ISRAEL JOURNAL OF MATHEMATICS **157** (2007), 1–46
DOI: 10.1007/s11856-006-0001-7

THE SUBADDITIVE ERGODIC THEOREM AND GENERIC STRETCHING FACTORS FOR FREE GROUP AUTOMORPHISMS

BY

VADIM KAIMANOVICH

School of Engineering and Science, International University-Bremen
P.O. Box 750 561, 28725 Bremen, Germany
e-mail: v.kaimanovich@iu-bremen.de

AND

ILYA KAPOVICH AND PAUL SCHUPP*

Department of Mathematics, University of Illinois at Urbana-Champaign
1409 West Green Street, Urbana, IL 61801, USA
e-mail: kapovich@math.uiuc.edu, schupp@math.uiuc.edu
http://www.math.uiuc.edu/˜kapovich/, http://www.math.uiuc.edu/People/schupp.html

ABSTRACT

Let F_k be a free group of rank $k \geq 2$ with a fixed set of free generators. We associate to any homomorphism ϕ from F_k to a group G with a left-invariant semi-norm a *generic stretching factor*, $\lambda(\phi)$, which is a non-commutative generalization of the translation number. We concentrate on the situation where ϕ: $F_k \to Aut(X)$ corresponds to a free action of F_k on a simplicial tree X, in particular, where ϕ corresponds to the action of F_k on its Cayley graph via an automorphism of F_k. In this case we are able to obtain some detailed "arithmetic" information about the possible values of $\lambda = \lambda(\phi)$. We show that $\lambda \geq 1$ and is a rational number with $2k\lambda \in \mathbb{Z}[1/(2k-1)]$ for every $\phi \in Aut(F_k)$. We also prove that the set of all $\lambda(\phi)$, where ϕ varies over $Aut(F_k)$, has a gap between 1 and $1+(2k-3)/(2k^2-k)$, and the value 1 is attained only for "trivial" reasons. Furthermore, there is an algorithm which, when given ϕ, calculates $\lambda(\phi)$.

* The second and the third author were supported by the NSF grant DMS#0404991 and the NSA grant DMA#H98230-04-1-0115.
Received April 28, 2005

Contents

1. Introduction

1.1. Random subgroup distortion and growth of random automorphisms. Let G be a finitely generated group with a word metric d_S determined by a finite generating set S and write $|g|_S := d_S(1, g)$ for $g \in G$. Recall that if $H = \langle T \rangle$ is a subgroup of G generated by a finite set T, then a function f is said to be a *distortion function* of H in G if for every $h \in H$ we have $|h|_T \leq f(|h|_S)$. The subgroup H is *quasi-isometrically embedded* in G if and only if for some (and hence for all) choices of S, T there is a linear distortion function for H in G, that is, if the ratio $|h|_T / |h|_S$ is bounded on $H \setminus \{1\}$.

The *translation number* of an element $g \in G$ is defined as

$$\lambda(g) = \lambda_S(g) = \lim_{n\to\infty} \frac{|g^n|_S}{n}$$

and the limit exists by the subadditivity of the sequence $|g^n|_S$. If g has infinite order, then the cyclic subgroup $\langle g \rangle$ is quasi-isometrically embedded in G if and only if $\lambda_S(g) > 0$ for some (and hence for any) finite generating set S of G.

The study of "random" or "generic" behavior is currently an increasingly active area of investigation in many different group-theoretic contexts. (See, for example, [23, 43, 25, 10, 11, 12, 13, 14, 2, 3, 4, 1, 46, 32, 33, 35, 21, 42].) In this paper we concentrate on algebraic and geometric consequences of subadditivity, specifically of Kingman's Subadditive Ergodic Theorem.

We investigate a "noncommutative analogue" of the notion of a translation number which is defined for a "mapped-in" subgroup $H = \phi(F)$, where ϕ: $F \to G$ is a homomorphism of a free nonabelian group $F = F(A)$ of finite rank into a group G with generating set S. Namely, there is a number $\lambda = \lambda(\phi, A, S) \geq 0$ such that for long "random" freely reduced words $w \in F(A)$ we have $|\phi(w)|_S/|w|_A \approx \lambda$. (Instead of the word metric d_S one could actually take an arbitrary semi-norm on G.)

Throughout this paper we fix the notation that $F = F(A)$ is the free group with basis $A = \{a_1, \dots, a_k\}$ where $k \geq 2$. For any $w \in F$ let $|w|$ denote the length of the unique freely reduced word over $A^{\pm 1}$ representing w. We identify the hyperbolic boundary ∂F with the set of all *geodesic rays* from $1 \in F$ in the Cayley graph $\Gamma(F, A)$ of F, that is, ∂F is the set of all semi-infinite freely reduced words over $A^{\pm 1}$ endowed with the standard topology. The space ∂F can be identified with the *space of ends* or the *hyperbolic boundary* of F.

The Borel σ-algebra $\mathcal{F}$ on ∂F is generated by the *cylinder sets* $Cyl_A(v), v \in F$, where $Cyl_A(v)$ consists of all infinite rays $\omega \in \partial F$ that begin with v. The *uniform* Borel probability measure μ_A on ∂F corresponding to A is defined by assigning equal weights to all cylinders based on the words on the same length. That is,

$$\mu_A(Cyl_A(v)) = \frac{1}{2k(2k-1)^{|v|-1}} \quad \forall v \in F \setminus \{1\}.$$

Note that although the boundary ∂F could be defined without referring to a particular generating set A, the uniformity of the measure μ_A *does* depend on the choice of A. The fact that the uniform measures corresponding to two different free generating sets may well be singular respect to each other is actually the cornerstone of our approach. (See [20, 34] for a detailed discussion of this phenomenon in the general context of word-hyperbolic and free groups and of the Patterson–Sullivan measures corresponding to geometric actions of such groups on Gromov-hyperbolic spaces.)

A ray $\omega \in \partial F$ can be thought of as a *non-backtracking edge-path* in $\Gamma(F, A)$ starting from the identity 1. We denote by ω_n the vertex on ω at distance n from 1. The measure space $(\partial F, \mu_A)$ then becomes the *space of sample paths* of the *non-backtracking simple random walk* (NBSRW) on the Cayley graph of F. This is the Markov chain on F whose transition probabilities $\pi_w, w \in F$ are equidistributed among the neighbors of w which are strictly further from the group identity. By choosing a random μ_A-distributed point $\omega \in \partial F$ we may think about ω_n as a "random" (with respect to the NBSRW) freely reduced word of length n in F.

In Section 2 we prove:

THEOREM A: *Let $F = F(a_1, \dots, a_k)$ with $k \geq 1$, and let μ_A be the uniform Borel probability measure on ∂F corresponding to the basis $A = \{a_1, \dots, a_k\}$.*

Let ϕ: $F \to G$ be a homomorphism to a group G endowed with a semi-norm, that is, a nonnegative function $|\cdot|_G$ on G satisfying $|gh|_G \leq |g|_G + |h|_G$ for all $g, h \in H$.

Then:

(1) *There exists a real number $\lambda \geq 0$ such that*

$$\lim_{n\to\infty} \frac{|\phi(\omega_n)|_G}{n} = \lambda$$

μ_A-a.e. and in $L^1(\partial F, \mu_A)$.

(2) *Suppose further that the image group $\phi(F)$ is non-amenable, and that the sequence $b_n = \#\{g \in \phi(F) : |g|_G \leq n\}$ grows at most exponentially. Then $\lambda > 0$.*

Note that the requirement of at most exponential growth of the b_n is automatically satisfied if the group G is finitely generated, and $|\cdot|_G$ is the word metric on G determined by a finite generating set.

Theorem A says that for a long "random" freely reduced element $w \in F$ we have $|\phi(w)|_G/|w| \approx \lambda$. For this reason we shall call the number $\lambda = \lambda(\phi, A, |\cdot|_G)$, whose existence is provided by part (1) of Theorem A, the *generic stretching factor* of ϕ with respect to the free basis A of F and the semi-norm $|\cdot|_G$.

We deduce Theorem A from the fact that the sample paths of the usual simple random walk on the group F asymptotically follow geodesics and the well-known results on the linear rate of escape of random walks on groups [38], [29]. We also give an alternative direct argument proof of part (1) of Theorem A applying Kingman's Subadditive Ergodic Theorem [37]. Part (2) of Theorem A can also be proved using the results of Cohen [15], Grigorchuck [22] and Woess [45] on co-growth in groups.

Example 1.1 (Stretching factors for isometric actions)*:* A typical example of a semi-norm $|\cdot|_G$ comes from isometric group actions on metric spaces. Namely, let X be a metric space with basepoint $x \in X$. For an isometry g of X define $|g|_x := d(x, gx)$. The triangle inequality implies that $|g_1g_2|_x \leq |g_1|_x + |g_2|_x$, so that $|\cdot|_x$ is a semi-norm on $G = Isom(X)$. Suppose $F = F(a_1, \dots, a_k)$ acts by isometries on X by a homomorphism ϕ: $F \to G$. It is easy to see that in this case $\lambda(\phi, A, |\cdot|_x)$ does not depend on the choice of a base-point $x \in X$ and is determined by the action ϕ and the choice of the basis A of F. In this case we shall denote $\lambda(\phi, A, |\cdot|_x)$ by $\lambda(\phi, A)$, or just by $\lambda(\phi)$ if the choice of A is fixed.

Example 1.2 (Random Subgroup Distortion)*:* Let $H \le G$ be finitely generated groups with finite generating sets $A \subset H$ and $S \subset G$, respectively. Denote the associated length functions by $|\cdot|_G$ and $|\cdot|_H$. Now H is a quotient of $F = F(A)$. Let ϕ: $F(A) \to G$ be composition of this quotient map with the inclusion of H into G. Then $|\phi(w)|_H \le |w|$ for any $w \in F$. If the group H is non-amenable then by Theorem A for a long “random” freely reduced word w in $F(A)$

$$\frac{|\phi(w)|_H}{|w|} \approx \lambda_1 > 0 \quad \text{and} \quad \frac{|\phi(w)|_G}{|w|} \approx \lambda_2 > 0,$$

and therefore

$$\frac{|\phi(w)|_H}{|\phi(w)|_G} \approx \frac{\lambda_1}{\lambda_2},$$

where the constants $\lambda_1, \lambda_2 > 0$ do not depend on w. Thus, informally speaking, Theorem A implies that any nonamenable finitely generated subgroup H of a finitely generated group G generically has linear distortion in G.

Example 1.3 (Normal Forms)*:* Let G be a nonamenable group with a finite generating set A and the associated length function $|\cdot|_G$. We will denote by $\overline{w}$ the element of G represented by a word w in the alphabet $A \cup A^{-1}$.

Let $L \subseteq (A \cup A^{-1})^*$ be a set of *normal forms* (not necessarily unique) for elements of G, that is $\overline{L} = G$. Consider, for instance, the case where L is an automatic language for G. By Theorem A there is $\lambda > 0$ such that for a random long freely reduced word $w \in F(A)$ we have $|\overline{w}|_G/|w| \approx \lambda$. Let $w_L \in L$ be a word representing the same element of G as w. Then

$$|w_L| \ge |\overline{w}|_G$$

and hence

$$\frac{|w_L|}{|w|} \ge \frac{|\overline{w}|_G}{|w|} \approx \lambda > 0.$$

Thus for a long random word $w \in F(S)$ when we take $\overline{w}$ to a normal form $w_L \in L$, the ratio $|w_L|/|w|$ is separated from zero. This conclusion applies to a number of experimental observations, such as those obtained by Dehornoy [17] in the case of braid groups.

Theorem A has implications regarding the growth of *random automorphisms.* Let G be a finitely generated group with a fixed word metric corresponding to a finite generating set S. Let $\phi \in Aut(G)$ be an automorphism. We define the *norm* of ϕ with respect to S as

$$||\phi|| = ||\phi||_S := \max_{s\in S} |\phi(s)|_S.$$

Then for any $g \in G$ we have $|\phi(g)|_S \le ||\phi||_S |g|_S$ and hence

$$||\phi|| = \sup_{g\in G, g\neq 1} \frac{|\phi(g)|_S}{|g|_S}.$$

For an individual ϕ the sequence $\log ||\phi^n||$ is subadditive and therefore the following limit (sometimes called the *growth entropy of* ϕ) exists:

$$\nu(\phi) := \lim_{n\to\infty} \frac{\log ||\phi^n||}{n}.$$

It turns out that this notion has a generalization for an arbitrary finitely generated subgroup of $Aut(G)$:

Theorem B: *Let G be a nontrivial finitely generated group with a word-metric d_S corresponding to a finite generating set S. Let $H \le Aut(G)$ be a noncyclic subgroup with a finite generating set T. Then:*

1. *There is $\nu = \nu(H) = \nu(H, T, S) \ge 0$ such that for a non-backtracking simple random walk ϕ_n on the Cayley graph of H with respect to T we have*
$$\lim_{n\to\infty} \frac{\log ||\phi_n||_S}{n} = \nu$$
almost surely and in L^1.
2. *If G has polynomial growth and H is non-amenable then $\nu(H, T, S) > 0$.*

Note that $\nu(H, T, S) > 0$ means that $||\phi_n||_S$ grows exponentially with n for a "random" automorphism ϕ_n.

Corollary C: *Let F be a free group of finite rank $k \ge 2$ and let $H \le Aut(F)$ be a finitely generated group of automorphisms of F such that the image H' of H in $Aut(F_{ab}) \cong GL(k, \mathbb{Z})$ is non-amenable. Then for any finite generating set S of F and for any finite generating set T of H we have $\nu(H, T, S) > 0$.*

By the Tits alternative a subgroup of $GL(k, \mathbb{Z})$ is either virtually solvable (and hence amenable) or it contains a free subgroup of rank two (and hence is nonamenable). Thus in the above corollary we could replace the assumption that H' is nonamenable by the requirement that H' is not virtually solvable.

1.2. Free actions on trees: Two interpretations of stretching factors. In the context of free and discrete isometric actions of free groups on $\mathbb{R}$-trees (cf. Example 1.1), generic stretching factors are related to Bonahon's notion [5, 6] of the intersection number between geodesic currents on hyperbolic surfaces. If G is a non-elementary word-hyperbolic group, a *geodesic current* on G is a G-invariant positive Borel measure on $\partial^2 G := \{(x,y) | x, y \in \partial G, x \neq y\}$. The space of all geodesic currents on G, endowed with the weak-$*$-topology, is denoted by $Curr(G)$. (See [7, 31] for a detailed discussion on the subject.)

Every nontrivial conjugacy class $[g]$ in G defines an associated "counting" current $\eta_{[g]}$ on G. When S is a closed surface of negative Euler characteristic and $G = \pi_1(S)$, Bonahon proved that the notion of geometric intersection number between free homotopy classes of essential closed curves on S (that is, between nontrivial conjugacy classes of G) extends to a bilinear continuous "intersection form"

$$i \colon Curr(G) \times Curr(G) \to \mathbb{R}.$$

Note that in this case $\partial G = \partial \mathbb{H}^2 = \mathbb{S}^1$. For every hyperbolic structure ρ on S there is an associated *Liouville current* $L_\rho \in Curr(G)$ (see [5]). Bonahon's construction has the following natural property: if ρ is as above and $[g]$ is a nontrivial conjugacy class in G then $i(L_\rho, \eta_{[g]}) = \ell_\rho(g)$. Here $\ell_\rho \colon G \to \mathbb{R}$ is the *length spectrum* of ρ. Thus $\ell_\rho(g)$ is equal to the translation length of g as an isometry of $\mathbb{H}^2 = \widetilde{(S,\rho)}$ and it is also equal to the ρ-length of the shortest curve of the free homotopy class of closed curves on S corresponding to $[g]$. It turns out that the intersection number $i(L_\rho, L_{\rho'})$ between Liouville currents corresponding to two hyperbolic structures ρ, ρ' can be interpreted as the generic stretching factor of a long random closed geodesic on (S,ρ) with respect to ρ'. Namely, let $p \in S$ and let v be a random unit tangent vector at p on (S,ρ). For every $n \geq 1$ let α_n be the geodesic of length n on (S,ρ) with origin p and with the tangent vector v at p. Let β_n be a geodesic from the terminus of α_n to p of length $\leq Diam(S,\rho)$. Then $\gamma_n = \alpha_n \beta_n$ is a closed curve on S. Bonahon's results imply that

$$\lim_{n\to\infty} \frac{\ell_{\rho'}([\gamma_n])}{\ell_\rho([\gamma_n])} = \lim_{n\to\infty} \frac{\ell_{\rho'}([\gamma_n)])}{n} = \frac{i(\rho,\rho')}{\pi^2 |\chi(S)|}.$$

It turns out that a version of this interpretation applies in the context of free groups acting on trees. Let F be a free group of finite rank $k \geq 2$ In [30, 31] Kapovich investigated a natural "intersection form" $I \colon FLen(F) \times Curr(F) \to \mathbb{F}$, where $FLen(F)$ is the space of hyperbolic length functions corresponding to free and discrete isometric actions of F on $\mathbb{R}$-trees. This form still has the natural property that for any nontrivial conjugacy class $[g]$ in F and any $\ell \in FLen(F)$ we have $I(\ell, \eta_{[g]} = \ell(g)$. Let A be a free basis of F and let $\ell \in FLen(F)$ be realized by a free and discrete isometric action $\phi : F \to Isom(X)$ of F on an $\mathbb{R}$-tree X. Let μ_A be the uniform measure on ∂F corresponding to A. The measure μ_A on ∂F determines a *uniform* current $\nu_A \in Curr(F)$ that is analogous to the Liouville current corresponding to a hyperbolic structure on a surface. As shown in [31], similarly to Bonahon's situation, we have

$$I(\ell, \nu_A) = \lambda_A(\phi).$$

Generic stretching factors are also related to the notion of the Hausdorff dimension of a measure with respect to a metric. If μ is a measure on a metric space (M, d), the *Hausdorff dimension of μ with respect to d*, denoted $\mathbf{HD}_d(\mu)$ (or just $\mathbf{HD}(\mu)$), is defined as the infimum of Hausdorff dimensions of subsets of (M, d) of full measure μ.

In [28] Kaimanovich proved that for the harmonic measure ν on ∂T associated to a regular Markov operator P with a positive rate of escape on a tree T with uniformly bounded vertex degrees we have

$$\mathbf{HD}(\nu) = h/c$$

where c is the rate of escape and h is the asymptotic entropy of P.

This result is relevant in our context. Indeed, let A be a free basis of F and let ϕ: $F \to Isom(X)$ be a free, discrete and minimal isometric action of F on an $\mathbb{R}$-tree X. Then X/F is a finite metric graph and X is the universal cover of this graph. Let $\Gamma(F, A)$ denote the Cayley graph of F with respect to A. The orbit map $w \mapsto wp$ (where $p \in X$ is a base-point) gives a quasi-isometry between the trees $\Gamma(F, A)$ and X which extends to a homeomorphism $\hat{\phi}$: $\partial\Gamma(F, A) \to \partial X$ where ∂X is metrized in the standard $CAT(-1)$ way: $d(\zeta, \xi) = e^{-d(p,[\zeta,\xi])}$ for $\zeta, \xi \in \partial X$. Let μ_A be the uniform probability measure on $\partial\Gamma(F, A) = \partial F$ corresponding to A and let μ'_A denote the push-forward of μ_A via $\hat{\phi}$ to ∂X.

Then the result of Kaimanovich [28] mentioned above implies that

$$\mathbf{HD}_d(\mu'_A) = \frac{\log(2k-1)}{\lambda_A(\phi)},$$

where $k \geq 2$ is the rank of F.

1.3. Main results about actions on trees.

Our first main result is:

Theorem D: *Let $F = F(A)$ be a free group of rank $k \geq 2$. Let ϕ: $F \to Aut(X)$ be a free simplicial action without inversion of F on a simplicial tree X.*

Then the following hold:

(1) *The generic stretching factor $\lambda = \lambda(\phi)$ is a rational number ≥ 1 with*

$$2k\lambda \in \mathbb{Z}\left[\frac{1}{2k-1}\right].$$

(2) *The number $\lambda(\phi)$ is algorithmically computable in terms of ϕ, provided X is the universal cover of a finite connected graph and ϕ is given by an isomorphism between F and the fundamental group of that graph.*

The most interesting case of the above theorem is where X is the Cayley graph of $F = F(A)$ and where the action of F on X is determined by an endomorphism of F.

Definition 1.4 (Generic stretching factor of an endomorphism)*:* Let $F = F(A)$ where $k \geq 2$ and $A = \{a_1, \dots, a_k\}$. Let ϕ: $F \to F$ be an endomorphism of F. Let $X = \Gamma(F, A)$ be the Cayley graph of F and consider the action θ: $F \to Isom(X)$ given by $\theta(w)x := \phi(w)x$, where $w \in F, x \in X$. The generic stretching factor $\lambda_A(\theta)$ corresponding to this action is called the *generic stretching factor of ϕ with respect to A* and is denoted $\lambda_A(\phi)$ or just $\lambda(\phi)$ if A is fixed.

Thus $\lambda(\phi)$ approximates the distortion $|\phi(w)|_A/|w|_A$ for a long random freely reduced word w in $A^{\pm 1}$. For instance, for the Nielsen automorphism $\phi \in Aut(F(a,b))$, $\phi(a) = ab, \phi(b) = b$ it turns out that $\lambda(\phi) = \frac{7}{6}$. If ϕ is an automorphism of $F(a_1, \dots, a_k)$, then the precise relationship between $\lambda(\phi)$ and the traditionally studied dynamical properties of ϕ is not very clear. Nevertheless, we are able to estimate the growth of $\lambda(\phi^n)$ for hyperbolic automorphisms. Recall that $\phi \in Aut(F)$ is *hyperbolic* if there exist $s > 1$ and $m \geq 1$ such that for any $w \in F$

$$s||w|| \leq \max\{||\phi^m(w)||, ||\phi^{-m}(w)||\}.$$

By a result of Brinkmann [9] an automorphism $\phi \in Aut(F)$ is hyperbolic if and only if ϕ does not have any nontrivial periodic conjugacy classes in F. We prove:

Theorem E: *Let $F = F(a_1, \dots, a_k)$ and let $\phi \in Aut(F)$ be a hyperbolic automorphism with parameters $s > 1$ and $m \geq 1$ as above. Then*

$$\liminf_{n\to\infty} \sqrt[n]{\lambda(\phi^n)} \geq s^{1/m} > 1.$$

It is obvious that any automorphism of a finitely generated group G equipped with a word metric is a quasi-isometry and indeed a bi-Lipschitz equivalence. However, from the geometric point of view, especially in light of various versions of the Marked Length Spectrum Rigidity Conjecture, it is natural to study finer features of quasi-isometries. Recall that a map $f\colon (X,d) \to (X',d')$ is called a *rough isometry* if there is $D>0$ such that for any $x,y\in X$ we have $|d'(f(x),f(y))-d(x,y)|\le D$. A map $f\colon (X,d)\to(X',d')$ is called a *rough similarity* if there are $\lambda>0$ and $D>0$ such that for any $x,y\in X$ we have $|d'(f(x),f(y))-\lambda d(x,y)|\le D$. It is interesting and natural to ask when an automorphism is a rough similarity or a rough isometry.

An automorphism ϕ of $F=F(A)$ is called a *relabelling automorphism* if it is induced by a permutation of the set $A=\{a_1,\dots,a_k\}^{\pm 1}$. We say that $\phi\in Aut(F)$ is *simple* if it is equivalent to a relabeling automorphism in $Out(F)$, that is, if ϕ is the composition of a relabeling automorphism and a conjugation. Note that being a simple automorphism has a nice geometric meaning. Let $F=F(a_1,\dots,a_k)$ be realized as the fundamental group of the metric graph Γ which is a bouquet of k circles of length 1 corresponding to the generators $a_1,\dots,a_k$. An automorphism ϕ is simple if and only if, after possibly a composition with an inner automorphism, ϕ is induced by an *isometry* of the graph Γ.

Let P_n be the uniform probability measure on the set of all elements of F of length n. A set $W\subseteq F$ is said to be *exponentially F-generic* if $\lim_{n\to\infty} P_n(W)=1$ and convergence to this limit is exponentially fast. Similarly, a subset $C\subseteq\mathcal{CR}$ of the set $\mathcal{CR}$ of all cyclically reduced words is *exponentially $\mathcal{CR}$-generic* if $\lim_{n\to\infty} P_n'(C)=1$ with exponentially fast convergence, where P_n' is the uniform discrete probability measure on the set of cyclically reduced words of length n.

Obviously, any simple automorphism is a rough isometry and a rough similarity. The converse is also true, that is, any automorphism which is a rough similarity must be simple. (This follows, for example, from Theorem 2 of [20] together with some standard results about Culler–Vogtmann outer space.) Here we obtain a strengthened "random rigidity" version of this fact:

Theorem F: *Let $F=F(a_1,\dots,a_k)$ be a free group of rank $k\ge 2$ with the standard word metric d corresponding to the free basis $\{a_1,\dots,a_k\}$. Put*

$$d_0 := 1+\frac{2k-3}{4k^2-2k}.$$

There exists an exponentially $\mathcal{CR}$-generic set $C\subseteq\mathcal{CR}$ with the following property.

For any $\phi\in Aut(F)$ the following conditions are equivalent:

(1) *The automorphism ϕ is simple.*
(2) *We have $\lambda(\phi) = 1$.*
(3) *We have $\lambda(\phi) < 1 + (2k-3)/(2k^2-k)$.*
(4) *The map ϕ: $(F, d) \to (F, d)$ is a rough isometry.*
(5) *The map ϕ: $(F, d) \to (F, d)$ is a rough similarity.*
(6) *For some $w \in C$ we have $||\phi(w)|| = ||w||$.*
(7) *For every $w \in C$ we have $||\phi(w)|| = ||w||$.*
(8) *For some $w \in C$ we have $||\phi(w)|| \leq d_0||w||$.*
(9) *For every $w \in C$ we have $||\phi(w)|| \leq d_0||w||$.*

This result shows, in particular, that the set of all possible values of $\lambda(\phi)$ (where $\phi \in Aut(F)$) has a *gap*, namely the interval $(1, 1 + (2k-3)/(2k^2-k))$. Moreover, in the above theorem we can choose d_0 to be any number such that $1 < d_0 < 1 + (2k-3)/(2k^2-k)$.

Theorem F introduces a new dimension for rigidity results related to Marked Length Spectra on hyperbolic groups. Indeed, it is well-known that if $\phi \in Aut(F)$ fixes the lengths of all conjugacy classes (that is of all cyclic words), then ϕ is a rough isometry of F. Theorem F shows that even if $\phi \in Aut(F)$ preserves the length of a single "random" cyclically reduced word w then ϕ is a rough isometry and indeed a simple automorphism. To prove Theorem F we need some rather different tools and ideas, both algebraic and probabilistic. The key ingredient there is the work of Kapovich–Schupp–Shpilrain [36] on the behavior of Whitehead's algorithm and the action of $Aut(F)$ on "random" elements of F.

Using Theorem F it is not hard to show that the set of generic stretching factors taken over *all* free actions of $F(a_1, \dots, a_k)$ on simplicial trees also has a gap. Thus we obtain:

Theorem G: *Let $F = F(a_1, \dots, a_k)$ where $k \geq 2$. Let ϕ: $F \to Aut(X)$ be a free minimal action on F on a simplicial tree X without inversions.*

Then exactly one of the following occurs:

(1) *There is a simple automorphism α of F such that X is $\phi \circ \alpha$-equivariantly isomorphic to the Cayley graph of F with respect to $\{a_1, \dots, a_k\}$. In this case $\lambda(\phi) = 1$.*
(2) *We have $\lambda(\phi) \geq 1 + 1/k(2k-1)$.*

For an automorphism $\phi \in Aut(F)$ the *conjugacy distortion spectrum* of ϕ is

$$I(\phi) := \Big\{ \frac{||\phi(w)||}{||w||} : w \in F - \{1\} \Big\}.$$

Kapovich proved in [30] that $I(\phi)$ is always a $\mathbb{Q}$-convex subinterval of $\mathbb{Q}$ (that is, a set closed under taking rational convex combinations) with rational endpoints. Here we obtain:

Corollary H: *Let $\phi \in Aut(F)$ be an arbitrary automorphism. Then the following hold.*

(1) *Either ϕ is simple and $I(\phi) = 1$ or, else, 1 belongs to the interior of $\overline{I(\phi)}$.*
(1) *There exists $w \in F, w \neq 1$ such that $||\phi(w)|| = ||w||$.*

Proof: Part (1) obviously implies part (2) since, by the above-mentioned result of [30], $I(\phi)$ is a $\mathbb{Q}$-convex subset of $\mathbb{Q}$.

To see that (1) holds, assume that ϕ is not simple. Hence ϕ^{-1} is not simple either. By Theorem F we have $\lambda(\phi) > 1$ and $\lambda(\phi^{-1}) > 1$. Part (2b) of Theorem D now implies that there exists $w_1, w_2 \in F, w_1 \neq 1, w_2 \neq 1$ such that

$$x := \frac{||\phi(w_1)||}{||w_1||} > 1 \quad \text{and} \quad y := \frac{||\phi^{-1}(w_2)||}{||w_2||} > 1.$$

By definition of $I(\phi)$ we have $x \in I(\phi)$. Also, with $u = \phi^{-1}(w_2)$ we have $y = ||u||/||\phi(u)|| > 1$ and so $1/y \in I(\phi)$. Since $x > 1$ and $1/y < 1$, the $\mathbb{Q}$-convexity of $I(\phi)$ implies that 1 belongs to the interior of $\overline{I(\phi)}$, as claimed. ∎

We also obtain an application of Theorem F concerning the notion of the *flux* of an automorphism that was introduced and studied by Myasnikov and Shpilrain in [41].

Definition 1.5 (Flux). [41]*:* Let G be a finitely generated group with a fixed word metric. Let $\phi \in Aut(G)$.

For each $n \geq 0$ define

$$flux_\phi(n) := \#\{g \in G : |g| \leq n, |\phi(g)| > n\}$$

and

$$flux(\phi) := \limsup \sqrt[n]{\frac{flux_\phi(n)}{\#\{g \in G : |g| \leq n\}}}.$$

The sequence $flux_\phi(n)$ and the number $flux(\phi)$ provide a certain dynamical "measure of activity" of an automorphism ϕ. As a corollary of our results in this paper we obtain:

Corollary I: *Let $F = F(a_1, \dots, a_k)$ be a free group of rank $k \geq 2$, equipped with the standard metric.*

Then for any $\phi \in Aut(F)$ we have:

$$flux(\phi) = \begin{cases} 0, & \text{if } \phi \text{ is a relabeling automorphism,} \\ 1, & \text{otherwise.} \end{cases}$$

1.4. Random elements in regular languages. By definition, a language L over the alphabet A is regular if and only if there is a deterministic finite automaton which accepts the language L. It is a basic fact of formal language theory that the class of languages accepted by nondeterministic finite automata (NDFA) is also the class of regular languages. (See Hopcroft and Ullman [27].) Nondeterministic automata are very useful because a NDFA accepting a language L may be much smaller than any deterministic automaton accepting L. Such an automaton is not unique and choosing some finite automaton accepting L is like choosing a presentation for a group. One can choose a "random" element in the regular language L is via a random walk in the transition graph of any "suitable" finite state automaton M accepting the language L. We make this precise in Section 8, where we associate to M a finite state Markov process M' with the set of states being the set of directed edges in the transition graph $\Gamma(M)$ of M. The sample space Ω of M' is the set of semi-infinite edge-paths in $\Gamma(M)$. Each path in $\Gamma(M)$ (finite or infinite) has a label that is a word (finite or infinite) over A. If $\omega \in \Omega$ is such an infinite path, we denote by $w_n = w_n(\omega)$ the label of the initial segment of length n of ω. Any initial probability distribution μ on the edge-set $E(\Gamma(M))$ defines a probability measure P_μ on Ω. We need to impose a natural assumption on M in order to guarantee that the Markov process M' is irreducible. This technical assumption, which is frequently satisfied in practice, is made precise in the definition of a *normal* automaton in Section 8. Again applying the Subadditive Ergodic Theorem, we have:

Theorem J: *Let M be a normal automaton over a finite alphabet A and let $L = L(M)$ be the language accepted by M.*

Let ϕ: $A^* \to G$ be a monoid homomorphism, where G is a group with a left-invariant semi-metric d_G. Then there exists a number $\lambda = \lambda(M, \phi, d_G) \geq 0$ such that for any initial distribution μ on $E(\Gamma(M))$ we have

$$(|\phi(w_n)|_G)/n \to \lambda \textit{ almost surely and in } L^1 \textit{ with respect to } P_\mu.$$

If the initial distribution μ is supported on the edges of $\Gamma(M)$ originating at the start states of M then the word w_n can be extended by a word of uniformly bounded length to get a word $w_n' \in L$. We can think of w_n' as a "random" element of L with respect to M and μ. Theorem J then implies that $|\phi(w_n')|_G/n \to \lambda$ as $n \to \infty$ almost surely and in L^1 with respect to μ.

2. Random words and random walks

Convention 2.1: Let $A = \{a_1, \dots, a_k\}$ be a *free generating set* of a *free group* $F = F(A)$ of finite rank $k > 1$. For $w \in F$ we denote by $|w|_A$, or simply $|w|$, the *freely reduced length* of w with respect to A. Let $d(w_1, w_2) = |w_1^{-1}w_2|$ be the associated left-invariant metric on F. Also, $||w||_A = ||w||$ denotes the *cyclically reduced length* of f with respect to A, that is, the length of any cyclically reduced word in the alphabet $A^{\pm 1}$ conjugate to f.

This convention, including the fixed choice of the free basis $A = \{a_1, \dots, a_k\}$ of F, is adopted for the remainder of the paper, unless specified otherwise.

Recall that a nonnegative function $|\cdot|_G$ on a group G is called a *semi-norm* if for all $g, h \in G$ we have $|gh|_G \le |g|_G + |h|_G$.

In this Section we shall prove Theorem A from the Introduction:

Theorem 2.2: *Let $F = F(A)$, and let $\mu_A = \mu_A(A)$ be the uniform Borel probability measure on ∂F corresponding to the generating set $A = \{a_1, \dots, a_k\}$ with $k \ge 2$. Let ϕ: $F \to G$ be a homomorphism to a group G endowed with a semi-norm $|\cdot|_G$. Then:*

(1) *There exists a real number $\lambda \ge 0$ such that*

$$\lim_{n\to\infty} \frac{|\phi(\omega_n)|_G}{n} = \lambda$$

for μ_A-a.e. $\omega \in \partial F$ and in the space $L^1(\partial F, \mu_A)$.

(2) *If the group $\phi(F)$ is non-amenable and the sequence*

$$b_n = \#\{g \in \phi(F) : |g|_G \le n\}$$

grows at most exponentially, then $\lambda > 0$.

The condition on b_n in the above theorem is always satisfied if G is a finitely generated group and $|.|_G$ is the word metric corresponding to some finite generating set of G.

Any geodesic ray $\omega \in \partial F$ can be identified with the *non-backtracking path* $\omega_0, \omega_1, \dots$ in F starting from the group identity. Then the measure space $(\partial F, \mu_A)$ becomes the *space of sample paths* of the *non-backtracking simple random walk* (NBSRW) on the Cayley graph of F starting from the identity of the group. This is the Markov chain on F whose transition probabilities $\pi_f, f \in F$ are equidistributed among the neighbors of f which are strictly further from the group identity. Therefore, the number λ above is the *linear rate of escape* of the ϕ-image of the non-backtracking simple random walk on F. We shall deduce

Theorem 2.2 from well-known analogous properties of the usual random walks on groups by using the fact that the simple random walk on the free group asymptotically follows uniformly distributed geodesic rays.

Let μ be a probability measure on a group G. By definition, the sample paths of the associated *random walk* (G,μ) are products $g_n = h_1 h_2 \cdots h_n$ of independent μ-distributed *increments* h_n. In other words, the measure $\mathbf{P}$ in the space of sample paths which describes the random walk (G,μ) is the image of the product measure $\mu \otimes \mu \otimes \cdots$ in the space of increments under the above product map.

The following statement is known as Kingman's Subadditive Ergodic Theorem [37]. (See also [19] for a short proof.)

Proposition 2.3 (Subadditive Ergodic Theorem): *Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and let $\mathcal{S}\colon \Omega \to \Omega$ be a measure-preserving operator, that is such that for any measurable set $Q \subseteq \Omega$ we have $\mu(Q) = \mu(\mathcal{S}^{-1}Q)$.*

Let $X_n\colon \Omega \to \mathbb{R}$ be a sequence of non-negative integrable random variables such that for any $n, m \geq 0$

$$X_{n+m}(\omega) \leq X_n(\omega) + X_m(\mathcal{S}^n\omega), \quad \text{a.e. } \omega \in \Omega.$$

Then there exists a $\mathcal{S}$-invariant random variable $\lambda\colon \Omega \to \mathbb{R}$ such that

$$\lim_{n\to\infty} \frac{X_n}{n} = \lambda$$

almost surely and in L^1 on Ω.

In particular, if $\mathcal{S}$ is ergodic then $\lambda = const$ on Ω.

A straightforward application of Kingman's Subadditive Ergodic Theorem gives:

Proposition 2.4 ([26]): *If the measure μ has a finite first moment $\sum |g|\mu(g)$ with respect to a semi-norm $|\cdot|$ on the group G, then there exists a number $c \geq 0$ (called the linear rate of escape of the random walk (G,μ) with respect to the semi-norm $|\cdot|$) such that $|g_n|/n \to c$ for $\mathbf{P}$-a.e. sample path (g_n) and in the space $L^1(\mathbf{P})$.*

The following claim, if slightly more general than the one formulated in [26], can be obtained in the same way by using the spectral characterization of amenability (or by showing that $c = 0$ implies vanishing of the asymptotic entropy of the random walk, and therefore amenability of the group [38]):

Proposition 2.5 ([26]): *Under the assumptions of Proposition 2.4, if the group G is non-amenable and the semi-norm $|\cdot|_G$ has exponentially bounded growth and the support of the measure μ generates the group G, then $c > 0$.*

Let now μ'_A be the probability measure on the free group F equidistributed on the set $A^{\pm 1}$, so that $\mu'_A(a_i^{\pm 1}) = 1/2k$ for $i = 1, 2, \dots, k$.

Proposition 2.6 (see [29] and the references therein): *For* **P**-*a.e. sample path (g_n) of the random walk (F, μ'_A)*

(1) *There exists a limit*

$$g_\infty = \lim_{n\to\infty} g_n \in \partial F,$$

and its distribution (i.e., the image of the measure **P** *under the map $(g_n) \mapsto g_\infty$) coincides with the uniform measure μ_A on ∂F.*

(2) *We have*

$$\lim_{n\to\infty} \frac{|g_n|}{n} = \theta = \frac{k-1}{k}$$

(so that the linear rate of escape of the random walk (F, μ'_A) is $\frac{k-1}{k}$).

(3) *We have*

$$d(g_n, (g_\infty)_{[\theta n]}) = o(n).$$

Here $[x]$ denotes the integer part of a number $x \in \mathbb{R}$ and $(g_\infty)_{[\theta n]}$ denotes the vertex at distance $[\theta n]$ from 1 on the unique geodesic ray from 1 to g_∞ in the Cayley graph of F with respect to A.

Proof of Theorem 2.2: Consider the random walk (F, μ'_A). Its image under the homomorphism ϕ is the random walk on the group $\phi(F)$ governed by the measure $\phi(\mu'_A)$. Denote its rate of escape with respect to the semi-norm $|\cdot|_G$ by c. Then the combination of Proposition 2.4 and Proposition 2.6 implies the first part of Theorem 2.2. Indeed, for **P**-a.e. sample path (g_n) the distance in F between g_n and $(g_\infty)_{[\theta n]}$ is sublinear, whence the distance in G (with respect to the semi-metric determined by the semi-norm $|\cdot|_G$) between the ϕ-images of these points is also sublinear. Since the distribution of g_∞ is μ_A, we arrive at the conclusion that the first part of Theorem 2.2 holds for the number $\lambda = c/\theta$.

The second part of Theorem 2.2 is now an immediate corollary of Proposition 2.5.

Here is another argument establishing part (1) of Theorem 2.2 as a direct consequence of the Subadditive Ergodic Theorem.

Let $\Omega = \partial F$. Recall that for $\omega \in \partial F$ we denote by ω_n the element of F that is at distance n from 1 along the geodesic ray ω in $\Gamma(F, A)$. Let $X_n \colon \partial F \to \mathbb{R}$

be defined as $X_n(\omega) := |\phi(\omega_n)|_G$. Also, let $\mathcal{S}\colon \partial F \to \partial F$ be the standard shift operator consisting in erasing the first letter of a semi-infinite freely reduced word representing an element of ∂F. It is well-known that $\mathcal{S}$ is stationary and ergodic.

Note that for any $\omega \in \partial F$ we have

$$\omega_{n+m} = \omega_n(\mathcal{S}^n\omega)_m.$$

Hence

$$|\phi(\omega_{n+m})|_G = |\phi(\omega_n)\phi((\mathcal{S}^n\omega)_m)|_G \le |\phi(\omega_n)|_G + |\phi((\mathcal{S}^n\omega)_m)|_G.$$

Thus the conditions of the Subadditive Ergodic Theorem are satisfied and part (1) of Theorem 2.2 follows. ∎

The following is Theorem B from the Introduction.

Theorem 2.7: *Let G be a nontrivial finitely generated group with a word-metric d_S corresponding to a finite generating set S. Let $H \le Aut(G)$ be a noncyclic finitely generated group with a finite generating set T. Then:*

1. *There is $\nu = \nu(H) = \nu(H, T, S) \ge 0$ such that for a non-backtracking simple random walk ϕ_n on the Cayley graph of H with respect to T we have*
$$\lim_{n\to\infty} \frac{\log ||\phi_n||_S}{n} = \nu$$
almost surely and in L^1.
2. *If G has polynomial growth and H is non-amenable then $\nu(H, T, S) > 0$.*

Proof: It is clear from the definition of $||\cdot||_S$ that for any $\phi, \psi \in Aut(G)$ we have

$$||\phi\psi||_S \le ||\phi||_S||\psi||_S$$

and hence

$$\log ||\phi\psi||_S \le \log ||\phi||_S + \log ||\psi||_S.$$

Also, for any $\phi \in Aut(G)$ we have $||\phi||_S \ge 1$ and so $\log ||\phi||_S \ge 0$. Thus $\log ||\cdot||_S$ is a semi-norm on $Aut(G)$ that uniquely extends to a left-invariant semi-metric on $Aut(G)$ and thus on $H \le Aut(G)$. Hence part (1) of Theorem 2.7 follows directly from part (1) of Theorem A.

To see that part (2) holds suppose that H is nonamenable and that G has polynomial growth. This implies that $(H, ||\cdot||_S)$ has at most exponential growth. Hence part (2) of Theorem 2.7 follows from part (2) of Theorem A. ∎

Remark 2.8: The requirement of G having polynomial growth in Theorem B is important and cannot be easily dispensed with. If G is a group and $g \in G$, denote by $ad(g) \in Aut(G)$ the inner automorphism of G defined by $ad(g)(x) = gxg^{-1}$ for every $x \in G$. Now let $G = F(a_1, \dots, a_k)$ and $H = Inn(F) \le Aut(F)$ be the (non-amenable!) group of inner automorphisms of F with the generating set $T = \{ad(a_1), \dots, ad(a_k)\}$. Then for any product ϕ_n of n elements of T we have $||\phi_n|| \le 2n+1$. Since $\lim_{n\to\infty}(\log 2n+1)/n = 0$, we see that $\nu(H, T, S) = 0$. Nevertheless, in some instances quotient group considerations still imply that $\nu(A) > 0$ even if G does not have polynomial growth, or, equivalently, G is not virtually nilpotent. A typical example is given by Corollary 2.9 below.

We obtain Corollary C from the Introduction:

Corollary 2.9: *Let F be a free group of finite rank $k > 1$ and let $H \le Aut(F)$ be a finitely generated group of automorphisms of F such that the image H' of H in $Aut(F_{ab}) \cong GL(k, \mathbb{Z})$ is nonamenable. Then for any finite generating set S of F and for any finite generating set T of H we have $\nu(H, T, S) > 0$.*

Proof: Let S' be the image of S in the abelianization $\mathbb{Z}^k = F_{ab}$ of F. For any $\phi \in Aut(F)$ the automorphism ϕ of F factors through to an automorphism ϕ' of F_{ab}. Clearly $||\phi||_S \ge ||\phi'||_{S'}$. Hence $\nu(H, T, S) \ge \nu(H', T', S')$, where T' is the image of T in $Aut(F_{ab})$. Since F_{ab} has polynomial growth, by Theorem B we have $\nu(H', T', S') > 0$ and hence $\nu(H, T, S) > 0$. ∎

3. Frequencies and cyclic words

The following convention is fixed until the end of the paper unless specified otherwise.

Convention 3.1: As before, let $k \ge 2$ and let $F = F(A)$ where $A = \{a_1, \dots, a_k\}$. Let $\widehat{A} = A^{\pm 1}$. We denote by $\mathcal{CR}$ the set of all cyclically reduced words in F.

A *cyclic word* is an equivalence class of nontrivial cyclically reduced words, where two nontrivial cyclically reduced words are equivalent if they are cyclic permutations of each other. If v is a cyclically reduced word, we denote by (v) the cyclic word defined by v. Recall that if u is a freely reduced word, then $|u|$ denotes the length of u and $|u|$ denotes the length of the cyclically reduced form of u. If $w = (v)$ is a cyclic word then $||w|| = ||v||$ is the length of w.

Note that the set of cyclic words is naturally identified with the set of nontrivial conjugacy classes of elements of F.

Definition 3.2 (Frequencies)*:* Let w be a cyclic word.

Let v be a nontrivial freely reduced word. We define $n_w(v)$, *the number of occurrences of v in w*, as follows. Let $w = (z)$. Take the smallest $p > 0$ such that $|z^{p-1}| \geq 2|v|$ and count the number of those $i, 0 \leq i < ||w||$ such that $z^p \equiv z_1 v z_2$ where $|z_1| = i$. By definition this number is $n_w(v)$. If $v = 1$, we define $n_w(1) := ||w||$.

There is a more graphical way of defining $n_w(v)$ for a nontrivial cyclic word w. We think of w as a cyclically reduced word written on a circle in a clockwise direction without specifying a base-point. Then $n_w(v)$ is the number of positions on the circle, starting from which it is possible to read the word v going clockwise along the circle (and wrapping around more than once, if necessary).

For any freely reduced word v we define *frequency* of v in w as:

$$f_w(v) := \frac{n_w(v)}{||w||}.$$

Also, if w is a nontrivial freely reduced word, and v is another nontrivial freely reduced word, we define $n_w(v)$, *the number of occurrences of v in w*, as follows. If $|w| = n > 0$ then by definition $n_w(v)$ is the number of those $i, 0 \leq i < n$ for which w decomposes as a freely reduced product $w = w'vw''$ with $|w'| = i$. Unlike the situation when w is a cyclic word, if $|v| \leq |w|$ then necessarily $n_w(v) = 0$.

Lemma 3.3: *Let w be a nontrivial cyclic word. Then:*

(1) *For any $m \geq 0$ and for any freely reduced word u with $|u| = m$ we have*

$$n_w(u) = \sum_{x \in \widehat{A}, |ux| = |u|+1} n_w(ux) = \sum_{x \in \widehat{A}, |xu| = |u|+1} n_w(xu),$$

and

$$f_w(u) = \sum_{x \in \widehat{A}, |ux| = |u|+1} f_w(ux) = \sum_{x \in \widehat{A}, |xu| = |u|+1} f_w(xu).$$

(2) *For any $m \geq 1$*

$$\sum_{|u|=m} n_w(u) = ||w|| \quad \textit{and} \quad \sum_{|u|=m} f_w(u) = 1.$$

(3) *For any $s > 0$ and any $u \in F$*

$$n_{w^s}(u) = s n_w(u) \quad \textit{and} \quad f_{w^s}(u) = f_w(u).$$

Proof: Parts (1) and (3) are obvious. We establish (2) by induction on m. For $m = 1$ the statement is clear. Suppose that $m > 1$ and that (2) has been established for $m - 1$.

We have

$$\sum_{|u|=m} n_w(u) = \sum_{\substack{|v|=m-1, x\in\widehat{A}:\\ |vx|=m}} n_w(vx) = \sum_{|v|=m-1} n_w(v) = ||w||,$$

as required. ∎

Definition 3.4 (Nielsen automorphisms)*:* A *Nielsen* automorphism of F is an automorphism τ of one of the following types:

1. There is some $i, 1 \le i \le k$ such that $\tau(a_i) = a_i^{-1}$ and $\tau(a_j) = a_j$ for all $j \neq i$.
2. There are some $1 \le i < j \le k$ such that $\tau(a_i) = a_j$, $\tau(a_j) = a_i$ and $\tau(a_l) = a_l$ when $l \neq i, l \neq j$.
3. There are some $1 \le i < j \le k$ such that $\tau(a_i) = a_i a_j$ and $\tau(a_l) = a_l$ for $l \neq i$.

It is a classical fact that the set of all Nielsen automorphisms generates $Aut(F)$.

The following proposition proved by Kapovich in [30] is crucial for our arguments.

Proposition 3.5: *Let $\phi \in Out(F)$ be an outer automorphism and let $p \ge 0$ be such that ϕ can be represented, modulo $Inn(F)$, as a product of p Nielsen automorphisms.*

Then for any freely reduced word $v \in F$ with $|v| = m$ there exists a collection of computable integers $c(u, v) = c(u, v, \phi) \ge 0$, where $u \in F$, $|u| = 8^p m$, such that for any nontrivial cyclic word w we have

$$n_{\phi(w)}(v) = \sum_{|u|=8^p m} c(u, v) n_w(u).$$

Corollary 3.6: *Let ϕ be an automorphism of F and let p be such that ϕ can be written as a product of p Nielsen automorphisms.*

There exists a collection of integers $e(v) = e(v, \phi) \ge 0$, where $v \in F, |v| = 8^p$, such that for any cyclic word w we have:

$$||\phi(w)|| = \sum_{|v|=8^p} e(v) n_w(v) \quad \textit{and} \quad \frac{||\phi(w)||}{||w||} = \sum_{|v|=8^p} e(v) f_w(v).$$

Moreover, there is an algorithm which, given ϕ and u, computes the numbers $e(v)$.

Proof: Since $||\phi(w)|| = \sum_{x\in\widehat{A}} n_{\phi(w)}(x)$, the statement follows directly from Proposition 3.5. ∎

The following well-known fact is a version of the so-called "Bounded Cancellation Lemma" (see [16]):

Lemma 3.7: *Let α be an injective endomorphism of F. There is $N = N(\alpha) > 0$ such that for any cyclically reduced word w the maximal terminal segment of $\alpha(w)$ that cancels in the product $\alpha(w)\cdot\alpha(w)$ has length at most N.*

4. Actions on trees

Let Γ be a finite connected graph with orientation $E\Gamma = E^+\Gamma \sqcup E^-\Gamma$. For $e \in E$ we use the following notation. The inverse edge of e is denoted by $\overline{e}$, $o(e)$ denotes the initial vertex of e and $t(e)$ denotes the terminal vertex of e.

Let F be a free group and let $\phi : F \to \pi_1(\Gamma, p)$ be an isomorphism between F and the fundamental group of Γ with respect to a vertex p. Let T be a maximal tree in Γ. For any vertex v, let $[p, v]_T$ denote the path in T from p to v. The choice of T defines a basis S_T of $\pi_1(\Gamma, p)$ as follows:

$$S_T := \{[p, o(e)]_T\ e\ [t(e), p]_T : e \in E^+(\Gamma - T)\}.$$

The ϕ-pullback of this basis $B_T := \phi^{-1}(S_T)$ is a basis of F referred to as the *geometric basis* of F determined by T.

Let $s_e := [p, o(e)]_T\ e\ [t(e), p]_T$ where $e \in E(\Gamma - T)$, so that $s_{\overline{e}} = s_e^{-1}$. Let $b_e = \phi^{-1}(s_e)$, where $e \in E(\Gamma - T)$, so that again $b_{\overline{e}} = b_e^{-1}$.

The following obvious lemma indicates the explicit correspondence between freely reduced words in S_T (or B_T) and reduced edge-paths in Γ.

Lemma 4.1:

(1) *Let γ be an edge-path in Γ from p to p. Let u be a word in S_T constructed from γ as follows: delete all the edges of T from Γ and replace each edge $e \in E^+(\Gamma - T)$ in γ by s_e and each edge $e \in E^-(\Gamma - T)$ in γ by $s_{\overline{e}}^{-1}$. Then $u = \gamma$ in $\pi_1(\Gamma, p)$ and u is a reduced word in S_T if and only if γ is a reduced path.*

(2) *Let u be a word in $S_T \cup S_T^{-1}$, where $\epsilon_i = \pm 1$.*

Construct the path γ from p to p as follows. First for each $e \in E^+(\Gamma - T)$

replace every s_e in u by e and replace every s_e^{-1} by $\bar{e}$. Then between every two consecutive e, e' insert the path $[t(e), o(e')]_T$. Finally, append the path $[p, o(e)]_T$ in front, for the first edge $e_0 \in E(\Gamma - T)$ of the resulting sequence, and append the path $[t(e_0'), p]_T$ at the end for the last edge $e_0' \in E(\Gamma - T)$ of the sequence.

Then γ is a path from p to p that is equal to u in $\pi_1(\Gamma, p)$ and that is reduced if and only if the word u over S_T is reduced.

(3) *Let γ be a closed edge-path in Γ. Let u be a word in $S_T^{\pm 1}$ obtained from γ as in (1). Then the loop at p corresponding to u in $\pi_1(\Gamma, p)$ is freely homotopic to γ in Γ and the word u is cyclically reduced if and only if the path γ is cyclically reduced.*

(4) *Let w be a cyclic word in $S_T^{\pm 1}$. Let γ be a circuit in Γ obtained as follows. Replace each occurrence of s_e in w by e and each occurrence of s_e^{-1} by $\bar{e}$; after that, between each two consecutive (in the cyclic order) edges e, e' insert the path $[t(e), o(e')]_T$. Then w and γ represent freely homotopic loops in Γ and the cyclic word w is reduced if and only if the circuit γ is reduced.*

Now suppose that Γ is endowed with the structure of a *metric graph*, that is, each edge e of Γ is assigned a *length* $\ell(e) > 0$ in such a way that $\ell(e) = \ell(\bar{e})$ for each edge e. Let $X = \widetilde{(\Gamma, p)}$ be the universal cover of Γ. Then X inherits the structure of a metric tree with an isometric action of $\pi_1(\Gamma, p)$ and, via ϕ, an action of F on X.

Let $\tilde{p}$ be a lift of p to X. For $g \in \pi_1(\Gamma, p)$ let $|g|_p := d_X(\tilde{p}, g\tilde{p})$. Also denote by $||g||_X$ the translation length of g when acting on X. Similarly, if w is a conjugacy class (or a cyclic word) in $\pi_1(\Gamma, p)$, we denote by $||w||_X$ the translation length of u with respect to X. For each freely reduced word $z = s_e^{\epsilon} s_{e'}^{\delta}$ of length two in $S_T^{\pm 1}$, where $\epsilon, \delta \in \{1, -1\}$, denote by r_z the length of the edge-path $[t(e^{\epsilon}), o(e'^{\delta})]_T$ in Γ. Let Z be the set of all freely reduced words of length two in S_T. For each $a = s_e \in S_T$ denote $e(a) :=$ and $e(a^{-1}) := \bar{e}$.

Lemma 4.2:

(1) *Let w be a reduced cyclic word in $S_T^{\pm 1}$. Then*

$$||w||_X = \sum_{a \in S_T^{\pm 1}} \ell(e(a)) n_w(a) + \sum_{z \in Z} r_z n_w(z).$$

(2) *Let u be a freely reduced word in S_T. Then*

$$|u|_p = \sum_{a \in S_T^{\pm 1}} \ell(e(a)) n_u(a) + \sum_{z \in Z} r_z n_u(z) + \ell([p, o(e)]_T) + \ell([t(e'), p]_T)$$

where e and e' are the last and the first edges of $\gamma(u)$ accordingly.

The following is Theorem D from the Introduction:

Theorem 4.3: *Let $F = F(a_1, \ldots, a_k)$, where $k \geq 2$, and let $A = \{a_1, \ldots, a_k\}$. Then for any free action ϕ of F on a simplicial tree X without inversions the generic stretching factor $\lambda(\phi) = \lambda_A(\phi)$ is a rational number with*

$$2k\lambda(\phi) \in \mathbb{Z}\left[\frac{1}{2k-1}\right].$$

Moreover, if X is given as the universal cover of a finite connected simplicial graph Γ and if the action ϕ is given via an explicit isomorphism between F and $\pi_1(\Gamma, p)$, then $\lambda(\phi)$ is algorithmically computable in terms of ϕ.

Proof: Recall that the definition of $\lambda(\phi, |\cdot|_x)$ does not depend on the choice of a point $x \in X$. Hence we may assume that x is a vertex of the minimal F-invariant subtree of X, and therefore, that the action of F on X is minimal. Let $\Gamma = X/F$ be the finite quotient graph. Choose an orientation on Γ, a maximal tree T in Γ. Choose a base-vertex p in Γ to be the image of $x \in X$ in Γ. Note that in both X and Γ every edge has length 1. Then there is a canonical isomorphism $\psi\colon F \to \pi_1(\Gamma, p)$. Let S_T and B_T be the geometric bases corresponding to T for $\pi_1(\Gamma, p)$ and F accordingly.

Fix a bijection between B_T and $A = \{a_1, \ldots, a_k\}$ and an automorphism α of F induced by this bijection of the two free bases of F.

Let $g = x_1 \cdots x_n \in F$ be a freely reduced word of length n in $F(a_1, \ldots, a_k)$. Let g' be a cyclically reduced word of length n over A obtained from g by changing the last letter of g if necessary. Thus $|g'g^{-1}|_A \leq 2$.

Let w' be the cyclic word over A defined by g'. Let w be the result of rewriting w' as the cyclic word in B_T. Then there is an integer $M \geq 1$ such that for each freely reduced word z in B_T of length at most 2,

$$n_w(z) = \sum_{|u|_A = M} c(u, z) n_w(u)$$

where $c(u, z) \geq 0$ are some integers independent of w. Let Z_i be the set of freely reduced words of length i in B_T, for $i = 1, 2$.

Then

$$\begin{aligned} ||g'||_X = ||w||_X &= \sum_{a \in Z_1} \ell(e(a)) n_w(a) + \sum_{z \in Z_2} r_z n_w(z) \\ &= \sum_{a \in Z_1} \sum_{|u|_A = M} \ell(a) c(u, a) n_{w'}(u) + \sum_{z \in Z_2} \sum_{|u|_A = M} r_z c(u, z) n_{w'}(u), \end{aligned}$$

It follows from Lemma 4.1 and Lemma 4.2 that if $h \in F$ is cyclically reduced over B_T then $|||h||_X - |h|_x| \le N$, where $N = N(x) > 0$ is some constant independent of h. On the other hand, by the Bounded Cancellation Lemma (Lemma 3.7) there exists a constant $N' > 0$ such that for any cyclically reduced word y over A, we have $|||y||_{B_T} - |y|_{B_T}| \le N'$. By construction g' is cyclically reduced over A and $|g'g^{-1}|_A \le 2$. Hence there exists a constant $L > 0$ such that for every g as above and each $u \in F$ with $|u|_A = M$ we have $||g|_x - ||g'||_X| \le L$ and $|n_g(u) - n_{w'}(u)| \le L$.

Therefore, there is another constant $L' > 0$ independent of f such that for every freely reduced word g of length n over A

$$(*).\quad \left| \sum_{a \in Z_1} \sum_{|u|_A = M} \ell(e(a)) c(u,a) f_g(u) + \sum_{z \in Z_2} \sum_{|u|_A = M} r_z c(u,z) f_g(u) - \frac{|g|_p}{n} \right| \le \frac{L'}{n}$$

If g_n is a freely reduced word of length n obtained by a non-backtracking simple random walk of length n on $F(a_1, \dots, a_k)$, then for each $u \in F(a_1, \dots, a_k)$ with $|u|_A = M$ we have

$$\lim_{n \to \infty} f_{g_n}(u) = \frac{1}{2k(2k-1)^{M-1}} \quad \text{almost surely.}$$

Therefore $(*)$ implies that

$$(**)\qquad \lambda(\phi) = \frac{1}{2k(2k-1)^{M-1}} \left[\sum_{a \in Z_1} \sum_{|u|_A = M} \ell(e(a)) c(u,a) + \sum_{z \in Z_2} \sum_{|u|_A = M} r_z c(u,z) \right].$$

Since $\ell(e(a)) = 1, c(u,a), c(u,z)$ and r_z are integers, it follows that $\lambda(\phi)$ is rational and, moreover, that

$$2k\lambda(\phi) \in \mathbb{Z}\left[\frac{1}{2k-1}\right].$$

Moreover, $\lambda(\phi)$ is computable in terms of an explicit isomorphism between F and $\pi_1(\Gamma, p)$. ∎

Remark 4.4: The formula $(**)$ for $\lambda(\phi)$ holds for an arbitrary structure of a metric graph on Γ, where the lengths of edges are allowed to be arbitrary positive real numbers and not necessarily 1. If the lengths of all edges of Γ are rational, then by $(**)$ $\lambda(\phi)$ is also rational. Moreover, if these length of the edges are given to us in some algorithmically describable form then $\lambda(\phi)$ is computable in terms of these lengths and of an an explicit isomorphism between F and $\pi_1(\Gamma, p)$.

5. Genericity

Convention 5.1: Recall that $\mathcal{CR}$ denotes the set of all cyclically reduced words in $F = F(a_1, \dots, a_k)$. If $S \subseteq F$ and $n \geq 0$ we denote

$$\rho(n, S) := \#\{w \in S : |w| \leq n\} \quad \text{and} \quad \gamma(n, S) := \#\{w \in S : |w| = n\}.$$

Let P_n be the uniform discrete probability measure on the set of all elements $w \in F$ with $|w| = n$. We extend P_n to F by setting $P_n(w) = 0$ for any $w \in F$ with $|w| \neq n$.

Similarly, let P'_n be the uniform discrete probability measure on the set of all cyclically reduced elements $w \in F$ with $||w|| = n$. We extend P_n to $\mathcal{CR}$ by setting $P'_n(w) = 0$ for any $w \in \mathcal{CR}$ with $||w|| \neq n$.

Thus $\gamma(n, F) = 2k(2k-1)^{n-1}$ for $n > 0$.

For a number sequence x_n with $\lim_{n\to\infty} x_n = x \in \mathbb{R}$ we say that the convergence is *exponentially fast* if there exist $0 < \sigma < 1$ and $D > 0$ such that for all $n \geq 1$ we have $|x_n - x| \leq D\sigma^n$.

Definition 5.2 (Genericity)*:* Let $S \subseteq \mathcal{W} \subseteq F$. We say that S is *exponentially $\mathcal{W}$-generic* if

$$\lim_{n\to\infty} \frac{\gamma(n, S)}{\gamma(n, \mathcal{W})} = 1$$

and the convergence is exponentially fast. The complement in $\mathcal{W}$ of an exponentially $\mathcal{W}$-generic set is called *exponentially $\mathcal{W}$-negligible.*

In practice we are only interested in the cases $\mathcal{W} = F$ and $\mathcal{W} = \mathcal{CR}$, the set of all cyclically reduced words in F. By definition $S \subseteq F$ is exponentially F-generic if and only if $\lim_{n\to\infty} P_n(S) = 1$ with exponentially fast convergence in this limit. Similarly $S \subseteq \mathcal{CR}$ is exponentially $\mathcal{CR}$-generic if and only if $\lim_{n\to\infty} P'_n(S) = 1$ with exponentially fast convergence. Here there is a simple criterion of being exponentially negligible [36] in F and $\mathcal{CR}$:

Lemma 5.3: *Let $\mathcal{W} = F$ or $\mathcal{W} = \mathcal{CR}$. Then for a subset $S \subseteq \mathcal{W}$ the following are equivalent:*

(1) *The set S is exponentially $\mathcal{W}$-negligible.*

(2) *We have*

$$\frac{\gamma(n, S)}{(2k-1)^n} \to_{n\to\infty} 0 \textit{ exponentially fast.}$$

(3) *We have*

$$\frac{\rho(n, S)}{(2k-1)^n} \to_{n\to\infty} 0 \textit{ exponentially fast.}$$

(4) *We have*

$$\limsup_{n\to\infty} \sqrt[n]{\rho(n,S)} < 2k-1.$$

(5) *We have*

$$\limsup_{n\to\infty} \sqrt[n]{\gamma(n,S)} < 2k-1.$$

Proposition 5.4: *Let $\epsilon > 0$ and let $m > 0$ be an integer. Then the set*

$$W(m,\epsilon) := \Big\{ w\in F : \text{ for any } u\ne 1 \text{ with } |u|=m \quad \text{we have } |f_w(u) - \frac{1}{2k(2k-1)^{m-1}}| < \epsilon \Big\}$$

is exponentially F-generic.

Proof: This is a straightforward corollary of Large Deviation Theory [18] applied to the finite state Markov chain generating the freely reduced words in F. We refer the reader to [36] for a more detailed discussion about large Deviation Theory and how it works in this particular case. ∎

It is not hard to deduce the following from Proposition 5.4.

Proposition 5.5: *Let $\epsilon > 0$ and let $m > 0$ be an integer. Then the set*

$$C(m,\epsilon) := \Big\{ w\in \mathcal{CR}\text{: for any } u\ne 1 \text{ with } |u|=m \text{ and for the cyclic word } (w) \quad \text{we have } |f_{(w)}(u) - \frac{1}{2k(2k-1)^{m-1}}| < \epsilon \Big\}$$

is exponentially $\mathcal{CR}$-generic.

We now give the definition of an "approximate" stretching factor, which will later be seen to be equivalent to the generic stretching factor of an automorphism introduced earlier.

Definition 5.6: Let ϕ: $F \to Aut(X)$ be a free simplicial action without inversions of $F = F(a_1,\dots,a_k)$ on a simplicial tree X.

We say that a number $\lambda \ge 0$ is a *approximate stretching factor* of ϕ if for every $p\in X$ and for any $\epsilon > 0$ the set

$$\Big\{ w\in F : |\frac{|w|_p}{|w|} - \lambda| \le \epsilon \Big\}$$

is exponentially generic in F.

Similarly, we say that a number $\lambda \ge 0$ is a *approximate conjugacy stretching factor* of ϕ if for any $\epsilon > 0$ the set

$$\{w \in \mathcal{CR} : |\frac{||w||_X}{||w||} - \lambda| \le \epsilon\}$$

is exponentially generic in $\mathcal{CR}$.

Proposition 5.7: *Let ϕ: $F \to Aut(X)$ be a free simplicial action of $F = F(a_1, \ldots, a_k)$ on a simplicial tree X.*

1. *There is at most one approximate stretching factor for ϕ.*
2. *There is at most one approximate conjugacy stretching factor for ϕ.*
3. *If λ is an approximate conjugacy stretching factor for ϕ then λ is also an approximate stretching factor for ϕ.*
4. *If λ is an approximate stretching factor for ϕ then λ is also an approximate conjugacy stretching factor for ϕ.*

Proof: Parts (1) and (2) are obvious.

We now establish (3). Indeed, suppose that λ is an approximate conjugacy stretching factor for ϕ. Let $\epsilon > 0$ and let S be the set of all cyclically reduced words w such that

$$|\frac{||w||_X}{||w||} - \lambda| \ge \epsilon/2.$$

Then S is exponentially $\mathcal{CR}$-negligible, so that

$$(\gamma(n,S))/(2k-1)^n \to_{n\to\infty} 0$$

exponentially fast. Let $p \in X$. Put $M = \max\{|a_i|_p : 1 \le i \le k\}$. Let $N > 0$ be an integer such that for any cyclically reduced word u we have $||u|_p - ||u||_X| \le N$.

Let S' be the set of all freely reduced words w in F that differ from an element of S in at most the last letter. Then $\gamma(n,S') \le 2k\gamma(n,S)$ and hence S' is exponentially F-negligible by Lemma 5.3.

Suppose $w \in F - S'$ is such that $(N+2M)/|w| < \epsilon/2$. Let u be a cyclically reduced word obtained from w by changing at most the last letter. Then $|u| = |w|$ and $u \in \mathcal{CR} - S$.

Thus $d_A(w,u) \le 2$ and hence $d_X(wp, up) \le 2M$. Thus $||u|_p - |w|_p| \le 2M$. Also $|||u||_X - |u|_p| \le N$. Therefore $|||u||_X - |w|_p| \le N + 2M$. Since $u \in \mathcal{CR} - S$, we have $|||u||_X - \lambda||u||| < \epsilon||u||$. Since $||u|| = |u| = |w|$, we have

$$||w|_p - \lambda|w|| < \epsilon|w| + N + 2M,$$
$$\Big|\frac{|w|_p}{|w|} - \lambda\Big| < \frac{\epsilon}{2} + \frac{N+2M}{|w|} < \epsilon.$$

The set $\{w \in F : (N+2M)/|w| \ge \epsilon/2\}$ is finite. Hence $S' \cup \{w \in F : (N+2M)/|w| \ge \epsilon/2\}$ is exponentially F-negligible and assertion (3) holds.

The proof of (4) is similar to that of (3) and we leave the details to the reader. ∎

THEOREM 5.8: *Let $F = F(a_1, \dots, a_k)$ and let $\phi : F \to Aut(X)$ be a free simplicial action of $F = F(a_1, \dots, a_k)$ on a simplicial tree X.*

Then the generic stretching factor $\lambda(\phi)$ is also an approximate conjugacy stretching factor (and thus by Proposition 5.7 an approximate stretching factor).

Proof: The proof is very similar to that of Theorem 4.3. Since the minimal F-invariant subtree of X contains the axes of all the nontrivial elements of F, we may again assume that the action of F on X is minimal.

Choose a vertex $x \in X$. Recall, that, using the notation from the proof of Theorem 4.3, for any $w \in F$

$$\Big|\sum_{a\in Z_1}\sum_{|u|_A=M} c(u,a)f_w(u) + \sum_{z\in Z_2}\sum_{|u|_A=M} r_z c(u,z) f_w(u) - (|g|_p/n)\Big| \le (L'/n).$$

It follows from Lemma 4.1 and Lemma 4.2 that if $w \in F$ is cyclically reduced over B_T then $|||w||_X - |w|_x| \le N$, where $N = N(x) > 0$ is some constant independent of w. On the other hand, by the Bounded Cancellation Lemma (Lemma 3.7) there exists a constant $N' > 0$ such that for any cyclically reduced word w over A, we have $|||w||_{B_T} - |w|_{B_T}| \le N'$. Hence for any cyclically reduced word w over A we have $|||w||_X - |w|_x| \le N''$ where $N'' = N''(x) > 0$ is some constant independent of w.

Therefore, for any cyclically reduced $w \in F$ over A with $||w|| = n$

$$(\dagger)\quad \Big|\sum_{a\in Z_1}\sum_{|u|_A=M} c(u,a)f_w(u) + \sum_{z\in Z_2}\sum_{|u|_A=M} r_z c(u,z) f_w(u) - \frac{||w||_X}{n}\Big| \le \frac{L'+N}{n}.$$

Let $\epsilon > 0$. We know that the set

$$C(M,\epsilon) := \Big\{w \in \mathcal{CR}\text{: for any } u \ne 1 \text{ with } |u| = M \text{ and for the cyclic word } (w) \text{ we have } \Big|f_{(w)}(u) - \frac{1}{2k(2k-1)^{M-1}}\Big| < \epsilon\Big\}$$

is exponentially $\mathcal{CR}$-generic.

Hence $(\dagger)$ implies that for any $w \in C(M,\epsilon)$

$$\Big|\frac{1}{2k(2k-1)^{M-1}}\Big(\sum_{a\in Z_1}\sum_{|u|_A=M} c(u,a) + \sum_{z\in Z_2}\sum_{|u|_A=M} r_z c(u,z)\Big) - \frac{||w||_X}{n}\Big| \le \frac{N_1}{n} + N_1\epsilon$$

for some constant $N_1 > 0$ independent of w and ϵ.

Thus by definition the number

$$\frac{1}{2k(2k-1)^{M-1}}\left(\sum_{a\in Z_1}\sum_{|u|_A=M} c(u,a)+\sum_{z\in Z_2}\sum_{|u|_A=M} r_z c(u,z)\right)$$

is an approximate conjugacy stretching factor for ϕ. In the proof of Theorem 4.3 we obtained the same formula for $\lambda(\phi)$. ∎

LEMMA 5.9: *Let $F = F(a_1,\dots,a_k)$ and let ϕ: $F \to Aut(X)$ be a free simplicial action of $F = F(a_1,\dots,a_k)$ on a simplicial tree X. Let $\mu \geq 0$.*

Suppose there exists an exponentially $\mathcal{CR}$-generic set S such that for any $w \in S$

$$\frac{||w||_X}{||w||} \geq \mu.$$

Then $\lambda(\phi) \geq \mu$.

Proof: Suppose, on the contrary, that $\lambda(\phi) < \mu$. Choose $\epsilon > 0$ such that $\lambda(\phi) + \epsilon < \mu$.

Then there is an exponentially $\mathcal{CR}$-generic set R of cyclically reduced words such that for any $w \in R$

$$\frac{||w||_X}{||w||} \leq \lambda + \epsilon.$$

The intersection $S \cap R$ is exponentially $\mathcal{CR}$-generic and hence nonempty. Take $w \in S \cap R$.

Then

$$\mu \leq \frac{||w||_X}{||w||} \leq \lambda + \epsilon < \mu,$$

yielding a contradiction. ∎

6. Whitehead's Peak Reduction and rigidity of free group automorphisms

We need to recall some definitions related to Whitehead's algorithm for solving the automorphic equivalence problem in a free group. We refer the reader to [39, 44] for a detailed exposition.

Definition 6.1 (Whitehead automorphisms)*:* A *Whitehead automorphism* of F is an automorphism τ of F of one of the following two types:

(1) There is a permutation t of $\widehat{A}$ such that $\tau|_{\widehat{A}} = t$. In this case τ is called a *relabeling automorphism* or a *Whitehead automorphism of the first kind.*

(2) There is an element $a \in \widehat{A}$, the *multiplier*, such that for any $x \in \widehat{A}$

$$\tau(x) \in \{x, xa, a^{-1}x, a^{-1}xa\}.$$

In this case we say that τ is a *Whitehead automorphism of the second kind.* (Note that we always have $\tau(a) = a$ in this case since τ is an automorphism of F.) To every such τ we associate a pair (S, a) where a is as above and S consists of all those elements of $\widehat{A}$, including a but excluding a^{-1}, such that $\tau(x) \in \{xa, a^{-1}xa\}$. We will say that (S, a) is the *characteristic pair* of τ.

Note that for any $a \in \widehat{A}$ the inner automorphism $ad(a)$ is a Whitehead automorphism of the second kind.

The following important result of Whitehead is known as the "peak reduction lemma":

Proposition 6.2: *Let u, v be cyclic words with $||u|| \le ||v||$ and let $\phi \in Aut(F)$ be such that $\phi(u) = v$. Then we can write ϕ as a product of Whitehead moves*

$$\phi = \tau_p \cdots \tau_1$$

so that for each $i = 1, \dots, p$

$$||\tau_i \cdots \tau_1(u)|| \le ||v||.$$

Moreover, if $||u|| < ||v||$ then the above inequalities are strict for all $i < p$.

Definition 6.3 (Weighted Whitehead graph)*:* Let w be a nontrivial cyclically reduced word in $\widehat{A}^*$. The *weighted Whitehead graph* Γ_w *of* w is defined as follows. Let (w) be the cyclic word defined by w. The vertex set of Γ_w is $\widehat{A}$. For every $x, y \in \widehat{A}$ such that $x \ne y^{-1}$ there is an undirected edge in Γ_w from x^{-1} to y labeled by the sum $\hat{n}_w(xy) := n_{(w)}(xy) + n_{(w)}(y^{-1}x^{-1})$.

There are $k(2k-1)$ undirected edges in Γ_w. Edges may have label zero, but there are no edges from a to a for $a \in \widehat{A}$. It is easy to see that we have $\Gamma_w = \Gamma_v$ for any cyclic permutation v of w or w^{-1}.

Convention 6.4: Let w be a fixed nontrivial cyclically reduced word. For two subsets $X, Y \subseteq \widehat{A}$ we denote by $X.Y$ the sum of all edge-labels in the weighted Whitehead graph Γ_w of w of edges from elements of X to elements of Y. Thus for $x \in \widehat{A}$ the number $x.\widehat{A}$ is equal to $n_w(x) + n_w(x^{-1})$, the total number of occurrences of $x^{\pm 1}$ in w.

The next lemma, which is Proposition 4.16 of Ch. I in [39], gives an explicit formula for the difference of the lengths of w and $\tau(w)$, where τ is a Whitehead automorphism.

Lemma 6.5: *Let w be a nontrivial cyclically reduced word and let τ be a Whitehead automorphism of the second kind with the characteristic pair (S, a). Let $S' = \widehat{A} - S$. Then*

$$||\tau(w)|| - ||w|| = S.S' - a.\widehat{A}.$$

The following important notion was introduced by Kapovich, Schupp and Shpilrain in [36].

Definition 6.6 (Strict Minimality)*:* A nontrivial cyclically reduced word w in F is *strictly minimal* if for every non-inner Whitehead automorphism τ of F of the second kind we have

$$||\tau(w)|| > ||w||.$$

The set of all strictly minimal elements in F is denoted SM.

An immediate consequence of the Peak Reduction Lemma is:

Proposition 6.7 ([36]): *Let $w \in F$ be a cyclically reduced strictly minimal element. Then w is of minimal length in its $Aut(F)$-orbit and for any $\phi \in Aut(F)$ we have*

$$|w| = ||w|| \leq ||\phi(w)|| \leq |\phi(w)|.$$

Theorem 6.8: *Put $c_0 := 1 + (2k-3)/(4k^2-2k)$. There exists an exponentially F-generic set $W \subseteq F$ with the following property.*

For any $\phi \in Aut(F)$ the following conditions are equivalent:

(1) *The automorphism ϕ is simple.*

(2) *We have $\lambda(\phi) = 1$.*

(3) *We have $\lambda(\phi) < 1 + (2k-3)/(2k^2-k)$.*

(4) *For some $w \in W$ we have $||\phi(w)|| = ||w||$.*

(5) *For every $w \in W$ we have $||\phi(w)|| = ||w||$.*

(6) *For some $w \in W$ we have $||\phi(w)|| \leq c_0||w||$.*

(7) *For every $w \in W$ we have $||\phi(w)|| \leq c_0||w||$.*

Proof: It is obvious that (1) implies (2) and that (2) implies (3).

We will now prove that (3) implies (1).

Let $\phi \in Aut(F)$.

Let $\epsilon > 0$ be arbitrary. Put $T(\epsilon)$ be the set of all cyclically reduced words w such that:

(1) For any $x \in \widehat{A}$ $|f_w(x) - 1/2k| \le \epsilon/2$.

(2) For any $x, y \in \widehat{A}$ with $x \neq y^{-1}$ $|f_w(xy) - 1/(2k(2k-1))| \le \epsilon/2$.

Then $T(\epsilon)$ is exponentially $\mathcal{CR}$-generic [36]. Moreover, every $w \in T(\epsilon)$ is strictly minimal [36], provided that $\epsilon < (2k-3)/(k(2k-1)(4k-3))$.

Suppose now that $\epsilon < \epsilon_0 := (2k-3)/(k(2k-1)(4k-3))$. Choose an arbitrary element $w \in T(\epsilon)$ and denote $n = ||w||$.

By strict minimality of w we have $||w|| \le ||\phi(w)||$. Moreover, by Proposition 6.2 (Peak Reduction Lemma) there is a decomposition $\phi = \tau_p\tau_{p-1}\cdots\tau_1$ such that each τ_i is a Whitehead move and for each $i = 1, \dots, p-1$

$$||\tau_i\tau_{i-1}\cdots\tau_1(w)|| \le ||\phi(w)||$$

with strict inequalities unless $||w|| = ||\phi(w)||$.

Suppose first that $||w|| = ||\phi(w)||$. Then all inequalities above are equalities and by strict minimality of w each τ_i is either inner or a relabeling automorphism. This implies that $\phi = \alpha\tau$ where α is inner and τ is a relabeling automorphism and that $\lambda(\phi) = 1$.

Suppose now that $||w|| < ||\phi(w)||$. Then the preceding argument shows that in fact for any $z \in T(\epsilon)$ we have $||z|| < ||\phi(z)||$ (since otherwise ϕ would be simple and so $||w|| = ||\phi(w)||$).

Denote $z_0 = z$ and $z_i = \tau_i\tau_{i-1}\cdots\tau_1(z)$ for $0 < i \le p$. Thus $z_p = \phi(z)$. Since $||z|| < ||\phi(z)||$, there is some $i, 1 \le i \le p$ such that τ_i is a non-inner Whitehead move of the second kind. Let j be the smallest i with this property. Then all τ_i with $i < j$ are either inner or relabeling automorphisms, $||z|| = ||z_i||$ and $z_i \in T(\epsilon)$. In particular, $z_{j-1} \in T(\epsilon)$ and z_{j-1} is strictly minimal.

Thus

$$n = ||z|| = ||z_{j-1}|| \le ||z_j|| = ||\tau_j(z_{j-1})|| < ||\phi(z)||.$$

Let (S, a) be the characteristic pair of τ_j and let $S' = \widehat{A} - S$. Since τ_j is non-inner, we have both $|S| \ge 2$ and $|S'| \ge 2$. Hence $|S||\ S'| \ge 2(2k-2)$ and there are at least $2(2k-2)$ edges between S and S' in the weighted Whitehead graph of z_{j-1}. Recall that $a.\widehat{A}$ is the total number of occurrences of $a^{\pm 1}$ in z.

By Lemma 6.5, we have $||\tau_j(z_{j-1})|| - ||z|| = S.S' - a.\widehat{A}$.

By assumption on z_{j-1} we have $a.\widehat{A} \le n(1/k + \epsilon)$ and so

$$||\tau_j(z_{j-1})|| - ||z_{j-1}|| = S.S' - a.\widehat{A} \ge 2n(2k-2)\Big(\frac{1}{k(2k-1)} - \epsilon\Big) - n\Big(\frac{1}{k} + \epsilon\Big).$$

Hence

$$||\phi(z)|| \ge ||z_j|| \ge n + 2n(2k-2)\Big(\frac{1}{k(2k-1)} - \epsilon\Big) - n\Big(\frac{1}{k} + \epsilon\Big)$$

and so, since $n = ||z||$,

$$\frac{||\phi(z)||}{||z||} \geq 1 + (4k-4)\Big(\frac{1}{k(2k-1)} - \epsilon\Big) - \Big(\frac{1}{k} + \epsilon\Big).$$

Note that the above inequality holds for any element $z \in T(\epsilon)$.

Since $T(\epsilon)$ is exponentially $\mathcal{CR}$-generic, this implies by Lemma 5.9 that

$$\lambda(\phi) \geq 1 + (4k-4)\Big(\frac{1}{k(2k-1)} - \epsilon\Big) - \Big(\frac{1}{k} + \epsilon\Big).$$

Since $0 < \epsilon < \epsilon_0$ was arbitrary, it follows that

$$\lambda(\phi) \geq 1 + (4k-4)\frac{1}{k(2k-1)} - \frac{1}{k} = 1 + \frac{2k-3}{2k^2-k} > 1.$$

This proves that (3) implies (1), so that (1), (2) and (3) are equivalent.

Choose $0 < \epsilon_1 < \epsilon_0$ such that

$$1 + (4k-4)\Big(\frac{1}{k(2k-1)} - \epsilon_1\Big) - \Big(\frac{1}{k} + \epsilon_1\Big) < c_0 = 1 + \frac{2k-3}{4k^2-2k}.$$

Put $W = T(\epsilon_1)$. The above argument shows that if for some $w \in W$ we have

$$\frac{||\phi(z)||}{||z||} < 1 + (4k-4)\Big(\frac{1}{k(2k-1)} - \epsilon_1\Big) - \Big(\frac{1}{k} + \epsilon_1\Big)$$

then ϕ is simple.

With this W we have proved that (5) implies (1). It is obvious that (1) implies (4)–(7) and that each of (4), (5), (7) implies (6). Thus statements (1), (4), (5), (6), (7) are equivalent.

We already know that (1), (2) and (3) are equivalent. This completes the proof of the theorem. ∎

The following statement, together with Theorem 6.8, implies Theorem F from the Introduction.

Corollary 6.9: *Let $F = F(a_1, \dots, a_k)$, where $k \geq 2$, and d be the word metric on F corresponding to the generating set $A = \{a_1, \dots, a_k\}$. Let $\phi \in Aut(F)$. Then the following conditions are equivalent:*

(1) *The automorphism ϕ is simple.*
(2) *The map ϕ: $(F, d) \to (F, d)$ is a rough isometry.*
(3) *The map ϕ: $(F, d) \to (F, d)$ is a rough similarity.*

Proof: It is obvious that (1) implies (2) and that (2) implies (3).

We will now show that (3) implies (1). Suppose that ϕ is a rough similarity, so that there exist $\lambda > 0$ and $D > 0$ such that for any $w \in F$

$$\lambda|w| - D \le |\phi(w)| \le \lambda|w| + D.$$

Then obviously $\lambda = \lambda(\phi)$. By Theorem 6.8 either ϕ is simple or $\lambda(\phi) > 1$.

Assume the latter. Put $\lambda_0 = (1+\lambda)/2$. Thus $1 < \lambda_0 < \lambda$.

Consider the ball B_n of radius n in F, where $n >> 1$. For any $w \in F$ with $|w| \ge n/\lambda_0$ we have

$$|\phi(w)| \ge \lambda|w| - D \ge \lambda n/\lambda_0 - D > n,$$

so that $\phi(w) \notin B_n$.

Thus only the elements of length $\le n/\lambda_0$ may be potentially taken to B_n by ϕ. The number of such elements is smaller than $\#B_n$ since $\lambda_0 > 1$ and $n/\lambda_0 < n$. This contradicts the fact that ϕ: $(F,d) \to (F,d)$ is a bijection. Therefore ϕ is simple, as required. ∎

The following is Theorem G from the Introduction:

Theorem 6.10: *Let $F = F(a_1, \dots, a_k)$ where $k \ge 2$. Let ϕ: $F \to X$ be a free minimal action on F on a simplicial tree X without inversions.*

Then exactly one of the following occurs:

(1) *There is a simple automorphism α of F such that X is $\phi\circ\alpha$-equivariantly isomorphic to the Cayley graph of F with respect to $\{a_1, \dots, a_k\}$. In this case $\lambda(\phi) = 1$.*

(2) *We have $\lambda(\phi) \ge 1 + 1/k(2k-1)$.*

Proof: Let $\Gamma = X/F$ and let T be a maximal tree in Γ and let $B = \{b_1, \dots, b_k\}$ be the geometric basis of F corresponding to T. Let $\psi \in Aut(F)$ be defined by $\alpha(b_i) = a_i$ for $i = 1, \dots, k$.

Note that because of Lemma 4.2 for any cyclic word w over B we have $||w||_X \ge ||w||_B$. Suppose first that α is not a simple automorphism. Then $\lambda(\alpha) \ge 1 + (2k-3)/k(2k-1)$.

Hence for every $\epsilon > 0$ there exists an exponentially $\mathcal{CR}$-generic set $R(\epsilon) \subseteq \mathcal{CR}$ such that for any $w \in R(\epsilon)$

$$\frac{||w||_B}{||w||_A} = \frac{||\alpha(w)||_A}{||w||_A} \ge 1 + \frac{2k-3}{k(2k-1)} - \epsilon.$$

Since $||w||_X \ge ||w||_B$, it follows that

$$\frac{||w||_X}{||w||_A} \ge 1 + \frac{2k-3}{k(2k-1)} - \epsilon.$$

Since $\epsilon > 0$ was arbitrary, it follows by Lemma 5.9 that

$$\lambda(\phi) \geq 1 + \frac{2k-3}{k(2k-1)} \geq 1 + \frac{1}{k(2k-1)},$$

as required.

Suppose now that α is a simple automorphism. We will assume that $\alpha = Id$, and it will be easily seen that the general case is similar.

If Γ is a wedge of k loop-edges then the statement of the theorem holds. Suppose Γ is not of this form. Then there exist edges $e, e' \in E(\Gamma - T)$, $e' \neq e^{-1}$, such that $[t(e), o(e')]_T$ has length at least 1. Let z be the freely reduced word of length 2 in B corresponding to the sequence ee'. Let $\epsilon > 0$ be arbitrary. Let $C(2, \epsilon/2) \subseteq \mathcal{CR}$ be defined as in Proposition 5.5. Thus $C(2, \epsilon/2)$ consists of all cyclically reduced words w' such that for the cyclic word $w = (w')$ and for every freely reduced word xy in A

$$|f_w(xy) - \frac{1}{2k(2k-1)}| \leq \epsilon/2.$$

Then $C(2, \epsilon/2)$ is exponentially $\mathcal{CR}$-generic. Let $w' \in C(2, \epsilon/2)$ be arbitrary and let $w = (w')$. Note that $||w'||_A = ||w'||_B = ||w||_A = ||w||_B$ and $||w||_X = ||w'||_X$. Then

$$||w||_X \geq ||w||_B + n_w(z) + n_w(z^{-1}) = ||w||_A + n_w(z) + n_w(z^{-1})$$

and so

$$\frac{||w'||_X}{||w'||_A} = \frac{||w||_X}{||w||_A} \geq 1 + f_w(z) + f_w(z^{-1}) \geq 1 + \frac{1}{k(2k-1)} - \epsilon.$$

Since $\epsilon > 0$ was arbitrary, Lemma 5.9 implies that $\lambda(\phi) \geq 1 + 1/k(2k-1)$, as required. ∎

7. Application to the geometry of automorphisms

Definition 7.1: Let $F = F(a_1, \ldots, a_k)$. An automorphism ϕ of F is said to be (s, m)-*hyperbolic*, where $s > 1$ and $m \geq 1$ is an integer, if for every nontrivial cyclic word w we have

$$s||w|| \leq \max\{||\phi^m(w)||, ||\phi^{-m}(w)||\}.$$

An automorphism is *hyperbolic* if it is (s, m)-hyperbolic for some $s > 1, m \geq 1$.

The following lemma is an easy consequence of the above definition:

Lemma 7.2: *Let $\phi \in Aut(F)$ be (s,m)-hyperbolic and let w be a cyclic word of minimal length in its $\langle\phi\rangle$-orbit. Then for any $n \geq 2$ we have*

$$||\phi^{mn}(w)|| \geq s^{n-1}||w||.$$

Proof:

By definition of hyperbolicity of ϕ we have

$$(\ddagger). \qquad ||\phi^{-m}(u)|| \leq ||u|| \Rightarrow s||u|| \leq ||\phi^{m}(u)||$$

Note that by the choice of w we have $||w|| \leq ||\phi^m(w)||$. Hence applying $(\ddagger)$ with $u = \phi^m(w)$ we get $s||\phi^m(w)|| \leq ||\phi^{2m}(w)||$. Then, using $(\ddagger)$, by induction on n we see that for any $n \geq 1$

$$||\phi^{m(n+1)}(w)|| = ||\phi^{mn+m}(w)|| \geq s||\phi^m(w)||.$$

This in turn implies that for any $n \geq 1$

$$||\phi^{m(n+1)}(w)|| = ||\phi^{mn+m}(w)|| \geq s^n||\phi^m(w)|| \geq s^{n-1}||w||.$$

This proves Lemma 7.2. ∎

The following is Theorem E from the Introduction:

Theorem 7.3: *Let ϕ be an (s,m)-hyperbolic automorphism of F. Then*

$$\liminf_{n\to\infty} \sqrt[n]{\lambda(\phi^n)} \geq \sqrt[m]{s} > 1.$$

Proof: Let $t \geq 2$ be an arbitrary integer. Let $w \in SM$ be a strictly minimal element. Since w is minimal in its $Aut(F)$-orbit, it is also minimal in its $\langle\phi\rangle$-orbit. Therefore by Lemma 7.2

$$||\phi^{tm}(w)|| \geq s^{t-1}||w|| \quad \text{and} \quad \frac{||\phi^{tm}(w)||}{||w||} \geq s^{t-1}.$$

Since SM is exponentially $\mathcal{CR}$-generic, Lemma 5.9 implies that $\lambda(\phi^{tm}) \geq s^{t-1}$.

Moreover, there is $D > 0$ such that for any cyclically reduced word u we have

$$||\phi^i(u)|| \geq D||u||, \quad \text{for all } 0 \leq i < m.$$

Let $n \geq 2m$ be an integer. Then we can write n as $n = mt + i$ where $t \geq 2$ and $0 \leq i < m$. For any $w \in SM$ we have

$$||\phi^n(w)|| = ||\phi^{mt+i}(w)|| \geq D||\phi^{mt}(w)|| \geq Ds^{t-1}||w||$$

and hence

$$\frac{||\phi^{tm}(w)||}{||w||} \geq Ds^{t-1}.$$

Since SM is exponentially $\mathcal{CR}$-generic, Lemma 5.9 again implies that for any $n \geq 2m$

$$\lambda(\phi^n) \geq Ds^{t-1} = \frac{D}{s}s^{\lfloor n/m \rfloor}.$$

This implies

$$\liminf_{n\to\infty} \sqrt[n]{\lambda(\phi^n)} \geq s^{1/m} > 1,$$

as claimed. ∎

We can now prove Corollary I from the Introduction:

Corollary 7.4: *Let $F = F(a_1, \ldots, a_k)$ be a free group of rank $k \geq 2$, equipped with the standard metric.*

Then for any $\phi \in Aut(F)$ we have

$$flux(\phi) = \begin{cases} 0, & \textit{if } \phi \textit{ is a relabeling automorphism,} \\ 1, & \textit{otherwise.} \end{cases}$$

Proof: Let $\lambda = \lambda(\phi)$ be the generic stretching factor.

Suppose first that $\lambda > 1$. Then the set

$$T := \left\{ w \in F : \frac{|\phi(w)|}{|w|} > \frac{\lambda+1}{2} \right\}$$

is exponentially F-generic.

Let $B(n)$ be the ball of radius n in F and let $w \in B(n) \cap T$ be such that $2n/(\lambda+1) \leq |w| \leq n$. Then

$$|\phi(w)| > |w|\frac{\lambda+1}{2} \geq \frac{2n}{\lambda+1}\frac{\lambda+1}{2} = n.$$

Hence for each $w \in [B(n) \cap T] - B(2n/(\lambda+1))$ we have $|\phi(w)| > n$. The size of $B(2n/(\lambda+1))$ is exponentially smaller than that of $B(n)$ since $2/(\lambda+1) < 1$. Hence by exponential genericity of T

$$\frac{\#[B(n) \cap T] - \#B(2n/(\lambda+1))}{\#B(n)} \longrightarrow_{n\to\infty} 1 \text{ exponentially fast.}$$

Hence

$$\lim_{n\to\infty} \frac{flux_\phi(n)}{\#B(n)} = 1$$

and therefore $flux(\phi) = 1$.

Suppose now that $\lambda(\phi) = 1$. By Theorem 6.8 this implies that $\phi = \alpha\tau$ where α is inner and τ is a relabeling automorphism.

If $\alpha = 1$, then obviously $flux(\phi) = 0$. Suppose now that α is nontrivial. Since τ acts as a permutation on each ball and each sphere in F, we can assume that $\tau = 1$ and $\phi = \alpha$. Thus there is $u \in F, u \neq 1$ such that for every $w \in F$, $\phi(w) = uwu^{-1}$. There are $\geq c_1(2k-1)^n$ elements f with $|w| = n$ such that the product uwu^{-1} is freely reduced as written, where $c_1 > 0$ is a constant independent of n and u. For each such element we have $|\phi(w)| > |w|$. Hence there is a constant $c_2 \in (0,1)$ independent of n and u such that for any $n > 0$

$$1 \geq \frac{flux_\phi(n)}{\#B(n)} \geq c_2 > 0.$$

Hence

$$1 \geq flux(\phi) = \lim_{n\to\infty} \sqrt[n]{\frac{flux_\phi(n)}{\#B(n)}} \geq \lim_{n\to\infty} \sqrt[n]{c_2} = 1.$$

Thus $flux(\phi) = 1$ and the proof is complete. ∎

8. Random elements in regular languages

The most reasonable way of choosing a “random” element in the regular language L is via a random walk in the transition graph of an automaton M accepting L. It turns out that the natural model of computation here is that of a *non-deterministic finite automaton* or NDFA. Such an automaton M over an alphabet A with *state set* Q is specified by a finite directed graph $\Gamma(M)$. The vertex set of $\Gamma(M)$ is the set Q of states of M and Q comes equipped with a distinguished nonempty subset I of *initial* or *start* states. The directed edges of $\Gamma(M)$ are labelled by elements of A and these edges are treated as transitions of M. If $q \in Q$ is a state and $a \in A$ is a letter, we allow multiple edges labelled a with origin q and we also allow the case when there are no such edges. Nondeterministic automata are thus by their nature “partial”. There is a distinguished subset of Q of *accepting* states. A word w over A is said to be *accepted* by M if there exists a directed path with label w in $\Gamma(M)$ from some initial state to an accepting state. The *language*, $L(M)$, accepted by M is the collection of all words accepted by M.

We will also use directed graph $\Gamma_1(M)$ defined as follows. The vertex set of $\Gamma_1(M)$ is the set of directed edges $E(\Gamma(M))$ of M. If $e_1, e_2 \in E(\Gamma(M))$ the pair (e_1, e_2) defines a directed edge from e_1 to e_2 in $\Gamma_1(M)$ if the terminus of e_1 is the origin of e_2, that is, e_1, e_2 is a directed edge-path in $\Gamma(M)$.

Definition 8.1 (Normal Automaton)*:* Let A be a finite alphabet. A *normal* automaton over a finite alphabet A is a nondeterministic finite state automaton M over A such that the following conditions hold:

(1) the automaton M has a nonempty set of accept states;
(2) the directed graph $\Gamma(M)$ has at least one edge;
(3) the directed graph $\Gamma(M)$ is *strongly connected*, that is for any two states q, q' of M there exists a directed edge-path from q to q' in $\Gamma(M)$.

The third condition in the above definition is the most important one as it is responsible for the irreducibility of a Markov chain naturally associated to a normal automaton:

Definition 8.2 (Associated Markov chain)*:* Let M be a normal automaton over a finite alphabet A. We define an *associated finite state Markov chain M'* as follows. The set of states of M' is the set E of directed edges of $\Gamma(M)$. If the origin of f is not the terminus of e we put the transition probability $p_{e,f} = 0$. If the origin of f is equal to the terminus of e we put $p_{e,f} = 1/m$, where m is the total number of outgoing directed edges from the terminus of e.

Convention 8.3: Note that the sample space Ω for the Markov chain M' defined above consists of all semi-infinite directed edge-paths

$$\omega = e_1, e_2, \ldots, e_n, \ldots$$

in the graph $\Gamma(M)$. Every such path has a label

$$w(\omega) = a_1 a_2 \cdots,$$

that is a semi-infinite word over the alphabet A. We will denote $w_n = w_n(\omega) := a_1 \cdots a_n$, the initial segment of length n of w. The set Ω comes equipped with the natural topology, where we think of Ω as the union of boundaries of rooted trees $(T_e)_e \in E$. The vertices of T_e are finite edge-path in $\Gamma(M)$ beginning with e. The Borel σ-algebra on Ω is generated by the following open-closed *cylinder sets $Cyl(\gamma)$*, where γ is a nonempty finite edge-path in $\Gamma(M)$:

$$Cyl(\gamma) := \{\omega \in \Omega : p \text{ is the initial segment of } \omega\}.$$

If we put an initial probability distribution μ on E, this defines a Borel probability measure P_μ on Ω. This measure is defined on the cylinder sets by the standard convolution formula. If $\gamma = e_1, \ldots, e_n$, where $n > 1$, then

$$P_\mu(Cyl(\gamma)) := \mu(e_1) p_{e_1,e_2} p_{e_2,e_3} \cdots p_{e_{n-1},e_n}.$$

If $n = 1$ then $P_\mu(Cyl(e)) := \mu(e)$.

LEMMA 8.4: *Let M be a normal automaton. Then the associated finite state Markov chain M' is irreducible. In particular, there is a unique stationary initial probability distribution μ_0 on the set of states E of M'. This distribution has the property $\mu_0(e) > 0$ for each $e \in E$.*

Proof: To show that M' is irreducible we have to prove that for any two edges $e, f \in E$ there is $n > 0$ such that the n-step transition probability $p_{e,f}^{(n)} > 0$. Since $\Gamma(M)$ is strongly connected, there exists a directed edge-path γ in $\Gamma(M)$ from the terminus of e to the origin of f. Then $e\gamma f$ is a directed edge-path in $\Gamma(M)$ that starts with e and ends with f. Hence $\Gamma_1(M)$ is strongly connected and therefore M' is irreducible.

The irreducibility of M' implies the existence and uniqueness of a positive stationary distribution μ_0 on E, as required. ∎

If we fix an initial probability distribution μ on E, this defines a probability measure P_μ on Ω.

LEMMA 8.5: *Let M be a normal automaton. Let M' be the associated finite state Markov chain and let μ_0 be the stationary initial distribution for M'. Let $Z \subseteq \Omega$ be a set such that $P_{\mu_0}(Z) = 0$. Then for any other initial distribution μ on E we have $P_\mu(Z) = 0$.*

Proof: Let μ and μ_0 be as above. Put

$$c := \max\left\{\frac{\mu(e)}{\mu_0(e)} : e \in E\right\}.$$

Note that $0 < c < \infty$ since $\mu_0(e) > 0$ for each $e \in E$. Consider an arbitrary cylinder set $Cyl(\gamma) \subset \Omega$, where $\gamma = e_1, e_2, \dots, e_n$. From the definitions of P_μ and P_{μ_0} we see that

$$P_\mu(Cyl(\gamma)) = \frac{\mu(e_1)}{\mu_0(e_1)} P_{\mu_0}(Cyl(\gamma)) \le cP_{\mu_0}(Cyl(\gamma)).$$

Hence for an arbitrary Borel set $Z \subseteq \Omega$ we have $P_\mu(Z) \le cP_{\mu_0}(Z)$. In particular, if $P_{\mu_0}(Z) = 0$ then $P_\mu(Z) = 0$. ∎

The previous two lemmas depend only on the automaton M being normal. Suppose now that $L = L(M)$. For each state q choose a shortest path from q to an accept state and let u_q be the word in A^* labelling that path. This is possible since $\Gamma(M)$ is strongly connected and the set of accept states is nonempty by the assumption on M. Note that u_q is the empty word if and only if q is an accept state. The lengths of u_q are bounded above by some constant depending on M. For a finite walk w_n denote $w_n' = w_n u_q$, where q is the state in which w_n ends. Note that if w_n begins in a state from I then $w_n' \in L$. Thus if μ is an distribution supported on the set of edges in $E(M)$ with initial vertices from I and w_n is obtained by performing n steps of the chain M' with initial distribution μ, then $w_n' \in L$ can be thought of as a "random" element of L.

We can now prove (a slight generalization of) Theorem J from the Introduction:

Theorem 8.6: *Let M be a normal automaton over the alphabet A and let $L = L(M)$ be the language accepted by M.*

Let ϕ: $A^* \to G$ be a monoid homomorphism, where G is a group with a left-invariant semi-metric d_G. Then there exists a number $\lambda = \lambda(M, \phi, d_G) \geq 0$ such that for any initial distribution μ on $E(M)$ we have

$$\lim_{n\to\infty} \frac{|\phi(w_n)|_G}{n} = \lim_{n\to\infty} \frac{|\phi(w_n')|_G}{n} = \lambda \text{ almost surely and in } L^1 \text{ with respect to } P_\mu.$$

Proof: Let μ_0 be the unique stationary initial distribution for M'. As before denote by $\mathcal{S}$: $\Omega \to \Omega$ the shift operator which erases the first edge of every $\omega = e_1, e_2, \cdots \in \Omega$. Stationarity of μ_0 means that $\mathcal{S}$: $(\Omega, P_{\mu_0}) \to (\Omega, P_{\mu_0})$ is a measure-preserving map. Since M' is irreducible and aperiodic, $\mathcal{S}$ is also ergodic.

As before, define X_n: $\Omega \to \mathbb{R}$ as

$$X_n(\omega) := |\phi(w_n(\omega))|_G.$$

Then again it is easy to see that $X_n \geq 0$, $X_{n+m}(\omega) \leq X_n(\omega) + X_m(\mathcal{S}^n\omega)$. Hence by the Subadditive Ergodic Theorem there is $\lambda \geq 0$ and there is a subset $Q \subseteq \Omega$ with $P_{\mu_0}(Z) = 0$ such that for any $\omega \notin Z$

$$\lim_{n\to\infty} \frac{|\phi(w_n(\omega))|_G}{n} = \lambda.$$

Let μ be an arbitrary initial distribution on E. Then by Lemma 8.5 we have $P_\mu(Z) = 0$. Thus

$$\frac{|\phi(w_n)|_G}{n} \to \lambda \text{ almost surely with respect to } P_\mu.$$

Note that by the left-invariance of d_G we have $|\phi(w)|_G \leq K|w|$ where $K = \max\{|\phi(a)|_G : a \in A\}$. Hence $X_n/n = |\phi(w_n)|_G/n \leq K$ and by the Dominated Convergence Theorem almost sure convergence of X_n/n implies L^1-convergence.

Since d_G is a seminorm on G and the length of any path w'_n differs from $|w_n|$ by at most a fixed constant, it is also true that $|\phi(w'_n)|_G$ differs from $|\phi(w_n)|_G$ by at most a fixed constant and thus it is also the case that

$$\lim_{n\to\infty} \frac{|\phi(w'_n(\omega))|_G}{n} = \lim_{n\to\infty} \frac{|\phi(w_n(\omega))|_G}{n} = \lambda. \qquad \blacksquare$$

There is substantial flexibility in the choice of the Markov chain M'. The proof of Theorem 8.6 goes through without change for any choice of transition probabilities in M' such that $p_{e,f} > 0$ whenever (e, f) is an edge of $\Gamma_1(M)$ and $p_{e,f} = 0$ whenever (e, f) is not an edge of $\Gamma_1(M)$.

9. Open Problems

Problem 9.1: Let ϕ be an arbitrary (not necessarily injective) endomorphism of $F = F(a_1, \dots, a_k)$. Is $\lambda(\phi)$ rational? Computable?

Problem 9.2: Let $\phi \in Aut(F)$. What can be said about the behavior of $\lambda(\phi^n)$ as $n \to \infty$? Same for $\sqrt[n]{\lambda(\phi^n)}$. How are these quantities connected with growth rates of different (or perhaps just top) strata from relative train-track representatives of ϕ?

It is clear that the asymptotics of $\lambda(\phi^n)$ should reflect the dynamical properties of ϕ. For example, it is not hard to see that for any Nielsen automorphism τ the stretching factor $\lambda(\tau^n)$ grows at most linearly and $\limsup_{n\to\infty} \sqrt[n]{\lambda(\tau^n)} = 1$. On the other hand, for hyperbolic automorphisms ϕ Theorem 7.3 implies that $\liminf_{n\to\infty} \sqrt[n]{\lambda(\phi^n)} > 1$, so that the sequence $(\lambda(\phi^n))_n$ grows exponentially.

Problem 9.3: Can one estimate (say in the sense of Large Deviations) the speed of convergence $|\phi(\omega_n)|_G/n \to \lambda(\phi)$?

We have seen that in the case of free group automorphisms for any $\epsilon > 0$

$$P_n\Big(\frac{|\phi(\omega_n)|}{n} \in (\lambda(\phi) - \epsilon, \lambda(\phi) + \epsilon)\Big) \to 1$$

with exponentially fast convergence as $n \to \infty$. Are there any other situations where the speed of convergence in Theorem A can be estimated?

Problem 9.4: Let $F = F(a_1, \dots, a_k)$ where $k \geq 2$. Consider the set

$$W = \{\lambda(\phi) : \phi\colon F \to Aut(X) \text{ is a free simplicial action of } F \text{ on some simplicial tree } X\}.$$

We know that $W \subseteq \mathbb{Q}$ and, moreover $2kW \subseteq \mathbb{Z}[\frac{1}{2k-1}]$.

Is W a discrete subset of $\mathbb{Q}$?

Problem 9.5: The notion of a generic stretching factor for $\phi \in Aut(F)$ depends on the choice of a free basis $b = (a_1, \dots, a_k)$ of F and, more generally, on the choice of a finite generating set S of F and the corresponding word metric d_S. Denote by $\lambda_S(\phi)$ the generic stretching factor of ϕ considered as a map $(F, d_S) \to (F, d_S)$.

One can define the following uniform constants:

$$\lambda'(\phi) := \inf\{\lambda_b(\phi) : b \text{ is a free basis of } F\}$$

and

$$\lambda''(\phi) := \inf\{\lambda_S(\phi) : S \text{ is a finite generating set of } F\}.$$

(Note that $\lambda''(\phi)$ can be defined in the same fashion for an automorphism ϕ of an arbitrary finitely generated group G.)

For $\phi \in Aut(F)$, are the constants $\lambda'(\phi)$ and $\lambda''(\phi)$ actually realized by some free bases and finite generating sets of F accordingly? That is, are the above infima actually minima? Are $\lambda'(\phi)$ and $\lambda''(\phi)$ algorithmically computable?

Similarly we can define

$$||\phi||' = \inf\{||\phi||_b : b \text{ is a free basis of } F\}$$

and

$$||\phi||'' = \inf\{||\phi||_S : S \text{ is a finite generating set of } F\}.$$

Since both of these constants are integers, they are clearly realizable by some b and S accordingly. Are these constants algorithmically computable?

References

[1] G. Arzhantseva and A. Ol'shanskii, *Genericity of the class of groups in which subgroups with a lesser number of generators are free,* (Russian) Matematicheskie Zametki **59** (1996), 489–496.

[2] G. Arzhantseva, *On groups in which subgroups with a fixed number of generators are free,* (Russian) Fundamentalnaya i Prikladnaya Matematika **3** (1997), 675–683.

[3] G. Arzhantseva, *Generic properties of finitely presented groups and Howson's theorem,* Constructive Approximation **26** (1998), 3783–3792.

[4] G. Arzhantseva, *A property of subgroups of infinite index in a free group,* Proceedings of the American Mathematical Society **128** (2000), 3205–3210.

[5] F. Bonahon, *Bouts des variétés hyperboliques de dimension* 3, Annals of Mathematics (2) **124** (1986), 71–158.

[6] F. Bonahon, *The geometry of Teichmller space via geodesic currents,* Inventiones Mathematicae **92** (1988), 139–162.

[7] F. Bonahon, *Geodesic currents on negatively curved groups,* in *Arboreal Group Theory (Berkeley, CA, 1988)*, Mathematical Sciences Research Institute Publications, 19, Springer, New York, 1991, pp. 143–168.

[8] A. Borovik, A. G. Myasnikov and V. Shpilrain, *Measuring sets in infinite groups,* in *Computational and Statistical Group Theory* (R. Gilman et al., eds.), Contemporary Mathematics, **298** (2002), pp. 21–42.

[9] P. Brinkmann, *Hyperbolic automorphisms of free groups,* Geometric and Functional Analysis **10** (2000), 1071–1089.

[10] C. Champetier, *Petite simplification dans les groupes hyperboliques,* Annales de la Faculté des Sciences de Toulouse Mathematics (6) **3** (1994), 161–221.

[11] C. Champetier, *Propriétés statistiques des groupes de présentation finie,* Advances in Mathematics **116** (1995), 197–262.

[12] C. Champetier, *The space of finitely generated groups,* Topology **39** (2000), 657–680.

[13] P.-A. Cherix and A. Valette, *On spectra of simple random walks on one-relator groups,* With an appendix by Paul Jolissaint, Pacific Journal of Mathematics **175** (1996), 417–438.

[14] P.-A. Cherix and G. Schaeffer, *An asymptotic Freiheitssatz for finitely generated groups,* L'Enseignement Mathématique (2) **44** (1998), 9–22.

[15] J. Cohen, *Cogrowth and amenability of discrete groups,* Journal of Functional Analysis **48** (1982), 301–309.

[16] D. Cooper, *Automorphisms of free groups have finitely generated fixed point sets,* Journal of Algebra **111** (1987), 453–456.

[17] P. Dehornoy, *Braid-based cryptography,* in *Group Theory, Statistics, and Cryptography*, Contemporary Mathematics, 360, American Mathematical Society, Providence, RI, 2004, pp. 5–33.

[18] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications,* Second edition. Applications of Mathematics, 38, Springer-Verlag, New York, 1998.

[19] Y. Derriennic, *Sur le théorème ergodique sous-additif,* Comptes Rendus de l'Académie des Sciences, Paris Série A–B **281** (1975), Aii, A985–A988.

[20] A. Furman, *Coarse-geometric perspective on negatively curved manifolds and groups,* in *Rigidity in Dynamics and Geometry* (Cambridge, 2000), Springer, Berlin, 2002, pp. 149–166.

[21] E. Ghys, *Groupes Aléatoires,* Seminar Bourbaki (2002/2003), Asterisque **294** (2004), 173–204.

[22] R. Grigorchuck, *Symmetrical random walks on discrete groups,* in *Multicomponent Random Systems*, Adv. Probab. Related Topics, Vol. 6, Dekker, New York, 1980, pp. 285–325.

[23] M. Gromov, *Hyperbolic Groups,* in *Essays in Group Theory* (G. M. Gersten, ed.), MSRI Publ. **8**, 1987, pp. 75–263.

[24] M. Gromov, *Asymptotic invariants of infinite groups,* in *Geometric Group Theory*, Vol. 2 (Sussex, 1991), London Mathematical Society Lecture Note Series, 182, Cambridge University Press, Cambridge, 1993, pp. 1–295.

[25] M. Gromov, *Random walks in random groups,* Geometric and Functional Analysis **13** (2003), 73–146.

[26] Y. Guivarc'h, *Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire,* in *Conference on Random Walks* (Kleebach, 1979), 3, Astérisque, **74** Soc. Math. France, Paris, 1980, pp. 47–98.

[27] J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages and Computation,* Addison-Wesley, Reading, 1979.

[28] V. Kaimanovich, *Hausdorff dimension of the harmonic measure on trees,* Ergodic Theory and Dynamical Systems **18** (1998), 631–660.

[29] V. Kaimanovich, *The Poisson formula for groups with hyperbolic properties,* Annals of Mathematics (2) **152** (2000), 659–692.

[30] I. Kapovich, *The frequency space of a free group,* Internat. J. Alg. Comput. (Grigorchuk's 50s anniversary issue) **15** (2005), 939–969
http://www.arxiv.org/math.GR/0311053.

[31] I. Kapovich, *Currents on free groups,* in *Topological and Asymptotic Aspects of Group Theory* (R. Grigorchuk, M. Mihalik, M. Sapir and Z. Sunik, Editors), AMS Contemporary Mathematics Series, Vol. 394, 2006, pp. 149–176.

[32] I. Kapovich, A. Myasnikov, P. Schupp and V. Shpilrain, *Generic-case complexity, Decision problems in group theory and Random walks,* Jouranl of Algebra **264** (2003), 665–694.

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

[34] I. Kapovich and T. Nagnibeda, *The Patterson-Sullivan embedding and minimal volume entropy for outer space,* Geometric and Functional Analysis, to appear.
http://arxiv.org/abs/math.GR/0504445

[35] I. Kapovich and P. Schupp, *Genericity, the Arzhantseva-Ol'shanskii method and the isomorphism problem for one-relator groups,* Mathematische Annalen **331** (2005), 1–19.

[36] I. Kapovich, P. Schupp, and V. Shpilrain, *Generic properties of Whitehead's algorithm and isomorphism rigidity of random one-relator groups,* Pacific Journal of Mathematics **223** (2006), 113–140.

[37] J.F.C. Kingman, *Subadditive ergodic theory,*With discussion by D.L.Bürkholder, Daryl Daley, H. Kesten, P. Ney, Frank Spitzer and J. M. Hammersley, and a reply by the author, The Annals of Probability **1** (1973), 883–909.

[38] V. Kaimanovich, and A. Vershik, *Random walks on discrete groups: boundary and entropy,* The Annals of Probability **11** (1983), 457–490.

[39] R. Lyndon and P. Schupp, *Combinatorial Group Theory*, Springer-Verlag, Berlin, 1977. Reprinted in the "Classics in mathematics" series, 2000.

[40] A. G. Myasnikov and V. Shpilrain, *Automorphic orbits in free groups,* Journal of Algebra **269** (2003), 18–27.

[41] A. G. Myasnikov and V. Shpilrain, *Some metric properties of automorphisms of groups,* Journal of Algebra, to appear;
http://www.sci.ccny.cuny.edu/˜shpil/papers.html

[42] Y. Ollivier, *Sharp phase transition theorems for hyperbolicity of random groups,* Geometric and Functional Analysis **14** (2004), 595–679.

[43] A. Yu. Ol'shanskii, *Almost every group is hyperbolic,* International Journal of Algebra and Computation **2** (1992), 1–17.

[44] J. H. C. Whitehead, *On equivalent sets of elements in free groups,* Annals of Mathematics **37** (1936), 782–800.

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

[46] A. Zuk, *On property (T) for discrete groups,* in *Rigidity in Dynamics and Geometry* (Cambridge, 2000), Springer, Berlin, 2002, pp. 473–482.