

Word-Mappings of level 3

G. Sénizergues

LaBRI and Université de Bordeaux **

Abstract. Sequences of numbers (either natural integers, or integers or rational) of level $k \in \mathbb{N}$ have been defined in [Fra05,FS06] as the sequences which can be computed by deterministic pushdown automata of level k . This definition has been extended to sequences of *words* indexed by *words* in [Sén07,FMS14]. We characterise here the sequences of level 3 as the compositions of two HDTOL-systems. Two applications are derived:

- the sequences of rational numbers of level 3 are characterised by polynomial recurrences
- the equality problem for sequences of rational numbers of level 3 is decidable.

Keywords: Iterated pushdown automata ; recurrent sequences of numbers; equivalence problems.

Version: January 25, 2023

1 Introduction

The class of pushdown automata of level k (for $k \geq 1$) has been introduced in [Gre70],[Mas74] as a generalisation of the automata and grammars of [Aho68],[Aho69],[Fis68] and has been the object of many further studies: see [Mas76], [ES77], [Dam82], [Eng83], [ES84], [EV86], [DG86], and more recently [Cau02], [KNU02], [CW03],[Fra05],[HO07],[CS21].

The class of *integer* sequences computed (in a suitable sense) by such automata was defined in [Fra05],[FS06] (we denote it by \mathbb{S}_k). The class of *word* mappings from a free monoid A^* into a free monoid B^* computed by such automata was defined in [Sén07] (as a straightforward extension of \mathbb{S}_k); we denote it by $\mathbb{S}_k(A^*, B^*)$.

Let $\mathcal{F}(S_k(A^*, \mathbb{N}))$ be the set of all the sequences of *rational* numbers which can be decomposed as $\frac{a_n - b_n}{a'_n - b'_n}$ for sequences $a, b, a', b' \in \mathbb{S}_k(A^*, \mathbb{N})$ (Definition 9 of [Sén07]).

These classes of number sequences fulfill many closure properties and generalize

** mailing adress:LaBRI and UFR info, Université de Bordeaux
 351 Cours de la libération -33405- Talence Cedex.
 email:{geraud.senizergues}@u-bordeaux.fr; fax: 05-56-84-66-69;
 URL:http://dept-info.labri.u-bordeaux.fr/~ges

some well-known classes of recurrent sequences or formal power series (see [FS06], sections 4,6,7).

This paper focuses on level 3, for words i.e. $\mathbb{S}_3(A^*, B^*)$. The class $\mathcal{F}(\mathbb{S}_3(\mathbb{N}, \mathbb{N}))$ contains all the so-called P-recurrent sequences of rational numbers, corresponding also to the D-finite formal power series (see [Sta80] for a survey and [PWZ96] for a thorough study of their algorithmic properties).

The main result of this paper consists in characterising the class $\mathbb{S}_3(A^*, B^*)$ in terms of Lindenmayer systems i.e. iterated homomorphisms.

Theorem 2:

The following properties are equivalent:

- 1- $f \in \mathbb{S}_3(A^*, B^*)$
- 2- There exists a finite family $(H_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \text{Hom}(C^*, C^*)$ which fulfils a system of recurrent relations in $(\text{Hom}(C^*, C^*), \circ, \text{Id})$, an element $h \in \text{Hom}(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$:

$$f(w) = h(H_1(w)(c)).$$

- 3- f is a composition of a DT0L sequence $g : A^* \rightarrow C^*$ by a HDT0L sequence $h : C^* \rightarrow B^*$.

As a corollary, the class $\mathbb{S}_3(A^*, \mathbb{N})$ is characterised by polynomial recurrences. The equality problem for two sequences in $\mathcal{F}(\mathbb{S}_3(A^*, \mathbb{N}))$ can thus be solved by a suitable reduction to polynomial ideal theory (Theorem 5 of [Sén07]).

This theorem 2, as well as a general version treating the class $\mathbb{S}_k(A^*, B^*)$, for every $k \geq 2$, was announced in Theorem 6 of the extended abstract [Sén07]. The full proof for $k = 3$ is given here for the first time.

The main difficulty is to prove that (1) \Rightarrow (2). The proof consists in constructing from the given pushdown-automaton of level 3, a sequence of homomorphisms that fulfils a monoidal recurrence, in the sense of [FMS14]. The main theorem of this reference (slightly reformulated here as Theorem 1 in section 2) then gives the conclusion.

Table of Contents

Word-Mappings of level 3	1
<i>G. Sénizergues</i>	
1 Introduction	1
2 Preliminaries	3
2.1 Sets-Relations	3
2.2 Abstract rewriting	4
2.3 Monoids	4
2.4 General Automata	6
2.5 Pushdown automata of level k	8
2.6 Regular sets of pushdowns of level k	13
2.7 Recurrences in monoids	13
2.8 Word-mappings of level 2	15
3 Word-mappings of level 3	16
3.1 Alphabets E and \mathcal{W}	17
3.2 Morphisms H_i^w	18
3.3 Mixed recurrence	20
3.4 Compositional recurrence	23
3.5 Termination	25
4 Applications	36
5 Examples and counter-examples	38
6 Perspectives	41

2 Preliminaries

We recall/introduce in §2.1-2.7 some notation and basic definitions which will be used throughout the text. In §2.8 we recall previous results on sequences of level 2, which are crucial for this work. For overviews on the links between automata-theory and number-theory, we forward the reader to [AS03,Rig04].

2.1 Sets-Relations

Given a set E , we denote by $\mathcal{P}(E)$ the set of its subsets and by $\mathcal{P}_f(E)$ the set of its *finite* subsets.

A *binary relation* from a set E into a set F is a subset R of $E \times F$. The domain and image of R are defined by:

$$\text{dom}(R) := \{x \in E \mid \exists y \in F, (x, y) \in R\}, \quad \text{im}(R) := \{y \in F \mid \exists x \in E, (x, y) \in R\}.$$

We denote by \circ the composition of binary relations: if $R \subseteq E \times F$, $R' \subseteq F \times G$ then:

$$R \circ R' := \{(x, z) \in E \times G \mid \exists y \in F, (x, y) \in R \wedge (y, z) \in R'\}$$

We denote by $\text{BR}(Q)$ the set of binary relations over Q .

A *function* from the set E into the set F is a binary relation $f \subseteq E \times F$ such that,

$$\forall(x, y), \forall(x', y') \in f, x = x' \Rightarrow y = y'$$

Note that, when using a functional notation, we still use the composition operator \circ as above i.e.

$$(f \circ g)(x) := g(f(x)).$$

We call *mapping* (or simply, *map*) from E to F any function $f : E \rightarrow F$ such that $\text{dom}(f) = E$.

The empty set \emptyset is also denoted by the symbol ε when it is viewed as the empty word.

2.2 Abstract rewriting

Let E be some set and $\rightarrow \subseteq E \times E$. The relations \rightarrow^n (for any natural integer n) and \rightarrow^* are defined from the binary relation \rightarrow as usual (see [Hue80]). A *derivation* with respect to the relation \rightarrow is a sequence :

$$D = (e_i)_{i \in I}$$

of elements $e_i \in E$, in dexted by $I = [0, n]$ or $I = \mathbb{N}$, such that, for every $i \in \mathbb{N}$, if $i+1 \in I$ then $e_i \rightarrow e_{i+1}$. The *length* of derivation D is the integer n (if $I = [0, n]$) or ∞ (if $I = \mathbb{N}$). The notation $e \rightarrow^\infty$ means that there exists some derivation of length ∞ that starts on e . The relation \rightarrow is called *noetherian* iff there exists no $e \in E$ such that $e \rightarrow^\infty$.

2.3 Monoids

We recall a monoid is a triple $\langle M, \cdot, \mathbf{1} \rangle$ such that, M is a set (the carrier of the monoid), \cdot is a composition law which is associative and $\mathbf{1}$ is a neutral element for \cdot (on both sides). Given two monoids $\mathbb{M}_1 := \langle M_1, \cdot, \mathbf{1}_1 \rangle$, $\mathbb{M}_2 := \langle M_2, \cdot, \mathbf{1}_2 \rangle$ a monoid-homomorphism from \mathbb{M}_1 to \mathbb{M}_2 is a map $h : M_1 \rightarrow M_2$ fulfilling: for every $x, y \in M_1$

$$h(x \cdot y) = h(x) \cdot h(y) \text{ and } h(\mathbf{1}_1) = \mathbf{1}_2.$$

We denote by $\text{HOM}(\mathbb{M}_1, \mathbb{M}_2)$ the set of all monoid-homomorphism from \mathbb{M}_1 to \mathbb{M}_2 . For every monoid \mathbb{M} , the set $\text{HOM}(\mathbb{M}, \mathbb{M})$, endowed with the composition law \circ and the identity map Id_M is a monoid.

Given a set X (that we see as an “alphabet”), we denote by X^* the set of all finite words labelled on this set X . We denote by \cdot the binary operation of concatenation over X^* and denote by ε the empty word. The structure $\langle X^*, \cdot, \varepsilon \rangle$ is the *free monoid* over the alphabet X . For every integer $k \geq 0$, by $X^{\leq k}$ we mean the set $\{u \in X^* \mid |u| \leq k\}$.

Partial monoids We call *partial-monoid* every triple $\mathbb{M} = \langle M, \cdot, \mathbf{1} \rangle$ such that, M is a set (the carrier of the partial-monoid), \cdot is a function from $M \times M$ into M such that: for every $x, y, z \in M$

$$[(x, y) \in \text{dom}(\cdot) \text{ and } (x \cdot y, z) \in \text{dom}(\cdot)] \Leftrightarrow [(y, z) \in \text{dom}(\cdot) \text{ and } (x, y \cdot z) \in \text{dom}(\cdot)]$$

and, if x, y, z fulfill the above two equivalent prerequisite, then

$$(x \cdot y) \cdot z = x \cdot (y \cdot z)$$

and, for every $x \in M$

$$x \cdot \mathbf{1} = \mathbf{1} \cdot x = x.$$

Note that, if $\langle M, \cdot, \mathbf{1} \rangle$ is a partial monoid, then

$$\langle \mathcal{P}(M), \cdot, \{\mathbf{1}\} \rangle$$

is a monoid (here, for $X, Y \subseteq M$, $X \cdot Y = \{x \cdot y \mid x \in X, y \in Y, (x, y) \in \text{dom}(\cdot)\}$). Let us consider the monoid \mathbb{M}^0 defined by

$$\mathbb{M}^0 := \langle \mathcal{P}_1(M), \cdot, \{\mathbf{1}\} \rangle$$

where $\mathcal{P}_1(M)$ denotes the set of subsets of M with at most one element. \mathbb{M}^0 is a sub-monoid of the monoid $\langle \mathcal{P}(M), \cdot, \{\mathbf{1}\} \rangle$. When \mathbb{M} is a partial-monoid which fails to be a monoid, there is a bijection between $\mathcal{P}_1(M)$ and $M \cup \{0\}$, where 0 is a new element not in M . Most assertions about a partial monoid \mathbb{M} can thus be deduced from the similar assertions about the monoid \mathbb{M}^0 .

Congruences An equivalence relation \sim over M is called a *right-regular* equivalence if and only if, for every $x, y, z \in M$

$$x \sim y \Rightarrow x \cdot z \sim y \cdot z.$$

(the notion of left-regular equivalence is defined analogously).

An equivalence over the monoid M is called a *congruence* iff it is both right-regular and left-regular.

Given an homomorphism $h : \mathbb{M}_1 \rightarrow \mathbb{M}_2$, the kernel of h is the congruence \equiv over \mathbb{M}_1 defined by:

$$x \equiv y \Leftrightarrow h(x) = h(y).$$

We shall use the following construction. Let $\mathbb{M}_1, \mathbb{M}_2$ be monoids and let us consider their free product $\mathbb{M}_1 * \mathbb{M}_2$. Let $h : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ be an homomorphism. Let $\hat{h} : \mathbb{M}_1 * \mathbb{M}_2 \rightarrow \mathbb{M}_2$ be the unique monoid-homomorphism such that:

$$\forall x_1 \in M_1, \hat{h}(x_1) = h(x_1), \quad \forall x_2 \in M_2, \hat{h}(x_2) = x_2.$$

The definition of \hat{h} shows that:

$$\forall x, y \in M_1, (x, y) \in \ker(\hat{h}) \Leftrightarrow (x, y) \in \ker(h).$$

Therefore $\ker(\hat{h})$ is an extension of the congruence $\ker(h)$ to the monoid $\mathbb{M}_1 * \mathbb{M}_2$. Starting with a congruence \equiv over \mathbb{M}_1 , we can define $\mathbb{M}_2 := \mathbb{M}_1 / \equiv$, consider the projection $h : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ and define an extension $\hat{\equiv}$ of \equiv over the monoid $\mathbb{M}_1 * (\mathbb{M}_1 / \equiv)$ by

$$\hat{\equiv} := \ker(\hat{h}). \quad (1)$$

We call $\hat{\equiv}$ the self-extension of \equiv . The monoid \mathbb{M}_1 / \equiv is isomorphic with $(\mathbb{M}_1 * (\mathbb{M}_1 / \equiv)) / \hat{\equiv}$, by the map

$$[x]_{\equiv} \mapsto [x]_{\hat{\equiv}}.$$

Intuitively, the elements of $\mathbb{M}_1 * (\mathbb{M}_1 / \equiv)$ are “mixed products” of “genuine” elements of M_1 with “fake” elements that are merely congruence-classes. When we map the genuine elements onto their classes, what we obtain is a product of classes. The value of this product is the class of the “mixed product” for the self-extension $\hat{\equiv}$.

This construction will be useful for extending some natural congruences over ordinary pushdowns of order 3 to pushdowns of order 3 *extended with undeterminates* (see subsection 3.5), where the undeterminates are classes of ordinary pushdowns modulo some congruence.

2.4 General Automata

General automaton We introduce in this paragraph a notion of automaton over an arbitrary memory structure, which is essentially the one defined by [Eng91]. Let us call *data-structure* every tuple $\mathbb{D} = \langle D, F, \text{OP}, \text{tops} \rangle$ consisting of a set D , a finite set F , a set OP of unary operations ($\forall \text{op} \in \text{OP}$, op is a total map : $D \rightarrow D$) and a total map $\text{tops} : D \rightarrow F$.

Definition 1 (Automaton over \mathbb{D}). *An automaton over the data-structure \mathbb{D} and the terminal alphabet B is a 4-tuple*

$$\mathcal{A} = (Q, B, \mathbb{D}, \delta)$$

where Q is a finite set of states,

$\mathbb{D} = \langle D, F, \text{OP}, \text{tops} \rangle$ is a data-structure,

the transition function, δ , is a map from $Q \times (B \cup \{\varepsilon\}) \times F$ into $\mathcal{P}(Q \times \text{OP}(\mathbb{D}))$.

A configuration is a triple $(q, u, w) \in Q \times B^* \times D$. The *direct computation relation* $\vdash_{\mathcal{A}}$, is defined as usual: it is a binary relation over the set of configurations. For every $q, q' \in Q, u, u' \in B^*, d, d' \in D$, we let

$$(q, u, d) \vdash_{\mathcal{A}} (q', u', d')$$

iff

- $u = \bar{b} \cdot u'$ for some $\bar{b} \in B \cup \{\varepsilon\}$
- $\exists \text{op} \in \text{OP}, (q', \text{op}) \in \delta(q, \bar{b}, \text{tops}(d))$ and $d' = \text{op}(d)$.

Pushdowns Given a data-structure \mathbb{D} and an alphabet Γ one constructs a new data-structure, denoted by $\mathbb{P}(\Gamma, \mathbb{D})$ and called the set of *pushdowns over* (Γ, \mathbb{D}) . It is defined as follows:

$$\mathbb{P}(\Gamma, \mathbb{D}) := \langle P(\Gamma, D), F', \text{OP}', \text{tops}' \rangle$$

where

$$\begin{aligned} P(\Gamma, D) &:= (\Gamma[D])^*, \quad F' := \Gamma \cdot F \cup \Gamma \cup \{\varepsilon\}, \quad \text{OP}' := \text{OP} \cup \{\text{pop}\} \cup \{\text{push}(h) \mid h \in \Gamma^+\}, \\ \text{tops}'(\gamma[d]w) &:= \gamma \cdot \text{tops}(d), \quad \text{tops}(\varepsilon) := \varepsilon. \end{aligned}$$

The maps $\text{push}(h) : P(\Gamma, D) \rightarrow P(\Gamma, D)$ and $\text{pop} : P(\Gamma, D) \rightarrow P(\Gamma, D)$ are defined by: for every $h \in \Gamma^+, d \in D, w \in D^*$: if $h = h_1 \cdots h_n$ with $h_i \in \Gamma$

$$\text{push}(h)(\gamma[d]w) := h_1[d] \cdots h_n[d]w, \quad \text{pop}(\gamma[d]w) := w,$$

$$\text{push}(h)(\varepsilon) := \varepsilon, \quad \text{pop}(\varepsilon) := \varepsilon.$$

A *pushdown automaton*, in the usual terminology, over a pushdown-alphabet Γ , appears to be an automaton over the data-structure $\mathbb{P}(\Gamma, \emptyset)$. A pushdown-automaton of level k , for some integer $k \geq 1$, is an automaton over the data-structure \mathbb{D}_k defined by:

$$\mathbb{D}_0 := \langle \{\emptyset\}, \emptyset, \emptyset, \emptyset \rangle \quad \text{and} \quad \forall i \in \mathbb{N}, \quad \mathbb{D}_{i+1} := \mathbb{P}(\Gamma, \mathbb{D}_i).$$

Derivations versus computations Let \mathbb{D} be some data-structure and

$$\mathcal{A} = (Q, B, \mathbb{P}(\Gamma, \mathbb{D}), \delta)$$

a pushdown-automaton i.e. an automaton over the data-structure $\mathbb{P}(\Gamma, \mathbb{D})$. We associate with \mathcal{A} a (possibly infinite) “alphabet”

$$V_{\mathcal{A}} = \{(p, \omega, q) \mid p, q \in Q, \omega \in P(\Gamma, D) - \{\varepsilon\}\}. \quad (2)$$

whose elements are called *variables*. The set of *productions* associated with \mathcal{A} , denoted by $P_{\mathcal{A}}$ is made of the set of all the following rules:

the *transition* rules:

$$(p, \omega, q) \rightarrow_{\mathcal{A}} \bar{a}(p', \omega', q)$$

if $(p, \bar{a}, \omega) \vdash_{\mathcal{A}} (p', \varepsilon, \omega')$ and $q \in Q$ is arbitrary,

$$(p, \omega, q) \rightarrow_{\mathcal{A}} \bar{a}$$

if $(p, \bar{a}, \omega) \vdash_{\mathcal{A}} (q, \varepsilon, \varepsilon)$.

the *decomposition* rule:

$$(p, \omega, q) \rightarrow_{\mathcal{A}} (p, \eta, r)(r, \eta', q)$$

if $\omega = \eta \cdot \eta', \eta \neq \varepsilon, \eta' \neq \varepsilon$ and $p, q, r \in Q$ are arbitrary. The one-step *derivation* generated by \mathcal{A} , denoted by $\rightarrow_{\mathcal{A}}$, is the smallest subset of $(V \cup \Sigma)^* \times (V \cup \Sigma)^*$

which contains $P_{\mathcal{A}}$ and is compatible with left-product and right-product. Finally, the *derivation* generated by \mathcal{A} , denoted by $\rightarrow_{\mathcal{A}}^*$, is the reflexive and transitive closure of $\rightarrow_{\mathcal{A}}$. These notions correspond to the usual notion of *context-free grammar* associated with the pushdown automaton \mathcal{A} . As soon as D is infinite, this variable alphabet is infinite, but all the usual properties of the relation $\rightarrow_{\mathcal{A}}$ and its links with $\vdash_{\mathcal{A}}$ remain true in this context (see [Har78, proof of theorem 5.4.3, pages 151-158]). In particular, for every $u \in \Sigma^*, p, q \in Q, \omega \in P(D)$

$$(p, \omega, q) \rightarrow_{\mathcal{A}}^* u \Leftrightarrow (p, u, \omega) \vdash_{\mathcal{A}}^* (q, \varepsilon, \varepsilon).$$

We usually assume that $P(D)$ and Q are disjoint, therefore, omitting the commas in (p, ω, q) does not lead to any confusion. The notions of *derivation* (mod $\rightarrow_{\mathcal{A}}$) is obtained by instantiating the general notion defined in subsection 2.2.

2.5 Pushdown automata of level k

pda Beside the usual notions of finite automaton and pushdown automaton, we shall consider here the notion of *pushdown automaton of level k* : this is an automaton (in the sense of Definition 1) over the data-structure \mathbb{D}_k (as defined in subsection 2.4). Let us describe in more details these automata:

Definition 2 (k -iterated pushdown store). Let Γ be a set. We define inductively the set of k -iterated pushdown-stores over Γ by:

$$0\text{-pds}(\Gamma) = \{\varepsilon\} \quad (k+1)\text{-pds}(\Gamma) = (\Gamma[k\text{-pds}(\Gamma)])^* \quad \text{it } k\text{-pds}(\Gamma) = \bigcup_{k \geq 0} k\text{-pds}(\Gamma).$$

The elementary operations that a k -pda can perform are:

- pop_j of level j (where $1 \leq j \leq k$), which consists of popping the leftmost letter of level j and all the bracket just on its right
- $\text{push}_j(h)$ of level j (where $1 \leq j \leq k, h \in \Gamma^+$), which consists of pushing successively all the symbols of h as new heads of the leftmost pushdown of level j , thus copying this pushdown in each bracket following each new head on its right (see example 1 below).

An operation of level j (whether pop_j or $\text{push}_j(h)$) applied on a pushdown where this leftmost pushdown of level j is ε , leaves the pushdown invariant.

Example 1. Let $\Gamma = \{A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2, C_3, D_1, D_2, D_3\}$. Let

$$\omega = A_1[A_2[A_3[\varepsilon]C_3[\varepsilon]]B_2[D_3[\varepsilon]C_3[\varepsilon]]B_1[B_2[B_3[\varepsilon]D_3[\varepsilon]]].$$

We shall (abusively) write:

$$\omega = A_1[A_2[A_3C_3]B_2[D_3C_3]]B_1[B_2[B_3D_3]].$$

i.e. we use a short notation where the bracketed empty words $[\varepsilon]$ that occur at the most internal level of the pushdown (here level 4) are removed.

The reading operation, applied on the above example gives:

$$\text{topsyms}(\omega) = A_1 A_2 A_3$$

The pop operations, applied on the above example give:

$$\text{pop}_1(\omega) = B_1[B_2[B_3 D_3]]$$

$$\text{pop}_2(\omega) = A_1[B_2[D_3 C_3]]B_1[B_2[B_3 D_3]]$$

$$\text{pop}_3(\omega) = A_1[A_2[C_3]B_2[D_3 C_3]]B_1[B_2[B_3 D_3]]$$

The push operations, applied on the above example gives:

$$\text{push}_1(AB)(\omega) = A[A_2[A_3 C_3]B_2[D_3 C_3]]B[A_2[A_3 C_3]B_2[D_3 C_3]]B_1[B_2[B_3 D_3]]$$

$$\text{push}_2(AB)(\omega) = A_1[A[A_3 C_3]B[A_3 C_3]B_2[D_3 C_3]]B_1[B_2[B_3 D_3]]$$

$$\text{push}_3(AB)(\omega) = A_1[A_2[ABC_3]B_2[D_3 C_3]]B_1[B_2[B_3 D_3]]$$

A transition of the automaton consists, given the word γ made of all the leftmost letters of the k -pushdown (the one of level 1, followed by the one of level 2, ..., followed by the one of level k), the state q and the leftmost letter b (or, possibly, the empty word ε) on the input tape, in performing one of the above elementary operations. More formally,

Definition 3 (operations on k -pds). Let $k \geq 1$, let $\text{POP}(k) := \{\text{pop}_j \mid j \in [1, k]\}$, $\text{PUSH}(k, \Gamma) := \{\text{push}_j(\gamma) \mid \gamma \in \Gamma^+, j \in [1, k]\}$, $\text{OP}(k, \Gamma) = \text{POP}(k) \cup \text{PUSH}(k, \Gamma)$, $\text{TOPSYMS}(k, \Gamma) := \Gamma^{\leq k} - \{\varepsilon\}$.

Definition 4 (k -pdas). A k -iterated pushdown automaton (abbreviated k -pda) over a terminal alphabet B is a 4-tuple

$$\mathcal{A} = (Q, B, \Gamma, \delta)$$

where

- Q is a finite set of states,
- Γ is a finite set of pushdown-symbols
- the transition function δ is a map from $Q \times (B \cup \{\varepsilon\}) \times \text{TOPSYMS}(k, \Gamma)$ into the set of finite subsets of $Q \times \text{OP}(k, \Gamma)$

The automaton \mathcal{A} is said *deterministic* iff, for every $q \in Q, \gamma \in \Gamma^{\leq k}, b \in B$

$$\text{Card}(\delta(q, \varepsilon, \gamma)) \leq 1 \text{ and } \text{Card}(\delta(q, b, \gamma)) \leq 1, \quad (3)$$

$$\text{Card}(\delta(q, \varepsilon, \gamma)) = 1 \Rightarrow \text{Card}(\delta(q, b, \gamma)) = 0. \quad (4)$$

Normalized automata In order to define a useful notion of map *computed* by a k -pda we introduce the following stronger condition: \mathcal{A} is called *strongly deterministic* iff, for every $q \in Q, \gamma \in \Gamma^{(k)} - \{\varepsilon\}$

$$\sum_{\bar{b} \in \{\varepsilon\} \cup B} \text{Card}(\delta(q, \bar{b}, \gamma)) \leq 1 \quad (5)$$

In other words, the automaton \mathcal{A} is *strongly deterministic* iff, the leftmost contents γ of the memory and the state q completely determine the transition of \mathcal{A} ,

in particular which letter b (or possibly the empty word) can be read. Therefore, such an automaton \mathcal{A} can accept at most one word w from a given configuration. We say that \mathcal{A} is *level-partitioned* iff Γ is the disjoint union of subsets $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ such that, in every transition of \mathcal{A} , every occurrence of a letter from Γ_i is at level i .

It is easy to transform any k -pushdown automaton \mathcal{A} into another one \mathcal{A}' which recognizes the same language and is level-partitioned. Moreover, if \mathcal{A} is strongly deterministic then \mathcal{A}' is strongly deterministic.

Variables As a particular case of the defining equation (2), we define the set $k - \text{var}(\Gamma)$ of *order k variables* over an alphabet Γ and a set of states Q :

$$k - \text{var}(\Gamma) = \{(p, \omega, q) \mid p, q \in Q, \omega \in k - \text{pds}(\Gamma) \setminus \{\varepsilon\}\}. \quad (6)$$

We name *variable-words* of level 3, the elements of $(k - \text{var}(\Gamma))^*$ i.e. the words over the alphabet $k - \text{vterm}(\Gamma)$.

Terms Given a denumerable alphabet Γ of pushdown symbols, we introduce another alphabet $\mathcal{U} = \{\Omega, \Omega', \Omega'', \dots, \Omega_1, \Omega_2, \dots, \Omega_n, \dots\}$ of *undeterminates*. We suppose that $\Gamma \cap \mathcal{U} = \emptyset$. We call a *term* of level k over the constant alphabet Γ and the alphabet of undeterminates \mathcal{U} , any $T \in k - \text{pds}(\Gamma \cup \mathcal{U})$ where every occurrence of an undeterminate U in T is followed by $[\varepsilon]$, in the rigorous bracketed notation (see example 1, first notation for ω).

Such a pushdown T , seen as a sequence of planar trees (as in fig.2 p.366 of [FS06]), has all the undeterminates at the leaves, as is the case for classical terms¹. We denote by $k - \text{term}(\Gamma, \mathcal{U})$ the set of all terms of level k over the constant alphabet Γ and the alphabet of undeterminates \mathcal{U} .

We denote an element of $k - \text{term}(\Gamma, \mathcal{U})$ by $T[\Omega_1, \Omega_2, \dots, \Omega_n]$ (resp. $T[\Omega, \Omega', \Omega'']$) provided that the only undeterminates appearing in T belong to $\{\Omega_1, \Omega_2, \dots, \Omega_n\}$ (resp. $\{\Omega, \Omega', \Omega''\}$). The notation $k - \text{vterm}(\Gamma, \mathcal{U})$ designates the set of *variable-terms* of level k , over the pushdown alphabet Γ and the set of undeterminates \mathcal{U} :

$$k - \text{vterm}(\Gamma, \mathcal{U}) = \{(p, \omega, q) \mid p, q \in Q, \omega \in k - \text{term}(\Gamma, \mathcal{U}) \setminus \{\varepsilon\}\}. \quad (7)$$

We name *variable-term-words* of level 3, the elements of $(k - \text{vterm}(\Gamma, \mathcal{U}))^*$ i.e. the words over the alphabet $k - \text{vterm}(\Gamma, \mathcal{U})$.

Graded alphabets A graded alphabet of height k is an alphabet $\Gamma = \bigcup_{i \in [1, k]} \Gamma_i$. We call the elements of Γ_i the letters of *level i* . Given $j \in [1, k]$, we call graded pushdown of level j over Γ , any element u of $j - \text{pds}(\Gamma)$ where every occurrence of a letter γ at level $k - j + i$ in u belongs to Γ_i . The notion of graded k -variable (over Γ, Q) is defined in the same way. Given two disjoint graded alphabets Γ, \mathcal{U} , we define in the same way the graded k -terms and the graded k -variable-terms. Given a graded alphabet Γ we denote again by $k - \text{pds}(\Gamma)$, $k - \text{var}(\Gamma)$ the

¹ However, no arity is attached here to the symbols of Γ , unlike what happens for the classical notion of term

corresponding sets of graded pushdowns and variable (hoping that the context will make clear that we use graded objects). We also denote by $k\text{-term}(\Gamma, \mathcal{U})$, $k\text{-vterm}(\Gamma, \mathcal{U})$ the sets of graded terms and variable terms.

Example 2. Let $\Gamma = \{A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2, C_3, D_1, D_2, D_3\}$, $\mathcal{U} = \{\Omega_1, \Omega'_1, \Omega_2, \Omega'_2, \Omega_3, \Omega'_3\}$ be graded alphabets: for every $i \in [1, 3]$,

$$\Gamma_i = \{A_i, B_i, C_i, D_i\}, \quad \mathcal{U}_i = \{\Omega_i, \Omega'_i\}.$$

Let

$$T = A_1[A_2[A_3[\varepsilon]\Omega_3[\varepsilon]]B_2[D_3[\varepsilon]C_3[\varepsilon]]]\Omega_1[\varepsilon], \quad T' = A_1[A_1[A_3[\varepsilon]]].$$

$$T'' = A_2[A_3[\varepsilon]B_3[\varepsilon]\Omega_3]\Omega_2, \quad T''' = A_3[\varepsilon]B_3[\varepsilon]\Omega_3$$

$$U = A_1[\Omega_2[A_3[\varepsilon]\Omega_3]], \quad U' = A_1[A_2[A_3[\varepsilon]\Omega_2]]$$

$$U'' = A_1[A_2[\varepsilon]B_2[\varepsilon]]$$

Here T is a graded 3-term,

T' is not a graded 3-term because A_1 occurs at level 2 in T ,

T'' is a graded 2-term, T''' is a graded 1-term.

U is not a graded term because Ω_2 occurs at a non-leaf position.

U' is not a graded 3-term because Ω_2 occurs at level 3 in U' .

U'' is a graded 3-term, though it has depth only 2.

For the same reasons pTq is a graded 3-variable-term while $pT'q$, pUq , $pU'q$ are not 3-variable-terms.

Substitutions Given $T[\Omega_1, \dots, \Omega_n] \in k\text{-term}(\Gamma, \mathcal{U})$, and $H_1, \dots, H_n \in k'\text{-term}(\Gamma, \mathcal{U})$, we denote by $T[H_1/\Omega_1, \dots, H_n/\Omega_n]$ the $(k + k' - 1)$ -term obtained by substituting H_i to Ω_i in T . Note that the new $(k + k' - 1)$ -pushdown thus obtained is really a term. The map $T \mapsto T[H_1/\Omega_1, \dots, H_n/\Omega_n]$ is extended as an homomorphism $(k\text{-vterm}(\Gamma, \mathcal{U}))^* \rightarrow ((k + k' - 1)\text{-vterm}(\Gamma, \mathcal{U}))^*$ by: for every $T \in k\text{-term}(\Gamma, \mathcal{U}) \setminus \{\varepsilon\}$, $p, q, \in Q$:

$$(p, T, q)[H_1/\Omega_1, \dots, H_n/\Omega_n] := (p, T[H_1/\Omega_1, \dots, H_n/\Omega_n], q).$$

When the alphabets Γ, \mathcal{U} are graded, if the substitution replaces every undetermined $\Omega_\lambda \in \mathcal{U}_{j_\lambda}$ by a term $H_\lambda \in (k - j_\lambda + 1)\text{-term}(\Gamma, \mathcal{U})$, then the result $T[H_\lambda/\Omega_\lambda, \lambda \in [1, A]]$ is again a (graded) k -term.

Example 3. Let $\Gamma = \{A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2, C_3, D_1, D_2, D_3\}$, $\mathcal{U} = \{\Omega_1, \Omega'_1, \Omega_2, \Omega'_2, \Omega_3, \Omega'_3\}$. Let

$$T = A_1[A_2[A_3[\varepsilon]\Omega_3[\varepsilon]]B_2[D_3[\varepsilon]C_3[\varepsilon]]]\Omega_1[\varepsilon].$$

Using the short notation:

$$T = A_1[A_2[A_3\Omega_3]B_2[D_3C_3]]\Omega_1.$$

Let

$$H_1 := B_1[B_2[\Omega_3]], \quad H_2 := C_2[A_3B_3\Omega'_3], \quad H_3 := C_3C_3C_3\Omega_3$$

Then

$$T[H_1/\Omega_1, H_2/\Omega_2, H_3/\Omega_3] = A_1[A_2[A_3C_3C_3C_3\Omega_3]B_2[D_3C_3]]B_1[B_2[\Omega_3]]$$

Let $w := (p, T, q)(q, A_1[\Omega_2], p)$. Then

$$w[H_1/\Omega_1, H_2/\Omega_2, H_3/\Omega_3] = (p, A_1[A_2[A_3C_3C_3C_3\Omega_3]B_2[D_3C_3]]B_1[B_2[\Omega_3]], q)(q, A_1[C_2[A_3B_3\Omega'_3]], p)$$

The following “substitution-principle” is straightforward and will be widely used in our proofs. Given some k -pda \mathcal{A} over a pushdown alphabet included in Γ , we extend the relations $\rightarrow_{\mathcal{A}}^*, \vdash_{\mathcal{A}}^*$ to the pushdown alphabet $\Gamma \cup \mathcal{U}$.

Lemma 1. *Let $\vec{\Omega} = (\Omega_1, \dots, \Omega_n)$, $w \in (k - \text{vterm}(\Gamma, \mathcal{U}))^*$, $w' \in \Gamma \hat{\cup} \mathcal{U}^*$. If*

$$w \rightarrow_{\mathcal{A}} w'$$

then

1- $w' \in (k - \text{vterm}(\Gamma, \mathcal{U}))^*$

2- *for every $\vec{H} \in (k' - \text{term}(\Gamma, \mathcal{U}))^n$,*

$$w[\vec{H}/\vec{\Omega}] \rightarrow_{\mathcal{A}} w[\vec{H}/\vec{\Omega}].$$

These properties still hold for every level-partitionned automaton \mathcal{A} and graded alphabets Γ, \mathcal{U} , provided that the levels of \mathcal{A} coincide with the graduations of the alphabets.

In other words:

1- The relation $\rightarrow_{\mathcal{A}}$ saturates the set $(k - \text{vterm}(\Gamma, \mathcal{U}))^*$.

2- Every substitution preserves the derivation relation.

By induction over the integer m , it also preserves all the relations $\rightarrow_{\mathcal{A}}^m$ for $m \geq 0$.

Point (1) is straightforward. The key-idea for point (2) is that, as $\Gamma \cap \mathcal{U} = \emptyset$, the symbols Ω_i can be copied or erased during the derivation, but they cannot influence the sequence of rules used in that derivation.

k-computable mappings

Definition 5 (k -computable mapping). *A mapping $f : A^* \mapsto B^*$ is called k -computable iff there exists a strongly deterministic k -pda \mathcal{A} , over a pushdown-alphabet Γ which is level-partitionned, such that Γ contains $k - 1$ symbols $\gamma_1 \in \Gamma_1, \dots, \gamma_i \in \Gamma_i, \dots, \gamma_{k-1} \in \Gamma_{k-1}$, a state q_0 , the alphabet A is a subset of Γ_k and for all $w \in A^*$:*

$$(q_0, f(w), \gamma_1[\gamma_2 \dots [\gamma_{k-1}[w]] \dots]) \vdash_{\mathcal{A}}^* (q_0, \varepsilon, \varepsilon).$$

One denotes by $\mathbb{S}_k(A^, B^*)$ the set of k -computable mappings from A^* to B^* .*

The particular case where $\text{Card}(A) = \text{Card}(B) = 1$ was studied in [FS06].

2.6 Regular sets of pushdowns of level k

The set $k\text{-pds}(\Gamma)$ of all pushdowns of level k over the alphabet Γ , can be considered as a language over the following *extended alphabet* $\hat{\Gamma}$:

$$\hat{\Gamma} := \Gamma \cup \{x, \bar{x}\},$$

where x (resp. \bar{x}) denotes an opening (resp. closing) square bracket. From this point of view the 3 – pds of example 1

$$\omega = A_1[A_2[A_3C_3]B_2[D_3C_3]]B_1[B_2[B_3D_3]].$$

is described by the word

$$A_1xA_2xA_3C_3\bar{x}B_2xD_3C_3\bar{x}\bar{x}B_1xB_2xB_3D_3\bar{x}\bar{x} \in \hat{\Gamma}^*.$$

One can check that, thus turned into a set of words, $k\text{-pds}(\Gamma)$ is a regular subset of $\hat{\Gamma}^*$. Let $\mathcal{A} = (Q, B, \Gamma, \delta)$ be some k – pda. For every states $p, q \in Q$, we define the language

$$P_{p,q}(\mathcal{A}) := \{w \in k\text{-pds}(\Gamma) \mid \exists u \in B^*, (p, w, q) \rightarrow_{\mathcal{A}}^* u\}.$$

Lemma 2. *Let $\mathcal{A} = (Q, B, \Gamma, \delta)$ be some k – pda. There exists a congruence of finite index $\equiv_{\mathcal{A}}$ over $\hat{\Gamma}^*$ such that:*

- $\equiv_{\mathcal{A}}$ saturates the language $k\text{-pds}(\Gamma)$
- for every $p, q \in Q$, $\equiv_{\mathcal{A}}$ saturates the language $P_{p,q}(\mathcal{A})$.

Proof. According to Theorem 1, p. 218 of [HO07], for every pda of order k , the set of co-accessible configurations (and, more generally, the set of all configurations that are co-accessible from a given regular set of k -pushdowns) is a regular subset of $k\text{-pds}(\Gamma)$. This shows that every set $P_{p,q}(\mathcal{A})$ is regular.

Let \sim (resp. $\sim_{\mathcal{A},p,q}$) be the syntactic congruence of the language $k\text{-pds}(\Gamma)$ (resp. of the language $P_{p,q}(\mathcal{A})$). Let us define

$$\equiv_{\mathcal{A}} := \sim \cap \left(\bigcap_{(p,q) \in Q \times Q} \sim_{\mathcal{A},p,q} \right).$$

Since each binary relation $\sim, \sim_{\mathcal{A},p,q}$ is a congruence of finite index, $\equiv_{\mathcal{A}}$ is a congruence of finite index (over the monoid $\hat{\Gamma}^*$) and it has the required saturation properties.

2.7 Recurrences in monoids

When considering mappings into *words* instead of integers, one is lead to consider recurrent relations based on the *concatenation* operation. Let us define a more general notion of mapping defined by recurrent relations based on the *product* operation in some arbitrary monoid $(M, \cdot, 1)$.

Definition 6 (recurrent relations in M). Given a finite set I and a family of mappings indexed by I , $f_i : A^* \rightarrow M$ (for $i \in I$), we call system of recurrent relations in \mathbb{M} over the family $(f_i)_{i \in I}$, a system of the form

$$f_i(aw) = \prod_{j=1}^{\ell(i,a)} f_{\alpha(i,a,j)}(w) \text{ for all } i \in I, a \in A, w \in A^*$$

where $\ell(a, i) \in \mathbb{N}$, $\alpha(i, a, j) \in I$ and the symbol \prod stands for the extension of the binary product in \mathbb{M} to an arbitrary finite number of arguments.

When the monoid \mathbb{M} is a finitely generated free monoid, such a system is called a system of *catenative* recurrent relations. The free monoid is used in the characterisation below of mappings of level 2, while the monoid $(\text{HOM}(C^*, C^*), \circ, \text{Id})$ will be suitable for studying mappings of level 3. As an intermediate tool, we shall use the following more general notion.

Definition 7 (regular recurrent relations). Let \equiv be a congruence of finite index on X^* . Let $\mathcal{C} = X^* / \equiv$, let I be a finite set and $f_i : X^* \rightarrow M$ (for $i \in I$) be a family of mappings. We call system of regular recurrent relations in \mathbb{M} over the family $(f_i)_{i \in I}$ a system of the form

$$f_i(xw) = \prod_{j=1}^{\ell(i,x,c)} f_{\alpha(i,x,j,c)}(u_{i,x,c}w) \text{ for all } i \in I, x \in X, c \in \mathcal{C}, w \in c$$

where $\ell(i, x, c) \in \mathbb{N}$, $\alpha(i, x, j, c) \in I$, $u_{i,x,c} \in X^*$.

The system is a system of *recurrent relations* (i.e. meets the conditions of Definition 6) when all the words $u_{i,x,c}$ have null length and the congruence \equiv is maximal i.e. has only one class. The system is called *noetherian* when the term-rewriting system consisting of all the oriented rules

$$f_i(xw) \rightarrow \prod_{j=1}^{\ell(i,x,c)} f_{\alpha(i,x,j,c)}(u_{i,x,c}w)$$

for all $i \in I, x \in X, c \in \mathcal{C}, w \in c$, is noetherian (in this rewriting system, both sides are understood as formal terms ²). The system is called *strict* when all the words $u_{i,x,c}$ are empty.

Clearly every strict system is also noetherian. In particular every system of recurrent relations is a noetherian system of regular recurrent relations, but the converse does not hold in general.

² i.e. both sides of the rules are elements of the free monoid generated by the (infinite) set of terms of depth 1 $\{f_i(w) \mid i \in I, w \in X^*\}$; the system is thus a so-called *monadic* semi-Thue system over this free monoid)

2.8 Word-mappings of level 2

Let us recall a notion which originated in computational biology, but turns out to be useful in general formal language theory (see [KRS97]).

Definition 8 (HDT0L sequences). *Let $f : A^* \rightarrow B^*$. The mapping f is called a HDT0L sequence iff there exists a finite alphabet C , a homomorphism $H : A^* \rightarrow \text{HOM}(C^*, C^*)$, an homomorphism $h \in \text{HOM}(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$*

$$f(w) = h(H^w(c)).$$

(here we denote by H^w the image of w by H). The mapping f is called a DT0L when $B = C$ and the homomorphism h is just the identity; f is called a HD0L when A is reduced to one element.

The following characterisation of word-mappings of level 2 is proved in [FMS14].

Theorem 1. *Let us consider a mapping $f : A^* \rightarrow B^*$. The following properties are equivalent:*

- 1- $f \in \mathbb{S}_2(A^*, B^*)$
- 2- *There exists a finite family $(f_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow B^*$ which fulfils a noetherian system of regular recurrent relations of order 1 in $\langle B^*, \cdot, \varepsilon \rangle$ and such that $f = f_1$*
- 3- *There exists a finite family $(f_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow B^*$ which fulfils a strict system of regular recurrent relations in $\langle B^*, \cdot, \varepsilon \rangle$ and such that $f = f_1$*
- 4- f is a HDT0L sequence.
- 5- *There exists a finite family $(f_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow B^*$ which fulfils a system of catenative recurrent relations and such that $f = f_1$.*

Proof. Lemma 27 of [FMS14] proves that:

$$(1) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (1).$$

Implication $(3) \Rightarrow (2)$ is obvious.

Let us suppose that (2) holds:

$$f_i(aw) = \prod_{j=1}^{\ell(i, a)} f_{\alpha(i, a, j)}(w) \text{ for all } i \in I, a \in A, w \in A^*.$$

Let us denote by $R_{i, a}(w)$ the righthand-side corresponding to $f_i(aw)$ on the lefthand-side. Let us denote by \mathcal{S} the rewriting system consisting of the set of rules

$$f_i(a \cdot \Omega) \rightarrow R_{i, a}(\Omega)$$

where Ω is a variable. Since this system \mathcal{S} is noetherian, there exists some irreducible expression $S_{i, a}(\Omega)$ obtained from $R_{i, a}(\Omega)$ by a derivation (mod \mathcal{S}). The new system of relations:

$$f_i(aw) = S_{i, a}(w) \text{ for all } i \in I, x \in X, c \in \mathcal{C}, w \in c$$

is a strict system of regular recurrent relations in $\langle B^*, \cdot, \varepsilon \rangle$ which is fulfilled by the family $(f_i)_{i \in I}$. Hence (2) \Rightarrow (3).

3 Word-mappings of level 3

Let us prove the main result of this paper.

Theorem 2. *Let us consider a mapping $f : A^* \rightarrow B^*$. The following properties are equivalent:*

- 1- $f \in \mathbb{S}_3(A^*, B^*)$
- 2- *There exists a finite family $(H_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \text{HOM}(C^*, C^*)$ which fulfils a system of recurrent relations in $\langle \text{HOM}(C^*, C^*), \circ, \text{Id} \rangle$, an element $h \in \text{HOM}(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$:*

$$f(w) = h(H_1(w)(c)).$$

- 3- *f is a composition of a DTOL sequence $g : A^* \rightarrow C^*$ by a HDTOL sequence $h : C^* \rightarrow B^*$.*

We prove this theorem through the more technical statement

Theorem 3. *Let us consider a mapping $f : A^* \rightarrow B^*$. The following properties are equivalent:*

- 1- $f \in \mathbb{S}_3(A^*, B^*)$
- 2- *There exists a finite family $(H_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \text{HOM}(C^*, C^*)$ which fulfils a noetherian system of regular recurrent relations in $(\text{HOM}(C^*, C^*), \circ, \text{Id})$, an element $h \in \text{HOM}(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$:*

$$f(w) = h(H_1(w)(c)).$$

- 3- *There exists a finite family $(H_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \text{HOM}(C^*, C^*)$ which fulfils a strict system of regular recurrent relations in $(\text{HOM}(C^*, C^*), \circ, \text{Id})$, an element $h \in \text{HOM}(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$:*

$$f(w) = h(H_1(w)(c)).$$

- 4- *f is a composition of a DTOL sequence $g : A^* \rightarrow C^*$ by a HDTOL sequence $h : C^* \rightarrow B^*$.*

- 5- *There exists a finite family $(H_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \text{HOM}(C^*, C^*)$ which fulfils a system of recurrent relations in $(\text{HOM}(C^*, C^*), \circ, \text{Id})$, an element $h \in \text{HOM}(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$:*

$$f(w) = h(H_1(w)(c)).$$

We first prove that $(1) \Rightarrow (2)$, then we prove that $(2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5)$ and finally that $(5) \Rightarrow (1)$.

Let us consider some $f \in \mathbb{S}_3(A^*, B^*)$ and let us show that it fulfils condition (2). Let $\mathcal{A} = (Q, B, \Gamma, \delta, q_0, Z_0)$ be a strongly deterministic 3-iterated pushdown automaton, such that $\mathcal{A} \subseteq \Gamma$ and \mathcal{A} computes f . We can normalize the automaton \mathcal{A} in such a way that:

(LP) \mathcal{A} is level-partitioned (see §2.5)

(RL) the only transitions that read a letter from B are of the form:

$$\delta(p, b, S) = (q, pop_1)$$

for some $p, q \in Q, b \in B, S \in \Gamma_1$.

(PI) the only push intructions occuring in δ are of the form $\text{push}_j(\gamma)$ with $\gamma \in \Gamma^2$.

Let us define from this automaton a finite alphabet $\hat{\mathcal{W}}$ and a family of mappings $H_i : \Gamma^* \rightarrow \text{HOM}(\hat{\mathcal{W}}^*, \hat{\mathcal{W}}^*)$ that is closely related to the computations of \mathcal{A} and fulfils a *noetherian system of regular recurrent relations of order 1 in* $\text{HOM}(\hat{\mathcal{W}}^*, \hat{\mathcal{W}}^*)$.

General idea The general idea that guides our proof of $(1) \Rightarrow (2)$ in Theorem 3 is as follows. Every 2 – pds atom $T[w]$ (where T has level 2 and w is a product of symbols of level 3) *acts* on the set of 3 – pds by:

$$S[u] \mapsto S[T[w]u].$$

We represent all the possible $S[u]$ by a finite alphabet \mathcal{W} , in such a way that the family of actions of $T[w]$ (for $T \in \Gamma_2$) is represented by a single homomorphism $H_i^w : \mathcal{W}^* \rightarrow \mathcal{W}^*$. The final outcome is that, to every word $w \in \Gamma_3^*$ is associated a finite family of homomorphisms $H_i^w : \mathcal{W}^* \rightarrow \mathcal{W}^*$ with $i \in I$, where I is a finite set of indices that encode the various letters $T \in \Gamma_2$ and various subsets of \mathcal{W} on which the action is restricted (with additional homomorphisms that do not depend on the argument w but ease the definition of a compositional recurrence that links the H_i^{aw} with the H_i^w for letters $a \in \Gamma_3$, see the defining equations (32,33)).

3.1 Alphabets E and \mathcal{W}

Congruences Let $\equiv_{\mathcal{A}}$ be a finite-index congruence over $\hat{\Gamma}^*$ given by Lemma 2 of subsection 2.6. For every integer $j \in [1, 3]$, we denote by $(\equiv_{\mathcal{A}, j})$ the restriction of $\equiv_{\mathcal{A}}$ over $(4 - j) - \text{pds}(\Gamma)$:

$$\forall w, w' \in (4 - j) - \text{pds}(\Gamma), \quad w \equiv_{\mathcal{A}, j} w' \Leftrightarrow w \equiv_{\mathcal{A}} w'.$$

The following facts are immediate:

$$\begin{aligned}
& \text{if } w \equiv_{\mathcal{A},1} w', \text{ then } \forall p, q \in Q, L(\mathcal{A}, pwq) = \emptyset \Leftrightarrow L(\mathcal{A}, pw'q) = \emptyset, \\
& \text{if } w \equiv_{\mathcal{A},2} w', \quad \quad \quad \text{then } \forall S \in \Gamma_1, S[w] \equiv_{\mathcal{A},1} S[w'], \\
& \text{if } w \equiv_{\mathcal{A},3} w', \quad \quad \quad \text{then } \forall T \in \Gamma_2, T[w] \equiv_{\mathcal{A},2} T[w'].
\end{aligned} \tag{8}$$

From now on, we abbreviate each congruence $\equiv_{\mathcal{A},i}$ by \equiv_i .

Alphabet E Let

$$E := 2 - \text{pds}(\Gamma) / \equiv_2.$$

In the sequel we use E as a graded alphabet of undeterminates, where the level of all the symbols is 2.

Alphabet W We define the finite alphabet

$$\mathcal{W} := \{(pS[e]q) \mid p, q \in Q, S \in \Gamma_1, e \in 2 - \text{pds}(\Gamma) / \equiv_2\}$$

Note that \mathcal{W} is a finite subset of the infinite alphabet $3 - \text{vterm}(\Gamma, E)$.

3.2 Morphisms H_i^w

We use letters $S, S_1, \dots, S_n, \dots$ to denote elements of Γ_1 , $T, T_1, \dots, T_n, \dots$ to denote elements of Γ_2 , $a, a_1, \dots, a_n, \dots$ to denote elements of Γ_3 , $\Omega, \Omega_1, \dots, \Omega_n, \dots$ to denote undeterminates other than those in E i.e. symbols which are not elements of Γ (see §2.5 of section 2). Let

$$I_0 := \mathcal{P}(\mathcal{W}) \times \Gamma_2.$$

Let $i \in I_0$: it has the form $i = (\mathcal{V}, T)$ where $\mathcal{V} \subseteq \mathcal{W}, T \in \Gamma_2$. We also denote by $\text{Alph}(i)$ the first component \mathcal{V} of i . For every $w \in \Gamma_3^*$, we define the homomorphism H_i^w by:
for every letter $W = (pS[e]q) \in \mathcal{W}$, if

$$W \in \mathcal{V} \tag{9}$$

and

$$(pS[T[w]\Omega]q) \rightarrow_{\mathcal{A}}^+ \prod_{j=1}^{\ell(i,W)} (p_{i,j} S_{i,j} [\Omega] q_{i,j}) \tag{10}$$

and

$$\forall j \in [1, \ell(i, W)], \forall t \in e, L(p_{i,j} S_{i,j} [t] q_{i,j}) \neq \emptyset \tag{11}$$

then

$$H_i^w : (pS[e]q) \mapsto \prod_{j=1}^{\ell(i,W)} (p_{i,j} S_{i,j} [e] q_{i,j}). \tag{12}$$

otherwise

$$H_i^w : (pS[e]q) \mapsto (pS[e]q). \quad (13)$$

We denote by H_i the mapping: $\Gamma^* \rightarrow \text{HOM}(\mathcal{W}^*, \mathcal{W}^*)$ defined by $w \mapsto H_i^w$.

Lemma 3. *For every $i = (\mathcal{V}, T) \in I_0$, word $w \in \Gamma^*$ and letter $W \in \mathcal{V}$, there exists at most one possible righthandside for the equation (10) that also satisfies equation (11).*

For every $i \in I_0$, we denote by $\text{Alph}(i, w)$ the set of letters $W \in \text{Alph}(i)$, such that there exists exactly one righthandside for the equation (10) that satisfies also equation(11). Let us remark that, if $w \in \Gamma^* \setminus \Gamma_3^*$, then $\text{Alph}(i, w) = \emptyset$. For this reason most of our further statements are trivial for words in $\Gamma^* \setminus \Gamma_3^*$.

Lemma 4. *For every $w, w' \in \Gamma_3^*$, if $w \equiv_3 w'$, then, for every $i \in I_0$,*

$$\text{Alph}(i, w) = \text{Alph}(i, w').$$

Proof. Suppose that $i = (\mathcal{V}, T)$. Let $W = pS[e]q$ and suppose that

$$W \in \mathcal{V}. \quad (14)$$

We observe that $pS[e]q \in \text{Alph}(i, w)$ iff

$$\forall u \in e, \quad L(\mathcal{A}, pS[T[w]u]q) \neq \emptyset \quad (15)$$

Let $w, w' \in \Gamma_3^*$ such that $w \equiv_3 w'$. By the three facts (9) of subsection 3.1, for every $u \in e$:

$$T[w]u \equiv_2 T[w']u$$

hence

$$S[T[w]u] \equiv 1S[T[w']u],$$

hence

$$L(\mathcal{A}, pS[T[w]u]q) = \emptyset \Leftrightarrow L(\mathcal{A}, pS[T[w']u]q) = \emptyset.$$

Using observation (15) we can conclude that, under the hypothesis that $(W \in \mathcal{V} \ \& \ w \equiv_3 w')$ it is true that

$$W \in \text{Alph}(i, w) \Leftrightarrow W \in \text{Alph}(i, w'). \quad (16)$$

Under the hypothesis that $W \notin \mathcal{V}$, we obtain that $W \notin \text{Alph}(i, w) \ \& \ W \notin \text{Alph}(i, w')$ which implies again that

$$W \in \text{Alph}(i, w) \Leftrightarrow W \in \text{Alph}(i, w'). \quad (17)$$

By (16,17) the lemma is proved.

Owing to Lemma 4, for every $i \in I_0$ and every class $d \in \Gamma_3^* / \equiv_3$, we denote by $\text{Alph}(i, d)$ the set $\text{Alph}(i, w)$ for any $w \in d$.

3.3 Mixed recurrence

Lemma 5. *Let $i \in I_0$ ($i = (\mathcal{V}, T)$), $d \in 1 - \text{pds}(\Gamma)/\equiv_3$, $a \in \Gamma_3$, $w \in 1 - \text{pds}(\Gamma_3)$ and $W \in \mathcal{W}$ ($W = pS[e]q$). There exists an integer $\ell(i, a, d, W) \in [1, 2]$ and for every $j \in [1, \ell(i, a, d, W)]$ there exist indices $\alpha(i, a, d, W, j), \beta(i, a, d, W, j) \in I_0$, words $\omega_{i,a,d,W,j} \in \Gamma_3^{\leq 2}$, letters $V_{i,a,d,W,j} \in \mathcal{W}$ and an alphabetic homomorphism $\Phi_{i,a,d,W,j} \in \text{HOM}(\mathcal{W}^*, \mathcal{W}^*)$ such that*

$$w \in d \Rightarrow H_i^{aw}(W) = \prod_{j=1}^{\ell(i,a,d,W)} H_{\alpha(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} \circ \Phi_{i,a,d,W,j} \circ H_{\beta(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w}(V_{i,a,d,W,j}) \quad (18)$$

and one of the following cases occurs:

Case 1.1: $\ell(i, a, d, W) = 1, \alpha(i, a, d, W, 1) = (\emptyset, T), \beta(i, a, d, W, 1) = (\emptyset, T)$.

Case 1.2: $\ell(i, a, d, W) = 1, \alpha(i, a, d, W, 1) = (\{V_{i,a,d,W,1}\}, T'), \Phi_{i,a,d,W,1} = \text{Id}_{\mathcal{W}^*}, \beta(i, a, d, W, 1) = (\emptyset, T)$ and

$$W[T[aw]e/e] \rightarrow_{\mathcal{A}} V_{i,a,d,W,1}[T'[\omega_{i,a,d,W,1}w]e/e] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W).$$

Case 1.3: $\ell(i, a, d, W) = 1, \alpha(i, a, d, W, 1) = (\{V_{i,a,d,W,1}\}, T')$,

$$e' = [T''[\omega_{i,a,d,W,j} \cdot w]]_{\equiv_{\mathcal{A},2}} \cdot e,$$

$$\Phi_{i,a,d,W,1} : p'S'[e'']q' \mapsto p'S'[e']q' \text{ if } e'' \neq e', \quad p'S'[e']q' \mapsto p'S'[e]q',$$

$\beta(i, a, d, W, 1) = (\text{Alph}(\Phi_{i,a,d,W,1}(H_{\alpha(i,a,d,W,1)}^{\omega_{i,a,d,W,1} \cdot w}(V_{i,a,d,W,1}))), T''), V_{i,a,d,W,1} = p'S[e']q$ for some $p' \in Q$, and

$$W[T[aw]e/e] \rightarrow_{\mathcal{A}} V_{i,a,d,W,1}[T'[\omega_{i,a,d,W,1}w]T''[\omega_{i,a,d,W,1}w]e/e'] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W)$$

Case 2: $\ell(i, a, d, W) = 2, \omega_{i,a,d,W,1} = \omega_{i,a,d,W,2} = a, \alpha(i, a, d, W, j) = (\{V_{i,a,d,W,j}\}, T), \Phi_{i,a,d,W,j} = \text{Id}_{\mathcal{W}^*}, \beta(i, a, d, W, j) = (\emptyset, T)$ and

$$W[T[aw]e/e] \rightarrow_{\mathcal{A}}^+ V_{i,a,d,W,1}[T[aw]e/e] \cdot V_{i,a,d,W,2}[T[aw]e/e] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W).$$

Let us sketch a proof of this lemma. In the case of an operation $push_1$, the value of $\ell(i, a, d, W)$ is 2 while the case pop_1 is impossible and for the four remaining cases ($\text{pop}_2, \text{pop}_3, \text{push}_2, \text{push}_3$) the value of $\ell(i, a, d, W)$ is 1. In the case of an operation $push_1$, the two letters $V_{i,a,d,W,1}, V_{i,a,d,W,2}$ depend on an intermediate state which is determined by the congruence class d .

In the case of an operation $push_2$, the composition of two non-trivial homomorphisms is really required in the righthandside. In the other cases we can choose $\Phi_{i,a,d,W,j} := \text{Id}_{\mathcal{W}^*}$ and $\beta(i, a, d, W, j) := (\emptyset, T)$, so that, for every u , $H_{\beta(i,a,d,W,j)}^u = \text{Id}_{\mathcal{W}^*}$.

Let us proceed now to the formal proof.

Proof. For $W \notin \text{Alph}(i, ad)^3$ we just choose $\ell(i, a, d, W) := 1, \omega(i, a, d, W, 1) := \varepsilon$ and $\alpha(i, a, d, W, 1) := (\emptyset, T), \Phi_{i,a,d,W,1} := \text{Id}_{\mathcal{W}^*}, \beta(i, a, d, W, 1) := (\emptyset, T), V_{i,a,d,W,1} := W$. These choices ensure that

$$H_{\alpha(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} \circ \Phi_{i,a,d,W,1} \circ H_{\beta(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} = \text{Id}_{\mathcal{W}^*}.$$

Equation (18) thus holds and Case 1.1. is realized.

Let us now suppose that $W \in \text{Alph}(i, ad)$. Since \mathcal{A} is strongly deterministic, there is at most one element $\bar{b} \in B \cup \{\varepsilon\}$ such that a transition $\delta(p, \bar{b}, STa)$ is defined. The fact $W \in \text{Alph}(i, ad)$ implies that $L(\mathcal{A}, pS[T[aw]u]q) \neq \emptyset$ for every $u \in e$, hence that there exists exactly one such transition. At last, the restriction (RL) that is assumed for \mathcal{A} shows that this transition has the form:

$$\delta(p, \varepsilon, STa) = (p_1, op) \text{ where } op \in PUSH(\Gamma) \cup \{\text{pop}_2, \text{pop}_3\}$$

We distinguish several cases corresponding to these five possible values of op .

pop₂: $pS[T[aw]e]q \rightarrow p_1S[e]q$.

Let us choose: $\ell(i, a, d, W) := 1, \alpha(i, a, d, W, j) := (\emptyset, T), \Phi_{i,a,d,W,j} := \text{Id}_{\mathcal{W}^*}, \beta(i, a, d, W, j) := (\emptyset, T), \omega_{i,a,d,W,1} := \varepsilon, V_{i,a,d,W,1} := p_1S[e]q$.

These choices ensure that the value of both sides of (18) is $p_1S[e]q$ and that Case 1.1 is realized.

pop₃: $pS[T[aw]e]q \rightarrow p_1S[T[w]e]q$.

Let us choose: $\ell(i, a, d, W) := 1, \alpha(i, a, d, W, j) := (\emptyset, T), \Phi_{i,a,d,W,j} := \text{Id}_{\mathcal{W}^*}, \beta(i, a, d, W, j) := (\emptyset, T), \omega_{i,a,d,W,1} := \varepsilon, V_{i,a,d,W,1} := p_1S[e]q$.

These choices ensure that equation (18) holds and that Case 1.2 is realized.

push₁(S_1S_2):

Since $L(\mathcal{A}, pS[T[aw]u]q) \neq \emptyset$ (for a given $u \in e$) and $pS[T[aw]u]q \rightarrow_{\mathcal{A}} p_1S_1[T[aw]e]S_2[T[aw]e]q$, there exists a unique $r \in Q$ such that $L(\mathcal{A}, p_1S_1[T[aw]u]r) \neq \emptyset$ and $L(\mathcal{A}, rS_2[T[aw]u]q) \neq \emptyset$.

This state r depends on d, e only and not really on w, u (by definition of the equivalences \equiv_j for $j \in [1, 3]$ and by the property that they are congruences).

We also have

$$pS[T[aw]u]q \rightarrow_{\mathcal{A}}^2 (p_1S_1[T[aw]e]r)(rS_2[T[aw]e]q).$$

Let us choose $\ell(i, a, d, W) := 2, V_{i,a,d,W,1} := p_1S_1[e]r, V_{i,a,d,W,2} := rS_2[e]q, \alpha(i, a, d, W, j) := (\{V_{i,a,d,W,j}\}, T), \Phi_{i,a,d,W,j} := \text{Id}_{\mathcal{W}^*}, \beta(i, a, d, W, j) := (\emptyset, T), \omega_{i,a,d,W,1} := \omega_{i,a,d,W,2} := a$. Equation (18) holds and Case 2 is realized.

push₂($T'T''$): In this case

$$pS[T[aw]e]q \rightarrow p_1S[T'[aw]T''[aw]e]q. \quad (19)$$

Let us choose: $\ell(i, a, d, W) := 1, e' := [T[aw]]_{\equiv 2} \cdot e, V_{i,a,d,W,1} := p_1S[e']q, \alpha(i, a, d, W, 1) := (\{p_1S[e']q\}, T'), \Phi_{i,a,d,W,1} : p'S'[e'']q' \mapsto p'S'[e'']q'$ if $e'' \neq e'$, $p'S'[e']q' \mapsto p'S'[e']q', \beta(i, a, d, W, 1) := (\text{Alph}(\Phi_{i,a,d,W,1}(H_{\alpha(i,a,d,W,1)}^{aw}(p_1S[e']q))), T''), \omega_{i,a,d,W,1} := a$.

Nota: the action of $\Phi_{i,a,d,W,1}$ on letter $p'S'[e'']q'$ with $e'' \neq e'$ has no influence on

³ i.e. with full rigor $\text{Alph}(i, [a]_{\equiv 3} \cdot d)$

equation (18).

In order to lighten the notation we abbreviate $\alpha(i, a, d, W, 1)$ as α , $\beta(i, a, d, W, 1)$ as β and $\Phi_{i,a,d,W,1}$ as Φ in the rest of this proof. Let us show that equation (18) holds. By definition of the mappings H_i we have:

$$p_1 S[T'[aw]e']q \rightarrow_{\mathcal{A}}^* H_{\alpha}^{aw}(p_1 S[e']q).$$

Applying the substitution $[T''[aw]e/e']$ on both sides of this derivation we obtain

$$p_1 S[T'[aw]T''[aw]e]q \rightarrow_{\mathcal{A}}^* (H_{\alpha}^{aw}(p_1 S[e']q))[T''[aw]e/e']. \quad (20)$$

Let us consider an arbitrary letter V occuring in the word $H_{\alpha}^{aw}(p_1 S[e']q)$: it has the form $V = p'S'[e']q'$ for some $p', q' \in Q, S' \in \Gamma_1$. By definition of the mappings H_i we have:

$$p'S'[T''[aw]e]q' \rightarrow_{\mathcal{A}}^* H_{\beta}^{aw}(p'S'[e]q').$$

This derivation can also be written as:

$$V[T''[aw]e/e'] \rightarrow_{\mathcal{A}}^* H_{\beta}^{aw}(V[e/e']) = H_{\beta}^{aw}(\Phi(V)).$$

Since the above derivation is valid for every letter V occuring in $H_{\alpha}^{aw}(p_1 S[e']q)$, we conclude that

$$H_{\alpha}^{aw}(p_1 S[e']q)[T''[aw]e/e'] \rightarrow_{\mathcal{A}}^* H_{\beta}^{aw}(\Phi(H_{\alpha}^{aw}(p_1 S[e']q))). \quad (21)$$

The product of the three derivations (19), (20) and (21) gives a derivation

$$pS[T[aw]e]q \rightarrow_{\mathcal{A}}^* H_{\beta}^{aw}(\Phi(H_{\alpha}^{aw}(p_1 S[e']q))). \quad (22)$$

Let us check now that the extremity of derivation (22) is nothing else than $H_i^{aw}(pS[e]q)$. We introduce, for every congruence class $f \in 2 - \text{pds}(\Gamma) \equiv 2$ the subalphabets:

$$\mathcal{W}_f := \{p'S'[f]q' \mid p', q' \in Q, S' \in \Gamma_1\},$$

$$\mathcal{W}_f^0 := \{p'S'[f]q' \mid p', q' \in Q, S' \in \Gamma_1, \forall u \in f, L(\mathcal{A}, p'S'[u]q') \neq \emptyset\}.$$

Since $pS[e]q \in \text{Alph}(i, aw)$, we know that

$$\forall u \in e, L(\mathcal{A}, pS[T[aw]u]q) \neq \emptyset.$$

Every derivation (modulo $\rightarrow_{\mathcal{A}}$) starting from $pS[T[aw]u]q$ has, as its first step, derivation (19). Hence

$$\forall u \in e, L(\mathcal{A}, p_1 S[T'[aw]T''[aw]u]q) \neq \emptyset.$$

By definition of e' , $\forall u \in e, T''[aw]u \in e'$. This ensures that,

$$\forall u' \in e', L(\mathcal{A}, p_1 S[T'[aw]u']q) \neq \emptyset.$$

Since $\alpha = (\{p_1S[e']q\}, T')$, the above non-emptiness assertion shows that

$$p_1S[e']q \in \text{Alph}(\alpha, aw).$$

By definition of H_α^{aw} we have

$$H_\alpha^{aw}(p_1S[e']q) \in (\mathcal{W}_{e'}^0)^*. \quad (23)$$

Let $V \in \mathcal{W}$ be factor of $H_\alpha^{aw}(p_1S[e']q)$. By (23) $V \in \mathcal{W}_{e'}^0$, in particular it has the form $V = p'S'[e']q'$. Since $\forall u \in e, T''[aw]u \in e'$, we then have

$$\forall u \in e, L(\mathcal{A}, p'S'[T''[aw]u]q') \neq \emptyset,$$

hence $p'S'[e]q' \in \text{Alph}(\beta, aw)$, or, equivalently, $\Phi(V) \in \text{Alph}(\beta, aw)$. It follows that

$$\Phi(H_\alpha^{aw}(p_1S[e']q) \in \text{Alph}(\beta, aw)^*$$

hence

$$H_\beta^{aw}(\Phi(H_\alpha^{aw}(p_1S[e']q))) \in (\mathcal{W}_e^0)^*. \quad (24)$$

By (22) and (24),

$$H_i^{aw}(pS[e]q) = H_\beta^{aw}(\Phi(H_\alpha^{aw}(p_1S[e']q)))$$

i.e. equation (18) is valid. The conditions of Case 1.3 are also valid.

push₃(ω): $pS[T[aw]e]q \rightarrow p_1S[T[\omega \cdot w]e]q$.

Let us choose: $\ell(i, a, d, W) := 1$, $V_{i,a,d,W,1} := p_1S[e]q$, $\omega(i, a, d, W, 1) := \omega$, $\alpha(i, a, d, W, 1) := (\{p_1S[e]q\}, T)$, $\Phi_{i,a,d,W,1} := \text{Id}_{\mathcal{W}^*}$, $\beta(i, a, d, W, 1) := (\emptyset, T)$. These choices ensure that equation (18) holds and Case 1.2 is realized.

Note that the above recurrence relations (18) involve both the product operation (over \mathcal{W}^*) and the composition operation (over $\text{HOM}(\mathcal{W}^*, \mathcal{W}^*)$). Therefore we call them the *mixed* recurrence relations. We shall transform these mixed recurrence relations into “purely” compositional recurrence relations in next paragraph.

3.4 Compositional recurrence

Let us define a new alphabet $\hat{\mathcal{W}}$ extending \mathcal{W} and a family of homomorphisms $\hat{\mathcal{W}}^* \rightarrow \hat{\mathcal{W}}^*$ that satisfy a system of compositional recurrence relations. Let

$$\hat{\mathcal{W}} := \mathcal{W} \cup (\mathcal{W} \times \mathcal{W} \times \{1, 2\}) \cup \{X\},$$

where X is a new letter not in $\mathcal{W} \cup (\mathcal{W} \times \mathcal{W} \times \{1, 2\})$. For every i, a, d we introduce the following homomorphisms $\hat{\mathcal{W}}^* \rightarrow \hat{\mathcal{W}}^*$:

$\psi_{i,a,d}$:

$$\begin{aligned} X &\mapsto X \\ W &\mapsto (W, W, 1) && \text{if } W \in \mathcal{W} \setminus \text{Alph}(i, ad) \\ W &\mapsto (V_{i,a,d,W,1}, W, 1) && \text{if } W \in \text{Alph}(i, ad) \text{ and } \ell(i, a, d, W) = 1 \\ W &\mapsto (V_{i,a,d,W,1}, W, 1) \cdot (V_{i,a,d,W,2}, W, 2) && \text{if } W \in \text{Alph}(i, ad) \text{ and } \ell(i, a, d, W) = 2 \\ (V, W, j) &\mapsto X && \text{if } V, W \in \mathcal{W}, j \in \{1, 2\} \end{aligned} \quad (25)$$

$\hat{\psi}_{i,a,d}$:

$$\begin{aligned} X &\mapsto X \\ V &\mapsto X \quad \text{if } V \in \mathcal{W} \\ (V, W, j) &\mapsto V \quad \text{if } V, W \in \mathcal{W}, j \in \{1, 2\} \end{aligned} \quad (26)$$

For every i, a, d, W, j we introduce the following homomorphisms $\hat{\mathcal{W}}^* \rightarrow \hat{\mathcal{W}}^*$:
 $\theta_{i,a,d,W,j}$:

$$\begin{aligned} X &\mapsto X \\ V &\mapsto X \quad \text{if } V \in \mathcal{W} \\ (V, W, j) &\mapsto V \quad \text{if } V \in \mathcal{W} \\ (V, W', j') &\mapsto (V, W', j') \quad \text{if } V, W' \in \mathcal{W}, j' \in \{1, 2\}, W' \neq W \text{ or } j' \neq j. \end{aligned} \quad (27)$$

$\hat{\theta}_{i,a,d,W,j}$:

$$\begin{aligned} X &\mapsto X \\ V &\mapsto (V, W, j) \quad \text{if } V \in \mathcal{W} \\ (V, V', j') &\mapsto (V, V', j') \quad \text{if } V, V' \in \mathcal{W}, j' \in \{1, 2\} \end{aligned} \quad (28)$$

We extend the homomorphisms H_i^w to $\hat{\mathcal{W}}^*$ by setting:

$$\begin{aligned} X &\mapsto X \\ (V, V', j') &\mapsto X \quad \text{if } V, V' \in \mathcal{W}, j' \in \{1, 2\} \end{aligned} \quad (29)$$

As well, we extend the homomorphisms $\Phi_{i,a,d,W,j}$ to $\hat{\mathcal{W}}^*$ by setting:

$$\begin{aligned} X &\mapsto X \\ (V, V', j') &\mapsto X \quad \text{if } V, V' \in \mathcal{W}, j' \in \{1, 2\} \end{aligned} \quad (30)$$

Lemma 6. *For every $i \in I_0, d \in 1 - \text{pds}(\Gamma) \equiv 3, a \in \Gamma, w \in 1 - \text{pds}(\Gamma)$ and every letter $W \in \text{Alph}(i, ad)$, there exists an integer $\ell(i, a, d, W) \in [1, 2]$ and for every $j \in [1, \ell(i, a, d, W)]$ there exist indices $\alpha(i, a, d, W, j), \beta(i, a, d, W, j) \in I_0$ and words $\omega_{i,a,d,W,j} \in \Gamma_3^{\leq 2}$ such that*

$w \in d \Rightarrow$

$$\begin{aligned} H_i^{aw} &= \psi_{i,a,d} \circ \prod_{W \in \text{Alph}(i, ad)} \left(\prod_{j=1}^{\ell(i,a,d,W)} \theta_{i,a,d,W,j} \circ H_{\alpha(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} \circ \Phi_{i,a,d,W,j} \circ H_{\beta(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} \circ \hat{\theta}_{i,a,d,W,j} \right) \\ &\quad \circ \hat{\psi}_{i,a,d}. \end{aligned} \quad (31)$$

In the above lemma the symbols \prod stand for the extension of the composition law to an arbitrary finite number of arguments. This law is non-commutative, so that the ordering of the five homomorphisms on the righthand-side is significant; however, for the particular operands of the symbols \prod , the result does *not* depend on their ordering).

Proof. For a letter $Y \in \hat{\mathcal{W}} \setminus \mathcal{W}$, both sides of equation (31) map Y on X . For a letter $Y \in \mathcal{W}$, one can check that both sides of equation (31) map Y on the same word, by using equations (18) and the definitions of ψ_* , $\hat{\psi}_*$, θ_* , $\hat{\theta}_*$.

We define a larger family of mappings by setting:

$$I := I_0 \cup \{(i, a, d, s) \mid s \in \{-1, +1\}\} \cup \{((i, a, d, W, j, s) \mid s \in \{-1, 0, +1\})\}.$$

The homomorphism H_i^w is already defined for $w \in \Gamma^*$, $i \in I_0$. We now define:

$$H_{i,a,d,1}^w = \psi_{i,a,d}, \quad H_{i,a,d,-1}^w = \hat{\psi}_{i,a,d}, \quad (32)$$

$$H_{i,a,d,W,j,1}^w = \theta_{i,a,d,W,j}, \quad H_{i,a,d,W,j,0}^w = \Phi_{i,a,d,W,j}, \quad H_{i,a,d,W,j,-1}^w = \hat{\theta}_{i,a,d,W,j}. \quad (33)$$

and we extend the map $\text{Alph}(*, *)$ which is already defined for arguments $(i, w) \in I_0 \times \Gamma_3^*$ by setting

$$\text{Alph}(i, w) := \emptyset \text{ for } i \in I \setminus I_0, w \in \Gamma_3^*.$$

With these new notations, relations (31) can be rewritten as: for every $w \in d$ & $i \in I_0$

$$\begin{aligned} H_i^{aw} &= H_{i,a,d,1}^w \\ &\circ \prod_{W \in \text{Alph}(i, ad)} \left(\prod_{j=1}^{\ell(i,a,d,W)} H_{i,a,d,W,j,1}^w \circ H_{\alpha(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} \circ H_{i,a,d,W,j,0}^w \circ H_{\beta(i,a,d,W,j)}^{\omega_{i,a,d,W,j} \cdot w} \circ H_{i,a,d,W,j,-1}^w \right) \\ &\circ H_{i,a,d,-1}^w. \end{aligned} \quad (34)$$

while it is obviously true that: for every $w \in d$ & $i \in I \setminus I_0$

$$H_i^{aw} = H_i^w. \quad (35)$$

Thus the family $(H_i)_{i \in I}$ fulfills a system of regular recurrent relations in the monoid $\text{HOM}(\mathcal{W}^*, \mathcal{W}^*)$. The aim of next paragraph is to prove that this system is *noetherian*.

3.5 Termination

Let us denote by \mathcal{C}_1 the rewriting system consisting of all rules of the form

$$H_i^{aw} \rightarrow \prod_{j=1}^J H_{i_j}^{w_j} \quad (36)$$

which belong to the set of equations (34) and $\bigcup_{j=1}^J \text{Alph}(i_j, w_j) \neq \emptyset$.

Let us note by \mathcal{C}_2 the rewriting system consisting of all rules of the form (36)

which belong to the set of equations (34) and $\bigcup_{j=1}^J \text{Alph}(i_j, w_j) = \emptyset$.

Similarly \mathcal{C}_3 corresponds to the set of equations (35) oriented from left to right and finally

$$\mathcal{C} := \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3.$$

This subsection proves termination of the system \mathcal{C} . Our proof-strategy consists in relating derivations (modulo $\rightarrow_{\mathcal{C}}$) with derivations (modulo $\rightarrow_{\mathcal{A}}$). We shall essentially show that, given a starting term H_i^w , and some starting derivation D of the form $\text{WD} \rightarrow_{\mathcal{A}}^* H_i^w(\text{WD})$, every derivation $H_i^w \rightarrow_{\mathcal{C}}^n \prod_{\lambda=1}^{\Lambda} H_{i_{\lambda}}^{w_{\lambda}}$ leads to a non-trivial *refinement* of size $J \geq n$ for the fixed derivation D . This shows that n cannot exceed the length of D . We introduce below a notion of *derivation-path* and notions of *types* for derivations and derivation-paths, which will allow a smooth inductive proof of the above refinement property (see Lemma 12).

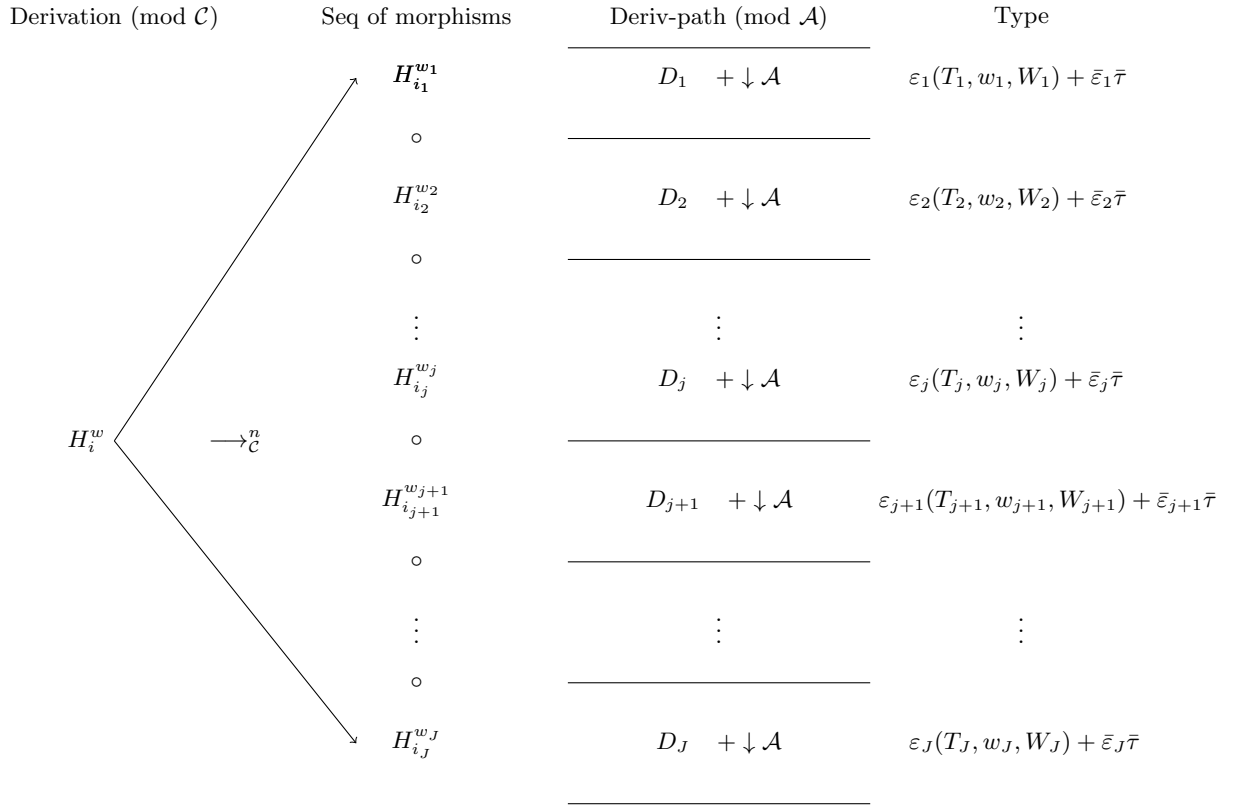


Fig. 1. Derivations (modulo \mathcal{C}) versus derivation-paths (modulo \mathcal{A}).

Figure 3.5 depicts the relationship between some \mathcal{C} -derivation $\rightarrow_{\mathcal{C}}^n$ starting from H_i^w and some refinement of length $J \geq n$ of the derivation $\text{WD} \rightarrow_{\mathcal{A}}^* H_i^w(\text{WD})$ (for some fixed well-chosen variable-term-word of order 3 WD); the numbers $\varepsilon_j, \bar{\varepsilon}_j$ take their value in $\{0, 1\}$ and $\sum_{j=1}^J \bar{\varepsilon}_j \geq n_1$ (where n_1 is the number of rules from \mathcal{C}_1 used in the \mathcal{C} -derivation).

Types and undeterminates We define a finite set of types \mathcal{T} by

$$\mathcal{T} := \{(T, w, W) \mid T \in \Gamma_2, w \in \Gamma_3^*, W \in \mathcal{W}\} \cup \{\varepsilon, \bar{\tau}\}$$

where $\bar{\tau}$ is just a new symbol.

We manipulate here terms (and variable-terms) of various levels in $[1, 3]$, with pushdown symbols in Γ and undeterminates in E . Since $E \subseteq \hat{\Gamma}^* / \equiv_2$ the self-extension $\hat{\equiv}_2$ is defined over $\hat{\Gamma}^* * E^*$, which includes all the sets $\ell - \text{term}(\Gamma, E)$ for $1 \leq \ell \leq 3$. In the following we keep the same notation \equiv_2 for what should be denoted by $\hat{\equiv}_2$ (see Definition (1) in §2.3).

Definition 9. Let D be a derivation (modulo $\rightarrow_{\mathcal{A}}$) over $(3 - \text{vterm}(\Gamma, E))^*$. Its type $\tau(D)$ is the following element of \mathcal{T} :

$\tau(D) := (T, w, W)$ if D has the form $L \cdot W[T[w]u/e] \cdot R \rightarrow_{\mathcal{A}}^* L \cdot H_i^w(W)[u/e] \cdot R$ where $W = pS[e]q, W \in \text{Alph}(i, w), u \in 2 - \text{term}(\Gamma, E), u \in e$ and $L, R \in (3 - \text{vterm}(\Gamma, E))^*$

$\tau(D) := \bar{\tau}$ if D has not the above form and $\ell(D) \geq 1$

$\tau(D) := \varepsilon$ if $\ell(D) = 0$.

Lemma 7. The binary relation τ introduced by Definition 9 is a map: every derivation has exactly one type.

Proof. Suppose that $A, B \in (3 - \text{vterm}(\Gamma, E))^*$ are the extreme points of a derivation $D : A \rightarrow_{\mathcal{A}}^* B$ that admits both types (T, w, W) and (T', w', W') . Let us show that these types are equal. This means

$$A = L \cdot W[T[w]u/e] \cdot R \rightarrow_{\mathcal{A}}^* L \cdot H_i^w(W)[u/e] \cdot R = B$$

and

$$A = L' \cdot W'[T'[w']u'/e'] \cdot R' \rightarrow_{\mathcal{A}}^* L' \cdot H_{i'}^{w'}(W')[u'/e'] \cdot R = B.$$

Let us assume that $W = pS[e]q, W' = p'S'[e']q'$

Case 1: $|L| = |L'|$.

The two lefthandsides of rules of $\rightarrow_{\mathcal{A}}$ occur at the same position of A , hence $W[T[w]u/e] = W'[T'[w']u'/e']$. It follows that $pS[T[w]u]q = p'S'[T'[w']u']q'$ hence that $p = p', q = q', S = S', T = T', w = w'$. Since $e = [T[w]u]_{\equiv_2}$ and $e' = [T'[w']u']_{\equiv_2}$, we also have $e = e'$, implying that $W = W'$.

Case 2: $|L| < |L'|$.

By equidivisibility of the monoid $(3 - \text{vterm}(\Gamma, E))^*$, there exists some $M \in (3 - \text{vterm}(\Gamma, E))^*$, such that:

$$A = L \cdot W[T[w]u/e] \cdot M \cdot W'[T'[w']u'/e'] \cdot R'$$

and

$$L' = L \cdot W[T[w]u/e] \cdot M, \quad R = M \cdot W'[T'[w']u'/e'] \cdot R'.$$

Let us consider the derivation (10) used for the definition of $H_i^w(pS[e]q)$:

$$(pS[T[w]\Omega]q) \rightarrow_{\mathcal{A}}^+ \prod_{j=1}^{\ell(i,W)} (p_{i,j} S_{i,j}[\Omega] q_{i,j}).$$

The word B has thus a prefix of the form:

$$P := L \cdot p_{i,1} S_{i,1}[u] q_{i,1}.$$

Since B is also obtained from A by rewriting the suffix $M \cdot W'[T'[w']u'/e'] \cdot R'$, and this derivation leaves invariant the prefix of length $|L| + 1$ of A , B has a prefix of the form

$$P' = L \cdot pS[T[w]u]q.$$

Note that the atomic pushdowns of order 3 $S_{i,1}[u], S[T[w]u]$ are different (because their level 2 are different). The two prefixes P, P' have same length and have a different last letter : this is impossible.

By symmetry $|L| > |L'|$ is also impossible.

Clearly the two other possible types ($\bar{\tau}$ and ε) are incompatible with a type (T, w, W) , and also mutually incompatible. Hence, every derivation has at most one type.

It is clear from the definition that it has at least one type.

Derivations and derivation-paths The *partial product* of two derivations is defined as usual and denoted by \otimes . We denote by a dot-symbol \cdot the left-operation of $(3\text{-vterm}(I, E))^*$ on derivations: it consists in concatenating the left operand to the left, at every step of the derivation. The right-operation is also denoted by \cdot . We call *derivation path* every sequence $(D_1, \dots, D_n, \dots, D_\ell)$ of derivations such that the product $D_1 \otimes \dots \otimes D_n \otimes \dots \otimes D_\ell$ is defined (equivalently, such that the product $D_i \otimes D_{i+1}$ is defined for every $i \in [1, \ell - 1]$). The *partial product* of two derivation-paths, denoted by \odot , is just the concatenation of the two derivation-paths, when this concatenation is itself a derivation-path, and it is otherwise undefined.

We define the ordering \preceq over derivation-paths as the least ordering, which is compatible with the (partial) product \odot and which fulfils that, for every pair of derivations D, D' , if the product $D \otimes D'$ is defined, then

$$(D \otimes D') \preceq (D, D').$$

The reader should notice the analogy between words and words of level 2 (on one hand), derivations and derivation-paths (on the other hand). A derivation-path can be interpreted as some *decomposition* of a derivation. An inequality $\vec{D} \preceq \vec{D}'$ means that both derivation-paths \vec{D}, \vec{D}' are decompositions of the *same* derivation (modulo $\rightarrow_{\mathcal{A}}$) and the later *refines* the former.

Typing the derivation-paths We extend the map τ into a map from the set of derivation-paths to the free commutative monoid generated by \mathcal{T} , $\mathbb{N}\langle\mathcal{T}\rangle$, in such a way that τ is an homomorphism (of partial monoids). Namely: for every derivation-path $(D_1, \dots, D_n, \dots, D_\ell)$:

$$\tau(D_1, \dots, D_n, \dots, D_\ell) := \sum_{n=1}^{\ell} \tau(D_n).$$

Let us use the notation:

$$t(i, w) := \sum_{W \in \text{Alph}(i, w)} (T, w, W)$$

for every $i \in I_0$ of the form (\mathcal{V}, T) and every $w \in \Gamma^*$,

$$t(i, w) = 0$$

for every $i \in I \setminus I_0$ and every $w \in \Gamma^*$. We also note $\text{al}(i, w) := \text{Card}(\text{Alph}(i, w))$. We denote by \leq the product ordering over $\mathbb{N}\langle\mathcal{T}\rangle$:

$$(\sum_{t \in \mathcal{T}} n_t \cdot t \leq \sum_{t \in \mathcal{T}} m_t \cdot t) \Leftrightarrow (\forall t \in \mathcal{T}, n_t \leq m_t).$$

Let us recall that, by Lemma 1, if D is a derivation (modulo $\rightarrow_{\mathcal{A}}^*$), $e \in E$ and $u \in 2 - \text{pds}(\Gamma, E)$, then $D[u/e]$ is also a derivation (modulo $\rightarrow_{\mathcal{A}}^*$).

Lemma 8. *Let us consider a derivation D (modulo $\rightarrow_{\mathcal{A}}^*$), some variable-term-words of order 3 L, R (over (Γ, E)), an undeterminate $e \in E$, a word $u \in \Gamma_3^*$ and two derivation-paths \vec{D}, \vec{D}' . Then*

- 1- $\tau(D[u/e]) = \tau(D)$
- 2- $\tau(L \cdot D \cdot R) = \tau(D)$
- 3- $\tau(\vec{D} \odot \vec{D}') = \tau(\vec{D}) + \tau(\vec{D}')$.

This lemma is an immediate consequence of the definitions of τ .

Lemma 9. *Let $H_i^{aw}(W) = \prod_{j=1}^{\ell(i, a, d, W)} H_{\alpha(i, a, d, W, j)}^{\omega_{i, a, d, W, j} \cdot w} \circ \Phi_{i, a, d, W, j} \circ H_{\beta(i, a, d, W, j)}^{\omega_{i, a, d, W, j} \cdot w}(V_{i, a, d, W, j})$ be a “mixed rule” with $i = (\mathcal{V}, T)$, $W \in \text{Alph}(i, w)$. Let D be a derivation (modulo $\rightarrow_{\mathcal{A}}$) of type (T, aw, W) .*

If the rule realizes Case 1.1, then there exists a derivation-path $\vec{D}' \succeq (D)$ such that:

$$\tau(\vec{D}') \geq \sum_{j=1}^{\ell(i, a, d, W)} t(\alpha(i, a, d, W, j), \omega_{i, a, d, W, j} \cdot w) + t(\beta(i, a, d, W, j), \omega_{i, a, d, W, j} \cdot w).$$

If the rule realizes Case 1.2 or 1.3 or 2, then there exists a derivation-path $\vec{D}' \succeq (D)$ such that:

$$\tau(\vec{D}') \geq (\sum_{j=1}^{\ell(i, a, d, W)} t(\alpha(i, a, d, W, j), \omega_{i, a, d, W, j} \cdot w) + t(\beta(i, a, d, W, j), \omega_{i, a, d, W, j} \cdot w)) + \bar{\tau}.$$

Proof. Let us consider a rule and a derivation D fulfilling the hypothesis of the lemma. By Lemma 5 we know that one of cases 1.1, 1.2, 1.3, 2 must occur.

Case 1.1:

Then $\sum_{j=1}^{\ell(i,a,d,W)} t(\alpha(i,a,d,W,j), \omega_{i,a,d,W,j} \cdot w) + t(\beta(i,a,d,W,j), \omega_{i,a,d,W,j} \cdot w) = 0$. Taking $\vec{D}' := (D)$, the conclusion of the lemma is true.

Case 1.2:

Let $D_0 : W[T[aw]e/e] \rightarrow_{\mathcal{A}} V_{i,a,d,W,1}[T'[\omega_{i,a,d,W,j}w]e/e] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W)[e/e]$. Let us decompose D as $D_1 \otimes D_2$ with:

$$D_1 : W[T[aw]e/e] \rightarrow_{\mathcal{A}} V_{i,a,d,W,1}[T'[\omega_{i,a,d,W,j}w]e/e]$$

$$D_2 : V_{i,a,d,W,1}[T'[\omega_{i,a,d,W,j}w]e/e] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W)[e/e] = H_{\alpha(i,a,d,W,1)}^{\omega_{i,a,d,W,1} \cdot w}(V_{i,a,d,W,1})[e/e].$$

Since the extremity of D_1 is not of the form $H_i^{aw}(W)[e/e]$, $\tau(D_1) = \bar{\tau}$. We also see that: $\tau(D_2) = (T', \omega_{i,a,d,W,j}w, V_{i,a,d,W,1})$. Let $\vec{D}' := (D_1, D_2)$. This derivation-path fulfils:

$$\vec{D}' \succeq (D)$$

and

$$\begin{aligned} \tau(\vec{D}') &= \tau(D_1) + \tau(D_2) \\ &= \bar{\tau} + (T', \omega_{i,a,d,W,j}w, V_{i,a,d,W,1}) \\ &= (T', \omega_{i,a,d,W,j}w, V_{i,a,d,W,1}) + \bar{\tau}. \end{aligned}$$

But the value of $t(\beta(i,a,d,W,1), \omega_{i,a,d,W,1} \cdot w)$ is 0, hence \vec{D}' fulfils the conclusion of the lemma.

A general derivation D of type (T, aw, W) , fulfilling this Case, must be of the form $D = L \cdot D_0[u/e] \cdot R$ for some $L, R \in (3 - \text{vterm}(\Gamma, E))^*$, $u \in 2 - \text{term}(\Gamma, E)$ and $u \equiv_2 e$. Since the substitution $[u/e]$ and the product by L , on the left, and by R , on the right, preserve the operation \otimes and preserve the type, we obtain the required conclusion for D .

Case 1.3:

Let $D_0 : W[T[aw]e/e] \rightarrow_{\mathcal{A}} V[T'[w']T''[w']e/e'] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W)$,

where, for ease of notation, we have used the abbreviation: $w' := \omega_{i,a,d,W,j}w$, $V := V_{i,a,d,W,1}$.

Let us also abbreviate $\alpha := \alpha(i,a,d,W,1)$; $\Phi := \Phi_{i,a,d,W,1}$ $\beta := \beta(i,a,d,W,1)$.

We define two derivations C_1, C_2 by:

$$C_1 : W[T[aw]e/e] \rightarrow_{\mathcal{A}} V[T'[w']T''[w']e/e']$$

$$C_2 : V[T'[w']T''[w']e/e'] \rightarrow_{\mathcal{A}}^* H_{\alpha}^{w'}(V)[T''[w']e/e']$$

The word $H_{\alpha}^{w'}(V)$ decomposes into letters as:

$$H_{\alpha}^{w'}(V) := W_1 \cdots W_{\lambda} \cdots W_{\Lambda}$$

where each W_{λ} has the form $W_{\lambda} = p_{\lambda}S_{\lambda}[e']q_{\lambda}$. Let us define

$$W'_{\lambda} := W_{\lambda}[e/e'].$$

We then have the derivations D_λ defined by:

$$D_\lambda : W_\lambda[T''[w']e/e'] = W'_\lambda[T''[w']e/e] \rightarrow_{\mathcal{A}}^* H_{\beta}^{w'}(W'_\lambda) \quad (37)$$

By definition of Case 1.3 in Lemma 5, $\text{Alph}(\beta) = \{\Phi(W_\lambda) \mid 1 \leq \lambda \leq A\}$ hence $\text{Alph}(\beta) = \{W'_\lambda \mid 1 \leq \lambda \leq A\}$. Let us denote by D_λ the derivation (37). We then have

$$D_0 = C_1 \otimes C_2 \otimes \bigotimes_{\lambda=1}^A D_\lambda \quad (38)$$

and

$$\tau(C_1) = \bar{\tau}, \quad \tau(C_2) = (T', w', V), \quad \tau(D_\lambda) = (T'', w', W'_\lambda) \quad (39)$$

Let us define

$$\vec{D}' := (C_1, C_2, D_1, \dots, D_\lambda, \dots, D_A)$$

By (38), $\vec{D}' \succeq (D_0)$. Since $\text{Alph}(\alpha, w') = \{V\}$ and $\text{Alph}(\beta, w') = \{W'_\lambda \mid 1 \leq \lambda \leq A\}$, we can write:

$$\begin{aligned} \tau(\vec{D}') &= \tau(C_1) + \tau(C_2) + \sum_{\lambda=1}^A \tau(D_\lambda) \\ &= \bar{\tau} + t(\alpha, w') + t(\beta, w') \\ &\geq t(\alpha, w') + t(\beta, w') + \bar{\tau} \end{aligned} \quad (40)$$

as required. We reduce to D_0 the treatment of a general derivation D of type (T, aw, W) , fulfilling Case 1.3, as we did for Case 1.2.

Case 2:

Let $D_0 : W[T[aw]e/e] \rightarrow_{\mathcal{A}} V_1[T[w']e/e]V_2[T[w']e/e] \rightarrow_{\mathcal{A}}^* H_i^{aw}(W)$ where, for ease of notation, we have used the abbreviation: $w' := \omega_{i,a,d,W,j} \cdot w$, $V_1 := V_{i,a,d,W,1}$, $V_2 := V_{i,a,d,W,2}$.

Let us also abbreviate $\alpha_1 := \alpha(i, a, d, W, 1)$; $\alpha_2 := \alpha(i, a, d, W, 2)$, $\beta_1 := \beta(i, a, d, W, 1)$, $\beta_2 := \beta(i, a, d, W, 2)$. We define three derivations C_1, D_1, D_2 by:

$$C_1 : W[T[aw]e/e] \rightarrow_{\mathcal{A}}^+ V_1[T[w']e/e]V_2[T[w']e/e]$$

$$D_1 : V_1[T[w']e/e] \rightarrow_{\mathcal{A}}^* H_{\alpha_1}^{w'}(V_1).$$

$$D_2 : V_2[T[w']e/e] \rightarrow_{\mathcal{A}}^* H_{\alpha_2}^{w'}(V_2).$$

We then have

$$D_0 = C_1 \otimes D_1 \otimes D_2 \quad (41)$$

and

$$\tau(C_1) = \bar{\tau}, \quad \tau(D_1) = (T, w', V_1), \quad \tau(D_2) = (T, w', V_2). \quad (42)$$

Let us define

$$\vec{D}' := (C_1, D_1, D_2)$$

Since $\text{Alph}(\alpha_1, w') = \{V_1\}, \text{Alph}(\alpha_2, w') = \{V_2\}$ and $\text{Alph}(\beta_1, w') = \text{Alph}(\beta_2, w') = \emptyset$, we can write:

$$\begin{aligned}\tau(\vec{D}') &= \tau(C_1) + \tau(D_1) + \tau(D_2) \\ &= \bar{\tau} + t(\alpha_1, w') + t(\alpha_2, w') \\ &= \left(\sum_{j=1}^2 (t(\alpha_j, w') + t(\beta_j, w')) \right) + \bar{\tau}\end{aligned}\tag{43}$$

as required. We reduce to D_0 the treatment of a general derivation D of type (T, aw, W) , fulfilling Case 2, as we did for Case 1.2.

We treat in next lemma the case of a rule of \mathcal{C}_1 .

Lemma 10. *Let us consider the compositional rule $H_i^{aw} \rightarrow \prod_{j=1}^J H_{i_j}^{w_j}$, given in (34), corresponding to a class d and an index $i \in I_0$ of the form $i = (\mathcal{V}, T)$ and suppose that $\bigcup_{j=1}^J \text{Alph}(i_j, w_j) \neq \emptyset$. Let \vec{D} be a derivation-path (modulo $\rightarrow_{\mathcal{A}}$) and let $\tau_0 \in \mathbb{N}\langle T \rangle$ such that*

$$\tau(\vec{D}) = t(i, aw) + \tau_0.$$

Then, there exists a derivation-path $\vec{D}' \succeq \vec{D}$ such that

$$\tau(\vec{D}') \geq \sum_{j=1}^J t(i_j, w_j) + \bar{\tau} + \tau_0.$$

Proof. Let \vec{D} fulfil the hypothesis of the lemma. It has the form

$$\vec{D} = \vec{E}_0 \odot \left(\bigodot_{\lambda=1}^{\text{al}(i, w)} (D_\lambda) \odot \vec{E}_\lambda \right)$$

with $\text{Alph}(i, w) = \{W_\lambda \mid 1 \leq \lambda \leq \text{al}(i, w)\}$,

$$\sum_{\lambda=0}^{\text{al}(i, w)} \tau(E_\lambda) = \tau_0; \quad \tau(D_\lambda) = (T, aw, W_\lambda) \text{ for } 1 \leq \lambda \leq \text{al}(i, w).$$

By Lemma 9, for every $\lambda \in [1, \text{al}(i, w)]$, there exists a derivation-path $\vec{D}'_\lambda \succeq (D_\lambda)$ such that:

$$\tau(\vec{D}'_\lambda) \geq \sum_{j=1}^{\ell(i, a, d, W_\lambda)} t(\alpha(i, a, d, W_\lambda, j), \omega_{i, a, d, W_\lambda, j} \cdot w) + t(\beta(i, a, d, W_\lambda, j), \omega_{i, a, d, W_\lambda, j} \cdot w).\tag{44}$$

Since the righthanside of the given rule

$$H_i^{aw} \rightarrow \prod_{j=1}^J H_{i_j}^{w_j}$$

contains, as disjoint factors, all the products $H_{\alpha(i,a,d,W_\lambda,j)}^{\omega_{i,a,d,W_\lambda,j} \cdot w} \circ H_{i,a,d,W_\lambda,j,0}^w \circ H_{\beta(i,a,d,W_\lambda,j)}^{\omega_{i,a,d,W_\lambda,j}}$, and no other factor with index i' and exponent w' such that $t(i', w') \neq 0$, we get that

$$\sum_{j=1}^J t(i_j, w_j) = \sum_{\lambda=1}^{\text{al}(i,w)} \left(\sum_{j=1}^{\ell(i,a,d,W_\lambda)} t(\alpha(i, a, d, W_\lambda, j), \omega_{i,a,d,W_\lambda,j} \cdot w) + t(\beta(i, a, d, W_\lambda, j), \omega_{i,a,d,W_\lambda,j} \cdot w) \right). \quad (45)$$

Let us define

$$\vec{D}' := \vec{E}_0 \odot \left(\bigodot_{\lambda=1}^{\text{al}(i,w)} \vec{D}'_\lambda \odot \vec{E}_\lambda \right).$$

This derivation-path fulfils $\vec{D}' \succeq \vec{D}$ and, by (44,45)

$$\tau(\vec{D}') \geq \left(\sum_{j=1}^J t(i_j, w_j) \right) + \tau_0. \quad (46)$$

Assuming that $\bigcup_{j=1}^J \text{Alph}(i_j, w_j) \neq \emptyset$, at least one index j cooresponds to some case other than case 1.1 in Lemma 10. Hence, at least one of the minorations (44) can be improved by adding $\bar{\tau}$ to its righhand-side. Hence

$$\tau(\vec{D}') \geq \left(\sum_{j=1}^J t(i_j, w_j) \right) + \tau_0 + \bar{\tau}.$$

Let us treat now the case of a rule of $\mathcal{C}_2 \cup \mathcal{C}_3$.

Lemma 11. *Let us consider the compositional rule $H_i^{aw} \rightarrow \prod_{j=1}^J H_{i_j}^{w_j}$, either given in (34), by a class d and an index $i \in I_0$ of the form $i = (\mathcal{V}, T)$ such that $\bigcup_{j=1}^J \text{Alph}(i_j, w_j) = \emptyset$ or given in (35). Let \vec{D} be a derivation-path (modulo $\rightarrow_{\mathcal{A}}$) and some $\tau_0 \in \mathbb{N}\langle T \rangle$ such that*

$$\tau(\vec{D}) = t(i, aw) + \tau_0.$$

Then, there exists a derivation-path $\vec{D}' \succeq \vec{D}$ such that

$$\tau(\vec{D}') \geq \sum_{j=1}^J t(i_j, w_j) + \tau_0.$$

Proof. Let us choose $\vec{D}' := \vec{D}$. Since $\bigcup_{j=1}^J \text{Alph}(i_j, w_j) = \emptyset$, in fact $\sum_{j=1}^J t(i_j, w_j) = 0$ and the lemma follows.

Lemma 12. Suppose that $H_i^w \rightarrow_{\mathcal{C}}^* \prod_{\lambda=1}^{\Lambda} H_{i_\lambda}^{w_\lambda}$ and that this derivation (modulo $\rightarrow_{\mathcal{C}}$) has n steps in $\rightarrow_{\mathcal{C}_1}$. Let \vec{D} be a derivation-path (modulo $\rightarrow_{\mathcal{A}}$) such that

$$\tau(\vec{D}) = t(i, w).$$

Then, there exists a derivation-path $\vec{D}' \succeq \vec{D}$ such that

$$\tau(\vec{D}') \geq \sum_{\lambda=1}^{\Lambda} t(i_\lambda, w_\lambda) + n \cdot \bar{\tau}.$$

Proof. We prove this lemma by induction on the length m of the given derivation (modulo $\rightarrow_{\mathcal{C}}$). If we suppose that $m = 0$, then $\Lambda = 1, i_\lambda = i, w_\lambda = w, n = 0$, so that the conclusion of the lemma does hold.

Let us consider a derivation of length $m + 1$ with n steps in $\rightarrow_{\mathcal{C}_1}$. It can be decomposed as

$$H_i^w \rightarrow_{\mathcal{C}}^* \prod_{\lambda=1}^{\Lambda} H_{i_\lambda}^{w_\lambda} \rightarrow_{\mathcal{C}} \prod_{\lambda=1}^{\ell-1} H_{i_\lambda}^{w_\lambda} \circ \prod_{j=1}^J H_{i_j}^{v_j} \circ \prod_{\lambda=\ell+1}^{\Lambda} H_{i_\lambda}^{w_\lambda}$$

where

$$H_{i_\ell}^{w_\ell} \rightarrow_{\mathcal{C}} \prod_{j=1}^J H_{i_j}^{v_j}. \quad (47)$$

Let \vec{D} fulfil the hypothesis of the lemma.

Case 1: The last step (47) uses $\rightarrow_{\mathcal{C}_1}$.

By induction hypothesis, there exists a derivation-path $\vec{D}'' \succeq \vec{D}$ such that

$$\tau(\vec{D}'') \geq \sum_{\lambda=1}^{\Lambda} t(i_\lambda, w_\lambda) + (n-1) \cdot \bar{\tau}.$$

Using Lemma 10 with $\tau_0 = \tau(\vec{D}'') - t(i_\ell, w_\ell)$, we obtain a derivation-path $\vec{D}' \succeq \vec{D}''$ such that

$$\begin{aligned} \tau(\vec{D}') &\geq \sum_{j=1}^J t(i_j, v_j) + \tau_0 + \bar{\tau} \\ &= \sum_{\lambda=1}^{\ell-1} t(i_\lambda, w_\lambda) + \sum_{j=1}^J t(i_j, v_j) + \sum_{\lambda=\ell+1}^{\Lambda} t(i_\lambda, w_\lambda) + n \cdot \bar{\tau} \end{aligned}$$

as required.

Case 2: The last step (47) uses $\rightarrow_{\mathcal{C}_2 \cup \mathcal{C}_3}$.

By induction hypothesis, there exists a derivation-path $\vec{D}'' \succeq \vec{D}$ such that

$$\tau(\vec{D}'') \geq \sum_{\lambda=1}^{\Lambda} t(i_\lambda, w_\lambda) + n \cdot \bar{\tau}.$$

Using Lemma 11 we can conclude as in Case 1.

For every $i \in I_0$ and $w \in I_3^*$, we choose an enumeration, without repetition, of the set $\text{Alph}(i, w)$:

$$\text{Alph}(i, w) = \{W_1, W_2, \dots, W_{\text{al}(i, w)}\},$$

and define the word

$$\text{WD}(i, w) = W_1 W_2 \dots W_{\text{al}(i, w)}$$

Lemma 13. *The relation $\rightarrow_{\mathcal{C}_2 \cup \mathcal{C}_3}$ is noetherian.*

This is straightforward.

Lemma 14. *The system of all relations (34-35) is a noetherian system of regular recurrent relations of order 1 in $\text{HOM}(\hat{\mathcal{W}}^*, \hat{\mathcal{W}}^*)$.*

Proof. Let us suppose that

$$\rightarrow_{\mathcal{C}} \text{ admits an infinite derivation} \quad (48)$$

The system \mathcal{C} is monadic, hence there exists such an infinite derivation starting on a letter H_i^w (for some $i \in I_0, w \in I_3^*$):

$$H_i^w \rightarrow_{\mathcal{C}}^\infty \quad (49)$$

By Lemma 13, this derivation (49) must use infinitely many steps in $\rightarrow_{\mathcal{C}_1}$. Let us consider a derivation

$$D_i : \text{WD}(i, w)[T[w]e/e; e \in 2 - \text{pds}(\Gamma)/\equiv_2] \rightarrow_{\mathcal{A}}^* H_i^w(\text{WD}(i, w))$$

Since $\tau(D_i) = t(i, w)$, by Lemma 12, for every $n \geq 0$,

$$\text{WD}(i, w)[T[w]e/e; e \in 2 - \text{pds}(\Gamma)/\equiv_2] \rightarrow_{\mathcal{A}}^{\geq n} H_i^w(\text{WD}(i, w)).$$

Since each step of a derivation modulo $\rightarrow_{\mathcal{A}}$ applies on only one letter, this would imply that there exists a single letter W_j such that, for every $n \geq 0$,

$$W_j[T[w]e_j/e_j] \rightarrow_{\mathcal{A}}^{\geq n} H_i^w(W_j).$$

Let $u_j \in 3 - \text{pds}(\Gamma)$ such that $u_j \in e_j$ and let us substitute u_j to e_j in the above derivation. For every $n \geq 0$,

$$W_j[T[w]u_j/e_j] \rightarrow_{\mathcal{A}}^{\geq n} H_i^w(W_j)[u_j/e_j]. \quad (50)$$

By definition of $H_i^w(W_j)$, since $W_j \in \text{Alph}(i, w)$,

$$\text{L}(\mathcal{A}, H_i^w(W_j)[u_j/e_j]) \neq \emptyset.$$

Hence there exists a fixed word $v_j \in B^*$ and a fixed integer $m \geq 0$ such that

$$H_i^w(W_j)[u_j/e_j] \rightarrow_{\mathcal{A}}^m v_j. \quad (51)$$

Finally, combining derivation (50) with derivation (51):

$$\forall n \geq m, W_j[T[w]u_j/e_j] \rightarrow_{\mathcal{A}}^{\geq n} v_j \in B^*. \quad (52)$$

Since \mathcal{A} is strongly deterministic, there is only one maximal \mathcal{A} -computation starting from $p_j S_j[T[w]u_j]$ and this computation is finite (it ends in the configuration with state q_j and empty pds). This entails that there are only finitely many derivations modulo $\rightarrow_{\mathcal{A}}$ starting from $p_j S_j[T[w]u_j]q_j$. This contradicts assertion (52). We have thus proved that hypothesis (48) is impossible.

Let us complete the proof of (1) \Rightarrow (2) in Theorem 3 by a suitable choice of alphabets A, C , initial letter $c \in C$, main homomorphism sequence $w \mapsto H_{i_1}^w : A^* \rightarrow \text{HOM}(C^*, C^*)$ and final homomorphism $h \in \text{HOM}(C^*, B^*)$ (while the alphabet B is the terminal alphabet of \mathcal{A}). We choose:

$$A := \Gamma_3, \quad C := \mathcal{W}$$

$$e_0 := [\varepsilon]_{\equiv_2}, \quad c := q_0 \gamma_1 [e_0] q_0, \quad i_1 := (\mathcal{W}, \gamma_2)$$

and, for all $p, q \in Q, S \in \Gamma_1, e \in 2 - \text{pds}(\Gamma)/\equiv_2$

$$h(pS[e]q) := s \text{ if } (e = [\varepsilon]_{\equiv_2} \text{ and } pS[\varepsilon]q \rightarrow_{\mathcal{A}}^* s), \quad (53)$$

$$h(pS[e]q) := \varepsilon \text{ otherwise} \quad (54)$$

We claim that, for every $w \in A^*$

$$f(w) = h(H_{i_1}^w(c)).$$

(Note that the right-hand side of (54) can be chosen arbitrarily, without affecting this crucial claim).

Let us prove that (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5):

It suffices to use the implications (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) from Theorem 1.

Let us sketch a proof that (5) \Rightarrow (1):

Proposition 70 of [FS06] asserts that, if $f \in \mathbb{S}_k(A^*, B^*), g \in \mathbb{S}_\ell(B^*, C^*)$ then $f \circ g \in \mathbb{S}_{k+\ell-1}(A^*, C^*)$, in the particular case where $|A| = |B| = |C| = 1$ (i.e., for integer sequences). By a suitable adaptation of the proof of this proposition, one can show that the same property holds for arbitrary finite alphabets A, B, C . The case where $k = \ell = 2$ gives the required implication (5) \Rightarrow (1).

4 Applications

= Polynomial automata

Definition 10 (Polynomial recurrent relations). *Given a finite index set $I = [1, n]$ and a family of mappings $f_i : A^* \rightarrow \mathbb{N}$ (for $i \in I$), we call system of polynomial recurrent relations a system of the form*

$$f_i(aw) = P_{i,a}(f_1(w), f_2(w), \dots, f_n(w)) \text{ for all } i \in I, a \in A, w \in A^*$$

where $P_{i,a} \in \mathbb{N}[X_1, X_2, \dots, X_n]$.

A similar definition can be given for mappings $f_i : A^* \rightarrow \mathbb{Z}$ (for $i \in I$) and polynomials $P_i \in \mathbb{Z}[X_1, X_2, \dots, X_n]$.

Theorem 2 specializes as follows in the particular case where B is reduced to one letter i.e. when the mapping f is a formal power series.

Corollary 1. *Let us consider a mapping $f : A^* \rightarrow \mathbb{N}$. The following properties are equivalent:*

- 1- $f \in \mathbb{S}_3(A^*, \mathbb{N})$
- 2- *There exists a finite family $(H_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \text{HOM}(C^*, C^*)$ which fulfils a system of recurrent relations in $(\text{HOM}(C^*, C^*), \circ, \text{Id})$, an element $h \in \text{HOM}(C^*, \mathbb{N})$ and a letter $c \in C$ such that, for every $w \in A^*$:*

$$f(w) = h(H_1(w)(c)).$$

- 3- *f is composition of a DTOL sequence $g : A^* \rightarrow C^*$ by a rational series $h : C^* \rightarrow \mathbb{N}$.*

- 4- *There exists a finite family $(f_i)_{i \in [1, n]}$ of mappings $A^* \rightarrow \mathbb{N}$ fulfilling a system of polynomial recurrent relations and such that $f = f_1$.*

Sketch of proof: (1) \Rightarrow (2) \Rightarrow (3):

this is exactly the sequence of implications (1) \Rightarrow (2) \Rightarrow (3) of Theorem 2, reformulated in the case where B is a unary alphabet.

(3) \Rightarrow (4):

Suppose that f fulfils condition (3): $g = g_1$ for some family $(g_i)_{i \in I}$ of maps $A^* \rightarrow C^*$ which fulfils the system of catenative recurrent relations

$$\forall i \in I, \forall a \in A, \forall w \in A^*, \quad g_i(aw) = \prod_{j=1}^{\ell(i, a)} g_{\alpha(i, a, j)}(w)$$

while the \mathbb{N} -rational series h fulfils:

$$\forall w \in C^*, \quad h(w) = L_0 \cdot H(w) \cdot C_0$$

for some $L_0 \in \mathbb{N}^{1 \times d}$, $H \in \text{HOM}(C^*, \mathbb{N}^{d \times d})$, $C_0 \in \mathbb{N}^{d \times 1}$.

Let us denote by $u_{i, k, \ell}(w)$ the entry with indices (k, ℓ) of the (d, d) matrix $H(g_i(w))$. Since we have

$$\forall i \in I, \forall a \in A, \forall w \in A^*, \quad H(g_i(aw)) = \prod_{j=1}^{\ell(i, a)} H(g_{\alpha(i, a, j)}(w))$$

and the formula expressing a matricial product are polynomials (of degree 2), we get some polynomials $P_{i, k, \ell}$ with undeterminates $X_{i', j', k'}$ ($i' \in I, j', k' \in [1, d]$) such that

$$\forall i \in I, \forall k, \ell \in [1, d], \forall a \in A, \forall i \in I, \quad u_{i, k, \ell}(aw) = P_{i, k, \ell}(\vec{u}(w)) \quad (55)$$

where $\vec{u}(w)$ is the $|I| \cdot d^2$ -tuple of integers $(u_{*, *, *}(w))$. Finally, if we note $M(u_i)$ the (d, d) matrix with coefficients $u_{i, *, *}$, and we note K the linear form $\vec{u} \mapsto L_0 \cdot M(u_i) \cdot C_0$ we get that:

$$\forall i \in I, \forall w \in A^*, \quad f_i(w) = K(\vec{u}(w)). \quad (56)$$

By (55) the sequences $u_{i,k,\ell}$ are polynomially recurrent and by (56), the sequences f_i are polynomially recurrent too.

(4) \Rightarrow (1):

Proposition 53 p. 384 of [FS06] asserts that, for every family $(f_i)_{i \in [1,n]}$ of mappings $A^* \rightarrow \mathbb{N}$ fulfilling a system of polynomial recurrent relations, f_1 belongs to $\mathbb{S}_3(A^*, \mathbb{N})$, in the particular case of a *unary* alphabet A . The extension of the proof, hence of the result, to an arbitrary finite alphabet A is easy.

□

Definition 11. Let \mathbb{S} be a set of mappings $A^* \rightarrow \mathbb{N}$. We denote by $\mathcal{D}(\mathbb{S})$ the set of mappings of the form:

$$f(w) = g(w) - h(w) \quad \text{for all } w \in A^*,$$

for some mappings $g, h \in \mathbb{S}$. We denote by $\mathcal{F}(\mathbb{S})$ the set of mappings of the form:

$$f(w) = \frac{g(w) - h(w)}{f'(w) - g'(w)} \quad \text{for all } w \in A^*,$$

for some mappings $f, g, f', g' \in \mathbb{S}$.

Using point 4 of corollary 1 we can prove the following

Theorem 4. The equality problem is decidable for formal power series in $\mathcal{F}(\mathbb{S}_3(A^*, \mathbb{N}))$.

Unclear how
to decide if
 $f(n)/g(n)$ is
a well-def. seq.

The method consists, in a way similar to [Sén99] or [Hon00], in reducing such an equality problem to deciding whether some polynomial belongs to the ideal generated by a finite set of other polynomials. Let us give a useful algebraic property of sequences of $\mathcal{F}(\mathbb{S}_3(\mathbb{N}, \mathbb{N}))$.

Theorem 5 (Th.11 of [CMP⁺20]). Let $(u_n)_{n \in \mathbb{N}} \in \mathcal{F}(\mathbb{S}_3(\mathbb{N}, \mathbb{N}))$. Then there exists some $k \in \mathbb{N}$ such that the sequences $u_n, u_{n+1}, \dots, u_{n+k-1}$ are algebraically dependent over \mathbb{Q} .

5 Examples and counter-examples

We examine here four examples of mappings $A^* \rightarrow \mathbb{N}$ and locate them in the classes $\mathbb{S}_k(A^*, \mathbb{N})$ or some related classes.

Example 4. The Fibonacci sequence F_n defined by

$$F_0 = 1, F_1 = 1, F_{n+2} = F_{n+1} + F_n \quad \text{for all } n \geq 0$$

is clearly in \mathbb{S}_2 since it fulfills a linear recurrence relation with coefficients in \mathbb{N} (i.e. a catenative recurrence where the alphabet has size 1).

Example 5. Let $G : \{0, 1\}^* \rightarrow \mathbb{N}$ be defined by

$$G(w) = F_{\nu(w)}$$

where $\nu(w)$ is the natural number expressed by w in base 2. Since $\sum_{n=0}^{\infty} \nu(w)w$ is a rational series, G fulfils point 3 of our characterisation of $\mathbb{S}_3(\{0, 1\}^*, \mathbb{N})$.

Example 6. Factorial sequence: $FC(n) := (n + 1)!$.
i.e. $FC(n)$ denotes “factorial $(n + 1)$ ”.
It has level 3.

Proof. Let us define the auxiliary sequence $L(n) := n + 2$. These sequences fulfill the polynomial recurrence relations:

$$L(0) = 2 \quad FC(0) = 1 \quad L(n + 1) = L(n) + 1 \quad FC(n + 1) = L(n) \cdot FC(n).$$

Hence FC is a sequence in \mathbb{S}_3 .

By Theorem 2, FC is the composition of two word-sequences of order 2.
An explicit such decomposition is the following:

$$\begin{aligned} u(0) &= b & v(0) &= ab & u(n + 1) &= u(n) \cdot v(n) & v(n + 1) &= a \cdot v(n) \\ U(\varepsilon) &= 1 & V(\varepsilon) &= 0 \\ U(a \cdot w) &= U(w) + V(w) & V(a \cdot w) &= V(w) & U(b \cdot w) &= U(w) & V(b \cdot w) &= V(w) \end{aligned}$$

We get:

$$u(n) = babaab \cdots ba^n b \quad v(n) = a^{n+1} b$$

$$U(u(n)) = (n + 1)!$$

We thus have decomposed FC into two recurrences of level 2 (going “through” words):

$$\forall n \in \mathbb{N}, \quad FC(n) = U(u(n)).$$

Nevertheless, FC is not the composition of two integer sequences of order 2: this can be deduced from the rate of growth of FC , which is strictly larger than any exponential function, but strictly smaller than any double-exponential function. The technically subtle point consists in showing that no composition of integer sequences in \mathbb{S}_2 can possess such a growth rate (work in preparation).

Example 7. Let us consider the sequence $(D_n)_{n \in \mathbb{N}}$ such that

$$D_0 = 0 \text{ and } \forall n \geq 1, \quad \underline{\underline{D_n = n^n}}.$$

This sequence belongs to \mathbb{S}_4 and does not belong to $\mathcal{F}(\mathbb{S}_3)$.

Proof for level 4

The maps $f(n) := n$ and $g(n) := n$ have level 1, hence $f \in \mathbb{S}_4$ and $g \in \mathbb{S}_3$. By Proposition 66 of [FS06], if $f \in \mathbb{S}_{k+1}$ and $g \in \mathbb{S}_k$ with $k \geq 3$, then $n \mapsto f(n)^{g(n)} \in \mathbb{S}_{k+1}$. It follows that, for the above maps f, g , $D = f^g \in \mathbb{S}_4$.

Decomposition

We take below the convention that $0^0 = 0$.

Let us exhibit two mappings $f \in \mathbb{S}_2(\mathbb{N}, \{a, b, c\}^*)$, $g \in \mathbb{S}_3(\{a, b, c\}^*, \mathbb{N})$ such that

$$\forall n \in \mathbb{N}, D_n = g(f(n)). \quad (57)$$

We define these mappings by two systems of recurrent relations. Let $f, A, C : \mathbb{N} \rightarrow \{a, b, c\}^*$ such that

$$\begin{aligned} f(0) &= b, & A(0) &= a, & C(0) &= c, \\ f(n+1) &= A(n) \cdot f(n) \cdot B(n), & A(n+1) &= A(n), & C(n+1) &= C(n). \end{aligned} \quad (58)$$

Let us introduce a finite alphabet $X = \{x, y\}$ and define mappings $H, K, K', P : \{a, b, c\}^* \rightarrow \text{HOM}(X^*, X^*)$ by the recurrence relations:

$$\begin{aligned} H(\varepsilon) &= \text{Id}, & K(\varepsilon) &= [x, xy], & K'(\varepsilon) &= [x, \varepsilon], & P(\varepsilon) &= [y, x], \\ H(au) &= H(u) \circ H(u), & H(bu) &= P(u) \circ H(u) \circ K'(u), & H(cu) &= H(u) \circ K(u), \\ K(au) &= K(u), & K(bu) &= K(u), & K(cu) &= K(u), \\ K'(au) &= K'(u), & K'(bu) &= K'(u), & K'(cu) &= K'(u), \\ P(au) &= P(u), & P(bu) &= P(u), & P(cu) &= P(u). \end{aligned} \quad (59)$$

where the bracketed notation $[w, w']$ designates the homomorphism $X^* \rightarrow X^*$ that maps x to w and y to w' . We finally define g by:

$$\forall u \in \{a, b, c\}^*, \quad g(u) = H(u)(x).$$

One can check that, for every integers $p, q, n \in \mathbb{N}$:

$$\begin{aligned} H(c^q) &= [x, x^q y], & H(bc^q) &= [x^q, x], & H(a^p bc^q) &= [x^{q^p}, x] \\ f(n) &= a^n bc^n. \\ g(a^n bc^n) &= H(a^n bc^n)(x) = x^{n^n} \\ g(f(n)) &= x^{n^n} \equiv n^n \end{aligned}$$

The system (58) shows that f has level 2. The system (59) is a system of recurrent relations in $\langle \text{HOM}(\{a, b, c\}^*, \{a, b, c\}^*), \circ, \text{Id} \rangle$, which implies, by point (2) of Theorem 2, that g has level 3.

This decomposition also proves that D has level 4: by an adaptation of Proposition 70 of [FS06] to arbitrary alphabets, the composition $D = f \circ g$ of a map of level 2 by a map of level 3, has level 4.

Proof for not level 3:

The proof of theorem 16 of [CMP⁺20] shows that, for every $k \geq 1$, the sequences $D_n, D_{n+1}, \dots, D_{n+k-1}$ are algebraically independent over \mathbb{Q} . Theorem 5 then shows that it cannot have level 3, in the strong sense that: $D \notin \mathcal{F}(\mathbb{S}_3(\mathbb{N}, \mathbb{N}))$.

6 Perspectives

All the open problems mentionned at the end of [FS06] remain unsolved at the moment and deserve interest. Let us mention some open problems which are specific to level 3.

1- The convolution operation is known to preserve \mathbb{S}_2 and it is shown in Proposition 67 of [FS06] that, for every integer $k \geq 2$, if $f \in \mathbb{S}_{k+1}, g \in \mathbb{S}_k$ then $f * g \in \mathbb{S}_{k+1}$.

We wonder if the level 3 is closed under convolution. The same question is open for all levels $k \geq 3$.

2- The equality problem for sequences in $\mathbb{S}_3(A^*, \mathbb{N})$ (for every finite alphabet A) is decidable ([Sén07]). Let us say that $(u(n))_{n \in \mathbb{N}}, (v(n))_{n \in \mathbb{N}}$ are *almost*-equal (which we denote by $u =_a v$) iff

$$\{n \in \mathbb{N} \mid u_n \neq v_n\} \text{ is finite.}$$

The following decision problem is not known to be decidable nor undecidable:

Instance: two sequences $u, v \in \mathbb{S}_3$

Question: $u =_a v$?

Note that the so-called “Skolem problem” (see below) is recursively reducible to this problem. Skolem problem (usually credited to [Sko34]) is the following decision problem:

Instance: two sequences $u, v \in \mathbb{S}_2$

Question: $\exists n \in \mathbb{N}, u(n) = v(n)$?

Skolem’s problem is reducible (by a many-one recursive reduction) to the almost-equivalence problem for sequences of level 3, by the following arguments.

Given $u, v \in \mathbb{S}_2$, let us consider the sequence

$$w(n) := \prod_{i=0}^n (u(i) - v(i)).$$

This sequence w belongs to \mathcal{D}_3 i.e. has the form

$$w(n) = w^+(n) - w^-(n)$$

where $w^+, w^- \in \mathbb{S}_3$ (see the sketch of proof below).

But

$$[\exists n \in \mathbb{N}, u(n) = v(n)] \Leftrightarrow [w^+ =_a w^-].$$

Sketch of proof: The sequence w fulfills the recurrence relation

$$w(n+1) = w(n) \cdot (u(n+1) - v(n+1)).$$

which has the form

$$w(n+1) = P(w(n))$$

for some polynomial P with coefficients in \mathcal{D}_3 . From Theorem 96 of [FS06], step 1 of the proof, it follows that this sequence w belongs to \mathcal{D}_3 . \square

3- A full proof of the extension of Theorem 2 to all levels $k > 1$, as stated by Theorem 6 of the extended abstract [Sén07], is in preparation and should appear soon.

References

- [Aho68] A.V. Aho. Indexed grammars-an extension of context-free grammars. *J. Assoc for Comput. Mach.* 15, pages 647–671, 1968.
- [Aho69] A.V. Aho. Nested stack automata. *J. Assoc for Comput. Mach.* 16, pages 383–406, 1969.
- [AS03] J.P. Allouche and J. Shallit. *Automatic sequences, Theory, applications, generalizations*. Cambridge university press, 2003.
- [Cau02] D. Caucal. On infinite terms having a decidable monadic theory. In *Proceedings MFCS 2002*, pages 165–176. Springer-Verlag, 2002.
- [CMP⁺20] Michaël Cadilhac, Filip Mazowiecki, Charles Paperman, Michał Pilipczuk, and Géraud Sénizergues. On polynomial recursive sequences. In *47th International Colloquium on Automata, Languages, and Programming*, volume 168 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 117, 17. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020.
- [CS21] A. Carayol and O. Serre. Higher-order recursion schemes and their automata models. In *Handbook of automata theory. Vol. II. Automata in Mathematics and Selected Applications*, pages 1295–1341. EMS Press, Berlin, 2021.
- [CW03] A. Carayol and S. Wöhrle. The Caucal hierarchy of infinite graphs in terms of logic and higher-order pushdown automata. In *FST TCS 2003: Foundations of software technology and theoretical computer science*, volume 2914 of *Lecture Notes in Comput. Sci.*, pages 112–123. Springer, Berlin, 2003.
- [Dam82] W. Damm. The IO- and OI-hierarchies. *TCS* 20, pages 95–207, 1982.
- [DG86] W. Damm and A. Goerdt. An automata-theoretical characterization of the OI-hierarchy. *Information and control* 71, pages 1–32, 1986.
- [Eng83] J. Engelfriet. Iterated pushdown automata and complexity classes. In *Proc. 15th Annu ACM Sympos*, pages 365–373. Theory Comput, Assoc. Comp Mach., 1983.
- [Eng91] Joost Engelfriet. Iterated stack automata and complexity classes. *Inform. and Comput.*, 95(1):21–75, 1991.
- [ES77] J. Engelfriet and E.M. Schmidt. IO and OI. I. *J. Comput. System Sci.*, 15(3):328–353, 1977.
- [ES84] J. Engelfriet and G. Slutski. Extended macro grammars and stack controlled machines. *Journal of Computer and System sciences* 29, pages 366–408, 1984.
- [EV86] J. Engelfriet and H. Vogler. Pushdown machines for the macro tree transducer. *TCS* 42, pages 251–368, 1986.
- [Fis68] M.J. Fischer. *Grammars with macro-like productions*. PhD thesis, Harvard University, 1968. See also: Proc. 9th Symp. on Switching and Automata Theory (1968) p. 131-142.
- [FMS14] J. Ferté, N. Marin, and G. Sénizergues. Word-mappings of level 2. *Theory Of Computing Systems*, 54:111–148., 2014.
- [Fra05] S. Fratani. *Automates à piles de piles ... de piles*. PhD thesis, Bordeaux 1 university, 2005.
- [FS06] S. Fratani and G. Sénizergues. Iterated pushdown automata and sequences of rational numbers. *Ann. Pure Appl. Logic*, 141(3):363–411, 2006.
- [Gre70] S. Greibach. Full AFL’s and nested iterated substitution. *Information and Control*, pages 7–35, 1970.
- [Har78] M.A. Harrison. *Introduction to Formal Language Theory*. Addison-Wesley, Reading, Mass., 1978.

- [HO07] Matthew Hague and C.-H. Luke Ong. Symbolic backwards-reachability analysis for higher-order pushdown systems. In *Foundations of Software Science and Computational Structures, 10th International Conference, FOS-SACS 2007*, volume 4423 of *Lecture Notes in Computer Science*, pages 213–227. Springer, 2007.
- [Hon00] J. Honkala. A short solution for the HDT0L sequence equivalence problem. *Theoret. Comput. Sci.*, 244(1-2):267–270, 2000.
- [Hue80] G. Huet. Confluent reductions: abstract properties and applications to term rewriting systems. *JACM*, 27(4):797–821, 1980.
- [KNU02] T. Knapik, D. Niwinski, and P. Urzyczyn. Higher-order pushdown trees are easy. In *FoSSaCs 2002*. LNCS, 2002.
- [KRS97] L. Kari, G. Rozenberg, and A. Salomaa. L systems. In *Handbook of formal languages, Vol. 1*, pages 253–328. Springer, Berlin, 1997.
- [Mas74] A.N. Maslov. The hierarchy of indexed languages of an arbitrary level. *Soviet. Math. Dokl.* 15, pages 1170–1174, 1974.
- [Mas76] A.N. Maslov. Multilevel stack automata. *Problemy Peredachi Informatsii* 12, pages 38–43, 1976.
- [PWZ96] M. Petkovšek, H.S. Wilf, and D. Zeilberger. *A = B*. A K Peters Ltd., Wellesley, MA, 1996.
- [Rig04] Michel Rigo. Automates et systèmes de numération. *Bull. Soc. Roy. Sci. Liège*, 73(5-6):257–270, 2004.
- [Sén99] G. Sénizergues. $T(A) = T(B)$? In P. van Emde Boas J. Wiedermann and M. Nielsen, editors, *Proceedings ICALP’99*, volume 1644, pages 665–675. Lecture Notes in Computer Science, 1999. Detailed version in technical report 1209-99, Pages 1-61. Can be accessed at URL, <http://www.labri.u-bordeaux.fr/~ges>.
- [Sén07] G. Sénizergues. Sequences of level $1, 2, 3, \dots, k, \dots$. In *Proceedings CSR’07*, volume 4649 of *LNCS*, pages 24–32. Springer-Verlag, 2007. Invited talk at CSR’07.
- [Sko34] Th. Skolem. Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen und diophantscher Gleichungen. *Congr. Scand à Stokolm*, 17:163–188, 1934.
- [Sta80] R.P. Stanley. Differentiably finite power series. *European Journal of Combinatorics* 1, pages 175–188, 1980.