JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 28, Number 1, January 2015, Pages 63–97
S 0894-0347(2014)00796-7
Article electronically published on April 17, 2014

MEASURE PRESERVING WORDS ARE PRIMITIVE

DORON PUDER AND ORI PARZANCHEVSKI

Contents

1. Introduction

This paper establishes a new characterization of primitive elements in free groups that is based on the distributions these elements induce on finite groups. Let $\mathbf{F}_k$ be the free group on k generators $X = \{x_1, \ldots, x_k\}$, and let $w = \prod_{j=1}^{r} x_{i_j}^{\varepsilon_j}$ ($\varepsilon_j = \pm 1$) be a word in $\mathbf{F}_k$. For every group G, w induces a *word map* from the Cartesian product G^k to G, by substitutions:

$$w : (g_1, \ldots, g_k) \mapsto \prod_{j=1}^{r} g_{i_j}^{\varepsilon_j}.$$

The word w is called *measure preserving* with respect to a finite group G if all the fibers of this map are of equal size. Namely, every element in G is obtained by substitutions in w the same number of times. We say that w is *measure preserving* if it is measure preserving w.r.t. every finite group. The last years have seen a great interest in word maps in groups, and the distributions they induce. We refer the reader, for instance, to [Sha09, LS09, AV11, PS13], and to the recent book [Seg09] and survey [Sha13]. Several authors have also studied words which are asymptotically measure preserving on finite simple groups; see e.g. [LS08, GS09, BK13].

Received by the editors January 31, 2013 and, in revised form, October 15, 2013, and November 23, 2013.
2010 *Mathematics Subject Classification.* Primary 20E05; Secondary 20E18, 20B30, 68R15, 06A11.

©2014 American Mathematical Society

The word w is called *primitive* if it belongs to some basis (free generating set) of $\mathbf{F}_k$. It is a simple observation (see 1.2 below) that primitive words are measure preserving, and several authors have conjectured that the converse is also true, namely, that measure preservation implies primitivity.[†] From private conversations we know that this has occurred to the following mathematicians and has been discussed among themselves: N. Avni, T. Gelander, M. Larsen, A. Lubotzky and A. Shalev. The question was independently raised in [LP10] and also in [AV11], alongside a generalization of it (see Section 8).

In [Pud13] the first author proved the conjecture for $\mathbf{F}_2$. Here we prove it in full:

Theorem 1.1. *A measure preserving word is primitive.*

A key ingredient of the proof is the extension of the problem from single words to (finitely generated) subgroups of $\mathbf{F}_k$. The concept of primitive words extends naturally to the notion of free factors: let H be a subgroup of the free group J (in particular, H is free as well). We say that H is a *free factor* of J, and denote this by $H \stackrel{*}{\leq} J$, if there is a subgroup $H' \leq J$ such that $H * H' = J$. Equivalently, $H \stackrel{*}{\leq} J$ iff some basis of H can be extended to a basis of J. (This in turn is easily seen to be equivalent to the condition that *every* basis of H extends to a basis of J.)

In order to generalize the notion of measure preservation to subgroups, we need to change our perspective of word maps a little. One can think of the word map w as the evaluation map from $\mathrm{Hom}(\mathbf{F}_k, G)$ to G, i.e., $w(\alpha) = \alpha(w)$ for $\alpha \in \mathrm{Hom}(\mathbf{F}_k, G)$. The identification of $\mathrm{Hom}(\mathbf{F}_k, G)$ with G^k depends on the chosen basis, and is due to the fact that a homomorphism from a free group is uniquely determined by choosing the images of the elements of a basis, and these images can be chosen arbitrarily.

In this perspective, w is measure preserving w.r.t. G if the element $\alpha_G(w)$ is uniformly distributed over G, where $\alpha_G \in \mathrm{Hom}(\mathbf{F}_k, G)$ is a homomorphism chosen uniformly at random. If w is primitive, then it belongs to some basis, and identifying $\mathrm{Hom}(\mathbf{F}_k, G)$ and G^k according to this basis gives

Observation 1.2. *A primitive word is measure preserving.*

We can now extend the notion of measure preservation from words to finitely generated subgroups (we write $H \leq_{fg} \mathbf{F}_k$ when H is a finitely generated subgroup of $\mathbf{F}_k$):

Definition 1.3. Let $H \leq_{fg} \mathbf{F}_k$. We say that H is *measure preserving* if for every finite group G and $\alpha_G \in \mathrm{Hom}(\mathbf{F}_k, G)$ a random homomorphism chosen with uniform distribution, $\alpha_G\big|_H$ is uniformly distributed in $\mathrm{Hom}(H, G)$.

This can be reformulated in terms of distributions of subgroups: observe the distribution of the random subgroup $\alpha_G(H) \leq G$, where $\alpha_G \in \mathrm{Hom}(\mathbf{F}_k, G)$ distributes uniformly. Then H is measure preserving if the distribution of $\alpha_G(H)$ is the same as that of the image of a uniformly chosen homomorphism from $\mathbf{F}_{\mathrm{rk}(H)}$ to G (where $\mathrm{rk}(H)$ denotes the rank of H).

As for single words, it is immediate that a free factor is measure preserving, and again it is natural to conjecture that the converse also holds. Since $1 \neq w \in \mathbf{F}_k$ is measure preserving iff $\langle w \rangle$ is measure preserving, this is an extension of

[†]It is interesting to note that there is an easy abelian parallel to this conjecture. A word $w \in \mathbf{F}_k$ belongs to a basis of $\mathbb{Z}^k \cong \mathbf{F}_k/\mathbf{F}_k'$ iff for any group G the associated word map is surjective. See [Seg09], Lemma 3.1.1.

the conjecture regarding words. In [Pud13] the first author proved the extended conjecture for subgroups of $\mathbf{F}_k$ of rank $\geq k-1$ (thus proving the conjecture for $\mathbf{F}_2$), but the techniques used in that paper are specialized for the proven cases. In this paper we introduce completely new techniques, which yield the extended conjecture in full:

Theorem 1.4. *A measure preserving subgroup is a free factor.*

In Section 8 we explain how this circle of ideas is related to the study of profinite groups and decidability questions. In fact, part of the original motivation for this study comes from this relation. In particular we have the following corollary (see also Corollary 8.1):

Corollary 1.5. *The set P of primitive elements in $\mathbf{F}_k$ is closed in the profinite topology.*

In plain terms, this amounts to the assertion that every non-primitive word in $\mathbf{F}_k$ is contained in a primitive-free coset of a finite index subgroup.

In order to prove Theorem 1.4, one needs to exhibit, for each non-primitive word $w \in \mathbf{F}_k$, some "witness" finite group with respect to which w is not measure preserving. Our witnesses are always the symmetric groups S_n. In fact, it is enough to restrict one's attention to the average number of fixed points in the random permutation $\alpha_{S_n}(w)$ (which we also denote by $\alpha_n(w)$). We summarize this in the following stronger version of Theorems 1.1 and 1.4:

Theorem (1.4′)**.** *Let $w \in \mathbf{F}_k$, and for every finite group G, let $\alpha_G \in \mathrm{Hom}(\mathbf{F}_k, G)$ denote a random homomorphism chosen with uniform distribution. Then the following are equivalent:*

(1) w is primitive.
(2) w is measure preserving: for every finite group G the random element $\alpha_G(w)$ has uniform distribution.
(3) For every $n \in \mathbb{N}$ the random permutation $\alpha_n(w) = \alpha_{S_n}(w)$ has uniform distribution.
(4) For every $n \in \mathbb{N}$, the expected number of fixed points in the random permutation $\alpha_n(w) = \alpha_{S_n}(w)$ is 1:

$$\mathbb{E}\left[\#\mathrm{fix}\left(\alpha_n(w)\right)\right] = 1.$$

(5) For infinitely many $n \in \mathbb{N}$,

$$\mathbb{E}\left[\#\mathrm{fix}\left(\alpha_n(w)\right)\right] \leq 1.$$

The analogue properties for f.g. subgroups are equivalent as well. For example, the parallel of property (4) for $H \leq_{fg} \mathbf{F}_k$ is that for every n, the image $\alpha_n(H) \subseteq S_n$ stabilizes on average exactly $n^{1-\mathrm{rk}(H)}$ elements of $\{1, \ldots, n\}$.

We already explained above the implication (1) ⇒ (2), and (2) ⇒ (3) ⇒ (4) ⇒ (5) is evident (recall that a uniformly distributed random permutation has exactly one fixed point on average). The only nontrivial, somewhat surprising part, is the implication (5) ⇒ (1) which is proven in this paper. It turns out that an effective bound can also be obtained:

Proposition 1.6. *A word w of length $\ell > 0$ is primitive iff $\mathbb{E}\left[\#\mathrm{fix}\left(\alpha_n(w)\right)\right] = 1$ for $n \leq \ell$.*

An analogue result holds for subgroups (see Corollary 6.6).

A key role in our proof is played by the notion of *primitivity rank*, an invariant classifying words and f.g. subgroups of $\mathbf{F}_k$, which was first introduced in [Pud13]: A primitive word $w \in \mathbf{F}_k$ is also primitive in every subgroup containing it (Claim 3.9(3)). However, if w is not primitive in $\mathbf{F}_k$, it may be either primitive or non-primitive in subgroups of $\mathbf{F}_k$ containing it. But what is the smallest rank of a subgroup giving evidence to the imprimitivity of w? Informally, how far does one have to search in order to establish that w is not primitive? Concretely,

Definition 1.7. The *primitivity rank* of $w \in \mathbf{F}_k$, denoted $\pi(w)$, is

$$\pi(w) = \min\left\{ \mathrm{rk}(J) \,\middle|\, \begin{array}{c} w \in J \le \mathbf{F}_k \ s.t. \\ w \text{ is } \mathbf{not} \text{ primitive in } J \end{array} \right\}.$$

If no such J exists, i.e. if w is primitive, then $\pi(w) = \infty$.

More generally, for $H \le_{fg} \mathbf{F}_k$, the *primitivity rank* of H is

$$\pi(H) = \min\left\{ \mathrm{rk}(J) \,\middle|\, \begin{array}{c} H \le J \le \mathbf{F}_k \ s.t. \\ H \text{ is } \mathbf{not} \text{ a free factor of } J \end{array} \right\}.$$

Again, if no such J exists, then $\pi(H) = \infty$. We call a subgroup J for which the minimum is obtained **H-critical**, and denote the set of H-critical subgroups by $\mathrm{Crit}(H)$. The set of w-critical subgroups of a word w is defined analogously.

Note that for $w \ne 1$, $\pi(w) = \pi(\langle w \rangle)$. Let us give a few examples: $\pi(w) = 0$ iff $w = 1$; $\pi(w) = \infty$ iff w is primitive, and $\pi(H) = \infty$ iff H is a free factor; $\pi(w) = 1$ if and only if w is a proper power, namely $w = v^d$ for some $v \in \mathbf{F}_k$ and $d \ge 2$, and then $\mathrm{Crit}(w) = \{\langle v^m \rangle : m \mid d,\ 1 \le m < d\}$ (assuming that v itself is not a power). By [Pud13, Lemma 6.8], $\pi\left(x_1^2 \ldots x_r^2\right) = r$ for every $1 \le r \le k$. We thus have that π takes all values in $\{0, 1, 2, \ldots, k\} \cup \{\infty\}$, and Claim 3.9(3) shows that these are all the values it obtains. The primitivity rank of a word or a subgroup is computable—this is shown in Section 4. The distribution of the primitivity rank is discussed in [Pud12] and [PW14].

In this paper we sometimes find it more convenient to deal with *reduced ranks* of subgroups: $\widetilde{\mathrm{rk}}(H) \overset{def}{=} \mathrm{rk}(H) - 1$. We therefore define analogously the *reduced primitivity rank*, $\widetilde{\pi}(\cdot) \overset{def}{=} \pi(\cdot) - 1$.

As mentioned above, our main result follows from an analysis of the average number of common fixed points of $\alpha_n(H)$ (where α_n denotes a uniformly distributed random homomorphism in $\mathrm{Hom}(\mathbf{F}_k, S_n)$). In other words, we count the number of elements in $\{1, \ldots, n\}$ stabilized by the images under α_n of *all* elements of H. Theorem 1.4′ follows from the main result of this analysis:

Theorem 1.8. *The average number of common fixed points of $\alpha_n(H)$ is*

$$\frac{1}{n^{\widetilde{\mathrm{rk}}(H)}} + \frac{|\mathrm{Crit}(H)|}{n^{\widetilde{\pi}(H)}} + O\left(\frac{1}{n^{\widetilde{\pi}(H)+1}}\right).$$

In particular, for a word w

$$\mathbb{E}\left[\#\mathrm{fix}\left(\alpha_n(w)\right)\right] = 1 + \frac{|\mathrm{Crit}(w)|}{n^{\widetilde{\pi}(w)}} + O\left(\frac{1}{n^{\widetilde{\pi}(w)+1}}\right).$$

We remark that $\mathrm{Crit}(H)$ is always finite (see Section 4). Table 1 summarizes the connection implied by Theorem 1.8 between the primitivity rank of w and the average number of fixed points in the random permutation $\alpha_n(w)$.

$\pi(w)$	Description of w	$\mathbb{E}[\#\text{fix}(\alpha_n(w))]$
0	$w = 1$	n
1	w is a power	$1 + \lvert \text{Crit}(w) \rvert + O\left(\frac{1}{n}\right)$
2	E.g. $[x_1, x_2], x_1^2 x_2^2$	$1 + \frac{\lvert \text{Crit}(w) \rvert}{n} + O\left(\frac{1}{n^2}\right)$
3		$1 + \frac{\lvert \text{Crit}(w) \rvert}{n^2} + O\left(\frac{1}{n^3}\right)$
$\vdots$		$\vdots$
k	E.g. $x_1^2 \ldots x_k^2$	$1 + \frac{\lvert \text{Crit}(w) \rvert}{n^{k-1}} + O\left(\frac{1}{n^k}\right)$
∞	w is primitive	1

TABLE 1. Primitivity Rank and Average Number of Fixed Points.

Theorem 1.8 implies the following general corollary regarding the family of distributions of S_n induced by word maps:

Corollary 1.9. *For a non-primitive $w \in \mathbf{F}_k$ the average number of fixed points in $\alpha_n(w)$ is strictly greater than* 1*, for large enough n.*

Corollary 1.9 is in fact the missing piece (5) $\Rightarrow$ (1) in Theorem 1.4′. In addition, it follows from this corollary that for every $w \in \mathbf{F}_k$ and large enough n, the average number of fixed points in $\alpha_n(w)$ is at least one.† In other words, primitive words generically induce a distribution of S_n with the fewest fixed points on average.

The results stated above validate completely the conjectural picture described in [Pud13]. Theorem 1.8 and its consequences, Corollaries 8.1, 1.5 and 1.9, are stated there as conjectures (Conjectures 1.10, 7.1, 7.2 and 8.2).

The analysis of the average number of fixed points in $\alpha_n(w)$ has its roots in [Nic94]. Nica notices that by studying the various quotients of a labeled cycle-graph (corresponding to w), one can compute a rational expression which gives this average for every large enough n. When $w = v^d$ with d maximal (so v is not a power), he shows that the limit distribution of the number of fixed points in $\alpha_n(w)$ (as $n \to \infty$) is $\delta(d) + O\left(\frac{1}{n}\right)$, where $\delta(d)$ is the number of divisors of d ([Nic94], Corollary 1.3).‡ Nica's result follows from Theorem 1.8: if $w \neq 1$ is a proper power, $w = v^d$ with $d \geq 2$ maximal, then $\lvert \text{Crit}(w) \rvert = \delta(d) - 1$, and if it is not a power, then $\widetilde{\pi}(w) \geq 1$.

The results of this paper have interesting implications in the study of expansion in random graphs: in [Pud12], the first author presents a new approach to showing that random graphs are nearly optimal expanders. A crucial ingredient in the proof is Theorem 1.8. More particularly, it was conjectured by Alon [Alo86] that the spectral gap of a random d-regular graph is a.a.s. arbitrarily close to $d - 2\sqrt{d-1}$, and this conjecture was generalized by Friedman [Fri03] to non-regular graphs. In [Fri08], Alon's conjecture is proved by highly sophisticated arguments, which are not applicable for the generalized conjecture (as far as is known). The results in [Pud12] give a simple proof which nearly recovers Friedman's results regarding

†It is suggestive to ask whether this holds for *all* n. Namely, is it true that for every $w \in \mathbf{F}_k$ and every n, the average number of fixed points in $\alpha_n(w)$ is at least 1? By results of Abért ([Abe06]), this statement turns out to be false.

‡Nica's result is in fact more general: the same statement holds not only for fixed points but for cycles of length L for every fixed L.

Alon's conjecture, and can be applied also for the generalized conjecture, giving the best results as of now regarding non-regular graphs.

2. Overview of the proof

The proof of our main theorem involves several structures of posets (partially ordered sets) on $\mathfrak{sub}_{fg}(\mathbf{F}_k)$, the set of finitely generated subgroups of $\mathbf{F}_k$. This set has, of course, a natural structure of a poset given by the relation of inclusion. However, there are other interesting partial orders defined on it: the relation of *algebraic extensions*, and the family of relations defined by *covers*. We introduce some notation: if $\preceq$ is some partial order on $\mathfrak{sub}_{fg}(\mathbf{F}_k)$, and $H, J \leq_{fg} \mathbf{F}_k$, we define the *closed interval*

$$[H,J]_{\preceq} = \{L \in \mathfrak{sub}_{fg}(\mathbf{F}_k) \mid H \preceq L \preceq J\}$$

and similarly the open interval $(H,J)_{\preceq} = \{L \mid H \precneqq L \precneqq J\}$, the half-bounded interval $[H,\infty)_{\preceq} = \{L \mid H \preceq L\}$, and so on (see also the glossary).

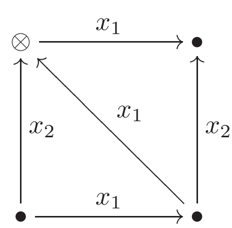
Algebraic Extensions. This notion goes back to [Tak51], and was further studied in [KM02, MVW07].

Definition 2.1. We say that J is an *algebraic extension* of H, denoted $H \leq_{alg} J$, if $H \leq J$ and H is not contained in any proper free factor of J.

The terminology comes from similarities (that go only to some extent) between this notion and that of algebraic extensions of fields (in this line of thought, J is a *transcendental extension* of H when $H \overset{*}{\leq} J$). We devote Section 4 to studying this relation. It is clearly reflexive and antisymmetric, but it is also transitive (Claim 4.1). In addition, it is very sparse: it turns out that $[H,\infty)_{alg}$, the set of algebraic extensions of H, is finite for every $H \leq_{fg} \mathbf{F}_k$, so in particular $(\mathfrak{sub}_{fg}(\mathbf{F}_k), \leq_{alg})$ is locally finite.† It is a simple observation that H-critical subgroups are in particular algebraic extensions of H, i.e. $\mathrm{Crit}(H) \subseteq [H,\infty)_{alg}$. In fact, they are the proper algebraic extensions of minimal rank.

X-cover. For every basis $X = \{x_1,\ldots,x_k\}$ of $\mathbf{F}_k$ there is a partial order denoted $\leq_{\overline{X}}$, which is based on the notion of quotients, or surjective morphisms, of *core graphs*. Introduced in [Sta83], core graphs provide a geometric approach to the study of free groups (for an extensive survey see [KM02], and also [MVW07] and the references therein). Given the basis X, Stallings associates with every $H \leq \mathbf{F}_k$ a directed and pointed graph denoted $\Gamma_X(H)$, whose edges are labeled by the elements of X. A full definition appears in Section 3, but we illustrate the concept in Figure 2.1. It shows the core graph of the subgroup of $\mathbf{F}_2$ generated by $x_1x_2^{-1}x_1$ and $x_1^{-2}x_2$, with $X = \{x_1,x_2\}$.

Figure 2.1. The core graph $\Gamma_X(H)$ where $X = \{x_1,x_2\}$ and $H = \langle x_1x_2^{-1}x_1, x_1^{-2}x_2\rangle \leq \mathbf{F}_2$.



†A *locally finite* poset is one in which every closed interval $[a,b] = \{x : a \leq x \leq b\}$ is finite.

The order $\leq_{\overrightarrow{X}}$ is defined as follows: for $H, J \leq \mathbf{F}_k$ one has $H \leq_{\overrightarrow{X}} J$ iff the associated core graph $\Gamma_X(J)$ is a quotient (as a pointed labeled graph) of the core graph $\Gamma_X(H)$ (see Definition 3.3). When $H \leq_{fg} \mathbf{F}_k$, $\Gamma_X(H)$ is finite (Claim 3.1(1)), and thus has only finitely many quotients. As it turns out that different groups correspond to different core graphs, this implies that $(\mathfrak{sub}_{fg}(\mathbf{F}_k), \leq_{\overrightarrow{X}})$ is locally finite too. We stress that we have here an infinite family of partial orders, one for every choice of basis for $\mathbf{F}_k$. Although the dependency on the basis makes these orders somewhat less universal, they turn out to be the most useful for our purposes.

The various relations between subgroups of $\mathbf{F}_k$ are the following:

$$J \in \mathrm{Crit}(H) \Rightarrow H \leq_{alg} J \Rightarrow H \leq_{\overrightarrow{X}} J \Rightarrow H \leq J$$

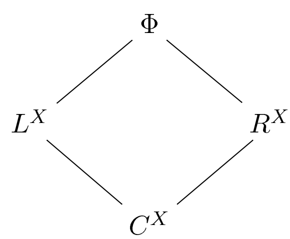
for any $H, J \leq \mathbf{F}_k$ and any basis X (see Sections 3 and 4).

Recall that the main theorems of this paper follow from Theorem 1.8, which estimates the expected number of common fixed points of $\alpha_n(H)$, where $H \leq_{fg} \mathbf{F}_k$ and α_n is a random homomorphism in $\mathrm{Hom}(\mathbf{F}_k, S_n)$. This result is achieved by studying a broader question: for every pair of $H, J \leq_{fg} \mathbf{F}_k$ such that $H \leq J$, we define for $n \in \mathbb{N}$

$$\Phi_{H,J}(n) = \text{The expected number of common fixed points of } \alpha_{J,n}(H), \tag{2.1}$$

where $\alpha_{J,n} \in \mathrm{Hom}(J, S_n)$ is a random homomorphism (chosen with uniform distribution). In this perspective, Nica finds $\lim_{n\to\infty} \Phi_{\langle w\rangle,\mathbf{F}_k}(n)$, and shows that it separates powers and non-powers. Theorem 1.8 shows that the first two terms in the expansion of $\Phi_{\langle w\rangle,\mathbf{F}_k}(n)$ yield w's primitivity rank, which in particular distinguishes powers ($\pi(w) = 1$) and primitives ($\pi(w) = \infty$). Furthermore, the same holds for subgroups using $\Phi_{H,\mathbf{F}_k}(n)$.

As remarked, in order to understand $\Phi_{H,\mathbf{F}_k}$ we turn to analyze the totality of functions $\Phi_{H,J}$, for various $H \leq J \leq \mathbf{F}_k$. We apply the machinery of Möbius inversions to the incidence algebra arising from the locally finite poset $(\mathfrak{sub}_{fg}(\mathbf{F}_k), \leq_{\overrightarrow{X}})$. The local finiteness of the order $\leq_{\overrightarrow{X}}$ allows us to "derive" the function Φ and obtain its "right derivation" R^X, its "left derivation" L^X, and its "two sided derivation" C^X (see Section 5). For instance, $\Phi_{H,J}$ can be presented as finite sums of R^X:



$$\Phi_{H,J} = \sum_{M \in [H,J]_{\overrightarrow{X}}} R^X_{H,M}$$

(here $[H,J]_{\overrightarrow{X}}$ is an abbreviation for $[H,J]_{\leq_{\overrightarrow{X}}}$, i.e. $[H,J]_{\overrightarrow{X}} = \{M \mid H \leq_{\overrightarrow{X}} M \leq_{\overrightarrow{X}} J\}$).

The proof of Theorem 1.8 is then based on a series of lemmas and propositions characterizing Φ and its three derivations:

- (Proposition 5.1) The right derivation R^X is supported on algebraic extensions; i.e. if $H \leq_{\overrightarrow{X}} M$ but M is not an algebraic extension of H, then $R^X_{H,M} \equiv 0$.
- (The discussion in Section 6) The random homomorphism $\alpha_{J,n} \in \mathrm{Hom}(J, S_n)$ can be encoded as a random covering space $\widehat{\Gamma}$ of the core graph $\Gamma_X(J)$, and $\Phi_{H,J}(n)$ can then be interpreted as the expected number of lifts of $\Gamma_X(H)$ into $\widehat{\Gamma}$.

- (Lemmas 6.3 and 6.4) The left derivation L^X is the expected number of *injective* lifts of the core graph $\Gamma_X(H)$ into the random covering $\widehat{\Gamma}$ of the core graph $\Gamma_X(J)$, and a rational expression can be computed for $L^X_{H,J}$.
- (Proposition 7.1 and Section 7.1) An analysis involving Stirling numbers of the rational expressions for L^X yields a combinatorial meaning for the two-sided derivation C^X. Using the classification of primitivity rank we then obtain a first-order estimate for the size of $C^X_{H,J}$.
- (Proposition 7.2) From C^X we return to R^X (by "left-integration"), obtaining that whenever $H \leq_{alg} M$ we have
$$R^X_{H,M} = \frac{1}{n^{\widetilde{\mathrm{rk}}(M)}} + O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(M)+1}}\right)$$
and by right integration of R^X, we obtain the order of magnitude of Φ, which was our goal.

The paper is arranged as follows: in Section 3 the notion of core graphs is explained in detail, as well as the partial order $\leq_{\vec{X}}$ and some of the results from [Pud13] which are used here. In Section 4 we survey the main properties of algebraic extensions of free groups. Section 5 is devoted to recalling Möbius derivations on locally finite posets and introducing the different derivations of Φ. In Section 6 we discuss the connection of the problem to random coverings of graphs and analyze the left derivation L^X. The proof of Theorem 1.8 is completed in Section 7 via the analysis of the two-sided derivation C^X and the consequence of the latter on the right derivation R^X. Finally, corollaries of our results to the field of profinite groups, and to decidability questions in group theory, are discussed in Section 8. We finish with a list of open problems naturally arising from this paper. For the reader's convenience, there is also a glossary of notions and notations at the end of this manuscript.

3. Core graphs and the partial order of covers

Fix a basis $X = \{x_1, \ldots, x_k\}$ of $\mathbf{F}_k$. Associated with every subgroup $H \leq \mathbf{F}_k$ is a directed, pointed graph whose edges are labeled by X. This graph is called *the (Stallings) core graph associated with* H and is denoted by $\Gamma_X(H)$. We recall the notion of the Schreier (right) coset graph of H with respect to the basis X, denoted by $\overline{\Gamma}_X(H)$. This is a directed, pointed and edge-labeled graph. Its vertex set is the set of all right cosets of H in $\mathbf{F}_k$, where the basepoint corresponds to the trivial coset H. For every coset Hw and every basis-element x_j there is a directed j-edge (short for x_j-edge) going from the vertex Hw to the vertex Hwx_j.[†]

The core graph $\Gamma_X(H)$ is obtained from $\overline{\Gamma}_X(H)$ by omitting all the vertices and edges of $\overline{\Gamma}_X(H)$ which are not traced by any reduced (i.e., non-backtracking) path that starts and ends at the basepoint. Stated informally, we trim all "hanging trees" from $\overline{\Gamma}_X(H)$. Formally, $\Gamma_X(H)$ is the induced subgraph of $\overline{\Gamma}_X(H)$ whose vertices are all cosets Hw (with w reduced), such that for some word w' the concatenation ww' is reduced, and $w \cdot w' \in H$. To illustrate, Figure 3.1 shows the graphs $\overline{\Gamma}_X(H)$

[†]Alternatively, $\overline{\Gamma}_X(H)$ is the quotient $H\backslash T$, where T is the Cayley graph of $\mathbf{F}_k$ with respect to the basis X, and F_k (and thus also H) acts on this graph from the left. Moreover, this is the covering space of $\overline{\Gamma}_X(F_k) = \Gamma_X(F_k)$, the bouquet of k loops, corresponding to H, via the correspondence between pointed covering spaces of a space Y and subgroups of its fundamental group $\pi_1(Y)$.

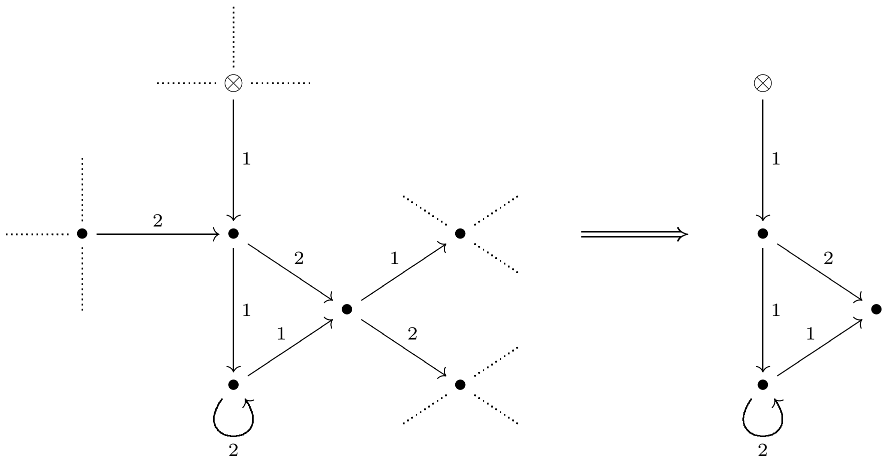
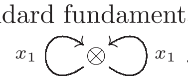


FIGURE 3.1. $\overline{\Gamma}_X(H)$ and $\Gamma_X(H)$ for $H = \langle x_1 x_2 x_1^{-3}, x_1^2 x_2 x_1^{-2} \rangle \leq \mathbf{F}_2$. The Schreier coset graph $\overline{\Gamma}_X(H)$ is the infinite graph on the left (the dotted lines represent infinite 4-regular trees). The basepoint "⊗" corresponds to the trivial coset H, the vertex below it corresponds to the coset Hx_1, the one further down corresponds to $Hx_1^2 = Hx_1x_2x_1^{-1}$, etc. The core graph $\Gamma_X(H)$ is the finite graph on the right, which is obtained from $\overline{\Gamma}_X(H)$ by omitting all vertices and edges that are not traced by reduced closed paths around the basepoint.

and $\Gamma_X(H)$ for $H = \langle x_1 x_2 x_1^{-3}, x_1^2 x_2 x_1^{-2} \rangle \leq \mathbf{F}_2$. Note that the graph $\overline{\Gamma}_X(H)$ is $2k$-regular: every vertex has exactly one outgoing j-edge and one incoming j-edge, for every $1 \leq j \leq k$. Every vertex of $\Gamma_X(H)$ has *at most* one outgoing j-edge and *at most* one incoming j-edge, for every $1 \leq j \leq k$.

If Γ is a directed pointed graph labeled by some set X, paths in Γ correspond to words in $\mathbf{F}(X)$ (the free group generated by X). For instance, the path (from left to right)

corresponds to the word $x_2^2 x_1 x_2^{-1} x_3 x_1^{-1}$. The set of all words obtained from closed paths around the basepoint in Γ is a subgroup of $\mathbf{F}(X)$ which we call the *labeled fundamental group* of Γ, and denote by $\pi_1^X(\Gamma)$. Note that $\pi_1^X(\Gamma)$ need not be isomorphic to $\pi_1(\Gamma)$, the standard fundamental group of Γ viewed as a topological space—for example, take $\Gamma =$ 

.

However, it is not hard to show that when Γ is a core graph, then $\pi_1^X(\Gamma)$ *is* isomorphic to $\pi_1(\Gamma)$ (e.g. [MVW07]). In this case the labeling gives a canonical identification of $\pi_1(\Gamma)$ as a subgroup of $\mathbf{F}(X)$. It is an easy observation that

$$\pi_1^X\left(\overline{\Gamma}_X(H)\right) = \pi_1^X(\Gamma_X(H)) = H. \tag{3.1}$$

This gives a one-to-one correspondence between subgroups of $\mathbf{F}(X) = \mathbf{F}_k$ and core graphs labeled by X. Namely, π_1^X and Γ_X are the inverses of each other in a bijection (Galois correspondence)

$$\left\{ \begin{array}{c} \text{Subgroups} \\ \text{of } \mathbf{F}(X) \end{array} \right\} \underset{\pi_1^X}{\overset{\Gamma_X}{\rightleftarrows}} \left\{ \begin{array}{c} \text{Core graphs} \\ \text{labeled by } X \end{array} \right\}. \tag{3.2}$$

Core graphs were introduced by Stallings [Sta83]. Our definition is slightly different, and closer to the one in [KM02, MVW07] in that we allow the basepoint to be of degree one, and in that our graphs are directed and edge labeled. We remark that it is possible to study core graphs from a purely combinatorial point of view, as labeled, pointed, and connected graphs satisfying

(1) No two equally labeled edges originate or terminate at the same vertex.
(2) Every vertex and edge are traced by some non-backtracking closed path around the basepoint.

Starting with this definition, every choice of an ordered basis for $\mathbf{F}_k$ then gives a correspondence between these graphs and subgroups of $\mathbf{F}_k$.

In this paper we are mainly interested in finite core graphs, and we now list some basic properties of these (proofs can be found in [Sta83, KM02, MVW07]).

Claim 3.1. *Let H be a subgroup of $\mathbf{F}_k$ with an associated core graph $\Gamma = \Gamma_X(H)$. The Euler Characteristic of a graph, denoted $\chi(\cdot)$, is the number of vertices minus the number of edges.*

(1) $\mathrm{rk}(H) < \infty \Longleftrightarrow \Gamma$ *is finite.*
(2) $\widetilde{\mathrm{rk}}(H) = -\chi(\Gamma)$.
(3) The correspondence (3.2) *restricts to a correspondence between* $\mathfrak{sub}_{fg}(\mathbf{F}_k)$ *and finite core graphs.*

Given a finite set of words $\{h_1, \ldots, h_m\} \subseteq \mathbf{F}(X)$ that generate a subgroup H, the core graph $\Gamma_X(H)$ can be algorithmically constructed as follows. Every h_i corresponds to some path with directed edges labeled by the x_j's (we assume the elements are given in reduced forms; otherwise we might need to prune leaves at the end of the algorithm). Merge these m paths to a single graph (bouquet) by identifying all their $2m$ end-points to a single vertex, which is marked as the basepoint. The labeled fundamental group of this graph is clearly H. Then, as long as there are two j-labeled edges with the same terminus (resp. origin) for some j, merge the two edges and their origins (resp. termini). Such a step is often referred to as *Stallings folding*. It is fairly easy to see that each folding step does not change the labeled fundamental group of the graph, that the resulting graph is indeed $\Gamma_X(H)$, and that the order of folding has no significance. To illustrate, we draw in Figure 3.2 a folding process by which we obtain the core graph $\Gamma_X(H)$ of $H = \langle x_1 x_2 x_1^{-3}, x_1^2 x_2 x_1^{-2} \rangle \leq \mathbf{F}_2$ from the given generating set.

A *morphism* between two core graphs is a map that sends vertices to vertices and edges to edges, and preserves the structure of the core graphs. Namely, it preserves the incidence relations, sends the basepoint to the basepoint, and preserves the directions and labels of the edges.

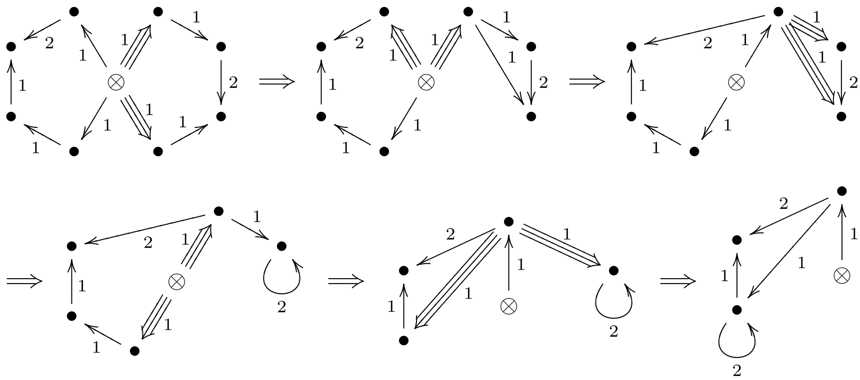


FIGURE 3.2. Constructing the core graph $\Gamma_X(H)$ of $H = \langle x_1x_2x_1^{-3}, x_1^2x_2x_1^{-2}\rangle \le \mathbf{F}_2$ from the given generating set. We start with the upper left graph which contains a distinct loop at the basepoint for each (reduced) element of the generating set. Then, at an arbitrary order, we merge pairs of equally labeled edges which share the same origin or the same terminus (here we mark by triple arrows the pair of edges being merged next). The graph at the bottom right is $\Gamma_X(H)$, as it has no equally labeled edges sharing the same origin or terminus.

As in Claim 3.1, each of the following properties is either proven in (some of) [Sta83, KM02, MVW07] or an easy observation:

Claim 3.2. *Let $H, J, L \le \mathbf{F}_k$ be subgroups. Then*

(1) A morphism $\Gamma_X(H) \to \Gamma_X(J)$ exists if and only if $H \le J$.
(2) If a morphism $\Gamma_X(H) \to \Gamma_X(J)$ exists, it is unique. We denote it by $\eta^X_{H\to J}$.
(3) Whenever $H \le L \le J$, $\eta^X_{H\to J} = \eta^X_{L\to J} \circ \eta^X_{H\to L}$.[†]
(4) If $\eta^X_{H\to J}$ is injective, then $H \overset{*}{\le} J$.[‡]
(5) Every morphism is an immersion (locally injective at the vertices).

A special role is played by *surjective* morphisms of core graphs:

Definition 3.3. Let $H \le J \le \mathbf{F}_k$. Whenever $\eta^X_{H\to J}$ is surjective, we say that $\Gamma_X(H)$ *covers* $\Gamma_X(J)$ or that $\Gamma_X(J)$ *is a quotient of* $\Gamma_X(H)$. We indicate this by $\Gamma_X(H) \twoheadrightarrow \Gamma_X(J)$. As for the groups, we say that H *X-covers* J and denote this by $H \le_{\overset{\twoheadrightarrow}{X}} J$.

By "surjective" we mean surjective on both vertices and edges. Note that we use the term "covers" even though in general this is *not* a topological covering map (a morphism between core graphs is always locally injective at the vertices, but it need not be locally bijective). In Section 6 we do study topological covering maps, and we reserve the term "coverings" for these.

[†]Points (1)–(3) can be formulated by saying that (3.2) is in fact an isomorphism of categories, given by the functors π_1^X and Γ_X.

[‡]But not vice versa: for example, consider $\langle x_1x_2^2\rangle \overset{*}{\le} \mathbf{F}_2$.

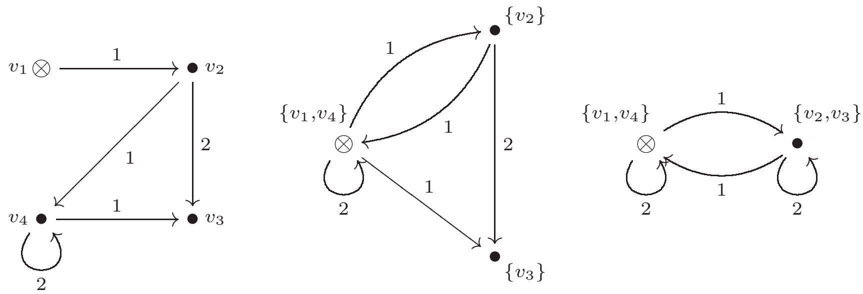


FIGURE 3.3. The left graph is the core graph $\Gamma_X(H)$ of $H = \langle x_1x_2x_1^{-3}, x_1^2x_2x_1^{-2}\rangle \le \mathbf{F}_2$. Its vertices are denoted by $v_1,\ldots,v_4$. The graph in the middle is the quotient ${}^{\Gamma_X(H)}/{}_P$ corresponding to the partition $P = \{\{v_1,v_4\},\{v_2\},\{v_3\}\}$. This is not a core graph as there are two 1-edges originating at $\{v_1,v_4\}$. In order to obtain a core quotient graph, we use the Stallings folding process (illustrated in Figure 3.2). The resulting core graph, $\Gamma_X\left(\pi_1^X\left({}^{\Gamma_X(H)}/{}_P\right)\right)$, is shown on the right and corresponds to the partition $\bar{P} = \{\{v_1,v_4\},\{v_2,v_3\}\}$.

For instance, $H=\langle x_1x_2x_1^{-3}, x_1^2x_2x_1^{-2}\rangle \le \mathbf{F}_k$ X-covers the group $J = \langle x_2, x_1^2, x_1x_2x_1\rangle$, the corresponding core graphs of which are the leftmost and rightmost graphs in Figure 3.3. As another example, a core graph Γ X-covers $\Gamma_X(\mathbf{F}_k)$ (which is merely a wedge of k loops) if and only if it contains edges of all k labels.

As implied by the notation, the relation $H \le_{\overrightarrow{X}} J$ indeed depends on the given basis X of $\mathbf{F}_k$. For example, if $H = \langle x_1x_2\rangle$ then $H \le_{\overrightarrow{X}} \mathbf{F}_2$. However, for $Y = \{x_1x_2, x_2\}$, H does not Y-cover $\mathbf{F}_2$, as $\Gamma_Y(H)$ consists of a single vertex and a single loop and has no quotients apart from itself.

It is easy to see that the relation "$\le_{\overrightarrow{X}}$" indeed constitutes a partial ordering of the set of subgroups of $\mathbf{F}_k$. We make a few other useful observations:

Claim 3.4. *Let $H, J, L \le \mathbf{F}_k$ be subgroups. Then*

(1) Whenever $H \le J$ there exists an intermediate subgroup M such that $H \le_{\overrightarrow{X}} M \overset{*}{\le} J$.

(2) If one adds the condition that $\Gamma_X(M)$ embeds in $\Gamma_X(J)$, then this M is unique.

(3) If $H \le_{\overrightarrow{X}} J$ and $H \le_{\overrightarrow{X}} L \le J$, then $L \le_{\overrightarrow{X}} J$.

(4) If H is finitely generated then it X-covers only a finite number of groups. In particular, the poset $(\mathfrak{sub}_{fg}(\mathbf{F}_k), \le_{\overrightarrow{X}})$ is locally finite.

Proof. Point (1) follows from the factorization of the morphism $\eta^X_{H\to J}$ to a surjection followed by an embedding. Indeed, it is easy to see that the image of $\eta^X_{H\to J}$ is a subgraph of $\Gamma_X(J)$ which is in itself a core graph. Namely, it contains no "hanging trees" (edges and vertices not traced by reduced paths around the basepoint). Let $M = \pi_1^X\left(\operatorname{im}\eta^X_{H\to J}\right)$ be the subgroup corresponding to this sub–core graph. (1) now follows from points (1) and (4) in Claim 3.2. Point (2) follows from the uniqueness of such factorization of a morphism. Point (3) follows from the fact that if $\eta^X_{H\to J} = \eta^X_{L\to J}\circ\eta^X_{H\to L}$ is surjective then so is $\eta^X_{L\to J}$. Point (4) follows from the fact that $\Gamma_X(H)$ is finite (Claim 3.1(1)) and thus has only finitely many quotients, and each quotient corresponds to a single group (by (3.2)). □

In [MVW07], the set of X-quotients of H

$$[H,\infty)_{\overrightarrow{X}} = \{J \mid H \leq_{\overrightarrow{X}} J\} \tag{3.3}$$

is called the *X-fringe* of H. Claim 3.4(4) states in this terminology that for every $H \leq_{fg} \mathbf{F}_k$ (and every basis X), $|[H,\infty)_{\overrightarrow{X}}| < \infty$. Note that $[H,\infty)_{\overrightarrow{X}}$ always contains the supremum of its elements, namely the group generated by the elements of X which label edges in $\Gamma_X(H)$ (which is $\pi_1^X\left(\operatorname{im}\eta^X_{H\to\mathbf{F}_k}\right)$). (We remark that in the special case of $H = \langle w\rangle$ for some $w \in \mathbf{F}_k$, the set $[\langle w\rangle,\infty)_{\overrightarrow{X}}$ appears also in [Tur96] and, in a very different language, in the aforementioned [Nic94].)

It is easy to see that quotients of $\Gamma_X(H)$ are determined by the partition they induce of the vertex set $V(\Gamma_X(H))$. However, not every partition P of $V(\Gamma_X(H))$ corresponds to a quotient core graph: in the resulting graph, which we denote by ${}^{\Gamma_X(H)}/_P$, two distinct j-edges may have the same origin or the same terminus. Then again, when a partition P of $V(\Gamma_X(H))$ yields a quotient which is not a core graph, we can perform Stallings foldings (as demonstrated in Figure 3.2) until we obtain a core graph. Since Stallings foldings do not affect π_1^X, the core graph we obtain in this manner is $\Gamma_X(J)$, where $J = \pi_1^X\left({}^{\Gamma_X(H)}/_P\right)$. The resulting partition $\bar{P}$ of $V(\Gamma_X(H))$ (as the fibers of $\eta^X_{H\to J}$) is the finest partition of $V(\Gamma_X(H))$ which gives a quotient core graph and which is still coarser than P. We illustrate this in Figure 3.3.

Thus, there is sense in examining the quotient of a core graph Γ "generated" by some partition P of its vertex set, namely, $\Gamma_X\left(\pi_1^X\left({}^{\Gamma}/_P\right)\right)$. The most interesting case is that of the "simplest" partitions: those which identify only a single pair of vertices. Before looking at these, we introduce a measure for the complexity of partitions: if $P \subseteq 2^{\mathcal{X}}$ is a partition of some set $\mathcal{X}$, let

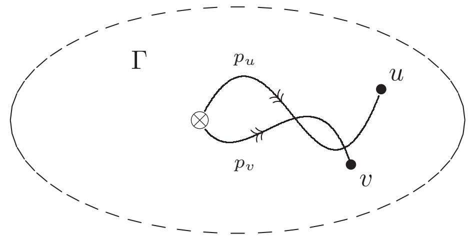
$$\|P\| \overset{def}{=} |\mathcal{X}| - |P| = \sum_{B\in P}(|B|-1). \tag{3.4}$$

Namely, $\|P\|$ is the number of elements in the set minus the number of blocks in the partition. For example, $\|P\| = 1$ iff P identifies only a single pair of elements. It is not hard to see that $\|P\|$ is also the minimal number of identifications one needs to make in $\mathcal{X}$ in order to obtain the equivalence relation P.

Definition 3.5. Let Γ be a core graph and let P be a partition of $V(\Gamma)$ with $\|P\| = 1$, i.e. having a single non-trivial block, of size two. Let Δ be the core graph generated from Γ by P. We then say that Δ is an *immediate quotient* of Γ.

Alternatively, we say that Δ is *generated by identifying a single pair* of vertices of Γ. For instance, the rightmost core graph in Figure 3.3 is an immediate quotient of the leftmost one.

The main reason that immediate quotients are interesting is their algebraic significance. Let $H, J \leq \mathbf{F}_k$ with $\Gamma = \Gamma_X(H), \Delta = \Gamma_X(J)$ their core graphs, and assume that Δ is an immediate quotient of Γ obtained by identifying the vertices $u, v \in V(\Gamma)$. Now let $w_u, w_v \in \mathbf{F}_k$ be the words corresponding to some paths p_u, p_v in Γ from the basepoint to u and v respectively (note that these paths are not unique). It is not hard to see that identifying u and v has the same effect as adding the word $w = w_u w_v^{-1}$ to H and considering the generated group. Namely, that $J = \langle H, w\rangle$.



The relation of immediate quotients gives the set of finite core graphs (with edges labeled by $1,\ldots,k$) the structure of a directed acyclic graph (DAG).[†] This DAG was first introduced in [Pud13], and is denoted by $\mathcal{D}_k$. The set of vertices of $\mathcal{D}_k$ consists of the aforementioned core graphs, and its directed edges connect every core graph to its immediate quotients. Every ordered basis $X=\{x_1,\ldots,x_k\}$ of $\mathbf{F}_k$ determines a one-to-one correspondence between the vertices of this graph and $\mathfrak{sub}_{fg}(\mathbf{F}_k)$.

In the case of finite core graphs, Δ is a quotient of Γ if and only if Δ is reachable from Γ in $\mathcal{D}_k$ (that is, there is a directed path from Γ to Δ). In other words, if $H\leq_{fg}\mathbf{F}_k$, then $H\leq_{\overset{\twoheadrightarrow}{X}} J$ iff $\Gamma_X(J)$ can be obtained from $\Gamma_X(H)$ by a finite sequence of immediate quotients. Thus, for any $H\leq_{fg}\mathbf{F}_k$, the subgraph of $\mathcal{D}_k$ induced by the descendants of $\Gamma_X(H)$ consists of all quotients of $\Gamma_X(H)$, i.e. of all (core graphs corresponding to) elements of $[H,\infty)_{\overset{\twoheadrightarrow}{X}}$. By Claim 3.4(4), this subgraph is finite. In Figure 3.4 we draw the subgraph of $\mathcal{D}_k$ consisting of all quotients of $\Gamma_X(H)$ when $H=\langle x_1x_2x_1^{-1}x_2^{-1}\rangle$. The edges of this subgraph (i.e. immediate quotients) are denoted by the dashed arrows in the figure.

It is now natural to define a distance function between a finite core graph and each of its quotients:

Definition 3.6. Let $H,J\leq_{fg}\mathbf{F}_k$ be subgroups such that $H\leq_{\overset{\twoheadrightarrow}{X}} J$, and let $\Gamma=\Gamma_X(H)$, $\Delta=\Gamma_X(J)$ be the corresponding core graphs. We define the X-*distance* between H and J, denoted $\rho_X(H,J)$ or $\rho(\Gamma,\Delta)$, to be the shortest length of a directed path from Γ to Δ in $\mathcal{D}_k$.

In other words, $\rho_X(H,J)$ is the length of the shortest series of immediate quotients that yields Δ from Γ. There is another useful equivalent definition for the X-distance. To see this, assume that Γ' is generated from Γ by the partition P of $V(\Gamma)$ and let $\eta:\Gamma\twoheadrightarrow\Gamma'$ be the morphism. For every $x,y\in V(\Gamma')$, let $x'\in\eta^{-1}(x),y'\in\eta^{-1}(y)$ be arbitrary vertices in the fibers, and let P' be the partition of $V(\Gamma)$ obtained from P by identifying x' and y'. It is easy to see that the core graph generated from Γ' by identifying x and y is the same as the one generated by P' from Γ. From these considerations we obtain that

$$\rho_X(H,J)=\min\left\{\|P\|\;\middle|\;\begin{array}{c}P\text{ is a partition of }V(\Gamma_X(H))\\ \text{such that }\pi_1^X(\Gamma_X(H)/P)=J\end{array}\right\}. \tag{3.5}$$

For example, if Δ is an immediate quotient of Γ, then $\rho_X(H,J)=\rho(\Gamma,\Delta)=1$. For $H=\langle x_1x_2x_1^{-1}x_2^{-1}\rangle$, $\Gamma_X(H)$ has four quotients at distance 1 and two at distance 2 (see Figure 3.4).

[†]that is, a directed graph with no directed cycles

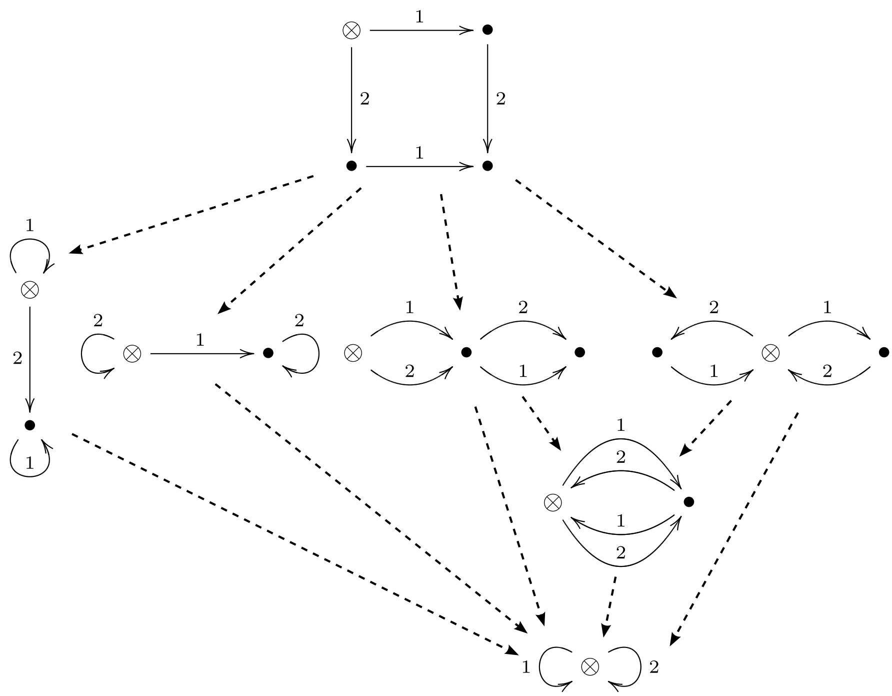


FIGURE 3.4. The subgraph of $\mathcal{D}_k$ induced by $[H,\infty)_{\vec{X}}$, that is, all quotients of the core graph $\Gamma = \Gamma_X(H)$, for $H = \langle x_1x_2x_1^{-1}x_2^{-1}\rangle$. The dashed arrows denote immediate quotients, i.e. quotients generated by merging a single pair of vertices. Γ has exactly seven quotients: itself, four immediate quotients, and two quotients at distance 2.

As mentioned earlier, by merging a single pair of vertices of $\Gamma_X(H)$ (and then folding) we obtain the core graph of a subgroup J obtained from H by adding some single generator (though not every element of $\mathbf{F}_k$ can be added in this manner). Thus, by taking an immediate quotient, the rank of the associated subgroup increases at most by 1 (in fact, it may also stay unchanged or even decrease). This implies that whenever $H \leq_{\vec{X}} J$, one has

$$rk(J) - rk(H) \;\leq\; \rho_X(H,J)\,. \tag{3.6}$$

In [Pud13] (Lemma 3.3), the distance is bounded from above as well:

Claim 3.7. *Let $H,J \leq_{fg} \mathbf{F}_k$ such that $H \leq_{\vec{X}} J$. Then*

$$rk(J) - rk(H) \;\leq\; \rho_X(H,J) \;\leq\; rk(J)\,.$$

We shall make use of the following theorem, which asserts that the lower bound is attained if and only if H is a free factor of J:

Theorem 3.8 ([Pud13, Theorem 1.1])**.** *Let $H,J \leq_{fg} \mathbf{F}_k$ and assume further that $H \leq_{\vec{X}} J$. Then $H \overset{*}{\leq} J$ if and only if*

$$\rho_X(H,J) = \mathrm{rk}(J) - \mathrm{rk}(H)\,.$$

In fact, the implication which is needed for our proof is trivial: as mentioned above, merging two vertices in $\Gamma_X(H)$ translates to adding some generator to H. If it is possible to obtain $\Gamma_X(J)$ from $\Gamma_X(H)$ by $rk(J) - rk(H)$ merging steps, this means we can obtain J from H by adding $rk(J)-rk(H)$ complementary generators to H, and hence $H \overset{*}{\leq} J$.† The other implication is not trivial and constitutes the essence of the proof of Theorem 1.1 in [Pud13]. The difficulty is that when $H \overset{*}{\leq}_{\overrightarrow{X}} J$, it is not *a priori* obvious why it is possible to find $rk(J) - rk(H)$ complementing generators of J from H, so that each of them can be realized by merging a pair of vertices in $\Gamma_X(H)$.

We finish this section with a classical fact about free factors that will be useful in the next section.

Claim 3.9. *Let H, J and K be subgroups of $\mathbf{F}_k$.*

(1) If $H \overset{*}{\leq} J$ and $K \leq J$, then $H \cap K \overset{*}{\leq} K$.
(2) If $H, K \overset{*}{\leq} J$, then $H \cap K \overset{*}{\leq} J$.
(3) If $H \overset{*}{\leq} J$, then H is a free factor of any intermediate group $H \leq M \leq J$.

Proof. Let Y be a basis of J extending a basis Y_0 of H. Then $\Gamma_Y(J)$ and $\Gamma_Y(H)$ are bouquets of $|Y|, |Y_0|$ loops, respectively. It is easy to check that $\Gamma_Y(H \cap K)$ is obtained from $\Gamma_Y(K)$ as follows: first, delete the edges labeled by $Y \setminus Y_0$; then, keep only the connected component of the basepoint; finally, trim all "hanging trees" (see the proof of Claim 3.4). Consequently, $\Gamma_Y(H \cap K)$ is embedded in $\Gamma_Y(K)$. Claim 3.2(4) then gives (1), and (2) and (3) follow immediately. □

In particular, the last claim shows that if $H \overset{*}{\leq} \mathbf{F}_k$, then $\pi(H) = \infty$ (see Definition 1.7). On the other hand, if H is not a free factor of $\mathbf{F}_k$, then obviously $\pi(H) \leq \mathrm{rk}(\mathbf{F}_k) = k$. Thus $\pi(H) \in \{0, 1, 2, \ldots, k\} \cup \{\infty\}$.

4. Algebraic extensions and critical subgroups

We now return to the sparsest partial order we consider in this paper, that of algebraic extensions. All claims in this section appear in [KM02, MVW07], unless specifically stated otherwise. We shall occasionally sketch some proofs in order to allow the reader to obtain better intuition and in order to exemplify the strength of core graphs.

Recall (Definition 2.1) that J is an algebraic extension of H, denoted $H \leq_{alg} J$, if $H \leq J$ and H is not contained in any proper free factor of J. For example, consider $H = \langle x_1 x_2 x_1^{-1} x_2^{-1} \rangle \leq \mathbf{F}_2$. A proper free factor of $\mathbf{F}_2$ has rank at most 1, and H is not contained in any subgroup of rank 1 other than itself (as $x_1 x_2 x_1^{-1} x_2^{-1}$ is not a proper power). Finally, H itself is not a free factor of $\mathbf{F}_2$ (as can be inferred from Theorem 3.8 and Figure 3.4). Thus, $H \leq_{alg} \mathbf{F}_2$. In fact, we shall see that in this case $[H, \infty)_{alg} = \{H, \mathbf{F}_2\}$.

We first show that "$\leq_{alg}$" is a partial order:

Claim 4.1. *The relation "$\leq_{alg}$" is transitive.*

Proof. Assume that $H \leq_{alg} M \leq_{alg} J$. Let $H \leq L \overset{*}{\leq} J$. By Claim 3.9(1), $L \cap M \overset{*}{\leq} M$. But $H \leq L \cap M$ and $H \leq_{alg} M$, so $L \cap M = M$, and thus $M \leq L$. So now $M \leq L \overset{*}{\leq} J$, and from $M \leq_{alg} J$ we obtain that $L = J$. □

†This relies on the well-known fact that a set of size k which generates $\mathbf{F}_k$ is a basis.

Next, we show that "$\leq_{alg}$" is dominated by "$\leq_{\overrightarrow{X}}$" for every basis X of $\mathbf{F}_k$. Namely, if $H \leq_{alg} J$, then $H \leq_{\overrightarrow{X}} J$. This shows, in particular, that the poset $(\mathfrak{sub}_{fg}(\mathbf{F}_k), \leq_{alg})$ is locally finite.

Claim 4.2. *If $H \leq_{alg} J$, then $H \leq_{\overrightarrow{X}} J$ for every basis X of $\mathbf{F}_k$.*

Proof. By Claim 3.4, there is an intermediate subgroup M such that $H \leq_{\overrightarrow{X}} M \overset{*}{\leq} J$, and from $H \leq_{alg} J$ it follows that $M = J$. □

Remark 4.3. It is natural to conjecture that the converse also holds, namely that if $H \leq_{\overrightarrow{X}} J$ for every basis X of $\mathbf{F}_k$, then $H \leq_{alg} J$. (In fact, this conjecture appears in [MVW07], Section 3.) This is, however, false: it turns out that for $H = \langle x_1^2 x_2^2 \rangle$ and $J = \langle x_1^2 x_2^2, x_1 x_2 \rangle$, $H \leq_{\overrightarrow{X}} J$ for every basis X of $\mathbf{F}_2$, but J is *not* an algebraic extension of H [PP12b]. However, there are bases of $\mathbf{F}_3$ with respect to which H does not cover J. Hence, it is still plausible that some weaker version of the conjecture holds, e.g. that $H \leq_{alg} J$ if and only if for every embedding of J in a free group F, and for every basis X of F, $H \leq_{\overrightarrow{X}} J$. It is also plausible that the original conjecture from [MVW07] holds for $\mathbf{F}_k$ with $k \geq 3$.
In a similar fashion, one can ask whether $H \leq J$ if and only if for some basis X of $\mathbf{F}_k$, $H \leq_{\overrightarrow{X}} J$.

Claim 4.2 completes the proof of the relations, mentioned in Section 2, between the different partial orders we consider in this paper: inclusion, the family $\leq_{\overrightarrow{X}}$, and algebraic extensions. Recall that H-critical subgroups are a special kind of algebraic extensions. Thus,

$$\mathrm{Crit}(H) \subseteq [H, \infty)_{alg} \subseteq [H, \infty)_{\overrightarrow{X}} \subseteq [H, \infty)_{\leq}.$$

Theorem 3.8 and Claim 4.2 give the following criterion for algebraic extensions:

Lemma 4.4. *Let $H \leq_{fg} \mathbf{F}_k$. The algebraic extensions of H are the elements of $[H, \infty)_{\overrightarrow{X}}$ which are* not *immediate quotients of any subgroup in $[H, \infty)_{\overrightarrow{X}}$ of smaller rank.*

Proof. Let $J \in [H, \infty)_{\overrightarrow{X}}$. If J is an immediate X-quotient of $L \in [H, \infty)_{\overrightarrow{X}}$ with $\mathrm{rk}(L) < \mathrm{rk}(J)$, then by Theorem 3.8 $H \leq L \overset{*}{\lneqq} J$, and hence J is not an algebraic extension of H. On the other hand, assume there exists some L such that $H \leq L \overset{*}{\lneqq} J$. By Claim 3.4(1), there exists M such that $H \leq_{\overrightarrow{X}} M \overset{*}{\leq} L \overset{*}{\lneqq} J$. By Claim 3.4(3), $M \overset{*}{\lneqq}_{\overrightarrow{X}} J$. From Theorem 3.8 it follows that there is a chain of immediate quotients $M = M_0 \leq M_1 \leq \cdots \leq M_r = J$ inside $[H, \infty)_{\overrightarrow{X}}$ with $\mathrm{rk}(M_{i+1}) = \mathrm{rk}(M_i) + 1$, and M_{r-1} is the group we have looked for. □

Since the subgraph of $\mathcal{D}_k$ induced by the vertices corresponding to $[H, \infty)_{\overrightarrow{X}}$, namely $\Gamma_X(H)$ and its descendants, is finite and can be effectively computed, Lemma 4.4 yields a straightforward algorithm to find all algebraic extensions of a given $H \leq_{fg} \mathbf{F}_k$ (this algorithm was first introduced in [Pud13]). This, in particular, allows one to find all H-critical subgroups, and thus to compute the primitivity rank $\pi(H)$: the subgroups constituting $\mathrm{Crit}(H)$ are those in $(H, \infty)_{alg}$ of minimal rank, which is $\pi(H)$. For instance, Figure 3.4 shows that for $H = \langle x_1 x_2 x_1^{-1} x_2^{-1} \rangle$ we have $H = \{H, \mathbf{F}_2\}$. Thus, $\mathrm{Crit}(H) = \{\mathbf{F}_2\}$ and $\pi(H) = 2$ (so $\widetilde{\pi}(H) = 1$).

We conclude this section with yet another elegant result from [KM02, MVW07] that will be used in the proof of Theorem 1.8. In the spirit of field extensions, it says that every extension of subgroups of $\mathbf{F}_k$ has a unique factorization to an algebraic extension followed by a free extension (compare this with Claim 3.4(1,2)):

Claim 4.5. *Let $H \leq J$ be free groups. Then there is a unique subgroup L of J such that $H \leq_{alg} L \overset{*}{\leq} J$. Moreover, L is the intersection of all intermediate free factors of J and the union of all intermediate algebraic extensions of H:*

$$L = \bigcap_{M\colon H \leq M \overset{*}{\leq} J} M = \bigcup_{M\colon H \leq_{alg} M \leq J} M. \tag{4.1}$$

In particular, the intersection of all free factors is a free factor, and the union of all algebraic extensions is an algebraic extension. Claim 4.5 is true in general, but we describe the proof only of the slightly simpler case of finitely generated subgroups. We need only this case in this paper.

Proof. By Claim 3.9 and rank considerations, the intersection in the middle of (4.1) is by itself a free factor of J. Denote it by L, so we have $H \leq L \overset{*}{\leq} J$. Clearly, L is an algebraic extension of H (otherwise it would contain a proper free factor). But we claim that L contains every other intermediate algebraic extension of H. Indeed, let $H \leq_{alg} M \leq J$. By Claim 3.9(1), $H \leq M \cap L \overset{*}{\leq} M$, so $M \cap L = M$, that is $M \leq L$. ☐

5. Möbius inversions

Let $(P, \leq)$ be a locally finite poset and let A be a commutative ring with unity. Then there exists an *incidence algebra*[†] of all functions from pairs $\{(x, y) \in P \times P \,|\, x \leq y\}$ to A. In addition to point-wise addition and scalar multiplication, it has an associative multiplication defined by convolution:

$$(f * g)(x, y) = \sum_{z \in [x,y]} f(x, z) g(z, y)$$

(where $x \leq y$ and $[x, y] = \{z \,|\, x \leq z \leq y\}$). The unit element is the diagonal

$$\delta(x, y) = \begin{cases} 1 & x = y \\ 0 & x \lneqq y \end{cases}.$$

Functions with invertible diagonal entries (i.e. $f(x, x) \in A^\times$ for all $x \in P$) are invertible w.r.t. this multiplication. Most famously, the constant ζ function, which is defined by $\zeta(x, y) = 1$ for all $x \leq y$, is invertible, and its inverse, μ, is called the *Möbius function* of P. This means that $\zeta * \mu = \mu * \zeta = \delta$; i.e., for every pair $x \leq y$

$$\sum_{z \in [x,y]} \mu(z, y) = (\zeta * \mu)(x, y) = \delta(x, y) = (\mu * \zeta)(x, y) = \sum_{z \in [x,y]} \mu(x, z).$$

Let f be some function in the incidence algebra. The function $f * \zeta$, which satisfies $(f * \zeta)(y) = \sum_{z \in [x,y]} f(z)$, is analogous to the right-accumulating function in calculus (for $g : \mathbb{R} \to \mathbb{R}$ this is the function $G(y) = \int_{z \in [x,y]} g(z)\, dz$). Thus, multiplying a function on the right by μ can be thought of as "right derivation." Similarly, one thinks of multiplying from the left by ζ and μ as left integration and left derivation, respectively.

Recall the function Φ (2.1), defined for every pair of free subgroups $H, J \leq_{fg} \mathbf{F}_k$ such that $H \leq J$: $\Phi_{H,J}(n)$ is the expected number of common fixed points

[†] The theory of incidence algebras of posets can be found in [Sta97].

of $\alpha_{J,n}(H)$, where $\alpha_{J,n} \in \mathrm{Hom}(J, S_n)$ is a random homomorphism chosen with uniform distribution. We think of Φ as a function from the set of such pairs (H, J) into the ring of functions $\mathbb{N} \to \mathbb{Q}$.

Let X be a basis of $\mathbf{F}_k$. We write Φ^X for the restriction of Φ to pairs (H, J) such that $H \le_{\overrightarrow{X}} J$. As "$\le_{\overrightarrow{X}}$" defines a locally finite partial ordering of $\mathfrak{sub}_{fg}(\mathbf{F}_k)$, there exists a matching Möbius function, $\mu^X = \left(\zeta^X\right)^{-1}$ (where $\zeta^X_{H,J} = 1$ for all $H \le_{\overrightarrow{X}} J$). Our proof of Theorem 1.8 consists of a detailed analysis of the left, right, and two-sided derivations of Φ^X:

$$\Phi^X$$

$$L^X \stackrel{def}{=} \mu^X * \Phi^X \qquad\qquad R^X \stackrel{def}{=} \Phi^X * \mu^X$$

$$C^X \stackrel{def}{=} \mu^X * \Phi^X * \mu^X$$

By definition, we have for every f.g. $H \le_{\overrightarrow{X}} J$:

$$\Phi_{H,J} = \sum_{M \in [H,J]_{\overrightarrow{X}}} L^X_{M,J} = \sum_{M,N\colon H \le_{\overrightarrow{X}} M \le_{\overrightarrow{X}} N \le_{\overrightarrow{X}} J} C^X_{M,N} = \sum_{N \in [H,J]_{\overrightarrow{X}}} R^X_{H,N}. \tag{5.1}$$

Note that (5.1) can serve as definitions for the three functions L^X, C^X, R^X: for instance, $L^X = \mu^X * \Phi^X$ is equivalent to $\zeta^X * L^X = \Phi^X$, which is the leftmost equality above.

We begin the analysis of these functions by the following striking observation regarding R^X. Recall (Claim 4.2) that if $H \le_{alg} J$, then $H \le_{\overrightarrow{X}} J$ for every basis X. It turns out that the function R^X is supported on algebraic extensions alone, and moreover, is independent of the basis X.

Proposition 5.1. *Let $H, J \le_{fg} \mathbf{F}_k$.*

(1) If $H \le_{\overrightarrow{X}} J$ but J is not *an algebraic extension of H, then $R^X_{H,J} = 0$.*
(2) $R^X_{H,J} = R^Y_{H,J}$ for every basis Y of $\mathbf{F}_k$, whenever both are defined.

Remark 5.2. The only property of Φ we use is that $\Phi_{H,L} = \Phi_{H,J}$ whenever $H \le L \overset{*}{\le} J$, which is easy to see from the definition of Φ. Therefore, the proposition holds for the right derivation of every function with this property. In particular, the proposition holds for every "statistical" function, in which the value of (H, J) depends solely on the image of H via a uniformly distributed random homomorphism from J to some group G.

Proof. We show both claims at once by induction on $|\,[H,J]_{\overrightarrow{X}}\,|$, the size of the closed interval between H and J. The induction basis is $H = J$. That $H \le_{alg} H$ is immediate. By (5.1), $R^X_{H,H} = \Phi_{H,H}$ and so $R^X_{H,H}$ is indeed independent of the basis X.

Assume now that $|\,[H,J]_{\overrightarrow{X}}\,| = r$ and that both claims are proven for every pair bounding an interval of size $< r$. By (5.1) and the first claim of the induction hypothesis,

$$R^X_{H,J} = \Phi_{H,J} - \sum_{N \in [H,J)_{\overrightarrow{X}}} R^X_{H,N} = \Phi_{H,J} - \sum_{N\colon H \le_{alg} N \lneq_{\overrightarrow{X}} J} R^X_{H,N}. \tag{5.2}$$

By Claim 3.4(3), $\{N \mid H \leq_{alg} N \lneqq_{\overrightarrow{X}} J\} = \{N \mid H \leq_{alg} N \lneqq J\}$, and the latter is independent of the basis X. Furthermore, by the induction hypothesis regarding the second claim, so are the terms $R^X_{H,N}$ in this summation. This settles the second point.

Finally, if J is *not* an algebraic extension of H, then let L be some intermediate free factor of J, $H \leq L \overset{*}{\lneqq} J$. As mentioned above, this yields that $\Phi_{H,J} = \Phi_{H,L}$. Therefore,

$$R^X_{H,J} = \Phi_{H,J} - \sum_{N \in [H,J)_{\overrightarrow{X}}} R^X_{H,N} = \underbrace{\Phi_{H,L} - \sum_{N \in [H,L]_{\overrightarrow{X}}} R^X_{H,N}}_{0 \text{ by definition}} - \sum_{N \in [H,J)_{\overrightarrow{X}} \setminus [H,L]_{\overrightarrow{X}}} R^X_{H,N}.$$

By Claim 4.5, all algebraic extensions of H inside the interval $[H,J]_{\overrightarrow{X}}$ are contained in L. Hence, every subgroup $N \in [H,J)_{\overrightarrow{X}} \setminus [H,L]_{\overrightarrow{X}}$ is not an algebraic extension of H, and by the induction hypothesis $R^X_{H,N}$ vanishes. The desired result follows. □

In view of Proposition 5.1 we can omit the superscript and write from now on $R_{H,J}$ instead of $R^X_{H,J}$. Moreover, we can write the following "basis independent" equation for every pair of f.g. subgroups $H \leq J$:

$$\Phi_{H,J} = \sum_{N\colon H \leq_{alg} N \leq J} R_{H,N}. \tag{5.3}$$

When $H \leq_{\overrightarrow{X}} J$ this follows from the proof above. For general $H \leq J$, there is some subgroup L such that $H \leq_{\overrightarrow{X}} L \overset{*}{\leq} J$ and every intermediate algebraic extension $H \leq_{alg} N \leq J$ is contained in L (see Claims 3.4 and 4.5). Therefore,

$$\Phi_{H,J} = \Phi_{H,L} = \sum_{N\colon H \leq_{alg} N \leq L} R_{H,N} = \sum_{N\colon H \leq_{alg} N \leq J} R_{H,N}.$$

It turns out that unlike the function R, the other two derivations of Φ, namely L^X and C^X, do depend on the basis X. However, the latter two functions have combinatorial interpretations. In the next section we show that $\Phi_{H,J}$ and $L^X_{H,J}$ can be described in terms of random coverings of the core graph $\Gamma_X(J)$, and that explicit rational expressions in n can be computed to express these two functions for given H, J (Lemmas 6.2 and 6.3 below). This, in turn, allows us to analyze the combinatorial meaning and order of magnitude of $C^X_{M,N}$ (Proposition 7.1).

Finally, using the fact that R is the "left integral" of C^X, that is, $R = \zeta^X * C^X$, we finish the circle around the diagram of Φ's derivations, and use this analysis of Φ, L^X and C^X to prove that for every pair $H \leq_{alg} J$, $R_{H,J}$ does not vanish and is, in fact, positive for large enough n. This alone gives Theorem 1.4. The more informative 1.8 follows from an analysis of the order of magnitude of $R_{H,J}$ in this case (Proposition 7.2).

6. Random coverings of core graphs

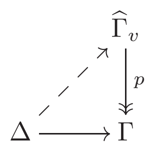
This section studies the graphs which cover a given core graph in the topological sense, i.e. $\widehat{\Gamma} \overset{p}{\twoheadrightarrow} \Gamma$ with p locally bijective. We call these graphs (together with their projection maps) *coverings* of Γ. The reader should not confuse this with our notion "covers" from Definition 3.3.

We focus on directed and edge-labeled coverings. This means we only consider $\widehat{\Gamma} \overset{p}{\twoheadrightarrow} \Gamma$ such that $\widehat{\Gamma}$ is directed and edge labeled, and the projection p preserves orientations and labels. When Γ is a core graph we do *not* assume that $\widehat{\Gamma}$ is a core graph as well. It may be disconnected, and it need not be pointed. Nevertheless, it is not hard to see that when Γ and $\widehat{\Gamma}$ are finite, for every vertex v in $p^{-1}(\otimes)$, the fiber over Γ's basepoint, we do have a valid core graph, which we denote by $\widehat{\Gamma}_v$: this is the connected component of v in $\widehat{\Gamma}$, with v serving as basepoint. Moreover, the restriction of the projection map p to $\widehat{\Gamma}_v$ is a core-graph morphism.

The theory of core-graph coverings shares many similarities with the theory of topological covering spaces. The following claim lists some standard properties of covering spaces, formulated for core graphs.

Claim 6.1. *Let Γ be a core graph, $\widehat{\Gamma} \overset{p}{\twoheadrightarrow} \Gamma$ a covering and v a vertex in the fiber $p^{-1}(\otimes)$.*

(1) The group $\pi_1^X(\Gamma)$ acts on the fiber $p^{-1}(\otimes)$, and these actions give a correspondence between coverings of Γ and $\pi_1^X(\Gamma)$-sets.
(2) In this correspondence, coverings of Γ with fiber $\{1,\ldots,n\}$ correspond to actions of $\pi_1^X(\Gamma)$ on $\{1,\ldots,n\}$, i.e., to group homomorphisms $\pi_1^X(\Gamma) \to S_n$.
(3) The group $\pi_1^X(\widehat{\Gamma}_v)$ is the stabilizer of v in the action of $\pi_1^X(\Gamma)$ on $p^{-1}(\otimes)$ (note that $\pi_1^X(\widehat{\Gamma}_v)$ and $\pi_1^X(\Gamma)$ are both subgroups of $\mathbf{F}(X)$).
(4) A core-graph morphism $\Delta \to \Gamma$ can be lifted to a core-graph morphism $\Delta \to \widehat{\Gamma}_v$ (i.e., the diagram



can be completed) if and only if $\pi_1^X(\Delta) \subseteq \pi_1^X(\widehat{\Gamma}_v)$. By the previous point, this is equivalent to saying that all elements of $\pi_1^X(\Delta)$ fix v.

We now turn our attention to random coverings. The vertex set of an n-sheeted covering of a graph $\Gamma = (V,E)$ can be assumed to be $V \times \{1,\ldots,n\}$, so that the fiber above $v \in V$ is $\{v\} \times \{1,\ldots,n\}$. For every edge $e = (u,v) \in E$, the fiber over e then constitutes a perfect matching between $\{v\} \times \{1,\ldots,n\}$ and $\{u\} \times \{1,\ldots,n\}$. This suggests a natural model for random n-coverings of the graph Γ. Namely, for every $e \in E$ choose uniformly a random perfect matching (which is just a permutation in S_n). This model was introduced in [AL02], and is a generalization of a well-known model for random regular graphs (see e.g. [BS87]).† Note that the model works equally well for graphs with loops and with multiple edges.

In fact, there is some redundancy in this model, if we are interested only in isomorphism classes of coverings (two coverings are isomorphic if there is an isomorphism between them that commutes with the projection maps). It is possible to obtain the same distribution on (isomorphism classes of) n-coverings of Γ with fewer random permutations: one may choose some spanning tree T of Γ, associate the identity permutation with every edge in T, and pick random permutations only for edges outside T.

†Occasionally these random coverings are referred to as random *lifts* of graphs. We shall reserve this term for its usual meaning.

We now fix some $J \leq_{fg} \mathbf{F}_k$, and consider random coverings of its core graph, $\Gamma_X(J)$. We denote by $\widehat{\Gamma}_X(J)$ a random n-covering of $\Gamma_X(J)$, according to one of the models described above. If $p : \widehat{\Gamma}_X(J) \to \Gamma_X(J)$ is the covering map, then $\widehat{\Gamma}_X(J)$ inherits the edge orientation and labeling from $\Gamma_X(J)$ via p^{-1}. For every i $(1 \leq i \leq n)$, we write $\widehat{\Gamma}_X(J)_i$ for the core graph $\widehat{\Gamma}_X(J)_{(\otimes,i)}$ (the component of $(\otimes, i)$ in $\widehat{\Gamma}_X(J)$ with basepoint $(\otimes, i)$).

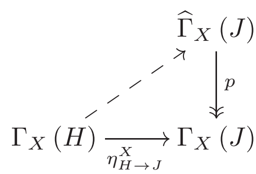
By Claim 6.1(2), each random n-covering of $\Gamma_X(J)$ encodes a homomorphism $\alpha_{J,n} \in \mathrm{Hom}(J, S_n)$, via the action of $J = \pi_1^X(\Gamma_X(J))$ on the basepoint fiber. Explicitly, an element $w \in J$ is mapped to a permutation $\alpha_{J,n}(w) \in S_n$ as follows: w corresponds to a closed path p_w around the basepoint of $\Gamma_X(J)$. For every $1 \leq i \leq n$, the lift of p_w that starts at $(\otimes, i)$ ends at $(\otimes, j)$ for some j, and $\alpha_{J,n}(w)(i) = j$.

By the correspondence of actions of J on $\{1, \ldots, n\}$ and n-coverings of $\Gamma_X(J)$, $\alpha_{J,n}$ is a uniform random homomorphism in $\mathrm{Hom}(J, S_n)$. This can also be verified using the "economical" model, as follows: choose some basis $Y = \{y_1, \ldots, y_{\mathrm{rk}(J)}\}$ for J via a choice of a spanning tree T of $\Gamma_X(J)$ and of orientation of the remaining edges, and choose uniformly at random some $\sigma_r \in S_n$ for every basis element y_r. Clearly, $\alpha_{J,n}(y_r) = \sigma_r$.

We can now use the coverings of $\Gamma_X(J)$ to obtain a geometric interpretation of $\Phi_{H,J}$, as follows: let $H \leq J \leq \mathbf{F}_k$ and $1 \leq i \leq n$. By 6.1(4), the morphism $\eta^X_{H\to J} : \Gamma_X(H) \to \Gamma_X(J)$ lifts to a core-graph morphism $\Gamma_X(H) \to \widehat{\Gamma}_X(J)_i$ iff $H = \pi_1^X(\Gamma_X(H))$ fixes $(\otimes, i)$ via the action of J on the fiber $\otimes \times \{1, \ldots, n\}$. Since this action is given by $\alpha_{J,n}$, this means that $\eta^X_{H\to J}$ lifts to $\widehat{\Gamma}_X(J)_i$ exactly when $\alpha_{J,n}(H)$ fixes i. Recalling that $\Phi_{H,J}(n)$ is the expected number of elements in $\{1, \ldots, n\}$ fixed by $\alpha_{J,n}(H)$, we obtain an alternative definition for it:

Lemma 6.2. *Let $\widehat{\Gamma}_X(J)$ be a random n-covering space of $\Gamma_X(J)$ in the aforementioned model from* [AL02]*. Then,*

$$\Phi_{H,J}(n) = \textit{The expected number of lifts of } \eta^X_{H\to J} \textit{ to } \widehat{\Gamma}_X(J).$$



Note that this characterization of $\Phi_{H,J}$ involves the basis X, although the original definition (2.1) does not. One of the corollaries of this lemma is therefore that the average number of lifts does *not* depend on the basis X.

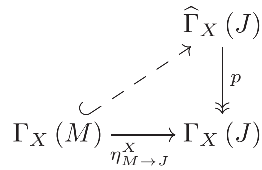
Recall (Section 5) the definition of the function L^X, which satisfies $\Phi_{H,J} = \sum_{M\in[H,J]_{\vec{X}}} L^X_{M,J}$ for every $H \leq_{\vec{X}} J$. It turns out that this derivation of Φ also has a geometrical interpretation. Assume that $\eta^X_{H\to J}$ does lift to $\widehat{\eta}_i : \Gamma_X(H) \to \widehat{\Gamma}_X(J)_i$. By Claim 3.4, $\widehat{\eta}_i$ decomposes as a quotient onto $\Gamma_X(M)$, where $M = \pi_1^X(\mathrm{im}\,\widehat{\eta}_i)$, followed by an embedding. Moreover, M lies in $[H,J]_{\vec{X}}$. On the other hand, if there is some $M \in [H,J]_{\vec{X}}$ such that $\Gamma_X(M)$ is embedded in $\widehat{\Gamma}_X(J)_i$, then such M is unique and $\widehat{\eta}_i$ lifts to the composition of $\eta^X_{H\to M}$ with this embedding. Consequently,

$$\begin{aligned}\Phi_{H,J}(n) &= \text{Expected number of lifts of } \eta^X_{H\to J} \text{ to } \widehat{\Gamma}_X(J)\\ &= \sum_{M\in[H,J]_{\overrightarrow{X}}} \text{Expected number of \emph{injective} lifts of } \eta^X_{M\to J} \text{ to } \widehat{\Gamma}_X(J).\end{aligned}$$

Taking the left derivations, we obtain:

Lemma 6.3. *Let $M \le_{\overrightarrow{X}} J$, and let $\widehat{\Gamma}_X(J)$ be a random n-covering space of $\Gamma_X(J)$ in the aforementioned model from* [AL02]. *Then,*

$$L^X_{M,J}(n) = \textit{The expected number of } \text{injective} \textit{ lifts of } \eta^X_{M\to J} \textit{ to } \widehat{\Gamma}_X(J).$$



Unlike the number of lifts in general, the number of injective lifts does depend on the basis X. For instance, consider $M = \langle x_1x_2\rangle$ and $J = \langle x_1, x_2\rangle = \mathbf{F}_2$. With the basis $X = \{x_1, x_2\}$, the probability that $\eta^X_{M\to J}$ lifts injectively to $\widehat{\Gamma}_X(J)_i$ equals $\frac{n-1}{n^2}$ (Lemma 6.4 shows how to compute this). However, with the basis $Y = \{x_1x_2, x_2\}$, the corresponding probability is $\frac{1}{n}$. We also remark that Lemma 6.3 allows a natural extension of L^X to pairs M, J such that M does not X-cover J.

Lemma 6.3 allows us to generalize the method used in [Nic94, LP10, Pud13] to compute the expected number of fixed points in $\alpha_n(w)$ (see the notations before Theorem 1.4′). We claim that for n large enough, $L^X_{M,J}(n)$ is a simple rational expression in n.

Lemma 6.4. *Let $M, J \le_{fg} \mathbf{F}_k$ such that $M \le_{\overrightarrow{X}} J$, and let $\eta = \eta^X_{M\to J}$ be the core-graph morphism. For large enough n,*

$$L^X_{M,J}(n) = \frac{\prod_{v\in V(\Gamma_X(J))} (n)_{|\eta^{-1}(v)|}}{\prod_{e\in E(\Gamma_X(J))} (n)_{|\eta^{-1}(e)|}}, \tag{6.1}$$

where $(n)_r$ is the falling factorial $n(n-1)\dots(n-r+1)$, and "large enough n" is $n \ge \max_{e\in E(\Gamma_X(J))} |\eta^{-1}(e)|$ (so that the denominator does not vanish).

Proof. Let v be a vertex in $\Gamma_X(J)$ and consider the fiber $\eta^{-1}(v)$ in $\Gamma_X(M)$. For every injective lift $\widehat{\eta} : \Gamma_X(M) \hookrightarrow \widehat{\Gamma}_X(J)$, the fiber $\eta^{-1}(v)$ is mapped injectively into the fiber $p^{-1}(v)$. The number of such injections is

$$(n)_{|\eta^{-1}(v)|} = n(n-1)\dots(n-|\eta^{-1}(v)|+1),$$

and therefore the number of injective lifts of $\eta|_{V(\Gamma_X(M))}$ into $V(\widehat{\Gamma}_X(J))$ is the numerator of (6.1).

We claim that any such injective lift has a positive probability of extending to a full lift of η: all one needs is that the fiber above every edge of $\Gamma_X(J)$ satisfy some constraints. To get the exact probability, we return to the more "wasteful" version of the model for a random n-covering of $\Gamma_X(J)$, the model in which we choose a random permutation for every edge of the base graph. Let $\widehat{\eta} : V(\Gamma_X(M)) \hookrightarrow V(\widehat{\Gamma}_X(J))$ be an injective lift of the vertices of $\Gamma_X(M)$ as above, and let e be some edge of $\Gamma_X(J)$. If $\widehat{\eta}$ is to be extended to $\eta^{-1}(e)$, the fiber above e in $\widehat{\Gamma}_X(J)$ must contain, for every $(u,v) \in \eta^{-1}(e)$, the edge $(\widehat{\eta}(u), \widehat{\eta}(v))$.

Thus, the random permutation $\sigma \in S_n$ which determines the perfect matching above e in $\widehat{\Gamma}_X(J)$ must satisfy $|\eta^{-1}(e)|$ non-colliding constraints of the form $\sigma(i) = j$. Whenever $n \geq |\eta^{-1}(e)|$ (which we assume), a uniformly random permutation in S_n satisfies such constraints with probability

$$\frac{1}{(n)_{|\eta^{-1}(e)|}}.$$

This shows the validity of (6.1). □

This immediately gives a formula for $\Phi_{H,J}$ as a rational function:

Corollary 6.5. *Let $H, J \leq_{fg} \mathbf{F}_k$ such that $H \leq_{\overrightarrow{X}} J$. Then, for large enough n,*

$$\Phi_{H,J}(n) = \sum_{M \in [H,J]_{\overrightarrow{X}}} L^X_{M,J}(n) = \sum_{M \in [H,J]_{\overrightarrow{X}}} \frac{\prod_{v \in V(\Gamma_X(J))} (n)_{\left|\left(\eta^X_{M\to J}\right)^{-1}(v)\right|}}{\prod_{e \in E(\Gamma_X(J))} (n)_{\left|\left(\eta^X_{M\to J}\right)^{-1}(e)\right|}}.$$

Since H X-covers every intermediate $M \in [H,J]_{\overrightarrow{X}}$, the largest fiber above every edge of $\Gamma_X(J)$ is obtained in $\Gamma_X(H)$ itself. Thus, "large enough n" in this corollary can be replaced by $n \geq \max_{e \in E(\Gamma_X(J))} \left| \left(\eta^X_{H\to J}\right)^{-1}(e) \right|$.

In fact, Corollary 6.5 applies, with slight modifications, to every pair of f.g. subgroups $H \leq J$: Lemma 6.2 holds in this more general case; that is, $\Phi_{H,J}$ is equal to the expected number of lifts of $\Gamma_X(H)$ to the random n-covering $\widehat{\Gamma}_X(J)$. The image of each lift (with the image of $\otimes$ as basepoint) is a core graph which is a quotient of $\Gamma_X(H)$, and so corresponds to a subgroup M such that $H \leq_{\overrightarrow{X}} M \leq J$. In explaining the rational expression in Lemma 6.4 we did not need M to cover J. Thus, for every $H \leq J$, both finitely generated,

$$\Phi_{H,J}(n) = \sum_{M \,:\, H \leq_{\overrightarrow{X}} M \leq J} \frac{\prod_{v \in V(\Gamma_X(J))} (n)_{\left|\left(\eta^X_{M\to J}\right)^{-1}(v)\right|}}{\prod_{e \in E(\Gamma_X(J))} (n)_{\left|\left(\eta^X_{M\to J}\right)^{-1}(e)\right|}}. \tag{6.2}$$

Corollary 6.5 yields in particular a straightforward algorithm to obtain a rational expression in n for $\Phi_{H,J}(n)$ (valid for large enough n). For example, consider $H = \langle x_1 x_2 x_1^{-1} x_2^{-1} \rangle$ and $\mathbf{F}_2 = \langle x_1, x_2 \rangle$. The interval $[H, \mathbf{F}_2]_{\overrightarrow{X}}$ consists of seven subgroups, as depicted in Figure 3.4. Following the computation in Corollary 6.5, we get that for $n \geq 2$ (we scan the quotients in Figure 3.4 top to bottom and in each row left to right),

$$\begin{aligned}\Phi_{H,\mathbf{F}_2}(n) &= \frac{(n)_4}{(n)_2 (n)_2} + \frac{(n)_2}{(n)_2 (n)_1} + \frac{(n)_2}{(n)_1 (n)_2} \\ &\quad + \frac{(n)_3}{(n)_2 (n)_2} + \frac{(n)_3}{(n)_2 (n)_2} + \frac{(n)_2}{(n)_2 (n)_2} + \frac{(n)_1}{(n)_1 (n)_1} \\ &= \frac{n}{n-1} = 1 + \frac{1}{n} + O\left(\frac{1}{n^2}\right).\end{aligned}$$

This demonstrates Theorem 1.8 and Table 1 for $H = \langle x_1x_2x_1^{-1}x_2^{-1}\rangle$ (recall the discussion following Lemma 4.4, where it is shown that $\pi(H) = 2$ and that $\mathrm{Crit}(H) = \{\mathbf{F}_2\}$).

The explicit computation of Φ yields an effective version of Theorem 1.4′:

Corollary 6.6. *Let $H \leq_{fg} \mathbf{F}_k$, and let ℓ denote the number of edges in $\Gamma_X(H)$. Then $H \overset{*}{\leq} \mathbf{F}_k$ iff $\Phi_{H,\mathbf{F}_k}(n) = n^{-\widetilde{\mathrm{rk}}\,H}$ for $n \leq \ell + \widetilde{\mathrm{rk}}\,H$. In particular, Proposition* 1.6 *follows.*

Proof. Assume that $\Phi_{H,\mathbf{F}_k}(n) = n^{-\widetilde{\mathrm{rk}}\,H}$ holds for $n \leq \ell + \widetilde{\mathrm{rk}}\,H$, and denote

$$\Phi'(n) = \sum_{M \in [H,\infty)_{\overset{\twoheadrightarrow}{X}}} \frac{(n)_{|V(\Gamma_X(M))|}}{\prod_{j=1}^{k}(n)_{|E_j(\Gamma_X(M))|}}, \tag{6.3}$$

where $E_j(\Gamma)$ are the j-edges in Γ. By Corollary 6.5, $\Phi'(n) = \Phi_{H,\mathbf{F}_k}(n)$ for $n \geq n_0 = \max_{j=1\ldots k}|E_j(\Gamma_X(H))|$, and in particular $\Phi'(n) = n^{-\widetilde{\mathrm{rk}}\,H}$ for $n_0 \leq n \leq \ell + \widetilde{\mathrm{rk}}\,H$. We proceed to show that $\Phi'(n) \equiv n^{-\widetilde{\mathrm{rk}}\,H}$, which implies $\Phi_{H,\mathbf{F}_k}(n) = n^{-\widetilde{\mathrm{rk}}\,H}$ for $n \geq n_0$. The conclusion then follows by Theorem 1.4′ (which is proved in the next section).

The number of j-edges in every quotient of $\Gamma_X(H)$ is at most $E_j(\Gamma_X(H))$, so that $\Phi'(n)\,g(n)$ is a polynomial for $g(n) = \prod_{j=1}^{k}(n)_{|E_j(\Gamma_X(H))|}$. We would like to establish

$$\Phi'(n)\,g(n)\,n^{\widetilde{\mathrm{rk}}(H)} \equiv g(n), \tag{6.4}$$

and we note that $\deg g = \ell$, and $\deg \Phi' \leq \max_{M\in[H,\infty)_{\overset{\twoheadrightarrow}{X}}} -\widetilde{\mathrm{rk}}(M) \leq 0$ follows from Claim 3.1(2) (assuming $H \neq id$). Therefore, the degrees of both sides of (6.4) are at most $\ell + \widetilde{\mathrm{rk}}\,H$, and it suffices to show they agree at $\ell + \widetilde{\mathrm{rk}}\,H + 1 = \ell + \mathrm{rk}\,H$ points. We already know that they agree for $n_0 \leq n \leq \ell + \widetilde{\mathrm{rk}}\,H$. For $0 \leq n < n_0$ it is clear that $g(n) = 0$. It turns out that the l.h.s. vanishes as well for these values of n. Expanding the l.h.s. gives

$$n^{\widetilde{\mathrm{rk}}\,H} \cdot \sum_{M\in[H,\infty)_{\overset{\twoheadrightarrow}{X}}} (n)_{|V(\Gamma_X(M))|} \prod_{j=1}^{k} (n - |E_j(\Gamma_X(M))|)_{|E_j(\Gamma_X(H))|-|E_j(\Gamma_X(M))|}, \tag{6.5}$$

and each term in the sum vanishes for $0 \leq n < n_0$: choose $1 \leq j \leq k$ for which $|E_j(\Gamma_X(H))| = n_0$. For each $M \in [H,\infty)_{\overset{\twoheadrightarrow}{X}}$ either $|E_j(\Gamma_X(M))| \leq n$, in which case $(n - |E_j(\Gamma_X(M))|)_{n_0 - |E_j(\Gamma_X(M))|} = 0$, or $|E_j(\Gamma_X(M))| > n$; as different j-edges must have different origins, the latter implies that $|V(\Gamma_X(M))| > n$, and hence $(n)_{|V(\Gamma_X(M))|}$ vanishes. □

Remark 6.7. The discussion in this section suggests a generalization of our analysis to finite groups G other than S_n. For any (finite) faithful G-set S, one can consider a random $|S|$-covering of $\Gamma_X(J)$. The fiber above every edge is chosen according to the action on S of a (uniformly distributed) random element of G. In this more general setting we also get a one-to-one correspondence between $\mathrm{Hom}(\mathbf{F}_k, G)$ and $|S|$-coverings. Although the computation of L^X and of Φ might be more involved, this suggests a way of analyzing words which are measure preserving w.r.t. G.

7. The proof of Theorem 1.8

The last major ingredient of the proof of our main result, Theorem 1.8, is an analysis of C^X, the double-sided derivation of Φ. Recall Definition 3.6 where the X-distance $\rho_X(H,J)$ was defined for every $H,J \leq_{fg} \mathbf{F}_k$ with $H \leq_{\overrightarrow{X}} J$.

Proposition 7.1. *Let $M,N \leq_{fg} \mathbf{F}_k$ satisfy $M \leq_{\overrightarrow{X}} N$. Then*

$$C^X_{M,N}(n) = O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(M)+\rho_X(M,N)}}\right).$$

Section 7.1 is dedicated to the proof of this proposition. Before getting there, we show how it practically finishes the proof of our main result. We do this with the following final step:

Proposition 7.2. *Let $H,N \leq_{fg} \mathbf{F}_k$ satisfy $H \leq_{alg} N$. Then*

$$R_{H,N}(n) = \frac{1}{n^{\widetilde{\mathrm{rk}}(N)}} + O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(N)+1}}\right).$$

Proof. Let X be some basis of $\mathbf{F}_k$. Recall that $R = \zeta^X * C^X$, i.e.

$$R_{H,N}(n) = \sum_{M\in[H,N]_{\overrightarrow{X}}} C^X_{M,N}(n).$$

For $M=N$ we have $C^X_{N,N}(n) = R_{N,N}(n) = \Phi_{N,N}(n) = n^{-\widetilde{\mathrm{rk}}(N)}$ (the last equality follows from the fact that m independent uniform permutations fix a point with probability n^{-m}). For any other M, i.e. $M \in [H,N)_{\overrightarrow{X}}$, the fact that N is an algebraic extension of H means that M is *not* a free factor of N and therefore, by Theorem 3.8 (and (3.6)), $\rho_X(M,N) \geq \widetilde{\mathrm{rk}}(N) - \widetilde{\mathrm{rk}}(M) + 1$. Proposition 7.1 then shows that

$$C^X_{M,N}(n) \in O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(M)+\rho_X(M,N)}}\right) \subseteq O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(N)+1}}\right).$$

Hence,

$$R_{H,N}(n) = C^X_{N,N}(n) + \sum_{M\in[H,N)_{\overrightarrow{X}}} C^X_{M,N}(n) = \frac{1}{n^{\widetilde{\mathrm{rk}}(N)}} + O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(N)+1}}\right).$$

□

The proof of Theorem 1.8 is now at hand. For every $H,J \leq_{fg} \mathbf{F}_k$ with $H \leq J$, by (5.3) and Proposition 7.2,

$$\begin{aligned}\Phi_{H,J}(n) &= \sum_{N\,:\,H\leq_{alg}N\leq J} R_{H,N}(n)\\ &= R_{H,H}(n) + \sum_{N\,:\,H\lneq_{alg}N\leq J} R_{H,N}(n)\\ &= \frac{1}{n^{\widetilde{\mathrm{rk}}(H)}} + \sum_{N\,:\,H\lneq_{alg}N\leq J} \frac{1}{n^{\widetilde{\mathrm{rk}}(N)}} + O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(N)+1}}\right).\end{aligned}$$

For $J = \mathbf{F}_k$ we can be more concrete. Recall that the H-critical groups, $\mathrm{Crit}(H)$, are the algebraic extensions of H of minimal rank, and this minimal rank is $\pi(H)$. Therefore,

$$\begin{aligned}\Phi_{H,\mathbf{F}_k}(n) &= \frac{1}{n^{\widetilde{\mathrm{rk}}(H)}} + \sum_{N \in (H,\infty)_{alg}} \frac{1}{n^{\widetilde{\mathrm{rk}}(N)}} + O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(N)+1}}\right) \\ &= \frac{1}{n^{\widetilde{\mathrm{rk}}(H)}} + \frac{|\mathrm{Crit}(H)|}{n^{\widetilde{\pi}(H)}} + O\left(\frac{1}{n^{\widetilde{\pi}(H)+1}}\right).\end{aligned}$$

This establishes our main results: Theorem 1.8, Theorem 1.4 and all their corollaries.

7.1. **The analysis of $C^X_{M,N}$.** In this subsection we look into C^X, the double-sided derivation of Φ, and establish Proposition 7.1, which bounds the order of magnitude of $C^X_{M,N}$. Recall that by definition $C^X = L^X * \mu^X$, which is equivalent to

$$L^X_{M,J} = \sum_{N \in [M,J]_{\overrightarrow{X}}} C^X_{M,N} \qquad (\forall M \leq_{\overrightarrow{X}} J). \tag{7.1}$$

We derive a combinatorial meaning of $C^X_{M,N}$ from this relation. To obtain this, we further analyze the rational expression (6.1) for $L^X_{M,J}$ and write it as a formal power series. Then, using a combinatorial interpretation of the terms in this series, we attribute each term to some $N \in [M,J]_{\overrightarrow{X}}$, and show that for every $N \in [M,J]_{\overrightarrow{X}}$, the sum of terms attributed to N is nothing but $C^X_{M,N}$. Finally, we use this combinatorial interpretation of $C^X_{M,N}$ to estimate its order of magnitude.

Rewriting $L^X_{M,J}$ as a power series in n^{-1}. Consider the numerator and denominator of (6.1): these are products of expressions of the type $(n)_r$. It is a classical fact that

$$(n)_r = \sum_{j=1}^{r} (-1)^{r-j} \begin{bmatrix} r \\ j \end{bmatrix} n^j$$

where $\begin{bmatrix} r \\ j \end{bmatrix}$ is the *unsigned Stirling number of the first kind.* That is, $\begin{bmatrix} r \\ j \end{bmatrix}$ is the number of permutations in S_r with exactly j cycles (see, for instance, [vLW01], Chapter 13).

We introduce the notation $[r]_j \overset{def}{=} \begin{bmatrix} r \\ r-j \end{bmatrix}$, which is better suited for our purposes. The cycles of a permutation $\sigma \in S_r$ constitute a partition P_σ of $\{1, \ldots, r\}$. We define $\|\sigma\| = \|P_\sigma\|$ (recall (3.4)), and it is immediate that $[r]_j$ is the number of permutations $\sigma \in S_r$ with $\|\sigma\| = j$. It is also easy to see that $\|\sigma\|$ is the minimal number of transpositions needed to be multiplied in order to obtain σ. Therefore, $[r]_j$ is the number of permutations in S_r which can be expressed as a product of j transpositions, but no less. In terms of this notation, we obtain

$$(n)_r = n^r \sum_{j=0}^{r-1} (-1)^j [r]_j n^{-j}.$$

The product of several expressions of this form, namely $(n)_{r_1}(n)_{r_2}\ldots(n)_{r_\ell}$, can be written as a polynomial in n whose coefficients have a similar combinatorial meaning, as follows. Let X be a set, and $\varphi : X \to \{1, \ldots, \ell\}$ some function with fibers of sizes $\left|\varphi^{-1}(i)\right| = r_i$ $(1 \le i \le \ell)$. We denote by

$$\mathrm{Sym}_{\varphi}(X) = \{\sigma \in \mathrm{Sym}(X) \,|\, \varphi \circ \sigma = \varphi\}$$

the set of permutations $\sigma \in \mathrm{Sym}(X)$ subordinate to the partition of X induced by the fibers of φ, i.e., such that $\varphi(\sigma(x)) = \varphi(x)$ for all $x \in X$. We define

$$[X]_j^{\varphi} = |\{\sigma \in \mathrm{Sym}_{\varphi}(X) \,:\, \|\sigma\| = j\}|,$$

the number of φ-subordinate permutations with $\|\sigma\| = j$. Put differently, $[X]_j^{\varphi}$ counts the permutations counted in $\left[|X|\right]_j$ which satisfy, in addition, that every cycle consists of a subset of some fiber of φ. With this new notation, one can write

$$(n)_{r_1}(n)_{r_2}\ldots(n)_{r_\ell} = \prod_{i=1}^{l}\left(n^{r_i}\sum_{m=0}^{r_i-1}(-1)^m [r_i]_m n^{-m}\right) = n^{|X|}\sum_{j=0}^{|X|}(-1)^j [X]_j^{\varphi} n^{-j}.$$

Turning back to (6.1), we let V_M and E_M denote the sets of vertices and edges, respectively, of $\Gamma_X(M)$. We denote by η the morphism $\eta_{M\to J}^X$, and use it implicitly also for its restrictions to V_M and E_M, which should cause no confusion. We obtain

$$L_{M,J}^X(n) = \frac{n^{|V_M|}\sum_{j=0}^{|V_M|}(-1)^j [V_M]_j^{\eta} n^{-j}}{n^{|E_M|}\sum_{j=0}^{|E_M|}(-1)^j [E_M]_j^{\eta} n^{-j}},$$

which by Claim 3.1(2) equals

$$L_{M,J}^X(n) = n^{-\widetilde{\mathrm{rk}}(M)}\frac{\sum_{j=0}^{|V_M|}(-1)^j [V_M]_j^{\eta} n^{-j}}{\sum_{j=0}^{|E_M|}(-1)^j [E_M]_j^{\eta} n^{-j}}. \tag{7.2}$$

Consider the denominator of (7.2) as a power series $Q(n^{-1})$. Its free coefficient is $[E_M]_0^{\eta} = 1$. This makes it relatively easy to get a formula for its inverse $1/Q(n^{-1})$ as a power series. In general, if $Q(x) = 1 + \sum_{i=1}^{\infty} a_i x^i$, then

$$\frac{1}{Q(x)} = \frac{1}{1-\sum_{i=1}^{\infty}(-a_i)x^i} = \sum_{t=0}^{\infty}\left(\sum_{i=1}^{\infty}(-a_i)x^i\right)^t =$$
$$= \sum_{t=0}^{\infty}\sum_{j_1,j_2,\ldots,j_t\ge 1}(-1)^t a_{j_1}\cdot\ldots\cdot a_{j_t} x^{\sum_{i=1}^t j_i}.$$

In the denominator of (7.2) we have $a_i = (-1)^i [E_M]_i^{\eta}$, and the resulting expression needs to be multiplied with the numerator $\sum_{j=0}^{|V_M|}(-1)^j [V_M]_j^{\eta} n^{-j}$. In total, we obtain

$$L_{M,J}^X(n) = \sum_{t=0}^{\infty}\sum_{\substack{j_0\ge 0\\ j_1,\ldots,j_t\ge 1}}(-1)^{t+\sum_{i=0}^t j_i}[V_M]_{j_0}^{\eta}\cdot[E_M]_{j_1}^{\eta}\cdot\ldots\cdot[E_M]_{j_t}^{\eta} n^{-\widetilde{\mathrm{rk}}(M)-\sum_{i=0}^t j_i}. \tag{7.3}$$

The combinatorial meaning and order of magnitude of $C^X_{M,N}$. The expression (7.3) is a bit complicated, but it presents $L^X_{M,J}(n)$ as a sum (with coefficients $\pm n^{-s}$) of terms with a combinatorial interpretation: the term $[V_M]^\eta_{j_0}\cdot[E_M]^\eta_{j_1}\cdot\ldots\cdot[E_M]^\eta_{j_t}$ counts $(t+1)$-tuples of η-subordinate permutations. The crux of the matter is that this interpretation allows us to attribute each tuple to a specific subgroup $N\in[M,J]_{\overrightarrow{X}}$. This is done as follows.

Let $(\sigma_0,\sigma_1,\ldots,\sigma_t)$ be a $(t+1)$-tuple of permutations such that $\sigma_0\in\mathrm{Sym}_\eta(V_M)$ and $\sigma_1,\ldots,\sigma_t\in\mathrm{Sym}_\eta(E_M)\setminus\{\mathrm{id}\}$ (we exclude $\mathrm{id}\in\mathrm{Sym}(E_M)$, which is the only permutation counted in $[E_M]^\eta_0$). Consider the graph $\Gamma={}^{\Gamma_X(M)}/_{\langle\sigma_0,\ldots,\sigma_t\rangle}$, which is the quotient of $\Gamma_X(M)$ by all identifications of pairs of the form $v,\sigma_0(v)$ ($v\in V_M$) and $e,\sigma_i(e)$ ($e\in E_M$, $1\le i\le t$).† Since Γ is obtained from $\Gamma_X(M)$ by identification of elements with the same η-image, η induces a well-defined morphism $\Gamma\to\Gamma_X(J)$. Thus, every closed path in Γ projects to a path in $\Gamma_X(J)$, giving $\pi^X_1(\Gamma)\le\pi^X_1(\Gamma_X(J))=J$. We denote $N=N_{\sigma_0,\sigma_1,\ldots,\sigma_t}=\pi^X_1(\Gamma)$. As usual (see Figures 3.2, 3.3), we can perform Stallings foldings on Γ until we obtain the core graph corresponding to N, $\Gamma_X(N)$. Obviously we have $M\le_{\overrightarrow{X}}N$, and by Claim 3.4(3) also $N\le_{\overrightarrow{X}}J$. Thus, we always have $N=N_{\sigma_0,\sigma_1,\ldots,\sigma_t}\in[M,J]_{\overrightarrow{X}}$. To summarize the situation,

(7.4)

$$\Gamma_X(M)\twoheadrightarrow\Gamma={}^{\Gamma_X(M)}/_{\langle\sigma_0,\ldots,\sigma_1\rangle}\xrightarrow{folding}\Gamma_X(N)\xrightarrow{\eta^X_{N\to J}}\Gamma_X(J)\,,\qquad \Gamma_X(M)\xrightarrow{\eta^X_{M\to N}}\Gamma_X(N)$$

Our next move is to rearrange (7.3) according to the intermediate subgroups $N\in[M,J]_{\overrightarrow{X}}$ which correspond to the tuples counted in it. For any $N\in[M,J]_{\overrightarrow{X}}$ we denote by $\mathcal{T}^X_{M,N,J}$ the set of tuples $(\sigma_0,\sigma_1,\ldots,\sigma_t)$ such that $N_{\sigma_0,\sigma_1,\ldots,\sigma_t}=N$, i.e.

$$\mathcal{T}^X_{M,N,J}=\left\{(\sigma_0,\sigma_1,\ldots,\sigma_t)\;\middle|\;\begin{array}{c}t\in\mathbb{N},\ \sigma_0\in\mathrm{Sym}_\eta(V_M)\\ \sigma_1,\ldots,\sigma_t\in\mathrm{Sym}_\eta(E_M)\setminus\{\mathrm{id}\}\\ \pi^X_1\left({}^{\Gamma_X(M)}/_{\langle\sigma_0,\sigma_1,\ldots,\sigma_t\rangle}\right)=N\end{array}\right\}.$$

The terms in (7.3) which correspond to a fixed $N\in[M,J]_{\overrightarrow{X}}$ thus sum to

$$\widetilde{C}^X_{M,J}(N)=\sum_{(\sigma_0,\sigma_1,\ldots,\sigma_t)\in\mathcal{T}^X_{M,N,J}}\frac{(-1)^{t+\sum_{i=0}^t\|\sigma_i\|}}{n^{\widetilde{\mathrm{rk}}(M)+\sum_{i=0}^t\|\sigma_i\|}},\tag{7.5}$$

and (7.3) becomes

$$L^X_{M,J}=\sum_{N\in[M,J]_{\overrightarrow{X}}}\widetilde{C}^X_{M,J}(N)\,.\tag{7.6}$$

The equation (7.6) looks much like (7.1), with $\widetilde{C}^X_{M,J}(N)$ playing the role of $C^X_{M,N}$. In order to establish equality between the latter two, we must show that $\widetilde{C}^X_{M,J}(N)$

†For the definition of the quotient of a graph by identifications of vertices see the discussion preceding Figure 3.3. Although we did not deal with merging of edges before, this is very similar to merging vertices. Identifying a pair of edges means identifying the pair of origins, the pair of termini and the pair of edges. In terms of the generated core graph (see Section 3), identifying a pair of edges is equivalent to identifying the pair of origins and/or the pair of termini.

does not depend on J. Fortunately, this is not hard: it turns out that

$$\widetilde{C}^X_{M,J}(N) = \widetilde{C}^X_{M,N}(N) \qquad \left(\forall N \in [M,J]_{\overrightarrow{X}}\right), \tag{7.7}$$

and the r.h.s. is, of course, independent of J. This equality follows from $\mathcal{T}^X_{M,N,J} = \mathcal{T}^X_{M,N,N}$, which we now justify. The only appearance J makes in the definition of $\mathcal{T}^X_{M,N,J}$ is inside $\eta = \eta^X_{M\to J}$, which is to be σ_i-invariant (for $0 \le i \le n$); i.e., σ_i must satisfy $\eta^X_{M\to J} \circ \sigma_i = \eta^X_{M\to J}$. If $(\sigma_0, \ldots, \sigma_t) \in \mathcal{T}^X_{M,N,J}$, then $\eta^X_{M\to N} \circ \sigma_i = \eta^X_{M\to N}$ follows from the fact that $\Gamma_X(N)$ is a quotient of ${}^{\Gamma_X(M)}/_{\langle\sigma_i\rangle}$. On the other hand, if $(\sigma_0, \ldots, \sigma_t) \in \mathcal{T}^X_{M,N,N}$, then we have $\eta^X_{M\to N} \circ \sigma_i = \eta^X_{M\to N}$, and hence also (see (7.4))

$$\eta^X_{M\to J} \circ \sigma_i = \eta^X_{N\to J} \circ \eta^X_{M\to N} \circ \sigma_i = \eta^X_{N\to J} \circ \eta^X_{M\to N} = \eta^X_{M\to J}.$$

Writing $\widetilde{C}^X_{M,N} \overset{def}{=} \widetilde{C}^X_{M,N}(N)$, we have by (7.1), (7.6), and (7.7)

$$C^X * \zeta^X = L^X = \widetilde{C}^X * \zeta^X$$

which shows that $C^X = \widetilde{C}^X$, as desired.

We approach the endgame. Let $(\sigma_0, \sigma_1, \ldots, \sigma_t) \in \mathcal{T}^X_{M,N,J} = \mathcal{T}^X_{M,N,N}$, and consider the partition P of $V(\Gamma_X(H))$, obtained by identifying v and v' whenever $\sigma_0(v) = v'$, or $\sigma_i(e) = e'$ for some $1 \le i \le t$ and edges e, e' whose origins are v and v', respectively. Since P can clearly be obtained by $\sum_{i=0}^t \|\sigma_i\|$ identifications, we have $\|P\| \le \sum_{i=0}^t \|\sigma_i\|$ (a strong inequality can take place—for example, one can have $\sigma_1 = \sigma_2$). Since $(\sigma_0, \sigma_1, \ldots, \sigma_t) \in \mathcal{T}^X_{M,N,J}$ we have $\pi_1^X\left({}^{\Gamma_X(H)}/_P\right) = N$, and thus by (3.5) we obtain

$$\rho_X(H,J) \le \|P\| \le \sum_{i=0}^t \|\sigma_i\|.$$

From (7.5) (recall that $\widetilde{C}^X_{M,J}(N) = \widetilde{C}^X_{M,N} = C^X_{M,N}$) we now have

$$C^X_{M,N}(n) = O\left(\frac{1}{n^{\widetilde{\mathrm{rk}}(M) + \rho_X(M,N)}}\right),$$

and Proposition 7.1 is proven.

8. Primitive words in the profinite topology

Theorem 1.4 has some interesting implications to the study of profinite groups. In fact, some of the original interest in the conjecture that is proven in this paper stems from these implications.

Let $\widehat{\mathbf{F}}_k$ denote the profinite completion of the free group $\mathbf{F}_k$. A basis of $\widehat{\mathbf{F}}_k$ is a set $S \subset \widehat{\mathbf{F}}_k$ such that every map from S to a profinite group G admits a unique extension to a continuous homomorphism $\widehat{\mathbf{F}}_k \to G$. It is a standard fact that $\mathbf{F}_k$ is embedded in $\widehat{\mathbf{F}}_k$, and that every basis of $\mathbf{F}_k$ is also a basis of $\widehat{\mathbf{F}}_k$ (see for example [Wil98]). An element of $\widehat{\mathbf{F}}_k$ is called *primitive* if it belongs to a basis of $\widehat{\mathbf{F}}_k$.

It is natural to ask whether an element of $\mathbf{F}_k$, which is primitive in $\widehat{\mathbf{F}}_k$, is already primitive in $\mathbf{F}_k$. In fact, this was conjectured by Gelander and by Lubotzky, independently. Theorem 1.4 yields a positive answer, as follows. An element $w \in \widehat{\mathbf{F}}_k$ is said to be *measure preserving* if for any finite group G, and a uniformly distributed random (continuous) homomorphism $\hat{\alpha}_G \in \mathrm{Hom}_{cont}(\widehat{\mathbf{F}}_k, G)$, the image $\hat{\alpha}_G(w)$ is uniformly distributed in G. By the natural correspondence $\mathrm{Hom}_{cont}(\widehat{\mathbf{F}}_k, G) \cong \mathrm{Hom}(\mathbf{F}_k, G)$, an element of $\mathbf{F}_k$ is measure preserving w.r.t. $\mathbf{F}_k$ iff it is so w.r.t.

$\widehat{\mathbf{F}}_k$. As in $\mathbf{F}_k$, a primitive element of $\widehat{\mathbf{F}}_k$ is easily seen to be measure preserving. Theorem 1.4 therefore implies that if $w \in \mathbf{F}_k$ is primitive in $\widehat{\mathbf{F}}_k$, then it is also primitive in $\mathbf{F}_k$. In other words,

Corollary 8.1. *If P denotes the set of primitive elements of $\mathbf{F}_k$, and $\widehat{P}$ the set of primitive elements of $\widehat{\mathbf{F}}_k$, then $P = \widehat{P} \cap \mathbf{F}_k$.*

As $\widehat{P}$ is a closed set in $\widehat{\mathbf{F}}_k$, this immediately implies Corollary 1.5, which states that P is closed in the profinite topology. In fact, there is also a direct proof to Corollary 1.5 from Theorem 1.8: one has to find, for every non-primitive word $w \in \mathbf{F}_k$, some $H \leq_{\text{f.i.}} \mathbf{F}_k$ such that the coset wH contains no primitives. By Theorem 1.8 there exists n so that w does not induce uniform distribution on S_n. For this n, $H = \bigcap_{\alpha:\mathbf{F}_k \to S_n} \ker(\alpha)$ has some primitive-free coset (as some of its cosets consist entirely of words which induce the uniform measure on S_n, while the others consist entirely of words which induce nonuniform measure on it).

This circle of ideas has a natural generalization. Observe the following five equivalence relations on the elements of $\mathbf{F}_k$:

- $w_1 \overset{A}{\sim} w_2$ if w_1 and w_2 belong to the same $\mathrm{Aut}\,\mathbf{F}_k$-orbit.
- $w_1 \overset{B}{\sim} w_2$ if w_1 and w_2 belong to the same $\overline{\mathrm{Aut}\,\mathbf{F}_k}$-orbit (where $\overline{\mathrm{Aut}\,\mathbf{F}_k}$ is the closure of $\mathrm{Aut}\,\mathbf{F}_k$ in $\mathrm{Aut}\,\widehat{\mathbf{F}}_k$).
- $w_1 \overset{C}{\sim} w_2$ if w_1 and w_2 belong to the same $\mathrm{Aut}\,\widehat{\mathbf{F}}_k$-orbit.
- $w_1 \overset{C'}{\sim} w_2$ if w_1 and w_2 have the same "statistical" properties, namely if they induce the same distribution on any finite group.
- $w_1 \overset{C''}{\sim} w_2$ if the evaluation maps $ev_{w_1}, ev_{w_2} : \mathrm{Epi}\,(F_k, G) \to G$ have the same images for every finite group G.

It is not hard to see that $(A) \Rightarrow (B) \Rightarrow (C) \Rightarrow (C') \Rightarrow (C'')$ (namely, that if $w_1 \overset{A}{\sim} w_2$, then $w_1 \overset{B}{\sim} w_2$, and so on). The only nontrivial implication is $(C') \Rightarrow (C'')$, which can be shown by induction on the size of G. In an unpublished manuscript, C. Meiri gave a one-page proof that (C), (C') and (C'') in fact coincide (in fact, these three coincide for all elements of $\widehat{\mathbf{F}}_k$).

From this perspective, our main result shows that in the case that w_1 is primitive, all five relations coincide, and it is natural to conjecture that they in fact coincide for all elements in $\mathbf{F}_k$.† Showing that $(A) \Leftarrow (B)$ would imply that $\mathrm{Aut}\,\mathbf{F}_k$-orbits in $\mathbf{F}_k$ are closed in the profinite topology, and the stronger statement $(A) \Leftarrow (C)$ would imply that words which lie in different $\mathrm{Aut}\,\mathbf{F}_k$-orbits can be told apart using statistical methods.

The analysis which is carried out in this paper does not suffice for the general case. For example, consider the words $w_1 = x_1x_2x_1x_2^{-1}$ and $w_2 = x_1x_2x_1^{-1}x_2^{-1}$. They belong to different $\mathrm{Aut}\,\mathbf{F}_2$-orbits, as $w_2 \in \mathbf{F}_2'$ but $w_1 \notin \mathbf{F}_2'$, but induce the same distribution on S_n for every n: their images under a random homomorphism are a product of a random permutation (σ) and a random element in its conjugacy class ($\tau\sigma\tau^{-1}$ for w_1, and $\tau\sigma^{-1}\tau^{-1}$ for w_2). However, while S_n do not distinguish between these two words, other groups do (in fact these words induce the same distribution on G precisely when every element in G is conjugate to its inverse; see [PS13] for a discussion of this).

†In [AV11], for example, the authors indeed ask whether $(C') \Rightarrow (A)$.

These questions also play a role in the theory of decidability in infinite groups. A natural extension of the word problem and the conjugacy problem is the following *automorphism problem*: given a group G generated by S, and two words $w_1, w_2 \in F(S)$, can it be decided whether w_1 and w_2 belong to the same $\mathrm{Aut}\, G$-orbit in G? Whitehead's algorithm [Whi36a, Whi36b] gives a concrete solution when $G = \mathbf{F}_k$. Showing that $(A) \Leftarrow (B)$ would provide an alternative decision procedure for $\mathbf{F}_k$.

More generally, and in a similar fashion to the conjugacy problem, it can be shown that if

(1) G is finitely presented
(2) $\mathrm{Aut}\, G$ is finitely generated
(3) $\mathrm{Aut}\, G$-orbits are closed in the profinite topology

then the automorphism problem in G is decidable. For the free group (1) and (2) are known, and (3) is exactly the conjectured coincidence $(A) \Leftrightarrow (B)$.

9. Open questions

We mention some open problems that naturally arise from the discussion in this paper.

- Section 8 shows how the questions about primitive elements can be extended to all $\mathrm{Aut}\, \mathbf{F}_k$-orbits in $\mathbf{F}_k$. (For example, is it true that $(A) \Leftrightarrow (B)$, and even the stronger equivalence $(A) \Leftrightarrow (C)$?) More generally, can statistical properties tell apart two subgroups $H_1, H_2 \leq_{fg} \mathbf{F}_k$ which belong to distinct $\mathrm{Aut}\, \mathbf{F}_k$-orbits? This would be a further generalization of Theorem 1.4.
- It is also interesting to consider words which are measure preserving w.r.t. other types of groups. For instance, does Theorem 1.4 still hold if we replace "finite groups" by "compact Lie groups," and study Haar-measure preserving words? Is there a single compact Lie group which suffices? Within finite groups, we showed that measure preservation w.r.t. S_n implies primitivity. Is it still true if we replace S_n by some other infinite family of finite groups (e.g. $\mathrm{PSL}_n(q)$, or solvable groups)?
- Is it true that
$$[H, \infty)_{\leq} = \bigcup_{\substack{X \text{ is a} \\ \text{basis of } \mathbf{F}_k}} [H, \infty)_{\overrightarrow{X}},$$
and under which assumptions does
$$[H, \infty)_{alg} = \bigcap_{\substack{X \text{ is a} \\ \text{basis of } \mathbf{F}_k}} [H, \infty)_{\overrightarrow{X}}$$
hold (see Remark 4.3)?
- The distribution induced by w on a finite group G is a class function, and so is a linear combination of the characters of G (for more on this point of view e.g. [AV11, PS13]). In particular, $\Phi_{\langle w \rangle, \mathbf{F}_k}(n) - 1$ is the coefficient of the *standard* character of S_n. The first nonzero term of $\Phi_{\langle w \rangle, \mathbf{F}_k} - 1$ encodes the primitivity rank and number of critical subgroups of w. Can the next terms be given an algebraic interpretation, and can they be estimated? (Such an estimate may contribute further to the study of expansion in graphs, which started in [Pud12].) What about the coefficients of other characters of S_n or of any other (family of) groups?

Glossary

		Reference	Remarks
$H \le_{fg} \mathbf{F}_k$	finitely generated		
$H \overset{*}{\le} J$	free factor		
$H \le_{alg} J$	algebraic extension	Definition 2.1	
$H \le_{\overset{\twoheadrightarrow}{X}} J$	H X-covers J	Definition 3.3	$H \overset{X}{\twoheadrightarrow} J$ in [Pud13]
$\mathfrak{sub}_{fg}\left(\mathbf{F}_k\right)$	the set of finitely generated subgroups of $\mathbf{F}_k$		
$[H, J]_{\preceq}$	$\{L \mid H \preceq L \preceq J\}$		$\preceq$ is either one of $\le, \overset{*}{\le}, \le_{alg}$ or $\overset{\twoheadrightarrow}{X}$ (standing for $\le_{\overset{\twoheadrightarrow}{X}}$)
$[H, J)_{\preceq}$	$\{L \mid H \preceq L \precneqq J\}$		
$[H, \infty)_{\preceq}$	$\{L \mid H \preceq L\}$		
$[H, \infty)_{\overset{\twoheadrightarrow}{X}}$	the X-quotients of H		$\mathcal{O}_X(H)$, or X-frigne in [MVW07]
$[H, \infty)_{alg}$	algebraic extensions of H		$AE(H)$ in [MVW07]
$\pi(H)$	primitivity rank of H	Definition 1.7	$\widetilde{\pi}(H) = \pi(H) - 1$
$\mathrm{Crit}(H)$	H-critical groups		
$\Gamma_X(H)$	X-labeled core graph of H		
$\rho_X(H, J)$	X-distance	Definition 3.6	$H \le_{\overset{\twoheadrightarrow}{X}} J$
$\eta^X_{H \to J}$	the morphism $\Gamma_X(H) \to \Gamma_X(J)$	Claim 3.2	$H \le J$
$\alpha_{J,n}$	a uniformly chosen random homomorphism in $\mathrm{Hom}(J, S_n)$		$J \le_{fg} \mathbf{F}_k$
$\Phi_{H,J}(n)$	the expected number of common fixed points of $\alpha_{J,n}(H)$	(2.1)	$H \le J$

Acknowledgments

It is a pleasure to thank our advisors Nati Linial and Alex Lubotzky for their support, encouragement and useful comments. We are grateful to Aner Shalev for supporting this research and for his valuable suggestions. We would also like to thank Uri Bader, Tsachik Gelander, Chen Meiri, Paul Nelson and Iddo Samet for their beneficial comments. We have benefited much from the mathematical open source community, and in particular from GAP [GAP4], and its free group algorithms package [Sie12]. The first author was supported by the ERC and by Adams Fellowship Program of the Israel Academy of Sciences and Humanities. The second author was supported by an Advanced ERC Grant.

References

[Abe06] Miklós Abért, *On the probability of satisfying a word in a group*, J. Group Theory **9** (2006), no. 5, 685–694, DOI 10.1515/JGT.2006.044. MR2253960 (2007j:20040)

[AL02] Alon Amit and Nathan Linial, *Random graph coverings. I. General theory and graph connectivity*, Combinatorica **22** (2002), no. 1, 1–18, DOI 10.1007/s004930200000. MR1883559 (2003a:05131)

[Alo86] Noga Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96, DOI 10.1007/BF02579166. Theory of computing (Singer Island, Fla., 1984). MR875835 (88e:05077)

[AV11] Alon Amit and Uzi Vishne, *Characters and solutions to equations in finite groups*, J. Algebra Appl. **10** (2011), no. 4, 675–686, DOI 10.1142/S0219498811004690. MR2834108

[BK13] Tatiana Bandman and Boris Kunyavskiĭ, *Criteria for equidistribution of solutions of word equations on* SL(2), J. Algebra **382** (2013), 282–302, DOI 10.1016/j.jalgebra.2013.02.031. MR3034482

[BS87] Andrei Broder and Eli Shamir, *On the second eigenvalue of random regular graphs*, 28th Annual Symposium on Foundations of Computer Science, IEEE, 1987, pp. 286–294.

[Fri03] Joel Friedman, *Relative expanders or weakly relatively Ramanujan graphs*, Duke Math. J. **118** (2003), no. 1, 19–35, DOI 10.1215/S0012-7094-03-11812-8. MR1978881 (2004m:05165)

[Fri08] Joel Friedman, *A proof of Alon's second eigenvalue conjecture and related problems*, Memoirs of the AMS **195** (September 2008), no. 910.

[GAP4] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.6.5* (2013).

[GS09] Shelly Garion and Aner Shalev, *Commutator maps, measure preservation, and T-systems*, Trans. Amer. Math. Soc. **361** (2009), no. 9, 4631–4651, DOI 10.1090/S0002-9947-09-04575-9. MR2506422 (2010f:20019)

[KM02] Ilya Kapovich and Alexei Myasnikov, *Stallings foldings and subgroups of free groups*, J. Algebra **248** (2002), no. 2, 608–668, DOI 10.1006/jabr.2001.9033. MR1882114 (2003a:20039)

[LP10] Nati Linial and Doron Puder, *Word maps and spectra of random graph lifts*, Random Structures Algorithms **37** (2010), no. 1, 100–135, DOI 10.1002/rsa.20304. MR2674623 (2011k:60026)

[LS08] Michael Larsen and Aner Shalev, *Characters of symmetric groups: sharp bounds and applications*, Invent. Math. **174** (2008), no. 3, 645–687, DOI 10.1007/s00222-008-0145-7. MR2453603 (2010g:20022)

[LS09] Michael Larsen and Aner Shalev, *Word maps and Waring type problems*, J. Amer. Math. Soc. **22** (2009), no. 2, 437–466, DOI 10.1090/S0894-0347-08-00615-2. MR2476780 (2010d:20019)

[MVW07] Alexei Miasnikov, Enric Ventura, and Pascal Weil, *Algebraic extensions in free groups*, Geometric group theory, Trends Math., Birkhäuser, Basel, 2007, pp. 225–253, DOI 10.1007/978-3-7643-8412-8_12. MR2395796 (2009f:20032)

[Nic94] Alexandru Nica, *On the number of cycles of given length of a free word in several random permutations*, Random Structures Algorithms **5** (1994), no. 5, 703–730, DOI 10.1002/rsa.3240050506. MR1300595 (95m:60017)

[PP12b] Ori Parzanchevski and Doron Puder, *Stallings graphs, algebraic extensions and primitive elements in F_2*, Mathematical Proceedings of the Cambridge Philosophical Society (2014). arXiv:1210.6574, to appear.

[PS13] Ori Parzanchevski and Gili Schul, *On the Fourier expansion of word maps*, Bull. London Math. Soc. **46** (2014), no. 1, 91–102, DOI 10.1112/blms/bdt068.

[Pud12] Doron Puder, *Expansion of random graphs: New proofs, new results* (2012). arXiv preprint arXiv:1212.5216.

[Pud13] Doron Puder, *Primitive words, free factors and measure preservation*, Israel Journal of Mathematics, posted on 2013, DOI 10.1007/s11856-013-0055-2.

[PW14] Doron Puder and Conan Wu, *Growth of primitives elements in free groups*, Journal of London Mathematical Society (2014). arXiv:1304.7979, to appear.

[Seg09] Dan Segal, *Words: notes on verbal width in groups*, London Mathematical Society Lecture Note Series, vol. 361, Cambridge University Press, Cambridge, 2009. MR2547644 (2011a:20055)

[Sha09] Aner Shalev, *Word maps, conjugacy classes, and a noncommutative Waring-type theorem*, Ann. of Math. (2) **170** (2009), no. 3, 1383–1416, DOI 10.4007/annals.2009.170.1383. MR2600876 (2011e:20017)

[Sha13] Aner Shalev, *Some results and problems in the theory of word maps* (L. Lovász, I. Ruzsa, V.T. Sós, and D. Palvolgyi, eds.), Erdős Centennial (Bolyai Society Mathematical Studies), Springer, 2013.

[Sie12] Christian Sievers, *Free Group Algorithms – a GAP package, Version 1.2.0* (2012).

[Sta83] John R. Stallings, *Topology of finite graphs*, Invent. Math. **71** (1983), no. 3, 551–565, DOI 10.1007/BF02095993. MR695906 (85m:05037a)

[Sta97] Richard P. Stanley, *Enumerative combinatorics. Vol. 1*, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 1997. With a foreword by Gian-Carlo Rota; Corrected reprint of the 1986 original. MR1442260 (98a:05001)

[Tak51] Mutuo Takahasi, *Note on chain conditions in free groups*, Osaka Math. J. **3** (1951), 221–225. MR0046362 (13,721e)

[Tur96] Edward C. Turner, *Test words for automorphisms of free groups*, Bull. London Math. Soc. **28** (1996), no. 3, 255–263, DOI 10.1112/blms/28.3.255. MR1374403 (96m:20039)

[vLW01] Jacobus H. van Lint and Richard M. Wilson, *A course in combinatorics*, 2nd ed., Cambridge University Press, Cambridge, 2001. MR1871828 (2002i:05001)

[Whi36a] John H. C. Whitehead, *On Certain Sets of Elements in a Free Group*, Proc. London Math. Soc. **S2-41**, no. 1, 48, DOI 10.1112/plms/s2-41.1.48. MR1575455

[Whi36b] John H. C. Whitehead, *On equivalent sets of elements in a free group*, Ann. of Math. **37** (1936), 768–800.

[Wil98] John S. Wilson, *Profinite groups*, London Mathematical Society Monographs. New Series, vol. 19, The Clarendon Press Oxford University Press, New York, 1998. MR1691054 (2000j:20048)

Einstein Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
E-mail address: doronpuder@gmail.com

School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540
E-mail address: parzan@ias.edu