Ideal Regular Languages and Strongly Connected Synchronizing Automata

Rogério Reis[a], Emanuele Rodaro[b]

[a]*DCC-Faculdade de Ciências, Universidade do Porto*
[b]*Dipartimento di Matematica, Politecnico di Milano*

Abstract

We introduce the notion of a reset left regular decomposition of an ideal regular language, and we prove that the category formed by these decompositions with the adequate set of morphisms is equivalent to the category of strongly connected synchronizing automata. We show that every ideal regular language has at least one reset left regular decomposition. As a consequence, every ideal regular language is the set of synchronizing words of some strongly connected synchronizing automaton. Furthermore, this one-to-one correspondence allows us to introduce the notion of reset decomposition complexity of an ideal from which follows a reformulation of Černý's conjecture in purely language theoretic terms. Finally, we present and characterize a subclass of ideal regular languages for which a better upper bound for the reset decomposition complexity holds with respect to the general case.

Keywords: Strongly connected synchronizing automaton; Černý's conjecture; Reset word; Ideal regular language.

1. Introduction

Since, in the context of this paper, we do not study automata as language recognizers, instead we are just interested on the action of its transition function δ on the set of states Q, we take, as a deterministic finite automaton (DFA), a tuple $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$, deliberately omitting the initial and final states from the definition. These automata are also referred in literature as semiautomata [3]. But, because, in some point, we also refer to an automaton as a language recognizer, we still call a DFA a tuple $\mathscr{B} = \langle Q', \Sigma', \delta', q_0, F \rangle$, and the language recognized by $\mathscr{B}$ is given by the set $L[\mathscr{B}] = \{u \in \Sigma^* : \delta'(q_0, u) \in F\}$. A DFA $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ is called synchronizing if there exists a word $w \in \Sigma^*$ "sending" all the states into a single state, i.e. $\delta(q, w) = \delta(q', w)$ for all $q, q' \in Q$. Any such word is said to be synchronizing (or reset) for the DFA $\mathscr{A}$. This notion has been widely studied since the work of Černý in 1964 [12] and his well known conjecture regarding an upper bound for the length of the shortest reset word. This conjecture states that any synchronizing

Email addresses: rvr@dcc.fc.up.pt (Rogério Reis), emanuele.rodaro@polimi.it (Emanuele Rodaro)

automata $\mathscr{A}$ with n states admits at least a reset word w with $|w| \leq (n-1)^2$. For more information on synchronizing automata we refer the reader to Volkov's survey [13]. In what follows, when there is no risk of ambiguity on the choice of the action δ of the automaton, we use the notation $q \cdot u$ instead of $\delta(q, u)$. We extend this action to a subset $H \subseteq Q$ in the obvious way $H \cdot u = \{q \cdot u : q \in H\}$ with the convention $\emptyset \cdot u = \emptyset$, and for a language $L \subseteq \Sigma^*$, we use the notation $H \cdot L = \{q \cdot u : q \in H, u \in L\}$. We say that $\mathscr{A}$ is *strongly connected* whenever for any $q, q' \in Q$ there is a word $u \in \Sigma^*$ such that $q \cdot u = q'$. In the realm of synchronizing automata this notion is crucial since it is well known that Černý's conjecture is true if and only if it is true for the class of strongly connected synchronizing automata (see for instance [14]).

In this paper we study the relationship between ideal regular languages and synchronizing automata. A language $I \subseteq \Sigma^*$ is called a *two-sided ideal* if $\Sigma^* I \Sigma^* \subseteq I$. Henceforth, we will only consider regular languages that are two-sided ideals, and for this reason we will simply refer to them as ideals. Denote by $\mathbf{I}_\Sigma$ the class of ideals on an alphabet Σ. For a given synchronizing automaton $\mathscr{A}$, let $\mathrm{Syn}(\mathscr{A})$ be the language of all the words synchronizing $\mathscr{A}$. It is easy to check that $\mathrm{Syn}(\mathscr{A}) = \Sigma^* \mathrm{Syn}(\mathscr{A}) \Sigma^*$ is a regular language that is also a two-sided ideal. This ideal is generated by the set of minimal synchronizing words $G = \mathrm{Syn}(\mathscr{A}) \setminus (\Sigma^+ \mathrm{Syn}(\mathscr{A}) \cup \mathrm{Syn}(\mathscr{A})\Sigma^+)$, i.e. $\mathrm{Syn}(\mathscr{A}) = \Sigma^* G \Sigma^*$. The set of generators G can be also obtained by applying to $\mathrm{Syn}(\mathscr{A})$ the bifix or infix operators defined by Pribavkina et al. [7, 9]. If G is finite, $\mathrm{Syn}(\mathscr{A})$ is called a *finitely generated ideal* and the corresponding automaton $\mathscr{A}$ is named *finitely generated synchronizing automaton* [6, 8, 10]. Maslennikova [4] observed that the minimal deterministic automaton $\mathscr{A}_I = \langle Q', \Sigma, \delta', q_0, \{s\}\rangle$ recognizing an ideal I is synchronizing with a unique final state s, that is fixed by all the elements of Σ. We will refer to such state as *the sink state* for $\mathscr{A}_I$. Furthermore, $\mathrm{Syn}(\mathscr{A}_I) = I$. Thus, each ideal has at least a synchronizing automaton for which I serves as the set of reset words. Therefore, for each ideal I, the set $\mathcal{SA}(I)$ of all the synchronizing automata $\mathscr{B}$ with $\mathrm{Syn}(\mathscr{B}) = I$, is non-empty. This simple observation led Maslennikova to introduce the notion of *reset complexity* of an ideal I as the number of states of the smallest automata in $\mathcal{SA}(I)$, and to show that the reset complexity can be exponentially smaller than the state complexity of the language. Gusev et al. [1] considered the special case of finitely generated synchronizing automata with the set of the reset words that is a principal ideal $P = \Sigma^* w \Sigma^*$ generated by a word $w \in \Sigma^*$. Moreover, the authors presented an algorithm to generate a strongly connected synchronizing automaton $\mathscr{B}_w$ with $\mathrm{Syn}(\mathscr{B}_w) = P$ with the same number of states of $\mathscr{A}_P$, and addressed the question whether, for any ideal I, there is always a strongly connected synchronizing automaton in $\mathcal{SA}(I)$. In Section 3 we answer affirmatively to this question by proving that any ideal I on a non-unary alphabet can serve as a set of the reset words for some strongly connected synchronizing automaton. However, to study and characterize the languages of the reset words of strongly connected synchronizing automata we need to introduce the following provisional class of *strongly connected ideals*:

Definition 1.1. *An ideal I is called strongly connected whenever $I = \mathrm{Syn}(\mathscr{A})$ for some strongly connected synchronizing automaton $\mathscr{A}$.*

The paper is organized as follows. In Section 2 we introduce the notion of a (reset) left regular decomposition of an ideal, and we prove that the strongly connected ideals are exactly the ideals admitting a reset left regular decomposition. We also present an

equivalence between the category of reset left regular decompositions and the category of the strongly connected synchronizing automata on the same alphabet. Using this equivalence, we prove, in Section 3, that each ideal is a strongly connected ideal. Thus, we can introduce the notion of reset regular decomposition complexity of an ideal, and give an equivalent formulation of Černý's conjecture via this notion. We present a general upper bound to this parameter, and show a better bound for the subclass of the ideals that are free from funnels. Finally, we state some open problems and some directions of further investigation.

2. Strongly connected ideals

In this section we explore a connection between strongly connected synchronizing automata and strongly connected ideals. We start by giving some definitions. An *homomorphism* $\varphi : \mathscr{A} \to \mathscr{B}$ between two DFAs, $\mathscr{A} = \langle Q, \Sigma, \delta\rangle$ and $\mathscr{B} = \langle T, \Sigma, \xi\rangle$, is a map $\varphi : Q \to T$ preserving the actions of the two automata, i.e. $\varphi(\delta(q, a)) = \xi(\varphi(q), a)$ for all $a \in \Sigma$, $q \in Q$. We temporarily denote the class of strongly connected ideals on some finite alphabet Σ by $\mathbf{SCI}_\Sigma$. We denote by $\mathbf{SCSA}_\Sigma$ the category of strongly connected synchronizing automata where arrows are homomorphisms. Note that any homomorphism between strongly connected automata is necessarily surjective. For $L \subseteq \Sigma^*$ and $u \in \Sigma^*$, we write $Lu = \{xu : x \in L\}$, $uL = \{ux : x \in L\}$. We recall that the *reverse* operator $\circ^R$ is the bijective map on Σ^* such that $u^R = u_k \dots u_2 u_1$, for any $u = u_1 u_2 \dots u_k$. This operator extends naturally to languages. To characterize the class $\mathbf{SCI}_\Sigma$ we use the following crucial notion of *reset left regular decomposition*.

Definition 2.1. *A left regular decomposition is a finite collection $\mathcal{D} = \{I_1, \dots I_m\}$ of disjoint non-empty left ideals I_i of Σ^* such that:*

i) For any $a \in \Sigma$ and $I_i \in \mathcal{D}$, there is a $I_j \in \mathcal{D}$ such that $I_i a \subseteq I_j$.

The decomposition $\{I_1, \dots I_m\}$ is called a reset left regular decomposition if it also satisfies the following closure condition:

ii) Let $I = I_1 \cup \dots \cup I_m$, if for any $u \in \Sigma^*$ there is an $i \in \{1, \dots, m\}$ such that $Iu \subseteq I_i$, then $u \in I$.

Substituting left ideals by right ideals and $I_i a, Iu$ by aI_i, uI, respectively, we get the dual notion of *(reset) right regular decomposition*. Note that if $\{I_1, \dots I_m\}$ is a reset left (right) regular decomposition, then the condition $Iu \subseteq I_i$ ($uI \subseteq I_i$) implies $u \in I_i$. For if $u \in I_j$ for some $j \in \{1, \dots, m\}$ with $i \neq j$, then we have both $Iu \subseteq I_i$ and $Iu \subseteq I_j$ that implies $I_i \cap I_j \neq \emptyset$, a contradiction. We say that an ideal I has a reset left (right) regular decomposition if there is a reset left (right) regular decomposition $\{I_1, \dots I_m\}$ such that $I = I_1 \cup \dots \cup I_m$. The *order* of $\{I_1, \dots I_m\}$ is the cardinality m of the family. Denote by $\mathbf{RLD}_\Sigma$ ($\mathbf{RRD}_\Sigma$) the category of the reset left regular decompositions, where an arrow $f : \{I_1, \dots I_m\} \to \{J_1, \dots J_\ell\}$ is a function among these two sets with the property that for any left ideal I of $\{I_1, \dots I_m\}$ we have $I \subseteq f(I)$. Note that, given a reset left regular decomposition $\{I_1, \dots I_m\}$, then $\{I_1^R, \dots I_m^R\}$ is a reset right regular decomposition. Thus, the reverse map $\circ^R$ is a bijection between the objects of $\mathbf{RLD}_\Sigma \to \mathbf{RRD}_\Sigma$. We have the following characterization.

Theorem 2.2. *An ideal I is strongly connected if and only if it has a reset left regular decomposition. Moreover $\boldsymbol{RLD}_\Sigma$ and $\boldsymbol{SCSA}_\Sigma$ are equivalent categories via the two functors $\mathcal{A}$ and $\mathcal{I}$ defined by:*

$$\begin{array}{rcl} \mathcal{A} : \boldsymbol{RLD}_\Sigma & \longrightarrow & \boldsymbol{SCSA}_\Sigma \\ \mathcal{D} = \{I_1, \dots I_m\} & \longmapsto & \mathcal{A}(\mathcal{D}) = \langle \mathcal{D}, \Sigma, \eta \rangle, \end{array}$$

with $\eta(I_i, a) = I_j$ for $a \in \Sigma$ if and only if $I_i a \subseteq I_j$, and if $f : \{I_1, \dots I_m\} \to \{J_1, \dots J_\ell\}$ then $\mathcal{A}(f)$ is the homomorphism $\varphi : \mathcal{A}(\{I_1, \dots I_m\}) \to \mathcal{A}(\{J_1, \dots J_\ell\})$ defined by $\varphi(I_i) = f(I_i)$;

$$\begin{array}{rcl} \mathcal{I} : \boldsymbol{SCSA}_\Sigma & \longrightarrow & \boldsymbol{RLD}_\Sigma \\ \mathscr{A} & \longmapsto & \mathcal{I}(\mathscr{A}) = \{\mathcal{I}(\mathscr{A})_q : q \in Q\}, \end{array}$$

where $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$, $\mathcal{I}(\mathscr{A})_q = \{u \in \Sigma^* : \delta(Q, u) = q\}$, and if $\varphi : \mathscr{A} \to \mathscr{B}$ is an arrow between $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ and $\mathscr{B} = \langle T, \Sigma, \xi \rangle$, then $\mathcal{I}(\varphi)$ is the arrow sending each $\mathcal{I}(\mathscr{A})_q$ into $\mathcal{I}(\mathscr{B})_{\varphi(q)}$.

Proof. Let us prove the first claim of the theorem. Let $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ be a strongly connected synchronizing automata with $\mathrm{Syn}(\mathscr{A}) = I$. For each $q \in Q$, let:

$$I_q = \mathcal{I}(\mathscr{A})_q = \{u \in I : Q \cdot u = q\}.$$

We claim that $\{I_q : q \in Q\}$ is a reset left regular decomposition for I. It is obvious that I_q are left ideals since for any $u \in I_q$ and $v \in \Sigma^*$, we get $Q \cdot vu \subseteq Q \cdot u = \{q\}$, i.e. $Q \cdot vu = \{q\}$. Let $q, q' \in Q$ with $q \neq q'$. Assume $I_q \cap I_{q'} \neq \emptyset$, and so consider $u \in I_q \cap I_{q'}$. By definition, we have $q = Qu = q'$, which is a contradiction. Hence, $I_q \cap I_{q'} = \emptyset$. Clearly $\uplus_{q \in Q} I_q \subseteq I$. Conversely if $u \in I$, since it is a reset word, then $Qu = q'$ for some $q' \in Q$, i.e. $u \in I_{q'}$ and so we have the decomposition $\uplus_{q \in Q} I_q = I$. Moreover, for any $a \in \Sigma$, if $u \in I_q$, then $Q \cdot ua = q \cdot a$, thus $I_q a \subseteq I_{q \cdot a}$ and so condition i) of the Definition 2.1 is fulfilled. Thus, it remains to prove that condition ii) is also satisfied. Let us assume that $Iw \subseteq I_{\overline{q}}$ for some $\overline{q} \in Q$. We claim that $q \cdot w = \overline{q}$ for any $q \in Q$, whence $w \in \mathrm{Syn}(\mathscr{A}) = I$. Take any $u' \in I$, thus $Q \cdot u' = q'$ for some $q' \in Q$. Since $\mathscr{A}$ is strongly connected, there is $u'' \in \Sigma^*$ such that $q' \cdot u'' = q$. Thus, $u = u'u'' \in I$ satisfies $Q \cdot u = q$. Since $Iw \subseteq I_{\overline{q}}$ we get $\overline{q} = Q \cdot (uw) = q \cdot w$, i.e. $q \cdot w = \overline{q}$.

Conversely, suppose that I has a reset left regular decomposition $\mathcal{D} = \{I_1, \dots, I_m\}$. We associate a DFA $\mathcal{A}(\mathcal{D}) = \langle \mathcal{D}, \Sigma, \eta \rangle$ in the following way. By condition i) of Definition 2.1, for any $I_i \in \mathcal{D}$ and $a \in \Sigma$ there is a $I_j \in \mathcal{D}$ with $I_i \cdot a \subseteq I_j$. Thus, we put $\eta(I_i, a) = I_j$. This function is well defined. Indeed, let j, k be two indices with $j \neq i$, such that $I_i \cdot a \subseteq I_j, I_k$. Then $I_i \cdot a \subseteq I_j \cap I_k$, from which we get $I_j \cap I_k \neq \emptyset$, which is a contradiction. Hence, $\mathcal{A}(\mathcal{D})$ is a well defined DFA. It is straightforward to check that $\eta(I_i, u) = I_k$ for $u \in \Sigma^*$ if and only if $I_i u \subseteq I_k$. Now, let us prove that $\mathcal{A}(\mathcal{D})$ is strongly connected. Take $I_i, I_j \in \mathcal{D}$ and let $w \in I_j$. Since I_j is a left ideal, then $I_i w \subseteq I_j$. Hence $I_i w \subseteq I_j$ implies $\eta(I_i, w) = I_j$ and so $\mathcal{A}(\mathcal{D})$ is strongly connected. We need to prove that $I \subseteq \mathrm{Syn}(\mathcal{A}(\mathcal{D}))$. Let $u \in I$. Since $\mathcal{D}$ is a decomposition, $u \in I_j$ for some $I_j \in \mathcal{D}$. Since I_j is a left ideal, we get $I_i u \subseteq I_j$ for any $I_i \in \mathcal{D}$. Hence $\eta(I_i, u) = I_j$ for all $I_i \in \mathcal{D}$, i.e. $u \in \mathrm{Syn}(\mathcal{A}(\mathcal{D}))$. Conversely, let $u \in u \in \mathrm{Syn}(\mathcal{A}(\mathcal{D}))$. By definition, $\eta(I_i, u) = I_j$ for some $I_j \in \mathcal{D}$ and for all $I_i \in \mathcal{D}$. Therefore, $I_i u \subseteq I_j$ that implies $Iu \subseteq I_j$ and so by ii) of Definition 2.1 we get $u \in I$.

Let us now prove the equivalence of the two categories. Let $\mathcal{D} = \{I_1, \ldots, I_m\}$, $\mathcal{C} = \{J_1, \ldots, J_\ell\}$ be two objects of $\mathbf{RLD}_\Sigma$ and let us prove first that if we have the arrow $f : \mathcal{D} \to \mathcal{C}$, then $\mathcal{A}(f) = \varphi$ is a homomorphism between $\mathcal{A}(\mathcal{D}) = \langle Q, \Sigma, \delta\rangle$ and $\mathcal{A}(\mathcal{C}) = \langle T, \Sigma, \eta\rangle$. Let I_i, $a \in \Sigma$ and $\delta(I_i, a) = I_j$, and put $J_h = f(I_i)$, $J_k = f(I_j)$. By definition we have $I_i a \subseteq I_j$. Since $\varphi(I_i) = J_h$, $\varphi(I_j) = J_k$, $I_i \subseteq J_h$ and $I_j \subseteq J_k$, then $I_i a \subseteq J_h a$ and $I_i a \subseteq I_j \subseteq J_k$ that yields $J_h a \subseteq J_k$. Hence

$$\varphi(\delta(I_i, a)) = \varphi(I_j) = J_k = \eta(J_h, a) = \eta(\varphi(I_i), a),$$

that shows that $\mathcal{A}(f) = \varphi$ is a homomorphism. Let $g : \{J_1, \ldots J_\ell\} \to \{S_1, \ldots, S_t\}$ be another arrow, then it is easy to check that $\mathcal{A}(g \circ f) = \mathcal{A}(g) \circ \mathcal{A}(f)$. Therefore $\mathcal{A} : \mathbf{RLD}_\Sigma \to \mathbf{SCSA}_\Sigma$ is a functor. Let us prove that $\mathcal{I} : \mathbf{SCSA}_\Sigma \to \mathbf{RLD}_\Sigma$ is also a functor. If $\varphi : \mathscr{A} \to \mathscr{B}$ is a homomorphism of the DFAs $\mathscr{A} = \langle Q, \Sigma, \delta\rangle$ and $\mathscr{B} = \langle T, \Sigma, \eta\rangle$, then, for any $q \in Q$ and $u \in \Sigma^*$ such that $\delta(Q, u) = \{q\}$, since φ is surjective, we get $\{\varphi(q)\} = \varphi(\delta(Q, u)) = \eta(T, u)$. Thus, $\mathcal{I}(\mathscr{A})_q \subseteq \mathcal{I}(\mathscr{B})_{\varphi(q)}$, whence $\mathcal{I}(\varphi) : \mathcal{I}(\mathscr{A}) \to \mathcal{I}(\mathscr{B})$ is the corresponding arrow in $\mathbf{RLD}_\Sigma$. Furthermore, if $\psi : \mathscr{B} \to \mathscr{C}$ is another arrow, using the previous fact it is easy to check that $\mathcal{I}(\psi \circ \varphi) = \mathcal{I}(\psi) \circ \mathcal{I}(\varphi)$, which completes the proof that $\mathcal{I}$ is a functor. With the previous construction, it is straightforward to check that $\mathcal{A}(\mathcal{I}(\mathscr{A})) \simeq \mathscr{A}$ and $\mathcal{I}(\mathcal{A}(\mathcal{D})) \simeq \mathcal{D}$, and it is also straightforward to verify that $\mathcal{I}\mathcal{A} = id_{\mathbf{RLD}_\Sigma}$ while the function which associates to each object $\mathscr{A}$ the arrow given by the isomorphism $\mathcal{A}(\mathcal{I}(\mathscr{A})) \simeq \mathscr{A}$, is a natural isomorphism between the functors $id_{\mathbf{SCSA}_\Sigma}$ and $\mathcal{A}\mathcal{I}$, whence $\mathbf{RLD}_\Sigma, \mathbf{SCSA}_\Sigma$ are equivalent categories. □

Henceforth, we will denote by $\{I_i\}_{i\in F}$ a collection of left (or right) ideals indexed by a minimal set F. The following corollary characterizes the case of strongly connected ideals on a unary alphabet.

Corollary 2.3. *Let I be an ideal over a unary alphabet $\{a\}$. Then I is strongly connected if and only if $I = \{a\}^*$.*

Proof. Since the alphabet is unary we have $I = a^* a^m a^*$ for some $m \geq 0$. Suppose that I is strongly connected, then by Theorem 2.2 there is a reset left regular decomposition $\{I_i\}_{i\in F}$ of I. Assume $a^m \in I_j$ for some $j \in F$. We claim $|F| = 1$. Indeed, since I_j is a left ideal we have $a^* a^m \subseteq I_j$, hence $I = a^* a^m a^* = a^* a^m \subseteq I_j$, i.e. $I = I_j$. Therefore, by Theorem 2.2 the only strongly connected synchronizing automaton having I as the set of reset words is the automaton with one state and a loop labelled by a. Hence $I = a^*$. On the other hand, if $I = a^*$ then I is the set of reset words of the synchronizing automaton with one state and a loop labelled by a, which is strongly connected, i.e. I is strongly connected. □

From this corollary we may assume, henceforth, that the ideals considered are taken over an alphabet Σ with $|\Sigma| > 1$.

Given a strongly connected ideal I with $\mathrm{Syn}(\mathscr{B}) = I$, for some strongly connected synchronizing automaton $\mathscr{B} = \langle Q, \Sigma, \delta\rangle$, there is an obvious way to calculate the associated reset left regular decomposition $\mathcal{I}(\mathscr{B})$ using Theorem 2.2. It is well known that I is recognized by the power automaton of $\mathscr{B}$ defined by

$$\mathcal{P}(\mathscr{B}) = \langle 2^Q, \Sigma, \delta, Q, \{\{q\} : q \in Q\}\rangle,$$

where 2^Q denotes the set of subsets of Q, the initial state is the set Q and the final set of states is formed by all the singletons $\{\{q\} : q \in Q\}$. Thus, for each $q \in Q$ we may consider the DFA $\mathcal{P}(\mathscr{B})_q = \langle 2^Q, \Sigma, \delta, Q, \{q\}\rangle$, where the associated reset left regular decomposition is given by $\mathcal{I}(\mathscr{B}) = \{L[\mathcal{P}(\mathscr{B})_q]\}_{q\in Q}$. A first and quite natural issue is to calculate the reset left regular decompositions of the reset words of the well known Černý series $\mathscr{C}_n = \langle\{1,\ldots,n\},\{a,b\},\delta_n\rangle$, where a acts like a cyclic permutation $\delta_n(i,a) = i+1$, for $i = 1,\ldots,n-1$ and $\delta_n(n,a) = 1$, while b fixes all the states except the last one: $\delta_n(i,b) = i$ for $i = 1,\ldots,n-1$ and $\delta_n(n,b) = 1$ (see Fig. 1).

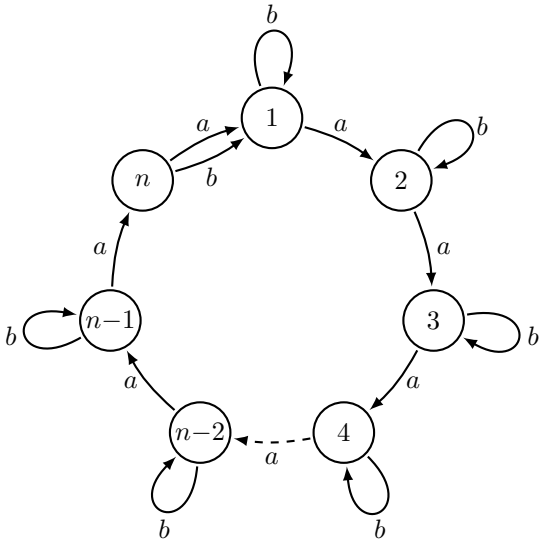

Figure 1: Černý's series $\mathscr{C}_n$

For example, in the case of $\mathscr{C}_4$ the associated reset left regular decomposition is:

$$\begin{aligned}
L[\mathcal{P}(\mathscr{C}_4)_1] &= (((a^*b)(b+ab+a^4)^*(a^3b+(a^2b(b+a^2)^*ab)))((b+ab^*a^3)+\\
&\quad +((ab^*ab)(b+a^2)^*)ab))^*(ab^*a^2b)(b+((ab^*ab^*)(a(a+b))))^*\\
L[\mathcal{P}(\mathscr{C}_4)_2] &= L[\mathcal{P}(\mathscr{C}_4)_1]ab^*\\
L[\mathcal{P}(\mathscr{C}_4)_3] &= L[\mathcal{P}(\mathscr{C}_4)_1]ab^*ab^*\\
L[\mathcal{P}(\mathscr{C}_4)_4] &= L[\mathcal{P}(\mathscr{C}_4)_1]ab^*ab^*a.
\end{aligned}$$

In general, for $\mathscr{C}_n$ it is straightforward to see that

$$|\delta_n(\{1,\ldots,n\},ux)| = 1 \text{ with } |\delta_n(\{1,\ldots,n\},u)| > 1, \text{ for some } u \in \{a,b\}^*, x \in \{a,b\}$$

holds if and only if $\delta_n(\{1,\ldots,n\},u) = \{n,1\}$ and $x = b$. Thus, for any word w such that $|\delta_n(Q,w)| = 1$, there is a prefix $w'b$ of w with $\delta_n(Q,w') = \{n,1\}$. Hence, in the general

case the decompositions are given by:

$$\begin{aligned} L[\mathcal{P}(\mathscr{C})_1] &= \{w \in \Sigma^* \,:\, \delta_n(\{1,\dots,n\}, w) = \{1\}\} \\ L[\mathcal{P}(\mathscr{C})_\ell] &= L[\mathcal{P}(\mathscr{C})_1](ab^*)^{\ell-1} \quad \text{for } \ell = 2,\dots,n-1 \\ L[\mathcal{P}(\mathscr{C})_n] &= L[\mathcal{P}(\mathscr{C})_1](ab^*)^{n-2}a. \end{aligned}$$

By Theorem 2.2 if the ideal I is strongly connected, then the set $\mathcal{R}(I)$ of all the reset left regular decompositions of I is non-empty. The following lemma shows a closure property of reset left (right) regular decompositions.

Lemma 2.4. *Let $\{I_i\}_{i\in F}$ be a reset left (right) regular decomposition of I and let $\{J_k\}_{k\in H}$ be a left (right) regular decomposition of an ideal J. If $I \subseteq J$, then the non-empty elements of $\{I_i \cap J_k\}_{i\in F, k\in H}$ form a reset left (right) regular decomposition of I.*

Proof. Let us consider just the left case. Let $T \subseteq F \times H$ be the set of all pairs of indices (i,j) for which $I_i \cap J_j \neq \emptyset$ and rename the set $\{I_i \cap J_k\}_{(i,k)\in T}$ by $\{S_j\}_{j\in T}$. It is clear that each S_j is a left ideal, that $S_j \cap S_t = \emptyset$ for $j \neq t$, and that $\uplus_{j\in T} S_j = I$. Condition i) is also verified. Take any S_j and suppose that $S_j = I_i \cap J_k$ for some $(i,k) \in T$, and let $a \in \Sigma$. Then $I_i a \subseteq I_s$, $J_k a \subseteq J_t$ for some $s \in F, t \in H$. Hence $(I_i \cap J_k)a = I_i a \cap J_k a \subseteq I_s \cap J_t = S_h$ for some $h \in T$, i.e. $S_j a \subseteq S_h$. To prove that reset condition ii) is also fulfilled, assume $Iu \subseteq S_t$ for some $t \in T$ and $u \in \Sigma^*$. Thus, $S_t = I_i \cap J_k$, for some $i \in F, k \in H$, hence $S_t \subseteq I_i$ which implies $Iu \subseteq I_i$. Hence, $u \in I$ since $\{I_i\}_{i\in F}$ is a reset left regular decompositions of I. □

Given $\mathcal{I}, \mathcal{J} \in \mathcal{R}(I)$ with $\mathcal{I} = \{I_i\}_{i\in F}$ and $\mathcal{J} = \{J_k\}_{k\in H}$, by Lemma 2.4, the family $\mathcal{I} \wedge \mathcal{J} = \{I_i \cap J_k\}_{i\in F, k\in H}$ is still a reset left regular decomposition. Thus we have the following immediate result.

Corollary 2.5. *The family of the reset left regular decompositions of a strongly connected ideal I is a $\wedge$-semilattice.*

Put $\|I\| = \min\{|u| : u \in I\}$. Since Černý's conjecture holds if and only if it holds for strongly connected synchronizing automata, we can reformulate it in pure language theoretic terms.

Proposition 2.6. *Černý's conjecture is true for strongly connected synchronizing automata if and only if, for any strongly connected ideal I and any reset left regular decomposition $\{I_i\}_{i\in F}$ of I, we have:*

$$|F| \geq \sqrt{\|I\|} + 1.$$

Proof. Suppose that Černý's conjecture is true for strongly connected synchronizing automata. Let I be a strongly connected ideal and let $\{I_i\}_{i\in F}$ be a reset left regular decomposition of I. Let $\mathcal{A}(\{I_i\}_{i\in F})$ be the synchronizing automata associated to this decomposition, as in Theorem 2.2. This automaton has $|F|$ states, hence there is a synchronizing word $u \in \mathrm{Syn}(\mathcal{A}(\{I_i\}_{i\in F})) = I$ with $|u| \leq (|F|-1)^2$. Thus $|F| \geq \sqrt{|u|}+1 \geq \sqrt{\|I\|}+1$.

Conversely, take any strongly connected synchronizing automata $\mathscr{A}$ with n states, and let $\mathcal{I}(\mathscr{A})$ be the associated reset left regular decomposition of $I = \mathrm{Syn}(\mathscr{A})$ as in

Theorem 2.2. Since the order of this decomposition is n, then $n \geq \sqrt{\|I\|} + 1$. Thus we have that there is a $u \in \mathrm{Syn}(\mathscr{A})$ with $|u| \leq (n-1)^2$ and Černý's conjecture holds for $\mathscr{A}$. □

3. All ideals are strongly connected

The notion of strongly connected ideals ($\mathbf{SCI}_\Sigma$) has been temporarily introduced in Section 2 to study the relationship between strongly connected synchronizing automata and ideals. In this section we show that $\mathbf{SCI}_\Sigma = \mathbf{I}_\Sigma$ for a non-unary alphabet Σ. As an immediate consequence we obtain that each ideal I has at least a strongly connected synchronizing automaton with set of reset words I. However, the number of states of such automaton is in general very big. Indeed, in Corollary 3.5 we prove an upper bound that is a double exponential with respect to the state complexity of the reverse of the ideal I. At the end of this section we show a particular subclass of ideals for which this bound is slightly better. Beside the better bound, this class is introduced to present a different way to obtain reset right (left) regular decompositions; this may shed some light in finding a general approach to build these decompositions.

Before we prove the main result of this section we introduce some notions that are crucial for the sequel. For this purpose, let us fix a synchronizing automaton $\mathscr{C} = \langle Q, \Sigma, \delta \rangle$ with n states and a sink state s. Note that for such an automaton $|Q \cdot u| = 1$ if and only if $Q \cdot u = \{s\}$. Fix a pair (H, u) with $u \in \Sigma^*$, and $H \subseteq Q$, and assume $u = u_1 \dots u_r$ for $u_1, \dots, u_r \in \Sigma$ and $r = |u|$. We use the standard notation $u[i, j]$, for $0 \leq i < j \leq r$, to indicate the factor $u_i u_{i+1} \dots u_j$ if $i > 0$, otherwise $u[0, j] = u_1 \dots u_j$ with the convention that $u[0, 0] = \epsilon$ and $u[i, i] = u_i$ if $i > 0$. There is a unique tuple $0 \leq i_1 < i_2 < \ldots < i_k = r$ of indices such that:

$$|H| = |H \cdot u[0, i_1]| > |H \cdot u[0, i_2]| > \ldots > |H \cdot u[0, i_k]|,$$

and for any $i_t < j \leq i_{t+1}$ with $1 \leq t \leq k-1$, we have $|H \cdot u[0, j]| = |H \cdot u[0, i_{t+1}]|$. In other words these indices pinpoint the longest prefixes $u[0, j]$ of u such that $|H \cdot u[0, j]| > |H \cdot u[0, j+1]|$. We call such tuple the *ladder decomposition* of the pair (H, u). The *ladder map* with respect to the word u is the function $\lambda_u : 2^Q \to 2^{2^Q}$ defined by

$$\lambda_u(H) = \{H \cdot u[0, i_1], H \cdot u[0, i_2], \dots, H \cdot u[0, i_k]\},$$

where $i_1 < i_2 < \ldots < i_k$ is the ladder decomposition of (H, u). Notice that the range of λ_u is contained in the set $\mathcal{L}(Q)$ formed by families of subsets $\{H_1, \dots, H_s\}$ with $|H_1| > |H_2| > \ldots > |H_s|$. Observe that we have the following upper bounds

$$|\mathcal{L}(Q)| \leq \prod_{i=1}^{|Q|} \left(\binom{|Q|}{i} + 1 \right) \leq 2^{n^2}. \tag{1}$$

where $n = |Q|$, and the "1" inside the formula is due to the fact that it is not mandatory to choose all the sets of possible sizes. We now introduce a partial internal binary operation $\star$ on the set $\mathcal{L}(Q)$. Let $\mathcal{V}_1 = \{T_1, \dots, T_m\} \in \mathcal{L}(Q)$ and $\mathcal{V}_2 = \{H_1, \dots, H_s\} \in \mathcal{L}(Q)$ with $|T_1| > |T_2| > \ldots > |T_m| \geq |H_1| > |H_2| > \ldots > |H_s|$, then:

$$\mathcal{V}_1 \star \mathcal{V}_2 = \begin{cases} \{T_1, \dots, T_{m-1}, H_1, \dots, H_s\}, & \text{if } |T_m| = |H_1| \\ \{T_1, \dots, T_m, H_1, \dots, H_s\}, & \text{otherwise.} \end{cases}$$

Lemma 3.1. *With the above notation for any $u, v \in \Sigma^*$ we have:*

$$\lambda_{vu}(T) = \lambda_v(T) \star \lambda_u(T \cdot v).$$

PROOF. It follows from the definitions. □

We introduce an analogous function that is the key to prove the main result of this section. Let $\mathbb{Z}_m$, $m \geq 2$, be the ring of the integers modulo m. For an integer $t \geq 1$, $[2^Q]_t$ denotes the set of subsets of Q of cardinality t. Let $\mathbb{T}_t = \mathbb{Z}_m([2^Q]_t \uplus \Sigma)$ be the free $\mathbb{Z}_m$-module on $[2^Q]_t \uplus \Sigma$. Let $H \in [2^Q]_t$, $a \in \Sigma$ and $p \in \mathbb{Z}_m([2^Q]_t \uplus \Sigma)$. We denote by $p(H)$ and $p(a)$ the coefficients in $\mathbb{Z}_m$ of p with terms H and a, respectively. Note that p can be decomposed as the sum of the two following terms

$$p\langle Q\rangle = \sum_{H \subseteq Q} p(H)H, \quad p\langle \Sigma\rangle = \sum_{a \in \Sigma} p(a)a.$$

Consider an element $u \in \Sigma^*$ and $H \subseteq Q$ with $|H| > 1$. The *last set* of the pair (H, u) is the smallest set $S \in \lambda_u(H)$ with $|S| \geq 2$. Hence, there is a maximal factor $u[i, j]$ of u such that $|S| = |H \cdot u[0, k]|$ for all $i \leq k \leq j$ and $S = H \cdot u[0, j]$. In case $|S| = |H|$ we may assume $i = 0$. The *tail* of (H, u) is the element of $\mathcal{T}(H, u) \in \mathbb{Z}_m([2^Q]_t \uplus \Sigma)$ with $t = |S| \geq 2$ defined by

$$\mathcal{T}(H, u) = \begin{cases} \sum_{k=i}^{j-1} (H \cdot u[0, k] + u[k + 1, k + 1]), & \text{if } u[0, j] = u \\ \sum_{k=i}^{j} (H \cdot u[0, k] + u[k + 1, k + 1]), & \text{otherwise.} \end{cases}$$

Put $\mathbb{T}(m) = \uplus_{t=2}^{m} \mathbb{T}_t$. This set is the disjoint union of free $\mathbb{Z}_m$-modules, and so we do not identify any pair of elements. For this reason, for each $2 \leq t \leq n$, we denote by 0_t the zero of $\mathbb{T}_t$. For an element $\mathcal{T} \in \mathbb{T}_t$ the integer $t \geq 2$ is called *the index* of $\mathcal{T}$ and it is denoted by $\mathrm{Ind}(\mathcal{T})$. We may endow $\mathbb{T}(m)$ with a structure of commutative semigroup by introducing an internal binary operation $\diamond$ defined in the following way. Let $\mathcal{T}_1 \in \mathbb{T}_i, \mathcal{T}_2 \in \mathbb{T}_j$, then

$$\mathcal{T}_1 \diamond \mathcal{T}_2 = \begin{cases} \mathcal{T}_1, & \text{if } i < j \\ \mathcal{T}_2, & \text{if } j < i \\ \mathcal{T}_1 + \mathcal{T}_2, & \text{if } i = j. \end{cases}$$

Note that $(\mathbb{T}(m), \diamond)$ is a commutative monoid with identity 0_n that has also a graded structure with respect to the semilattice $([2, n], \min)$, i.e. $\mathbb{T}_i \diamond \mathbb{T}_j \subseteq \mathbb{T}_{\min\{i,j\}}$. Using the tail of the pair (H, u), for any $H \subseteq Q$, $u \in \Sigma^*$, we define the *tail map* $\tau_u : 2^Q \to \mathbb{T}(m)$ by:

$$\tau_u(H) = \begin{cases} \mathcal{T}(H, u), & \text{if } |H| > 1 \\ 0_n, & \text{otherwise.} \end{cases}$$

We have the following lemma.

Lemma 3.2. *With the above notations for any $u, v \in \Sigma^*$ and $H \subseteq Q$ we have:*

$$\tau_{vu}(H) = \tau_v(H) \diamond \tau_u(H \cdot v).$$

PROOF. The equality in the statement clearly holds for H with $|H| = 1$, hence we can assume $|H| \geq 2$. We consider the following two cases.

- If $|H \cdot v| = 1$, then it is easy to see that $\tau_{vu}(H) = \tau_v(H)$ holds for every $u \in \Sigma^*$. Thus, since $\tau_u(H \cdot v) = 0_n$ for every $u \in \Sigma^*$, we get the statement $\tau_{vu}(H) = \tau_v(H) = \tau_v(H) \diamond 0_n = \tau_v(H) \diamond \tau_u(H \cdot v)$.
- If $|H \cdot v| \geq 2$. Let $(vu)[i, j]$ be the factor corresponding to the last set of (H, vu). Note that $\mathrm{Ind}(\mathcal{T}(H, v)) \geq \mathrm{Ind}(\mathcal{T}(H \cdot v, u))$. Therefore we consider two further cases. If $\mathrm{Ind}(\mathcal{T}(H, v)) > \mathrm{Ind}(\mathcal{T}(H \cdot v, u))$, then $i > |v|$, and so we get

  $$\tau_{vu}(H) = \tau_u(H \cdot v) = \tau_v(H) \diamond \tau_u(H \cdot v)$$

  whence in this case the statement of the lemma holds. Otherwise, we can assume $\mathrm{Ind}(\mathcal{T}(H, v)) = \mathrm{Ind}(\mathcal{T}(H \cdot v, u))$. Thus, in this case we get

  $$\mathcal{T}(H, v) = \sum_{k=i}^{|v|-1} (H \cdot v[0, k] + v[k + 1, k + 1])$$
  $$\mathcal{T}(H \cdot v, u) = \sum_{k=0}^{j-|v|} ((H \cdot v) \cdot u[0, k] + u[k + 1, k + 1])$$

  Since the two indices of the tails are the same, by a simple computation we get

  $$\tau_v(H) \diamond \tau_u(H \cdot v) = \mathcal{T}(H, v) + \mathcal{T}(H \cdot v, u) = \mathcal{T}(H, vu) = \tau_{vu}(H).$$

  which concludes the proof of the lemma.

□

For any sets A, B, $\mathrm{Hom}(A, B)$ denotes the set of all the maps $f : A \to B$, and, as usual, $\mathrm{Ker}(f) \subseteq A \times A$ denotes the kernel of the function f. We say that an equivalence relation $\sigma \subseteq A \times A$ has *finite index* whenever the quotient A/σ is finite. The following lemma shows a nice property shared by both the tail and the ladder map.

Lemma 3.3. *Using the previous notation, consider the following maps:*

1. $\mu : \Sigma^* \to \mathrm{Hom}(2^Q, \mathbb{T}(m))$ *defined by* $\mu(u) = \tau_u$,
2. $\psi : \Sigma^* \to \mathrm{Hom}(2^Q, \mathcal{L}(Q))$ *defined by* $\psi(u) = \lambda_u$.

Then, $\mathrm{Ker}(\mu), \mathrm{Ker}(\psi)$ *are left congruences on* Σ^*.

Proof. We prove that $\mathrm{Ker}(\mu)$ is a left congruence. Let $a \in \Sigma$ and $u, v \in \Sigma^*$ such that $\mu(u) = \mu(v)$. Hence, $\tau_u = \tau_v$ and so applying Lemma 3.2 two times we obtain

$$\tau_{au}(T) = \tau_a(T) \diamond \tau_u(T \cdot a) = \tau_a(T) \diamond \tau_v(T \cdot a) = \tau_{av}(T),$$

for any $T \subseteq Q$, whence $\tau_{au} = \tau_{av}$, i.e. $\mu(au) = \mu(av)$. Similarly, the other case follows from Lemma 3.1. □

We are now ready to prove the main theorem of this section.

Theorem 3.4. *Let* $I \subseteq \Sigma^*$ *with* $|\Sigma| \geq 2$ *be an ideal, then* I *is strongly connected.*

PROOF. Put $J = I^R$. Let $\mathscr{A}_J = \langle Q, \Sigma, \delta, q_0, \{s\}\rangle$ be the minimal DFA recognizing J and let μ be the map of Lemma 3.3 defined with respect to $\mathscr{A}_J$ and put $m = \|I\| + 1$, with $\|I\| = \min\{|u| : u \in I\}$. We claim that the equivalence classes $\{J_i\}_{i\in F}$ of the relation $\sim = (J \times J) \cap \mathrm{Ker}(\mu)$ form a reset right regular decomposition of J. Since $\mathrm{Hom}(2^Q, \mathbb{T}(m))$ is finite, then $\mathrm{Ker}(\mu)$ has finite index, whence $\sim$ has also finite index. Since $J = \mathrm{Syn}(\mathscr{A}_J)$, for any $H \subseteq Q$ and $u \in J$ we have $H \cdot u = \{s\}$. Hence, it is straightforward to check that $\tau_u = \tau_{uv}$ for any $v \in \Sigma^*$. Therefore the $\sim$-classes $\{J_i\}_{i\in F}$ are right ideals and form a finite partition of J. Furthermore, by Lemma 3.3, $\mathrm{Ker}(\mu)$ is a left congruence of Σ^*, and so, since J is an ideal, it is also a left congruence on J, hence, for any J_i and $a \in \Sigma$, we get $aJ_i \subseteq J_j$ for some $j \in F$. Thus, condition i) of Definition 2.1 is satisfied and $\{J_i\}_{i\in F}$ is a right regular decomposition. Let us now prove that condition ii) is also satisfied. Assume, contrary to our claim, that there are $i \in F$ and $v \in \Sigma^* \setminus J$ such that $vJ \subseteq J_i$. Let $H = Q \cdot v$. Since $\mathrm{Syn}(\mathscr{A}_J) = J$ we get $|H| > 1$. Thus, let

$$t = \min\{|H \cdot r| : r \in \Sigma^* \text{ such that } H \cdot r \neq \{s\}\}$$

and take any $S \in \{H \cdot r : r \in \Sigma^* \text{ with } |H \cdot r| = t\}$. Let $x \in \Sigma^*$ be the corresponding word such that $H \cdot x = S$ and put $u = vx$. Note that $u \in \Sigma^* \setminus J$, $uJ \subseteq vJ \subseteq J_i$ and $Q \cdot u = S$ with $|S| = t$. Since $\mathrm{Syn}(\mathscr{A}_J) = J$, then there is a synchronizing word $w \in J$ with $|w| \leq \|I\| < m$. Let T' be the last set of (S, w), and let w' be the maximal prefix of w such that $S \cdot w' = T'$. Then, there is a letter $a \in \Sigma$ such that $w'a$ is a prefix of w and $|T'a| = 1$. We consider two mutually exclusive cases.

i) Suppose $|T' \cdot b| = 1$ for any $b \in \Sigma$. It is not difficult to check that $\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'a)$. Since $|\Sigma| > 1$ consider a letter $b \in \Sigma$ with $b \neq a$. Since $Q \cdot uw' = T'$ and $|T' \cdot b| = 1$, we also have $\mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b)$. Since $uJ \subseteq J_i$ we have $uw, uw'bw \in J_i$ (being $w, w'bw \in J$). Hence we get

$$\mathcal{T}(Q, uw'a) = \mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b).$$

In particular, we get $\mathcal{T}(Q, uw'a)\langle\Sigma\rangle = \mathcal{T}(Q, uw'b)\langle\Sigma\rangle$, from which it follows $a = b$, a contradiction.

ii) Thus, we can assume that there is a letter $b \in \Sigma$, such that $|T' \cdot b| > 1$. Since $uw, uw'bw \in J_i$ (being $w, w'bw \in J$), we have $\mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw)$. Hence, by Lemma 3.2, we have

$$\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b) \diamond \mathcal{T}(T, w),$$

with $T = T' \cdot b$. Since $|T'| = t$ is minimal and $|T| > 1$ we have $|T| = |T'| = t$, hence $\mathrm{Ind}(\mathcal{T}(Q, uw')) = \mathrm{Ind}(\mathcal{T}(T, w)) = t$. Therefore, by the previous equality and the definition of the operation $\diamond$ we get

$$\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b) + \mathcal{T}(T, w),$$

In particular we have

$$\mathcal{T}(Q, uw)\langle Q\rangle = \mathcal{T}(Q, uw'bw)\langle Q\rangle = \mathcal{T}(Q, uw'b)\langle Q\rangle + \mathcal{T}(T, w)\langle Q\rangle. \qquad (2)$$

Furthermore, T' is the last set of $(Q, uw'a)$ and uw' is the maximal prefix of $uw'a$ such that $T' = Q \cdot uw'$. Since $|T'| = |T|$ we have that T is the last set of $(Q, uw'b)$

and $uw'b$ is the maximal prefix of $uw'b$ with $T = Q \cdot uw'b$. Thus, by the definition of tail we have $\mathcal{T}(Q, uw'a)\langle Q\rangle = \mathcal{T}(Q, uw'b)\langle Q\rangle$. We have already observed that $\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'a)$, hence by (2) we get

$$\mathcal{T}(T, w)\langle Q\rangle = 0. \tag{3}$$

By the minimality of $t = |T|$, we get that $\mathcal{T}(T, w) \in \mathbb{T}_t$. Therefore, if $0 = i_1 < i_2 < \ldots < i_\ell \leq |w|$ is the maximal set of indices such that $T = T \cdot w[0, i_j]$ for all $1 \leq j \leq \ell$, by the definition of tail and (3) we have in particular

$$0 = \mathcal{T}(T, w)(T) = \ell \mod m.$$

Since $\ell \geq 1$ we have that ℓ is a multiple of m. However $\ell \leq |w| < m$, a contradiction.

Therefore, $v \in J$. This concludes the proof that $\{J_i\}_{i\in F}$ is a reset right regular decomposition. Hence $\{J_i^R\}_{i\in F}$ is a reset left regular decomposition and, by Theorem 2.2, I is a strongly connected ideal. □

Since $\|I\| < n$ where n is the number of states of the minimal DFA recognizing I^R, the following corollary provides a better bound with respect to the one presented in [11, Corollary 3].

Corollary 3.5. *Let I be an ideal on a non-unary alphabet, and let n be the state complexity of I^R. There is a strongly connected synchronizing automaton $\mathscr{B}$ with N states and $\mathrm{Syn}(\mathscr{B}) = I$ such that:*

$$N \leq m^{k2^n} \left(\sum_{t=2}^{n} m^{\binom{n}{t}} \right)^{2^n},$$

where $k = |\Sigma|$ and $m = \|I\| + 1$.

Proof. By the proof of Theorem 3.4 I has a reset left regular decomposition $\{I_i\}_{i\in F}$, with $|F| \leq |\operatorname{Hom}(2^Q, \mathbb{T}(m))|$, where

$$\mathbb{T}(m) = \biguplus_{t=2}^{n} \mathbb{Z}_m([2^Q]_t \uplus \Sigma).$$

Hence we get the bound

$$|F| \leq \left(\sum_{t=2}^{n} m^{\binom{n}{t}+k} \right)^{2^n} \leq m^{k2^n} \left(\sum_{t=2}^{n} m^{\binom{n}{t}} \right)^{2^n}.$$

Let $\mathscr{B} = \mathcal{A}(\{I_i\}_{i\in F})$, where $\mathcal{A}(\cdot)$ is the functor in Theorem 2.2. Then $\mathscr{B}$ has $|F|$ states and $\mathrm{Syn}(\mathscr{B}) = I$. □

This last corollary gives a double exponential upper bound for the number of states of the associated strongly connected automaton with respect to the state complexity of the reverse of the ideal. This bound seems far from tight. Therefore, it is quite natural to look for better general constructions than the one given in Theorem 3.4, or to consider

the same task in particular classes of ideals. For instance, Gusev et al. [1] presented an algorithm that, given a principal ideal $I = \Sigma^* w \Sigma^*$ with $|w| = n$ in inputs, returns a strongly connected synchronizing automaton with $n+1$ states. In this case the bound is, thus, linear with respect to the state complexity of I^R, although it is not known whether it is tight. Even more recently, the same authors [2] proved that in the case where I is finitely generated, there is always a strongly connected synchronizing automaton with at most $2^{\|I\|}$ states, and this bound is tight for ideals of the form $\Sigma^{\geq n} = \{u \in \Sigma^* : |u| \geq n\}$ for any $n > 0$.

In the same manner as Maslennikova [4] has introduced the notion of reset complexity of an ideal I (denoted by $\mathrm{rc}(I)$) as the number of states of the smallest synchronizing automaton $\mathscr{A}$ with $\mathrm{Syn}(\mathscr{A}) = I$, we can also give a similar notion in the realm of strongly connected synchronizing automata or reset left regular decomposition. By Theorem 3.4 for any ideal I, the set $\mathcal{R}(I)$ of all its reset left regular decompositions is non-empty. Thus, we can define the *reset regular decomposition complexity of* I as the integer

$$\mathrm{rdc}(I) = \min\{|F| : \{I_i\}_{i \in F} \in \mathcal{R}(I)\}.$$

By the mapping introduced in Theorem 2.2, $\mathrm{rdc}(I)$ is also the number of states of the smallest strongly connected synchronizing automaton having I as the set of reset words. Furthermore, we clearly have $\mathrm{rc}(I) \leq \mathrm{rdc}(I)$. The construction of reset left regular decompositions of small cardinality seems a very hard task, actually this task is as hard as proving Černý's conjecture. The following theorem shows this fact by giving a purely language theoretic restatement of this longstanding conjecture.

Theorem 3.6. *Černý's conjecture holds if and only if for any ideal I we have:*

$$\mathrm{rdc}(I) \geq \sqrt{\|I\|} + 1,$$

where $\|I\| = \min\{|w| : w \in I\}$.

Proof. This a consequence of the fact that Černý's conjecture holds if and only if it holds for strongly connected automata, Proposition 2.6 and Theorem 3.4. □

Note that using $(n^3 - n)/6$ as the upper bound for the shortest reset word of a synchronizing automaton (see [5]) we have the lower bound $\mathrm{rdc}(I) \geq \sqrt[3]{6\|I\|}$. In general, a natural issue would be the study of bounds for $\mathrm{rdc}(I)$ depending on the state complexity of I or I^R. For instance, even lower bounds of the type $\mathrm{rdc}(I) \geq \sqrt{\|I\|}/c$ for some constant $c > 0$ would be a major breakthrough for this conjecture and all the theory of synchronizing automata.

These decompositions seem related to the maps described in Lemma 3.3. Indeed, we now show that for a subclass of the class of ideals we can improve the bound of Corollary 3.5 using the map ψ previously introduced in Lemma 3.3. Firstly, we need some definitions. Given a synchronizing DFA $\mathscr{B} = \langle Q', \Sigma, \delta' \rangle$ with a sink state s, we say that $\mathscr{B}$ has a *funnel* $\overline{q} \in Q' \setminus \{s\}$ if $\delta'(q, a) = s$ for some $a \in \Sigma$ and $q \neq s$, implies $q = \overline{q}$. In other words, every path going to the sink state passes from the state $\overline{q}$. We say that $\mathscr{B}$ is *free from funnels* whenever for any DFA $\mathscr{D}$ that is also a sub-automaton of $\mathscr{B}$, $\mathscr{D}$ has no funnel. For any $S \subseteq Q'$ the *induced sub-automaton* $\mathscr{B}[S] = \langle C, \Sigma, \delta'' \rangle$ of $\mathscr{B}$ is the DFA with set of states $C = \{\delta'(s, u) : s \in S, u \in \Sigma^*\}$, and δ'' is the restriction of δ' on C. We have the following theorem.

Theorem 3.7. *Let $I \subseteq \Sigma^*$ be an ideal such that the minimal DFA $\mathscr{A}_{I^R} = \langle Q, \Sigma, \delta, q_0, \{s\}\rangle$ recognizing I^R is free from funnels. Let $|Q| = n$ be the state complexity of I^R, then I has a reset left regular decomposition $\{I_i\}_{i\in F}$ with $|F| \leq 2^{n^2 2^n}$. In particular, there is a strongly connected synchronizing automaton $\mathscr{B}$ with $I = \mathrm{Syn}(\mathscr{B})$ and with a number of states less or equal to $2^{n^2 2^n}$.*

Proof. Let $J = I^R$, by Lemma 3.3 and following the same line of the proof of Theorem 3.4 we get that the equivalence classes $\{J_i\}_{i\in F}$ of the relation $\sim = (J \times J) \cap \mathrm{Ker}(\psi)$ is a finite collection of right ideals forming a partion of J satisfying condition i) of Definition 2.1. We claim that also condition ii) is satisfied. Indeed, assume, contrary to our claim, that there are $i \in F$ and $u \in \Sigma^* \setminus J$ such that $uJ \subseteq J_i$. Write $H = Q \cdot u$, clearly $|H| \geq 2$. Since $\psi(uJ) = \psi(J_i)$, then

$$\lambda_v(H) = \lambda_{v'}(H), \quad \forall v, v' \in J. \tag{4}$$

Consider a $v \in J$, note that $\{s\} \in \lambda_v(H)$ and let $S \in \lambda_v(H)$ be the last set of (H, v). Let x be a prefix of v such that $H \cdot x = S$. We claim that $|S| = 2$. Suppose to the contrary that $|S| > 2$. Note that, since $S = Q \cdot ux$, and $s \in Q$ is a sink, we have $s \in S$. Therefore, there are at least two distinct elements $q, q' \in S \setminus \{s\}$. We show that the right languages of q and q' with respect to $\mathscr{A}_J$ coincide. For if there would be a word $w \in \Sigma^*$ such that $q \cdot w = s$ but $q' \cdot w \neq s$, then $\lambda_{xwv}(H)$ would contain an element S' with $1 < |S'| < |S|$. However, by (4) and $xwv \in J$, this contradicts the fact that S is the last set of $\lambda_v(H) = \lambda_{xwv}(H)$. Therefore q, q' are equal in the minimal DFA $\mathscr{A}_J$, that is a contradiction. Hence, $S = \{q, s\}$. We claim that $q \in S$ is a funnel for $\mathscr{D} = \mathscr{A}_J[S]$. Indeed, suppose that there is a state $q' \neq s$ of $\mathscr{D}$ and $a \in \Sigma$ such that $q' \cdot a = s$ and $q' \neq q$. By definition of $\mathscr{D}$ there is a word $r \in \Sigma^*$ such that $q \cdot r = q'$. Consider the word $v' = xrav$ (recall $H \cdot x = S$). Clearly $v' \in J$ and $\lambda_{v'}(H)$ contains the set $\{q', s\} \neq S$. However, this contradicts (4). Thus $q' = q$, and so q is a funnel of $\mathscr{D}$, contradicting the statement of the theorem. Hence, $u \in J$ and so $\{J_i^R\}_{i\in F}$ is a reset left regular decomposition for I. By the upper bound stated in equation (1) we obtain $|F| \leq |\mathrm{Hom}(2^Q, \mathcal{L}(Q))| \leq 2^{n^2 2^n}$. The last statement is a consequence of Theorem 2.2. □

For completeness, we now characterize from a language theoretic point of view the ideals whose minimal DFAs are free from funnels. These are exactly the following ideals.

Definition 3.8 (free funnel ideal). *We say that $J \subseteq \Sigma^*$ is a* free funnel ideal *whenever the following property does not occur: there is a $y \in \Sigma^* \setminus J$ such that for any $u, v \in \Sigma^*$ with $yu, yv \notin J$, if there are some (maximal) non-empty subsets $\Sigma', \Sigma'' \subseteq \Sigma$ such that $yu\Sigma' \subseteq J$ and $yv\Sigma'' \subseteq J$, then $\Sigma' = \Sigma'' = \Lambda$, and for all $x \in \Sigma^*$, $yux\Lambda \subseteq J$ if and only if $yvx\Lambda \subseteq J$.*

Proposition 3.9. *$J \subseteq \Sigma^*$ is a free funnel ideal if and only if the minimal DFA recognizing J is free from funnels.*

Proof. Assume J is not a free funnel ideal, and thus, that the conditions described in Definition 3.8 hold. Let $\mathscr{A}_J = \langle Q, \Sigma, \delta, q_0, \{s\}\rangle$ be the minimal DFA recognizing J. We show that the induced sub-automaton $\mathscr{D} = \mathscr{A}_J[p]$, $p = q_0 \cdot y$, is not free from funnels ($y \in \Sigma^* \setminus J$ as in Definition 3.8). For this purpose fix any state $q_1 \neq s$ of $\mathscr{D}$ such that

there is a (maximal) non-empty subset $\Sigma' \subseteq \Sigma$ with $q_1 \cdot x = s$ for all $x \in \Sigma'$. We claim that q_1 is a funnel for $\mathscr{D}$. Indeed, let $q_2 \neq s$ be another state of $\mathscr{D}$ such that there is a (maximal) subset $\Sigma'' \subseteq \Sigma$ with $q_2 \cdot z = s$ for all $z \in \Sigma''$. By definition of $\mathscr{D}$ there are $u, v \in \Sigma^*$ such that $p \cdot u = q_1$, $p \cdot v = q_2$. Therefore, we have $yu\Sigma' \subseteq J$, $yv\Sigma'' \subseteq J$, and Σ', Σ'' are maximal with respect to the conditions $yu\Sigma' \subseteq J$, $yv\Sigma'' \subseteq J$, respectively. Therefore, by Definition 3.8 we get $\Sigma' = \Sigma'' = \Lambda$, and the condition $x \in \Sigma^*$, $yux\Lambda \subseteq J$ if and only if $yvx\Lambda \subseteq J$ ensures that both q_1 and q_2 have the same right languages with respect to $\mathscr{A}_J$. Hence, by the minimality of $\mathscr{A}_J$, we get $q_1 = q_2$, i.e. q_1 is a funnel for $\mathscr{D}$. Conversely, assume $\mathscr{A}_J = \langle Q, \Sigma, \delta, q_0, \{s\}\rangle$ is not free from funnels. Thus, let $\mathscr{D}$ be a subautomaton of $\mathscr{A}_J$ possessing a funnel q. Using the minimality of $\mathscr{A}_J$ it is straightforward to see that if we take any state $p \neq s$ of $\mathscr{D}$, then q is also a funnel for $\mathscr{D}[p]$. Let $y \in \Sigma^*$ with $p = q_0 \cdot y$. We show that J is not a free funnel ideal. Indeed, take any $u, v \in \Sigma^*$ with $yu, yv \notin J$ such that there are some (maximal) subsets $\Sigma', \Sigma'' \subseteq \Sigma$ with $yu\Sigma' \subseteq J$ and $yv\Sigma'' \subseteq J$. Consider the two states of $\mathscr{D}[p]$ $q_1 = q_0 \cdot yu$, $q_2 = q_0 \cdot yv$. Since q is a funnel for $\mathscr{D}[p]$, we get $q_1 = q = q_2$. Hence, by maximality we get $\Sigma' = \Sigma'' = \Lambda$, and since $q_1 = q_2$ we get that for all $x \in \Sigma^*$, $yux\Lambda \subseteq J$ if and only if $yvx\Lambda \subseteq J$. Thus, J is not a free funnel ideal, and this concludes the proof of the proposition. □

4. Conclusion and open problems

The main result of this paper is given by the combination of Theorem 2.2 and Theorem 3.4 which essentially reduce Černý's conjecture to find small reset regular decompositions of regular ideals over non-unary alphabets. Therefore it is clear how fundamental is the issue of understanding these decompositions from a pure language theoretic point of view. In particular, understanding why there are no regular reset decompositions of cardinality less or equal to $c\sqrt{\|I\|}$, $c > 0$, for any non-unary ideal I, would be a major breakthrough since it would give a quadratic upper bound for the shortest reset words.
We list here some open problems originated by the previous results, where I stands for an ideal on a non-unary alphabet.

1. To give a tight upper bound of $\mathrm{rdc}(I)$ with respect to the state complexity of I^R or I.
2. In case where I is finitely generated is it true that $\mathrm{rdc}(I) \geq \|I\| + 1$? The same problem for the case where I is a principal ideal language has been raised by Gusev et al. [1]. This would give a better bound for the shortest synchronizing word for the class of finitely generated synchronizing automata with respect to the bound obtained by Pribavkina et al. [10].
3. The proof of Theorem 3.4 uses the minimal DFA recognizing I^R. Is there a proof using another automaton associated to I? Maybe this could give smaller upper bounds.
4. Recall that $\mathcal{R}(I)$ is the set of all the reset left regular decompositions of I and the order of a decomposition $\mathcal{I} \in \mathcal{R}(I)$ is just the cardinality $|\mathcal{I}|$. We denote by $\mathcal{R}_k(I)$ the set of reset left regular decompositions of I of order $k \geq 1$.
   A quite natural question is whether $\sup\{k \geq 1 : \mathcal{R}_k(I) \neq \emptyset\} = \infty$? In particular, what happens if we consider I to be a finitely generated ideal or even a principal ideals? This would answer to the question whether, given a principal ideal P, there

are arbitrarily big strongly connected synchronizing automata having a P as the set of reset words.

5. By Theorem 2.2, there is a naive way to calculate $\mathcal{R}_k(I)$ by building all the strongly connected synchronizing automata with k states, and checking if their set of reset words coincide with I. Thus, it is natural to ask whether there is a more "efficient" way to perform this task without passing from the construction of all the automata with k states.

Acknowledgements

The authors thank E. Pribavkina for pointing out the unary case alphabet in Corollary 2.3. We also acknowledge M. Berlinkov for the useful comments. This work was partially supported by CMUP (UID/MAT/00144/2013), which is funded by FCT (Portugal) with national (MCTES) and european structural funds through the programs FEDER, under the partnership agreement PT2020. The second author acknowledges the support of the FCT project SFRH/BPD/65428/2009. The authors also thank the anonymous referees for the precious suggestions that have certainly improved the quality of the paper.

References

[1] Gusev, V., Maslennikova, M., Pribavkina, E., 2012. Principal ideal languages and synchronizing automata. in V. Halava, J. Karhumaki, Y. Matiyasevich (eds.) RuFiDimII, TUCS Lecture Notes 17.

[2] Gusev, V., Maslennikova, M., Pribavkina, E., 2013. Finitely generated ideal languages and synchronizing automata. In: J. Karhumäki, L. Zamboni (eds.) Proc. WORDS 2013. Vol. 8079 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg.

[3] Howie, J. M., 1991. Automata and Languages. Clarendon Press.

[4] Maslennikova, M., 2012. Reset complexity of ideal languages. In: M. Bieliková, G. Friedrich, G. Gottlob, S. Katzenbeisser, R. Špánek, G. Turán (eds.) Int. Conf. SOFSEM 2012, Proc. Volume II, Institute of Computer Science Academy of Sci- ences of the Czech Republic. pp. 33–44.

[5] Pin, J. E., 1983. On two combinatorial problems arising from automata theory. Ann Discrete Math. 17, 535–548.

[6] Pribavkina, E., Rodaro, E., 2009. Finitely generated synchronizing automata. In: Language and Automata Theory and Applications. Vol. 5457 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 672–683.

[7] Pribavkina, E., Rodaro, E., 2010. State complexity of prefix, suffix, bifix and infix operators on regular languages. In: Gao, Y., Lu, H., Seki, S., Yu, S. (Eds.), Developments in Language Theory. Vol. 6224 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 376–386.

[8] Pribavkina, E., Rodaro, E., 2011. Recognizing synchronizing automata with finitely many minimal synchronizing words is pspace-complete. In: Models of Computation in Context. Vol. 6735 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 230–238.

[9] Pribavkina, E. V., Rodaro, E., 2011. State complexity of code operators. International Journal of Foundations of Computer Science 22 (07), 1669–1681.

[10] Pribavkina, E. V., Rodaro, E., 2011. Synchronizing automata with finitely many minimal synchronizing words. Information and Computation 209 (3), 568 – 579.
URL http://www.sciencedirect.com/science/article/pii/S0890540110002063

[11] Reis, R., Rodaro, E., 2013. Regular ideal languages and synchronizing automata. In: WORDS 2013. Vol. 8079 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 205–216.

[12] Černý, J., 1964. Poznámka k homogénnym eksperimentom s konečnými automatami [in slovak]. Mat.-Fyz. Čas. Slovensk. Akad. Vied. 14, 208–216.

[13] Volkov, M. V., 2008. Synchronizing automata and the Černý conjecture. In C. Martín-Vide, F. Otto, H. Fernau (eds.), Languages and Automata: Theory and Applications. LATA 2008, Lect. Notes Comp. Sci, Berlin, Springer 5196, 11–27.

[14] Volkov, M. V., 2009. Synchronizing automata preserving a chain of partial orders. Theoretical Computer Science 410, 3513–3519.