Regular sets and counting in free groups

Elizaveta Frenkel
Alexei G. Myasnikov
Vladimir N. Remeslennikov

June 16, 2009

Contents

1 Introduction 1

2 Preliminaries 4
   2.1 Asymptotic densities 4
   2.2 Generating random elements and multiplicative measures 6
   2.3 The frequency measure 7
   2.4 Asymptotic classification of subsets 8
   2.5 Context-free and regular languages as a measuring tool 9

3 Schreier Systems of Representatives 10
   3.1 Subgroup and coset graphs 10
   3.2 Schreier transversals 13

4 Measuring subsets of $F$ 15

5 Comparing sets at infinity 19
   5.1 Comparing Schreier representatives 19
   5.2 Comparing regular sets 21

1 Introduction

In this paper we study asymptotic behavior of regular subsets in a free group $F$ of finite rank, compare their sizes at infinity, and develop techniques to compute the probabilities of sets relative to distributions on $F$ that come naturally from random walks on the Cayley graph of $F$. We apply these techniques to study cosets, double cosets, and Schreier representatives of
finitely generated subgroups of $F$ with an eye on complexity of algorithmic problems in free products with amalgamation and HNN extensions of groups.

During the last decade it has been realized that a natural set of algebraic objects usually can be divided into two parts. The large one (the regular part) consists of typical, ”generic” objects; and the smaller one (the ”singular” part) is made of ”exceptions”. Essentially, this idea appeared first in the form of zero-one laws in probability theory, number theory, and combinatorics. It became popular after seminal works of Erdős, that shaped up the so-called Probabilistic Method (see, for example, [1]). In finite group theory the idea of genericity can be traced down to a series of papers by Erdős and Turan in 1960-70’s (for recent results see, for example, [55]). In combinatorial group theory the concept of generic behavior is due to Gromov. His inspirational works [27, 28] turned the subject into an area of very active research, see for example, [2, 3, 4, 14, 7, 8, 13, 15, 16, 17, 8, 9, 10, 32, 33, 35, 36, 37, 38, 39, 42, 50, 51, 53, 59, 61]. It turned out that the generic objects usually have much simpler structure, while the exceptions provide most of the difficulties. For instance, generic finitely generated groups are hyperbolic [27, 51], generic subgroups of hyperbolic groups are free [26], generic cyclically reduced elements in free groups are of minimal length in their automorphic orbits [39], generic automorphisms of a free group are strongly irreducible [52], etc.

In practice, the generic-case analysis of algorithms is usually more important than the worst-case one. For example, knowing generic properties of objects one can often design simple algorithms that work very fast on generic elements. In cryptography, many successful attacks exploit the generic properties of random elements from a particular class, ignoring the existing hard instances [47, 48, 49, 54]. In the precise form the generic complexity of algorithmic problems appeared first in the papers [34, 35, 14, 13]. We refer the reader to a comprehensive survey [25] on generic complexity of algorithms.

In this paper we lay down some techniques that allow one to measure sets which appear naturally when computing with infinite finitely presented groups. Our main idea is to approximate a given set by some regular subsets and estimate the asymptotic sizes of the regular sets using powerful tools of random walks on graphs and generating functions. The particular applications we have in mind concern with the generic complexity of the Word and Conjugacy problems in free products with amalgamation and HNN extensions. In general, such problems can be extremely hard. In [45] Miller described a free product of free groups with finitely generated amalgamation where the Conjugacy problem is undecidable; while in [45] he gave similar examples in the class of generalized HNN-extension of free groups. However,
it has been proven in \cite{15,16} that on a precisely described set \( RP \) of "regular elements" in amalgamated free products and HNN extensions \( G \) the Conjugacy problem is decidable (under some natural conditions on the factors), furthermore, it is decidable in polynomial time. Namely, it was shown in \cite{15,16} that the group \( G \) (satisfying some natural assumptions) can be stratified into two parts with respect to the “hardness” of the conjugacy problem:

- the Regular Part \( RP \) consists of so-called regular elements for which the conjugacy problem is decidable in polynomial time by the standard algorithms (described in \cite{44,43}). Moreover, one can decide whether or not a given element is regular in \( G \);
- the Black Hole \( BH \) (the complement of \( RP \) in \( G \)) consists of elements in \( G \) for which either the standard algorithms do not work at all, or they are slow, or the situation is not quite clear yet.

The missing piece is to show that the set \( RP \) is, indeed, generic in \( G \). This is not easy, the complete proof, which will appear in \cite{24}, relies on the techniques developed in the present paper. Now, a few words on the structure of the paper. In Section \( 2 \) following \cite{14}, we describe some techniques for measuring subsets in a free group \( F \), the asymptotic classification of large and small sets, and approximations via context-free and regular sets.

In Section \( 3 \) we study, using graph techniques, Shreier system of representatives (transversal) of a finitely generated subgroup \( C \) in a free group \( F \) of finite rank. If \( S \) is a fixed Schreier transversal of \( C \) then \( s \in S \) is called \( stable \) (on the right) if \( sc \in S \) for any \( c \in C \). Intuitively, the stable representatives are "regular", they are easy to deal with.

In Section \( 4 \) we estimate the sizes of various subsets of \( F \). In particular, we show that \( S \) is regular and thick (see definitions in Section \( 2 \)), meanwhile the set \( S_{nst} \) of non-stable representatives from \( S \) is exponentially negligible. Furthermore, the set \( S_{nst} \) is exponentially negligible even relative to the set \( S \). Our approach here is to "approximate" the sets in question by regular sets and to measure sizes of the regular sets using tools of random walks on graphs and Perron-Frobenius techniques.

In Section \( 5 \) we develop a technique to compare sizes of different regular sets at "infinity" and give an asymptotic classification of regular subsets of \( F \) relative to a fixed prefix-closed regular subset \( L \subseteq F \). The main result describes when regular subsets of \( L \) are "large" or "small" at infinity in comparison to \( L \). Notice, in the case when \( L = F \), this result has been proven in \cite{14} (Theorem 3.2).
2 Preliminaries

In this section, following [14], we describe some techniques for measuring subsets in a free group $F$, the asymptotic classification of large and small sets, and approximations via context-free and regular sets.

2.1 Asymptotic densities

Let $F = F(X)$ be a free group with basis $X = \{x_1, \ldots, x_m\}$. We use this notation throughout the paper.

Let $R$ be a subset of the free group $F$ and $S_k = \{ w \in F \mid |w| = k \}$ the sphere of radius $k$ in $F$. The fraction

$$f_k(R) = \frac{|R \cap S_k|}{|S_k|}$$

is the frequency of elements from $R$ among the words of length $k$ in $F$. The asymptotic density $\rho(R)$ of $R$ is defined by

$$\rho(R) = \limsup_{k \to \infty} f_k(R).$$

$R$ is called generic if $\rho(R) = 1$, and negligible if $\rho(R) = 0$. If, in addition, there exists a positive constant $\delta < 1$ such that

$$1 - \delta^k < f_k(R) < 1$$

for all sufficiently large $k$ then $R$ is called exponentially generic. Meanwhile, if $f_k(R) < \delta^k$ for large enough $k$ then $R$ is exponentially negligible. In both the cases we refer to $\delta$ as a rate upper bound. Sometimes such sets are also called strongly generic or strongly negligible, but we refrain from this.

The Cesaro limit

$$\rho^c(R) = \lim_{n \to \infty} \frac{1}{n} (f_1 + \cdots + f_n).$$

(1)

gives another asymptotic characteristic, called Cesaro density, or asymptotic average density. Sometimes, it is more sensitive then standard asymptotic density $\rho$ (see, for example, [14], [60]). However, if $\lim_{k \to \infty} f_k(R)$ exists (hence is equal to $\rho(R)$) then $\rho^c(R)$ also exists and $\rho^c(R) = \rho(R)$. We will have to say more about $\rho^c(R)$ below.

Asymptotic density gives the first coarse classification of large (small) subsets:
Coarse classification

1) Generic sets;
2) Visible or thick sets: the set $R$ is visible if $\rho(R) > 0$;
3) Negligible sets.

Unfortunately, this classification is very coarse, it does not distinguish many sets which, intuitively, have different sizes.

All our results in this paper concern with the strong version of the asymptotic density $\rho$, when the actual limit $\lim_{k \to \infty} f_k(R)$ exists. This allows one to differentiate sets with the same asymptotic density with respect to their growth rates. Thus generic sets $R$ divide into subclasses of exponential, subexponential, superpolynomial, polynomial generic sets, with respect to the convergence rates of their frequency sequences $\{f_k(R)\}_{k \in \mathbb{N}}$. The same holds for negligible sets as well.

2.2 Generating random elements and multiplicative measures

One can use a no-return random walk $W_s$ ($s \in (0, 1]$) on the Cayley graph $C(F, X)$ of $F$ with respect to the generating set $X$, as a random generator of elements of $F$ (see [14]). We start at the identity element 1 and either do nothing with probability $s$ (and return value 1 as the output of our random word generator), or move to one of the $2m$ adjacent vertices with equal probabilities $(1 - s)/2m$. If we are at a vertex $v \neq 1$, we either stop at $v$ with probability $s$ (and return the value $v$ as the output), or move, with probability $\frac{1 - s}{2m - 1}$, to one of the $2m - 1$ adjacent vertices lying away from 1, thus producing a new freely reduced word $vx_i^\pm 1$. Since the Cayley graph $C(F, X)$ is a tree and we never return to the word we have already visited, it is easy to see that the probability $\mu_s(w)$ for our process to terminate at a word $w$ is given by the formula

$$\mu_s(w) = \frac{s(1 - s)^{|w|}}{2m \cdot (2m - 1)^{|w| - 1}}$$

for $w \neq 1$ (2)

and

$$\mu_s(1) = s.$$  (3)

For $R \subseteq F$ its measure $\mu_s(R)$ is defined by $\mu_s(R) = \sum_{w \in R} \mu_s(w)$. Recalculating $\mu_s(R)$ in terms of $s$, one gets

$$\mu_s(R) = s \sum_{k=0}^{\infty} f_k(1 - s)^k,$$
and the series on the right hand side is convergent for all \( s \in (0, 1) \). The ensemble of distributions \( \{\mu_s\} \) can be encoded in a single function

\[
\mu(R) : s \in (0, 1) \to \mu_s(R) \in \mathbb{R}.
\]

The argument above shows that for every subset \( R \subseteq F \), \( \mu(R) \) is an analytic function of \( s \).

It has been shown in \([14]\) that \( \mu(R) \) contains a lot of information about the asymptotic behaviour of the set \( R \). To see where this information comes from renormalise the measures \( \mu_s \) and consider the parametric family \( \mu^* = \{\mu^*_s\} \) of adjusted measures

\[
\mu^*_s(w) = \left( \frac{2m}{2m - 1} \cdot \frac{1}{s} \right) \cdot \mu_s(w).
\]

This new measure \( \mu^*_s \) is multiplicative in the sense that

\[
\mu^*_s(u \circ v) = \mu^*_s(u)\mu^*_s(v),
\]

where \( u \circ v \) denotes the product of non-empty words \( u \) and \( v \) such that \( |uv| = |u| + |v| \) (no cancelation in the product \( uv \)). Moreover, if we denote

\[
t = \mu^*_s(x_1^{\pm 1}) = \frac{1 - s}{2m - 1}
\]

then

\[
\mu^*_s(w) = t^{|w|}
\]

for every non-empty word \( w \). Assume now, for the sake of minor technical convenience, that \( R \) does not contain the identity element 1. It is easy to see that

\[
\mu^*_s(R) = \sum_{k=0}^{\infty} n_k(R)t^k
\]

is the generating function of the spherical growth sequence

\[
n_k(R) = |R \cap S_k|
\]

of the set \( R \) in variable \( t \) which is convergent for each \( t \in [0, 1) \).

The distribution \( \mu_s \) has the uncomfortably big standard deviation \( \sigma = \sqrt{\frac{1-s}{s}} \), which reflects the fact that \( \mu_s \) is strongly skewed towards "short" elements. The mean length of words in \( F \) distributed according to \( \mu_s \) is equal to \( L_s = \frac{1}{s} - 1 \), so \( L_s \to \infty \) when \( s \to 0 \). This shows that the
asymptotic behaviour of the set $R$ at "infinity" (when $L_s \to \infty$) depends on the behaviour of the function $\mu(R)$ when $s \to 0^+$.

Following [14], for a subset $R$ of $F$ we define a numerical characteristic

$$\mu_0(R) = \lim_{s \to 0^+} \mu(R) = \lim_{s \to 0^+} s \sum_{k=0}^{\infty} f_k(1-s)^k.$$ 

If $\mu(R)$ can be expanded as a convergent power series in $s$ at $s = 0$ (and hence in some neighborhood of $s = 0$):

$$\mu(R) = m_0 + m_1 s + m_2 s^2 + \cdots,$$

then

$$\mu_0(R) = \lim_{s \to 0^+} \mu_s(R) = m_0,$$

and an easy corollary from a theorem by Hardy and Littlewood [30, Theorem 94] asserts that $\mu_0(R)$ is precisely the Cesaro limit $\rho^c(R)$.

A subset $R \subseteq F$ is called smooth [14] if $\mu(R)$ can be expanded as a convergent power series in $s$ at $s = 0$.

### 2.3 The frequency measure

In this section we discuss the frequency measure, introduced in [14].

Let $W_0$ be the no-return non-stop random walk on the Cayley graph $C(F, X)$ of $F$ (like $W_s$ with $s = 0$), where the walker moves from a given vertex to any adjacent vertex away from the initial vertex 1 with equal probabilities $1/2m$. In this event, the probability $\lambda(w)$ that the walker hits an element $w \in F$ in $|w|$ steps (which is the same as the probability that the walker ever hits $w$) is equal to

$$\lambda(w) = \frac{1}{2m(2m - 1)^{|w|-1}}, \text{ if } w \neq 1, \text{ and } \lambda(1) = 1.$$

This gives rise to a measure called the frequency measure on $F$, or Boltzmann distribution, defined for subsets $R \subseteq F$ by

$$\lambda(R) = \sum_{w \in R} \lambda(w),$$

if the sum above is finite. One can view $\lambda(R)$ as the cumulative frequency of $R$ since

$$\lambda(R) = \sum_{k=0}^{\infty} f_k(R).$$
This measure is not probabilistic, since, for instance, \( \lambda(F) = \infty \), moreover, \( \lambda \) is additive, but not \( \sigma \)-additive.

A subset \( R \subseteq F \) is called \( \lambda \)-measurable, or simply measurable (since we do not consider any other measures in this paper) if \( \lambda(R) < \infty \). Every measurable set is negligible.

**Linear approximation.** If the set \( R \) is smooth then the linear term in the expansion of \( \mu(R) \) gives a linear approximation of \( \mu(R) \):

\[
\mu(R) = m_0 + m_1 s + O(s^2).
\]

In this case, \( m_0 = \mu_0(R) \) is the Cesaro density of \( R \). An easy corollary of [30, Theorem 94] shows that if \( \mu_0(R) = 0 \) then

\[
m_1 = \sum_{k=1}^{\infty} f_k(R) = \lambda(R).
\]

On the other hand, even without assumption that \( R \) is smooth, if \( R \) is measurable, then

\[
\mu_0(R) = 0 \text{ and } \mu_1 = \lim_{s \to 0^+} \frac{\mu(s)}{s} = \lambda(R).
\]

**2.4 Asymptotic classification of subsets**

In this section we describe a classification of subsets \( R \) in \( F \), according to the asymptotic behavior of the functions \( \mu(R) \).

Recall, that the function \( \mu(R) \) is analytic on \((0, 1)\) for every subset \( R \) of \( F \). \( R \) is smooth if \( \mu(R) \) can be analytically extended to a neighborhood of \( 0 \). The subset \( R \) is called rational, algebraic, etc, with respect to \( \mu \) if the function \( \mu(R) \) is rational, algebraic, etc.

**Asymptotic classification of sets.** The following subtler classification of sets in \( F \) (based on the linear approximation of \( \mu(R) \)) was introduced in [14]:

- **Thick subsets**: \( \mu_0(R) \) exists, \( \mu_0(R) > 0 \) and

\[
\mu(R) = \mu_0(R) + \alpha_0(s), \text{ where } \lim_{s \to 0^+} \alpha_0(s) = 0.
\]

- **Negligible subsets of intermediate density**: \( \mu_0(R) = 0 \) but \( \mu_1(R) \) does not exist.
• **Sparse negligible subsets**: \( \mu_0(R) = 0, \mu_1(R) \) exists and

\[
\mu(R) = \mu_1(R)s + \alpha_1(s) \quad \text{where} \quad \lim_{s \to 0^+} \frac{\alpha_1(s)}{s} = 0.
\]

• **Exponentially negligible sets**:

• **Singular sets**: \( \mu_0(R) \) does not exist.

For sparse sets, the values of \( \mu_1 \) provide a further and more subtle discrimination by size.

**Lemma 2.1.** [14] A subset is sparse in \( F \) if and only if it is measurable.

### 2.5 Context-free and regular languages as a measuring tool

The simple observation in Section 2.2 that \( \mu(R) \) is the generating function of the grows sequence \( \{n_k(R)\}_{k \in \mathbb{N}} \) allows one to apply a well-established machinery of generating functions of regular and context-free languages to estimate asymptotic sizes of subsets \( R \) in \( F \). We refer to [31] on regular and context-free languages, and to [21] on regular languages in groups.

**Algebraic sets and context-free languages.** If the set \( R \) is an (unambiguous) context-free language then, by a classical theorem of Chomsky and Schutzenberger [19], the generating function \( \mu^*(R) = \sum n_k(R)t^k \) and hence the function \( \mu(R) \), are algebraic functions of \( s \). Moreover, if \( R \) is regular then \( \mu(R) \) is a rational function with rational coefficients [22, 57].

It is well known that singular points of an algebraic function are either poles or branching points. Since \( \mu(R) \) is bounded for \( s \in (0, 1) \), this means that, for a context-free set \( R \), the function \( \mu(R) \) has no singularity at 0 or has a branching point at 0. A standard result on analytic functions allows us to expand \( \mu_s(R) \) as a fractional power series:

\[
\mu_s(R) = m_0 + m_1 s^{1/n} + m_2 s^{2/n} + \cdots,
\]

\( n \) being the branching index. This technique was used in [16, 17] for numerical estimates of generic complexity of algorithms.

If \( R \) is regular, than we actually have the usual power series expansion:

\[
\mu_s(R) = m_0 + m_1 s + m_2 s^2 + \cdots;
\]

in particular, \( \mu(R) \) can be analytically extended in the neighborhood of 0, so \( R \) is smooth.

The following gives an asymptotic classification of regular subsets of \( F \).
Theorem 2.2. [14, 6]

1) Every negligible regular subset of $F$ is strongly negligible.

2) A regular subset of $F$ is thick if and only if its prefix closure contains a cone.

3) Every regular subset of $F$ is either thick or strongly negligible.

3 Schreier Systems of Representatives

3.1 Subgroup and coset graphs

In this section for a given finitely generated subgroup of a free group we discuss its subgroup and coset graphs.

Let $F = F(X)$ be a free group with basis $X = \{x_1, \ldots, x_n\}$. We identify elements of $F$ with reduced words in the alphabet $X \cup X^{-1}$. Fix a subgroup $C = \langle h_1, \ldots, h_m \rangle$ of $F$ generated by finitely many elements $h_1, \ldots, h_m \in F$.

Following [33], we associate with $C$ two graphs: the subgroup graph $\Gamma = \Gamma_C$ and the coset graph $\Gamma^o = \Gamma^o_C$. We freely use definitions and results from [33] in the rest of the paper.

Recall, that $\Gamma$ is a finite connected digraph with edges labeled by elements from $X$ and a distinguished vertex (based-point) 1, satisfying the following two conditions. Firstly, $\Gamma$ is folded, i.e., there are no two edges in $\Gamma$ with the same label and having the same initial or terminal vertices. Secondly, $\Gamma$ accepts precisely the reduced words in $X \cup X^{-1}$ that belong to $C$. To explain the latter observe, that walking along a path $p$ in $\Gamma$ one can read a word $\ell(p)$ in the alphabet $X \cup X^{-1}$, the label of $p$, (reading $x$ in passing an edge $e$ with label $x$ along the orientation of $e$, and reading $x^{-1}$ in the opposite direction). We say that $\Gamma$ accepts a word $w$ if $w = \ell(p)$ for some closed path $p$ that starts at 1 and has no backtracking. One can describe $\Gamma$ as a deterministic finite state automata with 1 as the unique starting and accepting state.

For example, the graph $\Gamma$ for the subgroup generated by $x_1x_2x_1^{-1}$ is shown in the picture below.

![Diagram](Pic. 1.)
Given the generators $h_1, \ldots, h_m$ of the subgroup $C$, as words from $F(X)$, one can effectively construct the graph $\Gamma$ in time $O(n \log^* n)$ [?].

The coset graph (also known as the Schreier graph) $\Gamma^* = \Gamma_C^*$ of $C$ is a connected labeled digraph with the set $\{Cu \mid u \in F\}$ of the right cosets of $C$ in $F$ as the vertex set, and such that there is an edge from $Cu$ to $Cv$ with a label $x \in X$ if and only if $Cux = Cv$. One can describe the coset graph $\Gamma^*$ as obtained from $\Gamma$ by the following procedure. Let $v \in \Gamma$ and $x \in X$ such that there is no outgoing or incoming edge at $v$ labeled by $x$. We call such $v$ a boundary vertex of $\Gamma$ and denote the set of such vertices by $\partial \Gamma$. For every such $v \in \partial \Gamma$ and $x \in X$ we attach to $v$ a new edge $e$ (correspondingly, either outgoing or incoming) labeled $x$ with a new terminal vertex $u$ (not in $\Gamma$). Such vertices $u$ are called frontier vertices, we denote the set of frontier vertices of $\Gamma$ by $\partial^+ \Gamma$. Then we attach to $u$ the Cayley graph $C(F, X)$ of $F$ relative to $X$ (identifying $u$ with the vertex $1$ of $C(F, X)$), and then we fold the edge $e$ with the corresponding edge in $C(F, X)$ (that is labeled $x$ and is incoming to $u$). Observe, that for every vertex $v \in \Gamma^*$ and every reduced word $w$ in $X \cup X^{-1}$ there is a unique path $\Gamma^*$ that starts at $v$ and has the label $w$. By $p_w$ we denote such a path that starts at $1$, and by $v_w$ the end vertex of $p_w$. Here is the fragment of the graph $\Gamma^*$ for $C = \langle x_1 x_2 x_1^{-1} \rangle$:

Pic. 2.

**Lemma 3.1.** $\Gamma_C^*$ is the coset graph of $C$ in $F$.

**Proof.** See, for example, [33]. \qed

Notice that $\Gamma = \Gamma^*$ if and only if the subgroup $C$ has finite index in $F$. Indeed, $\Gamma = \Gamma^*$ if and only if for every vertex $v$ of $\Gamma$ and every label $x \in X$,
there is an edge in $\Gamma$ labeled by $x$ which exits from $v$, and an edge with label $x$ which enters $v$, but this is precisely the characterization of subgroups of finite index in $F$ [33, Proposition 8.3].

A spanning subtree $T$ of $\Gamma$ with the root at the vertex 1 is called geodesic if for every vertex $v \in V(\Gamma)$ the unique path in $T$ from 1 to $v$ is a geodesic path in $\Gamma$. For a given graph $\Gamma$ one can effectively construct a geodesic spanning subtree $T$ (see, for example, [33]).

From now on we fix an arbitrary spanning subtree $T$ of $\Gamma$. It is easy to see that the tree $T$ uniquely extends to a spanning subtree $T^*$ of $\Gamma^*$.

Let $V(\Gamma^*)$ be the set of vertices of $\Gamma^*$. For a subset $Y \subseteq V(\Gamma^*)$ and a subgraph $\Delta$ of $\Gamma^*$, we define the language accepted by $\Delta$ and $Y$ as the set $L(\Delta,Y,1)$ of the labels $\ell(p)$ of paths $p$ in $\Delta$ that start at 1 and end at one of the vertices in $Y$, and have no backtracking. Notice that the words $\ell(p)$ are reduced since the graph $\Gamma^*$ is folded. Notice, that $F = L(\Gamma^*, V(\Gamma^*), 1)$ and $C = L(\Gamma, \{1\}, 1) = L(\Gamma^*, \{1\}, 1)$.

Sometimes we will refer to a set of right (left) representatives of $C$ as the right (left) transversal of $C$. Furthermore, to simplify terminology, a transversal will usually mean a right transversal, if not said otherwise. Recall, that a transversal $S$ of $C$ is termed Schreier if every initial segment of a representative from $S$ belongs to $S$.

**Proposition 3.2.** Let $C$ be a finitely generated subgroup of $F$. Then:

1) for every spanning subtree $T$ of $\Gamma$ the set $S_{T^*} = L(T^*, V(\Gamma^*), 1)$ is a Schreier transversal of $C$.

2) for every Schreier transversal $S$ of $C$ there exists a spanning subtree $T$ of $\Gamma$ such that $S = S_{T^*}$.

**Proof.** The statement 1) follows directly from Lemma[5.1] To prove 2) notice that every reduced path $p \in \Gamma^*$ can be decomposed as $p = p_{\text{int}} \circ p_{\text{out}}$, where $p_{\text{int}}$ is a maximal reduced path in $\Gamma$, and $p_{\text{out}}$ is the tail of $p$ outside of $\Gamma$. This decomposition is unique. Moreover:

- if $v \in V(\Gamma)$ and $p$ is a reduced path from 1 to $v$ in $\Gamma^*$ then $p$ passes only through vertices of $\Gamma$;

- if $v \in V(\Gamma^*) \setminus V(\Gamma)$ and $p'$ and $p''$ are two paths from 1 to $v$, where $p' = p'_{\text{int}} \circ p'_{\text{out}}$, and $p'' = p''_{\text{int}} \circ p''_{\text{out}}$, then $p'_{\text{out}} = p''_{\text{out}}$.

Let $S$ be a Schreier transversal of $C$ in $F$ and $s \in S$. Suppose that the reduced path $p_s$ ends at some vertex $v_s$ in $\Gamma$. Then the whole path $p_s$ lies in
Let $T$ be a subgraph of $\Gamma$ generated by the union of all paths $p_s$, where $s \in S$ and $v_s \in \Gamma$. Since $S$ is a Schreier transversal $T$ is a maximal subtree of $\Gamma$. It is clear that $S = S_{T^*}$. Hence, the result.

Proposition 3.2 allows one to identify elements from a given Schreier transversal $S$ of $C$ with the vertices of the graph $\Gamma^*$, provided a maximal subtree of $\Gamma$ is fixed. We use this frequently in the sequel.

Corollary 3.3. The number of distinct Schreier transversals of $C$ in $F$ is finite and equal to the number of spanning subtrees of $\Gamma_C$.

3.2 Schreier transversals

In this section we introduce various types of representatives of $C$ in $F$ relative to a fix basis $X$ of $F$.

Definition 3.4. Let $S$ be a transversal of $C$.

- A representative $s \in S$ is called \textit{internal} if the path $p_s$ ends in $\Gamma$, i.e., $v_s \in V(\Gamma)$. By $S_{\text{int}}$ we denote the set of all internal representatives in $S$. Elements from $S_{\text{ext}} = S \setminus S_{\text{int}}$ are called the \textit{external} representatives in $S$.

- A representative $s \in S$ is called \textit{geodesic} if it has minimal possible length in its coset $Cs$. The transversal $S$ is \textit{geodesic} if every $s \in S$ is geodesic.

- A representative $s \in S$ is called \textit{singular} if it belongs to the generalized normalizer of $C$: 
  \[ N^*(C) = \{ f \in F | f^{-1} C f \cap C \neq \emptyset \}. \]
  All other representatives from $S$ are called \textit{regular}. By $S_{\text{sin}}$ and, respectively, $S_{\text{reg}}$ we denote the sets of singular and regular representatives from $S$.

- A representative $s \in S$ is called \textit{stable} (on the right) if $sc \in S$ for any $c \in C$. By $S_{\text{st}}$ we denote the set of all stable representatives in $S$, and $S_{\text{uns}} = S \setminus S_{\text{st}}$ is the set of all non-stable representatives from $S$.

In the following lemma we collect some basic properties of various types of representatives. Recall that the \textit{cone} defined by (or based at) an element $u \in F$ is the set $C(u)$ of all reduced words in $F$ that have $u$ as an initial segment. For $u, v \in F$ we write $u \circ v$ if there is no cancelation in the product $uv$, i.e., $|uv| = |u| + |v|$. In this case $C(u) = \{ w \in F \mid w = u \circ v, v \in F \}$. 

13
Proposition 3.5. Let $S$ be a Schreier transversal for $C$, so $S = S_{T^*}$ for some spanning subtree $T^*$ of $\Gamma^*$. Then the following hold:

1) $S_{\text{int}}$ is a basis of $C$, in particular, $|S_{\text{int}}| = |V(\Gamma)|$.
2) $S_{\text{ext}}$ is the union of finitely many coni $C(u)$, where $v_u \in \partial^+ \Gamma$.
3) $S_{\text{sin}}$ is contained in a finite union of double cosets $Cs_1s_2^{-1}C$ of $C$, where $s_1, s_2 \in S_{\text{int}}$.
4) $S_{\text{uns}}$ is a finite union of left cosets of $C$ of the type $s_1s_2^{-1}C$, where $s_1, s_2 \in S_{\text{int}}$.

Proof. 1) is well-known, see [12], for example. 2) follows immediately from the construction. To see 3) notice first that $S_{\text{sin}} \subseteq N^*_F(C)$ and $N^*_F(C)$ is the union of finitely many double cosets $CsC$, where $s \in S_{\text{sin}}$, and furthermore, every such coset has the form $CsC = Cs_1s_2^{-1}C$, where $s_1, s_2 \in S_{\text{int}}$ (see Lemma 5 in [12], or Propositions 9.8 and 9.11 in [33], or Theorem 2 in [32]).

To see 4) assume that $s \in S$ is not stable, so there exists an element $c \in C$ such that $sc \notin S$. Then $s = s_1 \circ t$, $c = t^{-1} \circ d$, $sc = s_1 \circ d$. We claim, that the terminal vertex of $s_1$ lies in $\Gamma$ (viewing $s$ as a path in $\Gamma^*$). Indeed, if not, then $s_1$, as well as $s_1 \circ d$, is in $S_{\text{ext}}$ - contradiction. Hence, $s_1 \in S_{\text{int}}$. Since $c = t^{-1} \circ d$ there is a closed path in $\Gamma$ with the label $t^{-1} \circ d$, starting at $1_C$. Let $s_2 \in S_{\text{int}}$ be the representative of $t^{-1}$. Then $s_2d = c_1 \in C$, hence $sc = s_1 \circ d = s_1s_2^{-1}c_1$, so $s \in s_1s_2^{-1}C$, as claimed.

Proposition 3.6. Let $S$ be a Schreier transversal for $C$, so $S = S_{T^*}$ for some spanning subtree $T^*$ of $\Gamma^*$. Then the following hold:

1) If $T^*$ is a geodesic subtree of $\Gamma^*$ (and hence $T$ is a geodesic subtree of $\Gamma$) then $S$ is a geodesic transversal.
2) If $C$ is a malnormal subgroup of $F$ then $S_{\text{sin}} = \emptyset$.
3) $S_{\text{sin}} \subseteq S_{\text{uns}}$.

Proof. 1) is straightforward (see also [33]).

2) If $C$ is malnormal then $N^*_F(C) = 1$, so $S_{\text{sin}} = \emptyset$.

3) If $s \in S_{\text{sin}}$ then $c = s^{-1}c_1s$ for some non-trivial $c, c_1 \in C$, so $c_1s = sc$. Since $sc \neq s$ we conclude that $sc \notin S$, hence $s \in S_{\text{uns}}$. 

14
4 Measuring subsets of $F$

Recall that a finite automaton $A$ is a finite labeled oriented graph (possibly with multiple edges and loops). We refer to its vertexes as states. Some of the states are called initial states, some accept or final states. We assume further that every edge of the graph is labeled by one of the symbols $x^\pm 1$, $x \in X$, where $F = F(X)$ is a free group of finite rank $m$. The language accepted by an automaton $A$ is the set $L = L(A)$ of labels on paths from initial to accept states. An automaton is said to be deterministic if, for any state there is at most one arrow with a given label exiting the state. A regular set is a language accepted by a finite deterministic automaton.

The following facts about regular sets are well known. Let $A$ and $B$ are regular subsets in $F$. Then:

- the sets $A \cup B$, $A \cap B$ and $A \setminus B$ are regular.
- The prefix closure $\overline{A}$ of a regular set $A$ is regular. Here, the prefix closure $\overline{A}$ is the set of all initial segments of all words in $A$.
- If $ab = a \circ b$ for any $a \in A, b \in B$ then $AB$ is regular.
- If $ab = a \circ b$ for any $a, b \in A$ then $A^*$ is regular.

The following notation is useful. For $u, v \in F$ define

$$c(u, v) = \frac{1}{2}(|u| + |v| - |uv|)$$

- the amount of cancelation in the product $uv$.

**Proposition 4.1.** Let $R_1$ and $R_2$ be subsets of $F$. Then the following statements hold:

1) If $R_1 \subseteq R_2$ and $R_2$ is negligible (exponentially negligible) then so is $R_1$.

2) If $R_1, R_2$ are negligible (exponentially negligible) then so is $R_1 \cup R_2$.

3) If $R_1$ and $R_2$ are negligible (exponentially negligible) then so is the set

$$R_1 \circ R_2 = \{r_1r_2 \mid r_i \in R_i, \ c(r_1, r_2) = 0\}.$$

4) If $R_1$ and $R_2$ are negligible (exponentially negligible) then so is the set

$$R_1 \circ t R_2 = \{r_1r_2 \mid r_i \in R_i, \ c(r_1, r_2) \leq t\}.$$
Proof. The proof is straightforward.

Definition 4.2. Let $R_1$ and $R_2$ be subsets of $F$ and $f : R_1 \to R_2$ a map. Then:

- $f$ is called \textit{d-isometry}, where $d$ is a non-negative real number, if for any $w \in R_1$
  \[|w| - d \leq |f(w)| \leq |w| + d.\]

- $f$ has \textit{uniformly bounded fibers} if there exists a constant $c$ such that every element $w \in R_2$ has at most $c$ pre-images in $R_1$.

Proposition 4.3. Let $R_1$ and $R_2$ be subsets of $F$. Then the following statements hold:

1) If $f : R_1 \to R_2$ is a surjective $d$-isometry and $R_1$ is negligible (exponentially negligible) then so is $R_2$.

2) If $f : R_1 \to R_2$ is a $d$-isometry with uniformly bounded fibers and $R_2$ is negligible (exponentially negligible) then so is $R_1$.

Proof. Notice that for $k > d$
\[f_k(R_2) \leq \sum_{j=k-d}^{k+d} f_j(R_1),\]
and 1) follows. Similarly,
\[f_k(R_1) \leq c \sum_{j=k-d}^{k+d} f_j(R_2)\]
for $k > d$ and 2) follows.

Proposition 4.4. Let $C$ be a finitely generated subgroup of infinite index of a free group $F$. Then every Schreier transversal of $C$ in $F$ is regular and thick.

Proof. By definition $S = S_{\text{int}} \cup S_{\text{ext}}$. By Proposition \ref{prop:finite_index} the set $S_{\text{int}}$ is finite and the set $S_{\text{ext}}$ is a non-empty finite union of cones. By Theorem \ref{thm:thick_cones} each cone in $S_{\text{ext}}$ is thick. Therefore, the set $S_{\text{ext}}$, as well as the set $S$, is thick. Clearly, every cone is regular, so is the set $S$.

Proposition 4.5. Let $C$ be a finitely generated subgroup of infinite index in $F$. Then the following hold:
1) $C$ is exponentially negligible in $F$ and one can find some upper bound $\delta < 1$ for the growth rate of $C$.

2) Every coset of $C$ in $F$ is exponentially negligible in $F$.

*Proof.* 1) follows from the Proposition 1 and Corollary 1 in \[6\].

2) follows from 1) above and 4) from Proposition 4.1. \hfill $\Box$

**Proposition 4.6.** Let $C$ be a finitely generated subgroup of infinite index in $F$. Then the following hold:

1) $C^* = \bigcup_{f \in F} C^f$ is exponentially negligible in $F$.

2) For every $c \in C$ the set conjugacy class $c^F = \{ f^{-1}cf | f \in F \}$ is exponentially negligible in $F$.

*Proof.* The statement 1) has been shown in \[14\] and also in Proposition 1.10 in \[5\]. The statement 2) is shown in Proposition 1.11 in \[5\]. \hfill $\Box$

**Proposition 4.7.** Let $C$ be a finitely generated subgroup of infinite index in $F$ and $S$ is a Schreier transversal of $C$ in $F$. If $S_0 \subseteq S$ is a exponentially negligible subset of $F$ then the set $\bigcup_{s \in S_0} Cs$ is exponentially negligible in $F$.

*Proof.* By Proposition 3.5 $S = S_{\text{int}} \cup S_{\text{ext}}$, where $S_{\text{int}}$ is a finite set and $S_{\text{ext}}$ is a union of finitely many cones $C(u), u \in \partial^+ \Gamma$. It suffices to prove the result for $S_0 \cap C(u)$ for a fixed $u \in \partial^+ \Gamma$. To this end we may assume from the beginning that $S_0 \subseteq C(u)$. If $s$ is the representative of $u$ in $S$ then every word from $C(u)$ contains $s$ as an initial segment. Since $s$ is not readable in $\Gamma$ the amount of cancelation $c(w, t)$ in the product $wt$, where $w \in C$ and $t \in C(u)$ does not exceed the length of $s$. Hence

$$CS_0 = C \circ \upharpoonright S_0$$

and the result follows from the statement 4) of Proposition 4.1. \hfill $\Box$

**Proposition 4.8.** Let $A$ and $B$ be finitely generated subgroups of infinite index in $F$. Then for any $w \in F$ the double coset $AwB$ is exponentially negligible in $F$.

*Proof.* Observe, that $AwB = AB^{w^{-1}} w$, so by the statement 2 of Proposition 4.5 it suffices to show that $AB^{w^{-1}}$ is exponentially negligible. Since $B^{w^{-1}}$ is just another finitely generated subgroup of infinite index in $F$ one can assume

17
from the beginning that \( w = 1 \). Let \( S \) be a geodesic Schreier transversal for \( A \) in \( F \). Then
\[
AB = \bigcup_{s \in S_0} As
\]
for some subset \( S_0 \subseteq S \). By Proposition 4.7 it suffices to show that the subset \( S_0 \) is exponentially negligible. Since the set \( S_{\text{int}} \) is finite we may assume that \( S_0 \subseteq S_{\text{ext}} \). Let \( T_A \) be the spanning subtree of \( \Gamma_A \) such that \( S = S_{T_A} \) and \( T_B \) be a spanning geodesic subtree of \( \Gamma_B \). Denote by \( d \) the maximum of the diameters of the trees \( T_A \) and \( T_B \). To describe the map \( \alpha \) choose an arbitrary element \( s \in S_0 \). Without loss of generality assume that \( |s| \geq d \), because there are only finitely many such \( s \) that have smaller length and by Proposition 4.5 they will not extremely change asymptotic size of \( AB \) since \( A \) of infinite index in \( F \). Then \( as = b \) for some \( a \in A \) and \( b \in B \). We claim that there exists an element \( b_s \in B \) such that \( |sb_s^{-1}| \leq 2d \). Indeed, the cancelation in the product \( as \) is at most \( d \) (see the argument in Proposition 4.7). Hence \( s \) and \( b \) have a common terminal segment \( t \) of length at least \( |s| - d \) (recall that \( |s| \geq d \)). It follows that in the graph \( \Gamma_B \) there exists a path from some vertex \( v \) to \( 1_B \) with the label \( t_v \). Then \( b_s = t_v t \in B \) and \( |sb_s^{-1}| = |st^{-1}t_v^{-1}| \leq 2d \). Hence \( s \) and \( b_s \) has a long common terminal segment and differ only on the initial segment of length at most \( 2d \). It follows that the map \( \alpha : s \to b_s \) gives a \( 2d \)-isometry \( \alpha : S_0 \to B \). Notice that \( \alpha \) has uniformly bounded fibers. Indeed, if \( \alpha(s_1) = \alpha(s_2) = b \) then \( s_1 \) and \( s_2 \) differ from each other, only on the initial segment of length at most \( 2d \). So there are at most \( (2d)^{|X|} \) such distinct elements. Since \( B \) is exponentially negligible by Proposition 4.3 the set \( S_0 \) is also exponentially negligible, as claimed. This proves the result. Notice, that the property being geodesic for Schreier transversal \( S \) for \( A \) in \( F \) is not crucial for our prove. Namely, for arbitrary Shreier transversal \( S \) all conclusions can be repeated with slightly different constant. \( \square \)

Now we can state the main result of the section.

**Theorem 4.9.** Let \( C \) be a finitely generated subgroup of infinite index in \( F \) and \( S \) a Schreier transversal for \( C \). Then the following hold:

1) The generalized normalizer \( N_F^*(C) \) of \( C \) in \( F \) is exponentially negligible in \( F \).

2) The set of singular representatives \( S_{\text{sin}} \) is exponentially negligible in \( F \).
3) The set $S_{\text{uns}}$ of unstable representatives is exponentially negligible in $F$.

Proof. To see 1) recall that the generalized normalizer $N^*_F(C)$ of $C$ in $F$ is a finite union of double cosets of $C$ in $F$. Therefore $N^*_F(C)$ is exponentially negligible in $F$ by Proposition 4.8.

2) follows immediately from 1).

To prove 3) observe that $S_{\text{uns}}$ is a finite union of left cosets of $C$ (see 3) in Proposition 3.5). Now the result follows from Proposition 4.5.

Theorem 4.9 can be strengthened as follows.

Corollary 4.10. Let $C$ be a finitely generated subgroup of infinite index in $F$. Then the sets

$$Sin(C) = \bigcup_S S_{\text{sin}}, \quad Uns(C) = \bigcup_S S_{\text{uns}},$$

where $S$ runs over all Schreier transversals of $C$, are exponentially negligible.

Proof. By Corollary 3.3 there are only finitely many Schreier transversals of $C$. Now the result follows from Theorem 4.9 and Proposition 4.11.

5 Comparing sets at infinity

5.1 Comparing Schreier representatives

In this section we give another version of Theorem 4.9. To explain we need a few definitions.

For subsets $R, L$ of $F$ we define their size ratio at length $k$ by

$$f_k(R, L) = \frac{f_k(R)}{f_k(L)} = \frac{|R \cap S_k|}{|L \cap S_k|}.$$

The size ratio $\rho(R, L)$ at infinity of $R$ relative to $L$ (or the relative asymptotic density) is defined by

$$\rho(R, L) = \limsup_{k \to \infty} f_k(R, L).$$

By $r_L(R)$ we denote the cumulative size ratio of $R$ relative to $L$:

$$r_L(R) = \sum_{k=1}^{\infty} f_k(R, L).$$
We say that $R$ is $L$-measurable, if $r_L(R)$ is finite. $R$ is called negligible relative to $L$ if $\rho(R, L) = 0$. Obviously, an $L$-measurable set is $L$-negligible.

A set $R$ is termed exponentially negligible relative to $L$ (or exponentially $L$-negligible) if $f_k(R, L) \leq q^k$ for all sufficiently large $k$.

The following result is simple, but useful.

**Proposition 5.1.** Let $R$ be an exponentially negligible set in $F$.

1) For any $w \in F$ the set $R$ is a exponentially negligible relative to the cone $C(w)$.

2) The set $R$ is exponentially negligible relative to any exponentially generic subset $T$ of $F$.

**Proof.** Observe, that $f_k(C(w)) = 1/2m(2m-1)^{|w|-1}$ is a constant. Since

$$f_k(R, C(w)) = \frac{f_k(R)}{f_k(C(w))}$$

it follows that $R$ is exponentially negligible relative to $C(w)$. This proves 1).

To prove 2) denote by $p$ and $q$ the corresponding rate bounds for $R$ and $T$, so $f_k(R) \leq p^k, f_k(T) \geq 1 - q^k$ for sufficiently large $k$. Then, for such $k$,

$$f_k(R, T) = \frac{f_k(R)}{f_k(T)} \leq \frac{p^k}{1 - q^k} = \left(\frac{p}{1 - q^k}\right)^k.$$ 

Since

$$\lim_{k \to \infty} \frac{p}{(1 - q^k)^k} = p$$

it follows that for any $\varepsilon > 0$

$$f_k(R, T) \leq (p + \varepsilon)^k$$

for sufficiently large $k$, as claimed.

**Corollary 5.2.** Let $C$ be a finitely generated subgroup of infinite index in $F$ and $S$ a Schreier transversal for $C$. Then the following hold:

1) The set of singular representatives $S_{\sin}$ is exponentially negligible in $S$. 

20
2) The set $S_{uns}$ of unstable representatives is exponentially negligible in $S$.

Proof. The statements of this corollary follow immediately from Theorem 4.9 and Propositions 3.5, 3.6 and 5.1.

5.2 Comparing regular sets

In this section we give an asymptotic classification of regular subsets of $F$ relative to a fixed prefix-closed regular subset $L \subseteq F$.

For this purpose we are going to describe how one can use a random walk on the finite automaton $B$ recognizing regular subset $R \subseteq L$ similar to the one in Section 2.3. It will be convenient to further put $B$ to special form consistent to $L$.

Recall Myhill-Nerode’s theorem on regular languages (see, for example, [21], Theorem 1.2.9.) For a language $R$ over an alphabet $A$ consider an equivalence relation $\sim$ on $A^*$ defined as follows: two strings $w_1$ and $w_2$ are equivalent if and only if for each string $u$ over $A$ the words $w_1u$ and $w_2u$ are either simultaneously in $R$ or not in $R$. Then $R$ is regular if and only if there are only finitely many $\sim$-equivalence classes.

Now, let $R \subseteq L$. Define an equivalence relation $\sim$ on $L$ such that $w_1 \sim w_2$ if and only if for each $u \in F$ the following condition holds: $w_1u = w_1 \circ u$ and $w_1u \in R$ if and only if $w_2u = w_2 \circ u$ and $w_2u \in R$.

The following is an analog of Myhill-Nerode’s theorem for free groups.

Lemma 5.3. Let $R \subseteq L \subseteq F$ and $L$ prefix-closed and regular. Then $R$ is regular if and only if there are only finitely many $\sim$-equivalence classes in $L$.

Proof. The proof is similar to the original one. We give a short sketch of the most interesting part of it. If the set of the equivalence classes is finite one can define an automaton $B$ on the set of equivalence class as states. If $x \in X \cup X^{-1}$ and $[w]$ is the equivalence class of some $w$ such that $w \circ x \in R$ then one connects the state $[w]$ with an edge labeled by $x$ to the state $[wx]$. The class $[\varepsilon]$, where $\varepsilon$ is the empty word, is the initial state, while a state $[w]$ is an accepting state if and only if $w \in R$. In this case $L(B) = R$.

Since $R$ is regular, we suppose that $B$ as in Lemma 5.3 and modify it in the next way. Without loss of generality we can assume that $A$ is in the normal form, i.e., it has only one initial state $I$ and doesn’t contain inaccessible states.
Pic. 3. Splitting the states of the automaton \( \mathcal{B} \).

Let \( S = [w] \) be a state of \( \mathcal{B} \). Denote by \( S^{pr} \) the uniquely defined state in \( \mathcal{A} \) which is the terminal state of the path with the label \( w \) in \( \mathcal{A} \), starting at \([\varepsilon]\). The state \( S^{pr} \) is well-defined, it does not depend on the choice of \( w \). We call \( S^{pr} \) the prototype of \( S \).

Since \( \mathcal{B} \) accepts only reduced words in \( X \cup X^{-1} \) one can transform \( \mathcal{B} \) to a form where the following hold:

a) \( \mathcal{B} \) has only one initial state \( I \) and one accepting state \( Z \).

b) For any state \( S \) of \( \mathcal{B} \), all arrows which enter \( S \) have the same label \( x \in X \cup X^{-1} \) and arrows exiting from \( S \) cannot have label \( x^{-1} \) (this can be achieved by splitting the states of \( \mathcal{B} \), see Pic. 3). We shall say in this situation that \( S \) has type \( x \).

c) For every state \( S \) of \( \mathcal{B} \) there is a direct path from \( S \) to the accept state \( Z \).

d) There are no arrows entering the initial state \( I \).

The final version of obtained automaton \( \mathcal{B} \) we will call an automaton consistent with \( \mathcal{A} \).

Now we are ready to define a no-return random walk on \( \mathcal{B} \) as it was claimed above. Namely, let \( \mathcal{B} \) be consistent with \( \mathcal{A} \) and let \( S \) be a state in \( \mathcal{B} \). Denote by \( \nu = \nu(S^{pr}) \) the number of edges exiting from the prototype state \( S^{pr} \) in \( \mathcal{A} \). The walker moves from \( S \) along some outgoing edge with the uniform probability \( \frac{1}{\nu} \). In this event, the probability that the walker hits an element \( w \in \mathcal{R} \in |w| \) steps (when starting at \([\varepsilon]\)) is the product of
frequencies of arrows in a direct path from the initial state $I$ to the accept state $Z$ with the label $w$. This gives rise to the measure $\lambda$ on $R$:

$$
\lambda_L(R) = \sum_{w \in R} \lambda_L(w) = \sum_{k=0}^{\infty} f'_k(R, L),
$$

where

$$
f'_k(R, L) = \sum_{w \in R \cap S_k} \lambda_L(w).
$$

Note that, generally speaking, $f'_k(R, L)$ differs from $f_k(R, L)$ defined in section 5.1. Indeed, walking in $\mathcal{B}$ we have different number of possibilities to continue our walk on the next step depending on way we chose. On the other hand, $f'_k(R, F(X)) = f_k(R, F(X))$.

Now we can use the tools of random walks to compute $\lambda_L(R)$. Notice, that $\lambda_L$ is multiplicative, i.e.,

$$
\lambda_L(uw) = \lambda_L(u)\lambda_L(v)
$$

for any $u, v \in R$ such that $uw = u \circ v$ and $uw \in R$. We say that $R$ is $\lambda_L$-measurable, if $\lambda_L(R)$ is finite. A set $R$ is termed exponentially $\lambda_L$-measurable if $f'_k(R, L) \leq q^k$ for all sufficiently large $k$.

The following result is simple, but useful.

Let $w \in F$. The set $C_L(w) = L \cap C(w)$ is called an $L$-cone. Obviously, $C_L(w)$ is a regular set. We say that $C_L(w)$ is $L$-small, if it is exponentially $\lambda_L$-measurable.

The following is the main result of this section.

**Theorem 5.4.** Let $R$ be a regular subset of a prefix-closed regular set $L$ in a free group $F$. Then either the prefix closure $\overline{R}$ of $R$ in $L$ contains a non-small $L$-cone or $\overline{R}$ is exponentially $\lambda_L$-measurable.

Before proving the theorem we establish a few preliminary facts. We fix a prefix-closed regular subset $L$ of $F$.

**Proposition 5.5.** Let $R_1$ and $R_2$ be subsets of $F$. Let also $P$ be one of the properties \{”to be $L$-measurable”, ”to be exponentially $L$-negligible”, ”to be $\lambda_L$-measurable”, ”to be exponentially $\lambda_L$-measurable”\}. Then the following hold:

1) If $R_1 \subseteq R_2$ and $R_2$ has property $P$ then so is $R_1$.

2) If $R_1, R_2$ have property $P$ then so is $R_1 \cup R_2$.
3) If $R_1$ and $R_2$ have property $P$ then so is the set

$$R_1 \circ R_2 = \{ r_1 r_2 \mid r_i \in R_i, \ c(r_1, r_2) = 0 \}.$$  

**Proof.** The proofs are easy.  

To strengthen the last statement in Proposition 5.5 we need the following notation. For a subset $T \subseteq F$ put $T^0 = T$ and define recursively $T_k^0 = T_k^0 \circ T$. Denote

$$T_k^\infty = \bigcup_{k=1}^\infty T_k^0.$$  

**Lemma 5.6.** Let $T$ be a regular set and a number $q$, $0 < q < 1$, such that $f_k'(T, L) \leq q^k$ for every positive integer $k$. Then the set $T_\infty$ is exponentially $\lambda_L$-measurable.  

**Proof.** Every word $w \in T_\infty$ of length $k$ comes in the form $w = u_1 \circ u_2 \circ \ldots \circ u_t$, where $u_i$’s are non-trivial elements from $T$ and $k = |u_1| + \ldots + |u_t|$. On the other hand, if $k = k_1 + \ldots + k_t$ is a partition of $k$ into a sum of positive integers and $u_1, \ldots, u_t$ are words in $T$ such that $u_i = k_i$, then $w = u_1 \ldots u_t \in T_\infty$. Since $f_k'(T_\infty, L)$ is multiplicative every partition of $k$ adds to $f_k'(T_\infty, L)$ a number $f_{k_1}'(T_\infty, L) \ldots f_{k_t}'(T_\infty, L)$, which is bounded from above by $q^{k_1+\ldots+k_t} = q^k$. If $p(k)$ is the number of all partitions of $k$ into a sum of positive integers then $f_k'(T_\infty, L) \leq p(k) q^k$. It is known (Hardy and Ramanujan) that

$$p(k) \sim \frac{e^{\pi \sqrt{2k}}}{4k\sqrt{3}}.$$  

Hence $f_k'(T_\infty, L) < q_t^k$, for some $0 < q < q_1 < 1$ and all sufficiently large $k$, so $T_\infty$ is exponentially $\lambda_L$-measurable, as claimed.  

**Proof of Theorem 5.4.** In the most part we follow the proof of Theorem 2.2 from [14]. Suppose that all $L$-cones in $R$ are non-small. Since $R \subseteq \overline{R}$ by Proposition 5.5 we can assume that $R$ itself is prefix-closed in $L$. We have to prove that $R$ is exponentially $\lambda_L$-measurable. Let $R = L(B)$ and $B$ consistent to $A$ (where $A$ recognize $L$).

It is convenient to further split $B$ into two parts. Denote by $B_1$ the automaton obtained from $B$ by removing all arrows exiting from $Z$.  

24
Let $B_2$ be the automaton formed by all states in $B$ that are accessible from the state $Z$, with the same arrows between them as in $B$; $Z$ is the only initial and accepting state of $B_2$.

We assign to arrows in $B_1$ and $B_2$ the same frequencies as to the corresponding arrows in $B$. If $R_1$ and $R_2$ are the languages accepted by $B_1$ and
$\mathcal{B}_2$ then, obviously, $R = R_1 \circ R_2$. By Proposition 5.5 to prove the theorem it suffices to show that $R_1$ and $R_2$ are exponentially $\lambda_L$-measurable.

**Claim.** The set $R_2$ is exponentially $\lambda_L$-measurable.

**Proof of the claim.** Notice, that for every $w \in R_1$ $w \circ R_2 \subseteq L(A) = R$ and $w \circ R_2$ is an $L$-cone. It is easy to see, that $R_2$ is exponentially $\lambda_L$-measurable if and only if so $w \circ R_2$ is.

Let $R_3 \subseteq R_2$ be the subset consisting of those non-trivial words $w \in R_2$, whose paths $p_w$ visit the state $Z$ of $\mathcal{B}_2$ only once. The set $R_3$ is regular - it is accepted by an automaton $\mathcal{B}_3$, which is obtained from $\mathcal{B}_2$ by splitting the state $Z$ into two separate states: the initial state $Z_1$ and an accepting state $Z_2$, in such a way that the edges exiting from $Z$ in $\mathcal{B}_2$ are now exiting from $Z_1$ and there no ingoing edges at $Z_1$, while there are no edges exiting from $Z_2$ and all those arrows incoming in $Z$ in $\mathcal{B}_2$ are now incoming into $Z_2$.

It follows immediately from the construction, that

$$R_2 = \{\varepsilon\} \bigcup R_3 \bigcup (R_3 \circ R_3) \bigcup (R_3 \circ R_3 \circ R_3) \bigcup \ldots = (R_3)_\infty$$

so

$$\lambda_L(R_2) = \lambda_L(R_3) + (\lambda_L(R_3))^2 + (\lambda_L(R_3))^3 + \ldots \quad (8)$$

By Proposition 5.6 it suffices to show that there is a number $q, 0 < q < 1$, such that $f'_k(R_3, L) \leq q^k$ for every $k$ (not only for sufficiently large $k$). It is not hard to see that this condition holds if $R_3$ is exponentially $\lambda_L$-measurable and $\lambda_L(R_3) < 1$, so it suffices to prove the latter two statements.

By our assumption all $L$-cones in $R = \mathcal{R}$ are $L$-small. If for every state $[w] = S$ in $\mathcal{B}_2$ and every $x \in X \cup X^{-1}$ there is an outgoing edge
labeled by $x$ at $[w]$ if and only if the same holds for the state $S^{pr}$ in $A$ (i.e., $B_2$ is $X$-complete relative to $A$) then for every given $w \in R_1$ one has $C(w) \cap R = w \circ R_2 = C(w) \cap L$, so $w \circ R_2$ is an $L$-cone. Hence it is $L$-small, i.e., exponentially $\lambda_L$-measurable, but then the set $\overline{R_2}$, hence $R_2$, is exponentially $\lambda_L$-measurable, as claimed.

This implies that for some state $S$ there are less then $\nu = \nu(S^{pr})$ arrows exiting from $S$. Consider a finite Markov chain $\mathcal{M}$ with the same states as in $B_3$ together with an additional dead state $D$. We set transition probabilities from $Z_2$ to $Z_2$ and from $D$ to $D$ being equal 1. Every arrow from a state $S$ in $B_3$ gives the corresponding transition from the state $S$ in $\mathcal{M}$ which we assign the transition probability $\frac{1}{\nu}$. If at some state $S$ of type $x$ in $B_3$ there is no exiting arrow labeled $y \in (X \cup X^{-1}) \setminus \{x^{-1}\}$, in $\mathcal{M}$ we make a transition from $S$ to $D$ with the transition probability $\frac{1}{\nu}$. This describes $\mathcal{M}$.

![Pic. 8. An automaton $\mathcal{M}$.

The states $Z_2$ and $D$ of Markov chain $\mathcal{M}$ are absorbing, and all other states are transient. The probability distribution on $\mathcal{M}$ concentrated at the initial state $Z_1$, converges to the steady state $P$ which is zero everywhere with the exception of the two absorbing states $Z_2$ and $D$. Obviously, $P(Z_2) = \lambda_L(R_3)$. Since $P(D) \neq 0$ we have $\lambda_L(R_3) = P(Z_2) < 1$, so one
of the required conditions on $R_3$ holds (for more details on this proof we refer to [14,41]). The other one follows directly from Corollary 3.1.2 in [41], which claims that in this case $R_3$ is exponentially $\lambda_L$-measurable. This proves the claim.

A similar argument shows that $R_1$ is exponentially $\lambda_L$-measurable. This proves the theorem.

Acknowledgements. The authors thank Alexandre Borovik for very fruitful discussions.

References

[1] N. Alon and J. Spencer, *The Probabilistic method*, Wiley, Interscience Series in Discrete Mathematics and Optimization, 2000.

[2] G. Arzhantseva and A. Olshanskii, *Generality of the class of groups in which subgroups with a lesser number of generators are free*, (Russian) Mat. Zametki 59 (1996), no. 4, pp. 489 - 496; translation in: Math. Notes 59 (1996), nos. 3-4, pp. 350 – 355.

[3] G. Arzhantseva, *On groups in which subgroups with a fixed number of generators are free*, (Russian) Fundam. Prikl. Mat. 3 (1997), no. 3, pp. 675 - 683.

[4] G. Arzhantseva, *Generic properties of finitely presented groups and Howson’s theorem*, Comm. Algebra 26 (1998), pp. 3783 - 3792.

[5] Ya. S. Averina and E. V. Frenkel, *The sparsity of cosets in a free group*, (Russian) Vestnik Omskogo Universiteta 3 (2004), no. 3, pp. 3 – 5.

[6] Ya. S. Averina and E. V. Frenkel, *On strictly sparse subsets of a free group*, (Russian) Siberian Electronic Mathematical Reports (2005), vol. 2 pp. 1 – 13, http://semr.math.nsc.ru

[7] O. Bogopolski, E. Ventura, *The mean Dehn function of abelian groups*, arXiv:math/0606273v1 [math.GR].

[8] J. Burillo, E. Ventura, *Counting primitive elements in free groups*, Geom. Dedicata 93 (2002), pp. 143 - 162.

[9] C. Champetier, *Propriétés statistiques des groupes de présentation finie*, Adv. Math. 116 (1995), pp. 197 - 262.
[10] C. Champetier, *The space of finitely generated groups*, Topology 39 (2000), pp. 657 - 680.

[11] P.-A. Cherix and G. Schaeffer, *An asymptotic Freiheitssatz for finitely generated groups*, Enseign. Math. 44 (1998), no. 2 pp. 9 - 22.

[12] G. Baumslag, A. G. Myasnikov and V. N. Remeslennikov, *Malnormality is decidable in free groups*, Intern. J. of Algebra and Computation, 9 (1999), no. 6, pp. 687 – 692.

[13] A. Borovik, A. Myasnikov, V. Shpilrain *Measuring sets in infinite groups*, Computational and Statistical Group Theory. Amer. Math. Soc., Contemporary Math. 298 (2002), pp. 21 – 42.

[14] A. V. Borovik, A. G. Myasnikov and V. N. Remeslennikov, *Multiplicative measures on free groups*, Intern. J. of Algebra and Computation, 13 (2003), no. 6, pp. 705 – 731.

[15] A. V. Borovik, A. G. Myasnikov and V. N. Remeslennikov, *The Conjugacy Problem in Amalgamated Products I: Regular Elements and Black Holes*, Intern. J. of Algebra and Computation, 17 (2007), no. 7, pp. 1301 - 1335

[16] A. V. Borovik, A. G. Myasnikov and V. N. Remeslennikov, *Algorithmic stratification of the conjugacy problem in Miller’s groups*, International Journal of Algebra and Computation vol. 17, nos. 5 & 6 (2007), pp. 963-997.

[17] Alexandre V. Borovik and Alexei G. Myasnikov *Quotient tests and random walks in computational group theory*, Topological and asymptotic aspects of group theory, Contemp. Math., Providence, RI, Amer. Math. Soc. (2006), pp. 31 – 45.

[18] S. Cleary, M. Elder, A. Rechnitzer, J. Taback, *Random subgroups of Thompsons group F*, arXiv:0711.1343v2 [math.GR].

[19] N. Chomsky and M. P. Schutzenberger, *The algebraic theory of context-free languages*, Computer Programming and Formal Systems (P. Bradford and D. Hirschberg, eds.), North-Holland, Amsterdam (1963), pp. 118 – 161.

[20] S.B.Cooper, *Computability Theory*, Chapman and Hall, CRC Mathematics, 2003.
[21] D. Epstein, J. Cannon, D. Holt, S. Levy, M. Paterson and W. Thurston, *Word Processing in Groups*, Jones and Bartlett, Boston, 1992.

[22] P. Flajolet and R. Sedgwick, *Analytic Combinatorics: Functional Equations, Rational and Algebraic Functions*, Res. Rep. INRIA RR4103 (2001), January, p. 98.

[23] E. Frenkel, A. G. Myasnikov and V. N. Remeslennikov, *Amalgamated products of groups II: Generation of random normal forms and estimates*, to appear.

[24] E.V. Frenkel, A. G. Myasnikov and V. N. Remeslennikov, *Amalgamated products of groups III: Generic complexity of algorithmic problems*, to appear.

[25] R. Gilman, A. G. Miasnikov, A. D. Myasnikov, A. Ushakov *Report on generic case complexity*, Herald of Omsk University, (2007), Special Issue pp. 103 – 110

[26] R. Gilman, A. Myasnikov, D. Osin, *Exponentially Generic Subsets of Groups*, to appear.

[27] M. Gromov, *Hyperbolic Groups*, Essays in Group Theory (G.M.Gersten, editor), MSRI publ. 8 (1987), pp. 75 - 263.

[28] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, Vol. 2 (Sussex, 1991), pp. 1295, London Math. Soc. Lecture Note Ser., 182, Cambridge Univ. Press, Cambridge, 1993.

[29] M. Gromov, *Random walks in random groups*, Geom. Funct. Analysis 13 (2003), no. 1, pp. 73 - 146.

[30] G. H. Hardy, *Divergent series*, Chelsea, 1991.

[31] J. Hopcroft, R. Motwani, J. Ullman, *Introduction to Automata Theory, Languages, and Computation*, 3rd ed., Addison-Wesley, Reading MA, 2006.

[32] T. Jitsukawa, *Malnormal subgroups of free groups*, Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001), Contemp. Math. 298, Amer. Math. Soc., Providence, RI (2002), pp. 83 – 95.
[33] I. Kapovich and A. G. Myasnikov, *Stallings foldings and subgroups of free groups*, J. of Algebra 248 (2002), pp. 608 – 668.

[34] I. Kapovich, A. Myasnikov, P. Schupp, V. Shpilrain *Generic-case complexity and decision problems in group theory*, J. of Algebra 264 (2003), pp. 665 – 694.

[35] I. Kapovich, A. Myasnikov, P. Schupp, V. Shpilrain *Average-case complexity for the word and membership problems in group theory*, Advances in Mathematics 190 (2005), pp. 343 – 359.

[36] I. Kapovich, P. Schupp and V. Shpilrain, *Generic properties of Whiteheads Algorithm and isomorphism rigidity of random one-relator groups*, Pacific J. Math. 223 (2006), no. 1, pp. 113 - 140.

[37] I. Kapovich and P. Schupp, *Genericity, the Arzhantseva-Olshanskii method and the isomorphism problem for one-relator groups*, Math. Ann. 331 (2005), no. 1, pp. 1 - 19.

[38] I. Kapovich and P. Schupp, *Delzants T-ivariant, one-relator groups and Kolmogorov complexity*, Comment. Math. Helv. 80 (2005), no. 4, pp. 911 - 933.

[39] I. Kapovich, I.Rivin, P. Schupp, V. Shpilrain, *Densities in free groups and $\mathbb{Z}^k$, visible points and test elements*, Math. Research Letters, 14 (2007), no. 2, pp. 263 – 284.

[40] J. G. Kemeny, J. L. Snell and A. W. Knapp, *Denumerable Markov chains*, D. van Nostrand, Princeton, 1966.

[41] J. G. Kemeny, J. L. Snell, *Finite Markov chains*, The University Series in Undergraduate Mathematics, Van Nostrand, Princeton, 1960

[42] E. G. Kukina, V. A. Roman’kov, *Subquadratic growth of the averaged Dehn function for free Abelian groups*, Siberian Mathematical Journal, Vol. 44, no. 4, (2003), pp. 605 - 610.

[43] R. C. Lyndon and P. Schupp, *Combinatorial group theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete vol. 89, Springer-Verlag, Berlin, Heidelberg, New York, 1977.

[44] W. Magnus, A. Karras and D. Solitar, *Combinatorial Group Theory*, Interschience Publishers, New York, 1966.
C. F. Miller III, *On group-theoretic decision problems and their classification*, Ann. of Math. Studies 68 (1971), Princeton University Press, Princeton.

C. F. Miller III, *Decision problems for groups – Survey and reflections*, Algorithms and Classification in Combinatorial Group Theory (G. Baumslag and C. F. Miller III, eds.), Springer (1992), pp. 1 – 60.

A. G. Myasnikov, A. Ushakov, *Random subgroups and analysis of the length-based and quotient attacks*, Journal of Mathematical Cryptology 1 (2007), pp. 15 – 47.

A. G. Myasnikov, V. Shpilrain, A. Ushakov, *Advanced course on Group-Based Cryptography*, Quaderns, 42, CRM, Barcelona, 2007.

A. G. Myasnikov, V. Shpilrain and A. Ushakov, *Random subgroups of braid groups: an approach to cryptanalysis of a braid group based cryptographic protocol*, PKC 2006, Lecture Notes Comp. Sc. 3958 (2006), pp. 302-314.

Y. Ollivier, *Critical densities for random quotients of hyperbolic groups*. C. R. Math. Acad. Sci. Paris 336 (2003), no. 5, pp. 391 - 394.

A. Yu. Olshanskii, *Almost every group is hyperbolic*, Internat. J. of Algebra and Computation 2 (1992), pp. 1 - 17.

I. Rivin, *Counting Reducible Matrices, Polynomials, and Surface and Free Group Automorphisms*, arXiv:math/0604489v2.

V. A. Roman’kov, *Asymptotic growth of averaged Dehn functions for nilpotent groups*, Algebra and Logic, vol. 46, (2007), no. 1, pp. 37 – 45.

D. Ruinsky, A. Shamir, and B. Tsaban, *Cryptanalysis of group-based key agreement protocols using subgroup distance functions*, Advances in Cryptology – PKC 2007, vol. 4450 of Lecture Notes in Computer Science, Berlin, Springer (2007), pp. 61 – 75.

A. Shalev, *Probabilistic group theory*, Groups St. Andrews 1997 in Bath, II, London Math. Soc., Lecture Notes Ser. 261, Cambridge Univ. Press, pp. 648 – 679.
[56] R. Sharp, *Local limit theorems for free groups*. Math. Ann. 321 (2001), no. 4, pp. 889 - 904.

[57] R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, 1999.

[58] N. Touikan, *A Fast Algorithm for Stallings' Folding Process*, Intern. J. of Algebra and Computation 16, (2006), no. 6 pp. 1031 – 1045.

[59] A. Martino, T. Turner, and E. Ventura. The density of injective endomorphisms of a free group. Preprint, CRM, Barcelona, 2006.

[60] W. Woess, *Cogrowth of groups and simple random walks*, Arch. Math. 41 (1983), pp. 363 – 370.

[61] A. Zuk, *On property (T) for discrete groups*. Rigidity in dynamics and geometry (Cambridge, 2000), Springer, Berlin, 2002, pp. 473 - 482.