The Formal Language Theory Column

BY

GIOVANNI PIGHIZZINI

Dipartimento di Informatica
Università degli Studi di Milano
20135 Milano, Italy
pighizzini@di.unimi.it

Words is a well-established international conference devoted to all aspects of Combinatorics on Words. The contribution by Jean Néraud collects and classifies several open problems and conjectures presented during more than 20 years. I am sure that it will be very stimulating for people interested in this area and, maybe, also for someone not familiar with these topics. I hope to see very soon discussions, progresses and solutions. As written in the article, the document should be gradually updated, so please help the author to do that.

Of course, any contribution to the Formal Language Column will be welcome!

Conferences *WORDS*, years 1997-2017: Open Problems and Conjectures

Jean Néraud
litis, Université de Rouen, France
jean.neraud@univ-rouen.fr neraud.jean@free.fr
http://neraud.jean.free.fr

Abstract

In connection with the development of the field of Combinatorics on Words, we present a list of open problems and conjectures which were stated in the context of the eleven international meetings WORDS, which held from 1997 to 2017.

Contents

Foreword

The first international conference WORDS was organized in 1997 in Rouen, France. Since then, a series of eleven meetings held: in [61, 62], we provided a summary of the contributions which were presented at the ten first of them, in connection with the development of the field of Combinatorics on Words.

The aim of the present paper is to bring some noticeable complementary information, with two key objectives:
- Beforehand, we provide a nomenclature of some of the challenging conjectures and problems which were stated at the occasion of the conferences WORDS.
- In another hand, with regards to the the state-of-the-field, we hardly wish to continually update the present study by including the most recent advances in the framework of the listed questions.

With regards to their corresponding scientific thematics, all the open questions and the conjectures which were stated, refer to the classification that we introduced in [61, 62]. To be more precise, according to the frequency of their presentations, questions and conjectures have been classified in three main domains: the topic of Patterns, that of Complexity and the one of Factorization. From a practical point of view, each statement is nomenclatured by referencing to its topic and to the year of the corresponding meeting WORDS; names of speakers, bibliographic references, and short introductions to the problematics are also provided.

The present document should be gradually updated: in view of this, please contact its author; clearly, some bibliographic references, if any, would be welcome.

Evidently, each of the numerous results, questions and conjectures which were presented during these eleven conferences WORDS plays a noticeable part in the state-of-the-field. From this point of view, we wish that our study will bring some valuable information to the researchers from the community.

1 The topic of patterns

Let Σ, A be two finite alphabets and let $p \in \Sigma^*$, $w \in A^* \cup A^\omega$. We say that the word w *encounters* p if a non-erasing morphism $h : \Sigma^* \longrightarrow A^*$ exists such that $w \in A^*h(p)(A^* \cup A^\omega)$; otherwise the word w *avoids* p or, equivalently said, is p-free. In this context p is refered as a *pattern*, moreover we impose that the morphism h satisfies $h(a) = a$ for any letter $a \in \Sigma \cap A$. The pattern p is *k-avoidable* if an infinite word avoiding p exists over a k-letter alphabet. From this point of view, it is well known that, on the alphabet $A = \{0, 1\}$, the infinite word of Thue-Morse have the fundamental property that it avoids any pattern of type $dXdXd$, for each letter $d \in A$ and any $X \in \Sigma^*$.

1.1 Avoidance of patterns

In the topic, avoidance of patterns is a central question: it has inspired lots of problems:

–WORDS 1997:

Authors: Roman Kolpakov, Gregory Kucherov and Yuri Tarannikov [58, 161–175].
For a natural $n \geq 2$, a word is *nth power-free* if it does not contain any nth power of a non-empty word as a factor. Given $A = \{0, 1\}$, denote by *PF(n)* the corresponding set of such words and set $\rho(n) = \underline{\lim}_{k\to\infty}(\frac{1}{k} \cdot \min\{|w|_1 : w \in PF(n) \cap A^k\})$ (the minimal density of the letter 1 in the words of $PF(n)$).
In their paper, the authors prove that:

$$(\forall n \geq 3)\ (\exists C > 0)\ \ \rho(n) \leq \frac{1}{n} + \frac{1}{n^3} + \frac{1}{n^4} + \frac{C}{n^5}.$$

The mapping ρ can be extended to real arguments: given a real $x \in \mathbb{R}$, denote by $PF(x)$ the set of the binary words that do not contain a factor of exponent not smaller than x. Actually, the authors proved that ρ is discontinuous to the right in each point of $\{7/3\} \cup \{n \in \mathbb{N} | n \geq 3\}$, moreover, they asked the following questions:

- **Question 1.1.97.1:** Does ρ has other discontinuity? What are they? Is ρ piece-wise constant?

- **Question 1.1.97.2:** If a pattern is not k-avoidable, but is $(k + 1)$-avoidable, what is the minimal frequency of a letter in an infinite word over $k+1$ letters that avoids that pattern?

- **Question 1.1.97.3:** Kirby Baker, Georges F. McNulty and Walter Taylor have shown that the pattern $abXbcYcaZbaTac$ is 4-avoidable, but not 3-avoidable [8]. What is the minimal proportion of the fourth letter needed to avoid that pattern?

–WORDS 2003:

Author: James Currie [30, 7–18].
The author reviews some results concerning words avoiding pattern. He recall a lot of open problems. Let's begin by two purely algorithmic questions:

- **Question 1.1.03.1:** Is it decidable whether p is k-avoidable, given a pattern p and an integer k?

- **Question 1.1.03.2:** Given a pattern p, what is the complexity of deciding whether p is avoidable?

With regard to k-avoidability itself, three open problems were stated:

- **Question 1.1.03.3:** Is there a patten that is 6-avoidable but not 5-avoidable?

- **Question 1.1.03.4:** Is *aabaacbaab* 3-D0L-avoidable (i.e. is there a ternary morphism g such that $g^{\omega}(a)$ avoids *aabaacbaab*)?

- **Conjecture 1.1.03.5:** If a pattern is k-avoidable then it is k-HD0L-avoidable (i.e. are there morphisms $f : \Sigma^* \longrightarrow A^*$, $g : \Sigma^* \longrightarrow \Sigma^*$, with $|\Sigma| = k$ such that $f(g^{\omega}(a))$ avoids p)?

The so-called probabilistic method is often use in tackling many problems in discrete mathematics [3, 3–5]. When trying to prove that a structure with certain properties exists, such a method consists in constructing a convenient probability space of structures and then, in showing that the desired properties hold in this space with a non-zero probability.

- **Question 1.1.03.6:** Explore the applications of the probabilistic method in the scope of pattern avoidance.

The so-called circular words were also concerned by some conjectures:

- **Conjecture 1.1.03.7:** If p is k-avoidable, then there are arbitrary long circular words on k letters that avoid p.

- **Conjecture 1.1.03.8:** If p is k-avoidable then there are circular words with length $|p|$ on k letters that avoid p.

- **Conjecture 1.1.03.9:** Let p be k-avoidable.

  1. If the number of p-free words on k letters of length n grows exponentially with n, then an integer N_0 exists such that, for every $n > N_0$, there are circular p-free words with length n on k letters.
  2. If the number of p-free words on k letters of length n grows polynomially with n, then the set of possible lengths for circular p-free words on k letters has density 0 in the set $\mathbb{N} \setminus \{0\}$.

- **Question 1.1.03.10:** The number of k-power-free binary words of length n grows polynomially with n for $k \leq 7/3$, but exponentially for $k > 7/3$ [39]. Examine analogous results for alphabets of arbitrary size.

- **Conjecture 1.1.03.11:** Extension of a result from [8]: the set of circular words over $\{0, 1, 2, 3\}$ avoiding the pattern $abXbcYcaZbaTac$ has density 0 in the set $\mathbb{N} \setminus \{0\}$.

Given an alphabet Σ and $w \in \Sigma^*$, the word w is maximal p-free if p encounters any word in $\Sigma w \cup w\Sigma$. The three following conjectures were stated:

- **Conjecture 1.1.03.12:** Let Σ be an alphabet and let $w \in \Sigma^*$ be p-free. Then w is a factor of some maximal p-free word over Σ.

Advances in problem solving:
In [11], Conjecture 1.1.03.12 was solved for a pattern of type $p = X^k$.

- **Conjecture 1.1.03.13:** Given an alphabet Σ and a pattern p, a maximal p-free word over Σ exists.

- **Conjecture 1.1.03.14:** Let Σ be an alphabet, $k \in [1, 2]$, and $w \in \Sigma^*$ be a k-power-free word. Then, in any case, w is a factor of some maximal k-power-free word over Σ.

–WORDS 2007:

Authors: Inna Mikhailova and Mikhail Volkov [6, 212–221].
The authors proved that every avoidable pattern can be actually avoided by an infinite sequence of palindromes over a fixed alphabet.

- **Question 1.1.07.1:** Is it possible to avoid an arbitrary pattern p by an infinite sequence of palindromes over each alphabet on which p is avoidable?

–WORDS 2011:

Authors: Helena Petrova and Arseny Shur [5, 168–178].

With respect to the prefix (suffix) order, any repetition-free language can be viewed as a poset whose diagram is a tree, each node generating a subtree and being a common prefix (suffix) of its descendants. The authors asked the three following questions:

- **Question 1.1.11.1:** Does a given word generate a finite or infinite subtree?

Advances in problem solving:

- In the case of a single word, in [11] it is shown that for all k-th power-free languages, the subtree generated by any word has at least one leaf.

- It has been shown in [20] that Question 1.1.11.1 is decidable for some power-free languages.

- **Question 1.1.11.2:** Are the subtrees generated by two given words isomorphic?

Actually, the authors proved that, in the langage of cube-free words, arbitrarily large finite subtrees may be generated.

- **Question 1.1.11.3 ([2, Problem 1.10.9] generalized to arbitrary words):** Can words generate arbitrarily large finite subtrees?

–WORDS 2013:

Authors: Tero Harju, Mike Müller [37, 154–160], [40, 29–38].

Let u_0, u_1 be two words over an alphabet A, and let $\beta \in \{0, 1\}^*$, called the *conducting sequence*, such that $|\beta| = |u_0| + |u_1|$, and such that the number of occurrences of the letter $i \in \{0, 1\}$ in β is the length of the word u_i: $|\beta|_i = |u_i|$. While forming the shuffle $w = u_0 \sqcup\!\sqcup_\beta u_1$ at step i ($i \in [1, |u_0| + |u_1|]$), the sequence β will choose the first unused letter from u_0 if $\beta(i) = 0$, or the first unused letter from u_1 if $\beta(i) = 1$ that is, the ith letter of w is $w(i) = u_{\beta(i)}(j)$, where $j = \mathrm{card}\{k \in [1, i] \mid \beta(k) = \beta(i)\}$ ($1 \leq k \leq i$). This definition can be extended to infinite words in a natural way (one requires that β contains infinitely many occurrences of both 0 and 1).

The authors proved that a ternary infinite square-free word exists, in such a way that it can be shuffled with itself to produce an infinite square-free word. They asked for the following questions:

- **Question 1.1.13.1:** Which square-free words u can be shuffled to obtain a square-free word $u \sqcup\!\sqcup_\beta u$?

- **Question 1.1.13.2:** Which words u can be shuffled to obtain a unique square-free word $u \sqcup\!\sqcup_\beta u$?
- **Question 1.1.13.3:** Which words w can be obtained in more than one way from a single word u using different conducting sequences?
- **Question 1.1.13.4:** Which square-free words w are themselves shuffles of square-free words: $w = u \sqcup\!\sqcup u$?
- **Question 1.1.13.5 (due to I. Petrykiewicz):** For any infinite ternary square-free word u, is there an infinite ternary square-free word w such that $u = u \sqcup\!\sqcup_\beta w$ for some infinite β?
- **Question 1.1.13.6:** Is there an infinite square-free word w such that $w = w \sqcup\!\sqcup_\beta w$ for some infinite β?
- **Question 1.1.13.7:** For each $n \geq 3$, is there a square-free word u of length n such that $u \sqcup\!\sqcup_\beta u$ is square free for some β?

–WORDS 2015:
Authors: Helena Petrova and Arseny Shur [52, 223–236].
As mentionned above, the set of square-free words over a given alphabet may be represented by a prefix tree T whose nodes are these square-free words. At WORDS 2015 the authors stated the following conjecture:

- **Conjecture 1.1.15.1:** In the tree T, the size of any minimal subtree of index n is $O(\log n)$.

1.2 The repetition threshold

The *repetition threshold* for k letters, commonly denoted by $RT(k)$, is the smallest rational number α such that there exists an infinite word whose finite factors have exponent at most α. For instance, every power in the Thue-Morse sequence has exponent at most 2, thus we have $RT(2) = 2$.
In the seventies, Françoise Dejean conjectured that, for every $k > 2$, the following holds:

$$RT(k) = \begin{cases} 7/4 & \text{if } k = 3 \\ 7/5 & \text{if } k = 4 \\ k/k-1 & \text{otherwise.} \end{cases}$$

Dejean's conjecture was partially solved by different authors. The final proof was completed in 2009 by James Currie and Narad Rampersad, for $15 \leq n \leq 26$, and independently by Michaël Rao, for $8 \leq k \leq 38$ [17, 3010–3018].

–WORDS 2005:

Author: Pascal Ochem [16, 388–392].
A word is α-free (resp. α^+-free) if it contains no factor that is an α'-power, for any rational $\alpha' \geq \alpha$ ($\alpha' > \alpha$).

- **Question 1.2.05.1 (stronger version of Dejean's conjecture):**
  - For every $k \geq 5$, there is an infinite $(k/k-1)^+$-free word over k letters with letter frequency $1/k+1$.
  - For every $k \geq 6$, there is an infinite $(k/k-1)^+$-free word over k-letter with letter frequency $1/k-1$.

Advances in problem solving:
- A partial solution, for $9 \leq k \leq 38$, was given by Michaël Rao [17, 3010–3018].
- The conjecture has been completely solved by Rao (private communication at WORDS 2015, see also [66]).

–WORDS 2011:

Authors: Golnaz Badkobeh and Maxime Crochemore [4, 37–43].
Starting with $RT(k)$, the definition of $FRT(k)$, the *finite repetition threshold* for k letters, stipulates that only a finite number of factors with exponent α may exist in the corresponding infinite word. In 2008, Jeffrey Shallit proved that $FRT(2) = 7/3$. In their presentation of WORDS 2011, Golnaz Badkobeh and Maxime Crochemore proved that $FRT(3) = RT(3) = 7/4$.

- **Conjecture 1.2.11.1:** We have $FRT(4) = RT(4) = 7/5$.

Advances in problem solving:
Conjecture 1.2.11.1 has been solved by Golnaz Badkobeh, Maxime Crochemore and Michaël Rao. In addition they proved that $FRT(k) = RT(k)$, for $k \leq 6$ (private communication at WORDS 2015).

1.3 On the number of different squares in a finite word

A natural question consists in examining the number of patterns that may appear in a finite word. From this point of view, Aviezri S. Fraenkel and Jamie Simpson focused on *dictinct* squares, defined as patterns of different shapes (not just translated of each other). At WORDS 1997, in the case of the sequence of Fibonacci words $(f_n)_{n\geq 0}$, they showed that the exact number of such squares is $2(f_{n-2}-1)$, for any integer $n \geq 5$ [58, 95–106]. In [27] they proved that the number of distinct squares in an arbitrary word of length n is bounded by $2n$.

–WORDS 2005:

Author: Lucian Illie [16, 373–376].

With regards to the number of distinct squares in a word, the author provided a refinement of $2n - O(\log n)$; in addition, he recalled the following conjecture:

- **Conjecture 1.3.05.1 (Square conjecture, due to A.S. Fraenkel and J. Simpson, [27]):** The number of different squares in a word of length n is bounded by n.

Advances in problem solving

- In the case of a binary alphabet in [36], the authors stated a stronger conjecture regarding the number of distinct squares in a binary word: the number of distinct squares is upper bounded by $\frac{2k-1}{2k+2}n$, where k is the least of the number of occurrences of each letter, the bound being tight.

- The best bound known so far is $\frac{11n}{6}$ [21].

–WORDS 2015:

Authors: Florin Manea and Shinnosuke Seki [52, 160–169].

Given a word w, define its *square density* by:

$\rho_{\text{sq}}(w) = |w|^{-1} \cdot \text{card}\{x^2 \in \Sigma^+ | x^2 \text{is a factor of } w\}$.

In their contribution to WORDS 2015, the authors proved that binary words have the largest square density; moreover, they asked the question of constructing a “square-density" amplifier:

- **Question 1.3.15.1:** Can we compute a mapping $f : \Sigma^* \longrightarrow \Sigma^*$ for which a constant $c > 1$ exists such that, for all $w \in \Sigma^*$, if $\rho_{\text{sq}}(w) \geq 1$ then we have $\rho_{\text{sq}}(f(w)) \geq c\rho_{\text{sq}}(w)$?

1.4 The “runs" conjecture

A *run* may be defined as some occurrence of a repetition of exponent at least 2 that is maximal, in the sense where it cannot be extended from left or right to obtain the same type pattern. Such objects play an important role in a lot of string matching algorithms.

–WORDS 2009:

Authors: Maxime Crochemore, Lucian Ilie and Liviu Tinta [17, 2931–2941].

These authors showed that, given a word of length n, the number of its runs is not greater than $1.029n$. This is a noticeable step in the proof of the so-called “runs" conjecture:

- **Conjecture 1.4.09.1 ("runs" conjecture, due to Kolpakov and Kucherov, [42]):** For a binary alphabet, given word of length n the number of its runs is bounded by n.

–WORDS 2015:
Authors: Štěpán Holub [53, 43–52].
Denote by $\rho(n)$ the maximal number of runs in a (binary) word of length n. The concept of "lost positions" is a recently introduced tool for counting the number of runs in binary words. By investigating the frequency of lost positions [35, 277–286] in prefixes of words, and by making use of an extensive computer search, the author proved that the asymptotic density of runs in binary words is less than $183/193 \approx 0.9482$; in addition, he formulated the following conjecture:

- **Conjecture 1.4.15.1:** The asymptotic upper bound of $\rho(n)/n$ is never reached.

1.5 The prefix-suffix square completion

–WORDS 2015:
Authors: Marius Dumitran and Florin Manea [52, 147–159].
The so-called *suffix-square duplication* allows to derive, from a word w, any word wx such that x is a suffix of w. The *suffix-square* completion, in turn, derives from w a word wx such that w has a suffix of type yxy. *Prefix-square* duplication (completion) may be defined in a similar way. In their talk at WORDS 2015, Marius Dumitran and Florin Manea made use of such operations for generating an infinite word that does not contain any repetition of exponent greater than 2. With regards to combinatorics properties of words, they asked the following questions:

- **Question 1.5.15.1:** What is the minimum exponent of a repetion which is avoidable by an infinite word constructed by iterated (prefix)-suffix duplication?

- **Question 1.5.15.2:** By applying prefix-suffix completion, can we construct words that avoid cubes, and every word containing squares?

- **Question 1.5.15.3:** Starting with a single word, does the language of finite words constructed by iterating prefix-suffix square completion remains semi-linear?

- **Question 1.5.15.4:** Draw studies of languages of finite words which are constructed by iterating prefix-suffix square completion, starting with special sets of initial words such as singleton sets, finite sets, regular sets, etc.

- **Question 1.5.15.5:** What is the minimum number of steps of square completion that are required to obtain a word from one of its factors?

1.6 Abelian patterns

An *abelian square* consists in a pattern which is obtained by applying a permutation on the letters of a square. Clearly, with every pattern, a corresponding abelian one can be associated.
In 1992, by constructing an *abelian square free* word over a four-letter alphabet, Veikko Keränen solved a famous open problem that was initially formulated by Erdös in 1961 [26, 41]. At WORDS 2007, he presented new abelian square-free morphisms and a powerful substitution over 4 letters [6, 190–200].

–WORDS 2003:

Author: James Currie [30, 7–18].

- **Question 1.6.03.1:** Which of the following patterns are avoidable in the abelian sense?
  01020312, 01020321, 01021303, 01023031, 010203013, 010213020.

- **Conjecture 1.6.03.2:** The number of abelian cube-free ternary words grows exponentially with length.

Given a n-letter alphabet, define the sequence Z_n recursively by: $Z_1 = 1$ and $Z_n = Z_{n-1} n Z_{n-1}$, for every $n > 1$.

- **Conjecture 1.6.03.3:** Let p be any pattern over a n-letter alphabet. Then p is abelian avoidable iff Z_n is p-free in the abelian sense.

- **Question 1.6.03.4:** Given pattern p and integer n, what is the complexity of deciding whether Z_n encounters p in the abelian sense?

Define respectively the *abelian repetitive threshold function*, and the *dual abelian repetitive threshold function* on $(1, 2]$ by:
$ART(n) = \inf\{s : y^s$ is avoidable on n letters in the abelian sense$\}$
$DART(r) = \min\{n \in \mathbb{N} : y^r$ is avoidable in the abelian sense on a n-letter alphabet$\}$.

- **Question 1.6.03.5:** What are the values of $ART(n)$ and $DART(r)$?

–WORDS 2013:

Two papers were concerned by open questions:

Authors: Mari Huova and Aleksi Saarela [37, 161–168].
Two words u, v are *k-abelian equivalent* if, for any string of length at most k, this word occurs as a factor in u as many times as in v. A word is a *strongly k-abelian nth-power* if it is k-abelian equivalent to some nth-power. In their contribution to WORDS 2013, the authors proved that strongly k-abelian nth-powers are unavoidable on any alphabet, moreover they formulated the following questions:

- **Question 1.6.13.1:** How many k-abelian equivalence classes of words of a given length contain an nth power?

- **Question 1.6.13.2:** How many words of a given length are strongly k-abelian nth powers?

- **Question 1.6.13.3:** What is the length of the longest word avoiding strongly k-abelian nth powers?

- **Question 1.6.13.4:** How many words avoid strongly k-abelian nth powers?

- **Question 1.6.13.5:** How many words of a given length contain a strongly k-abelian nth power?

- **Question 1.6.13.6:** How many words of a given length are strongly k-abelian nth powers?

Author: Michaël Rao [40, 39–46].
Given an integer $n \geq 2$, a word u is a k-abelian-n-power if we have $u = u_1 u_2 \cdots u_n$, where u_i and u_{i+1} are k-abelian equivalents for every $i \in \{1, \cdots n-1\}$.

- **Question 1.6.13.7:** Is there a pure morphic binary word avoiding 2-abelian cubes?

- **Question 1.6.13.8:** Can we avoid abelian-cubes of the form uvw, with $|u| = |v| = |w| \geq 2$, over a binary alphabet?

- **Question 1.6.13.9:** Is there a natural integer k such that 2-abelian-squares of period at least k can be avoided over a binary alphabet?

- **Question 1.6.13.10:** Is there a natural integer k such that abelian cubes of period at least k can be avoided over a binary alphabet?

The so-called additive powers consist in a generalization of abelian powers: given an alphabet $\Sigma \subseteq N$, an *additive kth power* is a word $p_1 \cdots p_k \in \Sigma^*$ such that $|p_1| = \cdots = |p_k|$ and $\sum(p_1) = \cdots = \sum(p_k)$, where $\sum(p_i)$ stands for the sum of the digits of the word p_i ($1 \leq i \leq k$). In 2011, Cassaigne, Currie, Schaeffer and Shallit proved that additives cubes are avoidable on $\{0, 1, 2, 3, 4\}$ [18]. At WORDS 2013, Rao asked the following question:

- **Question 1.6.13.11:** Are there infinite additive-cube-free words on the following alphabets: $\{0, 1, 2, 3\}$, $\{0, 1, 4\}$ and $\{0, 2, 5\}$?

–WORDS 2015:

Open questions were asked in the following contributions:

Authors: Gabriele Fici, Filippo Mignosi [52, 122–134], Gabriele Fici, Filippo Mignosi and Jeffrey Shallit [53, 29–42].
The authors focused of the maximum number of abelian squares that a word may contain. Actually, a word of length n which contains $O(n^2)$ distinct abelian squares exists [43]. At Words 2015, the authors stated the following conjectures:

- **Conjecture 1.6.15.1:** Assume that a word with length n, and containing k many distinct abelian-square factors, exists. Then a binary word of length n containing at least k many distinct abelian-square factors exists.

Two abelian squares are *inequivalent* if their Parikh vectors are different [28].

- **Conjecture 1.6.15.2 (due to T. Kosciumaka, J. Radoszewski, W. Rytter and T. Waleń [43]):** A word of length n contains $\Theta(n\sqrt{n})$ inequivalent abelian-squares.

Author: Michaël Rao [66].
Erdös formulated two fundamental problems:
(1) (1957,1961): Is there arbitrarily long abelian-square-free words over a finite alphabet?
(2) (1961): Is it possible to avoid long squares over a binary alphabet?

R.C. Entringer, D.E. Jackson and J.A. Schatz gave a positive answer to the second question [25]. In 2002 Mäkelä put similar questions with regards to the abelian squares or cubes over binary or ternary alphabets [51]. In his talk at WORDS 2015, Rao presented technics for deciding whether a morphic word avoids abelian and k-abelian repetitions: in particular, this allowed him to prove that long abelian squares are avoidable over a ternary alphabet. Then he asked the following questions:

- **Question 1.6.15.3:** Can we avoid long abelian cubes over two letters?

- **Question 1.6.15.4:** How to decide whether a morphic word avoids (long) abelian power?

- **Question 1.6.15.5 (due to S. Mäkelä):** Let h be the morphism onto $\{0, 1, 3, 4\}^*$ defined by $h(0) = 03$, $h(1) = 43$, $h(3) = 1$, $h(4) = 01$. Is there a morphism $g : \{0, 1, 3, 4\}^* \longrightarrow \{0, 1\}^*$ such that $g(h^\infty(0))$ has no long abelian cubes?

- **Question 1.6.15.6:** Find good heuristics to compute candidates for question 1.6.15.5.

- **Question 1.6.15.7:** Find a morphism avoiding abelian square on four letters which should be simpler than that of Keränen ?

- **Question 1.6.15.8:** What is the minimum k such that abelian squares of period at least k over three letters can be avoided ($2 < k < 6$)?

- **Question 1.6.15.9:** What is the minimum k such that 2-abelian squares of period at least k over two letters can be avoided ($2 < k < 60$)?

With regards to the so-called notion of *templates*, we refer the reader to [1]. From this point of view, Rao and M. Rosenfeld proved that it is possible to decide whether $h^\infty(a)$ realizes t, for any primitive morphism h whose matrix has no eigenvalue of norm 1, and for any template t. They formulated the following problems:

- **Question 1.6.15.10:** Is there a morphism over 5 letters with two eigenvalues of norm smaller than 1 and an abelian-square-free fixed point?

- **Question 1.6.15.11:** Is there a morphism on 3 letters with one eigenvalue of norm smaller than 1, and an abelian-cube-free fixed point?

- **Question 1.6.15.12:** How to decide whether eigenvalues of norm 1 may be allowed in the result that was mentioned above?

2 Complexity issues

In the literature, with a word several notions of complexity can be associated, the most famous one being the *factor complexity*: given a word w, this measure counts the number $p_w(n)$ of different factors of length n occuring in w. The famous characterization of Morse-Hedlund for ultimately periodic words led to introduce the infinite *Surmian words*, whose complexity is $p_w(n) = n + 1$, the best known example of them being certainly the famous *Fibonacci word* [54, 55].

2.1 The recurrence quotient

The *recurrence function* has been introduced by M. Morse and G.A. Hedlund [55]: given a factor u, with every non-negative integer n it associates the size $R_u(n)$ of the smallest window that contains every factor of length n of u.

–WORDS 1997:

Author: Julien Cassaigne [58, 3-12].
The *recurrence quotient* is defined as $\rho(u) = \limsup_{n\to\infty} \frac{R_u(n)}{n}$.
For a sturmian sequence of slope α, denote the recurrence quotient by $\rho(\alpha)$; the spectrum of values of ρ is the set S of the values taken by $\rho(\alpha)$ when α spans $[0,1] \setminus \mathbb{Q}$.

- **Question 2.1.97.1:** What is the Hausdorff dimension (see e.g. [31]) of S (or that of each of its intervals $S \cap [a, a+1]$)?

- **Question 2.1.97.2:** Draw a study of the recurrence quotients for other families of infinite words than sturmian words, such as words of complexity $2n + 1$, or infinite words in general.

2.2 The ratio $p(n)/n$

Alex Heinis proved that if $p(n)/n$ has a limit, then this limit is either equal to 1, or highter than and equal to 2 [32, 33]).

–WORDS 2001:

Author: Ali Aberkane [60, 31-46].
By using the so-called Rauzy graphs, at WORDS 2001 the author presented characterizations of the words such that the limit is 1.

- **Question 2.2.01.1:** Transform the preceding characterization into another one which makes use of a finite set of substitutions associated with rules governing their composition (i.e. S-adic system of representation).

- **Question 2.2.01.2:** Give a characterization of infine words whose complexity satisfies $\lim_n p(n)/n = 2$.

2.3 The balance function

–WORDS 2001:

Author: Boris Adamczewski [60, 47-75].
Boris Adamczewski defines the *balance function* by $\max_{a\in A} \max_{u,v\in F(w)}\{||u|_a - |v|_a|\}$.

With regards to the so-called *primitive substitutions*, the author investigated the connections between the asymptotic behavior of the balance function and the incidence matrices of such substitutions. Moreover, he showed that the Thue-Morse sequence is an example for which, the spectrum of the substitutions of order two is different of the spectrum of the initial substitutions.

- **Question 2.3.01:** Give an example of sequence for which the mentionned change of spectrum is really significant for the balance properties.

–WORDS 2013:

Author: Julien Cassaigne [37, 1–2].
A words is *balanced* if, for any pairs (u, v) of its factors with same length, and for any letter a, we have $||u|_a - |v|_a| \leq 1$ (where $|u|_a$ stands for the number of occurrences of the letter a in u). A classical characterization of Sturmian words is that they are the aperiodic 1-balanced sequences. For Arnoux-Rauzy words [7], whose complexity is $(|A| - 1)n + 1$, the following question can be asked (see also [12]):

- **Question 2.3.13:** Give characterizations of Arnoux-Rauzy words with a given balance.

2.4 Palindromic complexity, palindromic defect

The *palindromic complexity* of an (in)finite word is the function which counts the number $P(n)$ of different palindromes of length n that occur as factors of this word. Given a finite word w of length n, we have $P(n) \leq n + 1$ [22]: this leads to define the corresponding *palindromic defect* as $D(w) = n + 1 - P(n)$. In the case of an infinite word $\mathbf{u}$, set $D(\mathbf{u}) = \sup\{D(w)|w \in F(\mathbf{u})\}$.

–WORDS 2005:

Authors: Peter Baláži, Zuzana Maskóvá and Edita Pelantovà [16, 266-275].
The authors provide an estimate of $P(n)$ for uniformly recurrent words; denoting by $p(n)$ the classical factor complexity, this estimation is based on the equation: $P(n) + P(n + 1) = p(n + 1) - p(n) + 2$.

- **Question 2.4.05:** Describe the structure of the Rauzy graphs of words reaching the mentioned supremum.

–WORDS 2017:

Authors: Edita Pelantovà and Štěpán Starosta [14, 59–71].
A morphism ψ is of *Class P* if we have $\psi(a) = pp_a$ for any letter a, where p, p_a are

both (possibly empty) palindromes; morphisms of Class P' are defined as being conjugate of morphisms of Class P. Recall that, given a morphism, with fixed point $\mathbf{u}$, their common corresponding *language* consists in the set of all finite factors of $\mathbf{u}$. The main motivation for studying the preceding morphisms lays upon the following conjecture:

- **Conjecture 2.4.17.1 (Zero defect conjecture, due to Blondin-Massé, Brlek, Garon and Labbé, [13]):** Let $\mathbf{u}$ be an aperiodic fixed point of a primitive morphism whose language is closed under reversal. Then either we have $D(\mathbf{u}) = 0$ or we have $D(\mathbf{u}) = +\infty$.

A counterexample was given in [10]; however, in [46] the authors proved that the conjecture is true for some special class of the so-called *marked* morphism, which were defined as follows:

Given a morphism h, it is marked if two morphisms h_1, h_2 exist, both being conjugate to h, such that the set of the first (last) letters of the images of letters by h_1 (h_2) is the whole alphabet.

With regards to the general case, the authors suggest that a refinement of Conjecture 2.4.17.1 could be valid.

Words with zero palindromic defect are usually called *rich words*: with regards to this notion, some open problems were put in the presentation:

- **Question 2.4.17.2:** What is the number of rich words of a given length?

- **Question 2.4.17.3:** Can we decide whether two rich words are factors of a common rich word?

- **Conjecture 2.4.17.4 (Class P conjecture, due to A. Hof, O. Knill and B. Simon, [34]):** Let $\mathbf{u}$ be a fixed point of a primitive morphism. If $\mathbf{u}$ has infinitely many palindomic factors ($\mathbf{u}$ is *palindromic*, for short), then a morphism of class P' exists, whose fixed point has the same language as the word $\mathbf{u}$.

Advances in problems solving:
- Conjecture 2.4.17.4 was solved in the binary case ([67]).
- In [44], the authors proved that a ternary word $\mathbf{w}$ exists such that, it is a palindromic fixed point of a primitive morphism, although it is neither fixed by any morphism of class P'.
- The conjecture has been confirmed for the so-called marked morphisms [45].
- In [56], the conjecture has been confirmed for morphisms fixing a coding of a non-degenerate exchange of 3 intervals.

- **Conjecture 2.4.17.5:** Let $\mathbf{u}$ be a fixed point of a primitive morphism. Then the language of $\mathbf{u}$ is closed under reversal if and only if $\mathbf{u}$ is palindromic.

Advances in problems solving:
Conjecture 2.4.17.5 is true for marked morphisms [45].

2.5 Sets of sequences of a given complexity

The famous Arnoux-Rauzy words consist in a generalization of Sturmian sequences on a three-letter alphabet: they are in fact those of infinite sequences of complexity $2n+1$ that satisfy the following condition: exactly one left and one right factor exist for each length [7]. For any letter frequency, sequences of factor complexity $2n+1$ can be constructed by making use of coding some 3-interval exchange transformation. As shown in [68], such sequences are unbalanced and the question of finding balanced ternary sequences of complexity $2n+1$ for all letter frequency remains open.

–WORDS 2017:
Authors: Julien Cassaigne, Sébastien Labbé, and Julien Leroy [14, 144–156].
In 2015, based on the structure of Arnoux-Rauzy graphs, Julien Cassaigne introduced on $\mathbb{R}^3_{\geq 0}$ a bidimensional continued fraction algorithm, such as:
$F_C(x_1, x_2, x_3) = (x_1 - x_3, x_3, x_2)$ if $x_1 \geq x_3$
$F_C(x_1, x_2, x_3) = (x_2, x_1, x_3 - x_1)$ otherwise.
Some important properties of F_C were described at WORDS 2017: in particular the associated substitutions lead to obtain S-adic words with complexity $2n+1$ (S stands for a set of morphisms).

- **Question 2.5.17:** Can we find an analogue of F_C in dimension $d \geq 4$, generating S-adic sequences with complexity $(d-1)n+1$ for almost every vector of letter frequencies?

2.6 Abelian complexity

Let $A = \{a_1, \cdots, a_q\}$ be an alphabet and $w \in A^*$. Recall that the Parikh vector of w is $\psi(w) = (|w|_{a_1}, \cdots, |w|_{a_q})$ [63]. Given an infinite word $\mathbf{u}$, denote by $\Psi_n(\mathbf{u})$ the set of such vectors for the factors of length n of $\mathbf{u}$. The abelian complexity of $\mathbf{u}$ is the application onto $\mathbb{N}$ that is defined by $\rho^{ab}(n) = |\Psi_n(\mathbf{u})|$. Denote by $|u|_v$ the number of occurrences of a given word v as a factor of u. Given a positive integer ℓ, two words x and y are *ℓ-abelian equivalent* if $|x|_v = |y|_v$ for all words v of length $|v| \leq \ell$.

–WORDS 2017:
Authors: Idrissa Kaboré, Boukaré Kientéga [14, 132–143].
The so-called *ternary Thue-Morse word* is the infinite word $\mathbf{t_3}$ which is generated by the morphism μ_3 defined by $\mu_3(0) = 012$, $\mu_3(1) = 120$, $\mu_3(2) = 201$. The authors studied some properties of this words; in particular they proved that $\mathbf{t_3}$ satisfies the following conjecture:

- **Conjecture 2.6.17 (due to A. Parreau, M. Rigo, E. Rowland and E. Vandomme [64]):** Any k-automatic word admits a ℓ-abelian complexity function which is k-automatic.

3 Factorization of words. Equations

Some further important information can be obtained by decomposing a word into a convenient sequence of consecutive factors: $w = w_1 \cdots w_n$.

3.1 $\mathcal{F}$-factorization

The so-called $\mathcal{F}$-factorization has been introduced in [38]; it corresponds to the case where the preceding sequence $(w_1, \cdots, w_n)$ satisfies some given property $\mathcal{F}$, which is formally defined as follows:
Let $I = \{1, \cdots, k\}$ and Σ be two disjoint alphabets. Set $\mathcal{F} = (L, L_1, \cdots, L_k)$, with $L \subseteq I^*$ and $L_1, \cdots, L_k \subseteq \Sigma^*$. We say that the sequence of factors $(w_i, \cdots, w_n)$ is a $\mathcal{F}$-factorization if for all $j \in [1, n]$ we have $w_j \in L_{i_j}$ and $i_1 \cdots i_n \in L$. The factorization $\mathcal{F}$ is *regular* (*context-free*) if the languages $L, L_1, \cdots, L_k$ are regular (context-free).

–WORDS 1997:
Authors: Juhani Karhumäki, Wojciech Plandowski and Wojciech Rytter [58, 123–133].
Three fundamental properties of $\mathcal{F}$-factorizations were examined, namely *completeness*, *uniqueness* and *synchronization*.

- **Question 3.1.97.1:** Find efficient algorithms for the polynomial time solvable problems which were discussed in the paper.

- **Question 3.1.97.2:** Given a word, can its minimal and maximal regular $\mathcal{F}$-factorization (in the sense of the length of the sequence of indices) be found in polynomial time?

- **Question 3.1.97.3:** Could better algorithms be designed for the problems discussed in the paper if, in regular $\mathcal{F}$-factorizations, only finite languages are considered?

- **Question 3.1.97.4:** Is the completeness or the uniqueness undecidable, when context-free $\mathcal{F}$-factorizations are given by deterministic automata or by linear context-free grammars?

- **Question 3.1.97.5:** What is the complexity of the problem of determining whether a regular $\mathcal{F}$-factorization possesses synchronization property, when the parameters of the synchronization are not given? What about this problem for context-free $\mathcal{F}$-factorizations?

3.2 Periodicity

With the preceding notation, if for an integer $n \geq 2$, all the words $w_1, \cdots, w_{n-1}$ are equal, the word w_n being one of their prefixes, we say that the length of w_1 is a *period* of w.

–WORDS 2007:

Author: Kalle Saari [6, 273-279].
The author proved that the least period of a non-empty factor of the infinite Fibonacci word is a Fibonacci number. With regards to Sturmian words of a given slope, say α, the set $\Pi(\alpha)$ is defined as follows:

Let $[d_0 = 0, d_1 = 1, d_2, d_3, \cdots]$ the continued fraction expansion of α. Set $q_1 = q_0 = 1, q_{n+1} = d_{n+1}q_n + q_{n-1}$ $(n \geq 1)$ and :

$$\Pi(\alpha) = \bigcup_{n\geq 0}\{iq_n + q_{n-1} : i = 0, 1, \cdots, d_n\}$$

- **Conjecture 3.2.07:** Let t denote a Sturmian word with slope α. If a word is a nonempty factor of t, then its least period belongs to $\Pi(\alpha)$.

3.3 Quasiperiodicity

A word w is *quasiperiodic* if another word x exists such that any position in w falls within an occurrence of x as a factor of w (informally, w may be completely "covered" by some set of occurrences of the factor x).

–WORDS 2013:

Authors: Florence Levé and Gwenaël Richomme [37, 181–192].
A morphism is *strongly (resp. weakly) quasiperiodic* if it maps any (at least one)

non-quasiperiodic word to some quasiperiodic word. The authors provided algorithms for deciding whether a morphism is strongly quasiperiodic on finite and infinite words; in addition, they put the following questions:

- **Question 3.3.13.1:** Given a morphism f and a letter a such that a is the initial letter of $f(a)$, is it decidable that $f^{\omega}(a)$ is quasiperiodic?

- **Conjecture 3.3.13.2:** Let f be an morphism generating a quasiperiodic infinite word. If $f(a)$ is not a power of a then f is weakly quasiperiodic on any infinite word.

The authors define *weakly quasiperiodic morphisms* as those that map at least one non-quasiperiodic word to a quasiperiodic one (some partial answers are given in the paper).

- **Question 3.3.13.3:** Can we decide whether given a morphism, it is weakly quasiperiodic on finite (infinite) words?

3.4 Defect effect and independent systems of equations

The *combinatorial rank* of a set of words X, that we denote by $r(X)$, is the smallest number of words needed to express all strings of X as products of those words [57]. As a direct consequence of the famous theorem of defect [47, 50, 23, 48], if X is not a code (that is, if the words of X satisfy a nontrivial equation) then we have $r(X) \leq |X| - 1$.

–WORDS 1999:

Authors: Juhani Karhumäki and Ján Maňuch [59, 81–97].
Unformally, the so-called *X-factorization* of a bi-infinite word w consists in any sequence of words from X yielding w as their product. The authors stated the three following problems, which are connected to the famous critical factorization theorem [48, Chap. 8]:

- **Question 3.4.99.1:** Let X be a finite set of words, and w be a non-periodic bi-infinite word. Assume that w possesses k disjoint X-factorizations, with $k \leq |X|$. Is it true that we have $r(X) \leq |X| - k + 1$?

- **Question 3.4.99.2:** Let X be a code, and let w be a bi-infinite word. Assume that for $k \leq |X|$, w possesses k disjoint X-factorizations, such that at least one of them is non-periodic. Is it true that we have $r(X) \leq |X| - k + 1$?

- **Question 3.4.99.3:** Denote by $p(w)$ the smallest period of an word $w \in \Sigma^+$. Let $X \subseteq \Sigma^+$ satisfying $p(x) < p(w)$ for all $x \in X$. Is it true that w has at most $|X| + 1 - r(X)$ disjoint X-factorizations?

–WORDS 2001:
Authors: Tero Harju and Dirk Nowotka [60, 139–172].
Defect effect is strongly connected to *independent systems of equations*. Given an equation in three variables, say x, y, z, a solution α is *non-periodic* if $\alpha(x), \alpha(y), \alpha(z)$ are not powers of the same word [48, Chapt. 9]. A system of equations is *independent* if it is not equivalent to any of its proper subsets. An equation is *balanced* if the number of occurrences of each variable on the left- and the right-hand side is the same. In their presentation at WORDS 2001, the authors proved that every independent system of equations in three variables, with at least two equations and a non-periodic solution, actually consists in a balanced equation. They asked the following question, which was actually implicitely raised in 1983 by Culik II and Karhumäki [19]:

- **Question 3.4.01:** Does an independent system of three equations in three variables with a non-period solution exists?

–WORDS 2005:
Authors: Štěpán Holub and Juha Kortelainen [16, 363–372].
The authors studied the infinite system (S) of words equations :

$$\{x_0 u_1^i x_1 u_2^i x_2 \cdots u_m^i x_m = y_0 v_1^i y_1 v_2^i y_2 \cdots v_n^i y_n \ : i \geq 0\}$$

They stated the following questions:

- **Question 3.4.05.1:** Is there a positive integer k such that the system (S) is equivalent to one of its subsystems induced by k equations?
- **Question 3.4.05.2:** Is the system $\{u_1^i = v_1^i v_2^i \cdots v_n^i \ : i \geq 0\}$ equivalent to one of its subsystems induced by three equations?

3.5 The Post Correspondence Problem

The famous Post Correspondence Problem (*PCP* for short) consists in asking, given two morphisms h, g, whether or not the equation $h(x) = g(x)$ has a solution distinct of the empty word.
In the most general case, it is well known that this problem is undecidable [65]. In another hand, many studies were devoted to special cases of instances (eg. [24]).

–WORDS 2005:
Authors: Vesa Halava, Tero Harju, Juhani Karhumäki and Michel Latteux [16, 355-352].

The authors start from the following definitions: a morphism h is *marked* if for any pair of different letters a, b, the initial letters of the words $h(a)$ and $h(b)$ are different; two words u, v are *comparable* (denoted by $u \bowtie v$) is either u is a prefix of v, or v is a prefix of u. With such notions, special types of instances (h, g) may be defined: in particular (h, g) is called a *unique equality continuation instance* if, for any word u and any pair of different letters a, b, both the two conditions $h(ua) \bowtie g(ua)$ and $h(ub) \bowtie g(ub)$ imply $h(u) = g(u)$,

At WORDS 2005, the authors put the two following questions:

- **Question 3.5.05.1:** Is PCP decidable for unique equality continuation instances?

- **Question 3.5.05.2:** Is it decidable whether or not an instance of PCP satisfies the property of unique equality continuation instances?

3.6 The Palindromic Length

The so-called *palindromic length* of a word x is defined as the smallest number n such that x can be written as the concatenation of n palindromes.

–WORDS 2017:

Author: Aleksi Saarela [14, 203–213].
At WORDS 2017, the author firstly reminded the following conjecture:

- **Conjecture 3.6.17.1 (due to Frid, Puzynina and Zamboni, [29, p. 738]):** Every aperiodic infinite word has factors of arbitrarily high palindromic length.

Then, he proved that this conjecture is in fact equivalent to the following one:

- **Conjecture 3.6.17.1a:** Every aperiodic infinite word has prefixes of arbitrarily high palindromic length.

Next, he put the two following questions:

- **Question 3.6.17.2:** Are there words such that all of their prefixes have palindromic length at most n, but some of their factors have palindromic length $2n$?

- **Question 3.6.17.3:** In the binary case, give an improvement of the result of Lemma 10 in the paper.

In a classical way, the free monoid Σ^* can be extended to a free group, namely $(\Sigma \cup \Sigma^{-1})^*$. This leads to introduce the so-called *FG-palindromes* and the *FG-palindromic length* of a word. As an example the palindromic length of the word $abca$ is 4, however this word is the product of three FG-palindromes: $abca = aba \cdot a^{-2} \cdot aca$. Actually, Aleksi Saarela proved that the FG-palindromic length of a word can be much smaller that its palindromic length itself. This led him to state the following questions:

- **Question 3.6.17.4:** Does an aperiodic infinite word exists such that the FG-palindromic lengths of its factors are bounded by a constant?

- **Question 3.6.17.5 (due to V.G. Bardakov, V. Shpilrain and V. Tolstykh [9, Problem 2, p. 576]):** Find an algorithm for computing the FG-palindromic length.

3.7 Permutation on Words

They are lots of manners to construct permutations onto A^*, the best-known of them being automorphism or anti-automorphisms.

–WORDS 2017:

Authors: Niccolò Castruonovo, Robert Cori and Sébastien Labbé [14, 240–251]. Let $A = \{a, b\}$, and $A_n = \{w \in A^* : |w|_a = n, |w|_b = n + 1\}$ ($|w|_a$ denotes the number of occurrences of the letter a in w). At WORDS 2017, the authors proved that any word in A^* may be factorized as $u_1bu_2b \cdots u_pbwav_qav_{q-1} \cdots av_1$, where w and u_i, v_j are Dyck words ($1 \le i \le p, 1 \le i \le q$). Let $j = |u_1bu_2b \cdots u_pb|$, and θ be the morphism that is generated by $\theta(a) = b$, $\theta(b) = a$. Consider the map Γ_n onto A_n which, with each word $w = w_1 \cdots w_{2n+1}$ ($w_i \in A$, $1 \le i \le 2n + 1$), associates the word $\theta(w_1 \cdots w_{j-1})b\theta(w_{j+1} \cdots w_{2n+1})$.

The permutation Γ_n has particularly interesting combinatorial properties. In particular it can be extended into a permutation of A^* itself, as indicated in the following:

With the preceding notation let $w = u_1bu_2b \cdots u_pbtav_qav_{q-1} \cdots av_1$. If $p > q$ ($q < p$) then call *pivot* each of the $p - q$ ($q - p$) occurrences of b appearing just after (before) the words $u_{p-q+1}, u_{p-q+2}, \cdots, u_p$ ($v_{q-p+1}, v_{q-p+2}, \cdots, v_q$); if $p = q$ there are no pivots. The word $\Gamma(w)$ is obtained by substituting $\theta(c)$ to c for each occurrence of $c \in \{a, b\}$ that is not a pivot.

- **Conjecture 3.7.17:** The cycles of Γ containing words of odd lengths are also of odd lengths. Those containing words of even lengths, with an odd number of occurrences of a, are also of even lengths. Those containing words of even lengths, with an even number of occurrences of a, may have either odd or even length.

References

[1] Aberkane A., James Currie and Narad Rampersad, The number of ternary words avoiding abelian cubes grows exponentially, *J. Integer Seq.* **7** (2004) Art. 04.2.7.

[2] Allouche J.-P., Shallit J., Automatic Sequences: Theory, Applications, Generalizations, Cambridge University Press (2003).

[3] Noga Alon, Joel H. Spencer, The Probabilistic Method, 4th ed., Wiley (2016).

[4] Petr Ambrož, Štěpán Holub and Zuzana Masáková (Eds), WORDS 2011, 8th International Conference on Words, Prague, Czech Republic, 12-16th September 2011, EPTCS 63, Proceedings, http://eptcs.web.cse.unsw.edu.au/content.cgi?WORDS2011.

[5] Petr Ambrož, Štěpán Holub and Zuzana Masáková (Eds.), Special Issue, WORDS 2011, *Int. J. of Found. of Comput. Sci.* **23**(8) (2012) 1579-1727.

[6] Pierre Arnoux, Nicolas Bédaride and Julien Cassaigne (Eds), WORDS 2007, 6th International Conference on Words, Marseilles, France, September 17-21, 2007 Proceedings, document distribuated to the attendees of the conference.

[7] Arnoux P., Rauzy G., Représentation géométrique de suites de complexité $2n + 1$, *Bull. Soc. Math. France* **119**(2) (1991) 199-215.

[8] Baker K., McNulty G. F. and W. Taylor, Growth problems for avoidable words, *Theoret. Comput. Sci.* **69**(3) (1989) 319–345.

[9] Bardakov V.G., Shpilrain V. and V. Tolstykh, On the palindromic and primitive widths of a free group, *J. Algebra* **285**(2) (2005) 574–585.

[10] Bašić B., On highly potential words, *Eur. J. Combin.* **34**(6) (2013) 1028–1039.

[11] Bean D.. R. , A. Ehrenfeucht A. and G. McNulty (1979): Avoidable patterns in strings of symbols, *Pacific J. Math.* **85** (1979) 261–294.

[12] Berthé V., Cassaigne J. and W. Steiner, Balance properties of Arnoux-Rauzy words, *Intern. J. Algebra Comput.* **23**(4) (2013) 689–704.

[13] Blondin Massé A., Brlek S., Garon A. and S. Labbé, Combinatorial properties of f-palindromes in the Thue-Morse sequence, *Pure Math. Appl.* **19** (2–3) (2008) 39–52.

[14] Brelk S., Dolce F., Reutenauer C. and E. Vandomme (Eds.) Combinatorics on Words, 11th International Conference, WORDS 2017, Montréal, QC, Canada, September 11-15, 2017, Proceedings, *Lect. N. in Comput. Sci.* **10432** (2017), 1–263.

[15] Srečko Brlek and Christophe Reutenauer (Eds.), WORDS 2005, 5th International Conference on WORDS, Publications du Laboratoire de Combinatoire et d'Informatique Mathématique, Université du Québec à Montréal 36, ISBN 2-89276-376-2 LaCIM Montréal (2007) 1–341.

[16] Srečko Brlek and Christophe Reutenauer (Eds.), Combinatorics on words, *Theoret. Comput. Sci.* **380**(3) (2007), 219–410.

[17] Arturo Carpi and Clelia De Felice (Eds.), Combinatorics on Words (WORDS 2009), 7th International Conference on Words, Fisciano, Italy, *Theoret. Comput. Sci.* **412**(27) (2011), 2909-3032.

[18] Cassaigne J., Currie J. D., Schaeffer L. and J. Shallit, Avoiding three consecutive blocks of the same size and same sum, *J. of the ACM* **61**(2) (2014) Article No. 10.

[19] Culik II K., Karhumäki J., Systems of equations over a free monoid and Ehrenfeucht's conjecture, *Discrete Math.* **43** (2-3) (1983) 139–153.

[20] Currie J., On the structure and extendability of k-power free words, *European J. Combin.* **16** (1995) 111–124.

[21] Deza A., Franek F. and A. Thierry: How many double squares can a string contains? *Discrete Appl. Math.* **180** (2015) 52–69.

[22] Droubay X., Justin J. and G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, *Theoret. Comput. Sci.* **255**(1-2) (2001) 539–553.

[23] Ehrenfeucht A., Rozenberg G., Elementary homomorphisms and a solution to the D0L sequence equivalence problem, *Theoret. Comput. Sci* **7** (1978) 169–183.

[24] Ehrenfeucht A., Karhumäki J. and G. Rozenberg, The (generalized) Post Correspondence Problem with lists consisting of two words is decidable, *Theoret. Comput. Sci.* **21** (2) (1982) 119–144.

[25] Entriger R. C., Jackson D. E. and J. A. Schatz: On Nonrepetitive Sequences, *J. Combinat. Theory (A)* **16** (1974) 159–164.

[26] Erdös P., Some unsolved problems, Magyar Tud. Akad. Kutató Int. Közl. **6** (1961) 221–254, Partially reprinted in: Paul Erdös: The Art of Counting. Selected Writings (J. Spencer, ed.) MIT Press, Cambridge, MA (1973) 15–22.

[27] Fraenkel A., Simpson J.: How many squares can a string contain? *J. Combin. Theory, Series A* **82**(1) (1998) 112–120.

[28] Fraenkel A., Simpson J., and M. Paterson, On weak circular words in binary words, Combinatorial PatternMmatching, 8th Annual Symposium, CPM 97, Aarhus, Denmark, June/July 199, Proceedings, *Lect. N. in Comput. Sci. (lncs)* **1264** (1997) 76–82.

[29] Frid A.E., Puzynina S. and L.Q. Zamboni, On palindromic factorization on words, *Adv. in Appl. Math.* **50** (2013) 737–748.

[30] Tero Harju, Juhani Karhumäki and Antonio Restivo (Eds.), Combinatorics on Words, *Theoret. Comput. Sci.* **339**(1) (2005) 1-166.

[31] Harris J. W., Stocker H.: "Hausdorff Dimension", §4.11.3 in Handbook of Mathematics and Computational Science, New York: Springer-Verlag (1998) 113–114.

[32] Heinis A., Arithmetics and combinatorics of words of low complexity, Ph.D. Thesis, University of Leiden (2001).

[33] Heinis A., The $P(n)/n$-function for bi-infinite words, *Theoret. Comput. Sci.* **273**(1–2) (2002) 35–46.

[34] Hof A., Knill O. and B. Simon, Singular continuous spectrum for palindromic Schrödinger operators, *Commun. Math. Phys.* **174**(1) (1995) 149–159.

[35] Costas Iliopoulos, Simon Puglisi and Emine Yilmaz (Eds.), String Processing and Information Retrieval, 22nd International Symposium, SPIRE 2015, London, UK, September 1-4, 2015, Proceedings, *Lect. N. in Comput. Sci. (lncs)* **9309** (2015) 1–376.

[36] Jonoska N., Manea F. and S. Seki, A Stronger Square Conjecture on Binary Words, 40th International Conference on Current Trends in Theory and Practice of Computer Science, SOFCEM 2014, Nový Smokovec, Slovakia, January 26-29, 2014, Proceedings, *Lect. N. in Comput. Sci. (lncs)* **8327** (2014) 339–350.

[37] Juhani Karhumäki, Arto Lepistö and Luca Zamboni (Eds.), Combinatorics on Words, 9th International Conference, WORDS 2013, Turku, Finland, September 2013, Proceedings, *Lect. N. in Comput. Sci. (lncs)* **8079** (2013) 1–263.

[38] Karhumäki J., Mignosi F. and W. Plandowski, The expressibility of languages and relations by word equation, ICALP'97, 24th International Colloquium on Automata, Languages, and Programming, Bologna, Italy, July 7-11, 1997, Proceeding, *Lect. N. in Comput. Sci. (lncs)* **1256** (1997) 98–109.

[39] Karhumäki J., Shallit J. , Polynomial versus exponential growth in repetition-free binary words, *J. Combin. Theory*, Series A **105** (2004) 335–347.

[40] Juhani Karhumäki and Luca Zamboni (Eds.), Words 2013, *Theoret. Comput. Sci.* **601** (2015) 1–72.

[41] Keränen V., Abelian squares are avoidable on 4 letters, International Colloquium on Automata, Languages, and Programming, ICALP'92, Proceedings, *Lect. N. in Comput. Sci. (lncs)* **623** (1992) 41–52.

[42] Kolpakov R., Kucherov G., Finding maximal repetitions in a word in linear time, in: FOCS'99, 40th Annual Symposium on Foundations of Computer Science, October 17-18, 1999, Proceedings, IEEE Computer Society Press, New York (1999) 596–604.

[43] Kosciumaka T., Radoszewski J., Rytter W. and T. Waleń: Maximum number of distinct and nonequivalent nonstandard squares in a word, *Theoret. Comput. Sci.* **648** (2016) 84–95.

[44] Labbé S., A counter-example to a question of Hof, Knill and Simon, *Electron. J. Combin.* **21**(3) #P3.11 (2014).

[45] Labbé S., Pelantová E., Palindromic sequences generated from marked morphisms, *Eur. J. Combin.* **51** (2016) 200–214.

[46] Labbé S., Pelantová E. and Š. Starosta, On the zero-defect conjecture, *Eur. J. Comb.* **62** (2017) 132–146.

[47] Lentin A., Equations dans le monoïde libre, Gauthier-Villard (1972).

[48] M. Lothaire, Combinatorics on Words, Cambridge University Press, second edition, 1997. (First edition 1983).

[49] M. Lothaire, Algebraic Combinatorics on Words, Encyclopedia of Mathemetics and its Applications 90, Cambridge University Press, 2002 (First edition 1983).

[50] Makanin G.S., On the rank of the equations in four unknowns in a free semigroup *Mat. Sb. (N.S.)* **100** (1976) 285–311.

[51] Mäkelä S., Patterns in Words, Master Thesis, Univ. Turku, Finland (2002) (in Finnish).

[52] Florin Manea and Dirk Nowotka (Eds.), Combinatorics on Words, 10th International Conference, WORDS 2015, Kiel, Germany, September 14–17, 2015, Proceedings, *Lect. N. in Comput. Sci. (lncs)* **9304** (2015) 1–237.

[53] Florin Manea and Dirk Nowotka (Eds.), TCS Special Issue: Combinatorics on Words, WORDS 2015, *Theoretical Computer Science* **684**(1–2) (2017) 1–58.

[54] Morse M., and Hedlund G.A. Symbolic dynamics, *Amer. J. Math.* **60** (1938) 815–866.

[55] Morse M., and Hedlund G.A. Symbolic dynamics II. Sturmian trajectories *Amer. J. Math.* **62** (1940) 1–42.

[56] Masáková S., Pelantová E. and Š. Starosta, Exchange of three intervals: substitution and palindromicity, *Eur. J. Combin.* **62** (2017) 217–231.

[57] Néraud J., On the rank of the subsets of a free monoid, *Theoret. Comput. Sci.* **99** (2) (1992) 231-241.

[58] Jean Néraud (Ed.), WORDS, *Theor. Comput. Sci.* **218**(1) (1999) 1–216.

[59] Jean Néraud (Ed.), WORDS, *Theoret. Comput. Sci.* **273**(1-2) (2002) 1–306.

[60] Jean Néraud (Ed.), WORDS, *Theoret. Comput. Sci.* **307**(1) (2003) 1–217.

[61] Néraud J, On the Contribution of WORDS to the Field of Combinatorics on Words, *Lecture notes in Computer Sciences (lncs)* **9304** (2015) 27–34.

[62] Néraud J., Ten Conferences WORDS and their Contribution to the Field of Combinatorics on Words, *Bull. Eur. Ass. for Theoret. Comput. Sci. (BEATCS)* **118** (2016) 145-155.

[63] Parikh R.J. , On context-free languages, *J. of the Association for Computing Machinery* **13** (1966) 57–581.

[64] Parreau A., Rigo M., Rowland E. and E. Vandomme, A new approach to the 2-regularity of the *l*-abelian complexity of 2-automatic sequences, *Electron. J. Combin.* **22** (2015) #1.27.

[65] Post E., A variant of a recursively unsolvable problem, *Bull. Amer. Math. Soc.* **52** (1946) 264–268.

[66] Rao M., Last cases of Dejean's conjecture, slides, https://www.mccme.ru/ anromash/workshop/slides/rao.pdf.

[67] Tan Bo., Mirror substitutions and palindromic sequences, *Theoret. Comput. Sci.* **389** (2007) 118–124.

[68] Zorich A., Deviation for interval exchange transformations, *Ergodic. Theor. Dyn. Syst.* **17**(6) (1997) 1477–1499.