An extension of Kesten's criterion for amenability to topological Markov chains

Manuel Stadlbauer

Departamento de Matemática, Universidade Federal da Bahia, Av. Ademar de Barros s/n, 40170-110 Salvador, Brazil

Abstract

The main results of this note extend a theorem of Kesten for symmetric random walks on discrete groups to group extensions of topological Markov chains. In contrast to the result in probability theory, there is a notable asymmetry in the assumptions on the base. That is, it turns out that, under very mild assumptions on the continuity and symmetry of the associated potential, amenability of the group implies that the Gurevič-pressures of the extension and the base coincide whereas the converse holds true if the potential is Hölder continuous and the topological Markov chain has big images and preimages. Finally, an application to periodic hyperbolic manifolds is given.

1. Introduction and statement of main results

The motivation for the analysis of the change of pressure under group extensions stems from the attempt to relate two classical results from probability theory and geometry on the amenability of discrete groups. The probabilistic result was obtained by Kesten in [11] and characterises amenability in terms of the spectral radius of the Markov operator associated to a symmetric random walk, that is, a group G is amenable if and only if the spectral radius of the

---

E-mail addresses: manuel.stadlbauer@ufba.br, manuel.stadlbauer@gmail.com.

operator acting on $\ell^2(G)$ is equal to 1. The following counterpart in geometry was discovered by Brooks [3] using a completely different method. Assume that G is a Kleinian group acting on hyperbolic space $\mathbb{H}^{n+1}$ with exponent of convergence $\delta(G)$ bigger than $n/2$ and that $N \lhd G$ is a normal subgroup. Then the bottoms of the spectra of the Laplacians on $\mathbb{H}/G$ and $\mathbb{H}/N$ are equal if and only if G/N is amenable. Or equivalently, using the characterisation of the bottom of the spectrum in terms of the exponents of convergence, G/N is amenable if and only if $\delta(G) = \delta(N)$. More recently, these results were partially improved. Roblin [15] used conformal densities to prove that amenability implies $\delta(G) = \delta(N)$ if G is of divergence type and Sharp obtained in [19] the same statement for convex-cocompact Schottky groups using Grigorchuk's results on the co-growth of shortest representations (see [7]) applied to the Cayley graph of G.

In here, we consider group extensions of a topological Markov chain for a given potential function. That is, for a topological Markov chain (Σ_A, θ), a potential $\varphi : \Sigma_A \to (0, \infty)$ and a map $\psi : \Sigma_A \to H$ from Σ_A to a discrete group H, the group extension of $(\Sigma_A, \theta, \varphi)$ by ψ is defined by

$$T : \Sigma_A \times H \to \Sigma_A \times H, (x, g) \mapsto (\theta(x), g\psi(x))$$

and the lifted potential by $\hat{\varphi} : \Sigma_A \times H \to \mathbb{R}, (x, g) \mapsto \varphi(x)$, where it is throughout assumed that ψ is constant on the states $\mathcal{W}^1$ of Σ_A. This then gives rise to a natural notion of symmetry through the existence of an involution $\mathcal{W}^1 \to \mathcal{W}^1, w \mapsto w^{\dagger}$ such that $\psi([w^{\dagger}]) = \psi([w])^{-1}$ for all states $w \in \mathcal{W}^1$, where $[w]$ refers to the cylinder associated to w. This involution extends to finite words, leading to the notion of a weakly symmetric potential by requiring that there exists a sequence (D_n) with $\lim_n D_n^{1/n} = 1$ such that

$$\sup_{x \in [w], y \in [w^{\dagger}]} \frac{\prod_{j=0}^{n-1} (\varphi \circ \theta^j(x))}{\prod_{j=0}^{n-1} (\varphi \circ \theta^j(y))} \leq D_n,$$

for all $w \in \mathcal{W}^n$ and with $\mathcal{W}^n$ referring to the words of length n (see Section 3 for details). Note that this general framework establishes a connection between random walks on groups and the geodesic flow on the unit tangent bundle of $\mathbb{H}/N$ for a certain class of Kleinian groups G, since random walks can be recovered by assuming that Σ_A is a symmetric full shift equipped with a locally constant, symmetric potential whereas the relation to the geodesic flow is obtained through a group extension of the coding map associated with G as considered, e.g., in [1] or [12].

The main results in here extend Kesten's result for random walks to group extensions by replacing statements on the spectral radius by statements on the Gurevič pressure P_G. Furthermore, they reveal a certain asymmetry with respect to the method of proof and the requirements on the mixing properties of the base transformation θ. The first result, Theorem 4.1, essentially states that, if the potential is weakly symmetric and the group H is amenable, then $P_G(T) = P_G(\theta)$. This result is a consequence of Kesten's result, since the assumptions give rise to a construction of self-adjoint operators P_n on $\ell^2(H)$, for each $n \in \mathbb{N}$, whose spectral radii have to be equal to $\exp(n P_G(\theta))$ as a consequence of Kesten's theorem and the amenability of H. The arguments for obtaining the converse statement in Theorem 5.4 are more intricate and require that the potential is Hölder continuous and summable and that θ has the big images and preimages property. In this situation, we then have that $P_G(T) = P_G(\theta)$ implies that H is amenable. The proof is inspired by an argument of Day in [4] and relies on a careful analysis of the action of the Ruelle operator on the embedding of $\ell^2(H)$ into a certain subspace of $C(\Sigma_A \times H)$.

Finally, it is worth noting that the results of Kesten and Grigorchuk mentioned above can be easily deduced from Theorems 4.1 and 5.4 by choosing Σ_A to be a shift in the generators and φ to be a suitable, locally constant, symmetric potential. Furthermore, a reformulation in terms of Gibbs–Markov maps reveals that the results in here might be seen as an extension of Kesten's theorem for random walks on groups with independent, identically distributed increments to random walks with stationary, exponentially ψ-mixing increments. This generalisation then allows to apply the results to normal covers of Kleinian groups. That is, the theorem of Brooks extends to the class of essentially free Kleinian groups as defined below and, in particular, to a class of Kleinian groups with parabolic elements of arbitrary rank.

While writing this article, related results for locally constant potentials were independently obtained by Jaerisch [8] where it is shown that under this restriction, Theorem 5.4 is a consequence of a version of Kesten's result for graphs in [14].

2. Topological Markov chains

For a countable alphabet I and a matrix $A = (a_{ij} : i, j \in I)$ with $a_{ij} \in \{0, 1\}$ for all $i, j \in I$ and $\sum_j a_{ij} > 0$ for all $i \in I$, let the pair (Σ_A, θ) denote the associated one-sided topological Markov chain. That is,

$$\Sigma_A := \left\{ (w_k : k = 0, 1, \ldots) : w_k \in I, a_{w_k w_{k+1}} = 1 \ \forall i = 0, 1, \ldots \right\},$$
$$\theta : \Sigma_A \to \Sigma_A, \theta : (w_k : k = 1, 2, \ldots) \mapsto (w_k : k = 2, 3, \ldots).$$

A finite sequence $w = (w_1 \ldots w_n)$ with $n \in \mathbb{N}$, $w_k \in I$ for $k = 1, 2, \ldots, n$ and $a_{w_k w_{k+1}} = 1$ for $k = 1, 2, \ldots, n-1$ is referred to as a *word of length n*, and the set

$$[w] := \{(v_k) \in \Sigma_A : w_k = v_k \ \forall k = 1, 2, \ldots, n\}$$

as a *cylinder of length n*. The *set of admissible words of length n* will be denoted by $\mathcal{W}^n$, the length of $w \in \mathcal{W}^n$ by $|w|$ and the set of all admissible words by $\mathcal{W}^\infty = \bigcup_n \mathcal{W}^n$. Furthermore, since $\theta^n : [w] \to \theta^n([w])$ is a homeomorphism, the inverse exists and will be denoted by $\tau_w : \theta^n([w]) \to [w]$. For $a, b \in \mathcal{W}^\infty$ and $n \in \mathbb{N}$ with $n \geq |a|$, set

$$\mathcal{W}^n_{a,b} = \{(w_1 \ldots w_n) \in \mathcal{W}^n : (w_1 \ldots w_{|a|}) = a, \ w_n b \text{ admissible}\}.$$

As it is well known, Σ_A is a Polish space with respect to the topology generated by cylinders, and Σ_A is compact and locally compact with respect to this topology if and only if I is a finite set. Furthermore, recall that Σ_A is called *topologically transitive* if for all $a, b \in I$, there exists $n_{a,b} \in \mathbb{N}$ such that $\mathcal{W}^{n_{a,b}}_{a,b} \neq \emptyset$ and that Σ_A is called *topologically mixing* if for all $a, b \in I$, there exists $N_{a,b} \in \mathbb{N}$ such that $\mathcal{W}^n_{a,b} \neq \emptyset$ for all $n \geq N_{a,b}$. Moreover, a topological Markov chain is said to have *big images* or *big preimages* if there exists a finite set $\mathcal{I}_{\text{bip}} \subset \mathcal{W}$ such that for all $v \in \mathcal{W}$, there exists $\beta \in \mathcal{I}_{\text{bip}}$ such that $(v\beta) \in \mathcal{W}^2$ or $(\beta v) \in \mathcal{W}^2$, respectively. Finally, a topological Markov chain is said to have the *big images and preimages (b.i.p.) property* if the chain is topologically mixing and has big images and preimages (see [17]). Note that the b.i.p. property coincides with the notion of finite irreducibility for topological mixing topological Markov chains as introduced by Mauldin and Urbański [13].

We now consider a pair (Σ_A, φ) where $\varphi : \Sigma_A \to \mathbb{R}$ is a strictly positive function which we refer to as a *potential*. For $n \in \mathbb{N}$ and $w \in \mathcal{W}^n$, set

$$\Phi_n := \prod_{k=0}^{n-1} \varphi \circ \theta^k \quad \text{and} \quad C_w := \sup_{x, y \in [w]} \Phi_n(x)/\Phi_n(y). \tag{1}$$

The potential φ is said to have *(locally) bounded variation* if φ is continuous and there exists $C > 0$ such that $C_w \le C$ for all $n \in \mathbb{N}$ and $w \in \mathcal{W}^n$, and is called potential of *medium variation* if φ is continuous and, for all $n \in \mathbb{N}$, there exists $C_n > 0$ with $C_w \le C_n$ for all $w \in \mathcal{W}^n$ and $\lim_{n \to \infty} \sqrt[n]{C_n} = 1$. For positive sequences (a_n), (b_n) we frequently will write $a_n \ll b_n$ if there exists $C > 0$ with $a_n \le C b_n$ for all $n \in \mathbb{N}$, and $a_n \asymp b_n$ if $a_n \ll b_n \ll a_n$. A further, stronger assumption on the variation is related to local Hölder continuity. Therefore, recall that the n-th variation of a function $f : \Sigma_A \to \mathbb{R}$ is defined by

$$V_n(f) = \sup\{|f(x) - f(y)| : x_i = y_i, \ i = 1, 2, \ldots, n\}.$$

The function f is referred to as a *locally Hölder continuous function* if there exist $0 < r < 1$ and $C \ge 1$ such that $V_n(f) \ll r^n$ for all $n \ge 1$. Moreover, we refer to a locally Hölder continuous function with $\|f\|_\infty < \infty$ as a Hölder continuous function. We now recall the following well-known estimate. For $n \le m$, $x, y \in [w]$ for some $w \in \mathcal{W}^m$, and a locally Hölder continuous function f,

$$\left| \sum_{k=0}^{n-1} f \circ \theta^k(x) - f \circ \theta^k(y) \right| \ll \frac{1}{1-r} r^{m-n}. \tag{2}$$

In particular, the function $\exp f$ is a potential of bounded variation. For a given potential φ, the basic objects of thermodynamic formalism are partition functions. Since the state space might be countable, we consider partition functions Z_a^n for a fixed $a \in I$ which are defined by

$$Z_a^n := \sum_{\theta^n(x)=x, x \in [a]} \Phi_n(x).$$

Furthermore, we refer to the exponential growth rate of Z_a^n, that is to

$$P_G(\theta, \varphi) := \limsup_{n \to \infty} \log \sqrt[n]{Z_a^n} = \limsup_{n \to \infty} \frac{1}{n} \log Z_a^n,$$

as the *Gurevič pressure* of $(\Sigma_A, \theta, \varphi)$. This notion was introduced in [16] for topologically mixing systems where $\log \varphi$ is locally Hölder continuous. If $(\Sigma_A, \theta, \varphi)$ is transitive and φ is of medium variation, arguments in there combined with the decomposition of θ^p into mixing components, where p stands for the period of (Σ_A, θ), show that $P_G(\theta, \varphi)$ is independent of the choice of a and that

$$P_G(\theta, \varphi) = \lim_{n \to \infty, \mathcal{W}_{a,a}^n \ne \emptyset} \frac{1}{n} \log Z_a^n.$$

Furthermore, it is easy to see that $P_G(\theta, \varphi)$ remains unchanged by replacing $a \in \mathcal{W}^1$ with some $a \in \mathcal{W}^n$. Also recall that, if $\log \varphi$ is Hölder continuous and the system is topologically mixing, then a variational principle holds (see [16]).

We now recall the definitions of conformal and Gibbs measures related to a given potential φ. A Borel probability measure μ is called *φ-conformal* if

$$\mu(\theta(A)) = \int_A \frac{1}{\varphi} d\mu$$

for all Borel sets A such that $\theta|_A$ is injective. For $(w_1 \ldots w_{n+1}) \in \mathcal{W}^{n+1}$ and a potential of medium variation, it then immediately follows that

$$C_n^{-1} \mu(\theta([w_{n+1}])) \leq \frac{\mu([w_1 \ldots w_{n+1}])}{\Phi_n(x)} \leq C_n \mu(\theta([w_{n+1}])) \tag{3}$$

for all $x \in [w_1 \ldots w_{n+1}]$. Note that this estimate implies that $P_G(\theta, \varphi) = 0$ is a necessary condition for the existence of a conformal measure with respect to a potential of medium variation. Moreover, the above estimate motivates the following definitions. Assume that there exists a sequence $(B_n : n \in \mathbb{N})$ with $B_n \geq 1$ such that

$$B_n^{-1} \leq \frac{\mu([w])}{\Phi_n(x)} < B_n \tag{4}$$

for all $n \in \mathbb{N}$, $w \in \mathcal{W}^n$ and $x \in [w]$. If $\sup_n B_n < \infty$, then μ is called φ-*Gibbs measure*, and if $\lim_{n \to \infty} B_n^{1/n} = 1$, then μ is called *weak φ-Gibbs measure*. In order to introduce a further basic object, the Ruelle operator, we define the action of the inverse branches of τ_v on functions as follows. For $v \in \mathcal{W}^n$ and $f : \Sigma_A \to \mathbb{R}$, set

$$f \circ \tau_v : \Sigma_A \to \mathbb{R}, \quad x \mapsto \mathbf{1}_{\theta^n([v])}(x) f(\tau_v(x)),$$

that is $f \circ \tau_v(x) := f(\tau_v(x))$ for $x \in \theta^n([v])$ and $f \circ \tau_v(x) := 0$ for $x \notin \theta^n([v])$. The *Ruelle operator* is then defined by

$$L_\varphi(f) = \sum_{v \in \mathcal{W}^1} \varphi \circ \tau_v \cdot f \circ \tau_v,$$

where $f : \Sigma_A \to \mathbb{C}$ is an element of an appropriate function space such that the possibly infinite sum on the right hand side is well defined.

3. Group extensions of topological Markov chains

To introduce the basic object of our analysis, fix a countable discrete group G and a map $\psi : \Sigma_A \to G$ such that ψ is constant on $[w]$ for all $w \in \mathcal{W}^1$. Then, with $X := \Sigma_A \times G$ equipped with the product topology, the *group extension or G-extension* (X, T) of (Σ_A, θ) is defined by

$$T : X \to X, (x, g) \mapsto (\theta x, g \psi(x)).$$

Observe that (X, T) also is a topological Markov chain and that its cylinder sets are given by $[w, g] := [w] \times \{g\}$, for $w \in \mathcal{W}^\infty$ and $g \in G$. Furthermore, set $X_g := \Sigma_A \times \{g\}$ and

$$\psi_n(x) := \psi(x) \psi(\theta x) \cdots \psi(\theta^{n-1} x)$$

for $n \in \mathbb{N}$ and $x \in \Sigma_A$. Observe that $\psi_n : \Sigma_A \to G$ is constant on cylinders of length n and, in particular, that $\psi(w) := \psi_n(x)$, for some $x \in [w]$ and $w \in \mathcal{W}^n$, is well defined. Moreover, for $a, b \in \mathcal{W}^1$ and $n \in \mathbb{N}$, set

$$G_n(a, b) := \{\psi(w) : n \in \mathbb{N}, w \in \mathcal{W}^n, [a] \supset [w], \theta^n([w]) \supset [b]\}.$$

Note that (X, T) is topologically transitive if and only if, for $a, b \in \mathcal{W}^1$, $g \in G$, there exists $n \in \mathbb{N}$ with $g \in G_n(a, b)$, and that (T, X) is *topologically mixing* if and only if, for $a, b \in \mathcal{W}^1$ and $g \in G$, there exists $N \in \mathbb{N}$ (depending on a, b, g) such that $g \in G_n(a, b)$ for all $n > N$. Note that the base transformation (θ, Σ_A) of a topologically transitive group extension has to

be topologically transitive, and topologically mixing if the extension is topologically mixing, respectively. Furthermore, if (T, X) is topologically transitive, then $\{\psi(a) : a \in \mathcal{W}^1\}$ is a generating set for G.

Throughout, we now fix a topological mixing topological Markov chain (Σ_A, θ), and a topological transitive G-extension (T, X). Furthermore, we fix a (positive) potential $\varphi : \Sigma_A \to \mathbb{R}$ with $P_G(\theta, \varphi) = 0$. Note that φ lifts to a potential $\hat{\varphi}$ on X by setting $\hat{\varphi}(x, g) := \varphi(x)$. For ease of notation, we will not distinguish between $\hat{\varphi}$ and φ. Moreover, for $v \in \mathcal{W}^\infty$, the inverse branch given by $[v, \cdot]$ will be as well denoted by τ_v, that is $\tau_v(x, g) := (\tau_v(x), g\psi(v)^{-1})$. In order to distinguish between the Ruelle operator and the partition functions with respect to θ and T, these objects for the group extension will be written in calligraphic letters, that is, for $a \in \mathcal{W}, \xi \in [a] \times \{\mathrm{id}\}, (\eta, g) \in X$, and $n \in \mathbb{N}$,

$$\mathcal{L}(f)(\xi, g) := \sum_{v \in \mathcal{W}} \varphi \circ \tau_v(\xi) f \circ \tau_v(\xi, g), \qquad \mathcal{Z}_{a,g}^n := \sum_{\substack{T^n(x,g)=(x,g), \\ x \in [a]}} \Phi_n(x).$$

4. Extensions by amenable groups

In this section, we show that the Gurevič pressure remains unchanged under extension by an amenable group. In particular, it will turn out that this statement is true under very mild conditions. Also note that a similar result was proven in [19] for extensions of subshifts of finite type with respect to a hyperbolic potential. We first recall the definition of amenability using the Følner condition (see [5]). That is, G is referred to as an *amenable group* if and only if there exists a sequence (K_n) of finite subsets of G with $\bigcup_n K_n = G$ such that

$$\lim_{n \to \infty} |gK_n \triangle K_n|/|K_n| = 0 \quad \forall g \in G.$$

In here, $\triangle$ refers to the symmetric difference, and $|\cdot|$ to the cardinality of a set. We are now interested in the characterisation of amenability in terms of the Gurevič pressure of a group extension. As it will turn out, this is an extension of the following result of Kesten in [11]. Let m be a probability measure on G with $m(g^{-1}) = m(g)$ and assume that the support $\mathrm{supp}(m)$ of m is a generating set for G. Then the spectral radius of the operator P on the complex space $\ell^2(G)$ given by $Pf(\gamma) := \sum_{g \in G} f(\gamma g^{-1}) m(g)$ is equal to one if and only if G is amenable.

In order to obtain an extension of this result to shift spaces, one has to consider group extensions with symmetry. Namely, we say that (Σ_A, θ, ψ) is *symmetric* if there exists $\mathcal{W}^1 \to \mathcal{W}^1, w \mapsto w^\dagger$ with the following properties.

1. For $w \in \mathcal{W}^1, (w^\dagger)^\dagger = w$.
2. For $v, w \in \mathcal{W}^1$, the word (vw) is admissible if and only if $(w^\dagger v^\dagger)$ is admissible.
3. $\psi(v^\dagger) = \psi(v)^{-1}$ for all $v \in \mathcal{W}^1$.

Observe that this notion of symmetry can be extended to an involution of $\mathcal{W}^\infty$ by defining $(w_1 \ldots w_n)^\dagger := (w_n^\dagger \ldots w_1^\dagger)$. We refer to φ as a *weakly symmetric potential* if φ is continuous and there exists a sequence (D_n) with $\lim_{n \to \infty} \sqrt[n]{D_n} = 1$ such that, for all $n \in \mathbb{N}$ and $w \in \mathcal{W}^n$,

$$\sup_{x \in [w], y \in [w^\dagger]} \Phi_n(x)/\Phi_n(y) \le D_n.$$

If $\sup D_n < \infty$, then φ is referred to as a *symmetric potential*. Note that a weakly symmetric or symmetric potential is necessarily of medium variation with respect to $C_n := D_n^2$ or of bounded variation, respectively.

We now prove that amenability of G implies that $P_G(T) = P_G(\theta)$. For this purpose, we construct a family of self-adjoint operators on $\ell^2(G)$ for symmetric group extensions as follows. By transitivity, there exists $a \in \mathcal{W}^\infty$ with $\psi(a) = \mathrm{id}$. Furthermore, for $v \in \mathcal{W}^\infty$, we have that $ava^\dagger$ is admissible if and only if $(ava^\dagger)^\dagger = av^\dagger a^\dagger$ is admissible. Hence, for $n > |a|$,

$$\iota : \mathcal{W}^n_{a,a^\dagger} \to \mathcal{W}^n_{a,a^\dagger}, \qquad \iota(av) := av^\dagger$$

defines an involution. Since $\psi(a) = \mathrm{id}$, it follows that $\psi(v) = \psi(\iota v)^{-1}$ for all $v \in \mathcal{W}^n_{a,a^\dagger}$. We now assume that $P_G(\theta)$ is finite and fix $\xi \in [a^\dagger]$. For $n \in \mathbb{N}$ and $v \in \mathcal{W}^n_{a,a^\dagger}$, set

$$\pi(\xi, v) := \frac{1}{2} \left(\Phi_n(\tau_v(\xi)) + \Phi_n(\tau_{\iota v}(\xi)) \right).$$

This then gives rise to an operator $P_n : \ell^2(G) \to \ell^2(G)$ on the complex Hilbert space $\ell^2(G)$ by

$$P_n(f)(\gamma) := \sum_{v \in \mathcal{W}^n_{a,a^\dagger}} \pi(\xi, v) f(\gamma \psi(v)^{-1}),$$

where, for ease of notation, the fact that the operator P_n depends on ξ is omitted. Note that $P_n(1) = L_\varphi^n(\mathbf{1}_{[a]})(\xi) < \infty$ and, with $\langle f, g \rangle = \sum \overline{f(\gamma)} g(\gamma)$ referring to the standard inner product, $\langle \mathbf{1}_\gamma, P_n(\mathbf{1}_{\gamma^*}) \rangle = \langle P_n(\mathbf{1}_\gamma), \mathbf{1}_{\gamma^*} \rangle$, for $\gamma, \gamma^* \in G$. In particular, this implies that P_n is self-adjoint.

Combining $\langle \mathbf{1}_\gamma, P_n(\mathbf{1}_\gamma) \rangle = \langle \mathbf{1}_{\mathrm{id}}, P_n(\mathbf{1}_{\mathrm{id}}) \rangle$ for all $\gamma \in G$ with P_n being self-adjoint then gives that the spectral radius ρ_n of P_n satisfies (see, e.g., [9])

$$\rho_n = \limsup_{k \to \infty} \sqrt[k]{\langle \mathbf{1}_{\mathrm{id}}, P_n^k(\mathbf{1}_{\mathrm{id}}) \rangle},$$

and that, if P_n is a positive operator, then the $\limsup$ above is a limit. As an immediate corollary to Kesten's theorem on the spectral radius of the Markov operator associated with a symmetric random walk (see [11] and [9, Theorem 5]), we obtain the following theorem for group extensions under very mild conditions.

Theorem 4.1. *Assume that T is a topologically transitive, symmetric group extension of the topologically mixing topological Markov chain (Σ_A, θ), φ is weakly symmetric and $P_G(\theta)$ is finite. Then $P_G(T) = P_G(\theta)$, if G is an amenable group.*

Proof. Since we obviously have that $P_G(T) \le P_G(\theta)$, it remains to prove the reverse inequality. For $n \ge |a|$ and $g \in G$, set

$$m_n(g) := \frac{1}{P_n(1)} \sum_{v \in \mathcal{W}^n_{a,a^\dagger} : \psi(v) = g} \pi(\xi, v).$$

Note that $m_n(g)$ is well defined since $|P_G(\theta)| < \infty$ and that m_n is a symmetric probability measure on G, that is $m_n(g) = m_n(g^{-1})$. Moreover, note that the group generated by the support of m_n is a subgroup of G and hence is amenable. Since the Markov operator associated with the symmetric random walk given by m_n coincides with $P_n(\cdot)/P_n(1)$ we have by Kesten's theorem that $\rho_n = P_n(1)$. In order to prove the assertion it suffices to show that $\lim_n \log(\rho_n)/n = P_G(\theta)$ and $\limsup_n \log(\rho_n)/n \le P_G(T)$.

Step 1. For $n \geq |a|$, we have $\iota w \in \mathcal{W}^n_{a,a^\dagger}$ if and only if $w \in \mathcal{W}^n_{a,a^\dagger}$. Hence,

$$\rho_n = P_n(1) = \sum_{w \in \mathcal{W}^n_{a,a^\dagger}} \frac{1}{2} \left(\Phi_n(\tau_w(\xi)) + \Phi_n(\tau_{\iota w}(\xi)) \right) = \sum_{w \in \mathcal{W}^n_{a,a^\dagger}} \Phi_n(\tau_w(\xi)).$$

Choose $k \geq |a|$ and $v \in \mathcal{W}^k_{a,a^\dagger}$. For $n \geq k$, the medium variation property of φ gives

$$\rho_n = \sum_{w \in \mathcal{W}^n_{a,a^\dagger}} \Phi_n(\tau_w(\xi)) \geq C_k^{-1} C_n^{-1} \Phi_k(\tau_v(\xi)) Z_b^{n-k},$$

where $b := (a_{|a|})^\dagger$ for $a = (a_1 \ldots a_{|a|})$. Hence, $\limsup_n (\log \rho_n)/n = P_G(\theta)$.

Step 2. For ease of notation, we will write $x = y^{\pm k} z$ for $y^{-k} z \leq x \leq y^k z$. Then, for $w \in \mathcal{W}^n_{a,a^\dagger}$, weak symmetry implies

$$\Phi_n(\tau_w(\xi)) = \Phi_{n+|a|}(\tau_w(\xi))/\Phi_{|a|}(\xi) = (C_n D_{n+|a|})^{\pm 1} \Phi_n(\tau_{\iota w}(\xi)).$$

Hence, there exists $(\widehat{D}_n)$ with $\pi(\xi, w) = \Phi_n(\tau_w(\xi))\widehat{D}_n^{\pm 1}$ and $\lim_n \widehat{D}_n^{1/n} = 1$. By transitivity, there exist $l \in \mathbb{N}$, and $v \in \mathcal{W}^l_{a^\dagger a}$ with $\psi(v) = \mathrm{id}$. Hence, for $w_1, w_2, \ldots, w_k \in \mathcal{W}^n_{aa^\dagger}$, we have $w^* := (w_1 v w_2 v \ldots w_k v) \in \mathcal{W}^{k(n+l)}_{aa}$ and, for $x \in [w^*]$,

$$\Phi_{k(n+l)}(x) \geq \left(\widehat{D}_n^{-1} \inf\{ \Phi_l(y) : y \in [va] \} \right)^k \prod_{i=1}^k \pi(\xi, w_i).$$

This gives rise to the following estimate for $\langle 1_{\mathrm{id}}, P_n^k(1_{\mathrm{id}}) \rangle$ in terms of the partition function $\mathcal{Z}^{k(n+l)}_{a,\mathrm{id}}$ with respect to T as defined above.

$$\frac{1}{nk} \log\langle 1_{\mathrm{id}}, P_n^k(1_{\mathrm{id}}) \rangle \leq \frac{1}{n} \log \left(\widehat{D}_n^{-1} \inf\{ \Phi_l(y) : y \in [va] \} \right) + \frac{1}{nk} \log \mathcal{Z}^{k(n+l)}_{a,\mathrm{id}}.$$

Since P_n^2 is a positive operator with $\|P_n^2\| = \rho_n^2$, we obtain, by taking the limit for $k \to \infty$, $k \in 2\mathbb{N}$ and then for $n \to \infty$, that $\lim_n \log(\rho_n)/n \leq P_G(T)$. $\quad\square$

5. Kesten's theorem for group extensions

The essential ingredient of the proof of Theorem 4.1 is the fact that a symmetric probability measure defines a symmetric operator on $\ell^2(G)$. So, in order to prove the analogue of Kesten's result for group extensions of topological Markov chains, it remains to show that $P_G(\theta) = P_G(T)$ implies amenability, where one is tempted again to use the spectral radius formula applied to symmetric operators on $\ell^2(G)$ as, e.g., in [9, p. 478]. However, it will turn out that the key step in here is to carefully analyse an embedding of $\ell^2(G)$ and use an argument based on uniform rotundity to show that in the case of amenable groups, almost eigenfunctions are indicator functions.

The arguments of proof in here rely on stronger topological mixing properties in the base. That is, we will have to assume that the base has the b.i.p.-property. As a first consequence of this property, we obtain the existence of the following finite subset of $\mathcal{W}^\infty$.

Lemma 5.1. *Assume that (X, T) is a topologically transitive group extension of (Σ_A, θ) and that (Σ_A, θ) has the b.i.p.-property. Then there exist $n \in \mathbb{N}$ and a finite subset $\mathcal{J}$ of $\mathcal{W}^n$ such that for each pair (β, β') with $\beta, \beta' \in \mathcal{I}_{bip}$ there exists $w_{\beta,\beta'} \in \mathcal{J}$ such that $(w_{\beta,\beta'}) \in \mathcal{W}^n$ and $\psi_n(w_{\beta,\beta'}) = \mathrm{id}$.*

Proof. Let $p \in \mathbb{N}$ refer to the period p of the transitive topological Markov chain (X, T). Then, for each $a \in \mathcal{W}$, there exists $N_a \in \mathbb{N}$ such that $T^{ps}([a, g]) \supset [a, g]$ for all $s \geq N_a$ and $g \in G$. Hence, for each pair $(\beta, \beta') \in \mathcal{I}_{\text{bip}}^2$ with $(\beta'\beta)$ admissible, there is $N_{\beta, \beta'}$ such that for each $s \geq N_{\beta, \beta'}$ there exists $v_{\beta, \beta'} \in \mathcal{W}^{ps-2}$ such that $(\beta'\beta v_{\beta, \beta'}\beta')$ is admissible and $\psi_{ps}(\beta'\beta v_{\beta, \beta'}\beta') = \psi_{ps}(\beta v_{\beta, \beta'}\beta') = \text{id}$. Since $\mathcal{I}_{\text{bip}}$ is finite it follows that there exists k (given by $\max\{pN_{\beta, \beta'} : (\beta, \beta') \in \mathcal{I}_{\text{bip}}^2\}$) such that $v_{\beta, \beta'}$ can be always chosen to be an element of $\mathcal{W}^{k-2}$.

By possibly adding finitely many states we may assume without loss of generality that the subsystem of Σ_A with alphabet $\mathcal{I}_{\text{bip}}$ is topologically mixing. It then follows from this that there exists some $l \in \mathbb{N}$ such that each pair (β_0, β_l) in $\mathcal{I}_{\text{bip}}$ can be connected by an admissible word of the form

$$w_{\beta_0, \beta_l} := (\beta_0 v_{\beta_0, \beta_1} \beta_1 \beta_2 v_{\beta_2, \beta_3} \beta_3 \beta_4 \ldots v_{\beta_{l-1}, \beta_l} \beta_l).$$

The assertion follows with $\mathcal{J} := \{w_{\beta, \beta'} : \beta, \beta' \in \mathcal{I}_{\text{bip}}\}$. $\square$

A further important consequence of the b.i.p.-property is the existence of an invariant Gibbs measure. That is, if (Σ_A, θ) and φ are given such that (Σ_A, θ) has the b.i.p.-property, $\log \varphi$ is Hölder continuous and $\|L_\varphi(1)\|_\infty < \infty$, then there exist a $\exp(-P_G(\theta, \varphi)) \cdot \varphi$-Gibbs measure μ and a Hölder-continuous eigenfunction h such that $h d\mu$ is an invariant probability measure. It is also worth noting that (Σ_A, θ) has the Gibbs–Markov property as defined below with respect to μ (see [13,17]). Moreover, since the function $\log h$ is uniformly bounded from above and below, we assume from now on, without loss of generality, that $P_G(\theta, \varphi) = 0$ and $L_\varphi(1) = 1$.

The existence of μ then gives rise to the following definition of $\mathcal{H}_1$ and $\mathcal{H}_\infty$. Given a measurable function $f : X \to \mathbb{R}$, $g \in G$ and $p = 1$ or $p = \infty$, set $\|f\|_p^g := \|f(\cdot, g)\|_p$ and define

$$\llbracket f \rrbracket_p := \sqrt{\sum_{g \in G} (\|f\|_p^g)^2} \quad \text{and} \quad \mathcal{H}_p := \{f : X \to \mathbb{R} : \llbracket f \rrbracket_p < \infty\}.$$

Furthermore, set $\mathcal{H}_c := \{f \in \mathcal{H} : f \text{ is constant on } X_g \forall g \in G\}$ and $\rho := \exp(P_G(T, \hat{\varphi}))$.

Proposition 5.2. *The function spaces $(\mathcal{H}_1, \llbracket \cdot \rrbracket_1)$ and $(\mathcal{H}_\infty, \llbracket \cdot \rrbracket_\infty)$ are Banach spaces, the operators $\mathcal{L}_\varphi^k : \mathcal{H}_\infty \to \mathcal{H}_\infty$ are bounded and there exists $C \geq 1$ such that $\llbracket \mathcal{L}_\varphi^k \rrbracket_\infty \leq C$ for all $k \in \mathbb{N}$. Furthermore,*

$$\Lambda_k := \sup\left(\{\llbracket \mathcal{L}_\varphi^k(f) \rrbracket_1 / \llbracket f \rrbracket_1 : f \geq 0, f \in \mathcal{H}_c\}\right) \leq 1$$

and $\limsup_{k \to \infty} (\Lambda_k)^{1/k} \geq \rho$.

Proof. The proof that $(\mathcal{H}_p, \llbracket \cdot \rrbracket_p)$ are Banach spaces is standard and therefore omitted. In order to prove the uniform bound for $\llbracket \mathcal{L}_\varphi^k \rrbracket_\infty$, assume that $f \in \mathcal{H}_\infty$ and $k \in \mathbb{N}$. Using Jensen's inequality, we obtain

$$\llbracket \mathcal{L}_\varphi^k(f) \rrbracket_\infty^2 \leq \sum_{g \in G} \sup_{x \in \Sigma_A} \left(\sum_{v \in \mathcal{W}^k} \Phi_k \circ \tau_v(x) \|f\|_\infty^{g\psi_k(v)^{-1}}\right)^2$$

$$\leq \sum_{g \in G} \sup_{x \in \Sigma_A} \sum_{v \in \mathcal{W}^k} \Phi_k \circ \tau_v(x) \left(\|f\|_\infty^{g\psi_k(v)^{-1}}\right)^2$$

$$\leq C \sum_{v \in \mathcal{W}^k} \mu([v]) \llbracket f \rrbracket_\infty^2 = C \llbracket f \rrbracket_\infty^2,$$

where C is given by the Gibbs property of μ on Σ_A (see (4)). Hence, C is an upper bound for $[\![\mathcal{L}_\varphi^k]\!]_\infty$, independent of k.

Now assume that $f \in \mathcal{H}_c$. It then follows from the conformality of μ that

$$[\![\mathcal{L}_\varphi^k(f)]\!]_1 \le \sum_{v \in \mathcal{W}^k} [\![\Phi_k \circ \tau_v f \circ \tau_v]\!]_1 = \sum_{v \in \mathcal{W}^k} \left(\sum_{g \in G} \left(\int \Phi_k \circ \tau_v f \circ \tau_v(\cdot, g) d\mu \right)^2 \right)^{1/2}$$

$$= \sum_{v \in \mathcal{W}^k} \left(\sum_{g \in G} \mu([v])^2 f(x, g\psi_k(v)^{-1})^2 \right)^{1/2} = [\![f]\!]_1, \tag{5}$$

where $x \in \Sigma_A$ is arbitrary. Hence, $\Lambda_k \le 1$ for all $k \in \mathbb{N}$. Finally, observe that the Gibbs property of μ implies that

$$[\![\mathcal{L}_\varphi^n(\mathbf{1}_{X_{\mathrm{id}}})]\!]_1 \ge \|\mathcal{L}_\varphi^n(\mathbf{1}_{X_{\mathrm{id}}})\|_1^{\mathrm{id}} = \sum_{v : \psi_n(v)=\mathrm{id}} \mu([v]) \gg \sum_{x : T^n(x,\mathrm{id})=(x,\mathrm{id})} \Phi_n(x) \ge \mathcal{Z}_{a,\mathrm{id}}^n,$$

for all $a \in \mathcal{W}^1$. Hence, $\limsup_{n \to \infty} (\Lambda_n)^{1/n} \ge \rho$. $\square$

The proof of the following result is inspired by an argument of Day (see [4, Lemma 4]) which relies on the rotundity of $\ell^2(G)$. Recall that a Banach space $(B, \|\cdot\|)$ is *uniformly rotund* if for all $\delta > 0$ there exists $\epsilon > 0$ such that, for all f, g with $\|f - g\| \ge \delta$ and $\|f\| = \|g\| = 1$, it follows that $\|f + g\| \le 2 - \epsilon$. Note that the space $\mathcal{H}$ does not has this property but the closed subspace $\mathcal{H}_c$ which is isomorphic to $\ell^2(G)$. Since $\ell^2(G)$ is uniformly rotund, it follows that $\mathcal{H}_c$ has this property as well.

Lemma 5.3. *Assume that $\rho = 1$. Then there exists $n \in \mathbb{N}$ such that, for given $\epsilon > 0$, there exists $f \in \mathcal{H}_c$ with $f \ge 0$ and $[\![\mathcal{L}_\varphi^n(f) - f]\!]_1 \le \epsilon[\![f]\!]_1$.*

Proof. Let k be given by $\mathcal{J} \subset \mathcal{W}^k$ where $\mathcal{J}$ refers to the finite set given by Lemma 5.1 and assume that

$$\delta_1 := \inf\{[\![\mathcal{L}_\varphi^k(f) - f]\!]_1 / [\![f]\!]_1 : f \in \mathcal{H}_c, f \ge 0, f \neq 0\} > 0.$$

Step 1. We begin by showing that $\delta_1 > 0$ implies that there exists a finite set of words such that, for all $x \in \Sigma_A$ and $f \in \mathcal{H}_c$, the estimate (6) below holds and that two arbitrary words can be joined by elements of this finite set. In order to do so, note that it is possible to choose a finite set $\mathcal{W}^* \subset \mathcal{W}^k$ with $\sum_{v \in \mathcal{W}^k \setminus \mathcal{W}^*} \mu([v]) \le \delta_1/(4)$. For $f \in \mathcal{H}_c$, we hence have that

$$\left[\!\left[\sum_{v \in \mathcal{W}^k \setminus \mathcal{W}^*} \Phi_k \circ \tau_v f \circ \tau_v \right]\!\right]_1 \le \sum_{v \in \mathcal{W}^k \setminus \mathcal{W}^*} \mu([w]) [\![f]\!]_1 \le \delta_1 [\![f]\!]_1 / (4).$$

Using $L_\varphi(1) = 1$ and the $\triangle$-inequality then gives

$$\delta_1 [\![f]\!]_1 \le [\![\mathcal{L}_\varphi^k(f) - f]\!]_1 = \left[\!\left[\sum_{v \in \mathcal{W}^k} \Phi_k \circ \tau_v(f \circ \tau_v - f) \right]\!\right]_1$$

$$\le \sum_{v \in \mathcal{W}^*} [\![\Phi_k \circ \tau_v(f \circ \tau_v - f)]\!]_1$$

$$+ \left[\!\left[\sum_{v \notin \mathcal{W}^*} \Phi_k \circ \tau_v f \circ \tau_v \right]\!\right]_1 + \left[\!\left[\sum_{v \notin \mathcal{W}^*} \Phi_k \circ \tau_v f \right]\!\right]_1$$

$$\leq \sum_{v \in \mathcal{W}^*} \mu([v]) [\![(f \circ \tau_v - f)\mathbf{1}_{\theta^k([v]) \times G}]\!]_1 + \delta_1/2.$$

It follows from a convex sum argument that there exists $w_f \in \mathcal{W}^*$ with $[\![(f \circ \tau_{w_f} - f)$ $\mathbf{1}_{\theta^k([w_f]) \times G}]\!]_1 \geq \delta_1 [\![f]\!]_1 /2$. Furthermore, note that the b.i.p.-property implies that $\mu(\theta^k([v]))$ is uniformly bounded from below for all $v \in \mathcal{W}^k$. With $\hat{f}(g)$ referring to $f(x, g)$, we hence obtain that

$$[\![(f \circ \tau_{w_f} - f)\mathbf{1}_{\theta^k([w_f]) \times G}]\!]_1 = \mu(\theta^k([w_f])) \|\hat{f} - \hat{f}(\cdot \, \psi(w_f)^{-1})\|_{\ell^2(G)}$$

$$\gg \|\hat{f} - \hat{f}(\cdot \, \psi(w_f)^{-1})\|_{\ell^2(G)}.$$

For $x \in \Sigma_A$ and $a \in \mathcal{I}_{\text{bip}}$ with $[a] \subset \theta^k([w_f])$, we now choose $w_a, w_x \in \mathcal{J}$ such that $x \in \theta^k([w_x])$, $(w_a a)$ is admissible, and $u := (w_a w_x)$, $v := (w_f w_x)$ are in $\mathcal{W}^{2k}$. Since $\psi_k(w_a) = \psi_k(w_x) = \text{id}$ and $f \in \mathcal{H}_c$, we have $\psi_{2k}(u) = \text{id}$ and $\psi_{2k}(v) = \psi_k(w_f)$. Hence, the rotundity of $\mathcal{H}_c$ implies that there exists a uniform constant $\delta_2 > 0$ with

$$\frac{1}{2} \|f \circ \tau_u(x, \cdot) + f \circ \tau_v(x, \cdot)\|_{\ell^2(G)} \leq (1 - \delta_2)[\![f]\!]_1, \quad \forall x \in [a]. \tag{6}$$

Furthermore, we may choose $w_a \in \mathcal{J}$ above such that there exists $\beta \in \mathcal{I}_{\text{bip}}$ with $[w_a] \cup [w_f] \subset \theta([\beta])$. By substituting u and v with (wu) and (wv), respectively, for some admissible word w in $\mathcal{J}^{m-2}$ and $m \in \mathbb{N}$ to be specified later, we hence may assume that there exists a finite subset $\mathcal{W}^\dagger$ of $\mathcal{W}^{mk}$ with the following properties. For all $w_i \in \mathcal{W}^{n_i}$, $i = 1, 2$ and $f \in \mathcal{H}_c$, there exist $u(w_1, f, w_2)$ and $v(w_1, f, w_2)$ in $\mathcal{W}^\dagger$ such that

1. the estimate (6) holds for $u := u(w_1, f, w_2)$ and $v := v(w_1, f, w_2)$, with $x \in [w_2]$,
2. the first $k(m - 2)$ letters of $u(w_1, f, w_2)$ and $v(w_1, f, w_2)$ coincide,
3. $w_1 u(w_1, f, w_2) w_2$ and $w_1 v(w_1, f, w_2) w_2$ are admissible.

Step 2. In order to obtain the estimate in (8), we prove that the fluctuations of the potential can be absorbed into the uniform factor in (6). Observe that it follows from the Hölder continuity of $\log \varphi$ that we may choose m such that

$$\sqrt{1 - \delta_2} \leq \frac{\Phi_n \circ \tau_{w_1}(\tau_{u(w_1, f, w_2)}(x))}{\Phi_n \circ \tau_{w_1}(\tau_{v(w_1, f, w_2)}(x))} \leq \left(\sqrt{1 - \delta_2}\right)^{-1}, \tag{7}$$

for all $n \in \mathbb{N}$ and $w \in \mathcal{W}^n$. Moreover, since $|\mathcal{W}^\dagger| < \infty$, we have

$$2\alpha := \inf \left(\left\{ \Phi_{mk}(\tau_u(x)) : x \in \theta^{mk}([u]), u \in \mathcal{W}^\dagger \right\} \right) > 0.$$

By dividing each $u \in \mathcal{W}^\dagger$ into two words u_1 and u_2 and setting $\Phi_{mk}(\tau_{u_2}(x)) := \Phi_{mk}(\tau_u(x)) - \alpha$ and $\Phi_{mk}(\tau_{u_1}(x)) := \alpha$ for each $x \in \theta^{mk}([u])$, we may assume without loss of generality that $\Phi_{mk}(\tau_u(x)) = \alpha$ for all $x \in \theta^{mk}([u])$ and $u \in \mathcal{W}^\dagger$.

By combining the above considerations, we are now in a position to prove the main estimate with respect to $\mathcal{W}^\dagger$, for a given function $f \in \mathcal{H}_c$ and $(k + 1)$ finite words $w_i \in \mathcal{W}^{n_i}$, for $i = 0, 1, \ldots, k$. For a finite word w, set $f_w(x, g) := f(\tau_w(x), g\psi(w)^{-1})$, and define by

induction, for $j = 1, 2, \ldots, k$,

$$f_j := \sum_{(i_1, \ldots, i_{j-1}) \in \{1,2\}^{j-1}} f_{w_0 u_1^{(i_1)} w_1 \ldots u_{j-1}^{(i_{j-1})} w_{j-1}},$$

$$u_j^{(1)} := u(w_{j-1}, f_j, w_j), \qquad u_j^{(2)} := v(w_{j-1}, f_j, w_j).$$

We then have, where $N := n_0 + n_1 + \cdots + n_n + nmk$,

$$\left\| \sum_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} w_1 \ldots u_n^{(i_n)} w_n}(x)) f(\tau_{w_0 u_1^{(i_1)} w_1 \ldots u_n^{(i_n)} w_n})(x, \cdot) \right\|_{\ell^2(G)}$$

$$\leq \left(\max_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} \ldots u_n^{(i_n)} w_n}(x)) \right)$$

$$\times \left\| \sum_{(i_1, \ldots, i_n) \in \{1,2\}^n} \hat{f}_{w_0 u_1^{(i_1)} \ldots w_{n-1}} (\cdot \, \psi(w_n)^{-1} \psi(u_n^{(i_n)})^{-1}) \right\|_{\ell^2(G)}$$

$$= \left(\max_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} \ldots u_n^{(i_n)} w_n}(x)) \right) \left\| \sum_{i_n \in \{1,2\}} \hat{f}_n (\cdot \, \psi((u_n^{(i_n)}))^{-1}) \right\|_{\ell^2(G)}$$

$$\leq \left(\max_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} \ldots u_n^{(i_n)} w_n}(x)) \right) (2(1 - \delta_2))$$

$$\times \left\| \sum_{i_{n-1} \in \{1,2\}} \hat{f}_{n-1} (\cdot \, \psi(u_{n-1}^{(i_{n-1})})^{-1}) \right\|_{\ell^2(G)}.$$

Combining (7) with $\Phi_{mk}|_{[u]} = \alpha$ for all $u \in \mathcal{W}^\dagger$, we hence obtain that, with $\delta_3 := 1 - \sqrt{1 - \delta_2}$,

$$\left\| \sum_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} w_1 \ldots u_n^{(i_n)} w_n}(x)) f(\tau_{w_0 u_1^{(i_1)} w_1 \ldots u_n^{(i_n)} w_n})(x, \cdot) \right\|_{\ell^2(G)}$$

$$\leq \left(\max_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} w_1 \ldots u_n^{(i_n)} w_n}(x)) \right) (2(1 - \delta_2))^n [\![f]\!]_1$$

$$\leq \left(\sum_{(i_1, \ldots, i_n) \in \{1,2\}^n} \Phi_N(\tau_{w_0 u_1^{(i_1)} w_1 \ldots u_n^{(i_n)} w_n}(x)) \right) (1 - \delta_3)^n [\![f]\!]_1. \tag{8}$$

Step 3. We prove that $\delta_1 > 0$ implies that $\| \mathcal{L}_\varphi^{nmk}(f)(x, \cdot) \|_{\ell^2(G)}$ decays exponentially and show that this is a contradiction to the last statement in Proposition 5.2.

We now fix $f \in \mathcal{H}_c$ and $x \in \Sigma_A$. For $n \in \mathbb{N}$, we refer to $\mathcal{F}$ as the set of all subsets of $\{1, 2, \ldots, n\}$ and define, for each $\omega := \{k_1, k_2, \ldots, k_d\} \in \mathcal{F}$, a subset $\mathcal{V}_\omega$ of $\mathcal{W}^{nmk}$ as follows. A word $w = (w_1 \ldots w_n) \in \mathcal{W}^{nmk}$ is an element of $\mathcal{V}_\omega$ if and only if there exist $w_{k_j}^{(i_j)} \in \mathcal{W}^\dagger$, for $i_j = 1, 2$ and $j = 1, \ldots, d$, such that $w_{k_j} = w_{k_j}^{(1)}$ and

$$\left\| \sum_\star \Phi_{nmk}(\tau_v(x)) f(\tau_v(x, \cdot)) \right\|_{\ell^2(G)} \leq \left(\sum_\star \Phi_{nmk}(\tau_v(x)) \right) (1 - \delta_3)^{|\omega|} [\![f]\!]_1,$$

where $\star$ stands for the summation over all $v = (v_1 \ldots v_n)$ with $v_i = w_i$ for $i \notin \omega$ and $v_i \in \{w_i^{(1)}, w_i^{(2)}\}$ or $i \in \omega$. Observe that the construction of $u_i^{(1)}$ and $u_i^{(2)}$ above and the estimate

(8) imply that $\{\mathcal{V}_\omega : \omega \in \mathcal{F}\}$ is a covering of $\mathcal{W}^{nmk}$. Hence,

$$\mathcal{V}_\omega^* := \mathcal{V}_\omega \setminus \left(\bigcup_{\omega \subset \omega', \omega \neq \omega'} \mathcal{V}_{\omega'} \right)$$

defines a partition of $\mathcal{W}^{nmk}$. Furthermore, for $\omega := \{k_1, k_2, \ldots, k_d\}$, $j \in \{1, \ldots, d\}$ such that $k_j - 1 \notin \omega$ and $w_1 \ldots w_n \in \mathcal{V}_\omega^*$, we have

$$\sum_{v_1 \ldots v_{k_j-1} w_{k_j} \ldots w_n \in \mathcal{V}_\omega^*} \Phi_{mk}(\tau_{v_{k_j-1} w_{k_j} \ldots w_n}(x)) \leq 1 - 2\alpha,$$

since there exists at least one pair of elements $w^{(i)}$ ($i = 1, 2$) in $\mathcal{W}^\dagger$ such that $w_1 \ldots w^{(i)} w_{k_j} \ldots w_n \in \mathcal{V}_{\omega'}^*$ with $\omega' = \omega \cup \{k_j - 1\}$. Hence,

$$
\begin{aligned}
\|\mathcal{L}_\varphi^{nmk}(f)(x, \cdot)\|_{\ell^2(G)} &= \left\| \sum_{\omega \in \mathcal{F}} \sum_{w \in \mathcal{V}_\omega^*} \Phi_{nmk}(\tau_w(x)) \hat{f}(\cdot \, \psi_w^{-1}) \right\|_{\ell^2(G)} \\
&\leq \sum_{\omega \in \mathcal{F}} (1 - 2\alpha)^{n-|\omega|} (2\alpha)^{|\omega|} (1 - \delta_3)^{|\omega|} [\![f]\!]_1 \\
&= \sum_{j=0}^n \binom{n}{k} (1 - 2\alpha)^{n-k} (2\alpha(1 - \delta_3))^k [\![f]\!]_1 = (1 - 2\alpha\delta)^n [\![f]\!]_1 .
\end{aligned}
$$

By applying Jensen's inequality, we obtain $[\![\mathcal{L}_\varphi^{nmk}(f)]\!]_1 \leq (1 - 2\alpha\delta)^n [\![f]\!]_1$ which is a contradiction to $\rho = 1$ by Proposition 5.2. Hence, $\delta_1 = 0$. $\square$

We are now in a position to prove the converse to Theorem 4.1.

Theorem 5.4. *Assume that (Σ_A, θ) is a topological Markov chain with the b.i.p.-property, that (X, T) is a topologically transitive G-extension and that φ is a Hölder continuous potential with $\|L_\varphi(1)\|_\infty < \infty$. Then $P_G(T, \varphi) = P_G(\theta, \varphi)$ implies that the group G is amenable.*

Proof. We assume without loss of generality that $P_G(\theta, \varphi) = 0$. Now choose a finite subset K of G. It then follows by decomposition of T into mixing components that there exists $m \in \mathbb{N}$ such that m is a multiple of n in Lemma 5.3 and $K \subset \{\psi(v) : v \in \mathcal{W}^m\}$. In particular, there exists a finite subset $\mathcal{W}_K$ of $\mathcal{W}^m$ with

$$K = \{\psi(v) : v \in \mathcal{W}_K\}.$$

By Lemma 5.3, there exists a sequence of positive functions (f_k) in $\mathcal{H}_c$ with $\lim_{k \to \infty} [\![\mathcal{L}^m(f_k) - f_k]\!]_1 = 0$, $f_k \geq 0$ and $[\![f_k]\!]_1 = 1$ for all $k \in \mathbb{N}$. Note that this, in particular, implies that $\lim_{k \to \infty} [\![\mathcal{L}^m(f_k)]\!]_1 = 1$. In order to show that $[\![(f_k \circ \tau_v - f_k) \cdot \mathbf{1}_{\theta^m([v]) \times G}]\!]_1 \to 0$ for all $v \in \mathcal{W}_K$, assume that there exists $v \in \mathcal{W}_K$ with $\liminf_k [\![(f_k \circ \tau_v - f_k) \cdot \mathbf{1}_{\theta^m([v]) \times G}]\!]_1 > 0$. By the argument in the first step of proof of Lemma 5.3, this implies that $[\![\mathcal{L}^m(f_k)]\!]_1$ is bounded away from 1, which is a contradiction. Hence, $\liminf_k [\![(f_k \circ \tau_v - f_k) \cdot \mathbf{1}_{[\theta^m([v]), \cdot]}]\!]_1 = 0$ for all $v \in \mathcal{W}_K$, and, by taking a subsequence, we may assume without loss of generality that

$$\lim_{k \to \infty} [\![(f_k \circ \tau_v - f_k) \cdot \mathbf{1}_{\theta^m([v]) \times G}]\!]_1 = 0, \quad \forall v \in \mathcal{W}_K.$$

Furthermore, recall that the Hölder inequality implies, with $\hat{f}_k(g) := f_k(x, g)$ and $h \in G$, that

$$
\begin{aligned}
\| \hat{f}_k^2(\cdot) - \hat{f}_k^2(\cdot h) \|_1 &= \sum_{g \in G} |\hat{f}_k^2(g) - \hat{f}_k^2(gh)| \\
&= \sum_{g \in G} |\hat{f}_k(g) - \hat{f}_k(gh)| \cdot |\hat{f}_k(g) + \hat{f}_k(gh)| \\
&\leq \| \hat{f}_k(\cdot) - \hat{f}_k(\cdot h) \|_2 \cdot \| \hat{f}_k(\cdot) + \hat{f}_k(\cdot h) \|_2 \leq 2 \| \hat{f}_k(\cdot) - \hat{f}_k(\cdot h) \|_2. \quad (9)
\end{aligned}
$$

We now fix k to be specified later and use the following representation of $\hat{f}_k^2$. There exist $p \in \mathbb{N} \cup \{\infty\}$ and $\lambda_i > 0$ and $A_i \subset G$ with $A_i \subset A_{i+1}$, for $1 \leq i < p$, such that $\hat{f}_k^2 = \sum_{i=1}^p \lambda_i \mathbf{1}_{A_i}$. In particular, observe that $\sum_{i=1}^p \lambda_i |A_i| = 1$ and that $(A_i h \setminus A_i) \cap (A_j \setminus A_j h) = \emptyset$ by monotonicity of (A_i). Hence,

$$
\begin{aligned}
\left\| \hat{f}_k^2(\cdot) - \hat{f}_k^2(\cdot h) \right\|_1 &= \sum_{g \in G} \left| \sum_i \lambda_i (\mathbf{1}_{A_i}(g) - \mathbf{1}_{A_i h^{-1}}(g)) \right| \\
&= \sum_{g \in G} \sum_{i=1}^p \lambda_i \mathbf{1}_{A_i \triangle A_i h^{-1}}(g) = \sum_{i=1}^p \lambda_i \left| A_i \triangle A_i h^{-1} \right|. \quad (10)
\end{aligned}
$$

We are now in a position to prove the amenability of G. For $\epsilon > 0$, choose k such that

$$
[\![(f_k \circ \tau_v - f_k) \cdot \mathbf{1}_{[\theta^m(v), \cdot]}]\!] = \| \hat{f}_k(\cdot) - \hat{f}_k(\cdot \psi_m(v)^{-1}) \|_2 \leq \epsilon / |\mathcal{W}_K|
$$

for all $v \in \mathcal{W}_K$. Combining estimate (9) and the identity (10) then implies that

$$
\epsilon \geq \frac{1}{2} \sum_{i=1}^p \lambda_i \sum_{h \in K} |A_i h \triangle A_i|.
$$

Hence, there exists $1 \leq i \leq p$, such that $\sum_{h \in K} |A_i h \triangle A_i| \leq 2\epsilon |A_i|$. Note that the above argument shows that for each finite set K and $\epsilon > 0$, there exists a (K, ϵ)-Følner set A, that is A is finite and

$$
\sum_{h \in K} |Ah \triangle A| \leq \epsilon |A|.
$$

In order to prove amenability of G through the construction of a Følner sequence (B_n), choose a sequence of finite sets (K_n) with $K_n \nearrow G$ and assume, by induction, that B_{n+1} is a $(K_n \cup B_n, 1/n)$-Følner set. It is then easy to see that $\lim |g B_n \triangle B_n| / |B_n| = 0$ for all $g \in G$. $\quad \square$

It is worth noting that the proof is inspired by an argument in [6] where the identity (10) was used to derive a weak Følner condition. However, there is an alternative chain of arguments for the proof. For $n \in \mathbb{N}$ and $f \in \ell^2(G)$, set $Q_n(f)(g) := \sum_{v \in \mathcal{W}} \mu([v]) f(g \psi_v^{-1})$. It follows from Lemma 5.3 that $\|Q_n\| = 1$. Hence, by a result of Day ([4, Theorem 1(d)]), G is amenable. We now present two immediate applications of our results to the co-growth of groups and group extensions of Gibbs–Markov maps.

Co-growth. As a corollary of Theorems 4.1 and 5.4, we obtain the criteria for amenability in terms of the co-growth as introduced in [7]. For a set of generators $\mathcal{G} = \{\gamma_1, \gamma_1^{-1}, \ldots, \gamma_r^{-1}\}$ of G, let Σ_A be the subshift of finite type with $\mathcal{W} = \mathcal{G}$ and transition matrix (a_{gh}) given by $a_{gh} = 0$ if and only if (g, h) is equal to $(\gamma_i, \gamma_1^{-1})$ or $(\gamma_i^{-1}, \gamma_1)$ for some $i = 1, \ldots, r$. With

respect to the potential 1 and the cocycle $\psi|_{[g]} = g$, the Grigorchuk–Cohen criterion follows from $P_G(\theta) = \log(|\mathcal{G}| - 1)$. That is, G is amenable if and only if

$$\limsup_{n \to \infty} |\{w \in \mathcal{W}^n : \psi_n(w) = \mathrm{id}\}|^{1/n} = |\mathcal{G}| - 1 = 2r - 1.$$

Gibbs–Markov maps. The class of Gibbs–Markov maps was implicitly introduced in [2], and we recall its definition now. Let μ be a Borel probability measure on Σ_A such that, for all $w \in \mathcal{W}^1$, μ and $\mu \circ \tau_w$ are equivalent. Then (Σ_A, θ, μ) is called a *Gibbs–Markov map* if $\inf\{\mu(\theta([w])) : w \in \mathcal{W}^1\} > 0$ and there exists $0 < r < 1$ such that, for all $m, n \in \mathbb{N}$, $v \in \mathcal{W}^m$, $w \in \mathcal{W}^n$ with $(vw) \in \mathcal{W}^{m+n}$,

$$\sup_{x, y \in [w]} \left| \log \frac{d\mu \circ \tau_v}{d\mu}(x) - \log \frac{d\mu \circ \tau_v}{d\mu}(y) \right| \ll r^n. \tag{11}$$

Furthermore, we refer to a group extension of a Gibbs–Markov map as *weakly symmetric* if (Σ_A, θ, ψ) is symmetric and

$$\lim_{n \to \infty} \sup_{w \in \mathcal{W}^n} \sqrt[n]{\mu([w])/\mu([w^\dagger])} = 1.$$

Theorem 5.5. *Assume that (Σ_A, θ, μ) is a topologically mixing Gibbs–Markov map with the b.i.p.-property such that (X, T) is a topologically transitive G-extension.*

1. *If the group extension is weakly symmetric and G is amenable, then*

$$\limsup_{n \to \infty} (\mu(\{x \in \Sigma_A : \psi_n(x) = \mathrm{id}\})^{1/n}) = 1. \tag{12}$$

2. *If* (12) *holds, then G is amenable.*

Proof. Set $\Phi_n := d\mu/d\mu \circ \theta$, for $n \in \mathbb{N}$. Since $(d\mu/d\mu \circ \theta)(x) = (d\mu \circ \tau_v/d\mu)(\theta^n(x))$, for $x \in [v]$ and $v \in \mathcal{W}^n$, observe that the Gibbs–Markov property in (11) implies that Φ_n is of bounded variation and $\log \Phi_n$ is locally Hölder continuous. Furthermore, combining the last estimate in the proof of Proposition 5.2 and the finiteness of $\mathcal{J} \subset \mathcal{W}^k$ of Lemma 5.1, we obtain

$$\mathcal{Z}_{a,\mathrm{id}}^{n+k} \gg \mu(\{x : \psi_n(x) = \mathrm{id}\}) \gg \mathcal{Z}_{a,\mathrm{id}}^n,$$

for all $a \in \mathcal{W}^1$ and $n \in \mathbb{N}$. Hence, $P_G(T) = \limsup_n (1/n) \log \mu(\{x : \psi_n(x) = \mathrm{id}\})$ and the result follows from Theorems 4.1 and 5.4. $\square$

Note that the above result is an extension of Kesten's result to a class of random walks on groups with stationary increments.

6. An application to hyperbolic geometry

In order to apply the above results to normal covers of hyperbolic manifolds, we recall the following definition from [20]. A Kleinian group is called essentially free if there exists a Poincaré fundamental polyhedron F with faces $f_1, f_2, \ldots, f_{2n}$ and associated generators $g_1, g_2 \ldots g_{2n}$ of G with $g_i(f_i) = f_{i+n}$, $g_i^{-1}(f_{i+n}) = f_i$ and $g_i^{-1} = g_{i+n}$ for $i = 1, \ldots, n$, such that the following conditions are satisfied. In here, we refer to $\overline{(\cdot)}_{\overline{\mathbb{H}}}$ as the closure in $\overline{\mathbb{H}}$.

1. If $\overline{(f_i)}_{\overline{\mathbb{H}}} \cap \overline{\left(\bigcup_{j \neq i} f_j \right)}_{\overline{\mathbb{H}}} \neq \emptyset$ for some $i = 1, 2, \ldots, n$, then g_i, g_{i+n} are hyperbolic transformations, and $\overline{(f_{i+n})}_{\overline{\mathbb{H}}} \cap \overline{\left(\bigcup_{j \neq i+n} f_j \right)}_{\overline{\mathbb{H}}} \neq \emptyset$,

2. if $\overline{(f_i)}_{\mathbb{H}} \cap \overline{(f_j)}_{\mathbb{H}}$ is a single point p for some $j = 1, 2, \ldots, 2n$, then p is a parabolic fixed point,
3. if $f_i \cap f_j \neq \emptyset$ for some $j = 1, 2, \ldots, 2n$, then $g_i g_j = g_j g_i$.

Observe that this class comprises all non-cocompact, geometrically finite Fuchsian groups, the class of Schottky groups, and in general gives rise to geometrically finite hyperbolic manifolds which may have cusps of arbitrary rank.

We now proceed with the construction of the associated coding map. Fix a point $\mathbf{o}$ in the interior F, and denote by a_i the intersection of the shadow of f_i and the radial limit set $L_r(G)$ of G. Furthermore, denote by κ the inversion, acting on $\{f_i : i = 1, \ldots, 2n\}$, $\{g_i : i = 1, \ldots, 2n\}$ and $\{a_i : i = 1, \ldots, 2n\}$, which is defined by $\kappa f_i = f_{i+n}$, $\kappa g_i = g_{i+n}$ and $\kappa a_i = a_{i+n}$, respectively. Now let a be an atom of the partition α generated by $\{a_i : i = 1, \ldots, 2n\}$. Hence there exist $1 \le k \le 2n$ and $i_1, \ldots, i_k \in \{1, \ldots, 2n\}$ such that $a = \bigcap_{i=1}^{k} a_{i_l}$. Choose $g_a \in \{g_{i_l} : k = 1, \ldots, k\}$, and for $\kappa a := \bigcap_{i=1}^{k} \kappa a_{i_l}$, set $g_{\kappa a} = g_a^{-1}$. This then gives rise to the coding map Γ defined by

$$\Gamma : L_r(G) \to L_r(G), \theta(x) = g_a(x) \quad \text{for } x \in a, a \in \alpha$$

and a canonical symmetry given by $a^\dagger := \kappa a$. For further details of this construction, we refer to [20], where it is shown, that θ is well defined, α is a Markov partition and the underlying subshift of finite type is topologically mixing. In particular, $L_r(G)$ can be identified with a shift space Σ_A and Γ with the one-sided shift map.

In order to specify a potential adapted to the geometry of $\mathbb{H}$, recall that the Poisson kernel $\mathcal{K}$ with respect to the ball model is given by, for $x \in \partial \mathbb{H}$ and $\mathbf{o} \in \mathbb{H}$,

$$\mathcal{K}(\mathbf{o}, x) := \frac{1 - |g(\mathbf{o})|^2}{|g(\mathbf{o}) - x|^2}.$$

It is well known that $\log \mathcal{K}(g(\mathbf{o}), x)$, for $g \in G$, is equal to the orientated hyperbolic distance between $\mathbf{o}$ and the horocycle through $g(\mathbf{o})$ and x. Note that this horocyclic distance sometimes also is referred to as the Busemann cocycle. Furthermore, we recall the following for the potential given by

$$\varphi(x) = (\mathcal{K}(g_a(\mathbf{o}), x))^\delta, \tag{13}$$

for $x \in a, a \in \alpha$ and $\delta > 0$ to be specified later. If G is convex cocompact, then $\log \varphi$ is Hölder continuous with respect to the shift metric induced by Σ_A. In case that G is essentially free and contains parabolic elements, set

$$P := \overline{F} \cap \{p \in L(G) : p \text{ is a fixed point of a parabolic element in } G\}.$$

Then, if B is a subset of $L_r(G)$ which is bounded away from P and measurable with respect to some finite refinement by preimages of α, the potential associated with the first return map to B is Hölder continuous (see [10, Lemmata 3.3 & 3.4]). For ease of notation, let (Σ, θ) refer to the first return map to B if G contains parabolic elements and to (Σ_A, Γ) if G does not contain parabolic elements. Observe that in both cases the potential is Hölder continuous with respect to the shift metric, (Σ, θ) has the b.i.p.-property and can be identified with a maximal non-invertible factor of a Poincaré section of the geodesic flow on the unit tangent bundle $T^1(\mathbb{H}/G)$ of $\mathbb{H}/G$ (see [20]). Furthermore, since by construction B is bounded away from P, each return to the associated Poincaré section in $T^1(\mathbb{H}/G)$ corresponds to a return to a ball of bounded diameter with center $\mathbf{o}$ in $\mathbb{H}/G$.

Combining these observations, one then obtains the following relation between the finite words of Σ, the elements of G and the hyperbolic distance $d(\mathbf{o}, g(\mathbf{o}))$. Let $\mathcal{W}^n$ refer to the words of length n of Σ and $g_w \in G$ to the element in G defined by $\theta^n|_{[w]} = g_w$. Note that, by construction of Γ, G is isomorphic to the finite words with respect to the original partition α (see [20]). Hence, after a possible induction to B, the map $\mathcal{W}^\infty \to G$, $w \mapsto g_w$ is injective. Using the property of bounded returns, it follows that the map is almost onto in the sense that there exists a finite subset J of G such that

$$G = \bigcup_{h \in J, w \in \mathcal{W}^\infty} h g_w. \tag{14}$$

For $w \in \mathcal{W}^\infty$, set $\Phi_w := \sup_{x \in [w]} \Phi_{|w|}(x)$. As a further consequence of bounded returns, it then follows that, for $s > 0$,

$$\Phi_w^s \asymp e^{-s \delta d(\mathbf{o}, g_w(\mathbf{o}))}. \tag{15}$$

Now assume that N is a normal subgroup of G and recall that in this situation, the manifold $\mathbb{H}/N$ is called periodic with period G/N. Since $\mathbb{H}/N$ is a cover of $\mathbb{H}/G$, it follows that $\mathbb{H}/N$ is geometrically finite if and only if G/N is finite. The properties (15) and (14) above now allow relating the exponents of convergence of N and G in terms of the amenability of G/N. Therefore, recall that the Dirichlet series

$$\mathcal{P}(H, s) := \sum_{g \in H} e^{-s d(\mathbf{o}, g(\mathbf{o}))}$$

is referred to as the Poincaré series of the Kleinian group H, and that its abscissa of convergence $\delta(H)$ is called the exponent of convergence of H. In order to quantify $\mathcal{P}(N, s)$, we will now employ our results on group extensions. In order to do so, for $w \in \mathcal{W}$ and $x \in [w]$, set $\psi(x) := [g_w] \in G/N$ and

$$T : \Sigma \times G/N \to \Sigma \times G/N, \qquad (x, [g]) \mapsto (\theta(x), [g]\psi(x)).$$

In here, we only will make use of the estimates on $\mathcal{L}^n(\mathbf{1}_{X_{\mathrm{id}}})$, but it is worth noting that T is related to the geodesic flow on the periodic manifold. That is, T is a non-invertible factor of the base transformation of a Ambrose–Kakutani representation of the flow on the periodic manifold (see also, e.g. [1,18]), where the associated measure is the Liouville–Patterson measure of G.

As an application of Theorems 4.1 and 5.4 we now obtain the following partial refinement of a result of Brooks for convex-cocompact Kleinian groups in [3] to the class of essentially free Kleinian groups.

Theorem 6.1. *Let G be an essentially free Kleinian group and $N \lhd G$ a normal subgroup. Then $\delta(G) = \delta(N)$ if and only if G/N is amenable.*

Proof. Set $\delta = \delta(G)$ in the definition of φ and note that for this choice, the existence of a finite, invariant measure implies that $P_G(\theta, \varphi) = 0$ (see, e.g., [20]). Furthermore, as a consequence of (15) we have that φ is symmetric with respect to the involution generated by κ. Finally, it can easily be deduced from the connection between the group extension T and the geodesic flow on $T^1(\mathbb{H}/N)$ that T is topologically transitive. Hence, Theorems 4.1 and 5.4 are applicable and therefore, it remains to show that $P_G(T) < P_G(\theta)$ if and only if $\delta(N) < \delta = \delta(G)$.

So assume that $\delta(N) < \delta$. Hence, for $\epsilon > 0$ with $\delta(N) < (1-\epsilon)\delta$, $\mathcal{P}(N, (1-\epsilon)\delta) < \infty$. In particular, applying (14) and (15) gives that

$$\infty > \sum_{g \in N} e^{-(1-\epsilon)\delta(\mathbf{o}, g(\mathbf{o}))} \asymp \sum_{\substack{w \in \mathcal{W}^\infty, \\ [gw]=\mathrm{id}}} \Phi_w^{1-\epsilon} \geq \sum_{\substack{w \in \mathcal{W}^\infty, \\ [gw]=\mathrm{id}}} \Phi_w \|\varphi\|_\infty^{-\epsilon|w|}.$$

Since $\exp(-P_G(T))$ is the radius of convergence of $\sum_{[gw]=\mathrm{id}} \Phi_w x^{|w|}$, it follows from $\|\varphi\|_\infty < 1$ that $P_G(T) \leq \epsilon \log\|\varphi\|_\infty < 0 = P_G(\theta)$.

Now assume that $P_G(T) < 0$ and G does not contain parabolic elements. Then $\mathcal{W}$ is finite and, in particular, $\|1/\varphi\|_\infty < \infty$. Since $P_G(T) < 0$, we have that, for $1 < x < \exp(-P_G(T))$,

$$\infty > \sum_{\substack{w \in \mathcal{W}^\infty, \\ [gw]=\mathrm{id}}} \Phi_w x^{|w|} \geq \sum_{\substack{w \in \mathcal{W}^\infty, \\ [gw]=\mathrm{id}}} \Phi_w^{1-\epsilon} x^{|w|} \|1/\varphi\|_\infty^{-\epsilon|w|}$$

$$\asymp \sum_{\substack{w \in \mathcal{W}^\infty, \\ [gw]=\mathrm{id}}} e^{-(1-\epsilon)\delta(\mathbf{o}, g_w(\mathbf{o}))} x^{|w|} \|1/\varphi\|_\infty^{-\epsilon|w|}. \tag{16}$$

Hence, if $\epsilon < -P_G(T)/\log(\|1/\varphi\|_\infty)$, then $\mathcal{P}(N, (1-\epsilon)\delta) < \infty$ and, in particular, $\delta(N) < \delta$.

It remains to consider the case where $P_G(T) < 0$ and G contains parabolic elements. For $p \in P$, let G_p denote the stabiliser of p in G. By well known arguments (see, e.g., Lemma 3.2 in [21]), we have, for $s > k_p/2$ and $\ell > 0$, with k_p referring to the Abelian rank of G_p, that

$$\sum_{\substack{g \in G_p, \\ d(\mathbf{o}, g(\mathbf{o})) \geq 2\ell}} e^{-sd(\mathbf{o}, g(\mathbf{o}))} \asymp \sum_{n \geq e^\ell} \frac{1}{n^{2s-k_p+1}} \asymp \frac{1}{2s-k_p} e^{\ell(k_p-2s)}.$$

Since $\delta(G)$ is always bigger than $k_p/2$, we may choose $0 < \epsilon < 1 - k_p(2\delta)^{-1}$. For $s = (1-\epsilon)\delta$, we then have

$$\sum_{\substack{g \in G_p, \\ d(\mathbf{o}, g(\mathbf{o})) \geq 2\ell}} e^{-(1-\epsilon)\delta d(\mathbf{o}, g(\mathbf{o}))} \asymp e^{2\ell\delta\epsilon} \sum_{\substack{g \in G_p, \\ d(\mathbf{o}, g(\mathbf{o})) \geq 2\ell}} e^{-\delta d(\mathbf{o}, g(\mathbf{o}))}.$$

For $\Lambda > 0$, set $\mathcal{W}_\Lambda := \{w \in \mathcal{W}^1 : \inf_{x \in [w]} \varphi(x) \geq \Lambda\}$. It follows from the inducing process of θ, that $g_w \in G_p$ for some $p \in P$ and each $w \in \mathcal{W}_\Lambda$, if $\Lambda > 0$ is sufficiently small. The above estimate then implies the following uniform Lipschitz continuity. There exists $C \geq 1$ such that for arbitrary families $\{x_w : x \in [w], w \in \mathcal{W}_\Lambda\}$,

$$\frac{\sum_{w \in \mathcal{W}_\Lambda} \varphi(x_w)^{1-\epsilon}}{\sum_{w \in \mathcal{W}_\Lambda} \varphi(x_w)} - 1 \leq C\epsilon.$$

Combining the estimate with the argument in (16) then gives that $\delta(N) < \delta(G)$. $\quad\square$

Acknowledgments

References

[1] J. Aaronson, M. Denker, The Poincaré series of $\mathbb{C} \setminus \mathbb{Z}$, Ergodic Theory Dynam. Systems 19 (1) (1999) 1–20.
[2] J. Aaronson, M. Denker, M. Urbański, Ergodic theory for Markov fibred systems and parabolic rational maps, Trans. Amer. Math. Soc. 337 (2) (1993) 495–548.
[3] R. Brooks, The bottom of the spectrum of a Riemannian covering, J. Reine Angew. Math. 357 (1985) 101–114.
[4] M.M. Day, Convolutions, means, and spectra, Illinois J. Math. 8 (1964) 100–111.
[5] E. Følner, On groups with full banach mean value, Math. Scand. 3 (1955) 243–254.
[6] F.P. Greenleaf, Invariant Means on Topological Groups and Their Applications, in: Van Nostrand Mathematical Studies, No. 16, Van Nostrand Reinhold Co., New York, 1969.
[7] R.I. Grigorchuk, Symmetrical random walks on discrete groups, in: Multicomponent Random Systems, in: Adv. Probab. Related Topics, vol. 6, Dekker, New York, 1980, pp. 285–325.
[8] J. Jaerisch, Fractal models for normal subgroups of Schottky groups, arXiv:1106.0026v1, 2011.
[9] V.A. Kaimanovich, A.M. Vershik, Random walks on discrete groups: boundary and entropy, Ann. Probab. 11 (3) (1983) 457–490.
[10] M. Kesseböhmer, B.O. Stratmann, A multifractal formalism for growth rates and applications to geometrically finite Kleinian groups, Ergodic Theory Dynam. Systems 24 (1) (2004) 141–170.
[11] H. Kesten, Full Banach mean values on countable groups, Math. Scand. 7 (1959) 146–156.
[12] F. Ledrappier, O. Sarig, Fluctuations of ergodic sums for horocycle flows on $\mathbb{Z}^d$-covers of finite volume surfaces, Discrete Contin. Dyn. Syst. 22 (1–2) (2008) 247–325.
[13] R.D. Mauldin, M. Urbański, Gibbs states on the symbolic space over an infinite alphabet, Israel J. Math. 125 (2001) 93–130.
[14] R. Ortner, W. Woess, Non-backtracking random walks and cogrowth of graphs, Canad. J. Math. 59 (4) (2007) 828–844.
[15] T. Roblin, A Fatou theorem for conformal densities with applications to Galois coverings in negative curvature (Un théorème de Fatou pour les densités conformes avec applications aux revêtements galoisiens en courbure négative), Israel J. Math. 147 (2005) 333–357.
[16] O.M. Sarig, Thermodynamic formalism for countable Markov shifts, Ergodic Theory Dynam. Systems 19 (6) (1999) 1565–1593.
[17] O.M. Sarig, Existence of Gibbs measures for countable Markov shifts, Proc. Amer. Math. Soc. 131 (6) (2003) 1751–1758.
[18] O.M. Sarig, Invariant Radon measures for horocycle flows on Abelian covers, Invent. Math. 157 (3) (2004) 519–551.
[19] R. Sharp, Critical exponents for groups of isometries, Geom. Dedicata 125 (2007) 63–74.
[20] M. Stadlbauer, B.O. Stratmann, Infinite ergodic theory for Kleinian groups, Ergodic Theory Dynam. Systems 25 (4) (2005) 1305–1323.
[21] B. Stratmann, S.L. Velani, The Patterson measure for geometrically finite groups with parabolic elements, new and old, Proc. Lond. Math. Soc. 71 (1) (1995) 197–220.