Rational subsets of groups

L. Bartholdi[1] *P. V. Silva*[2,*]

[1]Mathematisches Institut
Georg-August Universität zu Göttingen
Bunsenstraße 3–5
D-37073 Göttingen, Germany
email: laurent.bartholdi@gmail.com

[2]Centro de Matemática, Faculdade de Ciências
Universidade do Porto
R. Campo Alegre 687
4169-007 Porto, Portugal
email: pvsilva@fc.up.pt

2010 Mathematics Subject Classification: 20F10, 20E05, 68Q45, 68Q70

Key words: Free groups, inverse automata, Stallings automata, rational subsets.

Contents

*The second author acknowledges support by Project ASA (PTDC/MAT/65481/2006) and C.M.U.P., financed by F.C.T. (Portugal) through the programmes POCTI and POSI, with national and E.U. structural funds.

Over the years, finite automata have been used effectively in the theory of infinite groups to represent rational subsets. This includes the important particular case of finitely generated subgroups (and the beautiful theory of Stallings automata for the free group case), but goes far beyond that: certain inductive procedures need a more general setting than mere subgroups, and rational subsets constitute the natural generalization. The connections between automata theory and group theory are rich and deep, and many are portrayed in Sims' book [53].

This chapter is divided into three parts: in Section 1 we introduce basic concepts, terminology and notation for finitely generated groups, devoting special attention to free groups. These will also be used in Chapter 24.

Section 2 describes the use of finite inverse automata to study finitely generated subgroups of free groups. The automaton recognizes elements of a subgroup, represented as words in the ambient free group.

Section 3 considers, more generally, rational subsets of groups, when good closure and decidability properties of these subsets are satisfied.

The authors are grateful to Stuart Margolis, Benjamin Steinberg and Pascal Weil for their remarks on a preliminary version of this text.

1 Finitely generated groups

Let G be a group. Given $A \subseteq G$, let $\langle A \rangle = (A \cup A^{-1})^*$ denote the subgroup of G *generated* by A. We say that $H \leqslant G$ is *finitely generated* and write $H \leqslant_{f.g.} G$ if $H = \langle A \rangle$ for some finite subset A of H.

Given $H \leqslant G$, we denote by $[G : H]$ the *index* of H in G, that is, the number of right cosets Hg for all $g \in G$; or, equivalently, the number of left cosets. If $[G : H]$ is finite, we write $H \leqslant_{f.i.} G$. It is well known that every finite index subgroup of a finitely generated group is finitely generated.

We denote by $\mathbb{1}$ the identity of G. An element $g \in G$ has *finite order* if $\langle g \rangle$ is finite. Elements $g, h \in G$ are *conjugate* if $h = x^{-1}gx$ for some $x \in G$. We use the notation $g^h = h^{-1}gh$ and $[g, h] = g^{-1}g^h$ to denote, respectively, conjugates and commutators.

Given an alphabet A, we denote by A^{-1} a set of *formal inverses* of A, and write $\widetilde{A} = A \cup A^{-1}$. We say that $\widetilde{A}$ is an *involutive alphabet*. We extend $^{-1} : A \to A^{-1} : a \mapsto a^{-1}$ to an involution on $\widetilde{A}^*$ through

$$(a^{-1})^{-1} = a, \quad (uv)^{-1} = v^{-1}u^{-1} \quad (a \in A,\ u, v \in \widetilde{A}^*)\,.$$

If $G = \langle A \rangle$, we have a canonical epimorphism $\rho : \widetilde{A}^* \twoheadrightarrow G$, mapping $a^{\pm 1} \in \widetilde{A}$ to $a^{\pm 1} \in G$. We present next some classical decidability problems:

Definition 1.1. Let $G = \langle A \rangle$ be a finitely generated group.

word problem: is there an algorithm that, upon receiving as input a word $u \in \widetilde{A}^*$, determines whether or not $\rho(u) = \mathbb{1}$?

conjugacy problem: is there an algorithm that, upon receiving as input words $u, v \in \widetilde{A}^*$, determines whether or not $\rho(u)$ and $\rho(v)$ are conjugate in G?

membership problem for $\mathcal{K} \subseteq 2^G$: is there for every $X \in \mathcal{K}$ an algorithm that, upon receiving as input a word $u \in \widetilde{A}^*$, determines whether or not $\rho(u) \in X$?

generalized word problem: is the membership problem for the class of finitely generated subgroups of G solvable?

order problem: is there an algorithm that, upon receiving as input a word $u \in \widetilde{A}^*$, determines whether $\rho(u)$ has finite or infinite order?

isomorphism problem for a class $\mathcal{G}$ of groups: is there an algorithm that, upon receiving as input a description of groups $G, H \in \mathcal{G}$, decides whether or not $G \cong H$? Typically, $\mathcal{G}$ may be a subclass of finitely presented groups (given by their presentation), or automata groups (see Chapter 24) given by automata.

We can also require complexity bounds on the algorithms; more precisely, we may ask with which complexity bound an answer to the problem may be obtained, and also with which complexity bound a witness (a normal form for the word problem, an element conjugating $\rho(u)$ to $\rho(v)$ in case they are conjugate, an expression of u in the generators of X in the generalized word problem) may be constructed.

1.1 Free groups

We recall that an equivalence relation $\sim$ on a semigroup S is a *congruence* if $a \sim b$ implies $ac \sim bc$ and $ca \sim cb$ for all $a, b, c \in S$.

Definition 1.2. Given an alphabet A, let $\sim$ denote the congruence on $\widetilde{A}^*$ generated by the relation

$$\{(aa^{-1}, 1) \mid a \in \widetilde{A}\}\,. \tag{1.1}$$

The quotient $F_A = \widetilde{A}^*/\sim$ is the *free group on A*. We denote by $\theta : \widetilde{A}^* \to F_A$ the canonical morphism $u \mapsto [u]_\sim$.

Free groups admit the following universal property: for every map $f : A \to G$, there is a unique group morphism $F_A \to G$ that extends f.

Alternatively, we can view (1.1) as a *confluent* length-reducing rewriting system on $\widetilde{A}^*$, where each word $w \in \widetilde{A}^*$ can be transformed into a unique *reduced* word $\overline{w}$ with no factor of the form aa^{-1}, see [9]. As a consequence, the equivalence

$$u \sim v \quad \Leftrightarrow \quad \overline{u} = \overline{v} \qquad\qquad (u, v \in \widetilde{A}^*)$$

solves the word problem for F_A.

We shall use the notation $R_A = \overline{\widetilde{A}^*}$. It is well known that F_A is isomorphic to R_A under the binary operation

$$u \star v = \overline{uv} \qquad (u, v \in R_A)\,.$$

We recall that the *length* $|g|$ of $g \in F_A$ is the length of the reduced form of g, also denoted by $\overline{g}$.

The letters of A provide a natural *basis* for F_A: they generate F_A and satisfy no nontrivial relations, that is, all reduced words on these generators represent distinct elements of F_A. A group is free if and only if it has a basis.

Throughout this chapter, we assume A to be a finite alphabet. It is well known that free groups F_A and F_B are isomorphic if and only if $\#A = \#B$. This leads to the concept of *rank* of a free group F: the *cardinality* of a basis of F, denoted by $\mathrm{rk}\, F$. It is common to use the notation F_n to denote a free group of rank n.

We recall that a reduced word u is *cyclically reduced* if uu is also reduced. Any reduced word $u \in R_A$ admits a unique decomposition of the form $u = vwv^{-1}$ with w cyclically reduced. A solution for the conjugacy problem follows easily from this: first reduce the words cyclically; then two cyclically reduced words in R_A are conjugate if and only if they are cyclic permutations of each other. On the other hand, the order problem admits a trivial solution: only the identity has finite order. Finally, the generalized word problem shall be discussed in the following section.

2 Inverse automata and Stallings' construction

The study of finitely generated subgroups of free groups entered a new era in the early eighties when Stallings made explicit and effective a construction [54] that can be traced back to the early part of the twentieth century in Schreier's coset graphs (see [53] and §24.1) and to Serre's work [46]. Stallings' seminal paper was built over *immersions of finite graphs*, but the alternative approach using finite inverse automata became much more popular over the years; for more on their link, see [26]. An extensive survey has been written by Kapovich and Miasnikov [20].

Stallings' construction for $H \leqslant_{f.g.} F_A$ consists in taking a finite set of generators for H in reduced form, building the so-called *flower automaton* and then proceeding to make this automaton deterministic through the operation known as *Stallings foldings*. This is clearly a terminating procedure, but the key fact is that the construction is independent from both the given finite generating set and the chosen folding sequence. A short simple automata-theoretic proof of this claim will be given. The finite inverse automaton $\mathcal{S}(H)$ thus obtained is usually called the *Stallings automaton* of H. Over the years, Stallings automata became the standard representation for finitely generated subgroups of free groups and are involved in many of the algorithmic results presently obtained.

Several of these algorithms are implemented in computer software, see e.g. CRAG [2], or the packages AUTOMATA and FGA in GAP [14].

2.1 Inverse automata

An automaton $\mathcal{A}$ over an involutive alphabet $\widetilde{A}$ is *involutive* if, whenever (p, a, q) is an edge of $\mathcal{A}$, so is (q, a^{-1}, p). Therefore it suffices to depict just the *positively labelled* edges (having label in A) in their graphical representation.

Definition 2.1. An involutive automaton is *inverse* if it is deterministic, trim and has a single final state.

If the latter happens to be the initial state, it is called the *basepoint*. It follows easily from the computation of the Nerode equivalence (see §10.2) that every inverse automaton is a minimal automaton.

Finite inverse automata capture the idea of an action (of a finite *inverse monoid*, their *transition monoid*) on a finite set (the vertex set) through partial bijections. We recall that a monoid M is inverse if, for every $x \in M$, there exists a unique $y \in M$ such that $xyx = x$ and $y = yxy$; then M acts by partial bijections on itself.
The next result is easily proven, but is quite useful.

Proposition 2.1. *Let $\mathcal{A}$ be an inverse automaton and let $p \xrightarrow{uvv^{-1}w} q$ be a path in $\mathcal{A}$. Then there exists also a path $p \xrightarrow{uw} q$ in $\mathcal{A}$.*

Another important property relates languages to morphisms. For us, a *morphism* between deterministic automata $\mathcal{A}$ and $\mathcal{A}'$ is a mapping φ between their respective vertex sets which preserves initial vertices, final vertices and edges, in the sense that $(\varphi(p), a, \varphi(q))$ is an edge of $\mathcal{A}'$ whenever (p, a, q) is an edge of $\mathcal{A}$.

Proposition 2.2. *Given inverse automata $\mathcal{A}$ and $\mathcal{A}'$, then $L(\mathcal{A}) \subseteq L(\mathcal{A}')$ if and only if there exists a morphism $\varphi : \mathcal{A} \to \mathcal{A}'$. Moreover, such a morphism is unique.*

Proof. ($\Rightarrow$): Given a vertex q of $\mathcal{A}$, take a successful path

$$\to q_0 \xrightarrow{u} q \xrightarrow{v} t \to$$

in $\mathcal{A}$, for some $u, v \in \widetilde{A}^*$. Since $L(\mathcal{A}) \subseteq L(\mathcal{A}')$, there exists a successful path

$$\to q_0' \xrightarrow{u} q' \xrightarrow{v} t' \to$$

in $\mathcal{A}'$. We take $\varphi(q) = q'$.

To show that φ is well defined, suppose that

$$\to q_0 \xrightarrow{u'} q \xrightarrow{v'} t \to$$

is an alternative successful path in $\mathcal{A}$. Since $u'v \in L(\mathcal{A}) \subseteq L(\mathcal{A}')$, there exists a successful path

$$\to q_0' \xrightarrow{u'} q'' \xrightarrow{v} t' \to$$

in $\mathcal{A}'$ and it follows that $q' = q''$ since $\mathcal{A}'$ is inverse. Thus φ is well defined.

It is now routine to check that φ is a morphism from $\mathcal{A}$ to $\mathcal{A}'$ and that it is unique.

($\Leftarrow$): Immediate from the definition of morphism. □

2.2 Stallings' construction

Let X be a finite subset of R_A. We build an involutive automaton $\mathcal{F}(X)$ by fixing a basepoint q_0 and gluing to it a *petal* labelled by every word in X as follows: if $x = a_1 \dots a_k \in X$, with $a_i \in \widetilde{A}$, the petal consists of a closed path of the form

$$q_0 \xrightarrow{a_1} \bullet \xrightarrow{a_2} \cdots \xrightarrow{a_k} q_0$$

and the respective inverse edges. All such intermediate vertices $\bullet$ are assumed to be distinct in the automaton. For obvious reasons, $\mathcal{F}(X)$ is called the *flower automaton* of X.

The automaton $\mathcal{F}(X)$ is almost an inverse automaton – except that it need not be deterministic. We can fix it by performing a sequence of so-called *Stallings foldings*. Assume that $\mathcal{A}$ is a trim involutive automaton with a basepoint, possessing two distinct edges of the form

$$p \xrightarrow{a} q, \quad p \xrightarrow{a} r \tag{2.1}$$

for $a \in \widetilde{A}$. The *folding* is performed by identifying these two edges, as well as the two respective inverse edges. In particular, the vertices q and r are also identified (if they were distinct).

The number of edges is certain to decrease through foldings. Therefore, if we perform enough of them, we are sure to turn $\mathcal{F}(X)$ into a finite inverse automaton.

Definition 2.2. The *Stallings automaton* of X is the finite inverse automaton $\mathcal{S}(X)$ obtained through folding $\mathcal{F}(X)$.

We shall see that $\mathcal{S}(X)$ depends only on the finitely generated subgroup $\langle X \rangle$ of F_A generated by X, being in particular independent from the choice of foldings taken to reach it.

Since inverse automata are minimal, it suffices to characterize $L(\mathcal{S}(X))$ in terms of H to prove uniqueness (up to isomorphism):

Proposition 2.3. *Fix $H \leqslant_{f.g.} F_A$ and let $X \subseteq R_A$ be a finite generating set for H. Then*

$$L(\mathcal{S}(X)) = \bigcap \{L \subseteq \widetilde{A}^* \mid L \textit{ is recognized by a finite inverse automaton with a basepoint and } \overline{H} \subseteq L\}\,.$$

Proof. ($\supseteq$): Clearly, $\mathcal{S}(X)$ is a finite inverse automaton with a basepoint. Since $X \cup X^{-1} \subseteq L(\mathcal{F}(X)) \subseteq L(\mathcal{S}(X))$, it follows easily from Proposition 2.1 that

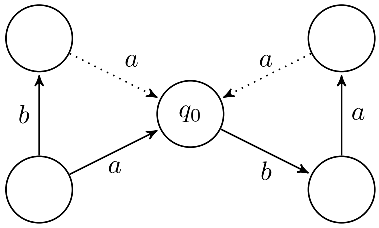
$$\overline{H} \subseteq L(\mathcal{S}(X))\,. \tag{2.2}$$

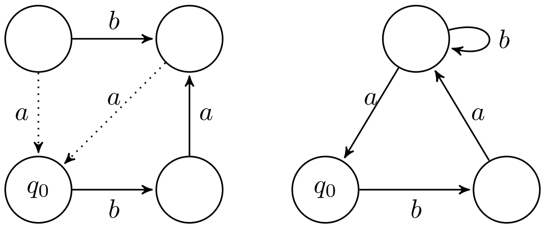
($\subseteq$): Let $L \subseteq \widetilde{A}^*$ be recognized by a finite inverse automaton $\mathcal{A}$ with a basepoint, with $\overline{H} \subseteq L$. Since $X \subseteq \overline{H}$, we have an automaton morphism from $\mathcal{F}(X)$ to $\mathcal{A}$, hence $L(\mathcal{F}(X)) \subseteq L$. To prove that $L(\mathcal{S}(X)) \subseteq L$, it suffices to show that inclusion in L is preserved through foldings.

Indeed, assume that $L(\mathcal{B}) \subseteq L$ and $\mathcal{B}'$ is obtained from $\mathcal{B}$ by folding the two edges in (2.1). It is immediate that every successful path $q_0 \xrightarrow{u} t$ in $\mathcal{B}'$ can be lifted to a successful path $q_0 \xrightarrow{v} t$ in $\mathcal{B}$ by successively inserting the word $a^{-1}a$ into u. Now $v \in L = L(\mathcal{A})$ implies $u \in L$ in view of Proposition 2.1. □

Now, given $H \leqslant F_A$ finitely generated, we take a finite set X of generators. Without loss of generality, we may assume that X consists of reduced words, and we may define $\mathcal{S}(H) = \mathcal{S}(X)$ to be the *Stallings automaton* of H.

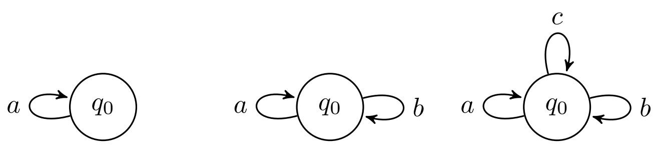
Example 2.1. Stallings' construction for $X = \{a^{-1}ba, ba^2\}$, where the next edges to be identified are depicted by dotted lines, is

$\mathcal{F}(X) =$ 



$= \mathcal{S}(X)$

A simple, yet important example is given by applying the construction to F_n itself, when we obtain the so-called *bouquet* of n circles:



$\mathcal{S}(F_1)$ $\mathcal{S}(F_2)$ $\mathcal{S}(F_3)$

In terms of complexity, the best known algorithm for the construction of $\mathcal{S}(X)$ is due to Touikan [56]. Its time complexity is $O(n \log^* n)$, where n is the sum of the lengths of the elements of X.

2.3 Basic applications

The most fundamental application of Stallings' construction is an elegant and efficient solution to the generalized word problem:

Theorem 2.4. *The generalized word problem in F_A is solvable.*

We will see many groups in Chapter 24 that have solvable word problem; however, few of them have solvable generalized word problem. The proof of Theorem 2.4 relies on

Proposition 2.5. *Consider $H \leqslant_{f.g.} F_A$ and $u \in F_A$. Then $u \in H$ if and only if $\overline{u} \in L(\mathcal{S}(H))$.*

Proof. ($\Rightarrow$): Follows from (2.2).

($\Leftarrow$): It follows easily from the last paragraph of the proof of Proposition 2.3 that, if $\mathcal{B}'$ is obtained from $\mathcal{B}$ by performing Stallings foldings, then $\overline{L(\mathcal{B}')} = \overline{L(\mathcal{B})}$. Hence, if

$H = \langle X \rangle$, we get

$$\overline{L(\mathcal{S}(H))} = \overline{L(\mathcal{F}(X))} = \overline{(X \cup X^{-1})^*} = \overline{H}$$

and the implication follows. □

It follows from our previous remark that the complexity of the generalized word problem is $O(n \log^* n + m)$, where n is the sum of the lengths of the elements of X and m is the length of the input word. In particular, once the subgroup X has been fixed, complexity is linear in m.

Example 2.2. We may use the Stallings automaton constructed in Example 2.1 to check that $baba^{-1}b^{-1} \in H = \langle a^{-1}ba, ba^2 \rangle$ but $ab \notin H$.

Stallings automata also provide an effective construction for bases of finitely generated subgroups. Consider $H \leqslant_{f.g.} F_A$, and let m be the number of vertices of $\mathcal{S}(H)$. A *spanning tree* T for $\mathcal{S}(H)$ consists of $m-1$ edges and their inverses which, together, connect all the vertices of $\mathcal{S}(H)$. Given a vertex p of $\mathcal{S}(H)$, we denote by g_p the T-*geodesic* connecting the basepoint q_0 to p, that is, $q_0 \xrightarrow{g_p} p$ is the shortest path contained in T connecting q_0 to p.

Proposition 2.6. *Let $H \leqslant_{f.g.} F_A$ and let T be a spanning tree for $\mathcal{S}(H)$. Let E_+ be the set of positively labelled edges of $\mathcal{S}(H)$. Then H is free with basis*

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

Proof. It follows from Proposition 2.5 that $L(\mathcal{S}(H)) \subseteq H$, hence $Y \subseteq H$. To show that $H = \langle Y \rangle$, take $h = a_1 \cdots a_k \in H$ in reduced form $(a_i \in \widetilde{A})$. By Proposition 2.5, there exists a successful path

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \cdots \xrightarrow{a_k} q_k = q_0$$

in $\mathcal{S}(H)$. For $i = 1, \ldots, k$, we have either $g_{q_{i-1}} a_i g_{q_i}^{-1} \in Y \cup Y^{-1}$ or $\overline{g_{q_{i-1}} a_i g_{q_i}^{-1}} = 1$, the latter occurring if $(q_{i-1}, a_i, q_i) \in T$. In any case, we get

$$h = a_1 \cdots a_k = \overline{(g_{q_0} a_1 g_{q_1}^{-1})(g_{q_1} a_2 g_{q_2}^{-1}) \cdots (g_{q_{k-1}} a_k g_{q_0}^{-1})} \in \langle Y \rangle$$

and so $H = \langle Y \rangle$.

It remains to show that the elements of Y satisfy no nontrivial relations. Let $y_1, \ldots, y_k \in Y \cup Y^{-1}$ with $y_i \neq y_{i-1}^{-1}$ for $i = 2, \ldots, k$. Write $y_i = g_{p_i} a_i g_{r_i}^{-1}$, where $a_i \in \widetilde{A}$ labels the edge not in T. It follows easily from $y_i \neq y_{i-1}^{-1}$ and the definition of spanning tree that

$$\overline{y_1 \cdots y_k} = g_{p_1} a_1 \overline{g_{r_1}^{-1} g_{p_2}} a_2 \cdots a_{k-1} \overline{g_{r_{k-1}}^{-1} g_{p_k}} a_k g_{r_k},$$

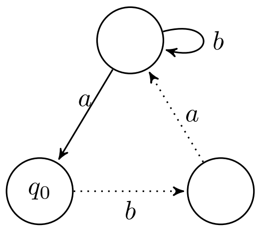
a nonempty reduced word if $k \geqslant 1$. Therefore Y is a basis of H as claimed. □

In the process, we also obtain a proof of the Nielsen-Schreier Theorem, in the case of finitely generated subgroups. A simple topological proof may be found in [36]:

Theorem 2.7 (Nielsen-Schreier). *Every subgroup of a free group is itself free.*

Example 2.3. We use the Stallings automaton constructed in Example 2.1 to construct a basis of $H = \langle a^{-1}ba, ba^2 \rangle$.

If we take the spanning tree T defined by the dotted lines in



then $\#E_+ \setminus T = 2$ and the corresponding basis is $\{ba^2, baba^{-1}b^{-1}\}$. Another choice of spanning tree actually proves that the original generating set is also a basis.

We remark that Proposition 2.6 can be extended to the case of infinitely generated subgroups, proving the general case of Theorem 2.7. However, in this case there is no effective construction such as Stallings', and the (infinite) inverse automaton $\mathcal{S}(H)$ remains a theoretical object, using appropriate cosets as vertices.

Another classical application of Stallings' construction regards the identification of finite index subgroups.

Proposition 2.8. *Consider $H \leqslant_{f.g.} F_A$.*

(i) *H is a finite index subgroup of F_A if and only if $\mathcal{S}(H)$ is a complete automaton.*

(ii) *If H is a finite index subgroup of F_A, then its index is the number of vertices of $\mathcal{S}(H)$.*

Proof. (i) ($\Rightarrow$): Suppose that $\mathcal{S}(H)$ is not complete. Then there exist some vertex q and some $a \in \widetilde{A}$ such that $q \cdot a$ is undefined. Let g be a geodesic connecting the basepoint q_0 to q in $\mathcal{S}(H)$. We claim that

$$Hga^m \neq Hga^n \quad \text{if} \quad m - n > |g| \,. \tag{2.3}$$

Indeed, $Hga^m = Hga^n$ implies $ga^{m-n}g^{-1} \in H$ and so $\overline{ga^{m-n}g^{-1}} \in L(\mathcal{S}(H))$ by Proposition 2.5. Since ga is reduced due to $\mathcal{S}(H)$ being inverse, it follows from $m - n > |g|$ that $\overline{gaa^{m-n-1}g^{-1}} = \overline{ga^{m-n}g^{-1}} \in L(\mathcal{S}(H))$: indeed, g^{-1} is not long enough to erase all the a's. Since $\mathcal{S}(H)$ is deterministic, $q \cdot a$ must be defined, a contradiction. Therefore (2.3) holds and so H has infinite index.

($\Leftarrow$): Let Q be the vertex set of $\mathcal{S}(H)$ and fix a geodesic $q_0 \xrightarrow{g_q} q$ for each $q \in Q$. Take $u \in F_A$. Since $\mathcal{S}(H)$ is complete, we have a path $q_0 \xrightarrow{u} q$ for some $q \in Q$. Hence $ug_q^{-1} \in H$ and so $u = ug_q^{-1}g_q \in Hg_q$. Therefore $F_A = \bigcup_{q \in Q} Hg_q$ and so $H \leqslant_{f.i.} F_A$.

(ii) In view of $F_A = \bigcup_{q \in Q} Hg_q$, it suffices to show that the cosets Hg_q are all distinct. Indeed, assume that $Hg_p = Hg_q$ for some vertices $p, q \in Q$. Then $g_pg_q^{-1} \in H$ and so $\overline{g_pg_q^{-1}} \in L(\mathcal{S}(H))$ by Proposition 2.5. On the other hand, since $\mathcal{S}(H)$ is complete, we have a path

$$q_0 \xrightarrow{g_pg_q^{-1}} r$$

for some $r \in Q$. In view of Proposition 2.1, and by determinism, we get $r = q_0$. Hence we have paths

$$p \xrightarrow{g_q^{-1}} q_0, \quad q \xrightarrow{g_q^{-1}} q_0 .$$

Since $\mathcal{S}(H)$ is inverse, we get $p = q$ as required. □

Example 2.4. Since the Stallings automaton constructed in Example 2.1 is not complete, it follows that $\langle a^{-1}ba, ba^2\rangle$ is not a finite index subgroup of F_2.

Corollary 2.9. *If $H \leqslant F_A$ has index n, then* $\operatorname{rk} H = 1 + n(\#A - 1)$.

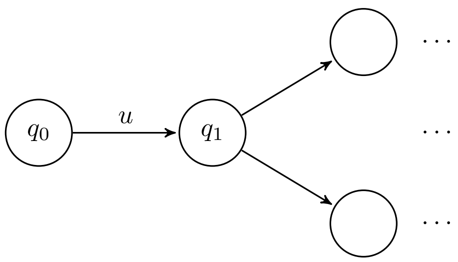
Proof. By Proposition 2.8, the automaton $\mathcal{S}(H)$ has n vertices and $n\#A$ positive edges. A spanning tree has $n-1$ positive edges, so $\operatorname{rk} H = n\#A - (n-1) = 1 + n(\#A - 1)$ by Proposition 2.6. □

Beautiful connections between finite index subgroups and certain classes of *bifix codes* — set of words none of which is a prefix or a suffix of another — have recently been unveiled by Berstel, De Felice, Perrin, Reutenauer and Rindone [6].

2.4 Conjugacy

We start now a brief discussion of conjugacy. Recall that the *outdegree* of a vertex q is the number of edges starting at q and the *geodesic distance* in a connected graph is the length of the shortest undirected path connecting two vertices.

Since the original generating set is always taken in reduced form, it follows easily that there is at most one vertex in a Stallings automaton having outdegree < 2: the basepoint q_0. Assuming that H is nontrivial, $\mathcal{S}(H)$ must always be of the form



where q_1 is the closest vertex to q_0 (in terms of geodesic distance) having outdegree > 2 (since there is at least one vertex having such outdegree). Note that $q_1 = q_0$ if q_0 has outdegree > 2 itself. We call $q_0 \xrightarrow{u}$ the *tail* (which is empty if $q_1 = q_0$) and the remaining subgraph the *core* of $\mathcal{S}(H)$.

Note that $\mathcal{S}(H)$, and its core, may be understood as follows. Consider the graph with vertex set $F_A/H = \{gH \mid g \in F_A\}$, with an edge from gH to agH for each generator $a \in A$. Then this graph, called the *Schreier graph* (see §24.1) of $H\backslash F_A$, consists of finitely many trees attached to the core of $\mathcal{S}(H)$.

Theorem 2.10. *There is an algorithm that decides whether or not two finitely generated subgroups of F_A are conjugate.*

Proof. Finitely generated subgroups G, H are conjugate if and only if the cores of $\mathcal{S}(G)$ and $\mathcal{S}(H)$ are equal (up to their basepoints). ☐

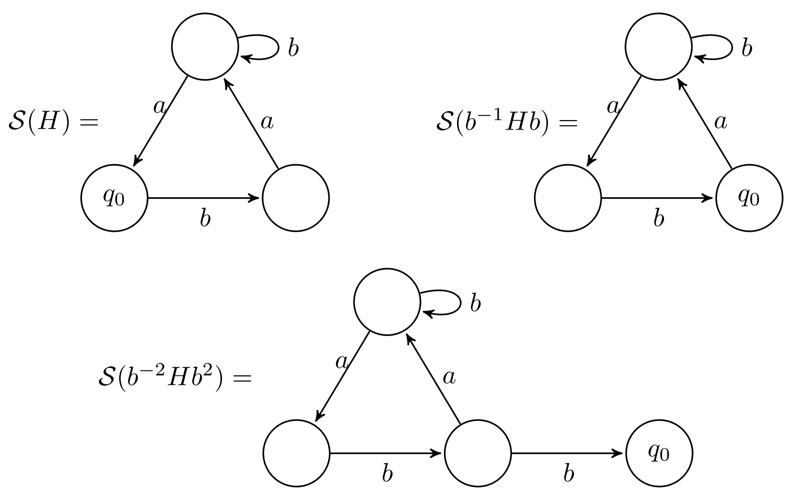
The Stallings automata of the conjugates of H can be obtained in the following ways: (1) declaring a vertex in the core C to be the basepoint; (2) gluing a tail to some vertex in the core C and taking its other endpoint to be the basepoint.

Note that the tail must be glued in some way that keeps the automaton inverse, so in particular this second type of operation can only be performed if the automaton is not complete, or equivalently, if H has infinite index. An immediate consequence is the following classical

Proposition 2.11. *A finite rank normal subgroup of a free group is either trivial or has finite index.*

Moreover, a finite index subgroup H is normal if and only if its Stallings automaton is *vertex-transitive*, that is, if all choices of basepoint yield the same automaton.

Example 2.5. Stallings automata of some conjugates of $H = \langle a^{-1}ba, ba^2 \rangle$:



We can also use the previous discussion on the structure of (finite) Stallings automata to provide them with an abstract characterization.

Proposition 2.12. *A finite inverse automaton with a basepoint is a Stallings automaton if and only if it has at most one vertex of outdegree 1: the basepoint.*

Proof. Indeed, for any such automaton we can take a spanning tree and use it to construct a basis for the subgroup as in the proof of Proposition 2.6. ☐

2.5 Further algebraic properties

The study of intersections of finitely generated subgroups of F_A provides further applications of Stallings automata. Howson's classical theorem admits a simple proof using the *direct product* of two Stallings automata; it is also an immediate consequence of Theorem 3.1 and Corollary 3.4(ii).

Theorem 2.13 (Howson). *If $H, K \leqslant_{f.g.} F_A$, then also $H \cap K \leqslant_{f.g.} F_A$.*

Stallings automata are also naturally related to the famous Hanna Neumann conjecture: given $H, K \leqslant_{f.g.} F_A$, then $\operatorname{rk}(H \cap K) - 1 \leqslant (\operatorname{rk} H - 1)(\operatorname{rk} K - 1)$. The conjecture arose in a paper of Hanna Neumann [34], where the inequality $\operatorname{rk}(H \cap K) - 1 \leqslant 2(\operatorname{rk} H - 1)(\operatorname{rk} K - 1)$ was also proved. In one of the early applications of Stallings' approach, Gersten provided an alternative geometric proof of Hanna Neumann's inequality [15].

A *free factor* of a free group F_A can be defined as a subgroup H generated by a subset of a basis of F_A. This is equivalent to saying that there exists a *free product decomposition* $F_A = H * K$ for some $K \leqslant F_A$.

Since the rank of a free factor never exceeds the rank of the ambient free group, it is easy to construct examples of subgroups which are not free factors: it follows easily from Proposition 2.6 that any free group of rank $\geqslant 2$ can have subgroups of arbitrary finite rank (and even infinite countable).

The problem of identifying free factors has a simple solution based on Stallings automata [50]: one must check whether or not a prescribed number of vertex identifications in the Stallings automaton can lead to a bouquet. However, the most efficient solution, due to Roig, Ventura and Weil [40], involves Whitehead automorphisms and will therefore be postponed to §23.2.7.

Given a morphism $\varphi : \mathcal{A} \to \mathcal{B}$ of inverse automata, let the *morphic image* $\varphi(\mathcal{A})$ be the subautomaton of $\mathcal{B}$ induced by the image by φ of all the successful paths of $\mathcal{A}$.

The following classical result characterizes the extensions of $H \leqslant_{f.g.} F_A$ contained in F_A. We present the proof from [32]:

Theorem 2.14 (Takahasi [55]). *Given $H \leqslant_{f.g.} F_A$, one can effectively compute finitely many extensions $K_1, \dots, K_m \leqslant_{f.g.} F_A$ of H such that the following conditions are equivalent for every $K \leqslant_{f.g.} F_A$:*

(i) *$H \leqslant K$;*

(ii) *K_i is a free factor of K for some $i \in \{1, \dots, m\}$.*

Proof. Let $\mathcal{A}_1, \dots, \mathcal{A}_m$ denote all the morphic images of $\mathcal{S}(H)$, up to isomorphism. Since a morphic image cannot have more vertices than the original automaton, there are only finitely many isomorphism classes. Moreover, it follows from Proposition 2.12 that, for $i = 1, \dots, m$, $\mathcal{A}_i = \mathcal{S}(K_i)$ for some $K_i \leqslant_{f.g.} F_A$. Since $L(\mathcal{S}(H)) \subseteq L(\mathcal{A}_i) = L(\mathcal{S}(K_i))$, it follows from Proposition 2.5 that $H \leqslant K_i$. Clearly, we can construct all $\mathcal{A}_i$ and therefore all K_i.

(i) $\Rightarrow$ (ii). If $H \leqslant K$, it follows from Stallings' construction that $L(\mathcal{S}(H)) \subseteq L(\mathcal{S}(K))$ and so there is a morphism $\varphi : \mathcal{S}(H) \to \mathcal{S}(K)$ by Proposition 2.2. Let $\mathcal{A}_i$

be, up to isomorphism, the morphic image of $\mathcal{S}(H)$ through φ. Since $\mathcal{A}_i = \mathcal{S}(K_i)$ is a subautomaton of $\mathcal{S}(K)$, it follows easily from Proposition 2.6 that K_i is a free factor of K: it suffices to take a spanning tree for $\mathcal{S}(K_i)$, extend it to a spanning tree for $\mathcal{S}(K)$, and the induced basis of K_i will be contained in the induced basis of K.

(ii) $\Rightarrow$ (i) is immediate. $\square$

An interesting research line related to this result is built on the concept of algebraic extension, introduced by Kapovich and Miasnikov [20], and inspired by the homonymous field-theoretical classical notion. Given $H \leqslant K \leqslant F_A$, we say that K is an *algebraic* extension of H if no proper free factor of K contains H. Miasnikov, Ventura and Weil [32] proved that the set of algebraic extensions of H is finite and effectively computable, and it constitutes the minimum set of extensions $K_1, \ldots, K_m$ satisfying the conditions of Theorem 2.14.

Consider a subgroup H of a group G. The *commensurator* of H in G, is

$$\mathrm{Comm}_G(H) = \{g \in G \mid H \cap H^g \text{ has finite index in } H \text{ and } H^g\}. \tag{2.4}$$

For example, the commensurator of $\mathrm{GL}_n(\mathbb{Z})$ in $\mathrm{GL}_n(\mathbb{R})$ is $\mathrm{GL}_n(\mathbb{Q})$.

The special case of finite-index extensions, $H \leqslant_{f.i.} K \leqslant F_A$ is of special interest, and can be interpreted in terms of commensurators. It can be proved (see [20, Lemma 8.7] and [52]) that every $H \leqslant_{f.g.} F_A$ has a maximum finite-index extension inside F_A, denoted by H_{fi}; and $H_{fi} = \mathrm{Comm}_{F_A}(H)$. Silva and Weil [52] proved that $\mathcal{S}(H_{fi})$ can be constructed from $\mathcal{S}(H)$ using a simple automata-theoretic algorithm:

(1) The standard minimization algorithm is applied to the core of $\mathcal{S}(H)$, *taking all vertices as final*.

(2) The original tail of $\mathcal{S}(H)$ is subsequently reinstated in this new automaton, at the appropriate vertex.

We present now an application of different type, involving transition monoids. It follows easily from the definitions that the transition monoid of a finite inverse automaton is always a *finite inverse monoid*. Given a group G, we say that a subgroup $H \leqslant G$ is *pure* if the implication

$$g^n \in H \Rightarrow g \in H \tag{2.5}$$

holds for all $g \in F_A$ and $n \geqslant 1$. If p is a prime, we say that H is *p-pure* if (2.5) holds when $(n, p) = 1$.

The next result is due to Birget, Margolis, Meakin and Weil, and is the only natural problem among applications of Stallings automata that is known so far to be PSPACE-complete [8].

Proposition 2.15. *For every $H \leqslant_{f.g.} F_A$, the following conditions hold:*

(i) *H is pure if and only if the transition monoid of $\mathcal{S}(H)$ is aperiodic.*

(ii) *H is p-pure if and only if the transition monoid of $\mathcal{S}(H)$ has no subgroups of order p.*

Proof. Both conditions in (i) are easily proved to be equivalent to the nonexistence in

$\mathcal{S}(H)$ of a cycle of the form

$$\begin{tikzcd} p \arrow[r, bend left, "u"] & q \arrow[l, bend left, "u^k"] \end{tikzcd} \qquad (k \geqslant 1, p \neq q)$$

where u can be assumed to be cyclically reduced. The proof of (ii) runs similarly. $\square$

2.6 Topological properties

We require for this subsection some basic topological concepts, which the reader can recover from Chapter 17.

For all $u, v \in F_A$, written in reduced form as elements of R_A, let $u \wedge v$ denote the longest common prefix of u and v. The *prefix metric* d on F_A is defined, for all $u, v \in F_A$, by

$$d(u,v) = \begin{cases} 2^{-|u \wedge v|-1} & \text{if } u \neq v \\ 0 & \text{if } u = v \end{cases}$$

It follows easily from the definition that d is an ultrametric on F_A, satisfying in particular the axiom

$$d(u,v) \leqslant \max\{d(u,w), d(w,v)\}\,.$$

The *completion* of this metric space is compact; its extra elements are *infinite reduced words* $a_1a_2a_3\ldots$, with all $a_i \in \widetilde{A}$, and constitute the *hyperbolic boundary* ∂F_A of F_A, see §24.1.5. Extending the operator $\wedge$ to $F_A \cup \partial F_A$ in the obvious way, it follows easily from the definitions that, for every infinite reduced word α and every sequence $(u_n)_n$ in F_A,

$$\alpha = \lim_{n\to+\infty} u_n \qquad \text{if and only if} \qquad \lim_{n\to+\infty} |\alpha \wedge u_n| = +\infty\,. \tag{2.6}$$

The next result shows that Stallings automata are given a new role in connection with the prefix metric. We denote by $\operatorname{cl} H$ the closure of H in the completion of F_A.

Proposition 2.16. *If $H \leqslant_{f.g.} F_A$, then $\operatorname{cl} H$ is the union of H with the set of all $\alpha \in \partial F_A$ that label paths in $\mathcal{S}(H)$ out of the basepoint.*

Proof. Since the topology of F_A is discrete, we have $\operatorname{cl} H \cap F_A = H$.

($\subseteq$): If $\alpha \in \partial F_A$ does not label a path in $\mathcal{S}(H)$ out of the basepoint, then $\{|\alpha \wedge h| : h \in H\}$ is finite and so no sequence of H can converge to α by (2.6).

($\supseteq$): Let $\alpha = a_1a_2a_3\cdots \in \partial F_A$, with $a_i \in \widetilde{A}$, label a path in $\mathcal{S}(H)$ out of the basepoint. Let m be the number of vertices of $\mathcal{S}(H)$. For every $n \geqslant 1$, there exists some word w_n of length $< m$ such that $a_1 \cdots a_n w_n \in H$. Now $\alpha = \lim_{n\to+\infty} a_1 \cdots a_n w_n$ by (2.6) and so $\alpha \in \operatorname{cl} H$. $\square$

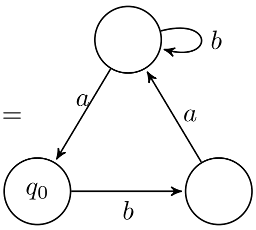
The *profinite topology* on F_A is defined in Chapter 17: for every $u \in F_A$, the collection $\{Ku \mid K \leqslant_{f.i.} F_A\}$ constitutes a basis of clopen neighbourhoods of u. In his seminal 1983 paper [54], Stallings gave an alternative proof of Marshall Hall's Theorem:

Theorem 2.17 (M. Hall). *Every finitely generated subgroup of F_A is closed for the profinite topology.*

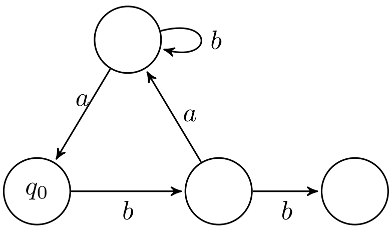
Proof. Fix $H \leqslant_{f.g.} F_A$ and let $u \in F_A \setminus H$ be written in reduced form as an element of R_A. In view of Proposition 2.5, u does not label a loop at the basepoint q_0 of $\mathcal{S}(H)$. If there is no path $q_0 \xrightarrow{u} \cdots$ in $\mathcal{S}(H)$, we add new edges to $\mathcal{S}(H)$ to get a finite inverse automaton $\mathcal{A}$ having a path $q_0 \xrightarrow{u} q \neq q_0$. Otherwise just take $\mathcal{A} = \mathcal{S}(H)$. Next add new edges to $\mathcal{A}$ to get a finite complete inverse automaton $\mathcal{B}$. In view of Propositions 2.8 and 2.12, we have $\mathcal{B} = \mathcal{S}(K)$ for some $K \leqslant_{f.i.} F_A$. Hence Ku is open and contains u. Since $H \cap Ku \neq \emptyset$ yields $u \in K^{-1}H = K$, contradicting Proposition 2.5, it follows that $H \cap Ku = \emptyset$ and so H is closed as claimed. □

Example 2.6. We consider the above construction for $H = \langle a^{-1}ba, ba^2 \rangle$ and $u = b^2$:

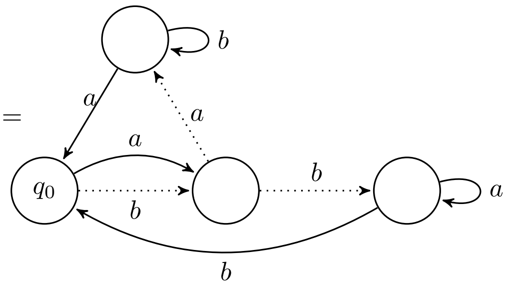
$\mathcal{S}(H) =$



$\mathcal{A} =$



$\mathcal{B} =$



If we take the spanning tree defined by the dotted lines in $\mathcal{B}$, it follows from Proposition 2.6 that

$$K = \langle ba^{-1}, b^3, b^2ab^{-2}, ba^2, baba^{-1}b^{-1} \rangle$$

is a finite index subgroup of F_2 such that $H \cap Kb^2 = \emptyset$.

We recall that a group G is *residually finite* if its finite index subgroups have trivial intersection. Considering the trivial subgroup in Theorem 2.17, we deduce

Corollary 2.18. *F_A is residually finite.*

We remark that Ribes and Zalessky extended Theorem 2.17 to products of finitely many finitely generated subgroups of F_A, see [38]. This result is deeply connected to the solution of Rhodes' Type II conjecture, see [37, Chapter 4].

If $\mathbf{V}$ denotes a pseudovariety of finite groups (see Chapter 16), the *pro-$\mathbf{V}$ topology* on F_A is defined by considering that each $u \in F_A$ has

$$\{Ku \mid K \trianglelefteq_{f.i.} F_A,\ F_A/K \in \mathbf{V}\}$$

as a basis of clopen neighbourhoods. The closure for the pro-**V** topology of $H \leqslant_{f.g} F_A$ can be related to an extension property of $\mathcal{S}(H)$, and Margolis, Sapir and Weil used automata to prove that efficient computation can be achieved for the pseudovarieties of finite p-groups and finite nilpotent groups [28]. The original computability proof for the p-group case is due to Ribes and Zalessky [39].

2.7 Dynamical properties

We shall mention briefly some examples of applications of Stallings automata to the study of endomorphism dynamics, starting with Gersten's solution of the subgroup orbit problem [16].

The subgroup orbit problem consists in finding an algorithm to decide, for given $H, K \leqslant_{f.g.} F_A$, whether or not $K = \varphi(H)$ for some automorphism φ of F_A. Equivalently, this can be described as deciding whether or not the automorphic orbit of a finitely generated subgroup is recursive.

Gersten's solution adapts to the context of Stallings automata Whitehead's idea to solve the orbit problem for words [59]. Whitehead's proof relies on a suitable decomposition of automorphisms as products of elementary factors (which became known as *Whitehead automorphisms*), and on using these as a tool to compute the elements of minimum length in the automorphic orbit of the word. In the subgroup case, word length is replaced by the number of vertices of the Stallings automaton.

The most efficient solution to the problem of identifying free factors [40], mentioned in §23.2.5, also relies on this approach: $H \leqslant_{f.g.} F_A$ is a free factor if and only if the Stallings automaton of some automorphic image of H has a single vertex (that is, a bouquet).

Another very nice application is given by the following theorem of Goldstein and Turner [17]:

Theorem 2.19. *The fixed point subgroup of an endomorphism of F_A is finitely generated.*

Proof. Let φ be an endomorphism of F_A. For every $u \in F_A$, define $Q(u) = \varphi(u)u^{-1}$. We define a potentially infinite automaton $\mathcal{A}$ by taking

$$\{Q(u) \mid u \in F_A\} \subseteq F_A$$

as the vertex set, all edges of the form $Q(u) \xrightarrow{a} Q(au)$ with $u \in F_A,\ a \in \widetilde{A}$, and fixing $\mathbb{1}$ as the basepoint. Then $\mathcal{A}$ is a well-defined inverse automaton.

Next we take $\mathcal{B}$ to be the subautomaton of $\mathcal{A}$ obtained by retaining only those vertices and edges that lie in successful paths labelled by reduced words. Clearly, $\mathcal{B}$ is still an inverse automaton, and it is easy to check that it must be the Stallings automaton of the fixed point subgroup of φ.

It remains to be proved that $\mathcal{B}$ is finite. We define a subautomaton $\mathcal{C}$ of $\mathcal{B}$ by removing exactly one edge among each inverse pair

$$Q(u) \xrightarrow{a} Q(au), \quad Q(au) \xrightarrow{a^{-1}} Q(u)$$

with $a \in A$ as follows: if a^{-1} is the last letter of $Q(au)$, we remove $Q(u) \xrightarrow{a} Q(au)$; otherwise, we remove $Q(au) \xrightarrow{a^{-1}} Q(u)$.

Let M denote the maximum length of the image of a letter by φ. We claim that, whenever $|Q(v)| > 2M$, the vertex $Q(v)$ has outdegree at most 1.

Indeed, if $Q(v)\xrightarrow{a^{-1}}Q(a^{-1}v)$ is an edge in $\mathcal{C}$ for $a \in A$, then a^{-1} is the last letter of $Q(v)$. On the other hand, if $Q(v)\xrightarrow{b}Q(bv)$ is an edge in $\mathcal{C}$ for $b \in A$, then b^{-1} is not the last letter of $Q(bv)$. Since $Q(bv) = \varphi(b)Q(v)b^{-1}$ and $|Q(v)| > 2|\varphi(b)|$, then b must be the last letter of $Q(v)$ in this case. Since $Q(v)$ has at most one last letter, it follows that its outdegree is at most 1.

Let $\mathcal{D}$ be a finite subautomaton of $\mathcal{C}$ containing all vertices $Q(v)$ such that $|Q(v)| \leqslant 2M$. Suppose that $p\longrightarrow q$ is an edge in $\mathcal{C}$ not belonging to $\mathcal{D}$. Since $p\longrightarrow q$, being an edge of $\mathcal{B}$, must lie in some reduced path, and by the outdegree property of $\mathcal{C}$, it is easy to see that there exists some path in $\mathcal{C}$ of the form

$$p'\longrightarrow p\longrightarrow q\longrightarrow r\longleftarrow r'$$

where p', r' are vertices in $\mathcal{D}$. Since there are only finitely many directed paths out of $\mathcal{D}$, it follows that $\mathcal{C}$ is finite and so is $\mathcal{B}$. Therefore the fixed point subgroup of φ is finitely generated. $\square$

Note that this proof is not by any means constructive. Indeed, the only known algorithm for computing the fixed point subgroup of a free group automorphism is due to Maslakova [31] and relies on the sophisticated *train track* theory of Bestvina and Handel [7] and other algebraic geometry tools. The general endomorphism case remains open.

Stallings automata were also used by Ventura in the study of various properties of fixed subgroups, considering in particular arbitrary families of endomorphisms [57, 30] (see also [58]).

Automata also play a part in the study of *infinite fixed points*, taken over the continuous extension of a monomorphism to the hyperbolic boundary (see for example [49]).

3 Rational and recognizable subsets

Rational subsets generalize the notion of finitely generated from subgroups to arbitrary subsets of a group, and can be quite useful in establishing inductive procedures that need to go beyond the territory of subgroups. Similarly, recognizable subsets extend the notion of finite index subgroups. Basic properties and results can be found in [5] or [43].

We consider a finitely generated group $G = \langle A \rangle$, with the canonical map $\pi : F_A \to G$. A subset of G is *rational* if it is the image by $\rho = \pi\theta$ of a rational subset of $\widetilde{A}^*$, and is *recognizable* if its full preimage under ρ is rational in $\widetilde{A}^*$.

For every group G, the classes $\operatorname{Rat} G$ and $\operatorname{Rec} G$ satisfy the following closure properties:

- $\operatorname{Rat} G$ is (effectively) closed under union, product, star, morphisms, inversion, subgroup generating.
- $\operatorname{Rec} G$ is (effectively) closed under boolean operations, translation, product, star, inverse morphisms, inversion, subgroup generating.

Kleene's Theorem is not valid for groups: $\operatorname{Rat} G = \operatorname{Rec} G$ if and only if G is finite. However, if the class of rational subsets of G possesses some extra algorithmic properties, then many decidability/constructibility results can be deduced for G. Two properties are particularly coveted for $\operatorname{Rat} G$:

- (effective) closure under complement (yielding closure under all the boolean operations);
- decidable membership problem for arbitrary rational subsets.

In these cases, one may often solve problems (e.g. equations, or systems of equations) whose statement lies far out of the rational universe, by proving that the solution is a rational set.

3.1 Rational and recognizable subgroups

We start by some basic, general facts. The following result is essential to connect language theory to group theory.

Theorem 3.1 (Anisimov and Seifert). *A subgroup H of a group G is rational if and only if H is finitely generated.*

Proof. ($\Rightarrow$): Let H be a rational subgroup of G and let $\pi : F_A \to G$ denote a morphism. Then there exists a finite $\widetilde{A}$-automaton $\mathcal{A}$ such that $H = \rho(L(\mathcal{A}))$. Assume that $\mathcal{A}$ has m vertices and let X consist of all the words in $\rho^{-1}(H)$ of length $< 2m$. Since A is finite, so is X. We claim that $H = \langle \rho(X) \rangle$. To prove it, it suffices to show that

$$u \in L(\mathcal{A}) \Rightarrow \rho(u) \in \langle \rho(X) \rangle \tag{3.1}$$

holds for every $u \in \widetilde{A}^*$. We use induction on $|u|$. By definition of X, (3.1) holds for words of length $< 2m$. Assume now that $|u| \geqslant 2m$ and (3.1) holds for shorter words. Write $u = vw$ with $|w| = m$. Then there exists a path

$$\to q_0 \xrightarrow{v} q \xrightarrow{z} t \to$$

in $\mathcal{A}$ with $|z| < m$. Thus $vz \in L(\mathcal{A})$ and by the induction hypothesis $\rho(vz) \in \langle \rho(X) \rangle$. On the other hand, $|z^{-1}w| < 2m$ and $\rho(z^{-1}w) = \rho(z^{-1}v^{-1})\rho(vw) \in H$, hence $z^{-1}w \in X$ and so $\rho(u) = \rho(vz)\rho(z^{-1}w) \in \langle \rho(X) \rangle$, proving (3.1) as required.

($\Leftarrow$) is trivial. □

It is an easier task to characterize recognizable subgroups:

Proposition 3.2. *A subgroup H of a group G is recognizable if and only if it has finite index.*

Proof. ($\Rightarrow$): In general, a recognizable subset of G is of the form NX, where $N \trianglelefteq_{f.i.} G$ and $X \subseteq G$ is finite. If $H = NX$ is a subgroup of G, then $N \subseteq H$ and so H has finite index as well.

($\Leftarrow$): This follows from the well-known fact that every finite index subgroup H of G contains a finite index normal subgroup N of G, namely $N = \bigcap_{g \in G} gHg^{-1}$. Since N has finite index, H must be of the form NX for some finite $X \subseteq G$. □

3.2 Benois' Theorem

The central result in this subsection is Benois' Theorem, the cornerstone of the whole theory of rational subsets of free groups:

Theorem 3.3 (Benois).

(i) *If $L \subseteq \widetilde{A}^*$ is rational, then $\overline{L}$ is also rational, and can be effectively constructed from L.*

(ii) *A subset of R_A is a rational language as a subset of $\widetilde{A}^*$ if and only if it is rational as a subset of F_A.*

We illustrate this in the case of finitely generated subgroups: temporarily calling "Benois automata" those automata recognizing rational subsets of R_A, we may convert them to Stallings automata by "folding" them, at the same time making sure they are inverse automata. Given a Stallings automaton, one intersects it with R_A to obtain a Benois automaton.

Proof. (i) Let $\mathcal{A} = (Q, \widetilde{A}, E, I, T)$ be a finite automaton recognizing L. We define a sequence $(\mathcal{A}_n)_n$ of finite automata with ε-transitions as follows. Let $\mathcal{A}_0 = \mathcal{A}$. Assuming that $\mathcal{A}_n = (Q, \widetilde{A}, E_n, I, T)$ is defined, we consider all instances of ordered pairs $(p, q) \in Q \times Q$ such that

$$\text{there exists a path } p \xrightarrow{aa^{-1}} q \text{ in } \mathcal{A}_n \text{ for some } a \in \widetilde{A}, \text{ but no path } p \xrightarrow{1} q. \qquad \text{(P)}$$

Clearly, there are only finitely many instances of (P) in $\mathcal{A}_n$. We define E_{n+1} to be the union of E_n with all the new edges $(p, 1, q)$, where $(p, q) \in Q \times Q$ is an instance of (P). Finally, we define $\mathcal{A}_{n+1} = (Q, \widetilde{A}, E_{n+1}, I, T)$. In particular, note that $\mathcal{A}_n = \mathcal{A}_{n+k}$ for every $k \geqslant 1$ if there are no instances of (P) in $\mathcal{A}_n$.

Since Q is finite, the sequence $(\mathcal{A}_n)_n$ is ultimately constant, say after reaching $\mathcal{A}_m$. We claim that

$$\overline{L} = L(\mathcal{A}_m) \cap R_A \,. \qquad (3.2)$$

Indeed, take $u \in L$. There exists a sequence of words $u = u_0, u_1, \ldots, u_{k-1}, u_k = \overline{u}$ where each term is obtained from the preceding one by erasing a factor of the form aa^{-1} for some $a \in \widetilde{A}$. A straightforward induction shows that $u_i \in L(\mathcal{A}_i)$ for $i = 0, \ldots, k$, since the existence of a path $p \xrightarrow{aa^{-1}} q$ in $\mathcal{A}_i$ implies the existence of a path $p \xrightarrow{1} q$ in $\mathcal{A}_{i+1}$. Hence $\overline{u} = u_k \in L(\mathcal{A}_k) \subseteq L(\mathcal{A}_m)$ and it follows that $\overline{L} \subseteq L(\mathcal{A}_m) \cap R_A$.

For the opposite inclusion, we start by noting that any path $p \xrightarrow{u} q$ in $\mathcal{A}_{i+1}$ can be lifted to a path $p \xrightarrow{v} q$ in $\mathcal{A}_i$, where v is obtained from u by inserting finitely many factors of the form aa^{-1}. It follows that

$$\overline{L(\mathcal{A}_m)} = \overline{L(\mathcal{A}_{m-1})} = \cdots = \overline{L(\mathcal{A}_0)} = \overline{L}$$

and so $L(\mathcal{A}_m) \cap R_A \subseteq \overline{L(\mathcal{A}_m)} = \overline{L}$. Thus (3.2) holds.

Since

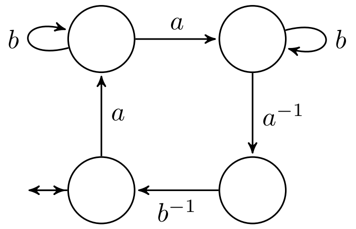
$$R_A = \widetilde{A}^* \setminus \bigcup_{a \in \widetilde{A}} \widetilde{A}^* aa^{-1} \widetilde{A}^*$$

is obviously rational, and the class of rational languages is closed under intersection, it follows that $\overline{L}$ is rational. Moreover, we can effectively compute the automaton $\mathcal{A}_m$ and a finite automaton recognizing R_A, hence the direct product construction can be used to construct a finite automaton recognizing the intersection $\overline{L} = L(\mathcal{A}_m) \cap R_A$.

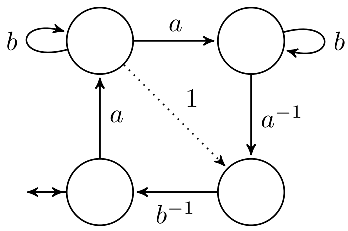
(ii) Consider $X \subseteq R_A$. If $X \in \operatorname{Rat} \widetilde{A}^*$, then $\theta(X) \in \operatorname{Rat} F_A$ and so X is rational as a subset of F_A.

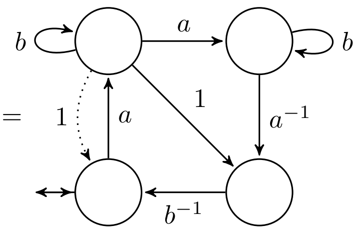
Conversely, if X is rational as a subset of F_A, then $X = \theta(L)$ for some $L \in \operatorname{Rat} \widetilde{A}^*$. Since $X \subseteq R_A$, we get $X = \overline{L}$. Now part (i) yields $\overline{L} \in \operatorname{Rat} \widetilde{A}^*$ and so $X \in \operatorname{Rat} \widetilde{A}^*$ as required. $\square$

Example 3.1. Let $\mathcal{A} = \mathcal{A}_0$ be depicted by



We get

$\mathcal{A}_1 =$


$\mathcal{A}_2 = \mathcal{A}_3 =$


and we can then proceed to compute $\overline{L} = L(\mathcal{A}_2) \cap R_2$.

The following result summarizes some of the most direct consequences of Benois' Theorem:

Corollary 3.4.
(i) F_A *has decidable rational subset membership problem.*
(ii) $\operatorname{Rat} F_A$ *is closed under the boolean operations.*

Proof. (i) Given $X \in \operatorname{Rat} F_A$ and $u \in F_A$, write $X = \theta(L)$ for some $L \in \operatorname{Rat} \widetilde{A}^*$. Then $u \in X$ if and only if $\overline{u} \in \overline{X} = \overline{L}$. By Theorem 3.3(i), we may construct a finite automaton recognizing $\overline{L}$ and therefore decide whether or not $\overline{u} \in \overline{L}$.

(ii) Given $X \in \operatorname{Rat} F_A$, we have $\overline{F_A \setminus X} = R_A \setminus \overline{X}$ and so $F_A \setminus X \in \operatorname{Rat} F_A$ by Theorem 3.3. Therefore $\operatorname{Rat} F_A$ is closed under complement.

Since $\operatorname{Rat} F_A$ is trivially closed under union, it follows from De Morgan's laws that it is closed under intersection as well. $\square$

Note that we can associate algorithms to these boolean closure properties of $\operatorname{Rat} F_A$ in a constructive way. We remark also that the proof of Theorem 3.3 can be clearly adapted to more general classes of rewriting systems (see [9]). Theorem 3.3 and Corollary 3.4 have been generalized several times by Benois herself [4] and by Sénizergues, who obtained the most general versions. Sénizergues' results [44] hold for *rational length-reducing left basic confluent* rewriting systems and remain valid for the more general notion of *controlled* rewriting system.

3.3 Rational versus recognizable

Since F_A is a finitely generated monoid, it follows that every recognizable subset of F_A is rational [5, Proposition III.2.4]. We turn to the problem of deciding which rational subsets of F_A are recognizable. The first proof, using rewriting systems, is due to Sénizergues [45] but we follow the shorter alternative proof from [48], where a third alternative proof, of a more combinatorial nature, was also given.

Given a subset X of a group G, we define the *right stabilizer* of X to be the submonoid of G defined by

$$R(X) = \{g \in G \mid Xg \subseteq X\}\,.$$

Next let

$$K(X) = R(X) \cap (R(X))^{-1} = \{g \in G \mid Xg = X\}$$

be the largest subgroup of G contained in $R(X)$ and let

$$N(X) = \bigcap_{g \in G} gK(X)g^{-1}$$

be the largest normal subgroup of G contained in $K(X)$, and therefore in $R(X)$.

Lemma 3.5. *A subset X of a group G is recognizable if and only if $K(X)$ is a finite index subgroup of G.*

In fact, the Schreier graph (see §24.1) of $K(X)\backslash G$ is the underlying graph of an automaton recognizing X, and $G/N(X)$ is the syntactic monoid of X.

Proof. ($\Rightarrow$): If $X \subseteq G$ is recognizable, then $X = NF$ for some $N \trianglelefteq_{f.i.} G$ and $F \subseteq G$ finite. Hence $N \subseteq R(X)$ and so $N \subseteq K(X)$ since $N \leqslant G$. Since N has finite index in G, so does $K(X)$.

($\Leftarrow$): If $K(X)$ is a finite index subgroup of G, so is $N = N(X)$. Indeed, a finite index subgroup has only finitely many conjugates (having also finite index) and a finite intersection of finite index subgroups is easily checked to have finite index itself.

Therefore it suffices to show that $X = FN$ for some finite subset F of G. Since N has finite index, the claim follows from $XN = X$, in turn an immediate consequence of $N \subseteq R(X)$. □

Proposition 3.6. *It is decidable whether or not a rational subset of F_A is recognizable.*

Proof. Take $X \in \operatorname{Rat} F_A$. In view of Lemma 3.5 and Proposition 2.8, it suffices to show that $K(X)$ is finitely generated and effectively computable.

Given $u \in F_A$, we have

$$u \notin R(X) \Leftrightarrow Xu \nsubseteq X \Leftrightarrow Xu \cap (F_A \setminus X) \neq \emptyset \Leftrightarrow u \in X^{-1}(F_A \setminus X),$$

hence

$$R(X) = F_A \setminus (X^{-1}(F_A \setminus X)) .$$

It follows easily from the fact that the class of rational languages is closed under reversal and morphisms, combined with Theorem 3.3(ii), that $X^{-1} \in \operatorname{Rat} F_A$. Since $\operatorname{Rat} F_A$ is trivially closed under product, it follows from Corollary 3.4 that $R(X)$ is rational and effectively computable, and so is $K(X) = R(X) \cap (R(X))^{-1}$. By Theorem 3.1, the subgroup $K(X)$ is finitely generated and the proof is complete. □

These results are related to the Sakarovitch conjecture [42], which states that every rational subset of F_A must be either recognizable or *disjunctive*: a subset X of a monoid M is disjunctive if it has trivial syntactic congruence, or equivalently, if any morphism $\varphi : M \to M'$ recognizing X is necessarily injective.

In the group case, it follows easily from the proof of the direct implication of Lemma 3.5 that the projection $G \to G/N$ recognizes $X \subseteq G$ if and only if $N \subseteq N(X)$. Thus X is disjunctive if and only if $N(X)$ is the trivial subgroup.

The Sakarovitch conjecture was first proved in [45], but once again we follow the shorter alternative proof from [48]:

Theorem 3.7 (Sénizergues). *A rational subset of F_A is either recognizable or disjunctive.*

Proof. Since the only subgroups of $\mathbb{Z}$ are the trivial subgroup and finite index subgroups, we may assume that $\#A > 1$.

Take $X \in \operatorname{Rat} F_A$. By the proof of Proposition 3.6, the subgroup $K(X)$ is finitely generated. In view of Lemma 3.5, we may assume that $K(X)$ is not a finite index subgroup. Thus $\mathcal{S}(K(X))$ is not complete by Proposition 2.8. Let q_0 denote the basepoint of $\mathcal{S}(K(X))$. Since $\mathcal{S}(K(X))$ is not complete, $q_0 \cdot u$ is undefined for some reduced word u.

Let w be an arbitrary nonempty reduced word. We must show that $w \notin N(X)$. Suppose otherwise. Since u, w are reduced and $\#A > 1$, there exist enough letters to make sure that there is some word $v \in R_A$ such that $uvwv^{-1}u^{-1}$ is reduced. Now $w \in N(X)$, hence $uvwv^{-1}u^{-1} \in N(X) \subseteq K(X)$ by normality. Since $uvwv^{-1}u^{-1}$ is reduced, it follows from Proposition 2.5 that $uvwv^{-1}u^{-1}$ labels a loop at q_0 in $\mathcal{S}(K(X))$, contradicting $q_0 \cdot u$ being undefined. Thus $w \notin N(X)$ and so $N(X) = 1$. Therefore X is disjunctive as required. □

3.4 Beyond free groups

Let $\pi : F_A \twoheadrightarrow G$ be a morphism onto a group G. We consider the *word problem submonoid* of a group G, defined as

$$W_\pi(G) = (\pi\theta)^{-1}(\mathbb{1}). \tag{3.3}$$

Proposition 3.8. *The language $W_\pi(G)$ is rational if and only if G is finite.*

Proof. If G is finite, it is easy to check that $W_\pi(G)$ is rational by viewing the Cayley graph of G (see §24.1) as an automaton. Conversely, if $W_\pi(G)$ is rational, then $\pi^{-1}(\mathbb{1})$ is a finitely generated normal subgroup of F_A, either finite index or trivial by the proof of Theorem 3.7. It is well known that the *Dyck language* $D_A = \theta^{-1}(\mathbb{1})$ is not rational if $\#A > 0$, thus it follows easily that $\pi^{-1}(\mathbb{1})$ has finite index and therefore G must be finite. $\square$

How about groups with context-free $W_\pi(G)$? A celebrated result by Muller and Schupp [33], with a contribution by Dunwoody [13], relates them to *virtually free groups*: these are groups with a free subgroup of finite index.

As usual, we focus on the case of G being finitely generated. We claim that G has a *normal* free subgroup F_A of finite index, with A finite. Indeed, letting F be a finite-index free subgroup of G, it suffices to take $F' = \bigcap_{g\in G} gFg^{-1}$. Since F has finite index, so does F', see the proof of Lemma 3.5. Taking a morphism $\pi : F_B \to G$ with B finite, we get from Corollary 2.9 that $\pi^{-1}(F') \leqslant_{f.i.} F_B$ is finitely generated, so F' is itself finitely generated. Finally, F' is a subgroup of F, so F' is still free by Theorem 2.7, and we can write $F' = F_A$.

We may therefore decompose G as a finite disjoint union of the form

$$G = F_A b_0 \cup F_A b_1 \cup \cdots \cup F_A b_m, \qquad \text{with } b_0 = 1. \tag{3.4}$$

Theorem 3.9 (Muller & Schupp). *The language $W_\pi(G)$ is context-free if and only if G is virtually free.*

Sketch of proof. If G is virtually free, the rewriting system implicit in (3.4) provides a rational transduction between $W_\pi(G)$ and D_A.

The converse implication can be proved by arguing geometrical properties of the Cayley graph of G such as in Chapter 24; briefly said, one deduces from the context-freeness of $W_\pi(G)$ that the Cayley graph of G is close (more precisely, quasi-isometric) to a tree. $\square$

It follows that virtually free groups have decidable word problem. In Chapter 24, we shall discuss the word problem for more general classes of groups using other techniques.

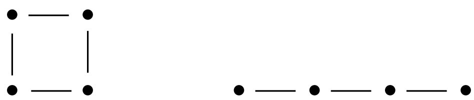
Grunschlag proved that every rational (respectively recognizable) subset of a virtually free group G decomposed as in (3.4) admits a decomposition as a finite union $X_0 b_0 \cup \cdots \cup X_m b_m$, where the X_i are rational (respectively recognizable) subsets of F_A, see [18]. Thus basic results such as Corollary 3.4 or Proposition 3.6 can be extended to virtually free groups (see [18, 47]). Similar generalizations can be obtained for free abelian groups of finite rank [47].

The fact that the strong properties of Corollary 3.4 do hold for both free groups and free abelian groups suggests considering the case of graph groups (also known as free partially abelian groups or right angled Artin groups), where we admit partial commutation between letters.

An *independence graph* is a finite undirected graph (A, I) with no loops, that is, I is a symmetric anti-reflexive relation on A. The *graph group* $G(A, I)$ is the quotient $F_A/\sim$, where $\sim$ denotes the congruence generated by the relation

$$\{(ab, ba) \mid (a, b) \in I\}.$$

On both extremes, we have $F_A = G(A, \emptyset)$ and the free abelian group on A, which corresponds to the complete graph on A. These turn out to be particular cases of *transitive* *forests*. We can say that (A, I) is a transitive forest if it has no induced subgraph of either of the following forms:



C_4 P_4

We recall that an induced subgraph of (A, I) is formed by a subset of vertices $A' \subseteq A$ and all the edges in I connecting vertices from A'.

The following difficult theorem, a group-theoretic version of a result on trace monoids by Aalbersberg and Hoogeboom [1], was proved in [23]:

Theorem 3.10 (Lohrey & Steinberg). *Let (A, I) be an independence graph. Then $G(A, I)$ has decidable rational subset membership problem if and only if (A, I) is a transitive forest.*

They also proved that these conditions are equivalent to decidability of the membership problem for finitely generated submonoids. Such a 'bad' $G(A, I)$ gives an example of a finitely presented group with a decidable generalized word problem that does not have a decidable membership problem for finitely generated submonoids.

It follows from Theorem 3.10 that any group containing a direct product of two free monoids has undecidable rational subset membership problem, a fact that can be directly deduced from the undecidability of the Post correspondence problem.

Other positive results on rational subsets have been obtained for graphs of groups, HNN extensions and amalgamated free products by Kambites, Silva and Steinberg [19], or Lohrey and Sénizergues [22]. Lohrey and Steinberg proved recently that the rational subset membership problem is recursively equivalent to the finitely generated submonoid membership problem for groups with two or more ends [24].

With respect to closure under complement, Lohrey and Sénizergues [22] proved that the class of groups for which the rational subsets form a boolean algebra is closed under HNN extension and amalgamated products over finite groups.

On the negative side, Bazhenova proved that rational subsets of finitely generated nilpotent groups do not form a boolean algebra, unless the group is virtually abelian [3]. Moreover, Romanʹkov proved in [41], via a reduction from Hilbert's 10th problem, that the rational subset membership problem is undecidable for free nilpotent groups of any class $\geqslant 2$ of sufficiently large rank.

Last but not least, we should mention that Stallings' construction was successfully generalized to prove results on both graph groups (by Kapovich, Miasnikov and Weidmann [21]) and amalgamated free products of finite groups (by Markus-Epstein [29]).

3.5 Rational solution sets and rational constraints

In this final subsection we make a brief incursion in the brave new world of rational constraints. Rational subsets provide group theorists with two main assets:

- A concept which generalizes finite generation for subgroups and is much more fit to stand most induction procedures.
- A systematic way of looking for solutions of the *right type* in the context of equations of many sorts.

This second feature leads us to the notion of *rational constraint*, when we restrict the set of potential solutions to some rational subset. And there is a particular combination of circumstances that can ensure the success of this strategy: if $\mathrm{Rat}\, G$ is closed under intersection and we can prove that the solution set of problem P is an effectively computable rational subset of G, then we can solve problem P with any rational constraint.

An early example is the adaptation by Margolis and Meakin of Rabin's language and Rabin's tree theorem to free groups, where first-order formulae provide rational solution sets [27]. The logic language considered here is meant to be applied to words, seen as models, and consists basically of unary predicates that associate letters to positions in each word, as well as a binary predicate for position ordering. Margolis and Meakin used this construction to solve problems in combinatorial inverse semigroup theory [27].

Diekert, Gutierrez and Hagenah proved that the existential theory of systems of equations with rational constraints is solvable over a free group [11]. Working basically on a free monoid with involution, and adapting Plandowski's approach [35] in the process, they extended the classical result of Makanin [25] to include rational constraints, with much lower complexity as well.

The proof of this deep result is well out of scope here, but its potential applications are immense. Group theorists are only starting to discover its full strength.

The results in [22] can be used to extend the existential theory of equations with rational constraints to virtually free groups, a result that follows also from Dahmani and Guirardel's recent paper on equations over hyperbolic groups with quasi-convex rational constraints [10]. Equations over graph groups with a restricted class of rational constraints were also successfully considered by Diekert and Lohrey [12].

A somewhat exotic example of computation of a rational solution set arises in the problem of determining which automorphisms of F_2 (if any) carry a given word into a given finitely generated subgroup. The full solution set is recognized by a finite automaton; its vertices are themselves structures named "finite truncated automata" [51].

References

[1] I. J. Aalbersberg and H. J. Hoogeboom. Characterizations of the decidability of some problems for regular trace languages. *Math. Systems Theory*, 22:1–19, 1989. 24

[2] Algebraic Cryptography Center. *CRAG – the Cryptography and Groups Software Library*, 2010. 4

[3] G. A. Bazhenova. On rational sets in finitely generated nilpotent groups. *Algebra and Logic*, 39(4):215–223, 2000. Translated from *Algebra i Logika*, 39:379–394, 2000. 24

[4] M. Benois. Descendants of regular language in a class of rewriting systems: algorithm and complexity of an automata construction. In *Rewriting techniques and applications*, volume 256 of *Lecture Notes in Comput. Sci.*, pages 121–132. Springer-Verlag, 1987. 21

[5] J. Berstel. *Transductions and context-free languages*. B. G. Teubner, 1979. 17, 21

[6] J. Berstel, C. De Felice, D. Perrin, C. Reutenauer, and G. Rindone. Bifix codes and sturmian words. preprint, 2010. `arXiv.org/pdf/1011.5369v2`. 10

[7] M. Bestvina and M. Handel. Train tracks and automorphisms of free groups. *Ann. Math.*, 135:1–51, 1992. 17

[8] J.-C. Birget, S. W. Margolis, J. C. Meakin, and P. Weil. PSPACE-complete problems for subgroups of free groups and inverse finite automata. *Theoret. Comput. Sci.*, 242(1-2):247–281, 2000. 13

[9] R. V. Book and F. Otto. *String-Rewriting Systems*. Springer-Verlag, 1993. 3, 20

[10] F. Dahmani and V. Guirardel. Foliations for solving equations in groups: free, virtually free, and hyperbolic groups. *J. Topology*, 3(2):343–404, 2010. 25

[11] V. Diekert, C. Gutierrez, and C. Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE-complete. *Inform. Comput.*, 202(2):105–140, 2005. 25

[12] V. Diekert and M. Lohrey. Word equations over graph products. *Internat. J. Algebra Comput.*, 18(3):493–533, 2008. 25

[13] M. J. Dunwoody. The accessibility of finitely presented groups. *Invent. Math.*, 81(3):449–457, 1985. 23

[14] The GAP Group. *GAP – Groups, Algorithms, and Programming, Version 4.4.12*, 2008. 4

[15] S. M. Gersten. Intersections of finitely generated subgroups of free groups and resolutions of graphs. *Inventiones Math.*, 71(3):567–591, 1983. 12

[16] S. M. Gersten. On whitehead's algorithm. *Bull. Amer. Math. Soc.*, 10(2):281–284, 1984. 16

[17] R. Z. Goldstein and E. C. Turner. Fixed subgroups of homomorphisms of free groups. *Bull. Lond. Math. Soc.*, 18(5):468–470, 1986. 16

[18] Z. Grunschlag. *Algorithms in geometric group theory*. PhD thesis, University of California at Berkeley, 1999. 23

[19] M. Kambites, P. V. Silva, and B. Steinberg. On the rational subset problem for groups. *J. Algebra*, 309(2):622–639, 2007. 24

[20] I. Kapovich and A. Myasnikov. Stallings foldings and subgroups of free groups. *J. Algebra*, 248(2):608–668, 2002. 4, 13

[21] I. Kapovich, R. Weidmann, and A. Miasnikov. Foldings, graphs of groups and the membership problem. *Internat. J. Algebra Comput.*, 15(1):95–128, 2005. 24

[22] M. Lohrey and G. Sénizergues. Rational subsets in HNN-extensions and amalgamated products. *Internat. J. Algebra Comput.*, 18(1):111–163, 2008. 24, 25

[23] M. Lohrey and B. Steinberg. The submonoid and rational subset membership problems for graph groups. *J. Algebra*, 320(2):728–755, 2008. 24

[24] M. Lohrey and B. Steinberg. Submonoids and rational subsets of groups with infinitely many ends. To appear in *J. Algebra*, 2009. 24

[25] G. S. Makanin. Equations in a free group. *Math. USSR Izv.*, 21:483–546, 1983. Translated from *Izv. Akad. Nauk. SSR, Ser. Math.*, 46:1199–1273, 1983. 25

[26] S. W. Margolis and J. C. Meakin. Free inverse monoids and graph immersions. *Internat. J. Algebra Comput.*, 3(1):79–99, 1993. 4

[27] S. W. Margolis and J. C. Meakin. Inverse monoids, trees and context-free languages. *Trans. Amer. Math. Soc.*, 335(1):259–276, 1993. 25

[28] S. W. Margolis, M. V. Sapir, and P. Weil. Closed subgroups in pro-V topologies and the extension problem for inverse automata. *Internat. J. Algebra Comput.*, 11(4):405–446, 2001. 16

[29] L. Markus-Epstein. Stallings foldings and subgroups of amalgams of finite groups. *Internat. J. Algebra Comput.*, 17(8):1493–1535, 2007. 24

[30] A. Martino and E. Ventura. Fixed subgroups are compressed in free groups. *Commun. Algebra*, 32(10):3921–3935, 2004. 17

[31] O. S. Maslakova. The fixed point group of a free group automorphism. *Algebra and Logic*, 42:237–265, 2003. Translated from *Algebra i Logika*, 42:422–472, 2003. 17

[32] A. Miasnikov, E. Ventura, and P. Weil. Algebraic extensions in free groups. In *Geometric group theory*, Trends Math., pages 225–253. Birkhäuser, 2007. 12, 13

[33] D. E. Muller and P. E. Schupp. Groups, the theory of ends, and context-free languages. *J. Comput. System Sci.*, 26(3):295–310, 1983. 23

[34] H. Neumann. On the intersection of finitely generated free groups. addendum. *Publ. Math. (Debrecen)*, 5:128, 1957. 12

[35] W. Plandowski. Satisfiability of word equations with constants is in PSPACE. In *Proc. 40th Ann. Symp. Found. Comput. Sci.*, pages 495–500. IEEE Press, 1999. 25

[36] K. Reidemeister. Fundamentalgruppe und Überlagerungsräume. *J. Nachrichten Göttingen*, pages 69–76, 1928. 8

[37] J. Rhodes and B. Steinberg. *The q-theory of finite semigroups*. Springer-Verlag, 2009. 15

[38] L. Ribes and P. A. Zalesskii. On the profinite topology on a free group. *Bull. Lond. Math. Soc.*, 25:37–43, 1993. 15

[39] L. Ribes and P. A. Zalesskii. The pro-p topology of a free group and algorithmic problems in semigroups. *Internat. J. Algebra Comput.*, 4(3):359–374, 1994. 16

[40] A. Roig, E. Ventura, and P. Weil. On the complexity of the Whitehead minimization problem. *Internat. J. Algebra Comput.*, 17(8):1611–1634, 2007. 12, 16

[41] V. Roman'kov. On the occurrence problem for rational subsets of a group. In V. Roman'kov, editor, *International Conference on Combinatorial and Computational Methods in Mathematics*, pages 76–81, 1999. 24

[42] J. Sakarovitch. *Syntaxe des langages de Chomsky, essai sur le déterminisme*. PhD thesis, Université Paris VII, 1979. 22

[43] J. Sakarovitch. *Eléments de théorie des automates*. Vuibert, 2003. 17

[44] G. Sénizergues. Some decision problems about controlled rewriting systems. *Theoret. Comput. Sci.*, 71(3):281–346, 1990. 21

[45] G. Sénizergues. On the rational subsets of the free group. *Acta Informatica*, 33(3):281–296, 1996. 21, 22

[46] J.-P. Serre. *Arbres, amalgames,* SL_2. Société Mathématique de France, 1977. Avec un sommaire anglais; rédigé avec la collaboration de Hyman Bass; *Astérisque* 46. 4

[47] P. V. Silva. Recognizable subsets of a group: finite extensions and the abelian case. *Bull. European Assoc. Theor. Comput. Sci.*, 77:195–215, 2002. 23

[48] P. V. Silva. Free group languages: rational versus recognizable. *RAIRO Inform. Théor. App.*, 38(1):49–67, 2004. 21, 22

[49] P. V. Silva. Fixed points of endomorphisms over special confluent rewriting systems. *Monatsh. Math.*, 161(4):417–447, 2010. 17

[50] P. V. Silva and P. Weil. On an algorithm to decide whether a free group is a free factor of another. *RAIRO Inform. Théor. App.*, 42:395–414, 2008. 12

[51] P. V. Silva and P. Weil. Automorphic orbits in free groups: words versus subgroups. *Internat. J. Algebra Comput.*, 20(4):561–590, 2010. 25

[52] P. V. Silva and P. Weil. On finite-index extensions of subgroups of free groups. *J. Group Theory*, 13(3):365–381, 2010. 13

[53] C. C. Sims. *Computation with finitely presented groups*. Cambridge University Press, 1994. 2, 4

[54] J. R. Stallings. Topology of finite graphs. *Inventiones Math.*, 71(3):551–565, 1983. 4, 14

[55] M. Takahasi. Note on chain conditions in free groups. *Osaka J. Math.*, 3(2):221–225, 1951. 12

[56] W. M. Touikan. A fast algorithm for Stallings' folding process. *Internat. J. Algebra Comput.*, 16(6):1031–1046, 2006. 7

[57] E. Ventura. On fixed subgroups of maximal rank. *Commun. Algebra*, 25(10):3361–3375, 1997. 17

[58] E. Ventura. Fixed subgroups of free groups: a survey. *Contemporary Math.*, 296:231–255, 2002. 17

[59] J. H. C. Whitehead. On equivalent sets of elements in a free group. *Ann. of Math. (2)*, 37(4):782–800, 1936. 16

Abstract. This chapter is devoted to the study of rational subsets of groups, with particular emphasis on the automata-theoretic approach to finitely generated subgroups of free groups. Indeed, Stallings' construction, associating a finite inverse automaton with every such subgroup, inaugurated a complete rewriting of free group algorithmics, with connections to other fields such as topology or dynamics.

Another important vector in the chapter is the fundamental Benois' Theorem, characterizing rational subsets of free groups. The theorem and its consequences really explain why language theory can be successfully applied to the study of free groups. Rational subsets of (free) groups can play a major role in proving statements (*a priori* unrelated to the notion of rationality) by induction. The chapter also includes related results for more general classes of groups, such as virtually free groups or graph groups.

Index