On the convergence of spherical multiergodic averages for Markov action groups

A.M. Mesón and F. Vericat

Abstract. In this note we study the norm convergence of multiple ergodic spherical averages from actions of word hyperbolic groups or more generaly strongly Markov groups. The objective is the generalization of the results by Bufetov, Khristoforov and Klimenko and by Pollicott and Sharp. We use the techniques that Walsh introduced for proving the norm convergence of multiple ergodic averages from measure-preserving transformations and for integer valuated polynomials.

M.S.C. 2010: 37C85
Key words: Multiergodic averages; Markov groups; complexity of systems.

1 Introduction

The classical Birkhoff ergodic theorem on a probability space $(X, \mathcal{B}, \mu)$ states the convergence of averages of real valuated maps on X along orbits of measure preserving transformations $T : X \to X$. One important subject is the extension of the ergodic theorem with action groups as dynamics. Among the most relevant contributions about the pointwise convergence of ergodic averages for action groups the following works can be mentioned. In [12] Nevo and Stein proved the pointwise ergodic convergence for finite measure preserving actions of the free group $F_{r,}$ $r \geq 2$,and in [5] Fujiwara established an ergodic convergence for exponentially mixing actions of word-hyperbolic groups. Lindenstrauss[11] considered actions of amenable groups Γ on a Lebesgue space (X, μ) and maps $\varphi : X \to \mathbf{R}$, to analyze averages of the form

$$\frac{1}{m_{\Gamma}(E)} \int_{E} \varphi\left(\gamma x\right) dm_{\Gamma}(\gamma),$$

where $E \subset \Gamma$ and m_{Γ} is the Haar measure on Γ. In that article was proved the convergence of these averages under the condition of the existence of adequate sequences (F_n) in Γ, called Følner sequences. Bufetov[2] proved an ergodic theorem of spherical averages for free semigroups. Later Bufetov and Series[4] applied this result to prove the ergodic convergence for a class of Fuchsian groups which includes surface groups, i.e. the fundamental groups of surfaces of genus $g \geq 2$.

Applied Sciences, Vol.17, 2015, pp. 62-74.

Let Γ be a finitely generated Markov groups Γ and let $S(n)$ be the sphere of radius n word metric, i.e.

$$S(n) = \{\gamma : |\gamma| = n\},$$

where $|\gamma|$, is the minimal number of generators needed to represent γ. If $\gamma \in \Gamma$, then by T_γ is denoted the transformation in X given by $T = T_\gamma(x) = \gamma x$, now for a map $\varphi \in L^\infty(X)$, can be considered the following average

$$\mathcal{S}_{N,}\varphi(x) := \frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}\varphi\left(T_\gamma(x)\right)$$

(When $cardS(n) = 0$ then we set $\mathcal{S}_{N,}\varphi(x) = 0$). Recently Bufetov, Khristoforov and Klimenko[3] analyzed these spherical averages and proved the pointwise convergence of Cesáro averages for the spherical sequence $\{\mathcal{S}_{N,}\varphi(x)\}$ .The word-hyperbolic groups are Markov groups, for this special case Pollicott and Sharp [15] gave a more concise proof of the main result of [3].

Multiple ergodic averages appeared as a dynamical version of the Szemeredi theorem in combinatorial number theory. The analogy was made by Furstenberg[6] who studied ergodic averages in a measure-preserving probability space $(X, \mathcal{B}, \mu, T)$ of the form

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu\left(A\cap T^{-n}A\cap ...\cap T^{-kn}A\right), \tag{1.1}$$

where $A \in \mathcal{B}$ and $j \in \mathbf{N}$. Furstenberg established that if $\mu(A) > 0$ then

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu\left(A\cap T^{-n}A\cap ...\cap T^{-jn}A\right) > 0.$$

This relevant result serves to prove by arguments from Ergodic Theory the Szemeredi theorem, which states that if S is a set of integers with positive upper density then S contains arithmetic progressions of arbitrary length. After this the task was the study of the convergence of averages like

$$\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{j}\varphi_i\left(T^{in}\right), \tag{1.2}$$

where $T : X \to X$ is an invertible measure preserving transformation and $\varphi_1, \varphi_2, ..., \varphi_j \in L^\infty(X)$. Furstenberg proved the L^2-convergence for the case $j = 2$, of course for the case $j = 1$ is obtained the L^2- Von Neumann ergodic theorem. For any j, Host and Kra have demonstrated the convergence in $L^2(X)$ of the multiergodic averages (1.2). The machinery used was mainly the theory of factors developed by Host-Kra-Ziegler and the approximation of multiergodic averages by averages on nyl-systems. A more general problem is to consider the following averages

$$\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{j}\varphi_i\left(T^{p_i(n)}x\right), \tag{1.3}$$

where p_i are polynomials valuated in $\mathbf{Z}$. The average (1.3) corresponds to the case of linear polynomials. Bergelson [1] proved the existence of $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{j}\varphi_i\left(T^{p_i(n)}\right)$ under the assumption that T ergodic. Later Leibman[10] proved convergence without assuming ergodicity. A more ambitious problem was to consider a nilpotent group G of measure preserving transformations of a probability space (X,μ) and study the norm convergence of averages

$$\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{j}T_1^{p_{i,1}(n)}...T_\ell^{p_{i,,\ell}(n)}\varphi_i, \tag{1.4}$$

where $T_1,...,T_\ell \in G$ and $p_{i,1},...,p_{i,\ell} \in \mathbf{Z}[n]$. The average (1.3) corresponds to the case of $T_1 = ... = T_\ell$, i.e. when G is a cyclic group. The L^2-convergence of (1.4) was recently proved by Walsh[16], who used for the demonstration an argument based doing induction on the complexity of systems of polynomial sequences in G. He also adapted results by a decomposition theory of functions developed by Gowers[8]. Another contribution in this direction is the work of Zorin-Kranich[17] who proved, using the Walsh argument, the convergence in norm of averages along Følner sequences in an amenable group.

The aim of this article is to study multiple ergodic spherical averages for hyperbolic, or more generally Markov action groups, in order to obtain a multiergodic version of the results in [3] and [15]. The novelty of our work with respect to the Walsh one is that we consider nonconmutative action groups and with respect to the Zorin-Kranich article is that we have no need to impose the existence Følner sequences.

The result to be proved herein is

Theorem 1.1: Let (X,μ) be a measure space and let Γ be a finitely generated Markov group. Let us consider actions h_1, h_2, ..., h_j of Γ on X, i.e. for any $\gamma\in\Gamma$ are defined measure-preserving maps $h_i(\gamma) : X \to X$, $i = 1,2,..,j$. If the actions commute in the sense that $h_i(\gamma_2)\,h_k(\gamma_1) = h_k(\gamma_1)\;\; h_i(\gamma_2)$, for any $\gamma_1,\gamma_2\in\Gamma$ and $i \neq k$ then for any $\varphi_1,\varphi_2,...,\varphi_j \in L^\infty(X)$, the multiergodic sequence

$$\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}\varphi_1\left(h_1(\gamma)x\right)\varphi_2\left(h_1(\gamma)h_2(\gamma)x\right)...\varphi_j(h_1(\gamma)...h_j(\gamma)x), \tag{1.5}$$

converges in L^2-norm.

In order to apply the Walsh techniques notice the following, if h_i, $i = 1,2,...,j$ is a measure-preserving action of Γ on X then can be defined maps $g_i : \Gamma \to G$, where G is a group of unitary operators on a space of maps, for our purposes we can take G a nilpotent group of unitary operators on $L^2(X)$. Thus each map g_i, $i = 1,2,...,j$, is defined as $g_i(\gamma)[\varphi](x) = \varphi(h_i(\gamma)(x))$. The maps g_i are ¨antihomomorphisms¨, i.e. $g_i(\gamma_1\gamma_2) = g_i(\gamma_2)\,g_i(\gamma_1)$. Thus Theorem 1.1 can be proved from this more general result:

Theorem 1.2. Let (X,μ) be a measure space and let Γ be a finitely generated Markov group. Let $g_1,\ g_2,...,\ g_j\ :\Gamma\rightarrow G$, with G a nilpotent group of unitary operators on $L^2(X)$. If the system $\mathbf{g}=(g_1,g_2,...,g_j)$ has finite complexity then the sequence

$$\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}g_1(\gamma)\varphi_1 g_2(\gamma)\varphi_2...g_j(\gamma)\varphi_j \tag{1.6}$$

converges in L^2-norm, for any $\varphi_1,\varphi_2,...,\varphi_j\in L^\infty(X)$.

The concept of complexity of a system of polynomial sequences was given by Walsh [16] who considered the maps $g_1,\ g_2,...,\ g_j:\mathbf{Z}\rightarrow G$, with

$$g_i(n)=T_1^{p_{i,1}(n)}...T_\ell^{p_{i,,\ell}(n)}, \tag{1.7}$$

where $T_1,...,T_\ell\in G$ and $p_{i,1},...,p_{i,\ell}\in\mathbf{Z}[n]$. The definition of complexity can be extended for non commutative groups. The innovative argument of [16] consists in using the complexity as induction parameter, so that is necessary that the system of maps $g_1,g_2,...,g_j$ have finite complexity. Systems with finite complexity can be reduced to the trivial system (1_G) in a finite number of steps. The polynomial systems are special systems of maps, for instance if $\Gamma=\mathbf{Z}$, then $g(n)=T_1^{p_1(n)}...T_\ell^{p,\ell(n)}$ with $T_1,...,T_\ell$ measure-preserving transformations and each $p_i\in\mathbf{Z}[n]$. Polynomial systems have finite complexity, this was proved for $\Gamma=\mathbf{Z}$ in [16], and for non commutative Γ in [17].

2 Markov groups

Let Γ be a group with a finite symmetric set of generators S. If $\mathcal{G}$ is a directed graph then $E[\mathcal{G}]$ denotes the set of edges of $\mathcal{G}$ and $P[\mathcal{G},v]$ denotes the set of finite paths in $\mathcal{G}$ starting in the vertex v. The group Γ is *a Markov group* if there is a labelling map

$$\lambda:E[\mathcal{G}]\rightarrow S,$$

and a distinguished vertex v_0, such that

i) λ can be lifted to a map $\lambda:P[\mathcal{G},v_0]\rightarrow\Gamma$ which is a bijection, so that for to a path of vertices $v_0,v_1,...,v_n$ can be assigned the value $\lambda(v_0v_1)\lambda(v_1v_2)...\lambda(v_{n-1}v_n)$.

ii) if γ is represented by the labelling $\lambda(v_0v_1)\lambda(v_1v_2)...\lambda(v_{n-1}v_n)$ then $|\gamma|=n$.

The word metric in a group Γ, with a finite symmetric set of generators S, is defined by $d(\gamma_1,\gamma_2)=\left|\gamma_1^{-1},\gamma_2\right|$, where $|\gamma|=|\gamma|_S$ is the minimal number of elements of S needed to represent γ. A group Γ is *word hyperbolic*, in Gromov sense, if there is a number $\delta>0$ if in any geodesic triangle every point in one side of it is within distance δ of the other sides. A group is hyperbolic if and only if its Cayley graph is hyperbolic.

Any word-hyperbolic group is a Markov group[7].

Recall that $S(n)=\{\gamma:|\gamma|=n\}$, a known fact is that a word hyperbolic group has exponential growth rate with respect to the word length, i.e. if $\rho=\lim_{n\rightarrow\infty}\frac{1}{n}\log cardS(n)$

then there are constants $C_2 > C_1 > 0$

$$C_1\rho^{-n} \leq cardS(n) \leq C_2\rho^{-n}.$$

Thus for ergodic averages for word hyperbolic actions groups can be considered weights ρ^{-n} instead of $\dfrac{1}{cardS(n)}$.

3 Complexity of systems

Let $\mathbf{g} = (g_1, g_2, ..., g_j)$ be an ordered tuple of maps from Γ to G, which is called a *system.* If $a, b \in \Gamma$ then derivative with respect to (a, b) of map $g : \Gamma \to G$ at the point γ is

$$D_{(a,b)}(g)(\gamma) := g(\gamma)^{-1} g(a\gamma b), \tag{3.1}$$

and $D_{(a,b)}\left(g^{-1}\right)(\gamma) := g(\gamma)g(a\gamma b)^{-1}$. Let $g, h : \Gamma \to G$, and set

$$\langle g, h\rangle_{(a,b)}(\gamma) := D_{(a,b)}\left(g^{-1}\right)(\gamma)T_{(a,b)}(h)(\gamma), \tag{3.2}$$

where $T_{(a,b)}(h)(\gamma) = h(a\gamma b)$. The $(a, b)-$*reduction* of the system $\mathbf{g} = (g_1, g_2, ..., g_j)$ is defined to be the system

$$\mathbf{g}^*_{(a,b)} = \left(g_1, g_2, ..., g_{j-1}, \langle g_j, g_1\rangle_{(a,b)}, ... \langle g_j, g_{j-1}\rangle_{(a,b)}\right). \tag{3.3}$$

.

The concept of complexity is given by induction in the following way. A system constituted only by the identity 1_G map has complexity ≤ 0, it is denoted by *cplx* $(\{1_G\}) \leq 0$. If *cplx* $\left(\mathbf{g}^*_{(a,b)}\right) \leq C - 1$, for any $a, b \in \Gamma$ then *cplx*$(\mathbf{g}) \leq C$. The property known as *cheating* says that if are added constants, repeated maps or repeated maps multiplied by a constant to a system it does not change its complexity. For example *cplx*$(g_1, c_1g_1, c_1, g_1, g_2, c_2g_2) =$ *cplx*(g_1, g_2).

A *filtration* on a group G is a sequence of subgroups

$$G_\bullet = G_0 \geq G_1 \geq ... \geq ...,$$

such that $G_0 = G$ and $[G_i, G_i] \subset G_{i+j}$. The length of the filtration is d when G_{d+1} is the identity. Let us denote by $G_{\bullet+t}$ the filtration given by $(G_{\bullet+1})_t = G_{i+t}$. If $G_\bullet$ has length d then $G_{\bullet+t}$ has length $d - t$.

A map $g : \Gamma \to G$ is *polynomial with respect to a filtration* $G_\bullet$ on G of length $d \in N \cup \{-\infty\}$, or just $G_\bullet-$polynomial, if $g = 1_G$, so $d = -\infty$ or $D_{(a,b)}(g)$ is $G_{\bullet+1}-$polynomial. When G is a nilpotent group a map $g : \Gamma \to G$ is a polynomial of scalar degree $\leq d$ if for any $a_1, b_1, ..., a_{d+1}, b_{d+1} \in \Gamma$

$D_{(a_1,b_1)}...D_{(a_{d+1},b_{d+1})}(g) \equiv 1_G$. If G is not nilpotent then the set of polynomials of degree$\leq d$, may be not a group.

The system $\mathbf{g}=(g_1,g_2,...,g_j)$, where $g_i:\Gamma\to G$ are polynomial maps and G is nilpotent, has finite complexity. For more details and justification of the mentioned results see[17].

The following concept was introduced by Walsh. We present herein an adapted definition for Markov groups. Let $\sigma\in L^{\infty}(X)$, with $\|\sigma\|_{\infty}\le 1$, $L\in\mathbf{N}$, $0<\varepsilon<1$. The function σ is *reducible* with respect to L and to a system $\mathbf{g}=(g_1,g_2,...,g_j)$, or just $(L,\mathbf{g})-$reducible if there is a number $M>0$ and functions $b_0,b_1,...,b_{j-1}\in L^{\infty}(X)$, $\|b_i\|_{\infty}\le 1$, $i=0,1,...,j-1$, such that fore any $\ell\le L$ and for any $\gamma\in S(\ell)$ holds

$$\left\|g_j(\gamma)\sigma-\frac{1}{M}\sum_{r=0}^{M-1}\frac{1}{cardS(r)}\sum_{\gamma'\in S(r)}\langle g_j,1_G\rangle_{(a,\gamma')}(\gamma)b_0\prod_{i=1}^{j-1}\langle g_j,g_i\rangle_{(a,\gamma')}(\gamma)b_i\right\|_{\infty}<\delta,$$

for any $a\ \in\Gamma$, and where δ is a constant depending on ε, to be specified later.

For $N\in\mathbf{N}$, let us denote

$$S^{\mathbf{g}}_{N'}[\varphi_1,\varphi_2,...,\varphi_j]=\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma'\in S(n)}\prod_{i=1}^{j}g_i(\gamma)\varphi_i \tag{3.4}$$

and for $N,N'\in\mathbf{N}$, let $S^{\mathbf{g}}_{N,N'}[\varphi_1,\varphi_2,...,\varphi_j]=S^{\mathbf{g}}_{N}[\varphi_1,\varphi_2,...,\varphi_j]-S^{\mathbf{g}}_{N'}[\varphi_1,\varphi_2,...,\varphi_j]$. Let $\phi:\mathbf{R}^{+}\to\mathbf{R}^{+}$, $\eta>0$, and let $C^{\eta,\phi}_{[2\eta^{-2}]},...,C^{\eta,\phi}_{1}$ be constants defined recursively as

$$C^{\eta,\phi}_{[2\eta^{-2}]}=1,\ \ C^{\eta,\phi}_{n-1}=\max\left\{C^{\eta,\phi}_{n},2\phi\left(C^{\eta,\varphi}_{n}\right)^{-1}\right\}.$$

For $\varepsilon>0$ let $C^{*}=C^{*}(\varepsilon):=C^{\eta,\phi}_{1}$ with $\eta=\dfrac{\varepsilon}{2^5 3}$ and $\phi(x)=\dfrac{\varepsilon^2}{2^3 3^3 x}$. In the definition of $L-$reducible of above will be taken $\delta=\varepsilon/16C^{*}(\varepsilon)$.

Proposition 3.1: Let $\mathbf{g}=(g_1,g_2,...,g_j)$ with $cplx(\mathbf{g})\le C$, for any $\varepsilon>0$ there exists a $K=K(\varepsilon,C)\in\mathbf{N}$ such that for any $M_*\in\mathbf{N}$ and for any function $F:\mathbf{N}\to\mathbf{N}$ there is a sequence

$$M_*\le M_1^{\varepsilon,C,F}\le...\le M_K^{\varepsilon,C,F}\le M^{*}=O_{M_*,\delta,c,\omega}(1),$$

in such a way, that if $\varphi_1,\varphi_2,...,\varphi_{j-1}\in L^{\infty}(X)$, $\|\varphi_i\|_{\infty}\le 1$, $i=1,...,j-1$ and $\varphi=\sum\limits_{t=0}^{k-1}\lambda_t\sigma_t$, where σ_t is $(L,\mathbf{g})-$reducible for any $L<F(M_*)$ and $\sum\limits_{t=0}^{k-1}|\lambda_t|<C^{*}$ then

$$\left\|S^{\mathbf{g}}_{N,N'}[\varphi_1,\varphi_2,...,\varphi_{j-1,}\varphi]\right\|_{L^2(X)}<\frac{\varepsilon}{4},$$

for $N\ge M_i$, $N'\le F(M_i)$, for some $i\in\{1.2,...,K\}$.

Proof. Let us consider a system $\mathbf{g}=(g_1,g_2,...,g_j)$ with $cplx(\mathbf{g})\le C$, and $\varepsilon>0$. Let M_t and $b_{0,t},b_{1,t},...,b_{j-1,t}\in L^{\infty}(X)$ be the integer and the functions corresponding to the $(L,\mathbf{g})-$reducibility of σ_t. Thus

$$S^{\mathbf{g}}_{N'}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\sum_{t=0}^{k-1}\lambda_t\sigma_t\right]=\frac{1}{N}\sum_{n=0}^{N-1}\sum_{t=0}^{k-1}\lambda_t\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}\prod_{i=1}^{j-1}g_i(\gamma)\varphi_i g_j(\gamma)\sigma_t.$$

If in this average $g_j(\gamma)\sigma_t$ is replaced by

$$\frac{1}{M_t}\sum_{r=0}^{M_t-1}\frac{1}{cardS(r)}\sum_{\gamma'\in S(r)}\langle g_j,1_G\rangle_{(a,\gamma')}(\gamma)b_0\prod_{i=1}^{j-1}\langle g_j,g_i\rangle_{(a_t,\gamma')}(\gamma)b_{i,t}$$

with $a_t\in\Gamma$, then is obtained an average with respect to a system of lower complexity, but with an error of $\delta=\varepsilon/(16C^*)$. Thus

$$\left\|S^{\mathbf{g}}_{N,N'}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\sum_{t=0}^{k-1}\lambda_t\sigma_t\right]\right\|_{L^2(X)}$$
$$\leq\sum_{t=0}^{k-1}\frac{1}{M_t}\sum_{r=0}^{M_t-1}\frac{|\lambda_t|}{cardS(r)}\sum_{\gamma'\in S(r)}\left\|S^{\mathbf{g}^*_{(a_t,\gamma')}}_{N,N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\ b_{0,t},b_{1,t},...,b_{j-1,t}\right]\right\|+\frac{\varepsilon}{8},$$

with $N,N'\leq L$. Recall that the coefficients λ_t were chosen such that $\sum\limits_{t=0}^{k-1}|\lambda_t|<C^*$. Using the complexity as induction parameter each term of the sum can be bounded for N, N' belonging to some integer interval of the form $\left[\overline{M},F(\overline{M})\right]$, but the bound is not uniform in the sense that the integer $\overline{M}$ may depend on the average. By a direct application of the Walsh techniques the norms can be uniformly bounded, with respect to the weights $\dfrac{|\lambda_t|}{cardS(r)}$, by a given δ.

Let F_1, F_2,..., $F_r:\mathbf{N}\to\mathbf{N}$, with $r=r(\varepsilon,C)$, be functions defined recursively by $F_1=F$ and $F_i=\max\limits_{1\leq M_i\leq N}\left\{F\left(M_i^{\delta,C-1,F}\right)\right\}$, where the M_i are given by the hypothesis of induction. Let i_1, i_2,..., $i_r\in\{1.2,...,K\}$ and define

$$\begin{aligned}
M^{(i_1)}&=M_{i_1}^{\delta,C-1,F_1}\\
M^{(i_1,i_2,)}&=\left(M_{i_1}^{\delta,C-1,F_2}\right)_{i_2}^{\delta,C-1,F_1}\\
&\vdots\\
M^{(i_1,i_2,...,i_s)}&=\left(\left(M_{i_1}^{\delta,C-1,F_2}\right)_{i_2,}^{\delta,C-1,F_1}...\right)_{i_s}^{\delta,C-1,F_s}.
\end{aligned}$$

Let $s=1\leq r$ and $(t,\gamma')\in\{0.1.,...,k-1\}\times S(M_t)$, by induction hypothesis there is some $i_1\in\{1.2,...,K\}$ such that for a given $\delta>0$ and for N, $N'\in\left[M^{(i_1)},F_1(M^{(i_1)})\right]$ holds

$$\left\|S^{\mathbf{g}^*_{(a_t,\gamma')}}_{N,N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\ b_{0,t},b_{1,t},...,b_{j-1,t}\right]\right\|_{L^2(X)}<\delta,$$

for at least $\dfrac{1}{K}$ of the pairs (t,γ'), with respect to the weights $\dfrac{|\lambda_t|}{cardS(r)}.\ \gamma'\in S(r)$. Continuing this process for the pairs (t,γ') with weights $\dfrac{|\lambda_t|}{cardS(r)}.\gamma'\in S(r)$, can be found an $i_2\in\{1.2,...,K\}$ such that for $\delta>0$ and for N, $N'\in\left[\ M^{(i_1,i_2)},F_1(\ M^{(i_1,i_2)})\right]$ holds

$$\left\|S^{\mathbf{g}^*_{(a_t,\gamma')}}_{N,N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\ b_{0,t},b_{1,t},...,b_{j-1,t}\right]\right\|_{L^2(X)}<\delta,$$

for at least a proportion of $\left(\dfrac{K-1}{K}\right)^{2} C^{*}$ of the pairs (t,γ'), with respect to the weights $\dfrac{|\lambda_t|}{cardS(M_t)}$. By repeatedly applying the process, one can find an $i_s \in \{1.2,...,K\}$, $s \leq r$, such that for any N, $N' \in \left[M^{(i_1,i_2,...,i_s)}, F_1(\ M^{(i_1,i_2,...,i_s)})\right]$ holds

$$\left\| S^{\mathbf{g}^*_{(a_t,\gamma')}}_{N,N'}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\ b_{0,t},b_{1,t},...,b_{j-1,t}\right]\right\|_{L^2(X)} < \delta,$$

for at least a proportion of $\left(\dfrac{K-1}{K}\right)^{r} C^{*}$ of the (t,γ'), with respect to the weights $\dfrac{|\lambda_t|}{cardS(r)}$. Thus we have

$$\sum_{t=0}^{k-1}\sum_{\gamma'\in S(M_t)}\frac{|\lambda_t|}{cardS(M_t)}\left\| S^{\mathbf{g}^*_{(a_t,\gamma')}}_{N,N'}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\ b_{0,t},b_{1,t},...,b_{j-1,t}\right]\right\|_{L^2(X)}$$

$$\leq \sum_{t=0}^{k-1}|\lambda_t|\,\delta + \left(\frac{K-1}{K}\right)^{r} C^{*} + \frac{\varepsilon}{8}.$$

Choosing r such that $\left(\dfrac{K-1}{K}\right)^{r} C^{*} < \dfrac{\varepsilon}{16}$, we obtain

$$\left\| S^{\mathbf{g}}_{N,N'}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\varphi\right]\right\|_{L^2(X)} < \frac{\varepsilon}{16}+\frac{\varepsilon}{16}+\frac{\varepsilon}{8}=\frac{\varepsilon}{4},$$

for any N, $N' \in \left[M^{(i_1,i_2,...,i_s)}, F_1(\ M^{(i_1,i_2,...,i_s)})\right]$ □

Lemma 3.2. Let $\varepsilon > 0$, and $\varphi_1,\varphi_2,...,\varphi_{j-1}$, with $\|\varphi_i\|_\infty \leq 1$. Let us assume that there is a map ψ and a constant $C \in [1, C^*]$ such that $\|\psi\|_\infty \leq 3C$. If

$$\left\| S^{\mathbf{g}}_{N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\psi\right]\right\|_{L^2(X)} > \varepsilon/6,$$

for some N, then there is a $(L,\mathbf{g})$ −reducible function σ, $L < N$, such that

$$C(\psi,\sigma) := \int \psi\sigma > 2\phi(C),$$

where ϕ is defined above Proposition 3.1.

Proof. Let $h_0 = \dfrac{S^{\mathbf{g}}_{N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\psi\right]\varphi_1}{\|\psi\|_\infty}$, and $h_i = \varphi_i$, $i \geq 1$.

We have $\dfrac{\left\| S_{N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\psi\right]\right\|^2_{L^2(X)}}{\|\psi\|_\infty} > 2\phi(C)$, and

$$C\left(\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}\prod_{i=1}^{j-1} g_i(\gamma)\varphi_i g_j(\gamma)\psi,\ \frac{\left\| S_{N}\left[\varphi_1,\varphi_2,...,\varphi_{j-1},\psi\right]\right\|^2_{L^2(X)}}{\|\psi\|_\infty}\right) =$$

$$C\left(\psi\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}g_j(\gamma)^{-1}h_0\prod_{i=1}^{j-1}g_i(\gamma)h_i\right).$$

So that if

$$\sigma:=\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}g_j(\gamma)^{-1}h_0\prod_{i=1}^{j-1}g_i(\gamma)h_i \tag{3.5}$$

then holds $C(\psi,\sigma)>2\phi(C)$. It remains to prove that σ is $(L,\mathbf{g})$ −reducible. Let $L<C_1N$, with $0<C_1<1$, if $\ell\le L$ and $\gamma'\in S(\ell)$ then

$$\left\|\sigma-\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}g_j\left(\gamma'\gamma\right)^{-1}h_0\prod_{i=1}^{j-1}g_i\left(\gamma'\gamma\right)h_i\right\|_\infty<\varepsilon/(16C^*),$$

since $\|S_N^{\mathbf{g}}[\varphi_1,\varphi_2,...,\varphi_{j-1},\psi]\|_{L^2(X)}\le 3C\le 3C^*$. If we change $\gamma\in S(1),...,\gamma\in S(N)$ by $\gamma\in S(1+\ell),...,\gamma\in S(N+l)$ then the magnitude of σ changes at most in $\dfrac{6\ell C^*}{N}<\varepsilon/(16C^*)$. Thus applying $g_j(\gamma)$ to σ we get

$$\left\|g_j(\gamma)\sigma-\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{cardS(n)}\sum_{\gamma\in S(n)}\langle g_j,1_G\rangle_{(a,\gamma)}(\gamma)h_0\prod_{i=1}^{j-1}\langle g_j,g_i\rangle_{(a,\gamma)}(\gamma)h_i\right\|_\infty<\varepsilon/(16C^*),$$

with $\gamma\in S(\ell)$, $\ell\le C_1N$, $C_1=\varepsilon/[96(C^*)^2]$. Therefore the definition of $(L,\mathbf{g})$ −reducibility is verified, with $M=N$, $b_i=h_i$. □

4 Gowers theory for decomposition of maps

One of the main ingredients in the Walsh techniques are the Gowers results[8] which, by an adequate adaptation of the Hahn-Banach theorem, give a decomposition of map in a ¨structured¨ component and a ¨random¨ component. More specifically, the problem posed in Gowers´s survey is when a real map ϕ can be written as $\phi=\phi_1+\phi_2$ with ϕ_1 with an structure enough strong such that the properties of ϕ_1 can be explicitly analyzed and ϕ_2 such that the properties of ϕ_1 are not affected by the perturbation by ϕ_2. In [8], as a corollary of the Hahn-Banach theorem, was obtained the following result: let $K_1,K_2,...,K_r$ open, convex subsets of a Hilbert space $(\mathcal{H},\langle\rangle)$, such that $0\in K_2$, for any i, let $c_1,c_2,...,c_r>0$ such that any $\varphi\in\mathcal{H}$ cannot be written as $\varphi=\sum_{i=1}^{r}c_i\varphi_i$, with $\varphi_i\in K_i$. Then there is a $\mathcal{L}\in\mathcal{H}$ such that $\langle\mathcal{L},\varphi\rangle\ge 1$ and $\langle\mathcal{L},\varphi_i\rangle\ge c_i^{-1}$, for any $i=1,2,...,r$.

Let $(\mathcal{H},\langle\cdot,\cdot\rangle)$ be a Hilbert space and let $\|\cdot\|$ be the norm induced by $\langle\cdot,\cdot\rangle$. If $\|\cdot\|_N$ is a norm in $\mathcal{H}$ equivalent to $\|\cdot\|$ then for $\varphi\in\mathcal{H}$ set

$$\|\varphi\|_N^*=\sup\left\{\langle\varphi,\psi\rangle:\|\psi\|_N\le 1\right\}. \tag{4.1}$$

The norm $\|\cdot\|_N^*$ is called the dual norm of $\|\cdot\|_N$ and is equivalent to $\|\cdot\|_N$. It can be considered a family of norms $\{\|\cdot\|_N\}$ indexed by $\mathbf{N}$. Walsh considered families $\{\|\cdot\|_N\}_{N\in\mathbf{N}}$ decreasing with N and such the family of dual norms also decreases with N.

Recall the definition of the constants $C^{\eta,\phi}_{[2\eta^{-2}]},...,C^{\eta,\phi}_1$ by $C^{\eta,\phi}_{[2\eta^{-2}]}=1,\ \ C^{\eta,\phi}_{n-1}=\max\left\{C^{\eta,\phi}_n,2\phi\left(C^{\eta,\varphi}_n\right)^{-1}\right\}$,with $\phi:\mathbf{R}^+\to\mathbf{R}^+$.

Proposition 4.1.[16] Let $\{\|\cdot\|_N\}_{N\in\mathbf{N}}$ be a family of norms in $\mathcal{H}$ with the above property, let $0<\delta,c<1$ and $\phi:\mathbf{R}^+\to\mathbf{R}^+$ be a decreasing function. Let $\omega:\mathbf{N}\to\mathbf{N}$ be a map such that $\omega(N)>N$ for any N. Then for any $M_*>0$ there is a sequence

$$M_*\leq M_1\leq...\leq M_{[2\delta^{-2}]}\leq M^*=O_{M_*,\delta,c,\omega}(1),$$

independent of the norms, such that for any $\varphi\in\mathcal{H}$ can be found some $i\in\left\{1,...,\left[2\delta^{-2}\right]\right\}$ and some $A,B\in\mathbf{N}$ with $M_*\leq A\leq M_i\leq\omega(M_i)\leq B$, in such a way that φ can be decomposed as $\varphi=\varphi_1+\varphi_2+\varphi_3$ with

$$\|\varphi_1\|_B<C^{\eta,\phi}_i,\ \ \|\varphi_2\|_A<\phi\left(C^{\eta,\varphi}_n\right),\quad \|\varphi_3\|<\delta.$$

Other result adapted by Walsh from Gowers theory is

Lemma 4.2. ([16],[8]) Let $\mathcal{H}_0$ be a bounded subset of a Hilbert space $\mathcal{H}$ and set

$$\|\varphi\|_{\mathcal{H}_0}=\inf\left\{\sum_{j=0}^{k-1}|\lambda_i|:\varphi=\sum_{j=0}^{k-1}c_j\sigma_j,\ \sigma_j\neq 0\right\}$$

If it is assumed that the norm $\|\|_{\mathcal{H}_0}$ is well defined and equivalent to $\|\|$, then its dual norm is given by

$$\|\varphi\|^*_{\mathcal{H}_0}=\sup\left\{\langle\varphi,\sigma\rangle:\sigma\in\mathcal{H}_0\right\}.$$

5 Norm convergence of spherical averages

Let us consider the following family of spaces $\Sigma_{L,\mathbf{g}}=\{\sigma:\sigma\text{ is }(L,\mathbf{g})-\text{reducible}\}$, and the norms $\|\cdot\|_L:=\|\cdot\|_{\Sigma_L}$ and $\|\cdot\|_L^*:=\|\cdot\|^*_{\Sigma_L}$. If $L_1>L_2$, then $\Sigma_{L_1}\subset\Sigma_{L_2}$. We have:

Theorem 5.1. Let $\varepsilon>0$ and $\mathbf{g}=(g_1,g_2,...,g_j)$ be a system with $cplx(\mathbf{g})\leq C$, then there exists a natural $K=K_{\varepsilon.C}$ such that for any function $F:\mathbf{N}\to\mathbf{N}$ and for any $M\in\mathbf{N}$, there are natural numbers $M_1^{\varepsilon,C,F},...,M_K^{\varepsilon,C,F}\geq M$ such that for any maps $\varphi_1,\varphi_2,...,\varphi_j\in L^\infty(X,\mu)$, with $\|\varphi_i\|_\infty\leq 1$, $i=1,,,.j$ there is a number $\bar{i}\in\{1,2,...,K_{\varepsilon.C}\}$ with the property that for any $N,\ N'\in[M_{\bar{i}},F(M_{\bar{i}})]$ holds

$$\left\|S^{\mathbf{g}}_{N',N'}\left[\varphi_1,\varphi_2,...,\varphi_j\right]\right\|_2<\varepsilon. \tag{5.1}$$

Proof. For any $M\ \in\mathbf{N}$ and for any function $F:\mathbf{N}\to\mathbf{N}$, by the Proposition 3.1 there is a sequence $M\leq M_1^{\varepsilon,C,F}\leq...\leq M_K^{\varepsilon,C,F}\leq M^*$, thus set $\omega(M)=F(M^*)$.

By the decomposition theorem φ_j can be written as

$$\varphi_j = \sum_{t=0}^{k-1} \lambda_t \sigma_t + \psi + \zeta,$$

with $\sum_{t=0}^{k-1} |\lambda_t| < C_i$, with $C_i \in [1, C^*]$, given by the Lemma 3.2 and $\sigma_t \in \Sigma_{B,\mathbf{g}}$ with $B \geq \omega(M) = F(M^*)$. Moreover, $\|\psi\|_{M_i}^* < \phi(C_i)$ and $\|\zeta\|_2 < \delta$, with $\delta = \dfrac{\varepsilon}{2^5 3}$.

To apply Lemma 3.2, a control on ψ is needed. Let $S = \{x \in X : |\zeta(x)| < C_i\}$, The issue is to see that the points outside S may be neglected. Since

$$\varphi_j = \sum_{t=0}^{k-1} \lambda_t \sigma_t + \psi + \zeta,$$

we have

$$\begin{aligned} |\psi| I_{S^C} &\leq \sum_{t=0}^{k-1} |\lambda_t| \, |\sigma_t| \, I_{S^C} \, |+ |\zeta|| \, I_{S^C} + |\varphi_j| \, I_{S^C} \\ &\leq \sum_{t=0}^{k-1} |\lambda_t| \, |\sigma_t| \, I_{S^C} + |\zeta| \, I_{S^C} \leq 3 \, |\zeta| \, I_{S^C}, \end{aligned}$$

where I_A is the characteristic function of A. Since $\sigma_t \in \Sigma_{\omega(M),\mathbf{g}}$ holds $\|\psi I_{S^C}\| \leq 3 \|\zeta\|_2$. By the Chebyshev inequality

$$\mu(S^c) \leq \frac{\|\zeta\|_2^2}{C_i^2} < \frac{\delta^2}{C_i^2}. \tag{5.2}$$

Let $\sigma \in \Sigma_{M,\mathbf{g}}$,replacing if necessary ψ by ψI_{S^C} and ζ by $\zeta + \psi I_{S^C}$, it can be assumed that $\|\zeta\|_\infty < 3C_i$. In this case we would have $\|\psi\|_{M_i}^* < 2\phi(C_i)$ and $\|\zeta\|_2 < 4\delta$. We have $\left|\left\langle \psi I_{S^C}, \sigma \right\rangle\right| < \phi(C_i) + 3 \|\zeta\|_2 \, \mu(S^c)^{1/2} \leq \phi(C_i) + 3\delta^2 + \delta/C_i < 2\phi(C_i)$. Thus, by the counter reciprocal of the lemma 3.2 for N, $N' \geq M$

$$\left\| S^{\mathbf{g}}{}_{N',N'} [\varphi_1, \varphi_2, ..., \zeta] \right\|_2 \leq \varepsilon/3.$$

By hypothesis $\|\varphi_j\|_\infty \leq 1$, and therefore

$$\|S^{\mathbf{g}}{}_{N'} [\varphi_1, \varphi_2, .., \psi I_{S^C} + \zeta]\|_2 \leq 3\delta + \|\zeta\|_2 < 8\delta, \text{ for any } N.$$

Thus $\left\| S^{\mathbf{g}}{}_{N',N'} [\varphi_1, \varphi_2, ..., \zeta] \right\|_2 < 16\delta = \varepsilon/6$. Moreover,

$$\left\| S^{\mathbf{g}}{}_{N',N'} \left[\varphi_1, \varphi_2, ..., \sum_{t=0}^{k-1} \lambda_t \sigma_t \right] \right\|_2 < \frac{\varepsilon}{3},$$

for any N, $N' \in [M_{\bar{i}}, F(M_{\bar{i}})]$, for some $M_{\bar{i}} \in [M, \omega(M)]$. Finally

$$\left\| S^{\mathbf{g}}{}_{N',N'} \left[\varphi_1, \varphi_2, ..., \sum_{t=0}^{k-1} \lambda_t \sigma_t \psi + \zeta \right] \right\|_2 < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

for any N, $N' \in [M_{\bar{i}}, F(M_{\bar{i}})]$, for some $M_{\bar{i}} \in [M, \omega(M)]$. $\square$

The result of Theorem 1.2 is obtained as follows: if the averages $S^{\mathbf{g}}{}_{N'}[\varphi_1, \varphi_2, ..., \varphi_j]$ were not convergent in norm, there would be and $\varepsilon > 0$ and a function $F : \mathbf{N} \to \mathbf{N}$ such that for any N

$$\left\| S^{\mathbf{g}}{}_{N',F(N)}[\varphi_1, \varphi_2, ..., \varphi_j] \right\|_2 > \varepsilon.$$

Theorem 1.1 will be proved as an immediate consequence of the following:

Proposition 5.2. Let $g_1, g_2, ..., g_j$ be antihomorphisms from Γ to G, and such that $g_i(\gamma_2)\, g_k(\gamma_1) = g_k(\gamma_1)\ g_i(\gamma_2)$, for any $\gamma_1, \gamma_2 \in \Gamma$. Then the system

$$\mathbf{G} = (g_1, g_1 g_2, ..., g_1 g_{2...} g_j)$$

has finite complexity.

Proof. Let $D_a(g)(\gamma) := g(\gamma)^{-1} g(\gamma a)$ and $T_a(g)(\gamma) := g(\gamma a)$, i.e., the derivative and the translation with one of the element the identity of Γ. If the map g is an antihomorphism, i.e., $g(\gamma_1\gamma_2) = g(\gamma_2)\, g(\gamma_1)$, then

$-D_a\left(g^{-1}\right)(\gamma) = g(\gamma) g(\gamma a)^{-1} = g(\gamma)\left(g(a)\, g(\gamma)\right)^{-1} = g(a)^{-1}$
$-T_a(g)(\gamma) := g(\gamma a) = g(a)\, g(\gamma).$

Thus, following [17], let $i < j$ and a fixed. Then

$$\begin{aligned}
&\langle\, g_1 g_2, ... g_j,\ g_1 g_2, ... g_i \rangle_a(\gamma) = D_a\left(\left(\, g_1 g_2, ... g_j\right)^{-1}\right)(\gamma)\, T_a\left(g_1 g_2, ... g_i\right)(\gamma) = \\
&= \left(\, g_1 g_2, ... g_j\right)^{-1}(a)\left(g_1 g_2, ... g_i\right)(a)\left(g_1 g_2, ... g_i\right)(\gamma) = \\
&= \left(\, g_1 g_2, ... g_j\right)^{-1}(a)\, g_1(a)\, g_1(\gamma) g_2(a)\, g_2(\gamma) ... g_i(a)\, g_i(\gamma) = \\
&= \left(g_1 g_2, ... g_i\right)(\gamma) g_{i+1}(\gamma)^{-1} ... g_j(\gamma)^{-1} g_{i+1}(a)^{-1} ... g_j(a)^{-1}.
\end{aligned}$$

Recall that if c is the constant map given $c \in G$, then $cplx(\mathbf{g},c) = cplx(\mathbf{g})$ holds. Hence

$$cplx(\mathbf{G}^*) = cplx(\mathbf{G}),$$

and the result follows by continuing the induction process on j. $\square$

Thus, recall that if h_i, $i = 1, 2, ..., j$ are measure-preserving actions of Γ on X, we can define maps g_i from Γ a nilpotent group of unitary operators on $L^2(X)$ by $g_i(\gamma)[\varphi](x) = \varphi(h_i(\gamma)(x))$, where each g_i result antihomomorphisms. Therefore, Theorem 1.1 is proved from Proposition 5.1.

Acknowledgements. The support of this work by Consejo Nacional de Investigaciones Científicas y Técnicas, Universidad Nacional de La Plata and Agencia Nacional de Promoción Científica y Tecnológica of Argentina is greatly appreciated. The second author is a member of *CONICET*.

References

[1] V. Bergelson, *Weakly mixing PET*, Ergod. Th. Dynam. Sys. **7** (1987), 337-349.

[2] A. Bufetov, *Convergence of spherical averages for actions of free groups*, Ann. of Math, **155** (2) (2002), 929-944.

[3] Bufetov, M. Khristoforov and A. Klimenko, *Césaro convergence of spherical averages for measure-preserving actions of Markov semigroups and groups*, to appear in IMRN.

[4] A. Bufetov and C. Series, *A pointwise ergodic theorem for Fuchsian groups*, Math Proc of the Cambridge Phil. Soc. **151**, 145-159 (2011),

[5] K. Fujiwara and A. Nevo, *Maximal and pointwise ergodic theorems for word-hyperbolic groups*, Erg. Th. & Dynam. Sys. **18** (1998), 843-858.

[6] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szmerédi on arithmetic progressions*, J. d'Analyse Math. **31** (1977), 204-256.

[7] E. Ghys and P. de la Harpe, *Sur le groupes hyperbolique d apres M. Gromov*, Progress in Math. **83**, Birkhauser, Zurich, 1990.

[8] W. T. Gowers, *Decomposition, approximate structure, transfernce and the Hahn-Banach theorem*, Bull London Math. Soc. **42** (2010), 573–606.

[9] B. Host and B. Kra, *Non conventional ergodical averages and nilmanifolds*, Ann. of Math, **161** (1) (2005), 397-488.

[10] A. Leibman, *Convergence of multiple ergodic averages along polynomials of several variables*, Israel Journal of Mathematics 146 (2005), 303-315

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

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

[13] A. Nevo, *Pointwise mergodic theorems for actions of groups*, in "Handsndbook of Dynamical Systems", Vol. 1B, 871-982, Elsevier B.V., Amsterdam 2006.

[14] C. Series, *Geometrical Markov coding of geodesics on surfaces with of constant negative curvature*, Erg. Th. Dynam. Sys. **6** (1986), 601-625.

[15] M. Pollicott and R. Sharp, *Ergodic theorems for actions of hyperbolic groups*, to appear in Proc. of the Amer. Math. Soc.

[16] M. Walsh, *Norm convergence of nilpotent ergodic averages*, Ann. of Math. **175** (2012), 1667–1688.

[17] P. Zorin-Kranich, *Norm convergence of nilpotent ergodic averages along Følner nets*, arXiv: 1111.7292.

Authors' address:

Alejandro M. Mesón and Fernando Vericat
IFLYSIB CONICET–UNLP and ***GAMEFI–UNLP***,
Calle 59, 789 (1900) La Plata, Argentina.
Email: meson@iflysib.unlp.edu.ar & vericat@iflysib.unlp.edu.ar