

MSc

2.º
CICLO

FCUP
2021

U. PORTO

Syntax of subgroups of a free group

Inês Ferreira Guimarães

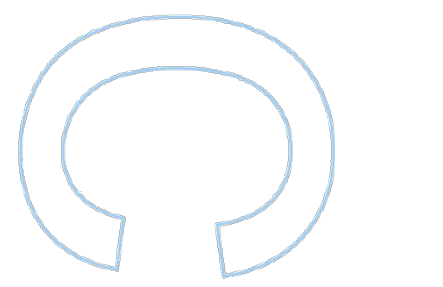
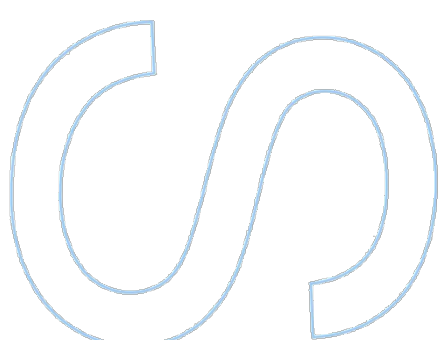
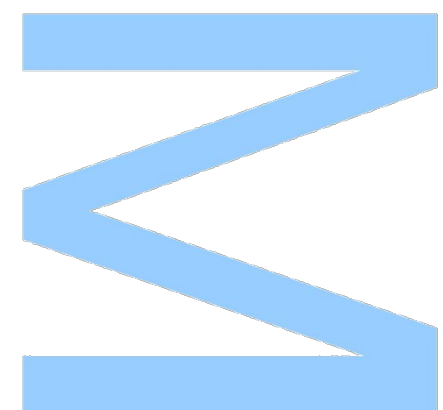
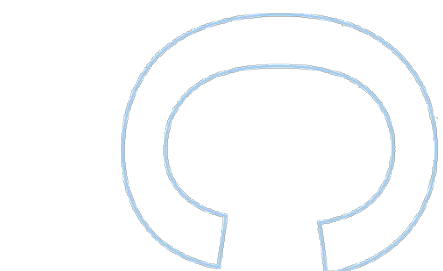
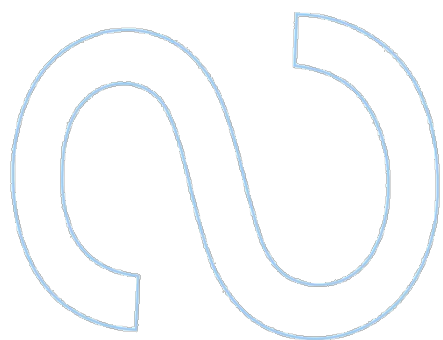
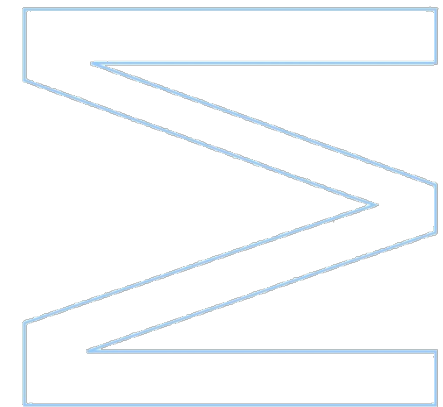
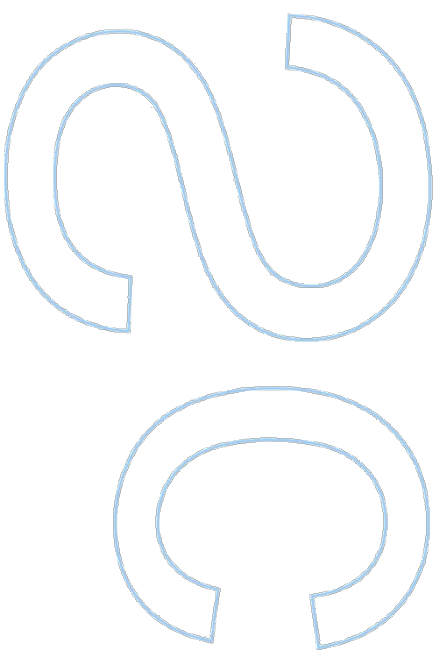
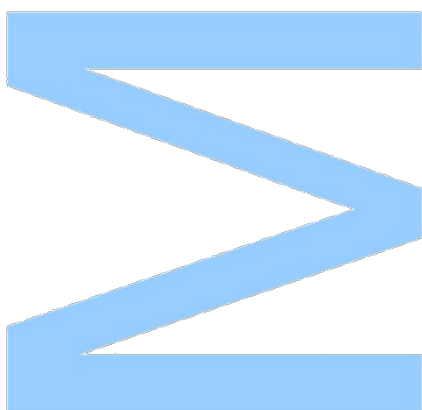
FC



Syntax of subgroups of a free group

Inês Ferreira Guimarães

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto em
Matemática
2021



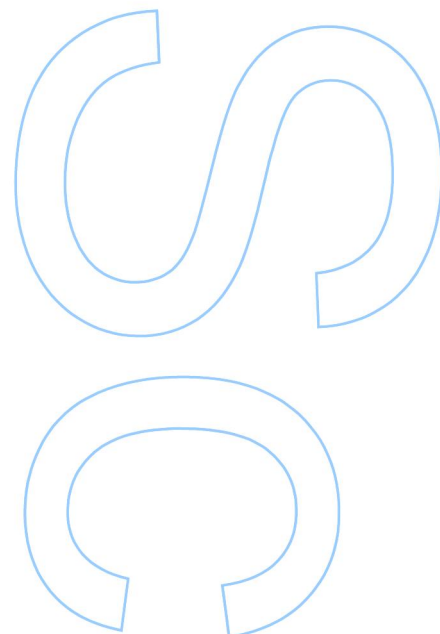
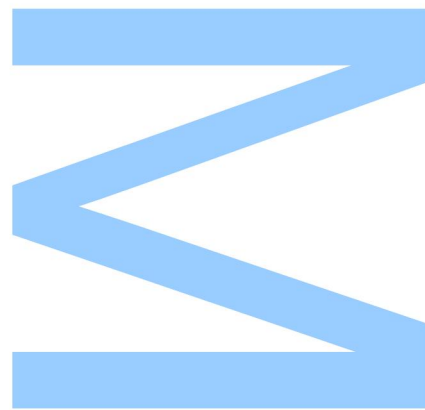
Syntax of subgroups of a free group

Inês Ferreira Guimarães

Mestrado em Matemática
Departamento de Matemática
2021

Orientador

Pedro Ventura Alves da Silva, Professor Catedrático,
Faculdade de Ciências da Universidade do Porto

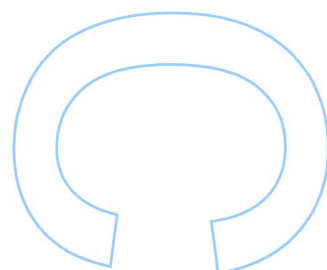
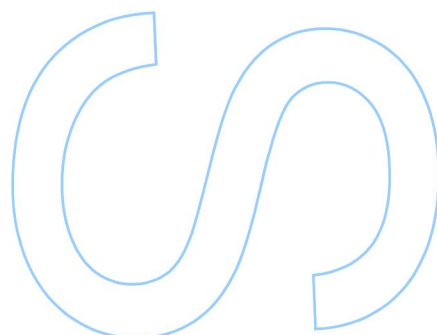
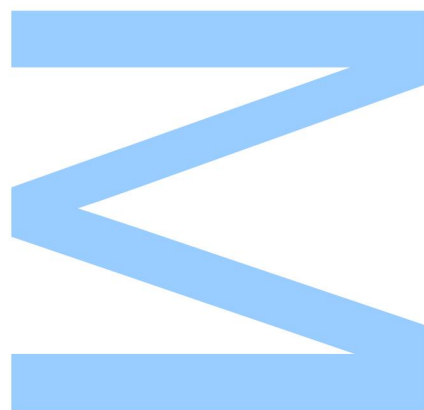




Todas as correções determinadas
pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____/____/____



To José Freitas and João Oliveira

Acknowledgments

First, I would like to thank my wonderful supervisor Pedro Silva for all the encouragement and dedication regarding this monograph. I feel very grateful for the opportunity to work with such a nice and funny person, with a talent for explaining and solving everything in the most clear, succinct and elegant way possible.

Second, I thank my parents for all their support and guidance throughout the years. They are the most important people in the world for me. I would also like to thank my brother and professor Samuel Lopes for showing me that scientific research can be an exciting endeavour.

More broadly, I feel extremely lucky to be surrounded by such an amazing group of professors, family members and friends. I am particularly grateful to João Matias, Bernardo Cunha, Miguel Barata, Rodrigo Girão and Nuno Arala for patiently listening to my complaints about life in college and making me feel understood. I would like to thank Luís Neves and José Nuno Oliveira for inspiring me to think deeply, to be brave and to stand up for my beliefs.

The creators of the GAP packages “Semigroups” and “SmgViz” also deserve an acknowledgment. These computational tools allowed me to test some conjectures and develop my intuition regarding the problem in hand.

Last but not least, I must thank my dog Delta for all the productive walks we took together. It was in those moments that I had my best ideas.

Abstract

In the study of finitely generated subgroups of free groups, it has been very productive to use structures known as Stallings automata [25], [11], [2]. The idea is to associate with each finitely generated subgroup $K \leq F$ of a free group F a specific finite inverse automaton $\mathcal{S}(K)$, called the Stallings automaton of K . One can then extract group-theoretic properties of K from automata properties of $\mathcal{S}(K)$. In [3], the authors establish a first connection between algebraic properties of a finitely generated subgroup K and algebraic properties of the transition monoid $M(K)$ of $\mathcal{S}(K)$. They prove that K is pure (respectively, p -pure) if and only if $M(K)$ is aperiodic (respectively, p -periodic). This dissertation takes inspiration from that paper to investigate further connections of this kind. Given a pseudovariety of finite groups $\mathbf{H}$, we denote by $\overline{\mathbf{H}}$ the pseudovariety of monoids all of whose subgroups lie in $\mathbf{H}$. We generalize the results obtained in [3], and characterize finitely generated subgroups K of a free group when $M(K) \in \overline{\mathbf{H}}$, for certain pseudovarieties of groups $\mathbf{H}$. We also provide a characterization of normal, malnormal and cyclonormal subgroups of a free group in terms of the transition monoids of their Stallings automata.

Keywords: free group, Stallings automaton, transition monoid

Resumo

Para estudar subgrupos finitamente gerados de um grupo livre, tem sido produtivo recorrer a estruturas conhecidas por autómatos de Stallings [25], [11], [2]. A ideia é associar a cada subgrupo finitamente gerado $K \leq F$ de um grupo livre F um determinado autómato inverso finito $\mathcal{S}(K)$, designado por autómato de Stallings de K . Com isto, é possível extrair propriedades do subgrupo K a partir de propriedades do autómato $\mathcal{S}(K)$. Em [3], os autores estabelecem uma primeira ligação entre propriedades algébricas de um subgrupo finitamente gerado K e propriedades algébricas do monóide de transição $M(K)$ de $\mathcal{S}(K)$. Eles provam que K é puro (respetivamente, p -puro) se e só se $M(K)$ é aperiódico (respetivamente, p -periódico). Esta dissertação é inspirada no referido artigo e tem como objetivo investigar mais conexões deste tipo. Dada uma pseudovarietade de grupos finitos $\mathbf{H}$, denotamos por $\overline{\mathbf{H}}$ a pseudovarietade de monóides cujos subgrupos pertencem a $\mathbf{H}$. Generalizamos os resultados obtidos em [3] e caracterizamos subgrupos finitamente gerados K de um grupo livre quando $M(K) \in \overline{\mathbf{H}}$, para certas pseudovarietades de grupos $\mathbf{H}$. Caracterizamos ainda subgrupos normais, malnormais e ciclonormais de um grupo livre em termos dos monóides de transição dos seus autómatos de Stallings.

Palavras-chave: grupo livre, autómato de Stallings, monóide de transição

Contents

1	Introduction	1
2	General monoid theory	3
2.1	Basic terminology	3
2.2	Transformation monoids	6
2.3	Free monoids and presentations	7
2.4	Green's relations	9
2.5	Regularity	13
2.6	Semilattices and partial orders	14
3	Inverse monoids	15
3.1	Basic properties	15
3.2	Natural partial order	17
3.3	Green's relations	19
3.4	Symmetric inverse monoids	19
3.5	Free inverse monoids	22
4	Automata and rational languages	25
4.1	Rational languages	25
4.2	Graphs and automata	26
4.3	Equivalent models	28
4.4	Kleene's Theorem	31
4.5	The transition monoid of an automaton	34
4.6	The syntactic monoid of a language	36
4.7	Minimal automata	39
4.8	Inverse automata	42
5	Varieties and pseudovarieties	45
5.1	Pseudovarieties of monoids	45
5.1.1	Examples	48
5.2	Varieties of languages	50
5.3	Pseudovarieties of algebras	53
6	Free groups and Stallings' construction	59
6.1	Basic properties of free groups	59
6.2	Decidability problems	63
6.3	Stallings' construction	65
6.4	Basic applications	67
7	Transition monoids of Stallings automata	73
7.1	Automorphisms and Stallings automata	73
7.2	Transition groups	81

7.3 Pseudovarieties of type $\overline{\mathbf{H}}$	83
7.3.1 Some characterizations	90
7.4 Non-examples	94
7.5 Conjugacy conditions	94
8 Conclusion	105
Bibliography	107
Index	109
Notation	113

1 Introduction

Group theory is a vast and beautiful subject that started being developed in the nineteenth century. By abstracting the notion of symmetry, groups not only appear in almost every branch of mathematics, namely geometry, topology and analysis, but also in physics, chemistry and material science.

Free groups are particularly ubiquitous in mathematics and play a central role in the study of groups, since every group is isomorphic to a quotient of some free group. The study of free groups from an algebraic point of view essentially began with the work of Jakob Nielsen [17]. He proved, among other results, that every finitely generated subgroup (which we abbreviate to “f.g. subgroup”) of a free group is itself free. Nielsen also published several results regarding the automorphism group of a free group.

In fact, since the early stages of group theory, mathematicians have been interested in understanding the subgroup structure of free groups. To this end, John Stallings [25] presented in a paper of 1983 a revolutionary approach. He developed a way of associating with each finitely generated subgroup of a free group a finite labeled graph, under the formalism of graph immersions. These graphs became known as Stallings automata and constitute a most powerful tool for studying f.g. subgroups of a free group.

Some years later, Alexei Miasnikov and Ilya Kapovich gave Stallings’ construction a more combinatorial flavour and collected numerous group-theoretic properties of f.g. subgroups of free groups based on combinatorial properties of their Stallings automata [11]. Besides being a very elegant theory, this approach displayed great benefits from an algorithmic point of view and had a great impact in tackling several decidability problems regarding free groups.

Additionally, the paper [3] written by Birget, Margolis, Meakin and Weil unveiled a new way of extracting properties of a finitely generated subgroup K of a free group, this time by looking at algebraic properties of its Stallings automaton $\mathcal{S}(K)$, namely its transition monoid $M(K)$. In this article, the authors prove that the condition of a f.g. subgroup K of a free group being pure (respectively, p -pure) is equivalent to the condition of $M(K)$ being aperiodic (respectively, p -periodic). Clearly, there are important instances of pseudovarieties of monoids involved in this result.

Indeed, finite automata and finite monoids are deeply related and constitute important objects not only in mathematics but also in computer science. Moreover, varieties and pseudovarieties contribute greatly to the classification of finite monoids and rational languages. Hence, given a f.g. subgroup K of a free group, there is a strong motivation to find further pseudovarietal properties of $M(K)$ (other than being aperiodic or p -periodic) that constitute algebraic properties of the subgroup K . More concretely, let $K \leq F$ be a finitely generated subgroup of a free group F . For what pseudovarieties of (inverse) monoids $\mathbf{V}$ can we guarantee that $M(K) \in \mathbf{V}$ implies $M(K\varphi) \in \mathbf{V}$ for every automorphism φ of F ? And what do those well-behaved pseudovarieties tell us about K itself? That is the main goal of this monograph.

Furthermore, inspired by the characterizations of normal, malnormal and cyclonormal subgroups in terms of their Stallings automata presented in [11], we were also motivated to

try to characterize these conjugacy conditions by inspecting the structure of the transition monoids of Stallings automata.

This dissertation is organized as follows. In Section 2 we present some basic concepts and results regarding monoids in general. We briefly discuss important classes of monoids such as transformation monoids, free monoids, regular monoids and semilattices. We also talk about Green's relations as a powerful tool for studying a monoid's structure. In Section 3 we focus on a specific class of monoids called inverse monoids, which will be very important to us since the transition monoid of a Stallings automaton is always inverse. Our main references were [9] and [18].

We proceed by discussing some fundamental results regarding automata and rational languages in Section 4. In particular, we define the crucial concept of transition monoid of an automaton and also introduce an important class of automata known as inverse automata; later we shall see that Stallings automata are inverse automata. This section was essentially based on [19].

Section 5 deals with pseudovarieties of finite algebras (namely monoids, groups and inverse monoids) and varieties of languages. We state and prove a major result known as Eilenberg's correspondence, which gives a way of classifying certain rational languages through their syntactic monoids. Even though this theorem will not play a role in understanding the main results of this monograph, we consider that it sheds light on why (pseudo)varieties are so important. Besides [19], in this section we relied on [18] and [1].

In Section 6, we finally focus on free groups and briefly discuss their automorphism groups. Next, to motivate Stallings' approach, we state some decidability problems concerning free groups. We present at last the construction and some basic applications of the Stallings automaton of a f.g. subgroup of a free group. These automata will be on the spotlight until the very end. Our main references for this section were [15], [11] and [2].

Finally, Section 7 is concerned with original work. We generalize the aforementioned results by Birget, Margolis, Meakin and Weil [3]. Despite the fact that our main focus was to find pseudovarieties of (inverse) monoids that are stable under all automorphisms of a free group, we also collect some well-known pseudovarieties that do not satisfy such property. We end by classifying some conjugacy conditions on f.g. subgroups based on the transition monoids of their Stallings automata.

In order for this dissertation to be accessible to a wider audience, the reader is only assumed to be familiar with some basic group theory.

2 General monoid theory

The theory of semigroups and monoids started being properly developed in the second half of the 20th century and has found deep connections with several other areas of mathematics ever since. In fact, besides classical fields such as geometry, number theory and functional analysis, finite semigroups are of great interest to computer science due to their connection with automata and language theory. In this monograph, monoids are one of the algebraic structures that are going to be in the spotlight, so we begin by introducing some basic definitions and results concerning them. In the current chapter, we will not focus on presenting concrete examples, but the concepts here shall be illustrated later on. For more details, we refer the reader to [9].

2.1 Basic terminology

A **semigroup** is a set equipped with an associative binary operation, often called product. A **monoid** is a semigroup with an identity element, usually denoted by 1. Many of the results we present in this section are valid in arbitrary semigroups and not only in monoids. However, since monoids constitute one of the main characters of this dissertation, whereas arbitrary semigroups will only be mentioned sporadically, we shall sacrifice some generality for clarity and conciseness.

Besides the identity element, which is unique, a monoid M may have another “special” element: $x \in M$ is called a **zero** if $xy = yx = x$ for all $y \in M$. If such an element exists, it is the only one with that property, and we usually denote it by 0. Furthermore, an element $x \in M$ satisfying $x^2 = x$ is called an **idempotent**, and we denote by $E(M)$ the set of all idempotents of M .

Given a monoid M and a nonempty subset $S \subseteq M$, we say that S is a **subsemigroup** of M if S is a semigroup on its own, considering that the operation on S is the restriction to S of the operation on M . If, in addition, we have $1 \in S$, then we call S a **submonoid** of M and write $S \leq M$ (or $S < M$ if the inclusion is proper). Moreover, if $S \leq M$ is group, we say that S is a **subgroup** of M . A nonempty subset $I \subseteq M$ is called a **right ideal** if $IM \subseteq I$; a **left ideal** if $MI \subseteq I$; and an **ideal** if it is both a right and left ideal, which is equivalent to the condition $MIM \subseteq I$.

Given two monoids M and N , a function $\varphi: M \rightarrow N$, $x \mapsto x\varphi$ is a **homomorphism** (or simply a **morphism**) if $1\varphi = 1$ and $(xy)\varphi = (x\varphi)(y\varphi)$ for all $x, y \in M$; the former condition means that the identity of M must be mapped to the identity of N . If φ is a bijective homomorphism, then we call it an **isomorphism**, and whenever there exists an isomorphism between the monoids M and N , we say that they are isomorphic and write $M \cong N$. Furthermore, we say that N is a **quotient** of M if there exists a surjective morphism $\varphi: M \rightarrow N$.

Yet there is another way to look at monoid quotients that better captures the idea of “lumping” elements together. If $\sigma \subseteq M \times M$ is an equivalence relation on M , we denote by $x\sigma$ the equivalence class of $x \in M$ and write $M/\sigma = \{x\sigma \mid x \in M\}$. When two elements

$x, y \in M$ lie in the same equivalence class, we sometimes write $x\sigma y$ instead of $x\sigma = y\sigma$. If σ satisfies the additional condition

$$(x\sigma y \wedge x'\sigma y') \Rightarrow (xx')\sigma(yy')$$

for all $x, x', y, y' \in M$, we say that σ is a **congruence** on M . This essentially means that the equivalence relation σ is compatible with the product on M , which allows us to define a monoid structure on the set M/σ . Indeed, the (well-defined) associative binary operation on M/σ is given by $(x\sigma)(y\sigma) = (xy)\sigma$, and the identity element is 1σ . In the next proposition, we shall verify that M/σ is a quotient of M and that any quotient is of this kind.

Given two sets M and N , we define the **kernel** of a function $\varphi: M \rightarrow N$ to be the equivalence relation on M given by

$$\ker \varphi = \{(x, y) \mid x\varphi = y\varphi\}.$$

If M, N are monoids and φ is a homomorphism, then $\ker \varphi$ is a congruence on M . Indeed, observe that an equivalence relation $\sigma \subseteq M \times M$ is a congruence if and only if the implication

$$x\sigma y \Rightarrow ((xz)\sigma(yz) \wedge (zx)\sigma(zy))$$

holds for all $x, y, z \in M$. Suppose, then, that $(x, y) \in \ker \varphi$ and $z \in M$. Since $x\varphi = y\varphi$ and φ is a homomorphism, we get

$$(xz)\varphi = (x\varphi)(z\varphi) = (y\varphi)(z\varphi) = (yz)\varphi,$$

so $(xz, yz) \in \ker \varphi$. We prove $(zx, zy) \in \ker \varphi$ in a similar manner. Hence, $\ker \varphi$ is a congruence on M , and we can speak of the quotient monoid $M/\ker \varphi$.

Proposition 2.1. *Let M be a monoid and let σ be a congruence on M . Then*

$$\begin{aligned} \pi: M &\longrightarrow M/\sigma \\ x &\longmapsto x\sigma \end{aligned}$$

is a surjective homomorphism. Moreover, if N is a monoid and $\varphi: M \rightarrow N$ is a surjective homomorphism, then $N \cong M/\ker \varphi$.

Proof. By definition of M/σ , it is clear that π is surjective. Given $x, y \in M$, we have

$$(xy)\pi = (xy)\sigma = (x\sigma)(y\sigma) = (x\pi)(y\pi),$$

and $1\pi = 1\sigma$ is the identity of M/σ , so π is a homomorphism. For the second part, let $\varphi: M \rightarrow N$ be a surjective homomorphism, write $\sigma = \ker \varphi$ and consider the function given

by

$$\begin{aligned}\phi: M/\sigma &\longrightarrow N \\ x\sigma &\longmapsto x\varphi.\end{aligned}$$

Observe that

$$x\sigma = y\sigma \Leftrightarrow (x, y) \in \ker \varphi \Leftrightarrow x\varphi = y\varphi,$$

and thus ϕ is well defined and injective. Finally, since φ is a surjective homomorphism, so is ϕ . Hence, ϕ is an isomorphism between M/σ and N , which concludes the proof. $\square$

The following result will be frequently (and, most of the times, implicitly) used in several proofs.

Theorem 2.2 (Homomorphism theorem). *Let σ be a congruence on a monoid M and let $\pi: M \rightarrow M/\sigma$, $x \mapsto x\sigma$ be the quotient morphism. If $\varphi: M \rightarrow N$ is a homomorphism from M to a monoid N and $\sigma \subseteq \ker \varphi$, then there exists a unique homomorphism $\phi: M/\sigma \rightarrow N$ satisfying $\varphi = \pi\phi$.*

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & N \\ \pi \downarrow & \nearrow \phi & \\ M/\sigma & & \end{array}$$

Proof. As in the previous proof, consider the function given by

$$\begin{aligned}\phi: M/\sigma &\longrightarrow N \\ x\sigma &\longmapsto x\varphi.\end{aligned}$$

Since, for all $x, y \in M$,

$$x\sigma = y\sigma \Rightarrow (x, y) \in \ker \varphi \Leftrightarrow x\varphi = y\varphi$$

and φ is a homomorphism, it follows that ϕ is well-defined and a homomorphism. Moreover, given $x \in M$,

$$x\pi\phi = (x\sigma)\phi = x\varphi,$$

and hence the equality $\varphi = \pi\phi$ holds. Regarding uniqueness, if $\phi': M/\sigma \rightarrow N$ is such that $\varphi = \pi\phi'$, then

$$(x\sigma)\phi' = (x\pi)\phi' = x(\pi\phi') = x\varphi = (x\sigma)\phi$$

for all $x \in M$. Therefore, $\phi = \phi'$ and the proof is finished. $\square$

Given a monoid M and a nonempty subset $X \subseteq M$, the submonoid of M **generated by** X is the intersection of all submonoids of M containing X ; we denote it by $\langle X \rangle$. More concretely,

$$\langle X \rangle = \{x_1x_2 \dots x_n \mid n \geq 0; x_1, \dots, x_n \in X\}.$$

We agree that the empty product (the result of multiplying no factors) is 1. We say that X **generates** M if $M = \langle X \rangle$, and M is said to be **finitely generated** when X can be chosen to be finite.

2.2 Transformation monoids

In group theory, Cayley's theorem states that every group is isomorphic to a subgroup of some symmetric group. We now present a similar result concerning monoids.

Given a nonempty set X , let $\mathcal{B}_X$ denote the set of all binary relations on X , which are nothing more than subsets of $X \times X$. Given $\sigma, \tau \in \mathcal{B}_X$, we define the **composite** relation

$$\sigma\tau = \{(x, z) \in X \times X \mid (x, y) \in \sigma \text{ and } (y, z) \in \tau \text{ for some } y \in X\}.$$

It is clear that the composition of relations is an associative operation and that the relation $\text{id}_X = \{(x, x) \mid x \in X\}$ is the identity element. Hence, we have just equipped $\mathcal{B}_X$ with a monoid structure.

Now, $\sigma \in \mathcal{B}_X$ is a **function** if

$$(x, y), (x, z) \in \sigma \Rightarrow y = z$$

for all $x, y, z \in X$. We define $\mathcal{PT}_X \leq \mathcal{B}_X$ to be the submonoid of all binary relations on X which are functions; its elements are called **partial transformations** on X . Given $\sigma \in \mathcal{B}_X$, we define the **domain**, **range** and **converse** of σ by

$$\begin{aligned} \text{dom } \sigma &= \{x \in X \mid (x, y) \in \sigma \text{ for some } y \in X\} \\ \text{ran } \sigma &= \{y \in X \mid (x, y) \in \sigma \text{ for some } x \in X\} \\ \sigma^{-1} &= \{(y, x) \in X \times X \mid (x, y) \in \sigma\}, \end{aligned}$$

respectively. Note that $(\sigma^{-1})^{-1} = \sigma$ and $\text{dom } \sigma^{-1} = \text{ran } \sigma$. Given $\alpha, \beta \in \mathcal{PT}_X$, it is easy to see that

$$\begin{aligned} \text{dom } \alpha\beta &= (\text{ran } \alpha \cap \text{dom } \beta)\alpha^{-1} \\ \text{ran } \alpha\beta &= (\text{ran } \alpha \cap \text{dom } \beta)\beta. \end{aligned}$$

When a partial transformation on X has domain X , we call it a **transformation** on X . We write $\mathcal{T}_X$ to denote the submonoid of $\mathcal{PT}_X$ consisting of all transformations on X , that is, $\mathcal{T}_X = \{\varphi \in \mathcal{PT}_X \mid \text{dom } \varphi = X\}$. Given $\varphi \in \mathcal{T}_X$, we refer to the range of φ as the **image** of φ and denote it by $\text{im } \varphi$. In case $\varphi: M \rightarrow N$ is a monoid homomorphism, it is easy to check that $M\varphi \leq N$, which means that the homomorphic image of a monoid is a submonoid of the codomain.

We are now ready to state the aforementioned theorem.

Theorem 2.3. *Let M be a monoid. Then there exists a nonempty set X such that M is isomorphic to a submonoid of $\mathcal{T}_X$.*

Proof. We choose $X = M$, seen as a set. Consider the function

$$\begin{aligned}\rho: M &\longrightarrow \mathcal{T}_M \\ x &\longmapsto \rho_x,\end{aligned}$$

where $\rho_x: M \rightarrow M$ is given by $y \mapsto yx$. Since $\text{im } \rho \leq \mathcal{T}_M$ and $\text{im } \rho \cong M/\ker \rho$, it suffices to show that ρ is an injective homomorphism. Given $x, y \in M$, we have $z\rho_{xy} = z(xy)$ and $z\rho_x\rho_y = (zx)y$ for every $z \in M$. As the operation on M is associative, we get $\rho_{xy} = \rho_x\rho_y$; and as ρ_1 is the identity function on M , we conclude that ρ is a homomorphism. Finally, suppose that $(x, y) \in \ker \rho$. Then, in particular, $1\rho_x = 1\rho_y$, so $x = y$ and ρ is injective. $\square$

The homomorphism ρ above is known as the **right regular representation** of M .

2.3 Free monoids and presentations

We will now present a construction of a very special kind of monoid in which we have the notion of a basis, just like in the case of vector spaces.

We begin by defining an **alphabet** to be a nonempty set A , and we designate its elements by **letters**. A **word** over the alphabet A is a finite sequence of letters of A . If w is a nonempty word, we may represent it by $w = a_1a_2 \dots a_n$ for some $n \geq 1$ and $a_i \in A$ ($1 \leq i \leq n$), in a unique way. We can then define the **length** of $w = a_1a_2 \dots a_n$ to be $|w| = n$. The **empty word** will be denoted by 1 , and we consider its length to be 0 , so $|1| = 0$.

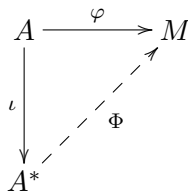
Let A^* be the set of all words over A and A^+ be the set of all nonempty words over A . We consider that $A \subseteq A^*$ by identifying each letter $a \in A$ with the word $a \in A^*$, obtaining thus an injective function $\iota: A \hookrightarrow A^*$. Accordingly, if $\Phi: A^* \rightarrow X$ is a function from A^* to a set X , then $\Phi|_A = \iota\Phi$.

The set A^* can be equipped with an associative binary relation called **concatenation**, given by

$$\begin{aligned}1w &= w = w1 \\ (a_1a_2 \dots a_n)(b_1b_2 \dots b_m) &= a_1a_2 \dots a_nb_1b_2 \dots b_m,\end{aligned}$$

where $m, n \geq 1$, $a_i, b_j \in A$ and $w \in A^*$. Hence, we have just equipped the set A^* with the structure of a monoid (the empty word being its identity) and, as we are about to testify, created a free monoid.

Theorem 2.4 (Universal property of A^*). *Let A be an alphabet and $\varphi: A \rightarrow M$ be a function from A to a monoid M . Then there exists a unique monoid homomorphism $\Phi: A^* \rightarrow M$ such that $\Phi|_A = \varphi$.*



Proof. We begin by showing that if such a homomorphism exists, then it must be unique. Indeed, if $\Phi: A^* \rightarrow M$ is a monoid homomorphism, then necessarily $1\Phi = 1$. Moreover, if $1 \neq w \in A^*$, then there exist $a_1, \dots, a_n \in A$ such that $w = a_1 \dots a_n$. Hence, the image of w is $w\Phi = (a_1\Phi) \cdots (a_n\Phi)$, which must be equal to $(a_1\varphi) \cdots (a_n\varphi)$, since $\Phi|_A = \varphi$. Therefore, there is only one possible way to define Φ , and note that it is indeed well defined because words on A admit a unique factorization into letters. Finally, as the operation on a monoid is associative, Φ is a homomorphism, and the claim follows. $\square$

The following corollary shows that, up to isomorphism, there exists a unique monoid satisfying the universal property mentioned above; therefore, we call A^* the **free monoid** over A . Similarly, A^+ is the **free semigroup** over A .

Corollary 2.5. *Given a nonempty set A , let F be a monoid such that $A \subseteq F$ and, for every function $\varphi: A \rightarrow M$ from A to a monoid M , there exists a unique monoid homomorphism $\Phi: F \rightarrow M$ extending φ . Then $F \cong A^*$.*

Proof. Let $\iota: A \hookrightarrow A^*$ and $\iota': A \hookrightarrow F$ denote the inclusion morphisms. Since both A^* and F satisfy the universal property with respect to A , there exist (unique) homomorphisms $\Phi: A^* \rightarrow F$ and $\Phi': F \rightarrow A^*$ such that $\iota\Phi = \iota'$ and $\iota'\Phi' = \iota$. It follows that $\iota\Phi\Phi' = \iota'\Phi' = \iota$, whence $\Phi\Phi': A^* \rightarrow A^*$ and id_{A^*} both extend ι . The uniqueness part of the universal property regarding A^* thus implies that $\Phi\Phi' = \text{id}_{A^*}$. We show that $\Phi'\Phi = \text{id}_F$ in a similar manner. Hence, Φ and Φ' are mutually inverse isomorphisms, yielding $F \cong A^*$ as we wanted. $\square$

It is important to observe that every monoid is isomorphic to a quotient of a free monoid. Indeed, let M be a monoid and let A be any set for which there exists a bijection $\varphi: A \rightarrow M$. By the universal property, there exists a homomorphism $\Phi: A^* \rightarrow M$ that satisfies $\Phi|_A = \varphi$. The surjectivity of φ entails the surjectivity of Φ , and hence $M \cong A^*/\ker \Phi$. This is why free monoids play such a unique and important role in monoid theory.

We mention one more result that follows from the universal property.

Proposition 2.6. *Let A be an alphabet and $\psi: A^* \rightarrow M$ be a homomorphism from A^* to a monoid M . If $\eta: N \rightarrow M$ is a surjective homomorphism from a monoid N to M , then there exists a homomorphism $\phi: A^* \rightarrow N$ satisfying $\phi\eta = \psi$.*

$$\begin{array}{ccc} & & N \\ & \nearrow \phi & \downarrow \eta \\ A^* & \xrightarrow{\psi} & M \end{array}$$

Proof. For each $a \in A$, the surjectivity of η allows us to pick $a\varphi \in a\psi\eta^{-1}$ in order to get a map $\varphi: A \rightarrow N$, $a \mapsto a\varphi$. In view of the universal property of A^* , we can extend φ to a homomorphism $\phi: A^* \rightarrow N$ satisfying $\phi\eta = \psi$, which concludes the proof. $\square$

We end this section by briefly mentioning monoid presentations. Given a monoid M and a set $R \subseteq M \times M$, there exists at least one congruence on M containing R , namely $M \times M$.

We define the **congruence generated by** R to be the intersection of all congruences on M containing R , and we denote it by $R^\sharp$. A monoid **presentation** is a formal expression of the type $\text{Mon}\langle A \mid R \rangle$, where A is an alphabet and $R \subseteq A^* \times A^*$. We say that $\text{Mon}\langle A \mid R \rangle$ is a presentation of a monoid M if $M \cong A^*/R^\sharp$. By the universal property, every monoid admits a presentation.

2.4 Green's relations

We will now look at some equivalence relations $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, $\mathcal{H}$ and $\mathcal{D}$ on a monoid that are extremely useful to analyse its structure. They are known as **Green's relations**, named after James Alexander Green [8], who introduced them seventy years ago.

Let M be a monoid and $x, y \in M$. We say that x and y are $\mathcal{R}$ -related if they generate the same right ideal:

$$x\mathcal{R}y \Leftrightarrow xM = yM;$$

they are $\mathcal{L}$ -related if they generate the same left ideal:

$$x\mathcal{L}y \Leftrightarrow Mx = My;$$

and they are $\mathcal{J}$ -related if they generate the same ideal:

$$x\mathcal{J}y \Leftrightarrow MxM = MyM.$$

Moreover, we define $\mathcal{H} = \mathcal{R} \cap \mathcal{L}$, so

$$x\mathcal{H}y \Leftrightarrow (x\mathcal{R}y \wedge x\mathcal{L}y),$$

and $\mathcal{D} = \mathcal{R}\mathcal{L}$, so

$$x\mathcal{D}y \Leftrightarrow \exists z \in M (x\mathcal{R}z \wedge z\mathcal{L}y).$$

Some simple computations allow us to verify that $\mathcal{D} = \mathcal{L}\mathcal{R}$ too, making $\mathcal{D}$ the smallest equivalence relation containing $\mathcal{R}$ and $\mathcal{L}$. In fact, all of Green's relations are equivalence relations; $\mathcal{R}$ is even a left congruence, that is,

$$x\mathcal{R}y \Rightarrow \forall z \in M \ zx\mathcal{R}zy,$$

and, similarly, $\mathcal{L}$ is a right congruence, which means that

$$x\mathcal{L}y \Rightarrow \forall z \in M \ xz\mathcal{L}yz.$$

If $\mathcal{K}$ is one of Green's relations, we denote by K_x the equivalence class of $x \in M$.

Observe that, in general, $\mathcal{H} \subseteq \mathcal{R}, \mathcal{L} \subseteq \mathcal{D} \subseteq \mathcal{J}$. In a finite monoid, however, Green's relation $\mathcal{D}$ and $\mathcal{J}$ coincide. Before we prove it, we need an auxiliary (yet extremely important) lemma, which states that in a finite monoid every element has an idempotent power.

Lemma 2.7. *Let M be a finite monoid and $x \in M$. Then there exists $n \geq 1$ such that x^n is an idempotent.*

Proof. Since M is finite, there exist $m > \ell \geq 1$ satisfying $x^{2^m} = x^{2^\ell}$, which yields

$$x^{2^m} = (x^{2^\ell})^{2^{m-\ell}} = (x^{2^m})^{2^{m-\ell}}.$$

Therefore, writing $y = x^{2^\ell}$ and $k = 2^{m-\ell} \geq 2$, we have $y = y^k$. If $k = 2$, then y is a power of x which is an idempotent, and the conclusion follows; otherwise, if $k > 2$, then

$$y^{k-1} = y(y^{k-2}) = y^k(y^{k-2}) = y^{2(k-1)},$$

and thus y^{k-1} is an idempotent power of x . In either case, the statement holds. $\square$

Proposition 2.8. *In a finite monoid, $\mathcal{D} = \mathcal{J}$.*

Proof. We already know that $\mathcal{D} \subseteq \mathcal{J}$, as $\mathcal{R}, \mathcal{L} \subseteq \mathcal{J}$ and $\mathcal{D}$ is the smallest equivalence relation containing $\mathcal{R}$ and $\mathcal{L}$. Thus, we only need to prove the other inclusion. Let M be a finite monoid. Taking $x, y \in M$ such that $x \mathcal{J} y$, there exist $a, b, c, d \in M$ satisfying $x = ayb$ and $y = cxd$. Hence, $x = ayb = (ac)x(db)$ implies $x = (ac)^k x (db)^k$ for all $k \geq 1$. Since M is finite, the previous lemma entails that there exist $m, n \geq 1$ for which $(ac)^m$ and $(db)^n$ are idempotents, whence $(ac)^{mn} = ((ac)^m)^n = (ac)^m$ and $(db)^{mn} = ((db)^n)^m = (db)^n$. Writing $e = (ac)^m$ and $f = (db)^n$, we get $(ac)^{mn} x (db)^{mn} = x = e x f$. Therefore,

$$(ac)^m x = e x = e(e x f) = e x f = x = x f = x (db)^n,$$

which gives $x d \mathcal{R} x$ and $c x \mathcal{L} x$. The last relation implies $c x d \mathcal{L} x d$ since $\mathcal{L}$ is a right congruence. Thus $y \mathcal{L} x d \mathcal{R} x$ entails $y \mathcal{D} x$. Since $x, y \in M$ where arbitrary $\mathcal{J}$ -related elements, we conclude that $\mathcal{J} \subseteq \mathcal{D}$, whence $\mathcal{J} = \mathcal{D}$. $\square$

We will now present some results concerning Green's relations which are of great importance to the study of a monoid's structure.

Proposition 2.9 (Green's Lemma). *Let M be a monoid and $x, y \in M$. If $x \mathcal{R} y$, with $x = ya$ and $y = xb$ for some $a, b \in M$, then*

$$\begin{aligned} \rho_a: L_y &\longrightarrow L_x \\ z &\longmapsto za \end{aligned}$$

is a bijection with inverse

$$\begin{aligned} \rho_b: L_x &\longrightarrow L_y \\ w &\longmapsto wb. \end{aligned}$$

Proof. Taking $z \in L_y$, there exist $c, d \in M$ satisfying $z = cy$ and $y = dz$, whence $za = cya$ and $ya = dza$. Since $x = ya$, it follows that $ya \mathcal{L} za$ and $za \in L_x$, which shows that ρ_a is

a well-defined mapping. Out of symmetry, so is ρ_b . Moreover, $zab = z$ for every $z \in L_y$, because

$$y = xb = yab \Rightarrow z = cy = cyab = zab.$$

A similar reasoning allows us to conclude that $wba = w$ for every $w \in L_x$. Therefore, $\rho_a\rho_b$ and $\rho_b\rho_a$ are the identity functions on L_y and L_x , respectively. $\square$

We obtain an analogous result by considering Green's $\mathcal{L}$ relation in place of $\mathcal{R}$.

Proposition 2.10 (Dual of Green's Lemma). *Let M be a monoid and $x, y \in M$. If $x\mathcal{L}y$, with $x = ay$ and $y = bx$ for some $a, b \in M$, then*

$$\begin{aligned}\lambda_a: R_y &\longrightarrow R_x \\ z &\longmapsto az\end{aligned}$$

is a bijection with inverse

$$\begin{aligned}\lambda_b: R_x &\longrightarrow R_y \\ w &\longmapsto bw.\end{aligned}$$

These two propositions yield the following result concerning $\mathcal{H}$ -classes.

Corollary 2.11. *Let M be a monoid and $x, y \in M$ be such that $x\mathcal{D}y$. Then $|H_x| = |H_y|$.*

Proof. Let $z \in M$ be such that $x\mathcal{R}z\mathcal{L}y$, with $x = za$ and $z = xb$ for some $a, b \in M$. By Green's Lemma, we have mutually inverse bijections $\rho_a: L_z \rightarrow L_x$ and $\rho_b: L_x \rightarrow L_z$. We claim that $\rho_a|_{H_z}$ is a bijection from H_z to H_x and, similarly, that $\rho_b|_{H_x}$ is a bijection from H_x to H_z . Indeed, taking $w \in H_z \subseteq L_z$, we already know that $w\rho_a = wa \in L_x$; hence, it remains to prove that $wa \in R_x$. Now, since $z\mathcal{H}w$, we have $x = za = wca$ for some $c \in M$, and $w = dz = dxb = dzab = wab$ for some $d \in M$. It follows that $x = wca = wabca$ and $wa = zea = xbea$ for some $e \in M$, whence $wa\mathcal{R}x$. We conclude that $H_z\rho_a \subseteq H_x$, and we show that $H_x\rho_b \subseteq H_z$ in a similar manner. Therefore, $\rho_a|_{H_z}$ and $\rho_b|_{H_x}$ are injective mappings into H_x and H_z , respectively.

Since $w = wab$, this entails that $w\rho_a\rho_b = w$, and hence $(\rho_a\rho_b)|_{H_z} = \text{id}_{H_z}$. The equality $(\rho_b\rho_a)|_{H_x} = \text{id}_{H_x}$ holds for an analogous reason, thereby implying that $\rho_a|_{H_z}$ and $\rho_b|_{H_x}$ are mutually inverse bijections, so $|H_x| = |H_z|$. Finally, working with the dual of Green's Lemma, a similar argument allows us to conclude that $|H_y| = |H_z|$, thus proving $|H_x| = |H_y|$. $\square$

Another important result regarding the structure of $\mathcal{H}$ -classes is presented next. Before we state it, a preliminary lemma is useful.

Lemma 2.12. *Let S be a semigroup. Then S is a group if and only if $Sx = S = xS$ for every $x \in S$.*

Proof. It is clear that every group satisfies the condition. Suppose now that S is such that $Sx = S = xS$ for all $x \in S$, and fix some $a \in S$. Since $Sa = S$, there exists $e \in S$ satisfying

$ea = a$. On the other hand, given $x \in S$, there exists $y \in S$ for which $x = ay \in aS$, so

$$ex = e(ay) = (ea)y = ay = x.$$

Similarly we show that there exists $f \in S$ satisfying $xf = x$ for all $x \in S$. But then $e = ef = f$, and thus e is such that $ex = x = xe$ for all $x \in S$, which means that S is a monoid with identity e . Finally, given $a \in S$, there exist $b, c \in S$ such that $ba = e = ac$, whence $c = ec = bac = be = b$. We conclude that every element in S has an inverse, so S is a group. $\square$

Theorem 2.13 (Green's Theorem). *Let M be a monoid and let H be an $\mathcal{H}$ -class of M . Then either $H^2 \cap H \neq \emptyset$ or H is a group.*

Proof. Suppose that $H^2 \cap H = \emptyset$ and $x, y \in H$ are such that $xy \in H$. By the proof of Corollary 2.11, we have bijections $\rho_y, \lambda_x: H \rightarrow H$, given by right multiplication by y and left multiplication by x , respectively. It follows that $Hy = H$ and $xH = H$. Applying this reasoning again, we conclude that, for every $h \in H$, the maps ρ_h and λ_h are also bijections on H . This implies that $H^2 = H$ and $hH = Hh = H$ for every $h \in H$. According to the previous lemma, H is a group. $\square$

Green's Theorem provides an easy way to detect whether or not an $\mathcal{H}$ -class is a group.

Corollary 2.14. *Let M be a monoid and let H be an $\mathcal{H}$ -class of M . Then H is a group if and only if it contains an idempotent.*

Proof. If H is a group, then it has an identity, which is an idempotent. Conversely, if $e \in H$ is an idempotent, then $e \in H^2 \cap H$; thus, by Green's Theorem, H is a group. $\square$

An $\mathcal{H}$ -class containing an idempotent is therefore called a **group $\mathcal{H}$ -class**. We present one last corollary.

Corollary 2.15. *The maximal subgroups of a monoid are precisely the $\mathcal{H}$ -classes containing an idempotent.*

Proof. Let M be a monoid and let $G \leq M$ be a subgroup satisfying $H_e \subseteq G$ for some $e \in E(M)$. Since a group has a unique idempotent (its identity element), e must be the identity of G . Consequently, given $g \in G$, it follows that $g = ge = eg$ and $gg^{-1} = e = g^{-1}g$, so $g \in H_e$. We conclude that $G \subseteq H_e$, whence H_e is a maximal subgroup of M . $\square$

It is common to think of a $\mathcal{D}$ -class as an “**egg-box**”, where a row of eggs is an $\mathcal{R}$ -class; a column of eggs is an $\mathcal{L}$ -class; and an egg compartment is an $\mathcal{H}$ -class. We depict a typical egg-box diagram of a $\mathcal{D}$ -class in Figure 1.

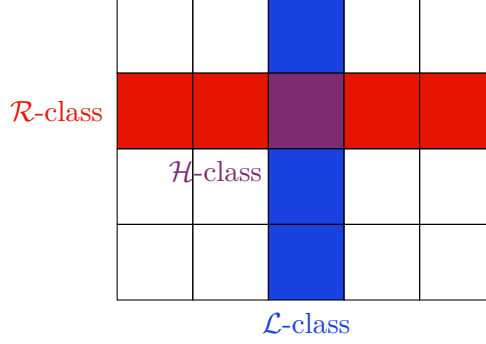


Figure 1: Egg-box diagram of a $\mathcal{D}$ -class

2.5 Regularity

In this section and in the succeeding one, we will equip our monoids with some extra properties, making them closer and closer to groups. We begin by inspecting a generalization of the concept of inverse element in a group.

Let M be a monoid and $x \in M$. We say that x is a **regular element** if there exists $y \in M$ such that $xyx = x$. The monoid M is said to be **regular** if all its elements are. Furthermore, $y \in M$ is called an **inverse** of $x \in M$ if $xyx = x$ and $yxy = y$. It is easy to check that if x is regular and $xyx = x$ for a certain $y \in M$, then yxy is an inverse of x . Therefore, every regular element of M has at least one inverse in M . This property of regularity plays well with Green's $\mathcal{D}$ relation, as the next lemma shows.

Lemma 2.16. *Let M be a monoid. If $x \in M$ is a regular element, then all elements in its $\mathcal{D}$ -class are.*

Proof. Let $y \in D_x$ be such that $x\mathcal{R}z\mathcal{L}y$ for some $z \in M$. Then there exist $a, b \in M$ for which $x = za$ and $z = xb$. Since x is regular, it follows that $xwx = x$ for some $w \in M$. Therefore,

$$z = xb = xwxb = zawxb = z(aw)z,$$

and z is a regular element. Similarly, the regularity of z and the fact that $z\mathcal{L}y$ allow us to conclude that y is regular. The lemma is thus proved. $\square$

Observe that in a regular $\mathcal{D}$ -class every $\mathcal{R}$ -class contains an idempotent, and the same applies to every $\mathcal{L}$ -class. Indeed, if D is a $\mathcal{D}$ -class and $x = xyx \in D$, then $x\mathcal{R}xy$ and $(xy)^2 = xy$; also, $x\mathcal{L}yx$ and $(yx)^2 = yx$.

In Corollary [2.11](#), we saw that $|H_x| = |H_y|$ whenever $x\mathcal{D}y$. In case x and y are idempotents, their $\mathcal{D}$ -class is regular, and we obtain a stronger result. A proof can be found in [\[9, Proposition 2.3.6\]](#).

Proposition 2.17. *Let M be a monoid. If $e, f \in E(M)$ are such that $e\mathcal{D}f$, then $H_e \cong H_f$.*

2.6 Semilattices and partial orders

In this section, we get a glimpse on how algebraic and order structures sometimes interact nicely between themselves.

Let $\leq$ be a **partial order** (reflexive, antisymmetric and transitive relation) on a set X , making $(X, \leq)$ a **partially ordered set** (or **poset**). We say that $(X, \leq)$ is a $\wedge$ -**semilattice** if, for every $x, y \in X$, there exists an element $z \in X$ satisfying

$$\begin{cases} z \leq x \\ z \leq y \\ \forall w \in X (w \leq x, w \leq y \Rightarrow w \leq z). \end{cases}$$

It is easy to see that such an element z must be unique, so we call z the **infimum** of x and y and denote it by $x \wedge y$.

For all $x, y, z \in X$, it holds that $x \wedge (y \wedge z) = (x \wedge y) \wedge z$, so we can consider $\wedge$ to be an associative binary operation on X , thereby introducing a semigroup structure on X . Moreover, since $x \wedge x = x$ and $x \wedge y = y \wedge x$ for all $x, y \in X$, all elements in X are idempotents and commute. A semigroup whose elements are idempotents is called a **band**, so any $\wedge$ -semilattice can be regarded as a commutative band from an algebraic point of view.

Thankfully, the converse is also true. Given a commutative band B , we define a relation $\leq$ on B by

$$x \leq y \Leftrightarrow x = xy.$$

Since the equality $x^2 = x$ always holds, it follows that $x \leq x$; if $x \leq y$ and $y \leq x$, then $x = xy = yx = y$; and if $x \leq y \leq z$, then $x = xy$ and $y = yz$ imply $x = xy = xyz = xz$, so $x \leq z$. On account of this, $(B, \leq)$ is a poset. Furthermore, given $x, y \in B$, we have $x \wedge y = xy$. Indeed, $xy \leq x$ since $xy = x(xy) = (xy)x$, and $xy \leq y$ can be deduced in a similar way. And $z \leq x, y$ implies $z = zy = zxy$, whence $z \leq xy$. Thus, we conclude that $(B, \leq)$ is a $\wedge$ -semilattice. From now on, the term “semilattice” will be used to refer to both algebraic and order structures.

We end this section by introducing a rather special order on the set of idempotents of a monoid. Given a monoid M , we define a relation $\leq$ on $E(M)$ by

$$e \leq f \Leftrightarrow e = ef = fe,$$

where $e, f \in E(M)$. This relation is reflexive because $e \leq e \Leftrightarrow e^2 = e$. Moreover, if $e \leq f$ and $f \leq e$, then $e = ef = fe = f$, so $\leq$ is antisymmetric. Finally, if $e \leq f \leq g$, then $eg = efg = ef = e$ and $ge = gfe = fe = e$, thereby implying $e \leq g$. We conclude that $\leq$ is a partial order, and we call it the **natural partial order** on $E(M)$.

3 Inverse monoids

In this section, we introduce a very special class of monoids called inverse monoids. Due to their similarities with groups, for many years they were known as “generalized groups” in Russian literature. Inverse monoids not only play an important role in semigroup theory but also show up in many other areas of mathematics, since they abstract the concept of partial symmetry in the same way that groups abstract the notion of symmetry. For more details, we refer the reader to [18].

3.1 Basic properties

We begin by defining what inverse monoids are. Then we present some equivalent characterizations and basic results concerning them.

Let M be a monoid. Recall that, given $x, y \in M$, we say that y is an inverse of x if $xyx = x$ and $yxy = y$; in that case, it is clear that x is an inverse of y too. Note that in an arbitrary monoid an element may have many inverses. We call M an **inverse monoid** if every element in M has exactly one inverse. For instance, every group is an inverse monoid. We denote by x^{-1} the unique inverse of x , which satisfies $x = xx^{-1}x$ and $x^{-1} = x^{-1}xx^{-1}$. However, it is often useful to consider alternative characterizations of inverse monoids. Below we state two of the most important ones.

Proposition 3.1. *Let M be a monoid. The following conditions are equivalent:*

- (i) M is an inverse monoid.
- (ii) M is regular and its idempotents commute.
- (iii) Every $\mathcal{R}$ -class and every $\mathcal{L}$ -class of M contains exactly one idempotent.

Proof. (i) $\Rightarrow$ (ii): Let M be an inverse monoid. Since every element in M has an inverse, in particular every element is regular, so M is regular. If $e, f \in E(M)$ and $x = (ef)^{-1}$, we infer that $efxef = ef$ and $x = xefx$; thus, $ef(fxe)ef = ef$ and $(fxe)ef(fxe) = fxe$. This shows that fxe is the inverse of ef , whence $x = fxe$. Therefore,

$$x^2 = fxe fxe = fxe = x,$$

and $x \in E(M)$ is its own inverse. But x is the inverse of ef , so $ef = x \in E(M)$, and we deduce that $fe \in E(M)$ in a similar manner. It follows that

$$\begin{aligned} ef(fe)ef &= efef = ef \\ fe(ef)fe &= fefe = fe, \end{aligned}$$

which implies that ef is the inverse of fe . Since inverses in M are unique and every idempotent is an inverse of itself, we conclude that $ef = fe$, so (ii) holds.

(ii) $\Rightarrow$ (iii): In section 2.5 we observed that in a regular monoid every $\mathcal{R}$ -class and every $\mathcal{L}$ -class have at least one idempotent. Suppose then that M is a regular monoid whose

idempotents commute and let $e, f \in E(M)$ be such that $e\mathcal{R}f$. There exists $x \in M$ for which $e = fx$, and thus $fe = f(fx) = fx = e$; out of symmetry, we get $ef = f$ too. As $ef = fe$ by assumption, we get $e = f$. One proves that $e\mathcal{L}f \Rightarrow e = f$ in a similar way.

(iii) $\Rightarrow$ (i): If every $\mathcal{R}$ -class contains an idempotent, so does every $\mathcal{D}$ -class; hence, all elements of the $\mathcal{D}$ -classes of M are regular, and thus M is regular. Now let $x \in M$ be such that $y, z \in M$ are inverses of x . Then $xyx = x = xzx$, whence $xy\mathcal{R}x\mathcal{R}xz$. But both xy and xz are idempotents, and therefore $xy = xz$. Moreover, since $yx\mathcal{L}x\mathcal{L}zx$, we deduce that $yx = zx$. On account of this,

$$y = yxy = zxy = zxx = z$$

and we conclude that inverses in M are unique. $\square$

When proving (i) $\Rightarrow$ (ii) in the above proposition, we showed that, in an inverse monoid M , the product of two idempotents is an idempotent, and hence $E(M)$ is a semigroup. Moreover, $E(M)$ is regular, commutative and contains $1 \in M$, so $E(M)$ is an inverse monoid. More precisely, $E(M)$ is a semilattice, and we call it the **semilattice of idempotents** of M . It is important to keep in mind that $xx^{-1}, x^{-1}x \in E(M)$ for all $x \in M$. Also, $xex^{-1} \in E(M)$ for all $e \in E(M)$ because

$$xex^{-1}xex^{-1} = xe(x^{-1}x)ex^{-1} = xe^2x^{-1}xx^{-1} = xex^{-1}.$$

As in the case of groups, if $x_1, x_2, \dots, x_n \in M$ are elements of an inverse monoid M , the equality

$$(x_1x_2 \cdots x_n)^{-1} = x_n^{-1} \cdots x_2^{-1}x_1^{-1}$$

holds. Indeed, it suffices to establish the result for $n = 2$, for then a simple induction argument establishes the general case. Taking $x, y \in M$, we have

$$(xy)y^{-1}x^{-1}(xy) = x(yy^{-1})(x^{-1}x)y = x(x^{-1}x)(yy^{-1})y = xy,$$

and we deduce that $y^{-1}x^{-1}(xy)y^{-1}x^{-1} = y^{-1}x^{-1}$ in an identical manner. Therefore, $(xy)^{-1} = y^{-1}x^{-1}$, and the general case now follows by induction.

Since inverse monoids are so close to groups, one may ask how to detect groups among inverse monoids. On the one hand, if M is a group, then M is an inverse monoid with a unique idempotent. Conversely, let M be an inverse monoid whose unique idempotent is 1. Since $xx^{-1}, x^{-1}x \in E(M)$ for all $x \in M$, it follows that $xx^{-1} = 1 = x^{-1}x$. Therefore, M is a group with identity 1, and we have just proved the subsequent proposition.

Proposition 3.2. *Let M be an inverse monoid. Then M is a group if and only if it has a unique idempotent.*

Before stating an important property of inverse monoids regarding homomorphisms, we need a preliminary lemma, named after Gérard Lallement [14].

Lemma 3.3 (Lallement’s Lemma). *Let σ be a congruence on a regular monoid M and let $x \in M$. If $x\sigma \in E(M/\sigma)$, then $x\sigma e$ for some $e \in E(M)$.*

Proof. Assuming that $x\sigma \in E(M/\sigma)$, let $y \in M$ be an inverse of x^2 and $e = xyx$. Then we deduce that

$$e^2 = xyx^2yx = xyx = e \in E(M).$$

From $e = xyx$ and $x\sigma = x^2\sigma$, we get $e\sigma = (xyx)\sigma = (x^2yx^2)\sigma$. Moreover, $x^2yx^2 = x^2$ by definition of y . Hence, we conclude that $e\sigma x^2\sigma x$, as we wanted. $\square$

In the language of homomorphisms, Lallement’s Lemma states that if M is a regular monoid and $\varphi: M \rightarrow N$ is a surjective homomorphism, then for every $e' \in E(N)$ there exists $e \in E(M)$ such that $e\varphi = e'$. We are ready to prove the following proposition.

Proposition 3.4. *Let M be an inverse monoid and N be any monoid. If $\varphi: M \rightarrow N$ is a surjective homomorphism, then N is an inverse monoid and $x^{-1}\varphi = (x\varphi)^{-1}$ for all $x \in M$.*

Proof. According to Proposition 3.1(ii), we are going to show that N is regular and its idempotents commute. Given $y \in N$, we have $y = x\varphi$ for some $x \in M$. Since

$$y(x^{-1}\varphi)y = (x\varphi)(x^{-1}\varphi)(x\varphi) = x\varphi = y,$$

it follows that N is regular. Given $e', f' \in E(N)$, it is a consequence of Lemma 3.3 that there exist $e, f \in E(M)$ satisfying $e' = e\varphi$ and $f' = f\varphi$. As e and f commute due to M being inverse, we conclude that

$$e'f' = (e\varphi)(f\varphi) = (ef)\varphi = (fe)\varphi = (f\varphi)(e\varphi) = f'e',$$

so N is inverse. To finish, we saw above that $(x\varphi)(x^{-1}\varphi)(x\varphi) = x\varphi$, and we also have $(x^{-1}\varphi)(x\varphi)(x^{-1}\varphi) = x^{-1}\varphi$. It follows that $x^{-1}\varphi = (x\varphi)^{-1}$. $\square$

To end this section, we call the reader’s attention to the fact that a submonoid of an inverse monoid need not be inverse. Indeed, $\mathbb{Z}_0^+$ is a submonoid of $\mathbb{Z}$ (for the operation of addition) which is not inverse. Thus, even though the condition of “being inverse” is preserved under homomorphisms, it does not behave so well with respect to submonoids. For that reason, we say that a nonempty set $N \subseteq M$ is an **inverse submonoid** of the inverse monoid M if $1 \in N$ and $xy, x^{-1} \in N$ whenever $x, y \in N$. Equivalently, a submonoid N of an inverse monoid M is an inverse submonoid if it is an inverse monoid on its own. In the context of inverse monoids, when we write $N \leq M$, we implicitly assume that N is an inverse submonoid of M .

3.2 Natural partial order

Recall that, given a monoid M , we defined the natural partial order on $E(M)$ by $e \leq f$ if $e = ef = fe$. We will now see that this order can be somewhat extended to all of M in case it is an inverse monoid.

Let M be an inverse monoid. We define a relation $\leq$ on M by

$$x \leq y \Leftrightarrow x = ey \text{ for some } e \in E(M).$$

Let us check that $\leq$ is a partial order. Reflexivity follows from the fact that $x = xx^{-1}x$ and $xx^{-1} \in E(M)$ for all $x \in X$. Regarding antisymmetry, suppose that $x, y \in M$ are such that $x \leq y$ and $y \leq x$. Then there exist $e, f \in E(M)$ satisfying $x = ey$ and $y = fx$. Consequently,

$$x = ey = efx = fex = fey = fx = y,$$

as e and f commute and $ex = e^2y = ey$. Lastly, if $x \leq y$ and $y \leq z$, then $x = ey$ and $y = fz$ for some $e, f \in E(M)$, whence $x = ey = e fz$. Since $E(M)$ is a monoid, we know that $ef \in E(M)$, which shows that $\leq$ is transitive, and therefore a partial order. The restriction of $\leq$ to $E(M)$ is precisely the natural partial order we had previously defined for idempotents.

Observe that this order is compatible with right and left multiplication, in the sense that $x \leq y$ implies $xz \leq yz$ and $zx \leq zy$ for all $z \in M$. Indeed, suppose that $x = ey \leq y$ for some $e \in E(M)$. Then $xz = e(yz) \leq yz$ and

$$zx = zey = zz^{-1}zey = (zez^{-1})zy \leq zy$$

because $zez^{-1} \in E(M)$. Furthermore, the natural partial order also respects inverses. Indeed, $x \leq y$ entails $x = ey$ for some $e \in E(M)$, and thus

$$x^{-1} = (ey)^{-1} = y^{-1}e = y^{-1}yy^{-1}e = (y^{-1}ey)y^{-1} \leq y^{-1}.$$

Note also that we might as well have defined $x \leq y \Leftrightarrow x = xx^{-1}y$ because $xx^{-1} \in E(M)$ and, if $x = ey$ for some $e \in E(M)$, then

$$xx^{-1}y = ey(ey)^{-1}y = eyy^{-1}ey = eyy^{-1}y = ey = x.$$

But one may ask why the idempotent in the definition of the natural partial order is placed on the left. What would happen if it appeared multiplying x on the right? It turns out that there would be no difference. Indeed, suppose that $x \leq y$ and $x = ey$ for some $e \in E(M)$. Then $x = eyy^{-1}y = y(y^{-1}ey)$, and $y^{-1}ey \in E(M)$. Conversely, if $x = ye$ for some $e \in E(M)$, then $x = yy^{-1}ye = (yey^{-1})y$, and $x \leq y$.

We collect the previous remarks in the next proposition.

Proposition 3.5. *Let M be an inverse monoid and $x, y \in M$. The following conditions are equivalent:*

- (i) $x = ey$ for some $e \in E(M)$.
- (ii) $x = xx^{-1}y$.
- (iii) $x = yf$ for some $f \in E(M)$.

3.3 Green's relations

We have already discussed Green's equivalence relations on an arbitrary monoid M . In case M is inverse, they enjoy some specific properties.

Proposition 3.6. *Let M be an inverse monoid and let $x, y \in M$, $e, f \in E(M)$. Then:*

- (i) $x\mathcal{R}y$ if and only if $xx^{-1} = yy^{-1}$.
- (ii) $x\mathcal{L}y$ if and only if $x^{-1}x = y^{-1}y$.
- (iii) $x\mathcal{R}y$ if and only if $x^{-1}\mathcal{L}y^{-1}$.
- (iv) $x\mathcal{H}y$ if and only if $xx^{-1} = yy^{-1}$ and $x^{-1}x = y^{-1}y$.
- (v) $e\mathcal{D}f$ if and only if there exists some $x \in M$ satisfying $e = xx^{-1}$ and $f = x^{-1}x$.
- (vi) Every $\mathcal{D}$ -class has the same number of $\mathcal{R}$ -classes and $\mathcal{L}$ -classes.

Proof. Items (i) and (ii) follow from Proposition 3.1(iii), together with the fact that $x\mathcal{R}xx^{-1}$ and $x\mathcal{L}x^{-1}x$ for all $x \in M$. Items (iii) and (iv) follow from (i) and (ii). Regarding (v), we have $e\mathcal{D}f$ if and only if there exists $x \in M$ satisfying $e\mathcal{R}x\mathcal{L}f$; by (i) and (ii), this is equivalent to $e = xx^{-1}$ and $x^{-1}x = f$. To prove (vi), let D be a $\mathcal{D}$ -class and denote by X and Y the sets of its $\mathcal{R}$ -classes and $\mathcal{L}$ -classes, respectively. By (iii), we have a well-defined function $\alpha: X \rightarrow Y$ that maps an $\mathcal{R}$ -class R to the $\mathcal{L}$ -class R^{-1} , whose inverse is $\beta: Y \rightarrow X$, $L \mapsto L^{-1}$. Therefore, α is a bijection, and $|X| = |Y|$. $\square$

Properties (i), (ii) and (iv) of the previous proposition have an interesting consequence. Let M be an inverse monoid and $N \leq M$ an inverse submonoid. If $\mathcal{K}$ is one of Green's relations $\mathcal{R}$, $\mathcal{L}$ or $\mathcal{H}$, then $x\mathcal{K}y$ holds in M if and only if $x\mathcal{K}y$ holds in N . In general, if M is an arbitrary monoid and N is a submonoid of M , we can only guarantee that $x\mathcal{K}y$ in N implies $x\mathcal{K}y$ in M . Moreover, property (vi) states that every “egg-box” representing a $\mathcal{D}$ -class (as in Figure 1) is square.

Note that Green's relations are of no interest in a group, the reason being that the equalities $\mathcal{R} = \mathcal{L} = \mathcal{H} = \mathcal{D} = \mathcal{J}$ always hold in that case. This means that, in a group, every “egg-box” has only room for a single egg.

3.4 Symmetric inverse monoids

We have seen before that every monoid M is isomorphic to a submonoid of $\mathcal{T}_X$ for some nonempty set X . If M is inverse, we can establish something more precise.

Recall that, given a nonempty set X , we designate by $\mathcal{PT}_X$ the monoid of all partial transformations on X ; such transformations need not be defined on all of X . If $\varphi \in \mathcal{PT}_X$ is such that

$$\forall x, y \in \text{dom } \varphi \ (x\varphi = y\varphi \Rightarrow x = y),$$

we say that φ is an **injective partial transformation**. We denote by $\mathcal{I}_X \subseteq \mathcal{PT}_X$ the set of all injective partial transformations on X , which is in fact a submonoid of $\mathcal{PT}_X$.

It is natural to ask what the idempotents of $\mathcal{I}_X$ look like. Let us show that $\alpha \in E(\mathcal{I}_X)$ if and only if $\alpha = \text{id}_A$ for some $A \subseteq X$. Given $A \subseteq X$, it is clear that $\text{id}_A \in E(\mathcal{I}_X)$. Assume

now that $\alpha \in E(\mathcal{I}_X)$. Given $x \in \text{dom } \alpha$, since $x\alpha^2 = x\alpha$ and α is injective, it follows that $x\alpha = x$. Writing $A = \text{dom } \alpha$, we conclude that $\alpha = \text{id}_A$, as we wanted.

Now that we know what the idempotents of $\mathcal{I}_X$ are, we can prove the following proposition.

Proposition 3.7. *Let X be a nonempty set. Then $\mathcal{I}_X$ is an inverse monoid.*

Proof. Given $\alpha \in \mathcal{I}_X$, we may regard α as a bijection between $\text{dom } \alpha$ and $\text{im } \alpha$. Indeed, let us write $A = \text{dom } \alpha$, $B = \text{im } \alpha$, and consider $\alpha: A \rightarrow B$. Then α admits an inverse $\alpha^{-1}: B \rightarrow A$, which satisfies $\alpha\alpha^{-1}\alpha = \alpha$, so $\mathcal{I}_X$ is regular. Now, as we have just remarked, given two idempotents in $\mathcal{I}_X$, they are of the form id_A and id_B for some $A, B \subseteq X$; therefore, they commute since

$$\text{id}_A \text{id}_B = \text{id}_{A \cap B} = \text{id}_B \text{id}_A.$$

According to Proposition 3.1(ii), we conclude that $\mathcal{I}_X$ is an inverse monoid. $\square$

We are ready to state the “inverse monoid version” of Cayley’s theorem for groups and Theorem 2.3 for arbitrary monoids. It is known as the **Wagner-Preston representation** [27, 20].

Theorem 3.8 (Wagner-Preston representation). *Let M be an inverse monoid. Then there exists a nonempty set X such that M is isomorphic to an inverse submonoid of $\mathcal{I}_X$.*

Proof. We choose $X = M$, seen as a set. Consider the function

$$\begin{aligned} \rho: M &\longrightarrow \mathcal{I}_M \\ x &\longmapsto \rho_x \end{aligned}$$

where $\rho_x \in \mathcal{PT}_M$ has domain Mx^{-1} and corresponds to right multiplication by x , that is, $y\rho_x = yx$ for all $y \in Mx^{-1}$. Let us verify that $\rho_x \in \mathcal{I}_M$, in order to prove that ρ is well defined. Suppose that, for $y, z \in M$, we have $y\rho_x = z\rho_x$. Since $y, z \in Mx^{-1}$, there exists $w \in M$ such that $y = wx^{-1}$, which implies $yxx^{-1} = wx^{-1} = y$; similarly, $zxx^{-1} = z$. Hence,

$$y = yxx^{-1} = (y\rho_x)x^{-1} = (z\rho_x)x^{-1} = zxx^{-1} = z,$$

so ρ_x is an injective partial transformation for every $x \in M$. Therefore, ρ is a well defined mapping.

Next we prove that ρ is a homomorphism. It is clear that $1\rho = \text{id}_M$. Taking $x, y \in M$, we have

$$z\rho_{xy} = z(xy) = (zx)y = z\rho_x\rho_y,$$

whenever $z \in \text{dom } \rho_{xy} \cap \text{dom } \rho_x\rho_y$; thus, it suffices to check that $\text{dom } \rho_{xy} = \text{dom } \rho_x\rho_y$ for all $x, y \in M$. On the one hand, $\text{dom } \rho_{xy} = My^{-1}x^{-1}$; on the other hand,

$$\text{dom } \rho_x\rho_y = (\text{im } \rho_x \cap \text{dom } \rho_y)\rho_x^{-1} = (Mx^{-1}x \cap My^{-1})\rho_x^{-1}.$$

Taking $z \in My^{-1}x^{-1}$, say $z = wy^{-1}x^{-1}$, we have

$$z\rho_x = wy^{-1}x^{-1}x = wy^{-1}yy^{-1}x^{-1}x = wy^{-1}x^{-1}xyy^{-1} \in Mx^{-1}x \cap My^{-1},$$

and hence $\text{dom } \rho_{xy} \subseteq \text{dom } \rho_x\rho_y$. Conversely, if $z \in (Mx^{-1}x \cap My^{-1})\rho_x^{-1}$, then $z\rho_x = wx^{-1}\rho_x$ for some $w \in M$, whence $z = wx^{-1} = zxx^{-1}$ since ρ_x is injective. Also, $zx = ty^{-1}$ for some $t \in M$, so

$$z = zxx^{-1} = ty^{-1}x^{-1} \in My^{-1}x^{-1},$$

which shows that $\text{dom } \rho_x\rho_y \subseteq \text{dom } \rho_{xy}$. The equality of domains now follows.

Finally, it remains to prove that ρ is injective, as the homomorphic image of an inverse monoid is an inverse monoid. That is due to the fact that if $x, y \in M$ satisfy $\rho_x = \rho_y$, then $x = 1\rho_x = 1\rho_y = y$. We conclude that ρ is an injective homomorphism, which establishes the theorem. $\square$

We end this section by collecting some properties regarding the natural partial order and Green's relations in a symmetric inverse monoid.

Proposition 3.9. *Let X be a nonempty set and $\alpha, \beta \in \mathcal{I}_X$. Then:*

- (i) $\alpha \leq \beta$ if and only if $\beta|_{\text{dom } \alpha} = \alpha$.
- (ii) $\alpha \mathcal{R} \beta$ if and only if $\text{dom } \alpha = \text{dom } \beta$.
- (iii) $\alpha \mathcal{L} \beta$ if and only if $\text{im } \alpha = \text{im } \beta$.
- (iv) $\alpha \mathcal{H} \beta$ if and only if $\text{dom } \alpha = \text{dom } \beta$ and $\text{im } \alpha = \text{im } \beta$.
- (v) $\alpha \mathcal{D} \beta$ if and only if $\alpha \mathcal{J} \beta$ if and only if $|\text{dom } \alpha| = |\text{dom } \beta|$.

Proof. (i) If $\alpha \leq \beta$, then there exists $\gamma \in E(\mathcal{I}_X)$ such that $\alpha = \gamma\beta$. As we have previously observed, this implies that $\alpha = \text{id}_A\beta$ for some $A \subseteq X$, whence

$$\text{dom } \alpha = (A \cap \text{dom } \beta)\text{id}_A^{-1} = A \cap \text{dom } \beta \subseteq \text{dom } \beta.$$

Furthermore, given $x \in \text{dom } \alpha \subseteq A$, we have $x\beta = x\text{id}_A\beta = x\alpha$, so $\beta|_{\text{dom } \alpha} = \alpha$ as claimed. On the other hand, if $\alpha = \beta|_{\text{dom } \alpha}$, then $\alpha = \text{id}_{\text{dom } \alpha}\beta$ and $\text{id}_{\text{dom } \alpha} \in E(\mathcal{I}_X)$. The property is thus established.

(ii) Since $\text{dom } \gamma\gamma^{-1} = \text{dom } \gamma$ for all $\gamma \in \mathcal{I}_X$, Proposition 3.6(i) yields

$$\alpha \mathcal{R} \beta \Leftrightarrow \alpha\alpha^{-1} = \beta\beta^{-1} \Leftrightarrow \text{id}_{\text{dom } \alpha} = \text{id}_{\text{dom } \beta} \Leftrightarrow \text{dom } \alpha = \text{dom } \beta.$$

(iii) Since $\text{dom } \gamma^{-1} = \text{im } \gamma$ for all $\gamma \in \mathcal{I}_X$, Proposition 3.6(ii) yields

$$\alpha \mathcal{L} \beta \Leftrightarrow \alpha^{-1}\alpha = \beta^{-1}\beta \Leftrightarrow \text{id}_{\text{im } \alpha} = \text{id}_{\text{im } \beta} \Leftrightarrow \text{im } \alpha = \text{im } \beta.$$

(iv) As $\mathcal{H} = \mathcal{R} \cap \mathcal{L}$, the conclusion follows from the two previous items.

(v) First we show that $\alpha\mathcal{J}\beta \Rightarrow |\text{dom } \alpha| = |\text{dom } \beta|$. If $\alpha\mathcal{J}\beta$, then $\alpha = \gamma\beta\delta$ for some $\gamma, \delta \in \mathcal{I}_X$, and thus $|\text{dom } \alpha| \leq |\text{dom } \beta|$. Out of symmetry we get $|\text{dom } \beta| \leq |\text{dom } \alpha|$, so $|\text{dom } \alpha| = |\text{dom } \beta|$. Now, if $\alpha, \beta \in \mathcal{I}_X$ are such that $|\text{dom } \alpha| = |\text{dom } \beta|$, then there exists a bijection $\gamma: \text{dom } \alpha \rightarrow \text{im } \beta$. Hence, $\text{dom } \gamma = \text{dom } \alpha$ and $\text{im } \gamma = \text{im } \beta$, which implies $\alpha\mathcal{R}\gamma$ and $\gamma\mathcal{L}\alpha$, by items (ii) and (iii). Therefore $\alpha\mathcal{D}\beta$, and we already know that $\mathcal{D} \subseteq \mathcal{J}$ in any monoid, so we have proven that

$$\alpha\mathcal{J}\beta \Rightarrow |\text{dom } \alpha| = |\text{dom } \beta| \Rightarrow \alpha\mathcal{D}\beta \Rightarrow \alpha\mathcal{J}\beta,$$

which yields the desired equivalences. □

3.5 Free inverse monoids

In Section [2.3](#), we constructed the so-called free monoid over a set. We now define a similar object in the context of inverse monoids.

Given an alphabet A , we start by creating a set $A^{-1} = \{a^{-1} \mid a \in A\}$ which consists of formal inverses of the letters in A , and we write $\tilde{A} = A \cup A^{-1}$. We declare $(a^{-1})^{-1} = a$ for all $a \in A$, so that the bijection $A \rightarrow A^{-1}$, $a \mapsto a^{-1}$ becomes an involution. Then we extend this involution to an involution of $\tilde{A}^*$ by declaring $1^{-1} = 1$ and $(a_1 \dots a_n)^{-1} = a_n^{-1} \dots a_1^{-1}$ for all $n \geq 1$ and $a_i \in \tilde{A}$.

The **Wagner congruence** on $\tilde{A}^*$ is denoted by ρ_A and is generated by the relation

$$\{(uu^{-1}u, u) \mid u \in \tilde{A}^*\} \cup \{(uu^{-1}vv^{-1}, vv^{-1}uu^{-1}) \mid u, v \in \tilde{A}^*\}.$$

In order to avoid too many parenthesis, by xyp_A we mean $(xy)\rho_A$. We define FIM_A to be the quotient $FIM_A = \tilde{A}^*/\rho_A$, and we begin by observing that it is an inverse monoid. Indeed, given $u \in \tilde{A}^*$, we have

$$(u\rho_A)(u^{-1}\rho_A)(u\rho_A) = uu^{-1}u\rho_A = u\rho_A,$$

so FIM_A is a regular monoid. Moreover, if $e\rho_A \in E(FIM_A)$, then

$$\begin{aligned} e\rho_A &= ee^{-1}e\rho_A = e(e^{-1}ee^{-1})e\rho_A \\ &= e(e^{-1}e)(ee^{-1})e\rho_A \\ &= e(ee^{-1})(e^{-1}e)e\rho_A \\ &= (ee^{-1})(e^{-1}e)\rho_A. \end{aligned}$$

Consequently, given $f\rho_A \in E(FIM_A)$, we deduce that

$$\begin{aligned} (e\rho_A)(f\rho_A) &= (ee^{-1})(e^{-1}e)\rho_A (ff^{-1})(f^{-1}f)\rho_A \\ &= (ff^{-1})(f^{-1}f)\rho_A (ee^{-1})(e^{-1}e)\rho_A \\ &= (f\rho_A)(e\rho_A), \end{aligned}$$

which proves that idempotents commute in FIM_A . In view of Proposition [3.1](#), we conclude that FIM_A is an inverse monoid.

The next theorem shows that FIM_A deserves to be called the **free inverse monoid** over A .

Theorem 3.10 (Universal property of FIM_A). *Let A be an alphabet and consider the morphism*

$$\begin{aligned}\theta: A &\longrightarrow FIM_A \\ a &\longmapsto a\rho_A.\end{aligned}$$

If $\varphi: A \rightarrow M$ is a function from A to an inverse monoid M , then there exists a unique monoid homomorphism $\Phi: FIM_A \rightarrow M$ such that $\theta\Phi = \varphi$.

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & M \\ \theta \downarrow & \nearrow \Phi & \\ FIM_A & & \end{array}$$

Proof. First, if M is an inverse monoid, then every function $\varphi: A \rightarrow M$ can be extended to a function $\tilde{\varphi}: \tilde{A} \rightarrow M$ by defining $(a^{-1})\tilde{\varphi} = (a\tilde{\varphi})^{-1}$ for every $a \in A$. We can then (uniquely) extend $\tilde{\varphi}$ to a monoid homomorphism $\phi: \tilde{A}^* \rightarrow M$, in view of the universal property of $\tilde{A}^*$. Now M being inverse implies $\rho_A \subseteq \ker \phi$, so ϕ induces a homomorphism $\Phi: FIM_A \rightarrow M$, according to Theorem [2.2](#). This homomorphism satisfies $\theta\Phi = \varphi$, and it is unique since $A\theta$ generates FIM_A as an inverse monoid. $\square$

As in Corollary [2.5](#), it follows that every inverse monoid satisfying the universal property with respect to a set A is isomorphic to FIM_A . It also holds that every inverse monoid is isomorphic to a quotient of a free inverse monoid.

4 Automata and rational languages

Automata theory plays a major role in computer science, being the study of an important model of computation. Nowadays, such abstract machines are fundamental to the design of many kinds of software and hardware. In mathematics, they are a most powerful tool to study certain algebraic structures, as we are about to see. On the other hand, the initial goal of formal language theory was to grasp syntactic regularities in the way humans communicate. With the emergence of computers, they became relevant to the development of programming languages. In this monograph, we will only be concerned with a specific class of formal languages — the rational languages — which are intimately related to automata. For more details, we refer the reader to [19].

4.1 Rational languages

In section 2.3, we constructed the so-called free monoid over an alphabet. Indeed, given an alphabet A , the free monoid over A is denoted by A^* and, as a set, it consists of all words which can be written using letters from A . But what if we want to consider only *some* words? That leads to the notion of a language.

Let A be an alphabet. A **language** over A (or **A -language**) is a subset $L \subseteq A^*$. We can combine A -languages using three **rational operators** known as **union**, **product** and **star**. The first ones are binary, whereas the latter is unary. More precisely, the union and product operators take as input a pair (K, L) of A -languages and output $K \cup L$ and $KL = \{uv \mid u \in K, v \in L\}$, respectively. Moreover, the star operator applied to L yields the A -language

$$L^* = \{u_1 \dots u_n \mid n \geq 0; u_1, \dots, u_n \in L\} = \bigcup_{n \geq 0} L^n,$$

under the convention $L^0 = \{1\}$. This notation is slightly misleading since L^* is not a free monoid in general. However, if $L \subseteq A$, then L^* is indeed the free monoid over L .

Example 4.1. Let $A = \{a, b, c\}$, $K = \{ab^2, bc\}$ and $L = \{a\}$. Then $K \cup L = \{a, ab^2, bc\}$, $KL = \{ab^2a, bca\}$ and $L^* = \{1, a, a^2, a^3, \dots\}$.

We say that an A -language is **rational** if it can be obtained from finite A -languages using the rational operators finitely many times. A decomposition of a rational language L in terms of finite languages and rational operators is called a **rational expression** for L .

Example 4.2. Let $A = \{a, b\}$. Then $L = \{a^2, ab, ba, b^2\}^* \{a, b\}$ is a rational language built from the finite A -languages $\{a^2, ab, ba, b^2\}$ and $\{a, b\}$. It consists of all words over $\{a, b\}$ with an odd number of letters.

In order to simplify the notation, we try to avoid too many brackets and commas so, for instance, the above rational expression becomes $L = (a^2 \cup ab \cup ba \cup b^2)^*(a \cup b)$.

We end this small section by introducing some more terminology concerning words in A^* . Given $u, v \in A^*$, we call u a **prefix** of v if $v = uw$ for some $w \in A^*$; a **suffix** of v if

$v = wu$ for some $w \in A^*$; and a **factor** of v if $v = wuz$ for some $w, z \in A^*$.

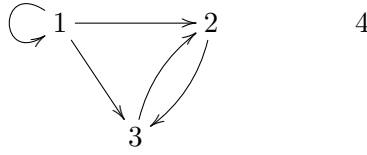
4.2 Graphs and automata

As we have already mentioned, automata will be one of the most important objects in this monograph. They are a kind of machine that has the ability to read words and recognize languages. But before we jump into the main concept, we begin by defining a more rudimentary object.

A (directed) **graph** is an ordered quadruple $\Gamma = (V, E, \alpha, \omega)$, where $V \neq \emptyset$ is called the set of **vertices**, E the set of **edges** and $\alpha, \omega: E \rightarrow V$ the **incidence functions**. Given an edge $e \in E$, we call $e\alpha$ its **tail** and $e\omega$ its **head**; we also say that e begins in $e\alpha$ and ends in $e\omega$. If $e\alpha = e\omega$, we call e a **loop**. Visually, each edge is represented by $e\alpha \xrightarrow{e} e\omega$.

It is usual to depict a graph as a set of points joined by arrows. The points illustrate the vertices and the arrows represent the edges entering or leaving a certain vertex. Given $v \in V$, the **outdegree** of v , written $d^+(v)$, is the number of edges that begin in v ; the **indegree** of v , written $d^-(v)$, is the number of edges that end in v ; and the **total degree** of v is $d(v) = d^+(v) + d^-(v)$.

Example 4.3. Below we present a directed graph with four vertices and five edges. We have $d^+(1) = 3$, $d^-(1) = 1$, $d^+(2) = d^+(3) = 1$, $d^-(2) = d^-(3) = 2$ and $d^+(4) = d^-(4) = 0$.



Let $\Gamma = (V, E, \alpha, \omega)$ and $\Gamma' = (V', E', \alpha', \omega')$ be two graphs. An **isomorphism** between Γ and Γ' is a pair (φ, ψ) , where $\varphi: V \rightarrow V'$ and $\psi: E \rightarrow E'$ are bijective functions that satisfy $\psi\alpha' = \alpha\varphi$ and $\psi\omega' = \omega\varphi$. Informally, two graphs are isomorphic if the relations between the vertices and the arrows are the same in each of them. An **automorphism** of a graph is an isomorphism from the graph to itself. When we are only concerned with the structure of graphs and do not care about the names of their vertices and edges, we consider isomorphic graphs to be the same.

We proceed with some more definitions. A **subgraph** of $\Gamma = (V, E, \alpha, \omega)$ is a graph $\Gamma' = (V', E', \alpha', \omega')$ satisfying $V' \subseteq V$, $E' \subseteq E$, $\alpha' = \alpha|_{E'}$ and $\omega' = \omega|_{E'}$. We say that Γ' is the subgraph **induced** by V' if

$$E' = \{e \in E \mid e\alpha, e\omega \in V'\};$$

in other words, it is the largest subgraph of Γ with vertex set $V' \subseteq V$. A **path** on Γ is a sequence of the type

$$v_0, e_1, v_1, e_2, \dots, v_{n-1}, e_n, v_n$$

where $v_i \in V$ ($0 \leq i \leq n$); $e_i \in E$, $e_i\alpha = v_{i-1}$ and $e_i\omega = v_i$ ($1 \leq i \leq n$). We say that this is

a path from v_0 to v_n with **length** n ; if $n = 0$, we call it the **trivial** path at v_0 . A graph is called **connected** if there exists a path from each vertex to any other vertex.

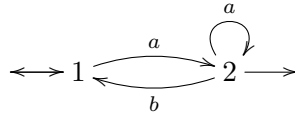
Let us define an equivalence relation $C \subseteq V \times V$ by

$$uCv \Leftrightarrow \text{there are paths from } u \text{ to } v \text{ and from } v \text{ to } u.$$

Each equivalence class C_v ($v \in V$) induces the largest connected subgraph containing v , which is called the **connected component** of v . For instance, the graph depicted in Example 4.3 has three connected components, namely the subgraphs induced by 1, $\{2, 3\}$ and 4. Observe that 1 does not belong to the connected component of 2 and 3 because there is no path from 2 or 3 to 1, as the only edge with head 1 is the loop.

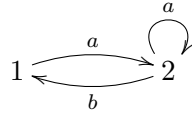
We are ready to present the main concept of this section. Informally speaking, an automaton is nothing more than a directed graph whose edges are labeled with letters from an alphabet. It may also have initial or terminal vertices, which are identified by an incoming arrow and an outgoing arrow, respectively.

Example 4.4. Below we present an automaton over the alphabet $\{a, b\}$, with initial vertex 1 and terminal vertices 1 and 2.



Formally, an **automaton** is a quintuple $\mathcal{A} = (Q, A, E, I, T)$ where A is an alphabet, Q a nonempty set, $E \subseteq Q \times A \times Q$ and $I, T \subseteq Q$. We call Q the set of **vertices** or **states**; E the set of **edges** or **transitions**; I and T the sets of **initial** and **terminal** states, respectively. If both Q and E are finite, $\mathcal{A}$ is a **finite automaton**. When it is possible to read every letter of A from any vertex in Q , we say that $\mathcal{A}$ is a **complete automaton**. The **underlying graph** of $\mathcal{A}$ is the labeled directed graph obtained from $\mathcal{A}$ by ignoring the designation of vertices as initial or terminal; we denote it by $\Gamma_{\mathcal{A}}$.

Example 4.5. If $A = \{a, b\}$ and $\mathcal{A} = (Q, A, E, I, T)$ is the automaton in Example 4.4, then $Q = \{1, 2\}$, $E = \{(1, a, 2), (2, a, 2), (2, b, 1)\}$, $I = \{1\}$ and $T = \{1, 2\}$. The underlying graph $\Gamma_{\mathcal{A}}$ is depicted by



An **isomorphism** between $\mathcal{A} = (Q, A, E, I, T)$ and $\mathcal{A}' = (Q', A, E', I', T')$ is a bijection $\varphi: Q \rightarrow Q'$ such that $I\varphi = I'$, $T\varphi = T'$ and

$$(p, a, q) \in E \Leftrightarrow (p\varphi, a, q\varphi) \in E'$$

holds for all $p, q \in Q$ and $a \in A$. More generally, a **morphism** between $\mathcal{A}$ and $\mathcal{A}'$ is a mapping $\varphi: Q \rightarrow Q'$ such that $I\varphi \subseteq I'$, $T\varphi \subseteq T'$ and $(p\varphi, a, q\varphi) \in E'$ whenever $(p, a, q) \in E$.

A **path** in an automaton $\mathcal{A} = (Q, A, E, I, T)$ is a sequence of the type

$$p_0, (p_0, a_1, p_1), p_1, (p_1, a_2, p_2), p_2, \dots, p_{n-1}, (p_{n-1}, a_n, p_n), p_n$$

where $n \geq 0$, $p_i \in Q$ ($0 \leq i \leq n$) and $a_i \in A$ ($1 \leq i \leq n$); it is nothing but a generalized version of a path in a graph. Geometrically, we represent it by

$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} p_n .$$

We call $p_1, \dots, p_{n-1}$ the **intermediate** vertices of the path. If $n > 0$, the **label** of such a path is the word $u = a_1 a_2 \dots a_n \in A^*$; if $n = 0$, we get the **trivial** path at p_0 and its label is the empty word $1 \in A^*$. If there exists a path between vertices $p, q \in Q$ labeled by $u \in A^*$, we represent it by $p \xrightarrow{u} q$; and when $p = q$, we say that u labels a **loop** at p .

The subsequent definitions make use of the fact that an automaton may have initial or terminal vertices. A path

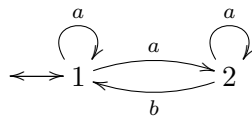
$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} p_n$$

in the automaton $\mathcal{A} = (Q, A, E, I, T)$ is called **successful** if $p_0 \in I$ and $p_n \in T$. We define the **language recognized by $\mathcal{A}$** as

$$L(\mathcal{A}) = \{u \in A^* \mid u \text{ is the label of a successful path in } \mathcal{A}\}.$$

A language $L \subseteq A^*$ is called **recognizable** if there exists a finite automaton $\mathcal{A}$ satisfying $L(\mathcal{A}) = L$. In simple terms, a recognizable language consists of all words that are “accepted” by some finite automaton. Two automata are called **equivalent** if they recognize the same language; for instance, isomorphic automata are always equivalent.

Example 4.6. The automaton in Example 4.4 is equivalent to the automaton



as they both recognize the language $(a^* \cup aa^*b)^*$.

4.3 Equivalent models

As we have seen, some automata have several initial states, while others have a unique one; some have many arrows with the same label leaving a vertex, whereas others do not even allow us to read some letters from certain states. In this section, we discuss different kinds of automata, only to realize that they are not so different after all.

We call an automaton $\mathcal{A} = (Q, A, E, I, T)$ **deterministic** if it has a unique initial state and

$$(p, a, q), (p, a, r) \in E \Rightarrow q = r$$

for all $p, q, r \in Q$ and $a \in A$. This implies that, given $u \in A^*$, there is at most one path labeled by u from the initial state to a terminal state. When $\mathcal{A}$ is deterministic, we define a partial function δ by

$$(p, a)\delta = q \Leftrightarrow (p, a, q) \in E$$

for all $p, q \in Q$ and $a \in A$. In other words, δ takes as input a state $q \in Q$ and a letter $a \in A$ and outputs the state reached after reading the letter a from vertex q , whenever that is possible. At the risk of abusing the notation, we write $\delta: Q \times A \rightarrow Q$ as in the case of transformations, despite the fact that the domain of δ may be a proper subset of $Q \times A$. We call it the **transition function** of $\mathcal{A}$.

As we shall see, these partial functions will be the key to establishing a connection between automata and monoids. Due to their importance, we often write $\mathcal{A} = (Q, A, \delta, q_0, T)$ instead of $\mathcal{A} = (Q, A, E, q_0, T)$ when dealing with deterministic automata. Note that we replace the set of initial states I by $q_0 \in Q$ to indicate that the unique initial state of $\mathcal{A}$ is q_0 . For example, the automaton $\mathcal{A} = (\{1, 2\}, \{a, b\}, \delta, 1, \{1, 2\})$ in Example 4.4 is deterministic and $(1, a)\delta = 2$, $(2, a)\delta = 2$, $(2, b)\delta = 1$. The automaton in Example 4.6 is nondeterministic, since there are two edges labeled by a leaving vertex 1.

Now, if we deviate our attention to languages again, a natural question arises: given a non-deterministic automaton, can we find an equivalent one that is deterministic? Fortunately, the answer is “yes”, and this can be achieved by considering the so-called **subset construction**, which we now explain. Given an automaton $\mathcal{A} = (Q, A, E, I, T)$, a subset $P \subseteq Q$ and a word $u \in A^*$, we define

$$Pu = \{q \in Q \mid \text{there exists a path } p \xrightarrow{u} q \text{ in } \mathcal{A} \text{ for some } p \in P\}.$$

Then, consider the automaton $\mathcal{A}' = (Q', A, E', I, T')$ given by

$$\begin{aligned} Q' &= 2^Q = \{P \mid P \subseteq Q\} \\ E' &= \{(P, a, Pa) \mid P \subseteq Q, a \in A\} \\ T' &= \{P \subseteq Q \mid P \cap T \neq \emptyset\}. \end{aligned}$$

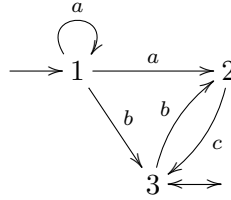
Note that $\mathcal{A}'$ is well defined because $I, Pu \in 2^Q$ for all $P \subseteq Q$ and $u \in A^*$. Moreover, it is a deterministic automaton: I is its unique initial state and, given $P \subseteq Q$ and $a \in A$, there is exactly one edge leaving P with label a , namely (P, a, Pa) . And since it is possible to read every letter of A from any state in $\mathcal{A}'$, we are facing a complete automaton. The equality $L(\mathcal{A}) = L(\mathcal{A}')$ follows from

$$\begin{aligned} u \in L(\mathcal{A}') &\Leftrightarrow \exists P \in T' \text{ such that } I \xrightarrow{u} P \text{ in } \mathcal{A}' \\ &\Leftrightarrow Iu \in T' \Leftrightarrow Iu \cap T \neq \emptyset \\ &\Leftrightarrow \exists p \in T, \exists i \in I \text{ such that } i \xrightarrow{u} p \text{ in } \mathcal{A} \\ &\Leftrightarrow u \in L(\mathcal{A}). \end{aligned}$$

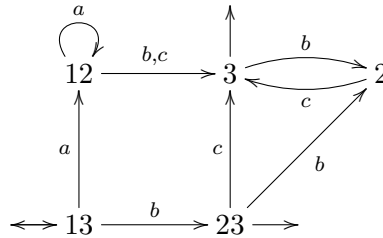
We have just proven that recognizable languages are precisely the languages recognized by finite deterministic automata. But note that the “determinization” $\mathcal{A}'$ of $\mathcal{A}$ constructed above has $2^{|Q|}$ states, whereas $\mathcal{A}$ has only $|Q|$. Fortunately, in most cases, we can simplify the subset construction and still obtain a deterministic automaton equivalent to $\mathcal{A}$, as we are about to verify.

Given an automaton $\mathcal{A} = (Q, A, E, I, T)$, a vertex $q \in Q$ is called **accessible** if it can be reached by a path departing from an initial state, and it is **co-accessible** if there is a path from q to a terminal vertex. Hence, eliminating vertices which are not accessible or vertices which are not co-accessible does not change the language recognized by the automaton. We call that process **trimming**, and when all vertices are both accessible and co-accessible, we say that the automaton is **trim**. Therefore, in the subset construction, we can always discard the vertex $\emptyset \in 2^Q$ (it is never co-accessible) and keep only the vertices of the type Iu for some $u \in A^*$.

Example 4.7. Applying the subset construction to the nondeterministic automaton



we obtain the deterministic and trim automaton



There is also a generalization of automata which is often useful, commonly named **automata with 1-transitions**, where edges can be labeled either by the empty word $1 \in A^*$ or by a letter $a \in A$. Thus, a new question arises: is an automaton with 1-transitions always equivalent to some deterministic automaton? Once again, the answer is positive. Let $\mathcal{A} = (Q, A, E, I, T)$ be an automaton with 1-transitions and define $\mathcal{A}' = (Q, A, E', I', T)$ for $I' = I1$ and

$$E' = \{(p, a, q') \in Q \times A \times Q \mid (p, a, q) \in E \text{ and } q' \in q1 \text{ for some } q \in Q\}.$$

Observe that $I \subseteq I'$ and $\mathcal{A}'$ no longer has 1-transitions, even though it may fail to be deterministic.

It is clear that $1 \in L(\mathcal{A}) \Leftrightarrow 1 \in L(\mathcal{A}')$. Now, if $n \geq 1$ and $a_1, \dots, a_n \in A$, then a path

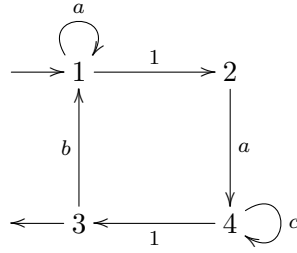
$$I' \ni p_1 \xrightarrow{a_1} p_2 \xrightarrow{a_2} \dots \xrightarrow{a_{n-1}} p_n \xrightarrow{a_n} t \in T$$

in $\mathcal{A}'$ yields a path

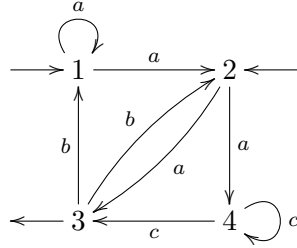
$$I \ni i \xrightarrow{1} p_1 \xrightarrow{a_1} q_1 \xrightarrow{1} \dots \xrightarrow{1} p_n \xrightarrow{a_n} q_n \xrightarrow{1} t \in T$$

in $\mathcal{A}$, and vice-versa. Therefore, $\mathcal{A}$ is equivalent to $\mathcal{A}'$. Applying the subset construction to $\mathcal{A}'$ if necessary, we end up with a deterministic automaton equivalent to $\mathcal{A}$.

Example 4.8. If we apply the above construction to the automaton with 1-transitions



we obtain the automaton



In conclusion, nondeterministic automata and automata with 1-transitions are all equivalent to deterministic automata.

4.4 Kleene's Theorem

We will soon verify that the recognizable languages over a finite alphabet are precisely the rational languages. First, we establish some closure properties of languages recognized by finite automata.

Proposition 4.9. *Let $K, L \subseteq A^*$ be recognizable languages. Then the languages $K \cup L$, $A^* \setminus K$, $K \cap L$, KL and K^* are also recognizable.*

Proof. Let $\mathcal{A} = (Q, A, E, I, T)$ and $\mathcal{A}' = (Q', A', E', I', T')$ be finite automata recognizing K and L , respectively. We will prove the statement by constructing finite automata recognizing each of the languages. Renaming the vertices in Q if necessary, we may assume that Q and Q' are disjoint.

It is easy to see that $\mathcal{U} = (Q \cup Q', A, E \cup E', I \cup I', T \cup T')$ is an automaton recognizing the language $K \cup L$. Indeed, any successful path in $\mathcal{A}$ or $\mathcal{A}'$ is still a successful path in $\mathcal{U}$;

conversely, a successful path in $\mathcal{U}$ must correspond to a successful path in $\mathcal{A}$ or $\mathcal{A}'$.

Regarding $A^* \setminus K$, we can apply the subset construction to $\mathcal{A}$ in order to obtain a complete and deterministic automaton recognizing K . Then, if we declare each terminal state to be nonterminal and each nonterminal state to be terminal, we get a complete and deterministic automaton recognizing the language $A^* \setminus K$.

As for the intersection $K \cap L$, we begin by defining the direct product of $\mathcal{A}$ and $\mathcal{A}'$ to be the automaton

$$\mathcal{A} \times \mathcal{A}' = (Q \times Q', A, E'', I \times I', T \times T')$$

where E'' is such that

$$((p, p'), a, (q, q')) \in E'' \Leftrightarrow ((p, a, q) \in E \wedge (p', a, q') \in E')$$

for all $p, q \in Q$, $p', q' \in Q'$ and $a \in A$. Now it is enough to observe that $(p, q) \xrightarrow{u} (p', q')$ is a successful path in $\mathcal{A} \times \mathcal{A}'$ if and only if $p \xrightarrow{u} q$ and $p' \xrightarrow{u} q'$ are successful paths in $\mathcal{A}$ and $\mathcal{A}'$, respectively. Therefore, $L(\mathcal{A} \times \mathcal{A}') = K \cap L$.

Consider now the product of languages KL . Let $\mathcal{P}$ be the automaton with 1-transitions given by

$$\mathcal{P} = (Q \cup Q', A, E \cup E' \cup (T \times \{1\} \times I'), I, T').$$

Then $w \in A^*$ labels a successful path in $\mathcal{P}$ if and only if there exist $u \in K$, $v \in L$ satisfying $w = uv$ and labeling a path

$$i \xrightarrow{u} t \xrightarrow{1} i' \xrightarrow{v} t'$$

in $\mathcal{P}$, for some $i \in I$, $t \in T$, $i' \in I'$ and $t' \in T'$. So $L(\mathcal{A} \times \mathcal{A}') = \{uv \mid u \in K, v \in L\} = KL$.

Finally, in order to construct an automaton recognizing the language K^* , we begin by considering a new vertex $r \notin Q$. Then we define the automaton with 1-transitions

$$\mathcal{S} = (Q \cup \{r\}, A, E \cup (\{r\} \times \{1\} \times I) \cup (T \times \{1\} \times \{r\}), r, r).$$

Since every successful path $I \ni i \xrightarrow{u} t \in T$ in $\mathcal{A}$ yields a successful path $r \xrightarrow{1} i \xrightarrow{u} t$ in $\mathcal{S}$, it follows that $K^* \subseteq L(\mathcal{S})$. On the other hand, every $u \in L(\mathcal{S})$ labels a path which factors into subpaths of the type $r \xrightarrow{1} i \xrightarrow{u} t$, for some $i \in I$ and $t \in T$. Therefore, $L(\mathcal{S}) \subseteq K^*$ and thus $L(\mathcal{S}) = K^*$. $\square$

Having established these closure properties of recognizable languages, we are ready to prove the following result, named after the mathematician Stephen Cole Kleene.

Theorem 4.10 (Kleene's Theorem). *Let A be a finite alphabet. A language $L \subseteq A^*$ is recognizable if and only if it is rational.*

Proof. If L is a rational language, then L is obtained from finite languages by applying the operators union, product and star, and it is very easy to construct an automaton recognizing a given finite language. In view of the previous proposition, the class of recognizable languages is closed under the rational operators, and hence every rational language is recognizable.

Suppose now that L is recognizable. Then we can take an automaton $\mathcal{A} = (Q, A, E, I, T)$ with $Q = \{1, 2, \dots, m\}$, for some $m \geq 1$, such that $L = L(\mathcal{A})$. Our goal is to find a rational expression for $L(\mathcal{A})$. Given $i, j \in Q$ and $k \in \{0, 1, \dots, m\}$, we define $L_{ij}^{(k)}$ to be the set of all words $u \in A^*$ labeling a path $i \xrightarrow{u} j$ such that $\ell \leq k$ for every intermediate vertex ℓ . Observe that

$$L = \bigcup_{i \in I} \bigcup_{t \in T} L_{it}^{(m)},$$

so it is enough to prove that every language of the form $L_{ij}^{(k)}$ is rational, and we will do so by induction on k . If $k = 0$, there are no intermediate vertices in a path from i to j , so either $i = j$ or we have a path $i \xrightarrow{a} j$ for some $a \in A$. Therefore, $L_{ij}^{(0)} \subseteq A \cup \{1\}$ is always a finite language, and thus rational. We proceed by assuming that $k \geq 1$. We aim to prove that

$$L_{ij}^{(k)} = L_{ij}^{(k-1)} \cup L_{ik}^{(k-1)} \left(L_{kk}^{(k-1)} \right)^* L_{kj}^{(k-1)}$$

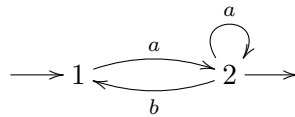
for all $i, j \in Q$, which will establish the result by the induction hypothesis. On the one hand, we have $L_{ij}^{(k-1)} \subseteq L_{ij}^{(k)}$, and it is also clear that $L_{ik}^{(k-1)} \left(L_{kk}^{(k-1)} \right)^* L_{kj}^{(k-1)} \subseteq L_{ij}^{(k)}$, so the opposite inclusion holds. On the other hand, if $u \in L_{ij}^{(k)}$ and $u \notin L_{ij}^{(k-1)}$, then there exist $x, y, z \in A^*$ such that $u = xyz$ labels a path

$$i \xrightarrow{x} k \xrightarrow{y} k \xrightarrow{z} j,$$

where k appears for the first time after reading x and for the last time before reading z . It follows that $x \in L_{ik}^{(k-1)}$, $y \in \left(L_{kk}^{(k-1)} \right)^*$ and $z \in L_{kj}^{(k-1)}$, thereby establishing the other inclusion. Applying this recursive relation a finite number of times, we end up with a rational expression for L , which concludes the proof. $\square$

The theorem above is only true under the assumption that A is a finite alphabet. This is a restriction that we will impose several times, for it is the only case we are truly interested in. Observe that the proof of Kleene's Theorem gives a method to derive a rational expression for the language recognized by a finite automaton, as we illustrate in the example that follows.

Example 4.11. If $\mathcal{A}$ is the automaton depicted by



then we obtain

$$\begin{aligned} L(\mathcal{A}) &= L_{12}^{(2)} \\ &= L_{12}^{(1)} \cup L_{12}^{(1)} \left(L_{22}^{(1)} \right)^* L_{22}^{(1)} \\ &= a \cup a(1 \cup a \cup ba)^*(1 \cup a \cup ba) \\ &= a(a \cup ba)^*. \end{aligned}$$

We end this section by presenting a simple, yet useful, result.

Proposition 4.12. *Let A and B be two finite alphabets. If $L \subseteq A^*$ is a rational language and $\varphi: A^* \rightarrow B^*$ is a homomorphism, then $L\varphi$ is also rational.*

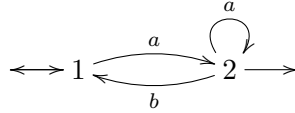
Proof. By Theorem 4.10, there exists a finite automaton $\mathcal{A} = (Q, A, E, I, T)$ recognizing L . If we replace each edge with label $a \in A$ in E by a sequence of edges labeled by $a\varphi$, we obtain an automaton (possibly with 1-transitions) recognizing the language $L\varphi$. Therefore, $L\varphi$ is a recognizable language, whence rational. $\square$

4.5 The transition monoid of an automaton

In this section, we establish a link between finite deterministic automata and finite monoids.

Let $\mathcal{A} = (Q, A, \delta, q_0, T)$ be a deterministic automaton, where $\delta: Q \times A \rightarrow Q$ is the so-called transition function, a partial transformation. We can consider an extension $\hat{\delta}: Q \times A^* \rightarrow Q$ of δ by letting $(q, u)\hat{\delta}$ be the state reached after reading the word $u \in A^*$ from $q \in Q$ by following the labels on the edges (whenever that is possible). Formally, we can create such a partial function $\hat{\delta}$ by induction over the length of a word $u \in A^*$. We begin by defining $(q, 1)\hat{\delta} = q$ for all $q \in Q$. To proceed, assume that $\hat{\delta}$ is already defined for words of length $n \geq 0$, and let $u \in A^*$ be such that $|u| = n + 1$. Then we can write $u = va$ for some $v \in A^n$ and $a \in A$. By hypothesis, $(q, v)\hat{\delta}$ is already determined, and we define $(q, w)\hat{\delta} = ((q, v)\hat{\delta}, a)\delta$. In this way, we obtain by induction a partial transformation $\hat{\delta}: Q \times A^* \rightarrow Q$ extending δ . By abuse of notation, we will always write δ instead of $\hat{\delta}$, and we also call $\hat{\delta}$ the transition function of the automaton.

Example 4.13. If $\delta: \{1, 2\} \times \{a, b\}^* \rightarrow \{1, 2\}$ is the transition function of the deterministic automaton



then $(1, ab)\delta = ((1, a)\delta, b)\delta = (2, b)\delta = 1$, which means that ab labels a loop at 1.

Given $u \in A^*$, we further define $\delta_u: Q \rightarrow Q$ by $q\delta_u = (q, u)\delta$, so that $\delta_u \in \mathcal{PT}_Q$ for all $u \in A^*$. It is easy to check that the map

$$\begin{aligned} \Delta: A^* &\longrightarrow \mathcal{PT}_Q \\ u &\longmapsto \delta_u \end{aligned}$$

is a monoid homomorphism; hence, the image of Δ is a submonoid of $\mathcal{PT}_Q$. We call it the **transition monoid** of $\mathcal{A}$ and denote it by $M(\mathcal{A})$.

We now exhibit a concrete way of computing the transition monoid of an automaton. Let $\mathcal{A} = (Q, A, \delta, q_0, T)$ be a finite deterministic automaton over a finite alphabet A and let $M_k = \{\delta_u \mid u \in A^k\}$ be the set of all partial transformations on Q induced by words of length $k \geq 0$. Since Q is finite, $\mathcal{PT}_Q$ and $M(\mathcal{A}) \leq \mathcal{PT}_Q$ are also finite, so there exists $n \geq 0$ such

that $M_n \subseteq \bigcup_{k=0}^{n-1} M_k$. This entails

$$M_{n+i} \subseteq \bigcup_{k=0}^{n-1} M_k$$

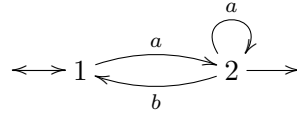
for all $i \geq 0$, as we now show by induction. The assertion is certainly true for $i = 0$. We then assume that it holds for all values less than $i > 0$. Given $u \in A^{n+i}$, we can write $u = va$ for some $v \in A^{n+i-1}$ and $a \in A$; hence, by hypothesis, $\delta_v \in M_j$ for some $0 \leq j \leq n-1$. Consequently,

$$\delta_u = \delta_{va} \in M_{j+1} \subseteq \bigcup_{k=0}^n M_k = \bigcup_{k=0}^{n-1} M_k,$$

by definition of n , and the claim is established by induction.

The above implies that $M(\mathcal{A}) = \bigcup_{k=0}^{n-1} M_k$. So, to compute $M(\mathcal{A})$, we only need to work out the partial transformations on Q induced by words of length less than n . Moreover, if $w \in A^*$ has $u \in A^*$ as a factor and $\delta_u = \delta_v$ with $|v| < |u|$, then we can replace every occurrence of u in w by v to obtain a word $w' \in A^*$ that satisfies $\delta_w = \delta_{w'}$ and $|w'| < |w|$. And if $\delta_u = \emptyset$ is the empty partial transformation, we can ignore all words having u as a factor. Finally, observe that $M(\mathcal{A})$ does not depend on either initial or terminal states.

Example 4.14. Let us compute the transition monoid of the deterministic automaton $\mathcal{A} = (\{1, 2\}, \{a, b\}, \delta, \{1\}, \{1, 2\})$ depicted by



We represent a partial transformation $\varphi \in \mathcal{PT}_{\{1,2\}}$ by $\varphi = (1\varphi \ 2\varphi)$ and, if $i \notin \text{dom } \varphi$, we replace $i\varphi$ by $-$. Since

$$\begin{array}{lll} \delta_1 = (1 \ 2) & \delta_a = (2 \ 2) = \delta_{a^2} = \delta_{aba} & \delta_b = (- \ 1) = \delta_{bab} \\ \delta_{ab} = (1 \ 1) & \delta_{ba} = (- \ 2) & \delta_{b^2} = (- \ -), \end{array}$$

we conclude that

$$M(\mathcal{A}) = \{(1 \ 2), (1 \ 1), (2 \ 2), (- \ 2), (- \ 1), (- \ -)\} = \bigcup_{k=0}^2 M_k.$$

We will now determine a finite presentation for $M(\mathcal{A})$. Recall that, given a relation R on a monoid M , $R^\#$ denotes the smallest congruence on M containing R .

Proposition 4.15. *Let $\mathcal{A} = (Q, A, \delta, q_0, T)$ be a deterministic automaton over a finite alphabet A . Let $n \geq 0$ and $\Delta: A^* \rightarrow \mathcal{PT}_Q$ be as above, and define a relation $R \subseteq A^* \times A^*$ by*

$$(u, v) \in R \Leftrightarrow (|u|, |v| \leq n \wedge \delta_u = \delta_v).$$

Then $\text{Mon}\langle A \mid R \rangle$ is a finite presentation for $M(\mathcal{A})$.

Proof. It is clear that R is finite and $R \subseteq \ker \Delta$, so $R^\# \subseteq \ker \Delta$. Since $\text{im } \Delta = M(\mathcal{A})$, the homomorphism Δ induces a surjective homomorphism $\tilde{\Delta}: A^*/R^\# \rightarrow M(\mathcal{A})$, and thus $|A^*/R^\#| \geq |M(\mathcal{A})|$. On the other hand, one easily proves by induction on word length that, for every $u \in A^*$, there exists $v \in \bigcup_{k=0}^{n-1} A^k$ such that $\delta_u = \delta_v$, which implies $(u, v) \in R^\#$. But then we get a well-defined surjection

$$\begin{aligned} M(\mathcal{A}) &\longrightarrow A^*/R^\# \\ \delta_u &\longmapsto uR^\# \end{aligned}$$

because, if $u, w \in A^*$ are such that $\delta_u = \delta_w$, then we can take $v \in A^*$ such that $|v| < n$ and $\delta_u = \delta_v = \delta_w$, so $uR^\# = vR^\# = wR^\#$. Consequently, we get $|M(\mathcal{A})| \geq |A^*/R^\#|$, and hence $|M(\mathcal{A})| = |A^*/R^\#|$. This shows that the surjective homomorphism $\tilde{\Delta}$ is, in fact, an isomorphism. We conclude that $M(\mathcal{A}) \cong A^*/R^\#$, and therefore $\text{Mon}\langle A \mid R \rangle$ is a finite presentation for $M(\mathcal{A})$. $\square$

4.6 The syntactic monoid of a language

In this section, we establish another connection between languages and monoids.

Let A be an alphabet and $L \subseteq A^*$ be a language. We define an equivalence relation $\sim_L$ on A^* by

$$u \sim_L v \text{ if } \forall x, y \in A^* (xuv \in L \Leftrightarrow xvy \in L).$$

Let us verify that $\sim_L$ is actually a congruence on A^* . Let $u, v \in A^*$ be such that $u \sim_L v$. For all $x, y, w \in A^*$, we have

$$x(wu)y \in L \Leftrightarrow (xw)uy \in L \Leftrightarrow (xw)vy \in L \Leftrightarrow x(wv)y \in L,$$

so $wu \sim_L wv$. By a similar argument, we conclude that $uw \sim_L vw$, thereby proving that $\sim_L$ is a congruence on A^* . We define the **syntactic monoid** of L to be $\text{Synt}(L) = A^*/\sim_L$. The next proposition shows, in particular, that L is a union of congruence classes of $\sim_L$. For that reason, we say that $\sim_L$ **saturates** L .

Proposition 4.16. *Let A be an alphabet and let $L \subseteq A^*$ be a language. Then $\sim_L$ is the greatest congruence on A^* saturating L .*

Proof. To prove that L is a union of congruence classes of $\sim_L$, it suffices to observe that, given $u \in L$ and $v \in A^*$ such that $v \sim_L u$, then $1u1 = u \in L$ implies $1v1 = v \in L$.

Next, if $\sigma \subseteq A^* \times A^*$ is a congruence that saturates L , then

$$\begin{aligned} u\sigma v &\Leftrightarrow \forall x, y \in A^* \ xuy \sigma xvy \\ &\Rightarrow \forall x, y \in A^* \ (xuy \in L \Leftrightarrow xvy \in L) \\ &\Leftrightarrow u \sim_L v. \end{aligned}$$

We conclude that $\sigma \subseteq \sim_L$, and the claim follows. $\square$

Let us proceed with some more definitions. Given two monoids M and N , we say that M **divides** N if M is a homomorphic image of a submonoid of N . Given a language $L \subseteq A^*$, we say that L is **recognized by the monoid** M if there exist $P \subseteq M$ and a homomorphism $\varphi: A^* \rightarrow M$ such that $L = P\varphi^{-1}$, which is the same as requiring $L = L\varphi\varphi^{-1}$. In this case, $\ker \varphi$ is a congruence on A^* that saturates L , so $\ker \varphi \subseteq \sim_L$. Consequently, if $N = \text{im } \varphi$, then there exists a surjective homomorphism $N \rightarrow \text{Synt}(L)$. This shows that the syntactic monoid $\text{Synt}(L)$ of a language L always divides any monoid recognizing L .

The following theorem allows us to determine whether or not a language is rational by looking at its syntactic monoid.

Theorem 4.17. *Let A be a finite alphabet and $L \subseteq A^*$ be a language. The following conditions are equivalent:*

- (i) L is a rational language.
- (ii) $\text{Synt}(L)$ is finite.
- (iii) L is recognized by a finite monoid.

Proof. (i) $\Rightarrow$ (ii): According to Theorem 4.10, there exists a finite deterministic automaton $\mathcal{A} = (Q, A, \delta, q_0, T)$ that recognizes L . Consider the map given by

$$\begin{aligned} \phi: M(\mathcal{A}) &\longrightarrow \text{Synt}(L) \\ \delta_u &\longmapsto u \sim_L. \end{aligned}$$

We check that ϕ is well defined. Let $u, v \in A^*$ be such that $\delta_u = \delta_v$, take $x, y \in A^*$ and suppose that $xuy \in L$. Then we have a path

$$q_0 \xrightarrow{x} p \xrightarrow{u} q \xrightarrow{y} t$$

in $\mathcal{A}$, for some $p, q \in Q$ and $t \in T$. Since $\delta_u = \delta_v$, it follows that $(p, u)\delta = (p, v)\delta = q$, so we also have the path

$$q_0 \xrightarrow{x} p \xrightarrow{v} q \xrightarrow{y} t$$

in $\mathcal{A}$, and therefore $xvy \in L$. It follows from symmetry that $u \sim_L v$, so ϕ is a well-defined surjective mapping. In fact, ϕ is a surjective homomorphism because $\delta_1\phi = 1 \sim_L$ and

$$(\delta_u\delta_v)\phi = (\delta_{uv})\phi = (uv) \sim_L = (u \sim_L)(v \sim_L) = (\delta_u\phi)(\delta_v\phi).$$

Since $M(\mathcal{A})$ is finite, we conclude that $\text{Synt}(L)$ is finite too.

(ii) $\Rightarrow$ (iii): Let $\varphi: A^* \rightarrow \text{Synt}(L)$ be the quotient morphism. Since L is a union of $\sim_L$ -classes, we deduce that $L = (L\varphi)\varphi^{-1}$, so (iii) follows from the assumption that $\text{Synt}(L)$ is finite.

(iii) $\Rightarrow$ (i): Let M be a finite monoid recognizing L via the homomorphism $\varphi: A^* \rightarrow M$, and let $P \subseteq M$ be such that $L = P\varphi^{-1}$. Consider the finite, deterministic and complete automaton $\mathcal{A} = (M, A, \delta, 1, P)$, where $(m, a)\delta = m(a\varphi)$ for all $m \in M$ and $a \in A$. Then

$(m, u)\delta = m(u\varphi)$ for all $m \in M$ and $u \in A^*$, so

$$u \in L \Leftrightarrow u\varphi \in P \Leftrightarrow (1, u)\delta \in P \Leftrightarrow u \in L(\mathcal{A}).$$

We conclude that L is recognized by the finite automaton $\mathcal{A}$, and thus it is a rational language. $\square$

The following property is similar to Proposition 4.12 and its proof makes use of the equivalence (i) $\Leftrightarrow$ (iii) above.

Proposition 4.18. *Let A and B be finite alphabets. If $L \subseteq B^*$ is a rational language and $\varphi: A^* \rightarrow B^*$ is a homomorphism, then $L\varphi^{-1}$ is also rational.*

Proof. Since L is rational, there exists a finite monoid M , a homomorphism $\psi: B^* \rightarrow M$ and a subset $P \subseteq M$ such that $L = P\psi^{-1}$. Consequently,

$$L\varphi^{-1} = P\psi^{-1}\varphi^{-1} = P(\varphi\psi)^{-1},$$

which shows that M recognizes $L\varphi^{-1}$ via the homomorphism $\varphi\psi$. We conclude that $L\varphi^{-1}$ is a rational language. $\square$

We end this section with an example.

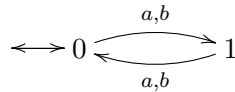
Example 4.19. Let $A = \{a, b\}$ and let $L \subseteq A^*$ be the language of all words with even length. Then the monoid (in fact, group) $\mathbb{Z}_2$ recognizes L via the homomorphism

$$\begin{aligned} \varphi: A^* &\longrightarrow \mathbb{Z}_2 \\ a &\longmapsto 1 \\ b &\longmapsto 1 \end{aligned}$$

since $L = 0\varphi^{-1}$. Moreover, given $u, v \in A^*$, we have

$$\begin{aligned} u \sim_L v &\Leftrightarrow \forall x, y \in A^* (|xuy| \text{ is even} \Leftrightarrow |xvy| \text{ is even}) \\ &\Leftrightarrow |u| \text{ and } |v| \text{ have the same parity.} \end{aligned}$$

Considering $\mathbb{Z}_2 = \{0, 1\}$ with addition modulo 2, the automaton $\mathcal{A}$ constructed in the proof of (iii) $\Rightarrow$ (i) in Theorem 4.6 is depicted by



Its transition monoid is $M(\mathcal{A}) = \{(0 \ 1), (1 \ 0)\} \simeq \mathbb{Z}_2$.

4.7 Minimal automata

Now that we know what transition and syntactic monoids are, it is natural to try to relate both, especially because the first ones are easy to compute, but the latter ones seem to involve more cumbersome work. In this section, we use automata to establish such a connection.

Given a rational language, there are many different automata that recognize it, even if we focus only on deterministic ones. Indeed, one can add states that are not accessible or co-accessible without changing the language recognized by an automaton. But can we speak of a “canonical” deterministic automaton that recognizes a certain language? As we are about to see, the answer is “yes”.

Let A be an alphabet and $L \subseteq A^*$ a nonempty rational language. We start by considering a deterministic automaton recognizing L . Then we delete all vertices which fail to be accessible or co-accessible, thereby obtaining a trim automaton $\mathcal{A} = (Q, A, \delta, q_0, T)$. Given $P, R \subseteq Q$, we denote by L_{PR} the set of all words labelling a path in $\mathcal{A}$ from a vertex in P to a vertex in R ; in particular, $L(\mathcal{A}) = L_{q_0T}$. Our goal is to dispose of all states that are somewhat redundant, so we define an equivalence relation $\sim$ on Q , known as the **Hopcroft equivalence**, by

$$p \sim q \Leftrightarrow L_{pT} = L_{qT}.$$

For all $q \in Q$, let $[q] = q \sim$ and consider the automaton $\mathcal{A}_m = \{Q_m, A, \delta_m, [q_0], T_m\}$, where $Q_m = \{[q] \mid q \in Q\}$, $T_m = \{[t] \mid t \in T\}$ and

$$\begin{aligned} \delta_m: Q_m \times A &\longrightarrow Q_m \\ ([q], a) &\longmapsto [(q, a)\delta]. \end{aligned}$$

To verify that δ_m is well defined, we must show that if $([q], a) \in \text{dom } \delta_m$ and $p \sim q$, then $L_{(p,a)\delta T} = L_{(q,a)\delta T}$. For that purpose, take $p, q \in Q$ and $a \in A^*$ satisfying $p \sim q$ and $(q, a)\delta = q' \in Q$. Then there exist $t \in T$, $u \in A^*$ and a path

$$q \xrightarrow{a} q' \xrightarrow{u} t$$

in $\mathcal{A}$, since $\mathcal{A}$ is trim. Now $au \in L_{q'T}$ implies $au \in L_{p'T}$, so $(p, a)\delta \in Q$ and there is a path

$$p \xrightarrow{a} p' \xrightarrow{u} t'$$

in $\mathcal{A}$, for some $t' \in T$ and $p' = (p, a)\delta$. Given $v \in L_{p'T}$, we get $av \in L_{pT} = L_{qT}$, and hence $v \in L_{q'T}$. Out of symmetry, we deduce that $L_{p'T} = L_{q'T}$, which means that $p' \sim q'$. We have thus proved that δ_m is indeed well defined, so $\mathcal{A}_m$ is a deterministic automaton.

Our goal now is to show that $L(\mathcal{A}) = L(\mathcal{A}_m)$. Observe that a path

$$p_1 \xrightarrow{u_1} p_2 \xrightarrow{u_2} \dots \xrightarrow{u_{n-1}} p_n$$

in $\mathcal{A}$ yields a path

$$[p_1] \xrightarrow{u_1} [p_2] \xrightarrow{u_2} \dots \xrightarrow{u_{n-1}} [p_n]$$

in $\mathcal{A}_m$, so we simultaneously conclude that $L(\mathcal{A}) \subseteq L(\mathcal{A}_m)$ and that $\mathcal{A}_m$ is trim. Next, we prove that $L(\mathcal{A}_m) \subseteq L(\mathcal{A})$. First we show that if $q \in Q$ and $u \in L_{[q]T_m}$, then $u \in L_{qT}$. We proceed by induction on u . Regarding the case $u = 1$, we have

$$1 \in L_{[q]T_m} \Leftrightarrow \exists t \in T [q] = [t] \Leftrightarrow \exists t \in T L_{qT} = L_{tT}.$$

Since $1 \in L_{tT}$ for all $t \in T$, the implication $1 \in L_{[q]T_m} \Rightarrow 1 \in L_{qT}$ holds. Suppose now that $u \in A^*$ is such that $u \in L_{[q]T_m}$ implies $u \in L_{qT}$, for all $q \in Q$. If $a \in A$ and $q \in Q$ are such that $au \in L_{[q]T_m}$, there is a path

$$[q] \xrightarrow{a} [r] \xrightarrow{u} [t]$$

in $\mathcal{A}_m$, for some $r \in Q$ and $t \in T$. Consequently, there exist $q', r' \in Q$ such that $L_{qT} = L_{q'T}$, $L_{rT} = L_{r'T}$ and $q' \xrightarrow{a} r'$ is a path in $\mathcal{A}$. As $u \in L_{rT} = L_{r'T}$ by hypothesis, it follows that $au \in L_{q'T} = L_{qT}$. Hence, by induction, the implication holds for all $q \in Q$ and $u \in A^*$. In particular, we get $L_{[q_0]T_m} \subseteq L_{q_0T}$, which means that $L(\mathcal{A}_m) \subseteq L(\mathcal{A})$. Therefore, $L(\mathcal{A}) = L(\mathcal{A}_m)$ and the automata $\mathcal{A}$ and $\mathcal{A}_m$ are equivalent.

The next proposition shows that if we apply the above construction to equivalent automata, we obtain isomorphic automata.

Proposition 4.20. *Let $\mathcal{A} = (Q, A, \delta, q_0, T)$ and $\mathcal{A}' = (Q', A, \delta', q'_0, T')$ be two equivalent, finite, deterministic and trim automata. Then $\mathcal{A}_m \cong \mathcal{A}'_m$.*

Proof. In order to simplify the notation, given $q \in Q$, $q' \in Q'$ and $u \in A^*$, let $qu = (q, u)\delta$ and $q'u = (q', u)\delta'$. Consider the function

$$\begin{aligned} \varphi: Q_m &\longrightarrow Q'_m \\ [q_0u] &\longmapsto [q'_0u]. \end{aligned}$$

Let us check that φ is well defined. First, note that every $q \in Q$ is of the type q_0u for some $u \in A^*$, since $\mathcal{A}$ is trim. Given $u, v \in A^*$ satisfying $[q_0u] = [q_0v]$, we must show that $[q'_0u] = [q'_0v]$. So assume that $L_{(q_0u)T} = L_{(q_0v)T}$ and take $w \in L_{(q'_0u)T}$. Since uw labels a successful path in $\mathcal{A}'$ and $L(\mathcal{A}) = L(\mathcal{A}')$, it follows that uw labels a successful path in $\mathcal{A}$, so $w \in L_{(q_0u)T} = L_{(q_0v)T}$. Consequently, $vw \in L(\mathcal{A}) = L(\mathcal{A}')$, and hence $w \in L_{(q'_0v)T}$, which implies that $L_{(q'_0u)T} \subseteq L_{(q'_0v)T}$. Out of symmetry, we conclude that $L_{(q'_0u)T} = L_{(q'_0v)T}$; therefore, φ is a well-defined function. Moreover, φ is surjective due to $\mathcal{A}'$ being trim.

We proceed by showing that φ is a morphism of automata. First observe that $[q_0]\varphi = [q'_0]$, so the initial state of $\mathcal{A}_m$ is mapped to the initial state of $\mathcal{A}'_m$. Next, given $q_0u \in T$, we have $u \in L(\mathcal{A}) = L(\mathcal{A}')$, and as a result $q'_0u \in T'$. This implies that $[q_0u]\varphi = [q'_0u] \in T'_m$, so $T_m\varphi = T'_m$. Finally, given $p, q \in Q$ and $a \in A$ satisfying $([p], a)\delta_m = [q]$, we must show that $([p]\varphi, a)\delta'_m = [q]\varphi$. If $u, v \in A^*$ are such that $p = q_0u$ and $q = q_0v$, then $[p]\varphi = [q'_0u]$ and $[q]\varphi = [q'_0v]$. And since we have $([q_0u], a)\delta_m = [q_0ua] = [q_0v]$, by the above reasoning we

conclude that $[q'_0ua] = [q'_0v]$. Therefore,

$$([p]\varphi, a)\delta'_m = ([q'_0u], a)\delta'_m = [q'_0ua] = [q'_0v] = [q]\varphi,$$

as we wanted.

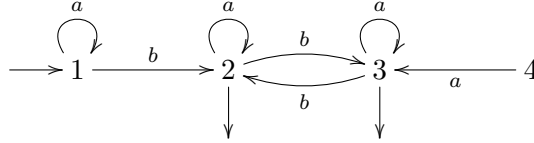
Consequently, φ is a surjective morphism from $\mathcal{A}_m$ to $\mathcal{A}'_m$ and, out of symmetry, there is an analogous morphism ψ from $\mathcal{A}'_m$ onto $\mathcal{A}_m$; in particular, $|Q_m| = |Q'_m|$. Note also that φ and ψ are mutually inverse bijections that respect initial and terminal vertices. To finish, observe that for all $u, v \in A^*$ and $a \in A$, we have

$$([q_0u]\varphi\psi, a)\delta_m = ([q'_0u]\psi, a)\delta_m = ([q_0u], a)\delta_m,$$

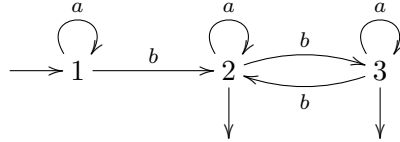
so $\varphi\psi$ fixes all edges in $\mathcal{A}_m$ and, similarly, $\psi\varphi$ fixes the edges in $\mathcal{A}'_m$. We conclude that φ and ψ are mutually inverse isomorphisms, whence $\mathcal{A}_m \cong \mathcal{A}'_m$. $\square$

This proposition shows that, up to isomorphism, there exists a unique deterministic automaton recognizing a given rational language $L \subseteq A^*$ with minimum number of vertices. For that reason, we call $\mathcal{A}_m$ the **minimal automaton** of L and denote it by $\min_L$.

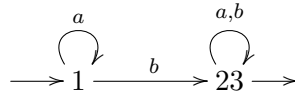
Example 4.21. Let us illustrate how one “minimizes” the deterministic automaton $\mathcal{A}$ below.



The automaton $\mathcal{A}$ fails to be trim because vertex 4 is not accessible from the initial state 1. If we remove it, though, we obtain a trim automaton:



The set of terminal states is $T = \{2, 3\}$ and $L_{2T} = (a \cup b)^* = L_{3T}$, so $2 \sim 3$. On the other hand, $1 \not\sim 2$ because the empty word belongs to $L_{2T} \setminus L_{1T}$. We conclude that the minimal automaton equivalent to $\mathcal{A}$ is



Finally, we prove that the syntactic monoid of a rational language is isomorphic to the transition monoid of the minimal automaton recognizing it.

Proposition 4.22. *Let $L \subseteq A^*$ be a rational language. Then $\text{Synt}(L) \cong M(\min_L)$.*

Proof. Let $\mathcal{A} = \min_L = (Q, A, \delta, q_0, T)$ and $\Delta: A^* \rightarrow \mathcal{PT}_Q, u \mapsto \delta_u$. Then $M(\mathcal{A}) \cong A^*/\ker \Delta$

and $\text{Synt}(L) = A^*/\sim_L$, so it is enough to prove that $\ker \Delta = \sim_L$. It follows from the proof of (i) $\Rightarrow$ (ii) in Theorem 4.17 that $\ker \Delta \subseteq \sim_L$, so we focus on the reverse inclusion. Suppose that $u, v \in A^*$ are such that $u \sim_L v$. Take any $p \in Q$ and assume that $(p, u)\delta = q$. Then there exists a path

$$q_0 \xrightarrow{x} p \xrightarrow{u} q \xrightarrow{y} t \in T$$

in $\mathcal{A}$, for some $x, y \in A^*$. This implies that $xuy \in L$, whence $xvy \in L$. Therefore, we also have a path

$$q_0 \xrightarrow{x} p \xrightarrow{v} q' \xrightarrow{y} t' \in T$$

in $\mathcal{A}$, for some $q' \in Q$. Now, if $w \in L_{qT}$, then $xuw \in L \Rightarrow xvw \in L$, and thus $w \in L_{q'T}$. Hence, $L_{qT} \subseteq L_{q'T}$ and, out of symmetry, we obtain $L_{qT} = L_{q'T}$. It follows from the minimality of $\mathcal{A}$ that $q = q'$, so $(p, u)\delta = (p, v)\delta$. Since $p \in Q$ was arbitrary, we conclude that $\delta_u = \delta_v$, thereby implying $\sim_L \subseteq \ker \Delta$ and completing the proof. $\square$

4.8 Inverse automata

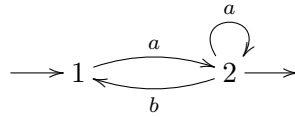
The class of automata that will be of most interest to us is the class of inverse automata, so we will now define these objects.

Let A be an alphabet and $\tilde{A} = A \cup A^{-1}$ be as in Section 3.5. We say that an automaton $\mathcal{A} = (Q, \tilde{A}, E, I, T)$ is **involutive** if, for all $p, q \in Q$ and $a \in A$, we have

$$(p, a, q) \in E \Leftrightarrow (q, a^{-1}, p) \in E.$$

These pairs of edges are considered inverses of each other, and an edge labeled by a letter in A (respectively, in A^{-1}) is said to be positive (respectively, negative). When we know beforehand that a certain automaton is involutive, we only draw the positive edges; the negative ones are like “ghost” edges that we visualize only in our heads. Moreover, we denote by $E^+ \subseteq E$ and $E^- \subseteq E$ the subsets of positive and negative edges, in that order. We define an **inverse automaton** to be a deterministic, trim and involutive automaton with a unique terminal state.

Example 4.23. Consider the involutive automaton over the alphabet $\{a, a^{-1}, b, b^{-1}\}$:



If E denotes its set of edges, then

$$\begin{aligned} E^+ &= \{(1, a, 2), (2, a, 2), (2, b, 1)\}, \\ E^- &= \{(2, a^{-1}, 1), (2, a^{-1}, 2), (1, b^{-1}, 2)\}. \end{aligned}$$

This automaton is not inverse because it fails to be deterministic, as there are two distinct edges labeled by a^{-1} leaving vertex 2.

We finish this section by collecting two properties of inverse automata that will be helpful later.

Proposition 4.24. *Every inverse automaton is a minimal automaton.*

Proof. Let $\mathcal{A} = (Q, \tilde{A}, \delta, q_0, t)$ be an inverse automaton. By definition, $\mathcal{A}$ is deterministic and trim, so we only need to show that $p \sim q$ implies $p = q$ for all $p, q \in Q$, where $\sim$ denotes the Hopcroft congruence. Suppose that $L_{pt} = L_{qt}$ and let $v \in L_{pt} = L_{qt}$. Since we have the paths $p \xrightarrow{v} t$ and $q \xrightarrow{v} t$ in $\mathcal{A}$, it follows that we also have the paths $t \xrightarrow{v^{-1}} p$ and $t \xrightarrow{v^{-1}} q$ in $\mathcal{A}$. By determinism, we conclude that $p = q$, which establishes the result. $\square$

In view of the preceding proposition, if $L \subseteq \tilde{A}^*$ is a language recognized by an inverse automaton $\mathcal{A}$, then $\text{Synt}(L) \cong M(\mathcal{A})$. And there is something more we can say about $M(\mathcal{A})$.

Proposition 4.25. *The transition monoid of an inverse automaton is an inverse monoid.*

Proof. Let $\mathcal{A} = (Q, \tilde{A}, \delta, q_0, t)$ be an inverse automaton. We already know that $M(\mathcal{A})$ is a submonoid of $\mathcal{PT}_Q$, so it is enough to prove that every $\delta_u \in M(\mathcal{A})$ is injective and $M(\mathcal{A})$ is closed under taking inverses, for that will imply that $M(\mathcal{A})$ is an inverse submonoid of $\mathcal{I}_Q$.

Suppose that $p, p' \in Q$ are such that $(p, u)\delta = (p', u)\delta = q \in Q$. If $u = 1$, it is clear that $p = p'$. Otherwise, there exist $n \geq 1$ and $a_1, a_2, \dots, a_n \in \tilde{A}^*$ such that $u = a_1 a_2 \dots a_n$ and u labels the paths

$$\begin{aligned} p &\xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_{n-1}} p_{n-1} \xrightarrow{a_n} q, \\ p' &\xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_{n-1}} p_{n-1} \xrightarrow{a_n} q \end{aligned}$$

in $\mathcal{A}$, for some $p_i \in Q$ ($1 \leq i < n$). Since $\mathcal{A}$ is involutive, we can traverse each path in the reverse direction, obtaining the paths

$$\begin{aligned} q &\xrightarrow{a_n^{-1}} p_{n-1} \xrightarrow{a_{n-1}^{-1}} \dots \xrightarrow{a_2^{-1}} p_1 \xrightarrow{a_1^{-1}} p, \\ q &\xrightarrow{a_n^{-1}} p_{n-1} \xrightarrow{a_{n-1}^{-1}} \dots \xrightarrow{a_2^{-1}} p_1 \xrightarrow{a_1^{-1}} p' \end{aligned}$$

labeled by u^{-1} . Now $\mathcal{A}$ being deterministic implies that $p = p'$, whence δ_u is an injective partial transformation.

Finally, by the above reasoning, every path $p \xrightarrow{u} q$ yields a path $q \xrightarrow{u^{-1}} p$ in $\mathcal{A}$, whence $\delta_u \delta_{u^{-1}} \delta_u = \delta_u$ and $\delta_{u^{-1}} \delta_u \delta_{u^{-1}} = \delta_{u^{-1}}$. This shows that $(\delta_u)^{-1} = \delta_{u^{-1}} \in M(\mathcal{A})$ for all $u \in \tilde{A}^*$, so we conclude that $M(\mathcal{A})$ is closed under taking inverses, which finishes the proof. $\square$

As a final remark, we sometimes say that the underlying alphabet of an involutive automaton is A instead of $\tilde{A}$; it is implicit that the inverse letters also belong to the alphabet of the automaton.

5 Varieties and pseudovarieties

The classification of finite monoids greatly benefits from the study of certain classes of monoids known as pseudovarieties. In 1976, Eilenberg [6] introduced the notion of a variety of languages and found a correspondence between pseudovarieties of monoids and varieties of rational languages, systematizing the classification of certain languages by analyzing their syntactic monoids. Moreover, we will resort to pseudovarieties in order to characterize finitely generated subgroups of free groups, which is the main goal of this dissertation. For more details, we refer the reader to [18], [19] and [1].

5.1 Pseudovarieties of monoids

We begin by defining the concept of a pseudovariety of finite monoids. Later on we shall present a generalization of some of these results to arbitrary algebraic structures.

We call a nonempty collection $\mathcal{C}$ of monoids a **class of monoids** if, for all $M \in \mathcal{C}$, every monoid isomorphic to M is also in $\mathcal{C}$. We say that $\mathcal{C}$ is closed under **arbitrary direct products** if

$$(\forall i \in I \ M_i \in \mathcal{C}) \Rightarrow \prod_{i \in I} M_i \in \mathcal{C}$$

for any indexing set I . If the above condition holds for finite indexing sets I , we say that $\mathcal{C}$ is closed under **finitary direct products**.

A **variety of monoids** is a class of monoids $\mathcal{V}$ closed under taking submonoids, homomorphic images and arbitrary direct products. This means that:

- (i) For all $M \in \mathcal{V}$, if $N \leq M$, then $N \in \mathcal{V}$.
- (ii) For all $M \in \mathcal{V}$, if $N = M\varphi$ for some monoid homomorphism $\varphi: M \rightarrow N$, then $N \in \mathcal{V}$.
- (iii) For all indexing sets I , if $M_i \in \mathcal{V}$ ($i \in I$), then $\prod_{i \in I} M_i \in \mathcal{V}$.

Equivalently, a variety of monoids $\mathcal{V}$ is a class closed under division (recall that M divides N if M is the homomorphic image of a submonoid of N) and arbitrary direct products. If a class of finite monoids $\mathbf{V}$ satisfies conditions (i), (ii) and is closed under finitary direct products, then we call $\mathbf{V}$ a **pseudovariety of finite monoids**.

In this monograph, we will only be concerned with pseudovarieties of finite monoids, so unless we are talking about free monoids, when we write “monoid”, we mean “finite monoid”.

Example 5.1. The class of all finite groups is a pseudovariety of monoids. Indeed, it is clear that the direct product of finitely many finite groups is a finite group. Next, let $\varphi: G \rightarrow M$ be a homomorphism from a group G to a monoid M . Since G is an inverse monoid, we already know that $G\varphi$ is an inverse monoid, and it follows from Lemma 3.3 that the unique idempotent of $G\varphi$ is 1, so $G\varphi$ is a group. Finally, regarding the closure under taking submonoids, let N be a submonoid of a group G . By Lemma 2.7, for every $x \in N$, there exists $n \geq 1$ such that $x^n \in N$ is an idempotent. Consequently, $x^n = 1$ and $x^{-1} = x^{n-1} \in N$, so N is a subgroup of G . However, the class of all finite inverse monoids is not a pseudovariety of monoids as a submonoid of an inverse monoid may not be inverse.

Given a set $\mathcal{C}$ of monoids, we may ask what is the smallest pseudovariety containing $\mathcal{C}$. As the intersection of all pseudovarieties containing $\mathcal{C}$ is again a pseudovariety containing $\mathcal{C}$, we have just found the answer to our question. We designate such pseudovariety by $V(\mathcal{C})$, and we can construct it in a more concrete way.

Proposition 5.2. *Let $\mathcal{C}$ be a set of monoids, and denote by $S(\mathcal{C})$, $H(\mathcal{C})$ and $P(\mathcal{C})$ the set of submonoids, homomorphic images and finitary direct products of elements in $\mathcal{C}$, respectively. Then $V(\mathcal{C}) = H(S(P(\mathcal{C})))$.*

Proof. To simplify the notation, let us write $HSP(\mathcal{C})$ instead of $H(S(P(\mathcal{C})))$. It follows from the definitions that $HSP(\mathcal{C}) \subseteq V(\mathcal{C})$, so it only remains to prove that $HSP(\mathcal{C})$ is a pseudovariety. The closure under homomorphic images follows from $H(HSP(\mathcal{C})) = HSP(\mathcal{C})$. Now, given $M \in HSP(\mathcal{C})$, there exist $M_1, \dots, M_n \in \mathcal{C}$, a submonoid $M' \leq M_1 \times \dots \times M_n$ and a surjective homomorphism $\varphi: M' \rightarrow M$. Consequently, every $N \leq M$ is the homomorphic image of $N\varphi^{-1} \leq M' \leq M_1 \times \dots \times M_n$, and thus $HSP(\mathcal{C})$ is closed under submonoids. Finally, given $M_1, \dots, M_n \in HSP(\mathcal{C})$ and $1 \leq i \leq n$, there exist monoids $M_{i1}, \dots, M_{im_i} \in \mathcal{C}$; submonoids $M'_i \leq M_{i1} \times \dots \times M_{im_i}$; and surjective homomorphisms $\varphi_i: M'_i \rightarrow M_i$. Since we can consider

$$M'_1 \times \dots \times M'_n \leq \prod_{i=1}^n \prod_{j=1}^{m_i} M_{ij}$$

and the φ_i induce a surjective homomorphism from $M'_1 \times \dots \times M'_n$ to $M_1 \times \dots \times M_n$, we conclude that $HSP(\mathcal{C})$ is closed under finitary direct products, which ends the proof. $\square$

We say that a pseudovariety $\mathbf{V}$ is **finitely generated** if there exists a finite set of monoids $\mathcal{C}$ such that $\mathbf{V} = V(\mathcal{C})$. In that case, $\mathbf{V}$ is actually generated by a single monoid. Indeed, if $\mathcal{C} = \{M_1, \dots, M_n\}$ generates $\mathbf{V}$, then so does $M_1 \times \dots \times M_n$, because the canonical projections $\pi_i: M_1 \times \dots \times M_n \rightarrow M_i$ are surjective homomorphisms.

There is yet another way to describe pseudovarieties of monoids, according to the “equations” they satisfy. Given a set of variables A , a **monoid identity** on A is an element $(u, v) \in A^* \times A^*$, which we usually indicate by a formal equality $u = v$. We say that a monoid M **satisfies the identity** $u = v$ if $u\varphi = v\varphi$ for every monoid homomorphism $\varphi: A^* \rightarrow M$. In that case, we write $M \models u = v$. Informally, this means that we obtain a true equality when we replace the variables of A by arbitrary elements of M in the identity $u = v$. If $(u_n = v_n)_n$ is a sequence of monoid identities, we say that M **ultimately satisfies** that sequence of identities if there exists some $p \geq 1$ such that $M \models u_n = v_n$ for every $n \geq p$. If $\mathcal{C}$ is the class of monoids ultimately satisfying the sequence of identities $(u_n = v_n)_n$, we say that $\mathcal{C}$ is **ultimately defined** by $(u_n = v_n)_n$.

Before we prove a major result due to Eilenberg and Schützenberger [7], which establishes a connection between pseudovarieties of monoids and identities, we need a preliminary lemma.

Lemma 5.3. *Let A be a finite alphabet. If $\sigma \subseteq A^* \times A^*$ is a congruence such that A^*/σ is finite, then there exists a finite relation $R \subseteq A^* \times A^*$ satisfying $\sigma = R^\#$.*

Proof. Since $A^*/\sigma = \{u\sigma \mid u \in A^*\}$ is finite, there exists $n \geq 1$ such that every congruence class has a representative $u \in A^*$ satisfying $|u| < n$. Consider then the finite relation

$$R = \{(u, v) \in \sigma : |u| \leq n, |v| < n\} \subset A^* \times A^*.$$

Since $R \subseteq \sigma$, it is clear that $R^\sharp \subseteq \sigma$, so we focus on the opposite inclusion.

First, we prove by induction on word length that, given $v \in A^*$, there exists $v' \in A^*$ such that $(v, v') \in R^\sharp$ and $|v'| < n$. If $|v| < n$, then we can take $v' = v$, so assume that $|v| = m \geq n$ and that the property holds for words with length less than m . We can write $v = wa$ for some $a \in A$ and $w \in A^{m-1}$. It follows from the induction hypothesis that there exists $w' \in A^*$ satisfying $|w'| < n$ and $(w, w') \in R^\sharp$. Now $R^\sharp$ being a congruence implies $(v, w'a) \in R^\sharp$, and since $|w'a| \leq n$, we have $(w'a, v') \in R$ for some $v' \in A^*$ satisfying $|v'| < n$. By transitivity of $R^\sharp$, we get $(v, v') \in R^\sharp$, which establishes the induction step.

By the above, given $(u, v) \in \sigma$, there exist $u', v' \in A^*$ such that $(u, u'), (v, v') \in R^\sharp$ and $|u'|, |v'| < n$. By transitivity of σ , we get $(u', v') \in \sigma$, which implies $(u', v') \in R \subseteq R^\sharp$. Hence $u R^\sharp u' R^\sharp v' R^\sharp v$, and the claim follows. $\square$

We are ready to state the big theorem.

Theorem 5.4 (Eilenberg and Schützenberger). *A class of finite monoids $\mathbf{V}$ is a pseudovariety of monoids if and only if $\mathbf{V}$ is ultimately defined by a sequence of identities $(u_n = v_n)_n$.*

Proof. Suppose first that $\mathbf{V}$ is a pseudovariety of monoids. Let $A = \{a_n \mid n \geq 1\}$ be a set of variables and let $A_k = \{a_1, \dots, a_k\}$ for all $k \geq 1$. Since all monoids in $\mathbf{V}$ are finite, we can consider a sequence of monoids $M_1, M_2, \dots$ such that every monoid in $\mathbf{V}$ is isomorphic to some M_i . Let $N_k = M_1 \times \dots \times M_k$. As A_k and N_k are finite, there is only a finite number $m_k \geq 1$ of homomorphisms from A_k^* to N_k . Hence, we can let $\varphi_k : A_k^* \rightarrow N_k^{m_k}$ be a homomorphism reproducing each of these m_k homomorphisms at each component. Since $(A_k^* / \ker \varphi_k) \cong \text{im } \varphi_k \leq N_k^{m_k}$ is finite, we deduce from the previous lemma that $\ker \varphi_k = R_k^\sharp$ for some finite $R_k \subset A_k^* \times A_k^*$. Now we can view $R = \cup_{k \geq 1} R_k$ as a sequence of identities over A , and our goal is to show that $\mathbf{V}$ is ultimately defined by R .

Given $M \in \mathbf{V}$, there exists some $\ell \geq 1$ such that $M \cong M_\ell$, and thus M_ℓ divides N_k for all $k \geq \ell$. As N_k satisfies the identities of R_k , it follows that M satisfies the identities of R_k for all $k \geq \ell$, which means that M ultimately satisfies R . Conversely, suppose that M is a monoid which satisfies the identities R_k , for all $k \geq \ell$, where ℓ is some positive integer. Let $p = \max\{\ell, |M|\}$ and take a surjective homomorphism $\phi : A_p^* \rightarrow M$. Since $M \models R_p$, we have $R_p \subseteq \ker \phi$, and hence $\ker \varphi_p = R_p^\sharp \subseteq \ker \phi$. As a result, we have a surjective homomorphism from $\text{im } \varphi_p$ to M , so it suffices to show that $\text{im } \varphi_p \in \mathbf{V}$. Observe that, by assumption, $M_k \in \mathbf{V}$ for all $k \geq 1$; therefore, $N_k = M_1 \times \dots \times M_k$ and $N_k^{m_k}$ also belong to $\mathbf{V}$. Now $\text{im } \varphi_p \leq N_k^{m_k}$ implies $\text{im } \varphi_p \in \mathbf{V}$, as we wanted.

On the other hand, let $\mathbf{V}$ be ultimately defined by the sequence of identities $(u_n = v_n)_n$ over a set of variables A . Given $M \in \mathbf{V}$ and $N \leq M$, every identity satisfied by M is also satisfied by N , whence $N \in \mathbf{V}$. Moreover, given $M_1, \dots, M_k \in \mathbf{V}$ and a monoid homomorphism $\varphi : A^* \rightarrow M_1 \times \dots \times M_k$, by composing φ with the canonical projections we

get homomorphisms $\varphi_i: A^* \rightarrow M_i$. Hence, for each $1 \leq i \leq k$, there exists $m_i \geq 1$ such that $u_n \varphi_i = v_n \varphi_i$ for all $n \geq m_i$. Taking $p = \max\{m_i \mid 1 \leq i \leq k\}$, it follows that $u_n \varphi = v_n \varphi$ for all $n \geq p$. Thus, $M_1 \times \cdots \times M_k$ satisfies $(u_n = v_n)_n$ for all $n \geq p$, and $\mathbf{V}$ is closed under finitary direct products. To finish, let $\eta: M \rightarrow M'$ be a surjective homomorphism and let $\psi: A^* \rightarrow M'$ be a homomorphism. In view of Proposition 2.6, there exists a homomorphism $\phi: A^* \rightarrow M$ satisfying $\phi\eta = \psi$. Taking $p \geq 1$ such that $u_n \phi = v_n \phi$ for all $n \geq p$, we get $u_n \psi = v_n \psi$ for all $n \geq p$, thereby showing that M' ultimately satisfies $(u_n = v_n)_n$. Therefore, $\mathbf{V}$ is closed under homomorphic images, which finishes the proof. $\square$

We now introduce a notation that will be useful to simplify the writing of pseudovariety identities. First, observe that in a finite monoid there exists $k \geq 1$ such that $x^k \in E(M)$ for all $x \in M$; we call k an **exponent** of M . If $\bar{n}$ denotes the least common multiple of the numbers $1, 2, \dots, n$, it follows easily that $n \geq k$ implies $x^{\bar{n}} \in E(M)$ for all $x \in M$. Consequently, $x^{\bar{n}}$ is the unique idempotent power of x whenever $n \geq k$. Following a convention of Schützenberger, in a pseudovariety defined by the identities $(u_n = v_n)_n$, we agree to replace every occurrence of $\bar{n}$ by the symbol ω . The usefulness of this notation will become evident in the succeeding section. If Σ is a set of monoid identities (which may feature the ω symbol), we denote by $\llbracket \Sigma \rrbracket$ the pseudovariety of monoids satisfying all the identities in Σ .

5.1.1 Examples

This subsection is entirely devoted to presenting important examples of pseudovarieties of monoids.

In Example 5.1, we saw that the class of all finite groups constitutes a pseudovariety of monoids, which we denote by $\mathbf{G}$. We claim that $\mathbf{G} = \llbracket x^\omega = 1 \rrbracket$. Indeed, on the one hand, every idempotent power of an element in a group must be the identity. On the other hand, if M is a finite monoid that satisfies the identity $x^\omega = 1$, then 1 is its unique idempotent and $x^{\omega-1}$ is the inverse of x . Consequently, M is a group, and the claim follows.

It is easy to see that the class of all finite commutative monoids constitutes a pseudovariety, which we denote by $\mathbf{Com}$, and clearly $\mathbf{Com} = \llbracket xy = yx \rrbracket$. Similarly, the pseudovariety $\mathbf{Sl}$ of all finite semilattices (commutative monoids where each element is an idempotent) is given by $\mathbf{Sl} = \llbracket x^2 = x, xy = yx \rrbracket$.

An important class of monoids is the class of aperiodic monoids. A monoid M is called **aperiodic** if, for every $x \in M$, there exists $n \geq 1$ such that $x^{n+1} = x^n$. Next we provide alternative characterizations of aperiodic monoids. In particular, we verify that this condition is reflected in the structure of $\mathcal{H}$ -classes. If $\mathcal{K}$ is one of Green's relations, we say that a monoid M is $\mathcal{K}$ -**trivial** if $\mathcal{K}$ is the identity relation on M .

Proposition 5.5. *Let M be a finite monoid. The following conditions are equivalent:*

- (i) M is aperiodic.
- (ii) There exists $n \geq 1$ such that $x^{n+1} = x^n$ for all $x \in M$.
- (iii) All subgroups of M are trivial.

(iv) M is $\mathcal{H}$ -trivial.

Proof. (i) $\Rightarrow$ (ii): Given $x \in M$, there exists $n_x \geq 1$ such that $x^{n_x+1} = x^{n_x}$. Taking $n = \max\{n_x \mid x \in X\}$, it follows that $x^{n+1} = x^n$ for every $x \in M$.

(ii) $\Rightarrow$ (iii): If K is a subgroup of M with identity e and $x \in K$, then $x^{n+1} = x^n \Rightarrow x = e$, so $K = \{e\}$.

(iii) $\Rightarrow$ (iv): Let $x, y \in M$ be such that $x\mathcal{H}y$, and take $a, b \in M$ satisfying $x = ya$ and $y = bx$. Then $x = bxa = b^n x a^n$ for every $n \geq 1$. Since M is finite, we know that there exist $m, p \geq 1$ such that $b^{m+p} = b^m$, and we may assume that m is as small as possible. After fixing m , we can fix p to be the smallest positive integer satisfying $b^m = b^{m+p}$. It follows that $\{b, \dots, b^{m+p-1}\}$ is the semigroup of all powers of b , and it is easy to check that $\{b^m, \dots, b^{m+p-1}\}$ is a subgroup of M . Condition (iii) then implies $p = 1$, and thus $b^{m+1} = b^m$. It follows that

$$x = b^m x a^m = b^{m+1} x a^m = b x = y,$$

so (iv) holds.

(iv) $\Rightarrow$ (i): Given $x \in M$, there exist $m, p \geq 1$ such that $\{x^m, \dots, x^{m+p-1}\}$ is a subgroup of M . Now $x^m \mathcal{H} x^{m+p-1}$ implies $x^m = x^{m+p-1}$, so $p = 1$. We conclude that $x^{m+1} = x^m$ and (i) holds. $\square$

In view of the preceding proposition, we deduce that the class $\mathbf{A}$ of all finite aperiodic monoids is the pseudovariety ultimately defined by the sequence of identities $(x^{n+1} = x^n)_n$, for $n \geq 1$. Consequently, $\mathbf{A} = \llbracket x^{\omega+1} = x^\omega \rrbracket$.

An important class of monoids is the class of completely regular monoids. A monoid M is called **completely regular** if it is a union of groups, that is, if every element belongs to some subgroup of M . Again, we present some equivalent conditions.

Proposition 5.6. *Let M be a finite monoid. The following conditions are equivalent:*

- (i) M is completely regular.
- (ii) For every $x \in M$, there exists $n \geq 2$ such that $x^n = x$.
- (iii) There exists $n \geq 2$ such that $x^n = x$ for all $x \in M$.

Proof. (i) $\Rightarrow$ (ii): Let K be a subgroup of M containing x . Then there exists $n \geq 1$ such that x^n is the idempotent (hence, the identity) of K , whence $x^{n+1} = x$.

(ii) $\Rightarrow$ (iii): Given $x \in M$, take $n_x \geq 1$ such that $x^{n_x+1} = x$. We prove that $x^{n_x m+1} = x$ for all $m \geq 1$ by induction on m , the case $m = 1$ being trivial. Suppose that $m \geq 2$ and $x^{n_x(m-1)+1} = x$. Then

$$x^{n_x m+1} = x^{n_x(m-1)+1} x^{n_x} = x^{n_x+1} = x,$$

which establishes the inductive step. Consequently, if $n \geq 1$ is, for instance, the least common multiple of the n_x ($x \in M$), we conclude that $x^{n+1} = x$ for all $x \in M$.

(iii) $\Rightarrow$ (i): If $x \in M$ and $n \geq 2$ are such that $x^n = x$, then the powers of x form a subgroup of M , whence we conclude that M is completely regular. $\square$

In view of the preceding proposition, we deduce that the class **CR** of all finite completely regular monoids is the pseudovariety of monoids ultimately defined by the sequence of identities $(x^{n+1} = x)_n$, for $n \geq 1$. Consequently, $\mathbf{CR} = \llbracket x^{\omega+1} = x \rrbracket$.

Finally, let $\mathcal{K}$ denote one of Green's relations $\mathcal{R}$, $\mathcal{L}$ or $\mathcal{J}$. Then the class **K** of all finite $\mathcal{K}$ -trivial monoids constitutes a pseudovariety of monoids. According to [19, Proposition 4.4], these pseudovarieties are given by

$$\begin{aligned}\mathbf{R} &= \llbracket (xy)^\omega x = (xy)^\omega \rrbracket \\ \mathbf{L} &= \llbracket y(xy)^\omega = (xy)^\omega \rrbracket \\ \mathbf{J} &= \llbracket (xy)^\omega = (yx)^\omega, x^{\omega+1} = x^\omega \rrbracket.\end{aligned}$$

5.2 Varieties of languages

We now establish a bijection between pseudovarieties of monoids and certain classes of languages. This correspondence gives a way of classifying rational languages by inspecting their syntactic monoids.

Given an alphabet A , a language $L \subseteq A^*$ and a word $u \in A^*$, we define the **left** (respectively, **right**) **quotient of L by u** as

$$\begin{aligned}u^{-1}L &= \{v \in A^* \mid uv \in L\} \\ Lu^{-1} &= \{v \in A^* \mid vu \in L\}.\end{aligned}$$

We say that a function $\mathcal{C}$ is a **class of rational languages** if $\mathcal{C}$ associates to every finite alphabet A a set of rational A -languages, which we denote by $A^*\mathcal{C}$. A **variety of rational languages** is a class of rational languages $\mathcal{V}$ such that the following conditions hold for all finite alphabets A and B :

- (i) If $K, L \in A^*\mathcal{V}$, then $K \cup L, A^* \setminus K \in A^*\mathcal{V}$.
- (ii) If $L \in A^*\mathcal{V}$ and $u \in A^*$, then $u^{-1}L, Lu^{-1} \in A^*\mathcal{V}$.
- (iii) If $L \in A^*\mathcal{V}$ and $\varphi: B^* \rightarrow A^*$ is a homomorphism, then $L\varphi^{-1} \in B^*\mathcal{V}$.

Note that, by induction, condition (i) implies that $L_1 \cup \dots \cup L_n \in A^*\mathcal{V}$ for all $n \geq 1$ and $L_1, \dots, L_n \in A^*\mathcal{V}$; and condition (ii) is equivalent to requiring $a^{-1}L, La^{-1} \in A^*\mathcal{V}$ for all $a \in A$. Moreover, it follows from (i) that $A^*\mathcal{V}$ is closed under intersection.

The following result is useful for establishing the varietal connection between monoids and languages.

Lemma 5.7. *Let A be an alphabet, $L \subseteq A^*$ and $u \in A^*$. If a monoid M recognizes L , then M also recognizes $u^{-1}L$ and Lu^{-1} .*

Proof. Let $\varphi: A^* \rightarrow M$ be a homomorphism and let $P \subseteq M$ be such that $L = P\varphi^{-1}$. Writing $Q = \{x \in M \mid (u\varphi)x \in P\}$, we have

$$v \in Q\varphi^{-1} \Leftrightarrow (u\varphi)(v\varphi) \in P \Leftrightarrow (uv)\varphi \in P \Leftrightarrow uv \in L.$$

Therefore, $Q\varphi^{-1} = u^{-1}L$ and M recognizes $u^{-1}L$. We prove that M recognizes Lu^{-1} in a similar way, by considering the set $Q' = \{x \in M \mid x(u\varphi) \in P\}$ instead. $\square$

We also need the next couple of lemmas.

Lemma 5.8. *Let M be a monoid that divides a monoid N and recognizes a language $L \subseteq A^*$. Then N also recognizes L .*

Proof. Let $\varphi: A^* \rightarrow M$ be a homomorphism and let $P \subseteq M$ be such that $L = P\varphi^{-1}$. Consider also $N' \leq N$ and a surjective homomorphism $\eta: N' \rightarrow M$. Then there exists a homomorphism $\psi: A^* \rightarrow N'$ satisfying $\varphi = \psi\eta$, whence $L = (P\eta^{-1})\psi^{-1}$. If $\iota: N' \hookrightarrow N$ denotes the canonical inclusion, then $L = (P\eta^{-1})\iota^{-1}\psi^{-1}$. Therefore, N recognizes L via the homomorphism $\psi\iota$. $\square$

Lemma 5.9. *Let $\mathbf{V}$ be a pseudovariety of monoids. Then $\mathbf{V}$ is generated by its syntactic monoids.*

Proof. Let $\mathcal{C}$ be the class of all syntactic monoids in $\mathbf{V}$, which are of the type $A^*/\sim_L$ for some language $L \subseteq A^*$. Using the notation of Proposition 5.2, our goal is to prove that $\mathbf{V} = \text{HSP}(\mathcal{C})$.

Given $M \in \mathbf{V}$, take a finite alphabet A and a surjective homomorphism $\varphi: A^* \rightarrow M$. Note that every language $x\varphi^{-1}$ is rational and $\text{Synt}(x\varphi^{-1}) \in \mathbf{V}$, as M recognizes $x\varphi^{-1}$; hence, by an observation made in Section 4.6, we deduce that $\text{Synt}(x\varphi^{-1})$ divides M . Consider now the finitary direct product

$$\Pi = \prod_{x \in M} \text{Synt}(x\varphi^{-1}),$$

and let $\theta: A^* \rightarrow \Pi$ reproduce the quotient morphisms $A^* \rightarrow \text{Synt}(x\varphi^{-1})$ in each component. If $u\theta = v\theta$, then $u \sim_{u\varphi\varphi^{-1}} v$ and $u \in u\varphi\varphi^{-1} \Rightarrow v \in u\varphi\varphi^{-1}$. This shows that $u\varphi = v\varphi$, and thus $\ker \theta \subseteq \ker \varphi$. Since $M \cong A^*/\ker \varphi$, we conclude that M is a homomorphic image of $A^*/\ker \theta \cong A^*\theta$. Finally, $A^*\theta \leq \Pi \in \mathbf{V}$ implies $M \in \mathbf{V}$, which concludes the proof. $\square$

The time has come for us to state and prove Eilenberg's correspondence theorem.

Theorem 5.10 (Eilenberg's correspondence). *There exists a bijection between pseudovarieties of monoids and varieties of rational languages.*

Proof. Let $\mathbf{V}$ be a pseudovariety of monoids and denote by $\mathcal{V}$ the class of all rational languages recognized by some monoid in $\mathbf{V}$. We claim that $\mathcal{V}$ is a variety of languages. Indeed, given $L \in \mathcal{V}$, it follows that $\text{Synt}(L)$ divides some monoid $M \in \mathbf{V}$; hence, $\mathcal{V}$ is the class of languages whose syntactic monoid lies in $\mathbf{V}$. By Proposition 4.18 and Lemma 5.7, it follows that $\mathcal{V}$ is closed under quotients and inverse homomorphic images. Suppose now

that $K, L \subseteq A^*$ are recognized by $M, N \in \mathbf{V}$, so that we have homomorphisms $\varphi: A^* \rightarrow M$, $\psi: A^* \rightarrow N$ and subsets $P \subseteq M$, $Q \subseteq N$ satisfying $K = P\varphi^{-1}$ and $L = Q\psi^{-1}$. Then $A^* \setminus K$ is recognized by M because $A^* \setminus K = (M \setminus P)\varphi^{-1}$. Finally, if $\theta: A^* \rightarrow M \times N$ is the homomorphism given by $u\theta = (u\varphi, u\psi)$, then

$$K \cup L = ((P \times N) \cup (M \times Q))\theta^{-1},$$

which shows that $M \times N$ recognizes $K \cup L$, and thus $K \cup L \in \mathcal{V}$.

Having shown that $\mathcal{V}$ is a variety of languages, let ξ be the function that maps every pseudovariety of monoids $\mathbf{V}$ to the variety $\mathbf{V}\xi$ of languages recognized by some monoid in $\mathbf{V}$. We wish to prove that ξ is a bijection, thereby establishing the theorem. Regarding injectivity, if $\mathbf{V}$ and $\mathbf{W}$ are distinct pseudovarieties of monoids, then they cannot have the same syntactic monoids, by Lemma 5.9. Therefore, we may assume that there exist an alphabet A and a rational language $L \subseteq A^*$ satisfying $\text{Synt}(L) \in \mathbf{V} \setminus \mathbf{W}$. But then we have $L \in \mathbf{V}\xi \setminus \mathbf{W}\xi$, so $\mathbf{V} \neq \mathbf{W}$, and ξ is an injection.

It remains to prove that ξ is surjective. Let $\mathcal{V}$ be a variety of languages and $\mathbf{V}$ the variety of monoids generated by the syntactic monoids of languages in $\mathcal{V}$. It is clear that $\mathcal{V} \subseteq \mathbf{V}\xi$, so we focus on the opposite inclusion. Given $L \in A^*(\mathbf{V}\xi)$, there exist $n \geq 1$, alphabets A_i and languages $L_i \in A_i^*\mathcal{V}$ ($1 \leq i \leq n$) such that $\text{Synt}(L)$ divides $\text{Synt}(L_1) \times \cdots \times \text{Synt}(L_n)$, by definition of $\mathbf{V}$. Since L is recognized by $\text{Synt}(L)$, it follows from Lemma 5.8 that there exists a homomorphism $\varphi: A^* \rightarrow \text{Synt}(L_1) \times \cdots \times \text{Synt}(L_n)$ satisfying $L = P\varphi^{-1}$, where $P = L\varphi$. Let $\pi_i: \text{Synt}(L_1) \times \cdots \times \text{Synt}(L_n) \rightarrow \text{Synt}(L_i)$ be the projection morphism on the i th component, and let $\eta_i: A_i^* \rightarrow \text{Synt}(L_i)$ be the quotient morphism ($1 \leq i \leq n$). Then there exist homomorphisms $\theta_i: A^* \rightarrow A_i^*$ such that $\varphi\pi_i = \theta_i\eta_i$. Since

$$L = P\varphi^{-1} = \bigcup_{x \in P} x\varphi^{-1}$$

and $A^*\mathcal{V}$ is closed under union, it suffices to show that $x\varphi^{-1} \in A^*\mathcal{V}$ for each $x \in P$. If $x = (x_1, \dots, x_n)$, then

$$\begin{aligned} u \in x\varphi^{-1} &\Leftrightarrow u\varphi = (x_1, \dots, x_n) \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} \quad (u\varphi\pi_i = x_i) \\ &\Leftrightarrow u \in \bigcap_{i=1}^n x_i(\varphi\pi_i)^{-1} \\ &\Leftrightarrow u \in \bigcap_{i=1}^n x_i(\theta_i\eta_i)^{-1} \\ &\Leftrightarrow u \in \bigcap_{i=1}^n (x_i\eta_i^{-1})\theta_i^{-1}. \end{aligned}$$

Therefore, it is enough to prove that $x_i\eta_i^{-1} \in A_i^*\mathcal{V}$ for all $1 \leq i \leq n$, as $A_i^*\mathcal{V}$ is closed under intersection.

Given $w \in x_i \eta_i^{-1}$, we define a relation on A_i^* by

$$R(w) = \{(u, v) \in A_i^* \times A_i^* \mid u w v \in L_i\}.$$

Then $w \eta_i = z \eta_i \Leftrightarrow w \sim_{L_i} z \Leftrightarrow R(w) = R(z)$. Since $u w v \in L_i \Leftrightarrow w \in u^{-1} L_i v^{-1}$, it follows that $z \in x_i \eta_i^{-1}$ if and only if

$$(\forall (u, v) \in R(w) \ z \in u^{-1} L_i v^{-1}) \wedge (\forall (u, v) \in (A_i^* \times A_i^*) \setminus R(w) \ z \notin u^{-1} L_i v^{-1}),$$

which entails

$$x_i \eta_i^{-1} = \left(\bigcap_{(u,v) \in R(w)} u^{-1} L_i v^{-1} \right) \setminus \left(\bigcap_{(u,v) \notin R(w)} u^{-1} L_i v^{-1} \right).$$

In view of Lemma 5.7, a monoid that recognizes L_i also recognizes all its left and right quotients, so there is only a finite amount of languages of the type $u^{-1} L_i v^{-1}$, since a finite monoid only possesses finitely many subsets. As $L_i \in A_i^* \mathcal{V}$ and $A_i^* \mathcal{V}$ is closed under the boolean operators, left quotients and right quotients, it follows that $x_i \eta_i^{-1} \in A_i^* \mathcal{V}$. We conclude that $L \in A^* \mathcal{V}$, whence $\mathcal{V} = \mathbf{V} \xi$, and ξ is a bijection as claimed. $\square$

This correspondence gives a way of classifying many rational languages. We give two examples, which actually appeared before Eilenberg's theorem and were his main source of inspiration. Let A be an alphabet and let $L \subseteq A^*$ be a language. We say that L is **star-free** if it can be obtained from finite languages using finitely many times the operators union, product and complement. It is called **piecewise testable** if it can be obtained from languages of the form

$$A^* a_1 A^* a_2 \cdots A^* a_n \ (n \geq 0, a_i \in A)$$

using finitely many times the operators union, intersection and complement. In 1965, Schützenberger [22] proved that star-free languages constitute the variety of languages corresponding to the pseudovariety of aperiodic monoids. And, in 1972, Imre Simon [24] proved that piecewise testable languages constitute the variety of languages corresponding to the pseudovariety of $\mathcal{J}$ -trivial monoids.

5.3 Pseudovarieties of algebras

In this monograph, besides dealing with pseudovarieties of monoids, we will also need the concepts of a pseudovariety of groups and a pseudovariety of inverse monoids. For that reason, we now generalize the notions presented in Section 5.1 to any kind of algebraic structure, and for that we need to present some basic definitions regarding universal algebra.

We begin by defining a **type** τ to be a pair $\tau = (\mathcal{O}, \alpha)$, where $\mathcal{O}$ is a set and $\alpha: \mathcal{O} \rightarrow \mathbb{Z}_0^+$ is a function. We call an element $f \in \mathcal{O}$ an **operation symbol** and $f\alpha \in \mathbb{Z}_0^+$ its **arity**. In the particular cases $f\alpha = 0$, $f\alpha = 1$ and $f\alpha = 2$, we say that f is a **constant symbol**, a **unary operation** and a **binary operation**, respectively. In the latter case, we usually represent f by

· and, instead of $(a, b) \cdot$, we write $a \cdot b$ or even ab . We denote by $\mathcal{O}_n$ the set of all operation symbols with arity n .

Given a type $\tau = (\mathcal{O}, \alpha)$, an **algebra of type τ** (or **τ -algebra**) is a pair $\mathcal{A} = (A, F^{\mathcal{A}})$, where A is a nonempty set and $F^{\mathcal{A}}$ is a function with domain $\mathcal{O}$ satisfying $fF^{\mathcal{A}}: A^n \rightarrow A$ for all $f \in \mathcal{O}_n$. We call A the **universe** of $\mathcal{A}$ and $fF^{\mathcal{A}}$ the **interpretation** of f in $\mathcal{A}$. To simplify the notation, we usually write $f^{\mathcal{A}}$ or f instead of $fF^{\mathcal{A}}$. Moreover, if the set of operation symbols is finite, say $\mathcal{O} = \{f_1, \dots, f_k\}$ for some $k \geq 0$, then we represent $\mathcal{A}$ by $(A, f_1, \dots, f_k)$; we call it an algebra with **signature** $(f_1\alpha, \dots, f_k\alpha)$. A **finite algebra** is an algebra $\mathcal{A}$ whose universe A is finite.

Example 5.11. A monoid can be regarded as an algebra $(M, \cdot, 1)$ with signature $(2, 0)$ that satisfies the equalities

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z \quad (*)$$

$$x \cdot 1 = x = 1 \cdot x \quad (**)$$

for all $x, y, z \in M$. Indeed, the operation on a monoid M is binary and associative, and we consider its unique identity element $1 \in M$ to be the interpretation of a constant symbol.

A group is an algebra $(G, \cdot, 1, {}^{-1})$ with signature $(2, 0, 1)$ that satisfies $(*)$, $(**)$ and also

$$x^{-1}x = 1 = xx^{-1}$$

for all $x \in G$.

An inverse monoid is an algebra $(M, \cdot, 1, {}^{-1})$ with signature $(2, 0, 1)$ that satisfies $(*)$, $(**)$ and also

$$\begin{aligned} xx^{-1}x &= x \\ (x^{-1})^{-1} &= x \\ xx^{-1}yy^{-1} &= yy^{-1}xx^{-1}. \end{aligned}$$

A proof can be found in [18, Theorem XII.1.1].

We proceed by defining the concepts of homomorphism, subalgebra and direct product of algebras, which generalize the definitions we already have for concrete algebraic structures, like monoids and groups. Let $\mathcal{A} = (A, F^{\mathcal{A}})$ and $\mathcal{B} = (B, F^{\mathcal{B}})$ be two algebras of type $\tau = (\mathcal{O}, \alpha)$. A **homomorphism** from $\mathcal{A}$ to $\mathcal{B}$ is a function $\varphi: A \rightarrow B$ such that

$$(a_1, \dots, a_n)f^{\mathcal{A}}\varphi = (a_1\varphi, \dots, a_n\varphi)f^{\mathcal{B}}$$

for all $n \geq 0$, $f \in \mathcal{O}_n$ and $a_1, \dots, a_n \in A$. We say that $\mathcal{A}$ is a **subalgebra** of $\mathcal{B}$ if $A \subseteq B$ and the inclusion mapping $A \hookrightarrow B$ is a homomorphism; in other words, for all $f \in \mathcal{O}_n$ and

$a_1, \dots, a_n \in A$, the equality

$$(a_1, \dots, a_n)f^A = (a_1, \dots, a_n)f^B$$

must hold. Finally, let I be a set and $\mathcal{A}_i = (A_i, F^{\mathcal{A}_i})$ be a family of algebras of type $\tau = (\mathcal{O}, \alpha)$ indexed by I . Given $a \in \prod_{i \in I} A_i$, we denote by ja the j -th component of a . We define the **direct product** of the family $(\mathcal{A}_i)_{i \in I}$ as the algebra $\mathcal{A} = \prod_{i \in I} \mathcal{A}_i$ whose universe is $\prod_{i \in I} A_i$ and where the equality

$$j((a_1, \dots, a_n)f^A) = (ja_1, \dots, ja_n)f^{\mathcal{A}_j}$$

holds for all $n \geq 0$, $f \in \mathcal{O}_n$, $j \in I$ and $a_1, \dots, a_n \in \prod_{i \in I} A_i$.

As in the case of monoids, a **class of τ -algebras** is a collection of τ -algebras that is closed under taking isomorphisms. We then define a **variety of τ -algebras** to be a class of τ -algebras closed under taking submonoids, homomorphic images and arbitrary direct products. And a **pseudovariety of finite τ -algebras** is a class of finite τ -algebras closed under taking submonoids, homomorphic images and finitary direct products. Sometimes we just write “pseudovariety of algebras” instead of “pseudovariety of finite algebras”; it is implicit that the algebras are finite.

We proceed by defining what identities are in this more general setting. Let X be a set of variables such that $X \cup \mathcal{O}_0 \neq \emptyset$. Consider the alphabet Y which consists of $X \cup \mathcal{O}$ together with the symbols “(”, “)” and “,”. Let $T(X)$ be the intersection of all subsets $Z \subseteq Y^+$ that contain $X \cup \mathcal{O}_0$ and satisfy

$$\forall t_1, \dots, t_n \in Y^+, \forall f \in \mathcal{O}_n \quad ((t_1, \dots, t_n) \in Z \Rightarrow (t_1, \dots, t_n)f \in Z).$$

Intuitively, $T(X)$ is obtained from $X \cup \mathcal{O}_0$ by successively applying finite combinations of operation symbols. We call $T(X)$ the set of all **terms** of type τ on X , and we can form an algebra $\mathcal{T}(X)$ of type τ with universe $T(X)$ by interpreting each $f \in \mathcal{O}$ as a function

$$\begin{aligned} T(X)^n &\longrightarrow T(X) \\ (t_1, \dots, t_n) &\longmapsto (t_1, \dots, t_n)f. \end{aligned}$$

It is designated by the **algebra of terms** in X of type τ . An **identity** of type τ on X is a pair $(u, v) \in T(X) \times T(X)$, usually written as $u = v$. We say that an algebra $\mathcal{A} = (A, F^{\mathcal{A}})$ satisfies the identity $u = v$ if $u\varphi = v\varphi$ for every homomorphism $\varphi: T(X) \rightarrow A$.

As the authors remark in [7], Theorem 5.4 remains true for pseudovarieties of algebras. Hence, a class of τ -algebras is a pseudovariety of τ -algebras if and only if it is ultimately defined by a sequence of identities $(u_n = v_n)_n$.

Having followed this more general approach, we already know what is meant by a **pseudovariety of finite groups** and a **pseudovariety of finite inverse monoids** (recall Example 5.11). Observe that in either case we have a unique binary operation, which is associative; a unique constant symbol, which is the identity; a unique unary operation; and

the equalities $(xy)^{-1} = y^{-1}x^{-1}$ and $(x^{-1})^{-1} = x$ hold for all elements x, y . It follows that $T(X) = \tilde{X}^*$, so an identity is a pair $(u, v) \in \tilde{X}^* \times \tilde{X}^*$ (where $\tilde{X} = X \cup X^{-1}$). Moreover, note that a group satisfies the identity $u = v$ if and only if it satisfies the identity $uv^{-1} = 1$; hence, we can always write a group identity in the form $u = 1$. Therefore, $\mathbf{H}$ is a pseudovariety of (finite) groups if and only if $\mathbf{H}$ is the class of finite groups ultimately satisfying a certain sequence of identities $(u_n = 1)_n$.

We proceed by making some remarks. Let $\mathbf{H}$ be a pseudovariety of groups. Then $\mathbf{H}$ is a pseudovariety of monoids because a direct product of monoids which are groups is a direct product of groups; a submonoid of a finite group is a subgroup; the homomorphic image of a finite group under a monoid homomorphism is a group (as we proved in Example 5.1); and a monoid homomorphism between two groups is a group homomorphism. We designate by $\overline{\mathbf{H}}$ the class of finite monoids all of whose subgroups lie in the pseudovariety of groups $\mathbf{H}$.

Proposition 5.12. *If $\mathbf{H}$ is a pseudovariety of groups, then $\overline{\mathbf{H}}$ is a pseudovariety of monoids.*

Proof. Given $M \in \overline{\mathbf{H}}$ and $N \leq M$, a subgroup of N is also a subgroup of M , so it belongs to $\mathbf{H}$; hence, $\overline{\mathbf{H}}$ is closed under taking submonoids. Next, let $\varphi: M \rightarrow M'$ be a surjective monoid homomorphism, and take a subgroup K' of M' . Consider a subsemigroup S of M that is minimal with respect to the property of being mapped onto K' by φ . Our goal is to prove that S is a group, for then we get that K' is the homomorphic image of a subgroup $S \in \mathbf{H}$, which shows that $K' \in \mathbf{H}$ and $M' \in \overline{\mathbf{H}}$. Given $x \in S$, we have $x\varphi \in K'$, and thus

$$\begin{aligned} (xS)\varphi &= (x\varphi)(S\varphi) = (x\varphi)K' = K' \\ (Sx)\varphi &= (S\varphi)(x\varphi) = K'(x\varphi) = K'. \end{aligned}$$

Since xS and Sx are subsemigroups of S , it follows that $xS = S = Sx$. Now Lemma 2.12 entails that S is a group, whence $\overline{\mathbf{H}}$ is closed under homomorphic images. Finally, take $M_1, M_2 \in \overline{\mathbf{H}}$; let K be a subgroup of $M_1 \times M_2$; and denote by $\pi_i: M_1 \times M_2 \rightarrow M_i$ the projection homomorphisms ($i \in \{1, 2\}$). Then $K\pi_i$ is a subgroup of M_i and K is a subgroup of $K\pi_1 \times K\pi_2 \in \mathbf{H}$. Hence $M_1 \times M_2 \in \overline{\mathbf{H}}$, and we conclude that $\overline{\mathbf{H}}$ is a pseudovariety of monoids. $\square$

As the group $\mathcal{H}$ -classes of a monoid are precisely its maximal subgroups (Corollary 2.15) and pseudovarieties of groups are closed under taking subgroups, we can also say that $\overline{\mathbf{H}}$ is the pseudovariety of monoids all of whose group $\mathcal{H}$ -classes belong to $\mathbf{H}$.

Let $\mathbf{V}$ be a pseudovariety of monoids, and denote by Inv the class of all finite inverse monoids. It is clear that $\mathbf{V} \cap \text{Inv}$ is a pseudovariety of inverse monoids, and we have the following result.

Proposition 5.13. *Let $\mathbf{R}$, $\mathbf{L}$ and $\mathbf{J}$ denote the pseudovarieties of $\mathcal{R}$ -trivial, $\mathcal{L}$ -trivial and $\mathcal{J}$ -trivial inverse monoids, respectively. Let also $\mathbf{Sl}$ be the pseudovariety of inverse monoids which are semilattices. Then $\mathbf{R} = \mathbf{L} = \mathbf{J} = \mathbf{Sl}$.*

Proof. Given $M \in \mathbf{R}$ and $x \in M$, from $xx^{-1}x = x$ we deduce that $xx^{-1}\mathcal{R}x$. Since $\mathcal{R}$ is the identity relation, it follows that $x = xx^{-1} \in E(M)$, whence every element in M is an

idempotent and commutes with every other element. Thus M is a semilattice, and a similar argument can be used to prove that an $\mathcal{L}$ -trivial or $\mathcal{J}$ -trivial inverse monoid is a semilattice too. Conversely, every semilattice is $\mathcal{R}$ -trivial and $\mathcal{L}$ -trivial, in view of Proposition 3.1(iii), and hence it is also $\mathcal{J}$ -trivial because all Green's relations coincide in a commutative monoid. This finishes the proof. $\square$

6 Free groups and Stallings' construction

We have already discussed free monoids, and free groups are of no less importance. They make an appearance in various fields of mathematics, namely algebra, topology, geometry and logic. The algebraic study of this class of groups essentially began with the work of Jakob Nielsen [17] in the twenties. Some decades later, John Stallings [25] introduced a revolutionary method to study finitely generated subgroups of free groups. Since then, the so-called Stallings automata became one of the most powerful tools for that purpose: not only they provided simple and elegant proofs of useful facts, but also boosted the algorithmics of free groups. For more information, we refer the reader to [15], [11] and [2].

6.1 Basic properties of free groups

In this section, we define what free groups are and briefly discuss their automorphism groups. First, we introduce the concept of a rewriting system.

Let A be an alphabet. A **rewriting system** on A is a binary relation $\mathcal{R} \subseteq A^* \times A^*$. Informally, it provides a way of replacing factors within words according to certain rules. Given $u, v \in A^*$, we write $u \rightarrow_{\mathcal{R}} v$ if there exist $(r, s) \in \mathcal{R}$ and $x, y \in A^*$ such that $u = xry$ and $v = xsy$; we call it an **elementary reduction**. We write $u \rightarrow_{\mathcal{R}}^* v$ if $u = v$ or if there exists a sequence of elementary reductions

$$u = u_0 \rightarrow_{\mathcal{R}} u_1 \rightarrow_{\mathcal{R}} \cdots \rightarrow_{\mathcal{R}} u_n = v,$$

for some $n \geq 1$ and $u_i \in A^*$; we call this a **reduction**. When $\mathcal{R}$ is implicit, we often omit it and write $u \rightarrow v$ and $u \rightarrow^* v$ instead.

We say that a rewriting system $\mathcal{R}$ is **length-reducing** if $(r, s) \in \mathcal{R}$ implies $|s| < |r|$; **noetherian** if every sequence of elementary reductions

$$u_0 \rightarrow u_1 \rightarrow u_2 \rightarrow \cdots$$

comes to an end; and **confluent** if, whenever we have $u \rightarrow^* v$ and $u \rightarrow^* w$, there exists $z \in A^*$ such that $v \rightarrow^* z$ and $w \rightarrow^* z$, as we depict in the diagram below:

$$\begin{array}{ccc} u & \xrightarrow{*} & v \\ \downarrow * & & \downarrow * \\ w & \xrightarrow{*} & z \end{array}$$

A word $u \in A^*$ is called **irreducible** if there exists no word $v \neq u$ satisfying $u \rightarrow^* v$. The set of all irreducible words of A^* (with respect to $\mathcal{R}$) is denoted by $\text{Irr}(\mathcal{R})$. Observe that if $\mathcal{R}$ is noetherian, then for all $u \in A^*$ there exists $v \in \text{Irr}(\mathcal{R})$ such that $u \rightarrow^* v$, since we cannot reduce a word indefinitely. If, in addition, $\mathcal{R}$ is also confluent, such $v \in \text{Irr}(\mathcal{R})$ is unique and we denote it by $v = \bar{u}$.

We define a length-reducing rewriting system on $\tilde{A}^* = (A \cup A^{-1})^*$ by

$$\mathcal{R}_A = \{(aa^{-1}, 1) \mid a \in \tilde{A}\}.$$

Let us prove that $\mathcal{R}_A$ is also confluent. Let $u, v, w \in A^*$ be such that we have $u \longrightarrow^* v$ and $u \longrightarrow^* w$ via the elementary reductions

$$\begin{aligned} u &= u_{00} \longrightarrow u_{01} \longrightarrow \cdots \longrightarrow u_{0n} = v \\ u &= u_{00} \longrightarrow u_{10} \longrightarrow \cdots \longrightarrow u_{m0} = w, \end{aligned}$$

for some $m, n \geq 0$. Our goal is to prove that there exist $u_{ij} \in \tilde{A}^*$ ($1 \leq i \leq m, 1 \leq j \leq n$) such that we have

$$\begin{array}{ccccccc} u_{00} & \longrightarrow & u_{01} & \longrightarrow & \cdots & \longrightarrow & u_{0n} \\ \downarrow & & \downarrow & & & & \downarrow \\ u_{10} & \dashrightarrow & u_{11} & \dashrightarrow & \cdots & \dashrightarrow & u_{1n} \\ \downarrow & & \downarrow & & & & \downarrow \\ \vdots & & \vdots & & & & \vdots \\ \downarrow & & \downarrow & & & & \downarrow \\ u_{m0} & \dashrightarrow & u_{m1} & \dashrightarrow & \cdots & \dashrightarrow & u_{mn} \end{array}$$

where the dotted arrows represent elementary reductions or equality itself. For that, it is enough to establish “local” confluence in the cases we have

$$\begin{array}{ccc} u & \longrightarrow & v_1 w_1 \\ \downarrow & & \downarrow \\ v_2 w_2 & \dashrightarrow & ? \end{array}$$

where $v_i, w_i \in \tilde{A}^*$ ($i \in \{1, 2\}$) and $u = v_1 a_1 a_1^{-1} w_1 = v_2 a_2 a_2^{-1} w_2$ for some $a_1, a_2 \in \tilde{A}$. Now, if the occurrences of $a_1 a_1^{-1}$ and $a_2 a_2^{-1}$ are completely disjoint, then there exist elementary reductions taking $v_1 w_1$ and $v_2 w_2$ to the word obtained from u by deleting both factors $a_i a_i^{-1}$ ($i \in \{1, 2\}$); if they partially overlap, say

$$\begin{aligned} u &= v'_1 b_1 a_1 a_1^{-1} w_1 \\ &= v_2 a_2 a_2^{-1} b_2 w'_2, \end{aligned}$$

then $v_1 w_1 = v_2 a_2 w_1 = v_2 a_1^{-1} w_1 = v_2 w_2$ (the other case being analogous); if they completely overlap, then also $v_1 w_1 = v_2 w_2$. In either case, we can complete the little diagram above, thereby showing that the general diagram can be completed too, and thus $\mathcal{R}_A$ is a confluent

rewriting system. Being noetherian and confluent, we conclude that, given $u \in \tilde{A}^*$, there exists a unique $\bar{u} \in \text{Irr}(\mathcal{R}_A)$ such that $u \longrightarrow^* \bar{u}$. Note that these irreducible words, also called **reduced words**, are precisely the ones that do not have any factor of the form aa^{-1} ($a \in \tilde{A}$).

We define an equivalence relation $\tau_A \subseteq \tilde{A}^* \times \tilde{A}^*$ by

$$(u, v) \in \tau_A \Leftrightarrow \bar{u} = \bar{v},$$

which is actually a congruence on $\tilde{A}^*$; then we define $F_A = \{u\tau_A \mid u \in \tilde{A}^*\}$. For now, F_A is just a set, but the next lemma shows that we can endow it with a group structure.

Lemma 6.1. *Let A be an alphabet. Then F_A becomes a group with the binary operation $(u\tau_A)(v\tau_A) = (uv)\tau_A$, for all $u, v \in \tilde{A}^*$.*

Proof. We begin by confirming that the operation is well defined. Suppose that $u\tau_A u'$ and $v\tau_A v'$. Then $\bar{u} = \bar{u}'$ and $\bar{v} = \bar{v}'$, so we deduce that

$$\overline{uv} = \overline{u'v'} = \overline{u'v'} = \overline{u'v'}.$$

Hence, $(uv)\tau_A = (u'v')\tau_A$ and the operation is indeed well defined. Next, it is immediate that this operation is associative, and $(1\tau_A)(u\tau_A) = u\tau_A = (u\tau_A)(1\tau_A)$, so $1\tau_A$ is the identity element. Finally, denoting by u^{-1} the formal inverse of u , we have

$$(u\tau_A)(u^{-1}\tau_A) = (uu^{-1})\tau_A = 1\tau_A$$

because $\overline{uu^{-1}} = 1$. Out of symmetry, we conclude that $u^{-1}\tau_A$ is the inverse of $u\tau_A$, and the statement follows. $\square$

We are ready to prove that F_A has the right to be called the **free group** over A .

Theorem 6.2 (Universal property of F_A). *Let A be an alphabet and consider the mapping*

$$\begin{aligned} \theta: A &\longrightarrow F_A \\ a &\longmapsto a\tau_A. \end{aligned}$$

If $\varphi: A \rightarrow G$ is a function from A to a group G , then there exists a unique group homomorphism $\Phi: F_A \rightarrow G$ such that $\theta\Phi = \varphi$.

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & G \\ \theta \downarrow & \nearrow \Phi & \\ F_A & & \end{array}$$

Proof. As in Theorem 3.10, we can extend φ to a function $\tilde{A} \rightarrow G$ and then to monoid homomorphism $\phi: \tilde{A}^* \rightarrow G$, in view of the universal property of $\tilde{A}^*$. Since G is a group, we

have $\tau_A \subseteq \ker \phi$, so ϕ induces a homomorphism $\Phi: F_A \rightarrow G$. This homomorphism satisfies $\theta\Phi = \varphi$, and it is unique since $A\theta$ generates F_A as a group. $\square$

Note that if $|A| = 1$, then $F_A \cong \mathbb{Z}$ (the infinite cyclic group), and that is the only case where F_A is commutative. Once again, every group is isomorphic to a quotient of a free group and every group satisfying the universal property with respect to a set A is isomorphic to F_A . In particular, if we define an operation $(u, v) \mapsto \overline{uv}$ on $\text{Irr}(\mathcal{R}_A)$, it is easy to check that it becomes a group, and the map

$$\begin{aligned} F_A &\longrightarrow \text{Irr}(\mathcal{R}_A) \\ u\tau_A &\longmapsto \overline{u} \end{aligned}$$

is an isomorphism. For that reason, we usually identify F_A and $\text{Irr}(\mathcal{R}_A)$, writing $\overline{u}$ or even u instead of $u\tau_A$; in particular, we consider $\tilde{A} \subseteq F_A$. Given a subgroup $K \leq F_A$, we write $\overline{K}$ to designate the set of all reduced words in $\tilde{A}^*$ representing the elements of K . Finally, given $u \in F_A$, we define $|u| = |\overline{u}|$, and we say that u is a prefix (suffix, factor) of $v \in F_A$ if $\overline{u}$ is a prefix (suffix, factor) of $\overline{v}$ in $\tilde{A}^*$.

As in the case of monoids, given a group G and a nonempty subset $X \subseteq G$, the subgroup of G **generated by** X is the intersection of all subgroups of G containing X . We denote it by $\langle X \rangle$, and it is also given by

$$\langle X \rangle = \{x_1^{\varepsilon_1} x_2^{\varepsilon_2} \dots x_n^{\varepsilon_n} \mid n \geq 0; x_i \in X; \varepsilon_i \in \{-1, 1\}\}.$$

We say that X **generates** G if $G = \langle X \rangle$, and G is said to be **finitely generated** when X can be chosen to be finite. In this monograph, we will only be concerned with finitely generated free groups, that is, groups of the type F_A for some finite set A ; thus, the reader shall always assume A to be finite. A subset $B \subseteq F_A$ is called a **basis** of F_A if B generates F_A and every nonempty reduced word $b_1^{\varepsilon_1} \dots b_n^{\varepsilon_n}$ over the alphabet B (where $\varepsilon_i \in \{-1, 1\}$) represents a nonidentity element in F_A . Equivalently, B is a basis if every element in F_A can be uniquely represented by a reduced word over the alphabet $\tilde{B} = B \cup B^{-1}$. It is clear that A is a basis of F_A , and the universal property still holds if we replace A by any other basis of F_A . The next result shows that all basis of a free group have the same cardinality.

Proposition 6.3. *Let A be an alphabet. If B is a basis of F_A , then $|B| = |A|$.*

Proof. Let $\mathbb{Z}_2$ denote the cyclic group of order 2. In view of the universal property of F_A , every function from A to $\mathbb{Z}_2$ can be (uniquely) extended to a group homomorphism from F_A to $\mathbb{Z}_2$. Conversely, every homomorphism $F_A \rightarrow \mathbb{Z}_2$ is determined by its restriction to A . Hence, we conclude that there are exactly $2^{|A|}$ homomorphisms from F_A to $\mathbb{Z}_2$. Similarly, since B is a basis of F_A , every homomorphism $F_A \rightarrow \mathbb{Z}_2$ is determined by its restriction to B , so there are exactly $2^{|B|}$ homomorphisms from F_A to $\mathbb{Z}_2$. It follows that $|A| = |B|$, as claimed. $\square$

We call $|A|$ the **rank** of F_A . It is a consequence of the preceding proposition that two free groups are isomorphic if and only if their basis have the same size.

We now discuss the automorphism group of a free group, that is, the group $\text{Aut}(F_A)$ whose elements are the automorphisms of F_A , for some finite alphabet A . Observe that an automorphism in $\text{Aut}(F_A)$ maps any basis of F_A to another basis of F_A , and it is completely determined by the images of the elements of a basis. The next result exhibits a finite generating set of $\text{Aut}(F_A)$.

Theorem 6.4. *Let A be an alphabet with at least two elements and F_A a free group over A . The automorphisms of the form*

$$\begin{array}{ll} \alpha_a: F_A \longrightarrow F_A & \beta_{ab}: F_A \longrightarrow F_A \\ a \longmapsto a^{-1} & a \longmapsto ab \\ x \longmapsto x \quad (x \in A \setminus \{a\}) & x \longmapsto x \quad (x \in A \setminus \{a\}), \end{array}$$

with $a, b \in A$ distinct letters, generate $\text{Aut}(F_A)$ as a group.

A proof can be found in [15, Proposition 4.1], and it was Nielsen [17] who first presented a set of generators of $\text{Aut}(F_A)$ very similar to the one above. For that reason, we designate such automorphisms by **elementary Nielsen automorphisms**. Moreover, we call α_a an automorphism of **type 1** and $\beta_{ab}, \beta_{ab}^{-1}$ automorphisms of **type 2**. Note that β_{ab}^{-1} is the automorphism whose restriction to A consists of replacing a by ab^{-1} and fixing all other letters of A . Such a manageable generating set of $\text{Aut}(F_A)$ will be absolutely crucial to derive our most important results.

6.2 Decidability problems

To motivate Stallings' construction, we briefly discuss some classical algorithmic problems regarding a free group F_A .

It is clear that $u, v \in \tilde{A}^*$ represent the same element in F_A if and only if $\bar{u} = \bar{v}$. Moreover, the reduced form of a word $u \in \tilde{A}^*$ can be achieved by successively eliminating finitely many factors of the type aa^{-1} ($a \in \tilde{A}$) in u . Hence, we have an algorithm to decide whether or not two elements in $\tilde{A}^*$ represent the same element in F_A . Put differently, there exists a finite set of instructions such that, for any input of two words $u, v \in \tilde{A}^*$, the output is “yes” in case they represent the same element in F_A and “no” otherwise. This decision problem is known as the **word problem** and, due to the existence of such an algorithm, it is considered a **decidable** problem.

Example 6.5. Let $A = \{a, b, c\}$. Since $\overline{abb^{-1}c^2a^{-1}a} = ac^2 = \overline{a^2a^{-1}c^2}$, we conclude that the words $abb^{-1}c^2a^{-1}a$ and $a^2a^{-1}c^2$ represent the same element in F_A .

Another classical decidability problem is the **conjugacy problem**: given $g, h \in F_A$, is there an algorithm that allows us to determine whether or not there exists some $x \in F_A$ satisfying $g = xhx^{-1}$? When such an equality is satisfied, we say that h is a **conjugate** of g . Note that the conjugacy relation on a group is always an equivalence relation.

Given $u \in \tilde{A}^*$, we say that u is **cyclically reduced** if uu is a reduced word, which is equivalent to saying that u is reduced and there exist no $a \in \tilde{A}$ and $v \in \tilde{A}^*$ satisfying

$u = ava^{-1}$. It is easy to see that every $u \in \text{Irr}(\mathcal{R}_A)$ can be uniquely written as $u = vxv^{-1}$ for some $x, v \in \text{Irr}(\mathcal{R}_A)$ with v cyclically reduced. We call v the **cyclic core** of u and write $v = c(u)$. Observe that if a word uw is cyclically reduced, then wu is cyclically reduced too, and we call it a **cyclic conjugate** of uw . This designation makes sense because $\overline{uw} = \overline{w^{-1}(wu)w}$, so the group elements corresponding to uw and wu are conjugate in F_A . Since a cyclic conjugate of a cyclically reduced word $a_1 \dots a_n$ ($a_i \in \tilde{A}$) is of the type $a_i \dots a_n a_1 \dots a_{i-1}$ for some $1 \leq i \leq n$, it follows that a word admits only finitely many cyclic conjugates. Once again, the relation of cyclic conjugacy is an equivalence relation. The following proposition provides a way of solving the conjugacy problem in F_A .

Proposition 6.6. *Let A be an alphabet and $g, h \in F_A$. Then g and h are conjugate in F_A if and only if $c(\bar{g})$ and $c(\bar{h})$ are cyclic conjugates.*

Proof. First note that g (respectively, h) is a conjugate of $c(\bar{g})$ (respectively, $c(\bar{h})$) in F_A . And if $c(\bar{g})$ and $c(\bar{h})$ are cyclic conjugates, then they are also conjugate in F_A . Since the conjugacy relation is transitive, it follows that g and h are conjugate in F_A .

Suppose now that g and h are conjugate in F_A . We may assume that $g \neq 1$ (so $c(\bar{g}) \neq 1$), and it is enough to prove that $v = c(\overline{aga^{-1}})$ is a cyclic conjugate of $u = c(\bar{g})$ for any $a \in \tilde{A}$. Let $x \in \text{Irr}(\mathcal{R}_A)$ be such that $\bar{g} = xux^{-1}$. If $x \neq 1$, then

$$v = c(\overline{aga^{-1}}) = c(\overline{axu(ax)^{-1}}) = u$$

and the conclusion follows, so we may assume that $x = 1$. Also, we are only interested in the case where $aua^{-1} \notin \text{Irr}(\mathcal{R}_A)$, for otherwise we obtain $u = v$ once again. Suppose then that au is not reduced, the other case being similar. Then $u = a^{-1}w$ for some $w \in \text{Irr}(\mathcal{R}_A)$, and thus

$$\overline{aga^{-1}} = \overline{aa^{-1}wa^{-1}} = \overline{wa^{-1}}.$$

Now u cannot end in a because u is cyclically reduced, and w cannot begin with a since u is reduced. We conclude that $\overline{aga^{-1}} = \overline{wa^{-1}}$ is cyclically reduced, whence $v = \overline{wa^{-1}}$ is a cyclic conjugate of $u = a^{-1}w$. $\square$

Accordingly, in order to determine if $g, h \in F_A$ are conjugate, we first compute $\bar{g}$ and $\bar{h}$, and then we verify whether or not $c(\bar{g})$ is a cyclic permutation of $c(\bar{h})$.

Example 6.7. Let $A = \{a, b, c\}$, $g = bcb^{-1}b^2ab^{-1}cc^{-1}$ and $h = a^{-1}c^{-1}aa^{-1}bac^2a$. Then we have

$$\begin{aligned}\bar{g} &= \overline{bcb^{-1}b^2ab^{-1}cc^{-1}} = bcbab^{-1} = b(cba)b^{-1} \\ \bar{h} &= \overline{a^{-1}c^{-1}aa^{-1}bac^2a} = a^{-1}c^{-1}bac^2a = (ca)^{-1}(bac)ca,\end{aligned}$$

so $c(\bar{g}) = cba$ and $c(\bar{h}) = bac$. Since cba is a cyclic permutation of bac , we conclude that g and h are conjugate in F_A .

Thinking in terms of cyclically reduced words also allows us to conclude that every

nonidentity element in a free group has infinite order, making it a **torsion-free** group. Indeed, given $1 \neq g \in F_A$, we can write $\bar{g} = xux^{-1}$ for some $x \in \text{Irr}(\mathcal{R}_A)$ and $u = c(\bar{g}) \neq 1$, whence

$$\overline{g^n} = \overline{(xux^{-1})^n} = \overline{xu^n x^{-1}} = xu^n x^{-1}.$$

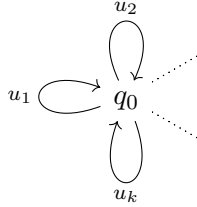
Consequently, if $n \geq 1$, then $\overline{g^n}$ is a nonempty reduced word, which shows that $g^n \neq 1$ in F_A , and g has infinite order. It is therefore trivial to decide whether or not an element has finite order in F_A .

However, there exists a decidability problem whose solution is not so obvious. Given a finitely generated subgroup $K \leq F_A$ and an element $g \in F_A$, is it decidable whether or not $g \in K$? This is known as the **generalized word problem**, and we shall address it in the upcoming section.

6.3 Stallings' construction

We are finally ready to establish the beautiful connection between finitely generated subgroups of free groups and finite automata.

Let F_A be a free group over a finite alphabet A . Let $K = \langle u_1, \dots, u_k \rangle \leq F_A$ be a finitely generated subgroup of F_A , where each generator u_i is seen as a (nonempty) reduced word in $\tilde{A}^*$. The notation $\leq_{f.g.}$ will be often used to indicate that a subgroup is finitely generated. We begin by constructing the so-called **flower automaton** $\mathcal{F}(u_1, \dots, u_k)$ of K by fixing a point q_0 , called the **basepoint** of the automaton, and gluing to it k “petals” labeled by each of the u_i , as well as the corresponding inverse edges, in order to obtain an involutive automaton over $\tilde{A}$:



We declare q_0 to be the unique initial state and the unique terminal state of the flower automaton. Our goal now is to turn it into an inverse automaton. For that purpose, whenever we encounter a pair of distinct edges $p \xrightarrow{a} q$ and $p \xrightarrow{a} r$ for some $a \in \tilde{A}$, we identify them, and we also identify the corresponding inverse edges (so q and r collapse into a single vertex if they are distinct). These identifications are known as **Stallings foldings**, and they are successively applied until we reach a deterministic automaton. The inverse automaton thus obtained is called a **Stallings automaton** of K .

Proposition 6.8. *Let F_A be a free group of finite rank and let $K \leq_{f.g.} F_A$. Then the language recognized by any Stallings automaton of K is the intersection of all languages $L \subseteq \tilde{A}^*$ containing $\overline{K}$ which are recognized by a finite inverse automaton with a basepoint.*

Proof. Let $K = \langle u_1, \dots, u_k \rangle$, where each u_i is seen as a (nonempty) reduced word over $\tilde{A}$, and denote by $\mathcal{S}(u_1, \dots, u_k)$ a Stallings automaton of K obtained from $\mathcal{F}(u_1, \dots, u_k)$.

Given $x \in \overline{K}$, we can write $x = \overline{u_{i_1}^{\varepsilon_1} \dots u_{i_m}^{\varepsilon_m}}$ for some $m \geq 1$; $1 \leq i_1, \dots, i_m \leq k$; and $\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}$. For all $1 \leq i \leq k$, we have $u_i \in L(\mathcal{F}(u_1, \dots, u_k))$, and since the flower automaton is an involutive automaton with a basepoint, it follows that

$$u_{i_1}^{\varepsilon_1} \dots u_{i_m}^{\varepsilon_m} \in L(\mathcal{F}(u_1, \dots, u_k)).$$

Now observe that any word labelling a successful path in an involutive automaton also labels a successful path in any automaton obtained from it through a sequence of foldings. In particular, we have $L(\mathcal{F}(u_1, \dots, u_k)) \subseteq L(\mathcal{S}(u_1, \dots, u_k))$. Moreover, if $\mathcal{A}$ is an inverse automaton over $\tilde{A}$ and $v, w, z \in \tilde{A}^*$ are such that $vw w^{-1}z \in L(\mathcal{A})$, then also $vz \in L(\mathcal{A})$, because ww^{-1} labels a loop at every state from where it can be read. It follows that

$$x = \overline{u_{i_1}^{\varepsilon_1} \dots u_{i_m}^{\varepsilon_m}} \in L(\mathcal{S}(u_1, \dots, u_k)),$$

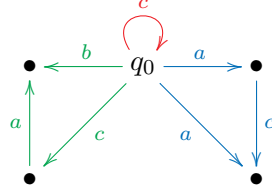
so $\mathcal{S}(u_1, \dots, u_k)$ is a finite inverse automaton with a basepoint whose language contains $\overline{K}$.

Now let $\mathcal{A}$ be a finite inverse automaton with a basepoint such that $L(\mathcal{A}) \supseteq \overline{K}$. We must show that $L(\mathcal{S}(u_1, \dots, u_k)) \subseteq L(\mathcal{A})$. Let Q and Q' be the states of $\mathcal{F}(u_1, \dots, u_k)$ and $\mathcal{A}$, respectively. Let also $q_0 \in Q$ and $q'_0 \in Q'$ be the basepoints. We begin by establishing a morphism from the flower automaton to $\mathcal{A}$. For each $1 \leq i \leq k$, let $u_i = a_{i1} \dots a_{im_i}$ be the factorization into letters of u_i . Given $q \neq q_0$, there exist a unique $1 \leq i \leq k$ and $n_i < m_i$ such that $w = a_{i1} \dots a_{in_i}$ labels a path $q_0 \xrightarrow{w} q$ in $\mathcal{F}(u_1, \dots, u_k)$. Since w is a prefix of $u_i \in L(\mathcal{A})$, there exists some $q' \in Q'$ such that we have a path $q'_0 \xrightarrow{w} q'$ in $\mathcal{A}$. Consider the function $\varphi: Q \rightarrow Q'$ given by $q_0\varphi = q'_0$ and $q\varphi = q'$ for $q \neq q_0$. It is easy to verify that φ induces an automaton morphism from $\mathcal{F}(u_1, \dots, u_k)$ to $\mathcal{A}$, and therefore $L(\mathcal{F}(u_1, \dots, u_k)) \subseteq L(\mathcal{A})$.

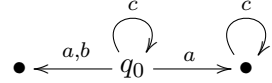
In order to show that this inclusion is preserved if we replace the flower automaton by the Stallings automaton, we consider a more general setting. Let $\mathcal{B}$ be an involutive automaton satisfying $L(\mathcal{B}) \subseteq L(\mathcal{A})$ and consider the automaton $\mathcal{B}'$ obtained by folding two edges $p \xrightarrow{a} q$ and $p \xrightarrow{a} r$ in $\mathcal{B}$. Every successful path in $\mathcal{B}'$ labeled by $v \in \tilde{A}^*$ can be lifted to a successful path in $\mathcal{B}$ labeled by a word w obtained from v by inserting factors $a^{-1}a$ if necessary. Then $w \in L(\mathcal{B}) \subseteq L(\mathcal{A})$ implies $v \in L(\mathcal{A})$ because $\mathcal{A}$ is inverse, whence $L(\mathcal{B}') \subseteq L(\mathcal{A})$. The desired conclusion now follows from the fact that $\mathcal{S}(u_1, \dots, u_k)$ is obtained from $\mathcal{F}(u_1, \dots, u_k)$ by successively applying one folding at a time. $\square$

In view of Propositions [4.20](#) and [4.24](#), this result assures that any two Stallings automata of K are isomorphic, so we can speak of *the* Stallings automaton of K , and we denote it by $\mathcal{S}(K)$. In other words, $\mathcal{S}(K)$ does not depend on the generating set of K nor on the order in which the foldings are made. However, $\mathcal{S}(K)$ depends on the basis A of the free group we are considering.

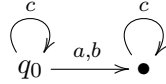
Example 6.9. Let $A = \{a, b, c\}$ and $K = \langle c, ba^{-1}c^{-1}, aca^{-1} \rangle \leq F_A$. Then the flower automaton $\mathcal{F}(c, ba^{-1}c^{-1}, aca^{-1})$ is depicted by



After folding the two blue edges labeled by a and the green and red edges labeled by c , we obtain the automaton



It remains to fold the edges labeled by a , so $\mathcal{S}(K)$ is given by



The argument presented in the last paragraph of the proof of Proposition 6.8 allows us to deduce that $\overline{L(\mathcal{F}(K))} = \overline{L(\mathcal{S}(K))}$, where $\mathcal{F}(K)$ is a flower automaton of $K \leq_{f.g.} F_A$. Since $\overline{L(\mathcal{F}(K))} \subseteq \overline{K}$ and $\overline{K} \subseteq \overline{L(\mathcal{S}(K))}$, we conclude that $\overline{L(\mathcal{S}(K))} = \overline{K}$. This is an important remark, so we will register it in a proposition.

Proposition 6.10. *Let F_A be a free group of finite rank and let $K \leq_{f.g.} F_A$. Then $\overline{L(\mathcal{S}(K))} = \overline{K}$.*

Consequently, the generalized word problem is decidable because, given $K \leq_{f.g.} F_A$ and $u \in \tilde{A}^*$, we have $\bar{u} \in K$ if and only if $\bar{u}$ labels a successful path in $\mathcal{S}(K)$. For instance, if K is the subgroup in Example 6.9 and $u = b^{-1}bcba^{-1}$, then $\bar{u} = cba^{-1} \in L(\mathcal{S}(K))$, so $\bar{u} \in K$. On the other hand, $ac \notin K$ because when we read ac from q_0 we end up at the rightmost vertex, which is not the terminal vertex.

From a computational point of view, the most efficient known algorithm to construct $\mathcal{S}(K)$ is due to Touikan [26].

6.4 Basic applications

In this section, we collect some basic applications of Stallings automata for studying finitely generated subgroups of free groups. Several others can be found in [11] and [2].

Stallings automata allow us to conclude that every finitely generated subgroup of a free group is itself free, a result known as Nielsen Theorem, for they provide a way to construct an explicit basis. Let A be a finite alphabet and let $K \leq_{f.g.} F_A$. Since $\mathcal{S}(K)$ is an inverse automaton with a basepoint q_0 , we write $\mathcal{S}(K) = (Q, A, E, q_0)$ instead of $\mathcal{S}(K) = (Q, \tilde{A}, E, q_0, q_0)$.

A **spanning tree** $T \subseteq E$ of $\mathcal{S}(K)$ is a subset consisting of $|Q| - 1$ positive edges and their inverses which are able to connect all the vertices in Q . Observe that a subset of E containing less than $|Q| - 1$ positive edges cannot possibly connect all the vertices. Hence, given a spanning tree T of $\mathcal{S}(K)$ and $q \in Q$, there exists a unique shortest path in T from q_0 to q : if there were two of them, we could erase an edge in T and still connect all the vertices in Q , which is not possible. We denote the label of this path by $g_q \in \tilde{A}^*$ and call it the **T -geodesic** from q_0 to q .

In general, a word is called a **geodesic** if it is the label of a shortest path between two vertices in $\mathcal{S}(K)$. Given $p, q \in Q$, we define the **geodesic distance** between p and q to be the length of a shortest path connecting these two vertices in $\mathcal{S}(K)$. The T -geodesics are the key to finding a basis of K .

Theorem 6.11. *Let $K \leq_{f.g.} F_A$ and let T be a spanning tree of $\mathcal{S}(K) = (Q, A, E, q_0)$. Then*

$$B = \{g_p a g_q^{-1} \mid (p, a, q) \in E^+ \setminus T\}$$

is a basis of K . In particular, K is itself a free group.

Proof. First observe that, for any $p, q \in Q$, we have $g_p, g_q \in \text{Irr}(\mathcal{R}_A)$, so $(p, a, q) \in E^+ \setminus T$ implies $g_p a g_q^{-1} \in \text{Irr}(\mathcal{R}_A)$. Since $B \subseteq L(\mathcal{S}(K))$, it follows from Proposition 6.10 that $B \subseteq \overline{K}$, and thus $B \subseteq K$ in the free group.

Now we verify that B generates K . Take $x \in K$ and let $x = a_1 \dots a_k$ be its factorization into letters, seen as a reduced word over $\tilde{A}$. By the same proposition, there exists a successful path

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \dots \xrightarrow{a_{k-1}} q_{k-1} \xrightarrow{a_k} q_k = q_0$$

in $\mathcal{S}(K)$. Given $1 \leq i \leq k$, either $(q_{i-1}, a_i, q_i) \in E^+ \setminus T$, which implies $g_{q_{i-1}} a_i g_{q_i}^{-1} \in B$; or $(q_i, a_i^{-1}, q_{i-1}) \in E^+ \setminus T$, which implies $g_{q_{i-1}} a_i g_{q_i}^{-1} \in B^{-1}$; or $(q_{i-1}, a_i, q_i) \in T$, which implies $g_{q_{i-1}} a_i g_{q_i}^{-1} = 1$. In any case, since $g_{q_0} = 1$, we get

$$x = \overline{(g_{q_0} a_1 g_{q_1}^{-1})(g_{q_1} a_2 g_{q_2}^{-1}) \dots (g_{q_{k-1}} a_k g_{q_0}^{-1})} \in \langle B \rangle$$

in the free group, and hence $K = \langle B \rangle$.

Finally, it remains to check that every nonempty reduced word $b_1 \dots b_k$ ($k \geq 1$) over the alphabet $\tilde{B}$ represents a nonidentity element in F_A . So take $b_i = g_{p_i} a_i g_{r_i}^{-1}$, for $1 \leq i \leq k$, satisfying $(p_i, a_i, r_i) \in E \setminus T$ and $b_i \neq b_{i-1}^{-1}$ ($1 < i \leq k$). Then we have

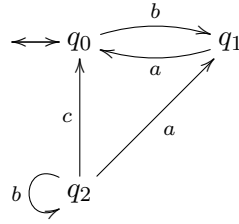
$$\overline{b_1 \dots b_k} = \overline{g_{p_1} a_1 g_{r_1}^{-1} g_{p_2} a_2 \dots a_{k-1} g_{r_{k-1}}^{-1} g_{p_k} a_k g_{r_k}^{-1}}$$

by definition of T and T -geodesics. Consequently, $\overline{b_1 \dots b_k}$ is a nonempty reduced word, whence B is a basis of K . $\square$

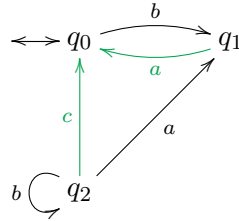
Note that every vertex in a flower automaton $\mathcal{F}(K)$ occurs in some successful path with reduced label, because the petals of $\mathcal{F}(K)$ are given by reduced words. This property still holds after the folding process, so every vertex in $\mathcal{S}(K)$ also appears in some successful path

labeled by a reduced word. Consequently, the basepoint q_0 is the only vertex of $\mathcal{S}(K)$ which may have outdegree 1; all the other vertices have outdegree at least 2. Therefore, in view of the previous proposition, we conclude that the class of Stallings automata is precisely the class of finite inverse automata with a basepoint where all the vertices which are not the basepoint have outdegree greater than 1. Indeed, we can take a spanning tree in an automaton of the latter kind and determine a basis of a subgroup K as in Proposition 6.11, so the automaton is precisely $\mathcal{S}(K)$.

Example 6.12. Let $A = \{a, b, c\}$ and let



be the Stallings automaton $\mathcal{S}(K)$ of a certain $K \leq_{f.g.} F_A$. Below we depict, in green, a spanning tree T of $\mathcal{S}(K)$:



According to Theorem 6.11, the set

$$B = \{g_{q_0} b g_{q_1}^{-1}, g_{q_2} a g_{q_1}^{-1}, g_{q_2} b g_{q_2}^{-1}\} = \{ba, c^{-1}a^2, c^{-1}bc\}$$

is a basis of K . Hence, $K \simeq F_A$ is a free group of rank 3.

We proceed with another application. Recall that, given a group G and $K \leq G$, the **index** of K in G is the number of right cosets Kg ($g \in G$) or, equivalently, the number of left cosets gK ($g \in G$), and we denote it by $[G : K]$. We write $\leq_{f.i.}$ to indicate that a subgroup is of finite index. Stallings automata can be used to determine the index of a finitely generated subgroup of a free group, as the next result shows.

Theorem 6.13. *Let $K \leq_{f.g.} F_A$ and $\mathcal{S}(K) = (Q, A, E, q_0)$. Then $[F_A : K] < \infty$ if and only if $\mathcal{S}(K)$ is a complete automaton. Moreover, if $K \leq_{f.i.} F_A$, then $[F_A : K] = |Q|$.*

Proof. First suppose that $K \leq_{f.g.} F_A$ is such that $\mathcal{S}(K)$ is not complete. Then there exist $p \in Q$ and $a \in \tilde{A}$ such that $(p, a, q) \notin E$ for all $q \in Q$. Consider a path $q_0 \xrightarrow{u} p$ in $\mathcal{S}(K)$ labeled by a reduced word $u \in \tilde{A}^*$. Our goal is to show that $Kua^m \neq Kua^n$ whenever $m - n > |u|$. Indeed, if $Kua^m = Kua^n$, then $w = ua^{m-n}u^{-1} \in K$, so $\bar{w} \in L(\mathcal{S}(K))$ by Proposition 6.10. The condition $m - n > |u|$ now implies that $\bar{w} = uaa^{m-n-1}u^{-1}$. It follows that $uaa^{m-n-1}u^{-1}$ labels a successful path in $L(\mathcal{S}(K))$, which is absurd because it is not

possible to read a from p , and hence we cannot read ua from q_0 either. Therefore, and in particular, the cosets $Kua^{k(|u|+1)}$ are all distinct for $k \geq 1$, which shows that the index of K in F_A is not finite.

Next, if $K \leq_{f.g.} F_A$ and $\mathcal{S}(K)$ is a complete automaton, let $q_0 \xrightarrow{g_q} q$ be a path of minimum length in $\mathcal{S}(K)$ for each $q \in Q$. Given a reduced word $w \in \tilde{A}^*$, there exists a path $q_0 \xrightarrow{w} q$ in $\mathcal{S}(K)$, for some $q \in Q$. This entails $wg_q^{-1} \in L(\mathcal{S}(K))$, so $wg_q^{-1} \in K$ (seen as a group element), and thus $w = wg_q^{-1}g_q \in Kg_q$. This shows that

$$F_A = \bigcup_{q \in Q} Kg_q,$$

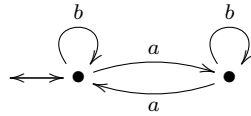
whence $[F_A : K] < \infty$ because Q is finite.

To finish, we show that if $K \leq_{f.i.} F_A$, then the cosets Kg_q ($q \in Q$) are all distinct. Indeed, if $p, q \in Q$ are such that $Kg_p = Kg_q$, then $g_pg_q^{-1} \in K$ and we get paths

$$q_0 \xrightarrow{\overline{g_pg_q^{-1}}} q_0 \quad \text{and} \quad q_0 \xrightarrow{\overline{g_pg_q^{-1}g_q}} q$$

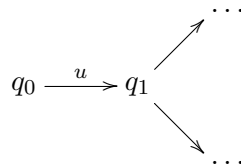
in $\mathcal{S}(K)$. But $\overline{g_pg_q^{-1}g_q} = \overline{g_p} = g_p$, as geodesics are reduced words, whence we get a path $q_0 \xrightarrow{g_p} q$ in $\mathcal{S}(K)$. This entails that $p = q$, and therefore $[F_A : K] = |Q|$, as we wanted. $\square$

Example 6.14. Let $A = \{a, b\}$ and consider the inverse automaton over $\tilde{A}$ given by



This is a finite inverse automaton with a basepoint without vertices of outdegree 1. Hence, it is the Stallings automaton of a certain subgroup $K \leq_{f.g.} F_A$. Since $\mathcal{S}(K)$ is complete, we deduce from the previous theorem that $[F_A : K] = 2$.

We now examine the conjugacy relation on subgroups. Recall that if G is a group and $H, K \leq G$, we say that H and K are **conjugate** subgroups if $H = gKg^{-1}$ for some $g \in G$. Given a nontrivial subgroup $K \leq_{f.g.} F_A$, we denote by $\Gamma(K)$ the underlying labeled graph of $\mathcal{S}(K)$. In view of the abstract characterization of Stallings automata, $\Gamma(K)$ is of the form



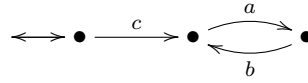
where q_1 is either q_0 or the closest vertex to q_0 having outdegree (strictly) greater than 2, in terms of the geodesic distance. We say that $q_0 \xrightarrow{u}$ is the **tail** of $\mathcal{S}(K)$ and the remaining graph is the **core** of $\mathcal{S}(K)$, which we denote by $\mathcal{C}(K)$. Note that the tail is empty if $q_1 = q_0$.

Consequently, given $K \leq_{f.g.} F_A$, its Stallings automaton $\mathcal{S}(K)$ is either equal to $\mathcal{C}(K)$ with a vertex chosen to be the basepoint; or it is obtained by “gluing” a tail $q_0 \xrightarrow{u}$ to $\mathcal{C}(K)$ and declaring q_0 to be the basepoint.

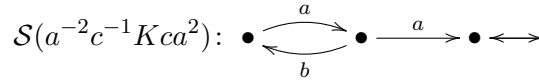
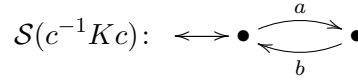
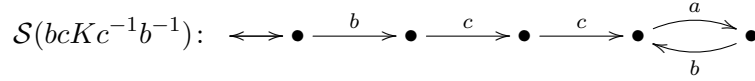
The next result should not be surprising taking into account the previous observations. A proof can be found in [11, Proposition 7.7].

Theorem 6.15. *Let A be a finite alphabet and $H, K \leq_{f.g.} F_A$. Then H and K are conjugate subgroups if and only if $\mathcal{C}(H) \cong \mathcal{C}(K)$.*

Example 6.16. Let $A = \{a, b, c\}$ and consider the Stallings automaton $\mathcal{S}(K)$



of a certain $K \leq_{f.g.} F_A$. Then we have



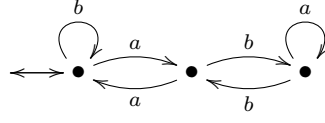
Recall that given a group G and $K \leq G$, we say that K is a **normal** subgroup of G , and write $K \trianglelefteq G$, if $gKg^{-1} = K$ for all $g \in G$. We say that a graph is **vertex-transitive** if, given two distinct vertices, there exists an automorphism of the graph taking one of the vertices to the other vertex. The following result is a consequence of Theorem 6.15.

Corollary 6.17. *Let $1 \neq K \leq_{f.g.} F_A$. Then $K \trianglelefteq F_A$ if and only if $[F_A : K] < \infty$ and $\mathcal{C}(K)$ is vertex-transitive.*

Proof. Let $\mathcal{S}(K) = (Q, A, \delta, q_0)$. In view of Theorem 6.15, K is normal in F_A if and only if $\mathcal{C}(K) \cong \mathcal{C}(gKg^{-1})$ for all $g \in F_A$. Thus, if $K \trianglelefteq F_A$, it is immediate that no $\mathcal{S}(gKg^{-1})$ can have a tail. This happens if and only if $\mathcal{S}(K)$ is a complete automaton which, by Theorem 6.13, is equivalent to the condition $[F_A : K] < \infty$. Consequently, every automaton $\mathcal{S}(gKg^{-1})$ can be obtained from $\mathcal{S}(K)$ by declaring one of the vertices to be the basepoint. Since $\mathcal{S}(gKg^{-1}) \cong \mathcal{S}(K)$ if and only if there exists an automorphism of $\mathcal{C}(K)$ taking q_0 to $(q_0, g^{-1})\delta$, this is equivalent to saying that $\mathcal{C}(K)$ is vertex-transitive, and the conclusion follows. $\square$

Example 6.18. Let $A = \{a, b\}$. The automaton in Example 6.14 is vertex-transitive and complete, so the corresponding subgroup is normal in F_A . However, if $K \leq_{f.g.} F_A$ is given

by the Stallings automaton



then K is not a normal subgroup of F_A because, despite being complete, $\mathcal{S}(K)$ is not vertex-transitive.

We end this section with a result concerning the intersection of finitely generated subgroups of a free group. Given a group G , we say that G has the **Howson property** if the intersection of any two finitely generated subgroups of G is again a finitely generated subgroup. It was Albert Howson [10] who first proved that free groups satisfy the Howson property, whence its name. A simple proof is given below, this time making use of Stallings automata.

Theorem 6.19. *Let H and K be two finitely generated subgroups of a free group F_A . Then $H \cap K \leq F_A$ is also finitely generated.*

Proof. Let q_0 and q'_0 be the basepoints of $\mathcal{S}(H)$ and $\mathcal{S}(K)$, respectively. By the proof of Proposition 4.9, we know that the product automaton $\mathcal{S}(H) \times \mathcal{S}(K)$ recognizes the language $L(\mathcal{S}(H)) \cap L(\mathcal{S}(K))$. It is easy to see that if $\mathcal{A}$ is the subautomaton of $\mathcal{S}(H) \times \mathcal{S}(K)$ induced by the connected component of (q_0, q'_0) , then also $L(\mathcal{A}) = L(\mathcal{S}(H)) \cap L(\mathcal{S}(K))$. Moreover, we can successively remove all the vertices in $\mathcal{A}$ with outdegree 1 which are not (q_0, q'_0) without changing the language it recognizes. So we obtain an equivalent automaton $\mathcal{B}$ which is an inverse automaton with a basepoint without vertices of outdegree 1 (except maybe the basepoint). Therefore, $\mathcal{B}$ is the Stallings automaton of some finitely generated subgroup of F_A . Observe now that

$$\begin{aligned} \overline{L(\mathcal{S}(H) \times \mathcal{S}(K))} &= L(\mathcal{S}(H) \times \mathcal{S}(K)) \cap \text{Irr}(\mathcal{R}_A) \\ &= L(\mathcal{S}(H)) \cap L(\mathcal{S}(K)) \cap \text{Irr}(\mathcal{R}_A) \\ &= L(\mathcal{S}(H)) \cap \text{Irr}(\mathcal{R}_A) \cap L(\mathcal{S}(K)) \cap \text{Irr}(\mathcal{R}_A) \\ &= \overline{L(\mathcal{S}(H))} \cap \overline{L(\mathcal{S}(K))} \end{aligned}$$

because $\overline{L(\mathcal{A}')} = L(\mathcal{A}') \cap \text{Irr}(\mathcal{R}_A)$ holds for every inverse automaton $\mathcal{A}'$ over A . So, in view of Proposition 6.10, the equalities

$$\overline{L(\mathcal{B})} = \overline{L(\mathcal{A})} = \overline{L(\mathcal{S}(H))} \cap \overline{L(\mathcal{S}(K))} = \overline{L(\mathcal{S}(H))} \cap \overline{L(\mathcal{S}(K))} = \overline{H} \cap \overline{K} = \overline{H \cap K}$$

allow us to deduce that $\mathcal{B} = \mathcal{S}(H \cap K)$. We conclude, in particular, that the subgroup $H \cap K$ is finitely generated. $\square$

7 Transition monoids of Stallings automata

We are finally ready to explore the connections between properties of a finitely generated subgroup $K \leq_{f.g.} F_A$ of a free group with $|A| \geq 2$ and properties of the transition monoid $M(K)$ of its Stallings automaton $\mathcal{S}(K)$. The departure point of this investigation is the article of Birget, Margolis, Meakin and Weil [3].

Given a group G , a subgroup $K \leq G$ is called **pure** if

$$\forall x \in G, \forall n \geq 1 \quad (x^n \in K \Rightarrow x \in K),$$

and **p -pure** (for a prime number p) if

$$\forall x \in G, \forall n \geq 1 \quad ((n, p) = 1 \Rightarrow (x^n \in K \Rightarrow x \in K)),$$

where (n, p) denotes the greatest common divisor of n and p . Recall also that a **p -group** is a group in which the order of every element is a power of a prime p . In [3], the authors prove that a subgroup $K \leq_{f.g.} F_A$ is pure (respectively, p -pure) if and only if $M(K)$ is aperiodic (respectively, **p -periodic**, i.e., every subgroup of $M(K)$ is a p -group).

Our goal is to find further pseudovarietal properties of $M(K)$ that constitute algebraic properties of K as a subgroup of F_A . By “algebraic properties” we mean properties involving K and F_A that are preserved under every automorphism of F_A . More precisely, given a free group F_A of finite rank, we want to find pseudovarieties $\mathbf{V}$ of finite (inverse) monoids satisfying

$$\forall K \leq_{f.g.} F_A, \forall \varphi \in \text{Aut}(F_A) \quad (M(K) \in \mathbf{V} \Rightarrow M(K\varphi) \in \mathbf{V}).$$

We say that a pseudovariety $\mathbf{V}$ satisfying this property is stable under all automorphisms of F_A .

In this section, we generalize the results mentioned in [3]. More precisely, we exhibit pseudovarieties other than the pseudovarieties of aperiodic and p -periodic monoids which are stable under all automorphisms of F_A , and we present some characterizations of finitely generated subgroups of F_A in terms of these pseudovarieties. We also extract some connections between conjugacy properties of K within F_A and properties of $M(K)$. Unless stated otherwise, the reader shall assume that A is a finite alphabet satisfying $|A| \geq 2$ and $a, b \in A$ are two distinct letters in the alphabet A .

7.1 Automorphisms and Stallings automata

We begin by exploring the effect of applying an automorphism of F_A to a subgroup $K \leq_{f.g.} F_A$ at the level of its Stallings automaton $\mathcal{S}(K)$.

Fix a finite alphabet A . Given $\varphi \in \text{Aut}(F_A)$, observe that φ induces a mapping $A \rightarrow \tilde{A}^*$, $a \mapsto \overline{a\varphi}$, which can be uniquely extended to a free monoid endomorphism $\phi: \tilde{A}^* \rightarrow \tilde{A}^*$. This will sometimes be useful.

Let $\mathcal{A}$ be an inverse automaton over A and let $w = w_1 w_2 \dots w_k \in \tilde{A}^*$, with $w_i \in \tilde{A}$ for

$1 \leq i \leq k$. Suppose that w labels a path

$$p_0 \xrightarrow{w_1} p_1 \xrightarrow{w_2} \dots \xrightarrow{w_k} p_k$$

in $\mathcal{A}$. We say that the path labeled by w is **reduced** if it does not contain any successive edges of the form

$$p_{i-1} \xrightarrow{w_i} p_i \xrightarrow{w_i^{-1}} p_{i-1}$$

for $1 \leq i \leq k$.

Proposition 7.1. *Let $\mathcal{S}(K) = (Q, A, \delta, q_0)$ be the Stallings automaton of $K \leq_{f.g.} F_A$ and let $\varphi \in \text{Aut}(F_A)$. Denote by $\phi: \tilde{A}^* \rightarrow \tilde{A}^*$ the monoid homomorphism induced by φ . Then $\mathcal{S}(K\varphi)$ is (isomorphic to) the inverse automaton obtained from $\mathcal{S}(K)$ by the following procedure:*

1. *Let $q \xrightarrow{a} q'$ be an edge in $\mathcal{S}(K)$ for some $q, q' \in Q$ and $a \in \tilde{A}$. Let $a\phi = a_1a_2\dots a_k$ ($k \geq 1$, $a_i \in \tilde{A}$) be the factorization of the word $a\phi$ into letters. Then replace that edge by the sequence of edges*

$$q = p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_k} p_k = q',$$

where $p_1, \dots, p_{k-1}$ are new vertices, and add the corresponding inverse edges.

2. *After completing all the edge replacements mentioned above, apply the necessary foldings in order to get an inverse automaton.*
3. *Successively eliminate every vertex with outdegree 1 which is not the basepoint.*

Proof. Let $\mathcal{A}$ be the automaton we obtain after following step 1; $\mathcal{A}'$ the automaton we get by the end of step 2; and $\mathcal{A}''$ the final automaton.

Since $\mathcal{A}''$ is inverse and the only vertex which may have outdegree 1 is the basepoint, we know that it is the Stallings automaton of some finitely generated subgroup of F_A . Thus, we only need to show that $\overline{K\varphi} = \overline{L(\mathcal{A}'')}$, in view of Proposition [6.10](#). We start by observing that $\overline{L(\mathcal{A}'')} = \overline{L(\mathcal{A}')} = \overline{L(\mathcal{A})}$. Indeed, regarding the latter equality, given $u \in L(\mathcal{A}')$, we know that u labels a successful path in $\mathcal{A}'$ which can be lifted to a successful path in $\mathcal{A}$ labeled by a word v obtained by inserting factors of the form aa^{-1} ($a \in \tilde{A}$) into u . Hence, $\bar{u} = \bar{v} \in \overline{L(\mathcal{A})}$ and $\overline{L(\mathcal{A}')} \subseteq \overline{L(\mathcal{A})}$. On the other hand, since the inverse automaton $\mathcal{A}'$ is obtained from $\mathcal{A}$ by simply folding edges, it is clear that $L(\mathcal{A}) \subseteq L(\mathcal{A}')$, so $\overline{L(\mathcal{A})} \subseteq \overline{L(\mathcal{A}')}$. As for the first equality, since $\mathcal{A}''$ is a subautomaton of $\mathcal{A}'$, we have $L(\mathcal{A}'') \subseteq L(\mathcal{A}')$, and any word accepted by $\mathcal{A}'$ gives rise to a word accepted by $\mathcal{A}''$ by deleting some factors of the form aa^{-1} ($a \in \tilde{A}$). Therefore, $\overline{L(\mathcal{A}')} \subseteq \overline{L(\mathcal{A}'')}$ and the chain of equalities follows.

Now, given $u \in \overline{K\varphi}$, there exists $v = v_1v_2\dots v_k \in \bar{K}$, with $v_i \in \tilde{A}$ for $1 \leq i \leq k$, such that $u = \bar{v\phi}$. We know that v labels a path

$$q_0 \xrightarrow{v_1} q_1 \xrightarrow{v_2} \dots \xrightarrow{v_k} q_k = q_0$$

in $\mathcal{S}(K)$ so, by step 1, we get a successful path

$$q_0 \xrightarrow{v_1\phi} q_1 \xrightarrow{v_2\phi} \cdots \xrightarrow{v_k\phi} q_k = q_0$$

in $\mathcal{A}$. We deduce that $v\phi \in L(\mathcal{A})$, which entails $u = \overline{v\phi} \in \overline{L(\mathcal{A})}$. Hence, $\overline{K\varphi} \subseteq \overline{L(\mathcal{A})}$.

As for the opposite inclusion, given $u = u_1u_2 \dots u_k \in L(\mathcal{A})$, with $u_i \in \tilde{A}$ for $1 \leq i \leq k$, we have a path

$$q_0 \xrightarrow{u_1} q_1 \xrightarrow{u_2} \cdots \xrightarrow{u_k} q_k = q_0$$

in $\mathcal{A}$, which can be modified to yield a reduced path $q_0 \xrightarrow{w} q_0$ in $\mathcal{A}$. By doing that, we get $\bar{u} = \bar{w}$, and there exists a successful path $q_0 \xrightarrow{v} q_0$ in $\mathcal{S}(K)$ for some $v \in \tilde{A}^*$ satisfying $w = v\phi$. Therefore, we obtain $w \in L(\mathcal{S}(K))\phi$, which implies that $\bar{u} = \bar{w} \in \overline{L(\mathcal{S}(K))\phi} = \overline{K\varphi}$, concluding the proof. $\square$

Recall that, by Theorem 6.4, $\text{Aut}(F_A)$ is generated by the automorphisms α_a for $a \in A$ (of type 1), carrying a to a^{-1} and fixing the other letters, and β_{ab} for $a, b \in A$ (of type 2), carrying a to ab and fixing the other letters. We also consider the automorphisms β_{ab}^{-1} ($a, b \in A$) to be of type 2. It is particularly easy to understand the effect of an automorphism of type 1 at the level of Stallings automata.

Lemma 7.2. *Let $K \leq_{f.g.} F_A$. Then $M(K\alpha_a) = M(K)$ for all $a \in A$.*

Proof. Let δ be the transition function of $\mathcal{S}(K)$ and θ the transition function of $\mathcal{S}(K\alpha_a)$. Then the functions of the type δ_{x^ε} and θ_{x^ε} , for $x \in A$ and $\varepsilon \in \{-1, 1\}$, generate $M(K)$ and $M(K\alpha_a)$ as monoids, respectively. Now it is enough to observe that $\theta_{a^\varepsilon} = \delta_{a^{-\varepsilon}}$ and $\theta_{x^\varepsilon} = \delta_{x^\varepsilon}$ for $x \in A \setminus \{a\}$. So the generating sets above are equal, and thus generate the same transition monoid. $\square$

Given a class of monoids $\mathcal{C}$ and $\varphi \in \text{Aut}(F_A)$, we say that $\mathcal{C}$ is **stable under** φ if the condition

$$M(K) \in \mathcal{C} \Rightarrow M(K\varphi) \in \mathcal{C}$$

holds for all $K \leq_{f.g.} F_A$.

Lemma 7.3. *Let $\mathcal{C}$ be any class of monoids. If $\mathcal{C}$ is stable under all automorphisms of types 1 and 2, then $\mathcal{C}$ is stable under all automorphisms of F_A .*

Proof. First observe that an automorphism of type 1 is its own inverse. Assume that $\mathcal{C}$ is stable under all the automorphisms α_a , β_{ab} and β_{ab}^{-1} . Let $K \leq_{f.g.} F_A$ be such that $M(K) \in \mathcal{C}$, and take any $\varphi \in \text{Aut}(F_A)$. By Theorem 6.4, we can write $\varphi = \varphi_1^{\varepsilon_1} \varphi_2^{\varepsilon_2} \cdots \varphi_k^{\varepsilon_k}$ ($k \geq 1$) where each φ_i is an elementary Nielsen automorphism and $\varepsilon_i \in \{-1, 1\}$. We show by induction on k that $M(K\varphi) \in \mathcal{C}$. If $k = 1$, then φ is an elementary Nielsen automorphism or the inverse of one, so the conclusion follows by assumption. Suppose now that $k > 1$ and the claim is satisfied for an automorphism which is the composite of $k - 1$ automorphisms of types 1 and 2, and let $K' = K\varphi_1^{\varepsilon_1} \cdots \varphi_{k-1}^{\varepsilon_{k-1}}$. Then, by the induction hypothesis, we obtain $M(K') \in \mathcal{C}$. Since $\varphi_k^{\varepsilon_k}$ is an automorphism of type 1 or 2, it follows that $M(K\varphi) = M(K'\varphi_k^{\varepsilon_k}) \in \mathcal{C}$, as we wanted. $\square$

Given $K \leq_{f.g.} F_A$, let $\mathcal{S}(K) = (Q, A, \delta, q_0)$ be the Stallings automaton of K . We proceed by analyzing the effect of an automorphism of type 2 on $\mathcal{S}(K)$. Out of convenience, we paint the vertices in Q **red**.

Let $a, b \in A$ be two distinct letters, and write $\beta = \beta_{ab}$. When no confusion arises, we will always write β instead of β_{ab} . Consider the automaton $\mathcal{B}$ obtained from $\mathcal{S}(K)$ by replacing each edge $q \xrightarrow{a} q'$ (and its inverse) with $q \xrightarrow{a} r \xrightarrow{b} q'$ (and their inverses), where $r \notin Q$ is a new vertex which we paint **blue**. This way, $\mathcal{B}$ becomes an involutive automaton. Observe that, by doing this, $\mathcal{B}$ is not necessarily deterministic, as we may encounter something like

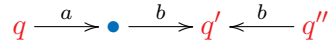


In fact, ambiguity only arises in these cases, which appear if and only if we find

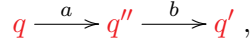


in $\mathcal{S}(K)$. After we perform all the foldings in $\mathcal{B}$ and delete all the vertices with outdegree 1 which are not the basepoint, we get an inverse automaton which is (isomorphic to) $\mathcal{S}(K^\beta)$, in view of Proposition [7.1](#).

Currently, our goal is to show that by doing only “first order” foldings, i.e., by replacing each occurrence of



in $\mathcal{B}$ with



we already get a deterministic automaton.

Consider the involutive automaton $\mathcal{S}(K)^\beta = (Q^\beta, A, \theta, q_0)$ obtained from $\mathcal{S}(K)$ by doing the following:

1. We keep all (positive) edges with label $x \in A \setminus \{a\}$.
2. If $q \xrightarrow{a} q' \xleftarrow{b} q''$ appears in $\mathcal{S}(K)$, we replace $q \xrightarrow{a} q'$ with $q \xrightarrow{a} q''$.
3. If $q \xrightarrow{a} q'$ appears in $\mathcal{S}(K)$, but it is not possible to read b^{-1} from q' , then we replace that edge with $q \xrightarrow{a} r \xrightarrow{b} q'$, where $r \notin Q$ is a new vertex.

Note that $Q \subseteq Q^\beta$ and $\mathcal{S}(K)^\beta$ is isomorphic to the automaton we get after performing all “first order” foldings in $\mathcal{B}$. We keep considering that the vertices in Q are painted red, and the ones in $Q^\beta \setminus Q$ are painted blue. When going from $\mathcal{S}(K)$ to $\mathcal{S}(K)^\beta$, we never change the starting point of an already existing (positive) edge; in particular, the amount of positive edges with some fixed label coming from a red vertex does not change, and it is never greater than 1. Second, every blue vertex in Q^β has outdegree (and also indegree, since $\mathcal{S}(K)^\beta$ is involutive) exactly two, because there is always one (and only one) positive edge labeled by a entering a blue vertex and one (and only one) positive edge labeled by b leaving it. We denote by $M(K)^\beta$ the transition monoid of $\mathcal{S}(K)^\beta$.

Let $q, q' \in Q^\beta$ be two distinct vertices. We have the following:

- (i) Given $x \in A$, we never encounter $q \xleftarrow{x} \bullet \xrightarrow{x} q'$ in $\mathcal{S}(K)^\beta$ due to the previous remarks.
- (ii) Given $x \in A \setminus \{a, b\}$, a pair of edges $q \xrightarrow{x} \bullet \xleftarrow{x} q'$ appears in $\mathcal{S}(K)^\beta$ if and only if it appears in $\mathcal{S}(K)$, and that never happens because $\mathcal{S}(K)$ is deterministic. Furthermore, if we had $q \xrightarrow{a} \bullet \xleftarrow{a} q'$ in $\mathcal{S}(K)^\beta$, then we would have

$$q \xrightarrow{a} p \xleftarrow{b} \bullet \xrightarrow{b} p' \xleftarrow{a} q'$$

in $\mathcal{S}(K)$, yielding $p = p'$ and $q = q'$, a contradiction.

- (iii) If $q, q' \in Q$, then $q \xrightarrow{b} \bullet \xleftarrow{b} q'$ cannot occur in $\mathcal{S}(K)$, so it cannot appear in $\mathcal{S}(K)^\beta$ either. Moreover, if $q \xrightarrow{b} q' \xleftarrow{b} \bullet$ appeared in $\mathcal{S}(K)^\beta$, then we would have

$$q \xrightarrow{b} q' \xleftarrow{b} \bullet \xleftarrow{a} q''$$

in $\mathcal{S}(K)^\beta$, for some $q'' \in Q$. Hence, $q \xrightarrow{b} q' \xleftarrow{a} q''$ would occur in $\mathcal{S}(K)$ and, as a consequence, we would get $\bullet \xleftarrow{a} q'' \xrightarrow{a} q'$ in $\mathcal{S}(K)^\beta$, which is absurd, according to (i).

- (iv) Finally, if $q, q' \in Q^\beta \setminus Q$ and we had $q \xrightarrow{b} \bullet \xleftarrow{b} q'$ in $\mathcal{S}(K)^\beta$, then we would necessarily find $p \xrightarrow{a} \bullet \xleftarrow{a} p'$ in $\mathcal{S}(K)^\beta$ for some distinct vertices $p, p' \in Q$, which cannot happen.

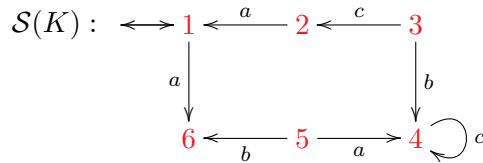
The previous observations show that $\mathcal{S}(K)^\beta$ is deterministic, whence inverse. We conclude that $\mathcal{S}(K^\beta) = (Q', A, \gamma, q_0)$ is the (inverse) automaton we obtain from $\mathcal{S}(K)^\beta$ by deleting all vertices with outdegree 1 which are not the basepoint.

Note that $\mathcal{S}(K^{\beta^{-1}})$ can be obtained by a similar procedure. This time, we let $\mathcal{B}'$ be the involutive automaton obtained from $\mathcal{S}(K)$ by replacing $q \xrightarrow{a} q'$ with $q \xrightarrow{a} r \xleftarrow{b} q'$, where $r \notin Q$ is a new vertex. Then, after we perform all the “first order” foldings in $\mathcal{B}'$, we arrive at $\mathcal{S}(K)^{\beta^{-1}}$, which is also the involutive automaton obtained from $\mathcal{S}(K)$ by doing the following:

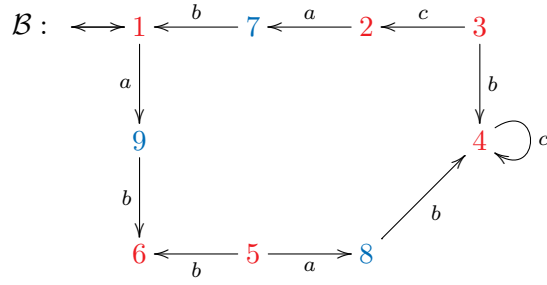
1. We keep all (positive) edges with label $x \in A \setminus \{a\}$.
2. If $q \xrightarrow{a} q' \xrightarrow{b} q''$ appears in $\mathcal{S}(K)$, then we replace $q \xrightarrow{a} q'$ with $q \xrightarrow{a} q''$.
3. If $q \xrightarrow{a} q'$ appears in $\mathcal{S}(K)$, but it is not possible to read b from q' , then we replace that edge with $q \xrightarrow{a} r \xleftarrow{b} q'$, where $r \notin Q$ is a new vertex.

After deleting all the vertices with outdegree 1 which are not the basepoint in $\mathcal{S}(K)^{\beta^{-1}}$, we end up with $\mathcal{S}(K^{\beta^{-1}})$. In this case, the notation we use is $\mathcal{S}(K)^{\beta^{-1}} = (Q^{\beta^{-1}}, A, \theta', q_0)$.

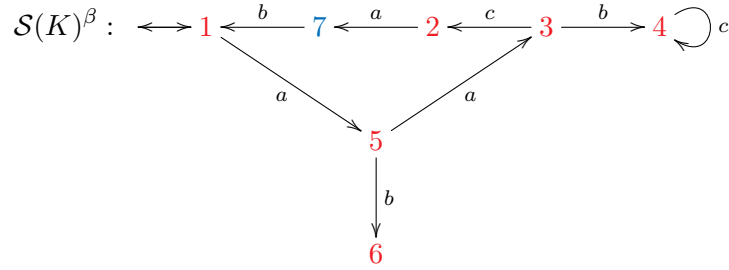
Example 7.4. Let $A = \{a, b, c\}$ and let $K \leq_{f.g.} F_A$ be given by



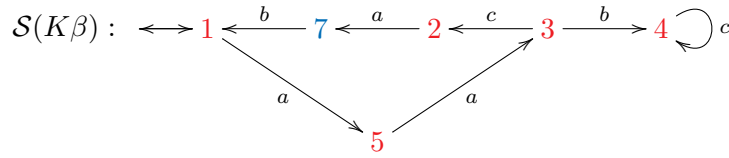
The automaton $\mathcal{B}$ described above is obtained by replacing every edge $i \xrightarrow{a} j$ with a pair of edges $i \xrightarrow{a} k \xrightarrow{b} j$, as we depict below.



After performing all the “first order” foldings in $\mathcal{B}$, we obtain an automaton isomorphic to $\mathcal{S}(K)^\beta$, which is achieved by making the following substitutions in $\mathcal{S}(K)$: replace the edge $2 \xrightarrow{a} 1$ with $2 \xrightarrow{a} 7 \xrightarrow{b} 1$; replace the edge $5 \xrightarrow{a} 4$ with $5 \xrightarrow{a} 3$; and replace the edge $1 \xrightarrow{a} 6$ with $1 \xrightarrow{a} 5$.



Finally, the Stallings automaton of $K\beta \leq_{f.g.} F_A$ is achieved by deleting vertex 6 in $\mathcal{S}(K)^\beta$, because it has outdegree 1 and it is not the basepoint.



We end this section with a lemma that we will use extensively. Let $\rho: \tilde{A}^* \rightarrow \tilde{A}^*$ be the unique free monoid homomorphism extending the function

$$\begin{aligned} \tilde{A} &\longrightarrow \tilde{A}^* \\ a &\longmapsto ab \\ a^{-1} &\longmapsto b^{-1}a^{-1} \\ x &\longmapsto x, \text{ if } x \in \tilde{A} \setminus \{a, a^{-1}\}. \end{aligned}$$

Similarly, let $\eta: \tilde{A}^* \rightarrow \tilde{A}^*$ be the extension of

$$\begin{aligned} \tilde{A} &\longrightarrow \tilde{A}^* \\ a &\longmapsto ab^{-1} \\ a^{-1} &\longmapsto ba^{-1} \\ x &\longmapsto x, \text{ if } x \in \tilde{A} \setminus \{a, a^{-1}\}. \end{aligned}$$

We have the following result.

Lemma 7.5. *Let $q, q' \in Q$ and $u \in \tilde{A}^*$. Given $K \leq_{f.g.} F_A$, let $\mathcal{S}(K) = (Q, A, \delta, q_0)$ and $\mathcal{S}(K)^\beta = (Q^\beta, A, \theta, q_0)$.*

- (a) *If $(q, u)\delta = q'$, then $(q, u\rho)\theta = q'$. As a consequence, $(q, \overline{u\rho})\theta = q'$.*
- (b) *If $(q, u)\theta = q'$, then $(q, \overline{u\eta})\delta = q'$.*

Proof. If $u = 1$, both items are trivial, so we carry on with $u \in \tilde{A}^+$.

- (a) We can write $u = u_1 u_2 \dots u_m$, where each u_i is either a letter, the word ab^{-1} or the word ba^{-1} , such that $u_i u_{i+1} \notin \{ab^{-1}, ba^{-1}\}$, for $1 \leq i \leq m-1$. Indeed, such factorization of u is always possible, for ab^{-1} and ba^{-1} have no letters in common, and it is unique. We then have a path

$$q = \textcolor{red}{p_0} \xrightarrow{u_1} \textcolor{red}{p_1} \xrightarrow{u_2} \dots \xrightarrow{u_{m-1}} \textcolor{red}{p_{m-1}} \xrightarrow{u_m} \textcolor{red}{p_m} = q'$$

in $\mathcal{S}(K)$, for some vertices $p_i \in Q$ and $0 \leq i \leq m$.

Given any $1 \leq i \leq m$, we analyse three possible cases.

- If $u_i \notin \{a^\varepsilon, (ab^{-1})^\varepsilon \mid \varepsilon = \pm 1\}$, then

$$(p_{i-1}, u_i)\delta = (p_{i-1}, u_i)\theta = (p_{i-1}, u_i\rho)\theta.$$

- If $u_i = a^\varepsilon$ with $\varepsilon \in \{-1, 1\}$, then

$$(p_{i-1}, u_i)\delta = (p_{i-1}, a^\varepsilon)\delta = (p_{i-1}, (ab)^\varepsilon)\theta = (p_{i-1}, u_i\rho)\theta.$$

- If $u_i = (ab^{-1})^\varepsilon$ with $\varepsilon \in \{-1, 1\}$, then

$$(p_{i-1}, u_i)\delta = (p_{i-1}, (ab^{-1})^\varepsilon)\delta = (p_{i-1}, a^\varepsilon)\theta = (p_{i-1}, (abb^{-1})^\varepsilon)\theta = (p_{i-1}, u_i\rho)\theta,$$

since $\theta_a = \theta_{abb^{-1}}$, by construction of $\mathcal{S}(K)^\beta$.

To finish, we show that $(q, u_1 \dots u_m)\delta = (q, (u_1 \dots u_m)\rho)\theta$ by induction on m . If $m = 1$, then

$$(q, u_1)\delta = (p_0, u_1)\delta = (p_0, u_1\rho)\theta = (q, u_1\rho)\theta,$$

so the claim holds. Proceed with $m > 1$ and suppose that the equality is satisfied for $m-1$. Then we obtain

$$\begin{aligned} (q, u_1 \dots u_m)\delta &= (p_{m-1}, u_m)\delta \\ &= (p_{m-1}, u_m\rho)\theta \\ &= ((q, u_1 \dots u_{m-1})\delta, u_m\rho)\theta \\ &= ((q, u_1\rho \dots u_{m-1}\rho)\theta, u_m\rho)\theta \\ &= (q, u_1\rho \dots u_m\rho)\theta \\ &= (q, (u_1 \dots u_m)\rho)\theta, \end{aligned}$$

since ρ is a homomorphism, and the statement follows.

- (b) Note that $(q, u)\delta = q' \Rightarrow (q, \bar{u})\delta = q'$ and $\bar{u}\bar{\eta} = \overline{u\eta}$, so we can assume that u is reduced. The proof is similar to the one above. Indeed, we can write $u = u_1 u_2 \dots u_m$, where each u_i is either a letter, the word ab or the word $b^{-1}a^{-1}$, such that $u_i u_{i+1} \notin \{ab, b^{-1}a^{-1}\}$, for $1 \leq i \leq m-1$. We get a path

$$q = p_0 \xrightarrow{u_1} p_1 \xrightarrow{u_2} \dots \xrightarrow{u_{m-1}} p_{m-1} \xrightarrow{u_m} p_m = q'$$

in $\mathcal{S}(K)^\beta$, for some vertices $p_i \in Q$ and $0 \leq i \leq m$.

Let $\varepsilon \in \{-1, 1\}$. Given any $1 \leq i \leq m$, again there are three possible cases.

- If $u_i \notin \{a^\varepsilon, (ab)^\varepsilon\}$, then

$$(p_{i-1}, u_i)\theta = (p_{i-1}, u_i)\delta = (p_{i-1}, u_i\eta)\delta = (p_{i-1}, \bar{u_i\eta})\delta.$$

- If $u_i = a^\varepsilon$, then

$$(p_{i-1}, u_i)\theta = (p_{i-1}, a^\varepsilon)\theta = (p_{i-1}, (ab^{-1})^\varepsilon)\delta = (p_{i-1}, u_i\eta)\delta = (p_{i-1}, \bar{u_i\eta})\theta.$$

- If $u_i = (ab)^\varepsilon$, then

$$(p_{i-1}, u_i)\theta = (p_{i-1}, (ab)^\varepsilon)\theta = (p_{i-1}, a^\varepsilon)\delta = (p_{i-1}, \bar{u_i\eta})\delta,$$

$$\text{since } u_i\eta = (ab^{-1}b)^\varepsilon.$$

Now we induct on m to show that $(q, u_1 \dots u_m)\theta = (q, \bar{u_1\eta} \dots \bar{u_m\eta})\delta$. If $m = 1$, then

$$(q, u_1)\theta = (p_0, u_1)\theta = (p_0, \bar{u_1\eta})\delta = (q, \bar{u_1\eta})\delta,$$

so the claim holds. Proceed with $m > 1$ and assume that the equality is satisfied for $m-1$. We get

$$\begin{aligned} (q, u_1 \dots u_m)\theta &= (p_{m-1}, u_m)\theta \\ &= (p_{m-1}, \bar{u_m\eta})\delta \\ &= ((q, u_1 \dots u_{m-1})\theta, \bar{u_m\eta})\delta \\ &= ((q, \bar{u_1\eta} \dots \bar{u_{m-1}\eta})\delta, \bar{u_m\eta})\delta \\ &= (q, \bar{u_1\eta} \dots \bar{u_m\eta})\delta, \end{aligned}$$

as we wanted.

Hence, we conclude that

$$\begin{aligned}
(q, u)\theta &= (q, u_1 \dots u_m)\theta \\
&= (q, \overline{u_1\eta} \dots \overline{u_m\eta})\delta \\
&= (q, \overline{u_1\eta \dots u_m\eta})\delta \\
&= (q, \overline{u_1\eta \dots u_m\eta})\delta \\
&= (q, \overline{u\eta})\delta.
\end{aligned}$$

□

When dealing with the automorphism β^{-1} instead of β , by switching the roles of ρ and η we get the following analogue of Lemma 7.5.

Lemma 7.6. *Let $q, q' \in Q$ and $u \in \tilde{A}^*$. Given $K \leq_{f.g.} F_A$, let $\mathcal{S}(K) = (Q, A, \delta, q_0)$ and $\mathcal{S}(K)^{\beta^{-1}} = (Q^{\beta^{-1}}, A, \theta', q_0)$.*

- (a) *If $(q, u)\delta = q'$, then $(q, u\eta)\theta' = q'$. As a consequence, $(q, \overline{u\eta})\theta' = q'$.*
- (b) *If $(q, u)\theta' = q'$, then $(q, \overline{u\rho})\delta = q'$.*

7.2 Transition groups

In this small section, we focus on finitely generated subgroups $K \leq_{f.g.} F_A$ of a free group with the property that $M(K)$ is a group. We begin with a simple lemma.

Lemma 7.7. *Let $\mathcal{A}$ be an inverse automaton over the alphabet A which is not complete. Then there exists a word $w \in \tilde{A}^*$ which cannot be read from any state in $\mathcal{A}$.*

Proof. Let $\mathcal{A}$ be an inverse and incomplete automaton, with set of states $Q = \{q_1, \dots, q_n\}$ for some $n \geq 1$, and transition function δ . We can assume that there exists a letter $a \in A$ such that $q_1 \notin \text{dom } \delta_a$. Since $\mathcal{A}$ is inverse, for each $1 \leq i \leq n$ there exists $u_i \in \tilde{A}^*$ such that $(q_i, u_i)\delta = q_1$. Consider the word

$$w = (u_1a)(u_1a)^{-1}(u_2a)(u_2a)^{-1} \dots (u_na)(u_na)^{-1}.$$

Since every factor of the type $(u_ja)(u_ja)^{-1}$ labels a loop at every state in $\text{dom } \delta_{(u_ja)(u_ja)^{-1}}$, and $q_i \notin \text{dom } \delta_{u_i a}$ for every $1 \leq i \leq n$, it follows that w is a word which cannot be read from any state in $\mathcal{A}$. □

Note that if there exists a word which cannot be read from any state in an automaton, then its transition monoid has a zero — the empty transformation ($\emptyset$). This remark yields the following corollary.

Corollary 7.8. *Consider $1 \neq K \leq_{f.g.} F_A$ with Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$. The following conditions are equivalent:*

- (i) $[F_A : K] < \infty$.

(ii) $\mathcal{S}(K)$ is complete.

(iii) $M(K)$ is a group.

Proof. (i) $\Leftrightarrow$ (ii): Follows from Theorem 6.13.

(ii) $\Rightarrow$ (iii): Let $1 \neq K \leq_{f.g.} F_A$ be such that $\mathcal{S}(K)$ is a complete automaton. In order to prove that $M(K)$ is a group, it suffices to show that $M(K)$ has a unique idempotent. Since $M(K)$ is a submonoid of the symmetric inverse monoid $\mathcal{I}_Q$, if $\delta_u \in E(M(K))$ then $\delta_u = \text{id}|_{\text{dom } \delta_u}$. But the completeness of $\mathcal{S}(K)$ implies $\text{dom } \delta_u = Q$, and thus $\delta_u = \text{id}_Q$. So the identity transformation is the only idempotent of $M(K)$, whence $M(K)$ is a group.

(iii) $\Rightarrow$ (ii): Suppose first that $M(K)$ is a group. By construction of $\mathcal{S}(K)$, there exist a cyclically reduced word $u \in \tilde{A}^*$ and $q \in Q$ such that u labels a loop at q . Hence, δ_u is not the empty transformation and, since $M(K)$ is finite, there exists $n \geq 1$ such that $\delta_u^n \in E(M(K))$. Hence, if $M(K)$ is a group, its unique idempotent is not the empty transformation, which means that every word in $\tilde{A}^*$ can be read from at least one state in $\mathcal{S}(K)$. In view of the previous lemma, we conclude that $\mathcal{S}(K)$ is complete. $\square$

We are ready to prove that every class of groups is stable under all automorphisms of a free group.

Theorem 7.9. *Let $\mathcal{C}$ be any class of groups. If $M(K) \in \mathcal{C}$, then $M(K\varphi) \in \mathcal{C}$.*

Proof. The case $K = 1$ is clear, so we focus on the case $K \neq 1$. Let $\mathcal{S}(K) = (Q, A, \delta, q_0)$ be the Stallings automaton of K and $\mathcal{S}(K\beta) = (Q', A, \gamma, q_0)$ be the Stallings automaton of $K\beta$. By the previous, and since $\mathcal{S}(K)$ is complete, $\mathcal{S}(K)^\beta$ is the automaton obtained from $\mathcal{S}(K)$ by simply replacing every occurrence of $q \xrightarrow{a} q' \xleftarrow{b} q''$ with $q \xrightarrow{a} q'' \xrightarrow{b} q'$. Moreover, since $\mathcal{S}(K)^\beta$ is also complete, it has no vertices with outdegree 1, whence $\mathcal{S}(K\beta) = \mathcal{S}(K)^\beta$.

Now, we know that $M(K\beta)$ is generated, as a group, by γ_a and all the γ_x with $x \in A \setminus \{a\}$. But $M(K\beta)$ is also generated by γ_{ab} and γ_x ($x \in A \setminus \{a\}$), since we can write $\gamma_a = \gamma_{ab}\gamma_b^{-1}$. Moreover, note that $\gamma_{ab} = \delta_a$ and $\gamma_x = \delta_x$ for all $x \in A \setminus \{a\}$. Since δ_a, δ_x ($x \in A \setminus \{a\}$) generate $M(K)$ as a group, we get $M(K) = M(K\beta)$. Similarly, observe that $\mathcal{S}(K\beta^{-1})$ is the automaton obtained from $\mathcal{S}(K)$ by replacing every occurrence of $q \xrightarrow{a} q' \xrightarrow{b} q''$ with $q \xrightarrow{a} q'' \xleftarrow{b} q'$. Hence, by a similar argument, we get $M(K) = M(K\beta^{-1})$, and therefore $M(K\beta^{\pm 1}) \in \mathcal{C}$. The conclusion now follows from Lemma 7.2 and Lemma 7.3. $\square$

When the transition monoid of $\mathcal{S}(K)$ is a group, we can actually characterize K in terms of the pseudovarieties $M(K)$ belongs to.

Theorem 7.10. *Let $K \leq_{f.i.} F_A$ with $\mathcal{S}(K) = (Q, A, \delta, q_0)$. Consider the pseudovariety of groups $\mathbf{V}$ ultimately defined by the sequence of identities $(u_n = 1)_n$ over the set of variables X . The following conditions are equivalent:*

(i) $M(K) \in \mathbf{V}$.

(ii) There exists $p \geq 1$ such that, for every group homomorphism $\psi: F_X \rightarrow F_A$, the condition $u_n\psi \in K$ holds for all $n \geq p$.

Proof. (i) $\Rightarrow$ (ii): By assumption, K has finite index in F_A , so the Stallings automaton $\mathcal{S}(K)$ is complete. This implies that if $v, w \in \tilde{A}^*$ are such that $\bar{v} = \bar{w}$, then for every $q \in Q$ we have

$$(q, v)\delta = (q, \bar{v})\delta = (q, \bar{w})\delta = (q, w)\delta.$$

Hence, the monoid homomorphism

$$\begin{aligned} \Delta: \tilde{A}^* &\longrightarrow M(K) \\ u &\longmapsto \delta_u \end{aligned}$$

induces a group homomorphism

$$\begin{aligned} \tilde{\Delta}: F_A &\longrightarrow M(K) \\ g &\longmapsto \delta_{\bar{g}}. \end{aligned}$$

By assumption, there exists $p \geq 1$ such that $M(K) \models u_n = 1$ for all $n \geq p$. This means that given any group homomorphism $\varphi: F_X \rightarrow M(K)$, we have $u_n\varphi = 1$ for all $n \geq p$. Fix such a $p \geq 1$ and let $\psi: F_X \rightarrow F_A$ be a homomorphism. Since $\psi\tilde{\Delta}: F_X \rightarrow M(K)$ is a homomorphism, it follows that $u_n\psi\tilde{\Delta} = 1$ for all $n \geq p$. This implies that $\delta_{\overline{u_n\psi}} = 1$ for all $n \geq p$, and thus $(q, \overline{u_n\psi})\delta = q$ for all $q \in Q$ and $n \geq p$. In particular, we have $(q_0, \overline{u_n\psi})\delta = q_0$ for all $n \geq p$, which shows that $u_n\psi \in K$ for all $n \geq p$, as we wanted.

(ii) $\Rightarrow$ (i): Let $\phi: \tilde{A}^* \rightarrow F_A$ be the quotient morphism. Take $p \geq 1$ satisfying (ii) and $n \geq p$. There exist $x_1, \dots, x_m \in X$ and $k_1, \dots, k_m \in \mathbb{Z}$ such that $u_n = x_1^{k_1} \cdots x_m^{k_m} \in F_X$. Our goal is to prove that, given any words $w_1, \dots, w_m \in \tilde{A}^*$, we have $\delta_{w_1}^{k_1} \cdots \delta_{w_m}^{k_m} = 1$, which is the same as proving that $(q, w_1^{k_1} \cdots w_m^{k_m})\delta = q$ for all $q \in Q$. Note that the condition $K \leq_{f.i.} F_A$ entails that $\mathcal{S}(K)$ is complete.

Write $w = w_1^{k_1} \cdots w_m^{k_m}$, take any $q \in Q$, and let $v \in \tilde{A}^*$ be such that $(q_0, v)\delta = q$. Since F_X is a free group, the function $X \rightarrow F_A$, $x_i \mapsto (vw_i v^{-1})\phi$ induces a homomorphism $\psi: F_X \rightarrow F_A$ satisfying $u_n\psi = (vwv^{-1})\phi$. Then $(vwv^{-1})\phi \in K$ and thus $(q_0, vwv^{-1})\delta = q_0$, by construction of $\mathcal{S}(K)$ and taking into account that it is a complete automaton. We conclude that $(q, w)\delta = q$, as we wanted. $\square$

Example 7.11. The pseudovariety of abelian groups is given by $\mathbf{Ab} = \llbracket xyx^{-1}y^{-1} = 1 \rrbracket$, and the pseudovariety of p -groups is given by $\mathbf{G}_p = \llbracket x^{p^\omega} = 1 \rrbracket$. Hence, the above theorem allows us to conclude that $M(K)$ is an abelian group if and only if K has finite index and contains all elements of the type $xyx^{-1}y^{-1}$ ($x, y \in F_A$); and it is a p -group if and only if it has finite index and there exists $m \geq 1$ such that $x^{p^m} \in K$ for all $x \in F_A$.

7.3 Pseudovarieties of type $\bar{\mathbf{H}}$

In this section, we prove that if $\mathbf{H}$ is a pseudovariety of finite groups, then the pseudovariety of monoids $\bar{\mathbf{H}}$ (see Proposition 5.12) is stable under all automorphisms of F_A . Then we characterize some finitely generated subgroups of F_A in terms of these pseudovarieties.

Recall that, given $K \leq_{f.g.} F_A$ with Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$, the automaton $\mathcal{S}(K^\beta) = (Q', A, \gamma, q_0)$ is obtained from $\mathcal{S}(K)^\beta = (Q^\beta, A, \theta, q_0)$ by successively deleting all vertices with outdegree 1 which are not the basepoint. We denote by $M(K)^\beta$ the transition monoid of $\mathcal{S}(K)^\beta$, and we paint the vertices in Q red and the ones in $Q^\beta \setminus Q$ blue, as usual. Let us begin with a simple lemma regarding the group $\mathcal{H}$ -classes of an inverse monoid.

Lemma 7.12. *Let M be an inverse monoid and denote by H_x the $\mathcal{H}$ -class of $x \in M$. Then H_x is a group if and only if $xx^{-1} = x^{-1}x$. Moreover, $y \in H_x$ if and only if $yy^{-1} = xx^{-1}$ and $y^{-1}y = x^{-1}x$.*

Proof. Let us start with the first statement. If $x \in M$ is such that H_x is a group, then x^{-1} is precisely the inverse of x seen as an element of the group H_x , and therefore $xx^{-1} = x^{-1}x$. On the other hand, since $x \mathcal{R} xx^{-1}$ and $x \mathcal{L} x^{-1}x$, the equality $x^{-1}x = xx^{-1}$ yields $x \mathcal{H} xx^{-1}$. Then H_x contains the idempotent xx^{-1} , and thus it is a group, in view of Corollary 2.14. The other statement follows from the characterization of Green's relations in an inverse monoid (Proposition 3.6). $\square$

Observe that if the elements of a monoid M are injective partial transformations, then $f, g \in M$ are in the same group $\mathcal{H}$ -class if and only if $\text{dom } f = \text{im } f = \text{dom } g = \text{im } g$. We will now show that every group $\mathcal{H}$ -class of $M(K)^\beta$ is the homomorphic image of a subgroup of some group $\mathcal{H}$ -class of $M(K)$.

Lemma 7.13. *Every group $\mathcal{H}$ -class of $M(K)^\beta$ divides some group $\mathcal{H}$ -class of $M(K)$.*

Proof. Let H be a group $\mathcal{H}$ -class of $M(K)^\beta$. We will analyse two cases and make use of Lemma 7.5 several times without explicitly mentioning it.

- **Case 1:** There is some $u \in \tilde{A}^+$ such that $\text{dom } \theta_{uu^{-1}} \subseteq Q$ and $H = H_{\theta_{uu^{-1}}}$.

Let $X = \text{dom } \theta_{uu^{-1}} = \text{dom } \theta_{u^{-1}u}$ and define

$$X' = \bigcap_{\substack{v \in \tilde{A}^* \\ \theta_v \in H}} \text{dom } \delta_{v\bar{v}\eta}.$$

Note that $X \subseteq X'$ because, given $q \in X$, we have $q \in \text{dom } \theta_v$ for every $v \in \tilde{A}^*$ satisfying $\theta_v \in H$, and thus $q \in \text{dom } \delta_{v\bar{v}\eta}$, which implies $q \in X'$. Furthermore, since $\mathcal{S}(K)$ is a finite automaton, there exists some $\ell \geq 0$ such that, for every $w \in \tilde{A}^*$, we can find a word $\tilde{w}$ satisfying $|\tilde{w}| \leq \ell$ and $\delta_w = \delta_{\tilde{w}}$. Hence, there is a finite amount of words $v_1, v_2, \dots, v_m \in \tilde{A}^*$ ($m \geq 0$) such that $\theta_{v_1}, \theta_{v_2}, \dots, \theta_{v_m} \in H$ and

$$X' = \bigcap_{i=1}^m \text{dom } \delta_{v_i\bar{v}_i\eta}.$$

Consider the element $e = \overline{v_1\eta} \overline{v_1\eta}^{-1} \dots \overline{v_m\eta} \overline{v_m\eta}^{-1}$. Then δ_e is an idempotent and $\text{dom } \delta_e = X'$. Now, consider the subgroup $H' \leq H_{\delta_e}$ given by

$$\delta_w \in H' \Leftrightarrow X\delta_w \subseteq X.$$

In fact, since $\delta_w \in H'$ is injective, we have $X\delta_w \subseteq X \Leftrightarrow X\delta_w = X$; and if δ_w satisfies this condition, then also $(X' \setminus X)\delta_w = X' \setminus X$, since $\text{dom } \delta_w = \text{dom } \delta_{w^{-1}}$. We claim that the function Φ given by

$$\begin{aligned}\Phi: H' &\longmapsto H \\ \delta_w &\longrightarrow \theta_{uu^{-1}}\theta_{w\rho}\end{aligned}$$

is a surjective group homomorphism.

First we show that Φ is well defined. Given $\delta_w \in H'$, we must prove that

$$\text{dom}(\theta_{uu^{-1}}\theta_{w\rho}) = X \quad (1)$$

$$\text{dom}(\theta_{w^{-1}\rho}\theta_{uu^{-1}}) = X. \quad (2)$$

We start with (1). It is clear that the left hand side of the equality is contained in X . Regarding the reverse inclusion, given $q \in X$, we have $q \in X' = \text{dom } \delta_w$, which implies $q \in \text{dom } \theta_{w\rho}$ and $q \in \text{dom}(\theta_{uu^{-1}}\theta_{w\rho})$. Now we focus on (2). If $q \in X$ then $q \in \text{dom } \delta_{w^{-1}}$, and thus $q \in \text{dom } \theta_{w^{-1}\rho}$. Consequently,

$$(q, w^{-1}\rho)\theta = (q, w^{-1})\delta \in X = \text{dom } \theta_{uu^{-1}}$$

because $X\delta_{w^{-1}} = X$. As for the other inclusion, if we take $q \in \text{dom}(\theta_{w^{-1}\rho}\theta_{uu^{-1}})$, then

$$q' = (q, w^{-1}\rho)\theta \in \text{dom } \theta_{uu^{-1}} = X \subseteq \text{dom } \delta_w.$$

Therefore, we can read w from q' in $\mathcal{S}(K)$, and if $(q', w)\delta = q''$, then $(q', w\rho)\theta = q''$ by Lemma 7.5. Since the automaton $\mathcal{S}(K)^\beta$ is inverse, the equality $q' = (q, w^{-1}\rho)\theta$ implies $q = (q', w\rho)\theta$. It follows that $(q', w)\delta = q \in X$, as $X\delta_w = X$. This completes the proof of (2).

In order for Φ to be indubitably well defined, we must also confirm that if $\delta_w = \delta_z \in H'$, then $\theta_{uu^{-1}}\theta_{w\rho} = \theta_{uu^{-1}}\theta_{z\rho}$. But we already know that

$$\text{dom}(\theta_{uu^{-1}}\theta_{w\rho}) = \text{dom}(\theta_{uu^{-1}}\theta_{z\rho}) = X;$$

and taking $q \in X \subseteq \text{dom } \delta_w = \text{dom } \delta_z$, we obtain

$$(q, uu^{-1}(w\rho))\theta = (q, w\rho)\theta = (q, w)\delta = (q, z)\delta = (q, z\rho)\theta = (q, uu^{-1}(z\rho))\theta,$$

which establishes the claim.

Next, we show that Φ is a homomorphism, which amounts to proving that

$$\theta_{uu^{-1}}\theta_{w\rho}\theta_{uu^{-1}}\theta_{z\rho} = \theta_{uu^{-1}}\theta_{w\rho}\theta_{z\rho}$$

for all $\delta_w, \delta_z \in H'$. Again, it suffices to check that these transformations have the same

domain. Indeed, if that is true, then for any vertex q in their domain, we will obtain

$$\begin{aligned} (q, uu^{-1}(w\rho)uu^{-1}(z\rho))\theta &= (((q, uu^{-1}(w\rho))\theta, uu^{-1}\theta, z\rho)\theta \\ &= ((q, uu^{-1}(w\rho))\theta, z\rho)\theta \\ &= (q, uu^{-1}(w\rho)(z\rho))\theta. \end{aligned}$$

It is clear that

$$\text{dom}(\theta_{uu^{-1}}\theta_{w\rho}\theta_{uu^{-1}}\theta_{z\rho}) \subseteq \text{dom}(\theta_{uu^{-1}}\theta_{w\rho}\theta_{z\rho}).$$

Regarding the other inclusion, let $q \in \text{dom}(\theta_{uu^{-1}}\theta_{w\rho}\theta_{z\rho})$. Then $q \in X \subseteq \text{dom} \delta_w$ and we can take $q' = (q, w\rho)\theta = (q, w)\delta \in X$, as $q \in X$ and $X\delta_w = X$. Since $q' \in \text{dom} \theta_{z\rho}$, we conclude that $q' \in \text{dom}(\theta_{uu^{-1}}\theta_{z\rho})$, whence Φ is indeed a group homomorphism.

Finally, we prove that Φ is surjective by showing that

$$\theta_v = \theta_{uu^{-1}}\theta_{(e\bar{v}\bar{\eta}e)\rho} = \Phi(\delta_{e\bar{v}\bar{\eta}e})$$

for every $\theta_v \in H$. Observe that $\delta_{e\bar{v}\bar{\eta}e} \in H_{\delta_e}$ because $\theta_{v\pm 1} \in H$, $\bar{v}\bar{\eta}^{-1} = \overline{(v\eta)^{-1}} = \overline{v^{-1}\eta}$ and $\text{dom} \delta_{e\bar{v}\bar{\eta}e} = X'$ by definition of X' and δ_e . Moreover, $\delta_{e\bar{v}\bar{\eta}e} \in H'$ since $q \in X$ implies

$$(q, e\bar{v}\bar{\eta}e)\delta = (q, \bar{v}\bar{\eta})\delta = (q, v)\theta \in X.$$

Hence, it remains to verify that $\theta_v = \theta_{uu^{-1}}\theta_{(e\bar{v}\bar{\eta}e)\rho}$. If these partial transformations have the same domain then, for any vertex q in their domain,

$$\begin{aligned} (q, uu^{-1}(e\bar{v}\bar{\eta}e)\rho)\theta &= (q, (e\rho)\bar{v}\bar{\eta}\rho(e\rho))\theta \\ &= (q, \bar{v}\bar{\eta}\rho)\theta \\ &= (q, v)\theta. \end{aligned}$$

Hence, it suffices to check that $\text{dom} \theta_v = \text{dom}(\theta_{uu^{-1}}\theta_{(e\bar{v}\bar{\eta}e)\rho})$. On the one hand, the equalities $\text{dom} \theta_v = X = \text{dom} \theta_{uu^{-1}}$ entail $\text{dom} \theta_v \supseteq \text{dom}(\theta_{uu^{-1}}\theta_{(e\bar{v}\bar{\eta}e)\rho})$. On the other hand, if $q \in \text{dom} \theta_v = \text{dom} \theta_{uu^{-1}}$, then $q \in \text{dom} \delta_e$, which implies that $q \in \text{dom} \theta_{e\rho}$. Also, $q \in \text{dom} \delta_{\bar{v}\bar{\eta}}$ yields $q \in \text{dom} \theta_{\bar{v}\bar{\eta}\rho}$. We conclude the proof by observing that $(q, \bar{v}\bar{\eta}\rho)\theta = (q, v)\theta \in X \subseteq \text{dom} \delta_e \subseteq \text{dom} \theta_{e\rho}$.

- **Case 2:** There is some $u \in \tilde{A}^+$ such that $\text{dom} \theta_{uu^{-1}} \not\subseteq Q$ and $H = H_{\theta_{uu^{-1}}}$.

We begin by recalling an important fact. Given a blue vertex $q \in Q^\beta \setminus Q$ in $\mathcal{S}(K)^\beta$ and a word $w \in \tilde{A}^*$, if $q \in \text{dom} \theta_w$, then the first letter of w must be a^{-1} or b ; and if $q \in \text{im} \theta_w$, then the last letter of w must be a or b^{-1} .

In this case, if $v \in \tilde{A}^*$ is such that $\theta_v \in H$, then the equalities $\theta_a = \theta_{abb^{-1}}$ and $\theta_{a^{-1}} = \theta_{bb^{-1}a^{-1}}$ allow us to write $\theta_v = \theta_{b\tilde{v}b^{-1}}$ for some $\tilde{v} \in \tilde{A}^*$; in particular, $\theta_u = \theta_{b\tilde{u}b^{-1}}$. Now, since

$$\theta_{b\tilde{u}b^{-1}b\tilde{u}^{-1}b^{-1}} \mathcal{R} \theta_{(b\tilde{u}b^{-1}b\tilde{u}^{-1}b^{-1})b} \mathcal{L} \theta_{b^{-1}(b\tilde{u}b^{-1}b\tilde{u}^{-1}b^{-1})b},$$

we get $\theta_{uu^{-1}} \mathcal{D} \theta_{(b^{-1}b\bar{u}b^{-1})(b^{-1}b\bar{u}b^{-1})^{-1}}$. In view of Proposition [2.17](#), we conclude that

$$H \cong H_{\theta_{b^{-1}b\bar{u}bb^{-1}\bar{u}^{-1}b^{-1}b}}.$$

Therefore, this case can be reduced to the previous one.

Observe that above we have covered all possibilities, since H_{θ_1} is either trivial (and the claim automatically follows), or there exists $x \in A$ satisfying $\theta_1 = \theta_{xx^{-1}}$ (and we fall into the previous cases). $\square$

We need one last lemma before proving the main theorem of this section.

Lemma 7.14. *Let $\mathcal{A}' = (P', A, \lambda', q_0)$ be an inverse automaton with a basepoint and let $\mathcal{A} = (P, A, \lambda, q_0)$ be a subautomaton obtained by deleting a vertex of outdegree 1 which is not the basepoint. Then every group $\mathcal{H}$ -class of $M(\mathcal{A})$ divides some group $\mathcal{H}$ -class of $M(\mathcal{A}')$.*

Proof. Let $P' \setminus P = \{p'\}$ and $p \xrightarrow{x} p'$, $p' \xrightarrow{x^{-1}} p$ ($p \in P$, $x \in \tilde{A}$) be the extra edges in $\mathcal{A}'$. Let also $u \in \tilde{A}^+$ be such that $H = H_{\lambda_{uu^{-1}}}$ is a group $\mathcal{H}$ -class of $\mathcal{A}$. We define the sets $Y = \text{dom } \lambda_{uu^{-1}} = \text{dom } \lambda_{u^{-1}u}$ and

$$Y' = \bigcap_{\substack{v \in \tilde{A}^* \\ \lambda_v \in H}} \text{dom } \lambda'_v.$$

It is clear that $Y \subseteq Y'$, and we can choose a finite amount of words $v_1, v_2, \dots, v_n \in \tilde{A}^*$ ($n \geq 0$) such that $\lambda_{v_1}, \lambda_{v_2}, \dots, \lambda_{v_n} \in H$ and $Y' = \bigcap_{i=1}^n \text{dom } \lambda'_{v_i}$. Consider the word

$$e = v_1 v_1^{-1} v_2 v_2^{-1} \dots v_n v_n^{-1} \in \tilde{A}^*.$$

Then λ'_e is an idempotent and $\text{dom } \lambda'_e = Y'$. As in the previous lemma, we split the proof in two cases and argue in a similar manner.

- **Case 1:** $p' \notin Y'$.

Let $H' \leq H_{\lambda'_e}$ be the subgroup defined by

$$\lambda'_w \in H' \Leftrightarrow Y \lambda'_w = Y.$$

We claim that the function Ψ given by

$$\begin{aligned} \Psi: H' &\longrightarrow H \\ \lambda'_w &\longrightarrow \lambda_{uu^{-1}} \lambda_{\bar{w}} \end{aligned}$$

is a surjective group homomorphism.

To confirm that Ψ is well defined, we first prove that, for any $\lambda'_w \in H'$,

$$\text{dom } (\lambda_{uu^{-1}} \lambda_{\bar{w}}) = Y \tag{3}$$

$$\text{dom}(\lambda_{\bar{w}^{-1}}\lambda_{uu^{-1}}) = Y. \quad (4)$$

Starting with (3), it is clear $\text{dom} \lambda_{uu^{-1}}\lambda_{\bar{w}} \subseteq Y$, so let us take $q \in Y$ and focus on the other inclusion. If $\lambda'_w \in H_{\lambda'_e}$ then, when we read w from q in $\mathcal{A}'$, whenever we visit vertex p' , we must visit vertex p immediately next, since $p' \notin Y'$ has outdegree one. That will produce a subpath of the form $p \xrightarrow{x} p' \xrightarrow{x^{-1}} p$ associated to a factor xx^{-1} of the word w . Consequently, by construction of $\mathcal{A}$, we will be able to read the word $\bar{w}$ from q in $\mathcal{A}$, as any “forbidden” subpath labeled by xx^{-1} will be avoided. Therefore,

$$\text{im} \lambda_{uu^{-1}} = Y \subseteq Y' \subseteq \text{dom} \lambda_{\bar{w}}, \quad (*)$$

as we needed. For (4), let $q \in \text{dom} \lambda_{\bar{w}^{-1}}\lambda_{uu^{-1}}$ and $q' = (q, \bar{w}^{-1})\lambda = (q, \bar{w}^{-1})\lambda'$. Since $q' \in \text{dom} \lambda_{uu^{-1}} = Y$ and $Y\lambda'_{\bar{w}} = Y\lambda'_w = Y$, it follows that $q \in Y$. Regarding the other inclusion, if $q \in Y$, then $q \in Y' \subseteq \text{dom} \lambda_{\bar{w}^{-1}}$ by (*). Finally,

$$q' = (q, \bar{w}^{-1})\lambda = (q, w^{-1})\lambda' \in Y = \text{dom} \lambda_{uu^{-1}}$$

because $Y\lambda'_{w^{-1}} = Y$. In order for Ψ to be well defined, it remains to prove that $\lambda'_w = \lambda'_z \in H'$ implies $\lambda_{uu^{-1}}\lambda_{\bar{w}} = \lambda_{uu^{-1}}\lambda_{\bar{z}}$. But we have just seen that

$$\text{dom}(\lambda_{uu^{-1}}\lambda_{\bar{w}}) = \text{dom}(\lambda_{uu^{-1}}\lambda_{\bar{z}}) = Y$$

and, taking any $q \in Y$, we obtain

$$\begin{aligned} (q, uu^{-1}\bar{w})\lambda &= (q, uu^{-1}\bar{w})\lambda' = (q, \bar{w})\lambda' = \\ &= (q, w)\lambda' = (q, z)\lambda' = (q, \bar{z})\lambda' = \\ &= (q, uu^{-1}\bar{z})\lambda' = (q, uu^{-1}\bar{z})\lambda, \end{aligned}$$

so the claim follows.

To prove that Ψ is a homomorphism, we must show that

$$\lambda_{uu^{-1}}\lambda_{\bar{w}}\lambda_{uu^{-1}}\lambda_{\bar{z}} = \lambda_{uu^{-1}}\lambda_{\bar{w}\bar{z}}$$

for all $\lambda'_w, \lambda'_z \in H'$, and it is enough to verify that these transformations have equal domains. On the one hand, it is immediate that

$$\text{dom}(\lambda_{uu^{-1}}\lambda_{\bar{w}}\lambda_{uu^{-1}}\lambda_{\bar{z}}) \subseteq \text{dom}(\lambda_{uu^{-1}}\lambda_{\bar{w}\bar{z}}).$$

As for the other inclusion, let $q \in \text{dom}(\lambda_{uu^{-1}}\lambda_{\bar{w}\bar{z}})$. Then

$$q \in Y \subseteq \text{dom} \lambda'_w = Y' \subseteq \text{dom} \lambda_{\bar{w}}$$

by (*), and we can write $q' = (q, \bar{w})\lambda = (q, \bar{w})\lambda' = (q, w)\lambda' \in Y$, as $q \in Y$ and $Y\lambda'_w = Y$. Since we have $q' \in Y \subseteq \text{dom} \lambda_{\bar{z}}$ by (*), we get $q' \in \text{dom}(\lambda_{uu^{-1}}\lambda_{\bar{z}})$, and the

conclusion follows.

To finish, we prove that Ψ is surjective by showing that

$$\lambda_v = \lambda_{uu^{-1}} \lambda_{\overline{eve}} = \Psi(\lambda'_{eve})$$

for every $\lambda_v \in H$. Observe that $\lambda'_{eve} \in H_{\lambda_e}$ since $\lambda_{v\pm 1} \in H$ and $\text{dom } \lambda'_{ev\pm 1e} = Y'$, by definition of Y' and λ'_e . Furthermore, $\lambda'_{eve} \in H'$ because $q \in Y$ entails

$$(q, eve)\lambda' = (q, v)\lambda' = (q, v)\lambda \in Y.$$

Thus, it remains to check that λ_v and $\lambda_{uu^{-1}} \lambda_{\overline{eve}}$ have the same domain. On the one hand, the equalities $\text{dom } \lambda_v = Y = \text{dom } \lambda_{uu^{-1}}$ yield $\text{dom } (\lambda_{uu^{-1}} \lambda_{\overline{eve}}) \subseteq \text{dom } \lambda_v$. For the opposite inclusion, given $q \in \text{dom } \lambda_v$, we have $q \in \text{dom } \lambda_{uu^{-1}} = Y = \text{dom } \lambda_e$, so $(q, uu^{-1})\lambda = q \in \text{dom } \lambda_e$ and $(q, uu^{-1}e)\lambda = q \in \text{dom } \lambda_v$. Finally, note that $(q, uu^{-1}ev)\lambda = (q, v)\lambda \in Y = \text{dom } \lambda_e$, and hence

$$q \in \text{dom } (\lambda_{uu^{-1}} \lambda_{eve}) \subseteq \text{dom } (\lambda_{uu^{-1}} \lambda_{\overline{eve}}),$$

which completes the proof.

- **Case 2:** We have $p' \in Y'$.

In this case, if $v \in \tilde{A}^+$ is such that $\lambda_v \in H$ then, since $p' \in \text{dom } \lambda'_v \cap \text{dom } \lambda'_{v^{-1}}$ has outdegree 1, we must have $v = x^{-1}\tilde{v}x$ for some $\tilde{v} \in \tilde{A}^*$. In particular, $u = x^{-1}\tilde{u}x$ for some $\tilde{u} \in \tilde{A}^*$. Similarly to what we obtained in Case 2 of Lemma 7.13, it follows that

$$\lambda_{x^{-1}\tilde{u}xx^{-1}\tilde{u}^{-1}x} \mathcal{R} \lambda_{(x^{-1}\tilde{u}xx^{-1}\tilde{u}^{-1}x)x^{-1}} \mathcal{L} \theta_{x(x^{-1}\tilde{u}xx^{-1}\tilde{u}^{-1}x)x^{-1}},$$

which implies $\lambda_{uu^{-1}} \mathcal{D} \lambda_{(xx^{-1}\tilde{u}x)(xx^{-1}\tilde{u}x)^{-1}}$. Now Proposition 2.17 allows us to conclude that

$$H \cong H_{\lambda_{xx^{-1}\tilde{u}x^{-1}x\tilde{u}^{-1}xx^{-1}}}$$

and this case can be reduced to the previous one, as we cannot read a word starting with x from p' .

Note that we have covered all possibilities above, as H_{λ_1} is either trivial or we can write $\lambda_1 = \lambda_{yy^{-1}}$ for some $y \in \tilde{A} \setminus \{x\}$, falling into the previous cases. $\square$

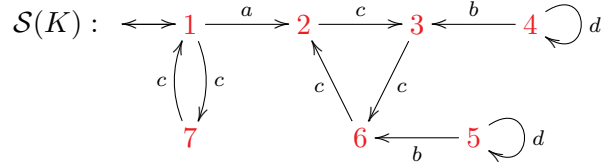
As we have remarked before, the automaton $\mathcal{S}(K\beta) = (Q', A, \gamma, q_0)$ can be obtained from $\mathcal{S}(K)^\beta = (Q^\beta, A, \theta, q_0)$ by successively deleting all vertices in $Q^\beta \setminus \{q_0\}$ with outdegree 1. Hence, in view of Lemma 7.14, it follows that every group $\mathcal{H}$ -class of $M(K\beta)$ divides some group $\mathcal{H}$ -class of $M(K)^\beta$. Applying then Lemma 7.13, we conclude that every group $\mathcal{H}$ -class of $M(K\beta)$ divides some group $\mathcal{H}$ -class of $M(K)$. Clearly, a similar result can be deduced for $M(K\beta^{-1})$ and $M(K)$, by invoking Lemma 7.6 instead of Lemma 7.5 throughout the proof of Lemma 7.13.

Since pseudovarieties are closed under division, the above discussion, together with Lemma 7.2 and Lemma 7.3, yields the following result.

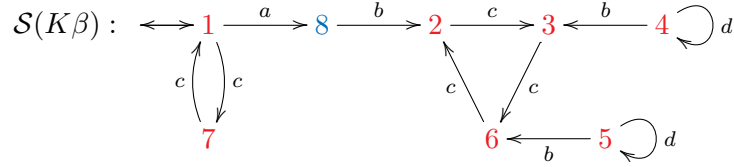
Theorem 7.15. *Let $\mathbf{H}$ be a pseudovariety of groups. Then the pseudovariety of monoids $\overline{\mathbf{H}}$ is stable under all automorphisms of F_A .*

Observe that if $\mathbf{H}$ is the trivial pseudovariety (i.e. the pseudovariety of groups consisting of only trivial groups), then $\overline{\mathbf{H}}$ is precisely the pseudovariety of aperiodic monoids, in view of Proposition 5.5. And if $\mathbf{H}$ is the pseudovariety of p -groups, then $\overline{\mathbf{H}}$ is the pseudovariety of monoids all of whose subgroups are p -groups. Therefore, Theorem 7.15 generalizes the results obtained in [3].

Example 7.16. Let $A = \{a, b, c, d\}$ and let $K \leq_{f.g.} F_A$ be given by the Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, 1)$ depicted by



Then $\mathcal{S}(K\beta) = \mathcal{S}(K)^\beta = (Q', A, \gamma, 1)$ can be portrayed by



Let $u = b^{-1}bc$. Using the notation in the proof of Lemma 7.13, consider the group $\mathcal{H}$ -class

$$H = H_{\gamma_{uu^{-1}}} = \{\gamma_u, \gamma_{u^2}, \gamma_{u^3}\} \cong \mathbb{Z}_3.$$

of $M(K\beta)$. Then we have

$$X = \text{dom } \gamma_{uu^{-1}} = \{2, 3, 6\}$$

$$X' = \text{dom } \delta_{\overline{u\eta}} \cap \text{dom } \delta_{\overline{u^2\eta}} \cap \text{dom } \delta_{\overline{u^3\eta}} = \text{dom } \delta_c \cap \text{dom } \delta_{c^2} \cap \text{dom } \delta_{c^3} = \{1, 2, 3, 6, 7\}.$$

Letting $e = cc^{-1}$, it follows that

$$H_{\delta_e} = H_{\delta_c} = \{\delta_c, \delta_{c^2}, \dots, \delta_{c^6}\} = H' \cong \mathbb{Z}_6$$

Hence, $H \cong \mathbb{Z}_3$ is a homomorphic image of $H' \cong \mathbb{Z}_6$.

7.3.1 Some characterizations

We proceed by characterizing finitely generated subgroups $K \leq_{f.g.} F_A$ satisfying $M(K) \in \overline{\mathbf{H}}$, for certain pseudovarieties of groups $\mathbf{H}$.

Let $S = (k_n)_{n \geq 1}$ be a sequence of positive integers and denote by $\mathbf{V}_S$ the pseudovariety of finite groups ultimately defined by the sequence $(x^{k_n} = 1)_n$. In other words, a finite group G belongs to $\mathbf{V}_S$ if and only if there exists $p \geq 1$ such that G satisfies the identity $x^{k_n} = 1$ for all $n \geq p$. Observe that, since an identity of the type $x_n = y_n$ in a group is equivalent to $x_n y_n^{-1} = 1$, any pseudovariety defined by a sequence encompassing a single letter is of the type $\mathbf{V}_S$ for some suitable S . Given two integers $m, n \geq 1$, we denote by (m, n) the greatest common divisor of m and n .

Theorem 7.17. *Let $K \leq_{f.g.} F_A$ have Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$, and let $S = (k_n)_{n \geq 1}$ be a sequence of positive integers. The following conditions are equivalent:*

- (i) $M(K) \in \overline{\mathbf{V}}_S$.
- (ii) $\exists p \geq 1, \forall q \in Q, \forall v \in \tilde{A}^*, \forall n \geq 1, \forall m \geq p \quad (q, v^n)\delta = q \Rightarrow (q, v^{(n, k_m)})\delta = q$.
- (iii) $\exists p \geq 1, \forall x \in F_A, \forall n \geq 1, \forall m \geq p \quad x^n \in K \Rightarrow x^{(n, k_m)} \in K$.

Proof. (i) $\Rightarrow$ (ii): Assume that (i) holds. Let $q \in Q$, $v \in \tilde{A}^*$ and $n \geq 1$ be such that $(q, v^n)\delta = q$. We define

$$O(v) = \bigcap_{m \geq 1} \text{dom } \delta_{v^m}$$

$$\ell = \min\{t \geq 1 \mid \forall o \in O(v) \ (o, v^t)\delta = o\},$$

and choose $s \geq 1$ satisfying

$$\forall o \in \text{dom } \delta_v \setminus O(v) \ (o, v^s)\delta \notin Q.$$

Then we define $r = \ell s + 1$ and $w = v^r$. Since we can only read w from vertices lying in $O(v)$, we deduce that δ_w belongs to a group $\mathcal{H}$ -class of $M(K)$. Moreover, $q \in O(v)$ implies

$$(q, w)\delta = (q, (v^\ell)^s v)\delta = (q, v)\delta,$$

so $(q, w^i)\delta = (q, v^i)\delta$ for all $i \geq 1$. If $O(v) = \{o_1, \dots, o_m\}$ ($m \geq 1$) and d is the order of δ_w , then δ_{w^d} is the restriction of the identity mapping fixing precisely $o_1, \dots, o_m$. Given $1 \leq i \leq m$, there is a smallest $d_i \geq 1$ such that $(o_i, v^{d_i})\delta = o_i$, so we get disjoint cycles

$$o_i = o_{i,0} \xrightarrow{v} o_{i,1} \xrightarrow{v} \dots \xrightarrow{v} o_{i,d_i-1} \xrightarrow{v} o_{i,0}$$

in $\mathcal{S}(K)$, where the vertices $o_{i,0}, o_{i,1}, \dots, o_{i,d_i-1} \in Q$ are all distinct. It is clear that d is the least common multiple of $d_1, \dots, d_m$; in particular, $d_i \mid d$ for all i . Assuming, without loss of generality, that $q = o_1$, we must have $d_1 \mid n$, since $(o_1, w^n)\delta = o_1 = (o_1, w^{d_1})\delta$. Furthermore, by hypothesis, there exists $p \geq 1$ such that $\delta_w^{k_m}$ is an idempotent for all $m \geq p$. This implies that $d_1 \mid k_m$ if $m \geq p$ because $d_1 \mid d$ and $d \mid k_m$. Thus, for all $m \geq p$, we obtain $d_1 \mid (n, k_m)$ and $(q, v^{(n, k_m)})\delta = (q, w^{(n, k_m)})\delta = q$, so (ii) follows.

(ii) $\Rightarrow$ (iii): Take p satisfying

$$\forall q \in Q, \forall v \in \tilde{A}^*, \forall n \geq 1, \forall m \geq p \quad (q, v^n) \delta = q \Rightarrow (q, v^{(n, k_m)}) \delta = q.$$

Let $x \in F_A$ and $n \geq 1$ be such that $x^n \in K$. We can assume that x is reduced, seen as an element of $\tilde{A}^*$, and write $x = u w u^{-1}$ for some $u, w \in \tilde{A}$ with w cyclically reduced. Then $(q_0, \overline{x^n}) \delta = (q_0, u w^n u^{-1}) \delta = q_0$. Letting $q = (q_0, u) \delta$, it follows that $(q, w^n) \delta = q$. Hence, $(q, w^{(n, k_m)}) \delta = q$ for all $m \geq p$, by assumption. This entails

$$q_0 = (q_0, u w^{(n, k_m)} u^{-1}) \delta = (q_0, \overline{x^{(n, k_m)}}) \delta,$$

so $x^{(n, k_m)} \in K$ for all $m \geq p$, and (iii) follows.

(iii) $\Rightarrow$ (ii): Take $p \geq 1$ satisfying

$$\forall x \in F_A, \forall n \geq 1, \forall m \geq p \quad x^n \in K \Rightarrow x^{(n, k_m)} \in K.$$

Let also $q \in Q$, $v \in \tilde{A}^*$ and $n \geq 1$ be such that $(q, v^n) \delta = q$. Take a word $u \in \tilde{A}^*$ satisfying $(q_0, u) \delta = q$. Then $(q_0, u v^n u^{-1}) \delta = q_0$ implies $(u v u^{-1})^n \in K$, and hence $(u v u^{-1})^{(n, k_m)} \in K$ for all $m \geq p$. Consequently, $(q_0, u v^{(n, k_m)} u^{-1}) \delta = q_0$, as we can surely read both $u v^{(n, k_m)}$ and $(v u^{-1})^{-1} = u v^{-1}$ from q_0 . It follows immediately that $(q, v^{(n, k_m)}) \delta = q$ for all $m \geq p$, as we wanted.

(ii) $\Rightarrow$ (i): Take p satisfying

$$\forall q \in Q, \forall v \in \tilde{A}^*, \forall n \geq 1, \forall m \geq p \quad (q, v^n) \delta = q \Rightarrow (q, v^{(n, k_m)}) \delta = q.$$

Let $H = H_{\delta_v}$ be a group $\mathcal{H}$ -class. If d is the order of δ_v in H then, for all $q \in \text{dom } \delta_v$, we obtain $(q, v^d) \delta = q$. It follows that $(q, v^{(d, k_m)}) \delta = q$ for all $m \geq p$, by hypothesis. Since $(d, k_m) \leq d$, we obtain that $d = (d, k_m)$, and thus $d \mid k_m$ if $m \geq p$. We conclude that the order of every element in H divides all the k_m for $m \geq p$, thereby showing that $H \in \mathbf{V}_S$. $\square$

Let $k \geq 1$ and denote by $\mathbf{B}_k$ the **Burnside pseudovariety** given by $\mathbf{B}_k = \llbracket x^k = 1 \rrbracket$. Considering the constant sequence $k_n = k$, we get the following corollary.

Corollary 7.18. *Let $K \leq_{f.g.} F_A$ have Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$. The following conditions are equivalent:*

- (i) $M(K) \in \overline{\mathbf{B}}_k$.
- (ii) $\forall q \in Q, \forall v \in \tilde{A}^*, \forall n \geq 1 \quad (q, v^n) \delta = q \Rightarrow (q, v^{(k, n)}) \delta = q$.
- (iii) $\forall x \in F_A, \forall n \geq 1 \quad x^n \in K \Rightarrow x^{(k, n)} \in K$.

Let π be a set of prime numbers and π' its complement in the set of primes. A π -number is a number all of whose prime factors lie in π . Out of convenience, we will also consider 1 to be a π -number, for every choice of π . A finite group G is called a π -**group** if its order factors into primes from π . By Lagrange and Cauchy theorems, this is equivalent to saying that

the order of every element in G is a π -number. We denote by $\mathbf{G}_\pi$ the pseudovariety of all finite π -groups. Observe that if $\pi = \{p_1, p_2, \dots\}$ and $S = \left((p_1 p_2 \cdots p_n)^{n!}\right)_n$, then $\mathbf{G}_\pi = \mathbf{V}_S$ [21, Proposition 7.1.16]. Hence, by Theorem 7.17, we get a characterization of subgroups $K \leq_{f.g.} F_A$ satisfying $M(K) \in \overline{\mathbf{G}}_\pi$. However, in this case, we can prove a somewhat simpler statement, using essentially the same ideas.

Corollary 7.19. *Let $K \leq_{f.g.} F_A$ with Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$. Let also π be a set of prime numbers with complement π' in the set of primes. Denote by Π and Π' the sets of π -numbers and π' -numbers, respectively. The following conditions are equivalent:*

- (i) $M(K) \in \overline{\mathbf{G}}_\pi$.
- (ii) $\forall q \in Q, \forall v \in \tilde{A}^*, \forall n \in \Pi' \quad (q, v^n)\delta = q \Rightarrow (q, v)\delta = q$.
- (iii) $\forall x \in F_A, \forall n \in \Pi' \quad x^n \in K \Rightarrow x \in K$.

Proof. (i) $\Rightarrow$ (ii): Assume that (i) holds. Let $q \in Q$, $v \in \tilde{A}^*$ and $n \in \Pi'$ be such that $(q, v^n)\delta = q$. Let $O(v) = \{o_1, \dots, o_m\}$, $w \in \tilde{A}^*$, $d \geq 1$ and d_i ($1 \leq i \leq m$) be as in the proof of Theorem 7.17. Once again, we must have $d_1 \mid n$, implying that $d_1 \in \Pi'$, and $d_1 \in \Pi$ because $d_1 \mid d$ and $d \in \Pi$. It follows that $d_1 = 1$ and $(q, v)\delta = (q, w)\delta = q$, so (ii) holds.

(ii) $\Rightarrow$ (iii): Let $x \in F_A$ and $n \in \Pi'$ be such that $x^n \in K$, and assume that x is a reduced word over $\tilde{A}^*$. Writing $x = u w u^{-1}$ for some $u, w \in \tilde{A}$ with w cyclically reduced, we get

$$(q_0, \overline{x^n})\delta = (q_0, u w^n u^{-1})\delta = q_0.$$

If $q = (q_0, u)\delta$, then $(q, w^n)\delta = q$, and thus $(q, w)\delta = q$ by assumption. We conclude that

$$q_0 = (q_0, u w u^{-1})\delta = (q_0, x)\delta,$$

whence $x \in K$.

(iii) $\Rightarrow$ (ii): Let $q \in Q$, $v \in \tilde{A}^*$ and $n \in \Pi'$ be such that $(q, v^n)\delta = q$. Picking a word $u \in \tilde{A}^*$ satisfying $(q_0, u)\delta = q$, we get $(q_0, u v^n u^{-1})\delta = q_0$, which implies $(u v u^{-1})^n \in K$, and thus $u v u^{-1} \in K$. The equalities $\delta(q, u^{-1} \overline{u v u^{-1}} u) = q$ and $\overline{u^{-1} u v u^{-1}} u = \bar{v}$ entail $(q, v)\delta = q$, as we wanted.

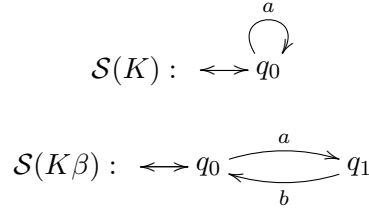
(ii) $\Rightarrow$ (i): Let $H = H_{\delta_v}$ be a group $\mathcal{H}$ -class, and let d be the order of δ_v in H . We can write $d = ef$ for some unique $e \in \Pi$, $f \in \Pi'$. If $w = v^e$, then $\delta_w \in H$ has order $f \in \Pi'$. Hence, for every $q \in \text{dom } \delta_w$, we obtain $(q, w^f)\delta = q$, which implies $(q, w)\delta = q$, by assumption. This means that w labels a loop at every vertex in $\text{dom } \delta_w$, so δ_w has order 1. Since δ_w has order f , it follows that $f = 1$, and the order of δ_v is simply $d = e \in \Pi$. We conclude that the order of every element in H is a π -number, thereby showing that $H \in \mathbf{G}_\pi$. $\square$

Again, note that these characterizations are compatible with the ones obtained in [3] in the particular cases of $M(K)$ being aperiodic or p -periodic (for some prime number p).

7.4 Non-examples

In this section we present some examples of pseudovarieties of monoids that do not satisfy the property of being stable under all automorphisms of a free group.

Denote by Inv the class of finite inverse monoids and let **Sl**, **J**, **R**, **L**, **Com** and **CR** be the pseudovarieties of semilattices, $\mathcal{J}$ -trivial, $\mathcal{R}$ -trivial, $\mathcal{L}$ -trivial, commutative and completely regular monoids, in that order. Let $A = \{a, b\}$, $\beta = \beta_{ab} \in \text{Aut}(F_A)$ and $K = \langle a \rangle \leq F_A$ be the subgroup of F_A generated by $a \in F_A$. Then the automata $\mathcal{S}(K) = (\{q_0\}, \{a, b\}, \delta, q_0)$ and $\mathcal{S}(K\beta) = (\{q_0, q_1\}, \{a, b\}, \gamma, q_0)$ are depicted by



Observe that $M(K) = \{\emptyset, 1\}$ and $M(K\beta) = \{(1\ 2), (1\ -), (-\ 2), (-\ 1), (2\ -), (-\ -)\} \simeq B_2^1$, where B_2^1 is the so-called **6-element Brandt monoid**. We have the following:

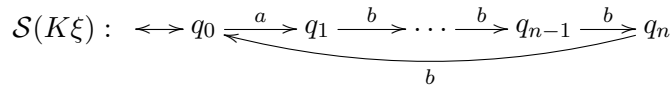
- $M(K\beta) \notin \mathbf{Sl}$, for γ_a is not an idempotent; however, $M(K) \in \mathbf{Sl}$.
- $M(K\beta) \notin \mathbf{Com}$, since $\gamma_a \gamma_b \neq \gamma_b \gamma_a$, but $M(K) \in \mathbf{Com}$.
- $M(K\beta) \notin \mathbf{CR} = \llbracket x^{\omega+1} = x \rrbracket$, since $\gamma_a^\omega = \emptyset \neq \gamma_a$; in contrast, $M(K) \in \mathbf{CR}$.

These observations allow us to conclude that the pseudovarieties **Sl** (and, consequently, **R**, **L** and **J**, in view of Proposition [5.13](#)), **Com** and **CR** are not stable under $\beta \in \text{Aut}(F_A)$.

Moreover, despite the fact that the pseudovariety $\mathbf{A} = \llbracket x^{\omega+1} = x^\omega \rrbracket$ of aperiodic monoids is stable under all automorphisms of F_A , the pseudovariety $\mathbf{A}_n = \llbracket x^{n+1} = x^n \rrbracket$ is not, for any $n \geq 1$. In fact, if $\xi \in \text{Aut}(F_A)$ is the automorphism given by

$$\begin{aligned} \xi : F_A &\longrightarrow F_A \\ a &\longmapsto ab^n \\ b &\longmapsto b, \end{aligned}$$

then $M(K\xi) \notin \mathbf{A}_n$. To see why, observe that $\mathcal{S}(K\xi) = (\{q_0, \dots, q_n\}, \{a, b\}, \theta, q_0)$ can be depicted by



So $\theta_b^{n+1} \neq \theta_b^n$, which shows that $M(K\xi) \notin \mathbf{A}_n$, even though $M(K) \in \mathbf{A}_n$.

7.5 Conjugacy conditions

In this section, we characterize normal, malnormal and cyclonormal subgroups of F_A using the transition monoid of the Stallings automaton.

Before we characterize the normal subgroups of F_A , we introduce the concept of a group action. Let X be a nonempty finite set and S_X be the symmetric group over X . Given a group G , a **right action** of G on X is a function $\varphi: X \times G \rightarrow X$ satisfying $(x, 1)\varphi = x$ and $(x, gh)\varphi = ((x, g)\varphi, h)\varphi$, for all $x \in X$ and $g, h \in G$. If $G \leq S_X$, then its elements are permutations on X , hence we get an action of G on X given by $(x, g) \mapsto xg$. For each $x \in X$, the **orbit** of x is $xG = \{xg \mid g \in G\}$ and the **stabilizer** of x is $G_x = \{g \in G \mid xg = x\} \leq G$. One easily checks that the orbits constitute a partition of X . We say that the action is **transitive** if it determines a single orbit. It is well known that $|xG| = [G : G_x]$, from which we derive the following result.

Lemma 7.20. *Let G be a finite group equipped with a transitive right action over a nonempty set X . The following conditions are equivalent:*

- (i) $G_x = \{1\}$ for all $x \in X$.
- (ii) $G_x = \{1\}$ for some $x \in X$.
- (iii) $|G| = |X|$.

Proof. The implication (i) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (iii): Let $x \in X$ be such that $G_x = \{1\}$ and define a function

$$\begin{aligned} \varphi: G &\longrightarrow X \\ g &\longmapsto xg. \end{aligned}$$

Then φ is a surjection since the group action is transitive. It is also injective, for if $xg = xh$ then $xgh^{-1} = x$, so $gh^{-1} = 1$ by (ii), and thus $g = h$. We conclude that φ is a bijection, thereby showing that $|G| = |X|$.

(iii) $\Rightarrow$ (i): Let $g \in G$ and $x \in X$ be such that $xg = x$. Since $|G| = |X|$, the surjection φ defined above is also injective. As $1\varphi = g\varphi$, it follows that $g = 1$ and, consequently, that $G_x = \{1\}$ for all $x \in X$. $\square$

We know, by Corollary 6.17, that $1 \neq K \leq_{f.g.} F_A$ is normal in F_A if and only if it is a subgroup of finite index and $\mathcal{C}(K)$, the core of $\mathcal{S}(K)$, is vertex-transitive. We are ready to state the aforementioned characterization of normal subgroups.

Theorem 7.21. *Let $1 \neq K \leq_{f.g.} F_A$ and $M(K)$ be the transition monoid of its Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$. Then K is normal in F_A if and only if $M(K)$ is a group of size $|Q|$.*

Proof. Suppose first that $K \trianglelefteq F_A$. Since K has finite index, $\mathcal{S}(K)$ is complete and $M(K)$ is a group, by Corollary 7.8. Moreover, given $p, q \in Q$, there exists $u \in \tilde{A}^*$ such that $(p, u)\delta = q$, as $\mathcal{S}(K)$ is connected, and hence $M(K) \leq S_Q$ acts transitively on Q . Let $\delta_u \in M(K)$ and $q \in Q$ be such that $(q, u)\delta = q$. Take an arbitrary vertex $p \in Q$ and consider an automorphism of $\mathcal{C}(K)$ inducing a bijection $\varphi: Q \rightarrow Q$ that maps q to p . Then we have

$$(p, u)\delta = (q\varphi, u)\delta = ((q, u)\delta)\varphi = q\varphi = p,$$

so δ_u fixes p too. Hence, if $\delta_u \in M(K)$ is a transformation with a fixed point, it follows that $\delta_u = 1$. We conclude that $M(K)$ satisfies property (i) of Lemma 7.20, thereby showing that $|M(K)| = |Q|$, by property (iii).

Assume now that $M(K) \leq S_Q$ is a group and $|M(K)| = |Q|$. As we remarked above, $M(K)$ acts transitively on Q by connectedness of $\mathcal{S}(K)$ so, by Lemma 7.20, the stabilizer of any state is trivial. Given any $p, q \in Q$, our goal is to define a bijection $\varphi: Q \rightarrow Q$ such that $p\varphi = q$ and, for all $r \in Q$ and $v \in \tilde{A}$, the equality $(r\varphi, v)\delta = ((r, v)\delta)\varphi$ holds; this will show that $\mathcal{C}(K)$ is vertex-transitive. For every $r \in Q$, take $u_r \in \tilde{A}^*$ satisfying $(p, u_r)\delta = r$ and define

$$\begin{aligned}\varphi: Q &\longrightarrow Q \\ r &\longmapsto (q, u_r)\delta.\end{aligned}$$

To check that φ is well defined, observe that if $v \in \tilde{A}^*$ is such that $(p, v)\delta = r$, then $(p, u_r v^{-1})\delta = p$ and $\delta_{u_r v^{-1}} = 1$ by property (i) of Lemma 7.20. Hence, $\delta_{u_r} = \delta_v$ and, in particular, $(q, u_r)\delta = (q, v)\delta$. Similarly one checks that φ is an injection, for if $r = (p, u_r)\delta$ and $s = (p, v)\delta$ satisfy $r\varphi = s\varphi$, then $(q, u_r)\delta = (q, v)\delta$ implies $\delta_{u_r} = \delta_v$, and

$$r = (p, u_r)\delta = (p, v)\delta = s.$$

Since Q is finite, we deduce that φ is a bijection. Finally, note that for all $r \in Q$ and $v \in \tilde{A}$, we have

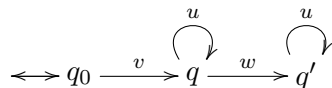
$$\begin{aligned}(r\varphi, v)\delta &= ((q, u_r)\delta, v)\delta \\ &= (q, u_r v)\delta \\ &= ((p, u_r v)\delta)\varphi \\ &= (((p, u_r)\delta, v)\delta)\varphi \\ &= ((r, v)\delta)\varphi.\end{aligned}$$

The desired conclusion now follows. $\square$

Given any group G and $K \leq G$, we say that K is **malnormal** if $gKg^{-1} \cap K = 1$ for all $g \in G \setminus K$. In the case of free groups, there is the following well-known result.

Lemma 7.22. *A subgroup $K \leq_{f.g.} F_A$ is malnormal if and only if, in its Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$, there do not exist two distinct vertices $q, q' \in Q$ and a nonempty reduced word u such that u labels a loop at both q and q' .*

Proof. We prove the contrapositive of the statement. Suppose first that $q, q' \in Q$ are distinct and $u \in \tilde{A}^+$ is a reduced word satisfying $(q, u)\delta = q$ and $(q', u)\delta = q'$. If $v, w \in \tilde{A}^*$ are such that $(q_0, v)\delta = q$ and $(q, w)\delta = q'$, then $vuv^{-1}, vwuw^{-1}v^{-1} \in K$. Visually, this phenomenon can be depicted as follows:



Letting $t = vw^{-1}v^{-1}$, we get

$$vuv^{-1} = t(vwuw^{-1}v^{-1})t^{-1} \in tKt^{-1}.$$

If $t \in K$, then $w \in v^{-1}Kv$ and, since $\Gamma(v^{-1}Kv)$ is a subgraph of $\Gamma(K)$, it follows that $(q, w)\delta = q$. Consequently, we get $q = q'$, which contradicts our original assumption, whence $t \notin K$ and K is not malnormal.

On the other hand, if K is not malnormal, let $v \in F_A \setminus K$ be such that $K \cap v^{-1}Kv \neq 1$, and assume that v is a reduced word seen as an element of $\tilde{A}^+$. Write $v^{-1} = wz$, where w is the smallest prefix of v^{-1} such that z^{-1} can be read in $\mathcal{S}(K)$ from q_0 . It follows from $wzKz^{-1}w^{-1} \cap K \neq 1$ that w can also be read from q_0 in $\mathcal{S}(K)$. Indeed, if we take $1 \neq y \in wzKz^{-1}w^{-1} \cap K$, say $y = wzzx^{-1}w^{-1}$ for some $x \in K$, then

$$\bar{y} = \overline{wzzx^{-1}w^{-1}} = \overline{wzzx^{-1}}w^{-1} \in L(\mathcal{S}(K)),$$

by minimality of w . Consequently, both $\mathcal{S}(w^{-1}Kw)$ and $\mathcal{S}(zKz^{-1})$ are subgraphs of $\mathcal{S}(K)$ with basepoints $q = (q_0, w)\delta$ and $q' = (q_0, z^{-1})\delta$ respectively. Now $zKz^{-1} = w^{-1}(v^{-1}Kv)w$ implies $zKz^{-1} \cap w^{-1}Kw \neq 1$. Hence, there exists a nonempty reduced word u such that u labels a loop at both q and q' , which are distinct vertices since $wz = v \notin L(\mathcal{S}(K))$. $\square$

We shall provide a characterization of malnormal subgroups of F_A using the transition monoids of Stallings automata. Before, we prove a simple lemma.

Lemma 7.23. *Let $K \leq_{f.g.} F_A$ and $x \in \tilde{A}^*$ be such that $\delta_x \in E(M(K))$. Then, for all $r \in \tilde{A}^*$ and $n \geq 1$, the condition $\delta_{rx^n r^{-1}} = \delta_{rxr^{-1}} \in E(M(K))$ holds.*

Proof. First observe that, if N is an inverse monoid, given $s \in N$ and $e \in E(N)$, we have $s^{-1}s \in E(N)$ and

$$\begin{aligned} (ses^{-1})(ses^{-1}) &= se(s^{-1}s)es^{-1} \\ &= see(s^{-1}s)s^{-1} \\ &= se(s^{-1}ss^{-1}) \\ &= ses^{-1}, \end{aligned}$$

because idempotents commute in N . It follows that $ses^{-1} \in E(N)$, which shows that idempotency is closed under conjugation in any inverse monoid. Since $M(K)$ is inverse and $\delta_{rxr^{-1}} = \delta_r \delta_x \delta_r^{-1}$, we conclude that $\delta_{rxr^{-1}} \in E(M(K))$.

To finish, given $n \geq 0$, it is enough to observe that

$$\begin{aligned} \delta_{rx^n r^{-1}} &= \delta_r \delta_x^n \delta_r^{-1} \\ &= \delta_r \delta_x \delta_r^{-1} \\ &= \delta_{rxr^{-1}} \end{aligned}$$

because, by assumption, δ_x is an idempotent. $\square$

Given a partially ordered set $(X, \leq)$ and $Y = \{y_1, \dots, y_n\} \subseteq X$ for some $n \geq 1$, we say that Y is a **chain** on X if there exists $\sigma \in S_n$ satisfying $y_{1\sigma} \leq \dots \leq y_{n\sigma}$. We say that a chain of the type $y_1 < y_2 < \dots < y_k$ ($k \geq 1$) has length k . We are ready to prove the theorem regarding the transition monoid of malnormal subgroups.

Theorem 7.24. *Given a nontrivial subgroup $K <_{f.g.} F_A$, consider the monoid homomorphism*

$$\begin{aligned} \Delta: \tilde{A}^* &\longrightarrow M(K) \\ u &\longmapsto \delta_u \end{aligned}$$

and let R_A be the set of all nonempty reduced words over $\tilde{A}$. Denote by E the set of idempotents of $R_A\Delta$. Consider the restriction of the natural partial order on $M(K)$ to E , and let k be the size of a maximal chain on E . Then K is malnormal if and only if $k = 2$ and $|E| = |Q| + 1$.

Proof. First observe that, by construction of $\mathcal{S}(K) = (Q, A, \delta, q_0)$, there exist a cyclically reduced word $u \in R_A$ and $q \in Q$ such that u labels a loop at q . Hence, such δ_u is not the empty transformation on Q . Since $M(K)$ is finite, there exists $n \geq 1$ such that $\delta_u^n \in E(R_A\Delta)$. So if $k = 1$, then $\emptyset \notin E(R_A\Delta)$ and, consequently, $\emptyset \notin M(K)$. This implies that $\mathcal{S}(K)$ is a complete automaton by Lemma 7.8, so $M(K)$ is a finite group. As $K \neq F_A$, it follows that $\mathcal{S}(K)$ has at least two vertices. Given $a \in A$, there exists some $n \geq 1$ such that $\delta_a^n = \text{id}_Q$, and clearly $a^n \in R_A$. Hence, if q and q' are two distinct vertices in $\mathcal{S}(K)$, the word a^n labels a loop at q and labels a loop at q' . This entails that K is not malnormal.

Now suppose that $k \geq 3$. Since every idempotent is a restriction of id_Q , there exist at least three distinct elements $\delta_u, \delta_v, \delta_w \in E(R_A\Delta)$ satisfying $\delta_u < \delta_v < \delta_w$. Accordingly, δ_w fixes at least two distinct vertices, say q and q' , which implies that $w \in R_A$ labels a loop both at q and q' . Hence, K is not malnormal.

All in all, if K is malnormal, then $k = 2$. Note that, in this case, $\emptyset \in E$; otherwise, by Lemma 7.7, $M(K)$ would be a group, and thus it would have a unique idempotent. So there exist distinct nonempty sets $Q_1, Q_2, \dots, Q_\ell \subseteq Q$ ($\ell \geq 1$) such that

$$E = \{\emptyset, \text{id}_{Q_1}, \dots, \text{id}_{Q_\ell}\},$$

yielding $|E| = \ell + 1$. Now, given any $q \in Q$, by construction of $\mathcal{S}(K)$ there exists some $u \in R_A$ satisfying $(q, u)\delta = q$, which means that δ_u fixes q . We can write $u = r x r^{-1}$ for some $r \in R_A \cup \{1\}$, so that $x \in R_A$ is cyclically reduced. As $M(K)$ is finite, there exists $n \geq 1$ for which $\delta_{u^n} = \delta_u^n$ is an idempotent. Hence, putting $v = r x^n r^{-1}$, we have $v \in R_A$ and $\delta_v \in E$, by Lemma 7.23 and δ_v also fixes q . This shows that $q \in Q_i$ for some $1 \leq i \leq \ell$, and therefore

$$Q = Q_1 \cup Q_2 \cup \dots \cup Q_\ell.$$

If $|Q| > 1$, malnormality of K implies that $|Q_i| = 1$, for $1 \leq i \leq \ell$, which gives $|Q| = \ell$ and $|E| = |Q| + 1$; if $|Q| = 1$, then $E = \{\emptyset, \text{id}_Q\}$ and $|E| = |Q| + 1$ too.

Finally, keeping the notation introduced in the preceding paragraph, suppose that $k = 2$ and $|E| = |Q| + 1$ (so $\ell = |Q|$). In order to prove that K is malnormal, we only need to show that, for all $1 \leq i, j \leq \ell$,

$$i \neq j \Rightarrow Q_i \cap Q_j = \emptyset.$$

Indeed, if that is the case, we get

$$|Q| = |Q_1| + |Q_2| + \cdots + |Q_\ell| \geq \ell.$$

So, under the hypothesis that $|Q| = \ell$, all the Q_i must be singletons, as they are nonempty. Hence, if K is not malnormal, then there exist distinct states $q, q' \in Q$ and $u \in R_A$ such that $(q, u)\delta = q$ and $(q', u)\delta = q'$. Writing $u = rxr^{-1}$ with x cyclically reduced, there exists $n \geq 1$ such that $v = rx^n r^{-1} \in R_A$ and $\delta_v = \text{id}_{Q_j} \in E$ for some $1 \leq j \leq \ell$. Since $q, q' \in Q_j$, it follows that $|Q_j| > 1$, which contradicts our assumption. Therefore, if all the Q_i ($1 \leq i \leq \ell$) are disjoint, then K is a malnormal subgroup of F_A .

We proceed by proving the disjointness claim by contradiction. Suppose that there exist $1 \leq i < j \leq \ell$ and $u, v \in R_A$ such that $\delta_u = \text{id}_{Q_i}$, $\delta_v = \text{id}_{Q_j}$ and $q \in Q_i \cap Q_j$. Note that, by assumption, $Q_i \neq Q_j$. As before, we can write $u = \tilde{u}x\tilde{u}^{-1}$ and $v = \tilde{v}y\tilde{v}^{-1}$, for some $\tilde{u}, \tilde{v} \in R_A \cup \{1\}$ and $x, y \in R_A$, so that both x and y are cyclically reduced words. Moreover, let $t \in R_A \cup \{1\}$ be the longest prefix shared by $\tilde{u}$ and $\tilde{v}$ and write $\tilde{u} = tr$, $\tilde{v} = ts$ for some appropriate $r, s \in R_A \cup \{1\}$. Then $u = trxr^{-1}t^{-1}$ and $v = tsys^{-1}t^{-1}$.

Since $M(K)$ is finite, there exists $n \geq 1$ such that $\delta_{x^n}, \delta_{y^n} \in E$, which implies

$$\delta_{rx^n r^{-1}}, \delta_{sy^n s^{-1}} \in E,$$

by Lemma [7.23](#). Letting $w = trx^n r^{-1}t^{-1}$ and $z = tsy^n s^{-1}t^{-1}$, we also have $\delta_w = \delta_u$ and $\delta_z = \delta_v$. Hence, we deduce that $\delta_w \neq \delta_z$ and $\delta_{rx^n r^{-1}} \neq \delta_{sy^n s^{-1}}$.

We now analyse three possible cases:

- **Case 1:** $r = 1$ and $s = 1$.

We have $u = txt^{-1}$ and $v = tyt^{-1}$, with $x \neq y$ nonempty cyclically reduced words. If $xy \notin R_A$, then we can write $x = \tilde{x}a$ and $y = a^{-1}\tilde{y}$ for some $a \in \tilde{A}$ and $\tilde{x}, \tilde{y} \in R_A \cup \{1\}$. In that case, the last letter of y cannot be a , and thus $xy^{-1} \in R_A$. So there exists $\varepsilon \in \{-1, 1\}$ such that $xy^\varepsilon \in R_A$. However, this yields a contradiction, for we would get a chain

$$\emptyset < \delta_{x^n y^{\varepsilon n}} < \delta_{x^n}$$

or

$$\emptyset < \delta_{x^n y^{\varepsilon n}} < \delta_{y^{\varepsilon n}}$$

in E , which violates the assumption that $k = 2$. Indeed, $\delta_{x^n y^{\varepsilon n}} \neq \emptyset$ because $x^n y^{\varepsilon n}$ labels a loop at $(q, t)\delta$, as x and y do so; and $\delta_{y^{\varepsilon n}} \neq \delta_{x^n}$ entails $\delta_{x^n y^{\varepsilon n}} < \delta_{x^n}$ or $\delta_{x^n y^{\varepsilon n}} < \delta_{y^{\varepsilon n}}$.

- **Case 2:** $r = 1$ and $s \neq 1$.

We have $u = txt^{-1}$ and $v = tsys^{-1}t^{-1}$, with $x \neq y$ nonempty cyclically reduced words. If $xsy^{-1} \notin R_A$, then we can write $x = \tilde{x}a$ and $s = a^{-1}\tilde{s}$ for some $a \in \tilde{A}$ and $\tilde{x}, \tilde{s} \in R_A \cup \{1\}$. In that case, the first letter of x cannot be a^{-1} , and thus $sys^{-1}x \in R_A$. Either way, we can use an argument similar to the above to get a chain

$$\begin{aligned} \emptyset < \delta_{x^n sy^n s^{-1}} < \delta_{x^n} \quad \text{or} \quad \emptyset < \delta_{x^n sy^n s^{-1}} < \delta_{sy^n s^{-1}} \\ \text{or} \\ \emptyset < \delta_{sy^n s^{-1} x^n} < \delta_{x^n} \quad \text{or} \quad \emptyset < \delta_{sy^n s^{-1} x^n} < \delta_{sy^n s^{-1}} \end{aligned}$$

in E , yielding a contradiction again. The case $r \neq 1$ and $s = 1$ is analogous.

- **Case 3:** $r \neq 1$ and $s \neq 1$.

We have $u = trxr^{-1}t^{-1}$ and $v = tsys^{-1}t^{-1}$, with $x \neq y$ nonempty cyclically reduced words. If $rxr^{-1}sys^{-1} \notin R_A$, then we can write $r = a\tilde{r}$ and $s = a\tilde{s}$ for some $a \in \tilde{A}$ and $\tilde{r}, \tilde{s} \in R_A \cup \{1\}$. However, that is not possible, for otherwise ta would be a longer prefix common to $\tilde{u}$ and $\tilde{v}$. So $rxr^{-1}sys^{-1} \in R_A$ and we get a chain

$$\begin{aligned} \emptyset < \delta_{rx^n r^{-1} sy^n s^{-1}} < \delta_{rx^n r^{-1}} \\ \text{or} \\ \emptyset < \delta_{rx^n r^{-1} sy^n s^{-1}} < \delta_{sy^n s^{-1}} \end{aligned}$$

in E , contradicting the assumption that $k = 2$.

In any case, we always arrive at a contradiction, which means that if $1 \leq i < j \leq \ell$, then $Q_i \cap Q_j = \emptyset$. This, as we remarked, concludes the proof of the theorem. $\square$

Example 7.25. Let $A = \{a, b, c, d\}$ and consider $K \leq_{f.g.} F_A$ given by

$$\mathcal{S}(K) : \quad \begin{array}{ccccccc} & & \overset{c}{\curvearrowright} & & & \overset{a}{\curvearrowright} & \\ & \longleftrightarrow & 1 & \xrightarrow{a,b} & 2 & \xrightarrow{d} & 3 & \xleftarrow{b,c} & 4 \end{array}$$

Using the notation of Theorem [7.24](#), we have a chain

$$\emptyset = \delta_{a^2} < (1 \text{ --- } -) = \delta_{c^2} < (1 \text{ --- } 3) = \delta_{ab^{-1}c}$$

in E with length 3, so we conclude that K is not a malnormal subgroup of F_A .

Given any group G and $K \leq G$, we say that K is **cyclonormal** if $gKg^{-1} \cap K$ is a cyclic group for every $g \in G \setminus K$. Before presenting some characterizations of cyclonormal subgroups of free groups, we introduce some notation. Let $K \leq_{f.g.} F_A$ with $\mathcal{S}(K) = (Q, A, \delta, q_0)$, and

let $p, q \in Q$. Denoting by $\phi: \tilde{A}^* \rightarrow F_A$ the quotient morphism, we define

$$L_{(p,q)} = \{u\phi \mid u \in \tilde{A}^* \text{ labels a loop at } (p, q) \text{ in } \mathcal{S}(K) \times \mathcal{S}(K)\} \leq F_A.$$

To simplify the notation, we write R_A instead of $\text{Irr}(\mathcal{R}_A)$. Note that if $|Q| = 1$, then K is cyclonormal; and if $|A| = 1$, then every subgroup of F_A is cyclonormal.

Lemma 7.26. *Let $K \leq_{f.g.} F_A$ be a subgroup of F_A and $\mathcal{S}(K) = (Q, A, \delta, q_0)$. The following conditions are equivalent:*

- (i) K is cyclonormal.
- (ii) For each pair of distinct vertices $(p, q) \in Q \times Q$, the subgroup $L_{(p,q)}$ is cyclic.
- (iii) Let $p, q \in Q$ be two distinct vertices. Then there exists $u \in R_A$ such that $(p, u)\delta = p$, $(q, u)\delta = q$, and for all $v \in R_A$ satisfying $(p, v)\delta = p$ and $(q, v)\delta = q$, the equality $v = \overline{u^m}$ holds for some $m \in \mathbb{Z}$.

Proof. (i) $\Leftrightarrow$ (ii): First we prove that if $v \in F_A \setminus K$ and $vKv^{-1} \cap K \neq 1$, then there exist distinct vertices $p, q \in Q$ such that $L_{(p,q)}$ and $vKv^{-1} \cap K$ are conjugate in F_A . As in the proof of Theorem 7.22, assume that v is a reduced word seen as an element of $\tilde{A}^+$. Write $v = wz$, where w is the smallest prefix of v such that z^{-1} can be read in $\mathcal{S}(K)$ from q_0 . Again, since $wzKz^{-1}w^{-1} \cap K \neq 1$, the minimality of w implies that w can also be read from q_0 in $\mathcal{S}(K)$. If $p = (q_0, w)\delta$ and $q = (q_0, z^{-1})\delta$, then $\mathcal{S}(w^{-1}Kw)$ and $\mathcal{S}(zKz^{-1})$ are subgraphs of $\mathcal{S}(K)$ with basepoints p and q , respectively. Note that $p \neq q$, for otherwise we would get a path

$$q_0 \xrightarrow{w} p = q \xrightarrow{z} q_0$$

in $\mathcal{S}(K)$, which contradicts the assumption that $v \notin K$. Now, since

$$w(w^{-1}Kw \cap zKz^{-1})w^{-1} = vKv^{-1} \cap K,$$

it follows that the subgroups $w^{-1}Kw \cap zKz^{-1}$ and $vKv^{-1} \cap K$ are conjugate in F_A . And, by the proof of Theorem 6.19, we know that

$$L_{(p,q)} = w^{-1}Kw \cap zKz^{-1} \cong vKv^{-1} \cap K.$$

So $L_{(p,q)}$ is cyclic if and only if $vKv^{-1} \cap K$ is cyclic.

On the other hand, given two distinct vertices $p, q \in Q$, we can take $w, z \in R_A$ such that $p = (q_0, w)\delta$ and $q = (q_0, z^{-1})\delta$. By a similar argument, we obtain

$$L_{(p,q)} = w^{-1}Kw \cap zKz^{-1} \cong wzKz^{-1}w^{-1} \cap K.$$

Now suppose that $wz \in K$. Then there exists $t \in K$ such that $tw = z^{-1}$ in F_A , and we can assume that t is reduced seen as a word in $\tilde{A}^*$. This implies that $(q_0, t)\delta = q_0$, and thus

$$p = (q_0, tw)\delta = (q_0, \overline{tw})\delta = (q_0, \overline{z^{-1}})\delta = (q_0, z^{-1})\delta = q,$$

which contradicts the assumption that $p \neq q$. Hence, we deduce that $wz \notin K$. We conclude that if $p, q \in Q$ are any two distinct vertices, then $L_{(p,q)}$ is isomorphic (conjugate) to a subgroup of the form $vKv^{-1} \cap K$ for some $v \in F_A \setminus K$.

Therefore, if K is cyclonormal and $(p, q) \in Q \times Q$ is a pair of distinct vertices, then $L_{(p,q)}$, being isomorphic to $vKv^{-1} \cap K$ for some $v \in F_A \setminus K$, must be cyclic. Conversely, suppose that $L_{(p,q)}$ is cyclic for every pair of distinct vertices $(p, q) \in Q \times Q$. Given $v \in F_A \setminus K$, either $vKv^{-1} \cap K = 1$ or $1 \neq vKv^{-1} \cap K \cong L_{(p,q)}$ for some distinct vertices $p, q \in Q$; in any case, $vKv^{-1} \cap K$ is cyclic.

(ii) $\Leftrightarrow$ (iii): We know that a word $w \in \tilde{A}^*$ labels a loop at (p, q) in $\mathcal{S}(K) \times \mathcal{S}(K)$ if and only if $(p, w)\delta = p$ and $(q, w)\delta = q$. Moreover, $L_{(p,q)}$ is a cyclic group if and only if there exists $u \in R_A$ such that $u\phi \in L_{(p,q)}$ and, for every $v \in \tilde{A}^*$ labelling a loop at (p, q) , the equality $v\phi = u^m\phi$ holds for some $m \in \mathbb{Z}$. Now $v\phi = u^m\phi$ is equivalent to $\bar{v} = \bar{u}^m$, and the claim follows. $\square$

Before stating the theorem concerning cyclonormality, we need to establish the following lemma.

Lemma 7.27. *Let A be a set of size $n > 2$ and $A_1, \dots, A_\ell \subseteq A$ ($\ell \geq 1$) be distinct nonempty subsets such that $\cup_{i=1}^\ell A_i = A$. If there are more than $\binom{n}{2}$ nonsingletons among the A_i , then there exist distinct $1 \leq i, j \leq \ell$ satisfying $|A_i \cap A_j| > 1$.*

Proof. We will proceed by contradiction. Assume that there exists a counterexample to the statement, so that for all distinct $1 \leq i, j \leq \ell$ we have $|A_i \cap A_j| \leq 1$. If there are subsets with more than 2 elements among the nonsingletons, say $A_1, \dots, A_k$ ($1 \leq k \leq \ell$), let $\mathcal{B}_i$ denote the collection of all 2-element subsets of such an A_i . Then, if we replace each A_i ($1 \leq i \leq k$) by the sets in $\mathcal{B}_i$, we obtain a counterexample to the theorem where all sets have exactly 2 elements. Indeed, if $i \neq j$ then no set in $\mathcal{B}_i$ can be equal to a set in $\mathcal{B}_j$, for otherwise we would have $|A_i \cap A_j| \geq 2$, and for the same reason no set in $\mathcal{B}_i$ can be equal to A_j . However, there are only $\binom{n}{2}$ subsets of A with only 2 elements, which contradicts the assumption that $k > \binom{n}{2}$. Hence, there cannot exist a counterexample to the statement, which means that there must exist distinct $1 \leq i, j \leq \ell$ such that $|A_i \cap A_j| > 1$. $\square$

We are ready to prove the following result.

Theorem 7.28. *Suppose that $|A| \geq 2$ and let $K \leq_{f.g.} F_A$ be a cyclonormal subgroup of F_A with Stallings automaton $\mathcal{S}(K) = (Q, A, \delta, q_0)$. Let E and $k \geq 1$ be as in Theorem 7.24. Then either $k = 2$ or $k = 3$. Moreover, if $|Q| = n > 2$, then $k = 2$ implies $|E| \leq \binom{n}{2} + 1$; and if $k = 3$, then $|E| \leq n + \binom{n}{2} + 1$.*

Proof. If $k = 1$, then $M(K)$ is a finite group and $\mathcal{S}(K)$ is a complete automaton, as we observed in the proof of Theorem 7.24. Consequently, given two distinct letters $a, b \in A$, there exist $m, n \in \mathbb{Z}$ such that a^m and b^n label loops at both p and q . However, there are no $u \in R_A$ and $s, t \in \mathbb{Z}$ satisfying $a^m = \bar{u}^s$ and $b^n = \bar{u}^t$, as the first letter of u would have to be simultaneously equal to a and b . Hence, if K is cyclonormal, then $k > 1$.

Now let $Q_1, Q_2, \dots, Q_\ell \subseteq Q$ ($\ell \geq 1$) be distinct nonempty sets satisfying

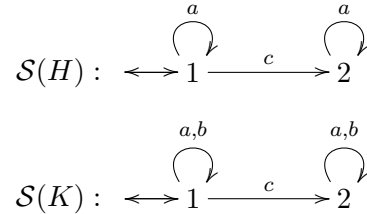
$$E = \{\emptyset, \text{id}_{Q_1}, \dots, \text{id}_{Q_\ell}\}.$$

Since K is cyclonormal, there cannot exist $i \neq j$ such that $|Q_i \cap Q_j| \geq 2$. Otherwise, we could find $\delta_v \neq \delta_w \in E$ fixing at least two vertices $p, q \in Q$, and we may assume that $\text{dom } \delta_w \not\subseteq \text{dom } \delta_v$. Suppose that there were $u \in R_A$ labelling a loop at both p and q and $s, t \in \mathbb{Z}$ satisfying $v = \overline{u^s}$ and $w = \overline{u^t}$. Then taking $r \in \text{dom } \delta_w \setminus \text{dom } \delta_v$, we would get $(r, u^{|t|})\delta = (r, u^t)\delta = (r, w)\delta = r$, whence $(r, u^{|t|s})\delta = r$. However, we would also get $(r, \overline{u^s})\delta = (r, v)\delta \notin Q$, which implies $(r, u^s)\delta \notin Q$ and $(r, u^{|t|s})\delta \notin Q$, a contradiction. This allows us to conclude that $k \leq 3$, for otherwise there would exist $\delta_v < \delta_w$ whose domains would have at least two vertices in common.

Moreover, if $|Q| = n > 2$, $k = 2$ and $|E| > \binom{n}{2} + 1$, then $\ell > \binom{n}{2}$. If there are no singletons among the Q_i , then Lemma 7.27 holds for $A = Q$ and $A_i = Q_i$ ($1 \leq i \leq \ell$); therefore, there exist distinct i and j such that $|Q_i \cap Q_j| \geq 2$, a contradiction since K is cyclonormal. Otherwise, assume that $Q_1, \dots, Q_t$ ($1 \leq t \leq \ell$) are singletons, say $Q_1 = \{q_1\}, \dots, Q_t = \{q_t\}$, for some $q_1, \dots, q_t \in Q$. Let $\sigma \in S_Q$ be a permutation on Q whose cycles have length greater than 2 (this is possible since $|Q| > 2$). For $1 \leq i \leq t$, let $B_i = \{q_i, q_i\sigma\}$. Then the nonsingular sets $B_1, \dots, B_t, Q_{t+1}, \dots, Q_\ell$ cover Q , the B_i are all distinct by the choice of σ , and no B_i is equal to a Q_j if $j \neq i$, as $k = 2$ implies $Q_i \not\subseteq Q_j$. By Lemma 7.27, we conclude that there exist $1 \leq i \leq t$ and $t+1 \leq j \leq \ell$ for which $|B_i \cap Q_j| \geq 2$, thereby implying $Q_i \subset Q_j$, which is absurd. The contradiction we arrived at followed from the assumption that $|E| > \binom{n}{2}$. We conclude that $|E| \leq \binom{n}{2} + 1$ as claimed.

Finally, if $|Q| = n > 2$, $k = 3$ and $\ell > n + \binom{n}{2}$, there are certainly more than $\binom{n}{2}$ nonsingletons among the Q_i . It follows from the previous lemma that there exist distinct $1 \leq i, j \leq \ell$ such that $|Q_i \cap Q_j| \geq 2$, which contradicts the fact that K is cyclonormal. We conclude that, in this case, $\ell \leq n + \binom{n}{2}$, and thus $|E| \leq n + \binom{n}{2} + 1$. $\square$

The following example shows that $M(K)$ and E alone are not sufficient to completely characterize cyclonormality. Let $A = \{a, b, c\}$ and consider the subgroups $H, K \leq_{f.g.} F_A$ with Stallings automata



It is clear that $M(K) = M(H)$ and, in both cases, $E = \{\emptyset, (1\ 2), (1\ -), (-\ 2)\}$. However, H is cyclonormal whereas K is not.

8 Conclusion

In this monograph, we witnessed once again the extraordinary power of Stallings automata to study finitely generated subgroups of a free group. We also concluded that it is worth to analyse the algebraic properties of the transition monoid of a Stallings automaton, for in many cases we can derive interesting conclusions regarding the finitely generated subgroup itself.

Given a free group F_A over a finite alphabet A , a subgroup $K \leq_{f.g.} F_A$ and a pseudovariety of finite groups $\mathbf{H}$, we proved that if $M(K) \in \overline{\mathbf{H}}$, then $M(K\varphi) \in \overline{\mathbf{H}}$ for every automorphism φ of the free group. Hence, all pseudovarieties of the type $\overline{\mathbf{H}}$ are stable under all automorphisms of F_A , making possible to extract properties of K from properties of $\mathbf{H}$. Indeed, we were able to provide explicit characterizations of K in the cases where a single variable appears in the identities defining $\mathbf{H}$. This generalizes some of the conclusions stated in [3].

We were also able to characterize normal and malnormal f.g. subgroups of a free group in terms of the transition monoids of their Stallings automata. A partial characterization of finitely generated cyclonormal subgroups was also achieved.

Regarding future directions of work, it would be interesting to determine more properties P (not necessarily related to pseudovarieties) satisfying the following condition: given a subgroup $K \leq_{f.g.} F_A$, if $M(K)$ satisfies property P , then $M(K\varphi)$ also satisfies property P , for all $\varphi \in \text{Aut}(F_A)$. The next step would be to link the algebraic property P of $M(K)$ to an algebraic property of K as a subgroup of F_A .

Moreover, it could also be interesting to study the effect of various operators acting on the lattice of f.g. subgroups of a free group at the level of the transition monoids of Stallings automata.

Finally, one could pursue these same goals regarding structures that generalize Stallings automata for some wider classes of groups [12], [16], [23], [13].

Bibliography

- [1] Almeida, J. (1994). *Finite Semigroups and Universal Algebra*. World Scientific.
- [2] Bartholdi, L., & Silva, P. V. (2010). Rational subsets of groups. Chapter 23 of the handbook *AutoMathA* (to appear), arXiv:1012.1532.
- [3] Birget, J. C., Margolis, S., Meakin, J., & Weil, P. (2000). PSPACE-complete problems for subgroups of free groups and inverse finite automata. *Theoret. Comput. Sci.*, 242(1–2), 247–281.
- [4] Cioabă, S. M. & Murty, M. R. (2009). *A First Course in Graph Theory and Combinatorics*. Hindustan Book Agency.
- [5] Eilenberg, S. (1974). *Automata, Languages, and Machines* (Vol. A). Academic Press.
- [6] Eilenberg, S. (1976). *Automata, Languages, and Machines* (Vol. B). Academic Press.
- [7] Eilenberg, S., & Schützenberger, M. (1976). On pseudovarieties. *Adv. Math.*, 19(3), 413–418.
- [8] Green, J. A. (1951). On the structure of semigroups. *Ann. Math.*, 54(1), 163–172.
- [9] Howie, J. M. (2003). *Fundamentals of Semigroup Theory*. Oxford University Press.
- [10] Howson, A. G. (1954). On the intersection of finitely generated free groups. *J. London Math. Soc.*, 1(4), 428–434.
- [11] Kapovich, I., & Myasnikov, A. (2002). Stallings foldings and subgroups of free groups. *J. Algebra*, 248(2), 608–668.
- [12] Kapovich, I., Weidmann, R., Miasnikov, A. (2005). Foldings, graphs of groups and the membership problem. *Internat. J. Algebra Comput.*, 15(1), 95–128.
- [13] Kharlampovich, O., Miasnikov, A., & Weil, P. (2017). Stallings graphs for quasi-convex subgroups. *J. Algebra*, 488, 442–483.
- [14] Lallement, G. (1966). Congruences et équivalences de Green sur un demigroupe régulier. *C. R. Acad. Sci. Paris*, 262, 613–616.
- [15] Lyndon, R. C., & Schupp, P. E. (2001). *Combinatorial Group Theory*. Springer.
- [16] Markus-Epstein, L. (2007). Stallings foldings and subgroups of amalgams of finite groups, *Internat. J. Algebra Comput.*, 17(8), 1493–1535.
- [17] Nielsen, J. (1924). Die Isomorphismengruppe der freien Gruppen. *Mathematische Annalen*, 91, 169–209.
- [18] Petrich, M. (1984). *Inverse Semigroups*. Wiley.

- [19] Pin, J.-E. (1986). *Varieties of Formal Languages*. London: Plenum.
- [20] Preston, G. B. (1954). Representations of inverse semi-groups. *J. London Math. Soc.*, 29(4), 411–419.
- [21] Rhodes, J., & Steinberg, B. (2010). *The q -theory of Finite Semigroups*. Springer.
- [22] Schützenberger, M. (1965). On finite monoids having only trivial subgroups. *Inform. and Control*, 8, 190–194.
- [23] Silva, P. V., Soler-Escriba, X., & Ventura, E. (2016). Finite automata for Schreier graphs of virtually free groups. *J. Group Theory*, 19(1), 25–54.
- [24] Simon, I. (1975). Piecewise testable events. In Brakhage, H. (Ed.), *Automata Theory and Formal Languages* (1st ed., pp. 214–222). Springer.
- [25] Stallings, J. (1983). Topology of finite graphs. *Invent. Math.*, 71, 551–565.
- [26] Touikan, N.W. (2006). A fast algorithm for Stallings’ folding process. *Int. J. Algebra Comput.*, 16, 1031–1046.
- [27] Wagner, V. V. (1952). Generalised groups. *Dokl. Akad. nauk SSSR*, 84, 1119–1122.

Index

- algebra, 54
 - τ -algebra, 54
 - finite, 54
 - interpretation, 54
 - signature, 54
 - subalgebra, 54
 - universe, 54
- alphabet, 7
 - letter, 7
- automaton, 27
 - basepoint, 65
 - complete, 27
 - deterministic, 28
 - edge, 27
 - loop, 28
 - equivalent, 28
 - finite, 27
 - flower, 65
 - inverse, 42
 - involutive, 42
 - minimal, 41
 - path, 28
 - intermediate vertex of, 28
 - label, 28
 - reduced, 74
 - successful, 28
 - trivial, 28
 - Stallings, 65
 - core, 70
 - tail, 70
 - state, 27
 - initial, 27
 - terminal, 27
 - subset construction, 29
 - trim, 30
 - trimming, 30
 - underlying graph of, 27
 - vertex, 27
 - accessible, 30
 - co-accessible, 30
 - with 1-transitions, 30
- automorphism, 26
 - elementary Nielsen, 63
 - graph, 26
 - of type 1, 63
 - of type 2, 63
- band, 14
- basis, 62
- binary relation, 6
 - composite, 6
- Burnside pseudovariety, 92
- Cayley's theorem, 6
- chain, 98
- class, 45
 - of τ -algebras, 55
 - of monoids, 45
 - of rational languages, 50
 - stable under an automorphism, 75
- concatenation, 7
- congruence, 4
 - generated by a relation, 9
 - saturation by, 36
 - Wagner, 22
- cyclic, 63
 - conjugate, 64
 - core, 64
 - reduction, 63
- direct product, 45
 - arbitrary, 45
 - finitary, 45
 - of τ -algebras, 55
- egg-box, 12
- Eilenberg's correspondence, 51
- finitely generated, 6
 - group, 62
 - monoid, 6
 - pseudovariety, 46
- free, 8
 - group, 61
 - inverse monoid, 23
 - monoid, 8
 - semigroup, 8

- function, 6
 - incidence, 26
 - transition, 29
- geodesic, 68
 - T -geodesic, 68
 - distance, 68
- graph, 26
 - connected, 27
 - connected component, 27
 - edge, 26
 - head, 26
 - loop, 26
 - tail, 26
 - path, 26
 - length, 27
 - trivial, 27
 - subgraph, 26
 - induced, 26
 - vertex, 26
 - indegree, 26
 - outdegree, 26
 - total degree, 26
 - vertex-transitive, 71
- Green's relations, 9
- group, 3
 - π -group, 92
 - p -group, 73
 - conjugate elements, 63
 - right action, 95
 - orbit, 95
 - stabilizer, 95
 - transitive, 95
 - subgroup, 3
 - p -pure, 73
 - conjugate, 70
 - cyclonormal, 100
 - generated by a set, 62
 - index of, 69
 - malnormal, 96
 - normal, 71
 - pure, 73
 - torsion free, 65
- group $\mathcal{H}$ -class, 12
- homomorphism, 3
 - τ -algebra, 54
 - monoid, 3
- Hopcroft equivalence, 39
- Howson property, 72
- ideal, 3
 - left, 3
 - right, 3
- identity, 46
 - τ -algebra, 55
 - monoid, 46
- isomorphism
 - automaton, 27
 - graph, 26
 - monoid, 3
- kernel, 4
- language, 25
 - A -language, 25
 - piecewise testable, 53
 - rational, 25
 - rational expression, 25
 - rational operator, 25
 - product, 25
 - star, 25
 - union, 25
 - recognizable, 28
 - recognized by a monoid, 37
 - recognized by an automaton, 28
 - star-free, 53
- monoid, 3
 - $\mathcal{K}$ -trivial, 48
 - p -periodic, 73
 - 6-element Brandt, 94
 - aperiodic, 48
 - completely regular, 49
 - division, 37
 - exponent, 48
 - idempotent element, 3
 - inverse, 15
 - submonoid, 17
 - inverse element, 13
 - regular, 13
 - regular element, 13
 - submonoid, 3
 - generated by a set, 5
 - syntactic, 36
 - transition, 34
 - zero element, 3

- morphism
 - τ -algebra, 54
 - automaton, 27
 - monoid, 3
- partial order, 14
 - natural, 14
- partial transformation, 6
 - converse, 6
 - domain, 6
 - injective, 19
 - range, 6
- partially ordered set, 14
- poset, 14
- presentation, 9
- problem, 63
 - conjugacy, 63
 - decidable, 63
 - generalized word, 65
 - word, 63
- pseudovariety, 45
 - of finite τ -algebras, 55
 - of finite groups, 55
 - of finite inverse monoids, 55
 - of finite monoids, 45
- quotient, 3
 - of a language, 50
 - left, 50
 - right, 50
 - of a monoid, 3
- rank, 62
- representation, 7
 - right regular, 7
 - Wagner-Preston, 20
- rewriting system, 59
 - confluent, 59
 - elementary reduction, 59
 - length-reducing, 59
 - noetherian, 59
 - reduction, 59
- semigroup, 3
 - subsemigroup, 3
- semilattice, 14
 - $\wedge$ -semilattice, 14
 - infimum, 14
 - of idempotents, 16
- spanning tree, 68
- Stallings folding, 65
- symbol, 53
 - constant, 53
 - operation, 53
 - arity, 53
 - binary, 53
 - unary, 53
- terms, 55
 - algebra of, 55
- transformation, 6
 - image, 6
- type, 53
- ultimately, 46
 - defined by, 46
 - satisfied, 46
- variety, 45
 - of τ -algebras, 55
 - of monoids, 45
 - of rational languages, 50
- word, 7
 - empty, 7
 - factor, 26
 - irreducible, 59
 - length, 7
 - prefix, 25
 - reduced, 61
 - suffix, 25

Notation

$(X, \leq)$, 14	$\mathbf{G}_p$, 83
$<_{f.g.}$, 98	$\mathbf{J}$, 50
A^* , 7	$\mathbf{L}$, 50
$A^*\mathcal{C}$, 50	$\mathbf{R}$, 50
A^+ , 7	$\mathbf{Sl}$, 48
A^{-1} , 22	$\mathbf{V}_S$, 91
B_2^1 , 94	$\mathcal{C}(K)$, 70
$E(M)$, 3	$\mathcal{D}$, 9
E^+ , 42	$\mathcal{F}(K)$, 67
E^- , 42	$\mathcal{F}(u_1, \dots, u_k)$, 65
FIM_A , 22	$\mathcal{H}$, 9
$F^{\mathcal{A}}$, 54	$\mathcal{I}_X$, 19
F_A , 61	$\mathcal{J}$, 9
G_x , 95	$\mathcal{L}$, 9
Inv , 56	$\mathcal{O}_n$, 54
$L(\mathcal{A})$, 28	$\mathcal{PT}_X$, 6
L^* , 25	$\mathcal{R}$, 9
$L_{(p,q)}$, 101	$\mathcal{R}_A$, 60
L_{PR} , 39	$\mathcal{S}(u_1, \dots, u_k)$, 65
Lu^{-1} , 50	$\mathcal{T}_X$, 6
$M(K)$, 73	$\cong$, 3
$M(K)^\beta$, 76	δ_u , 34
$M(\mathcal{A})$, 34	$\text{dom } \sigma$, 6
M/σ , 3	$\text{im } \varphi$, 6
Pu , 29	$\ker \varphi$, 4
$R^\sharp$, 9	$\langle X \rangle$, 5
R_A , 101	$\leq_{f.g.}$, 65
S_X , 95	$\leq_{f.i.}$, 69
$T(X)$, 55	$\llbracket \Sigma \rrbracket$, 48
$[G : K]$, 69	$\text{Mon}\langle A \mid R \rangle$, 9
$\Gamma(K)$, 70	$\min_L$, 41
$\Gamma_{\mathcal{A}}$, 27	$\models$, 46
$\Phi _A$, 7	ω , 48
$\mathcal{S}(K)$, 66	$\overline{K}$, 62
$\mathcal{S}(K)^\beta$, 76	$\overline{\mathbf{H}}$, 56
α_a , 63	$\overline{n}$, 48
β_{ab} , 63	$\overline{u}$, 61
$\mathbf{Ab}$, 83	$\mathbf{H}(\mathcal{C})$, 46
$\mathbf{A}$, 49	$\text{Irr}(\mathcal{R})$, 59
$\mathbf{A}_n$, 94	$\mathbf{P}(\mathcal{C})$, 46
$\mathbf{B}_k$, 92	$\text{Synt}(\mathbf{L})$, 36
$\mathbf{CR}$, 50	$\mathbf{S}(\mathcal{C})$, 46
$\mathbf{Com}$, 48	$\mathbf{V}(\mathcal{C})$, 46
$\mathbf{G}$, 48	$\text{ran } \sigma$, 6
$\mathbf{G}_\pi$, 93	$\text{Aut}(F_A)$, 63

id_X , 6	$d^-(v)$, 26
σ^{-1} , 6	$f^{\mathcal{A}}$, 54
$\sim$, 39	g_q , 68
$\sim_L$, 36	$u^{-1}L$, 50
$\sqsubseteq$, 71	$x\sigma y$, 4
$\widetilde{A}$, 22	$x \wedge y$, 14
$c(u)$, 64	xG , 95
$d(v)$, 26	x^{-1} , 15
$d^+(v)$, 26	$\mathcal{B}_X$, 6