Winding of geodesic rays chosen by a harmonic measure

Timothée Bénard

Abstract

We prove limit theorems for the homological winding of geodesic rays distributed via a harmonic measure on a Gromov hyperbolic space. We obtain applications to the inverse problem for the harmonic measure, and winding statistics for Patterson-Sullivan measures.

Contents

1 Introduction .................................................. 2
  1.1 The inverse problem for the harmonic measure ............ 2
  1.2 Homological winding of \( \nu \)-typical rays ................. 4
  1.3 Winding of Patterson-Sullivan measures .................. 8
  1.4 Strategy and organization of the paper .................... 9

2 Geodesic tracking of sample paths .......................... 11

3 The norm and the Busemann cocycle ....................... 13
  3.1 Busemann cocycle ...................................... 13
  3.2 Exit estimates for the norm ................................ 15

4 Limit laws for geodesic rays ................................ 15
  4.1 Law of large numbers .................................. 15
  4.2 Central limit theorem .................................. 17
  4.3 The covariance matrix ................................ 18
  4.4 Law of the iterated logarithm .......................... 19
  4.5 Principle of large deviations ......................... 20
  4.6 Gambler’s ruin ........................................ 21

5 Winding and Patterson-Sullivan measures .................. 22
  5.1 Winding on convex cocompact manifolds .................... 23
  5.2 Singularity of harmonic measures ....................... 26

6 Open questions .................................................. 27

*The author has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 803711).
1 Introduction

1.1 The inverse problem for the harmonic measure

Let \( \Omega \) be a proper geodesic hyperbolic metric space endowed with a basepoint \( o \), and \( \Gamma \) be a finitely generated subgroup of isometries of \( \Omega \). Given a non-elementary probability measure \( \mu \) on \( \Gamma \), it is known that for \( \mu \)-almost every \( b = (b_i)_{i \geq 1} \in \Gamma^{\mathbb{N}} \), the sample path \( (b_1 \ldots b_n.o)_{n \geq 0} \) on \( \Omega \) converges to a point \( \xi_b \) in the Gromov boundary \( \partial \Omega \). The distribution \( \nu \) of the exit point \( \xi_b \) is called the harmonic measure or the Furstenberg measure of generator \( \mu \). It is the only probability measure on \( \partial \Omega \) satisfying the stationarity equation

\[
\mu \ast \nu = \nu
\]

The properties of regularity of \( \nu \) are very much studied [33, 48] [37, 46, 47] [13, 10], and can be used to describe the asymptotic properties of \( \mu \)-sample paths on \( \Omega \) [9] [36] [8] [48].

The motivation for this text is to address the following inverse problem for the harmonic measure:

*How much information on \( \mu \) can be read in (the measure class of) its harmonic measure \( \nu \) on the boundary?*

We will shed some light on this question by examining the homological winding of \( \nu \)-typical geodesic rays on \( \Gamma \backslash \Omega \). As the statement of the main result requires a rather substantial introduction, we start with a few remarkable corollaries. The main result, as well as standard terminology concerning \( \mu \), are presented in Section 1.2.

(I) Expectation and covariance of \( \mu_{ab} \). In the context of random walks on hyperbolic groups, we see that \( \nu \) informs on the half lines spanned by the expectation and the covariance of \( \mu \) in the abelianization.

**Corollary 1.1.** Let \( (\Gamma, d_\Gamma) \) be a word hyperbolic group equipped with a word metric, \( \mu \) a non-elementary measure on \( \Gamma \), and \( \nu \) its harmonic measure on the Gromov boundary \( \partial \Gamma \). Write \( H_\Gamma \) the torsion-free part of the abelianization of \( \Gamma \), and \( \mu_{ab} \) the projection of \( \mu \) to \( H_\Gamma \).

- If \( \mu \) has finite first moment, then the measure class of \( \nu \) determines

\[
\lambda_\mu^{-1} E(\mu_{ab})
\]

where \( \lambda_\mu > 0 \) is the rate of escape of the \( \mu \)-walk on \( (\Gamma, d_\Gamma) \).
If \( \mu \) has finite second moment and centered projection to \( H_{\Gamma} \), then the measure class of \( \nu \) determines
\[
\lambda_{\mu}^{-1}\text{Cov}(\mu_{ab})
\]
In the above statement, \( \mathbb{E}(\mu_{ab}), \text{Cov}(\mu_{ab}) \) are respectively the expectation and the covariance of \( \mu_{ab} \) in the real vector space spanned by the free abelian group \( H_{\Gamma} \). These quantities would not make sense on the torsion part of \( \Gamma/\Gamma \).

It is remarkable that these formulas link objects which have very different origins. The rate of escape \( \lambda_\mu \) arises from hyperbolicity, and depends strongly on the word distance \( d_\Gamma \) on \( \Gamma \). In contrast, \( \mu_{ab} \) is independent from the metric and lives in a commutative setting.

Corollary 1.1 is a direct consequence of Theorem 1.4 below and the formulas for the drift and the limit covariance.

(II) Non-centered walks are singular at infinity. Assume \( \Omega \) is the real hyperbolic space \( \mathbb{H}^N \) of dimension \( N \geq 2 \), and \( \Gamma \) is a convex cocompact non-elementary discrete subgroup of orientation preserving isometries. An established question is to understand the properties of measures \( \mu \) for which \( \nu \) is in the class of Patterson-Sullivan on \( \partial \mathbb{H}^N \). Note that such measures do exist by a result of Connell-Muchnik \[18\], and can even be chosen with finite exponential moment by results of Lalley and Li \[44, 49\]. An active conjecture, commonly attributed to Kaimanovich-Le Prince, states that \( \mu \) cannot be finitely supported \[39, 42, 43\]. We prove here that \( \mu \) must be centered.

**Corollary 1.2.** Let \( \Gamma \leq \text{Isom}^+(\mathbb{H}^N) \) be a convex cocompact discrete subgroup, \( \mu \) a non-elementary probability measure on \( \Gamma \) with finite first moment in a word metric. If the harmonic measure of \( \mu \) is in the class of Patterson-Sullivan, then the projection of \( \mu \) to the torsion-free part of \( \Gamma/\Gamma \) is centered.

Corollary 1.2 is a direct consequence of Theorem 1.4 and Lemma 5.2 below.

(III) Walks with cusps are singular at infinity. Furstenberg’s discretization of the Brownian motion \[28, 29\] implies that any (possibly non-uniform) lattice of \( \text{PSL}_2(\mathbb{R}) \) is the support of a probability measure \( \mu \) whose harmonic measure is Lebesgue on \( \partial \mathbb{H}^2 \cong S^1 \). The proof gives \( \mu \) with finite first moment in the metric induced by \( \text{PSL}_2(\mathbb{R}) \). It came thus as a surprise when Guivarc’h-Le Jan \[34, 35\] proved twenty years later that \( \mu \) cannot have finite first moment in the word metric of \( \Gamma \) when \( \Gamma \) is the congruence subgroup \( \Gamma(2) \leq \text{PSL}_2(\mathbb{Z}) \). The result was extended step by step to any lattice \( \Gamma \leq \text{PSL}_2(\mathbb{R}) \) containing a parabolic element \[22, 30\]. Further generalisations were made in \[33, 31\], notably the case where \( \Gamma \leq \text{PSL}_2(\mathbb{R}) \) is any discrete finitely generated subgroup with a parabolic element, which was solved under the additional assumptions that \( \mu \) generates \( \Gamma \) as a semigroup and has superexponential moment \[31\]. We show those two assumptions are unnecessary.

**Corollary 1.3.** Let \( \Gamma \leq \text{PSL}_2(\mathbb{R}) \) be a finitely generated discrete subgroup containing a parabolic element. Let \( \mu \) be a non-elementary probability measure on \( \Gamma \) with finite first moment in a word metric. Then the Furstenberg measure of \( \mu \) on \( \partial \mathbb{H}^2 \) is singular with the class of Patterson-Sullivan.
The proof of Corollary 1.3 is given in Section 5.2. It relies on Theorem 1.5 and a result of Enriquez-Franchi-Le Jan [24].

1.2 Homological winding of $\nu$-typical rays

We address the inverse problem for the harmonic measure by reading information about $\mu$ in the winding statistics of $\nu$-typical geodesic rays on $\Gamma\setminus\Omega$. More precisely, for every $\xi \in \partial\Omega$, call $r_\xi : \mathbb{R}^+ \to \Omega$ a geodesic ray from $o$ to $\xi$ on $\Omega$. Letting $\xi$ vary with law $\nu$, we obtain a family of random segments $(r_\xi([0,t]))_{t \in \mathbb{R}^+}$ indexed by the time parameter. We prove limit laws for the winding of $(r_\xi([0,t]))_{t \in \mathbb{R}^+}$ on $\Gamma\setminus\Omega$. In those limit theorems, natural numerical quantities appear (drift, limit covariance) and have explicit formulas involving $\mu$. Those formulas only depend on $\nu$ so they yield numerical invariants for all possible generators $\mu$ of $\nu$. This strategy is motivated by the note of Guivarc’h-Le Jan [34]. It is also related to Gadre-Maher-Tiozzo [30] and Randecker-Tiozzo [53], as we discuss below.

Homological projection

The winding number of a geodesic segment $[o,z]$ on $\Gamma\setminus\Omega$ will be formally incarnated by a vector $i(z)$ where $i$ is a $\Gamma$-equivariant map from $\Omega$ to a vector space. Examples below justify the relation with homology and the term winding.

**Definition.** A measurable map $i : \Omega \to \mathbb{R}^d$ ($d \geq 1$) is $\Gamma$-equivariant if there exists a morphism $\pi : \Gamma \to \mathbb{R}^d$ satisfying for all $\gamma \in \Gamma$, $x \in \Omega$,

$$i(\gamma \cdot x) = \pi(\gamma) + i(x)$$

Note that $\pi$ is unique.

We consider two natural levels of regularity on $i$:

- We say $i$ is locally bounded if for every ball $B$ of $\Omega$, we have $\sup_B \| i \| < \infty$.
- We say $i$ is quasi-Lipschitz if there exists $R > 0$ such that for all $x, y \in \Omega$, we have

$$\| i(x) - i(y) \| \leq R d_\Omega(x, y) + R$$

**Examples.**

1. *(Abelianization of hyperbolic group).* Let $\Gamma$ be a word hyperbolic group, and $\pi : \Gamma \to \mathbb{R}^d$ any group morphism. Given any Cayley graph $\Omega$ of $\Gamma$, we can extend $\pi$ into $i : \Omega \to \mathbb{R}^d$, say by affine interpolation. The regularity of $i$ is always Lipschitz.

2. *(Homological winding on Gromov hyperbolic manifolds).* Let $\Omega$ be a simply connected complete Riemannian manifold which is Gromov hyperbolic, and $\Gamma$ a subgroup of isometries of $\partial\Omega$. Given a basepoint $o \in \Omega$ and a family $\omega_1, \ldots, \omega_d$ of $\Gamma$-invariant closed 1-forms on $\Omega$, we can define a $\Gamma$-equivariant map $i : \Omega \to \mathbb{R}^d$ by setting

$$i(z) = \left( \int_{[a,z]} \omega_1, \ldots, \int_{[a,z]} \omega_d \right)$$
where \([o, z]\) stands for any smooth path from \(o\) to \(z\).

Observe that \(i\) is smooth, whence locally bounded. It can be quasi-Lipschitz, for instance if \(\sup_{x \in \Omega, i \in \mathbb{R}} \|\omega_i\|_x < \infty\). However, if \(\Gamma \setminus \Omega = M\) is a manifold and some \(\omega_i\) has non zero integral on a loop \(l\) of \(M\) that may be shrunk arbitrarily by homotopy, then \(i\) cannot be quasi-Lipschitz. Indeed, first note that \(\omega_i\) has the same integral \(c \neq 0\) on all the homotopic deformations of \(l\) in \(M\), then for any \(\varepsilon > 0\), choose such a deformation of length at most \(\varepsilon^{-1}\) times into a path in \(\Omega\) with endpoints \(z, z'\) satisfying \(d(z, z') \leq 1\), and observe however that \(\|i(z) - i(z')\| \geq \varepsilon^{-1} |c| \neq 0\) can be made arbitrarily large by choosing \(\varepsilon\) small, thus causing \(i\) to be non quasi-Lipschitz. For instance, this occurs if \(M\) admits a cusp containing a loop with non trivial homology, as in the hyperbolic 3-horned sphere (but not the once-punctured torus).

When \(M = \Gamma \setminus \Omega\) is a manifold, the vector \(i(z)\) is related to the homological winding of the path \([o, z]\) projected to \(M\). Indeed, assume that \(\omega_1, \ldots, \omega_d\) project to a basis of \(H^1(M, \mathbb{R})\). This yields a representative \(\omega\) for any cohomology class in \(H^1(M, \mathbb{R})\), that we see as a \(\Gamma\)-invariant closed 1-form on \(\Omega\). Using that the integral pairing \(H^1(M, \mathbb{R}) \times H_1(M, \mathbb{R}) \to \mathbb{R}\) is non-degenerate (De Rham Theorem), we then have

\[i : \Omega \to H_1(M, \mathbb{R}), \quad z \mapsto \int_{[o, z]}\]

Here we use that we specified representatives for a basis of \(H^1(M, \mathbb{R})\) because the integral of an element of \(H^1(M, \mathbb{R})\) is not well defined on an arbitrary path in \(M\) or \(\Omega\). A priori, \(i\) depends on the choice of basepoint \(o\) and on the basis representatives \(\omega_i\). However, any other choice will lead to a \(\Gamma\)-equivariant map \(i'\) sharing the same morphism \(\pi : \Gamma \to H_1(M, \mathbb{R})\). In this sense, the construction is canonical. The theorems below will not distinguish \(i\) from \(i'\).

**Historical comments.** The homological winding for the geodesic flow on a negatively curved manifold has been thoroughly investigated from the point of view of invariant measures (say Liouville measure or Bowen-Margulis-Sullivan measures). For the compact case, see \([57, 54, 21]\). In presence of cusps, see \([35, 45]\). Homological winding is also related to the problem of counting closed geodesics as in \([41]\). A more precise form of winding, involving 1-forms which are only closed in a neighborhood of the cusps, is studied in \([25, 24, 27, 1]\).

Homological winding for rays chosen with a Furstenberg measure has not been much studied. It is hinted at in the seminal note of Guivarc’h-Le Jan \([34]\). It is also related to the works of Gadre-Maher-Tiozzo \([30]\) and Randecker-Tiozzo \([53]\) who consider a typical geodesic ray on a hyperbolic manifold and show that the non-oriented winding around a given cusp satisfies a law of large numbers with finite mean. In this paper, we establish a wide range of limit theorems for the homological winding of Furstenberg-typical geodesic rays. Contrary to \([30, 53]\) which rely on Kingman’s subadditive ergodic theorem, we compare directly geodesic rays with an abelian random walk (or martingale in the non-centered case) in the homology group. This new approach allows a precise control of the winding statistics and to free oneself from explicit geometrical computations, yielding with no extra cost the case of an arbitrary hyperbolic geodesic metric space. This level of generality,
notably describing abelianized geodesic rays in the context of a word hyperbolic group, does not seem to have prior history.

Statistics of $\nu$-typical rays

We now present the main result of the paper, describing the winding statistics of geodesic rays on $\Gamma \backslash \Omega$ chosen by a Furstenberg measure $\nu$ of generator $\mu$ supported by $\Gamma$. Several conditions on $\mu$ will come up:

- $\mu$ is non-elementary if the semigroup generated by its support contains two hyperbolic isometries with disjoint sets of fixed points.
- $\mu$ has finite $p$-th moment if $\sum_{\gamma \in \Gamma} d_{\Gamma}(e, \gamma)^p \mu(\gamma) < \infty$ where $d_{\Gamma}$ is any word distance on $\Gamma$.
- $\mu$ has finite exponential moment if $\exists s > 0$ such that $\sum_{\gamma \in \Gamma} e^{sd_{\Gamma}(e, \gamma)} \mu(\gamma) < \infty$.

We divide our result in two statements, assuming different levels of regularity for the projection.

Theorem 1.4 assumes $i$ to be quasi-Lipschitz. This situation occurs if $i$ is the projection of a hyperbolic group to its abelianization, or the homological projection for some convex cocompact hyperbolic manifold. We prove a law of large numbers (LLN), and a central limit theorem (CLT), a law of the iterated logarithm (LIL), a principle a large deviations (PLD), and a gambler’s ruin estimate (GR).

**Theorem 1.4** (Limit laws for geodesic rays I). Let $\Omega$ be a proper geodesic hyperbolic metric space with a basepoint, and $\Gamma$ a finitely generated subgroup of isometries of $\Omega$. Fix a quasi-Lipschitz $\Gamma$-equivariant map $i : \Omega \to \mathbb{R}^d$. Let $\nu$ be a Furstenberg measure on $\partial \Omega$ for some non-elementary probability measure $\mu$ on $\Gamma$.

- (LLN) If $\mu$ has finite first moment, then there is a vector $e_\nu \in \mathbb{R}^d$ such that for $\nu$-almost every $\xi \in \partial \Gamma$,
  $$\frac{i \circ r_\xi(t)}{t} \xrightarrow{t \to +\infty} e_\nu$$

- (CLT) If $\mu$ has finite second moment, then as $\xi$ varies with law $\nu$, we have the convergence in distribution,
  $$\frac{i \circ r_\xi(t) - te_\nu}{\sqrt{t}} \xrightarrow{\text{dist}} \mathcal{N}(0, A_\nu)$$
  for some positive semidefinite symmetric matrix $A_\nu \in M_d(\mathbb{R})$.

- (LIL) If $\mu$ has finite second moment, then for $\nu$-almost every $\xi \in \partial \Gamma$,
  $$\left\{ \text{Accumulation points of } \left( \frac{i \circ r_\xi(t) - te_\nu}{\sqrt{2t \log \log t}} \right)_{t \to +\infty} \right\} = A_\nu^{1/2}(\mathcal{B}(0, 1))$$
  where $\mathcal{B}(0, 1) = \{t \in \mathbb{R}^d : \sum_i t_i^2 \leq 1\}$.

- (PLD) If $\mu$ has finite exponential moment, then for all $\alpha > 0$, there exists $D, \delta > 0$ such that
  $$\nu\{ \xi : \|i \circ r_\xi(t) - te_\nu\| \geq \alpha t \} \leq De^{-\delta t}$$
• (GR) If \( \mu \) has finite exponential moment, \( d = 1 \), and \( A_\nu > 0 \), then for any \( k, l > 0 \) and any \( (k_s)_{s>0}, (l_s)_{s>0} \in \mathbb{R}^{R>0} \) such that \( k_s \sim sk, l_s \sim sl \), we have

\[
\nu \{ \xi : (i \circ r_\xi(t) - t e_\nu)_{t \geq 0} \text{ leaves } [-k_s, l_s] \text{ via } (l_s, +\infty) \} \xrightarrow{s \to +\infty} \frac{k}{k + l}
\]

**Remarks.** 1) We recall that the notation \( r_\xi \) refers to a choice of geodesic ray from the basepoint of \( \Omega \) to \( \xi \in \partial \Omega \). The theorem does not depend on the choice geodesics rays as any two geodesic rays on \( \Omega \) with same starting point and same exit point in the Gromov boundary remain at distance bounded by a constant depending only on \( \Omega \).

2) In the second statement, we wrote \( N(0, A_\nu) \) for the (possibly degenerate) centered gaussian measure on \( \mathbb{R}^d \) of covariance matrix \( A_\nu \).

3) The LLN and the LIL imply that the mean \( e_\nu \) and the covariance matrix \( A_\nu \) only depend on the measure class of \( \nu \). We will also give explicit formulas in terms of \( \mu \).

Theorem 1.5 below is more general, for it only assumes \( i \) to be locally bounded. As observed earlier in Example 2., it is particularly relevant when \( i \) is the homological projection to a hyperbolic manifold with a cusp. We prove a weak law of large numbers (WLLN), and a central limit theorem (CLT) encapsulating that of Theorem 1.4.

**Theorem 1.5** (limit laws for geodesic rays II). Let \( \Omega \) be a proper geodesic hyperbolic metric space with a basepoint, and \( \Gamma \) a finitely generated subgroup of isometries of \( \Omega \). Fix a locally bounded \( \Gamma \)-equivariant map \( i : \Omega \to \mathbb{R}^d \). Let \( \nu \) be a Furstenberg measure on \( \partial \Omega \) for some non-elementary probability measure \( \mu \) on \( \Gamma \).

• (WLLN) If \( \mu \) has finite first moment, then as \( \xi \) varies with law \( \nu \), we have the convergence in probability

\[
\frac{i \circ r_\xi(t)}{t} \xrightarrow{\text{prob} \quad t \to +\infty} e_\nu
\]

for some vector \( e_\nu \in \mathbb{R}^d \).

• (CLT) If \( \mu \) has finite second moment, then as \( \xi \) varies with law \( \nu \), we have the convergence in distribution,

\[
\frac{i \circ r_\xi(t) - t e_\nu}{\sqrt{t}} \xrightarrow{\text{dist} \quad t \to +\infty} N(0, A_\nu)
\]

for some positive semidefinite symmetric matrix \( A_\nu \in M_d(\mathbb{R}) \).

**Mean and covariance matrix.** The mean \( e_\nu \) and the covariance matrix \( A_\nu \) appearing in Theorems 1.4, 1.5 have explicit formulas in terms of the measure \( \mu \). As \( e_\nu, A_\nu \) only depend on \( \nu \), those formulas give numerical invariants on all possible \( \mu \) such that \( \mu \ast \nu = \nu \).
To express $e_\nu, A_\nu$ we need introduce the rate of escape $\lambda_\mu$ of the $\mu$-sample paths on $\Omega$:

$$\lambda_\mu = \lim_{n \to +\infty} \frac{d_\Omega(o, b_1 \ldots b_n o)}{n}$$

where $d_\Omega$ is the distance on $\Omega$, and $(b_i)_{i \geq 1} \in \Gamma^{\mathbb{N}^*}$ is $\mu^{\otimes \mathbb{N}^*}$-typical. An application of Kingman’s subadditive ergodic theorem to the probability space $(\Gamma^{\mathbb{N}^*}, \mu^{\otimes \mathbb{N}^*})$ endowed with the shift shows that if $\mu$ has finite first moment, then $\lambda_\mu$ is well defined and the convergence also holds in $L^1(\Gamma^{\mathbb{N}^*}, \mu^{\otimes \mathbb{N}^*})$. If $\mu$ is also non-elementary, then $\lambda_\mu$ is non-zero, see for instance [8, Proposition 3.3].

Letting $\mu_{ab} = \pi_* \mu$, we prove

$$e_\nu = \lambda_\mu^{-1} \mathbb{E}(\mu_{ab})$$

To describe $A_\mu$, we introduce the Busemann boundary $X$ of $(\Omega, d_\Omega)$, the Busemann cocycle $\sigma : \Gamma \times X \to \mathbb{R}$. We prove in Section 4.2 that uniformly in $x \in X$,

$$\lambda_\mu A_\nu = \lim_{n \to +\infty} \frac{1}{n} \int_{\Gamma} (\pi(\gamma) - \sigma(\gamma^{-1}, x)e_\nu)^t (\pi(\gamma) - \sigma(\gamma^{-1}, x)e_\nu) d\mu^n(\gamma)$$

(1)

where $v \mapsto ^tv$ refers to the transposition map. In particular, we get the complement information that

- $A_\nu$ coincides with the covariance matrix $\lambda_\mu^{-1} \text{Cov}(\mu_{ab})$ if $\mu_{ab}$ is centered (but not in general).
- $A_\nu$ is non-degenerate if $\langle \text{supp } \mu_{ab} \rangle_{\text{Vect}} = \mathbb{R}^d$ and the restriction of the stable length function $l : \Gamma \to \mathbb{R}^+, \gamma \mapsto \lim_{n \to +\infty} \frac{1}{n} d_\Omega(o, \gamma^n o)$ to the semigroup generated by the support of $\mu$ does not factor via $\pi$ (Lemma 4.1).

The connection between the statistics of the winding $i \circ r_\xi$ and that of the $\mu_{ab}$-random walk on $\mathbb{R}^d$ arises from the strategy of the proof, which is to approximate $\nu$-typical geodesic rays on $\Omega$ by $\mu$-sample paths then to project them to $\mathbb{R}^d$. It came thus as a surprise, while working on this paper, to find that $A_\nu$ may not be proportional to $\text{Cov}(\mu_{ab})$. Even worse: it is possible that $\text{Cov}(\mu_{ab})$ has rank $d - 1$ while $A_\nu$ has rank $d$. This is the case if $\Gamma$ is the free group $F_2 = \langle u, v \rangle$, $\pi : F_2 \to \mathbb{Z}^2, u, v \mapsto (1, 0), (0, 1)$, and $\mu = \frac{1}{3}(\delta_u + \delta_{uv} + \delta_{uv^{-1}})$, (see Lemma 4.2). More explanations on this peculiarity are given in the strategy of proof below.

### 1.3 Winding of Patterson-Sullivan measures

Applying Theorems 1.4, 1.5 to the context of Example 2 above, we obtain winding statistics for the geodesic flow directed by a Furstenberg measure on a general manifold $M$ with Gromov hyperbolic universal cover. The particular case of Patterson-Sullivan measures for convex cocompact discrete subgroups yields the following. Recall that $\mathbb{H}^N$ is the $N$-dimensional real hyperbolic space.
Corollary 1.6 (Homological winding for PS-measures). Let $\Gamma \leq Isom^+(\mathbb{H}^N)$ be a convex cocompact torsion-free non-elementary discrete subgroup, and $m$ be any probability measure on $\partial \mathbb{H}^N$ that is absolutely continuous with respect to the class of Patterson-Sullivan of $\Gamma$.

Then the homological projection of $m$-typical geodesic rays to $H_1(\Gamma \backslash \mathbb{H}^N, \mathbb{R})$ has sublinear escape, satisfies a central limit theorem with non-degenerate covariance matrix, the law of the iterated logarithm, and the gambler’s ruin estimate (for any 1-dimensional projection). It also satisfies the principle of large deviations if $m$ has $L^{1+\varepsilon}$-Radon Nikodym derivatives with respect to Patterson-Sullivan measures.

The proof is given in Section 5.1. It combines Theorem 1.4, the fact that PS-measures arise as Furstenberg measures for walks with finite exponential moment [49], and -to establish sublinear escape- the fact that PS-measures represent the geodesic projection of the Bowen-Margulis-Sullivan measure, which is symmetric and flow-invariant on $\Gamma \backslash \mathbb{H}^N$.

The CLT (for $m$ general), the (multi-dimensional) LIL, the GR, appear to be new, even in the most simple case of real hyperbolic surfaces, i.e. when $N = 2$.

The LLN, the PLD, and the CLT when $m$ is a Patterson-Sullivan measure, can be obtained by combining results already present in the literature. The seminal case of the CLT on a compact hyperbolic manifold follows from Sinai’s paper [57]. For the general case, the argument consists in coding the geodesic flow by a suspension flow above a shift of finite type [14, 19], then apply techniques of symbolic dynamics [54, 21, 61]. Our approach is different, we are directly able to simulate random geodesic rays projected in homology by a random walk on $H_1(\Gamma \backslash \mathbb{H}^N, \mathbb{R})$ with i.i.d. increments, and apply classical results for additive Markov chains.

We also signal to the reader only interested in Corollary 1.6 that the proof of this result, once isolated, is much shorter than the present paper. Indeed, it does not use the appendix whose purpose is to deal with the non-centered case in Theorems 1.4, 1.3.

1.4 Strategy and organization of the paper

The guiding principle to prove Theorems 1.4, 1.5 is simple. Let $(B, \beta) = (\Gamma^{\mathbb{N}*}, \mu^{\otimes \mathbb{N}*})$ denote the space of instructions. Typical $b \in B$ give a sample path $b_1 \ldots b_n.o$ on $\Omega$ that stays close to the geodesic ray $r_{\xi_b}$, whose distribution coincides with $\nu$ (for $b$ varying with law $\beta$). Hence, we aim to infer the behavior of $i \circ r_{\xi_b}(n\lambda_\mu)$ by comparison with that of $\pi(b_1 \ldots b_n).$ Unfortunately, $\pi(b_1 \ldots b_n)$ only approximates $i \circ r_{\xi_b}(t_{b,n})$ where $t_{b,n} = d_\Omega(o, b_1 \ldots b_n.o)$ varies at scale $\sqrt{n}$ around $n\lambda_\mu$. This approximation is not precise enough to prove second-order limit theorems. For this reason, we introduce the stopping times $\tau_{n}(b)$ at which $(b_1 \ldots b_{n},o)_{i \geq 1}$ leaves the ball of radius $n\lambda_\mu$, and we study how $\pi(b_1 \ldots b_{\tau_{n}(b)}(b))$ approximates $i \circ r_{\xi_b}(n\lambda_\mu)$. Then we reduce the limit theorems for $i \circ r_{\xi_b}(n\lambda_\mu)$ to limit theorems for the random variables $\pi(b_1 \ldots b_{\tau_{n}(b)})$, and we obtain the latter through a careful analysis of the stopping times $\tau_n$. The aforementioned reduction also requires some work, given that both sequences of variables ultimately do not yield the same covariance matrix. This discrepancy occurs in the non-centered case $\mathbb{E}(\pi_\bullet \mu) \neq 0$, where we also need take
into account the recentering parameter. This forces to deal with martingales instead of an i.i.d. additive Markov chain on $\mathbb{R}^d$. The complications brought up by this recentering can be observed in the formula for $A_\nu$.

The paper is organized as follows.

Section 2 explains how $\nu$-typical geodesic rays on $\Omega$ may be simulated with a small error by the $\mu$-sample paths on $\Omega$. More precisely, we recall various results of geodesic tracking from [60, 3, 17], and establish a uniform deviation inequality conditionally to a stopping time.

Section 3 recalls from [8] how the distance function can be replaced by a cocycle in order to study $\mu$-sample paths on $\Omega$. As a first application, we estimate the time of exit of sample paths from a ball, as well as the overshoot.

Section 4 is dedicated to the proof of Theorems 1.4, 1.5. It relies on the two previous sections and on Appendix A. We also comment on the limit covariance matrix $A_\nu$.

Section 5 starts by recalling the notion of Patterson-Sullivan measures associated to a geometrically finite discrete subgroup of isometries of $\mathbb{H}^N$. We then prove winding statistics for PS-measures in the convex cocompact case (Corollary 1.6), and establish singularity with harmonic measures in the presence of cusps (Corollary 1.3).

Section 6 records a few open questions leading to natural continuations of our study.

Appendix A establishes limit laws for martingales arising from cocycles. The interest is that our setting is very general, encapsulating that of [9, 8, 20], and deals with optimal 1 moment assumptions. We also obtain a stronger version of the CLT, in which the time variable $n$ is replaced by a family of stopping times $(\tau_n)$ approximating $n$ at scale $\sqrt{n}$.

Remark. We believe the assumption of properness on $\Omega$ is not essential in the paper, and could be replaced by only assuming separability. Without properness, the Gromov boundary does not identify with classes of geodesic rays, but with classes of $(1, Q)$-quasi geodesic rays, where $Q > 0$ is some fixed constant determined by the hyperbolicity constant of $\Omega$ [40, Remark 2.16]. Hence, we rather need define $r_\xi$ as a $(1, Q)$-quasi geodesic ray from $o$ to $\xi$. To adapt the proof, we point out that we only use the properness assumption on $\Omega$ is Section 3 while referring to Benoist-Quint’s paper [8] to introduce the horofunction boundary, the Busemann cocycle, and its properties. We believe [8] could be adapted almost verbatim to remove the properness assumption, using the general theory of horofunction compactification developed by Maher and Tiozzo in [50]. Asking that $\Omega$ be separable would guarantee as in [50] that this compactification is indeed a metric space. Not to burden the article, we do not pursue in this direction.

Acknowledgements. I am thankful to Péter Varjú for useful comments on a preliminary version of this text. I am also grateful to the anonymous referee for his meticulous reading and valuable remarks.

\footnote{at least for the CLT, LIL, PLD.}
2 Geodesic tracking of sample paths

We explain how the distribution of $\nu$-typical geodesic rays on $\Omega$ can be approximated by that of $\mu$-sample paths.

We keep the notations $(\Omega, d_\Omega, o, (r_\xi)_{\xi \in \partial \Omega}, \Gamma)$ of Section 1, let $\mu$ be a non-elementary probability measure on $\Gamma$ and $\nu$ its Furstenberg measure on $\partial \Omega$. We set $(B, \beta) = (\Gamma^{N^k}, \mu^{\otimes N^k})$ and recall that for $\beta$-typical $b \in B$, the sample path $(b_1 \ldots b_n, o)_{n \geq 0}$ on $\Omega$ converges to some point in the Gromov boundary $\xi_b \in \partial \Omega$, whose distribution under $\beta$ coincides with $\nu$. The corresponding geodesic ray $r_{\xi_b}$ is said to be asymptotic to the sample path $(b_1 \ldots b_n, o)_{n \geq 0}$. We also write $t_{b,n} = d_\Omega(o, b_1 \ldots b_n, o)$.

We start with a very general result stating that $\mu$-sample paths on $\Omega$ have a small chance to deviate from their asymptotic geodesic ray, regardless of the strategy deciding the time of observation. A strategy is incarnated by a stopping time on $(B, \beta)$, implicitly referring to the filtration $(\mathcal{B}_k)_{k \geq 1}$ where $\mathcal{B}_k$ is the sub $\sigma$-algebra of $B$ generated by $b_1, \ldots, b_k$.

**Theorem 2.1** (Deviation inequality with a stopping time). Let $\mu$ be a non-elementary probability measure on $\Gamma$ and $\nu$ its Furstenberg measure on $\partial \Omega$. There exists a family $(\varepsilon_R)_{R \geq 0} \in (0,1)^{\mathbb{R}^+}$ such that $\varepsilon_R \to 0$ as $R \to +\infty$, and for every $\beta$-a.e. finite stopping time $\tau : B \to \mathbb{N}^* \cup \{\infty\}$, every $R > 0$,

$$
\beta\{b : d_\Omega(b_1 \ldots b_{\tau(b)}, o, r_{\xi_b}(t_{b,\tau(b)})) > R\} \leq \varepsilon_R
$$

Moreover, if $\mu$ has finite exponential moment, then one can choose $\varepsilon_R = D e^{-\delta R}$ for some constants $D, \delta > 0$.

**Remarks.** 1) For $\tau = n$ constant, this result follows from [11, Proposition 2.11] and the remark that $d(b_1 \ldots b_n, r_{\xi_b}(t_{b,n})) \leq 2d(b_1 \ldots b_n, r_{\xi_b}(\mathbb{R}^+))$. We explained earlier in the strategy of proof why we need the stronger version above.

2) If $\mu$ has finite $p$-th moment where $p > 1$, the proof of Theorem 2.1 combined with the log$^{p-1}$-regularity of $\nu$ [8, Proposition 4.2] allows to choose $\varepsilon_R$ so that $\int_{R>0} \varepsilon_R R^{p-2}dR < \infty$.

**Proof of Theorem 2.1**. For $x \in \Omega$, $\xi \in \partial \Omega$, we let $F(x, \xi)$ be the image of any geodesic ray from $x$ to $\xi$. By hyperbolicity, the family $(F(x, \xi))_{x \in \Omega, \xi \in \partial \Omega}$ is coarsely $\Gamma$-equivariant, meaning that for $g \in \Gamma$, the distance between $gf(x, \xi)$ and $F(gx, g\xi)$ is bounded by $O(\delta_\Omega)$ where $\delta_\Omega$ is a hyperbolicity constant of $\Omega$.

Observe that for any $s \in \mathbb{R}^+$,

$$
d_\Omega(b_1 \ldots b_{\tau(b)}, o, r_{\xi_b}(t_{b,\tau(b)})) \leq d_\Omega(b_1 \ldots b_{\tau(b)}, o, r_{\xi_b}(s)) + d_\Omega(r_{\xi_b}(s), r_{\xi_b}(t_{b,\tau(b)}))
= d_\Omega(b_1 \ldots b_{\tau(b)}, o, r_{\xi_b}(s)) + |s - t_{b,\tau(b)}|
\leq 2d_\Omega(b_1 \ldots b_{\tau(b)}, o, r_{\xi_b}(s))
$$

where the last line uses that $|d_\Omega(x, z) - d_\Omega(y, z)| \leq d_\Omega(x, y)$ for all $x, y, z \in \Omega$.
Taking the infimum over $s \in \mathbb{R}^+$, we deduce
\[
\frac{1}{2}d_\Omega(b_1 \ldots b_{r(b)} \cdot o, r_{\xi_o}(t_{b, r(b)})) \leq d_\Omega(b_1 \ldots b_{r(b)} \cdot o, F(o, \xi_b)) + O(\delta_\Omega)
\]
\[
= d_\Omega(o, F(b_{r(b)}^{-1} \ldots b_1^{-1} \cdot o, \xi_{T_{r(b)}(b)}) + O(\delta_\Omega)
\]
\[
\leq (b_{r(b)}^{-1} \ldots b_1^{-1} \cdot o | \xi_{T_{r(b)}(b)})_o + O(\delta_\Omega)
\]
where $T : B \to B, (b_i)_{i \geq 1} \mapsto (b_{i+1})_{i \geq 1}$ is the one-sided shift, $(. | .)_o$ denotes the Gromov product based at the origin $o$. For the last inequality we use the fact that for $x, y \in \Omega$, $(x | y)_o \geq d_\Omega(o, [x, y]) - O(\delta_\Omega)$ where $[x, y]$ is the geodesic segment from $x$ to $y$ (see for instance [32, Chapter 2 §1.4]) and pass to the limit to allow $y \in \Omega \cup \partial \Omega$.

In particular, using that $T_{r(b)}b$ varies with law $\beta$ when $b$ varies with law $\beta$, and does so independently from $b_1, \ldots b_{r(b)}$, we get

\[
\beta\{b : d_\Omega(b_1 \ldots b_{r(b)} \cdot o, r_{\xi_o}(t_{b, r(b)})) > 2R\} \leq \beta \sup_{x \in \Omega} \{b : (x | \xi_o) > R - O(\delta_\Omega)\}
\]

Given $x \in \Omega, \xi, \eta \in \partial \Omega$, the definition of the hyperbolicity of $\Omega$ in terms of Gromov product [32] Chapter 2 §1.3] gives
\[
\frac{(x | \xi)_o}{(x | \eta)_o} \geq R \quad \implies \quad (\xi | \eta)_o \geq R - O(\delta_\Omega)
\]

Endow $\partial \Omega$ with a Gromov metric $d_{\partial \Omega}$ [32] Chapter 7 §3]. It satisfies that for some small enough $\varepsilon > 0$, one has $d_{\partial \Omega}(\xi, \eta) \leq \varepsilon^{-1}e^{-c(\xi | \eta)_o}$. In particular, the set $\{\xi \in \partial \Omega : (x | \xi)_o > R - O(\delta_\Omega)\}$ is included in a ball of $\partial \Omega$ whose radius is exponentially small in $R$. The fact that $\nu$ has no atom, and is even Hölder regular when $\mu$ has finite exponential moment [11] Corollary 2.13, finishes the proof.

We now recall various bounds for the distance separating a fixed typical sample path on $\Omega$ and its asymptotic geodesic ray.

**Theorem 2.2 (Geodesic tracking).** Let $\mu$ be a non-elementary probability measure on $\Gamma$ and $\nu$ its Furstenberg measure on $\partial \Omega$.

- For every $\varepsilon > 0$, there exists $R > 0$ such that for $\beta$-almost every $b \in B$,
  \[
  \liminf_{n \to +\infty} \frac{1}{n} \#\{i \leq n : d_\Omega(b_1 \ldots b_i \cdot o, r_{\xi_b}(t_{b,i})) \leq R\} \geq 1 - \varepsilon
  \]

- Assume $\mu$ has finite $p$-th moment where $p > 0$. Then for $\beta$-almost every $b \in B$,
  \[
  d_\Omega(b_1 \ldots b_n \cdot o, r_{\xi_b}(t_{b,n})) = o(n^{\frac{1}{p}})
  \]

- Assume $\mu$ has finite exponential moment. There exists a constant $D > 0$ such that for $\beta$-almost every $b \in B$, for large $n \geq 0$,
  \[
  d_\Omega(b_1 \ldots b_n \cdot o, r_{\xi_b}(t_{b,n})) \leq D \log n
  \]
By way of proof, we make a few comments.

**Remarks.** 1) Notice that the first bound does not require any moment condition on $\mu$. A similar result in the context of random walks on homogeneous spaces of arbitrary rank appears in [3, Theorem A bis] and the proof easily extends to our setting. We will not use this estimate, but we recall it for completeness.

2) For $\mu$ with finite first moment, Tiozzo establishes sublinear escape [60]. For any $p > 0$, the stronger estimates above is proven by Choi in [17, Theorem D].

3) The third bound assuming finite exponential moment follows directly from Theorem 2.1 and the Borel-Cantelli lemma.

### 3 The norm and the Busemann cocycle

The proof of Theorems 1.5, 1.4 requires a fine understanding of the distance $t_{b,n}$. However, it is not so handy to manipulate. We show here that $t_{b,n}$ can be replaced by a cocycle, up to a bounded error. The latter is much easier to deal with. A similar analysis is carried out by Benoist-Quint in [8, Sec. 2.3, 3.2, 4.3]. As a first application, we infer estimates for the exit time and overshoot of $t_{b,n}$ from a bounded interval.

We let $(\Omega, d_\Omega, o, \Gamma)$ be as in Section 1, and fix $\mu$ a non-elementary probability measure on $\Gamma$.

#### 3.1 Busemann cocycle

Every $z \in \Omega$ defines a so-called horofunction

$$h_z : \Omega \to \mathbb{R}, \quad y \mapsto d_\Omega(z, y) - d_\Omega(z, o)$$

and the map $\Omega \to C^0(\Omega, \mathbb{R})$ is injective continuous when $C^0(\Omega, \mathbb{R})$ is given the (metrizable) compact-open topology. The **Busemann compactification** $\overline{\Omega}^B$ of $\Omega$ is the closure of $\{h_z, z \in \Omega\}$ in $C^0(\Omega, \mathbb{R})$. The injection $z \mapsto h_z$ is an embedding that identifies $\Omega$ with a dense open set of $\overline{\Omega}^B$. Moreover, as every horofunction is 1-Lipschitz and vanishes at the basepoint, $\overline{\Omega}^B$ must be compact. We call **Busemann boundary** of $\Omega$ the set $X = \overline{\Omega}^B \setminus \Omega$. The $\Gamma$-action on $\Omega$ extends continuously to the boundary, via the formula: $\forall x \in \overline{\Omega}^B, \ z \in \Omega,$

$$h_{\gamma \cdot x}(z) = h_x(\gamma^{-1}z) - h_x(\gamma^{-1}o)$$

**Remark.** At the contrary of the Gromov boundary $\partial \Omega$, the Busemann boundary $X$ is not invariant by quasi-isometry. For instance, let $\Gamma = F_2 \times \mathbb{Z}/2\mathbb{Z}$ where $F_2$ is the group freely generated by two elements $\{u, v\}$. For the systems of generators $S_1 = \{e, u^{\pm 1}, v^{\pm 1}\} \times \{0, 1\}$, $S_2 = \{e, u^{\pm 1}, v^{\pm 1}\} \times \{1\}$, the Busemann boundaries of the associated Cayley graphs are respectively $\partial F_2$, $\partial F_2 \times \{0, 1\}$.

---

2To see this, call $\Omega_1, \Omega_2, H$ the Cayley graphs of $(\Gamma, S_1)$, $(\Gamma, S_2)$, $(F_2, \{u^{\pm 1}, v^{\pm 1}\})$, and $\partial_B \Omega_1, \partial_B \Omega_2, \partial_B H$ their Busemann boundaries. They come with natural projections $\Omega_i \to H, x \mapsto x$. On the one hand, we have $h_{2^{33}}(y) = \ldots$
However, the identity map $\Omega \to \Omega$ extends continuously to a surjective $\Gamma$-equivariant map $X \to \partial \Omega$.

We define the Busemann cocycle $\sigma : \Gamma \times \overline{\Omega}^B \to \mathbb{R}, (\gamma, x) \mapsto h_x(\gamma^{-1}.o)$. The term cocycle means that $\sigma$ satisfies the relation $\sigma(\gamma_2\gamma_1, x) = \sigma(\gamma_2, \gamma_1x) + \sigma(\gamma_1, x)$ for $\gamma_1, \gamma_2 \in \Gamma, x \in \overline{\Omega}^B$. It is also useful to observe that for all $\gamma \in \Gamma$,

$$\sup_{x \in \overline{\Omega}^B} |\sigma(\gamma, x)| \leq d_\Omega(o, \gamma.o)$$

(2)

as the $h_x$ are 1-Lipschitz and vanish at $o$. In particular, if $\mu$ has finite $k$-th moment, then the family of variables $(\sigma(., x))_{x \in \overline{\Omega}^B}$ is uniformly bounded in $L^k(\Gamma, \mu)$. The next proposition claims that the cocycle $\sigma$ gives a good approximation of the distance to $o$ along the $\mu$-trajectories. It assumes no moment condition on $\mu$.

**Proposition 3.1** (Cocycle approximation). *For every $\varepsilon > 0$, there exists $R > 0$ such that for all $x \in X$,*

$$\beta \left( b : \sup_{n \geq 1} |\sigma(b^{-1}_n \ldots b^{-1}_1, x) - t_{b,n}| \leq R \right) > 1 - \varepsilon$$

*Proof.* This is [8, Proposition 3.2] applied to $(\Omega, d_\Omega)$ and the image $\tilde{\mu}$ of $\mu$ by the inverse map $\gamma \mapsto \gamma^{-1}$. \hfill \qed

When $\mu$ has a finite second moment, the restriction of $\sigma$ to the boundary is known to be cohomologous to a cocycle with constant $\tilde{\mu}$-drift [8, Proposition 4.6]: there exist a cocycle $\sigma_0 : \Gamma \times X \to \mathbb{R}$ and a bounded function $\psi : X \to \mathbb{R}$ such that

- $\forall \gamma \in \Gamma, \forall x \in X$,
  $$\sigma(\gamma, x) = \sigma_0(\gamma, x) + \psi(\gamma x) - \psi(x)$$

- $\forall x \in X$,
  $$\sum_{\gamma \in \Gamma} \sigma_0(\gamma^{-1}, x) \mu(\gamma) = \lambda_\mu$$

It is worth noting that the definition of $\psi$ is explicit: letting $\nu_X$ be any $\mu$-stationary probability measure on $X$, we can set

$$\psi(x) = \int_X (x \mid y)_o d\nu_X(y)$$

where $(\cdot \mid .)_o$ refers to the Gromov product extended to $X$ (see [8, Section 2.3]).

**Remark.** It will be important for us to know that $\psi$ is not only bounded, but continuous. This property is not mentioned in [8]. However it can be deduced from the above formula, the fact that $(\cdot \mid .)_o$ is continuous, and Inequality (4.10) in [8]. Indeed the latter implies that the sub-integral $\sup_{x \in X} \int_{\{[x \mid .]_o \geq R\}} (x \mid y)_o d\nu_X(y) = o(1)$ as $R \to +\infty$.

$h^H_{\Omega}$ as long as $d_H(\Omega, \Gamma) \geq 1$, so $\partial_B \Omega_1 = \partial_B H = \partial F_2$. However, for every $p \in F_2$, the horofunctions $h^\Omega_{\{p, 0\}}$ and $h^{\Omega_2}_{\{p, 1\}}$ do not have the same parity on $\Gamma$, hence the distance between them is bounded below positively independently of $p$. This implies that any sequence converging to the Busemann boundary eventually remains in one leaf $F_2 \times \{\ast\}$, and the equality $\partial \Omega_2 = \partial F_2 \times \{0, 1\}$ follows.
3.2 Exit estimates for the norm

Given $s > 0$, set

$$\tau_s : B \rightarrow \mathbb{N} \cup \{\infty\}, \ b \mapsto \inf\{k \geq 1, t_{b,k} \geq s\lambda_\mu\}.$$ 

As a first application of the cocycle approximation from Section 3.1 we show that the overshoot $t_{b,\tau_s(b)} - s\lambda_\mu$ is essentially bounded, and that $\tau_s$ lands roughly in a window of size $\sqrt{s}$ around $s$.

**Proposition 3.2.** 1. (Overshoot control) Assume $\mu$ has finite first moment. Then

$$\sup_{s > 0} \beta(b : t_{b,\tau_s(b)} > s\lambda_\mu + R) \xrightarrow{R \rightarrow +\infty} 0$$

2. (Time control) Assume $\mu$ has finite second moment. We then have:

$$\forall \varepsilon > 0, \exists R > 0, \forall s > 1,$$

$$\beta(b : |\tau_s(b) - s| > R\sqrt{s}) \leq \varepsilon$$

**Proof.** Fix $x \in X$. According to the cocycle approximation from Proposition 3.1, we do not loose generality by replacing $t_{b,n}$ by $\sigma(b_{\gamma^{-1},x},,)$, both in the definition of $\tau_s$ and in the estimates of the proposition. The time estimate for $\sigma(b_{\gamma^{-1},x},)$ follows from Proposition A.18 applied to $\lambda^{-1}\sigma_0$. The overshoot control also follows from Proposition A.18 after replacing $\mu$ by a fixed convolution power $\mu^{*\sigma_0}$ to reduce to the case where the drift has a positive lower bound uniform on $X$. Indeed Proposition A.18 is applicable for $\mu^{*\sigma_0}$ if $n_0$ is large: for any $n_0 \geq 1$, the family $(\sigma(\cdot,x))_{x \in X}$ is uniformly bounded in $L^1(\Gamma,\mu^{*\sigma_0})$ and for large $n_0$, for all $x \in X$

$$\int_{\Gamma} \sigma(\gamma^{-1},x) d\mu^{*\sigma_0}(\gamma) \geq \frac{1}{2} n_0\lambda_\mu$$

due to Proposition 3.1, the convergence $n^{-1}t_{b,n} \rightarrow \lambda_\mu$ in $L^1(\Gamma^{\mathbb{N}^*},\mu^{\otimes\mathbb{N}^*})$, and the domination $\sup_{x \in X} |\sigma(\gamma,x)| \leq d_Q(o,\gamma,o)$.

4 Limit laws for geodesic rays

This section is dedicated to the proof of Theorem 1.4 and Theorem 1.5. Several parts of the proof can be extended to general limit theorems for cocycles. Those are postponed to Appendix A.

4.1 Law of large numbers

We establish the law of large numbers for $\nu$-typical geodesic rays on $\Omega$ projected to $\mathbb{R}^d$. We argue under various conditions of regularity on the projection $i$.

For $i$ quasi-Lipschitz, we prove the strong law of large numbers in Theorem 1.4. The proof uses the sublinear geodesic tracking from Theorem 2.2.
Proof of Theorem 1.4 (LLN). By Theorem 2.2 and the quasi-Lipschitz assumption on \( i \), we have that for \( \beta \)-almost every \( b \in B, n \geq 2 \),

\[ \| i(b_1 \ldots b_n) - i \circ r_{\xi_b}(t_{b,n}) \| = o(n) \]

or in other terms

\[ \| \pi(b_1 \ldots b_n) - i \circ r_{\xi_b}(t_{b,n}) \| = o(n) \]

as \( i(b_1 \ldots b_n) = \pi(b_1 \ldots b_n) + i(o) \). Note that the assumption of finite first moment on \( \mu \) for a word metric implies that its projection \( \pi_* \mu = \mu_{ab} \) has a finite first moment as a measure on \( \mathbb{R}^d \). By the law of large numbers for the \( \mu_{ab} \)-walk on \( \mathbb{R}^d \) and the definition of the rate of escape \( \lambda_{\mu} \), we know that

\[ \pi(b_1 \ldots b_n) = n \mathbb{E}(\mu_{ab}) + o(n) \]

\[ t_{b,n} = n \lambda_{\mu} + o(n) \]

Combining these estimates, we get

\[ \| n \mathbb{E}(\mu_{ab}) - i \circ r_{\xi_b}(n \lambda_{\mu}) \| = o(n) \]

which implies that \( i \circ r_{\xi_b}(n \lambda_{\mu}) = n \lambda_{\mu} e_\nu + o(n) \). The result follows for continuous times \( t \in \mathbb{R}^+ \) as \( i \) is quasi-Lipschitz.

\[ \square \]

For \( i \) locally bounded, we prove the weak law of large numbers in Theorem 1.5. The proof relies on the deviation inequality with stopping time from Theorem 2.1 and on the overshoot control in Proposition 3.2.

Proof of Theorem 1.5 (WLLN). For \( n \geq 0 \), recall the stopping time

\[ \tau_n : B \to \mathbb{N} \cup \{ \infty \}, b \mapsto \inf\{k \geq 1, t_{b,k} \geq n \lambda_{\mu} \}. \]

The deviation estimate from Theorem 2.1 guarantees that for every \( \varepsilon > 0 \), there exists \( R > 0 \) such that for all \( n \geq 0 \),

\[ \beta\{ b : d(b_1 \ldots b_{\tau_n(b)}, i(0), r_{\xi_b}(t_{b,\tau_n(b)})) \leq R \} \geq 1 - \varepsilon \]

Using Proposition 3.2, we may increase \( R \) to also have for every \( n \geq 0 \),

\[ \beta\{ b : d(b_1 \ldots b_{\tau_n(b)}, i(0), r_{\xi_b}(n \lambda_{\mu})) \leq R \} \geq 1 - \varepsilon \]

Using that \( i \) is locally bounded and \( \Gamma \)-equivariant, we can see that those \( b \) satisfy

\[ \| \pi(b_1 \ldots b_{\tau_n(b)}) - i \circ r_{\xi_b}(n \lambda_{\mu}) \| \leq C_R \text{ for some constant } C_R \text{ only depending on } (i, R). \]

Hence, it is enough to show the almost-sure convergence

\[ \frac{\pi(b_1 \ldots b_{\tau_n(b)})}{n} \to \lambda_{\mu} e_\nu = \mathbb{E}(\mu_{ab}) \] \quad (3)

The equivalence \( t_{b,n} \sim n \lambda_{\mu} \) implies \( \tau_n \sim n \). Hence (3) follows from the law of large numbers for \( \mu_{ab} \) on \( \mathbb{R}^d \).

\[ \square \]
4.2 Central limit theorem

We prove the central limit theorem from Theorem 1.5.

Proof of Theorem 1.5 (CLT). We want to show that, for \( b \) varying with law \( \beta \), the family of random variables

\[
\frac{i \circ r_{\xi_b}(s\lambda_b) - s\lambda_b e_\nu}{\sqrt{s}}
\]

converges to a (possibly degenerate) Gaussian distribution as \( s \to +\infty \). By Theorem 2.1 and Proposition 3.2 (overshoot control), for any \( \varepsilon > 0 \), there exists \( R > 0 \) such that for any \( s > 0 \),

\[
\beta(b : d(b_1 \ldots b_{\tau_b(b)}, o, r_{\xi_b}(s\lambda_b)) \leq R) \geq 1 - \varepsilon
\]

Using that \( i \) is locally bounded and \( \Gamma \)-equivariant, we deduce that, up to increasing \( R \), for any \( s > 0 \),

\[
\beta(b : |\pi(b_1 \ldots b_{\tau_b(b)}) - i \circ r_{\xi_b}(s\lambda_b)| \leq R) \geq 1 - \varepsilon
\]

Hence, the CLT for \( \nu \)-typical geodesic rays in (4) amounts to that of

\[
\frac{\pi(b_1 \ldots b_{\tau_b(b)}) - s\lambda_b e_\nu}{\sqrt{s}}
\]

Fix \( x \in X \). By Proposition 3.2 (overshoot control) and Proposition 3.1, the family of variables

\[
\frac{s\lambda_b - \sigma(b_{\tau_b(b)}^{-1} \ldots b_1^{-1}, x)}{\sqrt{s}}
\]

converges to zero in probability as \( s \to +\infty \). Hence we just need to check the CLT for

\[
\frac{\pi(b_1 \ldots b_{\tau_b(b)}) - \sigma(b_{\tau_b(b)}^{-1} \ldots b_1^{-1}, x)e_\nu}{\sqrt{s}}
\]

We introduce the cocycle \( \tilde{\sigma} : \Gamma \times X \to \mathbb{R}^d \) defined by

\[
\tilde{\sigma}(g, x) = \pi(g) + \sigma_0(g, x)e_\nu
\]

This cocycle satisfies \( \|\sigma(g, x)\| \ll |g|^\nu \) for \( |g|^\nu = \|\pi(g)\| + d(o, g, o) \) and has constant zero drift for \( \tilde{\mu} \). It is also continuous in the \( X \)-variable because \( \sigma \) is continuous and the function \( \psi \) relating \( \sigma \) and \( \sigma_0 \) is continuous. Finally Proposition 3.1 confirms that \( \sigma \) is \( (\tilde{\mu}, X) \)-controlled. These observations ensure we may apply our CLT with stopping time from the appendix (Proposition A.2) to \( \tilde{\sigma} \). It tells us that

\[
\frac{i \circ r_{\xi_b}(s\lambda_b) - s\lambda_b e_\nu}{\sqrt{s}} \overset{\text{dist}}{\underset{s \to +\infty}{\rightarrow}} N(0, \lambda_{\mu} A_\nu)
\]

where \( \lambda_{\mu} A_\nu \) is the limit, uniformly in \( x \in X \):

\[
\lambda_{\mu} A_\nu = \lim_{n \to \infty} \frac{1}{n} \int_G \tilde{\sigma}(g, x)^t \tilde{\sigma}(g, x) \, d\tilde{\mu}^* n(g)
\]
Using the cocycle relation and that $\tilde{\sigma}$ has zero drift for $\tilde{\mu}$, we can replace $\sigma_0$ in the above limit by the original Busemann cocycle $\sigma$:

$$\lambda_\mu A_\nu = \lim \frac{1}{n} \int_G (\pi(g) + \sigma(g, x) e_\nu) \left( (\pi(g) + \sigma(g, x) e_\nu) \right) d\tilde{\mu}^n(g)$$

This limit is still uniform in $x \in X$.

\[\square\]

### 4.3 The covariance matrix

We comment on the covariance matrix $A_\nu$ of the limit distribution in the CLT.

We start with a criterion of non-degeneracy for $A_\nu$. Recall that the stable length of an element $\gamma \in \Gamma$ is the quantity $l(\gamma) = \lim_{n \to +\infty} \frac{1}{n} d_\Omega(o, \gamma^n o)$. It is well defined by subadditivity. One can see it as an analog of the logarithm of the spectral radius for matrices in $SL_d(\mathbb{R})$.

**Lemma 4.1** (Non-degeneracy criterion). If $\det(A_\nu) = 0$, then $\langle \text{supp } \mu_{ab}\rangle_{\text{Vect}}$ is a proper subspace of $\mathbb{R}^d$ or there exists $\varphi \in (\mathbb{R}^d)^*$ such that $\varphi \circ \pi(g) = l(g)$ on the semigroup generated by the support of $\mu$.

**Proof.** Assume $\det(A_\nu) = 0$ and $\langle \text{supp } \mu_{ab}\rangle_{\text{Vect}} = \mathbb{R}^d$. By the first hypothesis, there exists a non zero linear form $\varphi$ such that $\varphi A_\nu = 0$. By Remark 2) following Proposition A.1, we can rewrite for every $n \geq 1$, any $\tilde{\mu}$-stationary measure $\nu_X$ on $X$:

$$\lambda_\mu A_\nu = \frac{1}{n} \int_{\Gamma \times X} (\pi(\gamma) + \sigma_0(\gamma, x) e_\nu) \left( (\pi(\gamma) + \sigma_0(\gamma, x) e_\nu) \right) d\tilde{\mu}^n(\gamma) d\nu_X(x).$$

Hence, $\varphi A_\nu \varphi = 0$ reads as

$$\int_{\Gamma \times X} (\varphi \circ \pi(\gamma) + \sigma_0(\gamma, x) \varphi(e_\nu))^2 d\tilde{\mu}^n(\gamma) d\nu_X(x) = 0.$$

Call $\Gamma_{\tilde{\mu}} \subseteq \Gamma$ the semigroup generated by the support of $\tilde{\mu}$. We deduce that for $\nu_X$-almost every $x \in X$, for all $\gamma \in \Gamma_{\tilde{\mu}}$, we have

$$\varphi \circ \pi(\gamma) = -\sigma_0(\gamma, x) \varphi(e_\nu).$$

As $\pi(\Gamma_{\tilde{\mu}})$ spans $\mathbb{R}^d$ by assumption and $\varphi \neq 0$, we must have $\varphi(e_\nu) \neq 0$, so we may assume $\varphi(e_\nu) = 1$. Using that the difference $|\sigma - \sigma_0|$ is bounded, we infer that for $\nu_X$-almost every $x$, for $\gamma \in \Gamma_{\tilde{\mu}}$, $n \geq 1$,

$$\varphi \circ \pi(\gamma) = \frac{1}{n} \varphi \circ \pi(\gamma^n) = -\frac{1}{n} \sigma_0(\gamma^n, x) = -\frac{1}{n} \sigma(\gamma^n, x) + o(1).$$

As the projection of $\nu_X$ to the Gromov boundary $\partial \Omega$ is the Furstenberg measure of $\mu$ and the latter has no atom, there exist in particular $x, y \in X$ with distinct projections on $\partial \Omega$ such that for all $\gamma \in \Gamma_{\tilde{\mu}}$, $n \geq 1$,

$$\varphi \circ \pi(\gamma) = -\frac{1}{n} \max[\sigma(\gamma^n, x), \sigma(\gamma^n, y)] + o(1).$$

18
However, given such $x, y$, there exists $C > 0$ such that
\[ d_{\Omega}(o, g.o) \geq \max(\sigma(g,x), \sigma(g,y)) \geq d_{\Omega}(o, g.o) - C \]
for all $g$ [8 (4.3)]. Letting $n$ go to infinity, we get $\varphi \circ \pi(\gamma) = -l(\gamma)$. As $l(\gamma^{-1}) = l(\gamma)$ and $-\varphi \circ \pi(\gamma) = \varphi \circ \pi(\gamma^{-1})$, the result follows.

We also give an example showing that the rank of $A_\nu$ may be strictly greater than the rank of the covariance matrix of $\mu_{ab}$.

**Lemma 4.2.** Let $\Gamma$ be the free group with two generators $u, v$, write $\Omega$ the Cayley graph associated to $\{u, v\}$, and let $i : \Omega \to \mathbb{R}^2$ continuous $\Gamma$-equivariant such that $i(u) = (1, 0)$, $i(v) = (0, 1)$. For $\mu = \frac{1}{3}(\delta_u + \delta_{uv} + \delta_{uv^{-1}})$ we have
\[ \det(Cov(\mu_{ab})) = 0 \neq \det(A_\nu) \]

**Proof.** We have $\det(Cov(\mu_{ab})) = 0$ because $\mu_{ab}$ is supported on an affine hyperplane of $\mathbb{R}^2$. Assume by contradiction that $\det(A_\nu) = 0$. By Lemma 4.1 there exists $\varphi \in (\mathbb{R}^2)^*$ such that $\varphi \circ \pi(g) = l(g)$ on the semigroup generated by the support of $\mu$. Writing $e_1 = (1, 0)$ and $e_2 = (0, 1)$, we get the equations
\[ \varphi(e_1) = 1, \quad \varphi(e_1 + e_2) = 2, \quad \varphi(e_1 - e_2) = 2. \]
These are not compatible, whence the contradiction.

## 4.4 Law of the iterated logarithm

We prove the law of the iterated logarithm in Theorem 1.4.

**Proof of Theorem 1.4 (LIL).** According to Theorem 2.2 and the second moment assumption on $\mu$, one has for $\beta$-almost every $b \in B$,
\[ d_{\Omega}(b_1 \ldots b_n.o, r_{\xi_b}(t_{b,n})) = o(\sqrt{n}) \]
By the quasi-Lipschitz assumption on $i$, we infer
\[ i \circ r_{\xi_b}(t_{b,n}) = \pi(b_1 \ldots b_n) + o(\sqrt{n}) \]
On the other hand, Proposition 3.1 yields that for a fixed $x \in X$, for $\beta$-almost every $b \in B$, we have
\[ t_{b,n}e_\nu = \sigma(b_n^{-1} \ldots b_1^{-1}, x)e_\nu + O(1) \]
Putting these together, we get
\[ \frac{i \circ r_{\xi_b}(t_{b,n}) - t_{b,n}e_\nu}{\sqrt{2t_{b,n} \log \log t_{b,n}}} = \frac{-\tilde{\sigma}(b_n^{-1} \ldots b_1^{-1}, x) + o(\sqrt{n})}{(1 + o(1))\sqrt{2\lambda_{n}n \log \log n}} \]
where \( \bar{\sigma} \) is the cocycle introduced in Section 4.2. We checked earlier in Section 4.2 that \( \bar{\sigma} \) satisfies the conditions of application of the general LIL for cocycles from Proposition 4.5. By this LIL, for \( \beta \)-almost every \( b \in B \),

\[
\left\{ \text{Accumulation points of } \left( \frac{i \circ r_{\xi_b}(t) - t e_\nu}{\sqrt{2t \log \log t}} \right)_{te\{t_{b,n}, n \in \mathbb{N}\}} \right\} = A_\nu^{1/2}(B(0,1)) \tag{7}
\]

We just need to check there are no other accumulation points. As \( \mu \) has finite second moment, an application of Borel-Cantelli Lemma yields for \( \beta \)-almost every \( b \in B \),

\[
|t_{b,n+1} - t_{b,n}| = o(\sqrt{n})
\]

In particular, for large \( n \),

\[
|t_{b,n+1} - t_{b,n}| \leq \sqrt{t_{b,n}}
\]

It follows that for \( b \) typical, the intervals

\[
I_n(b) = [t_{b,n} - \sqrt{t_{b,n}}, t_{b,n} + \sqrt{t_{b,n}}]
\]

cover a translate of \( \mathbb{R}_+ \). Moreover, the condition that \( i \) is quasi-Lischitz implies that for \( t \) varying inside \( I_n(b) \), the quantity \( \frac{i \circ r_{\xi_b}(t) - t e_\nu}{\sqrt{2t \log \log t}} \) is constant up to an error that tends to 0 as \( n \) goes to infinity. This implies that all the accumulation points of \( \left( \frac{i \circ r_{\xi_b}(t) - t e_\nu}{\sqrt{2t \log \log t}} \right)_{t \geq 2} \) are attainable along a sequence of times among the \( t_{b,n} \)'s. The LIL now follows from (7).

\[\square\]

### 4.5 Principle of large deviations

We prove the principle of large deviations in Theorem 1.4.

**Proof of Theorem 1.4 (PLD).** As \( i \) is quasi-Lipschitz, it is enough to prove that, for every \( \varepsilon > 0 \), there exist \( D, \delta > 0 \) such that for all \( n \geq 0 \),

\[
\beta\{b : \|i \circ r_{\xi_b}(n \lambda_\mu) - n \lambda_\mu e_\nu\| \geq n \varepsilon\} \leq De^{-\delta n}
\]

But we have the bound

\[
\|i \circ r_{\xi_b}(n \lambda_\mu) - n \lambda_\mu e_\nu\|
\leq \|i \circ r_{\xi_b}(n \lambda_\mu) - i \circ r_{\xi_b}(t_{b,n})\| + \|i \circ r_{\xi_b}(t_{b,n}) - i(b_1 \ldots b_n, o)\| + \|i(b_1 \ldots b_n, o) - n \lambda_\mu e_\nu\|
\leq |n \lambda_\mu - t_{b,n}| + d(r_{\xi_b}(t_{b,n}), b_1 \ldots b_n, o) + \|\pi(b_1 \ldots b_n) - n \lambda_\mu e_\nu\| + 1
\]

It is then enough to prove a PLD for each of the terms on the right-hand side. The first term is treated by the PLD for the distance to the origin \([11, \text{Theorem 1.1}]\), the second by the exponential deviation estimate in Theorem 2.1 for \( \tau = n \), and the third by the classical PLD applied to \( \mu_{ab} \). This concludes the proof. \[\square\]
4.6 Gambler’s ruin

We prove the gambler’s ruin estimate in Theorem 1.4. In particular \( d = 1 \).

Proof of Theorem 1.4 (GR). Let \( \varepsilon > 0 \). It is enough to show that

\[
\liminf_{s \to +\infty} \beta \{ b : (\pi \circ r_{\xi_b}(t) - te_\nu)_{t>0} \text{ meets } [l_s, +\infty) \text{ before } (-\infty, -k_s) \} > \frac{k}{k + l} - \varepsilon
\]

Indeed, the lower bound for the \( \lim \sup \) follows by symmetry, and one concludes letting \( \varepsilon \) go to \( 0^+ \).

Let us establish (8). Fix a (large) parameter \( C > 1 \) to be specified below, and \( x \in X \). Write \( M_n(b) = -\tilde{\sigma}(b_1^{-1} \ldots b_1^{-1}, x) = \pi(b_1 \ldots b_n) - \sigma_0(b_1^{-1} \ldots b_1^{-1}, x)e_\nu \) the martingale associated to \( x \). For \( s > C \), let \( E_s \) be the set of elements \( b \in B \) such that

- \( \forall j \in \mathbb{N}^*, |M_j(b) - (\pi(b_1 \ldots b_j) - t_{b,j}e_\nu)| \leq C \)
- \( \forall j \leq s^3, |\pi(b_1 \ldots b_j) - i \circ r_{\xi_b}(t_{b,j})| \leq C \log s \)
- \( \forall j \leq s^3, |t_{b,j+1} - t_{b,j}| \leq C \log s \)

According to Proposition 3.1, Theorem 2.2, and the exponential moment assumption on \( \mu \), we can choose \( C = C(\varepsilon) > 1 \) large enough so that for every \( s > C \), one has

\[
\beta(E_s) > 1 - \varepsilon
\]

We can also assume that \( i \) is \( C \)-quasi-Lipschitz. Let \( F_s \) be the set of \( b \in B \) such that \( M_n(b) \) meets \([l_s + 2C \log s, +\infty) \) before \((-\infty, -k_s + 3C^2 \log s + 3C] \) and during the time interval \([0, s^3] \). Lemmas A.16, A.15 from the appendix can be applied thanks to Proposition 3.1 and yield

\[
\beta(F_s) \underset{s \to +\infty}{\rightarrow} \frac{k}{k + l}
\]

Finally we check that for \( s > C \), \( b \in E_s \cap F_s \), the recentered projected geodesic ray \((\pi \circ r_{\xi_b}(t) - te_\nu)_{t>0} \) meets \([l_s, +\infty) \) before \((-\infty, -k_s) \). Indeed, let \( j_0 \leq s^3 \) minimal such that \( M_{j_0}(b) \geq l_s + 2C \log s \). Using the first two items in the definition of \( E_s \), we get that \( \pi \circ r_{\xi_b}(t_{b,j_0}) - t_{b,j_0}e_\nu \geq l_s \).

We also check that \((i \circ r_{\xi_b}(t) - te_\nu)_{t>0} \) does not meet \((-\infty, -k_s) \) for \( t < t_{b,j_0} \). Suppose by contradiction it is the case: \( \exists t_1 < t_{b,j_0} \),

\[
i \circ r_{\xi_b}(t_1) - t_1e_\nu < -k_s
\]

Let \( 0 \leq j_1 < j_0 \) such that \( t_{b,j_1} \leq t < t_{b,j_1+1} \) and observe that by the last item in the definition of \( E_s \), we have \(|t - t_{b,j_1}| \leq C \log s \). Using that \( i : \Omega \to \mathbb{R}^d \) is \( C \)-quasi-Lipschitz, we get

\[
i \circ r_{\xi_b}(t_{b,j_1}) - t_{b,j_1}e_\nu \leq -ks + 2C^2 \log s + 2C
\]

Then first two items in the definition of \( E_s \) yield \( M_{j_1}(b) \leq -k_s + 3C^2 \log s + 3C \). Hence a contradiction.

\( \square \)
The goal of this section is to prove Corollaries 1.6, 1.3. We first recall basic notions of hyperbolic geometry.

Let $\mathbb{H}^N$ be the the $N$-dimensional real hyperbolic space, $\Gamma \leq Isom^+(\mathbb{H}^N)$ a non-elementary discrete subgroup of orientation preserving isometries. If the orbits of $\Gamma$ in $\mathbb{H}^N$ are not too sparse, one can associate to $\Gamma$ a certain class of measures on the visual boundary $\partial \mathbb{H}^N \simeq \mathbb{S}^{n-1}$ called Patterson-Sullivan measures. We recall their definition, for a more general exposition see [56].

The discrete group $\Gamma$ comes with a limit set $\Lambda_{\Gamma} \subset \partial \mathbb{H}^N$ and a critical exponent $\delta_{\Gamma}$ defined by

$$\Lambda_{\Gamma} = \overline{\Gamma.z} \cap \partial \mathbb{H}^N$$

$$\delta_{\Gamma} = \lim_{R \to +\infty} \frac{1}{R} \log \frac{1}{R} \log \# B_{\mathbb{H}^N}(z, R) \cap \Gamma.z$$

where $z \in \mathbb{H}^N$ is any point, $\overline{\Gamma.z}$ is the closure of $\Gamma.z$ in $\overline{\mathbb{H}^N} = \mathbb{H}^N \cup \partial \mathbb{H}^N$, and $B_{\mathbb{H}^N}(z, R)$ denotes the ball of radius $R$ centered at $z$. As $\Gamma$ is non-elementary, we must have $\# \Lambda_{\Gamma} = +\infty$, $\delta_{\Gamma} > 0$. The number $\delta_{\Gamma}$ also corresponds to the exponent of convergence of the Poincaré series of $\Gamma$, that is the infimum of the positive numbers $s \in \mathbb{R}^+_+$ such that for some (hence any) $x, y \in \mathbb{H}^N$, one has

$$P_{\Gamma}(s, x, y) := \sum_{\gamma \in \Gamma} e^{-s \overline{d_{\mathbb{H}^N}(x, \gamma.y)}} < \infty$$

The value of $P_{\Gamma}(s, x, y)$ at $s = \delta_{\Gamma}$ may be finite or infinite (this does not depend on the choice of $x, y$), and $\Gamma$ is said to be of divergence type if it is infinite.

We also define the Busemann function by

$$\beta_{x}(z_1, z_2) = \lim_{z' \to x, z' \in \mathbb{H}^N} d(z', z_1) - d(z', z_2)$$

for $x \in \partial \mathbb{H}^N$, $z_1, z_2 \in \mathbb{H}^N$.

**Definition** (Patterson-Sullivan measures [52, 53]). Let $\Gamma \leq Isom^+(\mathbb{H}^N)$ be a non-elementary discrete subgroup of orientation preserving isometries. There exists a family of finite Borel measures $(\nu_z)_{z \in \mathbb{H}^N}$ on $\Lambda_{\Gamma}$ such that for all $z, z' \in \mathbb{H}^N$, $\gamma \in \Gamma$,

$$\gamma \ast \nu_z = \nu_{\gamma.z} \quad \text{and} \quad d\nu_{z}(\xi) = e^{-\delta_{\Gamma} \beta_{x}(z, z')} d\nu_{z'}(\xi)$$

Moreover, if $\Gamma$ is of divergence type, then this family is unique up to scalar multiplication. In this case, we call the measures $(\nu_z)_{z \in \mathbb{H}^N}$ the Patterson-Sullivan measures associated to $\Gamma$.

**Remark.** For $\Gamma$ of divergent type, the PS measures of $\Gamma$ are all equivalent to one another. Hence, they define without ambiguity a measure class.

Let $hull(\Lambda_{\Gamma})$ be the convex hull of $\Lambda_{\Gamma}$, in other terms $hull(\Lambda_{\Gamma})$ is the set of points in $\mathbb{H}^N$ belonging to an ideal polyhedral $n$-simplex with all vertices in $\Lambda_{\Gamma}$. In

---

This characterization is indeed equivalent due Caratheodory’s theorem applied to the projective model of $\mathbb{H}^N$, in which geodesics are straight lines.
particular, hull(ΛΓ) is compact because ΛΓ is. We say that Γ is **convex cocompact** if the action of Γ on $C(Γ) = \text{hull}(ΛΓ) \cap \mathbb{H}^N$ is cocompact. It is equivalent to ask that Γ is finitely generated and the orbit maps $θ_z : Γ \to \mathbb{H}^N, γ \mapsto γ. z$ for $z ∈ \mathbb{H}^N$ are quasi-isometric embeddings (for some/any word metric on Γ), see for instance [12, Section 1.8].

Convex cocompact subgroups are of divergence type, and one can realise their PS measures on the boundary $\partial \mathbb{H}^N$ as Furstenberg measures for a suitable probability measure on Γ:

**Fact 5.1 ([49]).** Let $Γ \leq Isom^+(\mathbb{H}^N)$ be a convex cocompact non-elementary discrete subgroup of orientation preserving isometries. For any $z ∈ \mathbb{H}^N$, there exists a probability measure $µ_z$ on Γ whose support generates Γ as a group, which has finite exponential moment, and such that $µ_z \ast ν_z = ν_z$.

### 5.1 Winding on convex cocompact manifolds

We now prove Corollary 1.6. We assume Γ non-elementary, torsion free, discrete, convex cocompact, and we fix $m$ a probability measure on the boundary that is absolutely continuous with respect to the class of Patterson-Sullivan. We aim to estimate the statistics of $m$-typical geodesic rays on $\mathbb{H}^N$ projected via the map $i : \mathbb{H}^N \to H_1(M, \mathbb{R})$ where $M = Γ \backslash \mathbb{H}^N$. We can assume without loss of generality that the basepoint $o$ is in $C(Γ)$. As $C(Γ)$ has compact projection, the norms of the 1-forms defining $i$ are bounded in $C(Γ)$, hence $i$ is Lipschitz on $C(Γ)$. We are thus in a position to apply Theorem 1.4 with $Ω = C(Γ)$ if $m$ is a Furstenberg measure for some probability measure $µ$ with moment conditions. Note this is not always the case, for instance if $m$ lacks regularity. Our strategy is to start with the case where $m = ν_o$ is the PS measure based at $o$, then deduce the general result using the Γ-equivariance of the $\{ν_z, z ∈ \mathbb{H}^N\}$ to approximate any $m$ by a combination of $ν_{γ.o}$.

We start by proving that $ν_o$-typical geodesic rays have zero drift in $H_1(M, \mathbb{R})$. Our argument relies on the dynamics of the geodesic flow for the Bowen-Margulis-Sullivan measure (denoted by $m_{BMS}$) that we now recall.

The BMS-measure can be defined as the unique probability measure on $T^1M$ which is invariant by the geodesic flow and has maximal entropy [31]. It is supported on the compact subset $T^1M_{\hat{K}(Γ)}$ where $K(Γ) = Γ \backslash C(Γ)$ is the **convex core**.

The BMS-measure also has an explicit construction in terms of Patterson-Sullivan measures. Indeed, recall Hopf coordinates allow to identify $T^1\mathbb{H}^N$ with $(\partial \mathbb{H}^N)^{(2)} × \mathbb{R}$ where $(2)$ refers to the direct product minus the diagonal. The identification is given by the map

$$ v \mapsto (v^− = \lim_{−∞} φ_tv, v^+ = \lim_{+∞} φ_tv, B_v(o, v_*) ) $$

where $φ_t$ is the time $t$ geodesic flow, $v_* ∈ \mathbb{H}^N$ stands for the basepoint of $v$. We define a measure $\tilde{m}_{BMS}$ on $T^1\mathbb{H}^N$ by setting in Hopf coordinates

$$ d\tilde{m}_{BMS}(x, y, s) = e^{δt B_{t}(o, v_*)} e^{δt B_{s}(o, v_*)} dν_o(x) dν_o(y) ds $$

...
Lemma 5.2 (zero drift). We have

$$e_{\nu_o} = 0$$

Proof. Extend $i$ to $T^1\mathbb{H}^N$ by precomposing it with the natural projection $T^1\mathbb{H}^N \to \mathbb{H}^N$. Introduce the cocycle

$$\sigma : T^1M \times \mathbb{R} \to H_1(M, \mathbb{R}), \quad (v, t) \mapsto i(\phi_t\tilde{v}) - i(\tilde{v})$$

where $\tilde{v} \in T^1\mathbb{H}^N$ lifts $v$. Note that the $\Gamma$-equivariance of $i$ guarantees that $\sigma$ is well defined. One can interpret $\sigma(v, t)$ has the winding of the path $[v, \phi_tv]$ on $M$. Also, $\sigma$ is a cocycle in the sense that

$$\sigma(v, s + t) = \sigma(v, s) + \sigma(\phi_sv, t)$$

Moreover, $\sigma(v, t)$ is uniformly bounded on the compact set $T^1M_{[K(\Gamma)]} \times [-1, 1]$.

The ergodicity of $(m_{BMS}, \phi_t)$ combined with Birkhoff Ergodic Theorem yields that for $m_{BMS}$-almost-every $v_0 \in T^1M$,

$$\frac{1}{n} (i(\phi_nv_0) - i(v_0)) = \frac{1}{n} \sum_{k=0}^{n-1} \sigma(\phi_kv_0, a_1) \longrightarrow \int_{T^1M} \sigma(v, 1) dm_{BMS}(v)$$

The integral on the right must be zero. Indeed, the measure $m_{BMS}$ is invariant under $\phi_1$ and the involution map $T^1M \to T^1M, v \mapsto \varepsilon(v)$ reversing orientation, so

$$\int_{T^1M} \sigma(v, 1) dm_{BMS}(v) = \frac{1}{2} \int_{T^1M} \sigma(v, 1) + \sigma(\varepsilon(\phi_1v), 1) dm_{BMS}(v) = 0$$

where the equality $\sigma(v, 1) + \sigma(\varepsilon(\phi_1v), 1) = 0$ comes from the computation $\sigma(\varepsilon(\phi_1v), 1) = i(\phi_1\varepsilon(\phi_1\tilde{v})) - i(\varepsilon(\phi_1\tilde{v})) = i(\tilde{v}) - i(\phi_1\tilde{v}) = -\sigma(v, 1)$.

It follows that for $m_{BMS}$-almost every vector $\tilde{v} \in T^1\mathbb{H}^N$, the path $[o, \phi_t\tilde{v}]$ has sublinear drift in $H_1(M, \mathbb{R})$. As the projection of $m_{BMS}$ to the boundary $\partial\mathbb{H}^N$ via $\tilde{v} \mapsto \lim_{t \to +\infty} \phi_t\tilde{v}$ is equivalent to $\nu_o$, we get nullity of the drift for $\nu_o$-typical directions, as required.

We also claim that the Gaussian central limit theorem for $\nu_o$-typical rays is non-degenerate.

Lemma 5.3 (definite covariance matrix). We have

$$\det(A_{\nu_o}) \neq 0$$
Proof. As $e_{\nu_o} = 0$, the formula (1) for the covariance matrix (established in Section 4.2) yields $A_{\nu_o} = \text{Cov}(\mu_{o,ab})$. If $A_{\nu_o}$ were degenerate, then $\mu_{o,ab}$ would be supported on some affine hyperplane of $H_1(M, \mathbb{R})$. As $\mathbb{E}(\mu_{o,ab}) = \lambda_o e_{\nu_o} = 0$, this hyperplane contains $o$. We now get a contradiction, since the support of $\mu_o$ generates $\Gamma$ as a group and $\pi(\Gamma)$ spans $H_1(M, \mathbb{R})$.

The discussion at the beginning of the section, combined with the two previous lemmas, justifies Corollary 1.6 in the case where $m = \nu_o$. We now extend the family of measures that are allowed to obtain the full statement of Corollary 1.6. We let $m$ be a probability measure on $\partial \mathbb{H}^N$ which is absolutely continuous with respect to the class of Patterson-Sullivan. The LLN and LIL stated in Corollary 1.6 for $m$-typical geodesic rays follow immediately from the case $m = \nu_o$, as they only refer to almost-sure behavior.

Let us check the central limit theorem and the gambler’s ruin for geodesic rays chosen via $m$. We use the following decomposition lemma.

**Lemma 5.4.** Let $\varphi : \partial \mathbb{H}^N \to \mathbb{R}_{\geq 0}$ be a smooth positive function (for $\partial \mathbb{H}^N \simeq \mathbb{S}^{N-1}$). Then there exist coefficients $(c_\gamma)_{\gamma \in \Gamma}$ such that $c_\gamma \in \mathbb{R}^+$ and for all $\xi \in \Lambda_{\Gamma}$,

$$\sum_{\gamma \in \Gamma} c_\gamma e^{-\delta_{\Gamma} B_\xi(\gamma_o,o)} = \varphi(\xi)$$

**Proof.** This follows from the appendix in [49]: just replace $R_0 = 1$ by $R_0 = \varphi$ in [49, Section A.3].

**Lemma 5.5.** $m$-typical geodesic rays satisfy a non-degenerate CLT, with same covariance matrix as $\nu_o$, as well the Gamler’s ruin estimate.

**Proof.** We have proven the CLT and the GR for $m \in \{\nu_{\gamma,o}, \gamma \in \Gamma\}$. Observe that they remain true (with same covariance matrix) for any convex combination of these measures, i.e. for measures of the form

$$\sum_{\gamma \in \Gamma} c_\gamma \nu_{\gamma,o} = \left(\sum_{\gamma \in \Gamma} c_\gamma e^{-\delta_{\Gamma} B_\xi(\gamma_o,o)}\right) \nu_o$$

where $c_\gamma \geq 0$, $\sum c_\gamma = 1$. By Lemma 5.4, the sum in the right hand side can be chosen to be any $\varphi$ positive smooth function on the boundary $\partial \mathbb{H}^N$. Now, if $m = \psi \nu_o$ with $\psi \in L^1(\nu_o)$ non negative, we can approximate $\psi$ in $L^1(\nu_o)$ by such functions $\varphi$ to extend the result to $m$.

It remains to check the PLD stated in Corollary 1.6 in case $m$ has $L^{1+\varepsilon}$ Radon-Nikodym derivatives with respect to Patterson-Sullivan measures.

**Lemma 5.6.** If $m = \psi \nu_o$ where $\psi \in L^{1+\varepsilon}(\partial \mathbb{H}^N, \nu_o)$ for some $\varepsilon > 0$, then $m$-typical geodesic rays satisfy the principle of large deviations.
Proof. Given $\alpha > 0$, set

$$E_t = \{ \xi : \| r_\xi (t) \| \geq \alpha t \}$$

By Hölder’s inequality,

$$m(E_t) = \int_{\Lambda_v} 1_{E_t}(\xi) \psi(\xi) d\nu_v(\xi) \leq \nu_v(E_t)^{\varepsilon/(1+\varepsilon)} \| \psi \|_{L^{1+\varepsilon}(\nu_v)}$$

whence exponential decay by the PLD for $\nu_v$ (consequence of Theorem 1.4 and Lemma 5.2).

5.2 Singularity of harmonic measures

To conclude the section, we combine Theorem 1.4 and a result of Enriquez-Franchi-Le Jan [24] to prove Corollary 1.3: Furstenberg measures for $L^1$-random walks on a finitely generated discrete subgroup $\Gamma$ of $PSL_2(\mathbb{R})$ containing a parabolic element are singular with Patterson-Sullivan measures.

Observe that we authorize $\Gamma \setminus \mathbb{H}^2$ to have only one cusp, of trivial homology. In order to apply Theorem 1.4, we need to generate cusps with non-trivial homology. This is the role of

Lemma 5.7. To prove Corollary 1.3, we can assume that $\Gamma$ is torsion free and $\Gamma \setminus [\Gamma, \Gamma]$ contains a parabolic element.

Proof. The idea is to replace $\mu$ by its induction on a finite index subgroup of $\Gamma$.

By Selberg’s Theorem, $\Gamma$ admits a normal subgroup $\Gamma'$ of finite index which is torsion-free. In particular, $M' = \Gamma' \setminus \mathbb{H}^2$ is a hyperbolic surface of finite type with a cusp. Its topology is characterized by its genus $g$ and its number of punctures $p \geq 1$. The first homology group is known to have dimension $2g + p - 1$. More precisely, a basis consists in the $2g$ standard loops of the genus $g$ compactification of $M'$, and any $p - 1$ loops which are homotopic to different cusps. Up to replacing $M'$ by a $2$-cover, one may assume $p \geq 2$. In this case, cusps have non-trivial homology in $M'$. This means that $\Gamma' \setminus \tilde{[\Gamma', \Gamma]}$ contains a parabolic element.

Let $\tau : B \to \mathbb{N} \cup \{ \infty \}, b \mapsto \inf\{ k \geq 1, b_1 \ldots b_k \in \Gamma' \}$ be the hitting time of $\Gamma'$ for the right $\mu$-walk on $\Gamma$. As $\Gamma' \setminus \tilde{\Gamma}$ is finite, $\tau$ must have finite first moment. Introduce the measure $\mu_\tau$, whose distribution is given by $b_1 \ldots b_{\tau(b)}$. By Wald’s formula, $\mu_\tau$ has finite first moment in any word metric on $\Gamma'$, hence on $\Gamma'$. It is also non-elementary, for instance by [17, Theorem A]. Finally, we know from [2, Appendix A.2] that $\mu_\tau$ and $\mu$ share the same Furstenberg measure on the boundary $\partial \mathbb{H}^2$. Summing up, we can replace $\mu$ by $\mu_\tau$ to prove Corollary 1.3, whence the announced reduction.

We now conclude the proof of Corollary 1.3.

Proof of Corollary 1.3. As in [35], followed by [30] [53], the idea is to show that the Furstenberg measure of $\mu$ and the Patterson-Sullivan measure drive geodesic rays with different winding statistics on $M = \Gamma \setminus \mathbb{H}^2$.

26
In view of Lemma 5.7, we may assume that \( \Gamma \) is torsion-free and let \( u \in \Gamma \setminus [\Gamma, \Gamma] \) be a parabolic element. \( u \) corresponds to a loop \( l_u \) around a cusp of \( M \) which is non trivial in \( H_1(M, \mathbb{R}) \). By the De Rham duality theorem, there exists a closed 1-form \( \omega \) on \( M \) with non zero integral on \( l_u \). We write \( \tilde{\omega} \) its lift to \( \mathbb{H}^2 \) and define the \( \Gamma \)-equivariant map \( i : \mathbb{H}^2 \to \mathbb{R} \),

\[
i(z) = \int_{[p,z]} \tilde{\omega}
\]

recording the homological winding around the cusp of \( u \). Observe that \( i \) is locally bounded (though not quasi-Lipschitz).

Write \( \nu \) the Furstenberg measure of \( \mu \) on \( \partial \mathbb{H}^2 \). By the weak law of large numbers in Theorem 1.5, we have for any \( \alpha > 0 \),

\[
\nu(\xi : |i \circ r_\xi(t) - te_\nu| \leq \alpha t) \to 1
\]

Let \( \delta_t \) the exponent of \( \Gamma \) (defined at the beginning of Section 5). Note that \( \delta_t \in (1/2, 1] \) since \( M \) has a cusp [59, p. 265]. Moreover \( \delta_t = 1 \) if and only if \( M \) has no funnel. Let \( D \subseteq \mathbb{H}^2 \) be a fundamental domain for \( \Gamma \), and \( \nu_D \) the geodesic projection of \( \tilde{m}_{BMS} \) on the boundary \( \partial \mathbb{H}^2 \). By [24], the winding of the geodesic flow directed by \( \nu_D \) satisfies a central limit theorem with renormalization in \( 1/t^{2\delta_t - 1} \): for any \( c \in \mathbb{R} \),

\[
\nu_D(\xi : i \circ r_\xi(t) \geq ct \frac{1}{2\delta_t - 1}) \to \mathcal{G}([c, +\infty))
\]

where is \( \mathcal{G} \) some atom-free probability distribution on \( \mathbb{R} \) (more precisely a two-sided stable law of exponent \( 2\delta_t - 1 \in (0, 1) \)) and independent of \( D \). Up to averaging on several domains \( D \), this is still true for some probability \( \nu_1 \) in the class of Patterson-Sullivan. In particular, for any \( \varepsilon > 0 \), there is \( \alpha > 0 \) such that

\[
\limsup_{t \to +\infty} \nu_1(\xi : |i \circ r_\xi(t) - te_\nu| \leq \alpha t) \leq \varepsilon
\]

This proves that \( \nu \) and \( \nu_1 \) are mutually singular.

\[\square\]

6 Open questions

We mention a few problems that were not discussed here, but might lead to interesting continuations of our study.

- **(Nilpotent case)** Let \( \Gamma \) be a non-abelian free group, \( N \) a simply connected nilpotent Lie group, and \( \pi : \Gamma \to N \) a group morphism. Let \( \nu \) be a Furstenberg measure on the Gromov boundary \( \partial \Gamma \), determined by a finitely supported non-elementary probability measure \( \mu \) on \( \Gamma \). What are the statistics of \( \nu \)-typical geodesic rays projected to \( N \)? Note that limit theorems on \( N \) for i.i.d random walks are available [55, 15, 23, 38, 4, 5].

- **(Inverse problem)** Let \( \mu_1, \mu_2 \) be two finitely supported non-elementary measures on a non-abelian free group \( \Gamma \). Is there a simple necessary and sufficient condition characterizing that \( \mu_1 \) and \( \mu_2 \) have the same Furstenberg measure?

\[\text{This means the limit of the push-forwards by the geodesic flow } \phi_t \text{ as } t \to +\infty\]
A Appendix : General limit laws for cocycles

We prove general limit theorems for cocycles. Similar results have been studied for the Iwasawa cocycle on linear groups in [9, 7] and the Busemann cocycle on hyperbolic groups [8], but they do not apply directly to the context of the paper. Our motivation is to propose an abstract setting, encapsulating both [9, 7, 8] and our needs, and which will hopefully apply to other situations.

We fix a compact metric space $X$ and a measurable group $G$ acting measurably on $X$.

We let $\sigma : \hat{G} \times X \to \mathbb{R}^d$ be a measurable cocycle and $|.| : G \to \mathbb{R}_{\geq 0}$ a subadditive measurable function such that for all $g \in G$,

$$\sup_{x \in X} \|\sigma(g, x)\| \leq |g|$$

We also fix a probability measure $\mu$ on $G$ such that $\int_G |g| d\mu(g) < \infty$.

Throughout Sections A.1-A.4, we assume that $\sigma$ has constant zero-drift for $\mu$:

$$\int_G \sigma(g, x) d\mu(g) = 0$$

Under suitable moment and regularity conditions, we establish a central limit theorem with stopping time, the law of the iterated logarithm, the large deviation principle and a gambler’s ruin estimate for the cocycle $\sigma$.

In Section A.5, the zero mean condition on $\sigma$ is replaced by assuming a uniform lower bound on the drift: for all $x \in X$,

$$\int_G \sigma(g, x) d\mu(g) \geq 1$$

and we study the first time for which $\sigma$ leaves a bounded interval under the $\mu$-random walk, as well as the overshoot.

Remarks. 1) If $\sigma(g,.)$ is continuous for every $g \in G$ (as will be assumed in most of this appendix), then one can always choose the function $|.|$ to be $|g| = \sup_{x \in X} \|\sigma(g, x)\|$. It is indeed measurable because the supremum over $X$ can be replaced by the supremum over a countable dense subset of $X$, and subadditivity follows from the cocycle relation.

2) Although formulated for $\sigma$ with constant zero drift, Sections A.1-A.4 yield limit theorems for cocycles $\sigma$ of the form

$$\sigma(g, x) = \sigma_\lambda(g, x) + \psi(gx) - \psi(x)$$

where $\sigma$ has constant drift $\lambda$ with $\lambda \in \mathbb{R}^d$ possibly non-zero, and $\psi : X \to \mathbb{R}^d$ is measurable and bounded. Indeed, notice that limit theorems for $\sigma_\lambda$ follow from the centered case by setting $G' = G \times \mathbb{Z}$ acting on $X$ via the $G$ component, $\sigma'((g, k), x) = \sigma_\lambda(g, x) - k\lambda$, and $\mu' = \mu \otimes \delta_1$. It is also obvious that a fixed bounded function $\psi$ does not play any role for Sections A.1-A.3, where the cocycle is renormalized by a
term that goes to $+\infty$. For A.4, one can argue as in the proof of GR in Theorem 1.4 to erase the influence of $\psi$.

We set $(B, \beta) = (G^\mathbb{N}, \mu^\otimes\mathbb{N})$, and write $B_k$ the $\sigma$-algebra of $B$ generated by the first $k$ coordinates. We say that $\mu$ has finite $p$-th moment if $\int_G |g|^p d\mu(g) < \infty$.

A.1 Central limit theorem

We introduce a notion of continuity on $\sigma$ with respect to the space variable in $X$. More precisely, we say $\sigma$ is $(\mu, X)$-controlled if

- for every $g \in G$, the function $\sigma(g, \cdot)$ is continuous.
- $\forall \varepsilon > 0, \exists R > 0, \forall n \geq 1 \ \forall x, y \in X$,
  \[
  \mu^sn(g \in G : |\sigma(g, x) - \sigma(g, y)| \leq R) \geq 1 - \varepsilon
  \]

Under this condition, we prove the central limit theorem for $\sigma$.

**Proposition A.1.** Assume $\mu$ has finite second moment, and $\sigma$ has constant zero drift and is $(\mu, X)$-controlled. Then,

1. (Covariance) Uniformly in $x \in X$, we have the convergence
   \[
   \frac{1}{n} \int_G \sigma(g, x)^t \sigma(g, x) d\mu^sn(g) \longrightarrow A
   \]
   where $A \in M_d(\mathbb{R})$ is some positive semi-definite symmetric matrix.

2. (CLT) Uniformly in $x \in X$, as $g$ varies with law $\mu^sn$,
   \[
   \frac{\sigma(g, x)}{\sqrt{n}} \xrightarrow{\operatorname{dist}}_{n \to \infty} \mathcal{N}(0, A)
   \]

**Remark.** 1) We see from the proof that the limit covariance matrix $A$ satisfies

\[
A = \frac{1}{n} \int_{G \times X} \sigma(g, x)^t \sigma(g, x) d\mu^sn(g) d\nu(x)
\]

for any $n \geq 1$, any $\mu$-stationary probability measure $\nu$ on $X$. One says that $\sigma$ has unique covariance.

2) Benoist and Quint prove a CLT where the $(\mu, X)$-control assumption is replaced by the condition that $\sigma$ is continuous with unique covariance [7, Theorem 3.4]. The $(\mu, X)$-control assumption seems to be more convenient. For instance, it applies directly to both the linear case and the hyperbolic case, whereas establishing unique covariance in the hyperbolic setting requires to go through the whole proof of the CLT, as in [8, Theorem 4.7].

For this paper, we actually need a stronger version of the CLT, dealing not only with $\mu^sn$ but with $\mu^*n$ where $\tau_n$ is a stopping time distributed around $n$ at scale $\sqrt{n}$. Let us define this notation.
Given $\tau : B \to \mathbb{N} \cup \{\infty\}$ be a $\beta$-a.e. finite stopping time, we define $\mu^{*\tau}$ as the distribution of the variable $b_{\tau(b)} \ldots b_1$ when $b$ varies with law $\beta$.

We say that a family $(\tau_n)_{n \geq 1}$ of stopping times is tight around $n$ if for all $\varepsilon > 0$, there exists $R > 0$ such that

$$\inf_{n \geq 1} \beta\{\tau_n \in [n - R\sqrt{n}, n + R\sqrt{n}]\} \geq 1 - \varepsilon$$

Proposition A.2 (CLT with stopping times). Assume $\mu$ has finite second moment, and $\sigma$ has constant zero drift and is $(\mu, X)$-controlled. Let $(\tau_n)_{n \geq 1}$ be a family of stopping times which is tight around $n$. Then uniformly in $x \in X$, as $g$ varies with law $\mu^{*\tau_n}$,

$$\frac{\sigma(g, x)}{\sqrt{n}} \xrightarrow{\text{dist}}_{n \to +\infty} \mathcal{N}(0, A)$$

We now engage in the proof of the CLT with no stopping time, i.e. Proposition A.1. We will combine methods used in [8] and [9]. Using Cramér-Wold device and the $(\mu, X)$-control assumption (for the uniformity), it is enough to check those statements when $\sigma$ is replaced by its image under a linear form. In other words, we may assume that $d = 1$. We start by a non-uniform version of the claims.

Lemma A.3. There exists $x_0 \in X$ and $A \in \mathbb{R}^+$ such that the following holds.

$$\frac{1}{n} \int_G \sigma(g, x_0)^2 \mu^{*n}(g) \longrightarrow A$$

and as $g$ varies with law $\mu^{*n}$,

$$\frac{\sigma(g, x_0)}{\sqrt{n}} \xrightarrow{\text{dist}}_{n \to +\infty} \mathcal{N}(0, A)$$

Proof. Let $x_0 \in X$ be a parameter and for $k \geq 1$, $b \in B$, introduce the martingale difference

$$X_k(b) = \sigma(b_k, b_{k-1} \ldots b_1 x_0)$$

We only need check that for some good choice of $x_0$, there exists $A \geq 0$ such that

(i) $W_n := \frac{1}{n} \sum_{k=1}^n \mathbb{E}(X_k^2 | B_{k-1}) \longrightarrow A$ in $L^1$

(ii) For all $\varepsilon > 0$, $W_{\varepsilon,n} := \frac{1}{n} \sum_{k=1}^n \mathbb{E}(X_k^2 1_{\{|X_k| > \varepsilon\}} | B_{k-1}) \longrightarrow 0$ in probability

Indeed, this is sufficient: Brown’s CLT for martingales [16] yields the convergence $\frac{1}{\sqrt{n}} \sum X_i \to \mathcal{N}(0, A)$ which is (10), and integrating item (i) yields (9) (use the zero drift condition and the cocycle relation).

Let us check (i). We observe that $W_n(b) = \frac{1}{n} \sum_{1 \leq k \leq n} M(b_{k-1} \ldots b_1 x_0)$ where $M : X \to \mathbb{R}^+$ is the bounded function defined by

$$M(y) = \int_G \sigma(g, y)^2 d\mu(g)$$

$$(y \in X)$$
Choose $\nu$ to be any ergodic $\mu$-stationary probability measure on $X$. In other terms, the dynamical system $(B \times X, \beta \otimes \nu, T^X)$ where $T^X : (b_1, x) \mapsto ((b_i)_{i \geq 1}, b_1 x)$ is measure preserving ergodic [9, Proposition 2.9]. By Birkhoff Ergodic Theorem, for all $x$ in a set $E$ of $\nu$-measure 1, $\beta$-almost every $b$, one has $W_n(b) \to \nu(M)$ where $\nu(M)$ is the $\nu$-average of $M$, that is

$$\nu(M) = \int_{G \times X} \sigma(g, x)^2 d\mu(g) d\nu(x) =: A$$

To get (i), choose $x_0$ in $E$, and note that almost sure convergence implies convergence in $L^1(B, \beta)$ for sequence of functions that are uniformly bounded.

For item (ii), observe the equality $W_{\epsilon, n}(b) = \frac{1}{n} \sum_{1 \leq k \leq n} u_n(b_{k-1} \ldots b_1 x_0)$ where

$$u_n(y) = \int_G \sigma(g, y)^2 1_{\{||\sigma(g, y)|| > \epsilon/\sqrt{n}\}} d\mu(g)$$

$$\leq \int_G |g|^2 1_{\{|g| > \epsilon/\sqrt{n}\}} d\mu(g)$$

By the $L^2$ assumption on $\mu$, we have $u_n(y) \to 0$ uniformly in $y$. Hence $W_{\epsilon, n} \to 0$. \qed

We now upgrade Lemma [A.3] to Proposition [A.4].

**Proof of Proposition [A.4].** Let $x_0 \in X$ and $A$ be as in Lemma [A.3]. The uniform CLT follows directly from (10) and the $(\mu, X)$-control assumption on $\sigma$.

To check the convergence of variances, we must prove that

$$\frac{1}{n} \int_G \sigma(g, x)^2 d\mu^\ast_n(g) \to A$$

uniformly in $x \in X$. Using the cocycle relation and the zero drift assumption, we can rewrite the left-hand side

$$\frac{1}{n} \int_G \sigma(g, x)^2 d\mu^\ast_n(g) = \frac{1}{n} \sum_{k=0}^{n-1} \mu^* k \ast \delta_x \left(\kappa_n, x\right) (M)$$

where $M$ is the function defined previously. To conclude, we need to check the following lemma.

**Lemma A.4.** The sequence of empirical averages $(\kappa_n, x(M))_{n \geq 1}$ converges to $A$ uniformly in $x \in X$.

**Proof of Lemma A.4.** We first observe that for any $\mu$-stationary probability measure $\nu'$ on $X$, the $\nu'$-average of $M$ is equal to $A$. Indeed, one can assume $\nu'$ ergodic. The proof of Lemma [A.3] shows that $\nu'(M)$ is the limit covariance in the CLT for $\sigma(., x'_0)$ where $x'_0$ is typical for $\nu'$. But we also note that the $(\mu, X)$-control assumption on $\sigma$ implies that a CLT for $\sigma(., x)$ implies the same CLT for $\sigma(., y)$ where $y \in X$ is any other spatial coordinate. Hence, the limit covariance is independent of $\nu'$. This justifies that $\nu'(M) = A$ for all stationary probability measure $\nu'$ on $X$. 31
To conclude, we now assume by contradiction that $\kappa_{\nu,x}(M)$ does not converge to $A$ uniformly in $x \in X$, i.e. we have some sequence $\kappa_{\nu,x}(M)$ at distance $\varepsilon > 0$ from $A$. But $M$ is bounded continuous and any accumulation point of $\kappa_{\nu,x}$ is stationary, whence a contradiction with the first paragraph.

This concludes the proof of Proposition A.1.

We now prove the CLT with stopping times.

Proof of Proposition A.2. Let $R > 0$ be some (large) fixed parameter and set, $q_n = [n + R\sqrt{n}]$, then $\tau'_n = \min(\tau_n, q_n)$. By the tightness assumption, $\tau'_n = \tau_n$ on a set of $\beta$-measure arbitrarily close to 1 as long as $R$ is large enough. Hence, it is sufficient to check the above convergence for the stopping times $\tau'_n$, or in other words, we may assume $\tau_n \leq q_n$. Now we can use the cocycle property to write

$$\frac{\sigma(b_{\tau_n(b)} \ldots b_1, x)}{\sqrt{n}} = \frac{\sigma(b_{q_n} \ldots b_1, x)}{\sqrt{n}} - \frac{\sigma(b_{q_n} \ldots b_{\tau_n(b) + 1}, b_{\tau_n(b)} \ldots b_1 x)}{\sqrt{n}}$$

By the CLT in Proposition A.1, the first term on the right-hand side converges in law to $N_{p_0, Aq}$, uniformly in $x \in X$. We only need to check that the second term goes to 0 in probability and uniformly in $x$. Denote by $T : B \rightarrow B, (b_i)_{i \geq 1} \mapsto (b_{i+1})_{i \geq 1}$ the one-sided shift. Observing that $T^\tau_n b$ has law $\beta$, we just need to check that as $a, b$ vary independently in $B$ with law $\beta$, we have the convergence in probability

$$\frac{\sigma(a_{q_n - \tau_n(b)} \ldots a_1, y)}{\sqrt{n}} \rightarrow 0$$

uniformly in $y \in X$. This follows from the tightness condition on the $\tau_n$ and the uniformity of the CLT in Proposition A.1.

A.2 Law of the iterated logarithm

We prove the law of the iterated logarithm for cocycles.

Proposition A.5. Assume $\mu$ has finite second moment, $\sigma$ has constant zero drift and is $(\mu, X)$-controlled. Let $A$ be the limit covariance matrix in the CLT (A.1). Then, for every $x \in X$, for $\beta$-almost every $b \in B$,

$$\left\{ \text{Accumulation points of } \left( \frac{\sigma(b_n \ldots b_1, x)}{\sqrt{2n \log \log n}} \right)_{n \geq 3} \right\} = A^{1/2}(B(0, 1))$$

where $B(0, 1) = \{ t \in \mathbb{R}^d, \sum_i t_i^2 \leq 1 \}$.

A similar LIL is presented in the book [9] under the assumption $\mu$ has finite exponential moment. To reach the optimal generality of $\mu$ having only finite second moment, we build on truncation techniques used by de Acosta [20] in his proof of the LIL for sum of i.i.d random variables.
The upper bound

We now prove the first half of Proposition A.5: the set of accumulation points of the renormalized cocycle is included in $A^{1/2}(B(0,1))$. For this purpose, we may reduce the problem to dimension 1.

**Lemma A.6.** We may assume without loss of generality that $d = 1$ and $A = 1$.

**Proof.** Fix $x \in X$. If the case $(d = 1, A = 1)$ holds, then for every linear form $\varphi$ on $\mathbb{R}^d$, for $\beta$-almost every $b \in B$, the accumulation points of $(a_n^{-1} \varphi \circ \sigma(b_n \ldots b_1, x))_{n \geq 3}$ are included in $\varphi A^{1/2}(B(0,1))$. Now use that the inclusion in $A^{1/2}(B(0,1))$ can be characterized via a countable set of linear forms on $\mathbb{R}^d$ to conclude. \hfill \square

From now on, we write $a_n = \sqrt{2n \log \log n}$.

**Lemma A.7.** Fix $r > 1$ arbitrary. It is enough to show that for all $x \in X$, for $\beta$-almost every $b \in B$,

$$\limsup \frac{1}{a_n} \sigma(b_n \ldots b_1, x) \leq r^3$$  \hfill (11)

**Proof.** The upper bound in the LIL follows by applying (11) to $\sigma$ and $-\sigma$, and by letting $r \to 1^+$ along a countable sequence. \hfill \square

**Lemma A.8.** Let $\delta > 0$. To prove (11), we may assume

$$\int_G \sigma(g, x)^2 d\mu(g) \leq 1 + \delta$$  \hfill (12)

**Proof.** Notice that for fixed $k \geq 1$, and $\beta$-almost every $b \in B$, the sequences

$$a_n^{-1} \sigma(b_n \ldots b_1, x), \quad a_n^{-1} \sigma(b_n \ldots b_1, x)$$

have the same accumulation points. Hence, we can replace $(\sigma, \mu)$ by $(\frac{1}{\sqrt{k}} \sigma, \mu^k)$ without loss of generality. By the assumption $A = 1$, and the uniform convergence of variances in Proposition A.1 there exists $k \geq 1$ such that for all $x \in X$,

$$\int_G \frac{1}{k} \sigma(g, x)^2 d\mu^*(g) \leq 1 + \delta$$

whence the result. \hfill \square

We now introduce truncations of $\sigma$. These are crucial to deal with the weak moment condition on $\sigma$.

Let $\alpha > 0$ be some parameter to be specified later. For $j \geq 1$, set $\text{LL}j = \log \log j$. We define $X_1^{(j)}$ as the $\alpha \text{LL}j$-truncation of $\sigma$, and call $X^{(j)}$ the corresponding recentered variable:

$$X_1^{(j)}(g, x) = \sigma(g, x)1_{|g| \leq \alpha \text{LL}j} \quad X^{(j)}(g, x) = X_1^{(j)}(g, x) - E_x(X_1^{(j)})$$
where $E_x$ refers to the $\mu \otimes \delta_x$-average.

For $(b, x) \in B \times X$, we set

$$S_n(b, x) = \sum_{j=1}^{n} X^{(j)}(b_j, b_{j-1} \ldots b_1, x)$$

$S_n(b, x)$ corresponds to a version of $\sigma(b_n \ldots b_1, x)$ where the $n$-increments arising from the cocycle relation have been truncated with increasing margin of error and (slightly) recentered. We first note that $S_n(b, x)$ is a good approximation of $\sigma(b_n \ldots b_1, x)$, at least for the purpose of proving the LIL:

**Lemma A.9.** For every $\alpha > 0$, for all $x \in X$, $\beta$-almost every $b \in B$,

$$\sigma(b_n \ldots b_1, x) = S_n(b, x) + o(a_n) \quad (13)$$

**Proof.** Using the inequality $|\sigma(g, x)| \leq |g|$ and the zero drift assumption, we can bound the difference

$$|\sigma(b_n \ldots b_1, x) - S_n(b, x)| \leq \sum_{j=1}^{n} |b_j|1_{|b_j| > \alpha \sqrt{\frac{1}{j}}} + \sum_{j=1}^{n} E_x(|g|1_{|g| > \alpha \sqrt{\frac{1}{j}}}) \quad (14)$$

The second moment assumption and [20, Lemma 2.3] imply that the right-hand side is $\beta$-almost surely of the form $o(a_n)$. \qed

We now play on the variance parameter $\delta$ and the truncation parameter $\alpha$ to prove the upper bound (11).

**Lemma A.10.** We can choose the parameters $\delta, \alpha > 0$ so that for some $c > 1$, all $n \geq 1$,

$$\sup_{x \in X} \mathbb{P}_x(S_n \geq r a_n) = O(\exp(-c LLn))$$

**Proof.** Using the relation $e^s \leq 1 + s + \frac{s^2}{2} e^{|s|}$ for any $s \in \mathbb{R}$ and the fact that $X^{(j)}$ has zero average, we get for $t > 0$, $x \in X$,

$$E_x(e^{tX^{(j)}}) \leq 1 + \frac{t^2}{2} E_x(|X^{(j)}|^2) \exp(t\alpha \frac{j}{\sqrt{LL}j})$$

The relation (12) allows to bound the second moment of $X^{(j)}$: for $j$ large enough (say $j \geq j_0$),

$$E_x(|X^{(j)}|^2) \leq \int_G \sigma(g, x)^2 d\mu(g) + \delta \leq 1 + 2\delta := a^2$$

hence

$$E_x(e^{tX^{(j)}}) \leq \exp\left(\frac{t^2a^2}{2} \exp(t\alpha \frac{j}{\sqrt{LL}j})\right)$$

We can now bound the exponential moments of $S_n$. For $n \geq 1$, $t > 0$, using that

$$E_x(e^{tX^{(j)}(b_j, b_{j-1} \ldots b_1, x)} | B_{j-1}) \leq \sup_{y \in X} E_y(e^{tX^{(j)}}),$$

34
we get

\[ \mathbb{E}_x(e^{tS_n}) \leq \prod_{j=1}^{n} \sup_{y \in X} \mathbb{E}_y(e^{tX(j)}) \]
\[ \leq \exp \left( \frac{t^2 a^2}{2} \sum_{j=1}^{n} \exp(t\alpha \frac{j}{\sqrt{Ll_j}}) \right) \]

So by Chebychev inequality

\[ \mathbb{P}_x(S_n > r a_n) \leq \exp(-t r a_n + \frac{t^2 a^2}{2} \sum_{j=1}^{n} \exp(t\alpha \frac{j}{\sqrt{Ll_j}})) \]

Setting \( t = ra^{-2}\sqrt{2(\text{LLn})/n} \), a direct computation yields

\[ \mathbb{P}_x(S_n > r a_n) \leq \exp(-(r/a)^2(2 - e^{ra^{-2}\alpha})\text{LLn}) \]

As \( \delta, \alpha \to 0^+ \), we have \( c := (r/a)^2(2 - e^{ra^{-2}\alpha}) \to r^2 > 1 \). This concludes the proof.

The rest of the proof is dedicated to showing that for \( \delta, \alpha \) as in Lemma A.10, we have the upper bound (11).

**Lemma A.11.** For all \( \varepsilon > 0 \), there exists \( n_0 \geq 0 \) such that for \( n \geq n_0 \), \( x \in X \),

\[ \min_{0 \leq k \leq n} \mathbb{P}_x(|S_n - S_k| \leq \varepsilon a_n) \geq 1/2 \]

**Proof.** Call \( r_n(b) \) the right-hand side of (14). We have

\[ |S_n - S_k| \leq |\sigma(b_n \ldots b_{k+1}, b_k \ldots b_1x)| + 2r_n(b) \]

As \( r_n(b) = o(a_n) \) almost-surely, we only need to check that for large \( n \), for all \( y \in X \),

\[ \min_{0 \leq l \leq n} \beta(|\sigma(b_l \ldots b_1, y)| \leq \frac{\varepsilon}{2} a_n) \geq 2/3 \]

For \( l \) greater than some \( n'_0 \), this follows from the CLT assumption. For \( l \leq n'_0 \), this follows from the \( L^2 \)-domination assumption.

Set \( S_n^* = \max_{k \leq n} S_k \).

**Lemma A.12.** For all \( c, \varepsilon > 0 \), for \( n \geq n_0 \), \( x \in X \),

\[ \mathbb{P}_x(S_n^* \geq (c + \varepsilon)a_n) \leq 2\mathbb{P}_x(|S_n| \geq ca_n) \]

**Proof.** It is enough to check that

\[ \mathbb{P}_x(S_n \geq ca_n \mid S_n^* \geq (c + \varepsilon)a_n) \geq \frac{1}{2} \]

This follows from Lemma A.11.
Lemma A.13. Choose $\delta, \alpha > 0$ as in Lemma [A.10]. For all $x \in X$, for $\beta$-almost every $b \in B$,
\[
\limsup \frac{1}{a_n} S_n(b, x) \leq r^3
\]

Proof. Fix some $s \in (1, r)$ and set $k_n = \lceil s^n \rceil$. Writing i.o. for infinitely often, we have
\[
\mathbb{P}_x(S_n \geq r^3 a_n \ i.o.) \leq \mathbb{P}_x(S_n^* \geq r^3 a_{k_n-1} \ i.o.) \leq \mathbb{P}_x(S_n^* \geq r^2 a_{k_n} \ i.o.)
\]
where the second inequality uses that $rk_{n-1} \geq k_n$ for $n$ large. Let us check that the right-hand side is 0. Using Lemma [A.12] and Borel Cantelli Lemma, we only need to check that
\[
\sum \mathbb{P}_x(S_{k_n} \geq ra_{k_n}) < \infty
\]
By Lemma [A.12]
\[
\mathbb{P}_x(S_{k_n} \geq ra_{k_n}) \leq \exp(-c\LL k_n) \ll n^{-c}
\]
whence the finitess. This shows that
\[
\mathbb{P}_x(S_n \geq r^3 a_n \ i.o.) = 0
\]

Conclusion of the proof. We only need to check that Equation (11) is satisfied. This follows from Lemma [A.13] and Lemma [A.9].

The lower bound

The proof of the reverse inclusion in the LIL follows verbatim the lines of the book [9, Section 12.4] given in the context of $\mu$-contracting cocycles with finite exponential moments. We choose not to repeat it.

A.3 Principle of large deviations

We show the principle of large deviations for cocycles. The proof is classic but also very short, so we include it for completeness.

Proposition A.14. Assume $\mu$ has finite exponential moment: $\int_{G} e^{\alpha|g|} d\mu(g) < \infty$ for some $\alpha > 0$, and $\sigma$ has constant zero drift. Then for all $\varepsilon > 0$ there exists $D, \delta > 0$ such that for any $n \geq 1$, $x \in X$,
\[
\beta(\cdot : \|\sigma(b_n \ldots b_1, x)\| \geq \varepsilon n) \leq De^{-\delta n}
\]

Proof. We can assume that $\sigma$ is $\mathbb{R}$-valued. Let $t \in (0, \alpha)$ be a parameter, $n \geq 1$, $x \in X$. Using the cocycle relation and Fubini Theorem,
\[
\beta(b : \sigma(b_n \ldots b_1, x) > \varepsilon n) \leq e^{-\varepsilon n} \int_{B} e^{t\sigma(b_n \ldots b_1, x)} d\beta(b) \leq e^{-\varepsilon n} \left( \sup_{y \in X} \int_{G} e^{t\sigma(g, y)} d\mu(g) \right)^n
\]
Using the bound $e^s \leq 1 + s + s^2 e^{|s|}$ for any $s \in \mathbb{R}$, we get for $y \in X$,
\[
\int_G e^{t \sigma(g,y)} d\mu(g) \leq 1 + t^2 C \leq \exp(t^2 C)
\]
where $C = \mathbb{E}_\mu(|g|^2 e^{t|g|})$ is finite. Combining inequalities, we get
\[
\beta(b : \sigma(b_n \ldots b_1, x) > \varepsilon n) \leq \exp(- (\varepsilon - tC)tn)
\]
Choosing $t < \varepsilon/C$, we have an exponential bound independent of $x$. Applying the argument to $-\sigma$, we get the desired bound for $|\sigma(b_n \ldots b_1, x)|$.

A.4 Gambler’s ruin

We show gambler’s ruin estimates for cocycles.

**Proposition A.15.** Assume $\mu$ has finite exponential moment, $\sigma$ has constant zero drift and is $(\mu,X)$-controlled, $d = 1$, and the limit variance $A$ in the CLT (A.1) is non-zero. Fix $k,l > 0$ and sequences $(k_s)_{s \geq 0}, (l_s)_{s \geq 0} \in \mathbb{R}^\mathbb{N}$ such that $k_s \sim sk$, $l_s \sim sl$. Then for every $x \in X$,
\[
\beta \{ b : (\sigma(b_n \ldots b_1, x))_{n \geq 1} \text{ meets } [l_s, +\infty) \text{ before } (-\infty, -k_s) \} \xrightarrow{s \to +\infty} \frac{k}{k + l}
\]

For the proof, we fix $x \in X$ and write $M_n(b) = \sigma(b_n \ldots b_1, x)$ the associated martingale.

We will need to know that $M_n$ does leave the window $(-k_s, l_s)$, and the exit time is bounded polynomially in $s$ with high probability.

**Lemma A.16.**
\[
\beta \{ b : (M_n(b))_{n \geq 1} \text{ leaves } (-k_s, l_s) \text{ for some } n \leq s^3 \} \xrightarrow{s \to +\infty} 1
\]

**Proof.**
\[
\beta \{ b : M_{[s^3]}(b) \in (-k_s, l_s) \} = \beta \{ b : s^{-3/2} M_{[s^3]}(b) \in (-s^{-3/2}k_s, s^{-3/2}l_s) \} \xrightarrow{s \to +\infty} 0
\]

by the CLT for $M_n$ proven in Proposition [A.1] □

**Proof of Proposition [A.15].** Set
\[
\tau_s := \inf \{ n \geq 1, (M_n(b))_{n \geq 1} \text{ meets } [l_s, +\infty) \text{ or } (-\infty, -k_s) \}
\]
and $\tau_s := \min([s^3], \tau_s')$. By the optional sampling theorem applied to the zero-mean martingale $M_n$, we have
\[
0 = \mathbb{E}(M_{\tau_s}) = \mathbb{E}(M_{\tau_s} 1_{M_{\tau_s} \geq l_s}) + \mathbb{E}(M_{\tau_s} 1_{M_{\tau_s} \leq -k_s}) + \mathbb{E}(M_{\tau_s} 1_{M_{\tau_s} \in (-k_s, l_s)})
\]
Let us estimate each term individually.

\[ \mathbb{E}(M_{\tau_s}1_{M_{\tau_s} \geq l_s}) = l_s\beta(M_{\tau_s} \geq l_s) + \mathbb{E}((M_{\tau_s} - l_s)1_{M_{\tau_s} \geq l_s}) \]

We show the overshoot is negligible compared to \( s \). We first notice the upper bound

\[ \mathbb{E}((M_{\tau_s} - l_s)1_{M_{\tau_s} \geq l_s}) \leq \mathbb{E}(\sup_{k \leq s^3} |M_{k+1} - M_k|) \]

\[ = \int_{\mathbb{R}^+} \beta(\sup_{k \leq s^3} |M_{k+1} - M_k| > t) \, dt \]

\[ = \int_{\mathbb{R}^+} 1 - \beta(\sup_{k \leq s^3} |M_{k+1} - M_k| \leq t) \, dt \]

\[ = \int_{\mathbb{R}^+} 1 - \beta(E_{\alpha^3}(t)) \, dt \]  

(15)

where for \( r > 0 \), we write \( E_r(t) := \{ b \in B : \sup_{k \leq r} |M_{k+1}(b) - M_k(b)| \leq t \} \). Arguing conditionally and observing that the cocycle relation gives \( M_{k+1}(b) - M_k(b) = \sigma(b_{k+1}, b_k \ldots b_1 x) \), we get

\[ \beta(E_{\alpha^3}(t)) = \beta(|M_{s^3+1} - M_{s^3}|) \leq t \mid E_{\alpha^3-1}(t))\beta(E_{\alpha^3-1}(t)) \geq (1 - e^{-\alpha t})\beta(E_{\alpha^3-1}(t)) \]

for some \( \alpha = \alpha(\mu, |.|) \) coming from the exponential moment assumption on \( \mu \). Iterating this inequality and plugging it into (15), we get

\[ \mathbb{E}((M_{\tau_s} - l_s)1_{M_{\tau_s} \geq l_s}) \leq \int_{\mathbb{R}^+} 1 - (1 - e^{-\alpha t})^{s^3} \, dt \]

Lemma A.17 below tells us that \( \sup_{s > 0} \int_{[s, +\infty)} 1 - (1 - e^{-\alpha t})^{s^3} \, dt < \infty \) for any \( \varepsilon > 0 \), whence we can finally control the overshoot by:

\[ \mathbb{E}((M_{\tau_s} - l_s)1_{M_{\tau_s} \geq l_s}) = o(s) \]

To sum up, we proved that

\[ \mathbb{E}(M_{\tau_s}1_{M_{\tau_s} \geq l_s}) = l_s\beta(M_{\tau_s} \geq l_s) + o(s) \]

Using either the same line of argument, or Lemma A.16, we get as well that

\[ \mathbb{E}(M_{\tau_s}1_{M_{\tau_s} \leq -l_s}) = -k_s\beta(M_{\tau_s} \leq -l_s) + o(s) \]

\[ |\mathbb{E}(M_{\tau_s}1_{M_{\tau_s} \in (-l_s, l_s)})| \leq \max(k_s, l_s)\beta(M_{\tau_s} \in (-l_s, l_s)) = o(s) \]

In conclusion,

\[ 0 = \mathbb{E}(M_{\tau_s}) \]

\[ = l_s\beta(M_{\tau_s} \geq l_s) - k_s\beta(M_{\tau_s} \leq -l_s) + o(s) \]

so as \( s \to +\infty \), one has \( \beta(M_{\tau_s} \geq l_s) \to \frac{k}{k+l} \) as desired. The result for \( \tau_s' \) follows because \( \beta(\tau_s \neq \tau_s') \to 0 \) as \( s \to +\infty \) by Lemma A.16. \( \square \)

We record the following estimate used in the proof of Proposition A.15 above.
Lemma A.17. For all \( \varepsilon > 0 \), one has
\[
\sup_{s > 0} \int_{[\varepsilon s, +\infty]} 1 - (1 - e^{-at})^{s_3} \, dt < \infty.
\]

Proof. The integral over the larger domain \( \int_{\mathbb{R}_+} 1 - (1 - e^{-at})^{s_3} \, dt \) is finite for every \( s > 0 \) and increases with \( s \). Hence, we only need to check the bound for \( s \) large enough. Now rewrite uniformly for \( t \in [\varepsilon s, +\infty) \),
\[
(1 - e^{-at})^{s_3} = \exp(s_3 \log(1 - e^{-at})) = \exp(s_3(e^{-at} + o(e^{-at}))) = 1 + s_3 e^{-at} + o(s_3 e^{-at})
\]
where \( o(.) \) goes to zero as \( s \to +\infty \). The result then follows because
\[
\sup_{s > 0} \int_{[\varepsilon s, +\infty]} s_3 e^{-at} \, dt < \infty.
\]
\end{proof}

A.5 Exit estimates for drifting cocycles

For this last section, we replace the zero mean condition on \( \sigma \) by assuming a uniform lower bound on the drift: for all \( x \in X \),
\[
\int_G \sigma(g, x) d\mu(g) \geq 1 \tag{16}
\]
For \( b \in B, x \in X \), we set
\[
S_n(b, x) = \sigma(b_n \ldots b_1, x)
\]
and for \( s > 0 \), we write
\[
\tau_s(b, x) = \inf\{n \geq 1, S_n(b, x) \geq s\}
\]
Foster’s criterion (see [26] or [6, Lemma 2.4]) implies that for any \( x \), the exit time \( \tau_s \) is finite, and even has bounded expectation \( \mathbb{E}_x(\tau_s) \leq s \). The goal of the section is to show that the overshoot \( S_{\tau_s} - s \) is essentially bounded. Assuming equality in (16), we also prove that \( \tau_s \) lands roughly in a window of size \( \sqrt{s} \) around \( s \). Throughout the section, given a formula \( I(n) \) representing an inequality involving (among others) a parameter \( n \in \mathbb{N}^* \), we will lighten notation by writing
\[
\mathbb{P}_x(I(\tau_s)) = \beta(b \in B : I(\tau_s(b, x)))
\]

Proposition A.18. 1. (Overshoot control) Assume that \( \sigma \) satisfies (16).
\[
\sup_{x \in X, s > 0} \mathbb{P}_x(S_{\tau_s} > s + R) \rightarrow 0 \quad \text{as } R \to +\infty.
\]
2. (Time control) Assume $\mu$ has finite second moment, and $\sigma$ is $(\mu, X)$-controlled and has constant drift equal to 1. Then we have: $\forall \varepsilon > 0, \exists R > 0,$ 

$$\sup_{x \in X, s \geq 1} \mathbb{P}_x(|\tau_s - s| > R \sqrt{s}) \leq \varepsilon$$

**Remark.** The control of the overshoot implies that, given any large enough interval $I$ in $\mathbb{R}^+$, the $\beta$-probability that the random sequence $(S_n)_{n \geq 0}$ intersects $I$ is close to 1. This can be interpreted as a weak form of renewal estimate for $(S_n)_{n \geq 0}$.

Let us begin with the overshoot estimate. Our first step is to use the cocycle property of $S$ to reduce the overshoot control for $S_{\tau_s}$ to that of a stronger control for $S_{\tau_1}$.

**Lemma A.19.** For $R > 0$, we have

$$\sup_{x \in X, s \geq 1} \mathbb{P}_x(S_{\tau_s} > s + R) \leq \sum_{n \geq 1} \sup_{y \in X} \mathbb{P}_y(S_{\tau_1} > n + R)$$

**Proof.** Fix $R > 0$. Let $n \geq 1$, and assume $H_n$ is a constant such that

$$\sup_{x \in X, s \in (1, n]} \mathbb{P}_x(S_{\tau_s} > s + R) \leq H_n$$

(17)

It is enough to show that

$$\sup_{x \in X, s \in (1, n+1]} \mathbb{P}_x(S_{\tau_s} > s + R) \leq \sup_{y \in X} \mathbb{P}_y(S_{\tau_1} > n + R) + H_n$$

(18)

Let $x \in X$, $s \in (n, n+1]$. In the event that $S_{\tau_s} > s + R$, the sequence $(S_k)_{k \geq 1}$ does not meet the interval $[s, s + R]$, in particular we get

$$\mathbb{P}_x(S_{\tau_s} > s + R) = \mathbb{P}_x(S_{\tau_1} > s + R) + \mathbb{P}_x(S_{\tau_s} > s + R; S_{\tau_1} < s)$$

$$= \mathbb{P}_x(S_{\tau_1} > s + R)$$

$$+ \mathbb{P}_x(S_{\tau_s} > s + R | S_{\tau_1} < s) \mathbb{P}_x(S_{\tau_1} < s)$$

$$\leq \mathbb{P}_x(S_{\tau_1} > n + R) + \mathbb{P}_x(S_{\tau_s} > s + R | S_{\tau_1} < s)$$

Let us bound above the second term. The cocycle property allows to write for $0 \leq k \leq n$

$$S_n(b, x) = S_{n-k}(T^kb, b_k \ldots b_1x) + S_k(b, x)$$

where we recall that $T : B \to B$, $(b_i)_{i \geq 1} \mapsto (b_{i+1})_{i \geq 1}$ is the one-sided shift. Applying this to $k = \tau_1(b, x)$ and $n = \tau_s(b, x)$, we have

$$\mathbb{P}_x(S_{\tau_s} > s + R | S_{\tau_1} < s) = \mathbb{P}_x(S_{\tau_s-\tau_1}(T^{\tau_1}b, b_{\tau_1} \ldots b_1x) > s + R - S_{\tau_1} | S_{\tau_1} < s)$$

$$= \mathbb{P}_x(S_{\tau_s-\tau_1}(T^{\tau_1}b, b_{\tau_1} \ldots b_1x) > s + R - S_{\tau_1} | S_{\tau_1} < s)$$

40
As $S_{\tau_1} \geq 1$, the lower bound $s + R - S_{\tau_1}$ is in the initial segment $[0, n + R]$ so we may apply the hypothesis (17) to get
\[
P_x(S_{\tau_1} > s + R \mid S_{\tau_1} < s) \leq H_n
\]
In sum, we have proven for any $x \in X$, $s \in (n, n + 1]$,
\[
P_x(S_{\tau_1} > s + R) \leq P_x(S_{\tau_1} > n + R) + H_n
\]
and the bound (18) follows, whence the lemma. \hfill \qed

We now need to show that the sum $\sum_{n \geq 1} \sup_{x \in X} P_x(S_{\tau_1} > n + R)$ appearing in Lemma A.19 is small as long as $R$ is large. We first check it is finite.

**Lemma A.20.** Assume that $\sigma$ satisfies (16). Then
\[
\sum_{n \geq 1} \sup_{x \in X} P_x(S_{\tau_1} > n) < \infty
\]

*Proof.* For $x \in X$, $n \geq 1$, using that $|\sigma(g, x)| \leq |g|$ for all $g \in G$ and that $\tau$ is a stopping time, we have
\[
P_x(S_{\tau_1} > n) = \sum_k P_x(S_k > n, \tau_1 = k)
\leq \sum_k P_x(S_k - S_{k-1} > n - 1, \tau_1 = k)
\leq \sum_k P_x(|b_k| > n - 1, \tau_1 = k)
\leq \sum_k P_x(|b_k| > n - 1, \tau_1 \geq k)
= \mu(|g| > n - 1) \sum_k P_x(\tau_1 \geq k)
= \mu(|g| > n - 1) \mathbb{E}_x(\tau_1)
\]
Foster’s criterion (see [26] or [6, Lemma 2.4]) yields the bound $\mathbb{E}_x(\tau_1) \leq 1$. It follows that
\[
\sum_{n \geq 2} \sup_{x \in X} P_x(S_{\tau_1} > n) \leq \mathbb{E}_\mu(|g|)
\]
which finishes the proof. \hfill \qed

We now conclude the proof of the overshoot estimate in Proposition A.18

*Proof of Proposition A.18 (overshoot control).* The estimate for the supremum involving only $s \in (0, 1]$ follows from Lemma A.20. Indeed, for $R > 1$,
\[
\sup_{x \in X, s \in (0, 1]} P_x(S_{\tau_1} > s + R) \leq \sup_{x \in X} P_x(S_{\tau_1} > R) \xrightarrow{R \to +\infty} 0
\]

Now assume $s > 1$. Using Lemma A.19, we just need to check that
\[
\sum_{n \geq 1} \sup_{x \in X} P_x(S_{\tau_1} > n + R) \xrightarrow{R \to +\infty} 0
\]
For $R = 0$, this sum is finite by Lemma A.20, and the result follows by dominated convergence (using again that for $n \geq 0$, $\sup_{x \in X} P_x(S_{\tau_1} > n + R)$ goes to 0 as $R$ goes to $+\infty$ by Lemma A.20).

We now deal with the time estimate in Proposition A.18.

**Proof of Proposition A.18 (time control).** Let $R > 0$ be a parameter to be specified below.

For $s > 1$, set $q_s = \lfloor s + R\sqrt{s} \rfloor$. We first show that for large $R > 0$, we have

$$\sup_{x \in X, s > 1} P_x(\tau_s > q_s) \leq \varepsilon/2$$

Indeed, if $\tau_s(b, x) > q_s$, then by definition $S_{q_s}(b, x) < s$, so $|S_{q_s}(b, x) - q_s| > R\sqrt{s}$. The central limit theorem for $S_n - n$ (Proposition A.1) guarantees this event only occurs on a set of measure at most $\varepsilon/2$ as long as $R$ is large enough (independently of $x$ and $s > 1$), hence the claim.

For $s > 1$, set $q'_s = \lfloor s - R\sqrt{s} \rfloor$. We now show that for large $R > 0$, we have

$$\sup_{x \in X, s > 1} P_x(\tau_s \leq q'_s) \leq \varepsilon/2$$

Using the cocycle property and the Markov property, we see that for every $\alpha \in (0, 1)$, there exists $R' > 0$ such that for any $x \in X$, $s > 1$,

$$P_x(S_{q'_s} > s - R' \mid \tau_s \leq q'_s) \geq 1 - \sup_{y \in X, n \geq 0} P_y(S_n \leq -R') \geq 1 - \alpha$$

where the inequality on the very right is justified by the central limit theorem for $S_n - n$ (Proposition A.1) and the choice of a large enough $R'$. This leads to

$$P_x(\tau_s \leq q'_s) \leq (1 - \alpha)^{-1} P_x(S_{q'_s} > s - R') \leq (1 - \alpha)^{-1} P_x(|S_{q'_s} - q'_s| \geq R\sqrt{s} - R')$$

Choosing $\alpha$ small enough, $R'$ accordingly, and applying the CLT for $S_n - n$ from Proposition A.1 bounds the right-hand side by $\varepsilon/2$ for $R, s$ large enough. We can increase $R$ even more to deal with all $s > 1$. 



**References**

[1] M. Babillot and M. Peigné. Asymptotic laws for geodesic homology on hyperbolic manifolds with cusps. *Bull. Soc. Math. Fr.*, 134(1):119–163, 2006.

[2] T. Bénard. Mesures de radon stationnaires pour une marche aléatoire sur $T^d \times \mathbb{R}$. *Annales de l’Institut Fourier (DOI: 10.5802/aif.3519)*, 2022.

[3] T. Bénard. Some asymptotic properties of random walks on homogeneous spaces. *Journal of Modern Dynamics*, 19:161–186, 2023.

[4] T. Bénard and E. Breuillard. The central limit theorem on nilpotent Lie groups. *To appear in Annals of Probability (arXiv:2302.06024)*.

[5] T. Bénard and E. Breuillard. Local limit theorems for random walks on nilpotent Lie groups. *Preprint arXiv:2304.14551*, 2023.
[6] T. Bénard and N. De Saxcé. Random walks with bounded first moment on finite-volume spaces. *Geom. Funct. Anal.*, 32(4):687–724, 2022.

[7] Y. Benoist and J.-F. Quint. Central limit theorem for linear groups. *Ann. Probab.*, 44(2):1308–1340, 2016.

[8] Y. Benoist and J.-F. Quint. Central limit theorem on hyperbolic groups. *Izv. Math.*, 80(1):3–23, 2016.

[9] Y. Benoist and J. F. Quint. *Random Walks on Reductive groups*. Springer International Publishing, 2016.

[10] Y. Benoist and J.-F. Quint. On the regularity of stationary measures. *Isr. J. Math.*, 226(1):1–14, 2018.

[11] A. Boulanger, P. Mathieu, C. Sert, and A. Sisto. Large deviations for random walks on Gromov-hyperbolic spaces. *Ann. Sci. Éc. Norm. Supér. (4)*, 56(3):885–944, 2023.

[12] B. Brown. Martingale central limit theorem. *Ann. Math. Stat.*, 42(59-66), 1971.

[13] I. Choi. Central limit theorem and geodesic tracking on hyperbolic spaces and Teichmüller spaces. (arXiv:2106.13017), 2021.

[14] C. Connell and R. Muchnik. Harmonicity of quasiconformal measures and Poisson boundaries of hyperbolic spaces. *Geom. Funct. Anal.*, 17(3):707–769, 2007.

[15] D. Constantine, J.-F. Lafont, and D. J. Thompson. The weak specification property for geodesic flows on CAT(−1) spaces. *Groups Geom. Dyn.*, 14(1):297–336, 2020.

[16] V. Gadre, J. Maher, and G. Tiozzo. Word length statistics and Lyapunov exponents for Fuchsian groups with cusps. *New York J. Math.*, 21:511–531, 2015.

[17] I. Gekhtman and G. Tiozzo. Entropy and drift for Gibbs measures on geometrically finite manifolds. *Trans. Am. Math. Soc.*, 373(4):2949–2980, 2020.
[32] E. Ghys and P. de la Harpe, editors. Sur les groupes hyperboliques d’après Mikhael Gromov. (On the hyperbolic groups à la M. Gromov), volume 83 of Prog. Math. Boston, MA: Birkhäuser, 1990.

[33] Y. Guivarc’h. Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire. (Products of random matrices and applications to geometric properties of subgroups of linear groups). Ergodic Theory Dyn. Syst., 10(3):483–512, 1990.

[34] Y. Guivarc’h and Y. Le Jan. Sur l’enroulement du flot géodésique. (On the winding of the geodesic flow). C. R. Acad. Sci., Paris, Sér. I, 311(10):645–648, 1990.

[35] Y. Guivarc’h and Y. Le Jan. Asymptotic winding of the geodesic flow on modular surfaces and continuous fractions. Ann. Sci. Éc. Norm. Supér. (4), 26(1):23–50, 1993.

[36] Y. Guivarc’h and É. Le Page. Spectral gap properties for linear random walks and Pareto’s asymptotics for affine stochastic recursions. Ann. Inst. Henri Poincaré, Probab. Stat., 52(2):503–574, 2016.

[37] M. Hochman and B. Solomyak. On the dimension of Furstenberg measure for $SL_2(\mathbb{R})$ random matrix products. Invent. Math., 210(3):815–875, 2017.

[38] R. Hough. The local limit theorem on nilpotent Lie groups. Probab. Theory Relat. Fields, 174(3-4):761–786, 2019.

[39] V. A. Kaimanovich and V. Le Prince. Matrix random products with singular harmonic measure. Geom. Dedicata, 150:257–279, 2011.

[40] I. Kapovich and N. Benakli. Boundaries of hyperbolic groups. In Combinatorial and geometric group theory. Proceedings of the AMS special session on combinatorial group theory, New York, NY, USA, November 4–5, 2000 and the AMS special session on computational group theory, Hoboken, NJ, USA, April 28–29, 2001, pages 39–93. Providence, RI: American Mathematical Society (AMS), 2002.

[41] A. Katsuda and T. Sunada. Closed orbits in homology classes. Publ. Math., Inst. Hautes Étud. Sci., 71:5–32, 1990.

[42] P. Kosenko. Fundamental inequality for hyperbolic Coxeter and Fuchsian groups equipped with geometric distances. Int. Math. Res. Not., 2021(6):4709–4728, 2021.

[43] P. Kosenko and G. Tiozzo. The fundamental inequality for cocompact fuchsian groups. (arXiv:2012.07417), 2020.

[44] S. P. Lalley. Regenerative representation for one-dimensional Gibbs states. Ann. Probab., 14:1262–1271, 1986.

[45] Y. Le Jan. The central limit theorem for the geodesic flow on noncompact manifolds of constant negative curvature. Duke Math. J., 74(1):159–175, 1994.

[46] F. Ledrappier. Une relation entre entropie, dimension et exposant pour certaines marches aléatoires. C. R. Acad. Sci., Paris, Sér. I, 296:369–372, 1983.

[47] F. Ledrappier and P. Lessa. Exact dimension of Furstenberg measures. (arXiv:2105.11712), 2021.

[48] J. Li. Decrease of Fourier coefficients of stationary measures. Math. Ann., 372(3-4):1189–1238, 2018.

[49] J. Li, F. Naud, and W. Pan. Kleinian Schottky groups, Patterson-Sullivan measures, and Fourier decay. Duke Math. J., 170(4):775–825, 2021.

[50] J. Maher and G. Tiozzo. Random walks on weakly hyperbolic groups. J. Reine Angew. Math., 742:187–239, 2018.

[51] J.-P. Otal and M. Peigné. Variational principle and Kleinian groups. Duke Math. J., 125(1):15–44, 2004.

[52] S. J. Patterson. The limit set of a Fuchsian group. Acta Math., 136:241–273, 1976.

[53] A. Randecker and G. Tiozzo. Cusp excursion in hyperbolic manifolds and singularity of harmonic measure. J. Mod. Dyn., 17:183–211, 2021.

[54] M. Ratner. The central limit theorem for geodesic flows on $n$-dimensional manifolds of negative curvature. Isr. J. Math., 16:181–197, 1974.

[55] A. Raugi. Théorème de la limite centrale sur les groupes nilpotents. Z. Wahrscheinlichkeitstheor. Verw. Geb., 43:149–172, 1978.

[56] T. Roblin. Ergodicity and equidistribution in negative curvature. Mém. Soc. Math. Fr., Nouv. Sér., 95-96, 2003.
[57] Y. G. Sinai. The central limit theorem for geodesic flows on manifolds of constant negative curvature. 
*Sov. Math., Dokl.*, 1:983–987, 1960.

[58] D. Sullivan. The density at infinity of a discrete group of hyperbolic motions. *Publ. Math., Inst. 
Hautes Étud. Sci.*, 50:171–202, 1979.

[59] D. Sullivan. Entropy, hausdorff measures old and new, and limit sets of geometrically finite kleinian 
groups. *Acta Mathematica*, 153:259–277, 1984.

[60] G. Tiozzo. Sublinear deviation between geodesics and sample paths. *Duke Math. J.*, 164(3):511–539, 
2015.

[61] S. Waddington. Large deviation asymptotics for Anosov flows. *Ann. Inst. Henri Poincaré, Anal. Non 
Linéaire*, 13(4):445–484, 1996.

Timothée Bénard
LAGA - INSTITUT GALILÉE, 99 AVENUE JEAN BAPTISTE CLÉMENT, 93430 VILLETANEUSE
Email address: benard@math.univ-paris13.fr