HAL open science

Statistical properties of subgroups of free groups

Frédérique Bassino, Armando Martino, Cyril Nicaud, Enric Ventura, Pascal Weil

To cite this version:

Frédérique Bassino, Armando Martino, Cyril Nicaud, Enric Ventura, Pascal Weil. Statistical properties of subgroups of free groups. Random Structures and Algorithms, 2013, 42, pp.349-373. 10.1002/rsa.20407 . hal-00450218v3

HAL Id: hal-00450218

https://hal.science/hal-00450218v3

Submitted on 22 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical properties of subgroups of free groups*

Frédérique Bassino, Frederique.Bassino@lipn.univ-paris13.fr
Université Paris 13, LIPN†

Armando Martino, A.Martino@upc.soton.ac.uk
University of Southampton‡

Cyril Nicaud, nicaud@univ-mlv.fr
Université Paris-Est, LIGM and CNRS§

Enric Ventura, enric.ventura@upc.edu
Universitat Politecnica de Catalunya¶

Pascal Weil, pascal.weil@labri.fr
Univ. Bordeaux, LaBRI, UMR 5800, F-33400 Talence, France
CNRS, LaBRI, UMR 5800, F-33400 Talence, France‖

*All authors benefitted from the support of the French-Spanish program PICASSO (project ACI-HF2006-0239). The first and third authors were supported by the ANR (projects BLAN 07-2_195422 GAMMA and ANR-2010-BLAN-0204 MAGNUM). The second and fourth authors were supported by MEC (Spain) and the EFRD (EC) through project number MTM2008-01550. The last author was supported by the ESF program AutoMathA and by the ANR (project ANR-2010-BLAN-0202 FREC).

†LIPN, Université Paris 13 and CNRS, Institut Galilée, 99, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France

‡School of Mathematics, University of Southampton, University Road, Southampton SO17 1BJ, United Kingdom

§Institut Gaspard Monge, Université Paris-Est, 77454 Marne-la-Vallée Cedex 2, France

¶Universitat Politecnica de Catalunya, Av. Bases de Manresa 61-73, 08240 Manresa, Catalunya, Spain

‖LaBRI, Univ. Bordeaux, 351 cours de la Libération, 33400 Talence, France.

Abstract

The usual way to investigate the statistical properties of finitely generated subgroups of free groups, and of finite presentations of groups, is based on the so-called word-based distribution: subgroups are generated (finite presentations are determined) by randomly chosen k-tuples of reduced words, whose maximal length is allowed to tend to infinity. In this paper we adopt a different, though equally natural point of view: we investigate the statistical properties of the same objects, but with respect to the so-called graph-based distribution, recently introduced by Bassino, Nicaud and Weil. Here, subgroups (and finite presentations) are determined by randomly chosen Stallings graphs whose number of vertices tends to infinity.

Our results show that these two distributions behave quite differently from each other, shedding a new light on which properties of finitely generated subgroups can be considered *frequent* or *rare*. For example, we show that malnormal subgroups of a free group are negligible in the graph-based distribution, while they are exponentially generic in the word-based distribution. Quite surprisingly, a random finite presentation generically presents the trivial group in this new distribution, while in the classical one it is known to generically present an infinite hyperbolic group.

Keywords: subgroups of free groups, finite group presentations, statistical properties, Stallings graphs, partial injections, malnormality

MSC: 20E05, 05A15, 20F69

1 Introduction

Statistical properties of elements and subgroups of free groups have evoked much interest in recent years, especially after Gromov's famous claim [14, 0.2.A] that "most" groups were hyperbolic, which led to precise statements and proofs by Ol'shanskiĭ [28] and by Champetier [8, 9]. Shortly thereafter, Ol'shanskiĭ and Arzhantseva [2, 1] pursued the study of the statistical properties of finite presentations of groups, that is, largely, of finitely generated normal subgroups of free groups. We refer the reader to the survey by Ollivier [27] for more details.

This interest encountered another historical trend in combinatorial group theory, namely the consideration of algorithmic problems, which leads naturally to an interest in the evaluation of the complexity of these algorithms (e.g. [5, 21, 29]) and in enumeration problems.

The search for innovative group-based cryptographic systems (see [25] for instance) only reinforced the study of complexity questions, and focused it on the investigation of the statistical properties of finitely generated subgroups of free groups, notably via the notion of generic complexity (see [18, 15]).

The usual method to approach statistical properties is to enumerate the objects under consideration, or more precisely, *representatives* for these objects, in a stratified way. For instance, if we wish to investigate k-generated subgroups of F_r (resp. finitely presented groups with r generators and k relators), we

proceed by enumerating lists of k-tuples of generators (resp. relators) over a fixed alphabet of r letters, so that at level n one has enumerated all such k-tuples whose elements have length at most n. In the situation we will consider, there are only finitely many objects of a given level n and it makes sense to ask what proportion of level n objects satisfy a given property. This gives us a number p_n between 0 and 1 for each n, associated to the given property and one can ask whether this sequence has a well defined limit. If the limit exists and is equal to 1, we would say that the property is *generic*, and take this to mean that most objects satisfy the property. At the other extreme, if the limit of the p_n equals 0, we would say that the property in question is *negligible* and conclude that it is rarely encountered amongst our objects.

A crucial observation, which is well worth mentioning in view of the intuitive weight carried by expressions such as *most objects* or *rarely encountered*, is that genericity and negligibility depend essentially on the choice of the stratification: different stratifications of the same objects, say finitely generated subgroups of free groups, will bring to light different insights on the statistical behavior of these objects. Concretely, different properties will appear to be generic or negligible.

Up to recently (namely the publication of [3]), the literature was unanimous in adopting the representation of finitely generated subgroups of free groups by k-tuples of generators, stratified by their maximal length – which we call the word-based distribution.

It is the purpose of this paper to question this unanimity. The basic idea is that there exists another very natural representation of finitely generated subgroups of free groups, by their Stallings graph ([32], see Section 2.1). Stratifying finitely generated subgroups by the size (number of vertices) of their Stallings graph – what we call the graph-based distribution – indeed sheds a different light on which properties of subgroups are frequent or rare. One of our main results is that malnormality and purity, which are generic in the word-based distribution, are negligible in the graph-based distribution (Section 4).

We also exhibit a property of finitely generated subgroups of F_r that is negligible in the word-based distribution and that has a non-zero, non-one asymptotic probability (namely e^{-r}) in the graph-based distribution (Section 5).

Finally we explore the possibility of using the graph-based distribution to discuss the statistical properties of finitely presented groups. The results there are disappointing: it turns out that finitely presented groups are generically trivial in this distribution – quite differently from the word-based distribution in which they are known to be generically infinite and hyperbolic (Section 6).

Sections 2 and 3 are devoted to preliminaries on genericity and to a review of the main features of the word-based and the graph-based distributions for finitely generated subgroups of free groups.

2 Preliminaries

Here we summarize standard facts about the Stallings graphs of subgroups (in Section 2.1) and we review the notions of generic and negligible properties.

Throughout the paper, A denotes an alphabet, that is, a finite non-empty set and $F(A)$ denotes the free group over A. The elements of $F(A)$ are represented by the reduced words written using letters from A and their formal inverses $\{a^{-1} \mid a \in A\}$. If $r \geq 1$, we often use the notation F_r instead of $F(A)$, to indicate that A consists of r letters. Throughout the paper, we will in fact assume that $r \geq 2$.

We denote by $[n]$ $(n \geq 1)$ the set $\{1, \ldots, n\}$.

2.1 Subgroup graph representation

Each finitely generated subgroup of $F(A)$ can be represented uniquely by a finite graph of a particular type, by means of the technique known as *Stallings foldings* [32]. This representation has been used by many authors, frequently using combinatorial, graph-theoretic notations that slightly differ from those used by Stallings. It is this formalism that we also use, which can also be found in [34, 17, 33, 22]. The procedure of Stallings foldings is informally described at the end of this section.

An *A-graph* is defined to be a pair $\Gamma = (V, E)$ with $E \subseteq V \times A \times V$, such that

- if $(u, a, v), (u, a, v') \in E$, then $v = v'$;
- if $(u, a, v), (u', a, v) \in E$, then $u = u'$.

The elements of V are called the *vertices* of Γ and the elements of E are its *edges*. We say that Γ is *connected* if the underlying undirected graph is connected. If $v \in V$, we say that v is a *leaf* if v occurs at most once in (the list of triples defining) E and we say that Γ is *v-trim* if no vertex $w \neq v$ is a leaf. Finally we say that the pair (Γ, v) is *admissible* if Γ is a finite, v-trim and connected A-graph. Then it is known (see [32, 34, 17, 22]) that:

- Stallings associated with each finitely generated subgroup H of $F(A)$ a unique admissible pair of the form $(\Gamma, 1)$, which we call the *graphical representation* or the *Stallings graph* of H and write $\Gamma(H)$;
- every admissible pair $(\Gamma, 1)$ is the graphical representation of a unique finitely generated subgroup of $F(A)$;
- if $(\Gamma, 1)$ is the graphical representation of H and u is a reduced word, then $u \in H$ if and only if u labels a loop at 1 in Γ (by convention, an edge (u, a, v) can be read from u to v with label a, or from v to u with label a^{-1});
- if $(\Gamma, 1)$ is the graphical representation of H, then $\mathsf{rank}(H) = |E| - |V| + 1$;
- finitely generated subgroups H and K are conjugates if and only if the cyclic cores of $\Gamma(H)$ and $\Gamma(K)$ (obtained by repeatedly deleting leaves and the edges they are adjacent to) are equal.

We informally remind the readers of the computation of the graphical representation of a subgroup generated by a subset $B = \{u_1, \ldots, u_k\}$. It consists in building an $(A \sqcup A^{-1})$-graph, changing it into a A-graph, then reducing it using *foldings*. First build a vertex 1. Then, for every word u of length n in B, build a loop with label u from 1 to 1, adding $n-1$ vertices. Change every edge (u, a^{-1}, v) labeled by a letter of A^{-1} into an edge (v, a, u). Then iteratively identify the vertices v and w whenever there exists a vertex u and a letter $a \in A$ such that either both (u, a, v) and (u, a, w) or both (v, a, u) and (w, a, u) are edges in the graph (the corresponding two edges are *folded*, in Stallings' terminology).

The resulting graph Γ is such that $(\Gamma, 1)$ is admissible and, very much like in the (1-dimensional) reduction of words, it does not depend on the order used to perform the foldings.

2.2 Negligibility and genericity

Let S be a countable set, the disjoint union of finite sets S_n $(n \geq 0)$, and let $B_n = \bigcup_{i \leq n} S_i$. Typically in this paper, S will be the set of Stallings graphs, of partial injections, of reduced words or of k-tuples of reduced words, and S_n will be the set of elements of S of size n. A subset X of S is *negligible* (resp. *generic*) if the probability for an element of B_n to be in X, tends to 0 (resp. to 1) when n tends to infinity; that is, if $\lim_n \frac{|X \cap B_n|}{|B_n|} = 0$ (resp. $= 1$).

Naturally, the negligibility or the genericity of a subset X of S depends on the layering of S into the S_n. In particular, if X and its complement are both infinite, then an appropriate partition of S into finite subsets S_n will make X negligible, another will make it generic, and indeed, another will be such that $\lim_n \frac{|X \cap B_n|}{|B_n|} = p$ for any fixed $0 < p < 1$.

Thus, any discussion of negligibility or genericity must clearly specify the distribution that is considered, that is, the choice of the partition $(S_n)_n$.

2.2.1 Rate of convergence

In general, we may be interested in the speed of convergence of $\frac{|X \cap B_n|}{|B_n|}$ – towards 0 if X is negligible and towards 1 if it is generic. One reason is that a higher speed of convergence indicates a higher rate of confidence that a randomly chosen element of S of size n will miss X if X is negligible, or will be in X if X is generic, even for moderately large values of n.

If a class $\mathcal{F}$ of functions tending to 0 is closed under max (of two elements), we say that a subset X is *$\mathcal{F}$-negligible* if $\frac{|X \cap B_n|}{|B_n|} = \mathcal{O}(f(n))$ for some $f \in \mathcal{F}$. We also say that X is *$\mathcal{F}$-generic* if the complement of X is $\mathcal{F}$-negligible. Note that $\mathcal{F}$-negligible (resp. $\mathcal{F}$-generic) sets are closed under finite unions and intersections.

Much of the literature is concerned with *exponential* negligibility or genericity, namely $\mathcal{F}$-negligibility or genericity where $\mathcal{F}$ is the class of functions e^{-cn} $(c > 0)$.

2.2.2 Balls versus spheres

The definition of negligibility and genericity above is given in terms of the balls B_n: the sets of elements of size at most n. It is sometimes more expedient to reason in terms of the proportion of elements of X in the spheres S_n: let us say, within the ambit of this section, that a set X is S-negligible (resp. S-generic) if the ratio $\frac{|X \cap S_n|}{|S_n|}$ tends to 0 (resp. 1). The definition of $\mathcal{F}$-S-negligibility or $\mathcal{F}$-S-genericity is analogous. We verify in this section that (exponential) S-negligibility implies (exponential) negligibility. The same holds of course for genericity.

Proposition 2.1 *An S-negligible (resp. S-generic) set is also negligible (resp. generic).*

If the structures under consideration grow fast enough, so that $\lim \frac{B_n}{B_{2n}} = 0$, *then the same result holds for exponential negligibility and genericity.*

The proof of this statement relies on the following technical lemma.

Lemma 2.2 *Let* (a_n) *and* (b_n) *be increasing sequences of positive real numbers.*

(1) ***(Stolz-Cesàro theorem)*** *If* $\lim b_n = \infty$ *and* $\lim \frac{a_{n+1}-a_n}{b_{n+1}-b_n} = 0$, *then* $\lim \frac{a_n}{b_n} = 0$.

(2) If $(\frac{b_n}{b_{2n}})$ *and* $(\frac{a_{n+1}-a_n}{b_{n+1}-b_n})$ *converge to* 0 *exponentially fast and if* $a_n \leq b_n$ *for each* n, *then* $(\frac{a_n}{b_n})$ *converges to* 0 *exponentially fast as well.*

Proof. (1) Since $\lim \frac{a_{n+1}-a_n}{b_{n+1}-b_n} = 0$, for each $\varepsilon > 0$, there exists n_0 such that $a_{n+1} - a_n \leq \varepsilon(b_{n+1} - b_n)$ for all $n \geq n_0$. Summing these inequalities for all integers between $n-1$ and n_0, we find that $a_n - a_{n_0} \leq \varepsilon(b_n - b_{n_0})$ for all $n > n_0$. Dividing by b_n and using the fact that $\lim b_n = \infty$, we conclude that $\frac{a_n}{b_n} < 2\varepsilon$ for all large enough n.

(2) Our hypothesis is now that there exists $c > 0$ such that $a_{n+1} - a_n \leq e^{-cn}(b_{n+1} - b_n)$ for all $n \geq n_0$. Summing these inequalities for the integers between n and $2n-1$, we find that $a_{2n} - a_n \leq e^{-cn}(b_{2n} - b_n)$ for all $n \geq n_0$. We now divide both sides by b_{2n} and use the fact that $\frac{a_n}{b_{2n}} \leq \frac{b_n}{b_{2n}}$ and that this sequence converges to 0 exponentially fast to conclude that $(\frac{a_{2n}}{b_{2n}})_n$ converges to 0 exponentially fast. Summing instead for the integers between n and $2n$ and dividing by b_{2n+1} shows that $(\frac{a_{2n+1}}{b_{2n+1}})_n$ converges to 0 exponentially fast as well. □

Proof of Proposition 2.1. Let $X \subseteq S$, $a_n = |X \cap B_n|$ and $b_n = |B_n|$. Then $a_n - a_{n-1} = |X \cap S_n|$ and $b_n - b_{n-1} = |S_n|$. The statement on (exponential) negligibility now follows directly from Lemma 2.2. The statement on genericity follows as well, since generic sets are the complements of negligible sets. □

3 The word-based and the graph-based distributions

In order to discuss the distribution of finitely generated subgroups of F_r, we need to fix a representation of these subgroups by means of discrete structures. In this paper we consider two such structures: a subgroup can be given by a tuple of generators (reduced words in F_r), or by its Stallings graph (Section 2.1). In the first case, the size of the representation is the pair (k, n) where k is the number of generators and n their maximal length – or n if k is fixed; in the second case, the size of the representation is the number n of vertices of the Stallings graph. In either case, there are only finitely many subgroups of each size.

We first review the literature on the word-based and the graph-based distributions (Sections 3.1 and 3.2), and then start the discussion of negligible or generic properties of subgroups (Section 3.3).

3.1 The word-based distribution

The distribution usually found in the literature (e.g. [18, 15, 16]) is in fact a distribution on the k-tuples $\vec{h} = (h_1, \ldots, h_k)$ of reduced words of length at most n, where k is fixed and n is allowed to grow to infinity; one then considers the subgroup H generated by $\vec{h}$. We call this distribution *word-based*.

Let us first record three elementary facts, which can also be found in [15][1]. We denote by R_n the set of reduced words of length at most n.

Fact 3.1 $|R_n| = \dfrac{r}{r-1}(2r-1)^n \left(1 - \dfrac{1}{r(2r-1)^n}\right)$. □

Proof. The number of reduced words of length $i \geq 1$ is $2r(2r-1)^{i-1}$, so the cardinality of R_n is

$$\begin{aligned} |R_n| = 1 + \sum_{i=1}^{n} 2r(2r-1)^{i-1} &= 1 + 2r\frac{(2r-1)^n - 1}{2r-2} \\ &= 1 + \frac{r}{r-1}((2r-1)^n - 1) \\ &= \frac{r}{r-1}(2r-1)^n \left(1 - \frac{1}{r(2r-1)^n}\right). \end{aligned}$$

□

Fact 3.2 Let $0 < \alpha < 1$. Exponentially generically, a reduced word in R_n has length greater than αn. □

Proof. The proportion of words in R_n, of length less than or equal to αn, is

$$\begin{aligned} \frac{|R_{\lfloor \alpha n \rfloor}|}{|R_n|} = \frac{\frac{r}{r-1}(2r-1)^{\lfloor \alpha n \rfloor}(1+o(1))}{\frac{r}{r-1}(2r-1)^n(1+o(1))} &= (2r-1)^{\lfloor \alpha n \rfloor - n}(1+o(1)) \\ &\leq (2r-1)^{(\alpha-1)n}(1+o(1)). \end{aligned}$$

[1] We choose to reiterate the proofs of these results, because we feel that our presentation exhibits more clearly their combinatorial underpinnings.

Since $\alpha - 1 < 0$, it converges to 0 exponentially fast. $\square$

Let $\vec{h} = (h_1, \ldots, h_k)$ be a tuple of reduced words and let $\mu > 0$ such that $\min |h_i| > 2\lceil \mu \rceil$. We denote by $\mathsf{Pref}_\mu(\vec{h})$ the set of prefixes of length at most $\lceil \mu \rceil$ of the h_i and h_i^{-1}.

Fact 3.3 Let $0 < \lambda < \frac{1}{2}$. Exponentially generically, a k-tuple $(h_1, \ldots, h_k)$ of elements of R_n, is such that $\min |h_i| > 2\lceil \lambda n \rceil$ and the prefixes of the h_i and h_i^{-1} of length $\lceil \lambda n \rceil$ are pairwise distinct. $\square$

Proof. The complement in R_n^k of the set of k-tuples described in the statement is the union of the set Y of k-tuples $\vec{h}$ where $\min |h_i| \leq 2\lceil \lambda n \rceil$, and of the set Z of k-tuples where $\min |h_i| > 2\lceil \lambda n \rceil$ and the set of prefixes of length $\lceil \lambda n \rceil$ of the h_i and h_i^{-1} has at most $2k-1$ elements. Since $2\lambda < 1$, the set Y is exponentially negligible by Fact 3.2 and we now concentrate on Z.

For each integer $2\lceil \lambda n \rceil < m \leq n$, let Z_m be the set of k-tuples in Z, such that $\min |h_i| = m$. Then

$$\begin{aligned} |Z_m| &\leq (2r(2r-1)^{\lceil \lambda n \rceil - 1})^{2k-1}\ k(2k-1)\ (2r-1)^{k(m-2\lceil \lambda n \rceil)} \\ &\leq (2r)^{2k-1}\ k(2k-1)\ (2r-1)^{k(m-2)-\lceil \lambda n \rceil + 1}. \end{aligned}$$

Summing these inequalities for all $2\lceil \lambda n \rceil < m \leq n$, we find

$$|Z| \ \leq\ (2r)^{2k-1}\ k(2k-1)\ (2r-1)^{k(n-1)-\lceil \lambda n \rceil + 1}.$$

As a result, the proportion of k-tuples in Z is at most

$$\frac{(2r)^{2k-1}\ k(2k-1)\ (2r-1)^{k(n-1)-\lceil \lambda n \rceil + 1}}{\frac{r^k}{(r-1)^k}(2r-1)^{kn}(1+o(1))} \ \leq\ C\ (2r-1)^{-\lambda n}(1+o(1))$$

for some constant C depending only on k and r. Thus, this proportion converges to 0 exponentially fast. $\square$

Remark 3.4 The small cancellation property, a closely related statement relative to common factors located anywhere in the words h_i and h_i^{-1} (not just at their extremities), is discussed in Lemma 4.5, in a variant of Arzhantseva and Ol'shanskiĭ's result on cyclic words [2, Lemma 3]. $\square$

Let $0 < \alpha < 1$ and $0 < \lambda < \frac{\alpha}{2}$, and for each n, let $Y_{\alpha,\lambda,n,k}$ be the set of k-tuples $\vec{h} = (h_1, \ldots, h_k) \in R_n^k$ such that $\min |h_i| > \alpha n$ and the prefixes of the h_i and h_i^{-1} of length $\lceil \lambda n \rceil$ are pairwise distinct. Facts 3.2 and 3.3 show that the proportion of elements of R_n^k in $Y_{\alpha,\lambda,n,k}$ converges to 1 exponentially fast: in the search for exponentially generic properties of subgroups, we can restrict our attention to the tuples in $Y_{\alpha,\lambda,n,k}$ and to the subgroups they generate.

The following observation is the basis for our exponential genericity proofs, in the context of the word-based distribution.

Fact 3.5 Let α, λ satisfy $0 < 2\lambda < \alpha < 1$. If $\vec{h} \in Y_{\alpha,\lambda,n,k}$ and $H = \langle \vec{h} \rangle$, then $\Gamma(H)$ consists of two parts:

- the vertices at distance at most $\lceil \lambda n \rceil$ from the distinguished vertex and the edges connecting them: this forms a tree with vertex set $\mathsf{Pref}_{\lambda n}(\vec{h})$, and edges $u \to^a ua$ if $a \in A$ and $u, ua \in \mathsf{Pref}_{\lambda n}(\vec{h})$; this tree, which we call the *central part* of $\Gamma(H)$, has $2k$ leaves;

- and for each $1 \leq i \leq k$, where $h_i = p_i m_i s_i$ and $|p_i| = |s_i| = \lceil \lambda n \rceil$, a path labeled m_i from the vertex p_i to the vertex s_i^{-1} (both are in the central part); we call these paths the *outer loops*. □

This leads to the following results. Propositions 3.6 and 3.8 first appeared in a paper by Jitsukawa [15]. They are direct consequences of earlier analogous results (counting cyclic words instead of words) due to Arzhantseva and Ol'shanskiĭ [2, Lemma 3].

Proposition 3.6 *Exponentially generically, a k-tuple of elements of R_n generates a subgroup of rank k.*

Proof. Let α, λ satisfy $0 < 2\lambda < \alpha < 1$. As observed above, it suffices to show that if $\vec{h} \in Y_{\alpha,\lambda,n,k}$, then $H = \langle \vec{h} \rangle$ has rank k. In that case, using Fact 3.5, we find that $\Gamma(H)$ is formed of a central part and k outer loops.

The central part is a tree and like all trees, the number of its edges is 1 less than the number of its vertices. With the notation of Fact 3.5, the number of additional vertices (resp. edges) in the outer loops is $\sum_i (|m_i| - 1)$ (resp. $\sum_i |m_i|$). Therefore, in $\Gamma(H)$, we have $|E| - |V| + 1 = k$ and hence $\mathsf{rank}(H) = k$ (see Section 2.1). □

Proposition 3.7 *Exponentially generically, a k-tuple $\vec{h}$ and an k'-tuple $\vec{h}'$ of elements of R_n generate subgroups that are distinct, have trivial intersection, and are such that $\langle \vec{h}, \vec{h}' \rangle = \langle \vec{h} \rangle * \langle \vec{h}' \rangle$.*

Proof. Since the first k components of a $(k + k')$-tuple of elements of R_n are independent from the k' last components, and since such a $(k + k')$-tuple exponentially generically generates a subgroup of rank $k + k'$ (Proposition 3.6), we find that a k-tuple and an k'-tuple of elements of R_n exponentially generically generate their free product. This in turn implies the other properties. □

Proposition 3.7 shows that two k-tuples of elements of R_n exponentially generically generate distinct subgroups. Proposition 3.8 is a little more precise.

Proposition 3.8 *Let α, λ satisfy $0 < 2\lambda < \alpha < 1$. The k-tuples $\vec{h}$ and $\vec{h}'$ in $Y_{\alpha,\lambda,n,k}$ generate distinct subgroups, unless $\vec{h}' = (h_{\sigma(1)}^{\varepsilon_1}, \ldots, h_{\sigma(k)}^{\varepsilon_k})$ for some permutation σ of $[k]$ and for $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, +1\}$.*

Proof. If $\langle \vec{h} \rangle = \langle \vec{h}' \rangle$, then the graphs $\Gamma\langle \vec{h} \rangle$ and $\Gamma\langle \vec{h}' \rangle$ are equal. In particular, their central parts, formed by the vertices at distance at most $\lceil \lambda n \rceil$ from the distinguished vertex, coincide. Fact 3.5 shows that if $\vec{h} \in Y_{\alpha,\lambda,n,k}$, then $\Gamma\langle \vec{h} \rangle$

completely determines all the k-tuples $\vec{h}' \in Y_{\alpha,\lambda,n,k}$ such that $\Gamma\langle\vec{h}'\rangle = \Gamma\langle\vec{h}\rangle$, and that they coincide with $\vec{h}$ up to the order of the elements and the direction in which the outer loops are read. $\square$

Proposition 3.8 shows that, if we consider the class $S_{\alpha,\lambda,n,k}$ of subgroups generated by k-tuples in $Y_{\alpha,\lambda,n,k}$, then each subgroup occurs the same number of times, namely $2^k k!$. Randomly choosing a k-tuple in $Y_{\alpha,\lambda,n,k}$ yields therefore a random subgroup in $S_{\alpha,\lambda,n,k}$, and the proportion of these subgroups among all subgroups generated by a k-tuple of words of length at most n tends to 1 exponentially fast.

3.2 The graph-based distribution

The uniform distribution on the set of size n Stallings graphs was analyzed by Bassino, Nicaud and Weil [3]. Here we summarize the principles of this distribution and the features which will be used in this paper.

We already noted that in Stallings graphs, each letter labels a partial injection on the vertex set: in fact, a Stallings graph can be viewed as a collection $(f_a)_{a\in A}$ of partial injections on an n-element set, with a distinguished vertex, and such that the resulting graph (with an a-labeled edge from i to j if and only if $j = f_a(i)$) is connected and has no vertex of degree 1, except perhaps the distinguished vertex. We may even assume that the n-element set in question is $[n] = \{1, \ldots, n\}$, with 1 as the distinguished vertex, see [3, Section 1.2] for a precise justification. In particular, the automorphism group of an admissible pair $(\Gamma, 1)$ is always trivial.

One shows [3, Corollary 2.7] that the probability that an A-tuple $(f_a)_{a\in A}$ of partial injections on $[n]$ induces a Stallings graph tends to 1 as n tends to infinity, and the problem of randomly generating a Stallings graph then reduces (via an efficient rejection algorithm, see [3, Section 3]) to the problem of efficiently generating a random partial injection on $[n]$. This view of a Stallings graph as an A-tuple of partial injections on $[n]$ is central in our analysis.

The maximal orbits of a partial injection f (equivalently: the connected components of the function graph of f) can be of two kinds: cycles – where each element is both in the domain and in the range of f – and sequences. The size of each of these components is defined to be the number of vertices which they contain. It is this combinatorial view of partial injections – as a disjoint union of cycles and sequences –, which is at the heart of the random generation algorithm, obtained using the so-called recursive method [26, 12].

The distribution of sizes of components is studied in [3, Section 3], as well as the distribution of cycles vs. sequences among size k components. The random generation algorithm consists in drawing a size of component, say k, according to the relevant distribution; then drawing whether this size k component is a cycle or a sequence; and finally drawing a partial injection on the remaining $n - k$ elements [3, Section 3.1]. This results in a partial injection on an n-element set, and we need only add a random numbering (1 through n) of the elements of that set.

However complex the method may seem, it guarantees a uniform distribution among all size n partial injections, it is easy to implement and its average time complexity is linear (in the RAM model; it is $\mathcal{O}(n^2 \log n)$ under the bit-cost assumption) [3, Section 3.3].

To further discuss partial injections and other combinatorial structures, we use the notion of *exponential generating series*, written EGS. If $\mathcal{S}$ is a class of finite discrete structures such that there are finitely many $\mathcal{S}$-structures of each size, let S_n be the number of $\mathcal{S}$-structures of size n. The EGS of $\mathcal{S}$ is the formal power series $S(z) = \sum_{n \geq 0} \frac{S_n}{n!} z^n$.

Let $I(z) = \sum_{n \geq 0} \frac{I_n}{n!} z^n$ be the EGS of partial injections. Bassino, Nicaud and Weil show the following [3, Section 2.1 and Proposition 2.10].

Proposition 3.9 *The EGS $I(z)$ of partial injections satisfies the following*

$$I(z) = \frac{1}{1-z} \exp\left(\frac{z}{1-z}\right) \quad \textit{and} \quad \frac{I_n}{n!} = \frac{e^{-\frac{1}{2}}}{2\sqrt{\pi}} e^{2\sqrt{n}} n^{-\frac{1}{4}} (1 + o(1)).$$

This result is obtained by means of deep theorems from analytic combinatorics. The same methods can be used to study the asymptotic behavior of particular parameters, such as the number of sequences of a partial injection. This parameter is directly connected with the number of edges in the Stallings graph formed by the partial injections f_a ($a \in A$), which leads to the following result [3, Lemma 2.11 and Corollary 4.1].

Proposition 3.10 *The expected number of sequences in a randomly chosen partial injection of size n is asymptotically equivalent to $\sqrt{n}$.*

The expected rank of a randomly chosen size n subgroup of F_r is asymptotically equivalent to $(r-1)n - r\sqrt{n} + 1$.

3.3 Negligible and generic properties of subgroups

Thus, in the discussion of statistical properties of finitely generated subgroups of a (fixed) free group F_r, we have two distributions at our disposal. One, the word-based distribution, is governed by two parameters: the number of generators and their maximum length, the former fixed and the latter allowed to tend to infinity. The other, the graph-based distribution, is governed by a single parameter: the size of the Stallings graph.

We first observe that our discussion of the graph-based distribution (as well as the results in [3]) is in terms of spheres rather than balls: as we saw in the Section 2.2.2, the (exponential) negligibility or genericity results obtained in that setting are sufficient. In contrast, the existing literature on the word-based distribution is in terms of balls, as is our description in Section 3.1 above.

The graph-based as well as the word-based distribution allow the discussion of properties of subgroups (of subgroups of a fixed rank k in the word-based case). There is of course no reason why a property that is generic or negligible in one distribution should have the same frequency in the other.

Our two distributions are indeed very different. How different is illustrated in Figure 1, which shows a "random" size 200 Stallings graph and the Stallings

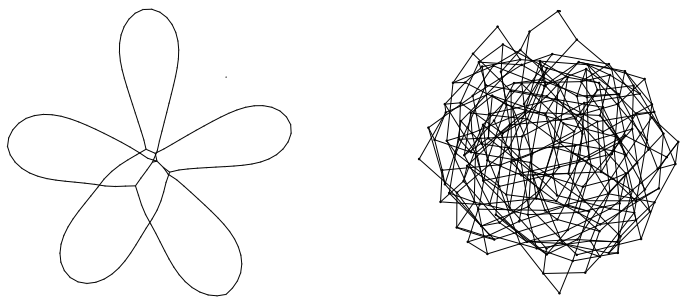

Figure 1: The Stallings graphs of two randomly generated subgroups of F_2. On the left, a subgroup generated by a random 5-tuple of words of length at most 40. On the right, a random Stallings graph of size 200. Only the shape of the graphs is depicted, vertices and edge labels and directions are not represented. The pictures have been generated by `neato`. Note that the scale (average distance between two vertices) is not the same on the two pictures.

graph of the subgroup of F_2 generated by a "random" 5-tuple of words of length at most 40 (which has close to 200 vertices). This figure provides the intuition to exhibit properties of subgroups that are negligible in one distribution and generic in the other.

It is not difficult to come up with such properties. It is the case, for instance, of the property *to have rank* ℓ, for a fixed integer $\ell \geq 1$. In the graph-based distribution, this property is negligible as a consequence of Proposition 3.10 (see [3, Corollary 4.2]). In contrast, it is exponentially generic in the word-based distribution *of* ℓ*-generated* subgroups, see Proposition 3.6. For the same reason, it is exponentially negligible in the word-based distribution *of* k*-generated* subgroups with $k \neq \ell$.

The properties of malnormality and purity, discussed in Section 4, provide more complex examples of this sort.

4 Malnormal and pure subgroups

Malnormality and purity are two important properties of subgroups. A subgroup H is *pure* if $x^n \in H$ and $n \neq 0$ implies $x \in H$. A pure subgroup is also called *closed under radicals* or *isolated*.

The subgroup H is *malnormal* if $H \cap H^g = 1$ for every $g \notin H$. Malnormal subgroups play an important role in the study of amalgamated products (e.g. [19, 6]) and in the characterization of their hyperbolicity [20]. The following is elementary from the definition.

Lemma 4.1 *If a subgroup is malnormal, then it is pure.*

Note that the converse statement does not hold: $\langle a, bab^{-1} \rangle$ is pure, yet not malnormal.

Both malnormality and purity have nice graphical characterizations, which imply that these properties are decidable for finitely generated subgroups of free groups. The result on malnormality is due to Kapovich and Myasnikov

[17] (following a decidability result in [4]), that on purity is due to Birget, Margolis, Meakin and Weil [5].

Proposition 4.2 *Let* $(\Gamma, 1)$ *be the graphical representation of a subgroup* H*.*

(1) H *is non-malnormal if and only if there exists a non-trival reduced word* u *and distinct vertices* $x \neq y$ *in* Γ *such that* u *labels loops at* x *and at* y*.*

(2) H *is non-pure if and only if there exists a non-trival reduced word* u*, an integer* $n \geq 2$ *and a vertex* x *in* Γ *such that* u^n *labels a loop at* x *but* u *does not.*

4.1 Genericity in the word-based distribution...

Jitsukawa shows that malnormality is a generic property in free groups [15, Theorem 4 and Lemma 6]. His arguments can be extended to show that it is exponentially generic.

Theorem 4.3 *Malnormality is exponentially generic in the word-based distribution.*

In view of Lemma 4.1, we also have the following result.

Corollary 4.4 *Purity is exponentially generic in the word-based distribution.*

We now proceed to prove Theorem 4.3. The proof relies on the two following lemmas, which provide an analogue of a small cancellation property for tuples of reduced words. These lemmas constitute a variant of the results of Gromov [14, 0.2.A] for tuples of cyclic words, proved also by Champetier [8, 9] and by Arzhantseva and Ol'shanskiĭ [2, Lemma 3].

Lemma 4.5 *Let* $0 < \beta < 1$*. The proportion of* k*-tuples* $\vec{h}$ *of reduced words in* F_r *of length at most* n*, such that one of the* h_i *contains two distinct occurrences of factors* v *and* w *of length at least* βn*, with* $v = w$ *or* $v = w^{-1}$*, converges to 0 exponentially fast.*

Lemma 4.6 *Let* $0 < \beta < 1$*. The proportion of* k*-tuples* $\vec{h}$ *of reduced words of length at most* n*, such that a word* v *of length at least* βn *has an occurrence in one of the* h_i *and* v *or* v^{-1} *has an occurrence in* h_j *for some* $j \neq i$*, converges to 0 exponentially fast.*

We can now prove Theorem 4.3.

Proof of Theorem 4.3. Let $0 < \alpha < 1$ and $0 < \lambda < \frac{\alpha}{4}$. Let $\vec{h} = (h_1, \ldots, h_k)$. Exponentially generically, we have $\min |h_i| > \alpha n$ and the prefixes of length $\lceil \lambda n \rceil$ of the h_i and h_i^{-1} are pairwise distinct (Facts 3.2 and 3.3). In addition, exponentially generically, no word of length at least $\frac{\alpha - 4\lambda}{2} n$ has distinct occurrences as a factor of the h_i and the h_i^{-1} (Lemmas 4.5 and 4.6).

Let us now assume that $\vec{h}$ satisfies all these properties. Then $\Gamma = \Gamma(\langle\vec{h}\rangle)$ is composed of a central part, which is a tree containing the distinguished vertex and all the vertices corresponding to the prefixes of the h_i and h_i^{-1} of length up to $\lceil\lambda n\rceil$, and of outer loops whose labels are factors of the h_i (or the h_i^{-1}, depending on the direction in which they are read), see Fact 3.5.

Any loop in Γ must visit the central part of Γ at least once, and run along at least one of the outer loops. Let us now assume that a word u labels two distinct loops in Γ. Up to conjugation of u, we can assume that the base point of the first loop is in the central part of Γ. Then u has a factor v of length $\lceil\alpha n\rceil - 2\lceil\lambda n\rceil$, which is a factor of some h_i or h_i^{-1}. The other occurrence of a loop labeled u reveals another path in Γ labeled v. This path may not be entirely in an outer loop, but if it is not, then it visits the central part of Γ only once, so it has a factor v' of length at least $\frac{\alpha-4\lambda}{2}n$ in an outer loop, and hence in one of the h_i or h_i^{-1}. This word v' has distinct occurrences as a factor of the h_i and the h_i^{-1}, a contradiction. □

4.2 ... and negligibility in the graph-based distribution

In contrast, we show that malnormality and purity are negligible in the graph-based distribution.

Theorem 4.7 *The probability that a random subgroup of size n is pure is $\mathcal{O}(n^{-\frac{r}{2}})$.*

By Lemma 4.1, this implies the following

Corollary 4.8 *The probability that a random subgroup of size n is malnormal is $\mathcal{O}(n^{-\frac{r}{2}})$.*

To prove Theorem 4.7, we observe that if H is a finitely generated subgroup of F_r and a cycle of length at least 2 in $\Gamma(H)$ is labeled by a power of some letter a, then H is not pure (Proposition 4.2). Therefore, if a subgroup is pure, then the partial injection determined by each letter in A has only sequences and length 1 cycles.

Thus Theorem 4.7 follows directly from the following proposition.

Proposition 4.9 *The probability that a size n partial injection has no cycle of length greater than or equal to 2 is asymptotically equivalent to $\frac{e}{\sqrt{n}}$.*

Our proof of Proposition 4.9 uses Hayman's theorem, discussed in Section 4.2.1 below.

Remark 4.10 There are many more reasons for a subgroup to fail to be pure, than those considered here. In terms of Proposition 4.2, we have considered only the words u that are equal to a letter of the alphabet. As a result, the probability of purity and that of malnormality are likely to be much smaller than the upper bounds given above. The open question here is whether purity and normality are exponentially negligible with respect to the graph-based distribution. □

4.2.1 H-admissible functions and Hayman's theorem

Hayman's theorem on the asymptotic behavior of the coefficients of certain power series requires a technical hypothesis called H-admissibility. Here we give only the technical definition and statement we will use, and we refer the readers to [11, Chapter VIII] for further details on this theorem and on saddlepoint asymptotics in general.

Let $f(z)$ be a function of the form $f(z) = e^{h(z)}$ that is analytic at the origin, with radius of convergence ρ. We denote by $[z^n]f(z)$ the coefficient of z^n in the power series development of f at the origin. Let

$$a(r) = rh'(r) \quad \text{and} \quad b(r) = r^2 h''(r) + rh'(r).$$

The function $f(z)$ is said to be *H-admissible* if there exists a function $\delta\colon]0,\rho[\longrightarrow]0,\pi[$ such that the following three conditions hold:

(H1) $\lim_{r\to\rho} b(r) = +\infty$.

(H2) Uniformly for $|\theta| \leq \delta(r)$

$$f(re^{i\theta}) \sim f(r)e^{i\theta a(r) - \frac{1}{2}\theta^2 b(r)} \quad \text{when } r \text{ tends to } \rho.$$

[That is, $f(re^{i\theta}) = f(r)e^{i\theta a(r)-\frac{1}{2}\theta^2 b(r)}(1+\gamma(r,\theta))$ with $|\gamma(r,\theta)| \leq \tilde{\gamma}(r)$ when $|\theta| \leq \delta(r)$ and $\lim_{r\to\rho}\tilde{\gamma}(r) = 0$.]

(H3) and uniformly for $\delta(r) \leq |\theta| \leq \pi$

$$f(re^{i\theta})\sqrt{b(r)} = o(f(r)) \qquad \text{when } r \text{ tends to } \rho.$$

Hayman's theorem [11, Theorem VIII.4] states the following.

Theorem 4.11 *Let $f(z) = e^{h(z)}$ be a H-admissible function with radius of convergence ρ and $\zeta = \zeta(n)$ be the unique solution in the interval $]0,\rho[$ of the saddlepoint equation*

$$\zeta\frac{f'(\zeta)}{f(\zeta)} = n.$$

Then

$$[z^n]f(z) = \frac{f(\zeta)}{\zeta^n\sqrt{2\pi b(\zeta)}}\left(1+o(1)\right).$$

where $b(z) = z^2h''(z) + zh'(z)$.

4.2.2 Proof of Proposition 4.9

Let $\mathcal{K}$ be the set of partial injections in which all the cycles have length 1 and let $\mathcal{J}$ the set of partial injections without any cycles (a subset of $\mathcal{K}$). The elements of $\mathcal{J}$ are known as *fragmented permutations*, see [11, Section II.4.2].

Let K_n and J_n be the number of size n elements of $\mathcal{K}$ and $\mathcal{J}$, and let $K(z)$ and $J(z)$ be the corresponding EGS. We want to show that $\frac{K_n}{I_n}$ is equivalent to $\frac{e}{\sqrt{n}}$.

The series $J(z)$ is studied in detail in [11, Example VIII.7, Proposition VIII.4]. There, it is shown in particular that $J(z)$ is H-admissible and that

$$J(z) = \exp\left(\frac{z}{1-z}\right) \quad \text{and} \quad \frac{J_n}{n!} = \frac{e^{-\frac{1}{2}}}{2\sqrt{\pi}} e^{2\sqrt{n}} n^{-\frac{3}{4}}(1+o(1)). \tag{1}$$

A partial injection in $\mathcal{K}$ consists of a set of length 1 cycles and a fragmented permutation. It follows that

$$K_n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} J_k,$$

so that

$$K(z) = \sum_{n=0}^{\infty} \frac{K_n}{n!} z^n = \left(\sum_{n=0}^{\infty} \frac{1}{n!} z^n\right)\left(\sum_{n=0}^{\infty} \frac{J_n}{n!} z^n\right) = e^z J(z) = \exp\left(z + \frac{z}{1-z}\right).$$

Now e^z is H-admissible: this can be verified directly, or by application of [11, Theorem VIII.5]. We already noted that $J(z)$ is H-admissible, and hence $K(z)$ is H-admissible as well, as the product of two H-admissible functions ([11, Theorem VIII.5] again).

The saddle-point equation $z\frac{K'(z)}{K(z)} = n$ (see Section 4.2.1) is

$$\frac{z(2-2z+z^2)}{(1-z)^2} = n,$$

$$i.e. \qquad z^3 - (n+2)z^2 + 2(n+1)z - n = 0.$$

Let $P_n(z)$ be the polynomial on the left hand side of this last equation. Examining the sign of the derivative of $P_n(z)$ on the interval $[0,1]$ and the values of P_n at 0 and 1, we find that P_n has a unique zero between 0 and 1, say ζ_n. Moreover

$$\zeta_n = 1 - \frac{1}{\sqrt{n}} + \frac{1}{2n} + \mathcal{O}\left(\frac{1}{n\sqrt{n}}\right). \tag{2}$$

This asymptotic development can be obtained using MAPLE, based on the application of the Cardan method to this degree 3 polynomial. We can also observe the following. Let $Q_n(z)$ be the polynomial defined by the identity

$$P_n(1-z) = 1 - z + (1-n)z^2 - z^3 = Q_n(z) - z^3.$$

The zero of $Q_n(z)$ in the interval $[0,1]$ is

$$\alpha_n = \frac{\sqrt{4n-3}-1}{2(n-1)} = \frac{1}{\sqrt{n}} - \frac{1}{2n} + \mathcal{O}(n^{-\frac{3}{2}})$$

and if $\beta_n = 1-\alpha_n$, we have $P_n(\beta_n) = -\alpha_n^3$, which is negative for n large enough.

Now let $\gamma_n = 1 - \frac{1}{\sqrt{n}} + \frac{1}{2n}$. Then

$$P_n(\gamma_n) = \frac{5}{4n} + \mathcal{O}(n^{-\frac{3}{2}}),$$

which is positive for n large enough. It follows that $\beta_n < \zeta_n < \gamma_n$, justifying the development in (2).

With the notation of Section 4.2.1, we also have

$$\begin{aligned} b(z) &= z^2 \frac{d^2}{dz^2} \log K(z) + z \frac{d}{dz} \log K(z) \\ &= \frac{z(2 - 2z + 3z^2 - z^3)}{(1-z)^3}, \\ \text{so that } b(\zeta_n) &= 2n^{3/2} + \mathcal{O}(n). \end{aligned}$$

Then we have

$$\begin{aligned} \log \zeta_n^n &= n \log \left(1 - \frac{1}{\sqrt{n}} + \frac{1}{2n} + \mathcal{O}\left(\frac{1}{n\sqrt{n}}\right)\right) = -\sqrt{n} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \\ \text{and } K(\zeta_n) &= \exp\left(\zeta_n + \frac{\zeta_n}{1-\zeta_n}\right) \sim e^{\frac{1}{2}} e^{\sqrt{n}}. \end{aligned}$$

By Theorem 4.11, we now have

$$[z^n]K(z) \sim \frac{K(\zeta_n)}{\zeta_n^n \sqrt{2\pi b(\zeta_n)}} \sim e^{\sqrt{n}} e^{\frac{1}{2}} e^{\sqrt{n}} \frac{1}{2\sqrt{\pi} n^{3/4}} \sim \frac{e^{\frac{1}{2}}}{2\sqrt{\pi}} n^{-3/4} e^{2\sqrt{n}}.$$

Proposition 4.9 follows since $[z^n]I(z) \sim \frac{e^{-\frac{1}{2}}}{2\sqrt{\pi}} e^{2\sqrt{n}} n^{-1/4}$ (Proposition 3.9) and hence, $\frac{K_n}{I_n} \sim \frac{e}{\sqrt{n}}$.

4.3 A remark on the Hanna Neumann conjecture

The Hanna Neumann Conjecture (HNC), recently established by Mineyev [23] after several decades of partial results, deals with the rank of the intersection of finitely generated subgroups of free groups (see also [24] for an alternative proof, purely in terms of groups and graphs). For convenience, let the *reduced rank* of a subgroup H, written $\widetilde{\mathsf{rk}}(H)$, be equal to

$$\widetilde{\mathsf{rk}}(H) = \max(0, \mathsf{rank}(H) - 1).$$

Mineyev's theorem states that, if H and K are finitely generated subgroups of F, then $\widetilde{\mathsf{rk}}(H \cap K) \leq \widetilde{\mathsf{rk}}(H)\widetilde{\mathsf{rk}}(K)$, as conjectured by Hanna Neumann. It also shows the stronger inequality conjectured by Burns [7], formerly known as the strengthened Hanna Neumann conjecture (SHNC):

$$\sum \widetilde{\mathsf{rk}}(H \cap K^g) \leq \widetilde{\mathsf{rk}}(H)\widetilde{\mathsf{rk}}(K),$$

where the sum runs over all subsets KgH (in $K\backslash F/H$) such that $H \cap K^g \neq 1$ and $K^g = g^{-1}Kg$.

We discuss here how our results show that the cases where this inequality is non-trivial are rare, in the sense that $\widetilde{\mathsf{rk}}(H \cap K)$ and $\sum \widetilde{\mathsf{rk}}(H \cap K^g)$ are generically equal to zero.

It was observed, initially by Stallings [32] and Gersten [13], that HNC and SHNC have natural interpretations in terms of Stallings graphs. If Γ is an A-labeled graph, let us denote by $\chi(\Gamma)$ the difference between the number of edges and the number of vertices of Γ: thus $\widetilde{\mathsf{rk}}(H) = \chi(\Gamma(H))$. Let $\Delta(H, K)$ be the graph obtained from $\Gamma(H)$ and $\Gamma(K)$ as follows: the vertices of $\Delta(H, K)$ are the pairs (u, v) such that u is a vertex of $\Gamma(H)$ and v is a vertex of $\Gamma(K)$; and the edges of $\Delta(H, K)$ are the triples $((u, v), a, (u', v'))$ such that (u, a, u') is an edge of $\Gamma(H)$ and (v, a, v') is an edge of $\Gamma(K)$.

Let Δ_1 be the connected component of $\Delta(H, K)$ containing $(1, 1)$ (where 1 denotes the origin of $\Gamma(H)$ and of $\Gamma(K)$), and let Δ_2 be the union of the connected components of $\Delta(H, K)$ which are not trees. Then HNC holds for H and K if and only if $\chi(\Delta_1) \leq \widetilde{\mathsf{rk}}(H)\widetilde{\mathsf{rk}}(K)$, and SHNC holds for H and K if and only if $\chi(\Delta_2) \leq \widetilde{\mathsf{rk}}(H)\widetilde{\mathsf{rk}}(K)$.

Now observe (as in Proposition 3.7) that a randomly chosen $(k+\ell)$-tuple of elements of R_n is composed of the juxtaposition of a randomly chosen k-tuple and a randomly chosen ℓ-tuple. Exponentially generically, such a k-tuple $\vec{h}$ and ℓ-tuple $\vec{h}'$ generate subgroups with trivial intersection: in particular, HNC holds exponentially generically in the word-based distribution.

In fact, with the same ideas as in the proof of Theorem 4.3, exponentially generically, there is no loop in $\Gamma(\langle\vec{h}\rangle)$ with an occurrence as a loop in $\Gamma(\langle\vec{h}'\rangle)$. Therefore SHNC holds exponentially generically.

As mentioned above, both HNC and SHNC are now known to hold, but it seems interesting to point out that exponentially generically, they hold for trivial reasons.

5 An intermediate property

In this section, we discuss an *intermediate* property of subgroups, that is a property such that the proportion of subgroups of size n with this property has a limit which is neither 0 nor 1 (respectively the negligible and the generic cases).

Theorem 5.1 *The probability that a random size n subgroup of F_r intersects trivially the conjugacy classes of the generators tends to e^{-r} when n tends to infinity.*

The discussion of this property is included here because we do not know many examples of such intermediate properties. Unfortunately, the property in question is geometric in the sense that it depends on the combinatorial parameters of the Stallings graph of the subgroup, and is not preserved under the automorphisms of F_r. It would be interesting to exhibit such a property that would be algebraic (preserved under automorphisms). One might think for instance of the property of avoiding the conjugacy classes of all the elements of some basis of F_r, or the property of avoiding all primitive words.

Remark 5.2 The property described in Theorem 5.1 is exponentially negligible in the word-based distribution. Indeed, if $\vec{h}$ is a k-tuple of reduced words of

length at most n, then $\Gamma(\langle\vec{h}\rangle)$ has exponentially generically k loops of length at least $\frac{n}{2}$ and no loop of length 1 (see the discussion in Section 3.1 with $\alpha = \frac{3}{4}$ and $\lambda = \frac{1}{8}$). $\square$

The rest of this section is devoted to the proof of Theorem 5.1. It is easily verified that a subgroup H contains a conjugate of letter $a \in A$ if and only if a labels a loop at some vertex of $\Gamma(H)$, that is, if and only if the corresponding partial injection has some fixpoint. Since the drawing of the partial injections corresponding to the different letters is independent, the theorem follows directly from the following proposition.

Proposition 5.3 *The probability that a size n partial injection has no fixpoint tends to $\frac{1}{e}$ when n tends to infinity.*

Remark 5.4 Note that $\frac{1}{e}$ is also the limit of the probability that a size n permutation has no fixpoint (a so-called *derangement*, see [10]). $\square$

Our proof of Proposition 5.3 again uses Hayman's theorem (Section 4.2.1). We also need the following technical result.

Proposition 5.5 *Let $f_0(z)$ be an H-admissible function with radius of convergence $\rho < \infty$. Then $f(z) = e^{-z} f_0(z)$ is H-admissible as well.*

Proof. Since f_0 is analytic at the origin, it is clear that $f(z)$ is analytic at the origin as well, with a radius of convergence equal to that of $f_0(z)$.

Let $h(z)$ be such that $f(z) = e^{h(z)}$. If $h(z) = h_0(z) - z$, then we have $f_0(z) = e^{h_0(z)}$.

Let $a_0(t) = th_0'(t)$, $a(t) = th'(t)$, $b_0(t) = t^2 h_0''(t) + a_0(t)$ and $b(t) = t^2 h''(t) + a(t)$. Then $a(t) = a_0(t) - t$ and $b(t) = b_0(t) - t$.

It is immediate that $\lim_{t\to\rho} b(t) = +\infty$ since this limit holds for b_0. That is, Condition (H1) holds.

We now verify Condition (H2). Let $\delta(t)$ be a positive function such that $\lim_{t\to\rho} \delta(t) = 0$; and such that, uniformly for $|\theta| \le \delta(t)$, and as t tends to ρ,

$$h_0(te^{i\theta}) = h_0(t) + i\theta a_0(t) - \frac{1}{2}\theta^2 b_0(t) + o(1).$$

Then

$$\begin{aligned} h(te^{i\theta}) &= h_0(te^{i\theta}) - te^{i\theta} \\ &= h_0(t) + i\theta a_0(t) - \frac{1}{2}\theta^2 b_0(t) + o(1) - te^{i\theta} \\ &= h(t) + i\theta a(t) - \frac{1}{2}\theta^2 b(t) + o(1) - te^{i\theta} + t + ti\theta - \frac{1}{2}t\theta^2. \end{aligned}$$

We now observe that, if $|\theta| \le \delta(t)$ and as t tends to ρ, then $|ti\theta| \le t\delta(t) = o(1)$ and similarly, $\frac{1}{2}t\theta^2 = o(1)$. Finally,

$$|t(1 - e^{i\theta})| = t\sqrt{(1-\cos\theta)^2 + \sin^2\theta} = t\sqrt{2(1-\cos\theta)} \le t|\theta| \le t\delta(t) = o(1).$$

Thus $h(te^{i\theta}) = h(t) + i\theta a(t) - \frac{1}{2}\theta^2 b(t) + o(1)$ uniformly for $|\theta| \le \delta(t)$, which concludes the verification of (H2).

Finally, we want to show that $\frac{f(te^{i\theta})\sqrt{b(t)}}{f(t)}$ tends to 0 when t tends to ρ, uniformly for $\delta(t) \le |\theta| \le \pi$. We have

$$\begin{aligned}\frac{f(te^{i\theta})\sqrt{b(t)}}{f(t)} &= \frac{f_0(te^{i\theta})e^{-te^{i\theta}}\sqrt{b(t)}}{f_0(t)e^{-t}} \\ &= \frac{f_0(te^{i\theta})\sqrt{b_0(t)}}{f_0(t)} e^{t(1-e^{i\theta})}\sqrt{1-\frac{t}{b_0(t)}}.\end{aligned}$$

Since f_0 is H-admissible, uniformly for $\delta(t) \le |\theta| \le \pi$ and as t tends to ρ,

$$\frac{f_0(te^{i\theta})\sqrt{b_0(t)}}{f_0(t)} = o(1).$$

Moreover, $\sqrt{1-\frac{t}{b_0(t)}} = 1 + o(1)$ since $\lim_{t\to\rho} b_0(t) = +\infty$. Finally, when $0 < t < \rho$, $|e^{t(1-e^{i\theta})}| = e^{t(1-\cos\theta)} \le e^{2\rho}$. This suffices to conclude that (H3) holds, and hence that $f(z)$ is H-admissible. □

Proof of Proposition 5.3. Let $\mathcal{L}$ be the set of partial injections without fixpoints (i.e., without size 1 cycles), let L_n be the number of size n elements of $\mathcal{L}$ and let $L(z)$ be the corresponding EGS. We want to show that $L_n = \frac{1}{e} I_n(1 + o(1))$.

The EGS $L(z)$ is computed using the standard calculus of enumeration of labeled structures (displayed in [11, Figure II-18]), which was already used to compute $I(z)$ in [3]: since the EGS of cycles is $\log(\frac{1}{1-z})$, the EGS of cycles of size at least 2 is $\log(\frac{1}{1-z}) - z$ and the EGS of non-empty sequences is $\frac{z}{1-z}$, we have

$$L(z) = \exp\left(\log(\frac{1}{1-z}) - z + \frac{z}{1-z}\right) = \frac{1}{1-z}\exp\left(\frac{z^2}{1-z}\right) = I(z)e^{-z}.$$

We already know that $I(z)$ is H-admissible [3, Lemma 2.8] and Proposition 5.5 shows that $L(z)$ is H-admissible as well.

The saddlepoint is the solution ζ_n in the open interval $]0,1[)$ of the equation $z\frac{L'(z)}{L(z)} = n$. An elementary computation shows that we need to solve the equation

$$\begin{aligned} & z^3 + (n-1)z^2 - (2n+1)z + n = 0, \\ \textit{i.e.,}\quad & (z+n+1)(1-z)^2 - 1 = 0. \end{aligned}$$

Letting $z = 0$ and $z = 1$ in this equation shows that there is a solution in the interval $(0,1)$; moreover, one verifies easily that $(z+n+1)(1-z)^2 - 1$ is monotonous on $(0,1)$, and hence our equation has exactly one solution in that interval, say, ζ_n. From $0 < \zeta_n < 1$, we deduce that $\frac{1}{n+2} < (1-\zeta_n)^2 < \frac{1}{n+1}$, and hence $1 - \sqrt{\frac{1}{n+1}} < \zeta_n < 1 - \sqrt{\frac{1}{n+2}}$. In particular, $\zeta_n = 1 - \frac{1}{\sqrt{n}} + \mathcal{O}\left(\frac{1}{n\sqrt{n}}\right)$.

It now follows from Theorem 4.11 that

$$[z^n]L(z) = \frac{L(\zeta_n)}{\zeta_n^n \sqrt{2\pi b(\zeta_n)}}(1+o(1)).$$

In view of the proof of Proposition 5.5, $b(t) = b_0(t) - t$, where b_0 is the corresponding function for the H-admissible function $I(z)$. Using [3, Equation (7), p. 392], we find that $b_0(t) = \frac{2t}{(1-t)^3}$, and hence $b(t) = \frac{t(1+3t-3t^2+t^3)}{(1-t)^3}$.

Elementary computations show that

$$\begin{aligned}
b(\zeta_n) &= 2n^{\frac{3}{2}}\left(1+\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right)\\
\text{and } \frac{1}{\sqrt{2\pi b(\zeta_n)}} &= \frac{n^{-\frac{3}{4}}}{2\sqrt{\pi}}\left(1+\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right).\\
\text{Moreover, } \zeta_n^{-n} &= \exp\left(-n\log\left(1-\frac{1}{\sqrt{n}}+\mathcal{O}\left(\frac{1}{n\sqrt{n}}\right)\right)\right)\\
&= \exp\left(\sqrt{n}+\frac{1}{2}+\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right).\\
\text{Finally, } L(\zeta_n) &= \sqrt{n}\exp\left(\sqrt{n}-2+\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right)\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right).
\end{aligned}$$

At last we have: $[z^n]L(z) = \dfrac{e^{-\frac{3}{2}}}{2\sqrt{\pi}} n^{-\frac{1}{4}} e^{2\sqrt{n}}\left(1+\mathcal{O}\left(\dfrac{1}{\sqrt{n}}\right)\right).$

Comparing with the estimate of $[z^n]I(z)$ in Proposition 3.9, we find the announced result, namely

$$\frac{L_n}{I_n} = \frac{[z^n]L(z)}{[z^n]I(z)} = \frac{1}{e}(1+o(1)).$$

□

6 Finitely presented groups

One of the motivations for the study of subgroup distributions has been the investigation of the statistical properties of *finitely presented groups*, see [14, 8, 9, 28, 2, 1]. Strictly speaking, this would require a notion of distribution of these groups, so that one would make a list of non-isomorphic groups and investigate the frequency of groups with certain properties within that list. No such notion is available, as far as the authors are aware and current literature operates rather with a notion of distribution of finite *presentations*.

Recall that a *finite presentation* is a pair (A, R), where A is a finite set (the alphabet of generators) and R is a tuple of elements of $F(A)$ (the relators). The resulting finitely presented group G, written $G = \langle A \mid R\rangle$, is the quotient $G = F(A)/N(R)$, where $N(R)$ is the *normal subgroup generated* by R. The usual approach of statistical properties of finitely presented groups is based on the uniform distribution on k-tuples of reduced (or cyclically reduced) words of length at most n.

Of course, different presentations may yield the same group, even if the alphabet of generators is fixed. We are not aware of an analogue of Proposition 3.8 above, which would state, say, that the distribution of finitely presented groups with k relators of length at most n resulting from the uniform distribution on k-tuples of reduced (or cyclically reduced) words of length at most n, is uniform, at least on a generic subset of k-tuples. However, partial results exist in this direction for one-relator groups (Kapovich, Schupp and Shpilrain [18], Sapir and Spakulova [30, 31]).

In this section, we want to discuss an idea that may seem reasonable in this context, but which turns out to be disappointing. If H is the subgroup generated by the tuple of relators R, then $N(R) = N(H)$, so the group $G = \langle A \mid R\rangle$ is also specified by the pair $\langle A \mid H\rangle$. Thus, instead of looking at the normal closure of a finite set of elements, we look at the normal closure of a finitely generated subgroup. Now, clearly, if one generates a list of subgroups H by listing k-tuples of generators (the word-based distribution discussed earlier in this article), then the distribution of groups produced by this process will be the same as if one were working with presentations.

The idea we wish to explore is to generate the subgroup H via its Stallings graph, that is, to use the graph-based distribution of subgroups. Precisely, we may present groups via pairs, $\langle A \mid \Gamma\rangle$ where A is an alphabet and Γ is a Stallings graph. This is *a priori* a more compact representation of the group (more compact in bit size, less convenient to LaTeX).[2] More importantly, as we have seen that the graph-based distribution of subgroups is different from the word-based distribution, we may anticipate a different distribution of finitely presented groups as well, which would give us different insights on finitely presented groups.

Now an interesting feature of the statistical study of group presentations by tuples of relators is that the groups produced are generically non-trivial, and in fact infinite. More strongly, if A and k are fixed and if the maximal length n of the relators in the k-tuple R tends to infinity, then generically $G = \langle A \mid R\rangle$ is such that every subgroup generated by $|A| - 1$ elements is free [2]. It is also known that G is generically hyperbolic (Ol'shanskiĭ [28] and Champetier [8, 9], proving a statement of Gromov [14]).

In sharp contrast, and somewhat disappointingly, generically, a finitely presented group of the form $\langle A \mid \Gamma\rangle$ is trivial.

Theorem 6.1 *Generically, the finitely presented group $\langle A \mid \Gamma\rangle$ is trivial. In other words, generically, the normal closure of a randomly chosen subgroup of F_r of size n, is F_r itself.*

The rest of Section 6 is devoted to the proof of Theorem 6.1. We note that if the lengths of the cycles of the partial injection induced by letter a in $\Gamma(H)$ are relatively prime, then a belongs to the normal subgroup $N(H)$, and hence $a = 1$ in $G = \langle A \mid H\rangle$. Thus it suffices to prove the following proposition.

[2] It would be more interesting to have a unique, discrete representation of finitely generated normal subgroups, but no such representation seems to be known. And distinct normal subgroups may well lead to isomorphic quotients.

Proposition 6.2 *Generically, the lengths of the cycles of a size n partial injection are relatively prime.*

Remark 6.3 Our proof that $\langle A \mid H \rangle$ is generically trivial relies on a rather rough upper bound: we show that generically with probability $1 - \mathcal{O}(n^{-\frac{1}{6}})$, each letter a is a product of conjugates of powers of a in H. We do not know whether $\langle A \mid H \rangle$ is exponentially generically trivial. See Remark 4.10 for a similar situation. □

6.1 The permutation case

We start with the case of permutations, which is interesting in and of itself. Observe that if the lengths of the orbits of a permutation are not relatively prime, then these lengths have a common prime divisor p, which is in particular a divisor of n. Let $\mathcal{P}_n^{(p)}$ be the set of size n permutations in which all the orbits have size a multiple of p.

Lemma 6.4 *Let $n \geq 2$ and let p be a prime divisor of n. Then*

$$|\mathcal{P}_n^{(p)}| \leq 2n!\, n^{\frac{1}{p}-1}$$

Proof. We fix p, so n is of the form $n = mp$ and we proceed by induction on m. If $m = 1$, that is, $p = n$, then $|\mathcal{P}_n^{(p)}|$ is the number of size n cycles, namely $(n-1)!$. We now assume that $m > 1$.

We enumerate the elements of $\mathcal{P}_n^{(p)}$ in terms of the size kp of the orbit of 1: to determine such a permutation, one needs to select the other $kp-1$ elements of that orbit, select a cycle on these kp elements, and select a permutation on the remaining elements, that is, an element of $\mathcal{P}_{n-kp}^{(p)}$. Thus, using the convention that $|\mathcal{P}_0^{(p)}| = 1$, we have

$$\begin{aligned}
|\mathcal{P}_n^{(p)}| &= \sum_{k=1}^{m} \binom{n-1}{kp-1} (kp-1)! |\mathcal{P}_{n-kp}^{(p)}| \\
&= \sum_{k=1}^{m} \frac{(n-1)!}{(n-kp)!} |\mathcal{P}_{n-kp}^{(p)}| \\
&= (n-1)! \sum_{j=0}^{m-1} \frac{|\mathcal{P}_{jp}^{(p)}|}{(jp)!}.
\end{aligned}$$

Isolating the term $j = 0$ and using the induction hypothesis, it follows that

$$|\mathcal{P}_n^{(p)}| \leq (n-1)! \left(1 + 2\sum_{j=1}^{m-1} (jp)^{\frac{1}{p}-1}\right) = (n-1)! \left(1 + 2p^{\frac{1}{p}-1} \sum_{j=1}^{m-1} j^{\frac{1}{p}-1}\right).$$

Since the map $x \mapsto x^{\frac{1}{p}-1}$ is non-increasing on positive reals, we have

$$(j+1)^{\frac{1}{p}-1} \leq \int_j^{j+1} x^{\frac{1}{p}-1} dx = p((j+1)^{\frac{1}{p}} - j^{\frac{1}{p}}).$$

Therefore, isolating the term $j = 1$,

$$\begin{aligned}\sum_{j=1}^{m-1} j^{\frac{1}{p}-1} &= 1 + \sum_{j=2}^{m-1} j^{\frac{1}{p}-1} \\ &= 1 + \sum_{j=1}^{m-2} (j+1)^{\frac{1}{p}-1} \\ &\leq 1 + p((m-1)^{\frac{1}{p}} - 1) \\ &\leq 1 + p(m^{\frac{1}{p}} - 1) \;\; = \;\; p^{1-\frac{1}{p}} n^{\frac{1}{p}} - p + 1.\end{aligned}$$

Now we have

$$\begin{aligned}|\mathcal{P}_n^{(p)}| &\leq (n-1)! \left(1 + 2p^{\frac{1}{p}-1} \left(p^{1-\frac{1}{p}} n^{\frac{1}{p}} - p + 1\right)\right) \\ &\leq (n-1)! \left(2n^{\frac{1}{p}} + 1 - 2p^{\frac{1}{p}} + 2p^{\frac{1}{p}-1}\right).\end{aligned}$$

Since $p \geq 2$, it holds

$$p^{\frac{1}{p}} - p^{\frac{1}{p}-1} = p^{\frac{1}{p}} \left(1 - \frac{1}{p}\right) \geq p^{\frac{1}{p}} \frac{1}{2} \geq \frac{1}{2}$$

and hence

$$|\mathcal{P}_n^{(p)}| \leq (n-1)! 2n^{\frac{1}{p}} = 2n! n^{\frac{1}{p}-1},$$

which concludes the proof. □

Proposition 6.5 *The probability that the lengths of the orbits of a size n permutation are not relatively prime is at most equal to $\frac{2}{\sqrt{n}} + 2n^{-\frac{2}{3}} \log_3 n$.*

Proof. Let Q_n be the set of size n permutations for which the lengths of the orbits are not relatively prime, and let $q_n = \frac{|Q_n|}{n!}$.

As we already observed, a common divisor of the lengths of the orbits of a size n permutation is also a divisor of n. Therefore, if n is prime, then Q_n is the set of size n cycles, so $|Q_n| = (n-1)!$, $q_n = \frac{1}{n}$ and we have the desired result.

If n is not prime, then every size n permutation in Q_n is in $\mathcal{P}_n^{(p)}$ for some prime divisor p of n. These sets are not pairwise disjoint, but the sum of their cardinalities is an upper bound for $|Q_n|$. For these values of p, $|\mathcal{P}_n^{(p)}| \leq 2n! n^{\frac{1}{p}-1}$ by Lemma 6.4. Separating the case $p = 2$ from the cases $p \geq 3$, we find that $q_n \leq \frac{2}{\sqrt{n}} + 2Dn^{-\frac{2}{3}}$, where D is the number of distinct odd prime divisors of n. Since $n \geq 3^D$, we have $D \leq \log_3 n$ and hence

$$q_n \leq \frac{2}{\sqrt{n}} + 2Dn^{-\frac{2}{3}} \leq \frac{2}{\sqrt{n}} + 2n^{-\frac{2}{3}} \log_3 n,$$

which concludes the proof. □

6.2 Proof of Proposition 6.2

Isolating the cycles in a size n partial injection, reveals a permutation (on a subset X of $[n]$) and a fragmented permutation (i.e., a cycle-less partial injection) on the complement of X.

The EGS $J(z) = \sum_n \frac{J_n}{n!} z^n$ of fragmented permutations was discussed in Section 4.2, where we noted in particular that $J(z) = \exp\left(\frac{z}{1-z}\right)$. Let us add the following observation.

Lemma 6.6 *The sequence $(J_n/n!)_{n>0}$ is increasing.*

Proof. Let $M_n = \frac{J_n}{n!}$, so that $J(z) = \sum_{n\geq 0} M_n z^n$. The equalities $\frac{d}{dz}J(z) = \frac{1}{(1-z)^2}J(z)$, and hence $(1-z)^2\frac{d}{dz}J(z) = J(z)$, yield the following recurrence relation, for all $n \geq 2$:

$$(n+1)M_{n+1} = (2n+1)M_n - (n-1)M_{n-1}.$$

It follows that, for all $n \geq 2$,

$$(n+1)(M_{n+1} - M_n) = nM_n - (n-1)M_{n-1} = n(M_n - M_{n-1}) + M_{n-1}.$$

The result follows by induction since $M_1 = 1$ and $M_2 = \frac{3}{2}$ (see for instance [11, Section II.4.2]). □

Specifying a size n partial injection whose permutation part (the union of the cycles) has size k, amounts to choosing k elements, choosing a permutation on these k elements, and choosing a fragmented permutation on the remaining $n-k$ elements: the number of such partial injections is

$$\binom{n}{k} k! J_{n-k} = n! \frac{J_{n-k}}{(n-k)!};$$

and the number of those in which the sizes of the cycles have a non-trivial gcd is at most equal to

$$2n! \frac{J_{n-k}}{(n-k)!} \left(\frac{1}{\sqrt{k}} + \frac{\log_3 k}{k^{\frac{2}{3}}} \right)$$

by Proposition 6.5. Moreover, summing the numbers of partial injections with permutation part of size k, we get

$$I_n = \sum_{k=0}^{n} n! \frac{J_{n-k}}{(n-k)!}.$$

We use these observations to show the following facts, which together suffice to establish Proposition 6.2.

Fact 6.7 The proportion of size n partial injections whose permutation part has size less than $n^{\frac{1}{3}}$ is $\mathcal{O}(n^{-\frac{1}{6}})$. □

Fact 6.8 The proportion of size n partial injections whose permutation part has size greater than $n^{\frac{1}{3}}$ and for which the sizes of the cycles has a non-trivial gcd, is $\mathcal{O}(n^{-\frac{1}{6}})$. $\square$

Proof of Fact 6.7. The proportion of size n partial injections whose permutation part has size less than $n^{\frac{1}{3}}$ is

$$\begin{aligned}\frac{1}{I_n}\sum_{k=0}^{\lfloor n^{\frac{1}{3}}\rfloor} n!\frac{J_{n-k}}{(n-k)!} &\leq (n^{\frac{1}{3}}+1)\frac{J_n}{I_n} \quad \text{by Lemma 6.6}\\ &\leq \mathcal{O}(n^{-\frac{1}{6}}).\end{aligned}$$

The last inequality holds since $\frac{J_n}{I_n} = \mathcal{O}(n^{-\frac{1}{2}})$ (compare the asymptotic equivalents of $\frac{J_n}{n!}$ given in Section 4.2.2 and of $\frac{I_n}{n!}$ in Proposition 3.9). $\square$

Proof of Fact 6.8. Here we use Proposition 6.5 and the fact that, for large enough integers, we have $\frac{1}{\sqrt{k}} + \frac{\log_3 k}{k^{\frac{2}{3}}} \leq \frac{2}{\sqrt{k}}$. The number of size n partial injections whose permutation part has size greater than $n^{\frac{1}{3}}$ and for which the sizes of the cycles has a non-trivial gcd, is bounded above by

$$\begin{aligned}\sum_{k=\lceil n^{\frac{1}{3}}\rceil}^{n} 2n!\frac{J_{n-k}}{(n-k)!}\left(\frac{1}{\sqrt{k}}+\frac{\log_3 k}{k^{\frac{2}{3}}}\right) &\leq 4n^{-\frac{1}{6}}\sum_{k=\lceil n^{\frac{1}{3}}\rceil}^{n} n!\frac{J_{n-k}}{(n-k)!}\\ &\leq 4n^{-\frac{1}{6}}\sum_{k=0}^{n} n!\frac{J_{n-k}}{(n-k)!} = 4n^{-\frac{1}{6}}I_n.\end{aligned}$$

Thus the proportion of these partial injections is at most $4n^{-\frac{1}{6}}$. $\square$

Acknowledgements

We want to thank Nicolas Pouyanne for fruitful discussions on the gcd of the lengths of cycles in a random permutation. We also thank the anonymous referees for their insightful remarks, pointing us to important references and spotting two computational mistakes.

References

[1] Goulnara N. Arzhantseva. A property of subgroups of infinite index in a free group. *Proc. Amer. Math. Soc.*, 128(11):3205–3210, 2000.

[2] Goulnara N. Arzhantseva and Alexander Yu. Ol′shanskiĭ. Generality of the class of groups in which subgroups with a lesser number of generators are free. *Mat. Zametki*, 59(4):489–496, 638, 1996.

[3] Frédérique Bassino, Cyril Nicaud, and Pascal Weil. Random generation of finitely generated subgroups of a free group. *Internat. J. Algebra Comput.*, 18(2):375–405, 2008.

[4] Gilbert Baumslag, Alexei Myasnikov, and Vladimir Remeslennikov. Malnormality is decidable in free groups. *Internat. J. Algebra Comput.*, 9(6):687–692, 1999.

[5] Jean-Camille Birget, Stuart Margolis, John Meakin, and Pascal Weil. PSPACE-complete problems for subgroups of free groups and inverse finite automata. *Theoret. Comput. Sci.*, 242(1-2):247–281, 2000.

[6] Steven A. Bleiler and Amelia C. Jones. The free product of groups with amalgamated subgroup malnormal in a single factor. *J. Pure Appl. Algebra*, 127(2):119–136, 1998.

[7] Robert G. Burns. On finitely generated subgroups of free products. *J. Austral. Math. Soc.*, 12:358–364, 1971.

[8] Christophe Champetier. *Propriétés génériques des groupes de présentation finie.* Ph.D. Thesis, Université Lyon 1 (1991).

[9] Christophe Champetier. Propriétés statistiques des groupes de présentation finie. *J. Adv. Math.*, 116(2):197–262, 1995.

[10] Louis Comtet. *Advanced combinatorics.* D. Reidel Publishing Co., Dordrecht, enlarged edition, 1974. The art of finite and infinite expansions.

[11] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[12] Philippe Flajolet, Paul Zimmerman, and Bernard Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theoret. Comput. Sci.*, 132(1-2):1–35, 1994.

[13] Stephen M. Gersten. Intersections of finitely generated subgroups of free groups and resolutions of graphs. *Invent. Math.*, 71(3):567–591, 1983.

[14] Mikhail Gromov. Hyperbolic groups. In *Essays in group theory*, S.M. Gersten ed., pages 75–265. Springer, 1987.

[15] Toshiaki Jitsukawa. Malnormal subgroups of free groups. In *Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001)*, volume 298 of *Contemp. Math.*, pages 83–95. Amer. Math. Soc., Providence, RI, 2002.

[16] Ilya Kapovich, Alexei Miasnikov, Paul Schupp, and Vladimir Shpilrain. Generic-case complexity, decision problems in group theory, and random walks. *J. Algebra*, 264(2):665–694, 2003.

[17] Ilya Kapovich and Alexei Myasnikov. Stallings foldings and subgroups of free groups. *J. Algebra*, 248(2):608–668, 2002.

[18] Ilya Kapovich, Paul Schupp, and Vladimir Shpilrain. Generic properties of Whitehead's algorithm and isomorphism rigidity of random one-relator groups. *Pacific J. Math.*, 223(1):113–140, 2006.

[19] Abraham Karrass and Donald Solitar. The free product of two groups with a malnormal amalgamated subgroup. *Canad. J. Math.*, 23:933–959, 1971.

[20] Olga Kharlampovich and Alexei Myasnikov. Hyperbolic groups and free constructions. *Trans. Amer. Math. Soc.*, 350(2):571–613, 1998.

[21] Stuart Margolis, Mark Sapir, and Pascal Weil. Closed subgroups in pro-$\mathbf{V}$ topologies and the extension problem for inverse automata. *Internat. J. Algebra Comput.*, 11(4):405–445, 2001.

[22] Alexei Miasnikov, Enric Ventura, and Pascal Weil. Algebraic extensions in free groups. In *Geometric group theory*, Trends Math., pages 225–253. Birkhäuser, Basel, 2007.

[23] Igor Mineyev. Submultiplicativity and the Hanna Neumann conjecture. *Annals of Math.*, 175(1): 393–414, 2011.

[24] Igor Mineyev. Groups, graphs, and the Hanna Neumann Conjecture. Preprint, 10 pages, `www.math.uiuc.edu/~mineyev/math/art/gr-gr-shnc.pdf`

[25] Alexei Myasnikov, Vladimir Shpilrain, and Alexander Ushakov. *Group-based cryptography.* Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser Verlag, Basel, 2008.

[26] Albert Nijenhuis and Herbert S. Wilf. *Combinatorial algorithms.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1978. For computers and calculators, Computer Science and Applied Mathematics.

[27] Yann Ollivier. *A January 2005 invitation to random groups.* Ensaios Matemáticos, 10. Sociedade Brasileira de Matemtica, 2005.

[28] Alexander Yu. Ol′shanskiĭ. Almost every group is hyperbolic. *Internat. J. Algebra Comput.*, 2(1):1–17, 1992.

[29] Abdó Roig, Enric Ventura, and Pascal Weil. On the complexity of the Whitehead minimization problem. *Internat. J. Algebra Comput.*, 17(8):1611–1634, 2007.

[30] Mark Sapir. Residual properties of 1-relator groups. Technical report, `arXiv:1001.2829v1`, 2010.

[31] Mark Sapir and Iva Spakulova. Almost all one-relator groups with at least three generators are residually finite. Technical report, `arXiv:0809.4693v2`, 2009.

[32] John R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.

[33] Nicholas W. M. Touikan. A fast algorithm for Stallings' folding process. *Internat. J. Algebra Comput.*, 16(6):1031–1045, 2006.

[34] Pascal Weil. Computing closures of finitely generated subgroups of the free group. In *Algorithmic problems in groups and semigroups (Lincoln, NE, 1998)*, Trends Math., pages 289–307. Birkhäuser Boston, Boston, MA, 2000.