A local inequality for Hankel operators on the sphere and its application

Quanlei Fang

CUNY-BCC

HFS, Gargnano, May 2013

This talk is based on a joint work with Jingbo Xia

INTRODUCTION

Basic setting:

$S = \{z \in \mathbf{C}^n : |z| = 1\}$, the unit sphere in $\mathbf{C}^n$.

$\sigma =$ the spherical measure on S with $\sigma(S) = 1$.

$H^2(S) =$ the Hardy space on S, which is the closure of the polynomials in $z_1, ..., z_n$ in $L^2(S, d\sigma)$.

$P =$ the orthogonal projection from $L^2(S, d\sigma)$ onto $H^2(S)$, which is given by the Cauchy integral formula.

Hankel operator $H_f : H^2(S) \to L^2(S, d\sigma) \ominus H^2(S)$ is defined by the formula

$$H_f = (1 - P)M_f | H^2(S).$$

Two kinds of problems

- "Two-sided" problems:
  - Concern H_f and $H_{\bar{f}}$ simultaneously.
  - By virtue of the relation

$$[P, M_f] = H_{\bar{f}}^* - H_f,$$

    "two-sided" problems are equivalent to the study of the commutator $[P, M_f]$.
- "One-sided" problems are
  - the study of H_f alone;
  - new challenges compared with the corresponding "two-sided" problems.

Normalized reproducing kernel for $H^2(S)$:

$$k_z(w) = \frac{(1-|z|^2)^{n/2}}{(1-\langle w, z\rangle)^n}, \quad |w| \le 1.$$

The main difference between the "one-sided" theory and "two-sided" theory is the inequality

$$\|(f - \langle fk_z, k_z\rangle)k_z\|^2 \le \|H_f k_z\|^2 + \|H_{\bar{f}} k_z\|^2.$$

This inequality is extremely important for the "two-sided" theory. So far, the challenge of the "one-sided" theory has been to find a way to get around this in each specific problem. That is, the difficulty in the "one-sided" theory is the recovery of the function-theoretic properties of $f - Pf$ from the operator-theoretic properties of H_f alone.

The boundedness and compactness of H_f were characterized in terms of mean oscillation:

Theorem 1. (Xia, 2008) Let $f \in L^2(S, d\sigma)$.
(1) If

$$\sup_{|z|<1} \|H_f k_z\| < \infty,$$

then $f - Pf \in \text{BMO}$. Therefore the Hankel operator H_f is bounded if and only if $f - Pf \in \text{BMO}$.
(2) If

$$\lim_{|z|\uparrow 1} \|H_f k_z\| = 0,$$

then $f - Pf \in \text{VMO}$. Therefore the Hankel operator H_f is compact if and only if $f - Pf \in \text{VMO}$.

Remark. This is the high-dimensional analogue of the classic theorems of Nehari and Hartman. The main difference between the case $n = 1$ and $n \geq 2$ is that there are no genuinely "one-sided" problems when $n = 1$. This is because in the circle case we always have

$$H_f = [M_{f-Pf}, P].$$

When $n \geq 2$, the main difficulty is to deal with Hankel operators H_f that cannot be expressed in the form of a commutator.

Later, the Schatten-class membership of H_f in the Hardy space of the sphere was also characterized.

For each $1 \le p < \infty$, the Schatten class $\mathcal{C}_p$ consists of operators A satisfying the condition $\|A\|_p < \infty$, where

$$\|A\|_p = \{\mathrm{tr}((A^*A)^{p/2})\}^{1/p}.$$

In terms of the s-numbers $s_1(A)$, $s_2(A)$, ..., $s_j(A)$, ... of A, we have

$$\|A\|_p = \left(\sum_{j=1}^{\infty} \{s_j(A)\}^p\right)^{1/p}.$$

Recall that

$$s_j(A) = \inf\{\|A + K\| : \mathrm{rank}(K) \le j - 1\}.$$

The Schatten-class result concerns Besov spaces and the Möbius-invariant measure on **B**.

Definition (a) For each $1 \le p < \infty$ and each $g \in L^2(S, d\sigma)$, denote

$$\mathcal{I}_p(g) = \iint \frac{|g(\zeta) - g(\xi)|^p}{|1 - \langle \zeta, \xi \rangle|^{2n}} d\sigma(\zeta) d\sigma(\xi).$$

(b) For each $1 \le p < \infty$, the Besov space $\mathcal{B}_p$ consists of those $g \in L^2(S, d\sigma)$ which satisfy the condition $\mathcal{I}_p(g) < \infty$.

Definition. Let dv be the volume measure on **B** with the normalization $v(\mathbf{B}) = 1$. Denote

$$d\lambda(z) = \frac{dv(z)}{(1 - |z|^2)^{n+1}}.$$

It is well known that $d\lambda$ is invariant under the Möbius transform of the ball.

For $g \in L^2(S, d\sigma)$, denote

$$\mathrm{Var}(g; z) = \|(g - \langle gk_z, k_z\rangle)k_z\|^2,$$

which is the "variance" of g with respect to the probability measure $|k_z|^2 d\sigma$.

Theorem 2. (Fang and Xia, 2009) Let $2n < p < \infty$ and $f \in L^2(S, d\sigma)$. Then the following are equivalent:
(a) H_f belongs to the Schatten class $\mathcal{C}_p$.
(b) $f - Pf \in \mathcal{B}_p$.
(c)

$$\int \|H_f k_z\|^p d\lambda(z) < \infty.$$

(d)

$$\int \mathrm{Var}^{p/2}(f - Pf; z) d\lambda(z) < \infty.$$

Remark. Theorem 2 is the high-dimensional analogue of Peller's characterization of Schatten-class membership of Hankel operators on the unit circle.

But there is a distinct difference between the case $n = 1$ and the case $n \geq 2$: Unlike Peller's classic result on the unit circle, in the case $n \geq 2$ there is a complete "cutoff line" for Schatten class Hankel operators at $p = 2n$.

Theorem 3. (Fang and Xia, 2009) Suppose that $n \geq 2$. Let $f \in L^2(S, d\sigma)$. If H_f is bounded and if $H_f \neq 0$, then there exists an $\epsilon = \epsilon(f) > 0$ such that

$$s_1(H_f) + ... + s_k(H_f) \geq \epsilon k^{(2n-1)/2n}$$

for every $k \in \mathbf{N}$.

Corollary. Suppose that $n \geq 2$. Let $f \in L^2(S, d\sigma)$. If H_f belongs to the Schatten class $\mathcal{C}_{2n}$, then $H_f = 0$.

A big hurdle that had to be overcome in order to prove Theorems 1 and 2 was the fact that, at the time, there was nothing comparable to the "two-sided" local inequality

$$\|(f - \langle fk_z, k_z\rangle)k_z\|^2 \leq \|H_f k_z\|^2 + \|H_{\bar{f}} k_z\|^2,$$

if one only had H_f to work with.

A local inequality and its application

To state the inequality, we need to introduce a sequence of contractions in the radial direction of the ball. For each pair of $j \in \mathbf{N}$ and $z \in \mathbf{B}$, define

$$\rho_j(z) = \begin{cases} (1 - 4^j(1 - |z|^2))^{1/2}(z/|z|) & \text{if} \quad 4^j(1 - |z|^2) < 1; \\ 0 & \text{if} \quad 4^j(1 - |z|^2) \geq 1. \end{cases}$$

In other words, if $4^j(1 - |z|^2) < 1$, then

$$\begin{cases} \rho_j(z)/|\rho_j(z)| = z/|z| & \text{and} \\ 1 - |\rho_j(z)|^2 = 4^j(1 - |z|^2) \end{cases}.$$

Some notation:

For $g \in L^2(S, d\sigma)$, we write

$$\mathrm{Var}(g; z) = \|(g - \langle gk_z, k_z \rangle)k_z\|^2, \quad z \in \mathbf{B},$$

which is the "variance" of g with respect to $|k_z|^2 d\sigma$.

For each $z \in \mathbf{B}$, define

$$m_z(w) = \frac{1 - |z|}{1 - \langle w, z \rangle},$$

$|w| \leq 1$. We call m_z a Schur multiplier.

Here is our one-sided local inequality:

Theorem 4. Given any $0 < \delta \leq 1/2$, there exists a constant $0 < C(\delta) < \infty$ which depends only on δ and the complex dimension n such that the inequality

$$(*) \qquad \mathrm{Var}^{1/2}(f - Pf; z) \leq C(\delta) \sum_{j=1}^{\infty} \frac{1}{2^{(1-\delta)j}} \|M_{m_{\rho_j(z)}} H_f k_{\rho_j(z)}\|$$

holds for all $f \in L^2(S, d\sigma)$ and $z \in \mathbf{B}$.

Since $\|M_{m_z}\| = \|m_z\|_\infty = 1$, a slightly weaker, but perhaps aesthetically more pleasing version of $(*)$ is

$$(**) \qquad \mathrm{Var}^{1/2}(f - Pf; z) \le C(\delta) \sum_{j=1}^{\infty} \frac{1}{2^{(1-\delta)j}} \|H_f k_{\rho_j(z)}\|,$$

$f \in L^2(S, d\sigma)$ and $z \in \mathbf{B}$. From $(**)$ we see immediately that if

$$\sup_{|z|<1} \|H_f k_z\| < \infty,$$

then $f - Pf \in \mathrm{BMO}$. Similarly, we can deduce from $(**)$ that the condition

$$\lim_{|z|\uparrow 1} \|H_f k_z\| = 0$$

implies $f - Pf \in \mathrm{VMO}$. In other words, $(**)$ recaptures Theorem 1, and indeed explains why Theorem 1 holds.

Our main application is to characterize the membership of the Hankel operator H_f in the Lorentz-like ideal $\mathcal{C}_p^+$, $2n < p < \infty$.

For each $1 \le p < \infty$, the formula

$$\|A\|_p^+ = \sup_{k \ge 1} \frac{s_1(A) + s_2(A) + \cdots + s_k(A)}{1^{-1/p} + 2^{-1/p} + \cdots + k^{-1/p}}$$

defines a symmetric norm for operators, where $s_1(A), \ldots, s_k(A), \ldots$ are the s-numbers of A. On any separable Hilbert space $\mathcal{H}$, the set

$$\mathcal{C}_p^+ = \{A \in \mathcal{B}(\mathcal{H}) : \|A\|_p^+ < \infty\}$$

is a norm ideal. It is well known that $\mathcal{C}_p^+$ contains the Schatten class $\mathcal{C}_p$ and that $\mathcal{C}_p^+ \neq \mathcal{C}_p$.

Let us also recall the general notation of symmetric gauge function.

Let $\hat{c}$ be the linear space of sequences $\{a_j\}_{j\in\mathbf{N}}$, where $a_j \in \mathbf{R}$ and for each sequence $a_j \neq 0$ only for a finite number of j's. A symmetric gauge function (also called symmetric norming function) is a map

$$\Phi : \hat{c} \to [0, \infty)$$

that has the following properties:

(a) Φ is a norm on $\hat{c}$.

(b) $\Phi(\{1, 0, \ldots, 0, \ldots\}) = 1$.

(c) $\Phi(\{a_j\}_{j\in\mathbf{N}}) = \Phi(\{|a_{\pi(j)}|\}_{j\in\mathbf{N}})$ for every bijection $\pi : \mathbf{N} \to \mathbf{N}$.

Such a Φ gives rise to the symmetric norm

$$\|A\|_{\Phi} = \sup_{k \geq 1} \Phi(\{s_1(A), \ldots, s_k(A), 0, \ldots, 0, \ldots\})$$

for operators. On any separable Hilbert space $\mathcal{H}$,

$$\mathcal{C}_{\Phi} = \{A \in \mathcal{B}(\mathcal{H}) : \|A\|_{\Phi} < \infty\}$$

is a norm ideal.

The term norm ideal refers to the following properties of $\mathcal{C}_\Phi$:

• For any $B, C \in \mathcal{B}(\mathcal{H})$ and $A \in \mathcal{C}_\Phi$, $BAC \in \mathcal{C}_\Phi$ and $\|BAC\|_\Phi \leq \|B\|\|A\|_\Phi\|C\|$.

• If $A \in \mathcal{C}_\Phi$, then $A^* \in \mathcal{C}_\Phi$ and $\|A^*\|_\Phi = \|A\|_\Phi$.

• For any $A \in \mathcal{C}_\Phi$, $\|A\| \leq \|A\|_\Phi$, and the equality holds when $\mathrm{rank}(A) = 1$.

• $\mathcal{C}_\Phi$ is complete with respect to $\|.\|_\Phi$.

For any unbounded operator X, we simply set $\|X\|_\Phi = \infty$ by convention.

The domain of definition of a symmetric gauge Φ is usually extended beyond the space $\hat{c}$. If $\{b_j\}_{j\in\mathbf{N}}$ is an arbitrary sequence of real numbers, then we define

$$\Phi(\{b_j\}_{j\in\mathbf{N}}) = \sup_{k\geq 1} \Phi(\{b_1, \ldots, b_k, 0, \ldots, 0, \ldots\}).$$

For our purpose we also need to deal with sequences indexed by sets other than $\mathbf{N}$. If W is a countable, infinite set, then we define

$$\Phi(\{b_\alpha\}_{\alpha\in W}) = \Phi(\{b_{\pi(j)}\}_{j\in\mathbf{N}}),$$

where $\pi : \mathbf{N} \to W$ is any bijection. The definition of symmetric gauge functions guarantees that the value of $\Phi(\{b_\alpha\}_{\alpha\in W})$ is independent of the choice of the bijection π.

In particular, associated with the ideal $\mathcal{C}_p^+$ is the symmetric gauge function Φ_p^+, which is defined as follows. Let $1 \leq p < \infty$. For each $\{a_j\}_{j\in\mathbf{N}} \in \hat{c}$, define

$$\Phi_p^+(\{a_j\}_{j\in\mathbf{N}}) = \sup_{k\geq 1} \frac{|a_{\pi(1)}| + \cdots + |a_{\pi(k)}|}{1^{-1/p} + \cdots + k^{-1/p}},$$

where $\pi : \mathbf{N} \to \mathbf{N}$ is any bijection such that $|a_{\pi(j)}| \geq |a_{\pi(j+1)}|$ for every $j \in \mathbf{N}$, which exists because $a_j = 0$ for all but a finite number of j's. Then

$$\mathcal{C}_p^+ = \mathcal{C}_{\Phi_p^+}.$$

More precisely, the relation between the norm $\|\cdot\|_p^+$ and symmetric gauge function Φ_p^+ is that

$$\|A\|_p^+ = \Phi_p^+(\{s_1(A), \ldots, s_k(A), \ldots\}).$$

In the language of symmetric gauge functions, Theorem 3 (2009) has the following interpretation:

Corollary. Suppose that $n \geq 2$. Let Φ be a symmetric gauge function. If there is an $f \in L^2(S, d\sigma)$ such that $H_f \in \mathcal{C}_\Phi$ and $H_f \neq 0$, then $\mathcal{C}_\Phi \supset \mathcal{C}_{2n}^+$.

In other words, $\mathcal{C}_{2n}^+$ is the smallest ideal of the form $\mathcal{C}_\Phi$ that contains any $H_f \neq 0, f \in L^2(S, d\sigma)$.

On the other hand, if $f \in \text{Lip}(S)$, then $H_f \in \mathcal{C}_{2n}^+$. So $\mathcal{C}_{2n}^+$ does contain plenty of nonzero Hankel operators.

Our characterization result involves the Bergman structure of the ball. For each $z \in \mathbf{B} \backslash \{0\}$, we have the Möbius transform

$$\varphi_z(w) = \frac{1}{1 - \langle w, z \rangle} \left\{ z - \frac{\langle w, z \rangle}{|z|^2} z - (1 - |z|^2)^{1/2} \left(w - \frac{\langle w, z \rangle}{|z|^2} z \right) \right\},$$

$w \in \mathbf{B}$. We define $\varphi_0(w) = -w$. The Bergman metric on $\mathbf{B}$ is given by the formula

$$\beta(z, w) = \frac{1}{2} \log \frac{1 + |\varphi_z(w)|}{1 - |\varphi_z(w)|}.$$

For $z \in \mathbf{B}$ and $a > 0$, we define the corresponding β-ball

$$D(z, a) = \{w \in \mathbf{B} : \beta(z, w) < a\}.$$

Definition. (i) Let $a > 0$. A subset Γ of $\mathbf{B}$ is said to be a-separated if $D(z, a) \cap D(w, a) = \emptyset$ for all $z \neq w$ in Γ.

(ii) Let $0 < a < b < \infty$. A subset Γ of $\mathbf{B}$ is said to be an a, b-lattice if it is a-separated *and* has the property $\cup_{z \in \Gamma} D(z, b) = \mathbf{B}$.

The simplest example of such a lattice is the following: Take any $a > 0$, and then take any subset M of $\mathbf{B}$ that is maximal with respect to the property of being a-separated. Then M is an $a, 2a$-lattice in $\mathbf{B}$.

With the above preparation, we can now state our characterization result.

Theorem 5. Let $2n < p < \infty$ be given. Let $0 < a < b < \infty$ be positive numbers such that $b \geq 2a$. Then there exist constants $0 < c \leq C < \infty$ which depend only on the given p, a, b and the complex dimension n such that the inequality

$$\begin{aligned} c\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z\in\Gamma}) & \\ &\leq \|H_f\|_p^+ \\ &\leq C\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z\in\Gamma}) \end{aligned}$$

holds for every $f \in L^2(S, d\sigma)$ and every a, b-lattice Γ in $\mathbf{B}$.

Corollary. Suppose that $2n < p < \infty$ and that Γ is an a, b-lattice in $\mathbf{B}$, $b \geq 2a$. Then for $f \in L^2(S, d\sigma)$, we have $H_f \in \mathcal{C}_p^+$ if and only if

$$\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z\in\Gamma}) < \infty.$$

Interpretation: For the purpose of determining the membership $H_f \in \mathcal{C}_p^+$, $2n < p < \infty$, the set of numbers

$$\{\mathrm{Var}^{1/2}(f - Pf; z) : z \in \Gamma\}$$

is just as good as the set of s-numbers

$$\{s_j(H_f) : j \in \mathbf{N}\}$$

of H_f.

An Open Problem

An Open Problem. How does one characterize the membership of H_f in the critical ideal $\mathcal{C}_{2n}^+$?

In the case $n \geq 2$, $\mathcal{C}_{2n}^+$ is the critical ideal because it is the smallest ideal of the form $\mathcal{C}_\Phi$ that contains nonzero Hankel operators.

It is natural to speculate that Theorem 5 holds in the case $p = 2n$.

Indeed the upper bound does hold in the case $p = 2n$.

But the lower bound is a different story altogether.

Because of the "decay rate" $1-\delta$ in the local inequality

$$\mathrm{Var}^{1/2}(f-Pf;z) \le C(\delta)\sum_{j=1}^{\infty}\frac{1}{2^{(1-\delta)j}}\|M_{m_{\rho_j(z)}}H_f k_{\rho_j(z)}\|,$$

our proof of the lower bound in Theorem 5 does not extend to the case $p=2n$.

Thanks for your attention!

Some Technical Aspects of the New Results

First of all, the proof of the local inequality

$$\mathrm{Var}^{1/2}(f - Pf; z) \leq C(\delta) \sum_{j=1}^{\infty} \frac{1}{2^{(1-\delta)j}} \|M_{m_{\rho_j(z)}} H_f k_{\rho_j(z)}\|$$

i.e., Theorem 4, involves various estimates of mean oscillation and is quite technical and tedious.

But many aspects of the proof of Theorem 5 are more interesting.

Note that Theorem 5 involves a lower bound and an upper bound. These two bounds are two completely different problems, and the lower bound is the more difficult of the two problems.

To prove the lower bound, one starts with a "working version" of the Bergman structure in the ball:

The formula $d(x,y) = |1 - \langle x, y\rangle|^{1/2}$ defines what is known as the anisotropic metric on the unit sphere S.

For $x \in S$ and $r > 0$, denote $B(x,r) = \{y \in S : d(x,y) < r\}$.

For each integer $k \geq 0$, let $\{u_{k,1}, \ldots, u_{k,m(k)}\}$ be a subset of S which is maximal with respect to the property

$$B(u_{k,j}, 2^{-k-1}) \cap B(u_{k,j'}, 2^{-k-1}) = \emptyset \quad \text{for all } 1 \leq j < j' \leq m(k).$$

The maximality of $\{u_{k,1}, \ldots, u_{k,m(k)}\}$ implies that

$$\cup_{j=1}^{m(k)} B(u_{k,j}, 2^{-k}) = S.$$

For each pair of $k \geq 0$ and $1 \leq j \leq m(k)$, define

$$T_{k,j} = \{ru : 1 - 2^{-2k} \leq r^2 < 1 - 2^{-2(k+1)}, u \in B(u_{k,j}, 2^{-k})\}.$$

We also define the index set

$$I = \{(k,j) : k \geq 0, 1 \leq j \leq m(k)\}.$$

The sets $T_{k,j}$, $(k,j) \in I$, give us the correct ball version of dyadic decomposition.

These sets are compatible with the properties of the radial contractions ρ_j.

These sets are also a “working version” of the Bergman lattice Γ.

The main application of the local inequality is that it enables us to establish

Proposition 4.7. Given any $2n < p < \infty$, there is a constant $C_{4.7}(p)$ such that the following holds true: Let $f \in L^2(S, d\sigma)$. For each $(k,j) \in I$, let $w_{k,j} \in T_{k,j}$ be such that

$$\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\| \geq \frac{1}{2} \sup_{w \in T_{k,j}} \|M_{m_w} H_f k_w\|.$$

Let $z_{k,j} \in T_{k,j}$, $(k,j) \in I$. Then we have

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j)\in I_m}) \leq C_{4.7}(p)\Phi(\{\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\|^p\}_{(k,j)\in I_m})$$

for every symmetric gauge function Φ and every $m \in \mathbf{N}$, where $I_m = \{(k,j) : k \leq m\}$.

We need the truncated index sets I_m because the proof of the lower bound involves cancellation that has to be done at finite stages.

Proposition 4.7 gets $\mathrm{Var}^{1/2}(f - Pf; z)$ involved and is a crucial step in the proof of the lower bound.

Obviously, another crucial step in the proof of the lower bound is to bring $\|H_f\|_p^+$ into action. This step involve modified kernel functions:

For each pair of $0 < t < \infty$ and $z \in \mathbf{B}$, define

$$\psi_{z,t}(\zeta) = \frac{(1 - |z|^2)^{(n/2)+t}}{(1 - \langle \zeta, z \rangle)^{n+t}},$$

$|\zeta| \leq 1$.

In terms of the Schur multiplier m_z and the normalized reproducing kernel k_z, we have the relation

$$\psi_{z,t} = (1 + |z|)^t m_z^t k_z.$$

We think of $\psi_{z,t}$ as a modified version of k_z. This modification improves the "decaying rate" of the kernel in the following sense:

Proposition 3.1. Given any positive number $0 < t < \infty$, there is a constant $C_{3.1}(t)$ that depends only on t and the complex dimension n such that the inequality

$$|\langle \psi_{z,t}, \psi_{w,t} \rangle| \le C_{3.1}(t) \left(\frac{(1 - |z|^2)^{1/2}(1 - |w|^2)^{1/2}}{|1 - \langle w, z \rangle|} \right)^{n+t}$$

holds for all $z, w \in \mathbf{B}$.

Proof. Given $0 < t < \infty$, it is known from Rudin's book that

$$\int \frac{d\sigma(\zeta)}{|1 - \langle \zeta, \gamma \rangle|^{n+t}} \le \frac{C(t)}{(1 - |\gamma|^2)^t}, \quad \gamma \in \mathbf{B}.$$

Given $z, w \in \mathbf{B}$, the key idea is to express $\langle w, z \rangle$ in the form

$$\langle w, z \rangle = v^2.$$

On the open unit disc, we have the power series expansion

$$\frac{1}{(1-u)^{n+t}} = \sum_{j=0}^{\infty} b_j u^j.$$

If $j \neq k$, then $\langle \zeta, z \rangle^j \perp \langle \zeta, w \rangle^k$ in $L^2(S, d\sigma)$. Therefore

$$\int \frac{d\sigma(\zeta)}{(1 - \langle \zeta, z \rangle)^{n+t}(1 - \langle w, \zeta \rangle)^{n+t}} = \sum_{j=0}^{\infty} b_j^2 \int \langle \zeta, z \rangle^j \langle w, \zeta \rangle^j d\sigma(\zeta).$$

Easy calculation shows that

$$\int \langle \zeta, z\rangle^j \langle w, \zeta\rangle^j d\sigma(\zeta) = \langle w, z\rangle^j \int |\zeta_1|^{2j} d\sigma(\zeta),$$

where ζ_1 denotes the first component of ζ. Since $\langle w, z\rangle = v^2$,

$$\begin{aligned}\int \frac{d\sigma(\zeta)}{(1-\langle \zeta, z\rangle)^{n+t}(1-\langle w, \zeta\rangle)^{n+t}} &= \sum_{j=0}^{\infty} b_j^2 \int (v\zeta_1)^j (v\bar{\zeta}_1)^j d\sigma(\zeta) \\ &= \int \frac{d\sigma(\zeta)}{(1-v\zeta_1)^{n+t}(1-v\bar{\zeta}_1)^{n+t}},\end{aligned}$$

consequently

$$\left| \int \frac{d\sigma(\zeta)}{(1-\langle \zeta, z\rangle)^{n+t}(1-\langle w, \zeta\rangle)^{n+t}} \right| \leq \int \frac{d\sigma(\zeta)}{|1-v\zeta_1|^{n+t}|1-v\bar{\zeta}_1|^{n+t}}.$$

It is elementary that if c is a complex number with $|c| \leq 1$ and if $0 \leq r \leq 1$, then

$$2|1 - rc| \geq |1 - c|.$$

Therefore

$$|1 - \langle w, z\rangle| = |1 - v^2| \leq 2|1 - v\zeta_1 \cdot v\bar{\zeta_1}| \leq 2|1 - v\zeta_1| + 2|1 - v\bar{\zeta_1}|.$$

Thus if we set

$$A = \{\zeta \in S : |1 - v\zeta_1| \geq (1/4)|1 - \langle w, z\rangle|\} \quad \text{and}$$
$$B = \{\zeta \in S : |1 - v\bar{\zeta_1}| \geq (1/4)|1 - \langle w, z\rangle|\},$$

then $A \cup B = S$.

This leads to

$$\left|\int \frac{d\sigma(\zeta)}{(1-\langle\zeta,z\rangle)^{n+t}(1-\langle w,\zeta\rangle)^{n+t}}\right| \leq \frac{4^{n+t}}{|1-\langle w,z\rangle|^{n+t}}\left(\int_A \frac{d\sigma(\zeta)}{|1-v\bar{\zeta}_1|^{n+t}} + \int_B \frac{d\sigma(\zeta)}{|1-v\zeta_1|^{n+t}}\right).$$

We have

$$\int \frac{d\sigma(\zeta)}{|1-v\bar{\zeta}_1|^{n+t}} \leq \frac{C(t)}{(1-|v|^2)^t} \quad \text{and} \quad \int \frac{d\sigma(\zeta)}{|1-v\zeta_1|^{n+t}} \leq \frac{C(t)}{(1-|v|^2)^t}$$

by applying the inequality from Rudin's book.

But $|v|^2 = |\langle w,z\rangle| \leq |w||z|$. Therefore $1-|v|^2 \geq (1/2)(1-|z|^2)$ and $1-|v|^2 \geq (1/2)(1-|w|^2)$.

The combination of the inequalities on the previous page gives us

$$\left| \int \frac{d\sigma(\zeta)}{(1-\langle \zeta, z\rangle)^{n+t}(1-\langle w, \zeta\rangle)^{n+t}} \right| \leq \frac{4^{n+t}2^{t+1}C(t)}{|1-\langle w, z\rangle|^{n+t}(1-|z|^2)^{t/2}(1-|w|^2)^{t/2}}.$$

Since

$$\psi_{z,t}(\zeta) = \frac{(1-|z|^2)^{(n/2)+t}}{(1-\langle \zeta, z\rangle)^{n+t}},$$

the proposition follows from the above inequality. □

Definition. (a) A partial sampling set is a finite subset F of the open unit ball $\mathbf{B}$ with the property that $\mathrm{card}(F \cap T_{k,j}) \leq 1$ for every $(k, j) \in I$.
(b) For any partial sampling set F and any $t > 0$, denote

$$R_F^{(t)} = \sum_{z \in F} \psi_{z,t} \otimes \psi_{z,t}.$$

The significance of Proposition 3.1 is that it enables us to prove

Proposition 3.3. For each $t > 0$, there is a constant $C_{3.3}(t)$ such that $\|R_F^{(t)}\| \leq C_{3.3}(t)$ for every partial sampling set F.

The significance of Proposition 3.3 is that it brings symmetric norm into the proof of lower bound:

Lemma 4.1. Let $0 < t < \infty$ and $2 \le p < \infty$. If $f \in L^2(S, d\sigma)$ and if H_f is bounded, then the inequality

$$\Phi(\{\|H_f\psi_{z,t}\|^p\}_{z\in F}) \le 2^{t(p-2)}C_{3.3}(t)\|(H_f^*H_f)^{p/2}\|_\Phi$$

holds for every symmetric gauge function Φ and every partial sampling set F, where $C_{3.3}(t)$ is the constant provided by Proposition 3.3.

Summarizing, the proof of the lower bound consists of two major steps, Proposition 4.7 (application of the local inequality) and Lemma 4.1 (partial sampling), and several other steps.

The proof of the upper bound in Theorem 5 involves different techniques.

Whereas the lower bound in Theorem 5 is a “one-sided” problem, the upper bound is inherently a “two-sided” problem. That is, for upper bound, it suffices to consider commutators of the form $[P, M_g]$. Therefore the upper bound is a little easier, although it does involve a different kind of decomposition.

Whereas for the lower bound we consider the decomposition $T_{k,j}$ of the ball $\mathbf{B}$, for the upper bound we need to decompose the sphere S.

Three ingredients in the proof of the upper bound:
A reverse Hölder’s inequality, a weak-type inequality, and interpolation.

For each $(k, j) \in I$, let

$$\begin{aligned} B_{k,j} &= B(u_{k,j}, 2^{-k+2}), \\ C_{k,j} &= B(u_{k,j}, 2^{-k+3}) \quad \text{and} \\ D_{k,j} &= B_{k,j} \times B_{k,j}, \end{aligned}$$

where the $u_{k,j}$'s are the same as before. For $g \in L^2(S, d\sigma)$ and $(k, j) \in I$ we define

$$J_t(g; k, j) = \left\{ \frac{1}{\sigma(C_{k,j})} \int_{C_{k,j}} |g - g_{C_{k,j}}|^t d\sigma \right\}^{1/t} \quad \text{and}$$

$$V_t(g; k, j) = \left\{ 2^{4nk} \iint_{D_{k,j}} |g(x) - g(y)|^t d\sigma(x) d\sigma(y) \right\}^{1/t},$$

$1 \le t < \infty$. These are mean oscillations.

Reverse Hölder's inequality:

Proposition 6.4. Suppose that $1 < p \leq t < \infty$. Then there is a constant $C_{6.4} = C_{6.4}(p, t, n)$ such that

$$\Phi_p^+(\{V_t(g; k, j)\}_{(k,j)\in I}) \leq C_{6.4}\Phi_p^+(\{J_1(g; k, j)\}_{(k,j)\in I})$$

for every $g \in L^2(S, d\sigma)$.

Interestingly, the proof of Proposition 6.4 involves symmetric gauge functions that are dual to the family Φ_p^+, $1 < p < \infty$.

For each $1 < p < \infty$, define

$$\Phi_p^-(\{a_j\}_{j\in\mathbf{N}}) = \sum_{j=1}^{\infty} \frac{|a_{\pi(j)}|}{j^{(p-1)/p}}, \quad \{a_j\}_{j\in\mathbf{N}} \in \hat{c},$$

where $\pi : \mathbf{N} \to \mathbf{N}$ is any bijection such that $|a_{\pi(1)}| \geq |a_{\pi(2)}| \geq \cdots \geq |a_{\pi(j)}| \geq \cdots$. Then Φ_p^- is a symmetric gauge function. In fact the pair of symmetric gauge functions $\Phi_{p/(p-1)}^+$ and Φ_p^- are dual to each other. The proof of Proposition 6.4 uses the following special property of Φ_p^-:

Lemma 6.1. Let $1 < p < \infty$. Let X, Y be countable sets and let $N \in \mathbf{N}$. Suppose that $T : X \to Y$ is a map that is at most N-*to*-1. That is, $\text{card}\{x \in X : T(x) = y\} \leq N$ for every $y \in Y$. Then for every set of real numbers $\{a_y\}_{y\in Y}$ we have

$$\Phi_p^-(\{a_{T(x)}\}_{x\in X}) \leq \max\{p, 2\}N^{1/p}\Phi_p^-(\{a_y\}_{y\in Y}).$$

Define the measure

$$d\mu(x,y) = \frac{d\sigma(x)d\sigma(y)}{|1-\langle x,y\rangle|^{2n}}$$

on $S \times S$. For each $1 < p < \infty$, let $L^p_{\mathrm{sym}}(S \times S, d\mu)$ be the collection of functions F on $S \times S$ which are L^p with respect to $d\mu$ and which satisfy the condition

$$|F(x,y)| = |F(y,x)|, \quad (x,y) \in S \times S.$$

For each $F \in L^p_{\mathrm{sym}}(S \times S, d\mu)$, define T_F to be the integral operator on $L^2(S, d\sigma)$ with the kernel function

$$K_F(x,y) = \frac{F(x,y)}{(1-\langle x,y\rangle)^n}.$$

A weak-type inequality for s-numbers:

Proposition 7.1. Given any $2 < p < \infty$, there is a constant $C_{7.1} = C_{7.1}(p, n)$ such that

$$\begin{aligned}\operatorname{card}\{j \in \mathbf{N} : s_j(T_F) > t\} &\leq \frac{C_{7.1}}{t^p} \iint \frac{|F(x,y)|^p}{|1 - \langle x, y\rangle|^{2n}} d\sigma(x) d\sigma(y) \\ &= \frac{C_{7.1}}{t^p} \int |F|^p d\mu\end{aligned}$$

for all $F \in L^p_{\text{sym}}(S \times S, d\mu)$ and $t > 0$.

Combining Propositions 6.4 and 7.1 with an interpolation of a rather ad hoc kind, we have

Proposition 7.2. Let $2 < p < \infty$. Then there is a constant $C_{7.2} = C_{7.2}(p, n)$ such that

$$\|[P, M_g]\|_p^+ \le C_{7.2}\Phi_p^+(\{J_2(g; k, j)\}_{(k,j)\in I})$$

for every $g \in L^2(S, d\sigma)$.

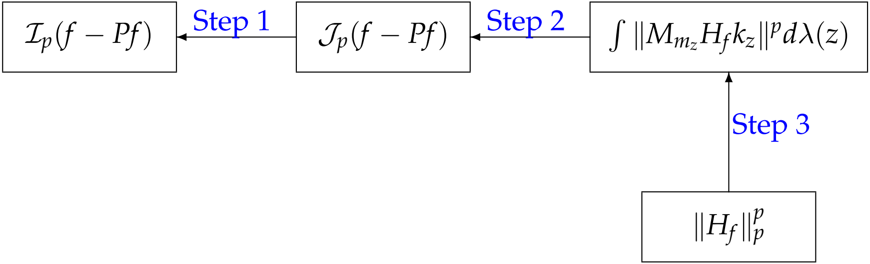
Proposition 7.2 represents the essential part in the proof of the upper bound in Theorem 5.

II The Methods for Schatten class membership results

Our methods have a very strong flavor of modern harmonic analysis. In fact, our work fits into the framework of “analysis on metric spaces”.

Recommended reading: Semmes’ “popular science” article: *An Introduction to Analysis on Metric Spaces, Notices of AMS, April 2003 Vol 50 Issue 4*

Basic Strategy for Proving Theorem 1.4



What Really Happens in Steps 2&3

$$\mathcal{J}_p(f - Pf) \xleftarrow{\text{Step 2}} \int \|M_{m_z} H_f k_z\|^p d\lambda(z) + \epsilon \mathcal{I}_p(f - Pf)$$

$$\int \|M_{m_z} H_f k_z\|^p d\lambda(z) \xleftarrow{\text{Step 3}} \|H_f\|_p^p + \delta \mathcal{I}_p(f - Pf)$$

The $\mathcal{I}_p(f - Pf)$ in the boxes need to be canceled out in the proof. This requires us to first prove the inequality in Theorem 1.4 for f satisfying the condition $\mathcal{I}_p(f - Pf) < \infty$. Then this condition is removed by a technique called "smoothing".

III. Major Steps in the Proofs

Step 1

The formula

$$d(\zeta, \xi) = |1 - \langle \zeta, \xi \rangle|^{1/2}, \quad \zeta, \xi \in S,$$

defines a metric on S. This is an anisotropic metric.

Many of our estimates involve balls with respect to this metric. We denote

$$B(\zeta, r) = \{x \in S : |1 - \langle x, \zeta \rangle|^{1/2} < r\}$$

for $\zeta \in S$ and $r > 0$. We have $\sigma(B(\zeta, r)) \approx r^{2n}$. Roughly speaking S has "dimension" $2n$ with respect to the metric d. That is why the number $2n$ is a "dividing line" for our problem.

Step 1

For each $k \geq 0$, let $\{\xi_{k,1}, \ldots, \xi_{k,\nu(k)}\}$ be a subset of S which is *maximal* with respect to the property

$$B(\xi_{k,i}, 2^{-k+1}) \cap B(\xi_{k,j}, 2^{-k+1}) = \emptyset \quad \text{if } \; i \neq j.$$

Denote

$$A_{k,j} = B(\xi_{k,j}, 2^{-k+3}),$$
$$B_{k,j} = B(\xi_{k,j}, 2^{-k+4}),$$
$$C_{k,j} = B(\xi_{k,j}, 2^{-k+5}),$$

$k \geq 0, 1 \leq j \leq \nu(k)$. The maximality of $\{\xi_{k,1}, \ldots, \xi_{k,\nu(k)}\}$ implies that

$$\cup_{j=1}^{\nu(k)} A_{k,j} = S.$$

Definition For $p \geq 1$ and $g \in L^2(S, d\sigma)$, write

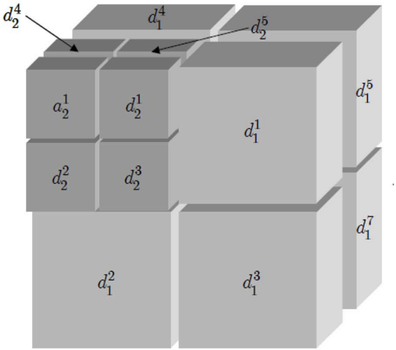
$$\mathcal{J}_p(g) = \sum_{k=0}^{\infty} \sum_{j=1}^{\nu(k)} \left(\frac{1}{\sigma(C_{k,j})} \int_{C_{k,j}} |g - g_{C_{k,j}}| d\sigma \right)^p .$$

Proposition 1. *Given any $p > 1$, there exists a constant $C(p) \in (0, \infty)$ which depends only on p and n such that for every $g \in L^2(S, d\sigma)$*

$$\mathcal{I}_p(g) \leq C(p) \mathcal{J}_p(g).$$

As we can see, $\mathcal{J}_p$ "takes the exponent p outside the integral", and the fact that $\mathcal{J}_p$ dominates $\mathcal{I}_p$ is a kind of "reverse Hölder's inequality".

The proof of this proposition is based on ideas adapted from the Janson-Wolff paper we mentioned earlier. In their paper, dyadic decomposition of cubes plays an essential role. Obviously, dyadic decomposition cannot be copied for sphere in a straightforward fashion. Instead, we use the covering scheme described above, which is the spherical equivalent of dyadic decomposition.



Step 2

A few more definitions:

For each $z \in \mathbf{B} \setminus \{0\}$, the formula

$$\varphi_z(w) = \frac{1}{1 - \langle w, z\rangle} \left\{ z - \frac{\langle w, z\rangle}{|z|^2} z - (1 - |z|^2)^{1/2} \left(w - \frac{\langle w, z\rangle}{|z|^2} z \right) \right\},$$

define a Möbius transform of the ball.

Let $d\lambda$ be the Möbius invariant measure on $\mathbf{B}$. Formula:

$$d\lambda(z) = \frac{dv(z)}{(1 - |z|^2)^{n+1}},$$

where dv is the volume measure on $\mathbf{B}$.

Define the Schur class function

$$m_z(w) = \frac{1-|z|}{1-\langle w, z\rangle},$$

$|w| \leq 1$. The reason why we need m_z will become clear in Step 3.

The main estimate in Step 2 is the following:

Proposition 2. *Suppose $p > 2n$. Let $\gamma > 0$ be given. Then there is a constant $C(\gamma)$ which depends only on n, p and γ such that for any $f \in L^2(S, d\sigma)$,*

$$\mathcal{J}_p(f - Pf) \leq C(\gamma) \int \|M_{m_z} H_f k_z\|^p d\lambda(z) + \gamma \mathcal{I}_p(f - Pf).$$

The proof of this proposition replies on estimates of mean oscillation. For any $f \in L^2(S, d\sigma)$, let

$$\mathrm{SD}(f;\zeta,r) = \left(\frac{1}{\sigma(B(\zeta,r))}\int_{B(\zeta,r)} |f - f_{B(\zeta,r)}|^2 d\sigma\right)^{1/2},$$

where

$$f_{B(\zeta,r)} = \frac{1}{\sigma(B(\zeta,r))}\int_{B(\zeta,r)} f d\sigma.$$

This is the mean oscillation of f over $B(\zeta, r)$. One can also think of this as the Standard Deviation of f over $B(\zeta, r)$, hence the notation. The key to the proof of Proposition 2 is a rather technical inequality which tells us how mean oscillation behaves under the combined action of P and Möbius transform.

Lemma. *There is a constant* C *such that the following estimate holds: Let* $0 < a < 1$ *and* $\zeta \in S$. *Set* $z = (1 - a^2)^{1/2}\zeta$. *If* $N \in \mathbf{N}$ *satisfies the condition* $2^N a \leq 4$, *then*

$$\sum_{k=N}^{\infty} \frac{1}{2^k} \mathrm{SD}((P(f \circ \varphi_z)) \circ \varphi_z; \zeta, 2^k a) \leq \frac{C}{2^{\epsilon N}} \sum_{j=1}^{\infty} \frac{j}{2^{(1-\epsilon)j}} \mathrm{SD}(f; \zeta, 2^j a)$$

for all $f \in L^2(S, d\sigma)$ *and* $0 < \epsilon \leq 1/2$.

This lemma takes care of the part of $\mathcal{J}_p(f - Pf)$ that cannot be controlled by $\int \|M_{m_z} H_f k_z\|^p d\lambda(z)$ in the proof of Proposition 2. This part gives rise to the term $\gamma \mathcal{I}_p(f - Pf)$, which will eventually be canceled out.

Step 3

The key ingredient in this step is something called a "Quasi-Resolution" of the identity operator on $H^2(S)$. To understand what a "Quasi-Resolution" is, let us see a real resolution for identity operator, on a different space. Recall that the Bergman space is

$$A^2(\mathbf{B}) = \{f \in L^2(\mathbf{B}, dv) : f \text{ is analytic on } \mathbf{B}\}.$$

The normalized reproducing kernel for $A^2(\mathbf{B})$ is

$$k_z^{\mathrm{Berg}}(w) = \frac{(1-|z|^2)^{(n+1)/2}}{(1-\langle w, z\rangle)^{n+1}}, \quad |w| < 1.$$

On the Bergman space we have

$$1 = \int k_z^{\text{Berg}} \otimes k_z^{\text{Berg}} d\lambda(z),$$

which "resolves" 1. This is what we call a resolution of the identity operator.

Unfortunately, on the Hardy space $H^2(S)$,

$$\int k_z \otimes k_z d\lambda(z)$$

is not a bounded operator. The reason for this is that, compared with k_z^{Berg}, k_z does not "decay" fast enough. So we need to modify it.

Let $t > 0$. For each $z \in \mathbf{B}$, define the function

$$\psi_{z,t}(w) = \frac{(1-|z|^2)^{(n/2)+t}}{(1-\langle w, z\rangle)^{n+t}},$$

$|w| \leq 1$. Then we have the relation

$$\psi_{z,t} = (1+|z|)^t m_z^t k_z.$$

Remark

- This is why we have m_z involved.
- Clearly $\psi_{z,t}$ is a modification of k_z.
- The difference between $\psi_{z,t}$ and k_z is that $\psi_{z,t}$ decays faster.

This faster decaying rate makes a huge difference:

Proposition 3. *For each $t > 0$, the self-adjoint operator*

$$R_t = \int \psi_{z,t} \otimes \psi_{z,t} d\lambda(z)$$

is bounded on $H^2(S)$. In other words, for any given $t > 0$, there exists a constant $0 < \beta(t) < \infty$ which depends only on t and n such that

$$\langle R_t h, h \rangle \leq \beta(t) \|h\|^2$$

for every $h \in H^2(S)$.

Once we obtain the above inequality, the undesirable effect of m_z^t can be handled by the following

Lemma 4. *There exists a constant* $0 < C < \infty$ *which depends only on* n *such that the inequality* $\|[P, M_{m_z^t}]\| \le Ct$ *holds for all* $z \in \mathbf{B}$ *and* $t > 0$.

Proposition 3 and Lemma 4 lead to the main estimate in Step 3:

Proposition 5. *Let* $p \ge 2$. *Then for all* $0 < t \le 1$ *and* $f \in L^2(S, d\sigma)$ *we have*

$$\int \|M_{m_z} H_f k_z\|^p d\lambda(z) \le 2^{2p}\beta(t)\|H_f\|_p^p + (Ct)^p \mathcal{I}_p(f - Pf).$$

where $\beta(t)$ *is the constant provided by Proposition* 3 *and* C *is a constant depending only on* n *and* p.

Combining the Previous Steps:

Proposition 6 *Let $2n < p < \infty$. Then there exists a constant $0 < C(p) < \infty$ which depends only on n and p such that the inequality*

$$\mathcal{I}_p(f - Pf) \le C(p)\|H_f\|_p^p$$

holds for every $f \in L^2(S, d\sigma)$ satisfying $\mathcal{I}_p(f - Pf) < \infty$.

Proof. Let $f \in L^2(S, d\sigma)$ and suppose $\mathcal{I}_p(f - Pf) < \infty$. Let $\gamma > 0$. Then by Step 1 and Step 2,

$$\mathcal{I}_p(f - Pf) \le C_2(\gamma) \int \|M_{m_z} H_f k_z\|^p d\lambda(z) + C_1 \gamma \mathcal{I}_p(f - Pf).$$

Pick a γ such that $C_1\gamma \le 1/2$. Then since $\mathcal{I}_p(f - Pf) < \infty$, we can cancel out $(1/2)\mathcal{I}_p(f - Pf)$ from both sides to obtain

$$(1/2)\mathcal{I}_p(f - Pf) \leq C_2(\gamma) \int \|M_{m_z} H_f k_z\|^p d\lambda(z).$$

Applying Step 3 on the right hand side,

$$\mathcal{I}_p(f - Pf) \leq 2C_2(\gamma)\{2^{2p}\beta(t)\|H_f\|_p^p + (Ct)^p \mathcal{I}_p(f - Pf)\}.$$

Now set t to be such that $2C_2(\gamma) \cdot (Ct)^p \leq 1/4$. Then, since $\mathcal{I}_p(f - Pf) < \infty$, we can cancel out $(1/2)\mathcal{I}_p(f - Pf)$ from both sides to obtain

$$(1/2)\mathcal{I}_p(f - Pf) \leq 2C_2(\gamma)2^{2p}\beta(t)\|H_f\|_p^p.$$

This completes the proof. □

Smoothing

We need to remove the *a priori* condition $\mathcal{I}_p(f - Pf) < \infty$ in Proposition 6. To accomplish this, we need a large class of functions for which the desired inequality holds. These are Lipschitz functions.

Let $\mathrm{Lip}(S)$ denote the collection of Lipschitz functions on S.

Lemma 7. *If* $g \in \mathrm{Lip}(S)$, *then* $\mathcal{I}_p(g) < \infty$ *for every* $p > 2n$.

What motivates the technique that we call "smoothing" is the convolution on groups. Although S is not a group when $n \geq 3$, there is a natural group acting on it.

Let $\mathcal{U}$ denote the collection of unitary transformations on $\mathbf{C}^n$. For each $U \in \mathcal{U}$, define the operator $W_U : L^2(S, d\sigma) \to L^2(S, d\sigma)$ by

$$(W_U g)(\zeta) = g(U\zeta),$$

$g \in L^2(S, d\sigma)$. By the invariance of σ, W_U is a unitary operator on $L^2(S, d\sigma)$. With the usual multiplication and the operator-norm topology, $\mathcal{U}$ is a compact group. Write dU for the Haar measure on $\mathcal{U}$. We define the operator

$$Y_\Phi g = \int \Phi(U) W_U g dU,$$

$g \in L^2(S, d\sigma)$. The meaning of the above integral is that

$$\langle Y_\Phi g, f \rangle = \int \Phi(U) \langle W_U g, f \rangle dU$$

for every $f \in L^2(S, d\sigma)$.

We need the following easy-to-prove facts:

Lemma 8. *If* $\Phi \in \mathrm{Lip}(\mathcal{U})$ *and* $g \in L^2(S, d\sigma)$*, then*

$$Y_\Phi g \in \mathrm{Lip}(S).$$

Lemma 9. *Let* $\Phi \in C(\mathcal{U})$ *be such that* $\|\Phi\|_1 \neq 0$*. Then*

$$\|H_{Y_\Phi f}\|_p \leq \|\Phi\|_1 \|H_f\|_p$$

for all $f \in L^2(S, d\sigma)$ *and* $1 \leq p < \infty$*.*

It is easy to construct a sequence of functions $\{\Phi_j\}$ on $\mathcal{U}$ with the following properties:

(1) $\Phi_j \geq 0$ on $\mathcal{U}$.

(2) $\int \Phi_j(U)dU = 1$.

(3) $\Phi_j \in \mathrm{Lip}(\mathcal{U})$.

(4) The sequence of operators $\{Y_{\Phi_j}\}$ converges to 1 strongly on $L^2(S, d\sigma)$.

Such a sequence of operators $\{Y_{\Phi_j}\}$ is usually called an approximate identity.

Proof of Theorem 1.4. Let $f \in L^2(S, d\sigma)$ be given and write

$$g = f - Pf.$$

Furthermore, for each $j \geq 1$ let

$$f_j = Y_{\Phi_j} f \quad \text{and} \quad g_j = f_j - Pf_j.$$

Because $[P, W_U] = 0$, we have $[P, Y_{\Phi_j}] = 0$. Therefore

$$g_j = Y_{\Phi_j} g \quad (**)$$

for every $j \geq 1$. Let $2n < p < \infty$ also be given.

By (3) and Lemmas 7 & 8, $\mathcal{I}_p(g_j) < \infty$. Therefore it follows from Proposition 6 that

$$\mathcal{I}_p(g_j) \le C\|H_{f_j}\|_p^p.$$

But by (1), (2) and Lemma 9, we have $\|H_{f_j}\|_p^p \le \|H_f\|_p^p$. Thus

$$\mathcal{I}_p(g_j) \le C\|H_f\|_p^p \quad \text{for every } \ j \ge 1.$$

By (4) and $(**)$, there is a subsequence $\{g_{j_\nu}\}$ of $\{g_j\}$ such that

$$\lim_{\nu\to\infty} g_{j_\nu}(\zeta) = g(\zeta) \quad \text{for } \ \sigma\text{-a.e. } \ \zeta \in S.$$

Applying Fatou's lemma, from the above two inequalities we obtain

$$\mathcal{I}_p(g) \le \liminf_{\nu\to\infty} \mathcal{I}_p(g_{j_\nu}) \le C\|H_f\|_p^p. \qquad \square$$

The proof of Theorem 1.6 is a long journey.

It Needs:

- The same "smoothing" technique
- The action of the n-dimensional torus $\mathbf{T}^n$ on S
- Function of the form $F(\zeta) = \int m(U)\psi(U\zeta)dU$ which has a single "Fourier frequency"
- Something we call "spherical calculus"
- Various estimates