On finite-index extensions of subgroups of free groups*

Pedro Silva, pvsilva@fc.up.pt
Centro de Matemática, Universidade do Porto †

Pascal Weil, pascal.weil@labri.fr
LaBRI, Université de Bordeaux, CNRS and CSE Department, IIT Delhi‡

Abstract

We study the lattice of finite-index extensions of a given finitely generated subgroup H of a free group F. This lattice is finite and we give a combinatorial characterization of its greatest element, which is the commensurator of H. This characterization leads to a fast algorithm to compute the commensurator, which is based on a standard algorithm from automata theory. We also give a sub-exponential and super-polynomial upper bound for the number of finite-index extensions of H, and we give a language-theoretic characterization of the lattice of finite-index subgroups of H. Finally, we give a polynomial time algorithm to compute the malnormal closure of H.

Keywords: free groups, subgroups, finite-index extensions

MSC: 20E05

This paper is part of the study of the lattice of finitely generated subgroups of a free group of finite rank F. Like most of the recent work on this topic, our paper makes crucial use of the graphical representation of the subgroups of F introduced in the seminal papers of Serre (1977 [8]) and Stallings (1983 [11]). This representation not only makes it easier to form an intuition and to prove properties of subgroups of F, but it also provides a convenient framework to efficiently solve algorithmic problems and compute invariants concerning these subgroups.

*The first author acknowledges support from 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 European Community structural funds. Both authors acknowledge support from ESF project AutoMathA.

†Faculdade de Ciências, Universidade do Porto, R. Campo Alegre 687, 4169-007 Porto, Portugal

‡LaBRI, 351 cours de la Libération, 33400 Talence, France.

The particular object of study in this paper is the lattice of extensions of a given finitely generated subgroup H of F, and more specifically of the sublattice of finite-index extensions of H.

It is elementary to verify that H has only finitely many finite-index extensions, and it is known that if K and L are finite-index extensions of H, then the subgroup they generate, namely $\langle K, L\rangle$, has finite index over H as well (Greenberg's theorem, see [11]). Therefore H has a maximum finite-index extension H_{fi}, which is effectively constructible, and the finite-index extensions of H form a full convex sublattice of the lattice of subgroups of F.

This paper contains a detailed discussion of the lattice of finite-index extensions of H. Our main contributions are the following.

We show that the maximum finite-index extension H_{fi} of H is the commensurator of H, and we give a combinatorial (graph-theoretic) characterization of H_{fi}. This characterization leads to efficient algorithms to compute all finite-index extensions of H, and to compute H_{fi} – the latter in time $\mathcal{O}(n \log n)$. We also give a rather tight upper bound on the number of finite-index extensions of H: there are at most $\mathcal{O}(\sqrt{n}\ n^{\frac{1}{2}\log_2 n})$ such extensions, where n is the number of vertices in the graphical representation of H. Note that this upper bound is sub-exponential but super-polynomial.

The consideration of the subgroups of the form H_{fi}, which have no proper finite-index extensions, leads us to the dual study of the lattice of finite-index subgroups of a given subgroup, and we give a combinatorial (language-theoretic) characterization of each such lattice.

Finally, we use our better understanding of the lattice of extensions of a subgroup of F, to give a polynomial time algorithm to compute the malnormal closure of a given subgroup.

As we already indicated, we use in a fundamental way the graphical representation of finitely generated subgroups of F, including a detailed study of the different steps of the computation of this representation (given a set of generators for the subgroup H), whose study was at the heart of an earlier paper by the authors [10]. It is particularly interesting to see that language-theoretic results and arguments play an important role in this paper: that is, we sometimes consider the graphical representation of a subgroup not just as an edge-labeled graph, but as a finite state automaton. Such considerations are present in almost all the results of this paper, but they become crucial at a rather unexpected juncture: the design of an efficient algorithm to compute the maximal finite-index extension H_{fi} of H. Indeed, the very low complexity we achieve is due to the possibility of using a standard automata-theoretic algorithm, namely the computation of the minimal automaton of a regular language.

Section 1 summarizes a number of well-known facts about free groups and the representation of their finitely generated subgroups, which will be used freely in the sequel (see [11, 12, 5, 7, 10] for more details). Section 2 is the heart of the paper: it starts with a technical study of the different steps of the algorithm to compute the graphical representation of a given subgroup, and a description of those steps which preserve finite-index (Section 2.2). These technical results are then used to characterize the maximal finite-index extension H_{fi} (Section 2.3),

to relate the computation of H_{fi} and the minimization of certain finite-state automata (Section 2.4), to evaluate the maximal number of finite-index extensions of a given subgroup (Section 2.5), and to describe an invariant of the lattice of finite-index subgroups of a given subgroup (Section 2.6).

Finally, we apply the same ideas in Section 3, to study the malnormal closure of a subgroup, and to show that it can be computed in polynomial time.

1 Subgroups of free groups and Stallings graphs

Let F be a finitely generated free group and let $A = \{a_1, \dots, a_r\}$ be a fixed basis of F. Let $\bar{A} = \{\bar{a}_1, \dots, \bar{a}_r\}$ be a disjoint copy of A and let $\tilde{A} = A \cup \bar{A}$: as usual, we extend the map $a \mapsto \bar{a}$ from the set A to all words by letting $\bar{\bar{a}} = a$ if $a \in A$ and $\overline{ua} = \bar{a}\bar{u}$ if $a \in \tilde{A}$ and $u \in \tilde{A}^*$. As usual again, the elements of F are identified with the *reduced words* over alphabet $\tilde{A}$, that is, the words that do not contain a sequence of the form $a\bar{a}$ ($a \in \tilde{A}$). If $u \in \tilde{A}^*$ is an arbitrary word, we denote by $\mathsf{red}(u)$ the corresponding reduced word, that is, the word obtained from u by repeatedly deleting all sequences of the form $a\bar{a}$ ($a \in \tilde{A}$).

A reduced word $u \in F$ is *cyclically reduced* if u cannot be written as $u = av\bar{a}$ with $a \in \tilde{A}$ and $v \in F$. Every reduced word u can be factored in a unique way in the form $u = xy\bar{x}$, with y cyclically reduced.

If H is a subgroup of F, an *extension* of H is any subgroup G containing H and we write $H \leq G$. If H is finitely generated, we also write $H \leq_{\mathsf{fg}} G$. If H has finite index in G, we say that G is a *finite-index extension* of H and we write $H \leq_{\mathsf{fi}} G$. Finally, we write $H \leq_{\mathsf{ff}} G$ if H is a free factor of G.

It is well known (since Serre's and Stalling's fundamental work [8, 11]) that every finitely generated subgroup $H \leq_{\mathsf{fg}} F$ admits a unique graphical representation of the form $\mathcal{A}(H) = (\Gamma(H), 1)$, where $\Gamma(H)$ is a finite directed graph with A-labeled edges and 1 is a designated vertex of $\Gamma(H)$, subject to the combinatorial conditions below. Here, a graph is a pair (V, E) where V is the set of *vertices* and $E \subseteq V \times A \times V$ is the set of *edges*; the *in-degree* (resp. *out-degree*) of a vertex $v \in V$ is the number of edges in E of the form (v', a, v) (resp. (v, a, v')); and the *degree* of v is the sum of its in- and out-degree. Every pair $\mathcal{A}(H)$ satisfies the following:

- the (underlying undirected) graph is connected;
- for each $a \in A$, every vertex is the source (resp. the target) of at most one a-labeled edge;
- and every vertex, except possibly 1, has degree at least 2.

Moreover, every pair $(\Gamma, 1)$ with these properties is said to be *admissible*, and it is the representation of a finitely generated subgroup of F. In addition, given a finite set of generators of H, the representation of H is effectively computable. We refer the reader to [11, 5, 7, 12, 10] for some of the literature on this construction and its many applications, and to Section 2.1 below on the construction of $\mathcal{A}(H)$.

We sometimes like to view the A-labeled graph $\Gamma(H)$ as a transition system over alphabet $\tilde{A}$: if p, q are vertices of $\Gamma(H)$, $a \in A$ and (p, a, q) is an edge of

$\Gamma(H)$, we say that a labels a path from p to q and $\bar{a}$ labels a path from q to p, written $p \xrightarrow{a} q$ and $q \xrightarrow{\bar{a}} p$. If $u \in \tilde{A}^*$ is a word (reduced or not) and $u = va$ ($a \in \tilde{A}$), we say that $p \xrightarrow{u} q$ (*u labels a path from p to q*) if $p \xrightarrow{v} p' \xrightarrow{a} q$ for some vertex p'. In particular, a reduced word is in H if and only if it labels a loop at vertex 1. Moreover, if we have a path $p \xrightarrow{u} q$, then we also have a path $p \xrightarrow{\mathsf{red}(u)} q$.

If $H \leq_{\mathsf{fg}} G \leq_{\mathsf{fg}} F$, then there is a *homomorphism* from $\mathcal{A}(H)$ into $\mathcal{A}(G)$, that is, a map φ from the vertex set of $\Gamma(H)$ to the vertex set of $\Gamma(G)$ such that

- $\varphi(1) = 1$ and
- if $p \xrightarrow{a} q$ in $\Gamma(H)$ (p, q vertices, $a \in \tilde{A}$), then $\varphi(p) \xrightarrow{a} \varphi(q)$ in $\Gamma(G)$.

It is not difficult to verify that this morphism, if it exists, is unique, and we denote it by φ_H^G. It is well known (see [8, 11, 5, 7]) that if φ_H^G is one-to-one, then $H \leq_{\mathsf{ff}} G$, and that $H \leq_{\mathsf{fi}} G$ if and only if the homomorphism φ_H^G is a *cover*, that is, it satisfies

- $p \xrightarrow{a} q$ in $\Gamma(H)$ (p, q vertices, $a \in \tilde{A}$) if and only if $\varphi_H^G(p) \xrightarrow{a} \varphi_H^G(q)$ in $\Gamma(G)$.

In that case, all sets of the form $(\varphi_H^G)^{-1}(q)$ (q a vertex of $\Gamma(G)$) have the same cardinality, which is equal to the index of H in G.

2 Finite-index extensions

It follows from the characterization of finite-index extensions by covers, that if $H \leq_{\mathsf{fi}} G$, then φ_H^G is onto[1]. Therefore H has only a finite number of finite-index extensions, and that number can be bounded above by the number of binary relations on $\Gamma(H)$: if that graph has n vertices, then H has at most 2^{n^2} finite-index extensions. We give a better upper bound in Section 2.5.

Moreover, the collection of finite-index extensions of H is effectively computable. In addition, it is elementary to use these graphical representations to show that the join of two finite-index extensions of H is again a finite-index extension (see Stalling's proof of Greenberg's theorem [11]). It follows that if $H \leq_{\mathsf{fg}} F$, then H admits an effectively computable maximum finite-index extension H_{fi}. The results of Section 2.4 below yield an efficient algorithm to compute the set of finite-index extensions of a given subgroup H, and its maximum finite-index extension H_{fi}.

Remark 2.1 We just observed that every finitely generated subgroup of F has a finite number of finite-index extensions. However, it usually has infinitely many finite-index subgroups. More precisely, every non-trivial subgroup $H \leq_{\mathsf{fg}} F$ admits a finite-index subgroup of index r for each $r \geq 1$. Indeed, let $\mathcal{A}(H) = (\Gamma(H), 1)$ and let Q be the vertex set of $\Gamma(H)$. Define Γ_r be the A-labeled graph with vertex set $Q \times \{1, \ldots, r\}$ and with the following edge set: for each edge (p, a, q) of $\Gamma(H)$, there is an edge $((p, i), a, (q, i+1))$ for each $1 \leq i < r$ and

[1] The converse is not true, see [7] for a detailed study of the extensions $H \leq_{\mathsf{fg}} G$ such that φ_H^G is onto.

an edge $((p,r),a,(q,1))$. Then Γ_r is an admissible graph, the map $\pi\colon (p,i)\mapsto p$ defines a cover from Γ_r to $\Gamma(H)$, and if H_r is the subgroup represented by $(\Gamma_r,(1,1))$, then H_r has index r in H. □

Finally, we note the following useful fact, which we will use freely in the sequel.

Lemma 2.2 *If $H \leq_{\mathsf{fi}} G$, $u \in F$, $p \xrightarrow{u} q$ in $\Gamma(H)$ and $\varphi_H^G(p) = \varphi_H^G(q)$, then $\mathsf{red}(u^m)$ labels a loop at p for some integer $m > 1$.*

Proof. Let x, y be such that $u = xy\bar{x}$ and y is cyclically reduced, and let p', q' be the vertices such that $p \xrightarrow{x} p' \xrightarrow{y} q' \xrightarrow{\bar{x}} q$. From $\varphi_H^G(p) = \varphi_H^G(q)$, we deduce that $\varphi_H^G(p') = \varphi_H^G(q')$ and that y labels a loop at $\varphi_H^G(p')$ in $\Gamma(G)$. Since φ_H^G is a cover, there exists a sequence of vertices $(p'_n)_{n\geq 0}$ of $\Gamma(H)$ such that $p'_0 = p'$ and $p'_n \xrightarrow{y} p'_{n+1}$. But $\Gamma(H)$ is finite, so there exists $m > 1$ such that $p'_m = p'$. Therefore y^m labels a loop at p' and $xy^m\bar{x} = \mathsf{red}(u^m)$ labels a loop at p. □

2.1 i-steps and finite-index extensions

Let $H \leq G$ be finitely generated subgroups of F and let $g_1, \ldots, g_n$ be reduced words such that $G = \langle H, g_1, \ldots, g_n\rangle$. Let $G_0 = H$ and let $G_i = \langle G_{i-1}, g_i\rangle$ $(1 \leq i \leq n)$. We may of course assume that $g_i \notin G_{i-1}$, so $G_{i-1} \neq G_i$.

Then $\mathcal{A}(G_i)$ is obtained from $\mathcal{A}(G_{i-1})$ by, first, adding sufficiently many new vertices and edges to create a new path from vertex 1 to itself, labeled by g_i; and second, by *reducing* the resulting graph, that is, repeatedly identifying vertices p and p' such that $q \xrightarrow{a} p$ and $q \xrightarrow{a} p'$ for some vertex q and some letter $a \in \tilde{A}$, see for instance [11, 12, 5, 10]. Depending on the length of prefixes of g_i and $\bar{g}_i$ that can be read from vertex 1 in $\Gamma(G_{i-1})$, this procedure amounts to one of the two following moves:

- a *reduced expansion*, or *re-step* (we write $\mathcal{A}(G_{i-1}) \longrightarrow_{\mathsf{re}}^{(p,w,q)} \mathcal{A}(G_i)$), that is, we add a new path labeled by a factor w of g_i, from some vertex p to some vertex q of $\Gamma(G_{i-1})$ in such a way that the resulting graph is admissible (needs no reduction);

- or an *i-step* (we write $\mathcal{A}(G_{i-1}) \longrightarrow_{\mathsf{i}}^{p=q} \mathcal{A}(G_i)$), that is, we identify a pair of vertices (p,q) of $\Gamma(G_{i-1})$, and we reduce the resulting graph.

We refer the readers to [10, Section 2] for a detailed analysis of these moves and we record the following observation

Lemma 2.3 *Let $G, H \leq_{\mathsf{fg}} F$. If $H \leq_{\mathsf{fi}} G$, then only i-steps are involved in the transformation from $\Gamma(H)$ to $\Gamma(G)$.*

Proof. Let $H = G_0 \leq G_1 \leq \ldots \leq G_n = G$ be as in the above discussion. Note that $H \leq_{\mathsf{fi}} G$ if and only if $G_{i-1} \leq_{\mathsf{fi}} G_i$ for each $1 \leq i \leq n$. If $\mathcal{A}(G_{i-1}) \longrightarrow_{\mathsf{re}} \mathcal{A}(G_i)$, then the homomorphism $\varphi_{G_{i-1}}^{G_i}$ is one-to-one, so $G_{i-1} \leq_{\mathsf{ff}} G_i$ and in particular, G_i is not a finite-index extension of G_{i-1}. □

2.2 Which i-steps yield finite-index extensions?

Let us say that H is *cyclically reduced* (with respect to the basis A) if every vertex of $\Gamma(H)$ has degree at least equal to 2. If H is not cyclically reduced, then the designated vertex 1 of $\Gamma(H)$ has degree 1 and $\Gamma(H)$ consists of two parts: $\mathsf{tail}(\Gamma(H))$, which contains the designated vertex 1 and all degree 2 vertices that can be connected to vertex 1 through other degree 2 vertices; and the rest of $\Gamma(H)$, which is called the *cyclic core* of $\Gamma(H)$, written $\mathsf{cc}(\Gamma(H))$. We let $\mathsf{t}(1)$ be the shortest word which labels a path from 1 to a vertex in $\mathsf{cc}(\Gamma(H))$ and let $\tau(1)$ be the vertex of $\mathsf{cc}(\Gamma(H))$ thus reached (if H is cyclically reduced, then $\mathsf{t}(1)$ is the empty word and $\tau(1) = 1$).

Lemma 2.4 *Let $H \leq_{\mathsf{fg}} F$ and let p be a vertex of $\Gamma(H)$. Then p is a vertex of $\mathsf{cc}(\Gamma(H))$ if and only if some cyclically reduced word u labels a path from p to p.*

Proof. By definition, if H is not cyclically reduced, then $\mathsf{tail}(\Gamma(H))$ consists of a single path from vertex 1 to vertex $\mathsf{t}(1)$ (excluding the latter vertex): it is therefore elementary to verify that no non-empty cyclically reduced word labels a loop at a vertex in $\mathsf{tail}(\Gamma(H))$.

Let now p be a vertex in $\mathsf{cc}(\Gamma(H))$: then p has degree at least 2, and if it has degree exactly 2, then neither of the two edges adjacent to it leads to a vertex in $\mathsf{tail}(\Gamma(H))$. Therefore, one can find distinct letters $a, a' \in \tilde{A}$ such that $p \xrightarrow{a} q$ and $p \xrightarrow{a'} q'$, with q and q' in $\mathsf{cc}(\Gamma(H))$ as well. Iterating this reasoning, one can show that there exist arbitrarily long paths within $\mathsf{cc}(\Gamma(H))$, starting from p and labeled by reduced words of the form au and $a'u'$. Since $\Gamma(H)$ is finite, vertices are repeated along these paths, and we consider the earliest such repetition after the initial p. If p itself is the first repeated vertex along the path labeled au, we have a loop $p \xrightarrow{au} p$ such that au is cyclically reduced, and we are done. The situation is similar if p is the first repeated vertex along the path labeled $a'u'$. Otherwise, let r and r' be the first repeated vertices along the two paths. Then $r \neq p$, $r' \neq p$, and $\mathsf{cc}(\Gamma(H))$ has paths of the form $p \xrightarrow{au} r$, $p \xrightarrow{a'u'} r'$, $r \xrightarrow{v} r$ and $r' \xrightarrow{v'} r'$ such that $auv\bar{u}\bar{a}, a'u'v'\bar{u}'\bar{a}'$ are reduced. Then the word $auv\bar{u}\bar{a}a'u'v'\bar{u}'\bar{a}'$ is cyclically reduced, and it labels a loop at p in $\Gamma(H)$. This concludes the proof. □

If p is a vertex of $\mathsf{cc}(\Gamma(H))$, we let $\widetilde{L}_p(H)$ be the language accepted by $\mathsf{cc}(\Gamma(H))$, seen as a finite state automaton with initial state p and all states final: that is, the set of (possibly non-reduced) words in $\tilde{A}^*$ that label a path in $\mathsf{cc}(\Gamma(H))$ starting at vertex p. Let then $L_p(H)$ be the set of reduced words in $\widetilde{L}_p(H)$ — which is also the set of all $\mathsf{red}(u)$ ($u \in \widetilde{L}_p(H)$), and also the set of prefixes of words in the subgroup represented by the pair $(\mathsf{cc}(\Gamma(H)), p)$. Let us first record the following elementary fact.

Fact 2.5 If $p \xrightarrow{u} q$ is a path in $\mathsf{cc}(\Gamma(H))$, then $L_q(H) = \{\mathsf{red}(\bar{u}x) \mid x \in L_p(H)\}$. □

We now refine the result of Lemma 2.3.

Proposition 2.6 *Let $H \leq_{\mathsf{fg}} F$ and let p, q be distinct vertices in $\Gamma(H)$. Let G be the subgroup of F such that $\mathcal{A}(H) \longrightarrow_{\mathsf{i}}^{p=q} \mathcal{A}(G)$. Then $H \leq_{\mathsf{fi}} G$ if and only if p, q are in $\mathsf{cc}(\Gamma(H))$ and $L_p(H) = L_q(H)$, if and only if p, q are in $\mathsf{cc}(\Gamma(H))$ and $\widetilde{L}_p(H) = \widetilde{L}_q(H)$.*

Proof. Let us first assume that $H \leq_{\mathsf{fi}} G$.

If $p, q \in \mathsf{tail}(\Gamma(H))$, then without loss of generality, there exists a reduced word labeling a path $1 \xrightarrow{u} p \xrightarrow{v} q$ entirely within $\mathsf{tail}(\Gamma(H))$. Then $\mathsf{red}(uv\bar{u})$ is an element of G and not of H. Let w be the longest common suffix of u and v, with $u = u'w$ and $v = v'w$, say $1 \xrightarrow{u'} p' \xrightarrow{w} p \xrightarrow{v'} q' \xrightarrow{w} q$. Then $\mathsf{red}(uv\bar{u}) = u'(wv')\bar{u}'$ and $\mathcal{A}(H) \longrightarrow_{\mathsf{i}}^{p'=q'} \mathcal{A}(G)$, so we may assume that $uv\bar{u}$ is reduced.

Let x, y be such that $v = xy\bar{x}$ and y is cyclically reduced, say, $p \xrightarrow{x} r \xrightarrow{y} s \xrightarrow{\bar{x}} q$. In particular, $r \in \mathsf{tail}(\Gamma(H))$. Since $H \leq_{\mathsf{fi}} G$, there exists $m \geq 2$ such that $uv^m\bar{u} = uxy^m\bar{x}\bar{u}$ (a reduced word) lies in H. It follows that the cyclically reduced word y^m labels a loop at r in $\Gamma(H)$, a contradiction in view of Lemma 2.4.

If $p \in \mathsf{tail}(\Gamma(H))$ and $q \in \mathsf{cc}(\Gamma(H))$, let $u \in F$ be such that $1 \xrightarrow{u} p$ is a path entirely within $\mathsf{tail}(\Gamma(H))$ and let v be a cyclically reduced word labeling a loop at q. Then $\mathsf{red}(uv\bar{u}) \in G \setminus H$ and $\mathsf{red}(uv^m\bar{u}) \in H$ for some $m > 1$.

If $uv^m\bar{u}$ is reduced, then the cyclically reduced word v^m labels a loop at vertex p, contradicting the fact that $p \in \mathsf{tail}(\Gamma(H))$. So we may assume that $uv^m\bar{u}$ is not reduced. However, since v is cyclically reduced, one of uv^m and $v^m\bar{u}$ is reduced. Let us assume that uv^m is not reduced (the other case is dual) and let $v^k v'$ be the longest common prefix of v^m and $\bar{u}$, so $v = v'v''$ and $u = u'\bar{v}'\bar{v}^k$ with $u', v' \in F$. Then $1 \xrightarrow{u'} r \xrightarrow{\bar{v}'\bar{v}^k} p$ for some vertex $r \in \mathsf{tail}(\Gamma(H))$ and $\mathsf{red}(uv^m\bar{u}) = u'(v''v')^m\bar{u}'$. It follows that the word $v''v'$, which is cyclically reduced, labels a loop at vertex r, a contradiction.

If $p, q \in \mathsf{cc}(\Gamma(H))$ and $\widetilde{L}_p(H) \neq \widetilde{L}_q(H)$, we consider (without loss of generality) a word $u \in \widetilde{L}_p(H) \setminus \widetilde{L}_q(H)$, with minimum length, say $u = va$ with $a \in \tilde{A}$. By definition, there exist paths $p \xrightarrow{v} p' \xrightarrow{a} p''$ and $q \xrightarrow{v} q'$, but no path $q' \xrightarrow{a} q''$ in $\mathsf{cc}(\Gamma(H))$. Observe that $\mathcal{A}(H) \longrightarrow_{\mathsf{i}}^{p'=q'} \mathcal{A}(G)$. If there is a path $q' \xrightarrow{a} q''$ in $\Gamma(H)$, then we also have $\mathcal{A}(H) \longrightarrow_{\mathsf{i}}^{p''=q''} \mathcal{A}(G)$ and since $p'' \in \mathsf{cc}(\Gamma(H))$ and $q'' \in \mathsf{tail}(\Gamma(H))$, we conclude to a contradiction as in the previous case. We now assume that there is no path $q' \xrightarrow{a} q''$ in $\Gamma(H)$.

We claim that there exists $w \in F$ such that aw is cyclically reduced and $p' \xrightarrow{aw} q'$ in $\mathsf{cc}(\Gamma(H))$. Let indeed $p'' \xrightarrow{z} q'$ be a path of minimal length in $\mathsf{cc}(\Gamma(H))$ (there exists one by connectedness). Since $p'' \in \mathsf{cc}(\Gamma(H))$, there exists a path $p'' \xrightarrow{b} r$ for some $b \in \tilde{A}$, $b \neq \bar{a}$, and as in the proof of Lemma 2.4, there exists a reduced word of the form bt labeling a loop at p''. Let $w = \mathsf{red}(btz)$: then we have a path $p'' \xrightarrow{w} q'$. By minimality of the length of z, $\bar{t}\,\bar{b}$ is not a prefix of z, so w starts with letter b, and hence aw is reduced. In fact, aw is cyclically reduced since there is no path $q' \xrightarrow{a} q''$.

Let $1 \xrightarrow{t} q'$ be a path in $\Gamma(H)$. Then $\mathsf{red}(taw\bar{t}) \in G$, and hence there exists $m > 1$ such that $\mathsf{red}(t(aw)^m\bar{t}) \in H$. Again, since there is no path $q' \xrightarrow{a} q''$, the word ta is reduced. By replacing m by a sufficiently large multiple, we find that ta is a prefix of $\mathsf{red}(t(aw)^m\bar{t})$, and hence that ta labels a path from 1 in $\Gamma(H)$: this contradicts the absence of a path $q' \xrightarrow{a} q''$.

Thus we have proved that, if $H \leq_{\mathsf{fi}} G$, then $p, q \in \mathsf{cc}(\Gamma(H))$ and $\widetilde{L}_p(H) = \widetilde{L}_q(H)$. The latter condition immediately implies that $L_p(H) = L_q(H)$. **We now assume that $p, q \in \mathbf{cc}(\Gamma(H))$ and $L_p(H) = L_q(H)$, and we show that $H \leq_{\mathsf{fi}} G$.** We first establish a technical fact.

Fact 2.7 Let $r_i \xrightarrow{z_i} s_{i+1}$ ($z_i \in F$, $0 \leq i \leq k$) be paths in $\mathsf{cc}(\Gamma(H))$, such that $r_i, s_i \in \{p, q\}$ for each $1 \leq i \leq k$. Then there exists a path $r_0 \xrightarrow{\mathsf{red}(z_0 \cdots z_k)} t$ in $\mathsf{cc}(\Gamma(H))$. □

Proof. The proof is by induction on k, and is trivial for $k = 0$. If $k > 0$, then there is a path $r_1 \xrightarrow{\mathsf{red}(z_1 \cdots z_k)} t$ in $\mathsf{cc}(\Gamma(H))$. Since $L_p(H) = L_q(H)$, there is also a path $s_1 \xrightarrow{\mathsf{red}(z_1 \cdots z_k)} t'$ for some $t' \in \mathsf{cc}(\Gamma(H))$, and therefore a path $r_0 \xrightarrow{\mathsf{red}(z_0 \cdots z_k)} t'$ as required. □

We want to show that G has finitely many H-cosets. Let $u \in G$: then u labels a loop at 1 in $\Gamma(G)$. Let $\mathcal{B}$ be the automaton obtained from $\Gamma(H)$ by identifying vertices p and q, but without performing any reduction. Then $\Gamma(G)$ is the result of the reduction of $\mathcal{B}$. In particular (say, in view of [10, Fact 1.4]), $u = \mathsf{red}(v)$ for some word $v \in \tilde{A}^*$ labeling a loop at 1 in $\mathcal{B}$. By definition of $\mathcal{B}$, the word v factors as $v = v_0 \cdots v_k$, in such a way that $\Gamma(H)$ has paths of the form $1 \xrightarrow{v_0} s_1$, $r_i \xrightarrow{v_i} s_{i+1}$ ($1 \leq i < k$) and $r_k \xrightarrow{v_k} 1$, and the vertices $r_1, s_1, \cdots, r_k, s_k$ are all equal to p or q. As observed in Section 1, $\Gamma(H)$ also has paths

$$1 \xrightarrow{\mathsf{red}(v_0)} s_1,\ r_i \xrightarrow{\mathsf{red}(v_i)} s_{i+1}\ (1 \leq i < k) \text{ and } r_k \xrightarrow{\mathsf{red}(v_k)} 1.$$

In particular, we have $\mathsf{red}(v_0) = \mathsf{t}(1)w_0$ and $\mathsf{red}(v_k) = w_k\bar{\mathsf{t}}(1)$ for some $w_0, w_k \in F$, and there are paths $1 \xrightarrow{\mathsf{t}(1)} \tau(1) \xrightarrow{w_0} s_1$ and $r_k \xrightarrow{w_k} \tau(1) \xrightarrow{\bar{\mathsf{t}}(1)} 1$. Note that the paths $\tau(1) \xrightarrow{w_0} s_1$, $r_i \xrightarrow{\mathsf{red}(v_i)} s_{i+1}$ ($1 \leq i < k$) and $r_k \xrightarrow{w_k} \tau(1)$ are set entirely within $\mathsf{cc}(\Gamma(H))$, since no reduced word-labeled path between vertices in $\mathsf{cc}(\Gamma(H))$ can visit a vertex in $\mathsf{tail}(\Gamma(H))$.

By Fact 2.7, there exists a path $\tau(1) \xrightarrow{\mathsf{red}(w_0 v_1 \cdots v_{k-1} w_k)} t$ for some vertex t in $\mathsf{cc}(\Gamma(H))$. Let h be a shortest-length word such that $t \xrightarrow{h} \tau(1)$ in $\mathsf{cc}(\Gamma(H))$. Then $z = \mathsf{t}(1)\,\mathsf{red}(w_0 v_1 \cdots v_{k-1} w_k)\ h\ \bar{\mathsf{t}}(1)$ labels a loop at vertex 1 in $\Gamma(H)$, so $\mathsf{red}(z) \in H$. By construction, we have $u = \mathsf{red}(\mathsf{t}(1)w_0 v_1 \cdots v_{k-1} w_k \bar{\mathsf{t}}(1))$, so $u \in H\,\mathsf{red}(\mathsf{t}(1)\ \bar{h}\ \bar{\mathsf{t}}(1))$. Since h was chosen to be a geodesic in $\mathsf{cc}(\Gamma(H))$, it can take only finitely many values, and this completes the proof that G has finitely many H-cosets. □

We note the following consequence of this proof.

Fact 2.8 Let $H \leq_{\mathsf{fg}} F$ and let $p, q \in \mathsf{cc}(\Gamma(H))$. Then $L_p(H) = L_q(H)$ if and only if $\widetilde{L}_p(H) = \widetilde{L}_q(H)$. □

2.3 The lattice of finite-index extensions of H

We further refine Proposition 2.6 as follows: we consider an extension $H \leq_{\mathsf{fi}} G$ and a pair (r, s) of vertices of $\Gamma(H)$, whose identification yields a finite-index extension of H. Then we show that identifying the vertices of $\Gamma(G)$ corresponding to r and s, also yields a finite-index extension of G.

Lemma 2.9 *Let $H \leq_{\mathsf{fg}} F$, let $H \leq_{\mathsf{fi}} G$ and let $\varphi\colon \Gamma(H) \to \Gamma(G)$ be the corresponding cover. Let p be a vertex of $\mathsf{cc}(\Gamma(H))$. Then $p \in \mathsf{cc}(\Gamma(H))$ if and only if $\varphi(p) \in \mathsf{cc}(\Gamma(G))$, and in that case, $\widetilde{L}_p(H) = \widetilde{L}_{\varphi(p)}(G)$ and $L_p(H) = L_{\varphi(p)}(G)$.*

Proof. Let $p \in \Gamma(H)$. By definition of a cover between finite graphs, if a word u labels a loop at p in $\Gamma(H)$, then u labels a loop at $\varphi(p)$ in $\Gamma(G)$; and if u labels a loop at $\varphi(p)$ in $\Gamma(G)$, then u^m labels a loop at p in $\Gamma(H)$, for some $m \geq 1$. Since the vertices of $\mathsf{cc}(\Gamma(H))$ (resp. $\mathsf{cc}(\Gamma(G))$) are those at which there is a loop labeled by a cyclically reduced word, and since any power of such a word is cyclically reduced as well, we conclude that $p \in \mathsf{cc}(\Gamma(H))$ if and only if $\varphi(p) \in \mathsf{cc}(\Gamma(G))$.

Let us now assume that $p \in \mathsf{cc}(\Gamma(H))$. If $p \xrightarrow{u} r$ in $\mathsf{cc}(\Gamma(H))$, then the φ-image of this path is a path $\varphi(p) \xrightarrow{u} \varphi(r)$, which is entirely contained in $\mathsf{cc}(\Gamma(G))$ by the previous paragraph. In particular, $\widetilde{L}_p(H) \subseteq \widetilde{L}_{\varphi(p)}(G)$.

Conversely, suppose that $\varphi(p) \xrightarrow{u} r'$ is a path in $\mathsf{cc}(\Gamma(G))$. By definition of a cover, u labels a path $p \xrightarrow{u} r$ for some vertex r such that $\varphi(r) = r'$. By the first part of the proof again, this path in entirely contained in $\mathsf{cc}(\Gamma(H))$, and hence $u \in \widetilde{L}_p(H)$. Thus $\widetilde{L}_p(H) = \widetilde{L}_{\varphi(p)}(G)$. □

Together with Proposition 2.6, Lemma 2.9 immediately implies the following statements.

Corollary 2.10 *Let $H \leq_{\mathsf{fg}} F$.*

1. *Let $H \leq_{\mathsf{fi}} G$ and let $\varphi\colon \Gamma(H) \to \Gamma(G)$ be the corresponding cover. If p, q are vertices of $\Gamma(H)$, $\mathcal{A}(H) \longrightarrow_{\mathsf{i}}^{p=q} \mathcal{A}(K)$ and $\mathcal{A}(G) \longrightarrow_{\mathsf{i}}^{\varphi(p)=\varphi(q)} \mathcal{A}(K')$, then $H \leq_{\mathsf{fi}} K$ if and only if $G \leq_{\mathsf{fi}} K'$.*
2. *$H \leq_{\mathsf{fi}} G$ if and only if $\mathcal{A}(G)$ is obtained from $\mathcal{A}(H)$ by identifying some pairs of vertices (p, q) in $\mathsf{cc}(\Gamma(H))$ such that $L_p(H) = L_q(H)$, and then reducing the resulting graph.*

The identification of all pairs of vertices (p, q) such that $L_p(H) = L_q(H)$ yields the minimum quotient of $\mathcal{A}(H)$ and hence the maximum finite-index extension H_{fi} of H. In addition, we find that H_{fi} is exactly the *commensurator* of H (the set $\mathsf{Comm}_F(H)$ of all elements $g \in F$ such that $H \cap H^g$ has finite index in both H and H^g), a fact that can also be deduced from [5, Lemma 8.7].

Theorem 2.11 *Let $H \leq_{\mathsf{fg}} F$ and let H_{fi} be its maximum finite-index extension.*

1. *$\mathcal{A}(H_{\mathsf{fi}})$ is obtained from $\mathcal{A}(H)$ by identifying all pairs of vertices p, q of $\mathsf{cc}(\Gamma(H))$ such that $L_p(H) = L_q(H)$. No reduction is necessary.*
2. *$H_{\mathsf{fi}} = \mathsf{Comm}_F(H)$.*

Proof. In view of Corollary 2.10, $\mathcal{A}(H_{\mathsf{fi}})$ is obtained from $\mathcal{A}(H)$ by identifying all pairs of vertices p, q of $\mathsf{cc}(\Gamma(H))$ such that $L_p(H) = L_q(H)$, and then by reducing the resulting graph $\mathcal{B}$. If $p \xrightarrow{a} r$ and $q \xrightarrow{a} s$ $(a \in \tilde{A})$ are paths in $\mathsf{cc}(\Gamma(H))$ and if $L_p(H) = L_q(H)$, then $L_r(H) = L_s(H)$ by Fact 2.5. Thus $\mathcal{B}$ is already reduced, which concludes the proof of the first statement.

The fact that $\mathsf{Comm}(H)$ is a subgroup and a finite-index extension of H is proved, for instance, in [5, Prop. 8.9]. Conversely, suppose that $H \leq_{\mathsf{fi}} G$ and $g \in G$. Since conjugation by g is an automorphism of G, we have $H^g \leq_{\mathsf{fi}} G$. Now the intersection of finite-index subgroups, again has finite index, so $H \cap H^g \leq_{\mathsf{fi}} G$ and hence $H \cap H^g \leq_{\mathsf{fi}} H$ and $H \cap H^g \leq_{\mathsf{fi}} H^g$. Thus $g \in \mathsf{Comm}_F(H)$, which concludes the proof. □

2.4 Computing finite-index extensions

Recall the notion of minimization of a deterministic finite-state automaton (see [6] for instance). Let $\mathcal{B} = (Q, i, E, T)$ be such an automaton, over alphabet B, with Q the finite set of states, $i \in Q$ the initial state, $E \subseteq Q \times B \times Q$ the set of transitions and $T \subseteq Q$ the set of accepting states, and let L be the language accepted by $\mathcal{B}$, that is, the set of words in B^* that label a path from p to a state in T. Then the minimal automaton of L is obtained by identifying the pairs of states (p, q) such that the automata (Q, p, E, T) and (Q, q, E, T) accept the same language.

In our situation, the alphabet is $\tilde{A}$ and $\widetilde{L}_p(H)$ is the language accepted by the automaton $\mathcal{B}_p$, whose states and transitions are given by $\mathsf{cc}(\Gamma(H))$, with initial state p and all states final. Therefore Corollary 2.10 and Theorem 2.11 show that the identification of two vertices $p, q \in \mathsf{cc}(\Gamma(H))$ yields a finite-index extension if and only if p and q are identified when minimizing $\mathcal{B}_{\tau(1)}$. Moreover, $\mathsf{cc}(\Gamma(H_{\mathsf{fi}}))$ is given by the states and transitions of the minimal automaton of $\widetilde{L}_{\tau(1)}(H)$.

The classical Hopcroft algorithm (see [6]) minimizes an n-state automaton in time $\mathcal{O}(n \log n)$, so we have the following result.

Proposition 2.12 *Let $H \leq_{\mathsf{fg}} F$, and let n be the number of vertices of $\Gamma(H)$.*

- $\mathsf{cc}(\Gamma(H_{\mathsf{fi}}))$ *is obtained by minimizing the automaton given by the vertices and edges of* $\mathsf{cc}(\Gamma(H))$*, with all states final (the initial state does not matter in that situation).*
- *One can compute* $\Gamma(H_{\mathsf{fi}})$ *in time* $\mathcal{O}(n \log n)$*.*
- *One can decide in time* $\mathcal{O}(n \log n)$ *whether identifying a given set of pair of vertices of* $\Gamma(H)$ *will produce a finite-index extension of* H*.*

Remark 2.13 It may be that for the particular automata at hand (over a symmetrized alphabet, with all states final), the complexity of Hopcroft's algorithm might be better than $\mathcal{O}(n \log n)$, even linear. It has also been observed that in many instances, Myhill's automata minimization algorithm exhibits a better performance than Hopcroft's, in spite of a $\mathcal{O}(n^2)$ worst-case complexity. Brzozowski's algorithm [2] also performs remarkably well in practice [3]. □

2.5 Counting finite-index extensions

Recall that, if Γ is an A-labeled graph, the *product* $\Gamma \times_A \Gamma$ (also called the *fiber product*, or the *pull-back*, of two copies of Γ) is the A-labeled graph whose vertex set is the set of pairs (p, q) of vertices of Γ and whose edges are the triples $((p, q), a, (p', q'))$ such that (p, a, q) and (p', a, q') are edges of Γ. This graph is not connected in general (the vertices of the form (p, p) form a connected component that is isomorphic to Γ), nor reduced. More precisely, there is a u-labeled path in $\Gamma \times_A \Gamma$ from (p, q) to (p', q'), if and only if Γ has paths $p \xrightarrow{u} p'$ and $q \xrightarrow{u} q'$.

If p, q are vertices of $\mathsf{cc}(\Gamma(H))$, we let $p \sim q$ if and only if $L_p(H) = L_q(H)$.

Proposition 2.14 *Let* $H \leq_{\mathsf{fg}} F$*.*

- *The relation* $\sim$ *is a union of connected components of* $\mathsf{cc}(\Gamma(H)) \times_A \mathsf{cc}(\Gamma(H))$*.*
- *Let* p, q *be vertices of* $\mathsf{cc}(\Gamma(H))$*. Then* $p \sim q$ *if and only if the first and the second component projections, from the connected component of* (p, q) *in* $\mathsf{cc}(\Gamma(H)) \times_A \mathsf{cc}(\Gamma(H))$ *to* $\mathsf{cc}(\Gamma(H))$ *are both covers.*

Proof. The first statement follows directly from Fact 2.5, which shows that if $p \sim q$ and there is a path $(p, q) \xrightarrow{u} (p', q')$, then $p' \sim q'$.

Let us now assume that $p \sim q$ and let us show that the first component projection is a cover from the connected component of (p, q) onto $\mathsf{cc}(\Gamma(H))$. Let (r, s) be a vertex in that connected component: then there exists $u \in F$ such that $p \xrightarrow{u} r$ and $q \xrightarrow{u} s$. Let $r \xrightarrow{a} r'$ $(a \in \tilde{A})$ be an edge in $\Gamma(H)$. Then $ua \in L_p(H)$, so $ua \in L_q(H)$, and hence (since $\Gamma(H) \times_A \Gamma(H)$ is deterministic), there exists an a-labeled path $s \xrightarrow{a} s'$. Therefore there exists an a-labeled path $(r, s) \xrightarrow{a} (r', s')$. Thus the first component projection is a cover. The proof concerning the second component projection is identical.

Conversely, suppose that the first and the second component projections, from the connected component of (p, q) in $\mathsf{cc}(\Gamma(H)) \times_A \mathsf{cc}(\Gamma(H))$ to $\mathsf{cc}(\Gamma(H))$

are covers, and let $u \in L_p(H)$. Then $\mathsf{cc}(\Gamma(H))$ has a path $p \xrightarrow{u} r$. It is an elementary property of covers that this path can be lifted to a path in $\mathsf{cc}(\Gamma(H)) \times_A \mathsf{cc}(\Gamma(H))$, of the form $(p,q) \xrightarrow{u} (r,s)$. The second component projection of that path yields a path $q \xrightarrow{u} s$ in $\mathsf{cc}(\Gamma(H))$, and hence $u \in L_q(H)$. □

Let $f(n)$ be the maximal number of finite-index extensions of a subgroup $H \leq_{\mathsf{fg}} F$ such that $\Gamma(H)$ has at most n vertices. By Proposition 2.14, every pair (p,q) such that $p \sim q$ is in the connected component of a pair of the form $(1,r)$ for some $r > 1$. Moreover, this connected component has elements of the form (i,j) for all $1 \leq i \leq n$, so the graph resulting from the identification of 1 and r (or from p and q) has at most $n/2$ vertices. Thus $f(1) = 1$ and $f(n) \leq n\, f(\lfloor n/2 \rfloor)$ for all $n \geq 2$. It follows that $f(n) \leq n^{\frac{1}{2}(1+\log_2 n)}$.

Proposition 2.15 *Let $H \leq_{\mathsf{fg}} F$. If $\mathsf{cc}(\Gamma(H))$ has n vertices, then H has at most $n^{\frac{1}{2}(1+\log_2 n)}$ finite-index extensions.*

Example 2.16 By means of lower bound, we consider the following example. Let $e_1, \ldots, e_k$ be the canonical basis of the vector space $\mathbb{Z}_2^k$, let φ be the morphism from the free group F over $A = \{a_1, \ldots, a_k\}$ into the additive group $\mathbb{Z}_2^k$, mapping a_i to e_i, and let $H = \ker \varphi$. Then H is normal and finite-index, so all its extensions have finite index and they are in bijection with the set of quotients of $\mathbb{Z}_2^k$, hence with the set of subgroups of $\mathbb{Z}_2^k$, or equivalently with the set of subspaces of $\mathbb{Z}_2^k$.

Let $\ell_{d,k}$ be the number of linearly independent d-tuples in $\mathbb{Z}_2^k$ ($d \geq 1$). Then $\ell_{1,k} = 2^k - 1$. If $d \geq 2$, a d-tuple $(x_1, \ldots, x_d)$ is linearly independent if and only if $(x_1, \ldots, x_{d-1})$ is linearly independent and x_d does not belong to the subspace generated by $x_1, \ldots, x_{d-1}$, so that $\ell_{d,k} = \ell_{d-1,k}(2^k - 2^{d-1})$. Now the set of cardinality d linearly independent subsets has $m_{d,k} = \ell_{d,k}/d!$ elements, and the number of dimension d subspaces of $\mathbb{Z}_2^k$ is

$$s_{d,k} = \frac{m_{d,k}}{m_{d,d}} = \frac{\ell_{d,k}}{\ell_{d,d}} = \frac{(2^k-1)(2^k-2)(2^k-4)\cdots(2^k-2^{d-1})}{(2^d-1)(2^d-2)(2^d-4)\cdots(2^d-2^{d-1})}.$$

Finally, the number of subspaces of $\mathbb{Z}_2^k$ is equal to $\sum_{d=0}^{k} s_{d,k}$, with $s_{0,k} = 1$.

We observe that for each $0 \leq i < d < k$, $\frac{2^k-2^i}{2^d-2^i} > 2^{k-d}$, so that $s_{d,k} > 2^{(k-d)d}$. By considering $d = \lfloor \frac{k}{2} \rfloor$, we find that $\sum s_{d,k} > 2^{k^2/4}$.

Finally, we note that $\Gamma(H)$ is the Cayley graph of $\mathbb{Z}_2^k$ with respect to the basis $e_1, \ldots, e_d$ (a graph known as the dimension k hypercube), so that $\Gamma(H)$ has $n = 2^k$ vertices. As a result, H has more than $n^{\frac{1}{4}\log_2 n}$ finite-index extensions. □

2.6 The lattice of finite-index subgroups of G

Let us call fi-*maximal* a subgroup $G \leq_{\mathsf{fg}} F$ which has no proper finite-index extension, that is (in view of Theorem 2.11), such that $G = \mathsf{Comm}_F(G)$. If G is fi-maximal, let $\mathsf{FIS}(G)$ be the set of all finite-index subgroups of G, that is,

the set of subgroups $H \leq_{\mathsf{fg}} F$ such that $H_{\mathsf{fi}} = G$. Note that distinct fi-maximal subgroups yield disjoint lattices of finite-index subgroups.

Fact 2.17 Let $G \leq_{\mathsf{fg}} F$ be non-trivial and fi-maximal. Then $\mathsf{FIS}(G)$ forms a convex sublattice of the lattice of subgroups of F, with greatest element G. This sublattice is always infinite (see Remark 2.1) and without a least element. □

Lemma 2.9 provides us with an invariant for every sublattice of the form $\mathsf{FIS}(G)$ (with G fi-maximal).

Proposition 2.18 *Let $H, K \leq_{\mathsf{fg}} F$ and let $\mathsf{t}_H(1)$ and $\mathsf{t}_K(1)$ be the labels of the tails of $\Gamma(H)$ and $\Gamma(K)$. Then $H_{\mathsf{fi}} = K_{\mathsf{fi}}$ if and only if $\mathsf{t}_H(1) = \mathsf{t}_K(1)$ and $\widetilde{L}_{\tau(1)}(H) = \widetilde{L}_{\tau(1)}(K)$, if and only if $\mathsf{t}_H(1) = \mathsf{t}_K(1)$ and $L_{\tau(1)}(H) = L_{\tau(1)}(K)$.*

Proof. If $H_{\mathsf{fi}} = K_{\mathsf{fi}}$, then $H, K \leq_{\mathsf{fi}} H_{\mathsf{fi}}$, and Lemma 2.9 shows directly that $\mathsf{t}_H(1) = \mathsf{t}_{H_{\mathsf{fi}}}(1) = \mathsf{t}_K(1)$, $\widetilde{L}_{\tau(1)}(H) = \widetilde{L}_{\tau(1)}(H_{\mathsf{fi}}) = \widetilde{L}_{\tau(1)}(K)$ and $L_{\tau(1)}(H) = L_{\tau(1)}(H_{\mathsf{fi}}) = L_{\tau(1)}(K)$.

We now prove the converse. More precisely, we show that if H and K are fi-maximal, $\mathsf{t}_H(1) = \mathsf{t}_K(1)$ and $L_{\tau(1)}(H) = L_{\tau(1)}(K)$, then $H = K$. First we note that $\mathcal{A}(H^{\mathsf{t}_H(1)}) = (\mathsf{cc}(\Gamma(H)), \tau(1))$ and in particular, $H^{\mathsf{t}_H(1)}$ is cyclically reduced and fi-maximal. Thus, it suffices to prove the expected result (namely, that $H = K$) under the hypothesis that H and K are cyclically reduced.

By Fact 2.5, the set of all $L_p(H)$ ($p \in \Gamma(H)$) coincides with the set $\{\{\mathsf{red}(\bar{u}x) \mid x \in L_1(H)\} \mid u \in L_1(H)\}$. In addition, since H is fi-maximal, if $u, v \in L_1(H)$ with $1 \xrightarrow{u} p$ and $1 \xrightarrow{u} q$ and $\{\mathsf{red}(\bar{u}x) \mid x \in L_1(H)\} = \{\mathsf{red}(\bar{v}x) \mid x \in L_1(H)\}$, then $p = q$. It also follows from the same fact that, again if H is fi-maximal, there is an edge (p, a, q) in $\Gamma(H)$ if and only if $L_q(H) = \{\mathsf{red}(\bar{a}x) \mid x \in L_p(H)\}$. Thus, the cyclically reduced fi-maximal subgroup H is entirely determined by the set $L_1(H)$. This concludes the proof. □

The pairs (t, L) that are equal to $(\mathsf{t}(1), \widetilde{L}_{\tau(1)}(G))$ for some subgroup $G \leq_{\mathsf{fg}} F$ are characterized as follows. Recall that an *i-automaton* [9] is a deterministic automaton (Q, i, E, T) over alphabet $\tilde{A}$ such that, $(p, a, q) \in E$ if and only if $(q, \bar{a}, p) \in E$ for all vertices p, q and $a \in A$. The automata $\mathcal{B}_p$ discussed in Section 2.4 are i-automata.

Proposition 2.19 *Let $t \in F$ and let $L \subseteq \tilde{A}^*$ be a rational language. The following conditions are equivalent.*

(1) There exists a subgroup $H \leq_{\mathsf{fg}} F$ such that $t = \mathsf{t}(1)$ and $L = \widetilde{L}_{\tau(1)}(H)$.

(2) There exists a fi-maximal subgroup $H \leq_{\mathsf{fg}} F$ such that $t = \mathsf{t}(1)$ and $L = \widetilde{L}_{\tau(1)}(H)$.

(3) L is accepted by an i-automaton with all states accepting and such that, for each state p, there exist transitions (p, a, q) and (p, b, r) for at least two distinct letters $a, b \in \tilde{A}$. In addition, if t is not the empty word, then $t = t'a$ for a letter $a \in \tilde{A}$ such that $\bar{a} \notin L$.

(3') *The minimal automaton of L is an i-automaton with all states accepting and such that, for each state p, there exist transitions (p, a, q) and (p, b, r) for at least two distinct letters $a, b \in \tilde{A}$. In addition, if t is not the empty word, then $t = t'a$ for a letter $a \in \tilde{A}$ such that $\bar{a} \notin L$.*

(4) *t and L satisfy the following conditions:*

- *(4.1)* *L is closed under taking prefixes;*
- *(4.2)* *if $u, v \in L$, then $u\bar{u}v \in L$;*
- *(4.3)* *if $uv\bar{v}w \in L$, then $uw \in L$;*
- *(4.4)* *if $ua \in L$ with $a \in \tilde{A}$, then $uab \in L$ for some $b \in \tilde{A}$ such that $b \neq \bar{a}$.*
- *(4.5)* *if t is not the empty word, then $t = t'a$ for a letter $a \in \tilde{A}$ such that $\bar{a} \notin L$.*

Proof. Conditions (1) and (2) are equivalent by Proposition 2.18.

Proposition 2.12 shows that (1) implies (3'), which in turn implies (3). Let us now assume that (3) holds and let Γ be the A-labeled graph induced by the states and transitions of the minimal automaton of L. The extra condition given shows that every vertex of Γ is visited by a loop labeled by a cyclically reduced word. It follows that, if G is the subgroup whose graphical representation is $\mathcal{A}(G) = (\Gamma, q_0)$, with q_0 the initial state, then G is cyclically reduced, $L = \widetilde{L}_{q_0}(G)$. The condition on the word t shows that Condition (1) holds with $H = G^t$.

Condition (3) easily implies Condition (4). Let us now assume that Condition (4) holds. By [9, Theorem 4.1], Properties (4.2) and (4.3) show that L is accepted by an i-automaton. Property (4.1) shows that all states of that automaton are final, and Property (4.4) shows that, for each state p, there exist transitions (p, a, q) and (p, b, r) for at least two distinct letters $a, b \in \tilde{A}$. Thus Condition (4) implies Condition (3), which concludes the proof. □

3 Malnormal closure

A subgroup H of F is *malnormal* if $H^g \cap H = 1$ for each $g \notin H$. Malnormality was proved decidable in [1], and a simple decision algorithm was given in [5], based on the following characterization [5, 4].

Proposition 3.1 *Let $H \leq_{\mathsf{fg}} F$. Then H is malnormal if and only if every connected component of $\Gamma(H) \times_A \Gamma(H)$, except for the diagonal complement, is a tree, if and only if, for every $p \neq q \in \mathsf{cc}(\Gamma(H))$, $L_p(H) \cap L_q(H)$ is finite.*

This yields directly an $\mathcal{O}(n^2 \log n)$ decision algorithm, where n is the number of vertices of $\mathsf{cc}(\Gamma(H))$. It also yields the following corollary.

Corollary 3.2 *Let $H \leq_{\mathsf{fg}} F$. If H is malnormal, then H is* fi*-maximal, that is, H has no proper finite-index extension.*

Proof. By Corollary 2.10, if H is not fi-maximal, then there exist vertices $p \neq q$ in $\mathsf{cc}(\Gamma(H))$ such that $L_p(H) = L_q(H)$, and hence such that $L_p(H) \cap L_q(H) = L_p(H)$ is infinite. In particular, H is not malnormal. □

It is shown in [7, Prop. 4.5] that for every finitely generated subgroup $H \leq_{\mathsf{fg}} F$, there exists a least malnormal extension H_{mal}, called the *malnormal closure* of H, that H_{mal} is finitely generated and effectively computable, and that the rank of H_{mal} is less than or equal to the rank of H [7, Corol. 4.14]. In fact, [7] shows that $\Gamma(H_{\mathsf{mal}})$ is obtained from $\Gamma(H)$ by a series of i-steps. The algorithm computing H_{mal} then consists in computing all the quotients of $\Gamma(H)$ and verifying, for each of them, whether it represents a malnormal subgroup. We now give a much better, polynomial-time algorithm.

Theorem 3.3 *Let $H \leq_{\mathsf{fg}} F_n$. The malnormal closure of H is computed by repeatedly applying the following construction: in $\Gamma(H)$, identify all pairs (p, q) of distinct vertices in $\mathsf{cc}(\Gamma(H))$, such that $L_p(H) \cap L_q(H)$ is infinite and reduce the resulting graph.*

Proof. Let $H_0 = H$ and let H_{i+1} be the subgroup of F such that $\Gamma(H_{i+1})$ is obtained from $\Gamma(H_i)$ by, first identifying all pairs p, q of vertices of $\mathsf{cc}(\Gamma(H_i))$ such that $L_p(H_i) \cap L_q(H_i)$ is infinite, and then reducing the resulting graph. Since $\Gamma(H_{i+1})$ has less vertices than $\Gamma(H_i)$, this defines a finite sequence of subgroups

$$H = H_0 < H_1 < \ldots < H_k,$$

where $k \geq 0$ and H_k is malnormal. We want to show that H_k is the least malnormal extension of H.

Let K be any malnormal subgroup of F such that $H \leq K$. By Takahasi's theorem (see for instance [7]), there exists a subgroup G such that $H \leq G \leq K$ such that G is obtained from H by a sequence of i-steps, and K is obtained from G by a sequence of re-steps (this fact can also be deduced from [10, Prop. 2.6]).

Let $p, q \in \mathsf{cc}(\Gamma(H))$ such that $L_p(H) \cap L_q(H)$ is infinite. It is elementary to verify that $L_p(H) \subseteq L_{\varphi_H^G(p)}(G) \subseteq L_{\varphi_G^K(\varphi_H^G(p))}(K)$. In particular, $\varphi_G^K(\varphi_H^G(p)) \cap \varphi_G^K(\varphi_H^G(q))$ is infinite. Since K is malnormal, it follows from Proposition 3.1 that $\varphi_G^K(\varphi_H^G(p)) = \varphi_G^K(\varphi_H^G(q))$. But φ_G^K is one-to-one by definition, so we have $\varphi_H^G(p) = \varphi_H^G(q)$. It follows that G is obtained from H_1 by a sequence of i-steps, and in particular, $H_1 \leq G \leq K$. Iterating this reasoning, we find that $H_k \leq K$, which concludes the proof. □

Corollary 3.4 *Let $H \leq F$. If $\Gamma(H)$ has n vertices, then one can compute $\Gamma(H_{\mathsf{mal}})$ in time $\mathcal{O}(n^3 \log n)$.*

Proof. According to the algorithm given in Proposition 3.3, we first need to compute the connected components of $\mathsf{cc}(\Gamma(H)) \times_A \mathsf{cc}(\Gamma(H))$ – done in time $\mathcal{O}(n^2 \log n)$ –, identifying which are trees and which are not – done in time $\mathcal{O}(n^2)$

–, identifying the vertices of $\mathsf{cc}(\Gamma(H))$ involved in a non-diagonal connected component and reducing the resulting graph to obtain $\Gamma(H_1)$ – which is done in time $\mathcal{O}(n \log n)$. Thus $\Gamma(H_1)$ is computed from $\Gamma(H_0)$ in time $\mathcal{O}(n^2 \log n)$.

This part of the algorithm is iterated k times, to compute $\Gamma(H_k) = \Gamma(H_{\mathsf{mal}})$, and we have $k < n$ since the number of vertices of the $\Gamma(H_i)$ forms a properly decreasing sequence. This concludes the proof. □

Acknowledgements

The authors thank A. Martino and E. Ventura for their help in the computation in Example 2.16.

References

[1] G. Baumslag, A. Miasnikov and V. Remeslennikov, Malnormality is decidable in free groups, *Internat. J. Algebra Comput.* **9** (1999) 687-692.

[2] J. Brzozowski. Canonical regular expressions and minimal state graphs for definite events, in *MRI Symposia Series* **12** 529561, Polytechnic Press, Polytechnic Institute of Brooklyn (1962).

[3] J.-M. Champarnaud and D. Ziadi. Canonical derivatives, partial derivatives and finite automaton constructions, *Theor. Comput. Sci.* **289** (2002) 137-163.

[4] T. Jitsukawa. Malnormal subgroups of free groups, in *Computational and statistical group theory*, pp. 83-95, *Contemp. Math.* **298**, Amer. Math. Soc., 2002.

[5] I. Kapovich and A.G. Myasnikov. Stallings foldings and subgroups of free groups, *J. Algebra* **248** (2002) 608-668.

[6] D. Kozen. *Automata and Computability*, Springer, 1997.

[7] A. Miasnikov, E. Ventura, P. Weil. Algebraic extensions in free groups, in *Algebra and Geometry in Geneva and Barcelona* (G.N. Arzhantseva, L. Bartholdi, J. Burillo, E. Ventura eds.), Trends in Mathematics, Birkhaüser (2007), pp. 225-253.

[8] J.-P. Serre. *Arbres, amalgames, SL_2*, Astérisque **46**, Soc. Math. France, 1977. English translation: *Trees*, Springer Monographs in Mathematics, Springer, 2003.

[9] P.V. Silva. On free inverse monoid languages, *Theoretical Informatics and Applications* **30** (1996) 349-378.

[10] P.V. Silva, P. Weil. On an algorithm to decide whether a free group is a free factor of another, *Theoretical Informatics and Applications* **42** (2008) 395-414.

[11] J. Stallings. The topology of graphs, *Inventiones Mathematicæ* **71** (1983) 551–565.

[12] P. Weil. Computing closures of finitely generated subgroups of the free group, in *Algorithmic Problems in Groups and Semigroups* (J.-C. Birget, S. Margolis, J. Meakin, M. Sapir eds.), Birkhaüser, 2000, pp. 289 - 307.