

ERGODIC THEOREMS FOR ACTIONS OF HYPERBOLIC GROUPS

MARK POLLICOTT AND RICHARD SHARP

University of Warwick and University of Manchester

ABSTRACT. In this note we give a short proof of a pointwise ergodic theorem for measure preserving actions of word hyperbolic groups, also obtained recently by Bufetov, Khristoforov and Klimenko. Our approach also applies to infinite measure spaces and one application is to linear actions of discrete groups on the plane.

0. INTRODUCTION

The well-known Birkhoff ergodic theorem for ergodic transformations $T : X \rightarrow X$ on a probability space (X, \mathcal{B}, μ) gives that the averages of an L^1 function along almost every orbit is equal to the integral of the function. If we view the transformation as a \mathbb{Z} -action, then it is natural to ask about pointwise ergodic theorems for other group actions.

For \mathbb{Z}^d , or nilpotent groups, pointwise ergodic theorems for L^1 functions are described in the books of Krengel [16] and Templeman [23]. For groups with exponential growth, results are considerably harder to obtain. For free groups, a mean ergodic theorem for L^2 functions was proved by Guivarc'h in the 1960s [13] but it was not until the work of Nevo and Stein [19] in the 1990s that pointwise ergodic theorems were obtained. Their result applied to L^p functions with $p > 1$ but subsequently Bufetov proved the stronger pointwise ergodic theorem for L^1 functions using an approach based on Markov operators [4]. (The latter was recently extended to the fundamental groups of compact surfaces with negative Euler characteristic by Bufetov and Series [6].) The extension of these pointwise theorems to more general word hyperbolic groups for L^2 functions was carried out by Fujiwara and Nevo, under strong mixing assumptions on the action [11]. Very recently, Bufetov, Khristoforov and Klimenko [5] showed that, independent of any mixing condition, there is pointwise convergence of appropriate Cesàro averages for L^∞ functions. (They also establish L^p convergence for L^p functions, for any $p \geq 1$.) In this note, we give a short proof of the almost everywhere convergence of (slightly modified) averages for Cesàro averages for L^∞ functions. A key ingredient of our approach is an observation of Calegari and Fujiwara [7] on the structure of the Markov matrix encoding a word hyperbolic group.

Let Γ be a finitely generated group and let Γ_0 be a finite generating set. (For convenience, we always assume our generating set is symmetric, i.e. $\gamma \in \Gamma_0$ implies that $\gamma^{-1} \in \Gamma_0$.)

Definition. We define the word length $|\gamma|$ of an element $\gamma \in \Gamma - \{e\}$, with respect to the generators Γ_0 , by

$$|\gamma| = \inf \{k \geq 1 : \gamma = \gamma_1 \dots \gamma_k \text{ where } \gamma_i \in \Gamma_0, 1 \leq i \leq k\}.$$

Definition. We say that Γ is *word hyperbolic* (or Gromov hyperbolic) if, for some finite generating set Γ_0 , the following holds. Let $|\cdot|$ denote word length with respect to Γ_0 and define a metric by $d(g, h) = |g^{-1}h|$. Then there exists $\delta \geq 0$ such that every geodesic triangle in this metric is δ -thin, i.e., every point on one side of the triangle is within δ of the other two sides.

We say that a word hyperbolic group is *non-elementary* if it does not contain a finite index cyclic subgroup.

For $n \geq 1$, let $N(n) = \#\{\gamma : |\gamma| = n\}$ denote the word length counting function. We let ρ denote the exponential growth rate with respect to word length, i.e., $\rho = \lim_{n \rightarrow +\infty} N(n)^{1/n}$. This limit always exists for word hyperbolic groups and, by [9], one has the stronger estimate

$$C_1 \rho^n \leq N(n) \leq C_2 \rho^n,$$

for some $C_2 > C_1 > 0$.

Let (X, \mathcal{B}, μ) be a probability space and let $\Gamma \times X \rightarrow X$ be an action with preserves μ . We then have the following general result.

Theorem 1. *Let Γ be a non-elementary word hyperbolic group and let Γ_0 be a symmetric set of generators. For $f \in L^\infty(X, \mu)$ we have that*

$$\frac{1}{N} \sum_{n=1}^N \left(\rho^{-n} \sum_{|\gamma|=n} f(\gamma x) \right)$$

converges for μ -a.e. $x \in X$ as $N \rightarrow +\infty$.

Remark. In the general spirit of subsequence ergodic theorems, we observe that similar results hold if we average over elements with word length indexed by squares or by primes (by adapting the proof to use [14]).

Definition. For $f \in L^1(X, \mathcal{B}, \mu)$ and $x \in X$ we can then define the *spherical averages*

$$\sigma_n f(x) = \frac{1}{N(n)} \sum_{|\gamma|=n} f(\gamma x),$$

for the average over elements $\gamma \in \Gamma$ of word length $|\gamma| = n$.

In particular, if $A \in \mathcal{B}$ satisfies $\mu(A) > 0$ then with $f = \chi_A$ we can write

$$\sigma_n \chi_A(x) = \frac{1}{N(n)} \#\{\gamma \in \Gamma : \gamma x \in A \text{ and } |\gamma| = n\}.$$

It is now natural to consider the Cesàro averages of these terms.

Definition. We will say that Γ with a symmetric set of generators Γ_0 has *purely exponential growth* if there exists $C > 0$ such that $N(n) \sim C\rho^n$, as $n \rightarrow +\infty$. (Here, $a(n) \sim b(n)$, as $n \rightarrow +\infty$, means that $\lim_{n \rightarrow +\infty} a(n)/b(n) = 1$.)

We have the following results on spherical averages.

Theorem 2.

- (1) Let Γ be a non-elementary word hyperbolic group and let Γ_0 be a symmetric set of generators. There exists $p \geq 1$ such that for any $f \in L^\infty(X, \mathcal{B}, \mu)$, we have $\frac{1}{N} \sum_{n=1}^N \sigma_{pn} f(x)$ converges as $N \rightarrow +\infty$ for μ -a.e. $x \in X$.
- (2) Let Γ be a non-elementary word hyperbolic group and let Γ_0 be a symmetric set of generators with purely exponential growth. Suppose that $\Gamma \times X \rightarrow X$ is a measure preserving action of Γ on a probability space (X, \mathcal{B}, μ) . Then, for any $f \in L^\infty(X, \mathcal{B}, \mu)$, we have $\frac{1}{N} \sum_{n=1}^N \sigma_n f(x)$ converges as $N \rightarrow +\infty$ for μ -a.e. $x \in X$.

In the case of fundamental groups of compact surfaces with negative Euler characteristic the result follows from [6]. The same methods can probably be applied to noncommutative spaces [2].

The method of proof is based on the natural association to these groups of directed graphs (section 2). This then allows the reduction of the statements of the ergodic theorems to those of known results for Markov operators (section 3).

2. HYPERBOLIC GROUPS AND STRONGLY MARKOV GROUPS

Let Γ be a (non-elementary) word hyperbolic group and let Γ_0 be a finite symmetric generating set. An essential feature of these groups is that their elements can be described in terms of a suitable directed graph. More precisely, we shall use the notion of strongly Markov groups introduced by Cannon [8]. A good reference is [12].

Definition. A finitely generated group Γ is *strongly Markov* if for every symmetric set of generators Γ_0 we can find a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and an edge labelling $\lambda : \mathcal{E} \rightarrow \Gamma_0$ such that:

- (a) there exists a distinguished vertex $* \in \mathcal{V}$ with no edge terminating at $*$;
- (b) the map from the set of paths starting at $*$ to $\Gamma - \{e\}$ given by multiplying the edge labels of the path is a bijection, i.e, we associate to a path $*, v_1, \dots, v_n$ such that we may write the element $\gamma = \lambda(*, v_1) \cdots \lambda(v_{n-1}, v_n)$;
- (c) the word length of γ is n .

Proposition 1 [12, Chapitre 9, Théorème 13]. *Any word hyperbolic group is strongly Markov.*

We can associate to the graph \mathcal{G} a transition matrix A indexed by $\mathcal{V} \times \mathcal{V}$ i.e., $A(v, v') = 1$ if $(v, v') \in \mathcal{E}$, and 0 otherwise.

Definition. A non-negative square matrix M is called *irreducible* if for all i, j there exists $n \geq 1$ such that $M^n(i, j) > 0$. It is called *aperiodic* if there exists $n \geq 1$ such that, for all i, j , $M^n(i, j) > 0$.

In the case of surface groups or even corners Kleinian groups, we have the following result. (A Kleinian group Γ is said to satisfy the even corners condition if Γ admits a fundamental domain R which is a finite sided polyhedron, possibly with infinite volume, such that $\bigcup_{\gamma \in \Gamma} \gamma \partial R$ is a union of hyperplanes.)

Proposition 2. *If Γ is the fundamental group of a compact surface with negative Euler characteristic, or, more generally, a convex co-compact Kleinian group satisfying the even corners condition then the matrix obtained from A by deleting the row and column indexed by the vertex $*$ is aperiodic.*

This follows from the existence, in these cases, of a Markov mapping on the ideal boundary originally introduced to code geodesics. In the case of surface groups, this follows from the work of Series [22] and Adler and Flatto [1], while the higher dimensional case was discussed by Bourdon [3]. With the exception of the vertex $*$, the vertices of the graph \mathcal{G} correspond to the regions of continuity of the map. We join two vertices with an edge if the image of the region corresponding to the first vertex contains the region corresponding to the second vertex. The important property of the Markov map is that it is locally eventually onto, which immediately implies the aperiodicity of the required matrix.

Remark (Extension to L^1 functions). Under the additional hypothesis that A is aperiodic then the conclusions of Theorem 1 and Theorem 2 hold for any $f \in L^1(X, \mathcal{B}, \mu)$, as is apparent from the proofs in the next section.

3. MARKOV OPERATORS AND ERGODIC THEOREMS

Let (Y, \mathcal{A}, ν) be a probability space. Let $\mathbb{P}(y, A)$ be a transition probability function, where $y \in Y$ and $A \in \mathcal{A}$, such that

- (1) $y \mapsto \mathbb{P}(y, A)$ is measurable, for each $A \in \mathcal{A}$;
- (2) $\mathbb{P}(y, \cdot) : \mathcal{A} \rightarrow \mathbb{R}$ is a probability measure for each $y \in Y$.

Let $\mathcal{P} : L^1(Y, \mathcal{A}, \nu) \rightarrow L^1(Y, \mathcal{A}, \nu)$ be the positive linear operator defined by

$$\mathcal{P}f(y) = \int f(\xi) d\mathbb{P}(\xi, \cdot)$$

In particular, \mathcal{P} is a *Markov operator*, i.e., $\mathcal{P}1 = 1$.

Definition. Let $\mathcal{I} \subset \mathcal{A}$ be the sigma algebra of sets A such that $y \in A$ iff

$$\mathbb{P}(y, A) = \begin{cases} 1 & \text{if } y \in A \\ 0 & \text{if } y \notin A \end{cases}$$

The following version of the Birkhoff ergodic theorem for Markov operators appears in Rosenblatt's book [20, p. 93] (cf. also [10]).

Proposition 3. *For $f \in L^1(Y, \mathcal{A}, \nu)$ we have that*

$$\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{P}^n f \rightarrow E(f|\mathcal{I}), \quad \nu\text{-a.e.}, \quad \text{as } N \rightarrow +\infty.$$

In order to apply this abstract theorem, we require the following result.

Lemma 1. *For non-elementary hyperbolic groups we can write*

$$A = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ C_{12} & B_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ C_{1l} & C_{2l} & \cdots & B_l \end{pmatrix}$$

where

- (1) the diagonal blocks B_1, \dots, B_l are irreducible matrices;

- (2) a subcollection B_{i_1}, \dots, B_{i_l} have spectral radius $\rho(B_{i_j}) = \rho$; and
 (3) for any i_r, i_s , for any $n \geq 1$ and any j_1, \dots, j_{n-1} we have that

$$C_{i_1, j_1} C_{j_1, j_2} \cdots C_{j_{n-1}, i_s} = 0,$$

i.e., there is no transition from the component represented by B_{i_r} to the component represented by B_{i_s} .

Remark. Parts (1) and (2) of Lemma 1 are standard results for non-negative matrices [21]. As observed in [7], part (3) is based on the fact that there always exists $C_1, C_2 > 0$ such that $C_1 \rho^n \leq N(n) \leq C_2 \rho^n$, for $n \geq 0$ [9]. In particular, if the conclusion of the lemma was not true then $N(n)$ would have an additional polynomial factor.

We recall the following lemma (cf. [15, Theorem 1.3.5], [21, Theorem 1.5]).

Lemma 2. *There is an eigenvector w_{i_r} with strictly positive entries such that $B_{i_r} w_{i_r} = \rho w_{i_r}$.*

Let $\mathcal{V}_{i_j} \subset \mathcal{V}$ be the indices corresponding to the block B_{i_j} . Let $\widehat{X}_{i_j} = X \times \mathcal{V}_{i_j}$ and $\widehat{\mu}_{i_j} = \mu \times \left\{ \frac{1}{|\mathcal{V}_{i_j}|}, \dots, \frac{1}{|\mathcal{V}_{i_j}|} \right\}$.

Definition. We can define an operator $\mathcal{Q}_{i_j} : L^1(\widehat{X}_{i_j}, \widehat{\mu}_{i_j}) \rightarrow L^1(\widehat{X}_{i_j}, \widehat{\mu}_{i_j})$ by

$$\mathcal{Q}_{i_j} F(x, v) = \sum_{v'} B_{i_j}(v, v') F(\lambda(v, v')x, v').$$

To keep track of the transitions between the rest of the vertices, we let A_0 be the matrix derived from A by setting to zero the entries corresponding the maximal components.

Let $\widehat{X} = X \times \mathcal{V}$ equipped with the product σ -algebra $\widehat{\mathcal{B}}$ and a measure $\widehat{\mu} = \mu \times \{1/|\mathcal{V}|, \dots, 1/|\mathcal{V}|\}$ on \widehat{X} .

Definition. We define an operator $\mathcal{R} : L^1(\widehat{X}, \widehat{\mu}) \rightarrow L^1(\widehat{X}, \widehat{\mu})$ by

$$\mathcal{R}F(x, v) = \sum_{v'} A_0(v, v') F(\lambda(v, v')x, v').$$

The following lemma is immediate.

Lemma 3. *The operator \mathcal{R} has spectral radius $\rho(\mathcal{R}) < \rho$.*

Let $\mathcal{V}_0 = \{v \in \mathcal{V} : A(*, v) = 1\}$ be the set of vertices which can be reached in one step from $*$. We require the following observation.

Lemma 4. *If we let $F(x, v) = f(x)$ then for $n \geq 0$,*

$$\sum_{|\gamma|=n+1} f(\gamma x) = \sum_{v \in \mathcal{V}_0} \sum_{j=1}^k \sum_{l=0}^n \sum_{d_0+d_1=n-l} \mathcal{R}^{d_1} \mathcal{Q}_{i_j}^l \mathcal{R}^{d_0} F(x, v) \quad (1)$$

This follows from a simple enumeration of the possible words that can occur, in light of Lemma 1. In particular, the words can have at most one subword in

a maximal component (of length l), preceded by a subword of length $d_0 \geq 0$ and followed by a subword of length $d_1 \geq 0$ outside of the maximal components.

Define a stochastic matrix P_{i_j} by $P_{i_j}(v, v') = \rho^{-1} B_{i_j}(v, v') w(v')/w(v)$ for $v, v' \in \mathcal{V}_{i_j}$.

Definition. We define a *normalized operator* $\mathcal{P}_{i_j} : L^1(\widehat{X}_{i_j}, \widehat{\mu}_{i_j}) \rightarrow L^1(\widehat{X}_{i_j}, \widehat{\mu}_{i_j})$ by

$$\mathcal{P}_{i_j} F(x, v) = \sum_{v'} P_{i_j}(v, v') F(\lambda(v, v')x, v').$$

The operator \mathcal{P}_{i_j} is Markov.

Let $\Delta(w_{i_j}) : L^1(\widehat{X}_{i_j}, \widehat{\mu}_{i_j}) \rightarrow L^1(\widehat{X}_{i_j}, \widehat{\mu}_{i_j})$ denote multiplication by w_{i_j} . We then have the following.

Lemma 5. For $l \geq 0$,

$$\mathcal{Q}_{i_j}^l = \rho^l \Delta(w_{i_j}) \mathcal{P}_{i_j}^l \Delta(w_{i_j})^{-1}. \quad (2)$$

Proof. We can write

$$\rho^n w_{i_j}(v_1) P_{i_j}(v_1, v_2) \cdots P_{i_j}(v_{k-1}, v_k) \cdot w_{i_j}^{-1}(v_k) = B_{i_j}(v_1, v_2) \cdots B_{i_j}(v_{k-1}, v_k).$$

Thus we have

$$\begin{aligned} & \mathcal{Q}_{i_j}^l F(x, v) \\ &= \sum_{v_1, \dots, v_l} B_{i_j}(v_1, v_2) \cdots B_{i_j}(v_{l-1}, v_l) F(\lambda(v_1, v_2) \cdots \lambda(v_{l-1}, v_l)x) \\ &= \sum_{v_1, \dots, v_l} w_{i_j}(v) P_{i_j}(v_1, v_2) \cdots P_{i_j}(v_{l-1}, v_l) w_{i_j}(v_l)^{-1} F(\lambda(v_1, v_2) \cdots \lambda(v_{l-1}, v_l)x) \\ &= \rho^l w_{i_j}(v) \mathcal{P}_{i_j}^l (F(\cdot, \cdot)/w_{i_j}(\cdot)) (x, v), \end{aligned}$$

where the summation is over all strings of indices such that

$$B_{i_j}(v, v_1) = B_{i_j}(v_1, v_2) = \cdots = B_{i_j}(v_{n-1}, v_n) = 1.$$

This completes the proof. \square

Proof of Theorem 1. Assume that $f \in L^\infty(X, \mu)$. It is clear that Theorem 1 is equivalent to the μ -a.e. convergence of

$$A_N(x) := \frac{1}{N} \sum_{n=1}^N \rho^{-n} \left(\sum_{|\gamma|=n+1} f(\gamma x) \right).$$

Using Lemma 4, we have that

$$A_N(x) = \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{l=0}^n \sum_{d_0+d_1=n-l} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) \mathcal{P}_{i_j}^l \Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F(x, v) \right).$$

Fix $m \geq 1$. Then, for $n > m$, we can split this summation as

$$\begin{aligned} & \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{r=0}^m \sum_{d_0+d_1=r} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) \mathcal{P}_{i_j}^{n-r} \Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F(x, v) \right) \\ & + \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{l=0}^{n-m-1} \sum_{d_0+d_1=n-l} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) \mathcal{P}_{i_j}^l \Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F(x, v) \right). \end{aligned} \quad (3)$$

For the first term in (3) we have that

$$\begin{aligned} & \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{r=0}^m \sum_{d_0+d_1=r} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) \mathcal{P}_{i_j}^{n-r} \Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F(x, v) \right) \\ & = \lim_{N \rightarrow +\infty} \sum_{r=0}^m \sum_{d_0+d_1=r} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) \left(\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{P}_{i_j}^{n-r} \right) \Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F(x, v) \\ & = G_m(x) := \sum_{r=0}^m \sum_{d_0+d_1=r} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) E \left(\Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F | \widehat{\mathcal{I}} \right) (x, v), \end{aligned}$$

by Proposition 3. Noticing that

$$G_m(x) = O \left(\sum_{r=0}^m \sum_{d_0+d_1=r} \|\rho^{-1}\mathcal{R}\|^r \|F\|_\infty \right) = O \left(\sum_{r=0}^m r \theta^r \right),$$

for μ -a.e. $x \in X$, where $\|R/\rho\| < \theta < 1$, we see that $G_m(x)$ converges μ -a.e., as $m \rightarrow +\infty$.

We can bound the second term in (3) using

$$\begin{aligned} & \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{l=0}^{n-m-1} \sum_{d_0+d_1=n-l} (\rho^{-1}\mathcal{R})^{d_1} \Delta(w_{i_j}) \mathcal{P}_{i_j}^l \Delta(w_{i_j})^{-1} (\rho^{-1}\mathcal{R})^{d_0} F(x, v) \right) \\ & = O \left(\sum_{l=0}^{n-m-1} \sum_{d_0+d_1=n-l} \|\rho^{-1}\mathcal{R}\|^{d_0+d_1} \|F\|_\infty \right) \\ & = O(m\theta^m). \end{aligned} \quad (4)$$

Given $\epsilon > 0$ we can choose m sufficiently large such that that expression in (4) is strictly smaller than ϵ for μ -a.e. $x \in X$. \square

Proof of Theorem 2. We can prove part (1) of Theorem 2 by studying the proof of Theorem 1. We choose p such that $B_{i_j}^p > 0$ for each $j = 1, \dots, k$. In particular, the matrix A^p can be used to model strings whose lengths are multiples of p .

We can deduce part (2) of Theorem 2 from Theorem 1 and the definition of purely exponential growth.

Remark (Sigma Finite Measure Spaces). Modified ‘‘ratio’’ versions of our results apply for the actions of a non-elementary word hyperbolic group on spaces with infinite measure. More precisely, suppose that $\Gamma \times X \rightarrow X$ is a measure preserving

action on a σ -finite measure space (X, \mathcal{B}, μ) . The analogue of Theorem 1 is that, for any $f, g \in L^1(X, \mathcal{B}, \mu)$ with $g > 0$,

$$\frac{\sum_{n=0}^{N-1} \rho^{-n} \sum_{|\gamma|=n} f(\gamma x)}{\sum_{n=0}^{N-1} \rho^{-n} \sum_{|\gamma|=n} g(\gamma x)}$$

converges for μ -a.e. $x \in X$, as $N \rightarrow +\infty$. In the particular case of the fundamental group of a compact surface group of genus $g \geq 2$ with the standard presentation and an ergodic action the limit is $\int f d\mu / \int g d\mu$. The analogue of Theorem 2 also holds.

Application to Linear actions. Consider a discrete co-compact subgroup Γ of the group $SL(2, \mathbb{C})$. There is a natural linear action of Γ on \mathbb{C}^2 given by

$$\gamma : (x_1, x_2) \mapsto (ax_1 + bx_2, cx_1 + dx_2), \text{ where } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

In [18] it was shown that orbits ordered by word length are distributed according to the measure $d\lambda / \|(x_1, x_2)\|_2$, where λ denotes Lebesgue measure on the plane (cf. also [17]). For any point $(x_1, x_2) \in \mathbb{C}^2 - \{(0, 0)\}$, and any functions $f, g \in L^1(\mathbb{C}^2, \lambda)$ with $g > 0$,

$$\frac{\sum_{n=0}^{N-1} \rho^{-n} \sum_{|\gamma|=n} f(\gamma \cdot (x_1, x_2))}{\sum_{n=0}^{N-1} \rho^{-n} \sum_{|\gamma|=n} g(\gamma \cdot (x_1, x_2))}$$

converges for λ -a.e. $x \in \mathbb{C}^2$, as $N \rightarrow +\infty$.

REFERENCES

1. R. Adler and L. Flatto, *Geodesic flows, interval maps, and symbolic dynamics*, Bull. Amer. Math. Soc. **25** (1991), 229–334.
2. C. Anantharaman-Delaroche, *On ergodic theorems for free group actions on noncommutative spaces*, Probability Theory and Related Fields **135** (2006), 520–546.
3. M. Bourdon, *Actions quasi-convexes d'un groupe hyperbolique, flot géodésique*, Thesis, Orsay 1993.
4. A. Bufetov, *Markov averaging and ergodic theorems for several operators*, Topology, ergodic theory, real algebraic geometry, Amer. Math. Soc. Transl. Ser. 2, 202, Amer. Math. Soc., Providence, RI, 2001, pp. 39–50.
5. A. Bufetov, M. Khristoforov and A. Klimenko, *Cesàro convergence of spherical averages for measure-preserving actions of Markov semigroups and groups*, preprint, arXiv:1101.5459 (2011).
6. A. Bufetov and C. Series, *A pointwise ergodic theorem for Fuchsian groups*, to appear, Math. Proc. Camb. Phil. Soc..
7. D. Calegari and K. Fujiwara, *Combable functions, quasimorphisms, and the central limit theorem*, Ergodic Theory Dynam. Systems **30** (2010), 1343–1369.
8. J. Cannon, *The combinatorial structure of cocompact discrete hyperbolic groups*, Geom. Dedicata **16** (1984), 123–148.
9. M. Coornaert, *Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov*, Pac. J. Math. **159** (1993), 241–270.
10. B. Farkas, T. Eisner, M. Hasse and R. Nagel, *Ergodic Theory - An Operator-theoretic approach*, 12th International Internet Seminar.
11. K. Fujiwara and A. Nevo, *Maximal and pointwise ergodic theorems for word-hyperbolic groups*, Ergod. Th. Dynam. Sys. **18** (1998), 843–858.
12. E. Ghys and P. de la Harpe, *Sur les groupes hyperboliques d'après Mikhael Gromov*, Progr. Math., 83, Birkhäuser, Boston, 1990.

13. Y. Guivarc'h, *Généralisation d'un théorème de von Neumann*, C. R. Acad. Sci. Paris Sr. A **268** (1969), 1010-1013.
14. R. Jones, J. Olsen, M. Wierdl, *Subsequence ergodic theorems for L^p contractions*, Trans. Amer. Math. Soc. **331** (1992), 837-850.
15. B. Kitchens, *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts*, Universitext, Springer-Verlag, Berlin, 1998.
16. U. Krengel, *Ergodic Theorems*, de Gruyter Studies in Mathematics, 6, Walter de Gruyter, Berlin-New, 1985.
17. F. Ledrappier, *Distribution des orbites des réseaux sur le plan réel plane*, C. R. Acad. Sci. Paris Sr. I Math **329** (1999), 61-64.
18. F. Ledrappier and M. Pollicott, *Ergodic properties of linear actions of 2×2 matrices*, Duke Math. J. **116** (2003), 353-388.
19. A. Nevo and E. Stein, *A generalization of Birkhoff's pointwise ergodic theorem*, Acta Math. **173** (1994), 135-154.
20. M. Rosenblatt, *Markov processes: structure and asymptotic behavior*, Die Grundlehren der mathematischen Wissenschaften, Band 184, Springer-Verlag,, New York, 1971.
21. E. Seneta, *Non-negative matrices and Markov chains*, Revised reprint of the second (1981) edition. Springer Series in Statistics, Springer, New York, 2006.
22. C. Series, *Geometrical Markov coding of geodesics on surfaces of constant negative curvature*, Ergod. Th. Dynam. Sys. **6** (1986), 601-625.
23. A. Templeman, *Ergodic theorems for group actions*, Kluwer Acad. Publ., Dordrecht, 1992.

MARK POLLICOTT, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WARWICK, COVENTRY, CV8 2AH, UK

E-mail address: mpollic@maths.warwick.ac.uk

RICHARD SHARP, SCHOOL OF MATHEMATICS, UNIVERSITY OF MANCHESTER, OXFORD ROAD, M13 9PL, UK

E-mail address: sharp@maths.man.ac.uk