Subgroups of free groups: a contribution to the Hanna Neumann conjecture[1]

J. Meakin
Dept of Mathematics and Statistics
University of Nebraska
Lincoln, Nebraska 68588-0323 – U.S.A.
jmeakin@math.unl.edu

P. Weil
LaBRI
Université Bordeaux-I and CNRS
351 cours de la Libération
33405 Talence Cedex – France.
Pascal.Weil@labri.u-bordeaux.fr

Résumé. On démontre que la conjecture forte de Hanna Neumann, sur le rang de l'intersection de sous-groupes de type fini d'un groupe libre, est vérifiée pour une large classe de sous-groupes, caractérisés par des propriétés géométriques. Un cas particulier de notre résultat entraîne que la conjecture est vraie pour les sous-groupes de type fini du groupe libre $F(A)$ (de base A) qui sont positivement engendrés, c'est-à dire engendrés par un ensemble fini de mots du monoïde libre de base A.

Abstract. We prove that the strengthened Hanna Neumann conjecture, on the rank of the intersection of finitely generated subgroups of a free group, holds for a large class of groups characterized by geometric properties. One particular case of our result implies that the conjecture holds for all positively finitely generated subgroups of the free group $F(A)$ (over the basis A), that is, for subgroups which admit a finite set of generators taken in the free monoid over A.

It has been known since Howson's 1954 paper [7] that the intersection of finitely generated subgroups of a free group is finitely generated. In fact, Howson gave an upper bound for the rank of the intersection $H \cap K$, in terms of the ranks of H and K.

In 1956, Hanna Neumann [12, 13] improved on Howson's bound, and formulated the following conjecture. If H is a free group, we call $\widetilde{r}(H) = \max(\mathrm{rk}(H) - 1, 0)$ the *reduced rank* of H. Hanna Neumann showed that if H and K are finitely generated subgroups of a free group, then $\widetilde{r}(H \cap K) \leq 2\widetilde{r}(H)\widetilde{r}(K)$, and she conjectured that

$$\widetilde{r}(H \cap K) \leq \widetilde{r}(H)\widetilde{r}(K).$$

[1] The authors gratefully acknowledge partial support from NSF grants number DMS-9970471 and DMS-0070701, and from 1999 INTAS grant number 1224 *Combinatorial and geometric theory of groups and semigroups and its applications to computer science.*

This conjecture is referred to as the **Hanna Neumann conjecture**. Hanna Neumann's proven bound was improved by Burns [2] in 1971. In 1991, Walter Neumann [14] formulated the following, stronger conjecture. Let H, K be finitely generated subgroups of a free group F and let $H\backslash F/K = \{HxK \mid x \in F\}$ be the set of double cosets. If $x \in F$, H^x denotes the conjugate $H^x = x^{-1}Hx$. Note that if $HxK = HyK$, then $\tilde{r}(H^x \cap K) = \tilde{r}(H^y \cap K)$.

Strengthened Hanna Neumann conjecture

If H and K are finitely generated subgroups of a free group F, then

$$\sum_{HgK \in H\backslash F/K} \tilde{r}(H^g \cap K) \leq \tilde{r}(H)\tilde{r}(K).$$

We will say that $SHN(H, K)$ *holds* if the above inequality is satisfied, and we say that SHN *holds* for H if $SHN(H, K)$ holds for every finitely generated subgroup K of $F(A)$.

Walter Neumann [14] extended Burns' bound to the strengthened version of the conjecture, and Tardos [18] proved that SHN holds for all subgroups of rank 2. Tardos [19] further improved on Burns' bound for the strengthened conjecture in 1996. As far as we are aware, the best bound known to date is due to Dicks and Formanek [4], and it states that

$$\sum_{HgK \in H\backslash F/K} \tilde{r}(H^g \cap K) \leq 2\tilde{r}(H)\tilde{r}(K) - 2(\tilde{r}(H) + \tilde{r}(K)) + 4,$$

proving in particular that SHN holds for subgroups of rank 3.

The aim of this paper is to prove that SHN holds for certain classes of subgroups H, characterized by geometric properties of the directed graph classically associated with a finitely generated subgroup of a free group. One particular case of our result implies that SHN holds for all positively finitely generated subgroups of the free group $F(A)$ (over the basis A), that is, for subgroups which admit a finite set of generators taken in the free monoid over A. This extends the well-known result that SHN holds for finite index subgroups (see [14]).

Our methods are elementary, and do not seem to relate in any obvious way with Dicks' alternative approach, whereby the problem is transformed into one on bipartite graphs [3]. Some of the results contained in this paper were announced in [11].

1 The graph-theoretical tools

It has become commonplace to use graph-theoretical tools to work with subgroups of a free group, and we will rely on this approach as well (Imrich [8, 9], Stallings [17], Gersten [5], Servatius [16], Nickolas [15], W. Neumann [14], Tardos [18], Dicks [3], etc.).

In this approach, one associates with each subgroup H of the free group $F(A)$ with basis A, a canonical *directed A-labeled graph* (that is, a directed graph whose edges are labeled with elements of A) denoted by $\Gamma(H)$.

Before we recall the definition of $\Gamma(H)$, let us make a few definitions concerning A-labeled graphs more precise. In a directed A-labeled graph, we consider paths which may traverse edges in either direction: the convention is that traversing an a-labeled edge ($a \in A$) in the backward direction carries the label a^{-1}. Thus every path has a label which is a word in the monoid $(A \sqcup A^{-1})^*$, freely generated by the disjoint union $A \sqcup A^{-1}$.

We say that a word $w \in (A \sqcup A^{-1})^*$ is *positive* if it contains no letter of A^{-1}, that is, if w lies in the free monoid A^*. A path is *positively labeled* if its label is a positive word.

A word in $(A \sqcup A^{-1})^*$ is said to be *(group-) reduced* if it does not contain any subword of the form aa^{-1} or $a^{-1}a$ ($a \in A$). The set of reduced words is in natural bijection with the free group $F(A)$. A path is said to be a *loop* if its initial and its final vertices coincide, and it is said to be *reduced* if its label is a reduced word. We say that a directed A-labeled graph is *connected* if given any pair of distinct vertices (u, v), there exists a path from u to v. Finally, such a graph is said to be *admissible* if it is connected and, for each $a \in A$ and each vertex v, there is at most one a-labeled edge starting at v, and at most one a-labeled edge into v.

The graph $\Gamma(H)$ can be constructed as follows: first we consider the graph $\Gamma_0(H)$, which represents $F(A)$ acting by right translations on the set of right H-cosets. The vertex set of $\Gamma_0(H)$ is the set $\{Hx \mid x \in F(A)\}$ and its edge set is $\{Hx \mid x \in F(A)\} \times A$: the pair (Hx, a) is an edge from Hx to Hxa, with label a. Next, $\Gamma(H)$ is the *H-core* of $\Gamma_0(H)$, that is, the least subgraph of $\Gamma_0(H)$ containing all the reduced loops around the vertex H. We denote by $V(H)$ and $E(H)$ the sets of vertices and edges of $\Gamma(H)$.

By construction, $\Gamma(H)$ is admissible. In addition, it is well-known [17] that $\Gamma(H)$ is finite if and only if H is finitely generated, that $\widetilde{r}(H) = |E(H)| - |V(H)|$, and that $\Gamma(H)$ uniquely characterizes the subgroup H. $\Gamma(H)$ can also be defined in terms of immersions on the bouquet of $|A|$ circles (Stallings [17]), or in terms of inverse automata (Margolis and Meakin [10]). Moreover, if H is given by a finite set of generators, $\Gamma(H)$ can be effectively constructed [17] (in fact in polynomial time [1]).

We say that a vertex v of a directed graph has *valence k* if v is adjacent to k edges in the underlying undirected graph. We call any vertex of valence less than 2 *extremal*, and any other vertex *internal*. For instance, in $\Gamma(H)$, there is at most one extremal vertex, and that is the vertex H. Iteratively removing extremal vertices in a directed A-labeled graph Γ yields the *core* of Γ, denoted by $c\Gamma$. If $c\Gamma$ is empty, we say that Γ is *contractible*: equivalently, Γ is a tree, or otherwise said, there is no reduced loop around any vertex of Γ.

In this framework, the intersection of subgroups of $F(A)$ can be represented in terms of pull-backs in the category of A-labeled graphs (i.e. the category of immersions into the bouquet of $|A|$ circles, Stallings [17]). More precisely, if H and K are subgroups of $F(A)$, we let $\Delta(H, K)$ be the directed A-labeled graph whose vertex set is $V(H) \times V(K)$, and whose a-labeled edges ($a \in A$) are of the form $((Hx, Ky), a)$, from vertex (Hx, Ky) to vertex (Hxa, Kya). It is well-known that $\Gamma(H \cap K)$ is isomorphic to the (H, K)-core of $\Delta(H, K)$. As $\Delta(H, K)$ may be disconnected, $\Gamma(H \cap K)$ is also isomorphic to the (H, K)-core of the connected component of $\Delta(H, K)$ containing the vertex (H, K).

Now proving the Hanna Neumann conjecture amounts to proving that the difference $|E|-|V|$ between the number of edges and the number of vertices in the connected component of the vertex (H,K) in $\Delta(H,K)$, is bounded above by $\tilde{r}(H)\tilde{r}(K) = (|E(H)|-|V(H)|)(|E(K)|-|V(K)|)$. It was also observed in [14, 18] that

$$\sum_{HgK \in H\backslash F(A)/K} \tilde{r}(H^g \cap K) = \sum(|E|-|V|)$$

where the second sum runs over all the connected components of $\Delta(H,K)$ which are not contractible. In particular, the strengthened Hanna Neumann conjecture SHN amounts to proving that $\sum(|E|-|V|) \leq \tilde{r}(H)\tilde{r}(K)$

Finally, let us note the following simple fact.

Lemma 1.1 *Let H be a finitely generated subgroup of $F(A)$ and let $\varphi\colon F(A) \to F(B)$ be an injective morphism. If f is a function such that, for each finitely generated subgroup K of $F(B)$, we have*

$$\sum_{\varphi(H)gK \in \varphi(H)\backslash F(B)/K} \tilde{r}(\varphi(H)^g \cap K) \leq f(\tilde{r}(H), \tilde{r}(K)),$$

then for each finitely generated subgroup K of $F(A)$ we have

$$\sum_{HgK \in H\backslash F(B)/K} \tilde{r}(H^g \cap K) \leq f(\tilde{r}(H), \tilde{r}(K)).$$

In particular, if SHN holds for $\varphi(H)$, then SHN holds for H.

Proof. For each subgroup K of $F(A)$, $\varphi(K)$ is a subgroup of $F(B)$ and $\tilde{r}(\varphi(K)) = \tilde{r}(K)$. Moreover, if $g \in F(A)$, then $\varphi(H^g) = \varphi(H)^{\varphi(g)}$ and φ maps $H\backslash F(A)/K$ injectively into $\varphi(H)\backslash F(B)/\varphi(K)$. The proof follows immediately from these elementary observations. $\square$

2 A result on the subgroups of a free group of rank two

In this section, we restrict our attention to a free group of rank two, and A denotes the 2-letter alphabet $A = \{a, b\}$.

For each vertex v of an admissible A-labeled graph Γ, let n_v be the number of edges out of v, and let $m_v = n_v - 1$. Then $m_v \in \{-1, 0, 1\}$, and we say that the vertex v is *positive* (resp. *zero*, *negative*) if m_v is equal to 1 (resp. 0, -1). In particular, if E and V are the sets of edges and of vertices of Γ, and if P and N are the sets of positive and negative vertices of V, then we have

$$|E| - |V| = \sum_{v \in V} m_v = |P| - |N|.$$

The aim of this section is to prove the following theorem.

Theorem 2.1 *Let H and K be finitely generated subgroups of $F(A)$. If $\Gamma(H)$ has n negative vertices, then*

$$\sum_{HgK \in H\backslash F(A)/K} \tilde{r}(H^g \cap K) \le (\tilde{r}(H) + n)\tilde{r}(K).$$

2.1 Basic paths and zig-zags

Let Δ be a directed A-labeled graph. An *acceptable* path in Δ is a positively labeled path ending in a negative vertex, such that all its interior vertices (the vertices other than the origin and the end of the path) are zero vertices. A *basic* path is an acceptable path starting in a positive vertex.

Lemma 2.2 *Let Δ be an admissible A-labeled graph. In an acceptable path of Δ, no vertex occurs twice. Moreover, two acceptable paths of Δ having a common zero vertex, coincide from that vertex on. In particular, there are at most two basic paths originating in a given positive vertex.*

Proof. Let $u \xrightarrow{p} v$ be an acceptable path, and let w be the vertex closest to the end of the path, occurring twice in that path. Note that w cannot be the last vertex of the path, which is negative and hence is not the origin of any edge. Then p can be factored as $p = p_1 p_2 p_3$, with p_2 and p_3 non-empty and

$$u \xrightarrow{p_1} w \xrightarrow{p_2} w \xrightarrow{p_3} v.$$

If the first edges of p_2 and p_3 have the same label, say a, then the (uniquely determined) vertex at the end of an a-labeled edge starting at w, occurs twice in the path p, contradicting the choice of w. So, say, the first edge of p_2 is labeled a and the first edge of p_3 is labeled b. In that case, w is a positive vertex, a contradiction which proves the first assertion.

The second assertion is immediate, since a zero vertex is the origin of a single edge. It follows that distinct basic paths have distinct first edges, whence the third assertion. □

Lemma 2.3 *Let Δ be an admissible A-labeled graph, without extremal vertices. Every edge into a negative vertex is the last edge of a basic path.*

Proof. Let v_0 be a negative vertex and let $v_1 \xrightarrow{a_1} v_0$ be an edge into v_0. Since v_0 is negative, we have $v_1 \ne v_0$. Suppose we have constructed an acceptable path $v_i \xrightarrow{a_i} v_{i-1} \cdots v_1 \xrightarrow{a_1} v_0$ $(i \ge 1)$ such that $v_0, v_1, \ldots, v_i$ are pairwise distinct.

If v_i is positive, we are done. Otherwise, v_i is a zero vertex and there is a positive edge into v_i, say $v_{i+1} \xrightarrow{a_{i+1}} v_i$ (since Δ has no extremal vertices).

If $v_{i+1} = v_j$ for some $0 \leq j \leq i$, then v_{i+1} occurs twice in an acceptable path, contradicting Lemma 2.2. Thus, as long as the v_i ($i \geq 1$) are all zero vertices, we are able to increase the length of our acceptable path. Since the graph Δ is finite, this process cannot go on forever and v_i is positive for some i, thus proving the lemma. □

We now collect basic paths together. We say that a sequence of vertices Z is a *zig-zag* if it is of the form $Z = (u_0, v_1, u_1, \ldots, v_m, u_m)$, where the v_i are pairwise distinct negative vertices, the u_i are positive, and there exist basic paths $u_i \longrightarrow v_i$ and $u_j \longrightarrow v_{j+1}$ for $0 < i \leq m$ and $0 \leq j < m$; we further require that no two of these $2m$ basic paths have a common edge.

In view of Lemma 2.2, this implies that the u_i ($1 \leq i \leq m-1$) are pairwise distinct, and all the basic paths out of these vertices end in $\{v_1, \ldots, v_m\}$. It also implies that $u_0, u_m \notin \{u_1, \ldots, u_{m-1}\}$, and that if $u_0 = u_m$, then all the basic paths out of u_0 end in $\{v_1, \ldots, v_m\}$.

Let R be a set of negative vertices. We say that a zig-zag *avoids* R if it contains no negative vertex in R. Zig-zags are partially ordered by inclusion: a zig-zag Z is *contained in* a zig-zag Z' if Z is a factor of Z', that is, if Z can be extended to the right or the left (or both) to yield Z'.

Lemma 2.4 *Let Δ be an admissible A-labeled graph without extremal vertices, and let R be a set of negative vertices of Δ. Let $Z = (u_0, v_1, u_1, \ldots, v_m, u_m)$ be a zig-zag avoiding R. Then Z is maximal among the zig-zags avoiding R if and only if all the basic paths out of u_0 and u_m end in $\{v_1, \ldots, v_m\} \cup R$.*

Proof. If Z is not maximal among the zig-zags avoiding R, then by definition of a zig-zag, one of u_0 and u_m is the origin of a basic path ending outside $\{v_1, \ldots, v_m\} \cup R$.

For the converse, we consider a zig-zag $Z = (u_0, v_1, u_1, \ldots, v_m, u_m)$, and we assume that one of u_0 and u_m, say u_m, is the origin of a basic path ending in a negative vertex $v_{m+1} \notin \{v_1, \ldots, v_m\} \cup R$. Since Δ has no extremal vertices, there are two edges ending in v_{m+1}. Therefore, by Lemma 2.3 and Lemma 2.2, v_{m+1} is the extremity of a basic path of the form $u_{m+1} \longrightarrow v_{m+1}$, having no edge in common with the basic path $u_m \longrightarrow v_{m+1}$.

Thus $(u_0, v_1, u_1, \ldots, v_m, u_m, v_{m+1}, u_{m+1})$ is a zig-zag avoiding R, which proves the lemma. □

2.2 Matchings

We now come to a technically important *matching* property.

Proposition 2.5 *Let Δ be an admissible A-labeled graph without extremal vertices, and let $v_1, \ldots, v_t$ be the negative vertices of Δ. Then there exist pairwise distinct positive vertices $u_1, \ldots, u_t$ such that, for each $1 \leq i \leq t$, there exists a basic path from u_i to v_i.*

Proof. Let N be the set of negative vertices of Δ and let $t = |N|$. The result is immediate if $t = 0$. We now assume that $t > 0$. Note that, by Lemma 2.3, if $R \subseteq N$ and $v \in N \setminus R$, then v is contained in a zig-zag avoiding R, and hence in one that is maximal among the zig-zags avoiding R.

We construct a collection of zig-zags $Z_1, \ldots, Z_k$ such that each negative vertex of Δ occurs in exactly one of the Z_i, and if a positive vertex u occurs in Z_i, then all the basic paths starting in u end in a negative vertex in some Z_j with $j \leq i$.

Let Z_1 be a maximal zig-zag and let R_1 be the set of negative vertices in Z_1. By Lemma 2.4 (applied with $R = \emptyset$), every basic path out of a positive vertex in Z_1 ends in R_1.

Suppose that we have constructed zig-zags $Z_1, \ldots, Z_i$ and sets $R_1, \ldots, R_i$ such that no negative vertex occurs in two different Z_j, each R_j is the set of the negative vertices in the Z_h, $h \leq j$, and if a positive vertex u occurs in Z_j, then all the basic paths starting in u end in R_j. If $R_i = N$, we are done.

Otherwise, let Z_{i+1} be a zig-zag that is maximal among the zig-zags avoiding R_i, and let R_{i+1} be the union of R_i and the set of negative vertices in Z_{i+1}. Then Lemma 2.4 shows that the basic paths originating in the positive vertices in Z_{i+1} end in R_{i+1}. This concludes the construction.

Now we verify that if a positive vertex u occurs in some Z_i, then it does not occur in any of the Z_j, $j \neq i$. Suppose indeed that u occurs in Z_i and in Z_j, with $i < j$. By definition of a zig-zag, there is a basic path from u to a negative vertex in Z_j, in contradiction with the property that the basic paths out of u all end in R_i.

It is now easy to extract, from each of the Z_i, a set of pairwise disjoint basic paths connecting the negative vertices with some of the positive vertices of Z_i. As no positive or negative vertex occurs in two different Z_i, the union of these matchings is a matching of all the negative vertices of Δ with some of its positive vertices via basic paths, which completes the proof. □

Investigating how such matchings behave under the operation of a pull-back, we prove the following.

Proposition 2.6 *Let Γ_1 and Γ_2 be admissible A-labeled graphs, such that Γ_2 has no extremal vertex. Let Δ be the pull-back of Γ_1 and Γ_2, and let Γ be the union of the non-contractible connected components of Δ. Finally let P_1 and N_1 (resp. P_2 and N_2, P and N) be the sets of positive and negative vertices of Γ_1 (resp. Γ_2, Γ). Then we have*

$$|P| - |N| \leq |P_1|(|P_2| - |N_2|).$$

Proof. Let $\mathfrak{M}$ be a matching of the negative vertices of Γ_2 with some of its positive vertices, as per Proposition 2.5: $\mathfrak{M} = \{(u_1 \to v_1), \ldots, (u_t \to v_t)\}$ where the $u_i \to v_i$ are basic paths of Γ_2, the positive vertices u_i are pairwise distinct, the negative vertices v_i are pairwise distinct, and all the negative vertices of Γ_2 occur in $\mathfrak{M}$. In particular, $P_2 - N_2$ is equal to the number of positive vertices of Γ_2 which do not occur in $\mathfrak{M}$.

We now *lift* $\mathfrak{M}$ in Γ. Formally, for each basic path $u \xrightarrow{p} v$ in $\mathfrak{M}$ and for each positive vertex of the form (u', u) in Γ (i.e. u' is a positive vertex in Γ_1 and (u', u) lies in Γ), we consider the longest prefix p' of p which labels a path in Δ (or equivalently in Γ), starting in (u', u).

Since we are assuming that (u', u) is positive, there are both an a-labeled and a b-labeled edge out of (u', u) in Δ, so the prefix p' is non-empty. Let (z', z) be the vertex reached after following the path labeled p' and starting at vertex (u', u). Then (z', z) is a negative vertex of Γ. Indeed, if there is an edge out of (z', z), say, with label a, then there are a-labeled edges out of z' in Γ_1 and out of z in Γ_2. In particular, $z \neq v$ and $p'a$ must be a prefix of p, contradicting the choice of p'. So the path $(u', u) \longrightarrow (z', z)$ labeled p' is a basic path of Γ.

Let $\mathfrak{N}$ be the collection of basic paths of Γ thus constructed. By definition, their initial positive vertices are pairwise distinct (i.e. we construct exactly one path for each positive vertex of Γ with second coordinate in $\mathfrak{M}$). Let us verify that the negative vertices at the end of these basic paths are also pairwise distinct.

Let us assume that $(u', u) \xrightarrow{p_1} (z', z)$ and $(w', w) \xrightarrow{p_2} (z', z)$ are constructed by lifting paths in $\mathfrak{M}$ as above. In particular, $u \xrightarrow{p_1} z$ and $w \xrightarrow{p_2} z$ are prefixes of basic paths in $\mathfrak{M}$. As such paths are pairwise disjoint, they must be prefixes of the same basic path, so that $u = w$ and $\mathfrak{M}$ contains a basic path $u \xrightarrow{p} v$ such that p_1 and p_2 are prefixes of p. If $p_1 \neq p_2$, then the vertex z occurs twice in the basic path $u \xrightarrow{p} v$, a contradiction. So $p_1 = p_2$ and hence, as Γ is admissible, the paths $(u', u) \xrightarrow{p_1} (z', z)$ and $(w', w) \xrightarrow{p_2} (z', z)$ are identical.

To conclude, we observe that $\mathfrak{N}$ establishes a bijection between a subset of P and a subset of N. If P' (resp. N') is the set of positive (resp. negative) vertices of Γ not in $\mathfrak{N}$, then $|P| - |N| = |P'| - |N'|$.

By construction of $\mathfrak{N}$, if u is a positive vertex occurring in $\mathfrak{M}$, then every positive vertex of the form (u', u) in Γ occurs in $\mathfrak{N}$. Contrapositively, if a positive vertex (u', u) of Δ does not occur in $\mathfrak{N}$, then u does not occur in $\mathfrak{M}$. Thus $|P'| \leq |P_1|(|P_2| - |N_2|)$. It follows that $|P| - |N| \leq |P_1|(|P_2| - |N_2|)$, which concludes our proof. □

2.3 Proof of Theorem 2.1

Let H, K be non-trivial finitely generated subgroups of $F(A)$. There exists an element $x \in F(A)$ such that $\Gamma(K^x)$ has no extremal vertex: it suffices to take for x the label of a path in $\Gamma(K)$ from vertex K to a vertex of the core $c\Gamma(K)$. Now we observe that if $Hg_1K, \ldots, Hg_rK$ is a list of the elements of $H\backslash F(A)/K$, then $Hg_1xK^x, \ldots, Hg_rxK^x$ is a list of the elements of $H\backslash F(A)/K^x$. Moreover, $H^{gx} \cap K^x = (H^g \cap K)^x$, so we get

$$\sum_{HgK^x \in H\backslash F(A)/K^x} \tilde{r}(H^g \cap K^x) = \sum_{HgK \in H\backslash F(A)/K} \tilde{r}(H^g \cap K).$$

Let $\Delta(H, K^x)$ be the pull-back of $\Gamma(H)$ and $\Gamma(K^x)$, and let Γ be the union of the non-contractible connected components of $\Delta(H, K^x)$. We let $P(H)$ and $N(H)$ (resp. $P(K^x)$ and

$N(K^x)$, P and N) be the set of positive and negative vertices of $\Gamma(H)$ (resp. $\Gamma(K^x)$, Γ). As $\Gamma(K)$ has no extremal vertex, we may apply Proposition 2.6 to $\Gamma_1 = \Gamma(H)$ and $\Gamma_2 = \Gamma(K^x)$ to get

$$|P| - |N| \leq |P(H)|(|P(K^x)| - |N(K^x)|) = (\tilde{r}(H) + n)\tilde{r}(K^x) = (\tilde{r}(H) + n)\tilde{r}(K).$$

But $|P| - |N| = \sum_{HgK \in H\backslash F(A)/K} \tilde{r}(H^g \cap K)$. So we have proved the required inequality, namely

$$\sum_{HgK \in H\backslash F(A)/K} \tilde{r}(H^g \cap K) \leq (\tilde{r}(H) + n)\tilde{r}(K).$$

3 Applications

3.1 Subgroups of a free group of rank two

In this section again, A denotes the two-letter alphabet, $A = \{a, b\}$. First we observe that Theorem 2.1 immediately implies the strengthened Hanna Neumann conjecture for an interesting class of subgroups of the free group of rank two.

Theorem 3.1 *Let H be a finitely generated subgroup of $F(\{a,b\})$. If $\Gamma(H)$ has no negative vertex, then SHN holds for H.*

Over a two-letter alphabet, negative vertices are a special case of valence 2 vertices, where the two incident edges on the given vertex are oriented towards that vertex. Let us now call such a vertex a *sink*. There are three other types of valence 2 vertices: *sources*, where the two incident edges are outgoing edges; *ab-vertices*, with an incoming a-labeled edge and an outgoing b-labeled edge; and the dually defined *ba-vertices*.

Theorem 3.2 *Let H be a finitely generated subgroup of $F(\{a,b\})$. If the core of $\Gamma(H)$ has no sink (resp. no source, no ab-vertex, no ba-vertex), then SHN holds for H.*

Proof. We know that for some $x \in F(\{a,b\})$, $\Gamma(H^x)$ is isomorphic to the core of $\Gamma(H)$. Now by Lemma 1.1, it suffices to show that SHN holds for H^x. Thus we may assume without loss of generality that $\Gamma(H)$ is equal to $c\Gamma(H)$.

The statement relative to the no-sink case is exactly Theorem 3.1. In order to prove the no-source (resp. no ab-vertex, no ba-vertex) statement, we consider the automorphism φ of $F(\{a,b\})$ given by $\varphi(a) = a^{-1}$ and $\varphi(b) = b^{-1}$ (resp. $\varphi(a) = a$ and $\varphi(b) = b^{-1}$, $\varphi(a) = a^{-1}$ and $\varphi(b) = b$). It is immediate that the graph $\Gamma(\varphi(H))$ has no negative vertices. Thus SHN holds for $\varphi(H)$, and hence it does for H as well by Lemma 1.1. □

Note that $\Gamma(H)$ may have a sink when its core does not: take for instance $H = \langle ab^{-1}ab^2a^{-1}\rangle$. So Theorem 3.2 is somewhat more general than Theorem 3.1, even in the no-sink case.

In the case where both $c\Gamma(H)$ and $c\Gamma(K)$ contain all four types of valence 2 vertices, Theorem 2.1 gives a bound on $\sum_{HgK\in H\backslash F(A)/K}\tilde{r}(H^g\cap K)$ which may be better, in many cases, than the bound obtained by Dicks (see the introduction), but it seems difficult to transcribe the geometric parameters of this bound (essentially the minimum of the numbers of sinks, sources, ab-vertices or ba-vertices in $c\Gamma(H)$ and $c\Gamma(K)$) into algebraic parameters.

3.2 Positively generated subgroups

In this section, A denotes any finite alphabet, possibly with more than two letters. We say that a subgroup H of a free group $F(A)$ is *positively generated* if it has a set of generators consisting of positive words.

As $\Gamma(H)$ is the union of the reduced loops around the vertex H in $\Gamma_0(H)$ (see Section 1), and as the labels of these loops are exactly the reduced words in H, it is not difficult to see that if $\{h_i\}_{i\in I}$ is a generating set of H, then $\Gamma(H)$ is the union of the reduced loops around the vertex H labeled by the h_i. In particular, if H is positively generated, then every vertex of $\Gamma(H)$ sits on a positively labeled loop around H.

Theorem 3.3 *Let H be a finitely, positively generated subgroup of $F(A)$. Then SHN holds for H.*

Proof. We already observed that every vertex of $\Gamma(H)$ is the origin of an edge. Thus, if A has 2 letters, then $\Gamma(H)$ has no negative vertex and SHN holds for H by Theorem 2.1.

In general, suppose that $A = \{a_1, \ldots, a_n\}$ with $n \geq 2$. For each $1 \leq i \leq n$, let $\varphi(a_i) = a^ib^i$. Then φ defines an injective morphism from $F(A)$ into $F(\{a, b\})$, which maps positive words to positive words. In particular, if H is a positively generated subgroup of $F(A)$, then $\varphi(H)$ is also positively generated. Therefore SHN holds for $\varphi(H)$, and hence SHN holds for H by Lemma 1.1. □

We note that being positively generated is not an *algebraic* property of a subgroup, in the sense that it is not preserved under automorphisms of the free group. For instance, $H = \langle abba^{-1}\rangle$ is positively generated. However, if φ is the automorphism φ defined by $\varphi(a) = ab^{-1}$ and $\varphi(b) = b$, then $\varphi(H) = \langle abab^{-1}\rangle$ is not positively generated. It is nevertheless interesting to note the following characterization of positively generated subgroups of $F(A)$.

Proposition 3.4 *Let H be a finitely generated subgroup of the free group $F(A)$. Then H is positively generated if and only if each vertex of $\Gamma(H)$ sits on a positively labeled loop around the vertex H, or equivalently, if and only if the directed graph $\Gamma(H)$ is strongly connected.*

Proof. We already noticed that if H is positively generated, then every vertex of $\Gamma(H)$ sits on a positively labeled loop around the vertex H. We now prove that this condition is sufficient for H to be positively generated.

Let $H_0 = \{1\}$ and let Γ_0 be the admissible graph reduced to the single vertex H (and no edge). We construct increasing sequences $(H_n)_{n\geq 0}$ and $(\Gamma_n)_{n\geq 0}$ of subgroups of $F(A)$ and subgraphs of $\Gamma(H)$ such that for each n, $\Gamma_n = \Gamma(H_n)$ and H_n is positively generated.

Let $n \geq 0$ and let us assume that H_n and Γ_n are constructed. If Γ_n does not contain all the edges of $\Gamma(H)$, we construct a larger graph Γ_{n+1} as follows. We choose an edge e of $\Gamma(H)$ which is not in Γ_n, say $e\colon v \to w$, with label $a \in A$. Next we choose positively labeled paths $H \xrightarrow{p} v$ and $w \xrightarrow{q} H$. Finally we let Γ_{n+1} be the union of Γ_n, paths p and q and edge e (with the corresponding vertices), and we let H_{n+1} be the subgroup generated by H_n and paq. Then H_{n+1} is positively generated, and it is easy to verify that $\Gamma(H_{n+1}) = \Gamma_{n+1}$.

Thus, as long as $\Gamma_n \neq \Gamma(H)$, we construct a new, larger subgraph Γ_{n+1} with a positively generated associated subgroup. For some n at most equal to the number of edges of $\Gamma(H)$, $\Gamma_n = \Gamma(H)$, so that $H = H_n$ is positively generated. $\square$

The above proof is in fact constructive. Together with the classical results on the complexity of a depth-first search in a graph, it yields the following corollary.

Corollary 3.5 *Given a finitely generated subgroup H of $F(A)$, one can decide whether H is positively generated in time $O(n \log n)$. If H is positively generated, one can exhibit a positive basis of H.*

It follows from a result of Héam [6, Prop. 14] that the finitely, positively generated subgroups are exactly the finitely generated subgroups which are the topological closure of a submonoid of A^* in the profinite topology of $F(A)$.

Note also that for a subgroup H of $F(\{a, b\})$, being positively generated is not equivalent to $\Gamma(H)$ having no sink. For example, let $H = \langle abab^{-1}, b^2a^{-1}\rangle$: H is not positively generated, yet $\Gamma(H)$ doesn't have negative vertices.

3.3 Concluding remarks

We remark that Theorem 3.3 extends the well-known result on the solution of the strengthened Hanna Neumann conjecture for finite index subgroups. Indeed, if H has finite index in $F(A)$, then $\Gamma(H)$ is a cover of the bouquet of A circles, that is, every vertex is the origin of an a-labeled edge for each letter $a \in A$ [17]. In view of Proposition 3.4, it follows that finite index subgroups are always positively generated.

A simple way to extend our results is by using Lemma 1.1. For the purpose of this discussion, say that a subgroup H of $F(\{a, b\})$ is *favorable* if it satisfies the hypothesis of Theorem 3.2, that is, if $c\Gamma(H)$ avoids one of the four types of valence 2 vertices, namely sources, sinks, ab-vertices and ba-vertices. Then SHN holds for any subgroup of a free group which is the pre-image by an injective morphism of a favorable subgroup.

Unfortunately, we do not know how to characterize the finitely generated subgroups of free groups which fall in this category. In fact, at this point, we do not even have an example of a finitely generated subgroup H of a free group such that no injective image of H in $F(\{a,b\})$ is favorable.

For instance, we observe that the class of favorable subgroups of $F(\{a,b\})$ is not preserved by automorphisms. Indeed let $H = \langle abab^{-1} \rangle$ and let φ be defined on $F(\{a,b\})$ by $\varphi(a) = ab$ and $\varphi(b) = b$. Then φ is an automorphism of $F(\{a,b\})$ and $\varphi(H) = \langle abba \rangle$. Thus $\varphi(H)$ is positively generated even though H is not favorable.

It would be of interest to characterize the class of subgroups of a free group that admit an injective morphism onto a favorable subgroup.

References

[1] J.-C. Birget, S. Margolis, J. Meakin, P. Weil. PSPACE-completeness of certain algorithmic problems on the subgroups of free groups, *Theoret. Comp. Sci.*, to appear.

[2] R.G. Burns. On the intersection of finitely generated subgroups of a free group, *Math. Z.* **119** (1971) 121–130.

[3] W. Dicks. Equivalence of the strengthened Hanna Neumann conjecture and the amalgamated graph conjecture, *Invent. Math.* **117** (1994) 373–389.

[4] W. Dicks and E. Formanek. The rank three case of the Hanna Neumann conjecture, to appear.

[5] S.M. Gersten. Intersections of finitely generated subgroups of free groups and resolutions of graphs, *Invent. Math.* **71** (1983) 567–591.

[6] P.-C. Héam. Some topological properties of rational sets, to appear.

[7] A.G. Howson. On the intersection of finitely generated free groups, *J. London Math. Soc.* **29** (1954) 428–434.

[8] W. Imrich. On finitely generated subgroups of free groups, *Arch. Math.* **28** (1977) 21–24.

[9] W. Imrich. Subgroup theorems and graphs, in C.H.C. Little (ed.) *Combinatorial Mathematics*, Lect. Notes Math. **622**, 1–27, Springer, 1977.

[10] S. Margolis, J. Meakin. Free inverse monoids and graph immersions, *Intern. J. of Algebra and Computation* **3** (1993) 79-100.

[11] J. Meakin, An invitation to inverse semigroups, in K.P. Shum and P.C. Yuen (eds.) *Proc. Hong Kong Conf. on Ordered Structures and Algebra of Computer Languages*, 91–115 , World Scientific, 1993.

[12] H. Neumann. On intersections of finitely generated subgroups of free groups, *Publ. Math. Debrecen* **4** (1956) 186–189.

[13] H. Neumann. On intersections of finitely generated subgroups of free groups. Addendum, *Publ. Math. Debrecen* **5** (1958) 128.

[14] W.D. Neumann. On intersections of finitely generated subgroups of free groups, in L.G. Kovács (ed.) *Groups - Canberra 1989*, Lect. Notes in Math. **1456**, 161–170, Springer, 1990.

[15] P. Nickolas. Intersections of finitely generated free groups, *Bull. Austr. Math. Soc.* **31** (1985) 339–348.

[16] B. Servatius. A short proof of a theorem of Burns, *Math. Z.* **184** (1983) 133–137.

[17] J.R. Stallings. Topology of finite graphs, *Invent. Math.* **71** (1983) 551–565.

[18] G. Tardos. On the intersection of subgroups of a free group, *Invent. Math.* **108** (1992) 29–36.

[19] G. Tardos. Towards the Hanna Neumann conjecture using Dicks' method, *Invent. Math.* **123** (1996) 95–104.