On Pansiot Words Avoiding 3-Repetitions

Irina A. Gorbunova
Ural Federal University
Ekaterinburg, Russia
i.a.gorbunova@gmail.com

Arseny M. Shur
Ural Federal University
Ekaterinburg, Russia
arseny.shur@usu.ru

The recently confirmed Dejean's conjecture about the threshold between avoidable and unavoidable powers of words gave rise to interesting and challenging problems on the structure and growth of threshold words. Over any finite alphabet with $k \geq 5$ letters, Pansiot words avoiding 3-repetitions form a regular language, which is a rather small superset of the set of all threshold words. Using cylindric and 2-dimensional words, we prove that, as k approaches infinity, the growth rates of complexity for these regular languages tend to the growth rate of complexity of some ternary 2-dimensional language. The numerical estimate of this growth rate is ≈ 1.2421.

Powers, integral and fractional, are the simplest and most natural repetitions in words. Any repetition over an arbitrary fixed alphabet is characterized by the set of all words over this alphabet, avoiding this repetition. The main question concerning such a set is whether it is finite or infinite. For fractional powers, this question is answered by Dejean's conjecture [5], which is now proved in all cases by the efforts of different authors, see [2–4, 8–11].

Recall that the *exponent* of a word w is the ratio between its length and its minimal period: $\exp(w) = |w|/\operatorname{per}(w)$. If $\exp(w) = \beta > 1$, then w is a *fractional power* (β-*power*). It is convenient to treat the notion of β-power as follows: a word w is a β-power if $\exp(w) \geq \beta$ while $(|w|-1)/\operatorname{per}(w) < \beta$, and a β^+-power if $\exp(w) > \beta$ while $(|w|-1)/\operatorname{per}(w) \leq \beta$. As usual, β^+ is treated as a "number", covering β in the usual $\leq$ order. A word is called β-*free* (where β can be a number with plus as well) if it contains no β-powers as factors. A β-power is k-*avoidable* if the number of k-ary β-free words is infinite. Dejean's conjecture states that a β-power is k-avoidable if and only if

$$\beta \geq (7/4)^+ \text{ and } k = 3, \ \ \beta \geq (7/5)^+ \text{ and } k = 4, \ \text{ or } \beta \geq (k/(k-1))^+ \text{ and } k = 2, k \geq 5.$$

The $(k/(k-1))^+$-free languages over k-letter alphabets, where $k \geq 5$, are called *threshold languages*; we denote them by T_k. We study structure and growth of these languages, aiming at the asymptotic properties as the size of the alphabet increases.

Any threshold language can be approximated from above by a series of regular languages consisting of words that *locally* satisfy the $(k/(k-1))^+$-freeness property. Namely, these words avoid all $(k/(k-1))^+$-powers w such that $|w| - \operatorname{per}(w) \leq m$, for some constant m. From our previous work [12], it is clear that the case $m = 3$ gives a lot of important structural information about the languages T_k. Here we study this case in details, using *cylindric representation* that captures the properties common for considered words over all alphabets.

1 Preliminaries

We study finite words and two-sided infinite words (*Z-words*) over finite k-letter alphabets Σ_k and over some special ternary alphabet introduced below. We also consider 2-dimensional words, which are just finite rectangular arrays of alphabetic symbols. Unlike to some commonly used models of 2-dimensional

words (cf. [7]), we do not use additional symbols to mark the borders of such a word. Factors of 2-dimensional words are also 2-dimensional words.

A (1- or 2-dimensional) language is *factorial*, if it is closed under taking factors of its words. A word *w avoids* a word *u* if *u* is not a factor of *w*. The set of all minimal (with respect to the factor order) words avoided by all elements of a factorial language *L* is called the *antidictionary* of *L*. All 1-dimensional languages with finite antidictionaries are regular.

We denote the antidictionary of the threshold language T_k by A_k. A word $u \in A_k$ can be factorized as $u = yzy$, where $|yz| = \mathrm{per}(u)$, $|u|/|yz| > k/(k-1)$, and all proper factors of u have the exponent at most $k/(k-1)$. If $|y| = m$, we call u an *m-repetition*.

The finite set $A_k^{(m)} \subset A_k$ consists of all *r*-repetitions with $r \leq m$. The notation $T_k^{(m)}$ is used for the (regular) language with the antidictionary $A_k^{(m)}$. Then, $T_k \subseteq T_k^{(m)}$. Since an infinite regular language contains arbitrary powers of some word, one has $T_k \subset T_k^{(m)}$. Clearly, $T_k = \bigcap_{m=1}^{\infty} T_k^{(m)}$.

The *combinatorial complexity* of a language *L* is a function $C_L(n)$ which returns the number of words in *L* of length *n*. This function serves as a natural quantitative measure of *L*. "Big" ["small"] languages have exponential [resp., subexponential] complexity. Exponential complexity can be described by means of the *growth rate* $\alpha(L) = \limsup_{n \to \infty} (C_L(n))^{1/n}$ (subexponential complexity is indicated by $\alpha(L) = 1$). For factorial languages, classical Fekete's lemma implies

$$\alpha(L) = \lim_{n \to \infty} (C_L(n))^{1/n} = \inf_{n \to \infty} (C_L(n))^{1/n}.$$

The growth rate of $T_k^{(m)}$ approximates the growth rate of T_k from above. It is easy to prove that $\lim_{m \to \infty} \alpha(T_k^{(m)}) = \alpha(T_k)$.

For regular languages, the growth rate equals the *index* (spectral radius of the adjacency matrix) of recognizing automaton, providing that this automaton is *consistent* (each vertex belongs to some accepting walk), and either deterministic, or non-deterministic but *unambiguous* (there is at most one walk with the given label between two given vertices); see [13].

In [10], Pansiot showed how to encode all words from the language $T_k^{(2)}$ with "characteristic" words over the alphabet $\{0,1\}$. This encoding played a big role in the proof of Dejean's conjecture; so, we refer to the elements of $T_k^{(2)}$ as to *Pansiot words*. These words can be equivalently defined by the following pair of conditions:

(P1) two closest occurrences of a letter are on the distance $k-1$, k, or $k+1$;

(P2) two closest occurrences of a letter are followed by different letters.

We also consider *Pansiot Z-words*, which are given by (P1), (P2) as well. Finite factors of Pansiot Z-words are exactly Pansiot words.

Now we introduce *cylindric representation* of Pansiot words. Imagine such a word (finite or infinite) as a rope with knots, which are representing letters. This rope is wound around a cylinder such that the knots at distance *k* are placed one under another (Fig. 1, a). By (P1), the knots labeled by two closest occurrences of the same letter appear on two consecutive winds of the rope one under another or shifted by one knot (Fig. 1, b). If we connect these closest occurrences by "sticks", we get three types of such sticks: vertical, left-slanted, and right-slanted (Fig. 1, b). We associate each letter in a Pansiot word with a stick going up from the corresponding knot, getting an encoding of this word by a *cylindric word* over the ternary alphabet $\Delta = \{\,\mathord{\updownarrow}, \nearrow, \searrow\,\}$. Since the sticks allow one to establish equality of letters in a Pansiot word, such a cylindric word [Z-word] uniquely represents the original word [resp., Z-word] up
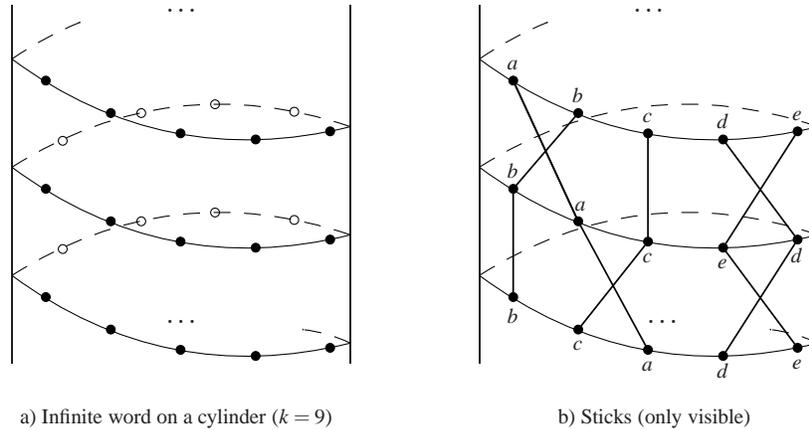
a) Infinite word on a cylinder ($k = 9$)   b) Sticks (only visible)

Figure 1: *Cylindric representation of Pansiot words.*

to the permutation of the alphabet. Note that cylindric words avoid squares of letters in view of (P2). Hence, cylindric Z-words are just infinite sequences of blocks ⌇✕ and ✕.

The feature of cylindric words is that they have an additional 2-dimensional structure, allowing one to capture structural properties of Pansiot words through 2-dimensional factors of cylindric words. We say that a Z-word W is *compatible* to a language L if all factors of W belong to L.

Theorem 1 ([12]). *For any integer $m \geq 3$, there exists a set S_m of 2-dimensional words of size $O(m) \times O(m)$ over Δ such that for any $k \geq 2m-3$, a Pansiot Z-word W over Σ_k is compatible to $T_k^{(m)}$ if and only if the corresponding cylindric Z-word has no 2-dimensional factors from S_m.*

This theorem states that cylindric words that encode the words from $T_k^{(m)}$ are defined by 2-dimensional avoidance properties. For example, cylindric words of the Pansiot words avoiding 3-repetitions are defined by the avoidance of the structures ⌇ and ⟍⟋. Indeed, any of these structures implies the existence of three successive letters (say, a, b, and c) in the encoded Pansiot word such that two occurrences of the factor *abc* appear one under another at the distance $2k$; since $(2k+3)/2k > k/(k-1)$, the encoded word contains a 3-repetition.

For a language L, let $\widehat{L}$ be its subset consisting of all factors of Z-words compatible to L. By [14, Theorem 3.1], $\alpha(\widehat{L}) = \alpha(L)$. Let $\mathsf{Cyl}_k^{(m)}$ be the set of all factors of cylindric Z-words encoding Pansiot Z-words compatible to $T_k^{(m)}$. Then clearly $\alpha(\mathsf{Cyl}_k^{(m)}) = \alpha(\widehat{T}_k^{(m)}) = \alpha(T_k^{(m)})$. Thus, the growth rates of threshold languages can be estimated through the study of cylindric words with simple avoidance properties that are independent of the size of the alphabet. In what follows, we refer to the elements of $\mathsf{Cyl}_k^{(m)}$ as cylindric factors.

The above considerations imply two natural conjectures: *for any fixed $m \geq 3$, the sequence $\{\alpha(T_k^{(m)})\}_5^\infty$ has a limit as k approaches infinity, and this limit is the "growth rate" of the 2-dimensional language defined by the same avoidance properties as* $\mathsf{Cyl}_k^{(m)}$. Through the computations of growth rates for the alphabets with $5, 6, \ldots, 60$ letters we observed in [12] that the sequence $\{\alpha(T_k^{(3)})\}$ demonstrates fast convergence to the limit ≈ 1.242096777.

In this paper, we confirm both conjectures for the case $m = 3$. The corresponding 2-dimensional language will be denoted by D; it consists of all rectangular words over Δ having no factors ⌇ and ⟍⟋.

In fact, the case $m = 3$ is the crucial one to approximate the growth rates of threshold languages, because in [12] it was shown that

- there is no 4- and 5-repetitions;

- m-repetitions with $m \geq 6$ do not affect significantly the growth rate, as far as we can check this by extensive computer-assisted studies based on the results of [13].

2   Two-dimensional languages

Combinatorial complexity $C_L(n,k)$ of a 2-dimensional language L is the function returning the number of $n \times k$ words in L. If L is factorial, then its *growth rate* is defined by the formula

$$\alpha(L) = \lim_{n,k \to \infty} (C_L(n,k))^{1/nk}. \tag{1}$$

The function $C_L(n,k)$ in this case is submultiplicative for each variable, and hence the existence of the limit (1) follows from the multivariate version of Fekete's lemma [1].

On the other hand, it is completely unclear how to calculate the growth rates of 2-dimensional languages. For the 1-dimensional case, the growth rate of a regular language can be found quite efficiently, see [13]. Here we give one idea how to estimate the growth rate of a 2-dimensional language. Since the limit (1) exists, we can take any "diagonal" subsequence of $C_L(n,k)$; we choose $\{(C_L(n,n))^{1/n^2}\}_1^\infty$. Applying Stolz's Theorem (see [6]) twice, we get

$$\alpha(L) = \lim_{n \to \infty} (C_L(n,n))^{1/n^2} = \lim_{n \to \infty} \left(\frac{C_L(n,n)}{C_L(n-1,n-1)} \right)^{1/(2n-1)} = \lim_{n \to \infty} \frac{(C_L(n,n)C_L(n-2,n-2))^{1/2}}{C_L(n-1,n-1)}$$

if the last two limits exist. Calculating the values of these sequences for the language D (see Table 1), we see that the last sequence has the best behaviour and allows one to suggest $\alpha(D) \approx 1.2421$. Thus, we get an additional support to the conjecture that $\alpha(D)$ is the limit of the sequence $\{\alpha(\text{Cyl}_k^{(3)})\}_5^\infty$. For the rest of the paper, we set $C(n,k) = C_D(n,k)$.

Table 1: *Approximation to the growth rate of the 2-dimensional language D.*

n	$(C_D(n,n))^{1/n^2}$	$\left(\frac{C_D(n,n)}{C_D(n-1,n-1)} \right)^{1/(2n-1)}$	$\frac{(C_D(n,n)C_D(n-2,n-2))^{1/2}}{C_D(n-1,n-1)}$
3	1.627251	1.438233	1.191687
4	1.525034	1.402991	1.318617
5	1.464419	1.362547	1.229958
...	...	...	...
27	1.280207	1.261332	1.242089
28	1.278823	1.260626	1.242080
29	1.277537	1.259972	1.242104
30	1.276337	1.259362	1.242102

3   Automata

Let us fix an arbitrary $k \geq 5$. We denote the set of all words of width k from D by D_k. It is natural to put $\alpha(D_k) = \lim_{n \to \infty}(C(n,k))^{1/nk}$; then $\lim_{k \to \infty} \alpha(D_k) = \alpha(D)$ as the iterative limit of the existing double limit. Note that D_k can be also viewed as a 1-dimensional regular language over the alphabet Δ^k. The automaton $\mathscr{A}$ recognizing D_k can be defined as follows:

(A1) the words of length k from $\mathsf{Cyl}_k^{(3)}$ (they coincide with the words of size $1 \times k$ from D_k) are the vertices;

(A2) an edge $u \to v$ exists if and only if the word ^u_v of size $2 \times k$ belongs to D_k; such an edge is labeled by v;

(A3) each vertex is both initial and terminal.

Note that $\mathscr{A}$ is an unambiguous nondeterministic automaton recognizing D_k as a language over Δ^k. The index of $\mathscr{A}$ (and the growth rate of D_k over Δ^k) equals $\alpha(D_k)^k$. The underlying graph of $\mathscr{A}$ is undirected due to vertical symmetry of the avoided factors. Let $P_u(n)$ be the number of walks of length n in $\mathscr{A}$, starting at the vertex u, $P(n) = \sum P_u(n)$ be the number of all walks of length n in $\mathscr{A}$. Then $P(n) = C(n+1,k)$.

For the language $\mathsf{Cyl}_k^{(3)}$, we build the *Rauzy graph* $\mathscr{R}$ of order $k+1$. The vertices of this graph are the words of $\mathsf{Cyl}_k^{(3)}$ of length $k+1$, and a directed edge connects a vertex u to v if and only if some word of $\mathsf{Cyl}_k^{(3)}$ of length $k+2$ has the prefix u and the suffix v. It is easy to see that the edges of $\mathscr{R}$ can be labeled such that $\mathscr{R}$ becames a deterministic *cover automaton* (all transitions are deterministic, all vertices are both initial and terminal), recognizing the language $\mathsf{Cyl}_k^{(3)}$. Deterministic cover automaton is a special case of unambiguous nondeterministic automaton; so, the index of $\mathscr{R}$ equals $\alpha(\mathsf{Cyl}_k^{(3)})$. Now consider the kth power $\mathscr{R}^k$ of $\mathscr{R}$. Note that in most cases the correctness of transition from some vertex u of $\mathscr{R}^k$ to some other vertex v can be checked using only k last symbols of u. The only exception is the case when the k-letter suffix of u begins and ends with $\times$: if u begins with $\mathbf{!}$, then the k-letter suffix of v can begin with both $\mathbf{!}$ and $\diagup$, while if u begins with $\diagdown$, then this suffix of v must begin with $\mathbf{!}$ to prevent the appearance of the avoided 2×2 factor. Let us require v to begin with $\mathbf{!}$ in any case and consider the automaton $\mathscr{B}$ such that

(B1) the words of length k from $\mathsf{Cyl}_k^{(3)}$ (the suffixes of length k of the vertices from $\mathscr{R}^k$) are the vertices;

(B2) an edge $u \to v$ exists if and only if (a) the automaton $\mathscr{R}^k$ contains the edge $au \to bv$ for some $a, b \in \Delta$, and (b) if u has the form $\times \cdots \times$, then v begins with $\mathbf{!}$; such an edge is labeled by v;

(B3) each vertex is both initial and terminal.

We will write $P'_u(n)$ for the number of walks of length n in $\mathscr{B}$, starting at u, and $P'(n) = \sum P'_u(n)$ for the number of all walks of length n in $\mathscr{B}$. If we denote the number of words of length nk in the language $\mathsf{Cyl}_k^{(3)}$ by $C'(n,k)$, then it is easy to see that $P'(n) \leq C'(n+1,k) \leq C(n+1,k)$.

4   Main result

Since the indices of automata depend only on their adjacency matrices, below we consider the automata $\mathscr{A}$ and $\mathscr{B}$ just as digraphs. Recall that they share the same set of vertices and any edge of $\mathscr{B}$ is contained in $\mathscr{A}$. The outdegrees of a vertex u in $\mathscr{A}$ and $\mathscr{B}$ are denoted respectively by $\deg^+_{\mathscr{A}}(u)$ and $\deg^+_{\mathscr{B}}(u)$. We say that the vertices u and v are *similar* if they coincide up to the first 11 letters. Similarity is an equivalence relation; we write $u \sim v$.

Remark 1. *The classes of $\sim$ are finite, since the cardinality of such a class is the number of words of length* 11 *over* Δ *that can be extended by the same suffix. The maximum cardinality of such a class is $N = 28$ independently of k, and is achieved on any suffix that begins with* ╱ .

The following two key lemmas hold for any $k \geq 12$ (this restriction is necessary only for the existence of 12th symbol in the label of the vertex).

Lemma 1. *For any vertex $u = u_1 \ldots u_k$ and any $a \in \Delta$ such that either $a \neq$ ╎ or $u_{12} \neq$ ╎ , there exists an edge $u \rightarrow x$ in $\mathscr{B}$ such that the 12th letter of x is a.*

Proof. Let $x = x_1 \cdots x_k$. We first show that if the condition of the lemma holds for some *i*th letter $(1 \leq i \leq k)$ then it also holds for any *j*th letter $(i < j \leq k)$. It suffices to check the case $j = i+1$. Indeed, the minimal structures avoided by the words from $\text{Cyl}_k^{(3)}$ are either factors of length 2, or the "vertical factor" ╎ of height 2, or the "square factor" ╲╱ of size 2×2. Thus, the possible values of x_{i+1} are determined by u_i, u_{i+1}, and x_i; each of these values together with u_{i+1} and u_{i+2} determine the possible values of x_{i+2}, and so on. There are only four possibilities for the factor $u_i u_{i+1}$. For each of them, we show that if the symbol x_i can take all possible values, then the same is true for x_{i+1}, see Fig. 2.
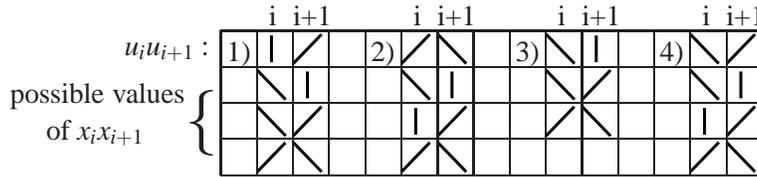

Figure 2: *Proving Lemma 1. If x_i can take any value, x_{i+1} can take any value as well.*

In order to prove the lemma we find, for each vertex u, the number i_u such that the i_uth letter of x can take any value required by the condition of the lemma. If $i_u \leq 12$ for any u, then we are done with the proof. So we examine all possible beginnings of u and try to build the word $x_1 \cdots x_{i_u}$ such that $x_{i_u} = a$ for any allowed $a \in \Delta$. Recall that the letter x_1 follows u_k in some cylinder word and hence, depends on u_k. In order to avoid the consideration of u_k (the restrictions involving u_k depend on k), we build the word $x_1 \cdots x_{i_u}$ for any $x_1 \in \Delta$. The word $x_1 \cdots x_{i_u}$ for all u that begin with ╎ and ╱ is shown in Fig. 3 (cases 1–3 and 4–11, respectively). The maximum value of i_u, namely 11, is achieved in case 9. If u begins with ╲, then its factor $u_2 \ldots u_{i_u}$ falls into one of the cases 1–11, so, we conclude that $i_u \leq 12$. ☐

Lemma 1 is used to prove another property of similarity.

Lemma 2. *If $u \sim v$ and $u \rightarrow x$ is an edge in $\mathscr{A}$, then there exists an edge $v \rightarrow y$ in $\mathscr{B}$ such that $x \sim y$.*

Proof. Let $u = u_1 \cdots u_k$, $x = x_1 \cdots x_k$, $v = v_1 \cdots v_k$, and we have to find the vertex $y = y_1 \cdots y_k$. Assume that we know only the letters $u_{12}, \ldots, u_k$, and x_{12}. Then we still can restore all possible values of the factor $x_{13} \cdots x_k$ independently of the letters $u_1, \ldots, u_{11}, x_1, \ldots, x_{11}$ (cf. the proof of Lemma 1).

Now consider all y's such that $v \rightarrow y$ is an edge in $\mathscr{B}$ and $y_{12} = x_{12}$. The set of all such y's is nonempty by Lemma 1. Since $v_{12} \cdots v_k = u_{12} \cdots u_k$ by similarity of u and v, the set of all possible values of the factor $y_{13} \cdots y_k$ coincides with such a set for the factor $x_{13} \cdots x_k$. Thus, we can pick up y so that the factor $y_{12} \cdots y_k$ equals $x_{12} \cdots x_k$ for the actual value of x. Then $x \sim y$, and the lemma is proved. ☐

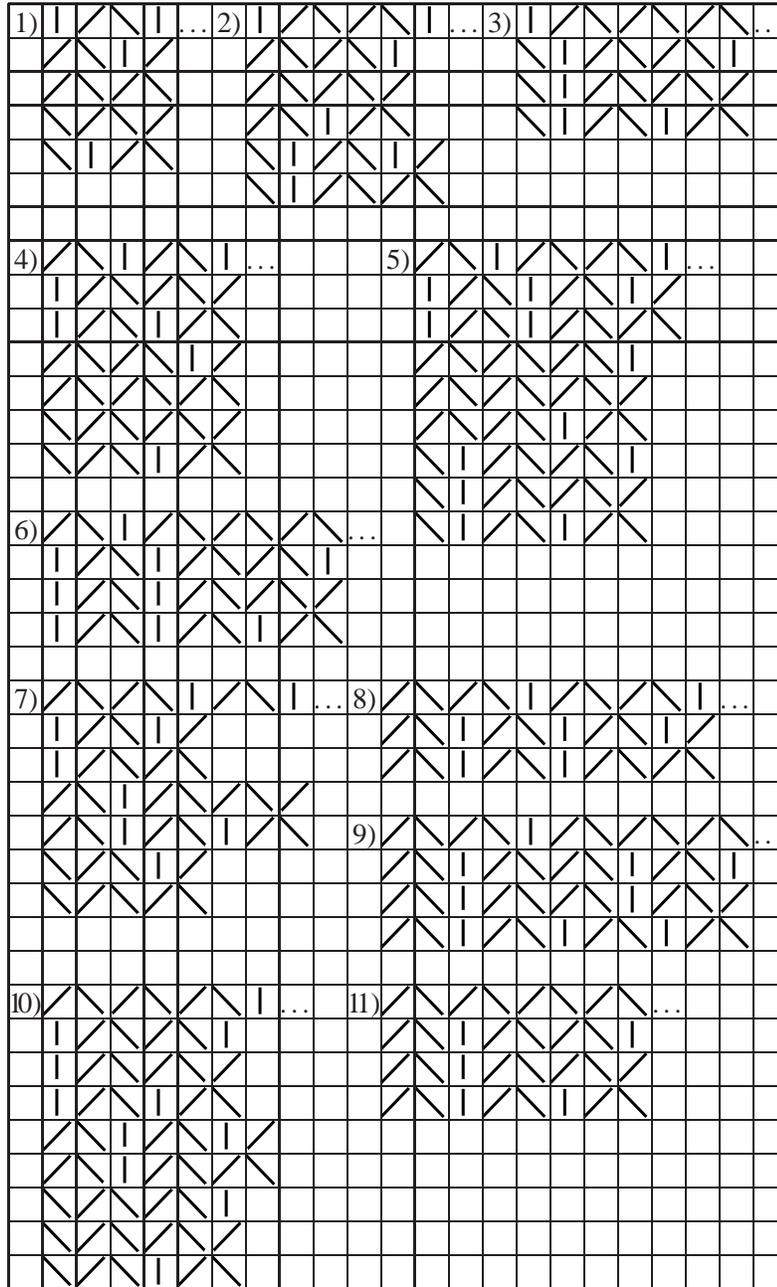**Theorem 2.** *The limit $\lim_{k \rightarrow \infty} \alpha(T_k^{(3)})$ exists and is equal to $\alpha(D)$.*

Figure 3: *Proving Lemma 1. Cases 1–11 represent different beginnings of the word u. Under each beginning, some possible beginnings of the word x are drawn. For each possible first letter of x, we exhibit such beginnings ending by all possible letters. In some cases, not all possible beginnings of x are drawn; for such missing beginnings, case 3 refers to case 2, case 6 to case 5, cases 8 and 9 to case 7, and case 11 to case 10.*

Proof. Recall that $\alpha(T_k^{(3)}) = \alpha(\mathsf{Cyl}_k^{(3)})$. Since the sequence $\big(C_{\mathsf{Cyl}_k^{(3)}}(n)\big)^{1/n}$ converges to $\alpha(\mathsf{Cyl}_k^{(3)})$, so does any its subsequence. Hence, $\alpha\big(\mathsf{Cyl}_k^{(3)}\big) = \lim_{n\to\infty}\big(C'(n,k)\big)^{1/nk}$. On the other hand, we know that $\alpha(D) = \lim_{k\to\infty}\alpha(D_k) = \lim_{k\to\infty}\lim_{n\to\infty}\big(C(n,k)\big)^{1/nk}$. Thus, let us estimate the ratio $\big(C'(n,k)/C(n,k)\big)^{1/nk}$. The upper bound $\big(C'(n,k)/C(n,k)\big)^{1/nk} \leq 1$ is trivial. In order to get the lower bound, we recall that $C(n+1,k) = P(n)$ and $C'(n+1,k) \geq P'(n)$.

Let us fix an arbitrary vertex u and consider the $\mathscr{A}$-*tree* (for u) defined as follows. The vertices of this tree are labeled by the vertices of $\mathscr{A}$, u being the label of the root. Any vertex labeled by v has $\deg_{\mathscr{A}}^+(v)$ children; the children are labeled by all forward neighbours of v in $\mathscr{A}$. Thus, there is a natural bijection between the set of vertices of level n in the $\mathscr{A}$-tree and the set of all walks from u of length n in the automaton $\mathscr{A}$. That is, nth level of the $\mathscr{A}$-tree contains exactly $P_u(n)$ vertices. The $\mathscr{B}$-*tree* is defined in the same way, using $\mathscr{B}$ instead of $\mathscr{A}$. The nth level of the $\mathscr{B}$-tree contains $P'_u(n)$ vertices.

Using Lemma 2 inductively, we get that the label of any vertex of nth level in the $\mathscr{A}$-tree is similar to the label of some vertex of nth level in the $\mathscr{B}$-tree. Let us start from the roots of the trees and inductively construct a total map μ from the $\mathscr{A}$-tree to the $\mathscr{B}$-tree satisfying the following conditions:

(1) if s is a level n vertex labeled by x, then $\mu(s)$ is a level n vertex labeled by some $y \sim x$;

(2) $\mu(\mathrm{parent}(s)) = \mathrm{parent}(\mu(s))$.

The existence of such a map is ensured by Lemma 2 and the structure of trees.

Now we take a level n vertex t from the $\mathscr{B}$-tree and estimate the size of the set $\mu^{-1}(t)$. Assume that $|\mu^{-1}(\mathrm{parent}(t))| = K$. If s is mapped to t, then $\mathrm{parent}(s) \in \mu^{-1}(\mathrm{parent}(t))$. All children of the vertex $\mathrm{parent}(s)$ are different. Hence, by Remark 1, at most N of these children can be mapped to t. Thus, $|\mu^{-1}(t)| \leq KN$. The case $n = 0$ gives us $|\mu^{-1}(t)| = 1$ whence we obtain $|\mu^{-1}(t)| \leq N^n$. Since μ is total, we have $P_u(n) \leq N^n P'_u(n)$. Summing up these inequalities for all vertices u, we finally get $P(n) \leq N^n P'(n)$.

Returning to combinatorial complexities, we can write

$$\frac{1}{N^n} \leq \frac{P'(n)}{P(n)} \leq \frac{C'(n+1,k)}{C(n+1,k)} \leq 1,$$

$$\left(\frac{1}{N^n}\right)^{1/(n+1)k} \leq \left(\frac{C'(n+1,k)}{C(n+1,k)}\right)^{1/(n+1)k} \leq 1.$$

Taking the limits of all sides as $n \to \infty$, we get

$$\left(\frac{1}{N}\right)^{1/k} \leq \frac{\alpha(T_k^{(3)})}{\alpha(D_k)} \leq 1.$$

Now we let $k \to \infty$ and use the squeese theorem to conclude that the limit $\lim_{k\to\infty}\alpha(T_k^{(3)})/\alpha(D_k)$ exists and is equal to 1 (recall that N is independent of k). Since the limit $\lim_{k\to\infty}\alpha(D_k) = \alpha(D)$ also exists, we have

$$\alpha(D) = \alpha(D)\cdot 1 = \lim_{k\to\infty}\alpha(D_k)\cdot\lim_{k\to\infty}\frac{\alpha(T_k^{(3)})}{\alpha(D_k)} = \lim_{k\to\infty}\frac{\alpha(T_k^{(3)})}{\alpha(D_k)}\cdot\alpha(D_k) = \lim_{k\to\infty}\alpha(T_k^{(3)}),$$

as desired. $\qquad\square$

Remark 2. *From the proof of the above theorem it is clear that the actual value of the constant N such that $P(n) \approx N^n P'(n)$ is much smaller than 28. Computations show that $N \approx 2.119$. Hence, the set D_k of 2-dimensional words of width k is not much bigger than the corresponding set $\mathsf{Cyl}_k^{(3)}$ of cylindric words.*

References

[1] S. Capobianco (2008): *Multidimensional cellular automata and generalization of Fekete's lemma*. Discrete Mathematics and Theoretical Computer Science 10(3), pp. 95–104.

[2] A. Carpi (2007): *On Dejean's conjecture over large alphabets*. Theoretical Computer Science 385, pp. 137–151, doi: 10.1016/j.tcs.2007.06.001.

[3] J. D. Currie & N. Rampersad (2009): *Dejean's conjecture holds for n ≥ 27*. RAIRO Theoretical Informatics and Applications 43, pp. 775–778, doi: 10.1051/ita/2009017.

[4] J. D. Currie & N. Rampersad (2011): *A proof of Dejean's conjecture*. Mathematics of Computation 80, pp. 1063–1070, doi: 10.1090/S0025-5718-2010-02407-X.

[5] F. Dejean (1972): *Sur un Theoreme de Thue*. Journal of Combinatorial Theory. Series A 13(1), pp. 90–99.

[6] G. M. Fichtenholz (2001): *Differential and integral calculus*, volume 1, Fizmatlit, Moscow.

[7] D. Giammarresi & A. Restivo, G. Rozenberg & A. Salomaa, editors (1997): *Two-dimensional languages*, Handbook of Formal Languages, volume 3, pp. 215–268, Springer, Berlin.

[8] M. Mohammad-Noori & J. D. Currie (2007): *Dejean's conjecture and Sturmian words*. European Journal of Combinatorics 28, pp. 876–890, doi: 10.1016/j.ejc.2005.11.005.

[9] J. Moulin-Ollagnier (1992): *Proof of Dejean's Conjecture for Alphabets with 5, 6, 7, 8, 9, 10 and 11 Letters*. Theoretical Computer Science 95(2), pp. 187–205, doi: 10.1016/0304-3975(92)90264-G.

[10] J.-J. Pansiot (1984): *A propos d'une conjecture de F. Dejean sur les répétitions dans les mots*. Discrete Applied Mathematics 7, pp. 297–311, doi: 10.1016/0166-218X(84)90006-4.

[11] M. Rao (2011): *Last Cases of Dejean's Conjecture*. Theoretical Computer Science 412(27), pp. 3010-3018, doi: 10.1016/j.tcs.2010.06.020.

[12] A. M. Shur & I. A. Gorbunova (2010): *On the growth rates of complexity of threshold languages*. RAIRO Theoretical Informatics and Applications 44, pp. 175–192, doi: 10.1051/ita/2010012.

[13] A.M. Shur (2010): *Growth rates of complexity of power-free languages*. Theoretical Computer Science 411, pp. 3209–3223, doi: 10.1016/j.tcs.2010.05.017.

[14] A.M. Shur (2008): *Comparing complexity functions of a language and its extendable part*. RAIRO Theoretical Informatics and Applications 42, pp. 647–655, doi: 10.1051/ita:2008021.