# Remarks on the Rademacher-Menshov Theorem

*Christopher Meaney* *

**Abstract**

We describe Salem's proof of the Rademacher-Menshov Theorem, which shows that one constant works for all orthogonal expansions in all $L^2$-spaces. By changing the emphasis in Salem's proof we produce a lower bound for sums of vectors coming from bi-orthogonal sets of vectors in a Hilbert space. This inequality is applied to sums of columns of an invertible matrix and to Lebesgue constants.

*Keywords:* Orthogonal expansion, Rademacher-Menshov Theorem, Bessel's inequality, bi-orthogonal, Lebesgue constants.
*MSC (2000): 42C15, 46C05.*
*Received 1 August 2006 / Accepted 18 October 2006.*

## 1 Introduction

Here we give an exposition of Salem's proof [14] of the Rademacher-Menshov Theorem. Although it is more elaborate than some proofs and over sixty years old, Salem's method makes it clear that one constant works for all orthogonal expansions in all $L^2$-spaces. Furthermore, some of the inequalities used in the proof lead to a general inequality concerning bi-orthogonal sets of vectors in Hilbert spaces (Proposition 2 in the next section.) In recent work [2] with Leonardo Colzani and Elena Prestini, we used the universal nature of the constant in the Rademacher-Menshov Theorem [13, 11] to produce some almost-everywhere convergence results for inverse Fourier Transforms. Theorem 1 below contains the basic idea used in that work.

*I am grateful to the organizers of the Murramarang conference for the opportunity to participate.

# 2 Statement of Main Results

**Proposition 2.1.** *There is a positive constant $C$ with the following property. For every positive measure space $(X, \mu)$, for every $n \geq 1$, and for every finite set $\{F_1, \dots, F_n\}$ of orthogonal functions in $L^2(X, \mu)$, the maximal function*

$$\mathcal{M}(x) = \max_{1 \leq m \leq n} \left| \sum_{j=1}^{m} F_j(x) \right| \tag{2.1}$$

*has norm*

$$\|\mathcal{M}\|_2 \leq C \log(n+1) \left( \sum_{j=1}^{n} \|F_j\|_2^2 \right)^{1/2}. \tag{2.2}$$

Suppose $H$ is a Hilbert space, with inner-product written as $\langle v, w \rangle$. We say that two sets of vectors $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$ are *bi-orthogonal* when

$$\langle v_j, w_k \rangle = 0, \qquad \forall j \neq k.$$

**Proposition 2.2.** *There is a positive constant $c$ with the following property. For every Hilbert space $H$ and every pair of bi-orthogonal sets $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$ in $H$,*

$$(\log n) \min_{1 \leq k \leq n} |\langle v_k, w_k \rangle| \leq c \max_{1 \leq m \leq n} \|w_m\| \max_{1 \leq k \leq n} \left\| \sum_{j=1}^{k} v_j \right\|. \tag{2.3}$$

These are proved in Section 4. In the next section we give some applications.

# 3 Consequences

## 3.1 Almost everywhere convergence.

Suppose that $(X, \mu)$ is a positive measure space and that $L^2(X, \mu) = \oplus_{n=1}^{\infty} \mathcal{H}_n$ is an orthogonal decomposition into closed subspaces $\mathcal{H}_n$. Let $P_n$ be projection onto $\mathcal{H}_n$. Each function $f \in L^2(X, \mu)$ has an orthogonal expansion

$$\sum_{n=1}^{\infty} P_n f,$$

which converges to $f$ in norm. The partial sum operators are

$$S_N f(x) = \sum_{n=1}^{N} P_n f(x), \qquad \forall N \geq 1, x \in X.$$

Proposition 2.1 says that

$$\left\| \max_{1\leq m\leq N} |S_m f| \right\|_2 \leq C \log(N+1) \|S_N f\|_2, \qquad \forall N \geq 1, f \in L^2(X,\mu).$$

Define the maximal function

$$S^* f(x) = \sup_{N\geq 1} |S_N f(x)|.$$

This is dominated by two pieces,

$$S^* f(x) \leq \sup_{m\geq 0} |S_{2^m} f(x)| + \sup_{m\geq 0} \left( \max_{2^m\leq n<2^{m+1}} |S_n f(x) - S_{2^m} f(x)| \right).$$

We can apply the Cauchy-Schwarz inequality to control the dyadic piece, as on pages 80–81 of [1],

$$\left| \sum_{n=2}^{2^m} P_n f(x) \right|^2 = \left| \sum_{k=1}^{m} \sum_{n=2^{k-1}+1}^{2^k} P_n f(x) \right|^2 \leq \left( \sum_{k=1}^{m} \frac{1}{k^2} \right) \left( \sum_{k=1}^{m} k^2 \left| \sum_{n=2^{k-1}+1}^{2^k} P_n f(x) \right|^2 \right).$$

This implies that

$$\left\| \sup_{m\geq 0} |S_{2^m} f| \right\|_2^2 \leq c \sum_{n=1}^{\infty} (\log(n+1))^2 \|P_n f\|_2^2.$$

For the other term, notice that if we have a non-negative sequence $(a_m)_{m=1}^{\infty}$ then

$$\sup_{m\geq 1} a_m^2 \leq \sum_{m=1}^{\infty} a_m^2.$$

We can use Proposition 2.1 to show that

$$\left\| \sup_{m\geq 0} \left( \max_{2^m\leq n<2^{m+1}} |S_n f - S_{2^m} f| \right) \right\|_2^2 \leq C^2 \sum_{m=0}^{\infty} \left(\log\left(2^m+1\right)\right)^2 \sum_{n=2^m}^{2^{m+1}-1} \|P_n f\|_2^2.$$

Combining these facts gives the general form of the Rademacher-Menshov Theorem.

**Theorem 3.1.** *There is a positive constant $\alpha$ so that for all $f \in L^2(X, \mu)$,*

$$\|S^* f\|_2 \leq \alpha \left( \sum_{n=1}^{\infty} (\log(n+1))^2 \|P_n f\|_2^2 \right)^{1/2}.$$

*If the right hand side is finite, then*

$$f(x) = \lim_{N \to \infty} S_N f(x), \quad \textit{almost everywhere on } X.$$

**Remark 3.1.** This method was used in [2, 9, 10].

## 3.2 Invertible Matrices.

Suppose we equip $\mathbb{C}^n$ with its usual inner product. If $A$ is an invertible $n \times n$ matrix with complex entries then the equation

$$A^{-1} A = I$$

can be viewed as saying that the columns of $A$ and the rows of $A^{-1}$ form a pair of bi-orthogonal sets in $\mathbb{C}^n$. In this case, Proposition 2.2 gives the following result.

**Theorem 3.2.** *Suppose that $\{a_1, \ldots, a_n\}$ are the columns of an $n \times n$ invertible matrix with complex entries $A$ and that $\{b_1, \ldots, b_n\}$ are the rows of $A^{-1}$. Then*

$$\log n \leq c \max_{1 \leq j \leq n} \|b_j\| \max_{1 \leq m \leq n} \|a_1 + \cdots + a_m\|,$$

*where $c$ is a positive constant independent of $n$ and $A$.*

## 3.3 Lebesgue Constants

This example follows the methods of another paper of Salem [15]. Suppose that $\{\phi_1, \ldots, \phi_n\}$ is an orthonormal subset of $L^2(X, \mu)$ consisting of essentially bounded functions, with

$$\|\phi_j\|_\infty \leq M, \qquad \forall 1 \leq j \leq n.$$

Define the maximal function

$$\Phi(x) = \max_{1 \leq m \leq n} \left| \sum_{j=1}^{m} \phi_j(x) \right| \leq \sum_{j=1}^{n} |\phi_j(x)|.$$

If $\Phi(x) = 0$ then $\phi_j(x) = 0$ for $1 \leq j \leq n$ and so the set of places where $\Phi(x) = 0$ can be discarded from $X$ without any effect on our calculations. Notice that for all $x$ where $\Phi(x) \neq 0$, we have

$$\frac{\left|\sum_{j=1}^{m} \phi_j(x)\right|}{\sqrt{\Phi(x)}} \leq \sqrt{\Phi(x)}, \qquad \forall 1 \leq m \leq n.$$

On the set where $\Phi(x) \neq 0$, define $g_j = \phi_j/\sqrt{\Phi}$ and $h_j = \phi_j\sqrt{\Phi}$. These give bi-orthogonal sets $\{g_1, \ldots, g_n\}$ and $\{h_1, \ldots, h_n\}$ in $L^2(X, \mu)$. Furthermore, $\langle g_j, h_k\rangle = \delta_{jk}$,

$$\left\|\sum_{j=1}^{m} g_j\right\|_2^2 \leq \|\Phi\|_1 \quad \text{and} \quad \|h_j\|_2^2 \leq M^2 \|\Phi\|_1 .$$

Proposition 2.2 says that

$$\log n \leq cM \|\Phi\|_1 .$$

**Theorem 3.3.** *There is a positive constant $\beta$ with the following property. Suppose that $\{\phi_1, \ldots, \phi_n\}$ is an orthonormal set in $L^2(X, \mu)$ consisting of essentially bounded functions, with*

$$M = \max_{1\leq j\leq n} \|\phi_j\|_\infty .$$

*Then*

$$\left\| \max_{1\leq m\leq n} \left|\sum_{j=1}^{m} \phi_j\right| \right\|_1 \geq \beta \log(n)/M.$$

**Remark 3.2.** This is a weak form of an inequality conjectured by Littlewood. For much stronger results in the case of characters on compact abelian groups see [8, 6]. For other orthonormal systems see [7, 3]. The inequality here can also be viewed as a special case of Theorem 1 of Olevskiĭ's book [12].

# 4 Proofs

Recall Bessel's inequality for orthogonal vectors in a Hilbert space (page 531 of [5].) Suppose that $\{v_1, \ldots, v_n\}$ is an orthogonal set of non-zero vectors in a Hilbert space $H$. Then $\{v_1/\|v_1\|, \ldots, v_n/\|v_n\|\}$ is an orthonormal set in $H$ and for every vector $w \in H$ we have

$$\sum_{j=1}^{n} \frac{|\langle w, v_j \rangle|^2}{\|v_j\|^2} \leq \|w\|^2 . \tag{4.1}$$

## 4.1 Proof of 2.1

Here we rework the proof published by Salem [14] in 1941 in a slightly more abstract setting.

### 4.1.1 The general set up.

Suppose that $H$ is a Hilbert space. Now let $V = L^2(X, \mu) \otimes H$ be the Hilbert space of $H$-valued $\mu$-measurable square-integrable functions on $X$. Let $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ be a bi-orthogonal pair of subsets of $H$ and define some elements of $V$ by multiplying terms,

$$p_k(x) = F_k(x) w_k, \qquad 1 \leq k \leq n.$$

Then $\{p_1, \ldots, p_n\}$ is an orthogonal subset of $V$ and (4.1) states that

$$\sum_{k=1}^{n} \frac{|\langle P, p_k \rangle_V|^2}{\|p_k\|_V^2} \leq \|P\|_V^2 , \qquad \forall P \in V. \tag{4.2}$$

Let $f_1 \geq f_2 \geq \cdots \geq f_n \geq f_{n+1} = 0$ be a decreasing sequence of characteristic functions of measurable subsets of $X$. For $G \in L^2(X, \mu)$ define an element of $V$ by

$$P_G(x) = G(x) \sum_{j=1}^{n} f_j(x) v_j. \tag{4.3}$$

The Abel transformation lets us rewrite this as

$$P_G(x) = G(x) \sum_{k=1}^{n} \Delta f_k(x) \sigma_k,$$

where $\sigma_k = \sum_{j=1}^{k} v_j$ and $\Delta f_k = f_k - f_{k+1}$, for $1 \leq k \leq n$. Notice that $\{\Delta f_1, \ldots, \Delta f_n\}$ is a set of characteristic functions of mutually disjoint subsets of $X$. For each $x \in X$, at most one of the terms $\Delta f_k(x)$ is non-zero. In particular,

$$\|P_G(x)\|_H^2 = |G(x)|^2 \sum_{k=1}^{n} \Delta f_k(x) \, \|\sigma_k\|_H^2 \, .$$

Integrating over $X$ gives

$$\|P_G\|_V^2 \le \|G\|_2^2 \max_{1\le k\le n} \|\sigma_k\|_H^2 \, .$$

Combining this with (4.2), we have

$$\sum_{k=1}^{n} \left| \int_X G\, f_k \, \frac{\overline{F_k}}{\|F_k\|_2} \, d\mu \right|^2 \frac{|\langle v_k, w_k\rangle|^2}{\|w_k\|_H^2} \le \|G\|_2^2 \max_{1\le k\le n} \|v_1 + \cdots + v_k\|_H^2 \, . \quad (4.4)$$

### 4.1.2 A specific case.

Following Salem, let us now assume that $H = L^2(0,1)$ and

$$v_k(t) = \sqrt{t}\, \sin(2\pi k t) \qquad \text{and} \qquad w_k(t) = \sin(2\pi k t)\,/\sqrt{t}, \qquad \forall 0 < t < 1, k \ge 1.$$

The usual estimates on Lebesgue constants (page 67 in [16]) show that

$$\|w_k\|_H^2 \le A \, \log(k+1) \qquad \text{and} \qquad \|v_1 + \cdots + v_k\|_H^2 \le B \, \log(k+1) \, , \qquad \forall 1 \le k \le n.$$

The constants $A$ and $B$ are independent of $k$ and $n$. Furthermore,

$$\langle v_k, w_k \rangle = \frac{1}{2}, \quad \forall 1 \le k \le n.$$

For this choice of $H$, inequality (4.4) becomes

$$\sum_{k=1}^{n} \frac{1}{\log(k+1)} \left| \int_X G\, f_k \, \frac{\overline{F_k}}{\|F_k\|_2} \, d\mu \right|^2 \le 2AB \, \|G\|_2^2 \log(n+1),$$

and the constant $2AB$ is independent of $X$, $\mu$, and $n$. Moving the logarithm term from the left hand side gives

$$\sum_{k=1}^{n} \left| \int_X G\, f_k \, \frac{\overline{F_k}}{\|F_k\|_2} \, d\mu \right|^2 \le 2AB \, \|G\|_2^2 \, (\log(n+1))^2. \quad (4.5)$$

### 4.1.3 Controlling the maximal function.

Define an integer-valued function $m(x)$ on $X$ by

$$m(x) = \min \left\{ m \ : \ \left| \sum_{k=1}^{m} F_k(x) \right| = \mathcal{M}(x) \right\}, \qquad \forall x \in X,$$

and let $f_k$ be the characteristic function of the subset $\{x \in X \ : \ m(x) \geq k\}$. For each $x \in X$ there is the partial sum

$$S_{m(x)}(x) = \sum_{k=1}^{m(x)} F_k(x) = \sum_{k=1}^{n} f_k(x) F_k(x).$$

For an element $G \in L^2(X, \mu)$, Cauchy-Schwarz gives

$$\left| \int_X G(x) \overline{S_{m(x)}(x)} \, d\mu(x) \right| = \left| \sum_{k=1}^{n} \|F_k\|_2 \int_X G \, f_k \frac{\overline{F_k}}{\|F_k\|_2} \, d\mu \right| \tag{4.6}$$

$$\leq \left( \sum_{k=1}^{n} \|F_k\|_2^2 \right)^{1/2} \left( \sum_{k=1}^{n} \left| \int_X G \, f_k \frac{\overline{F_k}}{\|F_k\|_2} \, d\mu \right|^2 \right)^{1/2}. \tag{4.7}$$

Using inequality (4.5) gives

$$\left| \int_X G(x) \overline{S_{m(x)}(x)} \, d\mu(x) \right| \leq \sqrt{2AB} \ \|G\|_2 \log(n+1) \left( \sum_{k=1}^{n} \|F_k\|_2^2 \right)^{1/2}.$$

This is true for all $G \in L^2(X, \mu)$ and so it follows that

$$\|\mathcal{M}\|_2 = \left\| S_{m(\cdot)} \right\|_2 \leq C \log(n+1) \left( \sum_{k=1}^{n} \|F_k\|_2^2 \right)^{1/2}.$$

This completes the proof of Proposition 2.1. For alternative proofs, see 2.3.1 on page 79 of [1] and Chapter 8 of [4].

## 4.2 Menshov's Result

In 1923 Menshov [11] showed that the logarithm term in Proposition 2.1 is best possible. The following is taken from page 255 of [4].

**Lemma 4.1.** *There is a positive constant $c_0$ so that for every $n \geq 2$ there is an orthonormal subset $\{\psi_1^n, \psi_2^n, \ldots, \psi_n^n\}$ in $L^2(0,1)$ for which the set*

$$\left\{x \in [0,1]\,;\ \max_{1\leq j\leq n}\left|\sum_{k=1}^{j} \psi_k^n(x)\right| > c_0\sqrt{n}\,\log(n)\right\}$$

*has Lebesgue measure greater than* $1/4$.

Notice that this means that the maximal function $\Psi^n(x) = \max_{1\leq j\leq n}\left|\sum_{k=1}^{j} \psi_k^n(x)\right|$ satisfies

$$\|\Psi^n\|_2^2 \geq c_0^2 \frac{n\,(\log(n))^2}{4}$$

and yet $\sum_{j=1}^{n} \left\|\psi_j^n\right\|_2^2 = n$.

## 4.3 Proof of Proposition 2.2

We use the set $\{\psi_1^n, \psi_2^n, \ldots, \psi_n^n\}$ as the orthonormal set in Salem's proof of Proposition 2.1. Keeping the earlier notation, fix a function $G$ on $[0,1]$ for which $|G(x)| = 1$ and

$$G(x)\overline{S_{m(x)}(x)} = \Psi^n(x) \geq 0.$$

Since $\Psi^n$ is nonnegative, inequality (4.6) becomes

$$\|\Psi^n\|_1 \leq \sqrt{n}\left(\sum_{k=1}^{n}\left|\int_0^1 G(x) f_k(x)\overline{\psi_k^n(x)}\,dx\right|^2\right)^{1/2}.$$

Put this back into inequality (4.4) to get

$$\frac{\|\Psi^n\|_1^2}{n}\,\frac{\min_{1\leq k\leq n}|\langle v_k, w_k\rangle|^2}{\max_{1\leq m\leq n}\|w_m\|_H^2} \leq \max_{1\leq k\leq n}\|v_1 + \cdots + v_k\|_H^2.$$

Lemma 4.1 shows that

$$\frac{\|\Psi^n\|_1^2}{n} \geq c_0^2\frac{(\log(n))^2}{16}$$

and so

$$\frac{c_0^2\,(\log(n))^2}{16}\min_{1\leq k\leq n}|\langle v_k, w_k\rangle|^2 \leq \max_{1\leq m\leq n}\|w_m\|_H^2 \max_{1\leq k\leq n}\|v_1 + \cdots + v_k\|_H^2.$$

This completes the proof of Proposition 2.2.

# References


[1] György Alexits, *Convergence problems of orthogonal series*, Translated from the German by I. Földer. International Series of Monographs in Pure and Applied Mathematics, Vol. 20, Pergamon Press, New York, 1961.

[2] Leonardo Colzani, Christopher Meaney, and Elena Prestini, *Almost everywhere convergence of inverse Fourier transforms*, Proc. Amer. Math. Soc. **134** (2006), no. 6, 1651–1660.

[3] Saverio Giulini, Paolo M. Soardi, and Giancarlo Travaglini, *A Cohen type inequality for compact Lie groups*, Proc. Amer. Math. Soc. **77** (1979), no. 3, 359–364.

[4] B. S. Kashin and A. A. Saakyan, *Orthogonal series*, Translations of Mathematical Monographs, vol. 75, American Mathematical Society, Providence, RI, 1989.

[5] Anthony W. Knapp, *Basic real analysis*, Cornerstones, Birkhäuser Boston Inc., Boston, MA, 2005.

[6] S. V. Konyagin, *On the Littlewood problem*, Izv. Akad. Nauk SSSR Ser. Mat. **45** (1981), no. 2, 243–265, 463.

[7] C. Markett, *Cohen type inequalities for Jacobi, Laguerre and Hermite expansions*, SIAM J. Math. Anal. **14** (1983), no. 4, 819–833.

[8] O. Carruth McGehee, Louis Pigno, and Brent Smith, *Hardy's inequality and the $L^1$ norm of exponential sums*, Ann. of Math. (2) **113** (1981), no. 3, 613–618.

[9] Christopher Meaney, *On almost-everywhere convergent eigenfunction expansions of the Laplace-Beltrami operator*, Math. Proc. Cambridge Philos. Soc. **92** (1982), no. 1, 129–131.

[10] Christopher Meaney, Detlef Müller, and Elena Prestini, *A. e. convergence of spectral sums on Lie groups*, to appear in Ann. l'Inst. Fourier.

[11] D. Menchoff, *Sur les séries de fonctions orthogonales. (Premiére Partie. La convergence.)*, Fundamenta math. **4** (1923), 82–105.

[12] A. M. Olevskiĭ, *Fourier series with respect to general orthogonal systems*, Springer-Verlag, New York, 1975.

[13] H. Rademacher, *Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen*, Math. Ann. **87** (1922), 112–138.

[14] Raphaël Salem, *A new proof of a theorem of Menchoff*, Duke Math. J. **8** (1941), 269–272.

[15] Raphaël Salem, *On a problem of Littlewood*, Amer. J. Math. **77** (1955), 535–540.

[16] Antoni Zygmund, *Trigonometric series. 2nd ed. Vols. I, II*, Cambridge University Press, New York, 1959.

**Christopher Meaney**, Department of Mathematics, Macquarie University, North Ryde, NSW 2109 Australia. `chrism@maths.mq.edu.au`