Central limit theorem on hyperbolic groups

Y. Benoist and J.-F. Quint

Abstract. We prove a central limit theorem for random walks with finite variance on Gromov hyperbolic groups.

Keywords: central limit theorem, hyperbolic groups, boundaries, cocycles, martingales, complete convergence, stationary measures.

§ 1. Introduction

1.1. Central limit theorem. Let \((M, d)\) be a metric space and let \(o\) be a point in \(M\). We write \(G := \text{Isom}(M)\) for the isometry group of \(M\) and put \(\kappa(g) := d(go, o)\) for \(g \in G\). Let \(\mu\) be a Borel probability measure on \(G\) with a finite first moment: \(\int_G \kappa(g) d\mu(g) < \infty\). Let \(g_1, \ldots, g_n, \ldots\) be independent random isometries of \(M\) chosen with law \(\mu\). We want to understand the behaviour of the sequence of random variables \(\kappa(g_n \cdot \cdots \cdot g_1)\). It is well known that this sequence satisfies the law of large numbers: there is a constant \(\lambda\), called the escape rate of \(\mu\), such that \(\lim_{n \to \infty} \frac{1}{n} \kappa(g_n \cdots g_1) = \lambda\) almost surely. In this paper we prove under certain additional assumptions that this sequence satisfies the central limit theorem if \(\mu\) has a finite second moment: \(\int_G \kappa(g)^2 d\mu(g) < \infty\). The additional assumptions require the metric space \((M, d)\) to be proper, quasiconvex and Gromov hyperbolic (see Definition 2.1) and the law \(\mu\) to be non-elementary and non-arithmetic (see Definition 3.1).

Theorem 1.1. Let \((M, d)\) be a proper quasiconvex Gromov hyperbolic metric space, let \(o\) be a point of \(M\), let \(\mu\) be a non-elementary non-arithmetic Borel probability measure with a finite second moment on the group \(G = \text{Isom}(M)\), and let \(\lambda\) be the escape rate of \(\mu\). Then the renormalized variables \(\frac{1}{\sqrt{n}} (\kappa(g_n \cdots g_1) - n\lambda)\) converge in law to a non-degenerate Gaussian law.

Here are some important examples of such hyperbolic spaces \(M\):

(i) metric trees;

(ii) Gromov hyperbolic groups \(\Gamma\) with left-invariant distance induced by a generating set \(S \subset \Gamma\);

(iii) universal coverings of compact Riemannian manifolds of negative curvature.

In the case when the finite second moment assumption is replaced by the finite exponential moment assumption \(\int_G e^{\alpha \kappa(g)} d\mu(g) < \infty\) for some \(\alpha_0 > 0\), the central limit theorem (Theorem 1.1) was proved by Björklund [1]. (It is also a generalisation of the earlier central limit theorem for free groups by Sawyer and Steger [2] and Ledrappier [3].)

AMS 2010 Mathematics Subject Classification. 20F67, 37D40, 60F05, 60G50, 60J50.

© 2016 Russian Academy of Sciences (DoM), London Mathematical Society, Turpion Ltd.
Hence the key point of our paper is to remove the finite exponential moment assumption. To do this, we adapt a method introduced in [4] for linear groups. It does not rely on the spectral gap property of transfer operators and, therefore, enables us to prove the central limit theorem without any further assumptions on the boundaries of $M$ or the support of $\mu$. Our central limit theorem seems to be new even for free groups since the law $\mu$ is only assumed to have a finite second moment.

1.2. Strategy. We want to prove the central limit theorem for the random variables $\kappa(g_n \cdots g_1)$. Let $X$ be the Busemann boundary of $M$ (see §2.3). Since the function $\kappa$ on $G$ is closely related to the Busemann cocycle $\sigma: G \times X \to \mathbb{R}$ (see §§2.4, 3.2), it suffices to prove a central limit theorem (Theorem 4.7) for the random variables $\sigma(g_n \cdots g_1, x)$ for every $x \in X$. Adding a suitable coboundary, we replace this cocycle $\sigma$ by another cocycle $\sigma_0$ whose ‘expected increase’ is finite, that is, $\int_G \sigma_0(g, x) d\mu(g) = \lambda$ for all $x$ in $X$. This enables us to use Brown’s classical central limit theorem for martingales (see [5]). The cocycle $\sigma_0$ is given by $\sigma_0(g, x) = \sigma(g, x) - \psi(x) + \psi(gx)$ for a certain bounded function $\psi$ on $X$ (Proposition 4.6). As in [4], we give an explicit formula for $\psi$ in terms of a $\tilde{\mu}$-stationary measure $\nu^*$ on $X$, where $\tilde{\mu}$ is the image of $\mu$ under the map $g \mapsto g^{-1}$. Namely, we put

$$\psi(x) = -2 \int_G (x|y)_o d\nu^*(y), \quad (1.1)$$

where $(x|y)_o$ is the Gromov product on $X$ (see §2.3). The main issue is to check that this integral is finite, that is, the stationary measure $\nu^*$ is log-regular, provided that the second moment of $\mu$ is finite (Proposition 4.2). As in [4], the key point is to prove the complete convergence of the sequence $\frac{1}{n}\kappa(g_n \cdots g_1)$ to the escape rate $\lambda$ (Proposition 4.1). This generalizes the Hsu–Robbins theorem.

Compared to the linear case in [4], new technical difficulties occur here because the Busemann coboundary $\sigma$ is a cocycle on the Busemann boundary $X$, whereas the behaviour of the random walk is much easier to describe on the Gromov boundary $\partial M$ (see Proposition 3.2). This forces us to change our point of view frequently, working sometimes with the Busemann boundary and sometimes with the Gromov boundary.

1.3. Plan. In §2 we recall without proof the basic definitions and properties of Gromov hyperbolic spaces and their boundaries.

In §3 we recall with short proofs the basic results concerning random walks on Gromov hyperbolic groups.

In §4 we prove the complete convergence to the escape rate, the log-regularity of the stationary measure on the Gromov boundary, the centerability of the Busemann cocycle on the Busemann boundary, and the central limit theorem (Theorem 4.7).

In §5 we prove an optimal version (Proposition 5.1 and Example 5.4) of the log-regularity of the stationary measure on the Gromov boundary in the case when $G$ acts cocompactly on $M$. This optimal version is not needed for the proof of the central limit theorem but seems interesting in its own right.

We thank M. Björklund for interesting discussions on this topic.
§ 2. Hyperbolic spaces and their boundaries

In this section we recall without proof the basic definitions and properties of Gromov hyperbolic spaces and their boundaries (see [6]–[10] for more details).

2.1. Gromov hyperbolic metric spaces. Let \((M, d)\) be a metric space and let \(o\) be a point of \(M\). The Gromov product is defined for all \(o, m_1, m_2\) in \(M\) by

\[
(m_1|m_2)_o := \frac{1}{2} (d(o, m_1) + d(o, m_2) - d(m_1, m_2)).
\]

We assume that \(M\) is a Gromov hyperbolic metric space: there is a constant \(\delta > 0\) such that for all points \(o, m_1, m_2, m_3\) in \(M\) we have

\[
(m_1|m_3)_o \geq \min((m_1|m_2)_o, (m_2|m_3)_o) - \delta.
\]  

(2.1)

**Definition 2.1.** A metric space \(M\) is said to be proper if all bounded closed subsets of \(M\) are compact. It is said to be quasiconvex if there is a constant \(C > 1\) such that any two points \(m\) and \(m'\) in \(M\) can be joined by a \(C\)-geodesic, that is, one can find a sequence \(m_1 = m, m_2, \ldots, m_n = m'\) of points in \(M\) such that the following inequalities hold for all \(i, j, 1 \leq i \leq j \leq n\):

\[
j - i - C \leq d(m_i, m_j) \leq j - i + C.
\]

In particular, if \(M\) is geodesic (that is, any two points of \(M\) can be joined by a geodesic arc), then \(M\) is quasiconvex.

2.2. Gromov boundary. Let \((M, d)\) be a proper quasiconvex Gromov hyperbolic space. The Gromov boundary \(\partial M\) is the set of equivalence classes \(\xi = [m_n]\) of sequences \((m_n)\) of points in \(M\) tending to infinity and satisfying \((m_n|m_p)_o \to_{n,p \to \infty} \infty\), where sequences \((m_n)\) and \((m'_n)\) are said to be equivalent if \((m_n|m'_p)_o \to_{n,p \to \infty} \infty\).

This is an equivalence relation because \(M\) is Gromov hyperbolic. Note that since \(M\) is proper and quasiconvex, every equivalence class \(\xi\) contains a representative \((m_n)\) which is a \(C\)-geodesic.

We extend the Gromov product to the Gromov boundary: for \(\xi, \xi'\) in \(\partial M\) and \(m\) in \(M\) we put

\[
(m|\xi)_o = (\xi|m)_o = \inf_{n \to \infty} \liminf (m|m_n)_o, \\
(\xi|\xi')_o = \inf_{n \to \infty} \liminf (m_n|m'_n)_o \in [0, \infty],
\]

(2.2) (2.3)

where the infimum is taken over all sequences \(m_n\) and \(m'_n\) such that \(\xi = [m_n]\) and \(\xi' = [m'_n]\).

We endow the union \(M^* := M \cup \partial M\) with the following topology. A basis of neighbourhoods of a point \(m\) in \(M\) is given by the balls

\[
B(m, r) := \{m' \in M \mid d(m, m') \leq r\}, \quad r > 0.
\]

A basis of neighbourhoods of a point \(\xi\) in \(\partial M\) is given by the sets

\[
\mathcal{V}(\xi, R) := \{\xi' \in M^* \mid (\xi|\xi')_o \geq R\}, \quad R > 0.
\]
Since $M$ is proper and quasiconvex, the resulting topological space $M^*$ is compact and metrizable. It is called the \textit{Gromov compactification} of $M$.

For all points $\xi_1$, $\xi_2$, $\xi_3$ in $M^*$, just like for all points in $M$, we have
\[
(\xi_1|\xi_3)_o \geq \min((\xi_1|\xi_2)_o, (\xi_2|\xi_3)_o) - \delta. \tag{2.4}
\]
The Gromov product does not have to be continuous on $M^*$, but the following rough continuity property holds: for any convergent sequences $\xi_n \to \xi$ and $\xi'_n \to \xi'$ in $M^*$ we have
\[
(\xi|\xi')_o \leq \liminf_{n \to \infty} (\xi_n|\xi'_n)_o \leq \limsup_{n \to \infty} (\xi_n|\xi'_n)_o \leq (\xi|\xi')_o + 2\delta. \tag{2.5}
\]
(The terms here may take values in the extended half-line $[0, \infty]$.)

\textbf{2.3. Busemann boundary.} Let $(M,d)$ be a proper quasiconvex Gromov hyperbolic space. Its \textit{Busemann compactification} $\overline{M}$ is the set of equivalence classes $x$ of sequences $(m_n)$ of points in $M$ such that for any $m$ in $M$ there is a limit
\[
h_x(m) := \lim_{n \to \infty} d(m,m_n) - d(o,m_n), \tag{2.6}
\]
and sequences $(m_n)$, $(m'_n)$ are said to be equivalent if they define the same limit function on $M$. This function $h_x$ is called the \textit{Busemann function} at $x$. Identifying every equivalence class $x \in \overline{M}$ with the corresponding Busemann function, we endow $\overline{M}$ with the topology of uniform convergence of compact subsets. Since $M$ is proper and all Busemann functions are 1-Lipschitzian and vanish at $o$, the topological space $\overline{M}$ is compact and metrizable. Moreover, since $M$ is quasiconvex, $\overline{M}$ contains $M$ as an open subset. The space $X := \overline{M}\setminus M$ is called the \textit{Busemann boundary} of $M$.

We extend the Gromov product to the Busemann boundary putting, for all $x$, $x'$ in $X$ and $m$ in $M$,
\[
(m|x)_o = (x|m)_o = \frac{1}{2}(d(m,o) - h_x(m)), \tag{2.7}
\]
\[
(x|x')_o := - \min_{m \in M} \frac{1}{2}(h_x(m) + h_{x'}(m)). \tag{2.8}
\]

We denote by $\pi: \overline{M} \to M^*$, $x \mapsto \pi_x$, the natural projection of the Busemann compactification onto the Gromov compactification. This is the unique continuous map such that $\pi_m = m$ for all $m$ in $M$. We also note that this map is surjective.

The Gromov products on the Busemann compactification $\overline{M}$ and the Gromov compactification $M^*$ are related by the following inequalities. \textit{There is a constant $C_0 > 0$ such that for all $x$, $y$ in $\overline{M}$ we have}
\[
(\pi_x|\pi_y)_o - C_0 \leq (x|y)_o \leq (\pi_x|\pi_y)_o + C_0. \tag{2.9}
\]
In particular, points $x$ and $y$ in $X$ have the same image $\pi_x = \pi_y$ in $\partial M$ if and only if $(x|y)_o = \infty$.

The following inequality enables us to control the distance function in terms of the Busemann functions. \textit{For any $x$ and $y$ in $\overline{M}$ with $\pi_x \neq \pi_y$ there is a constant $C_{x,y} > 0$ such that for all $m$ in $M$ we have}
\[
\max(h_x(m),h_y(m)) \geq d(o,m) - C_{x,y}. \tag{2.10}
\]
2.4. Isometries of hyperbolic spaces. Let \( G := \text{Isom}(M) \) be the group of isometries of \( M \). We denote by \( \sigma: G \times X \to \mathbb{R} \) the Busemann cocycle. This continuous cocycle on the Busemann boundary is given by the following formula for all \( g \in G \) and \( x \in X \):
\[
\sigma(g, x) := h_x(g^{-1} o). \tag{2.11}
\]

For every \( g \in G \) we define the length
\[
\kappa(g) := d(g o, o) \tag{2.12}
\]
and the stable length
\[
\ell(g) := \lim_{n \to \infty} \frac{1}{n} \kappa(g^n). \tag{2.13}
\]

The group \( G \) acts continuously on both compactifications \( \overline{M} \) and \( M^* \), and the projection \( \pi: \overline{M} \to M^* \) is \( G \)-equivariant.

We list some useful properties of the functions introduced above. The first two of them mean that \( \sigma \) is a cocycle. They are followed by certain relations between the length \( \kappa(g) \), the Busemann cocycle and the Gromov product. These relations are used in the proof of log-regularity of the stationary measure \( \nu_{\partial M} \). The last equality is a key formula that enables us to solve the cohomological equation (4.13).

**Lemma 2.2.** Let \((M,d)\) be a proper quasiconvex Gromov hyperbolic space. Then the following equalities hold for all isometries \( g, g' \) of \( M \) and any points \( x, y \) on the Busemann boundary \( X \):
\[
\begin{align*}
\sigma(gg', x) &= \sigma(g, g' x) + \sigma(g', x), \tag{2.14} \\
\sigma(g^{-1}, x) &= -\sigma(g, g^{-1} x), \tag{2.15} \\
\kappa(g) - \sigma(g, x) &= 2(g^{-1} o | x)_o, \tag{2.16} \\
\kappa(g) + \sigma(g, x) &= 2(g o | gx)_o, \tag{2.17} \\
\sigma(g, x) &= -2(x | g^{-1} y)_o + 2(gx | y)_o + \sigma(g^{-1}, y). \tag{2.18}
\end{align*}
\]

**Proof.** These equalities follow directly from the definitions. \( \square \)

The following lemma describes the dynamics of \( G \) on the Gromov boundary. It is a key ingredient in the proof of Proposition 3.2.

**Lemma 2.3.** Let \((M,d)\) be a proper quasiconvex Gromov hyperbolic space, let \( \nu \) be an atom-free Borel probability measure on the Gromov boundary \( \partial M \), and let \( g_n \) be a sequence of isometries of \( M \) such that there is a limit measure \( \nu' := \lim_{n \to \infty} (g_n)_* \nu \).

If the sequence \( g_n \) is unbounded, then \( \nu' \) is a Dirac mass. Conversely, if \( \nu' = \delta_\xi \), then \( \lim_{n \to \infty} g_n o = \xi. \)

**Proof.** Suppose that the sequence \( g_n \) is unbounded. Passing to a subsequence, we can assume that the limits \( \xi_+ := \lim_{n \to \infty} g_n o \) and \( \xi_- := \lim_{n \to \infty} g_n^{-1} o \) exist in \( M^* \) and belong to \( \partial M \). We claim that the sequence \( g_n \eta \) converges to \( \xi_+ \) for every \( \eta \neq \xi_- \) in \( \partial M \). Indeed, since the sequence \((\eta | g_n^{-1} o)_o\) is bounded, so is the sequence \( d(o, g_n^{-1} o) - (\eta | o) g_n^{-1} o \) and, therefore, we have \( \lim_{n \to \infty} (g_n \eta | g_n o)_o = \infty \). This proves
that the sequence $g_n \eta$ converges to $\xi_+$. Since $\nu(\{\xi_+\}) = 0$, the probability measure $\nu'$ must then coincide with the Dirac mass at $\xi_+$.

The converse holds for the same reasons. Indeed, if $\nu'$ is the Dirac mass at $\xi$, then $\xi$ must be the unique cluster point for the sequence $g_n o$. □

2.5. Hyperbolic groups. Let $G$ be a locally compact group which is generated by a compact neighbourhood $V$ of the identity $e$. We can assume that $V = V^{-1}$. Define a left-invariant metric $d_V$ on $G$ by $d_V(g, h) := \inf\{n \geq 0 \mid g^{-1} h \in V^n\}$ for all $g, h \in G$. This metric induces the discrete topology.

A locally compact group $G$ is said to be hyperbolic if it is generated by a compact neighbourhood $V$ of the identity $e$ and the distance $d_V$ is hyperbolic. We note that this property is independent of the choice of $V$.

By Corollary 2.6 in [11], a locally compact group $G$ is hyperbolic if and only if it has a continuous proper cocompact isometric action on a proper geodesic hyperbolic space $(M, d)$.

§ 3. Random walks on hyperbolic spaces

In this section we recall some basic results concerning random walks on Gromov hyperbolic groups (see [1], [12] and [13] for more details; a comparison with the linear case [14]–[19] may also be useful).

3.1. Stationary measure on the Gromov boundary. Let $(M, d)$ be a proper quasiconvex Gromov hyperbolic space, let $o \in M$, let $\mu$ be a Borel probability measure on the isometry group $G := \text{Isom}(M)$, and let $G_\mu$ be the smallest closed subgroup of $G$ such that $\mu(G_\mu) = 1$. Finally, let $g_1, \ldots, g_n, \ldots$ be independent random elements of $G$ chosen according to the law $\mu$. We want to understand the behaviour of the random variables $g_n \cdots g_1 o$ and $g_1 \cdots g_n o$. We write $\mu^* n$ for the $n$th convolution power of $\mu$, namely $\mu^* n = \mu * \cdots * \mu$.

Definition 3.1. We say that $\mu$ is non-elementary if the group $G_\mu$ is unbounded and the action of $G_\mu$ on the Gromov boundary $\partial M$ has no finite orbits.

We say that $\mu$ is non-arithmetic if one can find an integer $n \geq 1$ and elements $g, g'$ in the support of $\mu^* n$ such that $\ell(g) \neq \ell(g')$.

Let $(B, B, \beta, S)$ be the one-sided Bernoulli system associated with $\mu$ on $G$. That is, $B = G^{\mathbb{N}^*}$ is the set of sequences $b = (b_1, b_2, \ldots)$ with $b_n \in G$, $B$ is the natural $\sigma$-algebra on the product of spaces, $\beta = \mu^\otimes \mathbb{N}^*$ is the product of the measures $\mu$ over all coordinates, and $S \colon B \to B$ is the shift given by $Sb = (b_2, b_3, \ldots)$. For every $n \geq 1$ let $B_n$ be the $\sigma$-algebra generated by the first $n$ coordinates $b_1, \ldots, b_n$.

A Borel probability measure $\nu$ on $X$ or on $\partial M$ is said to be $\mu$-stationary if $\mu * \nu = \nu$.

Proposition 3.2. Let $(M, d)$ be a proper quasiconvex Gromov hyperbolic space, let $o \in M$, and let $\mu$ be a non-elementary Borel probability measure on the group $G := \text{Isom}(M)$. Then the following assertions hold.

a) For $\beta$-almost all $b \in B$ we have

$$\lim_{n \to \infty} \kappa(b_1 \cdots b_n) = \lim_{n \to \infty} \kappa(b_n \cdots b_1) = \infty.$$
b) For $\beta$-almost all $b \in B$ the limit $\xi_b := \lim_{n \to \infty} b_1 \cdots b_n o$ exists in $\partial M$ and we have $\xi_b = b_1 \xi_{Sb}$.

c) There is a unique $\mu$-stationary Borel probability measure on the Gromov boundary $\partial M$. It is given by $\nu_{\partial M} := \int_B \delta_{\xi_b} d\beta(b)$.

Proof. Let $\nu$ be a $\mu$-stationary probability measure on $\partial M$. Since $G_\mu$ has no finite orbits in $\partial M$, the measure $\nu$ is atom-free. By the martingale limit theorem, for $\beta$-almost all $b \in B$ there is a limiting probability measure $\nu_b := \lim_{n \to \infty} (b_1 \cdots b_n)_o \nu$.

Since $G_\mu$ is unbounded, the sequence $b_1 \cdots b_n$ is unbounded for $\beta$-almost all $b$. Hence, by Lemma 2.3, $\nu_b$ is the Dirac mass at some point $\xi_b \in \partial M$ and we have $\lim_{n \to \infty} b_1 \cdots b_n o = \xi_b$. Since $\nu = \int_B \delta_{\xi_b} d\beta(b)$, the $\mu$-stationary probability measure $\nu$ is unique. \qed

3.2. Random walk on the Busemann boundary. The following proposition compares the behaviour of the random variables $\kappa(b_n \cdots b_1)$ and $\sigma(b_n \cdots b_1, x)$, where $x$ is a point of the Busemann boundary $X$.

Proposition 3.3. Let $(M,d)$ be a proper quasiconvex Gromov hyperbolic space, let $o \in M$, and let $\mu$ be a non-elementary Borel probability measure on the group $G := \text{Isom}(M)$. Then the following assertions hold.

a) For every $\varepsilon > 0$ there is $T > 0$ such that for all $x \in X$ we have

$$\beta\left(\{b \in B \mid \sup_{n \geq 1} (\kappa(b_n \cdots b_1) - \sigma(b_n \cdots b_1, x)) \leq T\} \right) \geq 1 - \varepsilon$$  \hspace{1cm} (3.1)

and, therefore, for all $n \geq 1$,

$$\mu^n\left(\{g \in G \mid (\kappa(g) - \sigma(g, x)) \leq T\} \right) \geq 1 - \varepsilon. \hspace{1cm} (3.2)$$

b) For all $x \in X$ and for $\beta$-almost all $b \in B$ we have

$$\lim_{n \to \infty} \sigma(b_n \cdots b_1, x) = \infty.$$

Proof. a) By (2.9) and (2.16) it suffices to prove that for every $\varepsilon > 0$ there is $T > 0$ such that for all $\xi \in \partial M$ we have

$$\beta\left(\{b \in B \mid \sup_{n \geq 1} (b_1^{-1} \cdots b_n^{-1} o |\xi|_o) \leq T\} \right) \geq 1 - \varepsilon.$$  \hspace{1cm} (3.3)

By Proposition 3.2, the limit

$$\xi_b^- := \lim_{n \to \infty} b_1^{-1} \cdots b_n^{-1} o$$ \hspace{1cm} (3.4)

exists on $\partial M$ for $\beta$-almost all $b$ in $B$.

On the one hand, since $\mu$ is non-elementary, the $\tilde{\mu}$-stationary probability measure $\nu_{\partial M} = \int_B \delta_{\xi_b^-} d\beta(b)$ is atom free. Hence for every $\varepsilon > 0$ there is $R > 0$ such that for all $\xi \in \partial M$ we have

$$\beta\left(\{b \in B \mid (\xi_b^- |\xi|_o \leq R\} \right) \geq 1 - \frac{\varepsilon}{2}. \hspace{1cm} (3.5)$$
On the other hand, using (3.4), (2.4) and (2.5), we can see that for all $R > 0$ and for $\beta$-almost all $b \in B$ there is $T_{R,b} > 0$ such that for all $\xi \in \partial M$ with $(\xi_b - |\xi|)_{\partial M} \leq R$ we have
\[
\sup_{n \geq 1} (b_1^{-1} \cdots b_n^{-1} o |\xi|)_{\partial M} \leq T_{R,b}.
\] (3.6)
We now choose a constant $T > 0$ such that
\[
\beta(\{b \in B \mid T_{R,b} \leq T\}) \geq 1 - \frac{\varepsilon}{2}.
\] (3.7)
Then the desired inequality (3.3) follows from (3.5), (3.6) and (3.7).

b) This follows from (3.1) and Proposition 3.2(a). ∎

3.3. The escape rate of a random walk. We assume in this subsection that the first moment of $\mu$ is finite: $\int_G \kappa(g) d\mu(g) < \infty$. Then the limit
\[
\lambda := \lim_{n \to \infty} \frac{1}{n} \int_G \kappa(g) d\mu^n(g)
\]
exists by subadditivity and is called the escape rate of $\mu$.

We have seen in Proposition 3.2 that the Gromov boundary $\partial M$ supports a unique stationary measure. There may be more than one stationary measure on the Busemann boundary $X$, but according to the following Proposition 3.4(c) the Busemann cocycle $\sigma$ on the Busemann boundary $X$ has a unique average $\lambda$.

**Proposition 3.4.** Let $(M,d)$ be a proper quasiconvex Gromov hyperbolic space, let $o \in M$, let $\mu$ be a non-elementary Borel probability measure on $G := \text{Isom}(M)$ with
\[
\int_G \kappa(g) d\mu(g) < \infty,
\]
and let $\lambda$ be the escape rate of $\mu$. Then the following assertions hold.

a) For $\beta$-almost all $b \in B$ we have
\[
\lambda = \lim_{n \to \infty} \frac{1}{n} \kappa(b_n \cdots b_1).
\]
b) For every $x \in X$ and for $\beta$-almost all $b \in B$ we have
\[
\lambda = \lim_{n \to \infty} \frac{1}{n} \sigma(b_n \cdots b_1, x).
\]
c) For every $\mu$-stationary Borel probability measure $\nu$ on $X$ we have
\[
\lambda = \int_{G \times X} \sigma(g,x) d\mu(g) d\nu(x).
\]
d) The escape rate is positive: $\lambda > 0$.

**Proof.** a) This follows from Kingman’s subadditive ergodic theorem (see [20]) as applied to the sequence of integrable functions $a_n : B \to \mathbb{R}_+$ given by $a_n(b) = \kappa(b_n \cdots b_1)$, which satisfy the subadditivity condition $a_{n+m} \leq a_n \circ S^m + a_m$ for all $m, n \geq 1$. Since the dynamical system $(B, S, \beta)$ is ergodic, this theorem asserts that for $\beta$-almost all $b \in B$, the limit $\lim_{n \to \infty} \frac{1}{n} a_n(b)$ exists and is equal to $\lambda := \lim_{n \to \infty} \frac{1}{n} \int_B a_n(b) d\beta(b)$.

b) This follows from part a). Indeed, the sequence of differences $\kappa(b_n \cdots b_1) - \sigma(b_n \cdots b_1, x)$ is bounded for all $x \in X$ by Proposition 3.3(a).
c) We apply Birkhoff’s ergodic theorem to the measurable transformation $T$ of $B \times X$ given by $T(b, x) = (Sb, b_1 x)$, which preserves the measure $\beta \otimes \nu$, and to the function $f$ on $B \times X$ given by $f(b, x) = \sigma(b_1, x)$. It tells us that the limit of the Birkhoff averages

$$\tilde{f}(b, x) = \lim_{n \to \infty} \frac{1}{n} \sum_{1 \leq k \leq n} f(T^k(b, x))$$

exists for $\beta \otimes \nu$-almost all pairs $(b, x)$ and its $\beta \otimes \nu$-average is equal to that of $f$. We know from part b) that this limit is $\beta \otimes \nu$-almost surely equal to the constant $\lambda$. Hence,

$$\lambda = \int_{B \times X} f(b, x) \, d\beta(b) \, d\nu(x) = \int_{G \times X} \sigma(g, x) \, d\mu(g) \, d\nu(x).$$

d) By Proposition 3.3(b), the Birkhoff sums of $f$ tend to infinity for $\beta \otimes \nu$-almost all $(b, x) \in B \times X$:

$$\lim_{n \to \infty} \sum_{1 \leq k \leq n} f(T^k(b, x)) = \infty.$$

Hence, by Kesten’s lemma (see [21] or [17], Lemma II.2.2), the $\beta \otimes \nu$-average of $f$ is positive, and we obtain $\lambda > 0$. □

§ 4. The central limit theorem

In this section we begin the proof of the central limit theorem (Theorem 1.1). As in the linear case in [4], the key step is to establish a regularity property for the $\mu$-stationary measure on the Gromov boundary.

4.1. Complete convergence to the escape rate. As in [4], we need the following Proposition 4.1, which is an analogue of the Hsu–Robbins–Baum–Katz theorem (see [22], [23]) for the convergence to the escape rate in Proposition 3.4. When $p = 2$, it tells us that this convergence is complete.

**Proposition 4.1.** Suppose that $p > 1$, $(M, d)$ is a proper quasiconvex Gromov hyperbolic space, and $o \in M$. Let $\mu$ be a non-elementary Borel probability measure on $G$ with $\int_G \kappa(g)^p \, d\mu(g) < \infty$, and let $\lambda$ be the escape rate of $\mu$. Then for every $\varepsilon > 0$ there are constants $C_n = C_n(p, \varepsilon, \mu)$ such that $\sum_{n \geq 1} n^{p-2} C_n < \infty$ and the following inequalities hold for all $x$ in the Busemann boundary $X$:

$$\mu^*(\{ g \in G : |\sigma(g, x) - n\lambda| \geq \varepsilon n \}) \leq C_n,$$  \hfill (4.1)

$$\mu^*(\{ g \in G : |\kappa(g) - n\lambda| \geq \varepsilon n \}) \leq C_n.$$  \hfill (4.2)

**Proof.** Since the Busemann cocycle $\sigma$ has a unique average $\lambda$, (4.1) follows from Proposition 3.2 in [4]. By (2.10), for all points $x, y$ in $X$ with $\pi_x \neq \pi_y$ there is $C > 0$ such that for all $g \in G$ we have

$$\kappa(g) - C \leq \max(\sigma(g, x), \sigma(g, y)) \leq \kappa(g).$$  \hfill (4.3)

Then (4.2) with another constant $C_n$ follows from (4.1). □
4.2. Log-regularity of the stationary measure. The following proposition with $p = 2$ will be used to solve the cohomological equation (4.13).

**Proposition 4.2.** Suppose that $p > 1$, $(M,d)$ is a proper quasiconvex Gromov hyperbolic space, and $o \in M$. Let $\mu$ be a non-elementary Borel probability measure on $G$ with $\int_G \kappa(g)^p \, d\mu(g) < \infty$, and let $\nu$ be a $\mu$-stationary Borel probability measure on the Busemann boundary $X$. Then

$$\sup_{y \in X} \int_X (x|y)_o^{p-1} \, d\nu(x) < \infty. \quad (4.4)$$

**Remark 4.3.** Using (2.9), one can restate (4.4) as

$$\sup_{\eta \in \partial M} \int_{\partial M} (\xi|\eta)_o^{p-1} \, d\nu_{\partial M}(\xi) < \infty, \quad (4.5)$$

where $\nu_{\partial M}$ is the unique $\mu$-stationary Borel probability measure on the Gromov boundary $\partial M$.

**Remark 4.4.** When $\mu$ is assumed to have an exponential moment, one can prove that the stationary measure $\nu$ is much more regular: there is $t > 0$ such that

$$\sup_{y \in X} \int_X e^{t(x|y)_o} \, d\nu(x) < \infty. \quad (4.6)$$

**Lemma 4.5.** Under the hypotheses of Proposition 4.2 there are constants $a > 0$ and $C_n > 0$ such that $\sum_{n \geq 1} n^{p-2} C_n < \infty$ and the following inequalities hold for all $n \geq 1$ and $x,y \in X$:

$$\mu^*(\{g \in G \mid (go|gx)_o \leq an\}) \leq C_n, \quad (4.7)$$

$$\mu^*(\{g \in G \mid (go|y)_o \geq an\}) \leq C_n, \quad (4.8)$$

$$\mu^*(\{g \in G \mid (gx|y)_o \geq an\}) \leq C_n. \quad (4.9)$$

**Proof.** By Proposition 3.4, the escape rate $\lambda$ of $\mu$ is positive. We put $a = \lambda/2$. According to Proposition 4.1, there are constants $C_n$ such that $\sum_{n \geq 1} n^{p-2} C_n < \infty$ and, for all $n \geq 1$ and $x,y \in X$, there are subsets $G_{n,x,y} \subset G$ with $\mu^*(G_{n,x,y}) \geq 1 - C_n$ such that

$$|\kappa(g) - \lambda n|, \quad |\sigma(g,x) - \lambda n|, \quad |\sigma(g^{-1},y) - \lambda n|$$

are less than or equal to $\lambda n/4$ for any $g \in G_{n,x,y}$. We shall prove the inequalities (4.7)–(4.9) only for $n \geq n_0$, where $n_0$ is a large enough number to be chosen below. It suffices to check that for $n \geq n_0$ and $g \in G_{n,x,y}$ we have

$$(go|gx)_o \geq an, \quad (go|y)_o \leq an, \quad (gx|y)_o \leq an.$$  

We first notice that, in accordance with (2.17),

$$(go|gx)_o = \frac{1}{2} (\kappa(g) + \sigma(g,x)) \geq \frac{3\lambda n}{4}.$$  

This proves (4.7).
Using (2.16), we obtain

\[(go|y)_o = \frac{1}{2}(\kappa(g) - \sigma(g^{-1}, y)) \leq \frac{\lambda n}{4}.
\]

This proves (4.8).

Combining these two equalities and the bound

\[(go|y)_o \geq \min((go|gx)_o, (gx|y)_o) - \delta,
\]

we obtain for \(n \geq n_0 := 4\delta/\lambda\) that

\[(gx|y)_o \leq (go|y)_o + \delta \leq \frac{\lambda n}{2}.
\]

This proves (4.9). □

**Proof of Proposition 4.2.** Choose \(a\) and \(C_n\) as in Lemma 4.5. We claim that the following inequality holds for any \(n \geq 1\) and \(y \in X\):

\[\nu(\{x \in X \mid (x|y)_o \geq an\}) \leq C_n. \tag{4.10}\]

Indeed, since \(\nu = \mu^* \star \nu\), we obtain from (4.9) that

\[\nu(\{x \in X \mid (x|y)_o \geq an\}) = \int_X \mu^*(\{g \in G \mid (gx|y)_o \geq an\}) \, d\nu(x)
\leq \int_X C_n \, d\nu(x) = C_n.
\]

Then, splitting the integral (4.6) into integrals over the sets \(A_{n-1,y} \setminus A_{n,y}\), where

\[A_{n,y} := \{x \in X \mid (x|y)_o \geq an\},
\]

we obtain the upper bound

\[\int_X (x|y)_o^{p-1} \, d\nu(x) \leq \sum_{n \geq 1} a^{p-1} n^{p-1} (\nu(A_{n-1,y}) - \nu(A_{n,y}))
\leq a^{p-1} + a^{p-1} \sum_{n \geq 1} ((n + 1)^{p-1} - n^{p-1}) C_n \leq a^{p-1} + (p - 1)2^p a^{p-1} \sum_{n \geq 1} n^{p-2} C_n.
\]

It is finite, and this proves (4.4). □

### 4.3. Solving the cohomological equation.

**Proposition 4.6.** Let \((M,d)\) be a proper quasiconvex Gromov hyperbolic space, let \(o\) be a point of \(M\), and let \(\mu\) be a non-elementary Borel probability measure on \(G\) with a finite second moment. Then the Busemann cocycle \(\sigma\) on the Busemann boundary \(X\) is centerable, that is, one can find a bounded function \(\psi\) on \(X\) such that the cocycle \(\sigma_0\), which is defined for all \((g,x) \in G \times X\) by

\[\sigma_0(g,x) = \sigma(g,x) - \psi(x) + \psi(gx), \tag{4.11}\]

satisfies the following equality for all \(x \in X\):

\[\int_G \sigma_0(g,x) \, d\mu(g) = \lambda. \tag{4.12}\]
We define a function $\psi$ on $X$ by the following formula for all $x \in X$:

$$
\psi(x) = -2 \int_X (x | y)_o \, d\nu^*(y),
$$

where $\nu^*$ is a $\tilde{\mu}$-stationary probability measure on $X$. By (2.18) we have

$$
\sigma(g, x) = -2(x | g^{-1} y)_o + 2(gx | y)_o + \sigma(g^{-1}, y)
$$

for all $g \in G$ and $x, y \in X$. Integrating this equality over $G \times X$ with respect to the measure $d\mu(g) \, d\nu^*(y)$ and using the $\tilde{\mu}$-stationarity of $\nu^*$, we obtain

$$
\int_G \sigma(g, x) \, d\mu(g) = \psi(x) - \int_G \psi(gx) \, d\mu(g) + \lambda \quad (4.13)
$$

for all $x \in X$. Hence the cocycle $\sigma_0: (g, x) \mapsto \sigma(g, x) - \psi(x) + \psi(gx)$ satisfies (4.12). □

**4.4. The central limit theorem.** We restate our central limit theorem more precisely.

**Theorem 4.7.** Let $(M, d)$ be a proper quasiconvex Gromov hyperbolic space, let $o$ be a point of $M$, let $\mu$ be a non-elementary Borel probability measure on the group $G := \text{Isom}(M)$ with finite second moment $\int_G \kappa(g)^2 \, d\mu(g) < \infty$, and let $\lambda$ be the escape rate of $\mu$.

a) Then there is a Gaussian law $N_\mu$ on $\mathbb{R}$ such that for every continuous compactly supported function $F$ on $\mathbb{R}$ we have

$$
\int_G F\left(\frac{\sigma(g, x) - n \lambda}{\sqrt{n}}\right) \, d\mu^*(g) \xrightarrow{n \to \infty} \int_{\mathbb{R}} F(t) \, dN_\mu(t) \quad (4.14)
$$

uniformly with respect to $x$ in the Busemann boundary $X$, and

$$
\int_G F\left(\frac{\kappa(g) - n \lambda}{\sqrt{n}}\right) \, d\mu^*(g) \xrightarrow{n \to \infty} \int_{\mathbb{R}} F(t) \, dN_\mu(t). \quad (4.15)
$$

b) When $\mu$ is non-arithmetic, this Gaussian law $N_\mu$ is non-degenerate.

As already pointed out, the novelty here is that we do not assume the existence of a finite exponential moment for $\mu$ (see [1]). Not even in the case when $G$ is a free group with two generators and $d$ is the left-invariant distance on $M = G$ given by the length of a minimal word in these generators (word metric) has Theorem 4.7 been previously known in this generality (see [2] and [3]).

We shall use the following central limit theorem for martingales, which was proved by Brown [5] (see also [24]).

**Fact 4.8.** Suppose that $(\Omega, \mathcal{B}, \mathbb{P})$ is a probability space, $\mathcal{B}_0 \subset \cdots \subset \mathcal{B}_n \subset \cdots$ are sub-$\sigma$-algebras of $\mathcal{B}$ and $\varphi_{n,k}: \Omega \to \mathbb{R}$, $1 \leq k \leq n$, are $\mathcal{B}_k$-measurable square-integrable random variables with

$$
\mathbb{E}(\varphi_{n,k} | \mathcal{B}_{k-1}) = 0. \quad (4.16)
$$
Let $N_\Phi$ be a centred Gaussian law on $\mathbb{R}$ with variance $\Phi \geq 0$. Suppose that the random variables

$$W_n := \sum_{1 \leq k \leq n} \mathbb{E}(\varphi_{n,k}^2 | \mathcal{B}_{k-1})$$

converge to $\Phi$ in probability (4.17) and that, for all $\varepsilon > 0$,

$$W_{\varepsilon,n} := \sum_{1 \leq k \leq n} \mathbb{E}(\varphi_{n,k}^2 1_{\{|\varphi_{n,k}| \geq \varepsilon\}} | \mathcal{B}_{k-1}) \xrightarrow{n \to \infty} 0 \quad \text{in probability.}$$

Then the sequence $S_n := \sum_{1 \leq k \leq n} \varphi_{n,k}$ converges in law to $N_\Phi$.

**Proof of Theorem 4.7. a)** By Proposition 3.3, if the limit (4.14) exists for some $x \in X$, then the sequence (4.15) converges to the same limit, the limits (4.14) exist for all $x \in X$ and are equal to each other, and the convergence in (4.14) is uniform with respect to $x \in X$.

Since the cocycle $\sigma$ is centerable, it can be written as a sum of two cocycles: $\sigma = \sigma_0 + \sigma_1$, where $\sigma_0$ is given by (4.11) and $\sigma_1$ is a coboundary which is uniformly bounded by the constant $2\|\psi\|_\infty$. In particular, the cocycle $\sigma_1$ plays no role in the limit (4.14) and, therefore, we can replace $\sigma$ by $\sigma_0$ in (4.14).

As above, let $(B, \mathcal{B}, \beta)$ be the associated Bernoulli space. We want to find $x \in X$ such that the laws of the random variables $S_n$ on $B$, which are defined for $b \in B$ by

$$S_n(b) := \frac{1}{\sqrt{n}} \left( \sigma_0(b_n \cdots b_1, x) - n \lambda \right),$$

converge to some Gaussian law $N_\mu$.

We want to apply the central limit theorem for martingales (Fact 4.8) to the sub-$\sigma$-algebras $\mathcal{B}_k$ spanned by $b_1, \ldots, b_k$ and to the triangular array of random variables $\varphi_{n,k}$ on $B$, which are given for $b \in B$ by

$$\varphi_{n,k}(b) = \frac{1}{\sqrt{n}} \left( \sigma_0(b_k, b_{k-1} \cdots b_1 x) - \lambda \right), \quad 1 \leq k \leq n.$$ 

Since $S_n = \sum_{1 \leq k \leq n} \varphi_{n,k}$ by the cocycle property, it suffices to verify that the three hypotheses of Fact 4.8 hold for some choice of the constant $\Phi = \Phi_\mu \geq 0$. We keep the notation $W_n$ and $W_{\varepsilon,n}$ used in Fact 4.8.

First, since $\kappa$ is square-integrable, the functions $\varphi_{n,k}$ belong to $L^2(B, \beta)$ and (4.12) implies (4.16): for $\beta$-almost all $b \in B$ we have

$$\mathbb{E}(\varphi_{n,k} | \mathcal{B}_{k-1}) = \int_G \left( \sigma_0(g, b_{k-1} \cdots b_1 x) - \lambda \right) d\mu(g) = 0.$$

Second, we introduce the following continuous function on $X$:

$$x \mapsto M(x) = \int_G \left( \sigma_0(g, x) - \lambda \right)^2 d\mu(g).$$

Then for $\beta$-almost all $b \in B$,

$$W_n(b) = \frac{1}{n} \sum_{1 \leq k \leq n} M(b_{k-1} \cdots b_1 x).$$
Fix a \( \mu \)-ergodic \( \mu \)-stationary Borel probability measure \( \nu \) on \( X \). By Birkhoff’s ergodic theorem, the sequence \( W_n \) converges to
\[
\Phi_\mu := \int_X M(y) \, d\nu(y) \tag{4.19}
\]
in \( L^1(B, \beta) \) for \( \nu \)-almost all \( x \in X \). We choose any such point \( x \) in \( X \). Then, in particular, the hypothesis (4.17) holds.

Third, for every \( T > 0 \) we introduce the following continuous function \( M_T \) on \( X \):
\[
x \mapsto M_T(x) = \int_G (\sigma_0(g, x) - \lambda)^2 \mathbf{1}_{\{ |\sigma_0(g, x) - \lambda| \geq T \}} \, d\mu(g),
\]
and the following integral \( I_T \):
\[
I_T := \int_G \kappa_0(g)^2 \mathbf{1}_{\{ \kappa_0(g) \geq T \}} \, d\mu(g),
\]
where \( \kappa_0(g) := \kappa(g) + \lambda + 2\|\psi\|_\infty \), so that
\[
M_T(x) \leq I_T \underset{T \to \infty}{\longrightarrow} 0,
\]
and we find that for \( \varepsilon > 0 \) and for \( \beta \)-almost all \( b \) in \( B \),
\[
W_{\varepsilon, n}(b) = \frac{1}{n} \sum_{1 \leq k \leq n} M_{\varepsilon, \sqrt{n}}(b_{k-1} \cdots b_1 x) \leq I_{\varepsilon, \sqrt{n}} \underset{n \to \infty}{\longrightarrow} 0.
\]
In particular, the sequence \( W_{\varepsilon, n} \) converges to zero in probability. Hence the Lindeberg condition (4.18) holds. By Fact 4.8, the laws of \( S_n \) converge to the Gaussian law \( N_\mu \) with variance \( \Phi_\mu \).

b) It remains to verify that if \( \mu \) is non-arithmetic, then the Gaussian law \( N_\mu \) is not a Dirac mass. The variance \( \Phi_\mu \) of this Gaussian law is given by the following formula, which holds for all \( n \geq 1 \):
\[
\Phi_\mu = \frac{1}{n} \int_{G \times X} (\sigma_0(g, x) - n\lambda)^2 \, d\mu^*(g) \, d\nu(x).
\]
In particular, since the cocycle \( \sigma_1 \) is bounded by the constant \( 2\|\psi\|_\infty \) and \( \sigma \) is continuous, we obtain
\[
|\sigma(g, x) - n\lambda| \leq 2\|\psi\|_\infty \tag{4.20}
\]
for all \( g \) in the support of \( \mu^*n \) and all \( x \) in the support of \( \nu \). Using the bound (4.3), we can find a constant \( C > 0 \) such that for all \( n \geq 1 \) and for all \( g \) in the support of \( \mu^*n \) we have
\[
|\kappa(g) - n\lambda| \leq C. \tag{4.21}
\]
Then it follows from (2.13) that
\[
\ell(g) = n\lambda \tag{4.22}
\]
contrary to the non-arithmeticity of \( \mu \). □

Remark 4.9. It follows from this proof that the variance \( \Phi_\mu \), which is given by (4.19), does not depend on the choice of the \( \mu \)-stationary measure \( \nu \) on the Busemann boundary \( X \).
§ 5. Log-regularity of the stationary measure

In this section we give an alternative proof of the log\(^p\)-regularity of the \(\mu\)-stationary measure on the Gromov boundary.

5.1. Optimal log-regularity. In this alternative proof we assume that the isometry group \(G\) of the space \(M\) acts cocompactly on \(M\). On the other hand, we only assume the \(p\)th moment of \(\mu\) on \(G\) to be finite (rather than the \((p + 1)\)th moment as in Proposition 4.2). Here is the precise statement.

Proposition 5.1. Suppose that \(p > 0\), \((M, d)\) is a proper quasiconvex Gromov hyperbolic space, \(o \in M\), and the group \(G\) of isometries of \(M\) acts cocompactly on \(M\). Let \(\mu\) be a non-elementary Borel probability measure on \(G\) with \(\int_{G} \kappa(g)\, d\mu(g) < \infty\), and let \(\nu\) be the \(\mu\)-stationary Borel probability measure on the Gromov boundary \(\partial M\). Then

\[
\sup_{\eta \in \partial M} \int_{\partial M} (\xi | \eta)^p\, d\nu(\xi) < \infty. \tag{5.1}
\]

The alternative proof does not rely on martingales. It consists of three steps. The first step (Lemma 5.2) uses harmonic analysis on \(L^2(G)\) in the form of the spectral-gap characterization of non-amenability. The second step (Lemma 5.3) relies on the geometric properties of hyperbolic spaces: all geodesic triangles are \(\delta\)-thin. The last step (§ 5.4) uses an interpretation of the stationary measure as the image of the Bernoulli measure under the boundary map. Here are the details.

5.2. Spectral gap. To prove Proposition 5.1, we may (and will) assume that \(e\) belongs to the support of \(\mu\). Indeed, \(\nu\) is also \(\mu'\)-stationary, where \(\mu' := \frac{1}{2} (\mu + \delta_e)\).

We first notice that since \(\mu\) is a non-elementary probability measure on \(G\), the group \(G_{\mu}\) is non-amenable. Indeed, if \(G_{\mu}\) were amenable, there would be \(a\)-invariant probability measure on the Gromov boundary \(\partial M\). By Proposition 3.2 this would imply that the unique \(\mu\)-stationary measure \(\nu\) on \(\partial M\) is \(G_{\mu}\)-invariant and, therefore, \(\nu\) would be either a Dirac mass or a linear combination of two Dirac masses. This contradicts the fact that \(G_{\mu}\) has no finite orbits on \(\partial M\).

Hence we can apply the following lemma to our measure \(\mu\).

Lemma 5.2. Let \((M, d)\) be a proper metric space. Suppose that the group \(G\) of isometries of \(M\) acts cocompactly on \(M\). Let \(\mu\) be a non-elementary Borel probability measure on \(G\) such that \(G_{\mu}\) is non-amenable and the support of \(\mu\) contains \(e\). Let \(R > 0\). Then there exist \(A_0 > 0\) and \(a_0 < 1\) such that for all \(m, m' \in M\) and all \(n \geq 1\) we have

\[
\mu^*(\{ g \in G \mid d(gm, m') \leq R \}) \leq A_0 a_0^n. \tag{5.2}
\]

The cocompactness of the action of \(G\) is crucial in this lemma. Indeed, the conclusion of the lemma does not hold if we replace \((M, d)\) by a space consisting of isometric copies of the spaces \((M, \frac{k}{k}d)\), \(k \geq 1\), which are homothetic to the initial space.

Proof. Let \(o\) be a point in \(M\). Since \(G\) acts cocompactly on \(M\), every point of \(M\) is at a bounded distance from the \(G\)-orbit \(Go\) of the point \(o\). Hence we can assume
that the points \( m \) and \( m' \) belong to this orbit \( Go: m = ho \) and \( m' = h'o \) for some \( h, h' \) in \( G \).

Let \( \mu_G \) be a left-invariant measure on \( G \) and let \( \lambda_G \) be the left regular representation of \( G \) on \( L^2(G) \). We define a contraction \( \lambda_G(\mu) \) of \( L^2(G) \) by the following formula for all \( \varphi \in L^2(G) \) and \( g \in G \):

\[
\lambda_G(\mu)(\varphi): g' \mapsto \int_G \varphi(g^{-1}g') \, d\mu(g).
\]

By the spectral gap theorem of Kesten, Derrienic–Guivarc’h and Berg–Christensen (see [25], Theorem 4, and also [26], [27]) it follows from the non-amenability of \( G_\mu \) that the operator \( \lambda_G(\mu) \) has a spectral gap: there are \( C_0 > 0 \) and \( a_0 < 1 \) such that \( \|\lambda_G(\mu)^n\| \leq C_0 a_0^n \) for all \( n \geq 1 \). Let

\[
B_R := \{ g \in G \mid d(go, o) \leq R \}.
\]

We must estimate \( \mu^*(h'B_Rh^{-1}) \) from above. Using the inclusion \( B_RB_R \subset B_{2R} \), the Cauchy–Schwartz inequality and the spectral gap theorem, we find that for all \( n \geq 1 \)

\[
\mu_G(h'B_Rh^{-1})\mu^*(h'B_Rh^{-1}) \leq \langle \lambda_G(\mu)^n(1_{hB_2h^{-1}}), 1_{h'B_Rh^{-1}} \rangle_{L^2(G)} \\
\leq C_0 a_0^n \mu_G(hB_2h^{-1})^{1/2} \mu_G(h'B_Rh^{-1})^{1/2}.
\]

Since the Haar measure \( \mu_G \) is left-invariant and right semi-invariant, we obtain

\[
\mu^*(h'B_Rh^{-1}) \leq C_0 a_0^n \mu_G(B_{2R})^{1/2} \mu_G(B_R)^{-1/2}.
\]

This proves (5.2). \( \square \)

### 5.3. Thin triangles.

**Lemma 5.3.** Suppose that \( p > 0 \), \((M, d)\) is a proper quasiconvex Gromov hyperbolic space, \( o \in M \), and the group \( G \) of isometries of \( M \) acts cocompactly on \( M \). Let \( \mu \) be a non-elementary Borel probability measure on \( G \) such that \( \int_G \kappa(g)^p \, d\mu(g) < \infty \) and the support of \( \mu \) contains \( e \). Then there are constants \( A = A_{\mu, p} > 0 \) and \( a = a_{\mu, p} < 1 \) such that for all points \( \xi, \eta \) on the Gromov boundary \( \partial M \) and for all \( n \geq 1 \) we have

\[
\mu^*(\{ g \in G \mid (g\xi|\eta)_o \geq \kappa(g) \}) \leq Aa^n.
\]

Moreover, when \( p \geq 1 \), we also have

\[
\int_G \kappa(g)^{p-1} 1_{\{ (g\xi|\eta)_o \geq \kappa(g) \}} \, d\mu^*(g) \leq Aa^n.
\]

We recall some properties of triangles in \( M \), to be used in what follows. Since the metric space \((M, d)\) is proper, quasiconvex and hyperbolic, there are constants \( C > 1 \) and \( \delta > 0 \) with the following properties. Every triple \( x_1, x_2, x_3 \) of points in \( M^* \) is the set of vertices of a \( C \)-geodesic triangle, that is, a triangle whose sides are \( C \)-geodesics. Moreover, every \( C \)-geodesic triangle is \( \delta \)-thin, that is, one can divide each of its sides \([x_i, x_j]\) in two \( C \)-geodesic pieces: \([x_i, x_j] = [x_i, m_{i,j}] \cup [m_{i,j}, x_j]\) in such a way that the Hausdorff distance between any two pieces \([x_i, m_{i,j}]\) and \([x_i, m_{i,k}]\) that have a common vertex \( x_i \) does not exceed \( \delta \).
Proof of Lemma 5.3. We apply this property to the triples \((o, go, g\xi)\) and \((o, go, \eta)\).

Let \((m_i)_{i\geq 1}\) be a \(C\)-geodesic between the points \(o\) and \(\xi\), that is,
\[
m_1 = o, \quad \lim_{i \to \infty} m_i = \xi
\]
and for all \(i < j\) we have
\[
j - i - C \leq d(m_i, m_j) \leq j - i + C.
\]
Similarly, let \((m'_j)_{j\geq 1}\) be a \(C\)-geodesic from \(o\) to \(\eta\), and let \((m''_k)_{k\geq 1}\) be a \(C\)-geodesic from \(o\) to \(g\xi\).

We first prove (5.3). Consider the set
\[
S_{\xi, \eta} := \{ g \in G \mid (g\xi|\eta)_o \geq \kappa(g) \}.
\]
Put \(R = 2\delta + 6C\) and choose \(c > 1\) such that \(c^2a_0 < 1\), where \(a_0\) is the constant in Lemma 5.2. Let \(n \geq 1\). We claim that
\[
S_{\xi, \eta} \subset S_{\xi, \eta}^0 \cup S_{\xi, \eta}^1,
\]
where
\[
S_{\xi, \eta}^0 := \{ g \in G \mid \kappa(g) \geq c^n \},
\]
\[
S_{\xi, \eta}^1 := \{ g \in G \mid \text{there are } i, j \leq c^n \text{ such that } d(g m_i, m'_j) \leq R \}.
\]
Indeed, let \(g\) be an element of \(S_{\xi, \eta} \setminus S_{\xi, \eta}^0\). We choose a \(C\)-geodesic between \(o\) and \(go\) and apply the property stated above to the \(C\)-geodesic triangle with vertices \((o, go, g\xi)\). Since \(d(o, go) = \kappa(g)\), one can find a point \(gm_i\) and a point \(m''_k\) such that \(i \leq \kappa(g), \quad k \leq \kappa(g)\) and \(d(gm_i, m''_k) \leq \delta + 3C\).

We now apply the same property to the \(C\)-geodesic triangle with vertices \((o, go, \eta)\). Since \(k \leq \kappa(g)\) and \((g\xi|\eta)_o \geq \kappa(g)\), one can find a point \(m'_j\) such that
\[
j \leq \kappa(g) \quad \text{and} \quad d(m'_j, m''_k) \leq \delta + 3C.
\]
In particular, we have \(d(gm_i, m'_j) \leq R\) and, therefore, \(g\) belongs to \(S_{\xi, \eta}^1\). This proves the inclusion (5.5).

We need to estimate \(\mu^* n(S_{\xi, \eta})\). On the one hand, using the Chebyshev inequality and the finiteness of the \(p\)th moment, we have
\[
\mu^* n(S_{\xi, \eta}^0) \leq c^{-np} \int_G \kappa(g)^p \, d\mu^* n(g) \leq n^{p+1}c^{-np} \int_G \kappa(g)^p \, d\mu(g).
\]
On the other hand, using the bound (5.2) for at most \(c^2n\) pairs of points, we obtain
\[
\mu^* n(S_{\xi, \eta}^1) \leq A_0 c^{2n} a_0^n.
\]
This proves (5.3) as soon as the constant \(a < 1\) is chosen larger than \(c^{-p}\) and \(a_0 c^2\).
We now prove (5.4). For $p = 1$ this bound coincides with (5.3). Suppose that $p > 1$. The proof of (5.4) is similar to that of (5.3). Choose $c > 1$ such that $c^{p+1} a_0 < 1$. To estimate
\[ \int_{S_{\xi,\eta}} \kappa(g)^{p-1} d\mu^*(g), \]
we argue as above and get the bound
\[ \int_{S_{\xi,\eta}} \kappa(g)^{p-1} d\mu^*(g) \leq c^{-n} \int_G \kappa(g)^p d\mu(g) \leq n^p c^{-n} \int_G \kappa(g)^p d\mu(g) \]
and
\[ \int_{S_{\xi,\eta}} \kappa(g)^{p-1} d\mu^*(g) \leq c^{(p-1)n} \mu^*(S_{\xi,\eta}^1) \leq A_0 c^{(p+1)n} a_0^n. \]

This proves (5.4) as soon as the constant $a < 1$ is chosen larger than $c^{-1}$ and $a_0 c^{p+1}$.

5.4. The boundary map. We can now conclude the proof of Proposition 5.1.

Proof of Proposition 5.1. We first assume that $p \leq 1$. As above, let $(B, B, \beta, S)$ be the associated Bernoulli system. We write $b \mapsto \xi_b$ for the boundary map introduced in Proposition 3.2. For $b = (b_1, b_2, \ldots) \in B$ and $n \geq 0$ we put $\kappa_n(b) = \kappa(b_1 \cdots b_n)$. We need to estimate the following integral uniformly with respect to $\eta \in \partial M$:
\[ I_{p,\eta} := \int_{\partial M} (\xi|\eta)_o^p d\nu(\xi). \]

We have
\[ I_{p,\eta} = \int_B (\xi_b|\eta)_o^p d\beta(b) = \int_0^\infty \beta(\{ b \mid (\xi_b|\eta)_o^p \geq t \}) \, dt \]
\[ = \sum_{n \geq 0} \int_0^\infty \beta(\{ b \mid \kappa_n(b)^p \leq t < \kappa_{n+1}(b)^p, (\xi_b|\eta)_o^p \geq t \}) \, dt \]
\[ \leq \sum_{n \geq 0} \int_B \max(\kappa_{n+1}(b)^p - \kappa_n(b)^p, 0) 1_{\{ (\xi_b|\eta)_o \geq \kappa_n(b) \}} d\beta(b). \]

Since $p \leq 1$, for all $t > s > 0$ we have
\[ t^p - s^p \leq (t - s)^p. \]
Hence, putting $g = b_1 \cdots b_n$ and $b' = S^nb$ (so that $b'_1 = b_{n+1}$), we can continue the estimation as follows:
\[ I_{p,\eta} \leq \sum_{n \geq 0} \int_B \kappa(b'_1)^p \mu^*(\{ g \in G \mid (g\xi_{b'}|\eta)_o \geq \kappa(g) \}) \, d\beta(b'). \]

Using (5.3), we obtain that
\[ I_{p,\eta} \leq \sum_{n \geq 0} Aa^n \int_B \kappa(b'_1)^p d\beta(b'). \]
This yields the final bound
\[ I_{p,\eta} \leq \frac{A}{1-a} \int_G \kappa(g)^p \, d\mu(g). \]

When \( p > 1 \), the same computation works if we replace (5.6) by the following inequality for \( t > s > 0 \):
\[ t^p - s^p \leq 2^p (t-s)^p + 2^p s^{p-1} (t-s). \]
Hence an extra term occurs on the right-hand side of (5.7). It is equal to
\[ \sum_{n \geq 0} \int_B 2^p \kappa(b'_1) \int_G \kappa(g)^{p-1} 1_{\{g \in G | (g\xi_{|\eta})_o \geq \kappa(g)\}} \, d\mu^*\kappa(g) \, d\beta(b'). \]
We estimate it using (5.4) and finally obtain
\[ I_{p,\eta} \leq \frac{2^p A}{1-a} \int_G (\kappa(g)^p + \kappa(g)) \, d\mu(g). \]
This completes the proof of Proposition 5.1. □

5.5. The free semigroup. In this subsection we describe an example showing that Proposition 5.1 is optimal.

Let \( G \) be a free group with two generators \( u, v \). It acts cocompactly by isometries on the corresponding Cayley graph \( (M, d) \), which is a regular tree of valence 4. We denote its base point by \( o = e \). The boundary \( \partial M \) is the space of those infinite words \( \xi = (\xi_1, \xi_2, \ldots) \) in the alphabet \( \{u, v, u^{-1}, v^{-1}\} \) that are irreducible, that is, \( \xi_{i+1} \neq \xi_i^{-1} \) for all \( i \geq 1 \).

We define a probability measure \( \mu \) on \( G \) by
\[ \mu = \frac{1}{2} \left( \delta_v + \sum_{n \geq 1} p_n \delta_{u^n} \right), \]
where \( \sum_{n \geq 1} p_n = 1 \). The support of the unique \( \mu \)-stationary probability measure \( \nu \) on the Gromov boundary \( \partial M \) is contained in the set of infinite words \( \xi \) in the alphabet \( \{u, v\} \).

Choose the following point \( \eta \) on \( \partial M \):
\[ \eta = (u, u, u, \ldots). \]
Then for every \( \xi = (\xi_1, \xi_2, \ldots) \) in \( \partial M \), the Gromov product is given by
\[ (\xi|\eta)_o = \inf\{i \geq 0 \mid \xi_{i+1} \neq u\}. \]
Let \( p > 0 \). We want to estimate the integral
\[ I_{p,\eta} := \int_{\partial M} (\xi|\eta)_o^p \, d\nu(\xi). \]
Example 5.4. In this case we have the following equivalence:

\[ \sum_{n \geq 1} p_n n^p < \infty \iff I_{p, \eta} < \infty. \]  

(5.8)

Hence the converse of Proposition 5.1 holds in this case.

**Proof.** Indeed, in this case the associated Bernoulli space \( B \) is the space of sequences \( b = (b_1, b_2, \ldots) \), where \( b_i = v \) or \( u^n \), with the Bernoulli measure \( \beta = \mu^{\otimes \mathbb{N}} \), and the image \( \xi_b \) of \( b \) under the boundary map is the concatenation of the letters \( u, v \) occurring in \( b \).

We denote by \( B_1 \) the set of sequences in \( B \) whose first letter is a power of \( u \) and whose second letter is \( v \). Then the following lower bound holds:

\[ I_{p, \eta} = \int_B (\xi_b \mid \eta)^p_0 \, d\beta(b) \geq \int_{B_1} (\xi_b \mid \eta)^p_0 \, d\beta(b) = \frac{1}{4} \sum_{n \geq 1} p_n n^p. \]

This proves the converse implication in (5.8).

**Bibliography**

[1] M. Björklund, “Central limit theorems for Gromov hyperbolic groups”, *J. Theoret. Probab.* 23:3 (2010), 871–887.
[2] S. Sawyer and T. Steger, “The rate of escape for anisotropic random walks in a tree”, *Probab. Theory Related Fields* 76:2 (1987), 207–230.
[3] F. Ledrappier, “Some asymptotic properties of random walks on free groups”, *Topics in probability and Lie groups: boundary theory*, CRM Proc. Lecture Notes, vol. 28, Amer. Math. Soc., Providence, RI 2001, pp. 117–152.
[4] Y. Benoist and J.-F. Quint, “Central limit theorem for linear groups”, *Ann. Probab.* (to appear).
[5] B. M. Brown, “Martingale central limit theorems”, *Ann. Math. Statist.* 42:1 (1971), 59–66.
[6] M. Gromov, “Hyperbolic groups”, *Essays in group theory*, Math. Sci. Res. Inst. Publ., vol. 8, Springer, New York 1987, pp. 75–263.
[7] E. Ghys and P. de la Harpe (eds.), *Sur les groupes hyperboliques d’aprè s Mikhael Gromov*, Progr. Math., vol. 83, Birkhäuser Boston, Inc., Boston, MA 1990.
[8] M. Coornaert and A. Papadopoulos, *Symbolic dynamics and hyperbolic groups*, Lecture Notes in Math., vol. 1539, Springer-Verlag, Berlin 1993.
[9] J. Väisälä, “Gromov hyperbolic spaces”, *Expo. Math.* 23:3 (2005), 187–231.
[10] I. Holopainen, U. Lang, and A. Väähäkangas, “Dirichlet problem at infinity on Gromov hyperbolic metric measure spaces”, *Math. Ann.* 339:1 (2007), 101–134.
[11] P.-E. Caprace, Y. de Cornulier, N. Monod, and R. Tessera, “Amenable hyperbolic groups”, *J. Europ. Math. Soc.* 17:11 (2015), 2903–2947.
[12] V. A. Kaimanovich, “The Poisson formula for groups with hyperbolic properties”, *Ann. of Math.* (2) 152:3 (2000), 659–692.
[13] P. Haïssinsky, *Marches aléatoires sur les groupes hyperboliques*, Notes de cours, 2011, http://web.univ-ubs.fr/lmam(matheus/emilfest.html.
[14] E. Le Page, “Théorèmes limites pour les produits de matrices aléatoires”, *Probability measures on groups* (Oberwolfach 1981), Lecture Notes in Math., vol. 928, Springer, Berlin–New York 1982, pp. 258–303.
[15] H. Furstenberg and Y. Kifer, “Random matrix products and measures on projective spaces”, Israel J. Math. 46:1-2 (1983), 12–32.

[16] Y. Guivarc’h and A. Raugi, “Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence”, Z. Wahrsch. Verw. Gebiete 69:2 (1985), 187–242.

[17] P. Bougerol and J. Lacroix, Products of random matrices with applications to Schrödinger operators, Progr. Probab. Statist., vol. 8, Birkhäuser Boston, Inc., Boston, MA 1985.

[18] A. Furman, “Random walks on groups and random transformations”, Handbook of dynamical systems, vol. 1A, North-Holland, Amsterdam 2002, pp. 931–1014.

[19] Y. Benoist and J.-F. Quint, Random walk on reductive groups, 2013, http://www.math.u-bordeaux1.fr/~jquint.

[20] J. M. Steele, “Kingman’s subadditive ergodic theorem”, Ann. Inst. H. Poincaré Probab. Statist. 25:1 (1989), 93–98.

[21] H. Kesten, “Sums of stationary sequences cannot grow slower than linearly”, Proc. Amer. Math. Soc. 49 (1975), 205–211.

[22] P. L. Hsu and H. Robbins, “Complete convergence and the law of large numbers”, Proc. Nat. Acad. Sci. U. S. A. 33:2 (1947), 25–31.

[23] L. E. Baum and M. Katz, “Convergence rates in the law of large numbers”, Trans. Amer. Math. Soc. 120 (1965), 108–123.

[24] P. Hall and C. C. Heyde, Martingale limit theory and its application, Acad. Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York–London 1980.

[25] C. Berg and J. P. R. Christensen, “Sur la norme des opérateurs de convolution”, Invent. Math. 23:2 (1974), 173–178.

[26] H. Kesten, “Symmetric random walks on groups”, Trans. Amer. Math. Soc. 92:2 (1959), 336–354.

[27] Y. Derriennic and Y. Guivarc’h, “Théorème de renouvellement pour les groupes non moyennables”, C. R. Acad. Sci. Paris Sér. A-B 277 (1973), A613–A615.

Yves Benoist
Paris-Sud University 11, France
E-mail: yves.benoist@math.u-psud.fr

Jean-François Quint
Université Bordeaux 1, France
E-mail: Jean-Francois.Quint@math.u-bordeaux1.fr

Received 3/APR/14
Edited by A. V. DOMRIN