Statistics of matrix products in hyperbolic geometry

M. Pollicott R. Sharp

December 17, 2009

Abstract

We consider central limit theorems and their generalizations for matrix groups acting co-compactly or convex co-compactly on the hyperbolic plane. We consider statistical results for the displacement in the hyperbolic metric, the action on the boundary and the relationship with classical matrix groups.

1 Introduction

In this note, we want to consider the statistical properties of the action of discrete groups on negatively curved spaces. Consider, the upper half plane $\mathbb{H}^2$ with the Poincaré metric and let $\Gamma \subset PSL(2,\mathbb{R})$ be either a co-compact or a convex co-compact subgroup. In the first case, Γ may be given by the standard one-relator presentation

$$\Gamma = \left\langle a_1, \ldots, a_{2\mathfrak{g}} : \prod_{i=1}^{\mathfrak{g}} [a_{2i-1}, a_{2i}] = e \right\rangle,$$

where $\mathfrak{g} \geq 2$ is the genus of $\mathbb{H}^2/\Gamma$. In the second case, Γ is a free group, which we assume to be given by $\Gamma = \langle a_1, \ldots, a_k \rangle$. In either case, we can associate a finite directed graph $\mathcal{G}$ whose edges are labelled by the generators and their inverses such that mapping a finite path to the product of its labels gives a natural bijection between paths of length n and elements of Γ with word length n. (In the co-compact case, our results also hold when $\mathbb{H}^2$ is replaced by the universal cover of a compact surface with variable negative curvature.)

There is a well known isometric action $PSL(2,\mathbb{R}) \times \mathbb{H}^2 \to \mathbb{H}^2$ by linear fractional transformations. Fix $x \in \mathbb{H}^2$. The space Σ of infinite paths $\underline{i} = (i_n)_{n=0}^{\infty}$ in the directed graph $\mathcal{G}$ gives rise to a sequence of images

$$g_{i_0}x,\ g_{i_0}g_{i_1}x,\ g_{i_0}g_{i_1}g_{i_2}x,\ \ldots,\ g_{i_0}g_{i_1}\cdots g_{i_{n-1}}x,\ \ldots \in \mathbb{H}^2,\ n \geq 1,$$

where $g_{i_0}g_{i_1}\cdots g_{i_{n-1}} \in \Gamma$ has word length n. We shall consider properties of the sequence $(g_{i_0}\cdots g_{i_{n-1}}x)_{n=1}^{\infty}$ for μ-a.e. $\underline{i} \in \Sigma$, where μ is the Gibbs measure μ associated to some Hölder continuous function on Σ. (Natural choices for μ are the measure of maximal entropy, corresponding naturally to the weak star limit of evenly distributed measure on words of the same length, for the shift map $\sigma : \Sigma \to \Sigma$ and the Gibbs state for which $\pi_*(\mu)$ is in the Patterson-Sullivan measure class on $\partial\mathbb{H}^2$.) Given $g \in \Gamma$, we let $d(x, gx)$ denote the displacement of x in $\mathbb{H}^2$. There exists $\lambda_\mu > 0$ such that, for μ-a.e. $\underline{i} = (i_n)_{n=0}^\infty \in \Sigma$,

$$\lambda_\mu = \lim_{n\to+\infty} \frac{1}{n} d(x, g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x). \tag{1.1}$$

The following results are examples of the type of natural statistical results that we can obtain. They can be viewed as analogues of the more familiar results for independent identically distributed random variables. We begin with the Central Limit Theorem.

Theorem 1.1 (Central Limit Theorem)**.** *There exists $\sigma > 0$ such that for any $x \in X$ sequences and $y \in \mathbb{R}$ we have that*

$$\lim_{n\to+\infty} \mu\left\{\underline{i} \in \Sigma : \frac{1}{\sqrt{n}}\left(d(x, g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x) - n\lambda_\mu\right) \le y\right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt.$$

A closely related viewpoint is that of the boundary action of Γ on $\partial\mathbb{H}$. The sequence $g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x$ naturally converges to a limit point $\zeta = \zeta_{\underline{i}} \in \partial\mathbb{H}^2$, say, and we denote the resulting map, which is one-to-one μ-a.e., by $\pi : \Sigma \to \partial\mathbb{H}$. The following is an easy consequence of the above theorem.

Corollary 1.2 (Central Limit Theorem on the Boundary)**.** *There exists $\sigma > 0$ such that*

$$\lim_{n\to+\infty} \mu\left\{\underline{i} \in \Sigma : \frac{1}{\sqrt{n}}\left(\log|(g_{i_{n-1}} g_{i_{n-2}} \cdots g_{i_0})'(\zeta_{\underline{i}})| - n\lambda_\mu\right) \le y\right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt.$$

In this context we will also prove a (stronger) Local Central Limit Theorem. More precisely, we show the following.

Theorem 1.3 (Local Central Limit Theorem)**.** *There exists $\sigma > 0$ such that for any sequence $\epsilon_n > 0$, such that ϵ_n^{-1} grows subexponentially then*

$$\left|\frac{\sqrt{n}}{2\epsilon_n} \mu\left\{\underline{i} \in \Sigma : \log|(g_{i_{n-1}} g_{i_{n-2}} \cdots g_{i_0})'(\zeta_{\underline{i}})| - n\lambda_\mu \in (\xi - \epsilon, \xi + \epsilon_n)\right\} - \frac{1}{\sqrt{2\pi}\sigma} e^{-\xi^2/2\sigma^2 n}\right| \to 0,$$

as $n \to +\infty$, uniformly for $\xi \in \mathbb{R}$.

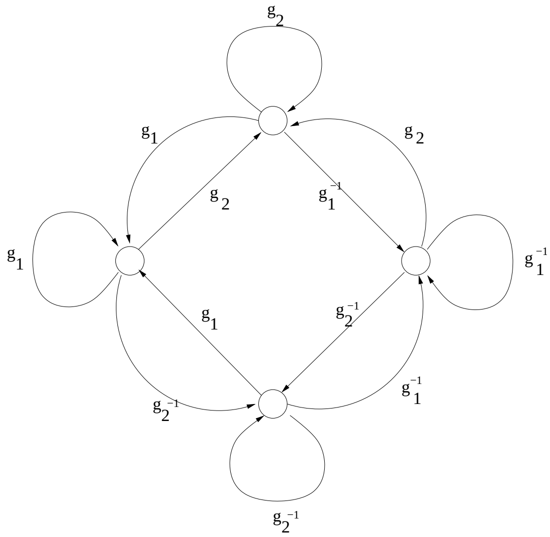
In another direction, we can consider a different type of distribution theorem. In a natural sense, the complement to Central Limit Theorems are Large Deviation results.

Theorem 1.4 (Large Deviations)**.** *Let $\epsilon > 0$. We have that*

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mu\left\{\underline{i} \in \Sigma : \left|\frac{1}{n} d(x, g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x) - \lambda_\mu\right| > \epsilon\right\} < 0.$$

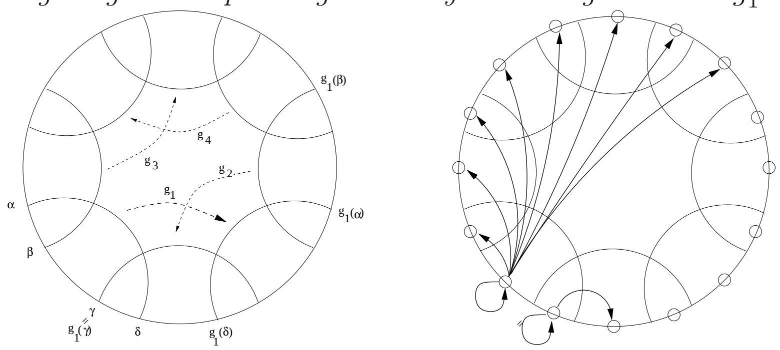
Let us consider some simple examples which illustrate groups to which these results apply.

Example 1.5 (Graph $\mathcal{G}$ for a free group)**.** *As a simple example, we could consider a free group $\Gamma = \langle a, b\rangle$ acting discretely and convex co-compactly by isometries on $\mathbb{H}^2$. In this case, $\mathcal{G}$ is a graph with 4 vertices whose edges are labelled by 4 generators and the space Σ codes the limit set in the boundary of $\mathbb{H}^2$.*



In this case, the measure of maximal entropy for Σ is the Markov measure associated to the matrix $P = \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix}$.

Example 1.6 (Graph $\mathcal{G}$ for genus 2 surface)**.** *We next consider a genus 2 co-compact surface group Γ acting discretely by isometries on $\mathbb{H}^2$. We can consider a graph with 16 vertices and 8 different edge weightings corresponding to the symmetric generators $g_1^{\pm 1}, g_2^{\pm 1}, g_3^{\pm 1}, g_4^{\pm 1}$.*



The figure on the left shows a boundary partition for the Poincaré disk and the action by the generators. The figure on the right shows part of the associated graph (the rest being clear by symmetry).

A natural Gibbs measure on μ is that which for which $\pi_*(\mu)$ is absolutely continuous (i.e., the same measure class as the Patterson-Sullivan measure).

These results complement the more classical approach to statistical properties of lattices acting on hyperbolic space through the use of convolutions of finitely supported measures cf. [4], [18], [31]. One advantage of our viewpoint is that it allows us to consider broader classes of measures and prove relatively deep statistical properties.

In section 2, we recall basic material about invariance principles and their consequences, including the statement of the Almost Sure Invariance Principle (Theorem 2.1). In section 3, we explain a connection with random matrix products. In section 4 we formulate the basic symbolic framework and present the proof of Theorem 2.1. In section 5, we prove the Local Central Limit Theorem for the boundary action (Theorem 1.3). Finally, in section 6 we present a proof of Theorem 1.4.

2 Invariance Principles

The Central Limit Theorem, and some related results, follow naturally from a more general invariance principle. Standard references for background material include [24] and [3].

We shall establish the following general result, which essentially says that, for appropriate $\lambda_\mu > 0$, sequences $d(g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x, x) - n\lambda_\mu$ are well approximated by a Brownian motion.

Theorem 2.1 (Almost Sure Invariance Principle)**.** *Let μ be the Gibbs measure for a Hölder continuous function on Σ. Then there exists $\lambda_\mu > 0$ such that, for any $x \in \mathbb{H}^2$, sequences*

$$\left(d(x, g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x) - n\lambda_\mu\right)_{n=1}^{\infty}. \tag{2.1}$$

associated to $\underline{i} \in \Sigma$ satisfy an Almost Sure Invariance Principle with respect to μ. More precisely, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a one-dimensional Brownian motion $W : \Omega \to C(\mathbb{R}^+, \mathbb{R})$, such that the random variable $\omega \mapsto W(\omega)(t)$ has mean zero and variance $\sigma^2 t > 0$, and sequences of random variables $\phi_n : \Sigma \to \mathbb{R}$ and $\psi_n : \Omega \to \mathbb{R}$ with the following properties:

1. *for some $\epsilon > 0$ we have $d(x, g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x) - n\lambda_\mu = \phi_n(\underline{i}) + O(n^{\frac{1}{2}-\epsilon})$ μ-a.e.;*
2. *the sequences $(\phi_n)_{n=1}^{\infty}$ and $(\psi_n)_{n=1}^{\infty}$ are equal in distribution;*
3. *for some $\epsilon > 0$, we have $\psi_n(\cdot) = W(\cdot)(n) + O(n^{\frac{1}{2}-\epsilon})$ $\mathbb{P}$-a.e..*

We now describe some consequences of the Almost Sure Invariance Principle (ASIP). Let $C([0,1], \mathbb{R})$ be the space of continuous functions on the interval $[0,1]$. Recall that the Brownian motion W induces the standard Wiener measure $\mathcal{W}$ on $C([0,1], \mathbb{R})$, defined by

$$\mathcal{W}\left(\{f(t) \,:\, f(t_i) - f(t_{i-1}) \le \alpha_i,\ i = 1, \cdots, k\}\right) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \int_{-\infty}^{\alpha_i} e^{-u^2(t_i - t_{i-1})/2} du,$$

where $0 = t_0 < t_1 < \cdots < t_k = 1$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ (cf. [3], p.68).

A particularly useful consequence of the ASIP is the Functional Central Limit Theorem (FCLT). This is also often called the Weak Invariance Principle, since it deals with continuous functions and weak star convergence.

Proposition 2.2 (Functional Central Limit Theorem)**.** *Let $\zeta_n : \Sigma \to C([0,1], \mathbb{R})$ be the map which associates to $\underline{i} \in \Sigma$ the piecewise linear function $\zeta_n(\underline{i})(\cdot)$, $n \ge 1$, on $[0,1]$ which interpolates the values*

$$\left(\frac{k}{n}, \frac{1}{\sigma\sqrt{n}}\left(d(x, g_{i_0} g_{i_1} \cdots g_{i_{k-1}} x) - k\lambda_\mu\right)\right), \ \textit{for } k = 1, \cdots, n.$$

Then ζ_n converges in distibution to W, as $n \to +\infty$.

Proof. The derivation of the FCLT from the ASIP is routine. We know from Theorem 2.1 that, for $0 \le t \le 1$,

$$\frac{1}{\sqrt{n}}\left((d(x, g_{i_0} g_{i_1} \cdots g_{i_{[nt]-1}} x) - [nt]\lambda_\mu) - \phi_{[nt]}(\underline{i})\right) = O\left(n^{-\frac{1}{2}+\epsilon}\right),$$

for μ-a.e. $\underline{i} \in \Sigma$, and

$$\frac{1}{\sqrt{n}}\left(\psi_{[nt]}(\omega) - W(\omega)(nt)\right) = O\left(n^{-\frac{1}{2}+\epsilon}\right),$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$. Since ϕ_n and ψ_n have the same distribution and the rescaling $\frac{1}{\sqrt{n}} W(\omega)(nt)$ is a Brownian motion with the same distribution as W, the result follows. □

A number of standard, and perhaps more familiar, results all follow from the FCLT. These include the Central Limit Theorem stated in the introduction (Theorem 1.1) and a number of other results recalled below as corollaries. A key ingredient in deriving these is the following classical result (cf. [3]).

Lemma 2.3 (Continuous Mapping Principle)**.** *If ξ_n is a sequence of random variables, taking values in $[0,1]$, which converges to ξ in distribution and $h : C([0,1],\mathbb{R}) \to \mathbb{R}$ is measurable and continuous (except possibly on a set of Wiener measure zero) then the sequence $h(\xi_n)$ converges to $h(\xi)$ in distribution.*

We begin with the proof of the Central Limit Theorem (Theorem 1.1).

Proof of Theorem 1.1. This follows from the choice of $h : C([0,1],\mathbb{R}) \to \mathbb{R}$ defined by $h(f) = f(1)$, for $f \in C([0,1],\mathbb{R})$. We can then write

$$\begin{aligned}\lim_{n\to+\infty} \mu\left\{\underline{i} \in \Sigma : \frac{1}{\sigma\sqrt{n}}\left(d(x, g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x) - n\lambda_\mu\right) \le y\right\} &= \mathcal{W}(\{f \in C([0,1],\mathbb{R}) : f(1) \le y\}) \\ &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{y} e^{-t^2/2\sigma^2}\,dt,\end{aligned}$$

as required. □

There are many natural corollaries that arise as direct consequences of the FCLT and the Lemma 2.3 (with suitable choices of functions). As a example, we consider the following.

Corollary 2.4. *For $y \ge 0$ we have that*

$$\begin{aligned}&\lim_{n\to+\infty} \mu\left\{\underline{i} \in \Sigma : \frac{1}{\sqrt{n}} \max_{1\le k\le n} d(x, g_{i_0} g_{i_1} \cdots g_{i_{k-1}} x) - n\lambda_\mu \le y\right\} \\ &= \frac{\sqrt{2}}{\sqrt{\pi}\sigma}\int_{-\infty}^{y} e^{-t^2/2\sigma^2}\,dt - 1.\end{aligned}$$

Proof. This follows from the choice $h : C([0,1],\mathbb{R}) \to \mathbb{R}$ defined by $h(f) = \sup_{0\le t\le 1} f(t)$. We can then write

$$\begin{aligned}
&\lim_{n\to+\infty} \mu(\{\underline{i}\in\Sigma : n^{-1/2}\max_{1\le k\le n} d(x, g_{i_0}g_{i_1}\cdots g_{i_{k-1}}x) - n\lambda_\mu \le y\}) \\
&= \mathbb{P}(\{\omega\in\Omega : \sup_{0\le t\le 1} W(\omega)(t)\le y\}) \\
&= 2\mathbb{P}(\{\omega\in\Omega : W(\omega)(1)\le y\}) - 1,
\end{aligned}$$

using a standard property of Brownian motion, which gives the required formula. □

Remark 2.5. By standard methods one can also prove the analogues of the Law of the Iterated Logarithm, and its functional version, [3], [24] and the Arcsine Law.

Remark 2.6. Similar invariance principles hold for periods of harmonic 1-forms. More precisely, suppose that Γ is co-compact and let η be a harmonic 1-form on $\mathbb{H}^2/\Gamma$ with lift $\widetilde{\eta}$ to $\mathbb{H}^2$. Let $\gamma_n(\underline{i})$ denote the geodesic arc joining x to $g_{i_0}g_{i_1}\cdots g_{i_{n-1}}x$. For each Gibbs measure μ on Σ, there exists $\kappa_\mu\in\mathbb{R}$ such that the sequence

$$\int_{\gamma_n(\underline{i})} \widetilde{\eta} - n\kappa_\mu$$

satisfies an ASIP. More precisely, there exists $\sigma > 0$ and a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and a one-dimensional Brownian motion $W : \Omega \to C(\mathbb{R}^+,\mathbb{R})$, such that the random variable $\omega \mapsto W(\omega)(t)$ has mean zero and variance $\sigma^2 t > 0$, and sequences of random variables $\phi_n : \Sigma \to \mathbb{R}$ and $\psi_n : \Omega\to\mathbb{R}$ with the following properties:

1. for some $\epsilon > 0$ we have $\int_{\gamma_n(\underline{i})} \widetilde{\eta} - n\kappa_\mu = \phi_n(\underline{i}) + O(n^{\frac{1}{2}-\epsilon})$ μ-a.e.;
2. the sequences $(\phi_n)_{n=1}^\infty$ and $(\psi_n)_{n=1}^\infty$ are equal in distribution;
3. for some $\epsilon > 0$, we have $\psi_n(\cdot) = W(\cdot)(n) + O(n^{\frac{1}{2}-\epsilon})$ $\mathbb{P}$-a.e..

3 Random Matrix Products

The results described in the preceding sections are clearly reminiscent of classical results on random products of matrices. Let $A_1,\cdots,A_k \in SL(2,\mathbb{R})$ be a finite set of matrices. Classically, these would be chosen randomly with respect to a Bernoulli probability $p = (p_1,\cdots,p_k)$. In 1960, Furstenberg and Kesten [12] showed that there exists a Lyapunov exponent λ such that for almost all $\underline{i} = (i_n)_{n=0}^\infty \in \{1,\cdots,k\}^{\mathbb{Z}^+}$ we have that

$$\lambda = \lim_{n\to+\infty} \frac{1}{n}\log \|A_{i_0}A_{i_1}\cdots A_{i_{n-1}}\|, \tag{3.1}$$

where $\|A\| = \sqrt{\mathrm{tr}(A^*A)}$. (A more modern approach would be to use the Kingman subadditive ergodic theorem.) The importance of having Bernoulli measures lies in the use of convolutions of measures [14].

The limit in (3.1) can be viewed as a non-commutative version of the Birkhoff ergodic theorem. As in the case of Birkhoff averages, for hyperbolic systems, say, one can ask for stronger results such as the Central Limit Theorem and more general Invariance Principles. Provided the matrices satisfy appropriate independence conditions, Le Page showed that for almost all $\underline{i}$ the sequences

$$\left(\log \|A_{i_0}A_{i_1}\cdots A_{i_{n-1}}\| - n\lambda\right)_{n=1}^{\infty}$$

satisfy a Central Limit Theorem (and other statistical results) [19].

In light of the above results for Bernoulli measures, it is interesting to consider a simple interpretation of our results in terms of matrix products. In our setting, we have more flexibility in the choice of measures. Let $\overline{\Gamma}$ be the subgroup of $SL(2,\mathbb{R})$ generated by $A_1,\ldots,A_k$. We impose the following two assumptions.

Assumption I. $-I \notin \overline{\Gamma}$ (so that $\overline{\Gamma} \cong \Gamma = \overline{\Gamma}/\{\pm I\}$).

Assumption II. Γ acts convex co-compactly on $\mathbb{H}^2$.

We recall the following simple result (cf. [2]).

Lemma 3.1. *Given $A \in SL(2,\mathbb{R})$ we have $2\cosh d(0, A0) = \|A\|^2$.*

In particular, writing $d = d(0, A0)$, we have $e^{2d} - e^d\|A\|^2 + 1 = 0$ and so

$$\begin{aligned} e^d = \frac{1}{2}\left(\|A\|^2 + \sqrt{\|A\|^4 - 4}\right) &= \frac{1}{2}\|A\|^2\left(1 + \sqrt{1 - 4\|A\|^{-4}}\right) \\ &= \|A\|^2\left(1 + O(\|A\|^{-2})\right) \end{aligned}$$

and thus

$$d(0, A0) = 2\log\|A\| + O(\|A\|^{-2}) = 2\log\|A\| + O(e^{-d}).$$

Since Γ is convex co-compact, $d(0, A0)$ is comparable to the word length of A (with respect to $A_1,\ldots,A_k$), so, for any $\underline{i} \in \Sigma$, one has

$$\log\|A_{i_0}A_{i_1}\cdots A_{i_n}\| = \frac{1}{2}d(0, A_{i_0}A_{i_1}\cdots A_{i_n}0) + O(e^{-n\epsilon}).$$

for some $\epsilon > 0$. Using this, we can recast Theorem 2.1 as:

Theorem 3.2 (Almost Sure Invariance Principle version 2)**.** *Assume that $A_1,\ldots,A_k$ satisfy Assumptions I and II. Let μ be the Gibbs measure associated to a Hölder continuous function on Σ. Then there exists $\lambda_\mu > 0$ such that sequences*

$$\left(\log \|A_{i_0}A_{i_1}\cdots A_{i_{n-1}}\| - n\lambda_\mu\right)_{n=1}^{\infty}. \tag{3.2}$$

associated to $\underline{i} \in \Sigma$ satisfy an ASIP. More precisely, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a one-dimensional Brownian motion $W : \Omega \to C(\mathbb{R}^+, \mathbb{R})$, such that the random variable $\omega \mapsto W(\omega)(t)$ has mean zero and variance $\sigma^2 t > 0$, and sequences of random variables $\phi_n : \Sigma \to \mathbb{R}$ and $\psi_n : \Omega \to \mathbb{R}$ with the following properties:

1. *for some $\epsilon > 0$ we have $\log\|A_{i_0}A_{i_1}\cdots A_{i_{n-1}}\| - n\lambda_\mu = \phi_n(\underline{i}) + O(n^{\frac{1}{2}-\epsilon})$ μ-a.e.;*
2. *the sequences $(\phi_n)_{n=1}^{\infty}$ and $(\psi_n)_{n=1}^{\infty}$ are equal in distribution;*
3. *for some $\epsilon > 0$, we have $\psi_n(\cdot) = W(\cdot)(n) + O(n^{\frac{1}{2}-\epsilon})$ $\mathbb{P}$-a.e..*

Remark 3.3. One can actually recover many of the classical results for random matrix products by considering free semi-groups, rather than groups.

4 Shifts, groups and geometry

In this section, we shall make a more precise connection between the groups Γ, their actions on $\mathbb{H}^2$, and the shift spaces Σ. This will provided a foundation for the proofs of the theorems announced in the preceding sections.

4.1 The shift maps

Let Γ be a discrete group of isometries of $\mathbb{H}^2$ and let $\Gamma_0 = \{a_1^{\pm 1}, \cdots, a_k^{\pm 1}\}$ be a (symmetric) set of generators. Denote the word length of $g \in \Gamma$ by

$$|g| = \min\{n : g = g_{i_1} \cdots g_{i_n} \text{ where } g_{i_1}, \cdots, g_{i_n} \in \Gamma_0\}.$$

Assume that either:

1. $\Gamma = \langle a_1, \ldots, a_k \rangle$ is a free group acting convex co-compactly on $\mathbb{H}^2$ and $\Gamma_0 = \{a_1^{\pm 1}, \ldots, a_k^{\pm 1}\}$ [13].
2. $\Gamma = \langle a_1, b_1, \cdots, a_{\mathfrak{g}}, b_{\mathfrak{g}} : \prod_{i=1}^{\mathfrak{g}} [a_i, b_i] = e \rangle$ is a surface group acting co-compactly on $\mathbb{H}^2$ and $\Gamma_0 = \{a_1^{\pm 1}, b_1^{\pm 1}, \cdots, a_{\mathfrak{g}}^{\pm 1}, b_{\mathfrak{g}}^{\pm 1}\}$.

Associated to Γ is its limit set $L_\Gamma \subset \partial \mathbb{H}^2$, defined as the set of accumlation points of Γx, for any $x \in \mathbb{H}^2$. In case 1, L_Γ is a Cantor set, while in case 2, $L_\Gamma = \partial \mathbb{H}^2$. Under either of the above assumptions, L_Γ is homeomorphic to the Gromov boundary $\partial \Gamma$ of Γ (viewed as an abstract group).

The main connection with symbolic dynamics comes from the following.

Lemma 4.1 (Cannon [8], Series [32], Adler-Flatto [1])**.** *Assume that we have a group Γ as above. There exists a directed graph $\mathcal{G}$ with edges labelled by Γ_0 such that elements g of Γ are in bijection with finite paths in $\mathcal{G}$ and the length of the path coincides with the word length $|g|$.*

We form a subshift of finite type by letting Σ denote the space of infinite paths $\underline{i} = (i_k)_{k=0}^\infty$ in $\mathcal{G}$ and let $\sigma : \Sigma \to \Sigma$ denote the shift map. We can associate to the graph $\mathcal{G}$ an incidence matrix A, where the entry $A(i,j) = 1$ if the ith edge leads to the jth edge (and 0 otherwise). In particular, we have the alternate formulation

$$\Sigma = \{\underline{i} = (i_k)_{k=0}^\infty : A(i_k, i_{k+1}) = 1, \forall k \geq 0\}.$$

We give Σ the metric $d_\Sigma(\underline{i}, \underline{j}) = 2^{-\mathfrak{n}(\underline{i}, \underline{j})}$, where $\underline{i}$ and $\underline{j}$ first differ in the $\mathfrak{n}(\underline{i}, \underline{j})$-th place and we use the convention $\mathfrak{n}(\underline{i}, \underline{i}) = -\infty$. Here, we can see by inspection that $\mathcal{G}$ is irreducible and aperiodic (i.e. there exists $N \geq 1$ such that each pair of vertices in $\mathcal{G}$ is joined by a path of length N), so that σ is topologically mixing.

Fix $x \in \mathbb{H}^2$. The formula

$$\pi(\underline{i}) = \lim_{n \to +\infty} g_{i_0} g_{i_1} \cdots g_{i_{n-1}} x$$

gives a well defined map $\pi : \Sigma \to L_\Gamma$ (which is independent of the choice of x).

4.2 Gibbs measures

For Hölder continuous functions $g : \Sigma \to \mathbb{R}$ we can associate a (σ-invariant) Gibbs measure $\mu = \mu_g$. This is the unique σ-invariant probability measure μ_g for which

$$h(\mu_g) + \int g d\mu_g \geq h(\nu) + \int g d\nu$$

for all σ-invariant probability measures ν. In particular, π is one-to-one μ-a.e.. (In fact, when L_Γ is a Cantor set then π is a homeomorphism.)

The next two examples provide particularly natural choices of Gibbs state on Σ.

Example 4.2. *The measure of maximal entropy μ_0 on Σ is an equilibrium measure for the function $\psi = 0$. If A denotes the incidence matrix of $\mathcal{G}$ then μ_0 is the Markov measure associated to the matrix P given by $P(i,j) = A(i,j)v_j/\rho v_i$, where ρ and v are the maximal eigenvalue and associated eigenvector guaranteed by the Perron-Frobenius Theorem. Clearly, μ_0 only depends on Γ as an abstract group.*

Let m_n be the probability measure on $\mathbb{H}$ given by equidistributing mass on the finite set of points $gx \in \mathbb{H}^2$ where $|g| = n$, i.e.,

$$m_n = \frac{1}{\#\{g \in \Gamma : |g| = n\}} \sum_{|g|=n} \delta_{gx}.$$

As $n \to +\infty$, we have that m_n converges in the weak star topology on the compactification $\mathbb{H}^2 \cup \partial\mathbb{H}^2$ to $\pi_*(\mu_0)$ supported on $\partial\mathbb{H}^2$.

Example 4.3. *There is a unique Gibbs state μ on Σ for which $\pi_*(\mu)$ is in the Patterson-Sullivan measure class on L_Γ. This class is defined in terms of its transformation properties under the action of Γ but in this setting is conveniently characterized as the δ-dimensional Hausdorff measure class, where $\delta > 0$ is the Hausdorff dimension of L_Γ. (In the co-compact case, this is just the Lebesgue measure class.).*

For $t > \delta$, let ν_t be the probability measure on $\mathbb{H}^2$ supported on the orbit $\{gx \in \mathbb{H}^2 : g \in \Gamma\}$ and defined by

$$\nu_t = \sum_{g\in\Gamma} e^{-td(x,gx)}\delta_{gx} / \sum_{g\in\Gamma} e^{-td(x,gx)}.$$

As $t \searrow \delta$ we have that ν_t converges in the weak star topology on the compactification $\mathbb{H}^2 \cup \partial\mathbb{H}^2$ to $\pi_*(\mu)$ supported on $\partial\mathbb{H}^2$.

In order to describe the displacements $d(x, gx)$, $g \in \Gamma$, we need to augment the directed graph $\mathcal{G}$ by adding an extra vertex "0" and a directed edge from each vertex to 0 (including a loop from 0 to itself 0). We call the resulting directed graph $\mathcal{G}^*$ and the associated subshift of finite type $\sigma : \Sigma^* \to \Sigma^*$. We also let $\Sigma_0 \subset \Sigma^*$ denote those paths which end by visiting 0 infinitely often. In this way, we associate an infinite path (in Σ_0) to each finite path in $\mathcal{G}$. Our new shift $\sigma : \Sigma^* \to \Sigma^*$ is not mixing or even transitive but this does not cause substantial problems [17], [26].

Lemma 4.4. *[17], [26] There exists a Hölder continuous function $r : \Sigma_0 \to \mathbb{R}$ such that*

$$d(x, g_{i_0} \cdots g_{i_{n-1}} x) = \sum_{k=0}^{n-1} r(\sigma^k \underline{i}^{(n)})$$

where $\underline{i}^{(n)} = (i_0, \cdots, i_{n-1}, 0, 0, \ldots)$ is the associated infinite path. Furthermore, r extends to a Hölder continuous function $r : \Sigma^* \to \mathbb{R}$.

In particular, we may now characterize μ in Example 4.3 as the Gibbs state associated to $-\delta r$ [17].

Corollary 4.5. *Let μ be a Gibbs measure on Σ. Then there exists $\lambda_\mu > 0$ such that*

$$\lambda_\mu = \lim_{n \to +\infty} \frac{1}{n} d(x, g_{i_0} ... g_{i_{n-1}} x),$$

for μ-a.e. $\underline{i} \in \Sigma$.

Proof. Let $\lambda_\mu = \int r d\mu$. By the Birkhoff Ergodic Theorem,

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} r(\sigma^k \underline{i}) = \int r d\mu,$$

for μ-a.e. $\underline{i} \in \Sigma$. Since r is Hölder on Σ^* (with exponent $\alpha > 0$, say), we have

$$\left| \sum_{k=0}^{n-1} r(\sigma^k \underline{i}) - \sum_{k=0}^{n-1} r(\sigma^k \underline{i}^{(n)}) \right| \leq C \sum_{k=0}^{n-1} \frac{1}{2^{k\alpha}} = O(1).$$

Combining these two observations with Lemma 4.3 gives the required convergence.

To see that $\lambda_\mu > 0$, recall that there exists $c > 0$ such that $d(x, gx) \geq c|g|$, for all $g \in \Gamma$. Thus, again using Lemma 4.3,

$$\frac{1}{n} \sum_{k=0}^{n-1} r(\sigma^k \underline{i}^{(n)}) \geq c > 0,$$

for all $n \geq 1$ and $\underline{i} \in \Sigma$. The result follows. □

Remark 4.6. For $\mu = \mu_0$, the drift λ_{μ_0} can be viewed as a constant which quantifies the average "geometric length to word length" ratio [26].

The following observation is useful.

Lemma 4.7. *For any choice of $\lambda \in \mathbb{R}$, the function $r - \lambda : \Sigma \to \mathbb{R}$ is not a coboundary, i.e., we cannot find a continuous function $u : \Sigma \to \mathbb{R}$ such that $r = u\sigma - u + \lambda$. In particular, $\sigma^2 > 0$.*

Proof. By Livsic's Theorem, the statement that $r - \lambda$ is a coboundary is equivalent to the identities

$$\sum_{k=0}^{n-1} r(\sigma^k \underline{i}) = n\lambda, \quad \text{whenever } \sigma^n \underline{i} = \underline{i}.$$

However, these sums are exactly the lengths of closed geodesics on $\mathbb{H}^2/\Gamma$ (see [25] for the convex co-compact case, [27] for the co-compact case). In the co-compact case, the fact that these lengths cannot lie in a discrete subgroup of $\mathbb{R}$ is equivalent to the well-known mixing of the geodesic flow [6]. In the convex co-compact case, the corresponding result was obtained by Dal'bo [9]. ☐

4.3 Transfer Operators and Pressure

In the proofs which follow, we shall need to use properties of a class of operators called transfer operators and a convex functional called topological pressure. We outline below the material we need; more details may be found in [23]. Let $g : \Sigma \to \mathbb{C}$ be a Hölder continuous function with Hölder exponent $\alpha > 0$. We define an associated transfer operator $L_g : C^\alpha(\Sigma) \to C^\alpha(\Sigma)$ by

$$L_g h(\underline{i}) = \sum_{\sigma(\underline{j})=\underline{i}} e^{g(\underline{j})} h(\underline{j}).$$

Now suppose that g is real-valued. For a (not necessarily Hölder) continuous function $g : \Sigma \to \mathbb{R}$, we define the topological pressure $P(g)$ by

$$P(g) = \sup \left\{ h(\nu) + \int g \, d\nu : \nu \text{ is a } \sigma\text{-invariant probability measure} \right\}.$$

If g is Hölder continuous then the associated Gibbs measure μ_g is uniquely defined by

$$P(g) = h(\mu_g) + \int g \, d\mu_g.$$

(This is consistent with the characterization of $\mu = \mu_g$ in section 4.2.)

There is a close relationship between the objects defined above. Suppose again that $g \in C^\alpha(\Sigma)$ is real-valued. Then $L_g : C^\alpha(\Sigma) \to C^\alpha(\Sigma)$ has $e^{P(g)}$ as a simple eigenvalue and the rest of the spectrum is contained in a disc of strictly smaller radius. We say that g is normalized if $L_g 1 = 1$, in which case 1 is the maximal eigenvalue, $P(g) = 0$, and $L_g^* \mu_g = \mu_g$.

Lemma 4.8. *For a normalized Hölder continuous function $g : \Sigma \to \mathbb{R}$, the eigenprojection associated to the eigenvalue 1 is $h \mapsto \int h d\mu_g$, so if $\int h d\mu_g = 0$ then*

$$\|L_g^n h\|_\infty \le C\theta^n,$$

for some $C \ge 0$ and $0 < \theta < 1$.

Finally, we discuss the pressure function $t \mapsto P(g + th)$, where h is a real-valued Hölder function and $t \in \mathbb{R}$.

Lemma 4.9. *The function $t \mapsto P(g+th)$ is real analytic and is strictly convex unless $h-c$ is a coboundary for some $c \in \mathbb{R}$. Furthermore, for $\xi \in \mathbb{R}$, we have*

$$\frac{d}{dt}P(g+th)\bigg|_{t=\xi} = \int h \, d\mu_{g+\xi h}.$$

4.4 Proof of Theorem 2.1

The ASIP which we have stated as Theorem 2.1 will follow from the next result.

Theorem 4.10. *Let $\sigma : \Sigma \to \Sigma$ be a mixing subshift of finite type and let μ be a Gibbs measure for a Hölder continuous function and let $f : \Sigma \to \mathbb{R}$ be a Hölder continuous function such that $\int f d\mu = 0$. Then, provided f is not a coboundary, the sequence*

$$\sum_{k=0}^{n-1} f(\sigma^k x), \quad n \geq 1,$$

satisfies an Almost Sure Invariance Principle with error term $O(n^{\frac{1}{2}-\epsilon})$.

This was obtained in [10] with an error term $O(n^{\frac{1}{2}-\epsilon})$ (which is sufficient for the CLT) but, using arguments from [24], the stronger error term is now standard (cf. [33]).

To deduce Theorem 2.1 from this, we first observe that, by Lemma 4.7, $r - \lambda : \Sigma \to \mathbb{R}$ is not a coboundary. Thus, by Theorem 4.10, the sequence of sums $\sum_{k=0}^{n-1} r(\sigma^k \underline{i}) - n\lambda$ satsisfies an ASIP. However, as in the proof of Corollary 4.5,

$$d(x, g_{i_0} \cdots g_{i_{n-1}} x) - n\lambda = \sum_{k=0}^{n-1} r(\sigma^k \underline{i}^{(n)}) - n\lambda = \left(\sum_{k=0}^{n-1} r(\sigma^k \underline{i}) - n\lambda \right) + O(1),$$

with the last error being smaller than $O(n^{\frac{1}{2}-\epsilon})$.

5 The Local Central Limit Theorem

In this section we turn to the proof of the stronger Local Central Limit Theorem for the action on the boundary, presented in the introduction as Theorem 1.3. We begin by a more detailed formulation of the result.

Given $\underline{i} \in \Sigma$ we can associate the corresponding limit point $\zeta_{\underline{i}} = \pi(\underline{i}) \in \partial \mathbb{H}^2$. In the case of $SL(2,\mathbb{R})$ there is natural conformal action on the boundary. If $g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2,\mathbb{R})$ then $g\zeta = \frac{\alpha\zeta+\beta}{\gamma\zeta+\delta}$. The action on the boundary has derivative $g'(\zeta) = (\gamma\zeta+\delta)^{-2}$.

Definition 5.1. *We say that, for a sequence $\epsilon_n > 0$, the reciprocals ϵ_n^{-1} grow subexponentially if $\lim_{n\to+\infty} -\frac{1}{n}\log \epsilon_n = 0$.*

Theorem 5.2 (Local Central Limit Theorem)**.** *There exists $\sigma > 0$ such that for any sequence $\epsilon_n > 0$, such that ϵ_n^{-1} grows subexponentially then,*

$$\left| \frac{\sqrt{n}}{2\epsilon_n} \mu \left\{ \underline{i} \in \Sigma \,:\, \log |(g_{i_{n-1}} g_{i_{n-2}} \cdots g_{i_0})'(\zeta_{\underline{i}})| - n\lambda_\mu \in (\xi - \epsilon_n, \xi + \epsilon_n) \right\} - \frac{1}{\sqrt{2\pi}\sigma} e^{-\xi^2/2\sigma^2 n} \right| \to 0,$$

as $n \to +\infty$, uniformly for $\xi \in \mathbb{R}$.

The proof of the local central limit theorem is similar the standard version for shifts of finite type (cf. [7], [30]); however, the presence of shrinking intervals require more careful estimates. Define $f : \Sigma \to \mathbb{R}$ by

$$f(\underline{i}) = \log |(g_{i_0}^{-1})'(\zeta_{\underline{i}})| - \lambda_\mu = \log |(g_{i_0}^{-1})'(\pi(\underline{i}))| - \lambda_\mu.$$

We can then write

$$\sum_{k=0}^{n-1} f(\sigma^k \underline{i}) = \sum_{k=0}^{n-1} \log |(g_{i_k}^{-1})'(\pi(\sigma^k \underline{i}))| - n\lambda_\mu = \log |(g_{i_{n-1}}^{-1} g_{i_{n-2}}^{-1} \cdots g_{i_0}^{-1})'(\pi(\underline{i}))| - n\lambda_\mu.$$

We shall need to use properties of transfer operators defined in Section 4. Recall that the measure μ is the Gibbs state for some Hölder continuous function $g : \Sigma \to \mathbb{R}$. Choose $\alpha > 0$ such that $g, f \in C^\alpha(\Sigma)$. We shall study the family $L_{g+itf} : C^\alpha(\Sigma) \to C^\alpha(\Sigma)$, $t \in \mathbb{R}$. We recall that g is normalized so that $L_g 1 = 1$ and $L_g^* \mu = \mu$. We need the following bound on L_{g+itf}^n.

Lemma 5.3. *There exist $C > 0$, $0 < \theta < 1$ and $\gamma > 0$ such that, for $|t| \geq 1$, we have*

$$\|L_{g+itf}^n\| \leq C \min\{\theta^n |t|^\gamma, 1\}.$$

Remark 5.4. This type of bound was first obtained by Dolgopyat [11] in his work on the rate of decay of correlations for geodesic flows. For co-compact groups, the above result (where the operator is acting on Hölder continuous functions) was proved in §6 of [27]. For convex co-compact groups, the analogous bound was established by Naud in [21]; however, the operators there were defined with respect to the boundary map and acted on C^1 functions. We could carry out our analysis in that context but, for simplicity of exposition, we continue to work with the shift space Σ.

Let $\chi : \mathbb{R} \to \mathbb{R}$ be a compactly supported function which is a C^k approximation to the indicator function of $[-1, 1]$. (The value of k will be chosen later.) Write $\chi_n^{(\xi)}(x) = \chi(\epsilon_n^{-1}(x - \xi))$. Then $\widehat{\chi}_n^{(\xi)}(t) = e^{i\xi t} \epsilon_n \widehat{\chi}(\epsilon_n t)$. Writing $f^n(\underline{i}) = \sum_{k=0}^{n-1} f(\sigma^k \underline{i})$ and

$$\rho(n, \xi) := \int \chi_n^{(\xi)}(f^n(\underline{i})) d\mu,$$

we define

$$A(n, \xi) := \left| \frac{\sigma\sqrt{n}}{\epsilon_n} \rho(n, \xi) - \frac{\int \chi(y) dy}{\sqrt{2\pi}\sigma} e^{-\xi^2/2\sigma^2 n} \right|.$$

We shall show the following.

Lemma 5.5. *Provided ϵ_n^{-1} grows at a subexponential rate, we have*

$$\lim_{n\to+\infty} \sup_{\xi\in\mathbb{R}} A(n,\xi) = 0.$$

We shall begin by obtaining a more useful formula for $A(n,\xi)$. Using Fourier inversion and Fubini's Theorem, we have

$$\begin{aligned}\frac{\sigma\sqrt{n}}{\epsilon_n}\rho(n,\xi) &= \frac{1}{2\pi}\frac{\sigma\sqrt{n}}{\epsilon_n}\int_{-\infty}^{\infty}\left(\int e^{itf^n(\underline{i})}d\mu(\underline{i})\right)\widehat{\chi}_n^{(\xi)}(t)dt \\ &= \frac{1}{2\pi}\sigma\sqrt{n}\int_{-\infty}^{\infty}\left(\int e^{itf^n(\underline{i})}d\mu(\underline{i})\right)e^{i\xi t}\widehat{\chi}(\epsilon_n t)dt.\end{aligned}$$

If we make the substitution $u = t\sigma\sqrt{n}$, then this becomes

$$\frac{\sqrt{n}}{2\epsilon_n}\rho(n,\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int e^{iuf^n(\underline{i})/\sigma\sqrt{n}}d\mu(\underline{i})\right)e^{i\xi u/\sigma\sqrt{n}}\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right)du.$$

Combining this with the standard identity

$$e^{-\xi^2/2\sigma^2 n} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{iu\xi/\sigma\sqrt{n}}e^{-u^2/2}du,$$

we obtain

$$2\pi A(n,\xi) = \left|\int_{-\infty}^{\infty} e^{it\xi/\sigma\sqrt{n}}\left\{\left(\int e^{iuf^n(\underline{i})/\sigma\sqrt{n}}d\mu(\underline{i})\right)\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right) - \left(\int \chi(y)dy\right)e^{-u^2/2}\right\}du\right|.$$

In particular, we have the bound

$$2\pi A(n,\xi) \le A_1(n,\xi) + A_2(n,\xi) + A_3(n,\xi),$$

where, given $\delta > 0$,

$$\begin{aligned}A_1(n,\xi) &= \left|\int_{-\delta\sigma\sqrt{n}}^{\delta\sigma\sqrt{n}} e^{iu\xi/\sigma\sqrt{n}}\left\{\left(\int e^{iuf^n(\underline{i})/\sigma\sqrt{n}}d\mu(\underline{i})\right)\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right) - \left(\int \chi(y)dy\right)e^{-u^2/2}\right\}du\right|, \\ A_2(n,\xi) &= \left|\int_{|u|\ge\delta\sigma\sqrt{n}} e^{iu\xi/\sigma\sqrt{n}}\left\{\left(\int e^{iuf^n(\underline{i})/\sigma\sqrt{n}}d\mu(\underline{i})\right)\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right)\right\}du\right|, \\ A_3(n,\xi) &= \left|\int_{|u|\ge\delta\sigma\sqrt{n}} e^{iu\xi/\sigma\sqrt{n}}\left(\int \chi(y)dy\right)e^{-u^2/2}du\right|.\end{aligned}$$

We shall proceed by estimating A_1, A_2 and A_3. First note that an elementary argument gives the following.

Lemma 5.6. *We have*

$$\lim_{n\to+\infty} \sup_{\xi\in\mathbb{R}} A_3(n,\xi) = 0.$$

Lemma 5.7. *We have*

$$\lim_{n\to+\infty}\sup_{\xi\in\mathbb{R}} A_1(n,\xi) = 0.$$

Proof. This proof is based on [30] (see also [7]). We have

$$A_1(n,\xi) = \left|\int_{-\delta\sigma\sqrt{n}}^{\delta\sigma\sqrt{n}} e^{iu\xi/\sigma\sqrt{n}}\left\{\left(\int L^n_{g+iuf/\sigma\sqrt{n}} 1 d\mu(\underline{i})\right)\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right) - \left(\int \chi d\mu\right)e^{-u^2/2}\right\}du\right|.$$

On the domain of integration, as $n\to+\infty$, we have

(i) $\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right)$ converges to $\widehat{\chi}(0) = \int\chi(y)dy$;

(ii) $\int L^n_{g+iuf/\sigma\sqrt{n}} 1 d\mu(\underline{i})$ converges to $e^{-u^2/2}$.

Furthermore, we have the bounds

$$\left|\int L^n_{g+iuf/\sigma\sqrt{n}} 1 d\mu(\underline{i})\right| \le e^{-u^2/4} \quad \text{and} \quad \left|\int L^n_{g+iuf/\sigma\sqrt{n}} 1 d\mu(\underline{i}) - e^{-u^2/2}\right| \le 2e^{-u^2/4}.$$

The result now follows from the Dominated Convergence Theorem. □

In order to estimate $A_2(n,\xi)$, we need to use Lemma 5.3. We shall also use the following result, which is each to prove using integration by parts.

Lemma 5.8. *If* $\chi:\mathbb{R}\to\mathbb{R}$ *is compactly supported and* C^k *then we have that* $\widehat{\chi}(u) = O(|u|^{-k})$, *as* $|u|\to+\infty$.

Lemma 5.9. *We have*

$$\lim_{n\to+\infty}\sup_{\xi\in\mathbb{R}} A_2(n,\xi) = 0.$$

Proof. Using Lemma 5.3, we have the bound

$$A_2(n,\xi) \le C\theta^n \int_{|u|\ge\delta\sigma\sqrt{n}} \left|\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right)\right| \left(\frac{|u|}{\sigma\sqrt{n}}\right)^\gamma du.$$

Splitting $[\delta\sigma\sqrt{n},\infty)$ into $[\delta\sigma\sqrt{n}, e^{\beta n})$ and $[e^{\beta n},\infty)$, for some small $\beta>0$, this gives us

$$A_2(n,\xi) \le \frac{C\theta^n\|\widehat{\chi}\|_\infty}{(\sigma\sqrt{n})^\gamma}\int_{\delta\sigma\sqrt{n}}^{e^{\beta n}} |u|^\gamma du + \int_{e^{\beta n}}^\infty \left|\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right)\right| \left(\frac{|u|}{\sigma\sqrt{n}}\right)^\gamma du.$$

The first term on the Right Hand Side is $O(\theta^n e^{\beta n(\gamma+1)})$, which, provided we choose $\beta>0$ sufficiently small, tends to zero exponentially fast. To estimate the second term, we use Lemma 5.8 to obtain

$$\begin{aligned}\int_{e^{\beta n}}^\infty \left|\widehat{\chi}\left(\frac{\epsilon_n u}{\sigma\sqrt{n}}\right)\right| \left(\frac{|u|}{\sigma\sqrt{n}}\right)^\gamma du &= O\left(\frac{n^{(k-\gamma)/2}}{\epsilon_n^k}\int_{e^{\beta n}}^\infty \frac{1}{u^{k-\gamma}}du\right)\\ &= O\left(\frac{n^{(k-\gamma)/2}}{\epsilon_n^k e^{(k-1-\gamma)\beta n}}\right),\end{aligned}$$

which tends to zero, as $n\to+\infty$, provided we choose $k>\gamma+1$. The uniformity in ξ is obvious from the proof. □

Combining Lemmas 5.6, 5.7 and 5.9, proves Lemma 5.5. Theorem 5.2 now follows by a standard approximation argument. More precisely, for $\delta > 0$, choose compactly supported C^k functions $\chi_1 \leq \mathbf{1}_{[-1,1]} \leq \chi_2$ such that

$$2 - \delta \leq \int \chi_1(x)dx \leq \int \chi_2(x)dx \leq 2 + \delta.$$

Then

$$\begin{aligned}
\frac{-\delta}{2\sqrt{2\pi}\sigma} &\leq \liminf_{n\to+\infty} \sup_{\xi\in\mathbb{R}} \left(\frac{\sqrt{n}}{2\epsilon_n} \mu\left\{\underline{i} \in \Sigma :\ f^n(\underline{i}) \in [\xi - \epsilon_n, \xi + \epsilon_n]\right\} - \frac{1}{\sqrt{2\pi}\sigma} e^{-\xi^2/2\sigma^2 n} \right) \\
&\leq \limsup_{n\to+\infty} \sup_{\xi\in\mathbb{R}} \left(\frac{\sqrt{n}}{2\epsilon_n} \mu\left\{\underline{i} \in \Sigma :\ f^n(\underline{i}) \in [\xi - \epsilon_n, \xi + \epsilon_n]\right\} - \frac{1}{\sqrt{2\pi}\sigma} e^{-\xi^2/2\sigma^2 n} \right) \\
&\leq \frac{\delta}{2\sqrt{2\pi}\sigma}.
\end{aligned}$$

6 Large Deviations

We will use the notation of the preceding section. In particular, μ will be the Gibbs measure associated to a Hölder continuous function $g : \Sigma \to \mathbb{R}$, which is assumed to be normalized so that $L_g 1 = 1$ and $L_g^* \mu = \mu$. Writing $f = r - \lambda_\mu$ and noting that

$$\sum_{k=0}^{n-1} f(\sigma^k \underline{i}) = d(x, g_{i_1} g_{i_2} \cdots g_{i_n} x) - n\lambda_\mu + O(1),$$

we see that Theorem 1.4 follows from the result below.

Theorem 6.1 (Large Deviations Theorem)**.** *Let $\epsilon > 0$. We have that*

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mu \left\{ \underline{i} \in \Sigma :\ \left| \frac{1}{n} \sum_{k=0}^{n-1} f(\sigma^j \underline{i}) \right| > \epsilon \right\} < 0.$$

Proof. We recall the straightforward proof (cf. [7], [16], [22]). For $t \in \mathbb{R}$, let $L_{g+tf} : C^\alpha(\Sigma) \to C^\alpha(\Sigma)$ be the transfer operator defined on $C^\alpha(\Sigma)$ by

$$L_{g+tf} h(\underline{i}) = \sum_{\sigma(\underline{j}) = \underline{i}} e^{g(\underline{j}) + tf(\underline{j})} h(\underline{j}).$$

Then, in particular, $\|L_{g+tf}^n 1\|_\infty \leq C e^{nP(tf)}$, for some constant $C > 0$, where $P(tf)$ denotes the pressure of the function $tf \in C^\alpha(\Sigma)$. Thus, for any $t \in \mathbb{R}$,

$$\begin{aligned}
\mu \left\{ \underline{i} \in \Sigma :\ \sum_{k=0}^{n-1} f(\sigma^k \underline{i}) > n\epsilon \right\} &\leq \int e^{t\left(\sum_{k=0}^{n-1} f(\sigma^k \underline{i}) - n\epsilon\right)} d\mu(\underline{i}) \\
&= e^{-tn\epsilon} \int L_{g+tf}^n 1(\underline{i}) d\mu(\underline{i}).
\end{aligned}$$

Similarly, for any $t \in \mathbb{R}$,

$$\mu \left\{ \underline{i} \in \Sigma : \sum_{k=0}^{n-1} f(\sigma^k \underline{i}) < -n\epsilon \right\} \leq \int e^{-t\left(\sum_{k=0}^{n-1} f(\sigma^k \underline{i}) + n\epsilon\right)} d\mu(\underline{i})$$

$$= e^{-tn\epsilon} \int L_{g-tf}^n 1(\underline{i}) d\mu(\underline{i}).$$

Therefore, one can bound the rate of convergence

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mu \left\{ \underline{i} \in \Sigma : \left| \frac{1}{n} \sum_{k=0}^{n-1} f(T^i \underline{i}) \right| > \epsilon \right\}$$

by the maximum of $\inf_{t \in \mathbb{R}} \{P(g+tf) - t\epsilon\}$ and $\inf_{t \in \mathbb{R}} \{P(g+tf) + t\epsilon\}$. Now,

$$\inf_{t \in \mathbb{R}} \{P(g+tf) - t\epsilon\} = P(g + \xi f) - \xi \epsilon,$$

where ξ is the unique real number such that

$$\frac{d}{dt} P(g+tf) \bigg|_{t=\xi} = \int f d\mu_{g+\xi f} = \epsilon.$$

In particular, using $\int f d\mu = 0$ and the fact that $P(g+tf)$ is strictly convex, $\xi = 0$ if and only if $\epsilon = 0$. Now, by definition,

$$P(g+\xi f) = h(\mu_{g+\xi f}) + \xi \int f d\mu_{g+\xi f} + \int g d\mu_{g+\xi f},$$

and so

$$P(g+\xi f) - \xi\epsilon = h(\mu_{g+\xi f}) + \int g d\mu_{g+\xi f} \leq P(g) = 0,$$

with equality if and only if $\xi = 0$. A similar calculation for $-\epsilon$ completes the proof. □

References

[1] R. Adler and L. Flatto, Geodesic flows, interval maps, and symbolic dynamics, *Bull. Amer. Math. Soc.* 25 (1991) 229–334.

[2] A. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics, 91. Springer-Verlag, New York, 1983.

[3] P. Billingsley, *Convergence of probability measures*, John Wiley, New York-London-Sydney 1968.

[4] M. Björklund, Central Limit Theorems for Gromov hyperbolic groups, *Preprint*, 2009.

[5] M. Bourdon, Immeubles hyperboliques, dimension conforme et rigidit de Mostow, *Geom. Funct. Anal.* 7 (1997) 245–268.

[6] R. Bowen, Symbolic dynamics for hyperbolic flows, *Amer. J. Math.* 95 (1973), 429–460.

[7] A. Broise, Transformations dilatantes de l'intervalle et theorèmes limites, *Asterisque* 238, 5-109, Societe mathematique de France, 1996.

[8] J. Cannon, The combinatorial structure of cocompact discrete hyperbolic groups, *Geom. Dedicata*, 16 (1984) 123-148.

[9] F. Dal'bo, Remarques sur le spectre des longueurs d'une surface et comptages, *Bol. Soc. Bras. Mat.* 30 (1999), 199-221.

[10] M. Denker and W. Philipp, Approximation by Brownian motion for Gibbs measures and flows under a function, *Ergodic Theory Dynam. Systems* 4 (1984), no. 4, 541-552.

[11] D. Dolgopyat, On decay of correlations in Anosov flows, *Ann. of Math.* 147 (1998), no. 2, 357-390.

[12] H. Furstenberg and H. Kesten, Products of random matrices. *Ann. Math. Statist.* 31 (1960) 457–469.

[13] E. Ghys and P. de le Harpe, *Sur les groupes hyperboliques d'après Mikhael Gromov* (Bern, 1988), 117-134, Progr. Math., 83, Birkhuser Boston, Boston, MA, 1990.

[14] Y. Guivarc'h and A. Raugi, Products of random matrices: convergence theorems. Random matrices and their applications (Brunswick, Maine, 1984), 31–54, *Contemp. Math.*, 50, Amer. Math. Soc., Providence, RI, 1986.

[15] H. Hennion, Limit theorems for products of positive random matrices, *Ann. Probab.* 25 (1997), no. 4, 1545-1587.

[16] Y. Kifer, Large deviations in dynamical systems and stochastic processs, *Trans. Amer. Math. Soc.* 321 (1990), 505-524.

[17] S. Lalley, Renewal theorems in symbolic dynamics, with applications to geodesic flows, non-Euclidean tessellations and their fractal limits, *Acta Math.* 163 (1989), no. 1-2, 1-55.

[18] F. Ledrappier, Some asymptotic properties of random walks on groups, in *Topics in probability and Lie groups: boundary theory*, pp. 117-152, CRM Proc. Lecture Notes, 28, Amer. Math. Soc., Providence, RI, 2001.

[19] E. Le Page, Théorèmes limites pour les produits de matrices aléatoires, in *Probability measures on groups* (Oberwolfach, 1981), pp. 258-303, Lecture Notes in Math., 928, Springer, Berlin-New York, 1982.

[20] A. Lubotzky, *Discrete Groups, Expanding graphs and invariant measures*, Progress in Mathematics 125, Birkhauser, Zurich, 1994.

[21] F. Naud, Expanding maps on Cantor sets and analytic continuation of zeta functions, *Ann. Sci. École Norm. Sup.* 38 (2005), 116-153.

[22] S. Orey and S. Pelikan, Deviations of trajectory averages and the defect in Pesin's formula for Anosov diffeomorphisms, *Trans. Amer. Math. Soc.* 315 (1989), 741-753.

[23] W. Parry and M. Pollicott, Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* 187-188 (1990), 1-268.

[24] W. Philipp and W. Stout, Almost sure invariance principles for partial sums of weakly dependent random variables, *Mem. Amer. Math. Soc.* 2 (1975) no. 161.

[25] M. Pollicott and R. Sharp, The circle problem on surfaces of variable negative curvature, *Monat. Math.* 123 (1997), 61-70.

[26] M. Pollicott and R. Sharp, Comparison theorems and orbit counting in hyperbolic geometry, *Trans. Amer. Math. Soc.* 350 (1998), 473-499.

[27] M. Pollicott and R. Sharp, Exponential error terms for growth functions on negatively curved surfaces, *Amer. J. Math*. 120 (1998), 1019-1042.

[28] M. Pollicott and R. Sharp, Distribution of ergodic sums for hyperbolic maps, in *Representation theory, dynamical systems, and asymptotic combinatorics*, 167–183, Amer. Math. Soc. Transl. Ser. 2, 217, Amer. Math. Soc., Providence, RI, 2006.

[29] D. Ruelle, Thermodynamic Formalism, Addison-Wesley, New York, 1978.

[30] J. Rousseau-Egele, Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux, *Ann. Probab.* 11 (1983), 772-788.

[31] S. Sawyer and T. Steger, The rate of escape for anisotropic random walks in a tree, *Probab. Th. and Related Fields* 76 (1987) 207-230.

[32] C. Series, Symbolic dynamics for geodesic flows, *Acta Math.* 146 (1981), no. 1-2, 103-128.

[33] C. Walkden, Invariance principles for iterated maps that contract on average, *Trans. Amer. Math. Soc.* 359 (2007), 1081-1097.