Measure Preserving Words are Primitive

Doron Puder* ∗ Ori Parzanchevski†

Einstein Institute of Mathematics
Hebrew University, Jerusalem
doronpuder@gmail.com parzan@math.huji.ac.il

February 16, 2012

Abstract

A word $w \in F_k$, the free group on $k$ generators, is called primitive if it belongs to some basis of $F_k$. Associated with $w$ and a finite group $G$ is the word map $w : G \times \ldots \times G \to G$ defined on the direct product of $k$ copies of $G$. We call $w$ measure preserving if given uniform measure on $G \times \ldots \times G$, the image of this word map induces uniform measure on $G$ (for every finite group $G$). It is easy to see that every primitive word is measure preserving, and several authors have conjectured that the two properties are, in fact, equivalent. Here we prove this conjecture. The main ingredients of the proof include random coverings of Stallings graphs, algebraic extensions of free groups and Möbius inversions. Our methods yield the stronger result that a subgroup of $F_k$ is measure preserving iff it is a free factor.

As an interesting corollary of this result we resolve a question on the profinite topology of free groups and show that the primitive elements of $F_k$ form a closed set in this topology.

1 Introduction

Let $F_k$ be the free group on $k$ generators $X = \{x_1, \ldots, x_k\}$, and let $w \in F_k$ be expressed in the letters of the basis $X$, as $w = \prod_{j=1}^r x_j^{\varepsilon_j}$, where $\varepsilon_j = \pm 1$. Associated with $w$ is a word map: for every group $G$, this map, which we also denote by $w$, sends $G \times G \times \cdots \times G$ to $G$ by $(g_1, \ldots, g_k) \mapsto \prod_{j=1}^r g_j^{\varepsilon_j}$. Alternatively, we think of the word map $w$ as the evaluation map from $\text{Hom}(F_k, G)$ to $G$, i.e., $w(\alpha) = \alpha(w)$ for $\alpha \in \text{Hom}(F_k, G)$. The identification of $\text{Hom}(F_k, G)$ with $G^k$ is due to the fact that a homomorphism from a free group is uniquely

*Supported by the ERC and by Adams Fellowship Program of the Israel Academy of Sciences and Humanities.
†Supported by Advanced ERC Grant.
determined by choosing the images of the elements of a basis, and these images can be chosen arbitrarily.

The last years have seen a great interest in word maps in groups, and research around them has been very active (see, for instance, [Sha09], [LS09]; for a recent book on the topic see [Seg09]). Our focus here is on the property of measure preservation: for every finite group $G$, we consider $G \times \ldots \times G$ as being equipped with the uniform probability distribution. Then, every $w \in \mathbb{F}_k$ induces some measure on $G$ via the associated word map. We say that the word $w$ preserves measure with respect to $G$ if this induced measure is again uniform. (In other words, if all fibers of the word map have the same size). We say that $w$ is measure preserving if it preserves measure with respect to every finite group $G$.

Measure preservation can be equivalently defined in the language of homomorphisms: $w$ is measure preserving iff for every finite group $G$, the element $\alpha_G(w)$ is uniformly distributed over $G$, where $\alpha_G \in \text{Hom} (\mathbb{F}_k, G)$ is a homomorphism chosen uniformly at random.

This concept was investigated in several recent works. See for example [LS08] and [GS09], where certain word maps are shown to be almost measure preserving, in the sense that the distribution induced by $w$ on finite simple groups $G$ tends to uniform as $|G| \to \infty$.

One finds in the literature an even stronger notion of measure preservation on a word $w$, where the image of $w$ is considered over compact groups $G$ w.r.t. their Haar measure. Our results make use only of the weaker condition that involves solely finite groups.

The notion of measure preservation has a strong connection to primitivity. An element $w$ of a free group $J$ is called primitive if it belongs to some basis (free generating set) of $J$. When $J$ is given with a basis $X$, this is equivalent to the existence of an automorphism of $J$ which sends $w$ to a given element of $X$. It is an easy observation that primitive words are measure preserving. The reason is that as mentioned, a homomorphism $\alpha_G \in \text{Hom} (\mathbb{F}_k, G)$ is determined by the images of the elements of a basis of $\mathbb{F}_k$, which can be chosen arbitrarily and independently.

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 discussed among themselves: T. Gelander, M. Larsen, A. Lubotzky and A. Shalev. The question was mentioned several times in the Einstein Institute Algebra Seminar, and was independently raised in [AV11] and also in [LP10].

In [Pud11] we proved the conjecture for $\mathbb{F}_2$. Here we prove it in full. A key ingredient of the proof is the extension of the problem from single words to (finitely generated) subgroups of $\mathbb{F}_k$. 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

\[\text{It is interesting to note that there is an easy abelian parallel to this conjecture. A word } w \in \mathbb{F}_k \text{ belongs to a basis of } \mathbb{Z}_k \cong \mathbb{F}_k/\mathbb{F}_k' \text{ iff for any group } G \text{ the associated word map is surjective. See [Seg09], Lemma 3.1.1.}\]
denote this by $H^* \leq J$, if there is a subgroup $H' \leq J$ such that $H^* H' = J$. Equivalently, $H \leq J$ 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 $J$.)

The concept of measure preservation can also be extended from words in the free group to finitely generated subgroups (we write $H \leq_{fg} F_k$ when $H$ is a finitely generated subgroups of $F_k$):

**Definition 1.1.** Let $H \leq_{fg} F_k$. We say that $H$ is measure preserving if for every finite group $G$ and $\alpha_G \in \text{Hom}(F_k, G)$ a random homomorphism chosen with uniform distribution, $\alpha_G|_H$ is uniformly distributed in $\text{Hom}(H, G)$.

In particular, $1 \neq w \in F_k$ is measure preserving iff $\langle w \rangle$ is measure preserving. As for single words, it is easy to see that if $H \leq_{fg} F_k$ is a free factor, then $H$ is measure preserving. In [Pud11] we conjectured that the converse also holds, and proved this conjecture for subgroups of $F_k$ of rank $\geq k - 1$ (thus proving the conjecture for $F_2$). Here we prove the extended conjecture in full:

**Theorem 1.2.** For every $w \in F_k$,

$$w \text{ is measure preserving } \iff w \text{ is primitive}$$

More generally, for $H \leq_{fg} F_k$,

$$H \text{ is measure preserving } \iff H^* \leq F_k$$

In Section 7 we explain how this circle of ideas is related to the study of profinite groups. In particular we have the following corollary (see also Corollary 7.1):

**Corollary 1.3.** The set $P$ of primitive elements in $F_k$ is closed in the profinite topology.

In order to prove Theorem 1.2 one needs to exhibit, for each non-primitive word $w \in F_k$, 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 Theorem 1.2

**Theorem (1.2').** Let $w \in F_k$, and for every finite group $G$, let $\alpha_G \in \text{Hom}(F_k, G)$ 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.
For every \( n \in \mathbb{N} \) the random permutation \( \alpha_n(w) = \alpha_{S_n}(w) \) has uniform distribution.

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:

\[
\mathbb{E} [\# \text{fix}(\alpha_n(w))] = 1
\]

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

\[
\mathbb{E} [\# \text{fix}(\alpha_n(w))] \leq 1
\]

We already explained above the implication \((1) \Rightarrow (2)\), and \((2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5)\) 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 \((5) \Rightarrow (1)\) which is proven in this paper.

With the proper adjustments, Theorem 1.2 applies to f.g. subgroups as well. For example, the parallel of property \((4)\) for \( H \leq_f \mathbb{F}_k \) is that for every \( n \), the image \( \alpha_n(H) \subseteq S_n \) stabilizes on average exactly \( n^{1 - \text{rk}(H)} \) elements of \( \{1, \ldots, n\} \) (where \( \text{rk}(H) \) denotes the rank of \( H \)). As stated, it turns out that this property alone is enough to yield that \( H \) is a free factor of \( \mathbb{F}_k \).

A key role in the proof is played by the notion of primitivity rank, an invariant classifying words and f.g. subgroups of \( \mathbb{F}_k \), which was first introduced in [Pud11]. A primitive word \( w \in \mathbb{F}_k \) is also primitive in every subgroup containing it (Claim 2.8(3)). However, if \( w \) is not primitive in \( \mathbb{F}_k \), it may be either primitive or non-primitive in subgroups of \( \mathbb{F}_k \) 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 1.4.** The primitivity rank of \( w \in \mathbb{F}_k \), denoted \( \pi(w) \), is

\[
\pi(w) = \min \left\{ \text{rk}(J) \mid w \in J \leq \mathbb{F}_k \text{ s.t. } w \text{ is not primitive in } J \right\}.
\]

If no such \( J \) exists, i.e. if \( w \) is primitive, then \( \pi(w) = \infty \). A subgroup \( J \) for which the minimum is obtained is called \( w \)-critical.

More generally, for \( H \leq_f \mathbb{F}_k \), the primitivity rank of \( H \) is

\[
\pi(H) = \min \left\{ \text{rk}(J) \mid H \leq J \leq \mathbb{F}_k \text{ s.t. } H \text{ is not a free factor of } J \right\}.
\]

Again, if no such \( J \) exists, then \( \pi(H) = \infty \), and a subgroup \( J \) for which the minimum is obtained is called \( H \)-critical.

Note that for \( w \neq 1 \), \( \pi(w) = \pi(\langle w \rangle) \). For instance, \( \pi(w) = 1 \) if and only if \( w \) is a proper power (i.e. \( w = v^d \) for some \( v \in \mathbb{F}_k \) and \( d \geq 2 \)). By Claim 2.8(3),
Table 1: Primitivity Rank and Average Number of Fixed Points.

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

In \(\mathbb{F}_k\) the primitivity rank takes values only in \(\{0, 1, 2, \ldots, k\} \cup \{\infty\}\) (the only word \(w\) with \(\pi (w) = 0\) is \(w = 1\)). Moreover, \(\pi (H) = \infty\) iff \(H \leq \mathbb{F}_k\). Finally, Lemma 23 from \cite{Pud11} yields that \(\pi\) can take on every value in \(\{0, \ldots, k\}\). It is interesting to mention that \(\pi (H)\) also generalizes the notion of compressed subgroups, as appears, e.g., in \cite{MVW07}: a subgroup \(H \leq \left< \mathbb{F}_k \right>\) is compressed iff \(\pi (H) \geq \text{rk}(H)\). Finally, in Section 3 we show how \(\pi (\cdot)\) can be computed.

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

As mentioned above, our main result follows from an analysis of the average number of common fixed points of \(\alpha_n (H)\) (where \(\alpha_n\) denotes a uniformly distributed random homomorphism in \(\text{Hom} (\mathbb{F}_k, S_n)\)). In other words, we count the number of elements in \(\{1, \ldots, n\}\) stabilized by the images under \(\alpha_n\) of all elements of \(H\). Theorem 1.5 follows from the main result of this analysis:

**Theorem 1.5.** For every \(H \leq \left< \mathbb{F}_k \right>\), denote by \(\text{Crit}(H)\) the set of \(H\)-critical subgroups of \(\mathbb{F}_k\). The average number of common fixed points of \(\alpha_n (H)\) is

\[
\frac{1}{n^{\text{rk}(H)}} + \frac{|\text{Crit}(H)|}{n^{\text{rk}(H)}} + O\left(\frac{1}{n^{\text{rk}(H)+1}}\right).
\]

In particular, the set of \(H\)-critical subgroups \(\text{Crit}(H)\) is finite. Theorem 1.5 is stated for subgroups, but it equally applies to single words: the average number of fixed points of the random permutation \(\alpha_n (\ell)\) is the same as the average number of common fixed points of \(\alpha_n ((\ell))\). This average is 1 + \(|\text{Crit}(\ell)| \cdot n^{-\pi(\ell)} + O\left(n^{-\pi(\ell)-1}\right)\), where \(\text{Crit}(\ell)\) denotes the set of \(\ell\)-critical subgroups. Table 1 summarizes the connection implied by Theorem 1.5 between the primitivity rank of \(\ell\) and the average number of fixed points in the random permutation \(\alpha_n (\ell)\).

Theorem 1.5 implies the following general corollary regarding the family of distributions of \(S_n\) induced by word maps:

\footnote{For example, it follows from this lemma that in \(\mathbb{F}_k\), for every \(1 \leq d \leq k\), \(\pi (x_1^2 \ldots x_d^2) = d\).}
Corollary 1.6. For an imprimitive $w \in F_k$ the average number of fixed points in $\alpha_n(w)$ is strictly greater than 1, for large enough $n$.

Corollary 1.6 is in fact the missing piece $(5) \Rightarrow (1)$ in Theorem 1.2. In addition, it follows from this corollary that for every $w \in F_k$ 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 [Pud11]. Theorem 1.5 and its consequences, Corollaries 7.1, 1.3 and 1.6, are stated there as conjectures (Conjectures 7, 24 and 27).

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 in $n$ which gives this average for every large enough $n$. When $w = u^d$ with $d$ maximal (so $u$ is not a power), he shows that the limit distribution of the number of fixed points in $\alpha_n(w)$ as $n \to \infty$ depends solely on $d$.

His results yield that the average number of fixed points in $\alpha_n(w)$ is $\delta(d) + O\left(\frac{1}{n}\right)$ where $\delta(d)$ is the number of (positive) divisors of $d$ ([Nic94], Corollary 1.3). This is equivalent to Theorem 1.5 for words which are powers, namely, for $w \in F_k$ with $\pi(w) = 1$ (see Table 1). Indeed, if $w = u^d$ as above with $d > 1$, then

$$\text{Crit}(w) = \{ (u^m) \mid 1 \leq m < d, m \mid d \},$$

so $|\text{Crit}(w)| = \delta(d) - 1$.

The first steps in extending these results to non-power words were made in [LP10], where the case $\pi(w) = 2$ was partially handled. The statement of Theorem 1.5 for this case was completely proven in [Pud11]. In [Pud11] the analysis was also extended to f.g. subgroups, and the statement of Theorem 1.5 was more generally proven for every $H \leq f_g F_k$ with $\text{rk}(H) \geq k - 1$. However, the technique used in these two papers is specialized for the proven cases, and so far could not be generalized to yield the general result.

In the current paper, better analysis of the average number of fixed points in $\alpha_n(w)$ is achieved by studying a broader question: for every pair of $H, J \leq f_g F_k$ such that $H \leq J$, we define for $n \in \mathbb{N}$

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

where $\alpha_{J,n} \in \text{Hom}(J, S_n)$ is a random homomorphism chosen with uniform distribution. The original function studied by Nica is the special case $\Phi_{\langle w \rangle, F_k}(n)$.

In addition, the current paper uses much more extensively various combinatorial structures of posets (partially order sets) on the set of f.g. subgroups of $F_k$. It is suggestive to ask whether this holds for all $n$. Namely, is it true that for every $w \in F_k$ 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 $L$. 

---

3It is suggestive to ask whether this holds for all $n$. Namely, is it true that for every $w \in F_k$ 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.

4Nica’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 $L$. 

---

6
This set has, of course, a natural structure of a lattice given by the relation of inclusion. However, there are other interesting partial orders defined on this set, some of which are much sparser and have the advantage of being locally finite. If \( \leq \) is some partial order on the subgroups of \( F_k \), and \( H, J \leq F_k \), we define the closed interval
\[
[H, J]_\leq = \{ L \leq F_k \mid H \preceq L \preceq J \}
\]
and similarly the open interval \( (H, J) \leq = \{ L \leq F_k \mid H \prec L \preceq J \} \), the half-bounded interval \( [H, \infty) \leq = \{ L \leq F_k \mid H \preceq L \} \), and so on (see also the glossary).

The sparsest partial order we consider here is that of Algebraic Extensions. This notion was introduced in [KM02] and further studied in [MVW07].

**Definition 1.7.** Let \( H, J \) be free groups. We say that \( J \) is an algebraic extension of \( H \), denoted \( H \leq_{\text{alg}} J \), 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 \( H \preceq J \)).

The relation of algebraic extension turns out to be a partial order on the set of subgroups of \( F_k \): it is clearly reflexive and antisymmetric, but it is also transitive (Claim 3.1). In addition, it is very sparse: it turns out that \( [H, \infty)_{\text{alg}} \), the set of algebraic extensions of \( H \), is finite for every \( H \leq_{fg} F_k \), so in particular this partial order is locally finite on the set of f.g. subgroups of \( F_k \). (These claims appear in [MVW07]. We repeat some of the arguments in Section 3.

It is a simple observation that \( H \)-critical subgroups are in particular algebraic extensions of \( H \), i.e. \( \text{Crit} (H) \subseteq [H, \infty)_{\text{alg}} \). In fact, they are the proper algebraic extensions of minimal rank (see Section 3).

Between the rather dense partial order of inclusion and the very sparse one of algebraic extensions, lies an infinite family of intermediate partial orders, one for every choice of basis for \( F_k \). Although the dependency on the basis makes these orders somewhat less universal, they turn out to be extremely useful. These partial orders are 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). Fix a basis \( X = \{ x_1, \ldots, x_k \} \) of \( F_k \). Associated with every \( H \leq F_k \) is a directed, pointed graph denoted \( \Gamma_X (H) \), whose edges are labeled by the elements of the basis \( X \). A full definition appears in Section 2, but we illustrate the concept in Figure 1.1. It shows the core graph of the subgroup of \( F_2 \) generated by \( x_1 x_2^{-1} x_1 \) and \( x_1^{-2} x_2 \).

The aforementioned relation holds for the pair \( H, J \) iff the associated core graph \( \Gamma_X (J) \) is a quotient (as a pointed labeled graph) of the core graph \( \Gamma_X (H) \)

---

1. A lattice is a partially ordered set in which every two elements have unique infimum and supremum.
2. A locally finite poset is one in which every closed interval \( [a, b] = \{ x : a \leq x \leq b \} \) is finite.
Figure 1.1: The core graph $\Gamma_X(H)$ where $H = \langle x_1x_2^{-1}x_1x_1^{-2}x_2 \rangle \leq F_2$.

(see Definition 2.3). We denote this by $H \leq \approx J$. When $H \leq_{fg} F_k$, $\Gamma_X(H)$ is finite (Claim 2.1(1)), and thus has only finitely many quotients. Furthermore, it turns out that different groups correspond to different core graphs. Thus, this partial order, restricted to f.g. subgroups, is too locally finite.

As claimed, these partial orders are indeed intermediate: if $H \leq_{alg} J$ then $H \leq \approx J$ for every basis $X$, and if $H \leq \approx J$ then $H \leq J$. (These claims appear in [MvW07], and also here in Sections 2 and 3). In [Pud11], the poset of f.g. subgroups of $F_k$ w.r.t. “$\leq$” was studied, and a distance function denoted $\rho_X(H, J)$ was introduced for $H, J \leq_{fg} F_k$ such that $H \leq \approx J$ (Definition 2.6 in the current paper). The following theorem ([Pud11], Theorem 1) is a key ingredient in the proof of Theorem 1.5:

**Theorem 1.8.** Let $H, J \leq_{fg} F_k$ and assume further that $H \leq \approx J$. Then $H \leq \approx J$ if and only if

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

The local-finiteness of the partial orders “$\leq_{alg}$” and “$\leq$” allows us to use machinery such as Möbius inversion on functions defined on the set of comparable pairs of f.g. subgroups. Indeed, we use this to “derive” the function $\Phi$ (recall (1.1)) and obtain its “right derivation” $R^X$, its “left derivation” $L^X$, and its “two sided derivation” $C^X$ (see Section 4). For instance, $\Phi_{H, J}$ can be presented as finite sums of $R^X$:

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

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

The proof of Theorem 1.5 (yielding our main result) is based on the following results:

- (Proposition 4.1) The right derivation $R^X$ is supported on algebraic extensions, i.e. if $H \leq_{\approx} M$ but $M$ is not an algebraic extension of $H$ then $R^X_{H, M} \equiv 0$.

- (The discussion in Section 5) The random homomorphism $\alpha_{J,n} \in \text{Hom}(J, S_n)$ can be encoded as a random covering-space of the core graph $\Gamma_X(J)$, and $\Phi_{H, J}(n)$ can then be interpreted as the expected number of lifts of $\Gamma_X(H)$. 

8
(Lemmas 5.3 and 5.4) The left derivation \( L^X \) is then the expected number of injective lifts of the core graph \( \Gamma_X(H) \) into the random covering of the core graph \( \Gamma_X(J) \), and a rational expression can be computed for \( L^X_{H,J} \).

(Proposition 6.1 and Section 6.1) An analysis involving Stirling numbers of the rational expressions for \( L^X \) yields a combinatorial meaning for the two-sided derivation \( C^X \). Using Theorem 1.8 and the classification of primitivity rank we obtain a first-order estimate for the size of \( C^X_{H,J} \).

(Proposition 6.2) From \( C^X \) we return to \( R^X \) (by “left-integration”), obtaining that whenever \( H \leq_{\text{alg}} M \) we have

\[
R^X_{H,M} = \frac{1}{n^r_k(M)} + O\left(\frac{1}{n^r_k(M)+1}\right)
\]

and by right integration of \( R^X \), we obtain the order of magnitude of \( \Phi^X \), which was our goal.

The paper is arranged as follows: in Section 2 the notion of core graphs is explained in details, as well as the partial order \( \leq_X \), and the background for Theorem 1.8. In Section 3 we survey the main properties of algebraic extensions of free groups. Section 4 is devoted to recalling Möbius derivations on locally-finite posets and introducing the different derivations of \( \Phi \). In Section 5 we discuss the connection of the problem to random coverings of graphs and analyze the left derivation \( L^X \). The proof of Theorem 1.5 is completed in Section 6 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 7. 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.

2 Core Graphs and the Partial Order “\( \leq_X \)”

Fix a basis \( X = \{x_1, \ldots, x_k\} \) of \( F_k \). Associated with every subgroup \( H \leq F_k \) is a directed, pointed graph whose edges are labeled by \( X \). This graph is called the 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 \( \Upsilon_X(H) \). This is a directed, pointed and edge-labeled graph. Its vertex set is the set of all right cosets of \( H \) in \( F_k \), 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 \( Hwx_j \).

\( \Gamma_X(H) \) is the quotient \( H \backslash T \), where \( T \) is the Cayley graph of \( F_k \) 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 \( \overline{\Gamma_X(F_k)} = \Gamma_X(F_k) \), 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 \( \pi_1(Y) \).

\footnote{Alternatively, \( \Gamma_X(H) \) is the quotient \( H \backslash T \), where \( T \) is the Cayley graph of \( F_k \) 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 \( \overline{\Gamma_X(F_k)} = \Gamma_X(F_k) \), 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 \( \pi_1(Y) \).}
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 2.1 shows the graphs \( \overline{\Gamma}_X (H) \) and \( \Gamma_X (H) \) for \( H = \langle x_1 x_2 x_1^{-3}, x_1^2 x_2 x_1^{-2} \rangle \leq F_2 \). 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 \( \overline{\Gamma}_X (H) \) has at most one outgoing \( j \)-edge and at most one incoming \( j \)-edge, for every \( 1 \leq j \leq k \).

Figure 2.1: \( \overline{\Gamma}_X (H) \) and \( \Gamma_X (H) \) for \( H = \langle x_1 x_2 x_1^{-3}, x_1^2 x_2 x_1^{-2} \rangle \leq F_2 \). 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 “\( \otimes \)” corresponds to the trivial coset \( H \), the vertex below it corresponds to the coset \( Hx_1 \), the one further down corresponds to \( Hx_1^2 = H x_1 x_2 x_1^{-1} \), 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.

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

\[
\bullet \rightarrow x_2 \bullet \rightarrow x_2 \bullet \rightarrow x_1 \bullet \rightarrow x_2 \bullet \rightarrow x_3 \bullet \rightarrow x_2 \bullet \rightarrow x_1 \bullet
\]

corresponds to the word \( x_2 x_1 x_2^{-1} x_3 x_2 x_1^{-1} \). The set of all words obtained from closed paths around the basepoint in \( \Gamma \) is a subgroup of \( F (X) \) which we call the \textit{labeled fundamental group} of \( \Gamma \), 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 \( \Gamma \) viewed as a topological space - for example, take \( \Gamma = x_1 \bigcirc \bigcirc x_1 \).

However, it is not hard to show that when \( \Gamma \) 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 \( F(X) \). It is an easy observation that

\[
\pi_1^X (\Gamma_X (H)) = \pi_1^X (\Gamma_X (H)) = H \tag{2.1}
\]

This gives a one-to-one correspondence between subgroups of \( F(X) = F_k \) and core graphs labeled by \( X \). Namely, \( \pi_1^X \) and \( \Gamma_X \) are the inverses of each other in a bijection (Galois correspondence)

\[
\left\{ \text{Subgroups of } F(X) \right\} \overset{\Gamma_X, \pi_1^X}{\leftrightarrow} \left\{ \text{Core graphs labeled by } X \right\} \tag{2.2}
\]

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. (This can be easily introduced into Stallings' framework, see [Pud11]). We remark that it is possible to study core graphs from a purely combinatorial point of view, as labeled pointed 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 \( F_k \) then gives a correspondence between these graphs and subgroups of \( F_k \).

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 2.1.** Let \( H \) be a subgroup of \( F_k \) 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. \( \text{rk} (H) < \infty \iff \Gamma \text{ is finite.} \)
2. \( \tilde{\text{rk}} (H) = -\chi (\Gamma) \).
3. The correspondence \( \text{(2.2)} \) restricts to a correspondence between finitely generated subgroups of \( F_k \) and finite core graphs.

Given a finite set of words \( \{ h_1, \ldots, h_m \} \subseteq F(X) \) 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 2.2 a folding process by which we obtain the core graph $\Gamma_X(H)$ of $H = \langle x_1x_2x_1^{-3}, x_1^2x_2x_1^{-2} \rangle \leq F_2$ from the given generating set.

Figure 2.2: Constructing the core graph $\Gamma_X(H)$ of $H = \langle x_1x_2x_1^{-3}, x_1^2x_2x_1^{-2} \rangle \leq F_2$ 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.

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.

As in Claim 2.1, each of the following properties is either proven in (some of) [Sta83, KM02, MVW07] or an easy observation:

Claim 2.2. Let $H, J, L \leq F_k$ be subgroups. Then

1. A morphism $\Gamma_X(H) \rightarrow \Gamma_X(J)$ exists if and only if $H \leq J$.
2. If a morphism $\Gamma_X(H) \rightarrow \Gamma_X(J)$ exists, it is unique. We denote it by $\eta^X_{H \rightarrow J}$.
3. Whenever $H \leq L \leq J$, $\eta^X_{H \rightarrow J} = \eta^X_{L \rightarrow J} \circ \eta^X_{H \rightarrow L}$.

Points (1)-(3) can be formulated by saying that $\eta^X_{H \rightarrow J}$ is in fact an isomorphism of cate-
If \( \eta_{H \to J}^X \) is injective, then \( H \preceq J \).

(5) Every morphism in an immersion (locally injective at the vertices).

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

**Definition 2.3.** Let \( H \leq J \leq F_k \). Whenever \( \eta_{H \to J}^X \) 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 \( \Gamma_X (H) \to \Gamma_X (J) \). As for the groups, we say that \( H X \)-covers \( J \) and denote this by \( H \leq J \).

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 5 we do study topological covering maps, and we reserve the term “coverings” for these.

For instance, \( H = \langle x_1 x_2 x_1^{-3}, x_1^2 x_2 x_1^{-2} \rangle \leq F_k \) \( X \)-covers the group \( J = \langle x_2, x_1^2, x_1 x_2 x_1 \rangle \), the corresponding core graphs of which are the leftmost and rightmost graphs in Figure [2.3]. As another example, a core graph \( \Gamma_X (F_k) \) (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 \( H \preceq J \) indeed depends on the given basis \( X \) of \( F_k \). For example, if \( H = \langle x_1 x_2 \rangle \) then \( H \preceq F_2 \). However, for \( Y = \{ x_1 x_2, x_2 \} \), \( H \) does not \( Y \)-cover \( F_2 \), 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 “\( \preceq \)" indeed constitutes a partial ordering of the set of subgroups of \( F_k \). We make a few other useful observations:

**Claim 2.4.** Let \( H, J, L \leq F_k \) be subgroups. Then

1. Whenever \( H \preceq J \) there exists an intermediate subgroup \( M \) such that \( H \preceq M \preceq J \).

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

3. If \( H \preceq J \) and \( H \preceq L \preceq J \), then \( L \preceq J \).

4. If \( H \) is finitely generated then it \( X \)-covers only a finite number of groups.
   In particular, the partial order “\( \preceq \)" restricted to f.g. subgroups of \( F_k \) is locally finite.

**Proof.** Point (1) follows from the factorization of the morphism \( \eta_{H \to J}^X \) to a surjection followed by an embedding. Indeed, it is easy to see that the image of \( \eta_{H \to J}^X \) is a sub-graph 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 given by the functors \( \pi_1^X \) and \( \Gamma_X \).

9But not vice-versa: for example, consider \( \langle x_1 x_2^2 \rangle \preceq F_2 \).
around the basepoint). Let $M = \pi_1^X (\im \eta_{\Gamma \to J})$ be the subgroup corresponding to this sub-core-graph. (1) now follows from points (1) and (4) in Claim 2.2. Point (2) follows from the uniqueness of such factorization of a morphism. Point (3) follows from the fact that if $\eta_{\Gamma \to J} = \eta_{\Gamma \to L} \circ \eta_{\Gamma \to H}$ is surjective then so is $\eta_{\Gamma \to J}$. Point (4) follows from the fact that $\Gamma_X (H)$ is finite (Claim 2.1(1)) and thus has only finitely many quotients, and each quotient correspond to a single group (by (2.2)).

Following [MVW07], we call the set of $X$-quotients of $H$, namely,

$$[H, \infty)_X = \left\{ J \mid H \leq_X J \right\}$$

(2.3)

the $X$-fringe of $H$. Claim 2.4(4) states in this terminology that for every $H \leq f_g F_k$ (and every basis $X$), $[H, \infty)_X < \infty$. Note that $[H, \infty)_X$ 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 $\pi_1^X (\im \eta_{\Gamma \to F_k})$).

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 2.2) until we obtain a core graph. Since Stallings foldings do not affect $\pi_1^X$, the core graph we obtain in this manner is $\Gamma_X (J)$, where $J = \pi_1^X (\Gamma_X (H)/P)$. The resulting partition $P$ of $V (\Gamma_X (H))$ (as the fibers of $\eta_{\Gamma \to F_k}$) 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 2.3.

Thus, there is sense in examining the quotient of a core graph $\Gamma$ “generated” by some partition $P$ of its vertex set, namely, $\Gamma_X (\pi_1^X (\Gamma)/P)$. 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 $P \subseteq 2^X$ is a partition of some set $X$, let

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

(2.4)

Namely, $\| P \|$ is the number of elements in the set minus the number of blocks in the partition. For example, $\| P \| = 1$ if $P$ 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 $X$ in order to obtain the equivalence relation $P$.

**Definition 2.5.** Let $\Gamma$ 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 $\Delta$ be the core graph generated from $\Gamma$ by $P$. We then say that $\Delta$ is an immediate quotient of $\Gamma$. 

14
Figure 2.3: The left graph is the core graph $\Gamma_{X}(H)$ of $H = \langle x_{1}x_{2}x_{1}^{-3}, x_{1}x_{2}x_{1}^{-2} \rangle \leq F_{2}$. 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 2.2). The resulting core graph, $\Gamma_{X}(\pi_{1}(\Gamma_{X}(H)/P))$, is shown on the right and corresponds to the partition $\bar{P} = \{\{v_{1}, v_{4}\}, \{v_{2}, v_{3}\}\}$.

Alternatively, we say that $\Delta$ is generated by identifying a single pair of vertices of $\Gamma$. For instance, the rightmost core graph in Figure 2.3 is an immediate quotient of the leftmost one.

The main reason that immediate quotients are interesting is their algebraic significance. Let $H, J \leq F_{k}$ with $\Gamma = \Gamma_{X}(H), \Delta = \Gamma_{X}(J)$ their core graphs, and assume that $\Delta$ is an immediate quotient of $\Gamma$ obtained by identifying the vertices $u, v \in V(\Gamma)$. Now let $w_{u}, w_{v} \in F_{k}$ be the words corresponding to some paths $p_{u}, p_{v}$ in $\Gamma$ 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).\footnote{That is, a directed graph with no directed cycles.}
This DAG was first introduced in [Pud11], and is denoted by \( D_k \). The set of vertices of \( D_k \) 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 \( F_k \) determines a one-to-one correspondence between the vertices of this graph and all finitely generated subgroups of \( F_k \).

It is easy to see that in the case of finite core graphs, \( \Delta \) is a quotient of \( \Gamma \) if and only if \( \Delta \) is reachable from \( \Gamma \) in \( D_k \) (that is, there is a directed path from \( \Gamma \) to \( \Delta \)). In other words, if \( H \leq_{fg} F_k \) then \( H \leq \Delta \) iff \( \Gamma_X \) (\( J \)) can be obtained from \( \Gamma_X(H) \) by a finite sequence of immediate quotients. Thus, for any \( H \leq_{fg} F_k \), the subgraph of \( D_k \) 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 the \( X \)-fringe \( [H, \infty) \). By Claim 2.4(4), this subgraph is finite. In Figure 2.4 we draw the subgraph of \( D_k \) consisting of all quotients of \( \Gamma_X(H) \) when \( H = \langle x_1 x_2 x_1^{-1} x_2^{-1} \rangle \). The edges of this subgraph (i.e. immediate quotients) are denoted by the dashed arrows in the figure.

Figure 2.4: The subgraph of \( D_k \) induced by \([H, \infty) \), that is, all quotients of the core graph \( \Gamma = \Gamma_X(H) \), for \( H = \langle x_1 x_2 x_1^{-1} x_2^{-1} \rangle \). The dashed arrows denote immediate quotients, i.e. quotients generated by merging a single pair of vertices. \( \Gamma \) has exactly seven quotients: itself, four immediate quotients, and two quotients at distance 2.

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

**Definition 2.6.** Let \( H, J \leq_{fg} F_k \) be subgroups such that \( H \leq X J \), 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 \( \Gamma \) to \( \Delta \) in \( D_k \).

In other words, \( \rho_X (H, J) \) is the length of the shortest series of immediate quotients that yields \( \Delta \) from \( \Gamma \). There is another useful equivalent definition for the \( X \)-distance. To see this, assume that \( \Gamma' \) is generated from \( \Gamma \) by the partition \( P \) of \( V (\Gamma) \) and let \( \eta : \Gamma \to \Gamma' \) 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 \( \Gamma' \) by identifying \( x \) and \( y \) is the same as the one generated by \( P' \) from \( \Gamma \). From these considerations we obtain that

\[
\rho_X (H, J) = \min \left\{ \| P \| \mid \begin{array}{l}
\text{P is a partition of } V (\Gamma_X (H)) \\
\text{such that } \pi^{X}_{1} (\Gamma_X (H)/P) = J
\end{array} \right\},
\]

(2.5)

For example, if \( \Delta \) is an immediate quotient of \( \Gamma \) then \( \rho_X (H, J) = \rho (\Gamma, \Delta) = 1 \). For \( H = \langle x_1 x_2 x_1^{-1} x_2^{-1} \rangle \), \( \Gamma_X (H) \) has four quotients at distance 1 and two at distance 2 (see Figure 2.3).

As aforementioned, 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 (thought not every element of \( F_k \) 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 \( H \leq X J \), one has

\[
\text{rk} (J) - \text{rk} (H) \leq \rho_X (H, J)
\]

(2.6)

In [Pud11] (Lemma 11), the distance is bounded from above as well:

**Claim 2.7.** Let \( H, J \leq_{fg} F_k \) such that \( H \leq X J \). Then

\[
\text{rk} (J) - \text{rk} (H) \leq \rho_X (H, J) \leq \text{rk} (J)
\]

Theorem 1.8 (the proof of which can be found in [Pud11]) then asserts that the lower bound is attained if and only if \( H \) is a free factor of \( J \):

\[
\rho_X (H, J) = \text{rk} (J) - \text{rk} (H) \iff H \leq^* J
\]

In fact, one of the implications of Theorem 1.8 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 \( \text{rk} (J) - \text{rk} (H) \) merging steps, this means we can obtain \( J \) from \( H \) by adding \( \text{rk} (J) - \text{rk} (H) \) extra generators to \( H \), hence clearly \( H \leq^* J \)\[1\]

\[1\]This relies on the well known fact that a set of size \( k \) which generates \( F_k \) is a basis.
The other implication is not trivial and constitutes the essence of the proof of Theorem 1 in [Pud11]. The difficulty is that when $H \preceq J$, 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 $\Gamma_X(H)$. As mentioned, this theorem is heavily relied upon in the proof of Theorem 1.5.

We finish this section with some classical facts about free factors that will be useful in the next section. We also give an easy graph-theoretical proof to exemplify the strength of core graphs.

**Claim 2.8.** Let $H, J$ and $K$ be subgroups of $\mathbb{F}_k$.

1. If $H \preceq J$ and $K \leq J$, then $H \cap K \leq K$.

2. If $H,K \preceq J$ then $H \cap K \leq J$.

3. If $H \preceq J$ 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.

1. 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 2.4). Consequently, $\Gamma_Y(H \cap K)$ is embedded in $\Gamma_Y(K)$ and by Claim 2.2(4), $H \cap K \leq K$.

2. This follows from (1) by the transitivity of free factorness.

3. Let $H \leq M \leq J$. Since $\Gamma_Y(H)$ is a bouquet, the morphism $\eta^{Y}_{H,M}$ must be injective, and we conclude again by Claim 2.2(4).

In particular, the last claim shows that if $H \preceq \mathbb{F}_k$ then $\pi(H) = \infty$ (see Definition 1.4). On the other hand, if $H$ is not a free factor of $\mathbb{F}_k$, then obviously $\pi(H) \leq rk(\mathbb{F}_k) = k$. Thus $\pi(H) \in \{0,1,2,\ldots,k\} \cup \{\infty\}$.

### 3 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 1.7) 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 $H = \langle x_1, x_2, x_1^{-1}, x_2^{-1} \rangle \leq F_2$. A proper free factor of $F_2$ 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 $F_2$ (as can be inferred from Theorem 1.8 and Figure 2.4). Thus, $H \leq_{alg} F_2$. In fact, we shall see that in this case $[H, \infty]_{alg} = \{H, F_2\}$.

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

**Claim 3.1.** The relation “$\leq_{alg}$” is a transitive partial order.

**Proof.** Assume that $H \leq_{alg} M \leq_{alg} J$. Let $H \leq L \leq J$. By Claim 2.8, $L \cap M \leq M$. But $H \leq L \cap M$ and $H \leq_{alg} M$, so $L \cap M = M$, and thus $L \leq M$. So now $M \leq L \leq J$, and from $M \leq_{alg} J$ we obtain that $L = J$. $\square$

Next, we show that “$\leq_{alg}$” is dominated by “$\leq_{x}$” for every basis $X$ of $F_k$. Namely, if $H \leq_{alg} J$ then $H \leq_{x} J$. This shows, in particular, that the poset of finitely generated subgroups of $F_k$ w.r.t. “$\leq_{alg}$” is locally-finite.

**Claim 3.2.** If $H \leq_{alg} J$ then $H \leq_{x} J$ for every basis $X$ of $F_k$.

**Proof.** By Claim 2.8, there is an intermediate subgroup $M$ such that $H \leq_{x} M \leq_{alg} J$, and from $H \leq_{alg} J$ it follows that $M = J$. $\square$

**Remark 3.3.** It is natural to conjecture that the converse also holds, namely that if $H \leq_{x} J$ for every basis $X$ of $F_k$ 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 $J = \langle x_1^2 x_2^2, x_1 x_2 \rangle$, $H \leq_{x} J$ for every basis $X$ of $F_2$, but $J$ is not an algebraic extension of $H$ [PP12]. However, there are bases of $F_3$ 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 $F$, $H \leq_{x} J$. It is also plausible that the original conjecture from [MVW07] holds for $F_k$ with $k \geq 3$.

In a similar fashion, one can ask whether $H \leq J$ if and only if for some basis $X$ of $F_k$, $H \leq_{x} J$.

Claim 3.2 establishes completely the relations, mentioned in Section 1, between the different partial orders we consider in this paper: inclusion, the big family based on $\leq_{x}$, and algebraic extensions. Recall that $H$-critical subgroups are a special kind of algebraic extensions. Thus:

$$\text{Crit}(H) \subseteq [H, \infty]_{alg} \subseteq [H, \infty]_{x} \subseteq [H, \infty]_{\leq}.$$ 

Theorem 1.8 and Claim 3.2 give the following criterion for algebraic extensions:

**Lemma 3.4.** Let $H \leq_{fg} F_k$. The algebraic extensions of $H$ are the elements of $[H, \infty]_{x}$ which are not immediate quotients of any subgroup in $[H, \infty]_{x}$ of smaller rank.
Proof. Let \( J \in [H, \infty)_X \). If \( J \) is an immediate \( X \)-quotient of \( L \in [H, \infty)_X \) with \( \text{rk}(L) < \text{rk}(J) \), then by Theorem 1.8 \( H \leq L \leq J \), hence \( J \) is not an algebraic extension of \( H \). On the other hand, assume there exists some \( L \) such that \( H \leq L \leq J \). By Claim 2.4 there exists \( M \) such that \( H \leq M \leq L \). By Claim 2.4(3), \( M \) is an immediate \( X \)-quotient of \( J \). From Theorem 1.8 it follows that there is a chain of immediate quotients \( M = M_0 \leq M_1 \leq \ldots \leq M_r = J \) with \( \text{rk}(M_{i+1}) = \text{rk}(M_i) + 1 \), and \( M_{r-1} \) is the group we have looked for.

Since the subgraph of \( D_k \) induced by the vertices corresponding to \([H, \infty)_X\), namely \( \Gamma_X(H) \) and its descendants, is finite and can be effectively computed, Lemma 3.4 yields a straightforward algorithm to find all algebraic extensions of a given \( H \leq_{fg} F_k \) (this algorithm was first introduced in [Pud11]). This, in particular, allows one to find all \( H \)-critical subgroups, and thus to compute the primitivity rank \( \pi(H) \): the subgroups constituting \( \text{Crit}(H) \) are those in \((H, \infty)_{\text{alg}} \) of minimal rank, which is \( \pi(H) \). For instance, Figure 2.4 shows that \( H \leq \{H, F_2\} \). Thus, \( \text{Crit}(H) = \{F_2\} \) and \( \pi(H) = 2 \) (so \( \tilde{\pi}(H) = 1 \)).

We conclude this section with yet another elegant result from [KM02, MVW07] that will be used in the proof of Theorem 1.5. In the spirit of field extensions, it says that every extension of subgroups of \( F_k \) has a unique factorization to an algebraic extension followed by a free extension (compare this with Claim 2.4(1.2)):

**Claim 3.5.** Let \( H \leq J \) be free groups. Then there is a unique subgroup \( L \) of \( J \) such that \( H \leq \text{alg} L \leq J \). Moreover, \( L \) is the intersection of all intermediate free factors of \( J \) and the union of all intermediate algebraic extensions of \( H \):  
\[
L = \bigcap_{M: H \leq M \leq J} M = \bigcup_{M: H \leq \text{alg} M \leq J} M \quad (3.1)
\]

In particular, the intersection of all free factors is a free factor, and the union of all algebraic extensions is an algebraic extension. Claim 3.5 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 2.8 and rank considerations, the intersection in the middle of (3.1) is by itself a free factor of \( J \). Denote it by \( L \), so we have \( H \leq L \leq J \). 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 \( H \leq \text{alg} M \leq J \). By Claim 2.8(1), \( H \leq M \cap L \leq M \), so \( M \cap L = M \), that is \( M \leq L \).
4 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\(^{12}\) of all functions from pairs \(\{(x, y) \in P \times P \mid x \leq y\}\) 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 \([x, y] = \{z \mid x \leq z \leq y\}\)). The unit element is the diagonal

\[\delta(x, y) = \begin{cases} 1 & x = y \\ 0 & x \not\leq y \end{cases}.\]

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

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

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

Recall the function \(\Phi\)\(^{13}\), defined for every pair of free subgroups \(H, J \leq_J F\) such that \(H \leq_X J\): \(\Phi_{H, J}(n)\) is the expected number of common fixed points of \(\alpha_{J,n}(H)\), where \(\alpha_{J,n} \in \text{Hom}(J, S_n)\) is a random homomorphism chosen with uniform distribution. We think of \(\Phi\) 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 \(F\). We write \(\Phi^X\) for the restriction of \(\Phi\) to pairs \((H, J)\) such that \(H \leq_X J\). As “\(\leq_X\)” defines a locally finite partial ordering of the set of finitely generated subgroups of \(F\), there exists a matching Möbius function, \(\mu^X = (\zeta^X)^{-1}\) (where \(\zeta^X_{H,J} = 1\) for all \(H \leq_X J\)). Our proof of Theorem\(^1\)

\(^{12}\)The theory of incidence algebra of posets can be found in [Sta97].
consists of a detailed analysis of the left, right, and two-sided derivations of $\Phi^X$:

$$
\Phi^X \\
L^X \overset{\text{def}}{=} \mu^X \ast \Phi^X \\
R^X \overset{\text{def}}{=} \Phi^X \ast \mu^X \\
C^X \overset{\text{def}}{=} \mu^X \ast \Phi^X \ast \mu^X
$$

By definition, we have for every f.g. $H \leq \overset{\sim}{X} J$:

$$
\Phi_{H,J} = \sum_{M \in [H,J]_X} L^X_{M,J} = \sum_{M,N: H \leq \overset{\sim}{X} M \leq \overset{\sim}{X} N \leq \overset{\sim}{X} J} C^X_{M,N} = \sum_{N \in [H,J]_X} R^X_{H,N} \quad (4.1)
$$

Note that (4.1) can serve as definitions for the three functions $L^X, C^X, R^X$: for instance, $L^X = \mu^X \ast \Phi^X$ is equivalent to $\zeta^X \ast L^X = \Phi^X$, which is the leftmost equality above.

We begin the analysis of these functions by the following striking observation regarding $R^X$. Recall (Claim 3.2) that if $H \leq_{\text{alg}} J$ then $H \leq \overset{\sim}{X} J$ 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 4.1.** Let $H, J \leq_{f.g} F_k$.

1. If $H \leq \overset{\sim}{X} J$ 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 $F_k$, whenever both are defined.

**Remark 4.2.** The only property of $\Phi$ we use is that $\Phi_{H,L} = \Phi_{H,J}$ whenever $H \leq \overset{\sim}{X} L \leq \overset{\sim}{X} J$, which is easy to see from the definition of $\Phi$. 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 $|[H,J]_X|$, the size of the closed interval between $H$ and $J$. The induction basis is $H = J$. That $H \leq_{\text{alg}} H$ is immediate. By (4.1), $R^X_{H,H} = \Phi_{H,H}$ and so $R^X_{H,H}$ is indeed independent of the basis $X$.

Assume now that $|[H,J]_X| = r$ and that both claims are proven for every pair bounding an interval of size $< r$. By (4.1) and the first claim of the induction hypothesis,

$$
R^X_{H,J} = \Phi_{H,J} - \sum_{N \in [H,J]_X} R^X_{H,N} = \Phi_{H,J} - \sum_{N: H \leq_{\text{alg}} \overset{\sim}{X} N \leq \overset{\sim}{X} J} R^X_{H,N} \quad (4.2)
$$

22
By Claim 2.4(3), \( \{ N \mid H \leq_{\text{alg}} N \leq_{X} J \} = \{ N \mid H \leq_{\text{alg}} N \leq_{X} J \} \), and the latter is independent of the basis \( X \). Furthermore, by the induction hypothesis regarding the second claim, so are the terms \( R_{H,N}^{X} \) 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 \( J, H \leq_{X} L \leq_{X} J \). As mentioned above, this yields that \( \Phi_{H,J} = \Phi_{H,L} \). Therefore,

\[
R_{H,J}^{X} = \Phi_{H,J} - \sum_{N \in (H,J)^{X}} R_{H,N}^{X} = \Phi_{H,L} - \sum_{N \in (H,L)^{X}} R_{H,N}^{X} - \sum_{N \in (H,J)^{X} \setminus (H,L)^{X}} R_{H,N}^{X} \leq 0 \text{ by definition}
\]

By Claim 3.5 all algebraic extensions of \( H \) inside the interval \( [H,J]^{X} \) are contained in \( L \). Hence, every subgroup \( N \in [H,J]^{X} \setminus [H,L]^{X} \) is not an algebraic extension of \( H \), and by the induction hypothesis \( R_{H,N}^{X} \) vanishes. The desired result follows.

In view of Proposition 4.1 we can omit the superscript and write from now on \( R_{H,J} \) instead of \( R_{H,J}^{X} \). Moreover, we can write the following “basis independent” equation for every pair of f.g. subgroups \( H \leq J \):

\[
\Phi_{H,J} = \sum_{N: H \leq_{\text{alg}} N \leq_{X} J} R_{H,N} \tag{4.3}
\]

When \( H \leq_{X} J \) this follows from the proof above. For general \( H \leq J \), there is some subgroup \( L \) such that \( H \leq_{X} L \leq_{X} J \) and every intermediate algebraic extension \( H \leq_{\text{alg}} N \leq_{X} J \) is contained in \( L \) (see Claims 2.4 and 3.5). Therefore,

\[
\Phi_{H,J} = \Phi_{H,L} = \sum_{N: H \leq_{\text{alg}} N \leq_{X} L} R_{H,N} = \sum_{N: H \leq_{\text{alg}} N \leq_{X} J} R_{H,N}.
\]

It turns out that unlike the function \( R \), the other two derivations of \( \Phi \), namely \( L_{X}^{X} \) and \( C_{X}^{X} \), 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 5.2 and 5.3 below). This, in turn, allows us to analyze the combinatorial meaning and order of magnitude of \( C_{M,N}^{X} \) (Proposition 6.1).

Finally, using the fact that \( R \) is the “left integral” of \( C_{X}^{X} \), that is \( R = \zeta_{X}^{X} \ast C_{X}^{X} \), we finish the circle around the diagram of \( \Phi \)’s derivations, and use this analysis of \( \Phi, L_{X}^{X} \) and \( C_{X}^{X} \) to prove that for every pair \( H \leq_{\text{alg}} J \), \( R_{H,J} \) does not vanish and is, in fact, positive for large enough \( n \). This alone gives Theorem 1.2. The more informative 1.5 follows from an analysis of the order of magnitude of \( R_{H,J} \) in this case (Proposition 6.2).
5 Random Coverings of Core Graphs

This section studies the graphs which cover a given core-graph in the topological sense, i.e., \( \hat{\Gamma} \rightarrow \Gamma \) with \( p \) locally bijective. We call these graphs (together with their projection maps) coverings of \( \Gamma \). The reader should not confuse this with our notion “covers” from Definition 2.3.

We focus on directed and edge-labeled coverings. This means we only consider \( \hat{\Gamma} \rightarrow \Gamma \) such that \( \hat{\Gamma} \) is directed and edge-labeled, and the projection \( p \) preserves orientations and labels. When \( \Gamma \) is a core-graph we do not assume that \( \hat{\Gamma} \) 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 \( \Gamma \) and \( \hat{\Gamma} \) are finite, for every vertex \( v \) in \( p^{-1}(\otimes) \), the fiber over \( \Gamma \)'s basepoint, we do have a valid core-graph, which we denote by \( \hat{\Gamma}_v \): this is the connected component of \( v \) in \( \hat{\Gamma} \), with \( v \) serving as basepoint. Moreover, the restriction of the projection map \( p \) to \( \hat{\Gamma}_v \) 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 5.1.** Let \( \Gamma \) be a core-graph, \( \hat{\Gamma} \rightarrow \Gamma \) 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 \( \Gamma \) and \( \pi_1^X(\Gamma) \)-sets.

2. In this correspondence, coverings of \( \Gamma \) with fiber \( \{1, \ldots, n\} \) correspond to actions of \( \pi_1^X(\Gamma) \) on \( \{1, \ldots, n\} \), i.e., to group homomorphisms \( \pi_1^X(\Gamma) \rightarrow S_n \).

3. The group \( \pi_1^X\left(\hat{\Gamma}_v\right) \) is the stabilizer of \( v \) in the action of \( \pi_1^X(\Gamma) \) on \( p^{-1}(\otimes) \) (note that \( \pi_1^X\left(\hat{\Gamma}_v\right) \) and \( \pi_1^X(\Gamma) \) are both subgroups of \( F(X) \)).

4. A core-graph morphism \( \Delta \rightarrow \Gamma \) can be lifted to a core-graph morphism \( \Delta \rightarrow \hat{\Gamma}_v \) (i.e., the diagram

\[
\begin{array}{ccc}
\hat{\Gamma}_v & \rightarrow \\
\downarrow^p & \\
\Gamma
\end{array}
\]

can be completed) if and only if \( \pi_1^X(\Delta) \subseteq \pi_1^X\left(\hat{\Gamma}_v\right) \). 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 $\Gamma$. 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 (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 $\Gamma$ with fewer random permutations: one may choose some spanning tree $T$ of $\Gamma$, associate the identity permutation with every edge in $T$, and pick random permutations only for edges outside $T$.

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

By Claim 5.1(2), each random $n$-covering of $\Gamma_X(J)$ encodes a homomorphism $\alpha_{J,n} \in \text{Hom}(J, S_n)$, 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 $(\varnothing, i)$ ends at $(\varnothing, j)$ 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 $\text{Hom}(J, S_n)$. This can also be verified using the “economical” model, as follows: choose some basis $Y = \{y_1, \ldots, y_k(J)\}$ 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 $H \leq J \leq \mathcal{F}$ and $1 \leq i \leq n$. By [5.1][13], the morphism $\eta_{H \rightarrow J}^X : \Gamma_X(H) \rightarrow \Gamma_X(J)$ lifts to a core-graph morphism $\Gamma_X(H) \rightarrow \hat{\Gamma}_X(J)_i$ iff $H = \pi_1^X(\Gamma_X(H))$ fixes $(\varnothing, i)$ via the action of $J$ on the fiber $\varnothing \times \{1, \ldots, n\}$. Since this action is given by $\alpha_{J,n}$, this means that $\eta_{H \rightarrow J}^X$ lifts to $\hat{\Gamma}_X(J)_i$ 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:

---

13Occasionally these random coverings are referred to as random lifts of graphs. We shall reserve this term for its usual meaning.
Lemma 5.2. Let \( \hat{\Gamma}_X (J) \) be a random \( n \)-covering space of \( \Gamma_X (J) \) in the aforementioned model from [AL02]. Then,

\[
\Phi_{H,J} (n) = \text{The expected number of lifts of } \eta_{H \to J}^X \text{ to } \hat{\Gamma}_X (J).
\]

\[
\hat{\Gamma}_X (J) \quad \overset{\text{p}}{\longrightarrow} \quad \Gamma_X (H) \quad \overset{\eta_{H \to J}^X}{\longrightarrow} \quad \Gamma_X (J)
\]

Note that this characterization of \( \Phi_{H,J} \) involves the basis \( X \), although the original definition (1.1) 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 4) the definition of the function \( L_X \), that satisfies \( \Phi_{H,J} = \sum_{M \in \mathbb{X}} L_{M,J}^X \) for every \( H \leq X \). It turns out that this derivation of \( \Phi \) also has a geometrical interpretation. Assume that \( \eta_{M \to J}^X \) does lift to \( \hat{\eta}_i : \Gamma_X (H) \to \hat{\Gamma}_X (J) \). By Claim 2.4, \( \hat{\eta}_i \) decomposes as a quotient onto \( \Gamma_X (M) \), where \( M = \pi_X^1 (\text{im} \hat{\eta}_i) \), followed by an embedding. Moreover, \( M \) lies in \( [H, J]_X \). On the other hand, if there is some \( M \in [H, J]_X \) such that \( \Gamma_X (M) \) is embedded in \( \hat{\Gamma}_X (J) \), then such \( M \) is unique and \( \hat{\eta}_i \) lifts to the composition of \( \eta_{M \to J}^X \) with this embedding. We obtain

\[
\Phi_{H,J} (n) = \text{Expected number of lifts of } \eta_{H \to J}^X \text{ to } \hat{\Gamma}_X (J) = \sum_{M \in \mathbb{X}} \text{Expected number of injective lifts of } \eta_{M \to J}^X \text{ to } \hat{\Gamma}_X (J).
\]

Taking the left derivations, we obtain:

Lemma 5.3. Let \( M \leq X \) and \( \hat{\Gamma}_X (J) \) be a random \( n \)-covering space of \( \Gamma_X (J) \) in the aforementioned model from [AL02]. Then,

\[
L_{M,J}^X (n) = \text{The expected number of injective lifts of } \eta_{M \to J}^X \text{ to } \hat{\Gamma}_X (J).
\]

\[
\hat{\Gamma}_X (J) \quad \overset{\text{p}}{\longrightarrow} \quad \Gamma_X (M) \quad \overset{\eta_{M \to J}^X}{\longrightarrow} \quad \Gamma_X (J)
\]

Unlike the number of lifts in general, the number of injective lifts does depend on the basis \( X \). For instance, consider \( M = \{x_1x_2\} \) and \( J = \{x_1, x_2\} = \mathbb{F}_2 \). With the basis \( X = \{x_1, x_2\} \), the probability that \( \eta_{M \to J}^X \) lifts injectively to \( \hat{\Gamma}_X (J) \) equals \( \frac{n-1}{n^2} \) (Lemma 5.4 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 5.3 allows a natural extension of \( L^X \) to pairs \( M, J \) such that \( M \) does not \( X \)-cover \( J \).

Lemma 5.3 allows us to generalize the method used in [Nic94, LP10, Pud11] to compute the expected number of fixed points in \( \alpha_n(w) \) (see the notations before Theorem (1.2')). We claim that for \( n \) large enough, \( L^X_{M,J}(n) \) is a simple rational expression in \( n \).

**Lemma 5.4.** Let \( M, J \leq \mathbb{F}_k \) such that \( M \leq \mathbb{X} \), and let \( \eta = \eta^X_{M \rightarrow J} \) be the core-graph morphism. For large enough \( n \),

\[
L^X_{M,J}(n) = \frac{\prod_{v \in V(\hat{\Gamma}_X(J))} (n|_{\eta^{-1}(v)})}{\prod_{e \in E(\hat{\Gamma}_X(J))} (n|_{\eta^{-1}(e)})} 
\tag{5.1}
\]

Here, “large enough \( n \)” is \( n \geq \max_{e \in E(\hat{\Gamma}_X(J))} |\eta^{-1}(e)| \) (so that the denominator does not vanish), and \((n)_r\) is the falling factorial \( n(n-1)\ldots(n-r+1)\).

**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 \( \hat{\eta} : \Gamma_X(M) \hookrightarrow \hat{\Gamma}_X(J) \), the fiber \( \eta^{-1}(v) \) is mapped injectively into the fiber \( p^{-1}(v) \). The number of such injections is

\[
(n|_{\eta^{-1}(v)}) = n(n-1)\ldots(n-|\eta^{-1}(v)|+1),
\]

and therefore the number of injective lifts of \( \eta \big|_{V(\Gamma_X(M))} \) into \( V(\hat{\Gamma}_X(J)) \) is the numerator of (5.1).

We claim that any such injective lift has a positive probability of extending to a full lift of \( \eta \): 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 \( \hat{\eta} : V(\Gamma_X(M)) \hookrightarrow V(\hat{\Gamma}_X(J)) \) be an injective lift of the vertices of \( \Gamma_X(M) \) as above, and let \( e \) be some edge of \( \Gamma_X(J) \). If \( \hat{\eta} \) is to be extended to \( \eta^{-1}(e) \), the fiber above \( e \) in \( \hat{\Gamma}_X(J) \) must contain, for every \( (u,v) \in \eta^{-1}(e) \), the edge \( (\hat{\eta}(u),\hat{\eta}(v)) \).

Thus, the random permutation \( \sigma \in S_n \) which determines the perfect matching above \( e \) in \( \hat{\Gamma}_X(J) \), 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 (5.1). \( \square \)

This immediately gives a formula for \( \Phi_{H,J} \) as a rational function:
Corollary 5.5. Let $H, J \leq_f G$ such that $H \leq \bar{J}$. Then, for large enough $n$,

$$\Phi_{H,J}(n) = \sum_{M \in [H,J]_{\bar{X}}} L^X_M(n) = \sum_{M \in [H,J]_{\bar{X}}} \prod_{v \in V(\Gamma_X(J))} \prod_{e \in E(\Gamma_X(J))} \frac{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}}{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}} \bigg| \prod_{v \in V(\Gamma_X(J))} \prod_{e \in E(\Gamma_X(J))} \frac{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}}{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}} \bigg|.$$

Since $H$ $X$-covers every intermediate $M \in [H,J]_{\bar{X}}$, the largest fiber above every edge of $\Gamma_X(J)$ is obtained in $\Gamma_X(H)$ itself. Thus, “large enough $n$” in the last corollary can be replaced by $n \geq \max_{e \in E(\Gamma_X(J))} \left| \frac{\eta_{H \rightarrow J}}{\eta_{H \rightarrow J}}(e) \right|$.

In fact, Corollary 5.5 applies, with slight modifications, to every pair of f.g. subgroups $H \leq J$: Lemma 5.2 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 $\Gamma_X(J)$. 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 $H \leq \bar{M} \leq J$. In explaining the rational expression in Lemma 5.4, we did not need $\bar{M}$ to cover $J$. Thus, for every $H \leq J$, both finitely generated,

$$\Phi_{H,J}(n) = \sum_{M : H \leq \bar{M} \leq J} \prod_{v \in V(\Gamma_X(J))} \prod_{e \in E(\Gamma_X(J))} \frac{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}}{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}} \bigg| \prod_{v \in V(\Gamma_X(J))} \prod_{e \in E(\Gamma_X(J))} \frac{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}}{\binom{n}{(\eta_{H \rightarrow J})^{-1}(e)}} \bigg|.$$

Corollary 5.5 yields in particular a straight-forward algorithm to obtain a rational expression in $n$ for $\Phi_{H,J}(n)$ (valid for large enough $n$). For example, consider $H = \langle x_1x_2x_1^{-1}x_2^{-1} \rangle$ and $F_2 = \langle x_1, x_2 \rangle$. The interval $[H,F_2]_{\bar{X}}$ consists of seven subgroups, as depicted in Figure 2.4. Following the computation in Corollary 5.5, we get that for $n \geq 2$ (we scan the quotients in Figure 2.4 top-to-bottom and in each row left-to-right):

$$\Phi_{H,F_2}(n) = \frac{(n)_1}{(n)_2(n)_2} + \frac{(n)_2}{(n)_2(n)_1} + \frac{(n)_2}{(n)_2(n)_1} + \frac{(n)_3}{(n)_2(n)_2} + \frac{(n)_3}{(n)_2(n)_1} + \frac{(n)_3}{(n)_2(n)_1} + \frac{(n)_4}{(n)_2(n)_2} + \frac{(n)_4}{(n)_2(n)_1}$$

$$= \frac{n}{n-1} + 1 + \frac{1}{n} + O\left(\frac{1}{n^2}\right).$$

This exemplifies Theorem 4.5 and Table 1 for $H = \langle x_1x_2x_1^{-1}x_2^{-1} \rangle$ (recall the discussion following Lemma 5.4, where it is shown that $\pi(H) = 2$ and that Crit$(H) = \{F_2\}$).

Remark 5.6. 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 $\text{Hom}(F_k, G)$ and $|S|$-coverings. Although the computation of $L^X$ and of $\Phi$ might be more involved, this suggests a way of analyzing words which are measure preserving w.r.t. $G$.
6 The proof of Theorem 1.5

The last major ingredient of the proof of our main result, Theorem 1.5, is an analysis of $C^X$, the double-sided derivation of $\Phi$. Recall Definition 2.6 where the $X$-distance $\rho_X(H, J)$ was defined for every $H, J \leq F_k$ with $H \leq X J$.

**Proposition 6.1.** Let $M, N \leq F_k$ satisfy $M \leq X N$. Then

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

Section 6.1 is dedicated to the proof of this proposition. But before getting there, we want to show how it practically finishes the proof of our main result. We do this with the following final step:

**Proposition 6.2.** Let $H, N \leq F_k$ satisfy $H \leq alg N$. Then

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

**Proof.** Let $X$ be some basis of $F_k$. Recall that $R = \zeta^X \ast C^X$, i.e.

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

For $M = N$ we have $C^X_{N, N} (n) = R_{N, N} (n) = \Phi_{N, N} (n) = n^{-rk(N)}$ (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. $M \in [H, N]_X$, the fact that $N$ is an algebraic extension of $H$ means that $M$ is not a free factor of $N$ and therefore, by Theorem 1.8 (and (2.6)), $\rho_X (M, N) \geq \frac{1}{2} (rk(N) - rk(M) + 1)$. Proposition 6.1 then shows that

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

Hence,

$$R_{H, N} (n) = C^X_{N, N} (n) + \sum_{M \in [H, N]_X} C^X_{M, N} (n) = \frac{1}{n^{rk(N)}} + O \left( \frac{1}{n^{rk(N) + 1}} \right).$$

\[\square\]

The proof of Theorem 1.5 is now at hand. For every $H, J \leq F_k$ with $H \leq J$, by (4.4) and Proposition 6.2

$$\Phi_{H, J} (n) = \sum_{N : H \leq alg N \leq J} R_{H, N} (n)$$

$$= R_{H, H} (n) + \sum_{N : H \leq alg N \leq J} R_{H, N} (n)$$

$$= \frac{1}{n^{rk(H)}} + \sum_{N : H \leq alg N \leq J} \frac{1}{n^{rk(N)}} + O \left( \frac{1}{n^{rk(N) + 1}} \right).$$
For $J = F_k$ we can be more concrete. Recall that the $H$-critical groups, $\text{Crit}(H)$, are its algebraic extensions of minimal rank (other than $H$ itself), and this minimal rank is $\pi(H)$. Therefore,

$$
\Phi_{H,F_k}(n) = \frac{1}{n^{\text{rk}(H)}} + \sum_{N \in (H,\infty)_{alg}} \frac{1}{n^{\text{rk}(N)}} + O\left(\frac{1}{n^{\text{rk}(N)+1}}\right)
$$

$$
= \frac{1}{n^{\text{rk}(H)}} + \frac{|\text{Crit}(H)|}{n^{\pi(H)}} + O\left(\frac{1}{n^{\pi(H)+1}}\right)
$$

This establishes our main results: Theorem 1.5, Theorem 1.2 and all their corollaries.

### 6.1 The Analysis of $C_{M,N}^X$

In this subsection we look into $C^X$, the double-sided derivation of $\Phi$, and establish Proposition 6.1 which deals with the order of magnitude of $C_{M,N}^X$. Recall that by definition $C^X = L^X * \mu^X$, which is equivalent to

$$
L_{X,M,J}^X = \sum_{N \in [M,J]_X} C_{M,N}^X \quad (\forall M \leq X J)
$$

We derive a combinatorial meaning of $C_{M,N}^X$ from this relation. To obtain this, we further analyze the rational expression (5.1) for $L_{X,M,J}^X$ 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 $N \in [M,J]_X$, and show that for every $N \in [M,J]_X$, the sum of terms attributed to $N$ is nothing but $C_{M,N}^X$. Finally, we use this combinatorial interpretation of $C_{M,N}^X$ to estimate its order of magnitude.

**Rewriting $L_{X,M,J}^X$ as a Power Series in $n^{-1}$**

Consider the numerator and denominator of (5.1): these are products of expressions of the type $(n)_r$. It is a classical fact that

$$
(n)_r = \sum_{j=1}^{r} (-1)^{r-j} \left(\begin{array}{c} r \\ j \end{array}\right) n^j
$$

where $\left(\begin{array}{c} r \\ j \end{array}\right)$ is the *unsigned Stirling number of the first kind*. That is, $\left(\begin{array}{c} r \\ j \end{array}\right)$ is the number of permutations in $S_r$ with exactly $j$ cycles (see, for instance, [vL W01], Chapter 13).

We introduce the notation $[r]_j := \left(\begin{array}{c} r \\ r-j \end{array}\right)$, which is better suited for our purposes. The cycles of a permutation $\sigma \in S_r$ constitute a partition $P_\sigma$ of $\{1, \ldots, r\}$. We define $\|\sigma\| = \|P_\sigma\|$ (recall (2.4)), 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 $\sigma$. 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 $|\varphi^{-1}(i)| = r_i$ ($1 \leq i \leq \ell$). We denote by

$$\text{Sym}_{\varphi}(X) = \{\sigma \in \text{Sym}(X) : \varphi \circ \sigma = \varphi\}$$

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

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

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

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

Turning back to (5.1), we let $V_M$ and $E_M$ denote the sets of vertices and edges, respectively, of $\Gamma_X(M)$. We denote by $\eta$ the morphism $\eta^X_{M \to J}$, and use it implicitly also for its restrictions to $V_M$ and $E_M$, which should cause no confusion. We obtain

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

which by Claim 2.1(2) equals

$$L^X_{M,J}(n) = n^{-\tilde{r}_k(M)} \frac{\sum_{j=0}^{|V_M|} (-1)^j [V_M]_j^\eta n^{-j}}{\sum_{j=0}^{|E_M|} (-1)^j [E_M]_j^\eta n^{-j}}. \quad (6.2)$$

31
Consider the denominator of (6.2) as a power series $Q(n^{-1})$. Its free coefficient is $|E_M|^n_1 = 1$. 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

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

In the denominator of (6.3) we have $a_i = (-1)^i |E_M|^\eta_j$, and the resulting expression needs to be multiplied with the numerator $\sum_{j=0}^{\infty} (-1)^{j} |V_M|^\eta_j n^{-j}$. In total, we obtain

$$L_M^X(n) = \sum_{t=0}^{\infty} \sum_{j_0 \geq 0} \cdots \sum_{j_{t-1} \geq 0} (-1)^{t+\sum_{i=0}^{j_t} |V_M|^\eta_j} \sum_{j_t \geq 0} |V_M|^\eta_{j_t} \cdots |V_M|^\eta_{j_0} n^{-\text{id}(M)} - \sum_{i=0}^{t} j_i \quad (6.3)$$

The Combinatorial Meaning and Order of Magnitude of $C^X_{M,N}$

The expression (6.3) is a bit complicated, but it presents $L_M^X(n)$ as a sum (with coefficients $\pm n^{-t}$) of terms with a combinatorial interpretation: the term $|V_M|^\eta_j \cdot |E_M|^\eta_{j_1} \cdots |E_M|^\eta_{j_t}$ counts $(t+1)$-tuples of $\eta$-subordinate permutations. The crux of the matter is that this interpretation allows us to attribute each tuple to a specific subgroup $N \in [M,J]_X$. This is done as follows.

Let $(\sigma_0, \sigma_1, \ldots, \sigma_t)$ be a $(t+1)$-tuple of permutations such that $\sigma_0 \in \text{Sym}_\eta(V_M)$ and $\sigma_1, \ldots, \sigma_t \in \text{Sym}_\eta(E_M) \setminus \{\text{id}\}$ (we exclude id $\in \text{Sym}(E_M)$, which is the only permutation counted in $|E_M|^\eta_j$). Consider the graph $\Gamma = \Gamma_X(M)/(\sigma_0, \ldots, \sigma_t)$, 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 $e, \sigma_i(e)$ ($e \in E_M, 1 \leq i \leq t$). Since $\Gamma$ is obtained from $\Gamma_X(M)$ by identification of elements with the same $\eta$-image, $\eta$ induces a well defined morphism $\Gamma \rightarrow \Gamma_X(J)$. Thus, every closed path in $\Gamma$ projects to a path in $\Gamma_X(J)$, giving $\pi_1^\eta(\Gamma) \leq \pi_1^\eta(\Gamma_X(J)) = J$. We denote $N = N_{\sigma_0, \sigma_1, \ldots, \sigma_t} = \pi_1^\eta(\Gamma)$. As usual (see Figures 2.2. 2.3), we can perform Stallings folding on $\Gamma$ until we obtain the core graph corresponding to $N$, $\Gamma_X(N)$. Obviously we have $M \leq X N$, and by Claim 2.4(3) also $N \leq X J$.

[14]For the definition of the quotient of a graph by identifications of vertices see the discussion preceding Figure 2.3. 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 2.1), identifying a pair of edges is equivalent to identifying the pair of origins and/or the pair of termini.
Thus, we always have \( N = N_{\sigma_0, \sigma_1, \ldots, \sigma_t} \in [M, J]_X \). To summarize the situation:

\[
\Gamma_X (M) \xrightarrow{\eta_{M \to N}} \Gamma = \Gamma_X (M)/\langle \sigma_0, \ldots, \sigma_t \rangle \xrightarrow{\text{folding}} \Gamma_X (N) \xrightarrow{\eta_{N \to J}} \Gamma_X (J) \quad (6.4)
\]

Our next move is to rearrange (6.3) according to the intermediate subgroups \( N \in [M, J]_X \) which correspond to the tuples counted in it. For any \( N \in [M, J]_X \) we denote by \( \mathcal{T}_{M,N,J}^X \) the set of tuples \( (\sigma_0, \sigma_1, \ldots, \sigma_t) \) such that \( N_{\sigma_0, \sigma_1, \ldots, \sigma_t} = N \), i.e.

\[
\mathcal{T}_{M,N,J}^X = \left\{ (\sigma_0, \sigma_1, \ldots, \sigma_t) \mid t \in \mathbb{N}, \sigma_0 \in \text{Sym}_\eta (V_M) \right. \\
\left. \sigma_1, \ldots, \sigma_t \in \text{Sym}_\eta (E_M) \setminus \{ \text{id} \} \right\} .
\]

The terms \( (6.3) \) which correspond to a fixed \( N \in [M, J]_X \) thus sum to

\[
\tilde{C}_{M,J}^X (N) = \sum_{(\sigma_0, \sigma_1, \ldots, \sigma_t) \in \mathcal{T}_{M,N,J}^X} (\frac{1}{\pi t} \frac{1}{\| \sigma_t \|}) (-1)^{t + \sum_{i=0}^t \| \sigma_i \|}, \quad (6.5)
\]

and (6.3) becomes

\[
L_{M,J}^X = \sum_{N \in [M, J]_X} \tilde{C}_{M,J}^X (N) \quad (6.6)
\]

The equation (6.6) looks much like (6.1), with \( \tilde{C}_{M,J}^X (N) \) playing the role of \( C^X_{M,N} \). In order to establish equality between the latter two, we must show that \( \tilde{C}_{M,J}^X (N) \) does not depend on \( J \). Fortunately, this is not hard: it turns out that

\[
\tilde{C}_{M,J}^X (N) = \tilde{C}_{M,N}^X (N) \quad (\forall N \in [M, J]_X), \quad (6.7)
\]

and the r.h.s. is, of course, independent of \( J \). This equality follows from \( \mathcal{T}_{M,N,J}^X = \mathcal{T}_{M,N}^X \), which we now justify. The only appearance \( J \) makes in the definition of \( \mathcal{T}_{M,N,J}^X \) is inside \( \eta = \eta_{M \to J} \), which is to be \( \sigma_i \)-invariant (for \( 0 \leq i \leq n \)), i.e., \( \sigma_i \) must satisfy \( \eta_{M \to J} \circ \sigma_i = \eta_{M \to J} \). If \( (\sigma_0, \ldots, \sigma_t) \in \mathcal{T}_{M,N,J}^X \) then \( \eta_{M \to N} \circ \sigma_i = \eta_{M \to N} \) follows from the fact that \( \Gamma_X (N) \) is a quotient of \( \Gamma_X (M)/\langle \sigma_i \rangle \). On the other hand, if \( (\sigma_0, \ldots, \sigma_t) \in \mathcal{T}_{M,N}^X \) then we have \( \eta_{M \to N} \circ \sigma_i = \eta_{M \to N} \), hence also (see (6.4))

\[
\eta_{M \to J} \circ \sigma_i = \eta_{N \to J} \circ \eta_{M \to N} \circ \sigma_i = \eta_{N \to J} \circ \eta_{M \to N} = \eta_{M \to J}.
\]

Writing \( \tilde{C}_{M,N}^X \defeq \tilde{C}_{M,N}^X (N) \), we have by (6.1), (6.6), and (6.7)

\[
C^X \ast \zeta^X = L^X = \tilde{C}_{M,N}^X \ast \zeta^X
\]
which shows that $O^X = \tilde{C}^X$, as desired.

We approach the endgame. Let $(\sigma_0, \sigma_1, \ldots, \sigma_t) \in T^X_{M,N,J}$ 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 $1 \leq i \leq t$ and edges $e, e'$ whose origins are $v$ and $v'$, respectively. Since $P$ can clearly be obtained by identifying, we have $\|P\| \leq \sum_{i=0}^t \|\sigma_i\|$. Since $(\sigma_0, \sigma_1, \ldots, \sigma_t) \in T^X_{M,N,J}$ we have $\pi_X (\Gamma_X (H)/P) = N$, and thus by (2.5) we obtain

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

From (6.5) (recall that $\tilde{C}^X_{M,N} (N) = \tilde{C}^X_{M,N} = C^{X}_{M,N}$) we now have

$$C^{X}_{M,N} (n) = O \left( \frac{1}{n^{\text{rk}(M)} + \rho_X (M,N)} \right),$$

and Proposition 6.1 is proven.

7 Primitive Words in the Profinite Topology

Theorem 1.2 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 $\widehat{F}_k$ denote the profinite completion of the free group $F_k$. A basis of $\widehat{F}_k$ is a set $S \subset \widehat{F}_k$ such that every map from $S$ to a profinite group $G$ admits a unique extension to a continuous homomorphism $\widehat{F}_k \to G$. It is a standard fact that $F_k$ is embedded in $\widehat{F}_k$, and that every basis of $F_k$ is also a basis of $\widehat{F}_k$ (see for example [Wil98]). An element of $\widehat{F}_k$ is called primitive if it belongs to a basis of $\widehat{F}_k$.

It is natural to ask whether an element of $F_k$, which is primitive in $\widehat{F}_k$, is already primitive in $F_k$. In fact, this was conjectured by Gelander and Lubotzky, independently. Theorem 1.2 yields a positive answer, as follows. An element $w \in \widehat{F}_k$ is said to be measure preserving if for any finite group $G$, and a uniformly distributed random (continuous) homomorphism $\hat{\alpha} : \widehat{F}_k \to G$. By the natural correspondence $\text{Hom}_{\text{cont}} (\widehat{F}_k, G) \cong \text{Hom} (F_k, G)$, an element of $\widehat{F}_k$ is measure preserving w.r.t. $F_k$ if it is so w.r.t. $\widehat{F}_k$. As in $F_k$, a primitive element of $\widehat{F}_k$ is easily seen to be measure preserving. Theorem 1.2 therefore implies that if $w \in F_k$ is primitive in $\widehat{F}_k$, then it is also primitive in $F_k$. In other words:

**Corollary 7.1.** Let $P$ denote the set of primitive elements of $F_k$, and let $\hat{P}$ denote the set of primitive elements of $\widehat{F}_k$. Then

$$P = \hat{P} \cap F_k.$$
As \( \hat{P} \) is a closed set in \( \hat{\text{F}}_k \), this immediately implies Corollary 1.3, which states that \( P \) is closed in the profinite topology.

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

- \( w_1 \overset{A}{\sim} w_2 \) if \( w_1 \) and \( w_2 \) belong to the same \( \text{Aut} \text{F}_k \)-orbit.
- \( w_1 \overset{B}{\sim} w_2 \) if \( w_1 \) and \( w_2 \) belong to the same \( \overline{\text{Aut} \text{F}_k} \)-orbit (where \( \overline{\text{Aut} \text{F}_k} \) is the closure of \( \text{Aut} \text{F}_k \) in \( \text{Aut} \hat{\text{F}}_k \)).
- \( w_1 \overset{C}{\sim} w_2 \) if \( w_1 \) and \( w_2 \) belong to the same \( \text{Aut} \hat{\text{F}}_k \)-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 \( ev_{w_1}, ev_{w_2} : \text{Epi}(F_k, G) \rightarrow G \) 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 \( w_1 \overset{\sim}{\sim} w_2 \) 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 \( \hat{\text{F}}_k \)).

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 \( \text{F}_k \). Showing that \((A) \iff (B)\) would imply that \( \text{Aut} \text{F}_k \)-orbits in \( \text{F}_k \) are closed in the profinite topology, and the stronger statement \((A) \iff (C)\) would imply that words which lie in different \( \text{Aut} \text{F}_k \)-orbits can be told apart using statistical methods.

The analysis which is carried out in this paper does not suffice to answer 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 \( \text{Aut} \text{F}_2 \)-orbits, as \( w_2 \in \text{F}_2^2 \), but \( w_1 \not\in \text{F}_2^2 \), but induce the same distribution on \( S_n \) for every \( n \): their images under a random homomorphism are a product of a random permutation \( (\sigma) \) and a random element in its conjugacy class \((\tau \sigma \tau^{-1} \text{ for } w_1, \text{ and } \tau \sigma^{-1} \tau^{-1} \text{ for } w_2)\).

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 \( \text{Aut} G \)-orbit in \( G \)? Whitehead’s algorithm [Whi36a, Whi36b] gives a concrete solution when \( G = \text{F}_k \). Showing that \((A) \iff (B)\) would provide an alternative decision procedure for \( \text{F}_k \).

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

\[ ^{15} \text{In [AV71], for example, the authors indeed ask whether } (C') \Rightarrow (A). \]
(1) $G$ is finitely presented

(2) $\text{Aut } G$ is finitely generated

(3) $\text{Aut } G$-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 $(A) \Leftrightarrow (B)$.

8 Open Problems

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

- Section 7 shows how the questions about primitive elements can be extended to all $\text{Aut } F_k$-orbits in $F_k$ (is it true that $(A) \Leftrightarrow (B)$, and even the stronger equivalence $(A) \Leftrightarrow (C)$?). More generally, can statistical properties tell apart two subgroups $H_1, H_2 \leq F_k$ which belong to distinct $\text{Aut } F_k$-orbits? This would be a further generalization of Theorem 1.2.

- It is also interesting to consider words which are measure preserving w.r.t. other types of groups. For instance, does Theorem 1.2 still hold if we replace “finite groups” by “compact Lie groups”? 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. $PGL_n(q)$)?

- Is it true that

$$[H, \infty] \leq \bigcup_{X \text{ is a basis of } F_k} [H, \infty]_X$$

and under which assumptions does the following hold

$$[H, \infty]_{alg} = \bigcap_{X \text{ is a basis of } F_k} [H, \infty]_X$$

(see Remark 3.3)?

- Finally, there are interesting questions regarding the distribution of the primitivity rank. It is known that primitive words are rare (in the sense that the proportion of primitive words among all words of length $t$ tends to 0 as $t \to \infty$, [BMS02]). Is it true, for example, that being of primitivity rank $k$ ($\pi(w) = k$) is generic, in the sense that the proportion of such words tends to 1 as $t \to \infty$? In addition, almost primitive words are words $w \in F_k$ such that $[(w), \infty]_{alg} = \langle w \rangle, F_k \rangle$ [FRSS98]. What is the proportion of these words inside the class of words with primitivity rank $k$?
Acknowledgments

It is a pleasure to thank our advisors Nati Linial and Alex Lubotzky for their support, encouragement and useful comments. We are also grateful to Aner Shalev for supporting this research and for his valuable suggestions. We would also like to thank Tsachik Gelander, Chen Meiri and Iddo Samet for their beneficial comments. We have benefited much from the mathematical open source community, and in particular from GAP [GAP08], and its free group algorithms package written by Christian Sievers.

Glossary

| Notation | Definition | Reference | Remarks |
|----------|------------|-----------|---------|
| $H \leq_{fg} F_k$ | finitely generated | | |
| $H \leq J$ | free factor | | |
| $H \leq_{alg} J$ | algebraic extension | Definition 1.7 | |
| $H \leq X J$ | $H$ $X$-covers $J$ | Definition 2.3 | $H \rightarrow X J$ in [Pud11] |
| $[H, J]_x$ | $\{L \mid H \leq L \leq J\}$ | | $\leq$ is either one of $\leq, \leq_{alg}$ or $\leq X$ (standing for $\leq \rightarrow X$) |
| $[H, J]_x$ | $\{L \mid H \leq L \leq J\}$ | | |
| $[H, \infty)_{x}$ | $\{L \mid H \leq L\}$ | | |
| $H, \infty)_{x}$ | the $X$-fringe of $H$ | $O_X(H)$ in [MVW07] | |
| $H, \infty)_{alg}$ | algebraic extensions of $H$ | $AE(H)$ in [MVW07] | |
| $\pi(H)$ | primitivity rank of $H$ | Definition 1.4 | |
| Crit($H$) | $H$-critical groups | | |
| $X \Gamma(H)$ | $X$-labeled core graph of $H$ | | |
| $\rho_X(H, J)$ | $X$-distance | Definition 2.0 | $H \leq X J$ |
| $\eta^{X}_{H \rightarrow J}$ | the morphism $\Gamma_X(H) \rightarrow \Gamma_X(J)$ | Claim 2.2 | $H \leq J$ |
| $\alpha_{J, n}$ | a uniformly chosen random homomorphism in $\text{Hom}(J, S_n)$ | | $J \leq_{fg} F_k$ |
| $\Phi_{H, J}(n)$ | the expected number of common fixed points of $\alpha_{J, n}(H)$ | (1.1) | $H \leq J$ |

References

[Abc06] Miklós Abert, On the probability of satisfying a word in a group, Journal of Group Theory 9 (2006), 685–694.
[AL02] A. Amit and N. Linial, *Random graph coverings i: General theory and graph connectivity*, Combinatorica 22 (2002), no. 1, 1–18.

[AV11] A. Amit and Uzi Vishne, *Characters and solutions to equations in finite groups*, Journal of Algebra and Its Applications 10 (2011), no. 4, 675–686.

[BMS02] A.V. Borovik, A.G. Myasnikov, and V. Shpilrain, *Measuring sets in infinite groups*, Contemporary Mathematics 298 (2002), 21–42.

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

[FRSS98] B. Fine, G. Rosenberger, D. Spellman, and M. Stille, *Test, generic and almost primitive elements in free groups*, Mat. Contemp 14 (1998), 45–59.

[GAP08] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.4.12*, 2008.

[GS09] Shelly Garion and Aner Shalev, *Commutator maps, measure preservation, and t-systems*, Trans. Amer. Math. Soc. 361 (2009), no. 9, 4631–4651.

[KM02] I. Kapovich and A. Myasnikov, *Stallings foldings and subgroups of free groups*, Journal of Algebra 248 (2002), no. 2, 608–668.

[LP10] Nati Linial and Doron Puder, *Words maps and spectra of random graph lifts*, Random Structures and Algorithms 37 (2010), no. 1, 100–135.

[LS08] Michael Larsen and Aner Shalev, *Characters of symmetric groups: sharp bounds and applications*, Inventiones mathematicae 174 (2008), no. 3, 645–687.

[LS09] Michael Larsen and Aner Shalev, *Words maps and waring type problems*, J. Amer. Math. Soc. 22 (2009), no. 2, 437–466.

[MVW07] Alexei Miasnikov, Enric Ventura, and Pascal Weil, *Algebraic extensions in free groups*, Geometric group theory (G.N. Arzhantseva, L. Bartholdi, J. Burillo, and E. Ventura, eds.), Trends Math., Birkhauser, 2007, pp. 225–253.

[Nic94] Alexandru Nica, *On the number of cycles of given length of a free word in several random permutations*, Random Structures and Algorithms 5 (1994), no. 5, 703–730.

[PP12] Ori Parzanchevski and Doron Puder, *A counter example to a conjecture of miasnikov, ventura and weil*, preprint, 2012.
[Pud11] D. Puder, *Primitive words, free factors and measure preservation*, submitted, 2011.

[Seg09] Dan Segal, *Words: notes on verbal width in groups*, London Mathematical Society, Lecture note Series 361, Cambridge University Press, Cambridge, 2009.

[Sha09] Aner Shalev, *Words maps, conjugacy classes, and a non-commutative waring-type theorem*, Annals of Math. 170 (2009), 1383–1416.

[Sta83] John R. Stallings, *Topology of finite graphs*, Inventiones mathematicae 71 (1983), no. 3, 551–565.

[Sta97] R.P. Stanley, *Enumerative combinatorics, vol. 1 (cambridge studies in advanced mathematics 49)*, Cambridge University Press, Cambridge, 1997.

[vLW01] J.H. van Lint and R.M. Wilson, *A course in combinatorics*, Cambridge Univ Pr, 2001.

[Whi36a] J.H.C. Whitehead, *On certain sets of elements in a free group*, Proc. London Math. Soc. 41 (1936), 48–56.

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

[Wil98] John S. Wilson, *Profinite groups*, Clarendon Press, Oxford, 1998.