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 generally 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 valued 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 valued 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 \(\Gamma\) on a Lebesgue space \((X, \mu)\) and maps \(\varphi : X \to \mathbb{R}\), to analyze averages of the form

\[
\frac{1}{m_\Gamma(E)} \int_E \varphi(\gamma x) \, dm_\Gamma(\gamma),
\]

where \(E \subset \Gamma\) and \(m_\Gamma\) is the Haar measure on \(\Gamma\). In that article was proved the convergence of these averages under the condition of the existence of adequate sequences \((F_\alpha)\) in \(\Gamma\), 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 \(g \geq 2\).
Let $\Gamma$ be a finitely generated Markov groups $\Gamma$ 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 $\gamma$. If $\gamma \in \Gamma$, then by $T_{\gamma}$ 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

$$S_N;\varphi(x) := \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card}S(n)} \sum_{\gamma \in S(n)} \varphi(T_{\gamma}(x))$$

(When $\text{card}S(n) = 0$ then we set $S_N;\varphi(x) = 0$). Recently Bufetov, Khristoforov and Klimenko[3] analyzed these spherical averages and proved the pointwise convergence of Cesáro averages for the spherical sequence $\{S_N;\varphi(x)\}$. 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

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

where $A \in \mathcal{B}$ and $j \in \mathbb{N}$. 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

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

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
(1.3) \[
\frac{1}{N} \sum_{n=0}^{N-1} \prod_{i=1}^{j} \varphi_i \left( T^{p_i(n)} x \right),
\]
where \(p_i\) are polynomials valued in \(\mathbb{Z}\). The average (1.3) corresponds to the case of linear polynomials. Bergelson [1] proved the existence of \(\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \prod_{i=1}^{j} \varphi_i \left( T^{p_i(n)} x \right)\) under the assumption that \(T\) is 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 \((X, \mu)\) and study the norm convergence of averages

(1.4) \[
\frac{1}{N} \sum_{n=0}^{N-1} \prod_{i=1}^{j} T_{i}^{p_i(n)} \varphi_i,
\]
where \(T_1, \ldots, T_l \in G\) and \(p_{1}, \ldots, p_{l} \in \mathbb{Z}[n]\). The average (1.3) corresponds to the case of \(T_1 = \ldots = T_l\), 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 noncommutative 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 \((X, \mu)\) be a measure space and let \(\Gamma\) be a finitely generated Markov group. Let us consider actions \(h_1, h_2, \ldots, h_j\) of \(\Gamma\) on \(X\), i.e. for any \(\gamma \in \Gamma\) are defined measure-preserving maps \(h_i(\gamma) : X \to X\), \(i = 1, 2, \ldots, j\). If the actions commute in the sense that \(h_i(\gamma_1) h_k(\gamma_2) = h_k(\gamma_1) h_i(\gamma_2)\), for any \(\gamma_1, \gamma_2 \in \Gamma\) and \(i \neq k\) then for any \(\varphi_1, \varphi_2, \ldots, \varphi_j \in L^\infty(X)\), the multiergodic sequence

(1.5) \[
\frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card}S(n)} \sum_{\gamma \in S(n)} \varphi_1(h_1(\gamma)x) \varphi_2(h_1(\gamma)h_2(\gamma)x) \ldots \varphi_j(h_1(\gamma)\ldots h_j(\gamma)x),
\]
converges in \(L^2\)-norm.

In order to apply the Walsh techniques notice the following, if \(h_i, i = 1, 2, \ldots, j\) is a measure-preserving action of \(\Gamma\) 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 \(g_i, i = 1, 2, \ldots, j\), 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 \((X, \mu)\) be a measure space and let \(\Gamma\) be a finitely generated Markov group. Let \(g_1, g_2, ..., g_j : \Gamma \to G\), with \(G\) a nilpotent group of unitary operators on \(L^2(X)\). If the system \(g = (g_1, g_2, ..., g_j)\) has finite complexity then the sequence

\[
\frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card}S(n)} \sum_{\gamma \in S(n)} g_1(\gamma) \varphi_1 g_2(\gamma) \varphi_2 ... g_j(\gamma) \varphi_j
\]

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 \(g_1, g_2, ..., g_j : \mathbb{Z} \to G\), with

\[
g_j(n) = T_1^{p_{1,1}(n)} ... T_\ell^{p_{1,\ell}(n)},
\]

where \(T_1, ..., T_\ell \in G\) and \(p_{i,1}, ..., p_{i,\ell} \in \mathbb{Z}[n]\). 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 \(\Gamma = \mathbb{Z}\), then \(g(n) = T_1^{p_{1,1}(n)} ... T_\ell^{p_{1,\ell}(n)}\) with \(T_1, ..., T_\ell\) measure-preserving transformations and each \(p_i \in \mathbb{Z}[n]\). Polynomial systems have finite complexity, this was proved for \(\Gamma = \mathbb{Z}\) in [16], and for non commutative \(\Gamma\) in [17].

2 Markov groups

Let \(\Gamma\) 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 \(\Gamma\) is a Markov group if there is a labelling map

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

and a distinguished vertex \(v_0\), such that

i) \(\lambda\) can be lifted to a map \(\lambda : P[\mathcal{G}, v_0] \to \Gamma\) 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 \(\gamma\) 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 \(\Gamma\), with a finite symmetric set of generators \(S\), is defined by \(d(\gamma_1, \gamma_2) = |\gamma_1^{-1} \gamma_2|\), where \(|\gamma| = |\gamma|_S\) is the minimal number of elements of \(S\) needed to represent \(\gamma\). A group \(\Gamma\) 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 \(\delta\) 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 \(\rho = \lim_{n \to \infty} \frac{1}{n} \log \text{card}S(n)\)

then there are constants \( C_2 > C_1 > 0 \)

\[
C_1 \rho^{-n} \leq \text{card} S(n) \leq C_2 \rho^{-n}.
\]

Thus for ergodic averages for word hyperbolic actions groups can be considered weights \( \rho^{-n} \) instead of \( \frac{1}{\text{card} S(n)} \).

### 3 Complexity of systems

Let \( g = (g_1, g_2, \ldots, g_j) \) be an ordered tuple of maps from \( \Gamma \) 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 \( \gamma \) is

\[
D_{(a,b)}(g)(\gamma) := g((\gamma)^{-1} g(a\gamma b)),
\]

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

\[
\langle g, h \rangle_{(a,b)}(\gamma) := D_{(a,b)}(g^{-1})(\gamma)T_{(a,b)}(h)(\gamma),
\]

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

\[
g^*_{(a,b)} = (g_1, g_2, \ldots, g_j, \langle g_j, g_1 \rangle_{(a,b)}, \ldots, \langle g_j, g_{j-1} \rangle_{(a,b)}).
\]

The concept of complexity is given by induction in the following way. A system constituted only by the identity \( 1_G \) map has complexity \( \leq 0 \), it is denoted by \( \text{cplx}(\{1_G\}) \leq C \). For any \( a, b \in \Gamma \) then \( \text{cplx}(g) \leq C \). The property known as cheating says that if and are added constants, repeated maps or repeated maps multiplied by a constant to a system it does not change its complexity. For example \( \text{cplx}(g_1, c_1 g_1, c_1 g_1, g_2, c_2 g_2) = \text{cplx}(g_1, g_2) \).

A filtration on a group \( G \) is a sequence of subgroups

\[
G_\bullet = G_0 \geq G_1 \geq \ldots \geq G_{\bullet},
\]

such that \( G_0 = G \) and \([G_1, G_i] \subseteq G_{i+1} \). 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 \( (G_{\bullet+1})_t = G_{t+1} \). 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 \( a_1, b_1, \ldots, d_1, b_{d+1} \in \Gamma \)

\[
D_{(a_1, b_1)} \cdots D_{(a_{d+1}, b_{d+1})}(g) = 1_G.
\]

If \( G \) is not nilpotent then the set of polynomials of degree \( \leq d \), may be not a group.
The system \( g = (g_1, g_2, ..., g_j) \), 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 \( \|\sigma\|_\infty \leq 1 \), \( L \in \mathbb{N} \), \( 0 < \varepsilon < 1 \). The function \( \sigma \) is reducible with respect to \( L \) and to a system \( g = (g_1, g_2, ..., g_j) \), or just \( (L, g) \) -reducible if there is a number \( M > 0 \) and functions \( b_0, b_1, ..., b_{j-1} \in L^\infty(X) \), \( \|b_i\|_\infty \leq 1 \), \( i = 0, 1, ..., j - 1 \), such that for any \( \ell \leq L \) and for any \( \gamma \in S(\ell) \) holds

\[
\left\| g_j(\gamma) \sigma - \frac{1}{M} \sum_{r=0}^{M-1} \chi_{\text{card} S(r)} \sum_{\gamma' \in S(r)} (g_j)_\gamma \left( \langle a, \gamma' \rangle(\gamma) b_0 \prod_{i=1}^{j-1} (g_j)_\gamma \left( \langle a, \gamma' \rangle(\gamma) b_i \right) \right) \right\|_\infty < \delta,
\]

for any \( a \in \Gamma \), and where \( \delta \) is a constant depending on \( \varepsilon \), to be specified later.

For \( N \in \mathbb{N} \), let us denote

\[
S^*_N, [\varphi_1, \varphi_2, ..., \varphi_j] = \frac{1}{N} \sum_{n=0}^{N-1} \chi_{\text{card} S(n)} \sum_{\gamma' \in S(n)} \prod_{i=1}^{j} g_i(\gamma) \varphi_i
\]

and for \( N, N' \in \mathbb{N} \), let \( S^*_N, [\varphi_1, \varphi_2, ..., \varphi_j] = S^*_N, [\varphi_1, \varphi_2, ..., \varphi_j] - S^*_N, [\varphi_1, \varphi_2, ..., \varphi_j] \).

Let \( \phi : \mathbb{R}^+ \to \mathbb{R}^+ \), \( \eta > 0 \), and let \( C_{[2n-1]}^n, ..., C_1^n \) be constants defined recursively as

\[
C_{[2n-1]}^n = 1, \quad C_{n-1}^n = \max \left\{ C_n^n, 2\phi(C_n^n)^{-1} \right\}.
\]

For \( 0 < \varepsilon \) let \( C^* = C^*(\varepsilon) := C_{[4]} \) with \( \eta = \frac{\varepsilon}{2\sqrt{3}} \) and \( \phi(x) = \frac{x^2}{2(3x)^r} \). In the definition of \( L \)-reducible of above will be taken \( \delta = \frac{\varepsilon}{16C^*(\varepsilon)} \).

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

\[
M_1 \leq M_1^{\varepsilon, C, F} \leq \ldots \leq M_K^{\varepsilon, C, F} \leq M^* = O_{M, \varepsilon}(1),
\]

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

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

for \( N \geq M_1 \), \( N' \leq F(M_t) \), for some \( i \in \{1, 2, ..., K\} \).

**Proof.** Let us consider a system \( g = (g_1, g_2, ..., g_j) \) with \( \text{cplx}(g) \leq C \), and \( \varepsilon > 0 \). Let \( M_t \) and \( b_0, b_1, ..., b_{j-1}, b_{j-1} \in L^\infty(X) \) be the integer and the functions corresponding to the \((L, g)\) -reducibility of \( \sigma_t \). Thus

\[
S^*_N, [\varphi_1, \varphi_2, ..., \varphi_{j-1}, \sum_{t=0}^{k-1} \lambda_t \sigma_t] = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{t=0}^{k-1} \lambda_t \chi_{\text{card} S(n)} \sum_{\gamma' \in S(n)} \prod_{i=1}^{j-1} (g_j \varphi_i) \sigma_i.
\]
If in this average \( g_j(\gamma)\sigma_i \) is replaced by

\[
\frac{1}{M_t} \sum_{r=0}^{M_t-1} \frac{1}{\text{card}S(r)} \sum_{\gamma' \in S(r)} \langle g_j, 1_G \rangle_{(a, \gamma')} (\gamma) b_0 \prod_{i=1}^{i=r-1} \langle g_j, g_i \rangle_{(a, \gamma')} (\gamma) b_{i,t}
\]

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

\[
\left\| S_{N,N'}^\delta \left[ \varphi_1, \varphi_2, \ldots, \varphi_{j-1}, \sum_{t=0}^{k-1} \lambda_t \sigma_t \right] \right\|_{L^2(X)} \leq \sum_{t=0}^{k-1} \frac{M_t}{\text{card}S(r)} \sum_{\gamma' \in S(r)} \left\| S_{N,N'}^\delta \left[ \varphi_1, \varphi_2, \ldots, \varphi_{j-1}, b_{0,t}, b_{1,t}, \ldots, b_{j-1,t} \right] \right\|_{L^2(X)} + \frac{\varepsilon}{8},
\]

with \( N, N' \leq L \). Recall that the coefficients \( \lambda_t \) were chosen such that \( \sum_{t=0}^{k-1} |\lambda_t| < C^* \).

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 \([M, M/F(M)]\), but the bound is not uniform in the sense that the integer \( M \) may depend on the average. By a direct application of the Walsh techniques the norms can be uniformly bounded, with respect to the weights \( \frac{|\lambda_t|}{\text{card}S(r)} \), by a given \( \delta \).

Let \( F_1, F_2, \ldots, F_r : \mathbb{N} \to \mathbb{N} \), with \( r = r(\varepsilon, C) \), be functions defined recursively by \( F_1 = F \) and \( F_i = \max_{i \leq M_i \leq N} \left\{ \text{F} \left( M_i^{\delta, C-1, F} \right) \right\} \), where the \( M_i \) are given by the hypothesis of induction. Let \( i_1, i_2, \ldots, i_r \in \{1, 2, \ldots, K\} \) and define

\[
M^{(i_1)} = M_i^{\delta, C-1, F_1},
M^{(i_1, i_2)} = \left( M^{(i_1, i_2, C-1, F_2)} \right)^{i_2},
\vdots
M^{(i_1, i_2, \ldots, i_s)} = \left( \left( M^{(i_1, i_2, C-1, F_2)} \right)^{i_2}, \ldots \right)^{i_s}.
\]

Let \( s = 1 \leq r \) and \((t, \gamma') \in \{0,1, \ldots, k-1\} \times S(M_t)\), by induction hypothesis there is some \( i_1 \in \{1, 2, \ldots, K\} \) such that for a given \( \delta > 0 \) and for \( N, N' \in [M^{(i_1)}], F_t(M^{(i_1)}) \) holds

\[
\left\| S_{N,N'}^\delta \left[ \varphi_1, \varphi_2, \ldots, C_{j-1} \right], b_{0,t}, b_{1,t}, \ldots, b_{j-1,t} \right\|_{L^2(X)} < \delta,
\]

for at least \( \frac{1}{K} \) of the pairs \((t, \gamma')\), with respect to the weights \( \frac{|\lambda_t|}{\text{card}S(r)} \), \( \gamma' \in S(r) \).

Continuing this process for the pairs \((t, \gamma')\) with weights \( \frac{|\lambda_t|}{\text{card}S(r)} \), \( \gamma' \in S(r) \), can be found an \( i_2 \in \{1, 2, \ldots, K\} \) such that for a given \( \delta > 0 \) and for \( N, N' \in [M^{(i_1, i_2)}], F_t(M^{(i_1, i_2)}) \) holds

\[
\left\| S_{N,N'}^\delta \left[ \varphi_1, \varphi_2, \ldots, C_{j-1} \right], b_{0,t}, b_{1,t}, \ldots, b_{j-1,t} \right\|_{L^2(X)} < \delta,
\]
for at least a proportion of \( \left( \frac{K-1}{K} \right)^r \) \( C^* \) of the pairs \((t, \gamma')\), with respect to the weights \( \frac{\lambda_i}{\text{card} S(M_i)} \). By repeatedly applying the process, one can find an \( i_s \in \{1, 2, \ldots, K\} \), \( s \leq r \), such that for any \( N, N' \in [M^{(i_1, i_2, \ldots, i_s)}, F_1(M^{(i_1, i_2, \ldots, i_s)})] \) holds

\[
\left\| S_{N,N'}^{\ell, \epsilon} [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, b_{0,t}, b_{1,t}, \ldots, b_{j-1,t}] \right\|_{L^2(X)} < \delta,
\]

for at least a proportion of \( \left( \frac{K-1}{K} \right)^r \) \( C^* \) of the \((t, \gamma')\), with respect to the weights \( \frac{\lambda_i}{\text{card} S(r)} \). Thus we have

\[
\sum_{t=0}^{k-1} \sum_{\gamma' \in S(M_i)} \frac{\lambda_i}{\text{card} S(M_i)} \left\| S_{N,N'}^{\ell, \epsilon} [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, b_{0,t}, b_{1,t}, \ldots, b_{j-1,t}] \right\|_{L^2(X)} \leq \sum_{t=0}^{k-1} \lambda_i \delta + \left( \frac{K-1}{K} \right)^r C^* + \frac{\varepsilon}{8}.
\]

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

\[
\left\| S_{N,N'}^{\ell, \epsilon} [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, \varphi] \right\|_{L^2(X)} < \frac{\varepsilon}{16} + \frac{\varepsilon}{8} = \frac{\varepsilon}{4},
\]

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

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

\[
\| S_N [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, \varphi] \|_{L^2(X)} > \varepsilon / 6,
\]

for some \( N \), then there is a \((L, g)\)–reducible function \( \sigma \), \( L < N \), such that

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

where \( \phi \) is defined above Proposition 3.1.

**Proof.** Let \( h_0 = S_N [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, \psi] / \|\psi\|_\infty \), and \( h_i = \varphi_i, \ i \geq 1 \).

We have \( \| S_N [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, \psi] \|_{L^2(X)}^2 > 2\phi(C) \), and

\[
C \left( \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card} S(n)} \sum_{\gamma \in S(n)} \prod_{i=1}^{j-1} g_i(\gamma) \varphi_i g_j(\gamma) \psi, \frac{\| S_N [\varphi_1, \varphi_2, \ldots, \varphi_{j-1}, \psi] \|_{L^2(X)}^2}{\|\psi\|_\infty} \right) =
\]
is verified, with $M$ by $L$ and written as $φ$. Gowers theory for decomposition of maps

\[
\| (4.1) \]

The problem posed in Gowers’s survey is when a real map $ϕ$ map in a structured component and a random component. More specifically, by an adequate adaptation of the Hahn-Banach theorem, give a decomposition of

One of the main ingredients in the Walsh techniques are the Gowers results[8] which, with since $C(3.5)$ holds then holds $C(3.5)$. So that if

\[
\sigma := \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card} S(n)} g_j(γ) h_0 \prod_{i=1}^{j-1} g_i(γ) h_i
\]

then holds $C(3.5)$. It remains to prove that $σ$ is $(L, g)$ - reducible. Let $L < C_1 N$, with $0 < C_1 < 1$, if $ℓ ≤ L$ and $γ' ∈ S(ℓ)$ then

\[
\left\| \sigma - \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card} S(n)} g_j(γ') h_0 \prod_{i=1}^{j-1} g_i(γ') h_i \right\| < ε/(16C^*),
\]

since $\| S^0_1 [ϕ_1, ϕ_2, ..., ϕ_{j-1}, ϕ_j] \|_{L^2(X)} ≤ 3C ≤ 3C^*$. If we change $γ ∈ S(1), ..., γ ∈ S(N)$ by $γ ∈ S(1 + ℓ), ..., γ ∈ S(N + ℓ)$ then the magnitude of $σ$ changes at most in $2ℓC^*/N < ε/(16C^*)$. Thus applying $g_j(γ)$ to $σ$ we get

\[
\left\| g_j(γ)σ - \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\text{card} S(n)} \prod_{\gamma ∈ S(n)} g_j(γ) h_0 \prod_{i=1}^{j-1} g_i(γ) h_i \right\| < ε/(16C^*),
\]

with $γ ∈ S(ℓ), ℓ ≤ C_1 N, C_1 = ε/[96(C^*)^2]$. Therefore the definition of $(L, g)$ – reducibility is verified, with $M = N, b_i = h_i$.

\[\Box\]

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 $ϕ = ϕ_1 + ϕ_2$ 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 $ϕ_2$ 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 $(H, ⟨⟩)$, such that $0 ∈ K_2$, for any $i$, let $c_1, c_2, ..., c_r > 0$ such that any $H ∈ H$ cannot be written as $ϕ = \sum_i c_i ϕ_i$, with $ϕ_i ∈ K_i$. Then there is a $L ∈ H$ such that $⟨L, ϕ⟩ ≥ 1$ and $⟨L, ϕ_i⟩ ≥ c_i^{-1}$, for any $i = 1, 2, ..., r$.

Let $(H, ⟨·, ·⟩)$ be a Hilbert space and let $\|\cdot\|$ be the norm induced by $⟨·, ·⟩$. If $\|\cdot\|_N$ is a norm in $H$ equivalent to $\|\cdot\|$ then for $ϕ ∈ H$ set

\[
\|ϕ\|_N^* = \sup \{⟨ϕ, ψ⟩ : \|ψ\|_N ≤ 1\}.
\]
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 $N$. Walsh considered families $\{ \| \cdot \|_N \}_{N \in \mathbb{N}}$ decreasing with $N$ and such the family of dual norms also decreases with $N$.

Recall the definition of the constants $C_{n-2}^{(n)}$, $C_0^{(n)}$ by $C_{n-2}^{(n)} = 1$, $C_0^{(n)} = \max\left\{ C_{n-2}^{(n)}, 2\phi(C_{n}^{(n)})^{-1} \right\}$, with $\phi : \mathbb{R}^+ \to \mathbb{R}^+$.

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

$$M_1 \leq M_1 \leq \ldots \leq M_{2n} \leq M^* = O_{M_0, \delta, c, \omega}(1),$$

independent of the norms, such that for any $\varphi \in \mathcal{H}$ can be found some $i \in \{ 1, \ldots, 2n \}$

and some $A, B \in \mathbb{N}$ with $M_1 \leq A \leq M_i \leq \omega(M_i) \leq B$, in such a way that $\varphi$ can be decomposed as $\varphi = \varphi_1 + \varphi_2 + \varphi_3$ with

$$\|\varphi_1\|_B < C_i^{(n)}, \|\varphi_2\|_A < \phi(C_i^{(n)}), \|\varphi_3\| < \delta.$$

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

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

If it is assumed that the norm $\| \cdot \|_{\mathcal{H}_0}$ is well defined and equivalent to $\| \cdot \|$, 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 $\Sigma_{L, g} = \{ \sigma : \sigma \text{ is } (L, g) \text{ - reducible} \}$, 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 $g = (g_1, g_2, \ldots, g_j)$ be a system with $\text{cpl}_L(g) \leq C$, then there exists a natural $K = K_{L, C}$ such that for any function $F : N \to \mathbb{N}$ and for any $M \in \mathbb{N}$, there are natural numbers $M_{K_{L, C}}^1, \ldots, M_{K_{L, C}}^j \geq M$ such that for any maps $\varphi_1, \varphi_2, \ldots, \varphi_j \in L^\infty(X, \mu)$, with $\| \varphi_i \|_\infty \leq 1, i = 1, \ldots, j$ there is a number $i \in \{ 1, 2, \ldots, K_{L, C} \}$ with the property that for any $N, N' \in [M_1, F(M)]$ holds

$$\| S_{N', N'}^g \varphi_1 \varphi_2 \ldots \varphi_j \|_2 < \varepsilon. $$

**Proof.** For any $M \in \mathbb{N}$ and for any function $F : N \to \mathbb{N}$, by the Proposition 3.1 there is a sequence $M \leq M_{K_{L, C}}^1 \leq \ldots \leq M_{K_{L, C}}^j \leq M^*$, thus set $\omega(M) = F(M^*)$. 


By the decomposition theorem \( \varphi_j \) can be written as
\[
\varphi_j = \sum_{t=0}^{k-1} \lambda_t \sigma_t + \psi + \zeta,
\]
with \( \sum_{t=0}^{k-1} |\lambda_t| < C_1 \), with \( C_1 \in [1, \omega^*] \), given by the Lemma 3.2 and \( \sigma_t \in \Sigma_{B, \mathfrak{g}} \) with \( B \geq \omega(M) = F(M^*) \). Moreover, \( \|\psi\|_{M_1} < \phi(C_1) \) and \( \|\zeta\|_2 < \delta \), with \( \delta = \frac{\varepsilon}{2^{3}} \).

To apply Lemma 3.2, a control on \( \psi \) is needed. Let \( S = \{x \in X : |\zeta(x)| < C_1 \} \), 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
\[
|\psi|_{I_{SC}} \leq \sum_{t=0}^{k-1} |\lambda_t| |\sigma_t| I_{SC} |+| \|\zeta\|_{I_{SC}} + |\varphi_j|_{I_{SC}}
\]
\[
\leq \sum_{t=0}^{k-1} |\lambda_t| |\sigma_t| I_{SC} |+| \|\zeta\|_{I_{SC}} \leq 3 |\zeta|_{I_{SC}},
\]
where \( I_A \) is the characteristic function of \( A \). Since \( \sigma_t \in \Sigma_{\omega(M), \mathfrak{g}} \) holds \( \|\psi|_{I_{SC}} \| \leq 3 \|\zeta\|_2 \). By the Chebyshev inequality
\[
(5.2)
\mu(S^c) \leq \frac{\|\zeta\|_2^2}{C_1^2} < \frac{\delta^2}{C_1^2},
\]
Let \( \sigma \in \Sigma_{M, \mathfrak{g}} \), replacing if necessary \( \psi \) by \( \psi_{I_{SC}} \) and \( \zeta \) by \( \zeta + \psi_{I_{SC}} \), it can be assumed that \( \|\zeta\|_\infty < 3C_1 \). In this case we would have \( \|\psi\|_{M_1} < 2\phi(C_1) \) and \( \|\zeta\|_2 < 4\delta \). We have \( |\langle \psi_{I_{SC}}, \sigma \rangle| < \phi(C_1) + 3 \|\zeta\|_2 \mu(S^c)^{1/2} \leq \phi(C_1) + 3\delta^2 + \delta/C_1 < 2\phi(C_1) \). Thus, by the counter reciprocal of the lemma 3.2 for \( N, N' \geq M \)
\[
\left\| S^g_{N', N} [\varphi_1, \varphi_2, \ldots, \zeta] \right\|_2 \leq \varepsilon/3.
\]
By hypothesis \( \|\varphi_j\|_\infty \leq 1 \), and therefore
\[
\left\| S^g_{N', N} [\varphi_1, \varphi_2, \ldots, \psi|_{I_{SC}} + \zeta] \right\|_2 \leq 3\delta + \|\zeta\|_2 < 8\delta, \text{ for any } N.
\]
Thus \( \left\| S^g_{N', N} [\varphi_1, \varphi_2, \ldots, \zeta] \right\|_2 < 16\delta = \varepsilon/6 \). Moreover,
\[
\left\| S^g_{N', N} \left[ \varphi_1, \varphi_2, \ldots, \sum_{t=0}^{k-1} \lambda_t \sigma_t \right] \right\|_2 < \frac{\varepsilon}{3},
\]
for any \( N, N' \in [M_T, F(M_T)] \), for some \( M_T \in [M, \omega(M)] \). Finally
\[
\left\| S^g_{N', N} \left[ \varphi_1, \varphi_2, \ldots, \sum_{t=0}^{k-1} \lambda_t \sigma_t \psi + \zeta \right] \right\|_2 < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\]
Theorem 1.1 will be proved as an immediate consequence of the following:

**Proposition 5.2.** Let $g_1, g_2, \ldots, g_j$ be antihomomorphisms from $\Gamma$ 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

$$G = (g_1, g_1 g_2, \ldots, g_1 g_2 \ldots g_j)$$

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 $\Gamma$. If the map $g$ is an antihomomorphism, i.e., $g(\gamma_1 \gamma_2) = g(\gamma_2) g(\gamma_1)$, then

$$-D_a(g^{-1})(\gamma) = g(\gamma) g(\gamma a)^{-1} = g(\gamma) (g(a) g(\gamma))^{-1} = g(a)^{-1}$$

$$-T_a(g)(\gamma) := g(\gamma a) = g(a) g(\gamma).$$

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

$$g_{1} g_{2} \ldots g_{i} (\gamma) = D_a \left( g_{1} g_{2} \ldots g_{i} \right)(\gamma) T_a(g_{1} g_{2} \ldots g_{i})(\gamma) =$$

$$= (g_{1} g_{2} \ldots g_{j})^{-1}(a)(g_{1} g_{2} \ldots g_{i})(a)(g_{1} g_{2} \ldots g_{i})(\gamma) =$$

$$= (g_{1} g_{2} \ldots g_{i})^{-1}(a)(a) g_{1}(\gamma) g_{2}(a) g_{2}(\gamma) \ldots g_{i}(a) g_{i}(\gamma) =$$

$$= (g_{1} g_{2} \ldots g_{j})(\gamma) g_{i+1}(\gamma)^{-1} \ldots g_{j}(\gamma)^{-1} g_{i+1}(a)^{-1} \ldots g_{j}(a)^{-1}.$$ 

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

$$cplx(G^*) = cplx(G),$$

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

Thus, recall that if $h_i$, $i = 1, 2, \ldots, j$ are measure-preserving actions of $\Gamma$ on $X$, we can define maps $g_i$ from $\Gamma$ 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**


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