ENTROPY AND DRIFT IN WORD HYPERBOLIC GROUPS

SÉBASTIEN GOUËZEL, FRÉDÉRIC MATHÉUS, FRANÇOIS MAUCOURANT

Abstract. The fundamental inequality of Guivarc'h relates the entropy and the drift of random walks on groups. It is strict if and only if the random walk does not behave like the uniform measure on balls. We prove that, in any nonelementary hyperbolic group which is not virtually free, endowed with a word distance, the fundamental inequality is strict for symmetric measures with finite support, uniformly for measures with a given support. This answers a conjecture of S. Lalley. For admissible measures, this is proved using previous results of Ancona and Blachère-Haïssinsky-Mathieu. For non-admissible measures, this follows from a counting result, interesting in its own right: we show that, in any infinite index subgroup, the proportion of non-distorted points is exponentially small. The uniformity is obtained by studying the behavior of measures that degenerate towards a measure supported on an elementary subgroup.

1. Main results

Let Γ be a finitely generated infinite group. Although the following discussion makes sense in a much broader context, we will assume that Γ is hyperbolic since all results of this article are devoted to this setting. There are two natural ways to construct random elements in Γ:

- Let d be a proper left-invariant distance on Γ (for instance a word distance). For large n, one can pick an element at random with respect to the uniform measure ρ_n on the ball $B_n = B(e, n)$ (where e denotes the identity of Γ).
- Let μ be a probability measure on Γ. For large n, one can pick an element at random with respect to the measure μ^{*n} (the n-th convolution of the measure μ). Equivalently, let $g_1, g_2, \dots$ be a sequence of random elements of Γ that are distributed independently according to μ. Form the random walk $X_n = g_1 \cdots g_n$. Then the distribution of X_n is μ^{*n}.

From a theoretical point of view, these methods share a lot of properties. From a computational point of view, the second method is much easier to implement in general groups since it does not require the computation of the ball B_n (note however that, in hyperbolic groups, simulating the uniform measure is very easy thanks to the automatic structure of the group). It is therefore of interest to find probability measures μ such that these two methods give equivalent results, in a sense that will be made precise below. This is the main question of Vershik in [Ver00]. In free groups (with the word distance coming from the usual set of generators), everything can be computed: if μ is the uniform measure on the generators, then μ^{*n} and ρ_n behave essentially in the same way. The situation is the

Date: January 20, 2015.

2010 *Mathematics Subject Classification.* 05C81 20F67 60J50 60B15.

same in free products of finite groups, again thanks to the underlying tree structure. However, in more complicated groups, explicit computations are essentially impossible, and it is expected that the methods always differ. Our main result confirms this intuition in a special class of groups: In hyperbolic groups which are not virtually free (i.e., there is no finite index free subgroup), if d is a word distance, the two methods are always different, in a precise quantitative way.

Remark 1.1. We emphasize that the question really depends on the choice of the distance d, since the shape of the balls B_n depends on d. For instance, for any symmetric probability measure μ on Γ whose support is finite and generates Γ, there exists a distance d (called the Green distance, see [BHM11]) for which the measures ρ_n and μ^{*n} behave in the same way. A famous open problem (to which our methods do not apply) is to understand what happens when Γ acts cocompactly on the hyperbolic space $\mathbb{H}^k$, and the distance d is given by $d(e,\gamma) = d_{\mathbb{H}^k}(O, \gamma \cdot O)$ where O is a base point in $\mathbb{H}^k$. In this case, it is also expected that the two methods are always different. Here are the main partial results in this context:

(1) The two methods are different for some symmetric measures with finite support ([LP07], see also Theorem 5.9 below).
(2) If, instead of a cocompact lattice, one considers a lattice with cusps, the two methods are always different [GLJ93].
(3) If, instead of a lattice, one considers a nice dense subgroup, there exist symmetric measures with finite support for which the two methods are equivalent [Bou12].

This question also makes sense in continuous time, for negatively curved manifolds. A conjecture of Sullivan asserts that, in this setting, the two methods coincide if and only if the manifold is locally symmetric, see [Led95].

One can give several meanings to the question "are the two methods equivalent?" Let us first discuss an interpretation in terms of behavior at infinity. The measures μ^{*n} converge in the geometric compactification $\Gamma \cup \partial\Gamma$ to a measure μ_∞, supported on the boundary, called the exit measure of the random walk, or its stationary measure. Geometrically, the random walk $(X_n)_{n\geqslant 1}$ converges almost surely to a random point on the boundary $\partial\Gamma$, the measure μ_∞ is its distribution. On the other hand, let ρ_∞ be the Patterson-Sullivan measure on $\partial\Gamma$ associated to the distance d, constructed in [Coo93] in this context. One should think of it as the uniform measure on the boundary (it is equivalent to the Hausdorff measure of maximal dimension on the boundary, for any visual distance coming from d). The measures ρ_n do not always converge to ρ_∞, but all their limit points are equivalent to ρ_∞, with a density bounded from above and from below (this follows from the arguments of [Coo93], see Lemma 2.13 below). A version of the question is then to ask if the measures μ_∞ and ρ_∞ are mutually singular: in this case, the random walk mainly visits parts of the groups that are not important from the point of view of the uniform measure.

Another version of the same question is quantitative: Does the random walk visit parts of the groups that are exponentially negligible from the point of view of the uniform measure? This is made precise through the notions of drift and entropy. Define

$$L(\mu) = \sum_{g\in\Gamma} \mu(g)|g|, \quad H(\mu) = \sum_{g\in\Gamma} \mu(g)(-\log \mu(g)), \tag{1.1}$$

where $|g| = d(e, g)$. The quantity $L(\mu)$ is the average distance of an element to the identity. The quantity $H(\mu)$, called the time one entropy of μ, is the average logarithmic weight of the points. They can both be finite or infinite. The functions L and H both behave in a subadditive way with respect to convolution: $L(\mu_1 * \mu_2) \leqslant L(\mu_1) + L(\mu_2)$ and $H(\mu_1 * \mu_2) \leqslant H(\mu_1) + H(\mu_2)$. It follows that the sequences $L(\mu^{*n})$ and $H(\mu^{*n})$ are subadditive. Hence, the following quantities are well defined:

$$\ell(\mu) = \lim L(\mu^{*n})/n, \quad h(\mu) = \lim H(\mu^{*n})/n. \tag{1.2}$$

They are called respectively the drift and the asymptotic entropy of the random walk. They also admit characterizations along typical trajectories. If $L(\mu)$ is finite, then almost surely $\ell(\mu) = \lim|X_n|/n$. In the same way, if $H(\mu)$ is finite, then almost surely $h(\mu) = \lim(-\log \mu^{*n}(X_n))/n$. The most intuitive characterization of the entropy is probably the following one: at time n, the random walk is essentially supported by $e^{h(\mu)n}$ points (see Lemma 2.4 for a precise statement). Let us also define the exponential growth rate of the group with respect to d, i.e.,

$$v = \liminf_{n\to\infty} \frac{\log|B_n|}{n}, \tag{1.3}$$

where B_n is the ball of radius n around e. In hyperbolic groups, it satisfies the apparently stronger inequality $C^{-1}e^{nv} \leqslant |B_n| \leqslant Ce^{nv}$, by [Coo93]. For large n, most points for μ^{*n} are contained in a ball $B_{(1+\varepsilon)\ell n}$, which has cardinality at most $e^{(1+2\varepsilon)\ell n v}$. Since the random walk at time n essentially visits e^{hn} points, we deduce the fundamental inequality of Guivarc'h [Gui80]

$$h \leqslant \ell v.$$

If this inequality is an equality, this means that the walk visits most parts of the group. Otherwise, it is concentrated in an exponentially small subset. Another version of our main question is therefore: Is the inequality $h \leqslant \ell v$ strict?

In hyperbolic groups, it turns out that the two versions of the question are equivalent, at least for finitely supported measures, and that they also have a geometric interpretation in terms of Hausdorff dimension. If μ is a probability measure on a group, we write Γ_μ^+ for the semigroup generated by the support of μ, and Γ_μ for the group it generates. When μ is symmetric, they coincide. We say that μ is admissible if $\Gamma_\mu^+ = \Gamma$. The following result is Corollary 1.4 and Theorem 1.5 in [BHM11] (see also [Haï13]) when the measure is symmetric, and is proved in [Tan14] when μ is not necessarily symmetric and d is a word distance.

Theorem 1.2. *Let Γ be a non-elementary hyperbolic group, endowed with a left-invariant distance d which is hyperbolic and quasi-isometric to a word distance. Let v be the exponential growth rate of (Γ, d). Let $d_{\partial\Gamma}$ be a visual distance on $\partial\Gamma$ associated to d. Consider an admissible probability measure μ on Γ, with finite support. Assume additionally either that the measure μ is symmetric, or that the distance d is a word distance. The following conditions are equivalent:*

1. *The equality $h = \ell v$ holds.*
2. *The Hausdorff dimension of the exit measure μ_∞ on $(\partial\Gamma, d_{\partial\Gamma})$ is equal to the Hausdorff dimension of this space.*

(3) *The measure* μ_∞ *is equivalent to the Patterson-Sullivan measure* ρ_∞.
(4) *The measure* μ_∞ *is equivalent to the Patterson-Sullivan measure* ρ_∞, *with density bounded from above and from below.*
(5) *There exists* $C > 0$ *such that, for any* $g \in \Gamma$,
$$|vd(e,g) - d_\mu(e,g)| \leqslant C,$$
where d_μ *is the "Green distance" associated to* μ, *i.e.,* $d_\mu(e,g) = -\log \mathbb{P}(\exists n, X_n = g)$ *where* X_n *is the random walk given by* μ *starting from the identity (it is an asymmetric distance in general, and a genuine distance if* μ *is symmetric).*

The different statements in this theorem go from the weakest to the strongest: since entropy is an asymptotic quantity, an assumption on h seems to allow subexponential fluctuations, so the assumption (1) is rather weak. On the other hand, (3) says that two measures are equivalent, so most points are controlled. Finally, in (5), all points are uniformly controlled. The equivalence between these statements is a strong rigidity theorem. The equivalence between (1) and (2) follows from a formula for the respective dimensions. The definition of a visual distance at infinity $d_{\partial\Gamma}$ involves a small parameter ε. In terms of this parameter, one has $HD(\mu_\infty) = h/(\varepsilon\ell)$ and $HD(\rho_\infty) = HD(\partial\Gamma) = v/\varepsilon$, so that these dimensions coincide if and only if $h = \ell v$.

In this theorem, the finite support assumption can be weakened to an assumption of superexponential moment (i.e., for all $M > 0$, $\sum_{g\in\Gamma} \mu(g) e^{M|g|} < \infty$), thanks to [Gou13]. The assumption that μ is symmetric or that d is a word distance is probably not necessary. However, the most important assumption in Theorem 1.2 is admissibility: it ensures that the random walk can see the geometry of the whole group (which is hyperbolic). For a random walk living in a strict (maybe distorted) subgroup, one would not be expecting the same nice behavior.

Our main theorem follows. It states that, in hyperbolic groups which are not virtually free, endowed with a word distance, the different equivalent conditions of Theorem 1.2 are never satisfied, uniformly on measures with a fixed support.

Theorem 1.3. *Let* Γ *be a hyperbolic group which is not virtually free, endowed with a word distance* d. *Let* Σ *be a finite subset of* Γ. *There exists* $c < 1$ *such that, for any symmetric probability measure* μ *supported in* Σ,
$$h(\mu) \leqslant c\ell(\mu)v,$$
where v *is the exponential growth rate of balls in* (Γ, d).

This theorem gives a positive answer to a conjecture of S. Lalley [Lal14, slide 16]. In the language of Vershik [Ver00], this theorem says that no finite subset of Γ is extremal. On the other hand, if one lets Σ grow, h/ℓ can converge to v:

Theorem 1.4. *Let* Γ *be a hyperbolic group, endowed with a left invariant distance* d *which is hyperbolic and quasi-isometric to a word distance. Let* ρ_i *be the uniform measure on the ball of radius* i. *Then* $h(\rho_i)/\ell(\rho_i) \to v$, *where* v *is the exponential growth rate of balls in* (Γ, d).

More precisely, we prove that $\ell(\rho_i) \sim i$ and $h(\rho_i) \sim iv$. The only difficulty is to prove the lower bound on $h(\rho_i)$: since h is defined in (1.2) using a subadditive sequence, upper bounds are automatic, but to get lower bounds one should show that additional cancellations do not happen later on. This difficulty already appears in [EK13], where the authors prove that the entropy depends continuously on the measure. Our proof of Theorem 1.4, given in Paragraph 2.5, also applies to this situation and gives a new proof of their result, under slightly weaker assumptions. There is nothing special about the uniform measure on balls, our proof also gives the same conclusion for the uniform measure on spheres, or for the measures $\sum e^{-s|g|}\delta_g / \sum e^{-s|g|}$ when $s \searrow v$.

Our main result is Theorem 1.3. It is a consequence of the three following results. Since their main aim is Theorem 1.3, they are designed to handle finitely supported symmetric measures. However, these theorems are all valid under weaker assumptions, which we specify in the statements as they carry along implicit information on the techniques used in the proofs.

The first result deals with admissible (or virtually admissible) measures.

Theorem 1.5. *Let Γ be a hyperbolic group which is not virtually free, endowed with a word distance. Let μ be a probability measure with a superexponential moment, such that Γ_μ^+ is a finite index subgroup of Γ. Then $h(\mu) < \ell(\mu)v$.*

The second result deals with non-admissible measures.

Theorem 1.6. *Let Γ be a hyperbolic group endowed with a word distance. Let μ be a probability measure with a moment of order 1 (i.e., $L(\mu) < \infty$). Assume that $\ell(\mu) > 0$ and that Γ_μ has infinite index in Γ. Then $h(\mu) < \ell(\mu)v$.*

Finally, the third result is a kind of continuity statement, to get the uniformity.

Theorem 1.7. *Let Γ be a hyperbolic group, endowed with a left-invariant distance which is hyperbolic and quasi-isometric to a word distance. Let Σ be a subset of Γ which does not generate an elementary subgroup. There exists a probability measure μ_Σ with finite support such that $\ell(\mu_\Sigma) > 0$ and*

$$\sup\{h(\mu)/\ell(\mu) \,:\, \mu \text{ probability}, \mathrm{Supp}(\mu) \subset \Sigma, \ell(\mu) > 0\} = h(\mu_\Sigma)/\ell(\mu_\Sigma).$$

The same statement holds if the maximum is taken over symmetric probability measures, the resulting maximizing measure being symmetric.

Theorem 1.3 is a consequence of these three statements.

Proof of Theorem 1.3 using the three auxiliary theorems. As in the statement of the theorem, consider a finite subset Σ of Γ. If Σ generates an elementary subgroup of Γ, all measures supported on Σ have zero entropy. Hence, one can take $c = 0$ in the statement of the theorem. Otherwise, by Theorem 1.7, there exists a symmetric measure μ_Σ with finite support that maximizes the quantity $h(\mu)/\ell(\mu)$ over μ symmetric supported by Σ. If $\Gamma_{\mu_\Sigma} = \Gamma_{\mu_\Sigma}^+$ has finite index, $h(\mu_\Sigma)/\ell(\mu_\Sigma) < v$ by Theorem 1.5. If it has infinite index, the same conclusion follows from Theorem 1.6. □

The three auxiliary theorems are non-trivial. Their proofs are independent, and use completely different tools. Here are some comments about them.

- At first sight, Theorem 1.5 seems to be the most delicate (this is the only one with the assumption that Γ is not virtually free). However, this is also the setting that has been mostly studied in the literature. Hence, we may use several known results, including most notably results of Ancona [Anc87], of Blachère, Haïssinsky and Mathieu [BHM11] and Tanaka [Tan14] (Theorem 1.2 above) and of Izumi, Neshveyev and Okayasu [INO08] on rigidity results for cocycles. The proof relies mainly on the fact that the word distance is integer valued, contrary to the Green distance (more precisely, we use the fact that the stable translation length of hyperbolic elements is rational with bounded denominator).
- In Theorem 1.6, the difficulty comes from the lack of information on the subgroup Γ_μ. If it has good geometric properties (for instance if it is quasi-convex), one may use the same kind of techniques as for Theorem 1.5. Otherwise, the random walk does not really see the hyperbolicity of the ambient group. The fundamental inequality always gives $h \leqslant \ell v_{\Gamma_\mu}$, where v_{Γ_μ} is the growth rate of the subgroup Γ_μ (for the initial word distance on Γ). If $v_{\Gamma_\mu} < v$, the result follows. Unfortunately, there exist non-quasi-convex subgroups of some hyperbolic groups with the same growth as the ambient group. However, a random walk does not typically visit all points of Γ_μ, it concentrates on those points that are not distorted (i.e., their distances to the identity in Γ and Γ_μ are comparable). To prove Theorem 1.6, we will show that in any infinite index subgroup of a hyperbolic group, the number of non-distorted points is exponentially smaller than e^{nv}.
- Theorem 1.7 is less simple than it may seem at first sight: it does not claim that μ_Σ is supported by Σ, and indeed this is not the case in general (see Example 5.4). Hence, the proof is not a simple continuity argument: We need to understand precisely the behavior of sequences of measures that degenerate towards a measure supported on an elementary subgroup. The proof will show that μ_Σ is supported by $K\cdot(\Sigma\cup\{e\})\cdot K$, where K is a finite subgroup generated by some elements in Σ.

A natural question is whether Theorem 1.3 holds for non-symmetric measures. For admissible measures, (i.e., $\Gamma_\mu^+ = \Gamma$), Theorem 1.5 holds. For non-symmetric measures such that Γ_μ has infinite index, Theorem 1.6 applies directly. However, since $\Gamma_\mu \neq \Gamma_\mu^+$ for general non-symmetric measures, there is another case to consider: the case of measures μ such that $\Gamma_\mu = \Gamma$ (or Γ_μ has finite index in Γ), but Γ_μ^+ is much smaller than Γ. In this case, it seems that our arguments do not suffice. We give in Section 6 two examples illustrating the new difficulties:

(1) One can not rely on growth arguments, as for Theorem 1.6. Indeed, there are subsemigroups Λ^+ with bad asymptotic behavior, for instance such that $\liminf |B_n \cap \Lambda^+|/|B_n| = 0$ and $\limsup |B_n \cap \Lambda^+|/|B_n| > 0$.
(2) The arguments of Theorem 1.5 work for finitely supported measures, or for measures with a superexponential moment, but also more generally for measures with a nice geometric behavior (they should satisfy so-called Ancona inequalities). In the non-symmetric situation, we give in Proposition 6.2 explicit examples of (non-admissible) measures with an exponential moment and a very nice geometric behavior, and such that nevertheless $h = \ell v$. So, arguments similar to those of Theorem 1.5 can not

suffice, one needs a new argument that distinguishes in a finer way between measures with finite support and measures with infinite support.

This article is organized as follows. In Section 2, we give more details on the notions of hyperbolic group, drift and entropy. We also prove Theorem 1.4 on the asymptotic entropy and drift of the uniform measure on large balls. The following three sections are then devoted to the proofs of the three auxiliary theorems. Finally, we describe in Section 6 what can happen in the non-symmetric setting. In particular, we show that in any torsion-free group with infinitely many ends, there exist (non-admissible, non-symmetric) measures with an exponential moment satisfying $h = \ell v$.

2. General properties of entropy and drift in hyperbolic groups

2.1. **Hyperbolic spaces.** In this paragraph, we recall classical properties of hyperbolic spaces. See for instance [GdlH90] or [BH99].

Consider a metric space (X, d). The Gromov product of two points $y, y' \in X$, based at $x_0 \in X$, is by definition

$$(y|y')_{x_0} = (1/2)[d(x_0, y) + d(x_0, y') - d(y, y')]. \tag{2.1}$$

The space (X, d) is hyperbolic if there exists $\delta \geqslant 0$ such that, for any x_0, y_1, y_2, y_3, the following inequality holds:

$$(y_1|y_3)_{x_0} \geqslant \min((y_1|y_2)_{x_0}, (y_2|y_3)_{x_0}) - \delta.$$

The main intuition to have is that, in hyperbolic spaces, configurations of finitely many points look like configurations in trees: for any k, for any subset F of X with cardinality at most k, there exists a map Φ from F to a tree such that, for all $x, y \in F$,

$$d(x, y) - 2k\delta \leqslant d(\Phi(x), \Phi(y)) \leqslant d(x, y).$$

Hence, a lot of distance computations can be reduced to equivalent computations in trees (which are essentially combinatorial), up to a bounded error. Up to δ, the Gromov product $(y|y')_{x_0}$ is, in the approximating tree, the length of the part that is common to the geodesics from x_0 to y and from x_0 to y'.

A space (X, d) is geodesic if there exists a geodesic between any pair of points. For such spaces, there is a convenient characterization of hyperbolicity. A geodesic space (X, d) is hyperbolic if and only if there exists $\delta \geqslant 0$ such that its geodesic triangles are δ-thin, i.e., each side is included in the δ-neighborhood of the union of the two other sides.

Assume that (X, d_X) and (Y, d_Y) are two geodesic metric spaces, and that they are quasi-isometric. If (X, d_X) is hyperbolic, then so is (Y, d_Y). Note however that this equivalence only holds for geodesic spaces.

Let (X, d) be a geodesic hyperbolic metric space. A subset Y of X is quasi-convex if there exists a constant C such that, for any $y, y' \in Y$, the geodesics from y to y' stay in the C-neighborhood of Y.

We will sometimes encounter hyperbolic spaces which are not geodesic, but only quasi-geodesic: there exist constants $C > 0$ and λ such that any two points can be joined by a (λ, C)-quasi-geodesic, i.e., a map f from a real interval to X such that $\lambda^{-1}|t' - t| - C \leqslant d(f(t), f(t')) \leqslant \lambda|t' - t| + C$. When the space is geodesic, a quasi-geodesic stays a bounded distance away from a true geodesic. Most properties that hold or can be defined using geodesics (for instance the notion of quasi-convexity) can be extended to this setting, simply replacing geodesics with quasi-geodesics in the statements.

Let (X,d) be a proper geodesic hyperbolic space. Its boundary at infinity ∂X is by definition the set of geodesics originating from a base point x_0, where two such geodesics are identified if they remain a bounded distance away. It is a compact space, which does not depend on x_0. The space $X \cup \partial X$ is also compact. If X is only quasi-geodesic, all these definitions extend using quasi-geodesics instead of geodesics.

Any isometry (or, more generally, quasi-isometry) of a hyperbolic space extends continuously to its boundary, giving a homeomorphism of ∂X.

The Gromov product may be extended to $X \cup \partial X$: we define $(\xi|\eta)_{x_0}$ as the infimum limit of $(x_n|y_n)_{x_0}$ for x_n and y_n converging respectively to ξ and η. The choice to take the infimum is arbitrary, one could also take the supremum or any accumulation point, those quantities differ by at most a constant only depending on δ. Hence, one should think of the Gromov product at infinity to be canonically defined only up to an additive constant. Heuristically, $(\xi|\eta)_{x_0}$ is the time after which two geodesics from x_0 to ξ and to η start diverging.

Let (X,d) be a proper geodesic (or quasi-geodesic) hyperbolic space. For any small enough $\varepsilon > 0$, one may define a visual distance $d_{\partial X,\varepsilon}$ on ∂X such that $d_{\partial X,\varepsilon}(\xi,\eta) \asymp e^{-\varepsilon(\xi|\eta)_{x_0}}$ (meaning that the ratio between these quantities is uniformly bounded from above and from below).

Let (X,d) be a proper hyperbolic metric space. One can define another boundary of X, the Busemann boundary (or horoboundary), as follows. Let x_0 be a fixed basepoint in X. To $x \in X$, one associates its horofunction $h_x(y) = d(y,x) - d(x_0,x)$, normalized so that $h_x(x_0) = 0$. The map $\Phi : x \mapsto h_x$ is an embedding of X into the space of 1-Lipschitz functions on X, with the topology of uniform convergence on compact sets. The horoboundary is obtained by taking the closure of $\Phi(X)$. In other words, a sequence $x_n \in X$ converges to a boundary point if $h_{x_n}(y)$ converges, uniformly on compact sets. Its limit is the horofunction h_ξ associated to the corresponding boundary point ξ (it is also called the Busemann function associated to ξ). We denote by $\partial_B X$ the Busemann boundary of X. There is a continuous projection $\pi_B : \partial_B X \to \partial X$, which is onto but not injective in general. The boundary $\partial_B X$ is rather sensitive to fine scale details of the distance d, while ∂X only depends on its quasi-isometry class.

Any isometry φ of X acts on horofunctions, by the formula $h_{\varphi(x)}(y) = h_x(\varphi^{-1}y) - h_x(\varphi^{-1}x_0)$. This implies that φ extends to a homeomorphism on $\partial_B X$, given by the same formula $h_{\varphi(\xi)}(y) = h_\xi(\varphi^{-1}y) - h_\xi(\varphi^{-1}x_0)$. Note that, contrary to the action on the geometric boundary, this only works for isometries of X, not quasi-isometries.

2.2. **Hyperbolic groups.** Let Γ be a finitely generated group, with a finite symmetric generating set S. Denote by $d = d_S$ the corresponding word distance. The group Γ is hyperbolic if the metric space (Γ, d_S) is hyperbolic. Since hyperbolicity is invariant under quasi-isometry for geodesic spaces, this notion does not depend on the choice of the generating set S. However, if one considers another left-invariant distance on Γ which is equivalent to d_S but not geodesic, its hyperbolicity is not automatic. Hence, one should postulate its hyperbolicity if it is needed, as in the statement of Theorem 1.2. We say that the pair (Γ, d) is a metric hyperbolic group if the group Γ is hyperbolic for one (or, equivalently, for any) word distance, and if the distance d is left-invariant, hyperbolic, and quasi-isometric to one (or equivalently, any) word distance. Such a distance d does not have to be geodesic, but it is quasi-geodesic since geodesics for a given word distance form a system of quasi-geodesics for d, going from any point to any point.

Let (Γ, d) be a metric hyperbolic group. The left-multiplication by elements of Γ is isometric. Hence, Γ acts by homeomorphisms on its compactifications $\Gamma \cup \partial\Gamma$ and $\Gamma \cup \partial_B\Gamma$. Moreover, any infinite order element $g \in \Gamma$ acts hyperbolically on $\Gamma \cup \partial\Gamma$: it has two fixed points at infinity g^- and g^+, the points in $\Gamma \cup \partial\Gamma \setminus \{g^-\}$ are attracted to g^+ by forward iteration of g, and the points in $\Gamma \cup \partial\Gamma \setminus \{g^+\}$ are attracted to g^- by backward iteration of g.

Definition 2.1. *Consider an action of a group Γ on a space Z. A function $c : \Gamma \times Z \to \mathbb{R}$ is a cocycle if, for any $g, h \in \Gamma$ and any $\xi \in Z$,*

$$c(gh, \xi) = c(g, h\xi) + c(h, \xi). \tag{2.2}$$

The cocycle is Hölder-continuous if Z is a metric space and each function $\xi \mapsto c(g, \xi)$ is Hölder-continuous.

There is a choice to be made in the definition of cocycles, since one may compose with g or g^{-1}. Our definition is the most customary. With this definition, the map $c_B : \Gamma \times \partial_B\Gamma \to \mathbb{R}$ given by $c_B(g, \xi) = h_\xi(g^{-1})$ is a cocycle, called the Busemann cocycle.

A subgroup H of Γ is nonelementary if its action on $\partial\Gamma$ does not fix a finite set. Equivalently, H is not virtually the trivial group or $\mathbb{Z}$. We say that a probability measure μ on Γ is nonelementary if the subgroup Γ_μ generated by its support is itself nonelementary.

Let μ be a probability measure on Γ. Since Γ acts by homeomorphisms on the compact space $\partial\Gamma$, it admits a stationary measure: there exists a probability measure ν on $\partial\Gamma$ such that $\mu * \nu = \nu$, i.e., $\sum_{g \in \Gamma} \mu(g) g_* \nu = \nu$. If μ is nonelementary, this measure is unique, and has no atom (see [Kai00]). It is also the exit measure of the corresponding random walk $X_n = g_1 \cdots g_n$: almost every trajectory $X_n(\omega)$ converges to a point $X_\infty(\omega) \in \partial\Gamma$, and moreover the distribution of X_∞ is precisely ν.

In the same way, since Γ acts on $\partial_B\Gamma$, it admits a stationary measure ν_B there. This measure is not unique in general, even if μ is nonelementary. However, all such measures project under π_B to the unique stationary measure on $\partial\Gamma$.

2.3. **The drift.** Let (Γ, d) be a metric hyperbolic group. Consider a probability measure μ on Γ, with finite first moment $L(\mu)$ (defined in (1.1)). The drift of the random walk has been defined in (1.2) as $\ell(\mu) = \lim L(\mu^{*n})/n$. Let $X_n = g_1 \cdots g_n$ be the position at time n of the random walk generated by μ (where the g_i are independent and distributed according to μ). Then, almost surely, $\ell(\mu) = \lim |X_n|/n$.

The drift also admits a description in terms of the Busemann boundary. The following result is well-known (compare with [KL11, Theorem 18]).

Proposition 2.2. *Let (Γ, d) be a metric hyperbolic group. Let μ be a nonelementary probability measure on Γ with finite first moment. Let ν_B be a μ-stationary measure on $\partial_B\Gamma$. Then*

$$\ell(\mu) = \int_{\Gamma \times \partial_B\Gamma} c_B(g, \xi) \, \mathrm{d}\mu(g) \, \mathrm{d}\nu_B(\xi). \tag{2.3}$$

Proof. Let X_n be the position of the random walk at time n. Using the cocycle property of the Busemann cocycle, we have

$$\begin{aligned}\int c_B(X_n(\omega),\xi)\,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}\nu_B(\xi) &= \int c_B(g_1\cdots g_n,\xi)\,\mathrm{d}\mu(g_1)\cdots\mathrm{d}\mu(g_n)\,\mathrm{d}\nu_B(\xi)\\ &= \sum_{k=1}^{n}\int c_B(g_k, g_{k+1}\cdots g_n\xi)\,\mathrm{d}\mu(g_k)\cdots\mathrm{d}\mu(g_n)\,\mathrm{d}\nu_B(\xi).\end{aligned}$$

Since the measure ν_B is stationary, the point $g_{k+1}\cdots g_n\xi$ is distributed according to ν_B. Hence, the terms in the above sum do not depend on k. We get

$$\int_{\Gamma\times\partial_B\Gamma} c_B(g,\xi)\,\mathrm{d}\mu(g)\,\mathrm{d}\nu_B(\xi) = \frac{1}{n}\int c_B(X_n(\omega),\xi)\,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}\nu_B(\xi). \tag{2.4}$$

We have $|c_B(X_n,\xi)|/n \leqslant |X_n|/n$, which converges in L^1 and almost surely to ℓ. Hence, the sequence of functions $c_B(X_n(\omega),\xi)/n$ is uniformly integrable on $\Omega\times\partial_B\Gamma$. Moreover, X_n converges almost surely to a point on the boundary $\partial\Gamma$, distributed according to the exit measure, which has no atom. It follows that, for all ξ, the trajectory $X_n(\omega)$ converges almost surely to a point different from $\pi_B(\xi)$. This implies that, almost surely, one has $c_B(X_n,\xi) = |X_n| + O(1)$, giving in particular $c_B(X_n,\xi)/n \to \ell$. The result follows by taking the limit in n in the equality (2.4). □

This formula easily implies that the drift depends continuously on the measure, as explained in [EK13].

Proposition 2.3. *Let (Γ, d) be a metric hyperbolic group. Consider a sequence of probability measures μ_i with finite first moment, converging simply to a nonelementary probability measure μ (i.e., $\mu_i(g)\to\mu(g)$ for all $g\in\Gamma$). Assume moreover that $L(\mu_i)\to L(\mu)$. Then $\ell(\mu_i)\to\ell(\mu)$.*

Proof. Let ν_i be stationary measures for μ_i on $\partial_B\Gamma$. Taking a subsequence if necessary, we may assume that ν_i converges to a limiting measure ν. By continuity of the action on the boundary, it is stationary for μ.

For each $g\in\Gamma$, the quantity $\int_{\partial_B\Gamma} c_B(g,\xi)\,\mathrm{d}\nu_i(\xi)$ converges to $\int_{\partial_B\Gamma} c_B(g,\xi)\,\mathrm{d}\nu(\xi)$ since $\xi\mapsto c_B(g,\xi)$ is continuous. Averaging over g (and using the assumption $L(\mu_i)\to L(\mu)$ to get a uniform domination), we deduce that

$$\sum_{g\in\Gamma}\mu_i(g)\int_{\partial_B\Gamma} c_B(g,\xi)\,\mathrm{d}\nu_i(\xi) \to \sum_{g\in\Gamma}\mu(g)\int_{\partial_B\Gamma} c_B(g,\xi)\,\mathrm{d}\nu(\xi).$$

Together with the formula (2.3) for the drift, this completes the proof. □

In this proposition, it is important that μ is nonelementary: the result is wrong otherwise. For instance, in the infinite dihedral group $\mathbb{Z}\rtimes\mathbb{Z}/2$, the measures $\mu_i = (1-2^{-i})\delta_{(1,0)} + 2^{-i}\delta_{(0,1)}$ have zero drift since the $\mathbb{Z}/2$ element symmetrizes everything in $\mathbb{Z}$, while the limiting measure $\mu = \delta_{(1,0)}$ has drift 1. The reason is the non-uniqueness of the stationary measure for μ on the boundary.

2.4. **The entropy.** Let Γ be a countable group. Consider a probability measure μ on Γ, with finite time one entropy $H(\mu)$ (defined in (1.1)). The entropy of the random walk has been defined in (1.2) as $h(\mu) = \lim H(\mu^{*n})/n$. Let $X_n = g_1 \cdots g_n$ be the position at time n of the random walk generated by μ (where the g_i are independent and distributed according to μ). Then, almost surely, $h(\mu) = \lim(-\log \mu^{*n}(X_n))/n$. The fundamental inequality (1.3) shows that if $h > 0$ then $\ell > 0$.

The entropy has several equivalent characterizations. The first one is in terms of the size of the typical support of the random walk: This support has size roughly e^{hn}. The following lemma follows from [Haï13, Proposition 1.13].

Lemma 2.4. *Consider a probability measure μ with $H(\mu) < \infty$ on a countable group. Let $h = h(\mu)$ be its asymptotic entropy. Let $\eta > 0$ and $\varepsilon > 0$.*

(1) *For large enough n, there exists a subset K_n of Γ with $\mu^{*n}(K_n) \geqslant 1 - \eta$ and $|K_n| \leqslant e^{(h+\varepsilon)n}$.*
(2) *For large enough n, there exists no subset K_n of Γ with $\mu^{*n}(K_n) \geqslant \eta$ and $|K_n| \leqslant e^{(h-\varepsilon)n}$.*

Another description is in terms of the Poisson boundary of the walk. To avoid general definitions, let us only state this description for measures on hyperbolic groups. The following proposition is a consequence of [Kai00].

Proposition 2.5. *Let Γ be a hyperbolic group. Let μ be a nonelementary probability measure on Γ with $H(\mu) < \infty$. Let ν be its unique stationary measure on $\partial\Gamma$. Define the Martin cocycle on $\Gamma \times \partial\Gamma$ by $c_M(g, \xi) = -\log(\mathrm{d}g_*^{-1}\nu / \mathrm{d}\nu)(\xi)$. Then*

$$h(\mu) \geqslant \int_{\Gamma\times\partial\Gamma} c_M(g,\xi)\,\mathrm{d}\mu(g)\,\mathrm{d}\nu(\xi), \tag{2.5}$$

with equality if μ has a logarithmic moment.

When μ has a logarithmic moment, this proposition has a very similar flavor to Proposition 2.2 expressing the drift of a random walk. Indeed, for symmetric measures, [BHM11] interprets Proposition 2.5 as a special case of Proposition 2.2, for a distance $d = d_\mu$ related to the random walk, the Green distance, which we defined in Theorem 1.2. This distance is hyperbolic if μ is admissible and has a superexponential moment, by [Anc87, Gou13]. It is not geodesic in general, but this is not an issue since we were careful enough to state Proposition 2.2 without this assumption. The Busemann cocycle for the Green distance is precisely the Martin cocycle.

An important difference between the formulas (2.3) for the drift and (2.5) for the entropy is that, in the latter situation, the cocycle c_M depends on the measure ν (and, therefore, on μ). This makes it more complicated to prove continuity statements such as Proposition 2.3 for the entropy. Nevertheless, Erschler and Kaimanovich proved in [EK13] that, in hyperbolic groups, the entropy also depends continuously on the measure. As $h(\mu) = \inf H(\mu^{*n})/n$ by subadditivity, it is easy to prove that when $\mu_i \to \mu$ one has $\limsup h(\mu_i) \leqslant h(\mu)$. The main difficulty to prove the continuity is to get lower bounds. We will need a slightly stronger (and more pedestrian) version of the results of [EK13] to prove Theorem 1.4. Although our argument may seem very different at first sight from the arguments in [EK13], the techniques are in fact closely related (an illustration is that we can recover with our techniques the result of Kaimanovich that, for measures with finite logarithmic moment, equality holds in (2.5), i.e., the Poisson boundary coincides with the geometric boundary, see Remark 2.11). Our main criterion to get lower bounds on the entropy is the following. We write $\mathbb{S}^k = \{g \in \Gamma \,:\, |g| \in (k-1, k]\}$ for the thickened sphere, so that the union of these spheres covers the whole group.

Theorem 2.6. *Let (Γ, d) be a metric hyperbolic group. Let μ_i be a sequence of nonelementary probability measures on Γ with $H(\mu_i) < \infty$. Let ν_i be the unique stationary measure for μ_i on $\partial\Gamma$. Assume that:*

(1) *The limit points of ν_i have no atom.*
(2) *The sequence*

$$h_i = \sum_k \sum_{g \in \mathbb{S}^k} \mu_i(g)(-\log(\mu_i(g)/\mu_i(\mathbb{S}^k))) \tag{2.6}$$

tends to infinity.

Then $\liminf h(\mu_i)/h_i \geqslant 1$.

The quantity h_i can be written

$$h_i = \sum_{g \in \Gamma} \mu_i(g)(-\log \mu_i(g)) - \sum_k \mu_i(\mathbb{S}^k)(-\log \mu_i(\mathbb{S}^k)).$$

The first term is the time one entropy $H(\mu_i)$ of the measure μ_i. In most reasonable cases, the second term is negligible. The theorem then states that the asymptotic entropy $h(\mu_i)$ is comparable to the time one entropy $H(\mu_i)$. In other words, if the measure is supported close to infinity, and sufficiently spread out in the group (this is the meaning of the assumption that the limit points of ν_i have no atom), then there are few coincidences and the entropy does not decrease significantly with time.

To prove this theorem, we will use the following technical lemma.

Lemma 2.7. *On a probability space (X, μ), consider a nonnegative function f with average 1. For any subset A of X,*

$$\int_X (-\log f) \geqslant \mu(A)\left(-\log \int_A f\right) - 2e^{-1}.$$

Proof. As the function $x \mapsto -\log x$ is convex, Jensen's inequality gives $\int(-\log f) \geqslant -\log(\int f)$. The last quantity vanishes when $\int f = 1$.

Let $B \subset X$. Write $a = \int_B f \,\mathrm{d}\mu/\mu(B)$. The measure $\mathrm{d}\mu/\mu(B)$ is a probability measure on B, and the function f/a has integral 1 for this measure. The previous inequality gives $\int_B (-\log(f/a)) \,\mathrm{d}\mu/\mu(B) \geqslant 0$, that is,

$$\int_B (-\log f) \,\mathrm{d}\mu \geqslant -\mu(B) \log a = -\mu(B) \log\left(\int_B f\right) + \mu(B) \log \mu(B).$$

The quantity $\mu(B)\log\mu(B)$ is bounded from below by $\inf_{[0,1]} x \log x = -e^{-1}$. Therefore,

$$\int_B (-\log f) \,\mathrm{d}\mu \geqslant -\mu(B) \log\left(\int_B f\right) - e^{-1}.$$

We apply this inequality to the complement A^c of A. As $-\log\big(\int_{A^c} f\big) \geqslant 0$, we get a lower bound $-e^{-1}$. Let us also apply this inequality to A, and add the results. We obtain

$$\int_X (-\log f)\, \mathrm{d}\mu \geqslant -\mu(A) \log\left(\int_A f\right) - 2e^{-1}. \qquad \square$$

We will use the notion of shadow, due to Sullivan and considered in this context by Coornaert [Coo93]. Let $C > 0$ be large enough. The shadow $\mathcal{O}(g, C)$ of $g \in \Gamma$ is $\{\xi \in \partial\Gamma : (g|\xi)_e \geqslant |g| - C\}$. In geometric terms (and assuming the space is geodesic), this is essentially the trace at infinity of geodesics originating from e and going through the ball $B(g, C)$. We will use the following properties of shadows [Coo93]:

(1) Their covering number is finite. More precisely, there exists $D > 0$ (depending on C) such that, for any integer k, for any $\xi \in \partial\Gamma$,

$$|\{g \in \mathbb{S}^k \,:\, \xi \in \mathcal{O}(g, C)\}| \leqslant D.$$

(2) The preimages of shadows are large. More precisely, for any $\eta > 0$, there exists $C > 0$ such that, for all $g \in \Gamma$, the complement of $g^{-1}\mathcal{O}(g, C)$ has diameter at most η (for a fixed visual distance on the boundary).

Proof of Theorem 2.6. Fix $\varepsilon > 0$. As the limit points of ν_i have no atom, there exists $\eta > 0$ such that any ball of radius η in $\partial\Gamma$ has measure at most ε for ν_i, for i large enough. We can then choose a shadow size C so that $g^{-1}\mathcal{O}(g, C)$ has for all g a complement with diameter at most η. This yields $\nu_i(g^{-1}\mathcal{O}(g, C)) \geqslant 1 - \varepsilon$.

By (2.5), the entropy of μ_i satisfies

$$h(\mu_i) \geqslant \sum_{g \in \Gamma} \mu_i(g) \int_{\partial\Gamma} \left(-\log \frac{\mathrm{d}g_*^{-1}\nu_i}{\mathrm{d}\nu_i}(\xi)\right) \mathrm{d}\nu_i(\xi).$$

The function $f_{i,g} = \frac{\mathrm{d}g_*^{-1}\nu_i}{\mathrm{d}\nu_i}(\xi)$ is nonnegative and has integral 1. For any $A \subset \partial\Gamma$, Lemma 2.7 gives

$$\begin{aligned}
\int_{\partial\Gamma} \left(-\log \frac{\mathrm{d}g_*^{-1}\nu_i}{\mathrm{d}\nu_i}(\xi)\right) \mathrm{d}\nu_i(\xi) &\geqslant -\nu_i(A) \log\left(\int_A \frac{\mathrm{d}g_*^{-1}\nu_i}{\mathrm{d}\nu_i}(\xi)\, \mathrm{d}\nu_i(\xi)\right) - 2e^{-1} \\
&= -\nu_i(A) \log(g_*^{-1}\nu_i(A)) - 2e^{-1} \\
&= -\nu_i(A) \log(\nu_i(gA)) - 2e^{-1}.
\end{aligned}$$

Let us take $A = g^{-1}\mathcal{O}(g, C)$, so that $\nu_i(A) \geqslant 1 - \varepsilon$. Summing over g, we get

$$h(\mu_i) \geqslant (1 - \varepsilon) \sum_{g \in \Gamma} \mu_i(g)(-\log \nu_i(\mathcal{O}(g, C))) - 2e^{-1}. \tag{2.7}$$

We split the sum according to the spheres $\mathbb{S}^k$. Let $\Sigma_k = \sum_{g\in\mathbb{S}^k} \nu_i(\mathcal{O}(g,C))$, it is at most D since the shadows have a covering number bounded by D. We have

$$\sum_{g\in\mathbb{S}^k} \mu_i(g)(-\log \nu_i(\mathcal{O}(g,C)))$$
$$= -\mu_i(\mathbb{S}^k) \sum_{g\in\mathbb{S}^k} \frac{\mu_i(g)}{\mu_i(\mathbb{S}^k)} \left[\log\left(\frac{\nu_i(\mathcal{O}(g,C))}{\Sigma_k \mu_i(g)/\mu_i(\mathbb{S}^k)}\right) + \log \Sigma_k + \log(\mu_i(g)/\mu_i(\mathbb{S}^k))\right].$$

The point of this decomposition is that the function on $\mathbb{S}^k$ given by $\varphi : g \mapsto \frac{\nu_i(\mathcal{O}(g,C))}{\Sigma_k \mu_i(g)/\mu_i(\mathbb{S}^k)}$ has integral 1 for the probability measure $\mu_i(g)/\mu_i(\mathbb{S}^k)$. By Jensen's inequality, the integral of $-\log\varphi$ is nonnegative. This yields

$$\sum_{g\in\mathbb{S}^k} \mu_i(g)(-\log \nu_i(\mathcal{O}(g,C))) \geqslant -\mu_i(\mathbb{S}^k)\log D + \sum_{g\in\mathbb{S}^k} \mu_i(g)(-\log(\mu_i(g)/\mu_i(\mathbb{S}^k)).$$

Summing over k, we deduce from (2.7) the inequality

$$h(\mu_i) \geqslant (1-\varepsilon)h_i - 2e^{-1} - \log D.$$

As h_i tends to infinity, this gives $h(\mu_i) \geqslant (1-2\varepsilon)h_i$ for large enough i, completing the proof. □

To apply the previous theorem, we need to estimate h_i. In this respect, the following lemma is often useful.

Lemma 2.8. *Let $R_i \geqslant 1$. The quantity h_i defined in* (2.6) *satisfies*

$$h_i \geqslant \sum_{|g|\leqslant R_i} \mu_i(g)(-\log \mu_i(g)) - \log(2+R_i).$$

Proof. In the definition of h_i, all the terms are nonnegative. Restricting the sum to those g with $|g| \leqslant R_i$, we get

$$h_i \geqslant \sum_{k\leqslant R_i} \sum_{g\in\mathbb{S}^k} \mu_i(g)(-\log(\mu_i(g)/\mu_i(\mathbb{S}^k)))$$
$$= \sum_{|g|\leqslant R_i} \mu_i(g)(-\log \mu_i(g)) - \sum_{k\leqslant R_i} \mu_i(\mathbb{S}^k)(-\log \mu_i(\mathbb{S}^k)).$$

A probability measure supported on a set with N elements has entropy at most $\log N$. The number $\mu_i(\mathbb{S}^k)$ for $0 \leqslant k \leqslant R_i$ are not a probability measure in general, let us add a last atom with mass $m = \mu_i(\bigcup_{k>R_i} \mathbb{S}^k)$. We are considering a space of cardinality R_n+2, hence

$$m(-\log m) + \sum_{k\leqslant R_i} \mu_i(\mathbb{S}^k)(-\log \mu_i(\mathbb{S}^k)) \leqslant \log(2+R_i),$$

completing the proof. □

Let us see how Theorem 2.6 implies a slightly stronger version of the continuity result for the entropy of Erschler and Kaimanovich [EK13].

Theorem 2.9. *Let Γ be a hyperbolic group. Consider a probability measure μ with finite time one entropy and finite logarithmic moment. Let μ_i be a sequence of probability measures converging simply to μ with $H(\mu_i) \to H(\mu)$. Then $h(\mu_i) \to h(\mu)$.*

The assumption $H(\mu_i) \to H(\mu)$ ensures that there is no additional entropy in μ_i coming from neighborhoods of infinity that would disappear in the limit. It is automatic if the support of μ_i is uniformly bounded or if μ_i satisfies a uniform L^1 domination, but it is much weaker. For instance, it is allowed that the μ_i have no finite logarithmic moment.

The main lemma for the proof is a lower bound on the entropy, following from Theorem 2.6.

Lemma 2.10. *Let Γ be a hyperbolic group. Consider a probability measure μ with finite time one entropy and finite logarithmic moment. Let μ_i be a sequence of measures converging simply to μ. Then $\liminf h(\mu_i) \geqslant h(\mu)$.*

Proof. Since the result is trivial if $h(\mu) = 0$, we can assume that $h(\mu) > 0$.

Let $\varepsilon > 0$. For large n, most atoms for μ^{*n} have a probability at most $e^{-(1-\varepsilon)nh(\mu)}$. Moreover, since μ has a finite logarithmic moment, $\log|X_n|/n$ tends almost surely to 0 by [Aar97, Proposition 2.3.1]. Therefore, the set

$$K_n = \{g \,:\, \mu^{*n}(g) \leqslant e^{-(1-\varepsilon)nh(\mu)},\ |g| \leqslant e^{\varepsilon n}\}$$

has measure tending to 1. In particular $\mu^{*n}(K_n) \geqslant 1 - \varepsilon$ for large n. We get

$$\begin{aligned}\sum_{|g|\leqslant e^{\varepsilon n}} \mu^{*n}(g)(-\log \mu^{*n}(g)) &\geqslant \sum_{g\in K_n} \mu^{*n}(g)(-\log \mu^{*n}(g)) \geqslant \sum_{g\in K_n} \mu^{*n}(g)(1-\varepsilon)nh(\mu) \\ &= \mu^{*n}(K_n)(1-\varepsilon)nh(\mu) \geqslant (1-\varepsilon)^2 nh(\mu).\end{aligned}$$

For each fixed n, the measures μ_i^{*n} converge to μ^{*n} when i tends to infinity. Hence, we get for large enough i the inequality

$$\sum_{|g|\leqslant e^{\varepsilon n}} \mu_i^{*n}(g)(-\log \mu_i^{*n}(g)) \geqslant (1-\varepsilon)^3 nh(\mu).$$

Letting ε tend to 0 (and, therefore, n to infinity), we deduce the existence of sequences $n_i \to \infty$ and $\varepsilon_i \to 0$ such that, for any i,

$$\sum_{|g|\leqslant e^{\varepsilon_i n_i}} \mu_i^{*n_i}(g)(-\log \mu_i^{*n_i}(g)) \geqslant (1-\varepsilon_i)^3 n_i h(\mu).$$

Let $\tilde{\mu}_i = \mu_i^{*n_i}$. Its stationary measure ν_i is also the stationary measure of μ_i, by uniqueness. Any limit point of ν_i is stationary for μ, and is therefore atomless since μ is nonelementary as $h(\mu) > 0$. The assumptions of Theorem 2.6 are satisfied by the sequence $\tilde{\mu}_i$. Moreover, Lemma 2.8 yields

$$h_i \geqslant (1-\varepsilon_i)^3 n_i h(\mu) - 2\varepsilon_i n_i \geqslant (1 - C\varepsilon_i) n_i h(\mu).$$

Theorem 2.6 ensures that $\liminf h(\tilde{\mu}_i)/h_i \geqslant 1$. As $h(\tilde{\mu}_i) = n_i h(\mu_i)$, this gives $\liminf h(\mu_i) \geqslant h(\mu)$ as desired. □

Proof of Theorem 2.9. For fixed n, the sequence μ_i^{*n} converges simply to μ^{*n}. Moreover, $H(\mu_i^{*n}) \to H(\mu^{*n})$ since there is no loss of entropy at infinity by assumption. Choose n such that $H(\mu^{*n}) \leqslant n(1+\varepsilon)h(\mu)$. We get $H(\mu_i^{*n})/n \leqslant (1+2\varepsilon)h(\mu)$ for large enough i. As $h(\mu_i) \leqslant H(\mu_i^{*n})/n$, this shows that $\limsup h(\mu_i) \leqslant h(\mu)$ (this is the classical semi-continuity property of entropy, valid in any group).

For the reverse inequality $\liminf h(\mu_i) \geqslant h(\mu)$, we apply Lemma 2.10. □

Remark 2.11. Let $h(\mu, \partial\Gamma) = \int_{\Gamma\times\partial\Gamma}(-\log \mathrm{d}g_*^{-1}\nu/\mathrm{d}\nu)(\xi)\,\mathrm{d}\mu(g)\,\mathrm{d}\nu(\xi)$ where ν is the stationary measure for μ on $\partial\Gamma$. In general, $h(\mu) \geqslant h(\mu, \partial\Gamma)$ with equality if and only if $(\partial\Gamma, \nu)$ is the Poisson boundary of (Γ, μ). A theorem of Kaimanovich [Kai00] asserts that, when μ has finite entropy and finite logarithmic moment, $h(\mu, \partial\Gamma) = h(\mu)$. We can recover this theorem using the previous arguments. Indeed, what the proof of Theorem 2.6 really shows is that $\liminf h(\mu_i, \partial\Gamma)/h_i \geqslant 1$. Hence, Lemma 2.10 proves that $\liminf h(\mu_i, \partial\Gamma) \geqslant h(\mu)$ if μ_i converges simply to a measure μ with a logarithmic moment. Taking $\mu_i = \mu$ for all i, we obtain in particular $h(\mu, \partial\Gamma) \geqslant h(\mu)$, as desired.

2.5. **A criterion to bound the entropy from below.** In order to prove Theorem 1.4 on the entropy of the uniform measure on balls, we want to apply Theorem 2.6. Thus, we need a criterion to check that limit points of stationary measures have no atom.

Lemma 2.12. *Let Γ be a hyperbolic group. Let μ_i be a sequence of probability measures on Γ. Assume that, on the space $\Gamma \cup \partial\Gamma$, the sequence μ_i converges to a limit ν which is supported on $\partial\Gamma$. Assume moreover that the limit points of $\check{\mu}_i$ (defined by $\check{\mu}_i(g) = \mu_i(g^{-1})$) have no atom. Then the stationary measures ν_i associated to μ_i also converge to ν.*

Proof. We fix a word distance d on Γ. Let f be a continuous function on $\Gamma \cup \partial\Gamma$. Let us show that, uniformly in $\xi \in \partial\Gamma$, the integral $\int f(g\xi)\,\mathrm{d}\mu_i(g)$ is close to $\int f(g)\,\mathrm{d}\mu_i(g)$. We estimate the difference as

$$\left|\int (f(g\xi) - f(g))\,\mathrm{d}\mu_i(g)\right| \leqslant \int |f(g\xi) - f(g)| 1((g\xi|g)_e > C)\,\mathrm{d}\mu_i(g) + 2\|f\|_\infty \int 1((g\xi|g)_e \leqslant C)\,\mathrm{d}\mu_i(g),$$

where C is a fixed constant. If C is large enough, $|f(x) - f(y)| \leqslant \varepsilon$ when $(x|y)_e > C$, by uniform continuity of f. Hence, the first integral is bounded by ε. For the second integral, we use the formula $(gx|g)_e = |g| - (x|g^{-1})_e$, valid for any $x \in \Gamma$ (it follows readily from the definition (2.1) of the Gromov product). This equality does not extend to the boundary since the Gromov product there is only well defined up to an additive constant D. Nevertheless, we get $(g\xi|g)_e \geqslant |g| - (\xi|g^{-1})_e - D$. Hence, the second integral is bounded by

$$\mu_i\{g \,:\, |g| - C - D \leqslant (\xi|g^{-1})_e\}. \tag{2.8}$$

If $|g|$ is large, the points g with $(\xi|g^{-1})_e \geqslant |g| - C - D$ are such that g^{-1} belongs to a small neighborhood of ξ in $\Gamma \cup \partial\Gamma$. As the limit points of $\check{\mu}_i$ are supported on $\partial\Gamma$ and have no atom, it follows that (2.8) converges to 0 when i tends to infinity, uniformly in ξ.

We have proved that

$$\sup_{\xi\in\partial\Gamma}\left|\int f(g\xi)\,\mathrm{d}\mu_i(g) - \int f(g)\,\mathrm{d}\mu_i(g)\right| \to 0.$$

By stationarity,

$$\int_{\xi\in\partial\Gamma} f(\xi)\,\mathrm{d}\nu_i(\xi) = \int_{\xi\in\partial\Gamma}\left(\int f(g\xi)\,\mathrm{d}\mu_i(g)\right)\mathrm{d}\nu_i(\xi).$$

Combining these equations, we get $\int f(\xi)\,\mathrm{d}\nu_i(\xi) - \int f(g)\,\mathrm{d}\mu_i(g) \to 0$. This shows that the limit points of ν_i and μ_i are the same. □

Let us now consider the uniform measure μ_i on the ball of radius i, as in Theorem 1.4. The next lemma follows from the techniques of [Coo93].

Lemma 2.13. *Let (Γ, d) be a metric hyperbolic group. Let ρ_i be the uniform measure on the ball of radius i. Let ρ_∞ be the Patterson-Sullivan of (Γ, d) constructed in* [Coo93] *(it is supported on $\partial\Gamma$ and atomless). Then the limit points of ρ_i are equivalent to ρ_∞, with a density bounded from above and from below.*

Proof. Let C be large enough. We will use the shadows $\mathcal{O}(g, C)$ as defined before the proof of Theorem 2.6. The main property of ρ_∞ is that it satisfies

$$K_0^{-1} e^{-v|g|} \leqslant \rho_\infty(\mathcal{O}(g, C)) \leqslant K_0 e^{-v|g|}, \tag{2.9}$$

where K_0 is a constant only depending on C and v is the growth of (Γ, d) (Proposition 6.1 in [Coo93]).

Let μ_i be the uniform measure on thickened spheres $S_i = \{g : i \leqslant |g| \leqslant i + L\}$, where L is large enough so that the cardinality of S_i grows like e^{iv}, see the proof of Theorem 7.2 in [Coo93]. Let us push μ_i to a measure $\tilde{\mu}_i$ on $\partial\Gamma$, by choosing for each $g \in S_i$ a corresponding point in its shadow. It is clear that μ_i and $\tilde{\mu}_i$ have the same limit points, since the diameter of the shadows tends uniformly to 0 when $i \to \infty$. We will prove that the limit points of $\tilde{\mu}_i$ are equivalent to ρ_∞. The same result follows for μ_i and then ρ_i.

The shadows of $g \in S_i$ have a covering number which is bounded from above by a constant D, and from below by 1 if C is large enough. Hence, the measures $\tilde{\mu}_i$ satisfy

$$K_1^{-1} e^{-iv} \leqslant \tilde{\mu}_i(\mathcal{O}(g, C)) \leqslant K_1 e^{-iv},$$

for any $g \in S_i$. This is comparable to $\rho_\infty(\mathcal{O}(g, C))$ by (2.9), up to a multiplicative constant K_2. Consider a limit $\tilde{\mu}$ of a sequence $\tilde{\mu}_{i_n}$, let us prove that it is uniformly equivalent to ρ_∞. We will only prove that $\tilde{\mu} \leqslant DK_2\rho_\infty$, the other inequality is proved in the same way. By regularity of the measures, it suffices to check this inequality on compact sets.

Let A be a compact subset of $\partial\Gamma$, and $\varepsilon > 0$. By regularity of the measure ρ_∞, there is an open neighborhood U of A with $\rho_\infty(U) \leqslant \rho_\infty(A) + \varepsilon$. Consider B a compact neighborhood of A, included in U, with $\tilde{\mu}(\partial B) = 0$ (such a set exists, since among the sets $B_r = \{\xi : d(\xi, A) \leqslant r\}$, at most countably of them many have a boundary with nonzero measure). For large enough i, the shadows $\mathcal{O}(g, C)$ with $g \in S_i$ which intersect B are contained in U. Therefore,

$$\tilde{\mu}_i(B) \leqslant \sum_{g\in S_i, \mathcal{O}(g,C)\cap B\neq\emptyset} \tilde{\mu}_i(\mathcal{O}(g, C)) \leqslant K_2 \sum_{g\in S_i, \mathcal{O}(g,C)\cap B\neq\emptyset} \rho_\infty(\mathcal{O}(g, C)) \leqslant DK_2\rho_\infty(U).$$

As $\tilde{\mu}(\partial B) = 0$, the sequence $\tilde{\mu}_{i_n}(B)$ tends to $\tilde{\mu}(B)$. We obtain $\tilde{\mu}(B) \leqslant DK_2\rho_\infty(U)$. As A is included in B, we get $\tilde{\mu}(A) \leqslant DK_2(\rho_\infty(A) + \varepsilon)$. Letting ε tend to 0, this gives $\tilde{\mu}(A) \leqslant DK_2\rho_\infty(A)$, as desired. □

Proof of Theorem 1.4. Let ρ_i be the uniform measure on the ball of radius i (which has cardinality in $[C^{-1}e^{iv}, Ce^{iv}]$). We wish to apply Theorem 2.6 to this sequence of measures. First, by Lemmas 2.12 and 2.13, the limit points of the stationary measures ν_i are equivalent to the Patterson-Sullivan measure. Therefore, they have no atom. Second, Lemma 2.8 shows that the quantity h_i in (2.6) satisfies $h_i \geqslant iv - \log C - \log(2+i)$. This tends to infinity. Hence, Theorem 2.6 applies, and gives $h(\rho_i) \geqslant (1-\varepsilon)iv$ for large i.

Using the fundamental inequality $h \leqslant \ell v$ and the trivial bound $\ell(\rho_i) \leqslant L(\rho_i) \leqslant i$, we get

$$(1-\varepsilon)iv \leqslant h(\rho_i) \leqslant \ell(\rho_i)v \leqslant iv.$$

It follows that $h(\rho_i) \sim iv$ and $\ell(\rho_i) \sim i$. □

Remark 2.14. Our technique also applies to estimate the entropy of other measures, for instance the measure $\mu_s = \sum e^{-s|g|}\delta_g / \sum e^{-s|g|}$ classically used in the construction of the Patterson-Sullivan measure. Indeed, μ_s converges when $s \searrow v$ to ρ_∞, which has no atom. Moreover, writing $Z_s = \sum e^{-s|g|}$, we have $H(\mu_s) = sL(\mu_s) + \log Z_s$. One checks that $\log Z_s$ is negligible with respect to $H(\mu_s)$, and that the quantity h_s from (2.6) is also equivalent to $H(\mu_s)$. Hence, Theorem 2.6 gives

$$H(\mu_s)(1+o(1)) \leqslant h_s(1+o(1)) \leqslant h(\mu_s) \leqslant \ell(\mu_s)v \leqslant L(\mu_s)v \leqslant H(\mu_s)(1+o(1)).$$

These inequalities show that $h(\mu_s)/\ell(\mu_s) \to v$.

Remark 2.15. One could imagine another strategy to find finitely supported measures μ_i for which $h(\mu_i)/\ell(\mu_i) \to v$. First, find a nice measure μ for which the stationary measure ν at infinity is precisely the Patterson-Sullivan measure (which implies that $h(\mu) = \ell(\mu)v$ since the Martin cocycle and the Busemann cocycle coincide). Let μ_i be a truncation of μ. Since it converges to μ, the continuity results for the drift and the entropy imply that $h(\mu_i)/\ell(\mu_i) \to h(\mu)/\ell(\mu) = v$.

We were not able to implement successfully this strategy. Given a measure ν, there is a general technique due to Connell and Muchnik [CM07] to get a measure μ on Γ with $\mu * \nu = \nu$. This technique requires a continuity assumption on $\xi \mapsto (\mathrm{d}g_*\nu/\mathrm{d}\nu)(\xi)$, which is not satisfied in our setting for $\nu = \rho_\infty$. However, in nice groups such as surface groups, this function is, for every g, continuous at all but finitely many points. The technique of [CM07] can be adapted to such a situation (in the proof of their Theorem 6.2, one should just take sets Y_n that avoid the discontinuities of the spikes we have already used). Unfortunately, the resulting measure μ (which satisfies $\mu * \nu = \nu$) has infinite moment and infinite entropy, and is therefore useless for our purposes.

3. Rigidity for admissible measures

In this section, we prove Theorem 1.5. Assume that (Γ, d) is a hyperbolic group endowed with a word distance, which is not virtually free. Let μ be a probability measure on Γ, with a superexponential moment, such that Γ_μ^+ is a finite index subgroup of Γ. We want to prove that $h(\mu) < \ell(\mu)v$. We argue by contradiction, assuming that $h(\mu) = \ell(\mu)v$. Assume first that $\Gamma_\mu^+ = \Gamma$.

Since we are assuming the equality $h(\mu) = \ell(\mu)v$, Theorem 1.2 implies that

$$|d_\mu(e,g) - vd(e,g)| \leqslant C.$$

As a warm-up, let us first deal with the baby case $C = 0$. Then the distances d_μ and d are proportional, hence they define the same Busemann boundary. The Busemann boundary $\partial_B\Gamma$ corresponding to d is totally discontinuous since the distance d takes integer values (it is a word distance). On the other hand, the Busemann boundary associated to the Green metric d_μ is known as the Martin boundary of the random walk (Γ, μ). By [Anc87] and [Gou13], it is homeomorphic to the boundary $\partial\Gamma$ of Γ. Since the group Γ is not virtually free, its boundary $\partial\Gamma$ is not totally discontinuous (see [KB02, Theorem 8.1]), hence a contradiction.

Let us now go back to the general situation, when C is nonzero (but still assuming $\Gamma_\mu^+ = \Gamma$). The argument is more complicated, but it still relies on the same facts: the boundary is not totally disconnected, while the word distance is integer valued (we will not use directly this fact, rather the fact that stable translation lengths are rational, see Lemma 3.4). These two opposite features will give rise to a contradiction.

In order to get rid of the constant C, we will need an homogenized version of the inequality $|d_\mu(e, g) - vd(e, g)| \leqslant C$. This is Lemma 3.1 below. The homogenized quantity associated to the distance d is called the stable translation length. For an element g of Γ, it is defined by $l(g) = \lim |g^n|/n$ (it exists by subadditivity).

Recall that we write $c_M(g, \xi)$ for the Martin cocycle associated to the random walk, defined in Proposition 2.5. It satisfies the cocycle relation of Definition 2.1. We will not use its probabilistic definition, but rather the fact that the Martin cocycle is the Busemann cocycle associated to the Green distance d_μ of Theorem 1.2. In other words, $c_M(g, \xi) = \lim_{x\to\xi} d_\mu(g^{-1}, x) - d_\mu(e, x)$ (and this limit exists).

Lemma 3.1. *For $g \in \Gamma$ with infinite order, $c_M(g, g^+) = vl(g)$.*

Proof. Recall that we are assuming that the equality $h(\mu) = \ell(\mu)v$ holds, therefore we have $|d_\mu(e, g) - vd(e, g)| \leqslant C$. It follows that the cocycle c_M corresponding to d_μ and the cocycle c_B corresponding to the distance d satisfy $|c_M - vc_B| \leqslant 2C$. Note that c_B is not defined on the geometric boundary, but on the horoboundary, so the proper way to write this inequality is $|c_M(g, \pi_B(\xi)) - vc_B(g, \xi)| \leqslant 2C$ for any $g \in \Gamma$ and any $\xi \in \partial_B\Gamma$.

Let $\xi \in \partial_B\Gamma$ with $\pi_B(\xi) \neq g^-$. Then $\lim c_B(g^n, \xi)/n = \lim h_\xi(g^{-n})/n = l(g)$. We choose ξ with $\pi_B(\xi) = g^+$, to get

$$\lim c_M(g^n, g^+)/n = \lim vc_B(g^n, \xi)/n \pm 2C/n = vl(g).$$

As g^+ is g-invariant, the cocycle equation for c_M on $\partial\Gamma$ gives $c_M(g, g^+) = c_M(g^n, g^+)/n$. This converges to $vl(g)$ when $n \to \infty$ by the previous equation. □

The proof of Theorem 1.5 uses the following general result on cocycles.

Proposition 3.2. *Let Γ be a hyperbolic group which is not virtually free. Let $c : \Gamma \times \partial\Gamma \to \mathbb{R}$ be a Hölder cocycle, such that any hyperbolic element g satisfies $c(g, g^+) \in \mathbb{Z}$. Then there exists a hyperbolic element $g \in \Gamma$ with $c(g, g^-) = c(g, g^+)$.*

Applied to the Busemann cocycle, this proposition implies that if a convex cocompact negatively curved manifold has a fundamental group which is not virtually free, then its length spectrum is not arithmetic, i.e., the lengths of its closed geodesics generate a dense subgroup of $\mathbb{R}$. This result is already known, see [Dal99, Page 205]. It is proved in this article using crossratios. This argument based on crossratios can be used to prove Proposition 3.2 in full generality. However, we will give a different, more direct, proof.

We will use the following topological lemma.

Lemma 3.3. *Let g be a hyperbolic element in a hyperbolic group Λ with connected boundary. There exists an arc I (i.e., a subset of $\partial\Lambda$ homeomorphic to $[0,1]$) joining g^- and g^+, invariant under an iterate g^i of g.*

Proof. We will use nontrivial results on the topology of $\partial\Lambda$. When it is connected, then it is also locally connected by [Swa96]. Hence, it is also path connected and locally path connected, see [HY61, Theorem 3-16]. Moreover, for any $\xi \in \partial\Lambda$, the space $\partial\Lambda \setminus \{\xi\}$ has finitely many ends by [Bow98b].

Consider g as in the statement of the lemma. Its action permutes the ends of $\partial\Lambda \setminus \{g^-\}$. Taking an iterate of g, we can assume it stabilizes the ends. If ξ is close to g^-, it is also the case of $g\xi$. As they belong to the same end, one can join them by a small arc J that avoids g^- (and g^+). Then $\bigcup_{n\in\mathbb{Z}} g^n J$ joins g^- to g^+, and it is invariant under g. However, it is not necessarily an arc if $g^i J$ intersects J in a nontrivial way for $i \neq 0$. To get a real arc, we will shorten J as follows.

As $g^n J$ converges to $g^{\pm}$ when n tends to $\pm\infty$, the arc J can only intersect finitely many $g^i J$. Let us fix a parametrization $u : [0,1] \to J$. The quantity

$$\inf\{|t-s| \ : \ s,t \in [0,1] \text{ and } \exists i \neq 0, u(t) = g^i u(s)\}$$

is realized by compactness (since i remains bounded), for some parameters s,t,i. Replacing s,t,i with $t,s,-i$ if necessary, we may assume $i > 0$. As g^- and g^+ are the only fixed points of g^i, we have $s \neq t$. Let $K = u([s,t])$, this is an arc between $\eta = u(s)$ and $g^i\eta = u(t)$. Moreover, $g^j K$ does not intersect K, except maybe at its endpoints for $j = \pm i$: otherwise, there exists x in the interior of K such that $g^j x$ also belongs to K, contradicting the minimality of $|s-t|$.

It follows that $\bigcup_{n\in\mathbb{Z}} g^{ni} K$ is an arc from g^- to g^+, invariant under g^i. □

Proof of Proposition 3.2. Let us consider the cocycle $\bar{c} = c \mod \mathbb{Z}$. The assumption of the proposition ensures that $\bar{c}(g, g^+) = 0$ for all hyperbolic elements g. In geometric terms, this would correspond to an assumption that the cocycle has vanishing average on all closed orbits. Hence, we may apply a version of Livsic's theorem, due in this context to [INO08] (Theorem 5.1). It ensures that the cocycle $\bar{c}$ is a coboundary: there exists a Hölder continuous function $\bar{b} : \partial\Gamma \to \mathbb{R}/\mathbb{Z}$ such that, for all $\xi \in \partial\Gamma$, for all $g \in \Gamma$,

$$\bar{c}(g,\xi) = \bar{b}(g\xi) - \bar{b}(\xi). \tag{3.1}$$

Recall that, since the group Γ is not virtually free, its boundary is not totally discontinuous (see [KB02, Theorem 8.1]). The stabilizer of a nontrivial component L of $\partial\Gamma$ is a subgroup Λ of Γ, quasi-convex hence hyperbolic, whose boundary is L (see the discussion on top of Page 55 in [Bow98a]).

Let us consider an infinite order element $g \in \Lambda$. Lemma 3.3 constructs an arc I from g^- to g^+ in $\partial\Lambda \subset \partial\Gamma$, invariant under an iterate g^i of g. Replacing g with g^i, we may assume $i = 1$.

The restriction of the function $\bar{b}$ to the arc I admits a continuous lift $b : I \to \mathbb{R}$, as I is simply connected. The function $F : \xi \mapsto c(g,\xi) - b(g\xi) + b(\xi)$ is well defined on

I, continuous, and it vanishes modulo $\mathbb{Z}$ by (3.1). Hence, it is constant. In particular, $c(g, g^-) = F(g^-) = F(g^+) = c(g, g^+)$. □

In order to apply Proposition 3.2, we will need the following result on stable translation lengths in hyperbolic groups ([BH99, Theorem III.Γ.3.17]).

Lemma 3.4. *Let (Γ, d) be a hyperbolic group with a word distance. Then there exists an integer N such that, for any $g \in \Gamma$, one has $Nl(g) \in \mathbb{Z}$.*

The combination of Lemma 3.1 and Lemma 3.4 shows that the cocycle $c' = Nc_M/v$ satisfies $c'(g, g^+) \in \mathbb{Z}$ for any hyperbolic element g. Moreover, this cocycle is Hölder-continuous since the Martin cocycle c_M is itself Hölder-continuous. This follows from [INO08] if μ has finite support, and from [Gou13] if it has a superexponential moment. Now, Proposition 3.2 implies the existence of a hyperbolic element g such that $c_M(g, g^+) = c_M(g, g^-)$. This is a contradiction since $c(g, g^+) = vl(g) > 0$ and $c(g, g^-) = -c(g^{-1}, g^-) = -vl(g) < 0$ again by Lemma 3.1. This concludes the proof of Theorem 1.5 when $\Gamma_\mu^+ = \Gamma$.

If Γ_μ^+ is a finite index subgroup of Γ, the same proof almost works in Γ_μ^+ to conclude that Γ_μ^+ is virtually free if $h = \ell v$, implying that Γ is also virtually free. The only difficulty is that the distance we are considering on Γ_μ^+ is not a word distance for a system of generators of Γ_μ^+. However, the only properties of the distance we have really used are:

(1) It is hyperbolic and quasi-isometric to a word distance (to apply Theorem 1.2).
(2) The stable translation lengths are rational numbers with bounded denominators.

These two properties are clearly satisfied for the restriction of the distance d to Γ_μ^+. Hence, the above proof also works in this case. This completes the proof of Theorem 1.5. □

Remark 3.5. If Λ is a quasi-convex subgroup of a hyperbolic group Γ, then the restriction to Λ of a word distance on Γ also satisfies the above two properties. Hence, Theorem 1.5 also holds in Λ for such a distance.

4. Growth of non-distorted points in subgroups

Our goal in this section is to prove Theorem 1.6 on the entropy of a random walk on an infinite index subgroup Λ of a hyperbolic group Γ. Since the geometry of such random walks is complicated to describe in general, our argument is indirect: we will show that, in any infinite index subgroup, the number of points that the random walk effectively visits is exponentially small compared to the growth of Γ. This is trivial if the growth $v_\Lambda = \liminf_{n\to\infty} \frac{\log|B_n \cap \Lambda|}{n}$ is strictly smaller than $v = v_\Gamma$. When $v_\Lambda = v$, on the other hand, we will argue that the random walk does not typically visit all of Λ, but only a subset made of non-distorted points. To prove Theorem 1.6, the main step is to show that, even when $v_\Lambda = v$, the number of such non-distorted points is exponentially smaller than e^{nv}. We introduce the notion of non-distorted points in Paragraph 4.1, prove this main geometric estimate in Paragraph 4.2, and apply this to random walks in Paragraph 4.4. Paragraph 4.3 is devoted to the case $v_\Lambda < v$, where unexpected phenomena happen even in distorted subgroups.

4.1. **Non-distorted points.** There are at least two different ways to define a notion of non-distorted point.

Definition 4.1. *Let Γ be a finitely generated group endowed with a word distance $d = d_\Gamma$, and let Λ be a subgroup of Γ.*

- *For $\varepsilon > 0$ and $M > 0$, we say that $g \in \Lambda$ is (ε, M)-quasi-convex if any geodesic γ from e to g spends at least a proportion ε of its time in the M-neighborhood of Λ, i.e.,*
$$|\{i \in [1, |g|] \,:\, d(\gamma(i), \Lambda) \leqslant M\}| \geqslant \varepsilon |g|.$$
We write $\Lambda_{QC(\varepsilon,M)}$ for the set of points in Λ which are (ε, M)-quasi-convex.
- *Assume additionally that Λ is finitely generated, and endowed with a word distance d_Λ. For $D > 0$, we say that $g \in \Lambda$ is D-undistorted if $d_\Lambda(e, g) \leqslant D d_\Gamma(e, g)$. We write $\Lambda_{UD(D)}$ for the set of D-undistorted points.*

Up to a change in the constants, these notions do not depend on the choice of the distance d. The first definition has the advantage to work for infinitely generated subgroups, but it may seem less natural than the second one. If Λ is a quasi-convex subgroup of a hyperbolic group Γ, then all its points are $(1, M)$-quasi-convex if M is large enough, and all its points are also D-undistorted for large enough D. In the general case, a quasi-convex point does not have to be undistorted: it may happen that the times i such that $d(\gamma(i), \Lambda) \leqslant M$ are all included in $[1, |g|/2]$, while between $|g|/2$ and $|g|$ one needs to make a huge detour to follow Λ, making $d_\Lambda(e, g)$ much larger than $d_\Gamma(e, g)$. On the other hand, an undistorted point is automatically quasi-convex, at least in hyperbolic groups:

Proposition 4.2. *Let Γ be a hyperbolic group, let Λ be a finitely generated subgroup of Γ, and let $D > 0$. There exist $\varepsilon > 0$ and $M > 0$ such that any D-undistorted point is also (ε, M)-quasi-convex, i.e., $\Lambda_{UD(D)} \subset \Lambda_{QC(\varepsilon,M)}$.*

Proof. Consider $g \in \Lambda$ which is not (ε, M)-quasi-convex, we have to show that $d_\Lambda(e, g)$ is much bigger than $n = d_\Gamma(e, g)$. The intuition is that, away from a Γ-geodesic from e to g, the progress towards g is much slower by hyperbolicity.

Let us consider a geodesic from e to g in Λ, with length $d_\Lambda(e, g)$. Replacing each generator of Λ by the product of a uniformly bounded number of generators of Γ, we obtain a path γ_Λ in the Cayley graph of Γ, remaining in the C_0-neighborhood of Λ (for some $C_0 > 0$) and with length $|\gamma_\Lambda| \leqslant C_0 d_\Lambda(e, g)$.

Let us consider a geodesic γ_Γ from e to g for the distance d_Γ. For each $x \in \Gamma$, we can consider its projection $\pi(x)$ on γ_Γ, i.e., the point on γ_Γ that is closest to x (if several points correspond, we take the closest one to e). This projection is 1-Lipschitz. In particular, the projection of γ_Λ covers the whole geodesic γ_Γ. For each $x_i \in \gamma_\Gamma$, let us consider the first point $y_i \in \gamma_\Lambda$ projecting to x_i.

Let us fix an integer L, large enough with respect to the hyperbolicity constant of Γ. Along γ_Γ, let us consider the points at distance kL from e, i.e., $x_0 = e, x_L, x_{2L}, \dots, x_{mL}$ with $m = \lfloor n/L \rfloor$. In particular, $|\gamma_\Lambda| \geqslant \sum_i d_\Gamma(y_{iL}, y_{(i+1)L})$. Moreover, a tree approximation shows that $d_\Gamma(y_{iL}, y_{(i+1)L}) \geqslant d_\Gamma(y_{iL}, x_{iL}) + L + d_\Gamma(x_{(i+1)L}, y_{(i+1)L}) - C_1$ (where C_1 only depends on the hyperbolicity constant of Γ). Choosing $L \geqslant C_1$, we get

$$|\gamma_\Lambda| \geqslant \sum_{i=0}^{m} d_\Gamma(x_{iL}, y_{iL}) \geqslant \sum_{i=0}^{m} (d_\Gamma(x_{iL}, \Lambda) - C_0).$$

Since we assume that g is not (ε, M)-quasi-convex, the set of indices i with $d(x_i, \Lambda) \leqslant M$ has cardinality at most εn. Taking $M \geqslant C_0$, the previous equation is bounded from below by

$$(m + 1 - \varepsilon n)M - (m + 1)C_0 \geqslant (n/L - \varepsilon n)M - nC_0/L.$$

Finally, we get

$$d_\Lambda(e, g) \geqslant |\gamma_\Lambda|/C_0 \geqslant n(1/L - \varepsilon)M/C_0 - n/L.$$

If ε is small enough and M is large enough so that $(1/L - \varepsilon)M/C_0 - 1/L > D$, we obtain $d_\Lambda(e, g) > Dn$, i.e., $g \notin \Lambda_{UD(D)}$, as desired. □

From this point on, we will mainly work with the notion of quasi-convex points, since counting results on such points imply results on undistorted points by the previous proposition.

4.2. **Non-distorted points in subgroups with $v_\Lambda = v$.** In this section, we show that there are exponentially few quasi-convex points in infinite-index subgroups of hyperbolic groups.

Theorem 4.3. *Let Γ be a nonelementary hyperbolic group endowed with a word distance. Let Λ be an infinite index subgroup of Γ. Then*

$$|B_n \cap \Lambda| = o(|B_n|). \tag{4.1}$$

Moreover, for all $\varepsilon > 0$ and $M > 0$, there exists $\eta > 0$ such that, for all large enough n,

$$|B_n \cap \Lambda_{QC(\varepsilon,M)}| \leqslant e^{-\eta n}|B_n|. \tag{4.2}$$

One may wonder why we put the estimate (4.1) in the statement of the theorem, while the main emphasis is on counting quasi-convex points. It turns out that this estimate is not trivial, and that its proof uses the same techniques as for the proof of (4.2). To illustrate that it is not trivial, let us remark that this estimate is not true without the hyperbolicity assumption. For instance, in $\Gamma = \mathbb{F}_2 \times \mathbb{Z}$ (with its canonical generating system, and the corresponding word distance), the infinite index subgroup $\Lambda = \mathbb{F}_2$ satisfies $|\Lambda \cap B_n|/|B_n| \geqslant c > 0$.

Theorem 4.3 is trivial if the growth rate v_Λ of Λ is strictly smaller than the growth rate v of Γ, since in this case $|B_n \cap \Lambda|$ itself is exponentially smaller than $|B_n|$. However, this is not always the case, even for finitely generated subgroups.

Consider for instance a compact hyperbolic 3-manifold which fibers over the circle, obtained as a suspension of a hyperbolic surface with a pseudo-Anosov. Its fundamental group Γ surjects into $\mathbb{Z} = \pi_1(\mathbb{S}^1)$. The kernel Λ of this morphism φ is the fundamental group of the fiber. It is finitely generated, with infinite index, and $|B_n \cap \Lambda| \sim c|B_n|/\sqrt{n}$, see [Sha98].

Heuristically, one can understand in this case why there are exponentially few quasi-convex points in Λ. Let us consider a geodesic of length n in Γ. It projects under φ to a path in $\mathbb{Z}$, which behaves roughly like a random walk. In particular, $e^{-nv}|\mathbb{S}^n \cap \Lambda|$ behaves like the probability that a random walk on $\mathbb{Z}$ comes back to the identity at time n. This is of order $1/\sqrt{n}$, in accordance with the rigorous results of [Sha98]. Such an element is quasi-convex if the random walk in $\mathbb{Z}$ spends a big proportion of its time close to the origin. A large deviation estimate shows that this is exponentially unlikely.

The proof of the theorem consists in making this heuristic precise, in the general case where the subgroup Λ is not normal (so that there is no morphism φ at hand). An important point in the proof is that a hyperbolic group is automatic, i.e., there exists a finite state automaton that recognizes a system of geodesics parameterizing bijectively the points in the group. Counting points in the group then amounts to a random walk on the graph of this automaton, while counting points in Λ amounts to a fibred random walk, on this graph times $\Lambda\backslash\Gamma$. As this space is infinite, the random walk spends most of its time outside of finite sets, i.e., far away from Λ.

To formalize this argument, we will reduce the question to Markov chains on graphs, where we will use the following probabilistic lemma.

Lemma 4.4. *Consider a Markov chain (X_n) on a countable set V, with a stationary measure m (i.e., $m(x) = \sum_y m(y)p(y,x)$ for all x). Let $\tilde V$ be the set of points $x \in V$ such that $\sum_{x\to y} m(y) = +\infty$, where we write $x \to y$ if there exists a positive probability path from x to y. Then, for all $x \in V$ and $x' \in \tilde V$,*

$$\mathbb{P}_x(X_n = x') \to 0 \text{ when } n \to \infty. \tag{4.3}$$

Take $x \in \tilde V$ and $\varepsilon > 0$. There exists $\eta > 0$ such that, for all large enough n,

$$\mathbb{P}_x(X_n = x \text{ and } X_i \text{ visits } x \text{ at least } \varepsilon n \text{ times in between}) \leqslant e^{-\eta n}. \tag{4.4}$$

Proof. In countable state Markov chains, a point x can be either transient, or null recurrent, or positive recurrent. Let us first show that points in $\tilde V$ are not positive recurrent, by contradiction. Otherwise, the points that can be reached from x form an irreducible class $\mathcal{C}$, which admits a stationary probability measure p. The restriction of m to $\mathcal{C}$ is an excessive measure. By uniqueness (see [Rev84, Theorem 3.1.9]), the measure m is proportional on $\mathcal{C}$ to p. In particular, it has finite mass there. This contradicts the assumption $\sum_{x\to y} m(y) = +\infty$.

Let us now show that, for all $x \in V$ and $x' \in \tilde V$, the probability $\mathbb{P}_x(X_n = x')$ tends to 0. Otherwise, conditioning on the first visit to x', we deduce that $\mathbb{P}_{x'}(X_n = x')$ does not tend to 0. This implies that x' is positive recurrent, a contradiction.

Let us now prove (4.4). Consider $x \in \tilde V$, it is either transient or null recurrent. If it is transient, the probability p to come back to x is < 1. Hence, the probability to come back εn times is bounded by $p^{\varepsilon n}$, and is therefore exponentially small as desired.

Assume now that x is null recurrent: almost surely, the Markov chain comes back to x, but the waiting time τ has infinite expectation. Let $\tau_1, \tau_2, \dots$ be the length of the successive excursions based at x. They are independent and distributed like τ, by the Markov property. The probability in (4.4) is bounded by $\mathbb{P}(\sum_{i=1}^{\varepsilon n} \tau_i \leqslant n)$, which is bounded for any M by $\mathbb{P}(\sum_{i=1}^{\varepsilon n} \tau_i 1_{\tau_i \leqslant M} \leqslant n)$. The random variables $\tau_i 1_{\tau_i\leqslant M}$ are bounded, independent and identically distributed. If M is large enough, they have expectation $> 1/\varepsilon$. A standard large deviation result then shows that $\mathbb{P}(\sum_{i=1}^{\varepsilon n} \tau_i 1_{\tau_i\leqslant M} \leqslant n)$ is exponentially small, as desired. □

We will also need the following technical lemma, which was explained to us by B. Bekka.

Lemma 4.5. *Let Λ be a subgroup of a group Γ. Assume that there exists a finite subset B of Γ such that $B\Lambda B = \Gamma$. Then Λ has finite index in Γ.*

Proof. We have by assumption $\Gamma = \bigcup_{i,j} b_i \Lambda b_j = \bigcup_{i,j} \Lambda_i b_i b_j$, where $\Lambda_i = b_i \Lambda b_i^{-1}$ is a conjugate of Λ (and has therefore the same index). A theorem of Neumann [Neu54] ensures that a group is never a finite union of right cosets of infinite index subgroups. Hence, one of the Λ_i has finite index in Γ, and so has Λ. □

Let Γ be a hyperbolic group, with a finite generating set S. Consider a finite directed graph $\mathcal{A} = (V, E, x_*)$ with vertex set V, edges E, a distinguished vertex x_*, and a labeling $\alpha : E \to S$. We associate to any path γ in the graph (i.e., a sequence of edges $\sigma_0, \sigma_1, \dots, \sigma_{m-1}$ where the endpoint of σ_i is the beginning of σ_{i+1}) a path in the Cayley graph starting from the identity and following the edges labeled $\alpha(\sigma_0)$, then $\alpha(\sigma_1)$, and so on. The endpoint of this path is $\alpha_*(\gamma) := \alpha(\sigma_0) \cdots \alpha(\sigma_{m-1})$. We always assume that any point can be reached by a path starting at x_*.

A hyperbolic group is *automatic* (see, for instance, [Cal13]): there exists such a graph with the following properties.

(1) For any path γ in the graph, the corresponding path $\alpha(\gamma)$ is geodesic in the Cayley graph.
(2) The map α_* induces a bijection between the set of paths in the graph starting from x_* and the group Γ.

In particular, the paths of length n in the graph originating from x_* parameterize the sphere $\mathbb{S}^n$ of radius n in the group. The existence of such a structure makes it for instance possible to prove that the growth series of a hyperbolic group is rational. We will use such an automaton to count the points in the subgroup Λ, and in particular the quasi-convex points.

We define a transition matrix A, indexed by V. By definition, A_{xy} is the number of edges from x to y. Hence, $(A^n)_{xy}$ is the number of paths of length n from x to y. In particular, the number of paths of length n starting from x_* is $\sum_y (A^n)_{x_* y}$. Write u for the line vector with 1 at position x_* and 0 elsewhere, and $\tilde{u}$ for the column vector with 1 everywhere. This number of paths reads $uA^n\tilde{u}$. Therefore, $|\mathbb{S}^n| = uA^n\tilde{u}$, proving the rationality of the growth function of the group. Let v be the growth rate of balls in Γ. It satisfies $|B_n| \leqslant Ce^{nv}$, by [Coo93]. Hence, the spectral radius of A is e^v, and A has no Jordan block for this maximal eigenvalue.

To understand the points of the infinite index subgroup Λ of Γ, we consider an extension $\mathcal{A}_\Lambda$ of $\mathcal{A}$, with fibers $\Lambda\backslash\Gamma$. Its vertex set V_Λ is made of the pairs $(x, \Lambda g) \in V \times \Lambda\backslash\Gamma$. For any edge σ in $\mathcal{A}$, going from x to y and with label $\alpha(\sigma)$, we put for any $g \in \Gamma$ an edge in $\mathcal{A}_\Lambda$ from $(x, \Lambda g)$ to $(y, \Lambda g\alpha(\sigma))$. A path γ in $\mathcal{A}$, from x to y, lifts to a path $\tilde{\gamma}$ in $\mathcal{A}_\Lambda$ originating from $(x, \Lambda e)$. By construction, its endpoint is $(y, \Lambda\alpha_*(\gamma))$. This shows that the paths in the graph $\mathcal{A}_\Lambda$ remember the current right coset of Λ.

The next lemma proves that the relevant components of this fibred graph are infinite.

Lemma 4.6. *Let $\tilde{x}_0 = (x_0, \Lambda g_0)$ belong to $\mathcal{A}_\Lambda$. Let $\mathcal{C}$ be the component of x_0 in $\mathcal{A}$ (i.e., the set of points that can be reached from x_0 and from which one can go back to x_0). Let $A_{\mathcal{C}}$ be the restriction of the matrix A to the points in $\mathcal{C}$. Assume that its spectral radius $\rho(A_{\mathcal{C}})$ is* *equal to e^v. Then, starting from $\tilde{x}_0$ in the graph $\mathcal{C}_\Lambda$ (the restriction of $\mathcal{A}_\Lambda$ to $\mathcal{C} \times \Lambda\backslash\Gamma$), one can reach infinitely many different points of $\mathcal{C}_\Lambda$.*

Proof. It suffices to show that one can reach infinitely many points whose component in $\mathcal{C}$ is x_0. Assume by contradiction that one can only reach a finite number of classes $(x_0, \Lambda g_i)$.

Given $w \in \Gamma$ and $C > 0$, let $Y_{w,C}$ be the set of points in Γ that have a geodesic expression in which, for any subword $\tilde{w}$ of this expression and for any a, b with length at most C, one has $w \neq a\tilde{w}b$. In other words, the points in $Y_{w,C}$ are those that never see w (nor even a thickening of w of size C) in their geodesic expressions. Theorem 3 in [AL02] proves the existence of C_0 such that, for any w, the quantity $|B_n \cap Y_{w,C_0}|/|B_n|$ tends to 0 (the important point is that C_0 does not depend on w).

The number of paths in $\mathcal{C}$ originating from x_0 grows at least like $c|B_n|$ since the spectral radius of $A_\mathcal{C}$ is e^v. These paths give rise to distinct points in Γ. Hence, there exists such a path γ_0 such that $\alpha_*(\gamma_0) \notin Y_{w,C_0}$. In particular, there exists a subpath γ_1 such that $\alpha_*(\gamma_1)$ can be written as a_1wb_1 with $|a_1| \leqslant C_0$ and $|b_1| \leqslant C_0$. We can choose a path from x_0 to the starting point of γ_1, with fixed length (since $\mathcal{C}$ is finite), and another path from the endpoint of γ_1 to x_0. Concatenating them, we get a path γ_2 from x_0 to itself with $\alpha_*(\gamma_2) = a_2wb_2$ and $|a_2|, |b_2| \leqslant C_1 = C_0 + 2\operatorname{diam}(\mathcal{C})$. By assumption, $\Lambda g_0\alpha_*(\gamma_2)$ is one of the finitely many Λg_i since we are returning to x_0. Hence, there exists $\lambda \in \Lambda$ such that $g_0a_2wb_2 = \lambda g_i$. This shows that $w \in B\Lambda B$, where B is the ball of radius $C_1 + \max_i d(e, g_i)$. As this holds for any w, we have proved that $B\Lambda B = \Gamma$. By Lemma 4.5, this shows that Λ has finite index in Γ, a contradiction. □

Lemma 4.7. *Let $K(n, \tilde{x}_0, \varepsilon_0)$ denote the set of paths in $\mathcal{A}_\Lambda$ starting at a point $\tilde{x}_0$, of length n, coming back to $\tilde{x}_0$ at time n, and spending a proportion at least ε_0 of the time at $\tilde{x}_0$. Consider $\tilde{x}_0 \in \mathcal{A}_\Lambda$ and $\varepsilon_0 > 0$. Then there exist $\eta > 0$ and $C > 0$ such that, for all $n \in \mathbb{N}$,*

$$|K(n, \tilde{x}_0, \varepsilon_0)| \leqslant Ce^{n(v-\eta)}.$$

Proof. Write $\tilde{x}_0 = (x_0, \Lambda g_0)$, let $\mathcal{C}$ be the component of x_0 in $\mathcal{A}$. If the spectral radius of the restricted transition matrix $A_\mathcal{C}$ is $< e^v$, we simply bound $|K(n, \tilde{x}_0, \varepsilon_0)|$ by the number of paths in $\mathcal{C}$ from x_0 to itself. This is at most $\|A_\mathcal{C}^n\|$, which is exponentially smaller than e^{nv} as desired.

Assume now that $\rho(A_\mathcal{C}) = e^v$. We will understand the number of paths in $\mathcal{C}$ (and in its lift $\mathcal{C}_\Lambda$) in terms of a Markov chain. The matrix $A_\mathcal{C}$ has a unique eigenvector q corresponding to the eigenvalue e^v, it is positive by Perron-Frobenius's theorem. By definition, $p(x,y) = e^{-v}A_{xy}q(y)/q(x)$ satisfies, for any $x \in \mathcal{C}$,

$$\sum_{y\in\mathcal{C}} p(x,y) = \frac{e^{-v}}{q(x)} \sum A_{xy}q(y) = 1.$$

This means that $p(x,y)$ is a transition kernel on $\mathcal{C}$. Denote by $(X_n)_{n\in\mathbb{N}}$ the corresponding Markov chain. By construction,

$$\mathbb{P}_x(X_n = y) = e^{-nv}(A^n)_{xy}q(y)/q(x).$$

Moreover, $(A^n)_{xy}$ is the number of paths of length n in $\mathcal{A}$ from x to y. Hence, up to a bounded multiplicative factor $q(y)/q(x)$, the transition probabilities of the Markov chain X_n

count the number of paths in the graph $\mathcal{C}$. Let m denote the unique stationary probability for the Markov chain on $\mathcal{C}$.

We lift everything to $\mathcal{C}_\Lambda$, assigning to an edge the transition probability of its projection in $\mathcal{C}$. The stationary measure m lifts to a stationary measure m_Λ, which is simply the product of m and of the counting measure in the direction $\Lambda\backslash\Gamma$. Denoting by X_n^Λ the Markov chain in $\mathcal{C}_\Lambda$, we have

$$e^{-nv}|K(n,\tilde{x}_0,\varepsilon_0)| = \mathbb{P}_{\tilde{x}_0}(X_n^\Lambda = \tilde{x}_0 \text{ and } X_i^\Lambda \text{ visits } \tilde{x}_0 \text{ at least } \varepsilon_0 n \text{ times in between}).$$

By Lemma 4.6, the Markov chain starting from $\tilde{x}_0$ can reach infinitely many points. Equivalently, since m is bounded from below, it can reach a set of infinite m_Λ-measure. Therefore, Lemma 4.4 applies, and shows that the above quantity is exponentially small. □

Proof of Theorem 4.3. Let us first prove (4.2). Counting the points in $\mathbb{S}^n \cap \Lambda_{QC(\varepsilon,M)}$ amounts to counting the paths of length n in $\mathcal{A}_\Lambda$, starting from $(x_*,\Lambda e)$ and spending a proportion at least ε of their time in the finite subset $F = V \times \Lambda B_M \subset V_\Lambda$. Such a path spends a proportion at least $\varepsilon_0 = \varepsilon/|F|$ of its time at a given point $\tilde{x} \in F$. Let k and $k+m$ denote the first and last visits to $\tilde{x}$ (with $m \geqslant \varepsilon_0 n$ since there are at least $\varepsilon_0 n$ visits). Such a path is the concatenation of a path from $(x_*,\Lambda e)$ to $\tilde{x}$ of length k (their number is bounded by the corresponding number of paths in $\mathcal{A}$, at most $\|A^k\| \leqslant Ce^{kv}$), of a path in $K(m,\tilde{x},\varepsilon_0)$, and of a path starting from $\tilde{x}$ of length $n-k-m$ (their number is again bounded by the number of corresponding paths in $\mathcal{A}$, at most $Ce^{(n-k-m)v}$). Hence, their number is at most $Ce^{(n-m)v}|K(m,\tilde{x},\varepsilon_0)|$. Summing over the points $\tilde{x} \in F$, over the at most n possible values of k, and the values of m, we get the inequality

$$|\mathbb{S}^n \cap \Lambda_{QC(\varepsilon,M)}| \leqslant Cne^{nv} \sum_{\tilde{x}\in F} \sum_{m=\varepsilon_0 n}^{n} e^{-mv}|K(m,\tilde{x},\varepsilon_0)|.$$

Lemma 4.7 shows that this is exponentially smaller than e^{nv}.

Let us now prove (4.1), using similar arguments. A point in $\mathbb{S}^n \cap \Lambda$ corresponds to a path of length n in $\mathcal{A}_\Lambda$, starting from $(x_*,\Lambda e)$ and ending at a point $(x,\Lambda e)$. We say that a component $\mathcal{C}$ in the graph $\mathcal{A}$ is maximal if the spectral radius of the corresponding restricted matrix $A_{\mathcal{C}}$ is e^v. Since the matrix A has no Jordan block corresponding to the eigenvalue e^v, a path in the graph encounters at most one maximal component. The paths in $\mathcal{A}_\Lambda$ whose projection in $\mathcal{A}$ spends a time k in non-maximal components give an overall contribution to $|\mathbb{S}^n \cap \Lambda|$ bounded by $Ce^{(n-k)v+k(v-\eta)} \leqslant Ce^{-\eta k}|B_n|$. Given $\varepsilon > 0$, their contribution for $k \geqslant k_0(\varepsilon)$ is bounded by $\varepsilon|B_n|$. Hence, it suffices to control the paths for fixed k. Let us fix the beginning of such a path, from $(x_*,\Lambda e)$ to a point $(x_0,\Lambda g_0)$ where x_0 is in a maximal component $\mathcal{C}$, and its end from $(x_1,\Lambda g_1)$ with $x_1 \in \mathcal{C}$ to a point $(x,\Lambda e)$. To conclude, one should show that the number of paths of length n from $(x_0,\Lambda g_0)$ to $(x_1,\Lambda g_1)$ is $o(e^{nv})$. This follows from the probabilistic interpretation in the proof of Lemma 4.7 and from (4.3). □

4.3. **Non-distorted points in subgroups with $v_\Lambda < v$.** Let Λ be a subgroup of a hyperbolic group Γ. Let v_Λ and v_Γ be their respective growths, for a word distance on Γ. If $v_\Lambda = v_\Gamma$, Theorem 4.3 proves that there is a dichotomy:

(1) Either Λ is quasi-convex (equivalently, Λ has finite index in Γ). Then $|B_n \cap \Lambda| \geqslant ce^{nv_\Lambda}$, and all points in Λ are quasi-convex.

(2) Or Λ is not quasi-convex (equivalently, it has infinite index in Γ). Then $|B_n \cap \Lambda| = o(e^{nv_\Lambda})$, and there are exponentially few quasi-convex points in Λ.

Consider now a general subgroup Λ with $v_\Lambda < v_\Gamma$. If it is quasi-convex, then (1) above is still satisfied: $|B_n \cap \Lambda| \geqslant ce^{nv_\Lambda}$ by [Coo93], and all points in Λ are quasi-convex. One may ask if these properties are equivalent, and if they characterize quasi-convex subgroups. This question is reminiscent of a question of Sullivan in hyperbolic geometry: Are convex cocompact groups the only ones to have finite Patterson-Sullivan measure? Peigné showed in [Pei03] that the answer to this question is negative. His counterexamples adapt to our situation, giving also a negative answer to our question.

Proposition 4.8. *There exists a finitely generated subgroup Λ of a hyperbolic group Γ endowed with a word distance, which is not quasi-convex, but for which $C^{-1}e^{nv_\Lambda} \leqslant |B_n \cap \Lambda| \leqslant Ce^{nv_\Lambda}$. Moreover, most points of Λ are quasi-convex: there exist ε and η such that*

$$|B_n \cap \Lambda \setminus \Lambda_{QC(\varepsilon,0)}| \leqslant Ce^{n(v_\Lambda - \eta)}. \tag{4.5}$$

Proof. The example is the same as in [Pei03], but his geometric proofs are replaced by combinatorial arguments based on generating series.

Let G be a finitely generated non-quasi-convex subgroup of a hyperbolic group $\tilde{G}$ (take for instance for $\tilde{G}$ the fundamental group of a hyperbolic 3-manifold which fibers over the circle, and for G the fundamental group of the fiber of this fibration). Let $H = \mathbb{F}_k$, with k large enough so that $v_H \geqslant v_G$. We take $\Lambda = G * H \subset \Gamma = \tilde{G} * H$. It is not quasi-convex, because of the factor G. Writing v_Λ for its growth, we claim that, for some $c > 0$,

$$|\mathbb{S}^n \cap \Lambda| \sim ce^{nv_\Lambda}. \tag{4.6}$$

We compute with generating series. Let $F_G(z)$ be the growth series for G, given by $F_G(z) = \sum_{n\geqslant 0} |\mathbb{S}^n \cap G| z^n$. Likewise, we define F_H and F_Λ. Since any word in Λ has a canonical decomposition in terms of words in G and H, a classical computation (see [dlH00, Prop. VI.A.4]) gives

$$F_\Lambda = \frac{F_G F_H}{1 - (F_G - 1)(F_H - 1)}. \tag{4.7}$$

Let $z_G = e^{-v_G} \geqslant z_H = e^{-v_H}$ be the convergence radii of F_G and F_H. At z_H, we have $F_H(z_H) = +\infty$, since the cardinality of spheres in the free group is exactly of the order of e^{nv_H}. When z increases to z_H, the function $(F_G(z) - 1)(F_H(z) - 1)$ takes the value 1, at a number $z = z_\Lambda$. Since this is the first singularity of F_Λ, we have $z_\Lambda = e^{-v_\Lambda}$. Moreover, the function F_Λ is meromorphic at z_Λ, with a pole of order 1 (since the function $(F_G - 1)(F_H - 1)$ has positive derivative, being a power series with nonnegative coefficients). It follows from a simple tauberian theorem (see, for instance, [FS09, Theorem IV.10]) that the coefficients of F_Λ behave like cz_Λ^{-n}, proving (4.6).

Let us estimate the number of non-quasi-convex points in Λ. Consider a word $w \in \Lambda$ of length n, for instance starting with a factor in G and ending with a factor in H. It can be written as $g_1 h_1 g_2 h_2 \cdots h_s$. Along a geodesic from e to w, all the words $g_1 h$ (with h prefix of h_1) belong to Λ. So do all the words $g_1 h_1 g_2 h$ with h prefix of h_2, and so on. Therefore, the proportion of time that the geodesic spends outside of Λ is at most $\sum |g_i|/n$. Such a point in $\Lambda \setminus \Lambda_{QC(\varepsilon,0)}$ satisfies $\sum |g_i| \geqslant (1-\varepsilon)n$ and $\sum |h_i| \leqslant \varepsilon n$. Assuming $\varepsilon \leqslant 1/2$, this gives $\sum|h_i| \leqslant (\varepsilon/2)\sum|g_i|$. In particular, for any $\alpha > 0$, we have $e^{\alpha(\sum|g_i|-2\varepsilon^{-1}\sum|h_i|)} \geqslant 1$. Let $u_n = |\mathbb{S}^n \cap \Lambda \setminus \Lambda_{QC(\varepsilon,0)}|$, its generating series satisfies the following equation (where we only write in details the words starting with G and ending in H, the other ones being completely analogous):

$$\begin{aligned}\sum u_n z^n &\leqslant \sum_{\ell\geqslant 1}\sum_{a_1+b_1+a_2+\dots+b_\ell=n} e^{\alpha(\sum a_i - 2\varepsilon^{-1}\sum b_i)}|\mathbb{S}^{a_1}\cap G||\mathbb{S}^{b_1}\cap H|\cdots|\mathbb{S}^{b_\ell}\cap H|z^n + \dots \\ &= \sum_{\ell\geqslant 1}\left[(F_G(e^\alpha z)-1)(F_H(e^{-2\alpha\varepsilon^{-1}}z)-1)\right]^\ell + \dots \\ &= \frac{F_G(e^\alpha z)F_H(e^{-2\alpha\varepsilon^{-1}}z)}{1-(F_G(e^\alpha z)-1)(F_H(e^{-2\alpha\varepsilon^{-1}}z)-1)}.\end{aligned}$$

This is the same formula as in (4.7), but the factor z has been shifted in F_G and F_H. Choose $\alpha > 0$ such that $e^\alpha z_\Lambda < z_G$, and then ε small enough so that $(F_G(e^\alpha z_\Lambda)-1)(F_H(e^{-2\alpha\varepsilon^{-1}}z_\Lambda)-1) < 1$. We deduce that the series $\sum u_n z^n$ converges for $z = z_\Lambda$, and even slightly to its right. It follows that u_n is exponentially small compared to z_Λ^{-n}. This proves (4.5). □

4.4. **Application to random walks in infinite index subgroups.** In this paragraph, we use Theorem 4.3 to prove Theorem 1.6 on random walks given by a measure μ on a hyperbolic group Γ, assuming that Γ_μ has infinite index in Γ.

Before proving Theorem 1.6, we give another easier result, pertaining to the case where μ has a finite moment for a word distance on Γ_μ (which should be finitely generated): In this case, the random walk typically visits undistorted points. This easy statement is not used later on, but it gives a heuristic explanation to Theorem 1.6.

Lemma 4.9. *Let Λ be a finitely generated subgroup of a finitely generated group Γ. Let d_Λ and d_Γ be the two corresponding word distances. Consider a probability measure μ on Λ, with a moment of order 1 for d_Λ (and therefore for d_Γ), with nonzero drift for d_Γ. Let X_n denote the corresponding random walk. There exists $D > 0$ such that $\mathbb{P}(X_n \in \Lambda_{UD(D)}) \to 1$.*

Proof. Almost surely, $d_\Gamma(e, X_n) \sim \ell_\Gamma n$, for some nonzero drift ℓ_Γ. In the same way, $d_\Lambda(e, X_n) \sim \ell_\Lambda n$. For any $D > \ell_\Lambda/\ell_\Gamma$, we get almost surely $d_\Lambda(e, X_n) \leqslant D d_\Gamma(e, X_n)$ for large enough n, i.e., $X_n \in \Lambda_{UD(D)}$. □

This lemma readily implies Theorem 1.6 under the additional assumption that Λ is finitely generated and that μ has a moment of order 1 for d_Λ. Indeed, for large n, with probability at least 1/2, the point X_n belongs to $B_{(\ell+\varepsilon)n} \cap \Lambda_{UD(D)}$, whose cardinality is bounded by $Ce^{(\ell+\varepsilon)n(v-\eta)}$ according to Theorem 4.3. Lemma 2.4 yields $h \leqslant (\ell+\varepsilon)(v-\eta)$, hence $h \leqslant \ell(v-\eta) < \ell v$, completing the proof.

However, the assumptions of Theorem 1.6 are much weaker: even when Λ is finitely generated, it is much more restrictive to require a moment of order 1 on Λ than on Γ, precisely because the Γ-distance is smaller than the Λ-distance on distorted points, which make up most of Λ. The general proof will not use undistorted points (which are not even defined when Λ is not finitely generated), but rather quasi-convex points: we will show that, typically, the random walk concentrates on quasi-convex points. With the previous argument, Theorem 1.6 readily follows from the next lemma.

Lemma 4.10. *Let Λ be a subgroup of a hyperbolic group Γ endowed with a word distance $d = d_\Gamma$. Let us consider a probability measure μ on Λ, with a moment of order 1 for d_Γ. There exist $\varepsilon > 0$ and $M > 0$ such that $\mathbb{P}(X_n \in \Lambda_{QC(\varepsilon,M)}) \geqslant 1/2$ for large enough n.*

Proof. The lemma is trivial if μ is elementary, since all the elements of $\Gamma_\mu \subset \Lambda$ are then quasi-convex. We may therefore assume that μ is non-elementary.

The random walk at time n is given by $X_n = g_1 \cdots g_n$, where g_i are independent and distributed like μ. We will show that most products $g_1 \cdots g_i$ (which belong to Λ) are within distance M of a geodesic from e to X_n (this amounts to the classical fact that trajectories of the random walk follow geodesics in the group), and moreover that they approximate a proportion at least ε of the points on this geodesic. This will give $X_n \in \Lambda_{QC(\varepsilon,M)}$ as desired. The second point is more delicate: we should for instance exclude the situation where, given a geodesic γ, one has $X_n = \gamma(a(n))$ where $a(n)$ is the smallest square larger than n. In this case, X_n follows the geodesic γ at linear speed, but nevertheless the proportion of γ it visits tends to 0. This behavior will be excluded thanks to the fact that, with high probability, the jumps of the random walk are bounded.

The argument is probabilistic and formulated in terms of the bilateral version of the random walk. On $\Omega = \Gamma^{\mathbb{Z}}$ with the product measure $\mathbb{P} = \mu^{\otimes \mathbb{Z}}$, let g_n be the n-th coordinate. The g_n are independent, identically distributed, and correspond to the increments of a random walk $(X_n)_{n\in\mathbb{Z}}$ with $X_0 = e$ and $X_n^{-1} X_{n+1} = g_{n+1}$. Almost surely, X_n converges when $n \to \pm\infty$ towards two random variables $\xi^\pm \in \partial\Gamma$, with $\xi^+ \neq \xi^-$ almost surely since these random variables are independent and atomless. Following Kaimanovich [Kai00], denote by $S(\xi^-, \xi^+)$ the union of all the geodesics from ξ^- to ξ^+. Let π be the projection on $S(\xi^-, \xi^+)$, i.e., $\pi(g)$ is the closest point to g on $S(\xi^-, \xi^+)$. It is not uniquely defined, but two possible choices are within distance C_0, for some C_0 only depending on Γ.

Let us choose $L > 0$ large enough (how large will only depend on the hyperbolicity constant of the space). Any measurable function is bounded on sets with arbitrarily large measure. Hence, there exists $K > 0$ such that, with probability at least $9/10$,

(1) For every $|k| \geqslant K$, the projections $\pi(X_k)$ are distant from $\pi(X_0)$ by at least L (and they are closer to ξ^+ if $k > 0$, and to ξ^- if $k < 0$).
(2) We have $d(e, S(\xi^-, \xi^+)) \leqslant K$.

As everything is equivariant, we deduce that, for all $i \in \mathbb{Z}$, the point X_i satisfies the same properties with probability at least $9/10$, i.e.,

$$d(X_i, S(\xi^-, \xi^+)) \leqslant K \text{ and, for all } |k| \geqslant K,\ d(\pi(X_i), \pi(X_{i+k})) \geqslant L. \tag{4.8}$$

Let n be a large integer. Write $m = \lfloor n/K \rfloor$. Among the integers $K, 2K, \dots, mK \leqslant n$, we consider the set $I_n(\omega)$ of those i such that X_i satisfies (4.8). We have $\mathbb{E}(|I_n|) \geqslant m \cdot 9/10$. As $|I_n| \leqslant m$, we get

$$\frac{9m}{10} \leqslant \mathbb{E}(|I_n|) \leqslant \frac{m}{10}\mathbb{P}(|I_n| < m/10) + m\mathbb{P}(|I_n| \geqslant m/10) = \frac{m}{10} + \frac{9m}{10}\mathbb{P}(|I_n| \geqslant m/10).$$

This gives $\mathbb{P}(|I_n| \geqslant m/10) \geqslant 8/9$. Let $\eta = 1/(20K)$. Let Ω_n be the set of ω such that $|I_n(\omega)| \geqslant \eta n + 1$, and X_0 and X_n satisfy (4.8), and $d(X_n, e) \leqslant 2\ell n$ (where ℓ is the drift of μ). It satisfies $\mathbb{P}(\Omega_n) \geqslant 1/2$ if n is large enough. This is the set of good trajectories for which we can control the position of many of the X_i.

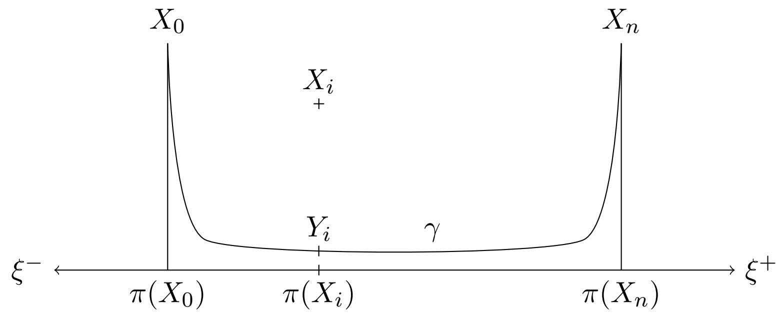

FIGURE 1. The projections on γ and S

Let $\omega \in \Omega_n$. We write Y_i for a projection of X_i on a geodesic γ from e to X_n. Let $\tilde{I}_n = I_n \setminus \{mK\}$, so that the elements of $\tilde{I}_n$ are at distance at least K of 0 and n. As X_0 and X_n satisfy (4.8), the projections $\pi(X_i)$ for $i \in \tilde{I}_n$ are located between $\pi(X_0)$ and $\pi(X_n)$, and are at a distance at least L of these points (see Figure 1). If L is large enough, we obtain $d(\pi(X_i), Y_i) \leqslant C_1$ by hyperbolicity, where C_1 only depends on Γ. This gives

$$d(Y_i, \Lambda) \leqslant d(Y_i, \pi(X_i)) + d(\pi(X_i), X_i) \leqslant C_1 + K,$$

thanks to (4.8) for X_i. When $i \neq j$ belong to $\tilde{I}_n$, we have $d(\pi(X_i), \pi(X_j)) \geqslant L$ again thanks to (4.8), hence $d(Y_i, Y_j) \geqslant L - 2C_1$. If L was chosen larger than $2C_1 + 1$, this shows that $Y_i \neq Y_j$. We have found along γ at least $|I_n| - 1$ distinct points, within distance $C_1 + K$ of Λ. Moreover, for large enough n,

$$|I_n| - 1 \geqslant \eta n \geqslant 2\ell n \cdot (\eta/2\ell) \geqslant d(e, X_n) \cdot (\eta/2\ell).$$

Let $\varepsilon = \eta/2\ell$ and $M = C_1 + K$. We have shown that, for $\omega \in \Omega_n$ (whose probability is at least 1/2), the point $X_n(\omega)$ belongs to $\Lambda_{QC(\varepsilon,M)}$. □

5. CONSTRUCTION OF MAXIMIZING MEASURES

In this section, we prove Theorem 1.7: Given any finite subset Σ in a hyperbolic group Γ, there exists a measure μ_Σ maximizing the quantity $h(\mu)/\ell(\mu)$ over all measures μ supported on Σ with $\ell(\mu) > 0$. To prove this result, we start with a sequence of measures μ_i supported on Σ such that $h(\mu_i)/\ell(\mu_i)$ converges to the maximum M of these quantities. We are looking for μ_Σ with $h(\mu_\Sigma)/\ell(\mu_\Sigma) = M$. Replacing μ_i with $(\mu_i + \delta_e)/2$ (this multiplies entropy and drift by 1/2, and does not change their ratio) and adding e to Σ, we can always assume $\mu_i(e) \geqslant 1/2$, to avoid periodicity problems.

Extracting a subsequence, we can ensure that μ_i converges to a limit probability measure μ. We treat separately the two following cases:

(1) Γ_μ is non-elementary.
(2) Γ_μ is elementary.

Let us handle first the easy case, where Γ_μ is non-elementary. In this case, the entropy and the drift are continuous at μ, by Proposition 2.3 and Theorem 2.9, both due to Erschler and Kaimanovich in [EK13]. Therefore, $h(\mu_i)/\ell(\mu_i)$ tend to $h(\mu)/\ell(\mu)$, since in this case $\ell(\mu) > 0$. One can thus take $\mu_\Sigma = \mu$.

The case where Γ_μ is elementary is much more interesting. Let us describe heuristically what should happen, in a simple case. We assume that $\mu_i = (1-\varepsilon)\mu + \varepsilon\nu$ where ν is a fixed measure, and ε tends to 0. The random walk for μ_i can be described as follows. At each jump, one picks μ (with probability $1-\varepsilon$) or ν (with probability ε), then one jumps according to the chosen measure. After time N, the measure ν is chosen roughly εN times, with intervals of length $1/\varepsilon$ in between, where μ is chosen. Thus, μ_i^{*N} behaves roughly like $(\mu^{*1/\varepsilon} * \nu)^{\varepsilon N}$.

When Γ_μ is finite, the measure $\mu^{*1/\varepsilon}$ is close, when ε is small, to the uniform measure π on Γ_μ. Therefore, μ_i^{*N} is close to $(\pi * \nu)^{\varepsilon N}$. We deduce $h(\mu_i) \sim \varepsilon h(\pi * \nu)$ and $\ell(\mu_i) \sim \varepsilon \ell(\pi * \nu)$. In particular, $h(\mu_i)/\ell(\mu_i) \to h(\pi * \nu)/\ell(\pi * \nu)$. One can take $\mu_\Sigma = \pi * \nu$.

When Γ_μ is infinite, it is virtually cyclic. Assuming that μ is centered for simplicity, the walk given by $\mu^{*1/\varepsilon}$ arrives essentially at distance $1/\sqrt{\varepsilon}$ of the origin, by the central limit theorem. Then, one jumps according to ν, in a direction transverse to Γ_μ, preventing further cancellations. Hence, the walk given by $(\mu^{*1/\varepsilon} * \nu)^{\varepsilon N}$ is at distance roughly $\varepsilon N/\sqrt{\varepsilon}$ from the origin, yielding $\ell(\mu_i) \sim \sqrt{\varepsilon}$. On the other hand, each step $\mu^{*1/\varepsilon}$ only visits $1/\varepsilon$ points, hence the measure $(\mu^{*1/\varepsilon} * \nu)^{\varepsilon N}$ is supported by roughly $(1/\varepsilon)^{\varepsilon N}$ points, yielding $h(\mu_i) \sim \varepsilon|\log \varepsilon|$. In particular, $h(\mu_i) = o(\ell(\mu_i))$. This implies that $h(\mu_i)/\ell(\mu_i)$, which tends to 0, can not tend to the maximum M. Therefore, this case can not happen.

The rigorous argument is considerably more delicate. One difficulty is that μ_i does not decompose in general as $(1-\varepsilon)\mu+\varepsilon\nu$: there can be in μ_i points with a very small probability (which are not seen by μ), but much larger than ε, the probability to visit a nonelementary subset of Γ. These points will play an important role on the relevant time scale, i.e., $1/\varepsilon$. Hence, we have to describe the different time scales that happen in μ_i.

For each $a \in \Sigma$, we have a weight $\mu_i(a)$, which tends to 0 if a is not in the support of μ. Reordering the a_k and extracting a subsequence, we can assume that $\Sigma = \{a_1, \dots, a_p\}$ with $\mu_i(a_1) \geqslant \dots \geqslant \mu_i(a_p)$ (and $a_1 = e$). Extracting a further subsequence, we may also assume that $\mu_i(a_k)/\mu_i(a_{k-1})$ converges for all k, towards a limit in $[0,1]$.

Let Γ_k be the subgroup generated by $a_1, \dots, a_k$. We consider the smallest r such that Γ_r is non-elementary. Then, we consider the biggest $s < r$ such that $\mu_i(r) = o(\mu_i(s))$. Roughly speaking, the random walk has enough time to spread on the elementary subgroup Γ_s, before seeing a_r. It turns out that the asymptotic behavior will depend on the nature of Γ_s (finite or virtually cyclic infinite).

We will decompose the measure μ_i as the sum of two components $(1-\varepsilon_i)\alpha_i + \varepsilon_i\beta_i$, where ε_i tends to 0, the measure α_i mainly lives on Γ_s, and the measure β_i corresponds to the remaining part of μ_i, on $\{a_{s+1}, \dots, a_p\}$. The precise construction depends on the nature of Γ_s:

- *If Γ_s is finite.* Let $\beta_i^{(0)}$ be the normalized restriction of μ_i to $\{a_{s+1}, \dots, a_p\}$. To avoid periodicity problems, we rather consider $\beta_i = (\delta_e + \beta_i^{(0)})/2$. We decompose $\mu_i = (1-\varepsilon_i)\alpha_i + \varepsilon_i\beta_i$, where α_i is supported on $a_1, \dots, a_s$. By construction, the probability of any element in the support of α_i is much bigger than ε_i.

- *If Γ_s is virtually cyclic infinite.* The group Γ_s contains a hyperbolic element g_0, with repelling and attracting points at infinity denoted by g_0^- and g_0^+. The elements of Γ_s all fix the set $\{g_0^-, g_0^+\}$. We take for α_i the normalized restriction of μ_i to those elements in Σ that fix $\{g_0^-, g_0^+\}$, and for β_i the normalized restriction of μ_i to the other elements. Once again, we can write $\mu_i = (1-\varepsilon_i)\alpha_i + \varepsilon_i\beta_i$.

In both cases, ε_i is comparable to the probability $\mu_i(a_r)$, and is therefore negligible with respect to $\mu_i(a_s)$. We will write $\mu_i = \mu_\varepsilon$ (and, in the same way, we will replace all indices i with ε, since the main parameter is $\varepsilon = \varepsilon_i$). The measure μ_ε converges to μ when ε tends to 0, while β_ε tends to a probability measure β, supported on $e, a_{s+1}, \dots, a_p$. If the measures μ_ε are symmetric to begin with, the measures α_ε and β_ε are also symmetric by construction.

To generate the random walk given by μ_ε, one can first independently choose random measures ρ_n: one takes $\rho_n = \alpha_\varepsilon$ with probability $1-\varepsilon$, and $\rho_n = \beta_\varepsilon$ with probability ε. Then, one chooses elements g_n randomly according to ρ_n, and one multiplies them: the product $g_1 \cdots g_n$ is distributed like the random walk given by μ_ε at time n.

We will group together successive g_k, into blocks where the equidistribution on Γ_s can be seen. More precisely, denote by $t_1, t_2, \dots$ the successive times where $\rho_n = \beta_\varepsilon$ (and $t_0 = 0$). They are stopping times, the successive differences are independent and identically distributed, with a geometric distribution of parameter ε (i.e., $\mathbb{P}(t_1 = n) = (1-\varepsilon)^{n-1}\varepsilon$), with mean $1/\varepsilon$. Write $L_N = g_{t_{N-1}+1} \cdots g_{t_N}$. By construction, the L_i are independent, identically distributed, and the random walk they define, i.e., $L_1 \cdots L_N$, is a subsequence of the original random walk $g_1 \cdots g_n$. Let λ_ε be the distribution of L_i on Γ, i.e.,

$$\lambda_\varepsilon = \sum_{n=0}^{\infty} (1-\varepsilon)^n \varepsilon \alpha_\varepsilon^{*n} * \beta_\varepsilon.$$

Lemma 5.1. *The measure λ_ε has finite first moment and finite time one entropy. Moreover, $\ell(\mu_\varepsilon) = \varepsilon\ell(\lambda_\varepsilon)$ and $h(\mu_\varepsilon) = \varepsilon h(\lambda_\varepsilon)$.*

Proof. As the mean of t_1 is $1/\varepsilon$, the random walk generated by λ_ε is essentially the random walk generated by μ_ε, but on a time scale $1/\varepsilon$. This justifies heuristically the statement.

For the rigorous proof, let us first check that λ_ε has finite first moment (and hence finite time one entropy). Since all the measures have finite support, we have $|L_1| \leqslant Ct_1$. Since a geometric distribution has moments of all order, the same is true for $|L_1|$.

The strong law of large numbers ensures that, almost surely, $t_N \sim N/\varepsilon$. Therefore, almost surely,

$$\ell(\lambda_\varepsilon) = \lim \frac{|L_1 \cdots L_N|}{N} = \lim \frac{|g_1 \cdots g_{t_N}|}{N} = \lim \frac{|g_1 \cdots g_{t_N}|}{t_N} \cdot \frac{t_N}{N} = \ell(\mu_\varepsilon) \cdot 1/\varepsilon.$$

This proves the statement of the lemma for the drift.

For the entropy, we use the characterization of Lemma 2.4. We will show that $h(\mu_\varepsilon) \leqslant \varepsilon h(\lambda_\varepsilon)$ and $h(\mu_\varepsilon) \geqslant \varepsilon h(\lambda_\varepsilon)$. Let K_n be a set of cardinality at most $e^{(h(\mu_\varepsilon)+\eta)n}$ which contains $g_1 \cdots g_n$ with probability at least $1/2$. Let $N = \varepsilon n$. With large probability, t_N is close to n, up to $\eta' n$ (where η' is arbitrarily small). Hence, with probability at least $1/3$, the point $L_1 \cdots L_N$ belongs to the $C\eta' n$-neighborhood of K_n, whose cardinality is at most

$$|K_n| \cdot e^{C'\eta' n} \leqslant e^{(h(\mu_\varepsilon)+\eta+C'\eta')n} = e^{(h(\mu_\varepsilon)+\eta+C'\eta')N/\varepsilon}.$$

As η and η' are arbitrary, this shows that $h(\lambda_\varepsilon) \leqslant h(\mu_\varepsilon)/\varepsilon$. The converse inequality is proved in the same way. □

The previous lemma shows that we should understand λ_ε. We define an auxiliary probability measure $\tilde{\alpha}_\varepsilon$ so that $\lambda_\varepsilon = \tilde{\alpha}_\varepsilon * \beta_\varepsilon$, by

$$\tilde{\alpha}_\varepsilon = \sum_{n=0}^{\infty} (1-\varepsilon)^n \varepsilon \alpha_\varepsilon^{*n}. \tag{5.1}$$

In this formula, most weight is concentrated around those n of the order of $1/\varepsilon$. Hence, we have to understand the iterates of α_ε in time $1/\varepsilon$. When Γ_s is finite, we will see that it has enough time to equidistribute on Γ_s (even though α_ε may give a very small weight to some elements, this weight is by construction much larger than ε, so that $1/\varepsilon$ iterates are enough to equidistribute). When Γ_s is virtually cyclic, we will see that the random walk has enough time to drift away significantly from the identity.

In both cases, we will need quantitative results on basic groups, but in weakly elliptic cases (i.e., the transition probabilities are not bounded from below). There are techniques to get quantitative estimates in such settings, especially comparison techniques (due for instance to Varopoulos, Diaconis, Saloff-Coste): one can compare weakly elliptic walks to elliptic ones (which we understand well) thanks to Dirichlet forms arguments: these arguments make it possible to transfer results from the latter to the former (modulo some loss in the constants, due to the lack of ellipticity). We will rely on such results when Γ_s is infinite. When it is finite, such techniques can also be used, but we will rather give a more elementary argument.

We start with the case where Γ_s is finite. We need to quantify the speed of convergence to the stationary measure in finite groups, with the following lemma.

Lemma 5.2. *Let Λ be a finite group. Let $\Sigma_\Lambda \subset \Lambda$ be a generating subset (it does not have to be symmetric). Let π_Λ be the uniform measure on Λ, and let $d(\mu, \pi_\Lambda)$ be the euclidean distance between a measure μ and π_Λ (i.e., $\left(\sum(\mu(g) - \pi_\Lambda(g))^2\right)^{1/2}$). For any $\delta > 0$, there exists $K > 0$ with the following property. Let $\eta > 0$. Consider a probability measure μ on Λ with $\mu(\sigma) \geqslant \eta$ for any $\sigma \in \Sigma_\Lambda \cup \{e\}$. Then, for all $n > K/\eta$,*

$$d(\mu^{*n}, \pi_\Lambda) \leqslant \delta.$$

In other words, the time to see the equidistribution towards the stationary measure is bounded by $1/\eta$, where η is the minimum of the transition probabilities on Σ_Λ.

Proof. Endow the space $\mathcal{M}(\Lambda)$ of signed measures on Λ with the scalar product corresponding to the quadratic form $|\nu|^2 = \sum \nu(g)^2$. Denote by $H = \{\nu : \sum \nu(g) = 0\}$ the hyperplane $\pi_\Lambda^\perp$ of zero mass measures. For any probability ρ, denote by M_ρ the left-convolution operator on $\mathcal{M}(\Lambda)$, that is $M_\rho(\nu) = \rho * \nu$. Since convolution preserves mass, H is M_ρ-invariant. Let us prove that the operator norm of M_ρ is bounded by 1. Indeed, put $u_\rho(g) = \sum_{h \in \Lambda} \rho(h)\rho(hg)$, this is a probability on Λ. We have

$$|M_\rho \nu|^2 = \sum_{g\in\Lambda} (M_\rho \nu(g))^2 = \sum_{(g,h_1,h_2)\in\Lambda^3} \rho(gh_1^{-1})\rho(gh_2^{-1})\nu(h_1)\nu(h_2)$$
$$= \sum_{(h_1,h_2)\in\Lambda^2} \nu(h_1)\nu(h_2)u_\rho(h_1h_2^{-1}) = \sum_{(g,h)\in\Lambda^2} \nu(h)\nu(g^{-1}h)u_\rho(g)$$
$$\leqslant \sum_{g\in\Lambda} |\nu|^2 u_\rho(g) = |\nu|^2.$$

This proves that $\|M_\rho\| \leqslant 1$. Now fix ρ_o to be the uniform probability on the set $\Sigma_\Lambda \cup \{e\}$. Notice that $u_{\rho_o}(g) > 0$ for any $g \in \Sigma_\Lambda \cup \{e\}$, since $\rho_o(e) > 0$. We claim that M_{ρ_o} restricted to H has an operator norm $c < 1$. Would it be not the case, there would exist $\nu \in H - \{0\}$ such that the previous inequalities would be equalities. Thanks to the equality case in the Cauchy-Schwarz inequality, this implies that, for any $g \in \Sigma_\Lambda$, the two measures $h \mapsto \nu(h)$ and $h \mapsto \nu(g^{-1}h)$ are positively proportional. Since their norm are equal, they must be equal. Since Σ_Λ generates Λ, ν is Λ-invariant and belongs to H, so it must be zero.

By assumption, the probability μ can be decomposed as

$$\mu = \eta\rho_o + (1-\eta)\nu,$$

where ν is some probability. This implies that M_μ restricted to H has operator norm at most $\eta c + (1-\eta)$. Therefore,

$$d(\mu^{*n}, \pi_\Lambda) = |\mu^{*n} - \pi_\Lambda| = |M_\mu^n(\delta_e - \pi_\Lambda)| \leqslant 2(1-(1-c)\eta)^n.$$

This inequality implies the result. □

We can now describe the asymptotic behavior of μ_ε when the group Γ_s is finite.

Lemma 5.3. *Assume that Γ_s is finite. Define a new probability measure $\lambda = \pi_{\Gamma_s} * \beta$ (it generates a non-elementary subgroup). When ε tends to 0, we have $h(\mu_\varepsilon) \sim \varepsilon h(\lambda)$ and $\ell(\mu_\varepsilon) \sim \varepsilon\ell(\lambda)$.*

Proof. The random variable t_1, being geometric of parameter ε, is of the order of $1/\varepsilon$ with high probability (i.e., for any $\delta > 0$, there exists $u > 0$ such that $\mathbb{P}(t_1 > u/\varepsilon) \geqslant 1-\delta$). Writing $\Sigma_s = \{a_1, \dots, a_s\}$ for the support of α_ε, we have $\min_{\sigma\in\Sigma_s} \alpha_\varepsilon(\sigma) = (1-\varepsilon)^{-1}\mu_\varepsilon(a_s)$, which is much bigger than ε by definition of s. Lemma 5.2 shows that the measures α_ε^{*n} are close to π_{Γ_s} for $n \geqslant u/\varepsilon$. This implies that $\tilde{\alpha}_\varepsilon$ (defined in (5.1)) converges to π_{Γ_s} when $\varepsilon \to 0$. As β_ε converges to β, this shows that λ_ε converges to λ.

The support of the measure λ contains Γ_s and $a_{s+1}, \dots, a_r$ (as the support of β contains $\{e, a_{s+1}, \dots, a_r\}$ by construction). Hence, Γ_λ contains the non-elementary subgroup Γ_r. It follows that the entropy and the drift are continuous at λ, by Proposition 2.3 and Theorem 2.9. We get $h(\lambda_\varepsilon) \to h(\lambda)$ and $\ell(\lambda_\varepsilon) \to \ell(\lambda)$. With Lemma 5.1, this completes the proof. □

We deduce from the lemma that $h(\mu_\varepsilon)/\ell(\mu_\varepsilon)$ tends to $h(\lambda)/\ell(\lambda)$. Hence, the measure $\mu_\Sigma = \lambda$ satisfies the conclusion of the theorem, at least in the non-symmetric case. In the symmetric case, where we are looking for a symmetric measure μ_Σ, the measure $\lambda = \pi_{\Gamma_s} * \beta$ is not an answer to the problem. However, $\lambda' = \pi_{\Gamma_s} * \beta * \pi_{\Gamma_s}$ is symmetric, and it clearly has the same entropy and drift as λ (since $\pi_{\Gamma_s} * \pi_{\Gamma_s} = \pi_{\Gamma_s}$). Hence, we can take $\mu_\Sigma = \lambda'$. This completes the proof of Theorem 1.7 when the group Γ_s is finite.

Example 5.4. Let $\Gamma = \mathbb{Z}/2 * \mathbb{Z}/4$, with $\Sigma = \{a, b, b^{-1}\}$ (where a is the generator of $\mathbb{Z}/2$ and b the generator of $\mathbb{Z}/4$), with the word distance coming from Σ. [MM07, Section 5.1] shows that the supremum over measures supported on Σ of $h(\mu)/\ell(\mu)$ is the growth v of the group (note that Γ is virtually free), and that it is not realized by a measure supported on Σ. This shows that, in Theorem 1.7, the fact that μ_Σ may need a support larger than Σ is not an artefact of the proof.

In this example, any symmetric measure on Σ is of the form $\mu_\varepsilon = (1-\varepsilon)\delta_a + \varepsilon\beta$ where β is uniform on $\{b, b^{-1}\}$. The above proof shows that, when ε tends to 0, $h(\mu_\varepsilon)/\ell(\mu_\varepsilon)$ converges to $h(\lambda)/\ell(\lambda)$ where $\lambda = \pi_{\Gamma_s} * \beta = \frac{1}{2}(\delta_e + \delta_a) * \frac{1}{2}(\delta_b + \delta_{b^{-1}})$ is the uniform measure on $\{b, b^{-1}, ab, ab^{-1}\}$.

It remains to treat the case where Γ_s is virtually cyclic infinite. Such a group surjects onto $\mathbb{Z}$ or $\mathbb{Z} \rtimes \mathbb{Z}/2$ (the infinite dihedral group), with finite kernel. From the point of view of the random walk, most things happen in the quotient. Hence, it would suffice to understand these two groups (separating in the case of $\mathbb{Z}$ the centered and non-centered cases). We will rather give direct arguments which do not use this reduction and which avoid separating cases. Let $t \leqslant s$ be the smallest index such that $\{a_1, \cdots, a_t\}$ generates an infinite group. Let $\eta = \eta(\varepsilon) = \mu_\varepsilon(a_t)$, this parameter governs the equidistribution speed on Γ_s (or, at least, on Γ_t, which has finite index in Γ_s since these two groups are virtually cyclic infinite). We will find the asymptotics of the entropy and the drift in terms of η/ε (which tends to infinity by definition of s). We start with the entropy (for which an upper bound suffices). Note that the random walk directed by α_ε does not live on Γ_s, but on a possibly bigger group since we have put in α_ε all the points that fix the set $\{g_0^-, g_0^+\}$ (this will be important in the control of the drift below). Let $\tilde{\Gamma}_s$ be the group they generate, it is still virtually cyclic (see, for instance, [GdlH90, Théorème 37 page 157]), and it contains Γ_s as a finite index subgroup.

Lemma 5.5. *There exists a constant C such that $h(\lambda_\varepsilon) \leqslant C \log(\eta/\varepsilon)$.*

Proof. Let K be the group generated by $\{a_1, \dots, a_{t-1}\}$. It is finite by definition of t. Let Σ' be the set of points among $a_t, \dots, a_p$ which stabilize $\{g_0^-, g_0^+\}$. The group $\tilde{\Gamma}_s$ is generated by K and Σ'. Let us consider the associated word pseudo-distance d', where we decide that elements in K have 0 length. This pseudo-distance is quasi-isometric to the usual distance, and it satisfies $d'(e, xk) = d'(e, x)$ for all $x \in \tilde{\Gamma}_s$ and all $k \in K$.

Let us first estimate the average distance to the origin for an element given by $\tilde{\alpha}_\varepsilon$. We decompose α_ε as the average of a measure supported on $\{a_1, \dots, a_{t-1}\} \subset K$, and of a measure supported on Σ' (the contribution of the latter has a mass $m(\varepsilon)$ bounded by $(p - t + 1)\eta \leqslant C\eta$). The measure α_ε^{*n} can be obtained by picking at each step one of these two measures (according to their respective weight), and then jumping according to a random element for this measure. When we use the first measure, the d'-distance to the origin does not change by definition. Hence, the distance to the origin is bounded by the number of choices of the second measure. We obtain

$$\begin{aligned}\mathbb{E}_{\tilde{\alpha}_\varepsilon}(d'(e,g)) &\leqslant \sum_{n=0}^{\infty}(1-\varepsilon)^n\varepsilon\sum_{i=0}^{n}\binom{n}{i}m(\varepsilon)^i(1-m(\varepsilon))^{n-i}\cdot Ci \\ &= Cm(\varepsilon)\sum_{n=0}^{\infty}(1-\varepsilon)^n\varepsilon\sum_{i=1}^{n}n\binom{n-1}{i-1}m(\varepsilon)^{i-1}(1-m(\varepsilon))^{n-i} \\ &= Cm(\varepsilon)\sum_{n=0}^{\infty}(1-\varepsilon)^n\varepsilon n = Cm(\varepsilon)(1-\varepsilon)/\varepsilon \leqslant C\eta/\varepsilon.\end{aligned}$$

A measure supported on the integers with first moment A has entropy bounded by $C\log A + C$ (see, for instance, [EK10, Lemma 2]). The proof also applies to virtually cyclic situations (the finite thickening does not change anything). Therefore, we get $H(\tilde{\alpha}_\varepsilon) \leqslant C\log(\eta/\varepsilon)+C$.

Finally,

$$H(\lambda_\varepsilon) = H(\tilde{\alpha}_\varepsilon * \beta_\varepsilon) \leqslant H(\tilde{\alpha}_\varepsilon) + H(\beta_\varepsilon) \leqslant C\log(\eta/\varepsilon)+C,$$

since the support of β_ε is uniformly bounded. As $\eta/\varepsilon \to \infty$, this gives $H(\lambda_\varepsilon) \leqslant C\log(\eta/\varepsilon)$. Finally, we estimate $h(\lambda_\varepsilon) = \inf_{n>0} H(\lambda_\varepsilon^{*n})/n \leqslant H(\lambda_\varepsilon)$ to get the conclusion of the lemma. □

For the drift, we need to be more precise since we need a lower bound to conclude. We will use a lemma giving lower bounds on the equidistribution speed in virtually cyclic infinite groups, using comparison techniques.

Lemma 5.6. *Let Λ be a virtually cyclic infinite group. Let $\Sigma_\Lambda \subset \Lambda$ be a finite subset generating an infinite subgroup of Λ. There exists a constant C with the following property. Let $\eta > 0$. Let μ be a probability measure on Λ with $\mu(e) \geqslant 1/2$ and $\mu(\sigma) \geqslant \eta$ for any $\sigma \in \Sigma_\Lambda$. Then, for all $n \geqslant 1$,*

$$\sup_{g\in\Lambda}\mu^{*n}(g) \leqslant C(\eta n)^{-1/2}.$$

The interest of the lemma is that C does not depend on the measure μ, and that we obtain an explicit control on μ^{*n} just in terms of a lower bound on the transition probabilities of μ.

Proof. We use the comparison method. Let ρ be the uniform measure on e, Σ_Λ and Σ_Λ^{-1}. The random walk it generates does not have to be transitive (since Σ_Λ does not necessarily generate the whole group Λ), but Λ is partitioned into finitely many classes where it is transitive (and isomorphic to the random walk on the group generated by Σ_Λ). Moreover, it is symmetric, and therefore reversible for the counting measure m on Λ. The Dirichlet form associated to ρ is by definition

$$\mathcal{E}_\rho(f,f) = \frac{1}{2}\sum_{x,y}|f(x)-f(y)|^2\rho(x^{-1}y),$$

for any $f : \Lambda \to \mathbb{C}$. As Λ has linear growth, the following Nash inequality holds (see, for instance, [Woe00, Proposition 14.1]).

$$\|f\|_{L^2}^6 \leqslant C\|f\|_{L^1}^4\mathcal{E}_\rho(f,f),$$

where all norms are defined with respect to the measure m on Λ. Let P_μ be the Markov operator associated to μ. It satisfies

$$\|f\|_{L^2}^2 - \|P_\mu f\|_{L^2}^2 = \langle f, f\rangle - \langle P_\mu f, P_\mu f\rangle = \langle (I - P_\mu^* P_\mu) f, f\rangle.$$

The operator $P_\mu^* P_\mu$ is the Markov operator associated to the symmetric probability measure $\nu = \check{\mu} * \mu$, which satisfies $\nu(\sigma) \geqslant \eta/2$ for $\sigma \in \Sigma_\Lambda \cup \Sigma_\Lambda^{-1}$ and $\nu(e) \geqslant 1/4$ (since $\mu(e) \geqslant 1/2$). Therefore, $\rho(g) \leqslant C\eta^{-1}\nu(g)$ for all g. We deduce

$$\begin{aligned}\|f\|_{L^2}^2 - \|P_\mu f\|_{L^2}^2 &= \sum \overline{f(x)}(f(x) - f(y))\nu(x^{-1}y) = \frac{1}{2}\sum |f(x) - f(y)|^2 \nu(x^{-1}y) \\ &\geqslant \frac{\eta}{2C}\sum |f(x) - f(y)|^2 \rho(x^{-1}y) = \frac{\eta}{C}\mathcal{E}_\rho(f, f).\end{aligned}$$

Combining this inequality with Nash inequality, we obtain

$$\|f\|_{L^2}^6 \leqslant C\eta^{-1}\|f\|_{L^1}^4 (\|f\|_{L^2}^2 - \|P_\mu f\|_{L^2}^2).$$

The operator P_μ^* satisfies the same inequality, for the same reason. Composing these inequalities, we obtain an estimate for the norm of P_μ^n from L^1 to L^∞ (this is [VSCC92, Lemma VII.2.6]), of the form

$$\|P_\mu^n\|_{L^1 \to L^\infty} \leqslant (C'\eta^{-1}/n)^{1/2}.$$

Applying this inequality to the function δ_e, we get the desired result. □

The previous lemma implies that, if C' is large enough, a neighborhood of size $(\eta n)^{1/2}/C'$ of the identity has probability for μ^{*n} at most $1/2$. Hence, the average distance to the origin is at least of the order of $(\eta n)^{1/2}$.

Now, we study the stationary measure for $\beta_\varepsilon * \tilde{\alpha}_\varepsilon$ on $\partial\Gamma$. We recall that g_0 is a hyperbolic element in Γ_s, fixed once and for all.

Lemma 5.7. *There exists a neighborhood U of $\{g_0^-, g_0^+\}$ in $\partial\Gamma$ such that the stationary measure ν_ε of $\beta_\varepsilon * \tilde{\alpha}_\varepsilon$ satisfies $\nu_\varepsilon(U) \to 0$.*

Proof. Let us first show that, for any neighborhood U of $\{g_0^-, g_0^+\}$, then $(\tilde{\alpha}_\varepsilon * \delta_z)(U^c)$ tends to 0, uniformly in $z \in \partial\Gamma$. This is not surprising since a typical element for $\tilde{\alpha}_\varepsilon$ is large in the virtually cyclic group $\tilde{\Gamma}_s$, and sends most points into U. To make this argument rigorous, we will use Lemma 5.6. The definition (5.1) shows that it suffices to prove that $(\alpha_\varepsilon^{*n} * \delta_z)(U^c)$ is small for $n \geqslant u/\varepsilon$.

The subgroup generated by g_0 has finite index in $\tilde{\Gamma}_s$. Hence, any element in $\tilde{\Gamma}_s$ can be written as $g_0^k \gamma_i$, for γ_i in a finite set. Thus, the measure α_ε^{*n} can be written as $\sum c_n(k,i)\delta_{g_0^k\gamma_i}$, for some coefficients $c_n(k,i)$. Lemma 5.6 (applied to $\Lambda = \tilde{\Gamma}_s$ with $\Sigma_\Lambda = \{a_1, \dots, a_t\}$) ensures that $\sup_{k,i} c_n(k,i) \leqslant C/(\eta n)^{1/2}$. When $n \geqslant u/\varepsilon$, this quantity tends to 0 since $\varepsilon = o(\eta)$. We have

$$(\alpha_\varepsilon^{*n} * \delta_z)(U^c) = \sum_{k,i} c_n(k,i) 1(g_0^k \gamma_i z \notin U).$$

As the element g_0 is hyperbolic, there exists C such that, for any $w \in \partial\Gamma$,

$$|\{k \in \mathbb{Z} \,:\, g_0^k w \notin U\}| \leqslant C.$$

The uniformity in w follows from the compactness of $(\partial\Gamma \setminus \{g_0^-, g_0^+\})/\langle g_0\rangle$. We obtain

$$(\alpha_\varepsilon^{*n} * \delta_z)(U^c) \leqslant \big(\sup_{k,i} c_n(k,i)\big) \sum_i |\{k \in \mathbb{Z} \,:\, g_0^k \gamma_i z \notin U\}| \leqslant C \sup_{k,i} c_n(k,i) \leqslant C/(\eta n)^{1/2}.$$

This shows that $(\alpha_\varepsilon^{*n} * \delta_z)(U^c)$ is small, as desired.

As $\tilde\alpha_\varepsilon * \delta_z(U^c)$ tends to 0 uniformly in z, we deduce that $(\tilde\alpha_\varepsilon * \nu_\varepsilon)(U^c)$ also tends to 0, and therefore that $(\tilde\alpha_\varepsilon * \nu_\varepsilon)(U)$ tends to 1.

Let $A = \{g_0^-, g_0^+\}$. We claim that, for all g such that $gA \cap A \neq \emptyset$, then $gA = A$. Indeed, if $g(g_0^-) \in A$ for instance, then $g^{-1} g_0 g$ is a hyperbolic element stabilizing g_0^-. It also stabilizes g_0^+, by [GdlH90, Théorème 30 page 154], i.e., $g_0 g(g_0^+) = g(g_0^+)$. Hence, $g(g_0^+)$ is a fixed point of g_0, i.e., $g(g_0^+) \in A$.

By definition of β_ε, the finitely many elements of its support do not fix A. They even satisfy $gA \cap A = \emptyset$ for all g in this support, by the previous argument. If U is small enough, we get $gU \cap U = \emptyset$, i.e., $g(U) \subset U^c$.

Finally,

$$\nu_\varepsilon(U^c) = (\beta_\varepsilon * \tilde\alpha_\varepsilon * \nu_\varepsilon)(U^c) \geqslant (\tilde\alpha_\varepsilon * \nu_\varepsilon)(U),$$

which tends to 1 when ε tends to 0. □

Lemma 5.8. *The drift $\ell(\lambda_\varepsilon)$ satisfies $\ell(\lambda_\varepsilon) \geqslant c \cdot (\eta/\varepsilon)^{1/2}$.*

Proof. Let ρ_ε be a stationary measure for λ_ε, on the Busemann boundary $\partial_B\Gamma$. By Proposition 2.2,

$$\ell(\lambda_\varepsilon) = \int c_B(g,\xi)\,\mathrm{d}\rho_\varepsilon(\xi)\,\mathrm{d}\lambda_\varepsilon(g),$$

where $c_B(g,\xi) = h_\xi(g^{-1})$ is the Busemann cocycle. As $\lambda_\varepsilon = \tilde\alpha_\varepsilon * \beta_\varepsilon$, this gives

$$\ell(\lambda_\varepsilon) = \int c_B(Lb,\xi)\,\mathrm{d}\rho_\varepsilon(\xi)\,\mathrm{d}\tilde\alpha_\varepsilon(L)\,\mathrm{d}\beta_\varepsilon(b).$$

With the cocycle relation (2.2), this becomes

$$\ell(\lambda_\varepsilon) = \int c_B(L, b\xi)\,\mathrm{d}\rho_\varepsilon(\xi)\,\mathrm{d}\tilde\alpha_\varepsilon(L)\,\mathrm{d}\beta_\varepsilon(b) + \int c_B(b,\xi)\,\mathrm{d}\rho_\varepsilon(\xi)\,\mathrm{d}\tilde\alpha_\varepsilon(L)\,\mathrm{d}\beta_\varepsilon(b).$$

The second integral is bounded independently of ε since the support of β_ε is finite. In the first integral, $\xi' = b\xi$ is distributed according to the measure $\tilde\rho_\varepsilon := \beta_\varepsilon * \rho_\varepsilon$, which is stationary for $\beta_\varepsilon * \tilde\alpha_\varepsilon$. Lemma 5.7 implies that its projection $(\pi_B)_*\tilde\rho_\varepsilon$ on the geometric boundary, which is again stationary for $\beta_\varepsilon * \tilde\alpha_\varepsilon$, gives a small measure to a neighborhood U of $\{g_0^-, g_0^+\}$.

As the limit set of $\tilde\Gamma_s$ is $\{g_0^-, g_0^+\}$, there exists a constant C such that, for all $\xi \notin \pi_B^{-1}U$ and $g \in \tilde\Gamma_s$, we have $|h_\xi(g^{-1}) - d(e,g)| \leqslant C$. For $\xi \in \pi_B^{-1}U$, we only use the trivial bound $h_\xi(g^{-1}) \geqslant -d(e,g)$, since horofunctions are 1-Lipschitz and vanish at the origin. We get

$$\begin{aligned}\ell(\lambda_\varepsilon) &\geqslant \int_{(L,\xi)\in\Gamma\times\pi_B^{-1}U^c} d(e,L)\,\mathrm{d}\tilde\alpha_\varepsilon(L)\,\mathrm{d}\tilde\rho_\varepsilon(\xi) - \int_{(L,\xi)\in\Gamma\times\pi_B^{-1}U} d(e,L)\,\mathrm{d}\tilde\alpha_\varepsilon(L)\,\mathrm{d}\tilde\rho_\varepsilon(\xi) - C \\ &= \left(\int d(e,L)\,\mathrm{d}\tilde\alpha_\varepsilon(L)\right)(\tilde\rho_\varepsilon(\pi_B^{-1}U^c) - \tilde\rho_\varepsilon(\pi_B^{-1}U)) - C.\end{aligned}$$

For small enough ε, we have $\tilde{\rho}_\varepsilon(\pi_B^{-1}U) \leqslant 1/4$ (and therefore $\tilde{\rho}_\varepsilon(\pi_B^{-1}U^c) \geqslant 3/4$). Moreover, Lemma 5.6 ensures that the average distance to the origin for the measure $\tilde{\alpha}_\varepsilon$ is at least $c \cdot (\eta/\varepsilon)^{1/2}$. Hence, the previous formula completes the proof. □

Combining Lemmas 5.5 and 5.8, we get

$$h(\lambda_\varepsilon)/\ell(\lambda_\varepsilon) \leqslant C \log(\eta/\varepsilon)/(\eta/\varepsilon)^{1/2}.$$

This tends to 0 since η/ε tends to infinity. We deduce from Lemma 5.1 that $h(\mu_\varepsilon)/\ell(\mu_\varepsilon)$ tends to 0. This is a contradiction since we were assuming that it converges to the maximum M, which is positive.

This concludes the proof of Theorem 1.7. □

The study of the case where Γ_s is virtually cyclic infinite gives in particular the following result.

Theorem 5.9. *Let (Γ, d) be a metric hyperbolic group. Let Σ be a finite subset of Γ which generates a non-elementary group. Let μ_i be a sequence of measures on Σ, with $h(\mu_i) > 0$, converging to a probability measure μ such that Γ_μ is infinite virtually cyclic. Then $h(\mu_i)/\ell(\mu_i) \to 0$.*

Note that the precise value of $\ell(\mu_i)$ depends on the choice of the distance, but if two distances are equivalent then the associated drifts vary within the same constants. Hence, the convergence $h(\mu_i)/\ell(\mu_i) \to 0$ does not depend on the distance.

We recover results of Le Prince [LP07]: In any metric hyperbolic group, there exist admissible probability measures with $h/\ell < v$. The construction of Le Prince is rather similar to the examples given by Theorem 5.9.

Example 5.10. We can use the above proof to also find an example where $h(\mu_\varepsilon)/\ell(\mu_\varepsilon) \to 0$ although μ_ε tends to a measure μ for which Γ_μ is finite and nontrivial. Consider $\Gamma = \mathbb{Z}/2 \times \mathbb{F}_2 = \{0, 1\} \times \langle a, b\rangle$, endowed with the probability measure μ_ε given by

$$\mu_\varepsilon(0, e) = \mu_\varepsilon(1, e) = 1/2 - \varepsilon - \varepsilon^2, \quad \mu_\varepsilon(0, a) = \mu_\varepsilon(0, a^{-1}) = \varepsilon, \quad \mu_\varepsilon(0, b) = \mu_\varepsilon(0, b^{-1}) = \varepsilon^2.$$

The measure μ_ε converges to $\mu = (\delta_{(0,e)} + \delta_{(1,e)})/2$. With the above notations, $\Gamma_\mu = \mathbb{Z}/2 \times \{e\}$ but $\Gamma_s = \mathbb{Z}/2 \times \langle a\rangle$ is virtually cyclic infinite (so that $h(\mu_\varepsilon)/\ell(\mu_\varepsilon) \to 0$) and $\Gamma_r = \Gamma$.

6. Examples for non-symmetric measures

In this section, we describe the additional difficulties that arise if one tries to prove Theorem 1.3 for non-symmetric measures. The main problem is that the random walk lives on the subsemigroup Γ_μ^+, which is not a subgroup any more. While many cases can be handled with the tools we have described in this article, one case can not be treated in this way: when the subsemigroup Γ_μ^+ has no nice geometric properties (it is not quasi-convex, it is not a subgroup), but $\Gamma_\mu = \Gamma$.

Let us first show that the growth properties of such a subsemigroup can be more complicated than what happens for subgroups. If Λ is a subgroup of Γ, either $|B_n \cap \Lambda| \asymp e^{nv}$, or $|B_n \cap \Lambda| = o(e^{nv})$ (the first case happens if and only if Λ has finite index in Γ, see the discussion at the beginning of Paragraph 4.3). Unfortunately, the behavior of semigroups can be more complicated.

Proposition 6.1. *In $\mathbb{F}_2$, there exists a subsemigroup Λ^+ such that* $\liminf|B_n \cap \Lambda^+|/|B_n| = 0$ *and* $\limsup|B_n \cap \Lambda^+|/|B_n| > 0$.

Proof. Let $\mathbb{S}^n_{a,a}$ denote the geodesic words in $\mathbb{F}_2 = \langle a, b\rangle$ of length n which start and end with a. Let n_j be a sequence tending very quickly to infinity. Let Λ^+ be the subsemigroup generated by $\bigcup \mathbb{S}^{n_j}_{a,a}$. Then $|B_{n_j} \cap \Lambda^+| \geqslant c|B_{n_j}|$. We claim that

$$|B_{n_j-1} \cap \Lambda^+|/|B_{n_j-1}| \to 0.$$

Indeed, the subsemigroup Λ^+_{j-1} generated by $\bigcup_{k<j} \mathbb{S}^{n_k}_{a,a}$ has a growth rate which is $< e^{nv}$, since some subwords such as $b^{n_{j-1}}$ are forbidden in this subsemigroup. Hence, if n_j is large enough with respect to n_{j-1}, we have $|\mathbb{S}^{n_j-1} \cap \Lambda^+| = |\mathbb{S}^{n_j-1} \cap \Lambda^+_{j-1}| = o(e^{(n_j-1)v})$. $\square$

In this example, most points in $\mathbb{S}^{n_j} \cap \Lambda^+$ are introduced by $\mathbb{S}^{n_j}_{a,a}$. This shows that Λ^+ is far from being quasi-convex. In particular, techniques based only on non-quasi-convexity and sub- or super-multiplicativity will never show that $|B_n \cap \Lambda^+| = o(|B_n|)$ for subsemigroups.

Now, we give an example of a well-behaved measure (apart from the fact that it is not symmetric, not admissible and not finitely supported) for which $h = \ell v$. The construction is done in free products. The idea is to forbid simplifications, so that we have an explicit control on the random walk at time n. To enforce this behavior, we will work in a free product $\Gamma_1 * \Gamma_2$, and consider a probability measure supported on elements of the form $g_1 g_2$ with $g_i \in \Gamma_i \setminus \{e\}$. The next statement applies to some non virtually free hyperbolic groups, for instance the free product of two surface groups. It also applies to some non-hyperbolic groups, more precisely to all finitely generated groups without torsion and with infinitely many ends, by Stallings' theorem. It would be of interest to extend it to all groups with infinitely many ends. For this, we would need to also handle amalgamated free products and HNN extensions.

Proposition 6.2. *Let Γ_1 and Γ_2 be two nontrivial groups, generated respectively by finite symmetric sets S_1 and S_2. Let $\Gamma = \Gamma_1 * \Gamma_2$ with the generating set $S = S_1 \cup S_2$ and the corresponding word distance. There exists on Γ a (nonsymmetric, nonadmissible) probability measure μ, with an exponential moment and nonzero entropy, satisfying $h(\mu) = \ell(\mu)v$.*

Proof. For $i = 1, 2$, let $\Gamma_i^* = \Gamma_i \setminus \{e\}$. We claim that

$$\sum_{g_1 \in \Gamma_1^*, g_2 \in \Gamma_2^*} e^{-v|g_1 g_2|} = 1, \tag{6.1}$$

where v is the growth rate of Γ.

Let $F_i(z)$ be the growth series of Γ_i, i.e., $F_i(z) = \sum_{g \in \Gamma_i} z^{|g|}$. The spheres $\mathbb{S}^n_i \in \Gamma_i$ satisfy $\mathbb{S}^{n+m}_i \subset \mathbb{S}^n_i \cdot \mathbb{S}^m_i$. Hence, the sequence $\log|\mathbb{S}^n_i|$ is subadditive. This implies that $\log|\mathbb{S}^n_i|/n$ converges to its infimum v_i, and moreover that $|\mathbb{S}^n_i| \geqslant e^{nv_i}$. We deduce that the radius of convergence of F_i is e^{-v_i}, and moreover $F_i(e^{-v_i}) = +\infty$.

Let $F(z)$ be the growth series of Γ. As in the proof of Proposition 4.8, it is given by

$$F(z) = \frac{F_1(z)F_2(z)}{1 - (F_1(z) - 1)(F_2(z) - 1)}.$$

Assume for instance $v_1 \geqslant v_2$. As $F_1(e^{-v_1}) = +\infty$, the function $(F_1(z)-1)(F_2(z)-1)$ takes the value 1 when z increases to e^{-v_1}, at a point which is precisely the radius of convergence e^{-v} of F. This shows that $(F_1(e^{-v})-1)(F_2(e^{-v})-1) = 1$. This is precisely the equality (6.1).

We define a probability measure μ on Γ as follows: for $(g_1, g_2) \in \Gamma_1^* \times \Gamma_2^*$, let

$$\mu(g_1 g_2) = e^{-v|g_1 g_2|}.$$

Since there is only one way to generate the word $g_1^1 g_2^1 \cdots g_1^n g_2^n$ using μ, we have

$$\mu^{*n}(g_1^1 g_2^1 \cdots g_1^n g_2^n) = e^{-v \sum_i |g_1^i g_2^i|}.$$

Denoting by X_n the position of the random walk at time n, it follows that $-\log \mu^{*n}(X_n) = v|X_n|$. Dividing by n and letting n tend to infinity, this gives $h(\mu) = \ell(\mu) v$. □

If one is interested in measures with finite support, one can only get the following approximation result. It has the same flavor as Theorem 1.4, but it is both stronger since it also applies to some non-hyperbolic groups, and weaker since the measures it produces are not admissible nor symmetric.

Proposition 6.3. *Let Γ_1 and Γ_2 be two nontrivial groups, generated respectively by finite symmetric sets S_1 and S_2. Let $\Gamma = \Gamma_1 * \Gamma_2$ with the generating set $S = S_1 \cup S_2$ and the corresponding word distance. Then*

$$\sup \left\{ h(\mu)/\ell(\mu) \ : \ \mu \text{ finitely supported probability measure in } \Gamma, \ell(\mu) > 0 \right\} = v.$$

Proof. Any element in Γ can be canonically decomposed as a word in elements of Γ_1 and Γ_2. Let $\mathbb{S}_{i,j}^p$ be the set of elements of length p that start with an element in Γ_i and end with an element in Γ_j. We have the decomposition

$$\mathbb{S}^p = \mathbb{S}_{1,1}^p \cup \mathbb{S}_{1,2}^p \cup \mathbb{S}_{2,1}^p \cup \mathbb{S}_{2,2}^p.$$

One term in this decomposition has cardinality at least $|\mathbb{S}^p|/4$. Hence, there exist i, j such that $\limsup \log|\mathbb{S}_{i,j}^p|/p = v$. Multiplying by fixed elements at the beginning and at the end to go from Γ_1 to Γ_i, and from Γ_j to Γ_2, we get

$$\limsup \log|\mathbb{S}_{1,2}^p|/p = v. \tag{6.2}$$

Let μ_p be the uniform probability measure on $\mathbb{S}_{1,2}^p$. By construction, there are no simplifications when one iterates μ_p. Hence, μ_p^{*n} is the uniform probability measure on $(\mathbb{S}_{1,2}^p)^{*n}$, whose cardinality is $|\mathbb{S}_{1,2}^p|^n$. We get $H(\mu_p^{*n}) = n \log|\mathbb{S}_{1,2}^p|$ and $L(\mu_p^{*n}) = np$. Therefore, $h(\mu_p) = \log|\mathbb{S}_{1,2}^p|$ and $\ell(\mu_p) = p$, giving

$$h(\mu_p)/\ell(\mu_p) = \log|\mathbb{S}_{1,2}^p|/p.$$

Together with (6.2), this proves the proposition. □

References

[Aar97] Jon Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997. MR1450400. Cited page 15.

[AL02] Goulnara N. Arzhantseva and Igor G. Lysenok, *Growth tightness for word hyperbolic groups*, Math. Z. **241** (2002), 597–611. MR1938706. Cited page 26.

[Anc87] Alano Ancona, *Negatively curved manifolds, elliptic operators, and the Martin boundary*, Ann. of Math. (2) **125** (1987), 495–536. MR890161. Cited pages 6, 11, and 19.

[BH99] Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999. MR1744486. Cited pages 7 and 21.

[BHM11] Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu, *Harmonic measures versus quasi-conformal measures for hyperbolic groups*, Ann. Sci. Éc. Norm. Supér. (4) **44** (2011), 683–721. MR2919980. Cited pages 2, 3, 6, and 11.

[Bou12] Jean Bourgain, *Finitely supported measures on $SL_2(\mathbb{R})$ which are absolutely continuous at infinity*, Geometric aspects of functional analysis, Lecture Notes in Math., vol. 2050, Springer, Heidelberg, 2012, pp. 133–141. MR2985129. Cited page 2.

[Bow98a] Brian H. Bowditch, *Boundaries of strongly accessible hyperbolic groups*, The Epstein birthday schrift, Geom. Topol. Monogr., vol. 1, Geom. Topol. Publ., Coventry, 1998, pp. 51–97. MR1668331. Cited page 20.

[Bow98b] ———, *Cut points and canonical splittings of hyperbolic groups*, Acta Math. **180** (1998), 145–186. MR1638764. Cited page 20.

[Cal13] Danny Calegari, *The ergodic theory of hyperbolic groups*, Geometry and topology down under, Contemp. Math., vol. 597, Amer. Math. Soc., Providence, RI, 2013, pp. 15–52. MR3186668. Cited page 25.

[CM07] Chris Connell and Roman Muchnik, *Harmonicity of quasiconformal measures and Poisson boundaries of hyperbolic spaces*, Geom. Funct. Anal. **17** (2007), 707–769. MR2346273. Cited page 18.

[Coo93] Michel Coornaert, *Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov*, Pacific J. Math. **159** (1993), 241–270. MR1214072. Cited pages 2, 3, 13, 17, 25, and 28.

[Dal99] Françoise Dal'bo, *Remarques sur le spectre des longueurs d'une surface et comptages*, Bol. Soc. Brasil. Mat. (N.S.) **30** (1999), 199–221. MR1703039. Cited page 19.

[dlH00] Pierre de la Harpe, *Topics in geometric group theory*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2000. MR1786869. Cited page 28.

[EK10] Anna Erschler and Anders Karlsson, *Homomorphisms to $\mathbb{R}$ constructed from random walks*, Ann. Inst. Fourier (Grenoble) **60** (2010), 2095–2113. MR2791651. Cited page 37.

[EK13] Anna Erschler and Vadim Kaimanovich, *Continuity of asymptotic characteristics for random walks on hyperbolic groups*, Funktsional. Anal. i Prilozhen. **47** (2013), 84–89. MR3113872. Cited pages 5, 10, 11, 14, and 32.

[FS09] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009. MR2483235. Cited page 28.

[GdlH90] Étienne Ghys and Pierre de la Harpe (eds.), *Sur les groupes hyperboliques d'après Mikhael Gromov*, Progress in Mathematics, vol. 83, Birkhäuser Boston Inc., Boston, MA, 1990, Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988. MR1086648. Cited pages 7, 36, and 39.

[GLJ93] Yves Guivarc'h and Yves Le Jan, *Asymptotic winding of the geodesic flow on modular surfaces and continued fractions*, Ann. Sci. École Norm. Sup. (4) **26** (1993), 23–50. MR1209912. Cited page 2.

[Gou13] Sébastien Gouëzel, *Martin boundary of measures with infinite support in hyperbolic groups*, Preprint, 2013. Cited pages 4, 11, 19, and 21.

[Gui80] Yves Guivarc'h, *Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire*, Conference on Random Walks (Kleebach, 1979) (French), Astérisque, vol. 74, Soc. Math. France, Paris, 1980, pp. 47–98, 3. MR588157. Cited page 3.

[Haï13] Peter Haïssinsky, *Marches aléatoires sur les groupes hyperboliques*, Géométrie ergodique, Monogr. Enseign. Math., vol. 43, Enseignement Math., Geneva, 2013, pp. 199–265. MR3220556. Cited pages 3 and 11.

[HY61] John G. Hocking and Gail S. Young, *Topology*, Addison-Wesley Publishing Co., 1961. MR0125557. Cited page 20.

[INO08] Masaki Izumi, Sergey Neshveyev, and Rui Okayasu, *The ratio set of the harmonic measure of a random walk on a hyperbolic group*, Israel J. Math. **163** (2008), 285–316. MR2391133. Cited pages 6, 20, and 21.

[Kai00] Vadim A. Kaimanovich, *The Poisson formula for groups with hyperbolic properties*, Ann. of Math. (2) **152** (2000), 659–692. MR1815698. Cited pages 9, 11, 16, and 30.

[KB02] Ilya Kapovich and Nadia Benakli, *Boundaries of hyperbolic groups*, Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001), Contemp. Math., vol. 296, Amer. Math. Soc., Providence, RI, 2002, pp. 39–93. MR1921706. Cited pages 19 and 20.

[KL11] Anders Karlsson and François Ledrappier, *Noncommutative ergodic theorems*, Geometry, rigidity, and group actions, Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 2011, pp. 396–418. 2807838. Cited page 9.

[Lal14] Steve Lalley, *Random walks on hyperbolic groups*, http://galton.uchicago.edu/∼lalley/Talks/paris-talkC.pdf, 2014. Cited page 4.

[Led95] François Ledrappier, *Applications of dynamics to compact manifolds of negative curvature*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 1195–1202. MR1404020. Cited page 2.

[LP07] Vincent Le Prince, *Dimensional properties of the harmonic measure for a random walk on a hyperbolic group*, Trans. Amer. Math. Soc. **359** (2007), 2881–2898 (electronic). MR2286061. Cited pages 2 and 40.

[MM07] Jean Mairesse and Frédéric Mathéus, *Random walks on free products of cyclic groups*, J. Lond. Math. Soc. (2) **75** (2007), 47–66. MR2302729. Cited page 36.

[Neu54] Bernhard H. Neumann, *Groups covered by finitely many cosets*, Publ. Math. Debrecen **3** (1954), 227–242 (1955). MR0072138. Cited page 25.

[Pei03] Marc Peigné, *On the Patterson-Sullivan measure of some discrete group of isometries*, Israel J. Math. **133** (2003), 77–88. MR1968423. Cited page 28.

[Rev84] Daniel Revuz, *Markov chains*, second ed., North-Holland Mathematical Library, vol. 11, North-Holland Publishing Co., Amsterdam, 1984. MR758799. Cited page 24.

[Sha98] Richard Sharp, *Relative growth series in some hyperbolic groups*, Math. Ann. **312** (1998), 125–132. MR1645953. Cited pages 23 and 24.

[Swa96] Gadde A. Swarup, *On the cut point conjecture*, Electron. Res. Announc. Amer. Math. Soc. **2** (1996), 98–100 (electronic). MR1412948. Cited page 20.

[Tan14] Ryokichi Tanaka, *Hausdorff spectrum of harmonic measure*, arXiv:1411.2312 [math.PR], 2014. Cited pages 3 and 6.

[Ver00] Anatoly Vershik, *Dynamic theory of growth in groups: entropy, boundaries, examples*, Uspekhi Mat. Nauk **55** (2000), 59–128. MR1786730. Cited pages 1 and 4.

[VSCC92] Nicholas Th. Varopoulos, Laurent Saloff-Coste, and Thierry Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992. MR1218884. Cited page 38.

[Woe00] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000. MR1743100. Cited page 37.

Sébastien Gouëzel, IRMAR, Univ. Rennes 1, 35042 Rennes Cedex, France
E-mail address: sebastien.gouezel@univ-rennes1.fr

Frédéric Mathéus, Univ. Bret. Sud, L.M.B.A., UMR 6205, BP 573, 56017 Vannes, France
E-mail address: Frederic.Matheus@univ-ubs.fr

François Maucourant, IRMAR, Univ. Rennes 1, 35042 Rennes Cedex, France
E-mail address: francois.maucourant@univ-rennes1.fr