Primitive digraphs with large exponents and slowly synchronizing automata*

D. S. Ananichev M. V. Volkov V. V. Gusev†

Abstract

We present several infinite series of synchronizing automata for which the minimum length of reset words is close to the square of the number of states. All these automata are tightly related to primitive digraphs with large exponent.

1 Background and structure of the paper

This paper has arisen from our attempts to find a theoretical explanation for the results of certain computational experiments in synchronization of finite automata. Recall that a (*complete deterministic*) *finite automaton* (DFA) is a triple $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$, where Q and Σ are finite sets called the *state set* and the *input alphabet* respectively, and $\delta : Q \times \Sigma \to Q$ is a totally defined function called the *transition function*. As usual, Σ^* stands for the collection of all finite words over the alphabet Σ, including the empty word 1. The function δ extends to a function $Q \times \Sigma^* \to Q$ (still denoted by δ) in the following natural way: for every $q \in Q$ and $w \in \Sigma^*$, we set $\delta(q, w) = q$ if $w = 1$ and $\delta(q, w) = \delta(\delta(q, v), a)$ if $w = va$ for some $v \in \Sigma^*$ and $a \in \Sigma$. Thus, via δ, every word $w \in \Sigma^*$ acts on the set Q.

A DFA $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ is said to be *synchronizing* if some word $w \in \Sigma^*$ brings all states to one particular state: $\delta(q, w) = \delta(q', w)$ for all $q, q' \in Q$. Any such word w is said to be a *reset word* for the DFA. The minimum length of reset words for $\mathscr{A}$ is called the *reset threshold* of $\mathscr{A}$.

Synchronizing automata serve as transparent and natural models of error-resistant systems in many applied areas (system and protocol testing, information coding, robotics). At the same time, synchronizing automata surprisingly arise in some parts of pure mathematics (symbolic dynamics, theory of substitution systems and others). Basics of the theory of synchronizing automata as well as its diverse connections and applications are discussed, for instance, in the recent surveys [19, 26]. Here we focus on only one aspect of the theory, namely, on the question of how the reset threshold of a DFA depends on the state number.

*A preliminary version of a part of the results of this paper has been published in [3].

†Supported by the Russian Foundation for Basic Research, grant 10-01-00793. V. V. Gusev was also supported by the program 'Erasmus Mundus Action 2 Partnerships—Triple I'.

For brevity, a DFA with n states will be referred to as an *n-automaton*. In 1964 Černý [9] constructed a series of synchronizing n-automata with reset threshold $(n-1)^2$. Soon after that he conjectured that these automata represent the worst possible case with respect to synchronization speed, i.e. that every synchronizing n-automaton can be reset by a word of length $(n-1)^2$. This hypothesis has become known as the *Černý conjecture*. In spite of its simple formulation and many researchers' efforts, the Černý conjecture remains unproved (and undisproved) for more than 45 years. Moreover, no upper bound of magnitude $O(n^2)$ for the reset threshold of a synchronizing n-automaton is known so far[1].

Why is the Černý conjecture so inaccessible? A detailed discussion of this important issue would go far beyond the scope of the present paper but one of the difficulties encountered by the theory of synchronizing automata is worth registering here. We mean the shortage of examples of *extremal* automata, i.e. n-automata having reset threshold $(n-1)^2$. In fact, the series found in [9] still remains the only known infinite series of extremal automata. Besides that, we know only a few isolated examples of such automata, the largest (with respect to the state number) being the 6-automaton discovered by Kari [13] in 2001. (See [26] for a complete list of known extremal automata.) Moreover, even n-automata whose reset threshold is close to $(n-1)^2$ have been very rare in the literature so far — besides the Černý series one can only refer to the series from [4]. With a very restricted number of examples, it has been difficult to verify various guesses and assumptions that arose when researchers were searching for approaches to the Černý conjecture. That is why the history of investigations in this area abounds in "false trails", i.e. auxiliary hypotheses that looked promising at first but were disproved after some time. (Cf. [6] for an analysis of a number of such "false trails".)

How can one find slowly synchronizing automata? Experiments (see, e.g., [20]) demonstrate that with probability very close to 1, a random automaton is reset by a word of length much less than the state number. Therefore it is impossible to encounter by chance an automaton whose reset threshold is close to the square of state number, and one has to reveal such automata via exhaustive search. It was such an exhaustive search experiment that served as a departure point of the present paper.

Our methodology and some results of the experiment are described in Section 2. We have noticed a similarity between the observable behaviour of the number of synchronizing automata with a fixed number of states as a function of their reset threshold and the well-studied behaviour of the number of primitive digraphs with a fixed number of vertices as a function of their exponent. We discuss this similarity in Section 3 after recalling the necessary concepts and facts from the theory of primitive digraphs. The main results of the paper are collected in Section 4. We show that slowly synchronizing automata revealed in our experiment represent some infinite series of such automata and that each

[1] Up today, the best upper bound on the reset threshold of a synchronizing n-automaton is the bound $\frac{n^3-n}{6}$ found by Pin [16] in 1983. A slightly better upper bound $\frac{n(7n^2+6n-16)}{48}$ has been recently published in [25] but the proof of this result contains an unclear place.

of these series is tightly related to a certain known series of primitive digraphs with large exponent. This connection between digraphs and automata allows us to provide transparent and uniform proofs for all statements concerning the reset threshold of series of slowly synchronizing automata—both for the previously known series and the series that have first appeared in the present paper. The proof technique is, to the best of our knowledge, new and appears to be of independent interest. In Section 5 we discuss further prospects of the suggested approach and formulate a few new conjectures.

2 Methodology and some results of the experiment

As mentioned in Section 1, finding automata whose reset threshold is close to the square of state number requires an exhaustive search. Since the quantity of n-automata drastically grows with n, such a search should be designed in a reasonable way. For instance, specifying a 9-automaton with two input letters is equivalent to specifying a pair of function on a 9-element set. There are $9^{18} \approx 1.50 \times 10^{17}$ such pairs, and if one will spend one nanosecond for calculating the reset threshold of each automaton defined this way, the exhaustive search would take around five years. Clearly, if an n-automaton with k input letters is specified by a k-tuple of function on an n-element set, each particular automaton is generated $n!k!$ times. However it is not possible to speed up the search by screening out isomorphic automata—even for $n = 9$ and $k = 2$ neither time nor memory would suffice to check whether the current automaton is isomorphic to one of the previously generated automata.

In order to optimize search, we have employed a string representation of initially-connected automata suggested in [2]. Recall that a DFA $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ is said to be *initially-connected* (from a state q_0), if one can reach any state from q_0 by applying a suitable word: for every $q \in Q$ there exists $w \in \Sigma^*$ such that $q = \delta(q_0, w)$. A DFA which is initially-connected from each of its states is called *strongly connected*. In general, a synchronizing automaton may fail to be strongly connected or initially-connected. However it is well known that one can restrict to strongly connected automata when dealing with issues related to the Černý conjecture. This is a consequence of the following easy result.

Proposition 1 ([27, Proposition 2.1]). *Let $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ be a synchronizing automaton and let S be the set of all states to which $\mathscr{A}$ can be reset. Then $\mathscr{S} = \langle S, \Sigma, \delta|_S \rangle$ is a strongly connected subautomaton in $\mathscr{A}$ and for every function $f : \mathbb{Z}^+ \to \mathbb{N}$ satisfying*

$$f(n) \geq \frac{n(n-1)}{2} \quad \textit{and} \quad f(n) \geq f(n-m+1) + f(m) \textit{ for } n \geq m \geq 1,$$

the fact that the reset threshold of the automaton $\mathscr{S}$ is bounded by $f(|S|)$ implies that the reset threshold for $\mathscr{A}$ does not exceed $f(|Q|)$.

In particular, taking $(n-1)^2$ as $f(n)$, we can conclude that if $\mathscr{A}$ is a counterexample for the Černý conjecture, then so is the strongly connected subautomaton $\mathscr{S}$. Similarly, if $\mathscr{A}$ has reset threshold close to the square of state number, then so does $\mathscr{S}$. Thus, restricting our search to initially-connected automata, we do not risk to overlook a counterexample for the Černý conjecture or any interesting slowly synchronizing automaton.

Now we describe the string representation from [2]. Let a DFA $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ be initially-connected from a state q_0. We fix a linear ordering of the input alphabet Σ and traverse the DFA by breadth-first search starting at q_0 and choosing the outgoing transitions according to the ordering. Let the state q_0 have number 0 and let all other states in Q be numbered in the order of their appearance in breadth-first search. For instance, the states of the DFA in Fig. 1 are numbered as follows: $\begin{array}{c|c|c|c} A & B & C & D \\ \hline 0 & 2 & 1 & 3 \end{array}$ provided that the input letters are ordered $a < b < c$ and the state A is chosen as q_0.

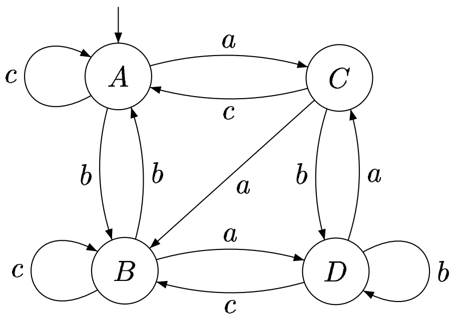

Figure 1: The DFA with the canonical string $[1, 2, 0, 2, 3, 0, 3, 0, 2, 1, 3, 2]$

We assign a string of length $|\Sigma|$ to each state q of $\mathscr{A}$; the i-th position of the string holds the number of the state to which q is sent under the action of the i-th letter. If we concatenate all these strings in the increasing order of the state numbers, we get a string of numbers from the set $\{0, 1, \ldots, |Q| - 1\}$ that has length $|Q||\Sigma|$ and uniquely determines $\mathscr{A}$. It is called the *canonical string* of the DFA $\mathscr{A}$. For instance, the canonical string of the DFA in Fig. 1 is $[1, 2, 0, 2, 3, 0, 3, 0, 2, 1, 3, 2]$.

It is not hard to see (cf. [2, Theorem 5]) that a string $[s_0, \ldots, s_{nk-1}]$ of numbers from the set $\{0, 1, \ldots, n-1\}$ is a canonical string of an initially-connected n-automaton with k input letters if and only if the following two conditions are satisfied:

(R1) for each i such that $s_i > 1$, there is $j < i$ with $s_j = s_i - 1$,

(R2) for each m such that $1 \leq m < n$, there is $j < mk$ with $s_j = m$.

Using this observation, we calculated reset thresholds of initially-connected n-automata with two input letters as follows. We used a 128-core grid of AMD Opteron 2.6 GHz processors. The grid belongs to the Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences; it runs under Linux, has 256 Gb of memory and the peak performance of 665.6 GFLOPS. One node of the gird generated relatively small portions of strings satisfying (R1) and (R2) and sent them to other nodes that worked on their portions of automata in parallel. The management program was written in C with MPI. Standard algorithms (cf. [19, 26]) were implemented to test whether the current automaton is synchronizing and to calculate its reset threshold. Both implementations were written in C. We notice that the synchronization test is very fast as it works on the digraph of pairs of states while the calculation of reset threshold works on the digraph of non-empty sets of states and in the worst case its running time exponentially depends on the size of the automaton under inspection[2]. However in practice the reset threshold was calculated fairly fast since, as mentioned in Section 1, it is small for an overwhelming majority of automata.

Presenting automata via their canonical strings drastically reduces the exhaustive search. (It is easy to see, for instance, that every n-automaton with 2 input letters is generated at most $2n$ times if presented this way.) Nevertheless, the search still remains quite large. For $n = 9$, say, the number of automata to be analyzed was 705 068 085 303. However, thanks to parallelization the computation for $n = 9$ took less than 24 hours.

As the result of the computation, we built an array that, for each possible value of reset threshold, contains the number of automata attaining this value. A part of results that was of outmost importance for us (namely, the part related to slowly synchronizing automata) had been double-checked with the package TESTAS [22].

Table 1 shows a part of the output array for the case $n = 9$. Here automata are counted up to isomorphism.

Table 1: Reset thresholds of synchronizing 9-automata with two input letters

N	64	63	62	61	60	59	58	57	56	55	54	53	52	51
# of automata with reset threshold N	1	0	0	0	0	0	1	2	3	0	0	0	4	4

Clearly, a unique automaton in the column corresponding to $N = 64$ is nothing but the 9-automaton from the Černý series. Then one observes a gap: no 9-automata with two input letters have reset threshold in the range from 59 to 63. This gap has been mentioned by Trahtman [22, 23] who has reported that for $n = 7, 8, 9, 10$ no n-automata with two input letters and reset threshold in the range from $n^2 - 3n + 4$ to $n^2 - 2n + 1$ have been registered in his experiments. The gap is followed by an "island" consisting of three values attained by 6 automata, and the "island" is followed by yet another gap. To the best of our knowledge, the second gap has not been reported in the literature up to now.

The behaviour just described — a unique extremal value followed by a gap which in turn is followed by a small "island" and yet another gap — persists also for automata with a larger number of states. The size of the "island" depends only on the parity of the state number and the sizes of the gaps grow

[2]It is known [15, Theorem 4] that the problem of computing the reset threshold of a given automaton is complete for the functional analogue $\mathsf{FP}^{\mathsf{NP}[\log]}$ of the complexity class $\mathsf{P}^{\mathsf{NP}[\log]}$ consisting of all decision problems solvable by a deterministic polynomial-time Turing machine that has an access to an oracle for an NP-complete problem, with the number of queries being logarithmic in the size of the input.

as linear functions of the state number. A similar behaviour is known for another value investigated in discrete mathematics, namely, for the number of primitive digraphs with a fixed vertex number and a given exponent. In the next section we recall these notions and discuss the similarity in more detail.

§3. Primitive digraphs and their exponents

A *digraph* (directed graph) is a pair $D = \langle V, E\rangle$, where V is a finite set and $E \subseteq V \times V$. The elements of V and E are called *vertices* and respectively *edges*. We notice that this definition allows loops but excludes multiple edges. If $v, v' \in V$ and $e = (v, v') \in E$, then we say that the edge e *starts* at v. We assume that the reader is acquainted with the basic notions of digraph theory such as directed path, directed cycle, isomorphism etc.

If $D = \langle V, E\rangle$ is a digraph, then its *incidence matrix* (referred to simply as *matrix* in the sequel) is a $V \times V$-matrix whose entry in row v and column v' is 1 if $(v, v') \in E$ and is 0 otherwise. Conversely, to each $n \times n$-matrix $P = (p_{ij})$ with non-negative real entries, one can assign a digraph $D(P)$ on $\{1, 2, \dots, n\}$ as the vertex set in which the pair (i, j) is an edge if and only if $p_{ij} > 0$. This correspondence between matrices and digraphs allows one to state in the language of digraphs a number of important notions and results from the classic theory of non-negative matrices (the Perron–Frobenius theory).

A digraph $D = \langle V, E\rangle$ is said to be *strongly connected* if for every pair $(v, v') \in V \times V$, there is a directed path from v to v'. A strongly connected digrpah D is called *primitive* if the greatest common divisor of the lengths of the directed cycles of D is equal to 1. In the literature such digraphs are sometimes called *aperiodic*. Our choice of the name is justified by the fact that a digraph has this property if and only if its matrix is primitive in the sense of the Perron–Frobenius theory, that is the matrix has a positive eigenvalue that is strictly greater than the absolute value of any of its other elgenvalues.

The t-th *power* of a digraph $D = \langle V, E\rangle$ is the digraph D^t with the same vertex set V in which a pair $(v, v') \in V \times V$ is an edge if and only if D has a directed path from v to v' of length precisely t. It is easy to see that if M is the matrix of D, then the digraph D^t is isomorphic to the digraph $D(M^t)$, where M^t is the usual t-th power of the matrix M. It is known (see, e.g., [18, p. 224]) that if D is a primitive digraph, then for some t the power D^t is a complete digraph (with loops), that is, in D^t each pair of vertices constitutes an edge. In matrix terms this means that each entry of the matrix M^t is positive. The least t with this property is called the *exponent* of the digraph D and is denoted by $\gamma(D)$. Exponents of digraphs have been intensively studied over the last 60 years and we refer to [7] for a survey of results accumulated in this area. In this paper we need only a few classic results collected in the following theorem. For brevity, in this theorem (and in the rest of the paper) a digraph with n vertices is called an *n-digraph*.

Theorem 1. (a) (Wielandt's theorem, see [10, 11, 28]) *If D is a primitive n-digraph, then $\gamma(D) \leq (n-1)^2 + 1$.*

(b) [11, Theorem 6 and Corollary 4] *If $n > 2$, then up to isomorphism, there is exactly one primitive n-digraph D with $\gamma(D) = (n-1)^2 + 1$, and exactly one with $\gamma(D) = (n-1)^2$. The matrices of the digraphs are*

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 1 & 0 & \dots & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 1 & 0 & \dots & 0 & 0 \end{pmatrix} \quad \text{respectively.} \qquad (1)$$

(c) [11, Theorem 7] *If $n > 4$ is even, then there is no primitive n-digraph D such that $n^2 - 4n + 6 < \gamma(D) < (n-1)^2$, and, up to isomorphism, there are either 3 or 4 primitive n-digraphs D with $\gamma(D) = n^2 - 4n + 6$, according as n is or is not a multiple of 3.*

(d) [11, Theorem 8] *If $n > 3$ is odd, then there is no primitive n-digraph D such that $n^2 - 3n + 4 < \gamma(D) < (n-1)^2$, and, up to isomorphism, there is exactly one primitive n-digraph D with $\gamma(D) = n^2 - 3n + 4$, exactly one with $\gamma(D) = n^2 - 3n + 3$, and exactly two with $\gamma(D) = n^2 - 3n + 2$. The matrices of these digraphs are:*

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \end{pmatrix}, \qquad (2)$$

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 1 & 0 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \end{pmatrix}. \qquad (3)$$

(e) [11, Theorem 8] *If $n > 3$ is odd, then there is no primitive n-digraph D such that $n^2 - 4n + 6 < \gamma(D) < n^2 - 3n + 2$, and, up to isomorphism, there are either 3 or 4 primitive n-digraphs D with $\gamma(D) = n^2 - 4n + 6$, according as n is or is not a multiple of 3.*

In Table 2, we compare the experimental data from Table 1 with the data that we can extract from Theorem 1. Both digraphs and automata are counted up to isomorphism.

There is an obvious similarity between the second and the third rows of Table 2. An analogous similarity is revealed when one compares the data for other sizes of automata/digraphs. We believe that the observed similarity is more than a mere coincidence and that, in contrary, it reflects some profound and perhaps yet hidden interconnections between primitive digraphs and synchronizing automata. Some of such interconnections have been discovered in the course of investigations related to the so-called Road Coloring Problem. We recall notions involved there.

Table 2: Exponents of primitive 9-digraphs with 9 vertices vs reset thresholds for 2-letter synchronizing automata with 9 states

N	65	64	63	62	61	60	59	58	57	56	55	54	53	52	51
# of digraphs with exponent N	1	1	0	0	0	0	0	1	1	2	0	0	0	0	3
# of automata with reset threshold N	0	1	0	0	0	0	0	1	2	3	0	0	0	4	4

Given a DFA $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$, its *digraph* $D(\mathscr{A})$ has Q as the vertex set and $(q, q') \in Q \times Q$ is an edge of $D(\mathscr{A})$ if and only if $q' = \delta(q, a)$ for some $a \in \Sigma$. It is easy to see that a digraph D is isomorphic to the digraph of some DFA if and only if each vertex of D has at least one outgoing edge. In the sequel, we always consider only digraphs satisfying this property. Every DFA $\mathscr{A}$ such that $D \cong D(\mathscr{A})$ is called a *coloring* of D. Thus, every coloring of D is defined by assigning non-empty sets of labels (colors) from some alphabet Σ to edges of D such that the label sets assigned to the outgoing edges of each vertex form a partition of Σ. Fig. 2 shows a digraph and two of its colorings by $\Sigma = \{a, b\}$.

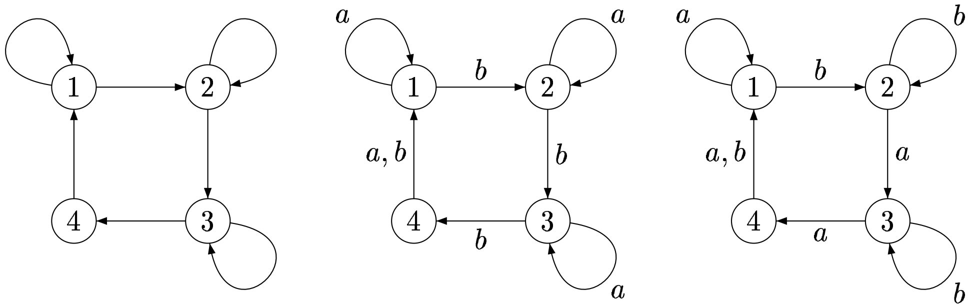

Figure 2: A digraph and two of its colorings

In 1977 Adler, Goodwyn, and Weiss [1] observed that the digraph of every strongly connected synchronizing automaton is primitive and conjectured that every primitive digraph has a synchronizing coloring. This conjecture, known as the Road Coloring Conjecture, has been recently proved by Trahtman [24]. It is easy to relate the reset threshold of a strongly connected synchronizing automaton with the exponent of its digraph.

Proposition 2. *Let $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ be a strongly connected synchronizing n-automaton with reset threshold t. Then*

$$\gamma(D(\mathscr{A})) \leq t + n - 1. \tag{4}$$

Proof. Let $w \in \Sigma^*$ be a reset word of length t and let p be such that $\delta(q, w) = p$ for each $q \in Q$. Now we take an arbitrary pair $(q', q'') \in Q \times Q$ and construct a directed path from q' to q'' of length precisely $t + n - 1$ in the digraph $D(\mathscr{A})$. Since the digraph $D(\mathscr{A})$ is strongly connected, it has a directed path from p to q''. Let ℓ be the minimum length of such a path. Since the path of minimum length visits no vertex twice, $\ell \leq n - 1$. Now we consider an arbitrary directed path of length $n-1-\ell$ starting with q'. From the endpoint of this path, we walk along the directed path of length t labelled by the word w in the automaton $\mathscr{A}$. Since w is a reset word, this path necessarily ends with p. It remains to walk along the directed path of length ℓ from p to q'' in order to obtain a directed path from q' to q'' of length $(n-1-\ell)+t+\ell = t+n-1$. □

Thus, colorings of primitive digraphs with large exponents yield automata with large reset thresholds. This observation is not yet sufficient to completely explain the similarity between the rows of Table 2 since many automata corresponding to non-zero entries in the third row cannot be obtained as colorings of primitive digraphs with large exponents. In Section 4 we present yet another way to produce slowly synchronizing automata from primitive digraphs.

§4. Series of slowly synchronizing automata

4.1. Overview. First of all, we would like to discuss what kind of automata we are interested in, in other words, what is the precise meaning of the expressions like "slowly synchronizing automata" or "automata whose reset threshold is close to the square of the state number". As mentioned in Section 1, the reset threshold of an n-automaton can reach $(n-1)^2$. The corresponding example (discovered by Černý [9]) is the automaton $\mathscr{C}_n = \langle\{1,2,\dots,n\},\{a,b\},\delta\rangle$ where the letters a and b act as follows:

$$\delta(i,a) = \begin{cases} i & \text{if } i < n, \\ 1 & \text{if } i = n; \end{cases} \quad \delta(i,b) = \begin{cases} i+1 & \text{if } i < n, \\ 1 & \text{if } i = n. \end{cases}$$

The automaton $\mathscr{C}_n$ is shown in Fig. 5(left). It is easy to see that if one adds a new state q_0 to the automaton $\mathscr{C}_{n-1}$ and then defines the action of the letters a and b at this added state in all possible ways such that at least one of the letters does not fix q_0, then one gets n^2-1 non-isomorphic initially connected n-automata with reset thresholds between $(n-2)^2$ and $(n-2)^2+1$. In a similar way one can "multiply" other $(n-1)$-automata whose reset threshold is close to $(n-2)^2$, thus obtaining families of n-automata.

Since we want to avoid considering such more or less trivial modifications, we focus on n-automata whose reset thresholds are between $(n-2)^2+2$ and $(n-1)^2$. (This explains, in particular, our choice of the range of reset thresholds in Tables 1 and 2.)

Our experiments show that the number of synchronizing automata in this range is not large and that their distribution with respect to the possible values of reset threshold clearly reveals the following pattern: an isolated extreme value — a gap — a small "island" — another gap — a "continent", see Table 1 and the discussion at the end of Section 2. We shall show that departing from primitive digraph with large exponent presented in Theorem 1, the following series of automata over 2-letter alphabet can be constructed:

- the series $\mathscr{C}_n$ that corresponds to the observed extreme value;
- the series $\mathscr{W}_n$, $\mathscr{D}'_n$, $\mathscr{D}''_n$, $\mathscr{E}_n$, and for odd n also the series $\mathscr{B}_n$ and $\mathscr{F}_n$ that correspond to all observed "island" values;

- the series $\mathscr{G}_n$ (for odd n) and $\mathscr{H}_n$ that correspond to the maximum observed "continental" values.

Fig. 3 demonstrates which series of digraphs give rise to the series of automata just listed. (For clarity, digraph series in Fig. 3 are presented by their "icons" rather than matrices from Theorem 1.) A solid arrow from an icon to a letter denoting a series of automata means that automata in the series are colorings of the corresponding digraphs; a dotted arrow indicates another way of producing automata from digraphs. Dotted frames embrace series of automata with the same value of reset threshold.

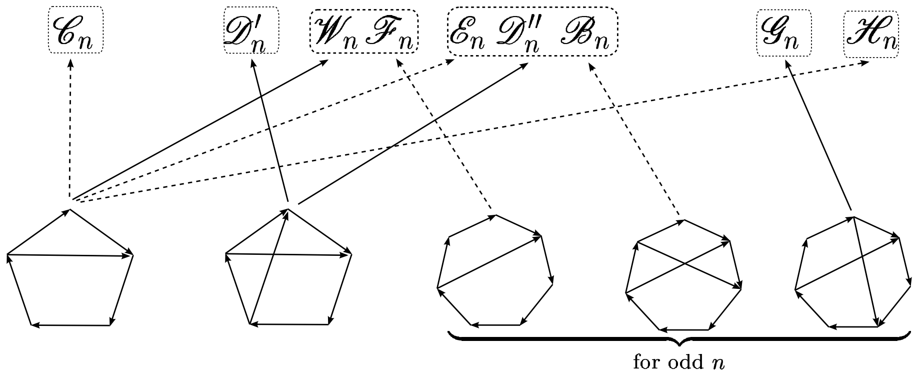

Figure 3: Connections between series of primitive digraphs with large exponents and series of slowly synchronizing automata

Thus, we are going to establish a number of results about automata series listed in Fig. 3. (Two of these results—namely, the two concerning the series $\mathscr{C}_n$ and $\mathscr{B}_n$—are known in the literature but our proofs are essentially novel.) The results are divided into 5 groups in the accordance with the "origin" of the series, that is, according to the type of digraphs that give rise to automata in the series.

4.2. Automata related to digraphs of the series W_n. The digraph W_n is the n-digraph with largest exponent that corresponds to the first matrix in (1). If we denote the vertices of W_n by $1, 2, \ldots, n$, then its edges are $(n, 1)$, $(n, 2)$ and $(i, i+1)$ for $i = 1, \ldots, n-1$. It is easy to see that up to isomorphism and renaming of letters, there is a unique coloring of the digraph W_n with two letters. We denote the resulting automaton by $\mathscr{W}_n$. The digraph W_n and the automaton $\mathscr{W}_n$ are shown in Fig. 4.

Theorem 2. *The automaton $\mathscr{W}_n$ is synchronizing and its reset threshold is equal to $n^2 - 3n + 3$.*

Proof. It is easy to see that the word $(ab^{n-2})^{n-2}a$ resets the automaton $\mathscr{W}_n$. The length of this word is equal to $(n-1)(n-2)+1 = n^2 - 3n + 3$. On the other hand, Theorem 1(b) and Proposition 2 imply that the reset threshold of $\mathscr{W}_n$ cannot be less than $((n-1)^2+1) - (n-1) = n^2 - 3n + 3$. $\square$

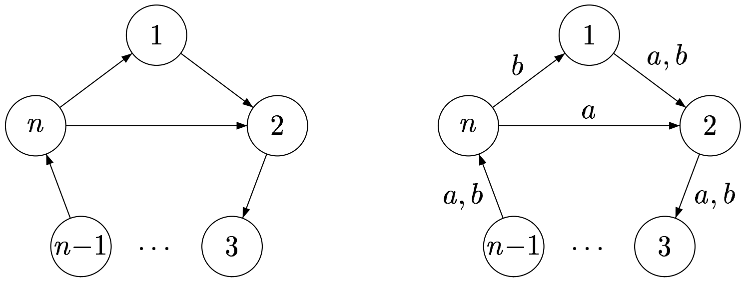

Figure 4: The digraph W_n and the automaton $\mathscr{W}_n$

The series $\mathscr{W}_n$ was discovered by the first author in 2008. His original proof of Theorem 2 relied on a game-theoretic technique from [4] and was rather difficult.

Now we show that also the automata in the series $\mathscr{C}_n$ are tightly related to the digraphs in the series W_n though the relation is less obvious. We notice that even though the automata $\mathscr{C}_n$ have been known for about 50 years and have been rediscovered several times, to the best of our knowledge, their relationship to the digraphs in the series W_n has not been observed previously.

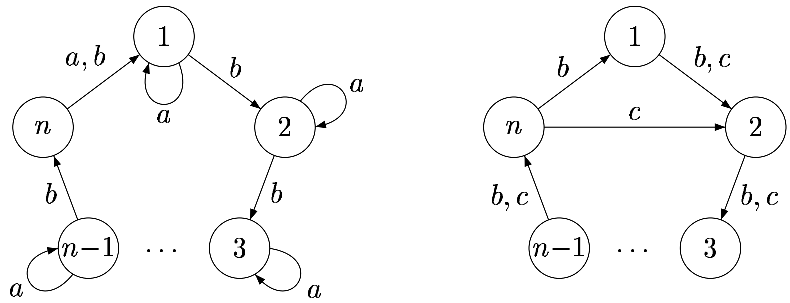

Figure 5: The automaton $\mathscr{C}_n$ and the automaton defined by the action of the words b and $c = ab$

We give a new simple proof of Černý's classical result.

Theorem 3 ([9, Lemma 1]). *The automaton $\mathscr{C}_n$ is synchronizing and its reset threshold is equal to $(n-1)^2$.*

Proof. It is easy to see that the word $(ab^{n-1})^{n-2}a$ resets the automaton $\mathscr{C}_n$. The length of this word is equal to $n(n-2)+1 = (n-1)^2$.

Now we invoke the following observation that will be used in some other proofs as well.

Proposition 3. *Let $\mathscr{A} = \langle Q, \{a, b\}, \delta\rangle$ be a synchronizing n-automaton with reset threshold t in which the letter a fixes all but one states and the letter b acts as a permutation of the set Q. Consider the automaton $\mathscr{B} = \langle Q, \{b, c\}, \zeta\rangle$ in which $\zeta(q, b) = \delta(q, b)$ and $\zeta(q, c) = \delta(q, ab)$ for all $q \in Q$. Then the automaton $\mathscr{B}$ is synchronizing and its reset threshold does not exceed $t - n + 2$.*

Proof. Let w be a reset word of the automaton $\mathscr{A}$ of length t. Since the letter b acts as a permutation of the set Q, the word w cannot end with b—otherwise we could obtain a shorter reset word by removing the last letter of w. Thus, $w = w'a$ for some word $w' \in \{a,b\}^*$. Let $q_1 \in Q$ be a unique state which is not fixed by the letter a and let $q_2 = \delta(q_1, a)$. The minimality of the length of the word w implies that the image of the set Q under the action of the word w' is equal to $\{q_1, q_2\}$.

Since the word a^2 acts on Q in the same way as the letter a, this word cannot occur in w as a factor—otherwise we obtain a shorter reset word by substituting the occurrence of a^2 in w by a. Therefore each occurrence of a in w, except the last one, is followed by an occurrence of the letter b. Hence the word w' can be written as a word in the generators b and ab. Now we substitute each occurrence of the factor ab in w' by an occurrence of the letter c so that we rewrite w' into a word v over the alphabet $\{b,c\}$. Since the words w' and v act on the set Q in the same way, vc is a reset word for the automaton $\mathscr{B}$. Thus, $\mathscr{B}$ is a synchronizing automaton; let s be its reset threshold.

Since b only permutes the states and each application of c can send to one state only one pair of states, the word vc that sends all states to a single state must contain at least $n-1$ occurrences of c. The length of v as a word over $\{b,c\}$ is not less than $s-1$ and v contains at least $n-2$ occurrences of c. Each occurrence of c in v corresponds to an occurrence of the factor ab in w', whence we conclude that the word w' has length at least $(s-1)+(n-2)$. Since the length of the word $w = w'a$ is equal to t, we obtain $t-1 \geq (s-1)+(n-2)$, whence $s \leq t-n+2$. □

Observe that in the sequel we will often use modifications of a given automaton $\mathscr{A}$ in the flavor of Proposition 3. In such modifications, we consider a new automaton on the same state set but with input letters c_1 and c_2 whose actions are defined by the actions of some words w_1 and w_2 respectively in the automaton $\mathscr{A}$. Slightly abusing terminology, we refer to the automaton arising this way as the automaton *defined by the actions of the words* $c_1 = w_1$ *and* $c_2 = w_2$.

We return to the proof of Theorem 3. It is easy to see that for the automaton $\mathscr{C}_n$, the automaton defined by the actions of the words b and $c = ab$ is isomorphic to the automaton $\mathscr{W}_n$, see Fig. 5(right). By Theorem 2, the reset threshold of $\mathscr{W}_n$ is n^2-3n+3. Applying Proposition 3, we conclude that the reset threshold of $\mathscr{C}_n$ cannot be less than $(n^2-3n+3)+(n-2) = n^2-2n+1 = (n-1)^2$. □

The next series in the family of automata related to the digraph W_n consists of the automata $\mathscr{E}_n = \langle\{1,2,\ldots,n\},\{a,b\},\delta\rangle$, where the letter a and b act as follows:

$$\delta(i,a) = \begin{cases} 2 & \text{if } i = 1, \\ 3 & \text{if } i = 2, \\ i & \text{if } i > 2; \end{cases} \qquad \delta(i,b) = \begin{cases} i+1 & \text{if } i < n, \\ 1 & \text{if } i = n. \end{cases}$$

The automaton $\mathscr{E}_n$ is shown in Fig. 6 (left).

Theorem 4. *The automaton* $\mathscr{E}_n$ *is synchronizing, and its reset threshold is equal to* n^2-3n+2.

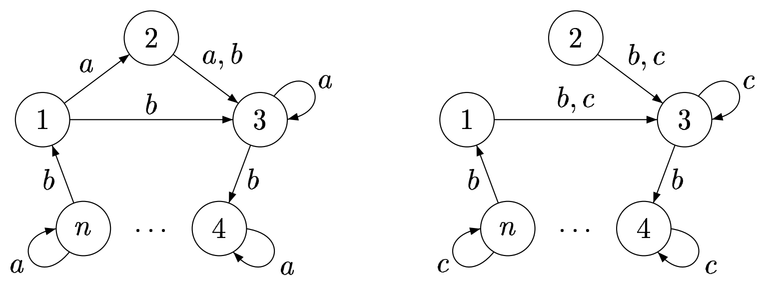

Figure 6: The automaton $\mathscr{E}_n$ and the automaton defined by the actions of the words b and $c = aa$

Proof. It is easy to verify that $(a^2b^{n-2})^{n-3}a^2$ is a reset word for the automaton $\mathscr{E}_n$. The length of this word is equal to $n(n-3)+2 = n^2-3n+2$.

Now let w be a reset word of minimum length for $\mathscr{E}_n$. We notice that in $\mathscr{E}_n$ the words bab and b^2 act in the same way and so do the words a^3 and a^2. Therefore neither bab nor a^3 can occur in the word w as a factor. Besides that, w cannot start with ab. Indeed, the image of the set of all states under the action of the word ab is equal to $\{1,3,\dots,n\}$ and thus coincides with the image of the letter b. Therefore would the word w start with ab, we could obtain a shorter reset word by substituting ab by b. Finally, w cannot end with ba. Indeed, if $w = w'a$, then the minimality of w implies that the image of the set of all states under the action of the word w' is equal to $\{2,3\}$. This set, however, is not contained in the image of the letter b, whence w' cannot end with b. Thus, every occurrence of the letter a in the word w happens within the factor a^2 and no occurrences of these factors in w can overlap.

Let $c = a^2$, then the word w can be rewritten into a word v over the alphabet $\{b,c\}$. The actions of b and c on the set $\{1,2,\dots,n\}$ define an automaton shown in Fig. 6 (right). Since the words w and v act on $\{1,2,\dots,n\}$ in the same way, v is a reset word for this automaton, and hence, for its subautomaton on the set $\{1,3,\dots,n\}$. It is easy to see that the latter subautomaton is isomorphic to the automaton $\mathscr{C}_{n-1}$. By Theorem 3 the length of v as a word over $\{b,c\}$ is at least $(n-2)^2$ and v contains at least $n-2$ occurrences of c. Since every occurrence of c in v corresponds to an occurrence of the factor a^2 in w, we conclude that the length of word w is not less than $(n-2)^2+(n-2) = n^2-3n+2$. $\square$

The proof of Theorem 4 shows that the automaton $\mathscr{E}_n$ arises from one of the "trivial" modifications of the automaton $\mathscr{C}_{n-1}$ that we discussed in Subsection 4.1. The last series of slowly synchronizing automata from the automata family related to the digraph W_n arises from a similar modification of the automaton $\mathscr{W}_{n-1}$. The series consists of the automata $\mathscr{H}_n = \langle\{1,2,\dots,n\},\{a,b\},\delta\rangle$, where the letter a and b act as follows:

$$\delta(i,a) = \begin{cases} n & \text{if } i = 1, \\ i & \text{if } 1 < i < n, \\ 1 & \text{if } i = n; \end{cases} \quad \delta(i,b) = \begin{cases} i+1 & \text{if } i < n-1, \\ 1 & \text{if } i = n-1, \\ 3 & \text{if } i = n. \end{cases}$$

The automaton $\mathscr{H}_n$ is shown in Fig. 7 (left).

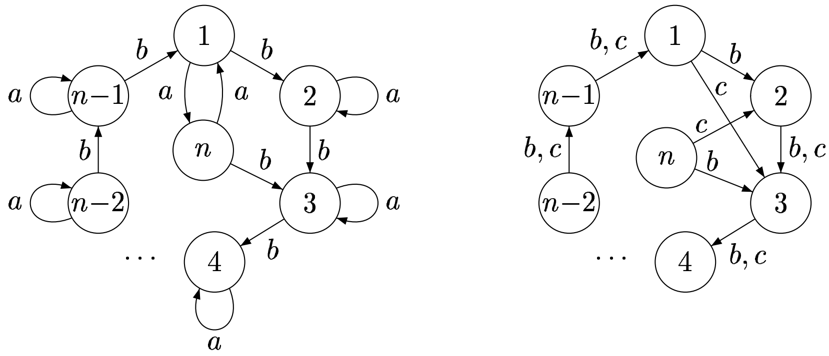

Figure 7: The automaton $\mathscr{H}_n$ and the automaton defined by the actions of the words b and $c = ab$

Theorem 5. *The automaton $\mathscr{H}_n$ is synchronizing and its reset threshold is equal to $n^2 - 4n + 6$.*

Proof. It is easy to check that the words $b(ab^{n-2})^{n-3}ab$ resets the automaton $\mathscr{H}_n$. The length of this word is equal to $1 + (n-1)(n-3) + 2 = n^2 - 4n + 6$.

Now let w be a reset word of minimum length for $\mathscr{H}_n$. Since the word a^2 acts in $\mathscr{H}_n$ as the identity transformation, it cannot occur as a factor in w Besides that, the word w neither starts nor ends with the letter a because this letter acts as a permutation of the state set of $\mathscr{H}_n$.

Let $c = ab$, then the word w can be rewritten into a word v over the alphabet $\{b, c\}$. The actions of b and c on the set $\{1, 2, \ldots, n\}$ define an automaton shown in Fig. 7 (right). Since the words w and v act on $\{1, 2, \ldots, n\}$ in the same way, v is a reset word for this automaton. We have noticed that the word w starts with the letter b, hence so does the word v. If we write $v = bv'$ for some $v' \in \{b, c\}^*$, then it is easy to see that v' is a reset word for the subautomaton on the set $\{1, 2, \ldots, n-1\}$. Since this subautomaton is isomorphic to the automaton $\mathscr{W}_{n-1}$, Theorem 2 implies that the length of v' as a word over $\{b, c\}$ is at least $(n-1)^2 - 3(n-1) + 3$. Besides that, v' contains at least $n-2$ occurrences of c because b only permutes the states and each application of c can send only one pair of states to a single state. Since every occurrence of c in v' corresponds to an occurrence of the factor ab in w, we can conclude that the length of the word w is not less than $1 + ((n-1)^2 - 3(n-1) + 3) + (n-2) = n^2 - 4n + 6$. □

4.3. Automata related to digraphs of the series D_n. The digraph D_n is the n-digraph with exponent $(n-1)^2$ that corresponds to the second matrix in (1). It can be obtained from the digraph W_n by adding the edge $(n-1, 1)$. It is easy to see that up to isomorphism and renaming of letters, there exist exactly two colorings of the digraph D_n with two letters. Fig. 8 shows the digraph D_n and two its colorings, the automata $\mathscr{D}'_n$ and $\mathscr{D}''_n$.

Theorem 6. *The automata $\mathscr{D}'_n$ and $\mathscr{D}''_n$ are synchronizing and its reset thresholds are equal to $n^2 - 3n + 4$ and $n^2 - 3n + 2$ respectively.*

Proof. It is not hard to verify that the word $(ab^{n-2})^{n-2}ba$ is a reset word for the automaton $\mathscr{D}'_n$ and the word $(ba^{n-1})^{n-3}ba$ is a reset word for the automaton

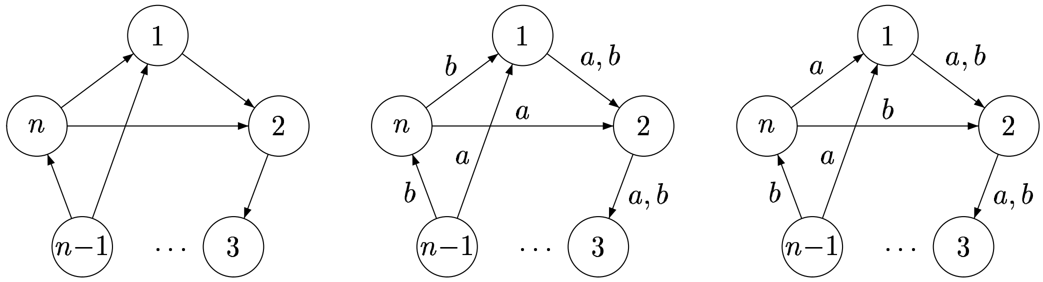

Figure 8: The digraph D_n and its colorings $\mathscr{D}'_n$ and $\mathscr{D}''_n$

$\mathscr{D}''_n$. The lengths of these words are equal $(n-1)(n-2)+2 = n^2-3n+4$ and $n(n-3)+2 = n^2-3n+2$ respectively.

Theorem 1(b) and Proposition 2 imply that the reset threshold for the colorings of the digraph D_n cannot be less than $(n-1)^2-(n-1) = n^2-3n+2$. This proves our theorem for the automaton $\mathscr{D}''_n$.

Now consider the automaton $\mathscr{D}'_n$. Here we shall make use of the following elementary result.

Lemma 1 ([17, Theorem 2.1.1]). *If k, ℓ are relatively prime positive integers, then $k\ell - k - \ell$ is the largest integer that is not expressible as a non-negative integer combination of k and ℓ.*

Let w be a reset word of minimum length for the automaton $\mathscr{D}'_n$. Since 2 is a unique state in $\mathscr{D}'_n$ that is a common end of two different edges with the same label, the minimality of w implies that w sends all states of the automaton to 2. Suppose that the length of w is equal to n^2-3n+2. Then the digraph D_n has a directed path of this length from 1 to 2. There is a unique edge starting at 1, namely, $(1,2)$, hence the path consists of this edge followed by a directed cycle of length n^2-3n+1. The digraph D_n has exactly three simple directed cycles: one of length n and two of length $n-1$. Every directed cycle consists of simple directed cycles whence the number n^2-3n+1 (as the length of a directed cycle in D_n) must be a non-negative integer combination of the numbers n and $n-1$ (the lengths of simple directed cycles). However this is impossible by Lemma 1 since $n^2-3n+1 = n(n-1)-n-(n-1)$.

Now suppose that the length of w is equal to n^2-3n+3. Then the digraph D_n has a directed path of this length from $n-1$ to 2. Since b acts as a permutation of the state set of the automaton $\mathscr{D}'_n$, the word w starts with the letter a. The state $n-1$ under the action of a goes to the state 1. Therefore D_n has also a directed path of length n^2-3n+2 from 1 to 2 but in the previous paragraph we have shown that this is impossible. Thus, the length of w cannot be less than n^2-3n+4. □

The series $\mathscr{D}'_n$ is of interest because for $n > 6$, the automata of this series have the largest reset threshold among all known automata except the ones from the Černý series $\mathscr{C}_n$ as well as the largest reset threshold among all known automata without loops. The series $\mathscr{D}''_n$ also possess an extremal property: the automata from this series have the largest reset threshold among all known automata in which no letter acts as a permutation of the state set.

There is one further series of slowly synchronizing automata related to the digraphs D_n; it consists of n-automata with reset threshold $n^2 - 4n + 6$. We do not present it here since one series with the same parameters has already been described above, see Theorem 5.

4.4. Automata related to digraphs of the series V_n. The digraph V_n is the n-digraph corresponding to the first matrix in (2). If we denote the vertices of V_n by $1, 2, \ldots, n$, then its edges are $(n, 1)$, $(n, 3)$ and $(i, i+1)$ for $i = 1, \ldots, n-1$. The digraph V_n is primitive only when n is odd, and in this case its exponent is equal to $n^2 - 3n + 4$. The digraphs of the series V_n give rise to the family of automata $\mathscr{F}_n = \langle\{1, 2, \ldots, n\}, \{a, b\}, \delta\rangle$ in which the letters a and b act as follows:

$$\delta(i, a) = \begin{cases} i & \text{if } i < n, \\ 2 & \text{if } i = n; \end{cases} \qquad \delta(i, b) = \begin{cases} i+1 & \text{if } i < n, \\ 1 & \text{if } i = n. \end{cases}$$

The automaton $\mathscr{F}_n$ is shown in Fig. 9 (left).

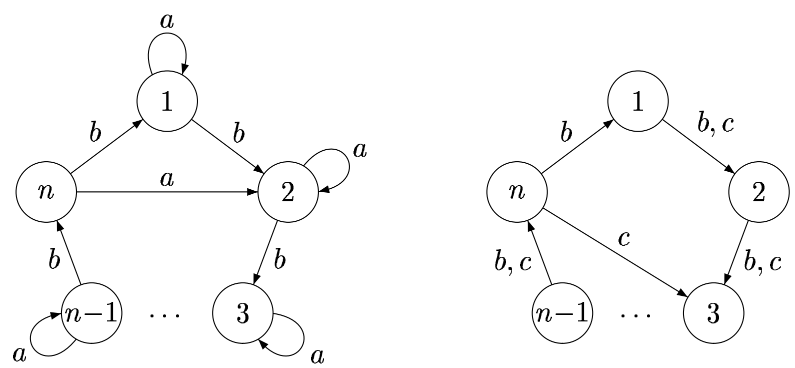

Figure 9: The automaton $\mathscr{F}_n$ and the automaton, defined by the actions of the words b and $c = ab$

Theorem 7. *For odd $n > 3$, the automaton $\mathscr{F}_n$ is synchronizing and its reset threshold is equal to $n^2 - 3n + 3$.*

Proof. It can be easily verified that for each odd $n > 3$ the word $(ab^{n-2})^{n-2}a$ is a reset word for $\mathscr{F}_n$. The length of this word is $(n-1)(n-2)+1 = n^2 - 3n + 3$.

Clearly, the automaton $\mathscr{F}_n$ satisfies the condition of Proposition 3. The action of the words b and $c = ab$ on the set $\{1, 2, \ldots, n\}$ defines an automaton shown in Fig. 9 (right); we denote this automaton by $\mathscr{V}$. It is easy to see that the automaton $\mathscr{V}$ is isomorphic to a coloring of the digraph V_n. Theorem 1(d) and Proposition 2 imply that the reset threshold for colorings of the digraph V_n cannot be less than $(n^2 - 3n + 4) - (n - 1) = n^2 - 4n + 5$. Applying Proposition 3, we conclude that reset threshold for $\mathscr{F}_n$ cannot be less than $(n^2 - 4n + 5) + (n - 2) = n^2 - 3n + 3$. $\square$

4.5. Automata related to digraphs of the series R_n. The digraph R_n is the n-digraph corresponding to the second matrix in (2). One obtains it from the digraph V_n by adding the edge $(n-1, 2)$. The digraph R_n is primitive only when n is odd, and in this case its exponent is equal to n^2-3n+3. The digraphs of the series R_n give rise to the family of automata $\mathscr{B}_n = \langle\{1, 2, \ldots, n\}, \{a, b\}, \delta\rangle$ in which the letters a and b act as follows:

$$\delta(i, a) = \begin{cases} i & \text{if } i < n-1, \\ 1 & \text{if } i = n-1, \\ 2 & \text{if } i = n; \end{cases} \quad \delta(i, b) = \begin{cases} i+1 & \text{if } i < n, \\ 1 & \text{if } i = n. \end{cases}$$

The automaton $\mathscr{B}_n$ is shown in Fig. 10 (left).

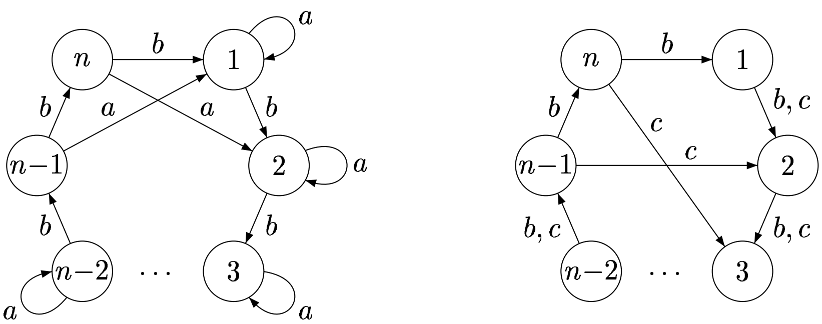

Figure 10: The automaton $\mathscr{B}_n$ and the automaton, defined by the action of the words b and $c = ab$

The series $\mathscr{B}_n$ (for odd $n > 3$) was published in [4] and up to recently, it remained the only infinite series of slowly synchronizing automata with two input letters in the literature besides the Černý series. The fact that the series $\mathscr{B}_n$ is related to the digraphs form the series R_n has not been reported earlier.

The next statement has been the main result of [4] where it has been proved by a game-theoretic method. Here we present a completely elementary proof similar to the proofs of Theorems 3 and 7.

Theorem 8 ([4, Theorem 1.1]). *If $n > 3$ is odd, the automaton $\mathscr{B}_n$ is synchronizing and its reset threshold is equal to $n^2 - 3n + 2$.*

Proof. For each odd $n > 3$ the word $(ab^{n-2})^{\frac{n-3}{2}} ab^{n-3} (ab^{n-2})^{\frac{n-3}{2}} a$ is easily seen to be a reset word for the automaton $\mathscr{B}_n$. The length of this word is equal to $(n-1)\frac{n-3}{2} + n - 2 + (n-1)\frac{n-3}{2} + 1 = n^2 - 3n + 2$.

Let w be a reset word of minimum length for the automaton $\mathscr{B}_n$ and let t be the length of w. Since the letter b acts as a permutation of the state set, the word w neither starts nor ends with b. In particular, $w = w'a$ for some word $w' \in \{a, b\}^*$. The minimality of the word w implies that the image of the state set under the action of the word w' is equal to either $\{1, n-1\}$ or $\{2, n\}$.

Since the word a^2 acts in $\mathscr{B}_n$ in the same way as the letter a, this word cannot occur in w as a factor. Further, the word b^n acts in $\mathscr{B}_n$ as the identity transformation and hence it also cannot occur as a factor in a reset word of minimum length. Thus, we conclude that $w = ab^{k_1}ab^{k_2}a\cdots ab^{k_m}a$, where $1 \le k_1, k_2, \ldots, k_m \le n-1$.

Let $c = ab$. Then the word w' and the word $v = cb^{k_1-1}cb^{k_2-1}c\cdots cb^{k_m-1}$ act on the set $\{1, 2, \ldots, n\}$ in the same way. Therefore the word vc is a reset word for the automaton $\mathscr{R}$ defined by the actions of the words b and $c = ab$ and shown in Fig. 10 (right). It is clear that the automaton $\mathscr{R}$ is isomorphic to a coloring of the digraph R_n. Theorem 1(d) and Proposition 2 imply that the reset threshold for colorings of the digraph R_n cannot be less than $(n^2-3n+3)-(n-1) = n^2-4n+4$ whence the length of v as a word over $\{b, c\}$, that is, $\sum_{i=1}^{m} k_i$, is not less than n^2-4n+3. Since $k_i \le n-1$ for all $i = 1, \ldots, m$, we have

$$m(n-1) \ge \sum_{i=1}^{m} k_i \ge n^2 - 4n + 3 = (n-3)(n-1), \tag{5}$$

whence $m \ge n-3$. The equality $m = n-3$ is only possible when all inequalities in (5) become equalities, that is when $k_i = n-1$ for all $i = 1, \ldots, m$. In this case $vc = (cb^{n-2})^{n-3}c$, but this word is not a reset word for $\mathscr{R}$ since, as it easy to see, this word permutes the states 2 and 3. Hence, $m \ge n-2$.

Since every occurrence of c in v corresponds to an occurrence of the factor ab in w', we conclude that the length of w' is at least $(n^2-4n+3)+(n-2) = n^2-3n+1$, whence the length of w is at least n^2-3n+2. $\square$

4.6. Automata related to digraphs of the series G_n. The digraph G_n is the n-digraph corresponding to the second matrix in (3). One obtains it from the digraph V_n by adding the edge $(n-2, 1)$. The digraph G_n is primitive only when n is odd, and in this case its exponent is equal to n^2-3n+2. Fig. 11 shows the automaton $\mathscr{G}_n$ which is one of possible colorings of the digraph G_n. This series is interesting for us because for odd n, its automata attain the maximal observed "continental" value of reset threshold.

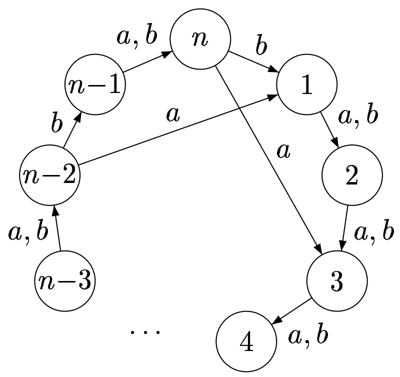

Figure 11: The automaton $\mathscr{G}_n$

Theorem 9. *For odd $n > 3$, the automaton $\mathscr{G}_n$ is synchronizing and its reset threshold is equal to n^2-4n+7.*

Proof. It is easy to see that for odd $n > 3$, the word $a^2(baba^{n-3})^{n-4}baba^2$ is a reset word for the automaton $\mathscr{G}_n$. The length of this word is equal to $2 + n(n-4) + 5 = n^2 - 4n + 7$.

The following arguments are quite similar to ones from the proof of Theorem 6. Theorem 1(d) and Proposition 2 imply that the reset threshold for colorings of the digraph G_n cannot be the less than $(n^2-3n+2)-(n-1) = n^2-4n+3$.

Now let w be a reset word of minimum length for the automaton $\mathscr{G}_n$. Since 3 is a unique state in $\mathscr{G}_n$ that is a common end of two different edges with the same label, the minimality of w implies that w sends all states of the automaton to 3. Suppose that the length of w is equal to $n^2 - 4n + 3$. Then the digraph G_n has a directed path of this length from 2 to 3. There is a unique edge starting at 2, namely, $(2, 3)$, hence the path consists of this edge followed by a directed cycle of length $n^2 - 4n + 2$. The digraph G_n has exactly three simple directed cycles: one of length n and two of length $n - 2$. Observe that n and $n - 2$ are relatively prime since n is odd. Every directed cycle consists of simple directed cycles whence the number $n^2 - 4n + 2$ (as the length of a directed cycle in G_n) must be a non-negative integer combination of the numbers n and $n - 2$ (the lengths of simple directed cycles). However this is impossible by Lemma 1 since $n^2 - 4n + 2 = n(n-2) - n - (n-2)$.

Suppose that the length of w is equal to $n^2 - 4n + 4$. Then the digraph G_n has a directed path of this length from 1 to 3. There is a unique edge starting at 1, namely, $(1, 2)$, hence the path consists of this edge followed by a directed path of length $n^2 - 4n + 3$ from 2 to 3. In the previous paragraph we have shown that G_n contains no directed path from 2 to 3 with this length.

Suppose that the length of w is equal to n^2-4n+5. The word w sends $n-2$ to 3. There are two edges starting at $n - 2$: the edge $(n-2, 1)$ labelled a and the edge $(n-2, n-1)$ labelled b. Since the letter b acts as a permutation on the state set of the automaton $\mathscr{G}_n$, the word w starts with the letter a. Therefore the first edge of the directed path from $n-2$ to 3 labelled by w is necessarily to edge $(n-2, 1)$ and this edge is followed by a directed path of length $n^2 - 4n + 4$ from 1 to 3. In the previous paragraph we have shown that G_n contains no directed path from 1 to 3 with this length.

Finally, let the length of w is $n^2 - 4n + 6$. The word w sends each of the states $n - 3$ and $n - 1$ to the state 3. If the second letter of the word w is a, then the directed path from $n - 3$ to 3 labelled by w starts with the edges $(n - 3, n - 2)$ and $(n - 2, 1)$ which are followed by a directed path of length $n^2 - 4n + 4$ from 1 to 3, and such a path is impossible. If the second letter of the word w is b, the directed path from $n - 1$ to 3 labelled by w starts with the edges $(n - 1, n)$ and $(n, 1)$, which again must be followed by an impossible directed path of length $n^2 - 4n + 4$ from 1 to 3.

Thus, we have proved that the reset threshold of the automaton $\mathscr{G}_n$ cannot be less than $n^2 - 4n + 7$. $\square$

§5. Discussion and new conjectures

5.1. Two conjectures. The constructions and the results presented in Section 4 witness that the interconnections between reset thresholds of automata with two input letters and exponents of primitive digraphs are sufficiently tight. This conclusion are also supported by recent results by the third author [12]. We believe that these interconnections deserve being further investigated. In order to make the future investigations be more concrete, we formulate a very general conjecture in the flavor of Theorem 1. This conjecture constitutes a strengthening of the Černý conjecture for the case of automata with two input letters and agrees with all theoretical and experimental results that we are aware of.

Conjecture 1. (a) (The Černý conjecture) *The reset threshold of every synchronizing n-automaton with two input letters does not exceed $(n-1)^2$.*

(b) *If $n > 6$, then up to isomorphism there exists exactly one synchronizing n-automaton with two input letters and reset threshold $(n-1)^2$, namely, the automaton $\mathscr{C}_n$.*

(c) *If $n > 6$, then there exists no synchronizing n-automaton with two input letters whose reset threshold is greater than $n^2 - 3n + 4$ but less than $(n-1)^2$.*

(d) *If $n > 7$ and n is odd, then up to isomorphism there exists exactly one synchronizing n-automaton with two input letters and reset threshold n^2-3n+4, namely, the automaton $\mathscr{D}'_n$, exactly two synchronizing n-automata with two input letters and reset threshold n^2-3n+3, namely, the automata $\mathscr{W}_n$ and $\mathscr{F}_n$, and exactly three synchronizing n-automata with two input letters and reset threshold n^2-3n+2, namely, the automata $\mathscr{E}_n$, $\mathscr{D}''_n$, and $\mathscr{B}_n$. There exists no synchronizing n-automaton with two input letters whose reset threshold is greater than n^2-4n+7 but less than n^2-3n+2.*

(e) *If $n > 8$ and n is even, then up to isomorphism there exists exactly one synchronizing n-automaton with two input letters and reset threshold n^2-3n+4, namely, the automaton $\mathscr{D}'_n$, exactly one synchronizing n-automaton with two input letters and reset threshold n^2-3n+3, namely, the automaton $\mathscr{W}_n$, and exactly two synchronizing n-automata with two input letters and reset threshold n^2-3n+2, namely, the automata $\mathscr{E}_n$ and $\mathscr{D}''_n$. There exists no synchronizing n-automaton with two input letters whose reset threshold is greater than n^2-4n+6 but less than n^2-3n+2.*

We also formulate a more special conjecture that can be treated as a quantitative form of the Road Coloring Conjecture mentioned in Section 3. Since now we know that every primitive digraph has a synchronizing coloring [24], the notion of reset threshold can be naturally extended to primitive digraphs. Namely, we call the *reset threshold* of a primitive digraph the minimum length of reset words for all synchronizing colorings of the digraph. This immediately leads to the question of how the reset threshold of a primitive digraph depends on the vertex number.

We notice that the digraphs of slowly synchronizing automata may admit colorings with small reset threshold. Fig. 2 illustrates this remark: the first coloring of the digraph shown in the left is the Černý automaton $\mathscr{C}_4$ whose shortest reset word has length 9 while the second coloring can be reset by the word a^3 of length 3. In this connection, the series W_n is of interest. In this series each digraph has a unique (up to isomorphism) coloring. Therefore the reset threshold of this coloring found in Theorem 2 coincides with the reset threshold of the digraph W_n and provides a lower bound for the problem under consideration. We conjecture that this bound is in fact tight.

Conjecture 2. *The reset threshold of every primitive n-digraph does not exceed $n^2 - 3n + 3$. If $n > 3$, then up to isomorphism there exists exactly one primitive n-digraph with reset threshold $n^2 - 3n + 3$, namely, the digraph W_n.*

Conjecture 2 has been presented in several talks of the second author since 2008 and some partial results towards its proof have already been published, see [8, 21]. It is clear that Conjecture 2 can be made more precise in the flavor of Conjecture 1: for instance, it is likely that the digraph D_n is the only (up to isomorphism) primitive n-digraph with reset threshold $n^2 - 3n + 2$, etc.

5.2. The role of the alphabet size. In our experiments we restrict ourselves to automata with two input letters. This restriction is caused by the fact that an increase in the alphabet size influences the number of automata much stronger than an increase in the state size. Therefore an exhaustive search through all automata with more than two input letters is far beyond our computational capacities even for automata with a modest number of states. Table 3 illustrates this fact. (The data in the table are calculated via a formula from [14].)

Table 3: The number of initially-connected automata with 2 and 3 input letters

# of states	7	8	9
2 input letters	256 182 290	12 665 445 248	705 068 085 303
3 input letters	500 750 172 337 212	572 879 126 392 178 688	835 007 874 759 393 878 655

Nevertheless, there are some reasons to expect that the behaviour of the function we are interested in (the number of synchronizing automata with a fixed number of states as a function of reset threshold) does not heavily depend on the alphabet size. For instance, Trahtman's experiments whose results have been reported in [22, 23], has revealed no 7-automaton with 3 or 4 input letters and with reset threshold larger than 32 and smaller than 36. Thus, the value of the gap between the maximum and the next to maximum possible value of reset threshold is the same as for 7-automata with two input letters.

We mention also the observation which, as far as we know, was first made in [5]: if there exists an upper bound of the form $O(n^2)$ for the reset threshold of synchronizing n-automata with two input letters, then a bound of the same magnitude (but probably with a worse constant) exists also for the reset threshold of synchronizing n-automata with any fixed size of the input alphabet.

References

[1] *R. L. Adler, L. W. Goodwyn, B. Weiss*, Equivalence of topological Markov shifts, Israel J. Math., 27 (1977), 49–63.

[2] *M. Almeida, N. Moreira, R. Reis*, Enumeration and generation with a string automata representation, Theor. Comput. Sci., 387 (2007), 93–102.

[3] *D. S. Ananichev, V. V. Gusev, M. V. Volkov*, Slowly synchronizing automata and digraphs, Mathematical Foundations of Computer Science, Lect. Notes Comput. Sci., 6281 (2010), 55–64.

[4] *D. S. Ananichev, M. V. Volkov, Yu. I. Zaks*, Synchronizing automata with a letter of deficiency 2, Theor. Comput. Sci., 376 (2007), 30–41.

[5] *M. Berlinkov*, Approximating the minimum length of synchronizing words is hard, Computer Science in Russia, Lect. Notes Comput. Sci., 6072 (2010), 37–47.

[6] *M. Berlinkov*, On a conjecture by Carpi and D'Alessandro, Int. J. Foundations Comp. Sci., 22 (2011), 1565–1576.

[7] *R. Brualdi, H. Ryser*, Combinatorial Matrix Theory, Cambridge University Press, 1991.

[8] *A. Carpi, F. D'Alessandro*, On the hybrid Černý-Road Coloring Problem and Hamiltonian paths, Developments in Language Theory, Lect. Notes Comput. Sci., 6224 (2010), 124–135.

[9] *J. Černý*, Poznámka k homogénnym eksperimentom s konečnými automatami, Matematicko-fyzikalny Časopis Slovensk. Akad. Vied, 14, no.3 (1964), 208–216 (in Slovak).

[10] *A. L. Dulmage, N. S. Mendelsohn*, The exponent of a primitive matrix, Can. Math. Bull., 5 (1962), 241–244.

[11] *A. L. Dulmage, N. S. Mendelsohn*, Gaps in the exponent set of primitive matrices, Ill. J. Math., 8 (1964), 642–656.

[12] *V. Gusev*, Lower bounds for the length of reset words in Eulerian automata, Reachability Problems, Lect. Notes Comput. Sci., 6945 (2011), 180–190.

[13] *J. Kari*, A counter example to a conjecture concerning synchronizing words in finite automata, Bull. European Assoc. Theor. Comput. Sci., 73 (2001), 146.

[14] *V. A. Liskovets*, The number of connected initial automata, Kibernetika, no.3 (1969), 16–19 (in Russian).

[15] *J. Olschewski, M. Ummels*, The complexity of finding reset words in finite automata, Mathematical Foundations of Computer Science, Lect. Notes Comput. Sci., 6281 (2010), 568–579.

[16] *J.-E. Pin*, On two combinatorial problems arising from automata theory, Ann. Discrete Math., 17 (1983), 535–548.

[17] *J. L. Ramírez Alfonsín*, The diophantine Frobenius problem, Oxford University Press, 2005.

[18] *V. N. Sachkov, V. E. Tarakanov*, Combinatorics of Nonnegative Matrices. American Mathematical Society, 2002.

[19] *S. Sandberg*, Homing and synchronizing sequences, Model-Based Testing of Reactive Systems, Lect. Notes Comput. Sci., 3472 (2005), 5–33.

[20] *E. Skvortsov, E. Tipikin*, Experimental study of the shortest reset word of random automata, Implementation and Application of Automata, Lect. Notes Comput. Sci., 6807 (2011), 290–298.

[21] *B. Steinberg*, The Černý conjecture for one-cluster automata with prime length cycle, Theor. Comput. Sci., 412 (2011) 5487–5491.

[22] *A. N. Trahtman*, An efficient algorithm finds noticeable trends and examples concerning the Černý conjecture, Mathematical Foundations of Computer Science, Lect. Notes Comput. Sci., 4162 (2006), 789–800.

[23] *A. N. Trahtman*, Notable trends concerning the synchronization of graphs and automata, Electr. Notes Discrete Math., 25 (2006), 173–175.

[24] *A. N. Trahtman*, The Road Coloring Problem, Israel J. Math., 172 (2009), 51–60.

[25] *A. N. Trahtman*, Modifying the upper bound on the length of minimal synchronizing word, Fundamentals of Computation Theory, Lect. Notes Comput. Sci, 6914 (2011), 173–180.

[26] *M. V. Volkov*, Synchronizing automata and the Černý conjecture, Languages and Automata: Theory and Applications, Lect. Notes Comput. Sci., 5196, (2008), 11–27.

[27] *M. V. Volkov*, Synchronizing automata preserving a chain of partial orders, Theor. Comput. Sci., 410 (2009), 3513–3519.

[28] *H. Wielandt*, Unzerlegbare, nicht negative Matrizen, Math. Z., 52, (1950), 642–648 (in German).