Università di Firenze, Università di Perugia, INdAM consorziate nel CIAFM

DOTTORATO DI RICERCA
IN MATEMATICA, INFORMATICA, STATISTICA

CURRICULUM IN MATEMATICA
CICLO XXXVIII

Sede amministrativa Università degli Studi di Firenze
Coordinatrice Prof.ssa Alessandra Sestini

An Algebraic Approach to the Study of Quantum Finite Automata

Settore Scientifico Disciplinare MATH-02/A

Dottorando:
Andrea Benso

Supervisore:
Prof. Flavio D'Alessandro

Coordinatrice:
Prof.ssa Alessandra Sestini

Anni 2022/2025

I deeply thank Prof. Flavio D'Alessandro for his supervision and continuous support throughout my research. His availability and guidance have been a constant reference point throughout my doctoral studies. I also thank the Committee for their comments and suggestions during the preparation of my thesis, which have improved its structure and clarity. Finally, I thank the University of Florence for the support that made this doctoral research possible.

Abstract

In this thesis we will study the decidability of the Intersection problem for the measure-once quantum finite automata introduced by Moore and Crutchfield in the 2000s. Specifically, we ask for which families of grammars it is possible to find conditions ensuring that, given a language recognized by such model of computation and a language generated by such families of grammars, it is decidable whether or not they have a nonempty intersection. Throughout the thesis, we will investigate the correlation between this problem and some decidability problems for matrices subsets on the Zariski topology, so as to apply the corresponding results to show the decidability of the problem for languages generated by monoidal context-free grammars and restricted matrix context-free grammars of finite index.

Prior Publication of Content:

Material covered in Chapter 4 has been previously published in:

- A. Benso, A. Carpi, F. D'Alessandro. On the commutative equivalence of algebraic structures and related problems, *Journal of Automata, Languages and Combinatorics*, Vol. 30, pp. 27–48, 2025.
- A. Benso, F. D'Alessandro, and P. Papi. Quantum automata and languages of finite index, in Proceedings of RP 2024, 26th Conference on Reachability Problems, Lecture Notes in Computer Science, Vol. 15050, pp. 88–103, Springer, Berlin, 2024.
- A. Benso, F. D'Alessandro, and P. Papi. On the Intersection Problem for Quantum Finite Automata, *Theoretical Computer Science*, Vol. 1053, 115454, 2025.

Contents

Introduction

Quantum finite automata (QFAs in short) have been introduced in the second half of '90s as a new model of language recognizer that, essentially, impose the quantum paradigm on classical finite automata. Thereafter, many papers proposed various models of QFAs—which basically change the types of measurements of the quantum state—and investigated their comparison with older models of computation. For example, some undecidable problems for probabilistic finite automata (PFAs) become decidable for QFAs.

In this regard, we will address one of the main issues in the mathematical theory of formal languages, that is the decidability of the *Emptiness problem:* it asks whether, given a model of computation $\mathcal{M}$, the language $L(\mathcal{M})$ accepted by $\mathcal{M}$ is equal to the empty set. A natural generalization of the previous problem is the *Intersection problem*: given a predefined family of languages $\mathcal{L}$ and taking arbitrarily $L \in \mathcal{L}$, is the intersection of L with $L(\mathcal{M})$ empty? Taking as $\mathcal{L}$ the family given by the sole free monoid as its element, one immediately obtains the Emptiness problem.

In this thesis, it is investigated the decidability of the Intersection problem, taking, as model of computation, the QFAs. In particular, we consider the model introduced by Moore and Crutchfield in [38], in which there is a sole measurement at the end of the computation, called for this reason also "*measure once*". The original material of this thesis has been obtained in a joint work with F. D'Alessandro and P. Papi [7, 8].

We now give a formal description of the problem. Let Σ be an alphabet and denote by Σ^* the free monoid generated by Σ. A *word* (over Σ) is any element of Σ^*. A *quantum finite automaton* (in the sense of [38]) is a quadruple

$$\mathcal{Q} = \langle s, \varphi, P, \lambda \rangle \,, \tag{1}$$

where $s \in \mathbb{R}^n$ is a row-vector of unit Euclidean norm, $\varphi : \Sigma^* \to O_n$ is a morphism from the free monoid generated by the input alphabet Σ of $\mathcal{Q}$ into

the group O_n of orthogonal $n \times n$-matrices over $\mathbb{R}$, P is a $n \times n$ orthogonal projection matrix and λ is a value in $\mathbb{R}$ called *threshold*. The languages recognized by $\mathcal{Q}$ with strict and nonstrict threshold λ are respectively defined as

$$|\mathcal{Q}_>| = \{w \in \Sigma^* \mid ||s\varphi(w)P||^2 > \lambda\}, \quad |\mathcal{Q}_\geq| = \{w \in \Sigma^* \mid ||s\varphi(w)P||^2 \geq \lambda\},$$

where $||\cdot||$ denotes the Euclidean norm of vectors. Since we are interested in effective properties which require the QFA to be effectively given, we consider the model of *rational quantum automaton*, i.e., where all the coefficients of the components of (1) are rational numbers. Let $\mathcal{L}$ be a given family of languages. The problem we tackle is the following:

$(L, \mathcal{Q})$ Intersection problem

INPUT: a language L in a family $\mathcal{L}$ of languages and a rational quantum automaton $\mathcal{Q}$.

QUESTION: does $L \cap |\mathcal{Q}_>| = \varnothing$ hold?

A first result has been proven by Blondel, Jeandel, Koiran, and Portier ([13]), who showed the decidability of the Emptiness problem formulated for the languages with strict threshold recognized by measure once QFAs. This result contrasts with the corresponding one for PFAs. Indeed, given a PFA, the Emptiness problem (appropriately reformulated for this class of automata) is undecidable (see [42, Theorem 6.17]). It also contrasts with the undecidability of the Emptiness problem for the languages with nonstrict threshold recognized by measure once QFAs ([13]). It has also been proven ([33]) that both the problems (strict and nonstrict) remain undecidable for the "*measure many*" QFAs, a model not computationally equivalent to the measure once, introduced by Kondacs and Watrous in [34]. Later, Bertoni, Choffrut and D'Alessandro ([11, 12]), showed that the Intersection problem is decidable for the families of linear and bounded semilinear languages. The aim of this thesis is to continue this investigation, by providing new conditions that make the Intersection problem decidable.

The main contributions are the following. In Section 4.1, we extend the results given in [12] by showing the decidability of the Intersection problem for languages generated by restricted matrix grammar, a remarkable subfamily of matrix context-free grammars of finite index. Precisely, since both linear and

bounded semilinear languages can be generated by such grammars, we prove that whenever the QFA is rational, Proposition 4.1 allows one to recover both the previous two cases as corollaries; moreover, it shows the decidability of the problem also for not context-free languages.

In Section 4.2, we investigate the Intersection problem for languages defined by *finite index context-free grammars*. These are grammars G where each word w in the language generated by G is obtained by some derivation δ, whose index is uniformly bounded, i.e., the number of variables in each sentential form of δ is bounded by an integer not depending on w. A remarkable result by Ginsburg and Spanier [24] (see also Nivat [39]) provides a characterization of such languages in terms of composition of grammars: precisely, each language of this type is generated by a grammar G given by the composition

$$G = \mathcal{G}_1 \circ \mathcal{G}_2 \circ \cdots \circ \mathcal{G}_k, \tag{2}$$

of families $\mathcal{G}_i,\ 1 \leq i \leq k,$ of linear context-free grammars. Here, we prove a statement of decidability for a subclass of these grammars called *monoidal*. A finite-index grammar G is monoidal if the grammars of (2) are minimal linear, and the terminal productions of G are of the form $X \to \varepsilon$ (see Definition 1.19).

We prove that, regardless of the length k of the composition (2) for G, if the Zariski closure of suitable monoids—called *monoids of cycles*—associated to the linear grammars of the lowest level $\mathcal{G}_k$ of (2), are (algebraic) irreducible, then one can effectively compute the Euclidean closure $\mathbf{C}l(\varphi(L))$ of $\varphi(L)$ (Proposition 4.6, Corollary 4.11).

We now find useful to provide a synthetic description of the proofs. Following the same strategy of [13], we first observe that the Intersection problem is equivalent to the inclusion problem

$$L \subseteq |\mathcal{Q}_{\leq}|. \tag{3}$$

Now, since the function

$$M \to ||sMP||^2,$$

—where M is an arbitrary matrix in $\mathbb{R}^{n\times n}$, and s, P are components of (1)— is continuous, it is sufficient to prove that, for all matrices M in the Euclidean closure $\mathbf{C}l(\varphi(L))$ of $\varphi(L)$, the condition

$$\|sMP\|^2 \leq \lambda$$

holds (see Section 3.2). Such condition is then effectively tested by expressing the property (3) in first-order logic of the field of reals, which, in turn, consists of effectively computing a representation of $\mathbf{C}l(\varphi(L))$ in terms of semialgebraic sets (a family of sets more general than algebraic, closed under the operation of product of sets). Afterwards, one applies to the constructed formula the Tarski-Seidenberg quantifier elimination to verify whether property (3) holds true or not.

In the proof of Proposition 4.1, given a language L generated by a restricted matrix grammar G, the semialgebraicity of $\mathbf{C}l(\varphi(L))$ and its effective computation is reduced to those of an algebro-combinatorial object associated with G and each of its variable A: the aforementioned *monoid of cycles of* A. Such monoid, denoted M_A, corresponds to the matrix image, under the morphism φ of (1), of the language of cycles in G associated with A (see Equation (3.6)). In the proof, we use a suitable decomposition of the derivations in these grammars to show that, for every variable A of G, M_A is a regular submonoid; from this, we are able to compute the effective representation of $\mathbf{C}l(\varphi(L))$ by applying Proposition 3.10 and an appropriate algorithm developed by Derksen, Jeandel, and Koiran in [19].

In the proof of Proposition 4.6, given a language L generated by a monoidal grammar G, we achieve the effective computation of $\mathbf{C}l(\varphi(L))$ by considering the property of (algebraic) irreducibility on the Zariski closure of a family of monoids of cycles canonically associated with G.

We finally describe the structure of this thesis. In Chapter 1 definitions and preliminary results in Formal Language Theory are presented; in particular, in Section 1.1 the concept of formal language is introduced, in Section 1.2 a useful combinatorial structure related to context-free languages is presented, and in Section 1.3 the families of languages on which we will state our main results are presented. In Chapter 2, in Section 2.1, topological notions concerning the matrices over the field of real numbers are introduced, then, in Sections 2.2, 2.3, definitions and properties of semialgebraic and irreducible algebraic sets are presented. In Chapter 3, in Section 3.1 the model of quantum finite automata is introduced and formally described; in Section 3.2, the problem is presented and an effective condition to which can be reduced its decidability is recalled; finally, its decidability for the case in which the language is the free monoid generated by the input alphabet is claimed and the case in which the language is generated by a linear context-free grammar is demonstrated.

In Chapter 4, the main results are presented. In Section 4.1, the case in which the language is generated by a restricted matrix context-free grammar is presented and implication of the previous results is discussed; in Section 4.2, the case in which the language is generated by a monoidal language is demonstrated and, in Section 4.3, special cases are studied and some basic points of the proofs on interesting examples are shown. Finally, concluding remarks and open problems are discussed.

Chapter 1

Definitions and preliminary results

In this first chapter we will present definitions and preliminary results concerning Formal Language Theory, with particular interest in combinatorial structures such as "*finite-state automata*" and "*context-free grammars*" that can generate them. We will then discuss the structure of particular classes of grammars useful to define classes of languages such as finite index languages, monoidal languages and finite index matrix languages. Some of the most interesting families of languages are introduced. For this chapter, we will refer to [9, 12, 17, 18, 20, 23, 25, 32, 43].

1.1 Preliminaries on formal languages

An alphabet A is a finite nonempty set whose elements are called *letters* or *symbols*. A *word* or *string* over the alphabet A is a finite sequence of elements of A. We denote by A^+ the set of all words over the alphabet A, i.e.

$$A^+ = \{a_1 \cdots a_n \,|\, n \in \mathbb{N}_+,\, a_i \in A\},$$

where $\mathbb{N}_+$ denotes the set of positive integers. Define on A^+ the *concatenation product* of words in the following way: for each $u, v \in A^+$, if $u = a_1 \cdots a_n$ and $v = b_1 \cdots b_m$, with $a_i, b_j \in A$, then

$$u \cdot v = uv = a_1 \cdots a_n b_1 \cdots b_m.$$

Let A^* be the following set:

$$A^* = A^+ \cup \{\varepsilon\},$$

where ε is an element not in A called *empty word*. Extend the concatenation product to A^* in the following way:

$$\forall\, u \in A^* \quad u \cdot \varepsilon = \varepsilon \cdot u = u.$$

It is straightforward to check that $(A^*, \cdot)$ is a monoid with identity ε. More precisely, A^* is the free monoid of base A, since A generates A^* and every element of A^* other than the empty word admits a unique factorization starting from the elements of A. The *length* of a word $w \in A^*$ is the integer $|w|$ inductively defined by $|\varepsilon| = 0$, $|wa| = |w| + 1$, $w \in A^*$, $a \in A$. For every $w \in A^*$ and $a \in A$, $|w|_a$ denotes the number of occurrences of the letter a in w. A word $u \in A^*$ is a *factor* of $w \in A^*$ if there exist $p, q \in A^*$ such that $w = puq$. If $w = uq$, for some $q \in A^*$ (resp. $w = pu$ for some $p \in A^*$), then u is called a *prefix* (resp. a *suffix*) of w. A *formal language*, or simply *language*, over the alphabet A is any subset $L \subseteq A^*$. The *rational operations* between languages over A, that are the *union* ($\cup$), the *product* ($\cdot$), and the *star* (*), are defined as follows. Let $L_1, L_2 \subseteq A^*$ be two languages over A. The union $L_1 \cup L_2$ of L_1 and L_2 is

$$L_1 \cup L_2 = \{w \in A^* \,|\, w \in L_1 \vee w \in L_2\}.$$

The product of L_1 and L_2, denoted by $L_1 \cdot L_2$ (or simply $L_1 L_2$), extends the concatenation product of words

$$L_1 \cdot L_2 = \{uv \,|\, u \in L_1,\ v \in L_2\}.$$

From the latter we obtain the definition of the *square* of a language L: $L^2 = L \cdot L$ and, by induction on n, of the nth power of L:

$$\forall\, n \geq 0, \quad L^{n+1} = L^n \cdot L = L \cdot L^n,$$

that is completed by setting $L^0 = \{\varepsilon\}$. By combining the operations of union and product, the star of a language L, denoted by L^*, is the union of all the powers of L:

$$L^* = \bigcup_{n \geq 0} L^n.$$

A language over A is said to be *rational* (or *regular*) if it is finite or it is obtained from finite sets by using finitely many times the rational operations. The family of rational languages over A is denoted by $\mathrm{Rat}(A)$.

To define a language L over a given alphabet A several approaches can be used. In the following we will show how it is possible to define a language through formal models. In particular, we will describe the languages defined by a *finite state automaton* by an *M-automaton* (for a given monoid M), and by *congruences* between words. Later, in Section 1.2, we will explain how other combinatorial structures called *grammars* can generate languages; for the latter, we will focus on the well known *context-free grammars*.

Definition 1.1. A finite (deterministic and complete) automaton $\mathcal{A}$ is a tuple $\langle A, S, \delta, s_0, S' \rangle$ consisting of an alphabet A, a finite set S of *states*, an *initial state* $s_0 \in S$, a set of *final states* $S' \subseteq S$ and a *transition total function* $\delta : S \times A \to S$.

The transition function is extended from $S \times A$ to $S \times A^*$ considering the map $\hat{\delta} : S \times A^* \to S$ defined as: for any $s \in S$, $u \in A^*$, $a \in A$

$$\begin{aligned} &\hat{\delta}(s, \varepsilon) = s, \\ &\hat{\delta}(s, a) = \delta(s, a), \\ &\hat{\delta}(s, ua) = \delta(\hat{\delta}(s, u), a). \end{aligned} \tag{1.1}$$

For simplicity, in the following we will identify $\hat{\delta}$ with δ. From equations (1.1), one may easily see that

$$\delta(s, uv) = \delta(\delta(s, u), v), \quad u, v \in A^*.$$

A word $u \in A^*$ is said to be *recognized* or *accepted* by the automaton $\mathcal{A}$ if and only if $\delta(s_0, u) \in S'$. The *language recognized* by $\mathcal{A}$ is then

$$L_{\mathcal{A}} = \{u \in A^* \mid \delta(s_0, u) \in S'\}.$$

Incidentally, we recall that one of the most important results in the theory, proven by Stephen Cole Kleene in the '50s, provides a characterization of rational languages in relation to finite state automata. This theorem states that a language is rational if and only if there exists a finite state automaton that recognizes it.

The previous argument can be stated also for monoids instead of alphabets. The definition of automata over a monoid M is very similar to that of automata over an alphabet. Formally, an automaton $\mathcal{A}_M$ over the monoid M is a tuple $\langle M, S, E, s_0, S' \rangle$, where S, s_0 and S' are defined as in Definition

1.1 and E is the set of *transitions*. Each transition is a triple (s, m, t), where $s, t \in S$ and $m \in M$; hence, E is a subset of $S \times M \times S$. In this sense, an automaton over a monoid can be described as a labeled graph: the set of nodes is the set of states S and a node s is connected to a node t, through an oriented arc from s to t labeled with an element $m \in M$ if and only if $(s, m, t) \in E$. For this reason, a transition (s, m, t) is also denoted by $s \xrightarrow{m} t$.

A *computation* is a path in the graph $\mathcal{A}_M$; it is *successful* if its source is the initial state s_0 and its destination is in S'. The *label* of a computation c in $\mathcal{A}_M$

$$c := s = s_1 \xrightarrow{m_1} s_2 \xrightarrow{m_2} \cdots s_n \xrightarrow{m_n} s_{n+1} = t,$$

is the product of the successive transition labels of $c : m_1 \cdots m_n$ and we write $c = s \xrightarrow{m_1 \cdots m_n} t$. The language recognized by $\mathcal{A}_M$ is then

$$L_{\mathcal{A}_M} = \{m \in M \mid \exists\, t \in S',\ s_0 \xrightarrow{m} t\}.$$

The definition of M-automaton is thus consistent with the previous Definition 1.1: an automaton over the alphabet A is an automaton over the monoid A^*. However, there is a difference: the set of labels of an automaton over a monoid is potentially infinite, which leads us to the following definition.

Definition 1.2. An automaton $\mathcal{A}_M = \langle M, S, E, s_0, S' \rangle$ over a monoid M is *finite* if the set of transitions E is finite.

It is natural to define the rational subsets of a monoid. Let M be a monoid. The family $\mathrm{Rat}(M)$ of *rational subsets of* M is the least family of subsets of M closed under the rational operations. In Chapters 3 and 4 we will use the following result (stated for the first time by Elgot and Mezei in the mid of sixties in [20]); similarly to the Kleene's Theorem, it gives a characterization of rational subsets of a monoid.

Theorem 1.3. *A subset of a monoid M is rational if and only if it is recognized by a finite automaton over M.*

We end this section introducing one of the most interesting class of formal languages, the so called *Dyck languages*. A Dyck language consists of "well-formed" words over a finite number of pairs of parentheses. The *restricted Dyck languages* $D_1'^*$ is formed by the words of parentheses which are "correct" in the usual sense. For example,

$$(()(()))$$

is a word of $D_1'^*$. For the *Dyck language* D_1^*, the interpretation of the parentheses is different. Two parentheses of the same type are rather considered as formal inverses for each other. A word is considered *correct* if and only if successive deletion of factors of associated parentheses yields the empty word. In this sense, an alternative way to define Dyck languages is to code a parenthesis as a letter a and its inverse parenthesis as $\bar{a}$. Thus

$$\bar{a}a\bar{a}aa\bar{a}$$

is a word of D_1^*.

Define now Dyck languages $D_n'^*$ and D_n^* for $n \geq 1$. Let $A_n = \{a_1, \ldots, a_n\}$ and $\bar{A}_n = \{\bar{a}_1, \ldots, \bar{a}_n\}$ be two alphabets of n letters. Each pair $a_k, \bar{a}_k$ can be considered as a pair of parentheses of the same type. Define $C_n := A_n \cup \bar{A}_n$. For $c \in C_n$, denote

$$\bar{c} = \begin{cases} \bar{a}_k & \text{if } c = a_k, \\ a_k & \text{if } c = \bar{a}_k. \end{cases}$$

Thus, $\bar{\bar{c}} = c$.

Definition 1.4. The *restricted Dyck congruence* $\sim_n'$ is the congruence of C_n^* generated by

$$a_k\bar{a}_k \sim \varepsilon \quad k = 1, \ldots, n. \tag{1.2}$$

The *Dyck congruence* $\sim_n$ is the congruence generated by Congruence (1.2) and by

$$\bar{a}_k a_k \sim \varepsilon \quad k = 1, \ldots, n. \tag{1.3}$$

Thus two words $w, w' \in C_n^*$ are congruent modulo $\sim_n'$ (resp. modulo $\sim_n$), and we write

$$w \equiv w' (\mathrm{mod}\ \sim_n') \quad (\text{resp. } w \equiv w' (\mathrm{mod}\ \sim_n))$$

if and only if w' can be obtained from w by a finite number of insertions or deletions of factors of the form $a_k\bar{a}_k$ (resp. $a_k\bar{a}_k$ or $\bar{a}_k a_k$).

Definition 1.5. The *restricted Dyck language* $D_n'^*$ is the class of ε in the congruence $\sim_n'$: $D_n'^* = [\varepsilon]_{\sim_n'}$. The *Dyck language* D_n^* is the class of ε in the congruence $\sim_n$: $D_n^* = [\varepsilon]_{\sim_n}$.

Example 1.6. Consider $n = 1$ and $C_1 = A_1 \cup \bar{A}_1 = \{a_1\} \cup \{\bar{a}_1\}$. In the following, when we will refer to C_1 we will write simply a instead of a_1 and b instead of $\bar{a}_1$. By the previous definition, the restricted Dyck language $D_1'^*$ and the Dyck language D_1^* can be written respectively as

$$D_1'^* = \{w \in C_1^* \mid |w|_a = |w|_b \text{ and } \forall\, u \text{ prefix of } w,\ |u|_a \geq |u|_b\},$$

$$D_1^* = \{w \in C_1^* \mid |w|_a = |w|_b\}.$$

1.2 Context-free languages

In this section we first define the family of *context-free languages*, denoted by CFL, and then discuss a combinatorial structuring of these languages useful for the proofs of main results. In order to define context-free languages, we first define context-free grammars. A context-free grammar is a quadruple $G = \langle V, \Sigma, P, S\rangle$, where:

- V is a finite set of objects called *variables* or *non-terminal symbols* of the grammar G; usually, the nonterminal symbols are denoted by uppercase letters as, for instance, A, B, etc.
- Σ is a finite set of symbols called *terminals* of the grammar G, moreover Σ is disjoint from V;
- P is a finite subset of $V \times (V \cup \Sigma)^*$, called the set of *productions* or *rules* of the grammar G; each production is usually denoted in the form $A \to \beta$, where A is a variable and β is a string of symbols of $(V \cup \Sigma)^*$;
- S is a variable called *start symbol* or *axiom* of the grammar.

To define the language generated by G, we first introduce the definition of two relations $\Rightarrow$ and $\overset{*}{\Rightarrow}$ between strings in $(V \cup \Sigma)^*$. If u, v are strings in $(V \cup \Sigma)^*$, then we set

$$u \Rightarrow v$$

if there exists a production $A \to \beta$ of G such that

$$u = \alpha A \gamma \quad \text{and} \quad v = \alpha\beta\gamma,$$

where $\alpha, \gamma \in (V \cup \Sigma)^*$. The words α, γ are called *contexts* of u and v, and the relation $\Rightarrow$ is called *atomic derivation*. Let $\overset{*}{\Rightarrow}$ be the relation defined as the

transitive and reflexive closure of $\Rightarrow$. This relation is called *derivation* of G. Equivalently, if $\alpha, \beta \in (V \cup \Sigma)^*$ are such that $\alpha \stackrel{*}{\Rightarrow} \beta$, then either $\alpha = \beta$ or there exist words $\alpha_0, \alpha_1, \ldots, \alpha_n$, $n \geq 1$, in $(V \cup \Sigma)^*$ such that

$$\alpha = \alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_n = \beta.$$

Finally, a string $\alpha \in (V \cup \Sigma)^*$ is said to be a *sentential form* of G if $S \stackrel{*}{\Rightarrow} \alpha$. The language $L(G)$ generated by the grammar G is the set of sentential forms of G with terminal symbols only, i.e.

$$L(G) = \{w \in \Sigma^* \mid S \stackrel{*}{\Rightarrow} w\}.$$

We will say that a language is *context-free* if there exists a context-free grammar that generates it. It is well known that rational languages can be generated by context-free grammars, indeed it can be shown that they are generated by so called *regular grammars*. A context-free grammar is said to be *right-linear* (resp. *left-linear*) if each of its productions $A \to \beta$ is such that $\beta \in (\Sigma^* V \cup \Sigma^*)$ (resp. $\beta \in (V\Sigma^* \cup \Sigma^*)$) and *regular* if it is right-linear or left-linear. Moreover, if each production of the grammar $A \to \beta$ has at most one occurrence of nonterminal symbols in the right hand side, i.e. $\beta \in (\Sigma^* V \Sigma^* \cup \Sigma^*)$, the grammar is said to be *linear*. The family of languages generated by linear grammars is denoted by LIN and its elements are called *linear* languages. Finally, if the grammar has a unique variable, it is said to be *minimal*.

Example 1.7. A first example of context-free languages are the Dyck languages defined in the previous section. Indeed, it can be shown that the restricted Dyck language D'^*_n, with $n \geq 1$, is generated by the grammar $G = \langle V, C_n, P, S \rangle$, where the productions of the grammar are

$$\begin{aligned} &S \to SS, \\ &S \to a_k S \bar{a}_k, \quad k = 1, \ldots, n, \\ &S \to \varepsilon. \end{aligned}$$

Similarly, the Dyck language D^*_n is generated by the grammar with produc-

tions

$$\begin{aligned} &S \to SS, \\ &S \to a_k S \bar{a}_k, \quad k = 1, \ldots, n, \\ &S \to \bar{a}_k S a_k, \quad k = 1, \ldots, n, \\ &S \to \varepsilon. \end{aligned}$$

A combinatorial structuring of context-free languages

Now, we recall a combinatorial structuring of an arbitrary context-free language based upon the notion of cycle introduced by Ginsburg and Spanier ([23]). Such decomposition follows a recursive scheme proposed by Incitti in [32] based upon the concept defined below. The idea of the following notation is to consider the set of all pairs of left and right contexts in the terminal alphabet of a self-embedding nonterminal symbol. In the next definition, the initial "C" is meant to suggest the term "cycle".

Definition 1.8. With each nonterminal symbol $A \in V$ associate the subset of $\Sigma^* \times \Sigma^*$ defined as

$$C_A = \{(\alpha, \beta) \in \Sigma^* \times \Sigma^* : A \stackrel{*}{\Rightarrow} \alpha A \beta\}.$$

C_A will be called the *set of cycles of* A. [1]

We define now the *sandwich* operation, denoted by $\diamond$, for languages. A similar operation will be defined for matrices in Chapter 2. If C_A is the set defined above and L' is an arbitrary language over Σ, we define the subset $C_A \diamond L'$ of Σ^* as

$$C_A \diamond L' = \{uwv \mid (u, v) \in C_A \text{ and } w \in L'\}. \tag{1.4}$$

As the construction of the structuring mentioned above proceeds by induction on the number of nonterminal symbols, we need to show how to recombine a grammar from simpler ones obtained by choosing an arbitrary non-axiom variable as the new axiom and by removing all the rules involving S. This is the reason for introducing the next notation.

[1] According to the terminology of [23], a *cycle* is a derivation of the form $A \stackrel{*}{\Rightarrow} \alpha A \beta$. With a minor abuse of language, we will call cycle the pair of contexts (α, β) too.

Definition 1.9. Let $G = \langle V, \Sigma, P, S\rangle$ be a context-free grammar. Let us define the set V' as $V' = V \setminus \{S\}$. For every $A \in V'$, define the context-free grammar $G_A = \langle V', \Sigma, P_A, A\rangle$, where

- V' is the set of variables of G_A,
- the set P_A consists of all the rules $B \to \gamma$ of G of the form

$$B \in V', \quad \gamma \in (V' \cup \Sigma)^*,$$

- A is the axiom of the grammar.

The language of all terminal words generated by the grammar G_A is denoted by L_A.

The next definition introduces the language of terminal words obtained in a derivation where S occurs at the start only.

Definition 1.10. Let $L'(G)$ denote the set of all the words of Σ^* which admit a derivation

$$S \Rightarrow \gamma_1 \Rightarrow \cdots \Rightarrow \gamma_\ell \Rightarrow w \tag{1.5}$$

where, for every $i = 1, \ldots, \ell$, $\gamma_i \in (V' \cup \Sigma)^*$.

If no ambiguity arises, in the sequel, the language $L'(G)$ is simply denoted L'. The language L' can be easily expressed in terms of the languages L_A for all $A \in V'$. Indeed, define $\mathcal{B}$ as the set of all strings $\beta \in (V' \cup \Sigma)^*$ such that

$$S \to \beta, \quad \beta \in (V' \cup \Sigma)^* \tag{1.6}$$

is a rule of the grammar G. Let us enumerate the elements of $\mathcal{B}$ as

$$\beta_1, \ldots, \beta_t \tag{1.7}$$

and, for every $j = 1, \ldots, t$, factorize β_j as

$$\beta_j = w_{j,1}A_{j,1}w_{j,2}A_{j,2}\cdots w_{j,\ell_j}A_{j,\ell_j}w_{j,\ell_j+1} \tag{1.8}$$

where $w_{j,1}, \ldots, w_{j,\ell_j+1} \in \Sigma^*$ and $A_{j,1}, A_{j,2}, \ldots, A_{j,\ell_j} \in V'$.

Lemma 1.11. *The language L' is the (finite) union of the languages of the form*

$$w_{j,1}L_{A_{j,1}}w_{j,2}L_{A_{j,2}}\cdots w_{j,\ell_j}L_{A_{j,\ell_j}}w_{j,\ell_j+1}, \tag{1.9}$$

where, $\beta_j = w_{j,1}A_{j,1}w_{j,2}A_{j,2}\cdots w_{j,\ell_j}A_{j,\ell_j}w_{j,\ell_j+1} \in \mathcal{B}$.

Proof. Let us consider a word $w \in L'$. Hence, according to (1.5), we have

$$S \Rightarrow \gamma_1 \Rightarrow \cdots \Rightarrow \gamma_\ell \Rightarrow w,$$

where, for every $i = 1, \ldots, \ell$, $\gamma_i \in (V' \cup \Sigma)^*$. Since γ_1 is a string of the set (1.7), there exists some j, with $1 \leq j \leq t$, such that

$$\gamma_1 = \beta_j = w_{j,1} A_{j,1} w_{j,2} A_{j,2} \cdots w_{j,\ell_j} A_{j,\ell_j} w_{j,\ell_j+1},$$

so that

$$w \in w_{j,1} L_{A_{j,1}} w_{j,2} L_{A_{j,2}} \cdots w_{j,\ell_j} L_{A_{j,\ell_j}} w_{j,\ell_j+1}.$$

Suppose now that w is a word of a language of the form (1.9), that is

$$w \in w_{j,1} L_{A_{j,1}} w_{j,2} L_{A_{j,2}} \cdots w_{j,\ell_j} L_{A_{j,\ell_j}} w_{j,\ell_j+1},$$

where, for every $s = 1, \ldots, \ell_j + 1$, $w_{j,s}$ are the words of Σ^* defined in (1.8). Hence w factorizes as

$$w = w_{j,1} \chi_{j,1} w_{j,2} \chi_{j,2} \cdots w_{j,\ell_j} \chi_{j,\ell_j} w_{j,\ell_j+1},$$

where, for each $s = 1, \ldots, \ell_j$, $\chi_{j,s} \in L_{A_{j,s}}$. Recall now that, according to the definition of the words $w_{j,s}$,

$$S \to \beta_j = w_{j,1} A_{j,1} w_{j,2} A_{j,2} \cdots w_{j,\ell_j} A_{j,\ell_j} w_{j,\ell_j+1}$$

is a production of the grammar G. On the other hand, since, for every $s = 1, \ldots, \ell_j$, $\chi_{j,s} \in L_{A_{j,s}}$, we have

$$A_{j,s} \stackrel{*}{\Rightarrow} \chi_{j,s} \tag{1.10}$$

where, by definition of the grammar $G_{A_{j,s}}$, all the sentential forms of the derivation do not have the variable S as a factor. This implies that the derivation

$$S \Rightarrow w_{j,1} A_{j,1} w_{j,2} A_{j,2} \cdots w_{j,\ell_j} A_{j,\ell_j} w_{j,\ell_j+1} \stackrel{*}{\Rightarrow} w$$

is a derivation of G of the form (1.5), and so $w \in L'$.

□

Proposition 1.12. *The language L is a finite union of languages of the form $C_S \diamond L''$ where*

$$L'' = w_1 L_{A_1} w_2 L_{A_2} \cdots w_\ell L_{A_\ell} w_{\ell+1},$$

and $\beta = w_1 A_1 w_2 A_2 \cdots w_\ell A_\ell w_{\ell+1}$ ranges over all rules in $\mathcal{B}$, with respect to the notation (1.6) and (1.8).

Proof. In order to prove that the language $C_S \diamond L''$ is included in L, it suffices to consider a word $w = \alpha u \beta$, with $u \in L'$ and $(\alpha, \beta) \in C_S$. One has

$$S \stackrel{*}{\Rightarrow} u \quad \text{and} \quad S \Rightarrow \gamma_1 \Rightarrow \cdots \Rightarrow \gamma_\ell = \alpha S \beta.$$

By gluing the two derivations in the right order one has

$$S \Rightarrow \gamma_1 \Rightarrow \cdots \Rightarrow \gamma_\ell = \alpha S \beta \stackrel{*}{\Rightarrow} \alpha u \beta.$$

Thus, $w \in L$. Let us prove the opposite inclusion. Let $w \in L$. Hence w admits a derivation of the form

$$S \Rightarrow \gamma_1 \Rightarrow \cdots \Rightarrow \gamma_\ell \Rightarrow w, \tag{1.11}$$

where $\ell \geq 0$ and, for every $i = 1, \ldots, \ell$, $\gamma_i \in (V \cup \Sigma)^*$. If the symbol S does not occur in the derivation except at the start of derivation, then $w \in L'$. Assume now that S appears in some sentential form γ_i, with $1 \leq i \leq \ell$. Thus $\gamma_i = \alpha_i S \beta_i$, with $\alpha_i, \beta_i \in (V \cup \Sigma)^*$. Hence the derivation (1.11) can be rewritten as

$$S \stackrel{*}{\Rightarrow} \alpha_i S \beta_i \Rightarrow \gamma_{i+1} \Rightarrow \cdots \Rightarrow \gamma_\ell \Rightarrow w. \tag{1.12}$$

By changing the order of the derivation, we can assume that $\alpha_i, \beta_i \in \Sigma^*$.

Let us now remark that, in derivation (1.12), we can assume that, for every $j = i+1, \ldots, \ell$, S is not a factor of γ_j, that is, the derivation that transforms the sentential form γ_i into the word w is such that, for every $j = i+1, \ldots, \ell$, S is not a factor of γ_j.

Indeed, otherwise, we would have

$$\gamma_i = \alpha_i S \beta_i \stackrel{*}{\Rightarrow} \gamma_j = \alpha_i \hat{\alpha}_j S \hat{\beta}_j \beta_i \stackrel{*}{\Rightarrow} w.$$

By replacing the ordered pair (α_i, β_i) with $(\alpha_i \hat{\alpha}_j, \hat{\beta}_j \beta_i)$, we obtain the claimed factorization for (1.12). Therefore, the previous two arguments imply that

$$w = \alpha_i u \beta_i,$$

where $(\alpha_i, \beta_i) \in C_S$ and u is a word generated by a derivation of the form

$$S \Rightarrow \gamma'_{i+1} \Rightarrow \cdots \Rightarrow \gamma'_\ell \Rightarrow u,$$

where, for every $j = i+1, \ldots, \ell$, S is not a factor of γ'_j and thus $u \in L'$. Hence $w \in C_S \diamond L''$. This completes the proof.

□

1.3 Bounded semilinear languages, matrix and monoidal languages

In this section, we introduce three families of languages: *bounded semilinear languages*, *finite index matrix languages* and *monoidal languages*. We will state our main results for these non-trivial families of languages in Chapter 4.

Bounded semilinear languages

The first non-trivial family we introduce is the family of *bounded semilinear languages*. A language L over an alphabet A is bounded semilinear if it is finite union of languages of the form

$$L = \{u_1^{n_1} \cdots u_k^{n_k} \,|\, (n_1, \ldots, n_k) \in R\}, \tag{1.13}$$

where $u_1, \ldots, u_k$ are fixed words over the alphabet A and $R \subseteq \mathbb{N}^k$ is a *linear* set, i.e., there exist $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_\ell \in \mathbb{N}^k$ such that

$$R = \{\mathbf{v}_0 + \lambda_1 \mathbf{v}_1 + \cdots + \lambda_\ell \mathbf{v}_\ell \,|\, \lambda_1, \ldots, \lambda_\ell \in \mathbb{N}\}. \tag{1.14}$$

Example 1.13. Let $G = \langle V, \Sigma, P, S\rangle$ be a context-free grammar where:

$$V = \{A\}, \qquad \Sigma = \{a, b\},$$
$$S = A, \qquad P = \{A \to aAb,\ A \to \varepsilon\}.$$

It is easily checked that the language generated by G is:

$$L(G) = \{a^n b^n \,|\, n \in \mathbb{N}\}.$$

Moreover, it is clear that L is a bounded semilinear language; indeed, $L = \{a^{n_1} b^{n_2} \,|\, (n_1, n_2) \in R\}$, where $R = \{\lambda_1(1, 1) \,|\, \lambda_1 \in \mathbb{N}\}$.

In order to properly define the other two families, we introduce the classes of grammars that generate them respectively: *restricted matrix grammars* and *finite index monoidal grammars*.

Matrix context-free languages

To define finite index matrix grammars we first recall the *matrix context-free grammars* (which we will henceforth simply call *matrix grammars*). Following [17], Ch. 1, Sec. 1.1, a matrix context-free grammar is a tuple $G = \langle V, \Sigma, M, S\rangle$, where V, Σ, and M are respectively the finite sets of nonterminals, terminals, and matrices (of rules), and $S \in V$ is the start symbol. Each *matrix rule* $m \in M$ is a finite sequence $m = (p_1, \ldots, p_s)$ where each p_i, $1 \leq i \leq s$, is a context-free production from V to $(V \cup \Sigma)^*$. We denote by P_M the set of all the context-free productions that appear in the matrices of M. For an arbitrary $m \in M$, we define the 1-*step derivation in* G, $x \Rightarrow_m y$, with $x, y \in (V \cup \Sigma)^*$ if

$$x = x_0 \underset{p_1}{\Rightarrow} x_1 \underset{p_2}{\Rightarrow} \cdots \underset{p_s}{\Rightarrow} x_s = y,$$

where $\Rightarrow_{p_i}$ denotes the standard context-free derivation relation, defined by the rule p_i. This is equivalent to say that $x = x_0, x_s = y$, and, for every $i = 1, \ldots, s$,

$$x_{i-1} = x'_{i-1} X_i x''_{i-1}, \quad x_i = x'_{i-1} \alpha_i x''_{i-1},$$

where $p_i = (X_i \to \alpha_i)$, and for some $x'_{i-1}, x''_{i-1} \in (V \cup \Sigma)^*$. In other words, a 1-step derivation in the matrix grammar G corresponds to a s-step derivation in the context-free grammar $\langle V, \Sigma, P_M, S\rangle$, i.e., using, one by one, all the context-free productions of m (w.r.t. the order the p_i's appear in m). Let M^* be the set of the finite sequences of matrices and $\alpha \in M^*$. If $\alpha = \varepsilon$, then $x \Rightarrow_\alpha x, x \in V^*$; if one has

$$\alpha = m_1 m_2 \cdots m_s \quad \text{and} \quad x_0 \underset{m_1}{\Rightarrow} x_1 \underset{m_2}{\Rightarrow} \cdots \; x_{s-1} \underset{m_s}{\Rightarrow} x_s,$$

with $x_i \in (V \cup \Sigma)^*$, $m_j \in M$, $1 \leq j \leq s$, $0 \leq i \leq s$, then we write $x_0 \Rightarrow_\alpha x_s$. The *language generated* by G is $L(G) = \{w \in \Sigma^* \mid S \Rightarrow_\alpha w\}$.

The following definition is instrumental. It corresponds to a specific case of the *simple matrix grammars* introduced by Ibarra in [28] (see also [31]).

Definition 1.14. A matrix grammar $G = \langle V, \Sigma, M, S\rangle$ is said to be *restricted of index* k, or simply *restricted*, if the set of nonterminals of G has the form $V = \{S\} \cup V'$, where $S \notin V'$ and $V' = V_1 \cup \cdots \cup V_k$, where $V_i \cap V_j = \varnothing$, for all $i \neq j$, and a matrix of rules of G can be only of one of the following forms

- $(X_1 \to u_1 Y_1 v_1, \ldots, X_k \to u_k Y_k v_k)$, for some $X_i, Y_i \in V_i,\ u_i, v_i \in \Sigma^*$, with $1 \leq i \leq k$;
- $(X_1 \to \varepsilon, \ldots, X_k \to \varepsilon)$, with $X_i \in V_i$, $1 \leq i \leq k$;
- $(S \to X_1 \cdots X_k)$, for some $X_i \in V_i$, with $1 \leq i \leq k$.

A languages is said to be a *restricted matrix language*, or simply a *restricted language*, if it is generated by a restricted matrix grammar.

Example 1.15. Let $L = \{a^n b^n c^n : n \in \mathbb{N}\}$ be the language over the alphabet $\Sigma = \{a, b, c\}$. Let G be the restricted matrix grammar with the matrices of rules

$$\begin{aligned} m_1 &= (S \to ABC), \\ m_2 &= (A \to aA, B \to bB, C \to cC), \\ m_3 &= (A \to \ \varepsilon, B \to \ \varepsilon, C \to \varepsilon). \end{aligned}$$

Note that the other components of G can be deduced from these matrices. One easily checks that $L = L(G)$. Indeed, each word $a^n b^n c^n, n \geq 0$, can be generated by using first the matrix m_1, then the matrix m_2 n times and finally the matrix m_3. Hence $L \subseteq L(G)$. The reverse inclusion is proved similarly.

Finite index languages

To define the finite index monoidal languages we first define a much more general family of languages: the *finite index languages*. Following [9], Sec. VII.5, we say that a context-free language L is *of index* k, if there exist a context-free grammar $G = \langle V, \Sigma, P, S \rangle$ with $L = L(G)$ and an integer $k \in \mathbb{N}$ such that the following property holds: for every word $w \in L(G)$, there exists some derivation $\delta = (S \overset{*}{\Rightarrow} w)$ such that the number of occurrences of variables in each sentential form of δ does not exceed k. The family of context-free languages of index k will be denoted by $Ind(k)$ and $\mathcal{L}(IND_{FIN})$ will denote the family of all the languages *of finite index*, that is, $\bigcup_{k \in \mathbb{N}} Ind(k)$. One immediately sees that the most simple languages in such a family are those of LIN, indeed, by definition, these are languages of index 1. A well-known result proven by Salomaa in [44] shows that Dyck languages cannot be

generated by finite index grammars, thus implying $\mathcal{L}(IND_{FIN}) \subset$ CFL. The importance of the family $\mathcal{L}(IND_{FIN})$ is due to the fact that, by a well-known result by Baron and Kuich [4] (up to a technical restriction) the unambiguous context-free languages whose characteristic series in commutative variables are rational, are exactly those in $\mathcal{L}(IND_{FIN})$.

Following Ginsburg and Spanier [24] (see also Nivat [39]), we recall a characterization of languages of $\mathcal{L}(IND_{FIN})$, that will be used in Chapter 4. To this purpose, we first recall that, given alphabets Σ_1 and Σ_2, and a family $\mathcal{F}$ of languages on Σ_2, a $\mathcal{F}$-*substitution* is a morphism $\theta : \Sigma_1^* \to \Pi(\Sigma_2^*)$ from the free monoid Σ_1^* into the multiplicative monoid $\Pi(\Sigma_2^*)$ of subsets of the free monoid Σ_2^* such that

$$\forall\, x \in \Sigma_1, \quad \theta(x) \in \mathcal{F}.$$

Let us define, recursively, the family $\{Qrt(k)\}_{k\in\mathbb{N}}$ of *quasi-rational languages of rank* k (or *bounded derivation languages of rank* k) as:

$$Qrt(k) = \begin{cases} LIN & \text{if } k = 1, \\ LIN \circ Qrt(k-1) & \text{if } k > 1, \end{cases} \tag{1.15}$$

where $LIN \circ Qrt(k-1)$ denotes the family of all the languages obtained as the images of linear languages, *via* $Qrt(k-1)$-substitutions, $k > 1$, i.e.

$$\{\theta(L) \,:\, L \in LIN, \quad \forall\, x \in \Sigma_1 \quad \theta(x) \in Qrt(k-1)\}.$$

The following holds ([24], Theorem 4.2, cf [9], Sec. VII.5, Theorem 5.2).

Theorem 1.16. *For any* $k \geq 1$, $Ind(k) = Qrt(k)$.

For our next use, we describe how quasi-rational languages can be characterized in terms of *composition of grammars*; in particular, the following argument shows that $Ind(2) = Qrt(2)$ (see Example 1.20 for an example). For this purpose, we recall the construction of the context-free grammar that generates the image $\theta(L)$ of a context-free language L over the alphabet Σ_1 under a *context-free substitution* $\theta : \Sigma_1^* \to \Pi(\Sigma_2^*)$, i.e., a $\mathcal{F}$-substitution, with $\mathcal{F}$=CFL.

We follow *verbatim* the argument of [9], Sec. II.2. Let $\mathcal{G}_1 = \langle V_1, \Sigma_1, P_1, S\rangle$ be the context-free grammar that generates L and let

$$\mathcal{G}_2 = \{G_x \,:\, x \in \Sigma_1\}$$

be the family of context-free grammars where, for every $x \in \Sigma_1$,

$$G_x = \langle V_x, \Sigma_2, P_x, S_x \rangle$$

is the grammar that generates $\theta(x)$. One may assume that $V_1 \cap V_x = \varnothing$, for all $x \in \Sigma_1$, and, for every $x, y \in \Sigma_1,\ x \neq y$, $V_x \cap V_y = \varnothing$. Let $\mathcal{V}_2$ be the set

$$\mathcal{V}_2 = \{S_x : x \in \Sigma_1\}$$

and let the copy morphism

$$\mathrm{c} : (V_1 \cup \Sigma_1)^* \to (V_1 \cup \mathcal{V}_2)^*, \tag{1.16}$$

defined as: $\mathrm{c}(z) = z$, if $z \in V_1$, and $\mathrm{c}(z) = S_z$, if $z \in \Sigma_1$. With respect to c, consider the grammar G_{c} that generates $\mathrm{c}(L)$. Such a grammar is defined as

$$G_{\mathrm{c}} = \langle V_1, \{S_x : x \in \Sigma_1\}, P_{\mathrm{c}}, S \rangle,$$

where $P_{\mathrm{c}} = \{A \to \mathrm{c}(\alpha) : (A \to \alpha) \in P_1\}$. Finally, define

$$H = \langle W, \Sigma_2, Q, S \rangle \tag{1.17}$$

as the context-free grammar where

$$W = V_1 \ \cup \bigcup_{x \in \Sigma_1} V_x, \qquad Q = P_{\mathrm{c}} \ \cup \bigcup_{x \in \Sigma_1} P_x.$$

Such grammar H is called the *composition* of $\mathcal{G}_1$ and $\mathcal{G}_2$ and denoted by

$$G = \mathcal{G}_1 \circ \mathcal{G}_2. \tag{1.18}$$

By the argument above, directly the following lemma holds.

Lemma 1.17. *If G is the grammar (1.18), then $\theta(L) = L(G)$.*

A *composition of k families of context-free grammars* $(\mathcal{G}_1, \ldots, \mathcal{G}_k)$, with $k \geq 1$ will be the context-free grammar

$$G = \mathcal{G}_1 \circ \mathcal{G}_2 \circ \cdots \circ \mathcal{G}_k,$$

obtained by iterating $(k-1)$ times the operation (1.18) of composition of grammars on the families $\mathcal{G}_1, \ldots, \mathcal{G}_k$. By (1.15), Theorem 1.16 can be formulated as follows.

Theorem 1.18. *Let L be an arbitrary language of $Qrt(k)$, $k \geq 1$. Then L is generated by a composition of k families of linear grammars.*

Monoidal languages

We are now able to define the monoidal languages.

Definition 1.19. A grammar $G = \langle V, \Sigma, P, S\rangle$ is said to be *monoidal of index* k if it is the composition

$$G = \mathcal{G}_1 \circ \mathcal{G}_2 \circ \cdots \circ \mathcal{G}_k,$$

of k families of minimal linear grammars where every terminal production $(X \to u),\ u \in \Sigma^*$, is such that $u = \varepsilon$.

A language is said to be *monoidal* if it is generated by a monoidal grammar of index k, for some $k \in \mathbb{N}$.

Example 1.20. Let $\mathcal{L}$ be the language over $\Sigma_2 = \{a, b\}$ given by

$$\mathcal{L} = \{a^n b^n : n \in \mathbb{N}\}^*.$$

Observe that $\mathcal{L}$ is the image of the language $\{x^n : n \in \mathbb{N}\}$, $\Sigma_1 = \{x\}$, under the substitution $\theta : \Sigma_1^* \to \Pi(\Sigma_2^*)$, with $\theta(x) = \{a^n b^n : n \in \mathbb{N}\}$. The language $\{x^n : n \in \mathbb{N}\}$ is rational, then there exists a left-linear (or right-linear) grammar that generates it; on the other hand, the language $\{a^n b^n : n \in \mathbb{N}\}$ is context-free as shown in Example 1.13; hence, θ is a context-free substitution and $\mathcal{L}$ is a context-free language of index 2. Indeed, $\mathcal{L}$ is generated by the composition $G = \mathcal{G}_1 \circ \mathcal{G}_2$ of the linear grammars $\mathcal{G}_1 = \langle V_1, \Sigma_1, P_1, S\rangle$ and $\mathcal{G}_2 = \langle V_2, \Sigma_2, P_2, \sigma\rangle$ where

$$V_1 = \{S\},\ \Sigma_1 = \{\sigma\},\ P_1 = \{S \to \varepsilon,\ S \to \sigma S\},$$

and

$$V_2 = \{\sigma\},\ \Sigma_2 = \{a, b\},\ P_2 = \{\sigma \to \varepsilon,\ \sigma \to a\sigma b\}.$$

Remark 1.21. We end this chapter by remarking that the families of restricted and monoidal languages are incomparable under inclusion; for an example it is sufficient to observe that the restricted language $\{a^n b^n c^n \mid n \in \mathbb{N}\}$ presented in Example 1.15 is not a context-free language (see [25], Theorem 6.2.3, pp. 191–192), indeed it cannot be generated by a context-free grammar, including thus also monoidal grammars. On the other hand, it can be shown (see [17], Lemma 1.5.6, pp. 73–74) that the monoidal language $\{a^n b^n \mid n \in \mathbb{N}\}^*$ illustrated above cannot be generated by a restricted grammar, so that it is not a restricted language.

Chapter 2

Preliminaries on semialgebraic sets

In this chapter we will introduce some topological and algebraic notions concerning the matrices over the field of real numbers and will present examples and preliminary results. Afterwards, we will discuss results which will be used in Chapter 4. For this chapter, we will refer to [5, 7, 8, 12, 13, 19, 22, 41].

2.1 Topology

In this section, we will recall the definitions of some topologies defined on $\mathbb{R}^m$: the *Euclidean topology* and the *Zariski topology*. In particular, we will define the Euclidean topology for square matrices with real entries of dimension n, then define the Zariski topology on $\mathbb{R}^n$ and present preliminary results and examples.

Euclidean topology

Let $M \in \mathbb{R}^{n\times n}$ be a matrix of dimension n and consider it as a vector $M \in \mathbb{R}^{n^2}$, i.e. if M is the matrix

$$M = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad a_{ij} \in \mathbb{R},$$

consider it as the vector $(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{n1}, \ldots, a_{nn}) \in \mathbb{R}^{n^2}$.

The *Euclidean norm* of M is defined as the value

$$\|M\| = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2}.$$

The *Euclidean distance* between two vectors $M, M' \in \mathbb{R}^{n^2}$ is the function

$$\begin{aligned} d : \mathbb{R}^{n^2} \times \mathbb{R}^{n^2} &\to \mathbb{R}_{\geq 0} \\ d(M, M') &\to \|M - M'\|. \end{aligned}$$

For any vector $M_0 \in \mathbb{R}^{n^2}$ and $\varepsilon > 0$ let $I_\varepsilon(M_0) = \{M \in \mathbb{R}^{n^2} : d(M, M_0) < \varepsilon\}$ be the *open ball* with center M_0 and radius ε. A set $A \subseteq \mathbb{R}^{n^2}$ is said to be *open* if for each $M \in A$ there exists $\varepsilon > 0$ such that $I_\varepsilon(M) \subseteq A$; the collection of all open sets forms a topology over $\mathbb{R}^{n^2}$, known as Euclidean topology, which is the topology induced by the Euclidean norm.

We recall the following property needed in the proofs of our main results. It will be applied in the particular case of orthogonal matrices.

Notation. Given a subset E of a finite dimensional vector space, we denote by $\mathbf{C}l(E)$ the topological closure of E respect to the topology induced by the Euclidean norm.

Given a k-tuple of matrices $(M_1, \ldots, M_k)$, denote by f the k-ary product $f(M_1, \ldots, M_k) = M_1 \cdots M_k$ and extend the notation to subsets $\mathcal{A}$ of k-tuples of matrices by posing $f(\mathcal{A}) = \{f(M_1, \ldots, M_k) \mid (M_1, \ldots, M_k) \in \mathcal{A}\}$. The following result will be applied in several instances in the next chapters. It says that since we are dealing with compact subsets, the two operators of matrix multiplication and the topological closure commute.

Theorem 2.1. *([12, Theorem 1]) Let $\mathcal{C}$ be a compact subset of matrices and let $\mathcal{A} \subseteq \mathcal{C}^k$ be a k-ary relation. Then we have $\mathbf{C}l(f(\mathcal{A})) = f(\mathbf{C}l(\mathcal{A}))$.*

Proof. Since the function f is continuous, the inverse image of $\mathbf{C}l(f(\mathcal{A}))$ is closed, i.e. $\mathbf{C}l(\mathcal{A}) \subseteq f^{-1}(\mathbf{C}l(f(\mathcal{A})))$ holds which yields $f(\mathbf{C}l(\mathcal{A})) \subseteq \mathbf{C}l(f(\mathcal{A}))$. Now we prove the opposite inclusion. Consider an element $A \in \mathbf{C}l(f(\mathcal{A}))$. It is the limit of a sequence $M_{1,n} \cdots M_{k,n}$ where $(M_{1,n}, \ldots, M_{k,n}) \in \mathcal{A}$ for $n \geq 0$. Because $\mathcal{C}$ is a compact set, there exists a subsequence $(M_{1,n_i}, \ldots, M_{k,n_i}) \in \mathcal{A}$, i.e. an infinite sequence of strictly indices n_i, which converges to a limit point $(A_1, \ldots, A_k) \in \mathbf{C}l(\mathcal{A})$. By continuity one has $f(A_1, \ldots, A_k) = A$ which shows $\mathbf{C}l(f(\mathcal{A})) \subseteq f(\mathbf{C}l(\mathcal{A}))$.

□

Consequently, if $\mathcal{A}$ is a binary relation defined over a compact subset of matrices $\mathcal{C}$—which is a direct product $\mathcal{A}_1 \times \mathcal{A}_2$, with $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{C}$—we have $\mathbf{C}l(\mathcal{A}_1\mathcal{A}_2) = f(\mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2))$. Moreover, as shown in the following lemma, the closure of the direct product of two subsets of matrices $\mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2)$ is equal to the direct product of the closures $\mathbf{C}l(\mathcal{A}_1) \times \mathbf{C}l(\mathcal{A}_2)$.

Lemma 2.2. *For any $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathbb{R}^{n\times n}$, we have $\mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2) = \mathbf{C}l(\mathcal{A}_1) \times \mathbf{C}l(\mathcal{A}_2)$.*

Proof. ($\subseteq$) Let $(x_1, x_2) \in \mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2)$. We will show that $x_1 \in \mathbf{C}l(\mathcal{A}_1)$ and $x_2 \in \mathbf{C}l(\mathcal{A}_2)$. Let $\mathcal{U}_1, \mathcal{U}_2 \subseteq \mathbb{R}^{n\times n}$ be two neighborhoods of x_1 and x_2 respectively. Thus, $\mathcal{U}_1 \times \mathcal{U}_2$ is a neighborhood of (x_1, x_2). Since $(x_1, x_2) \in \mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2)$, we have

$$(\mathcal{A}_1 \cap \mathcal{U}_1) \times (\mathcal{A}_2 \cap \mathcal{U}_2) = (\mathcal{A}_1 \times \mathcal{A}_2) \cap (\mathcal{U}_1 \times \mathcal{U}_2) \neq \varnothing$$

and so $\mathcal{A}_1 \cap \mathcal{U}_1, \mathcal{A}_2 \cap \mathcal{U}_2 \neq \varnothing$. Hence, by definition of closure, we obtain that $x_i \in \mathbf{C}l(\mathcal{A}_i)$, for $i = 1, 2$.

($\supseteq$) Let $(x_1, x_2) \in \mathbf{C}l(\mathcal{A}_1) \times \mathbf{C}l(\mathcal{A}_2)$, i.e. $x_i \in \mathbf{C}l(\mathcal{A}_i)$, for $i = 1, 2$. Consider two neighborhoods $\mathcal{U}_1, \mathcal{U}_2$ of x_1, x_2 respectively. By definition of closure, we have that $\mathcal{A}_i \cap \mathcal{U}_i \neq \varnothing$ for $i = 1, 2$. This implies that $(\mathcal{U}_1 \times \mathcal{U}_2) \cap (\mathcal{A}_1 \times \mathcal{A}_2) \neq \varnothing$ and so $(x_1, x_2) \in \mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2)$.

□

From Theorem 2.1 and Lemma 2.2, we obtain that for any two subsets of matrices $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{C}$, where $\mathcal{C}$ is a compact subspace, holds

$$\mathbf{C}l(\mathcal{A}_1\mathcal{A}_2) = f(\mathbf{C}l(\mathcal{A}_1 \times \mathcal{A}_2)) = f(\mathbf{C}l(\mathcal{A}_1) \times \mathbf{C}l(\mathcal{A}_2)) = \mathbf{C}l(\mathcal{A}_1)\ \mathbf{C}l(\mathcal{A}_2),$$

which yields the following:

Corollary 2.3. *The topological closure of the product of two sets of matrices included in a compact subspace is equal to the product of the topological closures of the two sets.*

Zariski topology

In order to define properly the Zariski topology let us give first the definition of *algebraic sets* over the field of real numbers.

Definition 2.4. A subset $\mathcal{A} \subseteq \mathbb{R}^n$ is *algebraic (over the field of real numbers)* if $\mathcal{A}$ is the zero set of a set $\mathcal{P}$ of polynomials in $\mathbb{R}[x_1, \ldots, x_n]$, i.e., for every vector $\boldsymbol{v} \in \mathbb{R}^n$,

$$\boldsymbol{v} \in \mathcal{A} \iff \forall\, p \in \mathcal{P} :\ p(\boldsymbol{v}) = 0. \tag{2.1}$$

In that case, we will write $\mathbf{V}(\mathcal{P}) = \mathcal{A}$. Conversely, for any subset $\mathcal{A} \subseteq \mathbb{R}^n$, the ideal

$$\mathbf{I}(\mathcal{A}) := \{p \in \mathbb{R}[x_1, \ldots, x_n] \,|\, p(\boldsymbol{v}) = 0 \text{ for all } \boldsymbol{v} \in \mathcal{A}\}$$

is called the *vanishing ideal* of $\mathcal{A}$.

Note that by the Hilbert's basis theorem (see [22], Ch. 1, Sec. 1, Theorem 1), one may assume that the set $\mathcal{P}$ is finite. Even more, since we are dealing with algebraic sets over $\mathbb{R}$, $\mathcal{P}$ can be reduced to a singleton since (2.1) can be assumed to be to the equation

$$\sum_{p \in \mathcal{P}} p(x_1, \ldots, x_n)^2 = 0.$$

Remark 2.5. Consider the constant $1 \in \mathbb{R}[x_1, \ldots, x_n]$. Thus, $\mathbf{V}(\{1\}) = \varnothing$. On the other hand, $\mathbf{V}(\{0\}) = \mathbb{R}^n$. Thus, $\varnothing$ and $\mathbb{R}^n$ are algebraic sets. Moreover, the family of algebraic sets is closed under finite unions and intersections. We prove the latter for the union and intersection of two sets, since the general case can be treated with a similar argument. Let $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathbb{R}^n$ be algebraic sets and $p_1, p_2 \in \mathbb{R}[x_1, \ldots, x_n]$ two polynomials such that, for every vector $\mathbf{v} \in \mathbb{R}^n$, $p_i(\mathbf{v}) = 0$ if and only if $\mathbf{v} \in \mathcal{A}_i$, for $i = 1, 2$. Let $p, q \in \mathbb{R}[x_1, \ldots, x_n]$ be the polynomials defined as $p = p_1 p_2$ and $q = p_1^2 + p_2^2$. For every $\mathbf{v} \in \mathbb{R}^n$ we have that $p(\mathbf{v}) = 0$ if and only if $p_1(\mathbf{v}) = 0 \vee p_2(\mathbf{v}) = 0$, i.e. $\mathbf{v} \in \mathcal{A}_1 \cup \mathcal{A}_2$; on the other hand, $q(\mathbf{v}) = 0$ if and only if $p_1(\mathbf{v}) = 0 \wedge p_2(\mathbf{v}) = 0$, i.e. $\mathbf{v} \in \mathcal{A}_1 \cap \mathcal{A}_2$. This shows that $\mathcal{A}_1 \cup \mathcal{A}_2$ and $\mathcal{A}_1 \cap \mathcal{A}_2$ are algebraic sets.

It is worth noting that the family of algebraic sets is not closed under complement and projection. Indeed, over the field of real numbers, any subset containing a transcendental number is not algebraic.

Example 2.6. Consider the algebraic set $\mathcal{A} = \{x \in \mathbb{R} \,|\, x^2 - 1 = 0\}$, then $\pi \in \mathbb{R} \setminus \mathcal{A}$ and so $\mathbb{R} \setminus \mathcal{A}$ is not algebraic. Let now $\mathcal{A}$ be the algebraic set defined as $\mathcal{A} = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 - y = 0\}$ and consider the projection

$$\begin{aligned} \pi_y :\ \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\to y. \end{aligned}$$

We have that $\pi \in \pi_y(\mathcal{A}) = \mathbb{R}_{\geq 0}$, and so $\pi_y(\mathcal{A})$ is not algebraic.

By Remark 2.5, the algebraic sets over $\mathbb{R}^n$ form the closed sets of a topology on $\mathbb{R}^n$, which is called the Zariski topology. Then, in the Zariski topology, a subset $\mathcal{A} \subseteq \mathbb{R}^n$ is *open* if and only if its complement $\mathbb{R}^n \setminus \mathcal{A}$ is closed (that is, algebraic).

2.2 Semialgebraic sets and properties

In this section we will define a family of subsets of $\mathbb{R}^n$ more general than algebraic sets: the family of *semialgebraic sets*, that enjoy extra closure properties and is therefore more robust. Indeed, this family is defined as the smallest family of sets in $\mathbb{R}^n$ that contains the algebraic sets as well as sets defined by polynomial inequalities, i.e. sets of the form $\{\boldsymbol{v} \in \mathbb{R}^n \mid f(\boldsymbol{v}) > 0\}$ for some polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$, and which is also closed under the boolean operations (complementation, finite unions, and finite intersections). We first introduce some terminology from Logic which is useful for the study of semialgebraic sets; to this purpose, we follow *verbatim* [5, Chapter 2].

A *(first-order) formula* is written starting from *atoms*, which are polynomial equations and inequalities, together with the logical connectives "*and*", "*or*", and "*negation*" (respectively denoted by $\wedge$, $\vee$ and $\neg$) and the existential and universal quantifiers "*exist*" and "*for all*" (respectively denoted by $\exists$, $\forall$). A formula has *free variables*, i.e. non-quantified variables, and *bound variables*, i.e. quantified variables. Formally, the formulas (with real coefficients) and the set of free variables $\mathrm{Free}(\Phi)$ of a formula Φ are defined as follows:

- an atom $f = 0$ or $f > 0$, where f is a polynomial in $\mathbb{R}[x_1, \ldots, x_n]$ is a formula with free variables $\{x_1, \ldots, x_n\}$;
- if Φ_1 and Φ_2 are formulas, then $\Phi_1 \wedge \Phi_2$ and $\Phi_1 \vee \Phi_2$ are formulas with $\mathrm{Free}(\Phi_1 \wedge \Phi_2) = \mathrm{Free}(\Phi_1 \vee \Phi_2) = \mathrm{Free}(\Phi_1) \cup \mathrm{Free}(\Phi_2)$;
- if Φ is a formula, then $\neg(\Phi)$ is a formula with $\mathrm{Free}(\neg(\Phi)) = \mathrm{Free}(\Phi)$;
- if Φ is a formula and $x \in \mathrm{Free}(\Phi)$, then $(\exists\, x)\Phi$ and $(\forall\, x)\Phi$ are formulas with $\mathrm{Free}((\exists\, x)\Phi) = \mathrm{Free}((\forall\, x)\Phi) = \mathrm{Free}(\Phi) \setminus \{x\}$.

A *quantifier free formula* is a formula in which no quantifier appears, neither $\exists$ nor $\forall$. A *basic formula* is a conjunction of atoms. The $\mathbb{R}$*-realization of a*

formula Φ with free variables contained in $\{x_1, \ldots, x_n\}$, denoted by $R(\Phi, \mathbb{R}^n)$, is the set of $\boldsymbol{v} \in \mathbb{R}^n$ such that $\Phi(\boldsymbol{v})$ is true. It is defined by induction on the construction of the formula, starting from atoms:

- $R(f = 0, \mathbb{R}^n) = \{\boldsymbol{v} \in \mathbb{R}^n \mid f(\boldsymbol{v}) = 0\}$;
- $R(f > 0, \mathbb{R}^n) = \{\boldsymbol{v} \in \mathbb{R}^n \mid f(\boldsymbol{v}) > 0\}$;
- $R(f < 0, \mathbb{R}^n) = \{\boldsymbol{v} \in \mathbb{R}^n \mid f(\boldsymbol{v}) < 0\}$;
- $R(\Phi_1 \wedge \Phi_2, \mathbb{R}^n) = R(\Phi_1, \mathbb{R}^n) \cap R(\Phi_2, \mathbb{R}^n)$;
- $R(\Phi_1 \vee \Phi_2, \mathbb{R}^n) = R(\Phi_1, \mathbb{R}^n) \cup R(\Phi_2, \mathbb{R}^n)$;
- $R(\neg\Phi, \mathbb{R}^n) = \mathbb{R}^n \setminus R(\Phi, \mathbb{R}^n)$;
- $R((\exists\, x)\Phi, \mathbb{R}^n) = \{\boldsymbol{v} \in \mathbb{R}^n \mid \exists\, v_0 \in \mathbb{R}, (v_0, \boldsymbol{v}) \in R(\Phi, \mathbb{R}^{n+1})\}$;
- $R((\forall\, x)\Phi, \mathbb{R}^n) = \{\boldsymbol{v} \in \mathbb{R}^n \mid \forall\, v_0 \in \mathbb{R}, (v_0, \boldsymbol{v}) \in R(\Phi, \mathbb{R}^{n+1})\}$.

Two formulas Φ and Ψ such that $\mathrm{Free}(\Phi) = \mathrm{Free}(\Psi) = \{x_1, \ldots, x_n\}$ are said to be $\mathbb{R}$*–equivalent* if $R(\Phi, \mathbb{R}^n) = R(\Psi, \mathbb{R}^n)$.

The first non-trivial property of semialgebraic sets is given by the following result (see [5], Ch. 2, Sec. 4, Theorem 2.92), which states that the family of semialgebraic sets is closed under projection.

Theorem 2.7. *([5, Theorem 2.92]) Given a semialgebraic set of* $\mathbb{R}^{n+1}$*, its projection to* $\mathbb{R}^n$ *is a semialgebraic set.*

The projection theorem (Theorem 2.7) implies that, since an algebraic set is a semialgebraic set, the projection of an algebraic set over $\mathbb{R}^{n+1}$ is a semialgebraic set over $\mathbb{R}^n$. For an example see Example 2.6. Indeed, here the projection of $\pi_y(\mathcal{A})$, where $\mathcal{A} = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 - y = 0\}$, is the semialgebraic set defined by the polynomial inequality $p(y) := y > 0$, i.e. $\pi_y(\mathcal{A}) = \{y \in \mathbb{R} \,|\, y > 0\}$.

We now present two important results (see [5], Ch.2, Sec. 5, Theorem 2.94 and Corollary 2.95), which are direct consequences of the projection theorem.

Theorem 2.8. *(Tarski-Seidenberg Quantifier Elimination Result) Let* Ψ *be a formula with coefficients in* $\mathbb{R}$*. Then there exists a quantifier free formula* Φ *with coefficients in* $\mathbb{R}$ *such that for every* $\boldsymbol{v} \in \mathbb{R}^n$*, the formula* $\Phi(\boldsymbol{v})$ *is true if and only if the formula* $\Psi(\boldsymbol{v})$ *is true.*

Proof. The proof of the theorem is by induction on the number of quantifiers, using as base case the elimination of an existential quantifier which is given by Theorem 2.7. Consider a formula $(\exists\, x)\Lambda(x, y_1, \ldots, y_n)$, where Λ is a quantifier free formula whose atoms are equations and inequalities involving polynomials in $\mathbb{R}[x, y_1, \ldots, y_n]$. Due to Theorem 2.7, there exists a quantifier free formula $\Theta(y_1, \ldots, y_n)$ whose atoms are equations and inequalities involving polynomials in $\mathbb{R}[x, y_1, \ldots, y_n]$ equivalent to $(\exists\, x)\Lambda(x, y_1, \ldots, y_n)$. Indeed, $R((\exists\, x)\Lambda(x, y_1, \ldots, y_n), \mathbb{R}^n)$, which is the projection of the semialgebraic set $R(\Lambda(x, y_1, \ldots, y_n), \mathbb{R}^{n+1})$ defined over $\mathbb{R}$, is semialgebraic and defined over $\mathbb{R}$, and semialgebraic sets defined over $\mathbb{R}$ are realizations of quantifier free formulas with coefficients in $\mathbb{R}$. Now, since $(\forall\, x)\Phi$ is equivalent to $\neg((\exists\, x)\neg\Phi(x))$, the theorem immediately follows by induction on the number of quantifiers. □

Corollary 2.9. *([5, Corollary 2.95]) Let Φ be a formula with real coefficients and free variables $x_1, \ldots, x_n$. The set $\{\boldsymbol{v} \in \mathbb{R}^n \mid \Phi(\boldsymbol{v})\}$ is semialgebraic.*

Due to Theorem 2.8 and Corollary 2.9, we have the equivalence of the two definitions below for semialgebraic sets.

Definition 2.10. A subset $\mathcal{A} \subseteq \mathbb{R}^n$ is *semialgebraic (over the field of real numbers)* if it satisfies one of the two equivalent conditions:

(*i*) $\mathcal{A}$ is the set of vectors satisfying a finite Boolean combination of predicates of the form $p(x_1, \ldots, x_n) > 0$ where $p \in \mathbb{R}[x_1, \ldots, x_n]$;

(*ii*) $\mathcal{A}$ is *first-order definable* in the theory of the structure whose domain are the reals and whose predicates are of the form $p(x_1, \ldots, x_n) > 0$ or $p(x_1, \ldots, x_n) = 0$ with $p \in \mathbb{R}[x_1, \ldots, x_n]$.

We now instantiate the definitions above to the setting of square matrices.

Definition 2.11. A set $\mathcal{A} \subseteq \mathbb{R}^{n\times n}$ of matrices is *algebraic* (resp., *semialgebraic*) if considered as a set of vectors of dimension n^2, it is algebraic (resp., semialgebraic).

We will adopt the following terminology. A set $\mathcal{A}$ of matrices will be called *effective algebraic* (resp., *effective semialgebraic*) if the polynomials (resp., the formula) defining $\mathcal{A}$ can be algorithmically computed.

In the following we rephrase two results of [13] by emphasizing the main features that serve our purpose.

Notation. We will use the following notation. The set of orthogonal matrices of size n with real (resp. rational) coefficients is denoted by $O_n(\mathbb{R})$ (resp. $O_n(\mathbb{Q})$). Moreover, given a subset E of a group, $\langle E\rangle$ and E^* denote the subgroup and the submonoid generated by E, respectively.

Theorem 2.12. *Let $S \subseteq \mathbb{R}^{n\times n}$ be a set of orthogonal matrices, then $\mathbf{C}l(S^*)$ is a group. In particular, for any subset $E \subseteq S$ satisfying $\langle E\rangle = \langle S\rangle$, we have $\mathbf{C}l(S^*) = \mathbf{C}l(\langle E\rangle)$.*

Proof. First, note that $\mathbf{C}l(S^*)$ is a compact set since it is bounded and closed. It is known that every compact subsemigroup of a compact group is a subgroup. Here is a self-contained proof in our setting. Let us now show that for any matrix $M \in \mathbf{C}l(S^*)$, $M^{-1} \in \mathbf{C}l(S^*)$. Consider the sequence of matrices $\{M^k\}_{k\in\mathbb{N}}$. Since $\mathbf{C}l(S^*)$ is compact, $\{M^k\}_{k\in\mathbb{N}}$ admits a subsequence $\{M^{\bar{k}}\}_{\bar{k}\in\mathbb{N}} \subseteq \{M^k\}_{k\in\mathbb{N}}$ which converges in $\mathbf{C}l(S^*)$. Hence, for any $\varepsilon > 0$ there exists $\bar{k} > 0$ and $\ell > \bar{k} + 1$ such that $\|M^{\bar{k}} - M^{\ell}\| < \varepsilon$. On the other hand, since M is an orthogonal matrix, we have

$$\|M^{\bar{k}} - M^{\ell}\| = \|M^{\bar{k}+1}(M^{-1} - M^{\ell-\bar{k}-1})\| = \|(M^{-1} - M^{\ell-\bar{k}-1})\|;$$

thus, for any matrix $M \in \mathbf{C}l(S^*)$ and $\varepsilon > 0$, there exists $M' \in \mathbf{C}l(S^*)$ such that

$$\|M^{-1} - M'\| < \varepsilon.$$

This implies that $M^{-1} \in \mathbf{C}l(S^*)$. Now, if $\langle S\rangle = \langle E\rangle$, we have that $S^* \subseteq \langle E\rangle$ and $G = \mathbf{C}l(S^*) \subseteq \mathbf{C}l(\langle E\rangle)$; on the other hand, $S \subseteq G$ implies $\mathbf{C}l(\langle E\rangle) \subseteq G$ and thus $\mathbf{C}l(S^*) = \mathbf{C}l(\langle E\rangle)$. This concludes the proof.

□

The main consequence of the next theorem is that the topological closure of a monoid of orthogonal matrices is algebraic ([13, Theorem 3.1]; [12, Theorem 7]).

Theorem 2.13. *Let $E = \{M_i\}_{i\in\Sigma}$, where Σ is a finite alphabet, be a set of orthogonal matrices of dimension n. Then $\mathbf{C}l(E^*)$ is algebraic, and if the matrices in E have rational coefficients, we can effectively compute a sequence of polynomials $f_1, \ldots, f_i, \ldots$ with rational coefficients such that*

1. *if $M \in \mathbf{C}l(E^*)$, $f_i(M) = 0$ for all i;*

2. there exists some k such that

$$\mathbf{C}l(E^*) = \{M \in O_n(\mathbb{Q}) \,|\, f_i(M) = 0,\ i = 1, \ldots, k\}.$$

Proof. We first prove that $\mathbf{C}l(E^*)$ is algebraic. By [41], Ch. 3, Sec. 4, Theorem 5, every compact group of real matrices is algebraic. In fact, letting $X = (x_{ij})_{i,j=1,\ldots,n}$ be a matrix of $n \times n$ free variables, the proof of algebraicity in [41] reveals that any compact group G of real matrices of size n is the zero set of

$$\mathbb{R}[X]^G = \{f \in \mathbb{R}[X] \,|\, f(I) = 0 \text{ and } f(gX) = f(X)\ \forall\, g \in G\},$$

i.e., G is the zero set of polynomials in $n \times n$ variables which vanish at the identity, denoted by I, and are invariant under the action of G. Due to Theorem 2.12, $G = \mathbf{C}l(E^*) = \mathbf{C}l(\langle E \rangle)$ is a compact group and then it is algebraic.

For the second part of the proof, let us show that $\mathbf{V}(\mathbb{Q}[X]^G) = G$, where

$$\mathbb{Q}[X]^G = \{f \in \mathbb{Q}[X] \,|\, f(I) = 0 \text{ and } f(M_iX) = f(X)\ \forall\, i \in \Sigma\}.$$

First, we show that

$$\mathbb{R}[X]^G = \{f \in \mathbb{R}[X] \,|\, f(I) = 0 \text{ and } f(M_iX) = f(X)\ \forall\, i \in \Sigma\};$$

indeed, a polynomial is invariant under the action of G if and only if it is invariant under the action of E. Since $E \subseteq G$, every polynomial invariant under the action of G is also invariant under the action of E. Conversely, assume that a polynomial is invariant under the action of E and let $g = M_{i_1} \cdots M_{i_k}$, where $M_{i_h} \in E$ for any $h = 1, \ldots, k$, be an element of E^*. Without loss of generality we can assume $g = E_1E_2$. By setting $Y = E_2X$, since X is a matrix of $n \times n$ free variable and E_2 is an orthogonal (and so invertible) matrix we have that Y is also a matrix of $n \times n$ free variable. Then we have

$$f(gX) = f(E_1E_2X) = f(E_1Y) = f(Y) = f(E_2X) = f(X),$$

from which we obtain that

$$f(gX) = f(X), \quad \forall\, g \in E^*.$$

Let now $\{g_k\}_{k\in\mathbb{N}}$ be a sequence in E^* and suppose $g \in G$ be its limit point, i.e. $\lim_{k\to\infty} g_k = g$. Then we have

$$f(gX) = f(\lim_{k\to\infty} g_kX) = \lim_{k\to\infty} f(g_kX) = \lim_{k\to\infty} f(X) = f(X),$$

where in the second equality we have used the fact that f, being a polynomial, is a continuous function. This shows that

$$f(gX) = f(X), \quad \forall\, g \in G.$$

We prove now that $G = \mathbf{V}(\mathbb{R}[X]^G) = \mathbf{V}(\mathbb{Q}[X]^G)$.
($\subseteq$) It is easily checked that $\mathbb{Q}[X]^G \subseteq \mathbb{R}[X]^G$ and so $\mathbf{V}(\mathbb{R}[X]^G) \subseteq \mathbf{V}(\mathbb{Q}[X]^G)$.
($\supseteq$) The converse inclusion follows from the fact that any polynomial $P \in \mathbb{R}[X]^G$ can be written as linear combination in $\mathbb{Q}[X]^G$. Indeed, let d be the degree of P and let E_d be the set of real polynomials in $n \times n$ variables with degree at most d. The set $V_d = E_d \cap \mathbb{R}[X]^G$ is a vector subspace of E_d. Indeed, for any $f, g \in V_d$ and $\alpha \in \mathbb{R}$, αf and $f + g$ have degree at most d and

- $\alpha f(I) = \alpha \cdot 0 = 0$;
- for each $M \in E$, $\alpha f(MX) = \alpha f(X)$;
- $(f + g)(I) = f(I) + g(I) = 0 + 0 = 0$;
- for each $M \in E$,

$$(f + g)(MX) = f(MX) + g(MX) = f(X) + g(X) = (f + g)(X).$$

Consider now a matrix $M \in E$. Each polynomial $f \in V_d$ is of the form

$$f(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_{11}^{\alpha_{11}^{i,j}} \cdots x_{1n}^{\alpha_{1n}^{i,j}} \cdots x_{n1}^{\alpha_{n1}^{i,j}} \cdots x_{nn}^{\alpha_{nn}^{i,j}}, \tag{2.2}$$

where, for all $i, j, l, m = 1, \ldots, n$, $\alpha_{lm}^{i,j} \in \mathbb{N}_0$ is the degree of the free variable x_{lm} related to the indices i, j. Being d the degree of f, it must be $\sum_{l=1}^{n} \sum_{m=1}^{n} \alpha_{lm}^{i,j} \leq d$ for any $i, j = 1, \ldots, n$. On the other hand,

$$(MX)_{ij} = \sum_{\ell=1}^{n} M_{i\ell} x_{\ell j}$$

and so $f(MX)$ is of the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} (\sum_{\ell=1}^{n} M_{1\ell} x_{\ell 1})^{\alpha_{11}^{i,j}} \cdots (\sum_{\ell=1}^{n} M_{1\ell} x_{\ell n})^{\alpha_{1n}^{i,j}} \cdots (\sum_{\ell=1}^{n} M_{n\ell} x_{\ell 1})^{\alpha_{n1}^{i,j}} \cdots (\sum_{\ell=1}^{n} M_{n\ell} x_{\ell n})^{\alpha_{nn}^{i,j}}, \tag{2.3}$$

rewriting $f(MX)$ by recombining for any i and j respect to variables

$$x_{11}^{\alpha_{11}^{i,j}} \cdots x_{1n}^{\alpha_{1n}^{i,j}} \cdots x_{n1}^{\alpha_{n1}^{i,j}} \cdots x_{nn}^{\alpha_{nn}^{i,j}}, \tag{2.4}$$

comparing the coefficients of (2.4) in the equation $f(MX) = f(X)$ and considering the equation $f(I) = 0$, we get a linear system where the variables are a_{ij} and the coefficients are M_{ij}, for $i, j = 1, \ldots, n$. Since V_d is a vector subspace, it is possible to find a basis from the linear system of equations

$$\begin{cases} f(I) = 0 \\ f(M_i X) = f(X), \quad \forall\, i \in \Sigma. \end{cases} \tag{2.5}$$

Moreover, since for any $i \in \Sigma$, $M_i \in O_n(\mathbb{Q})$, the polynomials $f_1^d, \ldots, f_{k_d}^d$ generated as solutions of the System (2.5) have rational coefficients and

$$V_d = E_d \cap \mathbb{R}[X]^G = \langle f_1^d, \ldots, f_{k_d}^d \rangle.$$

Clearly, for each $d \in \mathbb{N}_0$, $V_d \subset V_{d+1}$ and so $\langle f_1^d, \ldots, f_{k_d}^d \rangle \subset \langle f_1^{d+1}, \ldots, f_{k_{d+1}}^{d+1} \rangle$, so that

$$\mathbb{Q}[X]^G = V = \bigcup_{d=1}^{\infty} V_d = \mathbb{R}[X]^G$$

is a vector subspace and admits a basis of polynomials with rational coefficients

$$\langle f_1, \ldots, f_i, \ldots \rangle, \quad f_i \in \mathbb{Q}[X]^G.$$

Then, $\mathbf{V}(\mathbb{Q}[X]^G) \subseteq \mathbf{V}(\mathbb{R}[X]^G)$ and

$$G = \mathbf{V}(\mathbb{R}[X]^G) = \mathbf{V}(\mathbb{Q}[X]^G).$$

Due to Hilbert's basis Theorem, there exists a finite family of polynomials $\mathcal{F}$ in $\{f_1, \ldots, f_i, \ldots\}$ such that $\mathbf{V}(\mathbb{Q}[X]^G) = \mathbf{V}(\mathcal{F})$. This concludes the proof. □

Closure properties of algebraic and semialgebraic sets

In this paragraph we will recall some closure properties of the class of effective algebraic and semialgebraic sets of matrices investigated in [12]. First, we recall that the family of effective algebraic sets is closed under finite unions and intersections.

The *product* of two sets of matrices is defined as

$$\mathcal{A}_1 \mathcal{A}_2 = \{M_1 M_2 \mid M_1 \in \mathcal{A}_1, M_2 \in \mathcal{A}_2\}. \tag{2.6}$$

In line with the definition given in the case of languages in Chapter 1, we define the *sandwich* operation, denoted again by $\diamond$, whose first operand is a

set of pairs of matrices $\mathcal{A} \subseteq \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ and the second operand is a set of matrices $\mathcal{B} \subseteq \mathbb{R}^{n\times n}$, by setting

$$\mathcal{A} \diamond \mathcal{B} = \{XYZ \mid (X, Z) \in \mathcal{A} \text{ and } Y \in \mathcal{B}\}.$$

The next operation will be used in the following. Given a bijection

$$\pi : \{(i, j) \mid i, j \in \{1, \ldots, n\}\} \to \{(i, j) \mid i, j \in \{1, \ldots, n\}\} \tag{2.7}$$

and a matrix $M \in \mathbb{R}^{n\times n}$ denote by $\pi(M)$ the matrix $\pi(M)_{i,j} = M_{\pi(i,j)}$. Extend this operation to subsets of matrices $\mathcal{A}$. Write $\pi(\mathcal{A})$ to denote the set of matrices $\pi(M)$ for all $M \in \mathcal{A}$.

The last operation is the *sum* (sometimes called *direct sum*) of square matrices $M_1, \ldots, M_k$ whose result is the square block matrix

$$M_1 \oplus \cdots \oplus M_k = \begin{pmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & M_k \end{pmatrix}. \tag{2.8}$$

These notations extend to subsets of matrices in the natural way. Here we assume that all k matrices have the same size $n \times n$. Observe that if the matrices are orthogonal, so is their sum. Such matrices form a subgroup of orthogonal matrices of size $kn \times kn$.

Now we recall some closure properties of the class of semialgebraic sets of matrices (see [12], Sec. 2.4). Let us first consider the operation (2.6). We adopt the following conventions: we write $\exists^n X$ to mean that X is a vector of n bound variables. Furthermore, a vector of $n\times n$ variables can be interpreted as an $n \times n$ matrix of variables. Consider two semialgebraic sets of matrices, say $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively defined by first-order formulas $\Phi_1(X_1)$ and $\Phi_2(X_2)$ where X_1 and X_2 are two matrices of n^2 free variables. Then $\mathcal{A}_1\mathcal{A}_2$ is defined by the formula

$$\exists^{n\times n} X_1\ \exists^{n\times n} X_2 : X = X_1X_2 \wedge \Phi_1(X_1) \wedge \Phi_2(X_2),$$

where $X = X_1X_2$ is an abbreviation for the predicate defining X. This shows that the class of effective semialgebraic sets of matrices is closed with respect to the operation (2.6). The verification of the properties stated below is done similarly.

Proposition 2.14. *Let $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathbb{R}^{n\times n}$ be two sets of matrices, and let π be a one-to-one mapping as in (2.7).*

1) If $\mathcal{A}_1$ and $\mathcal{A}_2$ are algebraic (resp., effective algebraic) so are $\mathcal{A}_1 \cup \mathcal{A}_2$, $\mathcal{A}_1 \cap \mathcal{A}_2$, and $\pi(\mathcal{A}_1)$.

2) If $\mathcal{A}_1$ and $\mathcal{A}_2$ are semialgebraic (resp., effective semialgebraic), so are $\mathcal{A}_1 \cup \mathcal{A}_2$, $\mathcal{A}_1\mathcal{A}_2$ and $\pi(\mathcal{A}_1)$.

Proposition 2.15. *Let $\mathcal{A}_1 \subseteq \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ and $\mathcal{A}_2 \subseteq \mathbb{R}^{n\times n}$ be semialgebraic (resp., effective semialgebraic). Then $\mathcal{A}_1 \diamond \mathcal{A}_2$ is semialgebraic (resp., effective semialgebraic).*

Proposition 2.16. *If $\mathcal{A}_1, \ldots, \mathcal{A}_k \subseteq \mathbb{R}^{n\times n}$ are algebraic (resp., effective algebraic, semialgebraic, effective semialgebraic), then so is the set $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_k$.*

Proposition 2.17. *Let $\mathcal{A}$ be a semialgebraic (resp., effective semialgebraic) set of $kn \times kn$ matrices of the form (2.8). The set $\{X_1 \cdots X_k \mid X_1 \oplus \cdots \oplus X_k \in \mathcal{A}\}$ is semialgebraic (resp., effective semialgebraic).*

2.3 Algebraic irreducible sets

In this section we introduce two classical concepts from Algebraic Geometry: *irreducible algebraic sets* and *regular maps* between algebraic sets. We will then show a useful result that we will use in the proofs of the main results, giving a sufficient condition for the effectiveness property of algebraic sets.

We follow *verbatim* [22, Chapter 1]. Consider the topological space obtained by equipping $\mathbb{R}^m$ with the Zariski topology. Given a nonempty subset $\mathcal{Z}$ of $\mathbb{R}^m$, $\mathcal{Z}$ is said to be *irreducible* if $\mathcal{Z}$ cannot be written as the union of two closed, i.e. algebraic subsets of $\mathcal{Z}$, distinct from $\varnothing$ and $\mathcal{Z}$ itself. Such sets play a role in the theoretical developments presented later (see Chapter 4). Therefore, we recall a definition and some results.

The next notation is useful.

Notation. The Zariski closure of a set $\mathcal{A}$ will be denoted by $\overline{\mathcal{A}}$.

Let us now demonstrate an important characterization of algebraic sets that we will often use in the rest of this section.

Proposition 2.18. *([22, Proposition 1.1.12]) Let $\mathcal{Z} \subseteq \mathbb{R}^m$ be a nonempty algebraic set. Then $\mathcal{Z}$ is irreducible if and only if its vanishing ideal $\mathbf{I}(\mathcal{Z})$ is a prime ideal.*

Proof. Let $\mathcal{Z}$ be an irreducible set and $f, g \in \mathbb{R}[x_1, \ldots, x_n]$ be two polynomials such that $fg \in \mathbf{I}(\mathcal{Z})$. Then $\mathcal{Z} \subseteq \mathbf{V}(\{fg\}) = \mathbf{V}(\{f\}) \cup \mathbf{V}(\{g\})$. So we have

$$\mathcal{Z} = \mathcal{Z} \cap \mathbf{V}(\{fg\}) = \mathcal{Z} \cap (\mathbf{V}(\{f\}) \cup \mathbf{V}(\{g\})) = (\mathcal{Z} \cap \mathbf{V}(\{f\})) \cup (\mathcal{Z} \cap \mathbf{V}(\{g\})).$$

Under the hypothesis of irreducibility of $\mathcal{Z}$, it must be $\mathcal{Z} = \mathcal{Z} \cap \mathbf{V}(\{f\})$ or $\mathcal{Z} = \mathcal{Z} \cap \mathbf{V}(\{g\})$. Suppose $\mathcal{Z} = \mathcal{Z} \cap \mathbf{V}(\{f\})$; thus we have $\mathcal{Z} \subseteq \mathbf{V}(\{f\})$ and so $f \in \mathbf{I}(\mathcal{Z})$. This shows that $\mathbf{I}(\mathcal{Z})$ is a prime ideal.

Suppose now $\mathbf{I}(\mathcal{Z})$ be a prime ideal and let $\mathcal{Z}_1$ and $\mathcal{Z}_2$ be closed subsets such that $\mathcal{Z} = \mathcal{Z}_1 \cup \mathcal{Z}_2$. Suppose that $\mathcal{Z} \neq \mathcal{Z}_1$, $\mathcal{Z}_2$. This implies that $\mathbf{I}(\mathcal{Z}) \subsetneq \mathbf{I}(\mathcal{Z}_1)$ and $\mathbf{I}(\mathcal{Z}) \subsetneq \mathbf{I}(\mathcal{Z}_2)$. Thus, there exist two polynomials f, g such that

$$f \in \mathbf{I}(\mathcal{Z}_1) \setminus \mathbf{I}(\mathcal{Z}) \quad \text{and} \quad g \in \mathbf{I}(\mathcal{Z}_2) \setminus \mathbf{I}(\mathcal{Z}).$$

Consider the polynomial $h = fg$; it is readily checked that $h \in \mathbf{I}(\mathcal{Z}_1) \cap \mathbf{I}(\mathcal{Z}_2)$. However, since $\mathcal{Z} = \mathcal{Z}_1 \cup \mathcal{Z}_2$, we have that $\mathbf{I}(\mathcal{Z}_1) \cap \mathbf{I}(\mathcal{Z}_2) = \mathbf{I}(\mathcal{Z})$ and so $h \in \mathbf{I}(\mathcal{Z})$. Contradiction.

□

The next proposition shows that the cartesian product of two algebraic sets is still algebraic.

Proposition 2.19. *([22, Proposition 1.3.8]) Let $\mathcal{V} \subseteq \mathbb{R}^n$ and $\mathcal{W} \subseteq \mathbb{R}^m$ be nonempty algebraic sets. If $\mathcal{V}$ and $\mathcal{W}$ are irreducible, then $\mathcal{V} \times \mathcal{W}$ is irreducible.*

Proof. Due to Proposition 2.18, we must show that the vanishing ideal

$$\mathbf{I}(\mathcal{V} \times \mathcal{W}) \subseteq \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_m]$$

is a prime ideal.

Let $f_1, f_2 \in \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ be two polynomials such that $f_1 f_2 \in \mathbf{I}(\mathcal{V} \times \mathcal{W})$ and, for a fixed element $w \in W$, construct the two algebraic sets $\mathcal{V}_i(w) \subseteq \mathcal{V}$, $i = 1, 2$, defined as

$$\mathcal{V}_i(w) = \{v \in \mathcal{V} \mid f_i(v, w) = 0\}.$$

Since $f_1 f_2 \in \mathbf{I}(\mathcal{V} \times \mathcal{W})$, we have $\mathcal{V} = \mathcal{V}_1(w) \cup \mathcal{V}_2(w)$. By the irreducibility of $\mathcal{V}$, it must be $\mathcal{V} = \mathcal{V}_1(w)$ or $\mathcal{V} = \mathcal{V}_2(w)$. Consequently, we have $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2$, where $\mathcal{W}_i := \{w \in \mathcal{W} \,|\, \mathcal{V}_i(w) = \mathcal{V}\}$ for $i = 1, 2$. By noting that

$$\mathcal{W}_i = \{w \in \mathcal{W} \,|\, f_i(v, w) = 0 \text{ for all } v \in \mathcal{V}\},$$

we get the algebraicity of $\mathcal{W}_i$, for $i = 1, 2$. By the irreducibility of $\mathcal{W}$, we get that $\mathcal{W} = \mathcal{W}_1$ or $\mathcal{W} = \mathcal{W}_2$. In the first case, we have $f_1(v, w) = 0$ for all $v \in \mathcal{V}$ and $w \in \mathcal{W}$, and so $f_1 \in \mathbf{I}(\mathcal{V} \times \mathcal{W})$. In a very similar way, we get $f_2 \in \mathbf{I}(\mathcal{V} \times \mathcal{W})$. This shows that $\mathbf{I}(\mathcal{V} \times \mathcal{W})$ is a prime ideal, concluding the proof.

□

Let us now introduce the notion of regular map for algebraic sets. We will then show that some properties of algebraic sets are ensured under these maps.

Definition 2.20. Given nonempty algebraic sets $\mathcal{V} \subseteq \mathbb{R}^n$ and $\mathcal{W} \subseteq \mathbb{R}^m$, a map $f : \mathcal{V} \to \mathcal{W}$ is said to be *regular*, if there exist $p_1, \ldots, p_m \in \mathbb{R}[x_1, \ldots, x_n]$ (where the x_i are indeterminates) such that, for all $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathcal{V}$

$$f(\mathbf{x}) = (p_1(\mathbf{x}), \ldots, p_m(\mathbf{x})). \tag{2.9}$$

Remark 2.21. Any regular map f is continuous in the Zariski topology.

Let $\mathcal{Z} \subseteq \mathcal{W}$ be a closed set. Then there exists a family of polynomials $R \subseteq \mathbb{R}[y_1, \ldots, y_m]$ (where the y_j are the variables) such that $\mathcal{Z} = \mathbf{V}(R) \subseteq \mathbb{R}^m$; thus,

$$f^{-1}(\mathcal{Z}) = \mathbf{V}(\{g(p_1, \ldots, p_m) \,|\, g \in R\}) \subseteq \mathbb{R}^n.$$

Here, $g(p_1, \ldots, p_m)$ is the polynomial with variables $x_1, \ldots, x_n$ defined by replacing, for each $i = 1, \ldots, m$, the variable y_i with the polynomial p_i. Thus, for all $g \in R$, $g(p_1, \ldots, p_m) \in \mathbb{R}[x_1, \ldots, x_n]$ and so $f^{-1}(\mathcal{Z})$ is an algebraic set in $\mathbb{R}^n$.

We show the following well known lemma.

Lemma 2.22. *If $f : \mathcal{V} \to \mathcal{W}$ is a regular map and $\mathcal{V}$ is irreducible, then the Zariski closure $\overline{f(\mathcal{V})}$ of $f(\mathcal{V})$ is also irreducible.*

Proof. Let $p, q \in \mathbb{R}[y_1, \ldots, y_m]$ be two polynomials such that $pq \in \mathbf{I}(\overline{f(\mathcal{V})})$. By definition, for all $v \in V$, we have $p(f(v))q(f(v)) = 0$ and so

$$p(p_1, \ldots, p_m)q(p_1, \ldots, p_m) \in \mathbf{I}(\mathcal{V}).$$

By the irreducibility of $\mathcal{V}$ and applying Proposition 2.18, we have that $\mathbf{I}(\mathcal{V})$ is a prime ideal; this means that $p(p_1, \ldots, p_m) \in \mathbf{I}(\mathcal{V})$ or $q(p_1, \ldots, p_m) \in \mathbf{I}(\mathcal{V})$. Assume $p(p_1, \ldots, p_m) \in \mathbf{I}(\mathcal{V})$. In this case, $p(f(v)) = 0$ for all $v \in \mathcal{V}$, so that $f(\mathcal{V}) \subseteq \mathbf{V}(\{p\})$. However, $\mathbf{V}(\{p\}) \subseteq \mathbb{R}^n$ is a closed set and, hence, p also vanishes on $\overline{f(\mathcal{V})}$; that is, $p \in \mathbf{I}(\overline{f(\mathcal{V})})$. This concludes the proof.

□

We will now consider irreducible algebraic sets of the group $\mathrm{GL}_n(\mathbb{R})$ of invertible $n \times n$ matrices over $\mathbb{R}$ equipped with the Zariski topology. The next closure properties can be readily checked.

Proposition 2.23. *The family of irreducible algebraic sets is closed with respect to the operations (2.7), (2.8) and the (finite) Cartesian product of sets.*

Proof. Let $\mathcal{A}$ be an irreducible algebraic set. By Proposition 2.14, with π as in (2.7), $\pi(\mathcal{A})$ is algebraic so $\overline{(\pi(\mathcal{A}))} = \pi(\mathcal{A})$ and the claim comes from Lemma 2.22 since π is a regular map. For the operation (2.8) and Cartesian product, the fact that, given irreducible algebraic sets $\mathcal{A}_1, \ldots, \mathcal{A}_k$, then so is the set $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_k$ (resp., $\mathcal{A}_1 \times \cdots \times \mathcal{A}_k$) comes from Proposition 2.19.

□

The proof of the next lemma is almost immediate.

Lemma 2.24. *Let $H_1, \ldots, H_m$ be subsets of an arbitrary monoid M, every one of which contains the identity 1 of M. Then $(H_1 \cup \cdots \cup H_m)^* = (H_1 \cdots H_m)^*$.*

Proof. We prove the claim for $m = 2$, i.e., $(A \cup B)^* = (AB)^*$ since the general case is treated similarly. The inclusion $(AB)^* \subseteq (A \cup B)^*$ is obvious. For the reverse inclusion, assume $m = m_1 \cdots m_\ell,\ m_i \in A \cup B,\ 1 \leq i \leq \ell$. Then m can be written as $m = a_1 b_1 \cdots a_\ell b_\ell$, where, for every $i = 1, \ldots, \ell$ $a_i = m_i$ and $b_i = 1$ if $m_i \in A$ while $a_i = 1$ and $b_i = m_i$ if $m_i \in B$.

□

We now introduce an useful notation.

Notation. *Given a subset $\mathcal{A}$ of the group O_m of orthogonal matrices (over $\mathbb{R}$) of order $m \geq 1$, we write $\mathcal{A} \in (\mathcal{H})$ if there exists some $k \geq 1$ such that*

$$\mathcal{A} = \bigcup_{i=1}^{k} \mathcal{A}_i, \tag{2.10}$$

where, for every $i = 1, \ldots, k$, $\mathcal{A}_i$ is an irreducible, effective algebraic set that contains the identity matrix I.

Proposition 2.25. *Let $\mathcal{A}$ be a subset of O_m with $\mathcal{A} \in (\mathcal{H})$ and let Ψ be a regular map*

$$\Psi : O_m \longrightarrow O_n.$$

If, for every $i = 1, \ldots, k$, $\Psi(\mathcal{A}_i)$ contains I, then $\overline{\Psi(\mathcal{A})^*} \in (\mathcal{H})$.

Proof. Let $\mathcal{A} = \bigcup_{i=1}^{k} \mathcal{A}_i$ be a set as in (2.10). We may suppose that $k = 2$, i.e. $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$, since the general case can be treated similarly. Hence $\Psi(\mathcal{A}) = \Psi(\mathcal{A}_1 \cup \mathcal{A}_2)$ so, by Lemma 2.24,

$$\Psi(\mathcal{A})^* = (\Psi(\mathcal{A}_1)\Psi(\mathcal{A}_2))^*.$$

Let $H_1 = \overline{\Psi(\mathcal{A}_1)\Psi(\mathcal{A}_2)}$. By Lemma 2.22, $\overline{\Psi(\mathcal{A}_i)}$ $(i = 1, 2)$, is an algebraic irreducible set. Further, by hypothesis, $\overline{\Psi(\mathcal{A}_i)}$ contains the identity matrix I. Following [19], Sec. 3.1, by using *Gröbner bases techniques*, since $\mathcal{A}_i$ is effective algebraic and Ψ is a regular map, then one can effectively compute $\overline{\Psi(\mathcal{A}_i)}$. Now we observe that

$$\overline{\overline{\Psi(\mathcal{A}_1)} \cdot \overline{\Psi(\mathcal{A}_2)}} \;=\; \overline{\Psi(\mathcal{A}_1) \cdot \Psi(\mathcal{A}_2)}. \tag{2.11}$$

Indeed, taking the map m: $O_n \times O_n \to O_n$ defined, for every $X, Y \in O_n$, as $\mathrm{m}(X, Y) := X \cdot Y$, one has that m is a regular and thus continuous map (w.r.t. O_n endowed with the Zariski topology and $O_n \times O_n$ with the induced product topology). This implies, for arbitrary sets $\mathcal{X}_1, \mathcal{X}_2$, that $\overline{\mathrm{m}(\overline{\mathcal{X}_1}, \overline{\mathcal{X}_2})} = \overline{\mathrm{m}(\mathcal{X}_1, \mathcal{X}_2)}$, and thus Equation (2.11).

Let $H_2 = \overline{H_1 \cdot H_1}$. Since $I \in H_1$ then one has $I \in H_2$. By the very same argument used for H_1 one has that H_2 is an irreducible effective algebraic set such that $H_2 = \overline{\Psi(\mathcal{A})^2}$.

Finally, let us define recursively the family $\{H_i\}_{i \geq 1}$ of subsets of O_n, where, for every $i \geq 2$,

$$H_i = \overline{H_{i-1} \cdot H_1}.$$

By proceeding as in the previous cases, one has that, for every $i \geq 3$, H_i is an effective algebraic irreducible set such that $H_i = \overline{(\Psi(\mathcal{A}_1)\Psi(\mathcal{A}_2))^i}$. and containing I. Moreover, since, for each $i \geq 1$, $I \in H_i$, then one has the following ascending chain

$$H_1 \subseteq H_2 \subseteq \cdots H_i \subseteq \cdots \tag{2.12}$$

which, by the irreducibility of all of its terms, terminates, i.e., there exists some $d \in \mathbb{N}$ such that (cf Remark 2.26)

$$H_{d+1} = \overline{(\Psi(\mathcal{A}_1)\Psi(\mathcal{A}_2))^{d+1}} = \overline{(\Psi(\mathcal{A}_1)\Psi(\mathcal{A}_2))^d} = H_d. \tag{2.13}$$

Finally let $\mathcal{X} = \bigcup_{i\geq 1} H_i$. By (2.13), $\overline{\mathcal{X}} = H_1 \cup \cdots \cup H_d$. On the other hand, since, for every $i \geq 1$, $H_i = \overline{(\Psi(\mathcal{A}_1)\Psi(\mathcal{A}_2))^i}$, then it is readily checked that

$$\overline{\mathcal{X}} = H_1 \cup \cdots \cup H_d = \overline{\Psi(\mathcal{A})^*}.$$

This proves the claim.

□

Remark 2.26. The integer d of (2.13) can be bounded. Indeed, for an algebraic set $\mathcal{V} \subseteq \mathbb{R}^n$, denote by $\dim(\mathcal{V})$ its dimension (see [22, Definition 1.2.15]). By [22, Proposition 1.2.20], for any two algebraic sets $\mathcal{V}, \mathcal{W} \subseteq \mathbb{R}^n$, if $\mathcal{V}$ is irreducible and $\mathcal{V} \subset \mathcal{W}$, then $\dim(\mathcal{V}) < \dim(\mathcal{W})$. Since all the sets of (2.12) are irreducible algebraic subsets of $\mathrm{GL}_n(\mathbb{R})$, which has dimension n^2, one has $d \leq n^2$.

Chapter 3

Automata and quantum computing

In this chapter we will describe quantum finite automata and the problem on which we will state our main results in Chapter 4. Specifically, in the first section we will formally describe quantum finite automata. In the second section, we will introduce some central decidability problems in Theoretical Computer Science, with particular emphasis to the so called Intersection problem; then, we will observe that some of these problems become decidable for quantum finite automata, in contrast to the case of probabilistic automata. In the third section, we will present the first result for the Intersection Problem for non-trivial families of languages, showing that it is decidable for the family of linear languages. For this chapter, we will refer to [2, 3, 7, 8, 9, 10, 12, 13, 14, 19, 34, 37, 38, 42, 45].

3.1 Mathematical Background

In this section, we review the motivations behind quantum computing and the mathematical background defining its theoretical framework. We refer to [2, 13, 19]. In particular, we will follow the text by Ambainis and Yakaryilmaz [2, Sections 2,3] for the description of quantum finite automata. Quantum computing is a research area at the intersection of physics and computer science, investigating computation from a quantum mechanical viewpoint. Indeed, the design of a quantum computer would probably allow the resolution of certain computational problems much faster than classical computers. Two

well-known examples of such problems are factoring and discrete logarithm. These two number theoretic problems are considered to be very difficult for classical computers, however, they can be solved efficiently (in polynomial time) with quantum computing (see [45]). Since several widely used cryptosystems (such as RSA and Diffie-Hellman) rely on the difficulty of factoring or discrete logarithm, a quantum computer would be able to break those cryptosystems, threatening the foundations of cryptography. Given that finite automata are one of the most basic models of computation, it is thus natural to study them in the quantum setting.

Remark 3.1. In this regard, it is worth recalling a striking result obtained in a recent contribution of Mereghetti and Palano [35, 36]. Therein it is described a method for the physical implementation of measure-once quantum automata for the recognition of a well-known family of periodic languages.

The simplest way towards understanding quantum finite automata (QFAs) is by regarding them as a generalization of probabilistic finite automata (PFAs). In the following we will introduce PFAs and some undecidable problems for this computing model, then we will describe QFAs and prove in the next section that some of these previously undecidable problems become decidable in the quantum setting.

Probabilistic finite automata

A probabilistic finite automaton is a tuple $\mathcal{A} = \langle \Sigma, Q, \{X_\sigma\}_{\sigma\in\Sigma}, q_0, Q' \rangle$ where:

- Σ is a finite alphabet;
- $Q = \{q_1, \ldots, q_n\}$ is a finite set of states;
- $q_0 \in Q$ is the initial state;
- $Q' \subseteq Q$ is the set of final states;
- for each $\sigma \in \Sigma$, X_σ is a $n \times n$ stochastic matrix[1]: $(X_\sigma)_{ij}$ is the probability to transition from state q_i to state q_j when $\sigma \in \Sigma$ is the input letter. For instance, if each row in all X_σ contains exactly one 1 (and $n-1$ zeros)

[1] A matrix X of dimension n is *stochastic* if: (i) for each $i, j = 1, \ldots, n$, $X_{ij} \in [0, 1]$; (ii) for each $i = 1, \ldots, n$, $\sum_{j=1}^{n} X_{ij} = 1$.

we recover deterministic finite automata defined in Chapter 1. Another simple case is obtained when $|\Sigma| = 1$. In this case, the probabilistic finite automaton is a finite state (homogeneous) Markov chain.

In order to define the languages recognized by a PFA, we need to fix a *threshold* $\lambda \in [0, 1]$. A word $w = \sigma_1 \cdots \sigma_k \in \Sigma^*$ is accepted by $\mathcal{A}$ if the probability of ending up in Q' upon reading w is at least λ. This condition can be conveniently expressed in a matrix formalism. Let $s \in \mathbb{R}^n$ be the row vector of size n such that $s_i = 1$ if $q_i = q_0$ and $s_i = 0$ otherwise. Let $\pi \in \mathbb{R}^n$ be the column vector of size n such that $\pi_i = 1$ if $q_i \in Q'$ and $\pi_i = 0$ otherwise. Denoting by X_w the matrix $X_{\sigma_1} \cdots X_{\sigma_k}$, the word w is accepted with *strict* (resp., *nonstrict*) threshold λ, if $f_{\mathcal{A}}(w) := sX_w\pi > \lambda$ (resp., $f_{\mathcal{A}}(w) \geq \lambda$). Then, the languages recognized by $\mathcal{A}$ with strict and nonstrict threshold λ are respectively

$$\{w \in \Sigma^* \mid f_{\mathcal{A}}(w) > \lambda\} \quad \text{and} \quad \{w \in \Sigma^* \mid f_{\mathcal{A}}(w) \geq \lambda\}.$$

The languages recognized by PFAs with strict (resp., nonstrict) threshold form the class of *stochastic languages* (resp., *co-stochastic languages*). It turns out that one cannot decide whether the set of accepted words is empty, even if λ and the entries of the matrices X_σ are rational numbers. Indeed, the undecidability of all the following problems (is shown in [14, 42]):

1. Does there exist a word $w \in \Sigma^*$ such that $f_{\mathcal{A}}(w) \geq \lambda$?
2. Does there exist a word $w \in \Sigma^*$ such that $f_{\mathcal{A}}(w) \leq \lambda$?
3. Does there exist a word $w \in \Sigma^*$ such that $f_{\mathcal{A}}(w) = \lambda$?
4. Does there exist a word $w \in \Sigma^*$ such that $f_{\mathcal{A}}(w) > \lambda$?
5. Does there exist a word $w \in \Sigma^*$ such that $f_{\mathcal{A}}(w) < \lambda$?

Quantum finite automata

We now give the formal description of quantum finite automata (*cf* [2]). Let $n \in \mathbb{N}$ be a positive integer and consider a quantum system with n basis states, denoted by $|q_1\rangle, \ldots, |q_n\rangle$. A *state* of such a system is a linear combination of basis states with complex coefficients $z_\ell \in \mathbb{C}$

$$|\psi\rangle = z_1 |q_1\rangle + z_2|q_2\rangle + \cdots + z_n|q_n\rangle \tag{3.1}$$

satisfying $|z_1|^2 + \cdots + |z_n|^2 = 1$; in that case, the coefficients z_ℓ are called *amplitudes* and $|\psi\rangle$ is said to be a *superposition* of $|q_1\rangle, \ldots, |q_n\rangle$.

Example 3.2. Consider a quantum system consisting of 2 basis states $|q_1\rangle$ and $|q_2\rangle$. Some of the possible superpositions are $\frac{4}{5}|q_1\rangle + \frac{3}{5}|q_1\rangle$, $\frac{4}{5}|q_1\rangle - \frac{3}{5}|q_2\rangle$ and $\frac{1}{\sqrt{2}}|q_1\rangle + \frac{1}{\sqrt{2}}|q_2\rangle$.

In this sense, any superposition $|\psi\rangle$ can be viewed as a column vector consisting of amplitudes:

$$|\psi\rangle = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}.$$

Thus, the basis states $|q_\ell\rangle$ are n-dimensional vectors with 1 in the ℓ^{th} component and 0 everywhere else and (3.1) can be interpreted as a linear combination of vectors. The *length* of a vector $|\psi\rangle$ is $\|\psi\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$; hence, the requirement that $|z_1|^2 + \cdots + |z_n|^2 = 1$ is equivalent to require that $\|\psi\| = 1$, that is, a superposition is any vector of length 1.

Unitary transformations A *transformation* on a state is specified by a transformation matrix X. If the state before the transformation is $|\psi\rangle$, the state after the transformation is $X|\psi\rangle$. A transformation is valid if and only if

$$\|\psi\| = 1 \implies \|X|\psi\rangle\| = 1. \tag{3.2}$$

A transformation matrix satisfying (3.2) is said to be *unitary*.

A transformation X can be also specified by describing $X|q_1\rangle, \ldots, X|q_n\rangle$; thus, for any superposition $|\psi\rangle = z_1|q_1\rangle + z_2|q_2\rangle + \cdots + z_n|q_n\rangle$, we have

$$X\,|\psi\rangle = z_1 X|q_1\rangle + z_2 X|q_2\rangle + \cdots + z_n X|q_n\rangle.$$

Example 3.2 (continued). The action of a transformation X can be specified by prescribing that X maps

$$|q_1\rangle \to \frac{1}{\sqrt{2}}|q_2\rangle + \frac{1}{\sqrt{2}}|q_1\rangle \quad \text{and} \quad |q_2\rangle \to \frac{1}{\sqrt{2}}|q_1\rangle - \frac{1}{\sqrt{2}}|q_2\rangle, \tag{3.3}$$

thus determining how X acts on superpositions of $|q_1\rangle$ and $|q_2\rangle$; for example, (3.3) implies that X maps the superposition $\frac{4}{5}|q_1\rangle - \frac{3}{5}|q_2\rangle$ to

$$\frac{4}{5}\left(\frac{1}{\sqrt{2}}|q_1\rangle + \frac{1}{\sqrt{2}}|q_2\rangle\right) - \frac{3}{5}\left(\frac{1}{\sqrt{2}}|q_1\rangle - \frac{1}{\sqrt{2}}|q_2\rangle\right) = \frac{1}{5\sqrt{2}}|q_1\rangle + \frac{7}{5\sqrt{2}}|q_2\rangle.$$

Measurements To obtain information about a state, we have to measure it. The simplest measurement is by observing the superposition

$$|\psi\rangle = z_1|q_1\rangle + z_2|q_2\rangle + \cdots + z_n|q_n\rangle$$

with respect to $|q_1\rangle, \ldots, |q_n\rangle$; this gives $|q_\ell\rangle$ with probability $|z_\ell|^2$ ($\|\psi\| = 1$ guarantees that probabilities sum to 1). After the measurement, the state of the system changes to $|q_\ell\rangle$ and repeating the measurement gives the same state $|q_\ell\rangle$.

Example 3.2 (continued). The measurement of $\frac{4}{5}|q_1\rangle + \frac{3}{5}|q_2\rangle$ gives $|q_1\rangle$ with probability $\left(\frac{4}{5}\right)^2 = \frac{16}{25}$ and $|q_2\rangle$ with probability $\left(\frac{3}{5}\right)^2 = \frac{9}{25}$.

Partial measurement It may be the case that we only need to know whether the state $|q_\ell\rangle$ is an accepting state or not; thus, it is possible to perform just partial measurements. Consider a partition into disjoint subsets $\mathcal{P} = \{Q_1, \ldots, Q_k\}$ of $\{|q_1\rangle, \ldots, |q_n\rangle\}$. Then, after a measurement with respect to $\mathcal{P}$, a state $|\psi\rangle = z_1|q_1\rangle + z_2|q_2\rangle + \cdots + z_n|q_n\rangle$ is in Q_i with probability $p_i = \sum_{|q_\ell\rangle \in Q_i} |\alpha_\ell|^2$. The state after the measurement is

$$\sum_{|q_\ell\rangle \in Q_i} \frac{z_\ell}{\sqrt{p_i}}|q_\ell\rangle.$$

Such a measurement can tell us whether a quantum finite automaton accepts a string and, at the same time, preserves the part of the state which consists of accepting states.

Example 3.3. Consider the quantum system with basis states $\{|q_i\rangle,\ i = 1, \ldots, 4\}$ and the partition $Q_1 = \{|q_1\rangle, |q_2\rangle\}$ and $Q_2 = \{|q_3\rangle, |q_4\rangle\}$. Let $|\psi\rangle$ be the superposition defined as $\frac{1}{2}|q_1\rangle + \frac{1}{2}|q_2\rangle + \frac{1}{2}|q_3\rangle + \frac{1}{2}|q_4\rangle$. A partial measurement yields Q_1 with probability $\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$ and, after the measurement, $|\psi\rangle$ is then

$$\frac{1}{\sqrt{2}}|q_1\rangle + \frac{1}{\sqrt{2}}|q_2\rangle.$$

A QFA is thus described as a tuple

$$\mathcal{Q} = \langle \Sigma, Q, \{T_\sigma\}_{\sigma\in\Sigma}, q_0, R\rangle \,,$$

where

- Σ is the input alphabet;
- Q is a finite set of (classical) states. A (quantum) state of $\mathcal{Q}$ can be any superposition of basis states $\{|q\rangle \mid q \in Q\}$: $|\psi\rangle = \sum_{q\in Q} z_q |q\rangle$;
- $|q_0\rangle$, where $q_0 \in Q$, is the initial state;
- for each symbol $\sigma \in \Sigma$, we have a corresponding transformation T_σ on QFA's current state. In simpler models, T_σ is an unitary transformation, denoted by X_σ. In more general models, T_σ can be a sequence of unitary transformations and measurements, with the next operations in the sequence depending on the previous ones.
- R is a rule for determining how the QFA accepts a string. Typically, R is specified by a set of accepting states ($Q' \subseteq Q$) and measuring the final state of the QFA in the standard basis. If an accepting state $|q\rangle$, $q \in Q'$, is obtained, the automaton accepts; otherwise, the automaton rejects.

Example 3.4. Let p be an odd number. Consider the following QFA $\mathcal{Q}$ in one symbol alphabet $\Sigma = \{a\}$. The set of (quantum) basis states is $\{|q_1\rangle, |q_2\rangle\}$ and the initial state is $|q_1\rangle$. The transformation X_a that corresponds to reading a is defined by

$$\begin{aligned} X_a |q_1\rangle &= \cos\alpha\, |q_1\rangle + \sin\alpha\, |q_2\rangle \\ X_a |q_2\rangle &= -\sin\alpha\, |q_1\rangle + \cos\alpha\, |q_2\rangle \,, \end{aligned}$$

where $\alpha = \frac{2\pi}{p}$. The acceptance rule R is as follows. At the end of computation, we measure the state: if the result is $|q_1\rangle$, we accept; otherwise, we reject. After reading the first symbol a, the quantum state becomes

$$|\psi\rangle := X_a |q_1\rangle = \cos\alpha\, |q_1\rangle + \sin\alpha\, |q_2\rangle \,;$$

then, after reading the second symbol a, it becomes

$$\begin{aligned} X_a |\psi\rangle &= \cos\alpha\, X_a |q_1\rangle + \sin\alpha\, X_a |q_2\rangle \\ &= \cos\alpha\,(\cos\alpha\ |q_1\rangle + \sin\alpha\ |q_2\rangle) + \sin\alpha\,(-\sin\alpha\ |q_1\rangle + \cos\alpha\ |q_2\rangle) \\ &= \cos(2\alpha)\ |q_1\rangle + \sin(2\alpha)\ |q_2\rangle \,. \end{aligned}$$

It is possible to show that each next application of X_a also rotates the state in the plane formed by $|q_1\rangle, |q_2\rangle$ by an angle of $\alpha = \frac{2\pi}{p}$. Thus, after reading a^j, $j \in \mathbb{N}$, the state of $\mathcal{Q}$ is $\cos\left(\frac{2\pi j}{p}\right)|q_1\rangle + \sin\left(\frac{2\pi j}{p}\right)|q_2\rangle$. Therefore, $\mathcal{Q}$ accepts a^j with probability $\cos^2\left(\frac{2\pi j}{p}\right)$; that is, if p divides j, $\cos^2\left(\frac{2\pi j}{p}\right) = \cos^2 0 = 1$ and $\cos^2\left(\frac{2\pi j}{p}\right) < 1$ otherwise.

If the alphabet has more than one symbol, in order to preserve the natural left-to-right reading order of input strings, a standard convention in the study of QFAs is to represent quantum states as row vectors and define QFAs using right actions of the transition matrices. To this purpose, we will use the Dirac notation $\langle\psi|$ to denote the conjugate transpose of a quantum state $|\psi\rangle$.

Models

After first appearance, several models of QFAs have been introduced. Here, we consider the model proposed by Moore and Crutchfield in [38], however for completeness we introduce some of the most famous.

1. *Moore-Crutchfield quantum finite automaton* (MCQFA), [38]. In this model we measure the state of the QFA after reading the whole word and accept if the measurement gives a state in the set of final states. For this reason this model is also called "*measure-once*".

2. *Kondacs-Watrous quantum finite automaton* (KWQFA), [34]. In this model the transformations are still unitary matrices but the acceptance rule involves measurements after each step, thus allowing the possibility of halting the computation in the middle of input processing. In particular, the state set Q is partitioned into the set of accepting states Q_a, the set of rejecting states Q_r and the set of non-halting states Q_n, i.e., $Q = Q_a \sqcup Q_r \sqcup Q_n$. After reading each letter of the word, we perform a partial measurement whether the state is in Q_a, Q_r, or Q_n. If Q_a (Q_r) is obtained, the computation terminates and the input string is accepted (rejected). If Q_n is obtained, the computation continues by reading the next symbol; if the computation continues till the end of the word, the computation terminates and the input string is rejected. This model is also called "*measure many*".

3. *Latvian quantum finite automaton* (LQFA), [1]. This model can be regarded as an "intermediate" model between MCQFAs and KWQFAs.

Indeed, after reading each letter of the word a partial measurement is performed—not necessarily of the form presented in KWQFAs—and the acceptance probability is evaluated at the end of computation only, as in MCQFAs.

Linear representation of quantum finite automata over the field of real numbers

In this paragraph, we will show that in the measure-once model it is possible to assume the entries of the components defining the QFA being real instead of complex. As explained previously, in such a model the state of the QFA is measured after reading the whole word and acceptance takes place whenever a final state is observed; formally, the acceptance rule R is thus described by an Hermitian projection matrix, denoted by P. In this paragraph, following *verbatim* [13, p. 1466], we describe how a quantum finite automaton over real numbers can describe a quantum finite automaton in the general form.

Let $\mathcal{Q}$ be a QFA whose components have complex entries. By doubling the number of states, the behavior of $\mathcal{Q}$ can be simulated by another quantum finite automaton $\mathcal{Q}'$ with real entries. More precisely, let $Q = \{q_1, \ldots, q_n\}$ be the set of states of $\mathcal{Q}$. Consider the basis states $\langle q_\ell|$, $q_\ell \in Q$, as (row) vectors of dimension n with 1 in the ℓ^{th} component and 0 everywhere else. We replace each basis state $\langle q_\ell|$ by two states $\langle q_\ell^1|$ and $\langle q_\ell^2|$; in particular, $\langle q_\ell^1|$ (resp., $\langle q_\ell^2|$) is the $2n$-dimensional vector with 1 at $(2\ell-1)^{\text{th}}$ (resp., $(2\ell)^{\text{th}}$) component and 0 everywhere else. Let $\chi : \mathbb{C}^n \to \mathbb{R}^{2n}$ be the linear map that associates to any superposition of the basis states

$$\langle \psi| = \sum_{\ell=1}^{n} (a_\ell + ib_\ell) \langle q_\ell|$$

the linear combination

$$\chi(\langle \psi|) = \sum_{\ell=1}^{n} a_\ell \langle q_\ell^1| + \sum_{\ell=1}^{n} b_\ell \langle q_\ell^2| \,.$$

Let X_σ, $\sigma \in \Sigma$, be one of the matrices of $\mathcal{Q}$. The rows and columns of X_σ are indexed by elements of Q. Let $x_{ij} + iy_{ij}$ be the entry of X_σ at row i and column j, i.e., related to the basis states $\langle q_i|$ and $\langle q_j|$. Recall that a complex number $a + ib$ can be identified to the 2×2 matrix

$$\begin{pmatrix} a & -b \\ a & b \end{pmatrix}.$$

With a slight abuse of notation we also denote by χ the map that associates to each $n \times n$ matrix X_σ the corresponding $2n \times 2n$ matrix X'_σ obtained with the latter replacement. For any matrix with complex (real) coefficients X, denote by X^* (resp., $X^\top$) its conjugate transpose (resp., transpose). It can be checked that for any $v \in \mathbb{C}^n$ and for any $n \times n$ complex matrices X and Y the following relations hold:

- $\chi(Xv) = \chi(X)\chi(v)$;
- $\chi(XY) = \chi(X)\chi(Y)$;
- $\chi(X^*) = \chi(X)^\top$;
- $v^*v = \chi(v)^\top\chi(v)$.

Recalling that (i) unitary matrices, (ii) orthogonal matrices, (iii) complex matrices of orthogonal projection and (iv) real matrices of orthogonal projection are characterized, respectively, by the following conditions: (i) $XX^* = I$, (ii) $XX^\top = I$, (iii) $X = X^* = X^2$, and (iv) $X = X^\top = X^2$, it follows that χ maps unitary matrices to orthogonal matrices and complex matrices of orthogonal projection to real matrices of orthogonal projection. The quantum finite automaton $\mathcal{Q}'$ defined by the initial configuration $\langle s'| = \chi(s)$, the orthogonal matrices $X'_\sigma = \chi(X_\sigma)$ and the projection matrix $P' = \chi(P)$ satisfies $\chi(\langle s| X_w P) = \langle s'| X'_w P'$ for any word $w \in \Sigma^*$, so that

$$\| \langle s| X_w P\|^2 = \|\chi(\langle s| X_w P)\|^2, \quad \forall\, w \in \Sigma^*.$$

Due to argument above, we can reformulate the definition of QFAs as follows:

Definition 3.5. A *quantum finite automaton* $\mathcal{Q}$ is a quadruple $\langle s, \varphi, P, \lambda\rangle$ where $s \in \mathbb{R}^n$ is a (row) vector of unit norm, P is an orthogonal matrix of size n with $P^2 = P$, φ is a morphism

$$\varphi : \Sigma^* \longrightarrow O_n, \tag{3.4}$$

of the free monoid Σ^* into the group O_n of orthogonal $n \times n$-matrices in $\mathbb{R}^{n\times n}$ and the threshold λ has value in $\mathbb{R}$.

In a very similar way to the case of PFAs, defining for any $w \in \Sigma^*$ the value $f_{\mathcal{Q}}(w) := \|s\varphi(w)P\|^2$, the languages recognized by the quantum finite automaton $\mathcal{Q}$ with strict and nonstrict threshold λ are respectively

$$|\mathcal{Q}_>| = \{w \in \Sigma^* \mid f_{\mathcal{Q}}(w) > \lambda\} \quad \text{and} \quad |\mathcal{Q}_\geq| = \{w \in \Sigma^* \mid f_{\mathcal{Q}}(w) \geq \lambda\}.$$

3.2 Effectiveness issues

In this section, we will discuss decidability problems for quantum finite automata. In particular, we begin by introducing the Intersection problem and presenting a condition for its decidability. Later, since Intersection problem generalizes Emptiness problem (problems (4) and (5) of Section 3.1), we show that the latter becomes decidable in the quantum setting, in contrast to the case of PFAs.

Problems (1) through (5) presented in the previous section for PFAs, can be clearly considered also for quantum automata. The first three problems remain undecidable (see [13]). However, problems (4) and (5) become decidable in this setting, unlike the probabilistic case. Before proving this result, we first introduce the *Intersection problem*.

One of the main issue in Formal Language Theory concerns the decidability of the *Emptiness problem*: given a computational model $\mathcal{M}$, is the language $L(\mathcal{M})$ accepted by $\mathcal{M}$ empty? A natural generalization of this problem is the Intersection problem: one considers a predefined family $\mathcal{L}$ of effectively given languages and asks whether the intersection $L(\mathcal{M}) \cap L$ of $L(\mathcal{M})$ with an arbitrary language $L \in \mathcal{L}$ is empty. Then, taking L as the free monoid Σ^* over the input alphabet Σ of $\mathcal{M}$, i.e. $L = \Sigma^*$, one recovers the standard Emptiness problem; furthermore, if $\mathcal{M}$ is a PFA or QFA one recovers problems (1) and (5).

We are mainly interested in effective properties, which means that the quantum finite automaton must be effectively given. To this end, we consider the model of a *rational quantum automaton*, i.e., an automaton where all the coefficients defining $\mathcal{Q} = \langle s, \varphi, P, \lambda \rangle$ are rational numbers.

As pointed out in the Introduction, the problem we study can be formally defined as follows:

$(L, \mathcal{Q})$ Intersection problem

INPUT: a language L from an effectively given family $\mathcal{L}$ and a rational quantum automaton $\mathcal{Q}$.

QUESTION: is it true that $L \cap |\mathcal{Q}_{>}| = \varnothing$?

We state now a general condition that guarantees the decidability of the Intersection problem. It says that, given a formal language $L \subseteq \Sigma^*$ and considering its image under the morphism (3.4)—that is $\varphi(L) = \{\varphi(w) \,|\, w \in L\}$—if the topological closure $\mathbf{C}l(\varphi(L))$ of $\varphi(L)$ is effective semialgebraic, then the $(L, \mathcal{Q})$ Intersection problem is decidable.

Proposition 3.6. *([12, Proposition 1]) Let $\mathcal{Q}$ be a rational quantum automaton. Let $L \subseteq \Sigma^*$ be a formal language such that the set $\mathbf{C}l(\varphi(L))$ is effective semialgebraic. It is recursively decidable whether or not $L \cap |\mathcal{Q}_{>}| = \varnothing$ holds.*

Proof. We prove the inclusion $L \subseteq |\mathcal{Q}_{\leq}|$, which is equivalent to

$$\forall\, w \in L \implies f_{\mathcal{Q}}(w) \;=\; \|s\varphi(w)P\|^2 \;\leq\; \lambda.$$

By the continuity of the map f, the last is equivalent to

$$\forall\, X \in \mathbf{C}l(\varphi(L)) \implies \|sXP\|^2 \;\leq\; \lambda. \tag{3.5}$$

By hypothesis, there exists a first-order formula $\Phi(X)$ that defines $\mathbf{C}l(\varphi(L))$ as a semialgebraic set (see Definition 2.10). Hence (3.5) can be written as

$$\Upsilon \;\equiv\; \forall\, X : \Phi(X) \implies \|sXP\|^2 \;\leq\; \lambda.$$

This implies that $L \subseteq |\mathcal{Q}_{\leq}|$ if and only if Υ holds true. Finally, one verifies whether Υ holds, which can be achieved by applying the Tarski Seidenberg elimination result.

□

Proposition 3.6 shows that these problems are closely related to certain computational questions involving matrices. In particular, a crucial step in the construction of the decision procedure is the ability to effectively compute the Zariski closure of a finitely generated group of matrices over $\mathbb{Q}$.

Due to the latter proposition, one can demonstrate the following result that shows the decidability of the strict emptiness problems for quantum finite automata:

Theorem 3.7. *([13, Theorem 3.2]) The two following problems are decidable:*

(i) given a rational quantum automaton $\mathcal{Q} = \langle s, \varphi, P, \lambda \rangle$, decide whether there exists a word $w \in \Sigma^*$ such that $f_{\mathcal{Q}}(w) > \lambda$;

(ii) given a rational quantum automaton $\mathcal{Q} = \langle s, \varphi, P, \lambda \rangle$, decide whether there exists a word $w \in \Sigma^*$ such that $f_{\mathcal{Q}}(w) < \lambda$.

Proof. The statement follows directly from Theorem 2.13 and Proposition 3.6. □

Remark 3.8. In the proof of Theorem 3.7 we have bypassed the problem of explicitly computing a finite set of polynomials defining $\mathbf{C}l(\varphi(\Sigma^*))$. It is in fact possible by applying the algorithm of Derksen et al. ([19]). An alternative way is to apply Theorem 2.13. In fact, rerunning the proof we find a set of polynomials generating the vector subspace $V = \bigcup_{d \in \mathbb{N}} V_d$; however, from [40, Theorem 9], it is possible to compute a positive integer $d \in \mathbb{N}$ such that $V = V_1 \cup \cdots \cup V_d$. Thus, by solving the related linear system, we are able to compute a finite set of generators of V and thus the polynomials defining $\mathbf{C}l(\varphi(\Sigma^*))$.

We end this section by remarking that it has been proven (see [33]) that the problems (1),(2),(4),(5) are undecidable for the KWQFA model of quantum finite automata, introduced by Kondacs and Watrous and described in the previous section, in contrast with the above result for the MCQFAs.

3.3 The case of linear languages

In this section, we present the first non-trivial case for the decidability of the $(L, \mathcal{Q})$ Intersection problem. We will show that if L is generated by a linear context-free grammar, then the $(L, \mathcal{Q})$ Intersection problem is decidable. In other words, we will show that the problem is decidable for the family of languages $\mathcal{L} = LIN$. This result is the first generalization of Theorem 3.7 and has been proven in [12].

To our purpose, first we show that given a quantum finite automaton $\mathcal{Q}$ and a context-free language L, the closure $\mathbf{C}l(\varphi(L))$ of the image of L under the morphism (3.4) associated with $\mathcal{Q}$ is semialgebraic. To show the latter, we will combine the combinatorial structuring of context-free languages described in Section 1.2 and the closure properties of semialgebraic sets stated in

Section 2.2. Then, we will apply a result from [3] for automata over monoids to demonstrate that if L is a linear language, then $\mathbf{C}l(\varphi(L))$ is effective semialgebraic and conclude by appyling Proposition 3.6. For this section we will refer to [3, 7, 8, 9, 12].

The following construction is crucial. Let $G = \langle V, \Sigma, P, S\rangle$ be a context-free grammar. With an arbitrary nonterminal A of G, consider the set of cycles $C_A = \{(\alpha, \beta) \in \Sigma^* \times \Sigma^* : A \stackrel{*}{\Rightarrow} \alpha A \beta\}$ of A defined in Section 2.2 and associate the set of matrices M_A defined as

$$M_A = \{\varphi(u_1) \oplus \varphi(u_2)^\top \mid (u_1, u_2) \in C_A\}; \tag{3.6}$$

M_A will be called the *monoid of cycles of* A since the following holds.

Lemma 3.9. *Let A be a nonterminal symbol of G. M_A is a monoid and its closure $\mathbf{C}l(M_A)$ is an algebraic group.*

Proof. Observe first that M_A is a subsemigroup of the group $O_n \oplus O_n$. Indeed, if $\varphi(u_1) \oplus \varphi(u_2)^\top$ and $\varphi(v_1) \oplus \varphi(v_2)^\top$ are in M_A then we have

$$(\varphi(u_1) \oplus \varphi(u_2)^\top)(\varphi(v_1) \oplus \varphi(v_2)^\top) = \varphi(u_1 v_1) \oplus \varphi(u_2)^\top \varphi(v_2)^\top, \tag{3.7}$$

where $(u_1, u_2), (v_1, v_2) \in C_A$. On the other hand, $(u_1 v_1, v_2 u_2) \in C_A$ and the corresponding matrix of M_A

$$\varphi(u_1 v_1) \oplus \varphi(v_2 u_2)^\top \tag{3.8}$$

equals (3.7) since $\varphi(v_2 u_2)^\top = \varphi(u_2)^\top \varphi(v_2)^\top$. Since $(\varepsilon, \varepsilon) \in C_A$, the identity matrix $I \in M_A$ and M_A is a monoid. The second statement comes from Theorems 2.12 and 2.13.

□

Proposition 3.10. *Let L be a context-free language. The set $\mathbf{C}l(\varphi(L))$ is semialgebraic. Moreover, if for every $A \in V$, $\mathbf{C}l(M_A)$ is effective algebraic, then $\mathbf{C}l(\varphi(L))$ is effective semialgebraic.*

Proof. By Proposition 1.12, the language L is a finite union of languages of the form $C_S \diamond L''$ with

$$L'' = w_1 L_{A_1} w_2 L_{A_2} \cdots w_\ell L_{A_\ell} w_{\ell+1} \tag{3.9}$$

where, for every $1 \leq i \leq \ell + 1$, $w_i \in \Sigma^*$ and $A_i \in V'$ and L_{A_i} is the language generated by the grammar G_{A_i} obtained by removing the start symbol S from all the rules of the grammar G.

It suffices to show by induction on the number of nonterminal symbols that the subsets

$$\mathbf{C}l(\varphi(C_S) \diamond \varphi(L'')) \tag{3.10}$$

are semialgebraic. If $\mathrm{Card}(V) = 1$, i.e., the set of nonterminal symbols is reduced to S, then L is reduced to $C_S \diamond L'$ and L' is finite. We may further assume that there is a sole terminal rule $S \to w$, so that $L' = \{w\}$. Indeed, since the closure of a finite union of sets coincides with the union of the closures of the sets of the union, it can be checked that

$$\mathbf{C}l(\varphi(C_S) \diamond \varphi(L')) = \bigcup_{w \in L'} \mathbf{C}l(\varphi(C_S) \diamond \varphi(w)).$$

Since the union above is finite, due to Proposition 2.14, it suffices to show that $\mathbf{C}l(\varphi(C_S) \diamond \varphi(w))$ is semialgebraic. By Theorem 2.1, one has[2]

$$\mathbf{C}l(\varphi(L)) = \{X\varphi(w)Y^T \mid X \oplus Y \oplus \{\varphi(w)\} \in \mathbf{C}l(M_S \oplus \varphi(w))\}. \tag{3.11}$$

On the other hand, one has

$$\mathbf{C}l(M_S \oplus \varphi(w)) = \mathbf{C}l(M_S) \oplus \mathbf{C}l(\varphi(w)) = \mathbf{C}l(M_S) \oplus \varphi(w)$$

which, by Lemma 3.9 and Proposition 2.16, is algebraic. By (3.11) and Proposition 2.17, we conclude that $\mathbf{C}l(\varphi(L))$ is semialgebraic. Moreover, since any finite set of matrices is effective semialgebraic, under the hypothesis that $\mathbf{C}l(M_S)$ is effective semialgebraic we obtain that $\mathbf{C}l(\varphi(L))$ is effective semialgebraic for Proposition 2.17. The basis of the induction is thus proved.

Assume now that V contains more than one nonterminal symbol. By Theorem 2.1, Corollary 2.3 and (3.9), we get

$$\mathbf{C}l(\varphi(C_S) \diamond \varphi(L'')) = \{XZY^\top \mid X \oplus Y \oplus Z \in \mathbf{C}l(M_S) \oplus \mathbf{C}l(\varphi(L''))\} =$$

$$\{XZY^\top \mid X \oplus Y \oplus Z \in \mathbf{C}l(M_S) \oplus \mathbf{C}l(\varphi(w_1)\varphi(L_{A_1}) \cdots \varphi(w_\ell)\varphi(L_{A_\ell})\varphi(w_{\ell+1}))\}.$$

By Corollary 2.3, we have

$$\begin{aligned} &\mathbf{C}l(\varphi(w_1)\varphi(L_{A_1}) \cdots \varphi(w_\ell)\varphi(L_{A_\ell})\varphi(w_{\ell+1})) \\ &= \varphi(w_1)\mathbf{C}l(\varphi(L_{A_1})) \cdots \varphi(w_\ell)\mathbf{C}l(\varphi(L_{A_\ell}))\varphi(w_{\ell+1})). \end{aligned} \tag{3.12}$$

[2]By Proposition 2.14, if $\mathbf{C}l(M_S)$ is algebraic, then so is $\{X \oplus Y^\top : X \oplus Y \in \mathbf{C}l(M_S)\}$

On the other hand, since for any $A \in V'$ the context-free grammar G_A has $\mathrm{Card}(V) - 1$ variables, by applying the induction hypothesis to each G_A we have that $\mathbf{C}l(\varphi(L_A))$ is semialgebraic. Furthermore, we have that $\varphi(w_i)$ is algebraic for any $i = 1, \ldots, s$. Then, via Proposition 2.14, the subset (3.12) is semialgebraic. As for the basis of the induction, due to Lemma 3.9 the set $\mathbf{C}l(M_S)$ is algebraic, so that we obtain that the direct sum $\mathbf{C}l(M_S) \oplus \mathbf{C}l(\varphi(L''))$ is semialgebraic for Proposition 2.16. By applying Proposition 2.17 and Proposition 2.14 we conclude that $\mathbf{C}l(\varphi(L))$ is semialgebraic. The second part of the statement is obtained directly from the hypothesis following the same proof; indeed, $\mathbf{C}l(\varphi(L))$ is effective semialgebraic since are effective semialgebraic $\mathbf{C}l(M_S)$ and $\mathbf{C}l(\varphi(L_A))$ for each $A \in V'$.

□

In order to show that $\mathbf{C}l(M_S)$ is effective semialgebraic, let us introduce a useful theorem due to Anisimov and Seifert ([3, 9]) which states that a subset of a group is rational if and only if it is finitely generated. To this purpose, we recall the notion of *starheight* of a rational subset of a monoid and some result.

Let M be a monoid and define inductively subsets $\mathrm{Rat}_0(M) \subset \mathrm{Rat}_1(M) \subset \cdots$ by:

- $X \in \mathrm{Rat}_0(M)$ if and only if X is a finite subset of M;
- $X \in \mathrm{Rat}_{k+1}(M)$ if and only if X is a finite union of sets of the form $Y_1 Y_2 \cdots Y_n$,

where either Y_i is a singleton or $Y_i = Z_i^*$ for some $Z_i \in \mathrm{Rat}_k(M)$.

It can be shown that

$$\mathrm{Rat}(M) = \bigcup_{k \geq 0} \mathrm{Rat}_k(M).$$

The subsets in $\mathrm{Rat}_k(M) \setminus \mathrm{Rat}_{k-1}(M)$ are said to have starheight k.

Lemma 3.11. *Let G be a group and $H \in \mathrm{Rat}(G)$. Then, $H^{-1} = \{h^{-1} \mid h \in H\} \in \mathrm{Rat}(G)$.*

Proof. Since $H \in \mathrm{Rat}(G)$, there exists a non-negative integer $k \geq 0$ such that $H \in \mathrm{Rat}_k(G)$. The proof is by induction on k.

If $k = 0$, then $H \in \mathrm{Rat}_0(G)$ is finite. Thus H^{-1} is finite and $H^{-1} \in \mathrm{Rat}_0(G)$. This demonstrates the basis of induction.

Suppose now that $k > 0$. We show that if H is obtained by finite union of subsets in $\mathrm{Rat}_{k-1}(G)$, then $H^{-1} \in \mathrm{Rat}_k(G)$. We can suppose that $H = H_1 \cup H_2$, where $H_i \in \mathrm{Rat}_{k-1}(G)$ for $i = 1, 2$. By induction hypothesis $H_i^{-1} \in \mathrm{Rat}_{k-1}(G)$, so that we get

$$H_1^{-1} \cup H_2^{-1} = (H_1 \cup H_2)^{-1} = H^{-1} \in \mathrm{Rat}_k(G).$$

A very similar argument to the previous one shows that if H is obtained by finite product of sets in $\mathrm{Rat}_{k-1}(G)$, then $H^{-1} \in \mathrm{Rat}_k(G)$. Finally, suppose that $H = H_1^* = \bigcup_{i \geq 0} H_1^i$ for some $H_1 \in \mathrm{Rat}_{k-1}(G)$. By induction hypothesis, $H_1^{-1} \in \mathrm{Rat}_{k-1}(G)$, so that we obtain

$$H^{-1} = (\bigcup_{i \geq 0} H_1^i)^{-1} = \bigcup_{i \geq 0} (H_1^{-1})^i = (H_1^{-1})^* \in \mathrm{Rat}_k(G),$$

This concludes the proof.

□

Lemma 3.12. *Let G be a group and $H \in \mathrm{Rat}(G)$. Then, $\langle H \rangle \in \mathrm{Rat}(G)$.*

Proof. Since by Lemma 3.11 we have that $H^{-1} \in \mathrm{Rat}(G)$, it suffices to show that $\langle H \rangle = \{H \cup H^{-1}\}^*$. Clearly, $\{H \cup H^{-1}\}^* \subseteq \langle H \rangle$. To show the opposite inclusion, note that

$$\langle H \rangle = \bigcap_{F < G,\, H \subseteq F} F,$$

where the notation $F < G$ means that F is a subgroup of G. However, it can be verified that $\{H \cup H^{-1}\}^*$ is a subgroup of G, so that $\langle H \rangle \subseteq \{H \cup H^{-1}\}^*$.

□

Theorem 3.13. *([9, Theorem 2.7]) Let G be a group and H be a subgroup of G. Then H is finitely generated if and only if H is a rational subset of G.*

Proof. By Lemma 3.12, we have that any finitely generated group $\langle H \rangle$ is a rational subset of G. In order to prove the converse, we first consider the following situation. Let X be a subset of G such that

$$X = z_1 T_1^* z_2 T_2^* \cdots z_n T_n^* z_{n+1}, \tag{3.13}$$

with $z_1, \ldots, z_{n+1} \in G$ and $T_1, \ldots, T_n \subset G$. Defining

$$y_i = z_1 z_2 \cdots z_i, \qquad i = 1, \ldots, n+1, \tag{3.14}$$
$$S_i = y_i T_i y_i^{-1}, \qquad i = 1, \ldots, n,$$
$$X' = \{y_{n+1}\} \cup S_1 \cup \cdots \cup S_n, \tag{3.15}$$

we claim that

$$\langle X \rangle = \langle X' \rangle. \tag{3.16}$$

Indeed, observe that by Equation (3.13) we get $y_{n+1}, y_{n+1}^{-1} \in \langle X \rangle$. Further,

$$S_i = (z_1 \cdots z_i T_i z_{i+1} \cdots z_{n+1}) y_{n+1}^{-1}.$$

It follows that $S_i \subseteq \langle X \rangle$, which in turn implies $X' \subseteq \langle X \rangle$, and hence $\langle X' \rangle \subseteq \langle X \rangle$. Next

$$S_i^* (y_i T_i y_i^{-1})^* = y_i T_i^* y_i^{-1}.$$

Since $z_1 = y_1$ and $z_i = y_{i-1}^{-1} y_i$ for $2 \leq i \leq n+1$, we have

$$X = y_1 T_1^* y_1^{-1} y_2 T_2^* y_2^{-1} \cdots y_n T_n^* y_n^{-1} y_{n+1} = S_1^* S_2^* \cdots S_n^* y_{n+1}.$$

Thus $X \subseteq \langle X' \rangle$, that implies $\langle X \rangle \subseteq \langle X' \rangle$. This proves Equation (3.16).

Consider now a subgroup H of G such that $H \in \mathrm{Rat}(G)$. Since $H = \langle H \rangle$, H has a rational set of generators. We have to show that H has a set of generators of starheight 0. Let R be the rational set of generators of minimal starheight k and assume $k > 0$. Then

$$R = X_1 \cup X_2 \cup \cdots \cup X_r,$$

where each X_h, $(1 \leq h \leq r)$, has the form (3.13) and at least one X_h has starheight k. Set

$$R' = X_1' \cup X_2' \cup \cdots \cup X_r',$$

where each X_h' is deduced from X_h by equations (3.14) and (3.15). Then R' has starheight $k-1$. By Equation (3.16), each X_h is contained in $\langle R' \rangle$ and, conversely, each X_h' is contained in R. Thus $\langle R \rangle = \langle R' \rangle = H$ and R' is a set of generators of H of starheight $k-1$, in contradiction with the minimality of k. Thus $k = 0$ and the theorem is proved.

□

We are now able to demonstrate that, given a linear context-free grammar and a rational quantum automaton, the closures of the monoids of cycles associated to the grammar are effective algebraics.

Lemma 3.14. *Assume that L is a context-free linear language and the quantum automaton is rational. Then, for every variable A, $\mathbf{C}l(M_A)$ is effective algebraic.*

Proof. Due to Theorem 1.3, a subset of a monoid M is rational if it is recognized by some finite M-automaton. Further, by Theorem 3.13, the subgroup generated by a rational subset of a monoid has an effective finite generating set. Observe now that, if L is a context-free linear language, then M_A is a rational submonoid of the group of orthogonal matrices $O_n \oplus O_n$.

Indeed, it is recognized by the finite O_{2n}-automaton whose states are the nonterminal symbols, the transitions are of the form $B \xrightarrow{\varphi(a)\oplus\varphi(b)^\top} C$, where $B \to aCb$ is a rule of the grammar and where the initial and final states coincide with A. Hence $\langle M_A \rangle$ is rational and has an effective finite generating set and the claim follows by applying to $\langle M_A \rangle$ the algorithm of Derksen et al. [19].

□

Proposition 3.10 together with Lemma 3.14 finally yield.

Corollary 3.15. *([12, Proposition 22]) If L is a context-free linear language and the quantum automaton $\mathcal{Q}$ is rational, then its closure $\mathbf{C}l(\varphi(L))$ is an effective semialgebraic set. As a consequence, it is recursively decidable whether or not $L \cap |\mathcal{Q}_>| = \varnothing$ holds.*

An intermediate class of languages between the linear and the finite index ones is that of *metalinear* (or *ultralinear*) languages, denoted by $\mathcal{L}(Meta_{Lin})$. By definition, a language L is in $\mathcal{L}(Meta_{Lin})$ if there exists a context-free grammar G with $L = L(G)$ and an integer $k \in \mathbb{N}$ such that, in every derivation of G, the number of occurrences of variables in each sentential form does not exceed k; in that case, the context-free grammar G is said to be *metalinear of index k*. Formally, metalinear grammars are those context-free grammars $G = \langle V, \Sigma, P, S \rangle$ where all the productions are linear with the possible exception of productions of the form

$$S \to w_1 A_1 w_2 A_2 \cdots w_k A_k w_{k+1}, \tag{3.17}$$

where $A_i \in V$, $w_i \in \Sigma^*$ and whose occurrence implies that the start symbol S does not occur on the right-hand side of any production. The following theorem gives us a useful characterization of metalinear languages as finite union and finite products of linear languages.

Theorem 3.16. *A language L is metalinear if and only if it is a finite union of finite products of context-free linear languages.*

Proof. Let L be a metalinear language of index k and let $G = \langle V, \Sigma, P, S\rangle$ be a metalinear grammar generating L. We can suppose that there is a sole production of the form (3.17). By construction, the language L is equal to the language

$$w_1 L_{A_1} w_2 L_{A_2} \cdots w_k L_{A_k} w_{k+1},$$

where, for each $j = 1, \ldots, k$, L_{A_j} is the language generated by the contex-free grammar G_{A_j} obtained by removing the variable S by G (*cf* Chapter 1).

Since any other production of G is linear, L_{A_j} is a linear language for any $j = 1, \ldots, k$; hence, L is the product of linear languages.

Suppose now that L is the finite product of linear languages; the case in which it is finite union of finite products of linear languages can be treated in a very similar way.

Consider an integer $k \in \mathbb{N}$ and suppose $L = L_1 \cdots L_k$ where, for any $j = 1, \ldots, k$, L_j is the language generated by the linear context-free grammar $G_j = \langle V_j, \Sigma_j, P_j, S_j\rangle$. Let S be a symbol not in V_j for any $j = 1, \ldots, k$ and define

$$V = \{S\} \cup \bigcup_{j=1}^{k} V_j, \qquad \Sigma = \bigcup_{j=1}^{k} \Sigma_j,$$

$$P = \{p\} \cup \bigcup_{j=1}^{k} P_j,$$

where p is the production defined as $p : S \to S_1 \cdots S_k$. It is easily checked that the language L is generated by the metalinear grammar $G = \langle V, \Sigma, P, S\rangle$, concluding thus the proof.

□

Example 3.17. Consider the context-free grammar $G = \langle V, \Sigma, P, S\rangle$ where the set of variables is $V = \{S, A_1, \ldots, A_k\}$, $k \geq 1$, the set of terminal symbols is $\Sigma = \{a, b\}$ and the productions are of the form

$S \to A_1 A_2 \cdots A_k$,

$A_i \to a A_i b$, for $i = 1, \ldots, k$,

$A_i \to \varepsilon$, for $i = 1, \ldots, k$.

Since each derivation of the grammar is of the form

$$S \overset{*}{\Rightarrow} a^{n_1} A_1 b^{n_1} a^{n_2} A_2 b^{n_2} \cdots a^{n_k} A_k b^{n_k},$$

where $n_i \geq 0$ for any $i = 1, \ldots, k$, the language

$$L = \{a^{n_1} b^{n_1} a^{n_2} b^{n_2} \cdots a^{n_k} b^{n_k} \mid \ n_i \in \mathbb{N}, \ i = 1, \ldots, k\} = \{a^n b^n \mid n \in \mathbb{N}\}^k$$

generated by G is metalinear.

The following corollary gives us the first generalization of Corollary 3.15.

Corollary 3.18. *If $L \in \mathcal{L}(Meta_{Lin})$ and $\mathcal{Q}$ is a rational quantum automaton, then it is recursively decidable whether or not $L \cap |\mathcal{Q}_>| = \varnothing$ holds.*

Proof. By Theorem 3.16, L is a finite union of finite products of context-free linear languages. Thus the claim follows from Corollary 3.15, the closure properties of Section 2.2 and Proposition 3.6.

□

Example 3.19. Consider the linear language $L = \{a^n b^n \mid n \in \mathbb{N}\}$ generated by the linear context-free grammar defined in Example 1.13. Due to Corollary 3.15, given a rational quantum automaton $\mathcal{Q}$, we get that the $(L, \mathcal{Q})$ Intersection problem is decidable. Furthermore, due to Corollary 3.18 we get that, for any $k \in \mathbb{N}$, is also decidable the $(L^k, \mathcal{Q})$ Intersection problem. In the next chapter, by demonstrating that the problem is decidable for monoidal languages (under a certain condition on the grammar), we will show that the problem is decidable also for the language $L^* = \{a^n b^n \mid n \in \mathbb{N}\}^*$.

Chapter 4

The Intersection problem for languages of finite index

In this chapter the main results of this thesis are described. Material covered in this chapter has been previously published in [7, 8]. In particular, we will prove that the $(L, \mathcal{Q})$ Intersection problem is decidable for two families of languages: the languages generated by the class of restricted matrix context-free grammars and the languages generated by monoidal grammars, giving thus two non-trivial generalizations of the result given in [12]; moreover, by demonstrating that the family of bounded semilinear languages is (strictly) contained in the family of restricted languages, we will deduce the decidability of the problem also for this family of languages. Finally, we conclude the chapter by providing interesting examples of the results given in the two previous sections and by studying a special case. Specifically, we will study the problem under a restriction on the definition of quantum finite automata: we will study the case in which the transformations (the orthogonal matrices) defining the rational quantum automaton $\mathcal{Q}$ are commutative. In that case, we will show that the problem is recursively decidable for the family of Dyck languages, that are neither restricted nor monoidals (*cf* Chapter 1). For this chapter we will refer to [6, 7, 8, 12, 16, 19].

4.1 The case of restricted matrix languages

In this section, we will extend the decidability of the $(L, \mathcal{Q})$ Intersection problem to the family of languages generated by restricted grammars, giving thus a non-trivial generalization of Corollary 3.15. Later, we will show that any bounded semilinear language can be generated a restricted grammar, proving consequently that the $(L, \mathcal{Q})$ Intersection problem is decidable for this class of languages as well.

In the proof of the main result of this section below, given a language L generated by a restricted grammar of index k, we will construct a certain monoid generalizing the monoid of cycles (3.6) for the case of linear grammars; then, we will show that its closure is an algebraic group thus yielding the semialgebraicity of $\mathbf{C}l(\varphi(L))$; finally, we will show that it is accepted by a finite automaton over O_{2nk}, deducing then the effective semialgebraicity of $\mathbf{C}l(\varphi(L))$ and the decidability of the problem.

Proposition 4.1. *Let L be a language generated by a restricted grammar and let $\mathcal{Q}$ be a quantum automaton. Then one has:*

(i) $\mathbf{C}l(\varphi(L))$ *is semialgebraic;*

(ii) if the quantum automaton $\mathcal{Q}$ is rational, then $\mathbf{C}l(\varphi(L))$ is effective semialgebraic and the $(L, \mathcal{Q})$ Intersection problem is decidable.

Proof. Let L be a language generated by a grammar as in Definition 1.14. Let us prove (i). Set $\mathcal{X} = (V')^k$ and consider the sets

$$\mathcal{Y} = \{X_1 \cdots X_k \in \mathcal{X} : \exists\, m \in M : m = (S \to X_1 \cdots X_k)\}$$

and

$$\mathcal{Z} = \{X_1 \cdots X_k \in \mathcal{X} : \exists\, m \in M : m = (X_1 \to \varepsilon, \ldots, X_k \to \varepsilon)\}.$$

Denote by $(\Sigma^*)^{2k}$ the multiplicative monoid of $2k$-fold relations on Σ^*. It is easily checked that, given $w \in \Sigma^*$, $w \in L$ if and only if there exists some

$$\omega = (u_1, v_1, u_2, v_2, \ldots, u_k, v_k) \in (\Sigma^*)^{2k}$$

such that

$$w = u_1 v_1 u_2 v_2 \cdots u_k v_k$$

and ω belongs to the set $\mathcal{L}$ defined as

$$\mathcal{L} = \bigcup_{p\in\mathcal{Y},\, q\in\mathcal{Z}} \mathcal{L}_{pq}, \tag{4.1}$$

where, for $p = X_1 \cdots X_k,\ q = Y_1 \cdots Y_k$, $\mathcal{L}_{pq}$ is the subset of $(\Sigma^*)^{2k}$ defined as

$$\mathcal{L}_{pq} = \{\ (u_1, v_1, \ldots, u_k, v_k)\ :\ \exists\ \alpha \in M^* :\ X_1 \cdots X_k \Rightarrow_\alpha u_1 Y_1 v_1 \cdots u_k Y_k v_k\ \}.$$

Let $s = (q_1, \ldots, q_n)$ be a repetition-free sequence of elements of $\mathcal{X}$, i.e., for every positive integers i, j, if $i \neq j$, then $q_i \neq q_j$. Clearly, n is less than or equal to the cardinality of $\mathcal{X}$ and therefore the set $\mathcal{S}$ of all such repetition-free sequences is finite. Further, for every $p, q \in \mathcal{X}$, let $\mathcal{S}_{pq}$ be the set

$$\mathcal{S}_{pq} = \{s \in \mathcal{S}\ \mid\ s = (q_1, \ldots, q_n),\ q_1 = p,\ q_n = q\}.$$

Now, for every $p, q \in \mathcal{X}$, define the subset E_{pq} of $(\Sigma^*)^{2k}$ as:

$$E_{pq} = \{\ (u_1, v_1, \ldots, u_k, v_k)\ :\ \exists\ m \in M :\ X_1 \cdots X_k \Rightarrow_m u_1 Y_1 v_1 \cdots u_k Y_k v_k\ \}$$

and the subset C_q of $(\Sigma^*)^{2k}$ as:

$$C_q = \{\ (u_1, v_1, \ldots, u_k, v_k)\ :\ \exists\ \alpha \in M^* :\ X_1 \cdots X_k \Rightarrow_\alpha u_1 X_1 v_1 \cdots u_k X_k v_k\ \}.$$

For every $s = (q_1, \ldots, q_n) \in \mathcal{S}$, define the (possibly empty) subset $\mathcal{L}_s$ of $(\Sigma^*)^{2k}$ as:

$$\mathcal{L}_s = C_{q_1} \diamond E_{q_1 q_2} \diamond C_{q_2} \diamond \cdots \diamond C_{q_{n-1}} \diamond E_{q_{n-1} q_n} \diamond C_{q_n}, \tag{4.2}$$

where $(\diamond)$ corresponds to the extension of the operation (1.4) to $(\Sigma^*)^{2k}$, that is, for subsets $\mathcal{M}, \mathcal{M}'$ of $(\Sigma^*)^{2k}$, the tuple $(w_1, \ldots, w_{2k}) \in \mathcal{M} \diamond \mathcal{M}'$ if and only if there exist

$$(u_1, v_1, \ldots, u_k, v_k) \in \mathcal{M} \quad \text{and} \quad (u'_1, v'_1, \ldots, u'_k, v'_k) \in \mathcal{M}'$$

such that, for every $i = 1, \ldots, k$,

$$w_{2i-1} \ = \ u_i u'_i \quad \text{and} \quad w_{2i} = v'_i v_i.$$

By using an argument similar to that of [16, Theorem 2], one checks that

$$\mathcal{L}_{pq} = \bigcup_{s\in\mathcal{S}_{pq}} \mathcal{L}_s. \tag{4.3}$$

Let now consider the monoid morphism

$$\widehat{\varphi} : (\Sigma^*)^{2k} \to O_{2nk},$$

which maps every element $\omega = (u_1, v_1, \ldots, u_k, v_k)$ into the orthogonal matrix

$$\widehat{\varphi}(\omega) = \varphi(u_1) \ \oplus \ \varphi(v_1) \ \oplus \ \cdots \ \oplus \ \varphi(u_k) \ \oplus \ \varphi(v_k).$$

Since, by Theorem 2.1, one has

$$\mathbf{C}l(\varphi(L)) \ = \ \{M_1 \cdots M_{2k} \ : M_1 \oplus \cdots \oplus M_{2k} \in \mathbf{C}l(\widehat{\varphi}(\mathcal{L}))\},$$

to show that $\mathbf{C}l(\varphi(L))$ is semialgebraic, by Proposition 2.17, it is enough to show that $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$ is so. To this purpose, with an arbitrary $q \in \mathcal{X}$, associate the subset M_q of O_{2nk} defined as

$$\{\varphi(u_1) \oplus \varphi(v_1)^\top \oplus \cdots \oplus \varphi(u_k) \ \oplus \ \varphi(v_k)^\top \mid (u_1, v_1, \ldots, u_k, v_k) \in C_q\}. \quad (4.4)$$

By proceeding as in Lemma 3.9, one proves that M_q is a monoid and its closure $\mathbf{C}l(M_q)$ is an algebraic group. From this, by Proposition 2.14, it follows that $\mathbf{C}l(\widehat{\varphi}(C_q))$ is an algebraic set. This implies, via Theorem 2.1 and Proposition 2.15, that every set $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}_s))$ of (4.2) is semialgebraic. Then the semialgebraicity of $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$ follows, by applying Proposition 2.14, from the last condition and (4.1), (4.3).

(*ii*) Assume that the quantum automaton $\mathcal{Q}$ is rational. First we prove that $\mathbf{C}l(M_q)$ is effective algebraic. One proceeds as in the proof of Lemma 3.14. Indeed, since L is generated by a restricted grammar, then M_q is a regular submonoid of O_{2nk}. Precisely, it is recognized by the finite O_{2nk}–automaton whose states are the elements of $\mathcal{X}$, the initial and final states coincide with q, and the transitions are of the form

$$e = (s \xrightarrow{\ell(e)} t), \quad s, t \in \mathcal{X},$$

where, being $s = X_1 \cdots X_k, \ t = Y_1 \cdots Y_k$, there exists $m \in M$ such that

$$X_1 \cdots X_k \Rightarrow_m u_1 Y_1 v_1 \cdots u_k Y_k v_k,$$

with $\ell(e) \in O_{2nk}$ defined as

$$\ell(e) = \varphi(u_1) \oplus \varphi(v_1)^\top \oplus \cdots \oplus \varphi(u_k) \ \oplus \ \varphi(v_k)^\top.$$

Hence M_q is a regular monoid and the group $\langle M_q \rangle$ has an effective finite generating set by applying theorems 1.3 and 3.13. By applying to $\langle M_q \rangle$ the algorithm of Derksen et al. one gets an effective presentation of $\mathbf{C}l(M_q)$. From this, by Proposition 2.14, it follows that $\mathbf{C}l(\widehat{\varphi}(C_q))$ is effective algebraic as well. Finally, since all the steps of the proof (i) are effective, the fact that $\mathbf{C}l(\widehat{\varphi}(C_q))$ is effective algebraic implies that $\mathbf{C}l(\varphi(L))$ is so. The claim follows by applying Proposition 3.6.

□

We now discuss some implications of Proposition 4.1. First, remark that, by applying the proposition to the matrix grammars of index $k = 1$, we obtain Corollary 3.15. Moreover, note that the language of Example 1.15 is an instance of a bounded semilinear language. Indeed, the following lemma states that the family of bounded semilinear languages is contained in the family of restricted languages.

Lemma 4.2. *Every bounded semilinear language can be generated by a restricted grammar.*

Proof. It is enough to prove the claim for an arbitrary language L of the form (1.13). It is checked that L is generated by the restricted grammar of index k $G = \langle V, \Sigma, M, S \rangle$, where $V = \{S, X_1, \ldots, X_k\}$ and M is given by the matrix rules

- $(S \to X_1 \cdots X_k)$,
- $(X_1 \to \varepsilon, \ldots, X_k \to \varepsilon)$,
- $(S \to u_1^{\mathbf{v}_{01}} X_1 \cdots u_k^{\mathbf{v}_{0k}} X_k)$, with $\mathbf{v}_0 = (\mathbf{v}_{01}, \ldots, \mathbf{v}_{0k})$,
- For every $i = 1, \ldots, \ell$, $(X_1 \to u_1^{\mathbf{v}_{i1}} X_1, \ldots, X_k \to u_k^{\mathbf{v}_{ik}} X_k)$ with $\mathbf{v}_i = (\mathbf{v}_{i1}, \ldots, \mathbf{v}_{ik})$, $1 \leq i \leq k$,

where the vectors $\mathbf{v}_i$, $i = 0, \ldots, k$, define (1.14).

□

Proposition 4.1 and Lemma 4.2 lead to the following.

Corollary 4.3. *If L is bounded semilinear language then its closure $\mathbf{C}l(\varphi(L))$ is semialgebraic. Furthermore, if the quantum finite automaton $\mathcal{Q}$ is rational, then the $(L, \mathcal{Q})$ Intersection problem is decidable.*

Example 4.4. Consider the restricted language $L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ defined in Example 1.15. Let $\mathcal{Q} = (s, \varphi, P, \lambda)$ be a finite quantum finite automaton where $\varphi : \Sigma^* \longrightarrow O_n$ is the morphism (3.4) associated with $\mathcal{Q}$. Recall that $\mathbf{C}l(\mathcal{A})$ and $\overline{\mathcal{A}}$ denote, respectively, the Euclidean and the Zariski closure of a set $\mathcal{A}$.

The set $\mathbf{C}l(\varphi(L))$ is equal to

$$\mathbf{C}l(\{\varphi(a)^n\varphi(b)^n\varphi(c)^n : n \in \mathbb{N}\}) = \mathbf{C}l(\Psi(\{\varphi(a)^n \oplus \varphi(b)^n \oplus \varphi(c)^n : n \in \mathbb{N}\})) \tag{4.5}$$

where $(\oplus)$ is the operation defined in (2.8) and Ψ is the map defined, for every $X, Y, Z \in O_n$, as $\Psi(X \oplus Y \oplus Z) := XYZ$. By Theorem 2.1, the right-side term of (4.5) rewrites as

$$\Psi(\mathbf{C}l(\{\varphi(a)^n \oplus \varphi(b)^n \oplus \varphi(c)^n : n \in \mathbb{N}\})),$$

so that

$$\mathbf{C}l(\varphi(L)) = \Psi(\mathbf{C}l(\{g\}^*)) = \Psi(\mathbf{C}l(\langle\{g\}\rangle)),$$

where $g = \varphi(a) \oplus \varphi(b) \oplus \varphi(c)$. By Theorem 2.13, $\mathbf{C}l(\langle\{g\}\rangle) = \overline{\langle\{g\}\rangle}$ is algebraic. From this, by Proposition 2.17, it follows that $\mathbf{C}l(\varphi(L)) = \Psi(\overline{\langle\{g\}\rangle})$ is semialgebraic.

Now, if the automaton $\mathcal{Q}$ is a rational, then, by applying Derksen's algorithm to $\{g\}$, one computes $\overline{\langle\{g\}\rangle}$ as an algebraic set, which, in turn, implies that $\Psi(\overline{\langle\{g\}\rangle})$ is effective semialgebraic.

We observe that, not only, Proposition 4.1 provides a unitary scheme for the decidability results of [12] (i.e., Corollary 3.15 and Corollary 4.3) but it gives a non trivial extension of them. Indeed the following example provides languages, neither bounded nor context-free, defined by restricted grammars, for which the Intersection problem is now decidable.

Example 4.5. Let L_k, with $k \in \mathbb{N}$, be the language over $\Sigma = \{a, b\}$ defined as $L_k = \{u^k : u \in \Sigma^*\}$. It is easily checked that L_k can be generated by the restricted grammar $G = \langle V, \Sigma, M, S\rangle$, where $V = \{S, X_1, \ldots, X_k\}$, and M is given by the matrix rules $(S \to X_1 \cdots X_k)$, $(X_1 \to \varepsilon, \ldots, X_k \to \varepsilon)$, and, for every $\sigma \in \Sigma$, $(X_1 \to \sigma X_1, \ldots, X_k \to \sigma X_k)$. For a comprehensive description, we will rerun the proof of Proposition 4.1 for the language L_k in Section 4.3.

4.2 The case of monoidal languages

In this section we will give a second non-trivial extension of Corollary 3.15, not comparable with the first extension given in the previous section (*cf* Chapter 1). We will in fact show that the $(L, \mathcal{Q})$ Intersection problem is decidable for the languages generated by monoidal grammars. As for the proofs of the previous cases, in the proof the goal is to show that $\mathbf{C}l(\varphi(L))$ is effectively semialgebraic, so that by the application of Proposition 3.6 we can deduce the decidability of the problem; as for Corollary 3.10 and Proposition 4.1, to obtain the property of effective semialgebraicity of $\mathbf{C}l(\varphi(L))$ it is enough to show that for every variable A, the closure $\mathbf{C}l(M_A)$ of the monoid of cycles M_A is effectively algebraic; the main issue is therefore to show that $\mathbf{C}l(M_S)$ is effectively algebraic. In the proofs of the cases of linear grammars and restricted matrix grammars, we deduced the effective presentation of $\mathbf{C}l(M_S)$ by applying the algorithm of Derksen et al. to the group $\langle M_S \rangle$ generated by M_S, verifying first that the latter had an effective finite generating set; to do this, we constructed an automaton over the group of orthogonal matrices with rational coefficients that recognized it, so that by Theorem 1.3 it was a rational subset and thus finitely generated (with effective computable generators) due to Theorem 3.13. In this case we will use another approach, indeed we will deduce the effective presentation of $\mathbf{C}l(M_S)$ from the application of Proposition 2.25, verifying that this is the Zariski closure $\overline{\hat{\Psi}(H)^*}$ of the monoid generated by the image of a set H of the form (4.6) under a regular map $\hat{\Psi}$.

Proposition 4.6. *Let $(L, \mathcal{Q})$ be a pair where $\mathcal{Q}$ is a rational quantum automaton and L is a language generated by a monoidal grammar $G = \mathcal{G}_1 \circ \mathcal{G}_2 \circ \cdots \circ \mathcal{G}_k$. If the Zariski closures of the monoids of cycles of the grammars of $\mathcal{G}_k$ are irreducible, then $\mathbf{C}l(\varphi(L))$ is effective semialgebraic.*

Remark 4.7. Examples of monoids generated by a finite number of matrices whose Zariski closure is irreducible can be borrowed from [21]. In Theorem 1.1 of *loc.cit.* the authors determine the Zariski closure of a cyclic matrix semigroup over $\mathbb{C}$. By the proof of [27, Corollary 17], the above result implies that if we consider the cyclic matrix semigroup generated by a real matrix with eigenvalues greater than 1, its Zariski closure is irreducible over $\mathbb{R}$. Similarly, [21, Example 3.8] can be adapted to produce an example of a matrix semigroup M, generated by two matrices, such that the Zariski closure of M is

irreducible over $\mathbb{R}$.

We will do the proof of Proposition 4.6 for $k \leq 2$, i.e., for a monoidal grammar $G = \mathcal{G}_1 \circ \mathcal{G}_2$ of index 2, since, by (1.15), the general case is treated with the same argument. The proof is technically structured in the following lemmata.

First, we find useful to recall a notation introduced in Chapter 2.

Notation. Given a subset $\mathcal{A}$ of the group O_m of orthogonal matrices (over $\mathbb{R}$), we write

$$\mathcal{A} \in (\mathcal{H}) \tag{4.6}$$

if $\mathcal{A}$ is a finite union of irreducible, effective algebraic sets, every one of which contains the identity matrix I of O_m.

Lemma 4.8. *Let $G = \mathcal{G}_1$ be a monoidal grammar of index 1 and let $L = L(G)$. Under the assumption of Proposition 4.6, there exist a set $H \in (\mathcal{H})$ and a regular map Ψ such that $\mathbf{C}l(\varphi(L)) = \Psi(H)$. Moreover $\mathbf{C}l(\varphi(L))$ is effective semialgebraic and contains I.*

Proof. Set $G = \langle V, \Sigma, P, S\rangle$ and observe that $V = \{S\}$, so that G has a sole monoid of cycles M_S. Taking into account (3.11) and the form of the terminal productions of G, one has $\mathbf{C}l(\varphi(L_S)) = \Psi(\mathbf{C}l(M_S))$, where $\Psi(X \oplus Y) := XY^\top$. From this and the fact that, by hypothesis, $\mathbf{C}l(M_S)$ is irreducible, it follows the claim.

□

Assume now that $G = \mathcal{G}_1 \circ \mathcal{G}_2$ and let $\mathcal{G}_1 = \langle V_1, \Sigma_1, P_1, S\rangle$ be the first grammar of the composition G. Set $V_1 = \{S\}$. In the sequel, for every $A \in \Sigma_1$, we denote by L_A the language generated by the grammar $G_A \in \mathcal{G}_2$ (*cf* Chapter 1).

Let $\Sigma_1^* \times \Sigma_1^*$ be the multiplicative monoid of binary relations on Σ_1^* and let $\Pi(O_n \oplus O_n)$ be the multiplicative monoid of subsets of $O_n \oplus O_n$. Define now the morphism between the two monoids

$$\zeta : \ \Sigma_1^* \times \Sigma_1^* \ \longrightarrow \ \Pi(O_n \oplus O_n)$$

as: for every $(u, v) \in \Sigma_1^* \times \Sigma_1^*$, with $u = A_1 \cdots A_s,\ v = B_1 \cdots B_t,\ A_i, B_i \in \Sigma_1$

$$\zeta(u, v) \ = \ (\varphi(L_{A_1}) \cdots \varphi(L_{A_s}) \ \oplus \ \varphi(L_{B_1})^\top \cdots \varphi(L_{B_t})^\top) \tag{4.7}$$

where, for every $i = 1, \ldots, t$, $\varphi(L_{B_i})^\top$ denotes the set of matrices

$$\{m^\top \mid m \in \varphi(L_{B_i})\}. \tag{4.8}$$

The following lemma holds.

Lemma 4.9. *There exists an effective computable finite subset $\mathcal{R}$ of $\Sigma_1^* \times \Sigma_1^*$ such that $M_S = \zeta(\mathcal{R}^*)$.*

Proof. Let us first define the set

$$\widehat{C_S} = \{(\alpha, \beta) \in \Sigma_1^* \times \Sigma_1^* : (\alpha, \beta^\sim) \in C_S\},$$

where $\beta^\sim$ denotes the mirror image of β. Observe that $\widehat{C_S}$ is a submonoid of $\Sigma_1^* \times \Sigma_1^*$. Moreover one has

$$M_S = \zeta(\widehat{C_S}).$$

Indeed, let us check $M_S \subseteq \zeta(\widehat{C_S})$. If $\varphi(u) \oplus \varphi(v)^\top \in M_S$, then $S \stackrel{*}{\Rightarrow} uSv$. Since $G = \mathcal{G}_1 \circ \mathcal{G}_2$, the last derivation can be rewritten as

$$S \stackrel{*}{\Rightarrow} \alpha S \beta, \tag{4.9}$$

where (4.9) is a derivation of $\mathcal{G}_1$ and

$$\alpha \stackrel{*}{\Rightarrow} u, \qquad \beta \stackrel{*}{\Rightarrow} v, \tag{4.10}$$

where both the derivations of (4.10) are in the grammars of $\mathcal{G}_2$.

Set now $\alpha = A_1 A_2 \cdots A_s$, and $\beta = B_s B_{s-1} \cdots B_1$, with $A_i, B_i \in \Sigma_1 \cup \{\varepsilon\}, s \geq 0$. From (4.10), one then has

$$u = u_1 \cdots u_s, \qquad v = v_s \cdots v_1,$$

where, for every $i = 1, \ldots, s$, $u_i \in L_{A_i}$, and $v_i \in L_{B_i}$. From this, one has

$$\varphi(u) \in \varphi(L_{A_1}) \cdots \varphi(L_{A_s}), \qquad \varphi(v)^\top \in \varphi(L_{B_1})^\top \cdots \varphi(L_{B_t})^\top. \tag{4.11}$$

Now observe that, by (4.9), $(\alpha, \beta) \in C_S$ so that $(\alpha, \beta^\sim) \in \widehat{C_S}$. By (4.11) the latter implies $\varphi(u) \oplus \varphi(v)^\top \in \zeta(\widehat{C_S})$. Hence $M_S \subseteq \zeta(\widehat{C_S})$. The other inclusion is checked similarly. Finally observe that a finite set $\mathcal{R}$ of binary relations on Σ_1^* generating $\widehat{C_S}$ as a monoid, can be effectively computed from the productions of $\mathcal{G}_1$. The lemma is proved.

□

Lemma 4.10. $\mathbf{C}l(M_S) \in (\mathcal{H})$. *In particular,* $\mathbf{C}l(M_S)$ *is effective algebraic.*

Proof. By Lemma 4.9, there exists an effective computable finite subset $\mathcal{R}$ of $\Sigma_1^* \times \Sigma_1^*$ such that $M_S = \zeta(\mathcal{R}^*)$. First let us prove the following claim:

Claim. *Let* $\mathcal{R} = \{r\}$, *with* $r \in \Sigma_1^* \times \Sigma_1^*$. *Then* $\mathbf{C}l(\zeta(\mathcal{R}))$ *contains* I *and* $\mathbf{C}l(\zeta(\mathcal{R})) = \Psi(H)$ *for some* $H \in (\mathcal{H})$ *(cf (4.6)) and regular map* Ψ.

Let us write $r = (\alpha, \beta)$ with $\alpha = A_1 A_2 \cdots A_s$, $\beta = B_1 \cdots B_t$, $A_i, B_i \in \Sigma_1$, $s, t \geq 0$. We assume first that α and β have the same length, i.e., $s = t$. By (4.7), the fact that ζ is a monoid morphism and Corollary 2.3, one gets

$$\mathbf{C}l(\zeta(\mathcal{R})) = (\mathbf{C}l(\varphi(L_{A_1})) \cdots \mathbf{C}l(\varphi(L_{A_s})) \oplus \mathbf{C}l(\varphi(L_{B_1}))^\top \cdots \mathbf{C}l(\varphi(L_{B_s}))^\top),$$

taking into account that, for an arbitrary set $\mathcal{A}$ of orthogonal matrices, for the set (4.8), one has $\mathbf{C}l(\mathcal{A}^\top) = \mathbf{C}l(\mathcal{A})^\top$. By Lemma 4.8, the previous formula gives

$$\mathbf{C}l(\zeta(\mathcal{R})) = (\Psi(H_1) \cdots \Psi(H_s) \oplus \Psi(G_1)^\top \cdots \Psi(G_s)^\top),$$

where Ψ is a regular map and, for every $i = 1, \ldots, s$, $H_i, G_i \in (\mathcal{H})$.

Since each $\Psi(H_i)$ and $\Psi(G_i)$ of the last formula contain I, the same holds for $\mathbf{C}l(\zeta(\mathcal{R}))$. Set

$$H = H_1 \oplus \cdots \oplus H_s \oplus G_1 \oplus \cdots \oplus G_s$$

and observe that, by Proposition 2.23, $H \in (\mathcal{H})$. Then $\mathbf{C}l(\zeta(\mathcal{R})) = \Psi'(H)$, where Ψ' is the regular map defined as

$$\Psi'(X_1 \oplus \cdots \oplus X_s \oplus Y_1 \oplus \cdots \oplus Y_s) = (\Psi(X_1) \cdots \Psi(X_s) \oplus \Psi(Y_1)^\top \cdots \Psi(Y_s)^\top).$$

This proves the claim if $s = t$. Finally, if $s > t$, then we can write the word $\beta = B_1 \cdots B_t$ as $\beta = B_1 \cdots B_s$, with $B_j = \varepsilon$, for $j = t+1, \ldots, s$. Afterwards, set $L_{B_j} = \{\varepsilon\}$ and $\varphi(L_{B_j}) = \{I\}$, $t+1 \leq j \leq s$. Then one proceeds to prove the claim, by repeating the argument above for $s = t$.

$\diamond$

Since ζ is a monoid morphism, then one has

$$\mathbf{C}l(M_S) = \mathbf{C}l(\zeta(\mathcal{R}^*)) = \mathbf{C}l(\zeta(\mathcal{R})^*).$$

Now taking into account that, by Corollary 2.3, for an arbitrary set $\mathcal{X}$,

$$\mathbf{C}l(\mathcal{X}^*) = \mathbf{C}l(\mathbf{C}l(\mathcal{X})^*), \tag{4.12}$$

one has $\mathbf{C}l(M_S) = \mathbf{C}l(\mathbf{C}l(\zeta(\mathcal{R}))^*)$. By Theorems 2.12 and 2.13, one has

$$\mathbf{C}l(\mathbf{C}l(\zeta(\mathcal{R}))^*) = \overline{\mathbf{C}l(\zeta(\mathcal{R}))^*}, \tag{4.13}$$

where $\overline{\mathbf{C}l(\zeta(\mathcal{R}))^*}$ is the Zariski closure of $(\mathbf{C}l(\zeta(\mathcal{R})))^*$. Set $\mathcal{R} = \{r_1, \ldots, r_s\}$, with $r_i \in \Sigma_1^* \times \Sigma_1^*, s \geq 1$. Since, for every $i = 1, \ldots, s$, $\mathbf{C}l(\zeta(r_i))$ contains the identity, by Lemma 2.24, one gets

$$\mathbf{C}l(\zeta(\mathcal{R}))^* = (\mathbf{C}l(\zeta(r_1)) \cdots \mathbf{C}l(\zeta(r_s)))^*,$$

which, thus, implies

$$\mathbf{C}l(M_S) = \overline{(\mathbf{C}l(\zeta(r_1)) \cdots \mathbf{C}l(\zeta(r_s)))^*}.$$

By the previous claim, for every $i = 1, \ldots, s$, $\mathbf{C}l(\zeta(r_i)) = \Psi(H_i)$, where $H_i \in (\mathcal{H})$, and Ψ is a regular map not depending upon H_i. Hence one gets

$$\mathbf{C}l(M_S) \;=\; \overline{(\Psi(H_1) \;\cdots\; \Psi(H_k))^*}.$$

Observe that in the formula above, we can write $\Psi(H_1) \;\cdots\; \Psi(H_k)$ as $\widehat{\Psi}(H_1 \oplus \cdots \oplus H_k)$, where $\widehat{\Psi}$ is the regular map defined as $\widehat{\Psi}(X_1 \oplus \cdots \oplus X_k) = \Psi(X_1) \cdots \Psi(X_k)$. Hence we get

$$\mathbf{C}l(M_S) \;= \overline{(\widehat{\Psi}(H))^*},$$

where, again, $H = (H_1 \oplus \cdots \oplus H_k) \in (\mathcal{H})$ as one can easily check. Finally the claim of the lemma follows by applying Proposition 2.25.

We are now able to prove the main result of this section.

Proof of Proposition 4.6: As remarked before, it is enough to prove the claim for $k \leq 2$ since, by (1.15), the general case is treated with the same argument. Set $G = \langle V, \Sigma, P, S \rangle$ and observe that, for every $A \in V$, the fact that $\mathbf{C}l(M_A)$ is effective algebraic comes from Lemmata 3.14 and 4.10. The claim now follows from Proposition 3.10.

By Propositions 2.14 and 3.6, immediate corollaries of Proposition 4.6 are the following.

Corollary 4.11. *If the pair (L, Ω) satisfies the hypothesis of Proposition 4.6, then the (L, Ω) Intersection problem is decidable.*

Corollary 4.12. *Let $(L, \mathcal{Q})$ be a pair where $\mathcal{Q}$ is a rational quantum automaton and L is a finite union $L = \bigcup_{i\in\mathcal{J}} L_i$, of monoidal languages, every one of which is such that the corresponding pair $(L_i, \mathcal{Q})$ satisfies the hypothesis of Proposition 4.6. Then the $(L, \mathcal{Q})$ Intersection problem is decidable.*

As mentioned in the Example 3.19, a significant example for the application of Proposition 4.6 is the one given by the language $\mathcal{L} = \{a^n b^n \mid n \in \mathbb{N}\}^*$; indeed the latter can be generated by a monoidal grammar, but not by a metalinear grammar nor by a restricted matrix grammar. The description of the steps of Proposition 4.6 related to the language L is given in the following section.

4.3 Examples

In this section we will study the $(L, \mathcal{Q})$ Intersection problem for two relevant examples. In the first one, we will apply the proof of Proposition 4.1 to the restricted—but neither bounded nor context-free—language L_k presented in Example 4.5. Later, we will study the problem for the language

$$\{a^n b^n : n \in \mathbb{N}\}^* \;=\; \{a^{n_1} b^{n_1} \cdots a^{n_p} b^{n_p} \;:\; n_1, \ldots, n_p,\; p \in \mathbb{N}\}$$

presented in Example 1.20. Indeed, in Section 3.3 we studied the decidability of the $(L, \mathcal{Q})$ Intersection problem for $L = \{a^n b^n \mid n \in \mathbb{N}\}$ and extended the study to L^k for any $k \in \mathbb{N}$. Here, we will show the decidability of the $(L^*, \mathcal{Q})$ Intersection problem. To this purpose, we will rerun the proofs of propositions 2.25 and 4.6 of sections 2.3 and 4.2.

Example 4.5 (continued)

Let $k \in \mathbb{N}$ be a positive integer and consider the language L_k over the alphabet $\Sigma = \{a, b\}$ defined as

$$L_k = \{u^k : u \in \Sigma^*\}.$$

It is readily checked that L_k is generated by the restricted grammar $G = \langle V, \Sigma, M, S\rangle$, where $V = \{S, X_1, \ldots, X_k\}$ and the set of matrix rules M is given by

- $m_S : (S \to X_1 \cdots X_k)$;

- $m_\varepsilon : (X_1 \to \varepsilon, \ldots, X_k \to \ \varepsilon)$;

- $m_\sigma : (X_1 \to \sigma X_1, \ldots, X_k \to \sigma X_k)$, for every $\sigma \in \Sigma$.

Following the proof of Proposition 4.1, by setting $\mathcal{X} = \{X_1, \ldots, X_k\}^k$ the sets $\mathcal{Y}$ and $\mathcal{Z}$ are both equal to the singleton $\{X_1 \cdots X_k\}$. Denoting by $(\Sigma^*)^{2k}$ the multiplicative monoid of $2k$-fold relations on Σ^*, it can be checked that, given $w \in \Sigma^*$, $w \in L_k$ if and only if there exists some $\omega = (u_1, v_1, u_2, v_2, \ldots, u_k, v_k) \in (\Sigma^*)^{2k}$, such that

$$w = u_1 v_1 u_2 v_2 \cdots u_k v_k,$$

and ω belongs to the set

$$\begin{aligned}\mathcal{L} &= \{(u_1, v_1, \ldots, u_k, v_k) \mid \exists\, \alpha \in M^* \,:\, X_1 \cdots X_k \Rightarrow_\alpha u_1 X_1 v_1 \cdots u_k X_k v_k\} \\ &= \{(u_1, \varepsilon, \ldots, u_k, \varepsilon) \mid u_i \in \Sigma^*,\ i = 1, \ldots, k\}.\end{aligned}$$

Consider now the monoid morphism

$$\widehat{\varphi} : (\Sigma^*)^{2k} \to O_{2nk},$$

which maps every element $\omega = (u_1, v_1, \ldots, u_k, v_k)$ into the orthogonal matrix

$$\widehat{\varphi}(\omega) = \varphi(u_1) \ \oplus \ \varphi(v_1)^\top \ \oplus \ \cdots \ \oplus \ \varphi(u_k) \ \oplus \ \varphi(v_k)^\top.$$

Since, by Theorem 2.1, one has

$$\mathbf{C}l(\varphi(L_k)) \ = \ \{M_1 \cdots M_{2k} \ : M_1 \oplus \cdots \oplus M_{2k} \in \mathbf{C}l(\widehat{\varphi}(\mathcal{L}))\},$$

to show that $\mathbf{C}l(\varphi(L_k))$ is semialgebraic, by Proposition 2.17, it is sufficient to show that $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$ is so. To this purpose, note that, by proceeding as in Lemma 3.9, the subset $\widehat{\varphi}(\mathcal{L})$ of O_{2nk} is a monoid and its closure $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$ is an algebraic group. From this, follows the semialgebraicity of $\mathbf{C}l(\varphi(L))$.

Assume now that the quantum finite automaton $\mathcal{Q}$ is rational. We will prove that $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$ is effective algebraic. One proceeds as in the proof of Lemma 3.14. Indeed, since L_k is generated by a restricted grammar, then $\widehat{\varphi}(\mathcal{L})$ is a regular submonoid of O_{2nk}. Specifically, it is recognized by the finite O_{2nk}–automaton whose the unique state is $X_1 \cdots X_k$, that is also the initial and final state, and the transitions are of the loops

$$e_\sigma = (X_1 \cdots X_k \xrightarrow{\ell(e)_\sigma} X_1 \cdots X_k),$$

where, $\ell(e)_\sigma = \varphi(\sigma) \oplus I \oplus \cdots \oplus \varphi(\sigma) \ \oplus \ I$, for every $\sigma \in \Sigma$.

Hence $\widehat{\varphi}(\mathcal{L})$ is a regular monoid and the group $\langle \widehat{\varphi}(\mathcal{L}) \rangle$ has an effective finite generating set by applying theorems 1.3 and 3.13. By applying to $\langle \widehat{\varphi}(\mathcal{L}) \rangle$ the algorithm of Derksen et al. one gets an effective presentation of $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$, meaning that $\mathbf{C}l(\widehat{\varphi}(\mathcal{L}))$ is effective algebraic. From this, we get the effective semialgebraicity of $\mathbf{C}l(\varphi(L_k))$. The claim follows by applying Proposition 3.6.

Example 1.20 (continued)

Let $\mathcal{L}$ be the language of Example 1.20. Since $\mathcal{L} = L^*$, then

$$\mathbf{C}l(\varphi(\mathcal{L})) = \mathbf{C}l(\varphi(L^*)) = \mathbf{C}l(\varphi(L)^*),$$

where φ is the morphism (3.4) of the quantum finite automaton.

Recall that $\mathbf{C}l(\mathcal{A})$ and $\overline{\mathcal{A}}$ denote, respectively, the Euclidean and the Zariski closure of a set $\mathcal{A}$. By Theorems 2.12 and 2.13, one has $\mathbf{C}l(\varphi(L)^*) = \overline{\varphi(L)^*}$. This implies $\mathbf{C}l(\varphi(\mathcal{L})) = \overline{\Psi(G)^*}$, where

- $G = M^*$ is the monoid generated by the matrix $M \in O_n \oplus O_n$

$$M = \varphi(a) \oplus \varphi(b) = \begin{pmatrix} \varphi(a) & 0 \\ 0 & \varphi(b) \end{pmatrix};$$

- Ψ is the regular map defined as $\Psi(A \oplus B) := AB$.

We now prove that, if $\overline{G}$ is irreducible and effective algebraic, then the Zariski closure of the monoid generated by the image of G under the regular map Ψ is finite union of irreducible, effective algebraic sets, every one of which contains the identity matrix I, i.e.

$$\overline{\Psi(G)^*} \in (\mathcal{H}),$$

so as to obtain that $\mathbf{C}l(\varphi(\mathcal{L}))$ is effective algebraic as well.

For this purpose, observe first that $\Psi : O_n \oplus O_n \to O_n$ is a regular and thus a continuous map (*cf* Remark 2.21). This implies that

$$\Psi(\overline{G}) \subseteq \overline{\Psi(G)},$$

which, in turn, implies

$$\overline{\Psi(\overline{G})} \ = \ \overline{\Psi(G)}. \tag{4.14}$$

Let $H_1 = \overline{\Psi(\overline{G})}$. By Lemma 2.22, H_1 is an algebraic irreducible set. Moreover, being G a group, one has $I \in \Psi(G)$ and so $I \in H_1$. Since G is finitely generated as a monoid, then, it is an algebraic set by Theorem 2.13 and, by using Derksen's algorithm, one can effectively compute $\overline{G}$. Afterwards, following [19], Sec. 3.1, by using *Gröbner bases techniques*, since $\overline{G}$ is effective algebraic, then one can effectively compute H_1.

Let $H_2 = \overline{H_1 \cdot H_1}$. Observe first that $I \in H_1$ implies $I \in H_2$. Similarly to the previous case, one observes that, for arbitrary subsets $\mathcal{A}, \mathcal{B}$ of O_n

$$\overline{\overline{\mathcal{A}} \cdot \overline{B}} \;=\; \overline{\mathcal{A} \cdot \mathcal{B}}. \tag{4.15}$$

Indeed, taking the map m: $O_n \times O_n \to O_n$ defined, for every $X, Y \in O_n$, as $\mathrm{m}(X, Y) := X \cdot Y$, one has that m is a continuous map. This implies that $\overline{\mathrm{m}(\overline{\mathcal{A}}, \overline{\mathcal{B}})} = \overline{\mathrm{m}(\mathcal{A}, \mathcal{B})}$. By using the very same argument for H_1 one verifies that H_2 is an irreducible effective algebraic set. Moreover, by equations (4.14) and (4.15), one has $H_2 = \overline{\Psi(G)^2}$.

Finally, let us define recursively the family $\{H_i\}_{i \geq 1}$ of subsets of O_n, where, for every $i \geq 2$, $H_i = \overline{H_{i-1} \cdot H_1}$. We get an ascending chain

$$H_1 \;\subseteq\; H_2 \;\subseteq\; \cdots\; H_i \;\subseteq\; \cdots \tag{4.16}$$

of irreducible, effective algebraic subsets of O_n. By Remark 2.26, we obtain that

$$\dim H_1 \;\leq\; \dim H_2 \;\leq\; \cdots\; \dim H_i \;\leq\; \cdots \tag{4.17}$$

and, thus, that the ascending chain (4.16) terminates, i.e. there exists some integer $m \in \mathbb{N}$ such that

$$H_{m+1} \;=\; \overline{\Psi(G)^{m\,+\,1}} \;=\; \overline{\Psi(G)^m} \;=\; H_m. \tag{4.18}$$

Finally let $\mathcal{X} = \bigcup_{i \geq 1} H_i$. By (4.18), $\overline{\mathcal{X}} = H_1 \cup \cdots \cup H_m$. On the other hand, since, for every $i \geq 1$, $H_i = \overline{\Psi(G)^i}$, it is checked that

$$\overline{\mathcal{X}} = H_1 \cup \cdots \cup H_m = \overline{\Psi(G)^*}.$$

Hence, $\overline{\Psi(G)^*}$ is effective algebraic. Finally, we get the decidability of the $(\mathcal{L}, \mathcal{Q})$ Intersection by applying Proposition 3.6.

4.4 A special case: commutative transformations

We end this chapter by studying special cases of the $(L, \mathcal{Q})$ Intersection problem. In particular, we will observe that the problem becomes decidable for some cases of languages not treated before, under the assumption, although quite restrictive, that the transformations defining the quantum finite automaton are commutative, i.e. the orthogonal matrices associated to the input alphabet are contained in an abelian subgroup of O_n. Indeed, we will observe that under this condition the problem becomes decidable also for Dyck languages (*cf* Chapter 1).

Let $\mathcal{Q}$ be a rational quantum automaton and consider the case in which $\mathcal{Q}$ has only two basis states and the transformations are rotations, i.e. the morphism (3.4) is defined over SO_2 instead of O_2; in particular, the unitary transformations defining $\mathcal{Q}$ are commutative. In the following, we will show that in this case the problem is decidable also for the Dyck languages defined in Chapter 1. To this purpose, we will refer to [6] to recall some preliminary notions.

Let $A = \{a_1, \ldots, a_m\}$ be an alphabet of m letters. For an arbitrary word $w \in A^*$, the *Parikh vector of* w is the tuple $\psi(w)$ of $\mathbb{N}^m$ defined as

$$\psi(w) = (|w|_{a_1}, \ldots, |w|_{a_m});$$

this defines the function

$$\psi : A^* \to \mathbb{N}^m$$

which maps each word $w \in A^*$ into the Parikh vector of w. The map ψ is an epimorphism of the free monoid A^* onto the free commutative additive monoid $\mathbb{N}^m$, called the *Parikh morphism* (over A).

One can then introduce in A^* the equivalence relation $\sim$, called *Parikh equivalence*, defined, for all $u, v \in A^*$, as:

$$u \sim v \iff \psi(u) = \psi(v).$$

Thus, one has $u \sim v$ if the word v is obtained by rearranging the letters of u in a different order. Given a language $L \subseteq A^*$ and a word $w \in L$, we will denote the equivalence class of words in L Parikh equivalent to w by $[w]_{\sim_L}$.

Under our assumptions on the morphism (3.4), we obtain that for any word $w \in L$ and $w' \in [w]_{\sim_L}$, $\psi(w) = \psi(w')$. This leads us to the decidability of the $(L, \mathcal{Q})$ Intersection problem for the Dyck language over an alphabet of two symbols explained in the following.

Consider the Dyck language $\mathcal{L} = D_1^*$ defined in Example 1.6. By definition, for any $k \in \mathbb{N}$, the equivalence class $[a^k b^k]_{\sim_\mathcal{L}}$ of the word $a^k b^k \in \mathcal{L}$ is

$$[a^k b^k]_{\sim_\mathcal{L}} = \{w \in \{a, b\}^* \mid |w|_a = |w|_b = k\}.$$

Therefore, since

$$\mathcal{L} = \bigcup_{k \geq 0} [a^k b^k]_{\sim_\mathcal{L}}$$

we get that

$$\mathbf{C}l(\varphi(\mathcal{L})) = \mathbf{C}l(\varphi(L)), \tag{4.19}$$

where L is the linear language $\{a^k b^k \mid k \in \mathbb{N}\}$. Finally, since $\mathbf{C}l(\varphi(L))$ is effective semialgebraic for Corollary 3.15, we get the decidability of the $(\mathcal{L}, \mathcal{Q})$ Intersection problem from Equation (4.19) and Proposition 3.6.

It is possible to generalize the argument above to all Dyck languages, i.e., to D_m^* for any $m \in \mathbb{N}$. Let $m \in \mathbb{N}$ be a positive integer and consider two alphabets of m letters $A_m = \{a_1, \ldots, a_m\}$ and $\bar{A}_m = \{\bar{a}_1, \ldots, \bar{a}_m\}$; for simplicity of notation, we will denote each $\bar{a}_i$ as b_i. As explained in Chapter 1, the Dyck language D_m^* is the language in which each word w has, for any $i = 1, \ldots, m$, the same number of occurrences of a_i and b_i, i.e.,

$$w \in D_m^* \iff |w|_{a_i} = |w|_{b_i}, \quad i = 1, \ldots, m.$$

Consider now, for each $i = 1, \ldots, m$, the linear context-free language $L_i = \{a_i^{n_i} b_i^{n_i} \mid n_i \in \mathbb{N}\}$ and let L be the metalinear language

$$L = L_1 \cdots L_m = \{a_1^{k_1} b_1^{k_1} \cdots a_m^{k_m} b_m^{k_m} \mid k_1, \ldots, k_m \in \mathbb{N}\}.$$

In that case, for each word $w \in D_m^*$, there exists a tuple of m non-negative integers $(k_1, \ldots, k_m) \in \mathbb{N}^m$ such that

$$w \in [a_1^{k_1} b_1^{k_1} \cdots a_m^{k_m} b_m^{k_m}]_{\sim_{D_m^*}};$$

thus, we obtain that

$$D_m^* = \bigcup_{(k_1, \ldots, k_m) \in \mathbb{N}^m} [a_1^{k_1} b_1^{k_1} \cdots a_m^{k_m} b_m^{k_m}]_{\sim_{D_m^*}}.$$

Hence, under our assumptions on the morphism (3.4), we get

$$\mathbf{C}l(\varphi(D_m^*)) = \mathbf{C}l(\varphi(L)). \tag{4.20}$$

Again, since $\mathbf{C}l(\varphi(L))$ is effective semialgebraic for Corollary 3.18, we get the decidability of the $(D_m^*, \mathcal{Q})$ Intersection problem for any $m \in \mathbb{N}$ from Equation (4.20) and Proposition 3.6. The same argument shows the decidability of the Intersection problem also for the restricted Dyck languages $D_m'^*$.

Remark 4.13. It is worth noting that the argument above can be generalized to any quantum finite automaton $\mathcal{Q}$ with n basis states where the morphism (3.4) maps the monoid generated by the input alphabet Σ onto an abelian subgroup of O_n.

The commutative equivalence of languages

Parikh equivalence of languages is conceptually connected with the *commutative equivalence relation* described in [6]. Here, two languages L and L' over the same alphabet A are said to be *commutatively equivalent* if there exists a bijection $f : L \to L'$ such that, for every word $u \in L$, $u \sim f(u)$; thus, languages that are commutatively equivalent are Parikh equivalent but the vice-versa is not true. In the survey [6] we treated the relevance of this notion in Theoretical Computer Science. In fact, the study of the commutative equivalence of languages has significant applications in Theory of Codes and Formal Language Theory. In Theory of Codes, this notion is involved in the construction of (variable-length unique factorization) codes satisfying specific properties required in information transmission processes. In this setting, a well known conjecture (formulated by M. P. Schützenberger in 1956, [46]) states that every finite code is commutatively equivalent to a prefix one (*cf* [6, Sections 2, 3]). In Formal Language Theory, it is involved in the study of the asymptotic behaviour of the counting functions of languages. For an arbitrary language $L \subseteq A^*$, the *counting function* of L is the map that associates to any non-negative integer n the number of words in L with length n. Thus, since two commutatively equivalent languages have the same counting function, it is natural to ask under which conditions, given a context-free language L, this is *commutatively regular*, i.e. it can be effectively computed a finite automaton $\mathcal{A}$ such that the language $L_{\mathcal{A}}$ recognized by $\mathcal{A}$ is commutatively equivalent to L. Throughout the paper [6], we surveyed both properties on languages

and conditions on context-free grammars ensuring that, languages satisfying these properties and languages generated by such grammars respectively, are commutatively regular (see [6, Sections 4, 5, 6]).

Concluding remarks and open problems

We conclude by addressing some questions arising from this study and formulated in [7, 8].

[α] One could extend Proposition 4.6 to a broader class of quantum automata. In this context, the construction of the decidability procedure would require the computation of the Zariski closure of a group G of matrices that is *algebraically presented*, i.e. G admits a set of generators that is algebraic itself. It may be of interest to investigate conditions ensuring the extension of the algorithms of [19, 26, 40] for such groups.

[β] It is open whether the $(L, \mathcal{Q})$ Intersection problem is decidable for context-free languages. We conjecture that the answer to this problem is negative. An approach may be to reduce the problem to a so called *"birdie problem"* [47] (cf also [25], Ch. 8). In such a problem, assuming the feasibility of the computation of the Zariski closure of the groups above, one gets, as a consequence, the decidability of some unsolvable problem.

[γ] The $(L, \mathcal{Q})$ Intersection problem is decidable for bounded semi-linear languages. It may be of interest to get extensions of this result for the broader class of languages accepted by the reversal bounded non deterministic counter machines [29]. This extension could exploit a characterization of such languages provided by Ibarra in [30] in terms of some special combinatorial systems called RLGMC-grammars (Right-Linear Grammar with Multi-Counters). Essentially, a RLGMC-grammar is a regular grammar that associates with each derivation a weight, i.e. a vector in the additive monoid $\mathbb{N}^d$. Then a derivation is considered successful if the corresponding weight belongs to the monoid generated by $\mathbf{1^d} = (1, 1, \ldots, 1)$. Here the main crux is

the construction of a formula (see Proposition 3.6) based on the computation of the Zariski closure of the intersection of two finitely generated monoids of matrices. The intrinsic difficulty to *control* the accumulation points of the set – i.e., to exclude those one not coming from the effective functioning of the machines – makes such a computation a non trivial task.

Bibliography

[1] A. Ambainis, M. Beaudry, M. Golovkins, A. Kikusts, M. Mercer, and D. Thérien. *Algebraic results on quantum automata*. Theory of Computing Systems, 39(1):165–188, 2006.

[2] A. Ambainis, A. Yakaryilmaz. Automata and quantum computing. Handbook of Automata Theory, vol. II, pp. 1457–1493, European Mathematical Society Publishing House, 2021.

[3] A. V. Anisimov, F. D. Seifert. *Zur algebraischen Charakteristik der durch Kontext-freie Sprachen definierten Gruppen*. Elektron. Inform. Verarb. u. Kybern. 11, 695-702, 1975.

[4] G. Baron, W. Kuich, The Characterization of Nonexpansive Grammars by Rational Power Series, *Information and Control,* 48, 109–118 1981.

[5] S. Basu, R. Pollack, and M. -F. Roy. *Algorithms in Real Algebraic Geometry*. Springer, Berlin, 2003.

[6] A. Benso, A. Carpi, F. D'Alessandro. On the commutative equivalence of algebraic structures and related problems, *Journal of Automata, Languages and Combinatorics*, Vol. 30, pp. 27–48, 2025.

[7] A. Benso, F. D'Alessandro, and P. Papi. Quantum automata and languages of finite index, in Proceedings of RP 2024, 26th Conference on Reachability Problems, Lecture Notes in Computer Science, Vol. 15050, pp. 88–103, Springer, Berlin, 2024.

[8] A. Benso, F. D'Alessandro, and P. Papi. On the Intersection Problem for Quantum Finite Automata, *Theoretical Computer Science*, Vol. 1053, 115454, 2025.

[9] J. Berstel. *Transductions and Context-Free Languages*. Teubner, Stuttgart, 1979.

[10] A. Bertoni, C. Mereghetti, and B. Palano. Quantum Computing: 1-Way Quantum Automata, in Proceedings of DLT 2003, Developments in Language Theory, Lecture Notes in Computer Science, Vol. 6224, pp. 1-20, Springer, Berlin, 2003.

[11] A. Bertoni, C. Choffrut, and F. D'Alessandro. Quantum finite automata and linear context-free languages, in Proceedings of DLT 2013, Developments in Language Theory, Lecture Notes in Computer Science, Vol. 7907, pp. 82-93, Springer, Berlin, 2013.

[12] A. Bertoni, C. Choffrut, and F. D'Alessandro. On the decidability of the intersection problem for quantum automata and context-free languages, Internat. J. Found. Comput. Sci., 25, 1065-1081, 2014.

[13] V. D. Blondel, E. Jeandel, P. Koiran, and N. Portier. *Decidable and Undecidable Problems about Quantum Automata*. SIAM J. Comput., 34, 1464-1473, 2005.

[14] V. D. Blondel, V. Canterini. Undecidable *Problems for Probabilistic Automata of Fixed Dimension*. Theory Comput. Systems, 36, 231–245 (2003).

[15] A. Candeloro, C. Mereghetti, B. Palano, S. Cialdi, M.G.A. Paris, S. Olivares. *An enhanced photonic quantum finite automaton*. Applied Sciences, 11(18), 8768, 2021.

[16] C. Choffrut, F. D'Alessandro and S. Varricchio. Bounded rational languages of trace monoids, Theory of comput. sys., 46, 351-369, 2010.

[17] J. Dassow, G. Paun. *Regulated Rewriting in Formal Language Theory*. EATCS Monographs in Theoretical Computer Science, vol. 18, Springer, 1989.

[18] A. de Luca, F. D'Alessandro, Teoria degli Automi Finiti, Springer, 2013.

[19] H. Derksen, E. Jeandel, and P. Koiran. *Quantum automata and algebraic groups*. J. Symb. Comput., 39, 357-371, 2005.

[20] C. C. Elgot, J. E. Mezei. *On Relations Defined by Generalized Finite Automata*. IBM Journal of Research and Development, vol. 9(1), 47-68, 1965.

[21] F. Galuppi, M. Stanojkovski. *Toric varieties from cyclic matrix semigroups*. Rend. Istit. Mat. Univ. Trieste., 17, 1–17, 2021.

[22] M. Geck. *An Introduction to Algebraic Geometry and Algebraic Groups*. Oxford University Press, 2003.

[23] S. Ginsburg, *The mathematical theory of context-free languages*, Mc Graw-Hill, New York, 1966.

[24] S. Ginsburg, E. H. Spanier. *Derivation-bounded languages*. J. Comput. System Sci. 2 (3), 228-250, 1968.

[25] M. A. Harrison. *Introduction to Formal Language Theory*. Addison-Wesley Publishing Co., Reading, Mass 1978.

[26] E. Hrushovski, J. Ouaknine, A. Pouly, and J. Worrell. *On Strongest Algebraic Program Invariants*. J. ACM, 70, 1-22, 2023.

[27] E. Hrushovski, J. Ouaknine, A. Pouly, and J. Worrell. *Polynomial invariants for affine programs*. LICS '18-33rd Annual ACM/IEEE Symposium on Logic in Computer Science, ACM, New York, 2018, pp. 530–539.

[28] O. H. Ibarra. *Simple matrix languages*, Information and Control, 17, 359–394, 1970.

[29] O. H. Ibarra. *Reversal-bounded multicounter machines and their decision problems*, J. ACM, 25, 116–133, 1978.

[30] O. H. Ibarra. *Grammatical characterizations of NPDAs and VPDAs with counters*. Theoret. Comput. Sci., 746, 136-150, 2018.

[31] O. H. Ibarra, I. McQuillan. Techniques for Showing the Decidability of the Boundedness Problem of Language Acceptors, in Proceedings of DLT 2024, Developments in Language Theory, Lecture Notes in Computer Science, Vol.14791, pp.156–172, Springer, Göttingen, 2024.

[32] R. Incitti, The growth function of context-free languages, Theoret. Comput. Sci., 255, 601–605, 2001.

[33] E. Jeandel. *Indécidabilité sur les automates quantiques. Master's thesis.* ENS Lyon, 2002.

[34] A. Kondacs, J. Watrous. *On the power of quantum finite state automata.* In: Proceedings of the 38th Annual Symposium on Foundations of Computer Science, pp. 66–75, 1997.

[35] C. Mereghetti, B. Palano, S. Cialdi, V. Vento, M.G.A. Paris, S. Olivares. Photonic realization of a quantum finite automaton, Physical Review Research, 2(1), 013089, 2020.

[36] C. Mereghetti, B. Palano. Quantum finite automata: From Theory to Practice, SIGACT News, 52(3), 38–59, 2021.

[37] C. Mereghetti, B. Palano, and P. Raucci. *Latvian quantum finite state automata for unary languages.* International Journal of Foundations of Computer Science, 36(3), 419–455, 2025.

[38] C. Moore, J. Crutchfield. *Quantum automata and quantum grammars.* Theoret. Comput. Sci., 237, 275-306, 2000.

[39] M. Nivat. *Transductions des langages de Chomsky.* Ann. Inst. Fourier (Grenoble) 18, fasc. 1, 339-455, 1968.

[40] K. Nosan, A. Pouly, S. Schmitz, M. Shirmohammadi, and J. Worrell. On the computation of the Zariski closure of finitely generated groups of matrices, in Proceedings of ISSAC 2022, International Symposium on Symbolic and Algebraic Computation, ACM 2022, ISBN 978-1-4503-8688-3, pp. 129-138, 2022.

[41] A. Onishchik and E. Vinberg. *Lie Groups and Algebraic Groups.* Springer, Berlin, 1990.

[42] A. Paz. *Introduction to Probabilistic Automata.* Academic Press, New York, 1971.

[43] J. Sakarovitch. *Elements of Automata Theory.* Cambridge University Press, Cambridge, 2009.

[44] A. Salomaa. *On the index of context-free grammars and languages.* Inf. and Control. 14, 474-477, 1969.

[45] P. W. Shor. *Algorithms for quantum computation: discrete logarithms and factoring.* Proceedings 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA, pp. 124-134, 1994.

[46] M. P. Schützenberger. *On an application of semigroup methods to some problems in coding.* In: IRE Trans. on Information Theory I. T. 2, pp.47–60, 1956.

[47] J. S. Ullian. *Partial Algorithm Problems for Context Free Languages.* Inf. and Control. 11, 80-101, 1967.