LATVIJAS
UNIVERSITĀTE
ANNO 1919

UNIVERSITY OF LATVIA
FACULTY OF PHYSICS AND MATHEMATICS
DEPARTMENT OF MATHEMATICS

APERIODICITY IN FINITELY GENERATED BI-IDEALS AND BOUNDED BI-IDEALS

PhD Thesis

Author: **Inese Bērziņa**

Student ID ib09315

Scientific advisor: Prof. Jānis Buls

RIGA 2014

ESF
EIROPAS SOCIĀLAIS FONDS

EIROPAS SAVIENĪBA

LATVIJAS UNIVERSITATE
ANNO 1919

IEGULDĪJUMS TAVĀ NĀKOTNĒ

This work has been supported by the European Social Fund within the projects:
“Support for Doctoral Studies at University of Latvia - 2”,
“Support for Doctoral Studies at University of Latvia”,
project No. 2009/0223/1DP/1.1.1.2.0/09/APIA/VIAA/008.

Anotācija

Disertācijā pētīts aperiodiskums divu vienpusēji bezgalīgu vārdu klašu – galīgi ģenerētu bi-ideālu un ierobežotu bi-ideālu – kontekstā. Gan galīgi ģenerētus bi-ideālus, gan ierobežotus bi-ideālus ir ērti uzdot, izmantojot bāzes virkni. Ierobežotu bi-ideālu gadījumā tiek nofiksēta galīga vārdu kopa, no kuras tiek uzģenerēta bāzes virkne. Savukārt galīgi ģenerētu bi-ideālu gadījumā bāzes virkne ir periodiska, līdz ar to var uzskatīt, ka ir dota galīga bāze. Efektīvā ģenerēšana un zināmie aperiodiskuma nosacījumi kalpo par motivāciju apskatīt galīgi ģenerētu bi-ideālu potenciālo pielietojumu kriptogrāfijā – aperiodisku gadījuma skaitļu ģenerēšanā. Savukārt, ierobežoti bi-ideāli ir galīgi ģenerētu bi-ideālu dabīgs vispārinājums.

Disertācijā tiek piedāvāta konkrēta gadījuma skaitļu ģeneratora modifikācija. Oriģinālajā konstrukcijā tiek veikta filtrācija, izmantojot divas periodiskas virknes. Modificētajā gadījuma skaitļu ģeneratorā viena no periodiskajām virknēm tiek aizstāta ar galīgi ģenerētu bi-ideālu. Tiek parādīts, ka katrai periodiskai virknei var piekārtot bezgalīgi daudz galīgi ģenerētus bi-ideālus tā, ka filtrācijas rezultātā iegūtā virkne ir aperiodiska. Pierādīta arī universālo bi-ideālu eksistence – tie ir galīgi ģenerēti bi-ideāli, ar kuriem filtrējot patvaļīgu netriviālu periodisku virkni, vienmēr iegūst aperiodisku virkni.

Disertācijas galvenie rezultāti ir visu aritmētisko apakšvirkņu aperiodiskuma problēmas atrisinājums gan galīgi ģenerētiem bi-ideāliem, gan ierobežotiem bi-ideāliem, kā arī algoritms, kas nosaka, vai galīgi ģenerēts bi-ideāls ar uzdotu bāzi satur periodisku aritmētisku apakšvirkni. Papildus apskatīts speciālgadījums, kad galīgi ģenerēta bi-ideāla bāze satur tikai divus vārdus.

Pēdējā rezultātu nodaļā tiek ieviests pilnīgi ierobežota bi-ideāla jēdziens un parādīts, ka pilnīgi ierobežoto bi-ideālu klase sakrīt ar lineāri rekurento ierobežoto bi-ideālu klasi.

Atslēgas vārdi: galīgi ģenerēts bi-ideāls, ierobežots bi-ideāls, pilnīgi ierobežots bi-ideāls, universāls bi-ideāls, lineārā rekurence, aritmētiskā apakšvirkne, aperiodisks gadījuma skaitļu ģenerators.

MSC: 68R15

Abstract

The thesis explores aperiodicity property in the context of two subclasses of right infinite words. These two subclasses are finitely generated bi-ideals and bounded bi-ideals. Words of both these classes can be conveniently described using a basis sequence. In the case of bounded bi-ideals we fix a finite set of finite words and generate the basis sequence from the words of this fixed set. The basis sequence of a finitely generated bi-ideal is periodic; hence we can consider that a finite basis is given. The fact that a finitely generated bi-ideal can be effectively generated by a finite basis together with already known conditions of the aperiodicity of a finitely generated bi-ideal serves as a motivation to consider the possible use of finitely generated bi-ideals in cryptography – mainly for creating an aperiodic pseudo-number generators. Since bounded bi-ideals are a natural extension of finitely generated bi-ideals, the interest in this subclass of right infinite words increases too.

The thesis provides a modification of so-called shrinking generator, a pseudo-random number generator that uses two periodic sequences (S-sequence and A-sequence) to generate the third one. In the modified pseudo-random number generator one of the periodic sequences (S-sequence) of the shrinking generator is replaced by a finitely generated bi-ideal. It is proved that for each periodic A-sequence there exist infinitely many finitely generated bi-ideals such that the resulting shrunk sequence is aperiodic. In addition, the thesis contains the proof of the existence of universal bi-ideals – finitely generated bi-ideals that always generate aperiodic shrunk sequences when used as S-sequence in the modified shrinking generator with any periodic A-sequence containing both zeroes and ones.

The main results of the thesis are the necessary and sufficient condition of the aperiodicity of all arithmetical subsequences of finitely generated bi-ideal and bounded bi-ideal, as well as the algorithm for detecting whether a finitely generated bi-ideal, with a given basis, contains or does not contain a periodic arithmetical subsequence. The thesis also explores a special case when the basis of the finitely generate bi-ideal contains only two words.

Finally, the thesis contains the characterization of linearly recurrent bounded bi-ideals.

Keywords: finitely generated bi-ideal, bounded bi-ideal, completely bounded bi-ideal, universal bi-ideal, linear recurrence, arithmetical subsequence, aperiodic shrinking generator.

MSC: 68R15

Acknowledgements

Firstly, I want to express my gratitude to my scientific supervisor Jānis Buls for his help, support and encouragement I have received during these years. Secondly, I am also thankful to all the participants of the seminar “Combinatorics on words”, especially, to Edmunds for the fruitful discussions and advices he gave to me. In addition, the financial support of the University of Latvia and the European Social Fund has to be gratefully acknowledged. Finally, my greatest gratitude to my family, Andris and my best friends for their patience, understanding and support I have received all this time.

Contents

Notions

$\mathbb{N}$ the set of all non-negative integers,

$\mathbb{N}_+$ the set of all positive integers,

Σ_n the set $\{0, 1, 2, \ldots, n-1\}$ for some $n \in \mathbb{N}_+$,

$\overline{i, j}$ the set $\{i, i+1, i+2, \ldots, j-1, j\}$ for some $i, j \in \mathbb{N}$, $i \leq j$,

$\gcd(n, k)$ the greatest common divisor of numbers n and k,

$O(g(n))$ functions $f(x)$ such that for all n sufficiently large $\frac{f(n)}{g(n)} \leq c_2$ for some constant c_2,

$\Theta(g(n))$ functions $f(n)$ such that for all n sufficiently large, $c_1 \leq \frac{f(n)}{g(n)} \leq c_2$ for some constants c_1 and c_2 with $0 < c_1 < c_2 < +\infty$,

A^* the set of all finite words over an alphabet A,

A^+ the set of all finite non-empty words over an alphabet A,

A^ω the set of all infinite words over an alphabet A,

λ the empty word,

w^* the set $\{\lambda, w, ww, www, wwww, \ldots\}$, where w is a finite non-empty word,

u^n a finite word $\underbrace{uu\cdots u}_{n}$, where $n \in \mathbb{N}_+$ and u is a finite word,

u^ω the infinite word $uu\cdots u\cdots$, where u is a finite non-empty word,

$|u|$ the length of the finite word u,

$|u|_a$ the number of occurrences of the letter a in word the u,

$|u|_v$ the number of occurrences of the word v in word the u,

$alph(u)$ the set of all letters occurring in the word u,

$x[i, j]$ the factor of a word x starting in position i and ending in position j, where $i < j$, in other words, $x_i x_{i+1} \cdots x_{j-1} x_j$,

$x[i, j)$ the factor of a word x starting in position i and ending in position $j-1$, where $i + 1 < j$, in other words, $x_i x_{i+1} \cdots x_{j-1}$,

Pref(w)	the set of all prefixes of a word x,
F(w)	the set of all factors of a word x,
Suff(w)	the set of all suffixes of a word x,
$\mathcal{R}_{x,u}$	the set of all return words to u of x,
x_p^k	the arithmetical subsequence of x with starting position k and difference p,
$F_A(x)$	the arithmetical closure of an infinite word x,
$\langle u_0, u_1, \ldots, u_{m-1} \rangle$	a basis of a finitely generated bi-ideal,
$\langle u_0^{(n)}, u_1^{(n)}, \ldots, u_{m-1}^{(n)} \rangle$	a basis of a finitely generated bi-ideal that is obtained when the basis $\langle u_0, u_1, \ldots, u_{m-1} \rangle$ ir L-prolonged n times,
$(u_n^{(k)})_{n \geq 0}$	the basis sequence of a bi-ideal that is obtained when the basis sequence $(u_n)_{n \geq 0}$ is L-prolonged k times
$u_i^{(k)}$	the i-th element of the basis sequence that is obtained, when the initial basis sequence $(u_n)_{n \geq 0}$ is L-prolonged k times,
∂s	the differential sequence of the sequence of non-negative integers s,
$\mathbf{SHR}_y(x)$	the shrunk word of x by y,
FT$_s(u)$	a word that is obtained when the finite word u is filtered with the filter s,
FT$_s(L)$	a language that is obtained when all words of the language L are filtered with the filter s,
$K_x^{a,k}$	the set of all positions modulo $k \in \mathbb{N}_+$ of $a \in alph(x)$ in a word x , i.e., $K_x^{a,k} = \big\{ i \bmod k \big\| \, x[i] = a, i \in \mathbb{N} \big\}$.
$\breve{K}_x^{a,k}$	the set of all positions modulo $k \in \mathbb{N}_+$ of letters of the set $alph(x) \setminus \{a\}$ in a word x, i.e., $\breve{K}_x^{a,k} = \big\{ i \bmod k \big\| \, x[i] \in alph(x) \setminus \{a\}, i \in \mathbb{N} \big\}$.

Introduction

The history of combinatorics on words goes back to the beginning of 20^{th} century, when Axel Thue initiated systematic work on square-free words (Thue, 1906) and overlap-free words (Thue, 1912). The first book devoted to combinatorics on words (Lothaire, 1983) appeared in 1983 and was written by a group of authors who used pseudonym M. Lothaire. Since then the reasearch on combinatorics on words has grown rapidly and since year 2000 combinatorics on words in classification of *Mathematical Rewievs* has its own section under the chapter discrete mathematics related to computer science. Interesting facts about origins and a brief history of combinatorics on words are given in (Berstel and Perrin, 2007) and (Karhumäki, 2004).

The main objects of CoW are words – either finite or infinite sequences of symbols taken from a set called an alphabet. In this thesis we consider infinite words, known also as ω-words. Infinite periodic words are the trivial case of ω-words. We consider recurrent words, i.e., right infinite words such that each of its factors occurs in it infinitely many times. We use less known but equivalent notion of a recurrent word – a bi-ideal, i.e., the infinite word which contains as prefixes all elements of a bi-ideal sequence (Coudrain and Schützenberger, 1966), where bi-ideal sequence is a sequence of words such that each next element of the sequence is at least twice as long as the previous element and contains the previous element as both its prefix and suffix. Since the bi-ideal can be described using a basis sequence, then this equivalent notion of a recurrent word gives us some structure that appears to be useful in solutions of several problems. The elements of bi-ideal sequence are known also as Zimin's words (Zimin, 1982) and sesquipowers (Simon, 1988). In the thesis we focus on two subclasses of recurrent words – finitely generated bi-ideals and bounded bi-ideals. Finitely generated bi-ideals are generated by a periodic basis sequence, while bounded bi-ideals are generated by a basis sequence that contains only finite number of distinct words. In Chapter 1 we give a detailed survey on finitely generated and bounded bi-ideals.

Subword complexity

The subword complexity is a classical complexity measure of a word over a finite alphabet. It is a function that counts the number of factors of a given length in an infinite sequence. The complexity function $p_w(n)$ is clearly bounded by $|A|^n$ for each $n \in \mathbb{N}$, however not every function can be a complexity function. Either the sequence w is ultimately periodic and $p_w(n)$ is ultimately constant, or its subword complexity function grows at least like $n+1$ (Morse and Hedlund, 1938). Non-periodic sequences with minimal complexity $p_w(n) = n+1$ for all n exist and are called *Sturmian sequences* (Morse and Hedlund, 1940). Almost all sequences w over a finite alphabet A have high subword complexity, i.e., for all $n \geq 0$, $p_w(n) = Card(A)^n$ (see, e.g., (Allouche and Shallit, 2003) for the proof). The subword complexity is linear if there exists a constant C such that $p_w(n) \leq Cn$ for all $n \geq 0$. Cassaigne gave a characterization of infinite words with linear subword complexity using the first difference of the subword complexity, namely, an infinite word has a linear subword complexity if and only if the first difference of the subword complexity is bounded, i.e., the sequence $(p_w(n+1) - p_w(n))_{n\in\mathbb{N}}$ is bounded by some constant K (Cassaigne, 1996).

Many known sequences have relatively low subword complexity. For example, the subword complexity of an aperiodic k-automatic sequence is $\Theta(n)$ (Allouche and Shallit, 2003), but for pure morphic sequences $w = \sigma^\omega(a)$ the subword complexity function $p_w(n)$ can have only five asymptotic behaviours $\Theta(1)$, $\Theta(n)$, $\Theta(n \log n)$, $\Theta(n \log\log n)$ and $\Theta(n^2)$ (Pansiot, 1984). Clearly, as each morphic sequence is the image under a coding of a pure morphic word (Cobham, 1968), and since the coding can only decrease the subword complexity, the subword complexity of each morphic sequence is $O(n^2)$. (Deviatov, 2008) proved that for a morphic sequence w either $p_w(n) \in \Theta(n^{1+\frac{1}{k}})$ for some $k \in \mathbb{N}_+$, or $p_w(n) \in O(n \log n)$.

A natural extension of morphic words is S-adic sequences. An *adic representation* of an infinite word $w \in A^\omega$ is given by a sequence $(A_n)_{n\in\mathbb{N}}$ of alphabets, a sequence $(\sigma_n : A_{n+1}^* \to A_n^*)_{n\in\mathbb{N}}$ of morphisms, and a sequence $(a_n)_{n\in\mathbb{N}}$ of letters such that $a_i \in A_i$ for all non-negative integers i, $A_0 = A$, $\lim_{n\to+\infty} |\sigma_0\sigma_1\cdots\sigma_n(a_{n+1})| = +\infty$, and $w = \lim_{n\to+\infty} \sigma_0\sigma_1\cdots\sigma_n(a_{n+1}^\omega)$. When all the morphisms σ_n, $n \in \mathbb{N}$, belong to a given finite set S of morphisms, then w is called *S-adic*. The sequence $(\sigma_n)_{n\in\mathbb{N}}$ is called the *directive word* of w. It is clear that purely morphic sequences correspond to S-adic sequences with $|S| = 1$.

There is an open problem, called *S-adic conjecture*, to determine the link between having linear complexity (or, equivalently, affine complexity) and being an S-adic sequence. Host conjectured the existence of a strong notion of S-adicity that is equivalent to sub-affine complexity. More precisely, *S-adic conjecture* postulates: there exists a condition C such that a sequence has a linear complexity if and only if it is a S-adic sequence satisfying condition C for some finite set S of morphisms. Ferenczi solved a part of it by showing that minimal systems of sub-affine complexity are S-adic (Ferenczi, 1996). In terms of combinatorics on words, he proved that the set of factors of any uniformly recurrent sequence with linear complexity admits S-adic representation with S finite, satisfying $\lim_{n\to+\infty}\min_{a\in A_{n+1}}|\sigma_0\sigma_1\cdots\sigma_n(a)| = +\infty$.

Later Cassaigne showed the existence of a finite set S of morphisms over an alphabet $A' = A \cup \{l\}$, $l \notin A$, such that any sequence over the alphabet A is S-adic (Cassaigne, 2009). As each sequence is S-adic, it is clear that considering a particular condition C cannot be avoided. Moreover, Cassaigne's result implies that contrary to the case of morphic sequences any function which is the complexity function of some sequence is also a complexity function of some S-adic sequence. In fact, the set of all possible behaviours of the subword complexity functions of S-adic sequences is uncountable (Cassaigne, 2003). For more results on S-adic conjecture see, e.g., (Durand et al., 2013), (Leroy, 2012), (Leroy and Richomme, 2013).

An interesting property of infinite words, that implies the linearity of the subword complexity, is a linear recurrence (Durand et al., 1999). An infinite word is *linearly recurrent* if it is uniformly recurrent and there exists a constant K such that the return time to an arbitrary its factor u is bounded by $K|u|$, in other words, the gap between two consecutive occurrences of a factor of the length n does not exceed $K \cdot n$. We remark that the linear recurrence is a sufficient but not the necessary condition to have a linear subword complexity (see (Durand, 2003) for an example). Durand gave a S-adic charaterization of linearly recurrent words (Durand, 2003), (Durand, 2000). For morphic words properties uniform recurrence and linear recurrence are equivalent (see (Durand, 1998), (Durand et al., 1999),(Durand et al., 2013)). As finitely generated bi-ideals are morphic sequences (for construction see, e.g., (Cers, 2012)), then they are linearly recurrent and have a linear subword complexity. In Chapter 4 we give a characterization of linearly recurrent bounded bi-ideals by making a restriction on the basis sequence of the bounded bi-ideal.

Arithmetical complexity

Another complexity measure of infinite words (sequences) is an arithmetical complexity. It is a modification of the subword complexity and was introduced in (Avgustinovich et al., 2003).

Authors proposed to count not only all subwords of given length of the given word itself but also all factors of given length which occur in arithmetical subsequences of the given sequence. Clearly, as the given sequence is its arithmetical subsequence with starting position 0 and difference 1, then arithmetical complexity of the given sequence grows at least as fast as its subword complexity. In their seminal paper on this topic authors showed that there is no direct connection between the rate of growth of the subword complexity and the arithmetical complexity – if the subword complexity increases linearly, the arithmetical complexity can grow both linearly and exponentially. To prove it, they considered a family of D0L words with high arithmetical complexity and a family of Toeplitz words with low arithmetical complexity. More precisely, a morphism φ is called *symmetric* if for all $i \in A_q$ we have $\varphi(i) = \varphi(0) \oplus i^m$, where m is the length of $\varphi(0)$, and $\oplus$ is the addition modulo q, and $A_q = \{0, 1, \ldots, q-1\}$ for some $q \in \mathbb{N}_+$. A D0L word is *symmetric* if it is a fixed point of a symmetric morphism. If φ is a symmetric morphism over an alphabet with prime cardinality q, then for its fixed point w we have $f_w^A(n) \geq n^q$ (Avgustinovich et al., 2003). Nevertheless, the Toeplitz words that are generated by a single pattern with gaps constituting an arithmetical progression of a prime difference, dividing the length of the pattern, have a linear arithmetical complexity. Later Frid extended the result to symmetric D0L words with an arbitrary cardinality of the alphabet (Frid, 2003). A natural question that arose for arithmetical complexity was finding the non-periodic infinite words with minimal arithmetical complexity. It is interesting that the lower bound of the arithmetical complexity of the Sturmian words is $O(n^3)$ (Frid, 2005a), therefore Sturmian words are not even in the class of words with a linear complexity. The upper bound of the arithmetical complexity of the Sturmian words is also $O(n^3)$ (Cassaigne and Frid, 2007). Up to the set of factors, the uniformly recurrent non-periodic words whose arithmetical complexity grows linearly are Toeplitz words of a specific form (Frid, 2005b). The research on uniformly recurrent non-periodic words of the lowest arithmetical complexity is done in (Avgustinovich et al., 2006). Authors could not find a word of the minimal complexity, but they found a family of words with decreasing lower limits of the arithmetical complexity divided by n, which tend to be minimal. They proved that words of this family are essentially the only uniformly recurrent words of such low arithmetical complexity. Frid proved that the growth rate of the arithmetical complexity can behave as many sub-polynomial functions, namely, if $f_u(n)$ is the subword complexity of an infinite word u and $p \geq 3$ is a prime, then one can construct Toeplitz word on the same alphabet with the arithmetical complexity $a(n) = \Theta(n f_u(\lceil \log_p n \rceil))$ (Frid, 2006). It means that the variety of rates of growth of the arithmetical complexity is not less than the variety of the possible subword complexity rates of growth since each subword complexity function of a word can be included into a formula of the arithmetical complexity of another word.

We consider an aperiodicity problem related to arithmetical complexity, namely, the aperiodicity of all arithmetical subsequences of finitely generated bi-ideals and bounded bi-ideals. As each infinite word is its arithmetical subsequence, then each aperiodic infinite word has aperiodic arithmetical subsequence. However, it does not necessarily contain periodic arithmetical subsequences. For example, it is known that all arithmetical subsequences of the Thue-Morse word are aperiodic (Gelfond, 1968), while it does not hold for Toeplitz words of specific form (Frid, 2005b). In Chapter 3 we solve the aperiodicity problem of all arithmetical subsequences for finitely generated bi-ideals and bounded bi-ideals. We also present an efficient algorithm for checking, whether all arithmetical subsequences of a finitely generated bi-ideal are aperiodic.

Possible application in cryptography

Another problem we consider is a possible application of finitely generated bi-ideals in cryptography. We use them as a tool for obtaining aperiodic pseudo-random sequences with good statistical properties. As of today, the most convenient and most reliable way of generating the random numbers for stochastic simulations appears to be via deterministic algorithms with a solid mathematical basis (see, e.g., (L'Ecuyer, 1998)). These algorithms produce sequences of bits which are, in fact, not random at all, but seem to behave as if the bits were chosen independently at random.

A pseudo-random number generator is a structure $\mathfrak{G} = \langle Q, B, q_0, T, G \rangle$, where Q is a finite set of states, $q_0 \in Q$ is the initial state (or seed), the mapping $Q \xrightarrow{T} Q$ is the transition function, B is a finite set of symbols, and $Q \xrightarrow{G} B$ is the output function. This model is called a Moore machine in automata theory. In fact, this model is the specialised Moore machine.

The state of a generator is initially q_0 and evolves according to the recurrence $q_n = T(q_{n-1})$, for $n = 1, 2, 3, ...$, at step n the generator outputs the symbol $b_n = G(q_n)$.

Clearly, since the state space Q is finite, the sequence of states q_n is ultimately periodic; therefore, this approach is limited. One method for obtaining non-periodic sequences is to use the simplest chaotic system — the logistic map. Oishi and Inoue proposed the idea to use chaos in designing a pseudo-random generator (Oishi and Inoue, 1982). A simple non-periodic pseudo-random number generator which is based on a simple logistic map was introduced in (Sandri, 1992). Recently, Hu et. al proposed a true random number generator by combining congruential methods with prime numbers and higher order composition of logistic maps (Hu et al., 2009). It generates a 256-bit random number by computer mouse movement. For more information of using chaotic systems in generation of pseudo-random sequences, see, e.g., (Patidar et al., 2009), (Phatak and Suresh Rao, 1995).

We propose a method to generate an aperiodic pseudo-random number sequences based on modification of the shrinking generator, which was introduced in (Coppersmith et al., 1994)) and was considered a secure pseudo-random number generator for more than a decade. Normally, a shrinking generator uses two pseudo-random bit-sequences produced by LFSR's (see, e.g., (Schneier and Sutherland, 1995)) from which the resulting pseudo-random sequence is obtained by taking the subsequence of one of the sequences (called the A-sequence) corresponding to the positions of ones in the other sequence (called the S-sequence). Clearly, the resulting shrunk sequence is periodic with period $S \cdot A$. The periodicity of the shrunk sequence is used in crypt-analysis (see, e.g., (Caballero-Gil et al., 2009)). To avoid the periodicity of the shrunk sequence one of the sequences (S or A) should be replaced by an aperiodic sequence. In Chapter 2 we substitute the periodic S-sequence by a finitely generated bi-ideal to obtain an aperiodic shrunk sequence. First we prove that for each periodic A-sequence we can construct infinitely many finitely generated bi-ideals that generate an aperiodic pseudo-random sequence when used as the S-sequence. The resulting pseudo-random sequence has good statistical properties as indicated by the Diehard test suite. We also prove the existence of what we call universal bi-ideals – finitely generated bi-ideals that produce aperiodic pseudo-random sequences when used as the S-sequence in shrinking generator together with any non-trivial periodic A-sequence. We conjecture that for most non-trivial cases the resulting pseudo-random sequence is aperiodic.

Goals and objectives

The main objectives of this thesis is to explore fundamental properties of finitely generated bi-ideals and bounded bi-ideals, and their potential applications in cryptography.

The tasks of the thesis are associated with its goal:

- solve the aperiodicity problem of all arithmetical subsequences of finitely generated bi-ideals and bounded bi-ideals;

- give the characterization of linearly recurrent bounded bi-ideals;

- describe how finitely generated bi-ideals, bounded bi-ideals and completely bounded bi-ideals relate to other classes of right infinite words;

- explore possible applications of finitely generated bi-ideals and bounded bi-ideals in cryptography.

The scientific importance of the thesis

In the thesis I solve several problems regarding two fundamental properties of right infinite words. The first of these properties is the aperiodicity in recurrent words. I provide a necessary and sufficient condition of the aperiodicity of all arithmetical subsequences of a finitely generated bi-ideal. I give an algorithm for deciding whether a finitely generated bi-ideal, with a given basis, contains a periodic arithmetical subsequence or all its arithmetical subsequences are aperiodic. I also present a necessary and sufficient condition of the aperiodicity of all arithmetical subsequences of a bounded bi-ideal.

The second property is the linear recurrence. The linear recurrence of a uniformly recurrent word implies linearity of its subword complexity. It was known that for morphic words the linear recurrence is equivalent to its uniform recurrence, therefore finitely generated bi-ideals are linearly recurrent. Together with co-authors we give a characterization of linearly recurrent bounded bi-ideals. We introduce the notion of completely bounded bi-ideal and show that the class of completely bounded bi-ideals are exactly the class of linearly recurrent bounded bi-ideals.

Thesis also provides some research on possible applications of finitely generated bi-ideals in cryptography.

The structure of the thesis

- Chapter 1 gives preliminaries and background for the whole thesis. Section 1.1 contains some basic notions and results in combinatorics on words. In Section 1.2 I give some known and also new results on finitely generated bi-ideals and bounded bi-ideals, which are used later in the thesis.

- Chapter 2 provides some possible applications of bi-ideals in cryptography. We present modification of the shrinking generator which allows us to obtain an aperiodic pseudo-random sequences.

- Chapter 3 deals with aperiodicity problem of all arithmetical subsequences of a finitely generated bi-ideal and a bounded bi-ideal. First I present the necessary and sufficient condition of all arithmetical subsequences of a finitely generated bi-ideal. Next, I introduce an efficient algorithm for checking whether all arithmetical subsequences of a finitely generated bi-ideal, with given basis, are aperiodic. Then I give a necessary and sufficient condition of the aperiodicity of all arithmetical subsequences of a bounded bi-ideal. Finally, I show some connection of aperiodicity of all arithmetical subsequences of a finitely generated bi-ideal to the aperiodic shrinking generator introduced in Chapter 2.

- Chapter 4 contains a characterization of linearly recurrent bounded bi-ideals. We introduce the notion of a completely bounded bi-ideal and prove that a bounded bi-ideal is linearly recurrent if and only if it is a completely bounded bi-ideal.

Approbation

The results obtained during the thesis writing process have been presented at 5 international conferences and 6 domestic conferences (see full list on page 79). Here we mention only the most important conferences. Results on possible use of finitely generated bi-ideals in cryptography have been presented at the 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), held 2011 in Timisoara, Romania. Results on aperiodicity of all arithmetical subsequences of a finitely generated bi-ideal over a binary alphabet have been presented at the 14th Mons Days of Theoretical Computer Science in Louvain-La-Neuve, Belgium (2012). The characterization of linearly recurrent bounded bi-ideals have been presented at the 15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing in Timisoara, Romania (2013).

A list of author's publications is given at the end of the bibliography. The author's contribution to the papers "On a non-periodic shrinking generator" and "Bounded Bi-ideals and Linear Recurrence" is the proof of main results (Proposition 26, Theorem 30, and Theorem 65).

1 Preliminaries and Background

1.1 Preliminaries

Let $\mathbb{N}$ denote the set of all non-negative integers. Let $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. By Σ_n we denote the set $\{0, 1, \ldots, n-1\}$ for some $n \in \mathbb{N}_+$. Let $\overline{i, j}$ be the set $\{i, i+1, \ldots, j-1, j\}$, where i and j are two non-negative integers such that $i < j$.

Let A be a finite non-empty set called an *alphabet*. The elements of A are called *letters*. A string of letters $u = a_0 a_1 \cdots a_{n-1}$ from A is called a *finite word* of *length* n. We denote the length of a finite word u by $|u|$ and the number of occurrences of a letter $a \in A$ in a word u by $|u|_a$. We denote the empty word by λ and define $|\lambda| = 0$. By $alph(u)$ we denote the set of all letters that occur in the word u. If $Card(alph(u)) = 1$, we say that u is *trivial*. By A^* and A^+ we denote the sets of all finite words and all finite non-empty words over alphabet A, respectively. For finite words $u = a_0 a_1 \cdots a_n$ and $v = b_0 b_1 \cdots b_m$ we say that a word $uv = a_0 a_1 \cdots a_n b_0 b_1 \cdots b_m$ is the *concatenation of* u *and* v. A word w' is called a *factor* of $w \in A^*$ if there exist $u, v \in A^*$ such that $w = uw'v$. The word u (v, respectively) is called a *prefix* (*suffix*, respectively) of w. By u^n we denote the finite word $\underbrace{uu \cdots u}_{n}$, where $n \in \mathbb{N}_+$ and u is a finite word. Let u^* denote the set $\{\lambda, u, uu, uuu, \ldots, u^n, \ldots\}$, where u is a finite non-empty word.

A finite word $u = a_0 a_1 \cdots a_{n-1}$ is called *periodic* if there exists a positive integer p such that $a_i = a_{i+p}$ for all $i \in \overline{0, n-1-p}$. If $q \geq |u|$, then q is the period of u.

Theorem 1. (Fine and Wilf, 1965) *Let w be a finite non-empty word having periods p and q. If $|w| \geq p + q - \gcd(p, q)$, then w has also a period $\gcd(p, q)$.*

A total map $x : \mathbb{N} \to A$ is called a *(right) infinite word* and the set of all infinite words is denoted by A^ω. For all $i \geq 0$ we set $x_i = x(i)$ and write simply

$$x = x_0 x_1 \cdots x_n \cdots$$

The notion of suffix, prefix and factor generalizes straightforwardly to infinite words by setting that suffix v is an infinite word. Also concatenation extends naturally to the case when the right word is infinite. By $x[i,j]$ we denote a factor of a word x starting in the position i and ending in the position j, where $i < j$, in other words, $x[i,j] = x_i x_{i+1} \cdots x_{j-1} x_j$. Let $x[i,j)$ denote the factor of a word x starting in position i and ending in position $j-1$, where $i+1 < j$, i.e, $x[i,j) = x_i x_{i+1} \cdots x_{j-1}$. A non-negative integer i is called an *occurence* of a word u in a word x if $x[i; i+|u|) = x[i; i+|u|-1] = u$. If u is a factor of x, then we also say that u occurs or appears in x. We also write $u \searrow x$ if u appears in x. By $\mathrm{F}(w)$, Pref(w) and Suff(w) we denote the sets of all factors, prefixes and suffixes of a (finite or infinite) word w.

An infinite word $x = x_0 x_1 \ldots x_n \ldots$ is called *periodic* with a *period* p if $x_i = x_{i+p}$ for all $i \in \mathbb{N}$. If x is periodic with a period p and $v = x_0 x_1 \cdots x_{p-1}$, where $x_i \in A$ for all $i \in \mathbb{N}$, we write $x = v^\omega$. A word is called *non-periodic* if it is not periodic. A word x is called *ultimately periodic* if there exist words $u \in A^*$, $v \in A^+$ such that $x = uv^\omega$. Each periodic word $x = v^\omega$ is ultimately periodic, since it can be written in the form $x = uv^\omega$, where $u = \lambda$. A word is called *aperiodic* if it is not ultimately periodic.

An infinite word is called *recurrent* if each of its factors occurs in it infinite number of times. An infinite word x is called *uniformly recurrent* if for each non-negative integer n there exists non-negative integer m such that each x factor of length m contains as factors all factors of x of length n.

Let u be a non-empty factor of $x \in A^\omega$. A word $w \in A^+$ is called *a return word to u of x* if wu is a factor of x, u is a prefix of wu, and $|wu|_u = 2$. The set of all return words to u of x we denote by $\mathcal{R}_{x,u}$. If $x \in A^\omega$ is uniformly recurrent, then the difference between two consecutive occurrences of u in x is bounded, therefore $\mathcal{R}_{x,u}$ is finite. As finitely generated bi-ideals and bounded bi-ideals are uniformly recurrent, then $\mathcal{R}_{x,u}$ is finite for each finitely generated (or bounded) bi-ideal x and for each its factor u.

The infinite word $x_p^k = x_k x_{k+p} x_{k+2p} \cdots x_{k+np} \ldots$ is called an *arithmetical subsequence* of x starting with position k and having difference p. A factor of some x_p^k is called an *arithmetical subword* of x, but the set of all arithmetical subwords of x is called its *arithmetical closure* and is denoted by $F_A(x)$:

$$F_A(x) = \bigcup_{p \geq 1, k \geq 0} F\left(x_p^k\right) = \{\lambda\} \cup \left\{ x_k x_{k+p} \cdots x_{k+np} \middle| \, p \geq 1, k, n \geq 0 \right\}.$$

Theorem 2. (Van der Waerden, 1927) *For each infinite word w and positive integer n there exists a letter $a \in A$ such that $a^n \in F_A(w)$.*

Example 1. If $A = \{0, 1, 2\}$ and $x = 011222011222011222\cdots = (011222)^\omega$, then

$$x_1^0 = x,$$
$$x_4^2 = 102102102102\cdots = (102)^\omega.$$

A sequence of finite words $v_0, v_1, \ldots, v_i, \ldots$ is called a *bi-ideal sequence* if for each $i \geq 0$, $v_{i+1} \in v_i A^* v_i$ and $v_0 \neq \lambda$. If $v_0, v_1, \ldots, v_n, \ldots$ is a bi-ideal sequence, then there exists a unique sequence of finite words $u_0, u_1, \ldots, u_n, \ldots$ with $u_0 \neq \lambda$ called the *basis* of the bi-ideal sequence (v_n) such that

$$v_0 = u_0$$
$$v_{i+1} = v_i u_{i+1} v_i.$$

The infinite word one gets as a limit of this bi-ideal sequence $x = \lim_{n\to\infty} v_n$ is called a *bi-ideal* and the sequence (u_n) is called a basis of x or, equivalently, we say that (u_n) generates x. We also say that that (u_n) generates the bi-ideal sequence (v_n).

Definition 1. The bi-ideal is called *finitely generated* if its basis sequence (u_i) is periodic, i.e., there exists positive integer m such that for all $i, j \in \mathbb{N}$,

$$i \equiv j \pmod{m} \Rightarrow u_i = u_j.$$

In this case we say that the $m-$tuple $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ is a *finite basis* (or just a basis for short) of the finitely generated bi-ideal x. We also say that the basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ generates the bi-ideal sequence (v_n).

Example 2. If $2-$tuple $\langle 0, 1\rangle$ is a basis of the bi-ideal x, then

$$v_0 = u_0 = 0,$$
$$v_1 = v_0 u_1 v_0 = 010,$$
$$v_2 = v_1 u_0 v_1 = 0100010,$$
$$\cdots$$
$$x = 010001010100010\cdots.$$

Definition 2. If $(u_n)_{n\geq 0}$ is the basis of a bi-ideal x and there exists non-negative integer l such that, for each i, $|u_i| \leq l$, then the bi-ideal x is called *bounded*.

Proposition 3. (de Luca and Varricchio, 1999) *An infinite word x is recurrent if and only if it is a bi-ideal.*

Lemma 4. (de Luca and Varricchio, 1999) *Let $x \in A^\omega$ be an ultimately periodic word. If x is recurrent, then it is periodic.*

Due to the Proposition 3 and Lemma 4 in case of bi-ideals terms non-periodicity and aperiodicity are equivalent.

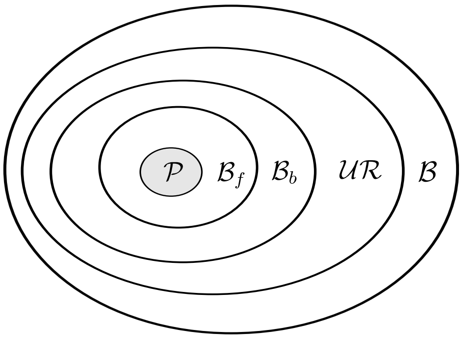


Figure 1.1: Hierarchy of the class of bi-ideals

(Buls and Lorencs, 2008) condsidered the hierarchy (see Figure 1.1):

$$\mathcal{P} \subset \mathcal{B}_f \subset \mathcal{B}_b \subset \mathcal{UR} \subset \mathcal{B},$$

where

$\mathcal{P}$	–	the class of periodic words,
$\mathcal{B}_f$	–	the class of finitely generated bi-ideals,
$\mathcal{B}_b$	–	the class of bounded bi-ideals,
$\mathcal{UR}$	–	the class of uniformly recurrent words,
$\mathcal{B}$	–	the class of bi-ideals.

For a deeper coverage of the above topics, see, for example, (Lothaire, 1983), (Lothaire, 2002), (Allouche and Shallit, 2003), (Buls and Lorencs, 2006), (Cers, 2012). In the rest of the work we mainly consider finitely generated bi-ideals and bounded bi-ideals.

1.2 More results on finitely generated bi-ideals and bounded bi-ideals

In this section we give some known and also new results on finitely generated bi-ideals and bounded bi-ideals that are used in proofs of our main results (in Chapters 2, 3, 4).

Lemma 5. *(see, e.g.,* (Buls and Lorencs, 2008)*) Let (v_n) be a bi-ideal sequence, then*

$$\forall m \leq n\ v_m \in \text{Pref}(v_n) \cap \text{Suff}(v_n).$$

(Buls and Lorencs, 2008) proved a necessary and sufficient condition of aperiodicity of a bi-ideal and also presented a criterion of aperiodicity of a finitely generated bi-ideal with a given basis.

Theorem 6. (Buls and Lorencs, 2008) *A bi-ideal x is periodic if and only if*

$$\exists n \in \mathbb{N} \exists u \exists v \left(v_n u \in v^* \wedge \forall i \in \mathbb{N}_+ u_{n+i} \in uv^* \right).$$

Theorem 7. (Buls and Lorencs, 2008) *Let (u_i) be a sequence of words, which contains every u_j infinitely often. The bi-ideal x generated by (u_i) is periodic if and only if*

$$\exists w \forall i \left(u_i \in w^* \right).$$

Later we will show (Lemma 15) that for each bounded bi-ideal there exists a basis sequence (u_i) such that each element of (u_i) occurs in (u_i) infinitely many times.

Theorem 8. (Buls and Lorencs, 2008) *A bi-ideal $x \in A^\omega$ generated by a m-tuple $\langle u_0, u_1, ..., u_{m-1} \rangle$ is periodic if and only if there exists a finite word $w \in A^+$ such that for all $i \in \overline{0, m-1}$*

$$u_i \in w^*.$$

Theorem 9. (Buls and Lorencs, 2008) *If $\bigcup_{i=0}^{m-1}$ Pref(u_i) or $\bigcup_{i=0}^{m-1}$ Suff(u_i) contains at least two words of the same length, then the bi-ideal with a basis $\langle u_0, u_1, \ldots, u_{m-1} \rangle$ is non-periodic.*

1.2.1 Change of Basis

(Lorencs, 2012) showed how to change the basis sequence of a finitely generated bi-ideal and proved that each finitely generated bi-ideal has countably many bases with the same number of basis words. In fact, his construction can be used for changing the basis sequence of an arbitrary bi-ideal.

Proposition 10. (Lorencs, 2012) *If x is a bi-ideal generated by a sequence $(u_n)_{n \geq 0}$, then the sequence $u'_0, u'_1, \ldots, u'_n, \ldots$, where $u'_i = u_0 u_{i+1}$, also generates x.*

Example 3. If x is a bounded bi-ideal with a basis sequence $0, 1, 00, 00, 00, \ldots$, then sequences $01, 000, 000, 000, 000, \ldots$ and $01000, 01000, 01000, 01000, 01000, \ldots$ are also basis sequences of x.

Proposition 11. (Lorencs, 2012) *If $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ is a basis of a finitely generated bi-ideal x, then the m-tuple $\langle u'_0, u'_1, \ldots, u'_{m-1}\rangle$, where $u'_i = u_0u_s$ and $s = i + 1 \bmod m$, also is a basis of x.*

Corollary 12. (Lorencs, 2012) *Every finitely generated bi-ideal x has countably many bases with the same number of basis words.*

Example 4. Let $\langle 0, 1, 2\rangle$ be a basis of a finitely generated bi-ideal x. Then 3-tuples $\langle 01, 02, 00\rangle$, $\langle 0102, 0100, 0101\rangle$, and $\langle 01020100, 01020101, 01020102\rangle$ are also bases of x.

If Proposition 10 or Corollary 11 is applied to some basis sequence of a bi-ideal x, then we say that basis words of the bi-ideal x are L-prolonged or simply that the basis sequence of the bi-ideal x (or the basis of a finitely generated bi-ideal) is L-prolonged. If x is a bi-ideal with a basis sequence $u_0, u_1, \ldots, u_n, \ldots$, then for all $n > 0$ the sequence

$$u_0^{(n)}, u_1^{(n)}, \ldots, u_m^{(n)}, \ldots,$$

where $u_i^{(n)} = u_0^{(n-1)}u_{i+1}^{(n-1)}$, is *the basis sequence of the bi-ideal x after n iterations of L-prolongation*. If x is a finitely generated bi-ideal with basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, then for all $n > 0$ the m-tuple

$$\left\langle u_0^{(n)}, u_1^{(n)}, \ldots, u_{m-1}^{(n)}\right\rangle,$$

where $u_i^{(n)} = u_0^{(n-1)}u_{i+1 \bmod m}^{(n-1)}$, is *the basis of the finitely generated bi-ideal x after n iterations of L-prolongation*.

Clearly, it is possible that $|u_i^{(n)}| < |u_i|$ for a basis word u_i and some integer n, however, $\lim_{n\to+\infty} \frac{|u_i^{(n)}|}{|u_i|} = +\infty$.

Lemma 13. *Let $(u_n)_{n\geq 0}$ be a basis sequence of a bi-ideal x. Let $(v_n)_{n\geq 0}$ be the bi-ideal sequence generated by $(u_n)_{n\geq 0}$. Then for each $n \in \mathbb{N}_+$ and each $i \in \mathbb{N}$*

$$u_i^{(n)} = v_{n-1}u_{i+n}.$$

Proof. If $n = 1$ then, by definition of L-prolongation, $u_i^{(1)} = u_0u_{i+1} = v_0u_{i+1}$ for each $i \in \mathbb{N}$. Assume that for all $n \leq k$ and for all $i \in \mathbb{N}$ we have $u_i^{(n)} = v_{n-1}u_{i+n}$. Let us prove it also holds for $n = k + 1$. By definition of L-prolongation and our assumption, for all $i \in \mathbb{N}$ we have

$$u_i^{(k+1)} = u_0^{(k)}u_{i+1}^{(k)} = v_{k-1}u_{0+k}v_{k-1}u_{i+1+k} = v_ku_{i+k+1}.$$

□

Corollary 14. *Let x be a bounded bi-ideal. Let $(u_n)_{n\geq 0}$ be a basis sequence of x such that each element of $(u_n)_{n\geq 0}$ occurs in $(u_n)_{n\geq 0}$ infinite number of times. Then for all $k \geq 1$ each element of a basis sequence $(u_n^{(k)})_{n\geq 0}$ occurs in $(u_n^{(k)})_{n\geq 0}$ infinitely often.*

Proof. It follows from Lemma 13. □

Lemma 15. *If x is a bounded bi-ideal, then there exists a basis sequence $(u_n)_{n\geq 0}$ of x such that each element of $(u_n)_{n\geq 0}$ occurs in $(u_n)_{n\geq 0}$ infinite number of times.*

Proof. Let x be a bounded bi-ideal with a basis sequence $(u_n)_{n\geq 0}$. Since the length of each basis word is bounded by some $\ell \in \mathbb{N}_+$, then there is at least one basis word u_i that occurs in $(u_n)_{n\geq 0}$ infinitely many times. Hence there is a non-negative integer δ such that each element of the sequence $(u_{\delta+k})_{k\geq 0}$ occurs in $(u_{\delta+k})_{k\geq 0}$ infinite number of times.

We L-prolong the basis δ times. Then by Lemma 13 we have

$$u_k^{(\delta)} = v_{\delta-1}u_{k+\delta}$$

for all $k \in \mathbb{N}$. Since $v_{\delta-1}$ is a common prefix of words $u_0^{(\delta)}, u_1^{(\delta)}, \ldots, u_n^{(\delta)}, \ldots$ and since for each $k \in \mathbb{N}$ the basis word $u_{k+\delta}$ occurs in the sequence $(u_{\delta+n})_{n\geq 0}$ infinitely many times, the basis sequence $(u_n^{(\delta)})_{n\geq 0}$ satisfies conditions of the lemma.

□

(Cers, 2010) solved the decision problem: given two bases, decide whether they generate the same finitely generated bi-ideal.

Definition 3. (Cers, 2010) We say a basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ of a finitely generated bi-ideal x is reducible if it can be changed by an application of any of the following reductions:

1. There is a word u and non-negative integers k_i, such that $u_i = u^{k_i}$ for all $i \in \overline{0, m-1}$. Then the 1-tuple $\langle u\rangle$ is also a basis of x.

2. There is a $T < m$ such that $m = k \cdot T$ for some $k \in \mathbb{N}$ and $u_i = u_{i+T}$ for all $i \in \overline{1, m-T-1}$. Then the T-tuple $\langle u_0, u_1, \ldots, u_{T-1}\rangle$ is also a basis of x.

3. There are words w_i such that $u_i = w_{m-1}w_i$ for all $i \in \overline{0, m-1}$. Then the m-tuple $\langle w_{m-1}, w_0, w_1, \ldots, w_{m-2}\rangle$ is also a basis of x.

Example 5. Here we give an example of each reduction:

1. if $\langle 0101, 01010101, 01\rangle$ is a basis of x, then $\langle 01\rangle$ also is a basis of x;

2. if $\langle 0,1,2,0,1,2\rangle$ is a basis of x, then $\langle 0,1,2\rangle$ also is a basis of x;

3. if $\langle 01020100, 01020101, 01020102\rangle$ is a basis of x, then $\langle 0102, 0100, 0101\rangle$ is also a basis of x.

A finite basis of a bi-ideal is called *irreducible* if it cannot be further reduced by any reduction of Definition 3.

Theorem 16. (Cers, 2010) *There is one and only one irreducible basis for any finitely generated bi-ideal.*

1.2.2 Length-differential Sequence of a Bi-ideal

Let x be a bi-ideal generated by a sequence $(u_n)_{n\geq 0}$. A sequence of integers $(k_n)_{n\geq 0}$, defined by $k_i = |u_i| - |u_{i+1}|$ for all $i \in \mathbb{N}$, is called a *length-differential sequence of the basis sequence* $(u_n)_{n\geq 0}$. Since the basis sequence of a finitely generated bi-ideal x is periodic, then the length-differential sequence of the basis sequence of x is periodic as well. Hence the m-tuple $\langle k_0, k_1, \ldots, k_{m-1}\rangle$, defined by

$$k_0 = |u_0| - |u_1|, k_1 = |u_1| - |u_2|, \ldots, k_{m-1} = |u_{m-1}| - |u_0|,$$

is called *the length-differential m-tuple of the basis* $\langle u_0, u_1, \ldots, u_{m-1}\rangle$.

We sometimes say that x is a finitely generated bi-ideal with a basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ and a *length-differential m-tuple* $\langle k_0, k_1, \ldots, k_{m-1}\rangle$.

Lemma 17. *Let $(u_n)_{n\geq 0}$ be a basis sequence of the bi-ideal x. Let $(k_n)_{n\geq 0}$ be the length-differential sequence of the basis sequence $(u_n)_{n\geq 0}$. Then for all $i, j \in \mathbb{N}$*

$$\left|u_j^{(i)}\right| - \left|u_{j+1}^{(i)}\right| = k_{i+j}.$$

Proof. By Lemma 13 we have $u_j^{(i)} = v_{i-1}u_{j+i}$ for all $i \in \mathbb{N}_+$, $j \in \mathbb{N}$. Hence

$$|u_j^{(i)}| - |u_{j+1}^{(i)}| = |v_{i-1}u_{j+i}| - |v_{i-1}u_{j+1+i}| = |u_{j+i}| - |u_{j+1+i}| = k_{i+j}.$$

□

Lemma 18. *If x is a finitely generated bi-ideal with basis $\langle u_0, u_1, ..., u_{m-1}\rangle$, $m \geq 1$, and length-differential m-tuple $\langle k_0, k_1, ..., k_{m-1}\rangle$, then for all $i \geq 0$, for all $j \in \overline{0, m-1}$*

$$|u_j^{(i)}| - |u_{j+1 \bmod m}^{(i)}| = k_{i+j \bmod m}.$$

Proof. Since the basis sequence $(u_n)_{n\geq 0}$ of a finitely generated bi-ideal is periodic with a period m, then the length-differential sequence also is periodic with the period m. Hence for all $i, j \geq 0$, $k_{i+j} = k_{i+j \bmod m}$. Next, by Lemma 13 we have

$$u_j^{(i)} = v_{i-1}u_{j+i} = v_{i-1}u_{i+(j \bmod m)} = u_{j \bmod m}^{(i)}.$$

From here and Lemma 17 for all $i, j \in \mathbb{N}$ we have

$$|u_j^{(i)}| - |u_{j+1 \bmod m}^{(i)}| = |u_j^{(i)}| - |u_{j+1}^{(i)}| = k_{i+j} = k_{i+j \bmod m}.$$

□

Example 6. Let $m = 3$. The bi-ideal x is generated by the basis $\langle u_0, u_1, u_2\rangle$, where $|u_0| = a$, $|u_1| = b$, $|u_2| = c$.

i	$\lvert u_0^{(i)}\rvert$	$\lvert u_1^{(i)}\rvert$	$\lvert u_2^{(i)}\rvert$	$\lvert u_0^{(i)}\rvert - \lvert u_1^{(i)}\rvert$	$\lvert u_1^{(i)}\rvert - \lvert u_2^{(i)}\rvert$	$\lvert u_2^{(i)}\rvert - \lvert u_0^{(i)}\rvert$
0	a	b	c	$a-b$	$b-c$	$c-a$
1	$a+b$	$a+c$	$2a$	$b-c$	$c-a$	$a-b$
2	$2a+b+c$	$3a+b$	$2a+2b$	$c-a$	$a-b$	$b-c$
3	$5a+2b+c$	$4a+3b+c$	$2(2a+b+c)$	$a-b$	$b-c$	$c-a$
4	$9a+5b+2c$	$9a+4b+3c$	$2(5a+2b+c)$	$b-c$	$c-a$	$a-b$
...	...	...	...	...	...	...

If we denote $k_0 = a - b$, $k_1 = b - c$, $k_2 = c - a$, then

i	$\lvert u_0^{(i)}\rvert - \lvert u_1^{(i)}\rvert$	$\lvert u_1^{(i)}\rvert - \lvert u_2^{(i)}\rvert$	$\lvert u_2^{(i)}\rvert - \lvert u_0^{(i)}\rvert$
0	k_0	k_1	k_2
1	k_1	k_2	k_0
2	k_2	k_0	k_1
3	k_0	k_1	k_2
4	k_1	k_2	k_0
...	...	...	...

Corollary 19. *If x is a bi-ideal with basis $\langle u_0, u_1\rangle$ and $k = ||u_0| - |u_1||$, then for each $i \in \mathbb{N}$*

$$\left||u_0^{(i)}| - |u_1^{(i)}|\right| = \left||u_1^{(i)}| - |u_0^{(i)}|\right| = k.$$

Lemma 20. *If x is a bounded bi-ideal, then there exists a basis sequence $(u_n)_{n\geq 0}$ such that*

1. *each element of $(u_n)_{n\geq 0}$ occurs in $(u_n)_{n\geq 0}$ infinitely often;*
2. *each element of the length-differential sequence $(k_n)_{n\geq 0}$ of the basis sequence $(u_n)_{n\geq 0}$ occurs in $(k_n)_{n\geq 0}$ an infinite number of times.*

Proof. Let x be a bounded bi-ideal. By Lemma 15, there is a basis sequence $(u_n)_{n\geq 0}$ of x such that each element of $(u_n)_{n\geq 0}$ occurs in $(u_n)_{n\geq 0}$ infinitely often. Let $(k_n)_{n\geq 0}$ be the length-differential sequence of $(u_n)_{n\geq 0}$, i.e.,

$$k_0 = |u_0| - |u_1|, k_1 = |u_1| - |u_2|, \dots, k_n = |u_n| - |u_{n+1}|, \dots.$$

As there is a finite number of distinct basis words in the sequence $(u_n)_{n\geq 0}$, then the number of distinct elements in the sequence $(k_n)_{n\geq 0}$ is finite as well. Hence there is at least one element of the sequence $(k_n)_{n\geq 0}$ that occurs in $(k_n)_{n\geq 0}$ infinitely often. Moreover, there exists non-negative integer δ such that each element of the sequence

$$k_\delta, k_{\delta+1}, k_{\delta+2}, \dots, k_{\delta+n}, \dots \tag{1.21}$$

occurs in it infinitely many times.

We recall that by Lemma 17 for each $i, j \in \mathbb{N}$, $|u_j^{(i)}| - |u_{j+1}^{(i)}| = k_{i+j}$. From here, if we L-prolong basis words δ times, then for each $j \in \mathbb{N}$ we have

$$|u_j^{(\delta)}| - |u_{j+1}^{(\delta)}| = k_{\delta+j}.$$

Let us consider the basis sequence $(u_n^{(\delta)})_{n\geq 0}$. Clearly, $(k_{\delta+n})_{n\geq 0}$ is the length-differential sequence of the basis sequence $(u_n^{(\delta)})_{n\geq 0}$. By (1.21) we already know that each element of $(k_{\delta+n})_{n\geq 0}$ occurs in $(k_{\delta+n})_{n\geq 0}$ infinite number of times. By Corollary 14, each element of $(u_n^{(\delta)})_{n\geq 0}$ occurs in $(u_n^{(\delta)})_{n\geq 0}$ infinitely often. Thus the basis sequence $(u_n^{(\delta)})_{n\geq 0}$ satisfies conditions of the lemma. □

Corollary 21. *If $(u_n)_{n\geq 0}$ is a basis sequence of a bounded bi-ideal that satisfies the conditions of Lemma 20, then for all $m \geq 0$ the sequence $(u_n^{(m)})_{n\geq 0}$ also satisfies the conditions of Lemma 20.*

Proof. According to Corollary 14 for each $m \geq 0$ the sequence $(u_n^{(m)})_{n\geq 0}$ satisfies the first condition of Lemma 20. By Lemma 17 we know that for each $m, j \in \mathbb{N}$

$$|u_j^{(m)}| - |u_{j+1}^{(m)}| = k_{m+j}.$$

Thus the length-differential sequence of x associated with $(u_n^{(m)})_{n\geq 0}$ is the sequence $(k_{m+j})_{j\geq 0}$, i.e., $k_m, k_{m+1}, k_{m+2}, \ldots, k_{m+j}, \ldots$. As each element of $(k_n)_{n\geq 0}$ occurs in $(k_n)_{n\geq 0}$ an infinite number of times, then also each element of $(k_{m+j})_{j\geq 0}$ occurs in $(k_{m+j})_{j\geq 0}$ infinitely often. □

Lemma 22. *Let x be a bounded bi-deal with a basis sequence $(u_n)_{n\geq 0}$. Let $(k_n)_{n\geq 0}$ be the length-differential sequence of $(u_n)_{n\geq 0}$. Let $\mathcal{K}_\infty$ be the set of all elements of the sequence $(k_n)_{n\geq 0}$ that occurs in $(k_n)_{n\geq 0}$ infinitely many times. Then for all $k \in \mathcal{K}_\infty$ there exists an infinite increasing sequence $(s_n)_{n\geq 0}$ of non-negative integers such that:*

1. $|u_{s_0}| = |u_{s_1}| = \cdots |u_{s_n}| = \cdots$;

2. $|k| = |k_{s_0}| = |k_{s_1}| = \cdots = |k_{s_n}| = \cdots$.

Proof. Clearly, for each $k \in \mathcal{K}_\infty$ there is an infinite increasing sequence of non-negative integers $(r_n)_{n\geq 0}$ such that

$$k = k_{r_0} = k_{r_1} = \cdots k_{r_n} = \cdots .$$

Since by definition of a bounded bi-ideal there is a non-negative integer l such that the length of each basis word does not exceed l, then there is also a finite number of distinct elements in the sequence $(u_{r_n})_{n\geq 0}$. Hence at least one of the basis words occurs in $(u_{r_n})_{n\geq 0}$ infinitely often. Thus we can choose a subsequence $(s_n)_{n\geq 0}$ of the sequence $(r_n)_{n\geq 0}$ such that

$$|u_{s_0}| = |u_{s_1}| = \cdots |u_{s_n}| = \cdots .$$

Clearly, as $(s_n)_{n\geq 0}$ is a subsequence of $(r_n)_{n\geq 0}$, then

$$|k| = |k_{s_0}| = |k_{s_1}| = \cdots = |k_{s_n}| = \cdots .$$

□

Lemma 23. *Let $(u_n)_{n\geq 0}$ be a basis sequence of a bounded bi-ideal x. Let $(k_n)_{n\geq 0}$ be the length-differential sequence of $(u_n)_{n\geq 0}$. Let $(s_n)_{n\geq 0}$ be an infinite increasing sequence of non-negative integers such that*

1. $|u_{s_0}| = |u_{s_1}| = \cdots |u_{s_n}| = \cdots$;

2. $|k_{s_0}| = |k_{s_1}| = \cdots = |k_{s_n}| = \cdots$.

Then for all $i \in \mathbb{N}$ the increasing sequence $(s'_n)_{n\geq 0}$ of non-negative integers, defined by

$$s'_0 = 0, s'_1 = s_{i+1} - s_i, \ldots, s'_j = s_{i+j} - s_i, \ldots,$$

satisfies such conditions

1. $|u^{(s_i)}_{s'_0}| = |u^{(s_i)}_{s'_1}| = \cdots = |u^{(s_i)}_{s'_n}| = \cdots$*;*

2. $|k_{s_0}| = |k'_{s'_0}| = |k'_{s'_1}| = \cdots = |k'_{s'_n}| = \cdots$*, where* $k'_j = |u^{(s_i)}_j| - |u^{(s_i)}_{j+1}|$ *for all* $j \in \mathbb{N}$*.*

Proof. Let $i \in \mathbb{N}$. According to Lemma 13 and definition of the sequence $(s'_n)_{n\geq 0}$, for each non-negative integer j,

$$u^{(s_i)}_{s'_j} = v_{s_i-1}u_{s_i+s'_j} = v_{s_i-1}u_{s_i+s_{i+j}-s_i} = v_{s_i-1}u_{s_{i+j}}.$$

From here, for each $j, j' \in \mathbb{N}$ we have

$$|u^{(s_i)}_{s'_j}| = |v_{s_i-1}u_{s_{i+j}}| = |v_{s_i-1}| + |u_{s_{i+j}}| = |v_{s_i-1}| + |u_{s_{i+j'}}| = |v_{s_i-1}u_{s_{i+j'}}| = |u^{(s_i)}_{s'_{j'}}|,$$

hence

$$|u^{(s_i)}_{s'_0}| = |u^{(s_i)}_{s'_1}| = \cdots = |u^{(s_i)}_{s'_n}| = \cdots .$$

According to Lemma 17, for each $j \in \mathbb{N}$

$$k'_{s'_j} = |u^{(s_i)}_{s'_j}| - |u^{(s_i)}_{s'_{j+1}}| = k_{s_i+s'_j} = k_{s_i+s_{i+j}-s_i} = k_{s_{i+j}} = k_{s_0},$$

whence

$$|k'_{s'_0}| = |k'_{s'_1}| = \cdots = |k'_{s'_n}| = \cdots .$$

□

1.2.3 Some other Relations in Bi-ideal Sequences

Next we give result which is very useful in proofs of our results in Chapter 2 and Chapter 3.

Lemma 24. *If x is a finitely generated bi-ideal generated by $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, then for all $p, T \in \mathbb{N}_+$ there exist infinitely many positive integers α and β, $\alpha < \beta$, such that*

$$|v_{\alpha m-1}| \equiv |v_{\beta m-1}| \pmod{p}, \tag{1.22}$$

$$|v_{\alpha m-1}|_1 \equiv |v_{\beta m-1}|_1 \pmod{T}. \tag{1.23}$$

Here v_i denotes the i-th element of the bi-ideal sequence generated by the basis sequence (u_n).

Proof. Let (v_n) be the bi-ideal sequence corresponding to the finitely generated bi-ideal x. We consider the subsequence $(v_{im-1})_{i\geq 1}$ of (v_n). Since (v_n) is an infinite sequence, $(v_{im-1})_{i\geq 1}$ is also an infinite sequence.

We partition $(v_{im-1})_{i\geq 1}$ into equivalence classes by their length modulus p:

$$\forall k \geq 1 \;\; A_t = \big\{ v_{km-1} \big| \, |v_{km-1}| \equiv t \,(\text{mod}\, p) \big\}.$$

Since $(v_{im-1})_{i\geq 1}$ is an infinite sequence, there exists an integer $\ell \in \{0, 1, \ldots, p-1\}$ such that $|A_\ell| = \infty$. For all $v_{k_1 m-1}, v_{k_2 m-1} \in A_\ell$ condition (1.22) holds.

Next, we partition $(v_{im-1})_{i\geq 1}$ further based on the number of ones modulo T:

$$\forall k \geq 1 \;\; B_t = \{ v_{km-1} | \; v_{km-1} \in A_\ell \wedge |v_{km-1}|_1 \equiv t \,(\text{mod}\, T) \}.$$

Since $|A_\ell| = \infty$, there exists an integer $s \in \{0, 1, \ldots, T-1\}$, such that $|B_s| = \infty$. For all $v_{k_1 m-1}, v_{k_2 m-1} \in B_s$ conditions 1.22 and 1.23 hold.

□

Corollary 25. *Let x be a bi-ideal generated by the m-tuple $\langle u_0, u_1, ..., u_{m-1} \rangle$. If*

$$|v_{\alpha m-1}| \equiv |v_{\beta m-1}| \pmod{p}$$

for some $\alpha, \beta, p \in \mathbb{N}_+$, $\alpha < \beta$, then for all $j \in \mathbb{N}_+$

$$|v_{\alpha m-1+j}| \equiv |v_{\beta m-1+j}| \pmod{p}.$$

Proof. By definition of a finitely generated bi-ideal, for each $k, l \in \mathbb{N}$

$$k \equiv l \pmod{m} \Rightarrow u_k = u_l,$$

therefore for each $j \in \mathbb{N}_+$ one gets

$$|v_{\alpha m-1+j}| = |v_{\alpha m-2+j} u_{j-1 \bmod m} v_{\alpha m-2+j}| = 2|v_{\alpha m-2+j}| + |u_{j-1 \bmod m}|$$

and

$$|v_{\beta m-1+j}| = |v_{\beta m-2+j} u_{j-1 \bmod m} v_{\beta m-2+j}| = 2|v_{\beta m-2+j}| + |u_{j-1 \bmod m}|.$$

The rest of the proof is left to the reader – it can be easily proved by induction on j.

□

2 Aperiodic shrinking generator and filtering problem

In this Chapter we consider two problems that historically were considered independently but obviously are connected. First, in Section 2.1 we show possible use of finitely generated bi-ideals in cryptography. We modify pseudo-random generator called *shrinking generator* (Coppersmith et.al., 1994) and obtain new pseudo-random generator, which produces aperiodic shrunk sequence and still has good statistical properties. In Section 2.2 we consider so called *filtering problem* (see (Berstel et al., 2006)), show its connection to shrinking generator and its solution for finitely generated and bounded bi-ideals.

2.1 Aperiodic shrinking generator

Definition 4. Let $x, y \in \{0,1\}^\omega$ be two infinite words with $|y|_1 = \infty$. The *shrunk sequence* of x by y is defined inductively:

$$w_0 := \begin{cases} x_0 & \text{if } y_0 = 1, \\ \lambda & \text{if } y_0 = 0, \end{cases}$$

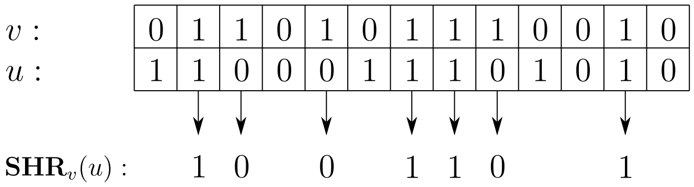
$$\forall i \geq 1 \; w_i := \begin{cases} w_{i-1}x_i & \text{if } y_i = 1, \\ w_{i-1} & \text{if } y_i = 0. \end{cases}$$

The infinite word $z = \lim_{i\to\infty} w_i$ is called a shrunk word of x by y and denoted by $\mathbf{SHR}_y(x)$.

Clearly, the shrunk word can also be defined for two finite words of equal length, namely, if $u, v \in \{0,1\}^n$, $n \in \mathbb{N}$, then $\mathbf{SHR}_v(u) = w_{n-1}$.

Example 7. Let $u = 1100011101010$ and $v = 0110101110010$, then $|u| = |v| = 13$, therefore $\mathbf{SHR}_v(u) = w_{12}$. We construct w_{12} inductively: $w_0 = \lambda$ since $v_0 = 0$; for $i \geq 1$ we have

$w_1 = w_0 1 = 1$, $w_2 = w_1 0 = 10$, $w_3 = w_2 = 10$, $w_4 = w_3 0 = 100, w_5 = w_4 = 100$, $w_6 = 1001, w_7 = 10011, w_8 = 100110, w_9 = 100110, w_{10} = 100110, w_{11} = 1001101$, and $w_{12} = 1001101$. Thus $\mathbf{SHR}_v(u) = w_{12} = 1001101$ (see also Figure 2.1 for a schematical representation of the shrinking operation).



v : 0 1 1 0 1 0 1 1 1 0 0 1 0

u : 1 1 0 0 0 1 1 1 0 1 0 1 0

$\mathbf{SHR}_v(u)$: 1 0 0 1 1 0 1

Figure 2.1: Construction of shrinking for finite words of equal length.

Further we show two approaches for generating aperiodic pseudo-random number sequences using our modified shrinking generator.

2.1.1 Construction of aperiodic shrunk words

In order to construct an aperiodic shrunk word, the finitely generated bi-ideal, which is used as S-sequence, has to be aperiodic. (Buls and Lorencs, 2008) obtained sufficient conditions for a finitely generated bi-ideal to be aperiodic (see Theorem 9), however, the aperiodicity of the bi-ideal (S-sequence) alone is not a sufficient condition for the shrunk sequence to be aperiodic. Next, we give two examples (without proof), where the resulting sequence is periodic.

Example 8. If $x = (1100)^\omega$ and y is a finitely generated bi-ideal with basis $\langle 01, 10 \rangle$, then $\mathbf{SHR}_y(x) = (10)^\omega$. Indeed, if we L-prolong the basis one time, then the new-obtained basis of the bi-ideal y is $\langle 0110, 0101 \rangle$. Now, we observe that $|1100| = |0110| = |0101| = 4$. Since the length of 1100 is equal to the length of new-obtained basis words 0110 and 0101, then $\mathbf{SHR}_y(x)$ can be expressed as an infinite concatenation of words $\mathbf{SHR}_{0110}(1100) = 10$ and $\mathbf{SHR}_{0101}(1100) = 10$. Clearly, the shrunk word $\mathbf{SHR}_y(x)$ is the periodic word $(10)^\omega$.

Example 9. If $x' = (01)^\omega$ and y' is a finitely generated bi-ideal with basis $\langle 101, 10001 \rangle$, then $\mathbf{SHR}_{y'}(x') = (0011)^\omega$. To prove it, we first L-prolong the basis one time. The new-obtained basis is $\langle 10110001, 101101 \rangle$. The period of x divides the length of both 10110001 and 101101, therefore the shrunk word $\mathbf{SHR}_y(x)$ can be expressed as an infinite concatenation of words $\mathbf{SHR}_{10110001}(01010101) = 0011$ and $\mathbf{SHR}_{101101}(010101) = 0011$. Hence $\mathbf{SHR}_y(x) = (0011)^\omega$.

In both examples condition of Theorem 9 is satisified, e.g., the bi-ideals used as the S-sequences are aperiodic, but the resulting shrunk word is periodic. Moreover, the period of the shrunk sequence can be smaller or larger than the period of the respective A-sequence.

In order to construct an aperiodic shrunk word, we have to put some additional restrictions on the basis of the finitely generated bi-ideal that will be used as the S-sequence.

Proposition 26. *Let $x = (u\bar{a}a^n)^\omega$, where $u \in \{0,1\}^*$, $a \in \{0,1\}$, $\bar{a} \in \{0,1\} \setminus \{a\}$, and $n \in \mathbb{N}_+$. Let $y \in \{0,1\}^\omega$ be a finitely generated bi-ideal with basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, $m \geq 2$, such that $u_0 = u'10v$, $u_1 = u''01v$, where $u', u'' \in \{0,1\}^*$ are two arbitrary finite words (possibly empty), and v is a binary word of length $|v| = n - 1$. Then $\mathbf{SHR}_y(x)$ is aperiodic.*

Proof. We suppose the contrary: For a periodic word $x = (u\bar{a}a^n)^\omega$, where $n \in \mathbb{N}_+$, $u \in \{0,1\}^*$, $a \in \{0,1\}$, and $\bar{a} \in \{0,1\} \setminus \{a\}$, there exists a finitely generated bi-ideal with basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, $m \geq 2$, such that $u_0 = u'10v$, $u_1 = u''01v$, $u', u'' \in \{0,1\}^*$ and v is a binary word of length $|v| = n - 1$, such that the shrunk word $\mathbf{SHR}_y(x)$ is ultimately periodic, namely, $\mathbf{SHR}_y(x) = w'w^\omega$ with $w' \in \{0,1\}^*$ and $w \in \{0,1\}^+$. Let $p = |u\bar{a}a^n|$, $T_1 = |w'|$ and $T = |w|$.

Now we consider the length of some terms of the bi-ideal sequence $(v_i)_{i\geq 0}$ of the bi-ideal y. We recall that by definition $v_0 = u_0$ and $v_i = v_{i-1}u_iv_{i-1}$ for all $i \geq 1$. According to Lemma 24 we can choose two distinct positive integers α and β, $\alpha + 1 < \beta$, such that

$$|v_{\alpha m-1}| \equiv |v_{\beta m-1}| \pmod{p},$$
$$|v_{\alpha m-1}|_1 \equiv |v_{\beta m-1}|_1 \pmod{T},$$
$$|v_{\alpha m-1}| \geq p \wedge |v_{\alpha m-1}|_1 \geq T \wedge |v_{\alpha m-1}|_1 > T_1,$$

therefore there exist $k, k_1 \in \mathbb{N}_+$ such that

$$|v_{\beta m-1}| - |v_{\alpha m-1}| = kp, \tag{2.11}$$

$$|v_{\beta m-1}|_1 - |v_{\alpha m-1}|_1 = k_1T. \tag{2.12}$$

Since by Definition 1 of a finitely generated bi-ideal we have $u_{\alpha m} = u_0 = u_{\beta m}$, then $v_{\alpha m} = v_{\alpha m-1}u_0v_{\alpha m-1}$ and $v_{\beta m} = v_{\beta m-1}u_0v_{\beta m-1}$. From here and (2.11) we easily obtain

$$|v_{\beta m}| - |v_{\alpha m}| = 2|v_{\beta m-1}| + |u_0| - (2|v_{\alpha m-1}| + |u_0|) = 2kp, \tag{2.13}$$

hence

$$|y[|v_{\beta m-1}| - |v_{\alpha m-1}|, |v_{\beta m}| - |v_{\alpha m}|)| = kp. \tag{2.14}$$

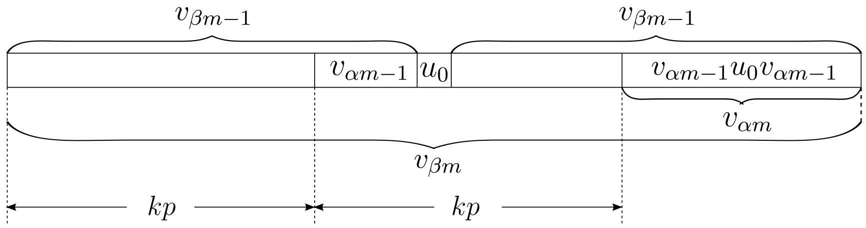


Figure 2.2: The structure of $v_{\beta m}$.

Further we observe that (see Figure 2.2)

$$\begin{aligned}|y[|v_{\beta m-1}|-|v_{\alpha m-1}|,|v_{\beta m}|-|v_{\alpha m}|)|_1 &= |v_{\alpha m-1}u_0|_1+|v_{\beta m-1}|_1-|v_{\alpha m-1}u_0v_{\alpha m-1}|_1\\ &= |v_{\beta m-1}|_1-|v_{\alpha m-1}|_1,\end{aligned}$$

which together with (2.12) imply

$$|y[|v_{\beta m-1}|-|v_{\alpha m-1}|,|v_{\beta m}|-|v_{\alpha m}|)|_1 = k_1T. \tag{2.15}$$

Moreover, as $\mathbf{SHR}_y(x)$ is ultimately periodic with preperiod T_1 and period T, then

$$\mathbf{SHR}_y(x)[i]=\mathbf{SHR}_y(x)[i+k_1T]$$

for all $i\in\overline{T_1,k_1T-1}$. From here we deduce the following: if there exists a positive integer $\delta\in\overline{1,kp-|v_{\alpha m-1}|}$ such that

$$|y[kp-\delta,kp)|_1=|y[2kp-\delta,2kp)|_1,$$

then

$$\mathbf{SHR}_{y[kp-\delta,kp)}(x[kp-\delta,kp))=\mathbf{SHR}_{y[2kp-\delta,2kp)}(x[2kp-\delta,2kp)). \tag{2.16}$$

Next we will show that indeed such integer δ exists. According to Lemma 5 and the fact $\alpha+1\le\beta$ we have $v_{\alpha m+1}\in\text{Suff}(v_{\beta m-1})$ and $v_{\alpha m+1}\in\text{Suff}(v_{\beta m})$ (see Figure 2.3). Since by Definition 1 of a finitely generated bi-ideal $u_{\alpha m}=u_0$ and $u_{\alpha m+1}=u_1$, then

$$v_{\alpha m+1}=v_{\alpha m}u_1v_{\alpha m}=v_{\alpha m-1}u_0v_{\alpha m-1}u_1v_{\alpha m-1}u_0v_{\alpha m-1}.$$

From here and (2.11) we obtain that $u_0=y[kp-|u_0|,kp)$ (see Figure 2.3), while (2.13) implies $u_1=y[2kp-|u_1|,2kp)$. Clearly, basis words u_0 and u_1 might be of different length. We also do not know the number of letter 1 in any of these words. Nevertheless, we have $u_0=u'10v$ and $u_1=u''01v$, where $v\in\{0,1\}^{n-1}$ and $u',u''\in\{0,1\}^*$.

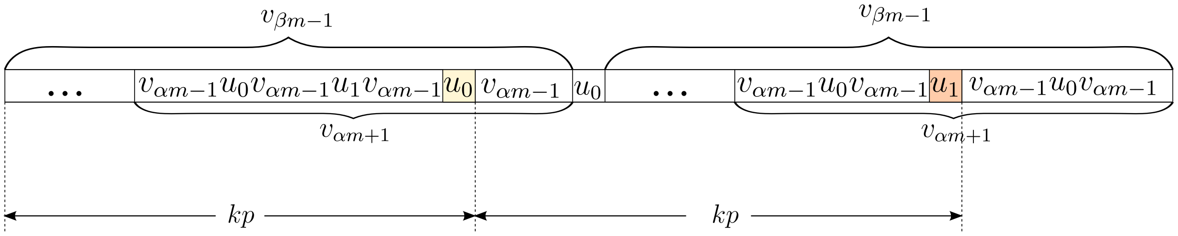

$v_{\beta m-1}$
$v_{\beta m-1}$
...
$v_{\alpha m-1}u_0v_{\alpha m-1}u_1v_{\alpha m-1}$
u_0
$v_{\alpha m-1}$
u_0
...
$v_{\alpha m-1}u_0v_{\alpha m-1}$
u_1
$v_{\alpha m-1}u_0v_{\alpha m-1}$
$v_{\alpha m+1}$
$v_{\alpha m+1}$
kp
kp

Figure 2.3: More detailed structure of $v_{\beta m}$.

Clearly, $|10v|_1 = |01v|_1$, therefore if we set $\delta = |10v| = |01v| = n+1$, then (2.16) holds, namely, $\mathbf{SHR}_{y[kp-|10v|,kp)}(x[kp-|10v|,kp)) = \mathbf{SHR}_{y[2kp-|10v|,2kp)}(x[2kp-|10v|,2kp))$.

Since $x = (u\bar{a}a^n)^\omega$ is periodic with period $p = |u\bar{a}a^n|$, then $x[kp-|10v|,kp)$ and also $x[2kp-|10v|,2kp)$ is a suffix of length $n+1$ of the word $u\bar{a}a^n$, i.e.,

$$x[kp-|10v|,kp) = x[2kp-|10v|,2kp) = \bar{a}a^n.$$

From here and (2.16) we have

$$\mathbf{SHR}_{10v}(\bar{a}a^n) = \mathbf{SHR}_{01v}(\bar{a}a^n). \tag{2.17}$$

Finally, we observe that (2.17) leads us to contradiction. Indeed, (2.17) together with equality $\mathbf{SHR}_v(a^{n-1}) = \mathbf{SHR}_v(a^{n-1})$ imply $\mathbf{SHR}_{10}(\bar{a}a) = \mathbf{SHR}_{01}(\bar{a}a)$, which does not hold since $\mathbf{SHR}_{10}(\bar{a}a) = \bar{a}$ and $\mathbf{SHR}_{01}(\bar{a}a) = a$. Thus $\mathbf{SHR}_y(x)$ is aperiodic. □

Corollary 27. *If x is a non-trivial infinite periodic word, then there exists an infinite number of finitely generated bi-ideals y, such that* $\mathbf{SHR}_y(x)$ *is aperiodic.*

Proof. It follows directly from Proposition 26 since we have put restrictions only on first two basis words u_0 and u_1. Moreover, the only restriction that we have put on the number of basis words is $m \geq 2$, therefore if we choose $m > 3$, then all basis words u_j, $3 \leq j \leq m-1$, can be chosen arbitrarily. □

Example 10. Let $x = (0110111)^\omega$. Following the notations in Proposition 26, we have $u = 011$, $a = 1$, $\bar{a} = 0$, $n = 3$. According to Proposition 26 in order to obtain aperiodic shrunk word we need to choose arbitrary 2-letter word $v \in \{0,1\}^2$ (since $n-1 = 3-1 = 2$). Let $v = 10$. As $u', u'' \in \{0,1\}^*$ can chosen arbitrarily, then let $u' = u'' = \lambda$. Then we have $u_0 = 1010$ and $u_1 = 0110$. We can fix any $m \geq 2$. Let $m = 5$, then we can arbitrarily choose the basis words u_2, u_3 and u_4. Let $u_2 = 0$, $u_3 = 1$ and $u_4 = 101$. Then $\mathbf{SHR}_y(x)$, where y is the finitely generated bi-ideal generated by the basis $\langle 1010, 0110, 0, 1, 101\rangle$, is aperiodic.

2.1.2 Statistics

One way of evaluating the fitness of a pseudo random generator for cryptographic applications is to check whether the produced bit-sequence appears random in the statistical sense, i.e., that it does not exemplify patterns that would be unexpected in a sequence of truly random and independent coin flips. The simplest of such tests is the frequency test, that checks if the number of ones is close to the number of zeroes. Many such tests can and have been constructed and several software packages for testing pseudo-random number generators are available. We used the well known Diehard battery of tests (Marsaglia, 1996) to asses the fitness of our generator. This test suite includes 18 main and several more additional tests, all of which a good generator is expected to pass.

While it was known that the shrinking generator has good statistical properties (Coppersmith et al., 1994), this did not necessitate that these properties would carry over to our construction. Still, we found that our shrinking generator passes *all* tests in the Diehard test suite. For the testing purposes a 32 bit LFSR was taken as the A-sequence and a bi-ideal with base words of lengths around 2KB (that were generated by cutting up another 32-bit LFSR) was used as the S-sequence. Additionally the first two base words were altered in a way so that conditions of Proposition 26 were satisfied, making the shrunk sequence aperiodic (the required changes are very small compared to a freely selected base). The number of base words was not limited, but the lengths of the tests were such, that around the first 20 base words were used while performing each test.

2.1.3 Universal bi-ideals

In Subsection 2.1.1 we showed how it is possible to construct aperiodic S-sequences for each periodic A-sequence such that the resulting shrunk words are aperiodic. Even though for each A-sequence there exists an infinite number of S-sequences such that the shrunk word is aperiodic, the choice of the S-sequence depends on the choice of the A-sequence. In order to simplify the choice of the sequences, it would be more convenient to use aperiodic bi-ideals (as S-sequences) such that for each non-trivial A-sequence the resulting shrunk word is aperiodic. In Theorem 30 we prove the existence of such bi-ideals.

Definition 5. A bi-ideal y is called universal if for all non-trivial periodic $x = u^\omega$, the shrunk word $\mathbf{SHR}_y(x)$ is aperiodic.

Lemma 28. *Let $a, b \in A$, $u \in A^*$ and $|aub| > T > 1$. If T is the least period of aub, then $au \neq ub$.*

Proof. If $u = \lambda$, then $aub = ab$. Since $T > 1$ then $a \neq b$. Therefore

$$au = a \neq b = ub.$$

The rest of the proof is by induction on the length of the word u. Since T is the period of aub, the period t of the word au has to be less than or equal to T, i.e., $t \leq T$.

(i) If $t = 1$, then $au = a^n$, where $n = |au|$. Since $T > 1$ is the period of the word aub, $b \neq a$. Therefore $au = a^n \neq ub$.

(ii) Let $u = vc$ and $t > 1$, i.e., $t > 1$ is the period of the word $au = avc$. By the induction assumption $av \neq vc$. From this

$$au = avc \neq vcb = ub.$$

□

Lemma 29. *Let $m \in \mathbb{N}$, $m \geq 2$. Let $u_0 = 1$, $u_1 = 10$, and $00 \notin F(u_i)$ for all $i \in \overline{2, m-1}$ (if $m > 2$). Then $00 \notin F(x)$, where x is the bi-ideal generated by the basis $\langle u_0, u_1, \ldots, u_{m-1} \rangle$.*

Proof. The proof is by induction. We denote by (v_n) the bi-ideal sequence generated by the basis $\langle u_0, u_1, \ldots, u_{m-1} \rangle$. Since $v_0 = 1$ and $v_1 = 1101$, then obviously $00 \notin \mathrm{F}(v_0)$ and $00 \notin \mathrm{F}(v_1)$. We assume that $00 \notin \mathrm{F}(v_i)$ for all $i \leq k$ and consider v_{k+1}.

We recall that by Definition 1 of a finitely generated bi-ideal $v_{k+1} = v_k u_j v_k$, where $j = k+1 \pmod{m}$. According to Lemma 5 we have $1 = v_0 \in \mathrm{Pref}(v_k) \cap \mathrm{Suff}(v_k)$. From here, assumption that $00 \notin \mathrm{F}(v_k)$, and the given $00 \notin \mathrm{F}(u_j)$, we obtain the necessary, namely, $00 \notin \mathrm{F}(v_{k+1})$.

□

Theorem 30. *Let $m \in \mathbb{N}$, $m \geq 2$. If $u_0 = 1$, $u_1 = 10$ and $00 \notin F(u_i)$ for all $i \in \overline{2, m-1}$ (if $m > 2$), then the bi-ideal generated by the basis $\langle u_0, u_1, \ldots, u_{m-1} \rangle$ is a universal bi-ideal.*

Proof. The beginning of the proof is similar to the proof of Proposition 26. Let y be the bi-ideal generated by the m-tuple $\langle u_0, u_1, \ldots, u_{m-1} \rangle$. Assume on contrary that y is not a universal bi-ideal. Then there exists a non-trivial periodic word $x = u^\omega$ with $|u| = p \geq 2$, such that $\mathbf{SHR}_y(x)$ is a ultimately periodic word with period T and pre-period T_1, i.e., $\mathbf{SHR}_y(x) = v'v^\omega$, where $|v| = T$ and $|v'| = T_1$.

According to Lemma 24 we can choose sufficiently large positive integers α, β, γ and γ', where $\alpha + 1 < \beta < \gamma$, such that $|v_{\alpha m-1}|_1 > T_1$ and

$$|v_{\alpha m-1}| \equiv |v_{\beta m-1}| \equiv |v_{\gamma m-1}| \equiv |v_{\gamma' m-1}| \pmod p,$$
$$|v_{\alpha m-1}|_1 \equiv |v_{\beta m-1}|_1 \equiv |v_{\gamma m-1}|_1 \equiv |v_{\gamma' m-1}|_1 \pmod T,$$
$$|v_{\gamma' m-1}| > |v_{\gamma m-1}| > |v_{\beta m-1}| > |v_{\alpha m-1}| > 3p,$$
$$|v_{\gamma' m-1}|_1 > |v_{\gamma m-1}|_1 > |v_{\beta m-1}|_1 > |v_{\alpha m-1}|_1 > T,$$

which implies

$$|v_{\beta m-1}| - |v_{\alpha m-1}| = kp, \tag{2.18}$$
$$|v_{\beta m-1}|_1 - |v_{\alpha m-1}|_1 = k_1 T. \tag{2.19}$$

for some $k, k_1 \in \mathbb{N}_+$.

Analogously, as in the proof of Proposition 26 we obtain $|v_{\beta m}| - |v_{\alpha m}| = 2kp$

$$|v_{\beta m}| - |v_{\alpha m}| = 2kp, \tag{2.110}$$
$$|y[|v_{\beta m-1}| - |v_{\alpha m-1}|, |v_{\beta m}| - |v_{\alpha m}|)| = kp, \tag{2.111}$$
$$|y[|v_{\beta m-1}| - |v_{\alpha m-1}|, |v_{\beta m}| - |v_{\alpha m}|)|_1 = k_1 T. \tag{2.112}$$

Again, as $\mathbf{SHR}_y(x)$ is ultimately periodic with preperiod T_1 and period T, then

$$\mathbf{SHR}_y(x)[i] = \mathbf{SHR}_y(x)[i + k_1 T]$$

for all $i \in \overline{T_1, k_1 T - 1}$, and the following holds: if there exists integer $\delta \in \overline{1, kp - |v_{\alpha m-1}|}$ such that

$$|y[kp - \delta, kp)|_1 = |y[2kp - \delta, 2kp)|_1,$$

then

$$\mathbf{SHR}_{y[kp-\delta,kp)}(x[kp - \delta, kp)) = \mathbf{SHR}_{y[2kp-\delta,2kp)}(x[2kp - \delta, 2kp)). \tag{2.113}$$

We will show that we can set $\delta = |u_1 v_{\alpha m-1} u_0| = |v_{\alpha m-1}| + 3$. Indeed, as $\alpha + 1 < \beta$ and $v_{\alpha m+1}$ is in suffix of both $v_{\beta m-1}$ and $v_{\beta m}$ (by Lemma 5), then (see Figure 2.3)

$$y[kp - |v_{\alpha m-1}| - 3, kp) = u_1 v_{\alpha m-1} u_0 = 10 v_{\alpha m-1} 1, \tag{2.114}$$
$$y[2kp - |v_{\alpha m-1}| - 3, 2kp) = u_0 v_{\alpha m-1} u_1 = 1 v_{\alpha m-1} 10. \tag{2.115}$$

Clearly, $|10v_{\alpha m-1}1|_1 = |1v_{\alpha m-1}10|_1$, therefore if we set $\delta = |v_{\alpha m-1}|+3$, then (2.113) holds.

From periodicity of x and (2.111) we have $x[kp-\delta, kp) = x[2kp-\delta, 2kp)$, therefore if we set $x[kp-\delta, kp) = abwcd$, where $a, b, c, d \in \{0,1\}$ and $w \in \{0,1\}^{\delta-4}$, then (2.113) can be rewritten as (see Figure 2.4)

$$\mathbf{SHR}_{10v_{\alpha m-1}1}(abwcd) = \mathbf{SHR}_{1v_{\alpha m-1}10}(abwcd). \tag{2.116}$$

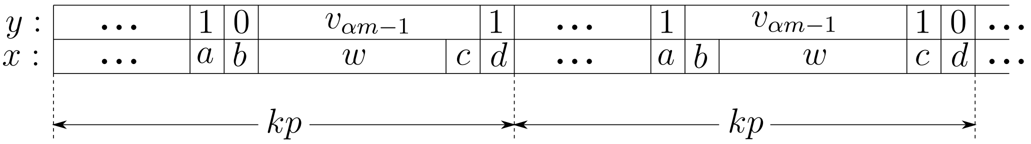


Figure 2.4: Structure of the shrinking

As one can see in Figure 2.4 , we have $\mathbf{SHR}_{10v_{\alpha m-1}1}(abwcd) = a\mathbf{SHR}_{v_{\alpha m-1}}(wc)d$ and $\mathbf{SHR}_{1v_{\alpha m-1}10}(abwcd) = a\mathbf{SHR}_{v_{\alpha m-1}}(bw)c$, thus $c = d$ and we can consider only

$$\mathbf{SHR}_{v_{\alpha m-1}}(wc) = \mathbf{SHR}_{v_{\alpha m-1}}(bw). \tag{2.117}$$

As α was chosen sufficiently large, then bwc contains both letters (zeros and ones), therefore its period $T' > 1$ and Lemma 28 can be applied. According to Lemma 28 words bw and wc are not equal. This implies the existence of integer $i \in \overline{1, |v_{\alpha m-1}|}$ such that in the word $w' = bwc$ we have

$$\forall j \leq i \ \ (w'[j-1] = w'[j] \wedge w'[i-1] \neq w'[i]) \tag{2.118}$$

From (2.117) we also have

$$\mathbf{SHR}_{v_{\alpha m-1}}(w'[0, |v_{\alpha m-1}|)) = \mathbf{SHR}_{v_{\alpha m-1}}(w'[1, |v_{\alpha m-1}|]). \tag{2.119}$$

Now, if i is the index satisfying (2.118), then $v_{\alpha m-1}[i] = 1$ leads to contradiction to (2.117). In other words, we can not have $v_{\alpha m-1}[i] = 1$ since it would return two distinct letters in the same position of two equal shrunk words (see Figure 2.5). Thus $v_{\alpha m-1}[i] = 0$. For the same reason – if $w'[t] \neq w'[t+1]$, $t \in \overline{i, |w'|-1}$, then $v_{\alpha m-1}[t] = 0$. As x is non-trivial and α was chosen sufficiently large, then there are several such indexes t in the word $w' = bwc$. The periodicity of x implies periodicity of such indexes t, therefore for all $\eta \in \mathbb{N}$

$$(w'[t] \neq w'[t+1] \wedge t + \mu p < |v_{\alpha m-1}| + 1) \Rightarrow v_{\alpha m-1}[t+\mu p] = v_{\alpha m-1}[t] = 0, \tag{2.120}$$

i.e., there are zeros in $v_{\alpha m-1}$ repeating periodically with period p.

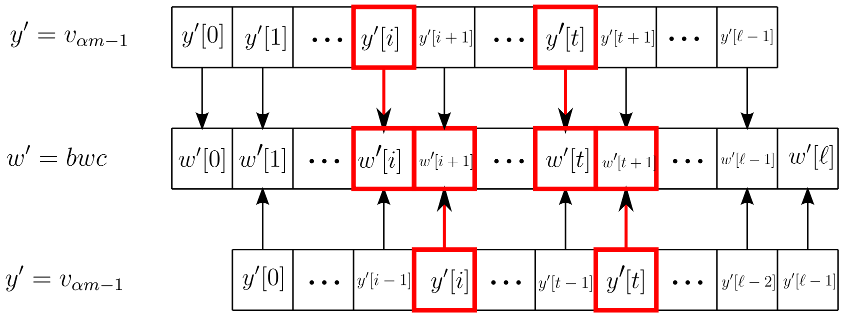

$y' = v_{\alpha m-1}$
$y'[0]$ $y'[1]$ … $y'[i]$ $y'[i+1]$ … $y'[t]$ $y'[t+1]$ … $y'[\ell-1]$
$w' = bwc$
$w'[0]$ $w'[1]$ … $w'[i]$ $w'[i+1]$ … $w'[t]$ $w'[t+1]$ … $w'[\ell-1]$ $w'[\ell]$
$y' = v_{\alpha m-1}$
$y'[0]$ … $y'[i-1]$ $y'[i]$ … $y'[t-1]$ $y'[t]$ … $y'[\ell-2]$ $y'[\ell-1]$

Figure 2.5: Here we have put in one Figure how the equal shrunk words in (2.117) are obtained. We denote $\ell = |v_{\alpha m-1}|$ and $y' = v_{\alpha m-1}$. As one can see, if $w'[i] \neq w'[i+1]$(or $w'[t] \neq w'[t+1]$), and $y'[i] = 1$ (or $y'[t] = 1$), then we obtain two distinct letters in the same positions of equal shrunk words, contradicting to (2.117).

Similarly, if we consider $v_{\gamma m-1}$ and $v_{\gamma' m-1}$ (instead of $v_{\alpha m-1}$ and $v_{\beta m-1}$), we obtain that there are zeros in $v_{\gamma m-1}$ that repeat periodically with period p. From here and $|v_{\gamma m-1}| > |v_{\alpha m-1}| > p$ we obtain the existence of an index $i_0 < p$ such that for all $\eta \in \mathbb{N}$

$$i_0 + \eta p \leq |v_{\gamma m-1}| \Rightarrow v_{\gamma m-1}[i_0] = v_{\gamma m-1}[i_0 + \eta p] = 0. \tag{2.121}$$

According to Lemma 5 $v_{\alpha m-1}$ is the prefix of $v_{\gamma m-1}$, hence for all $\eta \in \mathbb{N}$

$$i_0 + \eta p \leq |v_{\alpha m-1}| \Rightarrow v_{\alpha m-1}[i_0] = v_{\alpha m-1}[i_0 + \eta p] = 0.$$

Now, $\alpha < \beta < \gamma$ and $m \geq 2$ imply $|v_{\alpha m-1}| < |v_{\beta m-1}| < |v_{\beta m}| \leq |v_{\gamma m-1}|$. From here and (2.121) we have (see Figure 2.6)

$$v_{\gamma m-1}[i_0] = v_{\gamma m-1}[i_0 + kp] = v_{\gamma m-1}[i_0 + 2kp] = 0$$

and

$$v_{\gamma m-1}[i_0] = v_{\gamma m-1}[i_0 + (k-2)p] = v_{\gamma m-1}[i_0 + (2k-2)p] = 0. \tag{2.122}$$

As $i_0 < p$, $|v_{\alpha m-1}| > 3p$ and $p \geq 2$, then $kp - |v_{\alpha m-1}| < i_0 + (k-2)p < kp - 1$ and $2kp - |v_{\alpha m-1}| < i_0 + (2k-2)p < 2kp - 2$. Next, from (2.114) we know that

$$v_{\alpha m-1} = y[kp - |v_{\alpha m-1}| - 1, kp - 1),$$

while (2.115) implies

$$v_{\alpha m-1} = y[2kp - |v_{\alpha m-1}| - 2, kp - 2).$$

From here $y[i_0 + (k-2)p] = v_{\alpha m-1}[r] = 0$, while $y[i_0 + (2k-2)p] = v_{\alpha m-1}[r+1] = 0$ for some $r < |v_{\alpha m-1}| - 1$ (see Figure 2.6).

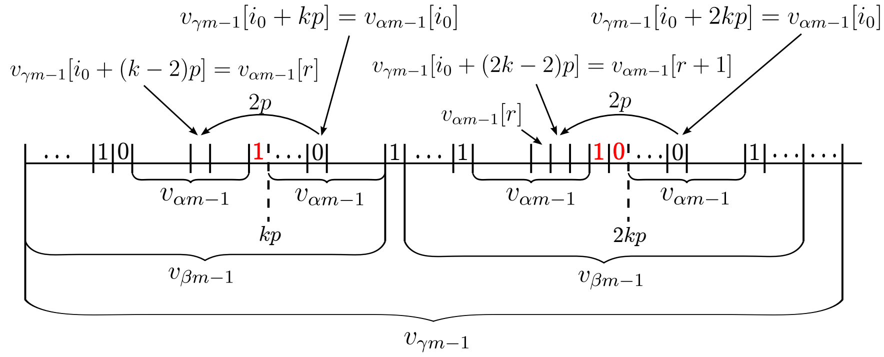

$v_{\gamma m-1}[i_0 + kp] = v_{\alpha m-1}[i_0]$
$v_{\gamma m-1}[i_0 + 2kp] = v_{\alpha m-1}[i_0]$
$v_{\gamma m-1}[i_0 + (k-2)p] = v_{\alpha m-1}[r]$
$v_{\gamma m-1}[i_0 + (2k-2)p] = v_{\alpha m-1}[r+1]$
$2p$
$v_{\alpha m-1}[r]$
$2p$
... 1 0 1 ... 0 1 ... 1 1 0 ... 0 1 ... ...
$v_{\alpha m-1}$
$v_{\alpha m-1}$
$v_{\alpha m-1}$
$v_{\alpha m-1}$
kp
$2kp$
$v_{\beta m-1}$
$v_{\beta m-1}$
$v_{\gamma m-1}$

Figure 2.6: The structure of $v_{\gamma m-1}$.

We have obtained that $v_{\alpha m-1}[r, r+1] = 00$ but according to Lemma 29 the finitely generated bi-ideal y does not contain 00 as a factor. Contradiction. Hence shrunk word $\mathbf{SHR}_y(x)$ is aperiodic for any non-trivial periodic word x. Thus y is a universal bi-ideal.

□

2.2 Filtering problem and its connection to the shrinking generator

In this Section we consider closure properties of regular languages, more precisely – operations that preserve regularity. It turns out that so called *filtering problem* (see (Berstel et al., 2006)) has connection to construction of shrinking generator.

As per (Berstel et al., 2006) a finite or infinite strictly increasing sequence of non-negative integers $(s_n)_{n\geq 0}$ is called a *filter*. Filtering finite word $u = a_0a_1\cdots a_{n-1}$ (respectively, infinite word $x = x_0x_1\ldots x_n\cdots$) by finite (or infinite) filter s consists in deleting the letters a_i (respectively, x_i) such that i is not in the set $\{s_0, s_1, \ldots\}$. By $\mathrm{FT}_s(u)$ we denote the word that is obtained, when u is filtered by s, i.e., $\mathrm{FT}_s(u) = a_{s_0}a_{s_1}\cdots a_{s_k}$, where k is the largest integer such that $s_k \leq n < s_{k+1}$. Similarly, if infinite word x is filtered with filter s we denote obtained word by $\mathrm{FT}_s(x)$, i.e., $\mathrm{FT}_s(x) = x_{s_0}x_{s_1}\cdots x_{s_k}\cdots$.

If L is a language, then by $\mathrm{FT}_s(L)$ we denote the set of all words of L filtered by s, .i.e.,

$$\mathrm{FT}_s(L) = \{\mathrm{FT}_s(u) \big| u \in L\}.$$

We say that filter *s preserves regularity* if, for every regular language L, the language $\mathrm{FT}_s(L)$ is also regular. Then the *filtering problem* is to characterize the regularity-preserving filters.

The *(first) differential sequence* of an integer sequence $(s_n)_{n\geq 0}$ is the sequence ∂s defined by

$$(\partial s)_n = s_{n+1} - s_n.$$

A sequence is called *syndetic* if its differential sequence is bounded. If S is infinite subset of $\mathbb{N}$, the *enumerating sequence* of S is the unique strictly increasing sequence $(s_n)_{n\geq 0}$ such that

$$S = \left\{ s_n \middle| \, n \geq 0 \right\}.$$

The differential sequence of this sequence is called the *differential sequence* of S. The *characteristic sequence* of a subset S of $\mathbb{N}$ is the sequence $(c_n)_{n\geq 0}$ defined by

$$c_n = \begin{cases} 1 & \text{if } n \in S, \\ 0 & \text{otherwise} . \end{cases}$$

Here we see the connection between shrinking generator and filtering – the S-sequence in shrinking generator is the characteristic sequence of S. Next, we give some more results and definitions on sequences and in Subsection **??** we solve filtering problem for filters which's characteristic sequence is finitely generated bi-ideal.

Proposition 31 (see, e.g., (Berstel et al., 2006))**.** *Let S be an infinite set of non-negative integers. The following conditions are equivalent:*

1. *the characteristic sequence of S is ultimately periodic,*

2. *the differential sequence of S is ultimately periodic.*

A sequence s of non-negative integers is called *ultimately periodic modulo p* if there exist two integers $m \geq 0$ and $r > 0$ such that, for each $n \geq m$, $u_n \equiv u_{n+r} \pmod{p}$. A sequence is called *cyclically ultimately periodic* if it is ultimately periodic modulo p for every $p > 0$.

A sequence s of non-negative integers is called *ultimately periodic threshold t* if there exist two integers $m \geq 0$ and $r > 0$ such that, for each $n \geq m$, $\min(u_n, t) = \min(u_{n+r}, t)$.

We say that (see Proposition 4.4. in (Berstel et al., 2006)) a sequence of non-negative integers is *residually ultimately periodic* if and only if it is cyclically ultimately periodic and ultimately periodic threshold t for all $t \geq 0$.

Proposition 32 ((Berstel et al., 2006), Prop. 5.4.)**.** *If s is a syndetic sequence of non-negative integers, then the following conditions are equivalent:*

1. s is residually ultimately periodic,

2. ∂s is residually ultimately periodic,

3. ∂s is ultimately periodic.

An integer sequence is called *differentially residually ultimately periodic* if its differential sequence is residually ultimately periodic.

Theorem 33. (Berstel et al., 2006) *A filter preserves regularity if and only if it is differentially residually ultimately periodic.*

Proposition 34. *If s is the enumerating sequence of a set $S \subset \mathbb{N}$ and the characteristic sequence of S is a uniformly recurrent word x, then s is regularity-preserving filter if and only if x is periodic.*

Proof. $\Rightarrow$: If s is regularity-preserving filter, then by Theorem 33 the sequence s is differentially residually ultimately periodic, i.e., its differential sequence ∂s is residually ultimately periodic.

By the definition of uniformly recurrent words we obtain that the distance between two following letters “1” in x is bounded, therefore ∂s is bounded, i.e., s is syndetic.

Now, since s is syndetic and ∂s is residually ultimately periodic, then by Proposition 32 differential sequence ∂s is ultimately periodic. Then by Propostion 31 the characteristic sequence, i.e., the uniformly recurrent word x is ultimately periodic.

Finally, Lemma 4 and Proposition 3 imply the periodicity of x.

$\Leftarrow$: If s the enumerating sequence of a set $S \subset \mathbb{N}$ but the characteristic sequence of S is infinite periodic word x, then, in order to prove that s is regularity-preserving filter, we have to show that it is differentially residually ultimately periodic, i.e., ∂s is residually ultimately periodic. By Proposition 31 ∂s is ultimately periodic. Clearly, then ∂s is also syndetic. Now, from Proposition 32 we obtain that ∂s is residually ultimately periodic.

□

Corollary 35. *If s is the enumerating sequence of a set $S \subset \mathbb{N}$ and the characteristic sequence of S is a finitely generated or bounded bi-ideal x, then s is regularity-preserving filter if and only if x is periodic.*

Proof. Finitely generated bi-idelas and bounded bi-ideals are subclasses of the class of uniformly recurrent words (Buls and Lorencs, 2006). □

3 Arithmetical subsequences of finitely generated bi-ideals and bounded bi-ideals

In Subsection 3.1.1 we give a necessary and sufficient condition of aperiodicity of all arithmetical subsequences of a finitely generated bi-ideal (Theorem 40). In Subsection 3.1.2 we give an algorithm for checking whether all arithmetical subsequences of a finitely generated bi-ideal are aperiodic. Section 3.2 contains a necessary and sufficient condition of the aperiodicity of arithmetical subsequences of a bounded generated bi-ideal. In Section 3.3 we give some connection of property having all arithmetical subsequences aperiodic and property being a universal bi-ideal.

3.1 Aperiodicity of all Arithmetical Subsequences of a Finitely Generated Bi-ideal

3.1.1 Necessary and Sufficient Condition of Aperiodicity of all Arithmetical Subsequences of a Finitely Generated Bi-ideal

Lemma 36 (See, e.g., (Avgustinovich et al., 2003))**.** *An arithmetical subsequence of a uniformly recurrent word is uniformly recurrent.*

Corollary 37. *If x is finitely generated bi-ideal or bounded bi-ideal and $x_p^l = 1^\omega$, where $l \geq p$ and $1^\omega = 111\cdots 1\cdots$, then $x_p^{l-p} = 1^\omega$.*

Proof. It follows from Lemma 36, Lemma 4 and the fact that both the class of finitely generated bi-ideals and the class of bounded bi-ideals are subclasses of the class of uniformly recurrent words.

□

Lemma 38. *If*

$$i_{0,0}, i_{0,1}, ..., i_{0,p-1}$$

$$i_{1,0}, i_{1,1}, ..., i_{1,p-1}$$

$$\dots\dots\dots\dots\dots$$

$$i_{k-1,0}, i_{k-1,1}, ..., i_{k-1,p-1}$$

are integers, where

1. $i_{0,0}, i_{1,0}, \dots, i_{k-1,0}$ *form a complete residue system modulo k,*

2. $\forall s \in \overline{0, k-1} \forall \sigma \in \overline{1, p-1} \, (i_{s,\sigma} \equiv i_{s,\sigma-1} - k \pmod p)$*),*

then it is possible to choose p numbers from them, which form complete residue system modulo p.

Proof. We assume assume on contrary the existence of an integer j such that none of the numbers in the list is congruent to j modulo p. According to the first condition of the theorem we can choose $s \in \overline{0, k-1}$ such that

$$i_{s,0} \equiv j \pmod k.$$

Hence, there exists $t \in \mathbb{Z}$ such that $j = i_{s,0} + tk$, thus

$$j = i_{s,0} + (\sigma + t'p)k$$

for some t' and σ, where $0 \leq \sigma < p$.

Observe that second condition of the theorem can be expressed as

$$i_{s,\sigma} \equiv i_{s,0} + \sigma k \pmod p,$$

therefore

$$j \equiv i_{s,\sigma} \pmod p,$$

which leads to a contradiction. □

Now we are ready to prove a necessary and sufficient condition of aperiodicity of all arithmetical subsequences of a finitely generated bi-ideal over an alphabet $\Sigma_n = \{0, 1, \dots, n-1\}$, $n \geq 2$.

By $K_x^{a,k}$ we denote the set of all positions modulo $k \in \mathbb{N}_+$ of $a \in alph(x)$ in a word x, i.e., $K_x^{a,k} = \left\{ i \bmod k \middle| \, x[i] = a, i \in \mathbb{N} \right\}$. Let $\breve{K}_x^{a,k}$ be the set of all positions modulo $k \in \mathbb{N}_+$ of letters of the set $alph(x) \setminus \{a\}$ in a word x, i.e., $\breve{K}_x^{a,k} = \left\{ i \bmod k \middle| \, x[i] \in alph(x) \setminus \{a\}, i \in \mathbb{N} \right\}$.

Theorem 39. *All arithmetical subsequences of a finitely generated bi-ideal $x \in \Sigma_n^\omega$, with $n \geq 2$, are aperiodic if and only if there exists a basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ of x such that $|u_0| \neq |u_1|$, and for each $a \in alph(x)$,*

$$Card\left(\breve{K}_{u_0}^{a,k}\right) = k,$$

where $k = ||u_0| - |u_1||$.

Proof. $\Rightarrow$: Let x be a finitely generated bi-ideal with all arithmetical subsequences aperiodic. Let $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ be a basis of x. Let $\langle k_0, k_1, \ldots, k_{m-1}\rangle$ be the length-differential m-tuple of the basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$.

Firstly, we observe that there exists at least one non-zero k_j, $j \in \overline{0, m-1}$. Indeed, otherwise all basis words $u_0, u_1, \ldots, u_{m-1}$ are of the same length. As by the bi-ideal construction x can be expressed in a form

$$x = v_0 u_* v_0 u_* v_0 u_* v_0 \cdots = u_0 u_* u_0 u_* u_0 u_* u_0 \cdots ,$$

where u_* is the corresponding basis word, then x contains periodic arithmetical subsequence x_p^l with $l \in \overline{0, |u_0| - 1}$ and $p = 2 \cdot |u_0|$.

We set $\mathcal{K} = \left\{|k_j| \middle| j \in \overline{0, m-1}, k_j \neq 0\right\}$. Now, if all arithmetical subsequences of the bi-ideal x are aperiodic, then so are all arithmetical subsequences with difference $k \in \mathcal{K}$ that start at an arbitrary position l.

Therefore for each $k \in \mathcal{K}$, each $l_k \in \overline{0, k-1}$, and each $a \in alph(x)$ we have

$$x_k^{l_k} \neq a^\omega. \tag{3.11}$$

Hence for each $k \in \mathcal{K}$, for each starting position $l_k \in \overline{0, k-1}$ of an arithmetical subsequence, and each letter $a \in alph(x)$, there exists non-negative integer δ_{l_k} such that

$$x[l_k + k\delta_{l_k}] \in alph(x) \setminus \{a\}. \tag{3.12}$$

Next, we choose $\gamma \in \mathbb{N}$ such that, for each letter $a \in alph(x)$, each $k \in \mathcal{K}$ and each $l_k \in \overline{0, k-1}$,

$$\left|u_0^{(\gamma)}\right| > l_k + k\delta_{l_k}.$$

Now, we observe that for all $k \in \mathcal{K}$, for all letters $a \in alph(x)$, we have

$$Card\left(\breve{K}_{u_0^{(\gamma)}}^{a,k}\right) = k. \tag{3.13}$$

Otherwise there would exist a position $l_k \in \overline{0, k-1}$ for some $k \in \mathcal{K}$ such that (3.12) does not hold.

We might have $\left||u_0^{(\gamma)}| - |u_1^{(\gamma)}|\right| = 0$. However, according to Lemma 18 we can L-prolong basis words several times, say $\xi \in \mathbb{N}_+$, until we have $|u_0^{(\gamma+\xi)}| \neq |u_1^{(\gamma+\xi)}|$. The basis

$$\langle u_0^{(\gamma+\xi)}, u_1^{(\gamma+\xi)}, \dots, u_{m-1}^{(\gamma+\xi)} \rangle$$

satisfies conditions of the Theorem.

$\Leftarrow$: Let x be a finitely generated bi-ideal with a basis $\langle u_0, u_1, \dots, u_{m-1} \rangle$ such that $|u_0| \neq |u_1|$, and for each $a \in alph(x)$

$$Card\left(\breve{K}_{u_0}^{a,k}\right) = k,$$

where $k = ||u_0| - |u_1||$.

We proceed by contradiction: suppose x contains an ultimately periodic arithmetical subsequence. Clearly, then there also exists an arithmetical subsequence, which contains only one letter. Without loss of generality we can assume that x contains arithmetical subsequence $x_p^l = 1^\omega$, where $p > 1$ (the case when $p = 1$ is trivial). According to Corollary 37 we can ensure that $l < p$. Hence for each $i \in \mathbb{N}$

$$x[l + i \cdot p] = 1,$$

therefore

$$Card\left(\breve{K}_x^{1,p}\right) < p. \tag{3.14}$$

Let $\sigma : \Sigma_n \mapsto \{0, 1\}$ be a morphism defined by

$$\sigma(a) = \begin{cases} 1, & \text{if } a = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Let $y = \sigma(x)$. Clearly, y is a finitely generated bi-ideal with a basis $\langle u'_0, u'_1, \dots, u'_{m-1} \rangle$ such that for all $i \in \overline{0, m-1}$, u'_i is obtained from u_i by replacing all letters $2, 3, \dots, n-1$ by 0. Clearly, $|u_i| = |u'_i|$ for all $i \in \mathbb{N}$. By conditions of the theorem, positions of zeros and ones form a complete residue system modulo k in u'_0, while (3.14) implies that positions of zeros do not form a complete residue system modulo p in the finitely generated bi-ideal y. Further we consider the bi-ideal y and obtain a contradiction by proving that positions of zeros form a complete residue system modulo p in y.

Let (v_i) be the bi-ideal sequence generated by the basis $\langle u_0, u_1, \dots, u_{m-1} \rangle$. Then by Lemma 24 there exist $\alpha, \beta \in \mathbb{N}$, $0 < \alpha < \beta$, such that

$$|v_{\alpha m-1}| \equiv |v_{\beta m-1}| \pmod{p}. \tag{3.15}$$

Moreover, we can choose α and β such that $v_{\alpha m-1}$ and hence $v_{\beta m-1}$ contain all possible positions of zeros modulo p.

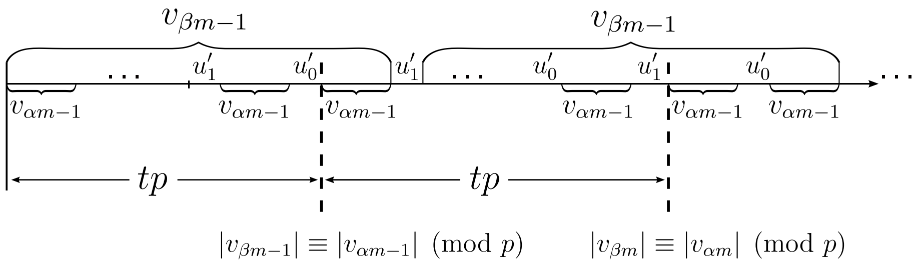


Figure 3.1: Construction of a bi-ideal y.

According to (3.15) there exists a non-negative integer t such that (see Figure 3.1)

$$|v_{\beta m-1}| - |v_{\alpha m-1}| = tp \tag{3.16}$$

The inequality $0 < \alpha < \beta$ implies $u'_0 \in \text{Pref}(v_{\alpha m-1})$ and $v_{\alpha m-1} \in \text{Pref}(v_{\beta m-1})$, therefore positions of zeros and ones form a complete residue system modulo k in words $v_{\alpha m-1}$ and $v_{\beta m-1}$. From here we obtain the existence of positions of zeros $i_{0,0}, i_{1,0}, \dots, i_{k-1,0}$, where $tp - |u'_0| - |v_{\alpha m-1}| \le i_{0,0}, i_{1,0}, \dots, i_{k-1,0} < tp - |u'_0|$, that form a complete residue system modulo k in the bi-ideal y. Hence positions $i_{0,0}, i_{1,0}, \dots, i_{k-1,0}$ satisfy the first condition of Lemma 38.

Observe, if for each $i \in \{i_{0,0}, i_{1,0}, \dots, i_{k-1,0}\}$ and for each integer $j \in \overline{0, k-1}$ there exists a position i_j, $tp - |u'_0| - |v_{\alpha m-1}| \le i_j < tp - |u_0|$, such that

$$i_j \equiv i - j \cdot k \pmod{p} \text{ and } y[i_j] = 0, \tag{3.17}$$

then positions $i_{0,0}, i_{1,0}, \dots, i_{k-1,0}$ together with (3.17) satisfy both conditions of Lemma 38. Hence, if (3.17) holds, then positions of zeros form a complete residue system modulo p in the bi-ideal y, contradicting to our assumption that y (and also x) contains periodic arithmetical subsequence 1^ω.

In the rest of the proof we show that, indeed, (3.17) holds. We consider the case when $|u'_1| = |u'_0| + k$ (the second case can be proved analogously) and prove that positions of zeros form a complete residue system modulo p in y.

Let $i \in \{i_{0,0}, i_{1,0}, \dots, i_{k-1,0}\}$. Then $y[i] = 0$. From (3.16) and the bi-ideal construction we obtain the following: If $y[i] = 0$, where $tp - |u'_0| - |v_{\alpha m-1}| \le i < tp - |u'_0|$, then (see Figure 3.2)

$$y[i + tp - k] = 0.$$

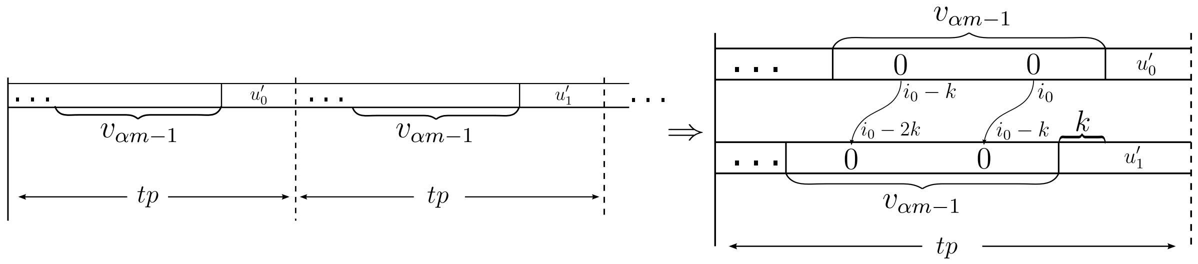


Figure 3.2: Shift for k positions of the factor $v_{\alpha m-1}$ in word y.

We set $i_0 = i$ and observe that i_0 satisfies (3.17) for $j = 0$.

Now, $i_0+tp-k \equiv i_0-k \pmod{p}$ and assumption that $v_{\alpha m-1}$ contains all possible positions of zeros modulo p imply the existence of a position i_1, where $tp-|u'_0|-|v_{\alpha m-1}| \leq i_1 < tp-|u'_0|$, such that

$$i_1 \equiv i_0 - k \pmod{p} \text{ and } y[i_1] = 0.$$

From (3.15) it follows that $y[i_1 + tp - k] = 0$. Observe, that

$$i_1 + tp - k \equiv i_1 - k \equiv i_0 - 2k \pmod{p}.$$

Further, $i_1 + tp - k \equiv i_0 - 2k \pmod{p}$ and assumption that $v_{\alpha m-1}$ contains all possible positions of zeros modulo p implies the existence of a position i_2, where $tp - |u'_0| - |v_{\alpha m-1}| \leq i_2 < tp - |u'_0|$, such that

$$i_2 \equiv i_0 - 2k \pmod{p} \text{ and } y[i_2] = 0.$$

Then (3.15) implies $y[i_2 + tp - k] = 0$.

Analogously, we can obtain that for each $j \in \overline{0, p-1}$ there exists a position

$$i_j \in \overline{tp - |u_j| - |v_{\alpha m-1}|, tp - |u_j| - 1}$$

such that (3.17) holds.

□

Corollary 40. *All arithmetical subsequences of a finitely generated bi-ideal $x \in \{0,1\}^\omega$ are aperiodic if and only if there exists basis $\langle u_0, u_1, ..., u_{m-1}\rangle$ such that $|u_0| \neq |u_1|$ and positions of zeros and ones form complete residue system modulo $k = ||u_0| - |u_1||$ in the word u_0.*

Proof. If $n = 2$, then $\Sigma_n = \{0, 1\}$. Therefore $\breve{K}^{0,k}_{u_0} = K^{1,k}_{u_0}$ and $\breve{K}^{1,k}_{u_0} = K^{0,k}_{u_0}$. In this case

$$Card\left(K^{1,k}_{u_0}\right) = Card\left(K^{0,k}_{u_0}\right) = k$$

is equivalent to the condition: positions of zeros and ones form a complete residue system modulo k in the word u_0. □

Corollary 41. *Let x be a finitely generated bi-ideal over alphabet Σ_n, $n \geq 2$. If there exists a basis $\langle u_0, u_1, ..., u_{m-1}\rangle$ of x such that $|u_0| \neq |u_1|$ and, for two distinct letters $a, b \in alph(x)$, positions of a and b form a complete residue system modulo $k = ||u_0| - |u_1||$ in the word u_0, then all arithmetical subsequences of x are aperiodic.*

Proof. We observe that the basis $\langle u_0, u_1, ..., u_{m-1}\rangle$ of the finitely generated bi-ideal x satisfies conditions of Theorem 39. Indeed, since $b \in alph(x) \setminus \{a\}$, then $Card\left(\breve{K}^{a,k}_{u_0}\right) = k$. For all letters $c \in alph(x) \setminus \{a\}$ we have $a \in alph(x) \setminus \{c\}$. Hence for all $c \in alph(x) \setminus \{a\}$ the cardinality of the set $\breve{K}^{c,k}_{u_0}$ equals k. □

In one direction we can prove even stronger result.

Theorem 42. *Let $x \in \Sigma_n^\omega$ be a finitely generated bi-ideal with basis $\langle u_0, u_1, ..., u_{m-1}\rangle$ and length-differential m-tuple $\langle k_0, k_1, ..., k_{m-1}\rangle$. If all arithmetical subsequences of x are aperiodic, then there exists a non-negative integer ξ such that for each $j \in \overline{0, m-1}$, each non-zero $k \in \{|k_0|, |k_1|, ..., |k_{m-1}|\}$, and each letter $a \in alph(x)$,*

$$Card\left(\breve{K}^{a,k}_{u_j^{(\xi)}}\right) = k.$$

Proof. In the proof of Theorem 39 we already proved the existence of $\gamma \in \mathbb{N}$ such that for all $k \in \{|k_j| \mid k_j \neq 0, j \in \overline{0, m-1}\}$, for all letters $a \in \Sigma_n$, we have (see (3.13))

$$Card\left(\breve{K}^{a,k}_{u_0^{(\gamma)}}\right) = k.$$

According to definition of L-prolongation, for each $j \in \overline{0, m-1}$

$$u_j^{(\gamma+1)} = u_0^{(\gamma)} u_{j+1 \bmod m}^{(\gamma)}.$$

As, for all $j \in \overline{0, m-1}$ the basis word $u_0^{(\gamma)}$ is a prefix of the basis word $u_j^{(\gamma+1)}$, then the basis

$$\langle u_0^{(\gamma+1)}, u_1^{(\gamma+1)}, \ldots, u_{m-1}^{(\gamma+1)}\rangle$$

satisfies conditions of the theorem. □

Further we consider two examples — in both of them the given basis of the finitely generated bi-ideal does not satisfy conditions of Theorem 39. However, while the basis in Example 11 can be made to satisfy the conditions of Theorem 39 by applying some iterations of L-prolongation, the basis in Example 12 can not be made to satisfy Theorem 39. This has to do with the fact that the bi-ideal in Example 12 has a periodic arithmetical subsequence.

Example 11. Let x be a finitely generated bi-ideal generated with basis $\langle 1, 1012\rangle$, then $|k_0| = |1-4| = 3$. We L-prolong basis words 3 times:

$$\begin{array}{llllll} u_0^{(1)} = 11012 & \rightarrow & u_0^{(2)} = 1101211 & \rightarrow & u_0^{(3)} = 11012111101211012, \\ u_1^{(1)} = 11 & \rightarrow & u_1^{(2)} = 1101211012 & \rightarrow & u_1^{(3)} = 11012111101211. \end{array}$$

Now, we consider the positions of each letter modulo $3 = \left||u_0^{(3)}| - |u_1^{(3)}|\right|$ in $u_0^{(3)}$:

$$\begin{array}{cccccccccccccccccccc} u_0^{(3)} & = & 1 & 1 & 0 & 1 & 2 & 1 & 1 & 1 & 1 & 0 & 1 & 2 & 1 & 1 & 0 & 1 & 2, \\ \text{pos. mod } 3 & & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1. \end{array}$$

As one can see $K^{0,3}_{u_0^{(3)}} = \{0,2\}$, $K^{1,3}_{u_0^{(3)}} = \{0,1,2\}$ and $K^{2,3}_{u_0^{(3)}} = \{1,2\}$. Hence

$$\begin{aligned} \breve{K}^{0,3}_{u_0^{(3)}} &= K^{1,3}_{u_0^{(3)}} \cup K^{2,3}_{u_0^{(3)}} = \{0,1,2\} \\ \breve{K}^{1,3}_{u_0^{(3)}} &= K^{0,3}_{u_0^{(3)}} \cup K^{2,3}_{u_0^{(3)}} = \{0,1,2\} \\ \breve{K}^{2,3}_{u_0^{(3)}} &= K^{0,3}_{u_0^{(3)}} \cup K^{1,3}_{u_0^{(3)}} = \{0,1,2\} \end{aligned}$$

Since $Card(\breve{K}^{0,3k}_{u_0^{(3)}}) = Card(\breve{K}^{1,3}_{u_0^{(3)}}) = Card(\breve{K}^{2,3}_{u_0^{(3)}}) = 3$, we see that the basis $\langle u_0^{(3)}, u_1^{(3)}\rangle$ satisfies conditions of Theorem 39, therefore all arithmetical subsequences of x are aperiodic.

Example 12. Let $\langle 1, 10121\rangle$ be a basis of a finitely generated bi-ideal x. Then $|k_0| = |1-5| = 4 = |5-1| = |k_1|$. We L-prolong basis words 2 times:

$$\begin{array}{lll} u_0^{(1)} = 110121 & \rightarrow & u_0^{(2)} = 11012111, \\ u_1^{(1)} = 11 & \rightarrow & u_1^{(2)} = 110121110121. \end{array}$$

The length of both basis words $\left|u_0^{(2)}\right| = 8$ and $\left|u_1^{(2)}\right| = 12$ divide by 4, therefore we consider positions of each letter modulo $k = 4$ in basis words $u_0^{(2)}$ and $u_1^{(2)}$:

$$\begin{array}{cccccccccccccc} u_0^{(2)} & = & 1 & 1 & 0 & 1 & 2 & 1 & 1 & 1 & , & & & \\ u_1^{(2)} & = & 1 & 1 & 0 & 1 & 2 & 1 & 1 & 1 & 0 & 1 & 2 & 1, \\ \text{pos. mod } 4 & & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3. \end{array}$$

We have $K^{0,4}_{u_0^{(2)}} = \{0,2\}$, $K^{1,4}_{u_0^{(2)}} = \{0,1,2,3\}$ and $K^{2,4}_{u_0^{(2)}} = \{0,2\}$, therefore

$$Card(\breve{K}^{1,4}_{u_0^{(2)}}) = Card(\{0,2\}) = 2 < 4.$$

Now, one could think of changing the basis again, however by bi-ideal construction we know that x is an infinite concatenation of basis words. As the length of both new obtained basis words $u_0^{(2)}$ and $u_1^{(2)}$ divide by 4, then L-prolongation will not give us new positions modulo 4 of letters 0 and 2. Thus $Card(\breve{K}^{1,4}_{u_0^{(2)}}) = Card(\breve{K}^{1,4}_x) = 2$. Hence x contains ultimately periodic arithmetical subsequence.

3.1.2 Efficiency

The examples in previous subsection lead to the following questions – given a finitely generated bi-ideal x with basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ that does not satisfy the conditions of Theorem 39:

1. How many times do we have to L-prolong basis words to check whether all arithmetical subsequences of x are aperiodic or not?
2. Is there another way to change the basis (without using L-prolongation)?

The answer to the second question can be found in Definition 3 and Theorem 16. As one can see, the only useful way to change the basis is to use inverse operation of reduction 3, which in fact is inverse operation of L-prolongation. We will give an efficient algorithm (Algorithm 47) for deciding whether periodic arithmetical subsequence exists in x instead of answering the first question.

Theorem 43. *Let $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, where $m \geq 2$, be a basis of a finitely generated bi-ideal $x \in \Sigma_n^\omega$, $n \geq 2$. Let k be an arbitrary positive integer, and let α, s be positive integers such that*

1. $|v_{\alpha m-1}| \equiv \left|v_{(\alpha+s)m-1}\right| \pmod{k}$,
2. $K^{a,k}_{v_{\alpha m-1}} = K^{a,k}_{v_{(\alpha+s)m-1}}$ *for all* $a \in alph(x)$.

Then for each $j \in \mathbb{N}_+$, for each letter $a \in alph(x)$

$$K^{a,k}_{v_{\alpha m-1+j}} = K^{a,k}_{v_{(\alpha+s)m-1+j}} = K^{a,k}_{v_{\alpha m-1}}.$$

Proof. We will prove the theorem for the positions of zeros, i.e., for all $j \in \mathbb{N}_+$

$$K^{0,k}_{v_{\alpha m-1+j}} = K^{0,k}_{v_{(\alpha+s)m-1+j}} = K^{0,k}_{v_{\alpha m-1}}.$$

The proof for positions of any other letter is analogous.

1. First we prove that for all $j \in \overline{1, sm}$

$$K^{0,k}_{v_{\alpha m-1+j}} = K^{0,k}_{v_{(\alpha+s)m-1+j}} = K^{0,k}_{v_{\alpha m-1}}.$$

Observe that by bi-ideal construction $v_{\alpha m-1+j} \in \text{Pref}(v_{(\alpha+s)m-1})$ for each $j \in \overline{1, sm}$, therefore for each $j \in \overline{1, sm}$

$$K^{0,k}_{v_{\alpha m-1}} \subseteq K^{0,k}_{v_{\alpha m-1+j}} \subseteq K^{0,k}_{v_{(\alpha+s)m-1}} = K^{0,k}_{v_{\alpha m-1}},$$

hence for all $j \in \overline{1, sm}$

$$K^{0,k}_{v_{\alpha m-1+j}} = K^{0,k}_{v_{(\alpha+s)m-1}} = K^{0,k}_{v_{\alpha m-1}}. \tag{3.18}$$

Further we need to prove that $K^{0,k}_{v_{(\alpha+s)m-1+j}} = K^{0,k}_{v_{\alpha m-1}}$ for all $j \in \overline{1, sm}$.

From Corollary 25 and the first condition of the theorem we obtain

$$\left|v_{\alpha m-1+j}\right| \equiv \left|v_{(\alpha+s)m-1+j}\right| \pmod{k} \tag{3.19}$$

for all $j \in \mathbb{N}$, but by construction of a finitely generated bi-ideal

$$\begin{aligned} v_{\alpha m-1+j} &= v_{\alpha m-2+j} u_{j-1} v_{\alpha m-2+j}, \\ v_{(\alpha+s)m-1+j} &= v_{(\alpha+s)m-2+j} u_{j-1} v_{(\alpha+s)m-2+j} \end{aligned} \tag{3.110}$$

for each $j \in \overline{1, sm}$.

Further the proof is by induction on j.

- If $j = 1$, then $v_{(\alpha+s)m-1+j} = v_{(\alpha+s)m}$ and $v_{\alpha m-1+j} = v_{\alpha m}$. From here and (3.18) one gets $K^{0,k}_{v_{\alpha m}} = K^{0,k}_{v_{\alpha m-1}}$.

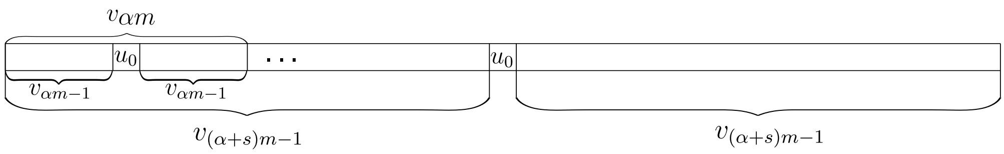


Figure 3.3: Construction of $v_{(\alpha+s)m}$.

Now we consider $K^{0}_{v_{(\alpha+s)m}}$ (see Figure 3.3). As $\left|v_{\alpha m-1}\right| \equiv \left|v_{(\alpha+s)m-1}\right| \pmod{k}$, then

$$\begin{aligned}K^{0,k}_{v_{(\alpha+s)m}} =&K^{0,k}_{v_{(\alpha+s)m-1}u_0v_{(\alpha+s)m-1}}\\ =&K^{0,k}_{v_{(\alpha+s)m-1}} \cup \{(i + |v_{(\alpha+s)m-1}|) \text{ mod } k \; \big| i \in K^{0,k}_{u_0}\}\cup\\ &\cup \{(i + |v_{(\alpha+s)m-1}| + |u_0|) \text{ mod } k \; \big| i \in K^{0,k}_{v_{(\alpha+s)m-1}}\}\\ =&K^{0,k}_{v_{\alpha m-1}} \cup \{(i + |v_{\alpha m-1}|) \text{ mod } k \; \big| i \in K^{0,k}_{u_0}\}\cup\\ &\cup \{(i + |v_{\alpha m-1}| + |u_0|) \text{ mod } k \; \big| i \in K^{0,k}_{v_{\alpha m-1}}\}\\ =&K^{0,k}_{v_{\alpha m-1}u_0v_{\alpha m-1}}\\ =&K^{0,k}_{v_{\alpha m}} = K^{0,k}_{v_{\alpha m-1}}.\end{aligned}$$

- Let $j < sm$. Assume that for all $t \leq j$ the following condition holds

$$K^{0,k}_{v_{(\alpha+s)m-1+t}} = K^{0,k}_{v_{\alpha m-1+t}} = K^{0,k}_{v_{\alpha m-1}}. \tag{3.111}$$

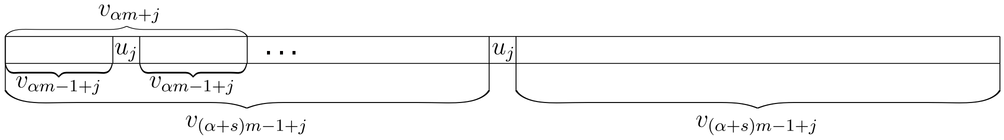


Figure 3.4: Construction of $v_{(\alpha+s)m+j}$.

Let us prove it also holds for $t = j + 1$ (see Figure 3.4). According to (3.19), $|v_{\alpha m-1+j}| \equiv \left|v_{(\alpha+s)m-1+j}\right| \pmod{k}$, therefore

$$\begin{aligned}K^{0,k}_{v_{(\alpha+s)m-1+j+1}} =&K^{0,k}_{v_{(\alpha+s)m+j}} = K^{0,k}_{v_{(\alpha+s)m-1+j}u_jv_{(\alpha+s)m-1+j}}\\ =&K^{0,k}_{v_{(\alpha+s)m-1+j}} \cup \{(i + |v_{(\alpha+s)m-1+j}|) \text{ mod } k \; \big| i \in K^{0,k}_{u_j}\}\cup\\ &\cup \{(i + |v_{(\alpha+s)m-1+j}| + |u_j|) \text{ mod } k \; \big| i \in K^{0,k}_{v_{(\alpha+s)m-1+j}}\}\\ =&K^{0,k}_{v_{\alpha m-1+j}} \cup \{(i + |v_{\alpha m-1+j}|) \text{ mod } k \; \big| i \in K^{0,k}_{u_j}\}\cup\\ &\cup \{(i + |v_{\alpha m-1+j}| + |u_j|) \text{ mod } k \; \big| i \in K^{0,k}_{v_{\alpha m-1+j}}\}\\ =&K^{0,k}_{v_{\alpha m-1+j}u_jv_{\alpha m-1+j}}\\ =&K^{0,k}_{v_{\alpha m+j}} = K^{0,k}_{v_{\alpha m-1}}.\end{aligned}$$

2. We have proved that the theorem holds for all $i \in \overline{1, sm}$. Let us prove it also holds for all $i > sm$. The proof is also inductive. We will show that for each $t \in \mathbb{N}$, for each $j \in \overline{1, sm}$

$$K^{0,k}_{v_{\alpha_t m-1+j}} = K^{0,k}_{v_{(\alpha_t+s)m-1+j}} = K^{0,k}_{v_{\alpha m-1}}, \tag{3.112}$$

where $\alpha_t = \alpha + t \cdot s$. One can see that (3.112) implies

$$K^{0,k}_{v_{\alpha m-1+t\cdot sm+j}} = K^{0,k}_{v_{(\alpha+s)m-1+t\cdot sm+j}} = K^{0,k}_{v_{\alpha m-1}},$$

for each $t \in \mathbb{N}$ and each $j \in \overline{1, sm}$. As each $i \in \mathbb{N}$ can be uniquely written as $i = t \cdot sm + j$ for some $t \in \mathbb{N}$ and $j \in \overline{1, sm}$, then (3.112) implies also

$$K^{0,k}_{v_{\alpha m-1+i}} = K^{0,k}_{v_{(\alpha+s)m-1+i}} = K^{0,k}_{v_{\alpha m-1}}, \tag{3.113}$$

where $i \in \mathbb{N}$.

- As $\alpha_0 = \alpha$, we have alredy proved the condition for $t = 0$ in Case (1).
- If $t = 1$, then $\alpha_1 = \alpha + s$. According to Case (1) if $j = sm$, then

  $$K^{0,k}_{v_{\alpha m-1+j}} = K^{0,k}_{v_{\alpha m-1+sm}} = K^{0,k}_{v_{(\alpha+s)m-1}}$$

  and

  $$K^{0,k}_{v_{(\alpha+s)m-1+j}} = K^{0,k}_{v_{(\alpha+s)m-1+sm}} = K^{0,k}_{v_{(\alpha+2s)m-1}},$$

  therefore

  $$K^{0,k}_{v_{(\alpha+2s)m-1}} = K^{0,k}_{v_{(\alpha+s)m-1}} = K^{0,k}_{v_{\alpha m-1}}. \tag{3.114}$$

  If we put $j = sm$ in (3.19), then we obtain

  $$|v_{(\alpha+2s)m-1}| \equiv |v_{(\alpha+s)m-1}| \equiv |v_{\alpha m-1}| \pmod{k}. \tag{3.115}$$

  Since $\alpha_1 = \alpha + s$, one can rewrite (3.114) and (3.115) as

  $$K^{0,k}_{v_{(\alpha_1+s)m-1}} = K^{0,k}_{v_{\alpha_1 m-1}} = K^{0,k}_{v_{\alpha m-1}}. \tag{3.116}$$

  and

  $$|v_{(\alpha_1+s)m-1}| \equiv |v_{\alpha_1 m-1}| \equiv |v_{\alpha m-1}| \pmod{k}, \tag{3.117}$$

  respectively. Now, from (3.117) and (3.116) we obtain that integers α_1 and s satisfy conditions of Theorem 43 (for positions of zeros). As we already proved theorem for $j \in \overline{1, sm}$ in Case (1), then

  $$K^{0,k}_{v_{\alpha_1 m-1+j}} = K^{0,k}_{v_{(\alpha_1+s)m-1+j}} = K^{0,k}_{v_{\alpha_1 m-1}}$$

  for each $j \in \overline{1, sm}$. We recall that $K^{0,k}_{v_{\alpha_1 m-1}} = K^{0,k}_{v_{\alpha m-1}}$ (according to (3.114)), therefore (3.112) holds for $t = 1$.

- Assume (3.112) holds for $t = l$ and consider the case for $t = l + 1$. By assumption we have

  $$K^{0,k}_{v_{\alpha_l m-1+j}} = K^{0,k}_{v_{(\alpha_l+s)m-1+j}} = K^{0,k}_{v_{\alpha m-1}},$$

  for all $j \in \overline{1, sm}$. From here, if $j = sm$, then

  $$K^{0,k}_{v_{\alpha_l m-1+sm}} = K^{0,k}_{v_{(\alpha_l+s)m-1+sm}} = K^{0,k}_{v_{\alpha m-1}},$$

  therefore

  $$K^{0,k}_{v_{(\alpha_l+s)m-1}} = K^{0,k}_{v_{(\alpha_l+2s)m-1}} = K^{0,k}_{v_{\alpha m-1}}.$$

  From here and the fact that $\alpha_{l+1} = \alpha + (l + 1) \cdot s = \alpha_l + s$ we have

  $$K^{0,k}_{v_{\alpha_{l+1} m-1}} = K^{0,k}_{v_{(\alpha_{l+1}+s)m-1}} = K^{0,k}_{v_{\alpha m-1}}. \tag{3.118}$$

  Identities $\alpha_{l+1}m - 1 = (\alpha + (l + 1)s)m - 1 = \alpha m - 1 + (l + 1) \cdot sm$ and $(\alpha_{l+1} + s)m - 1 = (\alpha + (l + 1)s + s)m - 1 = (\alpha + s)m - 1 + (l + 1) \cdot sm$ together with (3.19) imply

  $$\left|v_{\alpha_{l+1}m-1}\right| \equiv \left|v_{(\alpha_{l+1}+s)m-1}\right| \pmod{k}. \tag{3.119}$$

  Now, from (3.119) and (3.118) we obtain that integers α_{l+1} and s satisfy conditions of Theorem 43 (for positions of zeros). As we already proved theorem for $j \in \overline{1, sm}$ in Case (1), then

  $$K^{0,k}_{v_{\alpha_{l+1} m-1+j}} = K^{0,k}_{v_{(\alpha_{l+1}+s)m-1+j}} = K^{0,k}_{v_{\alpha_{l+1} m-1}}$$

  for each $j \in \overline{1, sm}$. From here, $\alpha_{l+1} = \alpha_l + s$, and assumption that (3.112) holds for $t = l$, we have

  $$K^{0,k}_{v_{\alpha_{l+1} m-1}} = K^{0,k}_{v_{(\alpha_l+s)m-1}} = K^{0,k}_{v_{\alpha m-1}},$$

  therefore (3.112) holds for $t = l + 1$ too.

We conclude the proof by recalling that (3.112) implies (3.113).

□

Corollary 44. *Let $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, $m \geq 2$, be a basis of a finitely generated bi-ideal $x \in \Sigma_n^\omega$, $n \geq 2$. Let k be an arbitrary positive integer and let α, s be positive integers such that*

1. $|v_{\alpha m-1}| \equiv \left|v_{(\alpha+s)m-1}\right| \pmod{k}$, *2.* $K^{a,k}_{v_{\alpha m-1}} = K^{a,k}_{v_{(\alpha+s)m-1}}$ *for all* $a \in alph(x)$.

Then for each letter $a \in alph(x)$*,*

$$K^{a,k}_x = K^{a,k}_{v_{\alpha m-1}}.$$

Lemma 45. *Let* $\langle u_0, u_1, ..., \ldots, u_{m-1}\rangle$ *be a basis of a finitely generated bi-ideal* x*. Then for each* $k \in \mathbb{N}_+$ *we can find* $\alpha, s \in \mathbb{N}_+$ *such that* $\alpha, s \leq k$ *and*

$$|v_{\alpha m-1}| \equiv |v_{(\alpha+s)m-1}| \pmod{k}.$$

Proof. By Pigeonhole principle we can choose from integers

$$m-1, 2m-1, \ldots, km-1, (k+1)m-1$$

two that are congruent modulo k. Clearly, $\alpha, s \leq k$. □

Lemma 46. *Let* $\langle u_0, u_1, ..., \ldots, u_{m-1}\rangle$ *be a basis of a finitely generated bi-ideal* $x \in \Sigma_n^\omega$*,* $n \geq 2$*. If* α *and* s *are positive integers such that* $|v_{\alpha m-1}| \equiv |v_{(\alpha+s)m-1}| \pmod{k}$*, where* $k \in \mathbb{N}_+$*, then there exists positive integer* $t \leq k$ *such that for each* $j \in \mathbb{N}$*, for each letter* $a \in alph(x)$*,*

$$K^{a,k}_{v_{(\alpha+t\cdot s)m-1}} = K^{a,k}_{v_{(\alpha+t\cdot s)m-1+j}}.$$

Proof. By Corollary 25, if $|v_{\alpha m-1}| \equiv |v_{(\alpha+s)m-1}| \pmod{k}$, then also

$$|v_{\alpha m-1}| \equiv |v_{(\alpha+s)m-1}| \equiv |v_{(\alpha+2s)m-1}| \equiv \cdots \equiv |v_{(\alpha+k\cdot s)m-1}| \pmod{k}.$$

Clearly, for each $a \in alph(x)$, $Card(K^{a,k}_x) \leq k$. From here and the fact that each of the given sets

$$K^{a,k}_{v_{\alpha m-1}}, K^{a,k}_{v_{(\alpha+s)m-1}}, K^{a,k}_{v_{(\alpha+2s)m-1}}, \ldots, K^{a,k}_{v_{(\alpha+(k-1)s)m-1}}, K^{a,k}_{v_{(\alpha+k\cdot s)m-1}} \tag{3.120}$$

is included in the next one, there are at most k distinct sets in (3.120). Hence, for each letter $a \in alph(x)$ there are $i_a, j_a \in \overline{0,k}$, $i_a < j_a$, such that $K^{a,k}_{v_{(\alpha+i_a s)m-1}} = K^{a,k}_{v_{(\alpha+j_a s)m-1}}$. If for each letter $a \in alph(x)$ we set $\alpha'_a = \alpha + i_a \cdot s$ and $s'_a = (j_a - i_a)s$, then according to Corollary 44 we have $K^{a,k}_x = K^{a,k}_{v_{\alpha'_a m-1}}$. Therefore for each $a \in alph(x)$, for each $j \in \mathbb{N}$

$$K^{a,k}_{v_{(\alpha+i_a s)m-1}} = K^{a,k}_{v_{(\alpha+i_a s)m-1+j}}$$

Each of the sets in (3.120) is included in the next one, therefore if we set $t = \max\{i_a \big| \, a \in alph(x)\}$, then $K^{a,k}_{v_{(\alpha+t\cdot s)m-1}} = K^{a,k}_{v_{(\alpha+t\cdot s)m-1+j}}$ for each $a \in alph(x)$.

□

Algorithm 47. *In order to check whether all arithmetical subsequences of a finitely generated bi-ideal $x \in \Sigma_n^\omega$, $n \geq 2$, with basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$ are aperiodic we need to follow these steps:*

1. *Find the length-differential m-tuple $\langle k_0, k_1, \ldots, k_{m-1}\rangle$ of the basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$.*
2. *If $k_j = 0$ for each $j \in \overline{0, m-1}$, then x contains periodic arithmetical subsequence. Otherwise, for a non-zero k_j, $j \in \overline{0, m-1}$:*
   (a) *Find $\alpha_j, s_j \in \mathbb{N}_+$ such that*
   $$|v_{\alpha_j m-1}| \equiv |v_{(\alpha_j+s_j)m-1}| \pmod{|k_j|}.$$
   (From Lemma 45 we know that $\alpha_j, s_j \leq |k_j|$.)

   (b) *Find $t < |k_j|$ such that, for each $a \in alph(x)$,*
   $$K^{a,k}_{v_{(\alpha_j+t\cdot s_j)m-1}} = K^{a,k}_{v_{(\alpha_j+(t+1)s_j)m-1}}.$$
   (From Lemma 46 we know that such integer $t < |k_j|$ exists)

   (c) *For each $a \in alph(x)$ find the cardinality of the set $\breve{K}^{a,k}_{v_{(\alpha_j+t\cdot s_j)m-1}}$. If for all letters $a \in alph(x)$, $Card(\breve{K}^{a,k}_{v_{(\alpha_j+t\cdot s_j)m-1}}) = |k_j|$, then all arithmetical subsequences of x are aperiodic.*

   (By definition of L-prolongation and Lemma 18 there exists a non-neagtive integer ξ such that $|u_0^{(\xi)}| \geq |v_{(\alpha_j+t\cdot s_j)m-1}|$ and $|u_0^{(\xi)}| - |u_1^{(\xi)}| = |k_j|$; thus the basis $\langle u_0^{(\xi)}, u_1^{(\xi)}, \ldots, u_{m-1}^{(\xi)}\rangle$ satisfies conditions of Theorem 39.)

   Otherwise, if there exists a letter $a \in alph(x)$ such that $Card(\breve{K}^{a,k}_{v_{(\alpha_j+t\cdot s_j)m-1}}) < |k_j|$, then x contains a periodic arithmetical subsequence.

   (By Corollary 25 and Corollary 44 we obtain the equality $\breve{K}^{a,k}_{v_{(\alpha_j+t\cdot s_j)m-1}} = \breve{K}^{a,k}_x$, therefore $Card(\breve{K}^{a,k}_x) < |k_j|$.)

Example 13. Let $x \in \{0, 1, 2\}^\omega$ be a finitely generated bi-ideal with basis $\langle 0, 121\rangle$.

1. The length-differential 2-tuple of the basis is $\langle -2, 2\rangle$.
2. We consider $k_0 = -2$.

(a) We generate v_{m-1}, v_{2m-1}, and v_{3m-1} if necessary:

$$\begin{aligned} v_{m-1} &= v_1 = v_0u_1v_0 = 01210 \\ v_{2m-1} &= v_3 = v_2u_3v_2 = (v_1u_2v_1)u_3(v_1u_2v_1) \\ &= (01210001210)121(01210001210). \end{aligned}$$

We see that

$$\begin{aligned} |v_{m-1}| &= 5 \equiv 1 \pmod{2}, \\ |v_{2m-1}| &= 2 \cdot 11 + 3 = 25 \equiv 1 \pmod{2}, \end{aligned}$$

therefore it is not necessary to generate v_{3m-1} and $\alpha = 1$, $s = 1$.

(b) Now we consider positions of each letter modulo 2:

$$K^{0,2}_{v_{m-1}} = \{0\}, K^{1,2}_{v_{m-1}} = \{1\}, K^{2,2}_{v_{m-1}} = \{0\}$$

and

$$K^{0,2}_{v_{2m-1}} = \{0,1\}, K^{1,2}_{v_{2m-1}} = \{1\}, K^{2,2}_{v_{2m-1}} = \{0\}.$$

Since Card$(K^{0,2}_{v_{2m-1}}) = 2 = |k_j|$, $K^{1,2}_{v_{m-1}} = K^{1,2}_{v_{2m-1}}$ and $K^{2,2}_{v_{m-1}} = K^{2,2}_{v_{2m-1}} = \{0\}$, then

$$K^{0,2}_{v_{2m-1}} = K^{0,2}_{v_{3m-1}}, K^{1,2}_{v_{2m-1}} = K^{1,2}_{v_{3m-1}}, K^{2,2}_{v_{2m-1}} = K^{2,2}_{v_{3m-1}},$$

therefore $n = 1$.

(c) Now, as $Card(\breve{K}^{0,2}_{v_{2m-1}}) = Card(\breve{K}^{1,2}_{v_{2m-1}}) = Card(\breve{K}^{2,2}_{v_{2m-1}}) = Card(\{0,1\}) = 2$, then all arithmetical subsequences of x are aperiodic.

Example 14. Let $x \in \{0,1\}^\omega$ be a finitely generated bi-ideal with basis $\langle 0, 101 \rangle$.

1. The length-differential 2-tuple of x is $\langle -2, 2 \rangle$.
2. We consider $k_0 = -2$.

(a) Since the basis words are of the same length in this and previous example, then $\alpha = 1$ and $s = 1$. Hence we generate only v_{m-1} and v_{2m-1}:

$$\begin{aligned} v_{m-1} &= v_1 = v_0u_1v_0 = 01010 \\ v_{2m-1} &= v_3 = v_2u_3v_2 = (v_1u_2v_1)u_3(v_1u_2v_1) \\ &= (01010001010)101(01010001010). \end{aligned}$$

(b) Now we consider positions of zeros and ones modulo 2:

$$K^{0,2}_{v_{m-1}} = \{0\}, K^{1,2}_{v_{m-1}} = \{1\}$$

and

$$K^{0,2}_{v_{2m-1}} = \{0,1\}, K^{1,2}_{v_{2m-1}} = \{1\}.$$

Since Card$(K^{0,2}_{v_{2m-1}}) = 2 = |k_j|$, and $K^{1,2}_{v_{m-1}} = K^{1,2}_{v_{2m-1}} = \{1\}$, then

$$K^{0,2}_{v_{2m-1}} = K^{0,2}_{v_{3m-1}}, K^{1,2}_{v_{2m-1}} = K^{1,2}_{v_{3m-1}},$$

and $n = 1$.

(c) Now, $\breve{K}^{0,2}_{v_{2m-1}} = K^{1,2}_{v_{2m-1}} = \{1\}$. Hence $Card(\breve{K}^{0,2}_{v_{2m-1}}) = 1 < 2$, therefore x contains periodic arithmetical subsequence.

3.1.3 Case of Two Basis Words

Here we give some more results on finitely generated bi-ideals with only two words in its basis.

Proposition 48 (See e.g. (Rosen, 2005, Thm. 4.6, p. 146)). *If* gcd$(a, m) = 1$*,* $b \in \mathbb{Z}$ *and*

$$x_1, x_2, ..., x_m$$

form complete residue system modulo m*, then integers*

$$y_1, y_2, ..., y_m,$$

where $y_i = ax_i + b$*, also form complete residue system modulo* m*.*

Proposition 49 (See e.g. (Reid, 1910, p. 37)). *If* $a \equiv b \pmod{m}$ *and* d *divides* m*, then*

$$a \equiv b \pmod{d}.$$

Proposition 50. *If* k *is an odd number, then these conditions are equivalent:*

1. *A bi-ideal* $x \in \Sigma_n^\omega$*,* $n \geq 2$*, has a basis* $\langle \tilde{u}_0, \tilde{u}_1 \rangle$ *such that* $k = ||\tilde{u}_0| - |\tilde{u}_1||$ *and for all* $a \in alph(x)$*,* $Card(\breve{K}^{a,k}_{\tilde{u}_0}) = k$*.*
2. *A bi-ideal* $x \in \Sigma_n^\omega$*,* $n \geq 2$*, has a basis* $\langle u_0, u_1 \rangle$ *such that* $k = ||u_0| - |u_1||$ *and for each* $a \in alph(x)$*,* $Card(\breve{K}^{a,s}_{u_0}) = s$*, where* $s =$ gcd$(|u_0|, |u_1|)$*.*

Proof. $(1) \Rightarrow (2)$: Let a be an arbitrary letter of the alphabet $alph(x)$. Since $\breve{K}_{\tilde{u}_0}^{a,k}$ is a complete residue system modulo k, then there exist non-negative integers $\mu_0, \mu_1, ..., \mu_{k-1}$ such that

$$i_0 = 0 + \mu_0 k,$$

$$i_1 = 1 + \mu_1 k,$$

$$\cdots$$

$$i_{k-1} = k - 1 + \mu_{k-1} k,$$

and for all $j \in \overline{0, k-1}$ we have

$$\tilde{u}_0[i_j] \in alph(x)\Sigma_n \setminus \{a\}.$$

Since $s = \gcd(|\tilde{u}_0|, |\tilde{u}_1|)$ is a divisor of k, then positions $i_0, i_1, ...i_{s-1}$ form a complete residue system modulo s too.

$(2) \Rightarrow (1)$: If $k = ||u_0| - |u_1||$ is odd, then integers

$$0, \quad 2, \quad 4, \quad 6, \quad ..., \quad 2(k-1)$$

form a complete residue system modulo k. We consider two cases.

- If $\gcd(|u_0|, |u_1|) = 1$, $d = |u_0| \bmod k$ and i is an arbitrary integer, then according to Proposition 48 integers

  $$i, \quad i+2d, \quad i+4d, \quad i+6d, \quad ..., \quad i+2(k-1)d$$

  form a complete residue system modulo k.

  Let a be an arbitrary letter of the alphabet $alph(x)$. The fact that $\breve{K}_{u_0}^{a,s}$ is a complete residue system modulo $s = 1$ implies that u_0 contains as factor at least one letter of the set $alph(x) \setminus \{a\}$. Hence there exists a non-negative integer i such that $u_0[i] \in alph(x) \setminus \{a\}$.

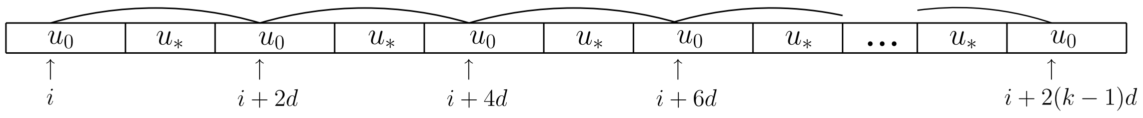
  

  Figure 3.5: Positions modulo k of $u_0[i]$ in the word x.

  Next, we observe that x can be expressed in the form (see Figure 3.5)

  $$x = v_0 u_* v_0 u_* v_0 u_* v_0 u_* ... = u_0 u_* u_0 u_* u_0 u_* u_0 u_* ...,$$

where u_* is the corresponding basis word (we write $*$, since we do not need to specify the index for this proof). We rewrite this factorization as

$$x = w_0w_1w_2w_3w_4...,$$

where $w_{2t} = u_0$ for all $t \in \mathbb{N}$, while w_{2t+1} is equal to the corresponding basis word u_*. For each $t \in \overline{0, k-1}$ we denote by i_t the absolute position of $w_{2t}[i] \in alph(x) \setminus \{a\}$ in the word x. Since $d = |u_0| \bmod k$ and $||u_0| - |u_1|| = k$, then $d = |u_1| \bmod k$. Hence we have

$$\begin{array}{lcll}
i_0 & \equiv & i & \pmod{k}, \\
i_1 & \equiv & i+2d & \pmod{k}, \\
i_2 & \equiv & i+4d & \pmod{k}, \\
i_3 & \equiv & i+6d & \pmod{k}, \\
 & \dots & & \\
i_{k-1} & \equiv & i+2(k-1)d & \pmod{k}.
\end{array}$$

From here and Proposition 48 it follows that positions $i_0, ..., i_{k-1}$ form a complete residue system modulo k. According to Lemma 18 there exists a non-negative integer η_a such that $\left||u_0^{(\eta_a)}| - |u_1^{(\eta_a)}|\right| = k$ and the length of the basis word $u_0^{(\eta_a)}$ exceeds all these positions $i_0, ..., i_{k-1}$. Thus $Card(\breve{K}^{a,k}_{u_0^{(\eta_a)}}) = k$. As $a \in alph(x)$ was chosen arbitrarily, then for each $a \in alph(x)$ such integer η_a exists. We set $\eta = \max\{\eta_0, \eta_1, \dots, \eta_{n-1}\}$. The basis $\langle \tilde{u}_0, \tilde{u}_1 \rangle$, with $\tilde{u}_0 = u_0^{(\eta)}$ and $\tilde{u}_1 = u_1^{(\eta)}$, satisfies the first condition of the proposition.

- The case when $\gcd(|u_0|, |u_1|) = s > 1$ we reduce to the first case by replacing blocks of basis words of the length s with one letter. We choose a position i, where $0 \le i < s$, and replace the respective block with a letter that occurs in i-th position of the block (see Figure 3.6), i.e., if $|u_0| = t \cdot s$ and $|u_1| = t' \cdot s$, where $t, t' \in \mathbb{N}_+$, then we replace u_0 and u_1 with words

  $$u_0' = u_0[i]u_0[i+s]u_0[i+2s]...u_0[i+(t-1)s]$$

  and

  $$u_1' = u_1[i]u_1[i+s]u_1[i+2s]...u_1[i+(t'-1)s],$$

  respectively.

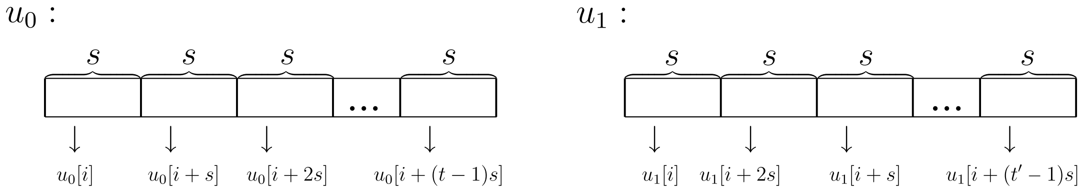


Figure 3.6: Replacing blocks by a single letter.

Let $a \in alph(x)$. We will show that in just obtained bi-ideal $x^{(i)}$ the set $\breve{K}^{a,k'}_{x^{(i)}}$ is a complete residue system modulo k', where $k' = \frac{k}{s}$. The fact that $\breve{K}^{a,s}_{u_0}$ is a complete residue system modulo s implies that u'_0 contains as factor at least one letter of the set $alph(x) \setminus \{a\}$.

Since gcd$(|u_0|, |u_1|) = s$ and $||u_0| - |u_1|| = k$ is odd, we obtain gcd$(t, t') = 1$. Moreover, $k' = |t - t'| = \frac{k}{s}$ is odd number. Case 1 implies that $\breve{K}^{a,k'}_{x^{(i_0)}}$ is a complete residue system modulo k'. As the position i was chosen arbitrary, it holds for all $i \in \overline{0, s-1}$.

Next, we observe that for all $i, j \in \overline{0, s-1}$:

$$i \neq j \implies \forall m, n \in \mathbb{N}\left(i + m \cdot s \not\equiv j + n \cdot s \pmod{k}\right). \tag{3.121}$$

Otherwise, if $i + m \cdot s \equiv j + n \cdot s \pmod{k}$ for some distinct $i, j \in \overline{0, s-1}$, then by Proposition 49

$$i \equiv j \pmod{s},$$

which leads to contradiction that inequality $i \neq j$ implies $i \not\equiv j \pmod{s}$ for all integers $i, j \in \overline{0, s-1}$.

By $x^{(i)}$ construction for each $i \in \overline{0, s-1}$ and each $n \geq 0$,

$$x^{(i)}[n] = x[i + n \cdot s].$$

From here and (3.121), for all $i, j \in \overline{0, s-1}$ the inequality $i \neq j$ implies that the position of an arbitrarily chosen letter $b \in alph(x) \setminus \{a\}$ from the word $x^{(i)}$ and the position of an arbitrarily chosen letter $c \in alph(x) \setminus \{a\}$ from the word $x^{(j)}$ are in different classes of residues modulo k in the word x. Since there are s distinct integers in list $0, 1, \ldots, s-1$ and since for all $i \in \overline{0, s-1}$ the set $\breve{K}^{a,k'}_{x^{(i)}}$ is a complete residue system modulo k', then we have $s \cdot k' = k$ distinct positions of $alph(x) \setminus \{a\}$ elements modulo k in word x, i.e., $\breve{K}^{a,k}_{x}$ is a complete residue system modulo k. Thus $Card(\breve{K}^{a,k}_{x}) = k$. As $a \in alph(x)$ was chosen arbitrarily and as the alphabet Σ_n is finite, there is a non-negative integer η

such that $Card(\breve{K}^{a,k}_{u_0^{(\eta)}}) = k$ for all $a \in alph(x)$. Moreover, according to Lemma 18 we can choose η such that $\left| |u_0^{(\eta)}| - |u_1^{(\eta)}| \right| = k$. The basis $\langle \tilde{u}_0, \tilde{u}_1 \rangle$, where $\tilde{u}_0 = u_0^{(\eta)}$ and $\tilde{u}_1 = u_1^{(\eta)}$, satisfies the first condition of the proposition.

□

Corollary 51. *Let $\langle u_0, u_1 \rangle$ be a basis of a finitely generated bi-ideal x. If* $\gcd(|u_0|, |u_1|) = 1$ *and $|u_0| - |u_1|$ is an odd number, then all arithmetical subsequences of x are aperiodic if and only if x contains at least two distinct letters.*

Remark 52. *If k is even, then Proposition 50 does not hold. We will show a counterexample.*

Example 15. Let x be a finitely generated bi-ideal generated with basis $\langle u_0, u_1 \rangle$, where $u_0 = 011$ and $u_1 = 1111111 = 1^7$. Then $k = |3 - 7| = 4$ and $s = \gcd(3, 7) = 1$. Since u_0 contains both "0" and "1" as factor, then positions of ones and zeros form complete residue system modulo $s = 1$ in u_0. We will show that positions of letter "0" do not form complete residue system modulo $k = 4$ in x, which implies that there does not exist basis word $u_0^{(i)}$ such that positions of ones and zeros form complete residue system modulo $k = 4$ in it.

We L-prolong basis words twice

$$\begin{aligned} u_0 &= 011 & &\to & u_0^1 &= 01^9 & &\to & u_0^{(2)} &= 01^9011011, \\ u_1 &= 1^7 & &\to & u_1^1 &= 011011 & &\to & u_1^{(2)} &= 01^901^9, \end{aligned}$$

and calculate $a_2 = |u_0^{(2)}| = 16$, $b_2 = |u_1^{(2)}| = 20$. Further we observe that the lengths of basis words of new basis $\langle u_0^{(2)}, u_1^{(2)} \rangle$ divide by 4.

Next, we find all positions of "0" in words $u_0^{(2)}$ and $u_1^{(2)}$:

$$u_0^{(2)}[0] = u_0^{(2)}[10] = u_0^{(2)}[13] = 0, \qquad u_1^{(2)}[0] = u_1^{(2)}[10] = 0.$$

Finally, one can see that neither in $u_0^{(2)}$ nor in $u_1^{(2)}$ exists a position of "0" which is congruent to 3 modulo 4. Since the length of both $u_0^{(2)}$ and $u_1^{(2)}$ divides by 4, then by bi-ideal construction there does not exist position of "0" which is congruent to 3 modulo 4 in the bi-ideal x.

3.2 Aperiodicity of all Arithmetical Subsequences of a Bounded Bi-ideal

Here we consider bounded bi-ideals and give a necessary and sufficient condition of aperiodicity of all arithmetical subsequences of a bounded bi-ideal (Theorem 53). The condition is basically the same as in Theorem 39, with the additional restriction 3, that, by definition, always holds in the case of finitely generated bi-ideals.

Theorem 53. *All arithmetical subsequences of a bounded bi-ideal $x \in \Sigma_n^\omega$, $n \geq 2$, are aperiodic if and only if there exists basis sequence $(u_i)_{i\geq 0}$ such that*

1. $|u_0| \neq |u_1|$*;*
2. *for each $a \in alph(x)$, $Card\left(\breve{K}_{u_0}^{a,k}\right) = k$, where $k = ||u_0| - |u_1||$;*
3. *there is an increasing infinite sequence $s_0 = 0, s_1, s_2, \ldots s_i, \ldots$ of integers that satisfies these conditions*
   (a) $|u_{s_0}| = |u_{s_1}| = \cdots = |u_{s_i}| = \cdots$*;*
   (b) $k = |k_{s_0}| = |k_{s_1}| = \cdots = |k_{s_i}| = \cdots$*, where $k_j = |u_j| - |u_{j+1}|$ for all $j \in \mathbb{N}$.*

Proof. $\Rightarrow$: By Lemma 20 we can assume that x is a bounded bi-ideal with basis sequence $(u_i)_{i\geq 0}$ such that

1. each element of $(u_i)_{i\geq 0}$ occur in $(u_i)_{i\geq 0}$ infinitely often;
2. each element of the length-differential sequence $(k_i)_{i\geq 0}$ of x (associated with $(u_i)_{i\geq 0}$) occurs in $(k_i)_{i\geq 0}$ an infinite number of times.

Similarly as in the case of finitely generated bi-ideals (see the proof of Theorem 39), the aperiodicity of all arithmetical subsequences of x implies the existence of a non-zero element in the sequence $(k_i)_{i\geq 0}$. Otherwise, all basis words are of the same length, and x contains a periodic arithmetical subsequence.

Let $\mathcal{K} = \left\{|k_j| \middle| k_j \neq 0 \wedge j \in \mathbb{N}\right\}$. Since each $k \in \mathcal{K}$ occurs in the sequence $(|k_i|)_{i\geq 0}$ infinitely often, then, according to Lemma 22, for each $k \in \mathcal{K}$ there exists an increasing infinite sequence $(s_i)_{i\geq 0}$ such that

1. $|u_{s_0}| = |u_{s_1}| = \cdots |u_{s_i}| = \cdots$;
2. $k = |k_{s_0}| = |k_{s_1}| = \cdots = |k_{s_i}| = \cdots$.

Similarly as in the proof of Theorem 39: if all arithmetical subsequences of the bi-ideal x are aperiodic, then for each $k \in \mathcal{K}$, each $l_k \in \overline{0, k-1}$ and each $a \in alph(x)$, we have $x_k^{l_k} \neq a^\omega$.

Hence, for each $k \in \mathcal{K}$, for each starting position $l_k \in \overline{0, k-1}$ of an arithmetical subsequence, and each letter $a \in alph(x)$, there exists non-negative integer δ_{l_k} such that

$$x[l_k + k\delta_{l_k}] \in alph(x) \setminus \{a\}. \tag{3.21}$$

Next, we recall that the number of distinct element in sequence $(k_i)_{i \geq 0}$ is finite, therefore $0 < Card(\mathcal{K}) < t$, where $t \in \mathbb{N}_+$. From here and (3.21), there exists a positive integer m such that

$$\left|u_0^{(s_m)}\right| > l_k + k\delta_{l_k}$$

Thus, for all $k \in \mathcal{K}$, for all letters $a \in alph(x)$, we have

$$Card\left(\breve{K}^{a,k}_{u_0^{(s_m)}}\right) = k.$$

Further we observe that, by Lemma 23, we have an increasing sequence $(s'_n)_{n \geq 0}$ of non-negative integers, defined by

$$s'_0 = 0, s'_1 = s_{m+1} - s_m, \ldots, s'_j = s_{m+j} - s_m, \ldots,$$

such that

$$|u^{(s_m)}_{s'_0}| = |u^{(s_m)}_{s'_1}| = \cdots = |u^{(s_m)}_{s'_i}| = \cdots$$

and

$$k = |k_{s_0}| = |k'_{s'_0}| = |k'_{s'_1}| = \cdots = |k'_{s'_i}| = \cdots .$$

Finally, we observe that the basis sequence $(u_i^{(s_m)})_{i \geq 0}$ together with the sequence $(s'_i)_{i \geq 0}$ satisfies the conditions of the theorem.

$\Leftarrow$: Let x be a bounded bi-ideal with a basis sequence $(u_n)_{n \geq 0}$ such that $|u_0| \neq |u_1|$, $k = ||u_0| - |u_1||$, $Card(K^{a,k}_{u_0}) = k$ for all $a \in alph(x)$, and there exists an increasing infinite sequence of non-negative integers $s_0, s_1, s_2, \ldots, s_i \ldots$ such that

1. $|u_{s_0}| = |u_{s_1}| = \cdots = |u_{s_i}| = \cdots$;

2. $k = |k_{s_0}| = |k_{s_1}| = \cdots = |k_{s_i}| = \cdots$, where $k_j = |u_j| - |u_{j+1}|$ for all $j \in \mathbb{N}$.

Assume on the contrary that x contains an ultimately periodic arithmetical subsequence. Clearly, then there also exists an arithmetical subsequence, which contains only one letter. Without loss of generality we can assume that x contains arithmetical subsequence $x^l_p = 1^\omega$, where

$p > 1$ (the case when $p = 1$ is trivial). From Corollary 37 we can ensure that $l < p$. Hence, for each $i \in \mathbb{N}$, $x[l + i \cdot p] = 1$. Therefore

$$Card\left(\breve{K}_x^{1,k}\right) < p. \tag{3.22}$$

Let $\sigma : \Sigma_n \mapsto \{0, 1\}$ be a morphism defined by

$$\sigma(a) = \begin{cases} 1, & \text{if } a = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Let $y = \sigma(x)$. Clearly, y is a bounded bi-ideal with a basis sequence $(u'_m)_{m \geq 0}$ such that, for all $j \in \mathbb{N}$, the word u'_j is obtained from u_j, by replacing all letters $2, 3, \dots, n-1$ by 0. Hence for each $j \in \mathbb{N}$ we have $|u'_j| = |u_j|$. Thus $|u'_{s_0}| = |u'_{s_1}| = \cdots = |u'_{s_i}| = \cdots$ and $k = |k'_{s_0}| = |k'_{s_1}| = \cdots = |k'_{s_i}| = \cdots$, where $k'_{s_i} = |u'_{s_i}| - |u'_{s_i+1}|$.

Then, by conditions of the theorem, positions of zeros and ones form a complete residue system modulo k in u'_0, while (3.22) implies that positions of zeros do not form a complete residue system modulo p in the bounded bi-ideal y. Further we consider the bi-ideal y and obtain a contradiction by proving that positions of zeros form a complete residue system modulo p in y.

Let (v_i) be the bi-ideal sequence of the bounded bi-ideal y. Observe that $s_0, s_1, \dots, s_i, \dots$ contains infinite subsequence $j_0, j_1, \dots, j_i$ such that

$$|v_{j_0-1}| \equiv |v_{j_1-1}| \equiv \cdots \equiv |v_{j_i-1}| \equiv \cdots \pmod{p} \tag{3.23}$$

Moreover, we can choose $\alpha, \beta \in \{j_0, j_1, \dots, j_i, \dots\}$ such that $1 < \alpha < \beta - 1$, $|v_{\alpha-1}| > p$ and v_α contains all possible positions of zeros modulo p.

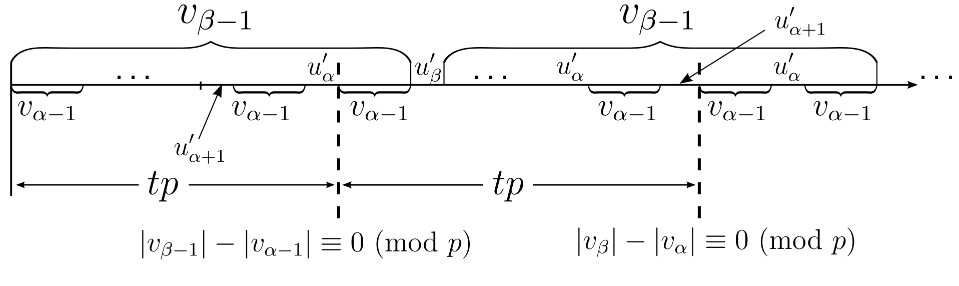


Figure 3.7: Construction of the bi-ideal y.

Now, from (3.23) there exists $t \in \mathbb{N}_+$ such that (see Figure 3.7)

$$|v_{\beta-1}| - |v_{\alpha-1}| = tp.$$

By bi-ideal construction $u'_0 \in \mathrm{Pref}(v_{\alpha-1})$ and $v_{\alpha-1} \in \mathrm{Pref}(v_{\beta-1})$, therefore, positions of zeros and ones in words $v_{\alpha-1}$ and $v_{\beta-1}$ form a complete residue system modulo k.

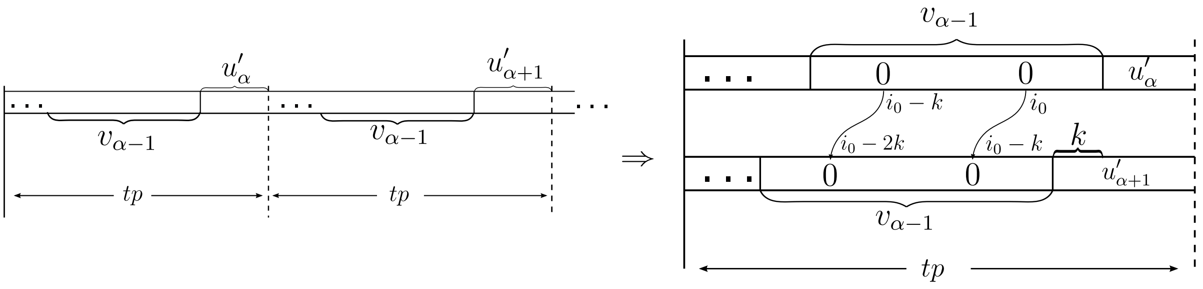


Figure 3.8: Shift for k positions of the factor $v_{\alpha-1}$ in the word y.

Next, we consider the case when $|u'_{\alpha+1}| = |u'_\alpha| + k$. Further the proof is analogous to the proof of Theorem 40 in case of finitely generated bi-ideals, i.e., we only replace $v_{\alpha m-1}$ with $v_{\alpha-1}$, $v_{\beta m-1}$ with $v_{\beta-1}$, u'_1 with $u'_{\alpha+1}$ and $|u'_0|$ with $|u'_\alpha| = |u'_\beta|$. See Figure 3.8.

□

3.3 Universal bi-ideals and arithmetical subsequences

In this section we give some connections between universal bi-ideals and aperiodicity of their arithmetical subsequences. We conjectured that a finitely generated bi-ideal is universal if and only if all its arithmetical subsequences are aperiodic.

It turned out that all arithmetical subsequences of the universal bi-ideals given in Theorem 30 are aperiodic.

Proposition 54. *If $u_0 = 1$, $u_1 = 10$ and $00 \notin F(u_i)$ for all $i \in \{2, 3, \ldots, m-1\}$ and y is the bi-ideal generated by the basis $\langle u_0, u_1, \ldots, u_{m-1}\rangle$, then all arithmetical subsequences are aperiodic.*

Proof. From Lemma 18 we obtain

$$\forall i \geq 0\, \forall j \in \overline{0, m-1}\ \left(|u_j^{(i)}| - |u_{j+1 \mod m}^{(i)}| = k_{i+j \mod m}\right),$$

therefore

$$|u_0^{(m)}| - |u_{1 \bmod m}^{(m)}| = k_0 = -1.$$

Since $m \geq 2$, then $u_0^{(m)}$ contains at least one zero and several letters one, therefore positions of ones and zeros form complete residue system modulo $|k_0| = 1$. □

We also proved that a universal bi-ideal does not contain an arithmetical subsequence 0^ω.

Proposition 55. *If finitely generated bi-ideal y has arithmetical subsequence $y_p^l = 0^\omega$ for some integers $l \geq 0$ and $p > 0$, then it is not universal bi-ideal.*

Proof. Assume on contrary that y is universal bi-ideal and it has arithmetical subsequence $y_p^l = 0^\omega$ for some integers $l \geq 0$ and $p > 1$. From Lemma 37 we can ensure that $l < p$. If $l = 0$ then we set $x = (01^{p-1})^\omega$, otherwise we set $x = \left(1^{l-1}01^{p-l}\right)^\omega$.

Clearly, $S_y(x) = 1^\omega$. □

However there are non-universal finitely generated bi-ideals that contains only aperiodic arithmetical subsequences.

Example 16. Let $\langle 101, 10001\rangle$ be a basis of a finitely generator bi-ideal y. In Example 9 we already proved that y is not a universal bi-ideal by showing that the shrunk sequence $\mathbf{SHR}_y(x)$, where $x = (01)^\omega$, is periodic.

We also showed that the basis of y after one iteration of L-prolongation is

$$\langle u_0^{(1)}, u_1^{(1)}\rangle = \langle 10110001, 101101\rangle.$$

Since $|u_0^{(1)}| - |u_0^{(1)}| = 8 - 6 = 2$ and since positions of ones and zeros form a complete residue system modulo 2 in the basis word $u_0^{(1)} = 10110001$, then according to Corollary 40 all arithmetical subsequences of y are aperiodic.

4 Bounded Bi-ideals and Linear Recurrence

The aim of this chapter is to give a charatcerization of linearly recurrent bounded bi-ideals. In Section 4.2 we introduce the notion of a completely bounded bi-ideal. Then in Section 4.3 we show that the class of completely bounded bi-ideals is exactly the class of linearly recurrent bounded bi-ideals.

4.1 Linear Recurrence

Let u be a non-empty factor of $x \in A^\omega$. A word $w \in A^+$ is called *a return word to u of x* if wu is a factor of x, u is a prefix of wu, and $|wu|_u = 2$. The set of all return words to u of x we denote by $\mathcal{R}_{x,u}$.

An infinite word $x \in A^\omega$ is called *linearly recurrent* if it is uniformly recurrent and there exists a constant $K \in \mathbb{N}$ such that for all $u \in \mathrm{F}(x)$ and all $w \in \mathcal{R}_{x,u}$ we have $|w| \leq K \cdot |u|$.

Theorem 56. (Durand et al., 1999) *Let x be an aperiodic linearly recurrent word with constant K. Then:*

1. *For all $n \in \mathbb{N}$ each subword of length n appears in each factor of length $(K+1)n$ in x.*

2. *The number of distinct factors of length n in x is less than or equal to Kn.*

3. *For all $u \in \mathrm{F}(x)$ and for all $w \in \mathcal{R}_{x,u}$ we have $(1/K)|u| < |w|$.*

4. *For all $u \in \mathrm{F}(x)$, Card$(\mathcal{R}_{x,u}) \leq K(K+1)^2$.*

An *adic representation* of an infinite word $x \in A^\omega$ is given by a sequence $(A_n)_{n\in\mathbb{N}}$ of alphabets, a sequence $(\sigma_n : A_{n+1}^* \to A_n^*)_{n\in\mathbb{N}}$ of morphisms, and a sequence $(a_n)_{n\in\mathbb{N}}$ of letters such that $a_i \in A_i$ for all non-negative integers i, $A_0 = A$, $\lim_{n\to+\infty} |\sigma_0\sigma_1 \cdots \sigma_n(a_{n+1})| = +\infty$, and $w = \lim_{n\to+\infty} \sigma_0\sigma_1\cdots\sigma_n(a_{n+1}^{\omega})$. When all the morphisms σ_n, $n \in \mathbb{N}$, belong to a given finite set S of morphisms, then x is called *S-adic*. If there exists an integer s_0 such that for all $r \in \mathbb{N}$, all $b \in A_r$ and all $c \in A_{r+s_0+1}$, the letter b has an occurrence in $\sigma_{r+1}\sigma_{r+2}\cdots\sigma_{r+s_0}(c)$, then we say that x is a primitive S-adic sequence (with constant s_0). A morphism σ is called *proper* if there are two letters r, l such that for all letters a, the image $\sigma(a)$ starts with letter l and ends with letter r. A sequence is called *primitive and proper S-adic* if it is primitive S-adic and all morphisms in S are proper.

Proposition 57. (Durand, 2003) *A sequence is linearly recurrent if and only if it is a primitive and proper S-adic sequence.*

According to Durand's result, in order to check whether a bounded bi-ideal x is linearly recurrent or is not linearly recurrent we should be able to choose a finite set S of proper morphisms and show that x is primitive and proper S-adic or to prove that there does not exist a finite set of proper morphisms S such that x is primitive and proper S-adic. We solve the problem without using the notion S-adicity. We put additional restrictions on the basis sequence (u_i). For our purposes it also more convenient to be able to put the restrictions on the basis sequence (u_i) before the bi-ideal is generated.

4.2 Completely Bounded Bi-ideals

Let $u_0, u_1, \ldots, u_n, \ldots$ be a sequence of finite words over an alphabet A. A subsequence

$$u_{i_0}, u_{i_1}, \ldots, u_{i_k}, \ldots \tag{4.21}$$

of the sequence $(u_n)_{n\geq 0}$ is called *constant* if the following conditions hold:

(i) $\forall j, j' \in \mathbb{N} \left(u_{i_j} = u_{i_{j'}}\right)$;

(ii) $u_k = u_{i_0} \Rightarrow \exists n(k = i_n)$.

A constant subsequence (4.21) is called *bounded* if it is finite or there is a positive integer l such that $i_n - i_{n-1} \leq l$ for all $n \in \mathbb{N}_+$. A constant subsequence (4.21) is called *boundless* if it is not bounded.

Definition 6. A bounded bi-ideal x is called *completely bounded* if there exists a basis sequence $(u_n)_{n\geq 0}$ of x which contains only bounded constant subsequences.

Since each bi-ideal has infinitely many basis sequences, then the “existence” condition in Definition 6 is crucial.

Example 17. Let x be a periodic word $(01)^\omega$. It is a completely bounded bi-ideal since its basis sequence

$$01, 01, 01, \dots, 01, \dots$$

has only one constant subsequence, and it is bounded. Nevertheless, the sequence $(u_i)_{i\geq 0}$ which is defined by

$$u_i = \begin{cases} 01 & \text{if } i = k^2 \text{ for a } k \in \mathbb{N}, \\ 0101 & \text{otherwise} \end{cases}$$

also is a basis sequence of x. Clearly, $(u_i)_{i\geq 0}$ contains a boundless constant subsequence.

Let x be a bounded bi-ideal generated by a sequence $(u_n)_{n\geq 0}$. As the length of each basis word of x is bounded by some $s \in \mathbb{N}$, the set

$$\mathcal{U} = \{u \big| \, \exists k(u = u_k)\}$$

is finite. From here there exists non-negative integer m such that

$$\mathcal{U} = \{u \big| \, \exists k \in \overline{0, m}(u = u_k)\}. \tag{4.22}$$

We denote $\mu_0 = \min\limits_{u\in\mathcal{U}} |u|$, and $\mu_1 = \max\limits_{u\in\mathcal{U}} |u|$.

If x is a completely bounded bi-ideal, then there exists a positive integer ℓ such that for all constant subsequences (4.21) of the basis sequence $(u_n)_{n\geq 0}$ of x we have

$$i_n - i_{n-1} \leq l \tag{4.23}$$

for all $n \in \mathbb{N}$.

In the sequel we use denotation m (l, respectively) for the smallest integer that satisfies (4.22) ((4.23), respectively) for a completely bounded bi-ideal with a given basis sequence $(u_i)_{i\geq 0}$.

We recall that by v_n we denote the n-th term of the bi-ideal sequence that is generated by the basis sequence (u_n), i.e., $v_0 = u_0$ and $v_{n+1} = v_n u_{n+1} v_n$ for all $n \in \mathbb{N}$.

Lemma 58. *Let x be a bounded bi-ideal with a basis sequence $(u_n)_{n\geq 0}$. Let $(v_n)_{n\geq 0}$ be the bi-ideal sequence generated by $(u_n)_{n\geq 0}$. Then $(2^{k+1} - 1)\mu_0 \leq |v_k| \leq (2^{k+1} - 1)\mu_1$ for all $k \in \mathbb{N}$.*

Proof. If $k = 0$, then $v_0 = u_0$, therefore

$$\mu_0 \leq |u_0| \leq \mu_1.$$

Now we assume that

$$(2^{k+1} - 1)\mu_0 \leq |v_k| \leq (2^{k+1} - 1)\mu_1$$

and consider the length of v_{k+1}. The equality $|v_{k+1}| = 2 \cdot |v_k| + |u_{k+1}|$ implies

$$|v_{k+1}| \geq 2(2^{k+1} - 1)\mu_0 + \mu_0 = (2^{k+2} - 1)\mu_0$$

and

$$|v_{k+1}| \leq 2(2^{k+1} - 1)\mu_1 + \mu_1 = (2^{k+2} - 1)\mu_1.$$

□

Lemma 59. *Let x be a bounded bi-ideal with a basis sequence $(u_n)_{n\geq 0}$. Let $(v_n)_{n\geq 0}$ be the bi-ideal sequence generated by $(u_n)_{n\geq 0}$. Then $|v_{k+n}| \leq 2^n|v_k| + (2^n - 1)\mu_1$ for all $k, n \in \mathbb{N}$.*

Proof. The proof is by induction. If $n = 1$, then

$$|v_{k+1}| = 2 \cdot |v_k| + |u_{k+1}| \leq 2 \cdot |v_k| + (2 - 1)\mu_1.$$

We assume that condition holds for $|v_{k+n}|$ and consider the length of v_{k+n+1}:

$$\begin{aligned} |v_{k+n+1}| &= 2 \cdot |v_{k+n}| + |u_{k+n+1}| \leq 2 \cdot (2^n|v_k| + (2^n - 1)\mu_1) + \mu_1 \\ &= 2^{n+1}|v_k| + (2^{n+1} - 1)\mu_1. \end{aligned}$$

□

From now we will consider only completely bounded bi-ideals.

Lemma 60. *Let x be a completely bounded bi-ideal with a basis sequence $(u_n)_{n\geq 0}$. Let $(v_n)_{n\geq 0}$ be the bi-ideal sequence generated by the sequence $(u_n)_{n\geq 0}$. If $u \in \mathrm{F}(v_n)$, $u \notin \mathrm{F}(v_{n-1})$, and $n \geq m + 1$, then $|u| > |v_{n-1-l}|$.*

Proof. Firstly, we observe that $v_n = v_{n-1}u_nv_{n-1}$ and $u \notin \mathrm{F}(v_{n-1})$ imply $u \setminus v'u_nv''$, where $v' \in \mathrm{Pref}(u) \cap \mathrm{Suff}(v_{n-1})$ and $v'' \in \mathrm{Suff}(u) \cap \mathrm{Pref}(v_{n-1})$. We can represent this condition with three alternative schemes (see Figure 4.1).

As one can see, it is possible to have $v' = \lambda$ or $v'' = \lambda$.

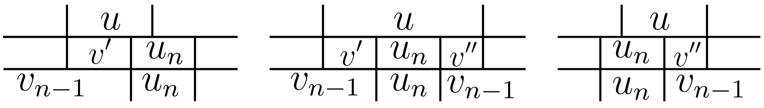


Figure 4.1: Three possibilities how word u can occur in $v'u_nv''$.

Definition of a completely bounded bi-ideal implies the existence of i, $1 \leq i \leq l$, such that $u_n = u_{n-i}$, but by bi-ideal construction $v_{n-1-i} \in \mathrm{Pref}(v_{n-1})$ and $v_{n-1-i} \in \mathrm{Suff}(v_{n-1})$. From here we obtain

$$v_n = v_{n-1}u_nv_{n-1} = w_1v_{n-i}w_2,$$

with $|w_1| = |w_2|$. For a schematical representation see Figure 4.2.

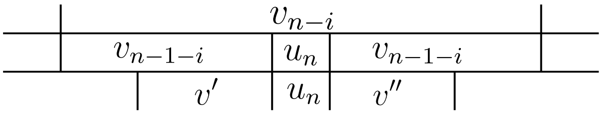


Figure 4.2: The structure of v_{n-i}.

Now, one can see that inequality $|u| \leq |v_{n-1-i}|$ implies $|v'| \leq |v_{n-1-i}|$ and $|v''| \leq |v_{n-1-i}|$. Thus $v'u_nv'' \diagdown v_{n-i} \in \mathrm{Pref}(v_{n-1})$. Contradiction, since $u \notin \mathrm{F}(v_{n-1})$. □

Corollary 61. *Let x be a completely bounded bi-ideal with a basis sequence $(u_n)_{n\geq 0}$. Let $(v_n)_{n\geq 0}$ be the bi-ideal sequence generated by $(u_n)_{n\geq 0}$. If $u \in \mathrm{F}(v_n)$, but u does not appear in v_{n-1}, then $|v_n| < 2^{\varkappa}(|u| + \mu_1)$, where $\varkappa = \max\{m+1, l+1\}$.*

Proof. (i) Let $n \geq m+1$. Then from Lemma 59 and Lemma 60 we obtain

$$\begin{aligned}|v_n| &\leq 2^{l+1}|v_{n-1-l}| + (2^{l+1}-1)\mu_1\\ &< 2^{l+1}|v_{n-1-l}| + 2^{l+1}\mu_1\\ &< 2^{l+1}|u| + 2^{l+1}\mu_1\\ &= 2^{l+1}\left(|u| + \mu_1\right)\\ &\leq 2^{\varkappa}\left(|u| + \mu_1\right).\end{aligned}$$

(ii) If $n \leq m$, then by Lemma 59 it follows that

$$\begin{aligned}|v_n| &\leq 2^n|v_0| + (2^n-1)\mu_1 < 2^n|u_0| + 2^n\mu_1\\ &\leq 2^n\mu_1 + 2^n\mu_1 \leq 2^{m+1}\mu_1 \leq 2^{m+1}\left(|u| + \mu_1\right)\\ &\leq 2^{\varkappa}\left(|u| + \mu_1\right).\end{aligned}$$

□

Lemma 62. *If* $a \geq 1$ *and* $b \geq 1$*, then* $a + b \leq ab + 1$.

Proof. From inequalities $a \geq 1$ and $b \geq 1$ we easily obtain

$$a - 1 \leq (a - 1)b,$$

$$a - 1 \leq ab - b,$$

$$a + b \leq ab + 1.$$

□

Corollary 63. *If* $a \geq 1$ *and* $b \geq 1$*, then* $a + b \leq 2ab$.

Proof. Lemma 62 implies

$$a + b \leq ab + 1 \leq ab + ab = 2ab.$$

□

Corollary 64. *Let* x *be a completely bounded bi-ideal with a basis sequence* $(u_n)_{n\geq 0}$*. Let* $(v_n)_{n\geq 0}$ *be the bi-ideal sequence generated by* $(u_n)_{n\geq}$*. If* $u \in \mathrm{F}(v_n)$ *and* $u \notin \mathrm{F}(v_{n-1})$*, then*

$$|v_n| < 2^{\varkappa+1}\mu_1|u|,$$

where $\varkappa = \max\{m + 1, l + 1\}$.

Proof. Corollary 61 and Corollary 63. □

4.3 Completely Bounded Bi-ideals and Linear Recurrence

In this subsection we show that the class of completely bounded bi-ideals is exactly the class of linearly recurrent bounded bi-ideals.

Theorem 65. *A bounded bi-ideal* x *is linearly recurrent if and only if it is completely bounded.*

Proof. $\Leftarrow$: At first we prove that a completely bounded bi-ideal x is linearly recurrent. Let x be a completely bounded bi-ideal, and $(u_i)_{i\geq 0}$ be its basis sequence that contains only bounded constant subsequences. Let $u \in \mathrm{F}(x)$. Then there exists a term v_n of the bi-ideal sequence such that $u \in \mathrm{F}(v_n)$. By construction of a completely bounded bi-ideal, x can be written as a factorization of v_n and basis words, i.e.,

$$x = v_n u'_1 v_n u'_2 \dots v_n u'_k \dots,$$

where $u'_s \in \mathcal{U}$ for all $s \in \mathbb{N}_+$.

Let $u[\mathbf{i},\mathbf{j})$ be an occurrence of u in x such that $u = x[i,j)$. Then there is $k \in \mathbb{N}$ and occurrence of $v_n u'_k v_n$ in x

$$v_n[\mathbf{i_1},\mathbf{i_2})u'_k[\mathbf{i_2},\mathbf{i_3})v_n[\mathbf{i_3},\mathbf{i_4}) = x[i_1,i_4)$$

such that $i_1 \leq i < i_3$. Otherwise, for $i \in [i_3,i_4)$ we would consider the occurrence of $v_n u'_{k+1} v_n$ in x instead of $v_n u'_k v_n$. So, $i_1 \leq i < i_3$ and we will find the next occurrence of u in x, e.g., $u[\mathbf{i'},\mathbf{j'}) = x[i',j')$. As u has an occurrence in v_n, then $i_3 \leq i' < i_4$. Clearly, $u[\mathbf{i},\mathbf{j})$ and $u[\mathbf{i'},\mathbf{j'})$ are two distinct occurrences of u in x and we can estimate the length of w, e.g., the length of the corresponding return word to u is

$$|w| \leq i' - i \leq |v_n u'_k| \leq |v_n| + \mu_1.$$

(i) If u does not appear in v_{n-1}, then, by Corollary 64, we have

$$|w| \leq |v_n| + \mu_1 < 2^{\varkappa+1}\mu_1|u| + \mu_1 \leq 2^{\varkappa+2}\mu_1|u|.$$

(ii) Observe, if $u \searrow v_{n-1}$, then we need to consider only the case when $u \searrow v_0$. Then

$$|w| \leq |v_0| + \mu_1 \leq 2\mu_1 \leq 2^{\varkappa+2}\mu_1|u|.$$

We conclude the proof by setting $K = 2^{\varkappa+2}\mu_1$. Then for each $u \searrow x$ and each return word $w \in \mathcal{R}_{x,u}$ we have

$$|w| \leq K \cdot |u|.$$

$\Rightarrow$: We assume on contrary that x is a linearly recurrent bounded bi-ideal that is not completely bounded. Then there exists a constant K such that for each factor w we have

$$|w| \leq K \cdot |r_w|,$$

where r_w is arbitrary return word to w in a bi-ideal x.

Let (u_i) be a basis sequence of x. Without loss of generality we can assume that the length μ_0 of the shortest basis word is greater than zero. Indeed, if $\mu_0 = 0$ (i.e., one of the basis words is the empty word λ), then we can L-prolong the basis words one time and consider new obtained basis sequence $(u_i^{(1)})$ instead of (u_i). By definition of L-prolongation, $u_0 \neq \lambda$ is a prefix of $u_i^{(1)}$ for all $i \in \mathbb{N}$. Thus none of the basis words in $(u_i^{(1)})$ is empty. Clearly, as x is not a completely bounded bi-ideal, then $(u_i^{(1)})$ contains a boundless constant subsequence.

Let $k_i = |u_i| - |u_{i+1}|$. Let $i \in \mathbb{N}$ and $\mathcal{K} = \{|k_i| | i \in \mathbb{N}\}$. According to definition of L-prolongation

$$k_i^{(1)} = |u_i^{(1)}| - |u_{i+1}^{(1)}| = |u_0 u_{i+1}| - |u_0 u_{i+2}| = k_{i+1}.$$

Hence, if $\mathcal{K}' = \{|k_i^{(1)}| | i \in \mathbb{N}\}$, then $\mathcal{K}' \subset \mathcal{K}$. From here, without loss of generality we can assume that the basis sequence (u_i) satisfies such conditions

$$\mu_1 < 2\mu_0, \tag{4.31}$$

where μ_0 (μ_1, respectively) is the length of the shortest (longest, respectively) basis word, and

$$0 \leq k_{\text{max}} < 0.1\mu_0, \tag{4.32}$$

where $k_{\text{max}} = \max\{||u_i| - |u_{i+1}||\,;\ i \in \mathbb{N}\}$.

Otherwise we could L-prolong basis words until we obtain a basis sequence that satisfies (4.31) and (4.32). Since x is not a completely bounded bi-ideal, then after any number of L-prolongations we obtain a basis sequence that contains a boundless constant subsequence.

The aperiodicity of x implies the existence of a constant n' such that for all $n'' \geq n'$ the n''-th term of the bi-ideal sequence $v_{n''}$ is not p-periodic for all $p \leq 3\mu_1$. Later this fact will help us to obtain a contradiction.

As the basis sequence of x contains a boundless constant subsequence (u_{i_k}), then there exists $n > \max\{2^m, n'\}$ such that $n = i_k$, $u_{i_k} = u_{i_{k+1}}$, and $i_{k+1} - i_k \geq K + 3$. In order to simplify calculations we assume that equality holds, namely, $i_{k+1} = n + K + 3$. We prove that for a word $w = v_n u_n v_n$ there exists return word r_w such that $|r_w| > K \cdot |w|$.

Further we divide the proof in two parts. First we prove that the word v_{n+K+2} does not contain $w = v_n u_n v_n$ as a factor. Afterwards we show that if w does not occur in v_{n+K+2}, then there exists a return word to w such that its length is greater than $K \cdot |w|$.

By the bi-ideal construction, it follows that v_{n+K+2} can be written as a factorization of v_n and basis words. Hence, if we denote each $u_i \in \mathcal{U} \setminus \{u_n\}$ by u_*, then v_{n+K+2} can be written in the form

$$v_{n+K+2} = v_n u_* v_n u_* v_n u_* \ldots v_n u_* v_n. \tag{4.33}$$

We use denotation u_* to point out the absence of u_n in factorization (4.33) of v_{n+K+2}, i.e., to point out the inequality $u_* \neq u_n$.

Firstly, we observe that $w \notin \text{Pref}(v_{n+K+2})$. Assume on contrary that $w \in \text{Pref}(v_{n+K+2})$. Then $|u_n| \neq |u_*|$ (otherwise the equality $u_* = u_n$ leads to contradiction). Hence we obtain a shift of v_n to the right of u_n and u_* (see Figure 4.3); therefore v_n is periodic with the length of the shift $|u'| < \mu_1 < 3\mu_1$. Contradiction. Analogously we can prove that $w \notin \mathrm{Suff}(v_{n+K+2})$.

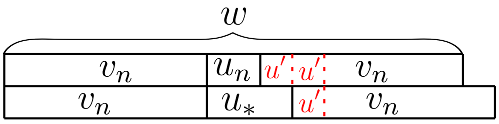


Figure 4.3: The word w is in prefix of v_{n+K+2}.

Next, we consider the case when $u_n = w[|v_n|, |v_n u_n|)$ occurs in $v_n u_* v_n u_* v_n$ so that u_n overlaps with one of the basis words u_*. As $u_n \neq u_*$, then these two conditions cannot hold at the same time:

a) $|u_*| = |u_n|$;

b) occurrences of words u_* and u_n in v_{n+K+2} start at the same position.

From here we obtain the shift of v_n to the right or to the left of basis words u_n and u_* (see Figure 4.4), therefore v_n is periodic with the length of the shift $|u'| < \mu_1 < 3\mu_1$. Contradiction.

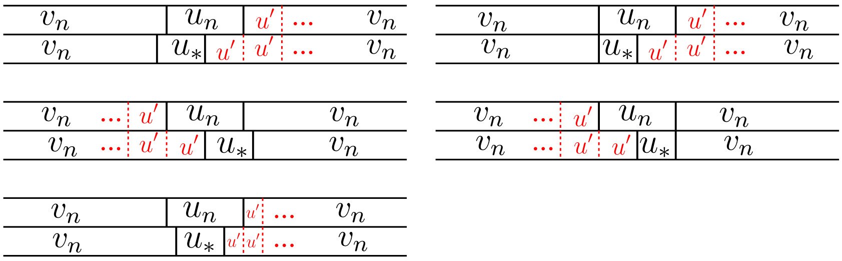


Figure 4.4: Word u_n overlaps with u_*; the case when $|u_n| > |u_*|$

It remains to consider the case when $u_n = w[|v_n|, |v_n u_n|)$ occurs in

$$v_n = (v_n u_* v_n u_* v_n)[|v_n u_*|, |v_n u_* v_n|).$$

Here we recall that $v_n = v_{n-1} u_n v_{n-1}$ and consider three subcases:

Case I

The word $u_n = w[|v_n|, |v_n u_n|)$ overlaps with $u_n = v_n[|v_{n-1}|, |v_{n-1} u_n|)$ so that

a) there is a shift of v_{n-1} to the left or to the right of u_n (see Figure 4.5). Then v_{n-1} is periodic with the length of the shift $|u'| < \mu_1 < 3\mu_1$. Contradiction.

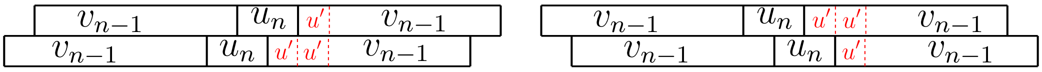


Figure 4.5: There is a shift, when u_n overlaps with u_n in v_n.

b) their occurrences coincide, namely,

$$w = v_n u_n v_n = v_{n-1} u_n v_{n-1} u_n v_{n-1} u_n v_{n-1}$$
$$= (v_n u_* v_{n-1} u_n v_{n-1} u_* v_n)[|v_n u_*| - |v_{n-1} u_n|, |v_n u_* v_n| + |u_n v_{n-1}|].$$

From here the equality of lengths of u_n and u_* would imply the equality of words themselves (which would leave to contradiction). Thus $|u_n| \neq |u_*|$ and we obtain shift of v_{n-1} to the right of u_* and u_n (see Figure 4.6). From here v_{n-1} is periodic with the length of

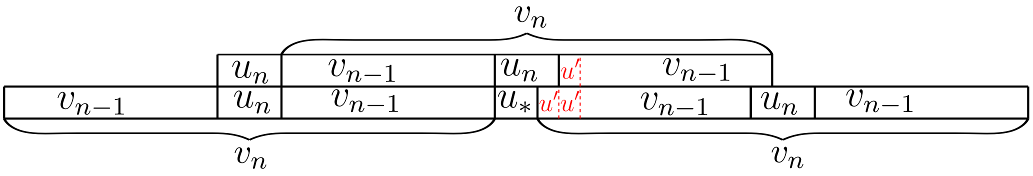


Figure 4.6: The case when $|u_n| > |u_*|$.

the shift $|u'| < k_{\text{max}} < \mu_1 < 3\mu_1$. Contradiction.

Case II

Let $v_{n-1} = v'v''v'''$ with $|v'| = |v'''| = 2\mu_1$. Let $u_n = w[|v_n|, |v_n u_n|)$ occur in $v' = v_{n-1}[0, 2\mu_1)$ or $v''' = v_{n-1}[|v_{n-1}| - 2\mu_1, |v_{n-1}|)$. Here we are not interested if u_n occurs in $v_{n-1} \in \text{Pref}(v_n)$ or in $v_{n-1} \in \text{Suff}(v_n)$. If u_n occurs in v' starting at position α, then, as

$$w[|v_n|, |v_n u_n v_{n-1}|) = u_n v_{n-1},$$

we obtain a shift of v_{n-1} of length $\alpha + |u_n|$ to the right of u_n (see Figure 4.7). Hence v_{n-1} is periodic with the length of the shift

$$\alpha + |u_n| \leq |v'| = 2\mu_1.$$

Contradiction.

Similarly, if u_n occurs in v''' ending in position $|v_{n-1}| - \beta$, then, since

$$w[|v_{n-1} u_n|, |v_n u_n|) = v_{n-1} u_n,$$

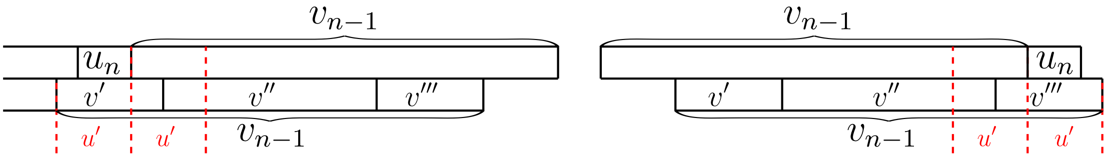


Figure 4.7: The word u_n occurs in v' or v'''.

we have shift of v_{n-1} of length $\beta + |u_n|$ to the left of u_n (see Figure 4.7), therefore v_{n-1} is periodic with a period

$$\beta + |u_n| \leq |v'''| = 2\mu_1.$$

Contradiction.

Case III

Let $v_{n-1} = v'v''v'''$, where $|v'| = |v'''| > \mu_1$. Let $u_n = w[|v_n|, |v_n u_n|)$ occur in the word $v'' = v_{n-1}[|v'|, |v'v''|)$. First we consider the case, when u_n occurs in v'' which is a factor of $v_{n-1} \in \mathrm{Pref}(v_n)$. If $u_n = w[|v_n|, |v_n u_n|)$ occurs in v_{n-1} starting in position α (see Figure 4.8), then v_n is both $\alpha + |u_*|$ and $\alpha + |u_n|$ periodic.

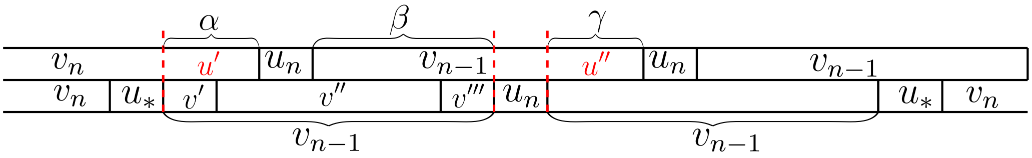


Figure 4.8: The word u_n occurs in v''.

a) If $|u_*| \neq |u_n|$, then according to Theorem 1: if

$$\begin{aligned}&\alpha + |u_*| + \alpha + |u_n| - gcd(\alpha + |u_*|, \alpha + |u_n|)\\ =&2\alpha + |u_n| + |u_*| - gcd(\alpha + |u_*|, \alpha + |u_n|) < |v_n|,\end{aligned}$$

then v_n is also $gcd(\alpha + |u_*|, \alpha + |u_n|)$ periodic. Indeed, we have

$$2\alpha + |u_n| + |u_*| - gcd(\alpha + |u_*|, \alpha + |u_n|) < 2|v_{n-1}| - 2\mu_1 + 2\mu_1 - 1 < |v_n|,$$

therefore v_n is

$$\begin{aligned}gcd(\alpha + |u_*|, \alpha + |u_n|) &= gcd(\min(\alpha + |u_*|, \alpha + |u_n|), ||u_*| - |u_n||)\\ &\leq ||u_*| - |u_n|| < \mu_1\end{aligned}$$

periodic. Contradiction.

b) If $|u_*| = |u_n|$, then we consider u' and u'' (see Figure 4.8), two suffixes of v_{n-1}, and compare their lengths α and γ:

$$\alpha = |u'| = |v_{n-1}| - \beta - |u_n| = \gamma.$$

Hence u_*u' and u_nu'' are two suffixes of v_{n-1} of the same length. From here and the equality of lengths of u_n and u_* we obtain that

$$u_n = u_*$$

as prefixes of the same length of equal words. Contradiction.

If $u_n = w[|v_n|, |v_nu_n|)$ occurs in $v_{n-1} \in \text{Suff}(v_n)$ ending in position $|v_{n-1}| - \alpha$ (see Figure 4.9), then v_n is both $\alpha + |u_*|$ and $\alpha + |u_n|$ periodic. Analogously as before we obtain:

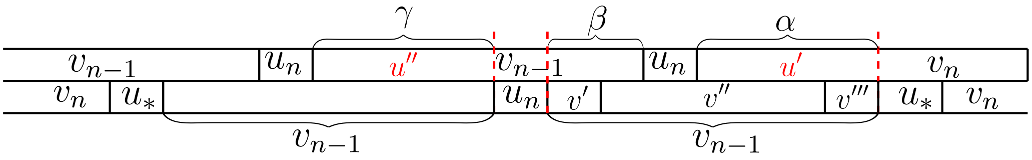


Figure 4.9: The word u_n occurs in v''.

a) If $|u_*| \neq |u_n|$, then v_n is also

$$gcd(\alpha + |u_*|, \alpha + |u_n|) < \mu_1$$

periodic, which leads to contradiction.

b) If $|u_*| = |u_n|$, then $|u''| = |u'|$ (see Figure 4.9), hence $u'' = u'$ as prefixes of v_{n-1} of equal length. From here $u_* = u_n$ as suffixes of equal length of the same word $u''u_n = u'u_* \in$ Pref(v_{n-1}).

We have proved that $w = v_nu_nv_n$ does not occur in v_{n+K+2}. It remains to show that this implies existence of a return word to w such that its length exceeds $K \cdot |w|$. As v_{n+K+2} does not contain w as a factor, there exists a return word r_w to w such that r_ww contains v_{n+K+2} as a factor. We need to prolong v_{n+K+2} for at least two symbols to obtain r_ww (if we add one symbol to the right or to the left, we can obtain maximum one occurrence of w). From here

$$|r_w| = |r_ww| - |w| \geq |v_{n+K+2}| + 2 - |w|. \tag{4.34}$$

Now we estimate the length of w. By Lemma 58:

$$|w| = |v_n u_n v_n| \leq 2|v_n| + \mu_1 \leq 2(2^{n+1} - 1)\mu_1 + \mu_1 < 2^{n+2}\mu_1. \tag{4.35}$$

Next, from (4.34), (4.35), and Lemma 58 we have

$$\begin{aligned}
|r_w| &\geq |v_{n+K+2}| + 2 - |w| \\
&> (2^{n+K+3} - 1)\mu_0 + 2 - 2^{n+2}\mu_1 \\
&= 2^{n+K+3}\mu_0 - \mu_0 + 2 - 2^{n+2}\mu_1 \\
&> 2^{n+K+2}\mu_1 - 2^{n+1}\mu_1 - 2^{n+2}\mu_1 \\
&= \left(2^K - \frac{3}{2}\right) \cdot 2^{n+2}\mu_1 \\
&> \left(2^K - \frac{3}{2}\right) \cdot |w|.
\end{aligned}$$

Finally, we conclude the proof by observing that for each integer $K > 1$ we have $2^K - \frac{3}{2} > K$. Hence

$$|r_w| > K \cdot |w|,$$

and x is not LR. Contradiction. □

Conclusions

The main results of the thesis are the necessary and sufficient condition of the aperiodicity of all arithmetical subsequences of finitely generated bi-ideals and bounded bi-ideals over an arbitrary alphabet, and the efficient algorithm for deciding whether all arithmetical subsequences of a finitely generated bi-ideal with given basis are or are not aperiodic. Thesis provides also results for a special case when the basis of a finitely generated bi-ideal contains only two words. It turns out that if the basis of a finitely generated bi-ideal contains only two basis words such that the greatest common divisor of their lengths is equal to 1, and the difference of their lengths is odd number, then all arithmetical subsequences of this bi-ideal are aperiodic if and only if it contains as factors at least two different letters.

Together with co-authors we made a research on possible use of finitely generated bi-ideals in cryptography by modifying so called shrinking generator. We replaced the S-sequence by finitely generated bi-ideals and found that for each A-sequence there exist infinitely many S-sequences such that resulting shrunk word is aperiodic. It turns out that we have to put some simple restrictions on suffixes of the first two basis words of a finitely generated bi-ideal. The practical testing done by co-author Edmunds Cers showed that resulting sequence performs well in statistical tests. We also gave some conditions on basis of finitely generated bi-ideal under which the resulting pseudo-random sequence is aperiodic for any periodic non-trivial A-sequence. The thesis provides some connection of aperiodicity of all arithmetical subsequences of a finitely generated bi-ideal to the aperiodic shrinking generator. A universal bi-ideal does not contain an arithmetical subsequence of the form $0000\cdots0\cdots$. We give an example of a finitely generated bi-ideal with all arithmetical subsequences aperiodic which is not a universal bi-ideal. Thus the aperiodicity of all subsequences of a finitely generated bi-ideal is not the sufficient condition for a finitely generated bi-ideal to be universal. The problem, is the aperiodicity of all subsequences of a finitely generated bi-ideal the necessary condition for a finitely generated bi-ideal to be universal, remains unsolved. Other problems for future work could be finding a general condition for a finitely generated bi-ideal to be universal or considering different modification of shrinking generator or other pseudo-random number generator. Author is convinced that if we substitute the S-sequence of the shrinking generator with a bounded bi-ideal, then we will have a larger class of universal bi-ideals than in already proposed construction.

Finally, while it was known that finitely generated bi-ideals are linearly recurrent, in case of bounded bi-ideals this problem was not solved. We introduce the notion of a completely bounded bi-ideal and prove that a bounded bi-ideal is linearly recurrent if and only if it is a completely bounded bi-ideal. This result helps us to describe better how finitely generated bi-ideals and bounded bi-ideals relate to other classes of right infinite words. Since the famous Thue-Morse word is linearly recurrent (as uniformly recurrent morphic infinite word), but it is not a bounded bi-ideal (Buls and Lorencs, 2006), then we conclude that the class of bounded bi-ideals intersects with the class of linearly recurrent words but neither of these classes is a proper subclass of another one (see Figure 4.10).

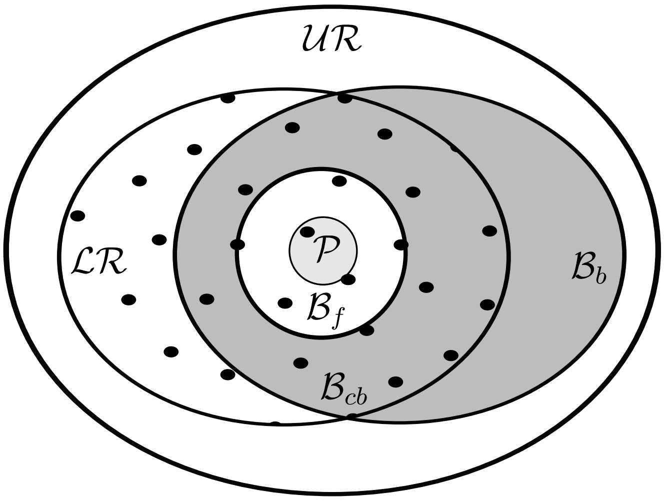


Figure 4.10: Hierarchy of uniformly recurrent words ($\mathcal{UR}$): the class of linearly recurrent words – $\mathcal{LR}$, the class of bounded bi-ideals – $\mathcal{B}_b$, the class of completely bounded bi-ideals – $\mathcal{B}_{cb}$, the class of finitely generated bi-ideals – $\mathcal{B}_f$, the class of periodic words – $\mathcal{P}$.

List of attended Conferences

International conferences

- 15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC 2013), September 23–26, 2013, Timisoara, Romania,
  presentation – "Bounded Bi-ideals and Linear Recurrence".

- 14th Mons Days of Theoretical Computer Science (JM2012), September 11–14, 2012, Louvain-La-Neuve, Belgium,
  presentation – "Arithmetical Subsequences of Finitely Generated Bi-ideals".

- 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC 2011), September 26–29, 2011, Timisoara, Romania,
  presentation – "On a Non-periodic Shrinking Generator".

- 1st International Conference of Students, Postgraduates and Young Scientists Theoretical and Applied Aspects of Cybernetics (TAAC), February 21–25, 2011, Kiev, Ukraine,
  presentation – "Merge of Right Infinite Words".

- 4th International conference Applied Information and Communication Technologies, April 22–23, 2010, Jelgava, Latvia.
  presentation – "Generalization of Toeplitz Words".

Domestic conferences and seminars

- 71st Conference of University of Latvia, February 28, 2013, Riga,
  presentation – "Aritmētisko apakšvirkņu aperiodiskuma problēma bi-ideālos".

- 2nd Joint Estonian-Latvian Theory Days, September 27–30, 2012, Medzabaki,
  presentation – "Aperiodic Modification of the Shrinking Generator".

- 70th Conference of University of Latvia, February 23, 2012, Riga,
  presentation – "Filtrācijas problēma".

- 69th Conference of University of Latvia, March 17, 2011, Riga,
  presentation – "Bezgalīgu vārdu filtrēšana".

- 8th Conference of Latvian Mathematical Society, April 9–10, 2010, Valmiera, presentation – “Toeplica vārdi”.
- 68th Conference of University of Latvia, March 4, 2010, Riga, presentation – “Toeplica vārdi”.

Bibliography

Allouche, J.-P. and J. Shallit. 2003. *Automatic sequences: Theory, applications, generalizations*, Cambridge University Press, Cambridge, UK.

Avgustinovich, S.V., J. Cassaigne, and A.E. Frid. 2006. *Sequences of low arithmetical complexity*, Theoret. Informatics Appl. **40**, 569–582.

Avgustinovich, S.V., D. Fon-Der-Flaass, and A.E. Frid. 2003. *Proceedings of the international colloquium on words, languages & combinatorics iii, kyoto, japan, march 14-18, 2000*, Words, languages & combinatorics.

Berstel, J., L. Boasson, O. Carton, B. Petazzoni, and J.-E. Pin. 2006. *Operations preserving regular languages*, Theoretical Computer Science **354**, no. 3, 405–420.

Berstel, J. and D. Perrin. 2007. *The origins of combinatorics on words*, European Journal of Combinatorics **28**, no. 3, 996–1022.

Buls, J. and A. Lorencs. 2006. *From bi-ideals to periodicity*, Proceedings of the 11th mons days of theoretical computer science, pp. 97–110.

———. 2008. *From bi-ideals to periodicity*, RAIRO-Theoretical Informatics and Applications **42**, no. 3, 467–475.

Caballero-Gil, P., A. Fuster-Sabater, and M. E Pazo-Robles. 2009. *New attack strategy for the shrinking generator*, Journal of Research and Practice in Information Technology **41**, 181–190.

Cassaigne, J. 1996. *Special factors of sequences with linear subword complexity*, in Developments in Language Theory II., World Sci. Publ., Singapore, 25–34.

———. 2003. *Constructing infinite words of intermediate complexity*, Developments in Language Theroy, Lect. Notes in Computer Science, Vol. **2450**, 173–184.

———. 2009. *s-adiques*, See https://www.lirmm.fr/arith/wiki/PytheasFogg/S-adiques.

Cassaigne, J. and A.E. Frid. 2007. *On arithmetical complexity of sturmian words*, Theoret. Comput. Sc. **380**, 304–316.

Cers, E. 2010. *An unique basis representation of finitely generated bi-ideals*, proceedings of the 13th mons theoretical computer science days (jm 2010), universite de picardie jules verne.

———. 2012. *Finitely generated bi-ideals and the semilattice of machine invariant ω-languages*, Ph.D. Thesis.

Cobham, A. 1968. *On the hartmanis-stearns problem for a class of tag macines*, Proceedings of the 9th annual symposium on switching and automata theory, ieee computer society, pp. 51–60.

Coppersmith, D., H. Krawczyk, and Y. Mansour. 1994. *The shrinking generator*, Proceedings of the 13th annual international cryptology conference on advances in cryptology, pp. 22–39.

Coudrain, M. and M.P. Schützenberger. 1966. *Une condition de finitude des monoides finiment engendres*, CR Acad. Sci., Paris, Ser. A **262**, 1149–1151.

de Luca, A. and S. Varricchio. 1999. *Finiteness and regularity in semigroups and formal languages*, Springer-Verlag, Berlin, Heidelberg.

Deviatov, R. 2008. *On subword complexity of morphic sequences*, Computer Science – Theory and Applications, Lect. Notes in Computer Science, Vol. **5010**, 146–157.

Durand, F. 1998. *A characterization of substitutive sequences using return words*, Discrete Math. **179**, 89–101.

———. 2000. *Linearly recurrent subshifts have a finite number of non-periodic subshift factors*, Ergod. Th. and Dynam. Sys. **20**, 1061–1078.

———. 2003. *Corrigendum and addendum to: Linearly recurrent subshifts have a finite number of non-periodic subshift factors*, Ergod. Th. and Dynam. Sys. **23**, 663–669.

Durand, F., B. Host, and Skau C. 1999. *Substitution dynamical systems, bratteli diagrams and dimension groups*, Ergod. Th. and Dynam. Sys. **19**, 953–993.

Durand, F., J. Leroy, and G. Richomme. 2013. *Do the properties of an s-adic representation determine factor complexity?*, J. of Integer sequences **16**, 1–30. Article 13.2.6.

Ferenczi, S. 1996. *Rank and symbolic complexity*, Ergodic Theory Dynam. Systems **16**, 663–682.

Fine, N.J. and H.S. Wilf. 1965. *Uniqueness theorem for periodic functions*, Proceedings of the american mathematical society, pp. 109–114.

Frid, A.E. 2003. *Arithmetical complexity of symmetric d0l words*, Theoret. Comput. Sc. **306**, 535–542.

———. 2005a. *A lower bound for the arithmetical complexity of sturmian words*, Siberian Electronic Mathematical Reports **2**, 14–22.

———. 2005b. *Sequences of linear arithmetical complexity*, Theoret. Comput. Sc. **339**, 68–87.

———. 2006. *On possible growths of arithmetical complexity*, Theoret. Informatics Appl. **40**, 443–458.

Gelfond, A.O. 1968. *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. **13**, 259–2651.

Hu, Y., X. Liao, K.-W. Wong, and Q. Zhou. 2009. *A true random number generator based on mouse movement and chaotic cryptography*, Chaos Solitons and Fractals **40**, 2286–2293.

Karhumäki, J. 2004. *Combinatorics on words: A new challenging topic*, Technical Report 645, Turku Centre for Computer Science.

L'Ecuyer, P. 1998. *Random number generation*, Handbook on simulation.

Leroy, J. 2012. *Some improvements pf the s-adic conjecture*, Advances in Applied Mathematics **48**, 79–98.

Leroy, J. and G. Richomme. 2013. *A combinatorial proof of s-adicity for sequences with linear complexity*, Integers **13**, 1–19. Article No. 5.

Lorencs, A. 2012. *The identity problem of finitely generated bi-ideals*, Acta Informatica **49**, no. 2, 105–115.

Lothaire, M. 1983. *Combinatorics on words. encyclopedia of mathematics and its applications 17*, Cambridge University Press, Cambridge, UK.

______. 2002. *Algebraic combinatorics on words. encyclopedia of mathematics 90*, Cambridge University Press, Cambridge, UK.

Marsaglia, G. 1996. *Diehard: a battery of tests of randomness*, See http://www.stat.fsu.edu/pub/diehard/.

Morse, M. and G.A. Hedlund. 1938. *Symbolic dynamics*, Amer. J.Math. **60**, 815–866.

______. 1940. *Symbolic dynamics ii*, Amer. J.Math. **62**, 1–42.

Oishi, S. and H. Inoue. 1982. *Pseudo-random number generators and chaos*, Transactions of the Institute of Electronics and Communication Engineers of Japan E **65**, 534–541.

Pansiot, J.-J. 1984. *Complexité des facteurs des mots infinis engendrés par morphismes itérés*, Automata, Languages and Programming, Lect. Notes in Computer Science, Vol. **172**, 380–389.

Patidar, V., K. K. Sud, and N. K. Pareek. 2009. *A pseudo random bit generator based on chaotic logistic map and its statistical testing*, Informatica **33**, no. 4, 441–452.

Phatak, S.C. and S. Suresh Rao. 1995. *Logistic map: A possible random number generator*, Physical Review E **51**, no. 4, 3670–3678.

Reid, L.W. 1910. *The elements of the theory of algebraic numbers*, The Macmillan Company, New York, US.

Rosen, K.H. 2005. *Elementary number theory and its applications*, Addison Wesley.

Sandri, G.H. 1992. *A simple nonperiodic random number generator: A recursive model for the logistic map*, Technical Report GL-TR-89-1066, Boston University College of Engineering and Center for Space Physics Boston.

Schneier, B. and P. Sutherland. 1995. *Applied cryptography: protocols, algorithms, and source code in c*, John Wiley & Sons, Inc. New York, NY, USA.

Simon, I. 1988. *Infinite words and a theorem of hindman*, Rev. Mat. Apl. **9**, 97–104.

Thue, A. 1906. *Über unendliche zeichenreihen*, Norske Vid. Selsk. Skr. I Math-Nat. Kl. 7, 1–22.

______. 1912. *Über die gegenseitige loge gleicher teile gewisser zeichenreihen*, Norske Vid. Selsk. Skr. IMath-Nat. Kl. Chris. 1, 1–67.

Van der Waerden, B.L. 1927. *Beweis einer baudet'chen vermutung*, Nieuw. Arch. Wisk. **15**, 212–216.

Zimin, A.I. 1982. *Blocking sets of terms*, Matematicheskii Sbornik **161**, no. 3, 363–375.

Author’s publications

Bērziņa I., Buls J., Bēts R. *Bounded Bi-ideals and Linear Recurrence.* Accepted for publication in: proceedings of the 15^{th} International symposium on symbolic and numeric algorithms for scientific computing (SYNASC 2013).

Bērziņa I., *Arithmetical Subsequences of Finitely Generated Bi-ideals*, 14th Mons Days of Theoretical Computer Science (JM2012), September 11–14, 2012, Louvain-La-Neuve, Belgium. Extended abstract (7 pages).

Bērziņa I., Bēts R., Buls J., Cers E. and Kuleša L. *On a non-periodic shrinking generator.* In: proceedings of the 13^{th} International symposium on symbolic and numeric algorithms for scientific computing (SYNASC 2011), IEEE Computer Society, 348–354, 2011. SCOPUS

Bērziņa I. *Merge of Right Infinite Words.* In: proceedings of 1st International conference of students, postgraduates and young scientists Theoretical and Applied Aspects of Cybernetics (TAAC), p. 8–11, Kiev (Ukraine), 2011.

Bērziņa I., *Generalization of Toeplitz Words*, In: proceedings of the 4-th International Scientific Conference Applied Information and Communication Technologies, 336–340, 2010. SCOPUS

Bērziņa I., *Arithmetical Subsequences of Finitely Generated Bi-ideals and Bounded Bi-ideals.* Submitted.