Defect Theorems
and
Infinite Words

Ján Maňuch

Defect Theorems and Infinite Words

Ján Maňuch

To be presented, with the permission of the Faculty of Mathematics and Natural Sciences of the University of Turku, for public criticism in the Auditorium of the Computer Science Department on June 28th, 2002, at 12 noon.

University of Turku
Department of Mathematics
FIN–20014 Turku, Finland
2002

Supervisor

Professor Juhani Karhumäki

Department of Mathematics
University of Turku
FIN–20014 Turku, Finland

Opponent

Professor Christian Choffrut

Université Denis Diderot – Paris VII, LIAFA
Tour 55–56, 1er étage, 2 place Jussieu
F–75251 Paris Cedex 05, France

Reviewers

Professor Wojciech Rytter

Department of Computer Science
Liverpool University
Peach Street, Liverpool L69 7ZF, U.K.

Professor Lila Kari

Department of Computer Science
University of Western Ontario
London, Ontario Canada N6A 5B7

ISBN 951–29–2288–3
ISSN 1239–1883
Painosalama Oy
Turku, Finland (2002)

To the country of thousands of lakes

Acknowledgments

I am immensely thankful to my supervisor, Professor Juhani Karhumäki, who invited me to Finland, introduced me to the many interesting problems of Combinatorics on Words, showed me how to cope with them, and guided me through all the steps of scientific life. He also took care of my well-being in Finland, gave me the honor to visit his cozy home several times and even let me experience his hobby, bird banding, in Finnish wilderness.

I would also like to thank the co-authors of the articles on which the thesis is based: Dr Wojciech Plandowski — my first experience of intense scientific cooperation, Dr Julien Cassaigne — who impressed me with his adventurous soul, and RNDr Pavol Ďuriš — with whom I spent many hours discussing the presence and the future of Mathematics, Computer Science, Artificial Intelligence and this World.

Many thanks to my reviewers, Professor Wojciech Rytter and Professor Lila Kari, for carefully reading my thesis. I am very grateful to Professor Christian Choffrut for accepting to be my opponent.

I have also benefited from the research cooperation between the University of Turku and the RWTH Aachen in Germany which resulted in solving two open problems.

Turku Centre for Computer Science together with the Department of Mathematics at the University of Turku provided me with excellent working conditions. Therefore I will use this opportunity to thank all the people who make these institutions work perfectly, starting from the secretaries, with special thanks to Leena Palmulaakso–Nylund, and ending with the directors of TUCS, Professor Ralph Back and Professor Timo Järvi, and the leader of the department, Professor Juhani Karhumäki.

Turku, June 2002

Ján Maňuch

Contents

Chapter 1

Introduction

In this work we attack different problems of Combinatorics on Words.

Combinatorics on Words is a rather new field of Theoretical Computer Science, although the first papers on words were written already at the beginning of 20th century by A. Thue, *cf.* [Th]. A pioneering paper on modern Combinatorics on Words is [LeS]. Actually, this paper deals with some problems related to several topics studied in this work, *cf.* mainly Section 3.2 and Chapter 5. The first monograph on Combinatorics on Words appeared as late as in the year 1983: [Lo] — a common project of several mathematicians. Recently, the book is followed by two new surveys on the field: [ChK] — studying combinatorial properties of words from the point of view of Formal Languages, and, just appeared, the new Lothaire: "Algebraic combinatorics on words".

One of the fundamental results on words is the *defect theorem*, *cf.* [Lo] and [BPPR]. Intuitively it states that if n words satisfy a non-trivial relation then these words can be expressed as products of at most $n-1$ words. Actually, as discussed in [ChK], for example, there does not exist just one defect theorem but several ones depending on restrictions put on the required $n-1$ words. It is also well-known that the non-trivial relation above can be replaced by a weaker condition, namely by the non-trivial one-way infinite relation, *cf.* [HK] or [Br].

The goal of Chapter 3 is to look for defect theorems for bi-infinite words. In a strict sense such results do not exist: the set $X = \{ab, ba\}$ of words satisfies a bi-infinite non-trivial relation since $(ab)^{\mathbb{Z}} = (ba)^{\mathbb{Z}}$, but there exists no word t such that $X \subseteq t^+$. However, we are going to prove several results which can be viewed as defect theorems for bi-infinite words.

To describe the results of Chapter 3, let w be a bi-infinite word, *i.e.,* an element of $\Sigma^{\mathbb{Z}}$, and X a finite subset of Σ^+. We say that w possesses an X-factorization if $w \in X^{\mathbb{Z}}$, and that w possesses two different X-factorizations, if it possesses two X-factorizations such that they do not match at least in one point of w. Further, the combinatorial rank of a set X, denoted by $\mathrm{rank}_{\mathrm{c}}(X)$, is the cardinality of the smallest set Y such that $X \subseteq Y^+$. We prove the following results:

- Section 3.1: If a non-periodic bi-infinite word w has two different X-factorizations then the combinatorial rank of X, denoted by $\operatorname{rank_c}(X)$, is at most $\operatorname{card}(X)-1$. Moreover, if $\operatorname{rank_c}(X) = \operatorname{card}(X)$ then the number of bi-infinite words with two different X-factorizations is at most $\frac{1}{2}\operatorname{size}(X)$.

- Section 3.2: Let $X = \{\alpha, \beta\}$ be a two-element code. If a bi-infinite word w possesses two different X-factorizations then either the X-factorizations are shift-equivalent and there exists a word $t \in \alpha\beta^+ \cup \alpha^+\beta$ such that $w = t^{\mathbb{Z}}$, or the primitive roots of α and β are conjugates. Moreover, there are at most two bi-infinite word possessing two different X-factorizations.

We want to emphasize that a restriction to *non-periodic* bi-infinite words is necessary, as shown by the example of $X = \{ab, ba\}$, and even more that the above theorems require to consider the *combinatorial rank*. The later restriction is quite interesting since in all previous defect theorems, see [ChK] and Section 2.2, any of the notions of the rank can be used to witness the defect effect.

The results in Section 3.2 are related to some considerations of [LeS] and to the main result of [LRLR]. In fact, as we realized recently, the main theorem of Section 3.2, Theorem 3.11, can be, after some effort, deduced from considerations of these two papers. However, our proof is self-contained and essentially shorter, and moreover formulated directly to yield a defect-type of theorem.

As argued in [HKP] defect theorems can be viewed as a weak dimension property of words. It is weak since a finite set X of words can satisfy several different, or independent as it is formalized in [HKP], relations without forcing a larger defect effect than 1, *i.e.,* a larger defect effect than is forced by a single relation. In Chapter 4 we ask to find conditions (on relations or sets of words) which yield a *cumulative defect effect*, *i.e.,* if the set X of n words satisfy k relations then X is of rank at most $n - k$.

There are only very few results known in this direction. The Graph Lemma, *cf.* Lemma 2.3 in Section 2.3, is such an example where the type of relations is restricted, *cf.* [ChK, HK]. A similar deep result is proved in [Br], extending ideas of [Ka1, Ka2, Ho], where it is shown that if X is a code and has unbounded synchronizing delay in both directions then the rank of X is at most $\operatorname{card}(X)-2$.

In Chapter 4, we interpret, in a natural way, a relation on words from X as a double X-factorization of some infinite word. We ask if the fact that a non-periodic bi-infinite word possesses k disjoint X-factorizations implies that $\operatorname{rank_c}(X) \leq \operatorname{card}(X) - k + 1$, *cf.* Problem 4.1. By our defect theorem for bi-infinite words (Theorem 3.3 in Section 3.1), the answer is “yes” in the case $k = 2$.

In Section 4.1 we prove that if X is a *prefix set* then the answer is affirmative also in the case $k = 3$.

Further, in Section 4.2, we consider the connections of the above problem in the case $k = \mathrm{card}(X)$ to the Critical Factorization Theorem, or more precisely, to the conjecture about its application stated in [Lo]. We will give several examples that the conjecture is false, and hence we are not able to obtain a positive answer to our problem in the case $k = \mathrm{card}(X)$. However, as a consequence of the application of the Critical Factorization Theorem we have that the number of disjoint X-factorizations of a non-periodic bi-infinite word is at most $\mathrm{card}(X)$.

The defect theorems motivated a research on words equations starting by a seminal paper of Makanin in 1976, *cf.* [Mak]. Despite the fact that many fundamental problems, such as the exact complexity of the satisfiability problem, *cf.* [Pl], or the maximal size of independent systems of equations in n variables, *cf.* [HKP], are not solved, one can say that there exists a deep and rich theory on word equations.

If *language equations*, as extensions of word equations, are considered the situation changes drastically: almost nothing is known about those. Recently, the commutation equation $XZ = ZX$ for languages has been studied in a number of papers, *cf.* [Ra, CKO, KPe, HP, Ka3, KLP] for a survey. In certain cases, for example when $\mathrm{card}(X) \leq 3$ or X is a code, it is completely solved: Z must be of the form $Z = \cup_{i \in I} \varrho(X)^i$ with $I \subseteq \mathbb{N}$, and $\varrho(X)$ being the primitive root of X, *i.e.,* the minimal set having the set X as its power. In these cases this characterization gives an affirmative answer to an old problem of Conway, *cf.* [Co], asking whether the (unique) maximal set Z commuting with a given rational X is also rational. Note also that, in these cases, the sets X and Z are expressible as unions of powers of a common set, *i.e.,* the commutation equation for languages in these particular cases causes a defect effect. As an example in [CKO] shows, this is not true in the general case even for the commutation equation.

In Chapter 5 we will consider the *conjugacy equation* $XZ = ZY$. Since, even the commutation equation seems to be a rather difficult problem solved only in special cases, we cannot expect the conjugacy equation to be easy. Hence, we will concentrate on the one of the simplest cases when both the sets X and Y are binary. We are able to solve this problem completely, *i.e.,* to characterize all binary sets X and Y for which there exists a non-empty set Z such that $XZ = ZY$, as well as to characterize such sets Z. However, even in this very restricted case we cannot witness a defect effect, *cf.* Example 5.2 in Chapter 5.

Finally, in Chapter 6 we look at infinite words from a different perspective. In [CuK] and [HKL] two new areas of investigation were introduced:

- the *descriptional complexity* of infinite words, *i.e.,* the comparative measure how complicated simple mechanisms are needed to generate particular infinite words;

- the *computational complexity* of infinite words, *i.e.,* the measure how much resources (such as time and space) are needed to generate a certain infinite word by a Turing machine.

The second paper concentrates on relations between these two complexities.

In [CuK, HKL, HK] several interesting problems are proposed. In Chapter 6 we solve a few of these problems, or in fact, in some cases, we show that they are equivalent to well-known hard open problems in the complexity theory of Turing machines.

In Section 6.3 we consider the open problem, proposed in [HKL], namely whether all infinite words generated by iterating deterministic generalized sequential machines, dgsm's for short, have logarithmic space complexity. As shown already in [HKL], the answer is "yes" if the dgsm has the maximal, *i.e.,* exponential, growth. We show that it is so also in the case when the dgsm has a smallest non-trivial growth ($\Theta(n \log n)$). In [Le] it is claimed that the answer to the problem is affirmative in general. Here we show that the problem is equivalent to an other hard open problem of complexity theory asking whether unary classes of languages P and DLOG are equivalent. Therefore, we believe that in the proof of [Le] some case must have been overlooked.

Another problem proposed in [HKL] is to find a concrete infinite word which cannot be generated in logarithmic space. In Section 6.4 we show that it is exactly as hard as the problem to find a concrete language, which does not belong to DSPACE(n).

Finally, in Section 6.5 we separate the classes of infinite words generated by double and triple D0L TAG systems as it was conjectured in [CuK].

A special attention is paid to the presentation. The definitions and proofs in this work are illustrated with the numerous figures, and we believe that they make the content of the work more comprehensive.

The thesis is based on the following papers [KMP, Man, KM, CKM, DM].

Chapter 2

Preliminaries

2.1 Basic definitions

In this section we fix our terminology and recall some basic notions and definitions of Combinatorics on Words. For undefined notions we refer a reader to [Lo] or [ChK]. We will pay the most of our attention to the notions of bi-infinite words and their factorizations which we will study in Chapters 3 and 4.

Let Σ be a finite non-empty set, called an *alphabet.* Elements of Σ are called *letters*, and finite sequences of letters are called *words.* The number of letters in the sequence, forming a word u, is the *length* of u, which we denote by $|u|$. In particular, the word of length 0 is called the *empty word*, and denoted by 1. The set of all words (resp. all non-empty words) over Σ is denoted by Σ^* (resp. Σ^+). The set Σ^* is naturally equipped with the operation *concatenation*, denote by ".", also called *product.* Obviously, each word has the unique representation as product of letters. Hence, Σ^* (resp. Σ^+) is the *free monoid* (resp. the *free semigroup*) generated by Σ.

We define three relations on words:

- u is a *prefix* of v, denoted by $u \leq v$, if there exists a word z such that $v = uz$;

- u is a *suffix* of v, if there exists a word z such that $v = zu$;

- u is a *factor* of v, if there exist words x and y such that $v = xuy$.

In the case when $u \neq v$, we call any of the above three relations *proper.* The relation "u is proper prefix of v" is denoted by $u < v$. We denote by $\mathrm{Pref}(v)$, $\mathrm{pref}(v)$, $\mathrm{Suff}(v)$, $\mathrm{suff}(v)$, $\mathrm{Fact}(v)$, the sets of all prefixes, proper prefixes, suffixes, proper suffixes, factors of a word v, respectively. All these notions can be generalized for the sets of words in a natural way.

Let u and v be words. If u is a prefix (resp. a suffix) of v then the word z such that $v = uz$ (resp. $v = zu$) is called the *left quotient* (resp. the *right quotient*) of v by u, denoted by $u^{-1}v$ (resp. vu^{-1}).

We say that words u and v are *left comparable* (resp. *right comparable*) if one of u and v is a prefix (resp. a suffix) of the other. Further, we say that a pair of words (u, v) *matches* a word w *at a position* (w_1, w_2) if $w = w_1w_2$, u and w_1 are right comparable, and v and w_2 are left comparable.

Let $a_1, \ldots, a_n \in \Sigma$ be the sequence of letters of a word v, *i.e.,* $v = a_1 \ldots a_n$. The *mirror image* of the word v is the word $a_n \ldots a_1$, denoted by v^{R}. The mirror image of a set $X = \{x_1, \ldots, x_m\}$ is the set $X^{\mathsf{R}} = \{x_1^{\mathsf{R}}, \ldots, x_m^{\mathsf{R}}\}$.

The sets of all *infinite* and *bi-infinite words* over Σ are denoted by $\Sigma^{\mathbb{N}}$ and $\Sigma^{\mathbb{Z}}$, respectively. Formally, an infinite word is a mapping $w : \mathbb{N} \to \Sigma$, and a bi-infinite word is a mapping $w : \mathbb{Z} \to \Sigma$. Usually, we write an infinite word w as

$$w = w_0w_1 \ldots , \qquad \text{with } w_i = w(i) \text{ for all } i \in \mathbb{N},$$

and similarly, we write a bi-infinite word w as

$$w = \ldots w_{-1}w_0w_1 \ldots , \qquad \text{with } w_i = w(i) \text{ for all } i \in \mathbb{Z}.$$

It is obvious that a bi-infinite word w and the bi-infinite word w' defined as $w'(k) = w(k_0 + k)$ for all $k \in \mathbb{Z}$ and a fixed $k_0 \in \mathbb{Z}$ represent the same word. Hence, by definition, we will consider w and w' as different representations of the same bi-infinite word.

Example 2.1. The bi-infinite word

$$w = (ab)^{\mathbb{Z}} = \ldots abab \ldots$$

has exactly two representations

$$w_1(n) = \begin{cases} a; & n \text{ is even,} \\ b; & n \text{ is odd,} \end{cases} \qquad \text{and} \qquad w_2(n) = \begin{cases} a; & n \text{ is odd,} \\ b; & n \text{ is even.} \end{cases} \tag{2.1}$$

A *factorization* of a word v is any sequence $(v_1, \ldots, v_k)$ of words such that

$$v = v_1v_2 \ldots v_k .$$

If the words $v_1, \ldots, v_k$ are elements of a set X, we say that the sequence $(v_1, \ldots, v_k)$ is an *X-factorization* of v. Similarly, an *X-interpretation* of v is any sequence $v_1, \ldots, v_k$ of words of X such that

$$pvs = v_1v_2 \ldots v_k$$

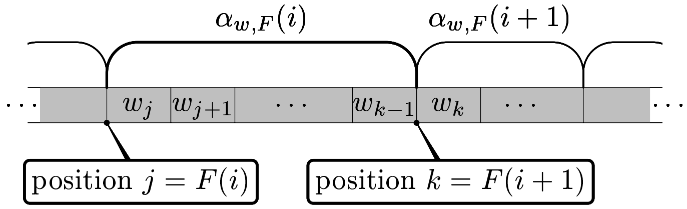


Figure 2.1: An illustration how the factorization F factorizes the word w.

for some words p and s satisfying $|p| < |v_1|$ and $|s| < |v_k|$.

A *factorization* of a bi-infinite word $w \in \Sigma^{\mathbb{Z}}$ is an increasing function $F : \mathbb{Z} \to \mathbb{Z}$. The range of F is called the *set of starting positions*, denoted by $F(\mathbb{Z})$. Indeed, the factorization F factorizes a bi-infinite word w into words:

$$\ldots, \alpha_{w,F}(-1), \alpha_{w,F}(0), \alpha_{w,F}(1), \ldots,$$

where

$$\alpha_{w,F}(i) = w_{F(i)}w_{F(i)+1}\ldots w_{F(i+1)-1}, \quad \text{for all } i \in \mathbb{Z},$$

i.e., the position $F(i)$ is the starting position of the factor $\alpha_{w,F}(i)$ in w, as depicted in Figure 2.1. Note that the way how a factorization factorizes a bi-infinite word depends on the representation of the bi-infinite word:

Example 2.2. Let $F(n) = 2n$ be a factorization and consider two representations (2.1) of the bi-infinite word $(ab)^{\mathbb{Z}}$. Then F factorizes w_1 into factors $\alpha_{w_1,F}(n) = ab$, and w_2 into factors $\alpha_{w_2,F}(n) = ba$, for all $n \in \mathbb{Z}$.

It is obvious that a factorization F and the factorization F' defined as $F'(k) = F(k_0+k)$, for all $k \in \mathbb{Z}$, and a fixed $k_0 \in \mathbb{Z}$ factorize the bi-infinite word w in the same way. Hence, similarly as for bi-infinite words, we will consider F and F' as different representations of the same factorization.

We say that two factorizations F_1 and F_2 are

- *different*, if they are not the same, *i.e.*, $F_1(\mathbb{Z}) \neq F_2(\mathbb{Z})$;
- *disjoint*, whenever the starting positions of all factors in F_1 are distinct from the ones in F_2, *i.e.*, $F_1(\mathbb{Z}) \cap F_2(\mathbb{Z}) = \emptyset$;
- *shift-equivalent* with respect to a representation of a bi-infinite word w, if there is a $k_0 \in \mathbb{Z}$ such that for all $k \in \mathbb{Z}$, $\alpha_{w,F_1}(k) = \alpha_{w,F_2}(k_0 + k)$.

Let us illustrate the above relations between factorizations on an example:

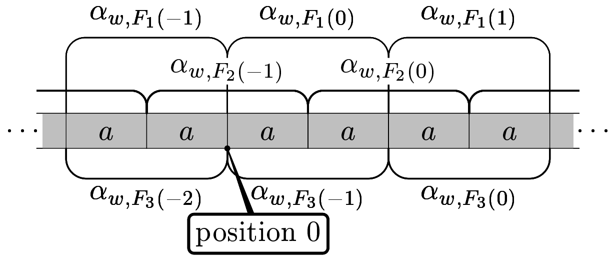


Figure 2.2: Three factorizations of the bi-infinite word $a^{\mathbb{Z}}$.

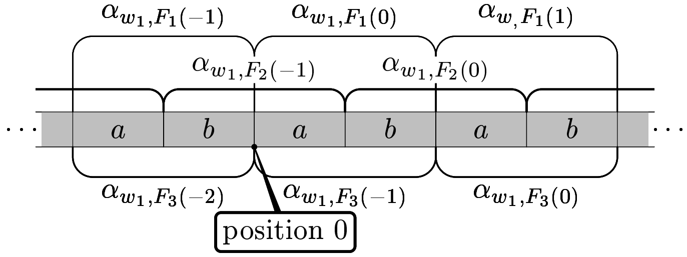


Figure 2.3: Three factorizations of the bi-infinite word $w_1 = (ab)^{\mathbb{Z}}$.

Example 2.3. Consider the bi-infinite word $w = a^{\mathbb{Z}}$ (it has only one representation) and the following three factorizations:

$$\begin{aligned} F_1(n) &= 2n\,, \\ F_2(n) &= 2n+1\,, \\ F_3(n) &= 2n+2\,. \end{aligned}$$

As it can be seen in Figure 2.2

- the factorizations F_1 and F_3 are the same, *i.e.*, different representations of the same factorizations. Indeed, $F_1(\mathbb{Z}) = 2\mathbb{Z} = F_3(\mathbb{Z})$;
- the factorizations F_1 and F_2 are disjoint, since $F_1(\mathbb{Z}) \cap F_2(\mathbb{Z}) = 2\mathbb{Z} \cap (2\mathbb{Z}+1) = \emptyset$.

In addition factorizations F_1 and F_2 are shift-equivalent with respect to w, since $\alpha_{w,F_1}(n) = aa = \alpha_{w,F_2}(n)$ for all $n \in \mathbb{Z}$. However, this is not true if we consider the bi-infinite word w_1, defined in Example (2.1), instead. Indeed, as Figure 2.3 shows, $\alpha_{w_1,F_1}(n) = ab$ and $\alpha_{w_1,F_2}(n) = ba$ for all $n \in \mathbb{Z}$.

Finally, we define a special type of factorizations, which we will study in more details in Chapter 3. Let X be a set of non-empty words. Let $\mathrm{Fact}(w, F)$

be the set of all the factors into which F factorizes w, *i.e.*,

$$\mathrm{Fact}(w, F) = \{\alpha_{w,F}(n);\ n \in \mathbb{Z}\}\,.$$

If $\mathrm{Fact}(w, F) \subseteq X$, we say that F is an X*-factorization* of the bi-infinite word w, and that w *possesses* an X-factorization F.

Next, we would like to define the mirror images of bi-infinite words and their factorizations in the way that the factors into which the mirror image of a factorization F factorizes the mirror image of a bi-infinite word w would be the mirror images of the factors into which F factorizes w. One way how to do it is as follows:

We define the *mirror image* w^{R} of a bi-infinite word w by the following formula

$$w^{\mathsf{R}}(n) = w(-n-1), \quad \text{for all } n \in \mathbb{Z},$$

and the mirror image F^{R} of a factorization F by the formula

$$F^{\mathsf{R}}(n) = -F(-n+1), \quad \text{for all } n \in \mathbb{Z}.$$

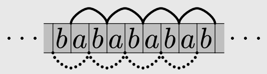
This definition is sound since F^{R} is also a growing function. Moreover, $F(\mathbb{Z}) = -F^{\mathsf{R}}(\mathbb{Z})$. The reader can check that the following conditions are satisfied for all $i \in \mathbb{Z}$

$$\begin{aligned} w^{\mathsf{R}}(F^{\mathsf{R}}(i)) &= w(F(-i+1)-1), \quad \text{and} \\ \alpha_{w^{\mathsf{R}},F^{\mathsf{R}}}(i) &= (\alpha_{w,F}(-i))^{\mathsf{R}}. \end{aligned} \tag{2.2}$$

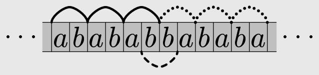
Therefore, $\mathrm{Fact}(w^{\mathsf{R}}, F^{\mathsf{R}}) = (\mathrm{Fact}(w, F))^{\mathsf{R}}$, and in particular, if F is an X-factorization of a bi-infinite word w then F^{R} is an X^{R}-factorization of w^{R}.

If a bi-infinite word possesses at least two different X-factorizations, we say that it is X*-ambiguous.* Let $\mathrm{Amb}(X)$ denote the set of all X-ambiguous bi-infinite words.

Example 2.4. Consider the set $X = \{ab, ba\}$. Then the bi-infinite word $(ab)^{\mathbb{Z}}$ is X-ambiguous, since it possesses two different X-factorizations:



On the other hand the bi-infinite word ${}^{\mathbb{N}}(ab)(ba)^{\mathbb{N}}$ is not X-ambiguous, it can be factorize over X in the unique way:



Indeed, any other factorization over a set containing only 2-letter words would require that $bb \in X$ (depicted with a dashed line).

Note that the above properties of bi-infinite words (possessing an X-factorization and being X-ambiguous) were, in fact, defined for a particular representation of a bi-infinite word. But, clearly, if a representation of a bi-infinite word possesses an X-factorization (resp. is X-ambiguous) then all representations of that bi-infinite word do so.

2.2 Defect theorems and ranks

The defect theorem is one of the fundamental results on words, *cf.* [Lo, BPPR]. Intuitively it states that if n words satisfy a non-trivial relation then these words can be expressed as products of at most $n-1$ words. Actually, as discussed in [ChK], for example, there does not exist just one defect theorem but several ones depending on restrictions put on the required $n-1$ words.

We say that words x and y *commute* if they satisfy the equation $xy = yx$. The following conditions are equivalent, *cf.* [Lo]:

- words x and y commute;
- words x and y satisfy a non-trivial equation;
- words x and y have a common power;
- there exists a word t such that $x, y \in t^*$;
- $\rho(x) = \rho(y)$ (see Section 2.4).

The above claim is one of the basic facts of the theory of Combinatorics on Words. It can be viewed as an example of a defect effect for $n = 2$. Indeed, the condition $x, y \in t^*$ expresses that the "dimension" of the set $\{x, y\}$ is 1.

Hence, defect theorems can be viewed as different dimension properties of sets of words. We have several ways how define "dimension". There are two main approaches, combinatorial and algebraic.

The *combinatorial rank* of $X \subseteq \Sigma^+$ is defined by the formula

$$\mathrm{rank_c}(X) = \min\{\mathrm{card}(Y); \quad X \subseteq Y^+\}.$$

In order to give algebraic definitions of the dimension of a set of words, we have to define the following properties. We call a submonoid M of Σ^*

- *stable* if for all $u, v, uw, wv \in M$ then also $w \in M$;
- *right unitary* if for all $u, uw \in M$ then also $w \in M$;

- *left unitary* if for all $v, wv \in M$ then also $w \in M$.

The above properties relate to the codes, prefix sets and suffix sets, respectively. We say that a set of words X is a *code* if it satisfies the following condition: for all integers $n, m \geq 1$ and words $x_1, \ldots, x_n, y_1, \ldots, y_m \in X$

$$x_1 \ldots x_n = y_1 \ldots y_m \quad \text{implies} \quad n = m \text{ and } x_i = y_i \text{ for } i = 1, \ldots, n.$$

Further, we say that a set of words X is a *prefix set* (resp. a *suffix set*) if no word of X is a prefix (resp. a suffix) of another. The relations between properties of submonoids of Σ^* and properties of sets of words can be expressed as follows:

Lemma 2.1. [BP] *A submonoid of* Σ^* *is*

- *stable if and only if it is a free submonoid if and only if its minimal generating set is a code;*
- *right unitary if and only if its minimal generating set is a prefix set;*
- *left unitary if and only if its minimal generating set is a suffix set.*

Note that the intersection preserves any of the above properties. Therefore, we can define *the smallest free* (resp. *right unitary*) *submonoid of* Σ^* *which contains* X by the formulas:

$$\mathrm{FM}(X) = \bigcap_{\substack{X \subseteq M \subseteq \Sigma^* \\ M \text{ is free}}} M,$$
$$\mathrm{RUM}(X) = \bigcap_{\substack{X \subseteq M \subseteq \Sigma^* \\ M \text{ is right unitary}}} M.$$

The minimal generating set of $\mathrm{FM}(X)$ (resp. $\mathrm{RUM}(X)$) is called the *free* (resp. *prefix*) *hull* of X, denoted by $\hat{X}_\mathrm{f}$ (resp. by $\hat{X}_\mathrm{p}$).

Finally, we can define free and prefix ranks:

$$\mathrm{rank}_\mathrm{f}(X) = \mathrm{card}(\hat{X}_\mathrm{f}) \quad \text{and} \quad \mathrm{rank}_\mathrm{p}(X) = \mathrm{card}(\hat{X}_\mathrm{p}).$$

Let us recall the defect theorems formulated for the free and prefix ranks.

Theorem 2.2. [BPPR, Lo] *For each finite set* $X \subseteq \Sigma^+$ *we have*

- $\mathrm{rank}_\mathrm{f}(X) \leq \mathrm{card}(X)$, *and moreover, the equality holds if and only if* X *is a code;*
- $\mathrm{rank}_\mathrm{p}(X) \leq \mathrm{card}(X)$, *and moreover, the equality implies that* X *is a code.*

The above defined ranks satisfy the following condition

$$\mathrm{rank_c}(X) \leq \mathrm{rank_p}(X) \leq \mathrm{rank_f}(X) \leq \mathrm{card}(X)\,,$$

The following example (based on Example 4.1 of [ChK]) shows the above ranks define the different properties of sets of words.

Example 2.5. Consider the set

$$X = \{aa, aaba, bac, cbb, bbaa, baa\}\,.$$

The only minimal non-trivial relation in X^+ is

$$aa.bac.bbaa = aaba.cbb.aa\,.$$

Since X^+ is a subset of $\mathrm{FM}(X)$ which is stable, we conclude that $\mathrm{FM}(X)$ contains the words ba, c and bb. Now the set

$$X_1 = \{aa, ba, c, bb, baa\}$$

is a code such that X_1^+ contains X^+ and the elements ba, c and bb are necessarily contained in $\mathrm{FM}(X)$, hence it is the free hull $\hat{X}_\mathrm{f}$ of X.

Obviously, the set $X_1^+ = \mathrm{FM}(X)$ is a subset of $\mathrm{RUM}(X)$. The set X_1 is not a prefix set, hence by the right unitary condition, we have that $a \in \mathrm{RUM}(X)$. Similarly, we obtain that the set

$$\{a, ba, c, bb\}$$

is the prefix hull $\hat{X}_\mathrm{p}$ of X. Clearly, the combinatorial rank of X is 3. Consequently, we can conclude that

$$3 = \mathrm{rank_c}(X) < \mathrm{rank_p}(X) < \mathrm{rank_f}(X) < \mathrm{card}(X) = 6\,.$$

2.3 Graph Lemma

In order to formulate one crucial lemma, we need some terminology. We associate a finite set $X \subseteq \Sigma^+$ with a graph $\mathcal{G}_X = (V_X, E_X)$, called *the dependency graph* of X, as follows: the set V_X of vertices of $\mathcal{G}_X$ equals to X, and the set E_X of edges of $\mathcal{G}_X$ is defined by the condition

$$(x, y) \in E_X \qquad \text{iff} \qquad xX^\mathbb{N} \cap yX^\mathbb{N} \neq \emptyset\,.$$

Then we have:

Lemma 2.3. [ChK, HK] *For each finite set $X \subseteq \Sigma^+$, the prefix (resp. combinatorial) rank of X is at most the number of connected components of $\mathcal{G}_X$.*

Note that in Lemma 2.3 the fact that the set X contains only non-empty words is crucial. Therefore, any time we will use the lemma, we have to be sure that all words occurring in the relations are non-empty.

Second, note that if the set X satisfies a finite relation

$$x_1 \dots x_n = y_1 \dots y_m \,,$$

then it can be easily extended to an infinite one. Hence, in such a case, the dependency graph of X contains also the edge (x_1, y_1).

Let size(X) be the sum of lengths of words of X, also called the *size* of the set X.

2.4 Periodicity

Let $v = a_1 \dots a_n$ be a word with $a_1, \dots a_n \in \Sigma$ as its sequence of letters. If there exists an integer $p \in \mathbb{N}$ such that for every integer $i = 1, \dots, n-p$, $a_i = a_{i+p}$, then we say that p is *a period* of v. The minimal period of v is called *the period* of v, and denoted by per(v).

Let w be an infinite word. If there exists a positive integer p such that for every $n \in \mathbb{N}$, $w(n) = w(n+p)$, we say that w is *periodic* with *a period* p. If there exist integers $p > 0$ and $n_0 \geq 0$ such that for every integer $n \geq n_0$, $w(n) = w(n+p)$, then we say that w is *ultimately periodic.*

Let w be a bi-infinite word. If there exists a positive integer p such that for every $n \in \mathbb{Z}$, $w(n) = w(n+p)$, we say that w is *periodic* with a period p. Let F be a factorization of w. We say that F is *periodic* if there exist an integer $k \geq 1$ and words $x_0, \dots, x_{k-1}$ such that

$$\alpha_{w,F}(i) = x_{i \bmod k}, \quad \text{for all } i \in \mathbb{Z}.$$

Every word v can be expressed in the form $v = u^n$, where $n \geq 1$. The word u satisfying this condition is called *a root* of v. The root of the minimal length is called *the primitive root* of v, denoted by $\rho(v)$. A word v is *primitive* if $\rho(v) = v$.

The following lemma claims that for a primitive word u the situation when u is an inner factor of uu, see Figure 2.4, cannot occur.

Lemma 2.4. [ChK] *If the word u satisfies the relation*

$$uu = pus \quad \textit{with } p, s \neq 1,$$

then u is non-primitive. Moreover, there is a primitive word t such that

$$u, p, s \in t^+.$$

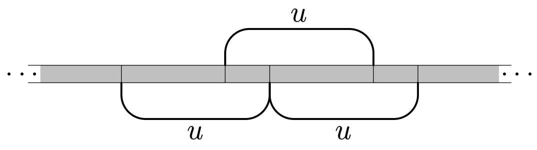


Figure 2.4: An illustration of the situation which cannot occur if u is primitive.

Lemma 2.5 (Fine and Wilf). [FW] *Let $u, v \in \Sigma^+$. The words u and v are powers of a common word if and only if the words $u^{\mathbb{N}}$ and $v^{\mathbb{N}}$ have a common prefix of a length at least $|u| + |v| - \gcd(|u|, |v|)$.*

As a corollary of Lemma 2.5 we have

Lemma 2.6. [LeS] *If non-empty words x, y and z satisfying the relation $x^m y = z^n$ (resp. $yx^m = z^n$) for some integers $m, n \geq 1$ then, either $|z| > (m-1)|x|$, or all the words x, y and z are powers of a common word.*

Proof. Consider the equation $x^m y = z^n$. If $|z| \leq (m-1)|x|$ then $|x| + |z| \leq |x^m|$, and thus the words x^m and z^n have a common prefix of a length at least $|x| + |z|$. By Lemma 2.5, the words x and z commute which implies that all the words x, y and z are powers of a common word. The proof for the equation $yx^m = z^n$ is essentially the same. □

In [LyS] a more intricate result was shown.

Lemma 2.7. [LyS] *If non-empty words x, y and z satisfying the relation $x^m y^n = z^p$ for some integers $m, n, p \geq 2$ then they are powers of a common word.*

However, the original proof in [LyS] is rather long and proved in a more general settings of free groups. Therefore, we refer a reader to a much shorter proof in [Sh].

2.5 Conjugacy

We say that two words u and v are *conjugates* if there exist words p and q such that

$$u = pq \quad \text{and} \quad v = qp\,.$$

We define a mapping $\mathsf{c} : \Sigma^* \to \Sigma^*$, called *cyclic permutation*, by the formulas

$$\begin{aligned} \mathsf{c}(1) &= 1\,, \\ \mathsf{c}(a_1 \ldots a_n) &= a_2 \ldots a_n a_1\,, \end{aligned}$$

where $a_1, \ldots, a_n \in \Sigma$. The equivalent definition of the conjugacy of x and y can be expressed as: words x and y are conjugates if there exists an integer k such that $x = \mathsf{c}^k(y)$.

As a consequence of Lemma 2.5 we have the following useful lemma and its corollary.

Lemma 2.8. [LeS] *Let $u, v \in \Sigma^+$. The primitive roots of u and v are conjugates if and only if the words $u^{\mathbb{N}}$ and $v^{\mathbb{N}}$ have a common factor of length at least $|u| + |v| - \gcd(|u|, |v|)$.*

Corollary 2.9. *If words u and v are conjugates then u is primitive if and only if v is primitive.*

The last, less obvious but very useful, characterization of a pair of conjugate words is formulated in the following lemma.

Lemma 2.10. [Lo] *Two non-empty words u and v satisfy the relation*

$$ut = tv$$

for some word t if and only if there exist words p and q such that pq is primitive and

$$u = (pq)^i, \quad v = (qp)^i, \quad \text{and} \quad t \in p(qp)^* \quad \text{for some } i \geq 1\,,$$

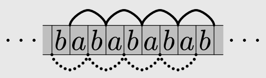
i.e., *u and v are conjugates.*

Chapter 3

Defect theorems for bi-infinite words

The goal of this chapter is to look for defect theorems for bi-infinite words. In a strict sense such results do not exist:

Example 3.1. The set $X = \{ab, ba\}$ is a code and of the combinatorial rank 2, although the bi-infinite word $(ab)^{\mathbb{Z}}$ possesses two disjoint, and even non-shift-equivalent, X-factorizations:



i.e., the set X satisfies a bi-infinite non-trivial relation $(ab)^{\mathbb{Z}} = (ba)^{\mathbb{Z}}$.

However, we are going to prove several results which can be viewed as defect theorems for bi-infinite words. In Section 3.1 we will prove that if a *non-periodic* bi-infinite word w has two different X-factorizations then the combinatorial rank $\mathrm{rank_c}(X)$ of X is less than $\mathrm{card}(X)$. In Section 3.2 we will refine this result for the sets X containing only two elements.

In first two sections we will also discuss the maximal number of X-ambiguous bi-infinite words, showing that if X is a finite set of the maximal combinatorial rank then this number is always finite, and moreover, it is at most 2, if X is a binary code.

3.1 The general case

In this section we prove a defect theorem for bi-infinite words. Frequently we illustrate our proofs by pictures. In these pictures a horizontal double line

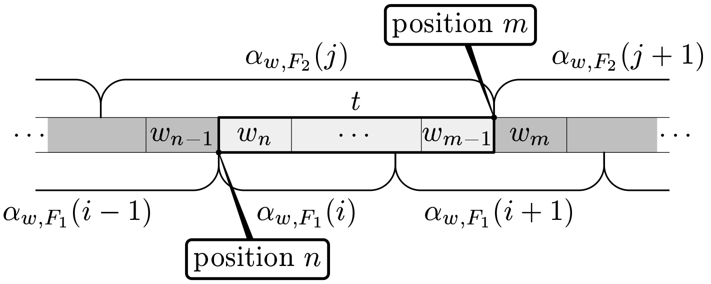


Figure 3.1: An illustration of $(F_1 \triangleright F_2, X)$-difference $t = w_n \dots w_{m-1}$. Note that $F_1(i) = n$ and $F_2(j+1) = m$.

expresses a bi-infinite word with two X-factorizations F_1, F_2. The sequences of words in the factorization F_1 are depicted **below the line** by consecutive arcs, similarly the sequences of words in F_2 are depicted by arcs, which are **above the line**. For example, in Figure 3.1 we consider words $\alpha_{w,F_1}(i), \alpha_{w,F_2}(j) \in X$, such that the words $\alpha_{w,F_1}(i)$ are factors of w defined by the factorization F_1 and the words $\alpha_{w,F_2}(j)$ are factors of w defined by F_2.

Consider a finite non-empty set $X \subseteq \Sigma^+$ and an X-ambiguous bi-infinite word w possessing X-factorizations F_1 and F_2. The set of starting positions of one factorization ($F_1(\mathbb{Z})$ or $F_2(\mathbb{Z})$) factorizes the bi-infinite word w into words of X. The set of starting positions of both factorizations, $F_1(\mathbb{Z}) \cup F_2(\mathbb{Z})$, factorizes w into some other words, which we call *X-differences.*

Formally, for every starting position $n \in F_1(\mathbb{Z})$ find the minimal starting position $m \in F_2(\mathbb{Z})$ such that $n \leq m$. Since F_2 is a growing function we know that such a starting position exists. We call the word

$$t = w_n w_{n+1} \dots w_{m-1}$$

an $(F_1 \triangleright F_2, X)$*-difference*, and we say that there is an *occurrence of* the $(F_1 \triangleright F_2, X)$-difference t in w at the position n, or, shortly, that n is an occurrence of the $(F_1 \triangleright F_2, X)$-difference t. The situation is depicted in Figure 3.1.

Similarly, for every starting position $n \in F_1(\mathbb{Z})$ find the maximal starting position $m \in F_2(\mathbb{Z})$ such that $m \leq n$. We call the word

$$t = w_m w_{m+1} \dots w_{n-1}$$

an $(F_1 \triangleleft F_2, X)$*-difference*, and we say that there is an *occurrence of* the $(F_1 \triangleleft F_2, X)$-difference t in w at the position n.

In Figure 3.2 we can see an example how factorizations F_1 and F_2 factorize the bi-infinite word w into different types of X-differences.

We define the following sets of X-differences:

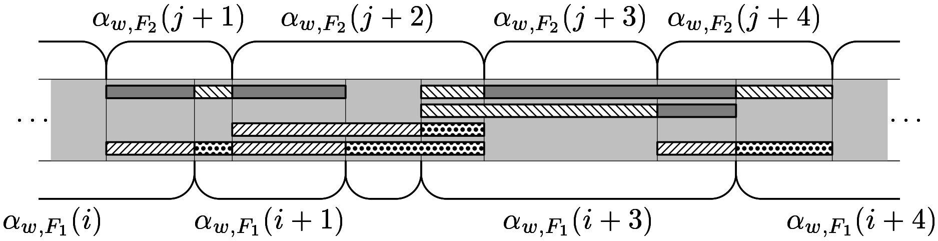


Figure 3.2: An example how factorizations F_1 and F_2 factorize the bi-infinite word w into $(F_1 \triangleright F_2, X)$-differences (dotted rectangles), $(F_1 \triangleleft F_2, X)$-differences (right-oblique hatched rectangles), $(F_2 \triangleright F_1, X)$-differences (gray rectangles) and $(F_2 \triangleleft F_1, X)$-differences (left-oblique hatched rectangles).

- $\mathrm{Diff}_X(w, F_1 \triangleright F_2)$ is the set of all $(F_1 \triangleright F_2, X)$-differences in w;
- $\mathrm{Diff}_X(w, F_1 \triangleleft F_2)$ is the set of all $(F_1 \triangleleft F_2, X)$-differences in w;
- $\mathrm{Diff}_X(w, \{F_1, F_2\}\triangleright) = \mathrm{Diff}_X(w, F_1 \triangleright F_2) \cup \mathrm{Diff}_X(w, F_2 \triangleright F_1)$.

For every position $n \in F_1(\mathbb{Z})$ there is an $(F_1 \triangleright F_2, X)$-difference t and an $(F_1 \triangleleft F_2, X)$-difference t' in w at the position n. Moreover, $tt' \in X$ is a factor of w defined by the factorization F_2. Note also that

$$\mathrm{Diff}_X(w, F_1 \triangleright F_2) \subseteq \mathrm{Pref}(X^+) \cap \mathrm{suff}(X) \quad \text{and}$$
$$\mathrm{Diff}_X(w, F_1 \triangleleft F_2) \subseteq \mathrm{pref}(X) \cap \mathrm{Suff}(X^+)\,.$$

Further, we define the following sets of occurrences:

- $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ is the set of all occurrences of $(F_1 \triangleright F_2, X)$-difference t in w;
- $\mathrm{Occ}_X(w, F_1 \triangleleft F_2, t)$ is the set of all occurrences of $(F_1 \triangleleft F_2, X)$-difference t in w.

Clearly, the sets $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ with $t \in \mathrm{Diff}_X(w, F_1 \triangleright F_2)$ (resp. $\mathrm{Occ}_X(w, F_1 \triangleleft F_2, t)$ with $t \in \mathrm{Diff}_X(w, F_1 \triangleleft F_2)$) form a decomposition of the set of starting positions $F_1(\mathbb{Z})$. Note also that

$$\mathrm{Occ}_X(w, F_1 \triangleright F_2, t) + |t| \subseteq F_2(\mathbb{Z}) \quad \text{and} \quad \mathrm{Occ}_X(w, F_1 \triangleleft F_2, t) - |t| \subseteq F_2(\mathbb{Z}).$$

Observation 3.1. If there is a position $n \in F_1(\mathbb{Z}) \cap F_2(\mathbb{Z})$ (hence, factorizations F_1 and F_2 are not disjoint) then there is an occurrence of the $(F_1 \triangleright F_2, X)$-difference (resp. of the $(F_1 \triangleleft F_2, X)$-difference) 1 at the position n. Obviously, we have

$$\mathrm{Occ}_X(w, F_1 \triangleright F_2, 1) = \mathrm{Occ}_X(w, F_1 \triangleleft F_2, 1) = F_1(\mathbb{Z}) \cap F_2(\mathbb{Z}).$$

Hence, if F_1 and F_2 are disjoint then $1 \notin \mathrm{Diff}_X(w, F_1 \triangleright F_2), \mathrm{Diff}_X(w, F_1 \triangleleft F_2)$.

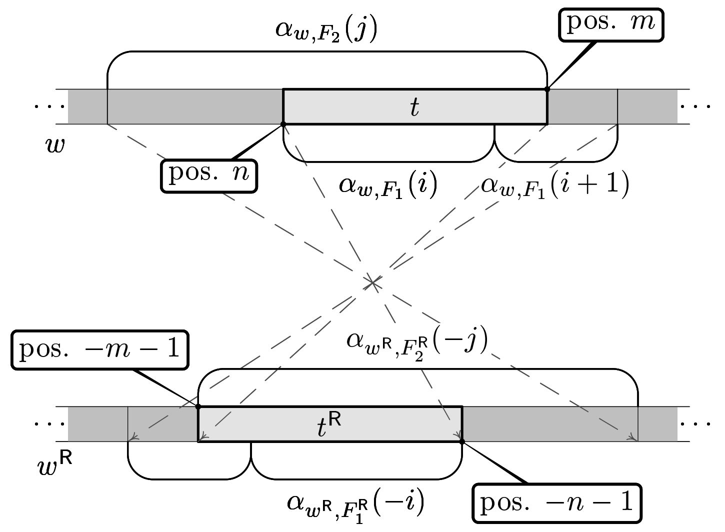


Figure 3.3: An illustration of correspondence between occurrences of the $(F_1 \triangleright F_2, X)$-difference t in w and the $(F_1^{\mathsf{R}} \triangleleft F_2^{\mathsf{R}}, X^{\mathsf{R}})$-difference t^{R} in w^{R}. Note that $m = n + |t|$ and, for example, by (2.2), $\alpha_{w^{\mathsf{R}},F_2^{\mathsf{R}}}(-j) = (\alpha_{w,F_2}(j))^{\mathsf{R}}$.

Observation 3.2. If we take the *mirror image* of bi-infinite word w, then an $(F_1 \triangleright F_2, X)$-differences t becomes the $(F_1^{\mathsf{R}} \triangleleft F_2^{\mathsf{R}}, X^{\mathsf{R}})$-difference t^{R}, *i.e.,*

$$\mathrm{Diff}_{X^{\mathsf{R}}}(w^{\mathsf{R}}, F_1^{\mathsf{R}} \triangleleft F_2^{\mathsf{R}}) = (\mathrm{Diff}_X(w, F_1 \triangleright F_2))^{\mathsf{R}}.$$

It is easy to check, see Figure 3.3, that we have the following equality for the sets of occurrences of the $(F_1 \triangleright F_2, X)$-difference t in w and the $(F_1^{\mathsf{R}} \triangleleft F_2^{\mathsf{R}}, X^{\mathsf{R}})$-difference t^{R} in w^{R}:

$$\mathrm{Occ}_{X^{\mathsf{R}}}(w^{\mathsf{R}}, F_1^{\mathsf{R}} \triangleleft F_2^{\mathsf{R}}, t^{\mathsf{R}}) = -\,\mathrm{Occ}_X(w, F_1 \triangleright F_2, t) - 1\,.$$

Observation 3.3. The following claims about the numbers of X-differences and their occurrences follow immediately:

- since $\mathrm{Diff}_X(w, F_1 \triangleright F_2) \subseteq \mathrm{Pref}(X^+) \cap \mathrm{suff}(X)$, there is only finitely many of $(F_1 \triangleright F_2, X)$-differences;
- there is infinitely many of occurrences of $(F_1 \triangleright F_2, X)$-differences;
- by pigeon hole principle the above claims imply that there is an $(F_1 \triangleright F_2, X)$-difference with infinitely many of its occurrences.

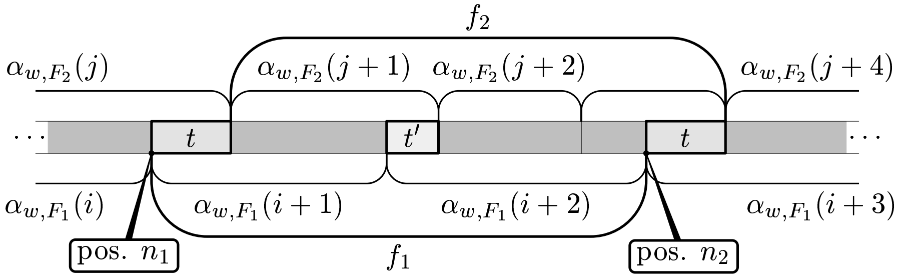


Figure 3.4: An example of a finite t-pair (f_1, f_2), $t \in \mathrm{Diff}_X(w, F_1 \triangleright F_2)$. There is an occurrence of an $(F_1 \triangleright F_2, X)$-difference t' between occurrences n_1 and n_2 of t, but since n_1 and n_2 are consecutive, we have necessarily $t' \neq t$.

Let w be a bi-infinite word possessing X-factorizations F_1 and F_2, and let $t \in \mathrm{Diff}_X(w, F_1 \triangleright F_2)$ be an $(F_1 \triangleright F_2, X)$-difference. We say that two occurrences of t, $n_1, n_2 \in \mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$, with $n_1 < n_2$, are *consecutive* if the is no occurrence $n \in \mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ such that $n_1 < n < n_2$.

Consider two consecutive occurrences $n_1 < n_2$ of an $(F_1 \triangleright F_2, X)$-difference t. Let the factor of the bi-infinite word w between the beginnings of occurrences of t at positions n_1 and n_2 be a word f_1, and similarly, the factor of w between the ends of the occurrences of t at positions n_1 and n_2 be a word f_2. We will call the pair of words (f_1, f_2) a *finite t-pair.* An example of a finite t-pair is in Figure 3.4.

Formally, a *finite t-pair* is a pair of words (f_1, f_2) such that

$$\begin{aligned} f_1 &= w_{n_1} w_{n_1+1} \dots w_{n_2-1}, \\ f_2 &= w_{n_1+|t|} w_{n_1+|t|+1} \dots w_{n_2+|t|-1}, \end{aligned}$$

where $n_1 < n_2$ are consecutive occurrences of an $(F_1 \triangleright F_2, X)$-difference t. Since, $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t) \subseteq F_1(\mathbb{Z})$ (resp. $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t) + |t| \subseteq F_2(\mathbb{Z})$) we have that $f_1 \in X^+$ (resp. $f_2 \in X^+$). Notice also that for any finite t-pair (f_1, f_2) we have that $f_1 t = t f_2$.

Further, assume that for an $(F_1 \triangleright F_2, X)$-difference t, the set of occurrences $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ has a maximum. Let it be n. Then, an *infinite t-pair* is a pair of infinite words (f_1, f_2) such that

$$\begin{aligned} f_1 &= w_n w_{n+1} \dots, \\ f_2 &= w_{n+|t|} w_{n+|t|+1} \dots . \end{aligned}$$

Similarly, we have that $f_1, f_2 \in X^{\mathbb{N}}$ and $f_1 = t f_2$. Note also that for every $t \in \mathrm{Diff}_X(w, F_1 \triangleright F_2)$ there is at most one infinite t-pair.

Finally, a *t-pair* is either a finite, or an infinite t-pair. In the same way one can define t-pairs also for $t \in \mathrm{Diff}_X(w, F_1 \triangleleft F_2)$, see Figure 3.5.

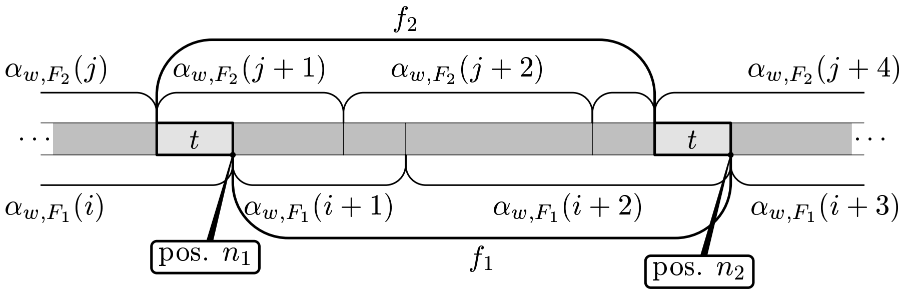


Figure 3.5: An example of a finite t-pair (f_1, f_2), $t \in \mathrm{Diff}_X(w, F_1 \lhd F_2)$.

An essential tool for proving the defect theorem for bi-infinite words is the Graph Lemma (Lemma 2.3). To use the Graph Lemma we have to be sure that all words involved are non-empty, hence we would like to exclude the cases when there are X-differences t with $t = 1$. By Observation 3.1, it is enough to assume that the X-factorizations of an X-ambiguous bi-infinite word are disjoint. First, let us deal with the case when the X-factorizations are not disjoint.

Lemma 3.1. *Let $X \subseteq \Sigma^+$ be a finite non-empty set and w an X-ambiguous bi-infinite word possessing two different joint X-factorizations F_1 and F_2. Then* $\mathrm{rank_c}(X) < \mathrm{card}(X)$.

Proof. The result follows by Lemma 2.3. Indeed, the parts of factorizations F_1 and F_2 to the right (respectively, to the left) from the place where they are joint form an infinite relation

$$\alpha_{w,F_1}(i)\alpha_{w,F_1}(i+1)\cdots = \alpha_{w,F_2}(j)\alpha_{w,F_2}(j+1)\ldots$$

over X (respectively,

$$\alpha_{w,F_1}(i-1)^{\mathsf{R}}\alpha_{w,F_1}(i-2)^{\mathsf{R}}\cdots = \alpha_{w,F_2}(j-1)^{\mathsf{R}}\alpha_{w,F_2}(j-2)^{\mathsf{R}}\ldots$$

over X^{R}). Since the factorizations are different, at least one of these two relations is non-trivial. □

In the case of disjoint X-factorizations we have the following crucial lemma:

Lemma 3.2. *Let $X \subseteq \Sigma^+$ be a finite non-empty set, w an X-ambiguous bi-infinite word possessing two disjoint X-factorization F_1 and F_2, and t an $(F_1 \rhd F_2, X)$-difference (resp. an $(F_1 \lhd F_2, X)$-difference). If there exist two different t-pairs (f_1, f_2) and (f_1', f_2') then* $\mathrm{rank_c}(X) < \mathrm{card}(X)$.

The situation considered in Lemma 3.2, in the case when both t-pairs are finite, is depicted in Figure 3.6.

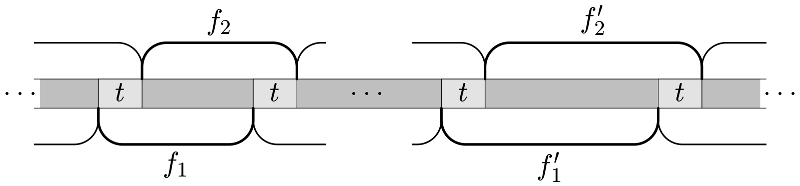


Figure 3.6: An illustration of the situation considered in Lemma 3.2: $t \in \mathrm{Occ}_X(w, F_1 \triangleright F_2)$ and $f_1, f'_1, f_2, f'_2 \in X^+$.

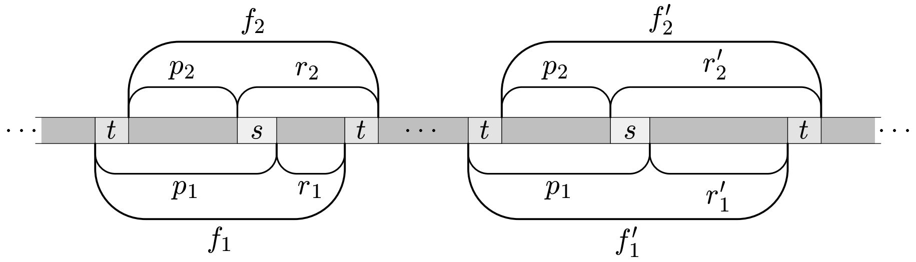


Figure 3.7: An illustration of the situation of the proof in the case $|p_1| > |tp_2|$.

Proof. We will consider only the case when $t \in \mathrm{Diff}_X(w, F_1 \triangleright F_2)$ and both t-pairs (f_1, f_2) and (f'_1, f'_2) are finite. The reader can check that in all other cases the proof is essentially the same (all what is needed is to change the notation in some places). By Observation 3.1, t is non-empty.

Let $p_i \in X^*$, be the longest common prefix of f_i, f'_i over the alphabet X and $r_i, r'_i \in X^*$ their corresponding suffices, *i.e.,* we have that $f_i = p_i r_i$, $f'_i = p_i r'_i$, for $i = 1, 2$. If $|p_1| = |tp_2|$ then the factorizations are not disjoint, a contradiction. Therefore, we will consider only two cases: either $|p_1| > |tp_2|$, or $|p_1| < |tp_2|$.

Case $|p_1| > |tp_2|$. Let $p_1 = tp_2 s$, for some $s \in \Sigma^+$. The situation is depicted in Figure 3.7. Since $|r_2| > |s|$ (resp. $|r'_2| > |s|$), both r_2 and r'_2 must be non-empty. Let words $x, x' \in X$ be the first letters of r_2, r'_2 over the alphabet X, *i.e.,* $r_2 = xq_2$ and $r'_2 = x'q'_2$, for some $q_2, q'_2 \in X^*$. By definition of words r_2 and r'_2, necessarily $x \neq x'$. We have the following three equations over the set $X \cup \{t, s\} \subseteq \Sigma^+$:

$$sr_1t = r_2 = xq_2\,, \qquad sr'_1t = r'_2 = x'q'_2\,, \qquad tp_2s = p_1\,.$$

Since $x \neq x'$, the dependency graph of $X \cup \{t, s\}$ (see Section 2.3) has at least 3 edges which do not form a triangle. Consequently, the number of connected components of the graph is at most $\mathrm{card}(X) - 1$. By Lemma 2.3, we obtain

$$\mathrm{rank_c}(X) \leq \mathrm{rank_c}(X \cup \{t, s\}) \leq \mathrm{card}(X) - 1\,.$$

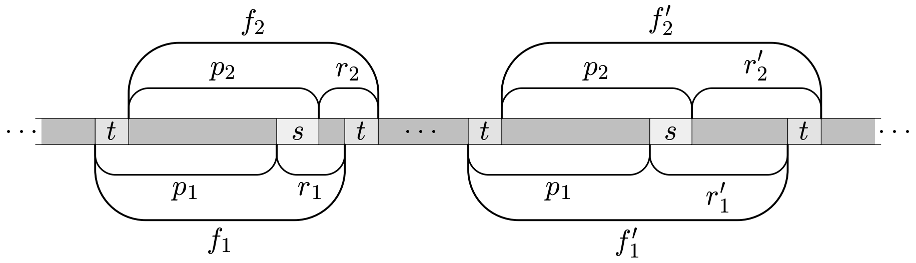


Figure 3.8: An illustration of the situation of the proof in the case $|p_1| < |tp_2|$.

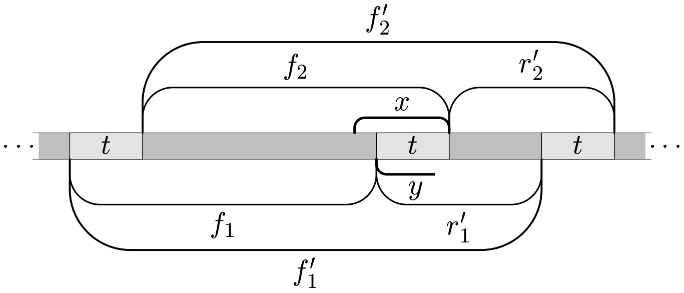


Figure 3.9: An illustration of the situation of the proof in the case when $|p_1| < |tp_2|$ and $r_1 = r_2 = 1$.

Case $|p_1| < |tp_2|$. Let $p_1 s = tp_2$, for some $s \in \Sigma^+$. The situation is depicted in Figure 3.8. If both r_1 and r_1' are non-empty, we obtain, as in the previous case, the following equations over $X \cup \{t, s\} \subseteq \Sigma^+$:

$$sr_2 = r_1 t = xq_1 t\,, \qquad sr_2' = r_1' t = x' q_1' t\,, \qquad tp_2 = p_1 s\,,$$

where $r_1 = xq_1$ and $r_1' = x'q_1'$ with $x, y \in X$ and $q_1, q_1' \in X^*$. By Lemma 2.3, we obtain a defect effect:

$$\mathrm{rank_c}(X) \leq \mathrm{rank_c}(X \cup \{t, s\}) \leq \mathrm{card}(X) - 1\,.$$

Hence, assume that, for instance, $r_1 = 1$, implying $t = sr_2$. This contradicts the definition of $(F_1 \triangleright F_2, X)$-differences, unless also $r_2 = 1$. We have that $p_1 = f_1$, $p_2 = f_2$ and $s = t$, see Figure 3.9. Let $x \in X$ be the last letter over the alphabet X of f_2 and $y \in X$ the first letter over X of r_1' (note that $r_1' \neq 1$, otherwise the t-pairs (f_1, f_2) and (f_1', f_2') are the same). Since (f_1, f_2) is a t-pair and t is an $(F_1 \triangleright F_2, X)$-difference, we have that $|t| < |x|$. Hence, there is an occurrence of t at the starting position of y, *cf.* Figure 3.9, which is a contradiction with the fact that (f_1', f_2') is a t-pair.

□

Remark 3.1. Note that the inequality

$$\mathrm{rank}_{\mathrm{c}}(X) \leq \mathrm{rank}_{\mathrm{c}}(X \cup \{t, s\}),$$

used in the proof above, does not hold in general for other types of rank. Indeed, consider any prefix code X over the alphabet $\{a, b\}$. If we take $t = a$ and $s = b$ then both the free and the prefix ranks of the set $X \cup \{t, s\}$ are equal to 2, while the free and prefix ranks of the set X are equal to $\mathrm{card}(X)$.

As a consequence of the observations and the lemmas above we obtain the defect theorem for bi-infinite words.

Theorem 3.3. *Let $X \subseteq \Sigma^+$ be a finite non-empty set and w an X-ambiguous bi-infinite word possessing two different X-factorizations F_1 and F_2. The combinatorial rank of X is less than $\mathrm{card}(X)$, or both the word w and the X-factorizations F_1 and F_2 are periodic. Moreover, if the combinatorial rank of X equals to $\mathrm{card}(X)$ then the number of X-ambiguous bi-infinite words is at most $\mathrm{size}(X) - \mathrm{card}(X)$, in particular, it is finite.*

Proof. If F_1 and F_2 are not disjoint the result follows by Lemma 3.1. Let t be any of the $(F_1 \triangleright F_2, X)$-differences with infinitely many occurrences (there is at least one, *cf.* Observation 3.3). This means that the set $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ is not bounded.

To prove that the bi-infinite word w is periodic we need to divide it completely into t-pairs. This can be done, *cf.* the definition of t-pairs, only if the set $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ is not bounded from the left, *i.e.,* it does not have a minimum. Without lost of generality we can assume that. Indeed, if the set $\mathrm{Occ}_X(w, F_1 \triangleright F_2, t)$ has a minimum then, since it is not bounded, it is not bounded from the right. In such a case the set $\mathrm{Occ}_{X^{\mathsf{R}}}(w^{\mathsf{R}}, F_1^{\mathsf{R}} \triangleleft F_2^{\mathsf{R}}, t^{\mathsf{R}})$ is not bounded from the left, hence we can consider the mirror image of w instead.

Therefore, assume that the bi-infinite word w is entirely divided into (finite and possibly one infinite) t-pairs. If any two of these t-pairs are not the same then, by Lemma 3.2, we have that $\mathrm{rank}_{\mathrm{c}}(X) < \mathrm{card}(X)$.

Now, assume that $\mathrm{rank}_{\mathrm{c}}(X) = \mathrm{card}(X)$. By Lemma 3.2, there is a unique t-pair (f_1^t, f_2^t), and therefore, the bi-infinite word, as well as both X-factorizations F_1 and F_2, are periodic with

$$w = (f_1^t)^{\mathbb{Z}} = (f_2^t)^{\mathbb{Z}}.$$

Further, since every X-ambiguous bi-infinite word w is periodic, every $(F_1 \triangleright F_2, X)$-difference t (resp. every $(F_2 \triangleright F_1, X)$-difference t) has infinitely many occurrences in w, *i.e.,* as above, there is the unique t-pair, which uniquely specifies the whole bi-infinite word w. Therefore, any two different X-ambiguous bi-infinite words w and w' do not contain any common X-difference:

$$\mathrm{Diff}_X(w, \{F_1, F_2\}_{\triangleright}) \cap \mathrm{Diff}_X(w', \{F_1, F_2\}_{\triangleright}) = \emptyset\,.$$

By Observation 3.3, $\mathrm{Diff}_X(w, \{F_1, F_2\}\triangleright) \subseteq \mathrm{Pref}(X^+) \cap \mathrm{suff}(X)$, hence there is at most

$$\mathrm{card}(\mathrm{Pref}(X^+) \cap \mathrm{suff}(X)) \leq \mathrm{card}(\mathrm{suff}(X)) \leq \mathrm{size}(X) - \mathrm{card}(X)$$

X-ambiguous bi-infinite words. □

We will need the following definition to state a corollary of Theorem 3.3 which we will use later. Consider a periodic bi-infinite word w and a set $X \subseteq \Sigma^+$ such that

$$w = (x_1 x_2 \dots x_k)^{\mathbb{Z}}$$

with $x_1, \dots, x_k \in X$. Let $[x_1 x_2 \dots x_k]_w^{\mathbb{Z}}$ denote the set of all X-factorizations F of w such that the sequence of factors of w by F is

$$\dots, x_1, x_2, \dots, x_k, x_1, x_2, \dots, x_k, \dots .$$

Obviously, the factorizations in $[x_1 x_2 \dots x_k]_w^{\mathbb{Z}}$ are periodic and pairwise shift-equivalent.

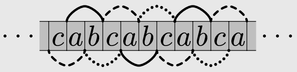
Corollary 3.4. *Let $X \subseteq \Sigma^+$ be a finite non-empty set such that* $\mathrm{rank_c}(X) = \mathrm{card}(X)$ *and w an X-ambiguous bi-infinite word possessing two different X-factorizations F_1 and F_2. For every t-pair (f_1, f_2) we have*

$$w = f_1^{\mathbb{Z}} = f_2^{\mathbb{Z}}, \qquad F_1 \in [f_1]_w^{\mathbb{Z}} \qquad \textit{and} \qquad F_2 \in [f_2]_w^{\mathbb{Z}} .$$

Theorem 3.3 deserves a few comments.

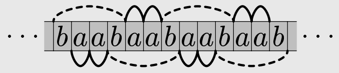
First, the possibility that the two factorizations are both periodic cannot be ruled out, as the following example shows:

Example 3.2. Let $X = \{ab, bc, ca\}$. Then we have $\mathrm{rank_c}(X) = \mathrm{rank_f}(X) = 3 = \mathrm{card}(X)$. Note also that the bi-infinite word $(abc)^{\mathbb{Z}}$ has two disjoint, but shift-equivalent, X-factorizations:



Second, as the following example shows, the combinatorial rank cannot be replaced by the free rank, for instance. This latter remark is quite interesting since in all previous defect theorems, see [ChK], either of our notions of the rank, or even some others, can be used to witness the defect effect.

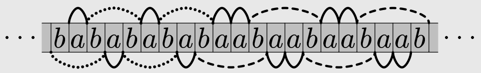
Example 3.3. Let $X = \{a, bab, baab\}$. The word $(baa)^{\mathbb{Z}}$ has two different X-factorizations, namely the ones depicted as



They are clearly shift-equivalent. On the other hand the word

$$w = \dots bababaabaab \cdots = {}^{\mathbb{N}}(ba)b(aab)^{\mathbb{N}}$$

also has two different X-factorizations, which, however, are not shift-equivalent



In the both cases above the two factorizations are disjoint. Clearly, $\mathrm{rank_c}(X) = 2$, since $X \subseteq \{a,b\}^+$, but for no word s the inclusion $X \subseteq s^+$ holds. On the other hand, since X is a prefix code we conclude that $\mathrm{rank_p}(X) = \mathrm{rank_f}(X) = 3$.

Finally, the upper bound "size(X) − card(X)" for the number of X-ambiguous bi-infinite words in the case that $\mathrm{rank_c}(X) = \mathrm{card}(X)$ can be, most likely, essentially improved. In fact, we conjecture that the upper bound is "card(X)". In the next section we will show that this conjecture is true if the cardinality of the set X is 2. The following example shows that we cannot expect a better upper bound.

Example 3.4. For arbitrary integer $n \geq 1$ let $X = \{a_1a_1, a_2a_2, \dots, a_na_n\}$ be a set of words over the alphabet $\{a_1, \dots, a_n\}$. Clearly, $\mathrm{rank_c}(X) = \mathrm{card}(X)$ and each of the periodic bi-infinite words $a_1^{\mathbb{Z}}, \dots, a_n^{\mathbb{Z}}$ has two disjoint shift-equivalent factorizations, hence $\mathrm{card}(\mathrm{Amb}(X)) \geq n = \mathrm{card}(X)$.

The next example shows that the estimation (*cf.* the end of proof of Theorem 3.3)

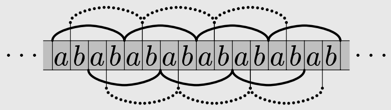
$$\mathrm{card}(\mathrm{Amb}(X)) \leq \mathrm{card}(\mathrm{Pref}(X^+) \cap \mathrm{suff}(X))$$

is not suitable for improving the upper bound.

Example 3.5. Let $n \geq 1$ be an integer and let $X = \{(ab)^n, (ba)^n\}$. We have

$$\mathrm{card}(\mathrm{pref}(X) \cap \mathrm{suff}(X)) = 4n - 2 = \mathrm{size}(X) - \mathrm{card}(X)\,,$$

with the set X being a prefix code, and so $\mathrm{rank_c}(X) = \mathrm{rank_p}(X) = \mathrm{rank_f}(X) = \mathrm{card}(X)$. However, similarly as in Example 2.4, there exists only one X-ambiguous bi-infinite word $w = (ab)^{\mathbb{Z}}$. In fact, it possesses $2n$ disjoint X-factorizations. For instance, if $n = 2$, then the bi-infinite word $w = (ab)^{\mathbb{Z}}$ has the following 4 factorizations:



Let us number the above factorizations $F_1, \ldots, F_4$ from up to down. Then

$$\begin{aligned}
\mathrm{Diff}_X(w, \{F_1, F_2\}\triangleright) &= \{a, bab\},\\
\mathrm{Diff}_X(w, \{F_1, F_3\}\triangleright) &= \{b, aba\},\\
\mathrm{Diff}_X(w, \{F_2, F_3\}\triangleright) &= \{ab\}, \text{ and}\\
\mathrm{Diff}_X(w, \{F_1, F_4\}\triangleright) &= \{ba\}.
\end{aligned}$$

Hence, the sets of X-differences in w are disjoint, although they relate to the same bi-infinite word w. This property is not considered in the proof of Theorem 3.3.

However, we can refine the analysis of the last part of the proof of Theorem 3.3 to get a slightly better bound.

Theorem 3.5. *Let $X \subseteq \Sigma^+$ be a finite non-empty set. If* $\mathrm{rank_c}(X) = \mathrm{card}(X)$ *then the number of X-ambiguous bi-infinite words is at most* $\frac{1}{2}\,\mathrm{size}(X)$.

Proof. The main idea of the proof is based on the fact that the sets of X-differences $\mathrm{Diff}_X(w, \{F_1, F_2\}\triangleright)$ are singletons only in a special case, which cannot appear more often than $\mathrm{card}(X)$ times.

Indeed, let k be the number of X-ambiguous bi-infinite words for which the set $\mathrm{Diff}_X(w, \{F_1, F_2\}\triangleright)$ is a singleton (here F_1 and F_2 are any two different X-factorizations of the bi-infinite word w). Then

$$\begin{aligned}
\mathrm{card}(\mathrm{Amb}(X)) &\le k + \frac{\mathrm{card}(\mathrm{Pref}(X^+) \cap \mathrm{suff}(X)) - k}{2}\\
&\le \frac{k}{2} + \frac{1}{2}(\mathrm{size}(X) - \mathrm{card}(X))\,.
\end{aligned}$$

Now, if we prove that $k \le \mathrm{card}(X)$, we have that $\mathrm{card}(\mathrm{Amb}(X)) \le \frac{1}{2}\,\mathrm{size}(X)$, and we are done.

Let us consider an X-ambiguous bi-infinite word w possessing X-factorizations F_1 and F_2 such that $\mathrm{Diff}_X(w, \{F_1, F_2\}\triangleright) = \{t\}$ is a singleton. Assume that there is an occurrence of $(F_1 \triangleright F_2, X)$-difference t in w at a position n. Let i, j be integers such that $F_1(i) = n$ and $F_2(j) = n + |t|$. If $|\alpha_{w,F_1}(i)| \le |t|$, see Figure 3.10, then there is an occurrence of an $(F_1 \triangleright F_2, X)$-difference s at the position $m = n + |\alpha_{w,F_1}(i)|$. But this is impossible since s is a proper suffix of t and $\mathrm{Diff}_X(w, F_1 \triangleright F_2)$ is a singleton.

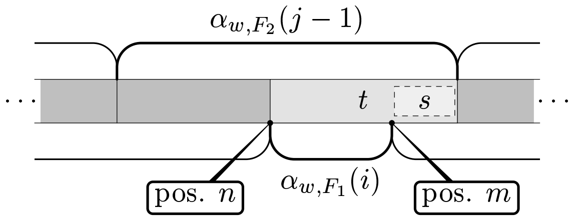


Figure 3.10: The case $|\alpha_{w,F_1}(i)| \leq |t|$ of the proof of Theorem 3.5.

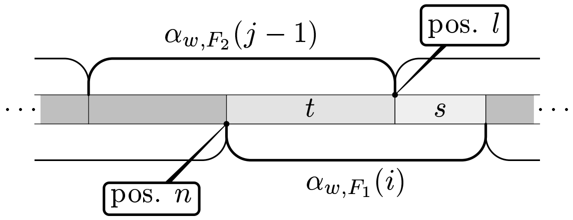


Figure 3.11: The case $|\alpha_{w,F_1}(i)| > |t|$ of the proof of Theorem 3.5.

Hence, assume that $|\alpha_{w,F_1}(i)| > |t|$. Then, the word s defined by the equation $\alpha_{w,F_1}(i) = ts$ is an $(F_2 \triangleright F_1, X)$-difference with an occurrence in w at the position $l = n + |t|$, see Figure 3.11. Therefore, we have that $s = t$ and

$$tt = \alpha_{w,F_1}(i) \in X\ .$$

Clearly, there is at most $\operatorname{card}(X)$ words t which satisfy the condition $tt \in X$, and hence also at most $\operatorname{card}(X)$ sets $\operatorname{Diff}_X(w, \{F_1, F_2\}\triangleright)$ which are singletons.

□

3.2 The two-element case

The restriction of Theorem 3.3 to two-element sets yields the following consequence.

Corollary 3.6. *Consider a binary set $X = \{\alpha, \beta\} \subseteq \Sigma^+$. Let w be an X-ambiguous bi-infinite word possessing two different X-factorizations F_1 and F_2. Then either the words α and β commute, or the bi-infinite word w and both the X-factorizations F_1 and F_2 are periodic.*

In this section we will refine the result above: we will characterize all binary sets X allowing the existence of an X-ambiguous bi-infinite word. We will also prove, as conjectured in general in Section 3.1, that for any binary code X there are at most $\mathrm{card}(X) = 2$ X-ambiguous bi-infinite words, and at most one X-ambiguous bi-infinite word such that its X-factorizations contain together both elements of X. Actually, in the second case, it can happen that the both X-factorizations consist of single elements of X, but then, necessarily, these two elements are different.

3.2.1 The defect theorem

Let X be a binary set containing non-empty words and let w be an X-ambiguous bi-infinite word. We will distinguish two cases. Either, the two factorizations of the X-ambiguous bi-infinite word w consist of only one element of X, or they they contain together both elements of X. In the first case, the situation is obvious: As an immediate consequence of Lemma 2.4 we have.

Claim 3.7. *Consider a unary set $Y = \{\alpha\}$ with $\alpha \in \Sigma^+$. Let w be an Y-ambiguous bi-infinite word possessing two different Y-factorizations F_1 and F_2. Then α is not primitive and the factorizations $F_1, F_2 \in [\alpha]_w^{\mathbb{Z}}$ are shift-equivalent.*

The other case is more interesting. In this case we will say that an X-ambiguous bi-infinite word is *proper*. Let us start with an auxiliary lemma and its two corollaries.

Lemma 3.8. *Let p and q be non-empty words such that pq is primitive and let $n \geq 1$ be an integer. If the pair (pq, qp) matches the word $p(qp)^n$ at a position (u, v) then one of the following conditions holds:*

(i) $u = p$ and $v = (qp)^n$;

(ii) $n = 1$ and there are a primitive word s and integers $i, j \geq 1$ such that

$$u = ps^i, \qquad v = s^j p \qquad \text{and} \qquad q = s^{i+j}\ ;$$

(iii) $u = (pq)^n$ and $v = p$.

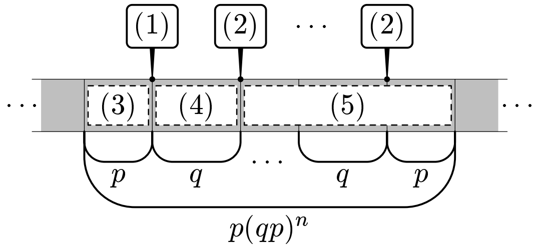


Figure 3.12: The all possibilities where the pair (pq, qp) can match the word $p(qp)^n$.

Similarly, if the pair (qp, pq) matches the word $p(qp)^n$ at a position (u, v) then one of the following conditions holds:

(i) $u = 1$ *and* $v = p(qp)^n$*;*

(ii) $n = 1$ *and* $|p| < |u| < |pq|$*;*

(iii) $u = p(qp)^n$ *and* $v = 1$*.*

Proof. We will prove only the first part of the claim, since the second one can be proved in the same way. Assume that the pair (pq, qp) matches the word $p(qp)^n$ at a position (u, v). If (pq, qp) matches the beginning (*i.e.,* $u = 1$) or the end of $p(qp)^n$ (*i.e.,* $v = 1$) then, clearly, $pq = qp$, a contradiction with the primitiveness of pq. Otherwise, there are 5 possibilities, see Figure 3.12:

(1) $u = p$ and $v = (qp)^n$;

(2) $u = (pq)^k$ and $v = p(qp)^{n-k}$ for an integer $1 \leq k \leq n$;

(3) the pair matches the first p, *i.e.,* u is a proper non-empty prefix of p;

(4) the pair matches the first q, *i.e.,* $u = pt_1$ where t_1 is a proper non-empty prefix of q;

(5) the pair matches anywhere after the first pq, but not at the end of any $(pq)^k$, with $1 \leq k \leq n$, *i.e.,* $|u| > |pq|$ and $u \notin (pq)^+$.

Let us analyze all the cases. The case (1) is the case *(i)* of the lemma. In the case (2), if $k < n$ then qp and $p(qp)^{n-k}$ are left comparable, and thus $pq = qp$, a contradiction. In the case (2) with $k = n$ we have the case *(iii)* of the lemma. In the case (3) the word qp of the pair (pq, qp) is an inner factor of the prefix pqp of $p(qp)^n$. Hence, by Lemma 2.4, we have that qp is non-primitive. By Corollary 2.9, this is a contradiction with the primitiveness of pq. Similarly, in

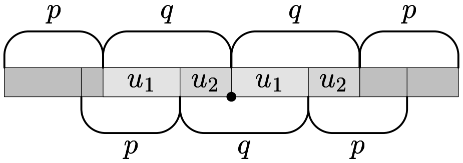


Figure 3.13: The situation in the case (4) with $n = 1$.

the case (5) the word pq of the pair (pq, qp) is an inner factor of a factor pqp of $p(qp)^n$, a contradiction.

Finally, in the case (4), if $n \geq 2$, the word qp of the pair (pq, qp) is an inner factor of the factor qpq of $p(qp)^n$ starting after the first p, again a contradiction. Hence, we have that $n = 1$. The situation is depicted in Figure 3.13. It follows, by Lemma 2.4, that q is non-primitive, and that there are a primitive word s and integers $i, j \geq 1$ such that $u_1 = s^j$, $u_2 = s^i$ and $q = s^{i+j}$. This is evidently, the case *(ii)* of the lemma. □

Lemma 3.8 has two straightforward corollaries.

Corollary 3.9. *Let p and q be non-empty words such that pq is primitive and $|p| = |q|$ and let $n \geq 1$ be an integer. Then the pair (pq, qp) does not match the word $p(qp)^n$ at any position.*

Proof. Assume that (pq, qp) matches $p(qp)^n$ at a position (u, v). By Lemma 3.8, we have to consider 3 cases.

Case *(i)*. We have that $u = p$ and pq are right comparable. Since $|p| = |q|$, this implies $p = q$, a contradiction with the primitiveness of pq.

Case *(ii)*. Since $u = ps^i$ and $pq = ps^{i+j}$ are right comparable, we have that also p and ps^j are right comparable. Since $|s^j| < |q| = |p|$, s^j is a suffix of p. Similarly, we have that s^i is a prefix of p. By the length argument, $p = s^i s^j = q$, a contradiction.

Case *(iii)*. The same as the case *(i)*. □

Corollary 3.10. *Let p and q be non-empty words such that pq is primitive and let $n \geq 1$ be an integer. Then $p(qp)^n$ and $q(pq)^n$ are not conjugates.*

Proof. Assume that $p(qp)^n$ and $q(pq)^n$ are conjugates. Then necessarily

- $|p| = |q|$; and
- $p(qp)^n$ is a factor of $q(pq)^n q(pq)^n$, and hence, the pair (pq, qp) matches the word $p(qp)^n$.

By Corollary 3.9, this is a contradiction. □

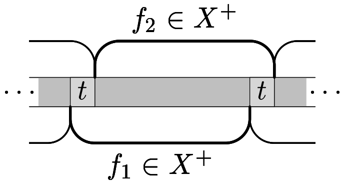


Figure 3.14: An illustration of a finite t-pair (f_1, f_2). In the picture $f_1 \in X^+$ (resp. $f_2 \in X^+$) is a product of factors of w by the factorization F_1 (resp. by the factorization F_2).

Now we are ready to prove the main result of this section.

Theorem 3.11. *Consider a binary set $X = \{\alpha, \beta\}$ with $\alpha, \beta \in \Sigma^+$. Let w be a proper X-ambiguous bi-infinite word possessing two different X-factorizations F_1 and F_2. Then at least one of the following conditions is satisfied:*

(i) α and β commute; or

(ii) the primitive roots of α and β are conjugates, $w = \alpha^{\mathbb{Z}} = \beta^{\mathbb{Z}}$, and $F_1 \in [\alpha]_w^{\mathbb{Z}}$ and $F_2 \in [\beta]_w^{\mathbb{Z}}$, or vice versa; or

(iii) there exists an integer $n \geq 1$ such that either $w = (\alpha\beta^n)^{\mathbb{Z}}$ and $F_1, F_2 \in [\alpha\beta^n]_w^{\mathbb{Z}}$ with α primitive, or $w = (\beta\alpha^n)^{\mathbb{Z}}$ and $F_1, F_2 \in [\beta\alpha^n]_w^{\mathbb{Z}}$ with β primitive.

Observation 3.4. Note that in the case *(iii)*, the factorizations F_1 and F_2 are necessarily shift-equivalent. Hence, by Lemma 2.4, in the case $w = (\alpha\beta^n)^{\mathbb{Z}}$ (resp. $w = (\beta\alpha^n)^{\mathbb{Z}}$), $\alpha\beta^n$ (resp. $\beta\alpha^n$) is non-primitive.

Proof. We can assume that α and β do not commute, otherwise we are in the case *(i)*. In particular, by Lemma 3.1, the X-factorizations F_1 and F_2 are necessarily disjoint. Our goal is to find a finite t-pair (f_1, f_2), see Figure 3.14. Then, by Corollary 3.4, we have an explicit characterization of the bi-infinite word w and its X-factorizations F_1 and F_2.

It is enough to prove the theorem in the case when α and β are both primitive. Indeed, in all other cases we replace the set X by the set $\bar{X} = \{\rho(\alpha), \rho(\beta)\}$, and the X-factorizations F_1 and F_2 by $\bar{X}$-factorizations $\bar{F}_1$ and $\bar{F}_2$ defined in a natural way such that $F_1(\mathbb{Z}) \subseteq \bar{F}_1(\mathbb{Z})$ and $F_2(\mathbb{Z}) \subseteq \bar{F}_2(\mathbb{Z})$, and the claim will follow. If $\bar{X}$-factorizations $\bar{F}_1$ and $\bar{F}_2$ will become joint, we have the case *(i)*. In the case *(iii)* with $w = (\rho(\alpha)\rho(\beta)^n)^{\mathbb{Z}}$ we have necessarily that $\rho(\alpha) = \alpha$, *i.e.*, α is primitive. Similarly, in the case *(iii)* with $w = (\rho(\beta)\rho(\alpha)^n)^{\mathbb{Z}}$, β is primitive.

Consider the factors of w defined by the X-factorization F_1. We have 3 possibilities: either $\mathrm{Fact}(w, F_1) = \{\alpha\}$, or $\mathrm{Fact}(w, F_1) = \{\beta\}$, or $\mathrm{Fact}(w, F_1) = \{\alpha, \beta\}$. In the first case, if $\alpha \in \mathrm{Fact}(w, F_2)$, see Figure 3.15, we have a contradiction with the primitiveness of α, by Lemma 2.4. Therefore, $\mathrm{Fact}(w, F_2) = \{\beta\}$,

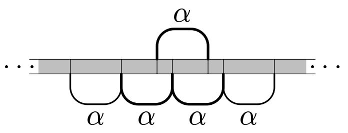


Figure 3.15: An illustration of the situation when $F_1 \in [\alpha]_w^{\mathbb{Z}}$ and F_2 produces α as a factor of w.

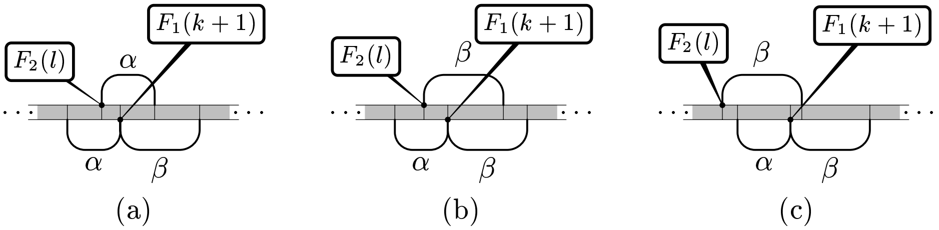


Figure 3.16: All possible coverings of the position $F_1(k+1)$ by a factor $\alpha_{w,F_2}(l)$.

and so, $F_1 \in [\alpha]_w^{\mathbb{Z}}$ and $F_2 \in [\beta]_w^{\mathbb{Z}}$. By Lemma 2.8, α and β are conjugates: the case *(ii)*. A similar argument can be applied in the second case.

Finally, consider the case $\text{Fact}(w, F_1) = \{\alpha, \beta\}$. Since, by Corollary 3.6, F_1 is periodic, there is an integer $k \in \mathbb{Z}$ such that $\alpha_{w,F_1}(k) = \alpha$ and $\alpha_{w,F_1}(k+1) = \beta$. Without loss of generality we can assume that $|\alpha| \leq |\beta|$. In Figure 3.16 we can see all the possibilities, (a), (b) and (c), how a factor $\alpha_{w,F_2}(l)$ of w defined by F_2, for some integer $l \in \mathbb{Z}$, covers the position $F_1(k+1)$, *i.e.,* $F_2(l) < F_1(k+1) < F_2(l+1)$.

Case (a). If we forget about the relation between the lengths of α and β, this case is symmetric to the case (b). Hence, it is enough to prove the result in the case (b) without using the assumption about the lengths.

Case (b). We have that $\alpha_{w,F_2}(l) = \beta$. If the factor $\alpha_{w,F_2}(l+1)$ is also β then,

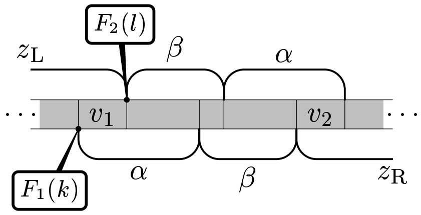


Figure 3.17: The situation in the case (b). Note that $z_{\mathrm{R}} = \alpha_{w,F_1}(k+2)$ and $z_{\mathrm{L}} = \alpha_{w,F_2}(l-1)$.

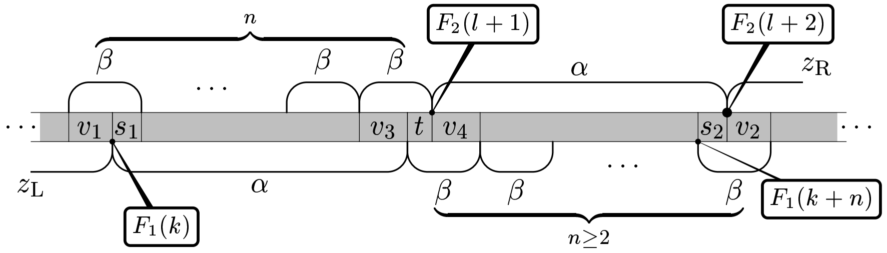


Figure 3.18: The situation in the case (b) when the sequences of β's exceed α's on both sides.

by Lemma 2.4, β is not primitive, a contradiction. Hence, we have the situation shown in Figure 3.17. Now, if $z_R = \alpha$ or $z_L = \alpha$ then $v_1 = v_2$ and we have a v_1-pair $(\alpha\beta, \beta\alpha)$. By Corollary 3.4, we are in the case *(iii)*. Thus assume that $z_R = z_L = \beta$.

We can repeat the above consideration taking $\alpha_{w,F_1}(k+3)$ as z_R and $\alpha_{w,F_2}(l-2)$ as z_L. Again, if z_R or z_L is equal to α, we arrive into the case *(iii)* with $w = (\alpha\beta^2)^{\mathbb{Z}}$. Otherwise, we will continue the same process until, either we find a t-pair $(\alpha\beta^n, \beta^n\alpha)$ for some integer $n \geq 1$, or the sequences of β's exceed α's, *i.e.,* $\alpha_{w,F_1}(k)$ on the left side and $\alpha_{w,F_2}(l+1)$ on the right side. Note that, by the length argument, this will happen on both sides at the same time. The situation in the later case is depicted in Figure 3.18.

Now again, if $z_L = \beta$ or $z_R = \beta$ then, since $|v_1| = |v_2|$, we have that $v_1 = v_2$, and hence we are again in the case *(iii)* with $w = (\alpha\beta^n)^{\mathbb{Z}}$. Thus assume that $z_L = z_R = \alpha$. We have $\beta = v_3 t = t v_4$, which, by Lemma 2.10, yields

$$v_3 = (pq)^{m_1}, \quad v_4 = (qp)^{m_1}, \quad t = p(qp)^{m_2}, \quad \text{and} \quad \beta = p(qp)^{m_1+m_2},$$

for some words p and q such that pq is primitive and some integers $m_1 \geq 1$ and $m_2 \geq 0$. We have the following two equations

$$\begin{aligned} \alpha_{w,F_1}(k) = \alpha = s_1\beta^{n-2}v_3 = s_1[p(qp)^{m_1+m_2}]^{n-2}(pq)^{m_1} \quad \text{and} \\ \alpha_{w,F_2}(l+1) = \alpha = v_4\beta^{n-2}s_2 = (qp)^{m_1}[p(qp)^{m_1+m_2}]^{n-2}s_2\,. \end{aligned} \tag{3.1}$$

We observe that α ends with pq and starts with qp. This means that the pair of words (pq, qp) matches the word $\alpha_{w,F_1}(k+n) = \beta = p(qp)^{m_1+m_2}$ at the position $F_2(l+2)$, *cf.* the bigger black point in Figure 3.18. By Lemma 3.8, we have 3 possibilities:

Case $s_2 = p$. Note that, by the length argument, $|s_1| = |s_2|$. Thus, since s_1 is a suffix of $\beta = p(qp)^{m_1+m_2}$, we have that $s_1 = p$. Equations (3.1) yield that p and q commute, a contradiction with the primitiveness of pq.

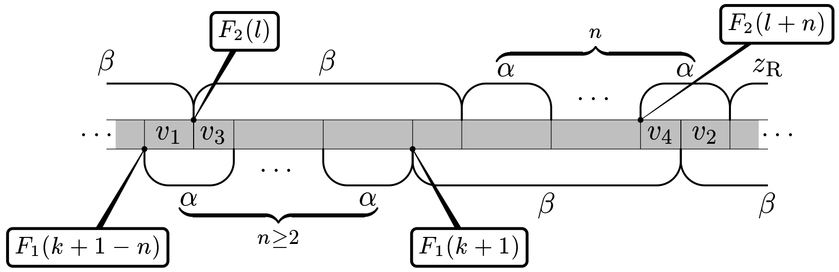


Figure 3.19: The situation in the case (c).

Case $s_2 = ps^i$ and $q = s^{i+j}$ for some word s and integers $i, j \geq 1$. As above, we have that $s_1 = s^ip$. Since $i < i + j$, Equations (3.1) yield a non-trivial equation over the set $\{p, s\}$, again a contradiction.

Case $s_2 = (pq)^{m_1+m_2}$ and $v_2 = p$. We have that $s_1 = (qp)^{m_1+m_2}$. To avoid the contradiction above, we have to assume that $m_2 = 0$. Then, Equations (3.1) will become identical. Note that we have $\alpha, \beta \in \{p, q\}^+$. Hence the following infinite equation, again *cf.* Figure 3.18,

$$v_2\alpha_{w,F_1}(k+n+1)\alpha_{w,F_1}(k+n+2)\cdots = \alpha\alpha_{w,F_2}(l+3)\alpha_{w,F_2}(l+4)\ldots \quad (3.2)$$

is an equation over the set $\{p, q\}$. Since $v_2 = p$ and, by Equations (3.1), $\alpha = (qp)^{m_1}[p(qp)^{m_1}]^{n-2}(pq)^{m_1}$, Equation (3.2) is non-trivial. So again, we have a contradiction.

We concluded that the case when $z_{\mathrm{R}} = \alpha$ cannot happen.

Case (c). Using the same considerations, as in the previous case, either we find a t-pair $(\alpha^n\beta, \beta\alpha^n)$: the case *(iii)* with $w = (\beta\alpha^n)^{\mathbb{Z}}$, for some integer $n \geq 1$, or we come to the situation depicted in Figure 3.19. In the second case we can immediately write the following equations

$$v_1v_3 = \alpha_{w,F_1}(k+1-n) = \alpha = \alpha_{w,F_2}(l+n) = v_4v_2\,, \quad \text{and} \quad (3.3)$$

$$\begin{aligned} v_3\alpha^{n-1}\beta &= v_3\alpha_{w,F_1}(k+2-n)\ldots\alpha_{w,F_1}(k+1) \\ &= \alpha_{w,F_2}(l)\ldots\alpha_{w,F_2}(l+n-1)v_4 = \beta\alpha^{n-1}v_4\,, \end{aligned} \quad (3.4)$$

for some words v_1, v_2, v_3 and v_4 such that $|v_1| = |v_2|$ and $|v_3| = |v_4|$, *cf.* Figure 3.19. We will distinguish two possibilities:

Case $z_{\mathrm{R}} = \alpha$. As it is shown in Figure 3.19, v_3v_1 is a prefix of $\alpha_{w,F_2}(l) = \beta = \alpha_{w,F_1}(k+2)$. Hence, the pair (v_3, v_1) matches the word $\alpha_{w,F_2}(l+n)\alpha_{w,F_2}(l+n+1) = \alpha\alpha = v_1v_3v_1v_3$. Since $\alpha = v_1v_3$ is primitive, the match must be at the position (v_1v_3, v_1v_3). Then, necessarily, $v_2 = v_3$, say equal to p, and $v_1 = v_4$,

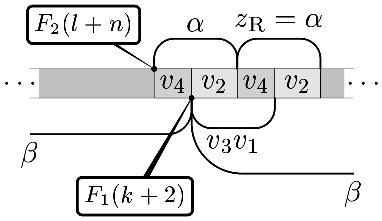


Figure 3.20: The situation in the case (c) with $z_R = \alpha$.

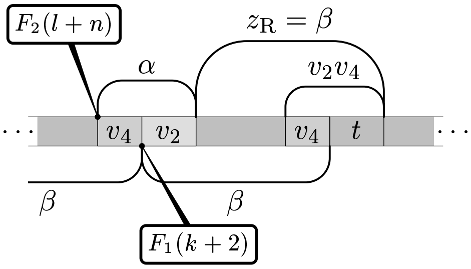


Figure 3.21: The situation in the case (c) with $z_R = \beta$.

say equal to q, *cf.* Figure 3.20. Moreover, we have that $|p| = |v_3| = |v_4| = |q|$. Consequently, Equation (3.4) implies that the words $v_3\alpha^{n-1} = p(qp)^{n-1}$ and $\alpha^{n-1}v_4 = q(pq)^{n-1}$ are conjugates. By Corollary 3.10, this is a contradiction.

Case $z_R = \beta$. The situation is depicted in Figure 3.21. By the length argument, it is obvious that $|t| = |v_2| = |v_1|$. Hence, since v_1 is a suffix of $\alpha_{w,F_2}(l-1) = \beta = \alpha_{w,F_2}(l+n+1)$, *cf.* Figure 3.19, we have that $t = v_1$. Similarly, since v_2v_4 (resp. v_4t) is a suffix of $\alpha_{w,F_1}(k+1) = \beta$ (resp. $\alpha_{w,F_2}(l+n+1) = \beta$) of the length $|\alpha|$, necessarily, $v_2v_4 = v_4t = v_4v_1$. We obtain the following system of equations with unknowns $Y = \{\alpha, \beta, v_1, v_2, v_3, v_4\}$:

$$\begin{aligned}
v_2v_4 &= v_4v_1 && \text{(by the argument above)},\\
\alpha &= v_1v_3 = v_4v_2 && \text{(by Equation (3.3))},\\
v_3\alpha^{n-1}\beta &= \beta\alpha^{n-1}v_4 && \text{(by Equation (3.4))},\\
v_2\beta &= v_2\alpha_{w,F_2}(l+n+1) && \\
&= \alpha_{w,F_1}(k+2)v_1 = \beta v_1 && (\textit{cf.}\ \text{Figure 3.21}).
\end{aligned}$$

The dependency graph associated with this system is connected, and hence all unknowns commute, in particular α and β, a contradiction. This completes the proof of the theorem. □

3.2.2 Comments on Theorem 3.11

Theorem 3.11 deserves a few comments.

The theorem is related to the main result of [LRLR]:

Theorem 3.12. [LRLR] *Let $X = \{x, y\} \subset A^+$ be a code. If $w \in X^+$ such that $|w| \geq |x| + 2|y|$ admits an X-interpretation disjoint from w then one of the following conditions is satisfied:*

(1) x and y are powers of two conjugates, and $w \in x^+ \cup y^+$;

(2) x and y are not powers of two conjugates, and there is a non-primitive word $z \in x^*y \cup xy^*$ such that w is a factor of a word in z^+.

Indeed, Theorem 3.11 can be deduced by the theorem above and by Theorem 3.15 which was stated and partially proved in [LeS], *cf.* Subsection 3.2.3 which contains the full proof of this theorem. However, our proof of Theorem 3.11 is self-contained and essentially shorter, and moreover formulated directly to yield a defect-type of theorem.

The number of different X-factorizations of an X-ambiguous bi-infinite word is very different in the cases *(i)–(iii)* of the theorem. In the case *(i)* there exist non-denumerably many such X-factorizations, in the case *(ii)* there are finitely many different X-factorizations, and if we consider all shift-equivalent X-factorizations as the same, then there are exactly two of them. Finally, in the case *(iii)* there are also finitely many different X-factorizations, which are all shift-equivalent. This actually means that in the case *(iii)* no bi-infinite word can be expressed in two different ways as a product of words from X. Hence, indeed, Theorem 3.11 shows a defect effect of a two-element set for bi-infinite factorizations.

In Theorem 3.11 we showed that if the words of X do not commute and their primitive roots are not conjugates then only the case *(iii)* is possible. But if they do not commute and are conjugates then the theorem allows either the case *(ii)* or the case *(iii)*. In the sequel we will prove that in this situation only the case *(ii)* is possible.

Lemma 3.13. *If α and β are different conjugates then all elements of the set $\alpha\beta^+ \cup \alpha^+\beta$ are primitive.*

Proof. Since α and β are conjugates, they are of the same length. Hence, they do not commute, otherwise they would be equal.

Assume, for instance, that $\alpha\beta^n$ for an $n \geq 1$ is not primitive. We have that $\alpha\beta^n = t^i$ for a primitive word t and an integer $i \geq 2$. If $n = 1$ and i is even then, immediately, $\alpha = \beta$, which is a contradiction. If $n = 1$ and i is odd, say $i = 2m+1$, we have that $\alpha = t^m p$ and $\beta = qt^m$, where $t = pq$. Since α and β are conjugates, by Corollary 3.10, we have a contradiction with the primitiveness of $t = pq$.

It remains to consider the case $n \geq 2$. By Lemma 2.6, $|t| > (n-1)|\beta|$. On the other hand, $i|t| = |\alpha| + n|\beta| = (n+1)|\beta|$ which implies that $n+1 > i(n-1)$. Since $n, i \geq 2$, this yields $n = i = 2$. We have $\alpha\beta^2 = t^2$. Since $|\alpha| = |\beta|$, it follows that there are words p and q such that $|p| = |q|$, $\beta = pq$ and $t = \alpha p = q\beta = qpq$. Hence, $p = q$ and $\alpha = \beta$, a contradiction. □

A slightly weaker variant of Lemma 3.13 was proved in [LeS]. It states, under the additional assumption that α and β are primitive, that all the words in $\alpha\beta^* \cup \alpha^*\beta$ are primitive. The lemma yields the following improvement of Theorem 3.11.

Theorem 3.14. *Consider a binary set $X = \{\alpha, \beta\}$ with $\alpha, \beta \in \Sigma^+$. Let w be a proper X-ambiguous bi-infinite word possessing two different X-factorizations F_1 and F_2. Then at exactly one of the following conditions is satisfied:*

(i) α and β commute; or

(ii) the primitive roots of α and β are different conjugates, $w = \alpha^{\mathbb{Z}} = \beta^{\mathbb{Z}}$, and $F_1 \in [\alpha]_w^{\mathbb{Z}}$ and $F_2 \in [\beta]_w^{\mathbb{Z}}$, or vice versa; or

(iii) α and β do not commute and there exists an integer $n \geq 1$ such that either $w = (\alpha\beta^n)^{\mathbb{Z}}$ and $F_1, F_2 \in [\alpha\beta^n]_w^{\mathbb{Z}}$ with α primitive, or $w = (\beta\alpha^n)^{\mathbb{Z}}$ and $F_1, F_2 \in [\beta\alpha^n]_w^{\mathbb{Z}}$ with β primitive.

Proof. If $\rho(\alpha) \neq \rho(\beta)$ then α and β do not commute, hence the cases *(i)* and *(ii)* are exclusive. Obviously, the cases *(i)* and *(iii)* do so. Thus, it suffices to show that the conditions

- the primitive roots of α and β are different conjugates;
- the case *(iii)*

cannot happen at the same time. Assume to the contrary that $\rho(\alpha)$ and $\rho(\beta)$ are conjugates and, for instance, $w = (\alpha\beta^n)^{\mathbb{Z}}$ with $\alpha = \rho(\alpha)$ primitive and $F_1, F_2 \in [\alpha\beta^n]_w^{\mathbb{Z}}$. Let k be an integer such that $\beta = \rho(\beta)^k$. Then, by Observation 3.4, $\alpha\beta^n = \rho(\alpha)\rho(\beta)^{nk}$ is non-primitive. By Lemma 3.13, it follows that $\rho(\alpha)$ and $\rho(\beta)$ are equal, a contradiction. □

Note that the case *(i)* characterize the situation when the primitive roots of α and β are equal, the case *(ii)* the situation when they are different conjugates, and finally, the case *(iii)* the situation when they are not even conjugates.

3.2.3 The maximal number of X-ambiguous bi-infinite words

Let $X = \{\alpha, \beta\} \subseteq \Sigma^+$ be a binary set. In Theorem 3.3 we have proved that if the combinatorial rank of the set X equals to $\mathrm{card}(X)$ then the number of X-ambiguous bi-infinite words is finite. In this section we will prove that in the two-element case, for each set X, there are at most two X-ambiguous bi-infinite words, and at most one proper X-ambiguous bi-infinite word.

In the case *(i)* when $\mathrm{rank}_c(X) = 1$, since the both elements of X are powers of a common word t, the only possible X-ambiguous bi-infinite word is $t^{\mathbb{Z}}$. The situation is also trivial in the case *(ii)* when the primitive roots of elements of X are conjugates: by Theorem 3.14, the only possible X-ambiguous bi-infinite word is $w = \alpha^{\mathbb{Z}} = \beta^{\mathbb{Z}}$.

Finally, consider the case *(iii)* when the primitive roots of α and β are not conjugates. By Claim 3.7, there are at most two non-proper X-ambiguous bi-infinite words: $\alpha^{\mathbb{Z}}$ with α non-primitive and $\beta^{\mathbb{Z}}$ with β non-primitive. By Theorem 3.11, any proper X-ambiguous bi-infinite word is of the form $(\alpha\beta^n)^{\mathbb{Z}}$, with α primitive, or $(\alpha^n\beta)^{\mathbb{Z}}$, with β primitive. Moreover, by Observation 3.4, the word $\alpha\beta^n$ or the word $\alpha^n\beta$ is non-primitive, respectively. Hence, the number of X-ambiguous (resp. proper X-ambiguous) bi-infinite words is equal to the number of non-primitive elements of the set $\alpha\beta^* \cup \alpha^*\beta$ (resp. $\alpha\beta^+ \cup \alpha^+\beta$).

As we stated in the previous subsection, *cf.* Lemma 3.13, if α and β are different conjugates then all the words in the set $\alpha\beta^+ \cup \alpha^+\beta$ are primitive. Now, we are interested in a similar result in the general case assuming only that α and β do not commute. Such a result was stated in [LeS] as follows:

Theorem 3.15. [LeS] *Let α and β be two different primitive words. Then at most one word in the set $\alpha\beta^+ \cup \alpha^+\beta$ is non-primitive.*

There is an outline of the proof of the claim in the end of [LeS] which refers to the proof of another theorem in the paper. Here, we will give a full proof in the form of the following two lemmas.

Lemma 3.16. *Let α and β be two different primitive words. Then for integers any $n, m \geq 0$ such that $n \neq m$, at most one of the words $\alpha\beta^n$ and $\alpha\beta^m$ is non-primitive.*

Proof. Assume on the contrary that both $\alpha\beta^n$ and $\alpha\beta^m$ are non-primitive with $m < n$. If $m = 0$ the claim is obvious, so assume that $m \geq 1$, implying $n \geq 2$. We have

$$\left.\begin{aligned} \alpha\beta^n &= s^i \\ \alpha\beta^m &= t^j \end{aligned}\right\} \quad \text{and therefore also} \quad s^i = t^j\beta^{n-m}, \tag{3.5}$$

for some primitive words s and t and integers $i, j \geq 2$. Now if $n - m \geq 2$ then, by Lemma 2.7, s, t and β are powers of a common word, and so are α and β, which is a contradiction. So we can assume $m = n - 1$, and thus equation (3.5) simplifies to $s^i = t^j\beta$.

By Lemma 2.6, we have

$$|s| > (n-1)|\beta| \geq |\beta|\,, \tag{3.6}$$

and similarly, by (3.5),
$$|s| > (j-1)|t|\,, \tag{3.7}$$

implying
$$|t| + |\beta| > (i-1)|s|\,. \tag{3.8}$$

Inequalities (3.7) and (3.8) implies

$$|t| + |\beta| \overset{(3.8)}{>} (i-1)|s| \overset{(3.7)}{>} (i-1)(j-1)|t|\,,$$

and hence, we obtain

$$|\beta| > [(i-1)(j-1)-1].|t|\,,$$

and similarly, by (3.6) and (3.8),
$$|t| > [(i-1)(n-1)-1].|\beta|\,.$$

Now, if $|t| > |\beta|$, the expression $(i-1)(j-1)-1$ is necessarily equal to 0, implying $i = j = 2$. Similarly, if $|t| < |\beta|$, we have that $i = n = 2$. Let us consider these two case separately.

Case $|t| > |\beta|$ and $i = j = 2$. The equation $\alpha\beta^n = t^2$ implies that $t = x\beta$ for some $x \neq 1$. Thus equation (3.5) yields to $s^2 = t^2\beta = x\beta x\beta\beta$, which implies that $|\beta|$ is an even integer, $|x\beta| < |s|$ and $\frac{3}{2}|\beta| < |s|$. Hence, we can write $s = x\beta y = z\beta_2\beta$ for some $y, z \neq 1$, where $|y| = |\beta_2| = \frac{|\beta|}{2}$, $\beta = \beta_1\beta_2$ and $|x| = |z|$. We can divide this equation into two parts: $x = z$ and $\beta y = \beta_2\beta$, where the second one, by Lemma 2.4, contradicts the primitiveness of β.

Case $|t| < |\beta|$ and $i = n = 2$. The inequality (3.6) simplifies to $|s| > |\beta| > |t|$. By the equations $\alpha\beta = t^j$ and $\alpha\beta^2 = s^2$ we can write $\beta = xt$ and $s = yt$ for some $x, y \neq 1$. Hence, Equation (3.5) yields $ytyt = t^j\beta$. We have that $|yt| = |s| = |t^j| + |\beta| - |s| < |t^j|$, so that we can write $ytz = t^j$, $z \neq 1$. Now, by Lemma 2.4, either t is not primitive, or t matches with some t in t^j in the above equation, but then we have $y = t^k$, and hence also $s = t^{k+1}$, so that the words t and s commute.

In both cases we arrive into a contradiction. □

Lemma 3.17. *Let α and β be two different primitive words. Then for any integers $n, m \geq 0$ such that $(n, m) \neq (1, 1)$, at most one of the words $\alpha\beta^n$ and $\alpha^m\beta$ is non-primitive.*

Proof. The cases $m = 0$ and $n = 0$ are trivial. The case $m = 1$ is a special case of Lemma 3.16. In the case $n = 1$ we can exchange α with β, while considering their mirror images, and we are again in the case $m = 1$. We will

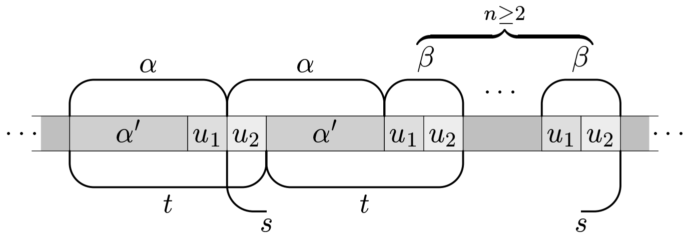


Figure 3.22: The situation in the case $\alpha^2\beta = t^2$ and $\alpha\beta^n = s^i$ with $n \geq 2$.

use this argument several times later on, and we will refer to it as to *the reverse argument*. Consider $n, m \geq 2$ and assume on the contrary that $\alpha\beta^n = s^i$ and $\alpha^m\beta = t^j$ for some integers $i, j \geq 2$ and primitive words s and t. By Lemma 2.6, we have

$$|s| > (n-1)|\beta| \geq |\beta|\,, \qquad |t| > (m-1)|\alpha| \geq |\alpha|\,. \tag{3.9}$$

Hence

$$\begin{aligned} |\alpha| = i|s| - n|\beta| > (in - i - n)|\beta|\,, \\ |\beta| = j|t| - m|\alpha| > (jm - j - m)|\alpha|\,, \end{aligned} \tag{3.10}$$

which implies that

$$\big[(i-1)(n-1) - 1\big] \cdot \big[(j-1)(m-1) - 1\big] < 1\,.$$

Therefore, we have that either $i = n = 2$, or $j = m = 2$. By the reverse argument, the first case is equivalent to the second one. Hence, let us consider only the case $j = m = 2$. If $|t| < |\beta|$ then, by (3.9), we obtain

$$|\alpha| \overset{(3.9)}{<} |t| < |\beta| \overset{(3.9)}{<} |s|\,.$$

Together with Inequalities (3.10) we have that $(i-1)(n-1)-1 < 1$ which implies that also $i = n = 2$. Now, by the reverse argument: the inequality $|s| > |\alpha|$ transforms to the inequality $|t| > |\beta|$. Therefore, without loss of generality, we can assume that $|t| > |\beta|$. We have the situation depicted in Figure 3.22, where $\beta = u_1u_2$ with $|u_1| = |u_2| = \frac{1}{2}|\beta|$ and $\alpha = \alpha'u_1 = u_2\alpha'$.

Since $u_2\alpha' = \alpha'u_1$, Lemma 2.10 gives us

$$\left.\begin{aligned} u_2 &= (pq)^k \\ u_1 &= (qp)^k \\ \alpha' &= p(qp)^l \end{aligned}\right\} \quad \text{and therefore} \quad \left\{\begin{aligned} \alpha &= p(qp)^{k+l} \\ \beta &= (qp)^k(pq)^k \end{aligned}\right. ,$$

where $k \geq 1$, $l \geq 0$ and pq is primitive. We may assume $p, q \neq 1$. Now considering the last occurrence of s in Figure 3.22, by (3.9), we can write $s =$

$s'\beta = s'(qp)^k(pq)^k$ for some word s'. We also have

$$|s| = |\alpha| + n|\beta| - (i-1)|s| \leq |\alpha| + n|\beta| - |s| \overset{(3.9)}{<} |\alpha| + |\beta|\,,$$

which yields

$$s'(qp)^k(pq)^k r = sr = \alpha\beta = \underbrace{p(qp)^{2k+l}}_{v}(pq)^k,$$

for some $r \neq 1$. The first occurrence of qp in s after s' must match with qp in w, otherwise qp is not primitive. But then, since $r \neq 1$, the first occurrence of pq in s after $s'(qp)^k$ matches with some qp in v, so we have that $pq = qp$, which is again a contradiction with the primitiveness of pq. □

Obviously, Lemmas 3.16 and 3.17 imply Theorem 3.15. Nevertheless, Theorem 3.15 is not directly applicable to our problem, since we cannot assume that α and β are primitive. As an immediate corollary of Lemma 2.7 and Theorem 3.15 we have

Corollary 3.18. *Let α and β be two different primitive words. Then at most one word in the set $\alpha^+\beta^+$ is non-primitive.*

This yields

Corollary 3.19. *Let α and β be two non-commuting words. Then*

- *at most* 1 *word in $\alpha\beta^+ \cup \alpha^+\beta$ is non-primitive;*
- *at most* 2 *words in $\alpha\beta^* \cup \alpha^*\beta$ are non-primitive.*

Proof. The first part of the claim follows by the relation

$$\alpha\beta^+ \cup \alpha^+\beta \subseteq \rho(\alpha)^+\rho(\beta)^+$$

and Corollary 3.18.

Consider the second part. Note that $\alpha\beta^* \cup \alpha^*\beta = \alpha\beta^+ \cup \alpha^+\beta \cup \{\alpha, \beta\}$. Hence, if at most one of the words α and β in non-primitive then the result follows by the first part of the claim. Otherwise, we have $\alpha\beta^+ \cup \alpha^+\beta \subseteq \rho(\alpha)^{2+}\rho(\beta)^{2+}$, where v^{2+} is an abbreviation for vv^+. Since α and β do not commute, by Lemma 2.7, all words in $\rho(\alpha)^{2+}\rho(\beta)^{2+}$ are primitive. Consequently, if α and β are both non-primitive then all the other words in $\alpha\beta^* \cup \alpha^*\beta$ are primitive. □

Finally, let us apply Corollary 3.19 to our problem. As a consequence of Corollary 3.19 and the considerations in the beginning of this section we obtain

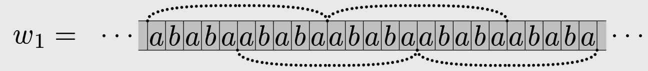
Theorem 3.20. *Consider a binary set $X = \{\alpha, \beta\}$ with $\alpha, \beta \in \Sigma^+$. There is at most one proper X-ambiguous bi-infinite word and at most two X-ambiguous bi-infinite words.*

The following example shows that the result is sharp.

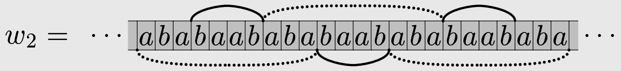
Example 3.6. There are exactly two types of the binary sets X such that $\mathrm{card}(\mathrm{Amb}(X)) = 2$, *cf.* proof of Corollary 3.19.

The obvious case is when both α and β are non-primitive and their primitive roots are not conjugates. Then $\alpha^{\mathbb{Z}}$ and $\beta^{\mathbb{Z}}$ are two different X-ambiguous bi-infinite words. Note that none of them is proper.

The less obvious case is when one of the words α and β is non-primitive, they do not commute and there is an integer $n \geq 1$ such that $\alpha\beta^n$ (resp. $\alpha^n\beta$) is non-primitive. As an example, take $\alpha = baab$ and $\beta = (ababa)^2$ non-primitive. Then $w_1 = \beta^{\mathbb{Z}}$ and $w_2 = (\alpha\beta)^{\mathbb{Z}}$ are different X-ambiguous bi-infinite words:



and



Note that only the bi-infinite word w_2 is proper.

3.2.4 Existence of an X-ambiguous bi-infinite word

Let X be a binary set containing non-empty words. In the previous subsection we proved that there is at most one proper X-ambiguous bi-infinite word and at most two X-ambiguous bi-infinite words. A natural question to ask is when such words exist. The answer is easy for the non-proper X-ambiguous bi-infinite words: such a word exists if and only if α or β is non-primitive. Hence, let us concentrate on the existence of a proper X-ambiguous bi-infinite word.

One can observe that there are sets X for which there is no proper X-ambiguous bi-infinite word. For example, take a set $X = \Sigma = \{a, b\}$. We say that a family of sets of words with the same cardinality k is *parameterizable* if it can be described in terms of k formulas with word and integer parameters, *cf.* Section 9.3 of [Lo] for details. Here, we are going to prove that the family of binary sets X for which there exists a proper X-ambiguous bi-infinite word is parameterizable.

In the case *(i)* of Theorem 3.11, when words of X are powers of a common word t, the bi-infinite word $t^{\mathbb{Z}}$ has infinitely many X-factorizations. In particular, in this case there is always an X-ambiguous bi-infinite word. In the case *(ii)*, when the primitive roots of the words of $X = \{\alpha, \beta\}$ are conjugates, the bi-infinite word $\alpha^{\mathbb{Z}} = \beta^{\mathbb{Z}}$ has exactly two different X-factorizations, so it is proper X-ambiguous.

Consider now the last case, the case *(iii)*, and a set $X = \{\alpha, \beta\}$. By Theorem 3.11, an X-ambiguous bi-infinite word is of the form $(\alpha\beta^n)^{\mathbb{Z}}$, where $\alpha\beta^n$ is non-primitive, or $(\alpha^n\beta)^{\mathbb{Z}}$, where $\alpha^n\beta$ is non-primitive, *i.e.,* there are integers $n \geq 1$, $i \geq 2$ and a primitive word $s \in \Sigma^+$ such that

$$\alpha\beta^n = s^i \quad \text{or} \quad \alpha^n\beta = s^i \, . \tag{3.11}$$

Conversely, if for some $n \geq 1$ and $i \geq 2$ at least one of equations (3.11) has a solution then, clearly, the bi-infinite word $(\alpha\beta^n)^{\mathbb{Z}}$ (resp. $(\alpha^n\beta)^{\mathbb{Z}}$) has exactly i shift-equivalent, but different X-factorizations. We formalize this as a lemma.

Lemma 3.21. *Let $X = \{\alpha, \beta\} \subseteq \Sigma^+$ be a binary set. Assume that the primitive roots of α and β are not conjugates. Then there is a proper X-ambiguous bi-infinite word if and only if one of the equations $\alpha\beta^n = s^i$ and $\alpha^n\beta = s^i$, with $n \geq 1$, $i \geq 2$, has a solution.*

We shall also give a characterization of the solutions of the equations (3.11). We need the following lemma.

Lemma 3.22. *All non-periodic solutions of the equation*

$$u_1u_2 = u_3(u_2u_3)^m, \quad m \geq 1 \tag{3.12}$$

are of the form

$$\begin{aligned} u_3 &= qp \, , \\ u_2 &= p(qp)^k, \\ u_1 &= u_3(u_2u_3)^{m-1}pq \, , \end{aligned} \tag{3.13}$$

where $p, q \in \Sigma^+$, $k \geq 0$.

Proof. It is easy to check that (3.13) is really a solution of equation (3.12). Now we shall prove that if equation (3.12) has a non-periodic solution, then it is of the form (3.13). We proceed by induction.

Consider first the case $m = 1$. We have the equation $u_1u_2 = u_3u_2u_3$. It is obvious that $|u_1| > |u_3|$, so we can write $u_1 = u_3t$. The equation transforms into $tu_2 = u_2u_3$, which has, by Lemma 2.10, the only solutions $t = pq$, $u_3 = qp$ and $u_2 = p(qp)^k$, $k \geq 0$. This implies that $u_1 = qppq$, so we have a solution of the form (3.13) for $m = 1$.

Consider now equation (3.12) with $m \geq 2$. Again we have that $|u_1| > |u_3|$, so we can substitute $u_1 = u_3t$ and equation (3.12) becomes $tu_2 = u_2u_3(u_2u_3)^{m-1}$. By Lemma 2.10, we have $t = uv$, $u_3(u_2u_3)^{m-1} = vu$, $u_2 = u(vu)^l$ for an integer $l \geq 0$. If $l \geq 1$, then $|vu| = |u_3(u_2u_3)^{m-1}| \geq 2|u| + |v| + |u_3|$. This implies that $u = u_3 = 1$, which leads to a periodic solution. Hence, consider the case $l = 0$. We have that $u_2 = u$, $u_1 = u_3u_2v$ and $vu_2 = u_3(u_2u_3)^{m-1}$. Now we can apply

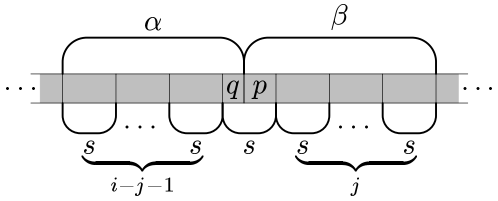


Figure 3.23: The situation when $|s| < |\beta^n|$ and $n = 1$.

induction hypothesis on the last equation and we obtain that all non-periodic solutions are of the form

$$u_3 = qp\,, \quad u_2 = p(qp)^k, \quad v = u_3(u_2u_3)^{m-2}pq\,, \quad k \geq 0\,,$$

which implies $u_1 = u_3u_2v = u_3(u_2u_3)^{m-1}pq$. We obtained exactly solution (3.13), which completes the proof. □

The following lemma gives us the characterization of solutions of equation (3.11), and hence also of sets X allowing an X-ambiguous bi-infinite word in the case *(iii)*.

Lemma 3.23. *Assume that words α and β do not commute. All solutions of the equation $\alpha\beta^n = s^i$ satisfying $n \geq 1$, $i \geq 2$ are*

$$\begin{aligned} \beta &= p(qp)^j, \\ s &= qp\beta^{n-1}, \\ \alpha &= s^{i-1}\beta^{-1}pq\,, \end{aligned} \tag{3.14}$$

where $p, q \in \Sigma^+$, $j \geq 0$ and $j < i$ if $n = 1$.

Proof. It is easy to check that (3.14) is a solution of equation (3.11). For the converse implication we analyze the following 3 cases:

Case $|s| > |\beta^n|$. Then we have that $\alpha = s^{i-1}q$ and $s = q\beta^n$ for some $q \neq 1$. This is solution (3.14) for $j = 0$, $p = \beta$.

Case $|s| < |\beta^n|$ and $n = 1$. The situation is depicted in Figure 3.23. According to the figure we can write

$$\beta = p(qp)^j, \quad s = qp\,, \quad \alpha = q(pq)^{i-j-1},$$

where $p, q \neq 1$ and $j < i$. Since

$$s^{i-1}\beta^{-1}pq = (qp)^{i-1}\left[p(qp)^j\right]^{-1}pq = (qp)^{i-j-1}q = \alpha\,,$$

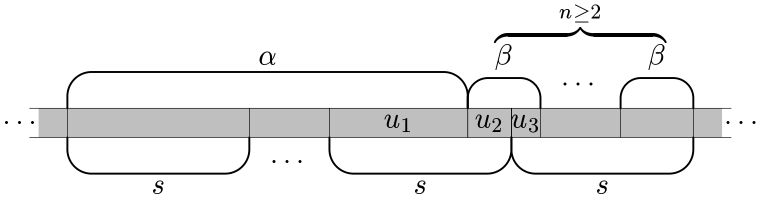


Figure 3.24: The situation when $|s| < |\beta^n|$ and $n \geq 2$.

we have the solution (3.14) with $n = 1$.

Case $|s| < |\beta^n|$ and $n \geq 2$. Since we are looking for non-periodic solutions, by Lemma 2.6, necessarily $|s| > |\beta^{n-1}|$. Hence, we have a situation shown in Figure 3.24. According to this figure we can write $\beta = u_2u_3$, $\alpha = s^{i-2}u_1$ and $u_1u_2 = s = u_3\beta^{n-1} = u_3(u_2u_3)^{n-1}$, which is equation (3.12). Now, Lemma 3.22 implies

$$\begin{aligned}
&\beta = u_2u_3 = p(qp)^{k+1} = p(qp)^j, \text{ for } j = k+1,\\
&s = u_1u_2 = u_3(u_2u_3)^{n-2}pqp(qp)^k = qp\beta^{n-2}\beta = qp\beta^{n-1}, \text{ and}\\
&\alpha = s^{i-2}u_1 = s^{i-2}u_3(u_2u_3)^{n-2}pq = s^{i-2}qp\beta^{n-2}\beta\beta^{-1}pq = s^{i-1}\beta^{-1}pq\,.
\end{aligned}$$

This is exactly solution (3.14). □

The following theorem summarizes the previous results.

Theorem 3.24. *Consider a binary set $X \subseteq \Sigma^+$. There exists a proper X-ambiguous bi-infinite word if and only if at least one of the following conditions is satisfied:*

(i) $X = \{p^n, p^m\}$, where $p \in \Sigma^+$ and $n, m \geq 1$,

(ii) $X = \{(pq)^n, (qp)^m\}$, where $p, q \in \Sigma^+$ and $n, m \geq 1$,

(iii) $X = \{\alpha, \beta\}$, where

$$\beta = p(qp)^j, \quad \alpha = (qp\beta^{n-1})^{i-1}\beta^{-1}pq\,,$$

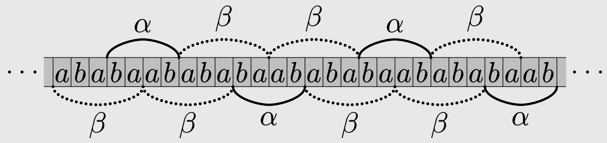
for $p, q \in \Sigma^+$, $n \geq 1$, $i \geq 2$, $j \geq 0$ and if $n = 1$ then $j < i$.

Notice, that in the last case of Theorem 3.24 the occurrence of β^{-1} can be eliminated, but we prefer this form for its simplicity. This theorem shows that the family of the two-element sets X, such that there exists a proper X-ambiguous bi-infinite word, is parameterizable. Such a characterization can be used to generate all such sets.

Example 3.7. Let us choose in (3.14) $p = a$, $q = b$, $n = 2$, $i = 2$ and $j = 2$. We have

$$\beta = ababa\,, \quad s = baababa\,, \quad \alpha = baab\,.$$

The bi-infinite word $(\alpha\beta^2)^{\mathbb{Z}}$ has two different shift-equivalent X-factorizations:


α
β
β
α
β
$\cdots$ $a\,b\,a\,b\,a\,a\,b\,a\,b\,a\,b\,a\,a\,b\,a\,b\,a\,b\,a\,a\,b\,a\,b\,a\,b\,a\,a\,b$ $\cdots$
β
β
α
β
β
α

Chapter 4

A cumulative defect effect: an example

In Section 2.2 we have recalled the fundamental result of Combinatorics on Words: if a set of n non-empty words satisfies a non-trivial relation then the rank, *i.e.,* the dimension, of the set is at most $n-1$. This property is called a defect effect. However, the dimension properties of words are rather weak, a system of k independent relations in n unknowns does not force usually a defect effect by k, *i.e.,* the rank of the set of unknowns is often greater than $n-k$, and sometimes even equal to $n-1$. The simplest example of this behavior can be found in [ChK]:

Example 4.1. [ChK] The system

$$xzy = yzx \qquad \text{and} \qquad xzzy = yzzx$$

of equations is independent, since the former has a solution

$$x = aba \quad y = a \quad \text{and} \quad z = b$$

which is not a solution of the later, while the later has a solution

$$x = abba \quad y = a \quad \text{and} \quad z = b$$

which is not a solution of the former. However, they have a common solution of degree two

$$x = a \quad y = a \quad \text{and} \quad z = b\,.$$

The more complicated and convincing examples can be found in [KPl]. [KPl] contains an example of a system of n^3 independent equations in $5n$ unknowns which forces only the minimal defect effect by 1.

This inspires an interesting problem area to find conditions (on relations or on sets of words) which imply a cumulative defect effect, *i.e.*, if the set X of n words satisfy k relations then X has a rank at most $n-k$.

There are only very few results in this direction. The Graph Lemma, *i.e.*, Lemma 2.3 of Section 2.3, is such an example where the type of relations is restricted. A similar deep result is proved in [Br], extending ideas of [Ka1, Ka2, Ho], where it is shown that if X is a code and has an unbounded synchronizing delay in both directions then the rank of X is at most $\mathrm{card}(X)-2$.

In this chapter we will show another example of a cumulative defect effect for bi-infinite words. We interpret, in a natural way, a relation on words from X as a double X-factorization of an infinite word. In fact, we consider only the case when a bi-infinite word possesses k disjoint X-factorizations which we interpret as $k-1$ non-trivial relations. We ask the following:

Problem 4.1. *Let X be a finite set of words and w a non-periodic bi-infinite word. Is it true that if w possesses k disjoint X-factorizations, for $k \leq \mathrm{card}(X)$, then the combinatorial rank of X is at most* $\mathrm{card}(X)-k+1$ *?*

Our starting point is the result proved in Section 3.1 (Theorem 3.3) stating that if a non-periodic bi-infinite word possesses two different X-factorizations then the rank of X is at most $\mathrm{card}(X)-1$. Hence, Problem 4.1 is solved affirmatively in the case $k=2$. As emphasized at the end of Section 3.1 it is essential to use the notion of the combinatorial rank and to assume that the bi-infinite word is non-periodic, *cf.* Examples 3.2 and 3.3. In the general case it is also necessary to assume that the X-factorizations are disjoint:

Example 4.2. Consider a finite set X of words such that $\mathrm{rank_c}(X)=\mathrm{card}(X)-1$. Hence, X is not a code and it satisfies a non-trivial relation

$$v = x_1 \ldots x_n = y_1 \ldots y_m$$

for some $x_1, \ldots, x_n, y_1, \ldots, y_m \in X$. Let x be an element of X not equal to v. Then the non-periodic bi-infinite word

$$w = {}^{\mathbb{N}}vxv^{\mathbb{N}}$$

has infinitely many different X-factorizations of the form

$${}^{\mathbb{N}}\{x_1 \ldots x_n, y_1 \ldots y_m\}x\{x_1 \ldots x_n, y_1 \ldots y_m\}^{\mathbb{N}},$$

but we have a defect effect by only 1.

We do not have either a counterexample or a proof for larger values of k. However, we are able to prove the following results. If a non-periodic bi-infinite word possesses 3 disjoint X-factorizations, where X is a **prefix** code, then the combinatorial rank of X is at most $\text{card}(X)-2$. As we shall see in Section 4.1, even this simple case seems to be quite complicated to prove. In the case $k=\text{card}(X)$, Problem 4.1 implies a contradiction: w is a non-periodic bi-infinite word with an X-factorization and $X\subseteq t^+$ for a non-empty word t, hence it is equivalent to the problem to show that a non-periodic bi-infinite word can possess at most $\text{card}(X)-1$ disjoint X-factorizations. In Section 4.2, we solve, based on the Critical Factorization Theorem and its application, a slightly weaker version of this problem, *i.e.,* we show that the maximal number of disjoint X-factorizations of a non-periodic bi-infinite word is $\text{card}(X)$.

The notions of independent relations is formalized in [HKP].

4.1 Bi-infinite words possessing 3 different X-factorizations

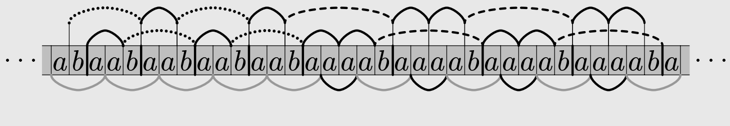
In this section we will show that in the case when X is a prefix set, if a non-periodic bi-infinite word possesses 3 different X-factorizations then we have a cumulative defect effect: the combinatorial rank of X is at most $\text{card}(X)-2$.

The following two examples show that, as we have seen in Chapter 3, the combinatorial rank is essential to obtain a defect effect for bi-infinite words. We will use these two examples later on to illustrate the proof of the main result of this chapter.

Example 4.3. Let $X=\{\alpha,\beta,\gamma,\delta\}$, where $\alpha=aa$, $\beta=baab$, $\gamma=baaaab$ and $\delta=aba$. The non-periodic bi-infinite word

$$w_1=\dots aab.aab.aaaab.aaaab\cdots={}^{\mathbb{N}}(aab)(aaaab)^{\mathbb{N}}$$

has three different X-factorizations: $F_1={}^{\mathbb{N}}(\beta\alpha)(\gamma\alpha\alpha)^{\mathbb{N}}$, $F_2={}^{\mathbb{N}}(\alpha\beta)(\alpha\alpha\gamma)^{\mathbb{N}}$ and $F_3={}^{\mathbb{N}}\delta(\alpha\delta)^{\mathbb{N}}$, which are pairwise non-shift-equivalent and are depicted as follows:


· · · a b a a b a a b a a b a a b a a b a a a a b a a a a b a a a a b a a a a b a · · ·

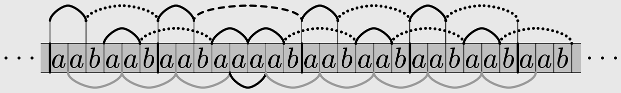
Clearly, these three factorizations are pairwise disjoint and also non-periodic. Moreover, $\text{rank}_\text{c}(X)=2$, since $X\subseteq\{a,b\}^+$, but for no word s the inclusion $X\subseteq s^+$ holds. On the other hand, since X is a prefix code we conclude that $\text{rank}_\text{p}(X)=\text{rank}_\text{f}(X)=4$.

Example 4.4. Let X be the same set as in the previous example. Take any non-periodic bi-infinite word w_2 in the set $\{aabaab, aabaaaab\}^{\mathbb{Z}}$. Any such bi-infinite word has 3 different X-factorizations: $F_1\in\{\alpha\beta,\alpha\gamma\}^{\mathbb{Z}}$, $F_2\in\{\alpha\beta,\alpha\alpha\beta\}^{\mathbb{Z}}$

and $F_3 \in \{\delta\delta, \delta\alpha\delta\}^{\mathbb{Z}}$, assuming that elements in the sets are selected in the same order as for w_2. For example, consider a part of w_2 in the form:

$$w_2 = \ldots aabaab.aabaaaab.aabaab.aabaab \ldots .$$

The corresponding parts of three X-factorizations are depicted as follows:


… a a b a a b a a b a a a a b a a b a a b a a b a a b a a b …

Again, three X-factorizations are pairwise disjoint, non-shift-equivalent and non-periodic, assuming that the bi-infinite word w_2 is so.

Examples 4.3 and 4.4 together with Example 3.3 and Theorem 3.3 show that in order to obtain the defect effect for bi-infinite words we have to use the combinatorial rank. It is also necessary to consider non-periodic X-factorizations or non-periodic bi-infinite words:

Example 4.5. In this example we show that for any positive integer k, there is a binary prefix set X without any defect effect and a periodic bi-infinite word with k disjoint X-factorizations.

Let $X = \{\alpha, \beta\}$, where $\alpha = a$ and $\beta = (ba)^{k-1}b$. Clearly, the bi-infinite word $w = (ab)^{\mathbb{Z}}$ has k disjoint X-factorizations of the form $(\alpha\beta)^{\mathbb{Z}}$. They are all shift-equivalent, but different. On the other hand, we have that $\mathrm{rank_c}(X) = 2 = \mathrm{card}(X)$.

To prove the defect theorem we need the following simple lemma.

Lemma 4.1. *Consider words $x, y, x', y', v \in \Sigma^+$ and $z_1, z_2, w_1, w_2 \in \Sigma^*$ satisfying equations*

$$xz_1 = vyw_1 \qquad \text{and} \qquad x'z_2 = vy'w_2\,. \tag{4.1}$$

If $y = x$, $y' = x'$ or $y = x'$, $y' = x$, i.e., *if $\{x, x'\} = \{y, y'\}$, then x and x' are left comparable,* i.e., *one is a prefix of the other.*

Proof. Consider, for example, the first case: $y = x$ and $y' = x'$. Without loss of generality we can assume $|x| \le |x'|$. If $|x'| \le |v|$ then $x' \le v$ and also $x \le v$, which implies that $x \le x'$, and we are done. Hence assume $|x'| > |v|$, *i.e., $v < x'$*. Now, if $|x| \le |v|$ then $x \le v \le x'$ and we are again done. Thus, the only case we have to consider is the case $|x'| \ge |x| > |v|$. We can substitute $x = v\bar{x}$, $x' = v\bar{x}'$ for some $\bar{x}, \bar{x}' \in \Sigma^+$. The equations (4.1) transforms into

$$\bar{x}z_1 = v\bar{x}w_1 \qquad \text{and} \qquad \bar{x}'z_2 = v\bar{x}'w_2\,.$$

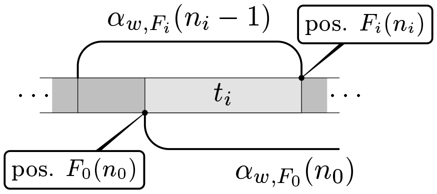


Figure 4.1: An illustration of the definition of t_i's.

We obtained the system of equations of the same type, but with $|\bar{x}| < |x|$ and $|\bar{x}'| < |x'|$. Hence, after a finite number of steps it must happen that the x's obtained, say $\tilde{x}$ and $\tilde{x}'$, are left comparable. Clearly, if $\bar{x}, \bar{x}'$ are left comparable then so are x, x'. Inductively, we obtain that x and x' are left comparable.

In the second case the proof is the same. □

Now, we can state and prove the main theorem of this chapter. Since the proof of the theorem and the two auxiliary lemmas is quite long and technical, we will illustrate the proof on examples. In fact, we will perform the proof on the bi-infinite words w_1 and w_2 defined in Examples 4.3 and 4.4.

Theorem 4.2. *Consider a prefix set $X \subseteq \Sigma^+$. Let w be a bi-infinite word over Σ with* 3 *different X-factorizations F_0, F_1, F_2. If the word w is non-periodic then the combinatorial rank of X is at most* $\operatorname{card}(X) - 2$.

Before we start to prove the theorem, let us define technical notions of X-differences, triples and minimal triples. Consider a bi-infinite word w possessing three disjoint X-factorizations F_0, F_1, F_2. Take an arbitrary factor $\alpha_{w,F_0}(n_0) \in X$ of w defined by the X-factorization F_0, and find, for $i = 1, 2$, the minimal starting point $F_i(n_i) \in F_i(\mathbb{Z})$ such that $F_i(n_i) \geq F_0(n_0)$, see Figure 4.1. Let us denote the word

$$w_{F_0(n_0)} w_{F_0(n_0)+1} \cdots w_{F_i(n_i-1)-1}$$

by t_i. We call the pair (t_1, t_2) an *X-difference*, or more precisely, an X-difference with respect to the triple (F_0, F_1, F_2), and we call the position $F_0(n_0)$ an *occurrence* of the X-difference (t_1, t_2). Note that t_1 and t_2 are always left comparable.

Assume that we have an occurrence $F_0(n_0)$ of an X-difference (t_1, t_2) followed by an occurrence $F_0(m_0)$ of an X-difference (t'_1, t'_2) in w. Figure 4.2 depicts such a situation, when $|t_1| \leq |t_2|$ and $|t'_1| \leq |t'_2|$. Consider the following 3 factors of w

$$\begin{aligned} f_0 &= \alpha_{w,F_0}(n_0)\alpha_{w,F_0}(n_0+1)\ldots\alpha_{w,F_0}(m_0-1)\,, \\ f_1 &= \alpha_{w,F_1}(n_1)\alpha_{w,F_1}(n_1+1)\ldots\alpha_{w,F_1}(m_1-1) \quad \text{and} \\ f_2 &= \alpha_{w,F_2}(n_2)\alpha_{w,F_2}(n_2+1)\ldots\alpha_{w,F_2}(m_2-1)\,. \end{aligned}$$

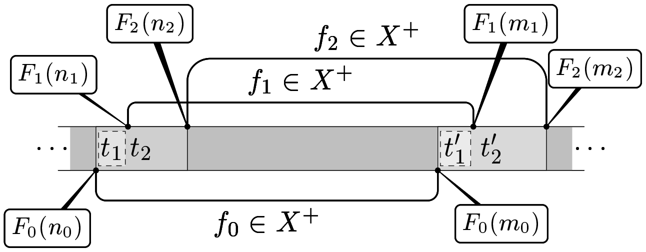


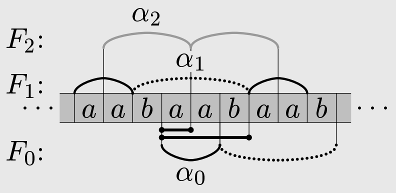
Figure 4.2: An illustration of a (t_1, t_2, t'_1, t'_2)-triple with $|t_1| \leq |t_2|$ and $|t'_1| \leq |t'_2|$..

The triple (f_0, f_1, f_2) is called a (t_1, t_2, t'_1, t'_2)-*triple*, and the position $F_0(n_0)$ is called an *occurrence* of the (t_1, t_2, t'_1, t'_2)-triple. Note that a (t_1, t_2, t'_1, t'_2)-triple (f_0, f_1, f_2) satisfies

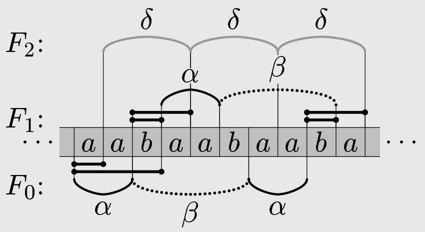
$$f_0 t'_1 = t_1 f_1 \qquad \text{and} \qquad f_0 t'_2 = t_2 f_2 \,. \tag{4.2}$$

We say that an occurrence $F_0(n)$ of an X-difference $(\tilde{t}_1, \tilde{t}_2)$ *occurs inside* of an occurrence $F_0(m)$ of a triple (f_0, f_1, f_2), if $F_0(m) \leq F_0(n) < F_0(m) + |f_0|$. Note that the definition above does not depend on the choice of the occurrences of the X-difference (resp. the triple), hence the definition applies also directly to X-differences (resp. triples). If there is no occurrence of X-difference (t_1, t_2) or (t'_1, t'_2) inside a (t_1, t_2, t'_1, t'_2)-triple, we say that the triple is *minimal*.

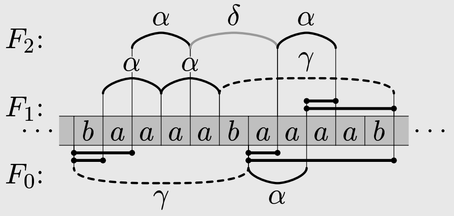
Example 4.6. Let us illustrate the previous definitions on the bi-infinite word w_1 from Example 4.3. Take as α_0, for instance, the first α in the factorization F_0 depicted in the figure of Example 4.3. Then $\alpha_1 = \beta$, $t_1 = aab$ and $\alpha_2 = \delta$, $t_2 = a$:



Hence, we have an occurrence of X-difference (aab, a) marked above with two black lines. Taking as α_0 the first γ in the factorization F_0 we find an occurrence of X-difference (b, ba). We have an (aab, a, b, ba)-triple $(\alpha\beta\alpha, \alpha\beta, \delta\delta\delta)$:



It contains inside of it occurrences of X-differences (b, ba) and (aab, a), hence it is not minimal. On other hand the following (b, ba, aab, a)-triple $(\gamma\alpha, \alpha\alpha\gamma, \alpha\delta\alpha)$ is minimal, since it contains inside of it only an occurrence of X-difference $(aaaab, a)$:



Indeed, here it is important to remember that the order of the factorizations is fixed.

The proof of Theorem 4.2 is similar to that of Theorem 3.3 assuming we have already proved the following two lemmas:

Lemma 4.3. *Consider a prefix set $X \subseteq \Sigma^+$. Let w be a bi-infinite word over Σ with 3 different X-factorizations F_0, F_1, F_2. If there are non-empty words t_1 and t_2 such that the bi-infinite word w possesses two different minimal (t_1, t_2, t_1, t_2)-triples then* $\mathrm{rank_c}(X) \leq \mathrm{card}(X) - 2$.

Lemma 4.4. *Consider a prefix set $X \subseteq \Sigma^+$. Let w be a bi-infinite word over Σ with 3 different X-factorizations F_0, F_1, F_2. If there are non-empty words t_1, t_2, t'_1 and t'_2 such that the bi-infinite word w possesses*

(i) a minimal (t_1, t_2, t_1, t_2)-triple without any occurrence of the X-difference (t'_1, t'_2) inside;

(ii) a minimal (t_1, t_2, t'_1, t'_2)-triple; and

(iii) a minimal (t'_1, t'_2, t'_1, t'_2)-triple

then $\mathrm{rank_c}(X) \leq \mathrm{card}(X) - 2$.

Rather technical Lemmas 4.3 and 4.4 can be proved in the similar way. Hence, we will give a full proof of Lemma 4.3 and, after that, we will point out the differences between the proofs of these two lemmas. Now, let us illustrate the situations these two lemmas deal with on an example.

Example 4.7. In Example 4.4 we have exactly the situation considered in Lemma 4.3. Any non-periodic bi-infinite word

$$w_2 \in \{aabaab, aabaaaab\}^{\mathbb{Z}}$$

contains exactly two different minimal (aab, a, aab, a)-triples $(\alpha\beta, \alpha\beta, \delta\delta)$ and $(\alpha\gamma, \alpha\alpha\beta, \delta\alpha\delta)$.

Further, Example 4.3 is an illustration of the case considered in Lemma 4.4. The bi-infinite word

$$w_1 = {}^{\mathbb{N}}(aab)(aaaab)^{\mathbb{N}}$$

contains:

(i) the minimal (aab, a, aab, a)-triple $(\alpha\beta, \alpha\beta, \delta\delta)$ without any occurrence of the X-difference $(aaaab, a)$ inside;

(ii) the minimal $(aab, a, aaaab, a)$-triple $(\alpha\gamma, \alpha\alpha\beta, \delta\alpha\delta)$; and

(iii) the minimal $(aaaab, a, aaaab, a)$-triple $(\alpha\alpha\gamma, \alpha\alpha\gamma, \alpha\delta\alpha\delta)$.

4.1.1 Proof of Lemma 4.3

Proof. Let us consider two different minimal (t_1, t_2, t_1, t_2)-triples (f_0, f_1, f_2) and (f'_0, f'_1, f'_2). Without loss of generality we can assume that $|f_0| \le |f'_0|$ and $t_1 < t_2$. Note that $t_1 \ne t_2$ since factorizations F_1, F_2 are disjoint.

Denote $t_0 = 1$ and let $s_1, s_2 \in \Sigma^+$ be such words that $t_1 = s_1$, $t_2 = t_1 s_2$. We define, for $0 \le a \le b \le 2$,

$$s_{(a,b]} = s_{a+1} \dots s_b .$$

Notice that for $a = b$, $s_{(a,b]} = 1$; for $0 \le a \le 2$, $t_a = s_{(0,a]}$; and for $a \le b \le c$, $s_{(a,c]} = s_{(a,b]} s_{(b,c]}$. Next, we define ±-notation: for arbitrary $a, b \in \{0, 1, 2\}$ let

$$s^+_{(a,b]} = \begin{cases} s_{(a,b]}, \text{ if } a < b, \\ 1, \text{ otherwise,} \end{cases} \qquad s^-_{(a,b]} = \begin{cases} s_{(b,a]}, \text{ if } b < a, \\ 1, \text{ otherwise.} \end{cases}$$

Equations (4.2) imply

$$f_a^{(\prime)} s_{(a,b]} = s_{(a,b]} f_b^{(\prime)}, \text{ for } a < b . \tag{4.3}$$

Equation (4.3), for each $a < b$, represents actually two equations: one with and one without the primed symbols, hence the notation $f^{(\prime)}$. Note that for $a = 0$

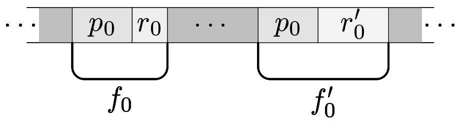


Figure 4.3: An illustration of the definition of p_i, r_i, r_i' for $i = 0$.

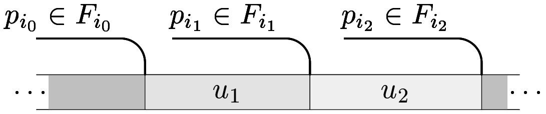


Figure 4.4: An illustration of u_i's.

we have only another transcription of Equations (4.2) and for $a = 1, b = 2$ we have

$$f_1 s_2 = s_2 f_2, \qquad f_1' s_2 = s_2 f_2' \,.$$

Using our ±-notation we can restate Equation (4.3) for any $0 \leq a, b \leq 2$:

$$s^{-}_{(a,b]} f^{(\prime)}_{a} s^{+}_{(a,b]} = s^{+}_{(a,b]} f^{(\prime)}_{b} s^{-}_{(a,b]} \,. \tag{4.4}$$

Indeed, it is easy to check that for $a < b$ and $b < a$ we get exactly Equation (4.3) and for $a = b$ a tautology $f^{(\prime)}_a = f^{(\prime)}_b$.

Let $p_i \in X^*$ be a common prefix of f_i, f_i' over the alphabet X and let $r_i, r_i' \in X^*$ be words such that $f_i = p_i r_i$ and $f_i' = p_i r_i'$ (see Figure 4.3).

Note that if $f_i = f_i'$ for any $i = 1, 2, 3$ then, by (4.2) and the choice of (t_1, t_2, t_1, t_2)-triples, $f_i = f_i'$ for all $i = 1, 2, 3$. This is impossible since X is a code and the triples are different. Thus, since $|f_0| \leq |f_0'|$ we have that $r_i' \neq 1$ for all i.

Let i_0, i_1, i_2 be the order of the ends of p_i's in the bi-infinite word w, as depicted in Figure 4.4. Note that since the X-factorizations are disjoint, the words u_1, u_2 are non-empty. Hence, we have that $|t_{i_0} p_{i_0}| < |t_{i_1} p_{i_1}| < |t_{i_2} p_{i_2}|$, where, we remind, $t_0 = 1$.

Example 4.7 (continued). Let us change the indexes of factorizations F_1 and F_2, so that the condition $t_1 < t_2$ is satisfied. Hence, we will consider (a, aab, a, aab)-triples $(\alpha\beta, \delta\delta, \alpha\beta)$ and $(\alpha\gamma, \delta\alpha\delta, \alpha\alpha\beta)$ with $s_1 = a$ and $s_2 = ab$. Then

$$\begin{aligned} p_0 &= \alpha, & r_0 &= \beta, & r_0' &= \gamma\,, \\ p_1 &= \delta, & r_1 &= \delta, & r_1' &= \alpha\delta\,, \\ p_2 &= \alpha, & r_2 &= \beta, & r_2' &= \alpha\beta\,. \end{aligned}$$

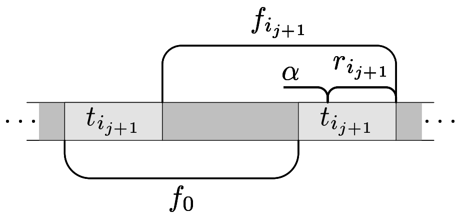


Figure 4.5: A hypothetical situation in which an α between the end of f_0 and the end of $f_{i_{j+1}}$ would exist.

Since $|p_0| < |t_1p_1| < |t_2p_2|$, the order of p_i's is $i_0 = 0$, $i_1 = 1$ and $i_2 = 2$, and $u_1 = ba$ and $u_2 = a$.

We have

$$t_{i_0}p_{i_0}u_1 = t_{i_1}p_{i_1}\,, \qquad t_{i_1}p_{i_1}u_2 = t_{i_2}p_{i_2}\,, \qquad t_{i_0}p_{i_0}u_1u_2 = t_{i_2}p_{i_2}\,. \tag{4.5}$$

Taking the first equation and multiplying both sides by $r_{i_1}s^-_{(i_0,i_1]}$ we obtain

$$s_{(0,i_0]}p_{i_0}u_1r_{i_1}s^-_{(i_0,i_1]} = t_{i_0}p_{i_0}u_1r_{i_1}s^-_{(i_0,i_1]} \overset{(4.5)}{=} t_{i_1}p_{i_1}r_{i_1}s^-_{(i_0,i_1]} = s_{(0,i_1]}f_{i_1}s^-_{(i_0,i_1]}\,.$$

This is equivalent to

$$s^-_{(i_0,i_1]}p_{i_0}u_1r_{i_1}s^-_{(i_0,i_1]} = s^+_{(i_0,i_1]}f_{i_1}s^-_{(i_0,i_1]} \overset{(4.4)}{=} s^-_{(i_0,i_1]}f_{i_0}s^+_{(i_0,i_1]} = s^-_{(i_0,i_1]}p_{i_0}r_{i_0}s^+_{(i_0,i_1]}\,,$$

hence $u_1r_{i_1}s^-_{(i_0,i_1]} = r_{i_0}s^+_{(i_0,i_1]}$. In the similar way we obtain

$$u_1r^{(\prime)}_{i_1}s^-_{(i_0,i_1]} = r^{(\prime)}_{i_0}s^+_{(i_0,i_1]}, \tag{4.6}$$

$$u_2r^{(\prime)}_{i_2}s^-_{(i_1,i_2]} = r^{(\prime)}_{i_1}s^+_{(i_1,i_2]}, \tag{4.7}$$

$$u_1u_2r^{(\prime)}_{i_2}s^-_{(i_0,i_2]} = r^{(\prime)}_{i_0}s^+_{(i_0,i_2]}. \tag{4.8}$$

If $r_{i_j} = 1$, for $j = 0,1$ then Equations (4.6) and (4.7) imply $s^+_{(i_j,i_{j+1}]} \neq 1$, and hence also $i_j < i_{j+1}$ and $|r_{i_{j+1}}| < |s_{(i_j,i_{j+1}]}| \leq |t_{i_{j+1}}|$. By the definition of the X-differences, we then obtain $r_{i_{j+1}} = 1$, otherwise there is an $\alpha \in X$ in the X-factorization $F_{i_{j+1}}$, which ends before the end of $f_{i_{j+1}}$ and after the end of f_0, as illustrated in Figure 4.5. But this is impossible by the definition of X-differences, see Figure 4.1.

We have three possibilities:

(i) $r_{i_0} = r_{i_1} = r_{i_2} = 1$, $i_0 < i_1 < i_2$;

(ii) $r_{i_0} \neq 1$, $r_{i_1} \neq 1$;

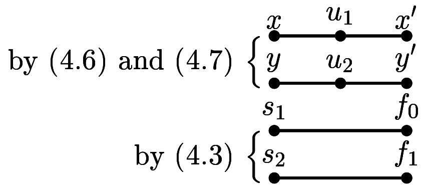


Figure 4.6: A part of the dependency graph in the case (ii).

(iii) $r_{i_0} \neq 1$, $r_{i_1} = r_{i_2} = 1$, $i_1 < i_2$.

Case (i). In this case there is an occurrence of the X-difference (t_1, t_2) inside the minimal (t_1, t_2, t_1, t_2)-triple (f'_0, f'_1, f'_2), which is a contradiction.

Case (ii). Let $x, x', y, y' \in X$ be the first letters of $r_{i_0}, r'_{i_0}, r_{i_1}, r'_{i_1}$, respectively. Hence, $r_{i_0}^{(\prime)} = x^{(\prime)}\bar{r}_{i_0}^{(\prime)}$ and $r_{i_1}^{(\prime)} = y^{(\prime)}\bar{r}_{i_1}^{(\prime)}$ with some $\bar{r}_{i_0}, \bar{r}'_{i_0}, \bar{r}_{i_1}, \bar{r}'_{i_1} \in X^*$. Clearly, since p_{i_0} and p_{i_1} are the maximal common prefixes, $x \neq x'$ and $y \neq y'$. Using this notation Equations (4.6) transform to

$$\begin{aligned} u_1 y \quad \overbrace{\bar{r}_{i_1} s^-_{(i_0,i_1]}}^{w_1} &= x \quad \overbrace{\bar{r}_{i_0} s^+_{(i_0,i_1]}}^{z_1}, \\ u_1 y' \quad \underbrace{\bar{r}'_{i_1} s^-_{(i_0,i_1]}}_{w_2} &= x' \quad \underbrace{\bar{r}'_{i_0} s^+_{(i_0,i_1]}}_{z_2}. \end{aligned}$$

Hence, by Lemma 4.1, the pair $\{x, x'\}$ is different from the pair $\{y, y'\}$.

The dependency graph contains at least 6 distinct edges, as shown in Figure 4.6, $\mathrm{card}(X) + 4$ vertices and, most importantly, at most $\mathrm{card}(X) - 2$ components. Hence, we can bound the combinatorial rank of X using the Graph Lemma:

$$\mathrm{rank}_\mathrm{c}(X) \leq \mathrm{rank}_\mathrm{c}(X \cup \{u_1, u_2, s_1, s_2\}) \leq \mathrm{card}(X) - 2\,.$$

Example 4.9 (continued). We have that $x = \beta$, $x' = \gamma$, $y = \delta$ and $y' = \alpha$, so $\{x, x'\} \neq \{y, y'\}$. By Equations (4.3), (4.6) and (4.7) the set X satisfies the following 6 equalities:

$$\begin{aligned} \alpha\beta s_1 &= aa.baab.a &&= a.aba.aba &&= s_1\delta\delta, \\ \delta\delta s_2 &= aba.aba.ab &&= ab.aa.baab &&= s_2\alpha\beta, \\ u_1\delta &= ba.aba &&= baab.a &&= \beta s_1, \\ u_1\alpha\delta &= ba.aa.aba &&= baaaab.a &&= \gamma s_1, \\ u_2\beta &= a.baab &&= aba.ab &&= \delta s_2, \\ u_2\alpha\beta &= a.aa.baab &&= aa.aba.ab &&= \alpha\delta s_2. \end{aligned}$$

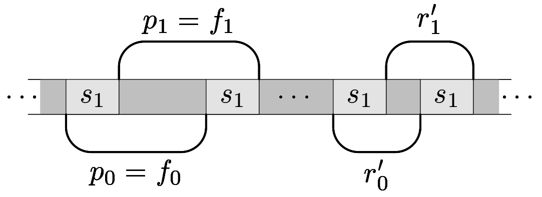
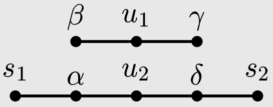


Figure 4.7: An illustration of Equations (4.11) and (4.12).

Hence, the dependency graph



has 2 components, so by the Graph Lemma, $\mathrm{rank}_c(X) \leq 2$.

Let us consider the remaining case.

Case (iii). Let us recall Equations (4.8), where we set $r_{i_2} = 1$:

$$u_1u_2s^-_{(i_0,i_2]} = r_{i_0}s^+_{(i_0,i_2]}, \qquad u_1u_2r'_{i_2}s^-_{(i_0,i_2]} = r'_{i_0}s^+_{(i_0,i_2]}. \tag{4.9}$$

Assume that $r^{(\prime)}_{i_0}$ starts with $x^{(\prime)} \in X$, where again x must be different from x'. Note that x and x' are connected in the dependency graph through u_1. If $|x| \leq |u_1u_2|$ then x and x' are left comparable, a contradiction to the prefix property of X. Thus, we have that $u_1u_2 < x$, which implies that $s^-_{(i_0,i_2]} \neq 1$, and also $i_1 < i_2 < i_0$. Therefore, $i_0 = 2$, $i_1 = 0$ and $i_2 = 1$. Equations (4.9) simplify to

$$u_1u_2s_2 = r_2, \qquad u_1u_2r'_1s_2 = r'_2, \tag{4.10}$$

where r_2 and r'_2 start with x and x', respectively.

Since $r_0 = r_1 = 1$, Equation (4.3) for $a = 0$, $b = 1$ implies

$$s_1p_1 = p_0s_1\,. \tag{4.11}$$

Again since $r_0 = r_1 = 1$, Equation (4.7) without primes gives $u_2 = s_1$. Hence, Equation (4.7) with primes simplifies to

$$s_1r'_1 = r'_0s_1\,. \tag{4.12}$$

Figure 4.7 illustrates the parts of factorizations F_{i_1} and F_{i_2} corresponding to Equations (4.11) and (4.12).

Let us analyze Equations (4.11) and (4.12). For $j = 0, 1$, let $\tilde{p}_j$ be the maximal common prefix of $p_j = f_j$ and r'_j over the alphabet X and let $p_j = \tilde{p}_j\tilde{r}_{p_j}$ and $r'_j = \tilde{p}_j\tilde{r}_{r'_j}$, for some $\tilde{p}_j, \tilde{r}_{p_j}, \tilde{r}_{r'_j} \in X^*$. There are two possibilities:

(a) $\tilde{p}_0$ ends later than $\tilde{p}_1$, *i.e.*, $\tilde{p}_0 = s_1\tilde{p}_1u$ for some $u \in \Sigma^+$;

(b) $\tilde{p}_1$ ends later than $\tilde{p}_0$, *i.e.*, $\tilde{p}_0u = s_1\tilde{p}_1$ for some $u \in \Sigma^+$.

Since the factorizations F_0 and F_1 are disjoint, u must be non-empty.

Case (iii.a). Equations (4.11) and (4.12) imply

$$u\tilde{r}_{p_0}s_1 = \tilde{r}_{p_1}, \qquad u\tilde{r}_{r'_0}s_1 = \tilde{r}_{r'_1}.$$

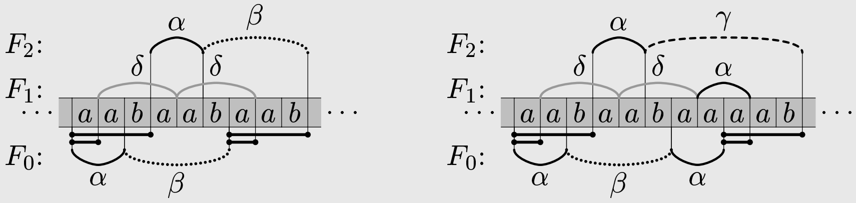
Note that both, $\tilde{r}_{p_1}$ and $\tilde{r}_{r'_1}$, are non-empty. Hence, we can assume that $\tilde{r}_{p_1}$ and $\tilde{r}_{r'_1}$ start with different symbols y and y', and so, y and y' are connected in the dependency graph through u. It is enough to show that the pair $\{y, y'\}$ is different from the pair $\{x, x'\}$, since after that the end of the proof is essentially the same as the one in the case (ii). Using Equations (4.3) and (4.10) we derive

$$\begin{aligned} \overbrace{u_1u_2\tilde{p}_1}^{v}\,\tilde{r}_{r'_1}s_2 &\overset{(4.10)}{=} && && r'_2\,, \\ \underbrace{u_1u_2\tilde{p}_1}_{v}\,\tilde{r}_{p_1}s_2 &= u_1u_2f_1s_2 &&\overset{(4.3)}{=} u_1u_2s_2f_2 &&\overset{(4.10)}{=} r_2f_2\,. \end{aligned}$$

Setting in Lemma 4.1, $v = u_1u_2\tilde{p}_1$ we obtain $\{y, y'\} \neq \{x, x'\}$.

Example 4.10. Consider again the set $X = \{\alpha, \beta, \gamma, \delta\}$ from Example 4.3. Recall that $\alpha = aa$, $\beta = baab$, $\gamma = baaaab$ and $\delta = aba$. Take any non-periodic word w_3 in the set $\{aabaab, aabaabaa\}^{\mathbb{Z}}$. It has 3 different X-factorizations in the sets: $\{\alpha\beta, \alpha\beta\alpha\}^{\mathbb{Z}}$, $\{\delta\delta, \delta\delta\alpha\}^{\mathbb{Z}}$ and $\{\alpha\beta, \alpha\gamma\}^{\mathbb{Z}}$. Note that this example is equivalent to Example 4.4, we have only changed order of X-factorizations.

The pieces of all 3 factorizations of w_3 can be illustrated as follows:


α
β
F2:
δ
δ
F1:
· · · a a b a a b a a b · · ·
F0:
α
β
α
γ
F2:
δ
δ
α
F1:
· · · a a b a a b a a a a b · · ·
F0:
α
β
α

We have two different (a, aab, a, aab)-triples $(\alpha\beta, \delta\delta, \alpha\beta)$ and $(\alpha\beta\alpha, \delta\delta\alpha, \alpha\gamma)$, with $s_1 = a$, $s_2 = ab$ and

$$\begin{array}{lll} p_0 = \alpha\beta, & r_0 = 1, & r'_0 = \alpha\,, \\ p_1 = \delta\delta, & r_1 = 1, & r'_1 = \alpha\,, \\ p_2 = \alpha, & r_2 = \beta, & r'_2 = \gamma\,. \end{array}$$

The order of p_i's is $i_0 = 2$, $i_1 = 0$ and $i_2 = 1$ with $u_1 = b$ and $u_2 = a = s_1$. This is the case (iii). It is easy to check that the set X satisfies Equations (4.11) and (4.12):

$$s_1\delta\delta = \alpha\beta s_1, \qquad s_1\alpha = \alpha s_1\,.$$

Next, we have

$$\begin{aligned} \tilde{p}_0 &= \alpha, & \tilde{r}_{p_0} &= \beta, & \tilde{r}_{r'_0} &= 1\,, \\ \tilde{p}_1 &= 1, & \tilde{r}_{p_1} &= \delta\delta, & \tilde{r}_{r'_1} &= \alpha\,. \end{aligned}$$

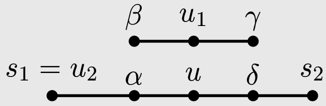
Hence, $|s_1\tilde{p}_1| < |\tilde{p}_0|$, *i.e.*, $\tilde{p}_0$ ends in w_3 later than $\tilde{p}_1$, so we are in the case (a) with $u = a$. Equations (4.11) and (4.12) imply

$$u\beta s_1 = \delta\delta, \qquad us_1 = \alpha,$$

hence, α and δ are connected through u in the dependency graph of the set X. On other hand Equations (4.10)

$$u_1u_2s_2 = \beta, \qquad u_1u_2\alpha s_2 = \gamma$$

imply that β and γ are connected through u_1. Hence again, the dependency graph has 2 components:



Finally, we arrive to the last case of the proof.

Case (iii.b). The proof is similar as in the previous case. Equations (4.11) and (4.12) imply

$$\tilde{r}_{p_0}s_1 = u\tilde{r}_{p_1}, \qquad \tilde{r}_{r'_0}s_1 = u\tilde{r}_{r'_1}\,. \tag{4.13}$$

Equation (4.3) implies that s_2 is a prefix of $f_1 = p_1 = \tilde{p}_1\tilde{r}_{p_1}$. We will show that x and x' are left comparable, and hence X is not a prefix.

First, assume that $s_2 < \tilde{p}_1$. Since, by (4.10), $x < r_2 = u_1u_2s_2$, it is enough to show that x' and $u_1u_2s_2$ are left comparable. Again by Equation (4.10) we have that $x' < r'_2 = u_1u_2r'_1s_2 = u_1u_2\tilde{p}_1\tilde{r}_{r'_1}s_2$. And since $s_2 < \tilde{p}_1$, x' and $u_1u_2s_2$ are left comparable.

Second, assume that $|\tilde{p}_1| < |s_2|$, *i.e.*, $s_2 = \tilde{p}_1\omega$ for some $\omega \in \Sigma^+$. By the definition of X-differences, we have that $|\tilde{p}_0\tilde{r}_{p_0}| = |p_0| = |f_0| = |f_2| > |s_1s_2|$. On the other hand, $|\tilde{p}_0| < |\tilde{p}_0u| = |s_1\tilde{p}_1| < |s_1s_2|$. Therefore, the word $\tilde{r}_{p_0}$ must be non-empty. If also $\tilde{r}_{r'_0} \neq 1$ then we can proceed as in the case (a) choosing y and y' to be the starting symbols of $\tilde{r}_{p_0}$ and $\tilde{r}_{r'_0}$ over the alphabet X.

Unfortunately, it can also happen that $\tilde{r}_{r'_0} = 1$. Let us consider this case. We have that $r'_0 = \tilde{p}_0$. The word $\tilde{p}_1$ in r'_1 depicted in Figure 4.7 should end after

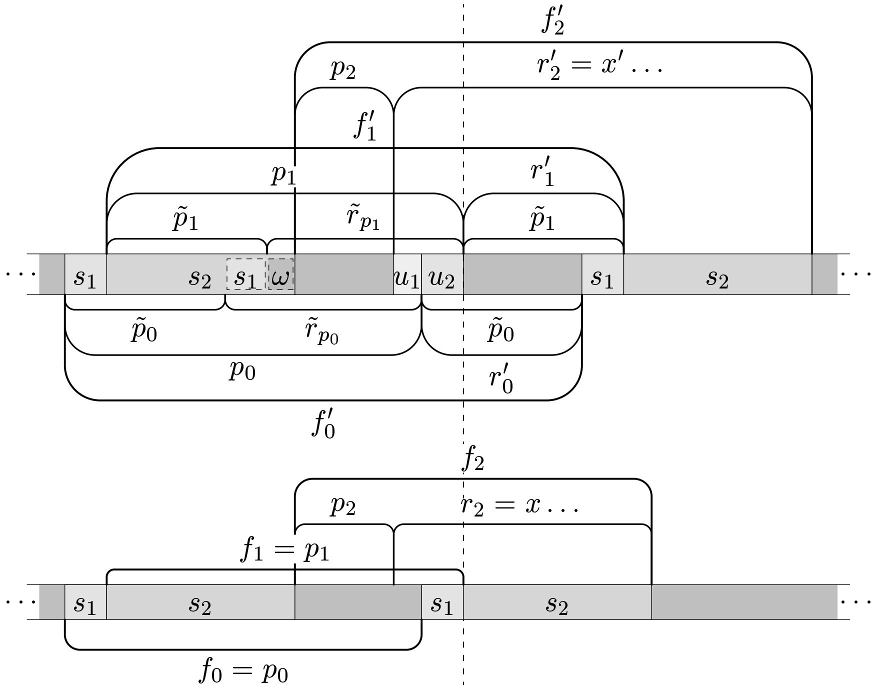


Figure 4.8: An illustration of triples (f_0, f_1, f_2) (down) and (f_0', f_1', f_2') (up) in the case (iii.b) when $\tilde{r}_{r_0'}$ is the empty word. Note that triples are same to the left from the dashed line, but they can differ to the right from it, and that $u_2 = s_1$.

the beginning of s_1. By the definition of X-differences, this is possible only if $r_1' = \tilde{p}_1$ and $\tilde{r}_{r_1'} = 1$. The second of Equations (4.13) implies $u = s_1$, so we have

$$s_1\tilde{p}_1 = \tilde{p}_0 s_1, \text{ by (4.12)}, \qquad s_1\tilde{r}_{p_1} = \tilde{r}_{p_0} s_1, \text{ by (4.13)}.$$

Thus, we can perform the same kind of analysis as we did for Equations (4.11) and (4.12). Then either $\tilde{r}_{p_0} < \tilde{p}_0$, or $\tilde{p}_0 < \tilde{r}_{p_0}$, or the words $\tilde{p}_0$ and $\tilde{r}_{p_0}$ are not left comparable. In the first case, we will show that x and x' are left comparable, as we wanted. In the second case, we obtain again equations of type (4.11) and (4.12), and we can continue inductively. Since the words in the new equations are shorter, we have to arrive to one of the other two cases after a final number of steps. In the third case, we can take the starting symbols of $\tilde{p}_0$ and $\tilde{r}_{p_0}$ over the alphabet X for the values of y and y' and proceed as in the case (a).

Hence, consider the first case $\tilde{r}_{p_0} < \tilde{p}_0$. Multiplying this relation by s_1 and using Equations (4.13) we obtain $\tilde{r}_{p_1} < \tilde{p}_1$. Since $\tilde{p}_1\omega = s_2 < f_1 = \tilde{p}_1\tilde{r}_{p_1}$, we have that $\omega < \tilde{r}_{p_1} < \tilde{p}_1$. Hence:

$$u_1u_2s_2 = u_1u_2\tilde{p}_1\omega < u_1u_2\tilde{p}_1\tilde{p}_1 < u_1u_2\tilde{p}_1\tilde{p}_1\omega = u_1u_2\tilde{p}_1 s_2 = u_1u_2r_1's_2\ ,$$

which together with Equation (4.10) gives that x' and $u_1u_2s_2$ are left comparable. This is, as we have seen before, a contradiction. □

The proof of Lemma 4.4 can be done in the same way as the proof above. The existence of a (t_1, t_2, t'_1, t'_2)-triple $(\bar{f}_0, \bar{f}_1, \bar{f}_2)$ implies that the set $X \cup \{t'_1, t'_2\}$ satisfies the following relations

$$\bar{f}_0 t'_1 = t'_1 \bar{f}_1 \qquad \text{and} \qquad \bar{f}_0 t'_2 = t'_2 \bar{f}_2 \, .$$

Hence, in the dependency graph of any set containing $X \cup \{t'_1, t'_2\}$ the words t'_1 and t'_2 are connected to some elements of X.

Now, let us consider the minimal (t_1, t_2, t_1, t_2)-triple (f_0, f_1, f_2) and the minimal (t_1, t_2, t'_1, t'_2)-triple (f'_0, f'_1, f'_2) instead of two different minimal (t_1, t_2, t_1, t_2)-triples, *cf.* the beginning of the proof of Lemma 4.3. We can follow the lines of the proof of Lemma 4.3 just changing t_1 to t'_1 (resp. t_2 to t'_2) at the ends of some equations. Any time we use Lemma 2.3 (the Graph Lemma) to show a defect effect by 2 we have to add the words t'_1 and t'_2 to the set of vertices of the dependency graph. But, as we have mentioned above, these two vertices are connected to some elements of X, hence Lemma 2.3 can be applied to force the same defect effect.

4.1.2 Proof of Theorem 4.2

Proof. Since X is a prefix set, we can assume that all three X-factorizations are pairwise disjoint. Lemmas 4.3 and 4.4 imply that if we assume that the combinatorial rank of X is at least $\mathrm{card}(X) - 1$ then any two minimal (t_1, t_2, t_1, t_2)-triples are equal, and there cannot occur all following three triples in w: a (t_1, t_2, t_1, t_2)-triple without any occurrence of X-difference (t'_1, t'_2) inside, a (t_1, t_2, t'_1, t'_2)-triple and a (t'_1, t'_2, t'_1, t'_2)-triple. Since t_1, t_2 are suffixes of words in X, there are only finitely many different X-differences. By the pigeon hole principle, there exists an X-difference (t'_1, t'_2), which occurs an infinite number of times in the word w. Each two consecutive occurrences define the minimal (t'_1, t'_2, t'_1, t'_2)-triple. If there are infinitely many occurrences to the right and also to the left from an arbitrary position in w then, clearly, all three X-factorizations and the bi-infinite word w are periodic, which is a contradiction.

Hence, without loss of generality we assume that there are occurrences of (t'_1, t'_2, t'_1, t'_2)-triple only to the right from a position n. Therefore, by the pigeon hole principle, there must be an X-difference (t_1, t_2) occurring an infinite number of times to the left from the position n in w. Clearly, a (t_1, t_2, t_1, t_2)-triple occurring to the left from the position n in w does not contain any occurrence of X-difference (t'_1, t'_2). Obviously, there is a (t_1, t_2, t'_1, t'_2)-triple in the word w, which is a contradiction. □

Note that in the proof of Theorem 4.2 we have used the fact that at least 1 of the X-factorizations is non-periodic, and not the fact that the bi-infinite word w is non-periodic. Hence, we have proved the following theorem:

Theorem 4.5. *Consider a prefix set $X \subseteq \Sigma^+$. Let w be a bi-infinite word over Σ with 3 disjoint X-factorizations F_0, F_1, F_2. If at least one of three X-factorizations is non-periodic then the combinatorial rank of X is at most* $\operatorname{card}(X) - 2$.

Nevertheless, under assumption that X is a prefix, this theorem is equivalent to Theorem 4.2. The following example shows that in Theorem 4.5, but not in Theorem 4.2, we have to put some assumptions on the set X, for example, that it is a code.

Example 4.11. Let $X = \{\alpha, \beta, \gamma\}$, where $\alpha = ababa$, $\beta = b$ and $\gamma = ababab$. Then the periodic bi-infinite word $w = (ab)^{\mathbb{Z}}$ has three disjoint X-factorizations of the form $\{\alpha\beta, \gamma\}^{\mathbb{Z}}$. We can choose them to be non-periodic and not shift-equivalent. The combinatorial rank of the set X is 2, so in this case the defect effect is only by 1.

4.2 A connection to the Critical Factorization Theorem

In this section we look at how the Critical Factorization Theorem is connected to Problem 4.1 in the cases $k = \mathrm{card}(X)$ and $k = \mathrm{card}(X) + 1$.

First, we will recall the Critical Factorization Theorem and its application, *cf.* Chapter 8 in [Lo]. We need a few definitions.

Let w be a finite word. We say that an integer $p \geq 1$ is *a local period* of w at the position (w_1, w_2), if there is a word z of the length p such that (z, z) matches w at the position (w_1, w_2). The minimal local period of w at the position (w_1, w_2) is called *the local period* of w at the position $|w_1|$, denoted by $\mathrm{lper}(w, |w_1|)$. Further, we say that the position $1 < i < |w|$ in w is *critical*, if $\mathrm{lper}(w, i) = \mathrm{per}(w)$.

Theorem 4.6 (Critical Factorization Theorem). *For any $w \in \Sigma^*$ with the period $p(w) > 1$ every sequence of $p(w) - 1$ consecutive positions contains a critical one.*

Let X be a set of non-empty words and let sequences of elements of X, $x_1, \ldots, x_n$ and $x'_1, \ldots, x'_m$, be two X-interpretations of w, *i.e.,* there are words p, s, p', s' such that

$$pws = x_1 \ldots x_n \quad \text{and}$$
$$p'ws' = x'_1 \ldots x'_m \,.$$

We say that X-interpretations $x_1, \ldots, x_n$ and $x'_1, \ldots, x'_m$ are *disjoint*, if for all integers $i \leq n$ and $j \leq m$, we have that $p^{-1}x_1 \ldots x_i \neq {p'}^{-1}x'_1 \ldots x'_j$.

The application of the Critical Factorization Theorem states, *cf.* [Lo]:

Proposition 4.7. *Let $w \in \Sigma^+$ and $X \subseteq \Sigma^+$ be a finite set satisfying* $\mathrm{per}(x) < \mathrm{per}(w)$ *for all $x \in X$. Then w has at most* $\mathrm{card}(X)$ *disjoint X-interpretations.*

Already in [Lo] it was noted that the bound in the proposition is close to the optimal: for each $n \geq 2$, words of the form $w \in (a^{2n-2}b)^+$ have exactly n disjoint X-interpretations where $X = \{a^n, b, aba, \ldots, a^{n-1}ba^{n-1}\}$ contains $n+1$ elements.

In [Lo] it was also conjectured that the exact value in Proposition 4.7 is $\mathrm{card}(X) - 1$. This conjecture is inspired by Schützenberger conjecture stated in [Sc] which in addition assumes that the set X satisfies an another algebraic property. It is out of scope of this work to restate the original conjecture, and so an interested reader is referred to [Sc, Pe2] for details. If the conjecture in [Lo] would be true then it would imply that a non-periodic bi-infinite word can possess at most $\mathrm{card}(X) - 1$ disjoint X-factorizations, which is also an immediate consequence of an affirmative answer to our Problem 4.1 in the case

$k = \mathrm{card}(X)$. However, the following examples show that the conjecture is false, and hence also, that the bound in Proposition 4.7 is optimal.

Example 4.12. For any integer $n \geq 2$ consider the set $X = \{\alpha_1, \dots, \alpha_{n-1}, \beta\}$ with $\alpha_i = a^iba^i$ and $\beta = ba^nb$. Note that $\mathrm{card}(X) = n$ and that $\mathrm{per}(\alpha_i) = i+1$ and $\mathrm{per}(\beta) = n+1$. The word $w = ba^nba^{n-1}b$ has the period $\mathrm{per}(w) = 2n+1$, hence Proposition 4.7 implies that w has at most n disjoint X-interpretations. The following list of X-interpretations of w shows that the word w has exactly $n = \mathrm{card}(X)$ of X-interpretations.

$$\begin{aligned}
\alpha_1\alpha_{n-1}\beta &= \quad a\,\overbrace{ba.a^{n-1}ba^{n-1}.b}^{w}\,a^nb \\
\alpha_2\alpha_{n-2}\alpha_1 &= \quad a^2\,\overbrace{ba^2.a^{n-2}ba^{n-2}.ab}^{w}\,a \\
&\dots \\
\alpha_i\alpha_{n-i}\alpha_{i-1} &= \quad a^i\,\overbrace{ba^i.a^{n-i}ba^{n-i}.a^{i-1}b}^{w}\,a^{i-1} \\
&\dots \\
\alpha_{n-1}\alpha_1\alpha_{n-2} &= a^{n-1}\,\overbrace{ba^{n-1}.aba.a^{n-2}b}^{w}\,a^{n-2} \\
\beta\alpha_{n-1} &= \quad \overbrace{ba^nb.a^{n-1}b}^{w}\,a^{n-1}
\end{aligned}$$

Hence, Proposition 4.7 is optimal.

Note that the X-interpretations above are parts of n disjoint shift-equivalent X-factorizations in

$$[\beta\alpha_{n-1}\alpha_1\alpha_{n-2}\alpha_2 \dots \alpha_1\alpha_{n-1}]_w^{\mathbb{Z}}$$

of the periodic bi-infinite word $w = (ba^nba^{n-1})^{\mathbb{Z}}$.

The above example shows that the conjecture stated in [Lo] is not valid. An interesting problem is to find the sufficient conditions in terms of notions of Combinatorics on Words such that the conjecture would turn true. Next, we will show that that two straightforward approaches in this direction do not give the desired result.

Note that in Example 4.12 the maximal ratio between the periods of elements of the set X and the period of the word w is $\frac{n+1}{2n+1} > \frac{1}{2}$. One possible additional assumption of the conjecture could be restricting the ratio above to $1/2$ or smaller. Unfortunately, the following example shows that the maximal ratio between the periods can be arbitrary small without decreasing the number of X-interpretations.

Example 4.13. Take an arbitrary integers $n \geq 1$ and $p \geq 1$. Consider the set $X = \{\alpha_i; \quad i = 1, \dots, n\}$ with $\alpha_i = a^i b(a^n b)^{p-1} a^i$ and the word $w = ((a^n b)^p a)^2$. Obviously,

$$\text{for every } i = 1, \dots, n \quad \operatorname{per}(\alpha_i) = n + 1\,, \quad \text{and} \quad \operatorname{per}(w) = p(n+1) + 1\,.$$

Thus, the ratio between the period of any element of X and the period of w is $\frac{n+1}{p(n+1)+1} < \frac{1}{p}$. On the other hand, for every $i = 0, \dots, n-1$ we have

$$w = a^i \alpha_{n-i} \alpha_{i+1} a^{-i}\,,$$

and thus, the word w had $n = \operatorname{card}(X)$ disjoint X-interpretations

$$\alpha_n \alpha_1, \alpha_2 \alpha_{n-1} \alpha_2, \dots, \alpha_{n-1} \alpha_2 \alpha_{n-1}, \alpha_n \alpha_1 \alpha_n\,.$$

The all examples above the combinatorial rank of the set X is 2. Therefore, we can ask if a condition on the combinatorial rank could help.

Example 4.14. Consider any set X over the alphabet Σ such that a word w has $\operatorname{card}(X)$ disjoint X-interpretations and the periods of elements of X are strictly smaller than the period of w. Let c be a new letter and let g and h be two morphisms from Σ to $\Sigma \cup \{c\}$ defined as follows: for every $a \in \Sigma$

$$g(a) = ac \qquad \text{and} \qquad h(a) = ca\,.$$

Let $\bar{X} = g(X) \cup h(X)$, $\bar{w} = g(w)$. Obviously, the word $\bar{w}$ has $\operatorname{card}(\bar{X}) = 2\operatorname{card}(X)$ disjoint X-interpretations. Moreover, $\operatorname{per}(\bar{w}) = 2\operatorname{per}(w)$ and, for every $x \in X$, $\operatorname{per}(g(x)) = \operatorname{per}(h(x)) = 2\operatorname{per}(x)$. Finally, $\operatorname{rank_c}(\bar{X}) = \operatorname{rank_c}(X) + 1$.

Example 4.14 shows that putting a condition on the combinatorial rank of X in the form of a lower bound by a constant is not sufficient. As an open problem remains the question if setting $\operatorname{rank_c}(X)$ close to $\operatorname{card}(X)$ would fix the problem.

Note that a more trivial counterexample to the problem stated in [Lo] was pointed out already in [Ha]: the word $w = ba^{n+1}b$ has n X-interpretations, where $X = \{a^i b a^{n+1-i}; \quad i = 1, \dots, n\}$. However, the maximal period of words in X, $n+1$, is almost equal to the period of w, $n+2$.

Finally, let us come back to Problem 4.1 posed in the beginning of this chapter. As a corollary of Proposition 4.7 we have the following result.

Corollary 4.8. *Let X be a set of non-empty words and w a non-periodic bi-infinite word. Then w can possess at most* $\operatorname{card}(X)$ *disjoint X-factorizations.*

Proof. Let $w_i \in \Sigma$, $i \in \mathbb{Z}$ be letters of the bi-infinite word w:

$$w = \ldots w_{-2} w_{-1} w_0 w_1 w_2 \ldots .$$

We define the sequence $\{u_i\}_{i \geq 0}$ of finite words as follows

$$u_i = w_{-i} \ldots w_{-1} w_0 w_1 \ldots w_i .$$

Clearly, $\mathrm{per}(u_{i+1}) \geq \mathrm{per}(u_i)$. This implies that the sequence $\{\mathrm{per}(u_i)\}_{i \geq 0}$ is non-decreasing. Assume that it is upper bounded, *i.e.,* there are positive integers j, p such that, for all $i \geq j$, $\mathrm{per}(u_i) = p$. Then the bi-infinite word w is periodic with a period p, which is a contradiction. Thus, there exists a positive integer j such that $\mathrm{per}(u_j)$ is greater than the periods of words in X. Assume that w possesses $\mathrm{card}(X) + 1$ disjoint X-factorizations. We can construct, in a natural way, $\mathrm{card}(X) + 1$ disjoint X-interpretations of the word u_j. But this together with Proposition 4.7 yields a contradiction. □

Chapter 5

Conjugacy of binary sets

A natural extension of word equations are language equations. Despite the fact that there is a rather rich theory on word equations, almost nothing is known about that of languages. Even the simples equation, *i.e.,* the commutation equation $XZ = ZX$ for languages, which has been recently studied in a number of papers, is solved only in some special cases, for example when $\mathrm{card}(X) \leq 3$ or when X is a code, *cf.* [CKO, Ra, KPe, HP, Ka3, KLP]. In all these cases Z must be of the form $Z = \cup_{i \in I} \varrho(X)^i$ with $I \subseteq \mathbb{N}$, and $\varrho(X)$ being the primitive root of X, *i.e.,* the minimal set having the set X as its power. Hence, we have that both sets X and Z are expressible in terms of one set using operations: the concatenation "$.$", the Kleene star "*" and the union "$\cup$". This can be viewed as a defect effect for languages. However, as an example of [CKO] shows, in general, a relation on languages, or even a commutation equation, do not always cause a defect effect:

Example 5.1. Consider the sets $X = \{a, aaa, b, ba, ab, aba\}$ and $Z = X \cup \{aa\}$. X and Z satisfy the commutation equation $XZ = ZX$, but they cannot be expressed as unions of powers of a common set.

In this chapter we study the conjugacy equation

$$XZ = ZY\,. \tag{5.1}$$

for languages. Since the commutation equation is a special case of the conjugacy equation, the conjugacy equation for languages cannot be easy. Therefore, we will study the conjugacy equation in the case when the sets X and Y are binary. Surprisingly, even in this very restricted case we cannot witness a defect effect.

Example 5.2. Consider the binary sets $X = \{ab, abaca\}$ and $Y = \{ba, caaba\}$, and the singleton set $Z = \{aba\}$. The sets X, Y and Z satisfy the conjugacy equation $XZ = ZY$, but they cannot be expressed in terms of two sets using the operations "$.$", "*" and "$\cup$".

In the word case the conjugacy equation has well-known solution and actually, as we have seen in Section 2.5, it is one of the several characterizations of two words which are conjugates. For languages we take (5.1) as a definition of conjugacy. We say that languages X and Y are *conjugates*, in symbols $X \sim Y$, if there exists a non-empty set Z such that (X, Y, Z) is a solution of (5.1). If this is the case we also say that X and Y are *conjugated via* Z, and we write $X \sim_Z Y$. Note that stricter definition of conjugacy of codes, corresponding to what we call word type solutions, was studied in [Pe1].

5.1 General considerations

In this section we will study some properties of the conjugacy equation in the general setting. We show that in special cases Equation (5.1) has only, so-called, word type solutions, while in general also other solutions are possible even for a unary set Z.

Let us recall that the conjugacy equation $xz = zy$ for non-empty words has a well known general solution, *cf.* Lemma 2.10:

$$\exists p, q \in \Sigma^* \text{ such that } x = pq,\ y = qp \text{ and } z \in (pq)^*p.$$

As is immediate to check the words can be replaced by languages (or finite languages) to obtain solutions of the conjugacy equation $XZ = ZY$ for languages: triples

$$X = PQ, \quad Y = QP, \quad \text{and} \quad Z = \bigcup_{i \in I} (PQ)^i P \tag{5.2}$$

for $P, Q \subseteq \Sigma^*$ and $I \subseteq \mathbb{N}$, are solutions. They are referred to as *word type solutions*. The conjugacy equation (5.1) has always word type solutions. In some cases these are the only possible solutions. For example, if the sets X, Y and Z are prefix codes, or if the sets X and Y are uniform, *i.e.,* consist of words of a fixed length, then Equation (5.1) has only word type solutions. This follows from the fact that the monoids of prefix codes, *cf.* [Pe1], and of uniform non-empty languages are free. Consequently, we formulate:

Proposition 5.1. *Consider prefix codes X, Y and Z such that $X, Y \neq \{1\}$. If the sets satisfy the conjugacy equation* (5.1) *then there exist prefix codes $P, Q \subseteq \Sigma^*$ and an integer $i \in \mathbb{N}$ such that $X = PQ$, $Y = QP$ and $Z = (PQ)^i P$.*

If we assume that the sets X and Y are uniform, we can decompose the set Z into uniform subsets, and clearly, (X, Y) is a solution of (5.1) for each such subset of the set Z as well. Therefore, we have the following proposition:

Proposition 5.2. *If sets X, Y and Z satisfy the conjugacy equation* (5.1) *and $X, Y \neq \{1\}$ are uniform then there exist uniform sets $P, Q \subseteq \Sigma^*$ and $I \subseteq \mathbb{N}$ such that $X = PQ$, $Y = QP$ and $Z = \cup_{i \in I} (PQ)^i P$.*

However, not all solutions are of the word type, even for a unary set Z:

Example 5.3. $Z = \{aa\}$, $X = Y = \{a, aaa\}$ is a solution of the conjugacy equation (5.1), which is not of the word type. However, this is not a minimal solution, since (X, Y) can be obtained as a union of two solutions $X_1 = Y_1 = \{a\}$ and $X_2 = Y_2 = \{aaa\}$ while keeping $Z = \{aa\}$. These "smaller" solutions are of the word type.

An example of a binary prefix code Z, which allows a minimal solution not being of the word type, is as follows:

Example 5.4. $Z = \{a, ba\}$, $X = \{a, ab, abb, ba, babb\}$, $Y = \{a, ba, bba, bbba\}$ is a solution of (5.1). This is a minimal solution, but not of the word type. Indeed, the only solutions contained in (X, Y) are: $X_1 = Y_1 = \{a, ba\}$, $X_2 = \{abb, babb\}$ and $Y_2 = \{bba, bbba\}$, and their union which does not form the whole (X, Y). Note that here X and Y are of different cardinality.

Now, let us study the basic properties of the conjugacy equation. Assume that sets X, Y and Z satisfy Equation (5.1). Then necessarily

$$\min_{x\in X} |x| + \min_{z\in Z} |z| = \min_{z\in Z} |z| + \min_{y\in Y} |y|,$$

and therefore also

$$\min_{x\in X} |x| = \min_{y\in Y} |y| . \tag{5.3}$$

Moreover, the sets

$$X_1 = \{x_1 \in X : \quad |x_1| = \min_{x\in X} |x|\} \quad \text{and} \quad Y_1 = \{y_1 \in Y : \quad |y_1| = \min_{y\in Y} |y|\}$$

are conjugated via $Z_1 = \{z_1 \in Z : \quad |z_1| = \min_{z\in Z} |z|\}$.

Further, if $1 \in X$ then $X_1 = Y_1 = \{1\}$, and so $1 \in Y$. Obviously, all languages containing the empty word are conjugated with each other via the set Σ^*. In the sequel, we assume that $1 \notin X$ and $1 \notin Y$, *i.e.*, $X_1, Y_1 \neq \{1\}$. Since all the sets X_1, Y_1 and Z_1 are uniform, by Proposition 5.2, necessarily

$$X_1 = PQ, \quad Y_1 = QP, \quad \text{and} \quad Z_1 = (PQ)^i P$$

for some non-negative integer i and uniform sets P and Q. Hence, we have the following proposition:

Proposition 5.3. *Let $X \sim_Z Y$ with $X, Y \subseteq \Sigma^+$ and Z non-empty. Let X_1 (resp. Y_1, Z_1) be the set of the elements of X (resp. Y, Z) of the minimal length. There exist uniform sets P and Q and an integer $i \geq 0$ such that $X_1 = PQ$, $Y_1 = QP$, and $Z_1 = (PQ)^i P$. In particular, if $|X_1| = 1$ or $|Y_1| = 1$ then P and Q must be singletons and $X_1 = \{(uv)^m\}$, $Y_1 = \{(vu)^m\}$, $Z_1 = \{(uv)^{k_1} u\}$ for some words u and v, where uv is primitive, and some integers $m \geq 1$ and $k_1 \geq 0$.*

We finish this section with the following simple but useful facts.

Proposition 5.4. *If sets X, Y and Z satisfy the conjugacy equation* (5.1) *then*

(i) for every positive integer n, $X^n Z = ZY^n$;

(ii) $Z \subseteq \mathrm{Pref}(X^+) \cap \mathrm{Suff}(Y^+)$.

5.2 Binary sets X and Y

In this and the following sections we will focus our study on the case when the sets X and Y are binary. We will be able to characterize completely all the solutions of the conjugacy equation in this case, *i.e.,* all triples (X, Y, Z) satisfying the conjugacy equation with X and Y binary.

Although, it might seem that the complete characterization in such a simple case is easy, we will need several technical lemmas to accomplish this task. We will divide them to several sections to make our considerations more comprehensive.

In this section we fix our notations, state a simple result which will be used several times later, and in the end, explain the logical structure of the following sections.

Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be binary sets with $|x_1| \leq |x_2|$ and $|y_1| \leq |y_2|$ and let Z be a non-empty set such that $XZ = ZY$. Note that, by (5.3), $|x_1| = |y_1|$.

We divide the set Z into the pairwise disjoint layers:

$$Z = Z_1 \uplus Z_2 \uplus Z_3 \uplus \dots ,$$

where for all $z_1, z_2 \in Z_i$, $|z_1| = |z_2|$ and for all $z_1 \in Z_i, z_2 \in Z_{i+1}$, $|z_1| < |z_2|$. Hence, for instance, the set Z_1 contains all shortest elements of the set Z.

We will need the following lemma which belongs to the folklore of the theory of Combinatorics on Words.

Lemma 5.5. *If a word z satisfies the equation $(uv)^k z = z(vu)^k$ with uv primitive, $v \neq 1$ and $k \geq 1$, then $z \in (uv)^* u$.*

Proof. Let $x = (uv)^k$ and $y = zv$. Then $xy = (uv)^k zv = z(vu)^k v = zv(uv)^k = yx$. The words x and y commute, and therefore have the same primitive root, uv. Let $y = (uv)^l$ for some $l \geq 1$ (note that $|y| \geq |v| > 0$). Then $z = (uv)^{l-1} u \in (uv)^* u$. □

Finally, let us describe the way how we are going to characterize all solutions, *i.e.,* all triples (X, Y, Z), of the conjugacy equations with X and Y binary. First, we will characterize all X and Y such that $XZ = ZY$ for some non-empty set

Z. In Sections 5.3 and 5.4 we will consider separately two main cases: the commutative case and the non-commutative case, depending on whether x_1 and x_2 commute or do not. Further, in the non-commutative case we distinguish two cases: $|x_1| = |x_2|$ and $|x_1| < |x_2|$, which, by Proposition 5.3, coincide with the cases $|y_1| = |y_2|$ and $|y_1| < |y_2|$, respectively. In the commutative case and the non-commutative case with $|x_1| = |x_2|$ we will immediately obtain also the characterization of the sets Z. In the last case the situation is more intricate, therefore, we will deal with this case in a separate section, namely, Section 5.5. In this case, by Lemma 5.10, there exists a word t such that either $Xt = tY$ or $tX = Yt$, where t represents the singleton set $\{t\}$. In Subsections 5.5.1–3 we will determine all possible Z's in terms of X and t satisfying (5.1) and one of the above conditions. Finally, in Section 5.6 we will combine all the results and obtain the characterization of all solutions of the conjugacy equations in the case when the sets X and Y are binary.

5.3 The commutative case

In this section we consider the case when x_1 and x_2 commute.

Lemma 5.6. *Let sets $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$, with $|x_1| \leq |x_2|$ and $|y_1| \leq |y_2|$, be conjugates via a non-empty set Z. If $x_1, x_2 \in t^+$, where t is primitive, then there is a word s such that $y_1, y_2 \in s^+$ and words t and s are conjugates,* i.e., *$t = uv$ and $s = vu$ for some words $u, v \in \Sigma^*$. Moreover, the set Z satisfies that $Z \subseteq (uv)^*u$.*

Proof. Take an arbitrary word $z \in Z$. By Proposition 5.4, for any integers $n > 0$ and $i = 1, 2$ there are integers $i_1, \ldots, i_n \in \{1, 2\}$ and $z' \in Z$ such that

$$zy_i^n = x_{i_1} \ldots x_{i_n} z' \in t^+ z' \,.$$

If we take $n \geq 2$ such that $|z| + 2|y_i| \leq n|x_1|$, then zy_i^2 is a prefix of t^ω. This implies that $z = t^{m_z} u_z$, for some integer $m_z \geq 0$ and some word u_z, a proper prefix of t. Let $t = u_z v_z$ and $s_z = v_z u_z$. Then $z \in (u_z v_z)^* u_z$ and y_i^2 is a prefix of s_z^ω. Note that since t is primitive, so is s_z. Since $|y_2| \geq |y_1| = |x_1| \geq |t| = |s_z|$, by Lemma 2.5, we have that y_1 (resp. y_2) commutes with s_z. Since s_z is primitive, we have $\rho(y_1) = \rho(y_2) = s_z$, *i.e.,* we can conclude that $y_1, y_2 \in s_z^+$.

Now, it suffices to prove that for all $z, \bar{z} \in Z$, $u_z = u_{\bar{z}}$ and $v_z = v_{\bar{z}}$. Since $s_z = \rho(y_1) = s_{\bar{z}}$, we have $u_z v_z u_{\bar{z}} = u_{\bar{z}} v_{\bar{z}} u_{\bar{z}} = u_{\bar{z}} v_z u_z$. Further, since $v_z \neq 1$, by Lemma 5.5, we have $u_{\bar{z}} \in (u_z v_z)^* u_z$, which implies $u_{\bar{z}} = u_z$. We are done. □

5.4 The non-commutative case

In what follows we will assume that x_1 and x_2, and similarly, y_1 and y_2, do not commute. As an immediate consequence of Proposition 5.3 we have that the lengths of x_1 and x_2 are equal if and only if the lengths of y_1 and y_2 do so:

Corollary 5.7. *Let sets $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$, with $|x_1| \le |x_2|$ and $|y_1| \le |y_2|$, be conjugates via a non-empty set Z. Words x_1 and x_2 have the same length, if and only if words y_1 and y_2 have so.*

Now, we will consider the simplest case when the sizes of words in X and Y are equal.

Lemma 5.8. *Let sets $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be conjugates via a non-empty set Z. If $|x_1| = |x_2| = |y_1| = |y_2|$ then there are words u, v and p such that $|u| = |v|$ and a set $I \subseteq \mathbb{N}$ such that one of the following conditions is satisfied:*

(i) $X = \{pu, pv\}$, $Y = \{up, vp\}$ *and* $Z = \bigcup_{i \in I} X^i p$,

(ii) $X = \{up, vp\}$, $Y = \{pu, pv\}$ *and* $Z = \bigcup_{i \in I} X^i \{u, v\}$.

Proof. Notice that the sets X and Y are uniform, so as a consequence of Proposition 5.2 there are sets $P, Q \subseteq \Sigma^*$ and $I \subseteq \mathbb{N}$ such that $X = PQ$, $Y = QP$ and $Z = \cup_{i \in I} (PQ)^i P$. Now, if $\operatorname{card}(X) = 2$ then either $\operatorname{card}(P) = 1$ and $\operatorname{card}(Q) = 2$ (the case *(i)* with $P = \{p\}$ and $Q = \{u, v\}$), or $\operatorname{card}(P) = 2$ and $\operatorname{card}(Q) = 1$ (the case *(ii)* with $P = \{u, v\}$ and $Q = \{p\}$). □

Observation 5.1. Note that in the case *(i)* of Lemma 5.8 we have $Xp = pY$, and similarly, in the case *(ii)* we have $pX = Yp$.

In the case when the lengths of words in X and Y are not all the same we need the following 2 lemmas:

Lemma 5.9. *Let sets $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$, with $|x_1| \le |x_2|$ and $|y_1| \le |y_2|$, be conjugates via a non-empty set Z. If $|x_2| \ne |y_2|$ then x_1 and x_2 commute.*

Proof. We will prove the claim only in the case $|x_2| < |y_2|$. By symmetry, the claim can be proved in the same way also in the case $|x_2| > |y_2|$. Hence, assume that $|x_2| < |y_2|$, and let z_1 be an element of Z of the minimal length. By Proposition 5.4, for any positive integer n, the word $w = x_1^n x_2 z_1$ belongs to the set ZY^{n+1}. Hence, we have $w = z' y_{i_1} \ldots y_{i_{n+1}}$ for some $i_1, \ldots, i_{n+1} \in \{1, 2\}$ and $z' \in Z$. As z_1 was chosen of minimal length, $|z'| \ge |z_1|$; recall also that $|y_{i_j}| \ge |y_1| = |x_1|$ and $|y_2| > |x_2|$. If for any $j \in \{1, \ldots, n+1\}$ we have $i_j = 2$ then

$$|w| = |z'| + |y_{i_1}| + \cdots + |y_{i_{n+1}}| \ge |z'| + n|y_1| + |y_2| > |z_1| + n|x_1| + |x_2|,$$

a contradiction since $|w| = n|x_1| + |x_2| + |z_1|$. Therefore $i_1 = \ldots = i_{n+1} = 1$, *i.e.*, $w = x_1^{n-1}x_1x_2z_1 = z'y_1^{n+1}$. By a similar argument, we obtain $x_1^{n-1}x_2x_1z_1 = z''y_1^{n+1}$ for some $z'' \in Z$. If we take an integer n such that $(n+1)|y_1| \geq |x_1x_2z_1|$, we find that the words $x_1x_2z_1$ and $x_2x_1z_1$ have the same length and are both suffixes of y_1^{n+1}, therefore are equal. Hence, x_1 and x_2 commute. □

Lemma 5.10. *Let sets $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be conjugates via a non-empty set Z. If $|x_1| = |y_1| < |x_2| = |y_2|$ then either x_1 and x_2 commute, or x_2 and y_2 are conjugates. Moreover, in the second case there exists a word t such that either $x_1 \sim_t y_1$ and $x_2 \sim_t y_2$,* i.e., *$Xt = tY$, or $y_1 \sim_t x_1$ and $y_2 \sim_t x_2$,* i.e., *$tX = Yt$.*

Proof. By Proposition 5.3, we know that there exist words u and v and integers k_1 and m such that uv is primitive, $x_1 = (uv)^m$, $y_1 = (vu)^m$ and $Z_1 = \{z_1\}$, where $z_1 = (uv)^{k_1}u$. Note that $x_1 \sim_{(uv)^iu} y_1$ for any $i \geq 0$. We have either $x_2z_1 = z_1y_2$, or $x_2z_1 = z'y_1$, for some $z' \in Z$. In the first case, we have immediately that x_2 and y_2 are conjugates via z_1, and we are done. In the second case, let Z' be the set of words in Z having the same length as z'.

We construct a sequence $\{z^{(i)}\}_{i\geq 1}$ in Z'. Let $z^{(1)} = z'$. For any $i \geq 1$ we have, either $x_1z^{(i)} = z^{(i+1)}y_1$, or $x_1z^{(i)} = z_1y_2$. First, assume that the second case never happens. We have $x_1^iz^{(j)} = z^{(i+j)}y_1^i$ for all $i \geq 1$ and $j \geq 1$. Hence all $z^{(j)}$ are suffixes of y_1^i for some big enough integer i, and therefore they are equal. Then $x_1z' = z'y_1$, and by Lemma 5.5, we have $z' \in (uv)^*u$. Using

$$x_2z_1 = z'y_1, \tag{5.4}$$

we obtain $x_2 \in (uv)^+$, hence x_1 and x_2 commute.

Now, assume that there is a non-negative integer n such that for all $i = 1, \ldots, n$, $x_1z^{(i)} = z^{(i+1)}y_1$ and $x_1z^{(n+1)} = z_1y_2$. These equalities imply that

$$x_1^{n+1}z' = x_1x_1^nz^{(1)} = x_1z^{(n+1)}y_1^n = z_1y_2y_1^n\,. \tag{5.5}$$

Equations (5.4) and (5.5) imply that

$$(uv)^{m(n+1)}x_2(uv)^{k_1}u = x_1^{n+1}x_2z_1 = z_1y_2y_1^{n+1} = (uv)^{k_1}uy_2(vu)^{m(n+1)}\,.$$

Now, if $m(n+1) \leq k_1$ then we have that $x_1 \sim_t y_1$ and $x_2 \sim_t y_2$ for $t = (uv)^{k_1-m(n+1)}u$. Otherwise, $y_1 \sim_t x_1$ and $y_2 \sim_t x_2$ for $t = (vu)^{m(n+1)-k_1-1}v$. In both cases we have that x_2 and y_2 are conjugates. □

5.5 Characterization of the sets Z in the cases $Xt = tY$ and $tX = Yt$.

In the above sections we have proved that if $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ are conjugates, with $|x_1| \leq |x_2|$ and $|y_1| \leq |y_2|$, then either

- x_1 and x_2 commute; or
- the lengths of all elements of X and Y are equal; or
- $|x_1| = |y_1| < |x_2| = |y_2|$.

In the first two case we have also characterized all Z's via which the sets are conjugated, *cf.* Lemmas 5.6 and 5.8. Such a characterization is missing in the last case, *cf.* Lemma 5.10. The goal of this section and its three subsections is to complete Lemma 5.10.

Assume that x_1 and x_2 do not commute and that $|x_1| = |y_1| < |x_2| = |y_2|$. Then, by Lemma 5.10 there is a word t such that either

(a) $Xt = tY$, *i.e.*, $x_1 t = t y_1$ and $x_2 t = t y_2$; or

(b) $tX = Yt$, *i.e.*, $t x_1 = y_1 t$ and $t x_2 = y_2 t$.

Since x_1 and y_1 (resp. x_2 and y_2) are conjugates via the same word t, by Lemma 2.10, we have that

$$\begin{aligned} x_1 = (uv)^m \qquad &\text{and} \qquad y_1 = (vu)^m, \\ x_2 = (pq)^i \qquad &\text{and} \qquad y_2 = (qp)^i, \end{aligned} \tag{5.6}$$

where uv and pq are primitive, $m, i \geq 1$ are integers, and,

$$\begin{aligned} &\text{(a)} \quad t = (uv)^M u = (pq)^N p, \\ &\text{(b)} \quad t = (vu)^M v = (qp)^N q, \end{aligned} \tag{5.7}$$

respectively, for some $M, N \geq 0$.

Before we start analyzing the above two case, let us define the minimal and maximal sets Z via which X and Y are conjugated. Assume that X and Y are conjugates. Let $\mathcal{Z}_{X,Y}$ be the class of all non-empty sets Z such that $X \sim_Z Y$. Obviously, if $X \sim_{Z_1} Y$ and $X \sim_{Z_2} Y$ for some $Z_1, Z_2 \in \mathcal{Z}_{X,Y}$ then also $X \sim_{Z_1 \cup Z_2} Y$. Consequently, there exists the unique *maximal* set $Z_{\text{MAX}} \in \mathcal{Z}_{X,Y}$ such that for all $Z' \in \mathcal{Z}_{X,Y}$, $Z' \subseteq Z_{\text{MAX}}$. Dually to the notion of the maximal Z_{MAX}, we call a set $Z \in \mathcal{Z}_{X,Y}$ *minimal*, if there are no $Z_1, Z_2 \in \mathcal{Z}_{X,Y}$ such that

$$Z_1, Z_2 \subset Z \qquad \text{and} \qquad Z = Z_1 \cup Z_2.$$

Of course, there might be several minimal sets Z. Clearly, all finite sets $Z \in \mathcal{Z}_{X,Y}$ can be expressed as unions of minimal sets Z.

Finally, let us prove one simple lemma for the later use.

Lemma 5.11. *Let $x_1 = (uv)^m$, and let $a, b \geq 0$ be integers. If $x_2 \sim_{(uv)^a u} y_2$ and $x_2 \sim_{(uv)^b u} y_2$ with $a \neq b$ then x_1 and x_2 commute. Similarly, if $x_2 \sim_{(uv)^a u} y_2$ and $y_2 \sim_{(vu)^b v} x_2$ then x_1 and x_2 commute.*

Proof. Without the lost of generality we can assume that $a > b$. We have

$$x_2(uv)^a u = (uv)^a u y_2 = (uv)^{a-b}(uv)^b u y_2 = (uv)^{a-b} x_2 (uv)^b u,$$

hence $x_2(uv)^{a-b} = (uv)^{a-b}x_2$. By the defect theorem x_2 and uv, and hence also, x_1 an x_2, commute.

The proof of the second claim is completely the same. □

5.5.1 The maximal $Z \in \mathcal{Z}_{X,Y}$ in the case $Xt = tY$

Consider binary codes $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ which satisfy the condition $Xt = tY$, and hence, also (5.6) and (5.7a). Let Z_{MAX} be the maximal solution of the equation $XZ = ZY$. It is easy to check that X^*t is a solution of $XZ = ZY$, hence we have that X^*t is a subset of Z_{MAX}.

We will show that $Z_{\text{MAX}} = X^*t$. Assume that it is not the case, and let Z_0 be the set of the shortest elements in $Z_{\text{MAX}} - X^*t$. Now, take an arbitrary element $z \in Z_0$. Similarly, as in the proof of Lemma 5.10 we can built the sequence of elements of Z_{MAX} as follows:

$$\begin{aligned} &z^{(1)} = z, \\ &x_1 z^{(l)} = z^{(l+1)} y_1, \quad \text{for } l = 1, \ldots, n'_z, \\ &x_1 z^{(n'_z+1)} = z' y_2, \quad \text{if } n'_z < \infty, \end{aligned} \tag{5.8}$$

where $n'_z \geq 0$ is an integer or infinity.

Case $n'_z = \infty$ for all $z \in Z_0$. Then, similarly as in the proof of Lemma 5.10, we have that all $z^{(l)}$'s are equal to $z = (uv)^\alpha u$. Since, by assumption, this is true for all $z \in Z_0$, the set Z_0 is a singleton. Note that $\alpha \neq M$, otherwise $z = t \in X^*t$.

We have that either $x_2 z = \bar{z} y_2$, or $x_2 z = \bar{z}^{(1)} y_1$. In the first case, since $|\bar{z}| = |z|$ and Z_0 is a singleton, we have that either $\bar{z} = z = (uv)^\alpha u$, which leads to a contradiction by Lemma 5.11, or $\bar{z} \in X^*t$. Let $\bar{z} = x_{i_1} \ldots x_{i_d} t$, for some integer $d \geq 0$. We have

$$x_2(uv)^{\alpha - M} = x_{i_1} \ldots x_{i_d} x_2 \,.$$

Since $\alpha \neq M$, this is a non-trivial equation. Applying the defect theorem we obtain that x_1 and x_2 commute.

Hence, assume that

$$x_2 z = \bar{z}^{(1)} y_1 \,. \tag{5.9}$$

Again, consider the following sequence of elements of Z_{MAX}:

$$\begin{aligned} x_1\bar{z}^{(l)} &= \bar{z}^{(l+1)}y_1, \quad \text{for } l = 1, \ldots, n, \\ x_1\bar{z}^{(n+1)} &= \bar{z}'y_2, \quad \text{if } n < \infty, \end{aligned} \tag{5.10}$$

where $n \geq 0$ is an integer or infinity.

In the case, $n = \infty$, we have, as above, that $\bar{z}^{(1)} = (uv)^{\beta}u$ for some integer $\beta \geq 0$. Then, by (5.9), $x_2 = (uv)^{m+\beta-\alpha}$, hence x_1 and x_2 commute, a contradiction.

Hence, assume that n is finite. By Equations (5.9) and (5.10) we have

$$x_1^{n+1}x_2z = \bar{z}'y_2y_1^{n+1}. \tag{5.11}$$

Since $|z| = |\bar{z}'|$ and Z_0 is a singleton, we have that either $\bar{z}' = z = (uv)^{\alpha}u$, or $\bar{z}' \in X^*t$. In the first case, if $\alpha < m(n+1)$ then Equation (5.11) implies that $y_2 \sim_{(vu)^{m(n+1)-\alpha-1}v} x_2$, which is a contradiction by Lemma 5.11. On the other hand, if $\alpha \geq m(n+1)$ then $x_2 \sim_{(uv)^{\alpha-m(n+1)}u} y_2$, hence by Lemma 5.11, we have that $\alpha - m(n+1) = M$. But then we can write z in the form: $z = x_1^{n+1}t \in X^*t$, a contradiction.

In the second case, let $\bar{z}' = x_{i_1} \ldots x_{i_d}t$ for some integer $d \geq 0$. By Equation (5.11)

$$x_1^{n+1}x_2(uv)^{\alpha-M} = x_{i_1} \ldots x_{i_d}x_2x_1^{n+1}.$$

If $(uv)^{\alpha-M} \neq x_1^{n+1}$ then this is a non-trivial equation, implying that x_1 and x_2 commute by the defect theorem. If $(uv)^{\alpha-M} = x_1^{n+1}$ then $z = x_1^{n+1}t \in X^*t$, a contradiction.

Case n'_z is finite for an element $z \in Z_0$. By Equation (5.8) we have

$$x_1^{n'_z+1}z = z'y_2y_1^{n'_z}, \tag{5.12}$$

where $|z'| < |z|$, hence $z' \in X^*t$. Let $z' = x_{i_1} \ldots x_{i_d}t$, for some integer $d \geq 0$. Therefore,

$$x_1^{n'_z+1}z = x_{i_1} \ldots x_{i_d}x_2x_1^{n'_z}t\,. \tag{5.13}$$

We can assume that $z \notin (uv)^*u$, otherwise we have that x_1 and x_2 commute by the defect theorem.

One can construct a similar sequence of elements of Z_{MAX} as (5.8) starting from zy_1, instead of x_1z. Since $z \notin (uv)^*u$, only the second case is possible: there is an integer $n''_z \geq 0$ and $z'' \in Z_{\mathrm{MAX}}$ such that, similarly as in (5.12),

$$zy_1^{n''_z+1} = x_1^{n''_z}x_2z''. \tag{5.14}$$

Again, $|z''| < |z|$, hence we can write z'' in the form $z'' = x_{j_1} \ldots x_{j_e} t$, for some integer $e \geq 0$. Putting Equations (5.12) and (5.14) together, we obtain

$$\begin{aligned} z' y_2 y_1^{n'_z + n''_z + 1} &= x_1^{n'_z + n''_z + 1} x_2 z'' , \quad \text{hence,} \\ x_{i_1} \ldots x_{i_d} t y_2 y_1^{n'_z + n''_z + 1} &= x_1^{n'_z + n''_z + 1} x_2 x_{j_1} \ldots x_{j_l} t , \quad \text{hence,} \\ x_{i_1} \ldots x_{i_d} x_2 x_1^{n'_z + n''_z + 1} &= x_1^{n'_z + n''_z + 1} x_2 x_{j_1} \ldots x_{j_l} . \end{aligned}$$

If $x_{i_1} = \cdots = x_{i_{n'_z+1}} = x_1$ then, by Equation (5.13), $z = x_{i_{n'_z+2}} \ldots x_{i_d} x_2 x_1^{n'_z} t \in X^* t$. Otherwise, we have a non-trivial equation, and by the defect theorem, x_1 and x_2 commute. In any case, we arrive to a contradiction.

We have proved the following result:

Claim 5.12. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be binary codes. If $Xt = tY$ then the maximal solution of the conjugacy equation $XZ = ZY$ is $Z_{\text{MAX}} = X^* t$.*

5.5.2 All sets $Z \in \mathcal{Z}_{X,Y}$ in the case $Xt = tY$

Consider binary codes $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ which satisfy the conditions $Xt = tY$, (5.6) and (5.7a). Recall that $\mathcal{Z}_{X,Y}$ is the class of all non-empty sets Z such that $X \sim_Z Y$. We have proved above that the maximal $Z \in \mathcal{Z}_{X,Y}$ is $Z_{\text{MAX}} = X^* t$. Next, we will show that all the minimal sets $Z \in \mathcal{Z}_{X,Y}$ are of the form $X^d t$ for some integer $d \geq 0$, and that all sets $Z \in \mathcal{Z}_{X,Y}$ are unions of the minimal ones. Obviously, all sets $Z \in \mathcal{Z}_{X,Y}$ are subsets of $X^* t$. It is enough to show the following lemma:

Lemma 5.13. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be binary codes such that $Xt = tY$. Let $Z \in \mathcal{Z}_{X,Y}$ be a solution of the conjugacy equation. If $x_{i_1} \ldots x_{i_d} t \in Z$ then $X^d t \subseteq Z$.*

Proof. Assume on the contrary that there is a $\bar{z} = x_{j_1} \ldots x_{j_d} t \notin Z$. Since $X^d Z = Z Y^d$, there exist $y_{k_1}, \ldots, y_{k_d} \in Y$ and $z' \in Z$, such that

$$x_{j_1} \ldots x_{j_d} z = z' y_{k_1} \ldots y_{k_d} .$$

Since, by Claim 5.12, $Z \subseteq X^* t$, we can write z' in the form $z' = x_{l_1} \ldots x_{l_e} t$, for some integer $e \geq 0$. Therefore, we have

$$\begin{aligned} x_{j_1} \ldots x_{j_d} x_{i_1} \ldots x_{i_d} t &= x_{l_1} \ldots x_{l_e} t y_{k_1} \ldots y_{k_d} , \quad \text{hence,} \\ x_{j_1} \ldots x_{j_d} x_{i_1} \ldots x_{i_d} &= x_{l_1} \ldots x_{l_e} x_{k_1} \ldots x_{k_d} . \end{aligned}$$

As the consequence of the defect theorem and the fact that x_1 and x_2 do not commute, we have that the equation above must be trivial, *i.e.*, $d = e$, $x_{j_r} = x_{l_r}$ and $x_{i_r} = x_{k_r}$, for all $r = 1, \ldots, d$. But this is a contradiction, since then $\bar{z} = z' \in Z$. □

Corollary 5.14. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be binary codes such that $Xt = tY$. All minimal sets $Z \in \mathcal{Z}_{X,Y}$ are of the form $X^d t$ for some integer $d \geq 0$, and all sets $Z \in \mathcal{Z}_{X,Y}$ can be expressed in the form $Z = \cup_{d \in I} X^d t$ for some index set I.*

5.5.3 The case $tX = Yt$

Consider binary codes $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ which satisfy the condition $tX = Yt$, (5.6) and (5.7b). Assume that $Z \in \mathcal{Z}_{X,Y}$ is a solution of the conjugacy equation $XZ = ZY$. Then $\tilde{Z} = tZt$ is a solution of the equation $YZ = ZX$. Indeed, we have

$$Y\tilde{Z} = YtZt = tXZt = tZYt = tZtX = \tilde{Z}X\,.$$

Since $\tilde{Z}$ is a solution of the conjugacy equation $YZ = ZX$ satisfying $Yt = tX$, Corollary 5.14 yields that there exists an index set I such that

$$tZt = \tilde{Z} = \bigcup_{d \in I} Y^d t = \bigcup_{d \in I} tX^d\,.$$

This implies that $\cup_{d \in I} X^d = Zt$, and therefore, for every index $d \in I$ and for every word $w \in X^d$, t is a suffix of w, and so $|t| \leq |w|$.

Let τ be the minimal integer such that for every word $w \in X^\tau$, $|w| \geq |t|$. The above implies that the index set I contains only the indexes greater or equal τ, *i.e.,* $I \subseteq \{\tau, \tau + 1, \ldots\}$. Hence, we can write that $\cup_{d \in I} X^{\tau+d} = Zt$, where I is now any index set. Moreover, since $tX^{\tau+d} = Y^{\tau+d}t$, the conditions $Z = \cup_{d \in I} X^{\tau+d} t^{-1}$ and $Zt = \cup_{d \in I} X^{\tau+d}$ are equivalent.

We showed that if Z is a solution of the conjugacy equation $XZ = ZY$, *i.e.,* if $Z \in \mathcal{Z}_{X,Y}$, then it can be expressed in the form $Z = \cup_{d \in I} X^{\tau+d} t^{-1}$. On the other hand, it is easy to check that if Z is in this form then it is a solution of the conjugacy equation $XZ = ZY$. Indeed, assume that $Z = \cup_{d \in I} X^{\tau+d} t^{-1}$. Then $Zt = \cup_{d \in I} X^{\tau+d}$, and we have

$$XZt = X \bigcup_{d \in I} X^{\tau+d} = \bigcup_{d \in I} X^{\tau+d} X = ZtX = ZYt\,,$$

and hence, $XZ = ZY$.

The following claim follows easily:

Claim 5.15. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ be binary codes such that $tX = Yt$. Let τ be the minimal integer such that for every word $w \in X^\tau$, $|w| \geq |t|$. All minimal sets $Z \in \mathcal{Z}_{X,Y}$ are of the form $X^{\tau+d} t^{-1}$ for some integer $d \geq 0$, the maximal $Z \in \mathcal{Z}_{X,Y}$ is $X^\tau X^* t^{-1}$, and all sets $Z \in \mathcal{Z}_{X,Y}$ are of the form $Z = \cup_{d \in I} X^{\tau+d} t^{-1}$, for some index set I.*

5.6 The complete characterization

Combining all lemmas proved above we obtain the following characterization of all solutions of the conjugacy equation in the case when the sets X and Y are binary.

Theorem 5.16. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ with $|x_1| \leq |x_2|$ and $|y_1| \leq |y_2|$ be conjugates via a non-empty set Z,* i.e., *$X \sim_Z Y$. Then at least one of the following conditions holds true:*

(i) $x_1x_2 = x_2x_1$, $y_1y_2 = y_2y_1$, i.e., *words x_1 and x_2 (resp. y_1 and y_2) commute, $|x_1| = |y_1|$, and moreover, the words x_1 and y_1 are conjugates,* i.e., *there are words u and v such that uv is primitive, $x_1 \in (uv)^+$ and $y_1 \in (vu)^+$. Finally, the set Z satisfies $Z \subseteq (uv)^*u$;*

(ii) there exists a word t such that, either $Xt = tY$ and $Z = \cup_{d \in I} X^d t$, for some index set I, or $tX = Yt$ and $Z = \cup_{d \in I} X^{\tau+d} t^{-1}$ for some index set I, where $\tau \geq 0$ is the minimal integer such that for every word $w \in X^\tau$, $|w| \geq |t|$.

Conversely, if the sets $X = \{x_1, x_2\} \subseteq \Sigma^+$, $Y = \{y_1, y_2\} \subseteq \Sigma^+$ and $Z \neq \emptyset$ with $|x_1| \leq |x_2|$ and $|y_1| \leq |y_2|$ satisfy either (i) or (ii), then $X \sim_Z Y$.

Proof. The first part of theorem is a consequence of several claims we have proved above: If $|x_1| = |x_2|$ or $|y_1| = |y_2|$, by Corollary 5.7 and Lemma 5.8, we are in the case (ii) with $t = p$. If $|x_1| < |x_2| = |y_2|$, by Lemma 5.10, we are also in the case (ii). Otherwise, by Lemma 5.9, x_1 and x_2 commute, and, by Lemma 5.6, we are in the case (i). Further, in the case (ii), the characterization of the set Z follows by Observation 5.1, Corollary 5.14 and Claim 5.15.

Conversely, assume that X and Y satisfy one of the above conditions. The case (i) is straightforward. In the case (ii) the result follows by Corollary 5.14 and Claim 5.15. □

Note that the notation $x \sim_z y$ means $xz = zy$, and therefore not necessarily implies that $y \sim_z x$. In fact, if words x, y, z satisfy both $x \sim_z y$ and $y \sim_z x$ then they all commute.

The following two corollaries are approaches to merge conditions (i) and (ii) of Theorem 5.16 into one to obtain a more compact form. In the first one we restrict the lengths of elements of X and Y.

Corollary 5.17. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$ with $|x_1| = |y_1|$ and $|x_2| = |y_2|$. Then X and Y are conjugates if and only if there exists a single word t such that $Xt = tY$ or $tX = Yt$.*

In the second one we consider the conjugacy via finite sets Z. In such case, similarly as we show that the lengths of the shortest elements of X and Y are equal, one can show that the same is true for the longest elements. Therefore, the following corollary is an immediate consequence of the previous one:

Corollary 5.18. *Let $X = \{x_1, x_2\} \subseteq \Sigma^+$ and $Y = \{y_1, y_2\} \subseteq \Sigma^+$. Then X and Y are conjugated via a finite non-empty set Z if and only if there exists a single word t such that $Xt = tY$ or $tX = Yt$.*

Chapter 6

On the computational complexity of infinite words

In Chapters 3 and 4 we have focused on combinatorial properties of infinite words. In [CuK] and [HKL] two new areas of investigation were introduced:

- the descriptional complexity of infinite words, *i.e.,* the comparative measure how complicated simple mechanisms are needed to generate particular infinite words;
- the computational complexity of infinite words, *i.e.,* the measure how much resources (such as time and space) are needed to generate a certain infinite word by a Turing machine.

The second paper concentrates on relations between these two complexities. Further results in this direction can be found in [HK].

In [CuK, HKL, HK] several interesting problems are proposed. In this chapter we will show that even some of the simplest problems proposed are equivalent to well-known hard open problems in the complexity theory of Turing machines.

In Section 6.1 we recall the definition of the computational complexity, while in Section 6.2 we define several simple methods for generating infinite words:

- iterating a morphism, the most commonly used method introduced already in [Th];
- iterating a deterministic generalized sequential machine (a dgsm for short), *i.e.,* a deterministic finite state transducer;
- double and triple D0L TAG systems.

In Section 6.3 we study an open problem, proposed in [HKL], namely whether all infinite words generated by iterating dgsm's have logarithmic space complexity. The problem has an affirmative answer in two special cases. First, it was shown in [HKL] that the greatest possible growth of a dgsm is exponential and that infinite words generated by such dgsm's have logarithmic space complexity. Second, here we show that the smallest non-trivial growth is $\Theta(n \log n)$ and that, similarly, dgsm's with such a growth generate infinite words which have logarithmic space complexity.

The general problem has been attacked in [Le], claiming that the answer is affirmative. On the other hand, in Section 6.3 we also show that this problem is equivalent to an other hard open problem asking whether unary classes of languages P and DLOG (denoted u-P and u-DLOG, respectively) are equivalent. One can easily observe that u-P = u-DLOG if and only if $\cup_{c>0} \mathrm{DTIME}(c^n) = \mathrm{DSPACE}(n)$.

In [HKL] another interesting problem is proposed: to find a concrete infinite word which cannot be generated in logarithmic space. It is mentioned already in [HK] that this problem is at least as hard as to prove $L \notin$ DLOG for some $L \in$ NP. In Section 6.4 we show that it is exactly as hard as the problem to find a concrete language, which does not belong to DSPACE(n). Note that even the problem to find a concrete language, which does not belong to $\mathrm{DSPACE}(\log n) =$ DLOG is a hard open problem.

Finally, in Section 6.5 we separate the classes of infinite words generated by double and triple D0L TAG systems as it was conjectured in [CuK].

6.1 The computational complexity of infinite words

The best way how to define the computational complexity of an object is to describe it in the terms of Turing machines. For example, the Kolmogorov complexity of a finite word is the size of the smallest Turing machine generating the word, *cf.* [Ko]. In the case of infinite words we will use the model of computation based on the *k-tape Turing machine*, which consists of

1. a finite state control;

2. k one-way infinite working tapes (we assume that there is a beginning on the left of the tape, but the tape is infinite to the right) each containing one two-way read/write head (*i.e.,* the head can move in both directions within the tape);

3. one infinite output tape containing one one-way write-only head.

We assume that the k-tape Turing machine starts in the initial state with all tapes empty and behaves as a usual Turing machine. We say that the k-tape Turing machine generates an infinite word $w \in \Sigma^{\mathbb{N}}$ if

1. in each step of the computation, the content of the output tape is a prefix of w;

2. for each prefix u of w, there is an integer n such that u is a prefix of the content of the output tape after n steps of the computation.

Let M be a k-tape Turing machine generating a word w. The *time* and *space complexities* of M are functions $T_M : \mathbb{N} \to \mathbb{N}$ and $S_M : \mathbb{N} \to \mathbb{N}$ defined as follows:

- $T_M(n)$ is the smallest number of steps of the computation of M when the prefix of w of length n is already written on the output tape;
- $S_M(n)$ is the space complexity of working tapes during first $T_M(n)$ steps of the computation, *i.e.,* the maximum of lengths of words written on working tapes.

Finally, for any integer function $s : \mathbb{N} \to \mathbb{N}$ we define the following complexity classes:

- GTIME$(s) = \{w \in \Sigma^{\mathbb{N}}$; there exists a k-tape Turing machine M generating w and $T_M(n) \leq s(n)$ for all $n \geq 1\}$;
- GSPACE$(s) = \{w \in \Sigma^{\mathbb{N}}$; there exists a k-tape Turing machine M generating w and $S_M(n) \leq s(n)$ for all $n \geq 1\}$;

It follows from the speed-up argument, as in ordinary complexity theory, that functions $s(n)$ and $c.s(n)$, where c is a constant, define the same space complexity classes of infinite words, *i.e.,* GSPACE(s) = GSPACE$(c.s)$.

6.2 Iterative devices generating infinite words

In this section we define several simple methods used for generating infinite words. The simplest and most commonly used method is to iterate a morphism $h : \Sigma^* \to \Sigma^*$: if h is non-erasing and for a letter $a \in \Sigma$, a is a prefix of $h(a)$, then there exists the limit

$$w = \lim_{n\to\infty} h^n(a)\,.$$

An illustrative scheme how the infinite word w is generated by the morphism h is depicted in Figure 6.1.

A natural generalization of this method is to use a more powerful mapping in the iteration: a *deterministic generalized sequential machine*, a *dgsm* for short. A dgsm τ is defined by

1. a finite set of states Q;
2. the initial state $q_0 \in Q$;

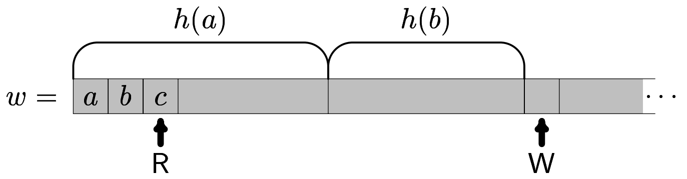


Figure 6.1: An illustration of the process of generating an infinite word w by iterating a morphism h. The reading head R is at the third position reading a letter c and the writing head W is at the position $|h(a)| + |h(b)| + 1$ prepared to write the word $h(c)$.

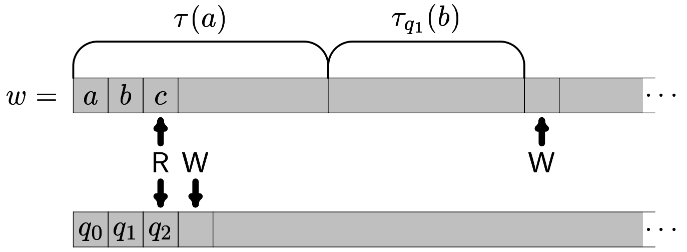


Figure 6.2: An illustration of the process of generating an infinite word w by iterating a dgsm τ. Note that τ_{q_1} is a variant of the dgsm τ with the initial state q_1.

3. an input alphabet Σ and an output alphabet Δ (we will assume $\Delta = \Sigma$ in this note to be able to iterate the dgsm τ);

4. a transition relation $\delta \subseteq Q \times \Sigma \times \Delta^* \times Q$, where δ is a partial function $Q \times \Sigma \to \Delta^* \times Q$.

A sequence of transitions

$$\alpha = (q_0, u_1, v_1, q_1)(q_1, u_2, v_2, q_2) \dots (q_{k-1}, u_k, v_k, q_k)$$

is a computation of τ with the input $I(\alpha) = u_1u_2 \dots u_k$ and the output $O(\alpha) = v_1v_2 \dots v_k$. Obviously, for an input $u \in \Sigma^*$ there exists at most one computation α of τ such that $I(\alpha) = u$. Hence, the mapping $\tau(u) = O(I^{-1}(u))$ is a well-defined partial function. As a convention we assume throughout that all dgsm's are non-erasing, *i.e.,* $\delta \subseteq Q \times \Sigma \times \Delta^+ \times Q$. A mechanism of generating an infinite word by iterating a dgsm τ is illustrated in Figure 6.2.

The further generalization of above methods leads to *double D0L TAG systems* which consist of two infinite one-way tapes each containing a one-way read-only head and a one-way write-only head. In each step of the generation both read-only heads read a symbol and move right to the next square while the write-only heads write the corresponding outputs to the first empty squares of these tapes. We assume that the infinite word generated by a double D0L TAG

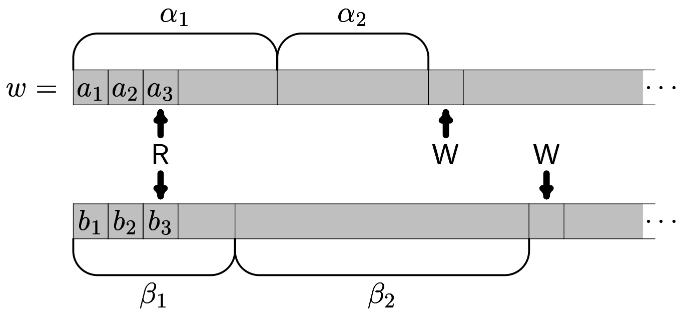


Figure 6.3: An illustration of the process of generating an infinite word w by a double D0L TAG system containing rewriting rules $\binom{a_1}{b_1} \to \binom{\alpha_1}{\beta_1}$ and $\binom{a_2}{b_2} \to \binom{\alpha_2}{\beta_2}$.

system is written on the first tape. A double D0L TAG system can be specified in the terms of rewriting rules of the form:

$$\begin{pmatrix} a \\ b \end{pmatrix} \to \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \text{ where } a, b \in \Sigma,\ \alpha, \beta \in \Sigma^+.$$

Figure 6.3 shows an idea how a double D0L TAG system works.

Assuming that in each rewriting rule, $|\beta| = 1$, we get a mechanism which iterates a dgsm. Finally, we can define *triple D0L TAG systems* by extending the number of tapes to three.

6.3 Do dgsm's have logarithmic space complexity?

In this section we study the following problem proposed in [HKL]:

- are all infinite words generated by iterating dgsm's in GSPACE$(\log n)$?

First, let us recall one result of [HKL] stating that an infinite word generated by a dgsm which has an exponential growth has logarithmic space complexity. Here, the growth of a dgsm τ is an integer function $g : \mathbb{N} \to \mathbb{N}$, where $g(n)$ is the length of $\tau^n(a)$. Note that for a dgsm the exponential growth is the maximal possible growth.

Next, we show that the smallest non-trivial growth of a dgsm is $\Theta(n \log n)$ and that dgsm's with the growth $\Theta(n \log n)$ generate infinite words which have logarithmic space complexity. More precisely:

Lemma 6.1. *If a dgsm has the growth $o(n \log n)$ then it generates an ultimately periodic infinite word. Such an infinite word can be generated in constant space.*

Proof. Let τ be a dgsm with the growth $o(n \log n)$ generating an infinite word $w = w_1 w_2 \ldots$, with $w_i \in \Sigma$. Let $\tau_q(z)$ (resp. $\sigma_q(z)$) be the output (resp. the last state) of the dgsm τ after reading the input z and starting in the state q.

We define two sequences of words and a sequence of states of the dgsm τ. Let a be the starting symbol, q_0 be the initial state and let $\tau(a) = \tau_{q_0}(a) = av$. Then put

$$\begin{array}{lll} u_1 = a, & v_1 = v, & q_1 = \sigma_{q_0}(a) \\ u_n = u_{n-1}v_{n-1}, & v_n = \tau_{q_{n-1}}(v_{n-1}), & q_n = \sigma_{q_{n-1}}(v_{n-1}). \end{array}$$

Observe that $\tau^n(a) = u_{n+1} = u_n v_n = \cdots = u_1 v_1 v_2 \ldots v_n$. This implies $1 + \sum_{j=1}^{n} |v_j| = |\tau^n(a)| = o(n \log n)$. Next, we estimate the length of the increment: $|v_n|$. Since the dgsm τ is non-erasing, we have that $|v_i| < |v_n|$ for all $i < n$. This implies

$$n|v_n| \leq \sum_{j=n+1}^{2n} |v_j| \leq \sum_{j=1}^{2n} |v_j| = o(2n \log 2n) = o(n \log n)\,.$$

Hence, $|v_n| = o(\log n)$, *i.e.,* for any constant c there exists an integer n such that $c^{|v_n|} < n$. If we take $c = \mathrm{card}(\Sigma)+1$, there must be a repetition among the words $v_1, \ldots, v_n$, say $v_i = v_{i+k}$ for some integers $i, i+k \leq n$ and $k > 0$. Then $v_j = v_{j+k}$ for all $j \geq i$, hence the infinite word $w = u_1 v_1 v_2 v_3 \ldots$ is ultimately periodic and it can be then generated in constant space, *cf.* [HKL], Lemma 3.1. □

Lemma 6.2. *An infinite word generated by a dgsm with the growth* $\Theta(n \log n)$ *has logarithmic space complexity.*

Proof. Let τ be a dgsm with the growth $\Theta(n \log n)$ generating an infinite word $w = w_1 w_2 \ldots$. Consider the sequences $\{u_n\}_{n \geq 0}$, $\{v_n\}_{n \geq 0}$, $\{q_n\}_{n \geq 0}$ defined in the previous proof.

We construct a Turing machine M generating w as follows. In the first step it writes u_1 on the output tape, v_1 on the first tape and it sets to the state q_1. In each step $n > 1$, it simulates the dgsm τ on the input written on the first tape starting in the state q_{n-1}. The output of the simulation of τ is written, at the same time, to the output tape and to some temporary tape, so that after the simulation it can be copied back to the first tape. Hence, in the end of the step n the first tape contains the word v_n. The last state of the simulation of τ in the step n is q_n, from which the simulation continues in the next step.

The space needed to generate the m-th letter of w is at most $|v_n|$, where n is an integer such that v_n contains the letter w_m, *i.e.,*

$$\sum_{i=1}^{n-1} |v_i| < m \leq \sum_{i=1}^{n} |v_i|\,.$$

One can show in the same way as in in the proof of Lemma 6.1 that $|v_n| = \mathcal{O}(\log n)$. Moreover, since $n \leq 1+\sum_{i=1}^{n-1} |v_i| \leq m$, we have that $|v_n| = \mathcal{O}(\log n) = \mathcal{O}(\log m)$. Hence, the Turing machine M works in logarithmic space. □

We have seen that the infinite words generated by iterating dgsm's with the maximal or the minimal non-trivial growth have logarithmic space complexity. Intuitively, one could expect that by combining the proof techniques in these two cases we could prove that all infinite words generated by iterating any dgsm's have such complexity, *i.e.,* to obtain an affirmative answer to the problem stated in the beginning of this section. In fact, such an attempt to prove this result can be found in [Le]. However, here, we prove that the problem is equivalent to the hard open problem whether classes of unary languages u-DLOG and u-P are equivalent.

This in some sense contradicts the result of [Le] that all infinite words generated by iterating dgsm's have logarithmic space complexity: if the result in [Le] is correct then, together with the following theorem, we have that D-EXPTIME = DSPACE(n), which is unlikely. Since [Le] gives only a sketch of the proof of the result, we are unable to check if it is correct, but we believe that some case has been overlooked in [Le].

Finally, let us prove our result claiming that the problem whether all infinite words generated by iterating dgsm's have logarithmic space complexity is a difficult one.

Theorem 6.3. *All infinite words generated by iterating dgsm's have logarithmic space complexity if and only if* u-P = u-DLOG.

Proof. First, let us assume that u-P = u-DLOG. Take a dgsm τ over a finite alphabet Σ generating an infinite word $w = w_1 w_2 \ldots$. We prove that the space complexity of w is $\mathcal{O}(\log n)$. It is obvious that there is a 1-tape Turing machine M generating the word w in quadratic time. Consider the languages $L_c = \{0^n;\ n \geq 1,\ w_n = c\}$ for all $c \in \Sigma$. Note that L_c is a unary language. We can easily construct a Turing machine recognizing L_c in quadratic time using the Turing machine M. By the assumption there exist Turing machines M_c recognizing the languages L_c in logarithmic space. Now, consider a 3-tape Turing machine, which runs M_c's to generate the n-th letter of w by using the third tape as a working tape. It stores the binary representation of n on the first tape and the position of the head of Turing machine M_c on the second tape. Before each run of any M_c it erases the working tape and writes "the position 1" on the second tape. It runs M_c for each letter c of the alphabet of τ until some M_c accepts, and then it writes the letter c on the output tape. In each step of the simulation of any M_c it checks whether the position represented on the second tape is the last one. Clearly such a machine generates the word w in logarithmic space.

Second, assume that all words generated by iterating a dgsm have logarithmic space complexity. Take a Turing machine M working in polynomial time, *i.e.,* $T(M) = \mathcal{O}(n^k)$, recognizing a language $L \subseteq 0^*$. We construct a 1-tape Turing machine M' with the tape divided into three layers. On the first layer it generates unary inputs in increasing order, on the second layer it simulates computations of M on the input stored on the first layer, and on the third layer it writes 1, if the computation ends in an accepting state, or 0, if it ends in a rejecting state. Before each simulation it erases the second and third layers of the tape.

Now, consider a dgsm τ which carries out the computations

$$C_i \to C_{i+1}, \text{ for } i \geq 0,$$

where C_i corresponds to the i-th configuration of M'. It also maps the starting letter \$ into the starting configuration of M'. Clearly, the iteration of τ will generate an infinite word: the sequence $W = \$C_0C_1C_2\ldots$ of all configurations of the computation of the Turing machine M'. By the assumption the infinite word W has logarithmic space complexity, *i.e.,* there exists a Turing machine M'' generating W in logarithmic space. Finally, we define a Turing machine M''' recognizing L, which on the input 0^n runs M'', but instead of writing the bits of W to the output tape, it compares its input with the input on the first layer of each generated configuration, and moreover, it checks the first letter on the third layer of each generated configuration. When the compared inputs coincide and the first letter on the third layer is 0 or 1 then the Turing machine M''' halts in the rejecting or in the accepting state, respectively. Otherwise, it continues in generating bits of the next configuration. Clearly, M''' recognizes the language L.

Now, it suffices to show that M''' works in logarithmic space. Let C_{i_n} be the configuration in which M' writes 0 or 1 on the first place of the third layer, while the first layer contains 0^n. Hence in the block of configurations $B_n = C_{i_{n-1}+1}\ldots C_{i_n}$, M' erases the second and the third layer of the tape, changes the input on the first layer to 0^n and runs the Turing machine M on this input. Since M works in time $\mathcal{O}(n^k)$ the length of any configuration in block B_n is at most $\mathcal{O}(n^k)$ and the number of configurations in B_n is at most $\mathcal{O}(n^k)$. Hence the length of the block B_n is at most $\mathcal{O}(n^{2k})$. The Turing machine M''' generates the first x blocks of configurations on the input 0^x until it halts. Hence it generates the prefix of the infinite word W of length

$$\sum_{n=0}^{x} |B_n| = \mathcal{O}(x^{2k+1})\,.$$

Since M'' works in logarithmic space, M''' will use space $\mathcal{O}(\log x)$ to carry out the computation on the input 0^x. □

6.4 Logarithmic space complexity

The second part of Problem 5.2 in [HK] asks to find a specific infinite word which cannot be generated in logarithmic space. We show that this problem is as hard as the problem to find a "natural" specific language which does not belong to DSPACE(n), and this is a hard open problem. (By a "natural" language we mean a language which is not obtain by diagonalization.)

Notation 6.1. Denote the n-th binary word in lexicographical order by $\mathrm{lex}(n)$. Note that for $n \geq 1$, $\mathrm{bin}(n) = 1\,\mathrm{lex}(n)$.

Definition 6.2. Let w be an infinite binary word and $L \subseteq \{0,1\}^*$ an binary language. We say that w *determines* the language L if for every positive integer n, the n-th letter of w is 1 if and only if $\mathrm{lex}(n)$ belongs to L.

Theorem 6.4. *Let w be an infinite binary word and L the language determined by the word w. Then the word w is in* GSPACE($\log n$) *if and only if L belongs to* DSPACE(n).

Proof. First, assume that w has logarithmic space complexity. Let M be a Turing machine generating w in logarithmic space. We construct a Turing machine M' recognizing the language L. Let $\mathrm{lex}(n)$ be the word on the input tape of M', where n is a positive integer. The length of the input is $\Theta(\log n)$. M' simulates M in the following way: it remembers only the last letter generated by M and counts the number of them on a special working tape. When this number is equal to n, it stops and accepts the input if and only if the last output letter was 1.

Since M uses only $\Theta(\log n)$ space to generate the first n output letters and the same space is needed for counting the number of output letters, M' works in space $\Theta(\log n)$, which is linear to the length of the input. Hence, $L \in$ DSPACE(n).

Next, assume that $L \in$ DSPACE(n). So we have a Turing machine M recognizing L in linear space. Let M' be a Turing machine such that it generates words in lexicographical order on the first working tape and runs M on each generated word. Depending on if the word was accepted or rejected it writes 1 or 0 on the output tape. Clearly, the length of the n-th word on the first working tape is $\Theta(\log n)$. M works in linear space, hence in space $\Theta(\log n)$. Therefore, M' uses logarithmic space to generate the n-th letter: $w \in$ GSPACE($\log n$). □

As a consequence we have that if we would be able to show about a specific infinite word that it does not belong to GSPACE($\log n$), then we would have also a specific language which does not belong to DSPACE(n), and vice versa. Note, that even the problem to show that a specific language does not belong to DSPACE($\log n$) is open.

6.5 Separation of double and triple D0L TAG systems

In the Section 6 of [HKL] is mentioned that the generation of infinite words by double D0L TAG systems is a very powerful mechanism, and that it is not known any concrete example of an infinite word which cannot be generated by this mechanism, although by a diagonalization argument such words clearly exist. In [CuK] (Conjecture 4) is conjectured that there exists an infinite word that can be generated by a triple D0L TAG system, but not by any double D0L TAG system. In what follows we are going to give the whole class of infinite words which cannot be generated by any double D0L TAG system. Combining this result with some results in [CuK] we can also give an affirmative answer to Conjecture 4 of [CuK]. First, let us fix some notation.

Notation 6.3. Let $w = c_1 \dots c_n$ be a word with $c_i \in \Sigma$. Then $\mathrm{symb}(w) = \{c_i,\ 1 \le i \le n\}$ denotes the set of all symbols occurring in the word.

Theorem 6.5. *Let $s : \mathbb{N} \to \mathbb{N}$ be an integer function such that $s(i) \in 2^{\omega(i)}$,* i.e., *$s(i)$ grows faster than exponentially. Then the infinite word*

$$w = 10^{s(1)}10^{s(2)}10^{s(3)}\dots$$

cannot be generated by any double D0L TAG system.

Proof. Assume that w can be generated by a double D0L TAG system, and let τ be such a system. Let M be the set of all symbols occurring on the second tape of τ and B the maximal number of symbols written on any tape in one step, *i.e.,*

$$B = \max\{\max(|\alpha|, |\beta|),\ \text{where} \begin{pmatrix} a \\ b \end{pmatrix} \to \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \text{is a rewriting rule of } \tau\}\,.$$

Let $w_i = 10^{s(i)}$. First, we show that for any constant $k \ge 1$, there is a positive integer j such that

$$s(j) + 1 = |w_j| > k|w_1 \dots w_{j-1}1| + 1\,. \tag{6.1}$$

Since $s(i) \in 2^{\omega(i)}$, for any number $c > 1$, there is an integer j such that

$$\begin{aligned} &|w_j| > c^j, \\ &|w_i| \le c^i \text{ for all } 1 \le i \le j-1. \end{aligned}$$

Then we have

$$k|w_1 \dots w_{j-1}1| \le k\sum_{i=0}^{j-1} c^i \le k \cdot \frac{c^j - 1}{c - 1}\,.$$

Taking $c = k+1$, we get $k|w_1 \dots w_{j-1}1| + 1 \le c^j < |w_j|$.

Consider that we are reading the first 0 of $w_j = 10^{s(j)}$ in the word w. Let u_1 (resp. v_1) be the word written on the first (resp. second) tape between the reading and the writing head. And let, recursively, u_l (resp. v_l) be the word added on the first (resp. second) tape after reading v_{l-1}. Note that for all $l \ge 1$, we have that $|v_1 \dots v_{l-1}| < |u_1 \dots u_l|$.

We define also a set $M_0 \subseteq M$ as follows. For $x \in M$, let $\binom{0}{x} \to \binom{\alpha}{\beta}$ be a rule of τ. Then, $x \in M_0$ if and only if $\alpha \in 0^+$. Hence, M_0 contains every symbol $x \in M$ such that when reading 0 on the first tape and the symbol x on the second tape, it writes only 0's on the first tape.

Assume that for some $i \ge 1$,

$$\sum_{l=1}^{i} |u_l| \le s(j)\,. \tag{6.2}$$

By (6.2), the words $u_1, \dots, u_i$ contain only 0's, hence when reading the words $v_1, \dots, v_{i-1}$, only 0's are written on the first tape. Since $|v_1 \dots v_{l-1}| < |u_1 \dots u_l|$, we have also that while reading the words $v_1, \dots, v_{i-1}$, only 0's are read from the first tape.

Consider the following oriented graph. The vertices are elements of M. There is an arc $x \to y$, if there is a rule $\binom{0}{x} \to \binom{\alpha}{\beta}$ in τ such that $y \in \mathrm{symb}(\beta)$ (see Notation 6.3). For $X \subseteq M$, let $\mathrm{clos}(X)$ be the set of all vertices of the graph to which we reach from any vertex of X following the arcs. If $\mathrm{clos}(\mathrm{symb}(v_1)) \subseteq M_0$ then since $u_1 \in 0^+$, we have, by induction, that $u_l \in 0^+$ and $\mathrm{symb}(v_l) \subseteq \mathrm{clos}(\mathrm{symb}(v_1)) \subseteq M_0$ for all $l \ge 1$. This is a contradiction since there must be u_l containing 1.

Hence, there must be an oriented path starting in a vertex of $\mathrm{symb}(v_1)$ and ending in a vertex of $M - M_0$ of length at most $\mathrm{card}(M_0)$. Let i_0 be the length of the shortest of such paths. If we prove that (6.2) holds for $i = \mathrm{card}(M) + 1 \ge \mathrm{card}(M_0) + 2 \ge i_0 + 2$ then, since during reading $v_1, \dots, v_{i_0+1}$ only 0's are read and written on the first tape, we have that $\mathrm{symb}(v_1), \dots, \mathrm{symb}(v_{i_0}) \subseteq M_0$ and $\mathrm{symb}(v_{i_0+1}) \cap (M - M_0) \ne \emptyset$. This implies that $1 \in \mathrm{symb}(u_{i_0+2})$ contradicting (6.2).

Now it suffices to prove that Equation (6.2) holds for $i = \mathrm{card}(M) + 1$. After reading one symbol, the system τ can write on any tape at most B symbols. Hence, we have that $|u_{l+1}|, |v_{l+1}| \le B|v_l|$ for all $l \ge 1$. We estimate the left hand side of (6.2):

$$\sum_{l=1}^{\mathrm{card}(M)+1} |u_l| \le |u_1| + \sum_{l=2}^{\mathrm{card}(M)+1} B^{l-1}|v_1| \le \max(|u_1|, |v_1|). \sum_{l=0}^{\mathrm{card}(M)} B^l\,. \tag{6.3}$$

Notice that $T = \sum_{l=0}^{\mathrm{card}(M)} B^l$ is a constant for τ.

Next, we estimate $|u_1|$ and $|v_1|$. Consider the situation when the reading heads are on the $(n+1)$-th symbols of both tapes. Then on each tape n symbols have already been read and hence at most Bn symbols written. So, there is at most $(B-1)n$ symbols between writing and reading head on each tape. This implies

$$\max(|u_1|, |v_1|) \leq (B-1)|w_1 \dots w_{j-1}1| \,. \tag{6.4}$$

Taking $k = T(B-1)$, we obtain

$$\sum_{l=1}^{\mathrm{card}(M)+1} |u_l| \overset{(6.3)}{\leq} T.\max(|u_1|, |v_1|) \overset{(6.4)}{\leq} T(B-1)|w_1 \dots w_{j-1}1| \overset{(6.1)}{<} s(j)$$

as desired. □

Let us recall one result proved in Examples 11 and 13 of [CuK].

Lemma 6.6. *Let $s : \mathbb{N} \to \mathbb{N}$ be an integer function which is computable,* i.e., *can be computed by a Turing machine. Then there exists an integer function $t : \mathbb{N} \to \mathbb{N}$ such that $t(n) \geq s(n)$ for all $n \geq 1$ and the word*

$$w = 10^{t(1)}10^{t(2)}10^{t(3)} \dots$$

can be generated by a triple D0L TAG system. Moreover, such a function t can be effectively computed.

The proof of the lemma is based on the following idea. Let M be a Turing machine computing unary strings $1^{s(1)}, 1^{s(2)}, \dots$ and let τ be a dgsm generating the sequence of configurations of the computation of M. We can easily extend the dgsm τ to a triple D0L TAG system by coding all letters generated by τ to 0, except for the last letters of the strings in the sequence $1^{s(1)}, 1^{s(2)}, \dots$, which are coded to 1. Together with our result we have the following corollary.

Corollary 6.7. *There exists an infinite word which can be generated by a triple D0L TAG system, but not by any double D0L TAG system.*

Hence, the inclusion “double D0L $\subseteq$ triple D0L” is proper, as conjectured in Conjecture 4 in [CuK].

Bibliography

[AL] Albert, M.H., Lawrence, J., *A proof of Ehrenfeucht's conjecture*, Theoret. Comput. Sci. **41**, no. 1, pp. 121–123 (1985).

[BP] Berstel, J., Perrin, D., *Theory of codes*, Pure and Applied Mathematics, Vol. **117**, Orlando etc.: Academic Press (1985).

[BPPR] Berstel, J., Perrin, D., Perrot, J.F., Restivo, A., *Sur le théorème du défaut*, J. Algebra **60**, no. 1, pp. 169–180 (1979).

[Br] Bruyère, V., *Codes*, Chap. 7 in: M. Lothaire, Algebraic combinatorics on words, Cambridge University Press (2002).

[ChK] Choffrut, C., Karhumäki, J., *Combinatorics of words*, in: G. Rozenberg and A. Salomaa (eds), Handbook of formal languages, Vol. **1**, Springer, Berlin, pp. 329–438 (1997).

[CKM] Cassaigne, J., Karhumäki, J., Maňuch, J., *On conjugacy of languages*, Theor. Inform. Appl. (to appear).

[CKO] Choffrut, C., Karhumäki, J., Ollinger, N., *The commutation of finite sets: a challenging problem*, in: Proc. of WORDS (Rouen, 1999), Theoret. Comput. Sci. **273**, no. 1–2, pp. 69–79 (2002).

[Co] Conway, J.H., *Regular algebra and finite machines*, Chapman and Hall Mathematics Series, London (1971).

[CuK] Culik K. II, Karhumäki, J., *Iterative devices generating infinite words*, Int. J. Found. Comput. Sci. **5**, no. 1, pp. 69-97 (1994).

[DM] Ďuriš, P., Maňuch, J., *On the computational complexity of infinite words*, Theoret. Comput. Sci. (to appear).

[Ei] Eilenberg, S., *Automata, languages and machines*, Pure and Applied Mathematics, Vol. 58, Academic Press, New York (1974).

[FW] Fine, N.J., Wilf, H.S., *Uniqueness theorems for periodic functions*, Proc. Amer. Math. Soc. **16**, pp. 109–114 (1965).

[Gu] Guba, V.S., *Equivalence of infinite systems of equations in free groups and semigroups to finite subsystems*, Mat. Zametki **40**, no. 3, pp. 321–324 (1986).

[Ha] Harju, T., *On factorizations of words*, Bulletin of the EATCS **24**, p. 217 (1984).

[HK] Harju, T., Karhumäki, J., *On the defect theorem and simplifiability*, Semigroup Forum **33**, no. 2, pp. 199–217 (1986).

[HKL] Hromkovič, J., Karhumäki, J., Lepistö, A., *Comparing descriptional and computational complexity of infinite words*, in: Proc. of Results and trends in theoretical computer science (Graz, 1994), Lecture Notes in Comput. Sci. **812**, Springer, Berlin, pp. 169–182 (1994).

[HKP] Harju, T., Karhumäki, J., Plandowski, W., *Independent systems of equations*, Chap. 14 in: M. Lothaire, Algebraic combinatorics on words, Cambridge University Press (2002).

[Ho] Honkala, J., *A defect property of codes with unbounded delays*, Discrete Appl. Math. **21**, no. 3, pp. 261–264 (1988).

[HP] Harju, T., Petre, I., *On commutation and primitive roots of codes*, TUCS Technical Report 402 (2001).

[Ka1] Karhumäki, J., *On three-element codes*, in: Proc. of Eleventh international colloquium on automata, languages and programming (Antwerp, 1984), Theoret. Comput. Sci. **40**, pp. 3–11 (1985).

[Ka2] Karhumäki, J., *A property of three-element codes*, Theoret. Comput. Sci. **41**, no. 2–3, pp. 215–222 (1985).

[Ka3] Karhumäki, J., *Combinatorial and computational problems of finite sets of words*, in: Proc. of MCU'01, Lecture Notes in Comput. Sci. **2055**, pp. 69–81 (2001).

[KLP] Karhumäki, J., Latteux, M., Petre, I., *The commutation with codes and ternary sets of words*, (a manuscript).

[KM] Karhumäki, J., Maňuch, J., *Multiple factorizations of words and defect effect*, Theoret. Comput. Sci. **273**, no. 1–2, pp. 81–97 (2002).

[KMP] Karhumäki, J., Maňuch, J., Plandowski, W., *A defect theorem for bi-infinite words*, Theoret. Comput. Sci. (to appear).

[Ko] Kolmogorov, A.N., *Three approaches to the quantitative definition of information*, Internat. J. Comput. Math. **2**, pp. 157–168 (1968).

[KPl] Karhumäki, J., Plandowski, W., *On the size of independent systems of equations in semigroups*, in: Proc. of 19th International Symposium on Mathematical Foundations of Computer Science (Košice, 1994), Theoret. Comput. Sci. **168**, no. 1, pp. 105–119 (1996).

[KPe] Karhumäki, J., Petre, I., *On the centralizer of a finite set*, in: Proc. of Automata, languages and programming (Geneva, 2000), Lecture Notes in Comput. Sci. **1853**, Springer, Berlin, pp. 536–546 (2000).

[Le] Leiss, E.L., *Language equations*, Monographs in Computer Science, Springer-Verlag, New York (1999).

[LeS] Lentin, A., Schützenberger, M.-P., *A combinatorial problem in the theory of free monoids*, in: Combinatorial Mathematics and its Applications (Proc. Conf., Univ. North Carolina, Chapel Hill, N.C., 1967), Univ. North Carolina Press, pp. 128–144 (1969).

[Lo] Lothaire, M., *Combinatorics on words*, Encyclopedia of Mathematics and its Applications **17**, Addison-Wesley Publishing Co., Reading, Mass. (1983).

[LRLR] Le Rest, E., Le Rest, M. *Sur la combinatoire des codes à deux mots*, Theoret. Comput. Sci. **41**, no. 1, pp. 61–80 (1985).

[LyS] Lyndon, R.C., Schützenberger, M.P., *The equation $a^M = b^N c^P$ in a free group*, Michigan Math. J. **9**, pp. 289–298 (1962).

[Mak] Makanin, G.S., *The problem of solvability of equations in a free semigroup*, Mat. Sb. **103(145)**, no. 2, pp. 147–236 (1977) (English transl. in Math. USSR Sb. **32**, pp. 129–198 (1979)).

[Man] Maňuch, J., *Defect effect of bi-infinite words in the two-element case*, Discrete Math. Theor. Comput. Sci. **4**, No. 2, pp. 273–290 (2001).

[Pe1] Perrin, D., *Codes conjugués*, Information and Control **20**, pp. 222–231 (1972).

[Pe2] Perrin, D., *Sur les groupes dans les monoides finis*, in: Proc. of Noncommutative structures in algebra and geometric combinatorics (Naples, 1978), Quad. Ric. Sci. **109**, CNR, Rome, pp. 27–36 (1981).

[Pl] Plandowski, W., *Satisfiability of word equations with constants is in* PSPACE, in Proc. of FOCS'99, IEEE, pp. 495–500 (1999).

[Ra] Ratoandromanana, B., *Codes et motifs*, RAIRO Inform. Theor. Appl. **23**, no. 4, pp. 425–444 (1989).

[Sc] Schützenberger, M.-P., *A property of finitely generated submonoids of free monoids*, in: Proc. of Algebraic theory of semigroups (Proc. 6th Algebraic Conf., Szeged, 1976), Colloq. math. Soc. János Bolyai **20**, North-Holland, pp. 545–576 (1979).

[Sh] Shyr, H.J., *Free monoids and languages*, CSLI Lecture Notes, Taichung (Taiwan), Hon Min Book Company (1991).

[Th] A. Thue, *Über unendliche zeichenreihen*, Norske Vid. Selsk. Skr., I Mat. Nat. KI., Kristiania **7**, pp. 1–22 (1906).

Appendix A

Proof of Lyndon and Schützenberger Lemma

Let us proof Lemma 2.7. Our proof is based on that in [Sh], however, the use of not very difficult results, *i.e.,* of Lemmas 2.6 and 3.8, will make our proof shorter.

Proof of Lemma 2.7. Let $m, n, p \geq 2$ be integers. Assume that the word equation $x^m y^n = z^p$ has a non-periodic solution. Lemma 2.6 yields immediately

$$\begin{aligned} &|z| > (m-1)|x| \geq |x|, \quad \text{and} \\ &|z| > (n-1)|y| \geq |y| \,. \end{aligned} \tag{A.1}$$

Hence, $4|z| > m|x| + n|y| = p|z|$, *i.e.,* we have that either $p = 2$, or $p = 3$.

Case $p = 3$. If $m \geq 3$ and $n \geq 3$ then, by (A.1), we have $|z| > 2|x|$ and $|z| > 2|y|$. Hence, as above,

$$3|z| > (m-1)|x| + (n-1)|y| + |x| + |y| = p|z| \,,$$

a contradiction. Therefore, without lost of generality we can assume that $m = 2$.

Again, by (A.1), we obtain

$$\begin{aligned} n|y| &= \qquad\qquad 3|z| - 2|x| > \quad 3|z| - 2|z| = |z| > (n-1)|y| \,, \quad \text{and} \\ 2|x| &= 3|z| - (n-1)|y| - |y| > 3|z| - |z| - |z| = |z| > |x| \,. \end{aligned}$$

We have the situation depicted in Figure A.1, *i.e.,* there are non-empty words x_1, x_2, y_1 and y_2 such that

$$x = x_1 x_2 \,, \qquad y = y_1 y_2 \,, \qquad \text{and} \qquad z = x x_1 = x_2 y_1 = y_2 y^{n-1}. \tag{A.2}$$

By the length argument we have that $|y_1| = 2|x_1|$. Since, by (A.2), both x_1 and y_1 are suffixes of z, there is a non-empty word y_0 such that $y_1 = y_0 x_1$. Note

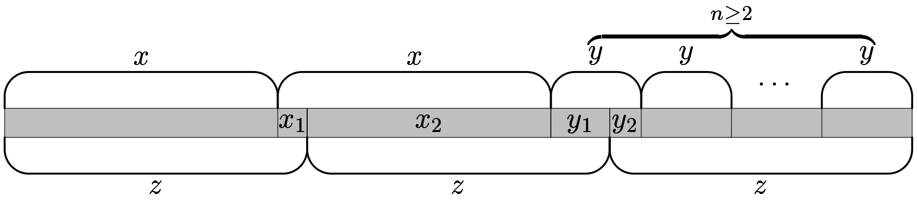


Figure A.1: The situation in the case $x^2y^n = z^3$.

that $|x_1| = |y_0|$. It follows that $x_1x_2 = x_2y_0$, *i.e.*, by Lemma 2.10, there are words p, q and integers $k \geq 1$ and $l \geq 0$ such that pq is primitive and

$$x_1 = (pq)^k, \qquad y_0 = (qp)^k, \qquad \text{and} \qquad x_2 = p(qp)^l\,.$$

Therefore,

$$p(qp)^{l+k}p(pq)^k = x_2y_0x_1 = x_2y_1 \overset{(\mathrm{A.2})}{=} y_2y^{n-1} = y_2[(qp)^k(pq)^ky_2]^{n-1}. \qquad (\mathrm{A.3})$$

Further, by (A.1) and (A.2) it follows that

$$|y_0| + |x_1| + |y_2| = |y| \overset{(\mathrm{A.1})}{<} |z| \overset{(\mathrm{A.2})}{=} |xx_1| = 2|x_1| + |x_2|\,,$$

which implies that $|y_2| < |x_2|$. Hence,

$$|y_2(qp)^k| < |x_2(qp)^k| = |p(qp)^{l+k}|\,.$$

Together with Equation (A.3) this yields that the pair (qp, pq) matches the word $p(qp)^{l+k}$. By Lemma 3.8 we have 3 possibilities. Either the pair matches the beginning of the word $p(qp)^{l+k}$, *i.e.*, $y_2(qp)^k = 1$; or it matches the end of the word, *i.e.*, $p(qp)^{l+k} = y_2(qp)^k$ and $p(pq)^k = (pq)^ky_2[(qp)^k(pq)^ky_2]^{n-2}$; or $l + k = 1$. The first case leads to a contradiction, since y_2 is non-empty. In the second case, we have a non-trivial equation over $\{p, q\}$, a contradiction with the primitiveness of pq. In the last case we have $l = 0$ and $k = 1$. Equation (A.3) simplifies to $pqpppq = y_2(qppqy_2)^{n-1}$. By the length argument we have that $n = 2$ and $|y_2| = |p|$. Obviously, this implies that $y_2 = p$ and $pq = qp$, again a contradiction with the primitiveness of pq.

Case $p = 2$. Assume that (x, y, z) is a solution of the equation $x^my^n = z^2$ such that z is of the minimal length. By (A.1), without lost of generality, there are words x_1 and x_2 such that

$$x = x_1x_2\,, \qquad \text{and} \qquad z = x^{m-1}x_1 = x_2y^n.$$

This implies $x_2^2y^n = x_2x^{m-1}x_1 = (x_2x_1)^m$. If $m \geq 3$ then the contradiction follows by the above considered cases. If $m = 2$ then, since $|x_2x_1| = |x| < |z|$, we have a contradiction with the minimality of $|z|$. □