

# Differentiating the entropy of random walks on hyperbolic groups

P. Mathieu \*

## Abstract

We show that the asymptotic entropy of a random walk on a non-elementary hyperbolic group, with symmetric and bounded increments, is differentiable and we identify its derivative as a correlation. We also prove similar results for the rate of escape.

## 1 Introduction

Let  $\Gamma$  be an infinite, countable, discrete group with neutral element  $id$ , and let  $\mu$  be a probability measure on  $\Gamma$ . The entropy of  $\mu$  is defined as

$$H(\mu) := \sum_{x \in \Gamma} (-\log \mu(x)) \mu(x).$$

Note that  $H(\mu)$  is non-negative and may be infinite.

Let  $\mu^n$  denote the  $n$ -th convolution power of  $\mu$ .

We assume that  $H(\mu) < \infty$ . It is easy to check that the sequence  $(H(\mu^n))_{n \in \mathbb{N}}$  is subadditive so that the following limit does exist:

$$h(\mu) := \lim_{n \rightarrow \infty} \frac{1}{n} H(\mu^n). \tag{1.1}$$

The quantity  $h(\mu)$  is called the **asymptotic entropy** of  $\mu$ .

The notion of asymptotic entropy was introduced by A. Avez in [2] in relation with random walk theory. In [3], Avez proved that, whenever  $h(\mu) = 0$ , then  $\mu$  satisfies the Liouville property: bounded,  $\mu$ -harmonic functions are constant. The converse was proved later, see [7] and [15].

Consider the random walk on  $\Gamma$  whose increments are distributed according to  $\mu$ . The Liouville property is equivalent to the triviality of the asymptotic  $\sigma$ -field of the random walk (its so-called Poisson boundary), see [7] and [15] again. In more general terms, the entropy plays a central role in the identification of the Poisson boundary of random walks in many examples. We refer in particular to [14] for groups with hyperbolic features. In this latter

---

\*Aix Marseille Université, CNRS, Centrale Marseille, LATP, UMR 7353, 13453 Marseille France, pierre.mathieu@cmi.univ-mrs.fr

case, the asymptotic entropy is also related to the geometry of the harmonic measure through a 'dimension-rate of escape- entropy' formula, see [6] and the references quoted therein.

The question of the regularity of  $h$  as a function of  $\mu$  was raised by A. Erschler and V. Kaimanovich in [9], where it is proved that, still for hyperbolic groups, the asymptotic entropy is continuous. If we restrict ourselves to measures  $\mu$  with fixed finite support, F. Ledrappier recently proved in [17] that  $h$  is Lipschitz continuous. We shall complement the result of [17] showing that, for a hyperbolic group  $\Gamma$  and restricting ourselves to symmetric measures  $\mu$  with a fixed finite support, the asymptotic entropy is differentiable (Theorem 2.2).

There is an analogy between the definition of  $h$  and the definition of the **rate of escape** of the random walk in some left-invariant metric. More precisely, it can be proved that  $h(\mu)$  coincides with the rate of escape of the corresponding random walk in the so-called Green metric, see the definition in part 2. We also give sufficient conditions on a metric ensuring that the rate of escape is differentiable (Theorem 2.1).

Our approach completely differs from the one in [17]. We start with the simple observation that the derivative of the mean position of a random walk is a correlation, see part 2.3. Thus the natural candidate to be the derivative of the rate of escape is some asymptotic covariance. These heuristics suggest a close connection between the differentiability of the rate of escape and the Central Limit Theorem and explain the statement of Theorem 2.1.

As for the entropy, one deduces the differentiability of  $h$  and the value of its derivative from Theorem 2.1 when choosing the right metric (namely the Green metric) and observing that fluctuations of the Green metric are of lower order - a fact that follows from the 'fundamental inequality' between entropy and rate of escape and which is true for random walks on non-amenable groups in general, see part 4.

The version of the Central Limit Theorem we need is a straightforward extension of [4]. We also rely a lot on the 'Green metric machinery' developed in [6].

Let us emphasize that we do not claim that our results are optimal. It is quite possible that the entropy and rate of escape are much more regular than differentiable. It is actually known that the entropy and rate of escape are analytic on the free group [16] and that the rate of escape is analytic for some Fuchsian groups [12]. One may hope analyticity holds for general hyperbolic groups (although it does not hold for all groups, see [18]). Anyway, we believe the interpretation of the derivative as a correlation is rather satisfactory, at least from an intuitive point of view. It clarifies the connection between the regularity of the rate of escape and the Central Limit Theorem, an observation that seems to be new in our context.

Let us finally mention that the interpretation of the derivative of a steady state (whatever it may mean) as some kind of correlation is a well known idea in theoretical physics or dynamical systems, where it is sometimes called 'linear response theory' or 'fluctuation-dissipation theory'. See [19] and other papers of the same author.

## 2 Definitions and results

### 2.1 Definitions

Let  $\Gamma$  be an infinite, countable, discrete group with neutral element  $id$ . Let  $d$  be a left-invariant proper metric on  $\Gamma$ . We assume that  $\Gamma$  is finitely generated and that  $d$  is equivalent to a word metric.

When the context is clear, we may also use the notation  $|x - y|$  to denote the distance between  $x$  and  $y$  and  $|x| = |x - id|$ .

We define the Gromov product of points  $x, y \in \Gamma$  with respect to the base point  $w \in \Gamma$  by

$$(x, y)_w := \frac{1}{2}(|x - w| + |y - w| - |x - y|).$$

We recall that the distance  $d$  is called **hyperbolic** if there exists a constant  $\tau$  such that

$$(x, y)_w \geq \min\{(x, z)_w, (z, y)_w\} - \tau, \tag{2.2}$$

for all  $x, y, z, w \in \Gamma$ .

The group  $\Gamma$  is called hyperbolic if any (equivalently some) word metric is hyperbolic. A hyperbolic group is called **non-elementary** if it is non amenable (which turns out to be equivalent to requiring  $\Gamma$  is not a finite extension of  $\mathbb{Z}$ ).

We refer to [10] for background material on hyperbolic groups.

From now on, we will assume that  $\Gamma$  is hyperbolic and non-elementary.

Following [6], we let  $\mathcal{D}(\Gamma)$  be the set of left-invariant proper metrics on  $\Gamma$  which are both equivalent to a word metric and hyperbolic. Note that these last two conditions are not redundant. Indeed there always exist non-hyperbolic (non-geodesic) metrics equivalent to any word metric on  $\Gamma$ , see [6], Proposition 2.3. There may also exist hyperbolic metrics on  $\Gamma$  that are not equivalent to a word metric.

We shall consider the following two compactifications of  $\Gamma$ . The **visual (Gromov) compactification** is obtained by considering all infinite sequences in  $\Gamma$  and identifying two such sequences, say  $(x_n)$  and  $(y_n)$ , if the Gromov product  $(x_n, y_n)_w$  tends to infinity. The **horofunction (Busemann) compactification** is constructed by identifying a point  $x \in \Gamma$  with the horo-function  $k_x : \Gamma \rightarrow \mathbb{R}$ ,  $k_x(y) = |y - x| - |w - x|$  and taking the closure for the topology of pointwise convergence. By Ascoli's theorem, this is indeed a compact. The group  $\Gamma$  acts by homeomorphisms on both compactifications. Up to equivariant homeomorphisms, the Gromov compactification does not depend on the choice of either the base point  $w$  or the choice of  $d \in \mathcal{D}(\Gamma)$ . The Busemann compactification is also independent of the choice of the base point but not of the choice of the metric. We shall say that  $d$  satisfies **Assumption (BA)** if, up to equivariant homeomorphisms, the Gromov and Busemann compactifications coincide.

Assumption **(BA)** is in particular satisfied by the class of metrics called **Green metrics**. These are constructed as follows. We call a probability measure  $\mu$  on  $\Gamma$  'symmetric' if  $\mu(x^{-1}) = \mu(x)$  for all  $x \in \Gamma$ . The support of  $\mu$  is the set of  $x \in \Gamma$  such that  $\mu(x)$  is not zero.

Let  $\mu$  be a probability measure on  $\Gamma$ . We assume that  $\mu$  is symmetric and that the support of  $\mu$  is finite and generates the whole group  $\Gamma$ . The Green function associated to  $\mu$  is defined by

$$G^\mu(x) = \sum_{n=0}^{\infty} \mu^n(x),$$

where  $\mu^n$  is the  $n$ -th convolution power of  $\mu$ .

We assumed that  $\Gamma$  is non-amenable. Therefore the sequence  $\mu^n(x)$  exponentially converges to zero so that the series defining  $G^\mu$  does converge. The Green distance between points  $x$  and  $y$  in  $\Gamma$  is then

$$d_G^\mu(x, y) := \log G^\mu(id) - \log G^\mu(x^{-1}y). \tag{2.3}$$

In [6], we proved that  $d_G^\mu$  belongs to  $\mathcal{D}(\Gamma)$ . Observe that  $d_G^\mu$  need not be geodesic. (As a matter of fact, it is not so difficult to deduce that  $d_G^\mu$  is equivalent to a word metric from the non-amenability of  $\Gamma$ . That  $d_G^\mu$  is hyperbolic is equivalent to a certain multiplicativity property of the Green function  $G^\mu$  which is the content of Ancona's classical - and difficult - theorem, see [1] and the proof in [21]).

We now give the definition of the **random walk** associated to a probability measure  $\mu$  on  $\Gamma$ . Because we will eventually use Radon-Nikodym transforms, it will be more convenient to work with the canonical construction on the set of trajectories on  $\Gamma$ , say  $\Omega = \Gamma^{\mathbb{N}}$ , where  $\mathbb{N} = \{0, 1, \dots\}$ . Given  $\omega = (\omega_0, \omega_1, \dots) \in \Omega$  and  $n \in \mathbb{N}$ , we define the maps  $Z_n$  and  $X_n$  from  $\Omega$  to  $\Gamma$  by  $Z_n(\omega) := \omega_n$ , and  $X_n(\omega) := (Z_{n-1}(\omega))^{-1}Z_n(\omega)$ . Thus  $Z_n(\omega)$  gives the position of the trajectory  $\omega$  at time  $n$ , while  $X_n(\omega)$  gives its increment also at time  $n$ . Following the usual usage in probability theory we often omit to indicate that random functions, as  $Z_n$  or  $X_n$ , depend on  $\omega$ .

We equip  $\Omega$  with the product  $\sigma$ -field (i.e. the smallest  $\sigma$ -field for which all functions  $Z_n$  are measurable). The law of the random walk with increments distributed like  $\mu$  is, by definition, the unique probability measure on  $\Omega$  under which  $Z_0 = id$  and the random variables  $(X_n)_{n \in \mathbb{N}}$  are independent and distributed like  $\mu$ . We denote it with  $\mathbb{P}^\mu$ . We also use the notation  $\mathbb{E}^\mu$  to denote the expectation with respect to  $\mathbb{P}^\mu$ . Observe that the law of  $Z_n$  under  $\mathbb{P}^\mu$  is  $\mu^n$ .

We recall that, given a probability measure  $\mu$  with finite support and given a left-invariant metric  $d \in \mathcal{D}(\Gamma)$ , the **rate of escape** of  $\mu$  in the metric  $d$  is defined by

$$\ell(\mu; d) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{x \in \Gamma} d(id, x) \mu^n(x)$$

and the **asymptotic entropy** of  $\mu$  is defined by

$$h(\mu) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{x \in \Gamma} (-\log \mu^n(x)) \mu^n(x).$$

Kingman's sub-additive theorem implies that

$$\ell(\mu; d) = \lim_{n \rightarrow \infty} \frac{1}{n} d(id, Z_n) \quad \text{and} \quad h(\mu) = \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mu^n(Z_n)$$

where both limits hold  $\mathbb{P}^\mu$  almost surely as well as in  $L^1(\Omega, \mathbb{P}^\mu)$ .

## 2.2 Differentiability of $\ell$ and $h$

In the sequel we shall study the derivatives of the rate of escape and the entropy of probability measures with a fixed finite support. A subset of  $\Gamma$ , say  $S$ , is 'symmetric' if  $x^{-1} \in S$  whenever  $x \in S$ . Let  $S$  be a finite symmetric subset of  $\Gamma$ . We assume that  $S$  generates the whole group  $\Gamma$ . Let  $\mathcal{P}_s(S)$  be the set of symmetric probability measures with support equal to  $S$ . Then  $\mathcal{P}_s(S)$  naturally identifies with an open subset of  $\mathbb{R}^d$  for some  $d$ . We use the differentiable structure it inherits this way.

**Regularity Assumption:** throughout the paper, we shall assume that the function  $\lambda \in [-1, 1] \rightarrow \log \mu_\lambda(a)$  has a derivative at  $\lambda = 0$  for all  $a \in S$ . Equivalently, we may write a first order expansion of  $\log \mu_\lambda(a)$  as  $\lambda$  tends to 0 in the form

$$\log \mu_\lambda(a) = \log \mu_0(a) + \lambda \nu(a) + \lambda o_\lambda(a), \tag{2.4}$$

where  $\nu(a)$  is the derivative of the function  $\lambda \rightarrow \log \mu_\lambda(a)$  at  $\lambda = 0$  and  $o_\lambda(a)$  converges to 0.

Observe that since  $S$  is finite, this is equivalent to requiring  $o_\lambda(a)$  to converge to 0 uniformly with respect to  $a \in S$ . We shall also repeatedly use the fact that  $\nu$  is bounded.

We shall use the shorthand notation  $\mathbb{P}^\lambda$  (resp.  $\mathbb{E}^\lambda$ ) instead of  $\mathbb{P}^{\mu_\lambda}$  (resp.  $\mathbb{E}^{\mu_\lambda}$ ).

From the condition that  $\mu_\lambda$  is a probability measure, one deduces that we must have

$$\sum_{a \in S} \nu(a) \mu_0(a) = 0.$$

We define the sequence  $M_0 = 0$  and, for  $n \geq 1$ ,

$$M_n = \sum_{j=1}^n \nu(X_j).$$

Note that the random process  $(M_n)_{n \in \mathbb{N}}$  is a centered martingale under  $\mathbb{P}^0$ .

Let  $d \in \mathcal{D}(\Gamma)$  and assume assumption **(BA)** holds. We shall see in Proposition 3.2 that the sequence  $(|Z_n|, M_n)$  satisfies a joint central limit theorem and that the asymptotic covariance of  $|Z_n|$  and  $M_n$  is given by

$$\sigma(\nu, \mu_0; d) := \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}^0[|Z_n| M_n]. \quad (2.5)$$

**Theorem 2.1** *Let  $d \in \mathcal{D}(\Gamma)$  satisfy **(BA)**. Then the function  $\lambda \rightarrow \ell(\mu_\lambda; d)$  is differentiable and its derivative satisfies*

$$\left. \frac{d}{d\lambda} \right|_{\lambda=0} \ell(\mu_\lambda; d) = \sigma(\nu, \mu_0; d). \quad (2.6)$$

**Theorem 2.2** *The function  $\lambda \rightarrow h(\mu_\lambda)$  has a derivative at  $\lambda = 0$  which satisfies*

$$\left. \frac{d}{d\lambda} \right|_{\lambda=0} h(\mu_\lambda) = \sigma(\nu, \mu_0; d_G^0), \quad (2.7)$$

where  $d_G^\lambda := d_G^{\mu_\lambda}$  is the Green metric associated to the probability  $\mu_\lambda$ .

In [5], we showed that the asymptotic entropy coincides with the rate of escape in the Green metric:

$$h(\mu) = \ell(\mu; d_G^\mu)$$

for all  $\mu \in \mathcal{P}_s(S)$ . Thus, using the result in Theorem 2.1, we may reformulate (2.7) as follows:

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} (h(\mu_\lambda) - h(\mu_0)) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} (\ell(\mu_\lambda; d_G^0) - \ell(\mu_0; d_G^0)).$$

In other words, as far as first order terms are concerned, the fluctuations of the Green metric do not contribute to the fluctuations of the entropy. This turns out to be a quite general statement for random walks on non-amenable groups, see part 4.

## 2.3 Heuristics

We give a simple - but not completely rigorous - way to guess why formula (2.6) should hold true. We provide these heuristics in order to clarify the scheme of the proofs, with the hope that this scheme can be adapted to other examples of random walks.

To compute the rate of escape, observe that

$$\mathbb{E}^\lambda[|Z_n|] = \mathbb{E}^0[|Z_n| \prod_{j=1}^n \frac{\mu_\lambda(X_j)}{\mu_0(X_j)}]. \quad (2.8)$$

Taking the derivative in (2.8), we get that

$$\frac{d}{d\lambda} \Big|_{\lambda=0} \mathbb{E}^\lambda[|Z_n|] = \mathbb{E}^0[|Z_n| \sum_{j=1}^n \nu(X_j)] = \mathbb{E}^0[|Z_n| M_n]. \quad (2.9)$$

Thus we see that a reasonable candidate to be the derivative of  $\ell(\mu_\lambda; d)$  is the limit of  $\frac{1}{n} \mathbb{E}^0[|Z_n| M_n]$  as  $n$  tends to  $+\infty$ .

Observe however that in order to turn this loose argument into a proof, one needs justify how to exchange the order between the limit in  $n$  and the derivation in  $\lambda$ .

On the one hand, we shall rely on a quantitative version of the law of large numbers for  $|Z_n|$  to show that the derivative of  $\ell(\mu_\lambda; d)$  is well approximated by the limit of the ratio  $(\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|])/\lambda n$  as soon as  $\lambda$  tend to  $0+$  and  $n$  tend to  $+\infty$  in such a way that the product  $\lambda n$  tends to  $+\infty$ , see Lemma 3.1. This follows from the fact that the function  $n \rightarrow \mathbb{E}^\lambda[|Z_n|]$  is almost additive, uniformly in  $\lambda$ .

Thus it is sufficient to describe the limit of  $\mathbb{E}^\lambda[|Z_n|]/\lambda n$  when  $\lambda n \rightarrow +\infty$ . We will actually choose  $\lambda$  and  $n$  such that  $\lambda\sqrt{n}$  tends to 1. Then  $\mathbb{E}^\lambda[|Z_n|]/\lambda n \sim \mathbb{E}^\lambda[|Z_n|]/\sqrt{n}$ , which is the scaling of the Central Limit Theorem.

More precisely, we show a joint C.L.T. for the random vector  $(|Z_n|, M_n)$  under  $\mathbb{P}^0$ , see Proposition 3.2. Let  $\sigma(\nu, \mu_0; d)$  denote the asymptotic covariance of  $|Z_n|$  and  $M_n$ .

Consider the Girsanov weight  $\prod_{j=1}^n \mu_\lambda(X_j)/\mu_0(X_j)$  in formula (2.8). Up to error terms of smaller order, it coincides with the exponential (in the sense of martingale theory) of the martingale  $(\lambda M_n)_{n \in \mathbb{N}}$ . With our choice of the scaling between  $\lambda$  and  $n$ , the asymptotic of  $\lambda M_n \sim M_n/\sqrt{n}$  is given by the Central Limit Theorem. Therefore the limit of the Girsanov weight is of the form  $e^{M - \frac{1}{2}\mathbb{E}[M^2]}$  for some Gaussian random variable  $M$ . Moreover, provided we can check some integrability conditions, the joint C.L.T. implies that the limit of  $\mathbb{E}^\lambda[|Z_n|] - n\ell(\mu_0; d)/\lambda n$  is then of the form  $\mathbb{E}[Z e^{M - \frac{1}{2}\mathbb{E}[M^2]}]$ , where  $(Z, M)$  is a Gaussian vector with covariance  $\mathbb{E}[ZM] = \sigma(\nu, \mu_0; d)$ . The integration by parts formula for Gaussian laws implies that (for any Gaussian vector)

$$\mathbb{E}[Z e^{M - \frac{1}{2}\mathbb{E}[M^2]}] = \mathbb{E}[ZM].$$

The next Theorem summarizes the part of this argument we just sketched that does not explicitly use the hyperbolicity of  $\Gamma$ .

**Theorem 2.3** *Let  $\Gamma$  be a finitely generated group;  $S$  a finite symmetric generating set;  $d$  a left-invariant proper metric on  $\Gamma$  and  $\lambda \in [-1, 1] \rightarrow \mu_\lambda \in \mathcal{P}_s(S)$  be a curve in  $\mathcal{P}_s(S)$  satisfying the Regularity Assumption. We further assume that:*

(i) the joint Central Limit Theorem holds for the vector  $(|Z_n|, M_n)$  under  $\mathbb{P}^0$  with asymptotic covariance  $\sigma$ ,  
(ii)

$$\sup_n \frac{1}{n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))^2] < +\infty.$$

Then we have

$$\lim_{n \rightarrow +\infty, \lambda \rightarrow 0} \frac{1}{\lambda n} (\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]) = \sigma,$$

along any sequence  $\lambda$  such that  $\limsup_{n \rightarrow +\infty} \lambda^2 n < +\infty$ .

Theorem 2.3 is proved in section 3.3.

### 3 Proofs of Theorems 2.1 and 2.3

The proofs are organized in the following way: in part 3.1, we recall some estimates on the mean distance  $\mathbb{E}^\mu[|Z_n|]$  from [6] and use them to show that Theorem 2.1 can be deduced from Theorem 2.3. In part 3.2, we recall results from [4] and show how they imply a slightly stronger version of the assumptions of Theorem 2.3. Parts 3.1 and 3.2 use the hyperbolicity of  $\Gamma$  in an essential way.

In part 3.3 we prove Theorem 2.3. Part 3.3 can be read independently of the preceding ones.

Let  $d \in \mathcal{D}(\Gamma)$  satisfy assumption **(BA)**. We will show that  $\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} (\ell(\mu_\lambda; d) - \ell(\mu_0; d))$  exists and equals  $\sigma(\nu, \mu_0; d)$ . In the proof it will be convenient to restrict ourselves to positive  $\lambda$ 's. This is no loss of generality since  $\sigma(\nu, \mu_0; d)$  is linear in  $\nu$ .

#### 3.1 Geometric input

**Lemma 3.1** *Let  $\lambda$  tend to  $0+$  and  $n$  tend to  $+\infty$  in such a way that the product  $\lambda n$  tends to  $+\infty$ . Then*

$$\frac{\ell(\mu_\lambda; d) - \ell(\mu_0; d)}{\lambda} - \frac{\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]}{\lambda n}$$

tends to 0.

*Proof:* Let  $\mu \in \mathcal{P}_s(S)$ .

The triangle inequality implies that the sequence  $a(n) := \mathbb{E}^\mu[|Z_n|]$  is sub-additive. Therefore we have  $\ell(\mu; d) = \inf_n \frac{a(n)}{n}$  and thus

$$\mathbb{E}^\mu[|Z_n|] \geq n\ell(\mu; d) \tag{3.10}$$

for all  $n$ .

We need a similar upper bound. It will follow from bounds on the lateral deviation of a trajectory of a random walk. Let us recall some results from [6].

In [6], Proposition 3.8, we showed that, for any  $\mu \in \mathcal{P}_s(S)$ , there exists a constant  $\tau_0$  such that for all integers  $m, n, k$ ,

$$\mathbb{E}^\mu[(Z_m, Z_{m+n+k})_{Z_{m+n}}] \leq \tau_0. \quad (3.11)$$

Applying this inequality with  $m = 0$  and using the fact that  $\mathbb{E}^\mu[|Z_{n+k} - Z_n|] = \mathbb{E}^\mu[|Z_k|]$ , we get that

$$\mathbb{E}^\mu[|Z_n|] + \mathbb{E}^\mu[|Z_k|] \leq 2\tau_0 + \mathbb{E}^\mu[|Z_{n+k}|]. \quad (3.12)$$

Thus the sequence  $b(n) := 2\tau_0 - \mathbb{E}^\mu[|Z_n|]$  is also sub-additive. Note that  $b(n)/n$  converges to  $-\ell(\mu; d)$ . As above this implies that

$$\mathbb{E}^\mu[|Z_n|] \leq n\ell(\mu; d) + 2\tau_0 \quad (3.13)$$

for all  $n$ .

Combining (3.10) and (3.13), we see that we have proved that

$$|\mathbb{E}^\mu[|Z_n|] - n\ell(\mu; d)| \leq 2\tau_0 \quad (3.14)$$

for all  $n$ .

A close inspection of the proof of (3.11), reveals that the constant  $\tau_0$  is locally uniform in  $\mathcal{P}_s(S)$  so that we may apply (3.14) with the same constant  $\tau_0$  to all measures  $\mu_\lambda$  for  $\lambda$  in a small enough neighborhood of 0. The statement of Lemma 3.1 immediately follows.  $\blacksquare$

Let us choose  $\lambda$  tending to 0+ and  $n$  tending to  $+\infty$  such that  $\lambda^2 n$  tends to 1. Thus Lemma 3.1 applies. In order to complete the proof of Theorem 2.1, it only remains to show that

$$\lim_{\lambda, n} \frac{\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]}{\lambda n} = \sigma(\nu, \mu_0; d). \quad (3.15)$$

## 3.2 Central Limit Theorems

In this part of the paper we recall some results from [4] on the Central Limit Theorem for  $|Z_n|$  and discuss their extension to a joint C.L.T. for  $(|Z_n|, M_n)$ .

Let  $d \in \mathcal{D}(\Gamma)$  satisfy assumption **(BA)**. Let  $\mu \in \mathcal{P}_s(S)$  be a finitely supported symmetric probability measure on  $\Gamma$ .

Mimicking the situation of the discussion preceding Theorem 2.1, we also let  $\nu$  be a real valued function defined on  $S$  and satisfying the centering condition:  $\sum_{a \in S} \nu(a)\mu(a) = 0$  and consider the sequence of random variables  $M_0 = 0$  and, for  $n \geq 1$ ,

$$M_n = \sum_{j=1}^n \nu(X_j).$$

### Proposition 3.2 .

(i) *The law of the two-dimensional random vector  $((|Z_n| - n\ell(\mu; d))/\sqrt{n}, M_n/\sqrt{n})$  under  $\mathbb{P}^\mu$  weakly converges as  $n$  tends to  $+\infty$  to a centered Gaussian law with some covariance matrix  $\Sigma^\mu$ .*

(ii) *The covariance matrix of  $((|Z_n| - n\ell(\mu; d))/\sqrt{n}, M_n/\sqrt{n})$  under  $\mathbb{P}^\mu$  converges to  $\Sigma^\mu$ . In particular, the sequence  $\frac{1}{n}\mathbb{E}^\mu[|Z_n|M_n]$  converges as  $n$  tends to  $+\infty$  and its limit is the non-diagonal term of  $\Sigma^\mu$ .*

*Proof:*

We recall the following classical version of the martingale central limit theorem: (see [13]).

**Lemma 3.3** *Let  $(\zeta_n)_{n \in \mathbb{N}}$  be a square integrable, centered martingale with respect to a filtration  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ , with stationary increments. Assume that*

$$\frac{1}{n} \sum_{j=1}^n \mathbb{E}[(\zeta_j - \zeta_{j-1})^2 | \mathcal{F}_{j-1}] \rightarrow \sigma^2 \quad (3.16)$$

*almost surely, where  $\sigma^2$  is a deterministic real. Then the law of  $(\frac{1}{\sqrt{n}}\zeta_n)$  weakly converges to a centered Gaussian law with variance  $\sigma^2$ .*

*Proof of Proposition 3.2:*

*Step 1:*

Following [4], we first prove a version of Proposition 3.2 where  $|Z_n| - n\ell(\mu; d)$  is replaced by an appropriate martingale approximation that we denote with  $(\chi_n)$ .

Let  $\partial\Gamma$  the Busemann boundary of  $\Gamma$ . We recall that the Gromov product extends to the boundary.

Under  $\mathbb{P}^\mu$ , almost any trajectory  $(Z_n)_{n \in \mathbb{N}}$  converges to a limiting point in  $\partial\Gamma$ , say  $Z_\infty$ . This follows from assumption **(BA)** since then  $\partial\Gamma$  can be identified with the Gromov boundary of  $\Gamma$  and one knows that random walk paths almost surely converge in the Gromov compactification of a hyperbolic group, see [1] or [14]. The law of  $Z_\infty$  is called the 'harmonic measure'. We denote it with  $\xi^\mu$ .

In [4] part 4, it is proved that there exists a bounded function  $u$  on  $\partial\Gamma$  such that the sequence

$$\chi_n := k(Z_n) - n\ell(\mu; d) + u(k) - u(Z_n^{-1}k)$$

is a martingale under  $\mathbb{P}^\mu$  for any  $k \in \partial\Gamma$ .

In the sequel, we shall assume that  $k$  is chosen according to the harmonic measure  $\xi^\mu$  and independent of the walk  $(Z_n)_{n \in \mathbb{N}}$ . It then follows that  $(\chi_n)_{n \in \mathbb{N}}$  has stationary increments. It is also proved in Theorem 9 in [4] that the Lindeberg condition (3.16) is satisfied.

**Remark 3.4** *The group  $\Gamma$  has a natural action on its boundary. For each  $k \in \partial\Gamma$ , the sequence  $(Z_n^{-1}k)_{n \in \mathbb{N}}$  is a Markov chain with values in  $\partial\Gamma$  started at  $k$ . This Markov chain has a unique invariant probability measure, namely the harmonic measure  $\xi^\mu$ . Moreover, the Markov chain  $(Z_n^{-1}k)_{n \in \mathbb{N}}$  has nice mixing properties: see Lemma 4 in [4]. By choosing  $k$  according to  $\xi^\mu$ , we are actually considering the chain in a stationary regime.*

*We shall not explicitly use these remarks but they lie at the heart of the proof of the central limit theorem in [4].*

On the other hand, the sequence  $(M_n)_{n \in \mathbb{N}}$  being a sum of independent bounded and centered random variables, is also a centered martingale with stationary increments under  $\mathbb{P}^\mu$  satisfying condition (3.16).

Thus we deduce from Lemma 3.3 that the law of the vector  $(\chi_n/\sqrt{n}, M_n/\sqrt{n})$  under  $\mathbb{P}^\mu$  converges to a centered Gaussian vector. Indeed, one may apply Lemma 3.3 to the martingale  $(a\chi_n + bM_n)_{n \in \mathbb{N}}$  for any  $a, b \in \mathbb{R}$ .

Let  $\Sigma^\mu$  be the limit covariance matrix. For  $(a, b) \in \mathbb{R}^2$ , we use the notation  $\Sigma^\mu(a, b)$  to denote the value of the quadratic form associated to  $\Sigma^\mu$  evaluated at  $(a, b)$ . We observe that since both martingales  $(\chi_n)$  and  $(M_n)$  have stationary increments, then  $\Sigma^\mu$  is also the covariance (under  $\mathbb{P}^\mu$ ) of the vector  $(\chi_n/\sqrt{n}, M_n/\sqrt{n})$  for any  $n \geq 1$ , that is

$$\frac{1}{n} \mathbb{E}^\mu[(a\chi_n + bM_n)^2] = \Sigma^\mu(a, b). \quad (3.17)$$

*Step 2:*

In the above claims, we wish to replace  $\chi_n$  by  $|Z_n| - n\ell(\mu; d)$ . We shall use the following

**Lemma 3.5** *There exists a constant  $C$  such that, for all  $D$  we have*

$$\mathbb{P}^\mu[(k, Z_n)_{id} \geq D] \leq C^{-1}e^{-CD}. \quad (3.18)$$

*Inequality (3.18) holds uniformly in  $k \in \partial\Gamma$ .*

*Proof of Lemma 3.5:*

The statement of the Lemma actually directly follows from arguments in [6].

One may for instance split the event  $(k, Z_n)_{id} \geq D$  into two, say  $A := ((k, Z_n)_{id} \geq D) \cap ((Z_\infty, Z_n)_{id} \geq \frac{D}{2})$  and  $B := ((k, Z_n)_{id} \geq D) \cap ((Z_\infty, Z_n)_{id} < \frac{D}{2})$ .

Let us show that  $\mathbb{P}^\mu[A] + \mathbb{P}^\mu[B] \leq C^{-1}e^{-CD}$ .

In the argument below  $\tau_1$  is a constant that depends on  $d$  and  $\mu$  only. We choose  $D$  large enough and how large depends only on the choice of the metric  $d$  and the measure  $\mu$ . In particular, we assume that  $D \geq 4\tau$ , where  $\tau$  is the hyperbolicity constant from (2.2).

Hyperbolicity implies that, on  $A$ , we also have  $(k, Z_\infty)_{id} \geq \frac{D}{2} - \tau_1 \geq \frac{D}{2}$ . We know from [6] Proposition 3.10 that  $\xi^\mu$  satisfies the doubling condition. Therefore the probability that  $(k, Z_\infty)_{id} \geq \frac{D}{2}$  can be compared to the harmonic measure of a ball of  $\partial\Gamma$  of radius of order  $e^{-C_1D}$  for some  $C_1$ , and since  $\xi^\mu$  is Ahlfors regular, see Theorem 1.1 in [6], we get that  $\mathbb{P}^\mu[A] \leq C^{-1}e^{-CD}$ .

On the event  $B$ , we have  $(Z_\infty, Z_n)_{id} < \frac{D}{2}$  and  $|Z_n| \geq D - \tau \geq \frac{3}{4}D$ . Therefore the distance between  $Z_n$  and any quasiruler from  $id$  to  $Z_\infty$  is larger than  $\frac{D}{4} - \tau_1$ . For large enough  $D$ , this last event has a probability bounded from above by  $C^{-1}e^{-CD}$  for some  $C$  as follows from the deviation inequality in Proposition 3.8 in [6]. Therefore  $\mathbb{P}^\mu[B] \leq C^{-1}e^{-CD}$ .  $\blacksquare$

Back to the proof of Proposition 3.2, we observe that  $k(x) = |x| - 2(k, x)_{id}$  for all  $k \in \partial\Gamma$  and  $x \in \Gamma$ . Therefore

$$|Z_n| - n\ell(\mu; d) - \chi_n = 2(k, Z_n)_{id} - (u(k) - u(Z_n^{-1}k)).$$

Using Lemma 3.5 and the fact that  $u$  is bounded, we get that

$$\mathbb{P}^\mu[||Z_n| - n\ell(\mu; d) - \chi_n| \geq D] \leq C^{-1}e^{-CD} \quad (3.19)$$

for some constant  $C$ .

As a by-product of (3.19), we get that  $\frac{1}{\sqrt{n}}(|Z_n| - n\ell(\mu; d) - \chi_n)$  converges to 0 in probability. Therefore the two sequences of vectors  $((|Z_n| - n\ell(\mu; d))/\sqrt{n}, M_n/\sqrt{n})$  and  $(\chi_n/\sqrt{n}, M_n/\sqrt{n})$  have the same limit in law. Thus we have proved that  $((|Z_n| - n\ell(\mu; d))/\sqrt{n}, M_n/\sqrt{n})$  satisfies the central limit theorem with asymptotic variance  $\Sigma^\mu$ .

We also deduce from (3.19) that

$$\sup_n \mathbb{E}^\mu[(|Z_n| - n\ell(\mu; d) - \chi_n)^2] < \infty, \quad (3.20)$$

and

$$\frac{1}{n} \mathbb{E}^\mu[(|Z_n| - n\ell(\mu; d) - \chi_n)^2] \rightarrow 0.$$

Combining (3.20) with (3.17), we see that the covariance matrix of  $((|Z_n| - n\ell(\mu; d))/\sqrt{n}, M_n/\sqrt{n})$  under  $\mathbb{P}^\mu$  converges to  $\Sigma^\mu$ . Also

$$\frac{1}{n} \mathbb{E}^\mu[|Z_n| M_n] = \frac{1}{n} \mathbb{E}^\mu[(|Z_n| - n\ell(\mu; d)) M_n]$$

converges as  $n$  tends to  $+\infty$  and its limit is the non-diagonal term of  $\Sigma^\mu$ . This concludes the proof of the Proposition. ■

For further references, we observe that (3.20) with (3.17) implies the following

**Lemma 3.6** *For all  $\mu \in \mathcal{P}_s(S)$ , we have*

$$\sup_n \frac{1}{n} \mathbb{E}^\mu[(|Z_n| - n\ell(\mu; d))^2] < \infty.$$

### 3.3 Proof of Theorem 2.3 and (3.15)

Here  $\Gamma$  is any finitely generated group.

We recall that we are assuming the joint Central Limit Theorem for the vector  $(|Z_n|, M_n)$  under  $\mathbb{P}^0$ ; namely the law of the two-dimensional random vector  $((|Z_n| - n\ell(\mu; d))/\sqrt{n}, M_n/\sqrt{n})$  under  $\mathbb{P}^0$  weakly converges as  $n$  tends to  $+\infty$  to a centered Gaussian law with some covariance matrix  $\Sigma$ . Let  $\sigma$  be the non-diagonal element of  $\Sigma$ .

We also assume that

$$\sup_n \frac{1}{n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))^2] < +\infty. \quad (3.21)$$

We wish to prove that

$$\lim_{n \rightarrow +\infty, \lambda \rightarrow 0} \frac{1}{\lambda n} (\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]) = \sigma, \quad (3.22)$$

along any sequence  $\lambda$  such that  $\limsup_{n \rightarrow +\infty} \lambda^2 n < +\infty$ . Without loss of generality, we may and will assume that  $\lambda^2 n$  converges to some limit  $\alpha \geq 0$ . Then  $\lambda \sim \sqrt{\alpha/n}$ .

We start dealing with the case  $\alpha \neq 0$ .

First note that

$$\frac{\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]}{\lambda n} = \frac{\mathbb{E}^\lambda[|Z_n| - n\ell(\mu_0; d)] - \mathbb{E}^0[|Z_n| - n\ell(\mu_0; d)]}{\lambda n}.$$

From the Central Limit Theorem for  $|Z_n|$  under  $\mathbb{E}^0$  and assumption (3.21), we get that

$$\frac{\mathbb{E}^0[|Z_n| - n\ell(\mu_0; d)]}{\lambda n} \sim \frac{\mathbb{E}^0[|Z_n| - n\ell(\mu_0; d)]}{\sqrt{\alpha}\sqrt{n}} \rightarrow 0.$$

In order to compute the limit of  $\mathbb{E}^\lambda[|Z_n| - n\ell(\mu_0; d)]/\lambda n$ , we write these terms in a form that is more amenable to the application of the Central Limit Theorem.

We have

$$\frac{1}{\lambda n} \mathbb{E}^\lambda[|Z_n| - n\ell(\mu_0; d)] = \frac{1}{\lambda n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d)) \prod_{j=1}^n \frac{\mu_\lambda(X_j)}{\mu_0(X_j)}]. \quad (3.23)$$

Let  $a \in S$ . Recall that we have a first order expansion of the function  $\lambda \rightarrow \mu_\lambda(a)$  in the form

$$\log \frac{\mu_\lambda(a)}{\mu_0(a)} = \lambda \nu(a) + \lambda o_\lambda(a), \quad (3.24)$$

where the function  $o_\lambda$  uniformly converges to 0 as  $\lambda$  goes to 0.

Because  $\mu_\lambda$  is a probability for all  $\lambda$ , it follows from (3.24) that  $\nu$  and  $o_\lambda$  must satisfy the following centering conditions:

$$\sum_{a \in S} \nu(a) \mu_0(a) = 0 \quad (3.25)$$

$$\text{and } \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \sum_{a \in S} (o_\lambda(a) + \frac{\lambda}{2} \nu^2(a)) \mu_0(a) = 0. \quad (3.26)$$

To see why (3.25) and (3.26) hold, note that  $\sum_{a \in S} \mu_\lambda(a) = 1$  for all  $\lambda$ . The expansion of  $\sum_{a \in S} \mu_\lambda(a)$  in terms of  $\lambda$  starts with  $\lambda \sum_{a \in S} \nu(a) \mu_0(a) + \lambda^2 \frac{1}{\lambda} \sum_{a \in S} (o_\lambda(a) + \frac{\lambda}{2} \nu^2(a)) \mu_0(a)$ , the rest being of order smaller than  $\lambda^2$ . Dividing by  $\lambda$  and letting  $\lambda$  tend to 0, one gets (3.25). Then dividing by  $\lambda^2$  and letting  $\lambda$  tend to 0, one gets (3.26).

Let us re-write (3.23) as

$$\frac{1}{\lambda n} \mathbb{E}^\lambda[|Z_n| - n\ell(\mu_0; d)] = \frac{1}{\lambda n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d)) e^{\lambda M_n - \lambda^2 A_n + R_n^\lambda}], \quad (3.27)$$

where

$$M_n := \sum_{j=1}^n \nu(X_j), \quad A_n := \frac{1}{2} \sum_{j=1}^n \nu^2(X_j) \quad \text{and} \quad R_n^\lambda = \lambda \sum_{j=1}^n o_\lambda(X_j) + \frac{\lambda}{2} \nu^2(X_j).$$

From the Law of Large Numbers for sums of i.i.d. random variables, it follows that  $\frac{1}{n} A_n$  almost surely converges to  $\frac{1}{2} \sum_{a \in S} \nu^2(a) \mu_0(a)$  and therefore  $\lambda^2 A_n$  converges to  $\frac{\alpha}{2} \sum_{a \in S} \nu^2(a) \mu_0(a)$ . We claim that  $R_n^\lambda$  converges to 0 in probability under  $\mathbb{P}^0$ .

The argument for this last claim goes as follows. Let  $Y_j^\lambda := \frac{1}{\lambda} o_\lambda(X_j) + \frac{1}{2} \nu^2(X_j)$ , so that  $R_n^\lambda = \lambda^2 \sum_{j=1}^n Y_j^\lambda$ . For a fixed  $n$ , the random variables  $(Y_j^\lambda)_{j=1}^n$  are independent and equally distributed. On the one hand (3.26) implies that  $\mathbb{E}^0[Y_1^\lambda]$  tends to 0 and therefore  $\mathbb{E}^0[R_n^\lambda]$  also converges to 0. On the other hand, the variance of  $Y_j^\lambda$  satisfies  $\lim_{\lambda \rightarrow 0} \lambda^2 \mathbb{V}^0[Y_1^\lambda] = 0$ . (Use the

fact that  $o_\lambda$  converges to 0.) Therefore the variance of  $R_n^\lambda$  is of lower order than  $\lambda^2 n$  and tends to 0. Thus we get that both the mean and variance of  $R_n^\lambda$  converge to 0.

We now use the Central Limit Theorem for  $((|Z_n| - n\ell(\mu_0; d))/\sqrt{n}, M_n/\sqrt{n})$ .

Ignoring for a moment that the function we integrate in (3.27) is not bounded, we pass to the limit using the relation  $\lambda^2 n \rightarrow \alpha$  and get that

$$\frac{1}{\lambda n} \mathbb{E}^\lambda[|Z_n| - n\ell(\mu_0; d)] \rightarrow \frac{1}{\sqrt{\alpha}} \mathbb{E}[Z e^{\sqrt{\alpha}M - \frac{\alpha}{2} \sum_{a \in S} \nu^2(a) \mu_0(a)}], \quad (3.28)$$

where  $(Z, M)$  is a centered Gaussian vector with variance  $\Sigma$ .

Observe that the variance of  $M_n$  equals  $n \sum_{a \in S} \nu^2(a) \mu_0(a)$  and therefore  $\sum_{a \in S} \nu^2(a) \mu_0(a) = \mathbb{E}[M^2]$ . Thus the right hand side of (3.28) equals

$$\frac{1}{\sqrt{\alpha}} \mathbb{E}[Z e^{\sqrt{\alpha}M - \frac{\alpha}{2} \mathbb{E}[M^2]}].$$

The integration by parts formula for Gaussian laws implies that (for any Gaussian vector and any  $\alpha$ )

$$\frac{1}{\sqrt{\alpha}} \mathbb{E}[Z e^{\sqrt{\alpha}M - \frac{\alpha}{2} \mathbb{E}[M^2]}] = \mathbb{E}[ZM] = \sigma.$$

Thus we are done with the proof of Theorem 2.1 once we justify that we may indeed pass to the limit in (3.27). In order to do so, it is sufficient to have bounds on the moments of the functions we integrate.

Hölder's inequality implies that

$$\begin{aligned} & \mathbb{E}^0 \left[ \left( \frac{1}{\lambda n} (|Z_n| - n\ell(\mu_0; d)) e^{\lambda M_n - \lambda^2 A_n + R_n^\lambda} \right)^{6/5} \right] \\ & \leq \mathbb{E}^0 \left[ \left( \frac{1}{\lambda n} (|Z_n| - n\ell(\mu_0; d)) \right)^{27/10} \mathbb{E}^0 \left[ e^{3\lambda M_n - 3\lambda^2 A_n + 3R_n^\lambda} \right]^{4/10} \right]. \end{aligned}$$

We already know from assumption (3.21) that  $\frac{1}{(\lambda n)^2} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))^2]$  is bounded in  $n$ . Let us prove that  $\mathbb{E}^0[e^{3\lambda M_n - 3\lambda^2 A_n + 3R_n^\lambda}]$  is also bounded in  $n$ .

Note that there exists a constant  $C$  such that

$$\begin{aligned} & \mathbb{E}^0[e^{3\lambda M_n - 3\lambda^2 A_n + 3R_n^\lambda}] \leq \mathbb{E}^0[e^{3\lambda M_n + 3R_n^\lambda}] \\ & = e^{3\mathbb{E}^0[R_n^\lambda]} \mathbb{E}^0[e^{3\lambda M_n + 3(R_n^\lambda - \mathbb{E}^0[R_n^\lambda])}] \leq e^C \mathbb{E}^0[e^{3\lambda M_n + 3(R_n^\lambda - \mathbb{E}^0[R_n^\lambda])}]. \end{aligned}$$

(We used the fact that  $\mathbb{E}^0[R_n^\lambda]$  is bounded for the last inequality.)

From the independence of the  $X_j$ 's follows that

$$\mathbb{E}^0[e^{3\lambda M_n + 3(R_n^\lambda - \mathbb{E}^0[R_n^\lambda])}] = \mathbb{E}^0[e^{3\lambda \nu(X_1) + 3\lambda^2 (Y_1^\lambda - \mathbb{E}^0[Y_1^\lambda])}]^n.$$

But the random variables  $\nu(X_1) + \lambda(Y_1^\lambda - \mathbb{E}^0[Y_1^\lambda])$  are centered and bounded (uniformly in  $\lambda$ ) i.e. there exists a number  $M$  such that  $|\nu(X_1) + \lambda(Y_1^\lambda - \mathbb{E}^0[Y_1^\lambda])| \leq M$  for all  $\lambda$  and all trajectory  $\omega$ . Therefore  $\mathbb{E}^0[e^{3\lambda \nu(X_1) + 3\lambda^2 (Y_1^\lambda - \mathbb{E}^0[Y_1^\lambda])}]^n$  is bounded whenever  $\lambda^2 n$  is also bounded. For this last step, we rely on the following classical Lemma, see part 7 of [8] for instance.

**Lemma 3.7** For all  $M$  and  $K$  there exist constants  $C_M$  and  $n_0$  s.t. for all random variable  $X$  with  $|X| \leq M$  and  $\mathbb{E}[X] = 0$  and for all  $\lambda$  and  $n \geq n_0$  s.t.  $\lambda^2 n \leq K$  then

$$\mathbb{E}[e^{\lambda X}]^n \leq e^{C_M}.$$

*Proof of Lemma 3.7:* Since  $X$  is bounded, the log-Laplace transform

$$\Lambda(\lambda) := \log \mathbb{E}[e^{\lambda X}]$$

is analytic in a neighborhood of 0. More precisely, we have:

Let  $y \in \mathbb{C}$ . Then

$$|e^y - 1 - y| \leq |e^{|y|} - 1 - |y|| \leq |y|^2 e^{|y|}.$$

Apply this to  $\lambda X$ ,  $\lambda \in \mathbb{C}$ :

$$|\mathbb{E}[e^{\lambda X}] - 1| \leq |\lambda|^2 M^2 e^{|\lambda|M} \leq \frac{1}{2},$$

if  $|\lambda| \leq \lambda_0$ . So  $|\mathbb{E}[e^{\lambda X}]| \geq \frac{1}{2}$  for  $|\lambda| \leq \lambda_0$ . So  $\Lambda$  is analytic in  $\{\lambda \text{ s.t. } |\lambda| \leq \lambda_0\}$  and  $|\Lambda(\lambda)| \leq c_0$  for some constant  $c_0$ , for all  $\lambda$  s.t.  $|\lambda| \leq \lambda_0$ . Note that  $\lambda_0$  and  $c_0$  depend only on  $M$ .

We have  $\Lambda(0) = 0$  and  $\Lambda'(0) = 0$  (because  $X$  is centered). So the function  $\lambda^{-2}\Lambda(\lambda)$  is also analytic in  $\{\lambda \text{ s.t. } |\lambda| \leq \lambda_0\}$ . By the maximum principle, for any  $\lambda$  such that  $|\lambda| \leq \lambda_0$ , we have

$$|\Lambda(\lambda)| \leq C|\lambda|^2$$

where  $C = \max_{z; |z|=\lambda_0} \frac{\Lambda(z)}{z^2} \leq \frac{c_0}{\lambda_0^2}$ .

The statement of the Lemma is thus proved with  $n_0$  chosen such that  $K/n_0 \leq \lambda_0^2$  and  $C_M = c_0 K/\lambda_0^2$ . ■

This completes the proof of (3.22) in the case  $\alpha \neq 0$ .

The case  $\alpha = 0$  is easier. As we did in (3.23) and (3.27), we start writing that

$$\frac{1}{\lambda n} (\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]) = \frac{1}{\lambda n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))(e^{\lambda M_n - \lambda^2 A_n + R_n^\lambda} - 1)].$$

Using similar arguments as for the case  $\alpha \neq 0$ , it is not difficult to show that

$$\begin{aligned} & \lim_{n \rightarrow +\infty, \lambda \rightarrow 0, \lambda^2 n \rightarrow 0} \frac{1}{\lambda n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))(e^{\lambda M_n - \lambda^2 A_n + R_n^\lambda} - 1)] \\ &= \lim_{n \rightarrow +\infty, \lambda \rightarrow 0, \lambda^2 n \rightarrow 0} \frac{1}{\lambda n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))\lambda M_n e^{-\lambda^2 A_n + R_n^\lambda}]. \end{aligned}$$

Observe that

$$\frac{1}{\lambda n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))\lambda M_n e^{-\lambda^2 A_n + R_n^\lambda}] = \frac{1}{n} \mathbb{E}^0[(|Z_n| - n\ell(\mu_0; d))M_n e^{-\lambda^2 A_n + R_n^\lambda}].$$

The limit of this last expression is given by the Central Limit Theorem and, with the notation we already used, it coincides with  $\mathbb{E}[ZM] = \sigma$ . Observe that, with our scaling satisfying  $\lambda^2 n \rightarrow 0$ , we have  $\lambda^2 A_n \rightarrow 0$ . The details are similar to the case  $\alpha \neq 0$ . ■

**End of the proof of Theorem 2.1**

Proposition 3.2 and Lemma 3.6 from part 3.2 show that the assumptions of Theorem 2.3 are satisfied. Thus we get that

$$\lim_{n \rightarrow +\infty, \lambda \rightarrow 0, \lambda^2 n \rightarrow 1} \frac{1}{\lambda n} (\mathbb{E}^\lambda[|Z_n|] - \mathbb{E}^0[|Z_n|]) = \sigma(\nu, \mu_0; d).$$

But we observed in part 3.1 that this convergence implies Theorem 2.1. ■

## 4 Proof of Theorem 2.2

We now explain how to deduce Theorem 2.2 from Theorem 2.1.

As in the proof of Theorem 2.1, we may and will restrict ourselves to positive  $\lambda$ 's.

We first recall that the entropy can be interpreted as a rate of escape in the Green metric:  $h(\mu) = \ell(\mu; d_G^\mu)$ . Thus we have:

$$h(\mu_\lambda) - h(\mu_0) = (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) + (\ell(\mu_\lambda; d_G^0) - \ell(\mu_0; d_G^0)). \quad (4.29)$$

By Theorem 2.1 and since  $d_G^0$  satisfies **(BA)**, once divided by  $\lambda$ , the second term in (4.29) converges to  $\sigma(\nu, \mu_0; d_G^0)$ . Thus the proof of Theorem 2.2 will be complete once we prove that

$$\lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) = 0. \quad (4.30)$$

It turns out the estimate (4.30) does not use the hyperbolicity of  $\Gamma$ . We have the following more general property:

**Proposition 4.1** *Let  $\Gamma$  be a finitely generated group. Assume  $\Gamma$  is non-amenable. Let  $\mu_0$  be a probability measure on  $\Gamma$  such that the support of  $\mu_0$  generates  $\Gamma$  (as a semi-group), and  $H(\mu_0) < \infty$ .*

*Consider a curve of probability measures on  $\Gamma$ , say  $\lambda \in [-1, 1] \rightarrow \mu_\lambda$ , satisfying the Regularity Assumption:*

$$\log \mu_\lambda(a) = \log \mu_0(a) + \lambda \nu(a) + \lambda o_\lambda(a),$$

*where  $\nu$  is bounded and  $o_\lambda(a)$  converges to 0 uniformly in  $a \in \Gamma$ .*

*Then*

$$\lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) = 0.$$

**Remark 4.2** *We do not assume any more that  $\mu_0$  or the  $\mu_\lambda$ 's are symmetric. Then the Green metric may not be a real distance. Indeed, although it still satisfies the triangle inequality, it may not be symmetric.*

*Thus in this part of the paper, the word 'metric' will refer to a function on  $\Gamma \times \Gamma$  that vanishes on the diagonal and satisfies the triangle inequality.*

*The interpretation of the asymptotic entropy as the rate of escape in the Green metric remains valid in this general framework, see [5].*

**Remark 4.3** *The assumption that  $H(\mu_0) < \infty$  implies that  $\mu_0$  has a finite first moment with respect to  $d_G^0$  (See Lemma 2.3 in [5]). The Regularity Assumption then implies that  $H(\mu_\lambda) < \infty$  and that  $\mu_\lambda$  also has a finite first moment with respect to  $d_G^0$ .*

*Proof:* We give two separate arguments for lower and upper bounds for  $h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)$ . Both arguments are based on the so-called 'fundamental inequality' that we first recall: let  $\mu$  be a probability measure on  $\Gamma$  with finite entropy and let  $d$  a left-invariant proper metric on  $\Gamma$ . We denote with  $v(d)$  the logarithmic volume growth of the metric  $d$ . The 'fundamental inequality' states that  $h(\mu) \leq v(d)\ell(\mu; d)$ , see [11], [20], [5] and the references quoted therein.

The 'fundamental inequality' in particular applies to any Green metrics  $d_G^\alpha$ . By a result in [5], we have  $v(d_G^\alpha) = 1$ . Therefore we get that

$$h(\mu_\lambda) = \ell(\mu_\lambda; d_G^\lambda) \leq \ell(\mu_\lambda; d_G^\alpha), \quad (4.31)$$

for all  $\lambda$  and  $\alpha$ .

Applying (4.31) with  $\alpha = 0$ , yields  $h(\mu_\lambda) \leq \ell(\mu_\lambda; d_G^0)$  and therefore

$$\limsup_{\lambda \rightarrow 0^+} \frac{1}{\lambda} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) \leq 0 \quad (4.32)$$

It remains to prove the lower bound. We first need review properties of the Green metric. Consider a probability measure  $\mu$  with finite entropy and whose support generates the whole group  $\Gamma$ . We recall that we defined the Green metric as

$$d_G^\mu(x, y) := \log G^\mu(id) - \log G^\mu(x^{-1}y),$$

where  $G^\mu(x) = \sum_{n=0}^{\infty} \mu^n(x)$  is the Green function of the random walk.

We may equivalently express  $d_G^\mu$  in terms of the hitting probabilities of the random walk: for a given trajectory  $\omega \in \Omega$  and  $z \in \Gamma$ , let

$$T_z(\omega) = \inf\{n \geq 0; Z_n(\omega) = z\}$$

be the hitting time of  $z$  by  $\omega$ . Observe that  $T_z(\omega)$  may be infinite.

Define  $F^\mu(z) := \mathbb{P}^\mu[T_z < \infty]$ . Then

$$d_G^\mu(id, z) = -\log F^\mu(z),$$

as can be easily checked using the Markov property.

In the sequel, we use the notation  $F^\lambda$  instead of  $F^{\mu_\lambda}$ .

**Lemma 4.4** *The function  $(\lambda, \alpha) \rightarrow \ell(\mu_\lambda; d_G^\alpha)$  is bounded on  $[0, 1]^2$ :*

$$\sup_{0 \leq \lambda \leq 1; 0 \leq \alpha \leq 1} \ell(\mu_\lambda; d_G^\alpha) < \infty.$$

*Proof of Lemma 4.4*

Let  $\mu$  and  $\mu'$  be probability measures on  $\Gamma$  and let  $d$  be a proper left-invariant metric on  $\Gamma$ . It is clear that

$$\ell(\mu; d) \leq \sum_{a \in \Gamma} d(id, a) \mu(a).$$

Also  $F^{\mu'}(a) \geq \mu'(a)$ , and therefore  $d_G^{\mu'}(id, a) \leq -\log \mu'(a)$ .

Applying these two inequalities to  $\mu = \mu_\lambda$  and  $\mu' = \mu_\alpha$ , we get that

$$\begin{aligned} \ell(\mu_\lambda; d_G^\alpha) &\leq - \sum_{a \in \Gamma} (\log \mu_\alpha(a)) \mu_\lambda(a) \\ &= - \sum_{a \in \Gamma} (\log \mu_0(a) + \alpha \nu(a) + \alpha o_\alpha(a)) \mu_\lambda(a). \end{aligned}$$

Since  $\nu$  and  $o_\lambda$  are bounded, we have  $\mu_\lambda(a) \leq e^C \mu_0(a)$  for some constant  $C$ . For the same reason, the term  $\sum_{a \in \Gamma} (\alpha \nu(a) + \alpha o_\alpha(a)) \mu_\lambda(a)$  is also controlled by a constant. Therefore

$$\ell(\mu_\lambda; d_G^\alpha) \leq e^C H(\mu_0) + C,$$

for some constant  $C$ . ■

We shall need the following estimate on  $T_z$ :

**Lemma 4.5** *Let  $\mu \in \mathcal{P}_s(S)$ . Then there exists a positive constant  $\kappa$  such that*

$$\sup_{z \in \Gamma} \mathbb{E}^\mu [e^{\kappa T_z}; T_z < \infty] < \infty.$$

*Proof of Lemma 4.5*

We use the non-amenability of  $\Gamma$ : there exists a constant  $\rho < 1$  such that, for all  $n$  and all  $z \in \Gamma$ , we have  $\mu^n(z) \leq \rho^n$ . Therefore

$$\begin{aligned} \mathbb{E}^\mu [e^{\kappa T_z}; T_z < \infty] &= \sum_n e^{\kappa n} \mathbb{P}^\mu [T_z = n] \\ &\leq \sum_n e^{\kappa n} \mathbb{P}^\mu [Z_n = z] = \sum_n e^{\kappa n} \mu^n(z) \\ &\leq \sum_n e^{\kappa n} \rho^n < \infty \end{aligned}$$

as soon as  $e^\kappa \rho < 1$ . ■

*End of the proof of Proposition 4.1: the lower bound.*

We use the shorthand notation  $F^\lambda(z) := F^{\mu^\lambda}(z)$ .

As in (3.23), we have

$$F^\lambda(z) = \mathbb{E}^0 \left[ \prod_{j=1}^{T_z} \frac{\mu_\lambda(X_j)}{\mu_0(X_j)}; T_z < \infty \right].$$

Let us apply Hölder's inequality with positive parameters  $(p, q, r)$  such that  $1/p + 1/q + 1/r = 1$  and with the notation  $\alpha = \lambda q$ . We assume that  $\alpha \leq 1$ . Thus

$$F^\lambda(z) \leq F^0(z)^{1/p} F^\alpha(z)^{1/q} \mathbb{E}^0 \left[ \left( \prod_{j=1}^{T_z} \frac{\mu_\lambda(X_j)}{\mu_0(X_j)} \left( \frac{\mu_0(X_j)}{\mu_\alpha(X_j)} \right)^{1/q} \right)^r; T_z < \infty \right]^{1/r}. \quad (4.33)$$

Let  $a \in S$ . Using the Regularity Assumption and the relation  $\alpha = \lambda q$ , we get

$$\frac{\mu_\lambda(a)}{\mu_0(a)} \left( \frac{\mu_0(a)}{\mu_\alpha(a)} \right)^{1/q} = e^{\lambda(o_\lambda(a) - o_\alpha(a))}.$$

Therefore, since  $\lambda \leq \alpha \leq 1$ , and remembering that  $o_\lambda$  uniformly converges to 0, we see that for all  $\varepsilon > 0$ , provided  $\alpha$  is small enough then

$$\frac{\mu_\lambda(a)}{\mu_0(a)} \left( \frac{\mu_0(a)}{\mu_\alpha(a)} \right)^{1/q} \leq \exp(\varepsilon \lambda). \quad (4.34)$$

Using (4.34) in equation (4.33), we see that

$$F^\lambda(z) \leq F^0(z)^{1/p} F^\alpha(z)^{1/q} \mathbb{E}^0[e^{\varepsilon \lambda r T_z}; T_z < \infty].$$

If we further assume that  $\varepsilon \lambda r \leq \kappa^0$ , where  $\kappa^0$  is the constant given by Lemma 4.5 when choosing  $\mu = \mu_0$ , then we have

$$F^\lambda(z) \leq C F^0(z)^{1/p} F^\alpha(z)^{1/q}, \quad (4.35)$$

for a new constant  $C$  that does not depend on  $z$ .

We evaluate inequality (4.35) at  $z = Z_n$ ; take the logarithm and take the expectation with respect to  $\mathbb{P}^\lambda$  to obtain

$$\mathbb{E}^\lambda[d_G^\lambda(id, Z_n)] \geq \frac{1}{p} \mathbb{E}^\lambda[d_G^0(id, Z_n)] + \frac{1}{q} \mathbb{E}^\lambda[d_G^\alpha(id, Z_n)] - \log C.$$

Now divide by  $n$  and let  $n$  tend to  $\infty$ , to get that

$$h(\mu_\lambda) = \ell(\mu_\lambda; d_G^\lambda) \geq \frac{1}{p} \ell(\mu_\lambda; d_G^0) + \frac{1}{q} \ell(\mu_\lambda; d_G^\alpha). \quad (4.36)$$

We choose  $r = \kappa^0/(\varepsilon \lambda)$  and  $\lambda$  small enough so that  $1/r + 1/q < 1$ . Then (4.36) becomes

$$h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0) \geq \frac{\lambda}{\alpha} (\ell(\mu_\lambda; d_G^\alpha) - \ell(\mu_\lambda; d_G^0)) - \frac{\varepsilon \lambda}{\kappa^0} \ell(\mu_\lambda; d_G^0). \quad (4.37)$$

We let  $\lambda$  tend to 0 in (4.37): by Lemma 4.4, we know that

$$\lambda (\ell(\mu_\lambda; d_G^\alpha) - \ell(\mu_\lambda; d_G^0)) \rightarrow 0$$

and  $\lambda \ell(\mu_\lambda; d_G^0) \rightarrow 0$ . Therefore

$$\liminf_{\lambda \rightarrow 0^+} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) \geq 0. \quad (4.38)$$

Using the inequality  $\ell(\mu_\lambda; d_G^\alpha) \geq h(\mu_\lambda)$  (which comes from the 'fundamental inequality'), we deduce from (4.37) that

$$\frac{1}{\lambda} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) \geq \frac{1}{\alpha} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) - \frac{\varepsilon}{\kappa^0} \ell(\mu_\lambda; d_G^0). \quad (4.39)$$

It follows from Lemma 4.4 that there exists a constant  $\ell_0$  such that  $\frac{1}{\kappa^0} \ell(\mu_\lambda; d_G^0) \leq \ell_0$  for all  $\lambda$ . By (4.38), the term  $h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)$  has a non-negative lim inf. Thus we deduce from (4.39) that

$$\liminf_{\lambda \rightarrow 0^+} \frac{1}{\lambda} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) \geq -\varepsilon \ell_0. \quad (4.40)$$

And since (4.40) holds for any small enough  $\varepsilon$ , we have

$$\liminf_{\lambda \rightarrow 0^+} \frac{1}{\lambda} (h(\mu_\lambda) - \ell(\mu_\lambda; d_G^0)) \geq 0. \quad (4.41)$$

■

**Remark 4.6** *F. Ledrappier and L. Shu recently adapted our strategy in a continuous setting: using martingales as here, they obtained differentiability results for the entropy and rate of escape of Brownian motions on the universal cover of negatively curved manifolds, see <http://front.math.ucdavis.edu/1309.5182>.*

## References

- [1] A. Ancona, Théorie du potentiel sur les graphes et les variétés. Ecole d'été de St-Flour XVIII, 1988, *Lecture notes in maths* **1427** (1990) 1–112.
- [2] A. Avez, Entropie des groupes de type fini. *C. R. Acad. Sc. Paris Sér A-B* **275** (1972), A 1363-A1366.
- [3] A. Avez, Théorème de Choquet-Deny pour les groupes à croissance non exponentielle. *C. R. Acad. Sci. Paris Sér. A* **279** (1974), 2528.
- [4] M. Bjorklund, Central limit theorems for Gromov hyperbolic groups. *J. Theoret. Probab.* **23** (2010), 871–887.
- [5] S. Blachère, P. Haissinsky, P. Mathieu, Asymptotic entropy and Green speed for random walks on countable groups. *Ann. Probab.* **36** (2008), 1134–1152.
- [6] S. Blachère, P. Haissinsky, P. Mathieu, Harmonic measures versus quasiconformal measures for hyperbolic groups. *Ann. Scient. Ec. Norm. Sup. série 4*, **t. 44** (2011), 683–721.
- [7] Y. Derriennic, Quelques applications du théorème ergodique sous-additif. *Astérisque* **74** (1980), 183-201.
- [8] R. Dobrushin, Perturbation methods of the theory of Gibbsian fields. Lectures on probability theory and statistics (Saint-Flour, 1994), *Lecture Notes in Math.* **1648**, Springer, Berlin (1996).

- [9] A. Erschler, V. Kaimanovich, Continuity of asymptotic characteristics for random walks on hyperbolic groups. *Funct. Anal. Appl.* **47** (2013), 152-156.
- [10] E. Ghys, P. de la Harpe (éditeurs), Sur les groupes hyperboliques d'après Mikhael Gromov, Progress in Mathematics, **83**, Birkhäuser Boston, Inc., Boston, MA, 1990.
- [11] Y. Guivarc'h, Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire. In *Conference on random walks (Kleebach, 1979)*, *Astérisque* **74** Soc. Math. France (1980), 47-98.
- [12] P. Haissinsky, P. Mathieu, S. Mueller, Renewal theory for random walks on surface groups. *Preprint 2013*. <http://front.math.ucdavis.edu/1304.7625>
- [13] I. Helland, Central limit theorems for martingales with discrete or continuous time. *Scand. Journ. Stat.* **9** (1982), 79-94.
- [14] V. Kaimanovich, The Poisson formula for groups with hyperbolic properties. *Ann. of Math.* **152** (2000), 659-692.
- [15] V. Kaimanovich, A. Vershik, Random walks on discrete groups: boundary and entropy. *Ann. Probab.* **11**(1983), 457-490.
- [16] F. Ledrappier, Analyticity of the entropy for some random walks. *Groups Geom. Dyn.* **6** (2012), 317-333.
- [17] F. Ledrappier, Regularity of the entropy for random walks on hyperbolic groups. *Ann. Probab.* **41** (2013), 3582-3605.
- [18] J. Mairesse, F. Mathéus, Randomly growing braids on three strands and the manta ray. *Ann. Appl. Probab.* **17** (2007), 502-536.
- [19] D. Ruelle, A review of linear response theory for general differentiable dynamical systems. *Nonlinearity* **22** (2009), 855-870.
- [20] A. Vershik, Dynamic theory of growth in groups: entropy, boundaries, examples. *Russian Mth. Surveys* **55** (2000), 667-733.
- [21] W. Woess, Random walks on infinite graphs and groups. Cambridge tracts in mathematics, **138**, Cambridge university press, Cambridge (2000).