Study of Primitive and Pseudo-Bordered Partial Words

A thesis submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

Ananda Chandra Nayak
(Roll Number: 136123006)

to the

Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati - 781 039, India

March 2019

In memory of
my mother
with eternal love and appreciation

DECLARATION

It is certified that the work contained in the thesis entitled **"Study of Primitive and Pseudo-Bordered Partial Words"** has been done by me, a student in the Department of Mathematics, Indian Institute of Technology Guwahati under the guidance of **Dr. Kalpesh Kapoor** for the award of Doctor of Philosophy and this work has not been submitted elsewhere for a degree.

March 2019

Ananda Chandra Nayak
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati - 781 039, India

CERTIFICATE

It is certified that the work contained in the thesis entitled **"Study of Primitive and Pseudo-Bordered Partial Words"** by **Ananda Chandra Nayak**, a student in the Department of Mathematics, Indian Institute of Technology Guwahati for the award of the degree of Doctor of Philosophy has been carried out under my supervision and this work has not been submitted elsewhere for a degree.

March 2019

Dr. Kalpesh Kapoor
Professor
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati - 781 039, India

Acknowledgements

The completion of this thesis would not have possible without the support and guidance from many people. I would like to take the opportunity to express my sincere gratitude to all of them. It has been a great privilege to spend my Ph.D. duration at Department of Mathematics, Indian Institute of Technology Guwahati.

First and foremost, I express my deepest appreciation to my thesis supervisor Dr. Kalpesh Kapoor for allowing me to pursue my Ph.D. research work under his supervision and introducing me to the area of word combinatorics and partial words. He patiently provided me the constant supervision, encouragement and advises necessary for me to complete my thesis work. Under his valuable guidance and advice, I have got exposure, not only in the area related to my research work but also other perspectives.

I want to convey my sincere gratitude to my doctoral committee members Dr. Gautam Kumar Das, Dr. Benny George K. and Dr. Bikas Bhattacharjya for their valuable suggestions to improve my research work and for reviewing my research work. My sincere thanks goes to Dr. Sagarmaya Dutta for helping me to review my work and for a fruitful discussion with him. I would like to thank Dr. Kalpana Mahalingam for allowing me to visit IIT Madras for two months to conduct my research work and introducing me to her group. During my visit to IIT Madras, I got the opportunity to work with Dr. Manasi Kulkarni who helped me to give a good shape to my thesis. I am thankful to Dr. Somanth Bera, Dr. Prithwineel Pal, Antaryami and Sasi for their immense help during my visit to IIT Madras. I would like to extend my heartful thanks to Dr. Prajakta Nimbhorkar, Swapna sir, Dr. B. N. B Roy for their encouragement and inspiring me to do research in theoretical computer science.

I thank Dr. Amit Kumar Srivastava for being a co-author of our papers and his valuable discussions on the subjects have been fruitful and have given me some insightness on the topic.

I am grateful to Ministry of Human Resource Development (MHRD), Government of India, for providing me financial assistance to pursue my Ph.D. at IIT Guwahati. I sincerely acknowledge IITG for providing me the necessary facilities which were quite essential for conducting my research work. I also thank the Head of Department of Mathematics, IIT Guwahati, for providing me the necessary facilities in the department to carry out my research work. I am also thankful to the staff members of Department of Mathematics, IIT Guwahati, for their support in administrative and technical works.

I wish to convey my heartiest thanks to my friends, my colleagues, my labmates for their moral support and company during my staying period in IIT Guwahati campus. I would like to thank Shiba, Pratap, Swarup, Rasmi Bhai, Himadri Sahu, Bibhuti, Ranjan, Ramesh, Pankaj, Pradeep Bhai, Subhasish, Chitta, Rahul, Prakash, Bala Subramani, Swarup Barik, Soumya, Saibal, Mohit, Gaurav, Anand Kamal, Sunil, Shibshankar, Himadri Nayak, Barun, Ashish, Maneesh, Tanushree, Madhu, Debu, Pratibha, Manjanna, Debasish, Mrutyunjay, Rajeev to name a few, for their help and support, and advices during my stay in IIT Guwahati. I am really lucky to have a friend like Ramesh Jallu, who is so kind and humble and I studied and prepared for the course work, comprehensive exam and also for other exams with him.

I would like to thank Muthuvel Murugan I, my roommate during my staying period in CMI, who helped me to understand the depth of the subjects like Algorithms, automata theory, and for his valuable suggestions throughout.

I would like to acknowledge my friends with whom I have spent the best days of my life during my school and college days. I would like to thank Amit, Baban, Chandra, Bikas, Sudhansu, Bibek, Dipu, Sambit, Saroj, Nab Bhai, Lalatendu, Pramod, Jyoti, Manas, Smruti, Lora Subrat, Satya to name a few, for their constant encouragement to carry out research.

Also I am very thankful to the colleagues and students of my college who have constantly provided moral support and encouraged me to complete my thesis work. Special thanks goes to Jnana, Goutam Bhai, Padma, Ranjita and Rekha.

I am also thankful to Nana, Bou, Bhaina and Didi for their blessings and support to complete my work.

Many thanks to my wife Yagnyaseni for her constant support to carry out my research. I would not have made it without her selfless love and tolerating me for a long period of being neglected because of my thesis. I am really indebted to her dedication, selfless love and affection.

Last but not the least, a special thanks to my family. I cannot express my feelings in words how grateful I am to my parents, my sisters for their sacrifices they have made on my behalf. Their support and prayers have strengthen me to overcome all the difficulties in my life. And finally, the one person who has made this all possible has been my mother Late Srimati Nayak. She has made an untold number of sacrifices for the entire family, and especially for me to continue my study. She has been the constant source of inspiration and her support and love towards me has made me the person today I am. I owe my deepest appreciation to her.

I thank the almighty who is above all of us, omnipresent God, for answering my prayers for giving me strength to trudge on each and every phase of my life.

March 2019 **Ananda Chandra Nayak**

Abstract

Let V be a finite and nontrivial alphabet. A word is a finite sequence of symbols drawn from V. A partial word is a word in which letters in some positions may not be known. Such a position is said to have a hole. A partial word w is said to be primitive if it is not contained in some power of a word. This thesis reports a theoretical investigation on the language of primitive partial words, a special type of bordered partial words, and palindromes in partial words.

We explore the position of the language of primitive partial words in conventional Chomsky hierarchy. We prove that the language of primitive partial words is not regular, not linear and not a deterministic context-free language. The pumping lemmata and the closure properties of language classes are used to prove these results. The language of nonprimitive partial words is shown to be a non-context-free language. We give a 2DPDA automaton to recognize the language of primitive partial words with one hole. We also prove that the language of nonprimitive words is an indexed language. We study several combinatorial properties such as denseness and reflectivity of the language of primitive partial words.

We study combinatorial properties that preserves primitivity in partial words with one hole under several point mutation operations such as deletion of a symbol, insertion of a symbol, substituting a symbol by another symbol, and exchanging two distinct consecutive symbols. The languages of non-del-robust primitive partial words with one hole and non-exchange-robust primitive partial words are proven to be non-context-free. A lower bound on the number of del-robust primitive partial words with one hole of length n over an alphabet of size k is given.

A partial word w is bordered if there exist words x, y, z such that w is contained in both xy and zx. For a given alphabet V, an involution θ is a mapping such that $\theta(\theta(u)) = u$ for all $u \in V^*$, and an antimorphic involution is an antimorphism, that is, $\theta(uv) = \theta(v)\theta(u)$ for all $u, v \in V^*$. We extend the concept of bordered partial words to pseudo-bordered partial words or θ-bordered partial words under the assumption that θ is either a morphic involution or an antimorphic involution. A partial word u is said to be θ-bordered if it is contained in xy and in $z\theta(x)$ for some $x, y, z \in V^+$. We present a comparative study of θ-bordered partial words and characterize the set of θ-bordered partial words when θ is an antimorphic involution. A necessary and sufficient condition for a partial word to be θ-unbordered which is based on θ-contained prefixes and suffixes is given. We do a counting of the number of θ-borders of a partial word $w = (u\theta(u))^i$ where u is unbordered as well as θ-unbordered partial word. The set of θ-unbordered partial words is shown to be a disjunctive language. It is proven that the set of all θ-unbordered partial words is regular when θ is an antimorphic involution, and the set of all θ-bordered partial words is not context-free when θ is a morphic involution. We also introduce the notion of θ-primitivity, θ-conjugacy and θ-commutativity in partial words. We provide a characterization of partial words x and y such that x is a θ-conjugate of y.

Lastly, we initiate a study of palindromes and pseudo-palindromes in partial words, and identify various properties in relation with primitive partial words. In particular, we give a bound on the number of palindromes in the conjugacy class of a primitive palindromic partial word. The number of θ-borders of a partial word u^j for some j when θ is a morphic involution is given. A lower bound on the number of θ-borders of some power of a θ-palindromic partial word with arbitrary number of holes is also given.

Contents

List of Abbreviations

lub	least upper bound
DNA	Deoxyriboneucleic Acid
CFL	Context-free Language
DCFL	Deterministic Context-free Language
CSL	Context-sensitive Language
DPDA	Deterministic Pushdown Automata
2DPDA	2-way Deterministic Pushdown Automata
2NPDA	2-way Nondeterministic Pushdown Automata
RAM	Random Access Machine
WK	Watson-Crick
gsm	generalized sequential machine

List of Symbols

$\mathbb{N}_0$	Natural numbers set $\cup \{0\}$
$\mathbb{Z}^+$	The set of positive integers
V	A finite alphabet
$\emptyset$	Empty set
ε	The empty string or null string
V^m	The set of all strings of length m over an alphabet V
V^*	$\bigcup_{i=0}^{\infty} V^i$
V^+	$V^* - \{\varepsilon\} = \bigcup_{i=1}^{\infty} V^i$
$w[i]$	The i^{th} symbol in a string w
$w[i..j]$	The substring of a string w that starts and ends at the indices i and j, respectively, that is, $w[i..j] = w[i]w[i+1]\ldots w[j]$
Q	The language of primitive words over an alphabet
Z	The language of nonprimitive words over an alphabet
$\Diamond$	Hole or do not know symbol
V_i^*	The language of partial words with at most i holes over an alphabet V
V_p^*	The language of partial words over an alphabet V
Q_p	The language of primitive partial words over an alphabet
$\overline{Q}_p$	The language of nonprimitive partial words over an alphabet
$\alpha(w)$	The set of symbols in w form the underlying alphabet
$\|w\|$	The number of symbols in w
$\|\alpha(w)\|$	The number of distinct symbols in w
Pref(w)	The set of all prefixes of the word w
Suff(w)	The set of all suffixes of the word w
$\sqrt{w}$	The root of a (partial)word w
$Conj(w)$	The set of all conjugates of a (partial)word w
$D(w)$	Domains of a partial word w
$H(w)$	The set of holes of a partial word w

$u \subset v$	The partial word u is contained in a partial word v
$u \uparrow v$	Partial words u and v are compatible
$u \vee v$	Least upper bound of partial words u and v
$Seq_{k,l}(i)$	The sequence of i relative to k and l
Q_p^i	The set of all primitive partial words with i holes
$\overline{Q}_p^i$	The set of nonprimitive partial words with i holes
$Q_p^{(i)}$	$\{u^i \mid u \in Q_p\}$
w^r	Reverse of a (partial) word w
X_k	The set of all (partial) words of length at most k
$V_1^*(s)$	$V_1^* - \{x\Diamond x \mid x \in V^*\}$
$Pref(w,i)$	The prefix of the word w of length i
$Suff(w,i)$	The suffix of the word w of length i
$L^{(i)}$	$L^{(i)} = \{w^i \mid w \in L\}$
L^i	$L^i = L \cdot L \cdots$ i times
$Q_p^1(m)$	$Q_p^1(m) = \{u^m \mid u \in Q_p^1\}$
Q_p^{1D}	The language of del-robust primitive partial words with one hole
$Q_p^{1\overline{D}}$	The language of primitive partial words with one hole which is not del-robust
$\mathcal{P}_{h,k}(n)$	The set of primitive partial words with h holes of length n over an alphabet of size k
$\mathcal{N}_{h,k}(n)$	The set of nonprimitive partial words with h holes of length n over an alphabet of size k
$T_{h,k}(n)$	Total number of partial words of length n with h holes over an alphabet of size k
$P_{h,k}(n)$	Number of primitive partial words with h holes of length n over an alphabet of size k
$N_{h,k}(n)$	Number of nonprimitive partial words with h holes of length n over an alphabet of size k
$Q_p^{1D}(n)$	The language of del-robust primitive partial words with one hole of length n
$Q_p^{1\overline{D}}(n)$	The language of non-del-robust primitive partial words with one hole of length n
Q_p^{1X}	The language of exchange-robust primitive partial words with one hole
$Q_p^{1\overline{X}}$	The language of non-exchange-robust partial words with one hole
$Q_p^{\overline{X}}$	The language of non-exchange-robust partial words
$one(x)$	$\{x_1bx_2 \mid x = x_1ax_2,\ x_1, x_2 \in V^*,\ a, b \in V,\ a \neq b\}$
$length(L)$	The set of lengths of the words in the language L

Q_p^{1S}	The language of subst-robust primitive partial words with one hole.
$Q_p^{1\overline{S}}$	The language of non-subst-robust primitive partial words with one hole
$\mathrm{CPref}(w)$	$\{x \in V^+ \mid \exists y \in V^*,\ w \subset xy\}$
$\mathrm{CSuff}(w)$	$\{x \in V^+ \mid \exists y \in V^*,\ w \subset yx\}$
$\mathrm{PCPref}(w)$	$\{x \in V^+ \mid \exists y \in V^+,\ w \subset xy\}$
$\mathrm{PCSuff}(w)$	$\{x \in V^+ \mid \exists y \in V^+,\ w \subset yx\}$
$u \leq_{cp} w$	$w \subset uv$ for some $u \in V^*, v \in V^+$
$u \leq_{cs} w$	$w \subset vu$ for some $u \in V^*, v \in V^+$
$u <_{cp} w$	$w \subset uv$ for some $u \in V^+, v \in V^+$
$u <_{cs} w$	$w \subset vu$ for some $u \in V^+, v \in V^+$
$u \leq_{cp}^{\theta_\Diamond} w$	$w \subset \theta(u)v$ for some $u, v \in V^*$
$u \leq_{cs}^{\theta_\Diamond} w$	$w \subset v\theta(u)$ for some $u, v \in V^*$
$\leq_d^{\theta_\Diamond}$	$\leq_{cp} \cap \leq_{cs}^{\theta_\Diamond}$
$u <_{cp}^{\theta_\Diamond} w$	$w \subset \theta(u)v$ for some $u \in V^*, v \in V^+$
$u <_{cs}^{\theta_\Diamond} w$	$w \subset v\theta(u)$ for $u \in V^*, v \in V^+$
$<_d^{\theta_\Diamond}$	$<_{cp} \cap <_{cs}^{\theta_\Diamond}$
$L_d^{\theta_\Diamond}(w)$	$\{u \in V^* \mid u <_d^{\theta_\Diamond} w\}$, the set of all θ-borders of a partial word w
$\nu_d^{\theta_\Diamond}(w)$	$\lvert L_d^{\theta_\Diamond}(w)\rvert$, the number of θ-borders of a partial word w
$D_{\theta_\Diamond}(i)$	$\{w \in V_\Diamond^+ \mid \nu_d^{\theta_\Diamond}(w) = i\}$, the set of partial words having i θ-borders
$D_\theta(1)$	The set of words which are θ-unbordered
$D_{\theta_\Diamond}(1)$	The set of all partial words which are θ-unbordered
$\alpha_{ub}(u)$	$\{v \in V^+ \mid uv \in D_\theta(1)\}$, θ-unbounded annihilator of a word u
$\alpha_{ub_\Diamond}(u)$	$\{v \in V_\Diamond^+ \mid uv \in D_{\theta_\Diamond}(1)\}$, θ-unbounded annihilator of a partial word u
$\alpha_{ub_\Diamond}(L)$	$\{u \in V_\Diamond^+ \mid Lu \subseteq D_{\theta_\Diamond}(1)\}$, θ-unbounded annihilator of a language L
$\mathcal{P}al_\Diamond$	The set of partial words which are also palindromes
$u \sqcup\!\sqcup v$	Shuffle of two words u and v, that is, if $u = u[1]\cdots u[n], v = v[1]\cdots v[n]$ then $u \sqcup\!\sqcup v = u[1]v[1]u[2]v[2]\cdots u[n]v[n]$
$\mathcal{P}al_{\theta_\Diamond}$	The set of all θ-palindromic partial words over an alphabet
Q_{p_θ}	The set of all θ-primitive partial words
$\rho_{\theta,\Diamond}(w)$	$\{x : x$ is a θ-primitive (total) word and $w \subset x\{x, \theta(x)\}^n,\ n \geq 0\}$.

$u \leq_c^{\theta_\Diamond} v$	There exist some $x \in V^*$ such that $v \subset ux$ and $v \subset \theta(x)u$.
$L_c^{\theta_\Diamond}(v)$	The set of all partial words that θ-commute with v, that is, $\{u : u \in V^*, u \leq_c^{\theta_\Diamond} v\}$.
$\nu_c^{\theta_\Diamond}(v)$	The number of partial words that θ-commute with v
$C_{\theta_\Diamond}(i)$	$\{v : v \in V_\Diamond^+, \nu_c^{\theta_\Diamond}(v) = i\}$ for $i \geq 1$.

List of Figures

List of Tables

Chapter 1

Combinatorics on Partial Words

1.1 Introduction

The origin of combinatorics on words goes back to the 20^{th} century and initiated by the work of Axel Thue on square-free words (words without having consecutive repetitions such as xx) [4]. A systematic study of words appeared in a series of Thue's papers where several combinatorial problems related to sequences of symbols were solved using the tools of discrete mathematics. The first infinite square-free word was given by Axel Thue in his work. His contributions were mainly concerned with repetitions in words, and usage of iterated morphisms for studying such repetitions.

The research in combinatorics on words has emerged due to development in discrete mathematics and theoretical computer science. The collective works of group of researchers under the pseudonym of Lothaire give an account of the topic in series of books; namely combinatorics on words [69], algebraic [70] and applied combinatorics on words [71]. It has numerous applications including the field of coding theory [5], formal language theory [47], graph theory [84] and semigroup theory [69]. Two fundamental properties of words, namely *periodicity* and *borderedness* play a vital role in many research areas including string algorithms [22, 26, 28, 42, 62, 63], theory of codes [5], data compression algorithm [89], computational biology [46], and serial data communication systems [23].

Primitive words are words that cannot be expressed as an integer power of another shorter word. Primitivity is one of the most fundamental and explored property of words. Primitive words play an important role in formal language theory, coding theory and combinatorics on words. The theory of primitive words has been extensively studied and many combinatorial properties have been identified; see for example [34, 54, 72, 78, 80, 81].

Partial words, a canonical extension of words, are words that may have some unknown symbols known as "holes" or "do not know" symbols and has been introduced by Berstel and Boasson [3]. The holes are represented by the symbol $\Diamond$. The motivation behind introducing partial words was stimulated by practical problems in Molecular biology such as in comparison of two genes. For example, alignment of two DNA sequences can be viewed as construction of two partial words which are compatible.

Another motivation for studying partial words is that partial words can be used to store the information in compact form known as universal partial words [24].

The book of Francine Blanchet-Sadri "*Algorithmic Combinatorics on Partial Words*" [9] gives an overview of works on partial words. Several concepts and results on partial words have been investigated, for example, conjugacy and commutativity [8,19], periodicity [7], Fine and Wilf's Theorem [7, 10, 17], unbordered partial words [14], avoidability of sets [12, 18, 20], critical factorization theorem [15] and equations on partial words [11]. Also partial words has been used as a tool for DNA encoding purposes [66]. The language of partial words has been investigated in several ways, for example, the language of partial words in relation to pcodes [6], punctured codes [65], and languages of partial words related to regular language of full words [30]. Recently, Blanchet-Sadri et al. [2] developed an approach that can compute the number of partial words having a set of strong and weak periods, and the number of partial words having a maximum border length. The problem of longest common compatible extension for each pair of positions (i, j) in a partial word has been studied in [16].

This thesis presents a theoretical investigation of problems on primitive partial words and a variation of bordered partial words known as pseudo-bordered partial words. In particular, following are the main goals of the thesis.

- Explore the position of the language of primitive partial words in the conventional Chomsky hierarchy.
- An indexed grammar for the language of nonprimitive words.
- Describe a Two-way Deterministic Pushdown Automata (2DPDA) that recognize the set of primitive partial words with one hole.
- Characterize and study combinatorial properties of the sets of primitive partial words which are robust under point mutation operations.
- Prove that the language of non-exchange-robust primitive partial words is not a context-free language over an alphabet.
- Initiate the study of θ-bordered partial words which is an extension of bordered partial words where θ is either a morphic or an antimorphic involution.
- Prove that the language of θ-unbordered partial words is a disjunctive language where θ is a morphic involution.
- Prove that the set of θ-bordered partial words is not a context-free language under the assumption of θ to be a morphic involution.
- Provide a characterization for θ-conjugacy and θ-commutativity of two partial words.
- Combinatorial properties of palindromes and θ-palindromes in partial words.

- Count the number of θ-borders of a partial word $u^j,\ j \geq 1$, where u is a primitive palindromic partial word with one hole under the assumption that θ to be a morphic involution.

1.2 Thesis Outline

The rest of the thesis is organized as follows.

Chapter 2: Preliminaries

In this chapter, we review the basic concepts on words and partial words. We mention the well-known results which are used later in the thesis. This chapter introduces the notations and definitions on partial words.

Chapter 3: The Language of Primitive Partial Words

In this chapter, we study the language of primitive partial words over a nontrivial alphabet and investigate the position of the language of primitive partial words. We prove that the language of primitive partial words is not regular, not linear and not a deterministic context-free language. The language of nonprimitive partial words is shown to be a non-context-free language. We describe a 2DPDA automaton that accepts the set of primitive partial words with one hole. We give an indexed grammar for the language of nonprimitive words over an alphabet and prove that the language of nonprimitive words is an indexed language.

Chapter 4: Robustness of Primitive Partial Words

In this chapter, we extend the notion of robustness of primitive words to primitive partial words. Under this setting, we consider various point mutation operations such as deletion of a symbol, insertion of a symbol, exchanging two distinct consecutive symbols and substitution of a symbol by another symbol. We give the structural characterization of each class of primitive partial words with one hole, and explore some important combinatorial properties related to each of the aforementioned classes. It is proved that the languages of non-del-robust primitive partial words with one hole and non-exchange robust primitive partial words are not context-free over a nontrivial alphabet. We also provide a lower bound on the number of non-del-robust primitive partial words with one hole of length n over an alphabet of size k.

Chapter 5: Pseudo-Bordered Partial Words

In this chapter, we consider the problem of θ-(un)bordered [1] partial words which is a combined extension of θ-(un)bordered words and bordered partial words where

[1] (un)bordered means either a bordered or an unbordered word.

θ is an involution. An involution θ is a mapping such that θ^2 is an identity, that is, $\theta^2(w) = w$ for every $w \in V^*$. A function θ over an alphabet V is said to be a morphism if $\theta(uv) = \theta(u)\theta(v)$ and an antimorphism if $\theta(uv) = \theta(v)\theta(u)$ for all words $u, v \in V^*$. A nonempty partial word w is said to be θ-bordered if there exist nonempty words x, y, z such that w is contained in xy and also contained in $z\theta(x)$. We identify several important combinatorial properties of θ-(un)bordered partial words under the assumption of θ to be (anti)morphic [2] involution. A necessary and sufficient condition for a partial word to be θ-unbordered is given in terms of contained prefixes and contained suffixes. For a given partial word u which is unbordered as well as θ-unbordered, we count the number of θ-borders of the partial word of the form $(u\theta(u))^i$ for $i \geq 1$. We also prove that the set of θ-unbordered partial words is a disjunctive language. The set of θ-unbordered partial words is shown to be a regular language under the assumption of θ being an antimorphic involution. It is proved that the language of θ-bordered partial words is not a context-free language when θ is a morphic involution. The θ-unbounded annihilator of a language of partial words is proved to be a post-plus language. We introduce the notion of θ-primitive partial words and study its basic properties. The concepts such as θ-conjugacy and θ-commutativity property of partial words is studied, and a characterization of partial words x and y such that x is a θ-conjugate of y is given.

Chapter 6: Palindromes and Pseudo-Palindromes

In this chapter, we study the notion of palindromes in partial words. Palindrome is a well studied topic in formal languages, automata theory and word combinatorics. It is known that the language of palindromes over an alphabet is a context-free language. A partial word u is said to be a palindrome if u is compatible to u^r where u^r is the reverse of u. For a partial word u, we provide necessary and sufficient condition for the concatenation of u and u^r to be nonprimitive. For an even length primitive palindromic partial word, there are exactly two palindromes in its conjugacy class. We introduce the notion of θ-palindromes in partial words and establish a relation between θ-palindromes and θ-bordered partial words for (anti)morphic involution θ. We count the number of θ-borders of a partial word u^j when θ is a morphic involution.

We conclude our thesis by giving a set of conclusions where the future directions of study are discussed.

[2] (anti)morphic means either a morphic involution or an antimorphic involution.

Chapter 2

Preliminaries

This chapter presents the definitions and basic concepts on total words and partial words that are used in the rest of this thesis. It states mathematical notations along with the existing results on words and partial words.

2.1 Words

An *alphabet* is a nonempty finite set whose elements are called as *symbols* or *letters*. We use V to denote an alphabet. If V is a singleton set then we call it trivial; otherwise we call it nontrivial. We assume V to be nontrivial throughout this thesis. A *total word* or *word* is a finite sequence of symbols drawn from V. We use the term "word" or "string" interchangeably. Formally, a word $w = a_0 \cdots a_{n-1}$ is a total function $w : \{0, \ldots, n-1\} \to V$ where $a_i \in V$ for all $i \in \{0, \ldots, n-1\}$.

The *length* of a word w over a given alphabet V is the total number of symbols that appear in w and is denoted by $|w|$. For a word w over an alphabet V, we use $\alpha(w)$ to denote the set of distinct symbols that appear in w, and $|\alpha(w)|$ is the number of distinct symbols in w. The word which does not contain any symbol is called the *empty word* and is denoted by ε; so $|\varepsilon| = 0$.

Example 2.1. Let $V = \{a, b, c\}$ be an alphabet. Then $w = bbaaa$ is a word of length 5 and $\alpha(w) = \{a, b\}$. □

For any $m \in \mathbb{N}_0$, V^m denotes the set of all strings of length m over V. We use V^* to denote the set of all finite length strings over V, and V^+ for the *set of all nonempty strings*. Therefore, $V^* = \cup_{m \geq 0} V^m$ and $V^+ = V^* \setminus \{\varepsilon\}$. A language L is a set of strings over an alphabet V, that is, $L \subseteq V^*$. For two words u and v, the *concatenation* of u and v is denoted by $u \cdot v$ (or simply uv) where '$\cdot$' is the juxtaposition or concatenation operator. This operation is associative, that is, for any $u, v, w \in V^*$, $(uv)w = u(vw)$. However, the operation is not commutative and in general for two strings u, v, $uv \neq vu$. Also, for any string $u \in V^*$, $u\varepsilon = \varepsilon u = u$. For a language $L \subseteq V^*$, we use $length(L) = \{|x| \mid x \in L\}$.

Prefix, Suffix and Factor

Let $w = xyz$ be a word where $x, y, z \in V^*$. We call y as a *factor* of the word w, and it is a *central factor* if $x \neq \varepsilon$ and $z \neq \varepsilon$. The word y is said to be a *prefix* (respectively, a suffix) of the word w if $x = \varepsilon$ (respectively, $z = \varepsilon$). The i^{th} symbol of a word w is denoted by $w[i-1]$ for $1 \leq i \leq |w|$. We use $w[i..j]$ to represent the factor that starts and ends at the positions i and j, respectively for $0 \leq i \leq j \leq |w|-1$, that is, $w[i..j] = w[i]w[i+1] \ldots w[j]$. Thus, $w[0..j-1]$ is the prefix of length j of a word w, and the word $w[j..|w|-1]$ of length $|w|-j$ is the suffix of w.

For a word w, we denote $Pref(w) = \{u \in V^+ \mid \exists v \in V^*, w = uv\}$ and $Suff(w) = \{u \in V^+ \mid \exists v \in V^*, w = vu\}$ as the set of all prefixes and suffixes, respectively. Similarly, the set of proper prefixes and proper suffixes of a word w is defined as $PPref(w) = \{u \in V^+ \mid \exists v \in V^+, w = uv\}$ and $PSuff(w) = \{u \in V^+ \mid \exists v \in V^+, w = vu\}$, respectively.

Primitive and Periodic Words

For a word u, the n^{th} *power* of u is defined inductively as $u^0 = \varepsilon$ and $u^n = u^{n-1} \cdot u$ for $n \geq 1$. A word u is said to be *primitive* if it cannot be expressed as an integer power of a shorter word, that is, $u = v^n$ implies that $u = v$ and $n = 1$. Note that ε is not a primitive word as it can be expressed as $\varepsilon^2 = \varepsilon$. The set of primitive words over an alphabet V is denoted as Q_V, or simply Q when V is understood. If $u = v^n$ and $v \in Q$ then v is called the *primitive root* of u and is denoted by $\sqrt{u}$. Observe that when u is a primitive word then $\sqrt{u} = u$. We use the notation Z_V, or simply Z when V is understood, to denote the set of all *nonprimitive words* over an alphabet V. For example, $w = abab = (ab)^2$ is a nonprimitive word, whereas $u = abb$ is a primitive word over the alphabet $\{a, b\}$.

For a word $w = a_0 \ldots a_{n-1}$ where $a_i \in V$ and $0 \leq i \leq n-1$, an integer p is said to be a period of w if $a_i = a_{i+p}$ for $0 \leq i < n-p$. A word can have several periods. The smallest period of a word w is called the *period* of w.

Example 2.2. Consider a word $w = abbabbabb$ over an alphabet $V = \{a, b\}$. The periods of w are 3 and 6. The smallest period is 3 and hence the *period* of w is 3. □

Conjugacy, Commutativity and Borders

Conjugacy and commutativity are two important word properties which are used to solve many word equations (see for example, [41] and [72]). For a group G operating on a set S, an element $x \in S$ is conjugate to another element $y \in S$ if there exists an element $z \in G$ such that $x = zy$. Similarly, two words u and v are said to conjugate if there exist $x, y \in V^*$ such that $u = xy$ and $v = yx$. We use $u \sim v$ to denote the conjugacy relation between two words u and v. Observe that for a word w, its conjugates are obtained by cyclic permutations of letters in w. The set of all conjugates of a word w is denoted by $Conj(w)$. The conjugacy relation $\sim$ is an equivalence relation. The number of conjugates in a conjugate class are different for a primitive and a nonprimitive word. For a primitive word w of length n, $Conj(w)$ has exactly n elements. For a nonempty word $w \in Z$, there exists a unique primitive word x such that $w = x^n$, and hence $|Conj(w)| = |x|$.

Example 2.3. Let $V = \{a, b\}$ be the alphabet. Let $w_1 = aabab$ be a primitive word and $w_2 = abab$ be a nonprimitive word. Then $Conj(w_1) = \{aabab, baaba, abaab, babaa, ababa\}$ and $Conj(w_2) = \{abab, baba\}$. □

Two nonempty words x and y are said to commute if $xy = yx$. The following example illustrates the commutative property of two words.

Example 2.4. Let $x = ab$ and $y = abab$ be two nonempty words. It is easy to observe that $x \cdot y = y \cdot x$ and hence x and y commute to each other. □

A word w is said to be bordered if there exists a nonempty proper prefix of w which is also a suffix. Formally, a word w is bordered if and only if there exist nonempty words x, y, z such that $w = xy = yz$ and y is said to be a border of w. A nonempty word which is not bordered is called as unbordered. Such words turn out to be primitive, that is, an unbordered word cannot be written as a power of another word. A word can have several borders. A border $x \in V^+$ of a word w is called a minimal border if for any $y \in V^+$ such that $|y| < |x|$ then y is not a border of w.

Example 2.5. Let $w = abababa$ be a word over the alphabet $V = \{a, b\}$. The nonempty borders of w are a, aba and $ababa$ where a is the minimal border of w. □

Morphisms

A *homomorphism* or *morphism* is a mapping that maps every letter to a string and is then extended to an arbitrary string by concatenation operation. Formally, a function $h : V_1^* \to V_2^*$ is a morphism if $h(xy) = h(x)h(y)$ for all $x, y \in V_1^*$. Note that $h(\varepsilon) = \varepsilon$.

Example 2.6. Let $V_1 = V_2 = \{a, b\}$ and h be a morphism such that $h(a) = ba$ and $h(b) = b$. For the word $w = baab$, we have $h(w) = h(baab) = h(b)h(a)h(a)h(b) = bbabab$. □

Definition 2.7. *Let $h : V_1^* \to V_2^*$ be a homomorphism. Then h is said to be*

(a) nonerasing if $h(a) \neq \varepsilon$ for all $a \in V_1$.

(b) uniform if $|h(a)| = |h(b)|$ for all $a, b \in V_1$.

(c) injective if $h(x) = h(y)$ then $x = y$ for all $x, y \in V_1^*$.

Palindrome

Let $w = a_0 \cdots a_{n-1}$ be a word. We use the symbol w^r to denote the reverse of a word w and $w^r = a_{n-1} \cdots a_0$. The word w is called a palindrome if $w = w^r$. For example, the word $w = abba$ is a palindrome.

Definition 2.8 (Reflective Language)**.** *A language L is said to be reflective if $xy \in L$ implies that $yx \in L$ for all $x,\ y \in V^*$.*

2.2 Partial Words

A partial word is a word in which letters at some positions may be unknown. Such positions are known as "holes" or "do not know" symbols. A partial word w of length n over an alphabet V can be defined as a partial function $w : \{0, 1, \ldots, n-1\} \rightarrow V$. For $i \in \{0, 1, \ldots, n-1\}$, if $w[i]$ is defined then we say that i belongs to domain of w denoted by $D(w)$; otherwise we say i belongs to the set of holes of w denoted by $H(w)$. We use the symbol $\Diamond$ to denote a hole and $\Diamond$ is not a symbol in alphabet V. A total word is a partial word with empty set of holes.

Example 2.9. Let $V = \{a, b\}$ be the alphabet. Let $w = ab\Diamond ba\Diamond b$ be a partial word. Then the domain and set of holes of w are $D(w) = \{0, 1, 3, 4, 6\}$ and $H(w) = \{2, 5\}$, respectively. □

A partial word can be seen as a total word over the extended alphabet $V \cup \{\Diamond\}$. We use $V_\Diamond$ for this extended alphabet. The concatenation of two partial words is defined as it is for total words.

We denote $V_\Diamond^*$ as the language of partial words with arbitrary number of holes and $V_\Diamond^+ = V_\Diamond^* \setminus \{\varepsilon\}$ is the set of nonempty partial words. The length of a partial word w is the number of symbols in w including the hole $\Diamond$. The notation $\alpha(w)$, for a partial word w, denotes the set of distinct symbols from the alphabet V excluding $\Diamond$, and $|\alpha(w)|$ is the number of distinct symbols from the alphabet V. Two partial words u and v are said to be equal if $H(u) = H(v)$ and $D(u) = D(v)$, and $u(i) = v(i)$ for all $i \in D(u)$. We define *containment* and *compatibility* which are useful in study of partial words.

Definition 2.10. [Containment & Compatibility] *Let u and v be two partial words over an alphabet V and $|u| = |v|$.*

(a) The partial word u is said to be contained in partial word v if $D(u) \subseteq D(v)$ and $u(i) = v(i)$ for all $i \in D(u)$. The containment relation is denoted by $\subset$.

(b) The partial words u and v are said to be compatible if there exists a partial word w such that $u \subset w$ and $v \subset w$. The notation $\uparrow$ is used to denote the compatibility relation between two partial words.

Note that containment relation is not symmetric whereas compatible relation is symmetric but not transitive. For two compatible partial words u and v, a partial word w can be constructed in a way such that the domain of w is the union of the domains of u and v. Such a word is called as the **least upper bound** (lub) of the partial words u and v and is denoted by $u \vee v$.

Example 2.11. Let $V = \{a, b\}$ be an alphabet and $u = a\Diamond b\Diamond ba, v = abb\Diamond\Diamond a$ be two partial words. Then $D(u) = \{0, 2, 4, 5\}$, $D(v) = \{0, 1, 2, 5\}$. Here u is compatible to v but u is not contained in v as $D(u) \nsubseteq D(v)$. The least upper bound of u and v is the partial word $abb\Diamond ba$. □

As partial words can be viewed as total words over the extended alphabet, the terms prefix, suffix and factor are defined as for total words.

Periods of a Partial Word

Unlike the periods of a total word, there are two different types of periods of a partial word; (i) strong period or period, and (ii) weak (local) period. Let $w = a_0 \dots a_{n-1}$ be a partial word of length n. A *strong period* of a partial word w is a positive integer p such that $a_i = a_j$ whenever $i, j \in D(w)$ and $i \equiv j \ mod \ p$ [1]. A *weak period* or *local period* of a partial word w is a positive integer p such that $a_i = a_{i+p}$ for all $0 \leq i < n - p$ whenever $i, i + p \in D(w)$. There may be several strong periods or local periods of a partial word. If a partial word w has a strong period p, we call w is p-periodic, and if w has a local period p, we call w is locally p-periodic. A period (local period) k of a partial word is said to be minimal if all other periods (local periods) are larger than k.

Example 2.12. Let $w = abc\Diamond\Diamond cacc$ be a partial word over the alphabet $V = \{a, b, c\}$. Here w is locally 3-periodic but is not 3-periodic. □

It is worthy to mention some important differences between words and partial words.

1. Every locally m-periodic word is m-periodic.
2. Even if the length of a partial word w is multiple of a local period of w, w need not to be a power of a shorter partial word.

The set of all strong periods of a partial word w is denoted by $\mathcal{P}(w)$ and the minimal strong period is denoted by $p(w)$.

Example 2.13. Consider a partial word $w = \Diamond aba\Diamond baa$ over the alphabet $V = \{a, b\}$. Then $\mathcal{P}(w) = \{3, 6\}$ and $p(w) = 3$. □

Primitive Partial Word

A partial word u is said to be primitive if there does not exist a word v such that $u \subset v^n$ with $n \geq 2$ [8]. Note that if u is primitive and $u \subset w$ then w is also primitive.

Example 2.14. A partial word $w = ab\Diamond a$ is a primitive partial word over the alphabet $V = \{a, b\}$. □

We use the symbols Q_p and $\overline{Q}_p$(or, Z_p) to denote the language of primitive partial words and language of nonprimitive partial words over an alphabet.

The root of a partial word is a set of total words. The root of a partial word u is defined as follows:

$$\sqrt{u} = \{v \mid v \text{ is a total primitive word and } u \subset v^n \text{ for some } n \geq 1\}$$

[1] $i \equiv j \ mod \ p$ means $i - j$ is divisible by p, i.e., i and j leaves the same remainder when divided by p.

Unlike for a total word, the root of a partial word is not unique. For example, consider a partial word $u = a\Diamond ab$ over the alphabet $V = \{a, b\}$. Then $u \subset abab = (ab)^2$ and also $u \subset aaab$. Hence $\sqrt{u} = \{ab, aaab\}$.

Next we recall the following rules which are useful to deal with the computation in partial words.

Lemma 2.15 ([9]). *Let u, v, w, x and y be partial words.*

(a) Multiplication: If $u \uparrow v$ and $x \uparrow y$ then $ux \uparrow vy$.

(b) Simplification: If $ux \uparrow vy$ and $|u| = |v|$ then $u \uparrow v$ and $x \uparrow y$.

(c) Weakening: If $u \uparrow v$ and $w \subset u$ then $w \uparrow v$.

Lemma 2.16 ([9]). *Let $u, v, x, y \in V_\Diamond^*$ such that $ux \uparrow vy$.*

(a) If $|u| \geq |v|$ then there exist partial words w and z such that $u = wz$, $v \uparrow w$ and $y \uparrow zx$.

(b) If $|u| \leq |v|$ then there exist partial words w and z such that $v = wz$, $u \uparrow w$ and $x \uparrow zy$.

Similar to total words, the concepts of conjugacy, commutativity and borders have been defined in partial words. Let u and v be two partial words. Then u and v are *conjugate* if there exist partial words x and y such that $u \subset xy$ and $v \subset yx$. For example, the partial words $u = ab\Diamond\Diamond b$ and $v = b\Diamond bab$ are conjugate because $u \subset xy$ and $v \subset yx$ for $x = ab$ and $y = bab$. Two partial words x and y are said to *commute* if $xy \uparrow yx$.

Blanchet-Sadri et al. [14] extended the definition of border to partial words. A nonempty partial word u is said to be bordered if one of its proper prefix is compatible to a suffix of u of the same length. A partial word u is unbordered if there exist no nonempty words x, y and z such that $u \subset xy$ and $u \subset zx$. Unbordered partial words also turn out to be primitive [8]. A nonempty word x is a border of a partial word u if there exist $y, z \in V^*$ such that $u \subset xy$ and $u \subset zx$. A border x of a partial word u is said to be minimal if for all words y with $|y| < |x|$ implies that y is not a border of u.

Example 2.17. For the partial word $w = ab\Diamond b\Diamond ab$, border of w is ab. Hence w is a bordered partial word. □

The combinatorial results such as commutativity and conjugacy are not true in case of partial words with arbitrary holes. To accommodate several results from total words to partial words, the structure of partial words through a sequence of positions need to be observed carefully. Thus, some special kinds of partial words have been defined and are taken from [9].

Definition 2.18. *Let k and l be two positive integers satisfying $k \leq l$ and $z = xy$ be a partial word with $|x| = k$ and $|y| = l$. For $0 \leq i < k + l$, we define the sequence of i relative to k, l as $\mathbf{seq}_{k,l}(i) = (i_0, i_1, \ldots, i_{n+1})$ where*

(a) $i_0 = i = i_{n+1}$

(b) *For* $1 \leq j \leq n$, $i_j \neq i$

(c) *For* $1 \leq j \leq n+1, i_j$ *is defined as,* $i_j = \begin{cases} i_{j-1} + k & \text{if } i_{j-1} < l, \\ i_{j-1} - l & \text{otherwise.} \end{cases}$

Definition 2.19 (***(k,l)-special*** partial word)**.** *Let k, l be two positive integers satisfying $k \leq l$ and let u be a partial word of length $k+l$. We say that u is* ***(k,l)-special*** *if there exists $0 \leq i < gcd(k,l)$ such that $seq_{k,l}(i) = (i_0, i_1, \ldots, i_{n+1})$ contains at least two positions that are holes of u while $u[i_0]\ u[i_1] \cdots u[i_{n+1}]$ is not 1-periodic.*

The following example illustrates the above definitions.

Example 2.20. Let $u = cac\Diamond c\Diamond cb\Diamond b$ be a partial word and let $k = 4$ and $l = 6$. As the gcd of 4 and 6 is 2, we compute the $seq_{4,6}(0)$ and $seq_{4,6}(1)$. Then we have, $seq_{4,6}(0) = (0, 4, 8, 2, 6, 0)$ and $seq_{4,6}(1) = (1, 5, 9, 3, 7, 1)$.

$$seq_{4,6}(0) = u(0)u(4)u(8)u(2)u(6)u(0) = cc\Diamond ccc$$

$$seq_{4,6}(1) = u(1)u(5)u(9)u(3)u(7)u(1) = a\Diamond b\Diamond\Diamond ba$$

The first sequence $seq_{4,6}(0)$ does not satisfy the conditions while the sequence $seq_{4,6}(1)$ has two holes and is also not 1-periodic. Hence the partial word $u = cac\Diamond c\Diamond cb\Diamond b$ is $(4,6)$-special. □

Definition 2.21 ($\{k,l\}$-special partial word)**.** *Let k, l be two positive integers such that $k \leq l$ and $u \in V_\Diamond^*$ with $|u| = k + l$. Then u is said to be* **$\{k,l\}$-special** *if there exists $0 \leq i < gcd(k,l)$ such that either of the conditions hold:*

(a) *$seq_{k,l}(i)$ has two positions that are holes and is not 1-periodic.*

(b) *$seq_{k,l}(i)$ has two consecutive positions that are holes.*

Remark 2.22. *If a partial word u is (k,l)-special then it is also $\{k,l\}$-special.*

Thus the example given in Example 2.20 is also $\{k,l\}$-special. But the converse does not hold true and shown in the following example.

Example 2.23. Let $u = cacac\Diamond ca\Diamond\Diamond$ be a partial word and $k = 4$ and $l = 6$. Thus, calculating $seq_{4,6}(0)$ and $seq_{4,6}(1)$, we have

$$seq_{4,6}(0) = u(0)u(4)u(8)u(2)u(6)u(0) = cc\Diamond ccc$$

$$seq_{4,6}(1) = u(1)u(5)u(9)u(3)u(7)u(1) = a\Diamond\Diamond aaa$$

Observe that $seq_{4,6}(1)$ contains two consecutive positions which are holes of u, and hence u is $\{4,6\}$-special. Since both of the sequences $seq_{4,6}(0)$ and $seq_{4,6}(1)$ are 1-periodic then u is not (k,l)-special. □

2.3 Results on Words

Several results on total words are known. We recall some of the existing results on primitive words.

Lemma 2.24 (Lyndon and Schützenberger [72]). *For every nonempty word w, there exists a unique primitive word x and a unique integer $n \geq 1$ such that w can be expressed as $w = x^n$.*

The following theorem shows the necessary and sufficient condition in which a nonempty word will have an identical prefix and suffix.

Theorem 2.25 (Lyndon and Schützenberger [72]). *Let x, y, z be nonempty words. Then $xy = yz$ if and only if there exist words $u \in V^+$, $v \in V^*$ such that $x = uv$, $z = vu$ and $y = (uv)^n u$ for some integer $n \geq 0$.*

The above theorem also shows an alternative way to check whether two words x and z are conjugates of each other. Two nonempty words x and z are said to conjugate if there exists a word $y \in V^*$ such that $xy = yz$. The next result deals with the commutativity of two words.

Theorem 2.26 (Lyndon and Schützenberger [72]). *Let $u, v \in V^+$. Then $uv = vu$ if and only if there exists a primitive word $x \in Q$ such that $u = x^m, v = x^n$ for some integers m, n.*

Next we recall an important result on periodicity which is also known as periodicity theorem of Fine and Wilf.

Theorem 2.27 (Fine and Wilf [41]). *Let u and v be nonempty words with $|u| = m, |v| = n$, and $d = gcd(m, n)$. If u^i and v^j for some $i, j \in N$ share a common prefix (or common suffix) of length at least $m + n - d$, then u and v are powers of a common primitive word of length d.*

Fine and Wilf's periodicity theorem can be restated in the following way. Let u be a word. If for some $p, q \in \mathbb{N}^+$, u is p-periodic and q-periodic and $|u| \geq p+q-gcd(p, q)$ then u is $gcd(p, q)$-periodic. The length $p + q - gcd(p, q)$ is optimal. For example, let $w = aaabaaa$ over $V = \{a, b\}$ and $|w| = 7$. The word w is having periods 4 and 6 but w does not have period $2 = gcd(4, 6)$. Hence $|w| = p_1 + p_2 - gcd(p_1, p_2)$ is optimal.

The next result deals with a word equation by Lyndon and Schützenberger which shows that the equation has only trivial solution.

Theorem 2.28 (Lyndon and Schützenberger [72]). *Let $u, v, w \in V^+$. The word equation $u^i v^j = w^k$ with $i, j, k \geq 2$ holds if and only if there exists a word z such that u, v and w are powers of z.*

The following result is a direct consequence of Theorem 2.28.

Corollary 2.29. *Let u and v be two primitive words such that $u \neq v$. Then $u^i v^j \in Q$ for all $i, j \geq 2$.*

Lemma 2.30 (Shyr and Yu [87]). *If $u, v \in Q$ with $u \neq v$ then there is at most one nonprimitive word in the language u^+v^+.*

The next proposition shows that a primitive word u cannot be a proper factor of uu.

Proposition 2.31 (Choffrut and Karhumäki [25]). *Let u be a word. Then u is primitive if and only if u is not a factor of uu, that is, $uu = xuy$ implies that either $x = \varepsilon$ or $y = \varepsilon$.*

The following theorem shows that if two words x and y are conjugates and x is a power of some primitive word then y is also same power of a primitive word.

Theorem 2.32 (Shyr and Thierrin [86]). *Let $x = u_1u_2$ and $y = u_2u_1$ be conjugates. Then $x = p^i$ for some $p \in Q$ if and only if $y = q^i$ for some $q \in Q$.*

2.4 Results on Partial Words

In this section, we recall some of the existing results on partial words which are used later in the thesis. The following result for partial words is an extension of Theorem 2.25.

Lemma 2.33 (Blanchet-Sadri and Luhmann [19]). *Let x, y be nonempty partial words, and z be a word. If $xz \uparrow zy$ then there exist words u, v such that $x \subset uv$, $y \subset vu$ and $z \subset (uv)^i u$ for some $i \geq 0$.*

Lemma 2.33 does not hold true if z is not a full word. For example, let $x = b$, $y = a$ and $z = \diamondsuit aa$. Then $xz = b\diamondsuit aa$ and $zy = \diamondsuit aaa$ and $xz \uparrow zy$. If there exist full words u and v such that $x \subset uv$, then it must be that $b = uv$. This makes it impossible for $y = a \subset vu$.

The next result deals with the conjugacy of two partial words similar to total words but under the assumption that the partial word $xz \vee zy$ is $|x|$-periodic.

Theorem 2.34 (Blanchet-Sadri and Luhmann [19]). *Let x, y and z be nonempty partial words. If $xz \uparrow zy$ and $xz \vee zy$ is $|x|$-periodic then there exist partial words u, v such that $x \subset uv$, $y \subset vu$ and $z \subset (uv)^n u$ for some $n \geq 0$.*

The above result is illustrated by the following example.

Example 2.35. Let $x = \diamondsuit ba$, $y = \diamondsuit b \diamondsuit$ and $z = b\diamondsuit ab\diamondsuit\diamondsuit\diamondsuit\diamondsuit$. Then we have

$$xz = \diamondsuit bab\diamondsuit ab\diamondsuit\diamondsuit\diamondsuit\diamondsuit$$

$$zy = b\diamondsuit ab\diamondsuit\diamondsuit\diamondsuit\diamondsuit\diamondsuit b\diamondsuit$$

$$xz \vee zy = bbab\diamondsuit ab\diamondsuit\diamondsuit b\diamondsuit$$

Observe that $xz \uparrow zy$ and $xz \vee zy$ is $|x|$-periodic. Putting $u = bb$ and $v = a$, we can verify that $x \subset uv, y \subset vu$, and $z \subset (uv)^n u$ for some $n \geq 0$. $\square$

Similar to the commutativity result of two words, the following theorem states the condition for two partial words u and v commute each other such that uv has at most one hole.

Theorem 2.36 (Berstel and Boasson [3]). *Let u and v be nonempty partial words such that uv has at most one hole. If $uv \uparrow vu$ then there exist a word $w \in V^+$ such that $u \subset w^m$ and $v \subset w^n$ for some $m, n \geq 1$.*

The above result does not hold for partial words having two or more holes. For example, consider the partial words $u = a\Diamond b$ and $v = \Diamond abb$. Then, $uv = a\Diamond b\Diamond abb \uparrow \Diamond abba\Diamond b = vu$, but it is not possible for u and v to be contained in powers of a common word.

The following result is about the commutativity equation in partial words having two or more holes.

Theorem 2.37 (Blanchet-Sadri and Luhmann [19]). *Let u and v be nonempty partial words such that neither uv nor vu is $\{|u|, |v|\}$-special. If $uv \uparrow vu$ then there exists a word w such that $u \subset w^m$ and $v \subset w^n$ for some integers m and n.*

The next two results provide other sufficient conditions for two words u and v to commute each other.

Lemma 2.38 (Blanchet-Sadri and Luhmann [19]). *Let u, v be nonempty words and w be a partial word with at most one hole. If $w \subset uv$ and $w \subset vu$ then $uv = vu$.*

The above result is not true in case of partial words having two or more holes. For example, consider the partial word $w = a\Diamond b\Diamond$ over the alphabet $V = \{a, b\}$. For $u = abb$ and $v = a$, we have $w \subset uv$ and $w \subset vu$ but $uv \neq vu$.

Lemma 2.39 (Blanchet-Sadri and Luhmann [19]). *Let u, v be nonempty words and w be a non $\{|u|, |v|\}$-special partial word. If $w \subset uv$ and $w \subset vu$ then $uv = vu$.*

The next result extends the well-known result of Fine and Wilf's periodicity theorem for partial words.

Theorem 2.40 (Blanchet-Sadri and Hegstrom [17]). *Let u be a partial word and u is locally p-periodic and locally q-periodic.*

(a) If u is having $2n$ holes and $|u| \geq (n+1)(p+q) - gcd(p, q)$, then u is $gcd(p, q)$-periodic.

(b) If u is having $(2n+1)$ holes and $|u| \geq (n+1)(p+q)$, then u is $gcd(p, q)$-periodic.

The bound $(n + 1)p + q - gcd(p, q)$ and $(n + 1)(p + q)$ turns out to be optimal. For example, $aaaabaaaa\Diamond aa$ has one hole, is locally 5-periodic and locally 8-periodic, has length $12 = 5 + 8 - 1$, but is not 1-periodic.

Similar to the language of primitive words, the language of primitive partial words is also closed under the cyclic permutation operation which is stated in the next result.

Proposition 2.41 (Balnchet-Sadri [8]). *Let u, v be partial words. If there exists a word $x \in Q$ and for some integer n such that $uv \subset x^n$ then $vu \subset y^n$ for some primitive word $y \in Q$.*

By Proposition 2.31, it is known that a word u is primitive if it is not a central factor of uu. Similar result also holds for partial words with one hole.

Proposition 2.42 (Balnchet-Sadri [8]). *Let u be a partial word with one hole. Then u is primitive if and only if $uu \uparrow xuy$ for some partial words x, y implies that $x = \varepsilon$ or $y = \varepsilon$.*

The above proposition does not hold in case of partial words having two or more holes. For example, let $u = a\Diamond b\Diamond$ be a primitive partial word. Then $uu = a\Diamond b\Diamond a\Diamond b\Diamond \uparrow xuy$ where $x = a$ and $y = \Diamond b\Diamond$. Hence u is compatible to a central factor of uu with $x \neq \varepsilon$ and $y \neq \varepsilon$ but u is a primitive partial word.

The following lemma is the extension of Proposition 2.42 and holds for primitive partial words having two or more holes.

Lemma 2.43 (Balnchet-Sadri [9]). *Let u be a partial word having two or more holes.*

(a) If $uu \uparrow xuy$ for some partial words x, y implies $x = \epsilon$ or $y = \epsilon$, then u is primitive.

(b) If $uu \uparrow xuy$ for some nonempty partial words x and y satisfying $|x| \leq |y|$, then the following hold:

  (i) If $|x| = |y|$, then u is not primitive.

  (ii) If u is not $(|x|, |y|)$-special, then u is not primitive (it is contained in a power of a word of length $|x|$).

  (iii) If u is $(|x|, |y|)$-special, then u is not contained in a power of a word of length $|x|$.

The following lemma states that for a partial word u with one hole having two or more distinct letters, at least one of u or ua is a primitive partial word.

Lemma 2.44 (Balnchet-Sadri [8]).

(i) Let u be a partial word with one hole such that $|\alpha(u)| \geq 2$. If a is any letter then u or ua is primitive.

(ii) Let u_1, $u_2 \in V_\Diamond^+$ such that u_1u_2 has one hole and $|\alpha(u_1u_2)| \geq 2$. If a is any letter then u_1u_2 or u_1au_2 is primitive.

Chapter 3

The Language of Primitive Partial Words

In this chapter, we investigate the position of the language of primitive partial words in conventional Chomsky hierarchy.

3.1 Introduction

The important role of primitive words across several research areas including formal language theory [47,84], coding theory [5,86], string matching [28], word combinatorics [69–71] and many others has drawn attention of researchers. Recall that, a nonempty word u is said to be primitive if it cannot be expressed as an integer power of another word, that is, if $u = v^n$ then it implies that $u = v$ and $n = 1$.

The language of primitive words Q has been extensively studied and many facts have been proved regarding the relation of Q with the traditional formal language classes. The language of primitive words Q is known to be not regular [33], not linear [50], not a Deterministic Context-free Language (DCFL) [68] but it is a Context-sensitive Language (CSL) [64]. However, it is still an open question whether the language of primitive words Q is context-free or not [34, 37, 67, 81]. There is a linear time algorithm based on Proposition 2.31 to test whether a given word u is primitive.

Organization. The outline of the rest of this chapter is as follows. The next section contains some fundamental concepts and describes the hierarchy of languages of partial words. Section 3.3 contains some basic results on primitive partial words. In Section 3.4, we show that the language of primitive partial words with at most one hole is a dense language. In Section 3.5, we prove that the language of primitive partial words is not regular, not linear and not a deterministic context-free language. In Section 3.6, we give a 2DPDA automaton that accepts the language of primitive partial words with one hole. An indexed grammar for the language of nonprimitive words is given in Section 3.7. Section 3.8 contains some conclusions.

3.2 Hierarchy of Partial Words

A partial word u is said to be primitive if there does not exist a word v such that $u \subset v^n,\ n \geq 2$. Note that if u is a primitive partial word and $u \subset v$, then v is primitive as well [8]. We denote V_i^* as the set of partial words over V with at most i holes, and the language of partial words is denoted by V_p^* or $V_\Diamond^*$. We use V_p^* and $V_\Diamond^*$ interchangeably. Observe that the following relation holds.

$$V_0^* \subset V_1^* \subset V_2^* \subset \cdots \subset V_i^* \subset \cdots$$

where $V_0^* = V^*$. Thus, $V_p^* = \bigcup_{i \geq 0} V_i^*$. Thus, we can have

$$V_p^* = V_0^* \cup V_1^* \cup V_2^* \cup V_3^* \cup \cdots$$

We use the notation Q_p^i to denote the language of primitive partial words with at most i holes. We know that $Q \cup \overline{Q} = V^*$ where $\overline{Q}$ is the set of all nonprimitive words.

The root of a partial word w is defined as follows,

$$\sqrt{w} = \{p \mid p \text{ is primitive and total and there exists } k \text{ such that } w \subset p^k\}.$$

Hence a partial word has a set of total words as root. For a language L of partial words, we define $\sqrt{L} = \{u \mid w \in L \text{ and } u \in \sqrt{w}\}$.

For example, consider a partial word $w = a\Diamond$ over the alphabet $V = \{a,\ b\}$. Since $a\Diamond \subset ab$ and $a\Diamond \subset a^2$, hence $\sqrt{a\Diamond} = \{ab, a\}$.

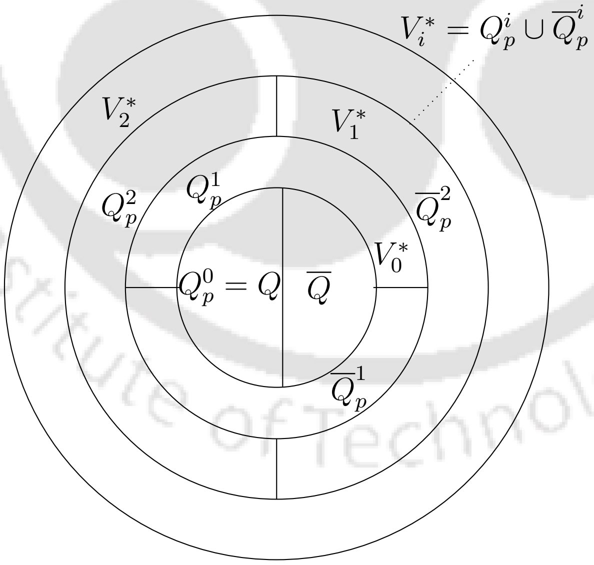

Figure 3.1: The hierarchy of languages of partial words

We denote the language of primitive partial words and the language of nonprimitive partial words as Q_p and $\overline{Q}_p$ respectively. Therefore,

$$Q_p = Q_p^0 \cup Q_p^1 \cup Q_p^2 \cup Q_p^3 \cup \cdots$$

and $Q_p \cup \overline{Q}_p = V_p^*$. The language of partial words with at most i holes can be viewed as hierarchy of formal languages. Figure 3.1 shows the hierarchy of languages of partial words.

3.3 Properties of Primitive Partial Words

In this section, we discuss some basic concepts such as bordered partial word and reverse of a partial word. We present some results in reference to these properties.

Let w be a partial word. The reverse of the partial word w is denoted by w^r and it is obtained by reversing the order of the symbols in w. For example, the reverse of $w = a\Diamond ba$ is $w^r = ab\Diamond a$. We see below a result that relates the reverse of a partial word with primitivity.

Proposition 3.1. *A partial word $w \in V_p^*$ is primitive if and only if w^r is also primitive.*

Proof. Let $w \in V_p^*$ be a primitive partial word. Suppose, the reverse of w, that is, w^r is not primitive. So, $w^r \subset z^k$ for some $z \in Q$ and $k \geq 2$. We can write $w = (w^r)^r$. So, $w = (w^r)^r \subset (z^k)^r = (z^r)^k$ for $k \geq 2$ which is a contradiction. Hence w^r is primitive.

The other direction can be proved similarly. □

The Proposition 2.41 shows that the language of primitive partial words Q_p is reflective.

Corollary 3.2. *The language Q_p is reflective.*

Corollary 3.3. *The language $\overline{Q}_p$ is reflective.*

Proof. Let $w = uv$ be a nonprimitive partial word. Let $w = uv \subset x^n$ for $n \geq 2$ but $vu \notin \overline{Q}_p$. If vu is primitive then by Proposition 2.41 uv is also primitive which is a contradiction. Hence, $vu \in \overline{Q}_p$. □

We denote the language of power of primitive partial words $Q_p^{(i)}$, and defined as

$$Q_p^{(i)} = \{u^i \mid u \in Q_p\}$$

In our next result, we show that the language $Q_p^{(i)}$ is closed under cyclic permutation.

Proposition 3.4. *Let u, v be two nonempty partial words. If $uv \in Q_p^{(i)}$ for $i \geq 1$ then $vu \in Q_p^{(i)}$.*

Proof. For $i = 1$, let $uv \in Q_p$. By using Corollary 3.2, we conclude that $vu \in Q_p$. For $i \geq 2$, let $uv \in Q_p^{(i)}$, that is, there exists a primitive word x such that $uv \subset x^i$. Let $x = yz$ for some $y \in V^+$ and $z \in V^*$. Hence $u \subset (yz)^m y$ and $v \subset z(yz)^n$ where $m + n + 1 = i$. Since $x = yz \in Q$ and Q is closed under the cyclic permutation, thus $zy \in Q$. Hence $vu \subset z(yz)^n(yz)^m y = (zy)^n zy(zy)^m = (zy)^{m+n+1} = (zy)^i$ which proves that $vu \in Q_p^{(i)}$. □

The following corollary immediately follows from Proposition 3.4.

Corollary 3.5. *The language $Q_p^{(i)}$ is reflective.*

Let $u = xyz$ be a nonempty partial word over an alphabet V where $x, z \in V_\Diamond^*$ and $y \in V_\Diamond^+$. The partial word y is called as factor of u. If $x \neq \varepsilon$ (respectively, $z \neq \varepsilon$), then x (*respectively*, z) is called the nonempty prefix (*respectively*, nonempty suffix) of u. Recall that, a partial word w is bordered if there exist words x, y and z such that $w \subset xy$ and $w \subset yz$.

Lemma 3.6. *Let u be a partial word such that $u \subset v^k$, $k \geq 2$ where $v \in Q$. Let w be a factor of u with $|w| > |v|$. Then w is bordered.*

Proof. Let u be a partial word and w is a factor of u. Hence $u = xwz$ for some x, $z \in V_\Diamond^*$. Consider the partial word wzx which is a conjugate of u. We know that the language of primitive partial words is reflective. Using Proposition 2.41, since u is contained in k^{th} power of v, the partial word wzx will also be contained in k^{th} power of some conjugate t of v. Hence we can write $w \subset t^j t'$ for $j \geq 1$ and t' is a nonempty prefix of t. Then w is contained in $t't''t'$ where $t't'' = t^j$ and hence w is bordered. □

By Proposition 2.42, if u is a primitive partial word with at most one hole and $uu \uparrow xuy$ then either $x = \varepsilon$ or $y = \varepsilon$. In our next result we prove that for $n \geq 2$, u^{n-1} is not compatible to a central factor of u^n where u is a primitive partial word with one hole.

Lemma 3.7. *Let u be a nonempty partial word with one hole. Then u is primitive if and only if u^{n-1} is not compatible to a central factor of u^n for all $n \geq 2$.*

Proof. $(\Rightarrow)$

We prove it by induction on n.

Base Case: Let $n = 2$. Assume that u is compatible to a central factor of u^2. Then there exist nonempty partial words t and s such that $u^2 \uparrow tus$ where $|u| = |t| + |s|$, and t and s are prefix and suffix of u^2, respectively. Thus there exist two partial words x and y such that $u = tx$, $u = ys$. Hence $uu = txys \uparrow tus$ implies that $u \uparrow xy$.

Now, $|u| = |x| + |y| = |t| + |x| = |y| + |s|$. Thus $|t| = |y|$ and $|x| = |s|$. Also $tx = ys$ and $|t| = |y|$ implies $t = y$ and $x = s$ which further implies $u = ys = yx$. Hence, $xy \uparrow yx$. Thus, by using Theorem 2.36 there exists a nonempty word v such that $x \subset v^m$ and $y \subset v^n$ for some $m, n \geq 1$. Hence, $u \subset v^{m+n}$, a contradiction.

Induction Hypothesis: Let us assume that the result is true for $2 \leq k \leq n-1$, that is, u^{k-1} is not compatible to a central factor u^k.

Induction Step: Now we prove the result for $n = k$, that is, we will prove that u^k is not compatible to a central factor of u^{k+1}. Assume that u^k is compatible to a central factor of u^{k+1}. Thus, $u^{k+1} = tvs$ such that $v \uparrow u^k$ for some nonempty partial words t, s with $|u| = |t| + |s|$ and t and s are prefix and suffix of u, respectively. Since $|t|, |s| < |u|$ then $u = ty = zs$ for some nonempty partial words $y, z \in V_\diamond^+$. Now, $u^{k+1} = uu^{k-1}u = tyu^{k-1}zs$. As $v \uparrow u^k$, we have $yu^{k-1}z \uparrow u^k$. Since y and z are nonempty u^{k-1} is compatible to a central factor of u^k which is a contradiction to the inductive hypothesis, which implies that u^k is not compatible to a central factor of u^{k+1}.

Hence if u is a primitive partial word then u^{n-1} cannot be compatible to a central factor of u^n for all $n \geq 2$.

($\Leftarrow$)

Assume that u is not a primitive partial word. Then there exists a nonempty word z such that $u \subset z^k$, $k \geq 2$. Now, $u^n \subset z^{nk} = zz^{k(n-1)}z^{k-1}$. Since $u^{n-1} \subset z^{k(n-1)}$ we have $u^n \uparrow zu^{n-1}z^{k-1}$, and thus u^{n-1} is compatible to a central factor of u^n, a contradiction.

□

However, Lemma 3.7 does not hold for partial words having two or more holes. For example, consider a partial word $w = a\diamond b\diamond \in Q_p$. We have $ww = a\diamond b\diamond a\diamond b\diamond \uparrow xwy$ where $x = a$ and for some $y \uparrow \diamond b\diamond$. Hence w is compatible to a central factor of ww.

3.4 Partial Primitivity and Density

We now explore the denseness property of the language of primitive partial words. We prove that the language of primitive partial words with at most one hole Q_p^1 is dense over an alphabet V in V_1^*.

A language $L \subseteq V^*$ is called a dense language if for every word $w \in V^*$, there exist words $x, y \in V^*$ such that $xwy \in L$ [55]. Let k be a positive integer. A language $L \subseteq V^*$ is called right k-dense (strictly right k-dense) if for every $u \in V^*$, there exists a word $x \in V^*$, $|x| \leq k$ $(1 \leq |x| \leq k)$, such that $ux \in L$, that is, $uX_k \cap L \neq \emptyset$ where $X_k = \{x \in V^* \mid |x| \leq k\}$. A language L is called right dense if it is right k-dense for all $k \geq 1$. It is known that the language of total primitive words Q is dense and given in the following result.

Lemma 3.8 ([78]). *The language of primitive words Q is right dense.*

A language L is said to be *minimal (strictly) right k-dense* if L is (strictly) right k-dense language and does not contain any proper (strictly) right k-dense language.

The definition of dense language, right k-dense, strictly right k-dense can be extended for partial words in similar manner.

Definition 3.9. *A language of partial words $L \subseteq V_\diamond^*$ is said to be dense over an alphabet V if for every partial word $w \in V_p^*$ there exist partial words x and y such that $xwy \in L$.*

We define a strict language of partial words with at most one hole. Let $V_1^*(s) = V_1^* \setminus \{x\diamond x \mid x \in V^*\}$.

Corollary 3.10. *Q_p^1 is right 1-dense in $V_1^*(s)$.*

Proof. The proof follows from Lemma 2.44. □

Theorem 3.11. *The language Q_p^1 is dense over an alphabet V in V_1^*.*

Proof. Consider a partial word w with at most one hole . Let $|w| = n,\ n \geq 1$. There are two different cases depending upon whether w is primitive partial word or nonprimitive partial word.

Case 1. Let w is a primitive partial word. By choosing $x = \varepsilon$ and $y = \varepsilon$, we have $xwy = w \in Q_p^1$.

Case 2. Let w be a nonprimitive partial word. Here we consider two possibilities depending on whether w is contained in power of a symbol from the alphabet or power of a word having two or more distinct letters.

(i) Let $w \subset a^n,\ n \geq 2$ for some symbol $a \in V$. It can be easily seen that $wb^n \in Q_p^1$ where a and b are two distinct letters. Here $x = \varepsilon$ and $y = b^n$ for some $b \neq a$ such that $xwy \in Q_p^1$.

(ii) Let $w \subset x^k$, where $|\alpha(x)| \geq 2$ and $k|n$. By choosing $x = \varepsilon$ and $y = b^n$, we have $xwy \in Q_p^1$.

Hence, for every $w \in V_1^*$, there exist $x,\ y \in V_1^*$ such that $xwy \in Q_p^1$. So Q_p^1 is dense over the alphabet V in V_1^*. □

Lemma 3.12. *Let L be a minimal right 1-dense language of primitive partial words with at most one hole over an alphabet V. Then the following statements hold:*

(a) $L \cap V_\Diamond \neq \emptyset$.

(b) *For every integer $n \geq 1$ and every $w \in V_1^*(s)$ such that $w \subset a^n,\ a \in V$ there exists $b \in V,\ b \neq a$ such that $wb \in L$.*

(c) *If L is minimal strictly right 1-dense and if $ua,\ ub \in L$ with*

$$u \in V^*\{\Diamond\}V^* \setminus \{x\Diamond x \mid x \in V^*\}$$

and $a,\ b \in V$ then $a = b$.

Proof. (a) Since L is minimal right 1-dense, then it is also right 1-dense. From the definition, there exists a partial word $x \in V_1^*,\ |x| \leq 1$ such that $\varepsilon x \in L$. Since ε is not primitive, so $|x| = 1$ and $x \in L \cap V_\Diamond$. Hence $L \cap V_\Diamond \neq \emptyset$.

(b) Since $w \subset a^n$ for $a \in V$, so w is not a primitive partial word. Since L is right 1-dense, there exists a partial word x with $|x| \leq 1$ such that $wx \in L$. This implies $x = b$ for some $b \in V$ and $b \neq a$.

(c) For the sake of contradiction, suppose that $a \neq b$. Consider the set $L - \{ub\}$. The set $L - \{ub\}$ is not strictly right 1-dense. So there exists $v \in V^*\{\Diamond\}V^* \setminus \{x\Diamond x \mid x \in V^*\}$ such that $\{v\}V \cap (L - \{ub\}) = \emptyset$. As L is minimal strictly right 1-dense, then $\{v\}V \cap L \neq \emptyset$. So it follows that, there exists some $y \in V$ such that $vy \in L$. From $vy \notin (L - \{ub\})$, it follows that $vy = ub$. Since $y \in V$, then $y = b$ and $v = u$. Now $\{u\}V \cap (L - \{ub\}) = \emptyset$. Since $ua \in L - \{ub\}$, we obtain a contradiction and hence $a = b$.

□

3.5 The Language Q_p in Chomsky Hierarchy

In this section, we shall prove that the language of primitive partial words Q_p is not regular, not linear as well as not a deterministic context-free Language (DCFL). Also we shall prove that the language of nonprimitive partial words $\overline{Q}_p$ over a nontrivial alphabet is not a context-free language.

The linear Diophantine equation is given in the next Lemma which is required to prove Theorem 3.14.

Lemma 3.13 ([33]). *For any fixed integer k, there exists a positive integer m such that the equation system $(k-j)x_j + j = m$, $j = 0, 1, \ldots, k-1$ has a nontrivial solution with appropriate positive integers $x_1, \ldots, x_j > 1$.*

In order to prove that the language of primitive partial words Q_p is not regular, we will use the pumping lemma for regular languages given in Appendix A.1.

Theorem 3.14. *Q_p is not a regular language.*

Proof. Suppose that the language of primitive partial words Q_p is a regular language. So there exists a finite state automaton and $n > 0$ is the pumping length depending upon the number of states in that automaton for Q_p.

Consider the partial word $w = a^n\Diamond a^m\Diamond a^m b, m > n$. Note that w is a primitive partial word over V where $|V| \geq 2$ and $a \neq b$. Since $w \in Q_p$ and $|w| \geq n$, then it must satisfy the other conditions of pumping lemma for regular languages. So there exists a decomposition of w into x, y, z such that $w = xyz, |y| > 0$, $|xy| \leq n$ and $xy^iz \in Q_p$ for all $i \geq 0$.

Let $x = a^k$, $y = a^{(n-j)}$, $z = a^{j-k}\Diamond a^m\Diamond a^m b$. Now choose $i = x_j$ and we know by Lemma 3.13 that for every $j \in \{0, 1, \ldots, n-1\}$, there exists a positive integer $x_j > 1$ such that $xy^{x_j}z = a^k a^{(n-j)x_j} a^{j-k}\Diamond a^m\Diamond a^m b = a^{(n-j)x_j+j}\Diamond a^m\Diamond a^m b = a^m\Diamond a^m\Diamond a^m b \subset (a^m b)^3 \notin Q_p$ which is a contradiction. Hence the language of primitive partial words Q_p is not regular. □

Next we prove that the language of primitive partial words Q_p is not linear. To prove this, we use the standard pumping lemma for linear languages given in Appendix A.2.

Theorem 3.15. *Q_p is not a linear language.*

Proof. Suppose that the language of primitive partial words Q_p is linear. Let $m \geq 1$ be an integer. Consider $z = a^m ba^{2m+m!} b\Diamond^m \in Q_p$ be a partial word over the alphabet $V = \{a, b\}$ and $a \neq b$. Since $|z| \geq m$, so there exists a factorization of z into u, v, w, x, y such that it satisfies the conditions of pumping lemma for linear languages.

Now, in the factorization of $z = uvwxy$, we have $|uvxy| \leq m$ which necessarily implies that $|uv|, |xy| \leq m$. So, $uv \in a^*$ and $xy \in \Diamond^*$. For $i = 1 + \frac{m!}{|vx|}$, we should have $z^i = uv^i wx^i y = a^{m+n'} ba^{2m+m!} b\Diamond^{m+n''} \in Q_p$ for some non-negative n' and n'' such that $n' + n'' = m!$. But $z^i = a^{m+n'} ba^{2m+m!} b\Diamond^{m+n''} = a^{m+n'} ba^{m+n''} a^{m+n'} b\Diamond^{m+n''} \subset (a^{m+n'} ba^{m+n''})^2 \notin Q_p$, which is a contradiction. Hence Q_p is not linear. □

In our next result, we show that the language of primitive partial words is not a deterministic context-free language and we use the closure properties of DCFL. In particular, the set of DCFLs are closed under complementation and the intersection of a deterministic context-free language and a regular language is a DCFL.

Theorem 3.16. *Q_p is not a deterministic context-free language.*

Proof. Suppose that the language of primitive partial words Q_p is deterministic context-free. As the set of DCFLs is closed under complementation, the complement of Q_p, that is, $\overline{Q}_p$ is also a DCFL.

Also, we know that the intersection of a DCFL with a regular language is also a DCFL. Therefore, $\overline{Q}_p \cap \{a^* b^* a^* b^*\} = \{a^n b^m a^n b^m \mid m, n \in N\}$ is also a DCFL. But the language $\{a^n b^m a^n b^m \mid m, n \in N\}$ is not a context-free language (CFL) which can be proved by using pumping lemma for CFLs. Hence it is a contradiction that the language of primitive partial words Q_p is a DCFL. Therefore the language Q_p is not deterministic context-free. □

In the following theorem, we prove that the language of nonprimitive partial words $\overline{Q}_p$ is not context-free. We prove it by using the pumping lemma for context-free languages which is given in Appendix A.3.

Theorem 3.17. *$\overline{Q}_p$ is not a context-free language.*

Proof. We prove it by contradiction. Let us assume that $\overline{Q}_p$ is a CFL. Let $n > 0$ be an integer that is guaranteed to exist by pumping lemma. Since $\overline{Q}_p$ is context-free, then it satisfies all the conditions of Lemma A.3. So for every partial word $z \in \overline{Q}_p$ with $|z| \geq n$, z can be factorized into $uvwxy$ such that $|vwx| \leq n$, $|vx| > 0$ and $uv^t wx^t y \in \overline{Q}_p$ for all $t \geq 0$.

Consider the partial word $z = a^{n+1} \Diamond^{n+1} a^{n+1} b^{n+1}$ where $a,\ b \in V$ and are distinct. It is easy to follow that $z \in \overline{Q}_p$ as $z \subset (a^{n+1} b^{n+1})^2$. There will be different possibilities depending on whether the substrings v and x contain one type of symbol or more than one alphabet symbol or hole.

Case 1. If v and x contain only one type of symbol. Either we shall have $vwx = a^p$ or $vwx = b^p$ or $vwx = \Diamond^p$ for $p \leq n$. Suppose $vwx = a^p$ from the first part of a's. Let $u = a^i$, $v = a^j$, $w = a^k$, $x = a^l$ and $y = a^m\Diamond^{n+1}a^{n+1}b^{n+1}$ such that $j + k + l = p$ and $j + l \geq 1$. Consider uv^twx^ty for $t = 2$.

$$\begin{aligned} uv^2wx^2y &= a^{i+2j+k+2l+m}\Diamond^{n+1}a^{n+1}b^{n+1} \\ &= a^{n+1+j+l}\Diamond^{n+1}a^{n+1}b^{n+1} \notin \overline{Q}_p \text{ as } j + l \geq 1 \end{aligned}$$

Other cases where vwx consists of only a's from the second part or only b's can be handled similar way. We consider vwx consist of only $\Diamond$'s separately.

Case 2. If v and x contain different types of symbols. We will have the following choices: (a) $vwx = a^p\Diamond^q$ (b) $vwx = \Diamond^pa^q$ (c) $vwx = a^pb^q$. We consider the first case and the other cases can be proved similarly. Let $vwx = a^p\Diamond^q$. Here we have three possibilities.

(i) $v = a^{p_1}$, $w = a^{p_2}$, $x = a^{p_3}\Diamond^q$

(ii) $v = a^{p_1}$, $w = a^{p_2}\Diamond^{q_1}$, $x = \Diamond^{q_2}$

(iii) $v = a^{p_1}\Diamond^{q_1}$, $w = \Diamond^{q_2}$, $x = \Diamond^{q_3}$

Let us consider subcase (i). Let $v = a^{p_1}$, $w = a^{p_2}$, $x = a^{p_3}\Diamond^q$ where $p_1+p_2+p_3 = p$. For $t = 2$, we have

$$uv^2wx^2y = a^{n+1+p_1}\Diamond^qa^{p_3}\Diamond^{n+1}a^{n+1}b^{n+1} \notin \overline{Q}_p.$$

Similarly the other two subcases can be proved.

Case 3. Let $vwx = \Diamond^p$. Suppose $u = a^{n+1}\Diamond^i$, $v = \Diamond^{p_1}$, $w = \Diamond^{p_2}$, $x = \Diamond^{p_3}$ and $y = \Diamond^ja^{n+1}b^{n+1}$ where $p_1 + p_2 + p_3 = p$, $p_1 + p_3 \geq 1$ and $i + j + p = n + 1$. Now $uv^twx^ty = a^{n+1}\Diamond^{i+p_2+j}a^{n+1}b^{n+1} \notin \overline{Q}_p$ for $t = 0$.

The other cases are similar to one of the above three cases. Hence the language of nonprimitive partial words $\overline{Q}_p$ over a nontrivial alphabet is not a context-free language. □

3.6 2DPDA for Q_p^1

In this section, we prove that the language of primitive partial words with one hole Q_p^1 is accepted by a two-way pushdown automaton (2DPDA). Let us recall the definition of 2DPDA.

A 2DPDA is the same as ordinary DPDA but with an additional ability to move its input head in both the directions. We use Z_0 as the bottom of stack and one left end and one right end symbol as $\vdash$ and $\dashv$, respectively.

Definition 3.18 ([88]). *A 2DPDA, P, consist of 7-tuples*

$$P = \langle S, I, T, \delta, s_0, Z_0, s_t \rangle,$$

where

(a) S is the states of the finite control.

(b) I is the input alphabet (excluding $\vdash$ and $\dashv$).

(c) T is the pushdown list alphabet (excluding Z_0).

(d) δ is a mapping on $(S - \{s_t\}) \times (I \cup \{\vdash, \dashv\}) \times (T \cup Z_0)$. The value of $\delta(s, a, A)$ is one of the forms $(s^{'}, d, push\ B)$, (s', d), or (s', d, pop) where $s^{'} \in S, B \in T$ and $d \in \{-1, 0, +1\}$. We assume a 2DPDA makes no moves from the final state s_t.

(e) $s_0 \in S$ is the initial state of the finite control.

(f) Z_0 is the special symbol that indicates the bottom of the pushdown list.

(g) s_t is one of the designated final state.

In the above definition $\delta(s, a, A)$ indicates the transition function of a machine when it is in the state s, has the symbol A on top of the stack and the input head has read a symbol a. The three possible actions for $\delta(s, a, A)$ are: $(s^{'}, d, push\ B)$ provided $B \neq Z_0$, $(s^{'}, d)$ and $(s^{'}, d,\ pop)$ if $A \neq Z_0$. In these transitions the machine enters in state s' and moves its input head in the direction d, where $d = -1$, $+1$ or 0, to indicate to move its head to left, right or remain stationary, respectively. The operation *push* B means add the symbol B on the top of pushdown list and *pop* means remove the topmost symbol from the pushdown list.

The following lemma shows that for a primitive partial word of length n, there are n primitive partial words in its conjugate class.

Lemma 3.19. *Let w be a primitive partial word with $|w| = n$. Then the cardinality of the class of conjugates is n.*

Proof. Let w be a primitive partial word over an alphabet V of length n. Thus every new partial word which is generated by cyclic permutation of w is a conjugate of w. Since w is primitive and we know that a cyclic permutation of a primitive partial word is also primitive, each of the conjugate of w is also primitive. Therefore, the number of such conjugates of w is n. This proves our claim. □

In [80], a 2DPDA that recognizes the language of primitive words Q is presented. We show that a 2DPDA can also be constructed for Q_p^1 by using the similar idea.

Theorem 3.20. *Q_p^1 is accepted by a 2DPDA.*

Proof. The informal idea is as follows: Let P be a 2DPDA. Let $\vdash w \dashv$ be the input partial word with at most one hole augmented with two end markers. If $w = \varepsilon$, P rejects. If not, then P moves its head towards $\dashv$. P skips the last symbol of w and pushes the remaining symbols of w onto its stack. Again P moves to $\dashv$, pushes all symbols of w onto the stack and pops one symbol. If we write the input $w = xw'y$ with $x, y \in V \cup \{\Diamond\}$ (assuming $|w| \geq 2$), the contents of the stack will be:

$$w'yxw'Z_0$$

The automaton P compares w with the pushdown contents one symbol at a time. If the symbols match (assuming that $a \neq b$ and $\Diamond = a$ for any $a \in V$), the head moves right and P pops the pushdown. If in this way the entire word w is completely scanned and is compatible to a factor of $w'yxw'$ P rejects (by assuming that the hole $\Diamond$ is compatible to any of the symbol in V).

If a mismatch is encountered P moves the input head back to $\vdash$ and pushes the symbols scanned during this move. Then P pops the topmost symbol from the pushdown and repeats the process. If the pushdown becomes empty then P accepts.

Hence the 2DPDA accepts w if and only if w is not compatible to a factor of $w'yxw'$. By Proposition 2.42 such a partial word is primitive. □

Observe that the idea used in above result does not work for the set of primitive partial words with two or more holes. We illustrate in the following example.

Example 3.21. Let $w = a\Diamond b\Diamond$ be a primitive partial word with two holes. By the method described in Theorem 3.20 we have $x = a, y = \Diamond$ and $w' = \Diamond b$. Now $w'yxw' = \Diamond b\Diamond a\Diamond b$ and w is compatible to a factor of $w'yxw'$. □

Next we show that the language of primitive partial words Q_p over an alphabet V is accepted by a Random Access Machine (RAM) model of computation in linear time. A RAM is a simple model of computation equipped with an input tape that is accessed sequentially during the computation. It requires the input and its length to be available before the RAM computation starts. For more details, see [48].

Next we prove the following result which shows the optimality to recognize a primitive partial word.

Lemma 3.22. *The language Q_p of primitive partial words over a nontrivial alphabet V is accepted by a RAM in linear time. The bound is optimal.*

Proof. In [9], a linear time algorithm has been given to recognize the set of primitive partial words for RAM model of computation. We use the standard adversary argument to prove the optimality. Consider an input w of length n and w is stored in the input array as $t[0], \ldots, t[n-1]$.

Let w be a partial word such that $w \subset a^n$ for $a \in V$. Suppose, RAM answers on w, without looking at each letter of the input. Let $i \in \{0, \ldots, n-1\}$ be a position such that RAM has not looked at position i of the input. RAM would therefore answer the same on partial word $w' = a^i b\Diamond a^{n-i-2}$ and, partial word w. But the former partial word is nonprimitive and the latter is a primitive partial word. Hence at least $|w|$, that is, n operations are required in the worst case. □

3.7 Indexed Grammar for Z

In this section, we study a new type of grammar for the language of nonprimitive words known as indexed grammar that was introduced by Alfred V Aho [1]. An indexed grammar is an extension of context-free grammar and is a proper subset of context sensitive grammar. The corresponding language recognized by an indexed grammar is known as indexed language. We present the definition of indexed grammar along with some examples.

Definition 3.23. *An indexed grammar is a 5-tuple $G = (N,\ T,\ F,\ P,\ S)$ where*

(a) N is the finite set of non-terminals.

(b) T is the finite set of terminal symbols.

(c) F is the set of flags or indices of which the elements are ordered pairs.

$$F = \{(A,\ \chi) \mid A \in N,\ \chi \in (N \cup T)^*\}$$

The elements of F are called index productions.

(d) P is the finite set of production rules in the form of ordered pairs

$$P = \{(B,\ \alpha) \mid B \in N,\ \alpha \in (NF^* \cup T)^*\}$$

(e) S is a special non-terminal which is the start symbol.

When there are no flags, an indexed grammar is equivalent to a context-free grammar. Unlike we expand only a nonterminal in context-free grammars, we expand nonterminals or flags in indexed grammar. The flags or indices $f \in F^*$ are successively applied on nonterminals in ordered pairs $(A,\ f)$ and produce the string of terminals, and does not have any effect on terminals. In the production rules, the flags $f \in F^*$ will be distributed over the nonterminals to generate the terminal strings. The language of an indexed grammar G is called as an indexed language and $L(G) = \{w \in T^* \mid S \xrightarrow[G]{*} w\}$.

While deriving the string $w \in L(G)$, the intermediate strings $\alpha \in (NF^* \cup T)^*$ are called as sentential forms. A sentential form α directly generates a sentential form β if either of the following holds:

1. Let the production $B \to X_1\eta_1X_2\eta_2 \dots X_k\eta_k$ and applying this production on $\alpha = \gamma B\sigma\varphi$, we have, $\gamma X_1\eta_1X_2\eta_2 \dots X_k\eta_k\sigma\varphi$ which gives $\beta = \gamma X_1\theta_1X_2\theta_2 \dots X_k\theta_k\varphi$ for $1 \le i \le k$, $\theta_i = \eta_i\sigma$ if $X_i \in N$ or $\theta_i = \varepsilon$ if $X_i \in T$.

or

2. Let $\alpha = \gamma Bf\sigma\varphi$ and $B \to X_1X_2 \dots X_k$ is an index production in the index f. Then we have $\beta = \gamma X_1\theta_1X_2\theta_2 \dots X_k\theta_k\varphi$ where for $1 \le i \le k$, $\theta_i = \sigma$ if $X_i \in N$ or $\theta_i = \varepsilon$ if $X_i \in T$.

In the first case, $\sigma \in F^*$ has been distributed over the production or index production nonterminals. In the second case, f has been consumed when the index production of f has been applied.

Next, we illustrate the concept of indexed grammar by the following examples.

Example 3.24 ([1])**.** Consider the language $L = \{a^n b^n c^n \mid n \geq 1\}$. It is known that the language L is context sensitive but not context-free. The language L can be generated by the following indexed grammar G:

$G = (\{S,\ A,\ B\}, \{a,\ b,\ c\}, \{f,\ g\}, P, S)$ where
$P = [S \rightarrow aAfc,\ A \rightarrow aAgc,\ A \rightarrow B\ ]$
$f = [B \rightarrow b], \qquad g = [B \rightarrow bB]$

The above grammar generates the strings of the form $a^n b^n c^n$ for $n \geq 1$. □

Let us consider another example.

Example 3.25. Consider the language $L_1 = \{a^n b^{n^2} a^n | n \geq 1\}$. The language L_1 can be generated by the following indexed grammar.

$G = (\{S,\ T,\ A,\ B,\ C\}, \{a, b\}, \{f, g\}, P, S)$

where

$$P = [\ S \rightarrow Tf,\ T \rightarrow Tg,\ T \rightarrow ABA\ ]$$

is the set of production rules, and

$$f = [A \rightarrow a, B \rightarrow b, C \rightarrow b], g = [A \rightarrow aA, B \rightarrow bBCC, C \rightarrow bC]$$

are the index production rules. □

Next we give an indexed grammar for the language of nonprimitive words. It is known that the language of nonprimitive words is not a context-free language.

Theorem 3.26. *The language of nonprimitive words Z is an indexed language.*

Proof. Consider the indexed grammar $G =< N,\ T,\ F,\ P,\ S >$ for the language of nonprimitive words Z where $N = \{S,\ A,\ B,\ C\}$, $T = \{a,\ b\}$, $F = \{f,\ g,\ h\}$ and

$P = \{S \rightarrow Af, A \rightarrow Ag \mid Ah \mid B, B \rightarrow CC, C \rightarrow CC\}$ is the set of production rules, and

$g = [C \rightarrow aC], h = [C \rightarrow bC], f = [C \rightarrow \varepsilon]$ are the index productions.

Next we prove that the above indexed grammar generates the language of nonprimitive words. We prove in both directions, that is, $L(G) \subseteq Z$ and $Z \subseteq L(G)$.

($\Rightarrow$) Let w be a word that is generated by the indexed grammar G. To generate w, the initial production rule $S \rightarrow Af$ is applied.

If $w = \varepsilon$, the production rules $A \rightarrow B$ and $B \rightarrow CC$ are subsequently applied. Then the distribution of index f over nonterminal C generates ε which is a nonprimitive word.

If $w \neq \varepsilon$, either one of the production rules $A \to Ag$ or $A \to Ah$ is applied. The only way to generate terminal symbols is to use the production rule $A \to B$. The repeated application of production rules $B \to CC$ and $C \to CC$ generates the number of copies of a word to be formed. Then the index production rules $g = [C \to aC]$ or $h = [C \to bC]$ is applied to generate terminals a or b respectively. Finally, f is applied on C to generate the word w and, $w \in Z$. Thus, $L(G) \subseteq Z$.

($\Leftarrow$) Let $w = x^n$ be a nonprimitive word with $n \geq 2$.

Depending upon the primitive word x, we apply the production rules and index production rules to generate the word x. To start with, the initial production rule $S \to Af$ is applied. Then the production rules $A \to Ag$ or $A \to Ah$ is applied repeatedly to generate the corresponding indices of word x. To generate n number of copies of x, the production rules $B \to CC$ and $C \to CC$ are applied. Then the distribution of indices over the nonterminals generate the word $w = x^n$. Thus, $Z \subseteq L(G)$.

Hence $Z = L(G)$ and Z is an indexed language. □

Let us illustrate the above grammar by deriving a string in the following example.

Example 3.27. Let $w = (abb)^2$ be a nonprimitive word. The following is the derivation steps for w.

$$
\begin{aligned}
S &\to Af \\
&\to Ahf \\
&\to Ahhf \\
&\to Aghhf \\
&\to Bghhf \\
&\to CghhfCghhf \\
&\to aChhfCghhf \\
&\to abChfCghhf \\
&\to abbCfCghhf \\
&\to abbCghhf \\
&\xrightarrow{*} abbabb = (abb)^2
\end{aligned}
$$

□

3.8 Conclusions

We discussed the language of primitive partial words and provided a hierarchy of languages of partial words. We proved that the language of primitive partial words is not regular, not linear and is not a deterministic context-free language. We also proved that the language of nonprimitive partial words is not a context-free language over a nontrivial alphabet. A 2DPDA automaton has been given to recognize the set of primitive partial words with one hole. An indexed grammar to recognize the language of nonprimitive words has been given.

Chapter 4

Robustness of Primitive Partial Words

In this chapter, we study several operations on primitive partial words that preserves primitivity. In particular, the operations that we consider are deletion of a symbol, insertion of a symbol, exchange of two distinct consecutive symbols and substitution of a symbol by another symbol from the underlying alphabet.

4.1 Introduction

When we study an algebraic structure, we are interested in investigating the operations which preserve some property of the underlying structure. Primitivity of words is an important property in the field of word combinatorics. Thus, it is worth considering the operations under which the property primitivity can be preserved. In [78], Păun et al. considered the effect of point mutation operations on primitive words such as deleting a symbol, inserting a symbol, substituting a symbol by another symbol and the operation of taking prefixes. The preservation of primitivity property has also been studied with respect to homomorphisms; [75, 76, 79].

In [31], Jürgen Dassow et al. have investigated the *edit distance* operation under which a word ww' is primitive where w is a given primitive word and w' is a modified mirror image or modified copy of w. For two words w and w', the *edit distance* of w and w' is the minimal number of changes, deletions and insertions of letters to transform w into w'. Blanchet-Sadri, Nelson and Tebbe [21] extended the work of Dassow et al. [31] to partial words with one hole and studied the notion of preservation of primitivity in partial words.

We state some of the existing results on the robustness of primitive words and some properties on primitive partial words that will be useful later in this chapter.

The following result shows that for any nonempty word u, appending two distinct letters a and b from the alphabet, at least one of ua or ub is primitive.

Lemma 4.1 ([78]). *For every word $u \in V^+$ and all symbols $a, b \in V$, where $a \neq b$, at least one of the words ua, ub is primitive.*

Lemma 2.44 shows the possibility of obtaining a primitive partial word by appending a symbol or removing a symbol in any nonempty partial word. Specifically, if u is a nonempty partial word with one hole, then at least one of the u or ua is primitive for $a \in V$.

The next result holds for partial words with one hole of some particular type.

Proposition 4.2 ([8]). *(a) Let u be a partial word with one hole which is not of the form $x\Diamond x$ for any word x. If a and b are distinct letters, then ua or ub is primitive.*

(b) Let u_1 and u_2 be partial words such that u_2u_1 has one hole and is not of the form $x\Diamond x$ for any word x. Then at most one of the words u_1au_2 with $a \in V$ is not primitive.

The following proposition shows the possibility of obtaining primitive word by deletion of a symbol in a primitive word.

Proposition 4.3 ([78]). *Every word $w \in Q$, $|w| \geq 2$, can be written in the form $w = u_1au_2$, for some u_1, $u_2 \in V^*$, $a \in V$, such that $u_1u_2 \in Q$.*

Organization. The rest of this chapter is organized as follows. In Section 4.2, we discuss the preservation of primitivity of partial words with one hole with respect to deletion operation. We call the set of primitive partial words with one hole which are robust to deletion operation as *del-robust primitive partial words*. We characterize the set of primitive partial words with one hole which are not del-robust and prove that the language of non-del-robust primitive partial words with one hole is not context-free over an alphabet. In Section 4.3, we explore the primitive partial words that are robust to exchange operation where we exchange two consecutive distinct symbols from the underlying alphabet. We prove that the language of non-exchange-robust primitive partial words is not context-free over an alphabet in Section 4.4. We investigate the preservation of primitivity of partial words with respect to substitution operation along with some combinatorial properties in Section 4.5. We finally conclude this chapter in Section 4.6.

4.2 Del-Robust Primitive Partial Words

In this section, we discuss the robustness of primitive partial words with one hole with respect to deletion of a symbol or a hole. This special class of primitive partial words with one hole is referred as del-robust primitive partial words with one hole.

We have the following result for primitive partial words with one hole that gives an alternative to obtain new primitive partial words by deletion of either a symbol or a hole.

Lemma 4.4. *Every primitive partial word with one hole w with $|w| \geq 4$, $w = v_1abv_2$ such that $v_1v_2 \neq x\Diamond x, x \in V^*$ can be written as $w = u_1au_2$ such that $u_1u_2 \in Q_p$.*

Proof. Given that $|w| \geq 4$ and $w = v_1abv_2$ for some $v_1, v_2 \in V_\Diamond^*$, and v_1v_2 has exactly one hole. Also v_1v_2 is not of the form $x\Diamond x$ for some $x \in V^*$. From Proposition 4.2(b), at least one of the v_1av_2 or v_1bv_2 is primitive. If v_1av_2 is primitive then $u_1 = v_1a$ and $u_2 = v_2$. If v_1bv_2 is primitive then $u_1 = v_1$ and $u_2 = bv_2$. □

The class of primitive partial words with one hole which remains primitive upon deletion of a symbol or a hole is a subclass of primitive partial words with one hole. A prefix (suffix) of length k of a partial word w is denoted as $\text{pref}(w, k)$ $(\text{suff}(w, k))$, respectively, where $k \in \{0, 1, \ldots, |w|\}$ and $\text{pref}(w, 0) = \text{suff}(w, 0) = \varepsilon$. We define the del-robust primitive partial words with one hole as follows.

Definition 4.5 (Del-Robust Primitive Partial Words)**.** *A primitive partial word w of length n with one hole is said to be del-robust if and only if the partial word*

$$\text{pref}(w, i)\ .\ \text{suff}(w, n - i - 1)$$

is a primitive partial word for all $i \in \{0, 1, \ldots, n - 1\}$.

Note that a del-robust primitive partial word must remain primitive on deletion of any symbol or the hole. The number of such partial words is infinite. For example, $a\Diamond bba$ and $a^m\Diamond b^n$ for $m,\ n \geq 2$ are del-robust primitive partial words with one hole.

We denote the set of all del-robust primitive partial words with one hole by Q_p^{1D} over an alphabet. It is obvious that the language of del-robust primitive partial words with one hole is a subset of Q_p^1 and hence $w \in Q_p^1$ for all $w \in Q_p^{1D}$ where Q_p^1 be the set of all primitive partial words with one hole. Let $Q_p^1(1) = Q_p^1 \cup \{\varepsilon\}$, and for all $n \geq 2$ let $Q_p^1(n) = \{u^n \mid u \in Q_p^1\}$.

Proposition 4.6. *Let m and n be two distinct positive integers. Then the sets $Q_p^1(m)$ and $Q_p^1(n)$ are disjoint.*

Next we study a subset of primitive partial words with one hole in which deletion of a symbol or a hole results in a nonprimitive partial word. We call such partial words as non-del-robust primitive partial words with one hole and the set is denoted by $Q_p^{1\overline{D}}$.

Definition 4.7 (Non-Del-Robust Primitive Partial Words)**.** *A primitive partial word w with one hole is said to be non-del-robust if and only if*

$$\text{pref}(w, i)\ .\ \text{suff}(w, n - i - 1)$$

is not a primitive partial word for some $i \in \{0, 1, \ldots, n - 1\}$.

Thus $Q_p^{1\overline{D}} = Q_p^1 \setminus Q_p^{1D}$ where '\' is the set difference operator. The number of such non-del-robust words is infinite. For example, $a\Diamond b$, $ba^n\Diamond$ for $n \geq 1$ are non-del-robust words.

Next we give complete structural characterization of those primitive partial words with one hole which are in the set Q_p^1 but not in Q_p^{1D}, that is, deletion of a symbol or hole from such words will result in nonprimitive partial words.

Theorem 4.8. *A primitive partial word w with one hole is in the set $Q_p^{1\overline{D}}$ if and only if $w \subset u^{k_1}u_1au_2u^{k_2}$ for some $u_1,\ u_2 \in V^*$ and $u = u_1u_2$, $a \in V$, $k_1,\ k_2 \geq 0$ and $k_1 + k_2 \geq 1$.*

Proof. The necessary and sufficient parts are proved as follows.

($\Leftarrow$) Let us consider a primitive partial word with one hole $w \subset u^{k_1}u_1au_2u^{k_2}$ where $u_1u_2 = u$ and $a \in V$. If the symbol in w which is contained in the letter a in the word $u^{k_1}u_1au_2u^{k_2}$ is deleted then the obtained partial word will be contained in exact integer power of u which is nonprimitive. Hence, w is a non-del-robust primitive partial word.

($\Rightarrow$) Let w be a primitive partial word with one hole which is non-del-robust. Therefore w can be written as $w = w_1cw_2$ for some $c \in V_\Diamond$ such that w_1w_2 is not a primitive partial word. Thus $w_1w_2 \subset u^n$ for some $u \in Q$ and $n \geq 2$. Hence, $w_1 \subset u^ru_1$ and $w_2 \subset u_2u^s$ for $r,\ s \geq 0$ and $r + s \geq 1$ such that $u_1u_2 = u$. Therefore for $r = k_1$ and $s = k_2$, $w = w_1cw_2 \subset u^{k_1}u_1au_2u^{k_2}$ where $c \subset a$.

□

Observe that the language of primitive partial words with one hole Q_p^1 is closed under cyclic permutation. Similar result also holds for the language of del-robust primitive partial words with one hole Q_p^{1D}.

Lemma 4.9. *The language of del-robust primitive partial words with one hole Q_p^{1D} is reflective.*

Proof. We prove it by contradiction. Let a partial word $w = xy \in Q_p^{1D}$ such that $yx \notin Q_p^{1D}$. Since $w \in Q_p^{1D}$, hence $w \in Q_p^1$. So $yx \subset u^{k_1}u_1au_2u^{k_2}$ for some k_1 and k_2 where $u = u_1u_2$, $a \in V$ and $k_1 + k_2 \geq 1$. We consider two cases depending upon the inclusion of the letter a either in y or in x.

Case 1 When the letter a is included in y, we consider two subcases.

Case 1.1 If u_1au_2 is completely in y, then $y \subset u^{k_1}u_1au_2u^ru_1'$ and $x \subset u_2'u^s$ where $u_1'u_2' = u$, $r + s + 1 = k_2$. Now $xy \subset u_2'u^su^{k_1}u_1au_2u^ru_1'$ which will not be del-robust after deletion of the symbol for which a is replaced and the obtained partial word will be nonprimitive. This is a contradiction to the assumption that $xy \in Q_p^{1D}$.

Case 1.2 If u_1au_2 is not completely in y, that is, some portion of u_2 is in y, then $y \subset u^{k_1}u_1au_2'$ and $x \subset u_2''u^{k_2}$ where $u = u_1u_2$ and $u_2 = u_2'u_2''$. Now $xy \subset u_2''u^{k_2}u^{k_1}u_1au_2'$ which will not result in a del-robust primitive partial word after deletion of the symbol for which a is replaced. Moreover, the partial word will be contained in $(u_2''u_1u_2')^{k_1+k_2+1}$ and xy is a nonprimitive partial word. Hence it is a contradiction.

Case 2 If a is included in x, we need to consider two cases with similar proofs as in the previous case.

Hence the language of del-robust primitive partial words with one hole Q_p^{1D} is reflective. □

The next result immediately follows from Lemma 4.9.

Corollary 4.10. *$Q_p^{1\overline{D}}$ is reflective.*

Next we prove that if a primitive partial word w with one hole is del-robust then its reverse w^r is also del-robust.

Lemma 4.11. *If $w \in Q_p^{1D}$ then $w^r \in Q_p^{1D}$.*

Proof. We prove it by contradiction. Let w be a del-robust primitive partial word with one hole. Suppose, the reverse of w, w^r is not del-robust, that is, $w \in Q_p^{1D}$ but $w^r \notin Q_p^{1D}$. Using the structural characterization of non-del-robust primitive partial words, we have $w^r \subset u^m u_1 a u_2 u^n$ where $u = u_1u_2$, $m + n \geq 1$. Now,

$$\begin{aligned} (w^r)^r &= w \subset (u^m u_1 a u_2 u^n)^r \\ (u^m u_1 a u_2 u^n)^r &= (u^r)^n (u_2)^r a^r (u_1)^r (u^r)^m \end{aligned}$$

Since $u = u_1u_2$, we have $u^r = (u_2)^r(u_1)^r$. Thus,

$$w \subset ((u_2)^r(u_1)^r)^n (u_2)^r a (u_1)^r ((u_2)^r(u_1)^r)^m.$$

It is clear that w is non-del-robust which is a contradiction to the assumption. Hence for $w \in Q_p^{1D}$, $w^r \in Q_p^{1D}$. □

Our next result is related to Theorem 4.8 which shows that a primitive partial word with one hole is non-del-robust if it is contained in a cyclic permutation of a nonprimitive word followed by a symbol.

Proposition 4.12. *A primitive partial word w with one hole is in the set $Q_p^{1\overline{D}}$ if and only if w is either contained in $u^n a$ or in its cyclic permutation for some $u \in Q_p$ where $a \in V$, $n \geq 2$, and $|\alpha(u)| \geq 2$.*

Proof. We prove the necessary and sufficient conditions as follows:

(⇒) Let $w \in Q_p^1$ be non-del-robust, that is, $w \in Q_p^{1\overline{D}}$. So w is contained in the word which is of the form $u^p u_1 c u_2 u^q$ for some $u = u_1u_2 \in Q$ and $c \in V$. By Corollary 4.10, $Q_p^{1\overline{D}}$ is reflective. Then consider the partial word w' which is a cyclic permutation of w and contained in the word $u_2 u^q u^p u_1 c$. Hence $w' \subset u_2 u^q u^p u_1 c = (u_2u_1)^{p+q+1} c$ is also in $Q_p^{1\overline{D}}$. Thus, w will also be contained in a word of the form $u^n a$ or in its cyclic permutation.

(⇐) Let w be either contained in the word $u^n a$ or its cyclic permutation. Deletion of that symbol in w for which the symbol a is replaced from $u^n a$ gives a nonprimitive partial word $w' \subset u^n$ which is nonprimitive. Hence, $w \in Q_p^{1\overline{D}}$.

Based on Proposition 2.42, there is an algorithm that recognizes whether a given partial word w with at most one hole is primitive and use the fact that if w is primitive and $ww \uparrow xwy$ then it implies that either $x = \varepsilon$ or $y = \varepsilon$. Observe that if a partial word with one hole w is in set $Q_p^{1\overline{D}}$ then there exists a cyclic permutation of w that contains a nonprimitive partial word of length $|w| - 1$. Also note that ww contains all the cyclic permutations of w.

In the following result we give a necessary and sufficient condition for a primitive partial word with one hole to be non-del-robust which follows from the Theorem 4.8.

Corollary 4.13. *Let w be a primitive partial word with one hole. Then w is non-del-robust if and only if ww contains at least one nonprimitive partial word factor of length $|w| - 1$.*

Proof. The necessary and sufficient conditions are proved as follows.

($\Rightarrow$) Let w be a primitive partial word with one hole. Let $w \in Q_p^{1\overline{D}}$. Since w is non-del-robust then $w \subset u^p u_1 a u_2 u^q$ for some $u \in Q$, $u_1 u_2 = u$, $a \in V$ and $p + q \geq 1$. So, $ww \subset u^p u_1 a u_2 u^q u^p u_1 a u_2 u^q$ which has a factor $u_2 u^q u^p u_1 = (u_2 u_1)^{p+q+1}$ of length $|w| - 1$. Since $p + q + 1 \geq 2$, so $(u_2 u_1)^{p+q+1}$ is a nonprimitive word. Hence a factor of ww contains one nonprimitive partial word of length $|w| - 1$.

($\Leftarrow$) Let the partial word ww has a nonprimitive factor of length $|w| - 1$ where $w \in Q_p^1$. So, $ww \subset u_1 v^n u_2$, where u_1, $u_2 \in V^*$, $|v^n| = |w| - 1$, $v \in Q$ and $n \geq 2$. There are two cases depending upon whether v^n entirely lies in w or partly in w.

Case 1 Let v^n, $n \geq 2$ is entirely lies in w. Since $|v^n| = |w| - 1$, so $w \subset v^n a$. The deletion of that particular symbol $c \in V_\Diamond$ for which the symbol a is replaced will give a nonprimitive partial word. Hence, w is non-del-robust.

Case 2 Let v^n be partially in w and $ww \subset u_1 v^n u_2$. Since the language of nonprimitive words Z is reflective, so $u_2 u_1 v^n \in Z$. So there exists a partial word x such that $xx \subset u_2 u_1 v^n$ where x is a cyclic permutation of w. If v^n is entirely in x, then either $x \subset a v^n$ or $x \subset v^n a$. In both cases the word x is a non-del-robust primitive partial word. As w is a cyclic permutation of x and Q_p is reflective, we conclude that w is a non-del-robust partial word.

□

4.2.1 $Q_p^{1\overline{D}}$ is not Context-free

In this section, we investigate the relation between the language of non-del-robust primitive partial words with one hole $Q_p^{1\overline{D}}$ and the language classes in Chomsky hierarchy. In particular, we show that $Q_p^{1\overline{D}}$ is not a context-free language (CFL) over a nontrivial alphabet V.

Theorem 4.14. *$Q_p^{1\overline{D}}$ is not a context-free language.*

Proof. We prove it by contradiction. Let us assume that $Q_p^{1\overline{D}}$ is a CFL. Let $n > 0$ be an integer which is the pumping length that is guaranteed to exist by pumping lemma. Since $Q_p^{1\overline{D}}$ is context-free, then it satisfies all the conditions of Lemma A.3. Consider a string $w = a^n b^n a^n b^n a^n b^n \Diamond$ where $a, b \in V$ and are distinct. It is clear that $w \in Q_p^{1\overline{D}}$ and of length at least n.

Hence, by the pumping lemma for CFL, w can be factorized into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq n$ and for all $i \geq 0$, $uv^ixy^iz \in Q_p^{1\overline{D}}$. There are several possibilities, that we consider below, depending on whether the substrings v and y contain more than one symbol or hole.

Case 1 When both v and y contain one type of symbol, that is, v does not contain both a's and b's, and same holds for y. Consider one such case. Let v and y contain only a's from the first set of a's. Let $vy = a^k$ for some $k > 0$. Let $u = a^j$, $vxy = a^p$ and $z = a^q b^n a^n b^n a^n b^n \Diamond$ such that $j \geq 0$ and $j + p + q = n$. Now for $i = 0$, we have $uv^ixy^iz = a^j a^{p-k} a^q b^n a^n b^n a^n b^n \Diamond = a^{j+p+q-k} b^n a^n b^n a^n b^n \Diamond = a^{n-k} b^n a^n b^n a^n b^n \Diamond \notin Q_p^{1\overline{D}}$ for $k > 1$ and for $k = 1$, we have $a^{n-1} b^n a^n b^n a^n b^n \Diamond \subset (a^{n-1} b^n a)^3 \notin Q_p^{1\overline{D}}$.

Similar cases can be handled if both v and y contain only symbol b.

Case 2 If v and y contain more than one type of symbol. There will be several cases depending upon whether v contains combinations of a's and b's and y contains only one type of symbol or v contains one type of symbol and y contains the combination of symbols or x contains combinations of a's and b's. Let us consider one such case.

Let $vxy = a^j b^k$ for some j and k such that $0 < j + k \leq n$. Observe that $j > 0$, $k > 0$ otherwise it will fall into Case 1. Suppose $u = a^l$, $v = a^{j_1}$, $x = a^{j_2}$, $y = a^{j_3} b^k$ and $z = b^p a^n b^n a^n b^n \Diamond$ such that $j_1 + j_2 + j_3 = j$, $l + j = n$, and $k + p = n$. For $i = 0$, the string $uv^ixy^iz = a^{l+j_2} b^p a^n b^n a^n b^n \Diamond \notin Q_p^{1\overline{D}}$ as $l + j_2 < n$ and $p < n$.

Similarly other cases in which v and y contain more than one symbol can be handled.

Case 3 Let us consider the last case. If $vxy = b^p \Diamond$ then there are following possibilities:

(i) If the symbol $\Diamond$ is in vy then $vy = b^l \Diamond$ and $x = b^{p-l}$. For $i = 0$, $uv^ixy^iz = a^n b^n a^n b^n a^n b^{n-p} \notin Q_p^{1\overline{D}}$.

(ii) If the symbol $\Diamond$ is in x then $v = b^l$, $y = \varepsilon$ and $x = b^{p-l} \Diamond$ and $l \geq 1$. Now, $uv^ixy^iz = a^n b^n a^n b^n a^n b^{n-l} \Diamond \notin Q_p^{1\overline{D}}$ for $i = 0$.

Observe that one of the above cases will occur. Since all the above cases result in a contradiction, the assumption that the language of non-del-robust primitive partial words with one hole $Q_p^{1\overline{D}}$ is context-free is not true. □

4.2.2 Counting n-Length Partial Words in Q_p^{1D}

In this section, we give a lower bound on the number of del-robust primitive partial words with one hole of length n. We use the following notations in counting. Denote by $Q_p^{1D}(n)$ (respectively, $Q_p^{1\overline{D}}(n)$) the set of del-robust (respectively, non-del-robust) primitive partial words with one hole of length n over an alphabet V.

We use some existing results and notations from [9] on counting of primitive words as well as primitive partial words. Let $\mathcal{P}_{h,k}(n)$ (respectively, $\mathcal{N}_{h,k}(n)$) denote the set of primitive (respectively, nonprimitive) partial words with h holes of length n over an alphabet of size k. Also, denote by $P_{h,k}(n)$ (respectively, $N_{h,k}(n)$) the number of primitive (respectively, nonprimitive) partial words with h holes of length n over an alphabet of size k. Let $T_{h,k}(n)$ denote the total number of partial words of length n with h holes over V. We have, from [13],

$$\begin{aligned} T_{h,k}(n) &= P_{h,k}(n) + N_{h,k}(n), \\ &= \binom{n}{h} k^{n-h} = \frac{n!}{h!(n-h)!} k^{n-h} \end{aligned}$$

Theorem 4.15 ([9]). $N_{1,k}(n) = nN_{0,k}(n)$.

A consequence of Theorem 4.15 is as follows.

Corollary 4.16 ([9]). $P_{1,k}(n) = n(P_{0,k}(n) + k^{n-1} - k^n)$.

Next we give an upper bound on the number of non-del-robust primitive partial words of length n over an alphabet of size k.

Theorem 4.17. *The following inequality holds:*

$$|Q_p^{1\overline{D}}(n)| \leq \begin{cases} [nk-(n-2)]\, N_{1,k}(n-1) - 2k(n-1) & \text{if } 2 \not| \text{ (n-2)}, \\ [nk-(n-2)]\, N_{1,k}(n-1) - 2k(n-1) - 2k^2 & \text{else.} \end{cases}$$

Proof. We give a constructive proof of the theorem. A non-del-robust primitive partial word with one hole of length n can be obtained in two ways; either by inserting a hole $\Diamond$ in a nonprimitive word w, $|\alpha(w)| \geq 2$ of length $(n-1)$ or inserting a symbol $a \in V$ in a nonprimitive partial word with one hole w of length $(n-1)$ with $|\alpha(w)| \geq 2$. But the first case results in repetition because all those partial words that will be generated by inserting $\Diamond$ in a word $w \in Z_{n-1}$ are already in the second case because the words are obtained from partial words by replacing holes. Thus we will consider only the second case, that is, inserting a symbol $a \in V$ in a partial word $w \in \mathcal{N}_{1,k}(n-1)$ to obtain non-del-robust primitive partial words with one hole of length n.

Take a partial word $w \in \mathcal{N}_{1,k}(n-1)$. The number of new partial words that will be generated by inserting a symbol $a \in V$ in w is

$$|\{w_1 a w_2 \mid w = w_1 w_2,\ w_1, w_2 \in V_\Diamond^*\}| = n - |w|_a$$

Now for a partial word $w \in \mathcal{N}_{1,k}(n-1)$, the total number of partial words generated by inserting any symbol $a \in V$ is

$$\sum_{a \in V} |\{w_1 a w_2 \mid w = w_1 w_2,\ w_1, w_2 \in V_\Diamond^*\}| = nk - (n-2)$$

where k is the size of the alphabet.

As we can observe that there are nonprimitive partial words with one hole which remains nonprimitive by inserting a symbol $a \in V$ and we need to discard such partial words from consideration. Such partial words are either of the form $\{w \mid w \subset a^n\}$ or $\{x \Diamond x \mid x \in V^+\}$.

First, we consider partial words with one hole w of the form $w \subset a^n$ for some $a \in V$. Since position of a hole generates different partial words, the number of such partial words that are contained in power of a particular symbol is $(n-1)$; and hence there are total $k(n-1)$ nonprimitive partial words over an alphabet of size k. Inserting the symbol $a \in V$ before and after $\Diamond$ gives two different partial words. Thus, the total number of partial words is $2k(n-1)$.

Also, there are some nonprimitive partial words which are contained in a^{n-1} but inserting a symbol b where $b \neq a$ gives a nonprimitive partial word. For example, $aa\Diamond aa$. Inserting $b \neq a$ either in suffix (respectively, prefix) gives us $aa\Diamond aab$ (respectively, $baa\Diamond aa$) which are nonprimitive. Thus, for a partial word with one hole which is contained in a^{n-1}, placing a letter either in suffix or prefix generate two nonprimitive partial words. For partial word $w \subset a^{n-1}$, $a \in V$, there are $2k$ ways to generate a nonprimitive partial word. Since $|V| = k$, there are total $2k^2$ ways to generate a nonprimitive partial words. Those case in which $(n-2)$ is divisible by 2, we have to subtract some extra partial words. Hence,

$$|Q_p^{1\overline{D}}(n)| \leq \begin{cases} [nk-(n-2)]\, N_{1,k}(n-1) - 2k(n-1) & \text{if } 2 \nmid \text{(n-2)}, \\ [nk-(n-2)]\, N_{1,k}(n-1) - 2k(n-1) - 2k^2 & \text{else.} \end{cases}$$

□

Now the lower bound on the count of number of del-robust primitive partial words with one hole of length n can be obtained by subtracting the number of non-del-robust primitive words with one hole of length n from the total number of primitive partial words of length n with one hole. This gives us

$$|Q_p^{1D}(n)| = P_{1,k}(n) - |Q_p^{1\overline{D}}(n)|$$

$$|Q_p^{1D}(n)| \geq P_{1,k}(n) - \begin{cases} [nk-(n-2)]\, N_{1,k}(n-1) - 2k(n-1) \text{ if } 2 \nmid (n-2), \\ [nk-(n-2)]\, N_{1,k}(n-1) - 2k(n-1) - 2k^2 \text{ else.} \end{cases}$$

4.3 Exchange-Robust Primitive Partial Words

In this section, we consider the set of primitive partial words with one hole which are robust to exchange operation. A partial word in exchange-robust class remains primitive when any two consecutive distinct symbols in a partial word are exchanged. For example, consider $w = abaaa\Diamond \in Q_p$, but exchanging position 2 and 3 we have $w' = aabaa\Diamond \subset (aab)^2 \notin Q_p$.

Definition 4.18 (Exchange-Robust Primitive Partial Words with One Hole)**.** *A primitive partial word $w = a_0a_1 \cdots a_{i+1}a_{i+2} \cdots a_{n-1}$ of length n with one hole is said to be exchange-robust if and only if*

$$\text{pref}(w,i)\ .\ a_{i+1}a_i\ .\ \text{suff}(w, n-i-2)$$

is a primitive partial word for all $i \in \{0,\ 1, \ldots n-2\}$.

Remark 4.19. *We consider the hole $\Diamond$ and a symbol from alphabet are distinct. If the symbol a and the hole $\Diamond$ are adjacent, we exchange a and $\Diamond$.*

We denote Q_p^{1X} as the set of all primitive partial words with one hole which are exchange-robust over an alphabet V. Clearly, the set of all exchange-robust primitive partial words with one hole is a subset of Q_p^1. There are infinitely many primitive partial words with one hole which are exchange-robust. For example, $a^n\Diamond b^n a,\ n \geq 2$ is exchange-robust.

Our next result concerns the exchange of two different symbols at consecutive places in a nonprimitive total word. We prove that the new word obtained by exchanging any two distinct consecutive symbols at any position in a nonprimitive word results a primitive word.

Lemma 4.20. *Let w be a total word with $|\alpha(w)| \geq 2$. If $w = x_1abx_2 \in Z$ with $a \neq b$ then $x_1bax_2 \in Q$.*

Proof. We prove it by contradiction. Let w be a nonprimitive word. Then there exists a unique primitive word u such that $w = u^m,\ m \geq 2$. We can express $w = u^{m_1}u_1abu_2u^{m_2}$ where $u_1abu_2 = u$ and $m_1,\ m_2 \geq 0,\ m_1 + m_2 \geq 1$. Assume for contradiction that $w' = u^{m_1}u_1bau_2u^{m_2} \notin Q$. As we know the languages Q and Z are reflective, then it is enough to consider the word $abu_2u^{m_2}u^{m_1}u_1$. Suppose $abu_2u^{m_2}u^{m_1}u_1 = v^m$ and $bau_2u^{m_2}u^{m_1}u_1 = y^n,\ m,\ n \geq 2$ and $y \in Q$. Let p be the common suffix of v^m and y^n. The words v^m and y^n have common suffix of length $m|v| - 2$ and $n|y| - 2$, respectively. We have $|p| = m|v| - 2 = n|y| - 2$. It is not possible to have $m = n = 2$ which is not feasible.

So at least one of m and n is strictly greater than 2. Without loss of generality, let us assume that $m \geq 3$ and $n \geq 2$. Now,

$$\begin{aligned} & 2|p| = m|v| + n|y| - 4 \\ \Rightarrow\ & |p| = \tfrac{m}{2}|v| + \tfrac{n}{2}|y| - 2 \\ \Rightarrow\ & |p| \geq |y| + |v| + \tfrac{1}{2}|v| - 2\ (\because m \geq 3 \text{ and } n \geq 2) \end{aligned}$$

Since $|v| \geq 2$, we obtain that $|p| \geq |y| + |v| - 1$. Hence by Theorem 2.27, v and y are powers of the same primitive word which is a contradiction. Thus $bau_2u^{m_2}u^{m_1}u_1 \in Q$ which implies that $w' = u^{m_1}u_1bau_2u^{m_2} \in Q$.

□

The above result does not hold for partial words having one or more holes. Consider a partial word $w = a\Diamond baab \in Z_p$. If we exchange b and a, we have $w' = a\Diamond abab \subset (ab)^3 \notin Q_p$.

Next we study the class of primitive partial words with one hole in which exchange of two distinct consecutive symbols results in a nonprimitive partial word. We call this set of partial words as non-exchange-robust primitive partial words with one hole. We denote the set of non-exchange-robust primitive partial words with one hole over an alphabet V as $Q_p^{1\overline{X}}$. Observe that $Q_p^{1\overline{X}} \cup Q_p^{1X} = Q_p^1$.

Definition 4.21 (Non-exchange-robust Primitive Partial Words)**.** *A primitive partial word with one hole is said to be non-exchange-robust if and only if exchange of two distinct consecutive symbols results in a nonprimitive partial word.*

In the following theorem, we give the structural characterization of non-exchange-robust primitive partial words with one hole.

Theorem 4.22. *A primitive partial word w with one hole is non-exchange-robust if and only if w is contained in some word of the form $u^{k_1}u_1abu_2u^{k_2}$, $a,\ b \in V$, $a \neq b$ where $k_1 + k_2 \geq 1$ and $u_1bau_2 = u$.*

Proof. We prove the necessary and sufficient conditions as follows:
($\Leftarrow$) Let w be a primitive partial word with one hole. Suppose $w = v_1xyv_2 \subset u^{k_1}u_1abu_2u^{k_2}$ where $a \neq b$ such that $v_1 \subset u^{k_1}u_1$, $v_2 \subset u_2u^{k_2}$, $x \subset a$ and $y \subset b$. If we exchange x and y, we get $w' = v_1yxv_2 \subset u^{k_1}u_1bau_2u^{k_2}$ such that $u_1bau_2 = u^m$ for $m \geq 1$. Hence $w' \subset u^k$, $k \geq 2$ where $k_1 + m + k_2 = k$ and thus w is not an exchange-robust primitive partial word.

($\Rightarrow$) Let $w \in Q_p^1$ which is not an exchange-robust partial word. Then there exists at least one consecutive positions where exchanging them makes the partial word nonprimitive. The partial word w can be written as either v_1abv_2 where $v_1,\ v_2 \in V_\Diamond^*$ or $v_1a\Diamond v_2$ or $v_1\Diamond av_2$ where $v_1,\ v_2 \in V^*$. In the first case, as we have exactly one hole, it is exactly in one among v_1 or v_2. Let $w' = v_1bav_2 \in Z_p$ that is $w' = v_1bav_2 \subset u^m$ for $m \geq 2$. Now $v_1 \subset u^iu_1$ and $v_2 \subset u_2u^j$ for $i,\ j \geq 0$ and $i + j \geq 1$. Combining both we have $v_1bav_2 \subset u^iu_1bau_2u^j$ where $u_1bau_2 = u^k$ for $k \geq 1$.

The other two cases can be handled similarly.

□

We have, $Q_p^{1\overline{X}} = Q_p^1 \setminus Q_p^{1X}$. There are primitive partial words of arbitrary length which are non-exchange-robust; for example, $(ab)^n\Diamond a(ab)^n$ for $n \geq 1$. We denote the set of exchange-robust (non-exchange-robust, respectively) primitive partial words with arbitrary number of holes by Q_p^X ($Q_p^{\overline{X}}$, respectively).

The set of $Q_p^{1\overline{X}}$ is not closed under the cyclic permutation unlike the language of del-robust primitive partial words with one hole. For example, consider a partial word $abbabb\Diamond ab \in Q_p^{1\overline{X}}$. One of the cyclic permutation of the partial word is $ababbabb\Diamond$, which is exchange-robust.

In Theorem 3.11, we prove that the language of primitive partial words with at most one hole is dense over an alphabet V. Next we show that the language of exchange-robust primitive partial words with one hole is not dense.

Proposition 4.23. *The language Q_p^{1X} is not right 1-dense.*

Proof. It is sufficient to find one partial word for which we cannot find any word which satisfies the condition. Let $w = x\Diamond x$ be a primitive partial word with one hole where $x \in V^*$. Let us assume that w is not an exchange-robust primitive partial word. Here both wa and wb are not primitive. Hence, we cannot find a word z with $|z| \leq 1$ for w such that $wz \in Q_p^{1X}$. Thus Q_p^{1X} is not right 1-dense.

□

For the above result, such partial word exists. For example, take $w = aaba\Diamond aaba \notin Q_p^{1X}$ and if we concatenate a or b at the right end of w then we obtain a nonprimitive partial word.

4.4 $Q_p^{\overline{X}}$ is not Context-free

In this section, we prove that the language of non-exchange-robust primitive partial words is not a context-free language over a given alphabet. In our proof, we use the fact that intersection of a context-free language and a regular language is also a context-free language. We also use the result that the family of context-free languages are closed under generalized sequential machine (gsm) mapping, and for details; see [36].

Theorem 4.24. *The language of non-exchange-robust partial words is not context-free over the alphabet $V = \{a,\ b\}$.*

Proof. Consider the regular language $R = ba^+ba^+ba^+ba^+$. Consider the language

$$L = \{ba^{n_1}ba^{n_2}ba^{n_3}ba^{n_4} \mid n_1,\ n_2,\ n_3,\ n_4 \geq 1,\ (|n_1 - n_3| \leq 1,\ |n_2 - n_4| \leq 1,\\ |(n_1 + n_2) - (n_3 + n_4)| = 0 \text{ or } 2) \text{ and } (n_1 \neq n_3 \text{ or } n_2 \neq n_4)\} \quad (4.1)$$

We claim that $Q_p^{\overline{X}} \cap R = L$.

The inclusion $Q_p^{\overline{X}} \cap R \supseteq L$ is easy to observe. For the converse, let us take a word $w = ba^{n_1}ba^{n_2}ba^{n_3}ba^{n_4} \in Q_p^{\overline{X}} \cap R$. As $w \in Q_p^{\overline{X}}$, then w can be represented as $w = u_1abu_2$ such that $u_1bau_2 \in Z$. We have the following possibilities of exchanging.

Case 1. $aba^{n_1-1}ba^{n_2}ba^{n_3}ba^{n_4}$

Case 2. $ba^{n_1-1}ba^{n_2+1}ba^{n_3}ba^{n_4}$

Case 3. $ba^{n_1+1}ba^{n_2-1}ba^{n_3}ba^{n_4}$

Case 4. $ba^{n_1}ba^{n_2-1}ba^{n_3+1}ba^{n_4}$

Case 5. $ba^{n_1}ba^{n_2+1}ba^{n_3-1}ba^{n_4}$

Case 6. $ba^{n_1}ba^{n_2}ba^{n_3-1}ba^{n_4+1}$

Case 7. $ba^{n_1}ba^{n_2}ba^{n_3+1}ba^{n_4-1}$

Observe that all the above cases are in the language $Q_p^{\overline{X}}$ only if we have

(1) $n_1 \neq n_3$ or $n_2 \neq n_4$ (otherwise $ba^{n_1}ba^{n_2}ba^{n_1}ba^{n_2} \notin Q$)

(2) $|n_1 - n_3| \leq 1,\ |n_2 - n_4| \leq 1,\ |(i+j)-(k+l)| = 0$ or 2 (otherwise the word $w' \in Q_p^{X}$

Hence the inclusion $Q_p^{\overline{X}} \cap R \subseteq L$.

As we know that a CFL is closed under the gsm mapping then using a sequential transducer (a gsm), the language $Q_p^{\overline{X}} \cap R$ can be translated into a new language

$$L' = \{a^{n_1}b^{n_2}c^{n_3}d^{n_4} \mid n_1,\ n_2,\ n_3,\ n_4 \geq 1,\ |n_1 - n_3| \leq 1,\ |n_2 - n_4| \leq 1,\ |(n_1+n_2)-(n_3+n_4)| = 0 \text{ or } 2 \text{ and } (n_1 \neq n_3 \text{ or } n_2 \neq n_4)\} \quad (4.2)$$

Now we prove that L' is not a context-free language. Assume for contradiction that L' is context-free. Let $N > 0$ be a constant which exists by the pumping lemma. As L' satisfies Ogden's lemma (see Appendix A.4), then every $w \in L'$, $|w| \geq N$ can be decomposed into $w = uvxyz$ such that the following conditions hold: (i) vxy contains at most N marked symbols, (ii) v and y have at least one marked symbol, (iii) and $uv^ixy^iz \in L'$ for all $i \geq 0$.

Consider a string $w = a^{n_1}b^{n_2}c^{n_3}d^{n_4}$ such that $n_1 = N,\ n_2 = N,\ n_3 = N+1$ and $n_4 = N-1$. As $|n_1 - n_3| \leq 1,\ |n_2 - n_4| \leq 1,\ |(n_1+n_2)-(n_3+n_4)| = 0$ and $n_1 \neq n_3,\ n_2 \neq n_4$ then $w \in L'$. Let us mark all the occurrences of b which are at least N of them. Now we can decompose $w = uvxyz$ in such a way that all the conditions of Ogden's lemma are satisfied.

Clearly, neither v nor y contain two different symbols. There are two cases depending on whether vy contains an occurrence of a or not.

(I) Suppose vy does not contain any occurrence of a. In this case, we have $u = a^Nb^{i_1}$, $v = b^{m_1}$, $x = b^{m_2}$, $y = b^{m_3}$ such that $m_1 + m_3 \geq 1$, $k_1 = m_1 + m_2 + m_3$ and $z = b^{N-(k_1+i_1)}c^{N+1}d^{N-1}$. For $i = 2$, $uv^2xy^2z = a^Nb^{N+(m_1+m_3)}c^{N+1}d^{N-1} = a^{p_1}b^{p_2}c^{p_3}d^{p_4}$ which is a contradiction as $|p_2 - p_4| \geq 2$.

(II) Suppose vy contains occurrences of a. Let $v = a^j$ and $y = b^k$ for $j,\ k \geq 1$. If $j < k$, then for a large value of i, we can have $w' = uv^ixy^iz = a^{p_1}b^{p_2}c^{p_3}d^{p_4}$ such that $|p_1 - p_3| > 1$ which is a contradiction. Therefore we must have $j \geq k$. Consider the word uv^ixy^iz which becomes $a^{N-j+ji}b^{N-k+ki}c^{N+1}d^{N-1}$. For $i = 5$, we have $w'' = a^{N+4j}b^{N+4k}c^{N+1}d^{N-1}$ where $|(N+4j)-(N+1)| = 4j-1 \geq 3$, $|(N+4k)-(N-1)| = 4k+1 \geq 5$ and $|(N+4j+N+4k)-(N+1+N-1)| = 4(j+k) \geq 8$ which is a contradiction.

Hence L' is not context-free. As we know that the family of context-free languages is closed under sequential transducers and intersection with regular languages, we conclude that $Q_p^{\overline{X}}$ is also not context-free.

□

4.5 Subst-Robust Primitive Partial Words

In this section, we study another subset of primitive partial words that remain primitive on substitution of a symbol by another symbol from the underlying alphabet.

We refer to the definition of substitute robust total words [78] and define symbol substitution in partial words as follows. Consider a nonempty partial word with one hole $x \in V_1^+$. We define

$$one(x) = \{x_1bx_2 \mid x = x_1ax_2,\ x_1,\ x_2 \in V_1^*,\ a, b\ \in V,\ a \neq b\}.$$

Let $L \subseteq V_\Diamond^*$ and $x \in L$. Then x is called subst-robust (with respect to L) if $one(x) \subseteq L$.

Definition 4.25 (Subst-Robust Primitive Partial Words)**.** *A primitive partial word w with one hole is said to be subst-robust if and only if $one(w) \subseteq Q_p^1$.*

Remark 4.26. *Since $\Diamond$ is considered as a place holder for any of the symbol from the given alphabet, we assume that only a symbol $a \in V$ can be substituted by another symbol $b \in V$ such that $a \neq b$.*

Proposition 4.27 ([78])**.** *If L consists of only subst-robust words, then $L = \{w \in V^* \mid |w| \in length(L)\}$.*

The above proposition does not hold in case of partial words with one or more holes. For example, let $L' = \{a\Diamond b, b\Diamond b, b\Diamond a, a\Diamond a\}$. Though L' is subst-robust, it does not contain all the partial words with one hole of length three.

We use the symbol Q_p^{1S} to denote the set of subst-robust primitive partial words with one hole which remain primitive on substitution of a symbol by another symbol from the given alphabet. There are infinitely many primitive partial words with one hole which are subst-robust. For example, $w = (ab)^n\Diamond$ for $n \geq 2$ is subst-robust. It is worth mentioning here that there are primitive partial words with one hole which are at the same time exchange-robust and subst-robust. An example of such partial word is

$$w_m = \Diamond aba^2b^2 \dots a^mb^m$$

for $m \geq 2$ over $V = \{a,\ b\}$.

Let $w = ab\Diamond a$ be a primitive partial word over the alphabet $V = \{a,\ b\}$. Substituting last occurrence of a by b will generate a nonprimitive partial word. Thus, there are primitive partial words with one hole in which substituting one symbol by another symbol generates a nonprimitive partial word, and we call that set of primitive partial words as non-subst-robust primitive partial words with one hole. We use the notation $Q_p^{1\overline{S}}$ to denote the set of non-subst-robust primitive partial words with one hole. Formally,

$$Q_p^{1\overline{S}} = \{w \mid w = u_1au_2 \in Q_p^1 \text{ and } w' = u_1bu_2 \notin Q_p^1, b \neq a\}.$$

Observe that $Q_p^{1S} = Q_p^1 \setminus Q_p^{1\overline{S}}$. Next, we characterize the set of primitive partial words with one hole which are not subst-robust.

Theorem 4.28. *A primitive partial word with one hole $w = xay$ where $x,\ y \in V_\Diamond^*$, $a \in V$ is not subst-robust if and only if it is contained in a word of the form $u^{k_1}u_1au_2u^{k_2}$ where $x \subset u^{k_1}u_1$, $y \subset u_2u^{k_2}$ with $k_1 + k_2 \geq 1$ such that $u_1bu_2 = u$ where $a \neq b$.*

Proof. ($\Rightarrow$) Let us assume that $w = xay$ is a non-subst-robust primitive partial word with one hole. Then there exists a position in w in which one symbol can be substituted by another symbol and makes it nonprimitive. Let $w' = xby$ be the nonprimitive partial word and hence $w' = xby \subset u^m$, $m \geq 2$ for some $u \in Q$. Therefore, we have $x \subset u^{k_1}u_1$, $y \subset u_2u^{k_2}$ such that $u_1bu_2 = u$ for $b \neq a$. Hence $w \subset u^{k_1}u_1au_2u^{k_2}$.

($\Leftarrow$) Let w be a primitive partial word with one hole and $w = xay \subset u^{k_1}u_1au_2u^{k_2}$, $k_1 + k_2 \geq 1$ where $x,\ y \in V_\Diamond^*$, $a \in V$ with $x \subset u^{k_1}u_1$ and $y \subset u_2u^{k_2}$. Also, it is given that substituting a symbol $b \neq a$, we have $u_1bu_2 = u$. If we substitute a symbol $b \neq a$ in $w = xay$, we get $w' = xby$ such that $w' = xby \subset u^{k_1}uu^{k_2} = u^{k_1+k_2+1}$ and $k_1 + k_2 + 1 \geq 2$. Hence $w = xay$ is not subst-robust. □

Next we show that the language of subst-robust primitive partial words with one hole is closed under cyclic permutation. We know that two partial words x and y are conjugate if there exist partial words u and v such that $x \subset uv$ and $y \subset vu$. A language L is closed under conjugacy relation if for every $w \in L$, the cyclic permutations of w are also in L.

Lemma 4.29. *The language Q_p^{1S} is closed under conjugacy relation.*

Proof. We prove it by contradiction. Let $w = v_1v_2$ be a primitive partial word with one hole such that $w \in Q_p^{1S}$. Suppose $w' = v_2v_1 \notin Q_p^{1S}$. Then $w = v_1v_2 \in Q_p^{1S}$ implies $w = v_1v_2 \in Q_p$. Since $w' = v_2v_1 \notin Q_p^{1S}$ then we can write $w' = v_2v_1 \subset u^{k_1}u_1au_2u^{k_2}$ such that $u_1bu_2 = u$ for some $b \in V$ and $b \neq a$. We consider two subcases depending on whether a is in v_1 or in v_2.

Case 1. If the symbol a is contained in v_2 then we consider the following possibilities.

Case 1.1 If entire u_1au_2 is from v_2 then $v_2 \subset u^{k_1}u_1au_2u^ru_1'$ and $v_1 \subset u_2'u^s$ where $u = u_1'u_2'$ and $r + s + 1 = k_2$. Now $v_1v_2 \subset u_2'u^su^{k_1}u_1au_2u^ru_1'$. Substituting a by b, we obtain a nonprimitive partial word which is a contradiction that $v_1v_2 \in Q_p^{1S}$.

Case 1.2 If a portion of u_2 is from v_2 then $v_2 \subset u^{k_1}u_1au_2'$ and $v_1 \subset u_2''u^{k_2}$ where $u_2 = u_2'u_2''$. Now $v_1v_2 \subset u_2''u^{k_2}u^{k_1}u_1au_2'$ which will result a nonprimitive partial word after the symbol a is substituted by the letter b. Moreover $v_1v_2 \subset (u_2''u_1bu_2')^{k_1+k_2+1}$ and v_1v_2 is not subst-robust primitive partial word.

Case 2. If the symbol a is not contained in v_2 then we can handle two different subcases similar to Case 1.

Hence the language of subst-robust primitive partial words Q_p^{1S} is closed under conjugacy relation. □

The following corollary is a consequence of Theorem 4.28.

Corollary 4.30. *A primitive partial word with one hole $w \in Q_p^{1\overline{S}}$ if and only if it is either contained in $u^n u' a$ or it's cyclic permutation for some $u \in Q_p$ and $u'b = u$ for $b \neq a$.*

Proof. The proof of necessary and sufficient conditions are as follows:
($\Rightarrow$) Let $w = xay$ be a primitive partial word with one hole which is not subst-robust. Then $w = xay \subset u^{k_1} u_1 a u_2 u^{k_2}$ for some $u \in Q$ such that $x \subset u^{k_1} u_1$, $y \subset u_2 u^{k_2}$, $u_1 b u_2 = u$ and $k_1 + k_2 + 1 \geq 2$. As the language $Q_p^{1\overline{S}}$ is reflective, then $yxa \subset u_2 u^{k_2} u^{k_1} u_1 a = (u_2 u_1 b)^{k_1+k_2} u_2 u_1 a = x^{k_1+k_2} x' a$ where $x = u_2 u_1 b$ and $x' = u_2 u_1$.

($\Leftarrow$) Let w be a partial word. If w is contained in $u^n u' a$ where $u'b = u$ or it's cyclic permutation then by substituting a by b where $b \neq a$, we obtain $w' \in u^{n+1}$ which is a nonprimitive partial word. Hence w is not a subst-robust primitive partial word. □

There is a linear time algorithm to test whether a given partial word with one hole is primitive which is based on the property that a primitive partial word u is not compatible to a proper factor of uu. In this chapter we defined several subclasses of primitive partial words with one hole that are robust to different operations such as del-robust primitive partial words, exchange-robust primitive partial words and subst-robust primitive partial words. It is easy to observe that whether a primitive partial word with one hole w of length n is del-robust can be tested in $O(n^2)$. Let u be a partial word of length n. A letter from each position of u can be removed, and primitivity of transformed partial word of u can be tested in $O(n)$ time. For a partial word of length n, that process can be repeated and the primitivity of partial word u can be tested in $O(n^2)$ time. Similarly, in other cases we also can have simple $O(n^2)$ time algorithm to test their respective robustness property.

4.6 Conclusions

We have discussed a special subclass of primitive partial words with one hole, referred to as del-robust primitive partial words. We have characterized such partial words and identified several properties. We have also proved that the language of non-del-robust primitive partial words with one hole $Q_p^{1\overline{D}}$ over a nontrivial alphabet is not a context-free language. We have given a lower bound on the number of such partial words of a given length.

We have also investigated the exchange-robust and substitute robust primitive partial words with one hole. The structural characterization of each of the aforementioned class of primitive partial words with respect to underlying operations have been discussed, and also some important combinatorial properties related to each of the class have been identified. We have shown that the language of non-exchange-robust primitive partial words is not a context-free language.

Chapter 5

Pseudo-Bordered Partial Words

In this chapter, we consider a variation of bordered partial words known as *pseudo-bordered partial words* for various pseudo-identity functions. Pseudo-bordered partial words is also known as θ-bordered partial words, where θ is assumed to be either a morphic involution or an antimorphic involution.

5.1 Introduction

Borderedness property of words is one of the fundamental notion in combinatorics on words and formal language theory. The set of all unbordered words is a proper subset of the set of all primitive words. The importance of bordered words have wide range of applications such as in string matching algorithms [27, 28], data compression algorithms [89], coding theory [5] and digital transmission systems [23].

One of the central problem in DNA computing is to find a good encoding scheme. In DNA computing, DNA strands are the basic blocks for proteins and carriers of information. The information encoded as a single strand of DNA contains basically the same information as its Watson-Crick(WK)-complement. The Watson-Crick complementarity property allows the information encoding strands to interact with each other to form a *double strand* in a process known as *base pairing*. The information equivalence between two strands in a biomolecule of DNA paved the road further by mathematical formalization of the Watson-Crick (WK) complementarity in DNA strands (wherein for a DNA alphabet $\triangle = \{A, T, C, G\}$, A is a WK-complement of T and vice versa, and G is a WK-complement of C and vice versa).

Motivated by the applications of primitivity, periodicity in several fundamental areas and close connection of the field of combinatorics on words with Mathematics initiated an interest to generalize these classical notions for pseudo-identity functions instead of identity functions. Several classical properties including primitivity, periodicity, conjugacy, borderedness property and palindromes have been extended for various pseudo-identity functions [29, 43, 44, 56, 57, 60, 61, 73, 74]. Bordered words have been studied with respect to homomorphisms, that is, characterization of unbordered preserving morphisms are given in [38].

Mathematically, the Watson-Crick complementarity has been modeled as an antimorphic involution θ over an alphabet V. An involution θ is a function such that θ^2 is an identity function, that is, $\theta^2(x) = x$ for all $x \in V^*$. A function $\theta : V^* \to V^*$ is said to be a morphism if $\theta(uv) = \theta(u)\theta(v)$, and an antimorphism if $\theta(uv) = \theta(v)\theta(u)$ for all $u,\ v \in V^*$. A DNA strand u over the DNA alphabet contains the same information as its WK-complement, denoted by $\theta(u)$. Basically, the WK-complement of a DNA strand is the reverse complement of the original strand. Recall from Chapter 2, a word w is bordered if and only if there exist words x, y, z such that $w = xy = yz$ and y is said to be a border of w. Kari and Mahalingam [58] defines a word w to be θ-bordered if there exists a word $x \in V^+$ that is a proper prefix of w, while $\theta(x)$ is a proper suffix of w. A word which is not θ-bordered is called θ-unbordered, and the set of all θ-unbordered words is denoted by $D_\theta(1)$.

In [66], Peter Leupold proposed the idea of using partial words as a tool to find the good encodings for DNA computing purposes. Along with the WK-complementarity for a DNA alphabet, he proposed that the "do not know symbol", $\Diamond$ is mapped to itself, that is, $\theta(\Diamond) = \Diamond$.

A nonempty partial word u is said to be bordered if there exists a nonempty prefix of u that is compatible with a nonempty suffix of u [14]. A partial word which is not bordered is called unbordered. Motivated by Peter Leupold's work in [66], and generalization of classical notion of (un)bordered[1] words to that of θ-(un)bordered[2] words, [58], we generalize the notion of (un)bordered partial words to θ-(un)bordered partial words. A nonempty partial word u is said to be θ-bordered if there exist nonempty words x, y, z such that $u \subset xy$ and $u \subset z\theta(x)$.

Organization. The outline of this chapter is as follows. We begin by reviewing basic concepts and notations of θ-bordered and θ-unbordered partial words in Section 5.2. In Section 5.3, we prove several combinatorial properties of θ-(un)bordered partial words that includes closure properties as well as characterization of the set of θ-unbordered partial words for antimorphic involution θ. In Section 5.4, we give necessary and sufficient conditions for a partial word to be θ-unbordered which is based on the concept of θ-contained prefixes and conatined suffixes. We also count the number of θ-borders of a partial word $w = (u\theta(u))^i$ where u is unbordered as well as θ-unbordered partial word. In Section 5.5, we prove that the set of all θ-unbordered partial words $D_{\theta_\Diamond}(1)$ is a disjunctive language. Section 5.6 discusses the classification of sets of θ-(un)bordered partial words and introduces new language classes in connection with θ-(un)bordered partial words. In Section 5.7, we define θ-primitive partial words for (anti)morphic[3] involution and discuss some basic properties of it. Section 5.8 defines the concept of θ-conjugacy and θ-commutativity for partial words, and provides a characterization such that one is a θ-conjugate of the other with concluding remarks in the Section 5.9.

[1]By (un)bordered word we mean either a bordered or an unbordered word.
[2]By θ-(un)bordered word we mean either a θ-bordered or a θ-unbordered word.
[3]By (anti)morphic involution, we mean either a morphic or an antimorphic involution.

5.2 Basics of θ-bordered Partial Words

In the following definition, we extend the definition of θ-(un)bordered words to θ-(un)bordered partial words. The notion of contained prefixes and contained suffixes are introduced to accommodate θ-(un)bordered partial words. We denote the set of contained prefixes and contained suffixes as *CPref* and *CSuff* respectively. Similarly, the set of proper contained prefixes and proper contained suffixes are denoted by *PCPref* and *PCSuff*, respectively. We define the notion of contained prefixes, contained suffixes, proper contained prefixes and proper contained suffixes as follows.

$$\text{CPref}(w) = \{x \in V^+ \mid \exists y \in V^*,\ w \subset xy\}$$

$$\text{CSuff}(w) = \{x \in V^+ \mid \exists y \in V^*,\ w \subset yx\}$$

$$\text{PCPref}(w) = \{x \in V^+ \mid \exists y \in V^+,\ w \subset xy\}$$

$$\text{PCSuff}(w) = \{x \in V^+ \mid \exists y \in V^+,\ w \subset yx\}$$

We denote the relations of a word u being a contained prefix, contained suffix, proper contained prefix and proper contained suffix of a partial word w by $u \leq_{cp} w$, $u \leq_{cs} w$, $u <_{cp} w$ and $u <_{cs} w$ respectively, and define them as follows.

Definition 5.1.

(a) For $w \in V_\diamondsuit^*$ and $u \in V^*$, $u \leq_{cp} w$ if and only if $w \subset uv$ for some $v \in V^+$.

(b) For $w \in V_\diamondsuit^*$ and $u \in V^*$, $u \leq_{cs} w$ if and only if $w \subset vu$ for some $v \in V^+$.

(c) For $w \in V_\diamondsuit^*$ and $u \in V^+$, $u <_{cp} w$ if and only if $w \subset uv$ for some $v \in V^+$.

(d) For $w \in V_\diamondsuit^*$ and $u \in V^+$, $u <_{cs} w$ iff $w \subset vu$ for some $v \in V^+$.

We illustrate the above definitions in the following example.

Example 5.2. Let $V = \{a, b\}$ be an alphabet and $u = a\diamondsuit b$ be a partial word over V. Then *CPref*$(u)= \{a, aa, ab, aab, abb\}$, *CSuff*$(u) = \{b, ab, bb, aab, abb\}$, *PCPref*$(u) = \{a, aa, ab\}$ and *PCSuff*$(u) = \{b, ab, bb\}$. □

In the following definition, we define the relations such as θ-contained prefix, θ-contained suffix, and θ-border of a partial word.

Definition 5.3. *Let θ be either a morphic or an antimorphic involution on $V_\diamondsuit^*$.*

(a) For $w \in V_\diamondsuit^*$ and $u \in V^*$, $u \leq_{cp}^{\theta_\diamondsuit} w$ if and only if $w \subset \theta(u)v$ for $v \in V^*$.

(b) For $w \in V_\diamondsuit^*$ and $u \in V^*$, $u \leq_{cs}^{\theta_\diamondsuit} w$ if and only if $w \subset v\theta(u)$ for $v \in V^*$.

(c) For a partial word $u \in V_\diamondsuit^*$, $x \in V^*$ is said to be a θ-border of u if $x \leq_d^{\theta_\diamondsuit} u$, that is, $u \subset xy$ and $u \subset z\theta(x)$.

(d) $\leq_d^{\theta_\Diamond} = \leq_{cp} \cap \leq_{cs}^{\theta_\Diamond}$.

(e) *For* $w \in V_\Diamond^+$ *and* $u \in V^*$, $u <_{cp}^{\theta_\Diamond} w$ *if and only if* $w \subset \theta(u)v$ *for* $v \in V^+$.

(f) *For* $w \in V_\Diamond^+$ *and* $u \in V^*$, $u <_{cs}^{\theta_\Diamond} w$ *if and only if* $w \subset v\theta(u)$ *for* $v \in V^+$.

(g) $<_d^{\theta_\Diamond} = <_{cp} \cap <_{cs}^{\theta_\Diamond}$.

(h) *For* $w \in V_\Diamond^*$, $u \in V^*$ *is said to be a proper* θ*-border of* w *if* $u <_d^{\theta_\Diamond} w$.

(i) *For* $w \in V_\Diamond^+$, *define by* $L_d^{\theta_\Diamond}(w) = \{u \in V^* \mid u <_d^{\theta_\Diamond} w\}$, *the set of all* θ*-borders of a partial word* w.

(j) $\nu_d^{\theta_\Diamond}(w) = |L_d^{\theta_\Diamond}(w)|$.

(k) *Denote by* $D_{\theta_\Diamond}(i) = \{w \in V_\Diamond^+ \mid \nu_d^{\theta_\Diamond}(w) = i\}$, *the set of all partial words with exactly* i θ*-borders for* $i \geq 1$.

From the above definitions, a partial word $u \in V_\Diamond^+$ is said to be θ-bordered if $u \subset xy$ and $u \subset z\theta(x)$ for $y, z \in V^+$. In Definition 5.3, if θ is an identity function then the relations $\leq_{cp}^{\theta_\Diamond}$, $\leq_{cs}^{\theta_\Diamond}$ and $\leq_d^{\theta_\Diamond}$ become contained prefix, contained suffix, and borders respectively. The set of all partial words which are not θ-bordered are called θ-unbordered and is denoted by $D_{\theta_\Diamond}(1)$. Note that ε is a θ-border of every nonempty partial word.

Example 5.4. Let θ be a morphic involution such that $\theta(a) = b, \theta(b) = a, \theta(\Diamond) = \Diamond$. Consider the partial word $u = a\Diamond ab$ over $V = \{a, b\}$. Then $L_d^{\theta_\Diamond}(u) = \{\varepsilon, a, aba\}$ and $\nu_d^{\theta_\Diamond}(u) = 3$. Hence $u \in D_{\theta_\Diamond}(3)$. □

5.3 Properties of θ-bordered Partial Words

In this section, we discuss some properties of θ-bordered and θ-unbordered partial words which includes the transitivity, closed under cyclic permutation property, denseness of the set of all θ-unbordered partial words, for morphic as well as antimorphic involution θ. We also provide a characterization of the set of θ-bordered partial words for an antimorphic involution θ.

Remark 5.5. *The set of borders and* θ*-borders of a partial word need not be the same.*

Example 5.6. Let $u = abb\Diamond a$ and θ be a morphic involution such that $\theta(a) = b$, $\theta(b) = a$, $\theta(\Diamond) = \Diamond$. Observe that a is a border of u while ab is a θ-border of u but neither a is a θ-border of u nor ab is a border of u. □

Remark 5.7. *A bordered partial word may not be a* θ*-bordered partial word and vice versa.*

Example 5.8. Let us consider the DNA alphabet $V = \{A,\ T,\ C,\ G\}$ and θ be (anti)morphic[4] involution such that $\theta(A) = T,\ \theta(T) = A,\ \theta(C) = G,\ \theta(G) = C$ and $\theta(\Diamond) = \Diamond$. The partial word $u = A\Diamond GT$ is θ-bordered but not bordered. Similarly the partial word $v = AG\Diamond A$ is bordered but not θ-bordered. □

The following observations follow directly from the definition.

Lemma 5.9. *Let θ be either a morphic or an antimorphic involution over an alphabet $V_\Diamond$.*

(a) A θ-bordered partial word $u \in V_\Diamond^+$ has length greater than or equal to 2.

(b) For $u \in V_\Diamond^+$, if $u \subset \{u_1, u_2, \ldots, u_n\}$ then $\theta(u) \subset \{\theta(u_1), \theta(u_2), \ldots, \theta(u_n)\}$.

Lemma 5.10 ([56]). *Let θ be a morphic or an antimophic involution and let $u \in V^+ \setminus D_\theta(1)$. Then there exists $v \in V^*$ with $|v| \leq \frac{|u|}{2}$ such that v is a proper θ-border of u.*

However, when θ is a morphic involution, the above lemma does not necessarily hold for partial words. For example, let $V = \{A, C, G, T\}$ and let θ be a morphic involution such that $\theta(A) = T$, $\theta(T) = A$, $\theta(G) = C$, $\theta(G) = C$ and $\theta(\Diamond) = \Diamond$. Let $u = AG\Diamond CG$ be a partial word. Then, the θ-borders of u are $L_d^{\theta_\Diamond}(u) = \{\varepsilon, AGC\}$ and there is no proper θ-border v of u such that $|v| \leq \frac{|u|}{2}$.

It has been shown in [58] that in case of total words, if $\theta(a) \neq a$ for all $a \in V$ then $a^+ \subseteq D_\theta(1)$. However, this result does not necessarily hold in case of partial words as any partial word ending or beginning with $\Diamond$ has a trivial 1-letter border.

The next result establishes a connection between the set of all θ-borders of a partial word u and the set of all θ-borders of $\theta(u)$.

Lemma 5.11. *Let $u \in V_\Diamond^+$ be a partial word. If θ is a morphic involution then $L_d^{\theta_\Diamond}(\theta(u)) = \theta(L_d^{\theta_\Diamond}(u))$ and if θ is an antimorphic involution then $L_d^{\theta_\Diamond}(u) = L_d^{\theta_\Diamond}(\theta(u))$.*

Proof. Let θ be a morphic involution, and let $v \in L_d^{\theta_\Diamond}(u)$. This implies that $u \subset vx$ and $u \subset y\theta(v)$ for some $x, y \in V^+$. Hence $\theta(u) \subset \theta(v)\theta(x)$ and $\theta(u) \subset \theta(y)\theta(\theta(v))$. Now $\theta(v)$ is a θ-border of partial word $\theta(u)$, that is, $\theta(v) \in L_d^{\theta_\Diamond}(\theta(u))$ and hence $\theta(L_d^{\theta_\Diamond}(u)) \subseteq L_d^{\theta_\Diamond}(\theta(u))$. Similarly, let $v \in L_d^{\theta_\Diamond}(\theta(u))$. Then $\theta(u) \subset vp$ and $\theta(u) \subset q\theta(v)$ for some $p, q \in V^+$. Hence we have $u \subset \theta(v)\theta(p)$ and $u \subset \theta(q)\theta(\theta(v))$ which implies that $\theta(v) \in L_d^{\theta_\Diamond}(u)$ and hence $L_d^{\theta_\Diamond}(\theta(u)) \subseteq \theta(L_d^{\theta_\Diamond}(u))$. Thus $L_d^{\theta_\Diamond}(\theta(u)) = \theta(L_d^{\theta_\Diamond}(u))$.

Let θ be an antimorphic involution and let $v \in L_d^{\theta_\Diamond}(u)$. This implies that $u \subset vx$ and $u \subset y\theta(v)$ for some $x, y \in V^+$. Hence $\theta(u) \subset \theta(x)\theta(v)$ and $\theta(u) \subset \theta(\theta(v))\theta(y) = v\theta(y)$. Thus $v \in L_d^{\theta_\Diamond}(\theta(u))$ and hence $L_d^{\theta_\Diamond}(u) \subseteq L_d^{\theta_\Diamond}(\theta(u))$. Similarly, let $v \in L_d^{\theta_\Diamond}(\theta(u))$. Then $\theta(u) \subset vx'$ and $\theta(u) \subset y'\theta(v)$ for some $x', y' \in V^+$ which implies $u \subset \theta(x')\theta(v)$ and $u \subset v\theta(y')$. Hence $v \in L_d^{\theta_\Diamond}(u)$ and $L_d^{\theta_\Diamond}(\theta(u)) \subseteq L_d^{\theta_\Diamond}(u)$. Thus, $L_d^{\theta_\Diamond}(u) = L_d^{\theta_\Diamond}(\theta(u))$. □

[4]By (anti)morphic involution, we mean either a morphic or an antimorphic involution.

Recall that a language of partial words $L \subseteq V_\Diamond^*$ is said to be reflective if $xy \in L$ implies that $yx \in L$ for all $x, y \in V_\Diamond^*$. A reflective language can be seen as a language that contains all the cyclic permutations of a word w whenever $w \in L$. From the Definition 5.3(k), $D_{\theta_\Diamond}(i)$ is the language of partial words having exactly i θ-borders. Next, we illustrate that $D_{\theta_\Diamond}(i)$ is not necessarily closed under the relation of cyclic permutation in the following example.

Example 5.12. Consider a partial word $u = ababa\Diamond a$ over $V = \{a, b\}$ such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(\Diamond) = \Diamond$. For θ to be a morphic involution, $L_d^{\theta_\Diamond}(u) = \{\varepsilon, ab, abab, ababab\}$ and hence $u \in D_{\theta_\Diamond}(4)$. Consider $u' = aababa\Diamond$ which is a cyclic permutation of u. Then $L_d^{\theta_\Diamond}(u') = \{\varepsilon, a\}$ and hence $u' \in D_{\theta_\Diamond}(2)$. Similarly, if θ is an antimorphic involution then for $w = aa\Diamond$, $w \in D_{\theta_\Diamond}(2)$ but one of its cyclic permutation $w' = a\Diamond a \in D_{\theta_\Diamond}(1)$. □

Hence for a given $i \geq 1$, $D_{\theta_\Diamond}(i)$ is not necessarily closed under the cyclic permutation with respect to either a morphic or an antimorphic involution.

The following result shows the transitivity of the relation $<_d^{\theta_\Diamond}$ for an antimorphic involution θ, and establishes a relation between partial words u, v, w such that $u <_d^{\theta_\Diamond} v$ and $v <_d^{\theta_\Diamond} w$ when θ is a morphic involution.

Proposition 5.13. *Let $u \in V^*$, $v \in V^+$ and $w \in V_\Diamond^+$ such that $u <_d^{\theta_\Diamond} v$, $v <_d^{\theta_\Diamond} w$. If θ is a morphic involution then u is a border of w, and if θ is an antimorphic involution then $u <_d^{\theta_\Diamond} w$.*

Proof. Since $u <_d^{\theta_\Diamond} v$, we have $v = ux$ and $v = y\theta(u)$ for $x, y \in V^+$. Also $v <_d^{\theta_\Diamond} w$ implies that $w \subset vp$ and $w \subset q\theta(v)$ for some $p, q \in V^+$.

If θ is a morphic involution, then $w \subset vp = uxp$ and $w \subset q\theta(v) = q\theta(y\theta(u)) = q\theta(y)u$. Hence u is a border of w.

If θ is an antimorphic involution, then $w \subset uxp$ and $w \subset q\theta(ux) = q\theta(x)\theta(u)$ which implies that $u <_d^{\theta_\Diamond} w$. □

Corollary 5.14. *The relation $<_d^{\theta_\Diamond}$ is transitive when θ is an antimorphic involution.*

We mention the following result from [58] that establishes a relation between two different θ-borders of a total word. In particular, for a total word w, if u and v are two distinct θ-borders of w then either u is a border of v or v is a border of u when θ is a morphic involution, and u (*respectively*, v) is a prefix of v(*respectively*, u) when θ is an antimorphic involution.

Lemma 5.15 ([58]). *Let u, v, w be such that u, $v \in V^+$, $u \neq v$ and $u <_d^\theta w$ and $v <_d^\theta w$. If θ is a morphic involution, then either $v <_d u$ or $u <_d v$. If θ is an antimorphic involution then either $v <_p u$ or $u <_p v$.*

The Lemma 5.15 need not hold for partial words as shown in the following example.

Example 5.16. Let $V = \{a, b\}$ and θ be (anti)morphic involution such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(\Diamond) = \Diamond$.

(a) Let θ be a morphic involution. Let $w = a\Diamond bbba$. Then $L_d^{\theta_\Diamond}(w) = \{\varepsilon, ab, aab\}$. We can choose $u = ab$ and $v = aab$. Observe that neither $v <_d u$ nor $u <_d v$.

(b) Let θ be an antimorphic involution and $w = ab\Diamond ab$. Then $L_d^{\theta_\Diamond}(w) = \{\varepsilon, a, ab, aba, abb, abba, abaa\}$. We can choose $u = aba$ and $v = abba$ and observe that neither $v <_{cp} u$ nor $u <_{cp} v$.

□

From the Example 5.16, it is clear that for a partial word $u \in V_\Diamond^+$, the set $L_d^{\theta_\Diamond}(u)$ is not necessarily a totally ordered set with respect to the relation $<_d$ when θ is a morphic involution. Also, if θ is an antimorphic involution, then $L_d^{\theta_\Diamond}(u)$ is not necessarily a totally ordered set with respect to $<_{cp}$, and $\theta(L_d^{\theta_\Diamond}(u))$ is not necessarily a totally ordered set with respect to $<_{cs}$.

The following lemma provides a characterization of θ-bordered partial words in case of an antimorphic involution θ. We provide both necessary and sufficient condition for a partial word to be θ-bordered.

Lemma 5.17. *Let θ be an antimorphic involution. A nonempty partial word $u \in V_\Diamond^+$ is θ-bordered if and only if $u \subset av\theta(a)$ for some $a \in V$ and some $v \in V^*$.*

Proof. ($\Rightarrow$) Let u be a θ-bordered partial word. Then there exist $x, y, z \in V^+$ such that $u \subset xy$ and $u \subset z\theta(x)$. Let $x = ap$ for some $a \in V$ and $p \in V^*$. Thus $\theta(x) = \theta(ap) = \theta(p)\theta(a)$ which implies $u \subset apy$, and $u \subset z\theta(p)\theta(a)$. Hence we can write $u \subset av\theta(a)$ for some $v \in V^*$.
The converse is obvious. □

Next we study some properties related to θ-unbordered partial words and prove that the set of θ-unbordered partial words is a dense set for an antimorphic involution θ.

Corollary 5.18. *Let θ be an antimorphic involution.*

(a) For all $u \in V_\Diamond^*$, $u \in D_{\theta_\Diamond}(1)$ if and only if $\theta(u) \in D_{\theta_\Diamond}(1)$.

(b) Let $\theta(a) \neq b$ for $a, b \in V$. Then $D_{\theta_\Diamond}(1)$ is a dense set.

(c) Let u be a partial word over V which does not begin with a hole and let $a, b \in V$ be such that $\theta(a) = b$ and $a \neq b$. Then either ua or ub is θ-unbordered.

(d) For all $x, z \in V_\Diamond^+$ and $y \in V_\Diamond^*$, we have $xyz \in D_{\theta_\Diamond}(1)$ if and only if $xz \in D_{\theta_\Diamond}(1)$.

Proof.

(a) Let us assume that $\theta(u) \notin D_{\theta_\Diamond}(1)$ and $u \in D_{\theta_\Diamond}(1)$. By Lemma 5.17, $\theta(u) \subset ay\theta(a)$ for some $y \in V^*$ and $a \in V$. Thus, $\theta(\theta(u)) = u \subset a\theta(y)\theta(a)$ which is a contradiction to the fact that u is θ-unbordered. Let $\theta(u) \in D_{\theta_\Diamond}(1)$ and $u \notin D_{\theta_\Diamond}(1)$. Then $u \subset bx\theta(b)$ for some $x \in V^*$ and $b \in V$. Hence $\theta(u) \subset b\theta(x)\theta(b)$ which shows that $\theta(u)$ is a θ-bordered partial word, a contradiction.

(b) Consider a partial word $y \in V_\Diamond^+$. We can choose $x = a$ and $z = b$ such that $xyz \in D_{\theta_\Diamond}(1)$ since $\theta(a) \neq b$. Hence $D_{\theta_\Diamond}(1)$ is a dense set.

(c) Let us assume that both ua and ub are θ-bordered partial words. Then by Lemma 5.17, there exist $\alpha, \beta \in V$ and $y, z \in V^*$ such that $ua \subset \alpha y \theta(\alpha)$ and $ub \subset \beta z \theta(\beta)$ which implies $a = \theta(\alpha)$ and $b = \theta(\beta)$. Since u does not begin with a hole, we have $u = cp$ for $c \in V$ and $p \in V_\Diamond^*$, that is, $u = cp \subset \alpha y$ and $u = cp \subset \beta z$ which implies $c = \alpha = \beta$ which further implies that $a = b$ which is a contradiction.

(d) Let us assume that $xz \notin D_{\theta_\Diamond}(1)$ and $xyz \in D_{\theta_\Diamond}(1)$. Then by Lemma 5.17, $xz \subset ap\theta(a)$ for $p \in V^*$ and $a \in V$. Thus, $xyz \in D_{\theta_\Diamond}(1)$ depends only on the first letter of x and last letter of z. Since $x, z \in V_\Diamond^+$, $xyz \in D_{\theta_\Diamond}(1)$ contradicts to the assumption that $xz \notin D_{\theta_\Diamond}(1)$ and vice versa.

□

The statement (c) of above Corollary 5.18 does not hold for a partial word that begins with a hole as illustrated by the following example.

Example 5.19. Let $V = \{a, b\}$ and θ be (anti)morphic involution such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(\Diamond) = \Diamond$, and let $u = \Diamond abb\Diamond$. Then both $ua = \Diamond abb\Diamond a$ and $ub = \Diamond abb\Diamond b$ are θ-bordered partial words. □

In Lemma 5.17, we have given the characterization of θ-bordered partial words for an antimorphic involution θ, and it is clear that the minimal θ-border of a θ-bordered partial word u is $a \in V$ which itself is a θ-unbordered word. However, the minimal θ-border of a θ-bordered partial word in case of a morphic involution need not be a θ-unbordered word. For example, if for $a, b \in V$, let $\theta(a) = b$, $\theta(b) = a$ and $\theta(\Diamond) = \Diamond$, and let $u = abb\Diamond a$. Then the minimal θ-border $ab = a \cdot \theta(a)$ is not θ-unbordered.

The next result relates the minimal border and other borders of a total word under the assumption of θ being a morphic involution.

Lemma 5.20 ([58]). *Let θ be a morphic involution and $\theta(a) = a$ for all $a \in V$. If x is the minimal border of u, then for all other borders $y \neq x$ of u, y is bordered.*

The above lemma does not necessarily hold in case of partial words.

Example 5.21. Let $V = \{a, b\}$ and θ be a morphic involution such that $\theta(a) = a, \theta(b) = b$ and $\theta(\Diamond) = \Diamond$. Consider a partial word $u = a\Diamond bbbaab$. The minimal θ-border of u is ab. However aab is also a θ-border of u but aab is θ-unbordered. □

5.4 Catenation of θ-bordered Partial Words

In this section, we give a necessary and sufficient condition for a partial word u to be θ-unbordered in terms of set of contained prefixes and set of contained suffixes. We also prove that for (anti)morphic involution θ, the set of θ-unbordered partial words $D_{\theta_\Diamond}(1)$ is a post-plus language. We give a count on the number of θ-borders of a partial word of the form $((u\theta(u))^i$ where u is a unbordered partial word as well as a θ-unbordered partial word and θ is assumed to be a morphic involution.

We have a characterization of θ-bordered partial words in Lemma 5.17 when θ is an antimorphic involution. The next lemma in similar lines provides a characterization of θ-unbordered partial word for (anti)morphic involution θ in terms of set of contained prefixes and contained suffixes. Note that, the characterization provided in Lemma 5.17 is stronger than the one provided latter when θ is an antimorphic involution.

Lemma 5.22. *Let θ be either a morphic or an antimorphic involution on $V_\Diamond^*$. Then for all $u \in V_\Diamond^+$ with $|u| \geq 2$, u is θ-unbordered if and only if $\theta(PCPref(u)) \cap PCSuff(u) = \emptyset$.*

Proof. We prove from both directions as follows.
($\Rightarrow$) Let u be a θ-unbordered partial word. Assume that $\theta(\text{PCPref}(u)) \cap \text{PCSuff}(u) \neq \emptyset$. Let $x \in \theta(\text{PCPref}(u)) \cap \text{PCSuff}(u)$. Then $u \subset \theta(x)u_2$ and $u \subset u_1x$ for $u_1, u_2 \in V^+$. Thus $u \notin D_{\theta_\Diamond}(1)$, a contradiction.

($\Leftarrow$) Let $\theta(\text{PCPref}(u)) \cap \text{PCSuff}(u) = \emptyset$. Assume that u is a θ-bordered partial word. Then there exist $v, x, y \in V^+$ such that $u \subset vx$ and $u \subset y\theta(v)$ which implies $\theta(v) \in \theta(\text{PCPref}(u)) \cap \text{PCSuff}(u)$, a contradiction. □

It is clear from the previous lemma that for an antimorphic involution θ, a partial word u is θ-bordered if and only if $\theta(\text{PCPref}(u)) \cap \text{PCSuff}(u) \neq \emptyset$.

In Corollary 5.23, we give a necessary and sufficient condition for concatenation of two θ-unbordered partial words to be θ-unbordered follows from Lemma 5.22.

Corollary 5.23. *Let θ be either a morphic or an antimorphic involution and let $u, v \in V_\Diamond^+$ be two θ-unbordered partial words. Then uv is θ-unbordered partial word if and only if $\theta(CPref(u)) \cap CSuff(v) = \emptyset$.*

Proof. ($\Leftarrow$) Let $u,\ v \in V_\Diamond^+$ and assume that $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) = \emptyset$. Let uv be a θ-bordered partial word. Then we have the following cases to consider.

Case 1: Let θ be an antimorphic involution. Then $uv \subset ay\theta(a)$ for some $y \in V_\Diamond^+$ and $a \in V$. Hence $\theta(a) \in \theta(\text{CPref}(u)) \cap \text{CSuff}(v)$ which is a contradiction. Hence uv is a θ-unbordered partial word.

Case 2: Let θ be a morphic involution. Since uv is a θ-bordered partial word, $uv \subset xy$ and $uv \subset z\theta(x)$ for some $y, z \in V^+$. Then we consider following cases.

Case 2.1 If $|x| \leq |u|$ and $|\theta(x)| \leq |v|$. In this case $x \in \text{CPref}(u)$ and $\theta(x) \in \text{CSuff}(v)$. Hence $\theta(x) \in \theta(\text{CPref}(u)) \cap \text{CSuff}(v)$ which is a contradiction.

Case 2.2 If $|x| \leq |u|$ and $|\theta(x)| > |v|$, that is, uv has a θ-border longer than v. Now the corresponding θ-contained suffix of uv is $\theta(u'v')$ for some θ-contained suffix of u where $u = u_1u_2$, $u_2 \subset \theta(u')$ and $v \subset v'$. By the definition of θ-border, $\theta(u')$ matches a contained prefix of u which shows that u is θ-bordered and hence a contradiction.

Case 2.3 If $|x| > |u|$, we can argue in the similar way to Case 2.2 and show that v is a θ-bordered partial word which is a contradiction.

Since all the cases lead to contradiction, we have that uv is a θ-unbordered partial word.

($\Rightarrow$) Let $u, v \in V_\diamond^+$ and let u, v be θ-unbordered partial words. Assume that uv is also θ-unbordered partial word and $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) \neq \emptyset$. Let $x \in \theta(\text{CPref}(u)) \cap \text{CSuff}(v)$. Then $x = \theta(u_1) = v_2$ for some $u = u'u'', v = v'v'', u' \subset u_1, v'' \subset v_2$ where $u', u'', v', v'' \in V_\diamond^+$. Now, $uv = u'u''v'v'' \subset u_1u_2v_1\theta(u_1)$ where $u'' \subset u_2$ and $v' \subset v_1$ which implies that uv is a θ-bordered partial word which is a contradiction. Hence $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) = \emptyset$. □

The following result provides a necessary and sufficient condition under which a partial word u^iv^j is θ-bordered.

Corollary 5.24. *Let θ be either a morphic or an antimorphic involution and let u, v be two partial words. Then $\theta(CPref(u)) \cap CSuff(v) \neq \emptyset$ if and only if $\theta(CPref(u^i)) \cap CSuff(v^j) \neq \emptyset$ for $i, j \geq 1$.*

Proof. ($\Rightarrow$) Let us assume that $\theta(\text{CPref}(u^i)) \cap \text{CSuff}(v^j) = \emptyset$ for $i, j \geq 1$. Then clearly $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) = \emptyset$, a contradiction. Hence if $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) \neq \emptyset$, then $\theta(\text{CPref}(u^i) \cap \text{CSuff}(v^j) \neq \emptyset$

($\Leftarrow$) Assume that $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) = \emptyset$ and $\theta(\text{CPref}(u^i)) \cap \text{CSuff}(v^j) \neq \emptyset$. Then there exists $x \in V^+$ such that $x \in \theta(\text{CPref}(u^i)) \cap \text{CSuff}(v^j)$. Then we have $x = \theta(\tilde{u}^k\tilde{u_1}) = \tilde{v_2}\tilde{v}^l$ where $\tilde{u} = \tilde{u_1}\tilde{u_2}$, $\tilde{v} = \tilde{v_1}\tilde{v_2}$, $u \subset \tilde{u}$, $v \subset \tilde{v}$ such that $\tilde{u_1} \in \text{CPref}(u)$ and $\tilde{v_2} \in \text{CSuff}(v)$, $\tilde{u_1}, \tilde{v_2} \in V^*$ and $k, l \geq 1$.

Let θ be a morphic involution. Since $x = \theta(\tilde{u}^k\tilde{u_1}) = \tilde{v_2}\tilde{v}^l$, $\theta(\tilde{u_1}) \in \theta(\text{CPref}(u)) \cap \text{CSuff}(v)$ which is a contradiction.

Let θ be an antimorphic involution. Then $x = \theta(\tilde{u}^k\tilde{u_1}) = \theta(\tilde{u_1})\theta(\tilde{u}^k) = \tilde{v_2}\tilde{v}^l$. Let $u = u_1u_2 \subset \tilde{u_1}\tilde{u_2} = \tilde{u}$ such that $u_2 \in V_\diamond^+$. Hence we can write $x = \theta(\tilde{u_1})\theta((\tilde{u_1}\tilde{u_2})^k) = \theta(\tilde{u_1})\theta((\tilde{u_1}\tilde{u_2})^{k-1})\theta(\tilde{u_2})\theta(\tilde{u_1}) = \theta(\tilde{u_1})y\theta(\tilde{u_1}) = \tilde{v_2}\tilde{v}^l$ where $y = \theta((\tilde{u_1}\tilde{u_2})^{k-1})\theta(\tilde{u_2})$ which implies that $\theta(\tilde{u_1}) \in \theta(\text{CPref}(u)) \cap \text{CSuff}(v)$ which is a contradiction. □

In the next result we prove that if a partial word x_1x_2 is θ-unbordered then the partial word $x_1x_2^2$ is θ-unbordered as well.

Proposition 5.25. *Let $x_1, x_2 \in V_\diamond^+$ and θ be either a morphic or an antimorphic involution. If x_1x_2 is a θ-unbordered partial word, then $x_1x_2^2$ is also a θ-unbordered partial word.*

Proof. Let θ be an antimorphic involution. Suppose, $x_1x_2^2$ is a θ-bordered partial word. Then by Lemma 5.17, there exists $a \in V$ and $y \in V^*$ such that $x_1x_2^2 \subset ay\theta(a)$. Since x_1, x_2 are both nonempty partial words, we have $x_1x_2 \subset ax\theta(a)$ for $x \in V^*$ which is a contradiction. Hence $x_1x_2^2$ is θ-unbordered.

Let θ be a morphic involution and $x_1x_2^2$ be θ-bordered. Then there exist $u, x, y \in V^+$ such that $x_1x_2^2 \subset ux$ and $x_1x_2^2 \subset y\theta(u)$. Then we have following cases to consider.

Case 1: $|u| \leq |x_1|$. Then there exists $\alpha \in V^*$ such that $x_1 \subset u\alpha$.

Case 1.1 If $|\theta(u)| \leq |x_2|$. Then for some $\beta \in V^*$, $x_2 \subset \beta\theta(u)$ which implies $x_1x_2 \subset u\alpha\beta\theta(u)$, a contradiction.

Case 1.2 If $|x_2| < |\theta(u)| \leq |x_2^2|$. Then for $x_2 = \beta\beta_1$, $\beta_1 x_2 \subset \theta(u)$ where $\beta_1 \in V_\Diamond^+$ and $\beta \in V_\Diamond^*$ which implies $\beta_1\beta\beta_1 \subset \theta(u)$. Let $\theta(u) = \alpha_1\alpha_2$ such that $\beta_1 \subset \alpha_1$ and $x_2 \subset \alpha_2$. Thus, $x_1x_2 \subset u\alpha\beta\beta_1 \subset u\alpha\beta\alpha_1 = \theta(\alpha_1)\theta(\alpha_2)\alpha\beta\alpha_1$, a contradiction.

Case 1.3 $|\theta(u)| > |x_2^2|$. Then for $x_1 = \beta\beta_1$, $\beta_1 x_2^2 \subset \theta(u)$ where $\beta_1 \in V_\Diamond^+$ and $\beta \in V_\Diamond^*$. This implies $x_1x_2 \subset u\alpha x_2 \subset \theta(u_1)\theta(u_2)\theta(u_2)\alpha x_2$ where $\theta(u) = u_1u_2u_2$ such that $\beta_1 \subset u_1$ and $x_2 \subset u_2$. Also, $x_1x_2 = \beta\beta_1x_2 \subset \beta u_1u_2$, a contradiction.

Case 2: $|x_1| < |u| \leq |x_1x_2|$. Then there exists $\alpha \in V^*$ such that $x_1x_2 \subset u\alpha$.

Case 2.1 If $|\theta(u)| \leq |x_2|$. Then for some $\beta \in V^*$, $x_2 \subset \beta\theta(u)$. Similarly, $x_1x_2 \subset x_1\beta\theta(u)$ and $x_1x_2 \subset u\alpha$, a contradiction.

Case 2.2 $|x_2| < |\theta(u)| \leq |x_2x_2|$. Then for $x_2 = \beta\beta_1$, $\beta_1x_2 \subset \theta(u)$ where $\beta_1 \in V_\Diamond^+$ and $\beta \in V_\Diamond^*$. Let $\theta(u) = u_1u_2$ where $\beta_1 \subset u_1$, $x_2 \subset u_2$ for some $u_1 \in V^+$, $u_2 \in V^*$. Thus, $x_1x_2 \subset u\alpha = \theta(u_1)\theta(u_2)\alpha$. Similarly, $x_1x_2 = x_1\beta\beta_1 \subset x_1\beta u_1$, a contradiction.

Case 2.3 $|x_2x_2| < |\theta(u)| < |x_1x_2x_2|$. Then for $x_1 = \beta\beta_1$ and $\theta(u) = u_1u_2u_2$, $\beta_1 \subset u_1$ and $x_2 \subset u_2$ where $\beta \in V_\Diamond^*$, $\beta_1 \in V_\Diamond^+$ and $u_1, u_2 \in V^+$. Thus, $x_1x_2 \subset u\alpha = \theta(u_1)\theta(u_2)\theta(u_2)\alpha$. Similarly, $x_1x_2 = \beta\beta_1x_2 \subset \beta u_1u_2$, a contradiction.

Case 3: $|x_1x_2| < |u| < |x_1x_2x_2|$. Then for $u = u_1u_2u_3$ and $x_2 = \beta\beta_1$, $x_1 \subset u_1$, $x_2 \subset u_2$ and $\beta \subset u_3$ where $\beta, \beta_1 \in V_\Diamond^+$.

Case 3.1 $|x_2| < |\theta(u)| \leq |x_2x_2|$. Then for $x_2 = ss_1$ and $\theta(u) = u'u''$ such that $s_1 \subset u'$ and $x_2 \subset u''$ where $s \in V_\Diamond^*$, $s_1 \in V_\Diamond^+$ for some $u' \in V^+$, $u'' \in V^*$. Also, $u = u_1u_2u_3 = \theta(u')\theta(u'')$. Now, since $x_2 \subset u_2$ as well as $x_2 \subset u''$, we will get that $|u_2| = |\theta(u'')|$ which will further imply that $|u_1u_3| = |\theta(u')|$ and hence $\theta(u') = u_1r$, $u_2 = rp$ and $\theta(u'') = pu_3$ where $r, p \in V^+$. Thus, $x_1x_2 = x_1ss_1 \subset x_1su'$. Similarly, $x_1x_2u_3 \subset u_1u_2u_3 = u_1rpu_3 = \theta(u')pu_3$ which, by length argument further implies that $x_1x_2 \subset \theta(u')p$, a contradiction.

Case 3.2 $|x_2x_2| < |\theta(u)| < |x_1x_2x_2|$. Then for $\theta(u) = u_1'u_2'u_2'$ and $x_1 = ss_1$, $s_1 \subset u_1'$ and $x_2 \subset u_2'$ where $s, s_1 \in V_\Diamond^+$. Also, $u = u_1u_2u_3 = \theta(u_1')\theta(u_2')\theta(u_2')$. Now, since $x_2 \subset u_2$ as well as $x_2 \subset u_2'$, i.e., $\theta(x_2) \subset \theta(u_2')$, we will get that $|u_1| > |\theta(u_1')|$, $|u_2| = |\theta(u_2')|$ and $|u_3| < |\theta(u_2')|$ which implies $u_1 = \theta(u_1')y$ and $\theta(u_2') = y'u_3$ for $y, y' \in V^+$. Thus, $u = \theta(u_1')yu_2u_3 = \theta(u_1')y'u_3y'u_3$ which implies that $y = y'$. Hence, $x_1x_2 = x_1\beta\beta_1 \subset u_1u_3\beta_1 = \theta(u_1')yu_3\beta_1$. Similarly, $x_1x_2 = ss_1x_2 \subset su_1'u_2' = su_1'\theta(y)\theta(u_3)$, a contradiction.

Since all the cases lead to contradiction, we have $x_1x_2^2 \in D_{\theta_\diamond}(1)$.

□

We generalize the result of Proposition 5.25 and prove that if x_1x_2 is θ-unbordered then $x_1x_2^k$ for $k > 1$ is also θ-unbordered.

Lemma 5.26. *Let $x_1, x_2 \in V_\diamond^+$ and θ be either a morphic or an antimorphic involution. If x_1x_2 is a θ-unbordered partial word, then $x_1x_2^k$ for $k \geq 2$ is also a θ-unbordered partial word.*

Proof. Let θ be an antimorphic involution. Suppose, $x_1x_2^k$ for $k \geq 2$ is a θ-bordered partial word. Then by Lemma 5.17, there exists $a \in V$ and $y \in V^*$ such that $x_1x_2^k \subset ay\theta(a)$. Since x_1, x_2 are both nonempty partial words, we have $x_1x_2 \subset ax\theta(a)$ for $x \in V^*$ which is a contradiction. Hence $x_1x_2^k$ is θ-unbordered.

Now, let θ be a morphic involution. We will prove the result by induction on k.

Base Case: For $k = 2$, $x_1x_2^2$ is θ-unbordered by Proposition 5.25.

Inductive hypothesis: Let us assume that the result holds for $k = n$, that is, $x_1x_2^n$ is θ-unbordered.

Inductive step: Next we prove the claim for $k = n+1$. Suppose $x_1x_2^{n+1}$ is a θ-bordered partial word. Then we have $x_1x_2^{n+1} \subset ux$ and $x_1x_2^{n+1} \subset y\theta(u)$ where $u, x, y \in V^+$.

Case 1: $|x_2^{n+1}| < |\theta(u)| < |x_1x_2^{n+1}|$. Then $\alpha_1 x_2^{n+1} \subset \theta(u)$ for some $x_1 = \alpha\alpha_1$ where $\alpha, \alpha_1 \in V_\diamond^+$.

Case 1.1: $|x_1x_2^n| < |u| < |x_1x_2^{n+1}|$. Then $x_1x_2^n\beta \subset u$ for some $\beta \in V_\diamond^+$ and $x_2 = \beta\beta_1$ where $\beta_1 \in V_\diamond^+$. We have $\theta(\alpha_1)\theta(x_2^n)\theta(x_2) \subset u$ which implies that $\theta(\alpha_1)\theta(x_2^n)\theta(\beta)\theta(\beta_1) \subset u$. Also $x_1x_2^{n-1}x_2\beta = x_1x_2^{n-1}\beta\beta_1\beta \subset u$. Thus $\theta(\alpha_1)\theta(x_2^n)\theta(\beta)\theta(\beta_1) \uparrow x_1x_2^{n-1}\beta\beta_1\beta$. Considering the length argument, we have $\theta(\alpha_1)\theta(x_2^n) \uparrow x_1x_2^{n-1}\beta$. By augmenting β_1 both sides, $\theta(\alpha_1)\theta(x_2^n)\beta_1 \uparrow x_1x_2^{n-1}\beta\beta_1 = x_1x_2^n = \alpha\alpha_1x_2^n$. Hence $\theta(\alpha_1x_2^n)\beta_1 \uparrow \alpha\alpha_1x_2^n$ which implies that there exists a partial word w such that $\theta(\alpha_1x_2^n)\beta_1 \subset w$ and $\alpha\alpha_1x_2^n \subset w$. Let $\alpha_1x_2^n \subset v$ for some $v \in V^+$. Then $v \in \theta(\mathit{CPref}(x_1x_2^n)) \cap \mathit{CSuff}(x_1x_2^n)$ which shows that $x_1x_2^n$ is θ-bordered and hence a contradiction.

Case 1.2: $|u| \leq |x_1x_2^n|$. Then $x_1x_2^i\beta \subset u$ for some $x_2 = \beta\beta_1$ where $\beta \in V_\diamond^+$, $\beta_1 \in V_\diamond^*$ and $i < n$. Thus $\theta(\alpha_1)\theta(x_2^{n+1}) \subset u$ which implies that $x_1x_2^i\beta \uparrow \theta(\alpha_1)\theta(x_2^{n+1}) = \theta(\alpha_1)\theta(x_2^n)\theta(\beta)\theta(\beta_1)$. Now, we have $x_1x_2^{i-1}\beta\beta_1\beta \uparrow \theta(\alpha_1)\theta(x_2^n)\theta(\beta)\theta(\beta_1)$ and by length argument, $x_1x_2^{i-1}\beta \uparrow \theta(\alpha_1)\theta(x_2^n)$. By augmenting $\beta_1x_2^{n-i}$ on both the sides, we get $x_1x_2^{i-1}\beta\beta_1\, x_2^{n-i} = x_1x_2^n = \alpha\alpha_1x_2^n \uparrow \theta(\alpha_1)\theta(x_2^n)\beta_1x_2^{n-i} = \theta(\alpha_1x_2^n)\beta_1x_2^{n-i}$. Hence there exists a partial word w such that $\alpha\alpha_1x_2^n \subset w$ and $\theta(\alpha_1x_2^n)\beta_1x_2^{n-i} \subset w$. Let $\alpha_1x_2^n \subset v$ and thus $v \in \theta(\mathit{CPref}(x_1x_2^n)) \cap \mathit{CSuff}(x_1x_2^n)$ which implies that $x_1x_2^n$ is θ-bordered and hence a contradiction.

Case 2: If $|\theta(u)| \leq |x_2^{n+1}|$ then by similar argument we can obtain a contradiction.

Since all the cases lead to contradiction, we conclude that $x_1x_2^{n+1} \in D_{\theta_\diamond}(1)$.

Hence $x_1x_2^k$ is θ-unbordered partial word for all $k > 1$. □

The following corollary is a direct consequence of Lemma 5.26.

Corollary 5.27. *Let $x_1, x_2 \in V_\diamond^+$ and θ be either a morphic or an antimorphic involution. If x_1x_2 is a θ-unbordered partial word, then for any $k > 1$, $x_1^kx_2$ is also a θ-unbordered partial word.*

A language $L \subseteq V^*$ is called a post-plus language if for $u,\ v \in V^*$, $uv \in L$ implies that $uv^+ \subseteq L$ [92]. Similarly, we define post-plus language in partial words as follows:

Definition 5.28. *A language $L \subseteq V_\diamond^+$ of partial words is said to be a post-plus language if for two partial words $u, v \in V_\diamond^*$, $uv \in L$ implies that $uv^+ \subseteq L$.*

Corollary 5.29. *For (anti)morphic involution θ, $D_{\theta_\diamond}(1)$ is a post-plus language.*

Proof. The proof directly follows from Lemma 5.26. □

In Proposition 5.31, for a partial word u which is unbordered and θ-unbordered as well, we provide the number of θ-borders of a partial word which consists of alternating pattern of u and $\theta(u)$ for a morphic involution θ such that θ is not identity on V. The following result is required for the proof of Proposition 5.31.

Lemma 5.30. *Let θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in V$ and let $u \in D_{\theta_\diamond}(1) \cap D_\diamond(1)$. Then for $w = (u\theta(u))^n$, $u_1 <_{cp} u$ and $u \subset u'$ for some u' and u_1, we have $(u'\theta(u'))^j u_1, (u'\theta(u'))^j u'\theta(u_1) \notin L_d^{\theta_\diamond}(w)$ for all $1 \leq j < n$.*

Proof. We prove it by contradiction.

Let us assume that $(u'\theta(u'))^j u_1 \in L_d^{\theta_\diamond}(w)$. Then there exist $\alpha, \beta \in V^+$ such that $w \subset (u'\theta(u'))^j u_1\alpha$ and $w = (u\theta(u))^n \subset \beta(\theta(u')u')^j\theta(u_1)$. Since $|u_1| < |u|$, $\theta(u_1) <_{cs} \theta(u)$, that is, $u_1 <_{cs} u$ and also $u_1 <_{cp} u$. This implies $u \notin D_\diamond(1)$, a contradiction.

Now, let us assume that $(u'\theta(u'))^j u'\theta(u_1) \in L_d^{\theta_\diamond}(w)$. Then there exist $\alpha', \beta' \in V^+$ such that $w \subset (u'\theta(u'))^j u'\theta(u_1)\alpha'$ and $w = (u\theta(u))^n \subset \beta'(\theta(u')u')^j\theta(u')u_1$. Since $|u_1| < |u|$, $u_1 <_{cs} \theta(u)$, that is, $\theta(u_1) <_{cs} u$ and also $u_1 <_{cp} u$. This implies $u \notin D_{\theta_\diamond}(1)$, a contradiction. □

In the following proposition, we count the number of θ-borders of a partial word of the form $(u\theta(u))^i$ where we assume that u is a unbordered partial word as well as a θ-unbordered partial word.

Proposition 5.31. *Let θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in V$, and let $u \in D_\diamond(1) \cap D_{\theta_\diamond}(1)$. Then $w = (u\theta(u))^i \in D_{\theta_\diamond}(m)$ for all $i \geq 1$ where*

$$m = 1 + k^l + k^{3l} + \ldots + k^{(2i-1)l}$$

and k is the size of the alphabet and $l = |u|_\diamond$.

Proof. Let $l = |u|_\diamond$ be a nonnegative integer.
Let $i = 1$, that is, $w = u\theta(u)$. Then clearly $\varepsilon \in L_d^{\theta_\diamond}(w)$. Observe that any u_j such that $u \subset u_j$ will be a θ-border of w. The number of such u_j's depend on the number of holes in u. Hence $\nu_d^{\theta_\diamond}(w) \leq 1 + k^l$.

Let $x \in L_d^{\theta_\diamond}(w)$ be such that $x \neq \varepsilon$ and $x \neq u_j$ for all j such that $u \subset u_j$. Then there exist $\alpha, \beta \in V^+$ such that $w = u\theta(u) \subset x\alpha$ and $w = u\theta(u) \subset \beta\theta(x)$.

Let $|x| < |u|$. Then $x <_{cp} u$ and $\theta(x) <_{cs} \theta(u)$, i.e, $x <_{cs} u$ which implies $u \notin D_\diamond(1)$, a contradiction. Now, let $|x| > |u|$. Then $x = u'x_2$ and $\theta(x) = x_1'\theta(u')$ where $u \subset u'$, for some $u', x_2, x_1' \in V^+$. Thus $w = u\theta(u) \subset x\alpha = u'x_2\alpha$ which implies $x_2 <_{cp} \theta(u)$, that is, $\theta(x_2) <_{cp} u$ since $|u'| = |u|$. Also $w = u\theta(u) \subset \beta\theta(x) = \beta\theta(u')\theta(x_2)$ which implies $\theta(x_2) <_{cs} \theta(u)$, that is, $x_2 <_{cs} u$ since $|x_2| < |u|$. Thus $u \notin D_{\theta_\diamond}(1)$, a contradiction. Hence $u\theta(u) \in D_{\theta_\diamond}(1 + k^l)$.

Let $i = 2$, that is, $w = u\theta(u)u\theta(u)$. Observe that any words α, β such that $u \subset \alpha$ and $u\theta(u)u \subset \beta$ will be θ-borders of w. Also by Lemma 5.30, $u'\theta(u')u_1$, $u'\theta(u')u'\theta(u_1) \notin L_d^{\theta_\diamond}(w)$ where $u \subset u'$ for any u' and $u_1 <_{cp} u$. Thus $\nu_d^{\theta_\diamond}(w) = 1 + k^{|u|_\diamond} + k^{|u\theta(u)u|_\diamond}$, that is, $(u\theta(u))^2 \in D_{\theta_\diamond}(1 + k^l + k^{3l})$.

Now, let $i > 2$, then we have $w = (u\theta(u))^i$. When $i > 2$, any words $\alpha_1, \alpha_2, \ldots, \alpha_i$ such that $u \subset \alpha_1$, $u\theta(u)u \subset \alpha_2$, $(u\theta(u))^2u \subset \alpha_3, \ldots, (u\theta(u))^{i-1}u \subset \alpha_i$ will be θ-borders of w. By Lemma 5.30, w will not have any border of the form $\beta u'$ where $(u\theta(u))^j \subset \beta$ and $u' <_{cp} u$ with $j < i$. Hence the number of θ-borders of $w = (u\theta(u))^n$ depends on the number of holes present in u, $u\theta(u)u$, $(u\theta(u))^2u, \ldots, (u\theta(u))^{i-1}u$. Thus $\nu_d^{\theta_\diamond}(w) = 1 + k^l + k^{3l} + \ldots + k^{(2i-1)l}$.

Hence $w = (u\theta(u))^i \in D_{\theta_\diamond}(m)$ where

$$m = 1 + k^l + k^{3l} + \ldots + k^{(2i-1)l}$$

□

5.5 Disjunctivity of the Set of θ-unbordered Partial Words

For a language $L \subseteq V^*$, the *principal congruence* P_L determined by L is defined as follows: for any $x, y \in V^*$ such that $x \neq y$, $x \equiv y(P_L)$ if and only if $uxv \in L \Leftrightarrow uyv \in L$ for all $u, v \in V^*$. The index of P_L is the number of equivalence classes of P_L. The language L is said to be *disjunctive* if P_L is the identity, that is, for any $x \neq y$ there exist $u, v \in V^*$ such that $uxv \in L$ and $uyv \notin L$ or vice versa. Every disjunctive language is dense and every dense language contains a disjunctive subset, [85]. It is clear from the definition of disjunctive language that every disjunctive language has infinitely many principal congruence classes whereas a regular language has finitely many principal congruence classes. Hence no disjunctive language can be regular.

The following lemma provides a necessary and sufficient condition for a language to be disjunctive.

Lemma 5.32. *[82] Let $L \subseteq V^*$. Then the following two statements are equivalent:*

(a) L is a disjunctive language.

(b) If $u, v \in V^+$, $u \neq v$, $|u| = |v|$, then $u \not\equiv v(P_L)$.

Recall that, a partial word can be viewed as a total word over an extended alphabet $V_\diamond$, and hence the definition of principal congruence and disjunctive language can be extended for partial words in a similar way.

Lemma 5.33. *Let $L \subseteq V_\diamond^*$. Then the following two statements are equivalent:*

(a) L is a disjunctive language.

(b) If $u, v \in V_\diamond^+$, $u \not\uparrow v$, $|u| = |v|$, then $u \not\equiv v(P_L)$.

Proof. $(a) \Rightarrow (b)$ It is straightforward.
$(b) \Rightarrow (a)$ We prove that if $|u| = |v|$ and $u \equiv v(P_L)$ then $u \uparrow v$.

Let $x \equiv y(P_L)$. Assume that for $a, b \in V$, $a \neq b$ and let $M = max\{|xb|, |yb|\}$. Then it is clear that both xba^M and yba^M are primitive. Now $xw \equiv yw(P_L)$ by setting $w = ba^M$ which implies that $(xw)^2 \equiv xwyw \equiv ywxw(P_L)$. Since $|xwyw| = |ywxw|$ and $xwyw \equiv ywxw(P_L)$, we get $xwyw \uparrow ywxw$. Thus, xw and yw are commutative partial words. However, xw and yw are primitive, which implies that $xw \uparrow yw$ which shows that $x \uparrow y$ by the length argument. Hence L is a disjunctive language. □

It has been shown that the set of all (total)words with exactly i borders, $D(i)$ is disjunctive [52] and the set of all words with exactly i θ-borders, $D_\theta(i)$ is disjunctive for morphic involution θ [57] for all $i \geq 1$. Here we prove that the set of all θ-unbordered partial words is a disjunctive language for a morphic involution. In order to prove the disjunctivity of the language of partial words $D_{\theta_\diamond}(1)$, we need several auxiliary lemmas as follows.

Lemma 5.34. *Let θ be a morphic involution and $a, b \in V$ such that $a \neq b$ and $\theta(a) \neq a$. Let $x, y \in V_\diamond^m$ for $m > 0$. Then*

(a) $a^{m+1}x\theta(b) \in D_{\theta_\diamond}(1)$.

(b) If $x = \theta(b)x''$ for $x'' \in V_\diamond^*$ and $k > m$, then $(a^k y\theta(b))(a^k x\theta(b)) \in D_{\theta_\diamond}(1)$.

Proof.

(a) Let us assume that $a^{m+1}x\theta(b) \notin D_{\theta_\diamond}(1)$. Then there exist $u, \alpha, \beta \in V^+$ such that $a^{m+1}x\theta(b) \subset u\alpha$ and $a^{m+1}x\theta(b) \subset \beta\theta(u)$. Then we have following two cases to consider.

Case 1: $|u| \leq m+1$. Then $u = a^n$ and $\theta(u) = x_2'\theta(b)$ where $n \leq m+1$ and $x = x_1x_2$ with $x_2 \subset x_2'$ and $x_1, x_2 \in V_\diamond^+$ which implies $u = a^n = \theta(x_2')b$ which further implies that $a = b$, a contradiction.

Case 2: $m+1 < |u| \leq 2m+1$. Then $u = a^{m+1}x_1'$ and $\theta(u) = a^j x'\theta(b)$ where for $x = x_1x_2$, $x \subset x'$, $x_1 \subset x_1'$, $x_1 \in V_\Diamond^+$, $x_2 \in V_\Diamond^*$ and $j < m+1$. This implies $u = a^{m+1}x_1' = \theta(a^j)\theta(x')b$ which further implies that $\theta(a) = a$, a contradiction.

Since all the cases lead to contradiction $a^{m+1}x\theta(b) \in D_{\theta_\Diamond}(1)$.

(b) Let us assume that $(a^k y\theta(b))(a^k x\theta(b)) \notin D_{\theta_\Diamond}(1)$. Then there exist $u', \alpha', \beta' \in V^+$ such that $(a^k y\theta(b))(a^k x\theta(b)) \subset u'\alpha'$ and $(a^k y\theta(b))(a^k x\theta(b)) \subset \beta'\theta(u')$. Then we have following cases to consider.

Case 1: $|u'| \leq k$. Then $u' = a^n$ for $n \leq k$ which implies $(a^k y\theta(b))(a^k x\theta(b)) \subset \beta'\theta(u') = \beta'\theta(a^n)$ which further implies $\theta(a) = \theta(b)$, i.e., $a = b$, a contradiction.

Case 2: $k < |u'| < m+k+1$. Then $u' = a^k y_1'$ where $y = y_1y_2$, $y_1 \subset y_1'$ and $y_1, y_2 \in V_\Diamond^+$. Also, $\theta(u') = a^n x'\theta(b) = a^n\theta(b)x'''\theta(b)$ where for $x = \theta(b)x''$, $x'' \subset x'''$ and $n < k$. Thus $(a^k y\theta(b))(a^k x\theta(b)) \subset u'\alpha' = \theta(a^n)b\theta(x''')b\alpha'$ which due to the fact that $n < k$ implies $a = b$, a contradiction.

Case 3: $|u'| = m+k+1$. Then $u' = a^k y'\theta(b)$ and $\theta(u') = a^k x'\theta(b)$ where $y \subset y'$ and $x \subset x'$. Thus $u' = a^k y'\theta(b) = \theta(a^k)\theta(x')b$ which implies $\theta(a) = a$, a contradiction.

Case 4: $m+k+1 < |u'| \leq 2k+m+1$. Then $u' = a^k y'\theta(b)a^j$ where $y \subset y'$ and $j \leq k$. Also, $\theta(u') = y_2'\theta(b)a^k x'\theta(b)$ where for $y = y_1y_2$, $y_2 \subset y_2'$, $y_1 \in V_\Diamond^*$, $y_2 \in V_\Diamond^+$ and $x \subset x'$. Thus $u' = a^k y'\theta(b)a^j = \theta(y_2')b\theta(a^k)\theta(x')b$ which implies $a = b$, a contradiction.

Case 5: $2k+m+1 < |u'| \leq 2k+2m+1$. Then $u' = a^k y'\theta(b)a^k x_1'$ where for $x = x_1x_2$, $x_1 \subset x_1'$, $x_1 \in V_\Diamond^+$, $x_2 \in V_\Diamond^*$ and $y \subset y'$. Also, $\theta(u') = a^n y'\theta(b)a^k x'\theta(b)$ where $n < k$ and $x \subset x'$. Thus, $(a^k y\theta(b))(a^k x\theta(b)) \subset \beta'\theta(u') = \beta'\theta(a^k)\theta(y')b\theta(a^k)\theta(x_1')$. Now, since $|ba^k x_1'| \leq |a^k xb|$, $b\theta(a^k)$ $\theta(x_1') \leq_{cs} a^k x\theta(b)$ and hence $a = b$, a contradiction.

Since all the cases lead to contradiction, $(a^k y\theta(b))(a^k x\theta(b)) \in D_{\theta_\Diamond}(1)$.

□

Next we prove the following result under the assumption that the alphabet contains at least three distinct letters and θ is a morphic involution.

Lemma 5.35. *Let θ be a morphic involution and V be an alphabet with $|V| > 2$ and let $a, b \in V$ such that $\theta(a) \neq a$, $a \neq b$ and $a \neq \theta(b)$. Let $x \not\uparrow y$, $x, y \in V_\Diamond^m$, $m > 0$, $x, y \in \theta(b)V_\Diamond^*$. Then for all $i \geq 1$,*

$$a^{m+1}y\theta(b)(\theta(a^{m+1}x\theta(b))a^{m+1}x\theta(b))^{i-1}\theta(a^{m+1}x\theta(b)) \in D_{\theta_\Diamond}(1).$$

Proof. Let us assume that $w = a^{m+1}y\theta(b)(\theta(a^{m+1}x\theta(b))a^{m+1}x\theta(b))^{i-1}\ \theta(a^{m+1}x\ \theta(b)) \notin D_{\theta_\Diamond}(1)$. Then there exist $u, \alpha, \beta \in V^+$ such that $w \subset u\alpha$ and $w \subset \beta\theta(u)$. Let $w = w'\theta(a^{m+1}x\theta(b))$ where $w' = a^{m+1}y\theta(b)(\theta(a^{m+1}x\theta(b))a^{m+1}x\ \theta(b))^{i-1}$. Then we have following cases to consider.

Case 1: $|u| \leq m+1$. Then $u = a^k$ for $1 \leq k \leq m+1$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^k)$$

which implies $\theta(a) = b$, a contradiction.

Case 2: $m+1 < |u| \leq 2m+1$. Then $u = a^{m+1}y_1'$ where for $y = y_1y_2$, $y_1 \in V_\Diamond^+$, $y_2 \in V_\Diamond^*$ and $y_1 \subset y_1'$ for some $y_1' \in V^+$. Also, since $x = \theta(b)x'$ for $x' \in V_\Diamond^*$,

$$w = w'\theta(a^{m+1})b\theta(x')b \subset \beta\theta(a^{m+1})\theta(y_1').$$

Since $|a^{m+1}y_1'| < |a^{m+1}bx'b|$, $\theta(a^{m+1})\theta(y_1') <_{cs} \theta(a^{m+1})b\theta(x')b$ which implies $\theta(a^{m+1})b\theta(x')b \subset \beta'\theta(a^{m+1})\theta(y_1')$ where $\beta' \in V^+$ which further implies that $\theta(a) = b$, a contradiction.

Case 3: $|u| = 2m+2$. Then $u = a^{m+1}y'\theta(b)$ where $y \subset y'$ for some $y' \in V^+$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1})\theta(y')b$$

which by length argument implies $\theta(x) \subset \theta(y')$, i.e., $x \subset y'$. Also, $y \subset y'$ and hence $x \uparrow y$, a contradiction.

Case 4: $2m+2 < |u| \leq 3m+3$. Then $u = a^{m+1}y'\theta(b)\theta(a^k)$ for $k \leq m+1$ and $y \subset y'$ for some $y' \in V^+$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1})\theta(y')ba^k$$

which implies $a = b$, a contradiction.

Case 5: $3m+3 < |u| \leq 4m+3$. Then $u = a^{m+1}y'\theta(b)\theta(a^{m+1})\theta(x_1')$ where for $x = x_1x_2$, $x_1 \in V_\Diamond^+$, $x_2 \in V_\Diamond^*$, $x_1 \subset x_1'$ and $y \subset y'$ for some x_1', $y' \in V^+$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1})\theta(y')ba^{m+1}x_1'.$$

Now, since $|a^{m+1}x_1'| \geq |xb|$, $\theta(a^k)b\theta(x')b \subset a^{m+1}x_1'$ where $k < m+1$, $x = \theta(b)x'$ for $x' \in V_\Diamond^*$ which implies $a = b$, a contradiction.

Case 6: $|u| = 4m+4$. Then $u = a^{m+1}y'\theta(b)\theta(a^{m+1})\theta(x'')b$ where $x \subset x''$ and $y \subset y'$ for some x'', $y' \in V^+$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1})\theta(y')ba^{m+1}x''\theta(b)$$

which implies $\theta(a) = a$, a contradiction.

Case 7: $4m+4 < |u| \leq 5m+5$. Then $u = a^{m+1}y'\theta(b)\theta(a^{m+1}x''\theta(b))a^k$ where $k \leq m+1$, $x \subset x''$ and $y \subset y'$ for some x'', $y' \in V^+$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1}y'\theta(b))a^{m+1}x''\theta(b)\theta(a^k)$$

which implies $\theta(a) = b$, a contradiction.

Case 8: $5m+5 < |u| \leq 6m+5$. Then $u = a^{m+1}y'\theta(b)\theta(a^{m+1}x''\theta(b))a^{m+1}x'_1$ where for $x = x_1x_2$, $x_1 \subset x'_1$, $x_1 \in V_\Diamond^+$, $x_2 \in V_\Diamond^*$, $x \subset x''$ and $y \subset y'$ for some x'_1, x'', $y' \in V^+$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1}y'\theta(b))a^{m+1}x''\theta(b)\theta(a^{m+1})\theta(x'_1).$$

Now, since $|ba^{m+1}x'_1| \leq |a^{m+1}xb|$, $\theta(b)\theta(a^{m+1})\theta(x'_1) \leq_{cs} \theta(a^{m+1}x\theta(b))$. This implies $\theta(a^{m+1}x\theta(b)) = \theta(a^{m+1})b\theta(x')b \subset \beta'\theta(b)\theta(a^{m+1})\theta(x'_1)$ where $\beta' \in V^*$ and $x' \in V_\Diamond^*$ which further implies $\theta(a) = \theta(b)$, i.e., $a = b$, a contradiction.

Case 9: $|u| = (2m+2)(2i-1)$. Then $u = a^{m+1}y'\theta(b)(\theta(a^{m+1}x'\ \theta(b))a^{m+1}x'\ \theta(b))^{i-1}$ where $i \geq 2$, $x \subset x'$ and $y \subset y'$ for some x', $y' \in V^+$. Thus,

$$\begin{aligned} w &= a^{m+1}y\theta(b)(\theta(a^{m+1}x\theta(b))a^{m+1}x\theta(b))^{i-1}\theta(a^{m+1}x\theta(b)) \\ &= a^{m+1}y\theta(b)\theta(a^{m+1}x\theta(b))(a^{m+1}x\theta(b)\theta(a^{m+1}x\theta(b)))^{i-1} \\ &\subset \beta\theta(a^{m+1}y'\theta(b))(a^{m+1}x'\theta(b)\theta(a^{m+1}x'\theta(b)))^{i-1} \end{aligned}$$

which implies $\theta(x) \subset \theta(y')$, i.e., $x \subset y'$ and also $y \subset y'$ which further implies $x \uparrow y$, a contradiction.

Case 10: $(2m+2)(2i-1) < |u| \leq (2m+1)(2i-1)+(m+1)$. Then $u = a^{m+1}y'\theta(b)(\theta(a^{m+1}x'\theta(b))a^{m+1}x'\theta(b))^{i-1}\theta(a^k)$ where $x \subset x'$, $y \subset y'$ for some x', $y' \in V^+$ and $k \leq m+1$. Thus,

$$w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1}y'\theta(b))(a^{m+1}x'\theta(b)\theta(a^{m+1}x'\theta(b)))^{i-1}a^k$$

which implies $a = b$, a contradiction.

Case 11: $(2m+2)(2i-1)+(m+1) < |u| \leq (2m+1)(2i-1)+(2m+1)$. Then $u = a^{m+1}y'\theta(b)(\theta(a^{m+1}x'\theta(b))a^{m+1}x'\theta(b))^{i-1}\theta(a^{m+1})\theta(x'_1)$ where for $x = x_1x_2$, $x_1 \subset x'_1$, $x_1 \in V_\Diamond^+$ and $x_2 \in V_\Diamond^*$ for some y', $x'_1 \in V^+$. Thus, $w = w'\theta(a^{m+1}x\theta(b)) \subset \beta\theta(a^{m+1}y'\theta(b))(a^{m+1}x'\theta(b)\theta(a^{m+1}x'\theta(b)))^{i-1}a^{m+1}x'_1$. Now, since $|a^{m+1}x'_1| < |a^{m+1}xb|$, $a^{m+1}x'_1 <_{cs} \theta(a^{m+1}x\theta(b))$, i.e., $\theta(a^{m+1}x\ \theta(b)) \subset \beta'a^{m+1}x'_1$ where $\beta' \in V^+$ which implies $a = \theta(a)$, a contradiction.

Since all the cases lead to contradiction,

$$a^m y\theta(b)(\theta(a^m x\theta(b))a^m x\theta(b))^{i-1}\theta(a^m x\theta(b)) \in D_{\theta_\Diamond}(1).$$

□

In Theorem 5.36, we prove that the language of θ-unbordered partial words is a disjunctive language which make use of the results Lemma 5.33, Lemma 5.34 and Lemma 5.35.

Theorem 5.36. *Let θ be a morphic involution and let V be an alphabet with $|V| > 2$ such that for $a, b \in V$, $a \neq b$, $a \neq \theta(b)$ and $\theta(a) \neq a$. Then the set of all θ-unbordered partial words $D_{\theta_\diamond}(1)$ is a disjunctive language.*

Proof. Let $i \geq 1$ and $x, y \in V_\diamond^n$, $x \not\uparrow y$, $m' = n + 2$, $n > 0$. Let $u = a^{m'}\theta(b)$ and

$$v = \theta(b)(\theta(a^{m'}\theta(b)x\theta(b))a^{m'}\theta(b)x\theta(b))^{i-1}\theta(a^{m'}\theta(b)x\theta(b)).$$

Since $a \notin \{b, \theta(a)\}$, by Lemma 5.34 $a^{m'}\theta(b)x\theta(b) \in D_{\theta_\diamond}(1)$ and by Proposition 5.31,

$$uxv = [a^{m'}\theta(b)x\theta(b)\theta(a^{m'}\theta(b)x\theta(b))]^i \in D_{\theta_\diamond}(m).$$

where $m(= 1 + k^l + k^{3l} + \cdots + k^{(2i-1)l}) > i \geq 1$ for $u' = a^{m'}\theta(b)x\theta(b)$, $l = |u|_\diamond$ and $k = |V| > 2$. Also, by Lemma 5.35

$$uyv = a^{m'}\theta(b)y\theta(b)(\theta(a^{m'}\theta(b)x\theta(b))a^{m'}\theta(b)x\theta(b))^{i-1}\theta(a^{m'}\theta(b)x\theta(b)) \in D_{\theta_\diamond}(1).$$

Thus for $x, y \in V_\diamond^*$ and $x \not\uparrow y$, we have $uxv \notin D_{\theta_\diamond}(1)$ and $uyv \in D_{\theta_\diamond}(1)$.

Therefore $x \not\equiv y(P_{D_{\theta_\diamond}(1)})$ for every $x \not\uparrow y$, $|x| = |y|$. Hence, by Lemma 5.33, $D_{\theta_\diamond}(1)$ is a disjunctive language. □

5.6 Pseudo-Bordered Partial Words and Language Classes

In this section, we study the languages of θ-unbordered and θ-bordered partial words with respect to the conventional language classes in Chomsky hierarchy. We prove that the set of all θ-unbordered partial words is regular when θ is an antimorphic involution, and the set of all θ-bordered partial words is not a context-free language when θ is a morphic involution. We introduce the concept of $D_{\theta_\diamond}(1)$-concatenate partial word which is an extension of the concepts defined in [52, 57]. We also extend the concept of θ-non-overlapped language, [57], for partial words and provide a necessary and sufficient condition for a language of partial words to be θ-non-overlapped when θ is assumed to be a morphic involution (Proposition 5.41).

Theorem 5.37. *$D_{\theta_\diamond}(1)$ is a regular language when θ is an antimorphic involution on $V_\diamond^*$.*

Proof. From Lemma 5.17, all the partial words of the form axb where $a, b \in V$, $x \in V_\diamond^*$, $b \neq \theta(a)$ are θ-unbordered. Also by Lemma 5.9, $a \in V_\diamond$ is also θ-unbordered. Hence $D_{\theta_\diamond}(1) = V_\diamond \cup L$ where $L = \bigcup_{a,\ b \in V} aV_\diamond^*b$, $b \neq \theta(a)$. Since V is a finite alphabet, L is regular. Hence $D_{\theta_\diamond}(1)$ is a regular language. □

Next we prove that the language of θ-bordered partial words is not a context-free language for a morphic involution θ.

Theorem 5.38. *Let θ be a morphic involution such that $\theta(a) \neq a$ for all $a \in V$. Then the language of θ-bordered partial words over V is not a context-free language.*

Proof. Since $\theta(a) \neq a$ for all $a \in V$, there exists $b \in V$ such that $\theta(a) = b$. Also, $\theta(\Diamond) = \Diamond$. Let L be a language of θ-bordered partial words.

Let us assume that L is context-free. Let $n > 0$ be the constant defined by pumping lemma for context-free languages. Since L is context-free, it must satisfy the conditions of the pumping lemma. Thus, for every partial word $z \in L$ with $|z| \geq n$, z can be factorized into $uvwxy$ such that $|vwx| \leq n$, $|vx| > 0$ and $uv^iwx^iy \in L$ for all $i \geq 0$.

Consider a partial word $z_1 = a^{n+1}\Diamond^{n+1}a^{n+1}$. Observe that z_1 is θ-bordered since $z_1 \subset (a^{n+1}b^{n+1})a^{n+1}$ and $z_1 \subset a^{n+1}\theta(a^{n+1}b^{n+1}) = a^{n+1}b^{n+1}a^{n+1}$. There are two cases depending upon whether vwx is a substring of $a^{n+1}\Diamond^{n+1}$ or $\Diamond^{n+1}a^{n+1}$.

Case 1: vwx is a substring of $a^{n+1}\Diamond^{n+1}$. Note that every θ-border w of z_i is of the form $w = ap$ for some $p \in V^*$ since z_i begins with a. We consider the following subcases depending on whether vwx contains $\Diamond$ or not.

Case 1.1 Suppose vwx does not contain any $\Diamond$, that is, vwx is a substring of a^{n+1} of z_1. Then for $i \geq 2$, we get $z_i = a^m\Diamond^{n+1}a^{n+1}$ where $m > n+1$. Hence for $m' > n+1$, $z_2 = uv^2wx^2y = a^{m'}\Diamond^{n+1}a^{n+1} \notin L$, a contradiction.

Case 1.2 Suppose vwx contains at least one $\Diamond$, that is, either v or x must include at least one $\Diamond$. Consider $z_0 = uwy = a^k\Diamond^j a^{n+1}$ where $k \leq n+1$, $j < n+1$. Since $j < n+1$, $a^k t$ with $|t| = j$ cannot be a θ-border of z_0 and hence $z_0 \notin L$, a contradiction.

Case 2: vwx is a substring of $\Diamond^{n+1}a^{n+1}$. The proof for this case is similar to that of Case 1 and leads to a similar contradiction.

Since all the cases lead to contradiction, the language of θ-bordered partial words L over V is not a context-free language. □

Let $V = \{a, b, c, d\}$ and θ be (anti)morphic involution such that $\theta(a) = b$, $\theta(b) = a$, $\theta(c) = d$, $\theta(d) = c$ and $\theta(\Diamond) = \Diamond$. Let $u = ad\Diamond d$. Clearly, $u \in D_{\theta_\Diamond}(1)$. Let us consider the decomposition of $u = u_1u_2$ such that $u_1 = ad$ and $u_2 = \Diamond d$. Then $u_1 \in D_{\theta_\Diamond}(1)$ but $u_2 = \Diamond d \notin D_{\theta_\Diamond}(1)$ as $\Diamond d \subset cd$. Hence, it is not necessary that any θ-unbordered partial word can be decomposed as a concatenation of two θ-unbordered partial words. However, if such a decomposition exists for a θ-unbordered partial word u then u is said to be $D_{\theta_\Diamond}(1)$-*concatenate* partial word. Similarly, a partial word u is said to be *completely* $D_{\theta_\Diamond}(1)$-*concatenate* if $u = u_1u_2$ for $u_1, u_2 \in V_\Diamond^+$, implies that $u_1, u_2 \in D_{\theta_\Diamond}(1)$.

In case of total words, the set of completely $D_\theta(1)$-concatenate words over an alphabet V is a regular language, [57]. However, no partial word $u \in V_\Diamond^+$ such that $|u| > 2$ and contains one or more holes can be completely $D_{\theta_\Diamond}(1)$-concatenate.

Proposition 5.39. *For (anti)morphic involution θ there does not exist any completely $D_{\theta_\Diamond}(1)$-concatenate partial word u such that $|u| > 2$ and u contains at least one hole.*

Proof. Let $u \in V_\Diamond^+$ be a partial word with at least one hole. Then there will always exist a decomposition $u = u_1u_2$ such that $u_1 = ay\Diamond \notin D_{\theta_\Diamond}(1)$ (or $u_1 = \Diamond ya$) for $y \in V_\Diamond^*$ and $u_2 \in V_\Diamond^+$. Hence u can never be a completely $D_{\theta_\Diamond}(1)$-concatenate partial word. □

In Theorem 5.23, we have given a necessary and sufficient condition for concatenation of two θ-unbordered partial word to be θ-unbordered. However, this result may not hold for catenation of partial words which are not θ-unbordered. Hence we extend the concept of *θ-non-overlapped language* for partial words, which is a generalization of concepts defined in [90] and [57].

Definition 5.40. *A pair of partial words $u, v \in V_\Diamond^+$ such that $u \not\uparrow v$ is said to be θ-non-overlapped if and only if $\theta(CPref(u)) \cap CSuff(v) = \emptyset$ and $\theta(CPref(v)) \cap CSuff(u) = \emptyset$. A language of partial words L is said to be θ-non-overlapped language if $L \subseteq D_{\theta_\Diamond}(1)$, $u, v \in L$ and $\theta(u) \not\uparrow v$ implies u, v are θ-non-overlapped.*

For a language L of partial words, by $L_\theta^{(2)}$, we denote a language $L_\theta^{(2)} = \{u\theta(u) | u \in L\}$.

Proposition 5.41. *Let $L \subseteq V_\Diamond^+$ and θ be a morphic involution. Then L is θ-non-overlapped language if and only if $L \subseteq D_{\theta_\Diamond}(1)$ and $L^2 \backslash L_\theta^{(2)} \subseteq D_{\theta_\Diamond}(1)$.*

Proof. Let $L \subseteq V_\Diamond^+$. Assume that L is θ-non-overlapped language. Then $L \subseteq D_{\theta_\Diamond}(1)$. Let $u, v \in L$ be such that $\theta(u) \not\uparrow v$, then $uv \in L^2 \backslash L_\theta^{(2)}$. Suppose $uv \notin D_{\theta_\Diamond}(1)$. Then there exist $w, x, y \in V^+$ such that $uv \subset wx$ and $uv \subset y\theta(w)$. Let $|w| > |u|$. Then for $w = w_1w_2$, let $v = v_1v_2$ be such that $u \subset w_1$ and $v_1 \subset w_2$ where $w_1, w_2 \in V^+$ and $v_1, v_2 \in V_\Diamond^+$. Thus $uv \subset wx = w_1w_2x$ and $uv \subset y\theta(w) = y\theta(w_1)\theta(w_2)$. Since $u \subset w_1$, we get $v \subset w_2x$ and $v \subset \alpha\theta(w_2)$ for $\alpha \in V^+$. Thus $v \notin D_{\theta_\Diamond}(1)$, a contradiction since $L \subseteq D_{\theta_\Diamond}(1)$. We will reach a similar contradiction if we assume that $|w| > |v|$. If $|w| \leq |u|$ and $|w| \leq |v|$, then $w \in \text{CPref}(u)$ and $w \in \theta(\text{CSuff}(v))$, a contradiction since L is a θ-non-overlapped language. Hence $uv \in D_{\theta_\Diamond}(1)$ and $L^2 \backslash L_\theta^{(2)} \subseteq D_{\theta_\Diamond}(1)$.

Conversely, assume that $L \subseteq D_{\theta_\Diamond}(1)$ and $L^2 \backslash L_\theta^{(2)} \subseteq D_{\theta_\Diamond}(1)$. Let $u, v \in L$ such that $\theta(u) \not\uparrow v$, that is, $uv \in L^2 \backslash L_\theta^{(2)}$. Suppose u, v are not θ-non-overlapped. Then either $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) \neq \emptyset$ or $\theta(\text{CPref}(v)) \cap \text{CSuff}(v) \neq \emptyset$. Let $\alpha \in \theta(\text{CPref}(u)) \cap \text{CSuff}(v)$. Thus there exist $u_1, u_2, v_1, v_2 \in V_\Diamond^+$ such that $u = u_1u_2$, $v = v_1v_2$, $u_1 \subset \theta(\alpha)$ and $v_2 \subset \alpha$. Thus $uv = u_1u_2v_1v_2 \subset \theta(\alpha)u_2v_1\alpha$ which implies $uv \notin D_{\theta_\Diamond}(1)$, a contradiction. We will reach a similar contradiction if we assume that $\theta(\text{CPref}(v)) \cap \text{CSuff}(u) \neq \emptyset$. Thus, $\theta(\text{CPref}(u)) \cap \text{CSuff}(v) = \emptyset$ and $\theta(\text{CPref}(v)) \cap \text{CSuff}(u) = \emptyset$ for every $u, v \in L$, $\theta(u) \not\uparrow v$ and $L \subseteq D_{\theta_\Diamond}(1)$, that is, L is θ-non-overlapped language. □

Recall the concept of θ-unbounded annihilator $\alpha_{ub}(u) = \{v \in V^+ | uv \in D_\theta(1)\}$ defined in [53] which is a generalization of unbounded annihilator defined in [93]. We extend the concept of θ-unbounded annihilator for a nonempty partial word u as follows;

$$\alpha_{ub_\diamond}(u) = \{v \in V_\diamond^+ | uv \in D_{\theta_\diamond}(1)\}.$$

In our next result, we prove that for a θ-unbordered partial word u, if we raise the power of the set of proper contained suffixes of u then it is a subset of θ-unbounded annihilator of u.

Proposition 5.42. *Let θ be (anti)moprhic involution on $V_\diamond^*$. If $u \in D_{\theta_\diamond}(1)$, then $(PCSuff(u))^+ \subseteq \alpha_{ub_\diamond}(u)$ where PCSuff(u) is the set of all proper contained suffixes of u.*

Proof. Let $u \in D_{\theta_\diamond}(1)$ and let $v = u_1u_2\cdots u_k$ for some $u_i \in \text{PCSuff}(u)$ and $1 \leq i \leq k$.

Let θ be a morphic involution. Suppose that $uv \notin D_{\theta_\diamond}(1)$. Then there exist $\alpha, \alpha', \beta' \in V^+$ such that $uv \subset \alpha\alpha'$ and $uv \subset \beta'\theta(\alpha)$. Then we consider following two cases:

Case 1: $|\alpha| > |v|$. Let $u \subset u'u''$ and $v = v'$ such that $\theta(\alpha) = u''v'$ and $u', u'' \in V^+$ which implies $u'' \in \text{CSuff}(u)$. Also, $uv \subset \alpha\alpha' = \theta(u'')\theta(v')\alpha'$ which implies $u'' \in \theta(\text{CPref}(u))$ which further implies that $u \notin D_{\theta_\diamond}(1)$, a contradiction.

Case 2: $|\alpha| \leq |v|$. We also have $v = u_1u_2\cdots u_k$ for $u_i \in \text{PCSuff}(u)$ where $1 \leq i \leq k$. Thus we have following two subcases:

Case 2.1 If $|\alpha| < |u_k|$. Then $u_k = u_{k'}u_{k''}$ and $\theta(\alpha) = u_{k''}$ for $u_{k'}, u_{k''} \in V^+$. Since $u_k \in \text{PCSuff}(u)$, for $u = u_1u_2$, $u_2 \subset u_k = u_{k'}u_{k''}$ and thus $u \subset u_1u_{k'}u_{k''}$ which implies $u_{k''} \in \text{CSuff}(u)$ where $u_1 \in V_\diamond^+$. Also, $uv \subset \alpha\alpha' = \theta(u_{k''})\alpha'$ implies that $u_{k''} \in \theta(\text{CPref}(u))$ and hence $u \notin D_{\theta_\diamond}(1)$, a contradiction.

Case 2.2 $|\alpha| \geq |u_k|$. Then $\theta(\alpha) = u_j''u_{j+1}\cdots u_k$ such that $u_j = u_j'u_j''$ where $u_j', u_j'' \in V^+$. Since $u_j \in \text{PCSuff}(u)$ and $u_j'' \in \text{CSuff}(u_j)$, we have $u_j'' \in \text{CSuff}(u)$. Also, $uv \subset \alpha\alpha' = \theta(u_j'')\theta(u_{j+1})\cdots\theta(u_k)\alpha'$ implies $u_j'' \in \theta(\text{CPref}(u))$ which further implies that $u \notin D_{\theta_\diamond}(1)$, a contradiction.

Since all the cases lead to contradiction, we have $(\text{PCSuff}(u))^+ \subseteq \alpha_{ub_\diamond}(u)$.

Now, let θ be an antimorphic involution. Assume that $uv \notin D_{\theta_\diamond}(1)$. Then by Lemma 5.17, $uv \subset ay\theta(a)$ for $y \in V^*$ and $a \in V$. Let $y = y_1y_2$ such that $u \subset ay_1$ and $v = y_2\theta(a)$ for $y_1, y_2 \in V^*$. This implies $a \in \text{CPref}(u)$. Also, $v = u_1u_2\cdots u_k = y_2\theta(a)$ implies that $\theta(a) \in \text{CSuff}(u_k)$ which along with the fact that $u_k \in \text{CSuff}(u)$ further implies that $a \in \theta(\text{CSuff}(u))$ and hence $u \notin D_{\theta_\diamond}(1)$, a contradiction. Thus $(\text{PCSuff}(u))^+ \subseteq \alpha_{ub_\diamond}(u)$. $\square$

Next we define θ-unbounded annihilator for a language of partial words in the following definition.

Definition 5.43. *Let $L \subseteq V_\Diamond^+$. The θ-unbounded annihilator of the language L is defined as;*

$$\alpha_{ub_\Diamond}(L) = \{u \in V_\Diamond^+ \mid Lu \subseteq D_{\theta_\Diamond}(1)\}$$

Proposition 5.44. *Let $L \subseteq V_\Diamond^+$ such that for every $u \in V_\Diamond^+$ contains at least one non-hole position and $V \subseteq \theta(CPref(L))$, then $\alpha_{ub_\Diamond}(L) = \emptyset$.*

Proof. For any $u \in V_\Diamond^+$, let $u \subset u'a$ for $u' \in V^*$ and $a \in V$. Since $V \subseteq \theta(\text{CPref}(L))$, there exists $v \in L$ such that $v \subset \theta(a)v'$ for $v' \in V^*$. Thus $vu \subset \theta(a)v'u'a \notin D_{\theta_\Diamond}(1)$ which implies $Lu \nsubseteq D_{\theta_\Diamond}(1)$ which further implies $\alpha_{ub_\Diamond}(L) = \emptyset$. □

Next we prove that for a language of partial words $L \subseteq V_\Diamond^+$, $\alpha_{ub_\Diamond}(L)$ is a post-plus language.

Theorem 5.45. *For any $L \subseteq V_\Diamond^+$, $\alpha_{ub_\Diamond}(L)$ is a post-plus language.*

Proof. Let $u, v \in V_\Diamond^*$ be such that $uv \in \alpha_{ub_\Diamond}(L)$. Then for all $w \in L$, we get $wuv \in D_{\theta_\Diamond}(1)$. Furthermore, by Proposition 5.26 $wuv^+ \subseteq D_{\theta_\Diamond}(1)$ which implies $Luv^+ \subseteq D_{\theta_\Diamond}(1)$ which further implies $uv^+ \subseteq \alpha_{ub_\Diamond}(L)$. Thus $\alpha_{ub_\Diamond}(L)$ is a post-plus language. □

5.7 θ-Primitive Partial Words

In this section, we extend the concept of primitivity to accommodate partial words under involution mappings. We first extend the notion of a period to θ-strong and θ-weak period of a partial word where θ is (anti)morphic involution.

Definition 5.46. *Let θ be (anti) morphic involution and u be a partial word over V.*

(a) A θ-(strong) period of a partial word u is a positive integer p such that $u(i) = u(j)$ or $u(i) = \theta(u(j))$ whenever $i, j \in D(u)$ and $i \equiv j \bmod p$. In such a case, we call u to be θ-p-periodic.

(b) A θ-weak period of u is a positive integer p such that $u(i) = u(i+p)$ or $u(i) = \theta(u(i+p))$ whenever $i, i+p \in D(u)$. In such a case, we call u to be weakly θ-p-periodic.

If a partial word is θ-p-periodic then it is weakly θ-p-periodic as well, but converse does not hold always. Let us illustrate the above definitions with the help of a following example.

Example 5.47. Let $V = \{A, C, G, T\}$ and let θ be (anti)morphic involution such that $\theta(A) = T$, $\theta(T) = A$, $\theta(G) = C$, $\theta(C) = G$ and $\theta(\Diamond) = \Diamond$. Let $u = A\Diamond\Diamond TC\Diamond TGA$, then u is θ-3,5,6,8-periodic. Similarly, $v = A\Diamond\Diamond TG\Diamond\Diamond CTG$ is weakly θ-2,3,5,8-periodic. However v is not θ-2-periodic. □

Next we define the concept of θ-power also known as pseudo-power for a partial word which is required to define θ-primitive partial word.

Definition 5.48. *Let θ be (anti)morphic involution. A partial word $w = u_1 u_2 \cdots u_n$ is said to be contained in θ-power of a word u if $u_1 \subset u$ and, either $u_i \subset u$ or $u_i \subset \theta(u)$ for all $i \in \{2, \ldots, n\}$. More specifically, a partial word $w \in V_\Diamond^+$ is said to be contained in pseudo-power of a nonempty word u relative to θ, if $w \subset u\{u, \theta(u)\}^*$.*

Definition 5.49 (θ-primitive partial word)**.** *A partial word w is said to be θ-primitive if there exists no nonempty word u such that $w \subset u\{u, \theta(u)\}^+$.*

We use the symbol Q_{p_θ} to denote the set of all θ-primitive partial words. If $w \subset u\{u, \theta(u)\}^*$ and u is a shortest such word, then u is said to be a θ-primitive root of w. Formally,

$$\rho_{\theta,\Diamond}(w) = \{x : x \text{ is a } \theta\text{-primitive (total) word and } w \subset x\{x, \theta(x)\}^n,\ n \geq 0\}.$$

It is known that every nonempty word has an unique primitive root, and also an unique θ-primitive root for a morphic as well as an antimorphic involution θ, [29]. However, in case of partial words, the primitive root of a partial word need not be unique, [8]. Similarly, the θ-primitive root of a partial word need not be unique as shown in the following example.

Example 5.50. Let $V = \{A, T, C, G\}$ and θ be an antimorphic involution such that $\theta(A) = T, \theta(T) = A, \theta(C) = G, \theta(G) = C, \theta(\Diamond) = \Diamond$. Let $w = A\Diamond CG\Diamond T$. It can be observed that $w \subset ATC\theta(ATC)$, $w \subset ACC\theta(ACC)$, $w \subset AGC\theta(AGC)$ and $w \subset AAC\theta(AAC)$. □

Proposition 5.51. *If a partial word $w \in V_\Diamond^+$ is θ-primitive then it is also primitive.*

Proof. Let w be a θ-primitive partial word. Suppose w is not primitive. Then there exists a word $u \in V^+$ such that $w \subset u^k$ with $k \geq 2$. Hence w is also contained in θ-power of u which is a contradiction. Thus, w is a primitive partial word. □

The converse of the above proposition need not be true as illustrated by the following example.

Example 5.52. Let $V = \{a, b\}$ and θ be an antimorphic involution such that $\theta(a) = b, \theta(b) = a$ and $\theta(\Diamond) = \Diamond$. Let $w = abb\Diamond \in Q_p$ but $w = abb\Diamond \subset a\theta(a)\theta(a)a$ and hence w is not a θ-primitive partial word. □

Since θ-primitive partial word is primitive, the Lemma 3.7 applies to θ-primitive partial words as well. However, a θ-primitive partial word x with one hole can be compatible to a central factor of a word in $\{x, \theta(x)\}^2 \backslash x^2$, as demonstrated by the following example.

Example 5.53. Let $V = \{A, C, G, T\}$ and θ be an antimorphic involution such that $\theta(A) = T, \theta(T) = A, \theta(G) = C, \theta(C) = G$ and $\theta(\Diamond) = \Diamond$. Then for $x = A\Diamond CG$, we can see that $\theta(x)\theta(x) = CG\Diamond TCG\Diamond T \uparrow CG\mathbf{A}\Diamond\mathbf{CG}AT$. Similarly, let $V = \{a, b, c\}$ and θ be an antimorphic involution such that $\theta(a) = b$ and vice versa, and $\theta(c) = c$. Then for $x = a\Diamond c$, $x\theta(x) = a\Diamond cc\Diamond b \uparrow a\mathbf{a}\Diamond\mathbf{c}\Diamond b$. □

The language of primitive partial words is closed under cyclic permutation. However, the class of θ-primitive partial words not necessarily closed under cyclic permutation which as shown in the following example.

Example 5.54. Consider the DNA alphabet $V = \{A, T, C, G\}$ and let θ be a morphic involution where $\theta(A) = T, \theta(T) = A, \theta(C) = G, \theta(G) = C$ and $\theta(\Diamond) = \Diamond$. Let $w = AC\Diamond GTC$ be a primitive partial word. Consider a cyclic permutation of w, $w' = CAC\Diamond GT \subset (CA)^2\theta(CA)$ which is not a θ-primitive partial word. □

5.8 θ-Conjugacy and θ-Commutativity

In this section, we define the concept of θ-conjugacy and θ-commutativity for partial words where θ is (anti)morphic involution.

Definition 5.55. *Let θ be either a morphic or an antimorphic involution on $V_\Diamond^*$. A partial word u is a θ-conjugate of another partial word w if there exists $v \in V_\Diamond^*$ such that $uv \uparrow \theta(v)w$.*

It is known that θ-conjugacy is a transitive relation for total words when θ is a morphic involution, [60]. However, this may not hold in case of partial words as demonstrated by the following example.

Example 5.56. Let $V = \{a, b, c\}$ and let θ be a morphic involution such that $\theta(a) = b, \theta(b) = a, \theta(c) = c$ and $\theta(\Diamond) = \Diamond$. Let $u = a\Diamond bc$, $w = \Diamond acb$ and $v = b\Diamond ab$. It is easy to see that for $x = acb$ and $y = c$, $u \subset xy$, $w \subset \theta(y)x$ and hence u is a θ-conjugate of w. Similarly, for $\alpha = b$ and $\beta = acb$, $w \subset \alpha\beta$ and $v \subset \theta(\beta)\alpha$ and hence w is a θ-conjugate of v. Now, if we assume that there exist $x', y' \in V_\Diamond^+$ such that $u \subset x'y'$ and $v \subset \theta(y')x'$, then it is easy to observe that x' must begin with a and $\theta(y')$ must begin with b, that is, y' must begin with a. Hence $x' = a$ and $y' = abc$, but then $v = b\Diamond ab \not\subset \theta(y')x' = baca$. Hence θ-conjugacy relation is not transitive for a morphic involution θ. □

The result from Theorem 2.34 provides a characterization of partial words such that x is a conjugate of y. It has been illustrated in the Example 2.35 which is required for a better understanding of Theorem 5.57.

In the following theorem, we provide a characterization of partial words x and y such that x is a θ-conjugate of y.

Theorem 5.57. *Let x and y be nonempty partial words and θ be a morphic involution on $V_\Diamond^*$. If there exists a partial word z such that $xz \uparrow \theta(z)y$ and $xz \vee \theta(z)y$ is $|x|$-θ-periodic, then there exist partial words u, v such that $x \subset uv$ and one of the following holds:*

(a) $y \subset v\theta(u)$ *and* $z \subset (\theta(u)\theta(v)uv)^i\theta(u)$ *for some* $i \geq 0$.

(b) $y \subset \theta(v)u$ *and* $z \subset (\theta(u)\theta(v)uv)^i\theta(u)\theta(v)u$ *for some* $i \geq 0$.

Proof. Let $m > 0$ be such that $m|x| > |z| > (m-1)|x|$. Let $x = x_1y_1$ and $y = y_2x_2$ where $|x_1| = |x_2| = |z| - (m-1)|x|$ and $|y_1| = |y_2|$. Let $z = x'_1y'_1x'_2y'_2 \cdots x'_{m-1}y'_{m-1}x'_m$ where $|x'_1| = |x'_2| = \cdots = |x'_m| = |x_1| = |x_2|$ and $|y'_1| = |y'_2| = \cdots = |y'_{m-1}| = |y_1| = |y_2|$. Note that such m always exist. Now, since $xz \uparrow \theta(z)y$,

$$\begin{array}{ccccccccccc} x_1 & y_1 & x'_1 & y'_1 & x'_2 & y'_2 & \cdots & x'_{m-1} & y'_{m-1} & x'_m & \uparrow \\ \theta(x'_1) & \theta(y'_1) & \theta(x'_2) & \theta(y'_2) & \theta(x'_3) & \theta(y'_3) & \cdots & \theta(x'_m) & y_2 & x_2 & \end{array}$$

By the length argument we get, $x_1 \uparrow \theta(x'_1)$, $y_1 \uparrow \theta(y'_1)$, $x'_1 \uparrow \theta(x'_2)$, $y'_1 \uparrow \theta(y'_2)$, $\ldots, x'_{m-1} \uparrow \theta(x'_m)$, $y'_{m-1} \uparrow y_2$, $x'_m \uparrow x_2$. Here, we have two different cases. If m is odd, then

$$x_1 \uparrow \theta(x'_1) \uparrow x'_2 \uparrow \theta(x'_3) \uparrow \cdots \uparrow \theta(x'_m) \uparrow \theta(x_2)$$

and

$$y_1 \uparrow \theta(y'_1) \uparrow y'_2 \uparrow \theta(y'_3) \uparrow \cdots \uparrow y'_{m-1} \uparrow y_2.$$

Similarly, if m is even then,

$$x_1 \uparrow \theta(x'_1) \uparrow x'_2 \uparrow \theta(x'_3) \uparrow \cdots \uparrow x'_m \uparrow x_2$$

and

$$y_1 \uparrow \theta(y'_1) \uparrow y'_2 \uparrow \theta(y'_3) \uparrow \cdots \uparrow \theta(y'_{m-1}) \uparrow \theta(y_2).$$

Also, $xz \vee \theta(z)y$ is $|x|$-θ-periodic. For $1 \leq i \leq |x_1|$, consider the partial word

$$\begin{array}{ccccccc} (x_1)(i) & (x'_1)(i) & (x'_2)(i) & \cdots & (x'_{m-1})(i) & (x'_m)(i) & \vee \\ (\theta(x'_1))(i) & (\theta(x'_2))(i) & (\theta(x'_3))(i) & \cdots & (\theta(x'_m))(i) & (x_2)(i) & \end{array}$$

It is clear that the above word is 1-θ-periodic, say with letter $a_i \in V_\diamondsuit$. Similarly, for $1 \leq j \leq |y_1|$, the partial word

$$\begin{array}{ccccccc} (y_1)(j) & (y'_1)(j) & (y'_2)(j) & \cdots & (y'_{m-2})(j) & (y'_{m-1})(j) & \vee \\ (\theta(y'_1))(j) & (\theta(y'_2))(j) & (\theta(y'_3))(j) & \cdots & (\theta(y'_{m-1}))(j) & (y_2)(j) & \end{array}$$

is 1-θ-periodic, say with letter $b_j \in V_\diamondsuit$. Now, let $u = a_1a_2 \cdots a_{|x_1|}$ and $v = b_1b_2 \cdots b_{|y_1|}$. Two cases arise depending on whether the value of m is even or odd.

Case 1. Let m be odd. Then $x_1 \subset u$, $x_2 \subset \theta(u)$, $y_1 \subset v$ and $y_2 \subset v$. Then $x = x_1y_1 \subset uv$, $y = y_2x_2 \subset v\theta(u)$ and $z = x'_1y'_1x'_2y'_2 \cdots x'_{m-1}y'_{m-1}x'_m \subset (\theta(u)\theta(v)uv) \cdots (\theta(u)\theta(v)uv)\theta(u) = (\theta(u)\theta(v)uv)^{\frac{m-1}{2}}\theta(u)$.

Case 2. Let m be even. Then $x_1 \subset u$, $x_2 \subset u$, $y_1 \subset v$ and $y_2 \subset \theta(v)$ and $z = x'_1y'_1x'_2y'_2 \cdots x'_{m-1}\ y'_{m-1}x'_m \subset (\theta(u)\theta(v)uv) \cdots (\theta(u)\theta(v)uv)\theta(u)\theta(v)u = (\theta(u)\theta(v)uv)^{\frac{m-2}{2}}\theta(u)\ \theta(v)u$.

□

5.8.1 θ-Commutativity

Definition 5.58. *Let θ be either a morphic or an antimorphic involution. A partial word $u \in V_\Diamond^+$ is said to θ-commute with a partial word $v \in V_\Diamond^+$ if $uv \uparrow \theta(v)u$.*

We define the θ-commutativity order of a partial word v as $u \leq_c^{\theta_\Diamond} v$ if and only if $v \subset ux$ and $v \subset \theta(x)u$ for some $x \in V^*$. By $L_c^{\theta_\Diamond}(v) = \{u : u \in V^*, u \leq_c^{\theta_\Diamond} v\}$, we denote the set of all partial words that θ-commute with v and $\nu_c^{\theta_\Diamond}(v) = |L_c^{\theta_\Diamond}(v)|$. For a positive integer $i \geq 1$, we define $C_{\theta_\Diamond}(i) = \{v : v \in V_\Diamond^+, \nu_c^{\theta_\Diamond}(v) = i\}$.

We illustrate θ-commutative partial words in the following example.

Example 5.59. Consider the DNA alphabet $V = \{A, T, C, G\}$ and θ be an antimorphic involution such that $\theta(A) = T, \theta(T) = A, \theta(C) = G, \theta(G) = C$ and $\theta(\Diamond) = \Diamond$. Let $v = A\Diamond G\Diamond$ and $x = CTAGA\Diamond G\Diamond$. Now $\theta(x) = \Diamond C\Diamond TCTAG$ and $vx = A\Diamond G\Diamond \cdot CTAGA\Diamond G\Diamond \uparrow \theta(x)v = \Diamond C\Diamond TCTAG \cdot A\Diamond G\Diamond$. Hence, v θ-commutes with x. □

In [59], it is shown that for all $a \in V$ such that $a \neq \theta(a)$, $a^+ \subseteq C_\theta(1)$ whenever θ is either a morphic or an antimorphic involution. This result does not hold in case of partial words which is illustrated in the following example.

Example 5.60. Let $V = \{a, b\}$ and θ be an antimorphic involution such that $\theta(a) = b, \theta(b) = a$ and $\theta(\Diamond) = \Diamond$. Consider the partial word $u = a\Diamond$ and observe that $u \subset a^2$. For $v = a, x = b$, we have $u \subset vx$ and $u \subset \theta(x)v$. Hence, $a \in L_c^{\theta_\Diamond}(u)$ and thus $u \notin C_{\theta_\Diamond}(1)$. □

In our next result, we provide a relation between θ-bordered partial words and θ-commutativity. In particular, we prove that if x and y are two θ-unbordered partial words such that $x \not\uparrow y$ then x and y do not θ-commute each other where θ is a morphic involution.

Lemma 5.61. *Let $x, y \in D_{\theta_\Diamond}(1)$ such that $x \not\uparrow y$ and θ be a morphic involution. Then $xy \not\uparrow \theta(y)x$.*

Proof. Let x, y be two θ-unbordered partial words and let $xy \uparrow \theta(y)x$. Then we have the following cases to consider.

Case 1. $|x| = |y|$. Then $x \uparrow y$ which is a contradiction.

Case 2. $|x| > |y|$. Then there exists a partial word $z \in V^+$ such that $x \uparrow \theta(y)z$ and $x \uparrow zy$, that is, $x \uparrow z\theta(\theta(y))$. Assume that $\theta(y) \subset u$ for some word $u \in V^+$. Hence $x \subset uz$ and $x \subset z\theta(u)$ which implies that x is a θ-bordered partial word, a contradiction.

Case 3. $|x| < |y|$. Then there exists $t \in V_\Diamond^+$ such that $\theta(y) \uparrow xt$ and $y \uparrow tx$. Now we have $\theta(y) \uparrow xt$ and hence $y \uparrow \theta(xt) = \theta(x)\theta(t)$. Suppose $t \subset u$ for some $u \in V^+$. Then $y \uparrow tx \subset ux$ and $y \uparrow \theta(x)\theta(t) \subset \theta(x)\theta(u)$. Hence y is θ-bordered partial word, a contradiction.

Since all the cases lead to contradiction we have $xy \not\uparrow \theta(y)x$. □

The result from Theorem 2.36 provides a characterization of partial words u and v that commutes with each other. On similar lines, in the following result, we provide a characterization for partial words u and v such that u θ-commutes with v for (anti)morphic involution θ.

Theorem 5.62. *Let $u, v \in V_\diamond^+$ such that u θ-commute with v, that is, $uv \uparrow \theta(v)u$ and $uv \vee \theta(v)u$ is $|v|$-periodic.*

1. *If θ is a morphic involution then $v \subset yx$, $u \subset (xy)^i x$ as well as $v \subset \theta(x)\theta(y)$ for $i \geq 0$ and $x \in V^+, y \in V^*$.*

2. *If θ is an antimorphic involution then $v \subset yx, u \subset (xy)^i x$ for $i \geq 0$ as well as $v \subset \theta(y)\theta(x)$ where $x \in V^+, y \in V^*$.*

Proof. We prove the result only for a morphic involution as the proof for an antimorphic is similar. Let $uv \uparrow \theta(v)u$ and $uv \vee \theta(v)u$ is $|v|$-periodic. Then by Theorem 2.34, there exist $x \in V^+$ and $y \in V^*$ such that $\theta(v) \subset xy$, $v \subset yx$ and $u \subset (xy)^i x$ for some $i \geq 0$. Furthermore, $\theta(v) \subset xy$ implies that $v \subset \theta(xy) = \theta(x)\theta(y)$ as θ is a morphic involution. □

5.9 Conclusions

We introduced the notion of θ-(un)bordered partial words which is a generalization of (un)bordered partial words, and extended the results on θ-bordered words to partial words for the case when θ is a morphic as well as an antimorphic involution. In particular, we provided the characterization of θ-unbordered partial words for antimorphic involution θ, and gave the necessary and sufficient condition for the concatenation of two θ-unbordered partial words to be θ-unbordered for morphic involution θ. We initiated a study of disjunctivity properties of set of θ-unbordered partial words, and proved that the set of all θ-unbordered partial words $D_{\theta_\diamond}(1)$ is a disjunctive language. We proved that, for an antimorphic involution θ, the set $D_{\theta_\diamond}(1)$ is regular, and the set of all θ-bordered partial words is not context-free for a morphic involution θ. We introduced the notion of θ-primitivity, θ-conjugacy and θ-commutativity in partial words and provided a characterization for a partial word to be θ-conjugate of another partial word.

Chapter 6

Palindromes and Pseudo-Palindromes

In this chapter, we study the notion of palindromes and pseudo-palindromes in partial words which are extension of palindromes in total words. We also explore some combinatorial properties related to palindromes and pseudo-palindromes.

6.1 Introduction

Palindromicity is a well studied property in formal language theory and word combinatorics. Study of palindromes in the field of word combinatorics has been motivated by the work of A. De Luca [32] and Lyndon and Schützenberger [72]. It is well-known that the language of palindromes over a finite alphabet is a linear context-free language [48]. A string is a palindrome if it reads forward the same as backward. Let $w = a_0 a_1 \dots a_{n-1}$ be a word of length n with $a_i \in V$ for $0 \le i \le n-1$. The reverse of w is the word $w^r = a_{n-1} \dots a_1 a_0$. Then the word w is said to be a palindrome if $w = w^r$. For example, the string "*madam*" is a palindrome.

Several combinatorial and algorithmic problems on palindromes have been taken into consideration such as unavoidability of primitive and palindromic words [40], palindromic richness [83] and involutively palindromes [61]. There is a linear time algorithm to count the number of distinct palindromes in a word [45]. The language of palindromic words has been studied with respect to the class of regular as well as context-free languages. In [51], Horváth, Karhumäki and Klejin have given the characterization of palindromic languages in relation to regular as well as context-free languages. Dömösi, Horváth, Ito and Katsura, [35], have investigated the languages which are primitive as well as palindromic, and proved that the language of primitive palindromes is not a context-free language. Also the language of primitive palindromes has been studied with respect to unavoidability class of languages. In [40], Fazekas, Leupold and Shikishima-Tsuji proved that the language of primitive words which are not palindromes are unavoidable for context-free languages.

A partial word w over an alphabet is said to be a palindrome if $w \uparrow w^r$. A partial word x is said to be a θ-palindrome if $x \uparrow \theta(x)$ where θ is either a morphic or an antimorphic involution. When θ is an identity function, θ-palindromes become palindromes in partial words under the assumption of θ being an antimorphic involution.

Organization. This chapter is organized as follows. We begin with Section 6.2 by defining the palindromes in partial words and study some properties in relation with primitive partial words. We provide a necessary and sufficient condition for the concatenation of a partial word w and its reverse w^r to be a nonprimitive partial word. The bound on the number of palindromes in the conjugacy class of a primitive partial word which is also a palindrome is studied in Section 6.3. We define θ-palindromes for partial word and establish a connection between θ-bordered and θ-palindromic partial words in Section 6.4. For a primitive partial word with one hole u, we count the number of θ-borders of u^j for some j when θ is a morphic involution. We give a lower bound on the number of θ-borders of some power of a θ-palindromic primitive partial word with arbitrary number of holes. We end this chapter with a concluding remark in Section 6.5.

6.2 Palindromes in Partial Words

In this section, we define palindromes in partial words and study the properties of palindromes in connection with primitive partial words.

Definition 6.1. *A partial word w is said to be a palindrome if $w^r \uparrow w$ where w^r is the reverse of w.*

Example 6.2. Let $w = ab\Diamond b\Diamond a$. Then $w^r = a\Diamond b\Diamond ba$ and $w \uparrow w^r$. Hence w is a palindromic partial word. □

We know that for a partial word w, there is a set of total words in which w can be contained. If at least one of the total words from that set is nonprimitive, then w is considered as a nonprimitive partial word. Similarly, the idea for a partial word w to be a palindrome is that among that set of words in which w is contained, if at least one word is found to be a palindrome, then the partial word w is considered to be a palindrome. Recall that a word is primitive if and only if it is not a nontrivial power of any other word. We use the symbol $\mathcal{P}al$ to denote the language of palindromic words. The result in Proposition 3.1 shows that a word w is primitive if and only if w^r is primitive.

Next we observe that the concatenation of a partial word w with its reverse w^r generates a nonprimitive partial word.

Proposition 6.3. *Let w be a palindromic partial word. Then ww^r and w^rw are nonprimitive partial words.*

Proof. Let $w \in V_\Diamond^*$ be a palindrome. We have $w \uparrow w^r$. By the definition of compatibility there exists a word v such that $w \subset v$ and $w^r \subset v$. Hence $ww^r \subset v \cdot v = v^2$ by the law of multiplication. Thus ww^r is nonprimitive.

Similarly, it is easy to follow that $w^r w$ is a nonprimitive partial word. $\square$

We use the symbol $\mathcal{P}al_\Diamond$ to denote the set of partial words which are also palindromes. We can observe that the language $\mathcal{P}al_\Diamond$ contains both primitive and nonprimitive partial words. For example, the palindromes $a\Diamond b\Diamond aa$ and $ab\Diamond\Diamond ba$ are primitive and nonprimitive partial words, respectively.

Lemma 6.4 ([91])**.** *Let $u,\ v \in V^+$. If $uv \in \mathcal{P}al \cap Q$ then either $u \notin \mathcal{P}al$ or $v \notin \mathcal{P}al$.*

But the above result does not hold in case of partial words.

Example 6.5. Let $u = a, v = b\Diamond$ be two partial words. Observe that $uv = ab\Diamond \in \mathcal{P}al_\Diamond$ but, both $u \in \mathcal{P}al_\Diamond$ and $v \in \mathcal{P}al_\Diamond$. $\square$

The following proposition shows that if a word uv which is the catenation of a nonempty palindrome u and a nonempty word v is a palindrome then vu is not a palindrome.

Proposition 6.6 ([91])**.** *Let u be a nonempty palindrome and $v \in V^+$. If $uv \in \mathcal{P}al \cap Q$ then $vu \notin \mathcal{P}al$.*

The above result need not be true for partial words as illustrated by the following example.

Example 6.7. Consider the partial words $u = a\Diamond$ and $v = ba$ over an alphabet $V = \{a, b\}$. Observe that u is a palindrome whereas v is not and, $uv = a\Diamond ba \in \mathcal{P}al_\Diamond \cap Q_p$. But $vu = baa\Diamond$ is also a palindrome. $\square$

Next we characterize the structure of palindromes in partial words. Note that the set of all palindromic partial words $\mathcal{P}al_\Diamond$ can contain even length as well as odd length palindromes.

Proposition 6.8. *(a) An even length palindromic partial word w can be expressed as $w = xy$ such that $|x| = |y|$ and $y \uparrow x^r$.*

(b) An odd length palindromic partial word w can be expressed as $w = xay$ such that $|x| = |y|$, $a \in V_\Diamond$ and $y \uparrow x^r$.

Observe that the language of palindromic partial words over an alphabet V is context-free. Next we give a context-free grammar for the language of palindromic partial words over a two letter alphabet, which can be generalized to any finite size alphabet.

Lemma 6.9. *The language of palindromic partial words $\mathcal{P}al_\Diamond$ is context-free over V where $V = \{a,\ b\}$.*

Proof. Consider the context-free grammar $G = \langle V,\ T,\ P,\ S\rangle$ for $\mathcal{P}al_{\Diamond}$ where $V = \{S,\ A,\ B\}$, $T = \{a,\ b,\ \Diamond\}$, S is the starting symbol and the set of production rules P are as follows:

$$S \longrightarrow aSa \mid bSb \mid aS\Diamond \mid \Diamond Sa \mid bS\Diamond \mid \Diamond Sb \mid \Diamond S\Diamond$$
$$S \longrightarrow a \mid b \mid \Diamond \mid \varepsilon$$

For completeness, we prove that the context-free grammar given above generates the palindromes in partial words over an alphabet $\{a, b\}$.

To prove our claim, we will show in both directions, that is, $L(G) \subseteq \mathcal{P}al_{\Diamond}$ and $\mathcal{P}al_{\Diamond} \subseteq L(G)$.

($\Rightarrow$) Let w be a partial word generated by the grammar G. We prove by using induction on the number of derivation steps from the set of production rules.

Base Case: If the number of derivation step is one then one of the four productions $S \to a, S \to b, S \to \Diamond, S \to \varepsilon$ will be used. Since $a, b, \Diamond$ and ε are palindromes, the base case holds.

Induction: Assume that it is true for all the partial words $x \in V_{\Diamond}^*$ which are derived by using n derivation steps, that is, x is a palindrome. Consider a partial word w which is derived by $(n+1)$-steps. Now w will be one of the form $w = \alpha x \beta$ for $\alpha, \beta \in \{a, b\}_{\Diamond}$ and $\alpha \uparrow \beta$. By induction hypothesis, x is derived by using n-steps and $x \uparrow x^r$. Thus, $w = \alpha x \beta \uparrow (\alpha x \beta)^r = \beta x^r \alpha = w^r$ and w is a palindrome. Hence, $L(G) \subseteq \mathcal{P}al_{\Diamond}$.

($\Leftarrow$) Let w be a partial word which is also a palindrome. We prove by using induction on the length of w.

Base Case: If $|w| = 0$ or $|w| = 1$. Then w is one of the form $w = \varepsilon,\ w = a$ or $w = b$, and w can be generated by using the rules $S \to \varepsilon,\ S \to a$ or $S \to b$.

Induction: Let us assume that $|w| \geq 2$. Since $w \uparrow w^r$, then the first and last symbols of w must be compatible. Then $w = \alpha x \beta$ for $\alpha, \beta \in \{a, b\}_{\Diamond}$ and $x \uparrow x^r$ by induction hypothesis. If $w = \alpha x \beta$ for $\alpha \uparrow \beta$ then by using induction hypothesis $G \overset{*}{\Rightarrow} x$. We can obtain $w = \alpha x \beta$ for $\alpha \uparrow \beta$ by using one of production rule from G which is of the form $S \to aSa, S \to aS\Diamond, S \to \Diamond Sa, S \to \Diamond S\Diamond$ for $a \in \{a, b\}$. Thus, w is a palindromic partial word and $w \in L(G)$. Hence $\mathcal{P}al_{\Diamond} \subseteq L(G)$.

We conclude that $L(G) = \mathcal{P}al_{\Diamond}$. □

Next we define the perfect shuffle of two words of same length and give a necessary and sufficient condition for a partial word to be an even length palindrome.

Definition 6.10. *Let $u = u_1u_2\cdots u_n$ and $v = v_1v_2\cdots v_n$ be two words of length n. The perfect shuffle of u and v is denoted as $u \sqcup\!\sqcup v$ and defined as $u \sqcup\!\sqcup v = u_1v_1u_2v_2\cdots u_nv_n$. For example, let $u = plnrm$ and $v = aidoe$, and then $u \sqcup\!\sqcup v = palindrome$.*

Proposition 6.11. *A partial word x is an even length palindrome if and only if there exist partial words y and z such that $x = y \sqcup\!\sqcup z$ and $z \uparrow y^r$.*

Proof. We prove the necessary and sufficient conditions as follows.
($\Leftarrow$) Let y and z be two partial words such that $x = y \sqcup z$ and $z \uparrow y^r$. Suppose $y = a_1a_2\cdots a_n$ and $z = b_1b_2\cdots b_n$. Thus $z = b_1b_2\cdots b_n \uparrow y^r = a_na_{n-1}\cdots a_1$ and $x = a_1a_2\cdots a_n \sqcup b_1b_2\cdots b_n = a_1b_1a_2b_2\cdots a_nb_n \subset uu^r$ for some $u \in V^*$. Hence x is an even length palindrome.
($\Rightarrow$) Let x be a partial word which is also an even length palindrome. Then $x \subset uu^r$ for some word $u \in V^*$. Since x is an even length palindrome, we can write $x = x_1x_2$ where $|x_1| = |x_2|$. We can construct y and z by taking the alternating positions of x. Hence we can write $x = y \sqcup z$ where $z \uparrow y^r$. □

In our next result, we identify a relation between palindromes and primitive roots in total words.

Theorem 6.12. *Let u and v be two total words. If u, uv and uv^r are palindromes, then v is also palindrome. Moreover, there exists a primitive word $x \in V^*$ such that x is also a palindrome for which $u,\ v \in x^*$.*

Proof. Let $u = u_1u_2\cdots u_m$ and $v = v_1v_2\cdots v_n$ where $u_i, v_j \in V$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. We have three cases to consider.

Case 1. If $m = n$. We have $uv = u_1u_2\cdots u_mv_1v_2\cdots v_m$ and $uv^r = u_1u_2\cdots u_mv_mv_{m-1}\cdots v_1$. Since uv^r is a palindrome, $u_1 = v_1, u_2 = v_2, \ldots, u_m = v_m$ which implies that $u = v$. Thus, v is a palindrome.

Case 2. If $m < n$. We have $uv = u_1u_2\cdots u_mv_1v_2\cdots v_mv_{m+1}\cdots v_n$ and $uv^r = u_1u_2\cdots u_mv_nv_{n-1}\cdots v_mv_{m-1}\cdots v_1$. Since uv^r is a palindrome, $u_1 = v_1, u_2 = v_2, \ldots, u_m = v_m$ and $v_nv_{n-1}\cdots v_{m+1}$ is a palindrome. Since uv is a palindrome,

$$u_1 = v_n = v_1$$
$$u_2 = v_{n-1} = v_2$$
$$u_m = v_{n+1-m} = v_m$$

and $v_1v_2\cdots v_{n-m}$ is a palindrome. Now, for $v = v_1v_2\cdots v_mv_{m+1}\cdots v_n$, we have $v_1v_2\cdots v_{n-m}$ is palindrome and $v_nv_{n-1}\cdots v_{m+1}$ is palindrome which implies that $v_{m+1}v_{m+2}\cdots v_n$ is also a palindrome.

Now we see that, for $i = 1$ to m,

$$v_i = v_{n+1-i}$$

and for $i > m$,

$$v_{m+1} = v_n = v_1 = v_{n-m}, v_{m+2} = v_{n-1} = v_2 = v_{n-m-1}$$

Similarly, we can show that $v_i = v_{n+1-i}$ for all $i > m$. Thus, v is a palindrome.

Case 3. If $m > n$, we can prove that v is a palindrome similar to Case 2.

Since uv is a palindrome,

$$uv = (uv)^r = v^r u^r = v^r u \tag{6.1}$$

From the Equation (6.1) and Theorem 2.25, there exist words p and q such that $v = pq$, $v^R = qp$ and $u = (pq)^e p$ for $e \geq 0$.

We have $v = v^r$ which implies that $pq = (pq)^R = q^r p^r = qp$. Thus there exists a word z such that $p,\ q \in z^*$. Since $v \neq \varepsilon$, $v = pq \neq \varepsilon$, hence $z \neq \varepsilon$. Let x be the primitive root of z. Because $p,\ q \in x^*$, $u = (pq)^e p \in x^+$ and $v = pq \in x^+$. Therefore $uv = x^m$ for some $m \geq 1$. Since uv is a palindrome, it implies that x^m is also a palindrome. Hence $x^m = (x^m)^r = (x^r)^m$ which proves that $x = x^r$. □

It is known that the primitive root of a palindromic total word is again a palindrome, that is, for $w \in V^+$, w is a palindrome if and only if $\sqrt{w}$ is also a palindrome [32]. But this does not hold in case of partial words. For a palindromic partial word w, all the roots in the set need not be palindromic. For example, consider a partial word $w = a\Diamond b\Diamond$ for which $w \subset \{aaba,\ aabb,\ abba,\ abbb\}$. The root $abba$ is a palindrome whereas the root $abbb$ is not a palindrome.

The following result gives a necessary and sufficient condition for the concatenation of a partial word and its reverse to be nonprimitive.

Theorem 6.13. *Let w be a nonempty partial word with $|\alpha(w)| \geq 2$. Then the word ww^r is nonprimitive if and only if at least one of the following conditions hold:*

(a) The partial word w is a palindrome.

(b) The partial word w has a strong period $2n$ and for some n, $n \mid |w|$ such that $w[1 \ldots 2n]$ is a palindrome, if it exists, that is, $2n \leq |w|$.

Proof. We prove the necessary and sufficient conditions as follows:
($\Leftarrow$) Suppose w is a palindrome. Therefore, $w \uparrow w^r$. Thus there exists a word x such that $w \subset x$ and $w^r \subset x$. So $ww^r \subset xx$ and ww^r is nonprimitive.

Suppose there exists n such that $n \mid |w|$ and $w[1 \ldots 2n]$ is a palindrome. Thus, $w[1 \ldots n] \uparrow w[2n \ldots n+1]$. Hence there exists a word $y[1..n]$ such that $w[1 \ldots n] \subset y[1 \ldots n]$ and $w[n+1 \ldots 2n] \subset y[1 \ldots n]^r$. Hence ww^r is contained in a nontrivial power of $y[1 \ldots n]y[1 \ldots n]^r$ and ww^r is nonprimitive.
($\Rightarrow$) Let us assume that ww^r is a nonprimitive palindrome. There are two different cases depending upon the degree of ww^r is even or odd.

Case 1. If the degree of ww^r is even, then $ww^r \subset x^k$, $k \geq 2$ and k is even. Thus, $w \subset x^{\frac{k}{2}}$ and $w^r \subset x^{\frac{k}{2}}$ which implies that $w \uparrow w^r$. Hence w is a palindrome.

Case 2. If the degree of ww^r is odd. Since ww^r is nonprimitive palindrome, then $ww^r \subset x^k$ where $x \in V^+$ and $|\alpha(x)| \geq 2$. The odd value of k implies that $|x|$ must be even. Considering the length of w, it can be seen that w must be bordered with overlap of length $\frac{|x|}{2}$. Also $|x|$ is a strong period of w as well as of w^r. Since ww^r is a palindrome, by looking at the center of ww^r, we have $w[1\ldots\frac{|x|}{2}] \uparrow w[\frac{|x|}{2}+1\ldots|x|]^r$. So $w[1\ldots\frac{|x|}{2}]\cdot w[\frac{|x|}{2}+1\ldots|x|]$ must be a palindrome. So $n=\frac{|x|}{2}$ and the conditions are satisfied.

□

6.3 Conjugates and Density of Palindromes

Recall that a language $L \subseteq V^*$ is said to be reflective if $xy \in L$ implies that $yx \in L$ for all $x,\ y \in V^*$. In Proposition 2.41, it has been shown that the language of primitive partial words is reflective. The cyclic permutation on a partial word can be considered as an operation. The cyclic permutations of a partial word can be obtained by cutting a prefix and pasting it at the end. An invariant property of primitive partial words is that primitive partial words are closed under cyclic permutation. The cyclic permutations of a primitive partial word forms the conjugacy class of that primitive partial word. It has been shown that the number of conjugates in the conjugacy class of a primitive partial word of length n is n in Lemma 3.19.

For a partial word w, if w is a palindrome, it is not necessary that all its cyclic permutations are also palindromes. For example, let $w = abb\Diamond$ be a palindromic partial word. Then $Conj(w) = \{abb\Diamond,\ \Diamond abb,\ b\Diamond ab,\ bb\Diamond a\}$. Consider one of its conjugate $w' = \Diamond abb$ which is not a palindrome. Thus, for a partial word, palindrome property is not closed under cyclic permutation.

Consider a partial word $w = abba\Diamond bba$ which is a nonprimitive and palindromic partial word. In its conjugacy class, there is a partial word $baabba\Diamond b$ which is also a palindrome. Thus, for a primitive partial word w with at most one hole which is also a palindrome, the set of cyclic permutations of w does not contain arbitrary number of palindromes. The following result gives a bound on the number of primitive palindromes in the set of cyclic permutations of a partial word with at most one hole. We consider two cases, viz., even and odd length primitive partial palindromes.

Theorem 6.14. *Let w be an even length primitive palindromic partial word with at most one hole. Then the conjugacy class of w contains exactly two primitive palindromes.*

Proof. Let $w = xy$ be an even length primitive palindromic partial word with at most one hole where $|x| = |y| = \frac{|w|}{2}$ and $y \uparrow x^r$. So $w = xy \uparrow xx^r$ and $yx \uparrow x^r x$ are palindromes in the given partial word. Suppose apart from these two partial words there is one more palindrome in the conjugacy class of w. Let $x'y'$ be the palindrome which is obtained by less than $\frac{|w|}{2}$ cyclic permutation steps where $y' \uparrow x'^r$. Let us assume that $x' = u$ and $y' \uparrow u^r$. Thus, uu^r is compatible to a factor of $xx^r x$. This gives a factorization of $x = x_1x_2x_3$ where x_1 is the nonempty prefix of x until the start of x', x_3 is the nonempty factor of y that is compatible to a suffix of u until the start of y', and x_2 is possibly the empty string between x_1 and x_3 in x.

We know that $u = (u^r)^r$. So $u \uparrow (x_2^r x_1^r x_1)^r = x_1^r x_1 x_2$. Similarly, we have $u^r \uparrow (x_2x_3x_3^r)^r = x_3x_3^r x_2^r$. If $x_2 = \varepsilon$ then we have $u \uparrow u^r$. So u is a palindrome and

$x'y' \uparrow uu^r$ implies that $x'y'$ is nonprimitive. If $x_2 \neq \varepsilon$ then it can be seen that $x'y'$ is compatible to an internal factor of $uu^r uu^r$. As $x'y' \uparrow uu^r$, so $uu^r \uparrow uu^r uu^r$. Hence by Proposition 2.42, $x'y'$ is nonprimitive which is a contradiction to the fact that the conjugacy class of a primitive partial word contains only primitive partial words. Therefore xy and yx are the only palindromes in the conjugacy class of w.

□

The above result does not hold for an even length primitive palindromic partial words having two or more holes as illustrated in the following example.

Example 6.15. Let $w = a\Diamond\Diamond b\Diamond a$ be a primitive partial word of even length which is also a palindrome with two holes. The conjugate class of w is

$$\mathrm{Conj}(w) = \{a\Diamond\Diamond b\Diamond a,\ aa\Diamond\Diamond b\Diamond,\ \Diamond aa\Diamond\Diamond b,\ b\Diamond aa\Diamond\Diamond,\ \Diamond b\Diamond aa\Diamond,\ \Diamond\Diamond b\Diamond aa\}$$

The partial words $a\Diamond\Diamond b\Diamond a$, $\Diamond aa\Diamond\Diamond b$, $b\Diamond aa\Diamond\Diamond$ and $\Diamond\Diamond b\Diamond aa$ in the set $\mathrm{Conj}(w)$ are palindromes. □

In case of total words, it is known that every odd length palindrome which is also primitive is unique in its conjugate class [36]. However, this is not true in case of partial words, and there may be more than one palindromes in the conjugacy class of an odd length primitive palindromic partial word. This is shown in the following example.

Example 6.16. Let $w = a\Diamond bba$ be a partial word over $V = \{a, b\}$. Observe that w is a primitive palindromic partial word and is of odd length. The conjugacy class of w is $\{a\Diamond bba,\ aa\Diamond bb,\ baa\Diamond b,\ bbaa\Diamond,\ \Diamond bbaa\}$, and the partial words $a\Diamond bba$ and $baa\Diamond b$ are two palindromes. □

In the following result, we prove that if the catenation of a partial word and its reverse is contained in a nontrivial power of some word, then the word is a palindrome.

Lemma 6.17. *Let u be a partial word such that $u \not\uparrow u^r$. If $uu^r \subset x^k$, $k \geq 2$ then x is a palindrome and k is an odd number.*

Proof. Suppose k is even. Then $uu^r \subset x^k = (x^{\frac{k}{2}})^2$ which implies that $u \uparrow u^r$. This is a contradiction. So k is odd. As $|uu^r|$ is even and k is odd, then $|x|$ is even. Let $x = yy'$ where $|y| = |y'|$. We can observe that a prefix of u of length $|y|$ is contained in y and a suffix of length $|y'|$ of u^r is contained in y'. Hence $y' = y^r$ and $x = yy^r$. Now $x = yy^r = x^r$ and thus x is a palindrome. □

Next we show that if an integer power of a partial word is a palindrome then the partial word is itself a palindrome.

Lemma 6.18. *For a partial word w, if w^n is a palindrome for $n \geq 1$ then w is a palindrome.*

Proof. We prove it using strong induction on n. For $n = 1$, the statement is true. Let us assume that it is true for all $k < n$ that is if w^k is a palindrome for all $k \leq n-1$, then w is a palindrome. Now we will prove it for n.

Suppose that w^n is a palindrome. We can write $w^n = w^{n-1}w = w^1 w^{n-1}$. Now, $w^n = ww^{n-1} \uparrow (w^n)^r = (w^r)^n = (w^r)(w^r)^{n-1}$. As $|w| = |w^r|$ and $w^{n-1} \uparrow (w^r)^{n-1}$ then by simplification, we have $w \uparrow w^r$. Hence w is a palindrome. □

The concatenation of two palindromic partial words need not be a palindrome. For example, let $u = ab\Diamond$ and $v = baab$. Both u and v are palindromes but uv is not a palindrome. In [39], Fazekas et al. presented a necessary and sufficient condition for the concatenation of two words to be a palindrome. We have the following observation.

Corollary 6.19. *Let u and v are two palindromic partial words. If u and v are contained in powers of some palindromic word z then uv is a palindrome.*

The converse of the above result is not true which is shown in the following example.

Example 6.20. Consider the partial words $u = a$ and $v = b\Diamond$ over the alphabet $V = \{a, b\}$. It is easy to see that both u and v are palindromes. Now $uv = ab\Diamond$ is a palindrome and $u \subset a$ and $v \subset b^2$. Thus, u and v are not contained in power of a palindromic word. □

Next we see the relation between palindromic partial words and bordered partial words. We know that a nonempty partial word w is called bordered if one of its proper prefixes is compatible with its suffix of the same length. Otherwise, if no nonempty words x, y and z exist such that $w \subset xy$ and $w \subset zx$ then w is called unbordered. If w is unbordered and $w \subset w'$, then w' is unbordered as well. A bordered partial word need not be a palindrome, whereas a palindrome is always bordered. For example, for the partial word $w = aba\Diamond$, w is a bordered partial word but not a palindrome.

Proposition 6.21. *Let w be a nonempty partial word. If w is a palindrome then it is bordered.*

Proof. Let $w = xyz$ be a palindromic partial word where $x, z \in V_\Diamond$ and $y \in V_\Diamond^*$. Since w is a palindrome, then $x \uparrow z$. Hence w is bordered. □

In our next result, we give a necessary and sufficient condition for a bordered partial word to be a palindrome.

Theorem 6.22. *A bordered partial word w is palindrome if and only if $w = xy$ or $w = xay$ for some x and y where $|x| = |y|$, $a \in V_\Diamond$ and $y \uparrow x^r$.*

Proof. We prove the necessary and sufficient conditions as follows:
($\Leftarrow$) Let $w = xy$ or $w = xay$ for some nonempty partial words x, y such that $|x| = |y|$ and $y \uparrow x^r$. In the first case, if $w = xy$ and $y \uparrow x^r$, then we can write $w = xy \uparrow xx^r$ and hence w is a palindrome and bordered. In the latter case, if $w = xay$ such that $|x| = |y|$, $a \in V_\Diamond$ and $y \uparrow x^r$. So $w \uparrow xax^r$. Hence w is a palindrome and bordered.

($\Rightarrow$) Let w be a bordered and palindromic partial word. From Proposition 6.8, we know that every palindromic partial word can be written as either $w = xy$ or $w = xay$ where $|x| = |y|$ and $y \uparrow x^r$ and $a \in V_\Diamond$. □

Next we prove that the language of palindromic partial words is dense over an alphabet V. Recall that a language of partial words $L \subseteq V_\Diamond^*$ is called dense if $V_\Diamond^* w V_\Diamond^* \cap L \neq \emptyset$, that is, for every partial word $w \in V_\Diamond^*$ there exist partial words x and y such that $xwy \in L$.

Lemma 6.23. *The language $\mathcal{P}al_\Diamond$ is dense over an alphabet V.*

Proof. Let $u \in V_\Diamond^*$. There are two possibilities depending on whether u is a palindrome or not.

Case I Let u be a palindrome. So we can choose $x = y = \varepsilon$ and we will have $xuy \in \mathcal{P}al_\Diamond$.

Case II Let us assume that u is not a palindrome. Taking $x = \varepsilon$ and $y = z$ where z is a partial word and $z \uparrow u^r$, we have $xuy \in \mathcal{P}al_\Diamond$.
Hence the language of partial palindromes $\mathcal{P}al_\Diamond$ is dense.

□

6.4 θ-Palindromic Partial Words

In this section, we introduce the notion of θ-palindromes for partial words under the assumption of θ to be (anti)morphic involution. The notation $\mathcal{P}al_\theta$ denotes the set of θ-palindromes in total words over an alphabet. We prove some combinatorial properties of θ-palindromic partial words in connection with θ-borders of a partial word. We count the number of θ-borders of a power of a palindromic primitive partial word with one hole assuming θ to be a morphic involution.

Definition 6.24. *A partial word $u \in V_\Diamond^*$ is said to be a θ-palindrome if and only if $u \uparrow \theta(u)$ where θ is either a morphic or an antimorphic involution.*

Example 6.25. Let θ be an antimorphic involution such that $\theta(a) = b, \theta(b) = a$ and $\theta(\Diamond) = \Diamond$. Consider a partial word $u = a\Diamond\Diamond bb$. Then $\theta(u) = \theta(b)\theta(b)\theta(\Diamond)$ $\theta(\Diamond)\theta(a) = aa\Diamond\Diamond b$ and hence $u \uparrow \theta(u)$. Thus, u is a θ-palindromic partial word. □

We use the notation $\mathcal{P}al_{\theta_\Diamond}$ for the set of all θ-palindromic partial words over an alphabet. Let us recall the following result which gives the number of θ-borders of a θ-palindromic word of length n.

Lemma 6.26 ([58]). *Let θ be an antimorphic involution and x be a θ-palindromic word of length n. Then $x \in D_\theta(n)$.*

However, the above result does not necessarily hold for partial words.

Example 6.27. Let θ be an antimorphic involution such that $\theta(a) = b, \theta(b) = a$ and $\theta(\Diamond) = \Diamond$. Let $u = a\Diamond\Diamond bb$ be a partial word of length 5. It is easy to see that $u \uparrow \theta(u)$. Now, $L_d^{\theta_\Diamond}(u) = \{\varepsilon, a, aa, aab, aaa, aaab, aabb\}$ and $u \in D_{\theta_\Diamond}(7)$. □

Consider an example. Let θ be an antimorphic involution such that $\theta(a) = b$, $\theta(b) = a$ and $\theta(\Diamond) = \Diamond$. For $u = ab\Diamond aa\Diamond b$, we have $\theta(u) = a\Diamond bb\Diamond ab$. Since $u \not\uparrow \theta(u)$, $u \notin \mathcal{P}al_{\theta_\Diamond}$. We have $L_d^{\theta_\Diamond}(u) = \{\varepsilon, a, ab, abb\}$. Let us consider another partial word $v = au\theta(a) = aab\Diamond aa\Diamond bb$. For v, we have $L_d^{\theta_\Diamond}(v) = \{\varepsilon, a, aa, aab, aabb\}$. Observe that, for $u \notin \mathcal{P}al_{\theta_\Diamond}$, $u \in D_{\theta_\Diamond}(4)$ and $v = au\theta(a) \in D_{\theta_\Diamond}(5)$. We generalize the result as follows.

Proposition 6.28. *Let θ be an antimorphic involution and $x \in V_\Diamond^*$ such that $x \notin \mathcal{P}al_{\theta_\Diamond}$. Then for all $i \in \mathbb{N}$, $x \in D_{\theta_\Diamond}(i)$ if and only if $ax\theta(a) \in D_{\theta_\Diamond}(i+1)$.*

Proof. ($\Leftarrow$) Let $v = ax\theta(a) \in D_{\theta_\Diamond}(i+1)$. Then $L_d^{\theta_\Diamond}(v) = \{\varepsilon, av_1, av_2, \ldots, av_i\}$ for all $v_i \in \mathrm{CPref}(x)$ which implies that $\theta(v_i) \in \mathrm{CSuff}(x)$. It is not possible to have $x \subset v_i$ because then $\theta(x) \subset v_i$ and $x \subset v_i$ will imply that $x \uparrow \theta(x)$, and hence $x \in \mathcal{P}al_{\theta_\Diamond}$ which is a contradiction. Hence for all $1 \leq j \leq i-1$, $av_j \in L_d^{\theta_\Diamond}(v)$ implies that $v_j <_d^{\theta_\Diamond} x$. Thus $L_d^{\theta_\Diamond}(x) = \{\varepsilon, v_1, v_2, \ldots, v_{i-1}\}$ and $x \in D_{\theta_\Diamond}(i)$.

($\Rightarrow$) Let x be a partial word with $x \notin \mathcal{P}al_{\theta_\Diamond}$ and $x \in D_{\theta_\Diamond}(i)$. Then $L_d^{\theta_\Diamond}(x) = \{\varepsilon, v_1, v_2, \ldots, v_{i-1}\}$ and for all $u \in L_d^{\theta_\Diamond}(x)$, $u <_d^{\theta_\Diamond} x$ implies that there exist $y, z \in V^+$ such that $x \subset uy$ and $x \subset z\theta(u)$. Let $v = ax\theta(a)$. We have $v \subset auy\theta(a) = auy_1$ and $v \subset az\theta(u)\theta(a) = z_1\theta(u)\theta(a)$ for some $y_1 = y\theta(a)$, $z_1 = az$. Observe that for all $u \in L_d^{\theta_\Diamond}(x)$, $au \in L_d^{\theta_\Diamond}(v)$. Suppose there exists a word w such that $w \in \mathrm{CPref}(x)$ such that $w \notin L_d^{\theta_\Diamond}(x)$ and $aw <_d^{\theta_\Diamond} v$. Then there exist some $p, q \in V^+$ such that $v \subset awp$ and $v \subset q\theta(w)\theta(a)$. If $|x| = |w|$ then $x \subset w$ and $x \subset \theta(w)$ which implies that $\theta(x) \subset \theta(\theta(w)) = w$. Hence $x \uparrow \theta(x)$ which shows that $x \in \mathcal{P}al_{\theta_\Diamond}$, a contradiction. If $|w| < |x|$ then $w \in \mathrm{CPref}(x)$ which implies that $x \subset wy_2$ and $x \subset z_2\theta(w)$ for some $y_2, z_2 \in V^+$ which is a contradiction to the assumption that $w \notin L_d^{\theta_\Diamond}(x)$. Hence $L_d^{\theta_\Diamond}(v) = \{\varepsilon, a, av_1, av_2, \ldots, av_{i-1}\}$ which implies that $v = ax\theta(a) \in D_{\theta_\Diamond}(i+1)$. □

Next we recall a result which talks about the number of θ-borders of a θ-primitive word which is a θ-palindrome as well.

Proposition 6.29 ([58]). *Let u be a θ-palindromic primitive word and $j \geq 1$ be an integer.*

(a) For a morphic involution θ, $\nu_d^\theta(u^j) = \nu_d^\theta(u) + j - 1$.

(b) For an antimorphic involution θ, $\nu_d^\theta(u^j) = |u^j| = j \times |u|$.

Both the above results are not true in case of partial words.

Example 6.30. Let $V = \{a, b\}$ and θ be (anti)morphic involution.

(a) Let θ be a morphic involution such that $\theta(a) = a, \theta(b) = b$ and $\theta(\lozenge) = \lozenge$. Let $u = b\lozenge a$ be a partial word. We have $u \uparrow \theta(u)$ and $L_d^{\theta_\lozenge}(u) = \{\varepsilon, ba\}$. Let $v = u^2$ and $L_d^{\theta_\lozenge}(v) = \{\varepsilon, ba, baa, bba\}$, and hence $\nu_d^{\theta}(u^2) \neq \nu_d^{\theta}(u) + 2 - 1$.

(b) Let θ be an antimorphic involution such that $\theta(a) = b, \theta(b) = a$ and $\theta(\lozenge) = \lozenge$. For $u = a\lozenge b$, $u \uparrow \theta(u)$ and hence $u \in \mathcal{P}al_{\theta_\lozenge}$. Let $v = u^2 = a\lozenge ba\lozenge b$. Then we have $L_d^{\theta_\lozenge}(u) = \{\varepsilon, a, aa, ab\}$ and $L_d^{\theta_\lozenge}(v) = \{\varepsilon, a, aa, ab, aab, abb, aaba, abba, aabaa, abbaa, aabab, abbab\}$. Observe that $\nu_d^{\theta_\lozenge}(u^2) \neq 2 \times |u|$.

□

To prove Proposition 6.31, we recall a result from Chapter 3 which states that a nonempty partial word u is primitive if and only if u^{n-1} is not compatible to a central factor of u^n for all $n \geq 2$.

We have the following result to count the number of θ-borders of some power of a partial word with one hole when θ is a morphic involution.

Proposition 6.31. *Let θ be a morphic involution and $j \geq 2$. Let $u \in V_\lozenge^+$ be a θ-palindromic primitive partial word with one hole, then*

$$\nu_d^{\theta_\lozenge}(u^j) = \nu_d^{\theta_\lozenge}(u) + k + k^2 + \cdots + k^{j-1}$$

where k is the size of the alphabet.

Proof. Let $u = a_1a_2\cdots a_n$ be a primitive partial word with one hole. Since u is θ-palindrome for a morphic involution θ, $u \uparrow \theta(u)$. Thus $a_1a_2\cdots a_n \uparrow \theta(a_1a_2\cdots a_n) = \theta(a_1)\theta(a_2)\cdots\theta(a_n)$ which implies that $a_i = \theta(a_i)$ for all $i \in \{1, 2, \ldots, n\}$. Hence $\nu_d^{\theta_\lozenge}(u) = \nu_d(u)$, where $\nu_d(u)$ is the number of borders of the partial word u.

For $j = 2$, $u^2 = uu$ will have exactly two holes and since u is primitive, by Lemma 3.7, u will not be compatible to a central factor of u^2. The borders of u^2 will include the borders of u. Moreover, u will be a proper prefix and as well as proper suffix of u^2. Since u has one hole, that hole can be replaced by any of the k symbols from the alphabet V. Thus $\nu_d^{\theta_\lozenge}(u^2) = \nu_d(u^2) = \nu_d(u) + k$.
For u^j, by Lemma 3.7 observe that u^{j-1} is not compatible to a proper factor of u^j . Hence, all of $u, u^2, \ldots, u^{j-1}$ are proper prefixes as well as proper suffixes of u^j. Also u^{j-1} has exactly $j-1$ holes. Thus,

$$\nu_d^{\theta_\lozenge}(u^j) = \nu_d^{\theta_\lozenge}(u) + k + k^2 + \cdots + k^{j-1}.$$

□

Corollary 6.32. *Let θ be a morphic involution and let u be a primitive partial word with one hole. If $u \in D_{\theta_\lozenge}(i) \cap \mathcal{P}al_{\theta_\lozenge} \cap Q_p$ then for $j \geq 2$,*

$$u^j \in D_{\theta_\lozenge}(i + \sum_{p=1}^{j-1} k^p)$$

where k is the size of the alphabet.

In our next result, we give a necessary and sufficient condition for a partial word to be a θ-palindrome under an antimorphic involution θ.

Lemma 6.33. *Let $w \in V_\diamond^+$ be a partial word and θ be an antimorphic involution. Then $w \in \mathcal{P}al_{\theta_\diamond}$ if and only if $w = xyz$ with $x \in V_\diamond^+$, $y \in V_\diamond^*$ such that $z \uparrow \theta(x)$ and $y \in \mathcal{P}al_{\theta_\diamond}$.*

Proof. We prove the necessary and sufficient condition as follows.

($\Leftarrow$) Let $w = xyz$ be a nonempty partial word such that $z \uparrow \theta(x)$ with $y \in \mathcal{P}al_{\theta_\diamond}$. Hence $y \uparrow \theta(y)$. Now $\theta(w) = \theta(xyz) = \theta(z)\theta(y)\theta(x) \uparrow xy\theta(x)$. Hence $w \uparrow \theta(w)$ and thus $w \in \mathcal{P}al_{\theta_\diamond}$.

($\Rightarrow$) Let $w \in V_\diamond^+$ be a partial word and $w \in \mathcal{P}al_{\theta_\diamond}$. Then $w \uparrow \theta(w)$. Since $|w| \geq 2$, we can write w as xyz with $x,\ z \in V_\diamond^+$, $y \in V_\diamond^*$ and $|x| = |z|$. From $w \uparrow \theta(w)$, we have $xyz \uparrow \theta(xyz) = \theta(z)\theta(y)\theta(x)$. Comparing both sides, we get $y \uparrow \theta(y)$ and $z \uparrow \theta(x)$ and hence $y \in \mathcal{P}al_{\theta_\diamond}$. □

In Proposition 6.31, we count the number of θ-borders of some integer power of a θ-palindromic primitive partial word with one hole. Next, we count the number of θ-borders of some power of a θ-palindromic primitive partial word with arbitrary number of holes and provide a lower bound.

Lemma 6.34. *Let u be a θ-palindrome and primitive partial word. Let $w = u^j$ for $j \geq 1$ and θ be a morphic involution. Then*

$$\nu_d^{\theta_\diamond}(w) \geq \nu_d^{\theta_\diamond}(u) + \sum_{i=1}^{j-1} k^{|u^i|_\diamond}$$

where k is the size of the alphabet.

Proof. We give a constructive proof of the lemma. Let θ be a morphic involution. Let $u = a_1 \ldots a_n$ be a θ-partial palindrome. Then $u \uparrow \theta(u)$, that is, $a_1 a_2 \ldots a_n \uparrow \theta(a_1)\theta(a_2)\ldots\theta(a_n)$ which implies that $a_i = \theta(a_i)$, $\forall i \in \{1, \ldots, n\}$. Let $L_d^{\theta_\diamond}(u) = \{\varepsilon,\ u_1,\ u_2, \ldots, u_k\}$. Let $w = u^{(j)} = u \ldots u$ (j *times*). If $j = 1$ then $\nu_d^{\theta_\diamond}(w) = \nu_d^{\theta_\diamond}(u)$ which satisfies the condition. For $j \geq 2$, observe that $u^1,\ u^2, \ldots, u^{j-1}$ are proper prefixes as well as proper suffixes of u^j.

Consider the case where u is a proper prefix and proper suffix of u^j for $j = 2$. The number of contained prefixes as well as contained suffixes of u is same as the number of words in the set in which u is contained. If u has i number of holes then u will be contained in the set of total words $S = \{u_1,\ u_2, \ldots, u_{k^i}\}$ where k is the size of the alphabet and $|S| = k^i$. Each of the element in the set S is a θ-border of uu. Also there may be some other proper prefixes of uu of length more than $|u|$ which are also a θ-border of uu. Hence

$$\nu_d^{\theta_\diamond}(u^2) \geq \nu_d^{\theta_\diamond}(u) + k^{|u|_\diamond}.$$

Similarly, u and uu are proper prefixes as well as proper suffixes of u^j for $j = 3$. We have $\nu_d^{\theta_\diamond}(u^3) \geq \nu_d^{\theta_\diamond}(u) + k^{|u|_\diamond} + k^{|uu|_\diamond}$. Hence in general, we have

$$\nu_d^{\theta_\diamond}(w) = \nu_d^{\theta_\diamond}(u^j) \geq \nu_d^{\theta_\diamond}(u) + \sum_{i=1}^{j-1} k^{|u^{(i)}|_\diamond}.$$

□

In the following lemma, we prove that for two nonempty total words u and v, if uv is a primitive word as well as a θ-palindrome then either u is not a θ-palindrome or v is not a θ-palindrome under the assumption of θ to be an antimorphic involution.

Lemma 6.35. *Let $u,\ v \in V^+$ and θ is an antimorphic involution. Then $uv \in \mathcal{P}al_\theta \cap Q$ implies that $u \notin \mathcal{P}al_\theta$ or $v \notin \mathcal{P}al_\theta$.*

Proof. We prove it by contradiction. Let us assume that both $u \in \mathcal{P}al_\theta$ and $v \in \mathcal{P}al_\theta$. Then $u = \theta(u)$ and $v = \theta(v)$. Also $uv \in \mathcal{P}al_\theta$ implies that $uv = \theta(uv) = \theta(v)\theta(u) = vu$. Then by Theorem 2.26, we know that both u and v are the powers of same word, that is, there exists a word $x \in Q$ such that $u = x^i$ and $v = x^j$ for $i,\ j \geq 1$. Now $uv = x^i \cdot x^j = x^{i+j} = x^k$ for $k \geq 2$ which is a contradiction to the assumption that $uv \in Q$. Hence either $u \notin \mathcal{P}al_\theta$ or $v \notin \mathcal{P}al_\theta$.

□

Lemma 6.36. *Let $u \in \mathcal{P}al_\theta$ and θ be either a morphic or an antimorphic involution. If $u = x(yx)^i,\ i \geq 1$ for some $x,\ y$ then $x,\ y \in \mathcal{P}al_\theta$.*

Proof. Let $u \in V^+$ be a θ-palindrome. Then $u = \theta(u)$. Given that $u = x(yx)^i$ for some $i \geq 1$. If θ is a morphic involution then $u = \theta(u)$ implies that $u = x(yx)^i = \theta(x)(\theta(y)\theta(x))^i$. Since $|x| = |\theta(x)|$ then $x = \theta(x)$ and $y = \theta(y)$. Hence $x,\ y \in \mathcal{P}al_\theta$.

If θ is an antimorphic involution then $u = x(yx)^i = \theta(x(yx)^i) = \theta(x)(\theta(x)\ \theta(y))^i$. Similarly, $|x| = |\theta(x)|$ implies that $x = \theta(x)$ and $y = \theta(y)$. Hence $x,\ y \in \mathcal{P}al_\theta$.

□

Corollary 6.37. *Let $u \in \mathcal{P}al_\theta$ and θ is either a morphic or an antimorphic involution. If $u = (xy)^i x$ for some $x,\ y$ then $x,\ y \in \mathcal{P}al_\theta$.*

Proof. It follows from Lemma 6.36. □

In the next result, we prove that for a given θ-palindromic primitive word u, if uv is a θ-palindromic primitive for some $v \in V^*$ then vu is not a θ-palindrome.

Proposition 6.38. *Let $u \in V^+$ be a θ-palindrome where θ is an antimorphic involution and $v \in V^*$.*

(a) If $uv \in \mathcal{P}al_\theta \cap Q$ implies that $vu \notin \mathcal{P}al_\theta$.

(b) If $vu \in \mathcal{P}al_\theta \cap Q$ then $uv \notin \mathcal{P}al_\theta$.

Proof. Consider the first statement. Given that $u \in \mathcal{P}al_\theta$, $uv \in \mathcal{P}al_\theta$ which implies that $u = \theta(u)$ and $uv = \theta(uv) = \theta(v)\theta(u) = \theta(v)u$. Then by Theorem 2.25, there exist $x_1, y_1 \in V^*$ such that $u = (x_1y_1)^{i_1}x_1, v = y_1x_1$ and $\theta(v) = x_1y_1$ for some $i_1 \geq 0$. As $u = \theta(u)$, by Lemma 6.36 $x_1, y_1 \in \mathcal{P}al_\theta$. If $y_1 = \varepsilon$, we have $uv = x_1^{i_1+2} \notin Q$ which is a contradiction. Similarly, if $x_1 = \varepsilon$, we have $uv = y_1^{i_1+2} \notin Q$ which is a contradiction.

Assume the contrary, that is, $vu \in \mathcal{P}al_\theta$. Hence $vu = \theta(vu) = \theta(u)\theta(v) = u\theta(v)$. Now $vu = u\theta(v)$ implies that there exist x_2, $y_2 \in V^*$ such that $v = x_2y_2$, $\theta(v) = y_2x_2$ and $u = (x_2y_2)^{i_2}x_2$ for $i_2 \geq 0$. As $u \in \mathcal{P}al_\theta$, by Lemma 6.36, $x_2, y_2 \in \mathcal{P}al_\theta$. We have the following two cases to consider.

Case 1. $v = y_1x_1 = x_2y_2$. We have $u = (x_1y_1)^{i_1}x_1 = (x_2y_2)^{i_2}x_2$. Since $|x_1y_1| = |y_1x_1|$, we conclude that $x_1y_1 = y_1x_1$. By Theorem 2.26, there exists a word p such that $x_1 = p^i, y_1 = p^j$ which implies that $uv = (x_1y_1)^{i_1}x_1y_1x_1 = p^{(i+j)(i_1+1)+i} \notin Q$, a contradiction.

Case 2. $\theta(v) = x_1y_1 = y_2x_2$. Also we have $u = (x_2y_2)^{i_2}x_2 = (x_1y_1)^{i_1}x_1$. By length argument $|x_2y_2| = |y_2x_2$, we conclude that $x_2y_2 = y_2x_2$. By Theorem 2.26, there exists a word q such that $x_2 = q^i, y_2 = q^j$ which implies that $uv = (x_2y_2)^{i_2}x_2x_2y_2 = q^{(i+j)i_2+2i+j} \notin Q$, a contradiction.

Therefore $uv \in \mathcal{P}al_\theta \cap Q$ implies that $vu \notin \mathcal{P}al_\theta$. Similarly, second statement holds true. □

6.5 Conclusions

In this chapter, we extended the notion of palindromes and θ-palindromes in partial words and identified several combinatorial properties that relate primitive partial words and palindromes. We provided a necessary and sufficient condition for the concatenation of a partial word and its reverse to be a nonprimitive partial word. For an even length primitive partial word w with at most one hole which is also a palindrome, we proved that there are exactly two palindromes in its conjugacy class. We counted the number of θ-borders of a partial word $u^j, j \geq 1$ where u is a primitive partial word with one hole as well as palindromic. Also the number of θ-borders of a θ-palindromic partial word raised to some integer power is given.

Chapter 7

Conclusions and Future Works

In this chapter, we summarize the thesis and its contributions. We also state some of the problems related to the thesis work which are open and are of interest for further research.

We have studied the language of primitive partial words with respect to Chomsky hierarchy. We have characterized several robust subclasses of language of primitive partial words under various point mutation operations. We have also considered a variation of bordered partial words known as pseudo-bordered partial words and identified some combinatorial properties related to pseudo-palindromes.

We began with Chapter 1 by giving the motivation for the Combinatorics on words and partial words, and stated the overview of the thesis. In Chapter 2, we presented the necessary fundamentals and existing results which were required for understanding this thesis work.

In Chapter 3, we studied the position of the language of primitive partial words Q_p over an alphabet V in conventional Chomsky hierarchy. We proved that the language of primitive partial words Q_p is not regular, not linear and not a deterministic context-free language. We also proved that the language of nonprimitive partial words $\overline{Q}_p$ is not a context-free language. But whether the language of primitive partial words is context-free or not is still open. We presented a 2DPDA automaton to recognize the language of primitive partial words with one hole. An indexed grammar to generate the set of nonprimitive words has been given.

In Chapter 4, we investigated the robustness of primitive partial words. Primitivity of partial words is an important property in combinatorics on partial words. We explored several point mutation operations such as deletion of a symbol, insertion of a symbol, exchanging two distinct consecutive symbols and substituting a symbol by another symbol. We characterized non-del-robust, non-exchange-robust and non-subst-robust primitive partial words with one hole. We also proved certain combinatorial properties related to each of the aforementioned classes. We provided a lower bound on the number of primitive partial words with one hole of length n over an alphabet of size k that are robust under deletion operation.

In Chapter 5, we extended the notion of bordered partial words to pseudo-bordered partial words under the assumption of an arbitrary involution function. We also investigated the properties of θ-bordered and θ-unbordered partial words for morphic as well as antimorphic involutions θ. We proved some basic properties of θ-(un)bordered partial words that includes closure properties as well as characterization of the set of θ-unbordered partial words for an antimorphic involution θ. It has been shown that the set of all θunbordered partial words $D_{\theta_\diamond}(1)$ is a disjunctive language. The language of θ-unbordered partial words has been proved to be regular under the assumption of θ being an antimorphic involution. Also we proved that the language of θ-bordered partial words is not a context-free language when θ is a morphic involution. The concepts such as θ-primitivity, θ-conjugacy and θ-commutativity are studied with respect to (anti)morphic involution θ. The characterization for two partial words x and y such that x is θ-conjugate of y has been provided.

In Chapter 6, we considered the notion of palindromes in partial words which is an extension of palindromes in total words. A necessary and sufficient condition has been given to show under what conditions the concatenation of a partial word and its reverse is nonprimitive. We provided a bound on the number of palindromes in the conjugacy class of an even length primitive palindromic partial word. We proved that the language of palindromes in partial words is context-free as well as dense language. We also introduced the notion of θ-palindromes where θ can be either a morphic or an antimorphic involution. We counted the number of θ-borders of a partial word u^j when θ is a morphic involution.

Some unsolved questions related to the problems addressed in the thesis are given below, which might be of interest for research in future.

- Whether the language of primitive partial words Q_p is context-free or not is one of the central problem to investigate.
- We conjecture that the set of primitive partial words with at least two holes is recognized by a 2NPDA.
- Construction of an Indexed grammar for the set of primitive words would be interesting.
- The complete characterization of several classes of primitive partial words with arbitrary number of holes which are del-robust, exchange-robust, subst-robust have to be investigated. Also one can explore primitive partial words under the operation of taking prefixes.
- We have simple $O(n^2)$ algorithm to test whether a given primitive partial word is either del-robust, exchange-robust or subst-robust. It would be interesting to find linear time algorithms for above robustness property on primitive partial words.
- Effect of homomorphisms on primitive partial words to preserve primitivity can be investigated.
- Further research on θ-(un)bordered partial words including disjunctivity properties of languages related to $D_{\theta_\diamond}(i)$.

- It will be interesting to study the extended Fine and Wilf's theorem, critical factorization theorem, Lyndon-Schutzenberger equations with respect to (anti)morphic involutions θ.

Summary of Our Work

Languages	**Context-free**	**Approach**
Q_p	?	-
$\overline{Q}_p$	No	Pumping Lemma
Q_p^{1D}	?	-
$Q_p^{1\overline{D}}$	No	Pumping Lemma
$Q_p^{\overline{X}}$	No	Ogden's Lemma and Closure properties of gsm
$D_{\theta_\diamond}(1)$	Yes	Grammar
$D_{\theta_\diamond}(i)$, $i \geq 2$	No	Pumping Lemma
$\mathcal{P}al_\diamond$	Yes	Grammar

Table 7.1: Summary of different languages and their position with respect to context-free class

Languages	**Reflective**
Q_p	Yes
$\overline{Q}_p$	Yes
Q_p^i	Yes
Q_p^{1D}	Yes
$Q_p^{1\overline{D}}$	Yes
$Q_p^{1\overline{X}}$	No
Q_p^{1S}	Yes
$D_{\theta_\diamond}(i)$	No

Table 7.2: Reflectivity of different languages

Appendix A

Pumping Lemmata

We recall the pumping lemmata for regular, linear and context-free languages which are used to show that the language of primitive partial words is not regular, not linear and the language of nonprimitive partial words is not context-free in Chapter 3. We also prove that the language of non-del-robust primitive partial words with one hole Q_p^{1D} and the language of θ-bordered partial words under the assumption θ to be morphic are not context-free languages by using pumping lemma. We also recall Ogden's lemma to show that the language of non-exchange-robust primitive partial words over an alphabet is not context-free in Chapter 4.

Lemma A.1 (Pumping Lemma for Regular Languages [48])**.** *For a regular language L, there exists an integer $n > 0$ such that for every word $w \in L$ with $|w| \geq n$, there exist a decomposition of w as $w = xyz$ such that the following conditions holds.*

(i) $|y| > 0$,

(ii) $|xy| \leq n$, *and*

(iii) $xy^iz \in L$ *for all* $i \geq 0$.

Lemma A.2 (Pumping Lemma for Linear Languages [49])**.** *Let L be a linear language. There exists an integer n such that any word $p \in L$ with $|p| \geq n$, admits a factorization $p = uvwxy$ satisfying the following conditions.*

(i) $|uvxy| \leq n$,

(ii) $|vx| > 0$, *and*

(iii) $uv^mwx^my \in L\ \forall m \in \mathbb{N}$.

Lemma A.3 (Pumping Lemma for Context-free Languages [48])**.** *Let L be a CFL. Then there exists an integer $n > 0$ such that for every $u \in L$ with $|u| \geq n$, u can be decomposed into $vwxyz$ such that the following conditions hold:*

(a) $|wxy| \leq n$.

(b) $|wy| > 0$.

(c) $vw^ixy^iz \in L$ *for all* $i \geq 0$.

Lemma A.4 (Ogden's Lemma [77]). *For every context-free language* L*, there exists a positive integer* N *such that every string* $z \in L$ *in which at least* N *symbols are distinguished, that is, marked can be written in the form* $z = uvwxy$ *for* $u, v, w, x, y \in V^*$ *such that the following conditions hold:*

(i) vwx *contains at most* N *marked symbols*

(ii) v *and* x *together have at least one marked symbols*

(iii) for all $i \geq 0$, $uv^iwx^iy \in L$.

References

[1] Alfred V Aho. Indexed grammars - an extension of context-free grammars. *Journal of the ACM (JACM)*, 15(4):647–671, 1968.

[2] Emily Allen, Francine Blanchet-Sadri, Michelle Bodnar, Brian Bowers, Joe Hidakatsu, and John Lensmire. Combinatorics on partial word borders. *Theoretical Computer Science*, 609:469–493, 2016.

[3] Jean Berstel and Luc Boasson. Partial words and a theorem of Fine and Wilf. *Theoretical Computer Science*, 218(1):135–141, 1999.

[4] Jean Berstel and Dominique Perrin. The origins of combinatorics on words. *European Journal of Combinatorics*, 28(3):996–1022, 2007.

[5] Jean Berstel, Dominique Perrin, and Christophe Reutenauer. *Codes and Automata*. Cambridge University Press, 2009.

[6] Francine Blanchet-Sadri. Codes, orderings, and partial words. *Theoretical Computer Science*, 329(1):177–202, 2004.

[7] Francine Blanchet-Sadri. Periodicity on partial words. *Computers and Mathematics with Applications*, 47(1):71–82, 2004.

[8] Francine Blanchet-Sadri. Primitive partial words. *Discrete Applied Mathematics*, 148(3):195–213, 2005.

[9] Francine Blanchet-Sadri. *Algorithmic Combinatorics on Partial Words*. CRC Press, 2007.

[10] Francine Blanchet-Sadri, Deepak Bal, and Gautam Sisodia. Graph connectivity, partial words, and a theorem of Fine and Wilf. *Information and Computation*, 206(5):676–693, 2008.

[11] Francine Blanchet-Sadri, D. Dakota Blair, and Rebeca V. Lewis. Equations on partial words. In *Mathematical Foundations of Computer Science*, pages 167–178. Springer, 2006.

[12] Francine Blanchet-Sadri, Naomi C Brownstein, Andy Kalcic, Justin Palumbo, and Tracy Weyand. Unavoidable sets of partial words. *Theory of Computing Systems*, 45(2):381–406, 2009.

[13] Francine Blanchet-Sadri and Mihai Cucuringu. Counting primitive partial words. *Journal of Automata, Languages and Combinatorics*, 15(3/4):199–227, 2010.

[14] Francine Blanchet-Sadri, C. D. Davis, Joel Dodge, Robert Mercaş, and Margaret Moorefield. Unbordered partial words. *Discrete Applied Mathematics*, 157(5):890–900, 2009.

[15] Francine Blanchet-Sadri and S. Duncan. Partial words and the critical factorization theorem. *Journal of Combinatorial Theory, Series A*, 109(2):221–245, 2005.

[16] Francine Blanchet-Sadri, Rachel Harred, and Justin Lazarow. Longest common extensions in partial words. In *International Workshop on Combinatorial Algorithms*, pages 52–64. Springer, 2015.

[17] Francine Blanchet-Sadri and Robert A. Hegstrom. Partial words and a theorem of fine and wilf revisited. *Theoretical Computer Science*, 270(1):401–419, 2002.

[18] Francine Blanchet-Sadri, Raphaël M. Jungers, and Justin Palumbo. Testing avoidability on sets of partial words is hard. *Theoretical Computer Science*, 410(8):968–972, 2009.

[19] Francine Blanchet-Sadri and D. K. Luhmann. Conjugacy on partial words. *Theoretical Computer Science*, 289(1):297–312, 2002.

[20] Francine Blanchet-Sadri, Robert Mercaş, Sean Simmons, and Eric Weissenstein. Avoidable binary patterns in partial words. *Acta Informatica*, 48(1):25–41, 2011.

[21] Francine Blanchet-Sadri, Sarah Nelson, and Amelia Tebbe. On operations preserving primitivity of partial words with one hole. In *Automata and Formal Languages*, pages 93–107, 2011.

[22] Robert S. Boyer and J. Strother Moore. A fast string searching algorithm. *Communications of the ACM*, 20(10):762–772, 1977.

[23] Piotr Bylanski and Derek George Woodward Ingram. *Digital transmission systems*. Peregrinus, 1976.

[24] H. Z. Q. Chen, S. Kitaev, T. Mütze, and B. Y. Sun. On universal partial words. *arXiv e-prints:1601.06456*, January 2016.

[25] Christian Choffrut and Juhani Karhumäki. Combinatorics of words, handbook of formal languages, vol. 1: word, language, grammar, 1997.

[26] Richard Cole, Lee-Ad Gottlieb, and Moshe Lewenstein. Dictionary matching and indexing with errors and don't cares. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 91–100. ACM, 2004.

[27] Maxime Crochemore and Dominique Perrin. Two-way string-matching. *Journal of the ACM (JACM)*, 38(3):650–674, 1991.

[28] Maxime Crochemore and Wojciech Rytter. *Jewels of Stringology: Text Algorithms*. World Scientific, 2002.

[29] Elena Czeizler, Lila Kari, and Shinnosuke Seki. On a special class of primitive words. *Theoretical Computer Science*, 411:617–630, 2010.

[30] Jürgen Dassow, Florin Manea, and Robert Mercaş. Regular languages of partial words. *Information Sciences*, 268:290–304, 2014.

[31] Jürgen Dassow, Gema M. Martín, and Francisco J. Vico. Some operations preserving primitivity of words. *Theoretical Computer Science*, 410(30):2910–2919, 2009.

[32] Aldo De Luca. On some combinatorial problems in free monoids. *Discrete Mathematics*, 38(2):207–225, 1982.

[33] Pál Dömösi and Géza Horváth. The language of primitive words is not regular: two simple proofs. *Bulletin of European Association for Theoretical Computer Science*, 87:191–194, 2005.

[34] Pál Dömösi, Sándor Horváth, Masami Ito, László Kászonyi, and Masashi Katsura. Formal languages consisting of primitive words. In *Fundamentals of Computation Theory*, pages 194–203. Springer, 1993.

[35] Pál Dömösi, Sándor Horváth, Masami Ito, and Masashi Katsura. Some remarks on primitive words and palindromes. In *Descriptional Complexity of Formal Systems*, pages 245–254, 2003.

[36] Pál Dömösi and Masami Ito. *Context-free Languages and Primitive Words*. World Scientific, 2014.

[37] Pál Dömösi, Masami Ito, and Solomon Marcus. Marcus contextual languages consisting of primitive words. *Discrete Mathematics*, 308(21):4877–4881, 2008.

[38] Chen-Ming Fan, Jen-Tse Wang, and Cheng-Chih Huang. Borderedness-preserving homomorphisms. *Information Processing Letters*, 114(10):529–534, 2014.

[39] Szilárd Zsolt Fazekas, Peter Leupold, and Kayoko Shikishima-Tsuji. Palindromes and primitivity. In *Automata and Formal Languages*, pages 184–196, 2011.

[40] Szilard Zsolt Fazekas, Peter Leupold, and Kayoko Shikishima-Tsuji. On non-primitive palindromic context-free languages. *International Journal of Foundations of Computer Science*, 23(06):1277–1289, 2012.

[41] Nathan J. Fine and Herbert S. Wilf. Uniqueness theorems for periodic functions. *Proceedings of the American Mathematical Society*, 16(1):109–114, 1965.

[42] Michael J Fischer and Michael S Paterson. String-matching and other products. Technical report, MASSACHUSETTS INST OF TECH CAMBRIDGE PROJECT MAC, 1974.

[43] Paweł Gawrychowski, Florin Manea, and Dirk Nowotka. Discovering hidden repetitions in words. In *The Nature of Computation. Logic, Algorithms, Applications*, volume 7921, pages 210–219.

[44] Pawel Gawrychowski, Florin Manea, Robert Mercaş, Dirk Nowotka, and Catalin Tiseanu. Finding pseudo-repetitions. *Leibniz International Proceedings in Informatics*, 20:257–268, 2013.

[45] Richard Groult, Élise Prieur, and Gwénaël Richomme. Counting distinct palindromes in a word in linear time. *Information Processing Letters*, 110(20):908–912, 2010.

[46] Dan Gusfield. *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*. Cambridge University Press, 1997.

[47] Michael A. Harrison. *Introduction to Formal Language Theory*. Addison-Wesley Longman Publishing Co., Inc., 1978.

[48] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. Introduction to automata theory, languages, and computation. *ACM SIGACT News*, 32(1):60–65, 2001.

[49] Géza Horváth and Benedek Nagy. Pumping lemmas for linear and nonlinear context-free languages. *CoRR*, abs/1012.0023, 2010.

[50] Sándor Horváth. Strong interchangeability and nonlinearity of primitive words. In *AMAST Workshop on Algebraic Methods in Language Processing*, volume 95, pages 173–178, 1995.

[51] Sándor Horváth, Juhani Karhumäki, and H. C. M. Kleijn. Results concerning palindromicity. *Journal of Information Processing and Cybernetics*, 23(8-9):441–451, 1987.

[52] S. C. Hsu, Masami Ito, and H. J. Shyr. Some properties of overlapping order and related languages. *Soochow Journal of Mathematics*, 15(1):29–45, 1989.

[53] C. C. Huang, Pei-Ching Hsiao, and C. J. Liau. A note of involutively bordered words. *Journal of Information and Optimization Sciences*, 31(2):371–386, 2010.

[54] Masami Ito, M. Katsura, H. J. Shyr, and S. S. Yu. Automata accepting primitive words. *Semigroup Forum*, 37(1):45–52, 1988.

[55] Masami Ito, Gabriel Thierrin, and S. S. Yu. Right k-dense languages. *Semigroup Forum*, 48(1):313–325, 1994.

[56] Lila Kari and Manasi S. Kulkarni. Pseudo-identities and bordered words. *Discrete Mathematics and Computer Science*, pages 207–222, 2014.

[57] Lila Kari and Manasi S. Kulkarni. Disjunctivity and other properties of sets of pseudo-bordered words. *Acta Informatica*, pages 1–20, 2015.

[58] Lila Kari and Kalpana Mahalingam. Involutively bordered words. *International Journal of Foundations of Computer Science*, 18(05):1089–1106, 2007.

[59] Lila Kari and Kalpana Mahalingam. Watson-crick conjugate and commutative words. In *International Workshop on DNA-Based Computers*, pages 273–283. Springer, 2007.

[60] Lila Kari and Kalpana Mahalingam. Watson-Crick bordered words and their syntactic monoid. *International Journal of Foundations of Computer Science*, 19(05):1163–1179, 2008.

[61] Lila Kari and Kalpana Mahalingam. Watson-Crick palindromes in DNA computing. *Natural computing*, 9(2):297–316, June 2010.

[62] Richard M Karp. *Complexity of computation*, volume 7. American Mathematical Soc., 1974.

[63] Donald E. Knuth, James H. Morris, JR., and Vaughan R. Pratt. Fast pattern matching in strings. *SIAM Journal on Computing*, 6(2):323–350, 1977.

[64] Yoshiyuki Kunimochi. A context sensitive grammar generating the set of all primitive words. In *Algebras, Languages, Computations and their Applications*, volume 1562, pages 143–145, 2007.

[65] Peter Leupold. Languages of partial words. *Grammars*, 7:179–192, 2004.

[66] Peter Leupold. Partial words for DNA coding. In *DNA Computing*, pages 224–234. Springer, 2005.

[67] Peter Leupold. Primitive words are unavoidable for context-free languages. In *Language and Automata Theory and Applications*, pages 403–413. Springer, 2010.

[68] Gerhard Lischke. Primitive words and roots of words. *Acta Universitatis Sapientiae, Informatica*, 3(1):5–34, 2011.

[69] M. Lothaire. *Combinatorics on Words.* Cambridge University Press, 1997.

[70] M. Lothaire. *Algebraic Combinatorics on Words*, volume 90. Cambridge University Press, 2002.

[71] M Lothaire. *Applied Combinatorics on Words*, volume 105. Cambridge University Press, 2005.

[72] Roger C. Lyndon and Marcel P. Schützenberger. The equation $a^M = b^N c^P$ in a free group. *The Michigan Mathematical Journal*, 9:289–298, 1962.

[73] Florin Manea, Mike Müller, and Dirk Nowotka. On the pseudoperiodic extension of $u^l = v^m w^n$. *Leibniz International Proceedings in Informatics*, 24, 2013.

[74] Florin Manea, Mike Müller, Dirk Nowotka, and Shinnosuke Seki. Generalised Lyndon-Schützenberger equations. In *Mathematical Foundations of Computer Science*, pages 402–413. Springer, 2014.

[75] Victor Mitrana. Primitive morphisms. *Information Processing Letters*, 64(6):277–281, 1997.

[76] Victor Mitrana. Some remarks on morphisms and primitivity. *Bulletin-European Association for Theoretical Computer Science*, 62:213–215, 1997.

[77] William Ogden. A helpful result for proving inherent ambiguity. *Theory of Computing Systems*, 2(3):191–194, 1968.

[78] Gheorghe Păun, Nicolae Santean, Gabriel Thierrin, and Sheng Yu. On the robustness of primitive words. *Discrete Applied Mathematics*, 117(1):239–252, 2002.

[79] Gheorghe Paun and Gabriel Thierrin. Morphisms and primitivity. *Bulletin-European Association for Theoretical Computer Science*, 61:85–88, 1997.

[80] Holger Petersen. The ambiguity of primitive words. In *Symposium on Theoretical Aspects of Computer Science*, pages 679–690. Springer, 1994.

[81] Holger Petersen. On the language of primitive words. *Theoretical Computer Science*, 161(1):141–156, 1996.

[82] C. M. Reis and H. J. Shyr. Some properties of disjunctive languages on a free monoid. *Information and Control*, 37(3):334 – 344, 1978.

[83] Antonio Restivo and Giovanna Rosone. Burrows–wheeler transform and palindromic richness. *Theoretical Computer Science*, 410(30-32):3018–3026, 2009.

[84] Grzegorz Rozenberg and Arto Salomaa. *Handbook of Formal Languages: Volume 3 Beyond Words*. Springer Science & Business Media, 2012.

[85] H. J Shyr. *Free Monoids and Languages*. Department of Mathematics, Soochow University, Taipei, Taiwan, 1979.

[86] H. J. Shyr and Gabriel Thierrin. Disjunctive languages and codes. In *Fundamentals of Computation Theory*, pages 171–176. Springer, 1977.

[87] H. J. Shyr and S. S. Yu. Non-primitive words in the language p^+q^+. *Soochow Journal of Mathematics*, 20(4):535–546, 1994.

[88] Richard Edwin Stearns, Juris Hartmanis, and Philip M Lewis. Hierarchies of memory limited computations. In *Switching Circuit Theory and Logical Design, 1965*, pages 179–190. IEEE, 1965.

[89] James A. Storer. *Data Compression: Methods and Theory*. Computer Science Press, Inc., 1988.

[90] S. S. Yu. d-minimal languages. *Discrete Applied Mathematics*, 89(1):243–262, 1998.

[91] S. S. Yu. Palindrome words and reverse closed languages. *International Journal of Computer Mathematics*, 75(4):389–402, 2000.

[92] S. S. Yu. Post-plus languages. *Theoretical Computer Science*, 255(1):85–105, 2001.

[93] S. S. Yu. Languages and codes. *Lecture Notes, Department of Computer Science, National Chung-Hsing University, Taichung, Taiwan*, 402, 2005.

List Of Publications

[1] **A. C. Nayak** and Kalpesh Kapoor. On the Language of Primitive Partial Words. In 9^{th} *International Conference on Language and Automata Theory and Applications*, volume 8977 of *Lecture Notes in Computer Science*, pages 436-445. Springer International Publishing, 2015.

[2] **A. C. Nayak** and Amit K. Srivastava. On Del-Robust Primitive Partial Words with One Hole. In 10^{th} *International Conference on Language and Automata Theory and Applications*, volume 9618 of *Lecture Notes in Computer Science*, pages 233-244. Springer International Publishing, 2016.

[3] **A. C. Nayak**, Amit K. Srivastava and Kalpesh Kapoor. On Exchange-Robust and Subst-Robust Primitive Partial Words. In 17^{th} *Italian Conference on Theoretical Computer Science*, volume 1720 of *CEUR workshop proceedings*, pages 190-202, 2016.

[4] Manasi S. Kulkarni, **A. C. Nayak** and Kalpana Mahalingam. Watson-Crick Partial Words. In *International Conference on Theory and Practice of Natural Computing*, pages 190-202, 2017.

Manuscripts Communicated / Under Preparation

[5] **A. C. Nayak** and Kalpesh Kapoor. Morphism and Primitive Partial Words.

[6] **A. C. Nayak** and Kalpesh Kapoor. Palindromes and Pseudo-Palindromes in Partial Words.

[7] Manasi S. Kulkarni, **A. C. Nayak** and Kalpana Mahalingam. On Pseudo-Bordered Partial Words.(Submitted)