ISSN 0012-9593

quatrième série - tome 48 fascicule 5 septembre-octobre 2015

ANNALES SCIENTIFIQUES de L'ÉCOLE NORMALE SUPÉRIEURE

Lewis BOWEN & Amos NEVO

Von Neumann and Birkhoff ergodic theorems for negatively curved groups

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Annales Scientifiques de l'École Normale Supérieure

Publiées avec le concours du Centre National de la Recherche Scientifique

Rédaction / *Editor*

Annales Scientifiques de l'École Normale Supérieure,
45, rue d'Ulm, 75230 Paris Cedex 05, France.
Tél. : (33) 1 44 32 20 88. Fax : (33) 1 44 32 20 80.
`annales@ens.fr`

Édition / *Publication*

Société Mathématique de France
Institut Henri Poincaré
11, rue Pierre et Marie Curie
75231 Paris Cedex 05
Tél. : (33) 01 44 27 67 99
Fax : (33) 01 40 46 90 96

Abonnements / *Subscriptions*

Maison de la SMF
Case 916 - Luminy
13288 Marseille Cedex 09
Fax : (33) 04 91 41 17 51
email : `smf@smf.univ-mrs.fr`

Tarifs

Europe : 515 €. Hors Europe : 545 €. Vente au numéro : 77 €.

ISSN 0012-9593

Directeur de la publication : Marc Peigné
Périodicité : 6 nos / an

Ann. Scient. Éc. Norm. Sup.
4^e série, t. 48, 2015, p. 1113 à 1147

VON NEUMANN AND BIRKHOFF ERGODIC THEOREMS FOR NEGATIVELY CURVED GROUPS

BY Lewis BOWEN* AND Amos NEVO†

Abstract. – We prove maximal inequalities for concentric ball and spherical shell averages on a general Gromov hyperbolic group, in arbitrary probability preserving actions of the group. Under an additional condition, satisfied for example by all groups acting isometrically and properly discontinuously on CAT(-1) spaces, we prove a pointwise ergodic theorem with respect to a sequence of probability measures supported on concentric spherical shells.

Résumé. – Pour tout groupe hyperbolique au sens de Gromov et pour toute action, préservant la mesure, sur un espace de probabilités, nous démontrons une inégalité maximale pour les moyennes sur des boules concentriques ou sur des anneaux sphériques concentriques de même épaisseur. Sous une hypothèse supplémentaire, valable par exemple pour les actions isométriques et proprement discontinues sur des espaces CAT(-1), nous démontrons de plus un théorème ergodique ponctuel pour une suite de mesures de probabilités à support dans des anneaux sphériques concentriques.

1. Introduction

1.1. Motivation and background

1.1.1. *The Arnol'd-Krylov problem.* – Given a dynamical system with invariant probability measure, von-Neumann's mean ergodic theorem [40] and Birkhoff's pointwise ergodic theorem [8] assert that the time evolution of the dynamical system distributes it evenly in the space. More concretely, under the sole assumption of ergodicity, namely the absence of invariant sets, when one samples the values of a function on the space at regular time intervals along the trajectory of the dynamical system, the average of these samples over time will converge (in the mean and pointwise) to the space average of the function, in accordance with Bolzmann's "ergodic hypothesis". Equivalently, the trajectories spend the right fraction of time in every subset of space, given by the measure of the subset.

* Lewis Bowen is supported in part by NSF grant DMS-0968762, NSF CAREER Award DMS-0954606 and BSF grant 2008274. † Amos Nevo is supported in part by ISF grant, and BSF grant 2008274.

The study of dynamical systems defined by a measure-preserving transformation is based on two main ingredients, namely that intervals are asymptotically invariant under translations, and that they satisfy the doubling property. Asymptotic invariance arguments play a crucial role in several proofs of Birkhoff's pointwise ergodic theorem, in Riesz's proof of von-Neumann's mean ergodic theorem, in the Krylov-Boglyobov proof of the existence of probability measures invariant under a continuous transformation of a compact space, in Calderón's transference principle reducing aspects of analysis of the orbits of the flow on phase space to analysis on the integers, and in Furstenberg's correspondence principle. The doubling property plays a crucial role in Wiener's covering argument which implies the Hardy-Littlewood maximal inequality, and thus also in the maximal ergodic theorem and the pointwise ergodic theorem. These arguments were extended over the years to actions by any finite number of commuting measure-preserving transformations.

Half a century ago, Arnol′d and Krylov [7] have raised the problem of establishing the following generalization of the classical ergodic theorems. Let Γ be a finitely generated group, and S a finite symmetric generating set. Consider the associated left invariant word metric on Γ, and let B_n and S_n denote the ball and the sphere of radius n with center e. Denote the uniform average on the ball by β_n, and on the sphere by σ_n. Given a measure-preserving action of Γ, consider sampling the values of a function on an orbit of Γ according to β_n. Does this averaging process converge in the mean and pointwise, and if so, what is its limit? Furthermore, when does the averaging process associated with σ_n converge ? Thus the Arnol′d-Krylov problem amounts to establishing ergodic theorems for *any choice* of finitely many measure-preserving transformations.

The Arnol′d-Krylov problem has proved to be very difficult and remains wide open. The only class of groups where a complete positive solution for β_n has been obtained is that of groups with polynomial volume growth. The sequence of balls in such groups satisfies the doubling condition, and the main ingredient in the relatively recent proof of the pointwise ergodic theorem is the fact that the sequence of word metric balls B_n is in fact asymptotically invariant under translations. This was established by Tessera [47] (see also [26]) and in a sharper form by Breuillard [18]. We refer to [42] and [3] for a detailed account of the proof of the pointwise ergodic theorem in this case.

The remarkable utility of asymptotically invariant sequences (exemplified very briefly above) resulted in the fact that much of the effort in developing ergodic theorems for countable groups has been devoted to generalizing the classical convergence results for a single transformation to ergodic actions of groups possessing an asymptotically invariant sequence. This class coincides with the class of amenable groups, and the averages studied most extensively have been uniform averages on a Følner sequence, often satisfying suitable regularity conditions introduced by Tempelman [46], generalizing the doubling condition. For more on the subject of ergodic theorems on amenable groups we refer to the foundational paper by Ornstein and Weiss [44], and for more recent results to Lindenstrauss's ergodic theorem for tempered Følner sequences [39] as well as the survey [49].

However, already for meta-Abelian (non-nilpotent) solvable groups, and certainly for non-amenable groups, the sequence of word-metric balls is not asymptotically invariant, so the methods developed for Følner sequences have not led to further progress on the Arnol′d-Krylov problem.

1.1.2. *Free groups.* – Turning to the problem of establishing ergodic theorems for non-amenable groups, we note that an important special case that figures prominently in the theory is that of free (non-Abelian) groups. In the case of the word metric associated with free generators, the symmetry inherent in the radial structure implies that the convolution algebra generated by σ_n is commutative, and this fact opens the door to a spectral approach to the problem. This approach was initiated already by Arnol′d and Krylov [7] who proved an equidistribution theorem for radial averages on dense free subgroups of isometries of the unit sphere $\mathbb{S}^2$ via a spectral argument similar to Weyl's equidistribution theorem on the circle. Guivarc'h has established a mean ergodic theorem for radial averages on the free group, using von-Neumann's original approach via the spectral theorem [36]. The pointwise ergodic theorem for the averages $\frac{1}{2}(\sigma_n+\sigma_{n+1})$ in general actions of the free group was proved in [41] for L^2-functions using spectral theory, and extended to function in L^p, $p > 1$, in [43], using more refined spectral methods.

Another successful approach to ergodic theorems on free groups and more general Markov groups is based on the theory of Markov operator. Grigorchuk [34] has applied the Hopf-Dunford-Scwartz operator ergodic theorem to deduce a pointwise ergodic theorem for the uniform averages $\mu_k = \frac{1}{k+1}\sum_{n=0}^{k}\sigma_n$ of the spheres. This approach was subsequently generalized by Bufetov to weighted averages of spheres on general Markov groups [20]. Using Rota's approach to the operator ergodic theorem via martingale theory, Bufetov [21] has extended pointwise convergence of the averages $\frac{1}{2}(\sigma_n + \sigma_{n+1})$ on the free group to the function space $L\log L$. Pointwise almost sure convergence for the uniform averages μ_k of the spheres on general Markov groups for bounded functions has been established recently in [22, Cor. 1], and under additional assumption also in [45]. It is a reflection of the difficulty of the Arnol′d-Krylov problem that in both of these results, the limit has not been identified, in general. Under a suitable mixing assumption on the action, it was shown in [32] that the limit is indeed the space average. Finally, we note that an ergodic theorem for actions of general word-hyperbolic groups on finite spaces was established in [12].

1.1.3. *Ergodic theorems for lattice subgroups.* – An important extension of the Arnol′d-Krylov problem is to consider balls and spheres defined by natural left-invariant metrics on the group, not necessarily given by word-metrics. Thus, when the group Γ is a discrete subgroup of a locally compact group, one can consider left-invariant metrics on G restricted to Γ. Extending the scope of the problem gives rise to many natural examples, of which we mention the following. When Γ is a lattice subgroup of a (non-compact) simple Lie group with finite center, consider (any) non-trivial linear representation $\tau : G \to SL_n(\mathbb{R})$, and restrict (any) norm on $SL_n(\mathbb{R})$ to $\tau(\Gamma)$. It was shown in [33] that in any ergodic action of Γ, the associated ball averages converge in norm and pointwise for any function $f \in L^p$, provided $1 < p < \infty$. Similar results hold more generally for irreducible lattices in S-algebraic groups, and we refer to [33] for further details. We remark that the methods of [33] rely in a crucial manner on spectral theory, namely on the unitary representation theory of the semisimple group G involved, and are thus limited only to those countable groups which arise as lattice subgroups of G.

More generally, when Γ acts isometrically and properly discontinuously on a locally compact metric space (X, d), one can consider the (pseudo) metrics given by

$d_x(\gamma_1, \gamma_2) = d(\gamma_1 x, \gamma_2 x)$. In the present paper our main concern will be the case where the metric space in question is hyperbolic, for example a CAT(-1) space. Before turning to describe our main results, let us note briefly that we will completely avoid any spectral considerations. Rather, the basic principle underlying our approach is the realization that one can utilize the amenable actions of a group in order to construct families of ergodic averages on it. Thus a remarkable feature of this approach is that it treats amenable and non-amenable groups on an equal footing, and in fact allows us to utilize many of the classical asymptotic invariance and doubling arguments in a much more general context. Indeed, the generality of our approach is underscored by the fact that for any countable group Γ, the Poisson boundary associated with a generating probability measure is an amenable ergodic action of Γ [50], so that any countable group admits such an action. For the case of free groups this approach is explained in full in [14], where the ergodic theorems of [43] and [21] are generalized. Some of the ingredients necessary in our approach were developed in [15], and in the present paper we will utilize and generalize the constructions appearing there.

We now proceed to introduce the necessary definitions and state the main results.

1.2. Basic definitions

Let Γ be a countable group and $\{\zeta_r\}_{r>0}$ a family of probability measures on Γ. Given a pmp (probability-measure-preserving) action $\Gamma \curvearrowright (X, m)$, we can associate to each ζ_r an operator $\pi_X(\zeta_r)$ on $L^p(X, m)$, acting by:

$$\pi_X(\zeta_r)(f)(x) = \sum_{g \in \Gamma} \zeta_r(g) f(g^{-1}x).$$

We also consider the associated maximal function:

$$\mathbb{M}_\zeta[f](x) := \sup_{r>0} \pi_X(\zeta_r)(|f(x)|).$$

We will usually suppress π_X from the notation and write simply $\pi_X(\zeta_r)f = \zeta_r f$. Let us recall the following definitions :

- $\{\zeta_r\}_{r>0}$ satisfies the *strong L^p maximal inequality* if there is a constant $C_p > 0$ such that $\|\mathbb{M}_\zeta[f]\|_p \leq C_p \|f\|_p$ for every $f \in L^p(X, m)$;
- $\{\zeta_r\}_{r>0}$ satisfies the *$L \log L$ maximal inequality* if there is a constant $C_1 > 0$ such that $\|\mathbb{M}_\zeta[f]\|_{L \log L} \leq C_1 \|f\|_1$ for every $f \in L^1(X, m)$;
- $\{\zeta_r\}_{r>0}$ is a *pointwise convergent family in L^p* if $\zeta_r(f)$ converges pointwise a.e. for every $f \in L^p(X, m)$;
- $\{\zeta_r\}_{r>0}$ is a *pointwise ergodic family in L^p* if $\zeta_r(f)$ converges pointwise a.e. to $\mathbb{E}[f|\Gamma]$, the conditional expectation of f on the sigma-algebra of Γ-invariant Borel sets (for every $f \in L^p(X, m)$).

The purpose of the present work is to establish maximal and pointwise ergodic theorems for natural geometric averages on word hyperbolic groups. Before describing the probability measures we will be interested in, let us recall the following definitions. Given a proper left-invariant metric d on Γ, the *Gromov product* of $x, y \in \Gamma$ relative to $z \in \Gamma$ is

$$(x|y)_z := \frac{1}{2}\Big(d(x, z) + d(y, z) - d(x, y)\Big).$$

The pair (Γ, d) is a *hyperbolic group* if for some $\delta \geq 0$,

$$(x|y)_w \geq \min\{(x|z)_w, (y|z)_w\} - \delta, \quad \forall x, y, w, z \in \Gamma. \tag{1.1}$$

Let $I \subset \mathbb{R}$ be an interval. A map $\gamma : I \to \Gamma$ is a (λ, c)-quasi-geodesic if $\lambda^{-1}|i-j| - c \leq d(\gamma(i), \gamma(j)) \leq \lambda|i-j| + c$ for every i, j. We say that (Γ, d) is *uniformly quasi-geodesic* if there exists a constant $c > 0$ such that for every pair of elements $x, y \in \Gamma$ there exists a $(1, c)$-quasi-geodesic from x to y.

1.3. Statement of main results

Our first result concerns maximal inequalities for radial averages. For $r > 0$, let β_r be a probability measure on Γ which is uniformly distributed on the closed ball $B(e, r)$ of radius r centered at the identity. In other words, $\beta_r(g) = |B(e, r)|^{-1}$ if $g \in B(e, r)$ and $\beta_r(g) = 0$ otherwise. Similarly, for a fixed $a > 0$, we denote by $\sigma_{r,a}$ the uniform probability measure on the spherical shell $S_{r,a}(e) = \{g \in \Gamma\,;\, r - a < |g| \leq r + a\}$. Finally we let $\mu_{r,a} = \frac{1}{r}\int_0^r \sigma_{s,a} ds$ be the uniform averages of the spherical shell measures. Our main maximal inequality is as follows.

THEOREM 1.1. – *Let (Γ, d) be a non-elementary uniformly quasi-geodesic hyperbolic group. Then the family of ball averages $\{\beta_r\}_{r>0}$ satisfies the $L \log L$-maximal inequality and the strong L^p maximal inequality for every $p > 1$. The same holds for the families $\{\sigma_{r,a}\}_{r>0}$ and $\{\mu_{r,a}\}_{r>0}$, provided a is larger than a fixed constant depending only on (Γ, d).*

Our second result considers the case of word metrics d_S, where S is a finite symmetric set of generators for Γ and $d_S(g, e) = |g|_S$ is the word length. We let σ_n denote the uniform probability measure on the sphere $S_n(e)$ of radius n and center e, and $\mu_n = \frac{1}{n+1}\sum_{k=0}^n \sigma_k$ their uniform averages. We then have

THEOREM 1.2. – *Let Γ be a word-hyperbolic group, S a symmetric set of generators. Then μ_n satisfies the strong maximal inequality in L^p, $1 < p < \infty$, and in $L \log L$, and is a pointwise (and mean) convergent family in these spaces.*

We remark that Theorem 1.2 improves on the L^2-maximal inequality for μ_n established in [32], extends the mean and pointwise convergence in L^p, $1 < p < \infty$, proved for μ_n in [22, Corollary 1] and [23, Thm. 1] to the function space $L \log L$, and generalizes the pointwise convergence for bounded functions established for μ_n (under an additional assumption) also in [45].

Let us turn now to the problem of pointwise (and mean) convergence of the balls and spherical shell averages on a hyperbolic group (Γ, d). We will require an additional assumption on the group, and we will comment on its prevalence and on the optimality of the ergodic theorem it gives rise to after stating the theorem.

We will say that the *Gromov boundary coincides with the horofunction boundary* if for every point ξ in the Gromov boundary of (Γ, d), every sequence $\{x_i\}_{i=1}^\infty \subset \Gamma$ converging to ξ and every $y \in \Gamma$, the limit

$$h_\xi(y) := \lim_{i \to \infty} d(x_i, y) - d(x_i, e)$$

exists and depends only on ξ, y. Our main pointwise ergodic theorem is:

THEOREM 1.3. – *Let (Γ, d) be a non-elementary uniformly quasi-geodesic hyperbolic group whose Gromov boundary coincides with its horofunction boundary. Then for each $a > 0$ larger than a fixed constant depending only on (Γ, d), there exists a family $\{\kappa_r\}_{r>0}$ of probability measures on Γ such that*

1. *each κ_r is supported on the spherical shell $S_{r,a}(e) = \{g \in \Gamma : d(e, g) \in (r - a, r + a]\}$,*
2. *$\{\kappa_r\}_{r>0}$ is a pointwise (and mean) ergodic family in L^p for every $p > 1$ and in $L \log L$.*

1.4. Comments on ergodic theorems for radial averages

As to the hypotheses of Theorem 1.3, let us note that the coincidence of the horofunction boundary and the Gromov boundary is a common phenomenon, but not a universal one. Let us give two examples where it is satisfied, and one where it may fail.

1. *CAT(−1) spaces*. Suppose Γ acts properly discontinuously by isometries on a CAT(−1) space (X, d_X). For $x \in X$ (with trivial stability group in Γ) define the metric d on Γ by $d(g, g') := d_X(gx, g'x)$. For a proof of the coincidence of the boundaries for this metric see [19, Chapter II.8, Theorem 8.13 and Chapter III.H].
2. *The Green metric*. If ζ is a finitely supported symmetric probability measure on Γ whose support generates Γ then the *Green metric* induced by ζ is defined as follows. Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables each with law ζ. Let $Z_0 = e$ and $Z_n = X_1 \cdots X_n$ (for $n \geq 1$) be the random walk on Γ induced by ζ. For $g, g' \in \Gamma$, let $d_\zeta(g, g') = -\log(p(g, g'))$ where $p(g, g')$ is the probability that $gZ_n = g'$ for some $n \geq 0$. This is the Green metric induced by ζ. The informative paper [9] explains why the horofunction boundary of (Γ, d_ζ) coincides with its Gromov boundary. This is based on work of Ancona [4, 5, 6] showing that the Martin boundary of (Γ, d_ζ) equals its Gromov boundary. In [10] it is shown that the Green metric is uniformly quasi-geodesic and Gromov hyperbolic if Γ is word hyperbolic. It is obviously proper and left-invariant. It follows that every non-elementary finitely generated word hyperbolic group has a metric d satisfying the hypotheses of Theorem 1.3.
3. *Word metrics*. By [30], the horofunction boundary of Γ with an arbitrary word metric admits a canonical finite-to-1 Γ-equivariant map onto the Gromov boundary. However, this map need not be a homeomorphism. For example, if $\Gamma = \Gamma_0 \times F$ where Γ_0 is word hyperbolic and F is a finite nontrivial group then for any word metric on Γ that is induced by a generating set of the form $\{(g, e_F) : g \in S\} \cup \{(e_{\Gamma_0}, f) : f \in F\}$ where S is a generating set for Γ_0, the horofunction boundary does not coincide with the Gromov boundary.

Let us now comment briefly on the optimality of Theorem 1.3 in the context of ergodic theorems for ball averages on hyperbolic groups. To begin with, recall the well-known fact that the ball averages on the free group, defined with respect to a free set of generators, actually fail to converge, in general. Indeed, whenever $L_0^2(X, m)$ contains an eigenfunction ϕ with $\pi_X(\sigma_n)\phi = (-1)^n\phi$, clearly the sequences $\pi_X(\sigma_n)\phi$ and $\pi_X(\beta_n)\phi$ do not converge, and the same holds true for spherical shells. Thus even when the horofunction boundary and the Gromov boundary coincide, as in the case of the free group acting on its Cayley tree, the ball averages (and the spherical shell averages) do not satisfy the ergodic theorem. This is not an isolated fact, and for example in the groups $\mathbb{Z}_p * \mathbb{Z}_q$ where $p \neq q$, the ergodic theorem likewise fails for ball (and spherical shell) averages, see the discussion following [42, Thm. 10.7].

Let us now point out however that in the cases just mentioned there exists a sequence of probability measures supported on the spherical shells $S_{n,1}(e)$ which is indeed a pointwise ergodic sequence. It is given by $\sigma'_n = \frac{1}{2}(\sigma_n + \sigma_{n+1})$ in these examples, see [41]. By the previous comment, the probability measures in question supported on the shell $S_{n,1}(e)$, are necessarily non-uniform in the examples mentioned. Thus Theorem 1.3 gives essentially the optimal result in this case, namely pointwise convergence for probability measures supported on spherical shells.

Finally, let us point out that in some cases, the ball averages associated with a hyperbolic metric d on Γ do in fact form a pointwise ergodic sequence in L^p, $1 < p < \infty$. Indeed, let G be a connected almost simple real Lie group of real rank one, and Γ a uniform lattice in G. Let $\mathcal{S}$ denote the symmetric space associated with G, and fix a point $p \in \mathcal{S}$ whose stabilizer in Γ is trivial. Let $D_{\mathcal{S}}$ be the G-invariant Riemannian metric on symmetric space, and define $d_{\mathcal{S}}(g,h) = D_{\mathcal{S}}(gp, hp)$ for $g, h \in \Gamma$. Then $d_{\mathcal{S}}$ is a hyperbolic metric on Γ, and the associated ball averages β_r are a pointwise (and mean) ergodic family in L^p, $1 < p < \infty$. This fact appears in [33], and its proof depends on detailed information regarding the unitary representation theory of the simple Lie group G.

1.5. A brief sketch

Let $\partial\Gamma$ denote the Gromov boundary of (Γ, d). Via the Patterson-Sullivan construction, there is a quasi-conformal probability measure ν on $\partial\Gamma$. So there are constants $C, \mathfrak{v} > 0$ such that

$$C^{-1}\exp(-\mathfrak{v} h_\xi(g^{-1})) \leq \frac{d\nu \circ g}{d\nu}(\xi) \leq C\exp(-\mathfrak{v} h_\xi(g^{-1}))$$

for every $g \in \Gamma$ and a.e. $\xi \in \partial\Gamma$ where

$$h_\xi(g^{-1}) := \inf \liminf_{n\to\infty} d(x_i, g^{-1}) - d(x_i, e)$$

where the infimum is over all sequences $\{x_i\} \subset \Gamma$ converging to ξ.

The *type* of the action $\Gamma \curvearrowright (\partial\Gamma, \nu)$ encodes the essential range of the Radon-Nikodym derivative. In [13], it is shown that this type is III_λ for some $\lambda \in (0,1]$. If $\lambda \in (0,1)$ then we set

$$R_\lambda(g,\xi) = -\log_\lambda\left(\frac{d\nu \circ g}{d\nu}(\xi)\right).$$

Using standard results, it can be shown that then we can choose ν so that $R_\lambda(g,\xi) \in \mathbb{Z}$ for every g and a.e. ξ. When $\lambda = 1$, we set $R_1(g,\xi) = +\log\left(\frac{d\nu\circ g}{d\nu}(\xi)\right)$.

In order to handle each case uniformly, we set $\mathbb{L} = \mathbb{R}$ if $\lambda = 1$ and $\mathbb{L} = \mathbb{Z}$ if $\lambda \in (0,1)$. Then we let Γ act on $\partial\Gamma \times \mathbb{L}$ by

$$g(\xi, t) = (g\xi, t - R_\lambda(g,\xi)).$$

This action preserves the measure $\nu \times \theta_\lambda$ where θ_1 is the measure on $\mathbb{R}$ satisfying $d\theta_1(t) = \exp(t)dt$ and, for $\lambda \in (0,1)$, θ_λ is the measure on $\mathbb{Z}$ satisfying $\theta_\lambda(\{n\}) = \lambda^{-n}$.

Given $a, b \in \mathbb{L}$, let $[a,b]_{\mathbb{L}} \subset \mathbb{L}$ denote the interval $\{a, a+1, \ldots, b\}$ if $\mathbb{L} = \mathbb{Z}$ and $[a,b] \subset \mathbb{R}$ if $\mathbb{L} = \mathbb{R}$. Similar considerations apply to open intervals and half-open intervals.

For any real numbers $r, T > 0$, and $(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}$, let

$$\Gamma_r(\xi, t) = \{g \in \Gamma : \ d(g, e) - h_\xi(g) - t \leq r, \ g^{-1}(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}\}$$

and

$$\mathfrak{B}_r(\xi, t) := \{g^{-1}(\xi, t) : \ g \in \Gamma_r(\xi, t)\}.$$

$\Gamma_r(\xi, t)$ is approximately equal to the intersection of the ball of radius r centered at the identity with the horoshell $\{g \in \Gamma : \ -t \leq h_\xi(g) \leq T - t\}$. Of course, Γ_r and $\mathfrak{B}_r$ depend on T, but we leave this dependence implicit.

Our first main technical result is that if T is sufficiently large then $\{\mathfrak{B}_r\}_{r>0}$ is *regular*: there exists a constant $C > 0$ such that for every $r > 0$ and a.e. $(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}$,

$$|\bigcup_{s \leq r} \mathfrak{B}_s^{-1} \mathfrak{B}_r(\xi, t)| \leq C |\mathfrak{B}_r(\xi, t)|.$$

Theorem 1.1 now follows from an extension of the general results in [15, 16]. The idea is that we can use the regularity of the sets $\mathfrak{B}_r$ and prove a maximal inequality for them, and thus for the equivalence relation on $\partial\Gamma \times [0, T)_{\mathbb{L}}$ given by the intersection of the Γ-orbits with this subset. We can then average this maximal inequality over $\partial\Gamma$ to obtain a maximal inequality for the resulting family of probability measures on Γ. By geometric arguments, the family we obtain is sufficiently close to being uniform averages on balls that this implies a maximal inequality for the uniform averages on balls.

The results of [15, 16] do not directly apply because the action of Γ on its boundary might not be essentially free. However, this action has uniformly bounded stabilizers. Using this hypothesis we generalize the needed theorems of [15, 16] in §2-2.2.

Next we let $\mathscr{S}_a = \{\mathfrak{S}_{r,a}\}_{r>0}$ be the family of subset functions on $\partial\Gamma \times [0, T)_{\mathbb{L}}$ defined by

$$\mathfrak{S}_{r,a}(\xi, t) := \mathfrak{B}_r(\xi, t) \setminus \mathfrak{B}_{r-a}(\xi, t)$$

and observe that $\mathscr{S}_a$ is also regular if $a, T > 0$ are sufficiently large.

Our second main technical result is that $\mathscr{S}_a$ is *asymptotically invariant* (modulo a minor technical issue) assuming that the horofunction boundary coincides with the Gromov boundary. To explain, we let E denote the equivalence relation on $\partial\Gamma \times [0, T)_{\mathbb{L}}$ given by $(\xi, t) E (\xi', t')$ if there exists $g \in \Gamma$ such that $(\xi, t) = g(\xi', t')$. The *full group* of E is the group of all (equivalence classes of) Borel automorphisms on $\partial\Gamma \times [0, T)_{\mathbb{L}}$ with graph contained in E, where two such automorphisms are equivalent if they agree almost everywhere. It is denoted by $[E]$. A subset $\Phi \subset [E]$ *generates* E if for a.e. $(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}$ and every (ξ', t') with $(\xi, t) E (\xi', t')$ there is an element $\phi \in \langle \Phi \rangle$ (the subgroup generated by Φ) such that $\phi(\xi, t) = (\xi', t')$. Finally, $\mathscr{S}_a$ being asymptotically invariant means that there exists a countable generating set $\Phi \subset [E]$ such that

$$\lim_{r \to \infty} \frac{|\mathfrak{S}_{r,a}(\xi, t) \vartriangle \phi(\mathfrak{S}_{r,a}(\xi, t))|}{|\mathfrak{S}_{r,a}(\xi, t)|} = 0$$

for every $\phi \in \Phi$ and a.e. (ξ, t).

Using asymptotic invariance, it now follows from general results of [15, 16] (as generalized in §2-2.2) that there is a pointwise *convergent* family $\{\kappa_r\}_{r>0}$ of probability measures on Γ and a constant $a > 0$ so that each κ_r is supported on the annulus

$$\{g \in \Gamma : \ d(e, g) \in [r - a, r + a]\}.$$

However, this is not the end of the proof because at this stage we only know that for any pmp action $\Gamma \curvearrowright (X, m)$ and any $f \in L^p(X, m)$ that $\kappa_r(f)$ converges almost everywhere. We have not yet identified what it converges to!

The issue is that even if $\Gamma \curvearrowright (X, m)$ is ergodic, it does not necessarily follow that the product action $\Gamma \curvearrowright (X \times \partial\Gamma \times \mathbb{L}, m \times \nu \times \theta_\lambda)$ is ergodic. To resolve this, first we show that $\Gamma \curvearrowright (\partial\Gamma, \nu)$ is weakly mixing (so $\Gamma \curvearrowright (X \times \partial\Gamma, m \times \nu)$ is ergodic). This uses the fact that Poisson boundary actions are weakly mixing [1] and that the action is equivalent to a Poisson boundary action [27]. From [13], it follows that $\Gamma \curvearrowright (\partial\Gamma, \nu)$ has type III_ρ and stable type III_τ for some $\rho, \tau \in (0, 1]$. From this it follows that the natural cocycle $\alpha : \tilde{E} \to \Gamma$ (where $\tilde{E}$ is the equivalence relation on $X \times \partial\Gamma \times [0, T)_{\mathbb{L}}$) is weakly mixing relative to a certain compact group. This is ultimately what is needed to invoke Theorems 2.2, 2.3 (which generalize [15, 16]) and thereby complete the proof.

Organization of the paper. – §2 discusses maximal and ergodic theorems for measured equivalence relations. This is used in §2.2 to obtain some general ergodic theorems which will be used to prove the main results. Then §3 reviews Gromov hyperbolic groups and sets some notation. §4 establishes the regularity of the averaging sets and proves Theorem 1.1 and Theorem 1.2. In §5, we prove asymptotic invariance of the averaging sets. The last section §5.4 uses asymptotic invariance to complete the proof of Theorem 1.3.

Convention. – We have not attempted to produce explicit estimates for the constants appearing in the statements of the main results, and throughout the paper we use the "variable constant convention", namely in different occurrences of a constant (even within the same argument) the values it assumes may be different.

2. Equivalence relations and ergodic sequences

2.1. An ergodic theorem for equivalence relations

The purpose of this section is to review and generalize the main theorems of [15, 16] so that we can later apply them to Gromov hyperbolic groups. To this end, let (B, ν) be a standard probability space and $E \subset B \times B$ a discrete measurable equivalence relation. Let $[E]$ denote the *full group* of E, namely the group of all measurable automorphisms of B with graph contained in E (discarding a null set). We assume that ν is E-invariant, namely that $\phi_*\nu = \nu$ for every $\phi \in [E]$.

We will obtain theorems for E first and then push them forward via a cocycle $E \to \Gamma$. We begin by discussing ergodic theorems for E, which utilize finite subset functions, defined next.

Let 2^B be the set of all subsets of B. A (finite) *subset function* for E is a map $\mathfrak{F} : B \to 2^B$ such that $\mathfrak{F}(\xi)$ is finite and $\mathfrak{F}(\xi) \subset [\xi]_E = \{\eta \in B : (\xi, \eta) \in E\}$ for almost every ξ. $\mathfrak{F}$ is called measurable if the set $\{(\xi, \eta) \in B \times B : \eta \in \mathfrak{F}(\xi)\}$ is measurable.

We let $\mathfrak{F}^{-1}$ be the subset function defined by

$$\mathfrak{F}^{-1}(\eta) = \{\xi \in B : \eta \in \mathfrak{F}(\xi)\}.$$

If $\mathfrak{F}, \mathfrak{G}$ are two subset functions on B then their product is defined by

$$\mathfrak{F}\mathfrak{G}(\xi) = \bigcup_{\eta \in \mathfrak{G}(\xi)} \mathfrak{F}(\eta).$$

The union, intersection and relative complement of two subset functions are defined pointwise. These constitute subset functions in their own right, namely they are measurable and finite for almost every $\xi \in B$.

We will be interested in one-parameter families $\mathscr{F} = \{\mathfrak{F}_r\}_{r>0}$ of set functions. Note that $\mathscr{F}$ denotes the family while each $\mathfrak{F}_r$ is a set function. We say that such a family is *measurable* if $\{(\xi, \eta, r) \in B \times B \times \mathbb{R}_{>0} :\ \eta \in \mathfrak{F}_r(\xi)\}$ is measurable.

We will be interested in averaging over such subset functions. First, let $\alpha : E \to \mathrm{Aut}(X, m)$ be a measurable cocycle into the automorphism group of a standard Borel space. In particular, we require $\alpha(\xi, \eta)\alpha(\eta, \xi') = \alpha(\xi, \xi')$ (for a.e. $\xi \in B$ and every $\eta, \xi' \in [\xi]_E$). Let E_α be the induced equivalence relation on $B \times X$. So $(\xi, u)E_\alpha(\xi', u')$ if and only if $\xi E \xi'$ and $\alpha(\xi', \xi)u = u'$.

Let $f \in L^p(B \times X, \nu \times m)$ for some p. Given a family of $\mathscr{F} = \{\mathfrak{F}_r\}_{r>0}$ of subset functions for E, we consider the averages

$$\mathbb{A}[f|\mathfrak{F}_r](\xi, u) := |\mathfrak{F}_r(\xi)|^{-1} \sum_{\xi' \in \mathfrak{F}_r(\xi)} f(\xi', \alpha(\xi', \xi)u)$$

and the maximal function

$$\mathbb{M}[f|\mathscr{F}](\xi, u) := \sup_{r>0} \mathbb{A}[|f||\mathfrak{F}_r](\xi, u).$$

Our assumption that the equivalence classes $[\xi]_E$ are almost always countable, and the subsets $\mathfrak{F}_r(\xi) \subset [\xi]_E$ almost always finite implies that the maximal function is measurable. We say :

– $\mathscr{F}$ satisfies the *weak* $(1, 1)$*-type maximal inequality* if there is a constant $C > 0$ such that for all $t > 0$ and all $f \in L^1(B \times X)$,

$$\nu \times m(\{(\xi, u) :\ \mathbb{M}[f|\mathscr{F}](\xi, u) > t\}) \le C \frac{\|f\|_1}{t};$$

– $\mathscr{F}$ satisfies the *strong* p*-type maximal inequality* if there is a constant $C_p > 0$ such that $\|\mathbb{M}[f|\mathscr{F}]\|_p \le C_p \|f\|_p$ for every $f \in L^p(B \times X)$;
– $\mathscr{F}$ is a *pointwise ergodic family in* L^p if for every $f \in L^p(B \times X)$, $\mathbb{A}[f|\mathfrak{F}_r]$ converges pointwise a.e. to $\mathbb{E}[f|E_\alpha]$, the conditional expectation of f on the σ-algebra of E_α-saturated sets (these are the measurable sets which are unions of E_α classes).

Next we provide conditions which imply the conditions above.

A family $\mathscr{F} = \{\mathfrak{F}_r\}_{r>0}$ of subset functions on B is *regular* if there exists a constant $C > 0$ (called a *regularity constant*) such that for every $r > 0$ and a.e. $\xi \in B$,

$$\left| \bigcup_{s \le r} \mathfrak{F}_s^{-1} \mathfrak{F}_r(\xi) \right| \le C |\mathfrak{F}_r(\xi)|.$$

A subset $\Phi \subset [E]$ *generates* E if for a.e. $\xi \in B$ and every $\eta \in [\xi]_E$, there is $\phi \in \langle\Phi\rangle$ (the subgroup generated by Φ) such that $\phi(\xi) = \eta$. The family $\mathscr{F}$ is *asymptotically invariant* if there exists a countable set $\Phi \subset [E]$ which generates E such that

$$\lim_{r\to\infty} \frac{|\mathfrak{F}_r(\xi) \vartriangle \phi(\mathfrak{F}_r(\xi))|}{|\mathfrak{F}_r(\xi)|} = 0 \quad \text{for a.e. } \xi, \forall\phi \in \Phi.$$

We now recall the following, which is part of [15, Theorems 2.4-2.6].

THEOREM 2.1. – *If $\mathscr{F}$ is regular then it satisfies the weak $(1,1)$-type maximal inequality and the strong p-type maximal inequality for all $p > 1$. If, in addition, it is asymptotically invariant then it is a pointwise ergodic family in L^p (for every $p \geq 1$).*

Most of the effort in this paper goes towards showing that certain subset functions on equivalence relations obtained from the action of a hyperbolic group on its boundary are both regular and asymptotically invariant. Next we explain how these results imply pointwise ergodic theorems for Γ. We also need to generalize previous results because the action of Γ on its boundary need not be essentially free.

2.2. From equivalence relations to ergodic sequences

We begin by recalling the definition of the Maharam extension of a general non-singular action, and the method of deriving ergodic theorems from it.

DEFINITION 2.1. – The *Maharam extension* of a measure-class-preserving action $\Gamma \curvearrowright (B,\nu)$ is the action $\Gamma \curvearrowright (B \times \mathbb{R}, \nu \times \theta)$ defined by

$$\gamma(\xi,t) := (\gamma\xi, t - R(g,\xi)), \quad R(g,\xi) := \log\left(\frac{d\nu\circ g}{d\nu}(\xi)\right)$$

and θ is the measure on $\mathbb{R}$ satisfying $d\theta(t) = e^t dt$. This action is measure-preserving.

If $\Gamma \curvearrowright (B,\nu)$ is ergodic then we say $\Gamma \curvearrowright (B,\nu)$ has *type III_1* if the Maharam extension is also ergodic.

DEFINITION 2.2. – A measure-class preserving action $\Gamma \curvearrowright (B,\nu)$ has *uniformly bounded stabilizers* if there is a constant $C > 0$ such that for a.e. $\xi \in B$, $|\mathrm{Stab}_\Gamma(\xi)| \leq C$ where $\mathrm{Stab}_\Gamma(\xi) = \{g \in \Gamma : g\xi = \xi\}$.

THEOREM 2.2 (The III_1 case). – *Let Γ be a countable group and $\Gamma \curvearrowright (B,\nu)$ a measure-class preserving action on a standard probability space with uniformly bounded stabilizers. Let $T > 0$, $\mathscr{F} = \{\mathfrak{F}_r\}_{r>0}$ be a measurable family of set functions for the equivalence relation E on $B \times [0,T]$ (induced from the Maharam extension $\Gamma \curvearrowright B \times \mathbb{R}$ as above) and let $\psi \in L^q(B,\nu)$ be a probability density (so $\psi \geq 0$, $\int \psi\, d\nu = 1$).*

Define probability measures ζ_r on Γ by

$$\zeta_r(g) = T^{-1}\int_0^T \int |\{w \in \Gamma : w(\xi,t) \in \mathfrak{F}_r(\xi,t)\}|^{-1} 1_{\mathfrak{F}_r(\xi,t)}(g^{-1}(\xi,t))\psi(\xi)\, d\nu(\xi)dt.$$

Let $p > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$.

– *If $\mathscr{F}$ is regular then $\{\zeta_r\}$ satisfies the strong L^p maximal inequality. If $\psi \in L^\infty(B,\nu)$ then $\{\zeta_r\}$ satisfies the $L\log L$ maximal inequality.*

- *If $\mathscr{F}$ is regular and asymptotically invariant then $\{\zeta_r\}$ is a pointwise convergent family in L^p (and if $\psi \in L^\infty(B,\nu)$ then it is pointwise convergent in $L\log L$).*
- *If $\mathscr{F}$ is regular, asymptotically invariant, $\Gamma \curvearrowright (B,\nu)$ is weakly mixing, type III_1 and stable type III_λ (where either $\lambda = 1$ or T is a positive integer multiple of $-\log(\lambda) > 0$) then $\{\zeta_r\}$ is a pointwise ergodic family in L^p (and if $\psi \in L^\infty(B,\nu)$ then it is pointwise ergodic in $L\log L$).*

We refer to [16] for background on type and stable type. We say that $\Gamma \curvearrowright (B,\nu)$ is *weakly mixing* if for any ergodic pmp action $\Gamma \curvearrowright (X,m)$, the product action $\Gamma \curvearrowright (B\times X, \nu\times m)$ is ergodic.

Proof. – To begin, let us assume that $\Gamma \curvearrowright (B,\nu)$ is essentially free. We will show that this result follows from [16, Theorems 3.1 and 5.1]. By Theorem 2.1, if $\mathscr{F}$ is regular then it satisfies the weak $(1,1)$-type maximal inequality and if it is both regular and asymptotically invariant then it is poinwise ergodic in L^p (for every $p \ge 1$).

Let $\alpha : E \to \Gamma$ be the cocycle $\alpha(\eta,\xi) = \gamma$ if $\gamma\xi = \eta$. For $(\eta,\xi) \in E$, let $\omega_r(\eta,\xi) = |\mathfrak{F}_r(\xi)|^{-1}$ if $\eta \in \mathfrak{F}_r(\xi)$ and $\omega_r(\eta,\xi) = 0$ otherwise. If $\mathscr{F}$ is regular then $\Omega = \{\omega_r\}$ satisfies the weak $(1,1)$-type maximal inequality and the strong L^p maximal inequality (in the sense of [16, §2.1]). If $\mathscr{F}$ is regular and asymptotically invariant, then Ω is a pointwise ergodic family in L^p (for every $p \ge 1$).

Let $K = \mathbb{R}/T\mathbb{Z}$ act on $B \times [0,T]$ by $k(\xi,t) = (\xi, t+k)$ where $t+k$ is taken modulo T. By [16, Theorem 5.1], if $\Gamma \curvearrowright (B,\nu)$ is weakly mixing, type III_1 and stable type III_λ (where either $\lambda = 1$ or $0 < T = -\log(\lambda) < \infty$) then α is weakly mixing relative to the action of K. The result now follows from [16, Theorem 3.1].

Let us suppose now that $\Gamma \curvearrowright (B,\nu)$ is not necessarily essentially free but does have uniformly bounded stabilizers. Let $\Gamma \curvearrowright (Y,p)$ be a nontrivial Bernoulli shift action. This action is essentially free, pmp and strongly mixing. It therefore enjoys the following multiplier property: if $\Gamma \curvearrowright (X,m)$ is any properly ergodic action then the product action $\Gamma \curvearrowright (X\times Y, m\times p)$ is also ergodic. In particular, it is not necessary for $\Gamma \curvearrowright (X,m)$ to be probability-measure-preserving.

Observe that the product action $\Gamma \curvearrowright (B\times Y, \nu\times p)$ is essentially free. Let $\Gamma \curvearrowright B\times Y\times\mathbb{R}$ denote the Maharam extension of $\Gamma \curvearrowright B\times Y$ and $\tilde{E}$ the induced equivalence relation on $B\times Y\times[0,T]$. Define the subset function $\widetilde{\mathfrak{F}}_r$ on $B\times Y\times[0,T]$ by $\widetilde{\mathfrak{F}}_r(\xi,y,t) = \{g(\xi,y,t) :\ g\in\Gamma, g(\xi,t)\in\mathfrak{F}_r(\xi,t)\}$. Note that

$$(2.1) \qquad \zeta_r(g) = \frac{1}{T}\iiint_0^T |\widetilde{\mathfrak{F}}_r(\xi,y,t)|^{-1} 1_{\widetilde{\mathfrak{F}}_r(\xi,y,t)}(g^{-1}(\xi,y,t))\psi(\xi)\ dt d\nu(\xi) dp(y).$$

Because $\Gamma \curvearrowright (B,\nu_B)$ has uniformly bounded stabilizers, there is a constant $C > 0$ such that

$$|\mathfrak{F}_r(\xi,t)| \le |\widetilde{\mathfrak{F}}_r(\xi,y,t)| \le C|\mathfrak{F}_r(\xi,t)| \ \text{ for a.e. } (\xi,y,t).$$

If $\mathscr{F}$ is regular, this implies $\widetilde{\mathscr{F}} := \{\widetilde{\mathfrak{F}}_r\}_{r>0}$ is also regular. Therefore, the essentially free case implies that $\{\zeta_r\}$ satisfies the strong L^p-maximal inequality and, if $q = \infty$, the $L\log L$-type maximal inequality.

Let us suppose now that $\mathscr{F}$ is asymptotically invariant. We will show that $\widetilde{\mathscr{F}}$ is asymptotically invariant. There exists a countable set $\Phi \subset [E]$ such that Φ generates E and

$$\lim_{r\to\infty} \frac{|\phi(\mathfrak{F}_r(\xi,t)) \vartriangle \mathfrak{F}_r(\xi,t)|}{|\mathfrak{F}_r(\xi,t)|} = 0$$

for a.e. (ξ, t) and every $\phi \in \Phi$.

Let $J : B \times Y \times [0,T] \to [0,1]$ be a Borel isomorphism and choose

$$L : B \times Y \times [0,T] \to \{1, 2, \dots, C\}$$

to satisfy: for a.e. (ξ, y, t)

– if $g \in \mathrm{Stab}_\Gamma(\xi, t)$ and $J(\xi, y, t) < J(g(\xi, y, t))$ then $L(\xi, y, t) < L(g(\xi, y, t))$;

– $\max\{L(g(\xi, h, t)) : \ g \in \mathrm{Stab}_\Gamma(\xi, t)\} = |\mathrm{Stab}_\Gamma(\xi, t)|$.

For each $\phi \in \Phi$ and $n \in \mathbb{Z}$ define $\tilde{\phi}_n \in [\tilde{E}]$ by $\tilde{\phi}_n(\xi, y, t) = (\xi', y', t')$ where $\phi(\xi, t) = (\xi', t')$, $(\xi, y, t)\tilde{E}(\xi', y', t')$ and $L(\xi, y, t) \equiv L(\xi', y', t') + n \mod |\mathrm{Stab}_\Gamma(\xi, t)|$. This is well-defined almost everywhere. Observe that for any $\phi \in \Phi$ and a.e. $(\xi, y, t) \in B \times Y \times [0,T]$ if $g \in \Gamma$ is such that $\phi(\xi, t) = g(\xi, t)$ then there exists $i \in \mathbb{Z}$ so that $\tilde{\phi}_i(\xi, y, t) = g(\xi, y, t)$. Therefore $\tilde{\Phi} := \{\tilde{\phi}_n : \ \phi \in \Phi, n \in \mathbb{Z}\}$ is generating.

This construction implies that for any $\phi \in \Phi, n \in \mathbb{Z}$ and a.e. $(\xi, y, t) \in B \times Y \times [0,T]$, $|\widetilde{\mathfrak{F}}_r(\xi, y, t)| \le C|\mathfrak{F}_r(\xi, t)|$ and

$$|\tilde{\phi}_n(\widetilde{\mathfrak{F}}_r(\xi, y, t)) \vartriangle \widetilde{\mathfrak{F}}_r(\xi, y, t)| \le C|\phi(\mathfrak{F}_r(\xi, t)) \vartriangle \mathfrak{F}_r(\xi, t)|.$$

So

$$\lim_{r\to\infty} \frac{|\tilde{\phi}_n(\widetilde{\mathfrak{F}}_r(\xi, y, t)) \vartriangle \widetilde{\mathfrak{F}}_r(\xi, y, t)|}{|\widetilde{\mathfrak{F}}_r(\xi, y, t)|} = 0$$

which implies $\widetilde{\mathscr{F}}$ is asymptotically invariant. So (2.1) and the essentially free case imply $\{\zeta_r\}$ is a pointwise convergent family in L^p (and in $L \log L$ if $q = +\infty$).

Next let us assume that $\Gamma \curvearrowright B$ is weakly mixing, type III_1 and stable type III_λ. Because $\Gamma \curvearrowright Y$ is weakly mixing and pmp, it follows immediately that the product action $\Gamma \curvearrowright B \times Y$ is weakly mixing and stable type III_λ. Because $\Gamma \curvearrowright B \times \mathbb{R}$ is ergodic, the action $\Gamma \curvearrowright (B \times \mathbb{R}) \times Y$ is also ergodic. But this is isomorphic to the Maharam extension of $\Gamma \curvearrowright B \times Y$. Therefore, $\Gamma \curvearrowright B \times Y$ has type III_1. The conclusion now follows from the essentially free case and (2.1). □

Definition 2.3. – Suppose $\Gamma \curvearrowright (B, \nu)$ is a measure-class-preserving action, $\lambda \in (0, 1)$ and the Radon-Nikodym derivatives satisfy

$$R_\lambda(g, \xi) := -\log_\lambda \left(\frac{d\nu \circ g}{d\nu}(\xi) \right) \in \mathbb{Z}$$

for a.e. $\xi \in B$ and $g \in \Gamma$. In this case, we consider the *discrete Maharam extension* which is the action $\Gamma \curvearrowright (B \times \mathbb{Z}, \nu \times \theta_\lambda)$ defined by

$$\gamma(\xi, t) := (\gamma\xi, t - R_\lambda(g, \xi))$$

where θ_λ is the measure on $\mathbb{Z}$ satisfying $\theta_\lambda(\{n\}) = \lambda^{-n}$. This action is measure-preserving.

If, in addition, $\Gamma \curvearrowright (B, \nu)$ is ergodic then $\Gamma \curvearrowright (B, \nu)$ has *type* III_λ if the discrete Maharam extension is also ergodic. (Type III_λ is also well-defined if the Radon-Nikodym derivatives do not satisfy the above condition: see [37] or [16] for background on type.)

Theorem 2.3 (The III_λ case). – *Let Γ be a countable group and $\Gamma \curvearrowright (B,\nu)$ a measure-class preserving action on a standard probability space with uniformly bounded stabilizers and Radon-Nikodym derivatives which satisfy*

$$R_\lambda(g,\xi) := -\log_\lambda\left(\frac{d\nu\circ g}{d\nu}(\xi)\right) \in \mathbb{Z}$$

for a.e. $\xi \in B$ and $g \in \Gamma$.

Let $\mathscr{F} = \{\mathfrak{F}_r\}_{r>0}$ be a measurable family of set functions for the equivalence relation E on $B \times \{0,1,\ldots,N-1\}$ (induced from the discrete Maharam extension $\Gamma \curvearrowright B \times \mathbb{Z}$ as above) and let $\psi \in L^q(B,\nu)$ be a probability density (so $\psi \geq 0$, $\int \psi \, d\nu = 1$).

Define probability measures ζ_r on Γ by

$$\zeta_r(g) = N^{-1}\sum_{t=0}^{N-1}\int |\{w \in \Gamma :\ w(\xi,t) \in \mathfrak{F}_r(\xi,t)\}|^{-1} 1_{\mathfrak{F}_r(\xi,t)}(g^{-1}(\xi,t))\psi(\xi)\, d\nu(\xi).$$

Let $p > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$.

- *If $\mathscr{F}$ is regular then $\{\zeta_r\}$ satisfies the strong L^p maximal inequality. If $\psi \in L^\infty(B,\nu)$ then $\{\zeta_r\}$ satisfies the $L\log L$ maximal inequality.*
- *If $\mathscr{F}$ is regular and asymptotically invariant then $\{\zeta_r\}$ is a pointwise convergent family in L^p (and if $\psi \in L^\infty(B,\nu)$ then it is pointwise convergent in $L\log L$).*
- *If $\mathscr{F}$ is regular, asymptotically invariant, $\Gamma \curvearrowright (B,\nu)$ is weakly mixing, type III_λ and stable type III_τ for $\tau = \lambda^m$ (some $m \in \mathbb{N}$ such that $(N/m) \in \mathbb{Z}$) then $\{\zeta_r\}$ is a pointwise ergodic family in L^p (and if $\psi \in L^\infty(B,\nu)$ then it is pointwise ergodic in $L\log L$).*

Proof. – The essentially free case follows from [16, Theorems 3.1 and 5.2]. The rest of the proof is analogous to the proof of Theorem 2.2 so we leave it to the reader. □

3. Gromov hyperbolic spaces

We now turn to establish some properties of Gromov hyperbolic spaces, the Gromov boundary and the horofunction boundary. These results will be applied to the case where (Γ, d) is a nonelementary uniformly quasi-geodesic hyperbolic group.

3.1. The Gromov boundary

Let $(\mathcal{X}, d_{\mathcal{X}})$ be a δ-hyperbolic space. A sequence $\{x_i\}_{i=1}^\infty$ in $\mathcal{X}$ is a *Gromov sequence* if

$$\lim_{i,j\to\infty}(x_i|x_j)_z = +\infty$$

for some (and hence, any) basepoint $z \in \mathcal{X}$. Two Gromov sequences $\{x_i\}_{i=1}^\infty$, $\{y_i\}_{i=1}^\infty$ are *equivalent* if $\lim_{i\to\infty}(x_i|y_i)_z = +\infty$ with respect to some (and hence any) basepoint z. It is an exercise to show that this defines an equivalence relation (assuming $(\mathcal{X}, d_{\mathcal{X}})$ is δ-hyperbolic). The *Gromov boundary* is the space of equivalence classes of Gromov sequences. We denote it by $\partial\mathcal{X}$, leaving the metric implicit. Let $\overline{\mathcal{X}}$ denote $\mathcal{X} \cup \partial\mathcal{X}$.

The Gromov product extends to $\partial\mathcal{X}$ as follows. Let $p, z \in \mathcal{X}$ and $\xi, \eta \in \partial\mathcal{X}$. Define

$$(\xi|p)_z := \inf\liminf_{i\to\infty}(x_i|p)_z, \quad (\xi|\eta)_z := \inf\liminf_{i\to\infty}(x_i|y_i)_z$$

where the infimums are over all sequences $\{x_i\}_{i=1}^\infty \in \xi, \{y_i\}_{i=1}^\infty \in \eta$. By [48, Lemma 5.11]

$$\limsup_{i\to\infty}(x_i|y_i)_z - 2\delta \le (\xi|\eta)_z \le \liminf_{i\to\infty}(x_i|y_i)_z \tag{3.1}$$

for any sequences $\{x_i\}_{i=1}^\infty \in \xi, \{y_i\}_{i=1}^\infty \in \eta$. These inequalities also hold if $\eta = p \in \mathcal{X}$ and y_i is any sequence with $\lim_{i\to\infty} y_i = p$. According to [48, Proposition 5.12], inequality (1.1) extends to $x, y \in \partial\mathcal{X}$.

In [19] it is shown that if $\epsilon > 0$ is sufficiently small and $\bar{d}_\epsilon : \overline{\mathcal{X}} \times \overline{\mathcal{X}} \to \mathbb{R}$ is defined by

$$\bar{d}_\epsilon(\xi,\eta) := e^{-\epsilon(\xi|\eta)_z}$$

then there exist a metric $\bar{d}$ on $\overline{\mathcal{X}}$ and constants $A, B > 0$ such that $A\bar{d}_\epsilon \le \bar{d} \le B\bar{d}_\epsilon$. Any such metric is called a *visual metric*. The topology on $\mathcal{X}$ induced by $d_\mathcal{X}$ agrees with the topology induced by $\bar{d}$. Moreover $\mathcal{X}$ is dense in $\overline{\mathcal{X}}$.

3.2. Quasi-conformal measures and horofunctions

Let $(\mathcal{X}, d_\mathcal{X})$ be a δ-hyperbolic metric space. Choose a basepoint $x_0 \in \mathcal{X}$.

LEMMA 3.1. – *Let $\xi \in \partial\mathcal{X}$ and suppose $\{y_i\}, \{z_i\} \subset \mathcal{X}$ are two sequences converging to ξ (w.r.t. the topology on $\overline{\mathcal{X}}$). Then for any $w \in \mathcal{X}$,*

$$\limsup_{i\to\infty}\left|d_\mathcal{X}(y_i,w) - d_\mathcal{X}(y_i,x_0) - \Big(d_\mathcal{X}(z_i,w) - d_\mathcal{X}(z_i,x_0)\Big)\right| \le 4\delta.$$

Proof. – Observe that

$$d_\mathcal{X}(y_i,w) - d_X(y_i,x_0) = d_\mathcal{X}(w,x_0) - 2(y_i|w)_{x_0}.$$

A similar statement holds for z_i in place of y_i. Thus,

$$\left|d_\mathcal{X}(y_i,w) - d_\mathcal{X}(y_i,x_0) - \Big(d_\mathcal{X}(z_i,w) - d_\mathcal{X}(z_i,x_0)\Big)\right| = 2\left|(y_i|w)_{x_0} - (z_i|w)_{x_0}\right|.$$

The lemma now follows from (3.1). □

For $\xi \in \partial\mathcal{X}$, define $h_\xi : \mathcal{X} \to \mathbb{R}$ by

$$h_\xi(z) := \inf\liminf_{i\to\infty} d_\mathcal{X}(z,y_i) - d_\mathcal{X}(y_i,x_0)$$

where the infimum is over all sequences $\{y_i\} \subset \mathcal{X}$ which converge to ξ. This is the *horofunction* associated to ξ (and the basepoint x_0). By the previous lemma, if $\{x_i\}$ is any sequence converging to ξ and $z \in \mathcal{X}$ is arbitrary then

$$\limsup_{i\to\infty}\left|h_\xi(z) - \Big(d_\mathcal{X}(x_i,z) - d_\mathcal{X}(x_i,x_0)\Big)\right| \le 4\delta. \tag{3.2}$$

Together with 3.1 this implies: if $\xi \in \partial\mathcal{X}$ and $z \in \mathcal{X}$ then

$$d_\mathcal{X}(z,x_0)/2 - h_\xi(z)/2 - 4\delta \le (z|\xi)_{x_0} \le d_\mathcal{X}(z,x_0)/2 - h_\xi(z)/2. \tag{3.3}$$

DEFINITION 3.2 (Quasi-conformal measure). – Suppose (Γ, d) is a Gromov hyperbolic group. A Borel probability measure ν on $\partial\Gamma$ is *quasi-conformal* if there are constants $\mathfrak{v}, C > 0$ such that for any $g \in \Gamma$ and a.e. $\xi \in \partial\Gamma$,

$$C^{-1}\exp(-\mathfrak{v}h_\xi(g^{-1})) \le \frac{d\nu\circ g}{d\nu}(\xi) \le C\exp(-\mathfrak{v}h_\xi(g^{-1})).$$

We will call $\mathfrak{v} > 0$ the *quasi-conformal constant* associated to ν.

It is well-known that if d comes from a word metric on Γ (or more generally, any geodesic metric) then there is a quasi-conformal measure on $\partial\Gamma$ [29]. More generally:

LEMMA 3.3. – *Let (Γ, d) be a non-elementary, uniformly quasi-geodesic, hyperbolic group. Then there exists a quasi-conformal measure ν on $\partial\Gamma$. Moreover, any two quasi-conformal measures are equivalent. Also if $\mathfrak{v}$ is the quasi-conformal constant of ν then there is a constant $C > 0$ such that*

1. *If $B(g,r)$ denotes the ball of radius r centered at $g \in \Gamma$ then*
$$C^{-1}e^{\mathfrak{v}r} \leq |B(g,r)| \leq Ce^{\mathfrak{v}r}, \quad \forall g \in \Gamma, r > 0.$$
2. $$C^{-1}e^{-\mathfrak{v}n} \leq \nu\left(\{\xi' \in \partial\Gamma : (\xi|\xi')_e \geq n\}\right) \leq Ce^{-\mathfrak{v}n}, \quad \forall n > 0, \xi \in \partial\Gamma.$$

Proof. – This follows immediately from [10, Theorem 2.3] and the fact that any non-elementary uniformly quasi-geodesic hyperbolic group is a proper quasi-ruled hyperbolic space by Lemma 3.4 below. □

The paper [10] contains many results for hyperbolic spaces under the assumption that these spaces are *quasi-ruled*. To be precise a metric space $(\mathcal{X}, d_{\mathcal{X}})$ is quasi-ruled if there are constants (τ, λ, c) such that $(\mathcal{X}, d_{\mathcal{X}})$ is (λ, c)-quasi-geodesic and for any (λ, c)-quasi-geodesic $\gamma : [a,b] \to \mathcal{X}$ and any $a \leq s \leq t \leq u \leq b$,
$$d_{\mathcal{X}}(\gamma(s),\gamma(t)) + d_{\mathcal{X}}(\gamma(t),\gamma(u)) - d_{\mathcal{X}}(\gamma(s),\gamma(u)) \leq 2\tau.$$

LEMMA 3.4. – *If $(\mathcal{X}, d_{\mathcal{X}})$ is $(1, c)$-quasi-geodesic then it is quasi-ruled.*

Proof. – If $\gamma : [a,b] \to \mathcal{X}$ is any $(1, c)$-quasi-geodesic then for any $a \leq s \leq t \leq u \leq b$,
$$d_{\mathcal{X}}(\gamma(s),\gamma(t)) + d_{\mathcal{X}}(\gamma(t),\gamma(u)) - d_{\mathcal{X}}(\gamma(s),\gamma(u)) \leq 3c.$$
So we may set $\tau = 3c/2$. □

The action of Γ on its boundary need not be essentially free. For example, if Γ_0 is word hyperbolic and F is a finite group then F, considered as a subgroup of $\Gamma_0 \times F$, acts trivially on the Gromov boundary of $\Gamma_0 \times F$. However, it does have uniformly bounded stabilizers (in the sense of Definition 2.2), a condition which is crucial to Theorems 2.2 and 2.3.

LEMMA 3.5. – *Let (Γ, d) be a non-elementary uniformly quasi-geodesic hyperbolic group and ν be a quasi-conformal measure on $\partial\Gamma$. Then there is a constant $C > 0$ such that ν-a.e. $\xi \in \partial\Gamma$, $|\mathrm{Stab}_\Gamma(\xi)| \leq C$.*

Proof. – Let $g \in \Gamma$ have infinite order. It is well-known that there exist distinct elements $\{g^-, g^+\} \subset \partial\Gamma$ such that $\lim_{n\to\infty} g^n = g^+$ and $\lim_{n\to\infty} g^{-n} = g^-$. Moreover if $\xi \in \partial\Gamma \setminus \{g^-, g^+\}$ then $\lim_{n\to\infty} g^n\xi = g^+$ and $\lim_{n\to\infty} g^{-n}\xi = g^-$.

Let $A \subset \partial\Gamma$ be the union of the points g^- and g^+ for all infinite-order elements $g \in \Gamma$. Because ν has no atoms (by Lemma 3.3) and A is a countable set, $\nu(A) = 0$. Moreover, if $\xi \in \partial\Gamma \setminus A$ and g is any infinite order element then $g \notin \mathrm{Stab}_\Gamma(\xi)$ since $\lim_{n\to\infty} g^n\xi = g^+ \neq \xi$.

Thus if $g \in \mathrm{Stab}_\Gamma(\xi)$ and $\xi \in \partial\Gamma \setminus A$ then g has finite order, i.e., $\mathrm{Stab}_\Gamma(\xi)$ is a torsion subgroup. Because the Tits alternative holds for hyperbolic groups [35], every torsion subgroup of Γ is finite. It is well-known (see e.g., [11], [19] or [17]) that there is a constant $C > 0$ such that for every finite subgroup $H < \Gamma$, $|H| \leq C$. This proves $|\mathrm{Stab}_\Gamma(\xi)| \leq C$. □

3.3. The type of the boundary action

Let $\partial\Gamma$ denote the Gromov boundary of Γ and let ν be a quasi-conformal measure on $\partial\Gamma$ with quasi-conformal constant $\mathfrak{v}$. By [13], $\Gamma \curvearrowright (\partial\Gamma, \nu)$ is type III_λ for some $\lambda \in (0,1]$.

LEMMA 3.6. – *Suppose $\lambda \in (0,1)$. Then there exists a quasi-conformal Borel probability measure ν' on $\partial\Gamma$ such that if*

$$R_\lambda(g,\xi) = -\log_\lambda\left(\frac{d\nu' \circ g}{d\nu'}(\xi)\right)$$

then $R_\lambda(g,b) \in \mathbb{Z}$ for a.e. $\xi \in \partial\Gamma$.

Proof. – By [2] $\Gamma \curvearrowright (\partial\Gamma, \nu)$ is amenable. So by [28], the orbit-equivalence relation on $\partial\Gamma$ is generated by a single measure-class-preserving Borel isomorphism $T : \partial\Gamma \to \partial\Gamma$. If $\lambda \in (0,1)$ then by [37, Proposition 2.2], there exists a Borel probability measure ν' on $\partial\Gamma$ which is equivalent to ν such that if

$$R_\lambda(g,\xi) = -\log_\lambda\left(\frac{d\nu' \circ g}{d\nu'}(\xi)\right)$$

then $R_\lambda(g,b) \in \mathbb{Z}$ for a.e. $\xi \in \partial\Gamma$ and every $g \in \Gamma$. A careful look at the proof reveals that ν' can be chosen so that the Radon-Nikodym derivatives between ν and ν' are bounded. More precisely, there is a constant $C > 0$ such that

$$C^{-1} \leq \frac{d\nu'}{d\nu} \leq C$$

almost everywhere. Therefore ν' is also quasi-conformal. □

We now assume that $\nu = \nu'$ satisfies the lemma above. By quasi-conformality, there exists a constant $C > 0$ such that

$$|R_\lambda(g,\xi) - \mathfrak{v}\log(\lambda)^{-1}h_\xi(g^{-1})| \leq C.$$

In order to streamline the exposition, let $\mathbb{L}$ denote either $\mathbb{R}$ or $\mathbb{Z}$ depending on whether $\lambda = 1$ or $\lambda \in (0,1)$. Also let

$$R_1(g,\xi) = R(g,\xi) = \log\left(\frac{d\nu \circ g}{d\nu}(\xi)\right).$$

By quasi-conformality,

$$|R_1(g,\xi) + \mathfrak{v}h_\xi(g^{-1})| \leq C$$

for some constant $C > 0$. So if we let $\mathfrak{v}_\lambda = -\mathfrak{v}\log(\lambda)^{-1}$ for $\lambda \in (0,1)$ and $\mathfrak{v}_1 = \mathfrak{v}$, then we can say

$$|R_\lambda(g,\xi) + \mathfrak{v}_\lambda h_\xi(g^{-1})| \leq C \tag{3.4}$$

for every $\lambda \in (0,1], g \in \Gamma$ and a.e. $\xi \in \partial\Gamma$.

Next we set some useful notation. Let θ_λ be the measure on $\mathbb{L}$ given by $d\theta_1(t) = e^t dt$ (if $\lambda = 1$) and $\theta_\lambda(\{n\}) = \lambda^{-n}$ if $\lambda \in (0,1)$. The *Maharam extension* of the action $\Gamma \curvearrowright (\partial\Gamma, \nu)$ is the action $\Gamma \curvearrowright (\partial\Gamma \times \mathbb{L}, \nu \times \theta_\lambda)$ given by

$$g(\xi,t) = (g\xi, t - R_\lambda(g,\xi)).$$

This action preserves the measure $\nu \times \theta_\lambda$.

Let $\mathbb{L}^+$ denote the set of positive elements of $\mathbb{L}$. Also, for $A < B \in \mathbb{L}$, we will let $[A,B)_{\mathbb{L}}$ denote the half-open interval in $\mathbb{L}$ from A to B. So if $\mathbb{L} = \mathbb{Z}$, then

$$[A,B)_{\mathbb{L}} = \{A, A+1, \ldots, B-1\}.$$

4. Volume growth and regularity

The purpose of this section is to prove Theorem 1.1 by applying Theorems 2.2, 2.3 and estimating the cardinality of the intersection of balls and horoshells.

4.1. Regularity of the averaging sets

Recall the definition of $R_\lambda(g,b)$ and the Maharam extension $\Gamma \curvearrowright \partial\Gamma \times \mathbb{L}$ from the previous section.

Definition 4.1. – Fix $a > 0$, $T \in \mathbb{L}^+$ and, for $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$, let

$$\begin{aligned} \Gamma_r(\xi,t) &= \{g \in \Gamma :\ d(e,g) - h_\xi(g) - t \le r, g^{-1}(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}\} \\ \mathfrak{B}_r(\xi,t) &= \{g^{-1}(\xi,t) :\ g \in \Gamma_r(\xi,t)\} \\ \mathfrak{S}_{a,r} &= \mathfrak{B}_r \setminus \mathfrak{B}_{r-a}. \end{aligned}$$

We let $\mathscr{B} = \{\mathfrak{B}_r\}_{r>0}, \mathscr{S}_a = \{\mathfrak{S}_{a,r}\}_{r>0}$ denote the corresponding families of subset functions. Although these definitions depend on T we will leave this dependence implicit in the notation.

We will show that $\mathscr{B}$ and $\mathscr{S}_a$ are regular if a, T are sufficiently large. We begin with an estimate of $|\mathfrak{B}_r|$.

Lemma 4.2. – *There exist constants $a_0, T_0 > 0$ such that if $T \ge T_0, a \ge a_0$ then for a.e. $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$*

$$C^{-1} e^{\mathfrak{v}r/2} \le |\mathfrak{B}_r(\xi,t)|, |\Gamma_r(\xi,t)|, |\mathfrak{S}_{r,a}(\xi,t)| \le C e^{\mathfrak{v}r/2} \quad \forall r \ge 2T + 2a$$

where the constant $C > 0$ may depend on T but not on ξ or r.

Proof. – Because stabilizers are uniformly bounded by Lemma 3.5, $C_0^{-1}|\Gamma_r(\xi,t)| \le |\mathfrak{B}_r(\xi,t)| \le C_0|\Gamma_r(\xi,t)|$ for some $C_0 > 1$. Because $\mathfrak{S}_{r,a} = \mathfrak{B}_r \setminus \mathfrak{B}_{r-a}$, the bound for $\mathfrak{B}_r$ implies the bound for $\mathfrak{S}_{r,a}$. So it suffices to estimate $|\Gamma_r(\xi,t)|$.

By (3.4) there is a constant $C_1 > 1$ such that for a.e. $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$, if $g \in \Gamma$ is such that $g^{-1}(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$, then

$$|R_\lambda(g^{-1},\xi)| \le C_1|h_\xi(g)| + C_1, \quad |h_\xi(g)| \le C_1|R_\lambda(g^{-1},\xi)| + C_1.$$

Moreover, $0 \le t \le T$ implies $|R_\lambda(g^{-1},\xi)| \le T$. We may assume $T > 1$. So,

$$\begin{aligned} B(e, r - T - C_1) \cap h_\xi^{-1}\left[\frac{t - T + C_1}{C_1}, \frac{t - C_1}{C_1}\right] &\subset \Gamma_r(\xi,t) \\ &\subset B\left(e, r + (C_1 + T)^2\right) \cap h_\xi^{-1}[-C_1 - C_1T, C_1 + C_1T] \end{aligned}$$

where $B(e,r)$ is the ball of radius r centered at the identity in Γ.

In [13, Lemma 6.3], it is shown that there is a constant $T_0 > 0$ such that if $T_2 \geq T_1$ are such that $T_2 - T_1 \geq T_0$, $\xi \in \partial\Gamma$ and $r \geq \max(|T_1|, |T_2|) - 2c$ then

$$C^{-1}e^{\mathfrak{v}(r+T_2)/2} \leq |B(e,r) \cap h_\xi^{-1}[T_1,T_2]| \leq Ce^{\mathfrak{v}(r+T_2)/2}$$

where $C > 0$ is a constant which may depend on T_1, T_2 but not on r, ξ. So the inclusions above imply the lemma. □

LEMMA 4.3. – *Let $T_0 > 0$ be as in Lemma 4.2. If $T > T_0$ then the family $\mathscr{B} = \{\mathfrak{B}_r\}_{r>0}$ is regular.*

Proof. – Fix $k_0, k_1, k_2 \in \partial\Gamma \times [0,T)_{\mathbb{L}}$ such that $k_1 \in \mathfrak{B}_r(k_0) \cap \mathfrak{B}_r(k_2)$. To make the notation simpler we will write $x \lesssim y$ if $x \leq y + C$ where C is a constant that may depend on T and (Γ, d) but not on r, k_0, k_1 or k_2. Of course, $x \gtrsim y$ means $y \lesssim x$ and $x \approx y$ means both $x \lesssim y$ and $y \lesssim x$.

Let $g_1 \in \Gamma_r(k_0)$ be such that $g_1^{-1}k_0 = k_1$ and let $g_2 \in \Gamma$ be such that $g_2^{-1}g_1 \in \Gamma_r(k_2)$ and $g_1^{-1}g_2k_2 = k_1$. Note $d(e,g_1) \lesssim r$ and $d(g_2^{-1}g_1, e) = d(g_1,g_2) \lesssim r$.

Let $k_0 = (\xi, t)$. Because ν is quasi-conformal and $g_1^{-1}k_0 = k_1 \in \partial\Gamma \times [0,T)_{\mathbb{L}}$, we must have $|h_\xi(g_1)| \lesssim \mathfrak{v}^{-1}T$ which implies $h_\xi(g_1) \approx 0$. Because $g_1^{-1}g_2k_2 = k_1 = g_1^{-1}k_0$, we have $k_2 = g_2^{-1}k_0$. Therefore, $h_\xi(g_2) \approx 0$ as well.

CLAIM. – *We have $d(g_2, e) \lesssim r$.*

Proof of claim. – By δ-hyperbolicity,

$$(e|g_2)_{g_1} \gtrsim \min\{(e|\xi)_{g_1}, (\xi|g_2)_{g_1}\}.$$

So either

$$2(e|g_2)_{g_1} = d(e,g_1) + d(g_2,g_1) - d(e,g_2) \gtrsim 2(e|\xi)_{g_1} \approx d(e,g_1) + h_\xi(g_1) \approx d(e,g_1)$$

which implies

$$d(e,g_2) \lesssim d(g_2,g_1) \lesssim r$$

or

$$2(e|g_2)_{g_1} = d(e,g_1)+d(g_2,g_1)-d(e,g_2) \gtrsim 2(\xi|g_2)_{g_1} \approx h_\xi(g_1)+d(g_2,g_1)-h_\xi(g_2) \approx d(g_2,g_1)$$

which implies

$$d(e,g_2) \lesssim d(e,g_1) \lesssim r.$$

This proves the claim. □

The claim implies $d(e,g_2) - h_\xi(g_2) - t \lesssim r$. Moreover, $g_2^{-1}k_0 = g_2^{-1}g_1k_1 = k_2$. So $g_2 \in \Gamma_{r+C_0}(k_0)$ (for some constant $C_0 > 0$ which may depend on T and (Γ, d) but not on r or the k_i's). Thus $k_2 \in \mathfrak{B}_{r+C_0}(k_0)$. Because k_0, k_1, k_2 are arbitrary, this establishes that for any $k_0 \in \partial\Gamma \times [0,T)_{\mathbb{L}}$,

$$\mathfrak{B}_{r+C_0}(k_0) \supset \bigcup_{s\leq r} \mathfrak{B}_s^{-1}\mathfrak{B}_r(k_0).$$

By Lemma 4.2, there is a constant $C_1 > 0$ such that $\left|\mathfrak{B}_{r+C_0}(k_0)\right| \leq C_1 \left|\mathfrak{B}_r(k_0)\right|$. Therefore,

$$\left|\bigcup_{s\leq r} \mathfrak{B}_s^{-1}\mathfrak{B}_r(k_0)\right| \leq C_1 \left|\mathfrak{B}_r(k_0)\right|.$$

Since C_1 does not depend on r or k_0, $\mathscr{B}$ is regular. □

LEMMA 4.4. – *Let W be a set. Let $\mathscr{F} = \{\mathfrak{F}_r\}_{r>0}$ be a regular family of subset functions on W. Suppose there is a constant $C > 0$ and a family $\mathscr{G} = \{\mathfrak{G}_r\}_{r>0}$ of subset functions on W that satisfies $\mathfrak{G}_r(w) \subset \mathfrak{F}_r(w)$ and $|\mathfrak{G}_r(w)| \geq C|\mathfrak{F}_r(w)|$ for every $w \in W$. Then $\mathscr{G}$ is regular.*

Proof. – Let $C_{\mathscr{F}}$ be a regularity constant for $\mathscr{F}$. Then for any $r > 0$ and $w \in W$,

$$|\bigcup_{s\leq r} \mathfrak{G}_s^{-1}\mathfrak{G}_r(w)| \leq |\bigcup_{s\leq r} \mathfrak{F}_s^{-1}\mathfrak{F}_r(w)| \leq C_{\mathscr{F}}|\mathfrak{F}_r(w)| \leq C^{-1}C_{\mathscr{F}}|\mathfrak{G}_r(w)|.$$ □

COROLLARY 4.5. – Let $a_0, T_0 > 0$ be as in Lemma 4.2. If $a > a_0$ and $T > T_0$ then the family $\mathscr{S}_a = \{\mathfrak{S}_{r,a}\}_{r>0}$ is regular. Moreover, suppose $\epsilon_0 > 0$ and

$$\epsilon : \partial\Gamma \times [0, T)_{\mathbb{L}} \times [0, \infty) \to [-\epsilon_0, \epsilon_0]$$

is any function. Define $\widetilde{\mathfrak{B}}_r(\xi, t) := \mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi, t)$ and $\widetilde{\mathfrak{S}}_{r,a}(\xi, t) := \widetilde{\mathfrak{B}}_r(\xi, t) \setminus \widetilde{\mathfrak{B}}_{r-a}(\xi, t)$ then $\widetilde{\mathscr{B}} := \{\widetilde{\mathfrak{B}}_r\}_{r>0}$ and $\widetilde{\mathscr{S}_a} := \{\widetilde{\mathfrak{S}}_{r,a}\}_{r>0}$ are regular.

Proof. – This follows immediately from Lemmas 4.2 - 4.4. □

COROLLARY 4.6. – Let $\psi \in L^q(\partial\Gamma, \nu)$ be a probability distribution (so $\psi \geq 0$ and $\int \psi \, d\nu = 1$). For $(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}$, let

$$\Gamma_{r,a}(\xi, t) := \{g \in \Gamma : \ g^{-1}(\xi, t) \in \mathfrak{S}_{r,a}(\xi, t)\}.$$

Define

$$\kappa^{\psi}_{r,a}(g) = \frac{1}{T}\int_0^T \int |\Gamma_{r,a}(\xi, t)|^{-1} 1_{\Gamma_{r,a}(\xi,t)}(g^{-1})\psi(\xi) \, d\nu(\xi)dt.$$

If a is sufficiently large then $\{\kappa^{\psi}_{r,a}\}_{r>0}$ satisfies the strong L^p maximal inequality for all $p > 1$ with $\frac{1}{p} + \frac{1}{q} \leq 1$. Moreover, if $\psi \in L^\infty(\partial\Gamma, \nu)$, then $\{\kappa^{\psi}_{r,a}\}_{r>0}$ satisfies the $L \log L$ maximal inequality.

Proof. – This follows from the previous corollary, Lemma 3.5 and Theorem 2.2. □

4.2. Bounding the ball averages

We now turn to show that when choosing $\psi = 1$, the integral $\kappa^1_{r,a}$ of the averaging sets (defined above) dominates the uniform measure β_r on the group. A preliminary step is to show that $\kappa^1_{r,a}$ dominates the measure uniformly distributed on a spherical shell $S_{r,a} = S_{r,a}(e)$.

To that end, for each $r, b > 0$, let $\zeta_{r,b}$ denote the probability measure on Γ which is distributed uniformly on the spherical shell $B(e, r - a/2 + b/2) \setminus B(e, r - a/2 - b/2)$. Then the following estimate holds.

PROPOSITION 4.7. – *For b sufficiently large (depending on (Γ, d)) and for a, T sufficiently large (depending on b and (Γ, d)) there is a constant $C > 0$ (which may depend on a, b, T) such that $\zeta_{r,b} \leq C\kappa^1_{r,a}$ for all $r > 0$.*

To prove Proposition 4.7, we need the following geometric lemmas.

LEMMA 4.8. – *There is a constant $C > 0$ such that for any $g \in \Gamma$ there exists $\eta \in \partial\Gamma$ with $|h_\eta(g)| \leq C$.*

Proof. – We have:

CLAIM 1. – *There is a constant* $C_0 > 0$ *such that for any* $g \in \Gamma$ *and* $r > 0$

$$\nu(\{\xi \in \partial\Gamma : (\xi|g)_e > r\}) \leq C_0 e^{-\mathfrak{v}r}.$$

Proof of Claim 1. – If $\xi, \eta \in \partial\Gamma$ satisfy $(\xi|g)_e, (\eta|g)_e > r$ then by (1.1),

$$(\xi|\eta)_e \geq \min\{(\xi|g)_e, (\eta|g)_e\} - \delta > r - \delta.$$

The claim now follows from Lemma 3.3. □

CLAIM 2. – *There is a constant* $R > 0$ *such that for any* $x, y, z \in \Gamma$ *there exists* $\eta \in \partial\Gamma$ *such that* $(\eta|x)_z, (\eta|y)_z \leq R$.

Proof of Claim 2. – It follows from Claim 1 that there is an $R > 0$ (independent of x, y, z) such that

$$\nu(\{\xi \in \partial\Gamma : (\xi|z^{-1}x)_e > R\}) + \nu(\{\xi \in \partial\Gamma : (\xi|z^{-1}y)_e > R\}) < 1.$$

Therefore, there is an $\xi \in \partial\Gamma$ such that $(\xi|z^{-1}x)_e, (\xi|z^{-1}y)_e \leq R$. But

$$(\xi|z^{-1}x)_e = (z\xi|y)_z, (\xi|z^{-1}y)_e = (z\xi|y)_z.$$

So set $\eta = z\xi$. □

Now let $g \in \Gamma$. Because $|h_\xi(g)| \leq d(e,g)$ (for any $g \in \Gamma, \xi \in \partial\Gamma$) we may assume without loss of generality, that $d(e,g) > c$ where $c > 0$ is the quasi-geodesicity constant. Let $\gamma : I \to \Gamma$ be a $(1,c)$-quasi-geodesic from e to g (where $I = [0,r]$ is some interval in the real line). Let $z = \gamma(d(e,g)/2)$. By Claim 2, there exist a constant $R > 0$ and $\eta \in \partial\Gamma$ such that $(\eta|g)_z, (\eta|e)_z \leq R$. Using (3.1), 3.2 we obtain:

$$2R \geq 2(\eta|g)_z = h_\eta(z) + d(g,z) - h_\eta(g) + O(\delta)$$
$$2R \geq 2(\eta|e)_z = h_\eta(z) + d(e,z) + O(\delta).$$

Because $d(g,z) = d(e,z) + O(c)$,

$$2R \geq |2(\eta|g)_z - 2(\eta|e)_z| \geq |h_\eta(g)| + O(\delta + c).$$ □

LEMMA 4.9. – *There exists a constant* $T_1 > 0$ *such that for any* $T > T_1$ *and any* $g \in \Gamma$,

$$C^{-1}\exp(-\mathfrak{v}d(e,g)/2) \leq \nu(\{\xi \in \partial\Gamma : |h_\xi(g)| \leq T\}) \leq C\exp(-\mathfrak{v}d(e,g)/2)$$

for some constant $C > 0$ *that is independent of* g *(but may depend on* T*).*

Proof. – By Lemma 4.8, there exist a constant $C_0 > 0$ and $\xi \in \partial\Gamma$ satisfying $|h_\xi(g)| \leq C_0$. Let $T_1 = 10\delta + C_0$ and $T > T_1$. In order to prove a lower bound, by Lemma 3.3, it suffices to prove

$$\{\eta \in \partial\Gamma : (\xi|\eta)_e \geq d(e,g)/2 + 2T\} \subset \{\eta \in \partial\Gamma : |h_\eta(g)| \leq T\}).$$

So suppose $\eta \in \partial\Gamma$ satisfies $(\xi|\eta)_e \geq d(e,g)/2 + 2T$. It suffices to show $|h_\eta(g)| \leq T$. By equation (3.3),

$$(\xi|g)_e + \delta \leq d(e,g)/2 - h_\xi(g)/2 + \delta < d(e,g)/2 + 2T \leq (\xi|\eta)_e.$$

Thus Gromov's inequality (1.1) and (3.3) implies

$$(\xi|g)_e + \delta \geq \min\{(\xi|\eta)_e, (\eta|g)_e\} = (\eta|g)_e \geq d(e,g)/2 - h_\eta(g)/2 - 4\delta.$$

So

$$d(e,g)/2 - h_\xi(g)/2 + \delta \ge (\xi|g)_e + \delta \ge d(e,g)/2 - h_\eta(g)/2 - 4\delta$$

implies $h_\eta(g) \ge h_\xi(g) - 10\delta \ge -C_0 - 10\delta$.

Similarly,

$$(\eta|g)_e + \delta \le d(e,g)/2 - h_\eta(g)/2 + \delta \le d(e,g)/2 + C_0/2 + 6\delta < d(e,g)/2 + 2T \le (\eta|\xi)_e.$$

Thus Gromov's inequality (1.1) and (3.3) implies

$$(\eta|g)_e + \delta \ge \min\{(\eta|\xi)_e, (\xi|g)_e\} = (\xi|g)_e \ge d(e,g)/2 - h_\xi(g)/2 - 4\delta.$$

So

$$d(e,g)/2 - h_\eta(g)/2 + \delta \ge (\eta|g)_e + \delta \ge d(e,g)/2 - h_\xi(g)/2 - 4\delta$$

implies

$$h_\eta(g) \le h_\xi(g) + 10\delta \le C_0 + 10\delta.$$

Thus $|h_\eta(g)| \le C_0 + 10\delta \le T$ as required.

Note

$$(\xi|g)_e = (1/2)(d(e,g) - h_\xi(g)) + O(\delta) \le d(e,g)/2 + C_0/2 + O(\delta).$$

Suppose $\eta \in \partial\Gamma$ satisfies $(\xi|\eta)_e > \max\{d(e,g)/2, (\xi|g)_e + \delta\}$. Then

$$(\eta|g)_e \ge \min\{(\eta|\xi)_e, (\xi|g)_e\} - \delta = (\xi|g)_e - \delta.$$

Similarly,

$$(\xi|g)_e \ge \min\{(\xi|\eta)_e, (\eta|g)_e\} - \delta.$$

Since $(\xi|\eta)_e > (\xi|g)_e + \delta$, we must have $(\xi|g)_e \ge (\eta|g)_e - \delta$. Thus $(\xi|g)_e = (\eta|g)_e + O(\delta)$. Since $(\xi|g)_e = (1/2)(d(e,g) - h_\xi(g)) + O(\delta)$ (and a similar formula holds for η), this implies $h_\xi(g) = h_\eta(g) + O(\delta)$.

To obtain the upper bound, suppose that $\eta, \xi \in \partial\Gamma$ satisfy $|h_\xi(g)|, |h_\eta(g)| \le R$ for some constant $R > 0$. Then

$$(\xi|\eta)_e \ge \min\{(\xi|g)_e, (\eta|g)_e\} - \delta \ge d(e,g)/2 - R/2 + O(\delta)$$

implies that for $R > 0$ large enough

$$\{\eta' \in \partial\Gamma : \ (\xi|\eta')_e \ge d(e,g)/2 - R\} \supset \{\eta' \in \partial\Gamma : \ |h_{\eta'}(g)| \le R\}).$$

Lemma 3.3 now implies the upper bound. □

Proof of Proposition 4.7. – Recall that by Lemma 3.3, $C^{-1}e^{\mathfrak{v}r} \le |B(e,r)| \le Ce^{\mathfrak{v}r}$ for all $r > 0$. Choose $b > \max\left\{a_0, \frac{1}{\mathfrak{v}}\log C\right\}$, $T_2 \ge \max(T_0, T_1)$, $T \ge 10T_2$ and $a > 2(b+T)$ where a_0, T_0 are as in Lemma 4.2 and T_1 is as in Lemma 4.9.

By definition, $\zeta_{r,b}$ is uniformly distributed on $B(e, r-a/2+b/2) \setminus B(e, r-a/2-b/2)$. It follows that

$$\begin{aligned}|B(e,r-a/2+b/2) \setminus B(e,r-a/2-b/2)| &\ge C^{-1}e^{\mathfrak{v}(r-a/2+b/2)} - Ce^{\mathfrak{v}(r-a/2-b/2)}\\ &= e^{\mathfrak{v}(r-a/2)}\left(C^{-1} - Ce^{-2\mathfrak{v}b}\right) \ge C(\mathfrak{v},b)e^{\mathfrak{v}(r-a/2)}.\end{aligned}$$

Therefore, for each $g \in \Gamma$,

$$\zeta_{r,b}(g) \le |B(e,r-a/2+b/2) \setminus B(e,r-a/2-b/2)|^{-1} \le C(\mathfrak{v},b)^{-1}e^{a/2}e^{-\mathfrak{v}r},$$

so that indeed $\zeta_{r,b}(g^{-1}) \le C' \exp(-\mathfrak{v}r)$ for some constant $C' > 0$ and all r sufficiently large, provided a and b satisfy the conditions above.

On the other hand, Lemma 4.2 implies $|\Gamma_{r,a}(\xi,t)| \le C\exp(\mathfrak{v}r/2)$ for some $C > 0$ (and every $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$). So,

$$\begin{aligned}\kappa^1_{r,a}(g) &= \frac{1}{T}\int_0^T \int |\Gamma_{r,a}(\xi,t)|^{-1} 1_{\Gamma_{r,a}(\xi,t)}(g^{-1})\, d\nu(\xi)dt\\ &\ge C^{-1}\exp(-\mathfrak{v}r/2)\frac{1}{T}\int_0^T \nu(\{\xi \in \partial\Gamma:\ g^{-1} \in \Gamma_{r,a}(\xi,t)\})\, dt.\end{aligned}$$

Definition 4.1 implies that $g \in \Gamma_{r,a}(\xi,t)$ if and only if:

$$r - a < d(e,g) - h_\xi(g) - t \le r, \quad t - R_\lambda(g^{-1},\xi) \in [0,T)_{\mathbb{L}}.$$

Because ν is quasi-conformal, there is a constant $\rho \ge 1$ such that

$$|R_\lambda(g^{-1},\xi)| \le \rho|h_\xi(g)| + \rho$$

for every $\xi \in \partial\Gamma, g \in \Gamma$. By choosing T larger if necessary, we may assume $T > 2\rho T_2 + 2\rho$.

Now suppose $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$, $g \in B(e, r - a/2 + b/2) \setminus B(e, r - a/2 - b/2)$, $\rho T_2 + \rho \le t < T - \rho T_2 - \rho$ and $|h_\xi(g)| \le T_2$. Then

$$\begin{aligned}r - a < r - a/2 - b/2 - T &\le r - a/2 - b/2 - T_2 - (T - \rho T_2 - \rho)\\ &\le d(e,g) - h_\xi(g) - t \le r - a/2 + b/2 + T_2 \le r.\end{aligned}$$

Also

$$|R_\lambda(g^{-1},\xi)| \le \rho T_2 + \rho \Rightarrow t + h_\xi(g) \in [0,T)_{\mathbb{L}}.$$

So $g \in \Gamma_{r,a}(\xi,t)$. Lemma 4.9 now implies

$$\begin{aligned}\kappa^1_{r,a}(g) &\ge C^{-1}\exp(-\mathfrak{v}r/2)\frac{1}{T}\int_0^T \nu(\{\xi \in \partial\Gamma:\ g^{-1} \in \Gamma_{r,a}(\xi,t)\})\, dt\\ &\ge C^{-1}\exp(-\mathfrak{v}r/2)\left(\frac{T - 2\rho T_2 - 2\rho}{T}\right)\nu(\{\xi \in \partial\Gamma:\ |h_\xi(g)| \le T_2\})\\ &\ge C^{-3}\exp(-\mathfrak{v}r)\end{aligned}$$

for some (possibly larger) constant $C > 0$. So $\zeta_{r,b}(g) \le C^4\kappa^1_{r,a}(g)$ as required. □

4.3. Proofs of Theorem 1.1 and Theorem 1.2

Proof of Theorem 1.1. – It follows from Proposition 4.7 that if $\Gamma \curvearrowright (X,m)$ is any probability-measure-preserving action and $f \in L^p(X,m)$ is nonnegative then $\zeta_{r,b}(f) \le C\kappa^1_{r,a}(f)$. By Corollary 4.6 there exist constants $C_p > 0$ ($p \ge 1$) such that $\|\mathbb{M}_\zeta[f]\|_p \le C\|\mathbb{M}_\kappa[f]\|_p \le CC_p\|f\|_p$ if $p > 1$ where

$$\mathbb{M}_\zeta[f] = \sup_{r>0}\zeta_{r,b}(|f|), \quad \mathbb{M}_\kappa[f] = \sup_{r>0}\kappa^1_{r,a}(|f|).$$

Similarly, if $f \in L\log L(X,m)$ then $\|\mathbb{M}_\zeta[f]\|_1 \le CC_1\|f\|_{L\log L}$. Now let β_r be the probability measure on Γ uniformly distributed over the ball of radius r centered at the identity. Let $\mathbb{M}_\beta[f] = \sup_{r>0}\beta_r(|f|)$. Because β_r can be represented as a convex linear combination of probability measures $\{\zeta_{t,b}\}_{t>0}$, it follows that $\mathbb{M}_\beta[f] \le \mathbb{M}_\zeta[f]$. Thus $\mathbb{M}_\beta[f] \le CC_p\|f\|_p$ if $f \in L^p(X,m)$ and $\mathbb{M}_\beta[f] \le CC_1\|f\|_1$ if $f \in L\log L(X,m)$ as claimed.

As to the averages $\sigma_{r,a}$, recall that by Lemma 3.3,

$$C^{-1}e^{\mathfrak{v}r} \leq |B(e,r)| \leq Ce^{\mathfrak{v}r}\ ,\ r>0.$$

So

$$\begin{aligned}|B(e,r+a)\setminus B(e,r-a)| &\geq C^{-1}e^{\mathfrak{v}(r+a)} - Ce^{\mathfrak{v}(r-a)}\\ &= \left(C^{-1}-Ce^{-2\mathfrak{v}a}\right)e^{\mathfrak{v}(r+a)} = C(\mathfrak{v},a)e^{\mathfrak{v}(r+a)}.\end{aligned}$$

Hence choosing a sufficiently large, $\sigma_{r,a} \leq C'(\mathfrak{v},a)\beta_{r+a}$ for some constant $C'(\mathfrak{v},a)$ as probability measures on Γ, and hence the maximal inequalities for β_r imply the maximal inequalities for $\sigma_{r,a}$. Since $\mu_{r,a}$ are convex combinations of $\sigma_{r,a}$, the strong maximal inequalities in L^p, $p>1$ and in $L\log L$ hold for them as well. □

We now consider a finite symmetric generating set $S\subset\Gamma$ and its associated word-metric. Recall that σ_n denotes the uniform probability measure on the sphere $S_n = S_n(e) = \{g\in\Gamma : |g| = n\}$ and β_n denotes the uniform probability measure on the ball $B_n = B_n(e) = \{g\in\Gamma : |g|\leq n\}$.

Before proceeding with the proof of Theorem 1.2, we claim that for word metrics Lemma 3.3 can be given a sharper form, as follows.

PROPOSITION 4.10. – *In a non-elementary word hyperbolic group* (Γ, S), *word metric spheres satisfy, for some constant* $C_0\geq 1$

$$C_0^{-1}e^{\mathfrak{v}n} \leq |S_n(e)| \leq C_0e^{\mathfrak{v}n}\ ,\ n\in\mathbb{N}. \tag{4.1}$$

REMARK 4.11. – Let us note that the estimate $C^{-1}e^{\mathfrak{v}n}\leq|B_n(e)|\leq Ce^{\mathfrak{v}n}$, $n\in\mathbb{N}$, for the growth of balls of geodesic hyperbolic metrics was established by Coornaert in [29, Thm. 7.2]. In fact, the discussion there establishes a stronger result, namely that there exists a fixed constant $c\geq 0$ such that annuli of fixed width c satisfy $C^{-1}e^{\mathfrak{v}n}\leq|S_{n,c}(e)|\leq Ce^{\mathfrak{v}n}$, $n\in\mathbb{N}$ (see the proof of Prop. 6.4 for the upper bound, and the proof of Théorème 7.2 for the lower bound). Curiously, in [25, Thm. 4.12], [12, Thm. 4.11] and [45, §3, Proof of Lem. 1] Coornaert's theorem is quoted as applying to the spheres S_n and not only to the balls B_n. In [24, Proof of Lemma 3.4.2] the number of elements in word-metric spheres is stated as $p(n)q^n + O(q_1^n)$ with $q_1 < q$ and $p(n)$ a polynomial, but this statement fails for certain hyperbolic groups, for example $\mathbb{Z}_a * \mathbb{Z}_b$, with $S = (\mathbb{Z}_a\cup\mathbb{Z}_b)\setminus\{e\}$, and $a>b>2$, see Remark 4.12 below.

We are not aware of a general proof of the validity of (4.1) for general hyperbolic groups. This estimate does hold for word metric spheres, but its proof requires an additional argument. Given the situation described in the foregoing remark, it may be useful to explain this argument briefly.

Proof of Proposition 4.10. – It was proved by Cannon that fixing a total order on the generating set S, the language $\mathscr{L}$ of lexicographically first geodesics is a prefix-closed regular language. Thus the number of elements of Γ of word-length n is the same as the number L_n of words of length n (with distinguished first letter) in a prefix-closed regular language (see [25] and [24] for a complete discussion). According to [31, Theorem V3], this implies the existence of an integer $D\geq 1$ such that for every $0\leq r<D$, one of the following alternatives holds.

Either for the residue class r modulo D, the sequence $\{L_{kD+r}\}_{k\in\mathbb{N}}$ is eventually zero, or the sequence is given by the following expression, for all $n \geq n_0$

$$L_n = P_r(n)\lambda_r^n + \sum_{j=1}^{N_r} p_{j,r}(n)\lambda_{j,r}^n \ , \ \ n \equiv r \ (\mathrm{mod}\ D)$$

with P_r a non-zero polynomial, $\lambda_r > |\lambda_{j,r}|$, and $p_{j,r}$ polynomials, $1 \leq j \leq N_r$. Thus when the sequence is not eventually zero, it satisfies the estimate

$$L_n = P_r(n)\lambda_r^n + O\left(\xi_r^n\right) \text{ with } 0 \leq \xi_r < \lambda_r \ , \ \ n \equiv r \ (\mathrm{mod}\ D).$$

Since Γ is non-elementary, and therefore infinite, when $L_n = |S_n|$ is the number of elements in the sphere of radius n, it is not eventually zero along any subseqeunce, so each of the exponents λ_r is strictly positive, $0 \leq r < D$. Furthermore, all the exponents λ_r are equal to one another in this case. Indeed, for all $n \in \mathbb{N}$ we have $S_{n+1} \subset S_n \cdot S$ so that the ratio between $|S_{n+1}|$ and $|S_n|$ is bounded by $|S|$, and similarly the ratio between $|S_{n+r}|$ and $|S_n|$ is bounded by $|S^r|$. Hence λ_r is constant for $0 \leq r < D$, and then looking at the same ratios again, we can conclude that all the polynomials $P_r(n)$ have the same degree, which we denote by m. Thus we can conclude that for some $C_1 > 1$ we have the estimate $C_1^{-1}n^m\lambda^n \leq |S_n| \leq C_1 n^m \lambda^n$ for every $n \in \mathbb{N}$. Let us denote $\log\lambda = \mathfrak{v}_1$. Since $|S_n| \leq |B_n| \leq Ce^{\mathfrak{v}n}$, we have $\mathfrak{v}_1 \leq \mathfrak{v}$. Conversely, since $|B_n| = \sum_{k=0}^n |S_k|$ we have

$$C^{-1}e^{\mathfrak{v}n} \leq |B_n| \leq C_1 \sum_{k=0}^{n} k^m e^{\mathfrak{v}_1 k} \leq C_2 n^m e^{\mathfrak{v}_1 n}$$

so that $\mathfrak{v}_1 \geq \mathfrak{v}$. We note in passing that the foregoing proof of the equality $\mathfrak{v}_1 = \mathfrak{v}$ does not really require Coornaert's result on the growth of balls, and follows simply from $\mathfrak{v} = \lim_{n\to\infty} \frac{1}{n}\log|B_n|$. However, our final argument, namely the fact that the degree m of the polynomials in question is in fact zero, does depend on the upper bound in Coornaert's result. Indeed, since $\mathfrak{v}_1 = \mathfrak{v}$, the fact that $m = 0$ follows from the inequality

$$Ce^{\mathfrak{v}n} \geq |B_n| \geq |S_n| \geq C_1^{-1}n^m e^{\mathfrak{v}n}.$$

The final result on the growth function of the spheres, then, is that there exist D positive constants $s_0, \dots s_{D-1}$ so that

$$|S_n| = s_r e^{\mathfrak{v}n} + O\left(e^{\mathfrak{k}n}\right) \text{ with } \mathfrak{k} < \mathfrak{v} \ , \ \ n \equiv r \ (\mathrm{mod}\ D).$$

This completes the proof of estimate (4.1) and Proposition 4.10. □

REMARK 4.12. –
1. We note that it is possible that the constants s_r appearing above will be different from one another. Indeed, this is the case in the example noted above, namely $\mathbb{Z}_a * \mathbb{Z}_b$, with $S = (\mathbb{Z}_a \cup \mathbb{Z}_b) \setminus \{e\}$, and $a > b > 2$. Here $D = 2$, $\lambda = \sqrt{(a-1)(b-1)}$, and $|S_{2k}| = 2[(a-1)(b-1)]^k$, $|S_{2k+1}| = (a+b-2)[(a-1)(b-1)]^k$.
2. For general hyperbolic metrics, it is not known whether the sequence $\left\{|S_n|\, e^{-\mathfrak{v}n}\right\}_{n\in\mathbb{N}}$ (where $\mathfrak{v}$ is the exponential rate of growth of the balls) has finitely many accumulation points or not. We thank Koji Fujiwara for bringing this problem to our attention.

Proof of Theorem 1.2. – It follows immediately from (4.1) that $\sigma_n \leq C'\beta_n$ as probability measures on the group Γ, and so $\{\sigma_n\}_{n=1}^\infty$ satisfies all the maximal inequalities satisfied by $\{\beta_n\}_{n=1}^\infty$. Since μ_n is a convex combination of the spheres σ_k, $0 \leq k \leq n$, it follows that $\{\mu_n\}_{n=1}^{\infty}$ satisfies the strong maximal inequalities satisfied by $\{\sigma_n\}_{n=1}^{\infty}$, namely in L^p, $p > 1$ and in $L \log L$. Pointwise almost sure convergence of $\pi_X(\mu_n)f$ for bounded functions $f \in L^\infty(X, m)$ has been established recently in [22, Cor. 1]. Because $L^\infty(X, m)$ is norm-dense in $L^p(X, m)$, standard arguments using the maximal inequality imply that $\pi_X(\mu_n)f$ converges almost surely for every $f \in L^p$, $1 < p \leq \infty$ and in $L \log L$. Finally, given pointwise convergence for L^p-functions, as well as the maximal inequality, norm convergence in L^p, $1 \leq p < \infty$ (and in $L \log L$), is a straightforward consequence of Lebesgue's dominated convergence theorem. This completes the proof of Theorem 1.2. □

5. Asymptotic invariance

In order to prove Theorem 1.3, we assume for the rest of the paper that the horofunction boundary coincides with the Gromov boundary of (Γ, d). This means that whenever $\{g_i\}_{i=1}^{\infty} \subset \Gamma$ converges to a point $\xi \in \partial\Gamma$ then the horofunction

$$h_\xi(g) := \lim_{i\to\infty} d(g_i, g) - d(g_i, e)$$

is well-defined. In particular, it depends on $\{g_i\}_{i=1}^{\infty}$ only through ξ.

Define $\mathfrak{B}_r$ and $\mathfrak{S}_{r,a}$ as in Definition 4.1 for some $T \in \mathbb{L}^+$ and $a > 0$ with $a \geq a_0, T \geq T_0$ (where a_0, T_0 are as in Lemma 4.2). Let $\Gamma \curvearrowright (\partial\Gamma \times \mathbb{L}, \nu \times \theta_\lambda)$ be the Maharam extension and let E be the induced equivalence relation on $\partial\Gamma \times [0, T)_{\mathbb{L}}$ (notational conventions are explained in §3.3). So $(\xi, t)E(\xi', t') \Leftrightarrow \exists g \in \Gamma$ such that $g(\xi, t) = (\xi', t')$.

Most of the work in proving Theorem 1.3 boils down to the next result the proof of which is the goal of this section.

Theorem 5.1. – *Let $a \geq a_0, T \geq T_0$ where a_0, T_0 are as in Lemma 4.2. For every $\epsilon_0, r > 0$ and $(\xi, t) \in \partial\Gamma \times [0, T]$ there exists $0 \leq \epsilon(\xi, t, r) < \epsilon_0$ such that if $\widetilde{\mathfrak{B}}_r(\xi, t) := \mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi, t)$ and $\widetilde{\mathfrak{S}}_{r,a} := \widetilde{\mathfrak{B}}_r(\xi, t) \setminus \widetilde{\mathfrak{B}}_{r-a}(\xi, t)$ then $\widetilde{\mathscr{B}} := \{\widetilde{\mathfrak{B}}_r\}_{r>0}$ and $\widetilde{\mathscr{S}}_a := \{\widetilde{\mathfrak{S}}_{r,a}\}_{r>0}$ are asymptotically invariant.*

5.1. The leafwise metric on the equivalence classes

To begin the proof, we need a leafwise metric on E: given $(\xi, t), (\xi', t') \in \partial\Gamma \times [0, T)_{\mathbb{L}}$ with $(\xi, t)E(\xi', t')$, let $d_\Gamma((\xi, t), (\xi', t'))$ be the minimum value of $d(g, e)$ over all $g \in \Gamma$ with $g(\xi, t) = (\xi', t')$. Most of the work in showing Theorem 5.1 boils down to the next two propositions.

Proposition 5.1. – *For $(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}$ and $r > 0$ let $\mathscr{N}_n(\mathfrak{B}_r(\xi, t))$ denote the radius-n neighborhood of $\mathfrak{B}_r(\xi, t)$ with respect to d_Γ. Then for any $n > 0$,*

$$\limsup_{\delta\to 0^+} \limsup_{r\to\infty} \frac{|\mathscr{N}_n(\mathfrak{B}_r(\xi, t))|}{|\mathfrak{B}_{r+\delta}(\xi, t)|} \leq 1.$$

Similarly,

$$\limsup_{\delta\to 0^+} \limsup_{r\to\infty} \frac{|\mathscr{N}_n(\mathfrak{S}_{r,a}(\xi, t))|}{|\mathfrak{B}_{r+\delta}(\xi, t)| - |\mathfrak{B}_{r-a-\delta}(\xi, t)|} \leq 1.$$

PROPOSITION 5.2. – *For any* $\epsilon_0, r > 0$ *and a.e.* $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$ *there exists* $0 \le \epsilon(\xi,t,r) < \epsilon_0$ *such that*

$$1 = \lim_{\delta\to 0^+} \limsup_{r\to\infty} \frac{|\mathfrak{B}_{r+\epsilon(\xi,t,r)+\delta}(\xi,t)|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|} = \lim_{\delta\to 0^-} \liminf_{r\to\infty} \frac{|\mathfrak{B}_{r+\epsilon(\xi,t,r)+\delta}(\xi,t)|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|}.$$

LEMMA 5.3. – *There exists a countable set* $\Phi \subset [E]$ *such that* Φ *generates* E *and for every* $\phi \in \Phi$ *there exists an* $n = n(\phi) > 0$ *such that* $d_\Gamma((\xi,t),\phi(\xi,t)) \le n$ *for a.e.* (ξ,t).

Proof. – For $n > 0$ let $G_n = \{((\xi,t),(\xi',t')) \in E : d_\Gamma((\xi,t),(\xi',t')) \le n\}$. Because d is locally finite, $(\partial\Gamma \times [0,T)_{\mathbb{L}}, G_n)$ is a bounded degree graph. By [38], this implies that the Borel edge-chromatic number of $(\partial\Gamma \times [0,T)_{\mathbb{L}}, G_n)$ is finite. That is, there exists a Borel map $\Omega_n : G_n \to A_n$ (where A_n is a finite set) such that if $((\xi,t),(\xi',t')), ((\xi',t'),(\xi'',t'')) \in G_n$ and $(\xi,t) \ne (\xi'',t'')$ then $\Omega_n((\xi,t),(\xi',t')) \ne \Omega_n((\xi',t'),(\xi'',t''))$. We can also assume without loss of generality that $\Omega_n((\xi,t),(\xi',t')) = \Omega_n((\xi',t'),(\xi,t))$.

For each element $a \in A_n$, define $\phi_a : \partial\Gamma \times [0,T)_{\mathbb{L}} \to \partial\Gamma \times [0,T)_{\mathbb{L}}$ as follows. If $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$ and there is a (ξ',t') such that $\Omega_n((\xi,t),(\xi',t')) = a$ then let $\phi_a(\xi,t) := (\xi',t')$. Otherwise let $\phi_a(\xi,t) := (\xi,t)$. Then $\phi_a \in [E]$. Moreover, if $((\xi,t),(\xi',t')) \in G_n$ then there is some $a \in A_n$ such that $\phi_a(\xi,t) = (\xi',t')$. Since $\bigcup_{n=1}^\infty G_n = E$, we have that $\Phi := \bigcup_{n=1}^\infty \{\phi_a : a \in A_n\}$ is generating. □

Proof of Theorem 5.1 given Propositions 5.1, 5.2. – Let Φ be the generating set from the previous lemma, $\phi \in \Phi$, $n = n(\phi)$ and $\epsilon(\xi,t,r)$ be as in Proposition 5.2. Then for any $(\xi,t) \in \partial\Gamma \times [0,T)_{\mathbb{L}}$,

$$\begin{aligned}\limsup_{r\to\infty} \frac{|\widetilde{\mathfrak{B}}_r(\xi,t) \vartriangle \phi(\widetilde{\mathfrak{B}}_r(\xi,t))|}{|\widetilde{\mathfrak{B}}_r(\xi,t)|} &\le \limsup_{r\to\infty} 2\frac{|\mathcal{N}_n(\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)) \setminus \mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|} \\ &= 2\left(\limsup_{r\to\infty} \frac{|\mathcal{N}_n(\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t))|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|} - 1\right) \\ &\le 2\left(\limsup_{\delta\searrow 0} \limsup_{r\to\infty} \frac{|\mathfrak{B}_{r+\epsilon(\xi,t,r)+\delta}(\xi,t)|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|} - 1\right) = 0.\end{aligned}$$

The last inequality is justified by Proposition 5.1 and the last equality follows from Proposition 5.2. Because $\phi \in \Phi$ is arbitrary, this proves $\widetilde{\mathscr{B}}$ is asymptotically invariant.

Recall that $\widetilde{\mathfrak{S}}_{r,a}(\xi,t) = \widetilde{\mathfrak{B}}_r(\xi,t) \setminus \widetilde{\mathfrak{B}}_{r-a}(\xi,t)$. By Lemma 4.2,

$$\begin{aligned}&\lim_{r\to\infty} \frac{|\widetilde{\mathfrak{S}}_{r,a}(\xi,t) \vartriangle \phi(\widetilde{\mathfrak{S}}_{r,a}(\xi,t))|}{|\widetilde{\mathfrak{S}}_{r,a}(\xi,t)|} \\ &\quad\le \lim_{r\to\infty} \frac{|\widetilde{\mathfrak{B}}_r(\xi,t) \vartriangle \phi(\widetilde{\mathfrak{B}}_r(\xi,t))| + |\widetilde{\mathfrak{B}}_{r-a}(\xi,t) \vartriangle \phi(\widetilde{\mathfrak{B}}_{r-a}(\xi,t))|}{|\widetilde{\mathfrak{B}}_r(\xi,t)|} \frac{|\widetilde{\mathfrak{B}}_r(\xi,t)|}{|\widetilde{\mathfrak{S}}_{r,a}(\xi,t)|} \\ &\quad\le C^2 \lim_{r\to\infty} \frac{|\widetilde{\mathfrak{B}}_r(\xi,t) \vartriangle \phi(\widetilde{\mathfrak{B}}_r(\xi,t))| + |\widetilde{\mathfrak{B}}_{r-a}(\xi,t) \vartriangle \phi(\widetilde{\mathfrak{B}}_{r-a}(\xi,t))|}{|\widetilde{\mathfrak{B}}_r(\xi,t)|} = 0.\end{aligned}$$

Since $\phi \in \Phi$ is arbitrary, this implies $\widetilde{\mathscr{S}}_a$ is asymptotically invariant. □

5.2. The key geometric argument

This section proves Proposition 5.1. We need a few geometric lemmas to begin.

LEMMA 5.4. – *Suppose* $\xi_1, \xi_2 \in \partial\Gamma$ *and* $\xi_1 \neq \xi_2$. *Then for any* $r \in \mathbb{R}$ *there are sets* $V_1 \subset \partial\Gamma$, $V_2 \subset \overline{\Gamma}$ $(= \Gamma \cup \partial\Gamma)$ *such that*

- $\xi_1 \in V_1, \xi_2 \in V_2$,
- V_1 *is open in* $\partial\Gamma$, V_2 *is open in* $\overline{\Gamma}$,
- $V_1 \cap V_2 = \varnothing$,
- $\forall v_2 \in V_2 \cap \Gamma,\ \forall \eta \in V_1,\ h_\eta(v_2) \geq r$.

Proof. – Let V_1 be an open neighborhood of ξ_1 whose closure does not contain ξ_2. Let $\{W_n\}_{n=1}^\infty$ be any sequence of decreasing open subsets of $\overline{\Gamma}$ such that $\cap_n W_n = \{\xi_2\}$ and $W_n \cap V_1 = \varnothing$. If the lemma is false then for each n we can find an $x_n \in W_n$ and an $\xi_n \in V_1$ such that $h_{\xi_n}(x_n) < r$. Observe

$$\lim_{n\to\infty} 2(\xi_n|x_n)_e = \lim_{n\to\infty} d(x_n, e) - h_{\xi_n}(x_n) = +\infty.$$

So if (ξ_∞, x_∞) is a limit point of $\{(\xi_n, x_n)\}_{n=1}^\infty$ in $\overline{\Gamma} \times \overline{\Gamma}$ then by equation (3.1), $(\xi_\infty|x_\infty)_e = +\infty$. In particular, every limit point of $\{x_n\}_{n=1}^\infty$ is contained in the closure of V_1. But the hypotheses on W_n imply $\lim_{n\to\infty} x_n = \xi_2$, a contradiction. □

LEMMA 5.5. – *Suppose that* $\{\xi_i\}_{i=1}^\infty \subset \partial\Gamma$ *and* $\lim_{i\to\infty} \xi_i = \xi_\infty$. *Fix* $C \in \mathbb{R}$ *and let* $H_i := \{x \in \Gamma :\ h_{\xi_i}(x) \leq C\}$ *(for* $1 \leq i \leq \infty$*). If* $x_i \in H_i$ *for all* i *then every limit point* y *of* $\{x_i\}_{i=1}^\infty$ *in* $\overline{\Gamma}$ *satisfies* $y \in H_\infty \cup \{\xi_\infty\}$.

Proof. – Without loss of generality we may assume $\{x_i\}_{i=1}^\infty$ converges in $\overline{\Gamma}$ to an element y. If $y \in \Gamma$ then $x_i = y$ for all i sufficiently large (since Γ is locally finite). So $h_{\xi_\infty}(y) = \lim_{i\to\infty} h_{\xi_i}(x_i) \leq C$ and $y \in H_\infty$.

To obtain a contradiction, suppose that $y \in \partial\Gamma$ but $y \neq \xi_\infty$. The previous lemma implies the existence of sets $V_\infty \subset \partial\Gamma$, $V_y \subset \overline{\Gamma}$ such that

- $\xi_\infty \in V_\infty, y \in V_y$,
- V_∞ is open in $\partial\Gamma$, V_y is open in $\overline{\Gamma}$,
- $V_\infty \cap V_y = \varnothing$,
- $\forall g \in V_y \cap \Gamma,\ \forall \eta \in V_\infty,\ h_\eta(g) \geq C + 1$.

For all n sufficiently large, $\xi_n \in V_\infty$. Therefore H_n has trivial intersection with V_y. Since $x_n \in H_n$, this implies $\lim_{n\to\infty} x_n \notin V_y$. But $\lim_{n\to\infty} x_n = y \in V_y$. This contradiction implies that if $\lim_{n\to\infty} x_n \in \partial\Gamma$ then $\lim_{n\to\infty} x_n = \xi_\infty$ as required. □

LEMMA 5.6. – *There exists a function* $\beta = \beta(r, n, t) \geq 0$ *such that if* $g, g' \in \Gamma$, $\xi \in \partial\Gamma$ *and*

$$d(g, e) \geq r, \quad d(g, g') \leq n, \quad |h_\xi(g)| \leq t$$

then

$$|(d(g, e) - h_\xi(g)) - (d(g', e) - h_\xi(g'))| \leq \beta(r, n, t).$$

Moreover, $\lim_{r\to\infty} \beta(r, n, t) = 0$ *for any* $n, t > 0$.

Proof. – The proof is by contradiction. Assuming no such function exists, there are constants $n, t, \epsilon_0 > 0$, elements $\xi_r \in \partial\Gamma$ and elements $g_r, g'_r \in \Gamma$ $(\forall r > 0)$ such that

– $d(g_r, e) \geq r$, $d(g_r, g'_r) \leq n$, $|h_{\xi_r}(g_r)| \leq t$,
– $|(d(g_r, e) - h_{\xi_r}(g_r)) - (d(g'_r, e) - h_{\xi_r}(g'_r))| \geq \epsilon_0$.

After passing to a subsequence if necessary, we may assume that the sequence $\{g_r^{-1}\xi_r\}_{r=1}^{\infty}$ converges to an element $\xi^* \in \partial\Gamma$. We claim $\{g_r^{-1}\}_{r=1}^{\infty}$ also converges to ξ^*. To see this, observe that for any $x, g \in \Gamma$ and $\xi \in \partial\Gamma$,

$$h_{g\xi}(x) = h_\xi(g^{-1}x) - h_\xi(g^{-1}).$$

Therefore,

$$h_{g_r^{-1}\xi_r}(g_r^{-1}) = h_{\xi_r}(e) - h_{\xi_r}(g_r) \leq t \; \forall r.$$

Since $\lim_{r\to\infty} d(g_r^{-1}, e) = +\infty$ the previous lemma implies $\lim_{r\to\infty} g_r^{-1} = \xi^*$ as claimed.

The claim implies that for any $x \in \Gamma$,

$$h_{\xi^*}(x) = \lim_{r\to\infty} d(g_r^{-1}, x) - d(g_r^{-1}, e) = \lim_{r\to\infty} d(e, g_r x) - d(e, g_r).$$

Since $d(g_r, g'_r) \leq n$ for all r and because (Γ, d) is locally finite we may assume after passing to a subsequence that there is a $y \in \Gamma$ such that $g_r^{-1}g'_r = y$ for all r. By setting $x = y$ in the equation above, we obtain:

$$\begin{aligned}\lim_{r\to\infty} d(e, g'_r) - d(e, g_r) = h_{\xi^*}(y) &= \lim_{r\to\infty} h_{g_r^{-1}\xi_r}(y) \\ &= \lim_{r\to\infty} h_{\xi_r}(g_r y) - h_{\xi_r}(g_r) = \lim_{r\to\infty} h_{\xi_r}(g'_r) - h_{\xi_r}(g_r).\end{aligned}$$

This contradicts the assumption $|(d(g_r, e) - h_{\xi_r}(g_r)) - (d(g'_r, e) - h_{\xi_r}(g'_r))| \geq \epsilon_0$. □

COROLLARY 5.7. – There is a constant $K > 0$ (depending only on $T, (\Gamma, d), \nu$) such that for any $r, n > 0$ and any $(\xi, t) \in \partial\Gamma \times [0, T)_{\mathbb{L}}$,

$$\mathcal{N}_n(\mathfrak{B}_r(\xi, t)) \subset \mathfrak{B}_{r+\beta(r-K,n,K)}(\xi, t).$$

Proof. – If $(\xi'', t'') \in \mathcal{N}_n(\mathfrak{B}_r(\xi, t)) \backslash \mathfrak{B}_r(\xi, t)$ then there exist $(\xi', t') \in \mathfrak{B}_r(\xi, t)$ and $g \in \Gamma$ such that $d(e, g) \leq n$ and $g(\xi', t') = (\xi'', t'')$. Since $(\xi', t') \in \mathfrak{B}_r(\xi, t)$, there is also a $\gamma \in \Gamma$ such that

$$\gamma^{-1}(\xi, t) = (\xi', t'), \quad d(\gamma, e) - h_\xi(\gamma) - t \leq r.$$

Because $t' = t - R_\lambda(\gamma^{-1}, \xi)$ and, by (3.4), $|R_\lambda(\gamma^{-1}, \xi) + \mathfrak{v}_\lambda h_\xi(\gamma)| < C$ (for some constant $C > 0$), we have

$$|h_\xi(\gamma)| \leq \mathfrak{v}_\lambda^{-1}(T + C) \Rightarrow d(\gamma, e) \leq r + T + \mathfrak{v}_\lambda^{-1}(T + C).$$

Let $f = \gamma g^{-1}$. So $d(f, \gamma) = d(g, e) \leq n$. Note $f^{-1}(\xi, t) = g\gamma^{-1}(\xi, t) = g(\xi', t') = (\xi'', t'')$. As above, this implies $|h_\xi(f)| \leq \mathfrak{v}_\lambda^{-1}(T + C)$. Since $(\xi'', t'') \notin \mathfrak{B}_r(\xi, t)$, $d(e, f) > r - T - \mathfrak{v}_\lambda^{-1}(T + C)$.

We now apply the previous lemma to f and γ to obtain

$$|d(e, f) - d(e, \gamma) - h_\xi(f) + h_\xi(\gamma)| \leq \beta(r - K, n, K)$$

where $K = T + \mathfrak{v}_\lambda^{-1}(T + C)$. Thus

$$|d(e, f) - h(f) - t| \leq |d(e, \gamma) - h_\xi(\gamma) - t| + \beta(r - K, n, K) \leq r + \beta(r - K, n, K).$$

This implies $(\xi'', t'') \in \mathfrak{B}_{r+\beta(r-K,n,K)}(\xi, t)$ as required. □

Proposition 5.1 follows from the corollary above and the fact that $\lim_{r\to\infty}\beta(r,n,t)=0 \quad \forall n,t$.

5.3. Proof of asymptotic invariance

In this section we prove Proposition 5.2 whose statement is recalled below.

PROPOSITION 5.2. – *For any* $\epsilon_0, r > 0$ *and a.e.* $(\xi,t)\in\partial\Gamma\times[0,T)_{\mathbb{L}}$ *there exists* $0\le\epsilon(\xi,t,r)<\epsilon_0$ *such that*

$$1=\lim_{\delta\to 0^+}\limsup_{r\to\infty}\frac{|\mathfrak{B}_{r+\epsilon(\xi,t,r)+\delta}(\xi,t)|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|}=\lim_{\delta\to 0^-}\liminf_{r\to\infty}\frac{|\mathfrak{B}_{r+\epsilon(\xi,t,r)+\delta}(\xi,t)|}{|\mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)|}.$$

Proof. – Let $(\xi,t)\in\partial\Gamma\times[0,T)_{\mathbb{L}}$, $0<a<b$, $l=b-a$ and $1\le N,m$ be integers such that N is divisible by 4. Suppose that for every $c\in[a+2l/N,b-2l/N]$,

$$\frac{|\mathfrak{B}_{c-2l/N}(\xi,t)|}{|\mathfrak{B}_{c+2l/N}(\xi,t)|}\le 1-1/m.$$

By Lemma 4.2,

$$\begin{aligned}Ce^{\mathfrak{v}b/2}\ge|\mathfrak{B}_b(\xi,t)|&=|\mathfrak{B}_a(\xi,t)|\prod_{j=0}^{(N/4-1)}\frac{|\mathfrak{B}_{a+(4j+4)l/N}(\xi,t)|}{|\mathfrak{B}_{a+4jl/N}(\xi,t)|}\\&\ge|\mathfrak{B}_a(\xi,t)|(1-1/m)^{-N/4}\ge C^{-1}e^{\mathfrak{v}a/2}(1-1/m)^{-N/4+3}.\end{aligned}$$

So $N\le\frac{-4\log(C^2e^{\mathfrak{v}l/2})-12\log(1-1/m)}{\log(1-1/m)}$.

Suppose now that $N>\frac{-4\log(C^2e^{\mathfrak{v}l/2})-12\log(1-1/m)}{\log(1-1/m)}$.

Then there exists $c\in[a+2l/N,b-2l/N]$ such that

$$\frac{|\mathfrak{B}_{c-2l/N}(\xi,t)|}{|\mathfrak{B}_{c+2l/N}(\xi,t)|}\ge 1-1/m.$$

For any $x\in[c-l/N,c+l/N]$,

$$\frac{|\mathfrak{B}_x(\xi,t)|}{|\mathfrak{B}_{x+l/N}(\xi,t)|}\ge\frac{|\mathfrak{B}_{c-2l/N}(\xi,t)|}{|\mathfrak{B}_{c+2l/N}(\xi,t)|}\ge 1-1/m$$

and

$$\frac{|\mathfrak{B}_{x-l/N}(\xi,t)|}{|\mathfrak{B}_x(\xi,t)|}\ge\frac{|\mathfrak{B}_{c-2l/N}(\xi,t)|}{|\mathfrak{B}_{c+2l/N}(\xi,t)|}\ge 1-1/m.$$

Now let $\epsilon_0>0$. By induction, for every $r>0$ and $j\ge 2$ there exist $c_{r,j},N_j>0$ such that:

- for every $x\in[c_{r,j}-1/N_j,c_{r,j}+1/N_j]$

$$\frac{|\mathfrak{B}_x(\xi,t)|}{|\mathfrak{B}_{x+1/N_j}(\xi,t)|}\ge 1-1/j,\text{ and }\frac{|\mathfrak{B}_{x-1/N_j}(\xi,t)|}{|\mathfrak{B}_x(\xi,t)|}\ge 1-1/j;$$

- $[c_{r,j+1}-1/N_{j+1},c_{r,j+1}+1/N_{j+1}]\subset[c_{r,j}-1/N_j,c_{r,j}+1/N_j]\subset[r,r+\epsilon_0]$.

Let $\delta_{r,\xi,t}$ be the only point in the nested intersection $\bigcap_j[c_{r,j}-1/N_j,c_{r,j}+1/N_j]$. Then $\epsilon(r,\xi,t):=\delta_{r,\xi,t}-r$ satisfies the conclusion of the proposition by construction. □

5.4. Proof of Theorem 1.3

In order to apply Theorems 2.2 and 2.3, we need to know the action $\Gamma \curvearrowright (\partial\Gamma, \nu)$ is weakly mixing, as well as its type and stable type. To prove this, we need the existence of a conformal measure on $\partial\Gamma$:

LEMMA 5.8. – *Let* (Γ, d) *be a non-elementary, uniformly quasi-geodesic, hyperbolic group whose horofunction boundary coincides with its Gromov boundary. Then there exists a conformal measure* ν_c *on* $\partial\Gamma$. *Thus* $\frac{d\nu_c \circ g}{d\nu_c}(\xi) = \exp(-\mathfrak{v} h_\xi(g^{-1}))$ *for a.e.* ξ *and every* $g \in \Gamma$.

Proof. – For $s > 0$ and $\gamma \in \Gamma$, let

$$Z_s(\gamma) := \sum_{g \in \Gamma} e^{-sd(\gamma, g)}.$$

By Lemma 3.3, there exist constants $C_0, a > 0$ so that if

$$N_k = |\{g \in \Gamma : \ ak \le d(e,g) < a(k+1)\}|$$

then

$$C_0^{-1} e^{\mathfrak{v} ak} \le N_k \le C_0 e^{\mathfrak{v} ak}.$$

So there is a constant $C_1 > 0$ such that

$$C_1^{-1} \sum_{k=0}^{\infty} N_k e^{-ska} \le Z_s(\gamma) \le C_1 \sum_{k=0}^{\infty} N_k e^{-ska}.$$

So if $s > \mathfrak{v}$,

$$\frac{C_0^{-1} C_1^{-1}}{1 - e^{a(\mathfrak{v} - s)}} \le Z_s(\gamma) \le \frac{C_0 C_1}{1 - e^{a(\mathfrak{v} - s)}}.$$

For $s > \mathfrak{v}$ let

$$m_s := \frac{1}{Z(s)} \sum_{g \in \Gamma} e^{-sd(e,g)} \delta_g$$

where δ_g is the Dirac measure concentrated on $\{g\} \subset \Gamma$. We consider these as measures on $\overline{\Gamma} = \Gamma \cup \partial\Gamma$. Let ν_c be any weak* limit of m_s as $s \searrow \mathfrak{v}$. Because $\lim_{s \searrow \mathfrak{v}} Z(s) = +\infty$, ν_c is supported on $\partial\Gamma$. An exercise left to the reader shows that ν_c is conformal as claimed.

□

REMARK 5.9. – If $\lambda \in (0,1)$ then we do not know whether there exists a conformal measure ν_c on $\partial\Gamma$ which in addition satisfies $\log_\lambda \left(\frac{d\nu_c \circ g}{d\nu_c}(\xi)\right) \in \mathbb{Z}$ for every $g \in \Gamma$ and a.e. $\xi \in \partial\Gamma$. However, Lemma 3.6 implies that ν_c is equivalent to a quasi-conformal measure ν which satisfies this condition.

LEMMA 5.10. – *Let* (Γ, d) *be a non-elementary uniformly quasi-geodesic hyperbolic group whose Gromov boundary coincides with its horofunction boundary and* ν *be a quasi-conformal measure on* $\partial\Gamma$. *Then* $\Gamma \curvearrowright (\partial\Gamma, \nu)$ *is weakly mixing, type* III_λ *and stable type* III_τ *for some* $\lambda, \tau \in (0,1]$.

Proof. – By Lemma 3.3 any two quasi-conformal measures on $\partial\Gamma$ are absolutely continuous to each other. So by Lemma 5.8 we may assume ν is conformal. So the Radon-Nikodym derivatives $\frac{d\nu\circ g}{d\nu}$ are continuous. By [27, Corollary 0.2], $\Gamma \curvearrowright (\partial\Gamma,\nu)$ is a factor of a Poisson boundary. By [1], the action of Γ on any of its Poisson boundaries is weakly mixing. This implies $\Gamma \curvearrowright (\partial\Gamma,\nu)$ is weakly mixing. The main theorem of [13] implies $\Gamma \curvearrowright (\partial\Gamma,\nu)$ is type III_λ and stable type III_τ for some $\lambda,\tau\in(0,1]$. □

We now combine Theorems 2.2, 2.3, Corollary 4.5 and Theorem 5.1 to prove Theorem 1.3.

Proof of Theorem 1.3. – Let $a,T \geq 0$ be sufficiently large so that the conclusions to Corollary 4.5 and Theorem 5.1 hold. Also let $\epsilon_0 > 0$.

We use notation as in §3.3. So let E be the equivalence relation on $\partial\Gamma\times[0,T)_{\mathbb{L}}$ induced by the partial action of Γ. Let $\epsilon(\xi,t,r)$, $\widetilde{\mathscr{S}_a} = \{\widetilde{\mathfrak{S}}_{r,a}\}_{r>0}$ be as in Theorem 5.1. By Corollary 4.5 and Theorem 5.1, $\widetilde{\mathscr{S}_a}$ is regular and asymptotically invariant.

Suppose that $\Gamma \curvearrowright (\partial\Gamma,\nu)$ has type III_1 (so $\mathbb{L}=\mathbb{R}$, $\lambda=1$). Let

$$\zeta_r(g) = T^{-1}\int_0^T\int |\{w\in\Gamma:\ w(\xi,t)\in\widetilde{\mathfrak{S}}_{r,a}(\xi,t)\}|^{-1}1_{\widetilde{\mathfrak{S}}_{r,a}(\xi,t)}(g^{-1}(\xi,t))\,d\nu(\xi)dt.$$

The previous lemma and Theorem 2.2 imply $\{\zeta_r\}$ is a pointwise ergodic family in $L\log L$. So it suffices to show that each ζ_r is supported on $B(e,r+\rho)\setminus B(e,r-\rho)$ where

$$\rho := \epsilon_0 + T + \mathfrak{v}_\lambda^{-1}(C+T) + a$$

is a constant independent of r.

We observe that ζ_r is supported on those $g\in\Gamma$ satisfying: there exists $(\xi,t)\in\partial\Gamma\times[0,T)_{\mathbb{L}}$ such that

$$g^{-1}(\xi,t)\in\widetilde{\mathfrak{S}}_{r,a}(\xi,t) = \mathfrak{B}_{r+\epsilon(\xi,t,r)}(\xi,t)\setminus\mathfrak{B}_{r-a+\epsilon(\xi,t,r-a)}(\xi,t).$$

Thus ζ_r is supported on

$$\Gamma_{r+\epsilon(\xi,t,r)}(\xi,t)\setminus\Gamma_{r-a+\epsilon(\xi,t,r-a)}(\xi,t)\subset\Gamma_{r+\epsilon_0}(\xi,t)\setminus\Gamma_{r-a}(\xi,t).$$

So let $g\in\Gamma_{r+\epsilon_0}(\xi,t)\setminus\Gamma_{r-a}(\xi,t)$. Then $g^{-1}(\xi,t)\in\partial\Gamma\times[0,T)_{\mathbb{L}}$. By definition (see §3.3)

$$g^{-1}(\xi,t) = (g^{-1}\xi, t - R_\lambda(g^{-1},\xi)).$$

So $|R_\lambda(g^{-1},\xi)|\leq T$. By (3.4), $|R_\lambda(g^{-1},\xi)+\mathfrak{v}_\lambda h_\xi(g)|\leq C$ where $C>0$ is a constant. Thus $|h_\xi(g)|\leq\mathfrak{v}_\lambda^{-1}(C+T)$. Since $g\in\Gamma_{r+\epsilon_0}(\xi,t)\setminus\Gamma_{r-a}(\xi,t)$,

$$r-a-\mathfrak{v}_\lambda^{-1}(C+T)\leq r-a+h_\xi(g) < d(e,g)\leq r+\epsilon_0+t+h_\xi(g)\leq r+\epsilon_0+T+\mathfrak{v}_\lambda^{-1}(C+T).$$

Thus $g\in B(e,r+\rho)\setminus B(e,r-\rho)$ as required.

This finishes the type III_1 case. The type III_λ case ($\lambda\in(0,1)$) is similar, using Theorem 2.3 instead of 2.2. □

BIBLIOGRAPHY

[1] J. Aaronson, M. Lemańczyk, Exactness of Rokhlin endomorphisms and weak mixing of Poisson boundaries, in *Algebraic and topological dynamics*, Contemp. Math. **385**, Amer. Math. Soc., Providence, RI, 2005, 77–87.

[2] S. Adams, Boundary amenability for word hyperbolic groups and an application to smooth dynamics of simple groups, *Topology* **33** (1994), 765–783.

[3] C. Anantharaman, J.-P. Anker, M. Babillot, A. Bonami, B. Demange, S. Grellier, F. Havard, P. Jaming, E. Lesigne, P. Maheux, J.-P. Otal, B. Schapira, J.-P. Schreiber, *Théorèmes ergodiques pour les actions de groupes*, Monographies de L'Enseignement Mathématique **41**, L'Enseignement Mathématique, Geneva, 2010.

[4] A. Ancona, Negatively curved manifolds, elliptic operators, and the Martin boundary, *Ann. of Math.* **125** (1987), 495–536.

[5] A. Ancona, Positive harmonic functions and hyperbolicity, in *Potential theory—surveys and problems (Prague, 1987)*, Lecture Notes in Math. **1344**, Springer, Berlin, 1988, 1–23.

[6] A. Ancona, Théorie du potentiel sur les graphes et les variétés, in *École d'été de Probabilités de Saint-Flour XVIII—1988*, Lecture Notes in Math. **1427**, Springer, Berlin, 1990, 1–112.

[7] V. I. Arnol′d, A. L. Krylov, Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain, *Dokl. Akad. Nauk SSSR* **148** (1963), 9–12.

[8] G. D. Birkhoff, Proof of the ergodic theorem, *Proc. Nat. Acad. Sci. USA* **17** (1931), 656–660.

[9] M. Björklund, Central limit theorems for Gromov hyperbolic groups, *J. Theoret. Probab.* **23** (2010), 871–887.

[10] S. Blachère, P. Haïssinsky, P. Mathieu, Harmonic measures versus quasiconformal measures for hyperbolic groups, *Ann. Sci. Éc. Norm. Supér.* **44** (2011), 683–721.

[11] O. V. Bogopol′skiĭ, V. N. Gerasimov, Finite subgroups of hyperbolic groups, *Algebra i Logika* **34** (1995), 619–622.

[12] L. Bowen, Invariant measures on the space of horofunctions of a word hyperbolic group, *Ergodic Theory Dynam. Systems* **30** (2010), 97–129.

[13] L. Bowen, The type and stable type of the boundary of a Gromov hyperbolic group, *Geom. Dedicata* **172** (2014), 363–386.

[14] L. Bowen, A. Nevo, Geometric covering arguments and ergodic theorems for free groups, *Enseign. Math.* **59** (2013), 133–164.

[15] L. Bowen, A. Nevo, Pointwise ergodic theorems beyond amenable groups, *Ergodic Theory Dynam. Systems* **33** (2013), 777–820.

[16] L. Bowen, A. Nevo, Amenable equivalence relations and the construction of ergodic averages for group actions, *J. Anal. Math.* **126** (2015), 359–388.

[17] N. Brady, Finite subgroups of hyperbolic groups, *Internat. J. Algebra Comput.* **10** (2000), 399–405.

[18] E. Breuillard, Geometry of locally compact groups of polynomial growth and shape of large balls, *Groups Geom. Dyn.* **8** (2014), 669–732.

[19] M. R. Bridson, A. Haefliger, *Metric spaces of non-positive curvature*, Grund. Math. Wiss. **319**, Springer, Berlin, 1999.

[20] A. I. Bufetov, Operator ergodic theorems for actions of free semigroups and groups, *Funktsional. Anal. i Prilozhen.* **34** (2000), 1–17.

[21] A. I. Bufetov, Convergence of spherical averages for actions of free groups, *Ann. of Math.* **155** (2002), 929–944.

[22] A. I. Bufetov, M. Khristoforov, A. Klimenko, Cesàro convergence of spherical averages for measure-preserving actions of Markov semigroups and groups, *Int. Math. Res. Not.* **2012** (2012), 4797–4829.

[23] A. I. Bufetov, A. Klimenko, Maximal inequality and ergodic theorems for Markov groups, *Tr. Mat. Inst. Steklova* **277** (2012), 33–48, translation: *Proc. Steklov Inst. Math.* **277** (2012), 27–42.

[24] D. Calegari, The ergodic theory of hyperbolic groups, in *Geometry and topology down under*, Contemp. Math. **597**, Amer. Math. Soc., Providence, RI, 2013, 15–52.

[25] D. Calegari, K. Fujiwara, Combable functions, quasimorphisms, and the central limit theorem, *Ergodic Theory Dynam. Systems* **30** (2010), 1343–1369.

[26] T. H. Colding, W. P. I. Minicozzi, Liouville theorems for harmonic sections and applications, *Comm. Pure Appl. Math.* **51** (1998), 113–138.

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

[28] A. Connes, J. Feldman, B. Weiss, An amenable equivalence relation is generated by a single transformation, *Ergodic Theory Dynam. Systems* **1** (1981), 431–450 (1982).

[29] M. Coornaert, Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov, *Pacific J. Math.* **159** (1993), 241–270.

[30] M. Coornaert, A. Papadopoulos, Horofunctions and symbolic dynamics on Gromov hyperbolic groups, *Glasg. Math. J.* **43** (2001), 425–456.

[31] P. Flajolet, R. Sedgewick, *Analytic combinatorics*, Cambridge Univ. Press, Cambridge, 2009.

[32] K. Fujiwara, A. Nevo, Maximal and pointwise ergodic theorems for word-hyperbolic groups, *Ergodic Theory Dynam. Systems* **18** (1998), 843–858.

[33] A. Gorodnik, A. Nevo, *The ergodic theory of lattice subgroups*, Annals of Math. Studies **172**, Princeton Univ. Press, Princeton, NJ, 2010.

[34] R. I. Grigorchuk, Ergodic theorems for the actions of a free group and a free semigroup, *Mat. Zametki* **65** (1999), 779–783; translation: *Math. Notes* **65** (1999), 779–783.

[35] M. Gromov, Hyperbolic groups, in *Essays in group theory*, Math. Sci. Res. Inst. Publ. **8**, Springer, New York, 1987, 75–263.

[36] Y. Guivarc'h, Généralisation d'un théorème de von Neumann, *C. R. Acad. Sci. Paris Sér.* **268** (1969), 1020–1023.

[37] Y. Katznelson, B. Weiss, The classification of nonsingular actions, revisited, *Ergodic Theory Dynam. Systems* **11** (1991), 333–348.

[38] A. S. Kechris, S. Solecki, S. Todorcevic, Borel chromatic numbers, *Adv. Math.* **141** (1999), 1–44.

[39] E. Lindenstrauss, Pointwise theorems for amenable groups, *Invent. Math.* **146** (2001), 259–295.

[40] J. von Neumann, Proof of the quasi-ergodic hypothesis, *Proc. Nat. Acad. Sci. USA* **18** (1932), 70–82.

[41] A. Nevo, Harmonic analysis and pointwise ergodic theorems for noncommuting transformations, *J. Amer. Math. Soc.* **7** (1994), 875–902.

[42] A. Nevo, Pointwise ergodic theorems for actions of groups, in *Handbook of dynamical systems. Vol. 1B*, Elsevier B. V., Amsterdam, 2006, 871–982.

[43] A. Nevo, E. M. Stein, A generalization of Birkhoff's pointwise ergodic theorem, *Acta Math.* **173** (1994), 135–154.

[44] D. S. Ornstein, B. Weiss, Entropy and isomorphism theorems for actions of amenable groups, *J. Analyse Math.* **48** (1987), 1–141.

[45] M. Pollicott, R. Sharp, Ergodic theorems for actions of hyperbolic groups, *Proc. Amer. Math. Soc.* **141** (2013), 1749–1757.

[46] A. A. Tempel′man, Ergodic theorems for general dynamical systems, *Dokl. Akad. Nauk SSSR* **176** (1967), 790–793; translation: *Soviet Math. Dokl.* **8** (1967), 1213–1216.

[47] R. Tessera, Volume of spheres in doubling metric measured spaces and in groups of polynomial growth, *Bull. Soc. Math. France* **135** (2007), 47–64.

[48] J. Väisälä, Gromov hyperbolic spaces, *Expo. Math.* **23** (2005), 187–231.

[49] B. Weiss, Actions of amenable groups, in *Topics in dynamics and ergodic theory*, London Math. Soc. Lecture Note Ser. **310**, Cambridge Univ. Press, Cambridge, 2003, 226–262.

[50] R. J. Zimmer, Amenable ergodic group actions and an application to Poisson boundaries of random walks, *J. Functional Analysis* **27** (1978), 350–372.

(Manuscrit reçu le 17 mars 2013 ;
accepté, après révision, le 2 décembre 2014.)

Lewis Bowen*
University of Texas at Austin
Mathematics Department
RLM 9.154
1 University Station, C1200
Austin, TX 78712, U.S.A.
E-mail: lpbowen@math.utexas.edu

Amos Nevo†
Technion - Israel Institute of Technology
Mathematics Department
Haifa, 32000, Israel
E-mail: anevo@tx.technion.ac.il