The spherical ergodic theorem revisited

Philippe Jaming
Université d'Orléans
Faculté des Sciences
MAPMO - Fédération Denis Poisson
BP 6759
F 45067 Orléans Cedex 2
France

Abstract

In this paper, we give a new proof of a result of R. Jones showing almost everywhere convergence of spherical means of actions of $\mathbb{R}^d$ on $L^p(X)$-spaces are convergent for $d \geq 3$ and $p > \dfrac{d}{d-1}$.

This is done by adapting the proof of the spherical maximal theorem by Rubio de Francia so as to obtain directly the ergodic theorem.

Key words: Spherical maximal theorem; ergodic averages
1991 MSC: 28D05;42B25

In memory of Martine Babillot (1959–2003)

1. Introduction

The aim of this paper is to give a new proof of a spherical ergodic theorem originally du to R. Jones. In order to give a precise statement, let us first give some notation. Througout this paper, d will be an integer with $d \geq 3$. For $r > 0$, we denote by $B(0,r)$ and $S(0,r)$ respectively the (Euclidean) ball and sphere of $\mathbb{R}^d$ centered at 0 and of radius

Email address: `Philippe.Jaming@univ-orleans.fr` (Philippe Jaming).

[1] This paper was part of Martine Babillot *et al*'s project to write a book on ergodic theory (*see* `http://www.univ-orleans.fr/mapmo` for a preliminary manuscript). Unfortunately this project seems unlikely to reach completion within the near future, therefore I have decided to make this results available in another form. I would like to express my gratitude to my colleagues for comments and careful reading of the original french manuscript.
I would also like to thank the anonymous referee for helpful comments that lead to an improvement of the presentation of this paper.

r. The Lebesgue measure on $\mathbb{R}^d$ is simply denoted $\mathrm{d}x$ and the uniform probability measure on $S(0,r)$ is denoted by σ_r. We will simply write $\mathbb{S}^{d-1} = S(0,1)$ and $\sigma = \sigma_1$. We will write $|E|$ for the Lebesgue measure of a subset $E \subset \mathbb{R}^d$.

For $\varphi \in \mathcal{S}(\mathbb{R}^d)$, the Schwartz class on $\mathbb{R}^d$, and $x \in \mathbb{R}^d$, we will write

$$\beta_r \cdot \varphi(x) := \frac{1}{|B(0,r)|} \int_{B(0,r)} \varphi(x+y)\,\mathrm{d}y$$

for the ball-averages, while the sphere averages are denoted by

$$\sigma_r \cdot \varphi(x) := \sigma_r * \varphi(x) = \int_{S(0,r)} \varphi(x+y)\,\mathrm{d}\sigma_r(y) = \int_{\mathbb{S}^{d-1}} \varphi(x+r\zeta)\,\mathrm{d}\sigma(\zeta).$$

We will further denote by φ^*_σ the corresponding maximal function:

$$\varphi^*_\sigma(x) = \sup_{r>0} |\sigma_r \cdot \varphi(x)|.$$

The following theorem was proved by E. M. Stein [10] in the case $d \geq 3$ and by J. Bourgain [1] for $d = 2$.

Theorem 1.1 (Spherical Maximal Theorem)
Let $d \geq 2$ and $p > \dfrac{d}{d-1}$. Then there exists a constant $C = C(p,d)$ such that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\|\varphi^*_\sigma\|_{L^p(\mathbb{R}^d)} \leq C\|\varphi\|_{L^p(\mathbb{R}^d)}. \tag{1.1}$$

Remark : — Note that, as $S(0,r)$ is of measure 0, one can not define $\sigma_r \cdot \varphi(x)$, and *a fortiori* φ^*_σ for an arbitrary L^p-function. Nevertheless, the validity *a priori* of Inequality (1.1) allows to extend the definition of φ^*_σ from $\mathcal{S}(\mathbb{R}^d)$ to L^p, provided $p > \dfrac{d}{d-1}$.

— As is well known, the hypothesis on d and p in Theorem 1.1 are sharp. For $d = 1$, any non-negative function φ will provide a counter-example, whereas for $d \geq 2$, a smoothed version of the characteristic function of a ball will do.

In order to state the Ergodic Theorem, let us introduce some further notation. Throughout the remaining of this paper, $(X, \mathcal{B}, \mu)$ will be a probability space and we will assume that $\mathbb{R}^d$ has a measure-preserving action on X. The action of $y \in \mathbb{R}^d$ on $x \in X$ is denoted by $y \cdot x$. The σ-sub-algebra of $\mathcal{B}$ of $\mathbb{R}^d$-invariant sets will be denoted by $\mathcal{I}$.

For $f \in L^1(X,m)$, the conditional expectation with respect to this σ-algebra is denoted $E(f|\mathcal{I})$. We will further write

$$\beta_r \cdot f(x) = \frac{1}{|B(0,r)|} \int_{B(0,r)} f(y \cdot x)\,\mathrm{d}y$$

and $\sigma_r \cdot f(x) = \displaystyle\int_{\mathbb{S}^{d-1}} f\big((r\zeta)\cdot x\big)\,\mathrm{d}\sigma(\zeta)$.

By some sophisticated arguments based on refinements of the proof of the spherical maximal theorem, the following ergodic theorem was then proved by R. Jones [5] in the case $d \geq 3$ and M. Lacey [6] for $d = 2$:

Theorem 1.2 (Spherical Ergodic Theorem)
Let $d \geq 2$ and $p \geq \dfrac{d}{d-1}$. Let $(X, \mathcal{B}, \mu)$ be a probability space and assume that $\mathbb{R}^d$ has

a measure-preserving action on X. Then, for $f \in L^p(X,m)$, $\sigma_r \cdot f$ converges almost everywhere to $E(f|\mathcal{I})$ as $r \to +\infty$.

Our aim here is to give a new proof of Jones' Theorem, *i.e.* the above theorem in the case $d \geq 3$. The main point here is that one may slightly modify Rubio de Francia's proof of the Spherical Maximal Theorem to obtain simultaneously a proof of the Spherical Ergodic Theorem. This proof is slightly simpler then Stein's original proof. Its main advantage is that it allows for a simpler proof of the Ergodic Theorem that we will present here.

A second ingredient is a lemma already used by Jones that allows to compare spherical averages to ball averages for which Wiener's Ergodic Theorem provides the result. As we will appeal to it, let us recall it now:

Theorem 1.3 (Wiener's Ergodic Theorem)
Let $(X,\mathcal{B},\mu)$ be a probability space and assume that $\mathbb{R}^d$ has a measure-preserving action on X. Let $p \geq 1$ be a real number. Then there exists a constant $C > 0$ such that, for every $f \in L^p(X,m)$,
(i) $\left\|\sup_{r>0}|\beta_r \cdot f|\right\|_p \leq \|f\|_p$ *and*
(ii) $\beta_r \cdot f \to E(f|\mathcal{I})$ *almost everywhere as* $r \to +\infty$.

In order to keep this paper both sufficiently self-contained and concise, we have decided to reproduce here only those elements of the proof of [8] that are specific to spherical averages.

The remaining of this article is organized as follows. In the next section, we complete this introduction by some further notations and preliminary results. The last section is then devoted to providing the proof of Theorem 1.2.

2. Preliminaries

2.1. *Further notations*

In the remaining of the paper, C will be a constant that depends only on the dimension d. As is usual, the exact value of C is irrelevant and may change from line to line. Results in this section may all be found *e.g.* in [3].

The *Fourier Transform* is defined for $\varphi \in \mathcal{S}(\mathbb{R}^d)$ by

$$\widehat{\varphi}(\xi) = \mathcal{F}\varphi(\xi) = \int_{\mathbb{R}^d} \varphi(x) e^{2i\pi\langle x,\xi\rangle}\,\mathrm{d}x$$

and this definition is then extended to L^2 and to bounded measures in the usual way. The *Inverse Fourier Transform* is denote by $\mathcal{F}^{-1}\varphi = \check{\varphi}$.

We will use the following fact: for $\rho \geq 0$, and $\theta \in \mathbb{S}^{d-1}$, $\widehat{\sigma}(\rho\theta) = 2\pi\rho^{1-d/2}J_{d/2-1}(2\pi\rho)$ where J_ν is the Bessel function of order ν. The following estimates are then classical

$$\widehat{\sigma}(\rho\theta) = O\big((1+\rho)^{-\frac{d-1}{2}}\big) \tag{2.1}$$

and

$$\rho\frac{\partial}{\partial\rho}\widehat{\sigma}(\rho\theta) = O\big((1+\rho)^{-\frac{d-1}{2}+1}\big). \tag{2.2}$$

The *Hardy-Littlewood Maximal Function* is defined by

$$\varphi_\beta^*(x) = \sup_{r>0} \beta_r \cdot \varphi(x).$$

Recall that φ_β^* is of weak type $(1,1)$ and of strong type (p,p), $p>1$.

2.2. *The transference principle*

Let us recall that results about actions of $\mathbb{R}^d$ on $\mathbb{R}^d$ by translations can be transferred to actions of $\mathbb{R}^d$ on a probability space $(X,\mathcal{B},\mu)$. The following fact and its proof is a particular case of [5, Theorem 2.2] that fits to our needs.

More precisely, let $(X,\mathcal{B},\mu)$ be a measure space and assume that $\mathbb{R}^d$ acts on X *i.e* that there is a map $\begin{array}{ccc}\mathbb{R}^d \times X & \to & X \\ (t,x) & \mapsto & t\cdot x\end{array}$ that satisfies $s\cdot(t\cdot x) = (s+t)\cdot x$. Assume that this map is (jointly) measurable and *measure preserving*, that is, for every $t\in\mathbb{R}^d$ and every $A\in\mathcal{B}$, $t\cdot A := \{t\cdot a \,:\, a\in A\}\in\mathcal{B}$ and $\mu(t\cdot A)=\mu(A)$.

Next, to a function f on X and $x\in X$, we naturally associate a function φ_x on $\mathbb{R}^d$ via the formula $\varphi_x(t) = f(t\cdot x)$. Note that, if $f\in L^p(X)$ then, for $R>0$ and $a\in\mathbb{R}^d$,

$$\int_X \int_{B(a,R)} |\varphi_x(t)|^p \,\mathrm{d}t \,\mathrm{d}\mu(x) = \int_{B(a,R)} \int_X |f(t\cdot x)|^p \,\mathrm{d}\mu(x)\,\mathrm{d}t = |B(a,R)| \|f\|_p^p.$$

In particular, φ_x is locally in $L^p(\mathbb{R}^d)$ for almost every x.

Now let $\{k_i\}_{i\in I}$ be a family of $L^1(\mathbb{R}^d)$-functions for which there is an R_0 such that $B(0,R_0)$ contains the support of each k_i, $i\in I$. Consider the operator T on $L^p(\mathbb{R}^d)$ defined by

$$T\varphi(t) = \sup_{i\in I} |k_i * f(t)|.$$

Note that T is sub-linear, commutes with translations, and is semi-local, that is, if φ is supported in $B(0,R)$, then $T\varphi$ is supported in $B(0,R+R_0)$.

Finally, assume that there is a constant $C>0$ such that, for every $\varphi\in L^p(\mathbb{R}^d)$, $\|T\varphi\|_p \le C_T\|\varphi\|$. Then T induces an operator $\bar{T}$ on $L^p(X,\mu)$ via the formula $Tf(x) = T\varphi_x(0)$ that satisfies $\|Tf\|_{L^p(X,\mu)} \le C_T \|f\|_{L^p(X,\mu)}$.

Proof. As the k_i's are in $L^1(\mathbb{R}^d)$ with compact support,

$$k_i * \varphi_x(0) = \int_{\mathbb{R}^d} k_i(t)\varphi_x(-t)\,\mathrm{d}t = \int_{B(0,R_0)} k_i(-t) f(t\cdot x)\,\mathrm{d}t$$

is well-defined provided x has been chosen so that $\varphi_x\in L^p_{loc}(\mathbb{R}^d)$. It follows that $Tf(x)$ is well-defined almost everywhere.

Further, note that

$$
\begin{aligned}
Tf(t\cdot x) &= \sup_{i\in I}\left|\int_{\mathbb{R}^d} k_i(-s)f\big(s\cdot(t\cdot x)\big)\,\mathrm{d}s\right|\\
&= \sup_{i\in I}\left|\int_{B(0,R_0)} k_i(-s)f\big((s+t)\cdot x)\big)\,\mathrm{d}s\right|\\
&= \sup_{i\in I}\left|\int_{B(t,R_0)} k_i(t-s)f(s\cdot x)\,\mathrm{d}s\right|.
\end{aligned}
$$

It follows that, if $t\in B(0,R)$, then

$$
\begin{aligned}
Tf(t\cdot x) &= \sup_{i\in I}\left|\int_{B(t,R_0)} k_i(t-s)\varphi_x(s)\chi_{B(0,R+R_0)}(s)\,\mathrm{d}s\right|\\
&= T[\varphi_x\chi_{B(0,R+R_0)}](t).
\end{aligned}
$$

But then, using the fact that the action of $\mathbb{R}^d$ on X is measure preserving,

$$
\begin{aligned}
\|Tf\|_p^p &= \int_X |Tf(x)|^p\,\mathrm{d}\mu(x) = \frac{1}{|B(0,R)|}\int_{B(0,R)}\int_X |Tf(t\cdot x)|^p\,\mathrm{d}\mu(x)\,\mathrm{d}t\\
&= \frac{1}{|B(0,R)|}\int_X\int_{B(0,R)} |Tf(t\cdot x)|^p\,\mathrm{d}t\,\mathrm{d}\mu(x)\\
&= \frac{1}{|B(0,R)|}\int_X\int_{B(0,R)} |T[\chi_{B(0,R+R_0)}\varphi_x](t)|^p\,\mathrm{d}t\,\mathrm{d}\mu(x)\\
&= \frac{1}{|B(0,R)|}\int_X\int_{\mathbb{R}^d} |T[\chi_{B(0,R+R_0)}\varphi_x](t)|^p\,\mathrm{d}t\,\mathrm{d}\mu(x).
\end{aligned}
$$

Finally, as T is bounded on $L^p(\mathbb{R}^d)$,

$$
\begin{aligned}
\|Tf\|_p^p &\leq C\frac{1}{|B(0,R)|}\int_X\int_{\mathbb{R}^d} |\chi_{B(0,R+R_0)}\varphi_x(t)|^p\,\mathrm{d}t\,\mathrm{d}\mu(x)\\
&= C\frac{1}{|B(0,R)|}\int_X\int_{B(0,R+R_0)} |f(t\cdot x)|^p\,\mathrm{d}t\,\mathrm{d}\mu(x)\\
&= C\frac{|B(0,R+R_0)|}{|B(0,R)|}\int_X |f(x)|^p\,\mathrm{d}\mu(x)
\end{aligned}
$$

using again the fact that the action is measure preserving.

The result follows by letting R go to infinity. $\square$

We have only presented a version of the transference principle that fits our needs. The operators under consideration need only to be semi-local and translation invariant. Further, $\mathbb{R}^d$ may be replaced by more general groups, the key property here being its *ameanability, see e.g.* [2] for developments on this theme.

2.3. *A comparison of spherical averages to ball-averages*

We will need the following Lemma:

Lemma 2.1 *Let $(X, \mathcal{B}, \mu)$ be a probability space and assume that $\mathbb{R}^d$ has a measure-preserving action on X. Let $p \geq 1$ be a real number.*

Let $k \in \mathcal{S}(\mathbb{R}^d)$ and assume k is radial and compactly supported. For $r > 0$, let us define the operator K_r on $L^p(X)$ by

$$K_r\varphi(x) = \int_{\mathbb{R}^d} \varphi\big((ry)\cdot x\big)k(y)\,dy$$

where $(ry)\cdot x$ denotes the action of $ry \in \mathbb{R}^d$ on $x \in X$.

Then, for every $\varphi \in L^p(X)$, $K_r\varphi$ converges almost everywhere to

$$\int k(y)\,\mathrm{d}y\,E(\varphi|\mathcal{I}) \tag{2.3}$$

as $r \to +\infty$.

Proof. Let us write $k(u) = k_0(|u|)$. By changing to polar coordinates, we obtain

$$K_r\varphi(x) = d|B(0,1)|\int_0^{+\infty} k_0(\rho)\rho^{d-1}\int_{\mathbb{S}^{d-1}} \varphi\big((r\rho\zeta)\cdot x\big)\,\mathrm{d}\sigma(\zeta)\,\mathrm{d}\rho$$

(recall that σ has been normalized to $\sigma(\mathbb{S}^{d-1}) = 1$). As k_0 is smooth and compactly supported, we may integrate by parts to get that $K_r\varphi(x)$ is equal to

$$\begin{aligned}
&= -\int_0^{+\infty} k_0'(\rho)\left(d|B(0,1)|\int_0^{\rho} t^{d-1}\int_{\mathbb{S}^{d-1}} \varphi\big((rt\zeta)\cdot x\big)\,\mathrm{d}\sigma(\zeta)\,\mathrm{d}t\right)\mathrm{d}\rho\\
&= -\int_0^{+\infty} k_0'(\rho)r^{-d}\int_{B(0,r\rho)} f(y\cdot x)\,\mathrm{d}y\,\mathrm{d}\rho
\end{aligned}$$

by changing back to usual coordinates. This may thus be rewritten as

$$K_r\varphi(x) = -|B(0,1)|\int_0^{+\infty} k_0'(\rho)\rho^d\beta_{r\rho}\cdot\varphi(x)\,\mathrm{d}\rho.$$

According to Wiener's Ergodic Theorem, $\|\sup_{r>0}|\beta_r\cdot f|\|_p \leq \|f\|_p$, in particular, $\sup_{r>0}|\beta_r\cdot f(x)| \leq c(x)$ with $c(x)$ finite for almost every x. Thus, as $k \in \mathcal{S}'$, $|k_0'(\rho)\rho^d\beta_{r\rho}\cdot\varphi(x)| \leq c(x)|k_0'(\rho)\rho^d| \in L^1(0,+\infty)$. Further, for almost every x, $\beta_{r\rho}\cdot\varphi(x) \to E(\varphi|\mathcal{I})$ when $r \to +\infty$. As $k \in \mathcal{S}'$, From Lebesgue's dominated convergence, one then obtains that

$$K_r\varphi(x) \to -|B(0,1)|\int_0^{+\infty} k_0'(\rho)\rho^d\,\mathrm{d}\rho\,E(\varphi|\mathcal{I}).$$

A second integration by parts and a new change to cartesian coordinates then gives (2.3). □

3. Proof of Theorem 1.1 and Theorem 1.2

As announced in the introduction, we will prove both theorems simultaneously. The proof of the maximal theorem is not new, as it is essentially Rubio de Francia's proof with an adaptation that allows to use the transference principal.

3.1. *The Littlewood-Paley decomposition*

We will here slightly modify the standard Littlewood-Paley decomposition. Let $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ be the Fourier transform of a $\mathcal{C}^\infty$-smooth radial compactly supported function. Assume further that $\psi_0(0) = 1$ and that, for $1 < j < \dfrac{d}{2}$,

$$\left(\frac{\partial}{\partial r}\right)^j \psi_0(0) = 0$$

where $\frac{\partial}{\partial r}$ is the radial derivation operator.

Such a function can be constructed in the following way: Let ψ be any function that is the Fourier transform of a $\mathcal{C}^\infty$-smooth radial compactly supported function and such that $\psi(0) = 1$. For $\xi \in \mathbb{S}^{d-1}$ and $r \geq 0$, we then define

$$\psi_0(r\xi) = \left(\sum_{j=0}^{d} a_j r^{2j}\right) \psi(r\xi)$$

where the a_j's are chosen inductively so as to have $\psi_0(0) = 1$ and then the required number of derivatives to vanish at 0.

Let us now define $\psi_1(\xi) = \psi_0(\xi/2) - \psi_0(\xi)$ and, for $j \geq 2$, $\psi_j(\xi) = \psi_1(2^{-j+1}\xi)$. Note that ψ_j is still radial, in $\mathcal{S}(\mathbb{R}^d)$ and is the Fourier transform of a compactly supported function. Moreover, there exists $c, \eta > 0$ such that,

$$\text{for } |\xi| < \eta, \quad |\psi_1(\xi)| \leq c|\xi|^{d/2}. \tag{3.1}$$

Finally, for every $\xi \in \mathbb{R}^d$,

$$\sum_{j=0}^{+\infty} \psi_j(\xi) = 1.$$

(Note that this sum is actually finite for ξ fixed).

Our aim is to get estimates for the maximal operator

$$\varphi_\sigma^* = \sup_{r>0} |\mathcal{F}^{-1}[\mathcal{F}\varphi(\cdot)\mathcal{F}\sigma(r\cdot)]|.$$

For this, we will do a Littlewood-Paley decomposition of this expression. More precisely, let $m_j = \widehat{\sigma}\psi_j$ and let σ_j be the inverse Fourier transform of m_j, $\widehat{\sigma}_j = m_j$. Let $\sigma_{j,r}(x) = r^{-d}\sigma_j(x/r)$. With obvious notations, we then have

$$\varphi_\sigma^* \leq \sum_{j=0}^{\infty} \varphi_{\sigma_j}^*. \tag{3.2}$$

The Spherical Maximal Theorem is then proved if we show that

$$\left\|\varphi_{\sigma_j}^*\right\|_{L^p(\mathbb{R}^d)} \leq C_j \|\varphi\|_{L^p(\mathbb{R}^d)}$$

with $\sum C_j < \infty$. This will be done in three different steps.

3.2. *Comparison of $\varphi^*_{\sigma_j}$ with other maximal functions*

Let P_t be the Poisson kernel on $\mathbb{R}^d$, that is

$$P_t(x) = \frac{c_d t}{(t^2 + |x|^2)^{\frac{d+1}{2}}}$$

where c_d is chosen so that $\displaystyle\int P_t(x)\,\mathrm{d}x = 1$. To P, we will associate the maximal function

$$\varphi^*_P(x) = \sup_{t>0} |P_t * \varphi(x)|.$$

The following lemma allows to compare $\varphi^*_{\sigma_j}$ and φ^*_P.

Lemma 3.1 *There exists a constant $C = C_d$ such that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and for $j \geq 0$,*

$$\varphi^*_{\sigma_j} \leq C 2^j \varphi^*_P.$$

Proof. It is enough to prove that

$$|\sigma_j| \leq C 2^j \frac{1}{(1+|x|)^{d+1}}.$$

Note that, $\sigma_0 = \sigma * \check{\psi}_0$ and for $j \geq 0$, $\sigma_j = \sigma * \check{\psi}_j$ where $\check{\psi}_j(x) = 2^{(j-1)d}\check{\psi}_1(2^{-j+1}x)$ where ψ_0 and ψ_1 are compactly supported $\mathcal{C}^\infty$ functions. Thus there exists C such that $\check{\psi}_0$ and $\check{\psi}_1$ are bounded by $\dfrac{C}{(1+|x|)^{d+1}}$. The proof of Lemma 3.1 is thus completed once we have proved the following lemma. $\square$

Lemma 3.2 *There exists a constant $C = C_d$ such that, for $j \geq 0$ and $x \in \mathbb{R}^d$,*

$$\int_{\mathbb{S}^{d-1}} \frac{2^{jd}}{(1+2^j|x-\xi|)^{d+1}}\,\mathrm{d}\sigma(\xi) \leq C \frac{2^j}{(1+|x|)^{d+1}}. \tag{3.3}$$

Proof. For $|x| > 2$, $|x-\xi| \geq |x|/2$ so that the left hand side of (3.3) is bounded by $2^{-j+d+1}|x|^{-d-1}$ which allows to conclude.

Let us now assume $|x| \leq 2$ and cut the integral into dyadic pieces. The left hand side of (3.3) is bounded by

$$2^{dj} \int_{|\xi - x| \leq 2^{-j}} \mathrm{d}\sigma(\xi) + \sum_{k=0}^{+\infty} 2^{dj} 2^{-(d+1)k} \int_{|\xi - x| \leq 2^{k-j+1}} \mathrm{d}\sigma(\xi).$$

It remains to notice that $\sigma\Big(\{\xi\ :\ |\xi - x| < r\}\Big) \leq C r^{d-1}$ to conclude. $\square$

Finally, the Poisson maximal function is bounded by the Hardy-Littlewood maximal function:

Lemma 3.3 *There exists a constant C such that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$,*

$$\varphi^*_P \leq C \varphi^*_\beta.$$

The proof of this fact is classical and can be found in any book on Hardy spaces. Let us however reproduce it here.

Proof. From invariance under translations and dilations, it is enough to prove that if φ is non-negative,

$$\int_{\mathbb{R}^d} \frac{\varphi(x)}{(1+|x|^2)^{\frac{d+1}{2}}}\,\mathrm{d}x \le C\varphi_\beta^*(0) \tag{3.4}$$

where C does not depend on φ. But this integral is bounded by

$$\int_{|x|\le 1} \varphi(x)\,\mathrm{d}x + \sum_{k=0}^{+\infty} 2^{-k(d+1)} \int_{2^k \le |x| \le 2^{k+1}} \varphi(x)\,\mathrm{d}x.$$

Further $\displaystyle\int_{|x|\le 1} \varphi(x)\,\mathrm{d}x$ is bounded by $|B(0,1)|\varphi_\beta^*(0)$ while the remaining integrals are bounded by

$$|B(0,2^{k+1})|\varphi_\beta^*(0) = 2^{(k+1)d}|B(0,1)|\varphi_\beta^*(0).$$

The estimate (3.4) follows immediately. □

Further, as is well known, the Hardy-Littlewood maximal function is of weak-type $(1,1)$. Grouping all results of this section, we thus get the following:

Proposition 3.4 *There exists a constant C such that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$, for $j \ge 0$,*

$$|\{x\ :\ \varphi_{\sigma_j}^*(x) \ge \alpha\}| \le C2^j \frac{\|\varphi\|_{L^1(\mathbb{R}^d)}}{\alpha}.$$

3.3. *The L^2-estimate of $\varphi_{\sigma_j}^*$*

Let us denote $\sigma_{j,r}(x) = r^{-d}\sigma_j(x/r)$, $\sigma_{j,r}\cdot\varphi = \sigma_{j,r} * \varphi$ so that $\widehat{\sigma_{j,r}\cdot\varphi}(\xi) = \widehat{\varphi}(\xi)m_j(r\xi)$. Let us write

$$G_j(\varphi)(x) = \left(\int_0^{+\infty} |\sigma_{j,r}\cdot\varphi(x)|^2 \frac{\mathrm{d}r}{r}\right)^{1/2}$$

for the associated Littlewood-Paley g-functional.

Let us further write $\tilde\sigma_{j,r}(x) = r\dfrac{\mathrm{d}}{\mathrm{d}r}\sigma_{j,r}(x)$, $\tilde\sigma_{j,r}\cdot\varphi = \tilde\sigma_{j,r} * \varphi$ and

$$\begin{aligned} g_j(\varphi)(x) &= \left(\int_0^{+\infty} |\tilde\sigma_{j,r}\cdot\varphi(x)|^2 \frac{\mathrm{d}r}{r}\right)^{1/2} \\ &= \left(\int_0^{+\infty} r\left|\frac{\mathrm{d}}{\mathrm{d}r}\sigma_{j,r}\cdot\varphi(x)\right|^2 \mathrm{d}r\right)^{1/2}. \end{aligned}$$

From Plancherel's Identity and Fubini's Theorem, we get that

$$\begin{aligned} \|G_j(\varphi)\|_{L^2(\mathbb{R}^d)}^2 &= \int_0^{+\infty} \int_{\mathbb{R}^d} |\widehat{\varphi}(\xi)m_j(r\xi)|^2\,\mathrm{d}\xi\,\frac{\mathrm{d}r}{r} \\ &= \int_{\mathbb{R}^d} |\widehat{\varphi}(\xi)|^2 \int_0^{+\infty} |m_j(r\xi)|^2\,\frac{\mathrm{d}r}{r}\mathrm{d}\xi. \end{aligned}$$

But $|m_j(u)|^2 = |\widehat{\sigma(u)}\psi_j(u)|^2 \le C(1+|u|)^{-d+1}|\psi_j(u)|^2$ by (2.1). It follows that

$$\int_0^{+\infty} |m_j(r\xi)|^2 \frac{\mathrm{d}r}{r} \le C \int_0^{+\infty} \frac{|\psi_j(r\xi)|^2}{(1+|r\xi|)^{d-1}} \frac{\mathrm{d}r}{r}$$
$$= C \int_0^{+\infty} \frac{\left|\psi_1\left(s\frac{\xi}{|\xi|}\right)\right|^2}{(1+s2^{j-1})^{d-1}} \frac{\mathrm{d}s}{s}$$

with the change of variable $s = 2^{-j+1} r|\xi|$. But then

$$\int_0^{+\infty} |m_j(r\xi)|^2 \frac{\mathrm{d}r}{r} \le C2^{-j(d-1)} \int_0^{+\infty} \frac{\left|\psi_1\left(s\frac{\xi}{|\xi|}\right)\right|^2}{s^d} \mathrm{d}s$$

and this last integral is finite by construction of ψ_1. As a consequence, we obtain

$$\|G(\varphi)\|_{L^2(\mathbb{R}^d)} \le C2^{-j(d-1)/2} \|\varphi\|_{L^2(\mathbb{R}^d)}. \tag{3.5}$$

In a similar way, using (2.2), we obtain

$$\|g(\varphi)\|_{L^2(\mathbb{R}^d)} \le C2^{-j(d-3)/2} \|\varphi\|_{L^2(\mathbb{R}^d)}. \tag{3.6}$$

We are now in a position to prove the following:

Proposition 3.5 *There exists a constant C such that, for every integer $j \ge 1$ and every $\varphi \in \mathcal{S}(\mathbb{R}^d)$,*

$$\left\|\varphi^*_{\sigma_j}\right\|_{L^2(\mathbb{R}^d)} \le C2^{-(d-2)j/2} \|\varphi\|_{L^2(\mathbb{R}^d)}.$$

Proof. As $\lim\limits_{r\to+\infty} \sigma_{j,r}(\varphi) = 0$, we get

$$\sigma_{j,r}(\varphi)(x)^2 = -2 \int_r^{+\infty} \sigma_{j,s}(\varphi)(x) s \frac{\mathrm{d}}{\mathrm{d}s} \sigma_{j,s}(\varphi)(x) \ \frac{\mathrm{d}s}{s}$$
$$= -2 \int_r^{+\infty} \sigma_{j,s}(\varphi)(x) \widetilde{\sigma}_{j,s}(\varphi)(x) \frac{\mathrm{d}s}{s}$$
$$\le 2 \int_0^{+\infty} |\sigma_{j,s}(\varphi)(x)| |\widetilde{\sigma}_{j,s}(\varphi)(x)| \frac{\mathrm{d}s}{s}.$$

From Cauchy-Schwarz, we deduce that

$$\sup_{r>0} |\sigma_{j,r}(\varphi)(x)|^2 \le 2G_j(\varphi)(x) g_j(\varphi)(x).$$

Integrating this inequality over $\mathbb{R}^d$ and appealing again to Cauchy-Schwarz, we obtain

$$\left\|\varphi^*_{\sigma_j}\right\|^2_{L^2(\mathbb{R}^d)} \le 2\|G_j(\varphi)\|_{L^2(\mathbb{R}^d)} \|g_j(\varphi)\|_{L^2(\mathbb{R}^d)} \le C_d 2^{-j(d-2)} \|\varphi\|^2_{L^2(\mathbb{R}^d)},$$

with (3.5) et (3.6). □

3.4. *The last step*

By interpolation between the strong type (2,2) estimate given in Proposition 3.5 and the weak type (1,1) estimate of Proposition 3.4 gives the existence, for each p with $1 < p \le 2$, of a constant C_p such that, for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and every $j \ge 1$

$$\left\|\varphi^*_{\sigma_j}\right\|_{L^p(\mathbb{R}^d)} \le C_p 2^{\frac{d-(d-1)p}{p} j} \|\varphi\|_{L^p(\mathbb{R}^d)}. \tag{3.7}$$

By interpolation between the weak type (1,1) estimate of Proposition 3.4 and the (trivial) strong type (∞,∞) estimate shows that the same is true for $j=0$.

Let us recall that σ_j is compactly supported. It follows that, for each $R>0$, the transference principle applies to

$$\varphi^*_{\sigma_j,R}(x) := \sup_{0<r<R} |\sigma_{j,r}\cdot\varphi|.$$

We thus get that there exists a constant $C>0$ such that, for every $R>0$, every $j\geq 0$, every $f\in L^p(X,m)$ $(1<p\leq 2)$,

$$\left\|\sup_{0<r<R}|\sigma_{j,r}\cdot f|\right\|_{L^p(X,m)} \leq C2^{\frac{d-(d-1)p}{p}j}\|f\|_{L^p(X,m)}.$$

As the left hand side does not depend on R, we thus get that

$$\left\|\sup_{r>0}|\sigma_{j,r}\cdot f|\right\|_{L^p(X,m)} \leq C2^{\frac{d-(d-1)p}{p}j}\|f\|_{L^p(X,m)}.$$

Finally, note that if $p>\dfrac{d-1}{d}$, then $\dfrac{d-(d-1)p}{p}<0$ so that, interpolating with the trivial (∞,∞) estimate, we get that, for each $p>\dfrac{d}{d-1}$, there exists $Q_p>0$ such that, for every $j\geq 0$, every $f\in L^p(X,m)$,

$$\left\|\sup_{0<r<R}|\sigma_{j,r}\cdot f|\right\|_{L^p(X,m)} \leq C_p 2^{-Q_p j}\|f\|_{L^p(X,m)}. \tag{3.8}$$

As the right hand side is independent on R, the Monotone Convergence Theorem implies that we may replace R by $+\infty$ in (3.8).

From this, we get that $\displaystyle\sum_{j=0}^{J}\sigma_{j,r}\cdot f$ is uniformly convergent in $L^p(X)$. This allows us to define $\sigma_r\cdot f$ as its limit. Moreover, we obtaint the following bound:

$$\begin{aligned}\left\|\sup_{0<r<+\infty}\left|\sigma_r\cdot f-\sum_{j=0}^{J}\sigma_{j,r}\cdot f\right|\right\|_{L^p(X)} &= \left\|\sup_{0<r<+\infty}\left|\sum_{j=J+1}^{+\infty}\sigma_{j,r}\cdot f\right|\right\|_{L^p(X)}\\ &\leq C\sum_{j=J+1}^{+\infty}2^{-Q_p j}\|f\|_{L^p(X,m)}\end{aligned} \tag{3.9}$$

and thus goes to 0 as $J\to+\infty$. Finally, from Lemma 2.1, we obtain that

$$\sum_{j=0}^{J}\sigma_{j,r}\cdot f \to \int_{\mathbb{R}^d}\sum_{j=0}^{J}\sigma_{j,r}(|x|)\,\mathrm{d}x\,E(f|\mathcal{I}) = E(f|\mathcal{I})$$

almost everywhere as $r\to+\infty$. Combining this with (3.9), one immediately obtains that $\sigma_r\cdot f\to E(f|\mathcal{I})$ as well. This completes the proof of the Spherical Ergodic Theorem when $d\geq 3$. □

4. Conclusion

In this paper, we have shown how to obtain the Spherical Ergodic Theorem from the proof of the Maximal Ergodic Theorem. The main feature is that, in order to appeal to the transference principle, one needs to use a compactly supported Littlewood-Paley decomposition that is well localized in frequency instead of a standard decomposition that has compactly supported Fourier transform.

Several results about maximal functions could thus be transformed into ergodic theorems. Let us mention a few. For instance, the case $d = 2$ (*i.e.* Lacey's Ergodic Theorem) could be obtained by adapting the proof of the Circular Maximal Theorem of [7]. One may also obtain Lacunary Ergodic Theorems by following the proofs in [9]. Let us for instance mention the following result which follows from the proof of their Theorem 1.1:

Corollary 4.1 *Let $d \geq 2$, $\alpha > 0$ and $p \geq 1 + [(d-1)(\alpha+1)]^{-1}$. Let $(X, \mathcal{B}, \mu)$ be a probability space and assume that $\mathbb{R}^d$ has a measure-preserving action on X. Let $\{t_j\}$ be a sequence such that $t_j \to +\infty$ and $\{t_j\} \subset \{2^k(1 + l^{-\alpha}) : k \in \mathbb{Z}_+, l \in \mathbb{Z}\}$. Then, for $f \in L^p(X, m)$, $\sigma_{t_j} \cdot f$ converges almost everywhere to $E(f|\mathcal{I})$ as $j \to +\infty$.*

More general results can also be obtained from Theorem I to IV of [9]. We refrain from introducing the lengthy notation needed to state those results.

We would also like to stress that a key ingredient in the proof is the decay estimate of $\widehat{\sigma}$. Such estimates are available for large classes of measures like the surface measure of the boundary of a smooth convex set with non-vanishing curvature. For maximal theorems that can be transformed into ergodic theorems with the method exposed in this paper, we refer *e.g.* to [**?**].

References

[1] J. Bourgain *Averages in the plane over convex curves and maximal operators.* J. Anal. Math. **47** (1986), 69–85.

[2] R. R. Coifman & G. Weiss *Transference methods in analysis.* American Mathematical Society, Providence, R.I., 1976. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, **31**.

[3] L. Grafakos *Classical and Modern Fourier Analysis* . Prentice Hall, 2004.

[4] A. Iosevich & E. Sawyer *Oscillatory integrals and maximal averages over homogeneous surfaces.* Duke Math. J. **82** (1996), 103–141.

[5] R. Jones *Ergodic averages on spheres.* J. Anal. Math. **61** (1993), 29–45.

[6] M. T. Lacey *Ergodic averages on circles.* J. Anal. Math. **67** (1995), 199–206.

[7] G. Mockenhaupt, A. Seeger & C. D. Sogge *Wave front sets, local smoothing and Bourgain's circular maximal theorem.* Ann. Math. **136** (1992), 207–218.

[8] J. L. Rubio de Francia *Maximal functions and Fourier transforms.* Duke Math. J. **53** (1986), 395–404.

[9] A. Seeger, T. Tao & J. Wright *Endpoint mapping properties of spherical maximal operators.* J. Inst. Math. Jussieu **2** (2003), 109–144.

[10] E. M. Stein *Maximal functions. I. Spherical means.* Proc. Nat. Acad. Sci. U.S.A. **73** (1976), 2174–2175.

[11] N. Wiener *The ergodic theorem.* Duke Math. J. **5** (1939), 1-18.