a **recognizable support**?

Substitution 101

$$\tau: \quad a \mapsto ab, \quad b \mapsto cb, \quad c \mapsto a$$

Fixed point $\mathbf{e} = abcbacbabacbabcbabacbabcbacbabcbabacbab\cdots$

Factor complexity $p(n)$ = number of distinct factors of length n.

Goal Compute $p(n)$. Is it $\mathcal{S}_\tau$-regular? $\mathcal{S}_\tau$-synchronized?

Substitution 101

$$\tau: \quad a \mapsto ab, \quad b \mapsto cb, \quad c \mapsto a$$

Fixed point $\mathbf{e} = abcbacbabacbabcbabacbabcbacbabcbabacbab\cdots$

Factor complexity $p(n)$ = number of distinct factors of length n.

Goal Compute $p(n)$. Is it $\mathcal{S}_\tau$-regular? $\mathcal{S}_\tau$-synchronized?

Rmk if you **already know** that $p(n) = 2n + 1$, imagine that you don't 😉

Abstract Numeration Systems (with zeros)

Definition (Lecomte, Rigo 2000) A **numeration system (NS)** is a tuple $(L, A, <, 0)$ with A the alphabet, ordered by $<$, minimal element $0 \in A$, and $L \subseteq A^*$ such that $\varepsilon \in L$ and $w \in L \Leftrightarrow 0w \in L \quad \forall w \in A^*$.

The **encoding** $\mathrm{rep}_{\mathcal{S}}(n)$ of $n \in \mathbb{N}$ is the nth element of $L \setminus 0^+L$ (in radix order).

The **valuation** $\mathrm{val}_{\mathcal{S}}(u)$ of $u \in L$ is $\mathrm{rep}_{\mathcal{S}}^{-1}(v)$ where $u \in 0^*v$.

Let $\langle ., . \rangle$ be the **canonical isomorphism** between $\bigcup_{n \geq 0} (A^n \times B^n)$ and $(A \times B)^*$.

Abstract Numeration Systems (with zeros)

Definition (Lecomte, Rigo 2000) A **numeration system (NS)** is a tuple $(L, A, <, 0)$ with A the alphabet, ordered by $<$, minimal element $0 \in A$, and $L \subseteq A^*$ such that $\varepsilon \in L$ and $w \in L \Leftrightarrow 0w \in L \quad \forall w \in A^*$.

The **encoding** $\mathrm{rep}_S(n)$ of $n \in \mathbb{N}$ is the nth element of $L \setminus 0^+L$ (in radix order).

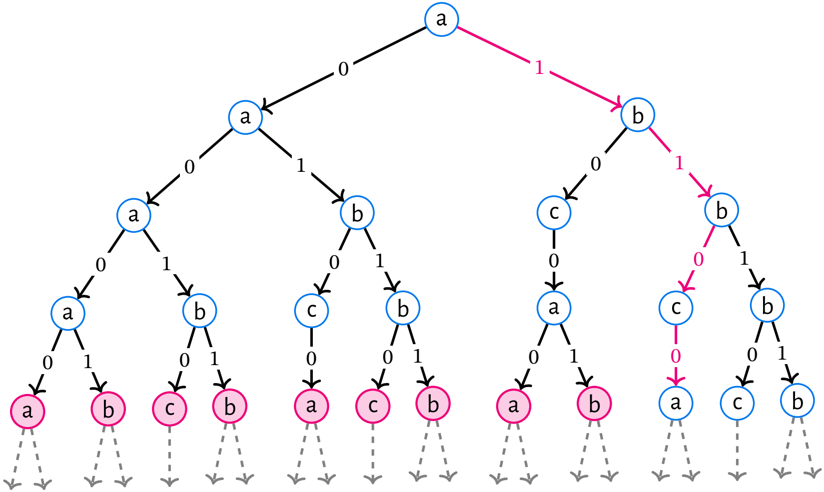
The **valuation** $\mathrm{val}_S(u)$ of $u \in L$ is $\mathrm{rep}_S^{-1}(v)$ where $u \in 0^*v$.

Let $\langle .,. \rangle$ be the **canonical isomorphism** between $\bigcup_{n \geq 0} (A^n \times B^n)$ and $(A \times B)^*$.

In this talk, consider every **ANS**:

- **prefix-closed** *i.e.* $uv \in L \Rightarrow u \in L, \quad \forall u, v;$
- **prolongeable** *i.e.* $u \in L \Rightarrow u0 \in L, \quad \forall u.$

Dumont-Thomas Numeration Systems



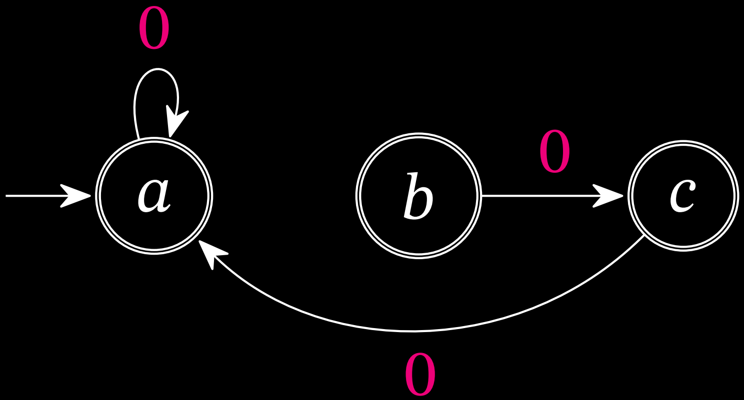
The **DTNS** for $\tau : a \mapsto ab,\ b \mapsto cb,\ c \mapsto a$ is build from the language L_τ of labels of the **derivation tree** of τ.

Dumont-Thomas numeration system

$$\tau : a \to ab$$
$$b \to cb$$
$$c \to a$$

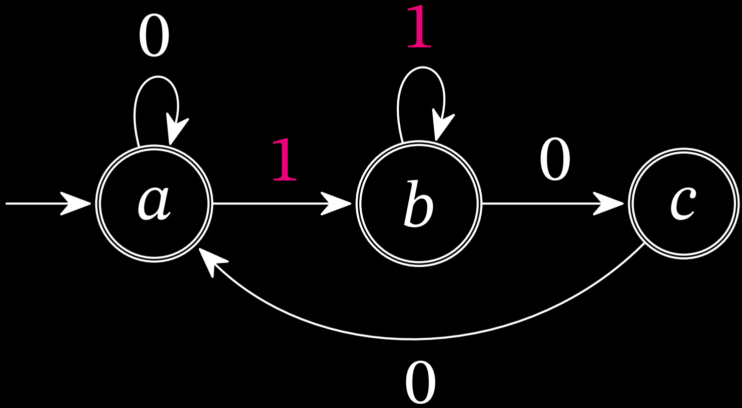
Dumont-Thomas numeration system

$$\tau : \begin{array}{l} a \to \overset{0\,1}{ab} \\ b \to cb \\ c \to a \end{array}$$



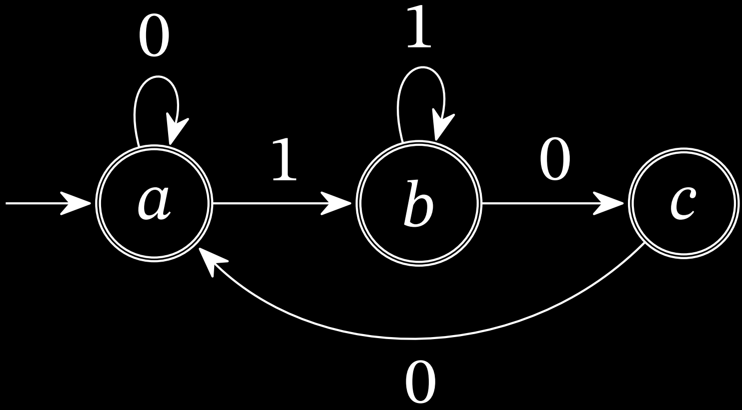
Dumont-Thomas numeration system

$$\begin{array}{rl} & \quad\;\; {\scriptstyle 0\,1} \\ \tau : a \to & ab \\ b \to & cb \\ c \to & a \end{array}$$



Dumont-Thomas numeration system

$$\tau : \begin{array}{l} \scriptstyle 0\ 1 \\ a \to ab \\ b \to cb \\ c \to a \end{array}$$



Dumont-Thomas numeration system

$$\tau : \begin{aligned} a &\to \overset{0}{a}\overset{1}{b} \\ b &\to cb \\ c &\to a \end{aligned}$$

As a DFA this automaton gives an **Abstract Numeration System** for radix order.

As a DFAO the same automaton gives an **automatic representation** for the fixpoint.

As it is **Pisot**, it has a **regular addition** [WORDS 2025].

$\mathcal{S}$-automatic and $\mathcal{S}$-synchronized Sequences

Definition A **DFAO** is a deterministic finite automaton with output. It computes a sequence $(f(n))_{n\geq 0}$ by reading $\mathrm{rep}_{\mathcal{S}}(n)$ and outputting the label of the reached state.

A sequence is **$\mathcal{S}$-automatic** if it is computed by a **DFAO**.

A predicate is **$\mathcal{S}$-synchronized** if it is computed by a **DFA** reading tuples.

Theorem (First-order closure (Büchi–Bruyère)) The set of $\mathcal{S}$-recognizable predicates is closed under **First-Order** operations (Boolean operations and existential quantification).

Minimal DFAO. Every $\mathcal{S}$-automatic sequence has a unique minimal **DFAO**, computable by standard minimization.

Walnut tool

101 predicates

- $\mathrm{feq}(i, j, n)$ positions i and j in $\mathbf{e}$ have the **same factor** of length n

$$\forall k \ (k < n \implies \mathbf{e}[i+k] = \mathbf{e}[j+k])$$

- $\mathrm{first}(i, n)$ position i is the **first occurrence** of a length-n factor

$$\forall j \ (\mathrm{feq}(i, j, n) \implies i \leqslant j)$$

Both are $\mathcal{S}_\tau$-synchronized, built by **Walnut** from the automatic sequence $\mathbf{e}$ via first-order formulas in $\mathcal{S}_\tau$.

$\mathcal{S}$-regular sequences

Definition A **linear representation** of dimension n over $\mathbb{K}$ and A is a triple (λ, μ, γ) with $\lambda \in \mathbb{K}^{1\times n}$, $\mu : A^* \to \mathbb{K}^{n\times n}$ a monoid morphism, $\gamma \in \mathbb{K}^{n\times 1}$. It defines $f(w) = \lambda\,\mu(w)\,\gamma$. A series is **recognizable** if it admits such a representation.

The **reduced** (minimal) representation is unique up to isomorphism.

The **support** $\mathrm{supp}(f)$ of a series f is the language $\Sigma^* \setminus f^{-1}(0)$.

$\mathcal{S}$-regular sequences

Definition A **linear representation** of dimension n over $\mathbb{K}$ and A is a triple (λ, μ, γ) with $\lambda \in \mathbb{K}^{1\times n}$, $\mu : A^* \to \mathbb{K}^{n\times n}$ a monoid morphism, $\gamma \in \mathbb{K}^{n\times 1}$. It defines $f(w) = \lambda\, \mu(w)\, \gamma$. A series is **recognizable** if it admits such a representation.

The **reduced** (minimal) representation is unique up to isomorphism.

The **support** $\mathrm{supp}(f)$ of a series f is the language $\Sigma^* \setminus f^{-1}(0)$.

$\mathcal{S}$-regular sequences are recognizable series evaluated on $\mathrm{rep}_{\mathcal{S}}(n)$.

$\mathcal{S}$-regular sequences

Definition A **linear representation** of dimension n over $\mathbb{K}$ and A is a triple (λ, μ, γ) with $\lambda \in \mathbb{K}^{1\times n}$, $\mu : A^* \to \mathbb{K}^{n\times n}$ a monoid morphism, $\gamma \in \mathbb{K}^{n\times 1}$. It defines $f(w) = \lambda\, \mu(w)\, \gamma$. A series is **recognizable** if it admits such a representation.

The **reduced** (minimal) representation is unique up to isomorphism.

The **support** $\operatorname{supp}(f)$ of a series f is the language $\Sigma^* \setminus f^{-1}(0)$.

$\mathcal{S}$-regular sequences are recognizable series evaluated on $\operatorname{rep}_{\mathcal{S}}(n)$.

Definition The **synchronized addition** $f \mathbin{\diamondplus} g$ between two series $f : \Sigma^* \to \mathbb{K}$ and $g : \Gamma^* \to \mathbb{K}$ is defined as

$$\left(f \mathbin{\diamondplus} g\right)(\langle u, v\rangle) = f(u) + g(v) \quad \forall (u, v) \in \Sigma^m \times \Gamma^m$$

Back to 101: counting first gives $p(n)$

From a synchronized predicate to a regular sequence

The projection

$$p(n) \;=\; \#\{\, i \;\big|\; \mathrm{first}(i,n) \,\}$$

is obtained by summing (projecting) a synchronized predicate over one variable — a standard operation yielding an $\mathcal{S}_\tau$-regular sequence.

Handled by **Walnut**.

$\mathcal{S}$-automatic

DFAO

$f(n) = \pi(\delta(q_0, \mathrm{rep}_{\mathcal{S}}(n)))$

$\mathcal{S}$-synchronized

$\mathcal{S}$-automatic

DFAO

$f(n) = \pi(\delta(q_0, \mathrm{rep}_{\mathcal{S}}(n)))$

DFA, closed by **FO formulaes**

$\langle \mathrm{rep}_{\mathcal{S}}(n), \mathrm{rep}_{\mathcal{S}}(f(n)) \rangle \in L(\mathcal{A})$

$\mathcal{S}$-regular

$\mathcal{S}$-synchronized

$\mathcal{S}$-automatic

DFAO

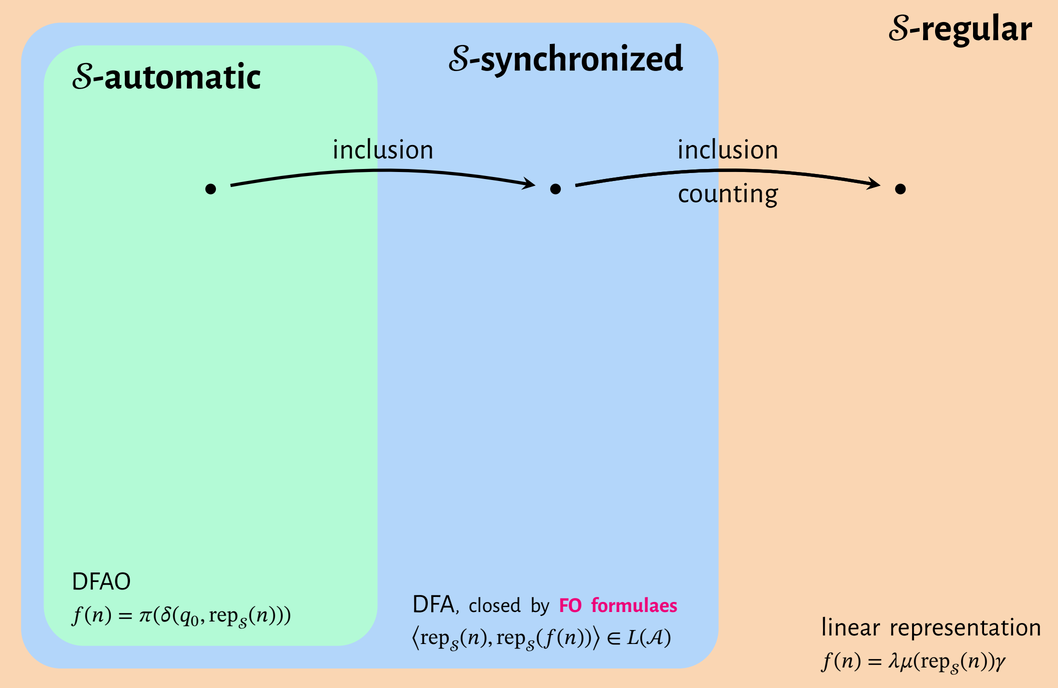
$f(n) = \pi(\delta(q_0, \mathrm{rep}_{\mathcal{S}}(n)))$

DFA, closed by **FO formulaes**

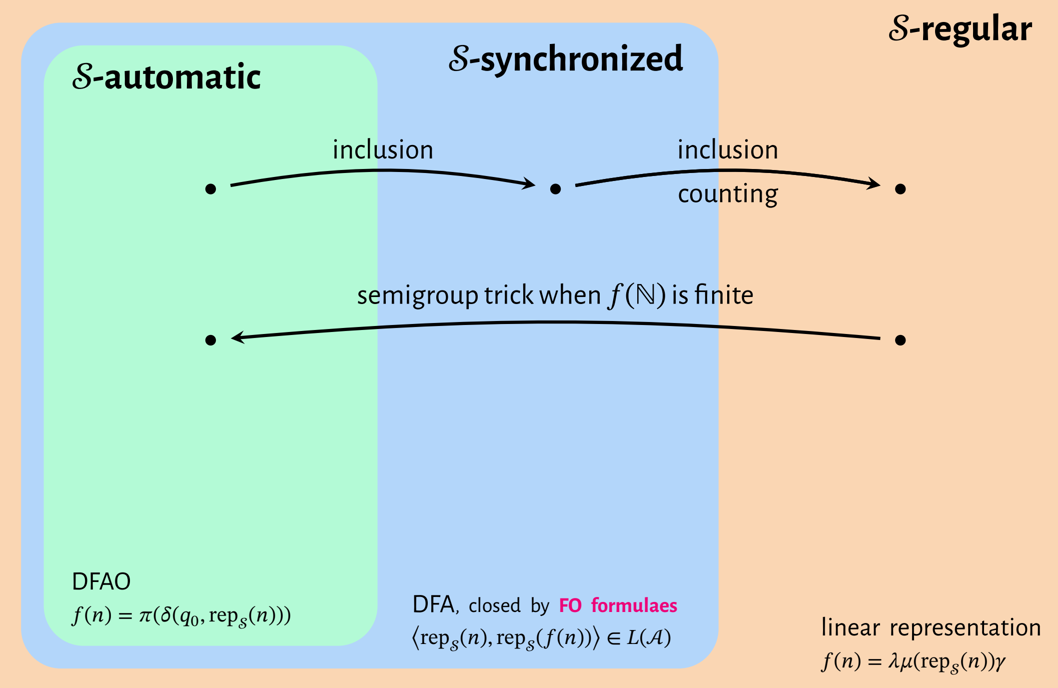
$\langle \mathrm{rep}_{\mathcal{S}}(n), \mathrm{rep}_{\mathcal{S}}(f(n)) \rangle \in L(\mathcal{A})$

linear representation

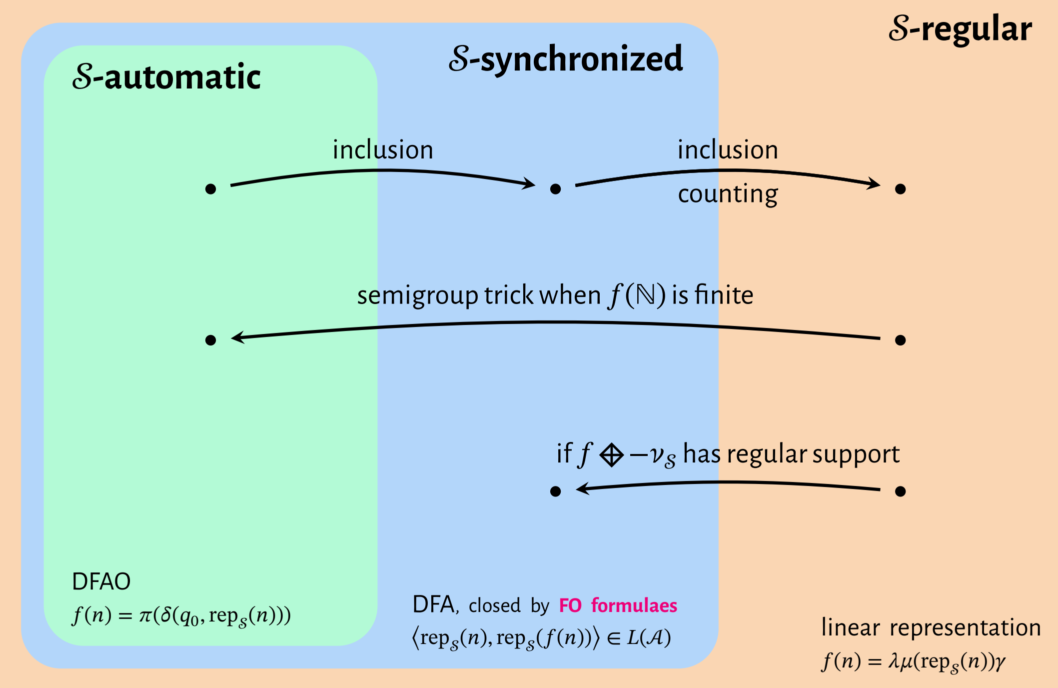
$f(n) = \lambda \mu(\mathrm{rep}_{\mathcal{S}}(n)) \gamma$



$\mathcal{S}$-regular
$\mathcal{S}$-synchronized
$\mathcal{S}$-automatic
inclusion
inclusion
counting
DFAO
$f(n) = \pi(\delta(q_0, \mathrm{rep}_{\mathcal{S}}(n)))$
DFA, closed by FO formulaes
$\langle \mathrm{rep}_{\mathcal{S}}(n), \mathrm{rep}_{\mathcal{S}}(f(n)) \rangle \in L(\mathcal{A})$
linear representation
$f(n) = \lambda \mu(\mathrm{rep}_{\mathcal{S}}(n)) \gamma$



$\mathcal{S}$-regular
$\mathcal{S}$-synchronized
$\mathcal{S}$-automatic
inclusion
inclusion
counting
semigroup trick when $f(\mathbb{N})$ is finite
DFAO
$f(n) = \pi(\delta(q_0, \mathrm{rep}_{\mathcal{S}}(n)))$
DFA, closed by FO formulaes
$\langle \mathrm{rep}_{\mathcal{S}}(n), \mathrm{rep}_{\mathcal{S}}(f(n)) \rangle \in L(\mathcal{A})$
linear representation
$f(n) = \lambda\mu(\mathrm{rep}_{\mathcal{S}}(n))\gamma$



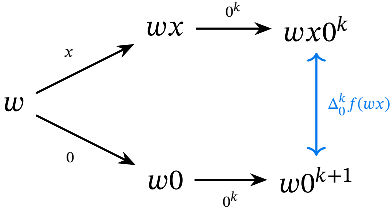
$\mathcal{S}$-regular
$\mathcal{S}$-synchronized
$\mathcal{S}$-automatic
inclusion
inclusion
counting
semigroup trick when $f(\mathbb{N})$ is finite
if $f \oplus -\nu_{\mathcal{S}}$ has regular support
DFAO
$f(n) = \pi(\delta(q_0, \mathrm{rep}_{\mathcal{S}}(n)))$
DFA, closed by FO formulaes
$\langle \mathrm{rep}_{\mathcal{S}}(n), \mathrm{rep}_{\mathcal{S}}(f(n)) \rangle \in L(\mathcal{A})$
linear representation
$f(n) = \lambda\mu(\mathrm{rep}_{\mathcal{S}}(n))\gamma$

The Distinguished Role of Letter 0

base-k numeration Appending 0 to a word **shifts in the numeration**.

$$n \xmapsto{\times k} kn \iff w \xmapsto{\cdot 0} w0$$

In any **ANS**, in a **linear representation** (λ, μ, γ), matrix $\mu(0)$ is the **transition matrix for the base step**.



Idea Measure how f varies when extending by 0^k vs. inserting one 0 first.

Linear Recurrent Sequences and $\mu(0)$

Definition $(u_k)_{k\geq 0}$ is **linearly recurrent** (LRS) if there exist $a_0, \ldots, a_{d-1}$ such that $u_{k+d} = a_{d-1}u_{k+d-1} + \cdots + a_0 u_k$ for all $k \geq 0$. The **minimal polynomial** of the sequence is the monic polynomial of smallest degree annihilating the recurrence.

Key fact: Recognizable series over $\{0\}^*$ are *exactly* the LRS:

$$u_k = \lambda\, \mu(0)^k\, \gamma$$

The matrix $\mu(0)$ plays the role of the **companion matrix** of the recurrence.

Remark In **reduced form**, the minimal polynomial of $\mu(0)$ *is* the minimal recurrence polynomial of the LRS. All LRS arising from $\Delta_0 f$ will share the **minimal polynomial** of $\mu(0)$.

The Discrete Derivative Δ_0

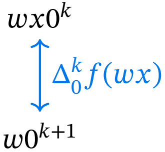
Definition The **discrete derivative** $\Delta_0 f$ of $f : A^* \to \mathbb{K}$ maps each word wx to the sequence $(\Delta_0^k f(wx))_{k\geq 0}$ where $w \in A^*$, $x \in A$, $k \geq 0$, and

$$\Delta_0^k f(wx) = f(wx \cdot 0^k) - f(w \cdot 0^{k+1}) \quad .$$

Dynamical intuition:

$\Delta_0^k f(wx)$ measures the **gap** between the two orbits starting from wx and $w0$ under the repeated action of 0.

At $k = 0$: $\Delta_0^0 f(wx) = f(wx) - f(w0)$, the **local difference** at the last letter.



Explicit Representation of $\Delta_0 f$

Theorem Let (λ, μ, γ) be a linear representation of f of dimension n, $w \in A^*, x \in A$.

$$\Delta_0^k f(wx) \;=\; \lambda\, \mu(w) \big(\mu(x) - \mu(0)\big)\, \mu(0)^k\, \gamma$$

In particular, $k \mapsto \Delta_0^k f(wx)$ is an **LRS** whose recurrence polynomial divides $\mathrm{minpoly}(\mu(0))$.

Theorem Given (λ, μ, γ) of dimension n, there exists a linear representation of $\Delta_0 f$ of dimension $2n$ with output matrix $\left(\frac{\Gamma'}{-\Gamma'}\right)$ with

$$\Gamma' \;=\; \big(\gamma \;\big|\; \mu(0)\gamma \;\big|\; \mu(0)^2\gamma \;\big|\; \cdots \;\big|\; \mu(0)^{2n-1}\gamma\big) \;\in\; \mathbb{K}^{2n\times 2n}$$

The output matrix has $2n$ columns: the first $2n$ terms of each output LRS, sufficient for **Berlekamp–Massey** ($\deg \mathrm{minpoly}(\mu(0)) \leq n$).

Discrete Integral I_0 and Closure

Definition The **discrete integral** $I_0 g$ is the **right inverse** of Δ_0:

$$\Delta_0(I_0 g) = g \quad .$$

Discrete Fundamental Theorem

(i) $I_0(\Delta_0 f) = f + c$ (additive constant),

(ii) $\Delta_0(I_0 g) = g$.

Closure Theorem The class of **recognizable series** is closed under Δ_0 and I_0, i.e. **$\mathcal{S}$-regular sequences** are closed under discrete derivation and integration.

From Derivative to Synchronized

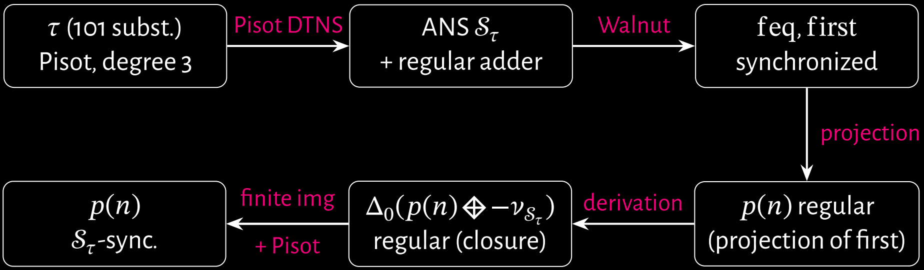
Lets use the **semigroup trick** with derivatives.

Theorem If $\Delta_0 f$ has **finite image**, then $\Delta_0 f$ is $\mathcal{S}$-**automatic**, it is a **sequence automaton** [WORDS 2025].

Combine everything together.

Corollary If $\mathrm{minpoly}(\mu(0))$ is a **Pisot polynomial**, then the support of $f \oplus -\nu_{\mathcal{S}}$ is recognizable and f is $\mathcal{S}$-**synchronized**.

Back to 101: Computing $p(n)$ Step by Step



- All automatic/synchronized steps are checkable by **Walnut**.
- An explicit **DFA** for $p(n)$ is **effectively** computed.

toolchain under preparation